Prof. Dr. H. Roos **WS 2002/03**

Mathematische Methoden der Theoretischen Physik

4stündige Vorlesung
mit 1stündigen Übungen

Inhaltsverzeichnis

Kapitel 0

Einleitung

Die meisten von Ihnen, so nehme ich an, studieren Physik im dritten Semester. Sie haben nun in den ersten zwei Semestern viel Mathematik gehört; erst jetzt fängt die Theoretische Physik an mit der Mechanik, und nun will ich Sie mit noch mehr Mathematik behelligen. Der Grund ist, daß Ihnen mit den Kenntnissen, die ich Ihnen vermitteln will, die Theoretische Physik, die in den kommenden Semestern auf Sie zukommt, mehr Spaß machen soll, weil Sie sie dann sehr viel besser verstehen, als wenn Sie die notwendigen mathematischen Methoden nur in der Kurzfassung kennenlernen, die in der jeweiligen Vorlesung gebracht werden kann.

Die Vorlesung Mathematische Methoden ist ein Gemischtwarenladen: sie berührt verschiedene Gebiete, die alle nicht in jeder Hinsicht exakt und eingehend behandelt werden. Viele Beweise werde ich nur andeuten oder Ihnen ganz als Übungsaufgabe überlassen; manches müssen Sie mir glauben oder, wenn Sie es genau wissen wollen, in der mathematischen Literatur nachlesen. (Jedes dieser Gebiete ist den Mathematikern eine mindestens einsemestrige Vorlesung wert.)

Was behandelt werden soll:

- Elemente der Funktionentheorie.
- Lineare topologische Räume, Hilberträume; Lebesgue-Integral.
- Fourier- und Laplace-Transformation.
- Distributionen.
- Operatoren im Hilbertraum; Spektraldarstellung.

Zur physikalischen Motivation

1. Messungen ergeben reelle Zahlen; wie nützlich aber komplexe Zahlen in der Physik sind, werden Sie im Laufe Ihres Studiums bald bemerken, wenn Sie es bisher noch nicht gesehen haben. Ich will im folgenden Kapitel einige Anwendungen skizzieren. Wir brauchen komplexe Zahlen und Funktionen unbedingt in späteren Kapiteln: bei der Fourier-Transformation und bei den komplexen Hilberträumen, die für die Quantenmechanik praktisch unerläßlich sind.

Generell:

— Komplex zu rechnen ist oft sehr bequem in der Physik; der Übergang zu den letztlich interessierenden reellen Größen ist meist ziemlich trivial.

— Analytische Funktionen, d.h. komplexwertige Funktionen einer komplexen Veränderlichen, die komplex differenzierbar sind, sind mathematisch „schön“; die Theorie dieser Funktionen, die Funktionentheorie, ist ein starkes mathematisches Werkzeug. In der Quantenfeldtheorie wird es wesentlich eingesetzt.

2. Warum interessieren lineare Räume? $\mathbb{R}^3$: trivial. $\mathbb{R}^n$: tritt in der Mechanik auf, z. B. $n = 3N$, wenn N Teilchen betrachtet werden.

Kann man (z. B. für kontinuierliche Systeme, $N \to \infty$) auch $\mathbb{R}^\infty$ definieren? Antwort: ja, gewissermaßen; man benutzt ihn zwar im allgemeinen nicht für kontinuierliche Systeme, aber man kommt dann zum Konzept des Hilbertraums, der in der Quantenmechanik eine zentrale Rolle spielt.

Viele Differentialgleichungen in der Physik sind linear; die Lösungsmenge einer homogenen linearen Gleichung ist ein linearer Raum.
Besipiel 1: Gewöhnliche lineare Dgl. n-ter Ordnung; Dimension des Lösungsraumes $= n$; z. B. wird in der Mechanik die Bewegung von N Teilchen durch $3N$ Differentialgleichungen zweiter Ordnung beschrieben, d. h. $n = 6N$.
Beispiel 2: Wellengleichung in einer Dimension (Notation $\partial_t \equiv \frac{\partial}{\partial t}$):

$$(\frac{1}{c^2}\partial_t^2 - \partial_x^2)f(t,x) = 0\,.$$

Sei $g(\xi)$ eine beliebige zweimal differenzierbare Funktion, dann ist $f(t,x) = g(x - ct)$ eine Lösung der Wellengleichung, Beweis durch Nachrechnen. Die Lösungsmannigfaltigkeit ist unendlich-dimensional.

Es interessieren also unter anderem diverse Funktionenräume von (komplexwertigen)

Funktionen; sie bilden lineare Räume über $\mathbb{C}$.

Die $n \times n$-Matrizen liefern lineare Abbildungen $\mathbb{R}^n \longrightarrow \mathbb{R}^n$; wenn man mit linearen Abbildungen allgemeinerer linearer Räume zu tun hat, spricht man von „linearen Operatoren".

Einfache Beispiele von linearen Operatoren auf Funktionenräumen:
a) Differentiation; Differentialoperatoren (z. B. der aus der obigen Wellengleichung).
b) Integraloperatoren, z. B. $f(x) \mapsto \int k(x-y)f(y)\,dy$; sie werden Ihnen u. a. bei der Lösung von Differentialgleichungen mithilfe von Greenschen Funktionen begegnen; physikalisches Beispiel: Beschreibung des Durchgangs einer Lichtwelle durch ein inhomogenes Medium mit dem Huygens'schen Prinzip.
c) Projektionsabbildungen, z. B. $f(x) \mapsto \chi_I(x)f(x)$, wobei χ_I die charakteristische Funktion des Intervalls I ist: $= 1$, wenn $x \in I$ und sonst $= 0$. „Projektion" heißt: eine wiederholte Anwendung dieser Operation bringt nichts Neues, die Projektion einer Projektion ist die ursprüngliche Projektion.

Für uns am wichtigsten: lineare Operatoren auf einem Hilbertraum. Sie sind für die Quantenmechanik unerläßlich.

Das ist ein Kapitel, das mathematische Mittel bereitstellt ohne physikalische Anwendungsbeispiele; doch diese Mathematik ist Voraussetzung für vieles, was in den späteren Kapiteln folgt, in denen physikalische Beispiele vorkommen werden.

3. Fourier-Transformation: tritt als Verallgemeinerung der Fourier-Analyse periodischer Funktionen überall in der Physik auf, Frequenzanalysen kommen überall vor. Sie ist wichtiges Hilfsmittel zur Lösung von Differentialgleichungen. Laplace-Transformation: ebenso; sie ist in gewisser Weise eine einfache Verallgemeinerung der Fourier-Transformation.

4. Distributionen: anderer Name: „verallgemeinerte Funktionen". Typischer, in der Physik oft auftretender Vertreter: δ-„Funktion". Sie ist keine Funktion, aber man kann die Funktionen derart verallgemeinern, daß man eine mathematisch wohldefinierte Menge von Objekten erhält, zu denen dann auch die δ-Funktion gehört, die man sogar differenzieren kann. Sie ist ebenfalls ein wichtiges Hilfsmittel zur Lösung von Gleichungen der Physik.

5. Operatoren im Hilbertraum: Standardwerkzeug der Quantenmechanik.

Zur Numerierung der Sätze und Gleichungen: Innerhalb der einzelnen Abschnitte werden Definitionen, Sätze, Theoreme u. s. w. zusammen mit den Gleichungen durchnumeriert, Definitionen etc. mit Kapitel- und Abschnittsnummer, die Gleichungen nur mit der Abschnittsnummer: z. B.

Satz 2.5.13 ... ...

$$a = b \tag{5.14}$$

Beispiel 2.5.15 ... .

Wenn Gleichungen aus früheren Kapiteln zitiert werden, dann mit Kapitelnummer, z. B. (I.3.23).

Literatur
zu den
Mathematischen Methoden der Physik

Literatur zu Kapitel 0

W. I. Smirnov, *Lehrgang der höheren Mathematik, Teil I, Kap. 6,*

Behnke, Sommer, *Funktionentheorie,*
ausführliches Standardlehrbuch

Peschl, *Funktionentheorie I,* Bi HTB 131

H. Fischer, H. Kaul, *Mathematik für Physiker 1,* Teubner Studienbuch 1998
enthält auh Vektoralgebra und -analysis und einges zu Differentialgleichungen

Literatur zu den übrigen Kapiteln

S. Großmann, *Funktionalanalysis,* Aula-Verlag, 4. Auflage, 1988
Paperback, geschrieben „im Hinblick auf Anwendungen in der Physik“

Hewitt, Stromberg, *Real and abstract analysis,* Springer 1969
umfangreiches Nachschlagewerk, gut für Maßtheorie und Lebesgue-Integral

N.I. Achieser, I.M. Glasmann, *Lineare Operatoren im Hilbertraum,* 8. Auflage 1981

F. Riesz, B. Sz.-Nagy, *Vorlesungen über Funktionalanalysis,* 1956
Standardwerk, enthält auch Abschnitte über Lebesgue-Integrale und Fourier-Transformation

J. Weidmann, *Lineare Operatoren in Hilberträumen,* 1976

Reed, Simon, *Methods of Mathematical Physics,*
vierbändig, anspruchsvoll; hier interessant: Vol II: Fourier analysis, self-adjointness

Gelfand-Shilov, *Verallgemeinerte Funktionen,*
Theorie der Distributionen, fünfbändig, hier interessant: Bd. I, II

H. Fischer, H. Kaul, *Mathematik für Physiker 2,* Teubner Studienbuch 1998

Paperback, ca. 750 S., überdeckt weitgehend den Stoff der Vorlesung und bringt noch mehr, z. B. partielle Differentialgleichungen

Ph. Blanchard, E. Brüning, *Distributionen und Hilbertraum-Operatoren,* Springer 1993

Kapitel 1

Lineare Räume

1.1 Lineare topologische Räume

Bekannt: linearer Raum: Menge X, in der eine Addition erklärt ist und die Multiplikation mit Größen aus einem Körper K, für uns üblicherweise $\mathbb{R}$ oder $\mathbb{C}$, und in der es ein Nullelement $\mathbf{0}$ gibt, so daß für $x, y \in X$ und $a, b \in K$

$$x + y \in X,\ a(x+y) = ax + ay,\ (a+b)x = ax + bx,\ 0x = a\mathbf{0} = \mathbf{0}$$

gilt. Ich schreibe im Folgenden wie üblich nur 0 statt $\mathbf{0}$.

Beispiele: $\mathbb{R}^n$, $\mathbb{C}^n$, Lösungsmenge einer homogenen linearen Gleichung, Menge $\mathcal{C}_0^0(\mathbb{R})$ aller stetigen komplexwertigen beschränkten Funktionen über der reellen Achse, die im Unendlichen verschwinden.
„Linearer Raum“ und „Vektorraum“ verwende ich als Synonyme; die Elemente eines linearen Raumes nenne ich deshalb auch Vektoren.

Im Weiteren werden uns immer wieder verschiedene Funktionenräume begegnen, u. a. auch die folgenden:

Definition 1.1.1 *i)* $\mathcal{C}^p(\mathbb{R}^n)$ *bezeichnet die Menge der komplexwertigen p-mal stetig partiell differenzierbaren Funktionen über dem Raum* $\mathbb{R}^n$.
ii) $\mathcal{C}_b^p(\mathbb{R}^n) = \{f \in \mathcal{C}^p(\mathbb{R}^n);\ f \text{ beschränkt}\}$.
iii) $\mathcal{C}_0^p(\mathbb{R}^n) = \{f \in \mathcal{C}_b^p(\mathbb{R}^n);\ f \to 0 \text{ für } |x| \to \infty\}$.
iv) $\mathcal{C}_{00}^p(\mathbb{R}^n) = \{f \in \mathcal{C}^p(\mathbb{R}^n);\ \operatorname{supp} f \text{ kompakt}\}$.

$\operatorname{supp} f$ meint: Träger von f, d. h. den Abschluß der Menge $\{x \in \mathbb{R}^n;\ f(x) \neq 0\}$. Definition 1.1.1 gilt auch für $p = \infty$. Wenn $p = 0$ ist, lasse ich den oberen Index manchmal weg.

Alle hier definierten Funktionenräume sind offensichtlich lineare Räume.

„Topologischer“ Raum: ich gebe hier keine vollständige Definition dessen, was „Topologie“ heißt; für uns soll dieser Begriff synonym sein mit der Aussage, daß ein Konvergenzbegriff in X definiert ist, daß also bekannt ist, wann ein Element „in einer kleinen Umgebung“ eines anderen liegt; und ich werde immer nur angeben, was im jeweiligen Fall „Konvergenz“ heißen soll, ohne zu begründen, daß diese Definition alle notwendigen Kriterien erfüllt.

Warum interessiert uns Konvergenz z. B. in irgend welchen Funktionenräumen? Antwort: fast alle Probleme der Physik führen auf Gleichungen, die nicht explizit exakt lösbar sind. Folge: man macht Näherungen; und dann muß man wohl wissen, ob die gefundene Näherungslösung der wahren Lösung in irgendeinem Sinne nahe ist bzw. ob eine Folge von Näherungen tatsächlich konvergiert. Wesentlich: je nach Fragestellung braucht man verschiedene Topologien.
Beispiel 1: Dichte $\rho(T)$ eines Stoffes als Funktion der Temperatur beim Phasenübergang flüssig–gasförmig, beschrieben durch eine bei der Phasenübergangstemperatur T_0 unstetige Funktion: sie springt bei steigender Temperatur bei T_0 auf einen sehr viel höheren Wert. Die Berechnung mit den Methoden der Statistischen Mechanik liefert eine Approximation durch eine steil bei T_0 ansteigende stetige Funktion $\rho_1(T)$. Hier darf man nicht nach $\sup |\rho_1(T) - \rho(T)|$ als Maß für die Konvergenz fragen, sonst bekommt man nie eine gute Näherung; sinnvoller wäre $\int |\rho_1(T) - \rho(T)|\, dT$.
Beispiel 2: Wenn ich die Höhe $h(t)$ einer startenden Rakete als Funktion der Zeit näherungsweise berechnen will, wäre es unbefriedigend, nur $\sup_{t\in\mathbb{R}} |h_n(t) - h(t)| \to 0$ zu fordern; denn dann könnten die Näherungen im Prinzip völlig unsinnige Geschwindigkeiten oder Beschleunigungen ergeben[1]. Hier möchte ich Konvergenz so interpretiert sehen, daß auch die ersten beiden Ableitungen der Näherungen den wahren Ableitungen nahe sind.

Eine einfache Möglichkeit, Konvergenz zu definieren: wenn auf X eine „Metrik“, d. h. eine Abstandsfunktion gegeben ist.
Notationen: $\mathbb{R}^+$ bezeichnet die positiven reellen Zahlen, $\mathbb{R}_0^+$ die nicht-negativen, d. h. einschließlich der 0. Überstreichung meint, daß auch die Werte ∞ zugelassen sind: $\overline{\mathbb{R}}_0^+ = \mathbb{R}_0^+ \cup \{+\infty\}$, $\overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty, -\infty\}$.

Definition 1.1.2 *Eine nicht-negative Funktion $d : X \times X \longrightarrow \mathbb{R}_0^+$ heißt Metrik auf X, wenn für alle $x, y, z \in X$ gilt:*
i) $d(x, y) = d(y, x)$,
ii) $d(x, y) = 0 \Longrightarrow x = y$,
iii) $d(x, y) \le d(x, z) + d(z, y)$ („Dreieicksungleichung“).

[1] Die erste Ableitung von h ergibt die Vertikalgeschwindigkeit, die zweite Ableitung die Vertikalbeschleunigung

Damit können wir durch $\{y \in X;\ d(x,y) < r\}$ Umgebungen eines Punktes x („Kreise um x“) definieren und sagen: x_n konvergiert gegen x, wenn $\lim_{n\to\infty} d(x_n, x) = 0$.
Einfache Beispiele: $\mathbb{C}$ mit $d(z_1, z_2) = |z_1 - z_2|$; $\mathcal{C}_b^0(\mathbb{R})$ mit $d(f,g) = \sup_{x\in\mathbb{R}} |f(x) - g(x)|$.

In diesen beiden Beispielen haben wir tatsächlich nicht nur metrische, sondern normierte Räume vor uns.

Definition 1.1.3 *Eine nicht-negative Funktion $x \in X \mapsto \|x\| \in \mathbb{R}_0^+$ heißt Norm auf dem linearen Raum X über dem Körper $\mathbb{C}$, wenn für alle $x, y \in X$ und alle $a \in \mathbb{C}$ gilt:*
i) $\|ax\| = |a|\,\|x\|$,
ii) $\|x\| = 0 \Longrightarrow x = 0$,
iii) $\|x + y\| \leq \|x\| + \|y\|$ *(Dreiecksungleichung).*

Einfache Beispiele für Normen sind der Betrag auf der Menge der komplexen Zahlen und der Betrag eines Vektors in $\mathbb{R}^n$. Weiteres Beispiel:

$$\|f\|_\infty := \sup_{x\in\mathbb{R}} |f(x)| \tag{1.4}$$

macht $\mathcal{C}_b^0(\mathbb{R})$ zu einem normierten Raum (Ü).
Zu einer Norm können wir sofort eine Metrik definieren: $d(x,y) := \|x - y\|$ (nur die Dreiecksungleichung ist zu überprüfen). Damit haben wir einen Konvergenzbegriff in normierten Räumen, der im Falle der bekannten Räume $\mathbb{R}^n$ und $\mathbb{C}$ mit dem üblichen Konvergenzbegriff übereinstimmt.

1.1.5 Anmerkung: Man kann auf $\mathbb{R}^n$ auch andere Normen definieren, z. B. $\|x\|_\infty = \max_{1\leq i\leq n} |x_i|$ oder $\|x\|_p = (\sum_{i=1}^n |x_i|^p)^{1/p},\ p \geq 1$. (Hier ist $x = (x_1, \dots, x_n) \in X$.)

Man kann sich vorstellen, daß man im $\mathbb{R}^n$ nur den Abstand der k-ten Koordinaten mißt: $p_k(x - y) := |x_k - y_k|$. Die Funktion $p_k : X \to \mathbb{R}_0^+$ ist nur eine „Halbnorm“, keine Norm, weil sie die zweite Bedingung von Definition 1.1.3 nicht erfüllt. Doch alle n Funktionen $p_k,\ k = 1, \dots, n$, zusammen genommen liefern uns ein vernünftiges Maß für Konvergenz. Das ist analog zum obigen Beispiel der Raketenbahn, bei der gleichzeitig auch Geschwindigkeit und Beschleunigung approximiert werden sollen: wir können dafür drei Halbnormen definieren: $p_k(f) := \sup |f^{(k)}(t)|,\ k = 0, 1, 2$, und verlangen, daß $f_n \to 0$ heißt: $p_k(f_n) \to 0\ \forall k = 0, 1, 2$. Wir brauchen später einen Funktionenraum mit einer ähnlichen Topologie; ich will ihn gleich hier definieren.

Definition 1.1.6 *Für $f \in \mathcal{C}^\infty(\mathbb{R})$ sei*

$$p_{mn}(f) := \sup_{x\in\mathbb{R}} |x^m\, f^{(n)}(x)|\,.$$

Raum der „temperierten Funktionen“:

$$\mathcal{S}(\mathbb{R}) := \{f \in \mathcal{C}^\infty(\mathbb{R});\ p_{mn}(f) < \infty\ \forall m, n \in \mathbb{N} \cup \{0\}\}\ .$$

Eine Folge $\{f_n;\ f_n \in \mathcal{S}(\mathbb{R})\}_{n\in\mathbb{N}}$ *heißt konvergent, wenn sie bezüglich aller Halbnormen* p_{mn} *eine Cauchy-Folge ist.*

Diese Funktionen fallen also samt ihren Ableitungen „stark“ ab, d. h. stärker als jede Potenz. Der Raum $\mathcal{S}(\mathbb{R})$ ist ein linearer Raum über $\mathbb{C}$, und die Funktionen p_{mn} sind tatsächlich Halbnormen; davon kann man sich leicht überzeugen.
Die Definition läßt sich sofort auf höhere Dimensionen erweitern. Man betrachtet Funktionen aus $\mathcal{C}^\infty(\mathbb{R}^N)$ und interpretiert die Indizes n und m von p_{mn} als Multiindizes: $n = (n_1, \ldots, n_N)\,, m = (m_1, \ldots, m_N)$ und setzt $x^n := \prod_{k=1}^N x^{n_k}$ und $f^{(n)} := \partial_{x_1}^{n_1} \ldots \partial_{x_N}^{n_N} f$. Das sup in der Definition von p_{mn} geht dann natürlich über $x \in \mathbb{R}^N$.

Im $\mathbb{R}^n$ haben wir noch eine weitere Struktur, die sich leicht verallgemeinern läßt: das Skalarprodukt.

Definition 1.1.7 *X sei linearer Raum über $\mathbb{C}$. Ein Skalarprodukt $(\cdot,\cdot)$ in X ist eine Abbildung $X \times X \longrightarrow \mathbb{C}$ mit*
i) $(x, y) = \overline{(y, x)}, \quad x, y \in X\,,$
ii) $(x, y + z) = (x, y) + (x, z), \quad (x, cy) = c(x, y) \quad \forall\, x, y, z \in X,\ c \in \mathbb{C}\,.$
(x, x) ist wegen i) für alle $x \in X$ reell, wenn es nicht negativ ist, heißt das Skalarprodukt positiv semidefinit, wenn zudem $(x, x) = 0 \Rightarrow x = 0$ gilt, heißt es positiv definit.

Einfache Folgerungen: $(\cdot,\cdot)$ ist antilinear im ersten Faktor, d. h. $(cx + y, z) = \overline{c}\,(x, z) + (y, z)$; es ist eine „Sesquilinearform“. (**Vorsicht:** das ist die Konvention der Physiker, die Mathematiker bevorzugen die Konvention, bei der $(\cdot,\cdot)$ linear im ersten und antilinear im zweiten Faktor ist.)
$(\cdot,\cdot) = 0$ wenn einer der Faktoren 0 ist; das folgt aus ii) mit $c = 0$. Wie aus dem $\mathbb{R}^n$ bekannt, kann $(x, y) = 0$ gelten, auch wenn $x \neq 0$ und $y \neq 0$, wir sagen dann auch im allgemeinen Fall, daß x und y „orthogonal“ sind.

Lemma 1.1.8 *Sei $(\cdot,\cdot)$ ein positiv definites Skalarprodukt auf X. Dann definiert $\|x\| := (x, x)^{1/2}$ eine Norm auf X.*

Beweis: Die beiden ersten Eigenschaften von Definition 1.1.3 sind trivial wegen der Definition des Skalarprodukts. Zum Beweis der Dreiecksungleichung benutzen wir die sehr wichtige **Cauchy-Schwarz-Ungleichung** (auch Schwarzsche Ungleichung genannt):

Lemma 1.1.9 $(\cdot,\cdot)$ *sei ein positiv definites Skalarprodukt. Dann gilt*

$$|(x,y)| \leq (x,x)^{1/2}(y,y)^{1/2}\,.$$

Damit folgt (mit der Notation $\|x\| := (x,x)^{1/2}$)

$$\|x+y\|^2 = (x+y,x+y) = (x+y,x)+(x+y,y) \leq \|x+y\|\;\|x\| + \|x+y\|\;\|y\|\;;$$

Wenn $x+y \neq 0$ ist, folgt die Dreiecksungleichung durch Division, wenn $x+y=0$ ist, gilt sie trivialerweise. ∎

Beweis der Schwarzschen Ungleichung: sie ist trivial, wenn einer der Faktoren $= 0$ ist. Sei $x \neq 0$, sei $z := y - \frac{(x,y)}{\|x\|^2}\,x$, dann ist $(x,z) = (x,y) - (x, \frac{(x,y)}{\|x\|^2}\,x) = 0$. Folglich ist

$$\|y\|^2 = \left(\frac{(x,y)}{\|x\|^2}\,x + z, \frac{(x,y)}{\|x\|^2}\,x + z\right) = \frac{|(x,y)|^2}{\|x\|^4}\,\|x\|^2 + \|z\|^2 \geq \frac{|(x,y)|^2}{\|x\|^2}\,,$$

d. h. die behauptete Gleichung ist richtig. ∎

(Die Größe $\frac{(x,y)}{\|x\|^2}\,x$ ist der Anteil von y in Richtung von x, und z ist der zu x orthogonale Anteil. Das ist wohlbekannt, wenn $X = \mathbb{R}^n$ und $(x,y) = \vec{x}\cdot\vec{y}$ das übliche Skalarprodukt ist; man übernimmt diese Redeweise für beliebige Räume.)

Räume mit positiv definitem Skalarprodukt sind also spezielle normierte Räume; sie heißen auch **unitäre Räume**.

1.1.10 Anmerkung: Der Beweis der Schwarzschen Ungleichung zeigt, daß Gleichheit möglich ist: dann und nur dann, wenn x und y parallel sind, d. h. $x = \mathrm{const}\, y$.

Korollar 1.1.11 *In einem unitären Raum ist die Abbildung $x \mapsto (y,x)$ normstetig.*

Vor dem Beweis zwei Vorbemerkungen: Bei einer linearen Abbildung $L: \;X \longrightarrow Y$ in normierten Räumen muß man nur die Stetigkeit bei $x = 0$ überprüfen[2]; und die Stetigkeit einer linearen Abbildung, d. h. $\|L(x) - L(y)\|_Y < \varepsilon$ falls $\|x-y\|_X < \delta(\varepsilon)$, folgt aus der Beschränktheit von L, d. h. $\|L(x)\|_Y \leq C\;\|x\|_X$ mit einer passenden Konstanten C: wenn L beschränkt ist, gilt $\|L(x)-L(y)\|_Y = \|L(x-y)\|_Y \leq C\;\|x-y\|_X$, man wähle also $\delta = \varepsilon/C$ und erhält die Stetigkeit. (Die Umkehrung gilt ebenfalls, ist aber nicht so einfach zu zeigen, doch brauchen wir sie im allgemeinen nicht.)

Beweis von 1.1.11: folgt sofort aus der Schwarzschen Ungleichung und der zweiten Vorbemerkung. Trivial: das Skalarprodukt ist natürlich auch im ersten Faktor stetig.

Hier folgt noch ein einfaches, aber nützliches Lemma:

[2]Denn $L(x) - L(y) = L(x-y) = L(x-y) - 0$; wenn y nahe bei x ist, dann ist $x-y$ nahe bei 0.

Lemma 1.1.12 *Sei $\{x_n \in X\}_{n\in\mathbb{N}}$ eine Cauchy-Folge, dann ist auch $\{\|x_n\|\}_{n\in\mathbb{N}}$ eine Cauchy-Folge. Wenn $\lim_{n\to\infty} x_n = x \in X$, dann gilt auch $\lim_{n\to\infty} \|x_n\| = \|x\|$.*

Beweis: Nach der Dreiecksungleichung ist

$$\|x_n\| = \|x_n - x_m + x_m\| \leq \|x_n - x_m\| + \|x_m\| \ ,$$

also $\|x_n\| - \|x_m\| \leq \|x_n - x_m\|$. Vertauschung von n und m ergibt

$$\Big| \|x_n\| - \|x_m\| \Big| \leq \|x_n - x_m\| \underset{n,m\to\infty}{\longrightarrow} 0\,. \tag{1.13}$$

Das ist die erste Aussage; die zweite folgt, wenn man in Gleichung (1.13) x_m durch x ersetzt.

Es fehlt uns hier noch ein Begriff: die **Vollständigkeit**. Ein topologischer Raum X heißt vollständig, wenn jede Cauchy-Folge $\{x_n \in X\}$ ein Limeselement $x \in X$ hat, so daß $x_n \to x$ gilt.
Beispiel: Der Raum $\mathcal{S}$ der temperierten Funktionen ist vollständig. (Eine Cauchy-Folge $f_n \in \mathcal{S}$ ist insbesonders eine p_{00}-Cauchy-Folge, konvergiert also punktweise; die Limesfunktion heiße f; sie ist stetig, weil die Konvergenz gleichmäßig ist (Konvergenz in der sup-Norm heißt gleichmäßige Konvergenz, glm. konvergente stetige Funktionen haben einen stetigen Limes); auch $x^k f_n$ konvergiert, und zwar gegen $x^k f$, man überlegt sich unschwer, daß f selbst wieder $\in \mathcal{S}$ ist und daß $f = \lim f_n$.)

Definition 1.1.14 *Ein vollständiger normierter Raum heißt* **Banachraum***; ein vollständiger unitärer Raum heißt* **Hilbertraum***.*

Beispiele:
Jeder Hilbertraum ist auch ein Banachraum. Der Raum $\mathcal{C}_0^0(\mathbb{R}^n)$ mit der sup-Norm ist ein Banachraum; er ist aber kein Hilbertraum, die sup-Norm läßt sich nicht durch ein Skalarprodukt erzeugen.
$\mathbb{C}^n$ mit dem Skalarprodukt $(a,b) = \sum_{i=1}^n \overline{a_i} b_i$ ist ein Hilbertraum (hier ist $a = (a_1, \ldots, a_n)$, $b = (b_1, \ldots, b_n)$). (Ihnen ist wohl der $\mathbb{R}^n$ vertrauter als der $\mathbb{C}^n$; doch ändert sich nicht viel, nur daß die Komponenten der Vektoren komplexe Zahlen sind und daß man beim Skalarprodukt den ersten Faktor komplex konjugiert nimmt, so bleibt das Skalarprodukt positiv definit.)
Hier ist ein weniger triviales Beispiel eines Hilbertraums:

1.1.15 Beispiel: Raum $\ell^2(\mathbb{C})$ der quadratsummablen Folgen

$$\ell^2(\mathbb{C}) := \Big\{(a_1, a_2, \ldots);\ a_k \in \mathbb{C}\,, \sum |a_k|^2 < \infty \Big\}$$

mit dem Skalarprodukt $(a,b) = \sum_{k=1}^\infty \overline{a_k} b_k$. (Man zeigt, daß diese Summe konvergiert, wenn $a, b \in \ell^2$ (Ü), etwas weniger trivial ist die Vollständigkeit.)

Beispiel eines unitären Raumes: Man definiere auf dem Raum $\mathcal{C}^0_{00}(\mathbb{R})$ der stetigen Funktionen mit kompaktem Träger

$$(f, g) := \int_{\mathbb{R}} \overline{f(x)} g(x)\, dx\,; \tag{1.16}$$

das Integral existiert trivialerweise und liefert ein positiv definites Skalarprodukt (Ü), das eine Norm definiert. Der Raum ist aber kein Hilbertraum, denn er ist nicht vollständig: $f = \lim f_n$ heißt ja, daß $\|f - f_n\|^2 = \int |f(x) - f_n(x)|^2\, dx \to 0$ gelten muß; punktweise Konvergenz ist nicht verlangt, nur „Konvergenz im Mittel“; und in diesem Sinne kann auch eine stückweise stetige Funktion als Limes auftreten (vergl. Beispiel 1: Dichte beim Phasenübergang). Aber: die Limesfunktionen, die hier auftreten können, haben selbst eine endliche Norm; man kann den Raum vervollständigen und erhält dann einen Hilbertraum. Zur Beschreibung dieses Raumes brauchen wir den Begriff des Lebesgue-Integrals, zu dem ich im nächsten Abschnitt etwas sagen will.

Nun eine nützliche Formel, mit der man das Skalarprodukt zweier Vektoren durch Normen berechnen kann: **Polarisationsidentität**

$$(x, y) = \frac{1}{4} \sum_{\nu=0}^{3} i^\nu \left\|x + (-i)^\nu y\right\|^2 . \tag{1.17}$$

Beweis einfach: durch direktes Ausrechnen.
Man könnte denken, daß man damit jeden normierten Raum zu einem unitären machen kann; das geht nicht:

Lemma 1.1.18 *In einem unitären Raum gilt die „Parallelogramm-Gleichung“*

$$\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2)\,;$$

sie ist notwendig und hinreichend dafür, daß Gleichung (1.17) ein positiv definites Skalarprodukt definiert.

Beweis: Übungsaufgabe.
Die Supremumsnorm aus Gleichung (1.4) erfüllt die Parallelogramm-Gleichung nicht, ebenso wenig wie die Normen $\|x\|_\infty$ und $\|x\|_p$, $p \neq 2$, aus Anmerkung 1.1.5.

Wir brauchen später noch einen weiteren wichtigen Begriff:

Definition 1.1.19 *X sei topologischer linearer Raum. Eine lineare Abbildung $f : X \longrightarrow \mathbb{C}$ heißt* **lineares Funktional** *über X.*

Die Menge X^* der linearen stetigen *Funktionale über X heißt der zu X (topologisch)* **duale Raum***.*

Beispiele für lineare Funktionale: i) X unitärer Raum: $x \mapsto (y,x)$; ii) $X = M_n(\mathbb{C}) =$ Raum der n-n-Matrizen mit komplexen Elementen: $A \mapsto \operatorname{tr} A =$ Spur von A (ich benutze die englische Abkürzung von „trace", um die Abkürzung „Sp" für „Spektrum" zu reservieren); iii) X = elektromagnetische Wellen, d. h. Lösungen der (linearen) Maxwell-Gleichungen: $(\vec{E}(\vec{x},t), \vec{B}(\vec{x},t)) \mapsto E_k(\vec{x}_0, t_0)$ = eine Komponente der elektrischen Feldstärke am Mittelpunkt Ihrer Radio-Antenne.
Diese Beispiele sind ungenau, denn ich sollte in jedem Fall die Topologie des Raumes charakterisieren, doch das ist hier weitgehend evident. Aber wichtig: der Zusatz „stetig" in der Definition ist nicht überflüssig; lineare Funktionale im $\mathbb{R}^n$ sind immer stetig, das gilt aber nicht in beliebigen unendlich-dimensionalen Räumen.

Lemma 1.1.20 *i) X^* ist ein linearer Raum.*
ii) Wenn X ein normierter Raum ist, definiert

$$\|\Phi\| := \sup_{x \in X, x \neq 0} \frac{|\Phi(x)|}{\|x\|}, \quad \Phi \in X^*$$

auf X^* eine Norm. X^* ist bezüglich dieser Norm vollständig.

Beweis: i): Wenn f und g stetige lineare Funktionale sind, dann offensichtlich auch $f + ag$.
Daß $\|\Phi\|$ eine Norm definiert, ist nicht schwer zu zeigen (Ü); auf den Beweis der Vollständigkeit verzichte ich, obwohl er nicht sehr kompliziert ist. ∎

Einfache Folgerung:

$$|\Phi(x)| \leq \|\Phi\| \; \|x\| \; . \tag{1.21}$$

Das ist eine Verallgemeinerung der Schwarzschen Ungleichung, denn in einem unitären Raum X ist die Abbildung $\Phi_y : x \mapsto (y,x)$ wegen der Stetigkeit des Skalarprodukts (Kor. 1.1.11) ein Element von X^*, und $\|\Phi_y\| = \|y\|$.

1.1.22 Beispiel: Sei $X = \mathcal{C}_0^0(\mathbb{R})$ mit $\|\cdot\|_\infty = \sup_{\mathbb{R}} |\cdot|$; dann sind

$$\Phi_1 : f \mapsto \int f(x) e^{-x^2} dx \,,$$

$$\Phi_2 : f \mapsto f(x_0), \; x_0 \text{ fest,}$$

stetige lineare Funktionale.

1.2 Lebesgue-Maß und Lebesgue-Integral

In diesem Abschnitt: unvollständige und mehr anschauliche Darstellung vorwiegend ohne Beweise. (Beweise finden sich z. B. im Buch von Hewitt und Stromberg.) Ziel: Vermittlung eines Eindrucks der notwendigen Begriffsbildungen; Sie müssen das Folgende nicht alles im Detail wissen. Aber:

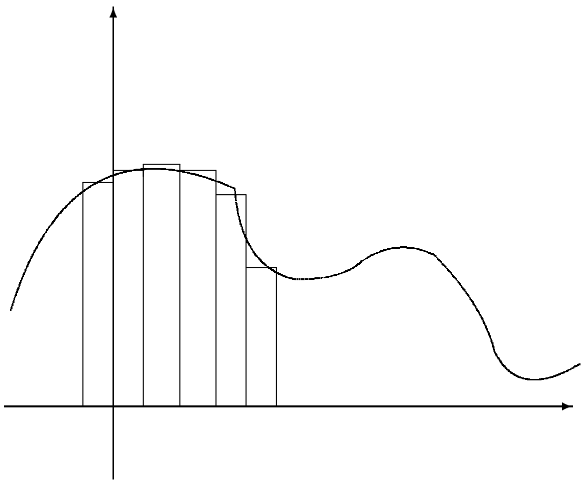
Wichtig zu merken: *Eine Riemann-integrierbare Funktion ist Lebesgue-integrabel, und es gilt*

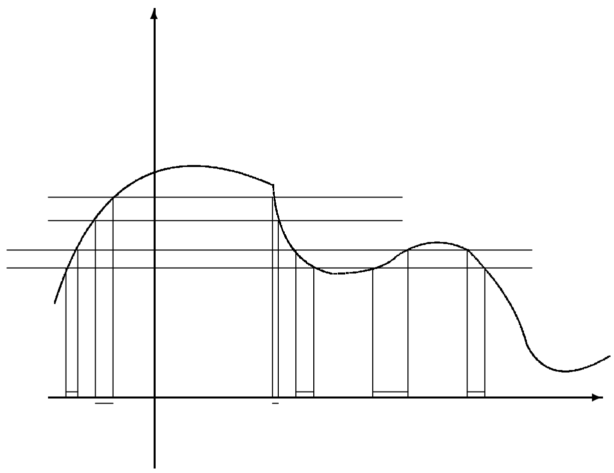
Riemann-Integral = Lebesgue-Integral, wenn die linke Seite existiert.
Nicht jede Lebesgue-integrable Funktion ist Riemann-integrabel.

Zur Definition des Lebesgue-Integrals und überhaupt wichtig: der Begriff des Maßes; er taucht bei der Spektralzerlegung von Operatoren im Hilbertraum in verallgemeinerter Form wieder auf.

Wir betrachten die Integration einer reellwertigen Funktion über $\mathbb{R}$; Verallgemeinerung später ziemlich leicht möglich. Riemann-Integral: Approximation durch Stufenfunktion, gewonnen durch Einteilung der Abszisse; $\int \approx \sum_{\text{Intervalle}}$ Intervall-Länge mal mittlerem Funktionswert.

Man kann genau so gut von einer Einteilung der Ordinate ausgehen und feststellen, in welchen Intervallen die Funktion etwa einen vorgegebenen konstanten Wert hat; Multiplikation dieses Wertes mit der Länge der gefundenen Intervalle und Summation über alle möglichen Werte ergibt ungefähr dasselbe wie die Riemann-Approximation. Das ist die Idee beim Lebesgue-Integral; man muß natürlich auch hier einen Grenzübergang durchführen.





Die Übertragung dieser Idee auf höhere Dimensionen oder auf Integration über andere Räume ist dann nicht schwer: man fragt, auf welcher Menge die Funktion ungfähr einen vorgegebenen Wert hat und mißt die Fläche, das Volumen, ... dieser Menge usw. Voraussetzung ist also, daß man den auftretenden Mengen ein **Maß** zuordnen kann.

Beschreibung des Maßbegriffs am Beispiel des Lebesgue-Maßes λ im $\mathbb{R}$. Für ein Intervall $I_{ab} = (a, b)$ soll λ die Länge angeben: $\lambda(I_{ab}) := b - a$. Das ist unser Ausgangspunkt. Die Längen disjunkter Intervalle werden addiert:

$$\lambda(\overset{\mathrm{d}}{\cup}\ I_n) = \sum \lambda(I_n)\,. \tag{2.1}$$

Diese Gleichung ist auch sinnvoll, wenn man abzählbar unendlich viele Intervalle betrachtet, vorausgesetzt, daß die Summe konvergiert oder daß man für λ auch den Wert ∞ zuläßt, man sagt, die Mengenfunktion λ sei „σ-additiv“. („σ-...“ meint in der Mathematik oft, daß eine Aussage sich auf abzählbar viele Objekte bezieht.)
Auch sinnvoll: Eine einpunktige Menge $A = \{x\}$ hat das Maß 0, denn man kann x in immer kleinere Intervalle einschließen, deren Längen $\to 0$ gehen. Das ist ein Beispiel einer „Nullmenge“.

Im allgemeinen kann man nicht jeder beliebigen Menge ein Maß zuordnen.

Definition 1.2.2 *Ein σ-Feld oder σ-Körper von Mengen ist eine Familie $\mathcal{F}$ von Mengen eines Raumes Ω, die abgeschlossen ist unter Komplementbildung und unter der Bildung endlicher Durchschnitte und abzählbarer Vereinigungen.*

(Der Begriff des Mengenkörpers ist nicht direkt vergleichbar mit dem Begriff des Körpers in der Algebra.)
Ω muß hier kein linearer Raum sein, „Raum“ heißt hier einfach: Menge.
Wenn man von den offenen Intervallen in $\mathbb{R}$ ausgeht, kann man durch Komplement-, Durchschnittsbildung und Vereinigungen einen σ-Körper erzeugen, den „Borelkörper“. Dazu gehören die leere Menge, $\mathbb{R}$ selbst, die abgeschlossenen Intervalle und alle einpunktigen Mengen (Ü). Seine Elemente heißen Borelmengen.

Definition 1.2.3 *Sei $\mathcal{F}$ ein σ-Körper von Teilmengen eines Raumes Ω. Eine Funktion $\mu: \ \mathcal{F} \longrightarrow \overline{\mathbb{R}}_0^+$ mit $\mu(\emptyset) = 0$, die für disjunkte abzählbare Vereinigungen Gleichung (2.1) erfüllt, heißt* **Maß***, das Tripel $(\Omega, \mathcal{F}, \mu)$ heißt* **Maßraum***, die Elemente von $\mathcal{F}$ μ-meßbare Mengen.*

Ausgehend von Intervallen $\subset \mathbb{R}$ bekommt man in der oben angedeuteten Weise das Lebesgue-Maß λ auf dem Borelkörper. Ich habe hier keine strenge Definition gegeben; Sie können das Lebesgue-Maß einer beliebigen Borelmenge berechnen, indem Sie seine folgenden schönen „Stetigkeits“-eigenschaften benutzen: es ist „regulär“, d. h. $\lambda(A) < \infty$ für kompakte Mengen, und $\lambda(A) = \inf\{\lambda(U);\ U \text{ offen},\ A \subset U\}$; $\lambda(U) =$

$\sup\{\lambda(G);\ G$ abgeschlossen, $G \subset U\}$ für alle offenen Mengen U.
(Man kann λ noch allgemeiner definieren, d. h. auf einem σ-Körper, der größer ist als der Borelkörper; davon werden wir hier keinen Gebrauch machen.)

Andere Beispiele für Maße:

1.2.4 Beispiel: Sei $\Omega = \mathbb{R}$ und $g : \mathbb{R} \longrightarrow \mathbb{R}_0^+$ eine nicht-negative stückweise stetige Funktion (die also Riemann-integrierbar ist). Dann definiert

$$\mu(U) := \int_U g(x)\,dx, \quad U \text{ offen}\,,$$

ein reguläres Maß. Wenn g stetig ist, ist $g(x)dx \approx \mu([x, x + dx])$.
Angenommen, g sei integrabel bis $-\infty$, dann ist $h(x) = \int_{-\infty}^{x} g(x)\,dx$ monoton wachsend; man kann das Riemann-Stieltjes-Integral $\int f\,dh$ bilden (gegeben durch Approximationssummen $\sum f(\xi_i)(h(x_{i+1}) - h(x_i));\ x_i \leq \xi_i \leq x_{i+1},\ \int f\,dh = \int f h'\,dx$, falls h differenzierbar). Man kann dann μ durch h definieren; es entspricht einer Bewertung von Intervallen nicht durch ihre Länge $b - a$, sondern durch $h(b) - h(a)$.

1.2.5 Beispiel: Sei $\Omega = \mathbb{R}$, $\mathcal{F} = \mathcal{F}_{\text{Borel}}$ und $\nu(n) = 1\ \forall n \in \mathbb{Z},\ \nu((n, n+1)) = 0$. Dadurch wird ein Maß definiert, das man nicht mehr in der Form $\int g(x)\,dx$ schreiben kann, bei dem gewisse einpunktige Mengen ein endliches Maß haben.

Wie schon erwähnt: es gibt Nullmengen, d. h. Mengen vom Maß 0. Beispiele: $A \subset \mathbb{R}$ mit $\operatorname{card} A = \aleph_0$ hat das Lebesgue-Maß 0; für eine Borelmenge A, auf der im Beispiel 1.2.4 $g = 0$ ist, ist $\mu(A) = 0$.

1.2.6 Notation: Eine Aussage gilt „μ-fast überall" (kurz: f. ü. oder englisch: almost everywhere, kurz: a. e.), wenn sie richtig ist mit Ausnahme auf einer μ-Nullmenge.

Beispiel: Eine stückweise stetige Funktion ist stetig λ-fast überall.

Definition 1.2.7 *Eine Funktion $f : \Omega \longrightarrow \overline{\mathbb{R}}$ heißt $\mathcal{F}$-meßbar, wenn für alle $a \in \mathbb{R}$ die Mengen $F = \{x \in \Omega;\ -\infty < f(x) < a\} \in \mathcal{F}$ sind.*
Wenn $\Omega = \mathbb{R}$ und $\mathcal{F} = \mathcal{F}_{\text{Borel}}$, sagt man , daß f Lebesgue-meßbar ist.
Eine komplexwertige Funktion $f : \Omega \longrightarrow \mathbb{C}$ heißt meßbar, wenn $\Re f$ und $\Im f$ meßbar sind.

(Für die $\mathcal{F}$-Meßbarkeit muß noch kein Maß vorgegeben sein; und „Lebesgue-meßbar" hat, genau genommen, auch noch nichts mit dem Lebesgue-Maß zu tun.)

Jetzt rede ich nur noch vom Lebesgue-Maß. Daß eine Funktion meßbar ist, ist offensichtlich Voraussetzung dafür, daß man das oben skizzierte Vorgehen für das Lebesgue-Integral durchführen kann. Etwa genauer: man beginnt mit reellwertigen positiven Funktionen, die auch den Wert ∞ annehmen können, und approximiert die zu integrierende Funktion durch „einfache" Funktionen, d. h. durch meßbaren Funktionen, die nur endlich viele Funktionswerte annehmen: man betrachtet alle möglichen Zerlegungen (der Menge der Funktionswerte)

$$\mathcal{Z}: \ a_i \in \mathbb{R}, \ 0 = a_0 < a_1 < a_2 < \ldots < a_{n+1} = \infty$$

und die Mengen

$$F_i = \{x \in \mathbb{R}; \ a_i \leq f(x) < a_{i+1}\}$$

und definiert

$$\text{L-}\int f(x)\,dx := \sup_{\mathcal{Z}} \sum_i a_i \lambda(F_i)\,.$$

Die a_i sind hier nicht irgendwelche Zwischenwerte, sondern die Untergrenzen der Werte-Intervalle; deshalb vergrößert sich die Summe, wenn man die Zerlegung $\mathcal{Z}$ verfeinert, d. h. die Definition des Integrals als Supremum ist sinnvoll. (Daß man manche Lebesgue-integrable Funktionen nicht Riemann-integrieren kann, liegt eventuell daran, daß die Riemann-Summen für diese Funktionen empfindlich von den gewählten Zwischenpunkten abhängen.)
Meßbarkeit einer Funktion ist natürlich noch nicht ausreichend dafür, daß sie integrabel ist, sie kann ja im Unendlichen zu langsam oder gar nicht abfallen, sie kann nicht-integrable Singularitäten haben; dann ist das sup $= \infty$. Eine meßbare reelle Funktion zerlegt man in den positiven und negativen Teil, $f = f^+ - f^-$ und setzt $\int f\,dx = \int f^+\,dx - \int f^-\,dx$, wenn nicht beide Summanden dem Betrag nach unendlich sind, entsprechend für komplexwertige Funktionen.
Mit analogem Vorgehen kann man das Integral für andere Räume und für andere Maße definieren.
Merke: eine Riemann-integrable Funktion ist meßbar, mehr noch: sie ist Lebesgue-integrabel, aber die Umkehrung gilt nicht.

Es ist nicht ganz trivial, zu zeigen, daß das Lebesgue-Integral eine lineare Operation ist; aber man kann sich überzeugen, daß es die üblichen Eigenschaften hat:

- $\int (f+g)\,dx = \int f\,dx + \int g\,dx$;
- $\int f\,dx$ existiert genau dann, wenn $\int |f|\,dx$ existiert; und $|\int f\,dx| \leq \int |f|\,dx$;
- Wenn $|f| \leq g$ und $\int g\,dx < \infty$, dann ist f integrabel.

Verschärfung des letzten Punktes: **Theorem der beschränkten Konvergenz**, (engl.: Lebesgue dominated convergence theorem) ein wichtiges und viel benutztes Theorem.

Theorem 1.2.8 *Sei $\{f_n\}_{n=1}^{\infty}$ eine Folge $\overline{\mathbb{R}}$-wertiger meßbarer Funktionen, und für jedes n gelte $|f_n(x)| \leq s(x)$ f. ü., wobei s eine nicht-negative Lebesgue-integrable Funktion ist: $s \geq 0$ und $\int s(x)\,dx < \infty$.*
Wenn f. ü. $\lim_{n\to\infty} f_n(x)$ existiert, dann existiert auch $\lim_{n\to\infty} \int f_n(x)\,dx$ und es gilt

$$\lim_{n\to\infty} \int f_n(x)\,dx = \int \lim_{n\to\infty} f_n(x)\,dx.$$

1.2.9 Anmerkung: Dieses Theorem gilt auch für andere Grundräume und andere Maße.

Man schreibt für das Lebesgue-Integral oft $d\lambda$ statt dx:

$$\int f(x)\,d\lambda \equiv \int f(x)\,d\lambda(x) \equiv \int f(x)\,\lambda(dx) \equiv \int f(x)\,dx\,,$$

ich werde vorwiegend die letzte Schreibweise benutzen.
(Zur Erklärung: oft definiert man für ein Maß μ auf der reellen Achse $\mu(x) := \mu(\{-\infty, x\})$, $x \in \mathbb{R}$, falls diese Größen endlich sind; im Beispiel 1.2.4 $\mu(U) := \int_U g(x)\,dx$ ist $\mu(x) = h(x)$, und für stetiges g ist $d\mu(x) = h'(x)dx = g(x)dx$. $d\lambda(x)$ schreibt man in Analogie zum Riemann-Stieltjes-Integral, $\lambda(dx)$ bezeichnet das (infinitesimal kleine) Maß des infinitesimalen Intervalls $(x, x+dx)$.)

1.2.10 Anmerkung: Der Maßbegriff läßt sich verallgemeinern: wenn μ_k Maße auf $\mathcal{F}$ sind, ist $\mu := \mu_1 - \mu_2$ eine Mengenfunktion mit allen in Def. 1.2.3 genannten Eigenschaften außer daß sie nicht $\mathbb{R}_0^+$-wertig ist: μ heißt reellwertiges Maß (engl.: signed measure); analog definiert man komplexwertige Maße. Wir brauchen später bei der Spektralzerlegung von Operatoren sogar operatorwertige Maße.

Definition 1.2.11

$$\mathcal{L}^2(\mathbb{R}) := \left\{ f : \mathbb{R} \longrightarrow \mathbb{C};\ f \text{ meßbar},\ \int_{\mathbb{R}} |f(x)|^2\,dx < \infty \right\}$$

heißt der Raum der quadratintegrablen Funktionen. Die Integrale sind im Sinne von Lebesgue gemeint; dabei ist also vorausgesetzt, daß die Funktionen f Lebesgue-meßbar sind im Sinne der Definition 1.2.7.

Genaue Bezeichnung: $\mathcal{L}^2(\mathbb{R}, \mathcal{F}_{\text{Borel}}, \lambda)$; wenn klar ist, was gemeint ist, lassen wir einige oder alle Argumente weg.
Auf $\mathcal{L}^2(\mathbb{R})$ kann man die Norm

$$\|f\|_2 := \left(\int_{\mathbb{R}} |f(x)|^2\,dx \right)^{\frac{1}{2}} \tag{2.12}$$

einführen; sie gehört zum Skalarprodukt $(f,g) = \int \overline{f(x)} g(x)\, dx$, Gleichung (1.16). Dieses Skalarprodukt ist für alle Elemente aus $\mathcal{L}^2(\mathbb{R})$ wohldefiniert, wie wir gleich sehen werden.

$\mathcal{L}^2(\mathbb{R})$ ist ein linearer Raum: Wenn $f, g \in \mathcal{L}^2(\mathbb{R})$, also meßbare Funktionen sind, ist auch $\overline{f} g$ meßbar.[3] Weiterhin ist

$$\left|\overline{f(x)} g(x)\right| = \sqrt{|f(x)|^2}\sqrt{|g(x)|^2} \leq \frac{1}{2}\left(|f(x)|^2 + |g(x)|^2\right)$$

(denn für reelle Zahlen ist $ab \leq \frac{1}{2}(a^2+b^2)$); also ist nach Theorem 1.2.8 $\overline{f} g$ integrabel, das gleiche gilt für $\overline{g} f$, folglich existiert das Skalarprodukt (f,g), und

$$\int |f(x)+g(x)|^2\ dx = \|f\|^2 + \int \overline{f(x)} g(x)\, dx + \int f(x)\overline{g(x)}\, dx + \|g\|^2 < \infty\,,$$

also $f+g \in \mathcal{L}^2(\mathbb{R})$. Daß $af \in \mathcal{L}^2(\mathbb{R})$ für $a \in \mathbb{C}$, ist trivial.

Man kann zeigen, daß man jede $\mathcal{L}^2$-Funktion im Sinne der obigen Norm durch Stufenfunktionen beliebig genau approximieren kann, sogar durch einfache Funktionen, d. h. Stufenfunktionen, die nur endlich viele Werte annehmen; oder man kann sie approximieren durch stetige Funktionen, sogar durch Funktionen aus $\mathcal{S}$ (ohne Beweis):

Satz 1.2.13 *Die Menge der λ-meßbaren einfachen Funktionen ist $\mathcal{L}^2$-dicht in $\mathcal{L}^2(\mathbb{R})$. Die Menge $\mathcal{S}(\mathbb{R})$ der temperierten Funktionen ist $\mathcal{L}^2$-dicht in $\mathcal{L}^2(\mathbb{R})$.*

(Ich habe Ihnen $\mathcal{S}(\mathbb{R})$ als Beispiel für einen vollständigen Raum genannt; das ist kein Widerspruch zu diesem Satz, denn die Aussagen beziehen sich auf verschiedene Topologien!) Dieser Satz macht plausibel:

Satz 1.2.14 $\mathcal{L}^2(\mathbb{R})$ *mit dem Skalarprodukt (1.16) ist ein Hilbertraum, d. h. nicht nur ein unitärer Raum, sondern auch vollständig.*

Ich war hier nicht genau: Wenn man eine Funktion $f \in \mathcal{L}^2(\mathbb{R})$ auf einer Nullmenge abändert zu einer Funktion $\tilde{f}$, dann ist offensichtlich $\left\|f - \tilde{f}\right\| = 0$, im $\mathcal{L}^2$-Sinn sind diese Funktionen ununterscheidbar, sollten also identifiziert werden. Genau genommen muß man also immer von Äquivalenzklassen von Funktionen sprechen, doch meinen wir das immer stillschweigend, wenn wir von $\mathcal{L}^2$-Funktionen reden; andernfalls würde das Skalarprodukt (1.16) nicht die Bedingung $\|f\| = 0 \Rightarrow f = 0$ erfüllen.

[3] Das erfordert einige, allerdings nicht sehr schwierige Beweisschritte: man zeigt zunächtst für reellwertige meßbare Funktionen f und g: i) $f + \text{const}$ ist meßbar, ii) die Mengen $\{x \in \mathbb{R};\ f(x) \geq g(x)\}$, $\{x \in \mathbb{R};\ f(x) > g(x)\}$ und $\{x \in \mathbb{R};\ f(x) = g(x)\}$ sind meßbar, iii) $f+g$ ist meßbar, iv) fg ist meßbar. Dann folgt die Meßbarkeit für $\overline{f} g$.

Zum Beweis von Satz 1.2.14: Wenn es eine meßbare Funktion f gibt, für die $\|f - f_m\| \to 0$ für $m \to \infty$, dann ist $f \in \mathcal{L}^2(\mathbb{R})$, das ist einfach zu zeigen. Denn

$$\begin{aligned}\frac{1}{2}\,|f(x)|^2 &\leq |f(x) - f_m(x)|^2 + |f_m(x)|^2 \ , \\ \frac{1}{2}\int |f(x)|^2\ dx &\leq \|f - f_m\|^2 + \|f_m\|^2 \ ;\end{aligned}$$

die rechte Seite geht gegen $\lim \|f_m\|$, vergl. Lemma 1.1.12. Doch es ist nicht-trivial zu zeigen, daß ein solches f für jede Cauchy-Folge existiert. (Es gibt Folgen f_n, die im $\mathcal{L}^2$-Sinn gegen 0 konvergieren, für die aber $f_n(x)$ für kein x konvergent ist! Beispiel in „Hewitt and Stromberg“, p. 157.)

Das Folgende dient zur Illustration, ist nicht wichtig für die Vorlesung.
Es gibt wilde meßbare Mengen. Beispiel: Cantor-Menge C. Sie ist „perfekt“, d. h. gleich der Menge ihrer Häufungspunkte, „nirgends dicht“, d. h. ihr Abschluß ($= C$) hat keine inneren Punkte (sie besteht sozusagen nur aus Randpunkten), sie hat überabzählbar viele Punkte, aber das Maß $\lambda(C) = 0$.
Eine Funktion f, die auf C den Wert 1 hat und sonst $= 0$ ist, ist also meßbar, und L-$\int f\,dx = 0$; mit Riemann hat man keine Chance.
Sei φ eine stetige, f. ü. differenzierbare Funktion. Wenn φ' Riemann-integrabel ist, gilt die wichtige Relation $\varphi(x) - \varphi(a) = \int_a^x \varphi'(t)\,dt$. Für Lebesgue-integrable Funktionen ist diese Aussage im allgemeinen falsch! Mithilfe der Cantor-Menge C kann man eine Funktion φ definieren, die stetig und monoton auf $[0, 1]$ von 0 bis 1 anwächst und f. ü. (nämlich außerhalb C) differenzierbar ist und die Ableitung 0 hat („Teufelstreppe“); für sie ist $\varphi(1) - \varphi(0) = 1 \neq 0 = \int_0^1 \varphi'\,dt$.[4]

Zum Schluß will ich noch einige andere auftretende Funktionenräume definieren:

Definition 1.2.15 *Sei* $p \geq 1$.

$$\begin{aligned}\mathcal{L}^p(\mathbb{R}^n) &:= \left\{ f : \mathbb{R}^n \longrightarrow \mathbb{C};\ f \textit{ Lebesgue-meßbar, } \int_{\mathbb{R}^n} |f(x)|^p\, d^n x < \infty \right\} , \\ \|f\|_p &:= \left(\int_{\mathbb{R}^n} |f(x)|^p\, d^n x \right)^{\frac{1}{p}} ; \\ \mathcal{L}^\infty &:= \{ f : \mathbb{R}^n \longrightarrow \mathbb{C};\ f \textit{ Lebesgue-meßbar, } \operatorname{ess\,sup}_{\mathbb{R}^n} |f(x)| < \infty \} , \\ \|f\|_\infty &:= \operatorname{ess\,sup}_{\mathbb{R}^n} |f(x)| .\end{aligned}$$

Hier ist $\operatorname{ess\,sup}_{\mathbb{R}^n} |f(x)|$ *das Infimum aller* $a \in \overline{\mathbb{R}}_0^+$, *für die* $|f(x)| < a$ *f. ü. gilt, d. h.* $= \inf \left\{ a \in \overline{\mathbb{R}}_0^+;\ \lambda(\{x \in \mathbb{R}^n;\ f(x) \geq a\}) = 0 \right\}$.

[4] Man kann beweisen, daß es Mengen in $\mathbb{R}$ gibt, die nicht λ-meßbar sind. Ihre Konstruktion ist kompliziert, und man kann zeigen, daß man sie nur mithilfe des Zermelo'schen Auswahlaxioms konstruieren kann. (Siehe Hewitt and Stromberg, p. 136.)

Ohne Beweis:

Theorem 1.2.16 *Die Menge $\mathcal{C}^0_{00}(\mathbb{R}^n)$ der stetigen Funktionen mit kompaktem Träger ist dicht in $\mathcal{L}^p(\mathbb{R}^n)$.*

„dicht“ meint hier: „dicht“ in der $\mathcal{L}^p$-Topologie. Wir können sogar noch weiter gehen und jede Funktion $f \in \mathcal{C}^0_{00}$ durch eine beliebig oft differenzierbare („beliebig glatte“) Funktion mit kompaktem Träger approximieren; ich führe es nur aus für Funktionen über $\mathbb{R}$:

Lemma 1.2.17 *Sei*

$$\psi_\delta(x) := \frac{1}{N}\begin{cases} \exp\left(\dfrac{-x^2}{\delta^2 - x^2}\right) & \text{für } |x| \leq \delta\,, \\ 0 & \text{sonst}\,, \end{cases} \tag{2.18}$$

N ist ein Normierungsfaktor, d. h. so zu wählen, daß $\int \psi_\delta \, dx = 1$ ist. f sei stetig mit kompaktem Träger. Dann ist

$$f_\delta(x) := \int f(\xi)\psi_\delta(x-\xi)\, d\xi \in \mathcal{C}^\infty_{00} \quad \forall \delta > 0$$

und $\lim_{\delta\to 0} f_\delta = f$ gleichmäßig in x.

Beweis: Ü.

Für Funktionen über $\mathbb{R}^n$ kann man analog vorgehen. Weil gleichmäßige Konvergenz natürlich auch $\mathcal{L}^p$-Konvergenz impliziert, gilt also auch

Korollar 1.2.19 *Die Menge $\mathcal{C}^\infty_{00}(\mathbb{R}^n)$ der beliebig oft differenzierbaren Funktionen mit kompaktem Träger ist dicht in $\mathcal{L}^p(\mathbb{R}^n)$.*

Notation: Statt $\mathcal{C}^\infty_{00}$ schreibe ich im folgenden $\mathcal{D}$.

Trivial: $\mathcal{D}(\mathbb{R}^n) \subset \mathcal{S}(\mathbb{R}^n)$, mehr noch:

Satz 1.2.20 *$\mathcal{D}(\mathbb{R}^n)$ ist dicht in $\mathcal{S}(\mathbb{R}^n)$, wobei hier die Konvergenz in $\mathcal{S}$ gemeint ist.*

Das überlegt man sich leicht (für $\mathbb{R}$), indem man analog zu ψ_δ Funktionen definiert, die in einem immer größer werdenden Intervall $= 1$ sind und außerhalb dieses Intervalls rasch beliebig glatt auf 0 abfallen:

$$\psi_R(x) := \begin{cases} 1 & \text{für } |x| \leq R\,, \\ \exp\left(-\dfrac{x^2 - R^2}{(R+1)^2 - x^2}\right) & \text{für } R < |x| < R+1\,, \\ 0 & \text{sonst.} \end{cases} \tag{2.21}$$

Wenn nun $f \in \mathcal{S}(\mathbb{R})$, dann sind die Funktionen $f\psi_R \in \mathcal{D}(\mathbb{R})$ und liefern die gewünschte Approximation.(Ü; wesentlich ist, daß sich $\psi_R^{(n)}$ unabhängig von R abschätzen läßt).

Theorem 1.2.16 und Lemma 1.2.17 sind für einen Physiker sehr beruhigend, denn er kann nie bis zu unendlichen Werten messen und glaubt meist ziemlich fest an die Differenzierbarkeit fast aller Funktionen (bis zum Beweis des Gegenteils).

Ich schließe mich dieser Haltung insofern an, als ich z. B. f' schreibe, ohne immer zu sagen oder zu zeigen, daß f eine differenzierbare Funktion ist; oder: $\int f\,dx$ meint oft automatisch, daß f integrierbar ist.

1.3 Hilberträume

Hilberträume werden in der Quantenmechanik benötigt. Quantenmechanische Teilchen haben *auch* Welleneigenschaften, der Zustand eines solchen Teilchens wird nicht mehr, wie in der Mechanik, durch seinen Ort und seinen Impuls bestimmt, sondern durch eine „Wellenfunktion“, die ihm zugeordnet ist. Das ist eine komplexwertige Funktion $\varphi(x,t)$ des Ortes $x \in \mathbb{R}^3$ und der Zeit t, die nicht direkt physikalisch interpretierbar und nicht direkt meßbar ist, doch ihr Betrag bestimmt die Wahrscheinlichkeit ΔW, das Teilchen zur Zeit t im kleinen Volumen d^3x an einem bestimmten Ort $x \in \mathbb{R}^3$ zu finden: $\Delta W = |\varphi(x,t)|^2\ d^3x$. Die Wahrscheinlichkeit, das Teilchen irgendwo zu finden, soll natürlich $= 1$ sein:

$$\int_{\mathbb{R}^3} |\varphi(x,t)|^2\ dx = 1\,.$$

Die Wellenfunktion muß also quadratintegrabel sein, d. h. $\in \mathcal{L}^2(\mathbb{R}^3)$. Wenn das Teilchen auf eine Raumdimension begrenzt ist, genügt der $\mathcal{L}^2(\mathbb{R})$, wenn man mit k Teilchen im $\mathbb{R}^3$ zu tun hat, braucht man den $\mathcal{L}^2(\mathbb{R}^{3k})$. Das sind alles Hilberträume, vergl. Satz 1.2.14, der für höhere Dimensionen entsprechend gilt. Wenn man Teilchen mit Spin betrachtet, braucht man wieder andere Hilberträume. Das ist die Motivation für diesen Abschnitt.

Ich wiederhole noch einmal: ein Hilbertraum $\mathcal{H}$ ist ein linearer Raum über $\mathbb{C}$, auf dem ein positiv definites Skalarprodukt definiert ist und der in der durch dieses Skalarprodukt gegebenen Norm-Topologie vollständig ist. Das heißt (ich benutze von jetzt an meist die Notation $\varphi, \psi, \chi, \ldots$ für die Hilbertraum-Vektoren):

$(\varphi, a\psi + \chi) = a(\varphi,\psi) + (\varphi,\chi)$;

$(\varphi,\psi) = \overline{(\psi,\varphi)}$;

$\|\varphi\|^2 \equiv (\varphi,\varphi) \geq 0,\ = 0$ nur, wenn $\varphi = 0$;

$\|\varphi_n - \varphi_m\| < \varepsilon$ für hinreichend große Indizes n, m impliziert, daß es ein Limeselement $\varphi \in \mathcal{H}$ gibt mit $\|\varphi_n - \varphi\| < \varepsilon$ für hinreichend großes n.

Wir haben ein anderes Beispiel erwähnt: den Raum $\ell^2(\mathbb{C})$ der quadratsummablen Folgen komplexer Zahlen (Beispiel 1.1.15); dieser Raum sieht aus wie $\mathbb{C}^n$ mit $n \to \infty$. Das gilt — in gewisser Weise — für fast jeden Hilbertraum, den man in der Physik betrachtet, auch — obwohl es Ihnen jetzt nicht einleuchtend erscheinen wird — für die Räume $\mathcal{L}^2(\mathbb{R}^n)$ der quadratintegrablen Funktionen!

Jeder endlich-dimensionale Vektorraum mit positiv-definitem Skalarprodukt läß sich durch eine orthogonale Basis normierter Vektoren — kurz „o. n. Basis“ — aufspannen. Das gilt — mit einiger Vorsicht — auch für einen Hilbertraum $\mathcal{H}$. („Normiert“ heißt auch hier: normiert auf 1, d. h. „Länge“ $\|\varphi\| = 1$.) Man hat auch hier den Orthogonalitätsbegriff: $\varphi \perp \psi$ heißt $(\varphi,\psi) = 0$. Damit kann man, wie üblich, den Anteil von ψ in

Richtung φ und den zu φ senkrechten Anteil definieren:

$$\psi_{\|} := \frac{(\varphi,\psi)}{(\varphi,\varphi)}\,\varphi = \text{ Projektion von } \varphi \text{ auf } \psi; \qquad \psi_{\perp} := \psi - \psi_{\|}\,.$$

Offensichtlich ist $(a\varphi)_{\|} = a\varphi$; die Abbildung $P : \psi \mapsto \psi_{\|}$ ist also eine Projektion, d. h. eine weitere Anwendung der Abbildung liefert nichts Neues: $P \circ P = P$. Und offensichtlich ist $(\psi_{\perp}, \varphi) = 0$.

Weiter: die Vektoren der Menge $\mathcal{D} = \{\varphi_\alpha\}$ heißen **linear unabhängig**, wenn für jede *endliche*(!) Auswahl von Vektoren aus $\mathcal{D}$ die Relation $\sum a_i\varphi_i = 0$ nur für $a_i = 0$ möglich ist.
Sei nun $\mathcal{D}$ eine endliche lineare Menge $\subset \mathcal{H}$. Dann kann man mit dem **Schmidtschen Orthogonalisierungsverfahren** ein System orthogonaler normierter Vektoren — kurz „Orthonormalsystem" — erzeugen, die $\mathcal{D}$ aufspannen:
Man wählt einen beliebigen Vektor $\varphi_\alpha \neq 0$ aus $\mathcal{D}$ und nennt ihn ψ_1', normiert: $\psi_1 := \psi_1'/\,\|\psi_1'\|$. Sei φ_β nicht parallel zu ψ_1 (wenn es einen solchen Vektor nicht gibt, bricht das Verfahren ab, alle Vektoren sind parallel zu ψ_1) und definiert $\psi_2' := \varphi_\beta - (\psi_1, \varphi_\beta)\psi_1$, $\psi_2 := \psi_2'/\,\|\psi_2'\|$. Dann sind ψ_1 und ψ_2 orthogonal. Man wähle nun φ_γ, das nicht linear von ψ_1, ψ_2 abhängig ist, und definiere

$$\psi_3' := \varphi_\gamma - (\psi_1, \varphi_\gamma)\psi_1 - (\psi_2, \varphi_\gamma)\psi_2 = \text{zu } \psi_1 \text{ und } \psi_2 \text{ orthogonaler Anteil von } \varphi_\gamma\,,$$

normiere, und gewinne so ψ_3. Und so weiter, bis es keinen Vektor in $\mathcal{D}$ mehr gibt, der nicht von dem bis dahin gewonnenen Orthonormalsystem aufgespannt wird.

Die ψ_n spannen die **lineare Hülle** L.H.$(\mathcal{D})$ von $\mathcal{D}$ auf, sie ist definiert als die Menge aller *endlichen* Linearkombinationen von Vektoren aus $\mathcal{D}$; jeder derartige Teilraum von $\mathcal{H}$ sieht also wie ein endlich-dimensionaler $\mathbb{C}^n$ aus. Wenn $\mathcal{D}$ eine unendliche Menge von Vektoren ist, kann man gegebenenfalls auch ein unendliches System orthonormaler Vektoren finden. Beispiel: man wähle in $\ell^2(\mathbb{C})$ die Menge der Folgen $e^{(i)} = (0, 0, \ldots, 0, 1, 0, \ldots)$, die nur an der i-ten Stelle eine 1 haben.
Doch muß nicht jedes Element von $\mathcal{D}$ eine endliche Linearkombination eines maximalen Orthonormalsystems sein, selbst wenn die Vektoren von $\mathcal{D}$ linear unabhängig sind!
Beispiel: Sei $\{\psi_n\}_{n=1}^{\infty}$ ein Orthonormalsystem, $a_k \in \mathbb{C}$ mit $\sum_1^\infty |a_k|^2 < \infty$ und

$$\varphi_k := \sum_{i=1}^{k} a_i\psi_i\,; \quad \varphi := \lim_{k\to\infty} \varphi_k \equiv \sum_1^\infty a_i\psi_i\,.$$

φ ist wohldefiniert, denn der Limes existiert: sei $k < l$,

$$\|\varphi_k - \varphi_l\|^2 = \left\|\sum_{k+1}^{l} a_i\psi_i\right\|^2 = \sum_{k+1}^{l} |a_i|^2$$

wird für sehr großes k (und l) beliebig klein; und die unendliche Summe ist normkonvergent: $\|\varphi\|^2 = \sum_1^\infty |a_i|^2 < \infty$. Aber φ ist *nicht* linear abhängig von ψ_n, $\varphi \notin$

$\overline{\text{L.H.}(\{\psi_n\})}$, denn man braucht *unendlich viele* ψ's zu seiner Darstellung („L.H." meint immer nur endliche Linearkombinationen!). Wohl aber gilt $\varphi_k \in \text{L.H.}(\{\psi_n\})$ und daher $\varphi \in \overline{\text{L.H.}(\{\psi_n\})}$. Dieses φ hat keinen zu L.H.$(\{\psi_n\})$ orthogonalen Anteil mehr; wenn man $\mathcal{D} = \text{L.H.}(\{\psi_n\}, \varphi)$ betrachtet, kann man das Orthogonalisierungsverfahren nicht über $\{\psi_n\}$ hinaus fortsetzen.

Das wird klar mit der **Besselschen Ungleichung** und ihrem Beweis:

$$\sum_{\psi\in\mathcal{E}} |(\psi,\varphi)|^2 \leq \|\varphi\|^2 \ , \quad \mathcal{E} \text{ Orthonormalsystem} \tag{3.1}$$

Zum Beweis betrachten wir beliebige *endliche* Summen über Vektoren aus $\mathcal{E}$:

$$\begin{aligned}
0 &\leq \left\|\varphi - \textstyle\sum_\psi (\psi,\varphi)\,\psi\right\|^2 \\
&= \left(\varphi - \textstyle\sum_\psi (\psi,\varphi)\,\psi, \varphi - \textstyle\sum_\psi (\psi,\varphi)\,\psi\right) \\
&= (\varphi,\varphi) - \textstyle\sum_\psi \overline{(\psi,\varphi)}(\psi,\varphi) - \textstyle\sum_\psi (\psi,\varphi)(\varphi,\psi) + \textstyle\sum_\psi \textstyle\sum_{\psi'} \overline{(\psi',\varphi)}(\varphi,\psi)(\psi',\psi) \\
&= (\varphi,\varphi) - \textstyle\sum_\psi |(\psi,\varphi)|^2 \,.
\end{aligned} \tag{3.2}$$

Die letzte Gleichheit gilt, weil $(\psi',\psi) = 0$, wenn $\psi' \neq \psi$. Weil das für alle endlichen Summationen gilt, deshalb auch für Summation über alle $\psi \in \mathcal{E}$. ■
Der Vektor $\varphi_\perp \equiv \varphi - \sum_{\psi\in\mathcal{E}} (\psi,\varphi)\,\psi$ ist der zu $\mathcal{E}$ senkrechte Anteil von φ.
Wir definieren also

Definition 1.3.3 *Eine Menge $\mathcal{B}$ von Vektoren in $\mathcal{H}$ heißt Basis von $\mathcal{H}$, wenn $\mathcal{H} = \overline{\text{L.H.}(\mathcal{B})}$, d. h. jedes $\varphi \in \mathcal{H}$ ist Norm-Limes einer Folge endlicher Linearkombinationen von Vektoren aus $\mathcal{B}$.*

„Basis" heißt nach dieser Definition nicht, daß die Basisvektoren linear unabhängig sein müssen; das ist natürlich der Fall, wenn $\mathcal{B}$ ein Orthonormalsystem ist, d. h. eine o. n. Basis
Im Fall einer o. n. Basis $\mathcal{B}$ (die man ja mit dem Schmidtschen Orthogonalisierungsverfahren aus einer beliebigen Basis konstruieren kann) kann man jeden Vektor als unendliche Summe schreiben:

$$\varphi = \sum_{\psi\in\mathcal{B}} (\psi,\varphi)\psi \,. \tag{3.4}$$

Die Konvergenz der Summe ist natürich im Sinne der Norm gemeint: zu jedem ε gibt es ein endliches $\mathcal{B}_1 \subset \mathcal{B}$, so daß

$$\left\|\varphi - \sum_{\psi\in\mathcal{B}_1} (\psi,\varphi)\psi\right\| < \varepsilon \,.$$

Orthogonale normierte Basis anders definiert: „vollständiges“ Orthonormalsystem

Definition 1.3.5 *Eine Menge $\mathcal{D} \subset \mathcal{H}$ heißt vollständig, wenn es keinen zu* L.H.$(\mathcal{D})$ *orthogonalen Vektor gibt.*

Dieser Begriff ist natürlich nicht identisch mit der (topologischen) Vollständigkeit von $\mathcal{H}$ selbst, er bezieht sich auf Teilmengen von $\mathcal{H}$ und hat nichts mit Konvergenz zu tun.

Satz 1.3.6 *Sei $\mathcal{E}$ ein Orthonormalsystem Dann sind folgende Aussagen äquivalent:*
(1) $\mathcal{E}$ ist vollständig.
(2) $\varphi = \sum_{\psi\in\mathcal{E}} (\psi, \varphi)\, \psi \quad \forall\, \varphi \in \mathcal{H}$.
(3) $\|\varphi\|^2 = \sum_{\psi\in\mathcal{E}} |(\psi, \varphi)|^2 \quad \forall\, \varphi \in \mathcal{H}$ (**Parsevalsche Gleichung**).
(4) $(\varphi, \chi) = \sum_{\psi\in\mathcal{E}} (\varphi, \psi)(\psi, \chi) \quad \forall\, \varphi, \chi \in \mathcal{H}$.
(5) $\overline{\text{L.H.}(\mathcal{E})} = \mathcal{H}$.

Beweis: (1) $\Leftrightarrow$ (2): (2) heißt kurz: $\varphi = \varphi_{||}$; wenn das für alle φ gilt, ist $\mathcal{E}$ nach Definition vollständig.
(2) $\Leftrightarrow$ (3): folgt aus Gleichung (3.2), wenn man beachtet, daß $\|\chi\| = 0 \Leftrightarrow \chi = 0$.
(2) $\Rightarrow$ (4), weil man wegen der Stetigkeit von $(\cdot,\cdot)$ Summationsbildung und Skalarproduktbildung vertauschen kann.
(4) $\Rightarrow$ (3) mit $\chi = \varphi$.
(2) $\Rightarrow$ (5) $\Rightarrow$ (1) gilt, weil $\varphi \perp \text{L.H.}(\mathcal{E}) \Leftrightarrow \varphi \perp \overline{\text{L.H.}(\mathcal{E})}$: $(\varphi, \chi_n) = 0$ impliziert $(\varphi, \lim \chi_n) = 0$. ■

Wegen der r.S. von (4) kann man suggestiv schreiben

$$\mathcal{E} \text{ ist vollständig} \quad \Longleftrightarrow \quad \sum_{\mathcal{E}} \psi)(\psi = \mathbb{1},$$

hier meint $\mathbb{1}$ den Einheitsoperator, der jeden Vektor auf sich selbst abbildet. Dabei habe ich $\psi)$ statt ψ geschrieben in Anlehnung an die Dirac-Schreibweise, die ich Ihnen später kurz vorstellen werde (dort schreibt man $\sum |\,\psi\rangle\langle\psi| = \mathbb{1}$).

Es ist klar: wenn es eine o. n. Basis mit endlich vielen Vektoren gibt, dann ist der Raum endlich dimensional,

$$\dim \mathcal{H} = \max\{\text{card}\, \mathcal{E};\ \mathcal{E}\ \textit{Orthonormalsystem}\}\,. \tag{3.7}$$

Dann gibt es natürlich keine o. n. Basis mit unendlich vielen Vektoren; und umgekehrt: ein Raum mit unendlicher o. n. Basis kann nicht endlich dimensional sein. Allgemein

gilt für zwei o. n. Basis $\mathcal{B}$ und $\mathcal{B}'$ in $\mathcal{H}$ immer: $\operatorname{card}\mathcal{B} = \operatorname{card}\mathcal{B}'$; und diese Gleichung ist so gemeint, daß man auch unendliche Kardinalzahlen unterscheidet (Beweis z. B. im Buch von Großmann). Dann ist es sinnvoll, Gleichung (3.7) auch für beliebige Hilberträume als Definition zu nehmen. In der Physik treten vorwiegend Räume abzählbarer Dimension auf; der Hilbertraum $\mathcal{H}$ heißt **separabel**, wenn $\dim\mathcal{H} = \aleph_0$ ist. (In der Mathematik gibt es noch andere Begriffe von „separabel“, Vorsicht!)
(Auch wenn $\dim\mathcal{H}$ überabzählbar ist, gilt Gleichung (3.4); n durchläuft zwar eine überabzählbare Menge, aber in der Summe können natürlich nur abzählbar viele Summanden $\neq 0$ sein.)

Die Aussage von Gleichung (3.7) zusammen mit der Existenz vollständiger Orthonormalsysteme zeigt: **Hilberträume gleicher Dimension sind isomorph**; es gibt eine isometrische, d. h. norm-erhaltende 1-1-Abbildung der Räume aufeinander, welche Skalarprodukte in Skalarprodukte überführt.[5] (Das sollte man sich einmal sauber überlegen: Ü!)

Wichtig zu merken: In unendlich-dimensionalen Räumen muß man zwischen $\text{L.H.}(\mathcal{D})$ und $\overline{\text{L.H.}(\mathcal{D})}$ unterscheiden.
Konvention: Ein **Unterraum** $\mathcal{H}_1$ eines Hilbertraumes $\mathcal{H}$ ist eine **abgeschlossene** lineare Untermannigfaltigkeit.
Wegen $\infty + \infty = \infty$ kann $\mathcal{H}$ zwei zueinander orthogonale Unterräume mit $\dim = \infty$ enthalten. Wenn $\mathcal{H}$ separabel ist, ist es dann sogar isomorph zu jedem dieser Räume!
Wenn $\mathcal{H}_1$ Unterraum von $\mathcal{H}$ ist, aber $\neq \mathcal{H}$, kann man eine o. n. Basis $\mathcal{B}_1$ von $\mathcal{H}_1$ finden, die aber natürlich keine Basis von $\mathcal{H}$ ist. $\varphi \neq \mathcal{H}_1$ läßt sich nicht durch Vektoren aus $\mathcal{H}_1$ approximieren; es hat einen zu $\mathcal{H}_1$ orthogonalen Anteil. Mit dem Schmidtschen Orthogonalisierungsverfahren kann man $\mathcal{B}_1$ zu einer o. n. Basis von $\mathcal{H}$ ergänzen. Damit kann man dann einen zu $\mathcal{H}_1$ orthogonalen Unterraum $\mathcal{H}_1^\perp = \mathcal{H}_2$ finden, so daß $\mathcal{H}$ von diesen beiden Räumen aufgespannt wird: $\mathcal{H} = \overline{\text{L.H.}(\mathcal{H}_1, \mathcal{H}_2)}$.

In der Quantenmechanik braucht man im allgemeinen nur separable Räume, es ist dann egal, welchen man wählt, denn alle sind isomorph. Ich habe am Anfang erwähnt, daß man den Raum $\mathcal{H} = \mathcal{L}^2(\mathbb{R}^3)$ benutzt, wenn man Quantenmechanik eines Teilchens treibt. Nicht ganz trivial: $\mathcal{L}^2(\mathbb{R}^3)$ ist ein separabler Raum. Man kann es sich überlegen, wenn man beachtet, daß jede $\mathcal{L}^2$-Funktion durch Stufenfunktionen approximierbar ist (Satz 1.2.13), auch durch solche, die nur rationale Werte annehmen über Quadern mit rationalen Eckpunktkoordinaten, und daß $(\aleph_0)^k = \aleph_0$ ist. (Vollständige Funktionensysteme: mehr in der QM-Vorlesung.)

