

MATHEMATIK

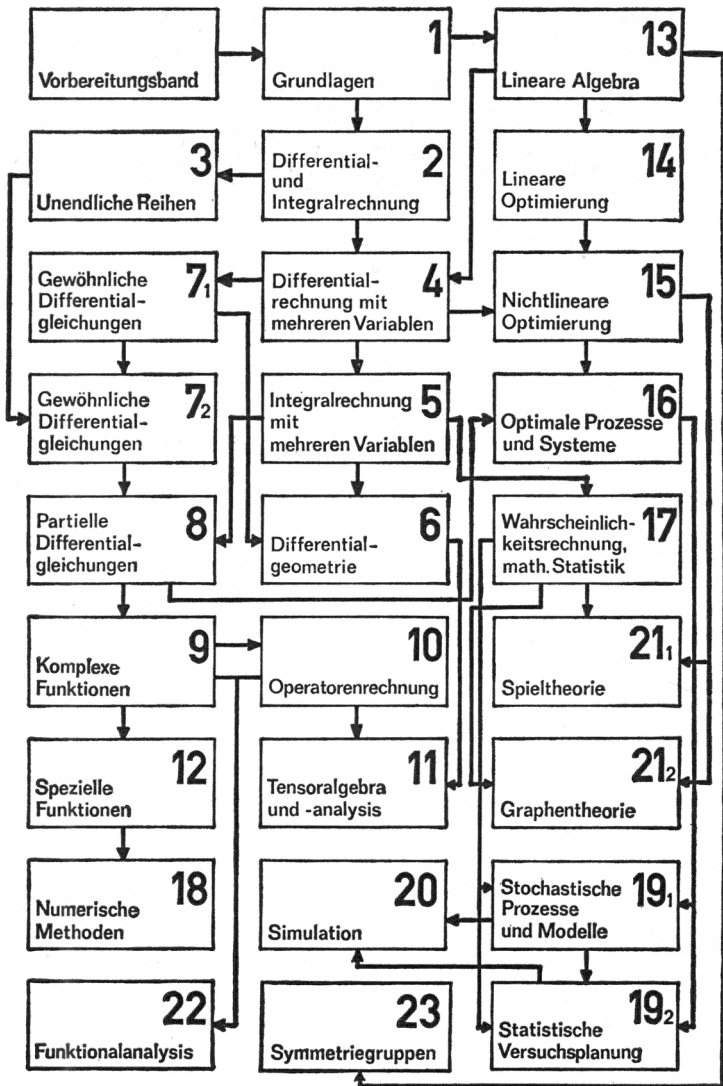
FÜR INGENIEURE
NATURWISSENSCHAFTLER
ÖKONOMEN
LANDWIRTE

22

GÖPFERT · RIEDRICH

Funktionalanalysis

Abhängigkeitsgraph



MATHEMATIK FÜR INGENIEURE, NATURWISSENSCHAFTLER,
ÖKONOMEN UND LANDWIRTE · BAND 22

Herausgeber: Prof. Dr. O. Beyer, Magdeburg · Prof. Dr. H. Erfurth, Merseburg
Prof. Dr. O. Greuel † · Prof. Dr. H. Kadner, Dresden
Prof. Dr. K. Manteuffel, Magdeburg · Doz. Dr. G. Zeidler, Berlin

PROF. DR. A. GÖPFERT
PROF. DR. T. RIEDRICH

Funktionalanalysis



BSB B. G. TEUBNER VERLAGSGESELLSCHAFT
1980

Verantwortlicher Herausgeber:

Dr. rer. nat. Günter Zeidler, Dozent an der Hochschule für Ökonomie „Bruno Leuschner“,
Berlin

Autoren:

Dr. sc. nat. Alfred Göpfert, ordentlicher Professor an der Technischen Hochschule
„Carl Schorlemmer“, Leuna-Merseburg

Dr. rer. nat. habil. Thomas Riedrich, ordentlicher Professor an der Technischen Universität
Dresden

Als Lehrbuch für die Ausbildung an Universitäten und Hochschulen der DDR anerkannt.

Berlin, Dezember 1979

Minister für Hoch- und Fachschulwesen

© BSB B. G. Teubner Verlagsgesellschaft, Leipzig, 1980

1. Auflage

VLN 294-375/16/80 · LSV 1034

Lektor: Dorothea Ziegler

Printed in the German Democratic Republic

Gesamtherstellung: INTERDRUCK Graphischer Großbetrieb Leipzig — III/18/97

Bestell-Nr. 665 985 5

DDR 10,— M

Vorwort

Mit der Formulierung und Präzisierung der Differential- und Integralrechnung (Analysis) im 17. bis zum 19. Jahrhundert war für die Naturwissenschaften, insbesondere für die Physik, ein äußerst tragfähiges mathematisches Fundament geschaffen worden. Die Erfolge, die mit der klassischen Differential- und Integralrechnung erzielt wurden, sind bis heute für den wissenschaftlich-technischen Fortschritt von großem Nutzen. Insbesondere werden gewöhnliche und partielle Differentialgleichungen zur Beschreibung und Modellierung deterministischer Prozesse in umfangreichem Maße angewandt.

In der Anfangszeit der Analysis stand die Untersuchung der einzelnen Lösungen von Differentialgleichungen, von einzelnen Funktionen (wie etwa $\sin x$ oder $\ln x$) im Vordergrund.

Ende des 19. und Anfang des 20. Jahrhunderts entwickelte sich durch das Zusammenwirken solcher Disziplinen wie der Theorie der Differentialgleichungen, der Theorie der Integralgleichungen und der Variationsrechnung in zunehmendem Maße die Auffassung einer Funktion als ein Einzelobjekt oder Element einer ganzen Funktionenmenge bzw. eines Funktionenraumes. Ein solches Element kann damit wieder als eine „unabhängige Variable“ betrachtet werden, die man ihrerseits in Funktionen (= *Abbildungen* oder *Operatoren*) einsetzt und das Verhalten solcher „Funktionen von Funktionen“ untersucht. Es stellte sich heraus, daß in Funktionenräumen Begriffe der anschaulichen, altbekannten (analytischen) Geometrie, wie Länge, Abstand, Winkel usw., in natürlicher Weise eingeführt werden konnten. Damit war eine „Geometrisierung der Analysis“ erreicht worden, die es ermöglichte, komplizierte analytische Sachverhalte in einer einfachen geometrischen Sprache auszudrücken.

In der Mitte der zwanziger Jahre unseres Jahrhunderts wurden (gemeinsam von Mathematikern und Physikern) die mathematischen Grundlagen der Quantentheorie geschaffen. Hier erkannte man die unabdingbare Notwendigkeit der funktional-analytischen Begriffe für die Aufstellung einer adäquaten Mathematisierung. Etwa gleichzeitig etablierte sich die Funktionalanalysis als selbständige mathematische Disziplin. Sie stellte Querverbindungen scheinbar weit voneinander entfernter mathematischer Gebiete her.

Heute ist die Funktionalanalysis, die sich auf Grund ihrer klaren, umfassenden Begriffsbildungen und der vielfachen Anwendungsmöglichkeiten zu einer umfangreichen Theorie entfaltet hat, für die Weiterentwicklung der mathematischen Grundlagen solcher aktueller Gebiete, wie Technologie, Ökonomie und Energiewirtschaft, speziell auch der Netzwerktheorie und Mikroelektronik unumgänglich.

Das Ziel des vorliegenden Bandes besteht darin, einige typische Anwendungen der Funktionalanalysis zu umreißen. Dies wollen wir erreichen, indem wir – aufbauend auf einigen Kenntnissen des Lesers aus mathematischen und physikalischen Grundvorlesungen sowie den vorangehenden Bänden, auf die an verschiedenen Stellen verwiesen wird – im ersten Kapitel an 5 Problemkreisen (Approximation/Fourierkoeffizienten; Stoßvorgang/Distributionen; Oszillator/Spektrum eines Operators, Wahl von Räumen; Verflechtungsmodell/Fixpunkteigenschaften eines Operators;

Steuerung/Prä-Hilbertraum-Methoden) motivieren, welche funktionalanalytischen Begriffsbildungen auftreten werden. Auf diese gehen wir dann in Kap. 2 und 3 systematisch ein. In den restlichen beiden Kapiteln werden ausgewählte Anwendungen der Begriffe des 2. und 3. Kapitels behandelt.

Der von dieser Reihe angesprochene Leserkreis wird naturgemäß sehr unterschiedliche mathematische Kenntnisse haben, so daß wir den Band als *Arbeitsbuch zur Funktionalanalysis* ansehen wollen, d. h., es muß das Verständnis der einzelnen Abschnitte durch wiederholtes Lesen, durch Studium der zugehörigen systematischen Kapitel bzw. Beispiele, durch eigenes Nachrechnen der Sachverhalte, durch Blicke in die angegebene Lehrbuchliteratur und schließlich vor allem bei neuen Begriffen und wichtigen Sätzen durch den Versuch, zugrunde liegende Dinge aus der gewöhnlichen Differential/Integralrechnung, linearen Algebra, Geometrie zu erkennen, erarbeitet werden. Solche zugrunde liegenden bekannten Sachverhalte sind etwa die Auflösungstheorie linearer Gleichungssysteme, der in der Funktionalanalysis die Fredholmsche Alternative (vgl. 3.2.4.) und mit weit größerem Anwendungsbereich die Spektralzerlegung eines Operators (vgl. 5.2.1.) gegenüberstehen, die Theorie der Fourier-Reihen, der in der Funktionalanalysis die Theorie der Orthogonalreihen entspricht (vgl. Satz 2.40), oder die Ableitungsbegriffe, die in Fréchet- und Gateaux-Differential (vgl. 4.2.) oder allgemeineren Begriffen ihr Gegenstück haben.

Besonders wichtig bei einer neu zu erlernenden mathematischen Disziplin sind vorgerechnete Beispiele, die aus den Anwendungen der Disziplin stammen. In dieser Hinsicht möchten wir insbesondere auf unsere Abschnitte 1.2.1., 1.2.2. und 1.2.5., auf die Berechnung des Spektrums für den einfachsten, in der Störungstheorie auftretenden Operator (Bsp. 3.3), auf die Lösungsdarstellung durch eine Reihe für ein elliptisches (verallgemeinertes) Randwertproblem [(5.73)ff.] und auf die (einfachen) Beispiele der Spektralzerlegung eines Operators (vgl. 5.2.3.) verweisen. Die Kapitel, die den Banachschen Fixpunktsatz (vgl. 1.2.4. und 4.3.2.), den Schauderschen Fixpunktsatz (vgl. Bem. 4.5), das Ritzsche Verfahren (vgl. 5.3.), die stark und beschränkt konvexen Funktionale (vgl. 4.2.2.2.) und die verschiedenen Anwendungen zur Physik (etwa 1.2.2., 1.2.3., 5.1.3.–5., 5.2.3. und 4.), zur Steuerungstheorie (vgl. 1.2.5.) und zur Störungsrechnung (vgl. 3.3.3.) betreffen, lassen die direkten Anwendungen der Funktionalanalysis (auch in der Numerik) erkennen. Sie zeigen auch, daß so abstrakte Abschnitte wie 1.3. (Lebesgue-Integrale) oder die Kapitel über die Sobolew-Räume oder über die Erweiterung von Operatoren durch die Anwendungen erforderlich sind.

Dabei möchten wir betonen, daß in diesem Band nur Ansätze für die allereinfachsten Anwendungen der Funktionalanalysis behandelt werden können. Inhaltsreiche und vor allem überzeugende (motivierende) Anwendungen der Funktionalanalysis auf Praxisprobleme sind bei ausführlicher mathematischer Darstellung (wie im Rahmen der MINÖL-Reihe erforderlich) sehr aufwendig und z. T. wesentlich komplizierter als die hier vorgeführten Beispiele. Dieser Tatsache stehen eine wachsende Zahl solcher inhaltsreicher Anwendungen in den Ingenieurwissenschaften, in der Physik, Ökonomie und Biologie gegenüber – wir erwähnen hier nur die interessanten Arbeiten in der Theorie nichtlinearer Schwingungen (man vergleiche hierzu die Tagungs-

berichte der VII. Internationalen Konferenz über Nichtlineare Schwingungen, Berlin 1975) oder in der Theorie der Kernreaktoren, die tiefliegende Hilfsmittel der Spektraltheorie von Halbgruppen linearer Operatoren benötigt (wir verweisen auf das grundlegende Werk von S. B. Schichow über Reaktortheorie, Moskau 1973). Somit ergibt sich in immer stärkerem Maße die Notwendigkeit, in der Mathematikausbildung von Ingenieuren und Naturwissenschaftlern Elemente der Funktionalanalysis einfließen zu lassen. Im internationalen Maßstab ist diese Tendenz bereits deutlich spürbar und stellt qualitativ höhere Ansprüche an die Erziehung und Ausbildung von Ingenieurstudenten. Diesen Forderungen zu entsprechen ist das Anliegen dieses Bandes. Ausdrücklich möchten wir noch auf den Ergänzungsband zum Taschenbuch der Mathematik von Bronstein und Semendjajew (19. Auflage), speziell den Abschnitt 8.1. – Funktionsanalysis –, hinweisen, den wir zum Gebrauch neben diesem Band besonders empfehlen.

Für wertvolle Hinweise danken die Verfasser vor allem Herrn Prof. Dr. Jentsch, TH Karl-Marx-Stadt, Herrn Prof. Dr. Möbius, TU Dresden, Herrn Prof. Dr. Stolle, WPU Rostock, Herrn Prof. Dr. Thämelt, Herrn Dr. sc. Bergmann und Herrn Dr. Lange, sämtlich TH Leuna-Merseburg; Herrn Dr. sc. Schirotzek, TU Dresden, sowie Herrn Doz. Dr. Zeidler, Hochschule für Ökonomie Berlin. Frau Dr. Böhlmann, Herrn Dr. Kayser, beide TU Dresden, und Herrn Dr. Rhodius, PH Dresden, danken wir für wichtige Hinweise bei den Korrekturen. Für die präzise und termingerechte Durchführung der Schreibarbeiten danken wir Frl. Augsten und Frau Hoffmann, TH Leuna-Merseburg. Dem Verlag sei für die gute Zusammenarbeit herzlich gedankt.

Dresden und Merseburg, im Februar 1980

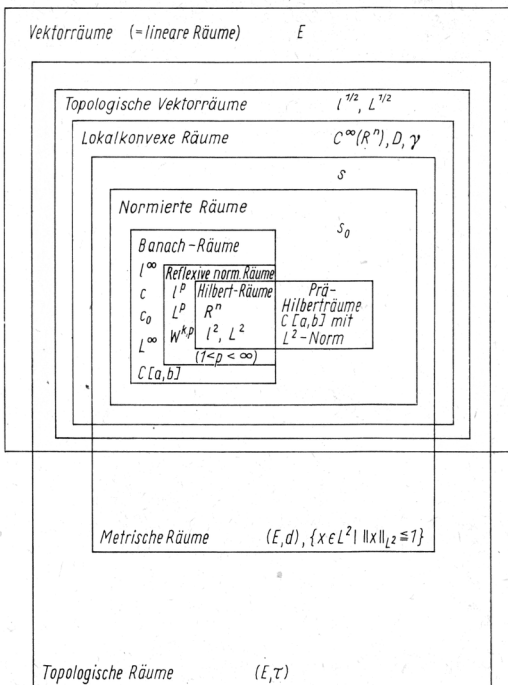
T. Riedrich
A. Göpfert

Inhalt

| | |
|--|----|
| Abhängigkeitsgraph für die einzelnen Raumtypen | 8 |
| Liste der auftretenden Räume | 8 |
| 1. Einführendes zur Anwendung der Funktionalanalysis | 9 |
| 1.1. Allgemeine Grundbegriffe | 9 |
| 1.2. Einführende Anwendungsbeispiele der Funktionalanalysis | 14 |
| 1.2.1. Ein Approximationsproblem | 14 |
| 1.2.2. Mathematische Beschreibung eines Stoßvorgangs | 17 |
| 1.2.3. Hamilton-Funktion und Hermitesche Differentialgleichung beim quantenmechanischen harmonischen Oszillator | 24 |
| 1.2.4. Ein volkswirtschaftliches Verflechtungsmodell als Fixpunktproblem | 28 |
| 1.2.5. Zeitoptimale Steuerung einer erzwungenen gedämpften Schwingung | 31 |
| 1.3. Meßbare Funktionen, Lebesgue-Integral | 36 |
| 2. Räume | 40 |
| 2.1. Vollständige metrische Räume, Banachräume | 40 |
| 2.1.1. Konvergenz von Folgen in metrischen Räumen. Abgeschlossene und offene Mengen. Vollständigkeit und Kompaktheit | 40 |
| 2.1.2. Banachräume | 45 |
| 2.2. Funktionsräume | 48 |
| 2.2.1. Räume stetiger und stetig differenzierbarer Funktionen | 48 |
| 2.2.2. Räume integrierbarer Funktionen (Lebesgue-Räume) | 50 |
| 2.2.3. Sobolew-Räume | 52 |
| 2.2.4. Folgenräume | 54 |
| 2.3. Lineare Funktionale, schwache Konvergenz, dualer Raum | 55 |
| 2.3.1. Lineare Funktionale | 55 |
| 2.3.2. Dualer Raum | 59 |
| 2.3.3. Fortsetzung stetiger linearer Funktionale. Satz von Hahn und Banach, Trennungssätze | 61 |
| 2.3.4. Schwache Konvergenz | 63 |
| 2.4. Hilberträume, Orthogonalentwicklungen | 64 |
| 2.4.1. Grundbegriffe, Beispiele | 64 |
| 2.4.2. Orthogonalentwicklungen | 65 |
| 2.4.3. Orthogonales Komplement, orthogonale direkte Summe | 68 |
| 3. Lineare Operatoren | 72 |
| 3.1. Das Rechnen mit linearen Operatoren | 73 |
| 3.2. Beschränkte lineare Operatoren in Banachräumen | 74 |
| 3.2.1. Spektrum und Resolvente | 75 |
| 3.2.2. Vollstetige lineare Operatoren | 81 |
| 3.2.3. Duale Operatoren | 83 |
| 3.2.4. Fredholmsche Alternative | 85 |
| 3.3. Lineare Operatoren in Hilberträumen | 86 |
| 3.3.1. Grundlegende Begriffe, Sätze und Beispiele | 86 |
| 3.3.1.1. Einführende Beispiele | 86 |
| 3.3.1.2. Die Matrixdarstellung eines linearen Operators | 87 |

| | |
|---|-----|
| Inhalt | 7 |
| 3.3.1.3. Der adjungierte Operator eines beschränkten Operators im Hilbertraum | 88 |
| 3.3.1.4. Der adjungierte Operator eines unbeschränkten Operators im Hilbertraum | 90 |
| 3.3.2. Vollstetige selbstadjungierte Operatoren im Hilbertraum | 92 |
| 3.3.3. Störungsrechnung | 92 |
| 4. Ausgewählte Anwendungen | 95 |
| 4.1. Distributionen | 95 |
| 4.1.1. Distributionen als lineare stetige Funktionale | 95 |
| 4.1.2. Rechenregeln | 96 |
| 4.2. Differentialrechnung und Anwendungen | 103 |
| 4.2.1. Ableitungsbegriffe | 103 |
| 4.2.2. Anwendungen der Ableitungsbegriffe | 105 |
| 4.2.2.1. Beziehungen zur Variationsrechnung | 105 |
| 4.2.2.2. Beziehungen zur konvexen Analysis | 108 |
| 4.3. Anwendungen von Fixpunktsätzen | 109 |
| 4.3.1. Gleichgewichtspunkte und Fixpunkte in Ökonomie und Spieltheorie | 109 |
| 4.3.2. Banachscher Fixpunktsatz und zugehöriges Iterationsverfahren | 110 |
| 5. Unbeschränkte Operatoren in Hilberträumen | 112 |
| 5.1. Halbbeschränkte Operatoren in Hilberträumen | 112 |
| 5.1.1. Der Satz von Friedrichs | 112 |
| 5.1.2. Der Fortsetzungsprozeß | 113 |
| 5.1.3. Einige Operatoren der Quantenmechanik | 116 |
| 5.1.4. Instationäre Zustände und Schrödinger-Gleichung | 118 |
| 5.1.5. Beziehungen zur quantenmechanischen Streuung, Unschärferelation | 120 |
| 5.1.6. Fortsetzung elliptischer Differentialoperatoren | 121 |
| 5.2. Spektralzerlegung selbstadjungierter Operatoren in Hilberträumen | 124 |
| 5.2.1. Vollstetige Operatoren | 124 |
| 5.2.2. Selbstadjungierte Operatoren in Hilberträumen | 125 |
| 5.2.3. Anwendungen auf die Quantenmechanik | 126 |
| 5.2.4. Eigendifferentiale | 128 |
| 5.3. Weitere Anwendungen von Operatoren | 130 |
| Literatur | 132 |
| Register | 134 |

Abhängigkeitsgraph für die einzelnen Raumtypen



Liste der auftretenden Räume

Räume (Abschnitt)

R^n, K^n (2.1.2. (1.1.))

$C(K), C^k(\Omega), C^\infty(\Omega), \tilde{C}^\infty(\Omega)$ (2.2.1.)

$L^p(\Omega), L^p_{loc}(\Omega), L^\infty(\Omega)$ (2.2.2.)

$W^{k,p}(\Omega), W^{k,p}(\Omega)$ (2.2.3.)

$l^p, l^\infty, c, c_0, s, s_0$ (2.2.4.)

$D, D', \mathcal{D}, \mathcal{D}'$ (4.1.1., 4.1.2. [1.1.2.])

L_∞ (1.3.)

Spezielle Hilberträume

$l^2, L^2, W^{m,2}, C^n$ (2.4.1.)

1. Einführendes zur Anwendung der Funktionalanalysis

1.1. Allgemeine Grundbegriffe

Vor einer systematischen, ausführlichen Darstellung der erforderlichen funktional-analytischen Begriffe in Abschnitt 2. und 3. soll hier nur über einige Grundbegriffe der Funktionalanalysis so viel gesagt werden, daß der mathematische Inhalt der ein-führenden Beispiele verständlicher wird.

Ein **Raum** bezeichnet innerhalb einer funktionalanalytischen Beschreibung stets eine geeignet gewählte Menge von gleichartigen Elementen. Deren mathematische Natur und Herkunft kann jeweils vollkommen unterschiedlich sein. So gibt es Räume von Zahlen, Funktionen, Vektoren, Maßen, Matrizen, Operatoren, von Systemen von Funktionen, Vektoren, Matrizen usw. Für die Anwendung der Funktional-analysis haben sich vor allem **lineare Räume** (vgl. Bd. 1) als nützlich erwiesen.

Unter einem **linearen Raum** versteht man eine Menge E von Elementen, f, g, h, \dots , für die die Rechenoperationen der Addition $f + g$ und der Multiplikation γf der Elemente f mit Zahlen γ (reell bzw. komplex) erklärt sind, welche die üblichen Rechen-regeln der Addition von Vektoren bzw. der Multiplikation von Vektoren mit Zahlen erfüllen. Aus diesem Grund nennt man einen linearen Raum auch verallgemeinernd einen **Vektorraum** und seine Elemente Vektoren. Wird die Multiplikation der Ele-mente von E nur mit **reellen** Zahlen zugelassen, so sprechen wir von einem **reellen Vektorraum**; sind alle **komplexen** Zahlen als Faktoren möglich, so nennen wir E einen **komplexen Vektorraum**. Eine Teilmenge $F \subseteq E$ heißt ein **linearer Teilraum** (linearer Unterraum) von E , wenn F bezüglich der in E erklärten Operationen „Addition“ und „Multiplikation mit einem (skalaren) Faktor“ selbst ein linearer Raum ist. In E gibt es genau ein **Nullelement** („Nullvektor“) o ; für dieses gelten die Gleichungen $f + o = f$ und $o \cdot f = o$ für alle $f \in E$. Eine endliche Menge f_1, \dots, f_n von Elementen eines linearen Raumes heißt **linear abhängig**, wenn es Zahlen $\gamma_1, \dots, \gamma_n$ gibt, die nicht alle gleich 0 sind, so daß

$$\sum_{k=1}^n \gamma_k f_k = o \quad (1.1)$$

gilt. Anderenfalls heißt die Menge f_1, \dots, f_n **linear unabhängig**.

Für die Anwendungen, insbesondere für Näherungsverfahren auf funktional-analytischer Grundlage, reicht jedoch die Einführung eines linearen Raumes allein nicht aus, weil mit der Festlegung der Rechenoperationen noch nichts über die relative Lage der Elemente des Raumes zueinander ausgesagt wird, z. B. darüber, ob zwei Elemente „weit“ voneinander entfernt sind oder „nahe“ beieinander liegen. Die Beschreibung derartiger Beziehungen gelingt z. B. mittels des Abstandsbegriffs, d. h., durch die Einführung einer **Metrik** genannten Abstandsfunktion (vgl. Bd. 1).

Definition 1.1: Unter einem **metrischen Raum** versteht man eine Menge X von Elemen-ten x, y, z, \dots , in welcher je zwei Elementen x, y eine nichtnegative Zahl $d(x, y)$, ihr **Abstand**, zugeordnet ist, wobei die folgenden Forderungen (Axiome, Postulate) erhoben werden:

D.1.1

(M1) $d(x, y) = 0$ genau dann, wenn $x = y$ gilt (Definitheit).

(M2) $d(x, y) = d(y, x)$ für alle $x, y \in X$ (Symmetrie).

(M3) $d(x, y) \leq d(x, z) + d(z, y)$ für alle $x, y, z \in X$ (Dreiecksungleichung).

Diese Forderungen sind einerseits allgemein genug, um die unterschiedlichsten Anwendungssituationen erfassen zu können, und sie sind andererseits speziell genug, um das Wesentliche eines Abstandsbegriffs zu enthalten. Stellen wir uns etwa $d(x, y)$ als die übliche Entfernung zweier Punkte x, y in der Ebene vor, so sind die Eigenschaften (M1)–(M3) unmittelbar ersichtlich.

Führt man nun in einem linearen Raum E eine Metrik d ein, so muß man zusätzlich darauf achten, daß die Rechenoperationen in E mit der Metrik d *verträglich* sein müssen. Man kann diese Verträglichkeitsforderungen in verschiedener Weise formulieren. Wir stellen sie in der Form der folgenden beiden Bedingungen:

(LM1) $d(f + h, g + h) = d(f, g)$ für alle $f, g, h \in E$ (Translationsinvarianz).

(LM2) Ist $\{\gamma_n\}$ eine Zahlenfolge mit $\lim_{n \rightarrow \infty} \gamma_n = 0$, so gilt $\lim_{n \rightarrow \infty} d(\gamma_n f, o) = 0$ für alle $f \in E$.

Die Forderung (LM1) sagt aus, daß sich der Abstand zweier Elemente (Vektoren) f und g nicht ändert, wenn beide Elemente um das Element (Vektor) h parallel verschoben werden. Die Forderung (LM2) drückt eine Stetigkeitseigenschaft aus.

Ist in einem linearen Raum eine Metrik gegeben, die außer den Forderungen (M1)–(M3) zusätzlich die Bedingungen (LM1) und (LM2) erfüllt, so heißt E ein **linearer metrischer Raum**.

Beispiel 1.1: Es sei $E = \mathbb{R}^n$ die Menge aller Zahlen- n -tupel $x = (\xi_1, \dots, \xi_n)$ reeller Zahlen ξ_j ($j = 1, \dots, n$) mit den üblichen Festlegungen der Rechenoperationen: Ist $x = (\xi_1, \dots, \xi_n)$, $y = (\eta_1, \dots, \eta_n)$, so sei $x + y = (\xi_1 + \eta_1, \dots, \xi_n + \eta_n)$, und für $\gamma \in \mathbb{R}$ sei $\gamma x = (\gamma \xi_1, \dots, \gamma \xi_n)$ (\mathbb{R} : Menge der reellen Zahlen). Wird der Abstand $d(x, y)$ zweier Elemente $x, y \in \mathbb{R}^n$ (abweichend vom üblichen Entfernungsbegriff) festgelegt durch $d(x, y) = \sum_{k=1}^n |\xi_k - \eta_k|$, so sind die Eigenschaften (M1), (M2), (M3), (LM1), (LM2) erfüllt.

S. Banach (der Begründer der polnischen funktionalanalytischen Schule) spezialisierte den Begriff des linearen metrischen Raumes in einer für viele Anwendungen geeigneten Weise durch die Einführung des Begriffs des normierten Raumes. Die Norm eines Elementes eines linearen Raumes verallgemeinert den Begriff der Länge eines Vektors.

D.1.2 Definition 1.2: Es sei E ein linearer Raum. Jedem $x \in E$ sei eine nichtnegative Zahl $\|x\|$, die **Norm** von x , zugeordnet. Dabei sollen die folgenden Eigenschaften (Normaxiome) (N1)–(N3) erfüllt sein (\mathbf{K} : Menge der komplexen Zahlen¹⁾):

(N1) $\|x\| = 0$ genau dann, wenn $x = o$ (Definitheit).

(N2) $\|\gamma x\| = |\gamma| \cdot \|x\|$ für alle $x \in E$ und alle $\gamma \in \mathbb{R}$ (bzw. $\gamma \in \mathbf{K}$) (positive Homogenität).

(N3) $\|x + y\| \leq \|x\| + \|y\|$ für alle $x, y \in E$ (Dreiecksungleichung).

Dann heißt E ein (reeller bzw. komplexer) **normierter Raum**.

Bemerkung 1.1: Im Hinblick darauf, daß auf demselben Vektorraum E verschiedene Normen eingeführt werden können (abhängig von der jeweiligen Anwendung) bezeichnet man genauer das Paar $(E, \|\cdot\|)$ als den (gegebenen) normierten Raum (vgl. Bsp. 1.2).

¹⁾ In der Literatur wird die Menge der komplexen Zahlen auch häufig mit \mathbf{C} bezeichnet.

Mittels der Gleichung

$$d(x, y) = \|x - y\| \quad (x, y \in E) \quad (1.2)$$

kann man in jedem normierten Raum eine Metrik einführen, die außer den Forderungen (M1)–(M3) auch die Forderungen (LM1), (LM2) erfüllt. (LM2) folgt z. B. so: $d(\gamma_n f, o) = \|\gamma_n f - o\| = |\gamma_n| \cdot \|f\| \rightarrow 0$, da $\{\gamma_n\}$, also auch $\{|\gamma_n|\}$ eine Nullfolge ist. Jeder normierte Raum ist daher mit der Metrik (1.2) ein linearer metrischer Raum.

Beispiel 1.2: Der Raum $E = \mathbb{R}^n$ wird mit der Festlegung $\|x\| = \sum_{j=1}^n |\xi_j|$ ($x = (\xi_1, \dots, \xi_n) \in \mathbb{R}^n$) zu einem normierten Raum. $\sum |\xi_j|$ ist eine Norm $\|x\|$, denn $\sum |\xi_j| = 0 \Leftrightarrow |\xi_j| = 0$ ($j = 1, 2, \dots, n$) $\Leftrightarrow x = (\xi_1, \dots, \xi_n) = o$, also gilt (N1); wegen $\|\gamma x\| = \sum |\gamma \xi_j| = |\gamma| \sum |\xi_j| = |\gamma| \cdot \|x\|$ gilt (N2); wegen der Dreiecksungleichung $\sum |\xi_j + \eta_j| \leq \sum (|\xi_j| + |\eta_j|)$ gilt (N3): $\|x + y\| = \sum |\xi_j + \eta_j| \leq \sum |\xi_j| + \sum |\eta_j| = \|x\| + \|y\|$. Wegen $\|x - y\| = \sum |\xi_j - \eta_j|$ ergibt sich über (1.2) die Metrik von Bsp. 1.1. Im \mathbb{R}^n wird auch durch $\left(\sum_{j=1}^n \xi_j^2\right)^{1/2}$ eine Norm (euklidische Norm) und durch (1.2) dann die zugehörige „euklidische“ Metrik erklärt.

Bemerkung 1.2: Es gibt für die Anwendungen wichtige metrische lineare Räume, deren Metrik nicht durch die Beziehung (1.2) aus einer passenden Norm hergeleitet werden kann. Man muß sogar noch allgemeinere Räume, die topologischen Vektorräume (s. [17] und Raum-Graph) einführen, um die allgemeinsten Zusammenhänge zwischen Vektorraumstruktur und Stetigkeitseigenschaften herstellen zu können. Die zugleich notwendigen und hinreichenden Bedingungen dafür, daß ein topologischer Vektorraum als ein normierter Raum aufgefaßt werden kann, wurden von dem sowjetischen Mathematiker A. N. Kolmogorow, der auch durch seine grundlegenden Arbeiten zur Wahrscheinlichkeitstheorie (s. Bd. 17) bekannt geworden ist, angegeben.

Die Erfolge der Funktionalanalysis liegen u. a. darin begründet, daß es ihr gelingt, komplizierte Fragen aus den **verschiedensten** Gebieten in einfacher, geometrisch faßbarer Form darzustellen und zu behandeln. Die Einführung des Raumbegriffs und des Abstandsbegriffs (Norm, Metrik) machen dies deutlich. Noch enger wird die Bindung an unsere gewohnten geometrischen Anschauungen, wenn auch Winkel zwischen Elementen (Vektoren), insbesondere die Beziehung des „Aufeinander senkrecht Stehens“ verfügbar sind. Dies wird durch die Einführung eines Skalarprodukts (vgl. Bd. 13, 2.3.1.) erreicht.

Definition 1.3: Es sei E ein linearer (komplexer) Raum, in dem zu je zwei Elementen x, y eine komplexe Zahl $\langle x | y \rangle$, das **Skalarprodukt** von x und y , erklärt ist, wobei die folgenden Eigenschaften (S1)–(S4) gelten sollen ($x, y, z \in E$): **D.1.3**

- (S1) $\langle x + z | y \rangle = \langle x | y \rangle + \langle z | y \rangle$.
- (S2) $\langle x | y \rangle = \overline{\langle y | x \rangle}$ (λ sei die zu λ konjugiert komplexe Zahl).
- (S3) $\langle x | \lambda y \rangle = \lambda \langle x | y \rangle$ ($\lambda \in \mathbb{K}$).
- (S4) $\langle x | x \rangle > 0$ für alle $x \neq o$.

Aus (S2) und (S3) folgt die Gleichheit $\langle \lambda x | y \rangle = \lambda \langle x | y \rangle$.

Ist der lineare Raum E reell, so fordert man, daß $\langle x | y \rangle$ reellwertig ist für alle $x, y \in E$. Die Forderung (S2) lautet dann

$$(S2') \quad \langle x | y \rangle = \langle y | x \rangle.$$

Einen linearen Raum E mit einem bestimmten Skalarprodukt $\langle \cdot | \cdot \rangle$ nennen wir einen **unitären Raum** oder auch **Prä-Hilbertraum**.

Beispiel 1.3: Es sei $E = \mathbb{R}^n$ (s. Bsp. 1.1). Mit

$$\langle x | y \rangle = \sum_{k=1}^n \xi_k \eta_k \quad (1.3)$$

wird E zu einem (reellen) Prä-Hilbertraum (n -dimensionaler euklidischer Raum).

Beispiel 1.4: E bestehe aus allen (reellwertigen) stetigen Funktionen $f(t)$ ($a \leq t \leq b$) der reellen Variablen $t \in [a, b]$. Mit den (üblichen) Festlegungen

$$\begin{aligned} (f+g)(t) &= f(t) + g(t) \quad (a \leq t \leq b), \\ (\gamma f)(t) &= \gamma f(t), \quad \gamma \in \mathbb{R}, \end{aligned} \quad (1.4)$$

wird E zu einem linearen Raum, den wir mit $C_{\mathbb{R}}[a, b]$ (vgl. Def. 2.15) bezeichnen. Wir definieren

$$\langle f | g \rangle = \int_a^b f(t) g(t) dt. \quad (1.5)$$

Dann sind die Forderungen (S1)–(S4) sämtlich erfüllt, und $C_{\mathbb{R}}[a, b]$ ist, versehen mit dem Skalarprodukt (1.5), ein reeller Prä-Hilbertraum.

Beispiel 1.5: E bestehe aus allen komplexwertigen stetigen Funktionen $f(t) = u(t) + i v(t)$ ($a \leq t \leq b$) der reellen Variablen $t \in [a, b]$. Mit den Festlegungen analog (1.4) im Bsp. 1.4 wird E zu einem linearen Raum. Bezeichnung: $C[a, b]$. Mit

$$\langle f | g \rangle = \int_a^b \overline{f(t)} g(t) dt \quad (1.6)$$

sind die Forderungen (S1)–(S4) von Def. 1.3 erfüllt, und E ist, versehen mit (1.6), ein komplexer Prä-Hilbertraum.

Beispiel 1.6: Es sei $E = K^n$ die Menge aller komplexen Vektoren $x = (\xi_1, \dots, \xi_n)$ ($\xi_j = \alpha_j + i\beta_j$; α_j, β_j reell, $j = 1, \dots, n$). K^n ist ein linearer Raum (Definition der Vektoroperationen wie im \mathbb{R}^n). Es seien p_1, p_2, \dots, p_n positive reelle Zahlen. Dann wird durch die Vorschrift

$$\langle x | y \rangle = \sum_{k=1}^n p_k \bar{\xi}_k \eta_k \quad (1.7)$$

ein Skalarprodukt auf K^n erklärt (Beweis als Aufgabe).

In einem Prä-Hilbertraum E läßt sich eine Norm dadurch gewinnen, daß man setzt

$$\|f\| = \sqrt{\langle f | f \rangle} \quad (f \in E). \quad (1.8)$$

Im \mathbb{R}^n (Bsp. 1.3) liefert die Gl. (1.8) die Beziehung $\|x\| = \sqrt{\langle x | x \rangle} = \sqrt{\sum_{k=1}^n \xi_k^2}$, was mit der üblichen Definition der Länge eines Vektors übereinstimmt. In jedem Prä-Hilbertraum E läßt sich so mit (1.8) und (1.2) eine Metrik einführen:

$$d(f, g) = \sqrt{\langle f - g | f - g \rangle} \quad (f, g \in E). \quad (1.9)$$

Der Beweis dafür, daß mit einem Skalarprodukt durch (1.8) tatsächlich eine Norm geliefert wird, ergibt sich aus den Eigenschaften des Skalarprodukts (Def. 1.3) und aus der folgenden *Schwarzschen Ungleichung*.

S.1.1 Satz 1.1 (Schwarz-Bunjakowskische Ungleichung): Es sei E ein Prä-Hilbertraum mit dem Skalarprodukt $\langle \cdot | \cdot \rangle$. Dann gilt im Sinne der Definitionsgleichung (1.8) die Ungleichung

$$|\langle f | g \rangle| \leq \|f\| \cdot \|g\|. \quad (1.10)$$

Das Gleichheitszeichen gilt in dieser Ungleichung genau dann, wenn f und g linear abhängig sind, d. h., wenn $f = \alpha g$ oder $g = \beta f$ gilt ($\alpha, \beta \in \mathbf{K}$).

Den Beweis findet man in [7], S. 28.

Der Nachweis der Normeigenschaften (N1)–(N3) für die durch (1.8) eingeführte Funktion $\sqrt{\langle f|f \rangle}$ läßt sich nun leicht führen: Die Eigenschaft (N1) und (N2) ergeben sich unmittelbar aus den Skalarprodukteigenschaften; die Dreiecksungleichung folgt so: Für $f + g = 0$ ist (N3) trivialerweise erfüllt. Für $f + g \neq 0$ erhalten wir mittels der Schwarzschen Ungleichung $\|f + g\|^2 = \langle f + g | f + g \rangle = \langle f + g | f \rangle + \langle f + g | g \rangle \leq \|f + g\| \|f\| + \|f + g\| \|g\| = \|f + g\| (\|f\| + \|g\|)$, woraus nach Division beider Seiten der erhaltenen Ungleichung durch $\|f + g\|$ sofort (N3) folgt.

Nicht jeder normierte Raum ist ein Prä-Hilbertraum. Genau dann, wenn für je zwei Elemente f und g des gegebenen Raumes die **Parallelogrammgleichung**

$$\|f + g\|^2 + \|f - g\|^2 = 2(\|f\|^2 + \|g\|^2) \quad (1.11)$$

gilt, existiert ein Skalarprodukt $\langle \cdot | \cdot \rangle$ auf diesem Raum, so daß $\|f\| = \sqrt{\langle f|f \rangle}$ für alle f besteht. In einem Prä-Hilbertraum kann man (1.11) sofort bestätigen.

Nun können wir den Begriff zweier „aufeinander senkrecht stehender“ Vektoren in beliebigen Prä-Hilberträumen einführen. Systeme paarweise aufeinander senkrecht stehender Vektoren werden für den Aufbau von Orthonormalsystemen in einem Prä-Hilbertraum (s. Def. 1.5) benötigt, die ihrerseits für die Lösung angewandter Probleme von Bedeutung sind.

Definition 1.4: Es sei $(E, \langle \cdot | \cdot \rangle)$ ein Prä-Hilbertraum. Zwei Elemente f und g aus E heißen **zueinander orthogonal** (senkrecht), wenn die folgende Gleichung gilt: **D.1.4**

$$\langle f | g \rangle = 0. \quad (1.12)$$

Definition 1.5: Es sei $(E, \langle \cdot | \cdot \rangle)$ ein Prä-Hilbertraum. Eine Folge f_1, f_2, \dots von Elementen von E heißt ein **Orthogonalsystem** in E , wenn die Gleichungen ($j = 1, 2, \dots$; $k = 1, 2, \dots$) **D.1.5**

$$\langle f_j | f_k \rangle = 0 \quad (j \neq k) \quad (1.13)$$

gelten. Eine Folge e_1, e_2, \dots von Elementen von E heißt ein **Orthonormalsystem (ONS)**, wenn die Gleichungen

$$\langle e_j | e_k \rangle = \begin{cases} 0 & (j \neq k) \\ 1 & (j = k) \end{cases} = \delta_{jk}, \quad \begin{matrix} j = 1, 2, \dots, \\ k = 1, 2, \dots, \end{matrix} \quad (1.14)$$

gelten. (Jedes e_j hat daher wegen (1.8) und (1.9) die Länge 1.)

Bemerkung 1.3: Jedes ONS ist ein Orthogonalsystem. Der Begriff des ONS verallgemeinert den Begriff eines orthonormierten Systems von Vektoren im R^n (bzw. im K^n).

Beispiel 1.7: Es seien $E = R^n$ und $\langle \cdot | \cdot \rangle$ das übliche Skalarprodukt (s. Def. 1.3). Dann bilden die Vektoren $e_j = (0, \dots, 0, 1, 0, \dots, 0)$ (j -te Koordinate gleich 1, sonst 0) ein ONS im R^n .

Beispiel 1.8: Es sei E der Raum aller reellen stetigen Funktionen $f(t), g(t), \dots$ auf dem Intervall $[0, 2\pi]$ mit Skalarprodukt $\langle f | g \rangle = \int_0^{2\pi} f(t) g(t) dt$ (s. Bsp. 1.4).

Dann bilden die Funktionen ($k = 1, 2, \dots$)

$$\frac{1}{\sqrt{2\pi}}, \frac{\cos t}{\sqrt{\pi}}, \frac{\sin t}{\sqrt{\pi}}, \frac{\cos 2t}{\sqrt{\pi}}, \dots, \frac{\cos kt}{\sqrt{\pi}}, \frac{\sin kt}{\sqrt{\pi}}, \dots$$

ein ONS (Beweis durch partielle Integration, vgl. Bd. 3).

1.2. Einführende Anwendungsbeispiele der Funktionalanalysis

Es wird jetzt an verschiedenen Problemen die funktionalanalytische Arbeitsweise erläutert. Neu auftretende Begriffe werden später (Kap. 2., 3.) systematisch behandelt.

1.2.1. Ein Approximationsproblem

Eine typische Aufgabenstellung, die in den Anwendungen wiederholt auftritt, ist die Frage nach der näherungsweise Ersetzung einer komplizierten, rechnerisch schwierig zu handhabenden Funktion durch eine möglichst einfache Funktion, deren Eigenschaften besser zu überblicken sind. Zur Veranschaulichung der Methoden der Funktionalanalysis behandeln wir das Problem der Approximation einer stetigen Funktion durch trigonometrische Polynome. Wir gehen dabei in einer für die Funktionalanalysis typischen Weise vor:

- A) Analytische Formulierung der gestellten Aufgabe.
- B) Herstellung einer abstrakt-funktionalanalytischen, aber **geometrisch motivierten** Fassung des Problems unter A).
- C) Lösung von A) auf der Grundlage der allgemeinen Methoden zur Lösung von B).

Wir führen nun für das gestellte Approximationsproblem die einzelnen Schritte näher aus.

A) Es sei $f(x)$ eine auf dem Intervall $[0, 2\pi]$ definierte stetige reellwertige Funktion. Wir betrachten für ein beliebiges, aber festes $n = 1, 2, \dots$ die Menge M aller trigonometrischen Polynome einer Ordnung, die höchstens gleich n ist:

$$T_n(x) = \frac{a_0}{2} + \sum_{k=1}^n (a_k \cos kx + b_k \sin kx) \quad (0 \leq x \leq 2\pi). \quad (1.15)$$

Unter allen $T_n(x)$ suchen wir solche, die die gegebene Funktion $f(x)$ möglichst gut approximieren. Damit diese Aufgabe sinnvoll ist, muß gesagt werden, was „approximieren“ bzw. „annähern“ im vorliegenden Fall heißen soll. Von den (unendlich) vielen Möglichkeiten wählen wir die folgende, die sich für praktische Aufgaben oft ausreichend gut bewährt hat: Gesucht seien diejenigen trigonometrischen Polynome $T_n(x)$, für die die **mittlere quadratische Abweichung**

$$Q = \left[\frac{1}{2\pi} \int_0^{2\pi} (f(x) - T_n(x))^2 dx \right]^{1/2} \quad (1.16)$$

einen kleinstmöglichen Wert annimmt (vgl. auch Bd. 3, 5.9.). Damit ist die Aufgabe **analytisch** formuliert (Schritt A): Die Koeffizienten $a_0, a_1, \dots, a_n; b_1, \dots, b_n$ in (1.15) sind so zu wählen, daß der Wert Q in (1.16) einen minimalen Wert erhält.

B) Man kann von der speziellen Natur der Funktionen $T_n(x)$ und $f(x)$ abstrahieren und diese Funktionen als Elemente oder „Punkte“ eines Funktionenraumes

ansehen. Alle $T_n(x)$ bilden dann eine gewisse Teilmenge M dieses Raumes. Als linearen normierten Raum wählen wir zweckmäßig den Raum $C_{\mathbf{R}}[0, 2\pi]$ mit der Norm

$$\|f\| = \left[\int_0^{2\pi} (f(x))^2 dx \right]^{1/2}. \quad (1.17)$$

Die zugehörige Metrik hat dann die Form [vgl. (1.2)]

$$d(f, g) = \|f - g\| = \left[\int_0^{2\pi} (f(x) - g(x))^2 dx \right]^{1/2} \quad (f, g \in C_{\mathbf{R}}[0, 2\pi]). \quad (1.18)$$

Der Vergleich der Formeln (1.16) und (1.18) ergibt die Beziehung

$$Q = \frac{1}{\sqrt{2\pi}} \|f - T_n\|. \quad (1.19)$$

Für die Minimierung von Q ist der konstante Faktor $\frac{1}{\sqrt{2\pi}}$ offenbar unwesentlich, mit Q wird auch $\sqrt{2\pi} Q$ minimal. Daher können wir das unter A) gestellte Problem auch so formulieren: „Von allen Elementen T_n der Menge M suchen wir diejenigen, die von dem gegebenen Element f einen kleinstmöglichen Abstand besitzen.“ In dieser Fassung besitzt die gestellte Aufgabe geometrischen Charakter, wie man sich z. B. in Bild 1.1 veranschaulichen kann.

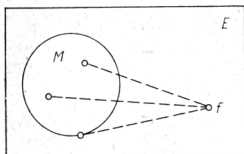


Bild 1.1

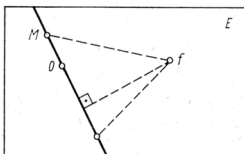


Bild 1.2

Berücksichtigt man die Tatsache, daß die Menge M selbst ein **linearer** Raum, ein Teilraum von $C_{\mathbf{R}}[0, 2\pi]$ ist (die Summe zweier trigonometrischer Polynome der Ordnung höchstens gleich n ist wieder ein solches Polynom, dasselbe gilt für die Multiplikation eines trigonometrischen Polynoms mit einer Zahl), so läßt sich die geometrische Veranschaulichung der gestellten Aufgabe noch präzisieren:

„Von allen Elementen des **linearen** Teilraums M von $C_{\mathbf{R}}[0, 2\pi]$ suchen wir diejenigen, die von dem gegebenen Element $f \in C_{\mathbf{R}}[0, 2\pi]$ einen kleinstmöglichen Abstand besitzen.“

Veranschaulicht man sich die Menge M mittels einer Geraden durch den Nullpunkt o (einfachster Fall eines linearen Teilraums), so erhält unsere geometrische Fassung des Problems die in Bild 1.2 dargestellte Form. Hieraus ist es möglich, Hinweise für einen Lösungsansatz zu gewinnen. Da nämlich im Raum $E = C_{\mathbf{R}}[0, 2\pi]$ ein Skalarprodukt und damit der Begriff des „Senkrechtstehens“ zur Verfügung steht, gelangt man (von Bild 1.2 ausgehend) zu der *Vermutung*, daß die gesuchten Elemente aus M die Eigenschaft haben müssen, auf dem Verbindungsvektor zu f

senkrecht zu stehen. Gesucht werden also zunächst alle Elemente $T_n \in M$ mit der Eigenschaft

$$\langle T_n | T_n - f \rangle = 0. \quad (1.20)$$

Es sei $T_n^{(0)}$ ein Element von M , das die obige Gl. (1.20) erfüllt (wobei zunächst die Frage offen bleibt, ob es ein solches Element tatsächlich gibt oder nicht). Wir vergleichen die Abstände zwischen f und einem beliebigen Element T_n von M und zwischen f und $T_n^{(0)}$. Es gilt nach den Rechenregeln für das Skalarprodukt, wegen (1.20) und wegen $\|T_n^{(0)} - T_n\|^2 \geq 0$:

$$\begin{aligned} \|f - T_n\|^2 &= \langle f - T_n | f - T_n \rangle = \langle f - T_n^{(0)} + (T_n^{(0)} - T_n) | f - T_n^{(0)} \\ &\quad + (T_n^{(0)} - T_n) \rangle \\ &= \|f - T_n^{(0)}\|^2 + 2\langle f - T_n^{(0)} | T_n^{(0)} - T_n \rangle + \|T_n^{(0)} - T_n\|^2 \\ &= \|f - T_n^{(0)}\|^2 - 2\langle f - T_n^{(0)} | T_n \rangle + \|T_n^{(0)} - T_n\|^2 \\ &\geq \|f - T_n^{(0)}\|^2 - 2\langle f - T_n^{(0)} | T_n \rangle. \end{aligned} \quad (1.21)$$

Folglich werden wir die gewünschte Ungleichung

$$\|f - T_n\|^2 \geq \|f - T_n^{(0)}\|^2$$

sicher dann erhalten, wenn $T_n^{(0)}$ so gewählt wird, daß über die Forderung (1.20) hinausgehend die Eigenschaft

$$\langle T_n | T_n^{(0)} - f \rangle = 0 \quad \text{für alle } T_n \in M \quad (1.22)$$

verlangt wird. Genügt $T_n^{(0)}$ der Bedingung (1.22), so folgt mittels derselben Rechnung wie in (1.21), daß die Gleichung

$$\|f - T_n\|^2 = \|f - T_n^{(0)}\|^2 + \|T_n^{(0)} - T_n\|^2 \quad (1.23)$$

für alle $T_n \in M$ gilt. Es folgt dann die Ungleichung

$$\|f - T_n\|^2 \geq \|f - T_n^{(0)}\|^2 \quad (T_n \in M),$$

d. h.,

$$\|f - T_n\| \geq \|f - T_n^{(0)}\| \quad (T_n \in M). \quad (1.24)$$

Ein Element $T_n^{(0)} \in M$, das der Bedingung (1.22) genügt, ist also ein Element aus M mit minimalem Abstand zu f und erfüllt die Forderung der Minimierung von (1.19).

C) Zur Lösung der unter A) gestellten Aufgabe verwenden wir die Ergebnisse von B), also insbesondere die gefundene Bedingung (1.22):

$$\langle T_n | T_n^{(0)} - f \rangle = 0 \quad (T_n \in M).$$

T_n besitze die Form (1.15), für $T_n^{(0)}$ gelte die Gleichheit

$$T_n^{(0)}(x) = \frac{\alpha_0}{2} + \sum_{k=1}^n (\alpha_k \cos kx + \beta_k \sin kx), \quad 0 \leq x \leq 2\pi. \quad (1.25)$$

Die Bedingung (1.22) lautet dann ausgeschrieben

$$\begin{aligned} &\int_0^{2\pi} \left(\frac{a_0}{2} + \sum_{k=1}^n (a_k \cos kx + b_k \sin kx) \right) \\ &\quad \times \left(\frac{\alpha_0}{2} + \sum_{j=1}^n (\alpha_j \cos jx + \beta_j \sin jx) - f(x) \right) dx = 0 \end{aligned} \quad (1.26)$$

für alle Werte $a_0, a_k; b_k$ ($k = 1, \dots, n$). Beachtet man, daß die Funktionen $1, \cos kx, \sin kx, \dots$ ein Orthogonalsystem bilden (s. Bsp. 1.8), so vereinfacht sich (1.26) wegen

$$\int_0^{2\pi} (\cos kx)^2 dx = \pi, \quad \int_0^{2\pi} (\sin kx)^2 dx = \pi \quad (k = 1, 2, \dots)$$

zur Gleichung

$$\pi \left(\frac{a_0 \alpha_0}{2} + \sum_k (a_k \alpha_k + b_k \beta_k) \right) - \frac{a_0}{2} \int_0^{2\pi} f(x) dx - \sum_k \left(a_k \int_0^{2\pi} f(x) \cos kx dx + b_k \int_0^{2\pi} f(x) \sin kx dx \right) = 0. \quad (1.27)$$

In dieser Gleichung sind die Koeffizienten $a_0; a_k, b_k$ ($k = 1, \dots, n$) beliebig wählbar! Wählen wir speziell z. B. $a_0 = 1, a_1 = a_2 = \dots = a_n = 0, b_1 = b_2 = \dots = b_n = 0$, so erhalten wir die Gleichung

$$\pi \left(\frac{\alpha_0}{2} \right) - \frac{1}{2} \int_0^{2\pi} f(x) dx = 0$$

und daraus

$$\alpha_0 = \frac{1}{\pi} \int_0^{2\pi} f(x) dx. \quad (1.28)$$

Setzen wir $a_k = 1, b_k = 0$ und alle übrigen $a_j = 0, b_j = 0$ ($j \neq k$), so erhalten wir entsprechend

$$\alpha_k = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos kx dx \quad (k = 1, \dots, n) \quad (1.29)$$

und analog

$$\beta_k = \frac{1}{\pi} \int_0^{2\pi} f(x) \sin kx dx \quad (k = 1, \dots, n); \quad (1.30)$$

also die bekannten Formeln für die **Fourierkoeffizienten** (vgl. Bd. 3 und 2.4.2.). Wenn wir $\alpha_0; \alpha_k, \beta_k$ gemäß (1.28)–(1.30) wählen, so erkennt man durch Einsetzen in (1.27) sofort, daß (1.27) erfüllt ist. Damit ist $T_n^{(0)}$ nach (1.25) eine Lösung von (1.22) und nach B) eine **Lösung** der eingangs gestellten Aufgabe A).

$T_n^{(0)}$ ist eindeutig bestimmt. Wäre auch $T_n^{(1)} \neq T_n^{(0)}$ Lösung, so gilt neben $\langle T_n | T_n^{(0)} - f \rangle = 0$ auch $\langle T_n, T_n^{(1)} - f \rangle = 0$ für alle $T_n \in M$. Folglich ist $\langle T_n | T_n^{(0)} - T_n^{(1)} \rangle = 0$ für alle $T_n \in M$, insbesondere für $T_n = T_n^{(0)} - T_n^{(1)}$. Dies heißt aber, daß $\langle T_n^{(0)} - T_n^{(1)} | T_n^{(0)} - T_n^{(1)} \rangle = 0$ gilt, also muß wegen (S4) in Def. 1.3 gelten $T_n^{(0)} - T_n^{(1)} = 0$, d. h., es gibt keine andere Lösung als $T_n^{(0)}$.

1.2.2. Mathematische Beschreibung eines Stoßvorganges

Wir betrachten einen einfachen Schwingkreis mit konstanter Selbstinduktion L und Kapazität C sowie mit Anfangsladung 0. Diesen wollen wir zum Zeitpunkt $t = 0$ durch einen **zeitlich punkthaften Spannungsstoß** der Größe S anregen. Uns interessiert

der Stromverlauf *nach* dem Stoß und die mathematische Beschreibung des Stoßes. Wir werden dabei auf den Begriff der verallgemeinerten Ableitung geführt, der von Sobolew erstmals in der Theorie der Differentialgleichungen konsequent genutzt wurde und samt dem Distributionenbegriff heute diese Theorie, die ein Teil der Funktionalanalysis ist, beherrscht.

Wir gehen von bekannten Tatsachen aus. Die Anregung des Schwingkreises erfolge durch Anlegung einer stetigen Spannung $E(t)$, die so verlaufe (S gegeben):

$$\begin{aligned} E(t) &= 0 \quad \text{für} \quad -\infty < t < 0, \\ \int_0^T E(t) dt &= S, \\ E(t) &= 0 \quad \text{für} \quad t \geq T; \end{aligned} \quad (1.31)$$

dabei nehmen wir, um der „momentanen“ (= punkthaften) Anregung nahezukommen, T als klein an. Damit kennt man den sich für $t \geq 0$ einstellenden Stromverlauf $I(t)$: er genügt der Gleichung

$$L\dot{I}(t) + \frac{1}{C} \int_0^t I(\tau) d\tau = E(t) \quad (1.32)$$

mit der Anfangsbedingung

$$I(t) = 0 \quad (t \leq 0). \quad (1.33)$$

(1.32) ist genau genommen eine Integro-Differentialgleichung. Wenn man aber die Stammfunktion

$$J(t) = \int_0^t I(\tau) d\tau \quad (1.34)$$

einführt, so ist (1.32), (1.33) gleichwertig dem Cauchyschen Anfangswertproblem

$$D_2[J] = L\ddot{J} + \frac{1}{C} J = E(t), \quad \dot{J}(0) = J(0) = 0. \quad (1.35)$$

(1): Da wir als erstes Ziel haben, den Stromverlauf nach einer punkthaften Anregung zu studieren, schreiben wir den Stromverlauf $I(t)$ bei der Anregung $E(t)$ von (1.31) auf [wir lösen also (1.32)/(1.33)]:

$$\begin{aligned} I(t) &= L^{-1} \int_0^t E(\tau) \cos(LC)^{-1/2} (t - \tau) d\tau \quad (t \geq 0), \\ I(t) &= 0 \quad \text{für} \quad t < 0 \end{aligned} \quad (1.36)$$

[vgl. Bd. 7 oder bei Verwendung der Laplace-Transformation zur Lösung von (1.32)/(1.33) Bd. 10], und betrachten (1.36) **nach** erfolgter Anregung, also für $t \geq T$.

Für solche t ist $E(t) = 0$, also:

$$I(t) = L^{-1} \int_0^T E(\tau) \cos(LC)^{-1/2} (t - \tau) d\tau \quad (t \geq T). \quad (1.37)$$

Nun hatten wir die Anregungszeit T klein angenommen, näherungsweise ist dann $\cos(LC)^{-1/2}(t - \tau)$ durch $\cos(LC)^{-1/2}t$ ersetzbar, weil in (1.37) nur über τ mit $0 \leq \tau \leq T$ integriert wird. Also gilt für $t \geq T$ genähert

$$I(t) \approx L^{-1} \cos(LC)^{-1/2} t \int_0^T E(\tau) d\tau \quad (1.38)$$

und wegen (1.31)

$$I(t) \approx L^{-1} (\cos(LC)^{-1/2} t) S, \quad t \geq T. \quad (1.39)$$

Zum Zeitpunkt T (Endzeitpunkt der Anregung) ist

$$I(T) \approx SL^{-1} \cos(LC)^{-1/2} T; \quad (1.40)$$

da T klein ist, ersetzen wir das \cos -Glied durch 1:

$$I(T) \approx SL^{-1}. \quad (1.41)$$

Die Formel (1.41) sagt aus, daß $I(t)$ im (kleinen) Anregungszeitintervall $0 \leq t \leq T$ von $I(0) = 0$ auf $I(T) \approx SL^{-1}$ „springt“. Die Herleitung (vgl. [11]) verrät, daß die Näherungen für $I(T)$ und $I(t)$ für $t \geq T$

$$I(T) \approx SL^{-1}, \quad I(t) \approx SL^{-1} \cos(LC)^{-1/2} t \quad (t \geq T) \quad (1.42)$$

um so genauer werden, je kleiner T ist, wenn nur (1.31) erhalten bleibt (die Größe des Spannungsstoßes). $E(t)$ kommt gar nicht mehr vor! Diese Überlegungen legen es nahe, den Spannungsverlauf $E(t)$ und die Anregungsdauer T ganz außerhalb der Betrachtung zu lassen und ein mathematisches Äquivalent für einen punkthaften Stoß der Größe S zum Zeitpunkt $t = 0$ zu finden, so daß in dessen Gefolge der Strom von $I(0) = 0$ auf *genau* SL^{-1} springt (bei $t = 0$), und dann [entsprechend (1.42)] gilt

$$\begin{aligned} I(t) &= SL^{-1} \cos(LC)^{-1/2} t \quad (t \geq 0), \\ I(t) &= 0 \quad (t < 0). \end{aligned} \quad (1.43)$$

Mit Benutzung der **Heaviside-Funktion**

$$\Theta(t) = \begin{cases} 0 & (t < 0), \\ 1 & (t \geq 0) \end{cases} \quad (1.44)$$

lautet der Stromverlauf bei punkthafter Anregung:

$$I(t) = SL^{-1} \Theta(t) \cos(LC)^{-1/2} t \quad (t \in \mathbf{R}). \quad (1.45)$$

$I(t)$ ist (an der Stelle 0) **nicht differenzierbar** (Bild 1.3), kann also [im Gegensatz zu $I(t)$ in (1.36)] nicht (für alle $t \in \mathbf{R}$) einer Differentialgleichung genügen. Dies ist der Ansatzpunkt für den Einsatz der Funktionalanalysis: Sie gestattet nämlich, daß auch in diesem Fall, also für $I(t)$ in (1.45), eine Differentialgleichung in einem verallgemeinerten Sinne „sehr ähnlich“ zu (1.32) [bzw. (1.35)] aufgeschrieben werden kann. Ist dies geschehen, so schließen wir folgendermaßen: Da ersichtlich in (1.32) bzw. (1.35) auf der rechten Seite der Differentialgleichung „der von außen auf das System Schwingkreis wirkende Einfluß“ steht, wie es nach den Kirchhoffschen Gesetzen sein muß, so müßte sich aus der angekündigten, über eine funktionalanalytische Methode gewonnenen „verallgemeinerten“ Differentialgleichung für (1.45) bei

geeigneter Schreibweise die gewünschte mathematische Beschreibung des äußeren Einflusses, also des punkthaften Spannungsstoßes, ablesen lassen.

