

Funktionalanalysis

Gernot Eichmann

Vorlesungsskriptum
Wintersemester 2023/24

Universität Graz
Institut für Physik

Revision: 31. Januar 2024

Literatur

- [Are15] T. Arens, F. Hettlich, Ch. Karpfinger, U. Kockelkorn, K. Lichtenegger, H. Stachel, *Mathematik*, Springer Spektrum (2015)
- [Beh20] J. Behrndt, M. Holzmann, P. Schlosser, Vorlesungsskriptum *Funktionalanalysis und partielle Differentialgleichungen*, WS 2020/21, TU Graz
- [Bla93] P. Blanchard, E. Brüning, *Distributionen und Hilbertraumoperatoren*, Springer (1993)
- [Gro14] S. Großmann, *Funktionalanalysis (im Hinblick auf Anwendungen in der Physik)*, Springer (2014)
- [Has13] S. Hassani, *Mathematical Physics (A Modern Introduction to its Foundations)*, Springer (2013)
- [Lan16] C. B. Lang, N. Pucker, *Mathematische Methoden der Physik*, Springer Spektrum (2016)
- [Nee22] J. van Neerven, *Functional Analysis*, Cambridge University Press (2022)
- [Ree81] M. Reed, B. Simon, *Methods of Modern Mathematical Physics I: Functional Analysis*, Academic Press (1981)
- [Sch21] W. Schweiger, Vorlesungsskriptum *Funktionalanalysis*, WS 2021/22, Universität Graz
- [Sta11] I. Stakgold, M. Holst, *Green's Functions and Boundary Value Problems*, Wiley (2011)
- [Wei00] J. Weidmann, *Lineare Operatoren in Hilberträumen (Teil 1: Grundlagen)*, Teubner (2000)
- [Wer05] D. Werner, *Funktionalanalysis*, Springer (2005)

Inhaltsverzeichnis

1	Funktionsräume	1
1.1	Vektorräume	1
1.2	Metrik und Norm	3
1.3	Hilberträume	10
1.4	Orthonormalbasen	13
2	Fourierreihen	20
2.1	Definition der Fourierreihe	20
2.2	Konvergenzkriterien	26
2.3	Weitere Anwendungen	33
3	Integraltransformationen	36
3.1	Fouriertransformation	38
3.2	Laplace-Transformation	47
4	Funktionale und Distributionen	52
4.1	Distributionen	54
4.2	Funktionalableitung	64
5	Lineare Operatoren in Hilberträumen	69
5.1	Grundbegriffe	70
5.2	Beschränkte Operatoren	73
5.3	Unbeschränkte Operatoren	88
5.4	Das Spektrum linearer Operatoren	96

Kapitel 1

Funktionenräume

1.1 Vektorräume

Die **Funktionalanalysis** ist in einem gewissen Sinn die Vereinigung der linearen Algebra mit der Analysis. Aus der linearen Algebra kennen Sie endlichdimensionale Vektorräume, lineare Abbildungen, Matrizen und Eigenwertprobleme. Die Analysis dreht sich um Funktionen und deren Eigenschaften wie z.B. Differenzierbarkeit und Integrierbarkeit. Die Funktionalanalysis beschäftigt sich mit unendlichdimensionalen Vektorräumen, deren Elemente Funktionen sind. Differenzial- und Integraloperatoren, die auf Funktionen wirken, sind dann lineare Abbildungen und können ebenso durch (unendlichdimensionale) Matrizen beschrieben werden, die auf Vektoren wirken, und Differenzialgleichungen werden zu Eigenwertproblemen.

Während sich Konzepte wie lineare Unabhängigkeit, Orthogonalität, Skalarprodukt und Norm sofort auf unendlichdimensionale Vektorräume übertragen lassen, stoßen wir in der Funktionalanalysis auf neue Probleme: während wir es bei endlichdimensionalen Vektorräumen mit endlichen Summen zu tun haben, sind es bei unendlichdimensionalen Vektorräumen unendliche Summen (d.h. Reihen) – somit müssen wir uns in letzterem Fall immer auch die **Konvergenzfrage** stellen. Betrachten wir als Beispiel für einen endlichdimensionalen Vektorraum den \mathbb{R}^n . Ein Element $\mathbf{x} \in \mathbb{R}^n$ können wir durch seine Komponenten x_i ausdrücken, und wir können den Betrag $|\mathbf{x}|$ des Vektors bilden:

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad |\mathbf{x}| = \sqrt{\sum_{i=1}^n x_i^2}. \quad (1.1)$$

Was passiert nun aber für $n \rightarrow \infty$? In diesem Fall müssten wir die Konvergenz der Partialsummen $\sum_{i=1}^n x_i^2$ untersuchen, um festzustellen, ob der Betrag auch tatsächlich existiert:

$$|\mathbf{x}| = \lim_{n \rightarrow \infty} \sqrt{\sum_{i=1}^n x_i^2} < \infty? \quad (1.2)$$

Unendlichdimensionale Vektorräume sind daher ungleich reichhaltiger und komplizierter.

Vektorraum. Auch wenn Sie den Begriff des Vektorraums bereits kennen, wollen wir ihn sicherheitshalber wiederholen, da er für alles Weitere eine zentrale Rolle spielt:

DEFINITION. Ein **Vektorraum** \mathbb{V} über einem Körper \mathbb{K} (\mathbb{R} oder \mathbb{C}) ist eine Menge, auf der eine Addition $+$: $\mathbb{V} \times \mathbb{V} \rightarrow \mathbb{V}$ sowie eine Skalarmultiplikation \cdot : $\mathbb{K} \times \mathbb{V} \rightarrow \mathbb{V}$ definiert sind. Dabei erfüllt die Addition folgende Eigenschaften ($\forall x, y, z \in \mathbb{V}$):

- (i) Assoziativität: $(x + y) + z = x + (y + z)$
- (ii) Kommutativität: $x + y = y + x$
- (iii) Existenz eines neutralen Elements $0 \in \mathbb{V}$, sodass $x + 0 = x$
- (iv) Existenz eines inversen Elements $(-x) \in \mathbb{V}$, sodass $x + (-x) = 0$

(In anderen Worten, \mathbb{V} bildet eine Abelsche Gruppe bezüglich der Addition.) Die Skalarmultiplikation erfüllt folgende Eigenschaften ($\forall x, y \in \mathbb{V}$ und $\lambda, \mu \in \mathbb{K}$):

- (i) $\lambda(x + y) = \lambda x + \lambda y$,
- (ii) $(\lambda + \mu)x = \lambda x + \mu x$,
- (iii) $(\lambda\mu)x = \lambda(\mu x)$,
- (iv) $1x = x$.

“Vektoren” sind also lediglich Objekte, die wir miteinander addieren und mit (reellen oder komplexen) Zahlen multiplizieren können – von einer *Länge* oder *Richtung* ist hier noch keine Rede. Diese Definition ist so allgemein, dass sie sofort auf Funktionen angewandt werden kann. Einige Beispiele für häufig verwendete Funktionsräume sind:

- Der Vektorraum $\mathcal{F}[a, b]$ der reell- bzw. komplexwertigen Funktionen auf dem reellen Intervall $[a, b] \subset \mathbb{R}$:

$$\mathcal{F}[a, b] := \{f \mid f: [a, b] \subset \mathbb{R} \rightarrow \mathbb{K}\}, \quad \text{mit } \mathbb{K} = \mathbb{R} \text{ oder } \mathbb{C}. \quad (1.3)$$

Die Vektoraddition $h = f + g$ (mit $f, g, h \in \mathcal{F}$) ist dabei **punktweise** erklärt, d.h. $h(x) = f(x) + g(x) \forall x \in [a, b]$. Wir benutzen also die Addition der Funktionswerte in \mathbb{K} , um die Addition für die Vektoren in \mathcal{F} zu erklären. Dasselbe gilt für die Skalarmultiplikation $h = \lambda f$ ($f, h \in \mathcal{F}$ und $\lambda \in \mathbb{K}$): $h(x) = \lambda f(x) \forall x \in [a, b]$. Man sieht hier auch die Analogie zu endlichdimensionalen Vektoren im \mathbb{R}^n bzw. \mathbb{C}^n , bei denen die ‘Funktionswerte’ eines Vektors \mathbf{f} seine Komponenten $f_i \in \mathbb{R}$ bzw. \mathbb{C} sind ($i = 1, \dots, n$). Bei einer Funktion im unendlichdimensionalen Vektorraum wird der diskrete Index i zum kontinuierlichen ‘Index’ x und die Komponenten f_i zu den Funktionswerten $f(x)$.

- Der Vektorraum $C[a, b]$ der reell- bzw. komplexwertigen und **stetigen** Funktionen auf dem Intervall $[a, b] \subset \mathbb{R}$:

$$C[a, b] := \{f \mid f: [a, b] \subset \mathbb{R} \rightarrow \mathbb{K} \text{ und } f \text{ stetig}\}. \quad (1.4)$$

$C[a, b]$ ist bezüglich der punktweise erklärten Vektoraddition und Skalarmultiplikation ebenfalls ein Vektorraum und ein Unterraum von $\mathcal{F}[a, b]$.

- Der Vektorraum $C^m(D)$ der m -fach **stetig differenzierbaren** Funktionen auf D , wobei D eine Teilmenge der reellen Zahlen \mathbb{R}^n bezeichnet. (Zur Erinnerung: m -fach stetig differenzierbar bedeutet, dass die Funktion m -fach differenzierbar ist und die m -te Ableitung stetig ist, denn die Stetigkeit der ersten $m - 1$ Ableitungen folgt bereits aus der Differenzierbarkeit.)
- Der Vektorraum $C^\infty(D)$ der **glatten**, d.h. beliebig oft stetig differenzierbaren Funktionen auf D .
- Der Vektorraum $C_0^m(D)$ der m -fach stetig differenzierbaren Funktionen auf D mit einem **kompakten Träger**, d.h. die nur auf einer beschränkten, abgeschlossenen Teilmenge von D ungleich Null sind.
- Der Vektorraum $C_0^\infty(D)$ der glatten Funktionen auf D mit einem kompakten Träger.
- Der **Schwartz-Raum** S der Funktionen $\in C^\infty$, die gemeinsam mit allen ihren Ableitungen im Unendlichen schneller verschwinden als jede Potenz. Typische Beispiele sind exponentiell abfallende Funktionen der Art $e^{-\alpha x^2}$. Die Schwartz-Funktionen werden später eine wichtige Rolle bei Distributionen spielen, denn mit ihrer Hilfe lassen sich Delta- und Stufenfunktionen differenzieren. Es gilt auch $C_0^\infty \subset S \subset C^\infty$.

1.2 Metrik und Norm

Metrik. Wir wollen die oben definierten Vektoren nun schrittweise mit weiteren Eigenschaften versehen, um auch die Konzepte von Länge und Richtung erfassen zu können. Dazu gehen wir zunächst einen Schritt zurück und führen den Begriff des *Abstands* ein. Eine minimale Forderung an einen Funktionenraum ist die Existenz einer Metrik, die es uns erlaubt, Abstände zu messen. Dazu benötigen wir noch keine Vektorraum-Struktur, denn eine Metrik kann für beliebige Mengen definiert werden:

DEFINITION. Ein **metrischer Raum** ist eine Menge X mit einer Abbildung (**Metrik**) $d : X \times X \rightarrow \mathbb{R}$, die jedem geordneten Paar (x, y) von Elementen aus X eine reelle Zahl $d(x, y)$ mit folgenden Eigenschaften zuordnet ($\forall x, y, z \in X$):

- (i) $d(x, y) \geq 0$ (Nichtnegativität)
- (ii) $d(x, y) = 0 \Leftrightarrow x = y$ (Eindeutigkeit)
- (iii) $d(x, y) = d(y, x)$ (Symmetrie)
- (iv) $d(x, z) \leq d(x, y) + d(y, z)$ (Dreiecksungleichung)

Tatsächlich kann (i) weggelassen werden, denn die Nichtnegativität folgt bereits aus den anderen drei Eigenschaften:

$$d(x, y) = \frac{1}{2} (d(x, y) + d(x, y)) \stackrel{\text{(iii)}}{=} \frac{1}{2} (d(x, y) + d(y, x)) \stackrel{\text{(iv)}}{\geq} \frac{1}{2} d(x, x) \stackrel{\text{(ii)}}{=} 0. \quad (1.5)$$

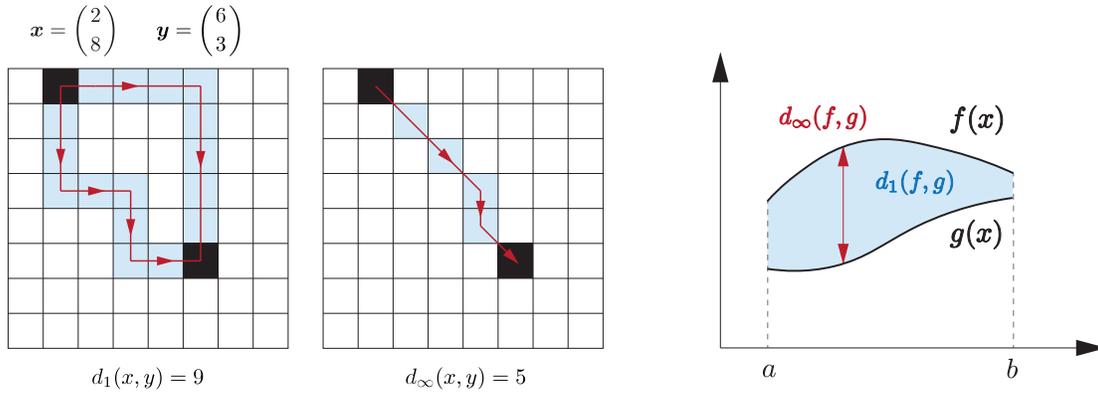


FIG. 1.1: Beispiele für die Manhattan-Metrik ($p = 1$) und Maximum-Metrik ($p = \infty$) im Raum \mathbb{R}^2 bzw. $C[a, b]$

Das sind alles naheliegende Forderungen, die man von einem Abstands begriff erwarten würde: Abstände sollen nicht negativ sein; aus (ii) folgt $d(x, x) = 0$; die Symmetrie (iii) ist sinnvoll, da bei einem geordneten Paar (x, y) die Reihenfolge eine Rolle spielt; und die Dreiecksungleichung (iv) ist eine plausible Forderung als Verallgemeinerung aus dem \mathbb{R}^n (und typischerweise auch jene Eigenschaft, die am schwierigsten zu beweisen ist). Einige Beispiele für metrische Räume sind:

- Der Raum $X = \mathbb{R}^n$ mit den **p -Metriken**:

$$d_p(x, y) = \left(\sum_{i=1}^n |x_i - y_i|^p \right)^{1/p}. \quad (1.6)$$

Der Fall $p = 1$ liefert die sogenannte **Manhattan-Metrik** (auch Taxi-Metrik genannt). $p = 2$ entspricht der üblichen **Euklidischen Metrik** $d_2(x, y) = |\mathbf{x} - \mathbf{y}|$ als Betrag der Differenz zweier Vektoren. Der Grenzfall $p \rightarrow \infty$ führt auf die **Maximum-Metrik**, die auch Tschebyschev- oder (in zwei Dimensionen) Schachbrett-Metrik genannt wird, siehe Fig. 1.1:

$$d_\infty(x, y) = \max_i |x_i - y_i|. \quad (1.7)$$

Verschiedene Abstands begriffe führen somit auch zu verschiedenen Geometrien.

- Ein Beispiel für einen Funktionenraum ist $X = C[a, b]$, der Raum der stetigen, reellwertigen Funktionen auf dem abgeschlossenen Intervall $[a, b]$. Auch hier kann man verschiedene p -Metriken definieren (siehe Fig. 1.1):

$$d_p(f, g) = \left(\int_a^b |f(x) - g(x)|^p dx \right)^{1/p}. \quad (1.8)$$

$d_1(f, g)$ ist die Fläche zwischen den Funktionen f und g , $d_2(f, g)$ die Verallgemeinerung der Euklidischen Metrik, und die Maximum-Metrik

$$d_\infty(f, g) = \max_{x \in [a, b]} |f(x) - g(x)| \quad (1.9)$$

entspricht dem Maximalwert der Differenz zwischen den beiden Funktionen.

- Die **diskrete Metrik** ist definiert durch $d(x, y) = 0$ für $x = y$ und $d(x, y) = 1$ für $x \neq y$. Zum Beweis der Dreiecksungleichung $d(x, z) \leq d(x, y) + d(y, z)$ unterscheidet man zwei Fälle: Für $x = z$ ist die Ungleichung erfüllt. Für $x \neq z$ muss $y \neq x$ oder $y \neq z$ gelten, da y nicht gleichzeitig mit zwei verschiedenen Elementen übereinstimmen kann, womit die Ungleichung ebenfalls erfüllt ist.

Nicht alle in der Physik wichtigen Räume sind metrische Räume, z.B. besitzt der Minkowski-Raum der speziellen Relativitätstheorie keine Metrik im eigentlichen Sinn, da die Minkowski-“Metrik” auch imaginäre Werte annehmen kann (raumartige Abstände) und $d(x, y) = 0$ für $x \neq y$ erlaubt ist (lichtartige Abstände).

Vollständigkeit. Die Existenz einer Metrik, mit der wir ein Maß für die ‘Nähe’ eingeführt haben, reicht bereits aus, um die Konvergenz von Folgen zu untersuchen und den Begriff der Vollständigkeit zu definieren. Als Grenzwert einer Folge $f_1, f_2, \dots, f_n, \dots$ von Elementen eines Raumes würde man anschaulich einen Vektor f verstehen, von dem sich die f_n immer weniger “unterscheiden”. Dazu muss man den Begriff “unterscheiden” quantifizieren, was eben durch die Metrik bewerkstelligt wird. Damit es einen solchen Grenzwert gibt, ist sicher auch notwendig, dass sich die f_n *untereinander* immer weniger unterscheiden. Das führt auf den Begriff der **Cauchy-Folge** (nach dem französischen Mathematiker [Augustin-Louis Cauchy](#), 1879–1857): Cauchy-Folgen sind solche Folgen, deren Elemente sich (bezüglich der Metrik d) beliebig nahe kommen.

DEFINITION: Sei X ein metrischer Raum mit Metrik d und $(f_n)_{n=1}^\infty$ eine Folge von Elementen $f_n \in X$.

- Die Folge $(f_n)_{n=1}^\infty$ heißt **Cauchy-Folge**, wenn

$$\forall \epsilon > 0 \quad \exists N(\epsilon) \in \mathbb{N} : d(f_n, f_m) < \epsilon \quad \forall n, m \geq N(\epsilon). \quad (1.10)$$

- $f \in X$ heißt **Grenzwert** der Folge $(f_n)_{n=1}^\infty$, wenn

$$\forall \epsilon > 0 \quad \exists N(\epsilon) \in \mathbb{N} : d(f_n, f) < \epsilon \quad \forall n \geq N(\epsilon). \quad (1.11)$$

- Hat eine Folge $(f_n)_{n=1}^\infty$ mit $f_n \in X$ einen Grenzwert $f \in X$, heißt sie **konvergent**, ansonsten **divergent**.

Eine Folge von Objekten f_n konvergiert also gegen den Wert f , wenn die numerische Folge $d(f_n, f)$ nach Null strebt. Eine notwendige Bedingung für die Konvergenz einer Folge ist, dass sie eine Cauchyfolge ist. Allerdings ist nicht jede Cauchyfolge auch konvergent, denn der Grenzwert muss nicht im Raum liegen – typische Beispiele sind Folgen von rationalen Zahlen, deren Grenzwerte keine rationalen Zahlen sind, z.B.

$$\lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n = e. \quad (1.12)$$

Wenn eine Cauchyfolge einen Grenzwert hat, ist er aber eindeutig – gäbe es nämlich mehrere Grenzwerte f und \tilde{f} , wäre ihr Unterschied Null: $d(f, \tilde{f}) \leq d(f, f_n) + d(f_n, \tilde{f}) < \epsilon_1 + \epsilon_2$.

Mathematisch besonders bequem sind Räume, in denen die Existenz eines Grenzwerts von Cauchyfolgen *gesichert* ist. Das führt auf folgende Definition:

DEFINITION: Ein metrischer Raum X mit Metrik $d(x, y)$ heißt **vollständig**, wenn in ihm jede Cauchyfolge konvergent ist.

Im Gegensatz zu den rationalen Zahlen \mathbb{Q} sind die reellen Zahlen \mathbb{R} vollständig, da jede Cauchyfolge in \mathbb{R} auch einen Grenzwert aus \mathbb{R} besitzt (auf dieser Tatsache beruhen wesentliche Aussagen der Analysis). Dasselbe gilt auch für die komplexen Zahlen. Die Konvergenz einer Folge und die Vollständigkeit eines Raumes hängt von der jeweiligen Metrik ab – Funktionenfolgen können bezüglich einer Metrik konvergent sein, bezüglich einer anderen Metrik aber divergent (Beispiele folgen weiter unten).

Norm. Während der Abstands begriff für beliebige Mengen eingeführt werden kann, benötigt ein *Längenbegriff* bereits die algebraische Struktur eines Vektorraums, dessen Elemente wir addieren und mit Skalaren multiplizieren können:

DEFINITION: Ein **normierter Raum** ist ein Vektorraum \mathbb{V} über einem Körper \mathbb{K} (\mathbb{R} oder \mathbb{C}) mit einer Abbildung (**Norm**) $\|\cdot\|: \mathbb{V} \rightarrow \mathbb{R}$, $x \mapsto \|x\|$, die jedem Element $x \in \mathbb{V}$ eine reelle Zahl mit folgenden Eigenschaften zuordnet ($\forall x \in \mathbb{V}$, $\lambda \in \mathbb{K}$):

- (i) $\|x\| \geq 0$ (Nichtnegativität)
- (ii) $\|x\| = 0 \Leftrightarrow x = 0$ (Eindeutigkeit)
- (iii) $\|\lambda x\| = |\lambda| \|x\|$ (absolute Homogenität)
- (iv) $\|x + y\| \leq \|x\| + \|y\|$ (Dreiecksungleichung)

Auch hier sind einige der Eigenschaften redundant, denn die Forderungen $\|x\| = 0 \Rightarrow x = 0$ sowie (iii) und (iv) genügen: aus $\lambda = 0$ in (iii) folgt $x = 0 \Rightarrow \|x\| = 0$; aus $\lambda = -1$ folgt $\| -x \| = \|x\|$; und mit $y = -x$ in (iv) folgt $\|x\| \geq 0$.

Ein normierter Raum ist gleichzeitig ein metrischer Raum, denn die Norm ‘induziert’ eine Metrik

$$d(x, y) = \|x - y\|. \quad (1.13)$$

Umgekehrt gilt das nicht immer, denn aus einer Metrik folgt nicht automatisch eine Norm: ein metrischer Raum muss kein Vektorraum sein, und $\|x\| = d(x, 0)$ erfüllt nicht zwangsläufig alle Eigenschaften einer Norm wie z.B. die Homogenität (etwa bei der diskreten Metrik).

Die p -Metriken in (1.6–1.7) und (1.8–1.9) gewinnt man aus der jeweiligen **p-Norm** bzw. **Maximumsnorm** im \mathbb{R}^n oder \mathbb{C}^n bzw. im Raum $C[a, b]$ der über $[a, b]$ stetigen Funktionen:

$$\blacksquare x \in \mathbb{R}^n \text{ bzw. } \mathbb{C}^n : \quad \|x\|_p = \left(\sum_{i=1}^n |x_i|^p \right)^{1/p}, \quad \|x\|_\infty = \max_i |x_i| \quad (1.14)$$

$$\blacksquare f \in C[a, b] : \quad \|f\|_p = \left(\int_a^b dx |f(x)|^p \right)^{1/p}, \quad \|f\|_\infty = \max_{x \in [a, b]} |f(x)|. \quad (1.15)$$

Natürlich spielt auch bei normierten Räumen das Kriterium der Vollständigkeit eine zentrale Rolle, deswegen trifft man die nach dem polnischen Mathematiker [Stefan Banach](#) (1892–1945) benannte Definition:

DEFINITION: Ein **Banachraum** ist ein vollständiger normierter Raum, d.h. ein normierter Raum, in dem jede Cauchyfolge bezüglich der induzierten Metrik $d(x, y) = \|x - y\|$ einen Grenzwert besitzt.

Einige Beispiele für Banachräume sind:

- Der \mathbb{R}^n oder \mathbb{C}^n mit Elementen $x = (x_1 \dots x_n)$ und Norm $\|x\|_2 = \sqrt{\sum_{i=1}^n |x_i|^2}$.
- Der **Folgenraum** ℓ^p ist die Menge der reellen oder komplexen ($\mathbb{K} = \mathbb{R}$ oder \mathbb{C}) unendlichen Folgen, die “p-summierbar” sind (mit $p > 0$):

$$\ell^p = \left\{ x = (x_i)_{i=1}^{\infty} \mid x_i \in \mathbb{K} \quad \text{und} \quad \sum_{i=1}^{\infty} |x_i|^p < \infty \right\}. \quad (1.16)$$

Er ist ein Banachraum bezüglich der p-Norm (zum Beweis siehe z.B. [Wei00]):

$$\|x\|_p := \left(\sum_{i=1}^{\infty} |x_i|^p \right)^{1/p}. \quad (1.17)$$

Die Elemente dieses unendlich-dimensionalen Vektorraums sind Folgen, z.B. gilt für die harmonische Folge $x = (1/i)_{i=1}^{\infty} \in \ell^p$ für $p > 1$, da die Norm $\|x\|_p$ durch die harmonische Reihe definiert wird und diese für $p \leq 1$ nicht mehr konvergiert. Ein wichtiger Spezialfall ist der **Hilbertsche Folgenraum** ℓ^2 der quadratsummierbaren Folgen mit der Norm $\|x\|_2$. Nach dem Satz von Fischer-Riesz ist jeder unendlichdimensionale separable Hilbertraum zu ℓ^2 isomorph (mehr dazu später).

- Die **Lebesgue-Räume** $L^p(a, b)$ der in p -ter Potenz integrierbaren und reell- oder komplexwertigen Funktionen ($\mathbb{K} = \mathbb{R}$ oder \mathbb{C}) auf dem Intervall $(a, b) \subset \mathbb{R}$,

$$L^p(a, b) = \left\{ f \mid f : (a, b) \subset \mathbb{R} \rightarrow \mathbb{K} \quad \text{und} \quad \int_a^b dx |f(x)|^p < \infty \right\}, \quad (1.18)$$

sind Banachräume bezüglich der p-Norm

$$\|f\|_p = \left(\int_a^b dx |f(x)|^p \right)^{1/p}. \quad (1.19)$$

Hierbei ist das Lebesgue-Integral gemeint; d.h. eigentlich besteht der Raum $L^p(a, b)$ aus Äquivalenzklassen von Funktionen mit derselben Norm, die sich nur auf einer Menge vom Maß Null unterscheiden, und diese werden miteinander identifiziert. Ein Spezialfall, dem wir noch öfters begegnen werden, ist der Raum $L^1(a, b)$ der **absolut integrierbaren Funktionen**. Insbesondere wichtig für die Quantenmechanik ist der Raum $L^2(a, b)$ der **quadratintegrierbaren Funktionen**; er ist ein Banachraum bezüglich der Norm $\|\cdot\|_2$. Die im Raum $L^p(a, b)$ mit der Norm $\|\cdot\|_p$ gültige Dreiecksungleichung trägt auch den Namen **Minkowski-Ungleichung**.

- Der Raum $C[a, b]$ der stetigen Funktionen mit der Norm $\|f\|_2 = \sqrt{\int_a^b dx |f(x)|^2}$ ist nicht vollständig und somit **kein** Banachraum.

Um das zu zeigen, müssen wir nur *eine* Cauchyfolge von Funktionen $f_n \in C[a, b]$ finden, die nicht gegen eine stetige Funktion konvergiert. Betrachten wir als Beispiel das Intervall $[a, b] = [-1, 1]$ und die Folge (siehe Fig. 1.2)

$$f_n(x) = \begin{cases} 0 & -1 \leq x \leq -\frac{1}{n} \\ \frac{1}{2}(1 + nx) & -\frac{1}{n} < x < \frac{1}{n} \\ 1 & \frac{1}{n} \leq x \leq 1 \end{cases} \quad (1.20)$$

Man kann leicht zeigen, dass das eine Cauchyfolge ist:

$$\begin{aligned} (\|f_n - f_m\|_2)^2 &= \int_{-1}^1 dx (f_n(x) - f_m(x))^2 \stackrel{n < m}{\leq} 2 \int_0^{1/n} dx \left(1 - \frac{1}{2}(1 + nx)\right)^2 \\ &= \frac{1}{2} \int_0^{1/n} dx (1 - nx)^2 = \frac{1}{2} \left(-\frac{1}{n}\right) \frac{1}{3} (1 - nx)^3 \Big|_0^{1/n} = \frac{1}{6n}. \end{aligned} \quad (1.21)$$

Die Ungleichung ist aus Fig. 1.2 unmittelbar klar. Wählen wir nun $N_\epsilon = 1/(6\epsilon^2)$ für ein beliebiges ϵ , dann gilt $\forall n, m \geq N_\epsilon$:

$$(\|f_n - f_m\|_2)^2 \leq \frac{1}{6n} \leq \frac{1}{6N_\epsilon} = \epsilon^2 \quad \Rightarrow \quad \|f_n - f_m\|_2 \leq \epsilon. \quad (1.22)$$

Die Folge ist also eine Cauchyfolge. Ihr Grenzwert ist allerdings keine stetige Funktion $\in C[-1, 1]$, sondern die Stufenfunktion

$$f(x) = \begin{cases} 0 & -1 < x < 0, \\ 1 & 0 < x < 1. \end{cases} \quad (1.23)$$

Bezüglich der Maximumsnorm $\|f\|_\infty = \max_{x \in [a, b]} |f(x)|$ ist der Raum $C[a, b]$ allerdings vollständig und somit ein Banachraum! Konvergiert nämlich eine Folge stetiger Funktionen $f_n \in C[a, b]$ auf einem abgeschlossenen Intervall $[a, b]$ gleichmäßig gegen f , ist f stetig. Das Kriterium der **gleichmäßigen Konvergenz** entspricht gerade der Maximumsnorm:

$$\lim_{n \rightarrow \infty} \|f_n - f\|_\infty = \lim_{n \rightarrow \infty} \max_{x \in [a, b]} |f_n(x) - f(x)| = 0. \quad (1.24)$$

Anschaulich bedeutet das, dass sich die Graphen von $f_n(x)$ auf dem Intervall $[a, b]$ gleichmäßig dem Graphen der Grenzfunktion f annähern. Dazu betrachtet man einen “ ϵ -Schlauch”, d.h. einen Parallelstreifen $f(x) - \epsilon < y < f(x) + \epsilon$ der Dicke 2ϵ (Fig. 1.2). Die gleichmäßige Konvergenz kann man dann so ausdrücken: Für jedes $\epsilon > 0$ existiert ein N_ϵ , sodass für alle $n > N_\epsilon$ der Graph von $f_n(x)$ innerhalb des ϵ -Schlauchs liegt.