Weiterer Punkt, an dem sich endlich-dimensionale von unendlich-dimensionalen Räum-

[5] Wegen der Polarisationsidentität (I.1.17) ist eine isometrische, d. h. normerhaltende Abbildung auch Skalarprodukt-erhaltend.

en unterscheiden: wenn $\varphi_n \to \varphi$, d. h. $\|\varphi_n - \varphi\| \to 0$, dann gilt auch

$$(\psi, \varphi_n) \longrightarrow (\psi, \varphi) \quad \forall\, \psi \in \mathcal{H}\,. \tag{3.8}$$

In endlich-dimensionalen Räumen ist auch das Umgekehrte richtig, nicht aber, wenn $\dim \mathcal{H} = \infty$:
Gegenbeispiel: sei $\{\varphi_n\}_{n=1}^{\infty}$ eine o. n. Basis Dann gilt für jedes $\psi \in \mathcal{H}$: $\psi = \sum(\varphi_n, \psi)\, \varphi_n$; die Summe ist konvergent, und $\|\psi\|^2 = \sum |(\varphi_n, \psi)|^2$; damit diese Summe konvergiert, ist notwendig: $|(\varphi_n, \psi)| \to 0$. Also: für die φ_n gilt Gleichung (3.8) mit $\varphi = 0$, aber natürlich ist $\|\varphi_n\| = 1 \neq 0$. Nach Lemma 1.1.12 zieht die Normkonvergenz die Konvergenz der Normen nach sich; $\lim \varphi_n = 0$ gilt also nicht.
Gleichung (3.8) definiert einen anderen, schwächeren Konvergenzbegriff:

Definition 1.3.9 *Eine Folge φ_n von Vektoren, welche die Gleichung (3.8) erfüllt, heißt* **schwach konvergent** *gegen φ.*

Schreibweise: $\varphi_n \xrightarrow{w} \varphi$ bzw. $\varphi = \text{w-lim}\, \varphi_n$, „w“ steht für „weak“.[6]

Wie das obige Beispiel zeigt, müssen dann die Normen von φ_n nicht gegen $\|\varphi\|$ konvergieren. Mit der Schwarzschen Ungleichung zeigt man, daß aus der Normkonvergenz die schwache Konvergenz folgt:

$$|(\varphi_\alpha - \varphi, \psi)| \leq \|\varphi_\alpha - \varphi\|\; \|\psi\| \longrightarrow 0\,.$$

Das obige Beispiel ist typisch für die schwache Konvergenz; die Folge muß gewissermaßen in immer höhere Dimensionen laufen können, sonst wäre sie bereits normkonvergent. Eine schwach konvergente Folge, die in einem endlich-dimensionalen Unterraum liegt, ist auch norm-konvergent. Beweis: In diesem Unterraum lassen sich alle Vektoren nach *endlich vielen* Basisvektoren e_k entwickeln, und Limiten lassen sich mit endlichen Summen vertauschen; $\lim_n(e_k, \varphi_n)$ existiert: $a_{nk} \equiv (e_k, \varphi_n) \longrightarrow a_k$; mit $\varphi = \sum a_k\, e_k$ folgt $(\varphi_n, \psi) \longrightarrow (\varphi, \psi)\ \forall\, \psi$ und $\|\varphi_n\|^2 = \sum |a_{nk}|^2 \longrightarrow \sum |a_k|^2 = \|\varphi\|^2$; Konsequenz:

$$\|\varphi_n - \varphi\|^2 = (\varphi_n, \varphi_n) - (\varphi_n, \varphi) - (\varphi, \varphi_n) + (\varphi, \varphi) \longrightarrow 0\,.$$

Diese Gleichung zeigt allgemein, daß eine schwach konvergente Folge $\{\varphi_n\}$ mit $\text{w-lim}\, \varphi_n = \varphi$, für die auch $\lim \|\varphi_n\| = \|\varphi\|$ gilt, norm-konvergent ist. Das ist der zweite Teil des folgenden Lemmas; der erste sagt, daß zwar die Normen, wie gesagt, nicht generell konvergieren, aber beschränkt sind:

[6] $\varphi_n \longrightarrow \varphi$ heißt immer: Konvergenz in der Norm; wenn es die Deutlichkeit erfordert, schreibe ich $\varphi_n \xrightarrow{n} \varphi$. Gelegentlich spricht man auch von starker Konvergenz und schreibt $\varphi_n \xrightarrow{s} \varphi$, s für strong; deshalb benutze ich die Abkürzung des englischen Wortes weak, um Mißverständnissen vorzubeugen.

Lemma 1.3.10 *Für eine schwach konvergente Folge $\{\varphi_n\}$ ist die Menge der Normen beschränkt: $\|\varphi_n\| \le M$ für eine passende Konstante M.*
Wenn zudem $\lim \|\varphi_n\| = \|\varphi\|$ gilt, folgt $\varphi = \lim \varphi_n$ im Sinne der Normkonvergenz.

Beweis: mit dem „Prinzip der gleichmäßigen Beschränktheit" (auch Banach-Steinhaus-Theorem genannt):[7] $\mathcal{T}$ sei eine Menge beschränkter linearer Transformationen eines Banachraums $\mathcal{X}$ in einen Banachraum $\mathcal{Y}$ (d. h. $\|T(x)\| \le C_T \|x\| \;\forall T \in \mathcal{T}, x \in \mathcal{X}$), und es sei darüber hinaus für jedes feste $x \in \mathcal{X}$ die Menge $\{T(x);\; T \in \mathcal{T}\}$ beschränkt (d. h. $\|T(x)\| \le C_x \;\forall T$); dann gibt es eine Konstante C, so daß

$$\|T(x)\| \le C\,\|x\| \quad \forall\, T \in \mathcal{T},\ \forall\, x \in \mathcal{X}\,.$$

(Beweis z. B. in Reed, Simon: Methods of modern mathematical physics, Vol. I.) Dieses Theorem wende man an auf die Abbildungen

$$T_n : \psi \in \mathcal{H} \mapsto T_n(\psi) = (\varphi_n, \psi)\,.$$

Das sind lineare beschränkte Abbildungen des Banachraums $\mathcal{H}$ nach $\mathbb{C}$; und weil (für jedes feste ψ) $\lim(\varphi_n, \psi) = \lim T_n(\psi)$ existiert, ist die Menge $\{T_n(\psi);\; n \in \mathbb{N}\}$ beschränkt, also gilt $|(\varphi_n, \psi)| \le C\,\|\psi\|$ mit einem von n unabhängigen C; mit $\psi = \varphi_n$ folgt die Behauptung. ■

Schließlich noch der wichtige **Satz von Riesz** über den zu einem Hilbertraum dualen Raum:

Theorem 1.3.11 *Jeder Hilbertraum ist zu sich selbst dual; d. h. jedes stetige lineare Funktional $\Phi : \mathcal{H} \longrightarrow \mathbb{C}$ ist gegeben durch einen Vektor $\varphi_\Phi \in \mathcal{H}$:*

$$\Phi(\psi) = (\varphi_\Phi, \psi) \quad \forall\, \psi \in \mathcal{H}\,. \tag{$*$}$$

φ_Φ ist durch Φ eindeutig festgelegt, die Abbildung $\Phi \mapsto \varphi_\Phi$ ist antilinear, und es gilt $\|\Phi\| = \|\varphi_\Phi\|$.

Beweis: Wir betrachten alle Vektoren, auf denen Φ verschwindet, den Kern von Φ: $\ker \Phi := \{\psi \in \mathcal{H};\; \Phi(\psi) = 0\}$. Man sieht leicht, daß $\ker \Phi$ ein Unterraum ist: linear und abgeschlossen, weil Φ stetig ist. Wenn $\ker \Phi = \mathcal{H}$ ist, setzen wir $\varphi_\Phi = 0$, andernfalls gibt es einen Vektor χ, der zu $\ker \Phi$ orthogonal ist, und wir definieren

$$\varphi_\Phi := \frac{\overline{\Phi(\chi)}}{\|\chi\|^2}\,\chi\,.$$

[7] Alternativer Beweis in Abschnitt 4.4.

Im ersten Fall ist $(*)$ trivial erfüllt, im zweiten Fall aber auch:
Für beliebiges ψ ist $\omega \equiv \Phi(\chi)\psi - \Phi(\psi)\chi \in \ker\Phi$, also $\perp \chi$; daraus folgt

$$(\chi, \omega) = 0 = \Phi(\chi)(\chi, \psi) - \Phi(\psi)(\chi, \chi)\,.$$

Division durch (χ, χ) ergibt $\Phi(\psi) = (\varphi_\Phi, \psi)$.
Wegen der Implikation $(\varphi_1, \psi) = (\varphi_2, \psi) \ \ \forall\,\psi \Rightarrow (\varphi_1 - \varphi_2, \psi) = 0 \ \forall\,\psi \Rightarrow \varphi_1 - \varphi_2 = 0$, folgt die Eindeutigkeit.
Mit der Schwarzschen Ungleichung erhalten wir $|\Phi(\psi)| = |(\varphi_\Phi, \psi)| \leq \|\varphi_\Phi\| \, \|\psi\|$, für $\psi = \varphi_\Phi$ gilt Gleichheit, also $\|\Phi\| = \|\varphi_\Phi\|$.
Die Antilinearität ist dann trivial. ■
Anmerkung: Der Beweis zeigt, daß $\ker\Phi$ sehr groß ist: weil φ_Φ eindeutig ist, kann es nur eine zu $\ker\Phi$ senkrechte Richtung geben; gebildet ausgedrückt: $\operatorname{codim} \ker\Phi = 1$ ($\operatorname{codim}\mathcal{H}_1$ ist die Kodimension von $\mathcal{H}_1$, d. h. die Dimension des zu $\mathcal{H}_1$ orthogonalen Unterraums). Das ist Ihnen im endlich-dimensionalen Raum vertraut: der Kern eines linearen Funktionals $\Phi : \vec{x} \mapsto \vec{a} \cdot \vec{x}$ definiert eine Hyperebene $\{\vec{x};\ \vec{a} \cdot \vec{x} = 0\}$ durch den Nullpunkt, d. h. eine lineare Mannigfaltigkeit mit Kodimension 1.

Mit dem Satz von Riesz und dem Lemma 1.3.10 kann man leicht zeigen, daß jede schwach konvergente Folge einen Limes in $\mathcal{H}$ hat (Übungsaufgabe).

Anhang: Die Dirac-Schreibweise

Man schreibt das Skalarprodukt mit eckigen Klammern anstatt mit runden Klammern: $(\cdot,\cdot) \equiv \langle \cdot | \cdot \rangle$. „Klammer" heißt auf Englisch „bracket"; daher bezeichnet man die Größen $| \cdot \rangle$ als „ket-Vektoren" und die Größen $\langle \cdot |$ als „bra-Vektoren"; die ket-Vektoren sind die Elemente unseres Hilbertraums, die bra-Vektoren sind Elemente des dualen Raums: $\langle \varphi |$ angewandt auf $| \psi \rangle$ liefert $\langle \varphi | \psi \rangle \in \mathbb{C}$; und nach dem Satz von Riesz sind das alle möglichen Elemente des Dualraums, man kann bra- und ket-Vektoren eineindeutig aufeinander abbilden.

Vorteile dieser Schreibweise: Wenn man einen Vektor eindeutig durch gewisse Eigenschaften $E_1,\ E_2,\ E_3 \ldots$ charakterisieren kann, bezeichnet man den Vektor einfach mit $| E_1,\ E_2, \ldots \rangle$. Die Entwicklung eines Vektors φ nach einer o. n. Basis $\{\psi_k\}$ schreibt man in der Form

$$| \varphi \rangle = \sum_k | \psi_k \rangle \langle \psi_k | \varphi \rangle \,.$$

(Ob man die komplexe Zahl $\langle \psi_k | \varphi \rangle$ vor oder hinter den Vektor schreibt, ist egal.) Und diese Gleichung kann man anders lesen:

$$| \varphi \rangle = \left(\sum_k | \psi_k \rangle \langle \psi_k | \right) | \varphi \rangle \,;$$

d. h. der Operator $(\sum_k | \psi_k \rangle \langle \psi_k |)$ wirkt auf den Vektor $| \varphi \rangle$ und reproduziert ihn, dieser Operator ist also der Eins-Operator. Die Vollständigkeit des Orthonormalsystem $\{\psi_k\}$ läßt sich also in der Form

$$\sum_k | \psi_k \rangle \langle \psi_k | = \mathbb{1}$$

ausdrücken.

Die Dirac-Schreibweise ist bei Physikern sehr beliebt; sie hat Nachteile bei einer rein mathematischen Diskussion von Hilbertraumoperatoren, siehe Kapitel 3. Ich benutze sie gelegentlich, werde sie aber in dieser Vorlesung nicht verwenden.

Kapitel 2

Fourier- und Laplace-Transformation

2.1 Fourier-Reihen

Ich beginnne mit einer kurzen Wiederholung von Dingen, die Sie hoffentlich wenigstens zum Teil schon wissen. Der Einfachheit halber rede ich über Funktionen über $\mathbb{R}$; die Verallgemeinerung auf $\mathbb{R}^n$ ist nicht schwer.

Zuerst: stetige periodische Funktionen mit Periode 2π. Sie lassen sich bekanntlich in eine Fourier-Reihe entwickeln:

$$f(x) = \sum_{n=0}^{\infty} (a_n \cos nx + b_n \sin nx)\,.$$

Äquivalent, aber oft nützlicher: sin und cos durch e-Funktionen ausgedrückt:

$$f(x) = \sum_{k=-\infty}^{\infty} c_k e^{ikx}\,. \tag{1.1}$$

Diese Gleichung gilt auch allgemeiner, nämlich für Funktionen aus $\mathcal{L}^2([0, 2\pi])$. (Hier sind auch wieder die *Lebesgue*-quadratintegrablen Funktionen gemeint.) Dieser Raum ist ein Hilbertraum mit dem Skalarprodukt

$$(f, g) = \int_0^{2\pi} \overline{f(x)} g(x)\, dx\,; \tag{1.2}$$

und es gilt

Lemma 2.1.3 *Die Menge* $\left\{e_k(x) := (2\pi)^{-1/2} e^{ikx};\ k \in \mathbb{Z}\right\}$ *ist eine o. n. Basis von* $\mathcal{L}^2([0, 2\pi])$.

Beweis: Weil $\int_0^{2\pi} e^{i(k-k')x}dx = 2\pi\delta_{0,k-k'}$ ist für $k, k' \in \mathbb{Z}$ (δ_{kl} ist das Kronecker-Symbol), sieht man sofort, daß die Funktionen $e_k(x)$ ein Orthonormalsystem bilden. Es bleibt die Vollständigkeit zu zeigen: $(e_k, f) = 0\ \forall k \in \mathbb{Z}$ muß $f = 0$ zur Folge haben. Ich erinnere dazu an den Weierstraßschen Approximationssatz:
Jede stetige Funktion läßt sich auf einem abgeschlossenen Intervall *gleichmäßig* durch Polynome approximieren.
Der Satz gilt auch für Funktionen mehrerer Variabler; und wir brauchen ihn für unseren Beweis in abgewandelter Form:

Satz 2.1.4 *Eine stetige* 2π-periodische *Funktion läßt sich gleichmäßig durch trigonometrische Polynome approximieren.*

Beweis: Statt der periodischen Funktion $h(\xi)$ betrachte ich $H(x, y) := rh(\xi)$ mit $x = r\cos\xi,\ y = r\sin\xi$; d. h. ich betrachte r und ξ als Polarkoordinaten in der Ebene. Weil h periodisch ist, ist $H(x, y)$ überall wohldefiniert; es ist stetig, also auf dem abgeschlossenen Einheitskreis durch Polynome gleichmäßig approximierbar. Daher ist auch $h(\xi) = H(\cos\xi, \sin\xi)$ durch Polynome in $\cos\xi$ und $\sin\xi$ gleichmäßig approximierbar.∎

Fortsetzung des Beweises von Lemma 2.1.3: Wäre also unser f mit $(e_k, f) = 0$ stetig und periodisch, wären wir fertig; denn gleichmäßige Approximierbarkeit impliziert natürlich $\mathcal{L}^2$-Approximierbarkeit, und kein Limeselement von L.H. $\{e_k;\ k \in \mathbb{Z}\}$ außer dem Nullelement kann senkrecht zu allen e_k sein. $f \in \mathcal{L}^2[0, 2\pi]$ beliebig: sei $F(x) := \int_0^x f(y)\,dy$, dann ist F stetig. (F existiert, denn eine lokal quadratintegrable Funktion ist auch absolut integrabel; anders: die charakteristische Funktion $\chi_{[0,x]}$ ist sicher $\in \mathcal{L}^2[0, 2\pi]$, und $F(x) = (\chi_{[0,x]}, f)$ ist wohldefiniert.) Weil nach Voraussetzung $F(2\pi) = \sqrt{2\pi}(e_0, f) = 0$ ist, gilt

$$F(0) = F(2\pi) = 0\,, \tag{1.5}$$

F ist also insbesondere periodisch. Für $k \neq 0$: mit partieller Integration und Gleichung (1.5) folgt für jede Konstante c (man beachte $e_0 \perp e_k$ für $k \neq 0$):

$$\sqrt{2\pi}(e_k, F - ce_0) = \int_0^{2\pi} e^{-ikx}F(x)\,dx = \frac{\sqrt{2\pi}}{ik}(e_k, f) = 0\,;$$

Mit $c = (e_0, F)$ ist auch $(e_0, F - ce_0) = 0$. $F - ce_0$ ist also stetig, periodisch und orthogonal zu allen e_k, wir können also Weierstraßschen Satz 2.1.4 anwenden und schließen, daß $F - ce_0 = F - \sqrt{2\pi}^{-1}c = 0$ ist, folglich $F' = f = 0$. (Anmerkung für Liebhaber mathematischer Genauigkeit: $F(x) := \int_0^x f(y)\,dy$ ist f. ü. differenzierbar, und $F' = f$ gilt f. ü., d. h. $f = F' = 0$ im $\mathcal{L}^2$-Sinn.) ∎

Die Berechnung der Fourier-Koeffizienten c_k für eine Funktion f ist klar: $f = \sum(e_k, f)e_k$, also

$$f(x) = \sum_{k=-\infty}^{\infty} c_k e^{ikx} \qquad \text{mit } c_k = \frac{1}{\sqrt{2\pi}}(e_k, f) = \frac{1}{2\pi}\int_0^{2\pi} e^{-ikx} f(x)\, dx\,. \tag{1.6}$$

Die Darstellung durch Fourier-Reihen gilt zunächst nur im $\mathcal{L}^2$-Sinn, d. h. die Reihe stellt die Funktion fast überall dar. Wenn f stetig differenzierbar ist — und das nimmt ein Physiker ja fast immer an — gilt die Darstellung in jedem Punkt (mit Ausnahme der Endpunkte, wenn f nicht periodisch ist): $f' \in \mathcal{L}^2$, also $= \sum b_k e_k$, also $f = \sum a_k e_k$ mit $a_k = b_k/(ik)$. Nun ist $\{b_k\} \in \ell^2$ und auch $c := \{1/(-ik)\} \in \ell^2$, also $(c, b) = \sum a_k$ absolut summierbar, also $\sum a_k e_k$ gleichmäßig konvergent, folglich stetig; und dann muß es punktweise mit f übereinstimmen.

Wenn f nicht in $[0, 2\pi]$ definiert ist, sondern $\in \mathcal{L}^2([0, a])$, kann man alles schnell übertragen, indem man zu $y = \frac{2\pi}{a}x$ übergeht: $g(y) := f(ay/2\pi)$, dann ist $g \in \mathcal{L}^2([0, 2\pi])$. Also:

$$f(x) = \sum_{-\infty}^{\infty} c_k e^{i(2k\pi/a)x} \qquad \text{mit } c_k = \frac{1}{a}\int_0^{a} e^{-i(2k\pi/a)x} f(x)\, dx\,. \tag{1.7}$$

Es sei noch die Parsevalsche Gleichung (siehe Satz 1.3.6) für diesen Fall notiert:

$$\frac{1}{a}\int_0^a |f(x)|^2\, dx = \sum_{k=-\infty}^{\infty} |c_k|^2 \tag{1.8}$$

Wenn $f \in \mathcal{L}^1([0, a])$, dann konvergieren die Integrale $\frac{1}{a}\int_0^a e^{-i(2k\pi/a)x} f(x)\, dx$ absolut; die Fourier-Koeffizienten c_k sind wohldefiniert; aber man kann in diesem Fall nicht garantieren, daß die damit gebildete Fourier-Reihe konvergiert, wenn $f \notin \mathcal{L}^2([0, a])$ (man nehme z. B. $f = x^{-\alpha}$, $x \in [0, a]$, mit $\frac{1}{2} < \alpha < 1$). Es gibt kompliziertere Methoden, mit denen man f auch in diesem Fall aus den c_k zurückgewinnen kann.

Noch eine triviale Anmerkung: statt $\left\{e^{ikx}\right\}$ kann man natürlich auch $\{\cos kx, \sin kx\}$ als orthogonale Basis nehmen.

Zum Schluß noch das sehr brauchbare **Riemann-Lebesgue-Lemma**:

Satz 2.1.9 *Sei* $f \in \mathcal{L}^1([0, 2\pi])$, *dann gilt* $\lim_{|k|\to\infty} c_k = 0$.

Beweis: Nach Theorem 1.2.16 läßt sich f durch eine stetige Funktion $\hat{f}$ approximieren, deren Fourier-Koeffizienten $\hat{c}_k$ heißen mögen. Dann ist auch

$$2\pi|c_k - \hat{c}_k| = \left|\int_0^{2\pi} (f(x) - \hat{f}(x))e^{-ikx} dx\right| \le \int_0^{2\pi} |f(x) - \hat{f}(x)|\, dx$$

klein (unabhängig von k!). $\hat{f}$ ist natürlich auch $\in \mathcal{L}^2$, nach der Parsevalschen Gleichung konvergiert also die Summe $\sum |\hat{c}_k|^2$, die $\hat{c}_k$ müssen also gegen 0 gehen, ebenso die c_k. ■

Es gibt eine naheliegende, sehr nützliche Verallgemeinerung dieses Satzes, die ich ohne Beweis anführe:

Satz 2.1.10 *Sei $f \in \mathcal{L}^1(\mathbb{R})$, dann gilt*

$$\lim_{|k|\to\infty} \int_{\mathbb{R}} f(x) e^{-ikx} dx = 0 \,.$$

Anwendungsbeispiel: schwingende Saite

Die Saite habe die Länge L, sie schwinge in einer Eben, ihre Auslenkung zur Zeit t am Ort x, $0 \leq x \leq L$, sei $\xi(t,x)$. Die Erhöhung der Spannung bei Auslenkung

bewirkt eine rücktreibende Kraft, die (für kleine Auslenkungen) proportional zur Krümmung der Saite angesetzt werden kann: $|K| \propto \left|\frac{\partial^2\xi}{\partial x^2}\right|$. Bei Auslenkung nach oben ist $\frac{\partial^2\xi}{\partial x^2} < 0$, und die Kraft weist nach unten, ist also negativ. Wir setzen also an: $K = k\frac{\partial^2\xi}{\partial x^2,\, k>0}$. Hierbei ist K die Kraft pro Länge; mit der Massendichte ρ (bezogen auf die Länge) definieren wir $D := k/\rho$ und erhalten als Bewegungsgleichung eine eindimensionale Wellengleichung:

$$\left(\frac{\partial^2}{\partial t^2} - D\frac{\partial^2}{\partial x^2}\right)\xi(t,x) = 0 \,.$$

Da die Saite an beiden Enden eingespannt ist, haben wir die Randbedingungen $\xi(t,0) = \xi(t,L) = 0$. Für alle t ist ξ also eine in x periodische Funktion, sie ist stetig (solange die Saite nicht reißt); sie läßt sich daher in eine Fourier-Reihe entwickeln:

$$\xi(t,x) = \sum_{n=1}^{\infty} c_n(t) \sin\frac{n\pi}{L}x \,. \tag{1.11}$$

Damit die Funktion $\xi(t,x)$ in allen Punkten, auch den Endpunkten $x = 0$ und $x = L$, durch die Fourier-Reihe dargestellt wird, muß sie auch in den Randpunkten stetig differenzierbar sein. Das läßt sich erreichen, indem sie in das Intervall $[-L,0]$ antisymmetrisch fortgesetzt wird: $\xi(t,x) := -\xi(t,-x)$ für $x \in [-L,0]$, und dann als $2L$-periodische Funktion betrachtet wird (daher der Faktor $2\pi/(2L) = \pi/L$ im Argument des Sinus). In der Fourier-Reihe einer antisymmetrischen Funktion treten keine Cosinus-Terme auf. Einsetzen in die Differentialgleichung liefert[1]

$$\sum \left(\ddot{c}_n(t) + D\left(\frac{\pi}{L}\right)^2 n^2 c_n\right) \sin\frac{n\pi}{L}x = 0 \,;$$

[1] Ich benutze die Physiker-Notation: $\dot{} = \frac{d}{dt}$.

skalare Multiplikation mit $\sin \frac{l\pi}{L} x$ liefert

$$\ddot{c}_l(t) + (\kappa l)^2 c_l(t) = 0\,,$$

wobei ich $\kappa^2 = D(\pi/L)^2$ gesetzt habe. Diese Gleichung ist leicht lösbar:

$$c_l(t) = c_{1l} e^{i\kappa l t} + c_{2l} e^{-i\kappa l t}\,.$$

Mit $c_{1l} = c_{2l} = 0$ für $l \neq k$ und $c_{1k} = -c_{2k} = a/2i$ bekommen wir die speziellen Lösungen

$$\xi(t,x) = a \sin \kappa k t \, \sin \frac{k\pi}{L} x\,,$$

das sind stehende Wellen, die „Eigenschwingungen" der Saite.
Allgemeine Lösungen: Die Anfangsbedingungen $\xi(0,x)$ und $\frac{\partial}{\partial t}\xi(0,x)$ seien vorgegeben (z. B. durch Streichen oder Zupfen der Saite); dann werden die Koeffizienten c_{ik} eindeutig bestimmt durch Fourier-Reihen dieser Funktionen:

$$\begin{aligned} \xi(0,x) &= \sum (c_{1k} + c_{2k}) \sin \frac{k\pi}{L} x\,, \\ \frac{\partial}{\partial t}\xi(0,x) &= \kappa \sum ik(c_{1k} - c_{2k}) \sin \frac{k\pi}{L} x\,. \end{aligned}$$

Zu jeder Eigenschwingung gehört eine Wellenlänge $2L/k$ und eine Eigenfrequenz $\nu_k = \kappa k/(\pi)$; die Darstellung von ξ durch eine Reihe, Gleichung (1.11), zeigt, welche Eigenschwingungen mit welchen Wellenlängen und Frequenzen in der Lösung enthalten sind, wir haben damit eine Frequenzanalyse gemacht.

2.2 Die Fouriertransformation

Was macht ein Physiker, wenn er eine nicht-periodische Funktion f auf der ganzen reellen Achse vor sich hat, für die er eine Frequenzanalyse sucht? Antwort: er betrachtet f im Intervall $[-\frac{a}{2}, \frac{a}{2}]$, schreibt Gleichung (1.7) für das verschobene Intervall auf:

$$
\begin{aligned}
f(x) &= \sum_{-\infty}^{\infty} c_k e^{i(2k\pi/a)x}\,, &\qquad (2.1)\\
c_k &= \frac{1}{a}\int_{-a/2}^{a/2} e^{-i(2k\pi/a)x} f(x)\,dx\,, &\qquad (2.2)
\end{aligned}
$$

und macht dann den Limes $a \to \infty$. Die Größe $y := 2k\pi/a$ durchläuft alle Punkte vom $\frac{2\pi}{a}\mathbb{Z}$, die für wachsendes a immer dichter liegen, im Limes hat man eine kontinuierliche Variable, die ganz $\mathbb{R}$ durchläuft. Die Fourier-Reihe für f kann man als

$$
f(x) = \sum_{k\in\mathbb{Z}} c_k e^{i(2k\pi/a)x} = \sum_{y\in\frac{2\pi}{a}\mathbb{Z}} \left(\frac{a}{2\pi}c_k\right) e^{iyx}\Delta y, \quad \Delta y = \frac{2\pi}{a}
$$

schreiben (Δy ist die Differenz zwischen aufeinander folgenden y-Werten) und als eine Riemannsche Approximationssumme an ein Integral interpretieren:

$$
f(x) = \sqrt{2\pi}^{-1}\int_{-\infty}^{\infty} c(y)e^{iyx}\,dy
$$

mit

$$
c(y) = \sqrt{2\pi}\,\text{„}\lim\left(\frac{a}{2\pi}c_k\right)\text{“} = \sqrt{2\pi}^{-1}\int_{-\infty}^{\infty} e^{-iyx} f(x)\,dx\,.
$$

Ich habe hier den Faktor $\sqrt{2\pi}$ eingeführt, um das Ganze symmetrisch zu schreiben. Übliche Notation: kontinuierliche Variable $k \in \mathbb{R}$ statt y, und $c(y) =: \tilde{f}(k)$,

$$
\begin{aligned}
f(x) &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \tilde{f}(k)e^{ikx}\,dk \equiv (\mathcal{F}_+\tilde{f})(x)\,, &\qquad (2.3)\\
\tilde{f}(k) &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-ikx} f(x)\,dx \equiv (\mathcal{F}_-f)(k)\,. &\qquad (2.4)
\end{aligned}
$$

Aufpassen: in der Literatur findet man oft die symmetrische, manchmal aber auch eine unsymmetrische Schreibweise.
$\mathcal{F}_-f$ heißt die Fourier-Transformierte von f; $\mathcal{F}_+\tilde{f}$ die inverse Fourier-Transformierte. Diese Bezeichnung ist natürlich erst gerechtfertigt, wenn wir die Richtigkeit dieser Formeln bewiesen haben, d. h. gezeigt, daß $\mathcal{F}_+ = \mathcal{F}_-^{-1}$ gilt. Dann benutzen wir die andere Notation:

$$
\tilde{f} = \mathcal{F}f, \qquad f = \mathcal{F}^{-1}\tilde{f}\,.
$$

Wenn f eine Funktion über dem $\mathbb{R}^n$ ist, muß auch $k \in \mathbb{R}^n$ sein und kx in den Gleichungen (2.3) und (2.4) durch das Skalarprodukt $\vec{k}\cdot\vec{x}$ ersetzt werden, die Integration läuft über $\mathbb{R}^n$ und der Vorfaktor lautet $(2\pi)^{-n/2}$. Das entspricht einer n-fachen Fourier-Transformation, eine für jede Komponente.

Zur Bequemlichkeit für das Folgende:
Notation: $\partial_x^m \equiv \frac{d^m}{dx^m}$, oder wenn $x \in \mathbb{R}^n$ und $m = (m_1, \ldots, m_n)$:

$$\partial_x^m \equiv \frac{\partial^{m_1}}{\partial x_1^{m_1}} \cdots \frac{\partial^{m_n}}{\partial x_n^{m_n}} .$$

Die besten Chancen, die obigen heuristischen Überlegungen streng durchzuführen, haben wir mit braven Funktionen, z. B. solchen aus $\mathcal{S}$ (siehe Definition 1.1.6). Wenn $f \in \mathcal{S}$, dann auch $x^k f \in \mathcal{S}$ und $\partial_x^k f \in \mathcal{S}$. Mehr noch:

Lemma 2.2.5 *Wenn $f \in \mathcal{S}(\mathbb{R}^n)$, dann folgt $\mathcal{F}f \in \mathcal{S}(\mathbb{R}^n)$.*
Die Abbildung $f \mapsto \mathcal{F}f$ ist stetig.

(Wie ich die Variable in $\mathbb{R}^n$ benenne, ist natürlich egal! Ich verwende, um im Einklang mit den obigen Formeln zu bleiben, im ersten Fall x und im zweiten k.)

Beweis: Um Schreibarbeit zu sparen, führe ich alles für $\mathbb{R}$ durch, mit der obigen Notation kann man fast alle Formeln auch als Formeln für den $\mathbb{R}^n$ lesen. Wir müssen das Supremum von $k^m \partial_k^n \tilde{f}$ abschätzen:

$$\begin{aligned} k^m \partial_k^n \tilde{f} &= \frac{1}{\sqrt{2\pi}} \int k^m (-ix)^n e^{-ikx} f(x)\, dx \\ &= \frac{1}{\sqrt{2\pi}} \int (-ix)^n \left(i^m \partial_x^m e^{-ikx} \right) f(x)\, dx \\ &= \frac{(-i)^{m+n}}{\sqrt{2\pi}} \int e^{-ikx} \partial_x^m (x^n f(x))\, dx\,. \end{aligned}$$

Die Vertauschung von ∂_k und Integration in der ersten Gleichheit ist erlaubt, weil $f \in \mathcal{S}$ und alle Integrale hervorragend konvergieren. Die dritte Gleichheit folgt mit m-facher partieller Integration, die Randterme verschwinden alle, weil sowohl f wie auch alle seine Ableitungen im Unendlichen verschwinden. Das Supremum über k ist jetzt leicht abzuschätzen, weil $\partial_x^m (x^n f(x))$ eine stark abfallende Funktion und daher integrabel ist: $p_{mn}(\tilde{f}) \leq (2\pi)^{-1/2} \int |\partial_x^m (x^n f(x))|\, dx < \infty$. Das gilt für alle m, n, also ist $\mathcal{F}f \in \mathcal{S}$.
Um die Stetigkeit zu zeigen:

$$p_{mn}(\tilde{f}) \leq (2\pi)^{-1/2} \int (1+x^2)^{-l} dx \sup |(1+x^2)^l \partial_x^m (x^n f)| \leq \sum c_{rs} p_{rs}(f)$$

mit endlich vielen passenden Paaren r, s und Koeffizienten c_{rs}. Die Halbnormen von $\mathcal{F}f$ sind also beschränkt durch Halbnormen von f, und daraus folgt die Stetigkeit (Ü). $\blacksquare$

(Anmerkung: Die Topologie von $\mathcal{S}$ wird manchmal auch mit den Halbnormen $\hat{p}_{lm} = \sup|(1+x^2)^l\partial_x^m(x^n f)|$ definiert, man überlegt sich leicht, daß beide *Mengen* von Halbnormen denselben Konvergenzbegriff liefern.)

Als Abfallprodukt der obigen Rechnung bekommen wir

$$\widetilde{x^n f} = i^n \partial_k^n \tilde{f}\,, \quad \widetilde{\partial_x^m f} = (ik)^m \tilde{f}\,. \tag{2.6}$$

Diese Relationen gelten nicht nur für Funktionen in $\mathcal{S}$, sondern immer dann, wenn die entsprechenden Fourier-Transformierten existieren.

Nun wollen wir zeigen, daß $\mathcal{S}$ durch die Fourier-Transformation *auf* sich abgebildet wird, und $\mathcal{F}^{-1}$ berechnen. Zur Vorbereitung und zur späteren Verwendung:

Lemma 2.2.7 *Sei $f(x) \geq 0$ eine integrable Funktion und $\int f(x)\,dx = 1$. Man definiere für beliebiges $a > 0$:*

$$f^{(a)}(x) := \frac{1}{a} f\left(\frac{x}{a}\right)\,.$$

Dann ist für jede stetige beschränkte Funktion g

$$\lim_{a\to 0}\int f^{(a)}(x)g(x)\,dx = g(0)\,.$$

Beweis:

$$\int f^{(a)}(x)g(x)\,dx = \int f(x/a)g(ax/a)\,d(x/a) = \int f(y)g(ay)\,dy\,,$$

und dieser Ausdruck konvergiert offensichtlich gegen $g(0)\int f(x)\,dx = g(0)$. ∎

(Die Funktionen $f^{(a)}$ selbst konvergieren nicht, jedenfalls gibt es keine gewöhnliche Limesfunktion, sondern nur eine „verallgemeinerte“: die Diracsche δ-Funktion, siehe Kap. 3.)

Sei nun $f \in \mathcal{S}(\mathbb{R})$ und $\tilde{f} = \mathcal{F}f = \mathcal{F}_- f$; ich möchte $\mathcal{F}_+\tilde{f}$ berechnen und führe dazu einen „konvergenzerzeugenden Faktor“ ein: sei wieder $a > 0$,

$$\begin{aligned} I_a(x) &:= \frac{1}{\sqrt{2\pi}}\int e^{-|k|a}\tilde{f}(k)e^{ikx}dk \\ &= \frac{1}{2\pi}\int dk\int dx' e^{-|k|a+ik(x-x')}f(x')\,. \end{aligned}$$

Uns interessiert letztlich $\lim_{a\to 0} I_a$. Doch solange $a > 0$ ist, können wir dank des konvergenzerzeugenden Faktors $e^{-|k|a}$ den Satz von Fubini anwenden und die Integrationsreihenfolge vertauschen. Die k-Integration läßt sich dann leicht ausführen (man teile das Integral auf in $\int_{-\infty}^{0} + \int_{0}^{\infty}$); Ergebnis:

$$I_a(x) = \frac{1}{\pi}\int dx' f(x') \frac{a}{a^2 + (x - x')^2} = \int dy f(x-y) \frac{1}{\pi a} \cdot \frac{1}{1 + (y/a)^2} \,.$$

Jetzt können wir Lemma 2.2.7 anwenden: hier ist $g(y) := f(x-y)$ offensichtlich stetig und beschränkt; und $\frac{1}{\pi}\frac{1}{1+y^2}$ hat die verlangten Eigenschaften: es ist ≥ 0 und $\int \frac{1}{1+y^2}\, dy = \arctan x\Big|_{-\infty}^{\infty} = \pi$. Also:

$$(\mathcal{F}_+ \tilde{f})(x) = \lim_{a\to 0} I_a(x) = f(x)\,;$$

die Vertauschung des Limes $a \to 0$ und der Integration ist für unsere braven Funktionen natürlich erlaubt. Damit haben wir Gleichung (2.3) für Funktionen aus $\mathcal{S}$ gezeigt, d. h.

$$\mathcal{F}_+\mathcal{F}_- = id \ \text{ auf } \mathcal{S}(\mathbb{R})\,.$$

2.2.8 Anmerkung: Für den Beweis, daß $\lim I_a(x) = f(x)$ ist, haben wir gebraucht: i) $\tilde{f}$ existiert, ist beschränkt und $e^{-|k|a}\tilde{f}(k)$ ist integrabel; ii) f ist integrabel, so daß wir Fubini anwenden können; iii) f ist stetig, so daß wir Lemma 2.2.7 anwenden können. Diese Bedingungen sind erfüllt, wenn $f \in \mathcal{L}^1 \cap \mathcal{C}^0$ ist. Auch dann können wir aus $\mathcal{F}f$ die Funktion f zurückgewinnen; allerdings dürfen wir nicht unbedingt den Limes $a \to 0$ unter das Integral I_a ziehen, $\mathcal{F}_+\tilde{f}$ muß nicht existieren.

Genauso sieht man auch $\mathcal{F}_-\mathcal{F}_+ = id$; also ist $\mathcal{F}_- \equiv \mathcal{F}$ eine Bijektion von $\mathcal{S}$ auf $\mathcal{S}$. Nach Lemma 2.2.5 ist diese Abbildung stetig; mit gleichen Argumenten folgt die Stetigkeit der Umkehrung. Damit haben wir das folgende Theorem bewiesen:

Theorem 2.2.9 *Die Abbildung $\mathcal{F} : \mathcal{S}(\mathbb{R}) \longrightarrow \mathcal{S}(\mathbb{R})$ ist eine bi-stetige Bijektion. Die Umkehrung $\mathcal{F}^{-1}$ ist durch Gleichung (2.3) gegeben.*

Wie für die Fourier-Reihen gilt auch im kontinuierlichen Fall die Parseval-Gleichung (vergl.Gleichung (1.8)), jedenfalls für $f \in \mathcal{S}$:

$$\frac{1}{\sqrt{2\pi}}\int |f(x)^2|\, dx = \frac{1}{\sqrt{2\pi}}\int |\tilde{f}(k)^2|\, dk\,. \tag{2.10}$$

Zum Beweis führen wir wieder einen konvergenzerzeugenden Faktor ein:

$$\begin{aligned}\int dk |\tilde{f}(k)|^2 &= \frac{1}{2\pi}\int dk \int dx \int dx' e^{ik(x-x')}\overline{f(x)} f(x') \\ &= \frac{1}{2\pi}\lim_{a\to 0}\int dx \int dx' \int dk e^{-a|x-x'|+ik(x-x')}\overline{f(x)} f(x')\,.\end{aligned}$$

Hier haben wir wieder Fubini benutzt. Statt über x und x' können wir über x und $y = x - x'$ integrieren:

$$\ldots = \frac{1}{2\pi}\lim_{a\to 0}\int dx \overline{f(x)} \int dy \int dk e^{-a|y|+iky} f(x-y)\,.$$

Wie oben können wir die k-Integration durchführen und mit Lemma 2.2.7 den Limes ausführen: die y-Integration verschwindet, und wir erhalten

$$\ldots = \int dx \overline{f(x)} f(x)\,.$$

■

Was ist mit beliebigen f? Wenn $f \in \mathcal{L}^1$, dann existiert $\tilde{f}$; wenn f zudem stetig ist, können wir es aus $\tilde{f}$ zurückgewinnen, siehe Anmerkung 2.2.8. Das geht sogar dann, wenn wir auf die Stetigkeit verzichten und nur verlangen, daß wir $f(x)$ wiederfinden für Punkte, an denen f nicht zu wild unstetig ist („Lebesgue-Punkte“, Näheres siehe Literatur, z. B. Hewitt, Stromberg).
Für $f \in \mathcal{L}^1$ gilt das Riemann-Lebesgue-Lemma 2.1.10; genauere Formulierung (ohne Beweis):

Satz 2.2.11 *Sei $f \in \mathcal{L}^1$, dann ist $\tilde{f} \in \mathcal{C}_0^0$ und $\left\|\tilde{f}\right\|_\infty \leq \frac{1}{\sqrt{2\pi}} \|f\|_1$.*

(Die Norm-Ungleichung ist leicht zu beweisen: Ü.)

Jetzt können wir sehr schnell den Fall $f \in \mathcal{L}^2$ erledigen: nach Korollar 1.2.19 ist $\mathcal{D}$ und damit erst recht $\mathcal{S}$ dicht in $\mathcal{L}^2$ (bezüglich der $\|\cdot\|_2$-Konvergenz). Zu $f \in \mathcal{L}^2$ gibt es also eine Folge $f_\alpha \in \mathcal{S}$ mit $\|f - f_\alpha\|_2 \to 0$. Zu f_α existiert die Fourier-Transformierte $\tilde{f}_\alpha$; und nach der Parsevalschen Gleichung(2.10) ist $\|f_\alpha - f_{\alpha'}\|_2 = \left\|\tilde{f}_\alpha - \tilde{f}_{\alpha'}\right\|_2$, das ist also eine Cauchy-Folge mit einem Limeselement $\in \mathcal{L}^2$ und $\|f\|_2 = \left\|\lim \tilde{f}_\alpha\right\|_2$. (Hier brauchen wir die Vollständigkeit von $\mathcal{L}^2$.) Wir *definieren* jetzt einfach $\mathcal{F}f := \lim \tilde{f}_\alpha$. Wegen der Stetigkeit von $\mathcal{F}$ auf $\mathcal{S}$ ist das eine konsistente Definition. Für die inverse Transformation können wir ebenso argumentieren; es gilt das **Theorem von Plancherel**:

Theorem 2.2.12 *Es gibt eine eindeutige isometrische lineare Abbildung $\mathcal{T}$ von $\mathcal{L}^2(\mathbb{R})$ auf $\mathcal{L}^2(\mathbb{R})$, die die Abbildung $\mathcal{F} : \mathcal{S}(\mathbb{R}) \longrightarrow \mathcal{S}(\mathbb{R}) \longrightarrow$ fortsetzt: für $f \in \mathcal{S}(\mathbb{R})$ ist $\mathcal{T}f = \mathcal{F}f = \tilde{f}$.*
Weiterhin gilt

$$\int \overline{(\mathcal{T}f)(k)} (\mathcal{T}g)(k)\, dk = \int \overline{f(x)} g(x)\, dx\,. \tag{2.13}$$

Beweis: Die Eindeutigkeit und die Bijektivität folgen aus den obigen Überlegungen; die Isometrie aus der obigen Gleichung $\|f\|_2 = \left\|\lim \tilde{f}_\alpha\right\|_2$. Gleichung (2.13) anders geschrieben: $(\mathcal{T}f, \mathcal{T}g) = (f, g)$, wobei $(\cdot,\cdot)$ das Skalarprodukt in $\mathcal{L}^2$ ist; diese Gleichung folgt aus $\|\mathcal{T}f\|_2 = \|f\|_2$ mit der Polarisationsidentität (I.1.17). ■

Ich werde im Folgenden immer $\mathcal{F}f$ schreiben, auch wenn $f \in \mathcal{L}^2$ ist.
Man kann für $\mathcal{L}^2$-Funktionen auch explizite Formeln angeben, ohne einen Limes durchzuführen. Man benutzt die triviale Relation $e^{ikx} = \frac{d}{dx}(e^{ikx}-1)/(ik)$ und die Tatsache, daß die Funktion $(e^{ikx}-1)/k \in \mathcal{L}^2(\mathbb{R}, dk)$, dann ist für $\tilde{f} \in \mathcal{L}^2(\mathbb{R}, dk)$ das Produkt $\tilde{f}(e^{ikx}-1)/k$ integrabel, weil es als Integrand bei der Berechnung des Skalarprodukts zweier $\mathcal{L}^2$-Funktionen auftritt. (Entsprechendes gilt bei Vertauschung von k und x):

$$\tilde{f}(k) = \frac{1}{\sqrt{2\pi}} \frac{d}{dk} \int \frac{e^{-ikx}-1}{-ix} f(x)\, dx\,, \tag{2.14}$$

$$f(x) = \frac{1}{\sqrt{2\pi}} \frac{d}{dx} \int \frac{e^{ikx}-1}{ik} \tilde{f}(k)\, dk\,. \tag{2.15}$$

Das Theorem 2.2.12 ist für die Quantenmechanik von Bedeutung. Wellenfunktionen $\psi(\vec{x}, t)$ sind $\mathcal{L}^2(\mathbb{R}^3, d^3x)$-Funktionen; das Theorem erlaubt den Übergang zu den Fourier-Transformierten, die ebenfalls wieder in einem Hilbertraum, dem $\mathcal{L}^2(\mathbb{R}^3, d^3k)$, liegen. In diesem Raum ist der allgemeine Formalismus der Quantenmechanik ebenfalls anwendbar, man hat jetzt nur eine andere Darstellung: da die neue unabhängige Variable $\vec{k}$ mit dem Impuls des Teilchens zu tun hat, spricht man von der „Ortsraum-" bzw. der „Impulsraum-Darstellung".

Bevor ich Beispiele bringe, will ich noch eine weitere nützliche Operation definieren: die **Faltung** zweier Funktionen. Für geeignete f und g definiert man

$$(f * g)(x) := \int_{\mathbb{R}} f(y) g(x-y)\, dy\,. \tag{2.16}$$

Man überprüft sofort durch Übergang zu einer anderen Integrationsvariablen, daß $f * g = g * f$ ist. „Geeignet" heißt hier, daß das Integral existieren muß; hinreichend ist, wie man zeigen kann, daß $f, g \in \mathcal{L}^1$, und dann ist auch $f * g \in \mathcal{L}^1$. Worauf es mir hier ankommt: es gilt

$$\frac{1}{\sqrt{2\pi}} \widetilde{f * g} = \tilde{f}\tilde{g}\,. \tag{2.17}$$

Denn

$$\begin{aligned}
\frac{1}{\sqrt{2\pi}} \widetilde{f * g} &= \frac{1}{2\pi} \int dx \int dy\, e^{-ikx} f(y) g(x-y) \\
&= \frac{1}{2\pi} \int dy\, f(y) \int dx\, g(x-y) e^{-ikx} \\
&= \frac{1}{2\pi} \int dy\, f(y) e^{-iky} \int dx\, g(x-y) e^{-ik(x-y)} \\
&= \frac{1}{2\pi} \int dy\, f(y) e^{-iky} \int du\, g(u) e^{-iku} \\
&= \tilde{f}\tilde{g}\,.
\end{aligned}$$

(Die Faltung haben wir im Lemma 1.2.17 verwendet, um eine Funktion zu glätten.)

Ich schreibe einige bisher bewiesene Eigenschaften (Gleichungen (2.6) und (2.17)) und eine weitere, betreffend $f_a(x) := f(x+a)$, noch einmal zusammen:

Lemma 2.2.18

$$\begin{aligned} \mathcal{F}(x^n f) &= i^n \partial_k^n \mathcal{F} f\,, \\ \mathcal{F}(\partial_x^m f) &= (ik)^m \mathcal{F} f\,, \\ \frac{1}{\sqrt{2\pi}} \mathcal{F} f * g &= \mathcal{F} f \cdot \mathcal{F} g\,, \\ \mathcal{F} f_a &= e^{ika} \mathcal{F} f\,. \end{aligned}$$

Die letzte Gleichung sagt, daß eine Translation einer Funktion um a im Fourier-Raum einer Multiplikation mit einem Phasenfaktor e^{ika} entspricht; Beweis: trivial, man führt $x+a$ als neue Integrationsvariable ein.

Als Ergänzung schreibe ich die Fourier-Transformation für Funktionen mehrerer Veränderlicher auf. Sei $\vec{x} \in \mathbb{R}^n$, dann transformiert man wie bisher bezüglich aller Komponenten von $\vec{x}$:

$$\begin{aligned} (\mathcal{F} f)(\vec{k}) = \tilde{f}(\vec{k}) &= (2\pi)^{-n/2} \int f(\vec{x}) e^{-i\vec{k}\cdot\vec{x}} d^n x\,, \\ (\mathcal{F}^{-1} \tilde{f})(\vec{x}) = f(\vec{x}) &= (2\pi)^{-n/2} \int \tilde{f}(\vec{k}) e^{i\vec{k}\cdot\vec{x}} d^n x\,. \end{aligned} \tag{2.19}$$

2.3 Beispiele und Anwendungen

Ich beginne mit zwei sehr einfachen Beispielen.

1. Sei $I = [-\pi, \pi]$ und χ_I die zu diesem Intervall gehörende charakteristische Funktion. Ihre Fourier-Transformierte ist leicht zu berechnen:

$$\sqrt{2\pi}(\mathcal{F}\chi_I)(k) = \int_{-\pi}^{\pi} e^{-ikx}dx = \frac{e^{-ik\pi} - e^{ik\pi}}{-ik} = \frac{2\sin k\pi}{k}\,.$$

χ_I ist $\mathcal{L}^2$-Funktion, dasselbe gilt für $\mathcal{F}\chi_I$ wie es nach der allgemeinen Theorie sein muß. Man kann χ_I wiedergewinnen, wenn man z. B. den Trick mit dem konvergenzerzeugenden Faktor anwendet; allerdings muß man von den Punkten $\pm\pi$ absehen. Man sieht auch: χ_I ist zwar eine Funktion mit kompaktem Träger, aber nicht überall differenzierbar. Das hat zur Folge, das die Fourier-Transformierte keine rasch abfallende Funktion ist: sie fällt nur wie $1/k$ ab. Das ist ein Phänomen, das man sich merken sollte: Unstetigkeiten (oder gar Singularitäten) im Endlichen führen zu langreichweitigen Fourier-Transformierten.

2. Sei $f(x) = x\chi_I(x)$. Auch hier ist die Berechnung leicht (mit einer partiellen Integration):

$$\begin{aligned}\sqrt{2\pi}(\mathcal{F}f)(k) &= \int_{-\pi}^{\pi} xe^{-ikx}dx \\ &= \frac{xe^{-ikx}}{-ik}\Big|_{-\pi}^{\pi} - \int_{-\pi}^{\pi} \frac{e^{-ikx}}{-ik}\,dx \\ &= \frac{2\pi i}{k}\cos k\pi - \frac{2i}{k^2}\sin k\pi\,.\end{aligned}$$

Auch hier fällt die Fourier-Transformierte nicht schnell ab. (Sie ist aber, wie es sein muß, $\in \mathcal{L}^2$: $1/k^2$ ist im Unendlichen integrabel, und die Singularitäten bei $k = 0$ heben sich heraus.) Man kann sich die Ausgangsfunktion f (wie auch im ersten Beispiel) periodisch fortgesetzt denken, und dann läßt sie sich in eine Fourier*reihe* entwickeln, zur Berechnung der Koeffizienten muß man das gleiche Integral berechnen. Aber: für die Reihe ist k eine ganze Zahl; und das obige Ergebnis vereinfacht sich zu $c_k = (-1)^k\frac{2\pi i}{k}$. Doch wenn man eine Funktion darstellen will, die außerhalb von I verschwindet, dann muß man die kontinuierliche Fouriertransformation machen,und hier sind für die Darstellung eben sehr viel mehr Wellenlängen nötig, nämlich kontinuierlich viele.

3. Das klassische Beispiel ist die Fourier-Transformierte für eine Gaußsche Glockenkurve $f(x) = e^{-x^2/2a^2}$. Standardmethode: man berechnet $\int \exp(-ikx - x^2/2a^2)dx = \sqrt{2\pi}ae^{-a^2k^2/2}$, indem man im Exponenten eine quadratische Ergänzung vornimmt und das Integral zurückführt auf $\int e^{-y^2}dy = \sqrt{\pi}$. Um zu rechtfertigen, daß man von einem komplexen Exponenten übergehen kann zu einem reellen, muß man sich überlegen, daß das Integral eine analytische Funktion von k ist. Ein anderer, sehr eleganter Weg nutzt

direkt die Analytizität von e^{-z^2}, $z \in \mathbb{C}$: da (mit $z = x+iy$) $z^2 = x^2 - y^2 + 2ixy$ ist, sieht man, daß e^{-z^2} auf jeder Geraden $\mathcal{C}_c : \{z;\ \Im z = c\}$ für $|x| = |\Re z| \to \infty$ rasch abfällt; das Integral $I_c = \int_{\mathcal{C}_c} e^{-z^2} dz$ existiert und ist wegen der Analytizität von e^{-z^2} unabhängig von c (der Integrationsweg läßt sich nach dem Cauchyschen Integralsatz beliebig verschieben). Also ist $I_c = I_0 = \int e^{-x^2}\, dx = \sqrt{\pi}$. Andererseits ist $I_c = e^{c^2} \int_{-\infty}^{\infty} e^{-x^2 - 2icx}\, dx$ bis auf eine Variablentransformation und den richtigen Vorfaktor gerade die gesuchte Fourier-Transformierte.

Mit der Fourier-Transformation ist es möglich, der folgenden formalen Rechnung einen wohldefinierten Sinn zu geben ($\Delta = \sum \frac{\partial^2}{\partial x_i^2}$):

$$(\Delta + \kappa^2) f(\vec{x}) = g(\vec{x}) \,, \qquad \text{also} \quad f(\vec{x}) = (\Delta + \kappa^2)^{-1} g(\vec{x}) \,.$$

Fourier-transformiert lauten diese Gleichungen

$$(-|\vec{k}|^2 + \kappa^2) \tilde{f}(\vec{k}) = \tilde{g}(\vec{k}) \,, \qquad \text{also} \quad \tilde{f}(\vec{k}) = (-|\vec{k}|^2 + \kappa^2)^{-1} \tilde{g}(\vec{k}) \,;$$

und somit ist

$$\begin{aligned} f(\vec{x}) &= (2\pi^{-3/2}) \int \frac{\tilde{g}(\vec{k})}{\kappa^2 - |\vec{k}|^2} e^{i\vec{k}\cdot\vec{x}}\, d^3k \\ &= \frac{1}{(2\pi)^3} \int d^3k \int d^3x' \, \frac{g(\vec{x'})}{\kappa^2 - |\vec{k}|^2} e^{i\vec{k}\cdot(\vec{x} - \vec{x'})} \,. \end{aligned}$$

Damit haben wir das Inverse des Differentialoperators definiert. Man beachte aber, daß es kein „lokaler“ Operator mehr ist, d. h. der Wert der linken Seite am Punkt $\vec{x}$ hängt nicht nur vom Wert von g in einer Umgebung von $\vec{x}$ ab: auf der rechten Seite gehen *alle* Funktionswerte von g in das Integral ein. Mit der Fourier-Transformierten ist es somit möglich, die obige partielle Differentialgleichung zu lösen; und das gilt auch für viele weitere Differentialgleichungen.

Anwendungsbeispiel: Wärmeleitung

Wärme ist eine Energieform, Wärmeleitung: Energiestrom. In einem kontinuierlichen Medium sind die interessanten Größen die Energie*dichte* ε und die Energiestrom*dichte* $\vec{j}$: Wärmeinhalt in einem Volumen V : $E_V = \int_V \varepsilon(\vec{x}, t)\, d^3x$; und diese Größe kann sich zeitlich ändern, wenn es einen Energiestrom durch die Oberfläche ∂V gibt:

$$\frac{d}{dt} E_V = - \int_{\partial V} \vec{j}(\vec{x}, t) \cdot d\vec{\sigma} \,.$$

Minuszeichen: wenn der Wärmestrom nach außen gerichtet ist, nimmt die Energie innen ab. Die rechte Seite läßt sich mit dem Gaußschen Satz in ein Volumenintegral über $\mathrm{div}\vec{j}$ umwandeln; und weil das für jedes V gelten soll, müssen die beiden Integranden gleich sein:

$$\partial_t \varepsilon + \mathrm{div}\vec{j} = 0 \,.$$

(Eine solche Gleichung heißt Kontinuitätsgleichung, sie drückt hier die Energieerhaltung aus; in der Elektrodynamik gilt eine ebensolche Gleichung für die Erhaltung der elektrischen Ladung.) Der Wärmeinhalt eines Körpers ist proportional zu seiner Temperatur, der Proportionalitätsfaktor ist die Wärmekapazität des Körpers; folglich setze ich

$$\varepsilon(\vec{x},t) = c(\vec{x})\rho(\vec{x})T(\vec{x},t)\,.$$

Hier ist c die spezifische Wärme pro Masse, ρ die Dichte, also $c\rho$ die spezifische Wärme pro Volumen; T ist die Temperatur. Ansatz für den Wärmestrom: er folgt dem Temperaturgefälle, genauer, er ist proportional zum Temperaturgradienten (Fouriersches Gesetz):

$$\vec{j}(\vec{x},t) = -\lambda(\vec{x})\mathrm{grad}T\,.$$

Einsetzen in die Kontinuitätsgleichung gibt die Wäremleitungsgleichung (man beachte div·grad $T = \Delta T$):

$$c(\vec{x})\rho(\vec{x})\partial_t T(\vec{x},t) = \vec{\partial}\cdot(\lambda(\vec{x})\vec{\partial}T(\vec{x},t)) = (\vec{\partial}\lambda(\vec{x}))\cdot(\vec{\partial}T(\vec{x},t)) + \lambda(\vec{x})\Delta T(\vec{x},t)\,. \tag{3.1}$$

Zur Vereinfachung betrachten wir ein eindimensionales Problem und nehmen an, daß es sich um ein homogenes Material handelt; dann sind c, ρ und λ nicht ortsabhängig, die Gleichung vereinfacht sich zu

$$\partial_t T(x,t) = \kappa\,\partial_x^2 T(x,t)\,, \quad \kappa = \frac{\lambda}{c\rho}\,. \tag{3.2}$$

Fragen wir erst nach einer stationären Lösung $T = T(x)$: dafür muß $\kappa T''(x) = 0$ gelten, d. h. $T(x)$ ist eine lineare Funktion, die Temperatur fällt linear zwischen dem warmen und dem kalten Ende ab.
Um die allgemeine Lösung zu finden, führen wir eine Fourier-Transformation bezüglich x durch, die Differentiation nach x geht über in eine Multiplikation mit ik, aus Gleichung (3.2) wird dann eine gewöhnliche Differentialgleichung:

$$\partial_t\tilde{T}(k,t) = -\kappa k^2\tilde{T}(k,t)\,,$$

die sich für jedes k einfach lösen läßt:

$$\tilde{T}(k,t) = c(k)e^{-\kappa k^2 t}\,.$$

Hier ist $c(k)$ eine beliebige Integrationskonstante (für festes k); wenn zur Zeit $t = 0$ eine Temperaturverteilung $T_0(x)$ gegeben ist, wählt man

$$c(k) = \widetilde{T_0}(k)\,,$$

und dann ist

$$T(x,t) = \frac{1}{\sqrt{2\pi}}\int\tilde{T}(k,t)e^{ikx}dk = \frac{1}{\sqrt{2\pi}}\int\widetilde{T_0}(k)e^{-\kappa k^2 t + ikx}dk$$

die gesuchte Lösung von Gleichung (3.2) zur Anfangsbedingung

$$T(x,0) = \frac{1}{\sqrt{2\pi}} \int \widetilde{T_0}(k) e^{ikx} dk = T_0(x)\,.$$

Anwendungsbeispiel: Wellengleichung in der Elektrodynamik
In einer Sendeantenne fließen Ströme, bauen sich Ladungen auf und wieder ab, die dann das abgestrahlte elektromagnetische Feld bestimmen. Mathematische Beschreibung mit den Maxwellgleichungen. Ich schreibe sie auf für das Vakuum, d. h. $\vec{D} = \varepsilon_0 \vec{E}$, $\vec{H} = \mu_0 \vec{B}$; $\rho(\vec{x},t)$ ist die Ladungsdichte und $\vec{j}(\vec{x},t)$ die Stromdichte am Ort $\vec{x}$ zur Zeit t, die wir uns vorgegeben denken:

$$\begin{array}{llll} (a) & \vec{\partial} \cdot \vec{E} = \dfrac{\rho}{\varepsilon_0}\,; & (b) & \vec{\partial} \times \vec{E} + \partial_t \vec{B} = 0\,; \\ (c) & \vec{\partial} \times \vec{B} = \mu_0(\vec{j} + \varepsilon_0 \partial_t \vec{E})\,; & (d) & \vec{\partial} \cdot \vec{B} = 0\,. \end{array}$$

Außerhalb der Antenne gilt natürlich $\vec{j} = 0$, $\rho = 0$; Differenzieren von (c) nach t liefert unter Benutzung von (b):

$$\mu_0 \varepsilon_0 \partial_t^2 \vec{E} = \vec{\partial} \times \partial_t \vec{B} = -\vec{\partial} \times (\vec{\partial} \times \vec{E})\,.$$

Mit

$$\vec{\partial} \times (\vec{\partial} \times \vec{E}) = \vec{\partial}(\vec{\partial} \cdot \vec{E}) - (\vec{\partial} \cdot \vec{\partial})\vec{E} \overset{(a)}{=} -\Delta \vec{E}$$

und mit $c^2 := \mu_0 \varepsilon_0$ bekommen wir daraus

$$\left(\frac{1}{c^2} \partial_t^2 - \Delta \right) \vec{E} = 0\,,$$

d. h. jede Komponente E_i des elektrischen Feldes erfüllt die Wellengleichung

$$\left(\frac{1}{c^2} \partial_t^2 - \Delta \right) f(\vec{x},t) = 0\,. \tag{3.3}$$

(Das Gleiche gilt für B_i; man differenziere (b) nach t und benutze (c).) Ich habe die Wellengleichung in einer Dimension schon ganz im Anfang der Vorlesung erwähnt, eine mögliche Lösung ist $f(\vec{x},t) = g(\vec{e} \cdot \vec{x} - ct)$, wobei $g(\xi)$ eine differenzierbare Funktion und $\vec{e}$ ein Einheitsvektor ist. Das beschreibt einen Wellenberg der Form g, der sich in Richtung $\vec{e}$ mit der Geschwindigkeit c fortpflanzt: die Punkte $\{\vec{x};\ \vec{e} \cdot \vec{x} - ct = \text{const} = \xi\}$ bilden eine Ebene, auf der f überall den Wert $g(\xi)$ hat, sie verschiebt sich mit der Geschwindigkeit c in $\vec{e}$-Richtung. Allgemeinere Lösungen findet man durch Fourier-Transformation: setze

$$f(\vec{x},t) = \int_{\mathbb{R}^4} \tilde{f}(\vec{k},\omega) e^{i(\vec{k}\cdot\vec{x} - \omega t)} d^3k\, d\omega\,.$$

Das Minuszeichen bei ω ist Konvention. Aus der Differentiation ∂_t bzw. $\vec{\partial}$ wird dabei Multiplikation mit $-i\omega$ bzw. $i\vec{k}$:

$$\left(\frac{\omega^2}{c^2} - \vec{k} \cdot \vec{k} \right) \tilde{f}(\vec{k},\omega) = 0\,;$$

und diese Gleichung läßt sich leicht lösen, wir müssen nur verlangen, daß

$$\omega^2 = c^2|\vec{k}|^2 \tag{3.4}$$

gilt, wo $f \neq 0$ ist. Das bedeutet, daß $\tilde{f}$ tatsächlich nur eine Funktion von $\vec{k}$ ist, $\tilde{g}(\vec{k}) = \tilde{f}(\vec{k}, \omega(\vec{k}))$. Man sieht: für jedes feste $\vec{k}$ ist

$$\tilde{g}(\vec{k})e^{i(\vec{k}\cdot\vec{x} - c^2|\vec{k}|^2 t)} \tag{3.5}$$

eine Lösung; und weil die Wellengleichung linear ist, kann man solche Lösungen superponieren:

$$f(\vec{x}, t) = \int_{\mathbb{R}^3} \tilde{g}(\vec{k})e^{i(\vec{k}\cdot\vec{x} - \omega t)} d^3k \tag{3.6}$$

mit $\omega = \omega(k) = c|\vec{k}|$ ist Lösung, die man durch Wahl von g an die Anfangsbedingungen anpassen kann; direkte Überprüfung durch Einsetzen. In (3.6) tritt keine Integration über ω mehr auf im Gegensatz zu unserem obigen Ansatz. Der Grund ist natürlich, daß $\tilde{f}$ nur von $\vec{k}$ abhängt; eine saubere Begründung des Weglassens der ω-Integration gebe ich später mit der Theorie der Distributionen. (Ich hätte in diesem Beispiel genauso vorgehen können wie beim letzten: nur bezüglich $\vec{k}$ transformieren und die entstehende gewöhnliche Differentialgleichung lösen; Ergebnis: Gleichung (3.6).)
Die Lösungen aus Gleichung (3.5) haben Ebenen $\vec{k} \cdot \vec{x} = \omega t + \text{const}$ als Flächen konstanter Phase, es sind Wellenlösungen, bei denen sich diese Flächen in Richtung $\vec{k}$ mit der Geschwindigkeit $\omega/|\vec{k}| = c$ bewegen, daher der Name „ebene Welle". (Ein Physiker spricht auch dann von ebener Welle, wenn er den Faktor $e^{-i\omega t}$ wegläßt.) $\vec{k}$ heißt der Wellenvektor, $\lambda = \frac{2\pi}{|\vec{k}|}$ ist die Wellenlänge, d. h. der Abstand der Ebenen gleicher Phase.
Die allgemeine Lösung Gleichung (3.6) ist eine Überlagerung ebener Wellen in alle möglichen Richtungen, kann also recht kompliziert sein. Wenn man an das Problem des von einer Antenne abgestrahlten Feldes denkt, erwartet man weit weg von der Antenne keine ebene Welle als Lösung, sondern eine Kugelwelle: Flächen konstanter Phasen sind Kugeln, mathematisch: $f(\vartheta, \varphi)e^{i(kr - \omega t)}/r$ mit $r = |\vec{x}|,\ k = |\vec{k}|$. Diese Form ist nur dann eine exakte Lösung der Wellengleichung, wenn f konstant ist. Man erhält sie aber mit einem winkelabhängigen Faktor, wenn man die exakte Lösung für sehr große r betrachtet. Um die Lösung zu vorgegebenem ρ und $\vec{j}$ zu finden, benutzt man die Methode der Greenschen Funktionen, über die ich im Kapitel über Distributionen etwas erzählen will.

Anwendungsbeispiel: quantenmechanisches freies Teilchen
Ein mikroskopisches Teilchen zeigt Welleneigenschaften: man kann Interferenzen erzeugen. In der Quantenmechanik wird es deshalb durch eine „Wellenfunktion" $\psi(\vec{x}, t)$ beschrieben. Wenn man eine ebene Welle betrachtet: $\psi = e^{i(\vec{k}\cdot\vec{x} - \omega t)}$, analog zur Lichtwelle, dann ordnet man ihr den Impuls $\hbar\vec{k}$ zu, $\hbar = h/(2\pi)$, h = Plancksches Wirkungsquantum (dafür kann man Gründe aus der Elektrodynamik anführen), und — wiederum in

Analogie zum Licht — die Energie $E = h\nu = \hbar\omega$. (Theoretiker unterscheiden sich von Experimentalphysikern oft um einen Faktor 2π, siehe auch $\omega = 2\pi\nu$.) Diese Zuordnung kann man verallgemeinern, wenn man beachtet, daß im vorliegenden Fall

$$E\psi = i\hbar\partial_t\psi, \qquad \vec{p}\psi = \frac{\hbar}{i}\vec{\partial}\psi \tag{3.7}$$

ist. Die ebene Welle ist bezüglich $\vec{x}$ keine quadratintegrable Funktion, man kann also nicht die früher erwähnte Wahrscheinlichkeitsinterpretation vornehmen, die man in der Quantenmechanik machen möchte. Doch das „Wellenpaket“

$$\psi(\vec{x}) = \int_{\mathbb{R}^3} f(\vec{k})e^{i\vec{k}\cdot\vec{x}}d^3x \tag{3.8}$$

ist quadratintegrabel, wenn f eine $\mathcal{L}^2$-Funktion ist, siehe Theorem 2.2.12. In der Mechanik gilt für ein freies Teilchen der Masse m die Relation $E = m|\vec{v}|^2/2 = |\vec{p}|^2/(2m)$; wenn man die Zuordnung (3.7) beibehält und in diese Relation einsetzt, bekommt man die Schrödingergleichung für ein freies Teilchen:

$$i\hbar\partial_t\psi(\vec{x},t) = -\frac{\hbar^2}{2m}\Delta\psi(\vec{x},t)\,.$$

Auch hier findet man die Lösung wieder durch Fourier-Transformation: man schreibt ψ wie in Gleichung (3.8) als Überlagerung ebener Wellen, wobei f natürlich noch von t abhängt; mit $f(\vec{k},t) = g(\vec{k})e^{-i\omega t}$ mit $\omega = \hbar|\vec{k}|^2/(2m)$ bekommt man dann eine Lösung; und die Anfangsbedingung $\psi(\vec{x},0) = \varphi(\vec{x})$ kann man durch Wahl von g erfüllen. (Genauere Durchführung: Ü.)

Wenn man, wie es in der Quantenmechanik getan wird, dem Impuls den Ableitungsoperator $\frac{\hbar}{i}\vec{\partial}$ zuordnet (vergl.(3.7)), dann sieht man, daß in der Darstellung (3.8) $\mathcal{F}\left(\frac{\hbar}{i}\vec{\partial}\psi\right) = \hbar\vec{k}\,\mathcal{F}(\psi)$ ist; daher der Name „Impulsraum-Darstellung“ für die Beschreibung mit der Fourier-Transformierten.

2.4 Das Wiener-Khinchin-Theorem

In vielen Experimenten erhält man als Ergebnis zunächst irgend welche zeitlich variierenden Signale, die man analysieren möchte, z. B. Spannung oder Strom in einem Netzwerk an einem bestimmten Punkt, die Lautstärke eines Schallsignals, die Schwingungsamplitude eines elektrisch angeregten piezoelektrischen Kristalls usw. Typisch für viele solche Zeitsignale: sie sind nicht periodisch und gehen nicht gegen 0 für $t \to \infty$ (wobei „∞" natürlich de facto durch die Geduld des Experimentators begrenzt ist). Eine solche Funktion kann man nicht Fourier-transformieren und dadurch nach ihrem Frequenzgehalt analysieren. Doch wesentliche Information steckt in der „Autokorrelationsfunktion" des Signals, das mit $U(t)$ bezeichnet werde:

$$G(\tau) := \lim_{T\to\infty} \frac{1}{T} \int_0^T U(t+\tau)U(t)\,dt\,. \tag{4.1}$$

Diese Größe ist in guter Approximation experimentell zugänglich, auch wenn die Lebensdauer bzw. Geduld des Experimentators nur endlich ist. $U(t)$ ist hier als reelles Signal betrachtet; für $t < 0$ setze ich $U = 0$. Im allgemeinen nimmt man an, daß $\overline{U}^t := \lim \frac{1}{T} \int_0^T U(t)\,dt = 0$ ist, andernfalls betrachte man $U - \overline{U}^t$. Überstreichen mit Index t meint die zeitliche Mittelwertbildung (nicht das komplex Konjugierte).

Anmerkung: wenn das Signal für $-\infty < t < \infty$ definiert ist, betrachtet man $G(\tau) = \lim_{T\to\infty} \int_{-T/2}^{T/2} U(t+\tau)U(t)\,dt$.
Wenn $U \in \mathcal{L}^2$, dann ist $G(\tau) = 0$, Beweis mit der Schwarzschen Ungleichung:

$$|G(\tau)|^2 \leq \lim_{T\to\infty} \frac{1}{T^2} \int_0^T |U(t+\tau)|^2\,dt \int_0^T |U(t)|^2\,dt = 0\,. \tag{4.2}$$

Nehmen wir im Folgenden an, daß der Limes in Gleichung (4.1) existiert. Dann gilt

$$G(0) = \lim_{T\to\infty} \frac{1}{T} \int_0^T |U(t)|^2\,dt\,;$$

das ist im allgemeinen $\neq 0$, und die Wurzel aus dieser Größe ist ein Maß für das Zeitmittel von U. Weiter:

$$|G(\tau)| \leq G(0)\,,$$

das folgt aus der Ungleichung in(4.2), der mittlere Ausdruck ergibt

$$\lim(1/T) \int |U|^2 dt = G(0)\,.$$

.

Um doch Fourier-transformieren zu können, schneiden wir U ab, wir setzen

$$U_T(t) := U(t)\chi_{[0,T]}(t)\,.$$

(Das ist dem Experimentator unmittelbar zugänglich.) Fourier-transformiert:

$$\widetilde{U_T}(\omega) = \frac{1}{\sqrt{2\pi}} \int_0^T U(t) e^{-i\omega t} dt \,.$$

Wir wollen auch $g(t) := U_T(t+\tau)$ tranformieren: mit Lemma 2.2.18:

$$\tilde{g}(\omega) = e^{i\omega\tau} \widetilde{U_T}(\omega) \,.$$

Nun beachte man, daß U reell ist und wende die Gleichung (2.13) aus dem Theorem von Plancherel auf g und $f(t) := (1/T)U_T(t)$ an: $\int \overline{\tilde{f}} \tilde{g} d\omega = \int \overline{f} g dt$, eingesetzt:

$$\begin{aligned} \int_{-\infty}^{\infty} e^{i\omega\tau} \frac{|\widetilde{U_T}(\omega)|^2}{T} d\omega &= \frac{1}{T} \int_{-\infty}^{\infty} U_T(t+\tau) U_T(t) \, dt \\ &= \frac{1}{T} \int_0^T U(t+\tau) U(t) \chi_{[0,T]}(t+\tau) \, dt \,; \end{aligned} \tag{4.3}$$

und der letzte Ausdruck konvergiert für jedes feste τ im Limes $T \to \infty$ gegen $G(\tau)$:

$$G(\tau) = \lim_{T\to\infty} \int_{-\infty}^{\infty} e^{i\omega\tau} \frac{|\widetilde{U_T}(\omega)|^2}{T} d\omega \,. \tag{4.4}$$

Die Größe

$$S(\omega) := \lim_{T\to\infty} \frac{2\pi}{T} |\widetilde{U_T}(\omega)|^2 = \lim_{T\to\infty} \frac{1}{T} \left| \int_0^T e^{-i\omega t} U(t) \, dt \right|^2$$

heißt das „Leistungsspektrum von $U(t)$“. Es gilt das **Wiener-Khinchin-Theorem**:

Satz 2.4.5

$$\begin{aligned} G(\tau) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega\tau} S(\omega) \, d\omega \\ S(\omega) &= \int_{-\infty}^{\infty} e^{-i\omega\tau} G(\tau) \, d\tau \,. \end{aligned}$$

Diese Gleichungen bedürfen noch einiger Kommentare. Die erste Gleichung des Satzes besagt, daß G die inverse Fourier-Transformierte von S ist (bis auf einen Faktor $\sqrt{2\pi}$); die zweite Gleichung folgt also als Umkehrung der ersten. Die erste Gleichung erhält man, indem man in Gleichung (4.4) Limes und Integral vertauscht und die Definition von S benutzt. Der Haken dabei ist allerdings, daß diese Vertauschung im üblichen Sinn nicht erlaubt ist: $S = \lim \ldots$ existiert nur als Distribution; und die Gleichungen im Satz 2.4.5 sind im distributiven Sinn zu lesen. Was das heißt, erkläre ich genauer im nächsten Kapitel. Hier möge es uns genügen, daß experimentell die Größe U_T/T ja nur für endliche Zeiten T zu messen ist; S ist also angenähert bestimmt, und diese Approximation ist eine wohldefinierte Größe, die man Fourier-transformieren kann; man kann die Gleichungen als gute Approximationen lesen, wenn man die Limiten nicht streng ausführt.

Zur physikalischen Bedeutung des Leistungsspektrums: betrachten wir den Fall, daß $U(t)$ die Spannung in einem Stromkreis mit Ohmschem Widerstand R ist. Dann ist $U(t)^2/R$ die momentane Leistung, und

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T\frac{U(t)^2}{R}\,dt=\frac{G(0)}{R}$$

ist die mittlere Leistung. Nach Satz 2.4.5 können wir schreiben

$$\frac{G(0)}{R}=\frac{1}{R}\int_0^\infty\frac{S(\omega)}{\pi}\,d\omega\,,$$

wobei ich ausgenutzt habe, daß $S(-\omega)=S(\omega)$ ist. (Allgemein gilt, wie man leicht kontrolliert, für reelles $f(t):\ \tilde{f}(-\omega)=\overline{\tilde{f}(\omega)}$; aber S ist nach Definition reell.)

Interpretation:

$(S(\omega)/\pi)\,d\omega$ ist die Leistung im Frequenzintervall $d\omega$ bei $R=1$,

$S(\omega)/\pi$ ist die spektrale Leistungs*dichte* bei $R=1$.

Zum Schluß noch eine allgemeine Anmerkung. Am Beispiel der Gaußschen Glockenkurve sieht man, daß eine Funktion mit kleinem Träger als Fourier-Transformierte eine stark ausgedehnte Funktion hat, und umgekehrt. Das gilt — cum grano salis — auch allgemein: Lokalisierung im Ortsraum geht auf Kosten der Lokalisierung im Impulsraum, eine ebene Welle ist im Impulsraum ganz scharf lokalisiert und im Ortsraum überhaupt nicht.
Es gibt eine neue Methode der Analyse, die Ende der achtziger Jahre entwickelt wurde: die „wavelet“-Transformation. Die Sinus- und Cosinuswellen, nach denen man mit der Fourier-Reihe entwickelt, sind keine $\mathcal{L}^2$-Funktionen. Es ist zwar bekannt, daß es im $\mathcal{L}^2$ orthonormale Basen gibt, nach denen man entwickeln kann, doch sind das im allgemeinen keine so einfachen Familien von Funktionen wie Sinus und Cosinus, bei denen man die Entwicklung physikalisch interpretieren kann. Die wavelets stehen hier genau dazwischen: sie werden erzeugt durch eine einzige Funktion, die vielfach differenzierbar sein und kompakten Träger haben kann, und sie bilden eine o. n. Basis, z. B. $\left\{2^{k/2}\psi(2^kx-l);\ k,\,l\in\mathbb{Z}\right\}$, wobei $\psi(x)$ das „Mutter-wavelet“ ist. Mehr noch: bei geschickter Wahl von ψ genügen wenige Entwicklungskoeffizienten für eine gute Approximation, man erreicht gewissermaßen eine gleichzeitige Lokalisierung im Orts- wie im „Frequenzraum“ (genauer: im Parameterraum). Die wavelets haben in Anwendungen, z. B. bei der Signalverarbeitung und -übertragung, aber auch in der Mathematik sehr schnell große Bedeutung gewonnen; wenn Sie im UB-Bestand unter dem Stichwort „wavelets“ suchen, finden Sie fast dreißig Bücher und Konferenzberichte. Ich bin kein Experte auf diesem neuen Gebiet und habe leider auch keine Zeit, in dieser Vorlesung näher darauf einzugehen.

2.5 Die Laplace-Transformation

Auch hier kann ich keine mathematische Vollständigkeit anstreben.

Wenn man von der Laplace-Transformation spricht, meint man meist die „einseitige Laplace-Transformation“, die man verwendet, wenn man mit Funktionen zu tun hat, die auf der negativen Halbachse verschwinden. Andernfalls definiert man die zweiseitige Laplace-Transformation und ersetzt die im Folgenden auftretenden Integrale $\int_0^\infty$ durch $\int_{-\infty}^\infty$. Wir betrachten hier die einseitige Laplace-Transformation.

Sei $f \in \mathcal{L}^2((0,\infty), dt)$ — wir benutzen hier t als unabhängige Variable —, dann können wir f fortsetzen durch $f(t) = 0$ für $t < 0$, und erhalten $f \in \mathcal{L}^2(\mathbb{R})$. Wenn $x > 0$ ist, gilt auch $e^{-tx} \in \mathcal{L}^2(0,\infty)$, und wenn wir wieder mit 0 auf die negative Achse fortsetzen: $\chi_{(0,\infty)}e^{-tx} \in \mathcal{L}^2(\mathbb{R})$. Dann ist $fe^{-tx} \in \mathcal{L}^1(\mathbb{R})$, die Fourier-Transformierte dieser Funktion existiert:

$$F_x(y) = \frac{1}{\sqrt{2\pi}} \int_0^\infty f(t)e^{-tx}e^{-ity}dt\,. \tag{5.1}$$

Das Integral konvergiert gleichmäßig in y und in x für $x \geq x_0$ mit beliebigem $x_0 > 0$. Das gleiche gilt, wenn wir e^{-tx} durch te^{-tx} ersetzen; das entspricht bis auf einen Faktor der Ableitung von F_x nach x oder nach y. Wir können also $F_x(y)$ als komplexwertige Funktion von $z = x+iy$ betrachten, sie ist wohldefiniert *und holomorph* für $\Re z \equiv x > 0$. Wir werden im Folgenden der üblichen Konvention für die Laplace-Transformierte folgend den Faktor $\sqrt{2\pi}^{-1}$ weglassen:

Definition 2.5.2 *Die für $\Re z > 0$ holomorphe Funktion*

$$F(z) := \int_0^\infty f(t)e^{-tz}dt \tag{5.3}$$

heißt die Laplace-Transformierte von f.

Notation: Die Laplace-Transformierte von f bezeichne ich mit $\widehat{f}$ oder $\mathcal{L}f$.
Diese Definition soll immer dann gelten, wenn das Integral (5.3) existiert, auch wenn $f \notin \mathcal{L}^2$ ist. Das heißt, daß wegen des dämpfenden Faktors e^{-tx} auch polynomial wachsende Funktionen Laplace-transformierbar sind. Die Definition soll auch auf den Fall erweitert werden, daß $\widehat{f}(z)$ nur in der Halbebene $\Re z > x_0$ holomorph ist für passend gewähltes $x_0 > 0$. Beispiele: sei $f(t) = 1\ \forall t > 0$, dann ist $\widehat{f} = \int_0^\infty e^{-tz}dz = 1/z$ eine wohldefinierte holomorphe Funktion in der rechten Halbebene. Wenn $f = e^{ct},\ c > 0$, ist $\widehat{f} = 1/(z-c)$ holomorph für $\Re z > c$.

Für die letzte Aussage können wir in gewissem Sinn auch die Umkehraussage beweisen:

Lemma 2.5.4 *f sei derart, daß $\widehat{f}$ analytisch in der Halbebene $\Re z > c$ ist, meromorph in einer kleinen Umgebung von $z = c$ mit einem Pol erster Ordnung bei c. Dann verhält sich $f(t)$ für große t wie e^{ct}.*

Es ist klar, daß $1/(z - c)$ gerade diese Voraussetzungen erfüllt.