(2): Wir beschreiben nun die funktionalanalytische Methode, der (bei $t = 0$) nicht differenzierbaren Funktion $I(t)$ von (1.45) eine „verallgemeinerte Ableitung“ zuzuordnen und mit diesem Ableitungsbegriff dann eine zu (1.32) ganz ähnliche „ver-

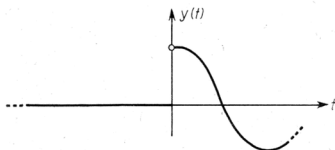


Bild 1.3

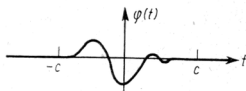


Bild 1.4

allgemeinerte“ Differentialgleichung aufzuschreiben. Wir müssen dafür zunächst die Gesamtheit der Funktionen φ , die für alle $t \in \mathbf{R}$ erklärt sind, betrachten, die jede der folgenden Eigenschaften haben (Bild 1.4):

- (a) $\varphi(t)$ beliebig oft differenzierbar.
- (b) Es gibt (abhängig von φ) eine Zahl $c > 0$, so daß gilt

$$\varphi(t) \equiv 0 \quad \text{für} \quad |t| \geq c.$$

(1.46)

Wenn noch festgelegt wird, wann eine Folge von solchen Funktionen konvergent heißt (vgl. 4.1.; wir machen aber hier in 1.2.2. keinen Gebrauch davon), so heißt die Gesamtheit der erwähnten Funktionen φ der **Grundraum** D (der Distributionen). Jede lineare (und stetige) Funktion (vgl. 2.3.), die für alle $\varphi \in D$ erklärt ist, heißt eine **Distribution**. Die Menge aller Distributionen sei mit D' bezeichnet. Der Begriff einer Distribution erscheint sehr abstrakt. Wir geben aber sogleich zwei klärende Beispiele:

Beispiel 1.9: Es sei f eine gewöhnliche, für alle $t \in \mathbf{R}$ erklärte stetig differenzierbare (oder wenigstens dem Betrage nach in jedem endlichen Intervall integrierbare [vgl. Def. 2.19, $L^1_{\text{loc}}(\mathbf{R})$]) Funktion. Wir betrachten für jedes $\varphi \in D$ das Integral

$$(f, \varphi) = \int_{-\infty}^{+\infty} f(t) \varphi(t) dt. \quad (1.47)$$

Es konvergiert [wegen (b)] und ist natürlich in φ linear ($((f, \alpha_1 \varphi_1 + \alpha_2 \varphi_2) = \alpha_1 (f, \varphi_1) + \alpha_2 (f, \varphi_2)$, $\alpha_1, \alpha_2 \in \mathbf{R}$, $\varphi_1, \varphi_2 \in D$) und stetig. Also haben wir durch f eine Distribution erzeugt, denn durch (1.47) ist jedem $\varphi \in D$ in linearer (stetiger) Weise eine Zahl (eben das Integral) zugeordnet: Diese Distribution können wir durch (f, \cdot) bezeichnen.

Es gelten die Deutungen:

f als Funktion: $t \rightarrow f(t) \quad (t \in \mathbf{R})$

$$f \text{ als Distribution: } \varphi \rightarrow (f, \varphi) = \int_{-\infty}^{+\infty} f(t) \varphi(t) dt \quad (\varphi \in D). \quad (1.48)$$

Man kann zeigen, daß [für $f \in L^1_{\text{loc}}(\mathbf{R})$] die Zuordnung zwischen f und der zugehörigen Distribution (f, \cdot) eineindeutig ist. **Daher unterscheidet man oft nicht zwischen f und (f, \cdot) [38].**

Beispiel 1.10. Jedem $\varphi \in D$ werde die Zahl $\varphi(0)$ zugeordnet. Damit ist eine Distribution δ_0 definiert, denn diese Definitionsvorschrift ist linear (und stetig). Wir schreiben in Anlehnung an die Schreibweise (1.47):

$$\delta_0 : (\delta_0, \varphi) = \varphi(0), \quad \varphi \in D. \quad (1.49)$$

Diese Distribution heißt **Dirac-Distribution**, und man kann zeigen, daß es keine Funktion $f \in L^1_{\text{loc}}(\mathbf{R})$ gibt, so daß diese Distribution durch ein Integral wie in (1.47) dargestellt werden könnte [38, 39].

Damit wissen wir, daß es „mehr“ Distributionen als gewöhnliche Funktionen gibt. Was ist der Sinn ihrer Definition? Ein wichtiger Zweck ist, daß man eine Distribution stets differenzieren kann. Man definiert nämlich für eine beliebige Distribution $f \in D'$, deren Wert für ein $\varphi \in D$ durch (f, φ) bezeichnet sei, die Ableitung f' als folgende Distribution:

$$f' : (f', \varphi) = -(f, \varphi') \quad (\varphi \in D). \quad (1.50)$$

In Worten: Unter f' verstehen wir die Distribution, deren Wert für jedes $\varphi \in D$ durch $-(f, \varphi')$ gegeben ist.

Auch hierfür geben wir zwei Beispiele an und kommen dann zur Distribution (I, \cdot) , die wir unserem Stromverlauf (1.45) zuordnen.

Beispiel 1.11. Wir nehmen eine stetig differenzierbare Funktion f , die für alle $t \in \mathbf{R}$ erklärt ist, und bilden ihre Ableitung f' an jeder Stelle $t \in \mathbf{R}$ sowie die f' zugeordnete Distribution, deren Wert für jedes φ durch

$$(f', \varphi) = \int_{-\infty}^{+\infty} f'(t) \varphi(t) dt$$

gegeben ist. Partielle Integration liefert unter Beachtung von (1.46 (b))

$$(f', \varphi) = - \int_{-\infty}^{+\infty} f(t) \varphi'(t) dt = -(f, \varphi'), \quad (1.51)$$

also gerade (1.50). Daher ist die allgemeine Definition (1.50) motiviert.

Beispiel 1.12: Wir nehmen diesmal die Heaviside-Funktion Θ von (1.44). Sie ist (bei $t = 0$) unstetig! Wir fassen sie als Distribution auf. Dann können wir ihre Ableitung berechnen (es ist gerade die Dirac-Distribution δ_0):

$$\begin{aligned} (\Theta', \varphi) &\stackrel{\text{def}}{=} -(\Theta, \varphi') = - \int_{-\infty}^{+\infty} \Theta(t) \varphi'(t) dt \\ &= - \int_0^{\infty} \varphi'(t) dt \stackrel{(1.46)}{=} -\varphi(t) \Big|_0^{\infty} = \varphi(0). \end{aligned} \quad (1.52)$$

Da dies für jedes $\varphi \in D$ gilt, ist im distributionellen Sinne $\Theta' = \delta_0$.

Nun kommen wir zu unserem Schwingkreis zurück und berechnen für I nach (1.45) den Ausdruck der linken Seite der Differentialgleichung (1.32) im distributionellen Sinne: Wir ersetzen also die Funktion [gemäß (1.48)] $\int_0^t I(\tau) d\tau$ durch ihre

Distribution $\left(\int_0^t I(\tau) d\tau, \cdot\right)$ und die Ableitung \dot{I} durch die distributionelle Ableitung: $\dot{I} \rightarrow (I, \varphi) = -(I, \dot{\varphi})$. Folglich ist

$$\begin{aligned} L(\dot{I}, \varphi) + \frac{1}{C} \left(\int_0^t I(\tau) d\tau, \varphi \right) \\ = -L(I, \dot{\varphi}) + \frac{1}{C} \int_{-\infty}^{+\infty} \varphi(t) \int_0^t \Theta(\tau) SL^{-1} \cos(LC)^{-1/2} \tau d\tau dt \\ = -L \int_{-\infty}^{+\infty} \Theta(t) SL^{-1} \cos(LC)^{-1/2} t \dot{\varphi}(t) dt + \frac{1}{C} \dots \\ = -S \int_0^{\infty} \cos(LC)^{-1/2} t \dot{\varphi}(t) dt + S \int_0^{\infty} (LC)^{-1/2} \sin(LC)^{1/2} t \varphi(t) dt \end{aligned}$$

und durch partielle Integration im ersten Term sowie wegen (1.46) folgt

$$\begin{aligned} &= -S \int_0^{\infty} (LC)^{-1/2} \sin(LC)^{-1/2} t \varphi(t) dt - S \cos(LC)^{-1/2} t \varphi(t) \Big|_0^{\infty} + S \dots \\ &= S \varphi(0) = S(\delta_0, \varphi); \end{aligned}$$

also erfüllt I von (1.45) im distributionellem Sinne die Gleichung [die zu (1.32) „ähnlich“ ist]

$$L(I, \cdot) + \frac{1}{C} \left(\int_0^t I(\tau) d\tau, \cdot \right) = S(\delta_0, \cdot), \quad \varphi \in D, \quad (1.53)$$

d. h., da rechts die Dirac-Distribution δ_0 , multipliziert mit S steht, **beschreibt $S\delta_0$ den Spannungsstoß der Größe S (bei $t = 0$) mathematisch.**

In der Schreibweise (1.35), D_2 aufgefaßt als distributionellen Differentialoperator, wobei die 2. distributionelle Ableitung analog (1.50) durch $f'' : (f'', \varphi) = (f, \varphi'')$, $\varphi \in D$, definiert wird (vgl. 4.1.), lautet das Ergebnis dann

$$D_2[J] = S\delta_0, \quad (1.54)$$

und J ist als die Distribution aufzufassen, die von $\int_0^t I(\tau) d\tau$ erzeugt wird.

Die gefundene mathematische Beschreibung eines Stoßvorganges war schon in Bd. 7 und Bd. 10 Untersuchungsgegenstand. In Bd. 7 wurde der punktförmige Stoß approximiert durch eine Folge kürzer werdender „Rechteckstöße“ ($T_n \rightarrow 0$, $T_n \geq 0$, $n = 1, 2, \dots$):

$$E_n(t) = \begin{cases} 0 & (t < 0), \\ \frac{S}{T_n} & (0 \leq t \leq T_n), \\ 0 & (t > T_n), \end{cases}$$

dann ist stets $\int_0^{T_n} \frac{S}{T_n} dt = S$ ($n = 1, 2, \dots$).

Der Stromverlauf I_n nach der Anregung E_n , also $\forall t \geq T_n$, ergibt sich auch nach (1.37):

$$I_n(t) = L^{-1} \int_0^{T_n} \frac{S}{T_n} \cos(LC)^{-1/2}(t - \tau) d\tau. \quad (1.55)$$

Nun interessiert natürlich der Grenzwert für $n \rightarrow \infty$ (also $T_n \rightarrow 0$). Die Regel von de l'Hospital ergibt wie in Bd. 7/1 unsere Lösung (1.43) für den Stromverlauf bei stoßförmiger Anregung: man faßt in (1.55) T_n als Variable auf. Dann folgt:

$$\lim_{T_n \rightarrow 0} SL^{-1} \cos(LC)^{-1/2}(t - T_n) = SL^{-1} \cos(LC)^{-1/2} t, \quad t \geq 0.$$

Diese Approximation ist auch für den Übergang von der Folge $\{E_n\}$ zur Dirac-Distribution anwendbar. Wir fassen dazu E_n als Distribution auf: $E_n \rightarrow (E_n, \cdot)$. Dann ist für jedes $\varphi \in D$ nach der de l'Hospitalischen Regel

$$\lim_{n \rightarrow \infty} \int_{-\infty}^{+\infty} E_n(t) \varphi(t) dt = \lim_{n \rightarrow \infty} \frac{1}{T_n} S \int_0^{T_n} \varphi(t) dt = S\varphi(0). \quad (1.56)$$

Also: Der Grenzwert der Folge der Distributionen (E_n, \cdot) ist $S\delta_0(\cdot)$. Der Konvergenzbegriff in (1.56) ist die **schwache Konvergenz einer Distributionenfolge** $\{E_n\}$ gegen die δ_0 -Distribution (s. auch 2.2.4.):

$$\{E_n\} \rightarrow S\delta. \quad (1.57)$$

Die δ_0 -Distribution wurde schon in Bd. 10 eingeführt. Dort heißt es, daß für sie folgende Regel gilt:

$$\int_{-\infty}^{+\infty} \varphi(\tau) \delta_0(\tau) d\tau = \varphi(0). \quad (1.58)$$

Dabei ist der Ausdruck links sowohl dort in Bd. 10 als auch hier bei uns als Vorschrift aufzufassen, jeder Funktion φ den Wert $\varphi(0)$ zuzuordnen. Wir hatten für φ dabei nur Elemente aus D zugelassen, die Zuordnungsvorschrift (wenn man nicht differenzieren will) ist aber auch für jedes nur stetige φ erklärt.

Wir wollen noch auf den Zusammenhang mit Greenschen Funktionen hinweisen (vgl. 4.1. und Bd. 7/2 und 8). Man gewinnt die Greensche Funktion $J(t, t_0)$ für unser Anfangswertproblem (1.35), indem man bei homogenen Anfangsbedingungen die Wirkung eines Einheitsstoßes zur Zeit $t_0 > 0$ studiert. Dieser wird durch die „verschobene“ δ -Distribution δ_{t_0} beschrieben, die jedem $\varphi \in D$ den Wert $\varphi(t_0)$ zuordnet. Also muß statt (1.54) jetzt

$$D_2[J] = \delta_{t_0} \quad (t_0 > 0) \quad (1.59)$$

erfüllt werden. Der wichtige Unterschied gegenüber (1.54) ist der Stoßzeitpunkt $t_0 > 0$. Also antwortet der (Zeitverschiebungen gegenüber invariante) Schwingkreis auch „um t_0 verschoben“ (vgl. 1.43):

$$I(t; t_0) = L^{-1} \Theta(t - t_0) \cos(LC)^{-1/2}(t - t_0), \quad (1.60)$$

und hieraus ergibt sich wieder [wie in (1.54)]

$$J(t; t_0) = \int_0^t I(\tau; t_0) d\tau, \quad t \in \mathbf{R}, \quad t_0 > 0. \quad (1.61)$$

Dies ist die gesuchte Greensche Funktion. Sie erfüllt die Bedingungen von Bd. 7/2, die ihr zugeordnete Distribution J erfüllt (1.59). Insbesondere wird die Lösung von (1.35) durch

$$J(t) = \int_0^t J(t; \tau) E(\tau) d\tau, \quad (1.62)$$

die Lösung von (1.32), (1.33) durch

$$I(t) = \int_0^t I(t; \tau) E(\tau) d\tau \quad (1.63)$$

dargestellt; $I(t; t_0)$ bzw. $J(t; t_0)$ gestatten also die Darstellung von $I(t)$ bzw. $J(t)$ als gewichtete [mit $E(\tau)$] Aufsummierung (das Integral) der Wirkungen $I(t; \tau)$ bzw. $J(t; \tau)$ der Stöße zu den Zeitpunkten $\tau \geq 0$.

1.2.3. Hamilton-Funktion und Hermiteische Differentialgleichung beim quantenmechanischen harmonischen Oszillator

Wir benutzen den linearen ungedämpften Oszillator, um in die Anwendung der Funktionalanalysis in der Quantenmechanik einzuführen.

Unter einem linearen harmonischen Oszillator versteht man bekanntlich die (eindimensionale) Bewegung eines Teilchens der Masse m unter dem Einfluß eines Kraftfeldes $K(x)$:

$$K(x) = -kx, \quad (1.64)$$

wenn $x(t)$ die Ortskoordinate des Teilchens und $k > 0$ eine gegebene Konstante ist. Die Anfangsbedingungen seien

$$\begin{aligned} x_0 &= x(0) = C_1 \quad (\text{Anfangsort}), \\ v_0 &= \dot{x}(0) = C_2 \quad (\text{Anfangsgeschwindigkeit}). \end{aligned} \quad (1.65)$$

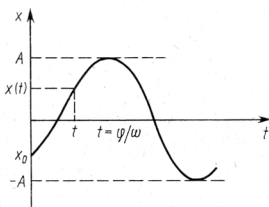


Bild 1.5

Die Kraft $K(x)$ bewirkt eine ungedämpfte Schwingung, deren mathematische Beschreibung durch das Newtonsche Kraftgesetz

$$m\ddot{x} = -kx \quad (1.66)$$

und die Anfangsbedingungen gegeben ist. Die Lösung dieses Anfangswertproblems (1.65), (1.66) ist (vgl. Bd. 3, 5.1.)

$$x(t) = A \cos(\omega t - \varphi), \quad (1.67)$$

wobei Amplitude A , Frequenz ω und Phase φ gegeben sind durch

$$\omega^2 = \frac{k}{m}, \quad A = \sqrt{x_0^2 + \frac{v_0^2}{\omega^2}}, \quad \sin \varphi = \frac{v_0}{A\omega}, \quad \cos \varphi = \frac{x_0}{A}. \quad (1.68)$$

Nach dem Energiesatz für unser Problem ist

$$E = E_K + E_P = \frac{m}{2} \dot{x}^2 + \frac{k}{2} x^2 \quad (1.69)$$

eine Konstante für $t \geq 0$. Folglich ist wegen (1.65) und (1.68)

$$E = \frac{m}{2} v_0^2 + \frac{k}{2} x_0^2 = \frac{m\omega^2}{2} A^2, \quad (1.70)$$

also ist die Amplitude A nur von der Gesamtenergie E abhängig. Da die Größe von φ nur eine Verschiebung der Kurve $x(t) = A \cos \omega t$ um φ/ω ausmacht, wollen wir sagen, die **Gesamtenergie** beschreibt den **Zustand** des Oszillators.

Wenn also zu festen k, m die Bedingungen (1.65) vorgegeben sind, berechnet man E nach (1.69), und nun schwingt das Teilchen ungedämpft mit der Amplitude A für diese Gesamtenergie. Dieser „Zustand“ ist **stationär**, denn er ändert sich nicht mit wachsender Zeit. Quantenmechanisch sind aber – wie wir gleich sehen werden – stationäre Zustände nur für **gewisse** Gesamtenergien E_n , $n = 0, 1, \dots$, möglich.

Für stationäre Zustände quantenmechanischer Systeme (als ein solches wollen wir unseren Oszillator ansehen) liefert die Quantenmechanik folgendes Herangehen:

1. Man schreibe die klassische Hamilton-Funktion $H(p_k, x_k)$ des Systems auf. Dann ersetze man H durch einen **Operator** \hat{H} , indem

- a) die Ortskoordinaten x_k durch die Multiplikation mit x_k ,
- b) die Impulskoordinaten $p_k (= m\dot{x}_k)$ durch den Differentiationsoperator

$$\frac{\hbar}{i} \frac{\partial}{\partial x_k} \quad (1.71)$$

ersetzt werden.

2. Man finde einen geeigneten Hilbertraum H (s. 2.4.), so daß der Definitionsbereich $D(\hat{H})$ von \hat{H} eine Teilmenge dieses Hilbertraumes ist. \hat{H} muß ein selbstadjungierter Operator (vgl. 3.3.) sein. [Dies ergibt sich bei 1a) von selbst, bei 1b) muß der gewonnene Operator der Differentiation erst erweitert werden (s. 5.1.2.).] Der erweiterte Operator heiße \tilde{H} .

3. Man schreibe das Eigenwertproblem (vgl. Bd. 13) auf bezüglich des Operators \tilde{H} :

$$\tilde{H}f = \lambda f, \quad \lambda \text{ reell}, \quad f \in D(\tilde{H}), \quad (1.72)$$

und bestimme das **Spektrum** (s. 3.2.1.) des Operators \tilde{H} . Insbesondere gilt dann

a) die **Eigenwerte** λ_n sind gerade die möglichen Energieniveaus E_n des quantenmechanischen Systems,

b) die zugehörigen linear unabhängigen **Eigenvektoren** f_n entsprechen (eindeutig) den zugehörigen stationären quantenmechanischen Zuständen,

c) erfolgt ein Übergang des Systems von dem stationären Zustand der Energie E_n zum stationären Zustand mit der Energie E_{n-1} , so findet gleichzeitig eine elektromagnetische Strahlung der Frequenz

$$\nu = \frac{1}{h} (E_n - E_{n-1}) \quad (1.73)$$

oder eine äquivalente Erscheinung mit entsprechender Energie- und Impulsübertragung statt.

Wir wollen für unser Beispiel dieses Programm genauer ausführen. Als Hilbertraum H wählen wir

$$H = L^2(\mathbf{R}) \quad (1.74)$$

(s. 2.4.1. und 1.3.).

1. Die Hamilton-Funktion lautet in unserem Beispiel

$$H(p, x) = \frac{1}{2m} p^2 + \frac{k}{2} x^2 \quad (1.75)$$

(sie ist für unser Beispiel einfach die Energie (1.69), wobei statt \dot{x} der Impuls $p = m\dot{x}$ als Variable stehen muß). Anwendung der Ersetzungsregel 1 a), 1 b) ergibt dann den Operator

$$\mathfrak{H} = \frac{1}{2m} \left(\frac{\hbar}{i} \frac{d}{dx} \right) \left(\frac{\hbar}{i} \frac{d}{dx} \right) + \frac{k}{2} x^2, \quad (1.76)$$

wobei an Stelle des Punktes ein Element des Definitionsbereiches $D(\mathfrak{H})$ von \mathfrak{H} zu stehen hat.

2. Als Definitionsbereich $D(\mathfrak{H})$ bietet sich an (s. 2.2.1.)

$$D(\mathfrak{H}) = \{f \mid f \in \dot{C}^\infty(\mathbf{R})\}. \quad (1.77)$$

Es ist dann

$$\mathfrak{H}f = -\frac{\hbar^2}{2m} \frac{d^2 f}{dx^2} + \frac{k}{2} x^2 \cdot f \quad (f \in D(\mathfrak{H})). \quad (1.78)$$

Den **Differentialoperator** rechts können wir identifizieren: bis auf einige Konstanten ist es der aus Band 7/2 bekannte **Hermiteische Differentialoperator**

$$D_2[y] = -y''(x) + x^2 y(x), \quad -\infty < x < +\infty. \quad (1.79)$$

3. \mathfrak{H} von (1.76) ist noch nicht selbstadjungiert [1, S. 137]. Wir müssen \mathfrak{H} durch seine Erweiterung $\tilde{\mathfrak{H}}$ (s. Kap. 3.) ersetzen. \mathfrak{H} selbst gestattet aber schon, viele Eigenschaften von $\tilde{\mathfrak{H}}$ auszurechnen, so z. B. die Eigenwerte λ_n und zugehörige Eigenfunktionen f_n : Es müssen alle $f_n \neq 0$ mit

$$\mathfrak{H}f_n = \lambda_n f_n, \quad n = 0, 1, 2, \dots, \quad (1.80)$$

gesucht werden. Steht statt \mathfrak{H} der Operator D_2 (aus 1.79), so kennen wir alle Eigenwerte $\tilde{\lambda}_n$ und Eigenfunktionen H_n [vgl. (2.69)]:

$$\tilde{\lambda}_n = 2n + 1, \quad n = 0, 1, 2, \dots; \quad H_n(x) = C_n e^{\frac{x^2}{2}} \frac{d^n}{dx^n} (e^{-x^2}). \quad (1.81)$$

Man bemerkt, daß der (ursprüngliche) Definitionsbereich für \mathfrak{H} nach (1.77) zu eng ist. Aber bei (1.80) müßte auch $\tilde{\mathfrak{H}}$ stehen, und es ist $D(\tilde{\mathfrak{H}}) \supseteq D(\mathfrak{H})$.

Wählt man die C_n so, daß

$$\int_{-\infty}^{+\infty} |H_n(x)|^2 dx = 1 \quad (1.82)$$

ist (**Normierung**), so ist jedem Eigenwert λ_n genau eine Eigenfunktion zugeordnet. Die $H_n(x)$ heißen normierte **Hermiteische Funktionen**. Wenn man den Operator \mathfrak{H} nun auf

$$\varphi_n(x) = D_n H_n(\alpha x), \quad n = 0, 1, 2, \dots; \quad D_n \text{ reell}, \quad \alpha = \sqrt[4]{m\hbar k^{-1}} \quad (1.83)$$

anwendet, so folgt

$$\hat{H}\psi_n(x) = \hbar\omega(n + \frac{1}{2})\psi_n(x), \quad n = 0, 1, 2, \dots, \quad (1.84)$$

d. h., die Funktionen in (1.83) sind gerade Eigenfunktionen des Eigenwertproblems (1.80) mit den aus (1.84) ablesbaren Eigenwerten

$$\lambda_n = \hbar\omega(n + \frac{1}{2}), \quad n = 0, 1, 2, \dots \quad (1.85)$$

Entsprechend 3a)–3c) heißt dieses Resultat für die quantenmechanische Betrachtung des Oszillators:

Mögliche Energieniveaus E_n für stationäre Zustände sind

$$E_n (= \lambda_n) = \hbar\omega(n + \frac{1}{2}), \quad (1.86)$$

und wenn der Oszillator auf ein niedrigeres Niveau übergeht, $E_n \rightarrow E_{n-1}$, so erfolgt eine Energieabgabe $E_n - E_{n-1}$, die einer elektromagnetischen Strahlung der Frequenz

$$\nu = \frac{1}{h}(E_n - E_{n-1}) \quad (1.87)$$

entspricht. Das entspricht der berühmten **Quantenhypothese von Planck** für den harmonischen Oszillator. Und die diesen quantenmechanisch möglichen Energieniveaus E_n für stationäre Zustände eineindeutig (das gilt, wenn die ψ_n normiert sind:

$\int_{-\infty}^{+\infty} |\psi_n(x)|^2 dx = 1$) zugeordneten Eigenfunktionen (1.83) nimmt man als mathematisches Äquivalent für den quantenphysikalischen stationären Zustand „Oszillator mit Gesamtenergie E_n “.

Schwingt das Teilchen mit der Energie E_n , so sagen wir, der harmonische Oszillator befinde sich – als quantenmechanisches System – im Zustand $\psi_n(x)$. Man kann auch sagen, daß die eindimensionalen Unterräume des $L^2_{\mathbf{R}}(\mathbf{R})$, die von den ψ_n erzeugt werden, die möglichen Zustände des quantenmechanischen Systems repräsentieren.

Wir wollen noch einen anderen Sachverhalt der Quantenmechanik am Oszillator funktional-analytisch deuten. Die normierten Eigenfunktionen $\psi_n(x)$ von (1.83) bilden ein vollständiges ONS [s. (1.14) und 2.4.2.]

$$\int_{-\infty}^{+\infty} \psi_n(x) \psi_m(x) dx = \delta_{mn} = \begin{cases} 1, & \text{falls } m = n, \\ 0, & \text{falls } m \neq n. \end{cases} \quad (1.88)$$

Dann kann man (s. 2.4.2.) jedes Element $g \in L^2_{\mathbf{R}}(\mathbf{R})$ in eine Reihe

$$g(x) = \sum_{\nu=0}^{\infty} c_{\nu} \psi_{\nu}(x) \quad (1.89)$$

entwickeln, wobei (1.89) in folgendem Sinne gilt:

$$\lim_{n \rightarrow \infty} \int_{-\infty}^{+\infty} |g(x) - \sum_{\nu=0}^n c_{\nu} \psi_{\nu}(x)|^2 dx = 0 \quad (1.90)$$

(die Reihe **konvergiert im quadratischen Mittel**).¹⁾ Die Koeffizienten dieser Reihenentwicklung kann man leicht berechnen, wenn unter Benutzung des **Skalarprodukts** [s. (1.5)]

$$\int_{-\infty}^{+\infty} g(x) h(x) dx = \langle g | h \rangle \quad (h, g \in L^2_{\mathbf{R}}(\mathbf{R})) \quad (1.91)$$

¹⁾ Das benutzte Integral ist das **Lebesgue-Integral**. Denn auch wenn g und die ψ_{ν} glatte Funktionen sind, muß (1.89) nur „bis auf eine Menge vom Maß 0“ (s. 1.3.) erfüllt sein.

wie folgt gerechnet wird: Man bildet (1.91) mit g von (1.89) und mit einem beliebigen ψ_k für h :

$$\begin{aligned}\langle g | \psi_k \rangle &= \left\langle \sum_{\nu=0}^{\infty} c_{\nu} \psi_{\nu} \mid \psi_k \right\rangle \\ &= \sum_{\nu=0}^{\infty} \langle c_{\nu} \psi_{\nu} \mid \psi_k \rangle \\ &= \sum_{\nu=0}^{\infty} c_{\nu} \langle \psi_{\nu} \mid \psi_k \rangle = c_k, \quad k = 0, 1, 2, \dots\end{aligned}\quad (1.92)$$

Damit sind die Koeffizienten bekannt, da die Vertauschung in (1.92) erlaubt ist [vgl. (2.42): das L^2 -Skalarprodukt ist (bezüglich jeder seiner beiden Variablen) stetig].

(1.89) heißt, daß **jedes** $g \in L^2_{\mathbf{R}}(\mathbf{R})$ als (Fourier-) Reihe dargestellt werden kann nach den $\psi_n(x)$, die die stationären Zustände des quantenmechanischen Systems darstellen. Um dies physikalisch auszunutzen, nehmen wir jetzt an, wir hätten die Erweiterung des Operators \mathfrak{H} zu $\bar{\mathfrak{H}}$ schon konstruiert. Das bedeutet, wir kennen den Definitionsbereich $D(\bar{\mathfrak{H}})$ als Teilmenge des $L^2_{\mathbf{R}}(\mathbf{R})$; $D(\bar{\mathfrak{H}})$ umfaßt $D(\mathfrak{H})$ und liegt weiterhin **dicht** in $L^2_{\mathbf{R}}(\mathbf{R})$. Wir betrachten ein beliebiges $g \in D(\bar{\mathfrak{H}})$, welches normiert sei: $\langle g | g \rangle = 1$. Ein solches g wollen wir **auch** als **Zustand** des quantenmechanischen Systems ansehen. Da es erst recht die Darstellung

$$g = \sum_{n=0}^{\infty} c_n \psi_n \quad (1.93)$$

gestattet, wollen wir sagen, daß der quantenmechanische Zustand g beschrieben wird durch Linearkombination (Superposition) der Eigenzustände ψ_n . Dabei gestatten die c_n folgende Interpretation: c_n ist die Wahrscheinlichkeitsamplitude und $|c_n|^2$ die Wahrscheinlichkeit, mit der der Eigenzustand (stationärer Zustand) ψ_n in g enthalten ist. Bei einer Energiemessung findet man dann mit der Wahrscheinlichkeit $|c_n|^2$ die Energie E_n .

Wir berechnen $\sum |c_n|^2$. Da das Orthonormalsystem $\{\psi_n(x)\}$ vollständig ist, gilt die **Vollständigkeitsrelation** [s. (2.66)], und es ist

$$\sum |c_n|^2 = \|g\|^2 = 1. \quad (1.94)$$

Die Wahrscheinlichkeit, irgend eines der Energieniveaus E_n zu messen, ist also 1. Man erhält somit mit Wahrscheinlichkeit 1 bei einer Energiemessung einen der Eigenwerte des Operators $\bar{\mathfrak{H}}$. [Diese einfache Interpretation geht nur, da \mathfrak{H} (und auch $\bar{\mathfrak{H}}$) ein reines Punktspektrum mit einfachen Eigenwerten hat.]

Die **mathematische Erwartung** (der quantenmechanische Mittelwert) bei der Energiemessung wäre dann (Bd. 17, Def. 2.28)

$$\sum_{n=0}^{\infty} \lambda_n |c_n|^2 \quad (1.95)$$

und ist endlich, da diese Reihe konvergent ist [36]. Es besteht ein enger Zusammenhang zur quadratischen Form $\langle \bar{\mathfrak{H}} g | g \rangle$ ($g \in D(\bar{\mathfrak{H}})$):

$$\langle \bar{\mathfrak{H}} g | g \rangle = \sum_{n=0}^{\infty} \lambda_n |c_n|^2. \quad (1.95')$$

Die **quadratische Form** $\langle \bar{\mathfrak{H}} g | g \rangle$ erweist sich im Definitionsbereich von $\bar{\mathfrak{H}}$ gerade als der quantenmechanische Mittelwert bei der Energiemessung im Zustand g . Auf Beziehungen zur Streuung gehen wir unten in 5.1.3. ein.

1.2.4. Ein volkswirtschaftliches Verflechtungsmodell als Fixpunktproblem

Im einfachsten volkswirtschaftlichen Verflechtungsmodell werden in einem bestimmten Zeitraum N Produktionszweige betrachtet, die je eine Produktart produ-

zieren. Der i -te Zweig produziere von seinem Produkt die Menge x_i , an den j -ten Zweig liefere er die Menge p_{ij} und an äußere Bedarfsträger die Menge a_i . Es sei p_{ij} der Gesamtproduktion des j -ten Zweiges proportional:

$$p_{ij} = m_{ij}x_j, \quad (1.96)$$

dann ergibt sich folgendes Gleichungssystem als Bilanz:

$$x_i = \sum_{j=1}^N m_{ij}x_j + a_i \quad (i = 1, \dots, N). \quad (1.97)$$

Entsprechend A) in 1.2.1. lautet die Aufgabe: Gesucht ist ein Vektor $x = (x_1, \dots, x_N) \in R^N$, so daß bei gegebener Verflechtungsmatrix $M = (m_{ij})$ und gegebenem output $a = (a_1, \dots, a_N) \in R^N$ folgende Gleichung erfüllt ist:

$$x = Mx + a. \quad (1.98)$$

Wir formulieren jetzt diese Aufgabe funktionalanalytisch entsprechend B) in 1.2.1. In (1.98) kommt die gesuchte Größe x sowohl „rechts“ vor als auch „isoliert“ auf der linken Seite. Wir fassen die rechte Seite in (1.98) als **Abbildungsvorschrift** auf, einem Vektor $x \in R^N$ den Vektor $Mx + a$ als Bild zuzuordnen. Wir sehen, daß (1.98) gelöst ist, wenn wir einen solchen Vektor $x^* \in R^N$ haben, daß das Bild $Mx^* + a$ von x^* gerade wieder x^* ist, also $Mx^* + a = x^*$ gilt. Ein solcher Vektor $x^* \in R^N$ heißt ein **Fixpunkt** der beschriebenen Abbildung. Ist allgemein A eine Abbildung, die einem Vektor $x \in R^N$ den Vektor $y = A(x) \in R^N$ zuordnet (man sagt auch, daß A den R^N „in sich“ abbildet), so heißt ein Vektor $x^* \in R^N$ ein Fixpunkt von A , wenn $A(x^*) = x^*$ ist.

Jetzt sind zwei Fragen zu beantworten: Welche Forderungen müssen an A gestellt werden, damit ein Fixpunkt existiert (die Abbildung A mit $A(x) = x + 1$ hat keinen Fixpunkt für $x \in \mathbf{R}$), und, wie kann ein existierender Fixpunkt berechnet werden? Eine Antwort gibt der sehr allgemeine Fixpunktsatz von Banach (s. 4.3.2.). In ihm ist A „kontrahierend“. Eine Abbildung A von R^N in sich heißt kontrahierend, wenn es eine Konstante k gibt mit $0 < k < 1$, so daß für je zwei Elemente $x, y \in R^N$ gilt (d sei die Metrik in R^N):

$$d(A(x), A(y)) \leq kd(x, y) \quad (x, y \in R^N). \quad (1.99)$$

(1.99) bedeutet, daß der Abstand der Bilder zweier Elemente aus R^N kleiner oder gleich ist dem mit einem festen Faktor k ($0 < k < 1$) multiplizierten Abstand der Urbilder. Da in (1.99) nur die Metrik d von R^N gebraucht wird, kann eine kontrahierende Abbildung auch in allgemeinen metrischen Räumen E (oder Teilmengen E_0 von E) definiert werden [vgl. (4.83)]. In unserem Beispiel ist $E_0 = E = (R^N, d)$, $Ax = Mx + a$, $x \in R^N$. Unten geben wir vier Möglichkeiten an, einen Faktor k für $A = M \cdot + a$ zu berechnen. Ist dann $0 < k < 1$ erfüllt, so weiß man, daß genau ein Fixpunkt x^* existiert, denn es gilt

Satz 1.2 (Fixpunktsatz von Banach): Eine kontrahierende Abbildung A einer nicht-leeren abgeschlossenen¹⁾ Teilmenge E_0 eines vollständigen metrischen Raumes E in sich hat genau einen Fixpunkt. Dieser kann mit folgendem Iterationsverfahren berechnet werden: Man wählt ein Anfangselement $x_0 \in E_0$. Dann bestimmt man sukzessive $x_1 = A(x_0)$, $x_2 = A(x_1)$, ..., $x_k = A(x_{k-1})$, ..., und die so gewonnene Folge $\{x_j\}$ ($j = 1, 2, \dots$) konvergiert gegen den Fixpunkt x^* . S.1.2

¹⁾ Vgl. (2.4).

Wir realisieren C) von 1.2.1. und geben vier Möglichkeiten zur Bestimmung eines Kontraktionsfaktors k für (1.99) an. Dazu sei $\|\cdot\|$ eine Norm in R^N . Folglich gilt mit der durch diese Norm festgelegten Metrik d :

$$d(A(x), A(y)) = \|A(x) - A(y)\| = \|(Mx + a) - (My + a)\| = \|M(x - y)\|. \quad (1.100)$$

Wir setzen $x - y = z \in R^N$, $z = (z_1, \dots, z_n)$.

a) Wahl der euklidischen Norm im R^N (s. Bsp. 1.2):

$$\begin{aligned} \|Mz\|_1^2 &= \sum_i \left(\sum_k m_{ik} z_k \right)^2 \leq \sum_i \left(\sum_k m_{ik}^2 \right) \left(\sum_k z_k^2 \right) \\ &= \|z\|_1^2 \left(\sum_i \sum_k m_{ik}^2 \right) = k_1^2 \|z\|_1^2, \quad k_1 = \left(\sum_{i,k} m_{ik}^2 \right)^{1/2}. \end{aligned} \quad (1.101)$$

b) Wahl der Norm $\|z\|_2 = \max_i |z_i|$ ($z \in R^N$):

$$\begin{aligned} \|Mz\|_2 &= \max_i \left| \sum_k m_{ik} z_k \right| \leq \max_k \sum_i |m_{ik}| |z_k| \\ &\leq \max_i \sum_k |m_{ik}| \max_k |z_k| = \left(\max_i \sum_k |m_{ik}| \right) \|z\|_2 \\ &= k_2 \|z\|_2, \quad k_2 = \max_i \sum_k |m_{ik}|. \end{aligned} \quad (1.102)$$

c) Wahl der Norm $\|z\|_3 = \sum_{i=1}^N |z_i|$:

$$\begin{aligned} \|Mz\|_3 &= \sum_i \left| \sum_k m_{ik} z_k \right| \leq \sum_i \sum_k |m_{ik}| |z_k| \\ &= \sum_k |z_k| \sum_i |m_{ik}| \leq \sum_k |z_k| \max_k \sum_i |m_{ik}| \\ &= \|z\|_3 \cdot k_3, \quad k_3 = \max_k \sum_i |m_{ik}|. \end{aligned} \quad (1.103)$$

d) Wahl der euklidischen Norm, aber andere Abschätzung der Matrix: Es gilt nämlich, wenn $\tilde{\lambda}$ der größte Eigenwert der Matrix $M^T M$ ist (es ist stets $\tilde{\lambda} \geq 0$)

$$\|Mz\|_1^2 \leq \tilde{\lambda} \|z\|_1^2, \quad (1.104)$$

und damit ist auch $\sqrt{\tilde{\lambda}}$ eine Konstante (k_4) gemäß (1.99).

Die für eine gegebene Matrix M berechenbaren Zahlen k_1, k_2, k_3, k_4 sind „Matrixnormen“ und heißen „euklidische Norm“ (k_1), „Zeilenbetragssummennorm“ (k_2), „Spaltenbetragssummennorm“ (k_3), „Spektralnorn“ (k_4). Die letztere ist auch die Norm von M als Norm einer linearen beschränkten Abbildung im Sinne der Funktionalanalysis (s. Def. 3.6); [7].

Bemerkung 1.4: Die Zahlen k_1, \dots, k_4 können bei einer gegebenen Matrix verschieden ausfallen.

Bemerkung 1.5: In Bd. 18, 2.3., Bsp. 2.5, wird ein lineares Gleichungssystem mit dem oben angegebenen Iterationsverfahren gelöst. Die dabei berechnete Kontraktionskonstante ist gerade k_3 . Im allgemeinen wird die Fixpunktmethod auf *nichtlineare* Gleichungen angewandt, z. B. werden mit Satz 1.2 im Lehrbuch [14] der Existenzsatz für gewöhnliche Differentialgleichungssysteme (s. auch Bd. 7/1) und auch der große Auflösungssatz (s. auch Bd. 4) bewiesen.

1.2.5. Zeitoptimale Steuerung einer erzwungenen gedämpften Schwingung

In der kybernetischen Betrachtungsweise von Systemen ist es üblich und zweckmäßig, sogenannte **Übertragungsglieder** zu betrachten, die ein **Eingangssignal** $x(t)$ ($t \geq 0$: Zeitparameter) in ein **Ausgangssignal** $y(t)$ verwandeln. Eine Frae der Steuerung entsteht, wenn für das Ausgangssignal bestimmte Vorgaben durch geeignete Wahl des Eingangssignals erzielt werden sollen. Wird verlangt, daß das Ausgangssignal $y(t)$ in kürzester Zeit T_0 einen bestimmten Endwert $y_{\text{end}} = y(T_0)$ erreichen soll, so liegt eine Aufgabe der **zeitoptimalen Steuerung** (s. Bd. 16) vor. Optimalitätsforderungen können aber auch hinsichtlich anderer Kehngrößen gestellt werden, z. B. in bezug auf minimalen Energieaufwand für das Eingangssignal.

Faßt man die Eingangssignale $x(t)$ als Elemente eines linearen Raumes E , die Ausgangssignale als Elemente eines linearen Raumes F auf, so läßt sich die Wirkung eines Übertragungsgliedes mittels eines **Übertragungsoperators** $U: E \rightarrow F$, d. h.:

$$(Ux)(t) = y(t) \quad (0 \leq t < \infty) \quad (1.105)$$

beschreiben. Für die Praxis interessant sind vor allem **lineare** und gleichzeitig **zeit-invariante** Übertragungsglieder. Sie sind durch folgende Eigenschaften gekennzeichnet:

$$1) U(c_1 x_1 + c_2 x_2) = c_1 Ux_1 + c_2 Ux_2 \quad (1.106)$$

für beliebige $x_1, x_2 \in E$; c_1, c_2 bel. Konstanten (Linearität).

$$2) \text{ Ist } y(t) = (Ux)(t) \text{ und ist } \bar{x}(t) = x(t - \tau), \text{ so gilt}$$

$$(U\bar{x})(t) = y(t - \tau) \quad (\text{Zeitinvarianz}). \quad (1.107)$$

Mit anderen Worten, ein **lineares** Übertragungsglied ist ein Umwandlungsmechanismus, für den das Superpositionsgesetz gilt. Die Zeitinvarianz bedeutet, daß das Übertragungsglied auf gleichartige, nur zeitlich gegeneinander verschobene Eingangssignale in gleicher Weise reagiert, abgesehen von einer (gleichgroßen) zeitlichen Verschiebung. Als Beispiel eines solchen Übertragungsgliedes kann das mathematische Modell eines Wassereinzugsgebietes dienen. Dabei bedeutet $x(t)$ die zur Zeit t durch Regen pro Zeiteinheit zugeführte Wassermenge und $y(t)$ die zur Zeit t an einem Abflußkanal pro Zeiteinheit abfließende Wassermenge.

Lineare zeitinvariante Übertragungsglieder lassen sich unter geeigneten Voraussetzungen durch einen Übertragungsoperator der folgenden Gestalt beschreiben:

$$(Ux)(t) = y(t) = \int_0^t h(t - \tau) x(\tau) d\tau \quad (t \geq 0). \quad (1.108)$$

Dabei ist $h(t)$ die sog. **Impulsantwortfunktion**, die als Reaktion (= Ausgangssignal) auf einen **Nadelimpuls** $x(t) = \delta(t)$ (δ -Distribution, s. 1.2.2. oder 4.1.) auftritt. Ist die Impulsantwortfunktion $h(t)$ bekannt, so ergibt sich für stetige Eingangssignale $x(t)$ aus Linearität, Zeitinvarianz und Stetigkeit (in einem geeigneten Sinne) des Übertragungsoperators die Darstellung (1.108) [s. auch (1.62)].

Bemerkung 1.6: In der Kybernetik bezeichnet man die Impulsantwortfunktion auch als **Stoßantwort** bzw. (bei geeigneter Festlegung der Dimension) als **Gewichtsfunktion** (s. [37, S. 857]).

Wir gehen im folgenden von der Darstellung (1.108) des Übertragungsoperators aus. Als Beispiel eines solchen Übertragungsgliedes betrachten wir eine schwingende

Punktmasse, die durch eine Zwangskraft $x(t)$ zu einer gedämpften Schwingung $y(t)$ aus der Ruhelage heraus ($y(0) = 0$, $\dot{y}(0) = 0$) angeregt wird. Das Ausgangssignal $y(t)$ (= Amplitude der Schwingung) bestimmt man aus dem Eingangssignal hierbei durch die Lösung der Differentialgleichung (s. Bd. 7/1)

$$m\ddot{y}(t) + \varrho\dot{y}(t) + ay(t) = x(t) \quad (1.109)$$

mit den Anfangsbedingungen $y(0) = 0$, $\dot{y}(0) = 0$. Die Schwingung erfolge schwach gedämpft, d. h., es gelte $\varrho^2 < 4am$. Dann ist (s. Bd. 7/1, S. 96–99, S. 133, Lösung der Aufgabe 3.21) die Lösungsformel

$$y(t) = \int_0^t \frac{1}{\omega m} e^{-\beta(t-\tau)} \sin \omega(t-\tau) x(\tau) d\tau \quad (t \geq 0), \quad (1.110)$$

wobei

$$\beta = \frac{\varrho}{2m} \quad \text{und} \quad \omega = \frac{1}{2m} \sqrt{4am - \varrho^2} \quad (1.111)$$

gilt.

Aus dem Vergleich der Formel (1.110) und der Gl. (1.108) erkennen wir, daß die Antwort $y(t)$ des schwingungsfähigen Systems „Punktmasse an Feder“ auf das Eingangssignal (Zwangskraft) $x(t)$ einem linearen, zeitinvarianten Übertragungsglied entspricht, wobei die Impulsantwortfunktion $h(t)$ durch den speziellen Ausdruck

$$h(t) = \frac{1}{\omega m} e^{-\beta t} \sin \omega t \quad (t \geq 0) \quad (1.112)$$

gegeben ist.

Bemerkung 1.7: Durch Vergleich der Ausdrücke (1.112) oben und den Betrachtungen in Bd. 7/1 (S. 98) erkennen wir, daß die Impulsantwortfunktion $h(t)$ über die Gleichung

$$G(t, \tau) = h(t - \tau)$$

mit der Greenschen Funktion $G(t, \tau)$ der Anfangswertaufgabe (1.109) zusammenhängt (s. auch 1.2.2.).

Im weiteren sei $h(t)$ stetig und in keinem Intervall $[0, t_0]$ identisch gleich null. Es sei nun $t_0 > 0$ ein fester Zeitpunkt. Dann gilt nach (1.108)

$$y(t_0) = \int_0^{t_0} h(t_0 - \tau) x(\tau) d\tau. \quad (1.113)$$

Wir benutzen (der grundlegenden Arbeit [20] folgend) die Schwarzsche Ungleichung (Satz 1.1) im Prä-Hilbertraum $C_{\mathbf{R}}[0, t_0]$ (s. Bsp. 1.4) der auf $[0, t_0]$ stetigen Funktionen mit dem Skalarprodukt $\langle f | g \rangle$ und der zugehörigen Norm:

$$\langle f | g \rangle = \int_0^{t_0} f(\tau) g(\tau) d\tau, \quad \|f\| = \left[\int_0^{t_0} (f(\tau))^2 d\tau \right]^{1/2}.$$

Es gilt wegen (1.113) und der Schwarzschen Ungleichung (vgl. S. 11)

$$|y(t_0)| = \left| \int_0^{t_0} h(t_0 - \tau) x(\tau) d\tau \right| \leq \left[\int_0^{t_0} (h(t_0 - \tau))^2 d\tau \right]^{1/2} \left[\int_0^{t_0} (x(\tau))^2 d\tau \right]^{1/2}. \quad (1.114)$$

Nach der Substitution $s = t_0 - \tau$ ergibt sich

$$\int_0^{t_0} (h(t_0 - \tau))^2 d\tau = - \int_{t_0}^0 (h(s))^2 ds = \int_0^{t_0} (h(s))^2 ds = \|h\|^2. \quad (1.115)$$

Also erhalten wir mittels (1.114)

$$|y(t_0)| = \left| \int_0^{t_0} h(t_0 - \tau) x(\tau) d\tau \right| \leq \|h\| \cdot \|x\|. \quad (1.116)$$

Das Gleichheitszeichen in dieser Ungleichung gilt genau dann (s. Satz 1.1), wenn die Funktionen unter dem Integralzeichen linear abhängig sind. In unserem Fall bedeutet dies, daß

$$x(\tau) = Kh(t_0 - \tau) \quad (0 \leq \tau \leq t_0), \quad (1.117)$$

K eine beliebige Konstante, gilt.

Aus (1.113) ergibt sich für ein $x(\tau)$ dieser Form [vgl. (1.115)]

$$y(t_0) = K \int_0^{t_0} (h(t_0 - \tau))^2 d\tau = K \|h\|^2. \quad (1.118)$$

Ist nun $y(t_0) = y_{\text{end}}$ ein vorgegebener, zu erreichender Wert, so folgt aus (1.118)

$$K = \frac{y_{\text{end}}}{\|h\|^2}, \quad (1.119)$$

und für das zugehörige Eingangssignal, das wir von jetzt ab mit $\hat{x}(\cdot)$ bezeichnen, ergibt sich mittels (1.117)

$$\hat{x}(\tau) = \frac{y_{\text{end}}}{\|h\|^2} h(t_0 - \tau) \quad (0 \leq \tau \leq t_0). \quad (1.120)$$

Es gilt für dieses spezielle Eingangssignal [s. (1.115)]

$$\|\hat{x}\| = \frac{|y_{\text{end}}|}{\|h\|^2} \|h\| = \frac{|y_{\text{end}}|}{\|h\|}. \quad (1.121)$$

Für jedes andere beliebige Eingangssignal $x(\tau)$ gilt nach (1.116) die Ungleichung

$$\frac{|y_{\text{end}}|}{\|h\|} \leq \|x\|, \quad (1.122)$$

wenn wir fordern, daß für dieses Eingangssignal ebenfalls die Beziehung $y(t_0) = y_{\text{end}}$ gilt. Aus (1.121) und (1.122) folgt

$$\|\hat{x}\| \leq \|x\|, \quad (1.123)$$

d. h., das Eingangssignal $\hat{x}(\tau)$ ist von allen Eingangssignalen $x(\tau)$, die zum Zeitpunkt t_0 den gleichen Wert des Ausgangssignals $y(t_0) = y_{\text{end}}$ liefern, dasjenige mit der kleinsten Norm. Da die Norm $\|x\| = \left[\int_0^{t_0} (x(\tau))^2 d\tau \right]^{1/2}$ als ein Maß für die Energie des Eingangssignals angesehen werden kann, lautet unser Ergebnis:

$$\hat{x}(\tau) = \frac{y_{\text{end}}}{\|h\|^2} h(t_0 - \tau)$$

erzielt den Wert y_{end} des Ausgangssignals („Schwingungsausschlag“) zum Zeitpunkt t_0 mit **minimalem Energieaufwand**. Damit ist eine spezielle Aufgabe der optimalen Steuerung gelöst. Uns interessiert im folgenden die Frage, wie das Eingangssignal $x(t)$ gewählt werden muß, um unter der **Energieaufwandsbeschränkung**

$$\|x\|^2 \leq E \quad (E > 0, \text{ gegeben}) \quad (1.124)$$

den vorgegebenen Wert $y(t_0) = y_{\text{end}}$ des Ausgangssignals in kürzester Zeit

$$t_0 \rightarrow \text{Minimum!} \quad (1.125)$$

zu erreichen.

Wir können die obigen Betrachtungen zur Lösung dieser Aufgabe verwenden. Es sei $x(t)$ ein Eingangssignal, welches unter der Bedingung (1.124) zu einem gewissen, nicht notwendig minimalen Zeitpunkt t_0 den geforderten Endwert $y(t_0) = y_{\text{end}}$ liefert. Dann gilt wegen (1.122), (1.124)

$$\|\dot{x}\| = \frac{|y_{\text{end}}|}{\|h\|} \leq \|x\| \leq \sqrt{E}. \quad (1.126)$$

Daraus folgt als **notwendige Bedingung**

$$\frac{|y_{\text{end}}|}{\sqrt{E}} \leq \|h\|. \quad (1.127)$$

Die rechte Seite $\|h\| = \left[\int_0^{t_0} (h(\tau))^2 d\tau \right]^{1/2}$ hängt von dem Endzeitpunkt t_0 ab; wir setzen

$$\alpha(t_0) = \|h\|. \quad (1.128)$$

Wegen (1.112) gilt daher für α :

$$\left. \begin{aligned} \alpha(0) &= 0, \\ 0 \leq t_1 < t_2 &\Rightarrow \alpha(t_1) < \alpha(t_2), \\ \alpha(t_0) &\text{ hängt stetig von } t_0 \text{ ab.} \end{aligned} \right\} \quad (1.129)$$

Es gibt somit genau einen oder keinen kleinsten Wert $t_0 = T_0 = \inf \left\{ t \geq 0 \mid \alpha(t) = \frac{|y_{\text{end}}|}{\sqrt{E}} \right\}$ mit

$$\|h\| = \alpha(T_0) = \frac{|y_{\text{end}}|}{\sqrt{E}}. \quad (1.130)$$

Wir setzen voraus, daß der Wert T_0 existiert. Nach den Eigenschaften (1.129) gilt für $t_0 < T_0$ die Ungleichung

$$\|h\| = \alpha(t_0) < \alpha(T_0) = \frac{|y_{\text{end}}|}{\sqrt{E}},$$

d. h., die notwendige Bedingung (1.127) ist für $t_0 < T_0$ nie erfüllt. Für $t_0 = T_0$ wird die Bedingung (1.127) und die Bedingung $y(t_0) = y_{\text{end}}$ durch die Lösung $\hat{x}(\tau)$ aus (1.120) erfüllt, d. h., T_0 ist die **minimale** Endzeit, zu der unter der Energiebeschränkung (1.124) ein Eingangssignal (eine Steuerung) existiert mit $x(T_0) = y_{\text{end}}$, und dieses Eingangssignal ist eindeutig bestimmt durch die Gleichung

$$x(\tau) = \hat{x}(\tau) = \frac{y_{\text{end}}}{\|h\|^2} h(T_0 - \tau) \quad (0 \leq \tau \leq T_0),$$

welches wegen der Gl. (1.130) die Form

$$\dot{x}(\tau) = \frac{E}{y_{\text{end}}} h(T_0 - \tau) \quad (0 \leq t \leq T_0) \quad (1.131)$$

hat. Also ist die gestellte Aufgabe lösbar [eben mittels des Eingangssignals (1.131)], und die **Minimalzeit** T_0 ermittelt man aus der Gleichung [vgl. (1.130)]

$$\int_0^{T_0} (h(t))^2 dt = \frac{y_{\text{end}}^2}{E} \quad (1.132)$$

(wobei T_0 der kleinste Wert ist, der diese Gleichung erfüllt). Für unser Beispiel der erzwungenen gedämpften Schwingung lautet diese Gl. (1.132)

$$\int_0^{T_0} \frac{e^{-2\beta t}}{\omega^2 m^2} \sin^2 \omega t dt = \frac{y_{\text{end}}^2}{E} \quad (1.133)$$

oder nach ausgeführter Integration und nach Einsetzen der Konstanten

$$\begin{aligned} \frac{y_{\text{end}}^2}{E} &= \frac{1 - e^{-\frac{\varrho}{m} T_0}}{2a\varrho} + \frac{1}{a} \frac{e^{-\frac{\varrho}{m} T_0}}{4am - \varrho^2} \\ &\quad \times (\varrho \sin^2 \omega T_0 + \sqrt{4am - \varrho^2} \sin \omega T_0 \cos \omega T_0). \end{aligned} \quad (1.134)$$

Aus (1.134) ist T_0 zu berechnen (durch ein Näherungsverfahren). Das zugehörige $x(t)$ hat die Form

$$\dot{x}(t) = \frac{E}{y_{\text{end}} \omega m} e^{-\beta(T_0-t)} \sin \omega(T_0 - t) \quad (0 \leq t \leq T_0). \quad (1.135)$$

Es gilt stets die Ungleichung

$$\|h\|^2 = \int_0^{T_0} (h(t))^2 dt \leq \int_0^{\infty} (h(t))^2 dt,$$

vorausgesetzt, daß das rechtsstehende Integral existiert. In unserem Beispiel hat es den Wert

$$\int_0^{\infty} (h(t))^2 dt = \frac{1}{2a\varrho}. \quad (1.136)$$

Zufolge der Gleichheit (1.132) folgt die Ungleichung $\frac{y_{\text{end}}^2}{E} \leq \frac{1}{2a\varrho}$ oder

$$|y_{\text{end}}| \leq \left[\frac{E}{2a\varrho} \right]^{1/2},$$

d. h., ein Endzustand $y_{\text{end}} = y(T_0)$ mit

$$|y_{\text{end}}| > \sqrt{\frac{E}{2a\varrho}} \quad (1.137)$$

ist durch **kein** Eingangssignal $x(\tau)$ mit $\|x\| \leq \sqrt{\frac{E}{2a\varrho}}$ realisierbar! Für große Werte von ϱ und a wird (bei festem E) der Wert von $\sqrt{\frac{E}{2a\varrho}}$ sehr klein, so daß dann bereits

relativ kleine Werte von $|y_{\text{end}}|$ nicht „ansteuerbar“ sind (dies ist von Interesse bei der Dimensionierung von Dämpfungsgliedern). Für $E = +\infty$ ist (1.137) nie erfüllt, d. h., **jeder** Endzustand y_{end} ist erreichbar, wenn beliebig viel Energie zur Verfügung steht.

Alle Betrachtungen lassen sich mit geringfügigen Modifikationen auch auf andere lineare zeitinvariante Systeme übertragen.

1.3. Meßbare Funktionen, Lebesgue-Integral

An den verschiedensten Stellen dieses Bandes sind die Grundbegriffe der Theorie des Lebesgue-Integrals unumgänglich. Sie werden hier – ausschließlich auf der Grundlage der vorangehenden Bände – nur soweit entwickelt, wie es die weiteren Darlegungen erfordern. Der Leser, der sich über dieses notwendige Minimum hinaus weiter informieren möchte, sei auf die Darstellungen in [27], [35], [19], [14] verwiesen.

Im einzelnen gehen wir nur auf die Theorie im R^1 ein. Unsere Darstellung ist so gewählt, daß der Übergang zur Lebesgue-Integration im R^n unmittelbar möglich ist. Zum Teil ist der hier gewählte Zugang an [31] bzw. an [33] orientiert, unterscheidet sich aber hinsichtlich der gewählten Funktionenmenge, von der man ausgeht (stetige Funktionen anstelle von Treppenfunktionen).

D.1.6 Definition 1.6: Es sei $[a, b]$ ein Intervall der Zahlengeraden, ($a < b$). Eine Teilmenge $A \subseteq [a, b]$ dieses Intervalls heißt eine Menge vom **Maß Null**, wenn es zu jedem $\varepsilon > 0$ eine Folge offener Intervalle J_n gibt ($J_n = (a_n, b_n)$), deren Gesamtlängensumme nicht größer ist als ε und deren Vereinigungsmenge die Menge A enthält:

$$\sum_{n=1}^{\infty} (b_n - a_n) \leq \varepsilon, \quad (1.138)$$

$$A \subseteq \bigcup_{n=1}^{\infty} J_n. \quad (1.139)$$

Beispiel 1.13: Eine Menge $A \subseteq [a, b]$, die aus endlich vielen Punkten x_1, \dots, x_n besteht, hat das Maß Null. Eine Menge $A = \{x_1, \dots, x_n, \dots\} \subseteq [a, b]$, die sich als Folge schreiben läßt (eine sog. abzählbar unendliche Menge), hat das Maß Null. Zum Beweis wählt man (es genügt, den zweiten Fall zu betrachten) zum gegebenen $\varepsilon > 0$ die Intervalle $J_n = (a_n, b_n)$ in der Form

$$J_n = \left(x_n - \frac{\varepsilon}{2^{n+1}}, x_n + \frac{\varepsilon}{2^{n+1}} \right) \quad (n = 1, 2, \dots).$$

Dann gilt $\sum_{n=1}^{\infty} (b_n - a_n) = \sum_{n=1}^{\infty} \frac{\varepsilon}{2^n} = \varepsilon \sum_{n=1}^{\infty} \frac{1}{2^n} = \varepsilon \frac{1}{1 - \frac{1}{2}} \frac{1}{2} = \varepsilon$ sowie $A = \{x_1, \dots, x_n, \dots\} \subseteq \bigcup_{n=1}^{\infty} J_n$.

Also sind die Bedingungen (1.138) und (1.139) für jedes $\varepsilon > 0$ erfüllbar. Somit hat A das Maß Null. Analog zu Bsp. 1.13 kann man zeigen, daß die Vereinigungsmenge einer Folge von Mengen vom Maß Null wieder das Maß Null hat. Das Intervall $[a, b]$ ($a < b$) hat andererseits nicht das Maß Null.

D.1.7 Definition 1.7: Es gelte eine gewisse Eigenschaft $P = P(x)$ für alle Punkte x des Intervalls $[a, b]$ mit Ausnahme der Punkte x , die zu einer Menge vom Maß Null gehören. Dann sagt man: **Die Eigenschaft P gilt für fast alle $x \in [a, b]$ oder: P gilt in $[a, b]$ fast überall.**

Zum Beispiel kann man durch Verwendung dieser Definition sagen, daß es in der Integrationstheorie auf Mengen vom Maß Null nicht ankommt, sie sind „vernachlässigbar klein“.

Definition 1.8: Eine (reellwertige) Funktion $f(x)$, definiert auf dem Intervall $[a, b]$, heißt **meßbare Funktion**, wenn es eine Folge stetiger Funktionen $\{g_n(x)\}$ ($a \leq x \leq b$) gibt, die fast überall in $[a, b]$ gegen $f(x)$ konvergiert; m. a. W., es gilt $f(x) = \lim_{n \rightarrow \infty} g_n(x)$ $f x \in [a, b] \setminus A$), wobei A eine Menge vom Maß Null ist. (Die Menge A hängt von (x) ab.) **D.1.8**

Man überzeugt sich leicht davon, daß die Summe zweier meßbarer Funktionen wieder eine meßbare Funktion ist und daß die Multiplikation mit einem (reellen) Zahlenfaktor nicht aus dem Bereich der meßbaren Funktionen herausführt, m. a. W., die meßbaren Funktionen bilden einen linearen Raum. Man bezeichnet ihn mit $S[a, b]$ oder mit $L_0[a, b]$.