Der Vollständigkeitsbeweis erfolgt so: Seien $x, p \in [a, b]$ und $\epsilon > 0$. Eine Funktion f ist stetig am Punkt p , wenn zu jedem ϵ ein $\delta > 0$ existiert, sodass für alle x mit $|x - p| < \delta$ gilt: $|f(x) - f(p)| < \epsilon$. Betrachten wir jetzt irgendeine Folge (f_n) stetiger Funktionen, die gleichmäßig gegen einen Grenzwert f konvergiert. Dann gibt es zu jedem ϵ ein N_ϵ , sodass die Funktion f_n für genügend große $n > N_\epsilon$ im ϵ -Schlauch um f liegt, d.h. $|f_n(x) - f(x)| < \epsilon$ für alle x und $n > N_\epsilon$. Jetzt können wir die Dreiecksungleichung anwenden:

$$|f(x) - f(p)| \leq \underbrace{|f(x) - f_n(x)|}_{< \epsilon} + \underbrace{|f_n(x) - f_n(p)|}_{< \epsilon} + \underbrace{|f_n(p) - f(p)|}_{< \epsilon} < 3\epsilon. \quad (1.25)$$

Der zweite Term auf der rechten Seite verschwindet für $\epsilon \rightarrow 0$, weil f_n eine stetige Funktion ist, und die beiden anderen Terme verschwinden wegen der gleichmäßigen Konvergenz. Somit muss auch f eine stetige Funktion sein. $C[a, b]$ ist also ein Banachraum bezüglich der Maximumsnorm.

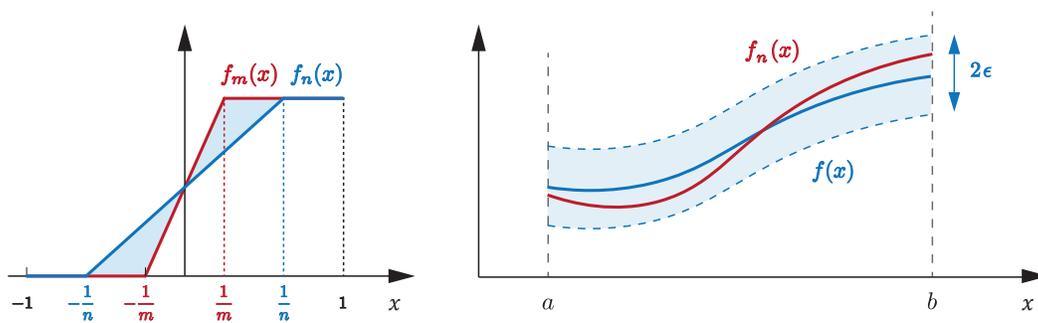


FIG. 1.2: Links: $C[a, b]$ ist kein Banachraum bezüglich $\|\cdot\|_2$. Rechts: gleichmäßige Konvergenz

Das ist das typische Schema, nach dem Vollständigkeitsbeweise verlaufen: sei $f_n \in \mathbb{V}$ und (f_n) irgendeine Cauchyfolge, die gegen ein Element f konvergiert, also $\lim_{n \rightarrow \infty} \|f_n - f\| = 0$. Wenn man dann aus den Eigenschaften des Raumes \mathbb{V} (z.B. Stetigkeit) und der Norm (z.B. Maximumsnorm) zeigen kann, dass auch $f \in \mathbb{V}$, hat man die Vollständigkeit bewiesen.

Was ist dann mit der Folge (1.20), die ja auch aus stetigen Funktionen besteht? Sie ist keine Cauchyfolge mehr bezüglich der Maximumsnorm! Aus Fig. 1.2 sieht man, dass der Maximalwert der Differenz immer bei $x = 1/m$ liegt, daher gilt

$$\|f_n - f_m\|_\infty \stackrel{n \leq m}{=} 1 - \frac{1}{2} \left(1 + \frac{n}{m}\right). \quad (1.26)$$

Die rechte Seite verschwindet aber nicht zwangsläufig für alle $n, m \rightarrow \infty$: nehmen wir z.B. $m = 2n$, wird die rechte Seite zu $1/4$, auch für beliebig große n und m , d.h. die Norm wird nicht beliebig klein.

Ein Raum lässt sich **vervollständigen**, indem man die Grenzwerte aller möglichen Cauchyfolgen hinzunimmt. Die Vervollständigung ist der **Abschluss** $\bar{X} = Y$. Ist $\bar{X} = Y$, dann liegt X **dicht** in Y . Die Elemente aus Y lassen sich dann beliebig gut durch jene aus X approximieren. Z.B. liegen die rationalen Zahlen dicht in den reellen Zahlen, also $\bar{\mathbb{Q}} = \mathbb{R}$. Bezüglich der L^2 -Norm $\|\cdot\|_2$ gilt $\overline{C[a, b]} = L^2[a, b]$, d.h. jede quadratintegrale Funktion (im Sinne von Äquivalenzklassen mit derselben Norm) kann beliebig gut durch stetige Funktionen approximiert werden, z.B. geht aus den stetigen Folgen (1.20) die unstetige Stufenfunktion hervor.

Im Allgemeinen ist der Abschluss (*closure*) $\bar{X} = Y$ einer Menge nicht dasselbe ist wie die Vervollständigung (*completion*) eines metrischen Raums. Während die Vervollständigung eine Metrik voraussetzt, benötigt der Abschluss lediglich eine Topologie, eine schwächere Bedingung für "Nähe". Außerdem bezieht sich der Abschluss einer Menge immer auf einen sie umgebenden Raum. Eine vollständige Teilmenge eines metrischen Raums Y ist aber immer auch abgeschlossen in Y , insbesondere gilt: ist Y ein vollständiger metrischer Raum, ist der *Abschluss der Teilmenge* X gleichbedeutend mit der *Vervollständigung des metrischen Unterraums* X . In diesem Sinne ist der Abschluss $\bar{X} = Y$ gleichzeitig die Vervollständigung des metrischen Raums.

Die Vollständigkeit eines Raumes hängt immer von der gewählten Norm bzw. Metrik ab. Verschiedene Normen induzieren auch verschiedene Konvergenzbegriffe:

- Die Maximumsnorm $\|\cdot\|_\infty$ entspricht der Definition der **gleichmäßigen Konvergenz**.
- Die L^p -Norm $\|\cdot\|_p$ für $p < \infty$ beschreibt die **Konvergenz im Mittel**, die auch **Normkonvergenz** oder **starke Konvergenz** genannt wird.
- Die **punktweise Konvergenz** $\lim_{n \rightarrow \infty} f_n(x) = f(x)$ für alle x (im Definitionsbereich) wird durch keine Norm erzeugt. Insbesondere macht sie für L^p -Funktionen auch wenig Sinn, da man den "Funktionswert" an einer Stelle ja beliebig abändern kann, ohne dass man es im Sinne des $L^p(a, b)$ mit einer anderen Funktion zu tun bekommt.

1.3 Hilberträume

Skalarprodukt. Besonders interessant wird die Struktur eines normierten Raums, wenn die Norm durch ein Skalarprodukt erzeugt wird. Während die Metrik den Abstand zwischen zwei Elementen einer Menge angibt und die Norm die Länge eines Vektors, ist das Skalarprodukt ein Maß für den Winkel zwischen zwei Vektoren. Analog zur Metrik und Norm wollen wir jetzt auch eine abstrakte Definition des Skalarprodukts einführen. Sie hat eine direkte Bedeutung in der Quantenmechanik, wo $|\langle f, g \rangle|^2$ der Übergangswahrscheinlichkeit zwischen zwei Quantenzuständen f und g entspricht.

DEFINITION: \mathbb{V} sei ein Vektorraum über $\mathbb{K} = \mathbb{R}$ oder \mathbb{C} . Ein **Skalarprodukt** auf \mathbb{V} ist eine Abbildung, die jedem geordneten Paar von Elementen (x, y) aus \mathbb{V} eine Größe $\langle x, y \rangle \in \mathbb{K}$ zuordnet und folgende Eigenschaften hat ($\forall x, y, z \in \mathbb{V}, \forall \lambda \in \mathbb{K}$):

- (i) $\langle x, y \rangle = \langle y, x \rangle^*$ (Hermitizität)
- (ii) $\langle x, \lambda y \rangle = \lambda \langle x, y \rangle$
- (iii) $\langle x, y + z \rangle = \langle x, y \rangle + \langle x, z \rangle$ } (Linearität im 2. Argument)
- (iv) $\langle x, x \rangle \geq 0$ (Nichtnegativität)
- (v) $\langle x, x \rangle = 0 \Leftrightarrow x = 0$ (Eindeutigkeit)

Ein Vektorraum mit einem Skalarprodukt wird **Innenproduktraum**, **Prä-Hilbertraum** oder **unitärer Raum** genannt.

Aus (i) und (ii) folgt $\langle \lambda x, y \rangle = \lambda^* \langle x, y \rangle$ und aus (i) und (iii) $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$. Das so definierte Skalarprodukt ist daher linear im zweiten Argument und **antilinear** im ersten Argument. Eine solche Abbildung heißt **Sesquilinearform**, bzw. mit der Einschränkung (i) auf die Hermitizität **hermitesche Form**.

Ist auf einem Vektorraum \mathbb{V} ein Skalarprodukt $\langle \cdot, \cdot \rangle$ definiert, dann induziert es eine Norm und eine Metrik durch

$$\|x\| := \sqrt{\langle x, x \rangle}, \quad d(x, y) = \|x - y\| = \sqrt{\langle x - y, x - y \rangle}. \quad (1.27)$$

Umgekehrt gilt das nicht zwangsläufig, d.h. die Existenz eines Skalarprodukts ist eine stärkere Forderung als die Existenz einer Norm. Man kann aber folgendes zeigen:

SATZ (Jordan-von Neumann): Zu einer Norm gibt es genau dann ein Skalarprodukt, wenn die Norm die **Parallelogrammgleichung**

$$\frac{1}{2} (\|x + y\|^2 + \|x - y\|^2) = \|x\|^2 + \|y\|^2 \quad (1.28)$$

erfüllt. Das Skalarprodukt lässt sich dann mittels **Polarisationsidentität** realisieren:

$$\langle x, y \rangle = \frac{1}{4} (\|x + y\|^2 - \|x - y\|^2) + \frac{i}{4} (\|ix + y\|^2 - \|ix - y\|^2). \quad (1.29)$$

Einige Beispiele:

- Die in (1.15) und (1.17) angeführten p-Normen im \mathbb{R}^n bzw. \mathbb{C}^n , im $C[a, b]$ und im Hilbertschen Folgenraum ℓ^2 sind als Normen zwar alle gleichberechtigt, doch nur die **2-Norm** $\|\cdot\|_2$ erfüllt die Parallelogrammgleichung und erlaubt damit eine Erweiterung zum Skalarprodukt:

$$\langle x, y \rangle = \sum_{i=1}^n x_i^* y_i \quad \forall x, y \in \mathbb{R}^n \text{ bzw. } \mathbb{C}^n, \quad (1.30)$$

$$\langle x, y \rangle = \sum_{n=1}^{\infty} x_n^* y_n, \quad \forall x, y \in \ell^2, \quad (1.31)$$

$$\langle f, g \rangle = \int_a^b dx f(x)^* g(x) \quad \forall f, g \in C[a, b]. \quad (1.32)$$

Beispielsweise ist für die Vektoren $x = (1, 1, 0, \dots, 0)$ und $y = (1, -1, 0, \dots, 0) \in \mathbb{R}^n$ die Parallelogrammgleichung für $p \neq 2$ verletzt:

$$\frac{1}{2} (\|x + y\|^2 + \|x - y\|^2) = 4, \quad \|x\|^2 + \|y\|^2 = 2 \cdot 2^{2/p}. \quad (1.33)$$

Im Fall des ℓ^2 ist die Summe in (1.31) wegen

$$\sum_{n=1}^{\infty} |x_n^* y_n| = \sum_{n=1}^{\infty} |x_n| |y_n| \leq \sum_{n=1}^{\infty} \frac{1}{2} (|x_n|^2 + |y_n|^2) = \frac{1}{2} (\|x\|^2 + \|y\|^2) < \infty \quad (1.34)$$

konvergent und erlaubt somit die Definition von $\langle x, y \rangle$ als Skalarprodukt, woraus man dessen Eigenschaften nachprüfen kann. Im Fall der stetigen Funktionen ist wegen $f, g \in C[a, b]$ auch $f(x)^* g(x)$ stetig, sodass das Integral (1.32) existiert.

- Auf dem Matrizenraum $\mathbb{C}^{m \times n}$ der komplexen $(m \times n)$ -Matrizen (analog für $\mathbb{R}^{m \times n}$) ist das **Frobenius-Skalarprodukt** durch die Spur definiert:

$$\langle A, B \rangle = \text{tr}(A^\dagger B) = \sum_{i=1}^m \sum_{j=1}^n a_{ij}^* b_{ij} \quad \forall A, B \in \mathbb{C}^{m \times n}. \quad (1.35)$$

Ein Skalarprodukt erfüllt außerdem die **Cauchy-Schwarz-Ungleichung**

$$|\langle x, y \rangle| \leq \|x\| \|y\|, \quad (1.36)$$

aus der man die Dreiecksungleichung $\|x + y\| \leq \|x\| + \|y\|$ für die Norm beweisen kann. Die Cauchy-Schwarz-Ungleichung garantiert, dass der Betrag der Größe

$$\cos \varphi := \frac{\langle x, y \rangle}{\|x\| \|y\|} \quad (1.37)$$

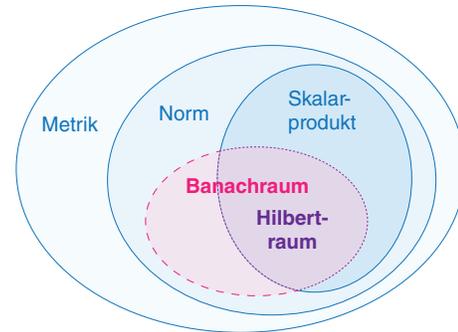
stets kleiner oder gleich 1 ist, sodass der Winkel φ zwischen zwei Vektoren für beliebige Innenprodukträume wohldefiniert ist. Im \mathbb{R}^3 lässt sich die Aussage der Ungleichung in Form einer Gleichung präzisieren:

$$|\mathbf{x} \cdot \mathbf{y}|^2 + |\mathbf{x} \times \mathbf{y}|^2 = |\mathbf{x}|^2 |\mathbf{y}|^2 \quad \forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^3. \quad (1.38)$$

Hilbertraum. Wenden wir das Kriterium der Vollständigkeit auf Innenprodukträume an, gelangen wir zum Begriff des Hilbertraums, der auf den deutschen Mathematiker [David Hilbert](#) (1862–1943) zurückgeht:

DEFINITION: Ein **Hilbertraum** ist ein Vektorraum mit einem Skalarprodukt, der bezüglich der Metrik $d(x, y) = \sqrt{\langle x - y, x - y \rangle}$ vollständig ist.

Anders gesagt: Ein Hilbertraum ist ein Banachraum mit einem Skalarprodukt. Wir werden meist nicht mit allgemeinen Banachräumen arbeiten, sondern nur mit jenen, deren Norm aus einem Skalarprodukt hervorgeht – eben den Hilberträumen, die für die Quantenmechanik von so zentraler Bedeutung sind. Die Verhältnisse sind in der nebenstehenden Abbildung skizziert.



Dazu einige Beispiele:

- Alle endlichdimensionalen Vektorräume sind automatisch Hilberträume, da man (i) auf ihnen immer ein Skalarprodukt definieren kann und (ii) zeigen kann, dass endlichdimensionale Vektorräume bezüglich der dadurch induzierten Metrik immer vollständig sind. Für unendlich-dimensionale Vektorräume gilt das nicht mehr.
- Uns werden vor allem unendlich-dimensionale Hilberträume interessieren. Die früher getroffenen Aussagen für Banachräume gelten auch für Hilberträume, sofern aus der Norm ein Skalarprodukt hervorgeht. Das ist für die Norm $\|\cdot\|_2$ der Fall, deswegen sind der \mathbb{R}^n und \mathbb{C}^n , der Hilbertsche Folgenraum ℓ^2 und der Raum $L^2(a, b)$ bezüglich dieser Norm auch Hilberträume. Die entsprechenden Skalarprodukte haben wir bereits in (1.30–1.32) aufgelistet.
- Der Raum $C[a, b]$ der stetigen Funktionen ist bezüglich der Norm $\|\cdot\|_2$ nicht vollständig und damit kein Banachraum (siehe Diskussion um Gl. (1.23)), also auch kein Hilbertraum. Bezüglich der Maximumsnorm $\|\cdot\|_\infty$ ist er zwar vollständig (und damit ein Banachraum), diese erzeugt aber kein Skalarprodukt. Um das zu sehen, genügt die Angabe eines Vektorpaares, das die Parallelogrammgleichung verletzt, z.B. $f(x) = x$ und $g(x) = 1$ in $C[0, 1]$. Dann ist

$$\|f\|_\infty = 1, \quad \|g\|_\infty = 1, \quad \|f + g\|_\infty = 2, \quad \|f - g\|_\infty = 1 \quad (1.39)$$

und infolgedessen

$$\frac{1}{2} (\|f + g\|_\infty^2 + \|f - g\|_\infty^2) \neq \|f\|_\infty^2 + \|g\|_\infty^2. \quad (1.40)$$

1.4 Orthonormalbasen

Orthogonalität. Die Existenz eines Skalarprodukts erlaubt es uns, den Begriff der Orthogonalität zweier Vektoren zu definieren. Ein Skalarprodukt $\langle f, g \rangle$ kann Null sein, wenn beide Faktoren f und g gleich Null sind; in diesem Fall sind f und g zueinander orthogonal:

DEFINITION: Sei \mathbb{V} ein Innenproduktraum. Zwei Vektoren $f, g \in \mathbb{V}$ heißen **orthogonal** ($f \perp g$), wenn ihr Skalarprodukt verschwindet:

$$f \perp g \quad \Leftrightarrow \quad \langle f, g \rangle = 0. \quad (1.41)$$

Für zwei Teilmengen $A, B \subset \mathbb{V}$ definiert man weiter:

- f ist orthogonal auf die Menge A ($f \perp A$), wenn $\langle f, a \rangle = 0 \quad \forall a \in A$.
- Die Mengen A und B sind orthogonal ($A \perp B$), wenn $\langle f, g \rangle = 0 \quad \forall f \in A, g \in B$.
- Das **orthogonale Komplement** A^\perp einer Menge A ist definiert durch

$$A^\perp = \{f \in \mathbb{V} \mid f \perp A\}. \quad (1.42)$$

Natürlich hängt der Begriff ‘Orthogonalität’ von der Wahl des Skalarprodukts ab. Somit lassen sich bereits einige Aussagen treffen:

- Für zwei linear abhängige Vektoren f und g (d.h. $g = \lambda f$) verschwindet das Skalarprodukt wegen $\langle f, g \rangle = \lambda \langle f, f \rangle = \lambda \|f\|^2 \neq 0$ nicht. Das bedeutet im Umkehrschluss: wenn $f \perp g$ gilt, sind f, g auch linear unabhängig.
- Für $A \subset \mathbb{V}$ folgt aus der Linearität des Skalarprodukts, dass A^\perp ein linearer Unterraum von \mathbb{V} ist: für alle $\lambda \in \mathbb{C}$ und $f, g \in A^\perp$ sind die Elemente λf und $f + g$ wieder in A^\perp .
- Das einzige Element, das zu allen anderen Elementen von \mathbb{V} orthogonal ist, ist der Nullvektor: $\{0\}^\perp = \mathbb{V}$ und $\mathbb{V}^\perp = \{0\}$.
- Aus $\|f+g\|^2 = \langle f+g, f+g \rangle = \|f\|^2 + 2\operatorname{Re}\langle f, g \rangle + \|g\|^2$ folgt der **Satz des Pythagoras**: $f \perp g \Rightarrow \|f+g\|^2 = \|f\|^2 + \|g\|^2$.

Einige Beispiele:

- Im Raum \mathbb{C}^n ist das Orthogonalkomplement des Einheitsvektors gegeben durch

$$\left\{ \left(\begin{array}{c} 1 \\ 0 \\ \vdots \\ 0 \end{array} \right) \right\}^\perp = \left\{ \left(\begin{array}{c} 0 \\ x_2 \\ \vdots \\ x_n \end{array} \right) \mid x_2, \dots, x_n \in \mathbb{C} \right\}. \quad (1.43)$$

- Im Raum $L^2(-a, a)$ sind die beiden Funktionen $f(x) = x$ und $g(x) = x^2$ zueinander orthogonal, denn es gilt

$$\langle f, g \rangle = \int_{-a}^a dx f(x)^* g(x) = \int_{-a}^a dx x^3 = 0. \quad (1.44)$$

DEFINITION: Sei \mathcal{H} ein Hilbertraum und $\mathcal{A}, \mathcal{B} \subset \mathcal{H}$ abgeschlossene Untervektorräume. Dann nennt man $\mathcal{H} = \mathcal{A} \oplus \mathcal{B}$ die **direkte Summe** von \mathcal{A} und \mathcal{B} , wenn

$$\mathcal{H} = \{a + b \mid a \in \mathcal{A}, b \in \mathcal{B}\} \quad \text{und} \quad \mathcal{A} \perp \mathcal{B}. \quad (1.45)$$

SATZ (**Projektionssatz**): Sei \mathcal{H} ein Hilbertraum und $\mathcal{M} \subset \mathcal{H}$ ein abgeschlossener Untervektorraum. Dann gilt

$$\mathcal{H} = \mathcal{M} \oplus \mathcal{M}^\perp, \quad (1.46)$$

d.h. jedes $f \in \mathcal{H}$ lässt sich eindeutig als $f = g + h$ schreiben mit $g \in \mathcal{M}$ und $h \in \mathcal{M}^\perp$. Dabei heißt g die **orthogonale Projektion** von f auf \mathcal{M} , und es gilt $\mathcal{M}^{\perp\perp} = \mathcal{M}$.

Beispiele:

- Im Raum $L^2(a, b)$ mit $a < c < b$ gilt $L^2(a, b) = L^2(a, c) \oplus L^2(c, b)$, wenn man letztere als Unterräume von $L^2(a, b)$ auffasst. So ist z.B.

$$g(x) = \begin{cases} f(x) & x \in (a, c), \\ 0 & x \in (c, b), \end{cases} \quad h(x) = \begin{cases} 0 & x \in (a, c), \\ f(x) & x \in (c, b) \end{cases} \quad (1.47)$$

$$\Rightarrow \langle g, h \rangle = \int_a^b dx g(x)^* h(x) = \int_a^c dx f(x)^* 0 + \int_c^b dx 0 f(x) = 0. \quad (1.48)$$

- Definiert man im $L^2(-a, a)$ die Unterräume der geraden und ungeraden Funktionen,

$$\begin{aligned} L_g^2(-a, a) &= \{f \in L^2(-a, a) \mid f(x) = f(-x)\}, \\ L_u^2(-a, a) &= \{f \in L^2(-a, a) \mid f(x) = -f(-x)\}, \end{aligned} \quad (1.49)$$

dann gilt $L^2(-a, a) = L_g^2(-a, a) \oplus L_u^2(-a, a)$. Offenbar sind die beiden Unterräume orthogonal,

$$\begin{aligned} \langle f_g, f_u \rangle &= \int_{-a}^0 dx f_g(x)^* f_u(x) + \int_0^a dx f_g(x)^* f_u(x) \\ &= \int_0^a dy f_g(-y)^* f_u(-y) + \int_0^a dx f_g(x)^* f_u(x) \\ &= - \int_0^a dy f_g(y)^* f_u(y) + \int_0^a dx f_g(x)^* f_u(x) = 0, \end{aligned} \quad (1.50)$$

und jedes $f \in L^2(-a, a)$ lässt sich als $f = f_g + f_u$ mit $f_g \in L_g^2(-a, a)$, $f_u \in L_u^2(-a, a)$ schreiben:

$$f_g(x) := \frac{1}{2} \{f(x) + f(-x)\}, \quad f_u(x) := \frac{1}{2} \{f(x) - f(-x)\}. \quad (1.51)$$

Orthonormalbasis. Vektoren im \mathbb{R}^3 kann man auf orthonormalen Basissystemen aufspannen und somit durch ihre Koeffizienten ausdrücken. Gilt dasselbe auch für Funktionen? Wir haben nun das nötige Rüstzeug, um diese Frage zu beantworten.

DEFINITION: Sei \mathcal{H} ein Hilbertraum. Eine Menge $\mathcal{S} = \{\varphi_i \mid i \in \mathcal{I}\} \subset \mathcal{H}$ mit einer endlichen oder (abzählbar oder überabzählbar) unendlichen Indexmenge \mathcal{I} heißt **Orthonormalsystem (ONS)**, wenn ihre Elemente paarweise orthogonal und auf Eins normiert sind:

$$\langle \varphi_i, \varphi_j \rangle = \delta_{ij} \quad \forall i, j \in \mathcal{I}. \quad (1.52)$$

Ein Orthonormalsystem \mathcal{S} heißt vollständiges Orthonormalsystem bzw. **Orthonormalbasis (ONB)**, wenn es *maximal* ist, d.h. wenn es nicht möglich ist, ein Element zur Menge \mathcal{S} hinzuzufügen und trotzdem die obige Eigenschaft zu erfüllen – oder anders gesagt: wenn der einzige Vektor $h \in \mathcal{H}$, der senkrecht auf allen φ_i steht, der Nullvektor ist: $\mathcal{S}^\perp = \{0\}$.

Beispiele:

- Die kartesischen Einheitsvektoren $\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ bilden eine mögliche ONB im Raum \mathbb{C}^3 .
- Die trigonometrischen Funktionen

$$\left\{ \frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{\pi}} \cos nx, \frac{1}{\sqrt{\pi}} \sin nx \right\}, \quad n \in \mathbb{N}_+ \quad (1.53)$$

die wir in Abschnitt 2 eingehend behandeln werden, bilden eine mögliche ONB im Raum $L^2(-\pi, \pi)$.

Um festzustellen, ob eine gegebene Menge \mathcal{S} ein ONS ist, muss man nur (1.52) überprüfen, während die Frage, ob ein ONS auch eine ONB ist, im Allgemeinen schwieriger zu beantworten ist. Betrachten wir als Beispiel den Hilbertschen Folgenraum ℓ^2 aus (1.16) und (1.31). Sei $\varphi_i = (0, 0, \dots, 0, 1, 0, \dots)$ mit $i \in \mathbb{N}$ jener Vektor, der nur an der i -ten Stelle eine 1 besitzt. Dann ist klar, dass die Menge $\mathcal{S} = \{\varphi_i \mid i \in \mathbb{N}\}$ ein ONS definiert. Um festzustellen, ob es sich auch um eine ONB handelt, überprüfen wir die Bedingung $\mathcal{S}^\perp = \{0\}$. Sei

$$x = (x_n)_{n=1}^\infty \in \mathcal{S}^\perp \quad \Rightarrow \quad 0 = \langle \varphi_i, x \rangle = \sum_{n=1}^\infty (\varphi_i)_n^* x_n = x_i \quad \forall i \in \mathbb{N}. \quad (1.54)$$

Das i -te Element der Folge x verschwindet also, doch weil i beliebig ist, muss jedes Element verschwinden und somit ist die Folge die Nullfolge $x = (0, 0, 0, \dots)$. \mathcal{S} ist also eine ONB.

Zunächst ist nicht einmal klar, ob jeder Hilbertraum überhaupt eine ONB besitzt. Die Existenz einer solchen Basis wird aber durch den folgenden Satz gesichert (dessen Beweis das **Lemma von Zorn** aus der Mengenlehre benutzt):

SATZ: Jeder Hilbertraum besitzt eine Orthonormalbasis.

Komponentendarstellung. Um eine Funktion $f \in \mathcal{H}$ in ihre Komponenten zu entwickeln, würden wir sie durch

$$f = \sum_{i \in \mathcal{I}} c_i \varphi_i, \quad c_i = \langle \varphi_i, f \rangle \quad \forall i \in \mathcal{I} \quad (1.55)$$

darstellen. Die Indexmenge $\mathcal{I} = \{i\}$ darf aber nicht nur endlich viele, sondern auch (abzählbar oder überabzählbar) unendlich viele Elemente i enthalten. Hat die Summe dann überhaupt eine Bedeutung? Der folgende Satz bejaht diese Frage:

SATZ: Sei \mathcal{H} ein Hilbertraum und $\mathcal{S} = \{\varphi_i \mid i \in \mathcal{I}\} \subset \mathcal{H}$ ein ONS. Dann gilt $\forall f \in \mathcal{H}$ die **Besselsche Ungleichung**

$$\sum_{i \in \mathcal{I}} |\langle \varphi_i, f \rangle|^2 \leq \|f\|^2. \quad (1.56)$$

Zum Beweis betrachten wir zunächst ein ONS $\mathcal{S} = \{\varphi_i\}_{i=1}^{\infty} \subset \mathcal{H}$ mit einer abzählbaren Indexmenge \mathcal{I} . Für einen gegebenen Vektor $f \in \mathcal{H}$ können wir die komplexen Zahlen $c_i = \langle \varphi_i, f \rangle$ bilden und die folgenden longitudinalen und transversalen Projektionen für beliebige $N \in \mathbb{N}$ konstruieren:

$$f_{\parallel} := \sum_{i=1}^N c_i \varphi_i, \quad f_{\perp} := f - f_{\parallel}. \quad (1.57)$$

Dann gilt $\langle \varphi_i, f_{\parallel} \rangle = c_i$, $\langle \varphi_i, f_{\perp} \rangle = \langle \varphi_i, f \rangle - \langle \varphi_i, f_{\parallel} \rangle = 0$ und $\langle f_{\parallel}, f_{\perp} \rangle = 0$. Aus $f = f_{\parallel} + f_{\perp}$ und $f_{\parallel} \perp f_{\perp}$ folgt $\|f\|^2 = \|f_{\parallel}\|^2 + \|f_{\perp}\|^2$ (Satz des Pythagoras) und somit

$$\|f_{\parallel}\|^2 = \sum_{i=1}^N |c_i|^2 \leq \|f\|^2, \quad (1.58)$$

also die Besselsche Ungleichung für eine endliche Indexmenge. (Damit lässt sich übrigens auch die Cauchy-Schwarz-Ungleichung (1.36) beweisen, indem man $N = 1$ und $\varphi_1 = g/\|g\|$ setzt.) Weil N aber beliebig war, gilt die Ungleichung auch für den Grenzwert $N \rightarrow \infty$, d.h. der Vektor $\lim_{N \rightarrow \infty} f_{\parallel}$ hat eine endliche Norm und ist somit konvergent.

Die Besselsche Ungleichung ist nicht nur für abzählbare, sondern auch für überabzählbare Indexmengen erfüllt: Da laut (1.58) $\|f\|^2$ eine endliche obere Schranke für beliebige endliche Summen der $|c_i|^2$ darstellt, kann es für jedes $n \in \mathbb{N}$ nur endlich viele (nämlich höchstens $n\|f\|^2$) Terme $|c_i|^2 \geq 1/n$ geben. Das ergibt mit $n \rightarrow \infty$ eine Abzählung aller $|c_i|^2$, d.h. selbst bei einer überabzählbaren Indexmenge können nur höchstens abzählbar unendlich viele Koeffizienten c_i in der Summe (1.56) ungleich Null sein.