Beweis: Die Laplace-Transformierte von $g(t) := e^{-ct} f(t)$ läßt sich leicht durch diejenige von f ausdrücken:

$$\widehat{g}(z) = \int_0^\infty f(t) e^{-ct} e^{-zt}\, dt = \widehat{f}(z + c)\,. \tag{5.5}$$

$\widehat{g}(z)$ ist dann analytisch für $\Re z > 0$, hat einen einfachen Pol bei 0 und ist analytisch in einem kleinen Kreis um 0 außer bei 0: $\widehat{g}(z) = a(z)/z$, wobei $a(z)$ analytisch in $\{\Re z > 0\} \cup \{|z| < \delta\}$ ist, $a(0) \equiv a \neq 0$, insbesondere existiert die Ableitung $a'(0)$. Bei $z = 0$ verhält sich also $\widehat{g}$ wie a/z, d. h. wie die Laplace-Transformierte der Konstanten a. Das legt nahe, $g(t)$ in der Form $g(t) = a(1 + b(t))$ zu schreiben; folglich ist

$$\widehat{b}(z) = \frac{\widehat{g}(z)}{a} - \frac{1}{z} = \frac{z\widehat{g} - a}{az} = \frac{a(z) - a}{az}\,.$$

Mit der l'Hospitalschen Regel folgt

$$\lim_{z \to 0} \widehat{b}(z) = \frac{a'(0)}{a}\,,$$

die rechte Seite existiert nach Voraussetzung. Die linke Seite ist gleich

$$\lim_{x \to 0} \int_0^\infty b(t) e^{-tx}\, dt\,,$$

und das heißt, daß $b(t)$ für alle $x > 0$ schwächer wachsen muß als e^{xt}. Folgerung: das Wachstum von $f(t) = e^{ct} g(t) = a(1 + b(t)) e^{ct}$ wird durch den Exponentialfaktor dominiert. ∎

Ich will eine Umkehrformel angeben und ableiten für den Fall, daß f hinreichend brav ist: sei $f \in \mathcal{L}^2((0, \infty), dt)$. Dann ist $\widehat{f}(x + iy) \in \mathcal{L}^2(\mathbb{R}, dy)$, denn diese Funktion kann nach (5.1) als Fourier-Transformierte von $f(t)e^{-tx}$ angesehen werden, und nach der Parsevalschen Gleichung (2.10) ist

$$\int_{-\infty}^\infty |\widehat{f}(x + iy)|^2 dy = 2\pi \int_0^\infty |f(t)|^2 e^{-2tx} dt \leq 2\pi \int_0^\infty |f(t)|^2 < \infty\,.$$

Dann können wir f — wenn es hinreichend brav ist — durch die inverse Fourier-Transformation wiedergewinnen:

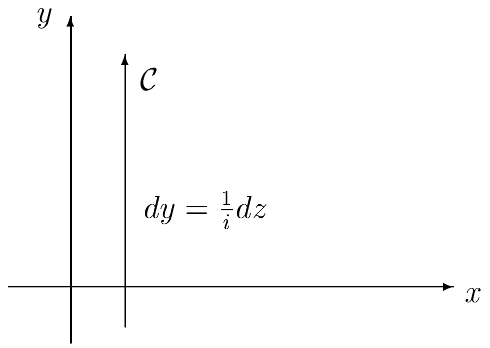
$$f(t) e^{-tx} = \frac{1}{2\pi} \lim_{R \to \infty} \int_{-R}^R \widehat{f}(x + iy) e^{ity} dy\,. \tag{5.6}$$

Multiplizieren wir mit e^{tx} und betrachten die rechte Seite als ein komplexes Wegintegral über den Weg

$$\mathcal{C}_x : \{z \in \mathbb{C};\ \Re z = x\} \ .$$

Auf diesem Weg ist $dz = i\,dy$, er liegt (für $x > 0$) ganz im Holomorphie-Gebiet von $\widehat{f}$:

$$f(t) = \frac{1}{2\pi i} \int_{\mathcal{C}_x} \widehat{f}(z) e^{tz} dz = (\mathcal{L}^{-1}\widehat{f})(t)\,. \tag{5.7}$$



Ich habe hier sofort den Limes $R \to \infty$ gemacht, das komplexe Wegintegral ist im Sinn von Gleichung (5.6) gemeint. Man beachte: es ist nach dem Cauchyschen Integralsatz unabhängig von x. Wenn man mit einer Funktion $\widehat{f}$ zu tun hat, die in der Halbebene $\Re z > x_0$ holmorph ist, muß man $x > x_0$ wählen.
Wenn f nicht so schön ist, wie oben angenommen, kann man, wie bei der Fourier-Transformation, durch kompliziertere Operationen f aus $\hat{f}$ wiedergewinnen.

Die Laplace-Transformierte ist, wie die Fourier-Transformierte, ein nützliches Werkzeug bei der Lösung von Differentialgleichungen. Für das folgende Anwendungsbeispiel brauche ich die Laplace-Transformierte für die Faltung zweier Funktionen. Seien f und g auf $\mathbb{R}^+$ definiert und $= 0$ gesetzt für $t < 0$. Dann ist, wie man sofort nachprüft,

$$(f * g)(t) = \int_0^t f(t')g(t-t')\,dt' = \int_0^t f(t-t')g(t')\,dt'\,, \tag{5.8}$$

und es gilt

$$\widehat{f * g} = \widehat{f}\,\widehat{g}\,. \tag{5.9}$$

Auch das ist leicht zu prüfen (vergl. (2.17)).

Man kann die Laplace-Transformation benutzen, um eine Integro-Differentialgleichung zu lösen, die bei einem einfachen Populationsmodell auftritt:

$$\partial_t n(x,t) = -\omega(x)\partial_x n(x,t) - \mu(x)n(x,t) + s(x) \int \sigma(x')n(x',t)\,dx'\,.$$

$\omega(x), \mu(x), s(x)$ und $\sigma(x)$ sind dabei bekannte Funktionen. Die Gleichung sieht kompliziert aus; ist aber vollständig lösbar, weil sie linear ist. (Das ist auch die wesentliche

Schwäche des Modells, ein realistisches Modell sollte durch eine nicht-lineare Gleichung beschrieben werden.) Die Darstellung und Lösung des Modells findet sich im Anhang A.

Kapitel 3

Distributionen

3.1 Einleitung und Definitionen

Anlaß zur Einführung von Distributionen war der Wunsch Dirac's und anderer Physiker nach der Dirac'schen „Delta-Funktion“: man hätte gern eine nicht-negative Funktion, die nur an einem Punkt $\neq 0$ ist und die integriert 1 ergibt: $\int \delta(x)dx = 1$. Man kann so etwas approximieren, z. B. durch Funktionen

$$d_a(x) := \begin{cases} a^{-1} & \text{für } |x| < \dfrac{a}{2}\,, \\ 0 & \text{sonst}\,. \end{cases} \tag{1.1}$$

Was kann man damit anfangen? Ich bringe ein Beispiel mit naiver (rein formaler) Rechnung; dabei benutze ich die dreidimensionale δ-Funktion $\delta(\vec{x})$, die nur für $\vec{x} = 0$ ungleich null ist und $\int \delta(\vec{x})\, d^3x = 1$ erfüllt.

In der Elektrostatik möchte man zu einer gegebenen Ladungsdichteverteilung $\rho(\vec{x})$ das elektrische Feld berechnen. Man erhält es aus den Maxwell-Gleichungen — ich führe die Rechnung für das Vakuum durch —

$$\vec{\partial} \cdot \vec{E} = \frac{\rho}{\varepsilon_0}\,, \qquad \vec{\partial} \times \vec{E} + \partial_t \vec{B} = 0\,.$$

Im statischen Fall ist $\partial_t \vec{B} = 0$; also ist $\vec{E}$ rotationsfrei und läßt sich als Gradient einer skalaren Funktion schreiben: $\vec{E} = -\operatorname{grad} \Phi$. Aus der ersten Gleichung folgt dann, daß das Potential $\Phi(\vec{x})$ die Laplace-Gleichung

$$\Delta\Phi = -\frac{1}{\varepsilon_0}\rho$$

erfüllt. Diese Gleichung ist also zu lösen.

Man beachte: ρ ist die Ladungs*dichte*, d. h. $\int_V \rho(\vec{x})d^3x$ ist die Ladung im Volumen V.

Wenn man nun eine Punktladung der Stärke q am Punkt x_0 vor sich hat, ist ρ nur an einem Punkt $\neq 0$, und die Gesamtladung ist $\int \rho(\vec{x})d^3x = q$, die Ladungsdichte ρ ist also δ-artig: $\rho(\vec{x}) = q\delta(\vec{x}-\vec{x_0})$. Wie in der Mechanik, wo man Massen*punkte* betrachtet, möchte man in der Elektrodynamik gern in zwei Richtungen idealisieren können: zum einen auf punktförmige Ladungen (das Elektron ist ja wirklich sehr klein), zum anderen auf kontinuierliche Ladungsdichten (man sieht ja, wenn man makroskopische Physik treibt, nicht in die Feinheit atomarer Strukturen hinein); die δ-Funktion liefert die Verbindung zwischen beiden Idealisierungen. Man kennt das Feld einer Punktladung (Coulomb-Gesetz) und demzufolge auch das zugehörige Potential:

$$\Phi(\vec{x}) = \frac{q}{4\pi\varepsilon_0|\vec{x} - \vec{x_0}|} \,. \tag{1.2}$$

Das sollte eine Lösung der Gleichung

$$\Delta\Phi = -\frac{q}{\varepsilon_0}\delta(\vec{x} - \vec{x_0}) \tag{1.3}$$

sein. (Für $\vec{x} \neq \vec{x_0}$ kann man direkt nachrechnen, daß $\Delta\Phi = 0$ ist.) Angenommen, Gleichung (1.3) werde durch (1.2) gelöst (Beweis später), dann kann man mit dieser Lösung für $q = 1$ — sie heiße $G(\vec{x}, \vec{x_0})$ — die Laplace-Gleichung für beliebiges ρ leicht lösen: man stellt sich vor, daß die Ladungsverteilung aufgebaut ist aus unendlich vielen infinitesimalen Punktladungen der Stärke $\rho(x)d^3x$; die Gleichung ist linear, man kann also die Felder und auch die Potentiale dieser Punktladungen einfach superponieren. Das suggeriert den Lösungsansatz

$$\Phi(\vec{x}) = \int G(\vec{x}, \vec{y})\rho(\vec{y})\, d^3y \,.$$

Eingesetzt in die Laplace-Gleichung: ich ziehe den Differentialoperator Δ unter das Integral, er wirkt auf das erste Argument von G,

$$\Delta\Phi = \int \Delta_x G(\vec{x}, \vec{y})\rho(\vec{y})\, d^3y = \int \delta(\vec{x} - \vec{y})\rho(\vec{y})\, d^3y \,.$$

Wegen der schönen Eigenschaften von δ kann man dieses Integral leicht lösen: der Integrand ist nur $\neq 0$ für $y = x$, nach dem Mittelwertsatz bekommen wir also

$$\ldots = \rho(\vec{x}) \int \delta(\vec{x} - \vec{y})d^3y = \rho(\vec{x}) \,,$$

wir haben also eine Lösung gefunden.

Wie ich hier gerechnet und argumentiert habe, ist das Ganze fauler Zauber; ich habe eine unmögliche Funktion benutzt, ich habe einfach Δ und $\int$ vertauscht, ich habe den Mittelwertsatz für eine nicht existierende Funktion angewandt. Aber alles das läßt sich mit der Theorie der Distributionen rechtfertigen: die Rechnungen sind korrekt, wenn sie „im distributiven Sinn“ gelesen werden.

Zur weiteren Motivation: man möchte sich beim Differenzieren nicht so leicht behindern lassen. Wenn

$$f(x) = \begin{cases} x & \text{für } x > 0\,, \\ 0 & \text{für } x < 0\,, \end{cases}$$

ergibt die Differentiation die Heaviside'sche Stufenfunktion

$$\Theta(x) = \begin{cases} 1 & \text{für } x > 0\,, \\ 0 & \text{für } x < 0\,, \end{cases}$$

Nochmaliges Differenzieren ergibt 0 für $x \neq 0$ und ∞ für $x = 0$. Daß man tatsächlich $\Theta'(x) = \delta(x)$ setzen kann, sieht man aus folgender Überlegung: weil die Sprunghöhe der Θ-Funktion $= 1$ ist, ergibt das Riemann-Stieltjes-Integral $\int d\Theta = 1$, also

$$1 = \int d\Theta = \int \Theta'(x)\, dx\,,$$

wie es für die Delta-Funktion sein muß. Wir werden sehen, daß man im Bereich der Distributionen beliebig differenzieren, Fourier-transformieren, Integrale und Summen mit Limiten vertauschen kann in Fällen, in denen es für Funktionen schlicht falsch wäre.

Man beachte einen wesentlichen Punkt: uns interessieren nie die Funktions*werte* der Delta-Funktion, sondern wenn es um Zahlenwerte geht, tritt sie nur im Integranden auf (Beispiel Punktladung: hier ist die Gesamtladung wesentlich, d. h. das Integral über die δ-Funktion). Und das Integral mit der Delta-Funktion ist linear in ρ im obigen Beispiel. Das ist der wesentliche Hinweis, wie man zu einer sauberen Definition kommt: die Delta-Funktion definiert ein lineares Funktional auf dem Raum der „Testfunktionen".

Definition 3.1.4 *$\mathcal{T}$ sei ein linearer topologischer Funktionenraum, er heiße Raum der Testfunktionen. Der zu $\mathcal{T}$ topologisch duale Raum, d. h. die Menge der stetigen linearen Funktionale über $\mathcal{T}$, heißt der Raum Distributionen über $\mathcal{T}$.*

In vielen mathematischen Texten wird der zu $\mathcal{T}$ topologisch duale Raum statt mit $\mathcal{T}^*$, wie ich es in der Definition 1.1.19 getan habe, mit $\mathcal{T}'$ bezeichnet; es ist die übliche Notation, die man benutzt, wenn man von Distributionenräumen spricht. Ich werde mich dieser Notation anschließen. Ich erinnere daran, daß der duale Raum selbst wieder ein linearer Raum ist (Lemma 1.1.20).

Uns interessiert $\mathcal{T} = \mathcal{S}(\mathbb{R})$ und $\mathcal{T} = \mathcal{D}(\mathbb{R})$. $\mathcal{S}'$ heißt der Raum der „temperierten Distributionen", es ist also die Menge der stetigen linearen Funktionale über der Raum $\mathcal{S}$ der temperierten Funktionen mit der in Definition 1.1.6 gegebenen Topologie. Weil wir im Folgenden diese Topologie oft brauchen, wiederhole ich noch einmal die Definition, die durch Halbnormen gegeben wird:

$$\begin{aligned} p_{mn}(f) \quad &:= \quad \sup_x \left| x^m f^{(n)}(x) \right|\,, \\ f_\alpha \underset{\alpha\to\infty}{\longrightarrow} f \quad &:\Longleftrightarrow \quad p_{mn}(f_\alpha - f) \underset{\alpha\to\infty}{\longrightarrow} 0\ \forall m,\, n\,. \end{aligned}$$

Man beachte: $\mathcal{D} \subset \mathcal{S}$, und deshalb $\mathcal{S}' \subset \mathcal{D}'$.
Später werde ich kurz auf die naheliegende Verallgemeinerung eingehen: $\mathcal{S}'(\mathbb{R}^n)$.

Beispiele:
δ-Distribution: $\delta : \ f \mapsto \delta(f) := f(0)$. Es ist leicht zu sehen, daß diese Abbildung stetig in der $\mathcal{S}$-Topologie ist: $f_\alpha \to f$ heißt $p_{mn}(f - f_\alpha) \to 0 \ \forall m, n$, also insbesondere für $m = 0$ und $n = 0$: $\sup_{x\in\mathbb{R}} |f - f_\alpha| \to 0$, folglich erst recht $|f(0) - f_\alpha(0)| \to 0$.
Da $\mathcal{S} \subset \mathcal{C}_0^0$ ist, können wir auf Beispiel 1.1.22 zurückgreifen: $\Phi_1(f) = \int f(x)e^{-x^2}dx$ und $\Phi_2(f) = f(x_0)$ sind auch lineare Funktionale über $\mathcal{S}$; die Stetigkeit in der $\mathcal{S}$-Topologie ist leicht zu überprüfen: die Topologie von $\mathcal{S}$ ist stärker als die von $\mathcal{C}_0^0$, ein Funktional, das in der schwächeren Topologie stetig ist, ist es erst recht in der stärkeren (die Menge der konvergenten Folgen ist kleiner in der stärkeren Topologie).
Merke: Abschwächen der Topologie im Bildraum oder Verstärken der Topologie im Urbildraum führt zu schwächerer Anforderung an die Stetigkeit.
Übliche Notation für Φ_2: $\delta_{x_0} : f \mapsto f(x_0)$. Das entspricht einer an die Stelle x_0 verschobenen δ-Distribution und ist die mathematisch saubere Formulierung dessen, was ich in der Einleitung mit $\delta(x - x_0)$ bezeichnet habe.
Beispiel Φ_1 läßt sich verallgemeinern: jede Lebesgue-meßbare Funktion g, die nicht stärker als polynomial anwächst, definiert eine Distribution

$$T_g : f \mapsto T_g(f) := \int \overline{g(x)} f(x)\, dx\,. \tag{1.5}$$

Die rechte Seite sieht aus wie das Skalarprodukt im $\mathcal{L}^2(\mathbb{R})$, doch weil die Funktion f außergewöhnlich brav ist und nicht nur quadratintegrabel, können wir für g sogar anwachsende Funktionen zulassen.
Distributionen dieser Art heißen „regulär“.
(Als Abkürzung schreibe ich „hpw“ für „höchstens polynomial wachsend“.)

In Analogie zu Distributionen der Art T_g aus Gleichung (1.5) schreibt man gern alle Distributionen *formal* als Integralkern:

$$T(f) = \int T(x)f(x)dx\,,$$

auch dann, wenn wie z. B. bei der Delta-Distribution keine *Funktion* $\delta(x)$ existiert. (Die Notation ist ein bißchen inkonsequent: man schreibt nicht die komplex konjugierte „Funktion“ $\overline{T(x)}$ in das Integral, wie man es in (1.5) macht; gegebenenfalls aufpassen!) Damit kann man z. B. die Distribution $\delta_{x_0} : f \mapsto f(x_0)$ sehr suggestiv schreiben:

$$\delta_{x_0}(f) = \int \delta(x - x_0)f(x)\, dx = \int \delta(x)f(x + x_0)\, dx = f(x_0)\,.$$

Ich werde diese Schreibweise auch benutzen, betone aber, daß sie **rein formal** gemeint ist. Der Grund ist, daß man viele Manipulationen mit Distributionen intuitiv anschaulicher schreiben kann. Aber: wenn man unsicher ist, was man tun darf, benutze man die

korrekte Formulierung als lineares Funktional! In einigen Fällen kann einen die formale Schreibweise zu falschen Schlüssen verleiten, im übrigen ist sie sehr bequem und macht die Nähe der Distributionen zu Funktionen deutlich. Wir werden später beweisen, daß alle temperierten Distributionen in wohldefiniertem Sinn durch gewöhnliche Funktionen approximiert werden können, sogar durch Funktionen aus $\mathcal{C}^\infty$.
Entsprechend dieser formalen Schreibweise werde ich auch kurz von der „Distribution g“ reden, wenn ich T_g meine.
Sie finden auch eine Funktionalschreibweise, die sich an die Notation für das Skalarprodukt anlehnt: $T(f) \equiv (T, f)$; und oft ist dann eine Variante nützlich, die der formalen Schreibweise nahe ist: $(T(x), f(x))$, d. h. mit Andeutung der Funktionsvariablen. Das ist z. B. bei δ_{x_0} nützlich: $(\delta(x - x_0), f(x)) = (\delta(x), f(x + x_0))$ sagt, bezüglich welcher Variable die Funktion als Testfunktion betrachtet wird, ohne zu suggerieren, daß man hier mit einem Integral über gewöhnliche Funktionen zu tun hat. Ich werde auch diese Schreibweise benutzen.

Wir wollen jetzt einen Konvergenzbegriff im Raum der Distributionen erklären. Das machen wir ähnlich wie beim Begriff der schwachen Konvergenz, die ich im Hilbertraum definiert habe (Definition 1.3.9).

Definition 3.1.6 *Sei $\{T_\alpha;\ \alpha \in \mathbb{N}\}$ eine Folge von Distributionen in $\mathcal{T}'$. Die Folge konvergiert gegen $T \in \mathcal{T}'$, wenn*

$$\lim_{\alpha\to\infty} T_\alpha(f) = T(f)$$

für alle Testfunktionen $f \in \mathcal{T}$ gilt.

Die Konvergenz muß nicht gleichmäßig in f sein, dieser Konvergenzbegriff ist relativ schwach. (Weder $\mathcal{S}$ noch $\mathcal{D}$ sind normierte Räume; daher kann man auch auf $\mathcal{S}'$ oder $\mathcal{D}'$ keine Norm einführen. Analog zur Definition 1.3.9 — man beachte, daß $\mathcal{H}$ zu sich selbst dual ist — kann man auf einem topologischen Raum $\mathcal{X}$ eine schwache Topologie einführen: $x_\alpha \longrightarrow x$, wenn $f(x_\alpha) \longrightarrow f(x)$ für alle $f \in \mathcal{X}^*$. Dasselbe kann man auf $\mathcal{X}^*$ machen: man braucht die Elemente von $\mathcal{X}^{**}$. Nun kann man sich $\mathcal{X}$ in $\mathcal{X}^{**}$ eingebettet denken und eine noch schwächere Topologie auf $\mathcal{X}^*$ definieren: $f_\alpha \longrightarrow f$, wenn $f_\alpha(x) = f(x)$ für alle $x \in \mathcal{X}$; man spricht dann von schwacher *-Konvergenz, englisch: weak* convergence.). Das ist die oben definierte Konvergenz in $\mathcal{T}'$.)

Beispiel: Die Funktionen d_a (Gleichung (1.1)) definieren eine Folge von Distributionen

$$T_{d_a} : f \mapsto \int d_a(x) f(x) = \frac{1}{a} \int_{-a/2}^{a/2} f(x)\, dx\,,$$

nach dem Mittelwertsatz ist die rechte Seite $= f(\xi)$ mit einem passenden $\xi \in [-a/2, a/2]$, sie konvergiert für $a \to 0$ offensichtlich gegen $f(0)$; also:

$$\lim_{a\to 0} T_{d_a}(f) = \delta(f) \quad \forall f \in \mathcal{S}\,.$$

T_{d_a} konvergiert also gegen δ; das ist die korrekte Darstellung dessen, was ich am Anfang in naiver Manier erzählt habe:

$$\lim_{a\to 0} d_a(x) = \delta(x)\,.$$

Diese Gleichung lesen wir jetzt als die formale Schreibweise der darüber stehenden.

Beispiel: Die (höchstens polynomial wachsenden) Funktionen g_α mögen fast überall gegen g konvergieren; dann gilt

$$\lim_{\alpha\to\infty} T_{g_\alpha} = T_g\,.$$

Beweis ziemlich trivial.

Wir haben ein weiteres Beispiel für die Approximation der Delta-Funktion kennengelernt: die Funktionen $f^{(a)}$ aus Lemma 2.2.7, die bei geeigneter Wahl sogar $\in \mathcal{S}$ sein können.

Entsprechend der formalen Schreibweise schreibe ich auch kurz $f^{(a)} \to \delta$.

Distributionen treten auch auf, wenn man Integrale „regularisieren", d. h. eigentlich nicht existenten Integralen einen Sinn geben will. Beispiel: $\int \frac{f(x)}{x}\,dx$ ist wohldefiniert, wenn f in einer Umgebung der 0 verschwindet (ich betrachte $f \in \mathcal{S}$, so daß es keine Schwierigkeiten für $x \to \infty$ gibt). Gesucht ist eine geänderte Definition, die es erlaubt, das Integral für jede Testfunktion $f \in \mathcal{S}$ hinzuschreiben, die aber für jedes f, das bei Null verschwindet, das alte Resultat ergibt. Eine solche Regularisierung ist nicht eindeutig. Eine Möglichkeit: Bildung des *Hauptwert-Integrals*

$$\text{P-}\!\int \frac{f(x)}{x}\,dx := \lim_{\varepsilon\to 0}\left(\int_{-\infty}^{-\varepsilon} \frac{f(x)}{x}\,dx + \int_{\varepsilon}^{\infty} \frac{f(x)}{x}\,dx\right)\,.$$

(Andere Schreibweise: $\text{P-}\!\int \equiv \fint$. P steht für 'principal value'.)
Anschaulich: wenn f bei 0 symmetrisch ist, dann ist der Integrand für $x \in (-\varepsilon, \varepsilon)$ antisymmetrisch, und die Singularität hebt sich heraus; wenn f antisymmetrisch ist, muß $f(0) = 0$ sein; man erwartet also, daß $\lim_{x\to 0} f(x)/x$ und damit das Integral existiert.
Man kann das Hauptwert-Integral als Distribution auffassen: $\mathrm{P}\frac{1}{x}$, denn es ist ein lineares stetiges Funktional. Anders definiert:

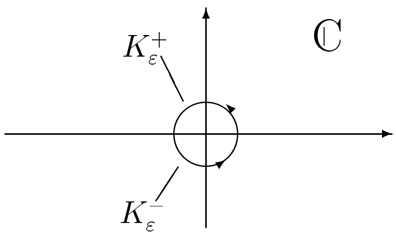
$$\left(\mathrm{P}\frac{1}{x}\right)(f) := \int_{-\infty}^{-1} \frac{f(x)}{x}\,dx + \int_{-1}^{1} \frac{f(x) - f(0)}{x}\,dx + \int_{1}^{\infty} \frac{f(x)}{x}\,dx\,. \tag{1.7}$$

Das mittlere Integal ist wohldefiniert: man benutze die Taylor-Entwicklung von f bei 0,

$$f(x) = f(0) + xf'(0) + f_R(x);$$

dann sieht man, daß der Integrand $(f(x) - f(0))/x$ bei 0 stetig ist, weil $f_R(0) = 0$ und $f'_R(0) = 0$, und somit $\lim_{x\to 0} f_R(x)/x = f'_R(0) = 0$ (l'Hospitalsche Regel).
Trivial: die Wahl der Grenzen ± 1 ist willkürlich, der Wert der rechten Seite von (1.7) bleibt derselbe, wenn man stattdessen $\pm\varepsilon$ wählt.

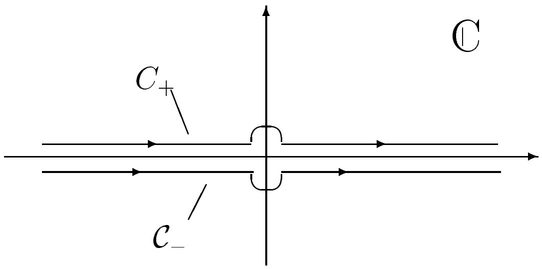
Wenn man Testfunktionen betrachtet, die in einer Umgebung der Null analytisch sind, kann man das Hauptwert-Integral noch anders ausdrücken. (Beachte: $f \in \mathcal{S}$ ist zwar beliebig oft differenzierbar bei 0, aber die Taylor-Reihe muß die Funktion nicht darstellen; f muß also nicht analytisch bei 0 sein. Gegenbeispiel: alle Ableitungen von e^{-x^2-1/x^2} verschwinden bei 0. Analytische Testfunktionen treten bei Fourier-Transformation von Funktionen $\in \mathcal{D}$ auf, siehe später.)



Sei K_ε^+ ein kleiner Halbkreis vom Radius ε um 0 in der oberen Halbebene. Weil f stetig ist, gilt

$$\int_{K_\varepsilon^+} \frac{f(z)}{z}\, dz \approx f(0) \int_{K_\varepsilon^+} \frac{dz}{z} = f(0) \int_0^\pi \varepsilon^{-1} e^{-i\varphi}\, d(\varepsilon e^{i\varphi}) = \pi i f(0)\,;$$

und das gleiche Resultat erhält man für das Integral über K_ε^- (Integration über φ von π
bis 2π.) Wenn man also den Integrationsweg für das Hauptwert-Integral einmal durch den Halbkreis K_ε^- zu einem Weg $\mathcal{C}_-$ ergänzt und einmal durch den negativ durchlaufenen Halbkreis K_ε^+ zu $\mathcal{C}_+$ und dann $\int_{\mathcal{C}_-} \frac{f(z)}{z}\, dz + \int_{\mathcal{C}_+} \frac{f(z)}{z}\, dz$ bildet, heben sich die Halbkreisanteile weg und übrig bleibt



$$\text{P-}\!\int \frac{f(x)}{x}\, dx = \frac{1}{2} \lim_{\varepsilon\to 0} \left(\int_{\mathcal{C}_-} \frac{f(z)}{z}\, dz + \int_{\mathcal{C}_+} \frac{f(z)}{z}\, dz \right) .$$

Man schreibt dafür formal

$$\text{P-}\!\int \frac{f(x)}{x}\, dx = \frac{1}{2} \int \left(\frac{f(x)}{x+i0} + \frac{f(x)}{x-i0} \right) dx\,. \tag{1.8}$$

Die Größen $(x \pm i0)^{-1}$ kann man als Distributionen über analytischen Testfunktionen auffassen:

$$\text{P}\frac{1}{x} = \frac{1}{2} \left(\frac{1}{x+i0} + \frac{1}{x-i0} \right) . \tag{1.9}$$

Wenn man die Differenz der Integrale über $\mathcal{C}_-$ und $\mathcal{C}_+$ bildet, bleibt nur das Integral über den vollen Kreis um 0 übrig, d. h. $2\pi i f(0)$, also ist

$$\frac{1}{x+i0} - \frac{1}{x-i0} = 2\pi i \delta(x)\,. \tag{1.10}$$

Integrale dieser Art treten z. B. bei quantenmechanischen Streuproblemen auf.

Ich komme später noch einmal auf die Regularisierung von Integralen zurück.

3.2 Operationen mit Distributionen

Als Funktionen betrachtet sind Distributionen meist viel schlimmer als unstetige Funktionen, doch als Distributionen sind sie so schön, daß man sie beliebig oft differenzieren kann!

Methode: Man definiert die Ableitung (und andere Operationen) durch Überwälzen auf die Testfunktionen, auf denen sie operieren.
Leitprinzip: Alle Definitionen, die Distributionen betreffend, sollen so gewählt werden,daß sie für reguläre Distributionen (d. h. Distributionen T_g, die durch normale Funktionen gegeben sind), mit den üblichen Definitionen für Funktionen übereinstimmen.

Die formale Schreibweise macht die Definition der Ableitung plausibel:
Mit formaler partieller Integration und Vernachlässigung der Randterme, weil f ja stark abfällt, erhalten wir

$$\int T'(x)f(x)\,dx = -\int T(x)f'(x)\,dx\,.$$

Dementsprechend:

Definition 3.2.1 *Die Ableitung T' einer Distribution $T \in \mathcal{S}'$ ist gegeben durch*

$$T'(f) := -T(f') \quad \forall f \in \mathcal{S}\,.$$

Wenn $f_\alpha \longrightarrow f$ in der Topologie von $\mathcal{S}$ gilt, dann folgt natürlich auch $f'_\alpha \longrightarrow f'$; daher ist, weil $T \in \mathcal{S}'$ ein stetiges Funktional ist, auch das so definierte T' stetig, $T' \in \mathcal{S}'$. Da die Testfunktionen beliebig oft differenzierbar sind, kann man die Prozedur beliebig oft wiederholen, also ist T beliebig oft differenzierbar.

Beispiel: Die Heaviside-Funktion

$$\Theta(x) = \begin{cases} 1 & x > 0\,, \\ 0 & x < 0 \end{cases}$$

definiert eine Distribution T_Θ. Es gilt $T'_\Theta = \delta$; denn:

$$T'_\Theta(f) = -T_\Theta(f') = -\int_0^\infty f'(x)\,dx = f(0)\,.$$

Nochmalige Ableitung ergibt

$$\delta'(f) = -f'(0)\,.$$

Wenn die Funktionen $f^{(a)} \in \mathcal{S}$ die δ-Funktion approximieren, dann gilt $f^{(a)\prime} \to \delta'$ (Beweis: Ü).

Kontrolle unseres Leitprinzips:

$$T'_g(f) = -T_g(f') = -\int \overline{g(x)} f'(x)\,dx = \int \overline{g'(x)} f(x)\,dx = T_{g'}(f)\,.$$

Als weiteres Beispiel betrachte ich die durch die Funktion

$$x_+^\lambda = \begin{cases} x^\lambda & x > 0\,, \\ 0 & x < 0 \end{cases}$$

gegebene Distribution. Das Integral $\int_0^\infty x^\lambda f(x)\,dx$ ist für Testfunktionen aus $\mathcal{S}$ wohldefiniert, wenn $\lambda > -1$ ist; x_+^λ definiert dann eine reguläre Distribution. Doch Ableitung von x^λ kann einen Exponenten im Intervall $(-2,-1)$ ergeben; dann muß das Integral regularisiert werden. Sei zunächst $\lambda > -1$; ich forme das Integral um:

$$\int_0^\infty x^\lambda f(x)\,dx = \int_0^1 x^\lambda (f(x) - f(0))\,dx + \frac{f(0)}{\lambda+1} + \int_1^\infty x^\lambda f(x)\,dx\,. \tag{2.2}$$

In dieser Form ist auch $\lambda \in (-2,-1)$ zugelassen, d. h. ich *definiere* nun x_+^λ für diese λ-Werte durch die rechte Seite von (2.2). Man kann das Spiel fortsetzen, indem man im ersten Integral weitere Terme $xf'(0)$, $\frac{1}{2}x^2 f''(0)$, usw., subtrahiert; dann definiert man damit x_+^λ als Distribution für immer stärker negative λ mit Ausnahme der Punkte $-1, -2, \ldots$. (Nicht jedes Integral läßt sich sinnvoll regularisieren.) Sei nun wieder $\lambda > -1$, Ableitung von x_+^λ:

$$((x_+^\lambda)', f(x)) = -(x_+^\lambda, f'(x)) = -\lim_{\varepsilon\to 0}\int_\varepsilon^1 x^\lambda f'(x)\,dx - \int_1^\infty x^\lambda f'(x)\,dx\,,$$

partiell integriert, wobei im ersten Term $f(x) - f(0)$ als Stammfunktion gewählt wird:

$$\begin{aligned} \ldots \;=\; & \lim_{\varepsilon\to 0}\left(-x^\lambda (f(x) - f(0))\,\Big|_\varepsilon^1 + \lambda \int_\varepsilon^1 x^{\lambda-1}(f(x) - f(0))\,dx\right) \\ & - x^\lambda f(x)\,\Big|_1^\infty + \lambda \int_1^\infty x^{\lambda-1} f(x)\,dx\,. \end{aligned}$$

Der Limes läßt sich in jedem der beiden Terme problemlos ausführen, die zwei Terme mit $f(1)$ heben sich heraus, und der Vergleich mit (2.2) mit $\lambda - 1$ anstelle von λ liefert

$$((x_+^\lambda)', f(x)) = \lambda (x_+^{\lambda-1}, f(x))\,,$$

wie man es erwarten würde, wenn x_+^λ eine gewöhnliche Funktion wäre. Man hat hier die Freiheit, die man bei jeder Regularisierung hat, so ausgenutzt, daß keine zusätzlichen Terme bei der Ableitung auftreten. Dies ist eine Art von Regularisierung, bei der man die Distribution x_+^λ als eine in λ analytische Funktion betrachten kann, die allerdings an den Stellen $\lambda = -1, -2, \ldots$ Pole hat: für jedes $f \in \mathcal{S}$ ist $x_+^\lambda(f)$ eine in λ analytische Funktion mit Ausnahme der Pole bei $\lambda = -1, -2, \ldots$.

Weiteres Beispiel:

$$(\ln|x|)' = \mathrm{P}\frac{1}{x}\,.$$

Beweis: Ü.

Wenn man an die δ-Funktion denkt, ist es ziemlich klar, daß man im allgemeinen keine zwei Distributionen miteinander multiplizieren kann. (Eine Definition $\delta^2(f) := f(0)^2$ ist nicht erlaubt, das wäre kein *lineares* Funktional.) Aber man kann eine Distribution mit einer nicht zu stark anwachsenden Funktion multiplizieren:

Definition 3.2.3 *Sei $p \in \mathcal{C}^\infty$ und höchstens polynomial wachsend und $T \in \mathcal{S}'$. Dann ist die Distribution pT definiert durch*

$$pT(f) := T(\overline{p}f) \quad \forall f \in \mathcal{S}\,.$$

Das ergibt das Erwartete, wenn $T(f) = T_g(f) = \int \overline{g}(x)f(x)\,dx$ eine reguläre Distribution ist: $pT_g = T_{pg}$.

Einfaches Beispiel zur Rechnung mit Distributionen: die Gleichung

$$f(x)\sin x = 0$$

hat in der Menge der stetigen Funktionen nur die Lösung $f = 0$. Sie kann gemäß Definition 3.2.3 als Gleichung in $\mathcal{S}'$ gelesen werden, dann gibt es noch weitere Lösungen: $f = \sum_{k\in\mathbb{Z}} a_k\delta_{k\pi}$ mit beliebigen Koeffizienten a_k; denn für Testfunktionen $g \in \mathcal{S}$ ist

$$(f\sin x)(g) = f(g\sin x) = \sum_{k\in\mathbb{Z}} a_k\delta_{k\pi}(g\sin x) = \sum_{k\in\mathbb{Z}} a_k g(k\pi)\sin k\pi = 0\,.$$

Merke: $xf(x) = 0$ wird durch $f(x) = \text{const}\,\delta(x)$ gelöst.
Wir hätten auch $f(x) = c(x)\delta(x)$ schreiben können; doch von der Funktion $c(x)$ ist nur der Wert $c(0)$ wesentlich: $f(g) = \int c(x)\delta(x)g(x)\,dx = c(0)g(0) = c(0)\delta(g)$. Also **merke** auch: $c(x)\delta(x) = c(0)\delta(x)$.
Wir haben eine derartige Gleichung schon einmal benutzt: bei der Suche nach Lösungen der Wellengleichung fanden wir

$$\left(\frac{\omega^2}{c^2} - \vec{k}\cdot\vec{k}\right)\tilde{f}(\vec{k},\omega) = 0\,;$$

sie wird gelöst durch $\tilde{f}(\vec{k},\omega) = g(\vec{k})\delta(\omega^2/c^2 - \vec{k}\cdot\vec{k})$.[1] Später werden wir sehen, wie wir mit einer solchen Distribution im Integral für die Fourier-Rücktransformation umzugehen haben.

[1] Man setze $\omega^2/c^2 - \vec{k}\cdot\vec{k} = x$, schreibe $\tilde{f}$ als $\tilde{f}(\vec{k},\omega) = h(\vec{k},x)$ und beachte $x\delta = 0$.

Weil man Distributionen differenzieren und mit regulären Funktionen multiplizieren kann, ist die Anwendung von recht allgemeinen Differentialoperatoren auf Distributionen wohldefiniert. Ich kann also nach Lösungen einer Differentialgleichung im Raum der Distributionen fragen.
Einfaches Beispiel: Die Gleichung $xy' = 0$ hat in der Menge der differenzierbaren Funktionen nur die Lösungen $y = \text{const}$, als Distributionsgleichung gelesen: $y = a_1 + a_2\Theta(x)$; denn $\Theta'(x) = \delta(x)$, und $x\delta(x) = 0$.

Man kann einer Distribution zwar keinen Funktionswert an einer Stelle x zuordnen, aber trotzdem von lokalen Eigenschaften reden: man kann definieren, wann eine Distribution in einem Gebiet $= 0$ ist und ihren Träger definieren. Zur Erinnerung: der Träger einer Funktion f ist der Abschluß der Menge der Punkte, auf denen f nicht verschwindet, anders:

$$\operatorname{supp} f = \overline{\{x \in \mathbb{R};\ f(x) = 0\}^c}\,.$$

Mit A^c bezeichne ich das Komplement der Menge A, Überstreichen bedeutet Abschluß der Menge. Dementsprechend möchte man $\operatorname{supp}\delta = \{0\}$ haben.

Definition 3.2.4 *Eine Distribution T verschwindet im (offenen) Gebiet G, $T = 0$ in G, wenn $T(f) = 0$ für alle Testfunktionen f mit $\operatorname{supp} f \subset G$.*

Das ist natürlich konsistent mit der üblichen Sprechweise, wenn $T = T_g$ ist. Doch das Beispiel $T = T_g$ zeigt auch, daß man nicht davon reden kann, daß „T=0 am Punkt x“ ist: g kann f. ü. $\neq 0$ sein, seine Nullstellen spielen dann im Integral und folglich für T_g keine Rolle. Wenn die Bedingung von Definition 3.2.4 nicht erfüllt ist, sage ich, daß $T \neq 0$ ist in G. (Das heißt, daß es ein f mit Träger in G gibt, auf dem T nicht verschwindet, es heißt *nicht*, daß T nicht auf einer Untermenge von G verschwinden könnte!)

Einfaches Beispiel: $\delta = 0$ auf $\mathbb{R}\setminus\{0\}$; und $\delta \neq 0$ in G, wenn G den Nullpunkt enthält.

Definition 3.2.5 *(i) Wenn $T \neq 0$ in allen Umgebungen eines Punktes x_0, heißt x_0 ein „wesentlicher Punkt von T“.*
(ii) Der Träger von T ist der Abschluß der Menge der wesentlichen Punkte von T.

Daraus kann man folgern: wenn $\operatorname{supp} T \cap \operatorname{supp} f = \emptyset$, dann ist $T(f) = 0$; und das führt zu

Lemma 3.2.6 *Der Träger einer Distribution T ist gegeben durch*

$$\begin{aligned}\operatorname{supp} T &= \cap\{F \subset \mathbb{R};\ F \text{ abgeschlossen}, T = 0 \text{ in } F^c\}\\ &= \left(\cup\{G \subset \mathbb{R};\ G \text{ offen}, T = 0 \text{ in } G\}\right)^c\,.\end{aligned}$$

Ich will den Beweis nicht vollständig führen.
Die zweite Gleichheit ist einfache Mengentheorie: $(\cup G)^c = \cap G^c$, und G^c ist eines der F's.
Wenn $T = 0$ in G, gilt offensichtlich $G \subset (\operatorname{supp} T)^c$, also $G^c \supset \operatorname{supp} T$; das zeigt, daß $\operatorname{supp} T \subset (\cup G)^c$. Um die Gleichheit zu zeigen, überlegt man sich zuerst, daß man für jedes Paar von offenen Mengen V und U mit $\overline{V} \subset U$ ein nicht-negatives $\varphi \in \mathcal{D}$ findet, das $= 1$ auf $\overline{V}$ und $= 0$ auf U^c ist. (Man suche erst ein stetiges φ und wende die Methode von Lemma 1.2.17 an.) Damit kann man eine „Zerlegung der Einheit" finden: man sucht eine lokal endliche Überdeckung von $\mathbb{R}$ (oder G) durch abzählbar viele offene Mengen U_i („lokal endlich": jedes x liegt in nur endlich vielen U_i) und zu jedem U_i ein $\varphi_i \in \mathcal{D}$, das auf U_i^c verschwindet, so, daß $\sum_i \varphi_i(x) = 1$ ist. Damit kann man f in eine Summe von Termen zerlegen, die ihre Täger in kleinen Umgebungen haben, und die lokale Definition von „wesentlicher Punkt" nutzen, um die obige Folgerung zu ziehen. Durchführung: Ü.

Damit ist es klar, daß $\operatorname{supp} \delta = \{0\}$. Ebenso: $\operatorname{supp} \delta' = \{0\}$, $\operatorname{supp} \delta^{(n)} = \{0\}$. Das gilt auch allgemeiner:

$$\operatorname{supp} T' = \operatorname{supp} T\,. \tag{2.7}$$

Denn man beachte, daß dort, wo f in einer Umgebung $= 0$ ist, auch $f' = 0$ gilt, also $\operatorname{supp} f' \subset \operatorname{supp} f$; somit ist $T'(f) = -T(f) = 0$ in G, wenn $T = 0$ in G. Die Gleichung folgt dann aus dem obigen Lemma.

Auch klar: $\operatorname{supp} T_g = \operatorname{supp} g$. Es gilt auch eine Umkehrung der ersten Aussage über den Träger der $\delta^{(n)}$:

Satz 3.2.8 *Für die Distribution T gelte $\operatorname{supp} T = \{0\}$, dann ist*

$$T = \sum_{\text{endlich}} c_n \delta^{(n)}\,.$$

Wenn außerdem $T \geq 0$ ist, folgt $T = \text{const}\,\delta$ mit positiver Konstante.

$T \geq 0$ heißt: $T(f) \geq 0$ für alle nicht-negativen f. Es ist klar, daß $\delta \geq 0$, aber $\delta^{(n)} \not\geq 0$, wenn $n > 0$, denn $\delta^{(n)}(f) = (-1)^n f^{(n)}(0)$, und die Ableitungen einer nicht-negativen Funktion können beliebiges Vorzeichen bei 0 haben.
Man beachte: Wenn f sich um $x = 0$ in eine Taylor-Reihe entwickeln läßt: $f(x) = \sum f^{(n)}(0)x^n/n!$, kann ich $f(x) = \delta_x(f) = \sum((-x)^n/n!)\delta^{(n)}(f)$ schreiben, d.h. $\delta_x = \sum((-x)^n/n!)\delta^{(n)}$; das ist keine endliche Summe, und deshalb kann ich die Distribution δ_x mit $\operatorname{supp} \delta_x = \{x\}$ darstellen durch Distributionen, die alle den Träger $\{0\}$ haben.

Der obige Satz wird des öfteren gebraucht, deshalb erwähne ich ihn; aber ich will ihn nicht in allen Einzelheiten beweisen. Man benutzt dazu den folgenden Satz, der die Struktur der temperierten Distributionen beschreibt:

Satz 3.2.9 *Jede Distribution $T \in \mathcal{S}'(\mathbb{R})$ läßt sich darstellen als*

$$T(f) = \int t(x) D^p f(x)\, dx \tag{2.10}$$

mit einer geeigneten stetigen Funktion t, die höchstens polynomial wächst und einem passenden Differentialoperator D^p der Ordnung p.

Das sieht sehr unerwartet aus; doch man beachte: $t(x)$ muß nur stetig, keineswegs differenzierbar sein, so daß man nur formal partiell integrieren kann. Diesen Satz will ich nicht beweisen. Zur Illustration: $\delta = \Theta' = t''$, wobei $t(x) = \int_0^x \Theta(y)\, dy$ eine stetige Funktion ist, $t = 0$ auf der negativen Halbachse, $= x$ für $x \geq 0$; also

$$\delta(f) = \int t(x) f''(x)\, dx = \int_0^\infty x f''(x)\, dx = -\int_0^\infty f'(x)\, dx = f(0)\,.$$

Jetzt kann ich den Beweis von Satz 3.2.8 skizzieren. Wir schreiben T in der Form (2.10). Weil $\operatorname{supp} T = \{0\}$, genügt es, das Integral über eine kleine Umgebung U_ε der 0 zu erstrecken, anders: $T(f) = T(h_\varepsilon f)$, wobei $h_\varepsilon \in \mathcal{S}$, $h_\varepsilon = 0$ für $|x| > \varepsilon$ und $= 1$ für $|x| < \varepsilon/2$. Man kann $f \mapsto \int_{U_\varepsilon} t(x) D^p f(x)\, dx$ auffassen als stetiges Funktional über dem Raum der p-mal stetig differenzierbaren Funktionen mit Träger in $[-1, 1]$, wobei man sich für den Konvergenzbegriff auf die Normen p_{mn} mit $n \leq p$ beschränkt; d. h. das Funktional T wird erweitert auf diese größere Menge von Testfunktionen. Es gilt $T(f) = 0$ für alle f, deren erste p Ableitungen in einer Umgebung von 0 verschwinden, und wegen der Stetigkeit der Erweiterung auch auf allen f, deren erste Ableitungen *im Punkt* $x = 0$ verschwinden. f werde Taylor-entwickelt bis zur p-ten Ordnung:

$$f(x) = \sum_{n \leq p} \frac{x^n}{n!} f^{(n)}(0) + f_R(x)\,,$$

wobei die ersten p Ableitungen von f_R bei 0 verschwinden; also gilt $T(f_R) = 0$. Folglich

$$T(f) = T(h_1 f) = \sum_{n \leq p} T(h_1 x^n / n!) f^{(n)}(0)\,,$$

also mit $c_n = (-1)^n T(h_1 x^n / n!)$:

$$T = \sum_{n \leq p} c_n \delta^{(n)}\,.$$

Die Form (2.10) macht auch die folgende Aussage plausibel:

Lemma 3.2.11 *Zu jeder Distribution $T \in \mathcal{S}'$ gibt es eine Konstante C und eine Zahl N, so daß*

$$|T(f)| \leq C q_N(f) \equiv C \sup_{m,n \leq N} p_{mn}(f) \quad \forall f \in \mathcal{S}\,.$$

Das hier definierte q_N ist ebenfalls eine Halbnorm; diese q_N's bilden eine aufsteigende Folge, die offensichtlich den gleichen Konvergenzbegriff definieren wie $\{p_{mn}\}$.

Beweis: T als lineares stetiges Funktional ist insbesondere stetig bei $0 \in \mathcal{S}$, zu $\varepsilon = 1$ gibt es eine $\tau_{\mathcal{S}}$-Umgebung U der 0, so daß $|T(f)| < 1$ für $f \in U$. Die Konvergenz ist durch die Halbnormen q_N bestimmt, d. h. es gibt eine durch ein q_N und ein ε' bestimmte Umgebung $U' = \{f;\, q_N(f) < \varepsilon'\}$, die in U enthalten ist. Auch auf ihr ist T durch 1 beschränkt. Für beliebiges $f \in \mathcal{S}$ ist $(\varepsilon'/2q_N(f))f \in U'$, also

$$\frac{\varepsilon'}{2q_N(f)}|T(f)| < 1\,,$$

und daraus folgt die Behauptung. ■

Zum Schluß will ich noch einmal auf die Regularisierung von Integralen zurückkommen: sie ist nicht eindeutig. Man kann zu einem Integral, das man bei $x = 0$ regularisiert hat, ja immer eine Distribution mit Träger $\{0\}$ addieren, ohne den Wert des Integrals für Funktionen zu ändern, die in einer Umgebung der Null $= 0$ sind, für die also auch alle Ableitungen bei 0 verschwinden. Wenn man ein Integral hat, das für alle Funktionen existiert, die eine einfache Nullstelle bei 0 haben (wie $\int (f(x)/x)\,dx$), dann kann man fordern, daß die Regularisierung auch für alle diese Funktionen dasselbe ergibt, und dann hat man nur noch eine Distribution $\mathrm{const}\,\delta$ frei.

Alles in den letzten beiden Abschnitten Gesagte läßt sich mit meist trivialen Änderungen auf Distributionen $\in \mathcal{S}(\mathbb{R}^n)$ übertragen; Beispiel Ableitung:

$$(\partial_{x_i}T)(f) = -T(\partial_{x_i}f)$$

u. s. w.

3.3 Die Faltung von Distributionen

Man kann zwar Distributionen nicht direkt multiplizieren, doch gibt es zwei andere Multiplikationsoperationen: das direkte Produkt und die Faltung. (Die Faltung zweier Distributionen ist allerdings nicht immer definiert.)

Ich werde im Folgenden häufig die Schreibweise $T(f) \equiv (T(x), f(x))$ benutzen, um anzudeuten, von welcher Variablen die Funktion abhängt, die als Argument der Distribution gemeint ist.

Seien $T_i \in \mathcal{S}(\mathbb{R})'$, dann kann man natürlich

$$T_1 \times T_2(f_1 \times f_2) = (T_1(x) \times T_2(y), f_1(x)f_2(y)) = T_1(f_1)T_2(f_2)$$

definieren, wenn $f_i \in \mathcal{S}(\mathbb{R})$. Das kann man erweitern auf beliebige Funktionen $f(x,y) \in \mathcal{S}(\mathbb{R}^2)$. Das will ich jetzt zeigen.

Lemma 3.3.1 *Sei $f(x,y) \in \mathcal{S}(\mathbb{R}^2)$ und $T_2 \in \mathcal{S}(\mathbb{R})'$. Man betrachte $f_x(y) := f(x,y)$ als Funktion von y, dann ist $f_x \in \mathcal{S}(\mathbb{R})$ und $g(x) := T_2(f_x)$ ist wohldefiniert; und es gilt $g \in \mathcal{S}(\mathbb{R})$.*

Beweis: Die erste Aussage, $f_x \in \mathcal{S}$, ist leicht zu sehen. Die Differenzierbarkeit von g folgt, weil T_2 linear und stetig ist, und weil $\lim(f(x+\Delta x,y)-f(x,y))/\Delta x = \partial_x f(x,y) \in \mathcal{S}(\mathbb{R})$ auch in der $\mathcal{S}$-Topologie konvergiert:

$$\lim \frac{1}{\Delta x}(T_2(f_{x+\Delta x})-T_2(f_x)) = \Big(T_2(y), \lim \frac{1}{\Delta x}(f(x+\Delta x,y) - f(x,y))\Big) = T_2(\partial_x f(x,y))\,.$$

Entsprechend für höhere Ableitungen. Abfalleigenschaften: wir benutzen Lemma 3.2.11,

$$\begin{aligned} p_{kl}(T_2(f_x)) &= \sup_x \left|x^k \partial_x^l T_2(f_x)\right| = \sup_x \left|x^k T_2(\partial_x^l f_x)\right| \\ &\le C \sup_x \left|x^k q_N(\partial_x^l f_x)\right| \\ &\le C \sum_{m,n\le N} c_{mn} \sup_{x,y} \left|x^k \partial_x^l y^m \partial_y^n f(x,y)\right| \\ &= \sum \gamma_{rs} p_{rs}(f) < \infty\,, \end{aligned}$$

weil die Summation nur über endlich viele Doppelindizes $r = (k,m)$ und $s = (l,n)$ geht. ■

Wir können also auf $T_2(f_x)$ die Distribution T_1 anwenden und damit das **direkte Produkt** definieren:

$$\Big(\,T_1(x) \times T_2(y), f(x,y)\,\Big) := \Big(\,T_1(x), \Big(\,T_2(y), f(x,y)\,\Big)\,\Big)\,.$$

Wir hätten statt mit T_2 auch mit T_1 beginnen können: $f_y(x) := f(x,y)$, und weil $T_1(f_y) \in \mathcal{S}(\mathbb{R})$, ist $T_2(T_1(f_y))$ wohldefiniert. Daß das Ergebnis dasselbe ist wie vorher, sieht man am schnellsten mit Satz 3.2.9:

$$T_1(T_2(f_x)) = \int dx \int dy\, t_1(x) D_x^p\, t_2(y) D_y^q\, f(x,y) = T_2(T_1(f_y))\,,$$

wobei man für die zweite Gleichheit die beiden Integrationen vertauschen muß, was aber problemlos möglich ist, weil $f \in \mathcal{S}(\mathbb{R}^2)$.

Wichtigstes Beispiel: $\delta^{[3]} = \delta \times \delta \times \delta \in \mathcal{S}'(\mathbb{R}^3)$, in formaler Schreibweise:

$$\delta^{[3]}(\vec{x}) = \delta(x)\delta(y)\delta(z)\,;$$

Illustration mit $\delta^{[2]}$: $g(x) = \delta(f_x) = f_x(0) = f(x,0)$, und damit $\delta^{[2]}(f) = \delta(g) = f(0,0)$.
(Den oberen Index, der die Dimension des Grundraums angibt,schreibt man meist in runden Klammern, ich wollte Verwechslungen mit der Ableitung vermeiden. Ich werde im Folgenden den Index weglassen, er ergibt sich meist aus dem Zusammenhang, besonders, wenn man die formale Schreibweise benutzt.)

Wenn wir $f(x,y) = h(x+y)$ mit $h \in \mathcal{S}(\mathbb{R})$ wählen, geht die obige Konstruktion im allgemeinen schief: f_x ist dann die Translation von h um x, ich benutze die Notation $f_x(y) = h(x+y) \equiv (\tau_x h)(y)$; man überlegt sich wie oben, daß $T_2(f_x) = (T_2(y), h(x+y))$ eine beliebig oft differenzierbare Funktion von x ist, doch im allgemeinen fällt sie nicht ab, wenn $|x| \to \infty$. Zwei sehr verschiedenartige Beispiele: $T_2 = \delta$ und $T_2 = T_g$ mit $g(y) = y^k$:
$(\delta(y), h(x+y)) = h(x)$ ist $\in \mathcal{S}$; $(T_g(y), h(x+y)) = \int y^k h(x+y)\, dy = \int (\xi - x)^k h(\xi)\, d\xi$ ist $\in \mathcal{C}^\infty$, aber wächst mit x polynomial an!

Es gibt also eine Klasse $\mathcal{T}_0 \subset \mathcal{S}'$ von Distributionen T_0, für die $T_0(\tau_x f) \in \mathcal{S}$ ist, falls $f \in \mathcal{S}$. Wenn der zweite Faktor des direkten Produkts aus $\mathcal{T}_0$ ist, können wir es auf $f(x+y)$ anwenden: $(T_1(x) \times T_0(y), f(x+y))$ ist wohldefiniert. Wenn der erste Faktor $\in \mathcal{T}_0$ ist, können wir mit dem ersten Faktor beginnen; und da f symmetrisch in x und y ist, kommt es auf die Reihenfolge der Faktoren nicht an. Dann definieren wir

Definition 3.3.2 *Die Faltung $T_1 * T_2$ zweier Distributionen $T_1, T_2 \in \mathcal{S}(\mathbb{R})'$ ist wohldefiniert, wenn einer der Faktoren in $\mathcal{T}_0$ ist:*

$$T_1 * T_2(f) := T_1(T_2(\tau_x f)) = \Big(\, T_1(x) \times T_2(y), f(x+y) \,\Big)\,, \quad f \in \mathcal{S}(\mathbb{R})\,.$$

Man beachte: $\quad T_1 * T_2 = T_2 * T_1$.
Ich kenne keine vollständige Charakterisierung der Menge $\mathcal{T}_0$, eine hinreichende Bedingung an T_0 ist jedoch leicht anzugeben:

Lemma 3.3.3 *Jede Distribution mit kompaktem Träger ist in $\mathcal{T}_0$.*

Beweis: soll nur für Testfunktionen aus $\mathcal{D}$ gegeben werden. Es genügt, nachzuweisen, daß $T_0(\tau_x f)$ kompakten Träger hat, die Differenzierbarkeit ist klar. Wenn $f \in \mathcal{D}$, dann ist $\operatorname{supp} \tau_x f = \operatorname{supp} f + x$, d. h. gleich dem um x verschobenen Träger; es ist folglich klar, daß $\operatorname{supp} T_0 \cap \operatorname{supp} \tau_x f = \emptyset$ für hinreichend großes x, d. h. $T_0(\tau_x f) = 0$, wenn x groß ist. Plausibel: wenn $f \in \mathcal{S}$, dann fällt $T_0(\tau_x f)$ stark ab. ■

Betrachten wir die Faltung für den Spezialfall, daß T_i gegeben sind durch höchstens polynomial wachsende Funktionen g_i, wobei $\operatorname{supp} g_2$ kompakt sein möge:

$$\begin{aligned}(T_{g_1} * T_{g_2})(f) &= \int dx\, \overline{g_1(x)} \int dy\, \overline{g_2(y)} f(x+y) \\ &= \int dx' \overline{\int dx\, g_1(x) g_2(x'-x)} f(x') \\ &= T_h(f)\,,\end{aligned}$$

wobei h die Faltung der *Funktionen* g_1 und g_2 ist: $h = g_1 * g_2$ (vergl. Gleichung (II.2.16)). Auch hier setzt also die Begriffsbildung für die Distributionen diejenige für Funktionen fort.

Auch wenn wir die formale Schreibweise benutzen, bekommen wir die zu (II.2.16) analoge Gleichung:

$$\int (T_1 * T_2)(x) f(x)\, dx = \int T_1(x) T_2(y) f(x+y)\, dy dx = \int T_1(x'-x) T_2(x) f(x')\, dx' dx\,,$$

und mit Vertauschung der Integrationen folgt:

$$(T_1 * T_2)(x) = \int T_1(x-y) T_2(y)\, dy\,.$$

Diese Vertauschung der Integrationen ist eine formale Operation; die resultierende Gleichung ist im distributiven Sinn zu verstehen: man multipliziere mit einer Testfunktion und integriere, dabei vertausche man die Integrationen so, daß die entstehenden Ausdrücke sinnvoll sind. Man geht also mit der formalen Vertauschung wieder rückwärts, im vorliegenden Fall erhält man unsere Ausgangsgleichung. Trotzdem: formale Gleichungen sind nützlich als Merk- und Rechenhilfe.

Einfaches, aber wichtiges Beispiel:

$$T * \delta = T\,, \tag{3.4}$$

denn $\delta(\tau_x f) = (\delta(y), f(x+y)) = f(x)$, also $(T * \delta)(f) = T(f)$. Im Raum der Distributionen gibt es also ein Eins-Element bezüglich der Faltung, im Funktionenraum existiert es nicht: $e * f = f\ \forall f$ ist nicht möglich mit einer gewöhnlichen Funktion e.

Wir können formal eine Distribution mit einer Funktion falten:

$$h(x) \equiv (T * g)(x) = \int T(x')g(x - x')\,dx' = \int T(x')(\tau_x g)(-x')\,dx'\,; \tag{3.5}$$

ich will zeigen, daß h für eine große Klasse von g´s tatsächlich eine Funktion ist.

Lemma 3.3.6 *Sei $g \in \mathcal{S}(\mathbb{R})$, dann ist $T(\tau_x g) \in \mathcal{C}^\infty(\mathbb{R})$ und wächst höchstens polynomial an.*

Beweis: Die Differenzierbarkeit von $T(\tau_x g)$ haben wir schon früher gezeigt, wir müssen die Wachstumseigenschaften prüfen. Ich zeige zunächst: wenn $f \in \mathcal{S}(\mathbb{R})$, dann ist $h(x,y) := f(x)g(x+y) \in \mathcal{S}(\mathbb{R}^2)$. Daß $h \in \mathcal{C}^\infty(\mathbb{R}^2)$, ist klar. Weiter:

$$\sup_{x,y} \left| x^k y^l \partial_x^m \partial_y^n f(x) g(x+y) \right| \leq \mathrm{const}\, q_N(f) q_M(g) < \infty\,,$$

wobei $N = \max(k,m),\ M = \max(l, m+n)$.
Betrachten wir nun $f(x)\,T(\tau_x g)$; wegen der Linearität von T und mit Lemma 3.3.1 folgt

$$f(x)\,T(\tau_x g) = (T(y), f(x)g(x+y)) = T(h_x) \in \mathcal{S}(\mathbb{R})$$

mit $h_x(y) = h(x,y) = f(x)g(x+y)$. Wenn aber $T(\tau_x g)$, mit einer stark abfallenden Funktion multipliziert, in $\mathcal{S}$ liegt, kann es nicht stärker als polynomial anwachsen. ■

Ich will die formale Schreibweise $T * g$ wie in Gleichung (3.5) benutzen:

$$h(x) = (T * g)(x) = (T(x'), g(x - x')) \equiv T((\tau_x g) \circ P) \tag{3.7}$$

mit der **Notation:** „Paritätsoperator" P: $(f \circ P)(x) = f(-x)$.

Jedes $g \in \mathcal{S}$ kann mit polynomial wachsenden Funktionen p multipliziert und integriert werden, d. h. $T_{pg} \in \mathcal{S}'$; es kann mit p gefaltet werden, und $p * g$ wächst höchstens polynomial. Mit stärker wachsenden Funktionen ist das nicht für alle $g \in \mathcal{S}$ möglich. Analog: man kann $T \in \mathcal{S}'$ auf alle $g \in \mathcal{S}$ anwenden und nach Lemma 3.3.6 mit g falten, wobei eine höchstens polynomial wachsende Funktion entsteht. Angesichts dessen kann man lax sagen: „Eine temperierte Distribution wächst höchstens polynomial an".

Nun zeigt man leicht, daß für $g \in \mathcal{D}$ (so daß $T_g \in \mathcal{T}_0$ ist) gilt

$$T * T_g = T_h \quad \text{mit } \overline{h(x)} = T((\tau_x \overline{g}) \circ P)\,. \tag{3.8}$$

Das entspricht also, abgesehen von der komplexen Konjugation, der Faltung $T * g$, vergl. (3.7). Beweis: Ü.

Jetzt haben wir das Werkzeug, um zu beweisen, daß man jede Distribution durch glatte Funktionen approximieren kann, wie wir es für δ gemacht haben (siehe Lemma 2.2.7):

Wir wählen $g \in \mathcal{D}$ so, daß $g \geq 0$ und $\int g\,dx = 1$, dann definiert $g_{(a)}(x) := (1/a)g(x/a)$ eine Folge von Funktionen $\in \mathcal{D}$, die gegen δ konvergieren, $\int g_{(a)}(x)f(x)\,dx \longrightarrow f(0)$. Genauer: für

$$\delta_{(a)} := T_{g_{(a)}} \qquad \text{d. h.} \quad \delta_{(a)}(f) := \int g_{(a)}(x)f(x)\,dx$$

gilt

$$\lim_{a\to 0} \delta_{(a)} = \delta\,,$$

und weiterhin:

$$\lim_{a\to 0} T * \delta_{(a)} = T\,. \tag{3.9}$$

Beweis für die letzte Aussage:

$$(T * \delta_{(a)})(f) = (\delta_{(a)} * T)(f) = \delta_{(a)}(T(\tau_x f)) = \int g_{(a)}(x)\,T(\tau_x f)\,dx \longrightarrow T(f)\,.$$

(Allgemein gilt, wie man sich leicht überlegt, daß $T_1 * T_2$ stetig in jedem Faktor ist.)

Satz 3.3.10 *Jede Distribution aus $\mathcal{S}'$ läßt sich durch reguläre Distributionen T_h mit $h \in \mathcal{D}$ approximieren (kurz: „läßt sich durch Funktionen aus $\mathcal{D}$ approximieren“).*

Beweis: Die Gleichungen (3.9) und (3.8) zeigen, daß wir T durch T_h mit glattem, höchstens polynomial wachsendem h approximieren können; wenn wir h mit glatten Funktionen $f_R \in \mathcal{D}$ abschneiden, die $= 1$ sind für $|x| < R$ (vergl. Gleichung (I.2.21)), dann wird T_h durch die $T_{f_R h}$ approximiert, und $f_R h \in \mathcal{D}$. ■

Für die Anwendung der Distributionen zur Lösung von Differentialgleichungen brauche ich noch die Wirkung eines Differentialoperators auf die Faltung. Sei $D = \sum_{k=0}^{n} a_k(x)\partial_x^k$. Die Funktionen a_k können komplexwertig sein. Der dazu „adjungierte“ Operator D^* ist gegeben durch

$$D^* = \sum_{k=0}^{n} (-1)^k \partial_x^k \overline{a_k(x)}\,. \tag{3.11}$$

Man beachte: die Ableitungen wirken auf alles, was dahinter steht, $\partial_x a(x)$ angewandt auf $f(x)$ heißt: $\partial_x(a(x)f(x))$. Grund für diese Form: es gilt für geeignete Funktionen (z. B. für $f, g \in \mathcal{S}$)

$$\int \overline{f(x)} Dg(x)\,dx = \int \overline{(D^* f(x))} g(x)\,dx\,. \tag{3.12}$$

Das folgt leicht durch mehrfache partielle Integration. Und ebenso sieht man

$$(DT)(f) = T(D^* f)\,. \tag{3.13}$$

Der Begriff des adjungierten Operators wird für uns später wichtig, wenn wir über Operatoren im Hilbertraum sprechen.

Lemma 3.3.14 *Sei D ein Differentialperator mit konstanten Koeffizienten. Dann gilt*

$$D(T_1 * T_2) = (DT_1) * T_2 = T_1 * (DT_2)\,.$$

Beweis: Wir nutzen die Definition der Faltung und Gleichung (3.13):

$$D(T_1 * T_2)(f) = (T_1 * T_2)(D^* f) = T_1(T_2(\tau_x D^* f))\,;$$

und da D^* sich auf die Variable von f bezieht, kann ich es mit τ_x vertauschen (die konstanten Koeffizienten werden von τ_x ja nicht berührt): $(\tau_x D^* f)(y) = D^*_y f(x+y)$, also

$$\ldots = T_1(T_2(D^* \tau_x f)) = T_1((DT_2)(\tau_x f)) = (T_1 * DT_2)(f)\,.$$

Die andere Form folgt wegen der Symmetrie der Faltung. ■

Man kann hier auch leicht einen formalen Beweis mit der formalen Schreibweise bringen (ich schreibe D_x, wenn ich andeuten will, daß D auf die Variable x wirken soll):

$$\begin{aligned} D(T_1 * T_2)(x) &= D\int T_1(x-y)T_2(y)\,dy = \int D_x T_1(x-y)T_2(y)\,dy \\ &= \int (DT_1)(x-y)T_2(y)\,dy = (DT_1) * T_2\,. \end{aligned}$$

Jetzt können wir die Methode der Greenschen Funktionen genau darstellen, die wir im Beipiel der Laplace-Gleichung für das Potential einer Ladungsverteilung in der Einleitung zu diesem Kapitel rein formal benutzt haben.

Satz 3.3.15 *Sei D ein Differentialoperator mit konstanten Koeffizienten in $\mathbb{R}^n$, und G sei Lösung der Gleichung*

$$DG = \delta\,.$$

Dann löst $f = G * g$ die Gleichung

$$Df = g\,.$$

δ meint hier natürlich die n-dimensionale Delta-Distribution. G heißt Greensche Funktion zum Differentialoperator D. f ist nach Gleichung (3.5) eine gewöhnliche Funktion.

Beweis: Wir benutzen Lemma 3.3.14 und Gleichung (3.4):

$$Df = D(G * g) = (DG) * g = \delta * g = g\,.$$

■

Wir haben gesehen, daß die Faltung einer Distribution mit einer Funktion eine gewöhnliche Funktion ergeben kann, wenn die Funktion, mit der gefaltet wird, hinreichend brav ist. Das sei hier vorausgesetzt; dann ist $G * g$ eine Lösungs*funktion.*

Die Greensche Funktion zu einem Differentialoperator ist nicht eindeutig; man kann ja immer eine Lösung der homogenen Differentialgleichung hinzuaddieren und erhält eine andere Greensche Funktion. Man versucht sie so zu wählen, daß sie die Randbedingungen für das vorliegende Problem erfüllt; dann erfüllt auch $G * g$ die Randbedingungen.

3.4 Die Fourier-Transformation von Distributionen

Wie bei der Differentiation definiert man die Fourier-Transformation auch durch Überwälzen der Operation auf die Testfunktionen. Dazu muß man natürlich die Wirkung der Fourier-Transformation auf die Testfunktionen kennen; im allgemeinen liegen sie in einem anderen Raum, so daß sich durch Fourier-Transformationauch eine andere Menge von Distributionen ergibt. Bekannt aus Theorem 2.2.9: $\mathcal{F}\mathcal{S} = \mathcal{S}$. Da $\mathcal{D} \subset \mathcal{S}$, ist auch $\mathcal{Z} := \mathcal{F}\mathcal{D} \subset \mathcal{F}\mathcal{S} = \mathcal{S}$; der folgende Satz liefert eine genaue Charakterisierung:

Satz 3.4.1 *Die Fourier-Transformierten von Testfunktionen aus $\mathcal{D}$ sind ganze analytische Funktionen, die bezüglich des Imaginärteils des Arguments exponentiell beschränkt sind, genauer:*

$$\mathcal{F}\mathcal{D} \equiv \mathcal{Z} :=$$
$$\left\{ \tilde{f} \in \mathcal{S};\ \tilde{f}(k) \textit{ ist analytisch in } k \in \mathbb{C} \textit{ und erfüllt die Bedingung } (*)\ \forall\, n = 0, 1, 2, \ldots \right\} :$$

$$\exists\, a > 0, C_n > 0 : \quad |k|^n |\tilde{f}(k)| \leq C_n e^{a|\Im k|} \tag{$*$}$$

Beweis: Sei $f \in \mathcal{D}$, dann verschwindet f außerhalb eines geeigneten Intervalls $[-a, a]$, und die Fourier-Transformation lautet

$$\tilde{f}(k) = \sqrt{2\pi}^{-1} \int_{-a}^{a} f(x) e^{-ikx}\, dx\,.$$

In diesem Ausdruck können wir k als komplexe Variable betrachten; wegen der endlichen Integrationsgrenzen existiert das Integral für *alle* $k \in \mathbb{C}$. Ableitung nach k ergibt einen zusätzlichen Faktor $-ix$ im Integranden, das ändert nichts an der Existenz des Integrals, d. h. die Funktion $\tilde{f}(k)$ ist eine ganze analytische Funktion.
Weil $\mathcal{F}(f^{(n)}) = (ik)^n \mathcal{F}(f)$ ist, folgt

$$|k|^n |\tilde{f}(k)| = \sqrt{2\pi}^{-1} \left| \int_{-a}^{a} f^{(n)}(x) e^{-i(\Re k + i\Im k)x}\, dx \right| \leq C_n e^{a|\Im k|}\,,$$

und das ist die Bedingung $(*)$.
Wir haben noch die Umkehrung zu zeigen: sei $\tilde{f}(k)$ eine ganze Funktion, die $(*)$ erfüllt. Dann definieren wir

$$f(x) := \sqrt{2\pi}^{-1} \int_{-\infty}^{\infty} \tilde{f}(k') e^{ik'x}\, dk'\,.$$

Die Bedingung $(*)$ garantiert, daß $\tilde{f}$ mit $\Re k$ stärker als jede Potenz abfällt; wir können also im Integranden einen Faktor $(ik)^n$ einfügen, und es ergibt sich die n-te Ableitung von f, d. h. $f(x)$ ist beliebig oft differenzierbar.

Jetzt bleibt noch zu zeigen, daß der Träger von f beschränkt ist.
Der Integrand ist analytisch in k', wir können den Integrationsweg auf eine zur reellen Achse parallele Gerade verschieben:

$$f(x) = \sqrt{2\pi}^{-1} \int_{-\infty}^{\infty} \tilde{f}(k' + i\kappa) e^{i(k'+i\kappa)x} \, dk' = \sqrt{2\pi}^{-1} e^{-\kappa x} \int_{-\infty}^{\infty} \tilde{f}(k' + i\kappa) e^{ik'x} \, dk' \,, \quad (4.2)$$

wobei das Integral wegen der Bedingung $(*)$ weiterhin konvergiert (und die unendlich fernen Stücke nichts beitragen). Wenden wir $(*)$ für $n = 0$ und $n = 2$ an, sehen wir, daß

$$|\tilde{f}(k)| \le e^{a|\Im k|} \min\{C_0, C_2/|k|^2\} \le \frac{Ce^{a|\Im k|}}{1 + |k|^2} \le \frac{Ce^{a|\Im k|}}{1 + (\Re k)^2}$$

ist (zur zweiten Abschätzung: solange $|k| < 1$, kann ich $C = 2(C_0 + C_2)$ wählen, wenn $|k| > 1$, ist der Korrekturfaktor $(1 + |k|^2)/|k|^2$ durch 2 beschränkt, die Wahl $C = 2(C_0 + C_2)$ ist also auch dann gut. Eingesetzt in (4.2):

$$|f(x)| \le \sqrt{2\pi}^{-1} e^{-\kappa x} \int_{-\infty}^{\infty} \frac{Ce^{a|\kappa|}}{1 + k'^2} \, dk' = C' e^{a|\kappa| - \kappa x} \,.$$

Sei nun $|x| > a$ und $c > 0$ beliebig. Ich wähle $\kappa = c|x|/x$, dann folgt

$$|f(x)| \le C' e^{(a-|x|)c} \,;$$

damit das für beliebige $c > 0$ gilt (wobei C' nicht von c abhängt!), muß $f(x) = 0$ sein, denn der Exponent ist nach Voraussetzung negativ. Daher ist $f \in \mathcal{D}$. Daß die Fourier-Transformation von f wieder die Ausgangsfunktion $\tilde{f}$ ergibt, ist klar, weil $\tilde{f}$ stetig ist. ■

Definition 3.4.3 *Die Fourier-Transformierte von $T \in \mathcal{S}'$ ist gegeben durch*

$$(\mathcal{F}T)(\tilde{f}) = T(\mathcal{F}^{-1}\tilde{f}) \,, \quad \tilde{f} \in \mathcal{S} \,.$$

Anders geschrieben: $\tilde{T}(\tilde{f}) = T(f)$. Wenn $T = T_g$ regulär ist, gegeben durch eine Fourier-transformierbare Funktion g, gilt

$$\mathcal{F}T_g = T_{\mathcal{F}g} \,;$$

das folgt mit der Parsevalschen Gleichung:

$$(\mathcal{F}T_g)(\mathcal{F}f) = T_g(f) = \int \overline{g(x)} f(x) \, dx = \int \overline{\tilde{g}(k)} \tilde{f}(k) \, dk = T_{\mathcal{F}g}(\mathcal{F}f) \,.$$

Beispiel:

$$\mathcal{F}\delta = \frac{1}{\sqrt{2\pi}} \,. \quad (4.4)$$

Denn $\delta(f) = f(0) = \sqrt{2\pi}^{-1}\int \tilde{f}(k)\,dk = (T_{\sqrt{2\pi}^{-1}}, \tilde{f}(k))$. Das entspricht genau dem, was man erwartet, wenn man die formale Schreibweise benutzt: $\sqrt{2\pi}^{-1}\int \delta(x)e^{-ikx}dx = \sqrt{2\pi}^{-1}e^{-ikx}\Big|_{x=0} = \sqrt{2\pi}^{-1}$. Umgekehrt ergibt die Fourier-Transformation (bzw. die inverse Fourier-Transformation) einer Konstante die δ-Distribution bis auf eine Konstante:

$$\int e^{ikx}\,dk = 2\pi\delta(x)\,. \tag{4.5}$$

(Weil δ reell ist, gilt natürlich auch $\int e^{-ikx}\,dk = 2\pi\delta(x)$.) Das ist die formale Schreibweise der invers Fourier-transformierten Gleichung (4.4). Hier sehen Sie auch leicht, wie man die formale Gleichung streng beweist: man wendet sie auf eine Testfunktion an:

$$\begin{aligned}\int dx \int dk\, e^{ikx} f(x) &= \int dk \int dx\, e^{ikx} f(x) = (2\pi)^{1/2}\int dk\,\tilde{f}(-k)\\ &= 2\pi f(0) = 2\pi\int \delta(x)f(x)\,dx\,.\end{aligned}$$

Mit der Vertauschung der Integrationen erhält man einen wohldefinierten Ausdruck, und so ist das linksstehende formale Integral zu interpretieren.

Wir können also im distributiven Sinn Funktionen integrieren (bzw. Fourier-transformieren), die normalerweise nicht integrabel (bzw. Fourier-transformierbar) sind. Weiteres Beispiel: x^n, betrachtet als Distribution:

$$\frac{1}{\sqrt{2\pi}}\int x^n e^{-ikx}\,dx = \frac{1}{\sqrt{2\pi}}(i\partial_k)^n\int e^{-ikx}\,dx = \sqrt{2\pi}i^n\delta^{(n)}(x) \tag{4.6}$$

Genauer Beweis ohne die formale Schreibweise: Ü.

Wenn man in Gleichung (4.5) die Rolle von k und x vertauscht, kann man sie auch folgendermaßen lesen (und das wird in der Quantenmechanik oft gemacht):
Sei $e_k = \sqrt{2\pi}^{-1}e^{ikx}$. Analog zum Fall $k \in \mathbb{Z}$, bei dem die e_k's eine o. n. Basis in $\mathcal{L}^2([0,2\pi])$ bilden, faßt man die e_k für kontinuierliches k, die ja nicht $\in \mathcal{L}^2(\mathbb{R})$ sind, als eine orthogonale Basis aus „uneigentlichen Vektoren“ auf: sie sind paarweise orthogonal und man kann durch „Linearkombination“ aus ihnen jeden Vektor aufbauen; sie sind natürlich nicht normierbar im gewöhnlichen Sinn, sondern gemäß (4.5) „auf die δ-Funktion normiert“:

$$(e_k, e_l) = \frac{1}{2\pi}\int e^{-ikx+ilx}\,dx = \delta(k-l)\,.$$

Nicht jede Funktion, die eine reguläre Distribution definiert, läßt sich als Funktion Fourier-transformieren, wohl aber als Distribution. Dazu ist es oft nützlich, einen „konvergenzerzeugenden Faktor“ $e_m(x)$ einzuführen. Wir verlangen $e_m \in \mathcal{S}$ und $\lim_{m\to 0} e_m = 1$ in der $\mathcal{S}$-Topologie, z. B. $e_m(x) = e^{-mx^2}$. Da T stetig ist, gilt

$$\lim_{m\to 0}(e_m T, f) = \lim_{m\to 0}(T, e_m f) = (T, f) \quad \forall f \in \mathcal{T}\,,$$

$\mathcal{T} = \mathcal{S}$ oder $= \mathcal{D}$; d. h. im distributiven Sinn ist

$$\lim_{m\to 0} e_m T = T\,.$$

Wenden wir Definition 3.4.3 an, so folgt

$$\begin{aligned}\lim_{m\to 0}\left(\mathcal{F}(e_m T), \tilde f\right) &= \lim_{m\to 0}(e_m T, \mathcal{F}^{-1}\tilde f)\\ &= (T, \mathcal{F}^{-1}\tilde f) = (\mathcal{F}T, \tilde f) \quad \forall\, \tilde f \in \mathcal{F}\mathcal{T},\end{aligned}$$

d. h.

$$\lim_{m\to 0}\mathcal{F}(e_m T) = \mathcal{F}T \qquad \text{im distributiven Sinn.}$$

($\mathcal{F}T$ kann sehr wohl wieder eine reguläre Distribution sein.)

Das werden wir im nächsten Beispiel durchführen: wir wollen die Fourier-Transformierte von x_+^λ berechnen. Ich will x_+^λ als Distribution über $\mathcal{D}$ betrachten; dann ist $\mathcal{F}(x_+^\lambda)$ Distribution über $\mathcal{Z}$.
Als e_m wähle ich eine Funktion aus $\mathcal{S}$, die für $x \geq 0$ mit e^{-mx} übereinstimmt; solche Funktionen sind unschwer zu konstruieren, ihr Verhalten für $x < 0$ ist für unsere Zwecke unerheblich. (Man nehme z. B. eine Funktion g analog f_R aus (I.2.21), die $= 0$ ist für stark negative x, dann glatt ansteigt und $= 1$ ist für $x > -1$; dann ist ge^{-mx} von der gewünschten Form.) Berechnen wir also für $-1 < \lambda < 0$ und $m > 0$

$$\begin{aligned}\mathcal{F}(x_+^\lambda e_m)(k) &= \frac{1}{\sqrt{2\pi}}\int_0^\infty x^\lambda e^{-mx}e^{-ikx}\,dx\\ &= \frac{1}{\sqrt{2\pi}}\int_0^\infty x^\lambda e^{-(m+ik)x}\,dx\,.\end{aligned}$$

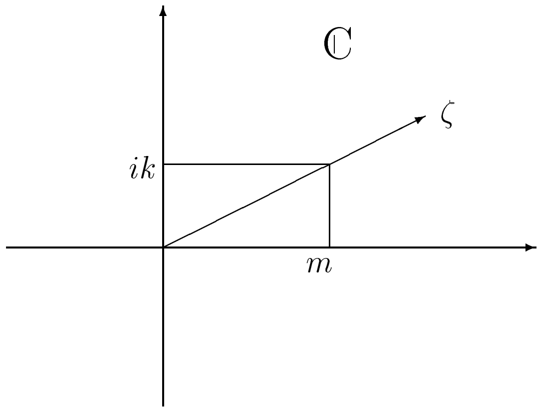
Ich führe eine neue (komplexe) Variable anstelle von x ein: $\zeta = (m+ik)x$. Dann ist

$$x = \frac{\zeta}{m+ik}\,, \qquad dx = \frac{d\zeta}{m+ik}\,,$$

und ζ läuft über eine von Null ausgehende Halbgerade $\mathcal{G}_+$ mit Steigung k/m relativ zur reellen Achse *in der rechten Halbebene*, weil $m > 0$; das Integral in der folgenden Formel ist also wohldefiniert:

$$\mathcal{F}(x_+^\lambda e^{-m|x|})(k) = \frac{1}{\sqrt{2\pi}}(m+ik)^{-\lambda-1}\int_{\mathcal{G}_+}\zeta^\lambda e^{-\zeta}\,d\zeta\,.$$

Der Integrand ist eine analytische Funktion in der rechten Halbebene, ich kann den Integrationsweg auf die reelle Achse verschieben. (Der Anteil über den unendlich fernen Bogen zwischen $\mathcal{G}_+$ und der reellen Achse verschwindet wegen des Faktors $e^{-\zeta}$, weil dort $\Re\zeta > 0$ ist.) Dann benutze ich die Darstellung $\int_0^\infty x^\lambda e^{-x}\,dx = \Gamma(\lambda+1)$ der Γ-Funktion (sie ist analytisch in der ganzen Ebene mit Ausnahme der Punkte $z = -1, -2, \ldots$, und $\Gamma(n+1) = n!$ für $n \in \mathbb{N}$). Der Vorfaktor läßt sich in der Form



$$-ie^{-i\pi\lambda/2}(k-im)^{-\lambda-1}$$

schreiben. Man beachte: nach Voraussetzung ist auch $-1 < (-\lambda-1) < 0$. Wenn wir nun den Limes $m \to 0$ machen, erhalten wir

$$\mathcal{F}(x_+^\lambda)(k) = -\frac{i}{\sqrt{2\pi}}e^{-i\pi\lambda/2}\Gamma(\lambda+1)\,(k-i0)^{-\lambda-1}\,. \tag{4.7}$$

Wenn man jetzt berücksichtigt, was ich nur gesagt, aber nicht gezeigt habe, daß nämlich x_+^λ in λ analytisch ist außerhalb der Punkte $-1, -2, \ldots$, dann kann man die obige Formel für alle diese Werte behaupten, denn auch die rechte Seite ist in diesem Bereich analytisch: $k - i0$ ist als Distribution $\in \mathcal{Z}'$ zu lesen, d. h. bei Anwendung auf eine analytische Testfunktion umgehe man bei der Integration den Nullpunkt in der oberen Halbebene; $k - i0$ wird also niemals 0, $(k-i0)^{-\lambda}$ ist folglich analytisch. Weil die analytische Fortsetzung links und rechts eindeutig ist, gilt (4.7) für alle $\lambda \neq -n$.