Definition 1.9 (Lebesgue-Integral für beschränkte meßbare Funktionen): Es sei $f(x)$ eine (reellwertige) beschränkte meßbare Funktion auf $[a, b]$; d. h., es gibt ein $M > 0$ mit $|f(x)| \leq M$ für $a \leq x \leq b$. Ist $\{g_n(x)\}$ eine Folge stetiger Funktionen, die auf $[a, b]$ fast überall gegen $f(x)$ konvergiert (Def. 1.7), so setzen wir **D.1.9**

$$\int_a^b f(x) dx = \lim_{n \rightarrow \infty} \left(\int_a^b g_n(x) dx \right). \quad (1.140)$$

Hierbei stehen rechts gewöhnliche Riemann-Integrale für stetige Funktionen (s. Bd. 2).

Bemerkung 1.8: Die obige Definition 1.9 ist korrekt, da gezeigt werden kann, daß der Grenzwert (1.140) *unabhängig* von der gewählten Folge $(g_n(x))$ ist.

Definition 1.10 (Nichtnegative summierbare Funktionen): Es sei $f(x)$ eine nichtnegative meßbare Funktion: $f(x) \geq 0$ ($a \leq x \leq b$). Wir bilden die Folge beschränkter meßbarer Funktionen **D.1.10**

$$f_n(x) = \begin{cases} f(x) & x \in [a, b] \text{ und } 0 \leq f(x) \leq n, \\ n & x \in [a, b] \text{ und } n < f(x). \end{cases} \quad (n = 1, 2, \dots)$$

Die Funktion $f(x)$ heißt (**Lebesgue-**) **summierbar** über $[a, b]$, wenn die Folge der Integrale $\left\{ \int_a^b f_n(x) dx \right\}$ nach oben beschränkt ist. Man setzt

$$\int_a^b f(x) dx = \lim_{n \rightarrow \infty} \int_a^b f_n(x) dx \left(= \sup_n \int_a^b f_n(x) dx \right) \quad (1.141)$$

und bezeichnet diesen Ausdruck als das **Lebesgue-Integral** von $f(x)$ über $[a, b]$ ¹⁾.

Definition 1.11 (Summierbare Funktionen beliebigen Vorzeichens): Es sei $f(x)$ eine (reellwertige) meßbare Funktion auf $[a, b]$. Wenn es zwei nichtnegative summierbare Funktionen $f_1(x), f_2(x)$ gibt, für die **D.1.11**

$$f(x) = f_1(x) - f_2(x) \quad (x \in [a, b]) \quad (1.142)$$

¹⁾ Henri Lebesgue 1875–1941

gilt, so heißt $f(x)$ **summierbar** ((L)-summierbar) **über** $[a, b]$. Die Zahl

$$\int_a^b f(x) dx = \int_a^b f_1(x) dx - \int_a^b f_2(x) dx \quad (1.143)$$

wird das (L)-Integral von $f(x)$ über $[a, b]$ genannt ((L)- bedeutet Lebesgue-).

Bemerkung 1.9: 1) Der Leser zeige als Übung, daß der Wert (1.143) des Integrals einer summierbaren Funktion von der speziellen Darstellung (1.142) unabhängig ist. 2) Statt „summierbar“ sagt man gelegentlich auch „integrierbar“.

Bemerkung 1.10: Komplexwertige summierbare Funktionen $f(x)$ erhält man genau in der Form

$$f(x) = u(x) + iv(x) \quad (a \leq x \leq b), \quad (1.144)$$

wobei $u(x), v(x)$ reellwertige summierbare Funktionen sind.

Das Lebesgue-Integral hat analoge Eigenschaften wie das Riemann-Integral (s. Kap. 2.). Insbesondere bildet die Menge aller (reell- oder komplexwertigen) summierbaren Funktionen einen linearen Raum, und die Zuordnung

$$f \rightarrow \int_a^b f(x) dx \quad (1.145)$$

ist linear, d. h.,

$$\int_a^b (f(x) + g(x)) dx = \int_a^b f(x) dx + \int_a^b g(x) dx, \quad (1.146)$$

$$\int_a^b (\lambda f(x)) dx = \lambda \int_a^b f(x) dx \quad (\lambda \in \mathbf{R} \text{ oder } \lambda \in \mathbf{K}). \quad (1.147)$$

Oben hatten wir Mengen mit dem Lebesgue-Maß Null betrachtet. Meßbare Mengen beliebigen Lebesgue-Maßes lassen sich über die Betrachtung ihrer Indikatorfunktion einführen. Ist $A = [a, b]$, so heißt die durch die Vorschrift

$$\chi_A(x) = \begin{cases} 0 & \text{für } x \notin A, \\ 1 & \text{für } x \in A \end{cases} \quad (a \leq x \leq b)$$

definierte Funktion χ_A die **charakteristische Funktion (Indikatorfunktion)** von A .

D.1.12 Definition 1.12: Eine Menge $A \subseteq [a, b]$ heißt **meßbar** (Lebesgue-meßbar), wenn ihre Indikatorfunktion χ_A summierbar ist. Die Zahl

$$\int_a^b \chi_A(x) dx \quad (1.148)$$

heißt das **Maß** ((L)-Maß) von A und wird mit $\text{mes } A$ bezeichnet.

Beispiel 1.14: Es sei $A = [c, d] \subseteq [a, b]$. Dann gilt $\text{mes } A = d - c$. Das Lebesgue-Maß ist also eine Verallgemeinerung des elementargeometrischen Längenbegriffs.

Satz 1.3: Sind $A \subseteq [a, b]$ und $B \subseteq [a, b]$ zwei meßbare disjunkte Teilmengen von $[a, b]$, d. h. gilt $A \cap B = \emptyset$, so ist auch $A \cup B$ meßbar, und es gilt (Additivität des Maßes)

$$\text{mes}(A \cup B) = \text{mes } A + \text{mes } B. \quad (1.149)$$

Im Rahmen der bisherigen Betrachtungen war das Intervall $[a, b]$ beliebig, aber fest. Man kann zeigen, daß alle obigen Definitionen, Sätze usw. von der Wahl eines solchen Intervalls unabhängig sind. Man erhält also allgemein beschränkte (L)-meßbare Mengen und summierbare Funktionen auf beschränkten Definitionsintervallen. Die Erweiterung der obigen Begriffe auf den Fall von Funktionen, die auf unbeschränkten Intervallen erklärt sind, wird wiederum mittels eines Grenzüberganges durchgeführt. Nichtnegative summierbare Funktionen erhält man durch die Forderungen:

$$f: (-\infty, +\infty) \rightarrow \mathbf{R}, \quad f(x) \geq 0 \quad (x \in \mathbf{R}),$$

$$f \text{ summierbar über jedes Intervall } [-n, n] \quad (n = 1, 2, \dots),$$

$$\lim_{n \rightarrow \infty} \int_{-n}^n f(x) dx \text{ existiert.}$$

Letzterer Grenzwert wird mit $\int_{-\infty}^{+\infty} f(x) dx$ bezeichnet.

Danach ist es möglich, wie in Def. 1.11 summierbare Funktionen beliebigen Vorzeichens auf $(-\infty, +\infty)$ einzuführen und unbeschränkte meßbare Mengen analog zu Def. 1.12 zu erklären. Die Betrachtungen im k -dimensionalen Raum \mathbf{R}^k verlaufen analog. Es können auch beliebige meßbare Mengen als Definitionsbereiche summierbarer Funktionen verwendet werden.

Schließlich führen wir noch folgenden wichtigen Satz an:

S.1.4

Satz 1.4: Ist $f(x)$ (komplex- oder reellwertig) summierbar über dem Intervall $[a, b]$ und gilt

$$\int_a^b |f(x)| dx = 0,$$

so ist $f(x) = 0$ fast überall in $[a, b]$.

Der große Fortschritt, der mit der Einführung des Lebesgue-Integrals erreicht wurde, liegt vor allem in der Möglichkeit, Grenzübergänge unter sehr allgemeinen Voraussetzungen durchführen zu können. Wir nennen hier nur zwei wichtige Sätze vom Typ „Grenzübergang unter dem Integralzeichen“:

Satz 1.5 (B. Levi): Es sei $\{f_n(x)\}$ eine nicht fallende Folge nichtnegativer summierbarer Funktionen auf dem Intervall $[a, b]$, und es existiere eine (von n unabhängige) reelle Zahl $C > 0$ mit

S.1.5

$\int_a^b f_n(x) dx \leq C \quad (n \leq 1, 2, \dots)$. Dann ist die Funktion $f(x) = \lim_{n \rightarrow \infty} f_n(x)$ über $[a, b]$ summierbar, und es gilt

$$\int_a^b f(x) dx = \int_a^b \left(\lim_{n \rightarrow \infty} f_n(x) \right) dx = \lim_{n \rightarrow \infty} \int_a^b f_n(x) dx. \quad (1.150)$$

Satz 1.6 (Lebesgue): Es sei $\{f_n(x)\}$ eine Folge summierbarer Funktionen auf $[a, b]$, die dort fast überall gegen eine (meßbare) Funktion $f(x) = \lim_{n \rightarrow \infty} f_n(x)$ ($x \in [a, b]$) mit evtl. Ausnahme einer Menge vom Maß Null) konvergiert. Es existiere eine über $[a, b]$ summierbare Funktion $g(x) \geq 0$ mit $|f_n(x)| \leq g(x)$ ($x \in [a, b]$) für $n = 1, 2, \dots$. Dann gilt (1.150).

S.1.6

2. Räume

2.1. Vollständige metrische Räume, Banachräume

2.1.1. Konvergenz von Folgen in metrischen Räumen. Abgeschlossene und offene Mengen. Vollständigkeit und Kompaktheit

Da zahlreiche Probleme zur Berechnung praxiswichtiger Größen, die bei der mathematischen Behandlung angewandter Aufgaben vorkommen, erfahrungsgemäß näherungsweise gelöst werden müssen, ist es erforderlich, exakte Maßstäbe an den Begriff „näherungsweise“ anzulegen. Dies geschieht erstens durch den in Kap. 1. eingeführten Begriff einer *Metrik*, die imstande ist, den Unterschied zwischen exakter Lösung und Näherungslösung zu erfassen, und zweitens mittels des Begriffs einer (bezüglich der gegebenen Metrik) *konvergenten Folge*. Die Existenz einer Folge von Näherungslösungen, die gegen die exakte Lösung konvergiert, sichert, daß die Näherungslösung *beliebig genau* gewählt werden kann.

D.2.1 Definition 2.1: Es sei (X, d) ein metrischer Raum (s. Def. 1.1). Eine Folge $\{f_n\}$ von Elementen aus X heißt **konvergent**, wenn ein Element $f \in X$ existiert mit

$$\lim_{n \rightarrow \infty} d(f_n, f) = 0, \quad (2.1)$$

d. h., wenn die (Zahlen-)folge der Abstände zwischen f und f_n eine Nullfolge bildet. Das Element f heißt der **Grenzwert** der Folge $\{f_n\}$, in Zeichen

$$f = \lim_{n \rightarrow \infty} f_n. \quad (2.2)$$

Bemerkung 2.1: Daß wir in obiger Definition von **dem** Grenzwert f einer Folge $\{f_n\}$ sprechen, ist dadurch begründet, daß eine konvergente Folge nur einen einzigen Grenzwert besitzt.

Gäbe es nämlich noch ein weiteres Element $g \in X$ mit $\lim_{n \rightarrow \infty} d(g, f_n) = 0$, so folgt mittels der Dreiecksungleichung (M3) und mittels (M2) (s. Def. 1.1) die Ungleichung

$$0 \leq d(f, g) \leq d(f_n, f) + d(f_n, g) \quad (n = 1, 2, \dots). \quad (*)$$

Führen wir in (*) den Grenzübergang $n \rightarrow \infty$ durch, so folgt, da sich Ungleichungen der Form „ \leq “ zwischen den Elementen konvergenter Zahlenfolgen auf deren Grenzwerte übertragen,

$$0 \leq d(f, g) \leq 0 + 0 = 0,$$

also ist $d(f, g) = 0$ und somit [nach (M1)] $f = g$.

D.2.2 Definition 2.2: Es sei $A \subseteq X$ eine Teilmenge des metrischen Raumes (X, d) . Die Menge A heißt **abgeschlossen**, wenn aus $f = \lim_{n \rightarrow \infty} f_n$ und $f_n \in A$ ($n = 1, 2, \dots$) folgt, daß auch $f \in A$ gilt. Mit anderen Worten, eine Menge A heißt abgeschlossen, wenn der Grenzwert jeder konvergenten Folge von Elementen aus A ebenfalls zu A gehört. Ist $A = \emptyset$ (leere Menge), so wird A definitionsgemäß als abgeschlossen bezeichnet.

Beispiel 2.1: Es sei $X = \mathbf{R}$ und $d(x, y) = |x - y|$ ($x, y \in \mathbf{R}$). Es sei A_1 die Menge aller rationalen Zahlen, A_2 die Menge $A_2 = [0, 1] = \{x \in \mathbf{R} \mid 0 \leq x \leq 1\}$. Die Menge A_1 ist nicht abgeschlossen; die Menge A_2 ist abgeschlossen (Beweis als Aufgabe).

Definition 2.3: Es sei $f \in X$ ein Element des metrischen Raumes (X, d) und r eine positive reelle Zahl. Weiter sei (s. Bd. 1, 7.8.) **D.2.3**

$$B(f; r) = \{g \in X \mid d(f, g) < r\},$$

$$K(f; r) = \{g \in X \mid d(f, g) \leq r\},$$

$$S(f; r) = \{g \in X \mid d(f, g) = r\},$$

dann heißt $B(f; r)$ die **offene Kugel**, $K(f; r)$ die **abgeschlossene Kugel** mit dem **Mittelpunkt** f und dem **Radius** r sowie $S(f; r)$ die **Kugeloberfläche** (Sphäre) dieser Kugel (vgl. Bsp. 2.9).

Bemerkung 2.2: Die Mengen $S(f; r)$ und $K(f, r)$ sind abgeschlossene Mengen im Sinne der Def. 2.2. Dies rechtfertigt die Bezeichnung „abgeschlossene Kugel“ für die Menge $K(f; r)$.

Definition 2.4: Es sei $G \subseteq X$ eine Teilmenge des metrischen Raumes (X, d) . Die Menge G heißt **offen**, wenn sie mit jedem ihrer Elemente f eine Kugel $B(f; r)$ enthält. Mit anderen Worten, G heißt **offen**, wenn es zu jedem $f \in G$ ein $r = r(f) > 0$ gibt mit $B(f; r) \subseteq G$. Die leere Menge \emptyset ist nach Definition eine offene Menge. **D.2.4**

Beispiel 2.2: Ist $r_0 > 0$, so ist die Menge $B(f_0; r_0)$ für jedes $f_0 \in X$ eine offene Menge (dies rechtfertigt die Bezeichnung „offene Kugel“).

Beweis: Ist $f \in B(f_0; r_0)$ ein beliebiges Element von $B(f_0; r_0)$, so gilt $d(f, f_0) < r_0$. Wir setzen $r = r_0 - d(f, f_0)$. Es gilt $r > 0$. Ist $h \in B(f; r)$ beliebig, so ist $d(h, f) < r$. Nach der Dreiecksungleichung (M3) folgt

$$d(h, f_0) \leq d(h, f) + d(f, f_0) < d(f, f_0) + r = d(f, f_0) + r_0 - d(f, f_0) = r_0.$$

Also ist $d(h, f_0) < r_0$, und somit gilt $h \in B(f_0, r_0)$. Jedes Element von $B(f; r)$ gehört somit zu $B(f_0; r_0)$, somit ist $B(f; r) \subseteq B(f_0, r_0)$. Es gibt also zu jedem Element f von $B(f_0; r_0)$ eine (offene) Kugel mit Mittelpunkt f , die ganz zu $B(f_0; r_0)$ gehört. Daher ist $B(f_0; r_0)$ eine offene Menge. ■

Beispiel 2.3: Es sei $X = \mathbf{R}$ und $d(x, y) = |x - y|$ ($x, y \in \mathbf{R}$). Die Menge $[0, 1]$ ist nicht offen, ebenso sind die Mengen $(0, 1) = \{x \in \mathbf{R} \mid 0 < x \leq 1\}$ und $[0, 1) = \{x \in \mathbf{R} \mid 0 \leq x < 1\}$ keine offenen Mengen (Beweis als Übungsaufgabe).

Der Zusammenhang zwischen den Begriffen „abgeschlossene Menge“ und „offene Menge“ wird durch den folgenden Satz geklärt:

Satz 2.1: Eine Teilmenge F eines metrischen Raumes (X, d) ist genau dann abgeschlossen, wenn die Komplementärmenge $G = X \setminus F$ eine offene Menge ist. **S.2.1**

(Beweis z. B. in [17].)

Die folgende Aussage über die Eigenschaften der Gesamtheit aller offenen bzw. abgeschlossenen Mengen eines metrischen Raumes gilt wie für Mengen des euklidischen \mathbf{R}^n (s. Bsp. 1.3).

Satz 2.2: A) Die Vereinigungsmenge beliebig vieler, der Durchschnitt je endlich vieler offener Mengen eines metrischen Raumes X sind stets (wieder) offene Mengen. **S.2.2**

B) Der Durchschnitt beliebig vieler, die Vereinigungsmenge je endlich vieler abgeschlossener Mengen eines metrischen Raumes X sind stets (wieder) abgeschlossene Mengen.

C) Die leere Menge \emptyset und der ganze Raum X sind beide sowohl abgeschlossene als auch offene Mengen.

Bemerkung 2.3: Die Eigenschaften A) und C) [bzw. B) und C)] sind der Ausgangspunkt der Theorie der topologischen Räume (vgl. [17]), die die Theorie der metrischen Räume als Spezialfall enthält.

Jede Teilmenge A eines metrischen Raumes X besitzt eine abgeschlossene Obermenge, z. B. X selbst. Der Durchschnitt aller die Menge A enthaltenden abgeschlossenen Mengen ist nach dem vorhergehenden Satz, Aussage B) selbst wieder abgeschlossen und ist (nach Definition des Durchschnitts) in jeder A enthaltenden abgeschlossenen Menge enthalten; also ist dieser Durchschnitt die kleinste abgeschlossene Menge, welche die gegebene Menge A enthält. Man bezeichnet diese Menge auch als *Abschließung* von A .

D.2.5 Definition 2.5: Es sei (X, d) ein metrischer Raum und $A \subseteq X$ eine Teilmenge von X . Die Menge $\bar{A} = \bigcap \{F \mid F \text{ abgeschlossen und } A \subseteq F\}$ bezeichnet man als **Abschließung** (abgeschlossene Hülle von A).

S.2.3 Satz 2.3: Es sei A eine Teilmenge des metrischen Raumes X . Ein Element f von X gehört genau dann zu \bar{A} , wenn es eine Folge $\{f_n\}$ aus A gibt mit $\lim_{n \rightarrow \infty} f_n = f$.

Die Abschließung einer Menge A stimmt also mit der Menge aller Grenzwerte konvergenter Folgen von Elementen aus A überein. Stets gilt die Enthaltenseinbeziehung

$$A \subseteq \bar{A}. \quad (2.3)$$

Eine Menge A ist genau dann abgeschlossen, wenn gilt

$$A = \bar{A}. \quad (2.4)$$

Beispiel 2.4: Es sei $X = \mathbf{R}$ und $d(x, y) = |x - y|$. Die Menge $A = \{x \in \mathbf{R} \mid x = 1/n; n = 1, 2, \dots\}$ ist nicht abgeschlossen, weil $\bar{A} = A \cup \{0\} \supsetneq A$ gilt. (Der Beweis dafür, daß \bar{A} die angegebene Form hat, ergibt sich mittels Satz 2.3; Aufg. für den Leser.)

Hinsichtlich ihrer Konvergenzeigenschaften können sich die Folgen in metrischen Räumen wesentlich unterscheiden. Es erweisen sich die Begriffe Cauchy-Folge, Kompaktheit, Vollständigkeit als sehr nützlich (s. auch Bd. 1):

D.2.6 Definition 2.6: Eine Folge von Elementen $\{f_n\}$ eines metrischen Raumes (X, d) heißt eine **Cauchy-Folge**, wenn es zu jedem $\varepsilon > 0$ ein $n_0 = n_0(\varepsilon)$ gibt mit $d(f_n, f_m) \leq \varepsilon$ für alle $n, m \geq n_0(\varepsilon)$.

S.2.4 Satz 2.4: Jede konvergente Folge ist eine Cauchy-Folge.

Beweis: Die Folge $\{f_n\}$ konvergiere, d. h., es existiert ein $f \in X$ mit: $\lim_{n \rightarrow \infty} d(f_n, f) = 0$. Bei vorgegebenem $\varepsilon > 0$ gibt es daher ein n_0 mit $d(f_n, f) \leq \varepsilon/2$ für alle $n \geq n_0$. Aus der Dreiecksungleichung (M3) erhalten wir für $n, m \geq n_0$

$$d(f_n, f_m) \leq d(f_n, f) + d(f, f_m) \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \quad \blacksquare$$

Die Umkehrung dieses Satzes gilt jedoch nicht! Mit anderen Worten, es gibt metrische Räume, in denen nicht jede Cauchy-Folge konvergiert (s. Bsp. 2.8).

Bemerkung 2.4: Es sei $\{f_n\}$ eine Cauchy-Folge in einem metrischen Raum (X, d) . Die Folge $\{f_n\}$ besitze eine konvergente Teilfolge $\{f_{n_j}\}$. Dann ist $\{f_n\}$ eine konvergente Folge, und es gilt

$$\lim_{n \rightarrow \infty} f_n = \lim_{j \rightarrow \infty} f_{n_j}.$$

Beweis: Wir setzen $f = \lim_{j \rightarrow \infty} f_{n_j}$ und geben ein $\varepsilon > 0$ beliebig vor. Da $\{f_n\}$ eine Cauchy-Folge ist, gibt es ein n_0 mit $d(f_n, f_m) \leq \varepsilon/2$ für $n, m \geq n_0$. Da die Teilfolge $\{f_{n_j}\}$ gegen f konvergiert, gibt es ein j_0 mit $d(f_{n_j}, f) \leq \varepsilon/2$ für $j \geq j_0$. Für $j \rightarrow \infty$ gilt aber $n_j \rightarrow \infty$, und es existiert ein $j_1 \geq j_0$ mit $n_j \geq n_0$ für $j \geq j_1$. Für $n \geq n_0$ gilt dann auf Grund der Dreiecksungleichung (M3) (weil $n_j \geq n_0$ und $j_1 \geq j_0$ ist)

$$d(f_n, f) \leq d(f_n, f_{n_{j_1}}) + d(f_{n_{j_1}}, f) \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

d. h. die Folge $\{f_n\}$ konvergiert gegen f . ■

Definition 2.7: Ein metrischer Raum (X, d) heißt **kompakt**, wenn jede Folge $\{f_n\}$ aus X eine konvergente Teilfolge besitzt. Eine Teilmenge eines metrischen Raumes heißt **kompakt**, wenn sie, als Teilraum aufgefaßt, ein kompakter metrischer Raum ist. **D.2.7**

Definition 2.8: Ein metrischer Raum (X, d) heißt **vollständig**, wenn jede Cauchy-Folge aus X eine konvergente Folge ist. **D.2.8**

Den Zusammenhang zwischen beiden Begriffen liefert

Satz 2.5: Jeder kompakte metrische Raum ist vollständig.

S.2.5

Beweis: Ist $\{f_n\}$ eine Cauchy-Folge des kompakten metrischen Raumes X , so besitzt diese (wegen der Kompaktheit) eine konvergente Teilfolge. Nach Bem. 2.4 ist $\{f_n\}$ daher selbst konvergent. ■

Die Umkehrung dieser Aussage gilt nicht! Dies zeigt

Beispiel 2.5: Der Raum $C[0, 1]$ (s. Bsp. 1.5) ist mit der Metrik $d(f, g) = \max_{0 \leq t \leq 1} |f(t) - g(t)|$ vollständig (s. [17] und Bd. 1: gleichmäßige Konvergenz). Die Folge $\{f_n\}$ mit $f_n(t) = n$ ($0 \leq t \leq 1$) besitzt aber keine gegen ein Element von $C[0, 1]$ konvergierende Teilfolge (Beweis als Aufgabe).

Hinsichtlich des Verhaltens von Teilmengen (als Teilräume aufgefaßt) gelten die folgenden Aussagen:

Satz 2.6: Es sei (X, d) ein vollständiger metrischer Raum. Eine Teilmenge $Y \subseteq X$ von X ist (als Teilraum von X aufgefaßt) genau dann vollständig, wenn Y abgeschlossen ist. **S.2.6**

Satz 2.7: Es sei (X, d) ein kompakter metrischer Raum. Eine Teilmenge $Y \subseteq X$ von X ist (als Teilraum von X aufgefaßt) genau dann kompakt, wenn Y abgeschlossen ist. **S.2.7**

Das einfachste Beispiel eines vollständigen metrischen Raumes ist die Menge \mathbf{R} der reellen Zahlen, versehen mit der Metrik $d(x, y) = |x - y|$. Die Eigenschaft der Vollständigkeit wird hier durch das bekannte Cauchysche Konvergenzkriterium geliefert. Betrachtet man den Teilraum $\mathbf{P} \subseteq \mathbf{R}$ dieses Raumes, der aus allen rationalen Zahlen besteht, so ist dieser Teilraum nicht vollständig.

Der Übergang von den rationalen Zahlen zu den reellen Zahlen, also von einem nicht vollständigen Raum zu einem vollständigen Raum, der den ersteren als *dichte* Teilmenge enthält („jede irrationale Zahl ist Grenzwert einer Folge rationaler Zahlen“), ist ein spezielles Beispiel für einen allgemeinen Sachverhalt, den man „Vervollständigung“ nennt.

D.2.9 Definition 2.9: Es sei (X, d) ein metrischer Raum. Die Teilmenge $A \subseteq X$ heißt **dicht** in X , wenn $\bar{A} = X$ gilt.

Beispiel 2.6: Die Menge \mathbf{P} der rationalen Zahlen liegt dicht in der Menge \mathbf{R} der reellen Zahlen (bezüglich der Metrik $d(x, y) = |x - y|$). Diese Tatsache ergibt sich daraus, daß jede reelle Zahl x als Grenzwert einer Folge rationaler Zahlen x_n (z. B. der nach der n -ten Stelle nach dem Komma abgebrochenen Dezimalbruchentwicklung x_n von x ; $n = 1, 2, \dots$) dargestellt werden kann.

D.2.10 Definition 2.10: Es seien (X, d) und (Y, ϱ) metrische Räume. Der Raum (X, d) heißt (eine) **Vervollständigung** von (Y, ϱ) , wenn folgende Bedingungen erfüllt sind:

- I) (Y, ϱ) ist ein Teilraum von (X, d) ; d. h., Y ist eine Teilmenge von X und $\varrho(f, g) = d(f, g)$ für $f, g \in Y$.
- II) Y ist eine dichte Teilmenge von X ; d. h., die Abschließung (in X !) von Y ist gleich X .
- III) Der Raum (X, d) ist vollständig.

S.2.8 Satz 2.8: Jeder metrische Raum besitzt eine Vervollständigung [17].

Bemerkung 2.5: Immer dann, wenn ein gegebener (nicht vollständiger) metrischer Raum (X_1, d_1) Teilraum eines vollständigen Raumes (X, d) ist, läßt sich eine Vervollständigung von (X_1, d_1) in einfacher Weise angeben. Als eine solche kann man nämlich die Abschließung $\bar{X}_1 \subseteq X$ von X_1 in X , versehen mit der auf X_1 eingeschränkten Metrik d , nehmen.

D.2.11 Definition 2.11: Zwei metrische Räume (X_1, d_1) und (X_2, d_2) heißen **isometrisch**, wenn es eine Abbildung $\varphi: X_1 \rightarrow X_2$ von X_1 auf X_2 (d. h., $\varphi(X_1) = X_2$) gibt mit

$$d_2(\varphi(f), \varphi(g)) = d_1(f, g) \quad (f, g \in X_1). \quad (2.5)$$

Jede solche Abbildung φ heißt eine **Isometrie** von X_1 auf X_2 .

Beispiel 2.7: Es sei $X_1 = X_2 = \mathbf{R}^n$ (bzw. $= K^n$) mit der euklidischen Metrik

$$d(x, y) = \left(\sum_{j=1}^n (\xi_j - \eta_j)^2 \right)^{1/2},$$

und es sei A eine **orthogonale** Matrix, d. h. $A^T A = A A^T = I$ (bzw. A sei eine **unitäre** Matrix $A^* A = A A^* = I$, wobei $A^* = \overline{A^T}$). Dann ist die durch A erklärte Abbildung eine Isometrie von X_1 auf X_2 . Der Beweis dieser Aussage ergibt sich sofort aus der Invarianz des Skalarprodukts bei der Anwendung orthogonaler (bzw. unitärer) Matrizen:

$$\begin{aligned} d(Ax, Ay) &= \langle Ax - Ay | Ax - Ay \rangle^{1/2} = \langle A(x - y) | A(x - y) \rangle^{1/2} \\ &= \langle A^T A(x - y) | x - y \rangle^{1/2} = \langle x - y | x - y \rangle^{1/2} = d(x, y). \end{aligned}$$

S.2.9 Satz 2.9: Je zwei Vervollständigungen eines metrischen Raumes sind isometrisch.

Identifiziert man also isometrische Räume untereinander, so kann man in diesem Sinne von „der“ Vervollständigung eines metrischen Raumes sprechen. Die Vervoll-

stündigung eines metrischen Raumes (X, d) hängt in entscheidendem Maße davon ab, **welche Metrik** auf der Trägermenge X dieses Raumes verwendet wird. Das folgende Beispiel soll dies zeigen: Versehen wir die Menge $X = C_{\mathbf{R}}[-1, 1]$ mit der Metrik $d(f, g) = \max_{-1 \leq t \leq 1} |f(t) - g(t)|$, so erhalten wir einen vollständigen metrischen Raum (X, d) . Verwenden wir aber in $X = C_{\mathbf{R}}[-1, 1]$ die Metrik

$$d_1(f, g) = \left\{ \int_{-1}^1 (f(t) - g(t))^2 dt \right\}^{1/2} \quad (f, g \in X),$$

so ist der entstehende metrische Raum (X, d_1) nicht vollständig und besitzt daher eine von (X, d) verschiedene (nicht zu (X, d) isometrische) Vervollständigung. Der Beweis für die Tatsache, daß (X, d_1) nicht vollständig ist, ergibt sich daraus, daß es in (X, d_1) Cauchy-Folgen gibt, die nicht konvergieren:

Beispiel 2.8: Im Raum $(C_{\mathbf{R}}[-1, 1], d_1)$ definieren wir eine Folge $\{f_n\}$ durch

$$f_n(t) = \begin{cases} 1 & \text{für } -1 \leq t \leq 0, \\ 1 - nt & \text{für } 0 < t \leq \frac{1}{n}, \\ 0 & \text{für } \frac{1}{n} < t \leq 1. \end{cases} \quad (n = 1, 2, \dots)$$

(Der Leser zeichne eine Skizze des Funktionsverlaufes von f_n .) Für $n \leq m$ gilt (kurze Zwischenrechnung)

$$d_1(f_n, f_m) \leq \frac{1}{\sqrt{3n}} \left(\frac{m-n}{m} \right) \leq \frac{1}{\sqrt{3n}} \quad (n = 1, 2, \dots; m \geq n).$$

Da $\frac{1}{\sqrt{3n}}$ ($n = 1, 2, \dots$) eine Nullfolge ist, ist $\{f_n\}$ eine Cauchy-Folge bezüglich d_1 . Es kann gezeigt werden, daß es kein Element $f \in C_{\mathbf{R}}[-1, 1]$ mit $\lim_{n \rightarrow \infty} d_1(f_n, f) = 0$ gibt. Die Cauchy-Folge $\{f_n\}$ ist somit nicht konvergent. Also ist der Raum $(C_{\mathbf{R}}[-1, 1], d_1)$ nicht vollständig. Seine Vervollständigung ist der Raum $L_{\mathbf{R}}^2[-1, 1]$, wobei sich in diesem Raum als Grenzwert der obigen Folge $\{f_n\}$ die Funktion

$$f(t) = \begin{cases} 1 & \text{für } -1 \leq t \leq 0, \\ 0 & \text{für } 0 < t \leq 1 \end{cases}$$

ergibt (s. 2.2.2.).

2.1.2. Banachräume

Definition 2.12: Einen normierten Raum $(E, \|\cdot\|)$ nennt man einen **Banach-Raum** **D.2.12** (auch: **Banachraum** bzw. **(B)-Raum**), wenn er bezüglich der durch die Norm induzierten Metrik $d(f, g) = \|f - g\|$ ($f, g \in E$) ein **vollständiger metrischer Raum** ist.

Bemerkung 2.6: Die Bezeichnung „Banach-Raum“ wurde zu Ehren von **Stefan Banach (1892–1945)** gewählt, dessen Buch „*Theorie des opérations linéaires*“ (1932) den Grundstein für alle weiteren Entwicklungen der Funktionalanalysis in normierten Räumen legte.

Der einfachste Banachraum ist der Raum R^n der n -dimensionalen Vektoren $x = (\xi_1, \dots, \xi_n)$ (bzw. $(\xi_1, \dots, \xi_n)^T$) (ξ_j reell, $j = 1, \dots, n$), versehen mit der **euklidischen Norm**

$$\|x\| = \left\{ \sum_{j=1}^n \xi_j^2 \right\}^{1/2} \quad (2.6)$$

bzw. der komplexe endlichdimensionale Raum K^n , dessen Elemente Vektoren $x = (\xi_1, \dots, \xi_n)$ (bzw. $(\xi_1, \dots, \xi_n)^T$) mit komplexen Koordinaten ξ_j ($j = 1, \dots, n$) sind, versehen mit der **euklidischen Norm**

$$\|x\| = \left\{ \sum_{j=1}^n |\xi_j|^2 \right\}^{1/2} = \left\{ \sum_{j=1}^n \xi_j \bar{\xi}_j \right\}^{1/2}. \quad (2.7)$$

Weitere wichtige konkrete Banachräume werden im Abschnitt 2.2. behandelt. Hier werden nur einige grundsätzliche Konstruktionen in Banachräumen besprochen: äquivalente Normen, Produktraum, Quotientenraum.

D.2.13 Definition 2.13: Es sei E ein linearer Raum (Vektorraum). Auf E seien zwei Normen $\|\cdot\|_1$ und $\|\cdot\|_2$ gegeben. Die Norm $\|\cdot\|_2$ heißt **äquivalent** zur Norm $\|\cdot\|_1$, wenn es Zahlen $m > 0$, $M > 0$ gibt, so daß die Ungleichungen

$$m\|x\|_1 \leq \|x\|_2 \leq M\|x\|_1$$

für alle $x \in E$ gelten.

Der Übergang zu einer äquivalenten Norm bringt oft gewisse Rechenvorteile. Beim Übergang von einer Norm $\|\cdot\|_1$ zu einer äquivalenten Norm $\|\cdot\|_2$ bleiben konvergente Folgen konvergent (genau darauf beruht Bemerkung 1.4, denn es erweist sich, daß in endlichdimensionalen Räumen alle Normen äquivalent sind). Dabei gilt: $(E, \|\cdot\|_2)$ ist genau dann ein Banachraum, wenn $(E, \|\cdot\|_1)$ ein Banachraum ist. Die Äquivalenz von Normen hat die üblichen (vgl. Bd. 1) Eigenschaften einer **Äquivalenzrelation**: Reflexivität, Symmetrie, Transitivität.

Eine lineare Abbildung S eines normierten Raumes E in den normierten Raum F (vgl. Def. 3.1) heißt ein **Normisomorphismus**, wenn S den Raum E auf F abbildet und $\|Sx\|_F = \|x\|_E$ für alle $x \in E$ gilt.

Beispiel 2.9: Im K^n sind folgende Normen (s. auch 1.2.4.) äquivalent:

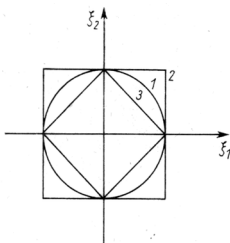


Bild 2.1

$$\begin{aligned} \|x\|_1 &= \left\{ \sum_{j=1}^n |\xi_j|^2 \right\}^{1/2}, \\ \|x\|_2 &= \max_{1 \leq j \leq n} (|\xi_j|), \\ \|x\|_3 &= \sum_{j=1}^n |\xi_j|. \end{aligned} \quad (2.8)$$

Dieselben Normen sind auch im (reellen) Raum R^n äquivalent. Dies erkennt man auch an der Gestalt der Einheitskugeln (für $n = 2$ s. Bild 1.2), d. h. der Mengen $\{x \in R^n \mid \|x\|_j \leq 1\}$ ($j = 1, 2, 3$).

Definition 2.14: Es seien endlich viele normierte Räume $(E_1, \|\cdot\|_1), (E_2, \|\cdot\|_2), \dots, (E_n, \|\cdot\|_n)$ gegeben. Weiter sei $E = E_1 \times E_2 \times \dots \times E_n$ das Produkt der Mengen E_1, \dots, E_n , d. h. die Menge aller n -Tupel $x = (\xi_1, \dots, \xi_n)$ mit $\xi_j \in E_j$ ($j = 1, \dots, n$). Bezüglich der Operationen ($y = (\eta_1, \dots, \eta_n) \in E$)

$$x + y = (\xi_1 + \eta_1, \dots, \xi_n + \eta_n),$$

$$\lambda x = (\lambda \xi_1, \lambda \xi_2, \dots, \lambda \xi_n), \quad \lambda \in \mathbf{R} \text{ bzw. } \lambda \in \mathbf{K},$$

wird E zu einem (reellen bzw. komplexen) Vektorraum. Versehen wir E mit der Norm

$$\|x\| = \sum_{j=1}^n \|\xi_j\|_j \quad (x \in E),$$

so nennen wir den Raum $(E, \|\cdot\|)$ den **Produktraum** der Räume $(E_j; \|\cdot\|_j)$ ($j = 1, \dots, n$).

Bezeichnung: $(E, \|\cdot\|) = \prod_{j=1}^n (E_j; \|\cdot\|_j)$.

Es seien $(E_j, \|\cdot\|_j)$ ($j = 1, \dots, n$) Banachräume. Dann ist auch der Produktraum $(E, \|\cdot\|) = \prod_{j=1}^n (E_j; \|\cdot\|_j)$ ein Banachraum.

Beispiel 2.10: Da der Übergang zu einer äquivalenten Norm nichts an der Struktur eines Banachraumes ändert, können wir mittels der Aussage von Bsp. 2.9 sagen, daß der Raum K^n der (n -fache) Produktraum der Räume K (mit der euklidischen Norm) ist.

Eine weitere Operation, die aus einem gegebenen Banachraum weitere Banachräume zu bilden gestattet, ist die Konstruktion von **Quotientenräumen**. Hierzu betrachten wir zunächst einen beliebigen linearen Raum E sowie einen linearen Teilraum $E_0 \subseteq E$. Wir führen eine (zweistellige) Relation \sim auf E durch die Gleichung

$$x \sim y \Leftrightarrow x - y \in E_0 \quad (2.9)$$

ein. Wie leicht zu sehen ist, erfüllt \sim die Eigenschaften einer Äquivalenzrelation (Reflexivität, Symmetrie und Transitivität). Die Menge

$$[x] = \{y \in E \mid x - y \in E_0\} = \{y \in E \mid x \sim y\} \quad (2.10)$$

bezeichnen wir als die zu x gehörende **Äquivalenzklasse** (Restklasse). In der Menge aller dieser Äquivalenzklassen führen wir die Vektorraumoperationen Addition und Multiplikation mit einem Zahlenfaktor durch die Definitionsgleichungen

$$\begin{aligned} [x] + [y] &= [x + y], \\ \lambda[x] &= [\lambda x] \end{aligned} \quad (x, y \in E; \lambda \in \mathbf{R} \text{ bzw. } \mathbf{K})$$

ein, die, wie man zeigen kann, eindeutig definierte Operationen in der Menge der Äquivalenzklassen liefern und den Axiomen für die Rechenoperationen eines Vektorraumes (s. 1.1.) genügen. Den auf diese Weise entstehenden Vektorraum nennt man den **Quotientenraum von E bezüglich E_0** (bzw. „nach E_0 “) und bezeichnet ihn mit dem Symbol

$$E/E_0. \quad (2.11)$$

Beispiel 2.11. Es sei $E = K^n$ und $E_0 = \{x \in E \mid \langle x \mid x_0 \rangle = 0\}$, wobei $x_0 \in K^n$ ein fest gegebener von 0 verschiedener Vektor aus K^n ist. E_0 ist ersichtlich ein linearer Teilraum von E (Beweis als Übung). Die Menge (Äquivalenzklasse) $[x]$ hat für beliebiges $x \in K^n$ folgendes Aussehen:

$$\begin{aligned} [x] &= \{y \in E \mid x - y \in E_0\} = \{y \in E \mid \langle x - y \mid x_0 \rangle = 0\} \\ &= \{y \in E \mid \langle x \mid x_0 \rangle = \langle y \mid x_0 \rangle\}. \end{aligned}$$

Ist $x \in K^n$ gegeben, so besteht $[x]$ also aus allen Elementen $y \in K^n$, für die das Skalarprodukt $\langle y | x_0 \rangle$ den konstanten Wert $c(x) = \langle x | x_0 \rangle$ hat. In anschaulicher Interpretation heißt dies, daß $[x]$ eine zu E_0 parallele (Hyper-) Ebene darstellt, deren Normalenvektor die Richtung von x_0 hat.

Ist $(E, \|\cdot\|)$ ein normierter Raum und E_0 ein abgeschlossener Teilraum von E , so kann man auf den Quotientenraum E/E_0 durch die Gleichung

$$\|[x]\|_{E/E_0} = \inf \{\|y\| \mid y \in [x]\} \quad (2.12)$$

eine Norm einführen (s. [17], [24]), womit der Quotientenraum E/E_0 zu einem normierten Raum wird.

S.2.10 Satz 2.10: *Ist $(E, \|\cdot\|)$ ein Banachraum, so ist $(E/E_0, \|\cdot\|_{E/E_0})$ ebenso ein Banachraum (Beweis s. [9]).*

2.2. Funktionenräume

Die für die Anwendungen wichtigsten normierten Räume sind **Funktionenräume**, d. h. lineare Räume, deren Elemente (Vektoren) Funktionen (mit reellen bzw. komplexen Werten) sind, die einen gemeinsamen Definitionsbereich besitzen. Ist X dieser gemeinsame Definitionsbereich, so sind die algebraischen Grundoperationen in einem Funktionenraum stets, wie üblich, durch die Gleichungen (punktweise)

$$\begin{aligned} (f+g)(x) &= f(x) + g(x) \quad (x \in X), \\ (\lambda f)(x) &= \lambda f(x) \quad (x \in X; \lambda \in \mathbf{R} \text{ bzw. } \lambda \in \mathbf{K}) \end{aligned}$$

erklärt. Hinzu kommen gewisse Eigenschaften, die die Funktionen des betrachteten Funktionenraumes auszeichnen, wie z. B. Stetigkeits-, Differenzierbarkeits- bzw. Integrierbarkeitseigenschaften. Einen gewissen Sonderfall stellen die **Folgenräume** dar, die als Funktionenräume, bestehend aus Funktionen mit dem gemeinsamen Definitionsbereich \mathbf{N} (Menge der natürlichen Zahlen), aufgefaßt werden können (s. 2.2.4.).

2.2.1. Räume stetiger und stetig differenzierbarer Funktionen

D.2.15 Definition 2.15: *Es sei $K \subseteq \mathbf{R}^n$ eine nichtleere Teilmenge des \mathbf{R}^n . Die Menge aller komplexwertigen (bzw. nur reellwertigen) Funktionen, die auf K stetig sind (s. Bd. 4) bezeichnet man mit $C(K)$ (bzw. mit $C_{\mathbf{R}}(K)$).*

S.2.11 Satz 2.11: *Es sei $K \subseteq \mathbf{R}^n$ nichtleer, abgeschlossen und beschränkt (K ist daher kompakt). Dann ist die Menge $C(K)$ (bzw. $C_{\mathbf{R}}(K)$) versehen mit der üblichen Vektorraumstruktur (s. oben, Beginn von 2.2.) bezüglich der Norm (Supremum-Norm, Maximum-Norm)*

$$\|f\|_{C(K)} = \max_{x \in K} |f(x)| = \sup_{x \in K} |f(x)| \quad (f \in C(K)), \quad (2.13)$$

ein Banachraum.

(Beweis und weitere Einzelheiten in [36].)

Im folgenden sei Ω ein Gebiet des \mathbf{R}^n , d. h. eine offene zusammenhängende¹⁾ Teilmenge des \mathbf{R}^n (z. B. der \mathbf{R}^n selbst). Mit $\partial\Omega = \bar{\Omega} \setminus \Omega$ bezeichnen wir den Rand

¹⁾ Dabei heißt eine Teilmenge des \mathbf{R}^n (allgemeiner: eines metrischen Raumes) zusammenhängend, wenn sie nicht die Vereinigungsmenge zweier disjunkter, nicht leerer, abgeschlossener Mengen ist.

des Gebietes Ω . Der Index \mathbf{R} weist stets darauf hin, daß nur reellwertige Funktionen zugelassen sind.

Ein geordnetes n -Tupel $\alpha = (\alpha_1, \dots, \alpha_n)$ von nichtnegativen ganzen Zahlen nennen wir einen **Multiindex**. Mit $|\alpha|$ bezeichnet man die zugehörige Summe $\sum_{j=1}^n \alpha_j$ der Komponenten des Multiindex α . Die Einführung eines Multiindex dient zur übersichtlichen Schreibweise partieller Ableitungen von Funktionen mehrerer Veränderlicher (s. 4.1.). Man setzt (s. Bd. 4) für eine Funktion f mit

$$f(x) = f(\xi_1, \dots, \xi_n) \quad (x = (\xi_1, \dots, \xi_n)):$$

$$\partial^\alpha f = \frac{\partial^{|\alpha|} f}{\partial \xi_1^{\alpha_1} \dots \partial \xi_n^{\alpha_n}}, \quad x^\alpha = \xi_1^{\alpha_1} \xi_2^{\alpha_2} \dots \xi_n^{\alpha_n}. \quad (2.14)$$

Definition 2.16: Es sei Ω ein beschränktes Gebiet im \mathbb{R}^n und $k = 0, 1, 2, \dots$ eine nicht-negative ganze Zahl. Die Menge aller komplexwertigen Funktionen, die auf der Abschließung $\bar{\Omega} (= \Omega \cup \partial\Omega)$ stetig sind und in Ω stetige partielle Ableitungen bis zur Ordnung k einschließlich besitzen und die sämtlich auf ganz $\bar{\Omega}$ stetig fortgesetzt werden können, bezeichnen wir mit $C^k(\bar{\Omega})$. D.2.16

Bemerkung 2.7: Für $k = 0$ gilt die Gleichung $C^k(\bar{\Omega}) = C(\bar{\Omega})$ (s. Def. 2.15). Die Forderung der stetigen Fortsetzbarkeit der partiellen Ableitungen auf ganz $\bar{\Omega}$ ist nicht unwesentlich. Ist z. B. im \mathbb{R}^2 die Funktion $f(x_1, x_2) = \sqrt{x_1} + \sqrt{x_2}$ auf dem Gebiet $\Omega: 0 < x_1 < 1, 0 < x_2 < 1$ gegeben, so läßt sich zwar $f(x_1, x_2)$ auf $\bar{\Omega}: 0 \leq x_1 \leq 1, 0 \leq x_2 \leq 1$ stetig fortsetzen (mit derselben Zuordnungsvorschrift), jedoch läßt sich z. B. die erste partielle Ableitung von $f(x_1, x_2)$ nach der ersten Variablen, die in Ω erklärte Funktion

$$\frac{\partial f}{\partial x_1} = \frac{1}{2\sqrt{x_1}},$$

nicht stetig auf $\bar{\Omega}$ (als Funktion mit Werten aus \mathbf{R}) fortsetzen, da für $x_1 \rightarrow 0$ die Funktion $\frac{\partial f}{\partial x_1}$ keinen (endlichen) reellen Grenzwert besitzt (Skizze!).

Satz 2.12: Die Menge $C^k(\bar{\Omega})$ ist bezüglich der Norm

S.2.12

$$\|f\|_{C^k(\bar{\Omega})} = \sum_{|\alpha| \leq k} \left(\max_{x \in \bar{\Omega}} |\partial^\alpha f(x)| \right) \quad (2.15)$$

ein Banachraum (das Symbol $\sum_{|\alpha| \leq k}$ bedeutet, daß über sämtliche Multiindizes α mit $|\alpha| \leq k$ zu summieren ist).

Definition 2.17: Mit $C^\infty(\Omega)$ bezeichnen wir die Menge aller im Gebiet $\Omega \subseteq \mathbb{R}^n$ beliebig oft differenzierbaren komplexwertigen Funktionen. Mit $C_0^\infty(\Omega)$ (oder auch mit $\dot{C}^\infty(\Omega)$) bezeichnen wir die Menge aller Funktionen, die Elemente von $C^\infty(\Omega)$ sind und deren Träger, d. h. die Menge $\text{supp } f = \{x \in \Omega \mid f(x) \neq 0\}$, beschränkt ist und ganz in Ω liegt. Mit $C^k(\Omega)$ bezeichnen wir die Menge aller im Gebiet $\Omega \subseteq \mathbb{R}^n$ insgesamt k -mal stetig differenzierbaren Funktionen. D.2.17

Die Untersuchung von Operatoren auf Funktionenräumen erfordert häufig die Kenntnis kompakter Mengen in diesen Räumen. Daher geben wir das folgende Kompaktheitskriterium (Satz von Arzelà und Ascoli [2], [17]) an:

S.2.13 Satz 2.13: Es sei Ω ein beschränktes Gebiet des \mathbb{R}^n (m. a. W., die Menge $\bar{\Omega}$ ist kompakt). Es sei weiter $M \subseteq C^k(\Omega)$ eine abgeschlossene Teilmenge des Raumes $(C^k(\Omega), \|\cdot\|_{C^k(\Omega)})$ [vgl. (2.15)]. Dafür, daß M eine kompakte Teilmenge $\bar{\Omega}$ von $C^k(\Omega)$ ist, ist notwendig und zugleich hinreichend, daß M beschränkt und gleichgradigstetig zur Ordnung k ist; m. a. W., M ist genau dann kompakt, wenn es eine feste Zahl $Q > 0$ gibt mit

$$\|f\|_{C^k(\bar{\Omega})} \leq Q \quad (f \in M)$$

und wenn es zu jedem $\varepsilon > 0$ ein $\delta = \delta(\varepsilon) > 0$ (welches **nicht** von $f \in M$ abhängt) gibt, so daß aus der Beziehung

$$\|x - x'\|_{\mathbb{R}^n} \leq \delta \quad (x, x' \in \Omega)$$

stets die Beziehung

$$|\partial^\alpha f(x) - \partial^\alpha f(x')| \leq \varepsilon$$

für jeden Multiindex α mit $|\alpha| = k$ und für alle $f \in M$ folgt.

2.2.2. Räume integrierbarer Funktionen (Lebesgue-Räume)

Eine wichtige Klasse von Funktionenräumen stellen die sog. Lebesgue-Räume dar. Ihre Elemente sind Funktionen (genauer: Mengen [s. Bem. 2.10] von (L)-fast überall übereinstimmenden Funktionen), die auf einem Gebiet $\Omega \subseteq \mathbb{R}^n$ meßbar sind und zusätzlich Integrierbarkeitseigenschaften besitzen.

D.2.18 Definition 2.18: Es sei p eine positive reelle Zahl. Der Raum $L^p(\Omega)$ ist die Menge aller auf dem Gebiet Ω definierten meßbaren komplexwertigen Funktionen $f(x)$ (genauer: die Menge aller Klassen zueinander (L)-äquivalenter Funktionen), für welche

$$\int_{\Omega} |f(x)|^p dx < +\infty \quad (2.16)$$

gilt. Die Funktionen, die Elemente von $L^p(\Omega)$ sind, nennt man die **zur p -ten Potenz über Ω absolut integrierbaren Funktionen**. Mit $L^\infty(\Omega)$ bezeichnet man die Menge aller auf dem Gebiet Ω definierten meßbaren komplexwertigen Funktionen $f(x)$, für welche eine (von der betrachteten Funktion abhängende) Konstante $a > 0$ existiert mit

$$\text{mes}(\{x \in \Omega \mid |f(x)| \geq a\}) = 0.$$

Jedes solche a heißt eine „wesentliche Schranke“ von $f(x)$, und $f(x)$ heißt dann **im wesentlichen beschränkt**. Man nennt $L^\infty(\Omega)$ den Raum der auf Ω fast überall beschränkten Funktionen.

S.2.14 Satz 2.14: Für $1 \leq p < +\infty$ ist der Raum $L^p(\Omega)$ bezüglich der Norm

$$\|f\|_p = \left(\int_{\Omega} |f(x)|^p dx \right)^{\frac{1}{p}} \quad (2.17)$$

ein Banachraum. Der Raum $L^\infty(\Omega)$ ist bezüglich der Norm

$$\begin{aligned} \|f\|_\infty &:= \text{vrai max}_{x \in \Omega} |f(x)| := \text{ess sup}_{x \in \Omega} |f(x)| \\ &:= \inf \{a > 0 \mid \text{mes}(\{x \in \Omega \mid |f(x)| \geq a\}) = 0\} \end{aligned} \quad (2.18)$$

ein Banachraum.

Bemerkung 2.8: Ist $\Omega = (a, b) \subseteq \mathbf{R}$ ein Intervall der reellen Zahlengeraden, so schreibt man für den Ausdruck $L^p(\Omega)$ auch $L^p(a, b)$. Betrachtet man Funktionen, die auf dem abgeschlossenen Intervall $[a, b]$ definiert und dort zur p -ten Potenz absolut integrierbar sind, so bezeichnet man den entstehenden Raum auch mit $L^p[a, b]$; dieser unterscheidet sich (im Sinne der Normisomorphie normierter Räume) nicht von $L^p(a, b)$. Entsprechendes gilt allgemein beim Übergang von $\Omega \subseteq \mathbf{R}^n$ zu $\bar{\Omega}$.

Bemerkung 2.9: Der Beweis dafür, daß durch den Ausdruck auf der rechten Seite von (2.17) eine Norm gegeben ist, ergibt sich aus den **Ungleichungen von Hölder bzw. Minkowski** (s. [17]):

Ist $p > 1$ und $p^{-1} + q^{-1} = 1$ und sind $f \in L^p(\Omega)$, $g \in L^q(\Omega)$, so ist $fg \in L^1(\Omega)$, und es gilt

$$\left| \int_{\Omega} f(x) g(x) dx \right| \leq \|f\|_p \cdot \|g\|_q. \quad (\text{Hölder})$$

Ist $p \geq 1$ und sind $f \in L^p(\Omega)$, $g \in L^p(\Omega)$, so ist

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p. \quad (\text{Minkowski})$$

Bemerkung 2.10: Aus Platzgründen können wir die exakte Definition der Räume $L^p(\Omega)$ als Räume von Klassen zueinander (L)-äquivalenter Funktionen nur beschreiben und nicht vollständig logisch durchkonstruieren. Zur Ergänzung seien aber die folgenden Bemerkungen angeführt. Mittels der Minkowskischen Ungleichung ergibt sich zunächst, daß die Menge aller Funktionen (komplexwertig, meßbar), die der Ungleichung (2.16) genügen, einen Vektorraum bilden; er werde mit \mathcal{L}^p bezeichnet. Die Menge N_0 aller Funktionen, die auf Ω fast überall gleich null sind, bildet einen linearen Teilraum von \mathcal{L}^p . Der Raum $L^p(\Omega)$ ist dann der Quotientenraum \mathcal{L}^p/N_0 [s. (2.11)]. Diese Quotientenraumbildung wird ausschließlich durch den Umstand veranlaßt, daß der Ausdruck (2.17) auf \mathcal{L}^p keine Norm liefert, da er für alle Elemente von N_0 gleich null ist (also auch für Funktionen, die nicht *überall* auf Ω gleich null sind). Das praktische Rechnen im Raum $L^p(\Omega)$ wird aber auf das Rechnen in \mathcal{L}^p zurückgeführt (man rechnet mit Funktionen, nicht mit den Elementen des Quotientenraumes).

Definition 2.19: Unter $L^p_{\text{loc}}(\Omega)$ ($0 < p < \infty$) versteht man die Menge aller auf dem Gebiet Ω definierten meßbaren komplexwertigen Funktionen (genauer: die Menge aller Klassen zueinander (L)-äquivalenten Funktionen) mit **D.2.19**

$$\int_{\Omega'} |f(x)|^p dx < +\infty$$

für jedes beschränkte Gebiet $\Omega' \subseteq \Omega$ (s. 1.2.2. und 4.1.).

Bemerkung 2.11: $L^p_{\text{loc}}(\Omega)$ ist ein Vektorraum (Funktionenraum). Ist Ω beschränkt, so gilt $L^p_{\text{loc}}(\Omega) = L^p(\Omega)$; sonst sind diese Räume verschieden. $L^1_{\text{loc}}(\Omega)$ wird auch als der Raum der „lokal integrierbaren Funktionen auf Ω “ bezeichnet. Zum Beispiel liegt $f(x) = x^{-1}$ nicht in $L^1_{\text{loc}}(\mathbf{R}^1)$.

Satz 2.15: Für $1 \leq p \leq \infty$ gilt die Beziehung ($\Omega \subseteq \mathbf{R}^n$, beliebiges Gebiet) **S.2.15**

$$L^p(\Omega) \subseteq L^1_{\text{loc}}(\Omega). \quad (2.19)$$

Über die Struktur der kompakten Mengen im Raum $L^p(\Omega)$ gibt der folgende Satz Auskunft.

S.2.16 Satz 2.16: Es sei $M \subseteq L^p(\Omega)$ eine abgeschlossene Teilmenge des Raumes $L^p(\Omega)$. Dafür, daß M kompakt ist, ist das gleichzeitige Bestehen der folgenden Bedingungen sowohl notwendig als auch hinreichend:

- (1) M ist beschränkt; d. h., es existiert ein $K > 0$ mit $\|f\|_p \leq K$ für alle $f \in M$;
- (2) Zu jedem $\varepsilon > 0$ gibt es ein $\delta > 0$ und eine abgeschlossene beschränkte Teilmenge $G \subseteq \Omega$ mit

$$(2.1) \quad \int_{\Omega \setminus G} |f(x)|^p dx \leq \varepsilon \quad (f \in M) \quad \text{und}$$

$$(2.2) \quad \int_{\Omega} |\hat{f}(x+h) - \hat{f}(x)|^p dx \leq \varepsilon$$

für alle $f \in M$ und alle $h \in \mathbb{R}^n$ mit $\|h\|_{\mathbb{R}^n} \leq \delta$, wobei $\hat{f}(x)$ die durch

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

erklärte Funktion bezeichnet.

Zum Verhältnis der Räume $C^k(\bar{\Omega})$ und $L^p(\Omega)$ ist zu sagen, daß für ein beschränktes Gebiet Ω die Beziehung

$$C^k(\bar{\Omega}) \subseteq L^p(\Omega)$$

in dem Sinne gilt, daß die von den Elementen von $C^k(\bar{\Omega})$ erzeugten Klassen [(L)-fast überall übereinstimmender Funktionen] Elemente von $L^p(\Omega)$ sind. Im Sinne der Abschlußbildung im Raum $L^p(\Omega)$ gilt die Gleichung $C^k(\bar{\Omega}) = L^p(\Omega)$, d. h., der Raum $C^k(\bar{\Omega})$ liegt dicht in $L^p(\Omega)$ (Ω beschränkt).

Ist das Gebiet Ω nicht beschränkt, so gilt in entsprechendem Sinn die folgende Aussage.

S.2.17 Satz 2.17: Die Menge $C^\infty(\Omega)$ (s. Def. 2.17) liegt dicht in $L^p(\Omega)$ für $1 \leq p < +\infty$.

2.2.3. Sobolew-Räume

In diesem Abschnitt führen wir Sobolew-Räume (ganzzahliger Ordnung $k \geq 0$) ein. Diese Räume sind Räume von (Klassen (L)-fast überall übereinstimmender Funktionen, die auf einem Gebiet $\Omega \subseteq \mathbb{R}^n$ definiert sind und gewisse Differenzierbarkeitseigenschaften besitzen; sie sind überdies Teilräume der Räume $L^p(\Omega)$ und haben sich insbesondere bei der theoretischen und numerischen Behandlung partieller Differentialgleichungen als nützlich erwiesen (s. auch 5.3.). Es gibt verschiedene Möglichkeiten, die Sobolew-Räume einzuführen:

- a) als Vervollständigung des Raumes $C^k(\Omega)$ bezüglich einer speziellen Metrik (s. u.);
- b) als Räume von Funktionen, deren verallgemeinerte Ableitungen (s. auch 4.1.) bis zur Ordnung k einschließlich existieren und zusätzlich Elemente des Raumes $L^p(\Omega)$ sind.

Es zeigt sich, daß diese verschiedenen Zugänge für $1 \leq p < +\infty$ stets die gleichen Räume liefern. Im einzelnen definieren wir folgende Begriffe (vgl. [2], [22], [35]).

D.2.20 Definition 2.20: Ist $f \in C^k(\Omega)$ und $1 \leq p < \infty$, so setzen wir

$$\|f\|_{k,p} = \left\{ \sum_{0 \leq |\alpha| \leq k} \|\partial^\alpha f\|_p^p \right\}^{\frac{1}{p}}, \quad k = 0, 1, 2, \dots; \quad (2.20)$$

ferner sei

$$\|f\|_{k,\infty} = \max_{0 \leq |\alpha| \leq k} \|\partial^\alpha f\|_\infty, \quad k = 0, 1, 2, \dots, \quad (2.21)$$

wobei $\|\cdot\|_p$ ($1 \leq p \leq \infty$) die Norm im Raum $L^p(\Omega)$ bezeichnet [s. (2.17), (2.18)].

Definition 2.21: Mit $H^{k,p}(\Omega)$ bezeichnet man die Vervollständigung der Menge $\{f | f \in C^k(\Omega), \|f\|_{k,p} < \infty\}$ bezüglich der durch $\|f\|_{k,p}$ erzeugten Metrik. **D.2.21**

Der Raum $H^{k,p}(\Omega)$ ist bereits ein Sobolew-Raum für ($1 \leq p < \infty$) (s. Satz 2.19). Seine Elemente sind aber nur sehr unkonkret erfaßbar. Die zweite Möglichkeit, Sobolew-Räume einzuführen, ist wesentlich anschaulicher. Allerdings kommt sie ohne Benutzung des Begriffs der verallgemeinerten Ableitung (= Distributionsableitung) nicht aus (s. auch 4.1.).