Die Besselsche Ungleichung sagt zwar, dass der Vektor $f_{\parallel} = \sum_{i \in \mathcal{I}} c_i \varphi_i$ als Summe einer konvergenten Reihe existiert, aber nicht, ob er tatsächlich gegen f konvergiert – wenn das ONS nicht alle Richtungen des Vektorraums berücksichtigt, kann es ja noch immer eine nichtverschwindende Komponente $f_{\perp} \neq 0$ geben. Nur bei einer ONB liefert die Entwicklungssumme auch alle Informationen über f :

SATZ: Sei \mathcal{H} ein Hilbertraum und $\mathcal{S} = \{\varphi_i \mid i \in \mathcal{I}\} \subset \mathcal{H}$ eine ONB. Dann gilt $\forall f \in \mathcal{H}$ die **Parsevalsche Gleichung**

$$\sum_{i \in \mathcal{I}} |\langle \varphi_i, f \rangle|^2 = \|f\|^2 \quad (1.59)$$

und jedes $f \in \mathcal{H}$ kann nach den Basisvektoren entwickelt werden:

$$f = \sum_{i \in \mathcal{I}} c_i \varphi_i, \quad c_i = \langle \varphi_i, f \rangle. \quad (1.60)$$

Bra-Ket-Notation. In der Quantenmechanik verwendet man oft die **Bra-Ket-Notation**, in der man statt f das Symbol $|f\rangle$ (“Ket”) benutzt und Skalarprodukte durch $\langle f, g\rangle = \langle f|g\rangle$ angibt. Dann schreibt sich die Koeffizientenentwicklung des Vektors $|f\rangle$ als

$$|f\rangle = \sum_{i \in \mathcal{I}} |\varphi_i\rangle \langle \varphi_i|f\rangle, \quad (1.61)$$

was sich besonders leicht in Form der beiden folgenden Beziehungen merken lässt:

$$\text{Orthonormalitätsrelation:} \quad \langle \varphi_i|\varphi_j\rangle = \delta_{ij}, \quad (1.62)$$

$$\text{Vollständigkeitsrelation:} \quad \sum_{i \in \mathcal{I}} |\varphi_i\rangle \langle \varphi_i| = \mathbf{1}. \quad (1.63)$$

Die Entwicklung in Koeffizienten entspricht dann einem “Einschieben der Eins” wie in (1.61) bzw.

$$\langle f|g\rangle = \sum_{i \in \mathcal{I}} \langle f|\varphi_i\rangle \langle \varphi_i|g\rangle, \quad \|f\|^2 = \langle f|f\rangle = \sum_{i \in \mathcal{I}} \langle f|\varphi_i\rangle \langle \varphi_i|f\rangle = \sum_{i \in \mathcal{I}} |\langle \varphi_i|f\rangle|^2. \quad (1.64)$$

Die Bra-Ket-Notation ist fürs praktische Rechnen zwar sehr bequem, kann aber insbesondere bei Matricelementen der Form $\langle f|A|g\rangle$ zu Verwirrungen bezüglich des Definitionsbereichs des Operators A führen, deshalb wollen wir hier lieber darauf verzichten.

Orthonormalbasen. Im Folgenden sammeln wir einige äquivalente Kriterien zur Bestimmung einer ONB aus einem gegebenen ONS. Ein Orthonormalsystem $\mathcal{S} = \{\varphi_i | i \in \mathcal{I}\} \subset \mathcal{H}$ ist eine Orthonormalbasis genau dann, wenn:

- (1) es maximal ist: $\mathcal{S}^\perp = \{0\}$, d.h. ist ein h orthogonal zu allen φ_i , dann gilt $h = 0$;
- (2) jedes $f \in \mathcal{H}$ durch $f = \sum_{i \in \mathcal{I}} \varphi_i \langle \varphi_i, f\rangle = \sum_{i \in \mathcal{I}} c_i \varphi_i$ eindeutig darstellbar ist;
- (3) die Vollständigkeitsrelation gilt: $\sum_{i \in \mathcal{I}} |\varphi_i\rangle \langle \varphi_i| = \mathbf{1}$;
- (4) die Parsevalsche Gleichung $\forall f \in \mathcal{H}$ gilt: $\|f\|^2 = \sum_{i \in \mathcal{I}} |\langle \varphi_i, f\rangle|^2 = \sum_{i \in \mathcal{I}} |c_i|^2$;
- (5) die 2. Parsevalsche Gleichung $\forall f, g \in \mathcal{H}$ gilt: $\langle f, g\rangle = \sum_{i \in \mathcal{I}} \langle f, \varphi_i\rangle \langle \varphi_i, g\rangle = \sum_{i \in \mathcal{I}} c_i^* d_i$.
- (6) die lineare Hülle von \mathcal{S} dicht in \mathcal{H} liegt.

Jedes dieser Kriterien ist äquivalent, d.h. man könnte statt (1) genausogut irgendein anderes zur Definition einer ONB heranziehen. Beispielsweise gilt:

- (1 \rightarrow 2): Der Vektor $h = f - \sum_i \varphi_i \langle \varphi_i, f\rangle$ ist orthogonal zu allen φ_j , daher gilt $h = 0$. (2 \rightarrow 1): Kann man jedes $f = \sum_i \varphi_i \langle \varphi_i, f\rangle$ darstellen, so auch ein h mit $\langle \varphi_i, h\rangle = 0$, die Summe ist also $h = 0$.
- (2 \rightarrow 3): Da $f = \mathbf{1}f = \sum_i \varphi_i \langle \varphi_i, f\rangle \forall f \in \mathcal{H}$ gilt, ist die Vollständigkeitsrelation erfüllt.
- (2 \rightarrow 4) und (2 \rightarrow 5): Einsetzen. (5 \rightarrow 4): Setze $g = f$.
- (4 \rightarrow 1): Gilt die Parsevalsche Gleichung und ist ein h orthogonal zu allen φ_i , dann sind alle Terme in der Summe Null $\Rightarrow \|h\| = 0 \Rightarrow h = 0$.

Separable Hilberträume. Bisher haben wir mit allgemeinen Indexmengen \mathcal{I} gearbeitet. Im Folgenden werden uns aber vor allem **separable Hilberträume** mit einer maximal abzählbar unendlichen Basis interessieren. Dazu definieren wir:

DEFINITION: Ein metrischer Raum X heißt **separabel**, wenn es eine *abzählbare* Teilmenge M gibt, die dicht in X liegt (d.h. $\overline{M} = X$).

In anderen Worten: separable Räume sind abgeschlossene Hüllen (jedes Element in X ist der Grenzwert einer Folge von Elementen in M) von abzählbaren Mengen.

SATZ: Ein Hilbertraum \mathcal{H} ist genau dann separabel, wenn er eine ONB aus abzählbar (endlich oder abzählbar unendlich) vielen Elementen besitzt. Außerdem:

- Ist \mathcal{H} separabel, dann ist jede ONB abzählbar und ihre lineare Hülle liegt dicht in \mathcal{H} .
- Besteht die ONB aus endlich vielen Elementen, ist \mathcal{H} isomorph zu \mathbb{C}^n bzw. \mathbb{R}^n .
- Besteht die ONB aus abzählbar unendlich vielen Elementen, ist \mathcal{H} isomorph zum Hilbertschen Folgenraum ℓ^2 (**Satz von Fischer-Riesz**).

Die letzte Aussage illustriert die Äquivalenz zwischen verschiedenen Zugängen zur Quantenmechanik, der ‘Wellenmechanik’ nach Schrödinger und ‘Matrizenmechanik’ nach Heisenberg. Im Hilbertschen Folgenraum ℓ^2 (vgl. (1.16) und (1.31)) mit der ONB $\mathcal{S} = \{\varphi_i \mid i \in \mathbb{N}\}$ nehmen die Komponentendarstellung und Parsevalsche Gleichung folgende Form an:

$$x = \sum_{i=1}^{\infty} \varphi_i \langle \varphi_i, x \rangle = \sum_{i=1}^{\infty} \varphi_i x_i = (x_1, x_2, x_3, \dots), \quad \|x\|^2 = \sum_{i=1}^{\infty} |\langle \varphi_i, x \rangle|^2 = \sum_{i=1}^{\infty} |x_i|^2. \quad (1.65)$$

Wir können dann das Rechnen mit Funktionen auf das Rechnen mit den unendlichen Folgen ihrer Koeffizienten zurückführen, d.h. unendlichdimensionalen Vektoren:

$$f \in \mathcal{H} \quad \longleftrightarrow \quad (\langle \varphi_1, f \rangle, \langle \varphi_2, f \rangle, \langle \varphi_3, f \rangle, \dots) \in \ell^2. \quad (1.66)$$

Später werden wir auch Differenzialoperatoren durch unendlichdimensionale Matrizen darstellen, die auf diese Folgen wirken. Es folgen einige Beispiele für separable Räume:

- Die Räume \mathbb{R}^n und \mathbb{C}^n ($n \in \mathbb{N}$) sind separabel, da die rationalen Zahlen \mathbb{Q}^n abzählbar sind und dicht in \mathbb{R}^n liegen.
- Die Folgenräume ℓ^p für $1 \leq p < \infty$ sind separabel (aber nur ℓ^2 ist ein Hilbertraum). Der Banachraum ℓ^∞ der beschränkten Folgen ist nicht separabel.
- Die Lebesgue-Räume L^p sind für $1 \leq p < \infty$ separabel, sofern der zugrundeliegende Maßraum ebenfalls separabel ist (wieder ist nur L^2 ein Hilbertraum).
- Der Raum $C[a, b]$ der stetigen Funktionen mit der Maximumsnorm $\|\cdot\|_\infty$ ist separabel, da die Menge der Polynomfunktionen $P[a, b] = \{p \mid p(x) = \sum_{n=0}^N a_n x^n, N < \infty\}$ dicht in $C[a, b]$ liegt (**Weierstraß’scher Approximationssatz**). Anders gesagt: jede stetige Funktion auf dem Intervall $[a, b]$ kann beliebig genau durch ein Polynom genähert werden, wobei die Konvergenz bezüglich $\|\cdot\|_\infty$ der gleichmäßigen Konvergenz entspricht.

Gram-Schmidt-Verfahren. Wir haben gesehen, dass jeder Hilbertraum eine ONB besitzt, und mit der Einschränkung auf separable Hilberträume ist diese ONB auch abzählbar. Wie konstruiert man dann eine solche ONB? Eine möglicher Algorithmus ist das **Gram-Schmidt-Verfahren**. Dazu starten wir von einer beliebigen Menge linear unabhängiger Vektoren $\{b_1, b_2, b_3, \dots\}$ eines Hilbertraums \mathcal{H} .

- **1. Schritt:** wähle den ersten Vektor

$$\varphi_1 = \frac{b_1}{\|b_1\|}. \quad (1.67)$$

- Der zweite Vektor soll aus b_2 gebildet werden und zu φ_1 orthogonal sein. Dazu muss man von b_2 den Anteil entfernen, der in die Richtung von φ_1 zeigt, was durch

$$\tilde{b}_2 = b_2 - \varphi_1 \langle \varphi_1, b_2 \rangle, \quad \varphi_2 = \frac{\tilde{b}_2}{\|\tilde{b}_2\|} \quad (1.68)$$

bewerkstelligt wird. Es gilt $\langle \varphi_1, \tilde{b}_2 \rangle = 0$, und wir haben den Vektor außerdem normiert. Vom dritten Vektor b_3 müssen wir die Anteile in die beiden Richtungen φ_1 und φ_2 entfernen, usw. Das allgemeine Verfahren für den **k -ten Schritt** ($k \geq 2$) lautet somit:

$$\tilde{b}_k = b_k - \sum_{i=1}^{k-1} \varphi_i \langle \varphi_i, b_k \rangle, \quad \varphi_k = \frac{\tilde{b}_k}{\|\tilde{b}_k\|}. \quad (1.69)$$

Auf diese Weise erhält man ein ONS $\mathcal{S} = \{\varphi_1, \varphi_2, \varphi_3, \dots\}$, welches den gleichen Raum aufspannt wie die ursprünglichen Vektoren $\{b_1, b_2, b_3, \dots\}$. Ist die Anzahl der Startvektoren groß genug, sodass ihre lineare Hülle dicht in \mathcal{H} liegt, ist \mathcal{S} auch eine ONB.

Als Beispiel betrachten wir die Menge der Monome $\{b_n(x) = x^{n-1} \mid n \in \mathbb{N}\}$ auf dem Intervall $[-1, 1]$ mit dem Skalarprodukt

$$\langle f, g \rangle = \frac{1}{2} \int_{-1}^1 dx f(x)g(x). \quad (1.70)$$

Dann gilt

$$\langle b_1, b_1 \rangle = 1, \quad \langle b_2, b_2 \rangle = \langle b_1, b_3 \rangle = \frac{1}{3}, \quad \langle b_3, b_3 \rangle = \frac{1}{5}, \quad \langle b_1, b_2 \rangle = \langle b_2, b_3 \rangle = 0. \quad (1.71)$$

Die orthogonalen Basisvektoren sind (bis auf die Normierung) die **Legendre-Polynome**, die wir später im Detail behandeln werden:

$$\begin{aligned} \varphi_1(x) &= 1, \\ \tilde{b}_2(x) &= x, \quad \|\tilde{b}_2\|^2 = \frac{1}{3}, \quad \varphi_2(x) = \sqrt{3}x, \\ \tilde{b}_3(x) &= x^2 - \frac{1}{3}, \quad \|\tilde{b}_3\|^2 = \frac{4}{45}, \quad \varphi_3(x) = \frac{\sqrt{5}}{2}(3x^2 - 1), \quad \dots \end{aligned} \quad (1.72)$$

Eine abzählbare ONB eines separablen Hilbertraums ist ein Spezialfall einer **Schauder-Basis**. Diese ist für einen normierten Raum X als Folge $(b_i)_{i \in \mathbb{N}}$ aus X definiert, sodass jeder Vektor $f \in X$ eine eindeutige Darstellung als konvergente Reihe $\sum_{n=1}^{\infty} \alpha_n b_n$ mit einer Folge $(\alpha_i)_{i \in \mathbb{N}}$ aus \mathbb{K} hat. Sie unterscheidet sich von der **Hamel-Basis**, von der verlangt wird, dass sich jeder Vektor als *endliche* Linearkombination der Basiselemente darstellen lässt. Für endlichdimensionale Vektorräume sind die beiden Konzepte äquivalent, für unendlichdimensionale allerdings nicht: Damit sich jeder Vektor als endliche Linearkombination darstellen lässt, muss die Hamel-Basis überabzählbar unendlich sein, was ihre konkrete Realisierung erschwert. Schauderbasen existieren dagegen nur für separable Hilberträume, können aber durch das Gram-Schmidt-Verfahren einfach konstruiert werden.

Kapitel 2

Fourierreihen

Als konkrete Anwendung des bisher Gesagten betrachten wir die **Fourieranalyse** oder Harmonische Analyse, ein zentrales Thema in den Natur- und Ingenieurwissenschaften. Die Theorie der Fourierreihen behandelt **periodische Funktionen**, die in der Natur häufig auftreten: von Schwingungen und Wellen zu Licht und Ton, der Erdrotation oder dem Herzschlag. Mit Hilfe der Fourierreihe lassen sich periodische Funktionen nach ihren Teilfrequenzen zerlegen. Typische Beispiele in der Physik sind die Normalmoden in der klassischen Mechanik, die Frequenzanalyse von elektromagnetischen Wellen oder quantenmechanischen Wellenpaketen, die Erzeugungs- und Vernichtungsoperatoren in der Quantenfeldtheorie usw.

Die Vorstellung, dass sich die Schwingungen einer Saite aus einer Grundschiwingung und harmonischen Oberschwingungen darstellen lässt, wurde bereits vom Schweizer Mathematiker und Physiker [Daniel Bernoulli](#) (1700–1782) vertreten. Der französische Mathematiker und Physiker [Jean Baptiste Joseph Fourier](#) (1768–1830) führte erstmals aus, dass periodische Temperaturverteilungen durch eine solche Reihe dargestellt werden können. Die mathematischen Grundbegriffe dieser Theorie wurden später vom deutschen Mathematiker [Peter Gustav Lejeune Dirichlet](#) (1805–1859) präzisiert.

2.1 Definition der Fourierreihe

Periodische Funktionen. Im Folgenden interessieren wir uns für Funktionen mit periodischem Verhalten:

DEFINITION: Eine Funktion $f: \mathbb{R} \rightarrow \mathbb{C}$ heißt **periodisch** mit Periode $2L > 0$, wenn

$$f(x + 2L) = f(x) \quad \forall x \in \mathbb{R}. \quad (2.1)$$

Daraus folgt $f(x) = f(x + 2nL) \forall x \in \mathbb{R}, n \in \mathbb{Z}$. Eine periodische Funktion ist also bereits eindeutig definiert, wenn man ihre Funktionswerte nur in einem beliebigen **Periodizitätsintervall** $[a, a + 2L]$ mit $a \in \mathbb{R}$ kennt; deswegen reicht es z.B. aus, f als eine Funktion $f: [-L, L] \rightarrow \mathbb{C}$ zu betrachten. Die Erweiterung auf ganz \mathbb{R} kann dann natürlich Diskontinuitäten in der Funktion bei ungeraden ganzzahligen Vielfachen von L erzeugen.

Bei periodischen Funktionen ist die Integration über eine volle Periode außerdem invariant gegen Verschiebungen des Integrationsintervalls,

$$\int_{a-L}^{a+L} dx f(x) = \int_{a-L}^{-L} dx f(x) + \int_{-L}^L dx f(x) + \int_L^{a+L} dx f(x) = \int_{-L}^L dx f(x), \quad (2.2)$$

da wegen der Periodizität gilt:

$$\int_{a-L}^{-L} dx f(x) = - \int_{-L}^{a-L} dx f(x) = - \int_{-L}^{a-L} dx f(x + 2L) = - \int_L^{a+L} dy f(y). \quad (2.3)$$

Typische Beispiele für periodische Funktionen sind die trigonometrischen Funktionen $\sin nx$, $\cos nx$ bzw. e^{inx} . Im Folgenden werden wir uns einfachheitshalber auf Perioden $2L = 2\pi$ beschränken, denn eine Funktion mit beliebiger Periode $2L$ kann durch die Transformation

$$\tilde{f}(x) = f\left(\frac{\pi}{L}x\right) \quad (2.4)$$

immer zu einer Funktion f mit Periode 2π transformiert werden. Zweckmäßigerweise kann man sich das Intervall $-\pi < x \leq \pi$ als um den Einheitskreis \mathcal{S} mit Kreisumfang 2π gewickelt vorstellen, sodass die Endpunkte $-\pi$ und π miteinander identifiziert werden (was wir im folgenden tun werden).

Fourierreihe. Wir betrachten zunächst den Hilbertraum $L^2(\mathcal{S}) = L^2(-\pi, \pi)$ der quadratintegralen Funktionen auf dem Einheitskreis \mathcal{S} . Dann gilt:

SATZ: Die Menge $E = \{\varphi_n \mid \varphi_n(x) = e^{inx}, n \in \mathbb{Z}\}$ bildet eine ONB im Raum $L^2(\mathcal{S})$ mit dem Skalarprodukt

$$\langle f, g \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} dx f(x)^* g(x). \quad (2.5)$$

Anstatt die Basiselemente mit $1/\sqrt{2\pi}$ zu versehen, haben wir das Skalarprodukt durch die Periode 2π dividiert; das garantiert, dass die L^2 -Norm der konstanten Funktion 1 gleich 1 ist und erfüllt natürlich noch immer alle Kriterien eines Skalarprodukts. Die Orthogonalität lässt sich einfach beweisen:

$$\langle \varphi_n, \varphi_m \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} dx e^{i(m-n)x} = \begin{cases} \frac{e^{i(m-n)x}}{2\pi i(m-n)} \Big|_{-\pi}^{\pi} = 0 & m \neq n \\ 1 & m = n \end{cases} = \delta_{mn}. \quad (2.6)$$

Das Resultat für $m \neq n$ folgt aus der 2π -Periodizität: $e^{i(m-n)x} = e^{i(m-n)(x-2\pi)}$. Der Beweis der Vollständigkeit ist etwas umständlicher [siehe z.B. Gro14]; dazu zeigt man mit Hilfe des Weierstraß'schen Approximationssatzes, dass $\langle \varphi_n, h \rangle = 0 \Rightarrow h = 0$ für $h \in L^2(\mathcal{S})$ gilt. Somit bildet E eine ONB des Hilbertraums $L^2(\mathcal{S})$. Da die φ_n eine abzählbar unendliche Basis bilden, ist $L^2(\mathcal{S})$ außerdem separabel.

DEFINITION: Die **Fourierreihe** einer Funktion $f \in L^2(\mathcal{S})$ ist mit der Orthonormalbasis $E = \{\varphi_n \mid \varphi_n(x) = e^{inx}, n \in \mathbb{Z}\}$ definiert durch

$$FR(f)(x) = \sum_{n=-\infty}^{\infty} c_n \varphi_n(x), \quad c_n = \langle \varphi_n, f \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} dx \varphi_n(x)^* f(x). \quad (2.7)$$

Da E eine ONB ist, folgt nach den Überlegungen der letzten Abschnitte, dass sich jede Funktion $f \in L^2(\mathcal{S})$ durch die eine Fourierreihe ausdrücken lässt – sie ist einfach die Projektion von f auf die ONB der φ_n . Dabei haben wir bewusst $FR(f)$ geschrieben, um die Funktion f von ihrer Fourierreihe zu unterscheiden und später allgemeinere Konvergenzbetrachtungen zu ermöglichen. Da $L^2(\mathcal{S})$ ein Hilbertraum ist, gilt die Parsevalsche Gleichung

$$\|f\|^2 = \sum_{n=-\infty}^{\infty} |c_n|^2 < \infty, \quad (2.8)$$

und somit konvergiert $FR(f)$ gegen f im Sinne der L^2 -Normkonvergenz $\|FR(f) - f\|^2 = 0$, d.h. der **Konvergenz im Mittel**:

$$\langle FR(f) - f, FR(f) - f \rangle = \sum_n |c_n|^2 - \sum_n c_n \underbrace{\langle f, \varphi_n \rangle}_{c_n^*} - \sum_n c_n^* \underbrace{\langle \varphi_n, f \rangle}_{c_n} + \|f\|^2 = 0. \quad (2.9)$$

Als Beispiel betrachten wir die **Rechteckfunktion** in Fig. 2.1:

$$f(x) = \begin{cases} 0 & -\pi < x \leq 0, \\ 1 & 0 < x \leq \pi. \end{cases} \quad (2.10)$$

Hieraus erhält man die Fourierkoeffizienten

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} dx e^{-inx} f(x) = \frac{1}{2\pi} \int_0^{\pi} dx e^{-inx} \stackrel{n \neq 0}{=} -\frac{e^{-inx}}{2\pi in} \Big|_0^{\pi} = \frac{1}{2\pi in} (1 - (-1)^n) = \begin{cases} \frac{1}{2} & n = 0 \\ 0 & n \text{ gerade} \\ \frac{1}{i\pi n} & n \text{ ungerade} \end{cases}$$

deren Absolutbeträge in Fig. 2.1 rechts geplottet sind. Die resultierende Fourierreihe lautet

$$FR(f)(x) = \frac{1}{2} + \frac{1}{i\pi} \sum_{k=0}^{\infty} \frac{1}{2k+1} \left(e^{i(2k+1)x} - e^{-i(2k+1)x} \right) = \frac{1}{2} + \frac{2}{\pi} \sum_{k=0}^{\infty} \frac{\sin(2k+1)x}{2k+1}. \quad (2.11)$$

An den Sprungstellen $x = n\pi$ liefert die Fourierreihe den Wert $1/2$, also den Mittelwert der beiden Limiten 0 und 1 der Rechteckfunktion.

Ein anderes klassisches Beispiel ist die **Sägezahnfunktion** (Fig. 2.1):

$$f(x) = x \quad -\pi < x \leq \pi. \quad (2.12)$$

Die Fourierkoeffizienten lauten

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} dx e^{-inx} x \stackrel{n \neq 0}{=} \frac{1}{2\pi} \frac{1}{(-in)} \left(x e^{-inx} \Big|_{-\pi}^{\pi} - \int_{-\pi}^{\pi} dx e^{-inx} \right) = \frac{i}{n} (-1)^n. \quad (2.13)$$

Das Integral über e^{-inx} über die volle Periode verschwindet, und $e^{-in\pi} = (-1)^n$. Für $n = 0$ ergibt sich $c_0 = 0$. Damit lautet die Fourierreihe

$$FR(f)(x) = \sum_{\substack{n=-\infty \\ n \neq 0}}^{\infty} \frac{i}{n} (-1)^n e^{inx} = \sum_{n=1}^{\infty} i (-1)^n \frac{e^{inx} - e^{-inx}}{n} = 2 \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} \sin nx. \quad (2.14)$$

An den Intervallgrenzen $x = \pm\pi$ liefert die Fourierreihe wieder den Mittelwert der beiden Limiten π und $-\pi$, nämlich Null.

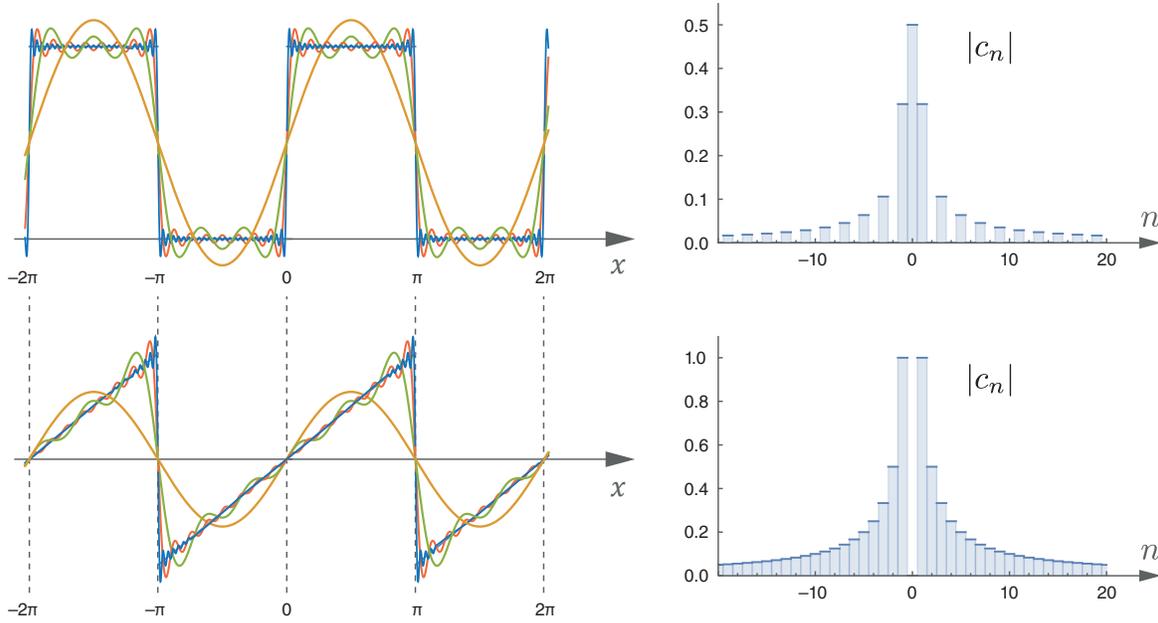


FIG. 2.1: Fourierreihe der Rechteckfunktion (oben) und Sägezahnfunktion (unten) jeweils mit $N = 1, 5, 15$ und 50 Termen. Rechts sind die Absolutbeträge $|c_n|$ der Fourierkoeffizienten aufgetragen, die in beiden Fällen wie $1/|n|$ abfallen.

Hierzu einige Bemerkungen:

■ Die aus den Basisfunktionen φ_n gewonnenen Partialsummen werden **Fourierpolynome** oder **trigonometrische Polynome** genannt. Das Fourierpolynom vom Grad N lautet

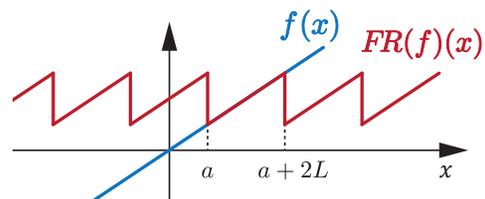
$$p_N(x) = \sum_{n=-N}^N c_n e^{inx}, \quad x \in \mathbb{R}. \tag{2.15}$$

■ Integrale über die trigonometrischen Funktionen e^{inx} , $\sin nx$ und $\cos nx$ für $n \neq 0$ verschwinden, wenn sie über ganzzahlige Vielfache des Periodizitätsintervalls gehen:

$$\int_a^{a+2k\pi} dx e^{inx} = \int_a^{a+2k\pi} dx \sin nx = \int_a^{a+2k\pi} dx \cos nx = 0 \quad (n, k \in \mathbb{N}, n \neq 0). \tag{2.16}$$

Wegen (2.2) kommt es dabei nicht auf die Lage des Integrationsintervalls an, das man beliebig verschieben kann.

■ Die Funktion f muss an sich nicht periodisch sein; die Fourierreihe erzeugt dann eine periodische Fortsetzung dieser Funktion im Periodizitätsintervall (siehe Figur rechts).