Und nun noch ein Beispiel im $\mathbb{R}^3$: die Funktion $1/|\vec{x}|$ ist (im $\mathbb{R}^3$) zwar bei 0 integrabel (man beachte, daß $d^3x = r^2\,dr\,d\Omega$), aber weder $\in \mathcal{L}^1$ noch $\in \mathcal{L}^2$. Wir können sie also nicht direkt Fourier-transformieren, wohl aber können wir es tun, wenn wir sie als (reguläre) Distribution betrachten.

Zur Berechnung benutze ich Polarkoordinaten $(r, \xi = \cos\vartheta, \varphi)$; wegen $\sin\vartheta\,d\vartheta = -d\xi$, $\xi \in [-1,1]$, schreibe ich $d^3x = r^2\,dr\,d\xi\,d\varphi$; außerdem $\vec{k}\cdot\vec{x} = kr\xi$. Die naive formale Rechnung

$$\begin{aligned}\int \frac{e^{-i\vec{k}\cdot\vec{x}}}{|\vec{x}|}\,d^3x &= 2\pi\int_0^\infty r\,dr\int_{-1}^1 d\xi e^{-ikr\xi}\\ &= 2\pi\int_0^\infty r\,dr\,\frac{1}{-ikr}e^{-ikr\xi}\,\Big|_{-1}^1\\ &= \frac{2\pi}{ik^2}\int_0^\infty d(kr)\left(e^{i(kr)} - e^{-i(kr)}\right)\end{aligned}$$

führt nicht zum Ziel, sie liefert nur den Hinweis, daß die Fourier-Transformierte vermutlich proportional $1/|\vec{k}|^2$ ist. Bessere Methode: man führe wieder einen konvergenzerzeugenden Faktor ein:

$$\begin{aligned}\int \frac{e^{-i\vec{k}\cdot\vec{x}}}{|\vec{x}|}e^{-\lambda|\vec{x}|}\,d^3x &= 2\pi\int_0^\infty r\,dr\int_{-1}^1 d\xi e^{-ikr\xi-\lambda r}\\ &= \frac{2\pi}{ik}\int_0^\infty dr\left(e^{(ik-\lambda)r}-e^{-(ik+\lambda)r}\right)\\ &= 4\pi\frac{1}{k^2+\lambda^2}\,;\end{aligned}$$

zusammen mit dem Faktor $(2\pi)^{-3/2}$ erhalten wir also im Limes $\lambda \to 0$

$$\mathcal{F}\left(\frac{1}{|\vec{x}|}\right) = (2\pi)^{-\frac{3}{2}}\,\frac{4\pi}{|\vec{k}|^2}\,. \tag{4.8}$$

Wir können diese Rechnung leicht so abändern, daß sie direkt die Definition der Fourier-Transformierten einer Distribution benutzt. Was wir zu machen haben, ist, den Ausdruck

$$(2\pi)^{-\frac{3}{2}}\int\frac{f(\vec{x})}{|\vec{x}|}\,d^3x$$

in der Form

$$\int \tilde{t}(\vec{k})\tilde{f}(\vec{k})\,d^3k$$

zu schreiben und dadurch $\tilde{t}$ zu bestimmen. Das ist aber leicht möglich, wenn wir in den ersten Ausdruck $f(\vec{x}) = \int d^3k\tilde{f}(\vec{k})e^{i\vec{k}\cdot\vec{x}}$ einsetzen, wie oben einen konvergenzerzeugenden Faktor hinzufügen und dadurch die Möglichkeit erhalten, die x-Integration und die k-Integration zu vertauschen. Dann können wir die obige Rechnung unter dem k-Integral ausführen und finden $\tilde{t} = \mathrm{const}\,k^{-2}$ wie oben.

Was hat man gewonnen, wenn man die Fourier-Transformierte einer Funktion als Distribution definieren kann?
Erste Antwort: z. B. kann man möglicherweise ein Integral $\int\overline{f(x)}h(x)$ dadurch lösen, daß man formal Gleichung (II.2.13) anwendet: $\ldots = \int\overline{\tilde{f}(k)}\tilde{h}(k)\,dk$; dieses Integral ist u. U. einfacher lösbar als das ursprüngliche, wenn z. B. $\tilde{f}$ delta-artige Terme enthält. Es ist wohldefiniert als $(T_{\tilde{f}},\tilde{h})$.
Zweite Antwort: Man kann manchmal, vergl. (4.8), Funktionen Fourier-transformieren, die im gewöhnlichen Sinne nicht transformierbar sind, und erhält als Transformierte wiederum Funktionen, die eigentlich keine inverse Transformation zulassen. Doch wenn man das Verfahren im distributiven Sinn wohldefinieren kann, darf man unbeschwert formal rechnen, darf Limiten und Integrale vertauschen etc., man muß nur kontrollieren, daß am Ende die entstehenden Ausdrücke wohldefiniert sind. Dritte Antwort: Mit der Fourier-Transformation haben wir ein systematisches Werkzeug, um Greensche Funktionen zu finden, so, wie wir die Lösungen verschiedener Differentialgleichungen durch Fourier-Transformation gefunden haben. Ich bringe als Beispiel die uns bereits bekannte Lösung von

$$\Delta G(\vec{x}) = \delta(\vec{x})\,,$$

vergl. die Einleitung zu diesem Kapitel, wo wir aber keinen sauberen Beweis gebracht haben. Durch Fourier-Transformation wird daraus

$$-|\vec{k}|^2 \widetilde{G}(\vec{k}) = (2\pi)^{-\frac{3}{2}},$$

also

$$G(\vec{x}) = -\mathcal{F}^{-1}\left(\frac{1}{(2\pi)^{3/2}|\vec{k}|^2}\right).$$

Die rechte Seite kennen wir aus Gleichung (4.8), also ist

$$G(\vec{x}) = -\frac{1}{4\pi|\vec{x}|}.$$

Ich brauche für eine spätere Anwendung auch die Greensche Funktion für den Differentialoperator $\Delta+\kappa^2$; ich will zeigen (analog zu der Übungsaufgabe $\Delta(1/r) = -4\pi\delta(\vec{x})$), daß

$$D\frac{e^{i\kappa|\vec{x}|}}{|\vec{x}|} \equiv (\Delta+\kappa^2)\frac{e^{i\kappa|\vec{x}|}}{|\vec{x}|} = -4\pi\delta(\vec{x}). \tag{4.9}$$

Ich setze wieder $r = |\vec{x}|$, in Polarkoordinaten lautet der Laplace-Operator

$$\Delta = \frac{1}{r}\,\partial_r^2\, r + \frac{L^2}{r^2},$$

wobei L^2 ein Differentialoperator bezüglich der Winkelkoordinaten ist (ohne konstanten Term, so daß $L^2 f(r) = 0$).[2] ($\partial_r^2\, r$ als Operator meint, daß die Differentiation auf r und auf die Funktion wirkt, auf die der Operator angewandt wird.) Damit sieht man sofort, daß für $r \neq 0$ gilt: $\Delta(e^{i\kappa r}/r) = (1/r)\partial_r^2 e^{i\kappa r} = -\kappa^2 e^{i\kappa r}/r$, also

$$D\frac{e^{i\kappa|\vec{x}|}}{|\vec{x}|} = 0 \quad \text{für } \vec{x} \neq 0. \tag{4.10}$$

Die Relation gilt natürlich auch für $-\kappa$ anstelle von κ. Um (4.9) zu zeigen, müssen wir die linke Seite auf eine Testfunktion f anwenden (man beachte, daß hier $D = D^*$):

$$\left(D\frac{e^{i\kappa|\vec{x}|}}{|\vec{x}|}, f(\vec{x})\right) = \left(\frac{e^{i\kappa|\vec{x}|}}{|\vec{x}|}, Df\right) = \int e^{-i\kappa|\vec{x}|}\frac{Df(\vec{x})}{|\vec{x}|}\, d^3x \overset{!}{=} -4\pi f(0).$$

Das Integral ist wohldefiniert, weil $Df \in \mathcal{S}(\mathbb{R}^3)$ und $d^3x/|\vec{x}| = r dr d\Omega$. Teilen wir das Integral auf in ein Integral über eine kleine Kugel um 0 mit Radius ε und den Rest:

$$I \equiv \int e^{-i\kappa|\vec{x}|}\frac{Df(\vec{x})}{|\vec{x}|}\, d^3x = \int_{K_\varepsilon} \ldots + \int_{\mathbb{R}^3\setminus K_\varepsilon} \ldots \equiv I_1(\varepsilon) + I_2(\varepsilon);$$

[2]Genau: $L^2 = \frac{1}{\sin\vartheta}\,\partial_\vartheta(\sin\vartheta\,\partial_\vartheta) + \frac{1}{sin^2\vartheta}\,\partial_\varphi^2$.

dann ist $\lim_{\varepsilon\to 0} I_1(\varepsilon) = 0$, und uns interessiert $I_2(\varepsilon)$. Der Integrationsbereich enthält nicht mehr den Punkt $\vec{x} = 0$, wir können also den gemäß (4.10) verschwindenden Term $f(\vec{x})D(e^{-i\kappa r}/r)$ im Integranden subtrahieren:

$$I_2(\varepsilon) = \int_{\mathbb{R}^3\setminus K_\varepsilon} \left(\frac{e^{-i\kappa r}}{r}\, \Delta f(\vec{x}) - f(\vec{x})\,\Delta\left(\frac{e^{-i\kappa r}}{r}\right)\right) d^3x\,,$$

die Terme mit κ^2 heben sich heraus. Jetzt wenden wir den zweiten Greenschen Satz an, um das Integral in ein Oberflächenintegral über $\partial(\mathbb{R}^3 \setminus K_\varepsilon)$ umzuwandeln. Diese Oberfläche besteht aus zwei Teilen: die unendlich ferne Kugel liefert nichts, weil f und seine Ableitungen stark abfallen, übrig bleibt die Oberfläche der kleinen Kugel K_ε mit nach innen weisender Normalen:

$$\ldots = \int_{\partial K_\varepsilon} \left(\frac{e^{-i\kappa r}}{r}\, \mathrm{grad} f - f\, \mathrm{grad}\left(\frac{e^{-i\kappa r}}{r}\right)\right) \cdot d\vec{\sigma}\,.$$

Mit $\vec{e}_r = \vec{x}/r,\ d\vec{\sigma} = -\vec{e}_r\,\varepsilon^2 d\Omega$ $(d\Omega = \sin\vartheta\, d\vartheta d\varphi)$ und

$$\mathrm{grad}\left(\frac{e^{-i\kappa r}}{r}\right) = -\frac{e^{-i\kappa r}}{r^2}\vec{e}_r - i\kappa\frac{e^{-i\kappa r}}{r}\vec{e}_r$$

finden wir schließlich

$$I_2(\varepsilon) = -\int_{\partial K_\varepsilon} \frac{e^{-i\kappa\varepsilon}}{\varepsilon}\, \mathrm{grad} f\cdot \vec{e}_r\, \varepsilon^2 d\Omega - i\kappa \int_{\partial K_\varepsilon} \frac{e^{-i\kappa\varepsilon}}{\varepsilon} f\, \varepsilon^2 d\Omega - \int_{\partial K_\varepsilon} f \frac{e^{-i\kappa\varepsilon}}{\varepsilon^2}\, \varepsilon^2 d\Omega\,.$$

Die ersten beiden Terme verschwinden für $\varepsilon \to 0$, der dritte liefert $-f(0)\int d\Omega = -4\pi f(0)$. ■

Als Anwendungsbeispiel will ich die Greensche Funktion für den eindimensionalen gedämpften getriebenen harmonischen Oszillator berechnen. Gesucht ist die Lösung der Gleichung

$$\ddot{x} + 2\gamma\dot{x} + \omega_0^2 x = F(t)$$

mit einer beliebigen zeitabhängigen Kraft F. Hier ist ω_0 die Eigenfrequenz des ungedämpften Oszillators, der γ-Term beschreibt die Dämpfung, die der Bewegung entgegenwirkt $(\gamma > 0)$ und proportional der Geschwindigkeit ist. Wir wollen hier den Fall „schwacher" Dämpfung behandeln: $\gamma < \omega_0$. Mithilfe der Greenschen Funktion $G(t)$, die dieselbe Gleichung mit $\delta(t)$ statt $F(t)$ löst, kann man dann die Bewegung des Oszillators für ein beliebiges $F(t)$ berechnen: $x = G * F$, siehe Satz 3.3.15.

Wir wenden die Fourier-Transformation auf die Gleichung an, bei der rechts $\delta(t)$ steht:

$$G(t) = \frac{1}{\sqrt{2\pi}} \int \tilde{G}(\omega) e^{-i\omega t}\, d\omega\,, \qquad -\omega^2\tilde{G} - 2i\gamma\omega\tilde{G} + \omega_0^2\tilde{G} = \frac{1}{\sqrt{2\pi}}\,.$$

Das ist einfach gelöst:

$$\tilde{G}(\omega) = -\frac{1}{\sqrt{2\pi}}\left(\omega^2 + 2i\gamma\omega - \omega_0^2\right)^{-1} = -\frac{1}{\sqrt{2\pi}}\Big(\,(\omega-\omega_+)(\omega-\omega_-)\,\Big)^{-1}\,,$$

wobei

$$\omega_\pm := -i\gamma \pm \omega_1, \qquad \omega_1 := \sqrt{\omega_0^2 - \gamma^2}\,.$$

Die Berechnung der inversen Fourier-Transformierten

$$G(t) = -\frac{1}{2\pi}\int \frac{e^{-i\omega t}\,dt}{(\omega-\omega_+)(\omega-\omega_-)}$$

läßt sich am einfachsten mit dem Residuensatz durchführen: der Integrand ist eine meromorphe Funktion mit einfachen Polen bei ω_+ und ω_-, die in der unteren Halbebene liegen. Wir integrieren von $-R$ bis R und machen den Limes $R \to \infty$; wenn $t < 0$ ist, können wir den Weg durch einen Halbkreis in der oberen Halbebene schließen, wegen des Terms $e^{-i\omega t}$ geht der zugehörige Beitrag im Limes $R \to \infty$ gegen 0; für $t > 0$ schließen wir in der unteren Halbebene.
Ergebnis: für $t < 0$ umschließt der Weg keinen Pol, also ist $G = 0$; für $t > 0$: der Weg wird in negativer Richtung durchlaufen, also ist

$$\begin{aligned} G(t) &= -\frac{1}{2\pi}(-2\pi i)\sum \text{Residuen} = i\frac{e^{-i\omega_+ t}}{\omega_+ - \omega_-} + i\frac{e^{-i\omega_- t}}{\omega_- - \omega_+} \\ &= \frac{i}{2\omega_1}e^{-\gamma t}\left(e^{-i\omega_1 t} - e^{i\omega_1 t}\right) = \frac{1}{\omega_1}e^{-\gamma t}\sin\omega_1 t\,. \end{aligned}$$

Man kontrolliert leicht, daß diese Greensche Funktion

$$G(t) = \begin{cases} \frac{1}{\omega_1}e^{-\gamma t}\sin\omega_1 t & \text{für } t > 0\,, \\ 0 & \text{für } t < 0 \end{cases}$$

die obige Gleichung erfüllt, sie ist eine stetige Funktion, die allerdings bei 0 einen Knick hat. Das ergibt beim zweimaligen Differenzieren die δ-Funktion.

3.5 Auf einer Fläche konzentrierte Distributionen

Wir haben solche Distributionen genau genommen schon kennengelernt: „Fläche“ meint eine Mannigfaltigkeit mit geringerer Dimension als die des betrachteten Raumes $\mathbb{R}^n$, im $\mathbb{R}^1$ also eine nulldimensionale Mannigfaltigkeit $\mathcal{M}_0$, die aus diskreten Punkten besteht. Und Summen von Delta-Distributionen sind auf einer solchen Mannigfaltigkeit konzentriert. In Anwendungen ist $\mathcal{M}_0$ oft durch die Nullstellen einer Funktion $a(x)$ gegeben: $\mathcal{M}_0 = \{x;\ a(x) = 0,\ a'(x) \neq 0\}$. Die Bedingung an a' garantiert, daß die Nullstellen isoliert liegen.
Frage: wie definiert man $\delta(a(x))$? Ein solcher Ausdruck ist uns schon einmal begegnet: $\delta(\omega^2/c^2 - k^2)$, ich komme nachher darauf zurück.

Vorübung: $a \in \mathbb{R},\ \delta(ax) = ?$
$\Big($Ich benutze die folgende **Notation**: $\int_I \ldots dx$ meint das nicht orientierte Integral über das Intervall $I = (a, b)$, $\int_I \ldots dx = \int_a^b \ldots dx$ wenn $a < b$, im Gegensatz zur Schreibweise $\int_a^b \ldots dx = -\int_b^a \ldots dx$, bei der die Integrationsrichtung das Vorzeichen bestimmt; noch genauer sollte ich $\int_I \ldots |dx|$ schreiben. Bei einer Variablentransformation $x \mapsto y = w(x)$ wird dann $dy = |w'(x)|dx$ (man denke sich die Absolutstriche bei den Differentialen hinzu); bei mehrdimensionalen Integralen tritt hier der *Absolutbetrag* der Jacobi-Determinante auf.$\Big)$
Formal ist leicht zu verstehen, was $\delta(ax)$ heißen soll:

$$\int \delta(ax) f(x)\, dx = \frac{1}{|a|} \int \delta(y) f(y/a)\, dy = \frac{1}{|a|} f(0)\,; \tag{5.1}$$

also

$$\delta(ax) = \frac{1}{|a|}\delta(x)\,. \tag{5.2}$$

Wenn ich die δ-Distribution durch glatte Funktionen d_α approximiere, kann ich die Rechnung in (5.1) exakt machen:

$$\lim_{\alpha\to 0} \int d_\alpha(ax) f(x)\, dx = \lim \frac{1}{|a|} \int d_\alpha(y) f(y/a)\, dy = \frac{1}{|a|} f(0)\,; \tag{5.3}$$

Wenn wir also $\delta(ax)$ durch $\lim d_\alpha(ax)$ definieren, bekommen wir exakt Gleichung (5.2). Ich betone, daß ich (5.2) nicht *bewiesen* habe, ich habe nur bewiesen, daß es eine konsistente *Definition* ist,$\delta(a\,\cdot\,) = \lim_{\alpha\to 0} T_{d_\alpha}$.

Gehen wir jetzt zurück zu $\delta(a(x))$. Der Ausdruck ist nur interessant, wenn $a(x)$ eine Nullstelle hat. Ich nehme an, daß a stetig differenzierbar ist, und setze zunächst voraus, daß es nur eine einfache Nullstelle gibt: $a(x_0) = 0$. Wie in (5.1) möchte man $y = a(x)$ als neue Variable im Integral einführen, dazu muß $a(x)$ invertierbar sein. Das ist in einer kleinen Umgebung von x_0 der Fall, weil dort $a'(x) \neq 0$, denn a' ist nach Voraussetzung stetig und $a'(x_0) \neq 0$. (Wir können ja das Integral auf eine kleine Umgebung des

Punktes beschränken, für den das Argument der Delta-Funktion verschwindet.) Es ist also $dy = a'(x)dx$, und wenn ich die Orientierung des Integrals nicht berücksichtige, $dx = |a'(x)|^{-1}dy$, wobei $x = a^{-1}(y)$ zu setzen ist (man beachte $x_0 = a^{-1}(0)$):

$$\int \delta(a(x))f(x)\,dx = \int_{|y|<\varepsilon} \delta(y)f(a^{-1}(y))\frac{dy}{|a'(a^{-1}(y))|} = \frac{f(x_0)}{|a'(x_0)|}\,. \tag{5.4}$$

($|y| < \varepsilon$ entspricht der Umgebung von x_0, in der a invertierbar ist; mehr als diese Umgebung brauchen wir ja nicht für die Berechnung des Integrals.) Wir erhalten also

$$\delta(a(x)) = \frac{1}{|a'(x_0)|}\,\delta(x - x_0)\,. \tag{5.5}$$

Das ist wieder im gleichen Sinne zu verstehen wie oben: wir können die Rechnung für approximierende Funktionen d_α ausführen und die linke Seite von (5.5) als Limes solcher Funktionen definieren, das ist dann wieder konsistent.
$a(x)$ kann mehrere Nullstellen x_i haben; wesentlich ist, daß es einfache Nullstellen sind, d. h. $a'(x_i) \neq 0$. Wir können dann eine Überdeckung von $\mathbb{R}$ durch offene Mengen U_i finden, so, daß in jedem U_i nur eine Nullstelle liegt. Dazu gibt es eine Zerlegung der Einheit durch Funktionen aus $\mathcal{S}$: $\varphi_i \geq 0$, $\varphi_i = 0$ auf U_i^c und $\sum_i \varphi_i(x) = 1$. Damit können wir auch ein beliebiges $f \in \mathcal{S}$ zerlegen: $f = \sum f_i$, $f_i(x) = f(x)\varphi_i(x)$, und für $\Big(\,\delta a(x), f_i(x)\,\Big)$ ist dann nur die Nullstelle x_i von Belang; mit (5.5):

$$\Big(\,\delta a(x), f_i(x)\,\Big) = \frac{f_i(x_i)}{|a'(x_i)|} = \frac{f(x_i)}{|a'(x_i)|}\,,$$

weil $\varphi_i(x_i) = 1$ nach Voraussetzung. Summation über alle i liefert schließlich

$$\delta(a(x)) = \sum_{\text{Nullstellen}} \frac{1}{|a'(x_i)|}\,\delta(x - x_i)\,. \tag{5.6}$$

Als Anwendungsbeispiel betrachte ich die Gleichung, die uns bei der Lösung der Wellengleichung durch Fourier-Transformation begegnet ist:

$$\left(\frac{\omega^2}{c^2} - \vec{k}\cdot\vec{k}\right)\tilde{f}(\vec{k},\omega) = 0\,.$$

Als Distributionsgleichung gelesen, hat sie die Lösung

$$\tilde{f}(\omega,\vec{k}) = h(\omega,\vec{k})\,\delta(a(\omega)) \qquad \text{mit } a(\omega) = \frac{\omega^2}{c^2} - |\vec{k}|^2\,.$$

Die Nullstellen sind $\omega = \pm c|\vec{k}|$, wobei uns aus physikalischen Gründen nur diejenige mit positivem ω interessiert; $a'(\omega) = 2\omega/c^2$ ist $\neq 0$ bei $\omega = c|\vec{k}|$. Daher ist

$$\tilde{f}(\omega,\vec{k}) = \frac{c}{2|\vec{k}|}h(\omega,\vec{k})\,\Big(\,\delta(\omega - c|\vec{k}|) + \ldots\,\Big)\,.$$

Den durch Pünktchen angedeuteten Term für die zweite Nullstelle wollen wir weglassen; die Fourier-Rücktransformation ergibt

$$f(t,\vec{x}) = (2\pi)^{-2}\int \tilde{f}(\omega,\vec{k})e^{i(\vec{k}\cdot\vec{x}-\omega t)}\, d^3k d\omega = \int \tilde{g}(\vec{k})e^{i(\vec{k}\cdot\vec{x}-c|\vec{k}|t)}\, d^3k$$

mit $\tilde{g}(\vec{k}) = \frac{c}{4\pi|\vec{k}|}\, h(c|\vec{k}|,\vec{k})$.
Das ist genau der Ausdruck, den ich mit etwas schwacher Argumentation früher in Gleichung (II.3.6) angegeben habe.

Die im letzten Beispiel auftretende Distribution kann man auch anders auffassen: als eine auf einer Hyperfläche konzentrierte Distribution. Mit $(\omega,\vec{k}) \in \mathbb{R}^4$, $P(\omega,\vec{k}) := \omega^2/c^2 - |\vec{k}|^2$ definiert die Gleichung $P = 0$ eine 3-dimensionale Hyperfläche, außerhalb derer $\delta(P)$ verschwindet.
Das führt uns zu einer allgemeinen Definition, mit der wir ohne formale Schreibweise auch die obigen Definitionen streng formulieren können.

Betrachten wir nun ganz allgemein Flächen im $\mathbb{R}^n$ („Hyper“ lasse ich weg), genauer: glatte Flächen $\mathcal{F}$, gegeben durch $P(\vec{x}) = 0$. „Glatt“ heißt, daß der Gradient $\vec{\partial}P$ in einer Umgebung von $\mathcal{F}$ wohldefiniert ist, die Fläche besitzt keine singulären Punkte; außerdem nehme ich an, daß $\vec{\partial}P \neq 0$.
Sei $d^{n-1}o$ das Oberflächenelement der $(n-1)$-dimensionalen Fläche $\mathcal{F}$, dann ist z. B.

$$(T,f) := \int_{\mathcal{F}} t(\vec{x})f(\vec{x})\, d^{n-1}o$$

eine auf $\mathcal{F}$ konzentrierte Distribution.

Wir wollen nun eine δ-artige, auf $\mathcal{F}$ konzentrierte Distribution definieren; dazu lassen wir uns vom Spezialfall leiten, daß $\mathcal{F}$ eine Koordinatenebene ist, z. B. $\{\vec{x};\ P(\vec{x}) = x_1 = 0\}$. Das Oberflächenelement ist dann $d^{n-1}o = dx_2 \dots dx_n$. Formal geschrieben, möchten wir definieren

$$(\delta_P, f(\vec{x})) = \int \delta(x_1)f(\vec{x})\, d^nx = \int f(0,x_2,\dots,x_n)\, dx_1 \dots dx_n\,. \tag{5.7}$$

Verallgemeinert: wir können neue (krummlinige) Koordinaten $u_1,\dots u_n$ einführen, so daß $\mathcal{F}$ gegeben ist durch $u_1 = 0$. Das heißt, wir setzen $u_1 = P(\vec{x})$, das ist möglich, weil $\vec{\partial}P \neq 0$; und $u_2,\dots u_n$ sind dann Koordinaten in der Fläche. Dann können wir in Anlehnung an (5.7) definieren

$$(\delta_P, f(\vec{x})) = \int_{\mathcal{F}} f(\vec{x})\,\Big|_{u_1=0}\, d^{n-1}o\,. \tag{5.8}$$

Ich muß noch dafür sorgen, daß wir hierbei keine Koordinatenwillkür einführen; wenn ich nicht u_i sondern $\tilde{u}_i = cu_i$ wählen würde, hätte ich $d^{n-1}\tilde{o} = c^{n-1}d^{n-1}o$, und auch δ_P würde um diesen Faktor vergrößert. Wir fordern also als Normierung

$$du_1\, d^{n-1}o = d^nx\,; \tag{5.9}$$

dann ist die Definition eindeutig. $d^{n-1}o$ ist hier als nicht orientiert zu betrachten, genauer müßten wir in (5.8) $|d^{n-1}o|$ schreiben.
Wenn wir nun die Fläche $\mathcal{F}$ durch $\widetilde{P}(\vec{x}) := cP(\vec{x}) = 0$ (mit $c > 0$) charakterisieren, wird $\tilde{u}_1 = cu_1$ und für $d^{n-1}\tilde{o}$ müssen wir nach (5.9) $d^{n-1}\tilde{o} = (1/c)d^{n-1}o$ setzen, also

$$\delta_{cP}(f) = \int_{\mathcal{F}} f(\vec{x})\,\Big|_{u_1=0}\; d^{n-1}\tilde{o} = \frac{1}{c}\delta_P(f)\,,$$

und das ist in Übereinstimmung mit unserer früherenGleichung (5.2). (Das ist also eine etwas andere bezüglich $\mathcal{F}$ delta-artige Distribution; ich habe deshalb nicht $\mathcal{F}$ als Index gewählt, sondern P bzw. cP.)

Man kann das Spiel wiederholen, wenn man Flächen geringerer Dimension betrachten will: Sei $\dim\mathcal{F} = n-k$ und $\mathcal{F} = \cap_{i=1}^{k}\mathcal{F}_i$, wobei $\mathcal{F}_i$ $(n-1)$-dimensionale Flächen sind, gegeben durch $P_i(\vec{x}) = 0$ mit $\vec{\partial}P_i \neq 0$. Dann können wir Koordinaten u_i einführen mit $u_i = P_i$ für $i = 1, \dots k$ und für das Oberflächenelement von $\mathcal{F}$ die Bedingung

$$du_1 \dots du_k d^{n-k}o = d^n x \tag{5.10}$$

stellen; die auf $\mathcal{F}$ konzentrierte Delta-Funktion ist gegeben durch

$$\delta_{P_1\dots P_k}(f) = \int_{\mathcal{F}} f(\vec{x})\,\Big|_{u_1=\dots=u_k=0}\; d^{n-k}o\,, \tag{5.11}$$

in voller Analogie zu Gleichung (5.8). Die Einschränkung an f habe ich nur der Deutlichkeit halber hingeschrieben; bei Integration über $\mathcal{F}$ interessieren ja nur die Werte von f auf $\mathcal{F}$, d. h. die durch diese Einschränkung gegebenen Werte.

Wenn man $k = n$ setzt, wird $\mathcal{F}$ nulldimensional, besteht also nur noch aus diskreten Punkten, und dann kann man streng Gleichung (5.6) begründen.

Für diese Begründung und um zu zeigen, daß δ_P und $\delta_{P_1\dots P_k}$ völlig koordinatenunabhängig definiert sind, sollte man die obigen Formeln als Gleichungen für Differentialformen bzw. als Integration über Differentialformen auffassen.

Exkurs über Differentialformen: siehe Anhang B.

Das (orientierte) Flächenelement auf $\mathcal{F}$ mit $\dim\mathcal{F} = n-k$ ist eine $(n-k)$-Form ω, die bis auf die Normierung eindeutig bestimmt ist. Ich will die Gleichungen (5.10) und (5.11) jetzt schreiben als

$$dx_1 \wedge \dots \wedge dx_n \;=\; dP_1 \wedge \dots \wedge dP_k \wedge \omega\,; \tag{5.12}$$

$$\delta_{P_1\dots P_k}(f) \;:=\; \int_{\mathcal{F}} f(\vec{x})\,|\omega|\,; \tag{5.13}$$

durch (5.12) ist ω vollständig festgelegt.

Wenn $k = n$ ist, ist $\omega = \omega^0$ eine Null-Form, d. h. eine Funktion; und Integration dieser Null-Form ω^0 über eine nulldimensionale Mannigfaltigkeit (die also nur aus diskreten Punkten besteht), bedeutet nur Summation der Funktionswerte von ω^0 an diesen Punkten. Betrachten wir den $\mathbb{R}^2$ (mit Koordinaten x, y); sei $P_1(x, y) = a(x) - y$ und $P_2(x, y) = y$. Die dadurch definierte nulldimensionale Menge $\mathcal{M}_0 = \{P_1 = 0\} \cap \{P_2 = 0\}$ ist die Menge der Nullstellen von $a(x)$, d. h. $\delta_{P_1 P_2}$ ist unsere früher formal als $\delta(a(x))$ geschriebene Distribution. Es gilt

$$dP_1 = a'(x)\,dx - dy, \qquad dP_2 = dy, \qquad dP_1 \wedge dP_2 = a'(x)\,dx \wedge dy\,.$$

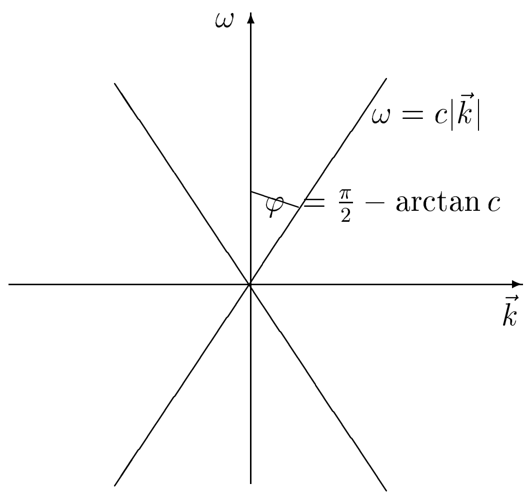
Wir müssen also $\omega^0 = a'(x)^{-1}$ setzen, um

$$dP_1 \wedge dP_2 \wedge \omega^0 = dx \wedge dy$$

zu erreichen; und dann liefert Gleichung (5.13)

$$\delta_{P_1 P_2}(f) = \sum_{\text{Nullstellen}} \frac{f(x_i, 0)}{|a'(x_i)|}\,.$$

Das entspricht genau Gleichung (5.6), wenn wir die y-Abhängigkeit der Testfunktion in der obigen Formel vergessen bzw. in (5.6) die Funktion $f(x)$ durch $f(x)h(y)$ mit $h(0) = 1$ ersetzen.



Gehen wir uns noch einmal zurück zum früheren Beispiel aus der Elektrodynamik: $P(\omega, \vec{k}) \equiv (\omega^2/c^2 - |\vec{k}|^2) = 0$ definiert eine dreidimensionale Hyperfläche, einen Doppelkegel mit der ω-Achse als Symmetrieachse und Spitze im Koordinatenursprung mit einem durch c bestimmten Öffnungswinkel. Die obere Hälfte ohne den Nullpunkt heiße $\mathcal{K}_+$, sie ist eine glatte Fläche. Ich benutze wieder Polarkoordinaten für den $\vec{k}$-Raum (ω ist jetzt natürlich eine Koordinate, keine 1-Form), das vierdimensionale Volumenelement lautet

$$\Omega^{(4)} \equiv d\omega \wedge dk_1 \wedge dk_2 \wedge dk_3 = k^2\,d\omega \wedge dk \wedge d\Omega\,.$$

($d\Omega$ sollte ich eigentlich $\Omega^{(2)}$ schreiben, gemeint ist die Winkel-2-Form $\sin\vartheta\,d\vartheta \wedge d\varphi$. Wie früher will ich nur positive Frequenzen betrachten, d. h. $\omega > 0$, daher die Beschränkung auf den nach oben weisenden Kegel.) Ich behaupte, daß für

$$\widetilde{\omega}^3 := (1/4)(k\,d\omega + (c^2 k^2/\omega)\,dk) \wedge d\Omega$$

die Relation $\Omega^{(4)} = dP \wedge \widetilde{\omega}^3$ gilt:

$$\begin{aligned} P &= \frac{\omega^2}{c^2} - k^2\,, \\ dP &= \frac{2\omega}{c^2}\,d\omega - 2k\,dk\,, \\ dP \wedge \widetilde{\omega}^3 &= 2\left(\frac{\omega}{c^2}\,d\omega - k\,dk\right) \wedge \frac{1}{4}\left(k\,d\omega + \frac{c^2k^2}{\omega}\,dk\right) \wedge d\Omega \\ &= \frac{1}{2}(-k^2\,dk \wedge d\omega + k^2\,d\omega \wedge dk) \wedge d\Omega = k^2\,d\omega \wedge dk \wedge d\Omega\,. \end{aligned}$$

Damit wird

$$\delta_P(f) = \int_{\mathcal{K}_+} f(\omega, \vec{k})\,|\widetilde{\omega}^3|\,,$$

und weil auf $\mathcal{K}_+$ die Relationen $\omega = ck$, und folglich $d\omega = c\,dk$ gelten, ist $\widetilde{\omega}^3\Big|_{\mathcal{K}_+} = (1/4)(ck\,dk + (c^2k^2/ck)\,dk) \wedge d\Omega = (c/2)k\,dk \wedge d\Omega$. Folglich ist

$$\begin{aligned} \delta_P(f) &= \frac{c}{2}\int_{\mathcal{K}_+} f(ck, \vec{k})k\,|dk \wedge d\Omega| \\ &= \frac{c}{2}\int_{\mathcal{K}_+} f(ck, \vec{k})\frac{k^2}{|\vec{k}|}\,|dk \wedge d\Omega| \\ &= \int_{\mathbb{R}^3} \frac{c}{2|\vec{k}|} f(c|\vec{k}|, \vec{k})\,d^3k\,. \end{aligned}$$

Die Rechnung mit unserer früheren Gleichung (5.6) liefert dasselbe:

$$\begin{aligned} \delta\left(\frac{\omega^2}{c^2} - k^2\right) &= \frac{c^2}{2\omega}\delta(\omega - ck) + \ldots\,, \\ \int f(\omega, \vec{k})\delta\left(\frac{\omega^2}{c^2} - k^2\right) d^3k d\omega &= \int \frac{c}{2|\vec{k}|} f(c|\vec{k}|, \vec{k})\,d^3k\,. \end{aligned}$$

(Hier ist die negative Nullstelle wieder weggelassen worden.)

Jetzt folgt ein weiteres Beispiel aus der Physik, bei dem eine auf einer Fläche konzentrierte Distribution auftritt.
Um das Strahlungsfeld einer Antenne mit vorgegebener zeitabhängiger Strom- und Ladungsdichte zu berechnen, muß man die inhomogene Wellengleichung

$$\left(\frac{1}{c^2}\partial_t^2 - \Delta\right) f(\vec{x}, t) = g(\vec{x}, t) \tag{5.14}$$

lösen können. $g(\vec{x}, t)$ ist hier eine bekannte Funktion, die durch die Strom- bzw. Ladungsdichte gegeben ist, je nachdem für welche Komponente des elektromagnetischen Feldes die gesuchte Funktion f steht. Die Lösung kann man mit der zugehörigen Greenschen Funktion $G(\vec{x}, t)$ finden:

$$\left(\frac{1}{c^2}\partial_t^2 - \Delta\right) G = \delta\,.$$

Die rechte Seite ist als 4-dimensionale Delta-Funktion zu lesen, in formaler Schreibweise: $\delta(\vec{x})\delta(t)$ (streng: $\delta^{[4]} = \delta^{[3]} \times \delta$). Wir führen die Fourier-Transformation dieser Gleichung bezüglich der Zeit durch (man beachte $\mathcal{F}\delta = \sqrt{2\pi}^{-1}$):

$$\begin{aligned} G(\vec{x},t) &= \frac{1}{\sqrt{2\pi}} \int \widetilde{G}(\vec{x},\omega) e^{i\omega t}\, dt\,, \\ \left(-\frac{\omega^2}{c^2} - \Delta\right) \widetilde{G}(\vec{x},\omega) &= \frac{1}{\sqrt{2\pi}} \delta(\vec{x})\,. \end{aligned}$$

Die Lösung dieser Gleichung kennen wir schon: wir müssen in Gleichung (4.9) $\kappa = \omega/c$ setzen und einen Faktor $\sqrt{2\pi}^{-1}$ ergänzen:

$$\widetilde{G}(\vec{x},\omega) = \frac{1}{\sqrt{2\pi}} \frac{e^{i\omega r/c}}{4\pi r}\,.$$

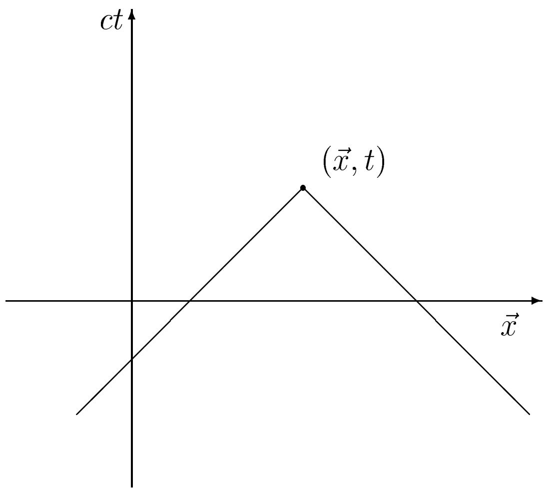
Durch Fourier-Rücktransformation erhält man unter Berücksichtigung von $\int e^{-i\omega\tau}\, d\omega = 2\pi\delta(\tau)$

$$G(\vec{x},t) = \frac{1}{2\pi} \frac{1}{4\pi r} \int e^{i\omega r/c} e^{-i\omega t}\, d\omega = \frac{1}{4\pi r} \delta\left(t - \frac{r}{c}\right)\,.$$

Die allgemeine Lösung von (5.14) ist dann gegeben durch $f = f_0 + G * g$, wobei f_0 die allgemeine Lösung der homogenen Gleichung ist; ausführlich geschrieben:

$$\begin{aligned} f(\vec{x},t) &= f_0(\vec{x},t) + \frac{1}{4\pi} \int \frac{g(\vec{x}',t')}{|\vec{x}-\vec{x}'|} \delta\left(t - t' - \frac{|\vec{x}-\vec{x}'|}{c}\right) d^3x'dt' \\ &= f_0(\vec{x},t) + \frac{1}{4\pi} \int \frac{g\left(\vec{x}', t - \frac{|\vec{x}-\vec{x}'|}{c}\right)}{|\vec{x}-\vec{x}'|} d^3x\,. \end{aligned} \tag{5.15}$$

Die Greensche Funktion (die ja genau genommen eine Distribution ist) ist konzentriert auf dem „Rückwärtslichtkegel", entsprechend der Tatsache, daß sich elektromagnetische Wellen mit Lichtgeschwindigkeit ausbreiten. Der Rückwärtslichtkegel ist die Menge aller Punkte $(\vec{x},t)$ in der vierdimensionalen Raum-Zeit, von denen ein Photon, das mit der Geschwindigkeit c fliegt, den räumlichen Nullpunkt $\vec{0}$ zur Zeit $t = 0$ erreichen kann. Die Delta-Funktion im obigen Integral ist konzentriert auf dem zum Punkt $(\vec{x},t)$ verschobenen Rückwärtslichtkegel.



Gleichung (5.15) besagt, daß der Wert von f zur Zeit t am Ort $\vec{x}$ nur von der Stärke des Quellterms g an denjenigen Raum-Zeit-Punkten $(\vec{x}',t')$ beeinflußt wird, von denen aus ein Lichtsignal den Ort $\vec{x}$ zur

Zeit t erreichen kann. Die Wirkung dessen, was am Ort $\vec{x}'$ zur Zeit t' durch Ströme und Ladungen verursacht wird, propagiert als elektromagnetische Welle, die sich genau mit Lichtgeschwindigkeit fortpflanzt.

Die Wellengleichung ist von zweiter Ordnung in der Zeitableitung, d. h. wenn man in G das Argument t durch $-t$ ersetzt, erhält man wieder eine Greensche Funktion. Gefaltet mit g ergibt sie eine Lösung, bei der der von $(\vec{x}, t)$ ausgehende Rückwärtslichtkegel durch den Vorwärtslichtkegel ersetzt ist, es ist die „avancierte Lösung“. Wir haben die physikalisch relevante „retardierte“ Lösung gewählt, bei der die Ursache vor der Wirkung liegt.

Im Anhang C finden sich einige Ergänzungen zu diesem Kapitel.

Kapitel 4

Operatoren im Hilbertraum

4.1 Einleitung: Operatoren als Observable

Die Motivation für dieses Kapitel läßt sich in einem Satz geben: Man braucht den Hilbertraum und Operatoren im Hilbertraum zur Beschreibung von „Zuständen“ und „Observablen“ in der Quantenmechanik.
Diese Begriffe kann man auch in der klassischen Mechanik einführen, doch wirken sie ein bißchen künstlich und unnötig; sie werden deutlicher, wenn man die Statistische Mechanik betrachtet; und sie sind für eine saubere Darstellung der Quantenmechanik unerläßlich.
Was ich Ihnen in diesem Abschnitt erzähle, werden Sie später genauer lernen. Ich bringe es als Motivation in sehr verkürzter Form; es ist nicht entscheidend für die Mathematik, die ich in den folgenden Abschnitten bringen will. Umgekehrt: diese Mathematik wird in Kurzform auch in der Quantenmechanik-Vorlesung gebracht; doch was ich in einem halben Semester erzählen will, ist dort auf wenige Vorlesungsstunden zusammengepreßt. Wichtig in meiner Darstellung (und mathematisch kein ganz leichtes Brot): das Spektraltheorem nebst Beweis. (Für diesen Beweis ist keine Zeit in der QM-Vorlesung.)
Wenn Sie mit der Hilbertraum-Mathematik ein bißchen vertraut sind, wird Ihnen das Verständnis der Quantenmechanik deutlich leichter fallen.

In der Theorie der Distributionen lernt man, wie man ohne schlechtes Gewissen Integrationen vertauschen, nicht integrierbare Funktionen mit Tricks Fourier-transformieren, nicht differenzierbare Funktionen differenzieren kann. Eine ähnliche Rolle spielt die saubere Behandlung der Hilbertraum-Operatoren und insbesonder die Spektralzerlegung für den laxen Umgang mit Operatoren und uneigentlichen Eigenvektoren in der Quantenmechanik, wie er besonders häufig vorkommt, wenn man die Dirac-Schreibweise benutzt (und wie er in vielen Quantenmechanik-Lehrbüchern ohne Kommentar gehandhabt wird).

„Observable“ heißt „beobachtbare Größe“. Beispiele aus der klasssischen Mechanik: Orte, Impulse, Energien, Gesamtenergie, Gesamtdrehimpuls usw. Welche Werte diese Größen haben, hängt vom speziellen Zustand des betrachteten Systems ab, in der Mechanik z. B. von den Anfangsbedingungen, welche die Bahnen der Körper (oder gebildet ausgedrückt: welche die Bahn des Systems im Phasenraum) festlegen, und im allgemeinen auch von der Zeit.

Beispiel für den „Zustand“ in der Statistischen Mechanik, welches der QM näher ist: wenn man sehr viele Teilchen, z. B. die Moleküle eines Gases, betrachtet, kann man die Anfangsbedingungen nicht mehr festlegen wie in der Mechanik; man macht statistische Aussagen über Orte und Geschwindigkeiten (bzw. Impulse $\vec{p}_i = m_i \vec{v}_i$): die Größe

$$\rho(\vec{x}_1, \ldots, \vec{x}_N, \vec{p}_1, \ldots, \vec{p}_N, t)\, d^{3N}x\, d^{3N}p$$

gibt die Wahrscheinlichkeit an, daß zur Zeit t Teilchen 1 am Ort $\vec{x}_1$ im kleinen Volumen d^3x_1 ist und den Impuls $\vec{p}_1$ (mit der Unschärfe d^3p_1) hat, Teilchen 2 bei $\vec{x}_2$ ist .... Damit es sich um eine Wahrscheinlichkeit handelt, muß $\int \rho(\vec{x}_1, \ldots, \vec{x}_N, \vec{p}_1, \ldots, \vec{p}_N, t)\, d^{3N}x\, d^{3N}p = 1$ gelten. (Man lernt in der Mechanik in der Hamiltonschen Formulierung, daß die Impulse die „natürlicheren“ Größen sind als die Geschwindigkeiten. Das gilt erst recht in der Quantenmechanik; wenn ich von kinetischer Energie rede, werde ich deshalb immer $|\vec{p}|^2/(2m)$ schreiben anstelle des Ihnen vielleicht vertrauteren $\frac{1}{2}m|\vec{v}|^2$.)
Der Ort des ersten Teilchens ist nicht scharf definiert; doch wir können seinen Mittelwert (seinen „Erwartungswert“) zur Zeit t berechnen, wenn wir den Zustand kennen, der durch $\rho(\ldots, t)$ definiert ist:

$$\langle \vec{x}_1 \rangle_\rho = \int \vec{x}_1 \rho(\vec{x}_1, \ldots, \vec{x}_N, \vec{p}_1, \ldots, \vec{p}_N, t)\, d^{3N}x\, d^{3N}p\,.$$

Ebenso berechnen wir die Erwartungswerte anderer Observabler in dem durch ρ gegebenen Zustand. Allgemein sind Observable Funktionen von $(\vec{x}_1, \ldots, \vec{x}_N, \vec{p}_1, \ldots, \vec{p}_N)$, denken Sie beispielsweise an die Energie:

$$E(\vec{x}_1, \ldots, \vec{x}_N, \vec{p}_1, \ldots, \vec{p}_N) = \sum \frac{|\vec{p}_i|^2}{2m_i} + V(\vec{x}_1, \ldots, \vec{x}_N)\,,$$

wenn die Wechselwirkung zwischen den Teilchen durch ein Potential V gegeben ist. Erwartungswert der Energie:

$$\langle E \rangle_\rho = \int E(\vec{x}_1, \ldots, \vec{x}_N, \vec{p}_1, \ldots, \vec{p}_N) \rho(\vec{x}_1, \ldots, \vec{x}_N, \vec{p}_1, \ldots, \vec{p}_N)\, d^{3N}x\, d^{3N}p\,.$$

(Wenn wir klassische Mechanik treiben, ist der Zustand zur Zeit t festgelegt durch Orte und Impulse der Teilchen zu dieser Zeit, ρ ist δ-artig; und $\langle E \rangle_\rho$ ist dann einfach der entsprechende Funktionswert der Funktion E.)

Auch in der Quantenmechanik müssen wir mit Wahrscheinlichkeitsaussagen leben. Der Zustand eines Systems wird mit der Wellenfunktion beschrieben, die der Schrödingergleichung gehorcht. Ihr Absolutquadrat muß als Wahrscheinlichkeitsdichte interpretiert werden, analog zur Statistischen Mechanik: $\rho(\vec{x},t)\,d^3x := (1/N)|\psi(x,t)|^2\,d^3x$ gibt die Wahrscheinlichkeit, ein Teilchen bei $\vec{x}$ im Volumen d^3x zu finden. N ist hier ein Normierungsfaktor, denn wir müssen wieder $\int \rho\,d^3x = 1$ haben: $N = \int |\psi(\vec{x},t)|^2\,d^3x$. Folgerung: notwendig ist $\psi \in \mathcal{L}^2(\mathbb{R}^3)$. Wie oben können wir auch hier den Erwartungswert des Ortes bilden, nehmen wir die l-te Koordinate:

$$\langle x_l\rangle_\psi = \frac{1}{N}\int x_l|\psi(\vec{x},t)|^2\,d^3x = \frac{1}{N}\int \overline{\psi(\vec{x},t)}x_l\psi(\vec{x},t)\,d^3x\,. \tag{1.1}$$

Der Grund für die letzte Umformung wird in Kürze klar werden.
Wir können uns ψ geschrieben denken als ein „Wellenpaket“, d. h. eine Überlagerung ebener Wellen, oder anders gesagt: wir können ψ durch seine Fourier-Transformierte $\widetilde{\psi}(\vec{k},t)$ ausdrücken. Bei einer ebenen Welle, und ebenso beim Wellenpaket, muß man den Wellenvektor $\vec{k}$ wie beim Licht als proportional dem Impuls betrachten. Heurististische Überlegung: der relativistische Zusammenhang von Energie und Impuls ist gegeben durch $E = c\sqrt{|\vec{p}|^2 + m^2c^2}$); Photonen sind Teilchen mit verschwindender Ruhmasse, und ihre Energie hängt bekanntlich von der Frequenz des Lichts ab: $E = \hbar\omega$; wegen $m = 0$ ist also $\hbar\omega = c|\vec{p}|$; andererseits hängt ω mit dem Wellenvektor zusammen: $\omega = c|\vec{k}|$, also müssen wir $\hbar\vec{k} = \vec{p}$ setzen. Diesen Zusammenhang übernimmt man in der Quantenmechanik. Konsequenterweise muß man dann $\widetilde{\psi}(\vec{k},t)$ als Wahrscheinlichkeitsdichte für den Impuls interpretieren: mit Wahrscheinlichkeit $(1/N)|\widetilde{\psi}(\vec{k},t)|^2\,d^3k$ hat das Teilchen den Impuls $\hbar\vec{k}$ mit Unschärfe $\hbar^3 d^3k$; der mittlere Impuls ist also

$$\langle p_l\rangle_\psi = \frac{1}{N}\int \hbar k_l|\widetilde{\psi}(\vec{k},t)|^2\,d^3k = \frac{1}{N}\int \overline{\widetilde{\psi}(\vec{k},t)}\hbar k_l\widetilde{\psi}(\vec{k},t)\,d^3k\,. \tag{1.2}$$

Ich habe auf der linken Seite den Index ψ und nicht $\widetilde{\psi}$ benutzt, denn zwischen ψ und $\widetilde{\psi}$ gibt es ja eine eineindeutige Beziehung. Man beachte außerdem, daß wegen des Theorems von Plancherel, Gleichung (II.2.13), $N = \int |\psi(\vec{x},t)|^2\,d^3x = \int |\widetilde{\psi}(\vec{k},t)|^2\,d^3k$ ist. Mit (II.2.13) können wir auch $\langle p_l\rangle_\psi$ leicht durch $\psi(\vec{x},t)$ ausdrücken:

$$\langle p_l\rangle_\psi = \frac{1}{N}\int \overline{\psi(\vec{x},t)}\,\mathcal{F}^{-1}\Big(\hbar k_l\widetilde{\psi}(\vec{k},t)\Big)\,d^3x = \frac{1}{N}\int \overline{\psi(\vec{x},t)}\frac{\hbar}{i}\partial_{x_l}\psi(\vec{x},t)\,d^3x\,. \tag{1.3}$$

In der „Ortsdarstellung“ ausgedrückt ist also der Impuls zu beschreiben durch den Differentialoperator $\frac{\hbar}{i}\vec{\partial}_x$. Wir können auch den Ort auffassen als beschrieben durch einen Operator $\vec{Q}$: Multiplikation mit $\vec{x}$, für die l-te Koordinate:

$$Q_l: \quad \psi(\vec{x},t) \mapsto (Q_l\psi)(\vec{x},t) = x_l\psi(\vec{x},t)\,.$$

Man beachte: $\mathcal{L}^2(\mathbb{R}^3)$ ist ein Hilbertraum, und wir können die Integrale in (1.1) und (1.3) als Skalarprodukte von ψ mit $Q_l\psi$ bzw. mit $\frac{\hbar}{i}\partial_{x_l}\psi$ lesen:

$$\langle x_l\rangle_\psi \equiv \langle Q_l\rangle_\psi = \frac{(\psi, Q_l\psi)}{(\psi,\psi)};\ \langle p_l\rangle_\psi = \frac{(\psi, \frac{\hbar}{i}\partial_{x_l}\psi)}{(\psi,\psi)}\,.$$

Beachten Sie, daß die Wellenfunktionen $\psi(\vec{x})$ und $c\psi(\vec{x})$ denselben Zustand beschreiben: bei der notwendigen Berücksichtigung der Normierung fällt der Faktor wieder heraus.

Wenn wir nun andere Observable, die Funktionen von $\vec{x}$ und $\vec{p}$ sind, wie z. B. Energie und Drehimpuls aus der Mechanik in die Quantenmechanik übersetzen wollen, müssen wir $\vec{x}$ durch den Multiplikationsoperator $\vec{Q}$ und $\vec{p}$ durch den Differentiationsoperator $\vec{P} := \frac{\hbar}{i}\vec{\partial}_x$ ersetzen; wir bekommen also im allgemeinen wieder Differentialoperatoren.
Weitere Beispiele:
Die Observable der Energie ist in der klassischen Mechanik im einfachsten Fall die Hamilton-Funktion $H(\vec{p},\vec{x}) = T + V = |\vec{p}|^2/2m + V(\vec{x})$ (T bezeichnet die kinetische Energie). Quantenmechanisch ist sie gegeben durch den „Hamilton-Operator“:

$$H = \frac{1}{2m}|\vec{P}|^2 + V(\vec{Q}) = -\frac{\hbar^2}{2m}\Delta + V(\vec{x})\,.$$

(Der Operator Q heißt in der Ortsdarstellung — ich erkläre hier im Eindimensionalen —: Multiplikation der Wellenfunktion mit x, Q^2 bedeutet natürlich Multiplikation mit x^2; dementsprechend: $V(Q)\ :\quad \psi(x) \mapsto (V(Q)\psi)(x) = V(x)\psi(x)$. Daher die Kurzschreibweise in der obigen Gleichung.) Der Drehimpuls ist klassisch gleich $\vec{x} \times m\vec{v} = \vec{x} \times \vec{p}$, quantenmechanisch:

$$\vec{L} = \vec{Q} \times \vec{P} = \frac{\hbar}{i}\vec{x} \times \vec{\partial}_x\,.$$

(Wenn Sie $|\vec{L}|^2$ berechnen, erhalten Sie bis auf einen Faktor $\hbar^2$ den Differentialoperator L^2, der auftritt, wenn man den Laplace-Operator in Polarkoordinaten schreibt.)
Es gibt quantenmechanische Observable, die kein klassisches Gegenstück haben, Standardbeispiel: der Spin des Elektrons. Um ihn zu beschreiben, muß man den $\mathcal{L}^2(\mathbb{R}^3)$ erweitern; und in diesem erweiterten Hilbertraum wird auch dem Spin ein geeigneter Operator zugeordnet.

Ich habe früher gezeigt, daß alle separablen Hilberträume zueinander isomorph sind. Das nutzt man in der Quantenmechanik aus, man wählt eine geeignete Darstellung, sei es die schon genannte Ortsdarstellung, die Impulsdarstellung mit der Fourier-transformierten Wellenfunktion oder noch andere. In allen Darstellungen haben wir die gleiche abstrakte Struktur:

- Quantenmechanische Zustände werden (bis auf die Normierung) beschrieben durch Vektoren in einem separablen Hilbertraum.
- Quantenmechanische Observable sind Operatoren in diesem Hilbertraum.
- Der Erwartungswert einer Observablen A in dem durch ψ beschriebenen Zustand ist
$$\langle A\rangle_\psi = \frac{(\psi, A\psi)}{(\psi,\psi)}\,. \tag{1.4}$$

Nicht alle Operatoren des Hilbertraums können als Observable interpretiert werden, sondern nur die „selbstadjungierten" Operatoren, die Definition folgt in Kürze. Aber auch andere Operatoren treten in der Quantenmechanik auf, auch wenn sie nicht Observable repräsentieren.

Eine statistische Aussage wie die über den Mittelwert $\langle x\rangle$ einer experimentell gemessenen Größe x gibt nur eine Teilinformation. Was man zudem wissen möchte, ist, wie sehr die Einzelmessungen um diesen Mittelwert herum streuen. $x-\langle x\rangle$ ist selbst eine statistisch schwankende Größe; doch ihr Mittelwert ist kein geeignetes Maß für die Schwankung, denn $\langle x-\langle x\rangle\rangle$ ist $=0$, weil die Mittelwertbildung eine lineare Operation und der Mittelwert einer Konstanten gleich der Konstanten ist. Besser folgt man dem Rezept von Gauß und definiert die mittlere quadratische Abweichung $\Delta x := \langle (x-\langle x\rangle)^2\rangle^{1/2} = (\langle x^2\rangle - \langle x\rangle^2)^{1/2}$. (Die zweite Gleichheit folgt aus der Linearität: Ü.)
Entsprechendes macht man in der Quantenmechanik, man definiert die „Unschärfe" einer Observablen A im Zustand, der durch ψ gegeben ist, als

$$\Delta_\psi A := \left(\langle A^2\rangle_\psi - \langle A\rangle_\psi^2\right)^{1/2} . \tag{1.5}$$

(Diese Begriffsbildung ist unabhängig von der Art der Mittelung, sei es bei endlich vielen Meßwerten durch Summation: $\langle x\rangle = (1/N)\sum_1^N x_i$, oder durch Integration wie in der Statistischen Mechanik: $\langle A(x,p)\rangle = \int A(x,p)\rho(x,p)\,dxdp$, oder wie in der Quantenmechanik durch Skalarproduktbildung.)

In der Statistischen Mechanik liegt die Unschärfe einer Observablen darin begründet, daß man aus begreiflichen praktischen Gründen bei 10^{20} Teilchen keine vollständige Information über den Zustand haben kann. Im Gegensatz dazu kann man solche Unschärfen in der Quantenmechanik prinzipiell nicht vermeiden; das drückt sich in der „Heisenbergschen Unschärferelation" aus, die besagt, daß das Produkt der Unschärfen von Ort und Impuls immer $\neq 0$ ist. Das hat seinen Grund letztlich darin, daß die Observablen nicht mehr alle miteinander vertauschbar sind. Nicht-kommutierende Größen kennen Sie ja schon aus dem $\mathbb{R}^n$: für Matrizen gilt im allgemeinen $AB \neq BA$, es gilt erst recht für Operatoren im Hilbertraum. Standardbeispiel:

$$PQ - QP = \frac{\hbar}{i}\mathbb{1}, \tag{1.6}$$

wobei wie oben $P = \frac{\hbar}{i}\partial_x$ und Q die Multiplikation mit x ist: der erste Term angewandt auf eine Funktion $\psi(x)$ meint ja $\frac{\hbar}{i}\partial_x(x\psi(x)) = \frac{\hbar}{i}(\psi(x) + x\psi'(x))$.
Man bezeichnet den „Kommutator" mit eckigen Klammern:

$$[A,B] \equiv AB - BA\,.$$

Wir werden uns also im Folgenden allgemein mit Operatoren im Hilbertraum zu befassen haben und insbesondere mit den selbstadjungierten Operatoren und ihrer Struktur (dazu liefert das Spektraltheorem die wesentliche Information); und dann interessieren uns Erwartungswerte $\langle A \rangle$ und Unschärfen ΔA solcher Operatoren. Schließlich will ich auch noch etwas zum Begriff des Tensorprodukts von Hilberträumen sagen, das man braucht, wenn man den Hilbertraum erweitern will, sei es zur Beschreibung mehrerer Teilchen oder zur Beschreibung des Spins.

4.2 Lineare Operatoren: Beispiele und Definitionen

Ich möchte zunächst an die Definition des Hilbertraums $\mathcal{H}$ erinnern: es ist ein linearer Raum über $\mathbb{C}$, in dem ein positiv-definites Skalarprodukt definiert ist, d. h. eine Abbildung $\varphi, \psi \mapsto (\varphi, \psi) \in \mathbb{C}$, die linear im zweiten und antilinear im ersten Faktor ist, für die $(\varphi, \psi) = \overline{(\psi, \varphi)}$ gilt und $(\varphi, \varphi) \geq 0, = 0$ nur für $\varphi = 0$; durch $\|\varphi\| := (\varphi, \varphi)^{1/2}$ ist eine Norm definiert, die einen Konvergenzbegriff liefert, bezüglich dessen der Raum vollständig ist: alle Cauchy-Folgen haben einen Limes im Raum.
Durch das Schmidtsche Orthogonalisierungsverfahren kann man eine o. n. Basis definieren, d. h. eine Menge $\mathcal{B} = \{\varphi_\alpha\}$ von Vektoren, für die $(\varphi_\alpha, \varphi_\beta) = \delta_{\alpha\beta}$ gilt und deren lineare Hülle in $\mathcal{H}$ dicht ist. Die Anzahl der Basiselemente ist $= \dim \mathcal{H}$. Ich erinnere daran, daß die lineare Hülle durch die *endlichen* Linearkombinationen gegeben ist; man beachte, daß in einem unendlich-dimensionalen Raum L.H.$\mathcal{B} \neq \mathcal{H}$.

Wenn ich in diesem Kapitel von „Hilbertraum“ rede, meine ich immer, ohne es explizit zu sagen, einen *separablen* Hilbertraum, d. h. einen Raum von endlicher oder abzählbar-unendlicher Dimension. Eine ganze Reihe von Aussagen, die wir in den folgenden Abschnitten beweisen wollen, gelten auch in beliebigen Hilberträumen.

Ich beginne nun mit einer allgemeinen Definition.

Definition 4.2.1 *$\mathcal{X}$ sei ein Vektorraum über $\mathbb{C}$. Ein linearer Operator in $\mathcal{X}$ ist gegeben durch eine lineare Mannigfaltigkeit $\mathcal{D} \subset \mathcal{X}$ und eine lineare Abbildung $\mathcal{D} \longrightarrow \mathcal{X}$. $\mathcal{D}$ heißt der Definitionsbereich des Operators.*

Man kann die Definition noch allgemeiner fassen, indem man Abbildungen zuläßt, die von $\mathcal{D} \subset \mathcal{X}$ in einen zweiten linearen Raum $\mathcal{Y}$ führen; doch das ist nicht das Thema dieses Kapitels.
Ich werde oft von Operatoren sprechen und fast immer *lineare* Operatoren meinen, ohne das immer dazuzusetzen. Ich weise ausdrücklich darauf hin, wenn einmal ein nicht-linearer Operator vorkommen sollte.
Beachten Sie bitte: *Zur vollständigen Definition eines Operators gehört immer die Angabe seines Definitionsbereichs.* Oft genug gibt es in der Quantenmechanik „natürliche“ Definitionsbereiche für die auftretenden Operatoren, dann werden sie in der Definition meist nicht angegeben; doch bei einer strengen Behandlung ist es nötig.
Wenn zwei Operatoren auf Bereichen $\mathcal{D}_1$ und $\mathcal{D}_2$ definiert sind, $\mathcal{D}_1 \subset \mathcal{D}_2$, und auf dem kleineren Bereich übereinstimmen, dann spricht man von einer Erweiterung des ersten Operators durch den zweiten.

Bevor ich irgend etwas zur obigen Definition kommentiere, will ich eine Reihe von Beispielen anführen, an denen ich die obige und weitere Definitionen illustrieren will.

Anstelle eines allgemeinen Vektorraumes $\mathcal{X}$ werde ich allerdings immer einen Hilbertraum $\mathcal{H}$ wählen.

4.2.2 Beispiel: Sei $\mathcal{H} = \mathbb{C}^n$. Jede $n \times n$-Matrix A mit komplexen Matrixelementen definiert eine lineare Abbildung von $\mathbb{C}^n$ in sich, d. h. einen linearen Operator mit $\mathcal{H}$ als Definitionsbereich.

4.2.3 Beispiele: Sei $\mathcal{H} = \ell^2(\mathbb{C})$ = Menge der quadratsummablen Folgen.

(1) $\mathcal{D}_{A_1} = \mathcal{H}$, und $A_1 : \ (a_1, a_2, \ldots) \mapsto (a_1, \frac{1}{2}a_2, \frac{1}{3}a_3, \ldots)$.

(2) $\mathcal{D}_{A_2}$ = Menge der abbrechenden Folgen (d. h. diejenigen, bei denen alle $a_n = 0$ für hinreichend große n), und $A_2 : \ (a_1, a_2, \ldots) \mapsto (a_1, 2a_2, 3a_3, \ldots)$.

In gewisser Weise ist A_2 das Inverse von A_1: es gilt offensichtlich $A_1A_2 = \mathbb{1}\big|_{\mathcal{D}_{A_2}}$, aber für das Inverse soll auch $A_1^{-1}A_1 = \mathbb{1}$ gelten. Doch die Anwendung von A_1 auf eine beliebige Folge liefert nicht unbedingt ein Element aus dem Definitionsbereich von A_2, so daß man A_2 nicht immer anwenden kann. Der Ausweg ist naheliegend: man muß den Definitionsbereich von A_2 erweitern. Er kann nicht ganz $\mathcal{H}$ sein, denn die Anwendung des Operators darf nicht aus $\mathcal{H}$ herausführen; doch kann man leicht einen „natürlichen“ Definitionsbereich angeben: $\mathcal{D}_{A_3} = \{\{a_k\} \in \ell^2;\ \sum |ka_k|^2 < \infty\}$, $A_3\,\{a_k\} = \{ka_k\}$. Dann ist offensichtlich $A_1A_3 = \mathbb{1}\big|_{\mathcal{D}_{A_3}}$ und weiterhin $A_1\mathcal{H} \subset \mathcal{D}_{A_3}$, genauer: $A_1\mathcal{H} = \mathcal{D}_{A_3}$; also ist A_3 auf alle Vektoren aus $A_1\mathcal{H}$ anwendbar und $A_3A_1 = \mathbb{1}$. Das sind typische Komplikationen, die im unendlich-dimensionalen Raum auftreten.
(Es ist oft angebracht, die Einschränkung oder die Erweiterung eines Operators vom ursprünglichen Operator strikt zu unterscheiden!)

4.2.4 Beispiele: In einem separablen Hilbertraum $\mathcal{H}$ mit o. n. Basis $\{e_n;\ n \in \mathbb{N}\}$ ist ein linearer Operator A mit Definitionsbereich $\mathcal{D}$ = L.H. $\{e_n\}$ vollständig festgelegt durch die Angabe seiner Wirkung auf die Basiselemente e_n, denn dann folgt $A\sum\gamma_ne_n = \sum\gamma_nAe_n$ wegen der Linearität. (Die Summen sind ja endlich, das ist wichtig! Aus der Linearitätsrelation $A(a\varphi + \psi) = aA\varphi + A\psi$ kann man induktiv nur auf die Vertauschbarkeit von A mit *endlichen* Summen schließen.)

(1) $V_1 : \ e_n \mapsto e_{n+1},\ n \in \mathbb{N}$.

(2) $V_2 : \ e_1 \mapsto 0,\ e_{n+1} \mapsto e_n,\ n > 0$.

Die Definitionsbereiche sind jeweils L.H. $\{e_n\}$; für den folgenden Operator können wir $\mathcal{D} = \mathcal{H}$ wählen:

(3) $P\psi := \sum_1^N (e_n, \psi)e_n$ ist linear, weil das Skalarprodukt linear im zweiten Faktor ist.

4.2.5 Beispiele: Sei $\mathcal{H} = \mathcal{L}^2(\mathbb{R})$. Für die ersten beiden Operatoren sei $\mathcal{D} = \mathcal{S}(\mathbb{R})$ ($\mathcal{D}$ meint hier natürlich den Definitionsbereich, nicht die Funktionen mit kompaktem Träger.)

(1) $Q_1 : \psi(x) \mapsto x\psi(x)$.

(2) $P_1 : \psi(x) \mapsto \frac{\hbar}{i}\partial_x\psi(x)$.

Für die folgenden Operatoren ist $\mathcal{D} = \mathcal{H}$:

(3) Sei $G(x, x')$ in beiden Variablen quadratintegrabel,

$$(A\psi)(x) := \int G(x, x')\psi(x')\, dx' .$$

(4) $U_t : \psi \mapsto \mathcal{F}^{-1}\left(e^{-i\hbar t k^2/(2m)}(\mathcal{F}\psi)(k)\right)$.

Das letzte Beispiel ist Ihnen bekannt: es beschreibt die Zeitentwicklung einer Wellenfunktion gemäß der freien Schrödingergleichung, siehe unser früheres Anwendungsbeispiel in Abschnitt 2.3.

Bei allen angegebenen Beispielen können Sie relativ leicht überprüfen, daß die Anwendung der Operatoren auf die Vektoren aus den angegebenen Definitionsbereichen wohldefiniert ist *und wieder Vektoren aus $\mathcal{H}$ liefert.* Wie schon erwähnt — und darauf werde ich noch eingehen —, kann man nicht alle diese Operatoren auf allen Vektoren aus $\mathcal{H}$ definieren; immerhin sind alle angegebenen Definitionsbereiche dicht in $\mathcal{H}$, und das ist es, was wir meist annehmen.

Mithilfe der in $\mathcal{H}$ definierten Norm können wir auch eine Operator-Norm definieren:

Definition 4.2.6 *Ein Operator A heißt* **beschränkt**, *wenn für eine passende Konstante*

$$||A\psi|| \leq \text{const}\, ||\psi|| \quad \forall\, \psi \in \mathcal{D}(A)$$

gilt. Die Größe

$$||A|| := \sup_{\psi\in\mathcal{D}(A),\psi\neq 0} \frac{||A\psi||}{||\psi||}$$

heißt die Norm von A.

Diese Definition ist ähnlich derjenigen für die Norm eines linearen Funktionals über einem normierten Raum, siehe Lemma 1.1.20. Es ist sofort zu sehen, daß alle Eigenschaften aus Defintion 1.1.3 erfüllt sind, die eine Norm haben muß.
Man beachte die triviale Folgerung: Wenn A beschränkt ist, gilt

$$||A\psi|| \leq ||A||\; ||\psi|| \quad \forall\, \psi \in \mathcal{D}_A . \tag{2.7}$$

Sehen wir uns die Beispiele 4.2.3 an: sie haben die Form $\{a_n\} \mapsto \{\gamma_n a_n\}$; und wenn die Folge $\{\gamma_n\}$ beschränkt ist, wie es im ersten Beispiel der Fall ist, folgt $\sum |\gamma_n a_n|^2 \leq \sup_k |\gamma_k|^2 \, \|\{a_n\}\|^2$, dann ist der Operator beschränkt durch $\sup |\gamma_k|$. Für A_1 ist $\gamma_n = 1/n$ und $\|A_1\| = \sup |\gamma_n| = 1$, denn für $n = 1$ wird das Supremum erreicht. Man sieht aus (2.7) auch, daß A_1 tatsächlich auf allen Folgen $\in \ell^2$ wohldefiniert ist und wieder quadratintegrable Folgen ergibt.
Der Operator A_2 mit $\gamma_n = n$ ist offensichtlich unbeschränkt: man wähle $a_n = \delta_{kn}$, dann ist $\{a_n\} \in \mathcal{D}_{A_2}$, $\|\{a_n\}\| = 1$ und $\|A\,\{a_n\}\| = k$; und das kann beliebig groß werden.

Die beiden Operatoren Q_1 und P_1 (siehe Beispiele 4.2.5) sind ebenfalls unbeschränkt: sei $\psi \in \mathcal{S}$, $\psi(x) > 0$, $\int |\psi|^2\, dx = 1$, dann gilt das gleiche für $\psi_y(x) := \psi(x-y)$; doch $\|Q_1\psi_y\|$ wird mit wachsendem y beliebig groß: $\|Q_1\psi_y\|^2 = \int (\xi+y)^2 |\psi(\xi|^2)\, d\xi \underset{y\to\infty}{\longrightarrow} \infty$. (Man sieht auch trivialerweise, daß auch wenn $\psi(x)$ quadratintegrabel ist, $\int |x\psi(x)|^2\, dx$ für $|x| \to \infty$ sehr wohl divergieren kann; die *Funktion* $(Q_1\psi)(x)$ ist zwar (als meßbare Funktion) wohldefiniert, aber nicht $\in \mathcal{L}^2(\mathbb{R})$; das zeigt direkt, daß Q_1 sich nicht auf ganz $\mathcal{H}$ fortsetzen läßt.) Entsprechend überlegt man sich, daß P_1 unbeschränkt ist: die $\mathcal{L}^2$-Norm der Ableitung einer auf 1 normierten Funktion aus $\mathcal{S}$ kann beliebig groß werden. Anders: P_1 sieht in der Impulsraumdarstellung aus wie Q_1 in der Ortsdarstellung, man kann dieselben Argumente verwenden.

Man überzeugt sich auch leicht, daß die Operatoren V_1, V_2 und P (Beispiele 4.2.4) beschränkte Operatoren sind (Ü: man berechne ihre Normen).
V_1 und V_2 sind nicht auf ganz $\mathcal{H}$ definiert, doch kann man sie stetig auf ganz $\mathcal{H}$ fortsetzen; das sagt das folgende Lemma:

Lemma 4.2.8 *Ein linearer beschränkter Operator A definiert eine stetige Abbildung von $\mathcal{D}_A$ in $\mathcal{H}$; jeder stetige lineare Operator ist beschränkt.*
Aus der Stetigkeit bei 0 folgt die Stetigkeit überall.

Der Beweis ist nicht sehr schwer, vergl. die Bemerkung nach Korollar 1.1.11.

Wenn nun A beschränkt ist und einen dichten Definitionsbereich $\mathcal{D}_A$ hat, der aber $\neq \mathcal{H}$ ist, können wir durch

$$A \lim \psi_\alpha := \lim A\psi_\alpha, \qquad \psi_\alpha \in \mathcal{D}_A\,, \tag{2.9}$$

den Operator fortsetzen: wenn $\lim \psi_\alpha$ existiert, dann wegen der Stetigkeit von A auch die rechte Seite, und $\lim \psi_\alpha = 0$ hat nach (2.7) $\lim A\psi_\alpha = 0$ zur Folge. (Das zu kontrollieren ist wichtig, sonst wäre die Erweiterung nicht linear; denn A muß den Null-Vektor auf sich abbilden.)

Die meisten Operatoren, die einem in der Quantenmechanik begegnen, sind unbeschränkt, d. h. sie sind nach dem letzten Satz von Lemma 4.2.8, der für jeden anderen Punkt als 0 natürlich ebenfalls gilt, überall unstetig. (Man beachte: die Ableitung, und damit P_1, ist auf $\mathcal{S}(\mathbb{R})$ stetig in der $\mathcal{S}$-Topologie, aber nicht in der $\mathcal{L}^2$-Topologie: wenn eine Folge von Funktionen $f_\alpha \in \mathcal{S}$ bezüglich der $\mathcal{L}^2$-Norm konvergiert, muß das noch lange nicht für die Ableitungen der f_α gelten.)[1]

Wichtige Bemerkung: Bei einem unbeschränkten Operator garantiert die Existenz von $\lim \psi_\alpha$ keineswegs die Existenz von $\lim A\psi_\alpha$, sie schließt sie aber auch nicht aus: wenn beide Limiten existieren und wenn A auf $\lim \psi_\alpha$ noch nicht definiert ist, kann man versuchen, A wie oben zu erweitern. Man kann also auch unbeschränkte Operatoren unter gewissen Umständen auf einen größeren Bereich fortsetzen. (Aber Vorsicht: $\lim \psi_\alpha = 0$ muß $\lim A\psi_\alpha = 0$ zur Folge haben!)[2]

Die obige Konstruktion bringt mich auf eine Eigenschaft, die für unstetige Operatoren ein gewisser schwacher, aber wichtiger Ersatz der Stetigkeit ist:

Definition 4.2.10 *Sei A ein linearer Operator mit folgender Eigenschaft:*

(C) Wenn für $\psi_\alpha \in \mathcal{D}_A$ sowohl $\lim \psi_\alpha \equiv \psi$ als auch $\lim A\psi_\alpha \equiv \varphi$ existieren, dann gilt $\psi \in \mathcal{D}_A$ und $A\psi = \varphi$.

Ein OperatorA mit der Eigenschaft (C) heißt **abgeschlossen**.

Man kann sich also vorstellen, daß der Operator nach dem Schema von (2.9) fortgesetzt ist, wo immer es geht. (Das heißt nicht, daß er bei ψ stetig ist: es kann andere Folgen ψ'_α geben, die gegen ψ konvergieren, für die aber $A\psi'_\alpha$ nicht konvergent ist.)
Diese Definition kann man sich veranschaulichen (und gut merken) mit dem Konzept des Graphen eines Operators. Der Graph einer Funktion $y = f(x)$ ist die Menge der Punkte $(x, f(x)) \in \mathbb{R}^2$. Analog hier: statt $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ betrachtet man $\mathcal{H} \times \mathcal{H}$, d. h. Paare $(\varphi; \psi)$ von Hilbertraumvektoren (ich habe ein Semikolon gesetzt, denn hier ist natürlich nicht das Skalarprodukt gemeint), und definiert

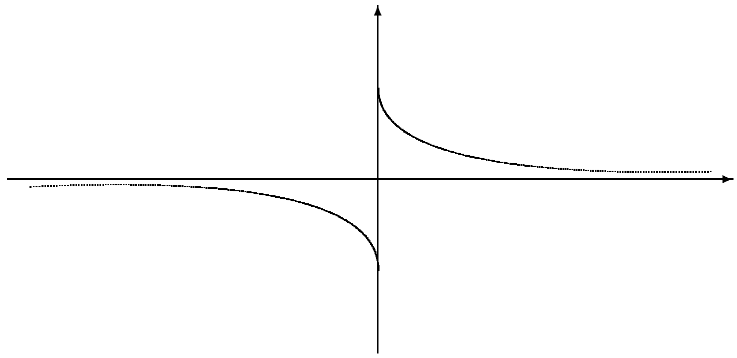
$$G(A) \equiv \text{ Graph von } A := \{(\psi; A\psi);\ \psi \in \mathcal{D}_A\} \subset \mathcal{H} \times \mathcal{H}\,. \tag{2.11}$$

Auf dem kartesischen Produkt $\mathcal{H} \times \mathcal{H}$ führt man die Produkttopologie ein: Konvergenz heißt: Norm-Konvergenz beider Komponenten; anders: $\|(\varphi; \psi)\| = \|\varphi\| + \|\psi\|$. Man sieht nun, daß ein Operator genau dann abgeschlossen ist, wenn sein Graph eine

[1]Unbeschränkte, und damit unstetige Operatoren gibt es natürlich nur, wenn $\dim \mathcal{H} = \infty$; dann gibt es an jedem Punkt $\varphi_0 \in \mathcal{H}$ unendlich viele orthogonale Richtungen, so daß man sich vorstellen kann, daß an jedem Punkt $A\varphi$ in einer oder mehrerer dieser Richtungen unstetig ist.

[2]Für einen linearen Operator A gilt $A \sum \psi_n = \sum A\psi_n$ für *endliche* Summen; das gilt, wenn A unbeschränkt ist, im allgemeinen aber nicht für unendliche Summen, denn unbeschränkte Operatoren sind überall unstetig.

in $\mathcal{H} \times \mathcal{H}$ abgeschlossene Menge ist. Die Tatsache, daß A linear ist, kommt darin zum Ausdruck, daß $G(A)$ eine lineare Mannigfaltigkeit ist.
Man kann natürlich jede Untermenge von $\mathcal{H} \times \mathcal{H}$ abschließen, auch jeden Graphen. Aber der Abschluß $\overline{G(A)}$ muß nicht mehr Graph eines Operators sein: dazu ist notwendig (und hinreichend), daß jedem Element φ der Abszisse, das im Graphen auftritt, genau ein Element ψ der Ordinate zugeordnet ist, so daß man $\psi = A\varphi$ definieren kann.
Schwache Analogie im $\mathbb{R} \times \mathbb{R}$ (hier kann ich allerdings nur ein Beispiel mit einer nichtlinearen Funktion konstruieren): man nehme $f(x) = \text{arc}\tan(1/x)$ mit Definitionsbereich $\mathbb{R} \setminus \{0\}$, die Funktion ist zwar beschränkt bei 0, läßt sich aber nicht eindeutig bei 0 definieren.



Nicht jeder Operator ist also entweder abgeschlossen oder abschließbar. Wenn $\overline{G(A)}$ wieder Graph eines Operators ist, heißt dieser Operator der Abschluß von A, er wird mit $\overline{A}$ bezeichnet. In der Quantenmechanik treten im allgemeinen keine nicht abschließbaren Operatoren auf; sie haben für den Physiker pathologische Eigenschaften.

Wir haben gesehen, daß man jeden beschränkten Operator auf ganz $\mathcal{H}$ fortsetzen kann; die umgekehrte Aussage ist nicht richtig: es gibt Operatoren, die auf ganz $\mathcal{H}$ definiert, aber nicht beschränkt sind. Immerhin gilt (ohne Beweis)

Satz 4.2.12 *Sei A abgeschlossen und $\mathcal{D}_A = \mathcal{H}$, dann ist A beschränkt.*

4.2.13 Anmerkung: Wegen der Fortsetzbarkeit der beschränkten Operatoren werden wir im Folgenden fast immer stillschweigend annehmen, daß ihr Definitionsbereich $= \mathcal{H}$ ist. Wenn ein Operator abschließbar ist, nehmen wir auch oft stillschweigend an, daß er abgeschlossen sei. Für die wichtige Klasse der selbstadjungierten Operatoren gibt es einen „natürlichen" Definitionsbereich; auch für sie werden wir nicht immer angeben, wo sie definiert sind.
In der Literatur werden oft Operatoren definiert, ohne daß ein Wort über ihren Definitionsbereich gesagt wird; dann ist entweder gemeint: sie sind auf ganz $\mathcal{H}$ definiert, oder: es gibt einen „natürlichen" Definitionsbereich, oder auch: der Leser wird sich schon was dabei denken; letzteres ist in manchen Physikbüchern üblich.
Ich werde auch nicht immer $\mathcal{D}(A)$ angeben, werde mich aber bemühen, es nur in den

ersten beiden Fällen zu tun.
Ich werde auch fast immer stillschweigend voraussetzen, daß die betrachteten Operatoren auf einer dichten Menge definiert sind.