Vorbereitend sei dazu bemerkt, daß für jedes feste $f \in C^k(\bar{\Omega})$ und beliebiges $\varphi \in \dot{C}^\infty(\Omega)$ nach den Regeln der partiellen Integration (Gaußscher Integralsatz!) immer gilt

$$\int_{\Omega} f(x) \partial^\alpha \varphi(x) \, dx = (-1)^{|\alpha|} \int_{\Omega} \varphi(x) \partial^\alpha f(x) \, dx,$$

wobei $\partial^\alpha \varphi, \partial^\alpha f$ die Ableitungen im üblichen Sinne sind. Diese Eigenschaft der „klassischen“ Ableitungen bildet die Grundlage für folgende Verallgemeinerung:

Definition 2.22: Es sei $f \in L^1_{\text{loc}}(\Omega)$. Eine Funktion g , die zu $L^1_{\text{loc}}(\Omega)$ (vgl. Bem. 2.11) gehört, heißt **schwache Ableitung oder verallgemeinerte Ableitung (Distributionsableitung)** von f bezüglich des Multiindex α , wenn die Gleichung **D.2.22**

$$\int_{\Omega} f(x) \partial^\alpha \varphi(x) \, dx = (-1)^{|\alpha|} \int_{\Omega} g(x) \varphi(x) \, dx \quad (2.22)$$

für jede Funktion $\varphi(x)$ gilt, die zu $\dot{C}^\infty(\Omega)$ gehört. Wir schreiben dann

$$g = D^\alpha f. \quad (2.23)$$

Wegen der vorbereitenden Bemerkung folgt, daß für jedes $f \in C^k(\bar{\Omega})$ die verallgemeinerte und die klassische Ableitung übereinstimmen (im Sinne der Gleichheit im Raum $L^p(\Omega)$). [Bezüglich der Motivation dieser Definition s. (1.51).]

Definition 2.23: Mit $W^{k,p}(\Omega)$ bezeichnen wir den linearen Raum

D.2.23

$$W^{k,p}(\Omega) = \{f \in L^p(\Omega) \mid D^\alpha f \in L^p(\Omega) \text{ für } 0 \leq |\alpha| \leq k\} \quad (2.24)$$

(andere Bezeichnung: $W^k_k(\Omega)$; $W^k_p(\Omega)$).

Zum Beispiel gehört die Funktion $f(x) = |x|^{-\alpha}$ ($x \in \Omega \subseteq \mathbb{R}^n$, Ω offen und beschränkt), wobei $|x| = \sqrt{\xi_1^2 + \dots + \xi_n^2}$ ist, für $0 < \alpha < \left(\frac{n}{2} - 1\right)$ zu $W^{1,2}(\Omega)$, falls $n \geq 3$ ist [6].

Satz 2.18: Der Raum $W^{k,p}(\Omega)$ ist, versehen mit der Norm

S.2.18

$$\|f\|_{k,p} = \left(\sum_{0 \leq |\alpha| \leq k} \|D^\alpha f\|_p^p \right)^{\frac{1}{p}}, \quad (2.25)$$

ein Banachraum.

D.2.24 Definition 2.24: Mit $\dot{W}^{k,p}(\Omega)$ (auch $W_0^{k,p}(\Omega)$) bezeichnet man die Abschließung der Menge $\dot{C}^\infty(\Omega)$ im Raum $W^{k,p}(\Omega)$ [versehen mit der Norm (2.25)].

$\dot{W}^{k,p}(\Omega)$ ist somit ein abgeschlossener linearer Teilraum von $W^{k,p}(\Omega)$.

Der folgende, wichtige Satz zeigt, daß beide geschilderten Möglichkeiten der Einführung von Sobolew-Räumen im wesentlichen äquivalent sind.

S.2.19 Satz 2.19 (N. Meyers, J. Serrin, 1964): Für $1 \leq p < \infty$ gilt die Gleichheit ($k = 0, 1, 2, \dots$)

$$H^{k,p}(\Omega) = \dot{W}^{k,p}(\Omega). \quad (2.26)$$

Bemerkung 2.12: Im Sinne der mengentheoretischen Enthaltenseinsrelation \subseteq gilt stets

$$\dot{C}^\infty(\Omega) \subseteq \dot{W}^{k,p}(\Omega) \subseteq W^{k,p}(\Omega) \subseteq L^p(\Omega). \quad (2.27)$$

Für $k = 0$ ist dabei speziell

$$\dot{W}^{0,p}(\Omega) = W^{0,p}(\Omega) = L^p(\Omega) \quad (1 \leq p < \infty). \quad (2.28)$$

Für $p = \infty$ gilt die Gl. (2.26) nicht! (Siehe [2].)

Für andere Varianten der Definition der Räume $H^{k,p}$ bzw. $\dot{W}^{k,p}$ vergleiche man das inhaltsreiche Buch [22] über Funktionenräume, S. 250–273. Für die Beziehungen zwischen diesen Räumen ergeben sich dann vom Satz 2.19 abweichende Resultate.

2.2.4. Folgenräume

In vielen funktionalanalytischen Betrachtungen lassen sich zur Darstellung eines Sachverhalts einfache unendlichdimensionale Räume, nämlich Folgenräume benutzen. Ihre Elemente (Vektoren) sind Folgen komplexer (bzw. reeller) Zahlen bzw. allgemeiner Folgen von Elementen eines Banachraumes. Die folgende Definition führt einige wichtige Folgenräume ein.

D.2.25 Definition 2.25: Im folgenden sei $x = \{\xi_n\}$ eine Folge komplexer Zahlen ξ_n ($n = 1, 2, \dots$). Wir setzen (s. [29]):

$$l^\infty = \left\{ x \mid \sup_n |\xi_n| < +\infty \right\} \quad (\text{Raum der beschränkten Folgen}),$$

$$c = \left\{ x \mid \lim_{n \rightarrow \infty} \xi_n \text{ existiert} \right\} \quad (\text{Raum der konvergenten Folgen}),$$

$$c_0 = \left\{ x \mid \lim_{n \rightarrow \infty} \xi_n = 0 \right\} \quad (\text{Raum der Nullfolgen}),$$

$$l^p = \left\{ x \mid \sum_{n=1}^{\infty} |\xi_n|^p < +\infty \right\} \quad (1 \leq p < \infty) \quad (\text{Raum der zur } p\text{-ten Potenz summierbaren Folgen}),$$

$$s = \left\{ x \mid \lim_{n \rightarrow \infty} n^k \xi_n = 0 \text{ für alle } k = 1, 2, \dots \right\} \\ (\text{Raum der rasch fallenden Folgen}),$$

$$s_0 = \{ x \mid \xi_n = 0 \text{ für alle } n, \text{ bis auf endlich viele} \} \\ (\text{Raum der finiten Folgen}).$$

Bemerkung 2.13: Es gelten die Enthaltenseinsbeziehungen

$$s_0 \subseteq s \subseteq l^p \subseteq c_0 \subseteq c \subseteq l^\infty \quad (2.29)$$

(ebenso richtig für $s_{0\mathbf{R}}, s_{\mathbf{R}}, l_{\mathbf{R}}, \dots$, also Räume reeller Zahlenfolgen).

Satz 2.20: 1) Mit der Norm $\|x\|_\infty = \sup_n |\xi_n|$ ist l^∞ ein Banachraum, ebenso c und c_0 mit der gleichen Norm. **S.2.20**

2) Mit der Norm $\|x\|_p = \left(\sum_{n=1}^{\infty} |\xi_n|^p \right)^{\frac{1}{p}}$ ist l^p ein Banachraum.

2.3. Lineare Funktionale, schwache Konvergenz, dualer Raum

2.3.1. Lineare Funktionale

Zu den einfachsten Funktionen, die man auf Vektorräumen betrachten kann, gehören die häufig auftretenden linearen Funktionen. Linearität ist hierbei eine Eigenschaft, die in den Anwendungen vor allem als „Superpositionsprinzip“ zutage tritt. Lineare Funktionen mit Werten in \mathbf{R} oder \mathbf{K} heißen lineare Funktionale oder Linearformen. Die Elemente von (normierten) Vektorräumen werden im folgenden mit x, y, z, \dots bezeichnet.

Definition 2.26: Es sei E ein Vektorraum. Eine Abbildung $f: E \rightarrow \mathbf{R}$ (bzw. \mathbf{K}) heißt **lineares Funktional (Linearform)**, wenn die folgenden Bedingungen erfüllt sind (s. Bd. 1, 8.4.): **D.2.26**

$$f(x + y) = f(x) + f(y) \quad (x, y \in E), \quad (2.30a)$$

$$f(\lambda x) = \lambda f(x) \quad (x \in E; \lambda \in \mathbf{R} \text{ bzw. } \mathbf{K}). \quad (2.30b)$$

Beispiel 2.12: Es sei $E = \mathbf{R}^n$ (bzw. \mathbf{K}^n) mit den üblichen Vektoroperationen, und es seien a_1, \dots, a_n feste reelle (bzw. komplexe) Zahlen. Für $x = (\xi_1, \dots, \xi_n) \in E$ setzen wir

$$f(x) = \sum_{j=1}^n a_j \xi_j. \quad (2.31)$$

$f(x)$ ist ein lineares Funktional auf E . Ist $e_1 = (1, 0, \dots, 0)$; $e_2 = (0, 1, 0, \dots, 0)$; ...; $e_n = (0, 0, \dots, 1)$ die ausgezeichnete Basis von E , so gilt ersichtlich die Gleichung

$$f(e_k) = a_k \quad (k = 1, \dots, n). \quad (2.32)$$

Wegen (2.30) ist

$$f(x) = \sum_{j=1}^n a_j \xi_j = \sum_{j=1}^n f(e_j) \xi_j. \quad (2.33)$$

Sind umgekehrt die Werte $f(e_j) = a_j$ ($j = 1, \dots, n$) vorgegeben, so definiert (2.31) ein lineares Funktional auf E mit (2.32). Mit anderen Worten, jedes lineare Funktional auf E hat die Form (2.31), und durch die Vorgabe der Werte $f(e_j)$ ($j = 1, \dots, n$) ist das lineare Funktional $f(x)$ eindeutig bestimmt.

Beispiel 2.13: Es sei $E = C[a, b]$ (s. Def. 2.15). Die Elemente von $C[a, b]$ bezeichnen wir jetzt mit x, y, \dots bzw. $x(t), y(t), \dots$ oder auch $x(\cdot), y(\cdot), \dots$. Es sei

$$f(x) = \int_a^b x(t) dt \quad (x \in E). \quad (2.34)$$

Es gilt $(x, y \in E; \lambda \in \mathbf{R}; \mathbf{K})$:

$$\begin{aligned} f(x+y) &= \int_a^b (x+y)(t) dt = \int_a^b (x(t)+y(t)) dt = \int_a^b x(t) dt + \int_a^b y(t) dt \\ &= f(x) + f(y); \end{aligned}$$

$$f(\lambda x) = \int_a^b (\lambda x)(t) dt = \int_a^b \lambda x(t) dt = \lambda \int_a^b x(t) dt = \lambda f(x);$$

also ist f ein lineares Funktional auf E .

Beispiel 2.14: Es sei $E = C[a, b]$ und $c \in [a, b]$ ein Punkt aus dem Intervall $[a, b]$. Wir setzen

$$f(x) = x(c) \quad (x \in E). \quad (2.35)$$

Wie durch Einsetzen in (2.30a), (2.30b) sofort ersichtlich wird, ist f ein lineares Funktional auf $E = C[a, b]$; s. auch Bsp. 1.10 und (1.58).

Beispiel 2.15: Es sei $E = L^2[a, b]$ (Bezeichnung der Elemente wie in Bsp. 2.13). Wir setzen

$$f(x) = \int_a^b x(t) dt. \quad (2.36)$$

Zufolge der Enthaltenseinsbeziehung

$$L^2[a, b] \subseteq L^1[a, b]$$

ist das Funktional f tatsächlich auf $L^2[a, b]$ definiert (der Leser überlege sich dies!) und (s. Bsp. 2.13) linear.

D.2.27 Definition 2.27: Ein (lineares) Funktional f auf dem normierten Raum $(E, \|\cdot\|)$ heißt stetig, wenn aus der Beziehung

$$\lim_{n \rightarrow \infty} x_n = x \quad (x_n, x \in E)$$

stets die Beziehung

$$\lim_{n \rightarrow \infty} f(x_n) = f(x) \quad (2.37)$$

folgt.

Bevor wir zu Beispielen stetiger linearer Funktionalen kommen, soll die Stetigkeits-eigenschaft noch eingehender diskutiert werden.

Hilfssatz 2.1: Ist $f: E \rightarrow \mathbf{R} (\mathbf{K})$ ein lineares Funktional auf dem normierten Raum $(E, \|\cdot\|)$ mit der Eigenschaft $(x_n \in E)$:

$$\lim_{n \rightarrow \infty} x_n = o \Rightarrow \lim_{n \rightarrow \infty} f(x_n) = 0,$$

so ist $|f(x)|$ auf jeder Kugel $K(o; r) = \{x \in E \mid \|x\| \leq r\}$ ($r > 0$) beschränkt.

Beweis: Wir zeigen zuerst, daß $|f(x)|$ auf der Kugeloberfläche $S = S(o; r) = \{x \in E \mid \|x\| = r\}$ beschränkt ist. Angenommen, dies wäre nicht der Fall, dann gibt es zu jedem $n = 1, 2, \dots$, ein

$x_n \in S$ mit $|f(x_n)| \geq n$. Mittels der Linearität von $f(x)$ folgt daraus $\left| f\left(\frac{1}{n} x_n\right) \right| \geq 1$. Wegen

$\left\| \frac{1}{n} x_n \right\| = \frac{1}{n} \|x_n\| = \frac{r}{n}$ ($n = 1, 2, \dots$) folgt durch Grenzübergang $\lim_{n \rightarrow \infty} \left(\frac{1}{n} x_n \right) = o$; aber es gilt

nicht $\lim_{n \rightarrow \infty} f\left(\frac{1}{n} x_n\right) = 0$; Widerspruch! Es gibt also ein $M > 0$ mit $|f(x)| \leq M$ ($x \in S$). Ist $x \in K(o; r)$, so gilt für $x \neq o$ die Relation $\left(\frac{1}{\|x\|} rx\right) \in S$ (weil $\left\|\frac{r}{\|x\|} x\right\| = \frac{r}{\|x\|} \|x\| = r$ gilt). Also ist $\left|f\left(\frac{r}{\|x\|} x\right)\right| \leq M$, und wir erhalten (Linearität von f) die Ungleichung (man beachte, daß $\frac{\|x\|}{r} \leq 1$ gilt)

$$|f(x)| \leq \frac{\|x\|}{r} M \leq M,$$

die auch für $x = o$ gilt. ■

Hilfssatz 2.2: Ist $f: E \rightarrow \mathbf{R}(\mathbf{K})$ ein lineares Funktional auf dem normierten Raum $(E, \|\cdot\|)$, welches auf einer gewissen Kugel $K(o; r)$ ($r > 0$) beschränkt ist, d. h., es existiere ein $M > 0$ mit $|f(x)| \leq M$ ($x \in K(o; r)$), so gilt die Ungleichung $|f(x)| \leq A\|x\|$ ($x \in E$), $A = Mr^{-1}$ auf ganz E , und das Funktional $f(x)$ ist auf dem (ganzen) Raum E stetig.

Beweis: Für gegebenes $x \in E$ ($x \neq o$) liegt das Element $\frac{r}{\|x\|} x$ in $K(o; r)$, daher gilt $\left|f\left(\frac{r}{\|x\|} x\right)\right| \leq M$, woraus sofort die Ungleichung $|f(x)| \leq \frac{M}{r} \|x\|$ folgt (die für $x = o$ trivialerweise wegen $f(o) = 0$ erfüllt ist). Ist weiter $\{x_n\}$ eine Folge aus E mit $\lim_{n \rightarrow \infty} x_n = x$, so erhalten wir aus der Abschätzung (f ist linear) $|f(x_n) - f(x)| = |f(x_n - x)| \leq \frac{M}{r} \|x_n - x\|$ ($n = 1, 2, \dots$) wegen $\lim_{n \rightarrow \infty} \|x_n - x\| = 0$ die Beziehung $\lim_{n \rightarrow \infty} f(x_n) = f(x)$. ■

Aus den beiden Hilfssätzen ergeben sich eine Reihe wichtiger Folgerungen, deren Beweis wir dem Leser überlassen.

Satz 2.21: Ist $f: E \rightarrow \mathbf{R}$ (bzw. $E \rightarrow \mathbf{K}$) ein lineares Funktional auf dem normierten Raum $(E, \|\cdot\|)$, so ist $f(x)$ genau dann stetig, wenn es an der Stelle $x = o$ stetig ist. S.2.21

Definition 2.28: Ein lineares Funktional $f: E \rightarrow \mathbf{R}(\mathbf{K})$ auf dem normierten Raum $(E, \|\cdot\|)$ heißt **beschränkt**, wenn es ein $M > 0$ gibt mit D.2.28

$$|f(x)| \leq M\|x\| \quad (x \in E). \quad (2.38)$$

Satz 2.22: Ein lineares Funktional $f: E \rightarrow \mathbf{R}(\mathbf{K})$ auf dem normierten Raum $(E, \|\cdot\|)$ ist genau dann stetig, wenn es beschränkt ist. S.2.22

Definition 2.29: Es sei $f: E \rightarrow \mathbf{R}(\mathbf{K})$ ein beschränktes lineares Funktional auf dem normierten Raum $(E, \|\cdot\|)$. Die Zahl D.2.29

$$\|f\| = \sup_{\|x\|=1} |f(x)| \quad (2.39)$$

heißt die **Norm** von f .

Satz 2.23: Ist $f: E \rightarrow \mathbf{R}(\mathbf{K})$ ein beschränktes lineares Funktional auf dem normierten Raum $(R, \|\cdot\|)$, so gilt die Ungleichung S.2.23

$$|f(x)| \leq \|f\| \|x\| \quad (x \in E). \quad (2.40)$$

Ist $M > 0$ eine Zahl, für die (2.38) gilt, so ist $\|f\| \leq M$.

Mit anderen Worten, die Zahl $\|f\|$ ist die kleinste **Dehnungskonstante** von f .

S.2.24 Satz 2.24: Ist $f: E \rightarrow \mathbf{R}$ (bzw. \mathbf{K}) ein beschränktes lineares Funktional auf dem normierten Raum $(E, \|\cdot\|)$, so gilt

$$\|f\| = \sup_{\|x\| \leq 1} |f(x)|.$$

Beispiel 2.16: Es sei $E = K^n$ der n -dimensionalen komplexe Vektorraum mit der Norm $\|x\| = \left\{ \sum_{j=1}^n |\xi_j|^2 \right\}^{1/2}$ ($x = (\xi_1, \dots, \xi_n)$). Mit gewissen komplexen Zahlen a_1, \dots, a_n bilden wir durch die Vorschrift $f(x) = \sum_{j=1}^n a_j \xi_j$ ein lineares Funktional $f(x)$ (Bsp. 2.12). Mittels der Schwarzschen Ungleichung folgt dessen Beschränktheit und damit Stetigkeit: für $x \in K^n$ ist

$$|f(x)| = \left| \sum_{k=1}^n a_k \xi_k \right| \leq \left\{ \sum_{k=1}^n |a_k|^2 \right\}^{1/2} \left\{ \sum_{k=1}^n |\xi_k|^2 \right\}^{1/2} = M \|x\|$$

mit $M = \left\{ \sum_{k=1}^n |a_k|^2 \right\}^{1/2}$. Ist $M > 0$ und wählt man für x speziell den Vektor x' mit den Koordinaten $\xi'_k = \frac{\bar{a}_k}{M}$ ($k = 1, \dots, n$), so gilt $\|x'\| = 1$ und

$$|f(x')| = \left| \sum_{k=1}^n a_k \xi'_k \right| = \left| \sum_{k=1}^n \frac{a_k \bar{a}_k}{M} \right| = \frac{1}{M} \sum_{k=1}^n |a_k|^2 = \frac{M^2}{M} = M.$$

Also ist $\|f\| = \sup_{\|x\|=1} |f(x)| \geq |f(x')| = M$. Aus der zuvor bewiesenen Ungleichung folgt für $\|x\| = 1$ die Abschätzung $|f(x)| \leq M$ und somit $\|f\| = \sup_{\|x\|=1} |f(x)| \leq M$. Insgesamt erhalten wir die Gleichheit

$$\|f\| = M = \left\{ \sum_{k=1}^n |a_k|^2 \right\}^{1/2}. \quad (2.41)$$

Beispiel 2.17: Es sei $E = C[a, b]$ mit der Maximum-Norm $\|x\| = \max_{a \leq t \leq b} |x(t)|$. Das lineare Funktional (s. Bsp. 2.13)

$$f(x) = \int_a^b x(t) dt \quad (x \in C[a, b])$$

ist stetig, wir geben sogar seine Norm an:

$$\begin{aligned} |f(x)| &= \left| \int_a^b x(t) dt \right| \leq \int_a^b |x(t)| dt \leq \int_a^b \sup_{a \leq t \leq b} |x(t)| dt \\ &= \int_a^b \|x\| dt = \|x\| \int_a^b dt = (b-a) \|x\| \quad (x \in E). \end{aligned}$$

Wir erhalten die Ungleichung $\|f\| \leq (b-a)$. Für die Funktion $x'(t) = 1$ ($a \leq t \leq b$) wird $\|x'\| = 1$ und $|f(x')| = \left| \int_a^b 1 dt \right| = b-a$ und somit $\|f\| = \sup_{\|x\| \leq 1} |f(x)| \geq b-a (= |f(x')|)$. Insgesamt erhalten wir die Beziehung $\|f\| = b-a$.

Beispiel 2.18: Es sei $E = L^2[a, b]$ mit der Norm $\|x\| = \left\{ \int_a^b (x(t))^2 dt \right\}^{1/2}$ und $g(\cdot) \in L^2[a, b]$. Das lineare Funktional

$$f(x) = \int_a^b x(t) g(t) dt \quad (2.42)$$

ist, wie man mittels der Schwarzschen Ungleichung erkennt, beschränkt und daher stetig. Analoge Betrachtungen wie in Bsp. 2.17 zeigen, daß die folgende Gleichung gilt:

$$\|f\| = \left\{ \int_a^b |g(t)|^2 dt \right\}^{1/2} = \|g\|_{L^2}. \quad (2.43)$$

Die letztgenannten Beispiele 2.17 und 2.18 sind Spezialfälle allgemeinerer grundlegender Aussagen über die Form stetiger linearer Funktionale im Raum $C[a, b]$ bzw. in einem Hilbertraum¹⁾.

Satz 2.25 (Stetige Linearform in $C_{\mathbf{R}}[a, b]$): Ist $E = C_{\mathbf{R}}[a, b]$, versehen mit der max-Norm, so läßt sich jedes stetige lineare Funktional $f(x)$ auf $C_{\mathbf{R}}[a, b]$ als Riemann-Stieltjes-Integral **S.2.25**

$$f(x) = \int_a^b x(t) dg(t) \quad (x \in C_{\mathbf{R}}[a, b]) \quad (2.44)$$

mit einer gewissen Belegungsfunktion $g(t)$ ($a \leq t \leq b$) von beschränkter Variation darstellen (die bis auf eine additive Konstante und fast überall bestimmt ist).

Bemerkung 2.14: Bezüglich der Definition des Riemann-Stieltjes-Integrals vgl. [17, S. 176/77].

Der folgende wichtige Satz stammt von F. Riesz.

Satz 2.26 (Stetige Linearformen im Hilbertraum): Ist H ein Hilbertraum¹⁾ mit dem Skalarprodukt $\langle \cdot | \cdot \rangle$ und $f(x)$ ein stetiges lineares Funktional auf H , so gibt es genau ein Element $g \in H$, so daß die Gleichheit **S.2.26**

$$f(x) = \langle g | x \rangle$$

für alle $x \in H$ gilt. Zusätzlich gilt die Beziehung

$$\|f\| = \|g\|_H = \sqrt{\langle g | g \rangle}. \quad (2.45)$$

2.3.2. Dualer Raum

Definition 2.30: Es sei E ein Vektorraum. Die Menge E^* aller linearen Funktionale auf E , die mittels der Operationen **D.2.30**

$$(f + g)(x) = f(x) + g(x) \quad (x \in E), \quad (2.46a)$$

$$(\lambda f)(x) = \lambda f(x) \quad (x \in E) \quad (2.46b)$$

($f, g \in E^*$, $\lambda \in \mathbf{K}$ bzw. \mathbf{R}) zu einem Vektorraum wird, nennt man den **algebraisch dualen** bzw. **algebraisch konjugierten Raum** von E .

Definition 2.31: Es sei $(E, \|\cdot\|_E)$ ein normierter Raum. Denjenigen linearen Teilraum E' von E^* , der aus allen **stetigen** linearen Funktionalen auf E besteht, versehen mit der Norm **D.2.31**

$$\|f\|_{E'} = \sup_{\|x\|_E \leq 1} |f(x)|, \quad (2.47)$$

nennt man den **Dualraum** von $(E, \|\cdot\|_E)$ (bzw. den **dualen** oder **konjugierten Raum** von $(E, \|\cdot\|_E)$).

¹⁾ (s. Def. 2.37 weiter unten).

S.2.27 Satz 2.27: Der Dualraum $(E', \| \cdot \|_{E'})$ eines normierten Raumes $(E, \| \cdot \|_E)$ ist stets ein Banachraum.

Beispiel 2.19: Es sei $E = K^n$ mit der Norm $\|x\| = \left\{ \sum_{j=1}^n |\xi_j|^2 \right\}^{1/2}$ ($x = (\xi_1, \dots, \xi_n) \in K^n$). Jedes $f \in E^*$ hat die Form

$$f(x) = \sum_{j=1}^n a_j \xi_j \quad (x \in K^n) \quad (2.48)$$

mit eindeutig bestimmten a_j ($j = 1, \dots, n$) (vgl. Bsp. 2.12) und ist stetig. Das heißt, im vorliegenden Fall gilt die Beziehung $E^* = E'$. Weiter ist [s. (2.41)] die Gleichung

$$\|f\| = \left\{ \sum_{j=1}^n |a_j|^2 \right\}^{1/2} \quad (2.49)$$

erfüllt. Ordnet man jedem $f \in E'$ den zugehörigen Vektor $(a_1, \dots, a_n) \in K^n$ zu, so erhält man eine eindeutige lineare Abbildung T von E' auf K^n , die wegen (2.49) sogar ein Normisomorphismus ist. (Beweis der Linearität und der Surjektivität der Abbildung T als Übung.) Im Sinne der Gleichsetzung normisomorpher Räume gilt also die Beziehung

$$(K^n)' = K^n. \quad (2.50)$$

Es sei noch bemerkt, daß – wiederum im Sinne der Gleichsetzung normisomorpher Räume – die Beziehung $(L^p(\Omega))' = L^q(\Omega)$ gilt, wobei $1 < p < \infty$ und $\frac{1}{q} + \frac{1}{p} = 1$ ist (q heißt der zu p konjugierte Exponent) (vgl. [24], S. 125).

Beispiel 2.20: Es sei E ein (komplexer) Hilbertraum. Nach dem Satz von F. Riesz hat jedes $f \in E'$ die Form

$$f(x) = \langle g | x \rangle \quad (x \in E) \quad (2.51)$$

mit einem eindeutig bestimmten $g \in E$, für welche die Gleichung

$$\|f\| = \|g\| \quad (2.52)$$

gilt. Die Zuordnung $T: f \rightarrow g$ ist eine eindeutige Abbildung von E' auf E , die aber nicht linear, sondern antilinear ist in folgendem Sinn (Beweis als Übung):

$$T(f_1 + f_2) = T(f_1) + T(f_2)$$

$$T(\lambda f) = \bar{\lambda} T f.$$

Wegen (2.52) ist T ein antilinearer Normisomorphismus von E' auf E . Ist E ein **reeller** Hilbertraum, so ist T linear, und es gilt (im Sinne der Identifizierung normisomorpher Räume) $E' = E$.

D.2.32 Definition 2.32: Es sei $(E, \| \cdot \|)$ ein Banachraum und $(E', \| \cdot \|_{E'})$ sein Dualraum. Den Dualraum von $(E', \| \cdot \|_{E'})$ bezeichnet man als **Bidualraum** (bidualen Raum) $(E'', \| \cdot \|_{E''})$ von $(E, \| \cdot \|)$.

S.2.28 Satz 2.28: Der biduale Raum $(E'', \| \cdot \|_{E''})$ enthält einen linearen Teilraum, der zu $(E, \| \cdot \|)$ normisomorph ist. Es gilt also (im Sinne der Gleichsetzung normisomorpher Räume) die Relation $E \subseteq E''$. Dabei wird jedem $x \in E$ dasjenige Element $I_x \in E''$ zugeordnet, das durch die Gleichung

$$I_x(f) = f(x) \quad (f \in E') \quad (2.53)$$

definiert ist (kanonische Einbettung).

Für einen Beweis dieses Satzes vgl. [17].

Definition 2.33: Der normierte Raum $(E, \|\cdot\|)$ heißt **reflexiv**, wenn die kanonische Einbettung $x \rightarrow I_x$ ($x \in E$) (s. Satz 2.28) ein Normisomorphismus von E auf E'' ist. D.2.33

Zu den reflexiven Räumen gehören die Räume R^n, K^n sowie die Räume $L^p(\Omega)$ ($1 < p < \infty$) und alle Hilberträume.

2.3.3. Fortsetzung stetiger linearer Funktionale. Satz von Hahn und Banach. Trennungssätze

Die Fortsetzung von Funktionen auf größere Definitionsbereiche spielt an vielen Stellen in der Funktionalanalysis eine Rolle. Die Forderung der Linearität und der Stetigkeit eines Funktionalen ermöglichen unter geringen zusätzlichen Voraussetzungen weitreichende Aussagen über die Existenz von Fortsetzungen.

Satz 2.29 (Hahn-Banach-Theorem). Es sei $(E, \|\cdot\|)$ ein normierter Raum und E_0 ein linearer Teilraum von E . Auf E_0 sei ein stetiges lineares Funktional f_0 definiert. Dann gibt es (mindestens) ein stetiges lineares Funktional f auf E , welches auf E_0 mit f_0 übereinstimmt ($f(x) = f_0(x)$ für $x \in E_0$) und für das die Gleichung S.2.29

$$\|f\| = \|f_0\| \quad (2.54)$$

gilt.

(Beweis s. [17].)

Bemerkung 2.15: Im Hinblick auf die Gl. (2.54) heißt f eine **normerhaltende Fortsetzung** (Erweiterung) von f_0 .

Eine wichtige Folgerung aus dem Hahn-Banach-Theorem ist der folgende Satz, der die Existenz nichttrivialer stetiger linearer Funktionale sichert.

Satz 2.30: Es sei $(E, \|\cdot\|)$ ein normierter Raum und $x_0 \neq 0$ ein (beliebiges) Element von E . Dann existiert ein stetiges lineares Funktional f auf E mit $\|f\| = 1$ und $f(x_0) = \|x_0\|$. (In der Ungleichung $|f(x)| \leq \|f\| \|x\|$, die generell für alle $x \in E$ gilt, steht also für $x = x_0$ das Gleichheitszeichen). S.2.30

Beweis (s. [17]): Mit E_0 bezeichnen wir die Menge $E_0 = \{x \in E \mid x = tx_0, t \text{ beliebig komplex}\}$. Dann ist E_0 ein linearer Teilraum von E . Durch die Vorschrift

$$f_0(x) = t \|x_0\| \quad (x = tx_0 \in E_0)$$

ist ein lineares Funktional auf E_0 gegeben. Es gilt

$$|f_0(x)| = |t| \|x_0\| = \|tx_0\| = \|x\| \quad (x \in E_0).$$

Hieraus ergibt sich die Gleichung $\|f_0\| = 1$, und aus Satz 2.29 folgt die Behauptung. ■

Für die Belange der Optimierungstheorie ist eine geometrische Form (bzw. Folgerung) des Hahn-Banach-Theorems wesentlich, die zu den sog. „Trennungssätzen“ gehört.

Definition 2.34: Es sei E ein reeller linearer Raum und $f \neq 0$ ein stetiges lineares Funktional auf E , sowie c eine (reelle) Konstante. Die Menge $[f = c] := \{x \in E \mid f(x) = c\}$ heißt die **Hyperebene** bezüglich f zum Niveau c . D.2.34

Beispiel 2.21: Es sei $E = R^3$ und $f(x) = \frac{\xi_1}{a} + \frac{\xi_2}{b} + \frac{\xi_3}{c}$, $x = (\xi_1, \xi_2, \xi_3)$, (a, b, c) reell und $\neq 0$.

Die Hyperebene $[f = 1]$ stimmt mit derjenigen Ebene im R^3 überein, die durch die Punkte $(a, 0, 0)$, $(0, b, 0)$, $(0, 0, c)$ geht.

Zwei konvexe Mengen des R^n kann man als „voneinander getrennt“ ansehen, wenn sie auf verschiedenen Seiten einer (Hyper-) Ebene liegen. In Bd. 15 werden Trennungssätze für konvexe Teilmengen des R^n aufgeführt. Die folgende Definition verallgemeinert diesen Sachverhalt auf konvexe Mengen in beliebigen linearen normierten Räumen. Eine Teilmenge M eines Vektorraumes heißt **konvex**, wenn mit je zwei Punkten $x, y \in M$ auch alle Punkte der Verbindungsstrecke $[x, y] = \{z = \lambda x + (1 - \lambda)y \mid 0 \leq \lambda \leq 1\}$ von x und y zu M gehören (s. Bd. 4, 2. Aufl., S. 12).

D.2.35 Definition 2.35: Es seien $(E, \|\cdot\|)$ ein reeller normierter Raum, $A \subseteq E$ und $B \subseteq E$ zwei konvexe Teilmengen von E . Man sagt, daß A und B **trennbar** sind, wenn es ein stetiges lineares Funktional f auf E und eine reelle Konstante c gibt mit

$$\begin{aligned} f(x) &\leq c \quad \text{für alle } x \in A, \\ f(x) &\geq c \quad \text{für alle } x \in B \quad (\text{s. Bild 2.2}). \end{aligned} \quad (2.55)$$

Falls sogar die Ungleichungen

$$\begin{aligned} f(x) &< c \quad \text{für alle } x \in A, \\ f(x) &> c \quad \text{für alle } x \in B \quad (\text{s. Bild 2.3}) \end{aligned} \quad (2.56)$$

gelten, so heißen die Mengen A und B **strikt trennbar**. $[f = c]$ heißt dann die A und B trennende bzw. strikt trennende Hyperebene.

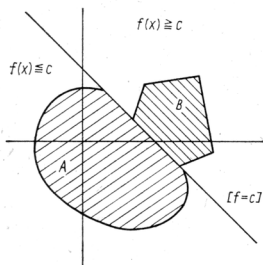


Bild 2.2

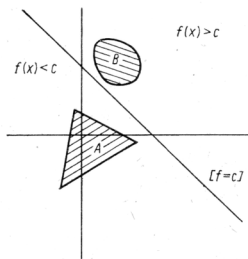


Bild 2.3

Ein für die Optimierung wichtiger Satz lautet dann:

S.2.31 Satz 2.31: Es seien $(E, \|\cdot\|)$ ein normierter Raum und $A \subseteq E$ eine konvexe Menge, die innere Punkte besitzt, d. h., es gibt eine offene Kugel (s. Def. 2.3), die ganz in A liegt. Ist x beliebiges Element von E , das nicht zu A gehört oder höchstens Randpunkt (d. h., in jeder Umgebung von x liegen sowohl Punkte von A als auch solche nicht aus A) von A ist, so sind A und $\{x\}$ trennbar. Sind A und B nichtleere offene konvexe Mengen mit leerem Durchschnitt, so sind A und B strikt trennbar.

(Für einen Beweis auf der Grundlage des Hahn-Banach-Theorems s. [32, S. 108].)

2.3.4. Schwache Konvergenz

In vielen Untersuchungen reichen Grenzübergänge „bezüglich der Norm“ in einem linearen Raum nicht aus [siehe z. B. (1.57)]. Neben der Norm-Konvergenz betrachtet man auch die sog. schwache Konvergenz. In diesem Kapitel werden nur die Definition und einige wesentliche Eigenschaften der schwachen Konvergenz angegeben.

Definition 2.36: Es sei $(E, \| \cdot \|)$ ein normierter Raum und $\{x_n\}$ eine Folge von Elementen von E . Die Folge $\{x_n\}$ heißt **schwach konvergent** gegen das Element $\bar{x} \in E$, wenn für jedes stetige lineare Funktional f auf E die Gleichheit $\lim_{n \rightarrow \infty} f(x_n) = f(\bar{x})$ gilt; abgekürzt $x_n \rightharpoonup \bar{x} \ (n \rightarrow \infty)$. **D.2.36**

Aus dieser Definition folgt sofort, daß jede (norm)konvergente Folge $\{x_n\}$ auch schwach konvergent ist. Die Umkehrung gilt jedoch i. allg. nicht. Hierzu betrachten wir das folgende Beispiel:

Beispiel 2.22: Es sei $(H, \langle \cdot | \cdot \rangle)$ ein Hilbertraum (s. Def. 2.37) und (e_n) ein vollständiges ONS (s. Def. 2.41) in H . Dann gilt

$$e_n \rightarrow 0 \quad (n \rightarrow \infty),$$

obwohl alle e_n die Norm 1 haben ($\|e_n\| = 1$) und daher die Folge $\{e_n\}$ in H nicht gegen 0 konvergiert.

Zum Beweis betrachte man ein beliebiges Element $x \in H$. Dann gilt $x = \sum_{n=1}^{\infty} \langle e_n | x \rangle e_n$. Für ein beliebiges stetiges Funktional f auf H folgt daraus (wegen der Stetigkeit von f) die Gleichung

$$f(x) = \sum_{n=1}^{\infty} \langle e_n | x \rangle f(e_n). \quad (*)$$

Nach dem Satz von F. Riesz (Satz 2.26) hat das Funktional $f(x)$ die Gestalt

$$f(x) = \langle y | x \rangle \quad (x \in H) \quad (**)$$

mit einem eindeutig bestimmten $y \in H$, für welches $\|f\| = \|y\|$ gilt. Einsetzen von (**) in (*) ergibt die Gleichheit

$$\langle y | x \rangle = \sum_{n=1}^{\infty} \langle e_n | x \rangle \langle y | e_n \rangle \quad (x \in H). \quad (***)$$

Setzen wir in (***) für x speziell $x = y$, so erhalten wir

$$\|y\|^2 = \langle y | y \rangle = \sum_{n=1}^{\infty} \langle e_n | y \rangle \langle y | e_n \rangle = \sum_{n=1}^{\infty} \langle e_n | y \rangle \overline{\langle e_n | y \rangle} = \sum_{n=1}^{\infty} |\langle e_n | y \rangle|^2.$$

Die rechtsstehende Reihe ist somit konvergent. Daher bilden (s. Bd. 3) ihre Glieder $|\langle e_n | y \rangle|^2$ eine Nullfolge. Daher ist aber auch die Zahlenfolge $f(e_n) = \langle y | e_n \rangle \ (n = 1, 2, \dots)$ eine gegen null konvergente Folge, womit alles gezeigt ist.

Satz 2.32: Es sei $(E, \| \cdot \|)$ ein normierter Raum und $\{x_n\}$ eine Folge aus E , die gegen das Element \bar{x} schwach konvergiert. Dann gilt die Ungleichung **S.2.32**

$$\|\bar{x}\| \leq \liminf \|x_n\|, \quad (2.57)$$

und es existiert eine Konstante $K > 0$ mit $\|x_n\| \leq K \ (n = 1, 2, \dots)$ (m. a. W., jede schwach konvergente Folge ist beschränkt).

In speziellen Räumen läßt sich die schwache Konvergenz genauer kennzeichnen, wie der folgende Satz am Beispiel des Raumes $C[a, b]$ zeigt. Der Raum $C[a, b]$ sei hierbei wie üblich mit der Maximum-Norm versehen.

S.2.33 Satz 2.33 (s. [17, S. 226]): Eine Folge $\{x_n(\cdot)\}$ aus $C[a, b]$ konvergiert genau dann schwach gegen ein Element $\bar{x}(\cdot)$ aus $C[a, b]$, wenn folgende Bedingungen beide (gleichzeitig) erfüllt sind

- 1) $|x_n(t)| \leq M$ ($t \in [a, b]$; $n = 1, 2, \dots$) für ein $M > 0$;
- 2) $\lim_{n \rightarrow \infty} x_n(t) = \bar{x}(t)$ ($t \in [a, b]$) (punktweise Konvergenz).

Sehr wichtig ist noch der Satz über die Charakterisierung der reflexiven Banachräume durch die schwache (Folgen-) Kompaktheit ihrer (Norm-) Kugeln, der auf W. F. Eberlein, L. Alaoglu, W. Schmuljan zurückgeht:

S.2.34 Satz 2.34 (s. [21, S. 50]): Ein Banachraum ist genau dann reflexiv, wenn seine abgeschlossene Einheitskugel $K = \{x \mid \|x\| \leq 1\}$ schwach folgenkompakt ist, d. h., wenn jede Folge aus K eine schwach konvergente Teilfolge enthält.

2.4. Hilberträume, Orthogonalentwicklungen

2.4.1. Grundbegriffe, Beispiele

Nach Def. 1.3 hatten wir den Begriff des Prä-Hilbertraumes H eingeführt. Dessen Skalarprodukt erzeugt eine Norm vermöge

$$\|x\| = \sqrt{\langle x \mid x \rangle} \quad (x \in H). \quad (2.58)$$

D.2.37 Definition 2.37: Ein Prä-Hilbertraum heißt ein **Hilbertraum**, wenn er bezüglich der durch die Norm (2.58) erzeugten Metrik (s. 1.1.) ein vollständiger metrischer Raum ist. Allgemeine Hilberträume bezeichnen wir mit dem Buchstaben H .

(D. Hilbert, 1862–1943, schuf die Grundlagen der Theorie linearer Operatoren in speziellen Hilberträumen.)

Bemerkung 2.16: Jeder Hilbertraum ist ein Banachraum, aber nicht umgekehrt. Ein Banachraum $(E, \|\cdot\|)$ ist genau dann ein Hilbertraum, wenn die Norm von E die sog. **Parallelogrammgleichung**

$$\|x - y\|^2 + \|x + y\|^2 = 2(\|x\|^2 + \|y\|^2) \quad (x, y \in E) \quad (2.59)$$

erfüllt. Ist die Gl. (2.59) erfüllt, so existiert auf E ein Skalarprodukt $\langle x \mid y \rangle$, für welches die Gl. (2.58) gilt (wobei auf der linken Seite dieser Gleichung die gegebene Norm von E steht).

Die rechts stehende Liste nennt Beispiele für Hilberträume mit dem jeweiligen Skalarprodukt. Alle aufgeführten Räume sind komplexe lineare Räume. Ist $z = a + ib$ eine komplexe Zahl, so bezeichnet $\bar{z} = a - ib$ die konjugierte komplexe Zahl.

Das Skalarprodukt $\langle x \mid y \rangle$ ist eine stetige Funktion von x und y . Dies zeigt der folgende Satz.

S.2.35 Satz 2.35: Es seien H ein Prä-Hilbertraum (speziell: ein Hilbertraum) und $\{x_n\}$ bzw. $\{y_n\}$ zwei Folgen aus H , die gegen bestimmte Elemente x bzw. y aus H konvergieren: $\lim_{n \rightarrow \infty} x_n = x$, $\lim_{n \rightarrow \infty} y_n = y$. Dann gilt

$$\lim_{n \rightarrow \infty} \langle x_n \mid y_n \rangle = \langle x \mid y \rangle. \quad (2.60)$$

| Elemente des Raumes $x, y; f, g$ | Bezeichnung des Raumes | Skalarprodukt $\langle x y \rangle$ bzw. $\langle f g \rangle$ |
|--|------------------------|---|
| n -tupel komplexer Zahlen ξ_j ($1 \leq j \leq n$) $x = (\xi_1, \dots, \xi_n)$, $y = (\eta_1, \dots, \eta_n)$ | C^n | $\sum_{j=1}^n \bar{\xi}_j \eta_j$ |
| Folgen komplexer Zahlen ξ_j ($j = 1, 2, \dots$) mit $\sum_{j=1}^{\infty} \xi_j ^2 < +\infty$, $x = (\xi_1, \xi_2, \dots, \xi_n, \dots)$, $y = (\eta_1, \eta_2, \dots, \eta_n, \dots)$ | l^2 | $\sum_{j=1}^{\infty} \bar{\xi}_j \eta_j$ |
| (Klassen) komplexwertiger quadratisch absolut summierbarer Funktionen f, g, \dots , die auf einem Gebiet $\Omega \subseteq R^n$ definiert sind $\int_{\Omega} f(x) ^2 dx < +\infty$ | $L^2(\Omega)$ | $\int_{\Omega} \overline{f(x)} g(x) dx$ |
| (Klassen fast überall übereinstimmender) komplexwertiger quadratisch summierbarer Funktionen f, g, \dots , die im Gebiet $\Omega \subseteq R^n$ verallgemeinerte Ableitungen bis zur Ordnung m besitzen ($m = 1, 2, \dots$; fest) und für die diese verallgemeinerten Ableitungen zu $L^2(\Omega)$ gehören | $W^{m,2}(\Omega)$ | $\sum_{0 \leq \alpha \leq m} \int_{\Omega} D^{\alpha} \overline{f(x)} D^{\alpha} g(x) dx$ |
| (Klassen fast überall übereinstimmender) komplexwertiger (bzw. reellwertiger) absolut quadratisch summierbarer Funktionen, die auf dem Intervall $[a, b]$ definiert sind | $L^2[a, b]$ | $\int_a^b \overline{f(x)} g(x) dx$ |

Beweis: Es gilt (Rechenregeln für $\langle x | y \rangle$, s. 1.1.)

$$0 \leq |\langle x | y \rangle| - \langle x_n | y_n \rangle = |\langle x - x_n | y \rangle + \langle x_n | y - y_n \rangle| \\ \leq |\langle x - x_n | y \rangle| + |\langle x_n | y - y_n \rangle| \leq \|x - x_n\| \|y\| + \|x_n\| \|y - y_n\| = \alpha_n$$

($n = 1, 2, \dots$) [Dreiecksungleichung für komplexe Zahlen und (1.10)]. Wegen $\|x - x_n\| \rightarrow 0$, $\|y_n - y\| \rightarrow 0$, $\|x_n\| \rightarrow \|x\|$ für $n \rightarrow \infty$ bilden die Zahlen α_n eine Nullfolge ($\lim_{n \rightarrow \infty} \alpha_n = 0$), und daher ist auch $|\langle x | y \rangle - \langle x_n | y_n \rangle|$ eine Nullfolge, woraus die Behauptung folgt. ■

2.4.2. Orthogonalentwicklungen

Definition 2.38: Es sei H ein Prä-Hilbertraum. Die Elemente $x, y \in H$ heißen (zu- **D.2.38**
einander) orthogonal, wenn $\langle x | y \rangle = 0$ gilt. Eine Teilmenge von H heißt ein Orthogonalsystem, wenn je zwei ihrer Elemente orthogonal sind. Ein Orthogonalsystem,

welches nur Elemente x mit $\|x\| = 1$ enthält, heißt ein **Orthonormalsystem** kurz: ein ONS.

S.2.36 Satz 2.36 (Pythagoras): Ist $G \subseteq H$ ein Orthogonalsystem im Prä-Hilbertraum H und sind z_1, \dots, z_n Elemente von G , so gilt

$$\|z_1 + z_2 + \dots + z_n\|^2 = \|z_1\|^2 + \|z_2\|^2 + \dots + \|z_n\|^2. \quad (2.61)$$

Beweis: Es gilt

$$\begin{aligned} \left\| \sum_{j=1}^n z_j \right\|^2 &= \left\langle \sum_{j=1}^n z_j \left| \sum_{r=1}^n z_r \right. \right\rangle = \sum_{j=1}^n \sum_{r=1}^n \langle z_r | z_j \rangle \\ &= \sum_{j=1}^n \langle z_j | z_j \rangle = \sum_{j=1}^n \|z_j\|^2. \quad \blacksquare \end{aligned}$$

S.2.37 Satz 2.37 (Schmidtsches Orthogonalisierungsverfahren): Ist x_1, x_2, \dots eine (endliche oder unendliche) Folge linear unabhängiger Vektoren eines Prä-Hilbertraumes H , so gibt es ein ONS e_1, e_2, \dots , welches den gleichen linearen Teilraum von H erzeugt, d. h., die Menge aller endlichen Linearkombinationen $\sum_{j=1}^n \alpha_j e_j$ stimmt mit der Menge aller endlichen Linearkombinationen $\sum_{k=1}^m \beta_k x_k$ überein ($n; m = 1, 2, \dots, \alpha_j, \beta_k$ beliebige komplexe Zahlen). Das ONS e_1, e_2, \dots kann auf die folgende Weise berechnet werden:

$$\begin{aligned} e_1 &= \frac{1}{\|x_1\|} x_1, \\ e_n &= \frac{x_n - \sum_{k=1}^{n-1} \langle e_k | x_n \rangle e_k}{\left\| x_n - \sum_{k=1}^{n-1} \langle e_k | x_n \rangle e_k \right\|} \quad \text{für } n = 2, 3, \dots \end{aligned} \quad (2.62)$$

Wie man leicht nachrechnet, gilt $\langle e_n | e_m \rangle = 0$ für $n \neq m$ sowie $\|e_n\| = 1$ ($n = 1, 2, \dots$). Mittels vollständiger Induktion ergeben sich die weiteren Aussagen des Satzes.

Wie wir bereits in den einführenden Beispielen gesehen haben, spielen die ONS bei der Darstellung von Elementen eines (Prä-) Hilbertraumes im Hinblick auf die Approximation allgemeiner Vektoren (= Funktionen, die Elemente eines Funktionenraumes sind) durch besonders übersichtliche und einfache Elemente eine wichtige Rolle (s. auch Bd. 12, Kap. 1.).

Im folgenden treten unendliche Reihen $\sum_{n=1}^{\infty} x_n$ auf, deren Glieder x_n Elemente eines Hilbertraumes sind. Die Konvergenz dieser Reihen wird analog zur Konvergenz von Zahlenreihen (Bd. 3) erklärt:

D.2.39 Definition 2.39: Es sei $\{x_k\}$ eine Folge von Elementen eines Hilbertraumes H . Der Ausdruck $\sum_{k=1}^{\infty} x_k$ bezeichnet einerseits die Folge der zugehörigen Partialsummen $s_n = \sum_{k=1}^n x_k$ ($n = 1, 2, \dots$) und wird **unendliche Reihe** genannt. Die unendliche Reihe

$\sum_{k=1}^{\infty} x_k$ heißt **konvergent**, wenn die Folge $\{s_n\}$ der Partialsummen konvergiert. In diesem Fall schreibt man $s = \lim_{n \rightarrow \infty} s_n = \sum_{k=1}^{\infty} x_k$ und nennt $s = \sum_{k=1}^{\infty} x_k$ andererseits auch den Wert der unendlichen Reihe $\sum_{k=1}^{\infty} x_k$. Eine unendliche Reihe, die nicht konvergent ist, heißt **divergent**.

Satz 2.38: Es sei H ein Hilbertraum, und die Folge $\{z_n\}$ sei ein Orthogonalsystem in H . **S.2.38**

Die unendliche Reihe $\sum_{n=1}^{\infty} z_n$ konvergiert genau dann, wenn $\sum_{n=1}^{\infty} \|z_n\|^2$ konvergiert. Gilt $\sum_{n=1}^{\infty} z_n = z$, so ist

$$\|z\|^2 = \sum_{n=1}^{\infty} \|z_n\|^2. \quad (2.63)$$

Beweis: Für $n > m$ gilt mit $s_k = \sum_{j=1}^k z_j$ ($k = 1, 2, \dots$) die Gleichheit (Satz des Pythagoras)

$$\|s_n - s_m\|^2 = \left\| \sum_{j=m+1}^n z_j \right\|^2 = \sum_{j=m+1}^n \|z_j\|^2 = \sum_{j=1}^n \|z_j\|^2 - \sum_{j=1}^m \|z_j\|^2.$$

Daher ist die Folge $\left\{ \sum_{k=1}^n z_k \right\} = \{s_n\}$ genau dann eine Cauchy-Folge in H (s. Def. 2.6), wenn die Folge $\left\{ \sum_{k=1}^n \|z_k\|^2 \right\}$ eine Cauchy-Folge (in \mathbf{R}) ist. Wegen der Vollständigkeit von H ergibt sich die erste der obigen Behauptungen. Es gelte jetzt $\sum_{k=1}^{\infty} z_k = z$, d.h., die Folge $\{s_n\} = \left\{ \sum_{k=1}^n z_k \right\}$ konvergiert gegen z . Nach Pythagoras (Satz 2.36) gilt $\|s_n\|^2 = \sum_{k=1}^n \|z_k\|^2$. Wegen $\lim_{n \rightarrow \infty} \|s_n - z\| = 0$ gilt auch $\lim_{n \rightarrow \infty} \|s_n\| = \|z\|$ (Dreiecksungleichung, Stetigkeit der Norm). Also ist auch $\lim_{n \rightarrow \infty} \|s_n\|^2 = \|z\|^2$, woraus die zweite Behauptung sofort folgt. ■

Definition 2.40: Es sei (e_n) ein ONS im Hilbertraum H . Ist x ein beliebiges Element von H , so heißt die Zahl **D.2.40**

$$\langle e_k | x \rangle \quad (k = 1, 2, \dots) \quad (2.64)$$

der k -te **Fourierkoeffizient** von x bezüglich des gegebenen ONS.

Bei der Entwicklung nach Orthogonalfunktionen (s. Bd. 11) ist wesentlich, daß das vorliegende ONS umfangreich genug ist, um alle Elemente des betrachteten Raumes approximieren zu können. Zum Beispiel bildet das System $\{e_1, e_2\}$ mit $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$ zwar ein ONS im (reellen Hilbertraum) R^3 . Der Vektor $x_0 = (1, 1, 1)$ hat jedoch von allen Linearkombinationen $\sum_{j=1}^2 \alpha_j e_j = \alpha_1 e_1 + \alpha_2 e_2 = (\alpha_1, \alpha_2, 0)$ einen Abstand $\left\| x_0 - \sum_{j=1}^2 \alpha_j e_j \right\| = \sqrt{(\alpha_1 - 1)^2 + (\alpha_2 - 1)^2 + 1} \geq 1$, kann also durch dieses ONS nicht beliebig genau approximiert werden. Die entscheidende Eigenschaft eines ONS, die eine solche Approximierbarkeit gewährleistet, ist die sog. **Vollständigkeit** des ONS.

D.2.41 Definition 2.41: Es sei (e_n) ein ONS im Hilbertraum H . Das ONS (e_n) heißt **vollständig**, wenn es keinen vom Nullvektor verschiedenen Vektor z gibt, der zu allen Vektoren e_n orthogonal ist.

Der entscheidende **Entwicklungssatz** lautet (s. [8, S. 130/132]):

S.2.39 Satz 2.39 (von der Orthogonalentwicklung): Es sei H ein Hilbertraum und (e_n) ein ONS in H . Dann sind die folgenden Aussagen gleichwertig:

- (1) Das ONS (e_n) ist vollständig.
- (2) Für jedes $x \in H$ gilt (**Fourierentwicklung von x**)

$$x = \sum_{n=1}^{\infty} \langle e_n | x \rangle e_n. \quad (2.65)$$

- (3) Für jedes $x \in H$ gilt (**Parsévalsche Gleichung**)

$$\|x\|^2 = \sum_{n=1}^{\infty} |\langle e_n | x \rangle|^2. \quad (2.66)$$

Beispiel 2.23: Es sei $H = L^2[0, 2\pi]$. Dann ist das folgende Funktionensystem ein vollständiges ONS:

$$\frac{e^{-int}}{\sqrt{2\pi}} \quad (n = 0, \pm 1, \pm 2, \dots). \quad (2.67)$$

Weitere Beispiele vollständiger ONS erhält man (Bd. 12, S. 14–19) durch Orthogonalisierung der Funktionenfolge $f_n(x) = x^n \sqrt{p(x)}$, wobei $p(x) \geq 0$ eine sog. Belegungsfunktion bezeichnet. Auf diese Weise erhält man z. B. im Raum $L^2_{\mathbf{R}}[-1, 1]$ und für $p(x) = 1$ die **normierten Legendreschen Polynome**

$$P_n(x) = \sqrt{\frac{2n+1}{2}} \frac{1}{2^n n!} \frac{d^n}{dx^n} (x^2 - 1)^n \quad (n = 0, 1, 2, \dots); \quad (2.68)$$

und im $L^2_{\mathbf{R}}(\mathbf{R})$ für $p(x) = e^{-x^2}$ die **normierten Hermiteischen Funktionen**

$$H_n(x) = \frac{(-1)^n}{\sqrt{2^n n!} \sqrt{\pi}} e^{\frac{x^2}{2}} \frac{d^n(e^{-x^2})}{dx^n} \quad (n = 0, 1, 2, \dots) \quad (2.69)$$

als ONS [s. auch (1.81)].

2.4.3. Orthogonales Komplement, orthogonale direkte Summe

D.2.42 Definition 2.42: Es sei H ein Prä-Hilbertraum und E eine Teilmenge von H . Die Menge aller Elemente $x \in H$, die auf allen Vektoren aus E senkrecht stehen, bezeichnet man mit $H \ominus E$; sie heißt das **orthogonale Komplement von E bezüglich H** (bzw. in H):

$$H \ominus E = \{x \in H \mid \langle x | y \rangle = 0 \text{ für alle } y \in E\}. \quad (2.70)$$

S.2.40 Satz 2.40: Unter den Voraussetzungen der obigen Definition ist $H \ominus E$ ein abgeschlossener linearer Teilraum von H .

Beweis: Ist $x \in H \ominus E$, so gilt für beliebiges komplexes λ : $\langle \lambda x | y \rangle = \lambda \langle x | y \rangle = 0$ für alle $y \in E$, also ist auch $\lambda x \in H \ominus E$. Sind $x_1, x_2 \in H \ominus E$, so auch $x_1 + x_2$, weil $\langle x_1 + x_2 | y \rangle = \langle x_1 | y \rangle + \langle x_2 | y \rangle = 0$ für alle $y \in E$ gilt. Gilt schließlich $x_n \in H \ominus E$ und $x = \lim_{n \rightarrow \infty} x_n$, so ist $\langle x | y \rangle = \lim_{n \rightarrow \infty} \langle x_n | y \rangle = 0$ für alle $y \in E$. ■

Beispiel 2.24: Es sei H ein Hilbertraum und (e_n) ein vollständiges ONS in H . Es sei $E = \{e_1, \dots, e_m\}$ die Teilmenge dieses ONS, die aus den ersten m Elementen e_r besteht.

Dann besteht $H \ominus E$ aus allen Elementen x der Form

$$x = \sum_{k=m+1}^{\infty} c_k e_k \quad \text{mit} \quad \sum_{k=m+1}^{\infty} |c_k|^2 < +\infty. \quad (2.71)$$

Denn $\langle x | e_r \rangle = \sum_{k=m+1}^{\infty} \bar{c}_k \langle e_k | e_r \rangle = 0$ für $r = 1, \dots, m$; also liegt jedes x der Gestalt (2.71) in $H \ominus E$. Ist umgekehrt $x \in H \ominus E$, so gilt nach dem Entwicklungssatz (Satz 2.39)

$$x = \sum_{k=1}^{\infty} \langle e_k | x \rangle e_k.$$

Weil $x \in H \ominus E$ ist, muß $\langle e_k | x \rangle = 0$ für $k = 1, \dots, m$ gelten. Also ist

$$x = \sum_{k=m+1}^{\infty} \langle e_k | x \rangle e_k = \sum_{k=m+1}^{\infty} c_k e_k \quad \text{mit} \quad c_k = \langle e_k | x \rangle$$

($k = m+1, m+2, \dots$). Aus dem Entwicklungssatz folgt weiter, daß $+\infty > \|x\|^2 = \sum_{k=m+1}^{\infty} |\langle e_k | x \rangle|^2 = \sum_{k=m+1}^{\infty} |c_k|^2$ gilt, woraus die Konvergenz der rechtsstehenden Reihe folgt. Mit anderen Worten, das orthogonale Komplement einer Menge endlich vieler Elemente eines (vollständigen) ONS besteht aus allen (Fourier-) Reihen, in denen nur die restlichen Elemente des ONS auftreten.

Definition 2.43: Es seien H ein Hilbertraum und H_1 bzw. H_2 abgeschlossene lineare D.2.43 Teilräume von H . Ist $x_1 \in H_1$ und $x_2 \in H_2$, so gelte stets $\langle x_1 | x_2 \rangle = 0$ (man sagt dann: die Teilräume H_1 und H_2 sind zueinander orthogonal). Die Menge

$$\{x \in H \mid x = x_1 + x_2, x_1 \in H_1, x_2 \in H_2\}$$

heißt die orthogonale direkte Summe von H_1 und H_2 und wird bezeichnet mit

$$H_1 \oplus H_2. \quad (2.72)$$

Satz 2.41: Unter den in Def. 2.43 getroffenen Voraussetzungen ist die orthogonale S.2.41 direkte Summe von H_1 und H_2 stets ein abgeschlossener linearer Teilraum von H .

Es ist in Analogie zu Def. 2.43 klar, wie die orthogonale direkte Summe $H_1 \oplus H_2 \oplus \dots \oplus H_n$ endlich vieler, paarweise orthogonaler linearer Teilräume zu definieren ist.

Satz 2.42: Es sei H_1 ein abgeschlossener linearer Teilraum eines Hilbertraumes H und S.2.42 $H_2 = H \ominus H_1$ das orthogonale Komplement von H_1 in H . Dann ist H die orthogonale direkte Summe von H_1 und H_2 : $H = H_1 \oplus H_2$.

Dieser und der folgende Satz zeigen, daß die Bildung des orthogonalen Komplements und der orthogonalen direkten Summe zueinander invers (komplementär) sind.

S.2.43 Satz 2.43: Es sei H ein Hilbertraum, und es gelte $H = H_1 \oplus H_2$, wobei H_1, H_2 zwei abgeschlossene lineare Teilräume von H bezeichnen. Dann gelten die Gleichungen

$$\begin{aligned} H_1 \cap H_2 &= \{0\}, \\ H_1 &= H \ominus H_2, \\ H_2 &= H \ominus H_1. \end{aligned} \quad (2.73)$$

Außerdem läßt sich jedes $x \in H$ auf genau eine Weise in der Form

$$x = x_1 + x_2 \quad (x_1 \in H_1, x_2 \in H_2) \quad (2.74)$$

darstellen.