Reelle Form der Fourierreihe. Wie wir am Beispiel der Rechteck- und Sägezahnfunktion gesehen haben, ist es oft zweckmäßiger, mit reellen Koeffizienten zu arbeiten, denn die Fourierreihe einer reellen Funktion ist wieder reell. Mit der Euler'schen Formel $e^{ix} = \cos x + i \sin x$ gelangen wir auf die reelle Form der Fourierreihe:

$$\begin{aligned} FR(f)(x) &= \sum_{n=-\infty}^{\infty} c_n e^{inx} = c_0 + \sum_{n=1}^{\infty} (c_n e^{inx} + c_{-n} e^{-inx}) \\ &= c_0 + \sum_{n=1}^{\infty} ((c_n + c_{-n}) \cos nx + i(c_n - c_{-n}) \sin nx) \\ &= \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx). \end{aligned} \quad (2.17)$$

Die Koeffizienten lauten dann

$$\begin{aligned} a_n = c_n + c_{-n} &= \frac{1}{2\pi} \int_{-\pi}^{\pi} dx \underbrace{(e^{-inx} + e^{inx})}_{2 \cos nx} f(x) = \frac{1}{\pi} \int_{-\pi}^{\pi} dx \cos nx f(x) \quad (n \geq 0), \\ b_n = i(c_n - c_{-n}) &= \frac{i}{2\pi} \int_{-\pi}^{\pi} dx \underbrace{(e^{-inx} - e^{inx})}_{-2i \sin nx} f(x) = \frac{1}{\pi} \int_{-\pi}^{\pi} dx \sin nx f(x) \quad (n > 0). \end{aligned} \quad (2.18)$$

Dabei wurde die Normierung von $a_0 = 2c_0$ so gewählt, um eine einheitliche Definition der Koeffizienten a_n zu gewährleisten. Dementsprechend stellt die Menge

$$\left\{ \frac{1}{\sqrt{2}}, \sin nx, \cos nx; \quad n = 1, 2, \dots \right\} \quad (2.19)$$

ebenfalls eine ONB des Hilbertraums $L^2(\mathcal{S})$ dar, allerdings mit dem Skalarprodukt

$$\langle f, g \rangle = \frac{1}{\pi} \int_{-\pi}^{\pi} dx f(x)^* g(x). \quad (2.20)$$

Die Parsevalsche Gleichung ergibt sich dann zu

$$\|f\|^2 = \frac{|a_0|^2}{2} + \sum_{n=1}^{\infty} (|a_n|^2 + |b_n|^2). \quad (2.21)$$

Beispielsweise lauten die entsprechenden Koeffizienten der Rechteckfunktion

$$a_0 = 1, \quad a_{n>0} = 0, \quad b_{n>0} = \begin{cases} 0 & n \text{ gerade} \\ \frac{2}{\pi n} & n \text{ ungerade} \end{cases} \quad (2.22)$$

und jene der Sägezahnfunktion

$$a_n = 0, \quad b_n = \frac{2}{n} (-1)^{n+1}. \quad (2.23)$$

In letzterem Fall liefert die Parsevalsche Gleichung ein nützliches Resultat:

$$\|f\|^2 = \frac{1}{\pi} \int_{-\pi}^{\pi} dx x^2 = \frac{2\pi^2}{3} = \frac{|a_0|^2}{2} + \sum_{n=1}^{\infty} (|a_n|^2 + |b_n|^2) = \sum_{n=1}^{\infty} \frac{4}{n^2} \Rightarrow \sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}. \quad (2.24)$$

Falls die Funktion f bestimmten Einschränkungen unterliegt, übertragen sich diese Einschränkungen mittels (2.7) und (2.18) auf die Fourierkoeffizienten. Für eine **reelle** Funktion $f(x) = f(x)^*$ gilt

$$c_n = c_{-n}^*, \quad a_n = 2 \operatorname{Re} c_n, \quad b_n = -2 \operatorname{Im} c_n, \quad (2.25)$$

und für eine **imaginäre** Funktion $f(x) = -f(x)^*$ gilt

$$c_n = -c_{-n}^*, \quad a_n = 2i \operatorname{Im} c_n, \quad b_n = 2i \operatorname{Re} c_n. \quad (2.26)$$

Für eine im Periodizitätsintervall $(-L, L)$ **symmetrische** Funktion $f(x) = f(-x)$ erfüllen die Koeffizienten

$$c_n = c_{-n}, \quad a_n = 2c_n, \quad b_n = 0 \quad (2.27)$$

und für eine **antisymmetrischen** Funktion $f(x) = -f(-x)$ gilt

$$c_n = -c_{-n}, \quad a_n = 0, \quad b_n = 2ic_n. \quad (2.28)$$

Allgemeines Periodizitätsintervall. Wie schon in (2.4) angemerkt, kann die Fourierreihe auch für ein allgemeines Periodizitätsintervall $(-\pi, \pi) \rightarrow (a, a + 2L)$ mit $a \in \mathbb{R}$ formuliert werden. Dazu trifft man folgende Ersetzungen:

$$e^{inx} \rightarrow e^{i\frac{n\pi x}{L}}, \quad \frac{1}{2\pi} \int_{-\pi}^{\pi} dx \rightarrow \frac{1}{2L} \int_a^{a+2L} dx, \quad (2.29)$$

bzw. für die reelle Darstellung

$$\cos nx \rightarrow \cos \frac{n\pi x}{L}, \quad \sin nx \rightarrow \sin \frac{n\pi x}{L}, \quad \frac{1}{\pi} \int_{-\pi}^{\pi} dx \rightarrow \frac{1}{L} \int_a^{a+2L} dx. \quad (2.30)$$

Dann wird die komplexe Form der Fourierreihe zu

$$FR(f)(x) = \sum_{n=-\infty}^{\infty} c_n e^{i\frac{n\pi x}{L}}, \quad c_n = \frac{1}{2L} \int_a^{a+2L} dx e^{-i\frac{n\pi x}{L}} f(x) \quad (2.31)$$

und die reelle Form zu

$$\begin{aligned} FR(f)(x) &= \frac{a_0}{2} + \sum_{n=1}^{\infty} \left(a_n \cos \frac{n\pi x}{L} + b_n \sin \frac{n\pi x}{L} \right), \\ a_n &= \frac{1}{L} \int_a^{a+2L} dx \cos \frac{n\pi x}{L} f(x) \quad (n \geq 0), \\ b_n &= \frac{1}{L} \int_a^{a+2L} dx \sin \frac{n\pi x}{L} f(x) \quad (n > 0). \end{aligned} \quad (2.32)$$

2.2 Konvergenzkriterien

Ein zentrales Thema in der Fouriertheorie bereits seit den Überlegungen J. Fouriers ist die Frage der **Konvergenz** von Fourierreihen. Welche Einschränkungen muss man an eine Funktion f stellen, um eine bestimmte Art der Konvergenz – z.B. Konvergenz im Mittel, punktweise Konvergenz oder gleichmäßige Konvergenz – zu erzielen?

Gehen wir dazu einen Schritt zurück und geben die Einschränkung an den Hilbertraum $L^2(S)$ auf. Im Prinzip können wir ja die zweite Gleichung in (2.7) als *Definition* für die Fourierkoeffizienten c_n benutzen und daraus die Reihe $FR(f)$ konstruieren:

$$c_n := \frac{1}{2\pi} \int_{-\pi}^{\pi} dx e^{-inx} f(x), \quad FR(f)(x) = \sum_{n=-\infty}^{\infty} c_n e^{inx}. \quad (2.33)$$

Damit das Integral existiert, muss lediglich $f \in L^1(S)$ erfüllt sein, d.h. die Funktion muss absolut integrierbar sein. Das ist eine schwächere Bedingung, da man allgemein zeigen kann, dass auf einem beschränkten Intervall $L^p \subset L^q$ für $p > q$ gilt. Ob und in welchem Sinn die so resultierende Reihe gegen f konvergiert, muss dann überprüft werden und wird zu weiteren Einschränkungen führen. Dazu definieren wir zweckmäßigerweise die Partialsummen

$$FR_N(f)(x) = \sum_{n=-N}^N c_n e^{inx}. \quad (2.34)$$

Das sind nichts anderes als die Fourierpolynome in (2.15), allerdings mit den Fourierkoeffizienten c_n aus (2.33) eingesetzt. Wir wollen nun untersuchen, in welchem Sinn $\lim_{N \rightarrow \infty} FR_N(f)$ gegen die ursprüngliche Funktion f konvergiert.

Konvergenz im Mittel. Für Funktionen im bisher besprochenen Hilbertraum $L^2(S)$ ist die Sache klar: die φ_n bilden eine ONB und somit gelten alle Aussagen aus Abschnitt 1.4. Es gilt die Parsevalsche Gleichung, und die Fourierreihe konvergiert **im Mittel** bezüglich der L^2 -Norm:

$$\lim_{N \rightarrow \infty} \sum_{n=-N}^N |c_n|^2 = \|f\|_2^2 < \infty \quad \Leftrightarrow \quad \lim_{N \rightarrow \infty} \|FR_N(f) - f\|_2 = 0. \quad (2.35)$$

Allgemeiner kann man zeigen, dass Konvergenz im Mittel auch für L^p -Räume mit $1 < p < \infty$ vorliegt (die dann im Allgemeinen aber keine Hilberträume mehr sind, denn nur L^2 hat ein Skalarprodukt).

Andererseits ist die Konvergenz im Mittel in der Praxis nicht immer nützlich. Der Raum $L^2(-\pi, \pi)$ der quadratintegrierbaren Funktionen im Sinne des Lebesgue-Integrals besteht ja aus Äquivalenzklassen von Funktionen, die sich um Funktionen auf Mengen vom Maß Null unterscheiden können. Die Funktionen innerhalb einer Klasse müssen lediglich im quadratischen Mittel den Abstand Null haben. Wir erhalten daraus keine Aussage, wie sich die Fourierreihe an einer einzelnen Stelle $x \in (-\pi, \pi)$ oder an den Endpunkten $x = \pm\pi$ verhält. Es macht dann auch keinen Sinn, vom ‘‘Funktionswert’’ an einer Stelle zu sprechen, denn diesen kann man immer abändern, ohne dass man es im Sinne des L^2 mit einer anderen Funktion zu tun bekommt. Wenn man z.B. sagt, dass eine Funktion $f \in L^2(-\pi, \pi)$ stetig ist, bedeutet das, dass unter all den Funktionen, die mit ihr im L^2 identifiziert werden, eine ist, die stetig ist.

Punktweise Konvergenz. Ein praxisrelevantes Kriterium für die Konvergenz einer Fourierreihe ist die **punktweise Konvergenz**:

$$\lim_{N \rightarrow \infty} FR_N(f)(x) = f(x) \quad \forall x \in \mathcal{S}. \quad (2.36)$$

Erinnern wir uns z.B. an die Rechteckfunktion in Fig. 2.1: An den Sprungstellen nimmt die Fourierreihe den Wert $1/2$ an. Die Rechteckfunktion wird daher nicht einmal punktweise durch ihre Fourierreihe dargestellt! Man könnte vermuten, das liegt an der Unstetigkeit der Rechteckfunktion, denn $FR_N(f)$ ist als Summe stetiger periodischer Funktionen sicher auch wieder stetig. Doch selbst bei stetigen Funktionen $f \in C(\mathcal{S})$ kann man nicht automatisch darauf schließen, dass sie punktweise durch ihre Fourierreihe dargestellt werden – ein erstes Gegenbeispiel wurde 1876 von **P. du Bois-Reymond** gefunden, und 1966 bewies **L. Carleson**, dass die Fourierreihen stetiger Funktion im Allgemeinen nur *fast überall* gegen diese Funktionen konvergieren. Für die punktweise Konvergenz ist Stetigkeit daher nicht ausreichend. Hinreichend wäre die Differenzierbarkeit, diese Bedingung ist allerdings wieder zu restriktiv. Wir treffen daher folgende Definition:

DEFINITION: Eine 2π -periodische Funktion f heißt **stückweise stetig differenzierbar** (wir schreiben $f \in D(\mathcal{S})$), falls eine endliche Unterteilung von \mathcal{S} existiert, sodass f eine stetige Ableitung in jedem Teilintervall hat und die links- und rechtsseitigen Ableitungen am Ende der Teilintervalle existieren.

Das ist sowohl “besser” als auch “schlechter” als Stetigkeit: $D(\mathcal{S})$ enthält zwar keine stetigen Funktionen mit Cusps wie z.B. $\sqrt{|x|}$, deren Ableitungen an diesen Stellen nicht existieren, andererseits aber einfache unstetige Funktionen wie die Rechteck- oder Sägezahnfunktion. Die Kombination $C(\mathcal{S}) \cap D(\mathcal{S})$ entspricht der Eigenschaft “stetig mit endlich vielen Knicken”. Noch stärkere Einschränkungen sind stetig differenzierbare Funktionen $f \in C^1(\mathcal{S})$ bis hin zu den m -fach stetig differenzierbaren Funktionen $f \in C^m(\mathcal{S})$. Damit lassen sich nun einige Aussagen über die Konvergenz von Fourierreihen treffen.

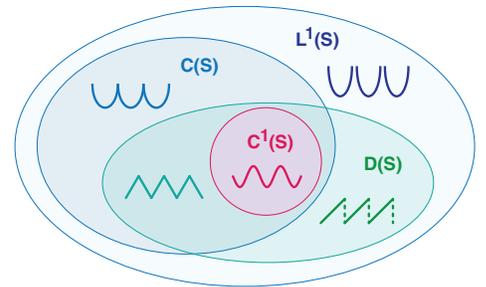
Eine notwendige Voraussetzung für die Konvergenz von $FR(f)$ ist sicher, dass die Folge $(c_n)_{n=-\infty}^{\infty}$ eine Nullfolge ist:

$$\lim_{|n| \rightarrow \infty} c_n = 0. \quad (2.37)$$

Für Funktionen aus dem Hilbertraum $L^2(\mathcal{S})$ ist das automatisch erfüllt, denn es gilt ja die Parsevalsche Gleichung $\sum_n |c_n|^2 = \|f\|^2 < \infty$. Für den Raum $L^1(\mathcal{S})$ gilt die Aussage ebenso, denn hier fällt sie mit dem Riemann-Lebesgue-Lemma zusammen:

LEMMA (**Riemann-Lebesgue**): Sei $f \in L^1(a, b)$, d.h. $\int_a^b dx |f(x)| < \infty$. Dann gilt

$$\lim_{|n| \rightarrow \infty} \int_a^b dx f(x) e^{inx} = 0. \quad (2.38)$$



Das sind nichts anderes als die Fourierkoeffizienten in (2.33), womit (2.37) erfüllt ist. Die Aussage gilt dann auch für die Integrale über die Sinus- und Kosinusfunktion:

$$\lim_{|n| \rightarrow \infty} \int_a^b dx f(x) \sin nx = \lim_{|n| \rightarrow \infty} \int_a^b dx f(x) \cos nx = 0. \quad (2.39)$$

Aus dem Abklingverhalten der Fourierkoeffizienten c_n für $n \rightarrow \infty$ lassen sich weitere Konvergenzaussagen treffen:

- Für $f \in C^m(\mathcal{S})$ erfüllen die Fourierkoeffizienten $\lim_{n \rightarrow \infty} c_n n^m = 0$, d.h. sie fallen schneller als $1/n^m$ ab. Der Beweis erfolgt durch wiederholte partielle Integration:

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} dx e^{-inx} f(x) = \frac{1}{2\pi} \frac{1}{(in)^m} \int_{-\pi}^{\pi} dx e^{-inx} f^{(m)}(x). \quad (2.40)$$

Die Randterme verschwinden wegen der Stetigkeit. Weil die m -te Ableitung $f^{(m)}$ stetig ist, ist sie auch integrierbar und somit verschwindet das Integral für $n \rightarrow \infty$ wegen des Riemann-Lebesgue-Lemmas, also $\lim_{n \rightarrow \infty} c_n n^m = 0$.

- Ebenfalls mit partieller Integration beweist man ein allgemeineres Kriterium [Sta11], nämlich: Für Funktionen $f \in C^{m-1}(\mathcal{S})$, deren Ableitungen $f^{(m)} \in D(\mathcal{S})$ erfüllen, klingen die Fourierkoeffizienten wie $c_n \sim 1/n^{m+1}$ ab. Je glatter eine Funktion ist, desto schneller fallen also ihre Fourierkoeffizienten ab. Typische Beispiele sind:
 - $m = 0$: Funktionen $f \in D(\mathcal{S})$ wie die Rechteckfunktion oder die Sägezahnfunktion, deren Koeffizienten c_n wie $1/n$ abfallen.
 - $m = 1$: Stetige Funktionen mit $f' \in D(\mathcal{S})$, wie die Dreiecksfunktion ($f(x) = |x|$ für $-\pi < x \leq \pi$); die Koeffizienten fallen wie $1/n^2$ ab.
 - $m \rightarrow \infty$: Bei glatten Funktionen fallen die Koeffizienten c_n schneller als jede Potenz ab (z.B. exponentiell, oder sie verschwinden ab einem gewissen n).

Für die punktweise Konvergenz ist die Einschränkung auf einen Funktionenraum eigentlich gar nicht nötig. Selbst bei Funktionen mit (integrablen) Singularitäten, deren Ableitungen dort nicht existieren, konvergiert die Fourierreihe an jenen Stellen punktweise gegen die Funktion, die ein Stück von der Problemstelle entfernt sind. Obwohl die Fourierkoeffizienten c_n durch die globale Formel (2.33) definiert sind, bestimmt nur das *lokale* Verhalten von f an der Stelle x_0 das Konvergenzverhalten:

SATZ (Dirichlet): Sei $f \in L^1(\mathcal{S})$ (d.h. 2π -periodisch und absolut integabel) und x_0 ein Punkt in \mathcal{S} , an dem sowohl die links- und rechtsseitigen Grenzwerte der Funktion f als auch die Grenzwerte ihrer Ableitungen existieren. Dann konvergiert die Fourierreihe (2.33) punktweise gegen den Wert $(f(x_0 + 0) + f(x_0 - 0))/2$. Falls zusätzlich $f \in C(\mathcal{S})$, konvergiert die Fourierreihe punktweise gegen $f(x_0)$.

Für $f \in D(\mathcal{S})$ ist diese Bedingung für jeden Punkt $x_0 \in \mathcal{S}$ erfüllt, d.h. die Fourierreihe konvergiert punktweise gegen $(f(x_0 + 0) + f(x_0 - 0))/2$. Falls zusätzlich $f \in C(\mathcal{S})$, konvergiert die Fourierreihe an jedem Punkt gegen $f(x_0)$.

Zum Beweis des obigen Satzes definieren wir zunächst den **Dirichlet-Kern** $D_N(x)$:

$$D_N(x) := \sum_{n=-N}^N e^{inx} = 1 + \sum_{n=1}^N (e^{inx} + e^{-inx}) = 1 + 2 \sum_{n=1}^N \cos nx. \quad (2.41)$$

Daraus sieht man, dass D_N symmetrisch und 2π -periodisch ist, $D_N(x) = D_N(-x)$ und $D_N(x + 2\pi) = D_N(x)$, und es gilt

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} dx D_N(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} dx + \frac{1}{\pi} \sum_{n=1}^N \underbrace{\int_{-\pi}^{\pi} dx \cos nx}_{=0} = 1. \quad (2.42)$$

Verwenden wir nun

$$\begin{aligned} \left(q^{\frac{1}{2}} - q^{-\frac{1}{2}}\right) \sum_{n=-N}^N q^n &= \left(q^{-N+\frac{1}{2}} + \dots + q^{N-\frac{1}{2}} + q^{N+\frac{1}{2}}\right) - \left(q^{-N-\frac{1}{2}} + q^{-N+\frac{1}{2}} + \dots + q^{N-\frac{1}{2}}\right) \\ &= q^{N+\frac{1}{2}} - q^{-(N+\frac{1}{2})} \end{aligned} \quad (2.43)$$

mit $q = e^{ix}$, so erhalten wir

$$\left(\sin \frac{x}{2}\right) D_N(x) = \frac{e^{\frac{ix}{2}} - e^{-\frac{ix}{2}}}{2i} \sum_{n=-N}^N e^{inx} = \frac{e^{ix(N+\frac{1}{2})} - e^{-ix(N+\frac{1}{2})}}{2i} = \sin \left[\left(N + \frac{1}{2}\right) x \right]. \quad (2.44)$$

Dann folgt für $x \neq 2k\pi$:

$$D_N(x) = \frac{\sin \left[\left(N + \frac{1}{2}\right) x \right]}{\sin \frac{x}{2}}. \quad (2.45)$$

Für $x = 2k\pi$ ergibt sich aus der Definition (2.41) $D_N(x) = 2N + 1$, was mit dem Grenzwert von (2.45) nach de l'Hospital übereinstimmt.

Betrachten wir nun die Partialsummen aus (2.34) für eine 2π -periodische Funktion f :

$$\begin{aligned} FR_N(f)(x) &= \sum_{n=-N}^N c_n e^{inx} = \sum_{n=-N}^N \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} ds e^{-ins} f(s) \right) e^{inx} = \frac{1}{2\pi} \int_{-\pi}^{\pi} ds D_N(x-s) f(s) \\ &= \frac{1}{2\pi} \int_{-\pi-x}^{\pi-x} du D_N(u) f(u+x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} du D_N(u) f(u+x). \end{aligned} \quad (2.46)$$

In der zweiten Zeile haben wir zuerst $s = u+x$ substituiert, dann die Symmetrie $D_N(-u) = D_N(u)$ ausgenützt und im letzten Schritt die Integralgrenzen um x verschoben, da sowohl f als auch D_N 2π -periodisch sind und wir über die gesamte Periode integrieren. Die Partialsummen können also durch Integration über den Dirichlet-Kern gewonnen werden. Wegen der Symmetrie $u \leftrightarrow -u$ können wir das Ergebnis auch so schreiben:

$$FR_N(f)(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} du D_N(u) \frac{f(x+u) + f(x-u)}{2}. \quad (2.47)$$

Die verbleibende Frage ist, ob die Partialsummen für $N \rightarrow \infty$ gegen die Funktion f konvergieren. Dazu bilden wir an einen gegebenen Punkt x_0 die Differenz

$$\begin{aligned} FR_N(f)(x_0) - f(x_0) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} du D_N(u) \frac{f(x_0+u) + f(x_0-u)}{2} - \frac{f(x_0)}{2\pi} \int_{-\pi}^{\pi} du D_N(u) \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} du D_N(u) \underbrace{\left(\frac{f(x_0+u) + f(x_0-u)}{2} - f(x_0) \right)}_{=: F(x_0, u)} \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} du \sin \left[\left(N + \frac{1}{2}\right) u \right] \frac{F(x_0, u)}{u} \frac{u}{\sin \frac{u}{2}}. \end{aligned} \quad (2.48)$$

Nach dem Riemann-Lebesgue-Lemma (2.39) verschwindet dieses Integral für $N \rightarrow \infty$, falls $F(x_0, u)/\sin \frac{u}{2}$ integrierbar ist. Die Funktion $u/\sin \frac{u}{2}$ ist im Intervall $(-\pi, \pi)$ beschränkt (ihr Limes für $u \rightarrow 0$ ergibt 2). Wir müssen also nur die Integrierbarkeit des Restterms

$$\frac{F(x_0, u)}{u} = \frac{1}{u} \left(\frac{f(x_0 + u) + f(x_0 - u)}{2} - f(x_0) \right) = \frac{f(x_0 + u) - f(x_0)}{2u} - \frac{f(x_0 - u) - f(x_0)}{-2u} \quad (2.49)$$

zeigen (was auch als **Konvergenzkriterium von Dini** bekannt ist). Dieser ist beschränkt, falls die links- und rechtsseitigen Ableitungen der Funktion f an der Stelle x_0 existieren; falls x_0 zusätzlich eine Sprungstelle ist, müssen wir den Funktionswert an dieser Stelle durch den Mittelwert der links- und rechtsseitigen Grenzwerte ersetzen, um die Existenz des Integrals zu gewährleisten:

$$f(x_0) = \frac{f(x_0 + 0) + f(x_0 - 0)}{2}. \quad (2.50)$$

Weil $FR_N(f)(x_0) - f(x_0)$ für $N \rightarrow \infty$ dann verschwindet, konvergiert die Fourierreihe gegen diesen Wert. Falls die Funktion an der Stelle x_0 stetig ist, konvergiert die Fourierreihe gegen den Funktionswert $f(x_0)$.

Damit haben wir den Satz bewiesen: Unsere einzigen Forderungen waren $f \in L^1(\mathcal{S})$ (sonst würden die Koeffizienten in (2.33) nicht existieren) und die Existenz der Grenzwerte der Funktion am Punkt x_0 sowie der Grenzwerte ihrer Ableitungen (sonst würde der Limes von (2.48) für $N \rightarrow \infty$ nicht verschwinden).

Gleichmäßige Konvergenz. Ein stärkeres Konvergenzkriterium als die punktweise Konvergenz ist die **gleichmäßige Konvergenz**:

$$\lim_{N \rightarrow \infty} \|FR_N(f) - f\|_\infty = \lim_{N \rightarrow \infty} \max_{x \in \mathcal{S}} |FR_N(f)(x) - f(x)| = 0. \quad (2.51)$$

Das ist die Maximumnorm, die wir schon früher besprochen haben. Anschaulich bedeutet das, dass sich die Graphen der Partialsummen $FR_N(f)(x)$ auf dem Definitionsintervall gleichmäßig dem Graphen der Grenzfunktion f annähern. Für reellwertige Funktionen betrachtet man dazu einen “ ε -Schlauch”, d.h. einen Parallelstreifen $f(x) - \varepsilon < y < f(x) + \varepsilon$ der Dicke 2ε (Fig. 1.2). Die gleichmäßige Konvergenz kann man dann so ausdrücken: Für jedes $\varepsilon > 0$ existiert eine Zahl $N_0(\varepsilon)$, sodass für alle $N \geq N_0(\varepsilon)$ der Graph von $FR_N(f)(x)$ im ε -Schlauch verläuft. Das ist eine stärkere Bedingung als die punktweise Konvergenz, da sie auf den globalen Eigenschaften der Funktion beruht.

Unter welchen Voraussetzungen an f konvergiert ihre Fourierreihe gleichmäßig? Eine nützliche Aussage ist die folgende:

Majorantenkriterium von Weierstraß: Sei $\sum_{n=0}^{\infty} a_n$ mit $a_n > 0$ eine konvergente Reihe und $|b_n(x)| \leq a_n \forall x \in A$ und fast alle n . Dann konvergiert die Reihe $\sum_{n=0}^{\infty} b_n(x)$ gleichmäßig auf $x \in A$.

Setzen wir $a_n = |c_n|$ und $b_n(x) = c_n e^{inx}$, wobei c_n die Fourierkoeffizienten der Funktion f sind, folgt daraus: Falls die Reihe $\sum_{n=0}^{\infty} |c_n|$ konvergiert, konvergiert die Fourierreihe $FR(f)$ gleichmäßig. Für welche Funktionen ist das der Fall?

SATZ: Sei $f \in C(\mathcal{S}) \cap D(\mathcal{S})$, d.h. f ist 2π -periodisch, stetig und stückweise stetig differenzierbar. Dann konvergiert ihre Fourierreihe gleichmäßig gegen f .

Zum Beweis erinnern wir uns daran, dass die Fourierreihe einer Funktion $f \in C(\mathcal{S}) \cap D(\mathcal{S})$ punktweise gegen die Funktion f konvergiert. Um die gleichmäßige Konvergenz zu überprüfen, betrachten wir die Fourierkoeffizienten d_n der Funktion f' und teilen das Intervall $(-\pi, \pi)$ in die Differenzierbarkeitsbereiche von f auf, sodass $x_0 = -\pi < x_1 < x_2 < \dots < x_N = \pi$ die Stellen sind, an denen f keine eindeutige Ableitung hat:

$$\begin{aligned} d_n &= \frac{1}{2\pi} \int_{-\pi}^{\pi} dx e^{-inx} f'(x) = \frac{1}{2\pi} \sum_{j=1}^N \int_{x_{j-1}}^{x_j} dx e^{-inx} f'(x) \\ &= \frac{1}{2\pi} \sum_{j=1}^N \left(f(x) e^{-inx} \Big|_{x_{j-1}}^{x_j} + in \int_{x_{j-1}}^{x_j} dx e^{-inx} f(x) \right) = \frac{1}{2\pi} f(x) e^{-inx} \Big|_{-\pi}^{\pi} + in c_n. \end{aligned} \quad (2.52)$$

Die Zwischenschritte erforderten die Stetigkeit von f und die Existenz der links- und rechtsseitigen Ableitungen am Ende der Teilintervalle. Der Randterm im letzten Ausdruck verschwindet ebenfalls wegen der Stetigkeit an den Rändern, $f(\pi) = f(-\pi)$. Damit gilt $d_n = in c_n$, wobei c_n die Fourierkoeffizienten der Funktion selbst bezeichnen, und insbesondere $d_0 = 0$.

Jetzt benutzen wir die Cauchy-Schwarz-Ungleichung $|\langle x, y \rangle| \leq \|x\| \|y\|$ aus Gl. (1.36) für den Hilbertschen Folgenraum ℓ^2 ,

$$\left| \sum_n a_n^* b_n \right| \leq \sqrt{\sum_n |a_n|^2} \sqrt{\sum_n |b_n|^2}, \quad (2.53)$$

und setzen $a_n = 1/|n|$, $b_n = |d_n|$. Von der Reihe $\sum_n 1/n^2$ wissen wir, dass sie konvergiert. Die Ableitung f' ist stückweise stetig und somit auch ein Element des $L^2(\mathcal{S})$, daher wissen wir aus der Parseval'schen Gleichung, dass die Reihe $\sum_n |d_n|^2$ konvergiert. Damit können wir schreiben:

$$\sum_{n=-\infty}^{\infty} |c_n| = |c_0| + \sum_{n \neq 0} \frac{|d_n|}{|n|} \leq |c_0| + \sqrt{\sum_{n \neq 0} \frac{1}{n^2}} \sqrt{\sum_{n \neq 0} |d_n|^2} < \infty. \quad (2.54)$$

Somit konvergiert $FR(f)$ gleichmäßig gegen f . Wir hätten den Satz auch nur für stetig differenzierbare Funktionen $f \in C^1(\mathcal{S})$ beweisen können, dann wäre die Intervallaufteilung in (2.52) nicht nötig gewesen.

Für gleichmäßig konvergente Reihen gelten außerdem eine Reihe weiterer Aussagen:

- Die Summe einer gleichmäßig konvergenten Reihe von stetigen Funktionen ist wieder eine stetige Funktion.
- Für gleichmäßig konvergente Reihen sind Summe und Integral vertauschbar.
- Wenn die Reihe, die man durch gliedweise Differenziation einer stetig differenzierbaren, konvergenten Reihe erhält, gleichmäßig konvergent ist, dann konvergiert sie gegen die Ableitung der Summenfunktion.

Fourierreihen unstetiger Funktionen können daher nicht überall gleichmäßig konvergieren, denn sie müssten sonst als Summe stetiger Funktionen wieder gegen eine stetige Funktion konvergieren und nicht die gewünschte unstetige Funktion darstellen. Insbesondere ist das für die Rechteck- und Sägezahnfunktion in (2.11) und (2.14) der Fall. Ihre Fourierkoeffizienten fallen nur wie $1/n$ ab, sodass die Reihen $\sum_n |c_n|$ nicht konvergieren. Das äußert sich in den Spitzen neben den Sprungstellen, die in Fig. 2.1 sichtbar sind. Dort schießen die Partialsummen der Fourierreihe um einen unangenehm hohen Betrag über das Ziel hinaus, der sich auch für $N \rightarrow \infty$ nicht ändert; die Fourierreihe konvergiert also nicht gleichmäßig. Dieses **Gibbs'sche Phänomen** ist bei allen unstetigen periodischen Funktionen zu beobachten. Man spricht auch von Unter- bzw. Oberschwingern der Fourierpolynome, was zu Artefakten in der Signalverarbeitung und Bildkompression führt.

Betrachten wir die Fourierreihe (2.11) der Rechteckfunktion. Man kann zeigen, dass ihr erstes Maximum für $x > 0$ bei $x_0 = \pi/(2(N+1))$ liegt. Daraus bestimmt sich der Wert bei x_0 zu

$$FR_N(f)(x_0) = \frac{1}{2} + \frac{1}{\pi} \sum_{k=0}^N \frac{\sin[(2k+1)x_0]}{(2k+1)x_0} 2x_0 \xrightarrow{N \rightarrow \infty} \frac{1}{2} + \frac{1}{\pi} \int_0^{\pi} dz \frac{\sin z}{z}. \quad (2.55)$$

Hier haben wir die Diskretisierung des Integrals über $(\sin z)/z$ benützt,

$$\int_0^a dz \frac{\sin z}{z} \approx \sum_{k=0}^N \Delta z \frac{\sin z_k}{z_k}, \quad \Delta z = z_{k+1} - z_k = \frac{a}{N+1}, \quad (2.56)$$

was mit $z_k = (2k+1)x_0$ auf $\Delta z = 2x_0$ und $a = \pi$ führt. Dementsprechend lautet die Differenz aus dem Maximum bei x_0 und dem Minimum bei $-x_0$

$$\frac{2}{\pi} \int_0^{\pi} dz \frac{\sin z}{z} \approx 1.17898, \quad (2.57)$$

d.h. die Differenz der Funktionswerte wird um fast 18% überschritten!