Man kann einen Operator, der nicht auf ganz $\mathcal{H}$ definiert ist, immer auf einen größeren Bereich erweitern zu einem Operator $\widehat{A}$; es genügt ja, für $\psi \notin \mathcal{D}_A$ einfach $\widehat{A}\psi := \widehat{\psi}$ mit beliebigem $\widehat{\psi}$ zu setzen und linear zu erweitern: $\widehat{A}(\varphi + c\psi) = A\varphi + c\widehat{\psi}\ \forall \varphi \in \mathcal{D}_A$. Dann haben wir wieder einen linearen Operator, der aber im allgemeinen keine schönen Eigenschaften hat; von einem „natürlichen" Definitionsbereich kann dann in den meisten Fällen nicht mehr die Rede sein. Für die Physik ist ein solcher Operator meist unbrauchbar. Manchmal ist es jedoch sinnvoll, Erweiterungen vorzunehmen, auch wenn sie willkürbehaftet sind; ich erwähne später ein Beispiel. (Wichtig ist, daß man bei allen Definitionen oder Erweiterungen von linearen Operatoren darauf achtet, daß man nicht implizit eine Definition $A\psi \neq 0$ trifft, wenn $\psi = 0$ ist; das kann einem manchmal in der Form unterlaufen, daß $\psi_\alpha \to 0$, aber $A \lim \psi_\alpha \neq 0$.) Ich füge eine Bemerkung an, die eigentlich trivial ist:

4.2.14 Anmerkung: Ein Operator ist — bei gegebenem dichtem Definitionsbereich — vollständig durch die Angabe von $\{(\varphi, A\varphi);\ \varphi \in \mathcal{D}_A\}$ festgelegt. Wenn dieser Operator also eine quantenmechanische Observable repräsentiert, genügt es, seine Erwartungswerte zu messen, um ihn vollständig zu kennen. Die Polarisationsidentität (I.1.17) läßt sich leicht etwas verallgemeinern:

$$(\psi, A\varphi) = \frac{1}{4}\sum_{\nu=0}^{3} i^\nu (\psi + (-i)^\nu \varphi, A(\psi + (-i)^\nu \varphi))\,. \tag{2.15}$$

Das folgt ebenfalls durch einfaches Ausrechnen, wesentlich ist die Linearität von A. Dann sind also durch die obige Angabe alle Matrixelemente $(\psi, A\varphi)$ für $\psi, \varphi \in \mathcal{D}_A$ bekannt (hier ist die Linearität von $\mathcal{D}_A$ wichtig); und diese Angaben legen $A\varphi$ für alle $\varphi \in \mathcal{D}_A$ fest, wenn $\mathcal{D}_A$ dicht ist:
Zu jedem $\chi \in \mathcal{H}$ gibt es eine Folge $\psi_\alpha \in \mathcal{D}_A$, für die $\lim \psi_\alpha = \chi$ ist; wegen der Stetigkeit des Skalarprodukts gilt dann $\lim(\psi_\alpha, A\varphi) = (\chi, A\varphi)$; und durch Angabe von $(\chi, A\varphi)$ für alle $\chi \in \mathcal{H}$ ist $A\varphi$ festgelegt.[3]

Wenn man nur beschränkte Operatoren betrachtet, ist es klar, daß man die Summe $A + B$ und das Produkt AB von Operatoren A und B definieren kann; man setzt $\mathcal{D}_A = \mathcal{D}_B = \mathcal{H}$ voraus, dann ist $B\psi$ wieder in $\mathcal{D}_A = \mathcal{H}$ und $AB\psi$ ist wohldefiniert. Außerdem kann man natürlich jeden Operator mit einer komplexen Zahl multiplizieren. Vornehm ausgedrückt: die beschränkten Operatoren des Hilbertraums bilden eine

[3] Aus $(\chi, A\varphi) = (\chi, \psi)\ \forall\, \chi \in \mathcal{H}$ folgt mit $\chi = A\varphi - \psi$: $\|A\varphi - \psi\| = 0$, also $A\varphi = \psi$.

normierte Algebra; sie ist vollständig (dazu später etwas mehr), es ist also ein Banachraum, mehr noch: eine „Banach-Algebra“, denn für die Norm eines Produktes gilt die folgende Ungleichung:

$$\|AB\| \leq \|A\| \, \|B\| \, . \tag{2.16}$$

Beweis: Ü.
Die Algebra der beschränkten Operatoren ist aber „nicht-kommutativ“: die Produkte AB und BA stimmen nicht überein. Das ist Ihnen ja schon von Matrizen, d. h. Operatoren im $\mathbb{C}^n$, bekannt. Außerdem kann alles Mögliche passieren, was man in der Algebra der komplexen Zahlen nicht gewohnt ist: möglich ist $A^2 = 0$, ohne daß $A = 0$ ist, oder $A^2 = \mathbb{1}$ oder $A^2 = A$, ohne daß $A = \pm\mathbb{1}$ bzw. $A = 0$ oder $= \mathbb{1}$. Auch dafür finden sich schon Matrixbeispiele.
Für P aus Beispiel 4.2.4(3) gilt $P^2 = P$.

Wenn man es mit unbeschränkten Operatoren zu tun hat, ist nur die Multiplikation mit einer komplexen Zahl ist uneingeschränkt möglich. Schon bei der Summe kann es Schwierigkeiten geben: auch wenn $\mathcal{D}_A$ und $\mathcal{D}_B$ dicht in $\mathcal{H}$ sind, kann $\mathcal{D}_A \cap \mathcal{D}_B = \{0\}$ sein. Erst recht ist es möglich, daß $B\mathcal{D}_B \not\subset \mathcal{D}_A$ gilt; und dann ergibt AB keinen Sinn. Doch oft gibt es für die Operatoren eingeschränkte Definitionsbereiche, genauer: eine dichte Menge $\mathcal{D}_0 \subset \mathcal{D}_A \cap \mathcal{D}_B$, so daß $B\mathcal{D}_0 \subset \mathcal{D}_0$ ist; dann kann man auf $\mathcal{D}_0$ Summe und Produkt bilden. Typisches Beispiel: die Operatoren Q_1 und P_1 aus den Beispielen 4.2.5 kann man zu selbstadjungierten Operatoren Q und P erweitern, sie sind in der Quantenmechanik wichtig als Observable von Ort und Impuls. Für sie kann man $\mathcal{D}_0 = \mathcal{S}(\mathbb{R})$ als gemeinsamen Definitionsbereich wählen; dieser Bereich ist invariant unter beiden Operatoren.

Auf diesem Bereich gilt erwähnte Relation $[P, Q] \equiv PQ - QP = \frac{\hbar}{i}\mathbb{1}$, die „kanonische Vertauschungsrelation“, die für die Quantenmechanik fundamental ist und die zur Heisenbergschen Unschärferelation führt.

4.3 Selbstadjungierte und unitäre Operatoren

In der Quantenmechanik möchte man die Erwartungswerte $\langle A \rangle$ von Observablen den Werten (besser: den Mittelwerten) gleichsetzen, die der Experimentator mißt; sie müssen also reelle Größen sein. Für einen beliebigen Operator ist aber $(\psi, A\psi)$ im allgemeinen keineswegs reell: $\overline{(\psi, A\psi)} = (A\psi, \psi)$, und das ist nur für bestimmte Operatoren gleich $(\psi, A\psi)$.

Definition 4.3.1 *Ein Operator, für den*

$$(\psi, A\psi) = (A\psi, \psi) \quad \forall\, \psi \in \mathcal{D}_A$$

gilt, heißt **hermitisch**.

Triviale Folgerung mit der verallgemeinerten Polarisationsidentität (2.15): für hermitische Operatoren gilt

$$(\psi, A\varphi) = (A\psi, \varphi) \quad \forall\, \psi, \varphi \in \mathcal{D}_A\,. \tag{3.2}$$

Es gibt einen weiteren Begriff, den ich von der Hermitizität unterscheide, obwohl diese Unterscheidung nicht in allen Büchern gemacht wird:

Definition 4.3.3 *Ein hermitischer Operator mit dichtem Definitionsbereich heißt* **symmetrisch**.

Quantenmechanische Observable sind durch symmetrische Operatoren gegeben; so ist garantiert, daß die Observablen auf einer hinreichend großen Menge von Zuständen gemessen werden können und daß die Erwartungswerte reell sind.

Man kann einen hermitischen Operator also von einer Seite des Skalarprodukts auf die andere bringen. Um eine solche Möglichkeit für alle Operatoren zu haben, braucht man einen (sehr wichtigen!) Begriff, der die komplexe Konjugation verallgemeinert: die Adjunktion. (Wir haben Ähnliches schon bei Differentialoperatoren kennengelernt.) Dazu muß ich an den Satz von Riesz (Theorem 1.3.11) erinnern, der besagt, daß jedes stetige Funktional auf $\mathcal{H}$ durch das Skalarprodukt mit einem passenden Vektor gegeben ist. Sei nun A beschränkt, d. h. stetig; dann ist $\Phi_{A,\psi} : \varphi \in \mathcal{H} \mapsto (\psi, A\varphi)$ eine stetige Abbildung. Es gibt also einen eindeutig bestimmten Vektor χ, der von A und ψ abhängt, so daß

$$(\psi, A\varphi) = (\chi, \varphi) \tag{3.4}$$

ist. Nach dem Satz von Riesz hängt χ antilinear von Φ ab, und dieses wiederum antilinear von ψ (das ist offensichtlich), also ist χ linear von ψ abhängig, die Abbildung

$\psi \mapsto \chi$ definiert also einen linearen Operator A^*: $\chi = A^*\psi$. Das gilt für alle ψ, also ist A^* auf ganz $\mathcal{H}$ definiert. Folglich:

$$(\psi, A\varphi) = (A^*\psi, \varphi)\,. \tag{3.5}$$

A^* heißt der zu A adjungierte Operator.

Diese Überlegungen hängen nur davon ab, daß $\varphi \in \mathcal{D}_A$ und daß $\Phi_{A,\psi}(\varphi) = (\psi, A\varphi)$ stetig in φ ist; sie gelten also auch für möglicherweise unbeschränkte Operatoren. Man braucht allerdings die Voraussetzung, daß $\mathcal{D}_A$ dicht ist: dann kann man ein stetiges $\Phi_{A,\psi}$ auf ganz $\mathcal{H}$ fortsetzen und den Satz von Riesz anwenden. Damit kommen wir zu der sehr wichtigen

Definition 4.3.6 *A habe einen dichten Definitionsbereich $\mathcal{D}_A$. Der zu A* **adjungierte** *Operator A^* ist gegeben durch*

$$\mathcal{D}_{A^*} := \{\psi;\ \varphi \mapsto (\psi, A\varphi)\ \textit{stetig}\,\forall\varphi \in \mathcal{D}_A\}\ ,$$

$$(\psi, A\varphi) = (A^*\psi, \varphi)\ \ \forall\psi \in \mathcal{D}_{A^*},\ \varphi \in \mathcal{D}_A\,.$$

Die Voraussetzung, daß $\overline{D_A} = \mathcal{H}$, ist notwendig (und hinreichend), damit A^* durch die letzte Gleichung wohldefiniert ist.

Für unbeschränkte Operatoren ist $\varphi \mapsto A\varphi$ unstetig, die Bedingung an $\psi \in \mathcal{D}_{A^*}$ also nichttrivial. Man sieht sofort, daß $\mathcal{D}_{A^*}$ eine lineare Mannigfaltigkeit ist. Natürlich ist $0 \in \mathcal{D}_{A^*}$; es kann durchaus vorkommen, daß das das einzige Element ist: $\mathcal{D}_{A^*} = \{0\}$. Für ein symmetrisches A ist $(\psi, A\varphi) = (A\psi, \varphi)$ für $\psi \in \mathcal{D}_A$, d. h. stetig in φ; demnach ist $\psi \in \mathcal{D}_{A^*}$ und $A^*\Big|_{\mathcal{D}_A} = A$. Aber im allgemeinen ist $\mathcal{D}_{A^*}$ größer als $\mathcal{D}_A$!
Man schreibt $A \subset B$, wenn $\mathcal{D}_A \subset \mathcal{D}_B$ und $A = B$ auf $\mathcal{D}_A$; wenn ich $A = B$ schreibe, meine ich, daß auch $\mathcal{D}_A = \mathcal{D}_B$ ist. Im allgemeinen gilt also für einen symmetrischen Operator

$$A \subset A^*\,. \tag{3.7}$$

Diese Gleichung kann man auch als Definition für „hermitisch" lesen.

Die in gewisser Weise „schönsten" und in der Quantenmechanik wichtigsten Operatoren sind diejenigen, bei denen diese Unterscheidung entfällt: die selbstadjungierten. Man kann sie analog zu den symmetrischen Matrizen im $\mathbb{R}^n$ vollständig „diagonalisieren"; was das heißt, und wie das zu machen ist, ist der wichtigste Teil dieses Kapitels: die „Spektralzerlegung" selbstadjungierter Operatoren, die ich in einem späteren Abschnitt bringen will.

Definition 4.3.8 *Ein Operator A, für den $A = A^*$ gilt, heißt* **selbstadjungiert**.

Selbstadjungierte Operatoren sind also insbesonders symmetrisch. Bei beschränkten Operatoren kann man auf diese Unterscheidung verzichten, wenn man, wie üblich, $\mathcal{H}$ als Definitionsbereich nimmt. In vielen Lehrbüchern der Quantenmechanik heißen auch die selbstadjungierten Operatoren hermitisch; doch das ist mathematisch nicht sauber. Man kann symmetrische Operatoren oft zu selbstadjungierten erweitern: Vergrößerung von D_A bedeutet, daß die Anforderungen steigen, die man an ψ stellen muß, damit $\psi \in \mathcal{D}_{A^*}$, d. h. $\mathcal{D}_{A^*}$ wird kleiner, enthält aber immer noch $\mathcal{D}_A$, solange die Erweiterung noch symmetrisch ist. Bei zu starker Erweiterung wird $\mathcal{D}_{A^*}$ zu klein, die Erweiterung ist dann nicht mehr symmetrisch. Wenn alles gut geht, findet man eine symmetrische Erweiterung, für die $\mathcal{D}_A = \mathcal{D}_{A^*}$ ist, und dann ist A nach Definition selbstadjungiert. Es gibt symmetrische Operatoren, die genau eine selbstadjungierte Erweiterung haben, sie heißen **wesentlich selbstadjungiert**. Unsere Beispiel-Operatoren Q_1 und P_1 (Beispiele 4.2.5) gehören dazu. die selbstadjungierten Operatoren Q (Ortsoperator in einer Dimension) und P (Impulsoperator) sind also durch ihre Wirkung auf dem eingeschränkten Definitionsbereich $\mathcal{D}_{Q_1} = \mathcal{D}_{P_1} = \mathcal{S}(\mathbb{R})$ vollständig bestimmt; es genügt also, die kanonische Vertauschungsrelation auf diesem „Kernbereich" zu formulieren. Für wesentlich selbstadjungierte Operatoren kann man den Definitionsbereich der selbstadjungierten Erweiterung als den „natürlichen" Definitionsbereich betrachten. Andere symmetrische Operatoren können verschiedenartige selbstadjungierte Erweiterungen haben; und deshalb ist es nicht unwesentlich, die Unterscheidung zwischen symmetrisch und selbstadjungiert zu machen (Beispiele in den Übungen). Das ist auch die Begründung dafür, daß quantenmechanische Observable durch selbstadjungierte (und nicht nur symmetrische) Operatoren gegeben sind. Die oben genannte Spektralzerlegung ist nur für selbstadjungierte, nicht für symmetrische Operatoren definiert; und „verschiedenartige selbstadjungierte Erweiterungen" bedeutet „verschiedenartige Spektralzerlegungen".
Vom folgenden Satz sollte man sich zumindestens die erste Aussage merken:

Satz 4.3.9 *(i) A^* ist ein abgeschlossener Operator (Def. 4.2.10).*
(ii) Wenn A^* dicht definiert ist, ist A^{**} der (minimale) Abschluß von A.

Einfache **Folgerung:** Ein selbstadjungierter Operator abgeschlossen.

Beweis: Sei $\varphi_\alpha \in \mathcal{D}_{A^*}$ und die Limiten $\lim \varphi_\alpha =: \varphi$ und $\lim A^*\varphi_\alpha =: \psi$ mögen existieren. Dann gilt für jedes $\chi \in \mathcal{D}_A$

$$(\varphi, A\chi) = \lim(\varphi_\alpha, A\chi) = \lim(A^*\varphi_\alpha, \chi) = (\psi, \chi)\,.$$

Die rechte Seite zeigt, daß $(\varphi, A\chi)$ stetig ist in $\chi \in \mathcal{D}_A$, d. h. es gilt $\varphi \in \mathcal{D}_{A^*}$ und $A^*\varphi = \psi$, A^* ist folglich abgeschlossen.
Die Aussage (ii) ist eine einfache Folgerung aus (i); ich will es hier aber nicht genauer

ausführen. ■

Daß Q_1 symmetrisch ist, sieht man leicht: sein Definitionsbereich ist dicht ($\mathcal{S}$ ist $\mathcal{L}^2$-dicht in $\mathcal{L}^2$, siehe Korollar 1.2.19), und

$$(\psi, Q_1\psi) = \int \overline{\psi(x)} x\psi(x)\, dx = \int \overline{x\psi(x)}\psi(x)\, dx = (Q_1\psi, \psi)\,, \tag{3.10}$$

weil x reell ist. Auch P_1 ist symmetrisch (Ü).
Aufgrund der obigen Gleichung kann man raten, wie man Q_1 zu einem selbstadjungierten Operator erweitern kann: der maximal mögliche sinnvolle Definitionsbereich ist offensichtlich durch diejenigen Funktionen $\psi(x)$ gegeben, für die $x\psi(x)$ noch quadratintegrabel ist; wir definieren

$$Q: \quad \psi(x) \mapsto x\psi(x) \quad \forall\, \psi \text{ mit } \int x^2|\psi(x)|^2\, dx < \infty\,. \tag{3.11}$$

Gleichung (3.10) gilt dann ebenfalls, so daß die Erweiterung symmetrisch ist. Mehr noch: Q ist selbstadjungiert.
Beweise für die Selbstadjungiertheit von Operatoren sind oft recht schwierig, ich werde sie in den seltensten Fällen geben; hier will ich es einmal tun.
Es sei $\varphi \in \mathcal{D}_{Q^*}$ und $\varphi^* = Q^*\varphi$; dann ist $(\varphi, Q\psi) = (\varphi^*, \psi)$ für $\psi \in \mathcal{D}_A$, oder mit Integralen geschrieben:

$$\int \overline{\varphi(x)} x\psi(x)\, dx - \int \overline{\varphi(x)^*}\psi(x)\, dx = \int (\overline{x\varphi(x)} - \overline{\varphi^*(x)})\psi(x)\, dx = 0\,.$$

Diese Gleichung gilt unter anderem für alle $\psi = \chi_I$, wobei I ein beliebiges Intervall ist, und auch für alle $e^{i\alpha(x)}\chi_I(x)$ mit meßbaren reellen Funktionen $\alpha(x)$. Wählt man $\alpha(x) = \operatorname{arc}(x\varphi(x) - \varphi^*(x))$, so ist der Integrand reell und ≥ 0; weil I beliebig ist, sieht man, daß $x\varphi(x) - \varphi^*(x)$ fast überall verschwindet, d. h. $x\varphi = \varphi^*$ im $\mathcal{L}^2$-Sinn, also ist $x\varphi \in \mathcal{L}^2$, d. h. $\varphi \in \mathcal{D}_Q$ und $Q^* = Q$.

Ich nenne einige einfache Eigenschaften der Adjunktion für den Fall von beschränkten Operatoren:

$$\begin{aligned}
(cA)^* &= \bar{c}A^*, \quad c \in \mathbb{C}\,, & (3.12)\\
A^{**} &= A\,, & (3.13)\\
(A+B)^* &= A^* + B^*\,, & (3.14)\\
(AB)^* &= B^*A^*\,. & (3.15)
\end{aligned}$$

Diese Eigenschaften sind leicht zu beweisen (Ü). Sie gelten auch für beliebige Operatoren, wenn man die notwendigen Einschränkungen wegen der Definitionsbereiche macht: (3.12) gilt immer (auf $\mathcal{D}_{A^*}$), (3.13) gilt (auf $\mathcal{D}_A$), wenn A abgeschlossen (und dicht definiert) ist (A^{**} ist abgeschlossen, siehe Satz 4.3.9); damit (3.14) und (3.15)

gelten, müssen die Definitionsbereiche von $A+B$ bzw. AB dicht sein.
Die Physiker pflegen derartige Relationen meist hinzuschreiben, ohne ein Wort über die Definitionsbereiche zu verlieren; gemeint ist dann: es gilt auf einem geeigneten (hoffentlich dichten) Bereich; zum Glück geht das im allgemeinen gut.
Man beachte, daß aus (3.15) sofort folgt:

Lemma 4.3.16 *Wenn A und B selbstadjungiert sind, ist AB dann und nur dann selbstadjungiert, wenn A und B kommutieren.*

Das bedeutet, daß in der Quantenmechanik das Produkt zweier Observabler nicht notwendig wieder eine Observable ist![4]

Die folgende Aussage gilt natürlich nur für beschränkte Operatoren:

$$\|A^*\| = \|A\| \,. \tag{3.17}$$

Beweis: In der Schwarzschen Ungleichung $|(\varphi, \chi)| \leq \|\varphi\| \, \|\chi\|$ gilt Gleichheit für $\chi = \varphi$, und daher kann man die Norm auch durch $\|\chi\| = \sup\{|(\varphi, \chi)|; \; \|\varphi\| \leq 1\}$ (3.18)

darstellen. Angewandt auf $\|A\psi\|$ folgt

$$\|\varphi\| = \sup\{|(\varphi, \psi)|; \; \|\psi\| \leq 1\} \tag{3.18}$$

darstellen. Folglich ist

$$\|A\| = \sup_{\psi \neq 0} \frac{\|A\psi\|}{\|\psi\|} = \sup_{\varphi, \psi \neq 0} \frac{|(\varphi, A\psi)|}{\|\varphi\| \, \|\psi\|} \,.$$

Weil $|\overline{(\varphi, A\psi)}| = |(A\psi, \varphi)| = |(\psi, A^*\varphi)|$ ist, folgt die Behauptung.

Allgemein gilt der Satz

Satz 4.3.19 *Wenn A dicht definiert ist und ein dicht definiertes Inverses A^{-1} hat, so daß A^* und $(A^{-1})^*$ existieren, gilt*

$$(A^{-1})^* = (A^*)^{-1} \,.$$

Beweis: Die Existenz von A^{-1} bedeutet: $A^{-1}A = \mathbb{1}$ auf $\mathcal{D}_A$ und $AA^{-1} = \mathbb{1}$ auf $\mathcal{D}_{A^{-1}}$. Entsprechend müssen wir zwei Gleichungen für die adjungierten Operatoren überprüfen. Sei $\varphi \in \mathcal{D}_{A^*}$ und $\psi \in \mathcal{D}_{A^{-1}}$; dann ist

$$(A^*\varphi, A^{-1}\psi) = (\varphi, AA^{-1}\psi) = (\varphi, \psi) \tag{3.20}$$

[4] Ein solches Produkt tritt bei der Definition des Drehimpulses auf: $\vec{L} = \vec{Q} \times \vec{P}$, in Komponenten: $L_i = Q_k P_l$, (i, k, l) zyklisch. Aber für $k \neq l$ vertauschen P_k und Q_l.

stetig in ψ und daher

$$A^*\varphi \in \mathcal{D}_{(A^{-1})^*}\,, \quad (A^{-1})^*A^*\varphi = \varphi\,,$$

weil (3.20) für eine dichte Menge von ψ's gilt. Ebenso zeigt man mit $\varphi \in \mathcal{D}_A$ und $\psi \in \mathcal{D}_{(A^{-1})^*}$ wegen

$$((A^{-1})^*\psi, A\varphi) = (\psi, A^{-1}A\varphi) = (\psi, \varphi)\,,$$

daß

$$(A^{-1})^*\psi \in \mathcal{D}_{A^*}\,, \quad A^*(A^{-1})^*\psi = \psi\,,$$

und damit folgt $(A^*)^{-1} = (A^{-1})^*$. ■

Nun möchte ich eine Klasse von Operatoren einführen, die wir unter anderem für die Spektralzerlegung brauchen. Es sind spezielle Observable, die Ja-Nein-Fragen entsprechen („Ist der Meßwert der Observablen A größer als 1?“); sie sind wichtig für viele axiomatische Begründungen der Quantenmechanik.

Definition 4.3.21 *Ein Operator P mit $\mathcal{D}_P = \mathcal{H}$, für den $P^2 = P$ gilt, heißt* **Projektor***; wenn P außerdem selbstadjungiert ist, heißt er orthogonaler Projektor.*

Man sagt meist „Projektor“ und meint „orthogonaler Projektor“; die „schiefen“, d. h. nicht-orthogonalen Projektoren spielen in der Quantenmechanik keine Rolle. Ich werde das Adjektiv auch oft weglassen. (Ich benutze gern den Buchstaben P für Projektoren, hier ist natürlich nicht $(\hbar/i)\partial_x$ gemeint.)

Die folgenden einfachen Folgerungen aus dieser Definition führen zu einer vollständigen Charakterisierung der Projektoren.
Ein orthogonaler Projektor $\neq 0$ hat die Norm 1 (schiefe Projektoren können beliebig große Normen haben): aus

$$\|P\varphi\|^2 = (P\varphi, P\varphi) = (\varphi, P\varphi) \leq \|\varphi\|\;\|P\varphi\|$$

folgt $\|P\| \leq 1$, und weil $P \neq 0$, gibt es ein ψ, so daß $\varphi = P\psi \neq 0$, damit wird der Wert 1 erreicht: $\|P\varphi\| = \|P^2\psi\| = \|\varphi\|$. $Q \equiv \mathbb{1} - P$ ist ebenfalls Projektor und $PQ = QP = 0$:

$$(\mathbb{1} - P)^2 = \mathbb{1} - 2P + P = \mathbb{1} - P = (\mathbb{1} - P)^*\,, \quad P(\mathbb{1} - P) = P - P^2 = 0\,.$$

Weil ein Projektor P beschränkt und daher stetig ist, ist $\mathcal{H}_1 \equiv P\mathcal{H}$ ein (abgeschlossener) Unterraum; für $\psi \in \mathcal{H}_1$ ist $P\psi = \psi$, denn ψ läßt sich in der Form $\psi = P\varphi$ schreiben, und $P\psi = P^2\varphi = P\varphi = \psi$.
Der Unterraum $\mathcal{H}_2 \equiv Q\mathcal{H}$ ist wegen $QP = 0$ zu $P\mathcal{H}$ orthogonal:

$$(Q\psi, P\varphi) = (\psi, QP\varphi) = 0\,.$$

Mehr noch: $\mathcal{H}_2$ ist das orthogonale Komplement von $\mathcal{H}_1$; denn $\varphi \in (P\mathcal{H})^\perp$ impliziert $0 = (\varphi, P\psi) = (P\varphi, \psi)$ für alle ψ, folglich $P\varphi = 0$, $\varphi = \varphi - P\varphi = (\mathbb{1} - P)\varphi$, also $\varphi \in Q\mathcal{H}$. (Wir haben hier die Selbstadjungiertheit gebraucht; und die dadurch bedingte Orthogonalität von $P\mathcal{H}$ und $(\mathbb{1} - P)\mathcal{H}$ ist der Grund für die Bezeichnung „orthogonaler“ Projektor.)
Zu jedem Projektor P gehört also ein Unterraum, der unter Anwendung von P punktweise invariant ist; P wirkt auf ihm wie der Einheitsoperator. Auf dem orthogonalen Komplement, das durch $(\mathbb{1} - P)\mathcal{H}$ gegeben ist, verschwindet P. Der Hilbertraum wird also durch P in die „direkte Summe“ zweier orthogonaler Unterräume $\mathcal{H}_1$, $\mathcal{H}_2$ zerlegt: jeder Vektor φ läßt sich eindeutig als Summe $\varphi = \psi_1 + \psi_2$ mit $\psi_i \in \mathcal{H}_i$ darstellen: $\varphi = P\varphi + (\mathbb{1} - P)\varphi$. (Man sagt dann, $\mathcal{H}$ sei „direkte Summe“ der $\mathcal{H}_i$ und schreibt $\mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_2$.)
Umgekehrt kann man jedem Unterraum $\mathcal{H}_1$ einen Projektor zuordnen: sei $\{\varphi_i;\ i \in I\}$ eine o. n. Basis von $\mathcal{H}_1$ (I ist eine Indexmenge, die endlich oder unendlich sein kann); ich kann sie durch eine o. n. Basis des orthogonalen Komplements zu einer o. n. Basis von ganz $\mathcal{H}$ ergänzen und jeden Vektor zerlegen in der Form

$$\psi = \sum_{i \in I} (\varphi_i, \psi)\varphi_i + \psi_R\,,$$

wobei $\psi_R \in \mathcal{H}_1^\perp$. Mit der Definition

$$P\psi := \sum_{i \in I} (\varphi_i, \psi)\varphi_i$$

gewinnt man, wie man leicht kontrolliert, einen orthogonalen Projektor, der genau $\mathcal{H}_1$ als invarianten Unterraum hat.

Diese eineindeutige Zuordnung von Projektoren zu invarianten Unterräumen, wobei der Projektor auf dem Orthokomplement verschwindet, muß man sich als die wesentliche Charakterisierung merken. Man sagt „P ist der Projektor auf $P\mathcal{H}$“. Zum ganzen Hilbertraum gehört die Identität als Projektor, und zum trivialen Unterraum $\{0\}$ der Null-Operator.
Wenn die zu P_1 und P_2 gehörenden Unterräume $\mathcal{H}_1$ und $\mathcal{H}_2$ zueinander orthogonal sind, nennt man die Projektoren zueinander **orthogonal**: $P_1P_2 = 0$, denn $P_k\varphi \in \mathcal{H}_k$ und $0 = (P_1\varphi, P_2\varphi) = (\varphi, P_1P_2\varphi)\ \forall\,\varphi$. $P := P_1 + P_2$ ist dann wieder ein Projektor:

$$(P_1 + P_2)^* = P_1 + P_2, \qquad (P_1 + P_2)^2 = P_1^2 + P_2^2 = P_1 + P_2\,; \tag{3.22}$$

er projiziert auf den durch die $\mathcal{H}_k$ erzeugten Unterraum $\overline{\mathrm{L.H.}(\mathcal{H}_1, \mathcal{H}_2)}$.
Ü: Wenn die beiden Projektoren P_1 und P_2 kommutieren, ist P_1P_2 wieder ein Projektor, sein invarianter Unterraum ist der Durchschnitt der zu den P_i gehörenden Unterräumen. Was ist mit $P_1 + P_2$?

Für manche Überlegungen ist es hilfreich, Projektoren als selbstadjungierte Operatoren zu betrachten, manchmal ist es besser, statt dessen in Unterräumen zu denken.

Wir werden die Projektoren für die Spektralzerlegung brauchen; „diagonalisieren“ heißt: man suche zu A geeignete Unterräume, in denen A in gewisser Weise „einfach“ ist, vielleicht sogar nur ein Vielfaches der Einheit; zudem soll A nicht aus diesen Unterräumen herausführen: $AP\mathcal{H} \subset P\mathcal{H}$. Das entspricht genau dem, was man durch Diagonalisierung von Matrizen erreicht.

(Selbstadjungierte) Projektoren können auch als Observablen auftreten: als Fragen. Wenn die Zustände, in denen einen Eigenschaft $\mathcal{E}$ gilt, durch Vektoren eines Unterraums $P\mathcal{H}$ gegeben sind, ist P die Observable für die Frage: „Liegt $\mathcal{E}$ im Zustand ψ vor?“ Wenn $P\psi = \psi$, ist die Antwort „ja“, wenn $P\psi = 0$, dann „nein“. Die Quantenmechanik läßt auch die Antwort „jein“ zu: genauer, die Antwort „ja mit Wahrscheinlichkeit α“, wenn $\langle P\rangle_\psi = \alpha$.

Ich brauche noch eine Klasse von Operatoren: isometrische und unitäre Operatoren. Sie sind keine Observablen, aber als Hilfsgrößen in der Quantenmechanik wichtig.

Berechen wir das Adjungierte des Operators V_1 aus Beispiel 4.2.4(1): $V_1\varphi_k = \varphi_{k+1}$ für eine o. n. Basis $\{\varphi_k;\ k \in \mathbb{N}\}$. Da der Operator beschränkt ist, wissen wir, daß auch V_1^* beschränkt ist, es genügt, das Adjungierte auf den Basiselementen zu kennen:

$$(\varphi_k, V_1 \sum_i a_i\varphi_i) = \sum_i a_i(\varphi_k, \varphi_{i+1}) = \begin{cases} a_{k-1} = (\varphi_{k-1}, \sum a_i\varphi_i) & \text{falls } k > 1\,, \\ 0 & \text{falls } k = 1\,. \end{cases}$$

Die linke Seite ist gleich $(V_1^*\varphi_k, \sum a_i\varphi_i)$. Daraus können wir ablesen: $V_1^*\varphi_k = \varphi_{k-1}$ für $k > 1$, $V_1^*\varphi_1 = 0$, d. h. $V_1^* = V_2$ aus Beispiel 4.2.4(2). So folgt $V_1^*V_1\varphi_k = \varphi_k$ für alle k, jedoch $V_1V_1^*\varphi_k = \varphi_k$ nur für $k > 1$ und $V_1V_1^*\varphi_1 = 0$; anders:

$$V_1^*V_1 = \mathbb{1}, \quad V_1V_1^* = P\,, \tag{3.23}$$

wo P der Projektor auf $\{\varphi_1\}^\perp$ ist. V_1 hat also kein dicht definiertes Inverses! Aus der ersten Gleichung schließen wir

$$\|V_1\varphi\|^2 = (\varphi, V_1^*V_1\varphi) = \|\varphi\|^2\ . \tag{3.24}$$

Ein Operator V, der Gleichung (3.24) für alle φ erfüllt, definiert eine normerhaltende Abbildung von $\mathcal{H}$ auf den Unterraum $\mathcal{H}_1 \equiv V\mathcal{H}$. Solche Operatoren heißen **isometrisch**. Wieder können wir mit der Polarisationsidentität (I.1.17) schließen, daß dann auch

$$(V\psi, V\varphi) = (\psi, \varphi) \quad \forall \psi, \varphi \in \mathcal{H}\,. \tag{3.25}$$

Ein isometrischer Operator kann auch durch die Gleichung $V^*V = \mathbb{1}$ charakterisiert werden. Da die Abbildung die Norm erhält, wird nur der Nullvektor auf den Nullvektor abgebildet, die Abbildung ist auf ihrem Wertebereich $\mathcal{H}_1$ invertierbar: $V^{-1} = V^*\Big|_{\mathcal{H}_1}$; und es gilt $VV^* = P$, wobei P der Projektor auf $\mathcal{H}_1$ ist.

Es ist natürlich denkbar, daß $P = \mathbb{1}$ ist. Dann existiert das Inverse auf ganz $\mathcal{H}$, und der Operator heißt **unitär**:

Definition 4.3.26 *Es sei $D_U = \mathcal{H}$, und es gelte*

$$U^*U = \mathbb{1}\,, \quad UU^* = \mathbb{1}\,;$$

dann heißt U unitär.

Ü: man rechne nach, daß der Operator U_t aus Beispiel 4.2.5(4) unitär ist.

So, wie man im Komplexen die reellen Zahlen $\{z \in \mathbb{C};\ \overline{z} = z = x\}$, auf den Einheitskreis $\mathcal{C}_1 \equiv \{z \in \mathbb{C};\ \overline{z}z = z\overline{z} = 1\}$ abbilden kann, z. B. durch $z = e^{ix}$, so kann man von selbstadjungierten Operatoren zu unitären kommen (und umgekehrt). Sei A beschränkt und selbstadjungiert. Dann kann man e^{itA} durch die Exponentialreihe definieren:

$$e^{itA} := \sum_{n=0}^{\infty} \frac{(it)^n}{n!}\, A^n\,. \tag{3.27}$$

Weil $\|A\| < \infty$, ist die Reihe konvergent in der durch die Operatornorm gegebenen Topologie: wenn N und M hinreichend groß sind, wird

$$\left\| \sum_N^M \frac{(it)^n}{n!}\, A^n \right\| \leq \sum_N^M \frac{|t|^n}{n!}\, \|A\|^n$$

beliebig klein. (Für diese Abschätzung habe ich mehrfach die Dreiecksungleichung und die Ungleichung (2.16) benutzt.) e^{itA} ist für reelle t ein unitärer Operator: Gleichung (3.27) zeigt, daß $(e^{itA})^* = e^{-itA}$; wegen $e^{it_1A}e^{it_2A} = e^{i(t_1+t_2)A}$ ist $e^{itA}e^{-itA} = e^{-itA}e^{itA} = \mathbb{1}$, d. h. $(e^{itA})^{-1} = e^{-itA}$. Die erste Gleichung kann man ganz genauso ableiten wie für die Exponentialfunktion gewöhnlicher komplexer Zahlen, wichtig dabei ist nur, daß die in den Exponenten stehenden Operatoren miteinander vertauschen (und daß die auftretende Doppelreihe wegen der absoluten Konvergenz beliebig umgeordnet werden kann); und wegen der absoluten Konvergenz sieht man auch leicht, daß $(e^{itA})^* = e^{-itA}$ ist.
(Es ist eine nützliche Übungsaufgabe, sich den Beweis für $e^ae^b = e^{a+b}$ einmal hinzuschreiben und sich zu überzeugen, daß für beschränkte Operatoren die entsprechende Gleichung $e^Ae^B = e^{A+B}$ genau dann gilt, wenn A und B vertauschen.)
Man kann hier die Analogie zur Exponentialfunktion für komplexe Zahlen noch weiter treiben: weil wir es mit einem beschränkten A zu tun haben, existiert

$$\frac{d}{dt}e^{itA} \equiv \lim_{\Delta t \to 0} \frac{1}{\Delta t}\left(e^{i(t+\Delta t)A} - e^{itA}\right) = iA\, e^{itA} \tag{3.28}$$

in der Norm-Topologie (d. h. $\left\| iAe^{itA} - \text{Differenzenquotient} \right\|$ wird beliebig klein). Beweis: Ü.

Auch für beliebig große reelle x liegt e^{ix} auf dem Einheitskreis. Dementsprechend kann man auch für nicht beschränkte selbstadjungierte Operatoren A den unitären Operator e^{itA} definieren; auch die Ableitungsformel kann man auf diesen Fall verallgemeinern. Das werden wir später formulieren, wenn uns die Spektralzerlegung zur Verfügung steht. (Dann kann man natürlich nicht mehr Norm-Konvergenz erwarten, einen passenden Konvergenzbegriff werden wir im nächsten Abschnitt behandeln.)

Zur physikalischen Bedeutung dieser Konstruktion: Analog zur klassischen Mechanik, in der die Dynamik durch die Hamilton-Funktion gegeben wird, ist die quantenmechanische Dynamik durch den selbstadjungierten Hamilton-Operator H mithilfe der Schrödingergleichung

$$i\hbar\frac{d}{dt}\psi_t = H\psi_t$$

bestimmt. Wenn man in (3.28) A durch H und t durch $-t/\hbar$ ersetzt, sieht man, daß die Schrödingergleichung formal durch

$$\psi_t = U_t\psi_0, \qquad U_t = e^{-\frac{i}{\hbar}tH},$$

gelöst wird. Wir werden sehen, daß diese Relation exakt und nicht nur formal gilt. Das mit dem im allgemeinen unbeschränkten H definierte U_t ist wohldefiniert, es ist der Zeitentwicklungsoperator. U_t aus Beispiel 4.2.5(4) beschreibt die Zeitentwicklung eines freien Teilchens.

Hier sehen Sie die Bedeutung der unitären Operatoren in der Quantenmechanik: Im Laufe der Zeitentwicklung sollen die physikalischen Interpretationen, z. B. die Gesamtwahrscheinlichkeit 1, ein Teilchen irgendwo zu finden, erhalten bleiben, d. h. es muß $\|\psi\| = \|\psi_t\|$ gelten, der Zeitentwicklungsoperator muß also isometrisch sein. Die Tatsache, daß U_t darüber hinaus unitär ist, d. h. ein Inverses hat, ist ein Ausdruck dessen, daß die Quantenmechanik wie die Mechanik invariant ist gegen Zeitumkehr. Zusammengenommen: Zeitentwicklungsoperatoren müssen unitär sein.

Unitäre Operatoren treten auch auf, wenn eine Symmetrie des Systems vorliegt: wenn sich z. B. ein Teilchen in einem kugelsymmetrischen Potential bewegt, ist das System rotationsinvariant. Das heißt nicht, daß alle Lösungen rotationsinvariant sind, aber zu jeder Lösung löst auch die „gedrehte" Lösung die Schrödingergleichung. (Das gilt ja genauso in der Mechanik, man nehme beispielsweise ein Kugelpendel, das ja als spezielle Lösungen auch Schwingungen in einer Ebene hat.) Diese Symmetrieoperation stellt man durch Operatoren $U(R)$, $R \in SO(3)$, dar: die „gedrehte" Lösung ist $U(R)\psi$.

„Darstellung" einer Gruppe oder einer Algebra ist ein wohldefinierter mathematischer Begriff; sie ist gegeben durch eine Menge linearer Operatoren eines linearen Raumes, die die Struktur der Gruppe bzw. der Algebra widerspiegeln. Für die Gruppe heißt das: der Darsteller der Einheit e muß der Einheitsoperator sein: $D(e) = \mathbf{1}$; die Multiplikation muß dargestellt sein: $D(ab) = D(a)D(b)$; und schließlich $D(a^{-1}) = D(a)^{-1}$. Wenn man eine Darstellung durch isometrische Operatoren hat, sind sie also automatisch unitär. Beispiel: Eine unitäre Darstellung der Drehungen in $\mathcal{L}^2(\mathbb{R}^3)$ ist gegeben durch

$$U(R) : \psi(\vec{x}) \mapsto \psi(R^{-1}\vec{x})\,.$$

Unabhängig davon, ob das System drehinvariant ist oder nicht, ist es wichtig, daß die Wahrscheinlichkeiten und Erwartungswerte, die man berechnen will, nicht von der speziellen Wahl des Koordinatensystems abhängen. Wenn in einem gedrehten Koordinatensystem also der Zustand ψ durch $U(R)\psi$ beschrieben wird, muß man auch die Observablen passend „gedreht“ beschreiben: Man ersetze A durch $U(R)AU(R)^*$, dann fällt das $U(R)$ in den Erwartungswerten heraus, weil es unitär ist, vergl. (3.25):

$$(U(R)\psi, U(R)AU(R)^*U(R)\psi) = (\psi, AU(R)^*U(R)\psi) = (\psi, A\psi)\,.$$

Wenn wir es mit einem Teilchen in einem kugelsymmetrischen Potential zu tun haben, muß der Hamiltonoperator in allen Koordinatensystemen gleich aussehen:

$$U(R)HU(R)^* = H\,.$$

Das können Sie mit der obigen Darstellung für $H = -\frac{\hbar^2}{2m}\Delta + V(|\vec{x}|)$ auch direkt kontrollieren.

4.4 Konvergenz von Operatoren

Um Approximationsrechnungen zu ermöglichen, die in allen Gebieten der Physik unerläßlich sind, brauchen wir Konvergenzbegriffe.

Ich habe früher definiert, was schwache Konvergenz auf dem Hilbertraum bedeutet (Definition 1.3.9): $\varphi_n \xrightarrow{w} \varphi$, wenn $(\psi, \varphi_n) \longrightarrow (\psi, \varphi)$ für alle ψ. Das ist eine schwächere Konvergenz als die gewöhnliche Konvergenz: $\varphi_n \longrightarrow \varphi \Longleftrightarrow \|\varphi - \varphi_n\| \longrightarrow 0$, die wir der Deutlichkeit halber auch Norm-Konvergenz oder starke Konvergenz nennen wollen.

Satz 4.4.1 *φ_n konvergiere schwach gegen φ.*

(i) Die Normen der φ_n sind beschränkt: $\|\varphi_n\| \leq M$ für ein passendes M.

(ii) Wenn außerdem $\lim \|\varphi_n\| = \|\varphi\|$ gilt, konvergiert φ_n in der Norm gegen φ.

Das ist unser früheres Lemma 1.3.10. Der dortige Beweis der ersten Aussage benutzte das hier nicht bewiesene Banach-Steinhaus-Theorem; deshalb will ich einen alternativen direkten Beweis angeben.

Beweis: Vorbemerkung: ich benutze wieder Gleichung (3.18)

$$\|\varphi\| = \sup_{\|\psi\| \leq 1} |(\varphi, \psi)| \,.$$

Die φ_n definieren lineare stetige Funktionale $\Phi_n : \ \psi \mapsto (\varphi_n, \psi)$; und die Behauptung, daß die Normen $\|\varphi_n\|$ beschränkt sind, ist gleichbedeutend mit der Aussage, daß die Abbildungen Φ_n auf der Einheitskugel beschränkt sind, siehe die Vorbemerkung. Wenn die Φ_n auf irgendeiner Kugel $K(\rho; \psi_0) = \{\psi; \ \|\psi - \psi_0\| < \rho\}$ beschränkt sind, dann auch auf der Einheitskugel: sei $|\Phi_n(\psi)| \leq M_1$ für alle $\psi \in K(\rho; \psi_0)$, dann ist $|\Phi_n(\psi)| \leq M_1/\rho$ auf $K(1; \psi_0)$; für $\|\chi\| \leq 1$ folgt daher $|\Phi_n(\psi_0 + \chi)| \leq M_1/\rho$; und

$$|\Phi_n(\chi)| \leq |\Phi_n(\psi_0)| + |\Phi_n(\psi_0 + \chi)|$$

ist beschränkt, weil nach Voraussetzung $\Phi_n(\psi_0) \longrightarrow (\varphi, \psi_0)$.
Nehmen wir also an, daß die Behauptung falsch sei, d. h. daß die Φ_n in keiner Kugel beschränkt seien. Dann gibt es in einer beliebigen Kugel zu jedem $k > 0$ ein ψ_k und ein Φ_{n_k} mit $|\Phi_{n_k}(\psi_k)| > 2k$. Da Φ_{n_k} stetig ist, gilt für alle ψ in einer hinreichend kleinen Kugel K_k um ψ_k $|\Phi_{n_k}(\psi)| > k$. In dieser kleinen Kugel finden wir zu $k+1$ ein $\Phi_{n_{k+1}}$ und eine noch kleinere Kugel K_{k+1} mit $|\Phi_{n_{k+1}}(\psi)| > k+1 \ \forall \psi \in K_{k+1}$. Wenn wir das Verfahren fortsetzen, erhalten wir eine Folge ineinander enthaltener Kugeln, die, weil der Hilbertraum vollständig ist, mindestens einen gemeinsamen Vektor χ enthalten, für den also $|\Phi_{n_k}(\chi)| > k \ \forall k$ gilt. Das ist aber im Widerspruch zur angenommenen schwachen Konvergenz: $\Phi_n(\chi) \longrightarrow (\varphi, \chi)$. ■

Auch für Operatoren wollen wir verschiedene Arten der Konvergenz definieren. Die nächstliegende ist die durch die Operator-Norm gegebene Konvergenz; doch können wir auch für jeden Vektor ψ die Folgen $\{A_n\psi\}$ betrachten und deren starke oder schwache Konvergenz fordern:

Definition 4.4.2 *(i) Eine Folge von Operatoren $\{A_n\}$ heißt* **norm-konvergent***, wenn $\|A_n - A_m\|$ für hinreichend große n und m beliebig klein wird.*

(ii) Eine Folge $\{A_n\}$ heißt **stark konvergent***, wenn für alle ψ aus einem gemeinsamen Definitionsbereich der A_n die Folgen $\{A_n\psi\}$ Cauchy-Folgen bezüglich der normalen Hilbertraum-Konvergenz sind.*

(iii) Eine Folge $\{A_n\}$ heißt **schwach konvergent***, wenn für alle ψ aus einem gemeinsamen Definitionsbereich der A_n und alle $\varphi \in \mathcal{H}$ die Folgen $\{(\varphi, A_n\psi)\}$ Cauchy-Folgen sind.*

Man überzeugt sich leicht, daß die Norm-Konvergenz die starke Konvergenz impliziert, und daß diese die schwache Konvergenz zur Folge hat. Man kann aber Beispiele konstruieren, die zeigen, daß keine zwei dieser Konvergenzbegriffe übereinstimmen.
Man kann sich weiterhin überlegen, daß jede konvergente Folge von (linearen) Operatoren gegen einen (linearen) Operator konvergiert, das gilt für alle drei Konvergenzbegriffe.
Wir schreiben $A_n \xrightarrow{n} A$ oder $\text{n-lim}\, A_n = A$ bzw. $A_n \xrightarrow{s} A$ oder $\text{s-lim}\, A_n = A$, $A_n \xrightarrow{w} A$ oder $\text{w-lim}\, A_n = A$.

Die Norm-Konvergenz ist eine sehr starke Konvergenz, sie ist im allgemeinen nur interessant, wenn man es mit beschränkten Operatoren zu tun hat. Der Limes-Operator ist dann auch wieder beschränkt (vergl. den folgenden Satz 4.4.3). Anders gesagt (das hatte ich schon früher erwähnt): die Menge der beschränkten Operatoren eines Hilberträums bildet einen Banach-Raum; und da man mit ihnen algebraische Operationen ausführen kann, bilden sie eine „Banach-Algebra“. Mehr noch: da auf dieser Menge eine „$*$-Operation“ existiert, nämlich die Adjunktion, und weil $\|A\| = \|A^*\|$ und $\|A^*A\| = \|A\|^2$ gilt, bilden sie eine „C^*-Algebra“. Doch diese Begriffe und weitere Konvergenzbegriffe, die man definieren kann, spielen für uns im Folgenden keine Rolle.

Analog zu Satz 4.4.1 gilt

Satz 4.4.3 *Wenn $\{A_n\}$ einen schwach konvergente Folge beschränkter Operatoren ist, sind die Normen der A_n beschränkt: $\|A_n\| \leq C$.*

Beweis: soll nur angedeutet werden. Weil nach Voraussetzung $A_n\psi$ für alle ψ eine schwach konvergente Folge ist, gilt nach Satz 4.4.1 $\|A_n\psi\| \leq M_\psi \ \forall n$. Andererseits gilt

$\|A_n\psi\| \leq \|A_n\| \, \|\psi\| \ \forall\psi \ \forall n$. Daraus folgt mit dem Banach-Steinhaus-Theorem (siehe den Beweis von Lemma 1.3.10 oder Lit.) die Behauptung. ■
(Ein direkterer Beweis findet sich im Buch von Achieser und Glasmann.)

Später brauche ich den Begriff des positiven Operators:
Der Operator A heißt **positiv**, wenn $(\varphi, A\varphi) > 0 \ \forall\varphi \in \mathcal{D}_A$. Man schreibt dann $A > 0$. Insbesondere ist $(\varphi, A\varphi)$ dann reell, d. h. A ist hermitisch. Wenn $(\varphi, A\varphi) \geq 0$ gilt, schreibt man $A \geq 0$. Wenn die Differenz $A - B > 0$ ist, sagt man, A sei größer als B (auf einem gemeinsamen Definitionsbereich) und schreibt $A > B$.
Einfaches Beispiel: Seien $\mathcal{H}_i$ Unterräume von $\mathcal{H}$ mit zugehörigen Projektoren P_i und $\mathcal{H}_1 \subset \mathcal{H}_2$. Dann gilt $P_2 \geq P_1 \geq 0$. (Alle Projektoren sind ≥ 0.)[5]

Satz 4.4.4 $\{A_n\}$ *sei eine monoton fallende Folge nicht-negativer beschränkter Operatoren. Dann gilt* $A_n \xrightarrow{s} A$ *mit einem geeigneten* $A \geq 0$.

Natürlich existiert der Limes auch, wenn die Folge durch eine andere Konstante $c\mathbb{1}$, $c \in \mathbb{R}$, nach unten beschränkt ist; und dann ist $A \geq c\mathbb{1}$.

Beweis: Nach Voraussetzung ist die Zahlenfolge $\{(\varphi, A_n\varphi)\}$ abnehmend, nach unten durch 0 beschränkt, also konvergent für alle φ, d. h. $\{A_n\}$ konvergiert schwach gegen ein A, das natürlich auch ≥ 0 sein muß. Die starke Konvergenz ergibt sich aus folgender Überlegung:
Für $n > m$ gilt $A_{mn} \equiv A_m - A_n \geq 0$. Dann definiert $(\psi, \varphi)_{mn} := (\psi, A_{mn}\varphi)$ eine positiv-semidefinite Sesquilinearform, für welche die Schwarzsche Ungleichung gilt:

$$|(\psi, A_{mn}\varphi)|^2 \leq (\varphi, A_{mn}\varphi)(\psi, A_{mn}\psi)\,;$$

und weil $A_m \leq A_1$ und $-(\psi, A_n\psi) \leq 0$, können wir den zweiten Faktor $(\psi, A_{mn}\psi) = (\psi, A_m\psi) - (\psi, A_n\psi)$ durch $\|A_1\| \, \|\psi\|^2$ abschätzen:

$$|(\psi, A_{mn}\varphi)| \leq \sqrt{\|A_1\|} \, \|\psi\| \sqrt{(\varphi, (A_m - A_n)\varphi)}\,.$$

Mit $\psi = (A_m - A_n)\varphi$ und Division durch $\|\psi\|$ folgt daraus

$$\|(A_m - A_n)\varphi\| \leq \sqrt{\|A_1\|} \sqrt{(\varphi, A_m\varphi) - (\varphi, A_n\varphi)}\,. \tag{4.5}$$

Da $(\varphi, A_n\varphi)$ konvergiert, wird die rechte Seite von (4.5) beliebig klein für hinreichend große n und m. Das gilt für alle φ, d. h. A_n ist eine starke Cauchy-Folge mit dem Limes $A \geq 0$. ■

[5] Bedeutung in der Quantenmechanik: wenn die Observable A positiv ist, gilt $\langle A\rangle_\varphi$ für alle φ, d. h. alle Meßwerte sind positiv.

Die Operator-Konvergenz (und der obige Satz) sind wichtig für die später zu behandelnde Spektralzerlegung. Allgemein braucht man sie, weil der Hilbertraum im allgemeinen unendlich-dimensional ist und man gern Operationen in diesem Raum auf endlich-dimensionale zurückführen will; das geht eben höchstens approximativ, d. h. man muß Folgen betrachten.

Ein Beispiel für eine norm-konvergente Folge haben wir schon kennengelernt: die Exponentialfunktion e^A für einen beschränkten Operator A. Der für uns wichtigste Konvergenzbegriff ist die starke Konvergenz. Dafür will ich jetzt Beispiele bringen.
Seien P_n paarweise orthogonale Projektoren: $P_k P_l = \delta_{kl} P_k$; sie gehören zu paarweise orthogonalen Unterräumen $\mathcal{H}_n$. Die Größen

$$Q_N := \sum_{n=1}^{N} P_n$$

sind dann ebenfalls Projektoren, und zwar zu den durch $\mathcal{H}_1, \ldots, \mathcal{H}_N$ aufgespannten Unterräumen (vergl. (3.22)). Wenn die $\mathcal{H}_n$ ganz $\mathcal{H}$ aufspannen, gilt

$$\text{s-lim}\, Q_N = \mathbb{1}\,. \tag{4.6}$$

Man beachte, daß für Folgen von Projektoren die schwache und die starke Konvergenz übereinstimmen, weil $(P\varphi, P\varphi) = (\varphi, P^2\varphi) = (\varphi, P\varphi)$.
Zum Beweis von (4.6): Für die $\mathcal{H}_n$ können wir orthonormale Basen $\left\{\varphi_j^n\right\}_{j\in I_n}$ (mit einer endlichen oder unendlichen Indexmenge I_n) finden; zusammengefaßt bilden die φ_j^n eine o. n. Basis für ganz $\mathcal{H}$. Jedes ψ können wir schreiben als $\psi = \sum_n \sum_j \alpha_{nj}\varphi_j^n$, wobei wegen der Orthonormalität der φ_j^n gilt: $\|\psi\|^2 = \sum_n \sum_j |\alpha_{nj}|^2$. Diese Doppelsumme ist also absolut konvergent; und deshalb können die Summen

$$\|(Q_N - Q_M)\psi\|^2 = (\psi, (Q_N - Q_M)\psi) = \sum_{n=M+1}^{N} (\psi, P_n\psi) = \sum_{n=M+1}^{N} \sum_j |\alpha_{nj}|^2$$

für hinreichend großes N und M beliebig klein gemacht werden.
Es ist klar, daß die Q_N nicht in der Norm konvergieren: $Q_N - Q_M$ ist wieder ein Projektor, hat also die Norm 1, egal wie groß N und M sind.

Mit Satz 4.4.4 kann man einen alternativen Beweis geben: Dieser Satz gilt offensichtlich auch für *aufsteigende* Folgen, die nach oben beschränkt sind. Wenn die zu zwei Projektoren $P^{(1)}$ und $P^{(2)}$ gehörigen Unterräume ineinander liegen: $\mathcal{H}^{(1)} \subset \mathcal{H}^{(2)}$, dann gilt $P^{(1)}P^{(2)} = P^{(2)}P^{(1)} = P^{(1)}$ und $P^{(1)} \leq P^{(2)}$, denn

$$\begin{aligned}(\varphi, (P^{(2)} - P^{(1)})\varphi) &= (\varphi, P^{(2)}(P^{(2)} - P^{(1)})P^{(2)}\varphi) = (P^{(2)}\varphi, (P^{(2)} - P^{(1)})P^{(2)}\varphi) \\ &= \left\|(P^{(2)} - P^{(1)})P^{(2)}\varphi\right\|^2 \geq 0\,.\end{aligned}$$

Also ist $Q_N \leq Q_{N+1} \leq \mathbb{1}$, und folglich konvergiert die Folge der Q_N stark gegen den Projektor zu dem von den $\mathcal{H}_N$ aufgespannten Raum.

Ich kann das hier Dargelegte auch anders ausdrücken: Wenn die $\mathcal{H}_N$ ganz $\mathcal{H}$ aufspannen, ist für jedes ψ $\psi = \sum P_n\psi$, wobei die Summanden paarweise orthogonal sind, also ist $\|\psi\|^2 = \sum_n \|P_n\psi\|^2$, und daher gilt $P_n\psi \longrightarrow 0\ \forall\psi$, folglich $P_n \xrightarrow{s} 0$ und $Q_N - Q_M = \sum_{M+1}^N P_n \xrightarrow{s} 0$.

Betrachten wir nun

$$A_N := \sum_{n=1}^N a_n P_n\,.$$

Das ist ein beschränkter Operator mit Norm $\max|a_n|$; denn mit der obigen Zerlegung von ψ nach den φ_j^n gilt

$$\|A_N\psi\|^2 = \sum_n \sum_j |a_n|^2 |\alpha_{nj}|^2 \le \max|a_n|^2\,\|\psi\|^2\ .$$

Wenn die $|a_n|$ für $n \to \infty$ beschränkt sind, sieht man mit der gleichen Art von Rechnung, daß die A_N für $N \to \infty$ stark konvergieren, den Limes-Operator nenne ich $A = \sum_n a_n P_n$:

$$\|(A - A_N)\psi\|^2 = \left\|\sum_{n>N} a_n P_n\psi\right\|^2 = \sum_{n>N} |a_n|^2\,\|P_n\psi\|^2\ , \tag{4.7}$$

und die rechte Seite wird wegen der Beschränktheit der a_n beliebig klein, weil $\sum_0^\infty \|P_n\psi\|^2 = \|\psi\|^2 < \infty$.

Wenn die a_n nicht beschränkt sind, kann ich ebenfalls A durch die obige Formel definieren; es ist aber jetzt kein beschränkter Operator mehr; sein Definitionsbereich ist gegeben durch

$$\mathcal{D}_A := \left\{\psi;\ \sum_n |a_n|^2(\psi, P_n\psi) < \infty\right\}\ .$$

Die Summe ist gleich $\sum|a_n|^2\,\|P_n\psi\|^2$, weil die P_n Projektoren sind; und daraus ersieht man, daß für $\psi \in \mathcal{D}_A$ die gleiche Rechnung wie in (4.7) durchgeführt werden kann; also

$$A_N = \sum_{n=1}^N a_n P_n \xrightarrow{s} A = \sum_{n=1}^\infty a_n P_n \quad \text{auf } \mathcal{D}_A\,. \tag{4.8}$$

(Wenn man die Folge der A_N auf einem Teilraum betrachtet, der von der Summe *endlich vieler* P_n´s aufgespannt wird, kommen nur endlich viele a_n ins Spiel; und die haben ein Maximum. Dann läuft die Argumentation also wie oben, auf diesem Teilraum existiert der starke Limes. Wenn man den Limes $A_N\varphi$ für ein beliebiges φ studieren will, muß man φ durch Vektoren aus den zuvor betrachteten Teilräumen approximieren („3ε–Argument“).)

Wenn die $\mathcal{H}_n$ nicht ganz $\mathcal{H}$ sondern nur einen Unterraum $\mathcal{H}'$ mit zugehörigem Projektor P' aufspannen, kann man natürlich entsprechend argumentieren: $Q_N \xrightarrow{s} P'$; und A kann man entweder so schreiben, daß $\sum P_n = P'$ ist, oder man kann noch einen Term $0\cdot(\mathbb{1} - P')$ hinzufügen und hat dann wieder den Fall wie oben.

Wenn die a_n reell sind, ist Gleichung (4.8) ein erster Fall von Spektralzerlegung; mehr noch: A mit dem oben angegebenen Definitionsbereich $\mathcal{D}_A$ ist selbstadjungiert, Beweis im nächsten Abschnitt. Man hat eine Darstellung von A durch einfachere Operatoren, nämlich die Projektoren P_n, die möglicherweise sogar auf endlich-dimensionale Räume projizieren; eingeschränkt auf $P_n\mathcal{H}$ sieht A höchst einfach aus: $A\Big|_{P_n\mathcal{H}} = a_n\mathbb{1}\Big|_{P_n\mathcal{H}}$.

Die Ableitung von e^{itA} nach t, Gleichung (3.28), ist ein weiteres Beispiel für eine normkonvergente Folge; ich habe danach bemerkt, daß man e^{itA} auch für unbeschränkte selbstadjungierte A definieren und ableiten kann, in der Ableitungsformel kann man natürlich auch nur starke Konvergenz erwarten.

4.5 Eigenwerte und Eigenvektoren

Gibt es zu einem Operator A einen Vektor $\varphi \in \mathcal{D}_A$, so daß

$$A\varphi = a\varphi$$

mit einer komplexen Zahl a gilt? Wenn ja, dann heißt a **Eigenwert** von A und φ der zugehörige **Eigenvektor**. Wenn A eine $n \times n$-Matrix ist, ist das ein homogenes Gleichungssystem, das lösbar ist, wenn $\det(A - a\mathbb{1}) = 0$ ist, diese Gleichung liefert die Eigenwerte der Matrix, die als Diagonalelemente bei der Diagonalisierung der Matrix auftauchen. Symmetrische Matrizen sind diagonalisierbar, haben also n Eigenwerte.

Wie sieht es aus, wenn $\dim \mathcal{H} = \infty$? Man kann leicht Beispiele von Operatoren konstruieren, die „diagonal“ sind:
Sei $\{\varphi_n\}$ eine o. n. Basis, $\mathcal{D}_A = \text{L.H.}\,\{\varphi_n\}$ und $A\varphi_n = a_n\varphi_n$ (vergl. Beispiele 4.2.3(1) und (2)). Dann ist trivialerweise jeder Basisvektor Eigenvektor von A. Es ist durchaus möglich, daß einige der a_n einander gleich sind, $a_{n_j} = b,\ j = 1 \dots k$, dann gibt es zu einem Eigenwert mehrere Eigenvektoren $\varphi_{n_1}, \dots, \varphi_{n_k}$; und trivialerweise ist dann auch jede Linearkombination $\sum \gamma_j \varphi_{n_j}$ Eigenvektor zum gleichen Eigenwert b. Anders gesagt: die Vektoren $\varphi_{n_1}, \dots, \varphi_{n_k}$ spannen einen Unterraum $\mathcal{H}_1$ mit Projektor P auf, der **Eigenraum** von A zum Eigenwert b ist. Wenn man auf diese Weise alle gleichen Eigenwerte zusammengefaßt hat, kann man mit den entsprechenden Eigenprojektoren schreiben

$$A = \sum b_r P_r\,. \tag{5.1}$$

Diese Summe ist im Sinne der starken Operatorkonvergenz gemeint; sie existiert, weil die P_r nach Konstruktion paarweise orthogonal sind (vergl. (4.8)).

Warum interessiert sich der Physiker für Eigenwerte und Eigenvektoren? Eigenvektoren eines Operators A, der eine Observable repräsentiert, definieren Zustände, in denen der Erwartungswert von A scharf ist: wenn $A\varphi = a\varphi$, folgt

$$\langle A \rangle_\varphi = \frac{(\varphi, A\varphi)}{(\varphi, \varphi)} = a\,;$$

und wegen $A^2\varphi = a^2\varphi$ ist $\langle A^2 \rangle_\varphi = a^2$, also

$$\Delta_\varphi A = \left(\langle A^2 \rangle_\varphi - \langle A \rangle_\varphi^2\right) = 0\,.$$

Wenn man A gleich dem Hamilton-Operator setzt, der die Observable „Energie“ darstellt, liefert ein Eigenvektor einen Zustand mit scharfer, wohldefinierter Energie; Atome können Übergänge zwischen solchen Zuständen scharfer Energie machen, die Energie-Differenz wird als Lichtquant mit scharf definierter Frequenz ausgesandt.
Es gilt auch die Umkehrung der obigen Aussage:
Wenn A ein hermitischer Operator ist, der im Zustand φ den Erwartungswert a mit

Schwankung 0 hat, gilt $A\varphi = a\varphi$.
Beweis: $\psi := A\varphi - a\varphi$ ist senkrecht auf φ, denn $(\varphi, \psi) = (\varphi, A\varphi) - a(\varphi, \varphi) = 0$, weil $\langle A\rangle_\varphi = a$ ist. Weiterhin gilt

$$0 = \left(\langle A^2\rangle_\varphi - \langle A\rangle_\varphi^2\right)(\varphi, \varphi) = (\varphi, A^2\varphi) - a^2(\varphi, \varphi)\,.$$

Der erste Term rechts ist $= (A\varphi, A\varphi) = (\psi + a\varphi, \psi + a\varphi) = (\psi, \psi) + a^2(\varphi, \varphi)$, also ist $(\psi, \psi) = 0$, d. h. $\psi = A\varphi - a\varphi = 0$.

Nun ein weniger triviales Beispiel eines Operators mit Eigenwerten: sei im $\mathcal{L}^2(\mathbb{R})$

$$H = -\frac{d^2}{dx^2} + x^2\,.$$

(Allgemein interessieren in der Quantenmechanik eines Teilchens, das sich nur in einer Dimension bewegen kann, die Eigenwerte des Hamilton-Operators $H = \frac{1}{2m}P^2 + V(Q)$. Wenn $V(x) = \frac{1}{2}\kappa x^2$ ist, haben wir den berühmten harmonischen Oszillator, der — in der einen oder anderen Form — in keiner Vorlesung der Theoretischen Physik fehlt, weil er zu den wenigen exakt lösbaren Beispielen gehört Mit der Federkonstanten $\kappa = 2$ und $m = 1/2$ ergibt sich der obige Operator, wenn man $\hbar = 1$ setzt, was Theoretiker gern tun; das bedeutet ja nur eine spezielle Wahl der Maßeinheit für die Energie).
Ich habe den Definitionsbereich nicht angegeben; es genügt uns, Vektoren aus $\mathcal{D}_0 \equiv \mathcal{S}(\mathbb{R})$ zu betrachten. Es zeigt sich später, daß alle Eigenvektoren in dieser Untermenge liegen. Ich möchte aber nicht auf alle mathematischen Feinheiten eingehen. Die folgenden Rechnungen führe ich alle auf $\mathcal{D}_0$ durch, ohne es jedesmal zu sagen.
Es ist nicht schwer, zu zeigen, daß $\varphi_g := e^{-x^2/2}$ ein Eigenvektor von H ist:[6]

$$\left(e^{-x^2/2}\right)' = -xe^{-x^2/2},\ \left(e^{-x^2/2}\right)'' = -e^{-x^2/2} + x^2e^{-x^2/2}\,,$$

also

$$H\varphi_g = \left(-\frac{d^2}{dx^2} + x^2\right)e^{-x^2/2} = e^{-x^2/2} = \varphi_g\,.$$

Definieren wir den Operator $P = \frac{1}{i}\frac{d}{dx}$ und den Operator Q der Multiplikation mit x, dann können wir H schreiben als

$$H = P^2 + Q^2 = (Q - iP)(Q + iP) + i(PQ - QP)\,.$$

Den zweiten Term brauchen wir, weil Q und P nicht vertauschen, er ist $= i[P, Q] = \mathbf{1}$, siehe Gleichung (1.6). Mit der Definition

$$a := \frac{1}{\sqrt{2}}(Q + iP), \quad a^+ := \frac{1}{\sqrt{2}}(Q - iP) = a^*\Big|_{\mathcal{D}_0} \tag{5.2}$$

[6] Der Grund, warum ich den Index g genannt habe, wird in Kürze klar: φ_g ist der „Grundzustand“ von H.Man kann φ_g raten: Fourier-Transformation der Eigenwertgleichung $(-\frac{d}{dx^2})\psi = E\psi$ ergibt eine Gleichung gleicher Art: $(-\frac{d}{dk^2})\tilde{\psi} = E\tilde{\psi}$; und wir wissen, daß die Gauß'sche Glockenkurve bei Fourier-Transformation eine Funktion derselben Art ergibt.

wird daraus

$$H = P^2 + Q^2 = 2a^+a + \mathbb{1}\,.$$

Anstatt nach den Eigenwerten von H zu suchen, können wir ebenso gut nach denen von $N := a^+a$ fragen; wenn $N\varphi = \nu\varphi$ ist, folgt $H\varphi = 2N\varphi + \varphi = (2\nu + 1)\varphi$, und umgekehrt ist jeder Eigenwert von H Eigenwert von N; insbesondere ist

$$N\varphi_g = \frac{1}{2}(H - \mathbb{1})\varphi_g = 0\,.$$

Die Eigenwerte von N lassen sich (fast) rein algebraisch mithilfe der Vertauschungsrelationen der Operatoren a, a^+ und N bestimmen. Aus der Vertauschungsrelation $[P, Q] = \frac{1}{i}\mathbb{1}$ berechnet man leicht

$$[a, a^+] = \mathbb{1}\,, \tag{5.3}$$

und daraus

$$[N, a] = a^+aa - aa^+a = (a^+a - aa^+)a = -a, \qquad [N, a^+] = \ldots = a^+\,. \tag{5.4}$$

Zunächst bemerken wir, daß $N \geq 0$ ist: $(\varphi, N\varphi) = (a\varphi, a\varphi) \geq 0 \ \forall\, \varphi \in \mathcal{D}_0$. Daher kann N nur nicht-negative Eigenwerte haben, denn $N\varphi = r\varphi$ impliziert $0 \leq (\varphi, N\varphi) = r(\varphi, \varphi)$.
Wenn es *einen* Eigenvektor gibt — und wir haben ja bereits einen gefunden —, können wir weitere Eigenvektoren konstruieren, wie das folgende Lemma zeigt. Basis ist die Gleichung (5.3):

Lemma 4.5.5 *Sei φ Eigenvektor von N zum Eigenwert r. Dann gilt*

$$Na\varphi = (r - 1)a\varphi, \qquad Na^+\varphi = (r + 1)a^+\varphi\,; \tag{5.6}$$

$$\|a\varphi\|^2 = r\,\|\varphi\|^2\,, \qquad \left\|a^+\varphi\right\|^2 = (r + 1)\,\|\varphi\|^2\,. \tag{5.7}$$

Beweis: $Na\varphi = [N, a]\varphi + aN\varphi$, mit (5.4): $\ldots = -a\varphi + a(r\varphi) = (r - 1)a\varphi$. Ebenso folgt die zweite Gleichung in (5.6).
Berechnung der Normen: $\|a\varphi\|^2 = (\varphi, a^+a\varphi) = r(\varphi, \varphi)$; und $\|a^+\varphi\|^2 = (\varphi, aa^+\varphi) = (\varphi, [a, a^+]\varphi) + (\varphi, a^+a\varphi) = (r + 1)(\varphi, \varphi)$. ■

Die erste Gleichung in (5.7) zeigt noch einmal, daß $r \geq 0$ sein muß. Weitere Folgerungen:
a) Zum Eigenwert r gibt es wegen (5.6) den Eigenwert $r+1$ mit Eigenvektor $a^+\varphi$, denn (5.7) zeigt, daß $a^+\varphi \neq 0$ ist. Ebenso folgt, daß alle Werte $r + k$, $k \in \mathbb{N}$, Eigenwerte sind.
b) Zum Eigenwert r gibt es den Eigenwert $r - 1$ dann und nur dann, wenn $a\varphi \neq 0$

ist; und das gilt genau dann, wenn $r \neq 0$ ist. Solange $a^k\varphi \neq 0$ bleibt für wachsendes k, kann man auf der „Eigenwertleiter" weiter hinabsteigen: $r-k$ ist dann auch Eigenwert.
c) Die Eigenwerte müssen ≥ 0 sein, daher muß man bei Absteigen mithilfe des Operators a irgendwann den Eigenwert 0 mit einem Eigenvektor φ_0 treffen, für den $a\varphi_0 = 0$ ist, siehe die erste Gleichung (5.7), dann bricht das Verfahren ab. Gäbe es ein $r \notin \mathbb{N} \cup \{0\}$, so gäbe es einen negativen Eigenwert $r-k < 0$ (mit $a^k\varphi \neq 0$) für passendes k. Nur für $r \in \mathbb{N}$ trifft man beim Hinabsteigen der Leiter einmal auf den Eigenwert 0. Also haben wir gezeigt: es gibt einen Eigenvektor φ_g zum Eigenwert 0, und deshalb sind alle Werte $0, 1, 2, \ldots$ Eigenwerte von N; und es gibt einen „Grundzustand" φ_g, der zum tiefstmöglichen Eigenwert gehört. Er ist gleichzeitig der Grundzustand von H.
Damit haben wir alle Eigenwerte von N und damit alle Eigenwerte von H gefunden. (Wir können an diesem Punkt noch nicht ausschließen, daß der Eigenwert 0 von N entartet ist; es könnte außer φ_g noch andere Eigenvektoren zum Eigenwert 0 geben; dazu gleich etwas mehr.) Man beachte: der niedrigste Eigenwert von H, die „Grundzustandsenergie", ist 1. Ein klassisches Teilchen mit $m = 1/2$ und der Hamiltonfunktion $H = p^2 + x^2$ könnte am Punkt $x = 0$ ruhen und hätte dann die Energie 0; ein quantenmechanisches Teilchen hat mindestens eine positive „Nullpunktsenergie".[7]
(Wenn man den Oszillator mit Masse m, Federkonstante κ und (klassischer) Eigenfrequenz $\omega = \sqrt{\kappa/m}$ behandelt und $\hbar$ nicht gleich 1 setzt, definiert man die Leiteroperatoren mit geeigneten Konstanten, so daß die Vertauschungsrelation (5.3) ebenfalls gilt; dann wird $H = \hbar\omega(N + \frac{1}{2}\mathbb{1}),\ N = a^+a$, mit den Eigenwerten $\hbar\omega(r + \frac{1}{2}),\ r \in \mathbb{N} \cup \{0\}$.)

Man kann Eigenvektoren für die höheren Eigenwerte auch als Funktionen in $\mathcal{L}^2$ angeben: ausgehend von $\varphi_g = e^{-x^2/2}$ bekommen wir die höheren Eigenvektoren durch Anwendung von $(a^+)^k = 2^{-k/2}(x - \frac{d}{dx})^k$. Das ergibt Funktionen der Form $\mathcal{P}_k(x)\, e^{-x^2/2}$ mit Polynomen k-ten Grades $\mathcal{P}_k$. Offensichtlich sind auch diese Eigenfunktkionen in $\mathcal{S}(\mathbb{R})$. Genaueres über diese Funktionen werden Sie in der Quantenmechanik lernen, unter anderem auch, daß diese Eigenfunktionen den ganzen Hilbertraum aufspannen. Die Begründung will ich nur kurz andeuten: Aus einer Menge von Polynomen *aller* Grade $n = 0, 1, 2, \ldots$ lassen sich durch Linearkombination die Potenzen $x^n,\ n = 0, 1, 2, \ldots$ erzeugen; und mit ihnen kann man nach Weierstraß jede stetige Funktion f mit kompaktem Träger approximieren, also auch die Funktion $fe^{+x^2/2}$. Daher lassen sich mit Linearkombinationen der $\mathcal{P}_k(x)\, e^{-x^2/2}$ alle stetigen Funktionen approximieren, die kompakten Träger haben, und diese sind dicht in $\mathcal{L}^2(\mathbb{R})$. Wir werden in Kürze sehen, daß Eigenvektoren symmetrischer Operatoren orthogonal sind, d. h. die normierten Eigenvektoren von H bilden ein orthogonales System, und hier, weil sie eine totale Menge bilden, sogar eine orthogonale Basis. Daraus folgt dann, daß wir alle Eigenwerte gefunden haben. Weiterhin folgt, daß $H = \sum_{n=0} \infty(2n+1)P_{\varphi_n},\ \varphi_n = (a^+)^n\varphi_g$. Jedes ψ läßt sich nach den φ_n entwickeln, $\psi = \sum \alpha_n\varphi_n$, Œ$\|\psi\|^2 = \sum |\alpha_n|^2 = 1$, daher ist $\langle H\rangle_\psi = \sum |\alpha_n|^2(2n+1) \leq \langle H\rangle_{\varphi_g}$.

[7] Wir zeigen in Kürze, daß für alle $\psi \in \mathcal{D}(H)$ gilt: $\langle H\rangle_\psi \leq \langle H\rangle_{\varphi_g}$.

Die Aussage, daß die Eigenvektoren eine Basis bilden, gilt — trivialerweise — auch für die Eigenvektoren unseres ersten Beispiels (5.1). Das muß nicht immer so sein. Es ist möglich, daß die Eigenvektoren eines Operators einen (endlich-dimensionalen oder unendlichen) Unterraum aufspannen; es ist natürlich auch möglich, daß der Operator gar keine Eigenwerte und Eigenvektoren hat! Einfaches Beispiel:

$$P\psi = \frac{1}{i}\frac{d}{dx}\psi(x) = k\psi(x)$$

hat die Lösungen $\psi(x) = \text{const}\, e^{ikx}$; doch das sind keine $\mathcal{L}^2$-Vektoren. Wir können sie als „uneigentliche Eigenvektoren" bezeichnen, sie sind, wie wir früher gesehen haben, „auf die δ-Funktion normiert": $(e^{ikx}, e^{ik'x}) = \delta(k-k')$, und bilden in gewisser Weise ein vollständiges System: wir können ja jeden $\mathcal{L}^2$-Vektor Fourier-transformieren, also nach diesen uneigentlichen Eigenvektoren „entwickeln".
Auch der Multiplikationsoperator Q hat keine eigentlichen Eigenwerte:

$$Q\psi = x\psi(x) = \xi\psi(x)$$

wird gelöst durch $\psi(x) = \delta(x - x_0)$ mit dem uneigentlichen Eigenwert $\xi = x_0$.

Die mathematisch saubere Behandlung dieser uneigentlichen Eigenwerte innerhalb des Hilbertraum-Formalismus ist der Inhalt der Spektraltheorie selbstadjungierter Operatoren.

Selbstadjungierte Operatoren, deren Eigenvektoren den ganzen Raum aufspannen, kann man — wie endliche Matrizen — in Diagonalform darstellen.

Satz 4.5.8 *Sei A symmetrisch, φ_1 und φ_2 seien Eigenvektoren von A zu verschiedenen Eigenwerten a_1 und a_2. Dann sind die a_i reell, und φ_1 und φ_2 orthogonal.*

Beweis: Aus $a_i(\varphi_i, \varphi_i) = (\varphi_i, A\varphi_i) = (A\varphi_i, \varphi_i) = \overline{(\varphi_i, A\varphi_i)}$ folgt, daß $a_i \in \mathbb{R}$.
$a_2(\varphi_1, \varphi_2) = (\varphi_1, A\varphi_2) = (A\varphi_1, \varphi_2) = a_1(\varphi_1, \varphi_2)$ kann wegen $a_1 \neq a_2$ nur gelten, wenn $(\varphi_1, \varphi_2) = 0$ ist. ■

Die Eigenräume zu verschiedenen Eigenwerten eines symmetrischen Operators A sind also orthogonal. Zu a_n gehöre der Eigenprojektor P_n; und $A\varphi = a_n\varphi$, wenn $\varphi \in P_n\mathcal{H}$. Wenn die Eigenvektoren ganz $\mathcal{H}$ aufspannen, ist $\sum P_n = \mathbb{1}$, also $A\psi = A\sum P_n\psi = \sum a_n P_n\psi$, d. h. A läßt sich schreiben als

$$A = \sum a_n\, P_n\,. \tag{5.9}$$

Das ist das genaue Analogon der Diagonaldarstellung einer Matrix: im Unterraum $P_n\mathcal{H}$ ist $A = a_n\mathbb{1}$. In dieser Form können wir also auch unser obiges H schreiben mit

$a_n = 2n+1$. Der „natürliche“ Definitionsbereich von A aus (5.9) ist gegeben durch

$$\mathcal{D}_A = \left\{\varphi \in \mathcal{H};\ \sum |a_n|^2 (\varphi, P_n\varphi) < \infty\right\} . \tag{5.10}$$

Es gilt sogar noch mehr:

Satz 4.5.11 *A habe die Form (5.9) mit paarweise orthogonalen Projektoren P_n und reellen a_n. Dann ist A selbstadjungiert auf dem durch Gleichung (5.10) gegebenen Definitionsbereich.*

Beweis: Trivialerweise ist A hermitisch; $\mathcal{D}_A$ ist dicht, also ist A symmetrisch. Wenn $\sum P_n = P' \neq \mathbb{1}$, ergänze man die Summen in (5.9) und (5.10) durch $P_0 = \mathbb{1} - P'$ und $a_0 = 0$, so daß $\sum_{n\geq 0} P_n = \mathbb{1}$. Seien $\left\{\varphi_j^n\right\}_{j\in I_n}$ orthonormale Basen von $P_n\mathcal{H}$, die zusammengenommen eine o. n. Basis von $\mathcal{H}$ bilden. Sei $\varphi = \sum_{n,j} \alpha_{nj}\varphi_j^n \in \mathcal{D}_A$, das bedeutet, daß die Folge $\{a_n\alpha_{nj}\}_{n,j}$ quadratsummabel ist; und sei $\psi = \sum \beta_{nj}\varphi_j^n \in \mathcal{D}_{A^*}$, dann ist $(\psi, A\varphi) = \sum \overline{\beta}_{nj} a_n \alpha_{nj}$ stetig in φ, daher muß die rechte Seite auf *alle* quadratsummablen Folgen α_{nj} erweiterbar sein. Das geht nur, wenn auch die Folge $\{a_n\beta_{nj}\}$ quadratsummabel ist: $\{\alpha_{nj}\} \mapsto \sum \overline{\beta}_{nj} a_n \alpha_{nj}$ definiert nach Konstruktion ein stetiges lineares Funktional auf $\ell^2(\mathbb{C})$,[8] jeder Hilbertraum ist aber selbstdual, also ist $\{a_n\beta_{nj}\} \in \ell^2(\mathbb{C})$. Das heißt, daß ψ bereits $\in \mathcal{D}_A$ ist, also $A = A^*$ gilt. ■

Die Darstellung (5.9), die man gewinnt, wenn ein symmetrischer Operator hinreichend viele Eigenvektoren hat, ist für diesen Fall bereits eine Form der Spektraldarstellung. Satz 4.5.11 und seine Verallgemeinerung für allgemeine Spektraldarstellungen ist ein gutes Werkzeug, um die Selbstadjungiertheit eines Operators nachzuweisen.

Wenn die Eigenvektoren nicht ganz $\mathcal{H}$ aufspannen, hat man diese schöne Darstellung nur in einem Unterraum; im orthogonalen Komplement hat ein selbstadjungierter Operator ähnliche Eigenschaften wie die beiden obigen Beispiele P und Q, es gibt nur uneigentliche Eigenwerte. Beispiel Q: der Operator ist selbstadjungiert, die uneigentlichen Eigenvektoren sind auch formal orthogonal; doch es gibt überabzählbar viele von ihnen, die Eigenwerte bilden ein Kontinuum. Das ist eine typische Komplikation, die im unendlich-dimensionalen Hilbertraum auftreten kann.
Die Eigenwerte eines selbstadjungierten Operators, der eine Observable darstellt, sind die möglichen *scharfen* Werte dieser Observablen. Am Beispiel von Q kann man sich klar machen, daß die uneigentlichen Eigenwerte Werte sind, die zwar nicht ganz scharf, aber doch *beliebig scharf* sein können, d. h. zu ε und x_0 gibt es ein φ, so daß $\langle Q\rangle_\varphi = x_0$ und $\Delta_\varphi Q < \varepsilon$ (Ü).
Wenn A nicht symmetrisch ist, können die Eigenwerte beliebig komplex sein.