Beweis: Wir zeigen nur die erste der Gln. (2.73) sowie die Gl. (2.74). Da H_1 und H_2 lineare Teilräume von H sind, gilt $0 \in H_1 \cap H_2$. Ist andererseits $x \in H_1 \cap H_2$, so gilt nach Def. 2.43 der orthogonalen direkten Summe, daß $\langle x | x \rangle = 0$ sein muß; d. h. aber $x = 0$. Somit ist $H_1 \cap H_2 = \{0\}$. Wegen $H = H_1 \oplus H_2$ gibt es für jedes $x \in H$ stets mindestens eine Darstellung der Form

$$x = x_1 + x_2, \quad x_1 \in H_1, x_2 \in H_2.$$

Gilt zusätzlich

$$x = x'_1 + x'_2, \quad x'_1 \in H_1, x'_2 \in H_2,$$

so wird wegen

$$x_1 + x_2 = x'_1 + x'_2 \quad \text{auch} \quad x_1 - x'_1 = x'_2 - x_2.$$

$x_1 - x'_1$ ist ein Element von H_1 , $x'_2 - x_2$ gehört zu H_2 . Da diese beiden Elemente gleich sind, gehören sie sowohl zu H_1 als auch zu H_2 , also zu $H_1 \cap H_2$, und daher gilt nach dem zuvor bewiesenen, daß $x_1 - x'_1 = x'_2 - x_2 = 0$ sein muß, woraus schließlich $x_1 = x'_1$, $x_2 = x'_2$ folgt. Es gibt also nur eine Zerlegung von x in der Form (2.74). ■

S.2.44 Satz 2.44: Es sei H_1 ein abgeschlossener linearer Teilraum des Hilbertraumes H . Zu jedem $x \in H$ gibt es genau ein Element $x_1 \in H_1$, welches von x einen minimalen Abstand (bezüglich H_1) besitzt (die Projektion von x auf H_1):

$$\|x - x_1\| = \inf_{y \in H_1} \|x - y\|. \quad (2.75)$$

Der Vektor $x - x_1 = x_2$ gehört dann zum orthogonalen Komplement $H \ominus H_1$.

(Zum Beweis s. [17].)

Zur Veranschaulichung des Satzes 2.44 betrachten wir im reellen Hilbertraum $H = L^2_{\mathbb{R}}[a, b]$ ($a = 0, b = 2\pi$) das vollständige ONS

$$\frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{\pi}} \cos t, \frac{1}{\sqrt{\pi}} \sin t, \dots, \frac{1}{\sqrt{\pi}} \cos nt, \frac{1}{\sqrt{\pi}} \sin nt, \dots \quad (0 \leq t \leq 2\pi).$$

H_1 bestehe aus allen trigonometrischen Polynomen der Ordnung $\leq n$:

$$y(t) = \frac{\alpha'_0}{2} + \sum_{k=1}^n (\alpha'_k \cos kt + \beta'_k \sin kt) \quad (0 \leq t \leq 2\pi). \quad (2.76)$$

Ist $x \in H$ beliebig, so läßt sich x (s. den Entwicklungssatz 2.39) in der Form

$$\begin{aligned} x(t) &= \frac{a_0}{\sqrt{2\pi}} + \sum_{k=1}^{\infty} \left(a_k \frac{\cos kt}{\sqrt{\pi}} + b_k \frac{\sin kt}{\sqrt{\pi}} \right) \\ &= \frac{\alpha_0}{2} + \sum_{k=1}^{\infty} (\alpha_k \cos kt + \beta_k \sin kt) \quad (0 \leq t \leq 2\pi) \end{aligned} \quad (2.77)$$

darstellen, wobei diese Reihe im Sinne des Raumes $L^2[0, 2\pi]$ gegen x konvergiert (die Folge der Partialsummen $s_n(t)$ konvergiert im quadratischen Mittel gegen $x(t)$, d. h.,

$$\lim_{n \rightarrow \infty} \int_0^{2\pi} (s_n(t) - x(t))^2 dt = 0).$$

Es gelten die Beziehungen

$$\alpha_0 = a_0 \sqrt{\frac{2}{\pi}}, \quad \alpha_k = \frac{a_k}{\sqrt{\pi}}, \quad \beta_k = \frac{b_k}{\sqrt{\pi}} \quad (k = 1, 2, \dots)$$

und

$$\begin{aligned} a_0 &= \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} x(t) dt, & a_k &= \frac{1}{\sqrt{\pi}} \int_0^{2\pi} x(t) \cos kt dt, \\ b_k &= \frac{1}{\sqrt{\pi}} \int_0^{2\pi} x(t) \sin kt dt \quad (k = 1, 2, \dots). \end{aligned}$$

Das Element $x_1 \in H_1$, welches gemäß Satz 2.44 den kürzesten Abstand zwischen H_1 und x realisiert, ist dann genau die n -te Partialsumme der Reihe (2.77):

$$x_1 = x_1(t) = \frac{\alpha_0}{2} + \sum_{k=1}^n (\alpha_k \cos kt + \beta_k \sin kt) \quad (0 \leq t \leq 2\pi).$$

Das Element $x_2 = x - x_1$ hat die Form

$$x_2 = x_2(t) = \sum_{k=n+1}^{\infty} (\alpha_k \cos kt + \beta_k \sin kt) \quad (0 \leq t \leq 2\pi)$$

und gehört (offensichtlich) zum orthogonalen Komplement von H_1 .

Zur Ergänzung der Ausführungen über Hilbert-Räume erwähnen wir noch folgenden Satz von F. Riesz und E. Fischer.

Satz 2.45: Es sei H ein Hilbert-Raum und $\{e_n\}$ ein vollständiges ONS in H . Konvergiert für eine Zahlenfolge $\{a_k\}$ die Reihe $\sum_{k=1}^{\infty} |a_k|^2$, dann gibt es genau ein Element $x \in H$, dessen Fourierkoeffizienten bezüglich $\{e_n\}$ gerade mit den gegebenen Werten a_k übereinstimmen; $a_k = \langle e_k | x \rangle$ ($k = 1, 2, \dots$).

Aus diesem Satz (zum Beweis s. [17]) folgt die Normisomorphie aller (separablen) Hilbert-Räume.

3. Lineare Operatoren

Im Zusammenhang mit der Lösung sog. „linearer Probleme“, d. h. von Problemen, in denen die wesentlichen Zusammenhänge ein lineares Verhalten zeigen, untersucht die Funktionalanalysis lineare Operatoren und Gleichungen, in denen lineare Operatoren auftreten. Die einfachste Grundaufgabe dieser Art ist die Frage nach den Lösbarkeitseigenschaften eines linearen Gleichungssystems mit endlich vielen Unbekannten ξ_1, \dots, ξ_n

$$\sum_{k=1}^n a_{jk} \xi_k = b_j \quad (j = 1, \dots, m), \quad (3.1)$$

wobei die Koeffizienten a_{jk} und die absoluten Glieder b_j gegebene (reelle oder komplexe) Zahlen sind.

Kennzeichnend für die funktionalanalytische Denkweise ist die Auffassung, daß der Koeffizientenmatrix (a_{jk}) ein Operator T entspricht, der jeden Vektor $x = (\xi_1, \dots, \xi_n)$ in einen Vektor $y = (\eta_1, \dots, \eta_m)$ überführt:

$$\eta_j = \sum_{k=1}^n a_{jk} \xi_k \quad (j = 1, \dots, m) \quad (3.2)$$

oder

$$y = T(x) \quad (\text{oder: } y = Tx). \quad (3.3)$$

Die Aufgabe, das lineare Gleichungssystem (3.1) zu lösen, kann also in der folgenden Weise formuliert werden:

Man bestimme alle x , die die Gleichung

$$Tx = b \quad (3.4)$$

($b = (b_1, \dots, b_m)$) erfüllen.

Der Operator T hat ersichtlich folgende Eigenschaften

$$\begin{aligned} \text{Additivität:} \quad T(x^{(1)} + x^{(2)}) &= T(x^{(1)}) + T(x^{(2)}), \\ \text{Homogenität:} \quad T(\lambda x) &= \lambda T(x) \quad (\lambda \text{ reell oder komplex}). \end{aligned} \quad (3.5)$$

Diese Tatsache veranlaßt die folgende Definition.

D.3.1 Definition 3.1: Eine Abbildung (ein Operator) $T: E \rightarrow F$, die einen linearen Raum E in einen linearen Raum F abbildet, heißt **linear**, wenn für alle $x^{(1)}, x^{(2)}, x$ aus E , und alle (reellen bzw. komplexen) Zahlen λ die Gleichungen (3.5) gelten.¹⁾

Ein Operator der Form (3.3) ist ersichtlich ein linearer Operator, der den Raum K^n (n -dimensionaler komplexer Vektorraum) in den Raum K^m abbildet. Die Auffassung des linearen Gleichungssystems (3.1) als lineare Operatorgleichung (3.4) (s. auch 1.2.4.) ist deshalb von so grundsätzlicher Bedeutung, weil sich die wichtigsten Ergebnisse über lineare Gleichungen mit endlich vielen Unbekannten auf allgemeinere Operatorengleichungen übertragen lassen, welche insbesondere solche Differential- bzw. Integralgleichungen umfassen, die in der Praxis häufig auftreten. Beispiele für lineare Operatoren werden in den folgenden Abschnitten eingeführt.

¹⁾ Vgl. auch Bd. 1, 8.4.

3.1. Das Rechnen mit linearen Operatoren

Ein linearer Operator $T: E \rightarrow F$, der den linearen Raum E in den linearen Raum F abbildet, legt folgende Teilräume fest:

den **Kern**¹⁾ von T , abgekürzt $\text{Ker } T$, der aus allen Nullstellen von T besteht, d. h. (o_F bezeichne das Nullelement von F)

$$\text{Ker } T = \{x \in E \mid Tx = o_F\}, \quad (3.6)$$

den **Rang** oder **Wertebereich** von T , abgekürzt $\text{Ran } T$ oder $R(T)$ oder $T[E]$, der aus allen Elementen von F besteht, die als Werte der Abbildung auftreten:

$$\text{Ran } T = \{y \in F \mid y = Tx \text{ für ein } x \in E\}. \quad (3.7)$$

Zufolge der Linearität von T sind $\text{Ker } T$ und $\text{Ran } T$ lineare Teilräume von E bzw. von F .

Schließlich muß man (vor allem für die Theorie unbeschränkter Operatoren) den Fall vorsehen, daß ein linearer Operator T nur auf einem linearen Teilraum E_0 von E definiert ist, $T: E_0 \rightarrow F$, deutlicher, $T: E_0 \subseteq E \rightarrow F$. In diesem Fall hebt man den **Definitionsbereich** $E_0 = D(T)$ von T besonders hervor.

Im folgenden seien T, S, U, \dots lineare Operatoren von E in F . Ist $F = E$, so spielt der **identische Operator** I , erklärt durch

$$Ix = x \quad (x \in E) \quad (3.8)$$

eine besondere Rolle.

Definition 3.2: Die Summe $T + S$ der (linearen) Operatoren T und S wird durch die **D.3.2** Gleichung

$$(T + S)(x) = T(x) + S(x) \quad (x \in E) \quad (3.9)$$

erklärt. Die Multiplikation des Operators T mit der (komplexen bzw. reellen) Zahl λ wird definiert durch die Gleichung

$$(\lambda T)(x) = \lambda T(x) \quad (x \in E). \quad (3.10)$$

Satz 3.1: Mittels der in Def. 3.2 eingeführten Operationen für lineare Operatoren **S.3.1** $T: E \rightarrow F$ wird die Menge dieser Operatoren zu einem Vektorraum (linearen Raum).

Definition 3.3: Sind E, F, G lineare Räume und $T: E \rightarrow F$; $S: F \rightarrow G$ lineare Operatoren, dann wird das **Produkt** ST der Operatoren T und S durch die Zuordnungsvorschrift **D.3.3**

$$(ST)(x) = S(T(x)) \quad (x \in E) \quad (3.11)$$

erklärt.

Satz 3.2: Das Produkt ST (im Sinne der Def. 3.3) ist ein linearer Operator von E in G : **S.3.2**

$$ST: E \rightarrow G.$$

¹⁾ Synonym: Nullraum von T : $N(T)$.

D.3.4 Definition 3.4: Es sei $T: E \rightarrow E$ ein linearer Operator, der E in sich abbildet. Man definiert (rekursiv) die **Potenzen** T^n von T für $n = 0, 1, 2, \dots$ durch die folgenden Gleichungen

$$\begin{aligned} T^0 &= I, \\ T^1 &= T, \\ T^{n+1} &= TT^n \quad (n = 1, 2, \dots), \end{aligned} \quad (3.12)$$

wobei in der Gl. (3.12) rechts das Produkt der Operatoren T und T^n im Sinne der Def. 3.3 steht.

Der aus der linearen Algebra bekannte Begriff der inversen Matrix wird durch die folgende Definition auf allgemeine lineare Operatoren übertragen.

D.3.5 Definition 3.5: Es sei $T: E \rightarrow F$ ein linearer Operator. Falls es einen linearen Operator $S: F \rightarrow E$ gibt, so daß folgende Gleichungen gelten

$$ST = I_E, \quad TS = I_F, \quad (3.13)$$

wobei I_E bzw. I_F die identischen Abbildungen von E bzw. F sind, so heißt S der zu T **inverse Operator (Umkehroperator, reziproker Operator)**, und man bezeichnet ihn mit

$$S = T^{-1}. \quad (3.14)$$

Bemerkung 3.1: Man weist leicht nach, daß es nur einen einzigen solchen Operator $S = T^{-1}$ geben kann. Hat T einen Umkehroperator T^{-1} , so ist T eine lineare eindeutige Abbildung von E auf F und T^{-1} eine lineare eindeutige Abbildung von F auf E (vgl. auch Bem. 3.3).

S.3.3 Satz 3.3: Es seien $T_1: E \rightarrow F$, $T_2: F \rightarrow G$ lineare Operatoren, die Umkehroperatoren T_1^{-1} , T_2^{-1} besitzen. Dann hat auch der Operator $T_2 T_1: E \rightarrow G$ einen Umkehroperator, und es gilt

$$(T_2 T_1)^{-1} = T_1^{-1} T_2^{-1}. \quad (3.15)$$

(Beweis der Sätze 3.1–3.3 als Übungsaufgabe.)

3.2. Beschränkte lineare Operatoren in Banachräumen

Für die physikalischen Anwendungen ist es vor allem erforderlich, lineare Operatoren in *Hilberträumen* zu betrachten. Zur Klärung einiger allgemeiner Begriffe ist es jedoch günstiger, zunächst lineare Operatoren in *Banachräumen* zu behandeln, soweit es die Theorie beschränkter Operatoren betrifft. Unbeschränkte lineare Operatoren werden wir von vornherein nur in Hilberträumen untersuchen (s. 3.3. und 5.). Im folgenden bezeichnet $(E, \|\cdot\|)$ einen komplexen Banachraum.

D.3.6 Definition 3.6: Es sei $T: E \rightarrow E$ eine lineare Abbildung von E in sich. T heißt **beschränkt**, wenn es eine Konstante $K > 0$ gibt mit

$$\|Tx\| \leq K\|x\| \quad (x \in E). \quad (3.16)$$

Das Infimum dieser Werte K , für die die Ungleichung (3.16) gilt, wird mit $\|T\|$ bezeichnet.

S.3.4 Satz 3.4: Eine lineare Abbildung (ein Operator) T ist genau dann stetig (d. h., aus $x_n \rightarrow x$ folgt stets $Tx_n \rightarrow Tx$), wenn T beschränkt ist.

Bemerkung 3.2: Ein völlig analoger Satz gilt für lineare Abbildungen eines normierten Raumes in einen anderen normierten Raum, wenn die Definition der Beschränktheit analog zur Gl. (3.16) getroffen wird; vgl. die entsprechende Aussage für lineare Funktionale (2.3.1.).

Beispiel 3.1: Es sei $C[a, b]$ (s. Bsp. 1.5) versehen mit der Maximum-Norm $\|x\| = \max_{a \leq t \leq b} |x(t)|$. Mit $K(s, t)$ werde eine für $a \leq s, t \leq b$ definierte stetige komplexwertige Funktion bezeichnet. Dann wird durch die Zuordnungsvorschrift [s. (3.3)]

$$(Tx)(s) = y(s) = \int_a^b K(s, t) x(t) dt \quad (a \leq s \leq b) \quad (3.17)$$

eine beschränkte lineare Abbildung von $C[a, b]$ in $C[a, b]$ definiert; T ist ein sog. **linearer Integraloperator** mit dem **Kern** $K(s, t)$. Das Wort „Kern“ wird hier in anderem Sinn als in (3.6), S.73, gebraucht. Als Übung zeige der Leser, daß durch die Vorschrift (3.17) tatsächlich eine Abbildung von $C[a, b]$ in $C[a, b]$ definiert ist. Die Beschränktheit von T erkennt man so:

$$\begin{aligned} |(Tx)(s)| &= |y(s)| = \left| \int_a^b K(s, t) x(t) dt \right| \leq \int_a^b |K(s, t)| |x(t)| dt \\ &\leq \int_a^b |K(s, t)| \max_{a \leq t \leq b} |x(t)| dt = \|x\| \int_a^b |K(s, t)| dt \quad (a \leq s \leq b). \end{aligned}$$

Durch Übergang zum Maximum folgt

$$\|Tx\| = \max_{a \leq s \leq b} |(Tx)(s)| \leq \|x\| \max_{a \leq s \leq b} \int_a^b |K(s, t)| dt \quad (x \in C[a, b]).$$

Also ist T beschränkt, denn es gilt für alle $x \in C[a, b]$ eine Ungleichung $\|Tx\| \leq M\|x\|$ mit

$$M = \max_{a \leq s \leq b} \int_a^b |K(s, t)| dt. \text{ Eine genauere Betrachtung zeigt, daß sogar } M = \|T\| \text{ gilt.}$$

3.2.1. Spektrum und Resolvente

Wesentliche Eigenschaften eines linearen Operators T treten erst zutage, wenn man ihn mit anderen linearen Operatoren vergleicht; speziell hat sich der Vergleich mit Vielfachen λI des identischen Operators als entscheidend erwiesen. I ist die durch

$$Ix = x \quad (x \in E) \quad (3.18)$$

erklärte (lineare, stetige) identische Abbildung [s. (3.8)]. Man untersucht hierzu die Menge (Schar, Familie) von Operatoren

$$\lambda I - T \quad (\lambda \text{ komplex}), \quad (3.19)$$

die für jedes fest gewählte komplexe λ einen linearen (stetigen, falls T stetig ist) Operator auf E (von E in sich) liefert.

Definition 3.7: Es sei T ein linearer stetiger Operator auf E mit Werten in E . Die Menge $\rho(T)$ aller komplexen Zahlen λ , für die die Abbildung $\lambda I - T$ surjektiv (d. h. eine eindeutige Abbildung von E auf sich: $T[E] = E$) ist und für die die Umkehrabbildung (die Reziproke)

$$R(\lambda; T) = R_\lambda(T) = (\lambda I - T)^{-1} \quad (3.20)$$

ein stetiger (linearer) Operator ist, heißt die **Resolventenmenge** $\rho(T)$ von T . Der inverse Operator $R_\lambda(T) = (\lambda I - T)^{-1}$ des Operators $\lambda I - T$ heißt die **Resolvente** von T im Punkt λ .

Bemerkung 3.3: Ist $T: E_1 \rightarrow E_1$ eine stetige lineare Abbildung, die den normierten Raum E_1 eindeutig (injektiv) auf sich abbildet, so existiert eine Umkehrabbildung $S: E_1 \rightarrow E_1$ der Abbildung T , d. h., die Gleichung $Tx = y$ ist äquivalent zur Gleichung $x = Sy$ für $x \in E_1, y \in E_1$. Der Operator S ist notwendig linear, aber nicht notwendig stetig. Ist aber zusätzlich E_1 ein Banachraum, dann muß S notwendig auch stetig sein (*Satz von Banach* über den inversen Operator). Eine komplexe Zahl λ gehört daher genau dann zur Resolventenmenge eines auf einem Banachraum definierten linearen stetigen Operators T , wenn die Abbildung $\lambda I - T$ eine injektive Abbildung von E auf sich ist. Die Gleichung

$$\lambda x - Tx = y \quad (3.21)$$

ist dann für jedes $y \in E$ lösbar mit genau einem $x \in E$, und dieses x hängt stetig von y ab (wenn y als variabel betrachtet wird).

D.3.8 Definition 3.8: Es sei $T: E \rightarrow E$ ein linearer, stetiger Operator. Die Komplementärmenge der Resolventenmenge $\varrho(T)$ (bezüglich \mathbf{K}), die Menge

$$\sigma(T) = \mathbf{K} \setminus \varrho(T) \quad (3.22)$$

heißt das **Spektrum** von T . Die Zahl

$$r(T) = \sup_{\lambda \in \sigma(T)} |\lambda| \quad (3.23)$$

heißt der **Spektralradius** von T .

S.3.5 Satz 3.5: Die Resolventenmenge $\varrho(T)$ ist eine (in \mathbf{K}) offene Menge, das Spektrum $\sigma(T)$ ist eine nichtleere, abgeschlossene und beschränkte Menge. Es gilt die Gleichung

$$r(T) = \lim_{n \rightarrow \infty} \sqrt[n]{\|T^n\|} = \inf_{n \in \mathbf{N}} \sqrt[n]{\|T^n\|} \quad (3.24)$$

sowie die Ungleichung

$$r(T) \leq \|T\|. \quad (3.25)$$

S.3.6 Satz 3.6: Für Zahlen μ, λ aus der Resolventenmenge gilt die sogenannte **1. Resolventengleichung**:

$$R(\lambda; T) - R(\mu; T) = (\mu - \lambda) R(\lambda; T) R(\mu; T). \quad (3.26)$$

Beweis: Mit $R(\lambda; T) = (\lambda I - T)^{-1}$, $R(\mu; T) = (\mu I - T)^{-1}$ gilt wegen der Vertauschbarkeit von $\lambda I - T$ und $\mu I - T$ die Gleichheit $(\mu I - T)(\lambda I - T)^{-1} [R(\lambda; T) - R(\mu; T)] = (\mu I - T) - (\lambda I - T) = (\mu - \lambda) I$. Multipliziert man beide Seiten letzterer Gleichung mit $R(\lambda; T) R(\mu; T)$, so erhält man unmittelbar die Gl. (3.26). ■

Mittels eines gegebenen stetigen linearen Operators $T: E \rightarrow E$, der den vorliegenden Banachraum in sich abbildet, kann man durch Linearkombination der Potenzen von T [s. (3.12)] mit komplexen Konstanten $\alpha_0, \alpha_1, \dots, \alpha_k, \dots$

$$\text{Operatorpolynome: } \sum_{k=0}^n \alpha_k T^k = \alpha_0 I + \alpha_1 T + \alpha_2 T^2 + \dots + \alpha_n T^n \quad (3.27)$$

bzw.

$$\text{Operatorpotenzreihen: } \sum_{k=0}^{\infty} \alpha_k T^k \quad (3.28)$$

gewinnen, wobei im letzteren Fall die Konvergenz geeignet erklärt werden muß. Es sind diese Ausdrücke Spezialfälle der allgemeineren **Operatorfunktionen**, auf deren nähere Beschreibung wir hier verzichten (s. aber 5.1.4.). Von Interesse in diesem Zusammenhang ist vor allem der sog. **Spektralabbildungssatz**, der aussagt, daß für eine komplexe Funktion $f(z)$, die in einer offenen Obermenge des Spektrums $\sigma(T)$ holomorph ist, die folgende Gleichung gilt:

$$\sigma(f(T)) = f(\sigma(T)). \quad (3.29)$$

Beispiel 3.2: Es sei $E = K^n$ der n -dimensionale komplexe Vektorraum mit der üblichen Norm und $T: K^n \rightarrow K^n$ eine lineare Abbildung. Dann ist T stetig (Beweis als Übungsaufgabe). Die Abbildung T läßt sich als **Matrix** darstellen: $T = (t_{ij})_{1 \leq i, j \leq n}$; ist $x \in E$ gegeben, $x = (\xi_1, \dots, \xi_n)$, so ist $Tx = (\eta_1, \dots, \eta_n)$ mit

$$\eta_k = \sum_{j=1}^n t_{kj} \xi_j \quad (k = 1, \dots, n). \quad (3.30)$$

Der Abbildung $\lambda I - T$ (komplex) entspricht die Matrix

$$(\lambda \delta_{ij} - t_{ij})_{1 \leq i, j \leq n} \quad \left(\delta_{ij} = \begin{cases} 0 & \text{für } i \neq j \\ 1 & \text{für } i = j \end{cases} \right)$$

und diese Abbildung besitzt, wie aus der linearen Algebra bekannt, genau dann eine Inverse $(\lambda I - T)^{-1}$ (diese ist dann automatisch stetig), wenn die Determinante der entsprechenden Matrix von null verschieden ist. Also gilt die Beziehung

$$\lambda \in \rho(T) \Leftrightarrow \det(\lambda \delta_{ij} - t_{ij}) \neq 0,$$

und deshalb ist

$$\lambda \in \sigma(T) \Leftrightarrow \det(\lambda \delta_{ij} - t_{ij}) = 0 \Leftrightarrow \det(t_{ij} - \lambda \delta_{ij}) = 0.$$

Eine komplexe Zahl gehört also genau dann zum Spektrum $\sigma(T)$, wenn die Gleichung

$$\begin{vmatrix} t_{11} - \lambda & t_{12} & \dots & t_{1n} \\ t_{21} & t_{22} - \lambda & \dots & t_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ t_{n1} & t_{n2} & \dots & t_{nn} - \lambda \end{vmatrix} = 0 \quad (3.31)$$

gilt. (3.31) ist aber die bekannte Gleichung zur Bestimmung der Eigenwerte λ_k der Matrix (t_{ij}) ; d. h., die Lösungen $\lambda_1, \lambda_2, \dots$ der Gl. (3.31) stellen die Eigenwerte von (t_{ij}) dar. Mit Berücksichtigung ihrer Vielfachheit hat die Gl. (3.31) n Lösungen $\lambda_1, \dots, \lambda_n$. Im vorliegenden Fall gilt somit

$$\sigma(T) = \{\lambda_1, \dots, \lambda_n\}. \quad (3.32)$$

Nicht in jedem Fall läßt sich das Spektrum eines linearen stetigen Operators in so einfacher Weise kennzeichnen wie im Bsp. 3.2, sondern $\sigma(T)$ zerfällt in mehrere qualitativ verschiedene Teile.

Definition 3.9: Es sei $(E, \|\cdot\|)$ ein Banachraum, $T: E \rightarrow E$ eine lineare beschränkte Abbildung. Eine komplexe Zahl λ heißt **Eigenwert** von T , wenn es ein Element $x \neq 0$ in E gibt mit **D.3.9**

$$Tx = \lambda x. \quad (3.33)$$

Das Element x heißt ein zugehöriger **Eigenvektor** zum Eigenwert λ .

Definition 3.10: Es sei $(E, \|\cdot\|)$ ein Banachraum; $T: E \rightarrow E$ ein linearer beschränkter Operator. Die Menge aller Eigenwerte von T bezeichnet man als das **Punktspektrum** **D.3.10** $\sigma_p(T)$. Die Menge aller $\lambda \in \sigma(T)$, die nicht in $\sigma_p(T)$ liegen und für die der Wertebereich

von $\lambda I - T$ in E dicht liegt, $\overline{(\lambda I - T)[E]} = E$, bildet das **kontinuierliche Spektrum** oder **Stetigkeitsspektrum** $\sigma_c(T)$ von T . Alle übrigen Punkte des Spektrums werden zum **Restspektrum** (auch: **Residualspektrum**) $\sigma_r(T) = \sigma(T) \setminus (\sigma_p(T) \cup \sigma_c(T))$ zusammengefaßt.

Bemerkung 3.4: Es gilt die Gleichung $\sigma(T) = \sigma_p(T) \cup \sigma_c(T) \cup \sigma_r(T)$, und die einzelnen Bestandteile dieser Vereinigungsmenge sind paarweise elementfremd. Für die Punkte λ , die zu $\sigma_c(T)$ oder zu $\sigma_r(T)$ gehören, ist die Abbildung $(\lambda I - T)$ zwar ein-eindeutig, aber die Umkehrabbildung $(\lambda I - T)^{-1}$ läßt sich nicht auf ganz E definie-ren, da in diesem Fall der Wertebereich von $(\lambda I - T)$ eine *echte* Teilmenge von E ist.

Beispiel 3.3: Es sei $E = C[0, 1]$; $\|x\| = \max_{t \in [0, 1]} |x(t)|$ ($x \in E$). Wir definieren den linearen Opera-tor $T: E \rightarrow E$ durch die Gleichung

$$(Tx)(t) = x(0) + \int_0^t x(s) ds \quad (0 \leq t \leq 1; x \in E). \quad (3.34)$$

(Man erhält T durch formale Integration (vgl. Bd. 7/1, 2.5.1.) der Differentialgleichung $x' = x$; s. auch [4, Kap. 3, § 2.2.], dort wird das Spektrum dieser Differentialgleichung nach einer Störung benötigt (vgl. 3.3.3. unten). Ein schwierigeres Beispiel zur Ermittlung des Spektrums geht über unseren Band hinaus; s. aber auch Kap. 5.)

Durch die Gl. (3.34) wird jeder auf dem Intervall $[0, 1]$ stetigen Funktion $x(\cdot)$ eine Funktion $Tx(\cdot)$ zugeordnet. (Die folgenden Tatsachen beweise der Leser als Übung.) Diese Funktion ist wieder stetig, d. h., T bildet den Raum E wieder in den Raum E ab. T ist linear und beschränkt; $\|T\| = 2$.

Zur Ermittlung des Spektrums betrachten wir den Operator $\lambda I - T$ für beliebiges komplexes λ . Es gilt für $x \in E$

$$(\lambda I - T)(x)(t) = \lambda x(t) - x(0) - \int_0^t x(s) ds \quad (0 \leq t \leq 1).$$

Wir geben ein $y \in E$ vor und untersuchen die Gleichung (mit der Unbekannten x):

$$(\lambda I - T)x = y. \quad (3.35)$$

Ausgeschrieben lautet diese Beziehung

$$\lambda x(t) - x(0) - \int_0^t x(s) ds = y(t) \quad (0 \leq t \leq 1). \quad (3.36)$$

Für $t = 0$ folgt aus dieser Gleichung

$$\lambda x(0) - x(0) = y(0), \quad \text{d. h.} \quad (\lambda - 1)x(0) = y(0). \quad (3.37)$$

Wir setzen zunächst voraus, daß $y(t)$ eine differenzierbare (nicht nur stetige) Funktion ist. Unter der Annahme, daß die Gl. (3.35) eine Lösung hat, gilt für $\lambda \neq 0$

$$x(t) = \frac{1}{\lambda} \left(x(0) + \int_0^t x(s) ds + y(t) \right) \quad (0 \leq t \leq 1).$$

Auf der rechten Seite stehen nur differenzierbare Funktionen, also ist auch $x(t)$ auf der linken Seite eine differenzierbare Funktion. Durch Differentiation beider Seiten von (3.36) erhalten wir mit

$$\lambda x'(t) - x(t) = y'(t) \quad (3.38)$$

eine gewöhnliche lineare Differentialgleichung erster Ordnung für die gesuchte Funktion $x(t)$, in der vor der höchsten Ableitung der Parameter λ auftritt! Mittels Standardmethoden (s. Bd. 7/1)

erhalten wir die allgemeine Lösung der Differentialgleichung (3.38):

$$x(t) = K e^{\frac{1}{\lambda} t} + \frac{1}{\lambda} \int_0^t y'(s) e^{-\frac{1}{\lambda}(s-t)} ds \quad (0 \leq t \leq 1), \quad K = x(0).$$

Ist $\lambda \neq 1$, so können wir aus (3.37) den Wert $K = x(0)$ bestimmen:

$$K = x(0) = \frac{y(0)}{\lambda - 1}$$

und erhalten

$$x(t) = \frac{y(0) e^{\frac{1}{\lambda} t}}{\lambda - 1} + \frac{1}{\lambda} \int_0^t y'(s) e^{-\frac{1}{\lambda}(s-t)} ds,$$

und aus dieser Formel mittels partieller Integration und Zwischenrechnung

$$x(t) = y(0) \frac{e^{\frac{1}{\lambda} t}}{\lambda(\lambda - 1)} + \frac{1}{\lambda} y(t) + \frac{1}{\lambda^2} \int_0^t y(s) e^{-\frac{1}{\lambda}(s-t)} ds \quad (0 \leq t \leq 1). \quad (3.39)$$

Wie man durch Einsetzen in die Gleichung (nach einer Zwischenrechnung) erkennt, stellt die rechte Seite der Gl. (3.39) für $\lambda \neq 0, \lambda \neq 1$ die Auflösung der Gl. (3.35) nach der Funktion $x(\cdot)$ dar. Man stellt dabei fest, daß die Voraussetzung, daß $y(t)$ differenzierbar sein soll, nirgends benötigt wird. Die Formel (3.39) stellt somit eine Lösung der Gl. (3.35) für eine beliebige (stetige) rechte Seite $y(t)$ aus E dar. Wie wir weiter unten nachweisen werden, ist es auch die einzige Lösung von (3.35). Da die Zuordnung $y(\cdot) \rightarrow$ „Ausdruck auf der rechten Seite von (3.39)“ ersichtlich einen linearen stetigen Operator in E definiert, der E in E abbildet, stellt die rechte Seite von (3.39) die Resolvente $R(\lambda; T)$ von T dar, d. h., für $\lambda \neq 0, \lambda \neq 1$ gilt $(0 \leq t \leq 1; y \in E)$:

$$\begin{aligned} R(\lambda; T)(y)(t) &= (\lambda I - T)^{-1}(y)(t) \\ &= y(0) \frac{e^{\frac{1}{\lambda} t}}{\lambda(\lambda - 1)} + \frac{1}{\lambda} y(t) + \frac{1}{\lambda^2} \int_0^t y(s) e^{-\frac{1}{\lambda}(s-t)} ds. \end{aligned} \quad (3.40)$$

Nachzutragen ist noch der Beweis dafür, daß (3.39) die einzige Lösung von (3.35) darstellt. Angenommen, es gäbe zwei Lösungen $x_1 \neq x_2$ der Gl. (3.35). Dann gilt $(0 \leq t \leq 1)$:

$$y(t) = \lambda x_1(t) - x_1(0) - \int_0^t x_1(s) ds = \lambda x_2(t) - x_2(0) - \int_0^t x_2(s) ds.$$

Für $t = 0$ folgt speziell

$$(\lambda - 1) x_1(0) = (\lambda - 1) x_2(0),$$

also $(\lambda \neq 1)$: $x_1(0) = x_2(0)$ und damit

$$\lambda(x_1(t) - x_2(t)) = \int_0^t (x_1(s) - x_2(s)) ds \quad (0 \leq t \leq 1).$$

Die rechte Seite letzterer Gleichung ist differenzierbar, also ist es auch die linke Seite, und wir erhalten durch Differentiation

$$\lambda \frac{d}{dt} (x_1(t) - x_2(t)) = x_1(t) - x_2(t)$$

bzw. mit der Abkürzung $x_1(t) - x_2(t) = u(t)$ $(0 \leq t \leq 1)$

$$\frac{du}{dt} = \frac{1}{\lambda} u.$$

Diese Differentialgleichung für u hat die allgemeine Lösung (s. Bd. 7/1) $u(t) = K e^{\frac{1}{\lambda} t}$ mit einer Konstanten K . Es gilt einerseits (s. o.) $u(0) = x_1(0) - x_2(0) = 0$ und andererseits $u(0) = K$. Also gilt $K = 0$ und damit ist $u(t) = 0$ ($0 \leq t \leq 1$), d. h. aber

$$x_1(t) = x_2(t) \quad (0 \leq t \leq 1).$$

Also stimmen je zwei Lösungen von (3.35) [bei gegebenem $y(\cdot)$] überein, und (3.39) liefert für $\lambda \neq 0$, $\lambda \neq 1$ die einzige Lösung von (3.35). Alle $\lambda \neq 0, 1$ gehören somit zu $\varrho(T)$.

Zur Bestimmung der einzelnen Bestandteile des Spektrums untersuchen wir zunächst die Frage nach den Eigenwerten von T . Es gelte also die Gleichung $Tx = \lambda x$ mit $x \neq 0$. Ausgeschrieben lautet diese Gleichung

$$x(0) + \int_0^t x(s) ds = \lambda x(t). \quad (3.41)$$

Für $\lambda = 0$ erhält man, wie durch Differentiation sofort ersichtlich ist, die Gleichung $x(t) = 0$ ($0 \leq t \leq 1$); also ist $\lambda = 0$ kein Eigenwert. Nach den obigen Betrachtungen ist nur noch der Wert

$\lambda = 1$ zu untersuchen. Die Gleichung $x(0) + \int_0^t x(s) ds = x(t)$ liefert nach Differentiation (erlaubt, weil die linke Seite differenzierbar ist) $x(t) = x'(t)$, woraus $x(t) = C e^t$ ($0 \leq t \leq 1$) folgt. Ist $C \neq 0$, so liegt, wie die Probe zeigt, tatsächlich ein Eigenvektor $x(\cdot)$ zum Eigenwert 1 vor. Somit gilt $\sigma_p(T) = \{1\}$. Die Untersuchung der Gleichung $(\lambda I - T)x = y$ führt für $\lambda = 0$ sofort auf die Beziehung

$$-x(0) - \int_0^t x(s) ds = y(t) \quad (0 \leq t \leq 1). \quad (3.42)$$

Auf der linken Seite steht eine differenzierbare Funktion von t [wenn wir voraussetzen, daß die Gl. (3.42) eine Lösung hat]. Also muß auch die Funktion $y(t)$ auf der rechten Seite von (3.42) differenzierbar sein ($0 \leq t \leq 1$). Da es aber stetige Funktionen $y(t)$ gibt, die nicht in jedem Punkt des Intervalls $[0, 1]$ differenzierbar sind (z. B. $y(t) = |t - \frac{1}{2}|$), ist die Gl. (3.42) nicht für jedes stetige $y(t)$ nach $x(t)$ auflösbar. Also gehört $\lambda = 0$ nicht zur Resolventenmenge von T . Eine genauere Untersuchung zeigt, daß für $\lambda = 0$ $\overline{(\lambda I - T[E])} = \overline{(-T[E])} = E$ gilt.

Somit gehört der Wert $\lambda = 0$ zum kontinuierlichen Spektrum von T . Zusammenfassend haben wir folgendes Ergebnis:

$$\begin{aligned} \varrho(T) &= \mathbb{K} \setminus \{0, 1\}; & \sigma(T) &= \{0, 1\}; & \sigma_p(T) &= \{1\}, & \sigma_c(T) &= \{0\}, & \sigma_r(T) &= \emptyset, \\ r(T) &= \sup_{\lambda \in \varrho(T)} |\lambda| = 1; & \|T\| &= 2. \end{aligned}$$

Im folgenden gehen wir zu einer speziellen Darstellung der Resolvente über, die vor allem für die Lösung von Integralgleichungen von Interesse ist.

S.3.7 Satz 3.7: Die Reihe $\sum_{n=0}^{\infty} \frac{1}{\lambda^n} T^n$ sei konvergent im Sinne der Operatornorm, das heißt, es existiert ein linearer stetiger Operator S auf E mit:

$$\lim_{k \rightarrow \infty} \left\| \sum_{n=0}^k \frac{1}{\lambda^n} T^n - S \right\| = 0.$$

Dann gilt die Gleichung

$$S = \sum_{n=0}^{\infty} \frac{1}{\lambda^n} T^n = \left(I - \frac{1}{\lambda} T \right)^{-1} \quad (\lambda \neq 0). \quad (3.43)$$

Beispiel 4.3: Die Entwicklung eines Polynoms nach Legendreschen Polynomen.

Wir wollen eine Formel für die Entwicklung von t^n nach Legendreschen Polynomen angeben und setzen an:

$$t^n = \sum_{k=0}^n a_k P_k(t). \quad (4.32)$$

Dann ergibt sich wegen (4.29)

$$a_k = \frac{2k+1}{2} \int_{-1}^{+1} t^n P_k(t) dt \quad \text{für } k = 0, 1, 2, \dots, n.$$

Eine kurze Rechnung ergibt

$$t^n = \frac{n!}{1 \cdot 3 \cdot 5 \dots (2n-1)} [P_n(t) + (2n-3) \frac{1}{2} P_{n-2}(t) + (2n-7) \frac{2n-1}{2 \cdot 4} P_{n-4}(t) + \dots]. \quad (4.33)$$

Diese Formel hat bereits Legendre angegeben.

Kurvenbild

Aus der Formel (4.31) können wir mit dem Satz von Rolle [Band 2] auf die Lage der Nullstellen von $P_n(t)$ schließen. Das Polynom $(t^2 - 1)^n = (t - 1)^n (t + 1)^n$ hat je eine n -fache Nullstelle bei $t = \pm 1$. Mithin hat seine 1. Ableitung je eine $(n-1)$ -fache bei ± 1 und eine einfache dazwischen. Die 2. Ableitung besitzt je eine $(n-2)$ -fache Nullstelle bei ± 1 und zwei einfache dazwischen. Schließlich besitzt die n -te Ableitung – also $P_n(t)$ – genau n einfache reelle Nullstellen zwischen -1 und $+1$.

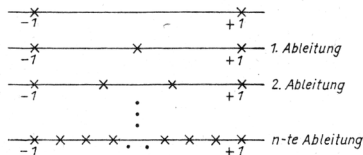


Bild 4.3. Lage der Nullstellen von $\frac{d^n}{dt^n} (t^2 - 1)^n$

Die Werte der Nullstellen (auf 6 Dezimalstellen genau) ergeben sich im Intervall $[0, 1]$ für die ersten 6 Polynome $P_n(t)$:

$$n = 1: t_1 = 0$$

$$n = 2: t_1 = 0,577350$$

$$n = 3: t_1 = 0 \quad t_2 = 0,774597$$

$$n = 4: t_1 = 0,339981 \quad t_2 = 0,861136$$

$$n = 5: t_1 = 0 \quad t_2 = 0,538469 \quad t_3 = 0,906180$$

$$n = 6: t_1 = 0,238619 \quad t_2 = 0,661209 \quad t_3 = 0,932470.$$

Die Kurven der Polynome $P_n(t)$ haben folgenden Verlauf:

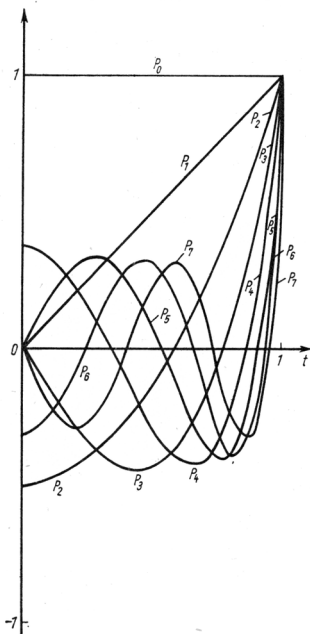


Bild 4.4. Kurvenverlauf der Legendreschen Polynome

4.2.3. Integraldarstellungen

Weitere Darstellungen der Legendreschen Kugelfunktionen erhält man – analog zum Vorgehen in Kapitel 3 –, wenn die Betrachtungen zu den Funktionen auf komplexe Veränderliche ausgedehnt und Sätze der Funktionentheorie genutzt werden. Insbesondere wird t im folgenden als komplexe Veränderliche aufgefaßt.

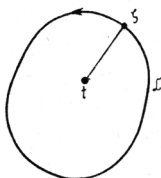
So erhalten wir durch Anwendung der Cauchyschen Integralformel [Band 9] auf die Rodriguessche Formel

$$P_n(t) = \frac{1}{2^n n!} \frac{d^n}{dt^n} [t^2 - 1]^n \quad (4.34)$$

eine erste Integraldarstellung für die Funktionen $P_n(t)$, t komplex:

$$P_n(t) = \frac{1}{2\pi i} \int_{\mathbb{C}} \frac{(\zeta^2 - 1)^n}{2^n (\zeta - t)^{n+1}} d\zeta = \frac{1}{2^{n+1} \pi i} \int_{\mathbb{C}} \frac{(\zeta^2 - 1)^n}{(\zeta - t)^{n+1}} d\zeta, \quad (4.35)$$

wobei der Integrationsweg \mathcal{C} in der komplexen ζ -Ebene den Punkt $\zeta = t$ einmal im positiven Sinne umläuft. Dieser Ausdruck stammt von Schläfli (1881).

Bild 4.5. Integrationsweg \mathcal{C}

Ebenso ergibt sich durch Anwendung der Cauchyschen Integralformel auf die Potenzreihenentwicklung

$$\frac{1}{\varrho} = \frac{1}{\sqrt{1-2zt+z^2}} = \sum_{n=0}^{\infty} P_n(t) z^n$$

eine weitere Integraldarstellung

$$P_n(t) = \frac{1}{2\pi i} \int_{\mathcal{C}} \frac{d\zeta}{\zeta^{n+1} \sqrt{1-2t\zeta+\zeta^2}}, \quad (4.36)$$

wobei \mathcal{C} eine Kurve ist, die den Nullpunkt der komplexen Zahlenebene einmal im positiven Sinne umläuft, ohne einen singulären Punkt von $(1-2t\zeta+\zeta^2)^{-\frac{1}{2}}$ zu umschließen.

Aufgabe 4.6: Man leite (4.35) aus (4.36) durch die Substitution $\sqrt{1-2t\zeta+\zeta^2} = \zeta z - 1$ für die neue Variable z her.

Wie betrachten wieder die Darstellung (4.35) und wählen \mathcal{C} als Kreis mit dem Mittelpunkt t und dem Radius $|t^2-1|^{\frac{1}{2}}$ ($t \neq \pm 1$), so daß längs \mathcal{C} gesetzt werden kann:

$$\zeta = t + \sqrt{t^2-1} e^{i\varphi}, \quad -\pi \leq \varphi \leq +\pi.$$

Die Wahl des Zweiges von $\sqrt{t^2-1}$ ist für die weiteren Betrachtungen ohne Bedeutung. Nach der Substitution erhalten wir für alle Werte von $t \neq \pm 1$

$$\begin{aligned} P_n(t) &= \frac{1}{2^{n+1}\pi i} \int_{-\pi}^{+\pi} \left(\frac{t^2-1+2t\sqrt{t^2-1}e^{i\varphi}+(t^2-1)e^{2i\varphi}}{\sqrt{t^2-1}e^{i\varphi}} \right)^n i d\varphi \\ &= \frac{1}{2^{n+1}\pi} \int_{-\pi}^{+\pi} (\sqrt{t^2-1}e^{-i\varphi}+2t+\sqrt{t^2-1}e^{i\varphi})^n d\varphi \\ &= \frac{1}{2\pi} \int_{-\pi}^{+\pi} (t+\sqrt{t^2-1}\cos\varphi)^n d\varphi. \end{aligned}$$

D.3.13 Definition 3.13: Es seien E, F normierte Räume; E', F' ihre Dualräume (s. 2.3.2.). Ist $T: E \rightarrow F$ eine lineare stetige Abbildung, so wird die Abbildung $T': F' \rightarrow E'$ durch die Gleichung

$$T'f = fT \quad (f \in F') \quad (3.49a)$$

gegeben, d. h., für jedes lineare stetige Funktional f auf F wird (ein lineares stetiges) Funktional $T'f$ auf E erklärt mittels der Beziehung

$$(T'f)(x) = f(Tx) \quad (x \in E). \quad (3.49b)$$

Die Zuordnung $T': f \rightarrow T'f$ heißt die zu T **adjungierte** oder **duale Abbildung** (T' ist der zu T **duale Operator**).

Bemerkung 3.6: Die Tatsache, daß $T'f$ zu E' gehört, ergibt sich sofort aus der Definitionsgleichung (3.49a): $T'f$ ist die Zusammensetzung (Hintereinanderausführung) zweier stetiger linearer Abbildungen und daher selbst stetig und linear.

S.3.14 Satz 3.14: Ist $T \in L(E, F)$, so ist $T' \in L(F', E')$, und es gilt

$$\|T'\| = \|T\|. \quad (3.50)$$

Beispiel 3.6: Es seien $E = R^n$, $F = R^m$ (mit der euklidischen Norm), $T \in L(E, F)$ eine gegebene lineare Abbildung. Bezüglich (gegebener) Basen in E bzw. F wird T dargestellt durch eine (m, n) -Matrix

$$A = [t_{jk}]_{\substack{1 \leq j \leq m; \\ 1 \leq k \leq n}}, \quad (3.51)$$

Die zu T duale Abbildung T' ist dann durch die (n, m) -Matrix

$$A' = [t'_{jk}]_{\substack{1 \leq j \leq n \\ 1 \leq k \leq m}} \quad \text{mit} \quad t'_{jk} = t_{kj}$$

gegeben, die $F' = F = R^m$ in $E' = E = R^n$ abbildet.

A' entsteht also aus A durch Vertauschung von Zeilen und Spalten, ist also die zu A **transponierte** Matrix (häufig mit A^T bezeichnet).

Beispiel 3.7: Es sei $K(s, t)$ ($a \leq t \leq b$; $c \leq s \leq d$) eine reellwertige stetige Funktion zweier Variabler. Der lineare Operator T :

$$(Tx)(s) = \int_a^b K(s, t) x(t) dt \quad (c \leq s \leq d) \quad (x \in L^2_{\mathbf{R}}[a, b])$$

bildet den Raum $L^2_{\mathbf{R}}[a, b]$ stetig in den Raum $L^2_{\mathbf{R}}[c, d]$ ab. Der zu T duale Operator T' ist dann durch die Gleichung

$$(T'x)(s) = \int_c^d K(t, s) x(t) dt \quad (a \leq s \leq b) \quad (x \in L^2_{\mathbf{R}}[c, d])$$

gegeben und bildet den Raum $F' = L^2_{\mathbf{R}}[c, d] = F$ in den Raum $E' = L^2_{\mathbf{R}}[a, b] = E$ ab. (Man beachte, daß entsprechend Bsp. 2.20 in 2.3.2. stetige lineare Funktionale auf $L^2_{\mathbf{R}}$ und Elemente von $L^2_{\mathbf{R}}$ identifiziert werden.) $K(t, s)$ heißt der zu $K(s, t)$ **transponierte Kern**.

Für den Übergang zur dualen Abbildung gelten die folgenden Rechenregeln (es gelte $T, S \in L(E, F)$):

- (1) $(T + S)' = T' + S'$;
- (2) $(\lambda T)' = \lambda T'$ (λ reell oder komplex);
- (3) $(I_E)' = I_{E'}$;
- (4) $(T^{-1})' = (T')^{-1}$;
- (5) ist $T \in L(E, F)$ und $S \in L(F, G)$, so ist $ST \in L(E, G)$, und es gilt $(ST)' = T'S'$.

längs dieser Wegeinschnitte ebenfalls gegen null gehen und somit die Deformation nach dem Cauchyschen Integralsatz erlaubt ist:

$$\begin{aligned} P_n(\cos \vartheta) &= \frac{1}{\pi} \int_{-\vartheta}^{+\vartheta} \frac{e^{i(n+1)\varphi}}{\sqrt{1 - 2e^{i\varphi} \cos \vartheta + e^{i2\varphi}}} d\varphi = \frac{1}{\pi} \int_{-\vartheta}^{+\vartheta} \frac{e^{i(n+\frac{1}{2})\varphi}}{\sqrt{2(\cos \varphi - \cos \vartheta)}} d\varphi \\ &= \frac{\sqrt{2}}{2\pi} \int_{-\vartheta}^{+\vartheta} \frac{\cos((n+\frac{1}{2})\varphi)}{\sqrt{\cos \varphi - \cos \vartheta}} d\varphi + i \frac{\sqrt{2}}{2\pi} \int_{-\vartheta}^{+\vartheta} \frac{\sin((n+\frac{1}{2})\varphi)}{\sqrt{\cos \varphi - \cos \vartheta}} d\varphi. \end{aligned}$$

Das zweite Integral verschwindet, da der Integrand ungerade in φ ist, somit wird

$$P_n(\cos \vartheta) = \frac{\sqrt{2}}{\pi} \int_0^{\vartheta} \frac{\cos(n+\frac{1}{2})\varphi}{\sqrt{\cos \varphi - \cos \vartheta}} d\varphi. \quad (4.39a)$$

Substituieren wir $\varphi = \pi - \varphi'$ und ersetzen ϑ durch $\pi - \vartheta'$, so ergibt sich nach Weglassen der Striche bei φ' und ϑ' aus (4.39a)

$$P_n(\cos \vartheta) = \frac{\sqrt{2}}{\pi} \int_0^{\pi} \frac{\sin(n+\frac{1}{2})\varphi}{\sqrt{\cos \vartheta - \cos \varphi}} d\varphi. \quad (4.39b)$$

Die letzten beiden Formeln stammen von Mehler (1872).

4.3. Zugeordnete Kugelfunktionen

Wir hatten im vorangegangenen Abschnitt für die Differentialgleichung (4.7) zunächst nur solche Lösungen gesucht, die lediglich von ϑ abhängen. Um allgemeinere Lösungen zu finden, die zudem auch von φ abhängen, machen wir den Separationsansatz $S_n(\vartheta, \varphi) = \Theta(\vartheta)\Phi(\varphi)$, wobei Θ nur von ϑ und Φ nur von φ abhängen soll. Setzen wir diesen Ansatz in (4.7) ein, so ergibt sich

$$n(n+1)\Theta\Phi + \frac{\Phi}{\sin \vartheta} \frac{\partial}{\partial \vartheta} \left(\sin \vartheta \frac{d\Theta}{d\vartheta} \right) + \frac{\Theta}{\sin^2 \vartheta} \frac{\partial^2 \Phi}{\partial \varphi^2} = 0$$

oder nach Multiplikation mit $\frac{\sin^2 \vartheta}{\Theta\Phi}$ und einer Umstellung, die die Trennung der Veränderlichen bedingt:

$$n(n+1)\sin^2 \vartheta + \frac{\sin \vartheta}{\Theta} \frac{d}{d\vartheta} \left(\sin \vartheta \frac{d\Theta}{d\vartheta} \right) = -\frac{1}{\Phi} \frac{d^2 \Phi}{d\varphi^2}.$$

Beide Seiten müssen somit konstant sein. Wir setzen die Konstante gleich m^2 . So erhält man für Φ die einfache gewöhnliche Differentialgleichung

$$\frac{d^2 \Phi}{d\varphi^2} + m^2 \Phi = 0,$$

deren allgemeine Lösung

$$\Phi_m(\varphi) = A_m \cos m\varphi + B_m \sin m\varphi \quad \text{mit konstantem } A_m, B_m \quad (4.40)$$

3.3. Lineare Operatoren in Hilberträumen

3.3.1. Grundlegende Begriffe, Sätze und Beispiele

3.3.1.1. Einführende Beispiele

Die linearen Operatoren, die mit der Diskussion physikalischer Probleme auftreten, sind häufig Operatoren, die auf einem dichten linearen Teilraum eines Hilbertraumes definiert und unbeschränkte Operatoren sind. Dabei heißt ein linearer Operator unbeschränkt, wenn er nicht beschränkt ist im Sinne der Definition 3.6, ein solcher Operator ist nicht stetig. Es sei z. B. E der lineare Raum aller stetig differenzierbaren (komplexwertigen) Funktionen, die auf dem Intervall $[0, 1]$ definiert sind. E ist ein dichter Teilraum des Hilbertraumes $L^2[0, 1]$. Auf E erklären wir den (linearen) Operator T durch die Zuordnungsvorschrift: jedem $x \in E$, d. h. jeder stetig differenzierbaren Funktion $x(t)$ ($0 \leq t \leq 1$) ordnen wir die stetige Funktion $x'(t) = \frac{dx}{dt}(t)$ zu:

$$(Tx)(t) = x'(t) \quad (0 \leq t \leq 1; x \in E). \quad (3.58)$$

Da jede stetige Funktion (hier: $x'(t)$) zu L^2 gehört, ist T eine Abbildung (ein Operator) von E in den Raum $L^2[0, 1]$,

$$T: E \rightarrow L^2[0, 1].$$

Man sagt auch: „ T ist ein linearer Operator im Hilbertraum $L^2[0, 1]$ mit $D(T) = E$ und $R(T) \subseteq L^2[0, 1]$ “.

Dieser Operator ist unbeschränkt. Das folgt z. B. daraus, daß die Elemente $x_n \in E$ mit $x_n(t) = t$ ($0 \leq t \leq 1$), $n = 1, 2, \dots$, die Norm $\|x_n\|_{L^2} = (2n + 1)^{-1/2} < 1$ ($n = 1, 2, \dots$) besitzen, für ihre T -Bilder jedoch $\|Tx_n\|_{L^2} = n(2n + 1)^{-1/2}$ gilt; für diese speziellen x_n folgt also $\|Tx_n\|_{L^2} \rightarrow +\infty$ für $n \rightarrow +\infty$, und dabei existiert keine endliche Konstante K derart, daß $\|Tx_n\|_{L^2} \leq K\|x_n\|$ gilt. Also ist T unbeschränkt.

Zum anderen betrachten wir eine stetige Funktion $\varphi(t)$ ($0 \leq t \leq 1$) und erklären den folgenden (linearen) Operator $S: L^2[0, 1] \rightarrow L^2[0, 1]$

$$(Sx)(t) = \varphi(t) x(t) \quad (0 \leq t \leq 1; x \in L^2[0, 1]). \quad (3.59)$$

S ist ein linearer Operator, der auf dem ganzen Hilbertraum $L^2[0, 1]$ erklärt ist. Es gilt für ein beliebiges $x \in L^2[0, 1]$

$$\begin{aligned} \|Sx\| &= \left\{ \int_0^1 |\varphi(t) x(t)|^2 dt \right\}^{1/2} = \left\{ \int_0^1 |\varphi(t)|^2 |x(t)|^2 dt \right\}^{1/2} \\ &\leq \left\{ \int_0^1 \left(\max_{0 \leq t \leq 1} |\varphi(t)|^2 \right) |x(t)|^2 dt \right\}^{1/2} \\ &= \left\{ \left(\max_{0 \leq t \leq 1} |\varphi(t)|^2 \right) \int_0^1 |x(t)|^2 dt \right\}^{1/2} = \left(\max_{0 \leq t \leq 1} |\varphi(t)|^2 \right)^{1/2} \left\{ \int_0^1 |x(t)|^2 dt \right\}^{1/2} \\ &= K \|x\| \end{aligned}$$

mit $K = \left(\max_{0 \leq t \leq 1} |\varphi(t)|^2 \right)^{1/2} = \max_{0 \leq t \leq 1} |\varphi(t)|$. Da $\varphi(t)$ stetig ist, gilt $0 \leq K < +\infty$.

Es gilt also $\|Sx\| \leq K\|x\|$ ($x \in L^2[0, 1]$), d. h., S ist ein beschränkter (stetiger) linearer Operator.

3.3.1.2. Die Matrixdarstellung eines linearen Operators

Es sei H ein Hilbertraum und $T: H \rightarrow H$ ein auf dem ganzen Raum definierter linearer Operator. Weiter bezeichne (e_n) ein vollständiges ONS (s. Def. 1.5) in H . Dann läßt sich jedes Bildelement Te_k nach den e_n entwickeln:

$$Te_k = \sum_{n=1}^{\infty} a_{nk} e_n \quad (3.60)$$

mit

$$a_{nk} = \langle e_n | Te_k \rangle \quad (n, k = 1, 2, \dots). \quad (3.61)$$

Ist x eine endliche Linearkombination von e_1, \dots, e_m ,

$$x = \sum_{j=1}^m \xi_j e_j \quad (\xi_j \text{ komplexe Zahlen}),$$

so gilt zufolge der Linearität von T und wegen (3.61) die Gleichung

$$\begin{aligned} Tx &= T\left(\sum_{j=1}^m \xi_j e_j\right) = \sum_{j=1}^m \xi_j Te_j = \sum_{j=1}^m \xi_j \sum_{n=1}^{\infty} a_{nj} e_n \\ &= \sum_{j=1}^m \sum_{n=1}^{\infty} a_{nj} \xi_j e_n = \sum_{n=1}^{\infty} \sum_{j=1}^m a_{nj} \xi_j e_n = \sum_{n=1}^{\infty} \left(\sum_{j=1}^m a_{nj} \xi_j\right) e_n. \end{aligned}$$

Bezeichnen wir das Element Tx mit y und stellen y durch seine Fourierreihe dar,

$$Tx = y = \sum_{n=1}^{\infty} \eta_n e_n,$$

so gilt daher

$$\eta_n = \sum_{j=1}^m a_{nj} \xi_j \quad (n = 1, 2, \dots)$$

oder

$$\eta_n = \sum_{j=1}^m \langle e_n | Te_j \rangle \xi_j.$$

Beachtet man noch, daß $\xi_j = \langle e_j | x \rangle$ und $\eta_n = \langle e_n | y \rangle = \langle e_n | Tx \rangle$ gilt, so erhält man die Beziehung

$$\langle e_n | Tx \rangle = \sum_{j=1}^m \langle e_n | Te_j \rangle \langle e_j | x \rangle \quad (n = 1, 2, \dots). \quad (3.62)$$

Ist der Operator T zusätzlich *beschränkt*, so gilt, wie man mittels eines Grenzübergangs zeigen kann, für jedes $x \in H$ die Gleichung

$$\langle e_n | Tx \rangle = \sum_{j=1}^{\infty} \langle e_n | Te_j \rangle \langle e_j | x \rangle \quad (n = 1, 2, \dots). \quad (3.63)$$

Zufolge der Gleichung $Tx = \sum_{n=1}^{\infty} \langle e_n | Tx \rangle e_n$ ist die Abbildung T durch die (komplexen) Zahlen $a_{nj} = \langle e_n | Te_j \rangle$ ($n, j = 1, 2, \dots$) festgelegt. Die (unendliche) Matrix $[a_{nj}]$ bezeichnen wir als die zum Operator T zugehörige **Koordinatenmatrix** bezüglich des gegebenen (vollständigen) ONS.

3.3.1.3. Der adjungierte Operator eines beschränkten Operators im Hilbertraum

Es sei T ein linearer, beschränkter Operator im Hilbertraum mit $D(T) = H$, der H wieder in H abbildet. Für jedes $y \in H$ und jedes $x \in H$ setzen wir

$$f_y(x) = \langle y | Tx \rangle. \quad (3.64)$$

Für festes y und variables x ist $f_y(x)$ ein lineares Funktional auf H . Mittels der Schwarzschen Ungleichung erhalten wir unter Benutzung der Beschränktheit von T

$$|f_y(x)| = |\langle y | Tx \rangle| \leq \|y\| \cdot \|Tx\| \leq \|y\| \|T\| \cdot \|x\| = K\|x\| \quad (x \in H).$$

Also ist das lineare Funktional f_y beschränkt (stetig). Nach dem Satz 2.26 (von Riesz) gibt es genau ein Element $z \in H$ mit

$$f_y(x) = \langle z | x \rangle \quad (x \in H), \quad (3.65)$$

wobei die Gleichung $\|f_y\| = \|z\|$ besteht. Das Element z werde mit T^*y bezeichnet (dies bringt zum Ausdruck, daß z durch y eindeutig bestimmt ist). Es gilt also die Gleichung [beachte (3.64) und (3.65)]

$$\langle y | Tx \rangle = \langle T^*y | x \rangle \quad (3.66)$$

für alle $x, y \in H$. In Abhängigkeit von y ist T^*y linear und stetig (Beweis als Aufgabe), d. h., die Zuordnung $y \rightarrow T^*y$ definiert einen (stetigen) linearen Operator auf H .