Differenziation und Integration. Auch beim Differenzieren von Fourierreihen muss man vorsichtig sein. So ergibt z.B. die gliedweise Differenziation der Fourierreihe (2.14) der Sägezahnfunktion

$$2 \sum_{n=1}^{\infty} (-1)^{n+1} \cos nx, \quad (2.58)$$

doch die Koeffizienten $2(-1)^{n+1}$ dieser Reihe bilden keine Nullfolge mehr, die Reihe kann also nicht konvergieren. Die Voraussetzungen für eine punktweise Konvergenz der Fourierreihe reichen also für die gliedweise Differenzierbarkeit nicht aus. Stattdessen gilt:

SATZ: Sei $f \in C(\mathcal{S}) \cap D(\mathcal{S})$, d.h. f ist 2π -periodisch, stetig und ihre Ableitung stückweise stetig. Dann kann $FR(f)$ gliedweise differenziert werden, und die so gewonnene Reihe stimmt mit der Fourierreihe der Ableitung der Funktion überein.

Das haben wir mit (2.52) schon bewiesen: die Fourierkoeffizienten der Funktion f' sind durch $d_n = in c_n$ gegeben, damit wird die Fourierreihe von f' zu

$$FR(f')(x) = \sum_{n=-\infty}^{\infty} in c_n e^{inx} = \frac{d}{dx} \Big|_{\text{termweise}} FR(f)(x). \quad (2.59)$$

Falls diese Reihe außerdem gleichmäßig konvergiert, dann ist die Grenzfunktion eine stetige Funktion, die auf dem Intervall $(-\pi, \pi)$ mit f' übereinstimmt.

Betrachten wir die Funktion $f(x) = x^2/2$ für $x \in (-\pi, \pi)$. Ihre periodische Fortsetzung ist stetig und ihre Ableitung stückweise stetig. Die zugehörige Fourierreihe

$$FR(f)(x) = \frac{\pi^2}{6} + 2 \sum_{n=1}^{\infty} \frac{(-1)^n}{n^2} \cos nx \quad (2.60)$$

kann daher gliedweise differenziert werden, und ihre Ableitung ergibt (2.14), was mit der Fourierreihe von $f'(x) = x$ für $x \in (-\pi, \pi)$ übereinstimmt.

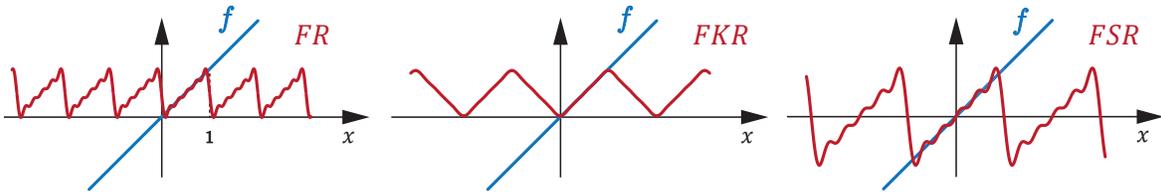


FIG. 2.2: Periodische Fortsetzung der Funktion $f(x) = x$ im Intervall $(0, 1)$ durch die Fourier-, Fourier-Kosinus- und Fourier-Sinus-Reihe mit jeweils 5 Termen.

Zur termweisen Integration einer Fourierreihe muss die Funktion f nur stückweise stetig sein. Dabei werden die Koeffizienten c_n mit Faktoren $1/n$ multipliziert, was die Konvergenz verbessert:

$$\sum_n \int dx c_n e^{inx} = c_0 x + \sum_{n \neq 0} \frac{c_n}{in} e^{inx} + c. \quad (2.61)$$

Allerdings ist die resultierende Funktion nur dann periodisch, wenn $c_0 = 0$ gilt.

2.3 Weitere Anwendungen

Fourier-Sinus- und Fourier-Kosinus-Reihe. In (2.27–2.28) haben wir gesehen, dass für eine symmetrische (gerade) Funktion $f(x) = f(-x)$ nur die Fourierkoeffizienten a_n überleben und für eine antisymmetrische (ungerade) Funktion $f(x) = -f(-x)$ nur die b_n . Da man jede Funktion in einen geraden und einen ungeraden Anteil aufspalten kann,

$$f(x) = \frac{f(x) + f(-x)}{2} + \frac{f(x) - f(-x)}{2} = f_+(x) + f_-(x), \quad (2.62)$$

trägt der gerade Anteil der Funktion nur zur Kosinusreihe bei und der ungerade nur zur Sinusreihe, z.B.

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} dx f_+(x) \cos nx = \frac{2}{\pi} \int_0^{\pi} dx f_+(x) \cos nx, \quad b_n = \frac{2}{\pi} \int_{-\pi}^{\pi} dx f_-(x) \sin nx. \quad (2.63)$$

Man führt daher spezielle Fourierreihen für gerade und ungerade Funktionen ein, die **Fourier-Kosinus-Reihe** für Funktionen, die im Periodizitätsintervall $(-L, L)$ gerade sind,

$$FKR(f)(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos\left(\frac{n\pi x}{L}\right), \quad a_n = \frac{2}{L} \int_0^L dx f(x) \cos\left(\frac{n\pi x}{L}\right), \quad (2.64)$$

und die **Fourier-Sinus-Reihe** für im Periodizitätsintervall $(-L, L)$ ungerade Funktionen:

$$FSR(f)(x) = \sum_{n=1}^{\infty} b_n \sin\left(\frac{n\pi x}{L}\right), \quad b_n = \frac{2}{L} \int_0^L dx f(x) \sin\left(\frac{n\pi x}{L}\right). \quad (2.65)$$

In beiden Fällen läuft die Integration nur über das halbe Intervall $(0, L)$, da der Beitrag der anderen Hälfte ja gleich groß ist. Damit lässt sich eine in einem Intervall $(0, L)$ gegebene Funktion auf mehrere Wege durch Fourierreihen darstellen: entweder durch die übliche Fourierreihe, oder als symmetrische bzw. antisymmetrische Erweiterung auf das Intervall $(-L, L)$ mittels *FKR* oder *FSR*.

Dabei können auch die Konvergenzeigenschaften verschieden sein, wie in Fig. 2.2 für die Funktion $f(x) = x$ für $0 \leq x < 1$ gezeigt. Die Koeffizienten der üblichen Fourierreihe für die Periode $(0, 1)$ lauten

$$a_0 = 1, \quad a_{n>0} = 0, \quad b_{n>0} = -\frac{1}{n\pi}, \quad (2.66)$$

jene der *FKR* für die Periode $(-1, 1)$ sind

$$a_0 = 1, \quad a_{n>0} = \begin{cases} 0 & n \text{ gerade} \\ -\frac{4}{n^2\pi^2} & n \text{ ungerade} \end{cases} \quad (2.67)$$

und jene der *FSR* für die Periode $(-1, 1)$

$$b_{n>0} = \frac{2}{n\pi} (-1)^{n+1}. \quad (2.68)$$

Die *FKR* konvergiert am schnellsten, da ihre Koeffizienten mit $1/n^2$ abfallen, während jene der anderen beiden Reihen nur mit $1/n$ verschwinden.

Diskrete Fouriertransformation. In der Praxis hat man es oft mit diskreten Funktionswerten $f(x_k) = f_k$ an äquidistanten Datenpunkten x_k zu tun, die man geeignet interpolieren will. Das kann man durch polynomiale Interpolation oder Splines bewerkstelligen, aber auch die Fourierreihe eignet sich dafür sehr gut. Zur Berechnung der Fourierkoeffizienten muss man allerdings das Integral in (2.33) diskretisieren, da die Funktion f ja nicht als geschlossener Ausdruck vorliegt:

$$c_n = \frac{1}{2\pi} \int_0^{2\pi} dx f(x) e^{-inx} \rightarrow \frac{1}{2\pi} \sum_{k=0}^{N-1} \Delta x f_k e^{-inx_k} = \frac{1}{N} \sum_{k=0}^{N-1} f_k e^{-ink \frac{2\pi}{N}}. \quad (2.69)$$

Hier haben wir das Intervall $(0, 2\pi)$ in N Teilintervalle mit $\Delta x = 2\pi/N$ zerlegt und das Integral durch eine Summe über die Funktionswerte an den äquidistanten Stützstellen $x_k = 2\pi k/N$ mit $k = 0, \dots, N-1$ genähert. Analog schreiben wir für die Fourierreihe:

$$f(x) = \sum_{n=0}^{\infty} c_n e^{inx} \rightarrow f_j = \sum_{n=0}^{N-1} c_n e^{inj \frac{2\pi}{N}}. \quad (2.70)$$

Als Resultat erhalten wir zwei Matrix-Vektor-Gleichungen,

$$f_j = \sum_{n=0}^{N-1} M_{jn} c_n, \quad c_n = \sum_{k=0}^{N-1} (M^{-1})_{nk} f_k, \quad M_{jn} = z_N^{jn}, \quad (M^{-1})_{nk} = \frac{1}{N} z_N^{-nk}, \quad (2.71)$$

wobei z_N die N -te Einheitswurzel bezeichnet:

$$z_N = e^{i \frac{2\pi}{N}}. \quad (2.72)$$

Dass die beiden Matrizen M und M^{-1} tatsächlich invers zueinander sind,

$$(MM^{-1})_{jk} = \frac{1}{N} \sum_{n=0}^{N-1} z_N^{(j-k)n} = \delta_{jk}, \quad (2.73)$$

kann man so sehen: für $j = k$ ist die Aussage trivial; für $j \neq k$ ist $Z = z_N^{j-k} \neq 1$ wieder eine Einheitswurzel mit $Z^N = e^{2i\pi(j-k)} = 1$, daher gilt

$$\sum_{n=0}^{N-1} Z^n = 1 + Z + Z^2 + \dots + Z^{N-1} = \frac{1 - Z^N}{1 - Z} = 0. \quad (2.74)$$

Hat man die Koeffizienten über $\mathbf{c} = M^{-1}\mathbf{f}$ bestimmt, kann man mittels (2.70) die Fourierreihe an beliebigen Punkten x auswerten. Insbesondere erhält man – trotz der Approximation – an den Stützstellen x_j wieder die ursprünglichen Funktionswerte zurück!

Da für jeden der Koeffizienten N Beiträge zu berechnen sind, wächst der numerische Aufwand allerdings wie $\mathcal{O}(N^2)$. Die **Fast Fourier Transformation (FFT)** ist eine Methode, um diesen Aufwand drastisch zu reduzieren. Nehmen wir dazu an, N sei eine Potenz von 2, also $N = 2, 4, 8, 16$ usw. Spalten wir zunächst die Summe von $k = 0 \dots 2N - 1$ mit N Termen in zwei Teilsummen mit jeweils $N/2$ Termen auf:

$$c_n = \sum_{k=0}^{N-1} (M^{-1})_{nk} f_k = \left(\sum_{k=0}^{N/2-1} + \sum_{k=N/2}^{N-1} \right) (M^{-1})_{nk} f_k = \sum_{k=0}^{N/2-1} \left[(M^{-1})_{nk} + (M^{-1})_{n, k+\frac{N}{2}} f_{k+\frac{N}{2}} \right]. \quad (2.75)$$

Wenn wir die Eigenschaften der Einheitswurzeln z_N ausnutzen,

$$z_N = z_{N/2}^{1/2}, \quad z_N^N = 1, \quad z_N^{N/2} = -1, \quad (2.76)$$

können wir für gerade $n = 2j$ schreiben:

$$(M^{-1})_{nk} = \frac{1}{N} z_N^{-2jk} = \frac{1}{N} z_{N/2}^{-jk}, \quad (M^{-1})_{n, k+\frac{N}{2}} = \frac{1}{N} z_N^{-2jk} z_N^{-jN} = \frac{1}{N} z_{N/2}^{-jk}, \quad (2.77)$$

und somit

$$c_{2j} = \frac{1}{N} \sum_{k=0}^{N/2-1} z_{N/2}^{-jk} (f_k + f_{k+\frac{N}{2}}) =: \sum_{k=0}^{N/2-1} z_{N/2}^{-jk} f_k^{(1)}. \quad (2.78)$$

Analog erhalten wir für ungerade $n = 2j + 1$

$$(M^{-1})_{nk} = \frac{1}{N} z_N^{-(2j+1)k} = \frac{1}{N} z_{N/2}^{-jk} z_N^{-k}, \quad (M^{-1})_{n, k+\frac{N}{2}} = \frac{1}{N} z_N^{-(2j+1)(k+\frac{N}{2})} = \frac{1}{N} z_{N/2}^{-jk} z_N^{-k}, \quad (2.79)$$

und damit

$$c_{2j+1} = \frac{1}{N} \sum_{k=0}^{N/2-1} z_{N/2}^{-jk} z_N^{-k} (f_k + f_{k+\frac{N}{2}}) =: \sum_{k=0}^{N/2-1} z_{N/2}^{-jk} f_k^{(2)}. \quad (2.80)$$

Das sind wieder Matrix-Vektor-Multiplikationen, aber jeweils nur mehr mit der *halben* Länge $N/2$ der Vektoren. Das Verfahren kann man weiter iterieren, wodurch die Zahl der Rechenschritte von $\mathcal{O}(N^2)$ auf $\mathcal{O}(N \log_2 N)$ reduziert wird. So verringert sich z.B. der Aufwand für $N = 2^{16}$ um einen Faktor $N^2 / (N \log_2 N) = 4096!$

Kapitel 3

Integraltransformationen

Im letzten Kapitel haben wir gesehen, dass sich periodische Funktionen auf einem Intervall $[-L, L]$ in Fourierreihen entwickeln lassen. Was passiert, wenn die Funktion nicht mehr periodisch ist – oder anders gesagt, wenn das Periodizitätsintervall auf die gesamte reelle Achse erweitert wird? Will man solche Funktionen durch Summen der Art $c_n e^{inx}$ darstellen, bemerkt man schnell, dass im Gegensatz zu periodischen Funktionen, wo n nur diskrete Werte annimmt, in diesem Fall n stetig variieren muss. In anderen Worten, man muss die Summen durch *Integrale* ersetzen.

Fouriertransformation. Betrachten wir dazu eine Funktion $f: \mathbb{R} \rightarrow \mathbb{C}$, die für $|x| \rightarrow \infty$ hinreichend schnell abfällt (was das genau bedeutet, werden wir später präzisieren) und die im Intervall $-L < x \leq L$ mit der Funktion f_L zusammenfällt. Falls f_L die früher besprochenen Bedingungen erfüllt (stetig und stückweise stetig differenzierbar), konvergiert ihre Fourierreihe punktweise gegen f_L , daher gilt für alle $x \in (-L, L)$:

$$f_L(x) = \sum_{n=-\infty}^{\infty} c_n e^{i \frac{n\pi x}{L}}, \quad c_n = \frac{1}{2L} \int_{-L}^L dx e^{-i \frac{n\pi x}{L}} f(x). \quad (3.1)$$

Definieren wir jetzt

$$p_n = n \frac{\pi}{L}, \quad \Delta p = \frac{\pi}{L}, \quad g(p) = \int_{-L}^L dx f(x) e^{-ipx}, \quad (3.2)$$

dann lässt sich (3.1) so schreiben:

$$c_n = \frac{\Delta p}{2\pi} g(p_n) \quad \Rightarrow \quad f_L(x) = \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} \Delta p g(p_n) e^{ip_n x}. \quad (3.3)$$

Für große L lässt sich das als Diskretisierung eines Riemannsches Integrals interpretieren, bei dem die p -Achse in Teilintervalle π/L zerlegt wurde. Der Grenzübergang $L \rightarrow \infty$ liefert dann die **Fouriertransformation**:

$$g(p) = \int_{-\infty}^{\infty} dx f(x) e^{-ipx} \quad \longleftrightarrow \quad f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dp g(p) e^{ipx}. \quad (3.4)$$

Die Funktion g wird Fouriertransformierte von f genannt. Die Normierung ist in der Literatur nicht einheitlich, man könnte den Faktor $1/(2\pi)$ genausogut symmetrisch aufteilen und stattdessen vor beide Integrale $1/\sqrt{2\pi}$ stellen; wichtig ist nur, dass das Produkt der Faktoren in (3.4) den Wert $1/(2\pi)$ ergibt. Das kann auch bei manchen anderen Beziehungen zu abweichenden Vorfaktoren führen, und darauf muss man achten, wenn man mit der Literatur vergleicht, Fouriertransformationen aus Tabellen nachschlägt oder eine mathematische Software verwendet.

Integraltransformationen. Die Fouriertransformation ist ein Spezialfall von allgemeinen Integraltransformationen der Art

$$g(p) = (Af)(p) = \int dx K(x, p) f(x), \quad (3.5)$$

die durch den jeweiligen **Kern** (engl. *kernel*) $K(x, p)$ der Transformation charakterisiert sind. A ist ein **linearer Operator**, der Funktionen aufeinander abbildet (mit solchen Operatoren werden wir uns in Kap. 5 eingehend beschäftigen) und zumindest im Prinzip umkehrbar ist, d.h für $f_1, f_2 \in \mathbb{V}$ und $\alpha, \beta \in \mathbb{C}$ gilt

$$A(\alpha f_1 + \beta f_2) = \alpha Af_1 + \beta Af_2, \quad f = A^{-1}g = A^{-1}Af. \quad (3.6)$$

Eine zentrale Frage betrifft seinen Definitions- und Wertebereich, denn um sicherzustellen, dass das Integral existiert, müssen bestimmte Forderungen an f und g gestellt werden.

Einige Beispiele für Integraltransformationen sind:

$$\blacksquare \text{ **Fouriertransformation:}** \quad g(p) = \mathcal{F}f(p) = \int_{-\infty}^{\infty} dx f(x) e^{-ipx}, \quad (3.7)$$

$$\blacksquare \text{ **Laplace-Transformation:}** \quad g(p) = \mathcal{L}f(p) = \int_0^{\infty} dx f(x) e^{-px}, \quad (3.8)$$

$$\blacksquare \text{ **Mellin-Transformation:}** \quad g(p) = \mathcal{M}f(p) = \int_0^{\infty} dx f(x) x^{p-1}. \quad (3.9)$$

Die Funktionen $g = \mathcal{F}f$, $\mathcal{L}f$ bzw. $\mathcal{M}f$ sind dann jeweils die Fourier-, Laplace- oder Mellin-Transformierten zu f . Oft verwendet man auch Sonderzeichen oder Großbuchstaben, z.B. $\hat{f} = \mathcal{F}f$ oder $F = \mathcal{L}f$. Insbesondere die Fourier- und Laplace-Transformation bieten oft elegante Wege zur Lösung von Differential- und Integralgleichungen. Nach der Transformation vom Originalbereich (“ x -Raum”) in den Bildbereich (“ p -Raum”) können diese unter Umständen auf algebraische Gleichungen zurückgeführt werden, deren Lösungen man dann wieder zurücktransformieren kann. Die Fouriertransformation spielt in der Quantenmechanik eine zentrale Rolle, wo sie den Übergang vom Ortsraum zum Impulsraum bewirkt, und findet außerdem vielfältigen Einsatz in der Spektralanalyse von Signalen und der Kompression von Audio- und Bilddateien. Andere Integraltransformationen wie die Mellin-, Hankel- und Hilbert-Transformationen sind bei der Behandlung bestimmter Differentialgleichungen nützlich, und die Radon-Transformation liefert die Grundlage der heutigen Computertomografie.

3.1 Fouriertransformation

Die Fouriertransformation (3.4) hat einige sehr angenehme mathematische Eigenschaften. Ihre natürliche Definitionsmenge ist der Raum $L^1(\mathbb{R})$. Obwohl solche Funktionen Pole, Sprungstellen usw. besitzen können, bügelt die Fouriertransformation diese Eigenschaften aus:

SATZ: Sei $f \in L^1(\mathbb{R})$ und die **Fouriertransformation** definiert durch

$$\mathcal{F}f(p) = \int_{-\infty}^{\infty} dx f(x) e^{-ipx}, \quad p \in \mathbb{R}. \quad (3.10)$$

Dann ist die Fouriertransformierte $\mathcal{F}f$ eine beschränkte, stetige Funktion, die im Unendlichen verschwindet (**Lemma von Riemann und Lebesgue**):

$$\lim_{|p| \rightarrow \infty} \mathcal{F}f(p) = 0. \quad (3.11)$$

Wir wollen diese Aussagen nacheinander beweisen:

- Für $f \in L^1(\mathbb{R})$ ist $\mathcal{F}f$ beschränkt: $\|\mathcal{F}f\|_{\infty} \leq \|f\|_1$, denn

$$|\mathcal{F}f(p)| \leq \int_{-\infty}^{\infty} dx |f(x) e^{-ipx}| = \int_{-\infty}^{\infty} dx |f(x)| = \|f\|_1 \quad \forall p \in \mathbb{R}. \quad (3.12)$$

- Für $f \in L^1(\mathbb{R})$ ist $\mathcal{F}f$ stetig: Sei $(p_k)_{k \in \mathbb{N}} \subset \mathbb{R}$ eine Folge mit $\lim_{k \rightarrow \infty} p_k = p \in \mathbb{R}$. Dann gilt

$$\lim_{k \rightarrow \infty} \mathcal{F}f(p_k) = \lim_{k \rightarrow \infty} \int_{-\infty}^{\infty} dx f(x) e^{-ip_k x} = \int_{-\infty}^{\infty} dx f(x) e^{-ipx} = \mathcal{F}f(p). \quad (3.13)$$

Wir benötigen also lediglich die Stetigkeit von e^{-ipx} , nicht jene der Funktion f selbst. Die Vertauschung des Limes mit der Integration ist wegen der Abschätzung $|f(x) e^{-ip_k x}| \leq |f(x)| \in L^1(\mathbb{R})$ erlaubt (**Satz von Lebesgue zur majorisierten Konvergenz**).

- Zum Beweis des Riemann-Lebesgue-Lemmas (dessen Spezialfall für ein endliches Intervall ist uns schon in (2.38) begegnet) betrachten wir eine Funktion $h \in C_0^{\infty}(\mathbb{R})$, d.h. eine glatte Funktion mit einem kompakten Träger. Für solche Funktionen gilt

$$\mathcal{F}h(p) = \int_{-\infty}^{\infty} dx h(x) e^{-ipx} = \int_{-\infty}^{\infty} dx h(x) \frac{d}{dx} \left(-\frac{1}{ip} e^{-ipx} \right) \stackrel{\text{p.I.}}{=} \frac{1}{ip} \int_{-\infty}^{\infty} dx h'(x) e^{-ipx} = \frac{1}{ip} \mathcal{F}(h')(p).$$

Bei der partiellen Integration verschwinden die Randterme wegen $h(|x| \rightarrow \infty) = 0$. Aus $\|\mathcal{F}h\|_{\infty} \leq \|h\|_1$ folgt dann die Aussage (3.11) für $h \in C_0^{\infty}(\mathbb{R})$:

$$|\mathcal{F}h(p)| = \frac{1}{|p|} |\mathcal{F}(h')(p)| \leq \frac{1}{|p|} \|h'\|_1 \quad \forall p \in \mathbb{R} \quad \Rightarrow \quad \lim_{|p| \rightarrow \infty} \mathcal{F}h(p) = 0. \quad (3.14)$$

Die Verallgemeinerung auf Funktionen $f \in L^1(\mathbb{R})$ beruht auf dem Umstand, dass $C_0^{\infty}(\mathbb{R})$ für $1 \leq p < \infty$ ein dichter Teilraum von $L^p(\mathbb{R})$ ist. Insbesondere lässt sich jedes $f \in L^1(\mathbb{R})$ durch eine Folge von Funktionen aus C_0^{∞} approximieren, d.h. für festes $\varepsilon > 0$ existiert ein $h \in C_0^{\infty}$, sodass

$$\|f - h\|_1 = \int_{-\infty}^{\infty} dx |f(x) - h(x)| < \varepsilon. \quad (3.15)$$

Dann gilt auch $\|\mathcal{F}(f - h)\|_{\infty} < \varepsilon$, also $|\mathcal{F}f(p) - \mathcal{F}h(p)| < \varepsilon \quad \forall p \in \mathbb{R}$, und für $|p| \rightarrow \infty$ folgt wegen $\mathcal{F}h(p) \rightarrow 0$ auch $\mathcal{F}f(p) \rightarrow 0$.

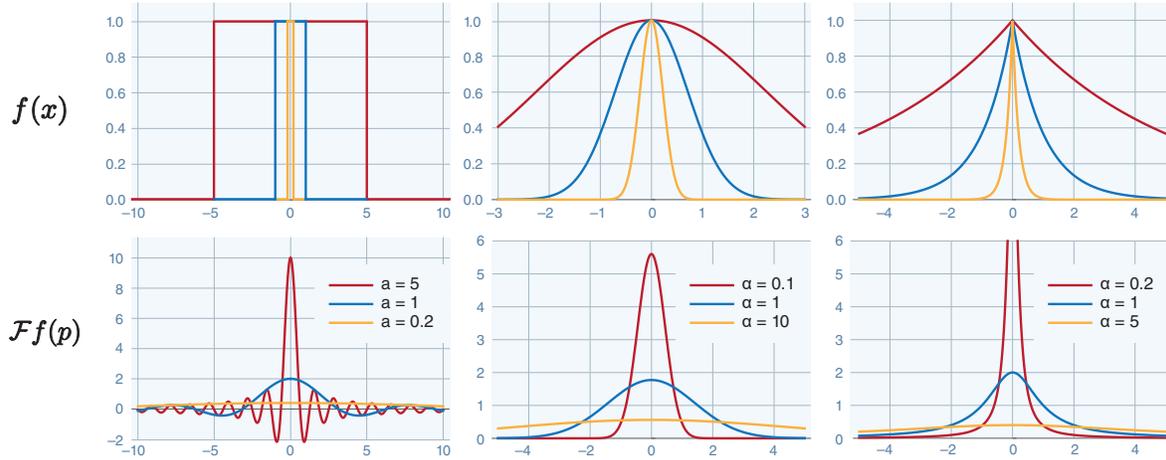


FIG. 3.1: Fouriertransformation des Rechteckimpulses, der Gaußfunktion und des exponentiell abfallenden Impulses

Beispiele. Sehen wir uns nun einige typische Beispiele zur Fouriertransformation an, um Intuition zu gewinnen.

■ Der Rechteckimpuls

$$f(x) = \begin{cases} 1 & -a < x \leq a, \\ 0 & \text{sonst} \end{cases}, \quad (3.16)$$

erfüllt $f \in L^1(\mathbb{R})$, sodass die Fouriertransformation wohldefiniert ist:

$$\mathcal{F}f(p) = \int_{-\infty}^{\infty} dx f(x) e^{-ipx} = \int_{-a}^a dx e^{-ipx} = \frac{e^{-ipa} - e^{ipa}}{-ip} = \frac{2 \sin ap}{p}. \quad (3.17)$$

Wenn wir annehmen, dass auch die Rücktransformation in (3.4) existiert, gilt

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dp \frac{2 \sin ap}{p} e^{ipx} = \frac{2}{\pi} \int_0^{\infty} dp \frac{\sin ap}{p} \cos px, \quad (3.18)$$

da der antisymmetrische Anteil bei der Integration wegfällt. Das liefert für $x = 0$ die nützliche Integrationsformel

$$\int_0^{\infty} dp \frac{\sin p}{p} = \frac{\pi}{2}. \quad (3.19)$$

Wie bei der Fourierreihe liefert die rücktransformierte Funktion an den Sprungstellen den Mittelwert der Funktionswerte:

$$f(\pm a) = \frac{1}{\pi} \int_0^{\infty} dp \frac{2 \sin(ap) \cos(ap)}{p} = \frac{1}{\pi} \int_0^{\infty} dp \frac{\sin(2ap)}{p} = \frac{1}{\pi} \int_0^{\infty} dx \frac{\sin x}{x} = \frac{1}{2}. \quad (3.20)$$

Was passiert für $a \rightarrow \infty$? Dann wird $f(x) = 1 \forall x \in \mathbb{R}$, was natürlich keine integrierbare Funktion mehr ist. Die Fouriertransformierte verschwindet dann für alle $p \neq 0$, während sie bei $p = 0$ eine unendlich hohe Spitze hat. Diese “**Dirac’sche Deltafunktion**” werden wir in Kap. 4 im Kontext von Distributionen behandeln.

■ Die **Gaußfunktion** $f(x) = e^{-\alpha x^2}$, $x \in \mathbb{R}$, $\alpha > 0$ ist nicht nur ein Element des $L^1(\mathbb{R})$, sondern auch der Prototyp einer Schwartz-Funktion, die gemeinsam mit allen ihren Ableitungen im Unendlichen schneller als jede Potenz verschwindet. In diesem Fall lauten die L^1 -Norm und die Fouriertransformation von f :

$$\|f\|_1 = \int_{-\infty}^{\infty} dx e^{-\alpha x^2} = \sqrt{\frac{\pi}{\alpha}}, \quad \mathcal{F}f(p) = \int_{-\infty}^{\infty} dx e^{-\alpha x^2} e^{-ipx} = \sqrt{\frac{\pi}{\alpha}} e^{-\frac{p^2}{4\alpha}}. \quad (3.21)$$

Die Fouriertransformierte einer Gaußfunktion ist also wieder eine Gaußfunktion!