[8] Hier sind die Elemente von $\ell^2(\mathbb{C})$ mit zwei Summationsindizes versehen; diese können im Prinzip natürlich zu einem reduziert werden.

Als Erwartungswerte eines Operators können alle Werte auftreten, die in der konvexen Hülle der eigentlichen und uneigentlichen Eigenwerte liegen (Ü).

4.5.12 Anmerkung: Eigenwerte unitärer Operatoren liegen auf dem Einheitskreis in $\mathbb{C}$. Wenn ψ Eigenvektor von U zum Eigenwert ζ ist, gilt $U^*\psi = \overline{\zeta}\psi$. Eigenvektoren von U zu verschiedenen Eigenwerten sind orthogonal. Diese Aussagen sind leicht zu beweisen (Ü).

Es folgt eine Reihe von <u>Ergänzungen</u>, teils mathematischen, teils physikalischen Inhalts.

E 1. Sei $A\varphi = a\varphi$, und B sei ein mit A vertauschbarer selbstadjungierter Operator, dann zeigt die einfache Rechnung

$$AB\varphi = BA\varphi = B(a\varphi) = a(B\varphi)\,,$$

daß auch $B\varphi$ ein Eigenvektor von A zum selben Eigenwert ist. Zwei Möglichkeiten sind denkbar: (i) der Eigenraum zu a ist eindimensional; dann ist $B\varphi$ ein Vielfaches von φ: $B\varphi = b\varphi$, also φ gleichzeitig Eigenvektor von B. (ii) $\dim P_a > 1$; dann kann man die Einschränkung von B auf $\mathcal{H}_a = P_a\mathcal{H}$ betrachten und hoffen, daß $B\Big|_{\mathcal{H}_a}$ Eigenwerte in $\mathcal{H}_a$ hat. (Das ist sicher der Fall, wenn $\dim P_a < \infty$.) Ein Eigenvektor von $B\Big|_{\mathcal{H}_a}$ ist natürlich auch Eigenvektor von B und von A.
Wenn es weitere Operatoren gibt, die mit A und B und miteinander vertauschen, kann man das Spiel noch weiter treiben. Manchmal kann man einen Vektor bis auf die Normierung dadurch charakterisieren, daß er gleichzeitiger Eigenvektor zu kommutierenden Operatoren A, B, C, $\ldots$ zu Eigenwerten a, b, c, $\ldots$ ist. In der Dirac-Schreibweise bezeichnet man ihn dann kurz mit $|\,a, b, c, \ldots\rangle$. Derartige Eigenwerte von Atom-Zuständen heißen auch „Quantenzahlen“ des Atoms im entsprechenden Zustand.

E 2. Die Verhältnisse bei kommutierenden Operatoren sind wie in der klassischen Mechanik: man kann sie alle „gleichzeitig messen“, genauer gesagt: gleichzeitig scharf messen, d. h. sie können gleichzeitig scharfe Werte annehmen ($\Delta_\varphi A = 0$, wenn φ Eigenvektor von A ist). Wie schon erwähnt, gilt für nicht vertauschbare Operatoren eine allgemeine Unschärferelation; ich will sie hier kurz ableiten.

Satz 4.5.13 *A und B seien zwei selbstadjungierte Operatoren. Dann gilt für φ aus einem geeigneten Definitionsbereich*

$$\Delta_\varphi A\,\Delta_\varphi B \geq \frac{1}{2}\,|\langle i[A,B]\rangle_\varphi|\ .$$

(Der Faktor i ist wegen der Betragsstriche natürlich nur Verzierung; Grund: $i[A,B]$ ist hermitisch.)

Beweis: Wir benutzen die Schwarzsche Ungleichung:

$$|\Im(A\varphi, B\varphi)|^2 \leq |(A\varphi, B\varphi)|^2 \leq \|A\varphi\|^2 \; \|B\varphi\|^2 \; ;$$

und weil

$$\Im(A\varphi, B\varphi) = \frac{1}{2i}\Big(\, (A\varphi, B\varphi) - (B\varphi, A\varphi) \,\Big) = \frac{1}{2i}(\varphi, (AB - BA)\varphi) \, ,$$

und $\|A\varphi\|^2 = (A\varphi, A\varphi) = (\varphi, A^2\varphi)$, folgt

$$\frac{1}{4}\, |\langle i[A,B]\rangle_\varphi|^2 \leq \langle A^2\rangle_\varphi \langle B^2\rangle_\varphi \, .$$

Ersetzt man A durch $A - \langle A\rangle_\varphi \mathbb{1}$ und B durch $B - \langle B\rangle_\varphi \mathbb{1}$, ändert sich die linke Seite nicht, und rechts erhält man $(\Delta_\varphi A)^2(\Delta_\varphi B)^2$. ■

Wählt man speziell $A = P$ und $B = Q$, also $[P,Q] = \frac{\hbar}{i}\mathbb{1}$, so kann die rechte Seite der Ungleichung in Satz 4.5.13 nie null werden, man erhält die Heisenbergsche Unschärfe-relation

$$\Delta P \Delta Q \geq \frac{\hbar}{2} \, .$$

E 3. Die Unschärfen, die von der Nicht-Vertauschbarkeit herrühren, können noch verstärkt werden, wenn man — wie beim Übergang von der klassischen Mechanik zur klassischen Statistischen Mechanik — gemischte Zustände betrachtet (das muß man spätestens dann tun, wenn man Quantenstatistik betreibt): mit Wahrscheinlichkeit r_i liege der Zustand ψ_i vor, d. h.

$$\langle A\rangle = \sum r_i \langle A\rangle_{\psi_i}, \quad r_i \geq 0, \; \sum r_i = 1 \, .$$

Man kann zeigen, daß man die ψ_i ohne Einschränkung paarweise orthogonal (und normiert) wählen kann. Dann kann man den Operator $\rho := \sum r_i P_{\psi_i}$ definieren, der die Wahrscheinlichkeiten r_i als Eigenwerte hat und in der o. n. Basis $\{\psi_1, \psi_2, \ldots\}$ diagonal ist. $\sum r_i$ ist dann die „Spur" der „Diagonalmatrix" ρ. Was „Spuroperatoren" sind, kann man allgemein definieren; und damit schreibt man dann

$$\mathrm{Spur}\, \rho = 1, \quad \langle A\rangle = \mathrm{Spur}\, (\rho A) \, .$$

Doch diese Zustände werden Ihnen frühestens in der QM II-Vorlesung begegnen.

E 4. Ich habe von der (unendlichen) „Matrix" ρ gesprochen. Vielfach werden Operatoren als unendliche Matrizen betrachtet: man nehme eine o. n. Basis $\{\varphi_i\}$; wenn $\psi = \sum a_i \varphi_i$ und $\chi = \sum b_i \varphi_i$, dann füge man in den Ausdruck $(\psi, A\chi)$ vor und hinter A die Identität $\sum P_{\varphi_i} = \mathbb{1}$ ein:

$$(\psi, A\chi) = \sum_{i,j} (\psi, \varphi_i)(\varphi_i, A\varphi_j)(\varphi_j, \chi) = \sum_{i,j} \overline{a_i} A_{ij} b_j \, ,$$

wobei $A_{ij} = (\varphi_i, A\varphi_j)$ ist. Alle Größen $(\psi, A\chi)$, die A ja vollständig festlegen, lassen sich also mit der unendlichen Matrix (A_{ij}) ausdrücken. Diese Art von Darstellung ist in vielen Lehrbüchern anzutreffen, insbesondere, wenn die Dirac-Schreibweise benutzt wird. Sie ist gut und schön für beschränkte Operatoren; bei unbeschränkten Operatoren muß man natürlich aufpassen: sie macht dann und nur dann Sinn, wenn die φ_i im Definitionsbereich von A liegen.

Genaugenommen braucht man sogar noch mehr: $\mathcal{D}(A) = \text{L.H.}\,\{\varphi_i\}$. A mag auf weiteren Elementen definiert sein, doch kann man die Matrixdarstellung für sie nicht benutzen. Grund: ein unbeschränkter linearer Operator ist nirgends stetig (aus der Stetigkeit bei irgendeinem $\psi \in \mathcal{D}(A)$ folgt wegen der Linearität Stetigkeit überall); daher definiert $A\Big|_{\text{L.H.}\{\varphi_i\}}$ den Operator *nicht* auf Limeselementen.

Wenn man die Dirac-Schreibweise benutzt, geht man noch weiter: man verwendet uneigentliche Eigenvektoren zu uneigentlichen Eigenwerten, z. B.

$$Q|\,x\rangle = x|\,x\rangle, \quad \text{mit} \quad \langle x|\,x'\rangle = \delta(x - x')\,.$$

Ein „Wellenpaket“ baut man dann als überabzählbare Überlagerung: $\psi_f = \int f(x)|\,x\rangle\,dx$ mit $f \in \mathcal{L}^2$; dann ist

$$(\psi_f, \psi_g) = \int dx\,dy\,\overline{f(x)}g(y)\langle x|\,y\rangle = \int dx\,\overline{f(x)}g(y)\,.$$

Sei nun $\psi_g \in \mathcal{D}(A)$:

$$(\psi_f, A\psi_g) = \int dx\,dy\overline{f(x)}\langle x|A|\,y\rangle g(y)\,;$$

A wird also dargestellt durch den Integralkern $\langle x|A|\,y\rangle$, wobei das letzte Integral definiert ist für geeignete g, nämlich solche, für die $\psi_g \in \mathcal{D}(A)$. (Der Integralkern wird im allgemeinen eine Distribution sein.)

E 5. Ich habe den Unterschied zwischen symmetrisch und selbstadjungiert betont. Daß das nicht nur mathematische Spitzfindigkeit ist, will ich illustrieren.
Sei $\mathcal{H}_1 = \mathcal{L}^2([0,1])$. Ich betrachte

$$P_0 = \frac{1}{i}\partial_x, \qquad \mathcal{D}(P_0) = \mathcal{D}_0 := \{\psi \in \mathcal{C}^\infty;\ \psi(0) = \psi(1) = 0\}\ .$$

Die Menge $\mathcal{D}_0$ ist dicht in $\mathcal{H}_1$, und P_0 ist hermitisch (also symmetrisch):

$$(\varphi, P_0\psi) = \int_0^1 \overline{\varphi(x)}\frac{1}{i}\psi'(x)\,dx = \frac{1}{i}\overline{\varphi(x)}\psi(x)\,\Big|_0^1 + \int_0^1 \overline{\left(\frac{1}{i}\varphi'(x)\right)}\psi(x)\,dx\,, \tag{5.14}$$

der erste Term der rechten Seite verschwindet wegen der Randbedingungen für φ und ψ, der zweite ist $= (P_0\varphi, \psi)$.
P_0 hat keine Eigenvektoren in $\mathcal{D}_0$:

$$\frac{1}{i}\partial_x\psi = a\psi \tag{5.15}$$

hat die Lösung $\psi(x) = \text{const}\, e^{iax}$, doch damit sind die Randbedingungen nicht erfüllbar.
Suchen wir nach P_0^*. Damit der kritische erste Term auf der rechten Seite von (5.14) verschwindet, genügt es ja, daß $\psi \in \mathcal{D}_0$ ist, die rechte Seite ist dann stetig in ψ für alle φ, für die φ' als Vektor von $\mathcal{H}_1$ definierbar ist. (Dazu muß φ „absolut stetig“ sein, was ich hier nicht definieren will; dann kann man φ' als lokal Lebesgue-integrable Funktion definieren und verlangen, daß φ' quadratintegrabel ist; ich schreibe nur kurz $\varphi' \in \mathcal{H}_1$.) Daher ist nach Definition

$$\mathcal{D}(P_0^*) = \{\varphi \in \mathcal{H}_1;\ \varphi' \in \mathcal{H}_1\}\ , \quad P_0^* = \frac{1}{i}\partial_x\,.$$

Gleichung (5.15) zeigt, daß P_0^* alle $a \in \mathbb{C}$ als Eigenwerte hat, denn $(e^{iax})' = ia\, e^{iax} \in \mathcal{L}^2([0,1])$ für alle $a \in \mathbb{C}$. Es ist also sicher nicht mehr symmetrisch.
Versuchen wir also, den Definitionsbereich von P_0 vorsichtig zu vergrößern und damit den des adjungierten Operators zu verkleinern:

$$P_1 = \frac{1}{i}\partial_x, \quad \mathcal{D}_1 = \{\psi \in \mathcal{C}^\infty;\ \psi(0) = \psi(1)\}\ .$$

Der kritische Term in (5.14) lautet dann $-i(\overline{\varphi(1)} - \overline{\varphi(0)})\psi(0)$, er verschwindet genau dann für alle ψ, wenn $\varphi(0) = \varphi(1)$ gilt; und das weist darauf hin, daß P_1 eine selbstadjungierte Erweiterung hat, wir müssen die Einschränkung auf $\psi \in \mathcal{C}^\infty$ noch etwas lockern. Jetzt finden wir Eigenvektoren e^{iax}, wenn wir $a = 2\pi k,\ k \in \mathbb{Z}$ setzen, denn dann ist $e^{ia} = 1$. Die Menge $\left\{e^{i2\pi kx};\ k \in \mathbb{Z}\right\}$ ist eine totale Menge; daraus folgt mit Satz 4.5.11, daß P_1 wesentlich selbstadjungiert ist, also eine eindeutige selbstadjungierte Erweiterung hat ($\mathcal{D} = \left\{\sum a_k\, e^{i2\pi kx};\ \sum |ka_k|^2 < \infty\right\}$).
Das ist aber nicht die einzige Möglichkeit. Der kritische Term

$$\overline{\varphi(1)}\psi(1) - \overline{\varphi(0)}\psi(0)$$

verschwindet auch dann, wenn

$$\varphi, \psi \in \mathcal{D}^{(\gamma)} = \left\{\psi;\ \psi' \in \mathcal{H}_1 \text{ und } \psi(1) = e^{i\gamma}\psi(0)\right\}$$

mit reellem γ. Man überzeugt sich wie im Fall P_1, daß man für jedes γ eine selbstadjungierte Erweiterung bekommt (mit Eigenwerten $2\pi k - \gamma$). Wir haben mit P_0 ein Beispiel für einen symmetrischen Operator, der viele selbstadjungierte Erweiterungen hat. Die allgemeine Theorie sagt einem, daß man hier alle möglichen Erweiterungen gefunden hat.

Wo treten solche Hilbertraume auf? Immer dann, wenn das Teilchen oder die Teilchen in einem endlichen Volumen eingesperrt sind. Das tritt auf zum einen in der Statistischen Mechanik: man hat mit Systemen zu tun, die in einem (wenn auch sehr großen) Kasten eingeschlossen sind. Zum anderen können die Halbleiter-Mikroelektroniker inzwischen auch sogenannte quantum dots herstellen, kleine Bereiche in einem Chip, in die man kontrolliert wenige Elektronen hineinbringen und dort halten kann, also reale mikroskopische Kästen. Es sind natürlich nicht so sehr Eigenwerte von P von Interesse wie von $P^2/(2m)$ bzw. H, doch auch hier sind die Eigenwerte von den Randbedingungen bestimmt. Welche selbstadjungierte Erweiterung man zu wählen hat, muß man versuchen, bei Mikro-Kästen mit physikalischen Argumenten zu entscheiden; in der Statistischen Mechanik hängen die physikalisch interessanten Ergebnisse im allgemeinen letztlich nicht an der speziellen Wahl.

4.6 Spektrum, Resolvente, Spektralscharen

Vorbemerkungen: 1. Wenn A den Eigenwert 0 hat, $A\varphi = 0$ und $\varphi \neq 0$, existiert kein Inverses; denn der inverse Operator müßte ja den Nullvektor auf φ abbilden, aber ein linearer Operator bildet 0 auf 0 ab.
2. Wenn $A\varphi \neq 0 \; \forall \varphi \in \mathcal{D}_A$, wenn dabei aber $\|A\varphi\|$ für $\|\varphi\| = 1$ beliebig klein werden kann, gibt es zwar eine eindeutige Abbildung des Wertebereichs $\mathcal{W}_A := A\mathcal{D}_A$ von A auf $\mathcal{D}_A$,[9] d. h. ein dicht definiertes Inverses A^{-1}, wenn $\mathcal{W}_A$ dicht ist; aber A^{-1} ist unbeschränkt: es muß ja beliebig kleine Vektoren $A\varphi$ auf Vektoren φ der Norm 1 abbilden.

Wenn also A den Eigenwert a hat, existiert $(A - a\mathbb{1})^{-1}$ nicht. Beispiel Multiplikationsoperator Q: $\|(Q - x_0\mathbb{1})\varphi\|$ kann für $\|\varphi\| = 1$ beliebig klein werden, man wähle $\varphi(x)$ mit sehr kleinem Träger um x_0 herum. Eigenwerte und uneigentliche Eigenwerte sind also dadurch charakterisiert, daß es für $A - a\mathbb{1}$ kein beschränktes Inverses gibt.

Definition 4.6.1 *Die Punkte $\lambda \in \mathbb{C}$, für die $(A - \lambda\mathbb{1})^{-1}$ als beschränkter, in ganz $\mathcal{H}$ definierter Operator existiert, heißen* **reguläre** *Punkte von A, sie bilden die* **Resolventenmenge** $\mathsf{P}(A)$, *ihr Komplement heißt das* **Spektrum** *von A, ich bezeichne es mit* $\operatorname{Sp} A$.
Für $\lambda \in \mathsf{P}(A)$ heißt $R(\lambda; A) := (A - \lambda\mathbb{1})^{-1}$ die **Resolvente** *von A.*

Man überlegt sich sofort, daß aus der Definition des Inversen folgt, daß

$$\mathcal{D}_A = \mathcal{D}_{A-\lambda 1} = R(\lambda; A)\mathcal{H} \quad \text{für } \lambda \in \mathsf{P}(A)\,. \tag{6.2}$$

Das Spektrum eines Hilbertraum-Operators kann nicht leer sein, aber es kann aus einem einzigen Punkt bestehen (Beispiel: $\operatorname{Sp} \mathbb{1} = \{1\}$) oder im anderen Extremfall die ganze komplexe Ebene umfassen (Beispiel: der „Leiteroperator" $a = 2^{-1/2}(Q - iP)$ aus (5.2) hat alle $\lambda \in \mathbb{C}$ als Eigenwerte, ebenso das Adjungierte P_0^* aus den Ergänzungen des letzten Abschitts). Uns interessieren im Folgenden vorwiegend selbstadjungierte und unitäre Operatoren und deren Spektren, die (wie die Eigenwerte dieser Operatoren) auf der reellen Achse bzw. auf dem Einheitskreis liegen.

Satz 4.6.3 *(i)* $\operatorname{Sp} A$ *ist eine abgeschlossene Menge.*
(ii) Wenn A selbstadjungiert ist, liegt das Spektrum auf der reellen Achse.
(iii) Sei $A = A^*$. λ ist ein Element des Spektrums genau dann, wenn

$$\inf_{\varphi \in \mathcal{D}_A, \varphi \neq 0} \|(A - \lambda\mathbb{1})\varphi\| \,/\, \|\varphi\| = 0\,. \tag{6.4}$$

[9] Denn keine zwei φ´s können dasselbe Bild haben: $A\varphi_1 = A\varphi_2$ heißt $A(\varphi_1 - \varphi_2) = 0$, also $\varphi_1 - \varphi_2 = 0$.

(iv) Für $A = A^*$ und $\Im\lambda \neq 0$ gilt $\mathcal{W}_{A-\lambda\mathbb{1}} = \mathcal{H}$.

Die erste Aussage sollte man kennen; ich will sie hier aber nicht beweisen, ein Beweis für selbstadjungiertes A findet sich in Achieser-Glasmann.
Zum Beweis der weiteren Aussagen brauche ich das folgende

Lemma 4.6.5 *Der Wertebereich $\mathcal{W}_A$ von A und der Nullraum (= Kern) von A^* sind orthogonal; genauer: wenn A dicht definiert ist, gilt*

$$\ker A^* = (\mathcal{W}_A)^\perp ,$$

d. h. $\mathcal{H} = \ker A^* \oplus \overline{\mathcal{W}_A}$.

Beweis: Sei φ orthogonal zu $\mathcal{W}_A$. Dann ist die Abbildung $\psi \mapsto (\varphi, A\psi) = 0$ offensichtlich stetig in $\psi \in \mathcal{D}_A$, d. h. $\varphi \in \mathcal{D}_{A^*}$ und $0 = (\varphi, A\psi) = (A^*\varphi, \psi)$ für alle ψ aus der dichten Menge $\mathcal{D}_A$, also $A^*\varphi = 0$. Daß jedes $\varphi \in \ker A^*$ zu $\mathcal{W}_A$ orthogonal ist, folgt trivialerweise. ■

Beweis von Satz 4.6.3:
(i): siehe Literatur. Für selbstadjungiertes A kann man es aus (iii) herleiten.
Sei A im Folgenden selbstadjungiert. Wenn λ nicht Eigenwert von A ist, also $(A - \lambda\mathbb{1})\varphi \neq 0$ für alle nicht verschwindenden $\varphi \in \mathcal{D}_A$, dann bildet $A-\lambda\mathbb{1}$ den Definitionsbereich eineindeutig auf $\mathcal{W}_{A-\lambda\mathbb{1}}$ ab: zu $\psi \in \mathcal{W}_{A-\lambda\mathbb{1}}$ gibt es ein $\varphi \in \mathcal{D}_A$ mit $(A-\lambda\mathbb{1})\varphi = \psi$.[10]
Wenn weiterhin (6.4) nicht gilt, d. h. es gibt ein c, so daß

$$\|(A - \lambda\mathbb{1})\varphi\| \,/\, \|\varphi\| \geq c > 0 \quad \forall\ \varphi \in \mathcal{D}_A,\ \varphi \neq 0\,, \tag{6.6}$$

folgt $\|(A - \lambda\mathbb{1})^{-1}\psi\| = \|\varphi\| \leq (1/c)\, \|(A - \lambda\mathbb{1})\varphi\| = (1/c)\, \|\psi\|$; Konsequenz:

$$\text{Wenn (6.6) gilt, ist } (A - \lambda\mathbb{1})^{-1} \text{ beschränkt auf } \mathcal{W}_{A-\lambda\mathbb{1}}. \tag{6.7}$$

Triviale Bemerkung: wenn (6.6) gilt, kann λ nicht Eigenwert sein.
Sei nun $\lambda = \xi + i\eta,\ \eta \neq 0$, dann ist (weil auch $A - \xi\mathbb{1}$ selbstadjungiert ist)

$$\begin{aligned} \|(A - \lambda\mathbb{1})\varphi\|^2 &= \|(A - \xi\mathbb{1})\varphi - i\eta\varphi\|^2 \\ &= \|(A - \xi\mathbb{1})\varphi\|^2 + i\eta(\varphi, (A - \xi\mathbb{1})\varphi) - i\eta((A - \xi\mathbb{1})\varphi, \varphi) + \eta^2\, \|\varphi\|^2 \\ &= \|(A - \xi\mathbb{1})\varphi\|^2 + \eta^2\, \|\varphi\|^2 \geq \eta^2\, \|\varphi\|^2 \ . \end{aligned}$$

Daraus folgt (mit $c = |\eta|$)

$$\text{Wenn } \Im\lambda \neq 0 \text{ ist, gilt (6.6).} \tag{6.8}$$

[10] Argumentation wie in Fußnote 9.

Sei $\Im\lambda \neq 0$, dann ist $\overline{\lambda}$ nicht Eigenwert von A, weil die Eigenwerte nur reell sein können. Sei $\lambda = \overline{\lambda}$ und (6.6) gelte, dann kann $\overline{\lambda}$ auch nicht Eigenwert sein. In beiden Fällen folgt daher $\ker(A - \lambda\mathbb{1})^* = \ker(A - \overline{\lambda}\mathbb{1}) = \{0\}$, d. h. nach Lemma 4.6.5:

$$\overline{\mathcal{W}_{A-\lambda 1}} = \mathcal{H} \quad \text{wenn (6.6) gilt.} \tag{6.9}$$

Aus (6.8) und (6.7) wissen wir also, daß in diesen beiden Fällen $(A-\lambda\mathbb{1})^{-1}$ beschränkt ist auf der Menge $\mathcal{W}_{A-\lambda 1}$, und diese Menge ist nach (6.9) dicht in $\mathcal{H}$; außerdem ist nach Satz 4.3.9 A ein abgeschlossener Operator. Ich behaupte, daß daraus folgt:

$$\mathcal{W}_{A-\lambda 1} = \mathcal{H} \quad \text{wenn (6.6) gilt.} \tag{6.10}$$

Denn sei $\psi \in \mathcal{H}$ beliebig, dann existiert $\psi_\alpha \in \mathcal{W}_{A-\lambda 1}$ mit $\lim \psi_\alpha = \psi$. Wegen (6.7) existiert auch $\lim \varphi_\alpha \equiv \lim(A - \lambda\mathbb{1})^{-1}\psi_\alpha =: \varphi$. φ_α ist aus $\mathcal{D}_A$ (siehe Gleichung (6.2)), $A\varphi_\alpha = ((A - \lambda\mathbb{1}) + \lambda\mathbb{1})\varphi_\alpha = \psi_\alpha + \lambda\varphi_\alpha$ konvergiert ebenfalls: $A\varphi_\alpha \to \psi + \lambda\varphi$. Weil A abgeschlossen ist, folgt $\varphi \in \mathcal{D}_A$ und $A\varphi = \psi + \lambda\varphi$, daher ist $\psi = (A - \lambda\mathbb{1})\varphi$, also gilt (6.10). Zusammen mit (6.7) bedeutet das, daß

$$\lambda \in \mathsf{P}(A) \quad \text{wenn (6.6) gilt.} \tag{6.11}$$

Wenn also $\Im\lambda \neq 0$ ist, folgt mit (6.8) und (6.11): $\lambda \in \mathsf{P}(A) \Longrightarrow \operatorname{Sp} A \subset \mathbb{R}$, d. h. die Aussage (ii). Umgekehrt gilt (6.6), wenn $\lambda \in \mathsf{P}(A)$ (denn dann existiert $(A - \lambda\mathbb{1})^{-1}$ als beschränkter Operator auf ganz $\mathcal{H}$; man benutze (6.2) und eine Rechnung wie für (6.7)). Mit anderen Worten:

$$\inf_{\varphi\in\mathcal{D}_A,\varphi\neq 0} \|(A - \lambda\mathbb{1})\varphi\| \,/\, \|\varphi\| \geq c > 0 \quad \Longleftrightarrow \quad \lambda \in \mathsf{P}(A)\,.$$

Das ist aber gleichbedeutend mit der Aussage (iii) des Satzes.
(iv) ist eine Konsequenz von (6.10). ■
Für unitäre Operatoren, $U^* = U^{-1}$, gilt der Satz analog; statt $\mathbb{R}$ lese man $C_1 := \{z \in \mathbb{C};\ |z| = 1\}$ (und statt $\Im z \neq 0$: $z \notin C_1$).

Bei selbstadjungierten (und bei unitären) Operatoren kann man das Spektrum in zwei Klassen teilen: das **diskrete** oder **Punktspektrum**, das aus allen Eigenwerten besteht, und den Rest, der bei diesen Operatoren das **kontinuierliche Spektrum** bildet. Wenn Sie an das Beispiel Q denken, wird diese Bezeichnung plausibel: Q ist selbstadjungiert, also $\operatorname{Sp} Q \subset \mathbb{R}$, da für alle $x \in \mathbb{R}$ das Inverse $(Q - x\mathbb{1})^{-1}$ unbeschränkt ist, folgt $\operatorname{Sp} Q = \mathbb{R}$; das Spektrum ist in diesem Fall eine kontinuierliche Menge, die Menge der uneigentlichen Eigenwerte. (Dasselbe gilt für P: Ü.)
Umgekehrt: wenn a ein Element des kontinuierlichen Spektrums eines selbstadjungierten Operators A ist, gibt es nach Satz 4.6.3(iii) Vektoren φ mit $\|\varphi\| = 1$, für die $\|(A - a\mathbb{1})\varphi\| \approx 0$ ist, d. h. $(A - a\mathbb{1})\varphi \approx 0$ oder $A\varphi \approx a\varphi$; φ ist dann ein „Beinahe-Eigenvektor“ zum „Beinahe-Eigenwert“ a. Das werde ich mithilfe der Spektralzerlegung später genauer darstellen.

Die obige Definition des kontinuierlichen Spektrums als Rest nach Herausnehmen der Eigenwerte ist nicht die allgemeine Definition, sondern eine Vereinfachung, die für selbstadjungierte und unitäre Operatoren gilt; im allgemeinen ist es dadurch charakterisiert, daß $(A - \lambda \mathbb{1})^{-1}$ unbeschränkt, aber dicht definiert ist. Für andere Operatoren gibt es neben dem diskreten und dem kontinuierlichen Spektrum noch einen dritten Anteil, der möglicherweise nicht leer ist, das Residualspektrum; ich gehe nicht darauf ein, weil es im Folgenden nicht vorkommt.
Zur Bedeutung des Spektrums für die Quantenmechanik: die Elemente des Spektrums eines selbstadjungierten Operators A sind die möglichen (beliebig) scharfen Meßwerte dieser Observablen; „beliebig scharf" bezieht sich darauf, daß man in der Quantenmechanik viele Einzelmessungen durchführen muß, um Mittelwerte zu erhalten, die mit den Wahrscheinlichkeitsaussagen der Theorie verglichen werden können. Die Werte, die sich bei einer Einzelmessung ergeben können, sind auch gerade die Punkte des Spektrums.

Man darf sich von der Bezeichnung „kontinuierliches Spektrum" und dem obigen Beispiel nicht irreführen lassen. Wenn P_n paarweise orthogonale Projektoren sind, für die $\sum_{n=1}^{\infty} P_n = \mathbb{1}$ gilt, ist $A = \sum_{n=1}^{\infty} \frac{1}{n} P_n$ ein selbstadjungierter Operator, der die Menge $\{1/n;\ n \in \mathbb{N}\}$ als Punktspektrum hat. $\operatorname{Sp} A$ ist abgeschlossen, der Nullpunkt gehört nicht zu den Eigenwerten, ist aber Häufungspunkt, muß also auch zum Spektrum gehören; er bildet das kontinuierliche Spektrum. Man kann sich direkt überlegen, daß 0 zum kontinuierlichen Spektrum gehört: seien $\varphi_n \in P_n \mathcal{H}$ und $\|\varphi_n\| = 1$, dann ist $\|A\varphi_n\| = 1/n$ beliebig klein, A^{-1} kann daher nicht beschränkt sein; jedes $\psi \in$ L.H. $\{\varphi_n;\ n \text{ groß}\}$ ist dann Beinahe-Eigenvektor zu 0.
Andere Punkte kann es im kontinuierlichen Spektrum nicht geben; denn ganz allgemein gilt

Lemma 4.6.12 *Sei $A = \sum a_n P_n$ mit paarweise orthogonalen Projektoren P_n, $\mathcal{D}_A = \left\{\varphi \in \mathcal{H};\ \sum |a_n|^2 (\varphi, P_n \varphi) < \infty\right\}$. Sein Spektrum besteht aus den Eigenwerten $\{a_n\}$ und deren Häufungspunkten.*

Beweis: Wir können ohne Einschränkung annehmen, daß $\sum P_n = \mathbb{1}$, andernfalls ist $P_0 \equiv \mathbb{1} - \sum P_n$ Eigenprojektor zum Eigenwert 0 und wir können ihn zur Summe hinzunehmen. Für $\varphi \in \mathcal{D}_A$ ist

$$\begin{aligned} \|(A - \lambda\mathbb{1})\varphi\|^2 &= \left\| \sum (a_n - \lambda) P_n \varphi \right\|^2 = \sum |a_n - \lambda|^2 \|P_n \varphi\|^2 \\ &\geq \inf_n |a_n - \lambda|^2 \|\varphi\|^2 \ ; \end{aligned}$$

und wenn $\lambda \notin \overline{\{a_n\}}$ ist das Infimum positiv, $\|(A - \lambda\mathbb{1})\varphi\| \geq \text{const} > 0$, also λ ein regulärer Punkt, denn $(A - \lambda\mathbb{1})^{-1}$ ist beschränkt auf $\mathcal{W}_{A-\lambda 1}$ und dieser Bereich ist dicht in $\mathcal{H}$, weil L.H. $\bigcup_n P_n \mathcal{H} \subset \mathcal{W}_{A-\lambda 1}$. ■

Wenn A selbstadjungiert und beschränkt ist, gilt $A \leq \|A\| \mathbb{1}$, denn $(\varphi, A\varphi)$ ist reell und $\leq \|A\| (\varphi, \varphi)$. Dann ist $B := a\mathbb{1} - A > c\mathbb{1}$, wenn $a > c + \|A\|$ ist, $c > 0$; folglich

$$\frac{\|(A - a\mathbb{1})\varphi\|}{\|\varphi\|} = \frac{\|B\varphi\|}{\|\varphi\|} \geq \frac{(\varphi, B\varphi)}{(\varphi, \varphi)} \geq c > 0\,.$$

Daher kann $\|(A - a\mathbb{1})\varphi\| \,/\, \|\varphi\|$ nicht beliebig klein werden, a ist regulärer Punkt von A. Das gleiche gilt für jedes $a < -c - \|A\|$. c kann beliebig klein gewählt werden, daher ist

$$\operatorname{Sp} A \subset [-\|A\|\,, \|A\|]\,. \tag{6.13}$$

Als weiteres Beispiel erwähne ich das Spektrum eines Projektors P: er hat die Eigenwerte 0 und 1 mit den zugehörigen Eigenprojektoren P und $\mathbb{1} - P$, die sich zu $\mathbb{1}$ addieren; weitere Punkte kann das Spektrum nach Lemma 4.6.12 nicht enthalten.

Der selbstadjungierte Operator $A = \sum a_n P_n$ mit reellen a_n aus Gleichung (5.9) hat das Punktspektrum $\{a_n;\ n \in \mathbb{N}\}$, sein kontinuierliches Spektrum besteht nur aus den Häufungspunkten des Punktspektrums.
Die Darstellung $\sum a_n P_n$ ist in diesem Fall eine Form der Spektralzerlegung von A. Wir nehmen Œ wieder an, daß $\sum P_n = \mathbb{1}$ ist. Weiterhin sei Œ $a_n < a_{n+1},\ n$ durchlaufe eine Indexmenge I, z. B. auch $I = \mathbb{Z}$. Um sie in der allgemeinen Form zu schreiben, führen wir die Projektoren $Q_k := \sum_{n=\leq k} P_n$ ein. Das ist eine wachsende Schar von Projektoren; ihre Unterräume sind ineinander geschachtelt und schöpfen ganz $\mathcal{H}$ aus. Ich kann sie auffassen als Größen, die eine projektorwertige Stufenfunktion charakterisieren:

$$\lambda \in \mathbb{R} \mapsto E_\lambda := Q_{k-1}, \quad \text{falls } a_{k-1} < \lambda \leq a_k\,. \tag{6.14}$$

Diese Stufenfunktion macht immer dann einen Sprung, wenn λ einen der Eigenwerte a_k erreicht hat und überschreitet; die „Größe" des Sprungs ist gleich dem zu diesem Eigenwert gehörenden Eigenprojektor P_k. Man beachte:

$$E_\lambda \xrightarrow{s} \mathbb{1} \ \text{ für } \lambda \to \infty\,.$$

Statt der obigen Darstellung kann ich auch schreiben

$$A = \sum_k a_k P_k = \sum a_k (Q_k - Q_{k-1}) = \int \lambda\, dE_\lambda\,. \tag{6.15}$$

Dieses Integral ist als Riemann-Stieltjes-Integral im Sinne der starken Konvergenz auf allen $\varphi \in \mathcal{D}_A$ zu verstehen (und das müssen Sie gut verstehen, weil wir es im Folgenden verallgemeinern werden).
Man mache es sich am besten klar, indem man zunächst gewöhnliche Funktionen betrachtet: $\int f(x)\, de(x)$, wobei f stetig und $e(x)$ eine Stufenfunktion ist, deren Niveaus

von 0 bis 1 ansteigen: $e(x) = q_{k-1}$, wenn $a_{k-1} < x \leq a_k$ ist; bei a_k macht $e(x)$ einen sprung der Größe $q_k - q_{k-1}$. Dann ist

$$\begin{aligned}\int f(x)\,de(x) &= \sum_k \int_{a_k-\varepsilon}^{a_k+\varepsilon} f(x)\,de(x) \\ &= \sum_k \int f(x)(q_k - q_{k-1})\delta(x-a_k)\,dx = \sum_k f(a_k)(q_k - q_{k-1})\,.\end{aligned}$$

Man kann das Integral ohne Benutzung der δ-Funktion durch Approximationssummen

$$\sum f(\xi_j)(e(x_j) - e(x_{j-1}))\,, \quad x_{j-1} \leq \xi_j \leq x_j$$

berechnen; bei einer günstigen Aufteilung des Integrationsintervalls und passender Wahl der Zwischenwerte ξ_j wird der Limes bereits bei endlicher Feinheit der Aufteilung erreicht: man wähle $x_j = \frac{1}{2}(a_{j+1} + a_j)$, $\xi_j = a_j$, dann ist $e(x_j) = q_j$ und die obige Summe $= \sum f(a_j)(q_j - q_{j-1})$. Eine weitere Verfeinerung bringt nichts Neues, weil die Differenzen $(e(x_j) - e(x_{j-1})$ großenteils verschwinden.
Ich habe etwas lax argumentiert: man sollte eine Einteilung in endlich viele Intervalle vornehmen; dann gilt diese Überlegung für den Fall endlich vieler Sprungstellen. Doch auch bei unendlich vielen Sprungstellen, die sich im Endlichen oder bei $\pm\infty$ häufen können, kann man, ausgehend von endlicher Intervall-Einteilung, den Limes erreichen, vorausgesetzt, die resultierende Summe $\sum f(a_j)(q_j - q_{j-1})$ konvergiert.
Genauso geht man vor, wenn man es mit dem projektorwertigen Maß E_λ statt des reellwertigen Maßes $e(x)$ zu tun hat. Man teilt die λ-Achse in endlich viele Intervalle ein und berechnet für jedes $\varphi \in \mathcal{D}_A$ den Limes der Approximationssummen

$$\sum_j \lambda_j\,(E_{\lambda_j} - E_{\lambda_{j-1}})\varphi\,.$$

Im vorliegenden Fall kann man den Limes genauso berechnen wie für $\int f\,de$, man schreibe λ_j statt x_j, λ statt $f(x)$ und $(E_{\lambda_j} - E_{\lambda_{j-1}})\varphi$ statt $(e(x_j) - e(x_{j-1})$.
Man kann im allgemeinen nur Konvergenz im Sinne der starken Topologie erwarten. Wesentlich für die Konvergenz ist, daß man mit jeder Aufteilung $\{\lambda_j\}$ eine Folge paarweise orthogonaler Projektoren $\hat{P}_j \equiv E_{\lambda_j} - E_{\lambda_{j-1}}$ bekommt, deren Summe $= \mathbb{1}$ ist! Dann ist $\left\|\hat{P}_j\varphi\right\|$ für große $|\lambda_j|$ für jedes φ irgendwann sehr klein. Nur wenn man es mit beschränkten oder unitären Operatoren zu tun hat, kann man Norm-Konvergenz erwarten.

Man führt nun allgemeinere projektorwertige Funktionen mit diesen Eigenschaften ein:

Definition 4.6.16 *$I = [a,b]$ sei ein Intervall der erweiterten reellen Achse, d. h. $a = -\infty$ und $b = \infty$ sind zugelassen. Eine einparametrige Schar von Projektoren $\{E_\lambda,\ \lambda \in I\}$, die den folgenden Bedingungen genügt, heißt* **Zerlegung der Einheit***:*

(1) $E_a = 0, \quad E_b = \mathbb{1}\,,$

(2) $E_\kappa E_\lambda = E_{\min(\kappa,\lambda)}$,

(3) E_λ *ist im Sinne der starken Konvergenz halbstetig von links:*

$$s\text{-}\lim_{\delta\lambda\searrow 0} E_{\lambda-\delta\lambda} = E_\lambda\,.$$

Die Bedingung (2) ist äquivalent zu

$$(2') \qquad E_\kappa \le E_\lambda\,, \quad \text{wenn } \kappa < \lambda\,.$$

Noch anders gesagt: die zu E_λ gehörenden Unterräume bilden eine aufsteigende Folge, die im Limes $\lambda \to b$ ganz $\mathcal{H}$ ausschöpft.
Die in Gleichung (6.14) definierte Schar erfüllt genau diese Bedingungen. Unser Ziel ist es, zu zeigen, daß man zu jedem unitären und zu jedem selbstadjungierten Operator genau eine Zerlegung der Einheit finden kann, so daß man eine Integraldarstellung wie auf der rechten Seite von (6.15) angeben kann. Diese Zerlegung der Einheit heißt dann die **Spektralschar** des Operators.[11]

Wir werden sehen: bei jedem Eigenwert des Operators hat die zugehörige Spektralschar einen Sprung, für reguläre Punkte κ ist E_λ in einer Umgebung von κ konstant (und folglich $dE_\lambda = 0$). Da es Operatoren gibt, die ein ausgedehntes kontinuierliches Spektrum haben, brauchen wir auch stetig anwachsende Stücke der Spektralschar; das ist möglich (auch im separablen Hilbertraum).

Damit Sie ein bißchen mit dieser Art von Integral vertraut werden, will ich den Konvergenzbeweis für das Integral $\int f(\lambda)\,dE_\lambda$ für den einfachen Fall andeuten, daß $f(x)$ eine beschränkte, stetige Funktion und das Integrationsintervall $[a,b]$ endlich ist. (Der Träger von dE_λ kann größer sein, d. h. E_λ braucht außerhalb von $[a,b]$ nicht konstant zu sein.) Nach Definition ist

$$\int_a^b f(\lambda)\,dE_\lambda = \lim \sum_Z f(\tilde\lambda_j)(E_{\lambda_j} - E_{\lambda_{j-1}})\,, \quad \lambda_{j-1} \le \tilde\lambda_j \le \lambda_j\,, \tag{6.17}$$

wobei der Limes immer feinerer Zerlegungen Z von $[a,b]$ gemeint ist. Entscheidend bei der Berechnung solcher Integrale ist, daß die Projektoren $\hat P_j := E_{\lambda_j} - E_{\lambda_{j-1}}$ paarweise orthogonal sind: $\hat P_j \hat P_k = \delta_{jk}\hat P_j$ (das folgt aus der Eigenschaft (2) von Definition 4.6.16); und daher ist

$$\left\|\sum a_j \hat P_j \varphi\right\|^2 = \sum |a_j|^2\ (\varphi, \hat P_j\varphi)\,. \tag{6.18}$$

Seien Z_1, Z_2 zwei Zerlegungen, dann gibt es eine gemeinsame Verfeinerung Z; und

$$\left\|(\sum_{Z_1}\ldots - \sum_{Z_2}\ldots)\varphi\right\| \le \left\|(\sum_{Z}\ldots - \sum_{Z_1}\ldots)\varphi\right\| + \left\|(\sum_{Z}\ldots - \sum_{Z_2}\ldots)\varphi\right\|\,.$$

[11] Ob man in (3) Halbstetigkeit von links oder von rechts verlangt, ist Konvention; viele Lehrbücher bevorzugen links (aber aufpassen beim Literaturstudium).

Sei $Z_1 = \{\lambda_i\}$, $Z = \{\lambda_{ik}\}$ mit $\lambda_{i-1} < \lambda_{ik} \le \lambda_i,\ k = 1, \ldots, m,\ \lambda_{i,m} = \lambda_i,\ \lambda_{i,0} = \lambda_{i-1}$.

$$\begin{aligned}\left\|\Big(\sum_Z \ldots - \sum_{Z_1} \ldots\Big)\varphi\right\|^2 &= \left\|\sum_i \sum_k (f(\tilde\lambda_{ik}) - f(\tilde\lambda_i))(E_{\lambda_{ik}} - E_{\lambda_{i,k-1}})\varphi\right\|^2 \\ &= \sum_i \sum_k \left|f(\tilde\lambda_{ik}) - f(\tilde\lambda_i)\right|^2 (\varphi, \hat P_{ik}\varphi)\,; \qquad (6.19)\end{aligned}$$

Hier habe ich $\hat P_{ik} = E_{\lambda_{ik}} - E_{\lambda_{i,k-1}}$ gesetzt und (6.18) ausgenutzt. f ist nach Voraussetzung stetig und beschränkt, Z enthält nur endlich viele Intervalle, daher können wir den ersten Faktor in (6.19) gleichmäßig in i, k durch ε abschätzen, wenn Z_1 fein genug gewählt ist; und

$$\sum_i \sum_k (\varphi, \hat P_{ik}\varphi) = \|\varphi\|^2 \ .$$

Die Normdifferenz der Approximationssummen wird also beliebig klein.
Wählt man $f(\lambda) = 1$, so ist die Approximationssumme in (6.17) $= E_b - E_a$, unabhängig von der Zerlegung. Ich muß genau genommen im Integral andeuten, wie ich die Grenzen meine: E_λ kann ja bei a oder b unstetig sein, dann muß ich sagen, ob $a+0$ oder $a-0$ (bzw. $b+0$ oder $b-0$) gemeint ist; z. B.

$$\int_{a+0}^b dE_\lambda = \lim_{\varepsilon\searrow 0}(E_b - E_{a+\varepsilon})$$

bzw.

$$\int_{a-0}^b dE_\lambda = \lim_{\varepsilon\searrow 0}(E_b - E_{a-\varepsilon}) = E_b - E_a$$

wegen der Halbstetigkeit von links. Ich schreibe für diese Differenz kurz

$$\int_\Delta dE_\lambda \equiv E(\Delta)\,,$$

wobei Δ offen, halboffen oder abgeschlossen sein kann.
Mit $\Delta = \mathbb{R}$ folgt wegen $E_{-\infty} = 0$ und $E_\infty = \mathbb{1}$

$$\int dE_\lambda = \mathbb{1} \qquad (6.20)$$

und

$$\|\varphi\|^2 = \Big(\varphi, \int dE_\lambda \varphi\Big) = \int d(\varphi, E_\lambda\varphi)\,. \qquad (6.21)$$

(Hier benutzt man die Tatsache, daß das Skalarprodukt stetig ist, die Skalarproduktbildung ist also vertauschbar mit der Bildung des Limes von Approximationssummen, d. h. mit der Integration.)
Hieraus folgt sofort, daß $(\varphi, (\mathbb{1} - E(\Delta))\varphi) < \varepsilon$ für hinreichend großes Δ; weil auch $(\mathbb{1} - E(\Delta))$ Projektor ist, ist die linke Seite $= \|(\mathbb{1} - E(\Delta))\varphi\|^2$,d. h. es gilt

$$\text{s-}\lim_{\Delta\to\mathbb{R}} E(\Delta) = \mathbb{1}.$$

Es ist klar, daß hier keine Normkonvergenz der Operatoren gelten kann; $\mathbb{1} - E(\Delta)$ ist Projektor und hat daher die Norm 1.
Wenn f in Gleichung (6.17) nicht beschränkt ist, d. h. eine Singularität hat, kann ich auch bei Stetigkeit in (a, b) die Differenz der Funktionswerte in (6.19) nicht mehr gleichmäßig abschätzen; dann muß man dafür sorgen, daß φ geeignet gewählt wird, so daß die Faktoren $(\varphi, \hat{P}_{ik}\varphi)$ die schlechten Werte herunterdrücken: der Definitionsbereich ist dann nicht mehr ganz $\mathcal{H}$. Das gleiche gilt, wenn die Integration über ein unendliches Intervall läuft und z. B. $f(x) = x$ ist.
Da man die Spektralschar E_λ benutzt, um durch Integration weitere Operatoren zu bilden, spricht man auch vom *Spektralmaß* E_λ.
Ich wiederhole noch einmal: entscheidend ist, daß man für jede beliebige Aufteilung Projektoren $\hat{P}_j \equiv E_{\lambda_j} - E_{\lambda_{j-1}}$ erhält, die paarweise orthogonal sind und deren Summe $= 1$ ist. Im Geiste der obigen Rechnung kann man sauber begründen, daß man formal

$$(dE_{\lambda'}\psi, dE_\lambda\varphi) = (\psi, dE_{\lambda'}dE_\lambda\varphi) = (\psi, dE_\lambda\varphi)\delta(\lambda - \lambda')d\lambda' \tag{6.22}$$

schreiben darf, d. h.

$$\begin{aligned}\left(\int g(\lambda')\, dE_{\lambda'}\psi, \int f(\lambda)\, dE_\lambda\right)\varphi \;&=\; \int \overline{g(\lambda')} f(\lambda)(\psi, dE_\lambda\varphi)\delta(\lambda - \lambda')d\lambda' \\ = \int \overline{g(\lambda)} f(\lambda)(\psi, dE_\lambda\varphi) \;&=\; \int \overline{g(\lambda)} f(\lambda)\, d(\psi, E_\lambda\varphi)\,.\end{aligned} \tag{6.23}$$

Diese Gleichung gilt nach unseren bisherigen Rechnungen jedenfalls dann, wenn f und g stetig und beschränkt sind und das Integrationsintervall endlich ist. Auf der rechten Seite steht ein gewöhnliches Riemann-Stieltjes-Integral mit einem komplexwertigen Maß $\mu_{\psi,\varphi}(\lambda) = (\psi, E_\lambda\varphi)$. (Man kann es sogar auf eine Summe von Integralen mit positivem Maß zurückführen, wenn man $(\psi, E_\lambda\varphi)$ mit der Polarisationsidentität (I.1.17) als entsprechende Summe schreibt.) Ich habe hier wieder die Stetigkeit des Skalarprodukts ausgenutzt, um Integration und Skalarproduktbildung zu vertauschen.

Sei $A = \sum a_k P_k$, P_k paarweise orthogonale Projektoren, mit Definitionsbereich $\mathcal{D}_A = \{\psi;\ \sum |a_k|^2(\psi, P_k\psi)\}$, dann überzeugt man sich leicht — siehe Gleichung (4.8) —, daß

$$A = \text{s-}\lim_{n\to\infty} \sum_{k\le n} a_k P_k\,.$$

Die a_k brauchen hierbei nicht beschränkt zu sein.
Entsprechendes gilt, wenn man statt einer Summe ein Integral hat. Wenn das Integral $\int f(\lambda)\, dE_\lambda$ über ein unendliches Intervall zu erstrecken ist und f nicht beschränkt ist, braucht man einen geeigneten Definitionsbereich:

$$\mathcal{D}_f = \left\{\varphi;\ \int |f(\lambda)|^2\, d(\varphi, E_\lambda\varphi) < \infty\right\}\,. \tag{6.24}$$

Ich will im Folgenden annehmen, daß f stetig ist; Verallgemeinerungen sind unschwer möglich. Betrachten wir $\int_a^b f(\lambda)\, dE_\lambda$ für $a, b \to \infty$: nach (6.23) ist

$$\left\|\int_a^b f(\lambda)\, dE_\lambda\varphi\right\|^2 = \int_a^b |f(\lambda)|^2\, d(\varphi, E_\lambda\varphi)\,;$$

und dieses Integral wird für große a, b beliebig klein, wenn $\varphi \in \mathcal{D}_f$ gewählt wird, denn $\int_{\mathbb{R}} |f(\lambda)|^2 \, d(\varphi, E_\lambda \varphi)$ ist ja konvergent. Das gleiche gilt, wenn $a, b \to -\infty$, daher folgt für $I_R(f) \equiv \int_{-R}^{R} f(\lambda) \, dE_\lambda$:

$$\|(I_R(f) - I_{R'}(f)) \, \varphi\| \longrightarrow 0 \quad \text{für } R, \, R' \to \infty \, ,$$

d. h. auf $\mathcal{D}_f$ ist der Operator

$$\text{s-}\lim_{R \to \infty} I_R(f) =: \int_{\mathbb{R}} f(\lambda) \, dE_\lambda \tag{6.25}$$

wohldefiniert.
Daher darf man (6.23) für geeignete ψ und φ auch für unbeschränkte Integrationsintervalle benutzen.

Beispiel: Sei $\mathcal{H} = \mathcal{L}^2(\mathbb{R})$, die Multiplikation mit der charakteristischen Funktion χ_I eines Intervalls I ist ein orthogonaler Projektor P_I (Beweis trivial). Dann ist

$$E_\lambda : \quad (E_\lambda \psi)(x) := \chi_{(-\infty, \lambda)}(x) \psi(x) \tag{6.26}$$

eine stark stetige Zerlegung der Einheit (d. h. stetig im Sinne der starken Konvergenz). Sie ist die Spektralschar für den Multiplikationsoperator mit x:

$$Q = \int \lambda \, dE_\lambda \qquad \text{mit } E_\lambda \text{ aus (6.26)} \, . \tag{6.27}$$

Beweis: Ü.
Dieses Beispiel zeigt ein Verhalten, das der Anschauung ziemlich zuwider läuft und nur im unendlich-dimensionalen Raum denkbar ist: der zum Projektor P_I gehörige Unterraum ist für ein beliebig kleines Intervall I unendlich-dimensional! Denn offensichtlich ist $\chi_{I_1} \chi_{I_2} = 0$, wenn $I_1 \cap I_2 = \emptyset$ d. h. $P_{I_1}\mathcal{H} \perp P_{I_2}\mathcal{H}$; und jedes Intervall I der Länge ε kann man ja in unendlich viele disjunkte Intervalle mit Längen $\varepsilon/2, \varepsilon/4, \varepsilon/8, \ldots$ einteilen.

Ergänzung: Eine eindimensionale Karikatur des Wasserstoffatoms.
Der Atomkern wird als unendlich schwer und ruhend betrachtet, das Elektron ist quantenmechanisch zu beschreiben, es bewegt sich im Kraftfeld des Kerns. Die Anziehung zwischen Kern und Elektron sei gegeben durch ein sehr einfaches Potential:

$$V(x) = \begin{cases} 0 & \text{für } |x| > a \, , \\ -V_0 & \text{für } |x| < a \, . \end{cases}$$

Der zugehörige Hamiltonoperator ist $H = \frac{1}{2m} P^2 + V(Q) = -\frac{\hbar^2}{2m} \partial_x^2 + V(x)$. Wir fragen nach dem Spektrum von H.

Suchen wir zuerst nach Eigenwerten. Wir müssen also die „zeitunabhängige Schrödingergleichung“ $H\psi = E\psi$ lösen; anders: definieren wir drei Gebiete, in denen $V = \text{const}$ ist: $= 0$ in $G_l \cup G_r$, $= -V_0$ in G_0, wobei

$$G_l = \{x; \; x < -a\} \, , \quad G_0 = \{x; \; -a < x < a\} \, , \quad G_r = \{x; \; x > a\} \; ;$$

dann müssen wir fragen, für welche Werte von E die Gleichung

$$\psi'' = \varepsilon\psi \qquad \text{mit } \varepsilon = \begin{cases} -\frac{2m}{\hbar^2}\,E & \text{für } x \in G_l \cup G_r\,, \\ -\frac{2m}{\hbar^2}\,(E+V_0) & \text{für } x \in G_0 \end{cases} \tag{6.28}$$

eine Lösung in $\mathcal{L}^2(\mathbb{R})$ hat. In jedem der drei Gebiete hat (6.28) als gewöhnliche Differentialgleichung die Lösungen $e^{\pm\sqrt{\varepsilon}x}$, wobei der Exponent möglicherweise imaginär ist. Doch wir suchen eine Lösung in $\mathcal{L}^2(\mathbb{R})$, müssen also passende Lösungsstücke aneinandersetzen. Damit ψ' im Definitionsbereich von ∂_x ist, muß es stetig sein, dann ist erst recht ψ stetig. Wir müssen daher die vier Anschlußbedingungen

$$\psi \text{ stetig bei } \pm a, \quad \psi' \text{ stetig bei } \pm a \tag{6.29}$$

für die Teillösungen erfüllen.

Wenn $E > 0$ ist, ist $\sqrt{\varepsilon}$ in allen drei Gebieten imaginär, und mit $e^{\pm i\sqrt{|\varepsilon|}x}$ läßt sich kein $\psi \in \mathcal{L}^2$ bilden. Es gibt also keine positiven Eigenwerte.

Wenn $E < 0$ ist, ist $\sqrt{\varepsilon}$ in G_l und G_r reell. In G_l können wir nur die Lösung $C_l e^{+\sqrt{\varepsilon}x}$ brauchen, in G_r nur $C_r e^{-\sqrt{\varepsilon}}$, sonst bekommen wir kein quadratintegrables ψ. Die Konstanten C_l, C_r und zwei Konstanten $C_\pm$ für die Lösung $C_+ e^{\sqrt{\varepsilon}} + C_- e^{-\sqrt{\varepsilon}}$ in G_0 müssen wir so wählen, daß die vier Bedingungen (6.29) erfüllt sind. Doch die vier Konstanten müssen dann einem *homogenen* System von vier linearen Gleichungen genügen; die Bedingung, daß dessen Determinante verschwindet, ist nur für gewisse Werte von E erfüllt: das sind die Eigenwerte.

Man sieht leicht, daß es für $E < -V_0$ keine Lösung gibt: $\sqrt{\varepsilon}$ ist dann auch in G_0 reell, und mit einer Überlagerung zweier reeller Exponentialfunktionen kann man keine stetigen Anschlüsse an die Exponentialfunktionen außen erreichen, bei denen auch noch die Ableitungen stetig sind.

Es bleibt also der Bereich $-V_0 < E < 0$; das ist der Energiebereich, in dem ein klassisches Teilchen in einem solchen Potentialtopf gefangen wäre. Jetzt ist $\sqrt{\varepsilon}$ imaginär in G_0. Man rechnet in G_0 bequemer mit der Lösung $A \sin \kappa x + B \cos \kappa x,\ \kappa = \sqrt{|\varepsilon|} = \sqrt{2m(V_0+E)}/\hbar$ und erhält die beiden Möglichkeiten $A \neq 0,\ B = 0$ und $A = 0,\ B \neq 0$. Im ersten Fall muß gelten

$$k \tan ka = \kappa\,, \quad k = \frac{1}{\hbar}\sqrt{-2mE}\,,$$

im zweiten Fall

$$k \cot ka = \kappa\,.$$

(Die Rechnungen sind einfach; man rechnet besser nicht die Determinante aus, sondern löst direkt, wobei man $C_l = 1$ setzen kann; ein Normierungsfaktor ist ja immer frei.)

Diese Gleichungen kann man graphisch oder numerisch lösen. Die erste hat immer mindestens eine Lösung, die zweite dann, wenn der Potentialtopf hinreichend groß ist ($a\sqrt{2mV_0/\hbar}$ muß $\geq \pi/2$ sein); es gibt aber immer nur endlich viele Lösungen; unser

Karikatur-Atom hat im Gegensatz zum realen Wasserstoffatom nur endlich viele scharfe Energie-Niveaus.

Damit haben wir aber noch nicht das ganze Spektrum. Ich behaupte, daß alle Werte $E > 0$ zum kontinuierlichen Spektrum gehören. Das will ich beweisen, indem ich für jedes $E > 0$ eine Folge von Beinahe-Eigenvektoren angebe.
Sei ψ_R eine glatte Funktion, die $= 1$ ist für $|x| < R$, die verschwindet für $|x| > R+1$ und dazwischen glatt abfällt. So eine Funktion habe ich explizit in Gleichung (I.2.21) angegeben. Für sie gilt ${\psi_R}' = 0$ in $|x| < R$ und

$$\|{\psi_R}'\|\,,\ \|{\psi_R}''\| \leq C \quad \text{unabhängig von } R\,, \tag{6.30}$$

$$N_R^2 \equiv \|\psi_R\|^2 > R\,. \tag{6.31}$$

Behauptung:

$$\varphi_R(x) := \frac{1}{N_R}\psi_R(x)e^{ikx} \text{ ist Beinahe-Eigenvektor von } H \text{ zu } E = \frac{(\hbar k)^2}{2m}\,,$$

wenn R sehr groß ist.
Beweis: φ_R ist normiert, wir müssen zeigen, daß $\|(H-E)\varphi_R\|$ beliebig klein werden kann.

$$\begin{aligned}
\left\|{\varphi_R}'' - \left(\frac{2m}{\hbar^2}\right)V\varphi_R + \left(\frac{2m}{\hbar^2}\right)E\varphi_R\right\| &= \left\|\frac{1}{N_R}\left({\psi_R}'' + 2ik{\psi_R}' - \left(\frac{2m}{\hbar^2}\right)V\psi_R\right)e^{ikx}\right\| \\
&\leq \frac{1}{N_R}\left(\|{\psi_R}''\| + \|2k{\psi_R}'\| + \frac{2m}{\hbar^2}\|V\psi_R\|\right) \\
&\leq \frac{1}{\sqrt{R}}\left((1+2|k|)C + \frac{2m}{\hbar^2}\sqrt{2a}V_0\right) \quad \xrightarrow[R\to 0]{} 0\,.
\end{aligned}$$

Bei der letzten Ungleichung habe ich (6.30) und (6.31) und die Tatsache verwendet, daß $V = 0$ ist für $|x| > a$. ∎
Für $E < -V_0$ gibt es keine Eigenwerte, diese Werte liegen aber auch nicht im kontinuierlichen Spektrum von H, sondern in $\mathsf{P}(H)$. Denn $\frac{1}{2m}P^2$ ist ein positiver Operator, also ist $H > -V_0$, folglich $H - E > -V_0 - E > c > 0$ nach Voraussetzung. Daraus folgt $0 < c < (\varphi, (H-E)\varphi) \leq \|\varphi\|\,\|(H-E)\varphi\|$, und das impliziert nach Satz 4.6.3, daß $E \notin \operatorname{Sp} H$.

Zur physikalischen Bedeutung dieses Spektrums:
Die (volle zeitabhängige) Schrödingergleichung $i\hbar\partial_t\psi_t = H\psi_t$ wird gelöst durch

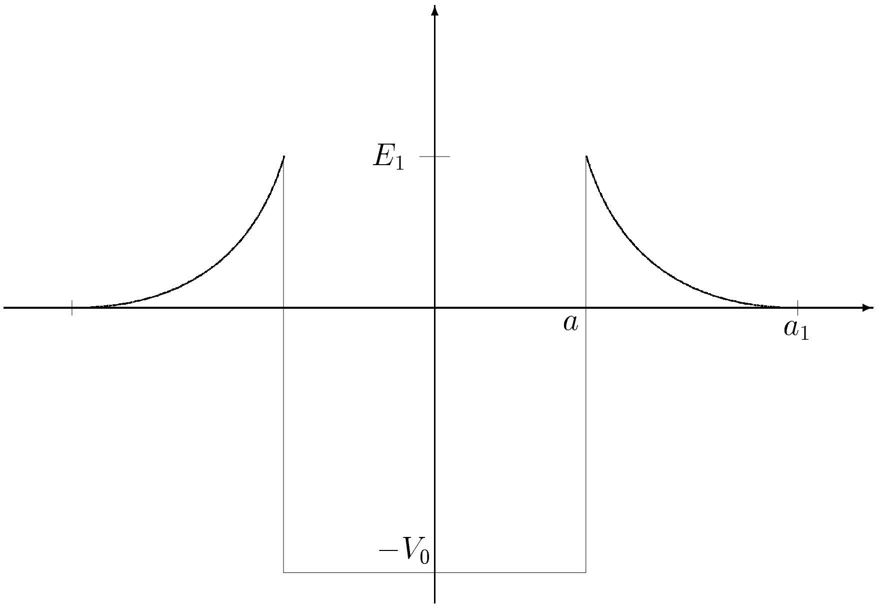
$$\psi_t = U_t\psi_0, \qquad U_t = e^{-\frac{i}{\hbar}Ht}\,.$$

Die Exponentialfunktion ist besonders bequem zu berechnen, wenn man sie auf einen Eigenvektor von H anwendet (man denke sie sich in eine Reihe entwickelt):

$$e^{-\frac{i}{\hbar}Ht}\varphi_E = e^{-\frac{i}{\hbar}Et}\varphi_E\,.$$

Setzt man also $\psi_0 = \varphi_E$, so ist die Zeitentwicklung von ψ_t trivial, es ändert sich nur ein Phasenfaktor, der in allen Erwartungswerten herausfällt. Das heißt, der dadurch definierte Zustand ist stationär, alle Erwartungswerte sind zeitunabhängig. Solchen stationären Zuständen entsprechen in der klassischen Mechanik gebundene Systeme wie z. B. Zentralstern – Planet. Details der Bahnbewegung („Gerade war das Elektron links oben, 3 Femtosekunden später rechts außen") kann die Quantenmechanik wegen der Unschärferelation prinzipiell nicht liefern. Man sieht an der Form des Zustandes in der Ortsdarstellung — exponentiell abfallend für große Enfernungen vom Zentrum, aber nicht $= 0$ —, daß es eine gewisse Wahrscheinlichkeit gibt, das Elektron außerhalb des klassisch erlaubten Bereichs zu finden; doch im wesentlichen ist es auf den Bereich des Potentialtopfs beschränkt.
Das ist ganz anders bei den Beinahe-Eigenvektoren, die zu positiven Energien gehören. Zum einen sind jetzt die Erwartungswerte bestenfalls *langsam* zeitlich veränderlich, aber nicht mehr zeitunabhängig, zum anderen ist der durch $\psi_R e^{ikx}$ gegebene Zustand räumlich weit ausgedehnt, das Elektron also nicht mehr an das Zentrum gebunden. Dafür hat er einen recht scharfen Impuls $\hbar k$. Diese φ_R sind Beinahe-Eigenvektoren für alle möglichen derartigen Hamiltonoperatoren, wenn nur das Potential von endlicher Reichweite und beschränkt ist; sie hängen gar nicht von der Form des Potentials ab. Wenn man der Faktor ψ_R wegläßt, sind es uneigentliche Eigenvektoren des freien Hamiltonoperators; doch weit weg vom Ursprung, wo das Potential verschwindet, sieht H ja aus wie der freie Operator. Die e^{ikx} bzw. daraus zusammengesetzte Wellenpakete beschreiben die Teilchen weit weg vom Zentrum.
Was hier physikalisch interessiert, sind Aussagen über die die Streuung von Elektronen an einem Kern: die Teilchen fliegen mit gut definiertem Impuls an, werden gestreut, und ihre Verteilung weit weg vom Streuzentrum wird registriert. Zur Berechnung der Streudaten benutzt man die uneigentlichen e^{ikx} bzw. aus ihnen zusammengesetzte Wellenpakete, die Streuzustände sind in gewisser Weise modifizierte φ_R's. Doch das ist nicht Thema dieser Vorlesung.

Denken wir uns das Potential ein wenig variiert:



In diesem Potential wäre ein klassisches Teilchen mit Energie $< E_1$ gebunden. Wir können mit den gleichen Argumenten wie oben zeigen, daß der Bereich $E > E_1$ zum kontinuierlichen Spektrum gehört; aber nicht nur das: offensichtlich bleibt die Argumentation auch für $0 < E < E_1$ richtig, φ_R ist Beinahe-Eigenvektor, wenn $R \gg a_1$. Ein Teilchen mit Energie in diesem Bereich ist also, entgegen dem klassischen Bild, nicht mehr fest gebunden: es kann den Potentialwall „durchtunneln"; unsere Überlegungen zum Spektrum liefern uns also hier den Hinweis auf den Tunneleffekt.

Zum Schluß noch eine mathematische Bemerkung. Ich habe stillschweigend vorausgesetzt, daß unser obiger Hamiltonoperator selbstadjungiert sei. Es ist klar, daß er symmetrisch ist. Die Selbstadjungiertheit dieses (und vieler anderen ähnlicher Operatoren) folgt aus einem sehr allgemeinen Satz von T. Kato, der besagt, daß $H_0 + V$ selbstadjungiert ist, wenn es für H_0 gilt (hier: $H_0 = P^2/(2m)$) und wenn V in bestimmtem Sinn durch H_0 beschränkt ist. Das ist zum Glück für viele Potentiale der Fall.

4.7 Funktionen unitärer Operatoren

Wenn man eine Darstellung $A = \int \lambda\, dE_\lambda$ hat, kann man (stetige) Funktionen von A definieren durch $f(A) := \int f(\lambda)\, dE_\lambda$. Das ist sinnvoll, man mache es sich im Fall $A = \sum a_n P_n$ klar: auf dem Unterraum $P_n\mathcal{H}$ (bzw. $E((\lambda, \lambda + \delta\lambda))\mathcal{H}$, $\delta\lambda$ klein,) ist $A = a_n \mathbb{1}$ (bzw. $A \approx \lambda\mathbb{1}$), d. h. $f(A)$ ist auf diesem Unterraum einfach $= f(a_n)P_n$ (bzw. $\approx f(\lambda)E((\lambda, \lambda + \delta\lambda))$). Aber zunächst haben wir diese Darstellung noch nicht und suchen für das Spektraltheorem zu einem selbstadjungierten A eine passende Zerlegung der Einheit $\lambda \mapsto E_\lambda$. E_λ kann als Funktion von A aufgefaßt werden:

$$E_\lambda = \int_{-\infty}^{\lambda} dE_{\lambda'} = \int \chi_{(-\infty,\lambda)}(\lambda')\, dE_{\lambda'} = \chi_{(-\infty,\lambda)}(A)\,.$$

Aber es ist nicht trivial, unstetige Funktionen von Operatoren zu definieren. Wir beginnen, weil es einfacher ist, mit unitären Operatoren.
Ziel dieses Abschnitts ist es, ausgehend von einem unitären Operator eine hinreichend große Klasse von operatorwertigen Funktionen zu bilden, in der auch Projektoren zu finden sind, mit denen wir eine passende Zerlegung der Einheit bilden können, die Spektralschar für den Ausgangsoperator ist.

Sei U unitär, dann ist $U^{-1} = U^*$ wohldefiniert und ich kann beliebige positive und negative Potenzen U^k, $k \in \mathbb{Z}$ bilden, es sind ebenfalls unitäre Operatoren. Dann sind auch beliebige „Quasipolynome"

$$\mathcal{Q}(U) = \sum_{k=-m}^{n} c_k\, U^k$$

wohldefiniert. Parallel dazu betrachten wir die entsprechenden Quasipolynome von e^{it}:

$$\mathcal{Q}(e^{it}) = \sum_{k=-m}^{n} c_k\, e^{ikt}\,.$$

Wir wollen gewisse Limiten von Quasipolynomen $\mathcal{Q}(e^{it})$ betrachten und schließlich Linearkombinationen solcher Limesfunktionen, und diese Operationen: Limesbildung und Bildung von Linearkombinationen, auf die entsprechenden Operatorfunktionen übertragen, um damit die oben erwähnten Funktionen von U zu gewinnen.

Beachten Sie: $\mathcal{Q}(e^{it})$ kann reell und sogar positiv sein (z. B. läßt sich $\cos^2 t$ als Quasipolynom schreiben). Wir brauchen im Folgenden eine Darstellung positiver Quasipolynome; darüber will ich zuerst sprechen.
Ich benutze folgende **Notation**: wenn $\mathcal{Q}(z) = \sum c_k z^k$, dann ist $\overline{\mathcal{Q}}(z) := \sum \overline{c}_k z^k$. Der Einheitskreis werde mit $\mathcal{C}_1 = \{z \in \mathbb{C};\ |z| = 1\}$ bezeichnet.
Für das Quasipolynom $\mathcal{Q}(z) = \sum_{k=-m}^{n} c_k\, z^k$ ist $\overline{\mathcal{Q}(z)} = \overline{\mathcal{Q}}(\overline{z}) = \sum_{k=-m}^{n} \overline{c}_k\, \overline{z}^k$; auf dem Einheitskreis ist $\overline{z} = z^{-1}$, also $\overline{\mathcal{Q}(z)} = \sum_{k=-m}^{n} \overline{c}_k\, z^{-k} = \sum_{l=-n}^{m} \overline{c}_{-l}\, z^l$.

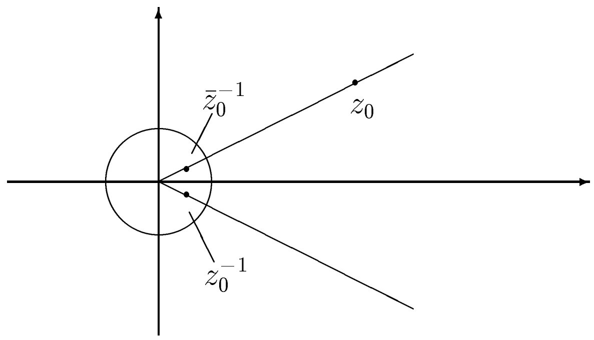
Man sieht: $\mathcal{Q}$ ist reell auf $\mathcal{C}_1$ genau dann, wenn $m = n$ und $\overline{c}_k = c_{-k}$ ist. In diesem Fall läßt sich $\mathcal{Q}$ für alle $z \in \mathbb{C}$ schreiben als

$$\mathcal{Q}(z) = c_0 + \sum_{k=1}^{n} (c_k\, z^k + \overline{c}_k\, z^{-k}) \,.$$

Daher ist

$$\mathcal{Q}(z^{-1}) = c_0 + \sum_{k=1}^{n} (c_k\, z^{-k} + \overline{c}_k\, z^{k}) = \overline{\mathcal{Q}}(z) \,.$$

Wir schließen daraus, daß die Nullstellen von $\mathcal{Q}$ „symmetrisch" zum Kreis $\mathcal{C}_1$ liegen: wenn z_0 Nullstelle ist, dann auch $1/\overline{z}_0$, denn aus $\overline{\mathcal{Q}(z_0)} = 0$ folgt $\overline{\mathcal{Q}}(1/z_0) = 0$, also $\overline{\mathcal{Q}(1/\overline{z}_0)} = 0$, und daher $\mathcal{Q}(1/\overline{z}_0) = 0$.
Multiplikation mit z^n macht aus dem Quasipolynom ein Polynom $2n$-ten Grades, das dieselben $2n$ „symmetrisch"



zu $\mathcal{C}_1$ liegenden Nullstellen $z_1, \dots, z_n, \overline{z}_1^{-1}, \dots, \overline{z}_n^{-1}$ besitzt. Es hat die Darstellung

$$z^n \mathcal{Q}(z) = c_n \prod_{k=1}^{n} (z - z_k)(z - \overline{z}_k^{-1}) = a_0 \prod_{k=1}^{n} (z - z_k)(1 - \overline{z}_k z), \quad a_0 = c_n \prod (-\overline{z}_k)^{-1} \,.$$

Definieren wir $q(z) := \prod_{k=1}^{n} (z - z_k)$, dann ist

$$\overline{q}(z^{-1}) = \prod (z^{-1} - \overline{z}_k) = z^{-n} \prod (1 - \overline{z}_k z) \,,$$

also

$$\mathcal{Q}(z) = a_0\, q(z)\, \overline{q}(z^{-1}) \,.$$

Sei nun $\mathcal{Q}(z) \geq 0$ auf $\mathcal{C}_1$. Aus $0 \leq \mathcal{Q}(1) = a_0\, q(1)\, \overline{q}(1) = a_0\, |q(1)|^2$ folgt $a_0 \geq 0$. Mit $\mathcal{P}(z) := \sqrt{a_0}\, q(z)$ bekommen wir die gesuchte Darstellung:

$$\mathcal{Q}(z) = \mathcal{P}(z)\, \overline{\mathcal{P}}(z^{-1}) \,,$$

und auf $\mathcal{C}_1$:

$$\mathcal{Q}(e^{it}) = \mathcal{P}(e^{it})\, \overline{\mathcal{P}}(e^{-it}) = \left| \mathcal{P}(e^{it}) \right|^2 \,. \tag{7.1}$$

Die Operatoren $\mathcal{Q}(U)$ sind beschränkt: $\|\mathcal{Q}(U)\| \leq \sum |c_k|$; alle anderen in diesem Abschnitt auftretenden Operatoren sind endliche Summen, Produkte und Limiten solcher $\mathcal{Q}$'s und daher ebenfalls beschränkt.

Zwischen den Quasipolynomen von e^{it} und von U bestehen enge Beziehungen:

Lemma 4.7.2 *(i) Aus* $\mathcal{Q}_3(e^{it}) = \mathcal{Q}_1(e^{it}) + \mathcal{Q}_2(e^{it})$ *und* $\mathcal{Q}_4(e^{it}) = \mathcal{Q}_1(e^{it})\,\mathcal{Q}_2(e^{it})$ *folgt*

$$\mathcal{Q}_3(U) = \mathcal{Q}_1(U) + \mathcal{Q}_2(U), \qquad \mathcal{Q}_4(U) = \mathcal{Q}_1(U)\,\mathcal{Q}_2(U)\,.$$

(ii)

$$\mathcal{Q}(U)^* = \sum_{k=-m}^{n} \overline{c}_k\, U^{-k} = \overline{\mathcal{Q}}(U^{-1})\,;$$

demnach ist $\mathcal{Q}(U)$ *genau dann selbstadjungiert, wenn* $m = n$ *und* $c_{-k} = \overline{c}_k$ *ist, d. h. wenn* $\mathcal{Q}(e^{it})$ *reell ist.*

(iii) Wenn

$$\mathcal{Q}(e^{it}) \geq 0, \qquad 0 \leq t < 2\pi\,, \tag{7.3}$$

gilt, ist $\mathcal{Q}(U)$ *nicht-negativ:* $(\varphi, \mathcal{Q}(U)\varphi) \geq 0\ \forall\, \varphi$.

Weil alle auftretenden Operatoren beschränkt sind, entfällt im folgenden die Unterscheidung zwischen symmetrisch und selbstadjungiert.

Beweis: Die ersten beiden Punkte sind trivial. Um die dritte Aussage zu beweisen, benutzen wir die Darstellung (7.1) des positiven Quasipolynoms:

$$\mathcal{Q}(e^{it}) = \overline{\mathcal{P}}(e^{-it})\,\mathcal{P}(e^{it}) = \left|\mathcal{P}(e^{it})\right|^2\,.$$

Deshalb gilt nach (i) und (ii) $\mathcal{Q}(U) = T^*T$ mit $T = \mathcal{P}(U)$, also $(\varphi, \mathcal{Q}(U)\varphi) = (T\varphi, T\varphi) \geq 0$. ■

4.7.4 Anmerkung: Wenn $\mathcal{Q}(e^{it}) \geq c \in \mathbb{R}$ ist, gilt $\mathcal{Q}(U) \geq c\mathbb{1}$; denn man betrachte $\tilde{\mathcal{Q}}$ mit $\tilde{c}_k = c_k$ für $k \neq 0$, $\tilde{c}_0 = c_0 - c$, dann gilt nach (iii) $\tilde{\mathcal{Q}}(e^{it}) \geq 0$ und $\tilde{\mathcal{Q}}(U) = \mathcal{Q}(U) - c\mathbb{1} \geq 0$.

Wir werden dieses Lemma in Kürze auf weitere Funktionen verallgemeinern, und auch das ist so einfach wie der obige Beweis. Doch es ist wichtig für unsere weitere Argumentation, insbesondere die dritte Eigenschaft: die Ordnungsrelation bleibt erhalten. Denn wenn $\mathcal{Q}_1(e^{it}) \geq \mathcal{Q}_2(e^{it})$, dann ist $\mathcal{Q} := \mathcal{Q}_1 - \mathcal{Q}_2$ ein auf $\mathcal{C}_1$ positives Quasipolynom, also gilt auch $\mathcal{Q}_1(U) \geq \mathcal{Q}_2(U)$.

Um allgemeinere Funktionen von U zu bilden, suche ich Funktionen von e^{it}, die sich in geeigneter Weise durch Quasipolynome approximieren lassen und hoffe, daß die entsprechenden Quasipolynome von U auch konvergieren.

4.7.5 Notation: Die Menge $\mathcal{K}$ bestehe aus allen auf dem Einheitskreis $|z| = 1$ stetigen reellen Funktionen $\varphi(e^{it})$ und allen stückweise stetigen, nach unten beschränkten Funktionen, die sich als Limiten monoton fallender, in jedem Punkt konvergenter Folgen stetiger reeller Funktionen $\varphi_n(e^{it})$ darstellen lassen.

Lemma 4.7.6 *Zu jeder Funktion $\psi(e^{it}) \in \mathcal{K}$ gibt es eine monoton fallende Folge reeller Quasipolynome $\mathcal{Q}_n(e^{it})$, die in jedem Punkt gegen $\psi(e^{it})$ konvergieren.*

Beweis: Nach dem (abgewandelten) Satz von Weierstraß (Satz 2.1.4) läßt sich jede reelle stetige 2π-periodische Funktion durch Polynome in $\sin t$ und $\cos t$ approximieren, anders: durch reelle Quasipolynome $\mathcal{Q}(e^{it})$. Approximieren wir also $\psi(e^{it})$ durch eine fallende Folge stetiger φ_n und diese wiederum durch Quasipolynome $\mathcal{Q}_n$. Wenn man sie so wählt, daß

$$\left|\mathcal{Q}_n(e^{it}) - \left(\varphi_n(e^{it}) + 3 \cdot 2^{-(n+2)}\right)\right| \leq 2^{-(n+2)}$$

gilt, sieht man aus den Folgerungen

$$\begin{aligned} \mathcal{Q}_n &\leq \varphi_n + 4 \cdot 2^{-(n+2)} = \varphi_n + 2^{-n}\,, \\ \mathcal{Q}_n &\geq \varphi_n + 2 \cdot 2^{-(n+2)} = \varphi_n + 2^{-(n+1)}\,, \end{aligned}$$

daß die $\mathcal{Q}_n$ auf dem Einheitskreis punktweise gegen ψ konvergieren und daß

$$\mathcal{Q}_{n+1}(e^{it}) \leq \varphi_{n+1}(e^{it}) + 2^{-(n+1)} \leq \varphi_n(e^{it}) + 2^{-(n+1)} \leq \mathcal{Q}_n(e^{it})\,.$$

■

Bilden wir jetzt $\mathcal{Q}_n(U)$. Das sind selbstadjungierte Operatoren, weil die $\mathcal{Q}_n$ auf dem Einheitskreis reell sind. Weil die $\mathcal{Q}_n(e^{it})$ eine monoton fallende Folge bilden, die durch $\inf\{\psi(e^{it});\ t \in [0, 2\pi)\}$ nach unten beschränkt ist, gilt dasselbe wegen Anmerkung 4.7.4 für die $\mathcal{Q}_n(U)$; nach Satz 4.4.4 haben sie einen starken Limes S, den ich gleich $\psi(U)$ setzen möchte. Dazu muß ich aber noch überprüfen, daß S nicht von der Wahl der approximierenden Folge abhängt. Sei $\mathcal{Q}'_n$ eine zweite derartige Folge. Dann findet man zu jedem m und jedem t ein n_0, so daß

$$\mathcal{Q}'_n(e^{it}) < \mathcal{Q}_m(e^{it}) + \frac{1}{m} \quad \text{für } n > n_0\,. \tag{7.7}$$

Wegen der Stetigkeit der Quasipolynome gilt das auch in einer kleinen Umgebung von t. Weil der Einheitskreis kompakt ist, genügen die Umgebungen von endlich vielen t's, um ihn zu überdecken; wählen wir das größte zugehörige n_0, so gilt Gleichung (7.7) für alle t. Das gleiche gilt, wenn wir die Rollen von $\mathcal{Q}_n$ und $\mathcal{Q}'_n$ vertauschen. Mit Lemma 4.7.2 auf die Operatorfunktionen übertragen:

$$\mathcal{Q}'_n(U) < \mathcal{Q}_m(U) + \frac{1}{m}\mathbb{1}, \quad \mathcal{Q}_n(U) < \mathcal{Q}'_m(U) + \frac{1}{m}\mathbb{1}$$

für alle $n > n_0$, und deshalb

$$T \equiv \text{s-lim}\, \mathcal{Q}'_n(U) \leq \mathcal{Q}_m(U) + \frac{1}{m}\mathbb{1}, \quad S \equiv \text{s-lim}\, \mathcal{Q}_n(U) \leq \mathcal{Q}'_m(U) + \frac{1}{m}\mathbb{1} \qquad \forall m\,;$$

folglich $S \leq T$ und $T \leq S$, also $T = S$.
Ich betone noch einmal: die Operator-Limites sind gemeint im Sinne der starken Operator-Konvergenz; denn das ist die in Satz 4.4.4 relevante Topologie.