D.3.14 Definition 3.14: Der Operator T^* , der durch die Gleichung (3.66) erklärt wird, heißt der zu T **adjungierte Operator** (auch: die **Adjungierte** von T).

Wichtige Eigenschaften des Übergangs zum adjungierten Operator sind die folgenden (T und S bezeichnen beschränkte lineare Operatoren, die auf H definiert sind):

- | | |
|--|-------------------------|
| (1) $(T + S)^* = T^* + S^*$, | (4) $I^* = I$, |
| (2) $(\lambda T)^* = \bar{\lambda} T^*$, | (5) $(T^*)^* = T$, |
| (3) $(ST)^* = T^* S^*$, | (6) $\ T^*\ = \ T\ $, |
| (7) $(T^{-1})^* = (T^*)^{-1}$ (falls einer dieser Operatoren existiert). | |

Zufolge dieser Rechenregeln gilt speziell für beliebige komplexe λ die Gleichung

$$(T - \lambda I)^* = T^* - \bar{\lambda} I$$

und

$$((T - \lambda I)^{-1})^* = (T^* - \bar{\lambda} I)^{-1},$$

falls eine der beiden Inversen als vorhanden vorausgesetzt wird, m. a. W., gehört λ zur Resolventenmenge $\varrho(T)$, so gehört $\bar{\lambda}$ zur Resolventenmenge $\varrho(T^*)$ (s. Def. 3.7) und umgekehrt. Wegen $\sigma(T) = \mathbf{K} \setminus \varrho(T)$ gilt dieselbe Aussage für das Spektrum: $\sigma(T^*)$ besteht genau aus den komplexen Zahlen, die zu den Zahlen aus $\sigma(T)$ konjugiert komplex sind.¹⁾

¹⁾ Im Hilbertraum übernimmt der adjungierte Operator die Rolle des dualen Operators (s. 3.2.3.). Auf den einfachen Zusammenhang zwischen beiden Begriffen gehen wir nicht weiter ein.

Beispiel 3.8: Es sei $A = [a_{jk}]$ eine komplexe (n, n) -Matrix. Dann wird durch die Zuordnung

$$\eta_j = \sum_{k=1}^n a_{jk} \xi_k \quad (j = 1, \dots, n)$$

eine (stetige) lineare Abbildung T des Hilbertraumes K^n in sich erklärt, die jedem n -dimensionalen komplexen Vektor $x = (\xi_1, \dots, \xi_n)$ einen ebensolchen Vektor $y = (\eta_1, \dots, \eta_n)$ zuordnet. (In Matrixschreibweise gilt $y' = Ax'$.) Die zugehörige adjungierte Transformation T^* wird durch die zu A hermitesch-konjugierte Matrix A^* geliefert, wobei $A = [a_{jk}]$ gilt mit

$$a_{jk}^* = \overline{a_{kj}}. \quad (3.67)$$

Mit anderen Worten, gilt $y = T^*x$, so ist

$$\eta_j = \sum_{k=1}^n \overline{a_{kj}} \xi_k \quad (j = 1, \dots, n) \quad (\text{vgl. Bsp. 3.6}).$$

Beispiel 3.9: Es sei T ein stetiger Integraloperator im Hilbertraum $L^2[a, b]$,

$$(Tx)(s) = \int_a^b \overline{K(s, t)} x(t) dt \quad (a \leq s \leq b; x \in L^2[a, b]).$$

Die adjungierte Abbildung T^* hat dann die Form

$$(T^*x)(s) = \int_a^b \overline{K(t, s)} x(t) dt \quad (a \leq s \leq b; x \in L^2[a, b]). \quad (3.68)$$

Der Kern $\overline{K(t, s)}$ heißt auch der zu $K(s, t)$ adjungierte Kern.

Wir geben schließlich noch einige Aussagen über vollstetige Operatoren in Hilberträumen an, darunter die Fredholmsche Alternative in der Fassung für diese Operatoren [s. auch Bemerkung bei (3.47)].

Satz 3.16: Ist H ein Hilbertraum und $T: H \rightarrow H$ ein linearer vollstetiger Operator, so ist auch T^* vollstetig. S.3.16

Satz 3.17: Ist H ein Hilbertraum und $T: H \rightarrow H$ ein linearer vollstetiger Operator, so gilt: S.3.17

(1) Das Spektrum von T^* besteht aus den zu den im Spektrum von T liegenden konjugiert komplexen Zahlen:

$$\sigma(T^*) = \{\lambda \in \mathbf{K} \mid \bar{\lambda} \in \sigma(T)\}.$$

(2) Der ganze Raum H läßt sich als orthogonale direkte Summe des Wertebereichs des Operators $T^* - \lambda I$ und des Nullraumes [vgl. (3.7), (3.6) und Def. 2.42] von $T - \lambda I$ für $\lambda \neq 0$ darstellen:

$$\begin{aligned} H &= (T^* - \bar{\lambda} I)[H] \oplus N(T - \lambda I) \quad (\lambda \neq 0) \\ N(T - \lambda I) &= \{x \in H \mid Tx - \lambda x = 0\}. \end{aligned} \quad (3.69)$$

(3) Die Nullräume $N(T - \lambda I)$ und $N(T^* - \bar{\lambda} I)$ haben für $\lambda \neq 0$ die gleiche, endliche Dimension.

Bemerkung 3.7: Die Eigenschaft (1) in obigem Satz gilt auch ohne die Forderung der Vollstetigkeit von T für lineares beschränktes T .

S.3.18 Satz 3.18 (Fredholmsche Alternative in der Fassung für Hilberträume): Es sei $T: \mathbf{H} \rightarrow \mathbf{H}$ ein vollstetiger linearer Operator. Dann gilt:

(1) Ist $\lambda \in \mathbf{K}$ kein Spektralwert des Operators T , so hat die Gleichung $Tx - \lambda x = y$ für jede rechte Seite $y \in \mathbf{H}$ (vorgegeben) genau eine Lösung. Diese Lösung hängt stetig von y ab.

(2) Ist $\lambda \neq 0$ ein Element des Spektrums des Operators T , so hat die Gleichung $Tx - \lambda x = y$ genau dann Lösungen, wenn der Vektor y zu allen Lösungen der Gleichung $T^*x - \lambda x = 0$ orthogonal ist. In diesem Fall hat die Gleichung $Tx - \lambda x = y$ unendlich viele Lösungen; diese haben die Gestalt

$$x = x_0 + \sum_{k=1}^n c_k x_k,$$

wobei x_0 eine spezielle Lösung der (inhomogenen) Gleichung $Tx - \lambda x = y$ bezeichnet und die x_k die (endlich vielen) linear unabhängigen Lösungen der (homogenen) Gleichung $Tx - \lambda x = 0$ sind.

3.3.1.4. Der adjungierte Operator eines unbeschränkten Operators im Hilbertraum

Für unbeschränkte Operatoren ist die Definition des adjungierten Operators komplizierter, da unbeschränkte Operatoren in der Regel nicht auf dem ganzen Raum definiert sind.

D.3.15 Definition 3.15: Es sei T ein auf einem dichten Teilraum $D(T)$ des Hilbertraumes \mathbf{H} definierter linearer Operator mit Werten in \mathbf{H} . Es sei $D(T^*)$ die Menge aller $x \in \mathbf{H}$, zu denen ein $y \in \mathbf{H}$ existiert, so daß die Gleichung

$$\langle x | Tz \rangle = \langle y | z \rangle \quad (3.70)$$

für alle $z \in D(T)$ gilt.

Für jedes $x \in D(T^*)$ setzen wir dann [wenn (3.70) erfüllt ist]

$$y = T^*x. \quad (3.71)$$

Der Operator T^* heißt der zu T adjungierte Operator.

Bemerkung 3.8: Es zeigt sich, daß $D(T^*)$ ein linearer Teilraum von \mathbf{H} ist und daß das Element y gemäß Formel (3.70) eindeutig festgelegt ist, so daß die Gl. (3.71) sinnvoll ist. Auf $D(T^*)$ ist dann T^* ein linearer Operator, der $D(T^*)$ in \mathbf{H} abbildet (aber nicht stetig zu sein braucht). Ist $D(T) = \mathbf{H}$ und ist T beschränkt, so gilt auch $D(T^*) = \mathbf{H}$, und T^* ist der zu T adjungierte Operator im Sinne der Def. 3.14.

Für die Anwendungen in der (Quanten-)Physik sind vor allem symmetrische, speziell selbstadjungierte Operatoren von Interesse. Wir bringen hier die Definitionen, Anwendungen folgen in Kap. 5.

D.3.16 Definition 3.16: Es sei T ein auf einem dichten Teilraum des Hilbertraumes \mathbf{H} definierter linearer Operator. T heißt **symmetrisch** (oder auch **hermitesch**), wenn $D(T) \subseteq D(T^*)$ und $T^*x = Tx$ für alle $x \in D(T)$ gilt; m. a. W.: T heißt symmetrisch, wenn die Gleichung

$$\langle x | Ty \rangle = \langle Tx | y \rangle \quad \text{für alle } x, y \in D(T) \quad (3.72)$$

gilt.

In diesem Zusammenhang ist der folgende Satz von E. Hellinger und O. Toeplitz von besonderem Interesse.

Satz 3.19: Es sei T ein auf $D(T) = H$ (Hilbertraum) definierter linearer Operator, S.3.19 der $D(T)$ in H abbildet. Falls die Gleichung

$$\langle x | Ty \rangle = \langle Tx | y \rangle$$

für alle $x, y \in H$ gilt, so ist T notwendig beschränkt (stetig).

Mit anderen Worten, ein überall auf H definierter symmetrischer linearer Operator ist (automatisch) beschränkt.

Eine im allgemeinen stärkere Forderung als die Symmetrie ist die Selbstadjungiertheit eines linearen Operators.

Definition 3.17: Ein auf einem dichten linearen Teilraum $D(T)$ eines Hilbertraumes H D.3.17 definierter linearer Operator heißt **selbstadjungiert**, wenn T symmetrisch ist und zusätzlich

$$D(T) = D(T^*)$$

gilt, d. h., wenn $T^* = T$ ist.

Der Nachweis für die Selbstadjungiertheit ist für unbeschränkte Operatoren relativ aufwendig, die Symmetrie läßt sich oft sehr viel leichter zeigen. Für beschränkte lineare Operatoren, die auf dem ganzen Raum (= Hilbertraum) definiert sind, fallen jedoch die Begriffe Symmetrie und Selbstadjungiertheit zusammen. Für unbeschränkte Operatoren kann man folgende Kriterien für Selbstadjungiertheit benutzen.

Satz 3.20: Es sei T ein auf einem dichten linearen Teilraum des Hilbertraumes H S.3.20 definierter linearer symmetrischer Operator. T ist genau dann selbstadjungiert, wenn die Bildmengen der Operatoren $T + iI$ und $T - iI$, also die Mengen $(T + iI)[D(T)]$ und $(T - iI)[D(T)]$ beide mit ganz H übereinstimmen.

Satz 3.21: Es sei T ein auf einem dichten linearen Teilraum des Hilbertraumes H S.3.21 definierter (linearer) selbstadjungierter Operator. Weiter sei S ein symmetrischer Operator mit $D(T) \subseteq D(S)$. Für gewisse reelle Zahlen δ und c mit $0 \leq \delta < 1$ und $c \geq 0$ sei die Ungleichung

$$\|Sx\| \leq \delta \|Tx\| + c\|x\| \quad (x \in D(T))$$

erfüllt. Dann ist der Operator $T + S$ mit $D(T + S) = D(T)$ selbstadjungiert.

Dieser Satz, der auf T. Kato zurückgeht, ist vor allem in der Quantenmechanik nützlich (Beweis: [36, S. 209]).

Ein Beispiel für einen selbstadjungierten, unbeschränkten Operator erhält man in folgender Weise.

Beispiel 3.10: Es sei $H = L^2(R)$ der Hilbertraum der (Klassen von) quadratisch summierbaren (komplexwertigen) Funktionen $x(t)$ ($-\infty < t < +\infty$) und $W^{2,2}(R)$ der zugehörige Sobolev-Raum der verallgemeinert zweimal differenzierbaren Funktionen. $W^{2,2}(R)$ liegt in H dicht [vgl. (2.27) und Satz 2.19]. Der Operator $T: W^{2,2}(R) \rightarrow H$, der durch die Vorschrift

$$Tx(t) = -x''(t) \quad (-\infty < t < +\infty; x \in W^{2,2}(R))$$

erklärt ist, ist selbstadjungiert in H (Beweis: [36, S. 287]).

3.3.2. Vollstetige selbstadjungierte Operatoren im Hilbertraum

Die Eigenschaft der Vollstetigkeit hat für die Gestalt eines auf dem ganzen Raum definierten selbstadjungierten Operators wesentliche Konsequenzen, die zuerst von D. Hilbert erkannt wurden.

S.3.22 Satz 3.22: *Es sei H ein (separabler)¹⁾ Hilbertraum und $T: H \rightarrow H$ ein selbstadjungierter, vollstetiger Operator. Dann besitzt T Eigenwerte $\lambda_1, \lambda_2, \dots, \lambda_n, \dots$, und es existiert in H vollständiges ONS von zugehörigen Eigenvektoren $e_1, e_2, \dots, e_n, \dots$:*

$$Te_n = \lambda_n e_n \quad (n = 1, 2, \dots).$$

Ist H unendlichdimensional, so gilt $\lim_{n \rightarrow \infty} \lambda_n = 0$.

Unter den Bedingungen des Satzes 3.22 hat die dem Operator T zugeordnete Matrix die folgende Gestalt [s. (3.61)]:

$$a_{nj} = \langle e_n | Te_j \rangle = \langle e_n | \lambda_j e_j \rangle = \lambda_j \langle e_n | e_j \rangle = \begin{cases} \lambda_n & \text{für } j = n, \\ 0 & \text{für } j \neq n. \end{cases}$$

Dem Operator T entspricht somit eine (i. allg. unendliche) Diagonalmatrix, in der in der Diagonale die Eigenwerte des Operators stehen; m. a. W., vollstetige selbstadjungierte Operatoren lassen eine „Hauptachsentransformation“ zu, wie dies von reellen symmetrischen Matrizen her bekannt ist. Die Bedingung der Selbstadjungiertheit ist dabei wesentlich. Es gibt nicht-selbstadjungierte vollstetige Operatoren, z. B. den Operator

$$(Tx)(s) = \int_0^s x(t) dt \quad (0 \leq s \leq 1) \quad x \in L^2_{\mathbb{R}}[0, 1]$$

im Hilbertraum $L^2_{\mathbb{R}}[0, 1]$, die keinen einzigen Eigenwert besitzen.

Wie für reelle symmetrische Matrizen gilt die folgende Aussage über die Eigenwerte und Eigenvektoren eines (nicht notwendig beschränkten) symmetrischen Operators.

S.3.23 Satz 3.23: *Es sei T ein symmetrischer Operator im Hilbertraum. Dann gilt*

- (I) *Jeder Eigenwert von T ist reell;*
- (II) *Eigenvektoren zu verschiedenen Eigenwerten sind zueinander orthogonal.*

Beweis. Ist $Tx = \lambda x$ mit $x \neq 0$, so folgt mittels der Symmetrie von T die Gleichheit $\lambda \langle x | x \rangle = \langle x | \lambda x \rangle = \langle x | Tx \rangle = \langle Tx | x \rangle = \langle \lambda x | x \rangle = \lambda \langle x | x \rangle$ mit $\|x\|^2 = \langle x | x \rangle > 0$. Folglich muß $\lambda = \lambda$ gelten; λ ist reell. Gilt zum anderen $Tx = \lambda x$ und $Ty = \mu y$ mit $x \neq 0$, $y \neq 0$ und $\lambda \neq \mu$, so sind λ, μ reell, und wir erhalten $\mu \langle x | y \rangle = \langle x | \mu y \rangle = \langle x | Ty \rangle = \langle Tx | y \rangle = \langle \lambda x | y \rangle = \lambda \langle x | y \rangle = \lambda \langle x | y \rangle$ also $\langle x | y \rangle = 0$. ■

3.3.3. Störungsrechnung

In der Störungsrechnung (allgemeiner: in der Störungstheorie) beschäftigt man sich mit der Frage des Verhaltens von Eigenwerten und Eigenvektoren (bzw. allgemeinen Operatoreigenschaften) eines Operators in Abhängigkeit von sog. „Störungen“ eines gegebenen festen Operators. Zum Beispiel sei T_0 ein selbstadjungierter Operator im Hilbertraum H , S ein weiterer selbstadjungierter

¹⁾ Ein Hilbertraum H heißt **separabel**, wenn es eine Folge $\{x_n\} \in H$ gibt, so daß zu jedem $x \in H$ eine Teilfolge $\{x_{n_k}\}$ von $\{x_n\}$ existiert, so daß $x = \lim_{k \rightarrow \infty} x_{n_k}$ ist.

Operator mit dem gleichen Definitionsbereich wie T_0 , der (wie T_0) Werte in H besitzt. Mit ε werde eine (beliebige) reelle Zahl bezeichnet. Der Operator

$$T_0 + \varepsilon S \quad (3.73)$$

kann (insbesondere für betragsmäßig kleine ε) als eine „Störung“ von T_0 angesehen werden. Der Operator T_0 besitzt einen Eigenwert λ_0 mit zugehörigem Eigenvektor x_0 ; ($x_0 \neq 0$)

$$T_0 x_0 = \lambda_0 x_0.$$

Man fragt nach den „gestörten“ Eigenwerten $\lambda(\varepsilon)$ bzw. Eigenvektoren $x(\varepsilon)$, die die Gleichungen

$$(T_0 + \varepsilon S) x(\varepsilon) = \lambda(\varepsilon) x(\varepsilon), \quad (3.74)$$

$$\lim_{\varepsilon \rightarrow 0} x(\varepsilon) = x_0, \quad \lim_{\varepsilon \rightarrow 0} \lambda(\varepsilon) = \lambda_0 \quad (3.75)$$

erfüllen. Die Forderung (3.75) besagt, daß für kleine Werte von ε der gestörte Eigenwert (bzw. der gestörte Eigenvektor) in der Nähe des ungestörten Eigenwertes (bzw. Eigenvektors) liegt und für $\varepsilon \rightarrow 0$ in diesen übergeht (Stabilitätseigenschaft). Derartige Aufgabenstellungen haben u. a. bei der quantenmechanischen Erklärung der Aufspaltung von Spektrallinien durch ein Magnetfeld große Bedeutung erlangt [4]. Es sei nämlich T_0 der selbstadjungierte Hamilton-Operator eines quantenmechanischen Systems, λ_0 ein Eigenwert (Energieniveau) des stationären Zustandes x_0 (s. 1.2.3.). Wirken nun äußere Felder, die durch den Störoperator S beschrieben werden, so besteht ein wichtiges Problem darin, zu λ_0 und x_0 „benachbarte“ [s. (3.74)] stationäre Zustände bzw. Energieniveaus zu finden, die zu $T_0 + \varepsilon S$ „gehören“; die Gl. (3.75) beschreiben die „Aufspaltung“ in benachbarte Zustände und Niveaus.

Im folgenden sei vorausgesetzt, daß der Operator T_0 ein vollständiges ONS von Eigenvektoren $e_0, e_1, e_2, \dots, e_n, \dots$ zu den Eigenwerten $\lambda_0, \lambda_1, \dots, \lambda_n, \dots$ besitzt. Es sei e_0 (bis auf Vielfache) der einzige Eigenvektor zum Eigenwert λ_0 . Mit den obigen Bezeichnungen sei $x_0 = e_0$.

Zur Lösung der Gl. (3.74) wird ein (formaler) Potenzreihenansatz bezüglich ε aufgestellt [der bereits (3.75) erfüllt]:

$$\begin{aligned} \lambda(\varepsilon) &= \lambda_0 + \mu_1 \varepsilon + \mu_2 \varepsilon^2 + \dots = \lambda_0 + \sum_{k=1}^{\infty} \mu_k \varepsilon^k, \\ x(\varepsilon) &= e_0 + \varepsilon z_1 + \varepsilon^2 z_2 + \dots = e_0 + \sum_{k=1}^{\infty} \varepsilon^k z_k, \end{aligned} \quad (3.76)$$

wobei die Koeffizienten μ_k und die Vektoren z_k aus (3.74) zu bestimmen sind. Für die weitere Rechnung fordert man noch die Gleichung

$$\langle x(\varepsilon) | x(\varepsilon) \rangle = 1, \quad (3.77)$$

die nur eine Normierung von $x(\varepsilon)$ beinhaltet. Ferner werden die Vektoren z_k nach den Vektoren des ONS (e_n) entwickelt:

$$z_k = \sum_{j=0}^{\infty} \alpha_j^{(k)} e_j \quad (k = 1, 2, \dots). \quad (3.78)$$

Einsetzen von (3.76), (3.77) und (3.78) in die Gl. (3.74) sowie anschließender Koeffizientenvergleich liefert die sog. **Formeln der Störungstheorie** für die wesentlichen Bestandteile $\mu_1, \mu_2; z_1, z_2$:

$$\mu_1 = \langle S e_0 | e_0 \rangle; \quad \mu_2 = \sum_{j=1}^{\infty} \frac{|\langle S e_0 | e_j \rangle|^2}{\lambda_0 - \lambda_j}, \quad (3.79)$$

$$z_1 = \sum_{j=1}^{\infty} \frac{\langle e_j | S e_0 \rangle}{\lambda_0 - \lambda_j} e_j, \quad (3.80)$$

$$\begin{aligned} z_2 &= \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} \frac{\langle e_k | S e_0 \rangle \langle e_j | S e_k \rangle}{(\lambda_0 - \lambda_j)(\lambda_0 - \lambda_k)} e_j - \sum_{j=1}^{\infty} \frac{\langle S e_0 | e_0 \rangle \langle e_j | S e_0 \rangle}{(\lambda_0 - \lambda_j)^2} e_j \\ &\quad - \frac{1}{2} \left(\sum_{j=1}^{\infty} \frac{|\langle e_j | S e_0 \rangle|^2}{(\lambda_0 - \lambda_j)^2} \right) e_0. \end{aligned}$$

Damit sind $\lambda(\varepsilon)$ und $x(\varepsilon)$ für betragsmäßig kleine ε näherungsweise berechenbar.

Tritt, im Gegensatz zu unserer bisherigen Voraussetzung, der Fall ein, daß zum Eigenwert λ_0 mehrere linear unabhängige Eigenvektoren gehören, so muß der Ansatz für $x(\varepsilon)$ modifiziert werden. Es gelte z. B.

$$\lambda_0 = \lambda_1 = \dots = \lambda_{n-1} \quad (\lambda_0 \text{ ist ein } n\text{-facher Eigenwert}),$$

dann wird der Ansatz

$$\lambda(\varepsilon) = \lambda_0 + \mu_1 \varepsilon + \mu_2 \varepsilon^2 + \dots, \quad (3.81)$$

$$x(\varepsilon) = \sum_{r=0}^{n-1} c_r e_r + \varepsilon z_1 + \varepsilon^2 z_2 + \dots$$

anstelle des Ansatzes (3.76) mit unbekannten Koeffizienten c_r ($r = 0, \dots, n-1$) gemacht.

Für die Bestimmung der Koeffizienten c_2 und des Wertes μ_1 erhält man [mittels Einsetzen in (3.74) und Koeffizientenvergleich] ein Matrizen-Eigenwertproblem, die sog. **Säkulargleichungen**:

$$\begin{bmatrix} \langle e_0 | S e_0 \rangle & \langle e_0 | S e_1 \rangle & \dots & \langle e_0 | S e_{n-1} \rangle \\ \langle e_1 | S e_0 \rangle & & & \\ \vdots & \vdots & & \vdots \\ \langle e_{n-1} | S e_0 \rangle & \langle e_{n-1} | S e_1 \rangle & \dots & \langle e_{n-1} | S e_{n-1} \rangle \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_{n-1} \end{bmatrix} = \mu_1 \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_{n-1} \end{bmatrix}. \quad (3.82)$$

Mit anderen Worten, μ_1 ergibt sich als Eigenwert und $(c_0, c_1, \dots, c_{n-1})$ als Eigenvektor der „Säkularmatrix“ $[\langle e_j | S e_k \rangle]$ ($0 \leq j, k \leq n-1$). Hat diese Matrix verschiedene Eigenwerte, so erfolgt die bereits erwähnte „Aufspaltung der Eigenwerte“ für den Eigenwert λ_0 des Operators T_0 bei Einschaltung der Störung εS .

Beispiel 3.11. Im reellen Hilbertraum $H = R^3$ seien T_0 bzw. S die durch die Matrizen

$$T_0 = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 1 & 1 \\ 1 & 1 & 3 \end{bmatrix} \quad \text{bzw.} \quad S = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

beschriebenen linearen Operatoren (selbstadjungiert). Der Operator T_0 besitzt die Eigenwerte $\lambda_0 = 3 + \sqrt{2}$; $\lambda_1 = 3 - \sqrt{2}$; $\lambda_2 = -1$ mit einem zugehörigen (vollständigen) ONS von Eigenvektoren.

$$e_0 = \begin{bmatrix} 1/2 \\ 1/2 \\ 1/\sqrt{2} \end{bmatrix}; \quad e_1 = \begin{bmatrix} 1/2 \\ 1/2 \\ -1/\sqrt{2} \end{bmatrix}; \quad e_2 = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \\ 0 \end{bmatrix}.$$

Nach den Formeln (3.79) ergibt sich durch die Entwicklung des Eigenwertes $\lambda_0(\varepsilon)$ des gestörten Operators $T_0 + \varepsilon S$ nach Potenzen von ε bis zu Gliedern mit ε^2 die Näherung (für kleine Werte

von $|\varepsilon|$) $\lambda_0(\varepsilon) \approx \lambda_0 + \mu_1 \varepsilon + \mu_2 \varepsilon^2 = 3 + \sqrt{2} + \frac{1}{2} \varepsilon + \frac{\sqrt{2}}{16} \varepsilon^2$. Für den zugehörigen (normierten)

Eigenvektor $x(\varepsilon)$ erhält man nach (3.80) näherungsweise

$$x(\varepsilon) \approx e_0 + \varepsilon z_1 + \varepsilon^2 z_2 = e_0 - \frac{\sqrt{2}}{8} \varepsilon e_1 - \frac{\varepsilon^2}{64} e_0 = \begin{bmatrix} \frac{1}{2} - \frac{\sqrt{2}}{16} \varepsilon - \frac{1}{128} \varepsilon^2 \\ \frac{1}{2} - \frac{\sqrt{2}}{16} \varepsilon - \frac{1}{128} \varepsilon^2 \\ \frac{1}{\sqrt{2}} + \frac{1}{8} \varepsilon - \frac{\sqrt{2}}{128} \varepsilon^2 \end{bmatrix}.$$

4. Ausgewählte Anwendungen

4.1. Distributionen

In Abschnitt 1.2.2. waren schon Distributionen als lineare (stetige) Funktionale über dem linearen Raum $\hat{C}^\infty_{\mathbb{R}^n}$ der über dem \mathbb{R}^n definierten, finiten, beliebig oft differenzierbaren Funktionen erklärt worden. Sie spielen in der modernen Theorie der Differentialgleichungen eine überragende Rolle [38], [39], [42]. Wir gehen jetzt genauer darauf ein.

4.1.1. Distributionen als lineare stetige Funktionale

Definition 4.1: Unter einer **Distribution** L verstehen wir ein lineares stetiges Funktional L über dem Grundraum $D = \hat{C}^\infty(\mathbb{R}^n)$ (s. Def. 2.17), wobei in D die Konvergenz einer Elementfolge $\{\varphi_m\}$ gegen φ gegeben ist durch [s. (2.14)] **D.4.1**

$$\{\varphi_m\} \rightarrow \varphi \Leftrightarrow \begin{cases} \{D^\alpha \varphi_m\} \rightarrow D^\alpha \varphi \quad (\alpha \geq 0, \text{ ganz});^1) \\ \text{es existiert unabhängig von } m \text{ eine beschränkte Menge} \\ U \subseteq \mathbb{R}^n \text{ mit } \text{supp } \varphi_m \subseteq U. \end{cases} \quad (4.1)$$

$$U \subseteq \mathbb{R}^n \text{ mit } \text{supp } \varphi_m \subseteq U. \quad (4.2)$$

D ist nichtleer [38, S. 69]. Die Menge der Distributionen bezeichnen wir mit D' ; D' ist nichtleer: denn jede lokalsummierbare (s. Def. 2.19) Funktion f liefert ein Element L_f aus D' , indem wir f das Integral

$$L_f \varphi = \int_{\mathbb{R}^n} f(x) \varphi(x) dx = (f, \varphi) \quad (\varphi \in D) \quad (4.3)$$

zuordnen. Dieses L_f ist bei festem f über D linear und stetig, also eine Distribution. Distributionen, die eine Darstellung (4.3) gestatten, heißen regulär, andernfalls singular. Die regulären Distributionen sind eineindeutig den lokalsummierbaren Funktionen zugeordnet, wenn man äquivalente Funktionen als gleich ansieht (s. 2.2.2.). Wenn wir in der Menge $L^1_{\text{loc}}(\mathbb{R}^n)$ jeweils f durch L_f ersetzen und bis auf Mengen vom Maß 0 gleiche Funktionen als gleich ansehen, die so gewonnene Menge sei \tilde{M} , so gilt $\tilde{M} \subseteq D'$. Da das bezüglich $0 \in \mathbb{R}^n$ gebildete Dirac-Funktional [s. (1.49)] δ_0 :

$$(\delta_0, \varphi) = L_{\delta_0}(\varphi) = \varphi(0) \quad (\varphi \in D) \quad (4.4)$$

auch ein lineares stetiges Funktional über D ist, was nicht durch ein lokalintegrables f repräsentierbar ist [38, S. 74], so ist $\tilde{M} \subsetneq D'$, weswegen Distributionen auch „verallgemeinerte Funktionen“ heißen. Man kann jeder von ihnen auch einen **Träger** (für Funktionen s. Def. 2.17) zuordnen: Wir sagen nämlich zunächst, daß eine verallgemeinerte Funktion L in einem Gebiet $G \subseteq \mathbb{R}^n$ (G habe nichtleeres Inneres, int $G \neq \emptyset$, s. Satz 2.31) **verschwindet**, wenn gilt $L\varphi = 0$ für alle $\varphi \in D$ mit $\text{supp } \varphi \subseteq \text{int } G$, und nun wird definiert:

Definition 4.2: Ein Punkt $x \in \mathbb{R}^n$ gehört zum **Träger** von L , $\text{supp } L$, falls in keiner Umgebung von x gilt $L = 0$. **D.4.2**

Beispielsweise ist daher $\text{supp } L_{\delta_0} = \{0\}$, und zwei Distributionen L_1, L_2 heißen **gleich**, wenn $L_1\varphi - L_2\varphi = 0$ gilt; sie heißen **gleich über** G (wobei G wie oben ein Gebiet mit nichtleerem Inneren ist), falls $L_1 - L_2 = 0$ über G gilt.

¹⁾ Gleichmäßige Konvergenz für jeden Multiindex α .

4.1.2. Rechenregeln

D.4.3 Definition 4.3 (Differentiationsregel): Ist die Distribution L gegeben, so ist die (distributionelle) Ableitung der Ordnung α von L die Distribution $D^\alpha L$ mit [s. auch (1.50), (2.23), (2.14)]

$$(D^\alpha L)(\varphi) = (-1)^{|\alpha|} L(D^\alpha \varphi) \quad (\varphi \in D). \quad (4.5)$$

Die Reihenfolge der Ableitungen (bei $n > 1$) spielt keine Rolle.

Beispiel 4.1: Es sei L die Delta-Distribution δ_0 . Dann ist

$$(D^\alpha \delta_0)(\varphi) = (-1)^{|\alpha|} \delta_0(D^\alpha \varphi) = (-1)^{|\alpha|} D^\alpha \varphi(0) \quad (\varphi \in D). \quad (4.5a)$$

Die Ladungsdichte für einen im Punkt $x = 0$ liegenden Dipol mit dem elektrischen Moment $+1$ auf einer Geraden (x -Achse) entspricht $-\delta'_0$ [38].

Ist f stetig differenzierbar bis auf die Stelle $\hat{x} \in \mathbf{R}$, an der endliche Grenzwerte von f und f' von rechts und links existieren, so berechnet sich für die dieser lokal-integrierbaren Funktion f zugeordneten Distribution L_f die 1. Ableitung wie folgt:

$$\begin{aligned} (L_f', \varphi) &= L_f'(\varphi) = \left(\int_{-\infty}^{+\infty} f(x) \varphi(x) dx \right)' = - \int_{-\infty}^{+\infty} f(x) \varphi'(x) dx \\ &= - \int_{-\infty}^{\hat{x}} f(x) \varphi'(x) dx - \int_{\hat{x}}^{+\infty} f(x) \varphi'(x) dx \\ &= -[f(\hat{x} - 0) \varphi(\hat{x})] + \int_{-\infty}^{\hat{x}} f'(x) \varphi(x) dx + f(\hat{x} + 0) \varphi(\hat{x}) \\ &\quad + \int_{\hat{x}}^{+\infty} f'(x) \varphi(x) dx \\ &= \int_{-\infty}^{+\infty} f'(x) \varphi(x) dx + (f(\hat{x} + 0) - f(\hat{x} - 0)) \varphi(\hat{x}), \end{aligned}$$

also:

$$(L_f)' = L_{f'} + (f(\hat{x} + 0) - f(\hat{x} - 0)) \delta_{\hat{x}}, \quad (4.5b)$$

bei der δ -Distribution steht gerade die Sprunghöhe von f an der Stelle \hat{x} als Faktor. Ist also f an der Stelle \hat{x} stetig, so ist $(L_f)' = L_{f'}$.

D.4.4 Definition 4.4 (Konvergenz einer Folge $\{L_m\}$ von Distributionen): $\{L_m\}$ heißt konvergent gegen die Distribution L , wofür wir kurz $\{L_m\} \rightarrow L$ schreiben, falls

$$\varphi\{L_m\} \rightarrow L\varphi \quad (\varphi \in D) \quad (4.6)$$

gilt (schwache Konvergenz der Funktionale über D).

Beispiel 4.2: $\{E_n\} \rightarrow S\delta_0$ von 1.2.2. bei (1.57).

Gilt für die Folge $\{L_m\}$ von Distributionen, daß $\{L_m \varphi\}$ für jedes $\varphi \in D$ konvergent ist gegen die Zahl $G(\varphi)$, so ist G eine Distribution (Satz von Banach-Steinhaus für D). Umgekehrt gilt: D ist dicht in D' eingebettet. Jede verallgemeinerte Funktion ist

also Grenzwert [im Sinne von (4.6)] regulärer Distributionen. Ferner ist D' mit der Konvergenz in (4.6) vollständig. $\{L_m\} \rightarrow L$ zieht auch die Konvergenz der Folge der Ableitungen $\{D^s L_m\} \rightarrow D^s L$ nach sich.

Beispiel 4.3: Der Hauptwert (Bd. 2, S. 231) von $\int_{\mathbf{R}} \frac{\varphi(x)}{x} dx$ ist in Abhängigkeit von $\varphi \in D$ eine

Distribution P : Denn er ist der (existierende) Grenzwert der Folge $\left\{ \int_{|x| > \frac{1}{k}} \frac{\varphi(x)}{x} dx \right\}_{k \geq 1}$ regulärer Distributionen: Man setze in (4.3)

$$f_k \begin{cases} = 0 & (|x| \leq \frac{1}{k}), \\ = \frac{1}{x} & (|x| > \frac{1}{k}). \end{cases}$$

P ist eine *singuläre* Distribution. Sie tritt in den Anwendungen (Quantenphysik) bei den **Formeln von Sochozki** auf:

$$\frac{1}{x + i0} = -i\pi \delta_0 + P, \quad \frac{1}{x - i0} = i\pi \delta_0 + P. \quad (4.6a, b)$$

Die Zeichen $+0$ und -0 auf der linken Seite sollen auf die Herkunft durch Grenzwertbildung im Sinne von (4.6) hinweisen: Betrachtet man nämlich die den Funktionen $\frac{1}{x + i\varepsilon_n}$ für $\varepsilon_n > 0$, $\lim_{n \rightarrow \infty} \varepsilon_n = 0$ zugeordneten regulären Distributionen, also

$$L_n \varphi = \int_{\mathbf{R}} \frac{1}{x + i\varepsilon_n} \varphi(x) dx,$$

so konvergiert die Folge $\{L_n \varphi\}$ für jedes φ :

$$\{L_n \varphi\} \rightarrow (-i\pi \varphi(0) + P\varphi) \quad (\varphi \in D),$$

und das ist gerade die rechte Seite von (4.6a) [39].

Definition 4.5 (Multiplikation einer Distribution L mit einer Funktion f): Ist $f \in C^\infty$, **D.4.5** so definiert man das Produkt $f \cdot L$ durch

$$(fL) \varphi = L(f\varphi) \quad (\varphi \in D). \quad (4.7)$$

Tatsächlich ist die rechte Seite der letzten Zeile sinnvoll, da $f \cdot \varphi \in D$.

Beispiel 4.4: Es sei $f \in C^\infty$. Wir bilden $f\delta_0$:

$$(f\delta_0) \varphi = \delta_0(f\varphi) = f(0) \varphi(0), \quad \text{also} \quad f\delta_0 = f(0) \delta_0. \quad (4.7a)$$

Ferner sei $f(x) \equiv x$. Wir bilden xP :

$$(xP) \varphi = P(x\varphi) = \text{Hauptwert} \int_{\mathbf{R}} x \frac{\varphi(x)}{x} dx = \int_{\mathbf{R}} 1 \varphi(x) dx = I, \quad (4.7b)$$

wobei I die von $f \equiv 1$ erzeugte Distribution ist.

Definition 4.6 (Faltung zweier Distributionen): Es sei L_1 eine beliebige Distribution **D.4.6** und L_2 eine Distribution mit kompaktem Träger G , d. h., $\text{supp } L_2$ ist eine kompakte Teilmenge $G \subseteq \mathbf{R}^n$. Dann versteht man unter der **Faltung** $L_1 * L_2$ von L_1 mit L_2 die Distribution

$$L_1 * L_2 : (L_1 * L_2) \varphi = L_1(L_2 \varphi(y + \cdot)) \quad (\varphi \in D). \quad (4.8)$$

Durch (4.8) ist tatsächlich eine Distribution definiert [39, S. 69], weil $L_2\varphi(y + \cdot)$ eine Funktion aus D ist.

Beispiel 4.5: Es seien L_1, L_2 reguläre Distributionen, die von den Funktionen $L_1(x), L_2(x)$ erzeugt werden: $(L_1, \varphi) = \int_{\mathbb{R}^n} L_1(x) \varphi(x) dx$, und $G = \text{supp } L_2$ sei kompakt. Dann ist

$$\begin{aligned}(L_1 * L_2) \varphi &= \int_{\mathbb{R}^n} L_1(y) \left[\int_{\mathbb{R}^n} L_2(x) \varphi(y+x) dx \right] dy \\ &= \int_{\mathbb{R}^n} L_1(y) \left[\int_{\mathbb{R}^n} L_2(\xi - y) \varphi(\xi) d\xi \right] dy\end{aligned}$$

(Satz von Fubini [36] für die Vertauschung der Integrationen)

$$= \int_{\mathbb{R}^n} \varphi(\xi) \left[\int_{\mathbb{R}^n} L_1(y) L_2(\xi - y) dy \right] d\xi,$$

und da das zweite Integral gerade die Faltung der beiden Funktionen $L_1(x), L_2(x)$ ist, können wir

$$\begin{aligned}(L_1 * L_2) \varphi &= \int_{\mathbb{R}^n} L_1(\cdot) * L_2(\cdot) \varphi(\xi) d\xi, \\ &= (L_1(\cdot) * L_2(\cdot)) \varphi\end{aligned}\tag{4.9}$$

schreiben, d. h., die Faltung zweier regulärer Distributionen ist (unter der Bedingung $\text{supp } L_2$ kompakt) gerade die von der Faltung der Funktionen erzeugte Distribution.

Beispiel 4.6: Es sei L eine beliebige Distribution. Dann existiert die Faltung mit der Delta-Distribution δ_0 , und es gilt

$$(L * \delta_0) \varphi = L(\delta_0(\varphi(y + \cdot))),$$

und nach Definition der Delta-Distribution folgt:

$$(L * \delta_0) \varphi = L(\varphi(y)) = L(\varphi), \quad (\varphi \in D),$$

und daher ist

$$L * \delta_0 = L.\tag{4.10}$$

Da die Faltung zweier Distributionen, wenn sie existiert, **kommutativ** ist, ist auch

$$\delta_0 * L = L.\tag{4.10'}$$

Die beiden letzten Formeln korrespondieren zu Bd. 10: „bezüglich der Faltung ist die Delta-Distribution das Einselement“.

Da Distributionen differenzierbar sind, kann auch die Faltung $L_1 * L_2$ zweier Distributionen **differenziert** werden:

$$D^\alpha(L_1 * L_2) = (D^\alpha L_1) * L_2 = L_1 * (D^\alpha L_2),\tag{4.11}$$

und das ergibt wegen (4.10) mit $L_1 = \delta_0$ die interessante Formel

$$(D^\alpha \delta_0) * L_2 = \delta_0 * D^\alpha L_2 = D^\alpha L_2;\tag{4.12}$$

Differentiation kann durch eine Faltung mit der entsprechenden Ableitung der δ_0 -Distribution (die wie δ_0 selbst kompakten, nämlich einpunktigen Träger hat) ersetzt werden. Man benötigt daher bei Benutzung der Faltung vom Differentiationskalkül nur die Ableitungen von δ_0 .

Hat weder L_1 noch L_2 kompakten Träger, so muß (4.11) nicht gelten! Dazu sei θ die der Heaviside-Funktion [s. (1.44)], 1 die der Funktion $f \equiv 1$ und 0 die der

Funktion $f \equiv 0$ zugeordnete Distribution. Dann (existiert $\Theta * I$ nicht, und es) gilt

$$\Theta' * I = \delta * I = I; \quad \Theta * I' = \Theta * 0 = 0,$$

wobei sich I' so berechnet ($\varphi \in D$):

$$\left(\int_{\mathbf{R}} 1 \cdot \varphi \, dx \right)' = - \int_{\mathbf{R}} 1 \cdot \varphi' \, dx = 0 \quad (\text{Integration!}).$$

Für einen partiellen Differentialoperator $P(D)$ mit konstanten Koeffizienten folgt wegen der Linearität aus (4.12)

$$P(D) = \sum_{|\alpha| \leq m} a_{\alpha} D^{\alpha}; \quad (a_{\alpha} = a_{\alpha_1 \alpha_2 \dots \alpha_n}),$$

$$P(D) L = (P(D) \delta_0) * L. \quad (4.13)$$

Eine Distribution G , die der distributionellen Differentialgleichung

$$P(D) G = \delta_0 \quad (4.14)$$

genügt, heißt eine „**Grundlösung**“ des Differentialoperators $P(D)$. In 1.2.2. hatten wir zum Differentialausdruck D_2 in (1.35)

$$D_2 J = P(D) J = L \ddot{J} + \frac{1}{C} J \quad (4.15)$$

eine Greensche Funktion J gefunden; die ihr zugeordnete Distribution genüge der distributionellen Gleichung $P(D) J = \delta_0$ als Funktion ihres ersten Arguments, als Funktion der Differenz ihrer Argumente genügt sie der Gl. (4.14) (und weil sie den homogenen Anfangsbedingungen genüge, war sie auch eindeutig bestimmt) und ist daher eine Grundlösung von $P(D)$.

Es sei nun E eine Distribution mit kompaktem Träger. Dann gilt folgende Darstellungsformel für Lösungen einer Differentialgleichung im distributionellem Sinne:

Satz 4.1: Ist $P(D)$ ein beliebiger Differentialausdruck mit konstanten Koeffizienten, S.4.1 und G eine Grundlösung von $P(D)$, also

$$P(D) G = \delta_0, \quad (4.16)$$

so gibt es eine Lösung L der distributionellen Differentialgleichung $P(D) L = E$ der Gestalt

$$L = E * G. \quad (4.17)$$

L heißt auch eine **verallgemeinerte Lösung** der Differentialgleichung $P(D) L = E$.

Zur Frage, wann eine verallgemeinerte Lösung eine klassische ist, vgl. z. B. [41/II], [39].

Die physikalische Bedeutung der Lösungsdarstellung (4.17) besteht in der Superposition $E * G$ der punktwisen Wirkungen des äußeren Einflusses E : G ist die Antwort des Systems auf einem punkthaften Einfluß der Stärke 1 an der Stelle 0, $G * E$ die richtige Überlagerung der mit E gewichteten punkthaften Wirkungen.

Die Lösungsdarstellung des homogenen Anfangswertproblems

$$L \ddot{J} + \frac{1}{C} J = E(t), \quad J(0) = \dot{J}(0) = 0$$

lautet (s. 1.2.2.; E stetige Funktion)

$$J(t) = \int_0^t J(t; \xi) E(\xi) d\xi;$$

und wenn wir berücksichtigen, daß $J(t; \xi)$ nur von der Argumentendifferenz $t - \xi$ abhängt, und $J(t; \xi) = G(t - \xi)$ schreiben:

$$J(t) = \int_0^t G(t - \xi) E(\xi) d\xi,$$

so entspricht dies wegen (4.9) gerade der Faltung

$$J = G * E$$

in (4.17). ($G(t - \xi)$ verschwindet für $\xi \geq t$).

Beweis von (4.17):

$$P(D) (E * G) = E * (P(D) G) = E * \delta_0 = E. \quad (4.17')$$

(4.11)
(4.16)
(4.10)

Wir wollen die Rechenregeln durch die Fouriertransformation ergänzen (und zugleich eine Möglichkeit angeben, zu einem gegebenen Differentialausdruck $P(D)$ eine Grundlösung G zu ermitteln). Damit die Fouriertransformierte FL einer Distribution L wieder eine Distribution ist und durch

$$FL : (FL) (\varphi) = L(F\varphi) \quad (\varphi \in D) \quad (4.18)$$

erklärt werden kann, muß $F\varphi$ in D liegen. Man weiß aber [s. auch (4.21)], daß $F\varphi$ eine beliebig oft differenzierbare Funktion ist, die nur dann einen kompakten Träger haben kann (also in D liegt), wenn gilt $\varphi \equiv 0$. Man muß daher eine Funktionenmenge \mathfrak{S} als Grundmenge zulassen, die umfassender als D ist. Als Grundmenge \mathfrak{S} wird jetzt die Menge der über dem R^n beliebig oft differenzierbaren Funktionen φ gewählt, für die gilt

$$\lim_{|x| \rightarrow \infty} (1 + |x|^2)^k |D^\alpha \varphi| = 0 \quad (4.19)$$

für jedes ganze k und für jeden Multiindex α . Es ist dann die lineare Abbildung F ein Isomorphismus von \mathfrak{S} auf \mathfrak{S} , der wegen (4.20) auch stetig ist.

Um lineare *stetige* Funktionale (sog. Distributionen schwachen Wachstums) über \mathfrak{S} bilden zu können, braucht man in \mathfrak{S} einen Konvergenzbegriff: Für Funktionen φ_k und φ aus \mathfrak{S} gilt:

$$(\varphi_k)_{k \geq 0} \rightarrow \varphi \Leftrightarrow (x^\beta D^\alpha \varphi_k)_{k \geq 0} \rightarrow x^\beta D^\alpha \varphi \quad (4.20)$$

für jedes α, β und für jedes $x \in R^n$ [vgl. zu x^β (2.14)].

Dann ist $\mathfrak{S} \supseteq D$, und aus der Konvergenz in D folgt die in \mathfrak{S} ; schließlich folgt, daß jede Distribution über \mathfrak{S} erst recht eine über D , also $\mathfrak{S}' \subseteq D'$, ist. Zu den Elementen aus \mathfrak{S}' gehören:

- Distributionen aus D' mit kompaktem Träger;
- Distributionen, die von lokalintegrablen Funktionen f erzeugt werden, wobei $\int |f(x)| (1 + |x|)^{-m} dx < \infty$ für ein gewisses $m \geq 0$ gilt (als f können somit z. B. $f(x) = \Theta(x)$, $f(x) \equiv 1$, $f(x) = e^{-|x|^2}$, $f(x) \in L^p(R^n)$ für $1 \leq p \leq \infty$, $f \in D$ oder $f(x)$ ein Polynom gewählt werden);

– jede Ableitung $D^\alpha L$ einer Distribution $L \in \mathcal{S}'$ und schließlich
 – jede Faltung $L_1 * L_2$ mit $L_1 \in \mathcal{S}'$, $\text{supp } L_2$ kompakt.
 (Jede Distribution aus \mathcal{S}' ist distributionelle Ableitung einer stetigen Funktion, die wie eine Potenz wächst [38, S. 112].)

Mit den Distributionen über \mathcal{S} , also den $L \in \mathcal{S}'$, werden die Fourier-Transformationen F ausgeführt. F bildet \mathcal{S}' auf \mathcal{S}' (linear und stetig) ab. Ebenso gilt das von F^{-1} . Die Distributionen aus \mathcal{S}' heißen auch „verallgemeinerte Funktionen schwachen Wachstums“.

Definition 4.7: Ist $L \in \mathcal{S}'$, so heißt FL **Fouriertransformierte** von L , und es gilt **D.4.7**

$$(FL)(\psi) = L(F\psi) \quad (\psi \in \mathcal{S}), \quad (4.21)$$

wobei

$$F\psi = \int_{\mathbb{R}^n} \psi(x) e^{i\langle \xi, x \rangle} dx \quad (\psi \in \mathcal{S})$$

aus \mathcal{S} ist¹⁾. F ist ein linearer stetiger Integraloperator über \mathcal{S} (und \mathcal{S}').

Eigenschaften: Es sei $L \in \mathcal{S}'$. Dann ist:

$$D^\alpha(FL) = F((ix)^\alpha L); \quad (4.22)$$

$$F(D^\alpha L) = (-i\xi)^\alpha F(L); \quad (4.23)$$

$$F\Theta = \pi \delta_0 + iP; \quad (4.24)$$

$$F\delta_0 = 1 \quad (\text{die von } f \equiv 1 \text{ erzeugte Distribution}); \quad (4.25)$$

$$FL\varphi(x + \hat{x}) = e^{i\langle \xi, \hat{x} \rangle} FL\varphi(x); \quad (4.26)$$

$$F1 = (2\pi)^n \delta_0, \quad \text{wobei } 1 \text{ die von } f(x) \equiv 1, x \in \mathbb{R}^n, \text{ erzeugte Distribution ist}; \quad (4.27)$$

$$F\delta_Q = e^{i\langle \xi, Q \rangle}, \quad \text{wobei } Q = (\hat{x}_1, \dots, \hat{x}_n)^T \text{ der charakteristische Punkt von } \delta_Q \text{ sei}; \quad (4.28)$$

$$F^{-1}(\psi) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \psi(\xi) e^{-i\langle \xi, x \rangle} d\xi = \frac{1}{(2\pi)^n} F(\psi)(-x); \quad (4.29)$$

$$F(\|x\|^{-2}) = 2\pi^2 \|\xi\|^{-1} \quad \text{bei } x, \xi \in \mathbb{R}^3; \quad (4.30)$$

$$F^{-1}(\|\xi\|^{-2}) = (4\pi \|x\|)^{-1} \quad \text{bei } x, \xi \in \mathbb{R}^3. \quad (4.31)$$

Wird schließlich die Distribution $L \in \mathcal{S}'$ von einer lokalintegrierbaren Funktion f erzeugt, so wird FL von Ff erzeugt.

Bemerkung 4.1: Ist G eine Grundleistung schwachen Wachstums (d. h. $G \in \mathcal{S}'$) von $P(D) = \sum_{\alpha} a_{\alpha} D^{\alpha}$, so gilt

$$P(D)G = \delta_0, \quad (4.32)$$

¹⁾ Manchmal (vgl. z. B. [36, S. 98] und (4.29)) steht $-i\langle \xi | x \rangle$ statt $i\langle \xi | x \rangle$ in der Definition von F und der Faktor $(2\pi)^{-n}$ bei F bzw. als $(2\pi)^{-n/2}$ bei F und F^{-1} .

²⁾ Also $F\left(\left(-\frac{1}{i}D\right)^{\alpha}L\right) = \xi^{\alpha}F(L)$, daher wird in der Physik statt D oft der Operator $-\frac{1}{i}D$ benutzt. Die Formel ergibt sich durch partielle Integration.

Fourier-Transformation liefert

$$F(P(D) \mathbf{G}) = \sum_{\alpha} a_{\alpha} (-i\xi)^{\alpha} F(\mathbf{G}) = F\delta_0 = \mathbf{I}, \quad (4.33)$$

$$P(-i\xi) F(\mathbf{G}) = \mathbf{I}. \quad (4.34)$$

Diese Gleichung für $F(\mathbf{G})$ ist in \mathfrak{S}' stets lösbar [38]. Es gibt also zu jedem Differentialausdruck $P(D)$ mit konstanten Koeffizienten eine Grundlösung schwachen Wachstums! Die Konstruktion der Lösung hängt von der Lage der Nullstellen von P ab. \mathbf{G} selbst ist über die inverse Transformation F^{-1} durch $F^{-1}(F(\mathbf{G}))$ prinzipiell bestimmbar [38, S. 134].

In 1.2.2. hatten wir für die Differentialgleichung

$$P(D) J = L\ddot{J} + \frac{1}{C} J = 0 \quad (4.35)$$

die Greensche Funktion zu homogenen Anfangsbedingungen gefunden:

$$J(t; t_0) = \sqrt{\frac{C}{L}} \Theta(t - t_0) \sin \frac{t - t_0}{\sqrt{LC}} = G(t - t_0). \quad (4.36)$$

Diese Funktion erzeugt wegen der Hinweise unter (4.20) eine Distribution \mathbf{G} aus \mathfrak{S}' , da $(t - t_0 = \tau$ gesetzt)

$$\int_{\mathbb{R}^1} |G(\tau)| (1 + |\tau|)^{-2} d\tau < \infty \quad (4.37)$$

gilt. \mathbf{G} löst (4.34) für $P(D)$ in (4.35).

Wir wollen jetzt umgekehrt eine Grundlösung nach (4.34) konstruieren. Dazu betrachten wir den Laplace-Operator $P(D) = \Delta$ in 3 Dimensionen. Eine Grundlösung \mathbf{G} muß also der Gl. (4.34) genügen:

$$P(-i\xi) F(\mathbf{G}) = \mathbf{I}, \quad (4.34')$$

wobei wegen $P(D) = \Delta$ gelten muß

$$P(-i\xi) = (-i\xi_1)^2 + (-i\xi_2)^2 + (-i\xi_3)^2 = -\|\xi\|^2. \quad (4.38)$$

Zunächst ist $F(\mathbf{G})$ gesucht, also ist die Gleichung

$$-\|\xi\|^2 F(\mathbf{G}) = \mathbf{I} \quad (4.39)$$

nach $F(\mathbf{G})$ aufzulösen. (4.39) ist eine Gleichung vom Typ (4.7). Da $-\|\xi\|^{-2}$ eine im \mathbb{R}^3 integrierbare Funktion ist (dies gilt auch für mehr als drei Raumdimensionen, nicht aber für $n = 2$, dort hat die Grundlösung mit $\frac{-1}{2\pi} \ln \|x\|$ gegenüber (4.43) für $n = 3$ auch eine ganz andere Gestalt; die Lösung von (4.39) hängt dann mit der singulären Distribution \mathbf{P} zusammen), gilt für die von $-\|\xi\|^{-2}$ erzeugten Distribution

$$\begin{aligned} -\|\xi\|^2 (-\|\xi\|^{-2}, \psi) &= (-\|\xi\|^{-2}, -\|\xi\|^2 \psi) \\ &= (\mathbf{I}, \psi) = \mathbf{I}, \end{aligned} \quad (4.40)$$

folglich löst $F(\mathbf{G}) = (-\|\xi\|^{-2}, \psi)$ die Gl. (4.39); daher

$$\mathbf{G} = F^{-1}(-\|\xi\|^{-2}, \psi) = (F^{-1}(-\|\xi\|^{-2}), \psi), \quad (4.41)$$

und aus (4.31) folgt

$$G = \left(-\frac{1}{4\pi \|x\|}, \psi \right), \quad (4.42)$$

daher ist eine Grundlösung gefunden; sie wird erzeugt von

$$G_3(x) = -\frac{1}{4\pi} \frac{1}{\|x\|} \quad (x \in R^3). \quad (4.43)$$

Dies ist die bekannte Fundamentallösung des Laplace-Operators Δ im R^3 .

Bemerkung 4.2: Eine lineare stetige Abbildung A eines Hilbertraumes E_1 auf einen Hilbertraum E_2 mit der Eigenschaft $\langle f | g \rangle_{E_1} = \langle Af | Ag \rangle_{E_2}$ für $f, g \in E_1$ heißt **unitär**. F und F^{-1} sind als Abbildungen mit $D(F) = L^2(R^n)$, $D(F^{-1}) = L^2(R^n)$ unitäre Abbildungen [36].

Dies kann man ausnutzen, um eine Variante des *Sobolewschen Einbettungssatzes* zu beweisen. Es ist im Sinne der Ersetzung von L^2 -Funktionen durch Distributionen aus \mathcal{D}' :

$$L^2(R^n) \subseteq \mathcal{D}'(R^n), \quad (4.44)$$

damit erst recht

$$W^{k,2}(R^n) \subseteq \mathcal{D}'(R^n), \quad (4.45)$$

wobei $W^{k,2}$ in 2.2.3. definiert wurde. F ist dann eine Abbildung von $W^{k,2}(R^n)$ auf $L_{2,p(x)}$, wobei $p(x) = (1 + \|x\|^{2k})^{1/2}$ und $L_{p(x)}^2(R^n) = \{f(x) \mid p(x)f(x) \in L^2(R^n)\}$ mit dem Skalarprodukt

$$\langle f | g \rangle_{L_{p(x)}^2} = \langle pf | pg \rangle_{L^2} \quad (4.46)$$

ist, und F ist sogar unitär, wenn in $L_{p(x)}^2$ zu einer äquivalenten Norm übergegangen wird. Ist $\mathcal{C}^l(R^n)$ die Vervollständigung von $\mathcal{C}^\infty(R^n)$ in der Norm

$$\|f\|_{\mathcal{C}^l} = \sum_{|\alpha| \leq l} \sup_{x \in R^n} |D^\alpha f(x)|, \quad (4.47)$$

so gilt

Satz 4.2 (Sobolewscher Einbettungssatz) [36]: Ist $k \geq 0$ ganz, $l > \frac{n}{2}$ eine natürliche Zahl, dann ist der Hilbertraum $H_1 = W^{l+k,2}(R^n)$ stetig in den Banachraum $\mathcal{C}^k(R^n)$ eingebettet. Es gibt daher eine Zahl c , so daß für alle $f \in W^{l+k,2}(R^n)$ gilt

$$\|f\|_{\mathcal{C}^k} \leq c \|f\|_{W^{l+k,2}}. \quad (4.48)$$

Beispiel 4.7: $n = 3$, $k = 0$, $l = 2$. Dann ist $l > \frac{n}{2}$, also sind die verallgemeinerten Ableitungen nullter Ordnung (d. h. die Funktionen f aus $W^{2,2}(R^3)$ selbst) stetige Funktionen (im Sinne der Gleichheit fast überall mit einer stetigen Funktion).

4.2. Differentialrechnung und Anwendungen

4.2.1. Ableitungsbegriffe

Die wichtigsten Sätze der gewöhnlichen Differentialrechnung (s. Bd. 4) beruhen auf dem Begriff des totalen Differentials: Eine Funktion f mit Definitionsbereich $D(f) = R^n$ heißt im Punkt \hat{x} total differenzierbar, falls

$$f(\hat{x} + h) - f(\hat{x}) = \nabla f(\hat{x})^T h + o(\|h\|) \quad (h \in R^n) \quad (4.49)$$

gilt. Dabei bedeutet $o(\|h\|)$, daß gilt

$$\lim_{\|h\| \rightarrow 0} \|h\|^{-1} o(\|h\|) = 0. \quad (4.50)$$

D.4.8 Definition 4.8: Der lineare Anteil des Funktionswertzuwachses in (4.49), der durch

$$df = \nabla f(\hat{x})^T h = \sum_{i=1}^n \frac{\partial}{\partial x_i} f(\hat{x}) h_i \quad (4.51)$$

gegeben ist, heißt das **totale oder Fréchet-Differential** von f an der Stelle \hat{x} .

Der Vektor $\nabla f(\hat{x})$ „erzeugt“ diese lineare Funktion. Er heißt (Fréchet-) Gradient von f an der Stelle \hat{x} . (4.50) heißt, daß das Fréchet-Differential den Funktionswertzuwachs $f(\hat{x} + h) - f(\hat{x})$ „von höherer als erster Ordnung“ bezüglich der Norm des Argumentzuwachses approximiert.

Bemerkung 4.3: Ist der Definitionsbereich $D(f)$ nicht der Gesamtraum, so muß stets $(\hat{x} + h) \in D(f)$ sein in (4.49), (4.50). Die eindeutige Bestimmtheit des Gradienten $\nabla f(\hat{x})$ ist noch gesichert, wenn der Definitionsbereich $D(f)$ die Form

$$\hat{x} + M = \{\hat{x} + y \mid y \in M\} \quad (4.52)$$

hat, wobei \hat{x} fest ist und M eine offene Nullumgebung ist (d. h., eine offene Menge, die den Nullpunkt enthält).

Es sei nun B ein reeller Banachraum: $B = (B, \|\cdot\|)$, und f ein über B definiertes Funktional; ist f ein über B lineares stetiges Funktional (s. Def. 2.29), so schreiben wir statt $f(x)$ auch (f, x) in Anlehnung an die Schreibweise von Distributionen [s. (4.4)] oder die Schreibweise von linearen stetigen Funktionalen als Skalarprodukte (s. Satz 2.26) in Hilberträumen.