Die L^1 -Norm kann man auf verschiedene Arten berechnen; eine Möglichkeit ist die Transformation auf Polarkoordinaten $x = r \cos \varphi$, $y = r \sin \varphi$ und $dx dy = r dr d\varphi$:

$$\begin{aligned} \left(\int_{-\infty}^{\infty} dx e^{-\alpha x^2} \right)^2 &= \int_{-\infty}^{\infty} dx e^{-\alpha x^2} \int_{-\infty}^{\infty} dy e^{-\alpha y^2} = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} dy e^{-\alpha r^2} = \int_0^{\infty} dr r \int_0^{2\pi} d\varphi e^{-\alpha r^2} \\ &= 2\pi \int_0^{\infty} dr r e^{-\alpha r^2} = 2\pi \left[-\frac{1}{2\alpha} e^{-\alpha r^2} \right]_0^{\infty} = \frac{\pi}{\alpha}. \end{aligned} \quad (3.22)$$

Gl (3.21) lässt sich mittels partieller Integration herleiten. Dazu bilden wir die Ableitung

$$(\mathcal{F}f)'(p) = \int_{-\infty}^{\infty} dx \underbrace{(-ix) e^{-\alpha x^2}}_{u'(x)} e^{-ipx} = \left[\underbrace{\frac{i}{2\alpha} e^{-\alpha x^2}}_{u(x)} e^{-ipx} \right]_{x=-\infty}^{\infty} - \frac{p}{2\alpha} \int_{-\infty}^{\infty} dx e^{-\alpha x^2} e^{-ipx} = -\frac{p}{2\alpha} \mathcal{F}f(p).$$

Das ist eine Differenzialgleichung mit der eindeutigen Lösung $\mathcal{F}f(p) = c e^{-\frac{p^2}{4\alpha}}$, wobei $c = \mathcal{F}f(0) = \sqrt{\pi/\alpha}$. Ein alternativer Beweis verläuft über die Funktionentheorie; dazu schreiben wir

$$\mathcal{F}f(p) = e^{-\frac{p^2}{4\alpha}} \int_{-\infty}^{\infty} dx e^{-\alpha(x + \frac{ip}{2\alpha})^2} = e^{-\frac{p^2}{4\alpha}} \int_{-\infty+i\omega}^{\infty+i\omega} dz e^{-\alpha z^2}, \quad \omega = \frac{p}{2\alpha}. \quad (3.23)$$

Nun können wir den Integralsatz von Cauchy benutzen: $e^{-\alpha z^2}$ hat keine isolierten Singularitäten und verschwindet für $|z| \rightarrow \infty$, daher können wir die Integrationskontur zurück auf die reelle Achse deformieren. Oder anders gesagt: innerhalb einer geschlossenen Kontur verschwindet das Integral über $e^{-\alpha z^2}$,

$$\left(\int_{-R}^{-R+i\omega} + \int_{-R+i\omega}^{R+i\omega} + \int_{R+i\omega}^R + \int_R^{-R} \right) dz e^{-\alpha z^2} = 0, \quad (3.24)$$

und weil der erste und dritte Beitrag im Limes $R \rightarrow \infty$ gegen Null gehen, heben sich die beiden restlichen Terme gerade gegenseitig auf:

$$\int_{-\infty+i\omega}^{\infty+i\omega} dz e^{-\alpha z^2} = \int_{-\infty}^{\infty} dz e^{-\alpha z^2} = \sqrt{\frac{\pi}{\alpha}}. \quad (3.25)$$

Hieraus können wir eine wichtige Eigenschaft der Fouriertransformation ablesen. Schreiben wir $f(x) = e^{-x^2/(4\sigma_x^2)}$ und $\mathcal{F}f(p) \sim e^{-p^2/(4\sigma_p^2)}$, dann gilt $\sigma_x \sigma_p = \frac{1}{2}$. Wenn wir $f(x)$ verbreitern, wird $\mathcal{F}f(p)$ schmaler und umgekehrt, aber das Produkt $\sigma_x \sigma_p$ bleibt konstant. Das ist nichts anderes als die **Heisenbergsche Unschärferelation!** In der Quantenmechanik wird der Zustand eines Systems durch eine Wellenfunktion $\psi(x)$ im Ortsraum bzw. $\mathcal{F}\psi(p)$ im Impulsraum beschrieben, die durch eine Fouriertransformation auseinander hervorgehen.

Die Breite der Wellenfunktion zeigt an, wie sehr der Ort des Teilchens lokalisiert ist bzw. wie scharf sein Impuls bekannt ist. Definiert man

$$\begin{aligned}\sigma_x^2 &= \langle (x - \langle x \rangle)^2 \rangle = \langle x^2 \rangle - \langle x \rangle^2, & \langle x^n \rangle &= \frac{\int_{-\infty}^{\infty} dx x^n |\psi(x)|^2}{\int_{-\infty}^{\infty} dx |\psi(x)|^2}, \\ \sigma_p^2 &= \langle (p - \langle p \rangle)^2 \rangle = \langle p^2 \rangle - \langle p \rangle^2,\end{aligned}\quad (3.26)$$

und analog für $\langle p^n \rangle$, dann gilt allgemein die Relation $\sigma_x \sigma_p \geq \frac{1}{2}$. Der Ort und Impuls eines Teilchens können nicht gleichzeitig beliebig scharf gemessen werden (damit p tatsächlich einen physikalischen Impuls beschreibt, muss man noch einen Faktor \hbar einfügen). Die mathematische Grundlage dieser Relation ist also die Fouriertransformation!

■ Auch ein **exponentiell abfallender Impuls** $f(x) = e^{-\alpha|x|}$, $x \in \mathbb{R}$, $\alpha > 0$ liegt im $L^1(\mathbb{R})$ und hat die Norm

$$\|f\|_1 = \int_{-\infty}^{\infty} dx e^{-\alpha|x|} = 2 \int_0^{\infty} dx e^{-\alpha x} = \frac{2}{-\alpha} e^{-\alpha x} \Big|_0^{\infty} = \frac{2}{\alpha}. \quad (3.27)$$

Seine Fouriertransformierte lautet

$$\begin{aligned}\mathcal{F}f(p) &= \int_{-\infty}^{\infty} dx e^{-\alpha|x|} e^{-ipx} = \int_{-\infty}^0 dx e^{-ipx+\alpha x} + \int_0^{\infty} dx e^{-ipx-\alpha x} \\ &= \frac{e^{\alpha x - ipx}}{\alpha - ip} \Big|_{-\infty}^0 - \frac{e^{-\alpha x - ipx}}{\alpha + ip} \Big|_0^{\infty} = \frac{1}{\alpha - ip} + \frac{1}{\alpha + ip} = \frac{2\alpha}{\alpha^2 + p^2}.\end{aligned}\quad (3.28)$$

■ Abschließend betrachten wir einen komplexen **Wellenzug** der Länge $2N\pi/\omega$:

$$f(x) = \begin{cases} e^{i\omega x} & |x| < \frac{N\pi}{\omega}, \\ 0 & \text{sonst.} \end{cases} \quad (3.29)$$

Die Fouriertransformation liefert ein ähnliches Resultat wie in (3.17), jetzt allerdings mit einem Maximum bei $p = \omega$:

$$\mathcal{F}f(p) = \int_{-N\pi/\omega}^{N\pi/\omega} dx e^{-i(p-\omega)x} = \frac{2 \sin \left[(p-\omega) \frac{N\pi}{\omega} \right]}{p-\omega}. \quad (3.30)$$

Fourier-Kosinus- und Sinus-Transformation. Da man laut (2.62) jede Funktion in einen geraden und einen ungeraden Anteil zerlegen kann, tragen dann auch jeweils nur die geraden (Kosinus) und ungeraden (Sinus) Anteile von e^{-ipx} zum Integral bei:

$$\begin{aligned}\mathcal{F}f(p) &= \int_{-\infty}^{\infty} dx f_+(x) e^{-ipx} + \int_{-\infty}^{\infty} dx f_-(x) e^{-ipx} = \int_{-\infty}^{\infty} dx f_+(x) \cos px - i \int_{-\infty}^{\infty} dx f_-(x) \sin px \\ &= 2 \int_0^{\infty} dx f_+(x) \cos px - 2i \int_0^{\infty} dx f_-(x) \sin px =: \mathcal{F}_c f_+(p) - i \mathcal{F}_s f_-(p).\end{aligned}\quad (3.31)$$

Dabei wird \mathcal{F}_c als Fourier-Kosinus-Transformation und \mathcal{F}_s als Fourier-Sinus-Transformation bezeichnet.

Rechenregeln und Ableitungen. Wir können einige nützliche Skalierungsregeln für die Fouriertransformation herleiten, wenn wir die Transformation auf $f_a(x) = f(x + a)$ bzw. $f_\lambda(x) = f(\lambda x)$ mit $a \in \mathbb{R}$ und $\lambda > 0$ anwenden:

$$\mathcal{F}(e^{iax} f)(p) = \int_{-\infty}^{\infty} dx f(x) e^{-i(p-a)x} = \mathcal{F}f(p - a), \quad (3.32)$$

$$(\mathcal{F}f_a)(p) = \int_{-\infty}^{\infty} dx f(x + a) e^{-ipx} = e^{iap} \int_{-\infty}^{\infty} dy f(y) e^{-ipy} = e^{iap} \mathcal{F}f(p), \quad (3.33)$$

$$(\mathcal{F}f_\lambda)(p) = \int_{-\infty}^{\infty} dx f(\lambda x) e^{-ipx} = \frac{1}{\lambda} \int_{-\infty}^{\infty} dy f(y) e^{-\frac{ipy}{\lambda}} = \frac{1}{\lambda} \mathcal{F}f\left(\frac{p}{\lambda}\right). \quad (3.34)$$

Die ersten beiden Gleichungen besagen, dass eine Translation im Argument einen Phasenfaktor in der Fouriertransformierten bewirkt. Wir können außerdem die Fouriertransformation innerhalb eines Integrals umschaukeln:

$$\begin{aligned} \int_{-\infty}^{\infty} dx g(x) \mathcal{F}f(x) &= \int_{-\infty}^{\infty} dx g(x) \int_{-\infty}^{\infty} dy f(y) e^{-ixy} = \int_{-\infty}^{\infty} dy f(y) \int_{-\infty}^{\infty} dx g(x) e^{-ixy} \\ &= \int_{-\infty}^{\infty} dy \mathcal{F}g(y) f(y) = \int_{-\infty}^{\infty} dx \mathcal{F}g(x) f(x). \end{aligned} \quad (3.35)$$

Hier haben wir den **Satz von Fubini** benutzt: bei einem Doppelintegral $\iint dx dy h(x, y)$ ist die Vertauschung der Integrationen erlaubt, wenn das Integral $\iint dx dy |h(x, y)|$ existiert – das ist hier erfüllt, da $f, g \in L^1(\mathbb{R})$.

Die Fouriertransformation verdankt ihre Bedeutung wesentlich dem Umstand, dass sie **Ableitungen** in *Multiplikationen* verwandelt. Das ist äußerst nützlich für das Lösen von partiellen Differenzialgleichungen, und in der Quantenmechanik basiert darauf die Identifikation des Ableitungsoperators im Ortsraum mit dem Impulsoperator:

SATZ:

- Sei $f \in L^1(\mathbb{R}) \cap C^{(m)}(\mathbb{R})$, d.h. f ist absolut integrierbar und m -fach stetig differenzierbar, und die m -te Ableitung $f^{(m)} \in L^1(\mathbb{R})$. Dann gilt

$$\mathcal{F}(f^{(m)})(p) = (ip)^m \mathcal{F}f(p). \quad (3.36)$$

- Sei $x^m f \in L^1(\mathbb{R})$. Dann ist $\mathcal{F}f \in C^{(m)}(\mathbb{R})$ und es gilt

$$(\mathcal{F}f)^{(m)}(p) = \mathcal{F}((-ix)^m f)(p). \quad (3.37)$$

Zum Beweis der ersten Aussage für $m = 1$ berechnen wir einfach die Fouriertransformation der Ableitung:

$$\mathcal{F}(f')(p) = \int_{-\infty}^{\infty} dx f'(x) e^{-ipx} = f(x) e^{-ipx} \Big|_{-\infty}^{\infty} + ip \int_{-\infty}^{\infty} dx f(x) e^{-ipx} = ip \mathcal{F}f(p). \quad (3.38)$$

Der Randterm verschwindet, da $f \in L^1(\mathbb{R})$ und somit $\lim_{|x| \rightarrow \infty} f(x) = 0$. Der allgemeine Fall folgt induktiv. Für die zweite Aussage im Fall $m = 1$,

$$(\mathcal{F}f)'(p) = \int_{-\infty}^{\infty} dx f(x) (-ix) e^{-ipx} = \mathcal{F}(-ixf)(p), \quad (3.39)$$

ist wegen $|\frac{d}{dp}(f(x)e^{-ipx})| = |xf(x)| \in L^1(\mathbb{R})$ die Vertauschbarkeit der Ableitung (die man ja als Limes schreiben kann) mit dem Integral nach dem Satz von Lebesgue von der majorisierten Konvergenz gewährleistet.

Fouriertransformation im Schwartz-Raum. Bis jetzt sind wir der Frage ausgewichen, unter welchen Bedingungen die Fouriertransformation eine **Umkehrtransformation** wie in (3.4) besitzt. Bisher wissen wir nur, dass die Fouriertransformierte $\mathcal{F}f$ einer Funktion $f \in L^1(\mathbb{R})$ stetig und beschränkt ist und dass $\mathcal{F}f$ für $|p| \rightarrow \infty$ verschwindet. Damit muss $\mathcal{F}f$ aber nicht mehr im $L^1(\mathbb{R})$ liegen (z.B. könnte die Fouriertransformierte im Unendlichen ja wie $1/p$ abfallen), und somit ist nicht klar, ob die Rücktransformation existiert. Wenn wir die beiden Ausdrücke in (3.4) einfach ineinander einsetzen, erhalten wir mit

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} dp g(p) e^{ipx} = \frac{1}{2\pi} \int_{-\infty}^{\infty} dp e^{ipx} \int_{-\infty}^{\infty} dy f(y) e^{-ipy} = \frac{1}{2\pi} \int_{-\infty}^{\infty} dy f(y) \int_{-\infty}^{\infty} dp e^{ip(x-y)} \quad (3.40)$$

kein sinnvolles Resultat, denn die Vertauschung der beiden Integrationen ist nach dem Satz von Fubini ja gar nicht erlaubt. Später werden wir dieser Formel durch eine Erweiterung des Definitionsbereichs der Fouriertransformation einen Sinn verleihen, was uns auf die Theorie der Distributionen führen wird.

Gehen wir zunächst den umgekehrten Weg und schränken die Definitionsmenge der Fouriertransformation ein, bis wir auf eine Menge gelangen, auf der die Umkehrung wohldefiniert ist. Das führt auf die folgende von [Laurent Schwartz](#) (1915–2002) eingeführte Definition:

DEFINITION: Der **Schwartz-Raum** ist die Menge der beliebig oft stetig differenzierbaren (glatten) Funktionen, die gemeinsam mit ihren Ableitungen im Unendlichen schneller als jede Potenz verschwinden:

$$S(\mathbb{R}) = \left\{ f \in C^\infty(\mathbb{R}) \mid \lim_{|x| \rightarrow \infty} x^\alpha f^{(\beta)}(x) = 0 \quad \forall \alpha, \beta \in \mathbb{N}_0 \right\}. \quad (3.41)$$

Der Schwartz-Raum ist ein Unterraum des $L^2(\mathbb{R})$, allerdings deutlich kleiner als dieser. Beispielsweise ist jeder Rechteckimpuls quadratintegrierbar, auch jede stückweise stetige Funktion, die außerhalb eines endlichen Intervalls $[a, b]$ verschwindet, oder Funktionen der Art $e^{-|x|}$, aber keine dieser Funktionen ist eine Schwartz-Funktion. Typische Schwartz-Funktionen sind die Gaußfunktion e^{-x^2} sowie die glatten $C_0^\infty(\mathbb{R})$ -Funktionen mit kompaktem Träger, wie z.B. $e^{1/(x^2-1)}$ für $|x| < 1$ und sonst Null. Falls f eine Schwartz-Funktion ist, sind auch $x^\alpha f$ und $f^{(\beta)}$ mit beliebigen $\alpha, \beta \in \mathbb{N}_0$ wieder Schwartz-Funktionen. Es gilt $C_0^\infty \subset S \subset C^\infty$, und man kann zeigen, dass für $1 \leq p < \infty$ der Schwartz-Raum $S(\mathbb{R})$ dicht im Raum der p -integrierbaren Funktionen $L^p(\mathbb{R})$ liegt.

SATZ: Die Fouriertransformation $\mathcal{F} : S(\mathbb{R}) \rightarrow S(\mathbb{R})$ ist eine bijektive Abbildung im Schwartz-Raum, und ihre **Umkehrtransformation** lautet

$$(\mathcal{F}^{-1}g)(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dp g(p) e^{ipx}. \quad (3.42)$$

Außerdem gilt die **Formel von Plancherel**:

$$\langle \mathcal{F}f, \mathcal{F}g \rangle_{L^2} = 2\pi \langle f, g \rangle_{L^2}. \quad (3.43)$$

Dass die Fouriertransformation einer Schwartz-Funktion wieder eine Schwartz-Funktion ergibt, sieht man so: $f \in S(\mathbb{R})$ ist eine stärkere Bedingung als die Voraussetzungen in (3.36–3.37), daher sind die dort getroffenen Aussagen für die Ableitungen sicher auch für Schwartz-Funktionen gültig. Dann ist aber

$$p^\beta (\mathcal{F}f)^{(\alpha)}(p) \stackrel{(3.37)}{=} p^\beta (-i)^\alpha \mathcal{F}(x^\alpha f)(p) = \frac{(-i)^\alpha}{i^\beta} (ip)^\beta \mathcal{F}(x^\alpha f)(p) \stackrel{(3.36)}{=} \frac{(-i)^\alpha}{i^\beta} \mathcal{F}\left((x^\alpha f)^{(\beta)}\right)(p). \quad (3.44)$$

Da auf der rechten Seite wieder eine Fouriertransformation steht, verschwindet dieser Ausdruck wegen (3.11) im Limes $|p| \rightarrow \infty$; somit ist $\mathcal{F}f$ eine Schwartz-Funktion. Die Existenz der Fouriertransformierten von $(x^\alpha f)^{(\beta)}$ für beliebige $\alpha, \beta \in \mathbb{N}_0$ ist für alle Schwartz-Funktionen garantiert.

Um zu beweisen, dass die Abbildung bijektiv ist, d.h. dass für jedes $g \in S(\mathbb{R})$ genau ein $f \in S(\mathbb{R})$ mit $\mathcal{F}f = g$ existiert, starten wir von (3.42) und schieben eine Gaußfunktion (3.21) ein:

$$h(x) := e^{-\frac{x^2}{2}} \Rightarrow h(0) = 1, \quad \mathcal{F}h(p) = \sqrt{2\pi} e^{-\frac{p^2}{2}} = \sqrt{2\pi} h(p), \quad \int_{-\infty}^{\infty} dp h(p) = \sqrt{2\pi}. \quad (3.45)$$

Dann gilt mit den Skalierungseigenschaften aus (3.33–3.34):

$$\begin{aligned} \frac{1}{2\pi} \int_{-\infty}^{\infty} dp \mathcal{F}f(p) e^{ipx} &= \lim_{\lambda \rightarrow 0} \frac{1}{2\pi} \int_{-\infty}^{\infty} dp \underbrace{\mathcal{F}f(p) e^{ipx}}_{\mathcal{F}f_\lambda(p)} \underbrace{h(\lambda p)}_{h_\lambda(p)} \stackrel{(3.35)}{=} \lim_{\lambda \rightarrow 0} \frac{1}{2\pi} \int_{-\infty}^{\infty} dp f(p+x) \underbrace{\mathcal{F}h_\lambda(p)}_{\mathcal{F}h(p/\lambda)/\lambda} \\ &= \lim_{\lambda \rightarrow 0} \frac{1}{2\pi} \int_{-\infty}^{\infty} dp f(\lambda p+x) \underbrace{\mathcal{F}h(p)}_{\sqrt{2\pi} h(p)} = f(x) \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dp h(p) = f(x). \end{aligned} \quad (3.46)$$

Zur Vertauschung des Limes mit dem Integral haben wir zweimal den Satz von Lebesgue zur majorisierten Konvergenz benutzt. Damit lässt sich f aus $\mathcal{F}f$ rekonstruieren, $\mathcal{F}^{-1}\mathcal{F}f = f$, womit die Abbildung injektiv ist. Dass die Fouriertransformation auch surjektiv ist, sieht man durch ihre wiederholte Anwendung:

$$\mathcal{F}^2 f(p) = \int_{-\infty}^{\infty} dx \mathcal{F}f(x) e^{-ipx} = 2\pi \mathcal{F}^{-1}\mathcal{F}f(-p) = 2\pi f(-p) \Rightarrow \frac{\mathcal{F}^4 f(p)}{(2\pi)^2} = f(p). \quad (3.47)$$

Wählt man $f = \mathcal{F}^3 g / (2\pi)^2 \in S(\mathbb{R})$, folgt aus $\mathcal{F}^{-1}\mathcal{F}f = f$ auch $\mathcal{F}\mathcal{F}^{-1}g = g$, und somit ist die Fouriertransformation bijektiv. Die Formel von Plancherel lässt sich dann einfach durch Einsetzen beweisen.

Der verbleibende Schönheitsfehler ist, dass wir uns weniger für den Schwartz-Raum interessieren und vielmehr für den Hilbertraum $L^2(\mathbb{R})$ der Quantenmechanik – wir wollen Hilbertraumvektoren mit Hilfe der Fouriertransformation vor- und rückwärts transformieren. Für unbeschränkte Definitionsbereiche ist $L^2(a, b) \subset L^1(a, b)$ nicht länger erfüllt, denn es gibt Funktionen $\in L^1(\mathbb{R})$, die nicht im $L^2(\mathbb{R})$ enthalten sind als auch umgekehrt. Die Erweiterung auf den $L^2(\mathbb{R})$ ist aber möglich, da der Schwartz-Raum $S(\mathbb{R})$ dicht im $L^2(\mathbb{R})$ liegt; somit kann man Funktionen aus $L^2(\mathbb{R})$ durch Funktionen aus $S(\mathbb{R})$ approximieren und anschließend einen Grenzübergang durchführen.

Fouriertransformation mehrerer Variablen. Alle bisher getätigten Aussagen lassen sich ohne Weiteres auf den Definitionsbereich \mathbb{R}^n erweitern:

$$f(\mathbf{x}) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} d^n p g(\mathbf{p}) e^{i\mathbf{p}\cdot\mathbf{x}}, \quad g(\mathbf{p}) = \int_{\mathbb{R}^n} d^n x f(\mathbf{x}) e^{-i\mathbf{p}\cdot\mathbf{x}}. \quad (3.48)$$

Zur Verallgemeinerung des Schwartz-Raums auf $S(\mathbb{R}^n)$ muss man in der Definition (3.41) folgende Ersetzungen treffen:

$$x^\alpha \rightarrow \prod_{i=1}^n x_i^{\alpha_i}, \quad f^{(\alpha)} \rightarrow \prod_{i=1}^n \frac{\partial^{\alpha_i}}{\partial x_i^{\alpha_i}} f, \quad \alpha = \sum_{i=1}^n \alpha_i, \quad \alpha_i \in \mathbb{N}_0. \quad (3.49)$$

Als Beispiel berechnen wir die Fouriertransformierte des **Yukawa-Potentials**

$$V(r) = \frac{e^{-mr}}{4\pi r}, \quad m > 0, \quad (3.50)$$

das in der Kern- und Elementarteilchenphysik eine wichtige Rolle spielt und z.B. die Wechselwirkung zwischen zwei Nukleonen durch den Austausch von Pionen beschreibt. Die Fouriertransformierte lautet

$$\hat{V}(\mathbf{p}) = \frac{1}{4\pi} \int d^3 x \frac{e^{-mr}}{r} e^{-i\mathbf{p}\cdot\mathbf{x}} = \frac{1}{4\pi} \int_0^\infty dr r^2 \int_{-1}^1 d(\cos\theta) \int_0^{2\pi} d\varphi \frac{e^{-mr}}{r} e^{-ipr \cos\theta}. \quad (3.51)$$

Hier haben wir Kugelkoordinaten verwendet und ohne Beschränkung der Allgemeinheit $\mathbf{p} = p \mathbf{e}_3$ gesetzt, sodass $\mathbf{p} \cdot \mathbf{x} = pr \cos\theta$. Damit können wir die Integration ausführen:

$$\begin{aligned} \hat{V}(\mathbf{p}) &= \frac{1}{2} \int_0^\infty dr r e^{-mr} \int_{-1}^1 dz e^{-iprz} = \frac{1}{2ip} \int_0^\infty dr e^{-mr} (e^{ipr} - e^{-ipr}) \\ &= \frac{1}{2ip} \int_0^\infty dr (e^{(-m+ip)r} - e^{-(m+ip)r}) = \frac{1}{2ip} \left[\frac{e^{(-m+ip)r}}{-m+ip} + \frac{e^{-(m+ip)r}}{m+ip} \right]_0^\infty \\ &= -\frac{1}{2ip} \left(\frac{1}{m+ip} - \frac{1}{m-ip} \right) = \frac{1}{p^2 + m^2} = \frac{1}{\mathbf{p}^2 + m^2}. \end{aligned} \quad (3.52)$$

Als Grenzfall für $m \rightarrow 0$ erhalten wir die Fouriertransformation des **Coulombpotentials** $1/(4\pi r)$ zurück, $\hat{V}(\mathbf{p}) = 1/\mathbf{p}^2$, die für sich genommen nicht wohldefiniert ist.

Lösen von Differenzialgleichungen. Fouriertransformationen sind insbesondere nützlich zur Lösung von **partiellen Differenzialgleichungen (PDGs)**. Dabei wird die Gleichung zusammen mit ihren Anfangsbedingungen fouriertransformiert, was unter Umständen auf eine gewöhnliche Differenzialgleichung führt, die einfach lösbar ist. Durch die Rücktransformation erhält man dann die Lösung der PDG. Auf diese Weise kann man Lösungen zu konkreten PDGs finden, aber in der Regel keine Aussagen bezüglich der Eindeutigkeit und stetigen Abhängigkeit der Lösung von den Anfangsdaten treffen.

Als Beispiel betrachten wir die **Wellengleichung** als Prototyp einer hyperbolischen PDG,

$$\left(\frac{\partial^2}{\partial t^2} - c^2 \frac{\partial^2}{\partial x^2} \right) u(x, t) = 0 \quad \forall x \in \mathbb{R}, t > 0 \quad (3.53)$$

mit den Anfangsbedingungen

$$u(x, 0) = u_0(x), \quad \frac{\partial u}{\partial t}(x, 0) = u_1(x) \quad \forall x \in \mathbb{R}. \quad (3.54)$$

Damit die Fouriertransformation existiert, setzen wir voraus, dass sowohl die Lösung u als auch die Anfangswerte u_0 und u_1 hinreichend regulär sind.

Zunächst wenden wir für ein fixes $t > 0$ die Fouriertransformation bezüglich der Ortsvariable x an. Bezeichnen wir die Fouriertransformierten von $u(x, t)$, $u_0(x)$ und $u_1(x)$ mit $U(p, t)$, $U_0(p)$ und $U_1(p)$, dann folgt aus (3.36) sofort die transformierte Gleichung

$$\left(\frac{\partial^2}{\partial t^2} + c^2 p^2 \right) U(p, t) = 0. \quad (3.55)$$

Das ist eine gewöhnliche Differenzialgleichung, deren Lösung bekannt ist (und durch Einsetzen überprüft werden kann):

$$U(p, t) = A(p) \sin(cpt) + B(p) \cos(cpt). \quad (3.56)$$

Setzen wir die Anfangsbedingungen ein,

$$U(p, 0) = U_0(p) = B(p), \quad \frac{\partial U}{\partial t}(p, 0) = U_1(p) = cpA(p), \quad (3.57)$$

dann lautet die Lösung

$$U(p, t) = \frac{U_1(p)}{cp} \sin(cpt) + U_0(p) \cos(cpt). \quad (3.58)$$

Die Lösung für $u(x, t)$ der ursprünglichen Gleichung folgt aus der inversen Fouriertransformation. Wegen der Linearität der Transformation können wir sie für beide Summanden separat durchführen. Für den zweiten Summanden ergibt sich aus der Skalierungseigenschaft (3.33):

$$\mathcal{F}^{-1} [U_0(p) \cos(cpt)] = \mathcal{F}^{-1} \left[U_0(p) \frac{e^{icpt} + e^{-icpt}}{2} \right] = \frac{1}{2} (u_0(x + ct) + u_0(x - ct)). \quad (3.59)$$

Um den ersten Summanden auf eine analoge Form zu bringen, müssen wir die Umkehrtransformation von $U_1(p)/(cp)$ finden. Dazu definieren wir

$$h(x) = \int_{-\infty}^x dy u_1(y) \quad \Rightarrow \quad h'(x) = u_1(x) \quad \Rightarrow \quad U_1(p) = \mathcal{F}h'(p) \stackrel{(3.36)}{=} ip \mathcal{F}h(p) \quad (3.60)$$

und bekommen somit

$$\frac{U_1(p)}{cp} \sin(cpt) = \frac{i}{c} \mathcal{F}h(p) \frac{e^{icpt} - e^{-icpt}}{2i} \stackrel{(3.33)}{=} \frac{1}{2c} \mathcal{F} (h(x + ct) - h(x - ct)). \quad (3.61)$$

Damit lautet die Lösung der Wellengleichung

$$u(x, t) = \frac{1}{2} (u_0(x + ct) + u_0(x - ct)) + \frac{1}{2c} \int_{x-ct}^{x+ct} dy u_1(y), \quad (3.62)$$

was auch als **d'Alembert'sche Lösungsformel** bekannt ist.

3.2 Laplace-Transformation

Die Laplacetransformation (3.8) wurde von **Pierre-Simon Laplace** (1749–1827) zunächst im Zusammenhang mit der Wahrscheinlichkeitstheorie betrachtet. Hier unterscheiden sich sowohl der Integralkern als auch der Definitionsbereich von der Fouriertransformation, weswegen auch die Menge der Funktionen, die eine Laplace-Transformation erlauben, eine andere ist:

DEFINITION: Eine Funktion $f : [0, \infty) \rightarrow \mathbb{C}$ heißt **exponentiell beschränkt** mit Schranke $\gamma \in \mathbb{R}$, falls eine Konstante $c \geq 0$ existiert, sodass

$$|f(t)| \leq c e^{\gamma t} \quad \forall t \in [0, \infty). \quad (3.63)$$

Dann ist die **Laplace-Transformation** folgendermaßen definiert:

$$\mathcal{L}f(p) = \int_0^{\infty} dt f(t) e^{-pt}, \quad p \in \mathbb{C}, \operatorname{Re} p > \gamma. \quad (3.64)$$

Die exponentielle Beschränktheit garantiert, dass das Integral existiert:

$$\int_0^{\infty} dt |f(t) e^{-pt}| = \int_0^{\infty} dt |f(t)| e^{-(\operatorname{Re} p)t} \leq c \int_0^{\infty} dt e^{-(\operatorname{Re} p - \gamma)t} = \frac{c}{\operatorname{Re} p - \gamma} < \infty. \quad (3.65)$$

Eigenschaften der Laplace-Transformation. Aus der Definition folgen sofort einige nützliche Eigenschaften. Durch Einsetzen sieht man, dass \mathcal{L} ein linearer Operator ist, d.h. es gilt $\mathcal{L}(f+g) = \mathcal{L}f + \mathcal{L}g$ und $\mathcal{L}(\lambda f) = \lambda \mathcal{L}f$. Aus (3.65) folgt, dass $\mathcal{L}f(p)$ für $\operatorname{Re} p \rightarrow \infty$ verschwindet. Ferner gilt

$$\mathcal{L}(t^n f(t))(p) = \int_0^{\infty} dt t^n f(t) e^{-pt} = (-1)^n \frac{d^n}{dp^n} \int_0^{\infty} dt f(t) e^{-pt} = (-1)^n \frac{d^n}{dp^n} \mathcal{L}f(p), \quad (3.66)$$

$$\mathcal{L}(f(t-a))(p) = \int_0^{\infty} dt f(t-a) e^{-pt} = \int_{-a}^{\infty} du f(u) e^{-p(a+u)} = e^{-ap} \mathcal{L}f(p), \quad (3.67)$$

$$\mathcal{L}(f(at))(p) = \int_0^{\infty} dt f(at) e^{-pt} = \frac{1}{a} \int_0^{\infty} du f(u) e^{-\frac{pu}{a}} = \frac{1}{a} \mathcal{L}f\left(\frac{p}{a}\right), \quad (3.68)$$

$$\mathcal{L}(e^{at} f(t))(p) = \int_0^{\infty} dt f(t) e^{-(p-a)t} = \mathcal{L}f(p-a). \quad (3.69)$$

Hier haben wir eine unsaubere Notation verwendet, um nicht für jede etwas andere Funktion ein neues Symbol einführen zu müssen; eigentlich sollten wir $g(t) = t^n f(t)$ und $\mathcal{L}g(p) = \dots$ usw. schreiben. In (3.67) benötigen wir die Funktionswerte $f(u)$ für $u < 0$, doch weil f nur auf der positiven reellen Achse $[0, \infty)$ definiert ist, benutzt man die Konvention, dass die Funktion für negative u durch Null fortgesetzt wird.