Damit haben wir $\psi(U)$ definiert für alle $\psi \in \mathcal{K}$. Man beachte: $\psi(U)$ ist selbstadjungiert, weil ψ und die approximierenden $\mathcal{Q}_n$ auf dem Einheitskreis reell, die $\mathcal{Q}_n(U)$ also selbstadjungiert sind. Man überprüft leicht das Analogon zu Lemma 4.7.2:

Lemma 4.7.8 *Für $\psi_k \in \mathcal{K}$ folgen aus $\psi_3(e^{it}) = \psi_1(e^{it})+\psi_2(e^{it})$, $\psi_4(e^{it}) = \psi_1(e^{it})\,\psi_2(e^{it})$ und $\psi_1(e^{it}) \geq \psi_2(e^{it})$ die entsprechenden Operatorrelationen*

$$\psi_3(U) = \psi_1(U) + \psi_2(U), \qquad \psi_4(U) = \psi_1(U)\,\psi_2(U), \qquad \psi_1(U) \geq \psi_2(U)\,.$$

Zum Beweis dieses Lemmas muß man berücksichtigen, daß Produkte von Quasipolynomen wieder Quasipolynome sind, daß Produkte stetiger reeller Funktionen wieder stetig und reell sind; weiterhin: wenn $\psi_1, \psi_2 \in \mathcal{K}$, $\psi_i = \lim \mathcal{Q}_n^{(i)}$, $Q_n^{(i)}$ fallend, dann ist offensichtlich auch $\lim Q_n^{(1)} Q_n^{(2)} = \psi_1\psi_2 \in \mathcal{K}$, und die ersten beiden Relationen folgen aus Lemma 4.7.2(i) und (ii). Die Ordnungsrelation folgt, weil man für alle $n \in \mathbb{N}$ ein ε_n findet, so daß $|\varepsilon_n| < 2/n$ und $\psi_1 - \psi_2 = \mathcal{Q}_n^{(1)} - \mathcal{Q}_n^{(2)} + \varepsilon_n \geq 0$, also $\mathcal{Q}_n \equiv \mathcal{Q}_n^{(1)} - \mathcal{Q}_n^{(2)} \geq -2/n$; jetzt kann man Lemma 4.7.2(iii) anwenden, und mit $\lim \mathcal{Q}_n = \psi_1 - \psi_2$ übertragen sich diese Relationen auch auf die Operatorfunktionen.

Jetzt kommt der letzte Schritt: wir bilden die lineare Hülle von $\mathcal{K}$ und ordnen der Funktion

$$\omega(e^{it}) = \sum_{k=1}^{n} c_k\, \psi_k(e^{it})$$

die Operatorfunktion

$$\omega(U) = \sum_{k=1}^{n} c_k\, \psi_k(U)$$

mit beliebigen komplexen Koeffizienten zu. Wieder müssen wir überprüfen, daß aus $\omega_1(e^{it}) = \omega_2(e^{it})$ auch $\omega_1(U) = \omega_2(U)$ folgt. Ich kann die c_k (und damit die $\omega(e^{it})$) in Real- und Imaginärteil zerlegen; dem entspricht die Zerlegung von $A = \omega(U)$ in $A = A_1 + iA_2$ mit den selbstadjungierten Operatoren $A_1 = \frac{1}{2}(A + A^*)$ und $A_2 = \frac{1}{2i}(A - A^*)$, man beachte $\psi(U)^* = \psi(U)$. Es genügt daher, reelle c_k zu betrachten. Summen mit *positiven* Koeffizienten führen nicht aus $\mathcal{K}$ heraus, sie sind durch fallende Folgen stetiger Funktionen approximierbar, das gilt aber nicht für solche mit negativen c_k's. Wenn wir die Terme mit positiven und negativen Koeffizienten jeweils zusammenfassen, können wir ω_i schreiben als

$$\omega_1(e^{it}) = \psi_1(e^{it}) - \psi_2(e^{it}) = \omega_2(e^{it}) = \psi_3(e^{it}) - \psi_4(e^{it})$$

mit $\psi_k \in \mathcal{K}$; oder gleichwertig damit:

$$\psi_1(e^{it}) + \psi_4(e^{it}) = \psi_3(e^{it}) + \psi_2(e^{it})\,.$$

Hier sind beide Seiten $\in \mathcal{K}$, folglich gilt die Gleichung auch, wenn wir e^{it} durch U ersetzen; und damit haben wir die Eindeutigkeit.
Wir haben die Zuordnung der Funktionen zu Operatorfunktionen also auf L.H.($\mathcal{K}$) fortgesetzt. Wieder überprüft man ohne Schwierigkeiten das folgende

Lemma 4.7.9 *Die Aussagen von Lemma 4.7.8 gelten auch für Funktionen aus* L.H.($\mathcal{K}$)*: die Relationen*

$$\psi_3(U) = \psi_1(U) + \psi_2(U), \qquad \psi_4(U) = \psi_1(U)\,\psi_2(U), \qquad \psi_1(U) \geq \psi_2(U)$$

gelten, wenn sie für diese Funktionen mit dem Argument e^{it} anstelle von U richtig sind.

4.7.10 Anmerkungen: (i) Alle Quasipolynome, auch die nicht reellen, liegen in L.H.($\mathcal{K}$), denn $e^{it} \in$ L.H.($\mathcal{K}$): man schreibe $z = \frac{1}{2}(z + \overline{z}) + \frac{1}{2i}i(z - \overline{z})$ als komplexe Linearkombination reeller Funktionen.

(ii) Summen- und Produktbildung führt nicht aus L.H.($\mathcal{K}$) heraus, vergl. die Bemerkungen zum Beweis von Lemma 4.7.8. L.H.($\mathcal{K}$) ist also eine Algebra von Funktionen.

(iii) Man beachte, daß der Funktion $\overline{\omega(e^{it})}$ der Operator $\omega(U)^*$ zugeordnet ist.

4.7.11 Anmerkung: Der Operator U vertauscht mit allen Operatoren $\omega(U)$, wenn $\omega \in$ L.H.($\mathcal{K}$) ist; das folgt sofort, wenn man die Relation $z\omega(z) = \omega(z)z$ mit $z = e^{it}$ in die entsprechende Operatorgleichung übersetzt.

4.8 Spektraldarstellung selbstadjungierter Operatoren

Die Überlegungen des letzten Abschnitts dienten der Vorbereitung für die Spektralzerlegung *unitärer* Operatoren. Das zu erreichen, ist jetzt das erste Ziel; daraus folgt dann die Spektraldarstellung für selbstadjungierte Operatoren relativ einfach. Ich beschreibe die dem Folgenden zugrunde Idee anhand von Funktionen auf dem Einheitskreis: χ_I^c sei eine charakteristische Funktion auf dem Einheitskreis; wir schreiben für $I \subset [0, 2\pi]$

$$\chi_I^c(e^{it}) := \begin{cases} 1 & t \in I \pmod{2\pi} \\ 0 & t \notin I \pmod{2\pi}\,. \end{cases} \tag{8.1}$$

Weiter definieren wir

$$\omega_0(e^{it}) := 0\,, \quad \omega_\lambda(e^{it}) := \chi_{[0,\lambda)}^c(e^{it}) \ \text{ für } 0 < \lambda \leq 2\pi\,. \tag{8.2}$$

Wenn wir nun $\omega_\lambda(e^{it_0}) \equiv \widetilde{\omega}(\lambda)$ mit festem t_0 als Funktion von λ betrachten, haben wir eine Stufenfunktion mit Sprung 1 bei $\lambda = t_0$, $\widetilde{\omega}(\lambda) = 0$ für kleine λ, und

$$\int_0^{2\pi} e^{i\lambda}\, d\widetilde{\omega}(\lambda) = e^{it_0}\,.$$

Man beachte, daß $\omega_t(e^{it}) = 0$ ist für $t < 2\pi$ (t liegt nicht im halboffenen Intervall $[0, t)$), deshalb ist ω_λ als Funktion von λ bei festem t *links*seitig stetig bei $\lambda = t$: $\omega_\lambda(e^{it}) = 0$ für $\lambda \leq t$; ω_λ hat also Eigenschaften, wie wir sie für Zerlegungen der Einheit gefordert haben: $\omega_0 = 0$, ω_λ wachsend mit λ, $\omega_{2\pi} = 1$ (denn $\omega_{2\pi}(e^{it}) = 1$, weil alle $t \pmod{2\pi}$ in $[0, 2\pi)$ liegen), und $\lambda \mapsto \omega_\lambda$ ist linksseitig stetig.

Charakteristische Funktionen haben Werte 0 und 1, es gilt also immer $\chi_I^{c\,2} = \chi_I^c$; *wenn* wir nun durch die Ersetzung $e^{it} \longrightarrow U$ alles auf Operatoren übertragen können, wird aus ω_λ ein Projektor und das Integral liefert uns — hoffentlich — die gesuchte Darstellung. Das müssen wir jetzt genauer untersuchen.

Erste Beobachtung: Wenn I abgeschlossen ist, nimmt χ_I^c, betrachtet als Funktion von t, an den Endpunkten des Intervalls den Wert 1 an, d. h. den höheren der beiden möglichen Funktionswerte; $\chi_I^c(e^{it})$ ist *oberhalb stetig*[12] und als Limes einer *fallenden* Folge stetiger Funktionen darstellbar, also $\in \mathcal{K}$. Das gilt auch, wenn I einpunktig ist. (Hier müssen wir allerdings ein bißchen aufpassen: wir müssen die Punkte $\lambda = 0$ und $\lambda = 2\pi$ identifizieren, also $\chi_{\{0\}}^c(e^{it}) = \chi_{\{2\pi\}}^c(e^{it}) = 1$ setzen, wenn $t \equiv 0 \bmod = 2\pi$ ist, und sonst $= 0$.) Da aber

$$\chi_{[0,\lambda)}^c = \chi_{[0,\lambda]}^c - \chi_{\{\lambda\}}^c$$

ist, gilt jedenfalls $\omega_\lambda = \chi_{[0,\lambda)}^c \in \text{L.H.}(\mathcal{K})$; und die Operatoren

$$E_\lambda := \omega_\lambda(U)$$

[12] Siehe Anhang zu diesem Abschnitt

sind wohldefiniert.

Zweite Beobachtung: Weil ω_λ reell ist und $\omega_\lambda^2 = \omega_\lambda$, sind die E_λ nach Lemma 4.7.9 orthogonale Projektoren. Aus $\omega_0(e^{it}) = 0$, $\omega_{2\pi}(e^{it}) = 1$ und der Tatsache, daß $\omega_\lambda(e^{it})$ für festes t eine mit λ monoton nicht fallende Funktion ist, folgt aus dem genannten Lemma

$$E_0 = 0\,, \quad E_{2\pi} = \mathbb{1}\,, \quad E_\kappa \leq E_\lambda\,, \quad \text{wenn } \kappa < \lambda\,. \tag{8.3}$$

Man beachte, daß für Projektoren P und Q die Relationen $P \leq Q$ und $PQ = P$ äquivalent sind, beide besagen, daß $P\mathcal{H} \subset Q\mathcal{H}$ (Beweis: Ü).
Wir haben also eine Zerlegung der Einheit konstruiert, wenn wir zeigen können, daß $\lambda \mapsto E_\lambda$ linkseitig stark stetig ist, d. h. für $\varepsilon \searrow 0$ muß gelten

$$E_\lambda - E_{\lambda-\varepsilon} = \omega_\lambda(U) - \omega_{\lambda-\varepsilon}(U) = \chi^c_{[\lambda-\varepsilon,\lambda)}(U) = \chi^c_{[\lambda-\varepsilon,\lambda]}(U) - \chi^c_{\{\lambda\}}(U) \longrightarrow 0. \tag{8.4}$$

Da $\chi^c_{\{\lambda\}} \in \mathcal{K}$, gibt es eine fallende Folge von Quasipolynomen $\mathcal{Q}_n$, die $\chi^c_{\{\lambda\}}$ approximieren; die $\mathcal{Q}_n$ sind stetig, deshalb ist dann für hinreichend kleines $\varepsilon < \varepsilon_0(n)$

$$\chi^c_{[\lambda-\varepsilon,\lambda]}(e^{it}) \leq \mathcal{Q}_n(e^{it}) + \frac{1}{n}\,,$$

(denn $\mathcal{Q}_n(e^{i\lambda}) \geq 1$, folglich $\mathcal{Q}_n(e^{it}) + 1/n > 1$ in einer kleinen Umgebung von $t = \lambda$,) und entsprechend gilt

$$\chi^c_{\{\lambda\}}(U) \leq \chi^c_{[\lambda-\varepsilon,\lambda]}(U) \leq \mathcal{Q}_n(U) + \frac{1}{n}\mathbb{1}\,.$$

Die rechte Seite geht für $n \to \infty$ gegen $\chi^c_{\{\lambda\}}(U)$ (im Sinne der starken Konvergenz), und wenn wir gleichzeitig ε gegen 0 gehen lassen, so daß immer $\varepsilon < \varepsilon_0(n)$ gilt, bleibt die Ungleichung bei Limesbildung richtig; und damit folgt die Relation (8.4).

Um die Spektraldarstellung zu gewinnen, wollen wir nun zeigen, daß geeignete Approximationssummen, gebildet analog einem gewöhnlichen Riemann-Stieljes-Integral, den Operator U approximieren.
Wir betrachten Zerlegungen

$$0 = \lambda_0 < \lambda_1 < \lambda_2 < \ldots < \lambda_n = 2\pi\,.$$

Dann ist

$$\left| e^{it} - \sum_{k=1}^{n} e^{i\lambda_{k-1}} \left(\omega_{\lambda_k}(e^{it}) - \omega_{\lambda_{k-1}}(e^{it}) \right) \right| \leq \max_r |\lambda_r - \lambda_{r-1}|\,; \tag{8.5}$$

denn wenn t im Intervall $[\lambda_{r-1}, \lambda_r)$ liegt, ist nur der Summand mit $k = r$ in der Summe $\neq 0$, die linke Seite ist gleich

$$\left| e^{it} - e^{i\lambda_{r-1}} \right| = 2 \left| \sin \frac{t - \lambda_{r-1}}{2} \right| \leq |\lambda_r - \lambda_{r-1}|\,,$$

und für $t = 2\pi$ ist die linke Seite von (8.5) $= 0$. (Denn nur der Term mit $k = 1$ verschwindet nicht, weil $\omega_\lambda(e^{2\pi i}) = 1 \; \forall \, \lambda \neq 0$, und $\omega_0(e^{it}) = 0$.)
Wenn die Zerlegung hinreichend fein ist, gilt also

$$\overline{\left(e^{it} - \sum_{k=1}^{n} e^{i\lambda_{k-1}} \left(\omega_{\lambda_k}(e^{it}) - \omega_{\lambda_{k-1}}(e^{it})\right)\right)} \left(e^{it} - \sum_{k=1}^{n} e^{i\lambda_{k-1}} \left(\omega_{\lambda_k}(e^{it}) - \omega_{\lambda_{k-1}}(e^{it})\right)\right) \leq \varepsilon^2 .$$

Man beachte, daß beide Seiten Funktionen $\in$ L.H.($\mathcal{K}$) sind — siehe die Anmerkung 4.7.10(ii): L.H.($\mathcal{K}$) ist eine Algebra —; wir können also die Ungleichung gemäß Lemma 4.7.9 auf die entsprechenden Operatoren übertragen:

$$0 \leq \left(U - \sum_{k=1}^{n} e^{i\lambda_{k-1}} \left(E_{\lambda_k} - E_{\lambda_{k-1}}\right)\right)^* \left(U - \sum_{k=1}^{n} e^{i\lambda_{k-1}} \left(E_{\lambda_k} - E_{\lambda_{k-1}}\right)\right) \leq \varepsilon^2 \mathbb{1} ,$$

oder anders:

$$\left\| U - \sum_{k=1}^{n} e^{i\lambda_{k-1}} \left(E_{\lambda_k} - E_{\lambda_{k-1}}\right) \right\| \leq \varepsilon .$$

Die Approximationssummen haben also für immer feinere Zerlegung einen Limes, der gleich U ist und den wir als Integral schreiben:

$$U = \int_0^{2\pi} e^{i\lambda} \, dE_\lambda . \tag{8.6}$$

Das ist die gesuchte Spektraldarstellung. In diesem Fall haben wir sogar Norm-Konvergenz des Integrals; doch das können wir in allgemeineren Fällen nicht erwarten, im allgemeinen sind die Integrale im Sinne der starken Konvergenz zu verstehen.

In ganz analoger Weise kann man vorgehen, wenn man von einem selbstadjungierten beschränkten Operator statt von einem unitären ausgeht: man setze $[-\|A\|, \|A\|]$ statt $[0, 2\pi]$ und betrachte Polynome und stetige reelle Funktionen in λ statt der Quasipolynome in e^{it} usw., das Ergebnis ist dann die Darstellung $A = \int \lambda \, dE_\lambda$. Doch werden wir später die Spektraldarstellung für beliebige, auch unbeschränkte selbstadjungierte Operatoren ausgehend von der Darstellung (8.6) ableiten.

Um unser Vorgehen noch einmal am denkbar einfachsten Beispiel zu erläutern, betrachte ich einen einen eindimensionalen Hilbertraum, in dem alle Operatoren durch Multiplikation mit einer komplexen Zahl gegeben sind: ein selbstadjungierter Operator A entspricht einer reellen Zahl r, die ich in der Form $r = \int \lambda \, \delta(\lambda - r) \, d\lambda = \int \lambda \, d\Theta(\lambda - r)$ schreiben kann. Die zugehörige Spektralschar $E_\lambda := \Theta(\lambda - r)$ hängt natürlich von dem „Operator“ r ab:

$$r \mapsto \Theta(\lambda - r) = \chi_{(-\infty,\lambda)}(r) .$$

(Hier genügte natürlich auch $\chi_{[-\|A\|,\lambda)}$.) Die charakteristische Funktion χ, angewandt auf den Operator $A = r$, pickt also im Integral den „richtigen“ Wert $\lambda = r$ heraus.
Wenn $\dim \mathcal{H} = \infty$, hat ein selbstadjungierter Operator A im allgemeinen nicht nur einen, sondern viele „Werte“: Eigenwerte und Beinahe-Eigenwerte, die durch $\chi_{(-\infty,\lambda)}(A)$

gewissermaßen automatisch herausgepickt werden.
Auf den Zusammenhang von Spektraldarstellung und Spektrum komme ich noch zurück, nur soviel hier schon: genau dann, wenn $\lambda_0 \in \operatorname{Sp} A$, ist E_λ bei λ_0 nicht konstant.

Wenn man in der Rechnung ab (8.5), die zu Gleichung (8.6) geführt hat, überall t bzw. $\lambda, \lambda_k, \ldots$ durch nt bzw. $n\lambda, n\lambda_k, \ldots$ ersetzt und U^n statt U setzt, bleibt alles richtig; es ist also

$$U^n = \int e^{in\lambda}\, dE_\lambda\,.$$

Man kann entsprechende Gleichungen für Funktionen von U aufstellen, die durch Polynome oder Quasipolynome approximierbar sind; wir können statt dessen umgekehrt vorgehen und mit der Spektraldarstellung Funktionen von U definieren durch

$$f(U) := \int f(e^{i\lambda})\, dE_\lambda\,. \tag{8.7}$$

Das Integral ist sinnvoll für alle beschränkten, stückweise stetigen Funktionen f, die an den Unstetigkeitsstellen von E_λ stetig sind. Für Quasipolynome und Funktionen aus L.H.($\mathcal{K}$) stimmt diese Definition mit der früheren überein.[13] Ich kann natürlich noch einfacher

$$A_g \equiv \int g(\lambda)\, dE_\lambda \tag{8.8}$$

schreiben; und ich kann diese Definition auch ausweiten auf unbeschränkte Funktionen g. Dann bekomme ich einen unbeschränkten Operator, dessen Definitionsbereich ich angeben muß:

$$\mathcal{D}_{A_g} = \left\{\varphi;\ \int |g(\lambda)|^2\, d(\varphi, E_\lambda\varphi) < \infty\right\}\,. \tag{8.9}$$

Das Integral in (8.9) ist $= \|A_g\varphi\|^2$, vergl. Gleichung (6.23); die Überlegungen aus Abschnitt 4.6 bei den Gleichungen (6.24) und (6.25) lassen sich sofort auf stückweise stetige Funktionen übertragen, die möglicherweise unbeschränkt sind. (Die Formeln gelten natürlich auch, wenn $\operatorname{supp} E_\lambda = \mathbb{R}$ ist; man kann ja auch das Intervall $[0, 2\pi]$ eineindeutig und stetig auf $\mathbb{R}$ abbilden und die transformierte Funktion und die transformierte Spektralschar betrachten, wie wir es in Kürze tun werden.)
Ich wiederhole noch einmal das in Abschnitt 4.6 Gesagte: weil diese Integrale $\int g(\lambda)\, dE_\lambda$ auf geeigneten Bereichen stark konvergieren, darf man Skalarproduktbildung und Integration vertauschen,

$$(\psi, A_g\varphi) = \int g(\lambda)\, d(\psi, E_\lambda\varphi)\,;$$

und man darf nach der formalen Regel

$$dE_\kappa\, dE_\lambda = \delta(\kappa - \lambda) d\kappa\, dE_\lambda$$

rechnen, z. B. wenn $A = \int \lambda\, dE_\lambda$ ist:

$$\|A\varphi\|^2 = (\varphi, \iint \kappa\, dE_\kappa\, \lambda\, dE_\lambda\varphi) = (\varphi, \int \lambda^2\, dE_\lambda\varphi) = \int \lambda^2\, d(\varphi, E_\lambda\varphi)\,; \tag{8.10}$$

[13] Man beachte, daß die Ordnungsrelationen für die $\mathcal{Q}(e^{it})$ sich auf die $\mathcal{Q}(U)$ übertragen!

oder

$$\int f(\lambda)\, dE_\lambda \int g(\lambda)\, dE_\lambda = \int f(\lambda) g(\lambda)\, dE_\lambda\,. \tag{8.11}$$

Die zweite Gleichung gilt wenigstens im Sinne der schwachen Konvergenz, wenn man den Operator auf $\varphi \in D_{A_g}$ anwendet und von links skalar mit $\psi \in \mathcal{D}_{A_f}$ multipliziert (für besonders brave φ kann sie natürlich auch im Sinne der starken Konvergenz gelten). Ich habe vorsichtshalber immer nur von starken Limiten gesprochen und $A = \int \lambda\, dE_\lambda$ als unbeschränkten Operator behandelt. Es kann natürlich sein — und die obigen Rechnungen bleiben trivialerweise richtig —, daß $E_\lambda = 0$ für $\lambda \le a$ und $E_\lambda = \mathbb{1}$ für $\lambda \ge b$ mit endlichen Werten a und b. Dann ist A tatsächlich beschränkt und $\|A\| \le \max(|a|, |b|)$, wie man leicht aus Gleichung (8.10) ersieht (Ü).

Die folgenden Aussagen sind nicht schwer zu beweisen:

Satz 4.8.12 *Der Operator $\int f(\lambda)\, dE_\lambda$ ist beschränkt durch M, wenn $|f(\lambda)| \le M$; er ist unitär, wenn $|f(\lambda)| = 1$; und er ist symmetrisch, wenn $f(\lambda)$ reell ist.*

Für die erste Aussage überlege man sich, daß

$$\|A\|^2 = \sup_{\varphi \neq 0} \frac{\|A\varphi\|^2}{\|\varphi\|^2} = \sup_{\varphi \neq 0} \frac{(\varphi, A^* A \varphi)}{\|\varphi\|^2}$$

ist und verwende Gleichung (8.11); auch die zweite Aussage ergibt sich aus dieser Gleichung. Die dritte Aussage ist leicht zu sehen, sie gilt in verschärfter Form:

Satz 4.8.13 *Sei $f(x)$ eine reelle stetige Funktion dann ist $A_f := \int f(\lambda)\, dE_\lambda$ ein selbstadjungierter Operator mit Definitionsbereich $\mathcal{D}_{A_f} = \{\varphi;\ \int |f(\lambda)|^2\, d(\varphi, E_\lambda \varphi) < \infty\}$.*

Man kann die Stetigkeitsforderung an f aufweichen, wenn E_λ stetig ist; aber wenigstens dort, wo E_λ Sprünge hat, muß f stetig sein, im übrigen Bereich kann es selbst unstetig sein, solange es Riemann-integrabel ist.

Beweis von Satz 4.8.13: Der Beweis ist ganz analog zum früheren Beweis von Satz 4.5.11, wo wir $A = \sum A_n P_n$ betrachtet hatten.
(i) A_f ist hermitisch, das ist trivial, denn $(\varphi, A_f \varphi) = \int f(\lambda)\, d(\varphi, E_\lambda \varphi)$ ist reell.
(ii) A_f ist symmetrisch, denn $\mathcal{D}_{A_f}$ ist dicht: Sei Δ ein endliches Intervall und $E(\Delta) = \int_\Delta dE_\lambda$; dann ist für $\varphi \in E(\Delta)\mathcal{H}$

$$\int |f(\lambda)|^2\, d(\varphi, E_\lambda \varphi) = \int_\Delta |f(\lambda)|^2\, d(\varphi, E_\lambda \varphi) < \infty\,,$$

d. h. $\varphi \in \mathcal{D}_{A_f}$. Wegen $\int dE_\lambda = \mathbb{1}$ ist aber klar, daß $\{E(\Delta)\varphi;\ \varphi \in \mathcal{H},\ \Delta \subset \mathbb{R} \text{ endlich}\}$ eine dichte Menge ist.
(iii) Nun müssen wir $\mathcal{D}_{A_f^*}$ berechnen. Für $\varphi \in \mathcal{D}_{A_f}$ und $\psi \in \mathcal{D}_{A_f^*}$ ist

$$(\psi, A_f \varphi) = \int f(\lambda)\, d(\psi, E_\lambda \varphi) = \int f(\lambda)\, d(E_\lambda \psi, \varphi) = ({A_f}^* \psi, \varphi)$$

stetig in φ, also fortsetzbar auf ganz $\mathcal{H}$, insbesondere auf $\psi_\Delta := \int_\Delta \overline{f(\lambda')}\, dE_{\lambda'}\psi$ für alle Δ. Nun ist

$$((A_f{}^*\psi, \psi_\Delta) = \int f(\lambda)\, d(E_\lambda\psi, \psi_\Delta) = \int_\Delta |f(\lambda)|^2\, d(\psi, E_\lambda\psi) = \|\psi_\Delta\|^2 ,$$

die linke Seite ist beschränkt durch $\|A_f{}^*\psi\|\ \|\psi_\Delta\|$, folglich ist

$$\|\psi_\Delta\| = \left(\int_\Delta |f(\lambda)|^2\, d(\psi, E_\lambda\psi)\right)^{\frac{1}{2}} \leq \|A_f{}^*\psi\|$$

unabhängig von Δ, der Limes $\Delta \to \mathbb{R}$ ist also möglich und somit $\psi \in \mathcal{D}_{A_f}$. ■

Aus unseren Überlegungen folgt sofort:

- A_f vertauscht mit den Projektoren E_λ, siehe Gleichung (8.11).
- A_f vertauscht (auf einem gemeinsamen dichten Definitionsbereich) mit A_g (vergleiche hierzu und zum ersten Punkt Anmerkung 4.7.11).
- Jedes unbeschränkte $A = \int \lambda\, dE_\lambda$ kann man im Sinne der starken Konvergenz durch beschränkte Operatoren $A_R := \int_{-R}^{R} \lambda\, dE_\lambda$ approximieren (Ü).

Besonders der letzte Punkt ist für die Physiker wichtig. Kein Experimentator kann unbeschränkte Werte messen; je nachdem, wie groß sein Meßbereich ist, müßte der Theoretiker beschränkte Operatoren mit hinreichend großer Norm betrachten. Doch die Idealisierung des Theoretikers auf unbeschränkte Observable ist gerechtfertigt, weil sie ja immer durch beschränkte approximiert werden können; die *Menge* $\{A_R;\ R \in \mathbb{R}\}$ beschränkter Operatoren enthält dieselbe Information, wie der unbeschränkte Operator A. Wenn Sie z. B. an den Ortsoperator Q denken, sehen Sie, daß es oft viel einfacher ist, mit der Idealisierung der unbeschränkten Observablen zu arbeiten.
In der Statistischen Mechanik und in der Quantenfeldtheorie geht man manchmal auch den anderen Weg: weil die beschränkten Operatoren alle relevante Information enthalten können, betrachtet man Operator-Algebren beschränkter Operatoren („C^*-Algebren“, „von Neumann-Algebren“), die mathematisch viel einfacher zu handhaben sind und für die es zahlreiche nützliche mathematische Aussagen gibt.

Nun fehlt uns noch der Beweis, daß man jeden selbstadjungierten Operator mithilfe einer Spektralschar darstellen kann.

Theorem 4.8.14 *Zu jedem selbstadjungierten Operator A gibt es eine Zerlegung der Einheit, seine „Spektralschar“$\{E_\lambda\}$, so daß*

$$A = \int \lambda\, dE_\lambda$$

ist. Der Definitionsbereich von A ist gegeben durch

$$\mathcal{D}_A = \left\{ \varphi \in \mathcal{H};\ \int \lambda^2\, d(\varphi, E_\lambda \varphi) < \infty \right\} .$$

Die Beweisidee ist, diese Darstellung auf die Darstellung (8.6) eines unitären Operators zurückzuführen. Im Komplexen kann man die reelle Achse eineindeutig auf den Einheitskreis abbilden, z. B. mit $x \mapsto g(x) := \exp(-2i \operatorname{arc\,cot} x)$. $\int \lambda\, dE_\lambda$ ist selbstadjungiert, $\int g(\lambda)\, dE_\lambda$ ein zughöriger unitärer Operator. Diese Abbildung ist nicht die einzige, eine andere, noch einfachere Möglichkeit ist $x \mapsto \frac{x-i}{x+i}$. Man versucht dementsprechend, eine eineindeutige Abbildung zwischen selbstadjungierten und unitären Operatoren zu finden: die **Cayley-Transformation**.
Aus Satz 4.6.3 wissen wir, daß $\pm i \notin \operatorname{Sp} A$ ist, wenn $A = A^*$, daher ist $R(\mp i; A) = (A \pm i\mathbb{1})^{-1}$ ein beschränkter Operator, und der Wertebereich von $A \pm i\mathbb{1}$ ist ganz $\mathcal{H}$ (siehe Satz 4.6.3 (iv)). An dieser Stelle brauchen wir die Selbstadjungiertheit von A; es genügt nicht, wenn A nur symmetrisch ist. Ein beliebiges $\varphi \in \mathcal{H}$ können wir also schreiben als

$$\varphi = (A + i\mathbb{1})\psi\,, \quad \psi \in \mathcal{D}_A \text{ passend;} \tag{8.15}$$

dann ist $\psi = R(-i; A)\varphi$, und wir können die folgende Abbildung definieren:

$$U :\ \varphi \mapsto (A - i\mathbb{1})\psi = (A - i\mathbb{1})R(-i; A)\varphi \equiv U\varphi\,. \tag{8.16}$$

Aus $\|(A \pm i\mathbb{1})\psi\|^2 = \|A\psi\|^2 + \|\psi\|^2$ (vergl. die Rechnung nach Gleichung (6.7)) sieht man, daß U isometrisch ist:

$$\|U\varphi\| = \|(A - i\mathbb{1})\psi\| = \|(A + i\mathbb{1})\psi\| = \|\varphi\|\ .$$

Weil $\mathcal{W}_U = \mathcal{W}_{A-i\mathbb{1}} = \mathcal{H}$ ist, existiert das Inverse von U, d. h. U ist unitär.
Dieses U heißt die **Cayley-Transformierte** von A:

$$U = (A - i\mathbb{1})(A + i\mathbb{1})^{-1}\,. \tag{8.17}$$

Die Abbildung $A \mapsto U$ ist eindeutig umkehrbar. Für jedes $\psi \in \mathcal{D}_A$ können wir die Gleichungen (8.15) und (8.16):

$$\begin{aligned} (A + i\mathbb{1})\psi &= \varphi\,, \\ (A - i\mathbb{1})\psi &= U\varphi \end{aligned}$$

nach $A\psi$ und $i\psi$ auflösen:

$$A\psi = \frac{1}{2}(U + \mathbb{1})\varphi\,, \quad i\psi = \frac{1}{2}(U - \mathbb{1})\varphi\,. \tag{8.18}$$

U kann nicht den Eigenwert 1 haben; denn $U\varphi = \varphi$ kann nur für $\varphi = 0$ gelten:

$$\varphi = U\varphi = (A - i\mathbb{1})R(-i; A)\varphi = \Big(-2i\mathbb{1} + (A + i\mathbb{1}) \Big) R(-i; A)\varphi = -2iR(-i; A)\varphi + \varphi\,,$$

und das hat zur Folge, daß $R(-i;A)\varphi = 0$ und daher $U\varphi = (A - i\mathbb{1})R(-i;A)\varphi = 0$ ist. Wir können also $U - \mathbb{1}$ auf einem geeigneten Bereich invertieren; die zweite Gleichung in (8.18) zeigt, daß $(U - \mathbb{1})^{-1}$ auf $\mathcal{D}_A$ existiert und daß

$$A = i(\mathbb{1} + U)(\mathbb{1} - U)^{-1}\,. \tag{8.19}$$

Das ist die gesuchte Umkehrformel.

Jetzt ist unser Theorem leicht zu beweisen.

Beweis von Theorem 4.8.14: Zu $A = A^*$ suchen wir die Cayley-Transformierte U, die gemäß Gleichung (8.6) eine Spektraldarstellung $U = \int_0^{2\pi} e^{i\lambda}\,dE_\lambda$ hat. Damit können wir die rechte Seite von (8.19) (mithilfe der Gleichung (8.11) für Produkte solcher Integrale) schreiben als

$$A = \int_0^{2\pi} i\frac{1 + e^{i\lambda}}{1 - e^{i\lambda}}\,dE_\lambda = -\int_0^{2\pi} \cot\frac{\lambda}{2}\,dE_\lambda\,. \tag{8.20}$$

Wir definieren nun eine Schar neuer Projektoren:

$$F_\kappa := E_\lambda\,, \quad \text{wenn } \kappa = -\cot\frac{\lambda}{2}\,.$$

κ läuft dann von $-\infty$ bis ∞. Man überprüft leicht, daß die F_κ wiederum eine Zerlegung der Einheit sind. Damit erhalten wir aus (8.20) die gewünschte Darstellung $A = \int_{-\infty}^{\infty} \kappa\,dF_\kappa$. ■

Was hat die *Spektral*darstellung mit dem *Spektrum* zu tun? Die Werte des Spektrums sind die möglichen scharfen Werte des betrachteten selbstadjungierten Operators A, nämlich seine Eigenwerte und Beinahe-Eigenwerte. Und die Antwort auf die obige Frage ist, daß das Integral $A = \int \lambda\,dE_\lambda$ genau über diese Werte, d. h. über $\operatorname{Sp} A$ zu erstrecken ist; anders:

$$dE_\lambda \neq 0 \quad \Longleftrightarrow \quad \lambda \in \operatorname{Sp} A\,. \tag{8.21}$$

Dabei habe ich eine formale Sprechweise benutzt:

$$dE_{\lambda_0} \neq 0 \quad :\Longleftrightarrow \quad \forall\varepsilon > 0: \quad \Delta E_{\lambda_0} \equiv E_{\lambda_0+\varepsilon} - E_{\lambda_0-\varepsilon} \neq 0\,.$$

Sei E_λ stetig bei λ_0 und $\Delta E_{\lambda_0} \neq 0$. Man kann ein $\psi \neq 0$ wählen mit $\psi = \Delta E_{\lambda_0}\psi$; und dann ist $A\psi = \int_{\lambda_0-\varepsilon}^{\lambda_0+\varepsilon} \lambda\,dE_\lambda\psi$, also

$$\|A\psi - \lambda_0\psi\|^2 = \left\| \int_{\lambda_0-\varepsilon}^{\lambda_0+\varepsilon} (\lambda - \lambda_0)\,dE_\lambda\psi \right\|^2 = \int_{\lambda_0-\varepsilon}^{\lambda_0+\varepsilon} (\lambda - \varepsilon)^2\,d(\psi, E_\lambda\psi) \leq \varepsilon^2\,\|\psi\|^2\,.$$

λ_0 ist also Beinahe-Eigenwert von A, d. h. $\lambda_0 \in \operatorname{Sp} A$.
dE_{λ_0} ist insbesondere dann $\neq 0$, wenn E_λ bei λ_0 unstetig ist, d. h. einen Sprung hat, also

$$\lim_{\varepsilon\searrow 0} \Delta E_{\lambda_0} = \lim_{\varepsilon\searrow 0} (E_{\lambda_0+\varepsilon} - E_{\lambda_0-\varepsilon}) = P_{\lambda_0} \neq 0$$

ist. Für jede Zerlegung des Integrationsintervalls $(\lambda_0 - \varepsilon, \lambda_0 + \varepsilon)$ ist dann natürlich $(E_{\lambda_k} - E_{\lambda_{k-1}})P_{\lambda_0} = 0$, außer wenn $\lambda_{k-1} < \lambda_0 \leq \lambda_k$; daher ist *für alle* $\varepsilon > 0$

$$\begin{aligned} AP_{\lambda_0}\varphi &= \int \lambda\, dE_\lambda\, P_{\lambda_0}\varphi = \int_{\lambda_0-\varepsilon}^{\lambda_0+\varepsilon} \lambda\, dE_\lambda\, P_{\lambda_0}\varphi \\ &\approx \lambda_0 \int_{\lambda_0-\varepsilon}^{\lambda_0+\varepsilon} dE_\lambda\, P_{\lambda_0}\varphi = \lambda_0\, P_{\lambda_0}\varphi \end{aligned}$$

(Genaue Rechnung durch eine Abschätzung wie im Fall für stetiges E_λ.) Der durch den Sprung definierte Operator P_{λ_0} ist folglich der zum Eigenraum zu λ_0 gehörende Projektor.

Damit haben wir die Richtung „$\Rightarrow$" in (8.21) gezeigt. (Rechnungen mit Spektraloperatoren werden oft klarer, wenn man an die zugehörigen Unterräume denkt.)

Sei nun $dE_{\lambda_0} = 0$, also E_λ konstant in einer kleinen Umgebung von λ_0. Dann braucht man das Integral $\int \lambda\, dE_\lambda$ nicht über diese Umgebung zu erstrecken, im verbleibenden Integrationsgebiet ist also $(\lambda - \lambda_0)^{-1}$ nicht singulär und beschränkt, folglich $R(\lambda_0; A) = \int (\lambda - \lambda_0)^{-1}\, dE_\lambda$ ein beschränkter Operator, definiert auf ganz $\mathcal{H}$, und das heißt, daß $\lambda \notin \operatorname{Sp} A$ ist; das beweist „$\Leftarrow$".

Die Spektraldarstellung ist also eine Zerlegung des Operators bezüglich der Eigenräume zu den möglichen scharfen oder beliebig scharfen Werten. Diese Überlegungen zeigen also, wie man zu einem beliebigen Wert im Spektrum eines selbstadjungierten Operators Eigenvektoren oder Beinahe-Eigenvektoren konstruiert.

<u>Ergänzende mathematische Bemerkung</u>: Damit die Cayley-Transformierte definiert werden kann, muß A selbstadjungiert sein; das ist erforderlich, damit $(A \pm i\mathbb{1})\mathcal{D}_A = \mathcal{H}$ ist (beide Vorzeichen sind notwendig). Wenn A symmetrisch, aber nicht selbstadjungiert ist, kann man im günstigsten Fall A abschließen zu $\bar{A}$, so daß $\bar{A}$ selbstadjungiert ist und $(\bar{A} \pm i\mathbb{1})\mathcal{D}_{\bar{A}} = \mathcal{H}$. Das geht genau dann, wenn A wesentlich selbstadjungiert ist; dann ist $\bar{A}$ die eindeutige selbstadjungierte Erweiterung, für die alle obigen Überlegungen durchführbar sind.

Im anderen Fall ist $(A + i\mathbb{1})\mathcal{D}_A$ oder $(A - i\mathbb{1})\mathcal{D}_A$ nicht dicht, d. h. — nehmen wir den ersten Fall — es gibt ein ψ, so daß

$$0 = (\psi, (A + i\mathbb{1})\varphi) = ((A^* - i\mathbb{1})\psi, \varphi) \qquad \forall\, \varphi \in \mathcal{D}_A\,;$$

weil D_A dicht ist, folgt $(A^* - i\mathbb{1})\psi = 0$, A^* hat also den Eigenwert i, kann also nicht hermitisch sein. Im zweitgünstigsten Fall hat A^* auch den Eigenwert $-i$, und die Entartungsgrade beider Eigenwerte sind gleich, sagen wir $= k$. Dann gibt es eine k-parametrige Schar von selbstadjungierten Erweiterungen von A, deren Spektren aber untereinander verschieden sind; und das heißt, wie wir gesehen haben, daß auch ihre Spektraldarstellungen verschieden sind.

Ich komme noch einmal auf die Dirac-Schreibweise zurück, die ich in den Abschnitten 1.3 und 4.5 erwähnt habe. Sie ist — wie schon früher gesagt — bequem, wenn man beispielsweise einen Vektor bis auf die Normierung eindeutig charakterisieren kann durch

die Angabe, daß er Eigenvektor der Operatoren $A, B, \ldots$ zu den Eigenwerten $a, b, \ldots$ sei; dann schreibt man einfach $|\, a, b, \ldots\rangle$. Bequem ist auch die Schreibweise $|\,\varphi\rangle\langle\psi|$: das ist ein Operator, der, angewandt auf $|\,\chi\rangle$, einen Vektor parallel zu φ liefert: $|\,\varphi\rangle\langle\psi|\,\chi\rangle$. Dementsprechend ist $|\,\varphi\rangle\langle\varphi|$ der Projektor auf $|\,\varphi\rangle$; wenn $\{|\,\varphi_k\rangle\}$ eine o. n. Basis ist, gilt also $\sum|\,\varphi_k\rangle\langle\varphi_k| = \mathbb{1}$. Ich habe auch schon erwähnt, daß man diese Schreibweise auch für uneigentliche Vektoren benutzt: $Q|\,x\rangle = x|\,x\rangle$. Und wenn ein selbstadjungierter Operator nur uneigentliche Eigenvektoren hat, die eine „Basis“ bilden, schreibt man $\sum|\,\lambda\rangle\langle\lambda| = \mathbb{1}$. Das ist in unsere Schreibweise zu übersetzen als $\int dE_\lambda = \mathbb{1}$; der uneigentliche Projektor $|\,\lambda\rangle\langle\lambda|$ ist $= dE_\lambda$ zu setzen und die Summe als Integral zu interpretieren. In diesem Sinne lautet die Spektraldarstellung des Ortsoperators formal geschrieben einfach $Q = \sum x|\,x\rangle\langle x|$; und im allgemeinen Fall eines Operators, der sowohl ein nicht leeres Punkt- wie auch ein nicht triviales kontinuierliches Spektrum hat, schreibt man nur eine formale Summe, die teilweise als Integral zu interpretieren ist.
Ich wiederhole: das ist oft bequem, doch verwischt es mathematische Unterschiede; und man sollte die Schreibweise nur benutzen, wenn man weiß, wie man alles streng machen kann und wo man vielleicht besonders aufpassen muß, damit sich durch die formale Schreibweise keine Fehler einschleichen.

Anhang: Oberhalb stetige Funktionen

Sei $\mathcal{X}$ ein topologischer Raum, f eine reellwertige Funktion auf X.

Definition: f heißt oberhalb stetig bei $x_0 \in X$, wenn es zu jedem $\varepsilon > 0$ eine Umgebung $\mathcal{U}_\varepsilon(x_0)$ gibt, für die

$$f(x) < f(x_0) + \varepsilon \quad \forall\, x \in \mathcal{U}_\varepsilon(x_0)\,.$$

f ist oberhalb stetig in seinem ganzen Definitionsbereich, wenn

$$\forall a \in \mathbb{R}: \quad \mathcal{N}_a := \{x \in \mathcal{X};\ f(x) < a\} \text{ ist offen.}$$

Wenn $\mathcal{X} = \mathbb{R}$: anschaulich gesagt, ist $f(x)$ oberhalb stetig, wenn es an seinen Sprungstellen den *oberen* Wert als Funktionswert annimmt.[14]

Mit dieser Formulierung zeigt man leicht:
Lemma: Sei $f(x) = \inf f_n(x)$ mit oberhalb stetigen f_n, dann ist f oberhalb stetig.

Beweis: Seien $\mathcal{N}_a^{(n)}$ die entsprechenden Mengen für die f_n, dann ist offensichtlich $\mathcal{N}_a^{(n)} \subset \mathcal{N}_a$ und $\bigcup \mathcal{N}_a^{(n)} \subset \mathcal{N}_a$. Sei $x \notin \bigcup \mathcal{N}_a^{(n)}$; dann gilt $f_n(x) \geq a$ für alle n, also auch $f(x) = \inf f_n(x) \geq a$; daher ist auch $x \notin \mathcal{N}_a$. Folgerung: $\bigcup \mathcal{N}_a^{(n)} = \mathcal{N}_a$. Die $\mathcal{N}_a^{(n)}$ sind nach Voraussetzung offen, also auch $\mathcal{N}_a$. ■

[14] Englische Bezeichnung: upper semi-continuous, kurz u.s.c.; Gegensatz: l.s.c = lower semi-continuous = unterhalb stetig.

4.9 Anwendungen und Ergänzungen

Mit der Spektraldarstellung lassen sich bequem Funktionen selbstadjungierter Operatoren konstruieren:

$$f(A) = \int f(\lambda)\, dE_\lambda, \qquad \mathcal{D}_{f(A)} = \left\{ \varphi;\ \int |f(\lambda)|^2\, d(\varphi, E_\lambda\varphi) < \infty \right\},$$

vergl. die Sätze 4.8.12 und 4.8.13.

Mit ihr lassen sich Aussagen über komplexe Funktionen auf Operatoren übertragen. Ich bringe ein Beispiel:
$f(x) = x,\ x \in \mathbb{R}$ ist eine unbeschränkte Funktion. Wenn man mit beschränkten Funktionen arbeiten will, kann man $r(x; z) := (x - z)^{-1}$ definieren; alle Funktionswerte von f lassen sich mit r berechnen, und r ist beschränkt, solange $\Im z \neq 0$ ist. Genauer: r ist analytisch in $\{z \in \mathbb{C};\ \Im z \neq 0\}$ und hat einen Pol bei $z = x$.
Für die analogen Aussagen über Operatoren betrachtet man einen (unbeschränkten) selbstadjungierten Operator A und seine Resolvente $(A - z\mathbb{1})^{-1}$. Die Aussagen des folgenden Satzes kennen wir bereits zu einem Teil; ich fasse damit eine Reihe von Eigenschaften der Resolvente zusammen.

Satz 4.9.1 *(i) Die Resolvente $R(z; A) = (A - z\mathbb{1})^{-1}$ ist für $z \notin \operatorname{Sp} A$ ein beschränkter Operator und ist analytisch in $z \in \mathbb{C}\backslash\operatorname{Sp} A$. Es gilt*

$$R(z; A)\mathcal{H} = \mathcal{D}(A)\,.$$

(ii) Wenn a ein isoliert liegender Eigenwert von A ist, hat R(z;A) einen Pol bei $z = a$ und es gilt

$$-\frac{1}{2\pi i}\oint R(z; A)\, dz = P_a = \text{ Projektor auf den Eigenraum zu } a\,,$$

wobei die Integration über eine geschlossene Kurve um a zu führen ist, die in der Resolventenmenge $\mathsf{P}(A) = \mathbb{C}\backslash\operatorname{Sp} A$ liegt und keinen weiteren Punkt des Spektrums umschließt. Weiter gilt für analytisches f

$$-\frac{1}{2\pi i}\oint f(z)R(z; A)\, dz = f(a)P_a = f(A)P_a\,.$$

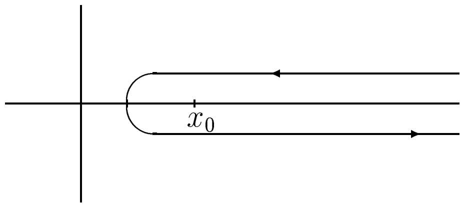
(iii) Sei Γ eine geschlossene Kurve, die $\operatorname{Sp} A$ im Inneren enthält, $f(z)$ sei analytisch in einer Umgebung von Γ. Dann gilt

$$-\frac{1}{2\pi i}\int_\Gamma f(z)R(z; A)\, dz = f(A)\,.$$

(iv) Wenn x im kontinuierlichen Spektrum von A liegt, existiert $R(x, A)$ als unbeschränkter Operator.

Eine operatorwertige Funktion $R(z)$ über $\mathbb{C}$ heißt analytisch, wenn alle Matrixelemente $(\varphi, R(z)\varphi)$ analytische Funktionen sind; das ist gleichbedeutend damit, daß $R(z)$ komplex differenzierbar ist und sich als Potenzreihe in z darstellen läßt.

Zum Beweis der Punkte (i), (ii) und (iii) für den Fall eines selbstadjungierten Operators kann man die Spektraldarstellung von A heranziehen; die Aussagen gehen dann in entsprechende Aussagen über analytische Funktionen über, die aus der Funktionentheorie bekannt sind; die Aussagen in (ii) entsprechen dem Cauchyschen Integralsatz. Anmerkung zu (iii): Die Formel gilt auch, wenn z. B. $\mathrm{Sp}\, A = \{x \in \mathbb{R};\ x \geq x_0\}$ und Γ die folgende Kurve ist:



Gemäß (i) existiert $R(z;A)^{-1}$ auf $\mathcal{D}(A)$, und $A = R(z;A)^{-1} + z\mathbb{1}$. $R(z;A)$ ist also ein beschränkter Operator, der alle Information über A enthält.

Ein wichtiges Beispiel für Funktionen von Operatoren, die mit der Spektraldarstellung konstruiert werden können, ist der **Satz von Stone**:

Theorem 4.9.2 *$A = \int \lambda\, dE_\lambda$ sei ein selbstadjungierter Operator mit Definitionsbereich $\mathcal{D}_A$. Dann ist*

$$\left\{U_t := e^{itA} = \int e^{it\lambda}\, dE_\lambda;\ t \in \mathbb{R}\right\}$$

eine Gruppe unitärer Operatoren, für die gilt

$$U_0 = \mathbb{1}\,, \quad U_t U_s = U_{t+s}\,, \quad \textit{und folglich } U_t^{-1} = U_{-t}\,.$$

Für $\varphi \in \mathcal{D}_A$ existiert

$$s\text{-}\lim_{t\to 0} \frac{1}{t}\,(U_t - \mathbb{1}) = iA\,;$$

und U_t ist die eindeutige Lösung der Gleichung (auf $\mathcal{D}_A$)

$$\frac{d}{dt} U_t = iAU_t$$

mit der Anfangsbedingung $U_0 = \mathbb{1}$.

Beweis: Die ersten Aussagen, die Unitarität von U_t und die Gruppeneigenschaft, folgen direkt aus der Spektraldarstellung.

Um die Differenzierbarkeit zu zeigen, schätzen wir für $\varphi \in \mathcal{D}_A$ den folgenden Ausdruck ab:

$$\left\| \left(\frac{1}{t}(U_t - \mathbb{1}) - iA \right) \varphi \right\|^2 = \int \left| \frac{1}{t}(e^{it\lambda} - 1) - i\lambda \right|^2 d(\varphi, E_\lambda \varphi) .$$

$\varphi \in \mathcal{D}_A$ impliziert, daß

$$\int_{\mathbb{R} \setminus \Delta} \lambda^2 \, d(\varphi, E_\lambda \varphi) < \varepsilon$$

für ein hinreichend großes endliches Intervall Δ. Und weil

$$\left| \frac{1}{t}(e^{it\lambda} - 1) \right|^2 = \lambda^2 \left| \frac{2}{\lambda t} \sin \lambda t \right|^2 \leq \lambda^2 \max_x \left| \frac{2 \sin x}{x} \right|^2 ,$$

kann man auch

$$\int_{\mathbb{R} \setminus \Delta} \left| \frac{1}{t}(e^{it\lambda} - 1) - i\lambda \right|^2 d(\varphi, E_\lambda \varphi)$$

beliebig klein machen, gleichmäßig in t. Es ist klar, daß $\left| \frac{1}{t}(e^{it\lambda} - 1) - i\lambda \right| < \varepsilon$ für hinreichend kleines t, gleichmäßig in $\lambda \in \Delta$, d. h. auch das Integral $\int_\Delta \ldots$ wird klein, also $\lim_{t \to 0} \|\ldots\| = 0$.

Man überzeugt sich leicht, daß $U_t \mathcal{D}_A \subset \mathcal{D}_A$. Die Differentialgleichung folgt dann auch wieder direkt aus der Spektraldarstellung.

Eindeutigkeit der Lösung: Sei $\widetilde{U}_t$ eine zweite Lösung mit derselben Anfangsbedingung. Es folgt

$$\frac{d}{dt} U_t^{-1} \widetilde{U}_t = \frac{d}{dt} U_{-t} \widetilde{U}_t = -iAU_{-t}\widetilde{U}_t + U_{-t}(iA\widetilde{U}_t) = 0 ,$$

weil A und U_t kommutieren. Das heißt, daß $U_{-t}\widetilde{U}_t = \text{const}\mathbb{1}$, und wegen der Anfangsbedingung muß const $= 1$ sein, d. h. $U_t = \widetilde{U}_t$. ■

Mit diesem Satz haben wir die mathematisch saubere Begründung dafür, daß wir die Lösung der Schrödingergleichung $i\hbar\partial_t\psi_t = H\psi_t$ in der Form $\psi_t = \exp^{-\frac{i}{\hbar}Ht} \psi_0$ schreiben dürfen.

Hier ist noch ein weiterer nützlicher Satz ohne Beweis:

Satz 4.9.3 *Ist T ein abgeschlossener Operator mit dichtem Definitionsbereich $\mathcal{D}_T$, so ist T^*T (mit geeignetem $\mathcal{D}_{T^*T} \subset \mathcal{D}_T$) ein selbstadjungierter Operator.*

Dieser Operator ist nicht-negativ, d. h. $T^*T = \int_0^\infty \lambda \, dE_\lambda$, dann existiert $|T| := \int_0^\infty \sqrt{\lambda} \, dE_\lambda$. Analog zur Polarzerlegung einer komplexen Zahl $z = |z|e^{i\varphi}$ kann man T mit einem geeigneten isometrischen Operator V darstellen als $T = V|T|$.

Es folgen zwei Ergänzungen mit physikalischem Hintergrund.

E.1 Die kanonischen Vertauschungsrelationen, die ich hier der Einfachheit halber nur in einer Dimension schreibe:
$$[P, Q] = \frac{\hbar}{i}\mathbb{1}\,,$$
sind die Grundlage der Quantenmechanik. P und Q sind „kanonisch konjugierte Observable“. Man muß bei den Definitionsbereichen aufpassen, wenn man mit unbeschränkten Operatoren zu tun hat: bei Darstellung in $\mathcal{L}^2(\mathbb{R})$ gilt die Relation jedenfalls auf $\mathcal{D}_0 = \mathcal{S}(\mathbb{R})$, und auf diesem Bereich sind P und Q wesentlich selbstadjungiert.

Man kann fragen, ob es nicht auch andere Darstellungen gibt, bei denen man mit beschränkten Operatoren auskommt, oder wenigstens nur einen unbeschränkten Operator hat. Man findet in der Literatur manchmal die Aussage, daß auch der Drehimpuls J bezüglich einer festen Achse und der Drehwinkel φ um diese Achse kanonisch konjugiert sind. In Polarkoordinaten geschrieben ist $J = \frac{\hbar}{i}\frac{\partial}{\partial\varphi}$; und formal ist

$$[J, \varphi] = \frac{\hbar}{i}\mathbb{1}\,.$$

Der Operator der Multiplikation mit $\varphi \in [0, 2\pi)$ ist natürlich beschränkt. Aber so naiv formuliert ist diese Aussage falsch. Ich will Ihnen zeigen, daß hier etwas faul ist.

Es ist schnell zu sehen, daß die Kommutator-Relation

$$[A, B] = c\mathbb{1}, \qquad c \in \mathbb{C} \tag{9.4}$$

nicht für beschränkte selbstadjungierte Operatoren A und B gelten kann.
Dazu erst eine Vorbemerkung: man rechnet leicht nach, daß

$$[A, BC] = [A, B]C + B[A, C]\,.$$

Das ist die Produktregel der Kommutatorbildung, sie erinnert, wie manches andere bei Kommutatoren, an die Differentiation. (Man merke sich: der Kommutator ist mit jedem Faktor zu bilden, *die Reihenfolge der Faktoren bleibt unverändert.* Nehmen wir nun an, es gäbe beschränkte $A = A^*$ und $B = B^*$, die (9.4) erfüllen. Dann folgt sofort mit der Produktregel
$$[A, B^n] = cn\, B^{n-1}\,,$$
denn die Gleichung gilt offensichtlich für $n = 1$, und der Rest folgt durch Induktion:

$$[A, B^n] = [A, B]B^{n-1} + B[A, B^{n-1}] = c\, B^{n-1} + B\,(n-1)cB^{n-2}\,.$$

Nehmen wir die Norm dieser Gleichung:

$$|c|n\,||B^{n-1}|| = ||[A, B^n]|| \leq 2\,||A||\;||B^n|| \leq 2\,||A||\;||B||\;||B^{n-1}||\,;$$

Nun ist $\|B^{n-1}\| = \|B\|^{n-1}$, weil B selbstadjungiert ist (das sieht man schnell mit der Spektraldarstellung), und $\|B\| \neq 0$, also

$$2\,||A||\;||B|| \geq |c|n \quad \forall\, n \in \mathbb{N}\,,$$

und das ist ein Widerspruch zur Beschränktheit von A.

Mindestens einer der Operatoren muß also unbeschränkt sein. Nehmen wir also an, wir hätten einen beschränkten selbstadjungierten Operator B und einen selbstadjungierten Operator A mit Definitionsbereich $\mathcal{D}_A$, eine dichte Menge $\mathcal{D}_0 \subset \mathcal{D}_A$, für die auch $B\mathcal{D}_0 \subset \mathcal{D}_A$ ist, und es gelte (9.4) auf $\mathcal{D}_0$. Wenn $\varphi, \psi \in \mathcal{D}_0$, folgt daraus (mit der Symmetrie von B und A)

$$(A\varphi, B\psi) - (B\varphi, A\psi) = c(\varphi, \psi)\,. \tag{9.5}$$

(Wenn wir die Vertauschungsrelation in dieser Form schreiben, können wir auf die Voraussetzung $B\mathcal{D}_0 \subset \mathcal{D}_A$ verzichten; man sagt, der „schwache Kommutator“ sei $= c\mathbb{1}$.) Diese Relation genügt zur Ableitung der Unschärferelation: (9.5) gilt auch für $\hat{A} := A - \langle A\rangle_\varphi$ und $\hat{B} := B - \langle B\rangle_\varphi$; mit $\psi = \varphi \in \mathcal{D}_0$:

$$2i\Im(\hat{A}\varphi, \hat{B}\varphi) = (\hat{A}\varphi, \hat{B}\varphi) - (\hat{B}\varphi, \hat{A}\varphi) = c(\varphi, \varphi)\,;$$

und damit folgt wie früher

$$\left|\frac{|c|}{2}(\varphi, \varphi)\right|^2 = \left|\Im(\hat{A}\varphi, \hat{B}\varphi)\right|^2 \le \left|(\hat{A}\varphi, \hat{B}\varphi)\right|^2 \le \left\|\hat{A}\varphi\right\|^2 \left\|\hat{B}\varphi\right\|^2 ,$$

und daher

$$\Delta_\varphi A\, \Delta_\varphi B \ge \frac{|c|}{2}\,.$$

Nun ist B nach Voraussetzung beschränkt, somit auch $\Delta_\varphi B$: sei $\|\varphi\| = 1$, dann gilt $(\Delta_\varphi B)^2 = (\varphi, B^2\varphi) - (\varphi, B\varphi)^2 \le \|B^2\| + \|B\|^2 \le 2\,\|B\|^2$; also

$$\Delta_\varphi A \ge \frac{|c|}{2\sqrt{2}\,\|B\|} > 0 \qquad \text{gleichmäßig in } \varphi \in \mathcal{D}_0\,.$$

A ist nach Voraussetzung selbstadjungiert, hat also Eigenvektoren oder wenigstens Beinahe-Eigenvektoren, für die $\Delta_\varphi A$ beliebig klein werden kann. Wir haben also wieder einen Widerspruch, wenn $\mathcal{D}_0 = \mathcal{D}_A$. Deshalb muß man $\mathcal{D}_0$ derart wählen, daß es zwar dicht ist, so daß man wohl Beinahe-Eigenvektoren aus $\mathcal{D}_0$ heraus approximieren kann, doch darf für diese approximierenden Vektoren φ_α die Folge $A\varphi_\alpha$ nicht konvergieren. Das ist tatsächlich möglich: man wähle z. B. $\mathcal{H} = \mathcal{L}^2[0,1]$, $Q = x$, $P = \frac{1}{i}\partial_x$ und $\mathcal{D}_0 = \{\psi;\ \psi \text{ absolut stetig}, \psi(0) = \psi(1) = 0\}$. Q ist dann beschränkt, $\|Q\| = 1$, und P ist symmetrisch, hat aber — siehe Ergänzung E 5 im Abschitt 4.5 — viele selbstadjungierte Erweiterungen. Man kann ganz allgemein zeigen, daß unter den obigen Voraussetzungen A auf $\mathcal{D}_0$ nur symmetrisch, aber nicht wesentlich selbstadjungiert sein kann. Es ist also falsch, zu sagen, es gäbe kanonisch konjugierte *Observable*, von denen eine beschränkt ist.

E.2 In der Mechanik kennt man das Noethersche Theorem (in Göttingen kommt man an Emmi Noether nicht vorbei; denn wenn man im Mathematischen Institut die

südliche Treppe benutzt, kommt man ja immer an ihr vorbei):
Zu jeder kontinuierlichen einparametrigen Schar von Symmetrien eines Systems gibt es eine Erhaltungsgröße.
Zum Beispiel: Wenn die Kräfte nicht von der Zeit abhängen, ist das System invariant gegenüber Zeittranslationen, die zugehörige Erhaltungsgröße ist die Gesamtenergie. Ist das Potential invariant gegenüber räumlichen Drehungen (das ist eine dreiparametrige Symmetriegruppe), dann ist der (dreikomponentige) Gesamtdrehimpuls erhalten; usw.

Zu diesem Theorem gibt es ein quantenmechanisches Analogon. Symmetrien werden dargestellt durch unitäre Operatoren. Nehmen wir als Beispiel ein System, das invariant ist unter Drehungen um die z-Achse; diese Drehungen entsprechen einer einparametrigen Schar $\{U_D(\varphi)\}$ von unitären Operatoren, die stetig von φ abhängen, $U_D(0) = \mathbb{1}$. Die Invarianz des Systems drückt sich darin aus, daß der Hamiltonoperator invariant ist:

$$U_D(\varphi) H U_D(\varphi)^{-1} = H\,. \tag{9.6}$$

Wir können nun auf einem geeigneten Bereich

$$\hat{L}_3 := \operatorname*{s-lim}_{\varphi\to 0} \frac{1}{\varphi}(U_D(\varphi) - \mathbb{1})$$

definieren; wir werden sehen, daß der Definitionsbereich von $\hat{L}_3$ dicht und daß $U_D(\varphi)$ stark stetig ist. Man sieht leicht, daß $-i\hat{L}_3$ symmetrisch ist, und man kann zeigen, daß es sogar selbstadjungiert ist. Dann ist nach dem Satz von Stone (Theorem 4.9.2) $U_D(\varphi)$ gegeben durch

$$U_D(\varphi) = e^{i\varphi(-i\hat{L}_3)} = e^{\varphi \hat{L}_3}\,.$$

Schreiben wir die Invarianzgleichung (9.6) als $U_D(\varphi)H = HU_D(\varphi)$ und differenzieren nach φ bei $\varphi = 0$: wir finden $i\hat{L}_3 H = H(i\hat{L}_3)$; d. h.

$$[H, \hat{L}_3] = 0\,.$$

Ein Operator A, der mit H vertauscht, ist aber als Erhaltungsgröße zu bezeichnen! Denn er vertauscht auch mit $U(t) = e^{-i\frac{t}{\hbar}H}$, und das bedeutet, daß alle Erwartungswerte von A zeitunabhängig sind:

$$\langle A\rangle_{\psi_t} = \frac{(U(t)\psi_0, AU(t)\psi_0)}{(\psi_0,\psi_0)} = \frac{(\psi_0, U(-t)AU(t)\psi_0)}{(\psi_0,\psi_0)} = \frac{(\psi_0, A\psi_0)}{(\psi_0,\psi_0)} = \langle A\rangle_{\psi_0}\,.$$

(Umgekehrt folgt aus der Tatsache, daß alle Erwartungswerte eines Operators zeitunabhängig sind, daß er mit $U(t)$ und daher auch mit H vertauschen muß.) Ich habe den obigen Operator natürlich nicht grundlos mit dem Buchstaben bezeichnet, den man üblicherweise für den Drehimpuls benutzt. Der unitäre Darstellungsoperator einer Drehung R ist in der Ortsdarstellung gegeben durch $(U^D(R)\psi)(\vec{x}) = \psi(R^{-1}\vec{x})$; im

Fall einer Drehung um die 3-Achse ist $(U_D(\varphi)\psi)(\vec{x}) = \psi(R_3(-\varphi)\vec{x})$, wobei die Drehmatrix gegeben ist durch

$$R_3(\varphi) = \begin{pmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Differenzieren bei $\varphi = 0$ ergibt

$$\frac{d}{d\varphi} R_3(\varphi)\Big|_{\varphi=0} = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} =: \tilde{I}_3\,.$$

Jetzt ist klar, daß $U_D(\varphi)$ stark stetig ist: $(U_D(\varphi)\psi)(\vec{x}) = \psi(R_3(-\varphi)\vec{x})$ hängt stetig von φ ab; und die Berechnung $\hat{L}_3$ ist einfach:

$$\begin{aligned} \frac{d}{d\varphi}(U_D(\varphi)\psi)(\vec{x})\Big|_{\varphi=0} &= \frac{d}{d\varphi}\psi(R_3(-\varphi)\vec{x})\Big|_{\varphi=0} = \vec{\partial}\psi(\vec{x})\cdot\frac{d}{d\varphi}(R_3(-\varphi)\vec{x})\Big|_{\varphi=0} \\ &= \vec{\partial}\psi(\vec{x})\cdot(-\tilde{I}_3\vec{x}) = (x_2\partial_{x_1} - x_1\partial_{x_2})\psi(\vec{x})\,; \end{aligned}$$

also ist

$$\hat{L}_3 = -\frac{i}{\hbar}\left(x_1\,\frac{\hbar}{i}\partial_{x_2} - x_2\,\frac{\hbar}{i}\partial_{x_1}\right) = -\frac{i}{\hbar}(\vec{Q}\times\vec{P})_3 \equiv -ih^{-1}L_3$$

im wesentlichen die 3-Komponente des Drehimpulsoperators $\vec{L} = \vec{Q}\times\vec{P}$. Aus dieser Rechnung sieht man auch, daß der Definitionsbereich von $\hat{L}_3$ dicht ist: er enthält offensichtlich $\mathcal{S}(\mathbb{R}^3)$, d. h. diejenige Menge, auf der $\vec{Q}$ und $\vec{P}$ wesentlich selbstadjungiert sind.

$\tilde{I}_3$ ist schiefsymmetrisch, daher ist $I_3 := i\tilde{I}_3$ ein selbstadjungierter Operator in $\mathbb{C}^3$, wir können auch auf ihn den Satz von Stone anwenden und schließen, daß

$$R_3(\varphi) = e^{-i\varphi I_3}$$

ist.
Mit den Drehungen um die 1- bzw. 2-Achse können wir ebenso verfahren; wir erhalten entsprechende Größen

$$\tilde{I}_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \qquad \tilde{I}_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}.$$

Die Drehungen um die verschiedenen Achsen vertauschen nicht miteinander. Das drückt sich darin aus, daß die $\tilde{I}_k$ nicht kommutieren: man rechnet leicht nach, daß für $I_k := i\tilde{I}_k$ die Relationen

$$[I_k, I_l] = i\,\varepsilon_{klm} I_m \tag{9.7}$$

gelten. Hierbei ist ε_{klm} der total antisymmetrische 3-3-Tensor, und ich habe wieder die Summationskonvention benutzt. Abgesehen von einem Faktor $\hbar$ gelten dieselben Relationen für den Drehimpuls:

$$[L_k, L_l] = i\hbar\,\varepsilon_{klm} L_m\,; \tag{9.8}$$

das überprüft man sofort mit den kanonischen Vertauschungsrelationen. Wir haben also nicht nur eine Darstellung der Drehgruppe durch die unitären Operatoren $\left\{U^D(R)\right\}$, sondern gleichzeitig eine Darstellung der „Lie-Algebra" der „Erzeugenden" I_k durch selbstadjungierte Operatoren $\hbar^{-1} L_k$. „Lie-Algebra" soll hier nur heißen: die multiplikative Verknüpfung der Algebra ist die Kommutatorbildung, die sich auf die Darstellung überträgt.

Die Kommutator-Relationen (9.8) sind das wesentliche Charakteristikum für den quantenmechanischen Drehimpuls, unabhängig von der speziellen Darstellung

(genau so, wie die kanonischen Vertauschungsrelationen charakteristisch für Ort und Impuls sind); sie gehen direkt auf die Struktur der zugehörigen Symmetriegruppe zurück. Diese Aussage sollte man sich merken, ich komme im nächsten Abschnitt noch einmal darauf zurück.

4.10 Direkte Summe und Tensorprodukt von Hilberträumen

Wenn man die Konfiguration oder den Phasenraum eines Systems von zwei Teilchen beschreiben will, braucht man zwei Koordinaten $\{x_1, x_2\}$, $x_i \in \mathbb{R}^3$ bzw. $\in \mathbb{R}^6$. Der Raum der Paare $\{x_1, x_2\}$ wird wieder ein linearer Raum, $\simeq \mathbb{R}^6$ bzw. $\simeq \mathbb{R}^{12}$, wenn man die Addition und Multiplikation mit einem Faktor komponentenweise erklärt:

$$\{x_1, x_2\} + \{y_1, y_2\} := \{x_1 + y_1, x_2 + y_2\}, \quad c\{x_1, x_2\} := \{cx_1, cx_2\}, \tag{10.1}$$

man kann ein Skalarprodukt einführen durch

$$(\{x_1, x_2\}, \{y_1, y_2\}) := (x_1, y_1) + (x_2, y_2). \tag{10.2}$$

(Das entspricht dem üblichen Skalarprodukt im $\mathbb{R}^6$ bzw. $\mathbb{R}^{12}$.)

Diese Konstruktion ist nicht auf Vektoren des $\mathbb{R}^n$ beschränkt, ich kann sie genauso gut für Vektoren ψ_i zweier (endlich oder unendlich-dimensionaler) Hilberträume $\mathcal{H}_i$ ausführen — man schreibe ψ_i in den Gleichungen (10.1) und (10.2) — und erhält einen neuen Hilbertraum $\mathcal{H}_S$ der Paare $\{\psi_1, \psi_2\}$. Es ist nicht schwer, zu kontrollieren, daß der Raum $\mathcal{H}_S$ in der aus (10.2) folgenden Norm vollständig ist. Man sieht sofort, daß man die ursprünglichen Räume in den neuen Raum einbetten kann:

$$\mathcal{H}_1 \simeq \mathcal{H}_1^S \equiv \{\{\psi_1, 0\};\ \psi_1 \in \mathcal{H}_1\};\ \mathcal{H}_2 \simeq \mathcal{H}_2^S \equiv \{\{0, \psi_2\};\ \psi_2 \in \mathcal{H}_2\},$$

und diese beiden Räume $\mathcal{H}_i^S \subset \mathcal{H}_S$ sind orthogonal zueinander, wie man an (10.2) sieht. $\mathcal{H}_S$ ist die lineare Hülle dieser beiden Unterräume (enthält natürlich sehr viel mehr als $\mathcal{H}_1^S \cup \mathcal{H}_2^S$).
Man nennt den entstehenden Hilbertraum die „direkte Summe“ der $\mathcal{H}_i$:

$$\mathcal{H}_S = \mathcal{H}_1 \oplus \mathcal{H}_2 .$$

Die Konstruktion entspricht $\mathbb{R}^3 \times \mathbb{R}^3 \simeq \mathbb{R}^6$; die Dimensionen addieren sich, was naürlich bei unendlich-dimensionalen Räumen nicht auffällt. Mit „$\times$“ bezeichnet man gewöhnlich das kartesische oder direkte Produkt:

$$\mathcal{H}_1 \times \mathcal{H}_2 = \{\{\psi_1, \psi_2\};\ \psi_i \in \mathcal{H}_i\} = \text{Menge der geordneten Paare.}$$

Das ist noch kein linearer Raum; aber mit (10.1) und (10.2) wird daraus ein linearer Raum mit Skalarprodukt. (Im Fall der $\mathbb{R}^n$ führt man für den so entstehenden Raum im allgemeinen kein neues Symbol ein.)

Wir kennen natürlich schon diese Situation, daß man den Hilbertraum in paarweise orthogonale Unterräume zerlegt:

Seien P_k orthogonale Projektoren: $P_kP_l = \delta_{kl}P_k$, $\sum P_k = \mathbb{1}$. Dann spannen die paarweise orthogonalen Unterräume $\mathcal{H}_k = P_k\mathcal{H}$ den ganzen Raum $\mathcal{H}$ auf: $\mathcal{H} \ni \varphi = \sum P_k\varphi$ ist eine Linearkombination von Vektoren aus $\mathcal{H}_k$. Hier haben wir im allgemeinen sogar unendlich viele orthogonale Unterräume. Die Umkehrung dieser Zerlegung, die Bildung einer direkten Summe unendlich vieler Räume, geht auch in diesem Fall:

$\mathcal{H}_1,\ \mathcal{H}_2,\ \ldots$ seien (endlich oder unendlich viele) Hilberträume mit Skalarprodukten $(\cdot,\cdot)_1, (\cdot,\cdot)_2,\ \ldots$. Die Menge der abbrechenden Folgen

$$\mathcal{D} = \{\Phi = (\varphi_1,\ \varphi_2,\ldots);\ \varphi_k \in \mathcal{H}_k,\ \text{nur endlich viele } \varphi_k \neq 0\}$$

wird zu einer linearen Mannigfaltigkeit, wenn man definiert

$$\lambda\Phi + \Psi := (\lambda\varphi_1 + \psi_1,\ \lambda\varphi_2 + \psi_2, \ldots)\,.$$

In $\mathcal{D}$ kann man ein Skalarprodukt definieren durch

$$(\Phi, \Psi) := \sum(\varphi_k, \psi_k)_k\,. \tag{10.3}$$

Die Summen sind endlich; und man sieht leicht, daß das Skalarprodukt positiv definit ist, $\mathcal{D}$ läßt sich also mit der zugehörigen Norm zu einem Hilbertraum vervollständigen. (Man überzeugt sich leicht, daß bei endlich vielen Summanden die Vervollständigung nicht notwendig ist; $\mathcal{D}$ ist bereits vollständig.)

Definition 4.10.4 *Die Vervollständigung von $\mathcal{D}$ in der mit (10.3) definierten Norm $\|\Phi\| = \sqrt{(\Phi, \Phi)}$ zu einem Hilbertraum $\mathcal{H}$ heißt die direkte Summe der $\mathcal{H}_k$: $\mathcal{H} = \bigoplus_k \mathcal{H}_k$.*

Wenn man eine unendliche direkte Summe betrachtet, enthält $\mathcal{H}$ dank der Vervollständigung auch unendliche Folgen $(\varphi_1,\ \varphi_2,\ldots)$ mit $\sum \|\varphi_k\|^2 < \infty$.
Man beachte, daß die Dimensionen der $\mathcal{H}_k$´s sehr unterschiedlich sein können.

Durch $I:\ \varphi_k \mapsto (0, 0, \ldots, \varphi_k, 0, \ldots)$ kann man $\mathcal{H}_k$ in $\mathcal{H}$ isomorph einbetten, $I(\mathcal{H}_k) \subset \mathcal{H}$; und die eingebetteten Räume sind paarweise orthogonal. Dann ist $\mathcal{D}$ die lineare Hülle von $\bigcup_k I(\mathcal{H}_k)$. Wenn man φ_k und $I(\varphi_k)$ identifiziert, also auch $(\varphi_1,\ \varphi_2, \ldots) = \sum \varphi_k$ schreiben kann, läßt sich das Skalarprodukt $(\sum_k \varphi_k, \sum_l \psi_l)$ formal definieren durch die Angabe: $(\cdot,\cdot)$ ist distributiv und $(\varphi_k, \psi_l) = \delta_{kl}(\varphi_k, \psi_k)$. Diese Definition entspricht genau Gleichung (10.3). Man hat also die Hilberträume $\mathcal{H}_k$ sozusagen orthogonal aneinandergesetzt, mit Linearkombinationen aufgefüllt und das Ganze durch Vervollständigung abgerundet.
Folgerung: Wenn in einem Hilbertraum eine Menge von paarweise orthogonalen Projektoren gegeben ist mit $\sum P_k = \mathbb{1}$, dann gilt $\mathcal{H} \simeq \bigoplus_k P_k\mathcal{H}$.

Einen Operator A in $\mathcal{H}_k$ kann man auf ganz $\mathcal{H}$ fortsetzen durch

$$I(A)\sum I(\varphi_l) := I(A\varphi_k)\,;$$

d. h. $I(A)$ wirkt auf der Einbettung $I(\mathcal{H}_k)$ wie A auf $\mathcal{H}_k$ und ist auf $I(\mathcal{H}_k)^\perp$ der Nulloperator. Natürlich gibt es auch andere Operatoren, die von einem Unterraum in einen anderen führen können.

(Um sehr genau zu sein: die oben definierte direkte Summe sollte als *orthogonale* direkte Summe bezeichnet werden. Man schreibt auch für endliches n $\mathcal{H} = \mathcal{H}_1 \oplus \ldots \oplus \mathcal{H}_n$, wenn es für jeden Vektor $\varphi \in \mathcal{H}$ eine *eindeutige* Zerlegung $\varphi = \varphi_1 + \ldots + \varphi_n$ gibt, ohne zu fordern, daß die $\mathcal{H}_k$ paarweise orthogonal sind.)

Der Schluß, den ich Ihnen hier suggeriert habe, daß $\mathcal{H}_1 \oplus \mathcal{H}_2$ der geeignete Hilbertraum ist, um zwei quantenmechanische Teilchen zu beschreiben, ist aber falsch! Für zwei Teilchen brauchen wir eine Wellenfunktion $\psi(\vec{x}_1, \vec{x}_2)$ von zwei Koordinaten, deren Absolutquadrat die Wahrscheinlichkeitsdichte liefert, ein Teilchen bei $\vec{x}_1$ und das andere bei $\vec{x}_2$ zu finden; ein solches ψ kann nicht als Summe zweier $\psi_i(\vec{x}_i)$ geschrieben werden, denn es muß ja normierbar sein: $\int |\psi(\vec{x}_1, \vec{x}_2)|^2\ d^3x_1 d^3x_2 = 1$. (Außerdem: ein Vektor $\{\psi(\vec{x}_1), 0\}$, der ja Element der direkten Summe ist, macht keinen Sinn, denn das zweite Teilchen soll ja vorhanden sein, also nicht beschrieben durch die Komponente 0.) Wenn die Teilchen keine Wechselwirkung miteinander haben, also völlig unabhängig voneinander sind, sieht man leicht, dass man mit $\psi(\vec{x}_1, \vec{x}_2) = \varphi_1(\vec{x}_1)\varphi_2(\vec{x}_2)$ eine vernünftig interpretierbare Funktion hat.

Wir nutzen also eine zweite Möglichkeit, aus dem kartesischen Produkt $\mathcal{H}_1 \times \mathcal{H}_2$ einen neuen Hilbertraum zu bilden. Interessanter als die Summe — und für die Beschreibung mehrerer quantenmechanischer Teilchen notwendig — ist das **Tensorprodukt** von Hilberträumen. (Manchmal spricht man auch vom kartesischen Produkt von Hilberträumen, dann meint man im allgemeinen nicht das kartesische, sondern das jetzt zu definierende Tensorprodukt.) Hier kann man allerdings ohne sehr großen mathematischen Aufwand nur von *endlich vielen* Räumen $\mathcal{H}_k$ ausgehen.[15]
Man betrachtet wieder Folgen $(\varphi_1,\ \varphi_2, \ldots, \varphi_n)$, $\varphi_k \in \mathcal{H}_k$, faßt sie aber nicht als Summen, sondern als Produkte auf und schreibt

$$\Phi \equiv (\varphi_1,\ \varphi_2, \ldots, \varphi_n) \equiv \varphi_1 \otimes \varphi_2 \otimes \ldots \otimes \varphi_n\,.$$

Konsequenterweise definiert man dann die Addition nicht komponentenweise, sondern bildet formale Summen:

$$\mathcal{D} = \left\{ \sum_{r=1}^{m} \varphi_1^{(r)} \otimes \varphi_2^{(r)} \otimes \ldots \otimes \varphi_n^{(r)};\ \varphi_k^{(r)} \in \mathcal{H}_k,\ m \in \mathbb{N} \right\}.$$

[15] Es liegt nahe, $\psi(\vec{x}_1, \vec{x}_2)$ als Element des $\mathcal{L}^2(\mathbb{R}^6)$ aufzufassen. Wir werden sehen, daß $\mathcal{L}^2(\mathbb{R}^6)$ gleich dem Tensorprodukt zweier $\mathcal{L}^2(\mathbb{R}^3)$-Räume ist. Man will in der Quantenmechanik aber nicht nur in der Ortsdarstellung rechnen; manchmal braucht man auch das Tensorprodukt sehr verschiedenartiger Räume, siehe später.

Mit den Definitionen

$$\alpha(\varphi_1 \otimes \varphi_2 \otimes \ldots \otimes \varphi_n) := \varphi_1 \otimes \ldots \otimes (\alpha\varphi_k) \otimes \ldots \otimes \varphi_n, \quad k \text{ beliebig},$$

$$\varphi_1 \otimes \ldots \otimes \varphi_k \otimes \ldots \otimes \varphi_n + \varphi_1 \otimes \ldots \otimes \psi_k \otimes \ldots \otimes \varphi_n := \varphi_1 \otimes \ldots \otimes (\varphi_k + \psi_k) \otimes \ldots \otimes \varphi_n\,,\ k = 1 \ldots n\,,$$

werden in $\mathcal{D}$ gewisse Elemente miteinander identifiziert, und $\mathcal{D}$ wird so zu einer linearen Mannigfaltigkeit. Setzt man in der ersten Gleichung $\alpha = 0$, so sieht man, daß $\Phi = \varphi_1 \otimes \varphi_2 \otimes \ldots \otimes \varphi_n$ mit dem Nullvektor identifiziert werden muß, sobald einer der Faktoren $= 0$ ist. Diese Definitionen sind so, wie man es für ein *Produkt* von φ's erwartet. Auch das Skalarprodukt wird als *Produkt* von Skalarprodukten definiert:

$$(\varphi_1 \otimes \ldots \otimes \varphi_n, \psi_1, \otimes \ldots, \otimes \psi_n) = \prod_{k=1}^{n} (\varphi_k, \psi_k)_k\,, \tag{10.5}$$

und dann distributiv auf $\mathcal{D}$ fortgesetzt. Es ist positiv definit.

Definition 4.10.6 *Die Vervollständigung von $\mathcal{D}$ in der zum obigen Skalarprodukt gehörenden Norm heißt* Tensorprodukt *der $\mathcal{H}_k$,*

$$\mathcal{H} = \bigotimes_{k=1}^{n} \mathcal{H}_k = \mathcal{H}_1 \otimes \ldots \otimes \mathcal{H}_n\,.$$

Nehmen wir $n = 2$: man sieht auch am Skalarprodukt, daß das Paar $\{\psi, 0\}$, das im Fall der direkten Summe einen nicht-trivialen Vektor darstellt (und das uns oben Interpretationsschwierigkeiten machte), jetzt mit dem Nullvektor zu identifizieren ist: $(\psi \otimes 0, \varphi_1 \otimes \varphi_2) = (\psi, \varphi_1)(0, \varphi_2) = 0$. (Man sieht an dieser Gleichung auch, daß man für eine ganz strenge Definition von $\mathcal{H}_1 \otimes \mathcal{H}_2$ eigentlich Äquivalenzklassen von Paaren einführen muß.)