D.4.9 Definition 4.9: f heißt im Punkte \hat{x} **Fréchet-differenzierbar**, falls ein (von \hat{x} abhängiges) Element $f^*(\hat{x})$ aus dem Dualraum B' existiert, so daß gilt

$$f(\hat{x} + h) = f(\hat{x}) + (f^*(\hat{x}), h) + o(\|h\|) \quad (h \in B). \quad (4.53)$$

Die Größe $(f^*(\hat{x}), h)$ stellt den linearen Anteil in der Zerlegung (4.53) dar und wird Fréchet-Differential von f an der Stelle \hat{x} genannt. $f^*(\hat{x}) \in B'$ heißt Fréchet-Gradient oder Fréchet-Ableitung von f an der Stelle \hat{x} . Wir behalten auch die im R^n übliche Bezeichnung bei: $f^*(\hat{x}) = \nabla f(\hat{x})$. Gilt Bem. 4.3, so ist in (4.53) $h \in M \subset B$ zu nehmen.¹⁾

Beispiel 4.8: Es sei H ein reeller Hilbertraum, A ein selbstadjungierter (s. Def. 3.17) linearer stetiger Operator von H in sich, g ein festes Element aus H . Dann sei

$$f(x) = \langle Ax \mid x \rangle - 2\langle g \mid x \rangle. \quad (4.54)$$

f ist überall auf dem Hilbertraum definiert. Wir wollen prüfen, ob das Funktional f Fréchet-differenzierbar ist. Es sei \hat{x} ein Element aus H . Wir bilden die Zerlegung (4.53):

$$\begin{aligned} f(\hat{x} + h) &= \langle A(\hat{x} + h) \mid \hat{x} + h \rangle - 2\langle g \mid \hat{x} + h \rangle \\ &= \langle A\hat{x} \mid \hat{x} \rangle - 2\langle g \mid \hat{x} \rangle + \langle A\hat{x} \mid h \rangle + \langle A\hat{x} \mid h \rangle - 2\langle g \mid h \rangle + \langle Ah \mid h \rangle, \end{aligned}$$

und weiter gilt insbesondere, da A selbstadjungiert ist,

$$f(\hat{x} + h) = f(\hat{x}) + 2\langle A\hat{x} \mid h \rangle - 2\langle g \mid h \rangle + \langle Ah \mid h \rangle. \quad (4.55)$$

¹⁾ Das Fréchet-Differential kann man natürlich für jedes $h \in B$ bilden.

Die Schwarzzsche Ungleichung und die Beschränktheit von A ergeben

$$\langle Ah | h \rangle \leq \|A\| \cdot \|h\|^2, \quad (4.56a)$$

folglich ist

$$2\langle A\hat{x} - g | h \rangle \quad (4.56b)$$

der lineare Anteil des Zuwachses $f(\hat{x} + h) - f(\hat{x})$, also ist f in \hat{x} Fréchet-differenzierbar, und es gilt

$$f^*(\hat{x}) = 2\langle A\hat{x} - g \rangle. \quad (4.56c)$$

Neben der Fréchet-Differenzierbarkeit studiert man noch einen weiteren Ableitungsbegriff: Wir wollen dazu in (4.53) für h setzen $h = \alpha s$, wobei α ein reeller Parameter und s ein festes Element aus B sei. Es sei wieder f ein Funktional über B . Ist f Fréchet-differenzierbar an der Stelle \hat{x} , so folgt aus (4.53)

$$\lim_{\alpha \rightarrow 0} \frac{f(\hat{x} + \alpha s) - f(\hat{x})}{\alpha} = (\nabla f(\hat{x}), s) + \lim_{\alpha \rightarrow 0} \frac{1}{\alpha} o(\|\alpha s\|) = (\nabla f(\hat{x}), s), \quad (4.57)$$

und dies ist gerade die Ableitung von f an der Stelle \hat{x} bezüglich der Geraden in s -Richtung.

Definition 4.10: Ist f ein Funktional mit $D(f) = \hat{x} + M = \{\hat{x} + s \mid s \in M\}$, M lineare Teilmenge von B , und existiert **D.4.10**

$$\lim_{\alpha \rightarrow 0} \frac{1}{\alpha} (f(\hat{x} + \alpha s) - f(\hat{x})) \stackrel{\text{def}}{=} \delta f(\hat{x}, s) \quad (4.58)$$

für jedes $s \in M$, so heißt $\delta f(\hat{x}, s)$ das Gateaux-Differential (1. Variation) von f an der Stelle \hat{x} in Richtung s ; f heißt dann Gateaux-differenzierbar an der Stelle \hat{x} .

Bei (4.57) folgte: ist f an \hat{x} Fréchet-differenzierbar, so auch Gateaux-differenzierbar. Umgekehrt gilt (mit $D(f) \equiv B$)

Satz 4.3: Existiert das Gateaux-Differential in einer Normumgebung von \hat{x} in B und ist $\delta f(x, s)$ gleichmäßig stetig in x und stetig bezüglich s , so stimmt es mit dem Fréchet-Differential von f an der Stelle \hat{x} überein [24, S. 310]. **S.4.3**

Beispiel 4.9: Ist A in (4.54) nur linear, selbstadjungiert und nicht beschränkt, so ist sein Definitionsbereich $D(A) \subsetneq H$ eine lineare in H dichte Teilmenge (s. Def. 3.17) und $D(A) = D(f)$. Wir wollen das Gateaux-Differential berechnen. Es ist für $\hat{x} \in D(A)$, $s \in D(A)$ wegen

$$\begin{aligned} & \frac{f(\hat{x} + \alpha s) - f(\hat{x})}{\alpha} \\ &= \frac{1}{\alpha} (2\langle A\hat{x} | \alpha s \rangle - 2\langle g | \alpha s \rangle + \alpha^2 \langle As | s \rangle) \\ & \delta f(\hat{x}, s) = 2\langle A\hat{x} - g | s \rangle. \end{aligned} \quad (4.59)$$

In diesem Falle ist das Gateaux-Differential $\delta f(\hat{x}, s)$ ein bezüglich s lineares stetiges Funktional. Dann heißt das „erzeugende“ Element $2\langle A\hat{x} - g \rangle$ der Gateaux-Gradient. Er sieht freilich genau so aus wie in (4.56c), aber diesmal ist die Gültigkeit der Zerlegung (4.53) nicht gegeben, wir haben nur aus (4.59) den Gradienten formal so abgelesen wie in (4.53). [(4.56a) gilt für A jetzt nicht.]

4.2.2. Anwendungen der Ableitungsbegriffe

4.2.2.1. Beziehungen zur Variationsrechnung

An einem wichtigen Beispiel soll erläutert werden, wie in der Variationsrechnung notwendige Bedingungen für relative Extremalstellen eines Funktionals f hergeleitet werden können (vgl. auch Lemma 5.1–5.3).

Über dem Einheitskreis $K = \{(x, y) \mid x^2 + y^2 \leq 1\}$ werde der reelle Hilbertraum $H = L^2_{\mathbf{R}}(K)$ betrachtet. Die Menge der über K zweimal stetig differenzierbaren Funktionen u , die auf dem Kreisrand ∂K gleich einer dort vorgegebenen Funktion $\varphi \in C^3_{\mathbf{R}}(\partial K)$ sind, ist dann eine Teilmenge \mathfrak{B} von H . Für jedes u aus \mathfrak{B} gilt also, wenn der Kreisrand durch Polarkoordinaten $r = 1$, $0 \leq \tau \leq 2\pi$ beschrieben wird:

$$u(\cos \tau, \sin \tau) = \varphi(\tau), \quad 0 \leq \tau \leq 2\pi. \quad (4.60)$$

Wir betrachten das Funktional f (Dirichlet-Integral):

$$f(u) = \iint_K (u_x^2 + u_y^2) \, dx \, dy, \quad D(f) = \mathfrak{B}. \quad (4.61)$$

Wir stellen die Aufgabe, relative (= lokale) Extremalstellen von f zu charakterisieren. Eine Stelle $\hat{u} \in \mathfrak{B}$ heißt relative Extremalstelle von f über \mathfrak{B} , wenn bei hinreichend kleinem $\varepsilon > 0$ für alle $u \in \mathfrak{B}$ mit $\|u - \hat{u}\|_H < \varepsilon$ gilt

$$f(u) \geq f(\hat{u}). \quad (4.62)$$

In der Variationsrechnung betrachtet man diese zu \hat{u} benachbarten u in der Form $u = \hat{u} + \alpha\eta$ mit $\eta \in M = \{\eta \in C^2_{\mathbf{R}}(K) \mid \eta = 0 \text{ auf } \partial K\}$ und für $|\alpha|$ hinreichend klein; „ \hat{u} wird variiert“. In der Tat ist bei vorgegebenem $\varepsilon > 0$ und $\eta \neq 0 \in H$

$$\|u - \hat{u}\| = \|\hat{u} + \alpha\eta - \hat{u}\| = |\alpha| \cdot \|\eta\| < \varepsilon,$$

falls nur $|\alpha| < \|\eta\|^{-1} \varepsilon$, und für solche α lautet (4.62)

$$f(\hat{u} + \alpha\eta) \geq f(\hat{u}), \quad \eta \in M. \quad (4.63)$$

M ist eine lineare Teilmenge von H . Ist nun \hat{u} eine relative Extremalstelle von f über \mathfrak{B} , so muß sie notwendig (4.63) erfüllen. Wir fassen daher die Menge $\hat{u} + M \subseteq H$ als neuen Definitionsbereich von f auf. Ist f an der Stelle \hat{u} Gateaux-differenzierbar in Richtung η (für jedes $\eta \in M$), so muß

$$\delta f(\hat{u}, \eta) = 0, \quad \eta \in M, \quad (4.64)$$

sein; denn wir brauchen nur die Vorzeichen des Quotienten

$$\frac{1}{\alpha} (f(\hat{u} + \alpha\eta) - f(\hat{u})), \quad \eta \text{ fest}, \quad (4.65)$$

zu studieren: wegen (4.63) ist (4.65) größer gleich null für (kleine $|\alpha|$ bei) $\alpha > 0$, kleiner gleich null für (kleine $|\alpha|$ und) $\alpha < 0$; da der Grenzwert $\delta f(\hat{u}, \eta)$ für $\alpha \rightarrow 0$ existiert (Gateaux-Differenzierbarkeit) und Ungleichungen beim Grenzübergang erhalten bleiben (s. Bem. 2.1), muß

$$0 \leq \delta f(\hat{u}, \eta) \leq 0 \quad (4.66)$$

sein; das ist gerade (4.64). (4.64) heißt auch: die **erste Variation** (erzeugt durch den Übergang von \hat{u} zu $\hat{u} + \alpha\eta$) von f an der Stelle \hat{u} muß verschwinden.

Wir berechnen jetzt das Gateaux-Differential $\delta f(\hat{u}, \eta)$ von f an der Stelle \hat{u} in η -Richtung:

$$\frac{1}{\alpha} \left[\iint_K ((\hat{u}_x + \alpha\eta_x)^2 + (\hat{u}_y + \alpha\eta_y)^2) \, dx \, dy - \iint_K (\hat{u}_x^2 + \hat{u}_y^2) \, dx \, dy \right] \quad (4.67)$$

$$= \iint_K [2(\hat{u}_x\eta_x + \hat{u}_y\eta_y) + \alpha(\eta_x^2 + \eta_y^2)] \, dx \, dy \quad (4.68)$$

und wegen der 1. Greenschen Formel (vgl. Bd. 8, S. 97),

$$\iint_K (\hat{u}_x v_x + \hat{u}_y v_y) dx dy = - \iint_K (v \Delta \hat{u}) dx dy + \int_{\partial K} v \frac{\partial \hat{u}}{\partial \vec{n}} d\tau, \quad (4.69)$$

wobei \vec{n} die (normierte) Außennormale bezüglich ∂K ist, folgt aus (4.68), wobei $v = \eta$ gesetzt werde,

$$\delta f(\hat{u}, \eta) = \lim_{\alpha \rightarrow 0} \frac{1}{\alpha} (f(\hat{u} + \alpha \eta) - f(\hat{u})) = -2 \iint_K \eta \Delta \hat{u} dx dy \quad (\eta \in M). \quad (4.70)$$

Dabei verschwand das Randintegral, weil η auf dem Rand verschwindet, und die Normalableitung von \hat{u} (unter unseren Voraussetzungen) existiert. Also existiert das Gateaux-Differential. Wegen (4.66) muß es verschwinden:

$$0 = \delta f(\hat{u}, \eta) = -2 \iint_K \eta \Delta \hat{u} dx dy \quad (\eta \in M).$$

Da nun M ein in $H = L^2_{\mathbf{R}}(K)$ sogar dichter (vgl. Satz 2.17) linearer Teilraum ist, muß $\Delta \hat{u}$ selbst verschwinden.¹⁾ Damit haben wir folgende *notwendige* Bedingung für \hat{u} erhalten, welches Lösung der Extremalaufgabe (4.62) war:

$$\Delta \hat{u} = 0. \quad (4.71)$$

Eine Lösung \hat{u} der Extremalaufgabe (der Klasse $C^2_{\mathbf{R}}(K)$) ist also *notwendig* Lösung des Dirichlet-Problems:

$$\begin{aligned} \Delta \hat{u} &= 0 \quad \text{im Kreisinnern,} \\ \hat{u}(x, y)|_{\partial K} &= \varphi(\tau) \quad \text{auf dem Kreisrand.} \end{aligned} \quad (4.72)$$

Gleichung (4.71) heißt **Eulersche Gleichung** zum Variationsproblem (4.62). (4.72) sind für eine relative Extremalstelle notwendige Bedingungen. Bei $\varphi \in C^2_{\mathbf{R}}(\partial K)$ gilt auch: das Dirichlet-Problem (4.72) ist eindeutig lösbar, die Lösung gehört zu $C^2_{\mathbf{R}}(K)$ und ist die eindeutige (sogar globale) Lösung des Variationsproblems, f von (4.61) über $D(f) = \mathfrak{B}$ zu minimieren:

$$f(u) = \text{Min!} \quad (u \in \mathfrak{B}). \quad (4.73)$$

Man kommt bei Verwendung anderer Beweismittel mit geringeren Regularitätsvoraussetzungen aus: **Dirichletsches Prinzip** (für den Einheitskreis: g sei über K stetig und in $K \setminus \partial K$ (stückweise) stetig differenzierbar, so daß $f(g)$ [s. (4.61)] existiert. Die Klasse aller Funktionen Φ , die in K stetig, in $K \setminus \partial K$ (stückweise) stetig differenzierbar sind und dieselben Randwerte wie g haben, enthält genau eine Funktion \hat{u} , so daß gilt

$$f(\hat{u}) \leq f(\Phi). \quad (4.74)$$

Diese Funktion \hat{u} löst das Dirichlet-Problem für $\Delta u = 0$ mit $u = g$ auf ∂K ; \hat{u} liegt dann in $C^2(K \setminus \partial K)$ (sogar in $C^\infty(K \setminus \partial K)$). Lit. s. z. B. [12], [28], [18], [41/III].

¹⁾ Das Funktional $\delta f(\hat{u}, \eta)$ von (4.70) ist in η linear und stetig auf M . Da M dicht liegt in H ist jedes Element von H Grenzwert einer Folge aus M . Somit kann $\delta f(\hat{u}, \eta)$ auf ganz H mit Wert 0 stetig fortgesetzt werden. Das nach dem Satz 2.26 von Riesz eindeutig existierende erzeugende Element $f^* \in H$ kann also nur das Nullelement $0 \in H$ sein. Das ist (4.71), da nach (4.70) das erzeugende Element $-2\Delta \hat{u}$ heißt; da \hat{u} in $C^2_{\mathbf{R}}(K)$ liegen und Randbedingungen genügen sollte, folgt daraus (4.72).

4.2.2.2. Beziehungen zur konvexen Analysis

Analysis beinhaltet das Studium von Räumen und von Abbildungen zwischen solchen Räumen unter Verwendung von Grenzwertbegriffen. Man spricht von konvexer Analysis, wenn bei der Gewinnung von Aussagen Konvexitätseigenschaften von Mengen oder Abbildungen eine wichtige Rolle spielen. Dies ist z. B. bei Fixpunktsätzen der Fall, die vom Schauderschen Typ sind:

S.4.4 Satz 4.4 (Schauderscher Fixpunktsatz [17], [30]: *Es sei E ein normierter Raum und $M \subseteq E$ eine nichtleere abgeschlossene und konvexe Teilmenge von E . Es sei $T: M \rightarrow E$ eine stetige Abbildung mit $T[M] \subseteq M$, und der Abschluß $\overline{T[M]}$ der Bildmenge $T[M]$ sei kompakt (s. Def. 2.7). Dann hat T (mindestens) einen Fixpunkt, d.h., es gibt ein $x \in M$ mit $Tx = x$.*

Man kann sich für E etwa den R^2 (oder den Banachraum $C[a, b]$) und für M die Einheitskreisscheibe im R^2 (oder eine kompakte konvexe Menge in $C[a, b]$, vgl. Satz 2.13) denken. Dieser Satz wird u. a. zur Lösung von Randwertproblemen ausgenutzt (vgl. z. B. [41/II]).

In der mathematischen Optimierung in allgemeinen Räumen spielt die konvexe Analysis eine besonders wichtige Rolle, denn für konvexe Optimierungsaufgaben können Hauptziele der Optimierungstheorie wie das Aufstellen von dualen Aufgaben, von Optimalitätsbedingungen, von Lösungsverfahren usw. recht zufriedenstellend gelöst werden (siehe z. B. Bd. 15 im R^n oder [41/III], [13]), während für nicht mehr konvexe Aufgaben sich die Resultate oft stark „am konvexen Fall orientieren“ (s. z. B. [45] oder [48]). Dabei spielen – als Gegenstände der konvexen Analysis – Trennungssätze für konvexe Mengen (s. 2.3.3.), Differenzierbarkeitsbegriffe (oft noch allgemeinere als oben behandelt) und (vor allem in der Steuereorie, vgl. [23]) die $W^{p,k}$ -Räume die Hauptrolle.

Wir wollen auf einige Beziehungen zwischen Fréchet-Differenzierbarkeit und „stark- und beschränkt-konvexen“ Funktionalen eingehen, die bei Abstiegsverfahren der Optimierung und bei der Lösung von Operatorgleichungen (vgl. auch [30]; s. 5.3. und z. B. in [18], [13]) wichtig sind.

Ein über einer offenen Menge D eines reellen Hilbertraumes gegebenes Funktional $f(x)$ heißt **stark konvex** (mit der Schranke $m > 0$) und **beschränkt konvex** (mit der Schranke $M > 0$), falls $\forall \bar{x}, \hat{x} \in D \subseteq H$ und $\forall \lambda_1, \lambda_2 \geq 0$ mit $\lambda_1 + \lambda_2 = 1$ gilt

$$\begin{aligned} \frac{1}{2} m \lambda_1 \lambda_2 \| \bar{x} - \hat{x} \|^2 &\leq \lambda_1 F(\bar{x}) + \lambda_2 F(\hat{x}) - F(\lambda_1 \bar{x} + \lambda_2 \hat{x}) \\ &\leq \frac{1}{2} M \lambda_1 \lambda_2 \| \bar{x} - \hat{x} \|^2. \end{aligned} \quad (4.75)$$

F ist damit erst recht streng konvex. F sei fernerhin stetig. Diese stark und beschränkten konvexen stetigen Funktionale lassen sich wie folgt durch Fréchet-Ableitungen charakterisieren:

S.4.5 Satz 4.5: *Das auf einer offenen konvexen Menge D erklärte stetige (reelle) Funktional F ist genau dann stark und beschränkt konvex, wenn F in allen Punkten von D eine Fréchet-Ableitung $\nabla F(x)$ besitzt und für diese die Monotonie-Bedingung gilt: (siehe z. B. in [13])*

$$m \| \bar{x} - \hat{x} \|^2 \leq \langle \nabla F(\bar{x}) - \nabla F(\hat{x}) | \bar{x} - \hat{x} \rangle \leq M \| \bar{x} - \hat{x} \|^2, \quad (\bar{x}, \hat{x} \in D). \quad (4.76)$$

Die Bedeutung dieser Klasse von konvexen Funktionalen liegt u. a. darin, daß der „Durchhang“ [d. i. die Differenz in der Mitte der Ungleichungskette (4.75)] nach oben und unten durch den Abstand $\|x - \hat{x}\|$ abgeschätzt wird oder (unter Benutzung des folgenden Satzes 4.6) daß der Graph des Funktional $z = F(x) - F(\hat{x})$ im Raum $R^1 \times H$ zwischen den beiden, sich in $x = \hat{x}$ berührenden Hyperparaboloiden

$$z = \langle \nabla F(\hat{x}) | x - \hat{x} \rangle + \frac{1}{2}m\|x - \hat{x}\|^2, \quad (4.77)$$

$$\bar{z} = \langle \nabla F(\hat{x}) | x - \hat{x} \rangle + \frac{1}{2}M\|x - \hat{x}\|^2 \quad (4.78)$$

liegt. Wenn also F auf ganz H definiert ist, so wächst es für $\|x - \hat{x}\| \rightarrow \infty$ stärker als z (stärker als linear) und „zwischen“ z und \bar{z} gegen $+\infty$.

Satz 4.6: *Das auf einer offenen konvexen Menge $D \subseteq H$ erklärte stetige Funktional F ist genau dann stark und beschränkt konvex, wenn F Fréchet-differenzierbar ist und folgende Bedingung gilt:* **S.4.6**

$$\begin{aligned} \frac{1}{2}m\|\hat{x} - \bar{x}\|^2 &\leq F(\hat{x}) - F(\bar{x}) - \langle \nabla F(\bar{x}) | \hat{x} - \bar{x} \rangle \\ &\leq \frac{1}{2}M\|\hat{x} - \bar{x}\|^2 \quad (\hat{x}, \bar{x} \in D). \end{aligned} \quad (4.79)$$

4.3. Anwendungen von Fixpunktsätzen

4.3.1. Gleichgewichtspunkte und Fixpunkte in Ökonomie und Spieltheorie

Bei der mathematischen Behandlung ökonomischer Modelle ergibt sich oft folgende Situation: Gegeben seien ein kompakter konvexer Bereich $Z \subseteq R^n$ (zulässiger Bereich, nichtleer) und n beteiligte Parteien (Spieler), denen je eine Funktion f_1, \dots, f_n zugeordnet ist (Auszahlungsfunktionen), die Z in R abbildet. Der ν -te Spieler ($\nu = 1, \dots, n$) möchte seine Auszahlung $f_\nu(z)$, $z \in Z$, maximieren, hat aber nur Einfluß auf die ν -te Komponente des Vektors $z \in Z$. Er muß sich daher mit der Aufgabe

$$\begin{aligned} f_\nu(z) = f_\nu(x_1, \dots, x_{\nu-1}, \xi, x_{\nu+1}, \dots, x_n) &= \max!, \\ z = (x_1, \dots, x_{\nu-1}, \xi, x_{\nu+1}, \dots, x_n) &\in Z \end{aligned} \quad (4.80)$$

beschäftigen, in der die restlichen $n - 1$ Komponenten des Vektors z fest sind. Wird also ein Punkt z aus Z fest angenommen (damit ist jedem Spieler eine Möglichkeit der Wahl seiner Komponente von z vorgeschlagen), so wird jeder Spieler die Aufgabe (4.80) lösen wollen. Sind die f_ν ($\nu = 1, \dots, n$) stetig, so existiert auch für jedes ν eine Lösung, i. allg. eine Menge F_ν von Lösungen, die von dem vorgegebenen z abhängt! Es wird somit jedem $z \in Z$ ein Vektor $(F_1(z), \dots, F_n(z))$ zugeordnet, dessen ν -te Komponente die Lösungen von (4.80) enthält. Wir wollen den Vektor der Lösungsmengen mit $F(z)$ bezeichnen. Es wird dann jedem $z \in Z$ das Bild $F(z)$ zugeordnet, es ist eine Teilmenge des R^n .

Wenn es nun gelingt, ein z^* zu finden, daß gilt

$$z^* \in F(z^*), \quad (4.81)$$

d. h. daß z_1^* ein Element von $F_1(z^*)$, z_2^* ein Element von $F_2(z^*)$ ist usw., so ist eine solche Wahl der von jedem Spieler beeinflussbare Komponente von z vorgeschlagen, daß die Auszahlungsfunktion f_ν für jedes ν im Sinne von (4.80) maximiert wird.

z^* heißt dann ein **Gleichgewichtspunkt** des n -Personenspieles, welches durch f_1, \dots, f_n und $Z \subseteq R^n$ beschrieben ist (s. auch Bd. 21/1).

Die Existenz eines solchen Gleichgewichtspunktes kann durch den Fixpunktsatz von Kakutani gesichert werden, wenn F gewisse Voraussetzungen erfüllt.

D.4.11 Definition 4.11: Es sei Z eine kompakte nichtleere Teilmenge des R^n und f eine Abbildung von Z in die Menge 2^Z der Teilmengen von Z . Für jedes $z \in Z$ sei die Menge $f(z)$ abgeschlossen in R^n . Dann heißt f stetig, falls die Menge (der Graph von f)

$$G(f) = \{(z, y) \mid z \in Z, y \in f(z)\} \quad (4.82)$$

abgeschlossen in $R^n \times R^n$ ist.

S.4.7 Satz 4.7 (Fixpunktsatz von Kakutani) [41/I]: Es sei F eine stetige (s. Def. 4.11) (Punkt-Mengen-) Abbildung der kompakten, konvexen nichtleeren Menge $Z \subseteq R^n$ in 2^Z (Menge aller Teilmengen von Z). Für jedes $z \in Z$ sei $F(z)$ konvex, abgeschlossen und nichtleer. Dann besitzt F einen Fixpunkt z^* ; d. h., es gibt ein $z^* \in Z$ mit $z^* \in F(z^*)$.

In [26] wird das ökonomische Modell von Arrow und Debrau mit Satz 4.7 behandelt.

Bemerkung 4.4: Für die Anwendungen ist neben der Existenz (mindestens) eines Gleichgewichtspunktes (Fixpunktes) die numerische Bestimmung von Bedeutung. Während bei Anwendbarkeit des Fixpunktsatzes von Banach (s. Satz 1.2) eine iterative Berechnung des (dann eindeutig bestimmten) Fixpunktes möglich ist (s. 4.3.2.), sind für Fixpunktsätze vom Schauderschen Typ (Satz 4.4, vgl. auch [41/I]) erst in jüngster Zeit konstruktive Beweise und iterative Berechnungsmöglichkeiten gefunden worden, die auf Lemke und Scarf zurückgehen, vgl. z. B. [25].

4.3.2. Banachscher Fixpunktsatz und zugehöriges Iterationsverfahren

In 1.2.4. hatten wir den Banachschen Fixpunktsatz (Satz 1.2) in Zusammenhang mit einem linearen Gleichungssystem (in der Form $x = Mx + a$, x und a aus R^N) genannt. Weitere Anwendungen findet man z. B. in [41/I]. Wir stellen jetzt den Beweis dieses Satzes dar, da er konstruktiv ist, d. h. ein konvergentes Verfahren [s. (4.84)] zur Bestimmung des Fixpunktes liefert, der unter den Bedingungen des Satzes existiert. Es sei d die Metrik eines vollständigen metrischen Raumes E und A eine Abbildung, die eine fest vorgegebene abgeschlossene Teilmenge $E_0 \subseteq E$ in sich abbildet und die auf E_0 kontrahierend ist:

$$d(A(x), A(y)) \leq kd(x, y), \quad (x, y \in E_0), \quad 0 < k < 1. \quad (4.83)$$

Wir untersuchen, unter welchen Bedingungen die Gleichung $x = A(x)$ für $x \in E_0$ durch sukzessive Approximation

$$x_0 \in E_0, \quad x_{n+1} = A(x_n), \quad n \geq 0, \text{ ganz}, \quad (4.84)$$

gelöst werden kann. Die Antwort gibt der

S.4.8 Satz 4.8 (Fixpunktsatz von Banach): Eine kontrahierende Abbildung A einer nicht-leeren abgeschlossenen Teilmenge E_0 (aus E) in sich hat genau einen Fixpunkt x^* . Die Folge $\{x_n\}$ mit $x_0 \in E_0$, $x_{n+1} = A(x_n)$ konvergiert gegen x^* für jedes $x_0 \in E_0$.

Beweis. Die Folge $\{x_n\}$ ist eine Cauchy-Folge (s. Def. 2.6). Denn es ist, indem abwechselnd (4.84) und (4.83) verwendet werden,

$$\begin{aligned} d(x_n, x_{n+1}) &= d(A(x_{n-1}), A(x_n)) \leq kd(x_{n-1}, x_n) \\ &= kd(A(x_{n-2}), A(x_{n-1})) \leq k^2 d(x_{n-2}, x_{n-1}) \leq \dots \\ &\dots = k^{n-1} d(A(x_0), A(x_1)) \leq k^n d(x_0, x_1). \end{aligned} \quad (4.85)$$

Für $0 \leq n < m$ ist dann (Dreiecksungleichung)

$$d(x_n, x_m) \leq d(x_n, x_{n+1}) + \dots + d(x_{m-1}, x_m),$$

und nach Einsetzen von (4.85) mit den entsprechenden Indizes:

$$\begin{aligned} d(x_n, x_m) &\leq k^n d(x_0, x_1) + k^{n+1} d(x_0, x_1) + \dots + k^{m-1} d(x_0, x_1) \\ &= d(x_0, x_1) k^n (1 + k + \dots + k^{m-n-1}) \\ &\leq d(x_0, x_1) (1 - k)^{-1} k^n. \end{aligned} \quad (4.86)$$

Da die Folge $\{k^n\}$ wegen $0 < k < 1$ gegen null konvergiert, ist $\{x_n\}$ wegen (4.86) eine Cauchy-Folge. Da E vollständig ist, liegt der (eindeutig bestimmte) Grenzwert x^* von $\{x_n\}$ in E , da alle x_n , $n \geq 0$, in E_0 liegen und E_0 abgeschlossen ist, liegt x^* in E_0 . Da A eine stetige Abbildung ist [wegen (4.83), s. Satz 3.4] ergibt Grenzwertbildung in der Beziehung $x_{n+1} = A(x_n)$, $n \geq 0$:

$$x^* = \lim_{n \rightarrow \infty} x_{n+1} = \lim_{n \rightarrow \infty} A(x_n) = A\left(\lim_{n \rightarrow \infty} x_n\right) = A(x^*). \quad (4.87)$$

Damit ist x^* als Lösung erkannt: $x^* \in E_0$ und $x^* = A(x^*)$.

Die Eindeutigkeit folgt so: Wären x^* , \bar{x} zwei verschiedene Fixpunkte, so müßte $d(x^*, \bar{x}) = d(Ax^*, A\bar{x}) \leq kd(x^*, \bar{x})$ sein; dies ist aber für $d(x^*, \bar{x}) \neq 0$ unmöglich, weil $k < 1$. Da es somit nur einen Fixpunkt gibt, ist dieser auch vom gewählten Anfangselement x_0 unabhängig. ■

Die Voraussetzungen des Banachschen Fixpunktsatzes lassen sich in verschiedenster Hinsicht abschwächen (vgl. hierzu [30]). Wir nennen repräsentierend die folgende Aussage.

Satz 4.9 (Verallgemeinerung des Banachschen Fixpunktsatzes): *Es sei E ein vollständiger metrischer Raum (Metrik d) und E_0 eine nichtleere abgeschlossene Teilmenge von E . Die Abbildung A bildet die Menge E_0 stetig in sich ab, und es sei eine gewisse Potenz A^n von A eine kontrahierende Abbildung (d. h., $d(A^n(x), A^n(y)) \leq kd(x, y)$, $x, y \in E_0$, $0 < k < 1$ fest). Dann hat A genau einen Fixpunkt $x^* \in E_0$. Für jedes $x_0 \in E_0$ konvergiert die Folge $\{x_n\}$ mit $x_{n+1} = A(x_n)$ ($n = 0, 1, 2, \dots$) gegen x^* .* **S.4.9**

5. Unbeschränkte Operatoren in Hilberträumen

5.1. Halbbeschränkte Operatoren in Hilberträumen

5.1.1. Der Satz von Friedrichs

Ein linearer Operator A , dessen Definitionsbereich $D(A)$ dicht in einem Hilbertraum H liegt, der $D(A)$ in H abbildet und symmetrisch ist,

$$\langle Au | v \rangle = \langle u | Av \rangle \quad (u, v \in D(A)), \quad (5.1)$$

und für den gilt

$$\text{es gibt ein reelles } \gamma \text{ mit } \langle Au | u \rangle \geq \gamma^2 \|u\|^2 \quad (u \in D(A)), \quad (5.2)$$

heißt **halbbeschränkt** (stark positiv definit).

Beispiel 5.1: Es sei $H = L^2_{\mathbb{R}}[0, 1]$, $D(A) = \{u \in C^2_{\mathbb{R}}[0, 1], u(0) = u(1) = 0\}$, und A sei der Operator

$$A = -\frac{d^2}{dx^2}. \quad (5.3)$$

A ist halbbeschränkt. Denn es ist $\overline{D(A)} = H$, A symmetrisch, wie durch partielle Integration folgt,

und (5.2) gilt wegen $u^2(x) = \left(\int_0^x u'(t) dt \right)^2 \leq \int_0^x dt \int_0^x u'^2(t) dt \leq \int_0^1 u'^2(t) dt$, $0 \leq x \leq 1$:

$$\|u\|_{L^2_{\mathbb{R}}}^2 \leq \int_0^1 (u'(t))^2 dt = \langle Au | u \rangle. \quad (5.4)$$

Bemerkung 5.1: Gleichwertig zu (5.2) ist, daß

$$g = \inf_{u \in D(A), u \neq 0} \frac{\langle Au | u \rangle}{\|u\|^2} > 0 \quad (5.2')$$

gilt, d. h., die **untere Grenze** g von A ist positiv.

Der zu A adjungierte Operator A^* ist definiert (vgl. Def. 3.15) für alle $v \in H$, für die ein v^* existiert mit

$$\langle v | Au \rangle = \langle v^* | u \rangle \quad (u \in D(A)). \quad (5.5)$$

Die Abhängigkeit $v \rightarrow v^*$ wird gerade mit $v^* = A^*v$ bezeichnet. Wenn man (5.5) mit (5.1) vergleicht, sieht man, daß

$$D(A^*) \supseteq D(A) \quad (5.6)$$

gilt. Wenn man also für A die Eigenschaft der Selbstadjungiertheit haben möchte ((5.1) zusammen mit $D(A) = D(A^*)$), so muß $D(A)$ wegen (5.6) vergrößert werden. Eine solche **Erweiterung** des Definitionsbereiches von A fordert daher, A auf gewissen weiteren Elementen in H zu definieren, und zwar soll dabei die untere Grenze [vgl. auch (5.2')]]

$$g = \inf \{ \langle Au | u \rangle \mid \|u\| = 1, u \in D(A) \} \quad (5.7)$$

von A (und damit die starke positive Definitheit) erhalten bleiben. Ist das gelungen, so hat man A fortgesetzt. Es gilt der folgende Satz von Friedrichs:

Satz 5.1 (Fortsetzung halbbeschränkter Operatoren): Jeder halbbeschränkte Operator A gestattet eine Fortsetzung \bar{A} , d. h., es ist $D(\bar{A}) \supseteq D(A)$, $Au = \bar{A}u$ ($u \in D(A)$), und dieser Operator \bar{A} ist selbstadjungiert, hat dieselbe untere Grenze wie A , und sein Wertebereich ist der gesamte Raum H . Es existiert sogar $(\bar{A})^{-1}$ und ist symmetrisch und beschränkt. S.5.1

Wir wollen zuerst diesen Fortsetzungsprozeß erläutern und dann auf zwei Anwendungen eingehen: Operatoren der Quantenmechanik, Lösung elliptischer Differentialgleichungen.

5.1.2. Der Fortsetzungsprozeß

Hätten wir die Fortsetzung (= Erweiterung) von A zu \bar{A} schon erledigt, so müßte für ein $\bar{u} \in D(\bar{A})$ das Bild $\bar{A}\bar{u}$ auch in H liegen. Wir nehmen so umgekehrt ein beliebiges $f \in H$ (es soll ja $R(\bar{A}) = H$ sein) und versuchen nach einer einheitlichen Verfahrensweise solche u zuzuordnen, daß für $f \in R(A)$ gerade $u \in D(A)$ gilt, also

$$Au = f \quad (5.8)$$

ist, und daß für die weiteren $f \in H \setminus R(A)$ auch u zugeordnet werden, die dann nicht mehr in $D(A)$ liegen, also $D(A)$ „erweitern“, und daß die gesamte Zuordnung die im Satz genannten Eigenschaften hat. Wir können sagen, daß wir A so erweitern wollen, daß (5.8) für jedes $f \in H$ lösbar wird. Daraus folgt auch, daß die ins Auge gefaßte Erweiterung „maximal“ ist, da H der größtmögliche Wertebereich eines „in H abbildenden“ Operators ist.

Wir führen einige Gedanken des Beweises für *reelle* Hilberträume hier durch, da sich interessante Beziehungen zur Variationsrechnung und konvexen Analysis und eine schöne Anwendung des Satzes von Riesz ergeben [28]; bezüglich des allgemeinen Falles vgl. etwa [16], [36].

Es sei zunächst $f \in R(A)$. Dann gibt es dazu eindeutig ein $u_f \in D(A)$ mit $Au_f = f$. Denn hätte die homogene Gleichung $Au = 0$ Lösungen $u \neq 0$, so widerspräche das (5.2). Für u_f gilt:

Lemma 5.1: u_f löst die Optimierungsaufgabe (das Variationsproblem)

L.5.1

$$F(u) = \langle Au | u \rangle - 2\langle u | f \rangle = \min_{u \in D(A)} \quad (5.9)$$

Beweis: Wir variieren u_f mit $\eta \neq 0$ beliebig aus $D(A)$. Dann ist

$$F(u_f + \eta) = F(u_f) + 2\langle Au_f - f | \eta \rangle + \langle A\eta | \eta \rangle \quad (5.10)$$

wie bei (4.55), da A symmetrisch ist; da u_f die Operatorgleichung (5.8) löst, und (5.2) gilt, folgt

$$F(u_f + \eta) > F(u_f) \quad (\eta \in D(A)). \quad (5.11)$$

Es ist also u_f sogar „globale“ Minimalstelle von F , dies ist auch klar, da F konvex über $D(A)$ ist. u_f ist ferner, wie aus (5.11) folgt, **eindeutige** Minimalstelle von F . Auch dies ist klar, denn F ist [wegen (5.2)] streng konvex (s. Bd. 15). ■

Es gilt auch die Umkehrung:

L.5.2 Lemma 5.2: Ist \hat{u} Minimalstelle von (5.9) bei gegebenem f , so löst \hat{u} die Gleichung $Au = f$.

Beweis: Da F Gateaux-differenzierbar ist [s. (4.58)], muß $\forall \eta \in D(A)$ das Gateaux-Differential $\delta F(u, \eta) = 2\langle Au - f | \eta \rangle$ an der Stelle $u = \hat{u}$ notwendig verschwinden. Wie schon bei (4.66) folgte, ist also $A\hat{u} - f = 0$ ($Au = f$ ist die **Eulersche Gleichung** zu (5.9)). ■

Wir lassen jetzt ein beliebiges $f \in H$ zu. Ist $f \in H \setminus R(A)$, so kann keine Minimalstelle von (5.9) existieren, denn sie würde die Eulersche Gleichung lösen, also $f \in R(A)$ nach sich ziehen. Es gilt aber, daß auch bei beliebigem festem f die untere Grenze J_f der Menge der F -Werte stets endlich ist, wenn u den Bereich $D(A)$ durchläuft:

$$J_f = \inf_{u \in D(A)} F(u) > -\infty. \quad (5.12)$$

Denn es ist, falls nur $\gamma^2 \|u\| - 2\|f\| \geq 1$ ist, wegen (5.2) und der Schwarzschen Ungleichung

$$\langle Au | u \rangle - 2\langle f | u \rangle \geq \|u\| (\gamma^2 \|u\| - 2\|f\|) \geq \|u\| \geq 0,$$

und im anderen Falle, d. h. $\gamma^2 \|u\| - 2\|f\| < 1$, ist

$$\begin{aligned} 2|\langle f | u \rangle| &\leq 2\|f\| \|u\| < \gamma^{-2}(1 + 2\|f\|)(2\|f\|) \\ \langle Au | u \rangle - 2\langle f | u \rangle &\geq \gamma^2 \|u\|^2 - 2\gamma^{-2}(1 + 2\|f\|)\|f\| \leq K \\ (K &= -2\gamma^{-2}(1 + 2\|f\|)\|f\|). \end{aligned} \quad (5.13)$$

Das Variationsproblem (5.9) hat also für jedes $f \in H$ einen endlichen Minimalwert, aber nur für $f \in W(A)$ eine Minimalstelle. Jetzt konstruieren wir durch Vergrößerung des zulässigen Bereiches $D(A)$ des Variationsproblems Minimalstellen für $f \in H \setminus R(A)$, wobei der Minimalwert J_f ungeändert bleibt(!):

Wir führen dazu in der Menge $D(A) \subseteq H$ ein neues Skalarprodukt ein durch

$$\langle u | v \rangle_{D(A)} = \langle Au | v \rangle \quad (u, v \in D(A)). \quad (5.14)$$

Wegen der Halbbeschränktheit ist das ein Skalarprodukt. Dieses induziert eine Norm in $D(A)$:

$$\|u\|_{D(A)}^2 = \langle u | u \rangle_{D(A)} = \langle Au | u \rangle, \quad (5.15)$$

und wegen der Halbbeschränktheit gilt (vgl. auch Def. 2.13)

$$\|u\|_{D(A)}^2 \geq \gamma^2 \|u\|^2 \quad (u \in D(A)). \quad (5.16)$$

Fundamentalfolgen in $D(A)$ im Sinne der Norm $\|\cdot\|_{D(A)}$ sind wegen (5.16) erst recht welche in H , und da H vollständig ist, gehört zu ihnen eindeutig ein Grenzelement. Wir nehmen zu $D(A)$ alle diese Grenzelemente von $\|\cdot\|_{D(A)}$ -Fundamentalfolgen hinzu. Die entstandene Menge heiße H_A und ist Abschließung von $D(A)$ im Sinne der $\|\cdot\|_{D(A)}$ -Norm. H_A ist selbst ein Hilbertraum (energetischer Raum), und es gilt natürlich $H_A \subseteq H$.

Die Norm eines Elements u in H_A ist, falls $u \in D(A) \subseteq H_A$ gilt:

$$\|u\|_{H_A}^2 = \langle Au | u \rangle; \quad (5.17)$$

falls $u \in H_A \setminus D(A)$, so ist u Grenzwert etwa der Fundamentalfolge $\{u_n\}$, also

$$\|u\|_{H_A}^2 = \lim_{n \rightarrow \infty} \langle Au_n | u_n \rangle. \quad (5.18)$$

Beispiel 5.2: Es seien $A, H, D(A)$ wie in Bsp. 5.1. Dann ist $(H \neq H_A)$

$$H_A = \{u \mid u \text{ absolut stetig}^1\} \text{ auf } [0, 1], u_x \in L^2_{\mathbf{R}}[0, 1], u(0) = u(1) = 0\} \subseteq H = L^2_{\mathbf{R}}[0, 1]. \quad (5.18')$$

Da zum zulässigen Bereich $D(A)$ nur Grenzwerte von Folgen aus $D(A)$ hinzugenommen werden, bleibt J_f wirklich ungeändert. Wir setzen nun in (5.9) für $u \in D(A)$ gerade (5.17) ein:

$$F(u) = \|u\|_{H_A}^2 - 2\langle u | f \rangle \quad (5.19)$$

und (5.19) kann sogar für alle $u \in H_A$ gelesen werden! Jetzt hat die Aufgabenstellung

$$F(u) = \|u\|_{H_A}^2 - 2\langle u | f \rangle = \min_{u \in H_A} \quad (5.20)$$

einen Sinn, und es gilt das bemerkenswerte

Lemma 5.3: Für beliebig fest vorgegebenes $f \in H$ existiert genau ein $u_f \in H_A$ mit $L.5.3$
 $F(u_f) = \inf F(u)$, und verschiedenen $f \in H$ entsprechen verschiedene $u_f \in H_A$.

Bemerkung 5.2: Für A gemäß (5.3) lautet $F(u)$ aus (5.9)

$$F(u) = - \int_0^1 uu_{xx} dx - 2 \int_0^1 fu dx,$$

und in der Form (5.19) hat es die Gestalt

$$F(u) = \int_0^1 u'^2 dx - 2 \int_0^1 fu dx, \quad (5.21)$$

und dies kann wegen $u \in H_A$ [s. (5.18')] ausgewertet werden.

Wir führen den Beweis von Lemma 5.3 an, da er eindrucksvoll den Satz von Riesz ausnützt:

Beweis von Lemma 5.3: Der Term $\langle u | f \rangle$ in (5.20) ist ein lineares beschränktes Funktional in H_A , denn es gilt wegen (5.16)

$$|\langle u | f \rangle| \leq \|u\| \|f\| \leq \gamma^{-1} \|u\|_{H_A} \|f\| \quad (5.22)$$

zunächst in $D(A)$ und wegen der Stetigkeit von $\langle \cdot | f \rangle$ und $\|\cdot\|_{H_A}$ auch für die Grenzelemente, also in H_A . Somit gibt es nach dem Satz von Riesz genau ein Element $u_f \in H_A$, so daß gilt

$$\langle u | f \rangle_H = \langle u | u_f \rangle_{H_A} \quad (u \in H_A), \quad (5.23)$$

¹⁾ $u(x)$ heißt absolut stetig auf $[a, b]$, wenn ein v existiert, $v \in L^1_{\mathbf{R}}[a, b]$, so daß $u(x) = \int_a^x v(t) dt + \text{const.}$ ($a \leq x \leq b$) gilt.

wenn mit $\langle \cdot | \cdot \rangle_{H_A}$ das Skalarprodukt in H_A bezeichnet wird. Damit ist für beliebiges festes $u \in H_A$

$$\begin{aligned} F(u) &= \|u\|_{H_A}^2 - 2\langle u | u_f \rangle_{H_A} + \langle u_f | u_f \rangle_{H_A} - \langle u_f | u_f \rangle_{H_A} \\ &= \langle u - u_f | u - u_f \rangle_{H_A} - \|u_f\|_{H_A}^2 \end{aligned} \quad (5.24)$$

und folglich, da der erste Summand in (5.24) nichtnegativ ist,

$$\min_{u \in H_A} F(u) = -\|u_f\|_{H_A}^2. \quad (5.25)$$

Da für $u \neq u_f$ gilt $\|u - u_f\|_{H_A}^2 > 0$, gibt es keine weitere Minimalstelle. ■

D.5.1 Definition 5.1: Das Element $u_f \in H_A$ (s. Lemma 5.3.) heißt **verallgemeinerte Lösung** von $Au = f$. Liegt u_f bereits in $D(A)$, so ist u_f gewöhnliche Lösung von $Au = f$. Die Gesamtheit der verallgemeinerten Lösungen sei der Definitionsbereich $D(\tilde{A})$ eines Operators \tilde{A} . Er wird wie folgt definiert: Zu $\tilde{u} \in D(\tilde{A})$ existiert gemäß Lemma 5.3 genau ein $\tilde{f} \in H$ mit der Eigenschaft, daß \tilde{u} verallgemeinerte Lösung von $Au = \tilde{f}$ ist; die Abbildung $\tilde{u} \rightarrow \tilde{f}$ definiert den Operator \tilde{A} : $\tilde{A}\tilde{u} = \tilde{f}$. Dieser Operator ist die **Friedrichs'sche Erweiterung** von A .

Die weiteren Aussagen des Satzes von Friedrichs sind erfüllt. Die Gleichheit der unteren Grenzen bedeutet

$$\inf_{u \in D(A), u \neq 0} \frac{\langle Au | u \rangle}{\|u\|^2} = \inf_{u \in D(\tilde{A}), u \neq 0} \frac{\|u\|_{H_A}^2}{\|u\|^2}. \quad (5.26)$$

5.1.3. Einige Operatoren der Quantenmechanik

In (1.71)–(1.73) waren die Hamilton-Operatoren als wichtige Hilfsmittel der Quantenmechanik genannt worden. Sie führen auf lineare selbstadjungierte Operatoren. Das frühere Oszillatorbeispiel ergab als Hamilton-Operator zunächst den Operator (1.78). Er war noch nicht selbstadjungiert. Dies kann aber gerade nach dem Satz von Friedrichs erreicht werden. Man studiert dazu die Operatoren folgender Art, die dem Hermiteschen Operator von (1.79) verwandt sind:

$$\begin{aligned} Cu &= -u'' + pu \quad \text{mit} \quad D(C) = \{u \in C^\infty(\mathbf{R})\} \subseteq L^2(\mathbf{R}), p(x) \text{ reell,} \\ \text{und} \quad p(x) &\rightarrow \infty \quad \text{für} \quad |x| \rightarrow \infty, p(x) \geq 1. \end{aligned} \quad (5.27)$$

Nachweis der Symmetrie und Halbbeschränktheit:

$$\begin{aligned} \langle Cu | v \rangle &= \int_{\mathbf{R}} \overline{(-u'' + pu)} v \, dx \\ &= \int_{\mathbf{R}} (\bar{u}' v' + p \bar{u} v) \, dx \\ &= \int_{\mathbf{R}} \bar{u} (-v'' + pv) \, dx = \langle Cv | u \rangle \quad (u, v \in D(C)), \end{aligned} \quad (5.28)$$

$$\langle Cu | u \rangle = \int_{\mathbf{R}} (|u'|^2 + p|u|^2) \, dx \geq \int_{\mathbf{R}} |u|^2 \, dx = \|u\|_{L^2(\mathbf{R})}^2 \quad (5.29)$$

und $\overline{D(C)} = H = L^2(\mathbf{R})$.

Wir denken uns C nach dem Satz von Friedrichs zu \bar{C} fortgesetzt. \bar{C} ist dann selbstadjungiert und hat reines Punktspektrum und leeres Stetigkeitsspektrum. Ist insbesondere $p(x) = x^2 + 1$, so ist folgender Zusammenhang zum Hermiteschen Operator H gegeben:

$$C = H + I, \quad H = D_2 \quad \text{von (1.79)} \quad (5.30)$$

(I ist der identische Operator), und (5.30) bleibt bei der Fortsetzung bestehen: $\bar{C} = \bar{H} + I$. Das Spektrum $\sigma(\bar{H})$ kennen wir bereits aus (1.81). Es gilt dann über dem Bereich $D(\bar{H})$, wenn $\lambda_j = 2j + 1$, $j = 0, 1, 2, \dots$ die Eigenwerte, $\{H_j\}$ die ortho-normierten Eigenfunktionen sind (sie bilden als Eigenfunktionen eines selbstadjungierten Operators mit reinem Punktspektrum und leerem Stetigkeitsspektrum ein vollständiges Orthogonalsystem, [36]):

$$D(\bar{H}) = \left\{ u \mid u \in H, \sum_{j=0}^{\infty} \lambda_j^2 |\langle u \mid H_j \rangle|^2 < \infty \right\}, \quad (5.31)$$

und dies gilt für jeden selbstadjungierten Operator mit reinem Punktspektrum, leerem Stetigkeitsspektrum und $\{\lambda_j\}$ nach der Größe der Beträge geordnet.

Wir wollen noch einen interessanten Zusammenhang zur Theorie der Distributionen (Kap. 4.1.) angeben.

Sind φ und u aus $D(C)$ in (5.27), so ist wegen der Selbstadjungiertheit von \bar{C} : $\bar{C} = \bar{C}^*$ somit

$$\langle \bar{C}u \mid \varphi \rangle = \langle \bar{C}^*u \mid \varphi \rangle = \langle u \mid \bar{C}\varphi \rangle \quad (\varphi \in \mathring{C}^\infty(\mathbf{R})). \quad (5.32)$$

Die rechte Seite kann für $u \in L^2(\mathbf{R})$ als Distribution L aufgefaßt werden ($\bar{C}\varphi = C\varphi$):

$$L: L\varphi = \langle u \mid C\varphi \rangle \quad (\varphi \in \mathring{C}^\infty(\mathbf{R})). \quad (5.33)$$

Diese Distribution wird für $u \in H_C$ gerade von $\bar{C}u$ erzeugt, und folglich kann der Operator \bar{C} im Sinne der Distributionen ausgerechnet werden.

Der Operator

$$Au = xu \quad \text{mit} \quad D(A) = \{u \in L^2(\mathbf{R}) \mid xu(x) \in L^2(\mathbf{R})\} \subseteq L^2(\mathbf{R}) \quad (5.34)$$

tritt ebenso in der Quantenmechanik auf und ist *selbstadjungiert* und *nicht beschränkt*:

Unbeschränktheit:

$$\|A\| = \sup_{\substack{\|u\|=1 \\ u \in D(A)}} \|Au\| = \sup_{\substack{\|u\|=1 \\ u \in D(A)}} \sqrt{\int_{\mathbf{R}} x^2 u^2(x) dx}. \quad (5.35)$$

Mit der Folge $\{u_n\}$ mit $u_n = 1$ für $x \in [n, n+1]$ und $u_n = 0$ sonst wird $\|Au_n\|$ beliebig groß.

Selbstadjungiertheit: Stets war $D(A) \subseteq D(A^*)$ für symmetrische Operatoren. Es sei $v \in D(A^*)$ und $A^*v = v^*$. Dann gilt nach der Definitionsformel für adjungierte Operatoren

$$\langle Au \mid v \rangle = \langle u \mid v^* \rangle \quad (u \in D(A)). \quad (5.36)$$

Dies bedeutet

$$\begin{aligned} 0 &= \int_{\mathbf{R}} [x(\bar{u}(x)) v(x) - \bar{u}(x) v^*(x)] dx \\ &= \int_{\mathbf{R}} \bar{u}(x) (xv(x) - v^*(x)) dx, \end{aligned} \quad (5.37)$$

insbesondere für alle $u \in \dot{C}^\infty(\mathbf{R})$, folglich fast überall $xv(x) = v^*(x)$; da $v^*(x) \in L^2(\mathbf{R})$, ist $xv(x) \in L^2(\mathbf{R})$, also $v \in D(A)$. Somit ist $D(A^*) \subseteq D(A)$, also $D(A) = D(A^*)$.

Auch der Operator B mit

$$Bu = -\frac{d^2u}{dx^2}, \quad D(B) = \{u \in \dot{C}^\infty(\mathbf{R})\} \subseteq L^2(\mathbf{R}) \quad (5.38)$$

tritt in der Quantenmechanik auf. Er fällt in die bei (5.27) behandelte Klasse. Als H_B im Sinne von (5.17) fungiert $W^{1,2}(\mathbf{R})$. Wird $D(B) = W^{2,2}(\mathbf{R})$ und $H = L^2(\mathbf{R})$ gewählt, so ist B selbstadjungiert über $W^{2,2}(\mathbf{R})$ und kann fortgesetzt werden zu \bar{B} mit $D(\bar{B}) = H_B = W^{1,2}(\mathbf{R})$.

Das Modell „Atomkern im Koordinatensprung, ein Elektron im Abstand r mit der Ladung e und der Masse m sowie der Kernladungszahl 1 des Atoms“ führt im R^3 auf den Hamilton-Operator

$$\mathfrak{H}u = -\left(\frac{\hbar^2}{2m} \Delta + \frac{e^2}{r}\right)u \quad \left(\hbar = \frac{h}{2\pi} 6,62 \cdot 10^{-27} \text{ erg s}\right). \quad (5.39)$$

In der Klammer steht ein elliptischer Differentialoperator. Etwa bei homogenen Dirichlet-Randvorgaben ist er halbbeschränkt (s. 5.1.6.), also über den Friedrichschen Formalismus fortsetzbar zu einem selbstadjungierten Operator $\tilde{\mathfrak{H}}$.

5.1.4. Instationäre Zustände und Schrödinger-Gleichung

Es sei $\tilde{\mathfrak{H}}$ der (selbstadjungierte) Hamilton-Operator (vgl. 1.2.3.) eines quantenmechanischen Systems. Befindet sich das System zur Zeit $t = 0$ im Zustand $\hat{u} \in D(\tilde{\mathfrak{H}})$, $\|\hat{u}\| = 1$, so gilt, daß sein Zustand zum Zeitpunkt $t > 0$ durch $u(t)$ mit $\|u(t)\| = 1$, $u(\cdot) \in D(\tilde{\mathfrak{H}})$ gekennzeichnet ist, wobei $u(t)$ für $t > 0$ der abstrakten Differentialgleichung (Schrödinger-Gleichung)

$$\dot{u}(t) = -i\hbar^{-1}\tilde{\mathfrak{H}}u(t) \quad (5.40)$$

genügt. Zusammen mit der Anfangsbedingung $u(0) = \hat{u}$ ist (5.40) ein **abstraktes Cauchysches Anfangswertproblem**.

Die Lösung dieses Cauchyschen Anfangswertproblems ist eindeutig bestimmt und gegeben durch [vgl. (5.51) und (5.99)]

$$u(t) = e^{-it\hbar^{-1}\tilde{\mathfrak{H}}} \hat{u}, \quad t \geq 0. \quad (5.41)$$

Der Operator rechts in (5.41) kann [vgl. (3.28)] als Potenzreihe nach den Potenzen von $\tilde{\mathfrak{H}}$ dargestellt werden:

$$e^{At} \hat{u} = \left(\sum_{k=0}^{\infty} \frac{1}{k!} A^k \right) \hat{u}, \quad A = -it\hbar^{-1}\tilde{\mathfrak{H}}, \quad (5.41')$$

wobei die Konvergenz gemäß (3.43) zu verstehen ist und die Reihe („in der Operator-norm“, also erst recht für jedes $\hat{u} \in H$) konvergiert [vgl. auch (5.51)].

Stationarität von \hat{u} bedeutet: Der Anfangszustand des Systems, charakterisiert durch \hat{u} , ändert sich nicht mit der Zeit. Da $e^{i\alpha(t)} \hat{u}$ mit $\alpha(t)$ reell, endlich, denselben eindimensionalen Unterraum wie \hat{u} , also denselben Zustand erzeugt, ist die Gleichung

$$e^{-it\hbar^{-1}\tilde{\mathfrak{H}}} \hat{u} = e^{i\alpha(t)} \hat{u} \quad (5.42)$$

notwendig und hinreichend für Stationarität.

Wir wollen (5.40) deuten. Für das bei (5.39) beschriebene Modell lautet (5.40) für Funktionen $u \in D(\mathfrak{H})$, wobei u jetzt noch von der Zeit abhängt,

$$i\hbar \dot{u} = -\frac{\hbar^2}{2m} \Delta u + V(x) u, \quad (5.43)$$

und da \mathfrak{H} über die Friedrichssche Fortsetzung eindeutig bestimmt ist, müssen die quantenmechanischen Informationen schon in (5.43) enthalten sein. Man kann Lösungen in Gestalt von Fourierreihen nach den orthonormierten Eigenfunktionen (= stationären Zuständen!) $\psi_k(x)$ und Eigenwerten λ_k des Operators $Lu = -\frac{\hbar^2}{2m} \Delta u + V(x) u$ erhalten, wenn man, wie bei einer parabolischen Differentialgleichung in dem zylindrischen Gebiet $[0, T] \times G$, ein Randanfangswertproblem stellt, etwa

$$\begin{aligned} u(0, x) &= \hat{u}(x) \quad (x \in G, \text{ Zeit } t = 0), \\ u(t, x) &= 0 \quad (t \geq 0, x \in \partial G). \end{aligned} \quad (5.44)$$

Es ist dann

$$u(t, x) = \sum_{k=1}^{\infty} \langle \psi_k | \hat{u} \rangle e^{-i\hbar^{-1} \lambda_k t} \psi_k(x) \quad (5.45)$$

eine formale Lösung (und sogar klassische Lösung bei $\hat{u}(x) \in \hat{C}^2(G)$ und hinreichend glattem Rand ∂G); [38], [14/4].

Diese Deutung von (5.40) bezog nicht den erweiterten Operator \mathfrak{H} ein. Dazu betrachten wir folgende, von einer Variablen t abhängige Menge von Elementen $u(t)$ eines Hilbertraumes H : $\{u(t) \mid u(t) \in H \text{ für jedes } t \in (\alpha, \beta) \text{ mit } -\infty \leq \alpha < \beta \leq +\infty\}$. $u(t)$ heißt auch **Funktion von t mit Werten in einem Hilbertraum**. Diese heißt stetig im Punkte $\gamma \in (\alpha, \beta)$, falls es zu jedem $\varepsilon > 0$ ein $\delta = \delta(\varepsilon) > 0$ gibt mit

$$\|u(t) - u(\gamma)\| < \varepsilon \quad \text{für } |\gamma - t| < \delta(\varepsilon); \quad (5.46)$$

sie heißt stetig differenzierbar in γ , falls ein Element $u'(\gamma) \in H$ existiert mit

$$\lim_{h \rightarrow 0} \left\| \frac{u(\gamma + h) - u(\gamma)}{h} - u'(\gamma) \right\| = 0 \quad (5.47)$$

und $u'(t)$ auch in einer Umgebung von γ existiert und in γ stetig ist.

Wir betrachten nun folgende Aufgabenstellung: Es seien ein selbstadjungierter Operator A und ein Element $\hat{u} \in D(A)$ gegeben. Weiter werde das Intervall $[0, \infty)$ betrachtet. Gesucht sei eine stetige und stetig differenzierbare Funktion¹⁾ $u(t)$, die für jedes $t \in [0, \infty)$ im Definitionsbereich von A liegt, bei $t = 0$ den Wert \hat{u} annimmt und so beschaffen ist, daß $u'(t)$ für jedes t gerade gleich $iAu(t)$ ist. Wir wollen also die folgende Aufgabe lösen:

Verallgemeinertes Cauchysches Anfangswertproblem

$$\begin{aligned} u(t) &\in D(A) \quad \text{für alle } t > 0, \quad u(0) = \hat{u}, \\ u'(t) &= iAu(t). \end{aligned} \quad (5.48)$$

Eine weitere Verallgemeinerung wäre ein von t abhängiger Operator A (vgl. [23], parabolische Differentialgleichungen).

¹⁾ In $t = 0$ sind dann (5.46), (5.47) rechtsseitig erfüllt.

S.5.2 Satz 5.2 (Existenz- und Unitätssatz): Aufgabe (5.48) besitzt genau eine Lösung. Sie läßt sich in der Form

$$u(t) = e^{itA} \hat{u}, \quad 0 \leq t < \infty, \quad (5.49)$$

darstellen. Die Operatoren e^{itA} sind unitär. Die Lösung hängt stetig von den Anfangswerten ab [36, S. 280].

Es sei A ein Operator mit reinem Punktspektrum, und das System der Eigenfunktionen φ_k zu den Eigenwerten $\lambda_1, \lambda_2, \dots$, sei vollständig. Dann ist für $y \in D(A)$

$$\begin{aligned} Ay &= \sum \langle \varphi_k | Ay \rangle \varphi_k = \sum \langle A\varphi_k | y \rangle \varphi_k \\ &= \sum \lambda_k \langle \varphi_k | y \rangle \varphi_k. \end{aligned} \quad (5.50)$$

Für den Operator e^{itA} gilt dann (für jedes $t \geq 0$)

$$e^{itA} y = \sum e^{it\lambda_k} \langle \varphi_k | y \rangle \varphi_k. \quad (5.51)$$

Diese Reihe konvergiert im Sinne der Norm. Es ist $D(e^{itA}) = R(e^{itA}) = H$. Man erkennt sofort, daß e^{itA} unitär ist:

$$\begin{aligned} \|e^{itA} y\|^2 &= \sum_j \sum_k e^{it(\lambda_k - \lambda_j)} \overline{\langle \varphi_j | y \rangle} \langle \varphi_k | y \rangle \langle \varphi_j | \varphi_k \rangle \\ &= \sum_j |\langle \varphi_j | y \rangle|^2 = \|y\|^2 \end{aligned} \quad (5.52)$$

wegen der Vollständigkeitsrelation. Unter Verwendung von (5.47) und der Tatsache, daß A mit e^{itA} vertauscht werden darf, erkennt man, daß $u(t)$ in (5.49) der Differentialgleichung (5.48) genügt.