Mit den obigen Relationen bekommt man schnell einen Grundstock an Rechenregeln, die für das praktische Arbeiten nützlich sind:

$$f(t) = 1 \quad \Rightarrow \quad \mathcal{L}f(p) = \int_0^{\infty} dt e^{-pt} = \frac{1}{p}, \quad (3.70)$$

$$f(t) = t^n \quad \stackrel{(3.66)}{\Rightarrow} \quad \mathcal{L}f(p) = (-1)^n \frac{d^n}{dp^n} \frac{1}{p} = \frac{n!}{p^{n+1}} = \frac{\Gamma(n+1)}{p^{n+1}}, \quad (3.71)$$

$$f(t) = t^n e^{at} \quad \stackrel{(3.69)}{\Rightarrow} \quad \mathcal{L}f(p) = \mathcal{L}(t^n)(p-a) = \frac{n!}{(p-a)^{n+1}}, \quad (3.72)$$

$$f(t) = e^{iat} \quad \stackrel{(3.69)}{\Rightarrow} \quad \mathcal{L}f(p) = \mathcal{L}(1)(p-ia) = \frac{1}{p-ia}. \quad (3.73)$$

Mit der Gammafunktion in (3.71) lässt sich die Laplace-Transformierte für t^n auch auf reelle $n > -1$ fortsetzen. Wegen der Linearität der Transformation kann man sofort die Laplace-Transformierten eine Reihe anderer Funktionen hinschreiben, z.B.

$f(t)$	$\cos at$	$\sin at$	e^{at}	$\cosh at$	$\sinh at$
$\mathcal{L}f(p)$	$\frac{p}{p^2 + a^2}$	$\frac{a}{p^2 + a^2}$	$\frac{1}{p-a}$	$\frac{p}{p^2 - a^2}$	$\frac{a}{p^2 - a^2}$

Auch für die später in Kap. 4.1 behandelten Stufen- und Deltafunktionen lassen sich die Laplace-Transformierten angeben:

$$f(t) = \delta(t-a) \quad \Rightarrow \quad \mathcal{L}f(p) = e^{-pa} \quad (a \geq 0), \quad (3.74)$$

$$f(t) = \Theta(t-a) \quad \Rightarrow \quad \mathcal{L}f(p) = \frac{e^{-pa}}{p} \quad (a > 0). \quad (3.75)$$

Wie sehen die Laplace-Transformationen von Ableitungen aus? Falls f stetig differenzierbar und auch die Ableitung f' vom exponentiellen Typ ist, bekommen wir

$$\mathcal{L}(f')(p) = \int_0^{\infty} dt f'(t) e^{-pt} = f(t) e^{-pt} \Big|_0^{\infty} + p \int_0^{\infty} dt f(t) e^{-pt} = p \mathcal{L}f(p) - f(0), \quad (3.76)$$

wobei der Randterm für $\operatorname{Re} p > \gamma$ wieder verschwindet. Die zweite Ableitung gewinnt man dann iterativ,

$$\mathcal{L}(f'')(p) = p \mathcal{L}(f')(p) - f'(0) = p^2 \mathcal{L}f(p) - pf(0) - f'(0), \quad (3.77)$$

bis hin zur n -ten Ableitung:

$$\mathcal{L}(f^{(n)})(p) = p^n \mathcal{L}f(p) - \sum_{k=0}^{n-1} f^{(k)}(0) p^{n-1-k}, \quad (3.78)$$

vorausgesetzt, f ist n -fach stetig differenzierbar und alle $f^{(n)}$ sind vom exponentiellen Typ. Ähnlich wie bei der Fouriertransformation verwandelt die Laplacetransformation Ableitungen in Multiplikationen, allerdings mit zusätzlichen Subtraktionen bei $t = 0$.

Lösen von Differenzialgleichungen. Die letzte Beobachtung führt, wie bei der Fouriertransformation, zu einem eleganten Verfahren zur Lösung von Differenzialgleichungen: durch Laplace-Transformation erhält man eine rein algebraische Gleichung in der Variablen p , die man nach $\mathcal{L}f(p)$ auflösen kann; die Rücktransformation zu $f(t)$ erfolgt dann z.B. mit Hilfe von Tabellen. Sehen wir uns als Beispiel folgende Differenzialgleichung an:

$$f''(t) + a_1 f'(t) + a_0 f(t) = b(t). \quad (3.79)$$

Bezeichnen wir die Laplace-Transformierten mit $F(p) = \mathcal{L}f(p)$ und $B(p) = \mathcal{L}b(p)$, lautet die transformierte Gleichung

$$\begin{aligned} p^2 F - p f(0) - f'(0) + a_1 (p F - f(0)) + a_0 F &= B \\ \Rightarrow \underbrace{(p^2 + a_1 p + a_0) F}_{=: P(p)} - \underbrace{[(p + a_1) f(0) + f'(0)]}_{=: Q(p)} &= B \end{aligned} \quad (3.80)$$

und die entsprechende Lösung

$$F(p) = \frac{B(p) + Q(p)}{P(p)}. \quad (3.81)$$

Die so entstehenden Ausdrücke sind oft rationale Funktionen, die durch Partialbruchzerlegung auf Terme der Art (3.70–3.73) zurückgeführt werden können. Ansonsten behilft man sich mit Formeltabellen. Betrachten wir als einfaches Beispiel $a_1 = -2$, $a_0 = 1$, $b(t) = t^3 e^t$ mit den Randbedingungen $f(0) = f'(0) = 0$. Dann ist

$$\begin{aligned} P(p) &= (p-1)^2, & B(p) &= \frac{6}{(p-1)^4}, & F(p) &= \frac{6}{(p-1)^6} & \stackrel{(3.72)}{\Rightarrow} & f(t) = \frac{t^5 e^t}{20}. \end{aligned} \quad (3.82)$$

Auch hier liefert das Verfahren die eindeutige Lösung, die den Randbedingungen genügt, macht aber keine allgemeinen Aussagen über die Existenz der Lösungen.

Die **Wärmeleitungsgleichung** ist eine parabolische partielle Differenzialgleichung. Betrachten wir als konkretes Beispiel die Gleichung

$$\left(\frac{\partial}{\partial t} - \frac{\partial^2}{\partial x^2} \right) u(x, t) = -f(x, t) = -(x + t^2) e^{-xt} \quad \forall x \in (0, 1), \quad t > 0, \quad (3.83)$$

mit der Anfangsbedingung $u(x, 0) = 1$ und den Randbedingungen $u(0, t) = 1$, $u(1, t) = e^{-t}$. Bezeichnen wir die Laplace-Transformierte von $u(x, t)$ für ein festes $x \in (0, 1)$ mit $U(x, p)$ und jene von $f(x, t)$ mit $F(x, p)$, dann lautet die transformierte Gleichung:

$$pU - 1 - \frac{\partial^2 U}{\partial x^2} = -\frac{x}{p+x} - \frac{2}{(p+x)^3}, \quad U(0, p) = \frac{1}{p}, \quad U(1, p) = \frac{1}{p+1}. \quad (3.84)$$

Eine partikuläre Lösung dieser gewöhnlichen Differenzialgleichung, die die Randbedingungen erfüllt und somit deren eindeutige Lösung darstellt, ist

$$U(x, p) = \frac{1}{p+x}, \quad (3.85)$$

wovon man sich leicht durch Einsetzen überzeugen kann. Die Rücktransformation ergibt

$$u(x, t) = e^{-xt}. \quad (3.86)$$

Umkehrtransformation. Wie sieht die Umkehrtransformation der Laplace-Transformation allgemein aus? Setzen wir dazu $p = x + iy \in \mathbb{C}$:

$$\mathcal{L}f(x + iy) = \int_0^{\infty} dt f(t) e^{-(x+iy)t} = \int_0^{\infty} dt f(t) e^{-xt} e^{-iyt} = \int_{-\infty}^{\infty} dt g(t) e^{-iyt}, \quad (3.87)$$

wobei

$$g(t) = \begin{cases} f(t) e^{-xt} & t > 0, \\ 0 & t < 0. \end{cases} \quad (3.88)$$

Das ist aber nichts anderes als die Fouriertransformation, von der wir die Umkehrung kennen! Somit können wir schreiben

$$f(t) = g(t) e^{xt} = \frac{1}{2\pi} \int_{-\infty}^{\infty} dy e^{(x+iy)t} \mathcal{L}f(x + iy) = \frac{1}{2\pi i} \int_{x-i\infty}^{x+i\infty} dz \mathcal{L}f(z) e^{zt}, \quad x = \operatorname{Re} p > \gamma.$$

Dieses Integral heißt **Bromwich-Integral** und kann mit den Methoden der Funktionentheorie (Residuenintegration) gelöst werden.

Faltung. Eine häufige Fragestellung ist folgende: Wir kennen die Laplace-Transformierten der Funktionen g und h und würden gerne die Funktion f bestimmen, die die Gleichung $\mathcal{L}f = \mathcal{L}g \mathcal{L}h$ löst, d.h. wir suchen die inverse Laplace-Transformierte des Produkts $\mathcal{L}g \mathcal{L}h$. Dazu schreiben wir

$$\mathcal{L}f(p) = \mathcal{L}g(p) \mathcal{L}h(p) = \int_0^{\infty} d\sigma g(\sigma) e^{-\sigma p} \int_0^{\infty} d\tau h(\tau) e^{-\tau p} = \int_0^{\infty} d\tau \int_0^{\infty} d\sigma g(\sigma) h(\tau) e^{-p(\sigma+\tau)}.$$

Durch eine Variablensubstitution $t = \sigma + \tau$ für festes τ wird das Intervall $0 < \sigma < \infty$ auf $\tau < t < \infty$ abgebildet und wir erhalten

$$\mathcal{L}f(p) = \int_0^{\infty} d\tau \int_{\tau}^{\infty} dt g(t - \tau) h(\tau) e^{-pt}. \quad (3.89)$$

Das Doppelintegral können wir dann aber genausogut für feste t und $0 < \tau < t$ ausführen, siehe Fig. 3.2:

$$\mathcal{L}f(p) = \int_0^{\infty} dt \int_0^t d\tau g(t - \tau) h(\tau) e^{-pt} = \int_0^{\infty} dt f(t) e^{-pt}. \quad (3.90)$$

Damit haben wir das Problem gelöst. Das obige Produkt wird **Faltung** oder **Konvolution** genannt,

$$f(t) = (g * h)(t) = \int_0^t d\tau g(t - \tau) h(\tau), \quad (3.91)$$

und somit gilt $\mathcal{L}g \mathcal{L}h = \mathcal{L}(g * h)$. Das Faltungsintegral ist kommutativ, d.h. es gilt $g * h = h * g$. Typische Anwendungen für die Faltung sind Daten, die bei der Messung durch ein ungenaues Messgerät verfälscht wurden.

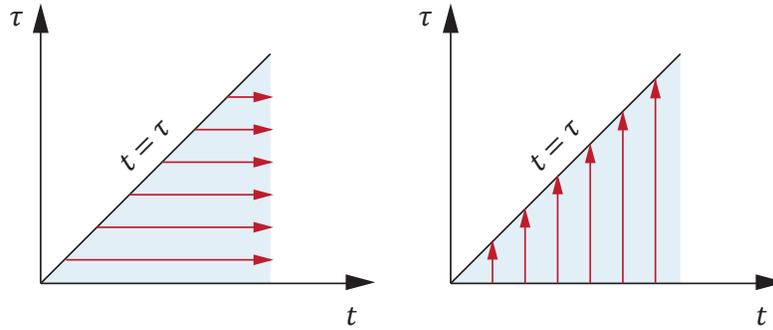


FIG. 3.2: Die Reihenfolge der Integrationen in t und τ kann vertauscht werden

Betrachten wir als Beispiel noch einmal die Differentialgleichung (3.79) mit der Lösung (3.81). Mit $F = \mathcal{L}f$, $B = \mathcal{L}b$ sowie $1/P = \mathcal{L}g$ und $Q/P = \mathcal{L}h$ lautet die Gleichung

$$\mathcal{L}f(p) = \mathcal{L}g(p) \mathcal{L}b(p) + \mathcal{L}h(p) \quad \Rightarrow \quad f(t) = \int_0^t d\tau g(t - \tau) b(\tau) + h(t). \quad (3.92)$$

Aus der Partialbruchzerlegung von $1/P$ und Q/P und Rücktransformation bekommt man die Funktionen g und h . Damit lassen sich die Lösungen von linearen gewöhnlichen Differentialgleichungen mit konstanten Koeffizienten zu jeder Homogenität in Form eines Faltungsintegrals angeben. Für das konkrete Beispiel (3.82) mit $b(t) = t^2 e^t$, $P(p) = (p-1)^2$ und $Q(p) = 0$ folgt $g(t) = t e^t$ und somit wie früher

$$f(t) = \int_0^t d\tau (t - \tau) e^{t-\tau} \tau^3 e^\tau = e^t \int_0^t d\tau \tau^3 (t - \tau) = \frac{t^5 e^t}{20}. \quad (3.93)$$

Auch für die **Fouriertransformation** ist die Faltung definiert:

$$\mathcal{F}g \mathcal{F}h = \mathcal{F}(g * h), \quad (g * h)(t) = \int_{-\infty}^{\infty} d\tau g(t - \tau) h(\tau), \quad (3.94)$$

wie man aus der analogen Vorgehensweise zu früher leicht nachrechnet. Aus der Umkehrung der Fouriertransformation folgt außerdem

$$\mathcal{F}g * \mathcal{F}h = 2\pi \mathcal{F}(gh). \quad (3.95)$$

Kapitel 4

Funktionale und Distributionen

Wir haben im letzten Kapitel die Fouriertransformationen verschiedener Funktionen wie des Rechteckimpulses, der Gaußfunktion und des exponentiell abfallenden Impulses untersucht. Auf $f_i(x=0) = 1/(2\pi)$ normiert lauten diese Funktionen

$$f_1(x) = \begin{cases} \frac{1}{2\pi} & -\frac{1}{\epsilon} < x \leq \frac{1}{\epsilon}, \\ 0 & \text{sonst,} \end{cases} \quad f_2(x) = \frac{1}{2\pi} e^{-\frac{\epsilon x^2}{4}}, \quad f_3(x) = \frac{1}{2\pi} e^{-\epsilon|x|}. \quad (4.1)$$

Deren Fouriertransformierten $g_i = \mathcal{F}f_i$ (wieder ausgewertet am Punkt x) lauten

$$g_1(x) = \frac{\sin \frac{x}{\epsilon}}{\pi x}, \quad g_2(x) = \frac{1}{\sqrt{\epsilon\pi}} e^{-\frac{x^2}{\epsilon}}, \quad g_3(x) = \frac{1}{\pi} \frac{\epsilon}{\epsilon^2 + x^2}. \quad (4.2)$$

In Fig. 3.1 haben wir gesehen, dass diese eine ausgeprägte Spitze entwickeln, die umso höher und schmaler wird, je breiter die ursprüngliche Funktion war (d.h. umso kleiner ϵ ist). Im Limes $\epsilon \rightarrow 0$ wird der Funktionswert bei $x = 0$ unendlich hoch und verschwindet überall sonst. Die Integrale dieser Funktionen sind allerdings konstant und unabhängig von ϵ ,

$$\int_{-\infty}^{\infty} dx g_i(x) = 1. \quad (4.3)$$

Dasselbe gilt für den komplexen Wellenzug (3.29),

$$f_4(x) = \begin{cases} \frac{1}{2\pi} e^{i\omega x} & |x| < \frac{\pi}{\epsilon\omega}, \\ 0 & \text{sonst,} \end{cases} \quad (4.4)$$

nur dass die Fouriertransformierte hier die Spitze bei $x = \omega$ ausbildet:

$$g_4(x) = \frac{1}{\pi} \frac{\sin \left[(x - \omega) \frac{\pi}{\epsilon\omega} \right]}{x - \omega}. \quad (4.5)$$

Das Integral ergibt aber nach wie vor 1, unabhängig von ϵ :

$$\int_{-\infty}^{\infty} dx g_4(x) = \frac{1}{\pi} \int_{-\infty}^{\infty} dx \frac{\sin \left[(x - \omega) \frac{\pi}{\epsilon\omega} \right]}{x - \omega} = \frac{1}{\pi} \int_{-\infty}^{\infty} dx \frac{\sin x}{x} = \frac{2}{\pi} \int_0^{\infty} dx \frac{\sin x}{x} = 1. \quad (4.6)$$

Diese Grenzfälle bieten eine Möglichkeit, die **Dirac'sche Deltafunktion** $\delta(x)$ einzuführen:

$$\int_a^b dx \delta(x - x_0) = \begin{cases} 1 & a < x_0 < b, \\ 0 & \text{sonst.} \end{cases} \quad (4.7)$$

Der Wert des Integrals ändert sich somit sprunghaft, sobald x_0 in den Integrationsbereich zu liegen kommt, egal, wie klein dieses Intervall (a, b) ist. Da nur der Punkt $x = x_0$ für die Integration relevant ist, gilt insbesondere

$$\int_a^b dx f(x) \delta(x - x_0) = \begin{cases} f(x_0) & a < x_0 < b, \\ 0 & \text{sonst.} \end{cases} \quad (4.8)$$

Die Deltafunktion projiziert also den Funktionswert heraus. Offenbar handelt es sich hier nicht um eine Lebesgue-Integration, denn sonst dürften einzelne Punkte vom Maß Null gar nicht beitragen!

Die Deltafunktion wurde ursprünglich von [Paul A. M. Dirac](#) (1902–1984) als kontinuierliche Version des Kronecker-Deltas δ_{ij} eingeführt, um die Normierung von Zustandsvektoren in der Quantenmechanik zu beschreiben. Heute findet sie breite Anwendung in der Physik und Mathematik, etwa bei der Ermittlung der Impulsantwort in der Akustik oder bei der Lösung von Differenzialgleichungen mit Hilfe von Greenfunktionen.

Die mathematische Sinnhaftigkeit dieses Objekts war umstritten, bis [Laurent Schwartz](#) (1915–2002) die Theorie der **Distributionen** entwickelte. Die Idee ist folgende: Eine reellwertige, stetige Funktion f auf \mathbb{R} weist jedem Punkt $x \in \mathbb{R}$ eine reelle Zahl $y = f(x)$ zu, den Funktionswert von f an der Stelle x . Im Prinzip könnte man statt diesem Funktionswert aber auch die reelle Zahl $f[\phi] := \int_{\mathbb{R}} dx f(x) \phi(x)$ für alle ϕ aus einer Menge K von Hilfsfunktionen angeben. Die Rolle dieser Hilfsfunktionen ist sekundär, denn sie dienen lediglich als Vehikel zur Beschreibung von f . In diesem Sinne kann man f als **Funktional** auf K betrachten, d.h. f weist jeder Funktion $\phi \in K$ eine reelle Zahl $f[\phi]$ zu.

Eine solche indirekte Beschreibung einer Funktion ist in den Ingenieurwissenschaften durchaus üblich, z.B. misst eine Messapparatur wie ein Voltmeter nicht den instantanen Wert $f(t_0)$ an einem Zeitpunkt t_0 , sondern einen gewichteten Mittelwert

$$\frac{1}{2T} \int_{t_0-T}^{t_0+T} dt f(t) \phi(t) \quad (4.9)$$

über eine kurze Zeitspanne $2T$, wobei ϕ charakteristisch für die Messapparatur ist. Ein anderes Beispiel ist die Fourier-Sinusreihe (2.65), die einer Funktion f ihre Koeffizienten b_n zuordnet; in diesem Fall sind die Hilfsfunktionen durch $\sin(n\pi x/L)$ gegeben. Daraus könnte man im Prinzip zwar die Funktionswerte $f(x)$ rekonstruieren, aber das ist in der Praxis unerheblich, sofern f eindeutig durch die b_n charakterisiert wird. Wenn wir für die Hilfsfunktionen aus K sehr glatte **Testfunktionen** wählen, erlaubt uns diese Sichtweise außerdem, die Menge der Funktionen f auf singuläre Objekte wie z.B. die Deltafunktion zu erweitern – diese lassen sich dann als Funktionale auf dem Raum solcher glatten Funktionen beschreiben.

4.1 Distributionen

Funktionale. Definieren wir zunächst allgemein den Begriff eines Funktionals:

DEFINITION: Sei \mathbb{V} ein Vektorraum über einem Körper \mathbb{K} (\mathbb{R} oder \mathbb{C}).

- Ein **Funktional** F ist eine Abbildung $F : \mathbb{V} \rightarrow \mathbb{K}$, die jedem $\phi \in \mathbb{V}$ eine Zahl $F[\phi] \in \mathbb{K}$ zuordnet.
- Ein **lineares Funktional** bzw. eine **Linearform** erfüllt zusätzlich die Linearitätseigenschaft $F[\alpha\phi + \beta\psi] = \alpha F[\phi] + \beta F[\psi]$ für beliebige $\phi, \psi \in \mathbb{V}$ und $\alpha, \beta \in \mathbb{K}$.

Die Menge aller linearen Funktionale auf dem Vektorraum \mathbb{V} bildet bezüglich der punktweisen Addition und Multiplikation mit Skalaren wieder einen Vektorraum, den **algebraischen Dualraum** \mathbb{V}^* :

$$\mathbb{V}^* := \{F \mid F : \mathbb{V} \rightarrow \mathbb{K} \text{ und linear}\}. \quad (4.10)$$

Der Dualraum hat in vielen Bereichen der Physik eine wichtige Bedeutung, z.B. im Kontext kontravarianter Vektoren.

Testfunktionen. Eine zentrale Rolle in der Theorie der Distributionen spielen lineare Funktionale auf dem Raum $C_0^\infty(\mathbb{R}^n)$. Obwohl sich alle folgenden Aussagen auf n Dimensionen verallgemeinern lassen, beschränken wir uns einfachheitshalber auf $C_0^\infty(\mathbb{R})$. Dieser Raum besteht aus allen glatten (beliebig oft differenzierbaren) Funktionen mit kompaktem Träger $\Omega \subset \mathbb{R}$, d.h. Funktionen, die außerhalb einer kompakten (beschränkten und abgeschlossenen) Menge Ω Null sind. Es ist gar nicht so einfach, solche Funktionen zu finden; ein typisches Beispiel ist (siehe Fig. 4.1)

$$\phi(x) = \begin{cases} e^{-\frac{1}{1-x^2}} & |x| < 1, \\ 0 & |x| \geq 1. \end{cases} \quad (4.11)$$

Hier verschwinden sowohl ϕ als auch alle Ableitungen $\phi^{(m)}$ bei $x = \pm 1$, denn es gilt

$$\lim_{x^2 \rightarrow 1^-} \frac{1}{(1-x^2)^m} e^{-\frac{1}{1-x^2}} \stackrel{(1-x^2=\frac{1}{z})}{=} \lim_{z \rightarrow \infty} z^m e^{-z} = 0 \quad \forall m. \quad (4.12)$$

Als Nächstes wollen wir den Raum $C_0^\infty(\mathbb{R})$ mit einem **Konvergenzbegriff** versehen. Da wir gerne hätten, dass sich die Glattheitseigenschaften auch in der Konvergenz widerspiegeln, definieren wir: Eine Folge $(\phi_n)_{n=0}^\infty$ mit $\phi_n \in C_0^\infty(\mathbb{R})$ konvergiert dann gegen ϕ , wenn

$$\lim_{n \rightarrow \infty} \max_{x \in \Omega} \left| \frac{\partial^\alpha}{\partial x^\alpha} (\phi_n(x) - \phi(x)) \right| = 0 \quad \forall \alpha \in \mathbb{N}. \quad (4.13)$$

Das entspricht einer gleichmäßigen Konvergenz nicht nur für die Folge selbst, sondern auch alle Ableitungen. Ein solcher Konvergenzbegriff wird durch keine Norm erzeugt, man benötigt dazu lediglich eine Topologie. Ausgestattet mit diesem Konvergenzbegriff nennt man die Menge $C_0^\infty(\mathbb{R})$ den Raum der **Testfunktionen** $\mathcal{D}(\mathbb{R})$ oder kurz \mathcal{D} .

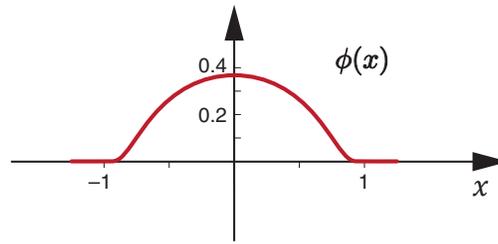


FIG. 4.1: Die Testfunktion ϕ in Gl. (4.11) ist glatt und verschwindet für $|x| \geq 1$.

Mit dieser Definition der Konvergenz können wir den Begriff der **Stetigkeit** auf Funktionale ausdehnen. Ein lineares Funktional F auf \mathcal{D} ist stetig, wenn für jede konvergente Folge $(\phi_n)_{n=0}^\infty \subset \mathcal{D}$ (im Sinne des obigen Konvergenzbegriffs) gilt:

$$\lim_{n \rightarrow \infty} |F[\phi_n] - F[\phi]| = 0. \quad (4.14)$$

Die Konvergenz der Funktionale selbst führt man dagegen einfach auf die Konvergenzeigenschaften im Körper \mathbb{K} zurück, d.h. man sagt: eine Folge von Funktionalen $(F_n)_{n=0}^\infty$ konvergiert dann gegen F , wenn (4.14) gilt – das wird auch als **schwache Konvergenz** bezeichnet.

Distributionen. Mit diesen Vorbemerkungen haben wir alle nötigen Begriffe gesammelt, um eine Distribution zu definieren:

DEFINITION: Eine **Distribution** ist ein stetiges lineares Funktional auf \mathcal{D} . Die Menge aller Distributionen

$$\mathcal{D}' = \{F: \mathcal{D} \rightarrow \mathbb{K} \mid F \text{ linear und stetig}\} \quad (4.15)$$

ist bezüglich der punktweisen Addition und Multiplikation mit Skalaren wieder ein Vektorraum und bildet den **topologischen Dualraum** von \mathcal{D} .

Die hier getroffene Unterscheidung zwischen algebraischem und topologischem Dualraum ist nur für unendlich-dimensionale Vektorräume relevant, da lineare Funktionale auf unendlich-dimensionalen Vektorräumen nicht notwendigerweise stetig sind. Auf einem endlich-dimensionalen Vektorraum ist jedes lineare Funktional stetig. Eine Topologie wird dann durch die euklidische Norm erzeugt, und der topologische Vektorraum ist identisch mit dem algebraischen Vektorraum und isomorph zum Vektorraum selbst.

Um Rechenregeln für Distributionen zu entwickeln, treffen wir noch eine weitere Definition: Eine Funktion f auf \mathbb{R} heißt **lokal integrierbar**, wenn für jeden kompakten Teilbereich $\Omega \subset \mathbb{R}$ gilt:

$$\int_{\Omega} dx |f(x)| < \infty. \quad (4.16)$$

Die Menge dieser Funktionen wird auch als $L_{\text{loc}}^1(\mathbb{R})$ bezeichnet (dabei identifiziert man wie früher wieder alle Funktionen, die sich nur auf Mengen vom Maß Null unterscheiden). Das Verhalten im Unendlichen spielt dann bei lokal integrierbaren Funktionen keine Rolle, z.B. ist die konstante Einsfunktion lokal integrierbar, aber nicht Lebesgue-integrierbar. Alle L^p -Funktionen sind lokal integrierbar, insbesondere auch stetige und stückweise stetige Funktionen. Die Funktion $f(x) = 1/|x|^n$ ist dagegen nur für $n < 1$ lokal integrierbar.

SATZ: Jede lokal integrierbare Funktion f auf \mathbb{R} erzeugt eine Distribution $F \in \mathcal{D}'$ via

$$F[\phi] = \int_{-\infty}^{\infty} dx f(x) \phi(x). \quad (4.17)$$

Der Integralkern der Distribution ist dann durch $f(x)$ gegeben. Lässt sich eine Distribution in dieser Form schreiben, heißt sie **regulär**, ansonsten **singulär**.

Zum Beweis müssen wir die Linearität und Stetigkeit von (4.17) zeigen. Die Linearität ist klar. Für Testfunktionen $\phi, \phi_n \in \mathcal{D}$, die außerhalb eines kompakten Gebiets Ω verschwinden, gilt dann

$$F[\phi_n] - F[\phi] = \int_{\Omega} dx f(x) (\phi_n(x) - \phi(x)) \Rightarrow |F[\phi_n] - F[\phi]| \leq \max_{x \in \Omega} |\phi_n(x) - \phi(x)| \int_{\Omega} dx |f(x)|. \quad (4.18)$$

Sei $(\phi_n)_{n=0}^{\infty}$ nun eine konvergente Folge im Sinne von (4.13). Dann verschwindet $\max_{x \in \Omega} |\phi_n(x) - \phi(x)|$ für $n \rightarrow \infty$, und für lokal integrierbare Funktionen bleibt $\int_{\Omega} dx |f(x)|$ endlich. Somit gilt auch (4.14), und damit ist die Stetigkeit von F bewiesen.

Zwei verschiedene Funktionen $f_1 \neq f_2$, die sich nicht nur auf Mengen vom Maß Null unterscheiden (dann werden sie ohnehin identifiziert) erzeugen auch zwei verschiedene Distributionen $F_1 \neq F_2$. Wenn $f_1(x_0) \neq f_2(x_0)$ nur in einer Umgebung Ω von x_0 gilt, gibt es nämlich sicher irgendein ϕ , sodass

$$F_1[\phi] - F_2[\phi] = \int_{\Omega} dx \underbrace{(f_1(x) - f_2(x))}_{\neq 0} \phi(x) \neq 0. \quad (4.19)$$

In diesem Sinn kann jede lokal integrierbare Funktion f genausogut als Distribution F interpretiert werden. Daher verwendet man üblicherweise auch dasselbe Symbol und schreibt

$$f[\phi] = \int_{-\infty}^{\infty} dx f(x) \phi(x). \quad (4.20)$$

Als Funktion hat f den Funktionswert $f(x)$ für ein $x \in \mathbb{R}$; als Distribution wirkt f auf Testfunktionen im Raum \mathcal{D} . Somit stellen Distributionen Verallgemeinerungen von Funktionen dar. Wegen der Ähnlichkeit mit dem Skalarprodukt verwendet man auch oft die (etwas verwirrende) Schreibweise $f[\phi] = \langle f, \phi \rangle$. Beispiele für reguläre Distributionen sind:

- Eine konstante Funktion $f(x) = c$ erzeugt die reguläre Distribution

$$F_c[\phi] = c \int_{-\infty}^{\infty} dx \phi(x). \quad (4.21)$$

- Die **charakteristische Funktion**

$$I_{\Omega}(x) = \begin{cases} 1 & x \in \Omega \\ 0 & \text{sonst} \end{cases} \quad (4.22)$$

ist stückweise stetig und erzeugt somit eine reguläre Distribution

$$I_{\Omega}[\phi] = \int_{-\infty}^{\infty} dx I_{\Omega}(x) \phi(x) = \int_{\Omega} dx \phi(x). \quad (4.23)$$

Für den Spezialfall $\Omega = [x_0, \infty)$ entspricht $I_{\Omega}(x) = \Theta(x - x_0)$ der **Heaviside-Funktion**.