Man kann sich leicht davon überzeugen, daß „$\otimes$“ eine assoziative Verknüpfung ist:

$$(\mathcal{H}_1 \otimes \mathcal{H}_2) \otimes \mathcal{H}_3 = \mathcal{H}_1 \otimes (\mathcal{H}_2 \otimes \mathcal{H}_3) \equiv \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \mathcal{H}_3\,.$$

Im Folgenden betrachte ich $n = 2$ und stelle Summe und Produkt einander gegenüber. Seien $\{e_k;\ k \in I_1\}$ und $\{f_k;\ k \in I_2\}$ orthogonale Basen von $\mathcal{H}_1$ bzw. $\mathcal{H}_2$, wobei I_1 und I_2 endliche oder unendliche Indexmengen sein können. Dann ist

$$\begin{aligned} &\{e_k,\ f_l;\ k \in I_1, l \in I_2\} \quad \text{orthogonale Basis von} \quad \mathcal{H}_1 \oplus \mathcal{H}_2\,, \\ &\{e_k \otimes f_l;\ k \in I_1, l \in I_2\} \quad \text{orthogonale Basis von} \quad \mathcal{H}_1 \otimes \mathcal{H}_2\,. \end{aligned}$$

Das erste ist trivial, das zweite fast trivial: jedes $\varphi_1 \otimes \varphi_2 \in \mathcal{D}$ ist Linearkombination von $e_k \otimes f_l$; die Orthogonalität dieser Vektoren folgt aus der *Produkt*struktur des Skalarprodukts (10.5). Demzufolge ist

$$\begin{aligned} \dim(\mathcal{H}_1 \oplus \mathcal{H}_2) &= \dim \mathcal{H}_1 + \dim \mathcal{H}_2\,, \\ \dim(\mathcal{H}_1 \otimes \mathcal{H}_2) &= \dim \mathcal{H}_1 \cdot \dim \mathcal{H}_2\,. \end{aligned}$$

4.10.7 Beispiel: Sei $\dim \mathcal{H}_1 = 3,\ \dim \mathcal{H}_2 = 2$, und

$$\vec{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}, \quad \vec{b} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}:$$

dann läßt sich $\vec{a} \oplus \vec{b}$ bzw. $\vec{a} \otimes \vec{b}$ darstellen als

$$\vec{a} \oplus \vec{b} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ b_1 \\ b_2 \end{pmatrix} \in \mathcal{H}_1 \oplus \mathcal{H}_2, \quad \vec{a} \otimes \vec{b} = \begin{pmatrix} a_1 b_1 \\ a_1 b_2 \\ a_2 b_1 \\ a_2 b_2 \\ a_3 b_1 \\ a_3 b_2 \end{pmatrix} \in \mathcal{H}_1 \otimes \mathcal{H}_2 \,.$$

Man prüft leicht nach, daß diese Darstellungen die richtigen Skalarprodukte ergeben.[16]

$\mathcal{H}_1$ läßt sich auch in das Tensorprodukt $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ isomorph einbetten: Man wähle ein festes $\psi \in \mathcal{H}_2$ mit $\|\psi\| = 1$ und definiere

$$I_\psi :\ \varphi \in \mathcal{H}_1 \mapsto \varphi \otimes \psi \in \mathcal{H} \,.$$

Wenn $\{\psi_l\}$ eine orthonormale Basis von $\mathcal{H}_2$ ist, dann sind I_{ψ_k} und I_{ψ_l} für $k \neq l$ zueinander orthogonale (isomorphe) Einbettungen; man erhält also eine direkte Summe von Kopien von $\mathcal{H}_1$ in $\mathcal{H}$, die Anzahl der direkten Summanden ist $\dim \mathcal{H}_2$. Man prüft ohne Schwierigkeiten (vergl. die obige Basis), daß man damit ganz $\mathcal{H}$ ausschöpft:

$$\mathcal{H}_1 \otimes \mathcal{H}_2 = \bigoplus_{k=1}^{\dim \mathcal{H}_2} I_{\psi_l}(\mathcal{H}_1); \qquad I_{\psi_l}(\mathcal{H}_1) \simeq \mathcal{H}_1 \,.$$

Man kann die Rollen von $\mathcal{H}_2$ und $\mathcal{H}_2$ natürlich vertauschen und $\mathcal{H}_2$ mehrfach im Tensorprodukt einbetten:

$$\mathcal{H}_1 \otimes \mathcal{H}_2 = \bigoplus_{k=1}^{\dim \mathcal{H}_1} I_{\varphi_l}(\mathcal{H}_2); \qquad I_{\varphi_l}(\mathcal{H}_2) \simeq \mathcal{H}_2 \,.$$

Das zeigt auch, daß $\mathcal{H}_1 \otimes \mathcal{H}_2$ und $\mathcal{H}_2 \otimes \mathcal{H}_1$ isomorph sind.

Erweiterung von Operatoren: A operiere auf $\mathcal{H}_1$ und B auf $\mathcal{H}_2$, dann definiert man

$$\begin{aligned} (A \otimes B)(\varphi \otimes \psi) :=&\ (A\varphi) \otimes (B\psi)\,, \\ (A \otimes B) \sum_{\text{endl.}} \alpha_k\, \Phi^{(k)} :=&\ \sum \alpha_k\, (A \otimes B)\Phi^{(k)}\,. \end{aligned} \tag{10.8}$$

[16]Für das Produkt habe ich hier eine Blockdarstellung gewählt, bei der jeder Block durch ein Element von $\vec{a}$ gekennzeichnet ist und in jedem Block alle Komponenten von $\vec{b}$ auftreten. Man könnte es natürlich auch genau umgekehrt machen, muß nur die einmal gewählte Konvention konsequent beibehalten.

Damit erhält man auch eine Einbettung der Operatoren auf $\mathcal{H}_1$ in diejenigen auf $\mathcal{H}$ durch $A \mapsto A \otimes \mathbb{1}_2$. Triviale, aber wichtige Relationen, die direkt aus (10.8) folgen:

$$(A \otimes B)\,(C \otimes D) = AB \otimes CD\,; (A \otimes \mathbb{1}_2)(\mathbb{1}_1 \otimes B) = A \otimes B = (\mathbb{1}_1 \otimes B)(A \otimes \mathbb{1}_2)\,. \quad (10.9)$$

Die folgende Relation ist auch leicht zu zeigen:

$$(A \otimes B)^* = A^* \otimes B^*\,, \quad (10.10)$$

Dazu schreibe man das Matrixelement $(\Psi, A \otimes B\Psi)$ mithilfe orthonormierter Basen $\{e_k\}$ und $\{f_l\}$ mit $\Psi = \sum \alpha_{kl} e_k \otimes f_l$, siehe oben.

Wenn A und B beschränkt sind, dann auch $A \otimes B$:

$$\|A \otimes B\| = \|A\|_1 \; \|B\|_2 \;. \quad (10.11)$$

Die erste Gleichung von (10.8) legt diese Behauptung nahe, doch um sie direkt zu beweisen braucht man die zweite Gleichung, und das wird kompliziert. Wir machen es anders: wir wollen zuerst zeigen, daß

$$\|A \otimes \mathbb{1}\| \le \|A\|_1\,, \quad \|\mathbb{1} \otimes B\| \le \|B\|_2\;.$$

Es genügt, die erste Gleichung zu beweisen, außerdem können wir annehmen, daß A selbstadjungiert ist; denn für Hilbertraumoperatoren gilt ja $\|C^*C\| = \|C\|^2$, und mit (10.10) folgt $\|A \otimes \mathbb{1}\|^2 = \|A^*A \otimes \mathbb{1}\|$. Sei A also selbstadjungiert, dann ist auch $A \otimes \mathbb{1}$ selbstadjungiert; und

$$\|A \otimes \mathbb{1}\| = \sup_{\mathcal{H}_1 \otimes \mathcal{H}_2 \ni \Psi \neq 0} \frac{(\Psi, A \otimes \mathbb{1}\Psi)}{(\Psi, \Psi)}\,.$$

Mit der obigen o. n. Basis$\{e_k \otimes f_l\}$ und $\Psi = \sum_{k,l} \alpha_{kl}\, e_k \otimes f_l$ ergibt sich

$$\begin{aligned}
(\Psi, A \otimes \mathbb{1}\Psi) &= \sum_{i,j,k,l} \overline{\alpha}_{ij} \alpha_{kl}\,(e_i, Ae_k)(f_j, f_l) \\
&= \sum_{i,k,l} \overline{\alpha}_{il} \alpha_{kl}\,(e_i, Ae_k) \\
&= \sum_l \left(\sum_i \alpha_{il} e_i, A \sum_k \alpha_{kl} e_k \right) \\
&\le \sum_l \|A\|_1 \left\| \sum_i \alpha_{il} e_i \right\|^2 = \|A\|_1 \sum_{il} |\alpha_{il}|^2 \\
&= \|A\|_1 \; \|\Psi\|^2\,,
\end{aligned}$$

und deshalb $\|A \otimes \mathbb{1}\| \le \|A\|_1$. Mit (10.9) folgt nun

$$\|A \otimes B\| \le \|A \otimes \mathbb{1}\| \; \|\mathbb{1} \otimes B\| \le \|A\|_1 \; \|B\|_2\;.$$

Andererseits ist für normierte φ_1, φ_2, für die man mit $\|A\varphi_1\|_1$ bzw. $\|B\varphi_2\|_2$ fast die Normen $\|A\|_1$ bzw. $\|B\|_2$ erreicht,

$$\|A \otimes B\| \geq \|A\varphi_1 \otimes B\varphi_2\| = \|A\varphi_1\|_1 \, \|B\varphi_2\|_2 \geq (\|A\|_1 - \varepsilon)(\|B\|_2 - \varepsilon)\,,$$

und ε kann durch geeignete Wahl der φ_i beliebig klein gemacht werden. Daraus folgt (10.11).

Man beachte: die wenigsten interessanten Operatoren auf $\mathcal{H}_1 \otimes \mathcal{H}_2$ haben diese Produktform!

4.10.12 Beispiel: Matrixdarstellung der Operatoren in $\mathcal{H} = \mathbb{C}^2 \otimes \mathbb{C}^2$.
Für die Vektoren wählen wir die Darstellung wie in Beispiel 4.10.7. Als Basis der beiden Faktoren nehmen wir

$$u_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad u_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \text{bzw. } \; v_1 = \ldots, \quad v_2 = \ldots\,.$$

Das ergibt die Basis $\{u_k \otimes v_l;\ k, l = 1, 2\}$ von $\mathcal{H}$,

$$u_1 \otimes v_1 = \left(\begin{array}{c} 1 \cdot v_1 \\ \hline 0 \cdot v_1 \end{array}\right) = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

usw.
Es ist bequem, diese „Blockschreibweise“ zu benutzen; wenn $u = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}$ ist:

$$u \otimes v = \left(\begin{array}{c} \alpha_1 v \\ \hline \alpha_2 v \end{array}\right).$$

Für $A = (a_{ik})$ können wir $A \otimes B$ in Blockschreibweise wie folgt angeben:

$$A \otimes B = \left(\begin{array}{c|c} a_{11}B & a_{12}B \\ \hline a_{21}B & a_{22}B \end{array}\right).$$

Bei Anwendung auf $u \otimes v = \left(\begin{array}{c} \alpha_1 v \\ \hline \alpha_2 v \end{array}\right)$ sieht man, daß B auf v operiert und A auf u. Der allgemeinste Operator hat die Form

$$\left(\begin{array}{c|c} K & L \\ \hline M & N \end{array}\right),$$

wobei $K, \ldots\ N$ 2×2-Matrizen sind; er ist im allgemeinen nicht mehr leicht als Summe von Tensorprodukten zu schreiben.

Dieses Beispiel ist auch für die Physik interessant, wie wir in Kürze sehen werden.

Wir haben gesehen, daß $\mathbb{R}^{2n} = \mathbb{R}^n \times \mathbb{R}^n$ mit der üblichen linearen Vervollständigung und dem üblichen Skalarprodukt in unserer Notation eigentlich als $\mathbb{R}^{2n} \simeq \mathbb{R}^n \oplus \mathbb{R}^n$ zu schreiben ist. Wenn man nun zu *Funktionen* über diesen Räumen übergeht, wird aus der Summe ein Produkt:

Satz 4.10.13

$$\mathcal{L}^2(\mathbb{R}^{2n}) \simeq \mathcal{L}^2(\mathbb{R}^n) \otimes \mathcal{L}^2(\mathbb{R}^n)\,.$$

Ich gebe nur einen *Plausibilitätsbeweis* für $n = 1$: Zum Beweis der Separabilität von $\mathcal{L}^2(\mathbb{R}^n)$ kann man heranziehen, daß Stufenfunktionen mit rationalen Werten, die konstant über Quadern mit rationalen Eckpunkten sind, eine dichte Menge bilden (vergl. Satz 1.2.13). Die $\mathcal{L}^2(\mathbb{R})$-Stufenfunktionen können wir zusammensetzen bzw. approximieren durch die charakteristischen Funktionen χ_I aller Intervalle I. Für das Rechteck $R = I_1 \times I_2$ ist $\chi_R(x_1, x_2) = \chi_{I_1}(x_1)\chi_{I_2}(x_2)$; und diesem Vektor $\in \mathcal{L}^2(\mathbb{R}^2)$ ordnen wir den Vektor $\chi_{I_1} \otimes \chi_{I_2} \in \mathcal{L}^2(\mathbb{R}) \otimes \mathcal{L}^2(\mathbb{R})$ zu. Die Abbildung, die der Funktion $(x_1, x_2) \mapsto f(x_1)g(x_2)$ den Vektor $f \otimes g$ zuordnet, werde linear fortgesetzt:

$$\sum f_i(x_1)g_i(x_2) \mapsto \sum f_i \otimes g_i\,.$$

Die Abbildung ist eindeutig umkehrbar denn die für die rechte Seite geforderten Äquivalenzrelationen (die Gleichungen vor (10.5)) gelten trivialerweise auch für die linke Seite. Sie ist isometrisch:

$$\begin{aligned}
\left\|\sum_i f_i(x_1)g_i(x_2)\right\| &= \sum_{ik} \int \overline{f_i(x_1)g_i(x_2)} f_k(x_1)g_k(x_2)\,dx_1dx_2 \\
&= \sum_{ik} (f_i, f_k)\,(g_i, g_k) = \sum_{ik} (f_i \otimes g_i, f_k \otimes g_k) \\
&= \left\|\sum_i f_i \otimes g_i\right\|^2 .
\end{aligned}$$

Damit haben wir dichte Mengen isomorph aufeinander abgebildet, wenn wir die Abbildung linear auf die lineare Hülle fortsetzen; eine isomorphe Abbildung ist natürlich stetig und läßt sich auf den Abschluß fortsetzen. ■

Aus dem Beweis von Satz 4.10.13 ergibt sich, daß man $\mathcal{L}^2$-Funktionen $f(x_1, x_2)$ durch Produktfunktionen approximieren kann! Er ist in der Quantenmechanik interessant, weil man oft versucht, durch Produktansätze weiterzukommen. Nehmen wir z. B. ein Helium-Atom: es hat zwei Elektronen, die man durch eine Wellenfunktion $\psi(\vec{x}, \vec{y}) \in \mathcal{L}^2(\mathbb{R}^6)$ beschreiben muß. Das Problem ist nicht streng lösbar, man muß nach Näherungsmethoden suchen; und dazu braucht man Ansätze, um die wahre Wellenfunktion zu approximieren. Wenn man nur die Wechselwirkung der Elektronen mit dem Kern berücksichtigt und die Wechselwirkung der Elektronen untereinander vernachlässigt, kann

man das so vereinfachte Problem mit einem Produktansatz lösen: $\psi(\vec{x},\vec{y}) = \psi_1(\vec{x})\psi_2(\vec{y})$. Satz 4.10.13 sagt nun, daß man die wahre Wellenfunktion im allgemeinen zwar nicht als Produkt schreiben, aber nach solchen Produkten entwickeln kann; und das liefert eine mögliche Näherungsmethode.
Die Wechselwirkung der Elektronen untereinander wird gegeben durch das Potential

$$V(\vec{x},\vec{y}) = \frac{e^2}{|\vec{x}-\vec{y}|}\,;$$

das ist ein Multiplikationsoperator in $\mathcal{L}^2(\mathbb{R}^6) \simeq \mathcal{L}^2(\mathbb{R}^3) \otimes \mathcal{L}^2(\mathbb{R}^3)$, der sich nicht als Produkt zweier Operatoren schreiben läßt.

Man braucht das Tensorprodukt von Hilberträumen auch, wenn man den $\mathcal{L}^2(\mathbb{R}^3)$ erweitern will, um Teilchen mit Spin (oder anderen inneren Freiheitsgraden, z. B. „Farbe“ und „Geschmack“ bei den Quarks) zu beschreiben.
Der Stern-Gerlach-Versuch, bei dem ein Strahl (neutraler) Silberatome durch ein stark inhomogenes Magnetfeld in zwei Teilstrahlen aufgespalten wird, hat gezeigt, daß man einem Elektron (und in den Silberatomen ist für diesen Versuch nur das Valenzelektron relevant) ein magnetisches Moment zuordnen muß, das man sich an einen „inneren Drehimpuls“, den Spin, gekoppelt denkt, genau wie dem Bahndrehimpuls des Elektrons ein magnetisches Moment des Stroms entspricht, der durch das umlaufende Elektron gebildet wird. Ich habe betont, daß das wesentliche Charakteristikum eines Drehimpulses die Vertauschungsrelationen (9.8) sind (d. h. die Darstellung der Lie-Algebra der Drehgruppe). Man beschreibt den Spin, den man als inneren Drehimpuls betrachtet, konsequenterweise durch Operatoren $S_i,\ i = 1,2,3$, die diese Relationen erfüllen. Frage: in welchem Hilbertraum? Die Tatsache, daß man im Stern-Gerlach-Versuch *zwei* Teilstrahlen sieht, interpretiert man so, daß der Elektronenspin zwei mögliche scharfe Werte in Magnetfeldrichtung haben kann. Um einen Operator mit zwei verschiedenen Eigenwerten zu beschreiben (hier: Spin in Magnetfeldrichtung $S_\parallel = \vec{S}\cdot\vec{B}/\left|\vec{B}\right|$), genügt ein zweidimensionaler Hilbertraum $\mathcal{H}_2 = \mathbb{C}^2$, aufgespannt durch zwei Vektoren u_+ und u_-, die den Zuständen „Spin nach oben“, (d. h. parallel zum Magnetfeld) und „Spin nach unten“ entsprechen. Man kann tatsächlich eine Darstellung der dreidimensionalen Drehgruppe im zweidimensionalen (komplexen) Raum finden: man setze $S_i = \frac{\hbar}{2}\sigma_i$, wobei

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},\quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix},\quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

die „Paulischen Spinmatizen“ sind. Es ist nicht schwer, die Relationen (9.8) zu überprüfen. Die Eigenwerte der S_i sind leicht zu berechnen: alle drei Operatoren — und auch $S_\parallel$ — haben die Eigenwertw $\pm\hbar/2$; d. h. „das Elektron hat den Spin $\frac{1}{2}$“.

Man will natürlich auch die räumlichen Freiheitsgrade des Elektrons beschreiben; dazu hat man $\mathcal{H}_1 = \mathcal{L}^2(\mathbb{R}^3)$. Zusammengenommen braucht man den Produktraum $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2 = \mathcal{L}^2(\mathbb{R}^3) \otimes \mathbb{C}^2$. Er wird aufgespannt durch die Vektoren $\{\psi(\vec{x}) \otimes u_+,\ \psi(\vec{x}) \otimes u_-\}$.

Wenn wir

$$u_+ = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad u_- = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

schreiben, hat ein allgemeiner Vektor $\varphi \in \mathcal{H}$ die Form

$$\varphi = \psi_+(\vec{x}) \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \psi_-(\vec{x}) \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

(mögliche Koeffizienten sind in den $\psi_\pm$ absorbiert), oder wenn ich die Symbole „$\otimes$" unterdrücke:

$$\varphi = \begin{pmatrix} \psi_+(\vec{x}) \\ \psi_-(\vec{x}) \end{pmatrix},$$

d. h. wir beschreiben das Elektron mit Spin durch eine zweikomponentige Wellenfunktion.

Die Operatoren zur Beschreibung des Spins wirken nur im zweiten Faktor des Hilbertraumprodukts, haben also die Form $\mathbb{1} \otimes S$, diejenigen zur Beschreibung von Ort oder Impuls wirken nur im ersten Faktor: $Q \otimes \mathbb{1}$, $P \otimes \mathbb{1}$; doch um die Aufspaltung des Strahls zu beschreiben, braucht man eine Wechselwirkung des (räumlichen) Magnetfelds mit dem Spin, d. h. einen Operator, der nicht mehr nur in einem Faktor wirkt. (Man benötigt in diesem Fall einen Operator der Form $\sum_{i=1}^{3} A_i \otimes B_i$.)

Im oben erwähnten Heliumatom hat man zwei Elektronen mit Spin $\frac{1}{2}$, man braucht hier als Hilbertraum $\mathcal{H} = \mathcal{L}^2(\mathbb{R}^6) \otimes \mathbb{C}^4 \simeq \mathcal{L}^2(\mathbb{R}^3) \otimes \mathbb{C}^2 \otimes \mathcal{L}^2(\mathbb{R}^3) \otimes \mathbb{C}^2$. Für die Beschreibung der Spinoperatoren ist Beispiel 4.10.12 nützlich.

Das letzte Thema, das ich ohne große Genauigkeit besprechen will, betrifft eine weitere Form des Hilbertraums, die in der Physik benötigt wird, den „Fockraum". In der Quantenmechanik behandelt man meist Systeme mit fest gegebener Zahl $1, 2, \dots\ n$ von Teilchen. In der relativistischen Quantentheorie ist wegen der berühmten Einsteinschen Relation $E = mc^2$ bekanntlich Masse und Teilchenzahl nicht notwendig erhalten; man braucht eine Möglichkeit, beliebig viele Teilchen zu beschreiben. Dasselbe trifft zu, wenn man — im Rahmen der nicht-relativistischen Physik — Statistische Mechanik treibt: ob ein Wassertropfen oder ein kleiner Kristall 10^{20} oder $10^{20} + 5$ Teilchen enthalten, ist experimentell nicht zu unterscheiden; man möchte sich nicht auf eine feste Teilchenzahl festlegen, sondern mit der mittleren Teilchenzahl arbeiten.
Ansatzweise haben wir die Struktur, auf die man geführt wird, schon beim harmonischen Oszillator kennengelernt, dessen Eigenwerte wir berechnet haben. Wir haben die Leiteroperatoren $a = \frac{1}{\sqrt{2}}(Q+iP)$ und $a^+ := \frac{1}{\sqrt{2}}(Q-iP)$ definiert, siehe Gleichung (5.2). Wie erwähnt, kann man sie entsprechend den Werten von Masse und Federkonstante so modifizieren daß sie die Relation (5.3): $[a, a^+] = \mathbb{1}$ erfüllen und $H = \hbar\omega(a^+a + \frac{1}{2}\mathbb{1})$ mit den Eigenwerten $\hbar\omega(n+\frac{1}{2})$, $n = 0, 1, 2, \dots$. Man kann das folgendermaßen beschreiben: die Zustände sind charakterisiert durch die Anzahl der „Energiequanten" $\hbar\omega$, die das

System enthält (die Grundzustandsenergie nachzuweisen ist möglich, aber schwierig). $N = a^+a$ zählt diese Anzahl, es ist der „Teilchenzahloperator". Die Eigenvektoren φ_0 (des Grundzustandes) und $(a^+)^n\varphi_0$ spannen den Hilbertraum auf; er hat die Struktur

$$\mathcal{H} = \bigoplus_{n=0}^{\infty} \mathcal{H}_n, \qquad \mathcal{H}_n = \left\{c(a^+)^n\varphi_0;\ c \in \mathbb{C}\right\}. \tag{10.14}$$

Das kann man verallgemeinern. Man tut es z. B. , wenn man einen Strahlungshohlraum behandelt, um das Plancksche Gesetz der schwarzen Strahlung abzuleiten. Der Hohlraum hat gewisse Eigenschwingungsmoden für die elektromagnetische Strahlung, gegeben durch Zustände ψ_k mit Energien $\varepsilon_k,\ k = 1, 2, \ldots$, er kann verschiedene „Photonen" enthalten. Man stellt sich nun vor, daß eine gewisse Anzahl n_k von Photonen der Sorte k mit Energie $\varepsilon_k = \hbar\omega_k$ vorhanden sind und fragt nach der Gleichgewichtsverteilung; die Zahl der Photonen ist natürlich nicht notwendig konstant, sie können von den Wänden absorbiert und emmitiert werden, man kennt nur die mittlere Energie. Doch will ich darauf nicht weiter eingehen, mir kommt es auf die Analogie mit dem harmonischen Oszillator an: Für jedes feste k kann ich die Zahl der Photonen der Zahl der Energiequanten des Oszillators gleichsetzen und kann Leiteroperatoren $a_k^+,\ a_k$ definieren, die ich jetzt besser „Erzeugungs- und Vernichtungsoperatoren" nenne, die mir aus dem Grundzustand Ω, dem Vakuum, das kein Photon enthält, Zustände mit n_k Photonen der Art k erzeugen: $(a^+)^{n_k}\Omega$. Diese Operatoren erfüllen die Vertauschungsrelationen

$$[a_k, a_l^+] = \delta_{kl}\mathbb{1}, \qquad [a_k, a_l] = [a_k^+, a_l^+] = 0\,.$$

(Das entspricht für $k = l$ den früheren Relationen (5.3), die Zahl der Photonen der Art k soll gleichzeitig meßbar sein mit der Zahlder l-Photonen, daher müssen die entsprechenden Operatoren für $k \neq l$ miteinander vertauschen.)

$$\psi_{n_1 \ldots n_r} = \prod_i (a_k^+)^{n_i}\Omega$$

ist dann ein Zustand mit n_i Photonen vom Typ $i,\ i = 1, \ldots r$.
Die a_k's sind die Vernichtungsoperatoren, Anwendung auf einen Zustand mit n_k Photonen vermindert diese Anzahl um 1. Bei Anwendung auf das Vakuum ist nichts zu Vernichtendes da: nach Definition ist $a_k\Omega = 0$. Das entspricht unserer früheren Gleichung $a\varphi_0 = 0$ für den Grundzustand des harmonischen Oszillators.

$\mathcal{H}_0 = \{c\Omega;\ c \in \mathbb{C}\}$ ist der vom Vakuum aufgespannte Unterraum,

$$\mathcal{H}_1 = \overline{\text{L.H.}\,\{\psi_k;\ k = 1, 2, \ldots\}} = \overline{\text{L.H.}\left\{a_k^+\Omega;\ k = 1, 2, \ldots\right\}}$$

ist der Raum der Ein-Photon-Zustände und

$$\mathcal{H}_n = \overline{\text{L.H.}\left\{\prod_{k:\sum n_k = n} (a_k^+)^{n_k}\Omega\right\}}$$

ist der n-Photonen-Raum. In

$$\mathcal{H} = \bigoplus_{n=0}^{\infty} \mathcal{H}_n$$

können wir dann Zustände mit beliebig vielen Photonen beschreiben.

Man kann sich überzeugen, daß

$$\mathcal{H}_n \simeq \mathrm{S} \bigotimes_{i=1}^{n} \mathcal{H}_1^{(i)} ,$$

wobei die $\mathcal{H}_1^{(i)}$ Kopien des Ein-Photon-Raumes sind: $\mathcal{H}_1^{(i)} \simeq \mathcal{H}_1$, vorausgesetzt, man identifiziert $\psi_k \otimes \psi_l$ und $\psi_l \otimes \psi_k$ miteinander; das soll durch das Symbol „S" angedeutet werden. Für das Produkt der Kopien schreibt man kurz $\mathcal{H}_1^{\otimes n}$.

Diese Konstruktion ist nicht auf Photonen beschränkt. Man kann von einem beliebigen Ein-Teilchen-Raum $\mathcal{H}_1$ ausgehen und damit den zugehörigen „Fockraum" definieren:

$$\mathcal{H}_F = \bigoplus_{n=0}^{\infty} \mathcal{H}_1^{\otimes n} ,$$

wobei $\mathcal{H}_0 = \{c\Omega;\ c \in \mathbb{C}\}$ der eindimensionale Vakuum-Unterraum ist. In diesem Raum kann man Operatoren $a^+(f),\ a(f)$ definieren, die ein Teilchen mit Wellenfunktion f erzeugen oder vernichten.

Ich habe hier die schon früher erwähnte Komplikation unberücksichtigt gelassen, die durch die Ununterscheidbarkeit identischer quantenmechanischer Teilchen entsteht. Im Photonenbeispiel ist sie durch das Symbol „S" berücksichtigt, Analoges muß man auch im Teilchenfall machen, wobei man aber zwischen Bosonen und Fermionen (Teilchen mit ganzzahligem oder halbzahligem Spin) unterscheiden muß. Doch das führt weit über den Rahmen dieser Vorlesung hinaus.

Anhang A

Ein einfaches lösbares Wachstumsmodell

Wir betrachten eine einzelne Spezies in einem gegebenen Gebiet, zum Beispiel die Bäume eines (Monokultur-)Waldes, und fragen nach der zeitlichen Entwicklung, die durch Wachstum, Reproduktion und Absterben bestimmt wird. Unser Modell ist unrealistisch insofern, als es „linear“ ist, die Endlichkeit der Nahrungsresourcen, die Konkurrenz innerhalb der Art und mit anderen Arten wird nicht berücksichtigt. Dank der Linearität ist es vollständig lösbar; und die Analyse eines verbesserten Modells mit Nichtlinearitäten, das nicht mehr streng lösbar ist, basiert auf den Ergebnissen des linearen Modells; doch das soll hier nicht mehr ausgeführt werden.

A.1 Beschreibung des Modells

Der Einfachheit halber rede ich von „Wald“ und „Bäumen“, obwohl man das Modell auch anders interpretieren kann, Sie können auch an die Entwicklung eines Fond-Vermögens denken, das durch Zinsen wächst, sich durch Erwerb neuer Papiere „reproduziert“ und durch Dividendenausschüttung teilweise „abstirbt“.

Uns interessiert die Anzahl $n(x,t)$ von Bäumen des Alters oder der Größe x zur Zeit t. Ohne Reproduktion und Absterben gilt, wenn man x als Alter interpretiert:

$$n(x+\Delta t, t+\Delta t) = n(x,t)\,,$$

die Anzahl der zur Zeit t gerade x Jahre alten ist nach Δt Jahren gleich der Anzahl der $x+\Delta t$ alten Exemplare. Anders:

$$n(x, t+\Delta t) - n(x,t) = n(x-\Delta t, t) - n(x,t)\,;$$

Division durch Δt ergibt im Limes $\Delta t \to 0$

$$\partial_t n = -\partial_x n \,.$$

Wenn man x als „Größe“ oder Biomasse pro Baum interpretiert, muß man $x - \Delta t$ durch $x - \omega(x)\Delta t$ ersetzen, wo ω die von der Größe abhängige Wachstumsgeschwindigkeit ist; die rechte Seite der obigen Gleichung wird dann

$$\frac{\omega}{\omega\Delta t}(n(x-\omega\Delta t, t) - n(x,t)) \longrightarrow -\omega(x)\,\partial_x n \,.$$

Berücksichtigung des Absterbens: üblicherweise setzt man die Zahl der absterbenden Exemplare einer gegebenen Größe proportional der Zahl der vorhandenen dieser Größe:

$$\partial_t n(x,t) = -\omega(x)\partial_x n(x,t) - \mu(x)n(x,t)\,;$$

hier ist $\mu(x)$ die Sterberate.

Zur Reproduktion tragen alle Bäume mit der Größe entsprechender Samenrate bei; die Sämlingszahl ist also proportional zu $\int \sigma(x)n(x,t)$. σ bezeichne eine effektive Samenrate, die die Anzahl der *überlebenden* Sämlinge ergibt. Wir wollen in langen Zeiträumen rechnen — Bäume werden sehr alt — und dementsprechend die jahreszeitlichen Schwankungen ausgemittelt denken, also eine kontinuierliche Reproduktion annehmen. Die Sämlinge seien nicht alle gleich groß, es gibt also eine sehr schmale Verteilung $s(x)$, die nahe $x = 0$ konzentriert ist, die wir aber als stetig annehmen wollen, um nicht Differentialgleichungen mit unstetigen Funktionen zu haben. Unsere Ausgangsgleichung ist also

$$\partial_t n(x,t) = -\omega(x)\partial_x n(x,t) - \mu(x)n(x,t) + s(x)\int \sigma(x')n(x',t)\,dx'\,;$$

oder anders geschrieben:

$$\begin{aligned}
\partial_t n(x,t) &= (Dn)(x,t) + s(x)I_\sigma(n)\,, & (1.1)\\
(Dn)(x,t) &:= (-\omega(x)\partial_x - \mu(x))n(x,t)\,, & (1.2)\\
I_\sigma(n) &:= \int \sigma(x)n(x,t)\,dx\,. & (1.3)
\end{aligned}$$

Die Größen ω, μ, σ und s sind als bekannte Funktionen zu betrachten, über die wir folgende Voraussetzungen machen:

$$0 < \omega(x) < \omega_{\max} < \infty \text{ for } x \in J = [0, x_m), \text{ wobei } x_m = \infty \text{ zugelassen ist}\,. \qquad (1.4)$$

Ein Sämling beginnt sofort mit dem Wachsen, daher die erste Ungleichung; und sehr alte Bäume wachsen schließlich kaum noch, daher die zweite. Endliches x_m bedeutet, daß es ein Höchstalter für die Bäume gibt, aber das muß mathematisch nicht sein.

Die Sterblichkeitsrate ist ebenfalls nicht-negativ, sie soll für Bäume ab einer gewissen Größe strikt positiv sein:

$$\mu(x) \geq 0;\ \mu(x) > 0 \text{ für } x > x_0\,. \tag{1.5}$$

Für die Samenproduktion nehmen wir auch an, daß sie bei sehr alten Bäumen wieder zurückgeht, bei sehr jungen kann sie $= 0$ sein, doch es gibt samenproduzierende Exemplare:

$$0 \leq \sigma(x) \leq \sigma_{\max} < \infty;\ \operatorname{supp}\sigma \neq \emptyset\,. \tag{1.6}$$

Die Sämlingsverteilung liegt nahe bei $x = 0$, aus technischen Gründen brauchen wir Differenzierbarkeit:

$$s(x) \geq 0, \quad \operatorname{supp} s \subset \mathbb{R}^+ \text{ kompakt},\ s \text{ differenzierbar.} \tag{1.7}$$

Schließlich nehmen wir an, daß alle diese Funktionen im Intervall J stetig sind.

Um ein Gefühl für die möglichen Lösungen zu bekommen, vergessen wir einmal den Sterbe- und den Samenterm und nehmen ω als konstant an: $\partial_t n = -\omega\partial_x n = -\frac{\partial}{\partial(\omega^{-1}x)}n$, mit der Variablen $y = \omega^{-1}x$ geschrieben: $\partial_t n(y,t) = -\partial_y n(y,t)$, und die Lösung ist einfach: $n = f(y-t)$ mit einer beliebigen differenzierbaren Funktion f; umgeschrieben in x: $n(x,t) = f(\omega^{-1}x - t) \equiv g(x - \omega t)$, d. h. die Anfangsverteilung $n(x,0) = g(x)$ verschiebt sich mit der Geschwindigkeit ω. Wenn $\omega(x)$ nicht konstant ist, müssen wir statt $y = \omega^{-1}\int^x dx'$ die Größe $\gamma(x) := \int^x \omega(x')^{-1}\,dx'$ nehmen: $n(x,t) = f(\gamma(x) - t)$, d. h. eine Verschiebung der Anfangsverteilung mit einer zusätzlichen Verzerrung durch die unterschiedlichen Geschwindigkeiten für unterschiedliche x.
Vergessen wir alle Terme außer dem Sterbeterm: $\partial_t n = -\mu(x)n$, dann zerfällt die Verteilung exponentiell: $n(x,t) = n(0,x)\exp(-\mu(x)t)$. Zusammengenommen haben wir also eine Verschiebung der Anfangsverteilung mit Verzerrung, dabei ein exponentielles Abnehmen, doch durch den Samenterm werden kontinuierlich neue schmale „Anfangsverteilungen“ nachgeliefert, die das gleiche Schicksal haben.

A.2 Der „freie“ Entwicklungsoperator

Um die Lösung unserer Gleichung zu finden, benötigen wir zunächst die Lösung der Gleichung ohne Samenterm:

$$\partial_t n = Dn\,, \quad D \equiv -\omega(x)\partial_x - \mu(x)\,. \tag{2.1}$$

Formal können wir die Lösung zu gegebener Anfangsverteilung $n_0(x)$ sofort schreiben als

$$n(x,t) = (L_t n_0)(x)\,, \quad L_t = e^{tD} \tag{2.2}$$

Die Exponentialfunktion kann man durch die Exponentialreihe zu definieren suchen, man muß kontrollieren, auf welche Funktionen sie angewandt werden darf, so daß man eine konvergente Reihe erhält. Wir gehen anders vor, wir wollen den Zeitentwicklungsoperator L_t hier explizit beschreiben: Wir definieren

$$\gamma(x) := \int_0^x \omega(x')^{-1}\, dx', \quad \lambda(x) := \int_0^x \omega(x')^{-1}\mu(x')\, dx', \quad \text{für } x \in J. \tag{2.3}$$

γ und λ sind monoton wachsende Funktionen, denn μ ist nicht-negativ und ω ist > 0, die Integrale sind wohldefiniert. γ ist sogar strikt monoton auf J, d. h. die Umkehrfunktion γ^{-1} existiert.

Sei f eine differenzierbare Funktion auf $\mathbb{R}$, dann ist

$$n(x,t) = e^{-\lambda(x)} f(\gamma(x) - t) \tag{2.4}$$

eine Lösung von Gleichung (2.1). Das kann man durch Differenzieren ohne Schwierigkeit verifizieren. Die zugehörige Anfangsverteilung ist $n_0(x) = e^{-\lambda(x)} f(\gamma(x))$; zu gegebenem n_0 müssen wir also $f_{n_0} = (e^{\lambda} n_0) \circ \gamma^{-1}$ wählen.

Wir definieren nun $x_t := \gamma^{-1}(\gamma(x) - t)$ für $x \geq \gamma^{-1}(t)$, und $x_t := 0$ sonst, dann können wir also die freie Zeitentwicklung durch

$$\begin{aligned} L_t : n_0(x) \mapsto (L_t n_0)(x) &= e^{-\lambda(x)} f_{n_0}(\gamma(x) - t) \\ &= e^{-\lambda(x)+\lambda(x_t)}\, n_0(x_t) \end{aligned} \tag{2.5}$$

explizit beschreiben mithilfe bekannter Funktionen und ihrer Integrale. Zu gegebener Anfangsverteilung gibt es eine eindeutige Lösung; L_t ist ein linearer Operator, weil die Gleichung (2.1) linear ist.

γ und daher auch γ^{-1} sind strikt monoton wachsend, deshalb gilt

$$x_t = \gamma^{-1}(\gamma(x) - t) < \gamma^{-1}(\gamma(x)) = x \text{ wenn } t > 0\,;$$

und weil auch λ strikt monoton wächst, sobald x hinreichend groß ist, so daß $\mu(x) \neq 0$, ist der Exponent in (2.5) negativ und abnehmend. (Wenn ω und μ konstant sind, lautet der Exponent einfach $-\mu t$.) Wie nicht anders zu erwarten, stirbt der Wald aus, wenn wir keine Samenproduktion haben:

$$(L_t n_0)(x) \leq n_0(x_t) \longrightarrow 0 \text{ wenn } t \to \infty\,.$$

(x_t wird $= 0$ für hinreichend große t, und für $x \leq 0$ müssen wir $n_0 = 0$ setzen.)

A.2.6 Anmerkung: Damit $L_t n_0$ eine differenzierbare Funktion ist, müssen wir voraussetzen, daß n_0 bei $x = 0$ nebst seiner ersten (rechtsseitigen) Ableitung verschwindet; sonst wäre die Lösung bei $x_0 = \gamma^{-1}(t) > 0$ nicht differenzierbar.

A.3 Allgemeine Lösungen

Wir betrachten jetzt die volle Gleichung

$$\partial_t n = Dn + sI_\sigma(n)\,.$$

Erinnern wir uns an die qualitative Beschreibung des erwarteten Verhaltens der Lösung: Verschiebung mit Verzerrung bei gleichzeitiger exponentieller Dämpfung und ständiges Nachliefern schmaler Verteilungen mit demselben Schicksal. Zu einer Zeit t finden wir eine Superposition aller Verteilungen, die zu Zeiten $t' < t$ begannen und sich mit L_t entwickelten. Das suggeriert den folgenden Ansatz:

$$n = L_t n_0 + \int_0^t k(t') L_{t-t'}\, s\, dt'\,, \tag{3.1}$$

wobei $k(t)$ ein noch zu bestimmender Gewichtsfaktor ist. Wir wollen (3.1) nach t differenzieren. Dazu beachte man, daß $\partial_t L_t n_0 = DL_t n_0$ ist, denn $L_t n_0$ löst ja die freie Gleichung, und daß D ein auf die Variable x wirkender Differentialoperator ist, den wir mit der t'-Integration vertauschen können:

$$\begin{aligned} \partial_t n &= D[L_t n_0] + k(t)s + \int_0^t k(t') D[L_{t-t'} s]\, dt' \\ &= Dn + k(t)s\,. \end{aligned}$$

Wir bekommen also eine Lösung unserer Gleichung, wenn wir $k(t)$ so wählen, daß

$$k(t) = I_\sigma(n(\,\cdot\,, t))\,. \tag{3.2}$$

n hängt selbst von $k(t)$ ab, wir müssen also ein Selbstkonsistenzproblem lösen:

$$\begin{aligned} k(t) &= I_\sigma(L_t n_0) + \int_0^t dt'\, k(t') \int dx\, \sigma(x)(L_{t-t'} s)(x) \\ &\equiv l_0(t) + \int_0^t dt'\, k(t') l(t-t') \qquad (3.3) \\ \text{mit}\ \ l_0(t) &:= I_\sigma(L_t n_0)\,, \\ l(t) &:= \int dx\, \sigma(x)(L_t s)(x) = I_\sigma(L_t s)\,. \end{aligned}$$

Die hier auftretenden Funktionen l_0 und l sind im Prinzip bekannt, weil wir n_0, s und L_t kennen. Definieren wir $f * g = \int_0^t f(t') g(t-t')\, dt'$, dann müssen wir also die Konvolutionsgleichung

$$k = l_0 + k * l\,. \tag{3.4}$$

lösen. Das ist formal leicht möglich: ein Physiker, der diese Gleichung sieht, wird versuchen, durch Iterieren weiterzukommen: das k auf der rechten Seite kann ja gemäß der Gleichung durch die rechte Seite ersetzt werden: $k = l_0 + (l_0 + k * l) * l = l_0 + l_0 * l +$

$k * l * l$; und diese Prozedur läßt sich fortsetzen, so daß wir schließlich eine unendliche Reihe

$$k = l_0 + l_0 * l + l_0 * l * l + \ldots \tag{3.5}$$

bekommen, die man — hoffentlich! — durch wenige Anfangsglieder approximieren kann. Aber hier müssen wir mathematisch sauber arbeiten, und wir haben das Werkzeug dazu: Die $\mathcal{L}^1$-Norm. Wenn wir damit dieser Reihe einen guten Sinn geben können, haben wir tatsächlich eine Lösung gefunden.

Man beachte, daß – wie man leicht kontrolliert — $\|f * g\|_1 \leq \|f\|_1 \|g\|_1$ gilt. Also ist $\|l_0 * l^{*n}\|_1 \leq \|l_0\|_1 \|l\|_1^n$, und die Reihe konvergiert im Sinne der $\mathcal{L}^1$-Norm, wenn $\|l\|_1 < 1$ ist. Das ist aber immer dann der Fall, wenn die Samenproduktion hinreichend klein ist: $\|l\|_1 = \|I_\sigma(L_t s)\|_1$. Um genau zu sein: man kann zeigen, daß $I_\sigma(L_t n_0) \in \mathcal{L}^1$ für eine vernünftige Anfangsbedingung ($n_0(0) = 0,\ n_0'(0) = 0,\ I_\sigma(n_0) < \infty$), insbesondere für $n_0 = s$. Damit ist $\|I_\sigma(L_t s)\|_1 < \infty$; und durch Wahl eines hinreichend kleinen Vorfaktors für σ kann man erreichen, daß es < 1 ist. Dann ist

$$k = l_0 + l_0 * l_1 \quad \text{mit } l_1 = \sum_{n=1}^{\infty} l^{*n} \in \mathcal{L}^1 \quad \text{und} \quad k \in \mathcal{L}^1\,.$$

Einsetzen in den Ansatz (3.1) zeigt dann, daß die Verteilung ausstirbt, weil $(L_t s)(x) \to 0$ für alle x (man benutze den Mittelwertsatz der Integralrechnung). Folglich gilt

Satz A.3.6 *Die Verteilung $n(x,t)$ stirbt aus für $t \to \infty$, wenn die Samenproduktion unter einer kritischen Schwelle liegt, die durch $\|I_\sigma(L_t s)\|_1 = 1$ gegeben ist.*

Man kann nach stationären Lösungen unserer Gleichung fragen: man setze $\partial_t n = 0$. Sie existieren *für hinreichend kleine Samenproduktion(!)*, und das hat mich zunächst verwirrt, als ich anfing, mich mit diesem Modell zu befassen, denn das scheint ein eklatanter Widerspruch zu dem obigen Satz zu sein. Die Lösung des Rätsels ist, daß für die stationären Lösungen im allgemeinen $n_0(0) \neq 0$ gilt; und damit darf man sie nicht als Anfangsbedingungen für die allgemeine Lösung nehmen, vergl. Anmerkung A.2.6. Für die stationären Lösungen gibt es eine *obere* Grenze der Samenproduktion, die gerade mit der *unteren* Grenze aus dem Satz übereinstimmt; für den Grenzwert gilt $n(0) = 0$, und die Verteilung überlebt (aber jede noch so kleine Schwankung läßt sie aussterben).

Was macht man, wenn die Samenproduktion höher ist? Wir interessieren uns für positive Zeiten, und dann ist die (einseitige) Laplace-Transformierte das Werkzeug, das sich anbietet: $\widehat{f}(z) := \int_0^\infty f(t)e^{-zt}\,dt$; denn dann gilt $\widehat{f * g} = \widehat{f}\widehat{g}$, (Gleichung (5.9)). Dann tranformiert sich Gleichung (3.4) in

$$\widehat{k} = \widehat{l_0} + \widehat{k}\widehat{l}\,, \tag{3.7}$$

und das ist jetzt eine algebraische Gleichung, die leicht zu lösen ist:

$$\widehat{k} = \frac{\widehat{l_0}}{1 - \widehat{l}}. \tag{3.8}$$

Man muß natürlich auf die Nullstellen des Nenners achten! Wir nehmen wieder an, daß n_0 so gewählt ist, daß $l_0 \in \mathcal{L}^1$ ist. Auch $l(t)$ ist eine positive $\mathcal{L}^1$-Funktion (die Sämlingsverteilung s erfüllt die gleichen Bedingungen wie n_0); folglich ist $\widehat{l}(z)$ analytisch in der rechten Halbebene $\Re z > 0$, stetig für $\Re z \geq 0$, strikt abnehmend mit $\Re z$, also beschränkt durch $\widehat{l}(0)$, reell für reelle Argumente; weiterhin ist $\widehat{l}(0) = \|l\|_1$ (wir setzen $l(t) = 0$ für $t < 0$). Wenn also $\|l\|_1 > 1$ ist, gibt es ein $c_1 \in \mathbb{R}^+$, so daß $\widehat{l}(c_1) = 1$.
$\widehat{l_0}$ hat dieselben schönen Eigenschaften, daher ist $\widehat{k}$ aus Gleichung (3.8) analytisch für $\Re z > c_1$, mehr noch: meromorph für $\Re z > 0$ mit genau einem Pol bei $z = c_1$. Der Pol ist von erster Ordnung: denn $\widehat{l}(x)$ ist eine strikt abnehmende Funktion, daher ist $\widehat{l}'(c_1) \neq 0$, also $1 - \widehat{l}(z) = (z - c_1)\widehat{l}'(c_1) + \ldots$. Damit können wir das Lemma 2.5.4 anwenden; wir erhalten das Ergebnis

Lemma A.3.9 *$k(t)$, definiert als inverse Laplace-Transformierte von $\widehat{k}$ aus Gleichung (3.8) existiert und hat die Form*

$$k(t) = k_1(t)e^{c_1 t},$$

wo k_1 eine nicht-negative Funktion ist, die für $t \to \infty$ schwächer als exponentiell wächst.

Daß die inverse Laplace-Transformierte existiert, folgt aus den schönen Eigenschaften von $\widehat{l_0}$, der Nenner geht im Unendlichen gegen 1, so daß man die Umkehrformel (II.5.7) anwenden kann. $k(t)$ wächst also exponentiell; und Faltung mit $L_t s$ gemäß unserem Ansatz (3.1) ergibt, wie man sich sofort überzeugt, ebenfalls eine exponentiell wachsende Funktion.
Die Lösung mithilfe der Laplace-Transformierten ist natürlich nicht auf den Fall großer Samenproduktion beschränkt. Wenn $\widehat{l}(0) \neq 1$ ist, bleibt der Nenner in (3.8) regulär, dann ist $k(t)$ ebenso wie l_0 eine $\mathcal{L}^1$-Funktion, und die Aussage von Satz A.3.6 gilt.

Wir haben das Modell also vollständig gelöst. Das Ergebnis ist das, was jeder Populationsdynamiker von einem linearen Modell erwartet:
Es gibt einen Schwellenwert für die Samenproduktion, definiert durch

$$\|I_\sigma(L_t s)\|_1 = 1\,,$$

unterhalb dessen die Population ausstirbt; oberhalb des Schwellenwertes explodiert die Population exponentiell.

In einem realistischeren Modell muß man, wenn man nur eine Spezies betrachtet, die Wachstumsgeschwindigkeit, die Sterberate und die Samenproduktion selbst als abhängig von der Populationsdichte ansetzen, d. h. als Funktionale von $n(x,t)$. Damit bekommen wir sofort eine *nicht-lineare* Integro-Differentialgleichung, die — auch bei einfachsten Ansätzen — nicht mehr geschlossen lösbar ist. Man kann, basierend auf den Lösungen für das lineare Modell, noch einige allgemeine Aussagen machen; doch das gehört nicht in diese Vorlesung.

Anhang B

Exkurs über Differentialformen

Eine gute Einführung in die Theorie der Differentialformen finden Sie z. B. in dem — für Mechanik-Hörer auch sonst sehr empfehlenswerten — Buch von V. I. Arnold: Mathematical Methods of Classical Mechanics, Springer 1980.

Vorbemerkung: Alle auftretenden Funktionen werden als ausreichend oft differenzierbar angenommen, alle Kurven, Flächen oder Hyperflächen der Einfachheit halber als glatt, einfach zusammenhängend und doppelpunktfrei.

B.1 1-Formen

$f' = \lim_{\Delta x \to 0} \frac{\Delta f}{\Delta x} = \frac{df}{dx}$ impliziert *nicht*, daß df und dx „infinitesimal klein" sind (Physiker-Argumentation); es folgt nur, daß ihr Quotient gleich dem Limes des Quotienten von Δf und Δx ist. Wenn ich zu gegebenem dx definiere: $df := f'\,dx$, dann kann ich df und dx als endliche Größen ansehen, deren Quotient das Gewünschte ergibt. Diese Differentiale werden als endliche Größen aufgefaßt, bei denen aber immer nur Quotienten interessieren. Beispiel Kettenregel: Sei $x = x(t)$, $dx = x'(t)\,dt$, dann ist

$$df = \frac{df}{dx}\,dx = \frac{df}{dx}\frac{dx}{dt}\,dt\,, \quad \text{also } \frac{df}{dt} = \frac{df}{dx}\,x'\,.$$

Eine bessere — und vor allem verallgemeinerungsfähige — Einführung ist die folgende: Der Wert einer differenzierbaren Funktion f in der Nähe des Punktes x unterscheidet sich von $f(x)$ durch einen Term, der linear in der Differenz der Argumente ist, und einen kleinen Korrekturterm:

$$f(x+h) = f(x) + L_f(x;h) + O(h^2)\,. \tag{1.1}$$

Dabei ist die Abbildung $L_f(x;\cdot) : h \mapsto L_f(x;h)$ linear in h, gegeben durch $L_f(x;h) = f'(x)\,h$. Die Menge dieser linearen Abbildungen $\{L_f(x;\cdot)\,; x \in \text{Definitionsbereich von } f\}$

nennt man kurz df.
„Differentialformen", genauer: „1-Formen" oder Differentialformen erster Stufe sind also Mengen von Linearformen.
Spezielle Beispiele:

$$dx: \ (x, h) \mapsto h; \quad df: \ (x, h) \mapsto f'(x)\, h\,.$$

Das ist eine kompliziert klingende Definition; der Grund dafür ist die Möglichkeit der Verallgemeinerung. Wir können hier entsprechend dem speziellen Beispiel einfach überall dx statt h schreiben, dann bekommen wir wieder $df = f'(x)\, dx$.
Da die Mannigfaltigkeit $\mathbb{R}$, auf der die betrachteten Funktionen definiert sind, eindimensional ist, kann man alle 1-Formen durch eine einzige, z. B. durch dx, ausdrücken, es gibt (an jedem Punkt x) nur eine unabhängige Linearform; die allgemeine 1-Form hat also die Gestalt

$$g(x)\, dx\,. \tag{1.2}$$

Man sieht: 1-Formen lassen sich mit glatten Funktionen multiplizieren; z. B. $h(x)\, df = h(x) f'(x)\, dx$.
Der Quotient $\frac{df}{dx}$ ist nach der komplizierten Definition genaugenommen als $\frac{df(x;h)}{dx(x;h)}$ zu lesen; da man aber — siehe oben — dx mit h identifizieren kann, können wir bei der Kurzschreibweise bleiben.

Über Differentialformen erster Stufe kann man integrieren:

$$\int_I g\, dx\,, \qquad \int_I g\, df\,, \qquad \ldots\,, \qquad I \subset \mathbb{R}\,.$$

Genauer: 1-Formen kann man über eindimensionale Mannigfaltigkeiten integrieren, siehe weiter unten.
Es ist präziser, das, was unter einem Integral steht, als Differentialform aufzufassen, als zu sagen, man „integriere über eine Funktion". (Die letztere Auffassung kann in Extremfällen zu Beweisen verleiten, daß $1 = 0$ ist.)

Verallgemeinerung auf den $\mathbb{R}^n$: Grundlage ist wieder Gleichung (1.1). x und h sind jetzt als n-dimensionale Vektoren zu lesen; und die 1-Form df ist an jedem Punkt $\vec{x}$ wiederum durch eine Linearform gegeben: $\vec{h} \stackrel{\text{linear}}{\longrightarrow} L_f(\vec{x}; \vec{h})$. (Man kann noch weiter verallgemeinern: x kann ein Punkt einer n-dimensionalen Mannigfaltigkeit $\mathcal{M}$ sein, h ein Vektor aus dem Tangentialraum an $\mathcal{M}$ bei x, der isomorph ist zum $\mathbb{R}^n$.)
Hier gibt es nun n linear unabhängige 1-Formen:

$$dx_i: \ (\vec{x}, \vec{h}) \mapsto h_i\,; \quad i = 1, \ldots, n\,.$$

$L_f(\vec{x}; \vec{h})$ gibt die Änderung von f bei $\vec{x}$ in der Richtung von $\vec{h}$ in linearer Näherung an (ich benutze wieder die Notation $\frac{\partial f}{\partial x_i} = \partial_i f$):

$$L_f(\vec{x}; \vec{h}) = \sum_{i=1}^{n} \partial_i f(\vec{x})\, h_i\,;$$

oder kurz:

$$df = \sum_{i=1}^{n} \partial_i f \, dx_i \,. \tag{1.3}$$

Im Folgenden werde ich immer die Kurzform benutzen.
Die allgemeinste 1-Form läßt sich dann schreiben als

$$\omega^1 = a_1(\vec{x})\, dx_1 + a_2(\vec{x})\, dx_2 + \ldots + a_n(\vec{x})\, dx_n \,. \tag{1.4}$$

Analog zum eindimensionalen Fall, bei dem wir die Ableitung von f durch Division der Differentiale gewonnen haben, können wir aus (1.3) die partiellen Ableitungen erhalten: wir setzen $dx_i = h_i = 0$ für $i \neq k$ und dividieren durch dx_k.

Man benutzt hier gern die „Summationskonvention“: über doppelt auftretende Indizes wird summiert: $df = \sum_{i=1}^{n} \partial_i f \, dx_i \equiv \partial_i f \, dx_i$.
Sei $\mathcal{C}$ eine Kurve im $\mathbb{R}^n$, $\mathcal{C} \,:\, x = x(t)$. Die Änderung von f längs der Kurve ist proportional zu $\lim \frac{1}{\Delta t} \Big(f(x(t+\Delta t)) - f(x(t)) \Big) = \partial_i f \, x_i'$. Mit Differentialformen geschrieben: df ist jetzt einzuschränken auf $\mathcal{C}$,

$$dx_i \Big|_{\mathcal{C}} = x_i' \, dt; \qquad df \Big|_{\mathcal{C}} = \partial_i f \, dx_i \Big|_{\mathcal{C}} = \partial_i f \, x_i' \, dt \,.$$

Division durch dt ergibt die wohlbekannte Kettenregel

$$\frac{df(x(t))}{dt} = \frac{\partial f}{\partial x_i} \frac{dx_i}{dt} \,.$$

(In allen diesen Formeln ist über i zu summieren!)

Auch eine allgemeine 1-Form wie in Gleichung (1.4) kann ich über eine eindimensionale Mannigfaltigkeit, d. h. über eine Kurve integrieren:
sei $\mathcal{C} := \{\vec{x}(t) \in \mathbb{R}^n;\; t_0 \leq t \leq t_1\}$, dann ist $dx_i = x_i' \, dt$ auf $\mathcal{C}$, und

$$\int_{\mathcal{C}} \omega^1 = \sum_i \int_{t_0}^{t_1} a_i(\vec{x}(t)) \frac{dx_i}{dt} \, dt \,. \tag{1.5}$$

Was unter dem Integral steht, ist, über i summiert, $= \omega^1 \Big|_{\mathcal{C}}$.
Spezialfall: es gibt ein f, so daß $a_i = \partial_i f$; dann ist $\omega^1 = \partial_i f \, dx_i = df$ und

$$\int_{\mathcal{C}} \omega^1 = \int_{\mathcal{C}} df = f(\vec{x}(t_1)) - f(\vec{x}(t_0)) \,. \tag{1.6}$$

B.2 Höhere Differentialformen

Für die Integration über eine k-dimensionale Mannigfaltigkeit braucht man Differentialformen k-ter Stufe oder k-Formen.

Beginnen wir mit dem $\mathbb{R}^2$: das Flächenelement $dx_1\, dx_2$ ist nach dem, was wir oben gesagt haben, als Bilinearform aufzufassen: $(\vec{h_1}, \vec{h_2}) \mapsto (h_1)_1(h_2)_2$. Man beachte: das ist eine Vorzeichen-behaftete Größe. Entsprechendes gilt für andere 2-Formen. Wir wollen ihre Eigenschaften so festlegen, daß wir den Anschluß an die bekannte Integration über 2-dimensionale Flächen im $\mathbb{R}^n$ finden; insbesondere soll eine Form $dx_1\, dx_1$ verschwinden. Wir könnten das Flächenelement $dx_1\, dx_2$ auffassen als Vektorprodukt $\vec{e}_1 dx_1 \times \vec{e}_2 dx_2$. Das ist ein Vektor parallel zur Normalen, wenn wir den $\mathbb{R}^2$ in den $\mathbb{R}^3$ einbetten, also das korrekte Oberflächenelement mit Orientierung. (Damit erreichten wir natürlich auch $\vec{e}_1 dx_1 \times \vec{e}_1 dx_1 = 0$.) Im $\mathbb{R}^3$: das Spatprodukt $(\vec{a}, \vec{b}, \vec{c})$ dreier Vektoren gibt das orientierte Volumen des durch die Vektoren aufgespannten Parallelepipeds an; das orientierte Volumenelement könnten wir schreiben als $(\vec{e}_1 dx_1, \vec{e}_2 dx_2, \vec{e}_3 dx_3)$; wesentlich ist: es ist, wie das Vektorprodukt, antisymmetrisch gegenüber paarweiser Vertauschung der Argumente. Dementsprechend definieren wir, indem wir die Vektorauffassung fallenlassen und wieder zu den Differentialen in Kurzschreibweise übergehen, das **äußere Produkt** von Differentialen als ein antisymmetrisches Produkt, geschrieben mit dem Keil $\wedge$ (englisch: wedge):

$$dx_1 \wedge dx_2 = -\, dx_2 \wedge dx_1\,, \qquad dx_1 \wedge dx_2 \wedge dx_3\,, \qquad \ldots$$

Diese Produkte können wir auch noch mit Funktionen multiplizieren und addieren. Dadurch erhalten wir Differentialformen zweiter, dritter, … Stufe.
Es ist klar, daß $dx_1 \wedge (dx_2 \wedge dx_3) = +(dx_2 \wedge dx_3) \wedge dx_1$, allgemein: das Vorzeichen bei Vertauschung hängt von der Zahl der Differentiale in den beiden Faktoren ab, anders gesagt: von der Stufe der beiden Formen,

$$\omega^k \wedge \omega^l = (-1)^{kl} \omega^l \wedge \omega^k\,, \tag{2.1}$$

wenn ω^k eine k-Form und ω^l eine l-Form ist.
Funktionen werden auch als 0-Formen bezeichnet, und eigentlich sollten wir nicht $f\, dx_1$ sondern besser $f \wedge dx_1$ schreiben; nach (2.1) ist $f \wedge dx_1 = dx_1 \wedge f$; und daher wird der Keil hier oft weggelassen.
Allgemeine Multiplikation: mit (2.1) und Distributivgesetz und der Regel $f(\vec{x}) \wedge g(\vec{x}) = f(\vec{x})g(\vec{x})$ für Null-Formen; z. B.

$$\begin{aligned} & \Big(\, a(\vec{x}) \wedge dx_1 + b(\vec{x}) \wedge dx_2 \,\Big) \wedge \Big(\, c(\vec{x}) \wedge dx_1 \wedge dx_3 \,\Big) \\ = \;& a(\vec{x}) \wedge c(\vec{x}) \wedge dx_1 \wedge dx_1 \wedge dx_3 + b(\vec{x}) \wedge c(\vec{x}) \wedge dx_2 \wedge dx_1 \wedge dx_3 \\ = \;& -b(\vec{x})c(\vec{x}) \wedge dx_1 \wedge dx_2 \wedge dx_3\,, \end{aligned}$$

weil $dx_1 \wedge dx_2 = -dx_2 \wedge dx_1$ und $dx_1 \wedge dx_1 = -dx_1 \wedge dx_1 = 0$. Das äußere Produkt einer k-Form mit einer l-Form ergibt also eine $(k+l)$-Form.

Im $\mathbb{R}^n$:

0-Formen: $f(\vec{x}),\ g(\vec{x}),\ \ldots$

1-Formen: $\sum a_i(\vec{x}) \wedge dx_i \equiv a_i(\vec{x})\, dx_i$,

2-Formen: $a_{ij}(\vec{x})\, dx_i \wedge dx_j$,

k-Formen: $a_{i_1 \dots i_k}\, dx_{i_1} \wedge \dots \wedge dx_{i_k}$,

n-Formen: $c(\vec{x})\, dx_1 \wedge \dots \wedge dx_n$.

Abgesehen von verschiedenen $c(\vec{x})$ und Vertauschungen der Differentiale, die nur ein anderes Vorzeichen ergeben, gibt es nur eine einzige n-Form im $\mathbb{R}^n$, sie entspricht dem mit c multiplizierten Volumenelement.
Gegeben eine k-Form ω^k und eine k-dimensionale Mannigfaltigkeit $\mathcal{M}$, dann ist $\int_{\mathcal{M}} \omega^k$ ein sinnvoller Ausdruck; man schreibt $\vec{x}$ und dx_i als Funktionen der Parameter, mit denen man die Mannigfaltigkeit beschreibt. Ich mache es am Beispiel einer Fläche $\mathcal{F}$ im $\mathbb{R}^3$ deutlich. $\mathcal{F}$ sei gegeben durch $\vec{x} = \vec{x}(\lambda_1, \lambda_2)$, die Tangentialvektoren $\vec{\tau}^{\,i}$ an die Koordinatenlinien $\lambda_i = \text{const}$ sind $\vec{\tau}^{\,i} = \frac{d\vec{x}}{d\lambda_i}$, das Flächenelement $d\vec{o}$ ist $=$ $\vec{\tau}^{\,1} \times \vec{\tau}^{\,2}\, d\lambda_1 d\lambda_2$. Das ist Ihnen aus den Mathematischen Methoden I bekannt. Betrachten wir nun

$$\int_{\mathcal{F}} \omega^2 \equiv \int_{\mathcal{F}} \Big(a_1(\vec{x})\, dx_2 \wedge dx_3 + a_2(\vec{x})\, dx_3 \wedge dx_1 + a_3(\vec{x})\, dx_1 \wedge dx_2 \Big) \ .$$

(Das ist wegen (2.1) die allgemeinste 2-Form im $\mathbb{R}^3$.) Es ist

$$\begin{aligned} dx_i \Big|_{\mathcal{F}} &= \frac{\partial x_i}{\partial \lambda_1}\, d\lambda_1 + \frac{\partial x_i}{\partial \lambda_2}\, d\lambda_2 \,, \\ dx_1 \wedge dx_2 \Big|_{\mathcal{F}} &= \left(\frac{\partial x_1}{\partial \lambda_1} \frac{\partial x_2}{\partial \lambda_2} - \frac{\partial x_1}{\partial \lambda_2} \frac{\partial x_2}{\partial \lambda_1} \right) d\lambda_1 \wedge d\lambda_2 \\ &= (\vec{\tau}^{\,1} \times \vec{\tau}^{\,2})_3\, d\lambda_1 \wedge d\lambda_2 = do_3 \,, \end{aligned}$$

folglich erhalten wir

$$\int_{\mathcal{F}} \omega^2 = \int_{\mathcal{F}} a_i(\vec{x})\, do_i = \int_{\mathcal{F}} \vec{a} \cdot (\vec{\tau}^{\,1} \times \vec{\tau}^{\,2})\, d\lambda_1 \wedge d\lambda_2 \,,$$

d. h. einen Ausdruck, den man mithilfe der Parametrisierung mit den λ_i berechnen kann. Das Integral $\int \dots\, d\lambda_1 \wedge d\lambda_2$ ist als gewöhnliches Integral $\int \dots\, d\lambda_1 d\lambda_2$ zu lesen. Bekanntlich ist der Wert des Integrals auf der rechten Seite unabhängig von der Parametrisierung von $\mathcal{M}$.

Daß die k-Formen die richtigen Integranden für Integration über k-dimensionale Mannigfaltigkeiten sind, wird auch deutlich aus dem Transformationsverhalten bei einer Koordinatentransformation $x'_i = \xi_i(\vec{x})$:

$$\begin{aligned} dx'_i &= \partial_{x_j} \xi_i\, dx_j \,, \\ dx'_1 \wedge \dots \wedge dx'_n &= \det \left(\frac{\partial x'_i}{\partial x_j} \right) dx_1 \wedge \dots \wedge dx_n \,; \end{aligned}$$

die auftretende Determinante ist die Jacobi-Determinante, die bekanntlich bei Variablentransformation in einem Volumen-Integral auftritt. Entsprechendes gilt für k-dimensionale Integrale.

Ich fasse dementsprechend eine Differentialform ω als eine abstrakte, koordinatenunabhängige Größe auf, die wir je nach Wahl der Koordinaten in der Mannigfaltigkeit durch Funktionen dieser Koordinaten und die Koordinatendifferentiale ausdrücken können.

B.3 Die äußere Ableitung

Die Differentiation einer Funktion $f(x)$ können wir auffassen als Abbildung der 0-Form f auf die 1-Form $df = f'\,dx = f' \wedge dx$. Verallgemeinert: wir definieren eine Abbildung d, die **äußere Ableitung**, die eine k-Form in eine $(k+1)$-Form überführt, durch

$$\begin{aligned} d(\omega^k \wedge \omega^l) &= d\omega^k \wedge \omega^l + (-1)^k \omega^k \wedge d\omega^l \quad \text{(Kettenregel)}, && (3.1)\\ dd &= 0\,. && (3.2) \end{aligned}$$

Zur Illustration:

$$d\omega^1 \equiv d\Big(\, a_i(\vec{x}) \wedge dx_i \,\Big) = da_i \wedge dx_i + a_i \wedge ddx_i = da_i \wedge dx_i = \partial_k a_i \wedge dx_k \wedge dx_i\,. \quad (3.3)$$

Beispiel $\mathbb{R}^3$: Sei $\omega^1 = \vec{f}\cdot d\vec{x}$, und $dx_i \wedge dx_k = do_l$, wenn (ikl) zyklisch sind; (3.3) lautet dann

$$d\omega^1 = \partial_k f_i \, dx_k \wedge dx_i = \sum_{ikl\ \text{zykl}} (\partial_i f_k - \partial_k f_i)\, do_l = \mathrm{rot}\vec{f}\cdot d\vec{o}\,. \quad (3.4)$$

Ü: Wählen Sie $\omega^2 = \vec{f}\cdot d\vec{o}$ und zeigen Sie, daß $d\omega^2 = \mathrm{div}\vec{f}\,dx_1 \wedge dx_2 \wedge dx_3$ ist.

Eine $(k+1)$-Form $d\omega^k$ kann ich über eine $(k+1)$-dimensionale Mannigfaltigkeit $\mathcal{M}$ integrieren, die k-Form ω^k über den Rand $\partial\mathcal{M}$ von $\mathcal{M}$. Dann gilt der verallgemeinerte **Satz von Stokes**:

$$\int_{\partial\mathcal{M}} \omega^k = \int_{\mathcal{M}} d\omega^k\,.$$

Mit ω^1 aus (3.4) erhalten Sie den üblichen Satz von Stokes, mit ω^2 aus der Übungsaufgabe den Satz von Gauß.
(In dem erwähnten Buch von V. I. Arnold bezeichnet der Autor, ein sehr bekannter russischer Mathematiker, den verallgemeinerten Stokes'schen Satz als „Satz von Newton–Leibniz–Gauß–Green–Ostrogradski–Stokes–Poincaré".)

Wenn für die k-Form ω^k gilt, daß $d\omega^k = 0$ ist, verschwindet die Integration über ω^k über eine geschlossene k-Mannigfaltigkeit (die der Rand einer $(k+1)$-Mannigfaltigkeit ist) nach dem Stokes'schen Satz. Ein solches ω^k heißt **geschlossen**. Wenn speziell ω^k durch Anwendung von d auf eine $(k-1)$-Form $\widetilde{\omega}^{k-1}$ entsteht, $\omega^k = d\widetilde{\omega}^{k-1}$, ist ω^k wegen

$dd = 0$ geschlossen. Eine solche Form heißt **exakt**. Im $\mathbb{R}^n$ gilt auch das Umgekehrte: alle geschlossenen Formen sind exakt. Ein spezieller Fall dieser Aussage ist, daß man ein rotationsfreies Vektorfeld $\vec{A}(\vec{x})$ im $\mathbb{R}^3$ als Gradienten eines Skalarfeldes schreiben kann: man wähle $\omega^1 = \vec{A} \cdot d\vec{x}$ und wende (3.4) an; das zeigt, daß ω^1 geschlossen, folglich exakt ist: $\omega^1 = df = \vec{\partial} f \cdot d\vec{x}$.

Anhang C

Ergänzungen

C.1 Stetigkeit der Differentiation

Lemma C.1.1 *Wenn $T_\alpha \to T$, $T_\alpha, T \in \mathcal{S}'$, dann gilt auch $T'_\alpha \to T'$.*

Beweis: Für alle $f \in S$ ist

$$\lim T'_\alpha(f) = -\lim T_\alpha(f') = -T(f) = T'(f)\,.$$

■

Folgerungen und Beispiele:

(i) Wenn $\sum_1^\infty T_n$ konvergiert, darf man gliedweise differenzieren!

(ii) Beispiel: $g_\nu(x) := (1/\nu)\sin \nu x$ konvergiert punktweise gegen 0, die Ableitungen $g'_\nu(x) = \cos \nu x$ konvergieren nicht. Aber $T_{g'_\nu} = (T_{g_\nu})'$ konvergiert gegen 0 nach dem obigen Lemma; direkt kontrolliert:

$$\int g'_\nu(x) f(x)\, dx = \int \cos \nu x f(x)\, dx = -\frac{1}{\nu}\int \sin \nu x f(x)\, dx \longrightarrow 0\,.$$

(iii) Die Funktion $g(x)$ sei 2π-periodisch und integrabel, ihre Fourier-Koeffizienten $c_n = \int_0^{2\pi} g(x) e^{-inx}\, dx$ sind wohldefiniert, doch im allgemeinen Fall ist g(x) nicht notwendig durch die Fourier-Reihe gegeben, die ja auch nicht immer konvergieren muß. Aber:

$$g(x) = \sum_{-\infty}^{\infty} c_n e^{inx} \quad \text{im distributiven Sinn.} \tag{1.2}$$

Denn nach der Parsvalschen Ungleichung ist $\int |g(x)|^2\,dx = \sum |c_n|^2$, d. h. $\{c_n\}_{n\in\mathbb{Z}} \in \ell^2(\mathbb{C})$. Deshalb existiert die Funktion

$$G(x) := \sum_{-\infty}^{-1} \frac{c_n}{in} e^{inx} + c_0 x + \sum_{1}^{\infty} \frac{c_n}{in} e^{inx}\,; \tag{1.3}$$

und nach einem Satz aus der Theorie der Fourier-Transformationen gilt $G' = g$ fast überall. Fassen wir also (1.3) als Distributionsgleichung auf, dürfen wir gliedweise differenzieren und erhalten Gleichung (1.2).
Wir können auch anders argumentieren: seien $\gamma_n(f)$ die Fourier-Koeffizienten einer 2π-periodischen Testfunktion f, dann gilt (wieder mit der Parsevalschen Gleichung)

$$\begin{aligned} T_g(f) &= \int \overline{g(x)} f(x)\,dx = \sum \overline{c}_n \gamma_n(f) \\ &= \sum \overline{c}_n \int e^{-inx} f(x)\,dx = \int \overline{\sum c_n e^{inx}} f(x)\,dx\,. \end{aligned}$$

Die letzte Gleichheit ist hier als formale Schreibweise für eine Distribution zu lesen.

C.2 Grundlösung der Wärmeleitungsgleichung

Wir hatten früher die Lösung der Wärmeleitungsgleichung in einer Dimension

$$\partial_t u(x,t) = \kappa\, \partial_x^2 u(x,t) \tag{2.1}$$

durch Fourier-Transformation bezüglich x gefunden: $\partial_t \tilde{u} = -\kappa k^2 \tilde{u}$, und folglich $\tilde{u} = c(k) e^{-\kappa k^2 t}$. Die freie Funktion $c(k)$ war durch die Anfangsbedingung $u(x,0) = u_0(x)$ zu bestimmen: $c(k) = (\mathcal{F}u_0)(k)$, Einsetzen und Rücktransformation ergab die Lösung. Man kann die Lösung mithilfe einer „Grundlösung“ expliziter angeben:

Setzten wir $c(k) = (4\pi)^{-1}$, so erhalten wir eine spezielle Lösung von (2.1):

$$u_1(x,t) = \frac{1}{\sqrt{4\pi\kappa t}}\, e^{-\frac{x^2}{4\kappa t}}\,.$$

Die Konstante ist so gewählt, daß $\int u_1(x,1)\,dx = 1$ ist, wir können Lemma 2.2.7 mit $a = \sqrt{t}$ anwenden und finden $u(x,0) = \delta(x)$. Wir lesen $u_1(\cdot,t)$ also als Distributionslösung; wegen Lemma 3.3.14: $D(T_1 * T_2) = (DT_1) * T_2$ und $(\delta * u_0)(x) = u_0(x)$ löst dann $(u_1 * u_0)(x,t)$ die Gleichung (2.1) mit der gewünschten Anfangsbedingung. Ausführlich geschrieben:

$$u(x,t) = \frac{1}{\sqrt{4\pi\kappa t}} \int_{-\infty}^{\infty} e^{-\frac{(x-x')^2}{4\kappa t}} u_0(x')\,dx'\,.$$

Das ist ein schnell konvergierendes Integral, das sich gut approximieren läßt; und es entfällt die Notwendigkeit der Fourier-Hin- und -Rücktransformation von u_0 bzw. $\tilde{u}$.

C.3 Die Vollständigkeit der Distributionenräume

Theorem C.3.1 *Die Räume $\mathcal{D}'$ und $\mathcal{S}'$ sind vollständig bezüglich der $\mathcal{S}$-Topologie, d. h. jede Cauchy-Folge von Distributionen definiert wiederum eine Distribution.*

Beweis: siehe Literatur, z. B. Gelfand-Shilov, Bd. I, Anhang.

C.4 Kramers-Kronig-Relationen

Viele Prozesse, viele Geräte in der Physik versucht man als lineare Vorgänge oder Operationen zu verstehen. Beispiele:
(i) Ein linearer Verstärker macht aus einem Eingangssignal ein Ausgangssignal, das linear vom Eingangssignal abhängt, sehr allgemein geschrieben:

$$f_a(t) = \int G(\tau) f_e(t-\tau)\, d\tau = (G * f_e)(t)\,. \tag{4.1}$$

Das Ausgangssignal $f_a(t)$ sollte zur Zeit t natürlich nicht von Werten von $f_e(t')$ zu späteren Zeiten $t' > t$ abhängen; es genügt also, über das Intervall $(0, \infty)$ zu integrieren, anders: die Kausalität, d. h. das Prinzip, daß die Ursache nicht vor der Wirkung liegen kann, verlangt, daß $G(t) = 0$ für $t < 0$. Es ist hier natürlich nicht ausgeschlossen, daß $G(t)$ delta-artige Anteile enthält.
(ii) In einem homogenen Medium vereinfacht sich die allgemeine Beziehung $\vec{D} = \varepsilon_0 \vec{E} + \vec{P}$ zwischen dieelektrischer Verschiebung und elektrischem Feld im einfachsten Fall dadurch, daß man die Polarisation $\vec{P}$ proportional zu $\vec{E}$ ansetzt: $\vec{D} = \varepsilon \vec{E}$. Das entspricht der Vorstellung, daß sich die Elektronenhüllen der Atome unter dem Einfluß des elektrischen Feldes gegenüber den Atomkernen verschieben, wobei die Größe der Verschiebung proportional zur Feldstärke ist. Wenn es sich um ein sehr hochfrequentes elektrisches Feld handelt, werden die Elektronen des Materials nicht mehr trägheitsfrei, d. h. praktisch instantan folgen können, die Polarisation hängt von der Frequenz ω ab:

$$\vec{D}(\omega) = \varepsilon_0 (1 + \chi(\omega)) \vec{E}(\omega)\,. \tag{4.2}$$

Ich hätte hier besser $\tilde{D}$ und $\tilde{E}$ geschrieben, lassen Sie mich für die Fourier-Rücktransformierte dieselben Symbole verwenden; außerdem habe ich eine mögliche $\vec{x}$-Abhängigkeit unterdrückt:

$$\frac{1}{\varepsilon_0} \vec{D}(\vec{x}, t) = \vec{E}(\vec{x}, t) + \sqrt{2\pi}^{-1} \int_0^\infty \chi(\tau) \vec{E}(\vec{x}, t-\tau)\, d\tau\,; \tag{4.3}$$

dem Kausalitätsprinzip entsprechend habe ich auch hier nur von 0 ab integriert. Mit $\tilde{G}(\omega) = \sqrt{2\pi}^{-1}(1 + \chi(\omega))$ (und folglich $G(t) = \delta(t) + \sqrt{2\pi}^{-1}\chi(t)$) hat Gleichung (4.3) hinsichtlich der Zeitabhängigkeit dieselbe Form wie (4.1).

Das Kausalitätsprinzip hat meßbare Folgen für $\tilde{G}(\omega)$!
Man kann annehmen, daß die „Gedächtniseffekte“, die durch G beschrieben werden, nicht zu weit in die Vergangenheit zurückreichen, d. h. daß $G(t)$ mit t hinreichend stark abfällt und somit integrierbar ist. Weil in

$$\tilde{G}(\omega) = \sqrt{2\pi}^{-1} \int_0^\infty G(t) e^{i\omega t}\, dt$$

nur positive t wesentlich sind, können wir ω als komplexe Variable mit positivem Imaginärteil betrachten und bekommen ein gleichmäßig in ω konvergierendes Integral, mehr noch: $\tilde{G}(\omega)$ ist analytisch in der oberen Halbebene und stetig auf der reellen Achse, weil G nach unseren Annahmen integrierbar ist (das gilt ja auch, wenn $G(t)$ eine Delta-Funktion enthält). Aus physikalischen Gründen wollen wir auch annehmen, daß $\tilde{G}(\omega)$ auf $\mathbb{R}$ integrabel ist. Nach dem Cauchyschen Integralsatz gilt für eine analytische Funktion

$$f(z_0) = \frac{1}{2\pi i} \oint \frac{f(z)}{z - z_0}\, dz\,,$$

wobei die Integration über eine geschlossene Kurve im Analytizitätsgebiet (oder in seinem Abschluß) geht, die den Punkt z_0 enthält. Die Verschiebung des Integrationswegs auf den Rand ist erlaubt, wenn der Integrand dort noch stetig ist. Nehmen wir also $\Im\omega > 0$ und führen die Integration über die reelle Achse und den unendlich fernen Halbkreis in der oberen Halbebene, der natürlich wegen des Faktors $e^{-\Im\omega t}$ nichts beiträgt:

$$\tilde{G}(\omega) = \frac{1}{2\pi i} \int_{-\infty}^{\infty} \frac{\tilde{G}(\omega')}{\omega' - \omega}\, d\omega'\,,$$

oder, wenn wir $\Im\omega$ gegen 0 gehen lassen,

$$\tilde{G}(\omega) = \frac{1}{2\pi i} \int_{-\infty}^{\infty} \frac{\tilde{G}(\omega')}{\omega' - \omega - i0}\, d\omega'\,.$$

$1/(\omega' - \omega - i0)$ ist hier natürlich als Distribution aufzufassen, wir können sie erweitern auf analytische Funktionen, die auf der reellen Achse noch stetig und integrierbar sind. Erinnern wir uns an die Gleichungen (III.1.9) und (III.1.10): $(x+i0)^{-1} \pm (x-i0)^{-1} = P\frac{1}{x}$ bzw. $= 2\pi i\delta(x)$, so ist

$$\frac{1}{\omega' - \omega - i0} = P\frac{1}{\omega' - \omega} + i\pi\delta(\omega' - \omega)\,.$$

Einsetzen in das obige Integral und Umordnen (rechts taucht ja auch ein Term $(1/2)\tilde{G}(\omega)$) auf) ergibt schließlich

$$\tilde{G}(\omega) = \frac{1}{\pi i}\, \mathrm{P}\text{-}\!\!\int_{-\infty}^{\infty} \frac{\tilde{G}(\omega')}{\omega' - \omega}\, d\omega'\,. \tag{4.4}$$

Aufgeteilt in Real- und Imaginärteil:

$$\Re\tilde{G}(\omega) = \frac{1}{\pi}\, \mathrm{P}\text{-}\!\!\int_{-\infty}^{\infty} \frac{\Im\tilde{G}(\omega')}{\omega' - \omega}\, d\omega'\,, \tag{4.5}$$

$$\Im\tilde{G}(\omega) = -\frac{1}{\pi}\, \mathrm{P}\text{-}\!\!\int_{-\infty}^{\infty} \frac{\Re\tilde{G}(\omega')}{\omega' - \omega}\, d\omega'\,. \tag{4.6}$$

Das sind die **Kramers-Kronig-Relationen**, manchmal auch Dispersionsrelationen genannt.

Wir sehen also, daß die Fourier-Transformierte einer Funktion, die für $t < 0$ verschwindet, notwendig komplex sein muß und daß Real- und Imaginärteil diesen Bedingungen genügen müssen. Denken wir wieder an das Beispiel mit der Polarisation, so folgt also, daß $\varepsilon(\omega) = 1 + \chi(\omega)$ komplex sein muß. In der Elektrodynamik lernt man, daß der Imaginärteil von ε verantwortlich ist für die Dämpfung der Wellen im Material, während der Realteil mit dem Brechungsindex zusammenhängt ($n = \sqrt{\mu_{\rm rel} \Re \varepsilon_{\rm rel}}$). Beides sind meßbare Größen, so daß man die Gleichungen (4.5) und (4.6) direkt überprüfen kann.

Index