Nachweis der stetigen Abhängigkeit von den Anfangswerten: Es seien $u_0(t)$ und $u_1(t)$ die Lösungen von (5.48) mit $u_0(0) = \hat{u}$ bzw. $u_1(0) = \hat{u}$. Dann ist wegen (5.52)

$$\|u_0(t) - u_1(t)\| = \|u_0 - u_1\| \quad (t > 0), \quad (5.53)$$

und somit ist die Lösungsdifferenz klein, wenn die Differenz der Anfangswerte klein ist.

Die Gl. (5.48) ist die abstrakte Fassung der Schrödinger-Gleichung der Quantentheorie [vgl. (5.40)].

5.1.5. Beziehungen zur quantenmechanischen Streuung. Unschärferelation

Es sei $\tilde{\mathfrak{H}}$ der selbstadjungierte Hamilton-Operator eines quantenmechanischen Systems; $f \in D(\tilde{\mathfrak{H}})$ mit $\|f\| = 1$ sei der Zustand des Systems. In (1.95') war $\langle \tilde{\mathfrak{H}} f | f \rangle$ gerade der quantenmechanische Mittelwert der Energiemessung. Die Streuung wird dann (wie in der Wahrscheinlichkeitsrechnung, vgl. auch Bd. 17) durch

$$\sigma^2(f|\tilde{\mathfrak{H}}) = \|\tilde{\mathfrak{H}} f - \langle \tilde{\mathfrak{H}} f | f \rangle f\|^2 \quad (5.54)$$

definiert. Eine „scharfe“ Energiemessung ist für einen Zustand möglich genau dann, wenn gilt $\sigma^2 = 0$:

$$\tilde{\mathfrak{H}} f = \langle \tilde{\mathfrak{H}} f | f \rangle f. \quad (5.55)$$

Dies bedeutet, daß f Eigenwert von $\tilde{\mathfrak{H}}$, also stationärer Zustand ist. Man gelangt mit dem Streuungsbegriff (5.54) in folgender Weise funktionalanalytisch zur Heisenbergschen Unschärferelation: Es sei f mit $\|f\| = 1$ der Zustand eines quantenmechanischen Systems, und A, B seien zwei beobachtbare Größen, also zwei lineare selbstadjungierte Operatoren. Es sei

$$f \in D(A) \cap D(B), \quad Af \in D(B), \quad Bf \in D(A). \quad (5.56)$$

Dann ist ([36, S. 474])

$$\sigma(f|A) \sigma(f|B) \geq \frac{1}{2} |\langle (BA - AB)f | f \rangle|. \quad (5.57)$$

Ist nun $(BA - AB)f = \varrho f$ (ϱ komplexe Zahl), so folgt

$$\sigma(f|A) \sigma(f|B) \geq \frac{1}{2} |\varrho| \langle f | f \rangle = \frac{1}{2} |\varrho|. \quad (5.58)$$

(5.58) mit $\varrho \neq 0$ zieht nach sich, daß ein f [welches (5.56) genügt] nur existieren kann, wenn $\sigma(f|A)$ und $\sigma(f|B)$ nicht verschwinden. f kann dann nicht Eigenelement von A oder B sein, bzw. im f entsprechenden Zustand des quantenmechanischen Systems kann weder A noch B scharf gemessen werden.

Die praktische Bedeutung liegt darin, daß für einen Zustand f , für den die beobachtbare Größe A eine Messung mit kleiner Streuung $\sigma(f|A)$ erlaubt, die Streuung $\sigma(f|B)$ die Relation (5.58) erfüllen muß. Für den Fall des Oszillators (A, B seien die der Impuls- bzw. Ortskoordinate zugeordneten Operatoren)

$$Bf = xf, \quad D(B) = \{f | f \in L^2(\mathbf{R}), xf \in L^2(\mathbf{R})\}, \quad (5.59)$$

$$Af = \frac{\hbar}{i} \frac{df}{dx}, \quad D(A) = W^{1,2}(\mathbf{R}) \quad (5.60)$$

ist dann für $f \in \dot{C}^\infty(\mathbf{R})$

$$(BA - AB)f = \frac{\hbar}{i} x \frac{df}{dx} - \frac{\hbar}{i} \frac{d}{dx} (xf) = \frac{\hbar}{i} f \quad (5.61)$$

in die rechte Seite von (5.57) einzusetzen, so daß sich

$$\sigma(f|A) \sigma(f|B) \geq \frac{1}{2} \hbar, \quad (5.62)$$

also die bekannte **Unschärferelation** ergibt.

5.1.6. Fortsetzung elliptischer Differentialoperatoren

T sei ein partieller linearer selbstadjungierter Differentialausdruck 2. Ordnung, der in einem beschränkten Gebiet Ω mit stückweise glattem Rand $\partial\Omega$ elliptisch ist:

$$Tu \equiv - \sum_{i,k} \frac{\partial}{\partial x_i} \left(a_{ik}(x) \frac{\partial}{\partial x_j} u \right) + cu, \quad c = c(x_1, \dots, x_n) \geq 0, \quad (5.63)$$

mit $a_{ik} \in C^1_{\mathbf{R}}(\bar{\Omega})$, $c \in C^0_{\mathbf{R}}(\bar{\Omega})$, $\sum a_{ik}(x_1, \dots, x_n) \lambda_i \lambda_k \geq \mu \sum_{i=1}^n \lambda_i^2$, $(a_{ik}) = A$ (symmetrisch, $\mu > 0$).

Als μ kann das Infimum μ^* der Eigenwerte der quadratischen Form $\lambda^T A(x) \lambda \forall x$ dienen, falls $\mu^* > 0$ ist. Als Randbedingung sei $u(x) = 0$ ($x \in \partial\Omega$) gestellt. Damit ist

$$D(T) = \{u | u \in C^2_{\mathbf{R}}(\bar{\Omega}), u = 0 \text{ für } x \in \partial\Omega\} \quad (5.64)$$

dicht in $L^2_{\mathbf{R}}(\Omega)$. T ist dann symmetrisch, wie aus der *Greenschen Formel*¹⁾

$$\int_{\Omega} \dots \int (v Tu - u Tv) dx = - \int_{\partial\Omega} \dots \int \left(v \frac{\partial}{\partial N} u - u \frac{\partial}{\partial N} v \right) df, \quad (5.65)$$

folgt, wobei $\frac{\partial}{\partial N}$ die Konormalenableitung

$$\frac{\partial}{\partial N} = \sum \alpha_i \frac{\partial}{\partial x_i}, \quad \alpha_i = \sum_k a_{ki} \cos(n, x_k), \quad (5.66)$$

¹⁾ Vgl. Bd. 5 und 8.

und n die Außennormale bezüglich $\partial\Omega$ ist. T ist aber auch halbbeschränkt in $L^2_{\mathbf{R}}(\Omega)$, denn es ist (partielle Integration)

$$\langle Tu | u \rangle = \int_{\Omega} \dots \int \left(- \sum_{i,k} \frac{\partial}{\partial x_i} \left(a_{ik} \frac{\partial}{\partial x_k} u \right) u + cuu \right) dx \quad (5.67)$$

$$= \int_{\text{part. Int}} \dots \int_{\Omega} \left(\sum_{i,k} a_{ik} \frac{\partial u}{\partial x_k} \right) u_{x_i} + cuu \, dx + 0, \quad (5.68)$$

und wegen der Elliptizität $\lambda^T A \lambda \geq \mu \|\lambda\|^2$ folgt

$$\langle Tu | u \rangle \geq \mu \int_{\Omega} \dots \int \sum |u_{x_i}|^2 dx + \int_{\Omega} \dots \int c(x_1, \dots, x_n) u^2 dx \quad (5.69)$$

und hieraus wegen der Ungleichung von Friedrichs und $c \geq 0$ die Halbbeschränktheit.

Die **Ungleichung von Friedrichs** lautet ([36], [28])

$$\int_{\Omega} \dots \int u^2(x) dx \leq \text{const} \int_{\Omega} \dots \int \sum u_{x_i}^2(x) dx. \quad (5.70)$$

Die Bildung der Fortsetzung nach Friedrichs gestattet leicht, Aussagen über das Eigenwertproblem zum 1. Randwertproblem zu machen: Es lautet

$$Tu - \lambda u = 0 \quad \text{im Gebiet } \Omega \text{ mit } \partial\Omega = S, \quad u(S) = 0. \quad (5.71)$$

Zu T sei \bar{T} die Friedrichssche Fortsetzung (\bar{T} ist ein Operator mit reinem Punktspektrum, der Raum H_T ist gleich $\dot{W}_{\mathbf{R}}^{1,2}$). \bar{T} hat eine lineare selbstadjungierte vollstetige Inverse G . Wendet man G auf die mit \bar{T} gebildete Eigenwertgleichung an, so erhält man

$$u - \lambda Gu = 0, \quad u(S) = 0. \quad (5.72)$$

Dieses EWP ist vollständig zu übersehen, da G *vollstetig* ist. Die Eigenwerte (λ_j^{-1}) von (5.72) häufen sich also höchstens bei 0, sind reell, haben endliche Vielfachheit, Eigenfunktionen zu verschiedenen Eigenwerten stehen orthogonal zueinander, und es gilt der Fredholmsche Alternativsatz (s. 3.3.2. und 3.3.1.3.).

Es ist interessant, daß das System der Eigenfunktionen **vollständig** ist, denn G besitzt eine Inverse: $G\varphi = 0 \Rightarrow \varphi = 0$. Deshalb gilt, daß wirklich unendlich viele Eigenfunktionen da sind, da der Grundraum $L^2_{\mathbf{R}}$ unendlichdimensional ist. Ein zugehöriges vollständiges ONS kann man zur Lösung der Gleichung

$$\bar{T}u - \lambda u = g \quad (5.73)$$

bei gegebenem $g \in L^2_{\mathbf{R}}(\Omega)$ anwenden:

Es sei λ nicht Eigenwert und $\{\lambda_j\}$ das System der Eigenwerte, $\{\varphi_j\}$ das System der Eigenfunktionen als ONS. Für die gesuchte Lösung u von $\bar{T}u - \lambda u = g$ machen wir folgenden Ansatz:

$$u = \sum_{j=1}^{\infty} \langle \varphi_j | u \rangle \varphi_j. \quad (5.74)$$

Da G ein beschränkter linearer Operator ist, folgt

$$Gu = \sum \langle \varphi_j | u \rangle G\varphi_j = \sum \langle \varphi_j | u \rangle \lambda_j^{-1} \varphi_j. \quad (5.75)$$

Da die inhomogene Operatorgleichung erfüllt sein soll, ergibt Einsetzen in die mittels G umgeformte Gleichung $u - \lambda Gu = Gg$:

$$\sum_j \langle \varphi_j | u \rangle \varphi_j - \lambda \sum_j \varphi_j \frac{1}{\lambda_j} \langle \varphi_j | u \rangle = Gg,$$

und nach Skalarmultiplikation mit φ_i , $i = 1, 2, \dots$, erhält man

$$\sum_j \langle \varphi_i | \varphi_j \rangle \langle \varphi_j | u \rangle - \lambda \sum_j \langle \varphi_j | u \rangle \frac{1}{\lambda_j} \langle \varphi_i | \varphi_j \rangle = \langle \varphi_i | Gg \rangle.$$

Die Orthonormalität des Systems $\{\varphi_i\}$ ergibt

$$\langle \varphi_i | u \rangle (1 - \lambda \lambda_i^{-1}) = \langle \varphi_i | Gg \rangle \quad (i = 1, 2, \dots);$$

durch Koeffizientenvergleich mit dem Ansatz folgt

$$u = \sum_i \frac{\langle \varphi_i | Gg \rangle}{1 - \lambda \lambda_i^{-1}} \varphi_i \quad (\lambda \neq \lambda_i^{-1}),$$

und wegen der Symmetrie von G und $G\varphi_i = \lambda_i^{-1}\varphi_i$ erhält man schließlich

$$u = \sum_i \frac{\langle \varphi_i | g \rangle}{\lambda_i - \lambda} \varphi_i. \quad (5.76)$$

Im Falle, daß $\lambda = \lambda^*$ Eigenwert ist, muß laut Theorie (Fredholmscher Alternativsatz, Satz 3.15) mit der rechten Seite Gg der umgerechneten Eigenwertgleichung gelten (damit eine Lösung existiert)

$$\langle Gg | \varphi^* \rangle = 0 \text{ für alle } \varphi^* \text{ mit } \varphi^* - \lambda G\varphi^* = 0. \quad (5.77)$$

Dies heißt:

$$0 = \langle Gg | \varphi^* \rangle = \langle g | G\varphi^* \rangle = \left\langle g \left| \frac{1}{\lambda^*} \varphi^* \right. \right\rangle = \langle g | \varphi^* \rangle, \quad (5.78)$$

also $g \perp \varphi^*$. Das ist die Lösbarkeitsbedingung für elliptische Differentialgleichungen bei Vorhandensein von Eigenwerten bei der homogenen 1. Randwertaufgabe.

Das Spektrum von \bar{T} (erst recht das von T) ist ein reines Punktspektrum, da das Lösen von $\bar{T}u - \lambda u = g$ mit dem Lösen von $u - \lambda Gu = Gg$ gleichwertig war. Nun gehört in der Gleichung $\mu u - Gu = 0$ zwar $\mu = 0$ als Häufungspunkt der Eigenwerte zum Spektrum, ist aber nicht Eigenwert, da $Gu = 0$ nach sich zieht $u = 0$. Da für \bar{T} die μ^{-1} die Eigenwerte sind, hat \bar{T} reines Punktspektrum.

Ist also λ nicht Eigenwert λ_i , so ist $\bar{T}u - \lambda u = g$ für jedes g lösbar. Folglich ist für $\lambda \neq \lambda_i$ die **Resolvente** $R_\lambda = (\bar{T} - \lambda I)^{-1}$ überall erklärt. Sie ist auch beschränkt:

Die Lösung von $\bar{T}u - \lambda u = g$ für λ (nicht Eigenwert) lautete:

$$u = \sum_{k=1}^{\infty} \frac{\langle \varphi_k | g \rangle}{\lambda_k - \lambda} \varphi_k = R_\lambda g, \quad (5.76')$$

weil $R_\lambda = (\bar{T} - \lambda I)^{-1}$. Da $\{\varphi_k\}$ vollständig ist, gilt für g ,

$$g = \sum_{k=1}^{\infty} \langle \varphi_k | g \rangle \varphi_k,$$

die Vollständigkeitsrelation [= Parsevalsche Gleichung, vgl. (2.66)]:

$$\|g\|^2 = \sum_{k=1}^{\infty} |\langle \varphi_k | g \rangle|^2.$$

In der Lösungsdarstellung (5.76) ergibt Normbildung

$$\|u\|^2 \leq \sum_{k=1}^{\infty} \frac{|\langle \varphi_k | g \rangle|^2}{(\lambda_k - \lambda)^2} \underbrace{\|\varphi_k\|^2}_{=1}. \quad (5.79)$$

Nun sei $\min_{\lambda_i} |\lambda - \lambda_i| = d$ (λ fest, λ nicht Eigenwert). Dann ergeben die beiden letzten Resultate

$$\|R_\lambda g\|^2 = \|u\|^2 \leq \sum_{k=1}^{\infty} \frac{|\langle \varphi_k | g \rangle|^2}{d^2} = \frac{1}{d^2} \|g\|^2 \quad (g \in L^2_\mathbf{R}(\Omega)),$$

also

$$\|R_\lambda\| \leq d^{-1}. \quad (5.80)$$

Folglich gehören alle λ zur Resolventenmenge, nur die Eigenwerte nicht. (Das gilt auch für komplexe λ , $g \in L^2(\Omega)$, s. [24].) Eigenwertprobleme für elliptische Differentialoperatoren treten z. B. bei Eigenschwingungen einer eingespannten Membran auf [35/2].

5.2. Spektralzerlegung selbstadjungierter Operatoren in Hilberträumen

5.2.1. Vollstetige Operatoren

In separablen Hilberträumen H gilt für vollstetige, vom Nulloperator verschiedene, symmetrische Operatoren $T: H \rightarrow H$ der Entwicklungssatz (s. auch Sätze 2.39 und 3.22): Ist $\lambda = 0$ nicht Eigenwert, so ist das zu den Eigenwerten $\lambda_j \neq 0$ ($j = 1, 2, \dots$) gehörige Orthonormalsystem von Eigenfunktionen $\{\varphi_j\}$ vollständig. Für jedes $u \in H$ gilt

$$u = \sum_{j=1}^{\infty} \langle \varphi_j | u \rangle \varphi_j, \quad Tu = \sum_{j=1}^{\infty} \lambda_j \langle \varphi_j | u \rangle \varphi_j. \quad (5.81)$$

Wir formen die Reihendarstellung von u um, um eine Gestalt zu erhalten, die auch allgemeine (d. h. T nur selbstadjungiert, H beliebiger Hilbertraum) Gültigkeit hat. Dazu führen wir folgende Operatoren E_λ ein, wobei λ ein reeller Parameter sei:

$$E_\lambda u = \sum_{\lambda_j \leq \lambda} \langle \varphi_j | u \rangle \varphi_j \quad (u \in H). \quad (5.81')$$

Offenbar ist E_λ für $\lambda < -\|T\|$ der Nulloperator, denn der Spektralradius von T ist $\|T\|$. Ebenso erkennt man $E_\lambda = I$, falls $\lambda > \|T\|$ ist, denn wegen der Vollständigkeit ist für $\lambda > \|T\|$

$$E_\lambda u = \sum_{j=1}^{\infty} \langle \varphi_j | u \rangle \varphi_j = Iu = u.$$

Deshalb heißt die Schar der Operatoren E_λ eine **Zerlegung der Einheit**. Diese Operatoren E_λ haben bemerkenswerte Eigenschaften:

a) Sie sind (für jedes λ) linear, symmetrisch, idempotent, d. h., $E_\lambda(E_\lambda u) = E_\lambda u$ für jedes $u \in H$, und es ist $E_\lambda(E_\mu(u)) = E_{\min(\lambda, \mu)}(u)$. E_λ ist Projektionsoperator von H in $R(E_\lambda)$; es gilt überdies $\|E_\lambda\| = 1$, ferner ist E_λ positiv (denn $\langle E_\lambda u | u \rangle = \|E_\lambda u\|^2 \geq 0$).

b) Die Funktion $\varrho(\lambda) = \langle E_\lambda u | u \rangle$ für jedes (beliebige) $u \in H$, und $-\infty < \lambda < +\infty$ ist eine Treppenfunktion mit (höchstens) abzählbar unendlich vielen Sprungstellen:

$$\begin{aligned} \varrho(\lambda) = \langle E_\lambda u | u \rangle &= \left\langle \sum_{\lambda_j \leq \lambda} \langle \varphi_j | u \rangle \varphi_j \middle| \sum_{k=1}^{\infty} \langle \varphi_k | u \rangle \varphi_k \right\rangle \\ &= \sum_{\lambda_j \leq \lambda} |\langle \varphi_j | u \rangle|^2, \end{aligned} \quad (5.82)$$

und für $\mu \geq \lambda$ ist

$$\varrho(\mu) - \varrho(\lambda) = \sum_{\lambda < \lambda_j \leq \mu} |\langle \varphi_j | u \rangle|^2 \geq 0; \quad (5.83)$$

es ist also $\varrho(\lambda)$ eine monoton wachsende Funktion (und ϱ ist rechtsseitig stetig: $\lim_{\lambda \rightarrow \lambda_0 + 0} \varrho(\lambda) = \varrho(\lambda_0)$).

Für stetige Funktionen $f(\lambda)$ existiert somit das Stieltjes-Integral [14/2] mit der Belegungsfunction $\varrho(\lambda)$:

$$\int_a^b f(\lambda) d\varrho(\lambda). \quad (5.84)$$

Es gilt nun der umgeformte Entwicklungssatz [vgl. (5.81)]:

Satz 5.3: Für jedes $u \in H$ (0 ist nicht Eigenwert von T , sonst s. Satz 5.4) ist S.5.3

$$\langle Tu | u \rangle = \int_{-\|T\| - 0}^{\|T\|} \lambda d\varrho(\lambda) = \int_{-\|T\| - 0}^{\|T\|} \lambda d\langle E_\lambda u | u \rangle. \quad (5.85)$$

Man schreibt kurz (und diese Kurzschreibweise läßt sich ebenso durch einen Zerlegungsprozeß wie bei einem Stieltjes-Integral erklären [16]):

$$T = \int_{-\|T\| - 0}^{\|T\|} \lambda dE_\lambda; \quad Tu = \int_{-\|T\| - 0}^{\|T\|} \lambda d(E_\lambda u). \quad (5.86)$$

Da im Stieltjes-Integral in (5.84) nur die *Sprungstellen* von ϱ einen Anteil zum Integral liefern (weil ϱ eine Treppenfunktion nach (5.82) ist), wird T in (5.86) genau durch seine Eigenwerte, also sein Punktspektrum ($\sigma(T) = \sigma_p(T) \cup \{0\}$, vgl. Satz 3.13) bestimmt. Der Satz heißt daher auch der **Spektralsatz**, die Darstellung (5.85) bzw. (5.86) die **Spektralzerlegung** von T und die Schar $\{E_\lambda\}$ die **Spektralschar** von T .

5.2.2. Selbstadjungierte Operatoren in Hilberträumen

Analog gilt für selbstadjungierte nicht notwendig beschränkte (vgl. Def. 3.17 und Abschnitt 5.1.1.) Operatoren T :

Satz 5.4 (Allgemeiner Spektralsatz in Hilberträumen¹⁾): Ist T ein selbstadjungierter Operator mit $D(T) \subseteq H$, $R(T) \subseteq H$, so gibt es eine eindeutig durch T festgelegte Spektralschar $\{E_\lambda\}$ mit den Eigenschaften S.5.4

- wie im vollstetigen Fall.
- $\varrho(\lambda) = \langle E_\lambda u | u \rangle$ ist für jedes $u \in H$ monoton wachsend, rechtsseitig stetig, und es ist $\lim_{\lambda \rightarrow -\infty} \varrho(\lambda) = 0$, $\lim_{\lambda \rightarrow +\infty} \varrho(\lambda) = \|u\|^2$ (ϱ wird i. allg. nicht mehr Treppenfunktion sein). Ferner ist für jedes reelle μ : $E_{\mu+0}u = \lim_{\lambda \rightarrow \mu+0} E_\lambda u = E_\mu u$ für jedes $u \in H$ und $\lim_{\lambda \rightarrow -\infty} \|E_\lambda u\| = 0$, $\lim_{\lambda \rightarrow +\infty} \|E_\lambda u\| = \|u\|$ für jedes $u \in H$.

¹⁾ Ein Beweis wurde zuerst von J. v. Neumann gegeben, s. a. [1].

c) Für alle $u \in D(T)$ gilt die Spektraldarstellung

$$\langle Tu | u \rangle = \int_{-\infty}^{+\infty} \lambda d\langle E_\lambda u | u \rangle \quad (5.87)$$

oder auch analog abgekürzt wie in (5.86) geschrieben.

Interessanterweise kann auch der Definitionsbereich $D(T)$ durch das Stieltjes-Integral mit $\varrho(\lambda)$ charakterisiert werden; es ist nämlich

$$u \in D(T) \Leftrightarrow \int_{-\infty}^{+\infty} \lambda^2 d\varrho(\lambda) < \infty, \quad (5.88)$$

und es ist $E_\lambda u \in D(T)$ für jedes $u \in H$ und jedes λ sowie

$$\langle Tu | v \rangle = \int_{-\infty}^{+\infty} \lambda d\langle E_\lambda u | v \rangle \quad (u \in D(T), v \in H). \quad (5.89)$$

Die Kenntnis der Spektralschar $\{E_\lambda\}$ eines selbstadjungierten Operators T gestattet, die früher vorgenommenen Anwendungen der Funktionalanalysis in der Quantenmechanik zu verallgemeinern, denn wir hatten uns in Kap. 4. stets auf Operatoren mit reinem Punktspektrum bezogen. Ganz ähnlich wie bei beschränkten Operatoren lassen sich über die Spektralschar $\{E_\lambda\}$ jetzt die Begriffe Spektrum von T , Größe a Resolventenmenge usw. erklären (vgl. [16], [1]; s. auch bei (5.110)).

5.2.3. Anwendungen auf die Quantenmechanik

A sei ein selbstadjungierter Operator, der eine beobachtbare Größe a der Quantenmechanik beschreibt. Befindet sich nun das quantenmechanische System im Zustand a , so ist der quantenmechanische Mittelwert bezüglich der beobachtbaren Größe a durch $\langle Au | u \rangle$ definiert. Es ist daher

$$\langle Au | u \rangle = \int_{-\infty}^{+\infty} \lambda d\langle E_\lambda u | u \rangle, \quad (5.90)$$

wenn $\{E_\lambda\}$ die zugehörige Spektralschar ist; es muß folglich, $\varrho(\lambda)$ als Verteilungsfunktion einer Zufallsgröße a betrachtet, sein:

$$P(A; [\alpha, \beta]; u) = \langle (E_\beta - E_\alpha) u | u \rangle, \quad (5.91)$$

wenn links die Wahrscheinlichkeit dafür steht, daß die Messung der dem Operator A zugrunde liegende physikalische Größe a im Zustand u zwischen α und β liegt. Die Sprungpunkte von E_λ (d. h. λ mit $\lim_{\lambda \rightarrow \lambda-0} E_\lambda u \neq E_\lambda u$) sind gerade Werte von a , denen

stationäre Zustände (normierte Eigenvektoren von A) entsprechen, denen also eine Wahrscheinlichkeit $\langle (E_{\lambda-0} - E_\lambda) u | u \rangle > 0$ zukommt.

Als Beispiel studieren wir den Operator A , der der Ortskoordinate t zugeordnet wird: Ist der Definitionsbereich $D(A) = \{u \in L^2(\mathbf{R}) \mid tu(t) \in L^2(\mathbf{R})\}$, so ist A selbstadjungiert, jedoch nicht beschränkt [s. (5.35)]. A hat keine Eigenwerte, denn wäre für fast alle t

$$(A - \lambda I)u(t) = (t - \lambda)u(t) = 0, \quad (5.92)$$

so folgt (λ fest, reell) $u = o \in L^2(\mathbf{R})$. Trotzdem muß der Spektralsatz gelten. Wir erraten die zugehörige Spektralschar:

$$E_\lambda u = \chi(-\infty, \lambda) u, \quad (5.93)$$

wobei $\chi(-\infty, \lambda)$ die charakteristische Funktion (s. Text vor Def. 1.12) für das Intervall $(-\infty, \lambda)$ ist. Tatsächlich ist auch

$$\begin{aligned} \int_{-\infty}^{+\infty} \lambda \, d\langle E_\lambda u \mid u \rangle &= \int_{-\infty}^{+\infty} \lambda \, d \left(\int_{-\infty}^{+\infty} \chi(-\infty, \lambda) \bar{u}(\xi) u(\xi) \, d\xi \right) \\ &= \int_{-\infty}^{+\infty} \lambda \, d \left(\int_{-\infty}^{\lambda} |u(\xi)|^2 \, d\xi \right) \\ &= \int_{-\infty}^{+\infty} \lambda |u(\lambda)|^2 \, d\lambda \\ &= \int_{-\infty}^{+\infty} t |u(t)|^2 \, dt \\ &= \langle tu \mid u \rangle = \langle Au \mid u \rangle. \end{aligned} \quad (5.94)$$

$\varrho(\lambda)$ ist somit in diesem Beispiel von der Gestalt

$$\varrho(\lambda) = \int_{-\infty}^{\lambda} |u(\xi)|^2 \, d\xi$$

und offenbar monoton wachsend. Es kann keine Eigenwerte geben, denn

$$\begin{aligned} &\lim_{\delta \rightarrow +0} \|(E_\lambda - E_{\lambda-\delta}) u\|^2 \\ &= \lim_{\delta \rightarrow +0} \|\chi(-\infty, \lambda) - \chi(-\infty, \lambda - \delta)\| u\|^2 \\ &= \lim_{\delta \rightarrow +0} \|\chi(\lambda - \delta, \lambda) u\|^2 = \lim_{\delta=0} \int_{\lambda-\delta}^{\lambda} |u(\xi)|^2 \, d\xi = 0, \end{aligned} \quad (5.95)$$

weil das Lebesgue-Integral stetig von der Länge des Integrationsintervalls abhängt. Auch (5.88) läßt sich explizieren, denn wegen

$$\begin{aligned} D(A) &= \left\{ u \mid \int_{-\infty}^{+\infty} \lambda^2 \, d\|E_\lambda u\|^2 < \infty \right\} \\ &= \left\{ u \mid \int_{-\infty}^{+\infty} \lambda^2 \, d \int_{-\infty}^{\lambda} |u(\xi)|^2 \, d\xi < \infty \right\} \\ &= \left\{ u \mid \int_{-\infty}^{+\infty} \lambda^2 |u(\lambda)|^2 \, d\lambda < \infty \right\} \end{aligned}$$

erhält man gerade $tu(t) \in L^2$ wie in (5.34).

Quantenmechanisch gedeutet heißt (5.95), daß der beobachtbaren Größe A kein stationärer Zustand zukommt. Eine Wahrscheinlichkeit ist aber wie in (5.91) für jedes (abgeschlossene)¹⁾ Intervall (und einen Zustand $u \in D(A)$) angebar:

$$\begin{aligned} P(A; [\alpha, \beta], u) &= \langle (E_\beta - E_\alpha) u | u \rangle \\ &= \int_{-\infty}^{+\infty} \chi[\alpha, \beta] |u(\xi)|^2 d\xi \\ &= \int_{\alpha}^{\beta} |u(\xi)|^2 d\xi. \end{aligned} \quad (5.96)$$

Man erkennt so auch $0 \leq P \leq 1$, da u als Zustand stets mit $\|u\| = 1$ anzusehen war.

Als quantenmechanische Streuung der beobachtbaren Größe im Zustand u ergibt sich wegen (φ stückweise stetige Funktion)

$$\begin{aligned} \left\| \int_{-\infty}^{+\infty} \varphi(\lambda) d(E_\lambda u) \right\|^2 &= \int_{-\infty}^{+\infty} |\varphi(\lambda)|^2 d\|E_\lambda u\|^2: \\ \sigma^2(u|A) &= \|(A - \langle Au | u \rangle I) u\|^2 = \left\| \int_{-\infty}^{+\infty} (\lambda - \langle Au | u \rangle) d(E_\lambda u) \right\|^2 \\ &= \int_{-\infty}^{+\infty} |\lambda - \langle Au | u \rangle|^2 d\|E_\lambda u\|^2. \end{aligned} \quad (5.97)$$

Man erkennt weiter leicht, daß im Falle $A = \tilde{\mathfrak{H}}$ von 1.2.3. aus (5.96) wieder die Interpretation der $|c_j|^2$ wie im Text nach (1.94) folgt. Denn man setzt die Spektralschar aus (5.81) ein und erhält mit $\alpha = \lambda_j$ (Eigenwert) wie in (5.82)

$$P(\tilde{\mathfrak{H}}; [\lambda_{j-0}, \lambda_j], u) = -\langle (E_{j-0} - E_j) u | u \rangle = |\langle u | \varphi_j \rangle|^2 = |c_j|^2. \quad (5.98)$$

Schließlich kommen wir nochmals zur zeitabhängigen Schrödinger-Gleichung. Jetzt könnte im Beweis des Existenz- und Unitätssatzes auf die dort gemachte Einschränkung eines reinen Punktspektrums verzichtet werden, wenn mit der Spektraldarstellung eines Operators wie in (5.87) gerechnet wird.

Dem Operator e^{-itA} entspricht dann mit $\{E_\lambda\}$ als Spektralschar von A

$$e^{itA} y = \int_{-\infty}^{+\infty} e^{it\lambda} d(E_\lambda y). \quad (5.99)$$

5.2.4. Eigendifferentiale

Wenn man die Spektralschar zu einem Operator kennt, so ist also die quantenmechanische Interpretation vollständig durchführbar. Das Auffinden der Spektralschar ist aber in jedem Fall ein schwieriges Problem. Wir haben auch nur für den vollstetigen Fall allgemein und für den Ortsoperator die Spektralschar angegeben (5.81), (5.93). Als weiteres Hilfsmittel zur Analyse von Operatoren dienen die sog. *Eigenpakete*, die auch *Wellenpakete* oder *Eigendifferentiale* genannt werden. Sie stehen in enger Beziehung zur Spektralschar. Wir betrachten wieder einen selbstadjungierten (eventuell sogar nur symmetrischen) Operator A mit $D(A) \subseteq H$, $R(A) \subseteq H$. Wir fassen eine feste Stelle $\lambda_0 \in \mathbf{R}$ ins Auge. Gibt es ein vom Parameter $\lambda \in \mathbf{R}$ normstetig abhängiges Element $\phi_\lambda \in D(A)$,

¹⁾ Unwesentlich, da ϱ stetig ist für den Operator A .

welches für $\lambda = \lambda_0$ mit dem Nullelement zusammenfällt und dessen Bild unter A sich durch Stieltjes-Integration über die reelle Achse von λ_0 (Bezugspunkt) bis λ darstellen läßt:

$$A\Phi_\lambda = \int_{\lambda_0}^{\lambda} \mu \, d\Phi_\mu, \quad (5.100)$$

so heißt Φ_λ ($-\infty < \lambda < +\infty$) ein **Eigenpaket**. Für einen Eigenvektor φ zum Eigenwert λ von A würde $A\varphi = \lambda\varphi$ gelten; (5.100) bedeutet, daß $A\Phi_\lambda$ mit *allen* Φ_μ (von λ_0 bis λ) darstellbar ist.

Das Rechnen mit Stieltjes-Integralen gestattet, (5.100) umzuformen:

$$A\Phi_\lambda = \lambda\Phi_\lambda - \int_{\lambda_0}^{\lambda} \Phi_\mu \, d\mu, \quad (5.101)$$

und diese Gleichung zeigt, daß Φ_λ nicht Eigenvektor ist für festes λ . Man kann aber (5.101) auch in Differentialform schreiben, denn

$$\int_{\lambda_0}^{\lambda} \mu \, d(\Phi_\mu - \Phi_{\mu_0}) = (\lambda\Phi_\lambda - \lambda\Phi_{\lambda_0}) - \int_{\lambda_0}^{\lambda} \Phi_\mu \, d\mu, \quad (5.102)$$

und formal in 1. Näherung (das Integral rechts verschwindet von höherer als 1. Ordnung) ist

$$A(d\Phi_\lambda) = \lambda(d\Phi_\lambda), \quad (5.103)$$

woraus der Name **Eigendifferential** abgeleitet ist.

Der Zusammenhang zur Spektralschar $\{E_\lambda\}$ von A , wenn A selbstadjungiert ist, ist in folgender Weise herstellbar [16]: Man löst die Sprungstellen aus $\{E_\lambda\}$ heraus (denn Φ_λ ist normstetig, $\{E_\lambda\}$ hat Sprungstellen). Dies geschieht in folgender Weise: Da $\varrho(\lambda) = \langle E_\lambda u | u \rangle$ rechtsseitig stetig und monoton wachsend ist, läßt es sich als Summe einer Treppenfunktion $\tau(\lambda)$ und einer stetigen monotonen Funktion $\sigma(\lambda)$ darstellen [14]. Entsprechend sei

$$E_\lambda = T_\lambda + S_\lambda \quad (5.104)$$

mit

$$T_\lambda = \sum_{\lambda_j \leq \lambda} (E_{\lambda_j} - E_{\lambda_{j-0}}), \quad (5.105)$$

$$S_\lambda = E_\lambda - \sum_{\lambda_j \leq \lambda} (E_{\lambda_j} - E_{\lambda_{j-0}}) = E_\lambda - T_\lambda, \quad (5.106)$$

wobei die λ_j die Eigenwerte von A sind. Dann gilt [16]

Satz 5.5: Ist A selbstadjungierter Operator mit $D(A) \subseteq H$ und $u \in H$ ein beliebiges Element, so ist **S.5.5**

$$\Phi_\lambda = (S_\lambda - S_{\lambda_0}) u \quad (5.107)$$

ein Eigenpaket von A in $D(A)$ bezüglich λ_0 . Ist Φ_λ ein Eigenpaket von A bezüglich λ_0 , so kann es in jedem Intervall $-\infty < \alpha \leq \lambda \leq \beta < \infty$ mit $\alpha \leq \lambda_0 \leq \beta$ durch

$$\Phi_\lambda = (S_\lambda - S_{\lambda_0})(\Phi_\beta - \Phi_\alpha) \quad (5.108)$$

dargestellt werden.

Eigenpakete kann man im allgemeinen einfacher finden als die Spektralscharen. Ist z. B. A ein symmetrischer Differentialoperator mit leerem Punktspektrum¹⁾ und $\varphi(x, \lambda) \neq 0$ eine Lösungsschar von $Au = \lambda u$ mit $-\infty < \lambda < +\infty$, so muß $\|\varphi\|_{L_2} = \infty$ für jedes λ sein, denn sonst wäre φ Eigenfunktion für festes λ . Dann kann man versuchen, Eigenpakete in der Form

$$\Phi_\lambda = \int_{\lambda_0}^{\lambda} \varphi(x, \mu) \, d\mu \quad (5.109)$$

oder auch in Form eines Stieltjes-Integrals zu finden.

¹⁾ Zum Beispiel $A\psi = i\hbar^{-1} \frac{d}{dx} \psi$, dann ist $A\psi = \lambda\psi$ durch $\psi(x, \lambda) = e^{-i\hbar\lambda x}$ lösbar für jedes λ , aber $\|\psi\|_{L^2(\mathbf{R})} = +\infty$. Durch einen Integrationsprozeß ähnlich (5.109) werden dann „Wellenpakete“ in der Quantenmechanik aufgebaut. (siehe z. B. [10, Aufg. Nr. 2]).

Beispiel 5.3: Der schon bei (5.92) behandelte Multiplikationsoperator A hat leeres Punktspektrum, und jedes reelle λ gehört zum Stetigkeitsspektrum.

Die Spektralschar war nach (5.93)

$$\{E_\lambda\} = \{\chi(-\infty, \lambda)\}, \quad \lambda \in \mathbb{R} \quad (5.110)$$

wobei auch rechts Operatoren stehen.

Wir bestimmen ein Eigenpaket zu $\lambda_0 = 0$. Gemäß (5.104) ist $E_\lambda = S_\lambda$, gemäß (5.107) ist

$$\begin{aligned} \Phi_\lambda &= (S_\lambda - S_{\lambda_0})u = [\chi(-\infty, \lambda) - \chi(-\infty, 0)]u \\ &= +(\operatorname{sgn} \lambda) \chi(\lambda, 0)u, \end{aligned} \quad (5.111)$$

also

$$\Phi_\lambda = \begin{cases} +(\operatorname{sgn} \lambda)u & \text{für } x \text{ zwischen } 0 \text{ und } \lambda \\ 0 & \text{für die sonstigen } x. \end{cases} - \infty < \lambda < \infty. \quad (5.112)$$

Wir wenden zur Kontrolle A auf Φ_λ an; Φ_λ liegt offenbar in $D(A)$. Es sei $\lambda > 0$. Dann ist bei festem x

$$\begin{aligned} \int_{\lambda_0}^{\lambda} \mu \, d\Phi_\mu &= \lambda \Phi_\lambda - \int_0^{\lambda} \Phi_\mu \, d\mu = +\lambda \chi(\lambda, 0)u - \int_0^{\lambda} \chi(\mu, 0)u \, d\mu \\ &= +\lambda \chi(\lambda, 0)u - \left(\int_x^{\lambda} d\mu \right) \chi(\lambda, 0)u \\ &= +\lambda \chi(\lambda, 0)u - (\lambda - x) \chi(\lambda, 0)u \\ &= x \chi(\lambda, 0)u \\ &= A\Phi_\lambda. \end{aligned} \quad (5.113)$$

5.3. Weitere Anwendungen von Operatoren

Wir nennen hierzu noch Problemkreise, die in den naturwissenschaftlichen und ingenieurtechnischen Anwendungen große Bedeutung haben.

(1) Ergodensätze

Die Ergodentheorie befaßt sich mit der Existenz von Grenzwerten zeitlicher Mittel und Darstellungsmöglichkeiten dieser Grenzwerte. Ein einfaches Beispiel ist:

Zum Zeitpunkt $t = 0$ sei der Zustand eines Systems durch $x \in H$, $\|x\| = 1$ dargestellt. In den folgenden diskreten Zeitpunkten $n = 1, \dots$ werde der Zustand durch $U^n x$ beschrieben, wobei U ein in H unitärer Operator sei. Dann besagt eine Variante des Ergodensatzes von von Neuman, daß im Normsinne gilt

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^{N-1} U^n x = Px, \quad (5.114)$$

wobei P die (orthogonale) Projektion von x in den Raum der Eigenvektoren von U zum Eigenwert 1 ist. Links in (5.114) steht der Grenzwert des zeitlichen Mittels über die Zustände längs einer Trajektorie. Einen kurzen Beweis von F. Riesz findet man in [15]. Weiteres siehe z. B. in [29], [8, Bd. 2].

(2) Ritzsches Verfahren

Oft lassen sich Bestimmungsgleichungen für eine Größe x eines reellen separablen Hilbertraumes H in der Form $Tx = y$ ($y \in H$, y gegeben) schreiben, wobei T ein linearer stetiger selbstadjungierter halbbeschränkter Operator mit $D(T) = H$ ist. Dann ist $R(T) = H$, und es existiert für festes $y \in H$ eine eindeutige Lösung x^* (s. [40]).

x^* ist (s. 5.1.2.) eindeutige Minimalstelle von

$$F(x) = \langle Tx \mid x \rangle - 2\langle x \mid y \rangle \quad (5.115)$$

Jede Minimalfolge, d. i. eine Folge $\{x_n\}$ mit

$$\lim_{n \rightarrow \infty} F(x_n) = \inf \{F(x) \mid x \in H\} \quad (5.116)$$

erfüllt wegen (5.115)

$$\lim_{n \rightarrow \infty} F(x_n) = \inf_{x \in H} F(x) = \min_{x \in H} F(x) = F(x^*)$$

und konvergiert unter unseren Voraussetzungen an T sogar selbst (in der Norm von H) gegen x^* . Eine solche Minimalfolge gewinnt man folgendermaßen (Ritzsches Verfahren): Es sei $\{q_j\}$ ein vollständiges ONS in H . Man betrachtet nun eine echt aufsteigende Folge endlichdimensionaler Räume $H_1 \subseteq H_2 \subseteq \dots \subseteq H$, dabei ist H_n der lineare Teilraum von H , der aus den ersten n Elementen des ONS gebildet werden kann. Auf jedem H_n bestimmt man ein minimales Element x_n von F . x_n ist unter unseren Voraussetzungen an T eindeutig bestimmt, wird durch Lösen eines linearen Gleichungssystems gewonnen, und es gilt

$$F(x^*) \leq \dots \leq F(x_n) \leq F(x_{n-1}) \leq \dots \leq F(x_2) \leq F(x_1). \quad (5.117)$$

Die Folge $\{x_n\}$ ist eine Minimalfolge (Beweis s. [40]).

Ist T nur symmetrisch und halbbeschränkt aber nicht beschränkt, so ist $D(T)$ in der Regel nur eine dichte Teilmenge in H . Dann gilt: Für jedes $y \in R(T)$ existiert eindeutig eine Lösung $x_0 \in D(T)$ von $Tx = y$, und x_0 minimiert F über $D(T)$ eindeutig; jede Minimalfolge $\{x_n\}$ ($x_n \in D(T)$) von F konvergiert gegen x_0 (in der Norm von H). Das Ritzsche Verfahren zur Gewinnung einer Minimalfolge muß jetzt $D(T) \neq H$ berücksichtigen und ist in [40] erklärt. Für weitere Anwendungen und die Verfahren von Trefftz und Galerkin s. [42].

(3) Finite-Elemente-Methode

In (2) wurde zur Approximation der Lösung einer Operatorgleichung ein ONS benutzt. Es bildet eine *Basis* des benutzten Raumes. Als Basis kann man auch stückweise polynomiale Funktionen benutzen (finite Elemente), die in der modernen Numerik eine hervorragende Rolle spielen. Man gewinnt sie, indem das Grundgebiet $\bar{\Omega}$ des zugrunde liegenden Funktionenraumes geeignet trianguliert wird, und dann (etwa durch Vorgabe von Funktionswerten in den Eckpunkten und auf Punkten der Seitenflächen der durch die Triangulation entstandenen Teilbereiche) auf jedem Teilbereich Polynome festgelegten Grades bestimmt werden. Zum Verhalten des Ritzschen Verfahrens bei solchen Basisfunktionen, angewendet auf das erste Randwertproblem der elliptischen Differentialgleichung $\Delta u = -f$ in Ω , $u = 0$ auf $\partial\Omega$, $u \in W^{1,2}(\Omega)$, $f \in L_2(\Omega)$ vgl. [41/II, S. 47]. Weiteres z.B. in [51].

Literatur

- [1] *Achieser, N. I.; Glasmann, I. M.*: Theorie der linearen Operatoren im Hilbertraum. (Übers. a. d. Russ.) 6. Auflage. Berlin: Akademie-Verlag 1975.
- [2] *Adams, R. A.*: Sobolev Spaces. New York, San Francisco, London: Academic Press 1975.
- [3] *Багриновский, К. А.; Берланд, И. Я.*: Математические вопросы построения системы модели. Новосибирск, Наука 1976.
- [4] *Baumgärtel, H.*: Endlichdimensionale analytische Störungstheorie. Berlin: Akademie-Verlag 1972.
- [5] *Brehmer, S.*: Hilbert-Räume und Spektralmaße. Berlin: Akademie-Verlag 1979.
- [6] *Choquet-Bruhat, Y.; Dewitt-Morette, C.; Dillard-Bleick, M.*: Analysis, Manifolds and Physics. Amsterdam, New York, Oxford: North-Holland-Publ. Co. 1977.
- [7] *Collatz, L.*: Funktionalanalysis und Numerische Mathematik. Berlin, Göttingen, Heidelberg: Springer-Verlag 1964.
- [8] *Dieudonné, J.*: Grundzüge der modernen Analysis. Band 1. (Übers. a. d. Franz.) Berlin: VEB Deutscher Verlag der Wissenschaften 1971
- [9] *Dunford, N.; Schwartz, J. T.*: Linear Operators I. New York: Interscience Publishers 1958.
- [10] *Flügge, S.*: Rechenmethoden der Quantentheorie. Berlin, Heidelberg, New York: Springer-Verlag 1965.
- [11] *Focke, J.*: Distributionen und Heaviside-Kalkül. Wiss. Ztsch. d. KMU, Nat. wiss. Reihe, **11** (1962) 1, S. 627–639.
- [12] *Fučík, S.; Nečas, J.; Souček, V.*: Einführung in die Variationsrechnung. Teubner-Texte zur Mathematik. (Übers. a. d. Tschech.) Leipzig: BSB B. G. Teubner Verlagsgesellschaft 1977.
- [13] *Göpfert, A.*: Mathematische Optimierung in allgemeinen Vektorräumen. Leipzig: BSB B. G. Teubner Verlagsgesellschaft 1973.
- [14] *Günther, P.; Beyer, K.; Gottwald, S.; Wünsch, V.*: Grundkurs Analysis, Bände 2, 4. Leipzig: BSB B. G. Teubner Verlagsgesellschaft 1974.
- [15] *Халмош, П. Р.*: Конечномерные векторные пространства. (Перев. из англ.) Москва: Физматгит 1963.
- [16] *Hellwig, G.*: Differentialoperatoren der mathematischen Physik. Berlin, Göttingen, Heidelberg: Springer Verlag 1964.
- [17] *Kantorowitsch, L. W.; Akilow, G. P.*: Funktionalanalysis in normierten Räumen. (Übers. a. d. Russ.) Berlin: Akademie-Verlag 1964.
- [18] *Kluge, R.*: Theory of nonlinear operators, constructive aspects. Abh. d. Akad. d. Wiss. d. DDR, N 1, Berlin, Akademie-Verlag (1977) S. 125–163.
- [19] *Kolmogorow, A. N.; Fomin, S.*: Reelle Funktionen und Funktionalanalysis. (Übers. a. d. Russ.) Berlin: VEB Deutscher Verlag der Wissenschaften 1975.
- [20] *Kranc, G. M.; Sarachik, P. E.*: An application of functional analysis to the optimal control problem. Transactions of the ASME, Journal of Basic Engineering (Series D) **85** (1963) S. 143–150.
- [21] *Крейн, С. Г., и др.*: Функциональный анализ. СМБ Москва: Наука 1972.
- [22] *Kufner, A.; John, O.; Fučík, S.*: Function spaces. Prague: Academia 1977.
- [23] *Лионс, Ж.-Л.*: Оптимальное управление системами, описываемыми уравнениями с частными производными. (Перев. из франц.) Москва: Изд. Мир 1972.
- [24] *Ljusternik, L. A.; Sobolew, W. I.*: Elemente der Funktionalanalysis. (Übers. a. d. Russ.) 5. Auflage. Berlin: Akademie Verlag 1975.
- [25] *Lüthi, H. J.*: Komplementaritäts- und Fixpunktalgorithmen in der mathematischen Programmierung, Spieltheorie und Ökonomie, Lecture Notes in Economics and Math. Systems Nr. 129. Berlin, Heidelberg, New York: Springer-Verlag 1976.
- [26] *Макаров, В. Л.; Рубинов, А. М.*: Математическая теория экономической динамики и равновесия. Москва: Наука 1973.
- [27] *v. Mangoldt, H.; Knopp, K.*: Einführung in die Höhere Mathematik, IV. Band von F. Lössch. Stuttgart: S. Hirzel Verlag 1973.

- [28] *Michlin, S. G.*: Lehrgang der mathematischen Physik. (Übers. a. d. Russ.) Berlin: Akademie-Verlag 1975.
- [29] *Руд, М.; Саймон, Б.*: Методы современной математической Физики I, II, (перев. из англ.) Москва: Изд. Мир 1977.
- [30] *Riedrich, T.*: Vorlesungen über nichtlineare Operatorengleichungen. Teubner-Texte zur Mathematik. Leipzig: BSB B. G. Teubner Verlagsgesellschaft 1976.
- [31] *Riesz, F.; Sz.-Nagy, B.*: Vorlesungen über Funktionalanalysis. (Übers. a. d. Ung.) 3. Auflage. Berlin: VEB Deutscher Verlag der Wissenschaften 1973.
- [32] *Rolewicz, S.*: Funktionalanalysis und Steuerungstheorie. (Übers. a. d. Poln.) Berlin, Heidelberg, New York: Springer-Verlag 1976.
- [33] *Шулов, Г. Е.*: Математический анализ, специальный курс. Москва: Физматгиз 1967.
- [34] *Schmeidler, W.*: Integralgleichungen mit Anwendungen in Physik und Technik, Bd. I, Lineare Integralgleichungen. Leipzig: Geest u. Portig 1955.
- [35] *Smirnow, W. I.*: Lehrgang der höheren Mathematik, Teile II, IV, V. (Übers. a. d. Russ.) 6. Auflage. Berlin: VEB Deutscher Verlag der Wissenschaften 1973.
- [36] *Triebel, H.*: Höhere Analysis. Berlin: VEB Deutscher Verlag der Wissenschaften 1972.
- [37] Wörterbuch der Kybernetik, 4. Auflage, Hrsg. G. Klaus, H. Liebscher. Berlin: Dietz-Verlag 1976.
- [38] *Wladimirow, W. S.*: Gleichungen der mathematischen Physik. (Übers. a. d. Russ.) Berlin: VEB Deutscher Verlag der Wissenschaften 1971.
- [39] *Владимиров, В. С.*: Обобщённые функции в математической физике. Москва: Изд. Наука 1976.
- [40] *Wulich, B. S.*: Einführung in die Funktionalanalysis, Teil 2. Leipzig: B. G. Teubner Verlag 1962.
- [41] *Zeidler, E.*: Vorlesungen über nichtlineare Funktionalanalysis, Bände I–III. Teubner-Texte zu Mathematik. Leipzig: BSB B. G. Teubner Verlagsgesellschaft 1976/77.

Ergänzende Literaturangaben

- [42] *Bronstein, I. N.; Semendjajew, K. A.*: Taschenbuch der Mathematik. 19. Auflage. Hrsg. G. Grosche u. V. Ziegler. Ergänzende Kapitel. Leipzig: BSB B. G. Teubner Verlagsgesellschaft 1979.
- [43] *Gelfand, I. M.; Schilow, G. E.*: Verallgemeinerte Funktionen (Distributionen) Bände I–III (Übers. a. d. Russ.). Berlin: VEB Deutscher Verlag der Wissenschaften; I (2. Aufl.), II (2. Aufl.), III (1. Aufl) 1967, 1969 bzw. 1964.
- [44] VII. Internationale Konferenz über Nichtlineare Schwingungen 8.–13. September 1975 in Berlin. Bände I, 1; I, 2. Abh. d. Akad. d. Wiss. d. DDR N 3, N 4 Berlin: Akademie-Verlag 1977.
- [45] *Ioffe, A. D.; Tichomirow, V. M.*: Theorie der Extremalaufgaben (Übers. a. d. Russ.). Berlin: VEB Deutscher Verlag der Wissenschaften 1979.
- [46] *Jürgens, K.*: Lineare Integraloperatoren, Stuttgart: B. G. Teubner 1970.
- [47] *Kecs, W.; Teodorescu, P. P.*: Introducere in teoria distributiilor cu aplicatii in tehnică. Bucuresti: Editura Tehnică 1975.
- [48] *Kluge, R.*: Nichtlineare Variationsungleichungen und Extremalaufgaben. Berlin: VEB Deutscher Verlag der Wissenschaften 1979.
- [49] *Красносельский, М. А., и др.*: Приближенное решение операторных уравнений. Москва: Наука 1969.
- [50] *Шихов, С. Б.*: Вопросы математической теории реакторов. Москва: Атомиздат 1973.
- [51] *Strang, G.; Fix, G. J.*: An analysis of the finite element method. Englewood Cliffs: Prentice-Hall 1973.

Register

Abbildung, adjungierte 84

–, duale 84

–, injektive 76

–, kontrahierende 29, 110

–, lineare 72

–, surjektive 75

–, unitäre 103

– von endlichem Rang 82

abgeschlossene Hülle 42

Ableitung, distributionelle 96

–, schwache 53

–, verallgemeinerte 53

Ableitungsbegriffe 103

Abschließung einer Menge 42

absolut stetig 115

Abstand 9

Abweichung, mittlere quadratische 14

Adjungierte 88

Anfangswertproblem, abstraktes Cauchysches 118

Arzelà-Ascoli, Satz von 50, 83

Ausgangssignal 31

Banach, Satz von 76, 81

Banach, Stefan 45

Banachraum 45

–, reflexiver 64

Bidualraum 60

Cauchy-Folge 42

Definitionsbereich 73

Dipol 96

Dirac-Distribution 21

Dirichletsches Prinzip 107

Distribution 20, 95

–, Multiplikation einer, mit einer Funktion 97

–, reguläre bzw. singuläre 95

Distributionen, Faltung zweier 97

–, Konvergenz einer Folge von 96

– schwachen Wachstums 100

Distributionenfolge, schwache Konvergenz einer 23

Distributionsableitung 53

divergent 67

Dreiecksungleichung 9

Dualraum 59

Eigendifferential 128, 129

Eigenpaket 128, 129

Eigenschwingungen einer Membran 124

Eigenvektor 77

Eigenwert 77

Eigenwerte, Aufspaltung der 94

Eigenwertproblem 25

Eingangssignal 31

Entwicklungssatz 68

Ergodentheorie, -sätze 130

Erreichbarkeit 36

Erwartung, mathematische 28

Erweiterung 61

–, maximale 113

ess sup 50

Eulersche Gleichung 107

Fixpunkt 29

Fixpunktsatz von Banach 29, 110

– – Kakutani 110

Folge, konvergente 40

Folgenräume 54

Fortsetzbarkeit, stetige 49

Fourierentwicklung 68

Fourierkoeffizienten 17, 67

Fouriertransformation 100

Fouriertransformierte 101

Fréchet-Ableitung 104

Fréchet-Differential 104

Fredholmsche Alternative 85

Friedrichs, Satz von 113

–, Ungleichung von 122

Friedrichssche Erweiterung 116

Funktion, charakteristische 38

–, meßbare 37

– schwachen Wachstums, verallgemeinerte 101

–, summierbare 37

–, verallgemeinerte 95

Funktional, beschränkt konvexes 108

–, beschränktes lineares 57

–, lineares 55

–, stark konvexes 108

–, stetiges lineares 56

Funktionenraum 48

Gateaux-Differential 105

Gewichtsfunktion 31

Gleichungspunkt des n -Personenspiels 110

gleichgradigstetig 50

Greensche Funktion 23, 32

Grenzwert 40

Grundlösung 99

Grundraum 20

Hahn-Banach-Theorem 61

Heaviside-Funktion 19

Hermiteische Funktion 26

- Hermitesche Funktionen, normierte 68
- Hermiteischer Differentialoperator 26
- Hilbertraum 64
- Hilbertraum, Funktion mit Werten in einem 119
- , separabler 92
- Hölder, Ungleichung von 51
- Hyperebene 61
- , strikt trennende bzw. trennende 62
- Hyperparaboloid 109
- Impulsantwortfunktion 31
- Indikatorfunktion 38
- Integralgleichung 81
 - zweiter Art, lineare 82
- Integraloperator 82
 - , linearer 75
- Isometrie 44
- Iterationsverfahren 29
- kanonische Einbettung 60
- Kern 73, 82
 - , adjungierter 89
 - eines Integraloperators 75
 - , transponierter 84
- Kompaktheit 43
 - , schwache 64
- Kompaktheitskriterium 50
- konvergent im Sinne der Operatornorm 80
- Konvergenz, schwache 63
- konvergiert im quadratischen Mittel 71
- Koordinatenmatrix 87
- Kugel, abgeschlossen bzw. offene 41
- Laplace-Operators, Fundamentallösung des 103
- Lebesgue-Integral 37
- Lebesgue-Maß 38
- Lebesgue-Räume 50
- Legendresche Polynome, normierte 68
- linear abhängig bzw. unabhängig 9
- linearer Teilraum 9
- Linearform 55
- Lösung, stetige Abhängigkeit der 90
- , verallgemeinerte 99, 116
- Maß 38
 - Null 36
- Matrix, orthogonale 44
 - , transponierte 84
 - , unitäre 44
- Matrixnormen 30
- Maximum-Norm 48
- Menge, abgeschlossene 40
 - , meßbare 38
 - , offene 41
 - , zusammenhängende 48
- mes A 38
- Metrik 9
 - , euklidische 11
- Minimalfolge 131
- Minkowski, Ungleichung von 51
- Multiindex 49
- Nadelimpuls 31
- Neumannsche Reihe 81
- Norm 10, 57
 - , äquivalente 46
 - einer linearen beschränkten Abbildung 30
 - , euklidische 30, 46
- normerhaltende Fortsetzung 61
- Normisomorphismus 46
- Nullelement 9
- Nullraum 73
- Nullvektor 9
- Operator, adjungierter 88, 90
 - , beschränkter linearer 74
 - , dualer 84
 - , halbbeschränkter 112
 - , hermitescher 90
 - , identischer 73
 - , inverser 74
 - , kompakter 82
 - , linearer 72
 - , selbstadjungierter 91
 - , – vollstetiger 92
 - , stark positiv definit 112
 - , stetiger linearer 81
 - , symmetrischer 90
 - , vollstetiger linearer 82
- Operatoren, Fortsetzung halbbeschränkter 113
- , Schar der 124
- Operatorfunktion 77
- Operatorpolynom 76
- Operatorpotenzreihe 76
- Operators, Erweiterung eines 112
 - , Spektrum eines vollstetigen 83
 - , untere Grenze eines 112
- orthogonal 13
- orthogonale direkte Summe 68, 69
 - Teilräume 69
- Orthogonalentwicklung 68
- orthogonales Komplement 68
- Orthogonalsystem 13, 65
- Orthonormalsystem (ONS) 13, 66
- Oszillator, harmonischer 24
- Parallelogrammgleichung 13, 64
- Parsevalsche Gleichung 68
- Polyome, trigonometrische 14
- Pär-Hilbertraum 11

- Produktraum 47
- Projektion 70
- Punktspektrum 77
- quadratische Form
- Quantenhypothese von Planck 27
- Quotientenraum 47
- Rang 73
- Raum 9
 - , algebraisch dualer 59
 - der beschränkten Folgen 54
 - – finiten Folgen 54
 - – konvergenten Folgen 54
 - – lokal integrierbaren Funktionen 51
 - – Nullfolgen 54
 - – rasch fallenden Folgen 54
 - – zur p -ten Potenz summierbaren Folgen 54
 - , dualer 59
 - , kompakter metrischer 43
 - , konjugierter 59
 - , linearer 9
 - , – metrischer 10
 - , metrischer 9
 - , normierter 10
 - , unitärer 11
 - , vollständiger metrischer 43
- Räume integrierbarer Funktionen 50
 - , isometrische 44
- reflexiv 61
- Reihe, unendliche 66, 67
- Residualspektrum 78
- Resolvente 75, 123
- Resolventengleichung 76
- Resolventenmenge 75
- Restspektrum 78
- Riesz, F. 59
- Ritzsches Verfahren 130, 131
- Satz von E. Hellinger und O. Toeplitz 90
- Säkulargleichungen 94
- Schauderscher Fixpunktsatz 108
- Schmidtsches Orthogonalisierungsverfahren 66
- Schrödingler-Gleichung 118
- Schwarz-Bunjakowskische Ungleichung 12
- Skalarprodukt 11
- Skalarproduktes, Stetigkeit des 64
- Sobolew-Räume 52
- Sobolewischer Einbettungssatz 103
- Sochzki, Formeln von 97
- Spaltenbetragssummennorm 30
- Spannungsstoß 14
- Spektralabbildungssatz 77
- Spektralnrm 30
- Spektralradius 76, 124
- Spektralsatz 125
- Spektralschar 125
- Spektralzerlegung 125
- Spektrum 76
 - , kontinuierliches 78
- Steuerung, zeitoptimale 31
- Störfunktion 82
- Störungstheorie, Formeln der 93
- Stoßantwort 31
- Stoßvorgang 17
- Streuung 120
- supp 49
- Teilmenge, dichte 44
- Träger 49, 95
- Trennungssätze 61, 62
- Übertragungsglieder, lineare zeitinvariante 31
- Übertragungsoperator 31
- Umkehroperator, stetiger 81
- Unschärferelation, Heisenbergsche 120, 121
- Variation, erste 105, 106
- Vektorraum 9
 - , komplexer bzw. reeller 9
 - , topologischer 11
- Vervollständigung 44
- vollständig 43, 68, 97
- Vollständigkeitsrelation 28, 123
- Vollständigkeit 82
- vrai max 50
- Wellenpakete 128
- Wertebereich 73
- Zeilenbetragssummennorm 30
- Zerlegung der Einheit 124
- Zustand, quantenmechanischer 25, 28
 - , stationärer 25