Der Nutzen des Distributionsbegriffs zeigt sich in den **singulären Distributionen**, die nicht in der Form (4.17) geschrieben werden können. Hier schließt sich der Bogen zur Einleitung dieses Kapitels, denn das klassische Beispiel für eine singuläre Distribution ist die **Delta-Distribution** δ_{x_0} , definiert durch

$$\delta_{x_0}[\phi] = \phi(x_0). \quad (4.24)$$

Sie weist also einfach der Testfunktion ϕ ihren Funktionswert bei x_0 zu. Offenbar handelt es sich um eine Distribution: es gilt die Linearität, $\delta_x[\alpha\phi_1 + \beta\phi_2] = \alpha\phi_1(x) + \beta\phi_2(x)$, und für eine konvergente Folge $(\phi_n)_{n=0}^\infty$ gilt

$$\lim_{n \rightarrow \infty} |\delta_x[\phi_n] - \delta_x[\phi]| = \lim_{n \rightarrow \infty} |\phi_n(x) - \phi(x)| = 0 \quad \forall x, \quad (4.25)$$

was genau der gleichmäßigen Konvergenz entspricht; somit ist auch die Stetigkeit erfüllt.

Die Delta-Distribution ist nicht regulär, sonst müsste es eine lokal integrierbare Funktion f geben, sodass

$$\int_{-\infty}^{\infty} dx f(x) \phi(x) = \phi(0) \quad \forall \phi \in \mathcal{D}. \quad (4.26)$$

Um zu zeigen, dass dies nicht der Fall ist, müssen wir nur *eine* Testfunktion ϕ finden, die das nicht bewerkstelligt. Betrachten wir konkret die Funktion ϕ aus (4.11) und definieren $\phi_a(x) = \phi(x/a)$, sodass $\phi_a(0) = 1/e$ und $|\phi_a(x)| \leq 1/e$. Dann wäre

$$\int_{-\infty}^{\infty} dx f(x) \phi_a(x) = \frac{1}{e} \quad (4.27)$$

unabhängig von a . Andererseits ist

$$\left| \int_{-\infty}^{\infty} dx f(x) \phi_a(x) \right| = \left| \int_{-a}^a dx |f(x)| e^{-\frac{a^2}{a^2-x^2}} \right| \leq \frac{1}{e} \int_{-a}^a dx |f(x)|. \quad (4.28)$$

Wenn f lokal integrierbar ist, muss im Limes $a \rightarrow 0$ aber gelten:

$$\lim_{a \rightarrow 0} \int_{-a}^a dx |f(x)| = 0 \quad \Rightarrow \quad \lim_{a \rightarrow 0} \left| \int_{-\infty}^{\infty} dx f(x) \phi_a(x) \right| = 0. \quad (4.29)$$

Das ist aber ein Widerspruch zu (4.27)! Die Funktion f kann also nicht lokal integrierbar sein, und somit ist die Delta-Distribution eine singuläre Distribution.

In Analogie zu regulären Distributionen hat es sich eingebürgert, die Delta-Distribution ebenfalls in der Form (4.17) zu schreiben:

$$\delta_{x_0}[\phi] = \int_{-\infty}^{\infty} dx \delta(x - x_0) \phi(x) = \phi(x_0). \quad (4.30)$$

Damit entspricht die **Dirac'sche Deltafunktion** $\delta(x - x_0)$ formal dem Integralkern der Distribution. Solche "verallgemeinerten Funktionen" sind aber keine Funktionen im üblichen Sinn, und auch das Integral kann weder als Lebesgue- noch als Riemann-Integral verstanden werden. Die Eigenschaften der Distribution δ_{x_0} übertragen sich aber auf das Objekt $\delta(x - x_0)$, und in diesem Sinne erlangen auch die Formeln (4.7–4.8) eine Bedeutung – solange man im Auge behält, dass das zugrundeliegende Objekt eine Distribution ist.

Eigenschaften von Distributionen. Eine sehr nützliche Eigenschaft von Distributionen macht sich in ihren **Ableitungen** bemerkbar, denn Distributionen darf man beliebig oft differenzieren! Betrachten wir eine reguläre Distribution F mit Integralkern $f(x)$. Falls auch f' lokal integrierbar ist, erzeugt es wieder eine reguläre Distribution

$$F'[\phi] = \int_{-\infty}^{\infty} dx f'(x) \phi(x) = - \int_{-\infty}^{\infty} dx f(x) \phi'(x) = -F[\phi']. \quad (4.31)$$

Wegen $\phi \in \mathcal{D}$ verschwindet der Randterm, somit konnten wir die Ableitung auf die Testfunktion ϕ überwälzen. Da dies eine glatte Funktion ist, existieren auch beliebig hohe Ableitungen von Distributionen. Nicht-reguläre Distributionen haben keine Integraldarstellung, deswegen fordert man die obige Eigenschaft als *Definition* der Ableitung:

DEFINITION: Die m -te Ableitung einer Distribution $F \in \mathcal{D}'$ ist definiert durch

$$F^{(m)}[\phi] := (-1)^m F[\phi^{(m)}] \quad \forall \phi \in \mathcal{D}(\mathbb{R}). \quad (4.32)$$

Diese Ableitung wird auch als **distributionelle Ableitung** bezeichnet.

Man kann sich klarmachen, dass $F^{(m)}$ linear und stetig ist und somit wieder eine Distribution definiert. Einige Beispiele:

- Die Ableitung der Heaviside-Distribution ergibt die Delta-Distribution:

$$\begin{aligned} \Theta'_{x_0}[\phi] &= -\Theta_{x_0}[\phi'] = - \int_{-\infty}^{\infty} dx \Theta(x - x_0) \phi'(x) = - \int_{x_0}^{\infty} dx \phi'(x) \\ &= -\phi(x) \Big|_{x_0}^{\infty} = \phi(x_0) = \delta_{x_0}[\phi]. \end{aligned} \quad (4.33)$$

Für den "Integralkern" lässt sich das auch so schreiben:

$$\frac{d}{dx} \Theta(x - x_0) = \delta(x - x_0). \quad (4.34)$$

- Die Ableitung der Delta-Distribution ergibt

$$\delta'_{x_0}[\phi] = -\delta_{x_0}[\phi'] = -\phi'(x_0). \quad (4.35)$$

Das **Produkt** zweier Distributionen ist nicht unbedingt wieder eine Distribution, wie man schon anhand von regulären Distributionen sehen kann: z.B. ist die durch $f(x) = 1/|\sqrt{x}|$ definierte Funktion lokal integrierbar, ihr Produkt mit sich selbst aber nicht mehr. Die Tatsache, dass Distributionen nicht ohne Weiteres multipliziert werden können, schränkt die Brauchbarkeit der Distributionentheorie in nichtlinearen Problemen stark ein. Für $F \in \mathcal{D}'(\mathbb{R})$ und Funktionen $u \in C^\infty(\mathbb{R})$ lässt sich ein solches Produkt aber bilden,

$$(u \cdot F)[\phi] := \int_{-\infty}^{\infty} dx u(x) f(x) \phi(x) = F[u\phi], \quad (4.36)$$

und seine Ableitung $(u \cdot F)' = u'F + uF'$ erfüllt die Produktregel, wie man mittels (4.32) und (4.36) überprüfen kann.

Rechenregeln für die Deltafunktion. Im Folgenden stellen wir einige häufig verwendete Rechenregeln für die Dirac'sche Deltafunktion zusammen. Es gilt

$$\int_a^b dx f(x) \delta(x - x_0) = \begin{cases} f(x_0) & a < x_0 < b, \\ 0 & \text{sonst.} \end{cases} \quad (4.37)$$

$$\int_a^b dx f(x) \delta^{(m)}(x - x_0) = \begin{cases} (-1)^m f^{(m)}(x_0) & a < x_0 < b, \\ 0 & \text{sonst.} \end{cases} \quad (4.38)$$

Multipliziert man eine Funktion mit einer Deltafunktion, kann man den Funktionswert durch $f(x_0)$ ersetzen:

$$f(x) \delta(x - x_0) = f(x_0) \delta(x - x_0) \quad \Rightarrow \quad x \delta(x) = 0. \quad (4.39)$$

Aus (4.38) folgt außerdem $-x \delta'(x) = \delta(x)$. Die Deltafunktion ist die Ableitung der Heaviside-Stufenfunktion:

$$\delta(x - x_0) = \frac{d}{dx} \Theta(x - x_0). \quad (4.40)$$

Oft benötigt man die Deltafunktion einer Funktion f . In diesem Fall muss man über alle Nullstellen der Funktion summieren,

$$\delta(f(x)) = \sum_{\substack{x_i \\ f(x_i)=0}} \frac{\delta(x - x_i)}{|f'(x_i)|}, \quad (4.41)$$

was die folgenden Spezialfälle liefert:

$$\delta(-x) = \delta(x), \quad \delta(ax) = \frac{\delta(x)}{|a|}, \quad \delta(x^2 - x_0^2) = \frac{\delta(x - x_0) + \delta(x + x_0)}{2|x_0|}. \quad (4.42)$$

Die obigen Relationen lassen sich herleiten, indem man sie über Testfunktionen integriert, z.B.

$$\int_a^b dx f(x) \phi(x) = \int_a^b dx f(x) \int_{-\infty}^{\infty} dy \delta(y - x) \phi(y) = \int_{-\infty}^{\infty} dy \phi(y) \int_a^b dx f(x) \delta(x - y) \stackrel{(4.37)}{=} \int_a^b dy \phi(y) f(y), \quad (4.43)$$

wobei die Vertauschung der Integrale unter einer geeigneten Limesbildung erlaubt ist. In der Schreibweise (4.37) muss die Funktion f selbst daher keine Testfunktion sein. Die zweite Identität in (4.42) folgt aus

$$\int_{-\infty}^{\infty} dx \delta(ax) \phi(x) \stackrel{ax=y}{=} \int_{-\infty}^{\infty} dy \frac{\delta(y)}{|a|} \phi\left(\frac{y}{a}\right) = \frac{\phi(0)}{|a|} = \int_{-\infty}^{\infty} dx \frac{\delta(x)}{|a|} \phi(x), \quad (4.44)$$

wobei der Betrag $|a|$ daraus resultiert, dass sich für $a < 0$ die Integralgrenzen umkehren. Für (4.41) spaltet man das Integral in Regionen um die Nullstellen x_i von f auf und entwickelt die Funktion dort in eine Taylorreihe $f(x) = f(x_i) + (x - x_i) f'(x_i) + \dots$, wobei $f(x_i) = 0$ ist:

$$\int_{-\infty}^{\infty} dx \delta(f(x)) \phi(x) = \sum_i \int_{x_i - \epsilon_i}^{x_i + \epsilon_i} dx \delta((x - x_i) f'(x_i)) \phi(x) \stackrel{(4.44)}{=} \sum_i \int_{x_i - \epsilon_i}^{x_i + \epsilon_i} dx \frac{\delta(x - x_i)}{|f'(x_i)|} \phi(x) = \sum_i \int_{-\infty}^{\infty} dx \frac{\delta(x - x_i)}{|f'(x_i)|} \phi(x).$$

Fouriertransformation von Distributionen. Auch die Fouriertransformation lässt sich sofort auf Distributionen verallgemeinern. Für eine reguläre Distribution f gilt

$$\begin{aligned}\mathcal{F}f[\phi] &= \int_{-\infty}^{\infty} dp \mathcal{F}f(p) \phi(p) = \int_{-\infty}^{\infty} dp \int_{-\infty}^{\infty} dx f(x) e^{-ipx} \phi(p) \\ &= \int_{-\infty}^{\infty} dx f(x) \underbrace{\int_{-\infty}^{\infty} dp \phi(p) e^{-ipx}}_{=\mathcal{F}\phi(x)} = f[\mathcal{F}\phi].\end{aligned}\quad (4.45)$$

Hier muss man allerdings vorsichtig sein, denn die Fouriertransformierte einer Testfunktion $\phi \in \mathcal{D}'$ ist nicht notwendigerweise wieder eine Testfunktion. Wir wissen aber von früher, dass die Fouriertransformation eine Bijektion im Schwartz-Raum ist, deswegen betrachten wir hier konkret Testfunktionen $\phi \in S(\mathbb{R})$ im Schwartz-Raum. Der zugehörige Dualraum $S'(\mathbb{R})$ der Funktionale f ist der Raum der **temperierten Distributionen**. Man kann außerdem zeigen, dass folgendes gilt:

$$\mathcal{D} \subset S \subset C^\infty \quad \longleftrightarrow \quad (C^\infty)' \subset S' \subset \mathcal{D}', \quad (4.46)$$

d.h. der Raum S' der temperierten Distributionen bildet einen Unterraum von \mathcal{D}' . Somit lässt sich die Fouriertransformation für alle temperierten Distributionen definieren:

DEFINITION: Die Fouriertransformierte $\mathcal{F}f$ einer temperierten Distribution $f \in S'(\mathbb{R})$ ist definiert durch

$$\mathcal{F}f[\phi] = f[\mathcal{F}\phi] \quad \forall \phi \in S(\mathbb{R}). \quad (4.47)$$

Damit können wir jetzt auch die Fouriertransformierte der Delta-Distribution berechnen:

$$\mathcal{F}\delta_{x_0}[\phi] = \delta_{x_0}[\mathcal{F}\phi] = \mathcal{F}\phi(x_0) = \int_{-\infty}^{\infty} dp \phi(p) e^{-ipx_0} =: g_{x_0}[\phi]. \quad (4.48)$$

Die Fouriertransformierte des Integralkerns $\delta(x-x_0)$ lautet daher $g_{x_0}(p) = e^{-ipx_0}$. Für $x_0 = 0$ folgt, dass die Fouriertransformierte von $\delta(x)$ gleich 1 ist. Damit muss die Umkehrtransformation von e^{-ipx_0} wieder $\delta(x-x_0)$ ergeben:

$$\mathcal{F}^{-1}g_{x_0}[\phi] = g_{x_0}[\mathcal{F}^{-1}\phi] = \int_{-\infty}^{\infty} dp e^{-ipx_0} \mathcal{F}^{-1}\phi(p) = \phi(x_0) = \delta_{x_0}[\phi]. \quad (4.49)$$

Analog lauten die Fouriertransformation von $f_{p_0}(x) := \frac{1}{2\pi} e^{ip_0x}$ und deren Umkehrung

$$\mathcal{F}f_{p_0}[\phi] = f_{p_0}[\mathcal{F}\phi] = \frac{1}{2\pi} \int_{-\infty}^{\infty} dx e^{ip_0x} \mathcal{F}\phi(x) = \phi(p_0) = \delta_{p_0}[\phi], \quad (4.50)$$

$$\mathcal{F}^{-1}\delta_{p_0}[\phi] = \delta_{p_0}[\mathcal{F}^{-1}\phi] = \mathcal{F}^{-1}\phi(p_0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dx \phi(x) e^{ip_0x} = f_{p_0}[\phi]. \quad (4.51)$$

Mit den Rechenregeln für die Deltafunktion und einer mathematisch etwas unsauberen Schreibweise (die in der Physik aber allgegenwärtig ist) gehen diese Relationen noch viel schneller von der Hand:

$$\mathcal{F}(\delta(x - x_0))(p) = \int_{-\infty}^{\infty} dx \delta(x - x_0) e^{-ipx} = e^{-ipx_0}, \quad (4.52)$$

$$\mathcal{F}^{-1}(e^{-ipx_0})(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dp e^{ip(x-x_0)} = \delta(x - x_0), \quad (4.53)$$

$$\mathcal{F}\left(\frac{1}{2\pi} e^{ip_0x}\right)(p) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dx e^{-i(p-p_0)x} = \delta(p - p_0), \quad (4.54)$$

$$\mathcal{F}^{-1}(\delta(p - p_0))(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dp \delta(p - p_0) e^{ipx} = \frac{e^{ip_0x}}{2\pi}. \quad (4.55)$$

Aus (4.53) bzw. (4.54) erhalten wir eine sehr nützliche Darstellung der Deltafunktion:

$$\delta(x - x_0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dp e^{ip(x-x_0)}. \quad (4.56)$$

Auf diese Weise ergibt auch die frühere Gleichung (3.40) Sinn, denn deren rechte Seite liefert einfach wie erwartet $f(x)$.

Erweiterung auf \mathbb{R}^n . Obwohl wir uns bisher auf den Raum \mathbb{R} beschränkt haben, lässt sich alles Gesagte problemlos auf den \mathbb{R}^n erweitern. In diesem Fall ist $\mathcal{D}'(\mathbb{R}^n) = C_0^\infty(\mathbb{R}^n)$, und das Analogon der Konvergenzeigenschaft (4.13) erfordert die Definition eines n -dimensionalen Multiindex $\alpha = (\alpha_1, \dots, \alpha_n)$, sodass

$$\frac{\partial^\alpha}{\partial x^\alpha} \longrightarrow D^\alpha := \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}}, \quad |\alpha| = \alpha_1 + \dots + \alpha_n. \quad (4.57)$$

Falls irgendeine Komponente von α Null ist, wird die Ableitung nach dieser Variable übergangen, z.B. für $n = 3$ und $\alpha = (2, 0, 5)$:

$$D^\alpha = \frac{\partial^7}{\partial x_1^2 \partial x_3^5}. \quad (4.58)$$

Die Dirac'sche Deltafunktion ist dann das Produkt von n Deltafunktion, z.B. im \mathbb{R}^3 :

$$\delta^3(\mathbf{x} - \mathbf{x}_0) = \delta(x_1 - x_{01}) \delta(x_2 - x_{02}) \delta(x_3 - x_{03}). \quad (4.59)$$

Die Wirkung unter einem dreidimensionalen Integral entspricht dann dem Produkt der einzelnen Faktoren:

$$\delta^3(\mathbf{x} - \mathbf{x}_0) = \frac{1}{(2\pi)^3} \int_{\mathbb{R}^3} d^3p e^{ip \cdot (\mathbf{x} - \mathbf{x}_0)}, \quad \int_{\mathbb{R}^3} d^3x f(\mathbf{x}) \delta^3(\mathbf{x} - \mathbf{x}_0) = f(\mathbf{x}_0). \quad (4.60)$$

Greenfunktionen. Distributionen bieten eine leistungsfähige Methode zur Lösung allgemeiner inhomogener Differenzialgleichungen mit gegebenen Randbedingungen, nämlich über die nach [George Green](#) (1793–1841) benannten Greenfunktionen.

SATZ: Sei $Au = f$ eine lineare Differenzialgleichung mit dem linearen Differenzialoperator

$$A = \sum_{k=1}^n a_k(x) \frac{d^k}{dx^k} \quad (4.61)$$

und der Inhomogenität f . Dann erhält man eine partikuläre Lösung u_p durch Faltung der Inhomogenität mit der zu A zugehörigen **Greenfunktion** G :

$$u_p(x) = (G * f)(x) = \int_{-\infty}^{\infty} dy G(x-y) f(y), \quad AG(x-y) = \delta(x-y). \quad (4.62)$$

Davon kann man sich durch Einsetzen leicht überzeugen:

$$Au_p(x) = \int_{-\infty}^{\infty} dy AG(x-y) f(y) = \int_{-\infty}^{\infty} dy \delta(x-y) f(y) = f(x). \quad (4.63)$$

In anderen Worten: Sobald die Greenfunktion bekannt ist, ist die Differenzialgleichung für jede beliebige Inhomogenität gelöst!

Betrachten wir als Beispiel eine lineare, inhomogene Differenzialgleichung zweiter Ordnung zusammen mit der entsprechenden Gleichung für die Greenfunktion $G(z)$, mit $z = x - y$:

$$u''(x) + u(x) = f(x), \quad G''(z) + G(z) = \delta(z). \quad (4.64)$$

Letztere kann auf verschiedene Arten gelöst werden, z.B. analog wie früher in (3.79) über eine Laplace-Transformation, oder durch Variation der Konstanten aus der Lösung der homogenen Gleichung $u_h(x) = c_1 \cos x + c_2 \sin x$ mit der Inhomogenität $\delta(z)$. Das Resultat ist

$$G(z) = \Theta(z) \sin z, \quad (4.65)$$

wie man durch Einsetzen unter Verwendung der Regeln (4.39–4.40) überprüfen kann. Damit ergibt sich die partikuläre Lösung der Differenzialgleichung zu

$$u_p(x) = \int_{-\infty}^{\infty} dy \Theta(x-y) \sin(x-y) f(y) = \int_{-\infty}^x dy \sin(x-y) f(y), \quad (4.66)$$

und die Gesamtlösung, bei der die Randbedingungen noch zu spezifizieren sind, lautet

$$u(x) = c_1 \cos x + c_2 \sin x + \int_{-\infty}^x dy \sin(x-y) f(y). \quad (4.67)$$

Ein Beispiel in drei Raumdimensionen sind die **Maxwellgleichungen** der Elektrostatik:

$$\nabla \cdot \mathbf{E} = \rho, \quad \nabla \times \mathbf{E} = 0. \quad (4.68)$$

Sie bestimmen das zu einer Ladungsverteilung $\rho(\mathbf{x})$ zugehörige elektrische Feld $\mathbf{E}(\mathbf{x})$. Aus der zweiten Gleichung folgt, dass sich $\mathbf{E} = -\nabla\phi$ als Gradient eines Potential $\phi(\mathbf{x})$ schreiben lässt, was auf die **Poissongleichung** führt:

$$\Delta\phi = -\rho. \quad (4.69)$$

Die homogene Version dieser Gleichung ist die **Laplace-Gleichung** $\Delta\phi = 0$. Die zugehörige Greenfunktion $G(\mathbf{x} - \mathbf{y}) = G(\mathbf{z})$ entspricht dem Potential einer punktförmigen Ladungsverteilung im Ursprung:

$$\Delta G(\mathbf{z}) = -\delta^3(\mathbf{z}). \quad (4.70)$$

Um $G(\mathbf{z})$ zu berechnen, führen wir eine Fouriertransformation in allen drei Variablen durch. Bezeichnen wir die Fouriertransformierte mit $\hat{G}(\mathbf{p})$, dann folgt aus den Eigenschaften (3.36) und (4.52)

$$\mathbf{p}^2 \hat{G}(\mathbf{p}) = 1, \quad (4.71)$$

und somit

$$\hat{G}(\mathbf{p}) = \frac{1}{\mathbf{p}^2} \Rightarrow G(\mathbf{z}) = \frac{1}{(2\pi)^3} \int d^3p \hat{G}(\mathbf{p}) e^{i\mathbf{p}\cdot\mathbf{z}}. \quad (4.72)$$

In drei Dimensionen ist der Ausdruck $1/\mathbf{p}^2$ integrierbar. Zur Berechnung des Integrals verwenden wir analog zum früheren Beispiel (3.50) Kugelkoordinaten mit $d^3p = dp p^2 d(\cos\theta) d\varphi$, $\mathbf{p} \cdot \mathbf{z} = |\mathbf{p}||\mathbf{z}| \cos\theta = pr \cos\theta$ und $\cos\theta = w$, wobei wir ohne Beschränkung der Allgemeinheit $\mathbf{z} = r \mathbf{e}_3$ gesetzt haben. Daraus bekommen wir das **Coulombpotential** zurück:

$$G(\mathbf{z}) = \frac{1}{(2\pi)^2} \int_0^\infty dp \int_{-1}^1 dw \underbrace{e^{iprw}}_{\frac{2 \sin(pr)}{pr}} = \frac{1}{2\pi^2 r} \int_0^\infty dx \underbrace{\frac{\sin x}{x}}_{\frac{\pi}{2}} = \frac{1}{4\pi r}. \quad (4.73)$$

Das deckt sich auch mit unserem früheren Resultat (3.52). Das Potential einer elektrischen Punktladung ist also gerade das Coulombpotential, und es gilt

$$\Delta \frac{1}{4\pi|\mathbf{z}|} = -\delta^3(\mathbf{z}). \quad (4.74)$$

Damit lässt sich die partikuläre Lösung der Poissongleichung für eine allgemeine Ladungsverteilung $\rho(\mathbf{x})$ angeben:

$$\phi_p(\mathbf{x}) = \frac{1}{4\pi} \int d^3x' \frac{\rho(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|}. \quad (4.75)$$

4.2 Funktionalableitung

Das zentrale Thema der Funktionalanalysis ist die Untersuchung linearer Abbildungen. Viele Probleme in der Physik sind allerdings **nichtlinear**. In diesem Abschnitt werden wir kurz auf nichtlineare Funktionale eingehen und uns insbesondere mit der Frage beschäftigen, wie man die Extremwerte eines nichtlinearen Funktional findet.

Was ist die kürzeste Verbindung zwischen zwei Punkten $A = (x_1, y_1)$ und $B = (x_2, y_2)$ in der Ebene? Der Abstand ist die Bogenlänge, ein Funktional der Form

$$S[y] = \int_A^B ds = \int_A^B \sqrt{dx^2 + dy^2} = \int_{x_1}^{x_2} dx \sqrt{1 + y'^2}. \quad (4.76)$$

Offenbar ist das ein nichtlineares Funktional, das jeder Kurve von A nach B , also jeder Funktion $y(x)$, eine Zahl zuordnet. Welche Kurve muss man wählen, damit der Weg am kürzesten wird? Dazu müssen wir den Extremwert des Funktional finden.

Das erinnert an die Extremwertsuche von Funktionen in der Analysis. Die Extremwerte einer stetig differenzierbaren Funktion $f: (a, b) \rightarrow \mathbb{R}$ sind die stationären Punkte, an denen $f'(x) = 0$ gilt. Diese können als Minima, Maxima oder Sattelpunkte vorliegen. Für reelle Funktionen mehrerer Variablen $f: \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$, die in allen Variablen stetig differenzierbar sind, werden die stationären Punkte durch $\partial f / \partial x_i = 0$ für $i = 1 \dots n$ bestimmt, d.h. durch einen verschwindenden Gradienten $\nabla f = 0$. Der Gradient am Ort \mathbf{x} ergibt sich aus der linearen Näherung der Funktion in der Umgebung von \mathbf{x} :

$$f(\mathbf{x} + \mathbf{h}) = f(\mathbf{x}) + \sum_{i=1}^n \frac{\partial f}{\partial x_i} h_i + r(\mathbf{h}) = f(\mathbf{x}) + \nabla f \cdot \mathbf{h} + r(\mathbf{h}). \quad (4.77)$$

Wenn f an der Stelle \mathbf{x} differenzierbar ist, verschwindet der Fehler $r(\mathbf{h})$ für $|\mathbf{h}| \rightarrow 0$ schneller als linear, d.h. $\lim_{|\mathbf{h}| \rightarrow 0} r(\mathbf{h})/|\mathbf{h}| = 0$. Dasselbe Prinzip wollen wir jetzt auf Funktionale verallgemeinern, deren Argumente *Funktionen* sind.

DEFINITION: Sei \mathbb{V} ein normierter Vektorraum über \mathbb{R} und $F: \mathbb{V} \rightarrow \mathbb{R}$ ein Funktional. F heißt an der Stelle $\phi \in \mathbb{V}$ **differenzierbar**, wenn es ein stetiges lineares Funktional $\delta F[\phi]$ gibt, sodass gilt:

$$F[\phi + h] = F[\phi] + \delta F[\phi][h] + r(h), \quad \lim_{\|h\| \rightarrow 0} \frac{|r(h)|}{\|h\|} = 0. \quad (4.78)$$

Für Funktionenräume schreibt man $\delta F[\phi][h]$ dann in der folgenden Form, wobei der Integral-kern $\delta F[\phi]/\delta \phi(x)$ als **Funktionalableitung** bezeichnet wird:

$$\delta F[\phi][h] = \int_{-\infty}^{\infty} dx \frac{\delta F[\phi]}{\delta \phi(x)} h(x). \quad (4.79)$$

Falls F ein Funktional über dem \mathbb{R}^n ist, also einfach eine Funktion, dann reduziert sich die Funktionalableitung auf den Gradienten in (4.77). Im Fall von Funktionenräumen entspricht die Funktionalableitung der Verallgemeinerung des Gradienten, wobei die Summe über einen diskreten Index i zu einem Integral über eine kontinuierliche Variable x wird. Sehen wir uns dazu einige Beispiele an:

- Bei einem linearen Funktional $F[\phi] = \int_{-\infty}^{\infty} dx J(x) \phi(x)$ lässt sich die Funktionalableitung unmittelbar ablesen:

$$F[\phi + h] = \int_{-\infty}^{\infty} dx J(x) \phi(x) + \int_{-\infty}^{\infty} dx J(x) h(x) \Rightarrow \frac{\delta F[\phi]}{\delta \phi(x)} = J(x). \quad (4.80)$$

- Für ein Funktional der Form $F[\phi] = \int dx J(x) \phi(x)^n$ gilt

$$(\phi + h)^n = \sum_{k=0}^n \binom{n}{k} \phi^{n-k} h^k = \phi^n + n \phi^{n-1} h + \mathcal{O}(h^2), \quad (4.81)$$

woraus man die Funktionalableitung ablesen kann:

$$F[\phi + h] = \int_{-\infty}^{\infty} dx J(x) \phi(x)^n + \int_{-\infty}^{\infty} dx \underbrace{J(x) n \phi(x)^{n-1}}_{\delta F[\phi] / \delta \phi(x)} h(x) + \mathcal{O}(h^2). \quad (4.82)$$

- Analog für ein Funktional $F[\phi] = \int_{-\infty}^{\infty} dx J(x) f(\phi(x))$:

$$F[\phi + h] = \int_{-\infty}^{\infty} dx J(x) [f(\phi) + f'(\phi) h + \mathcal{O}(h^2)] \Rightarrow \frac{\delta F[\phi]}{\delta \phi(x)} = J(x) f'(\phi(x)). \quad (4.83)$$

- Oder:

$$F[\phi] = \int_0^{\infty} dx \phi(x) = \int_{-\infty}^{\infty} dx \Theta(x) \phi(x) \Rightarrow \frac{\delta F[\phi]}{\delta \phi(x)} = \Theta(x). \quad (4.84)$$

- Auch für die Delta-Distribution $F[\phi] = \delta_z[\phi] = \phi(z) = \int_{-\infty}^{\infty} dx \delta(x - z) \phi(x)$, eine singuläre Distribution, für welche die Integralschreibweise (4.79) formal nicht existiert, lässt sich die Funktionalableitung bilden:

$$F[\phi + h] = F[\phi] + \int_{-\infty}^{\infty} dx \delta(x - z) h(x) \Rightarrow \frac{\delta F[\phi]}{\delta \phi(x)} = \frac{\delta \phi(z)}{\delta \phi(x)} = \delta(x - z). \quad (4.85)$$

- Analog kann man weitere Beispiele ausrechnen, z.B.

$$F[\phi] = f(\phi(z)) \Rightarrow \frac{\delta F[\phi]}{\delta \phi(x)} = f'(\phi(x)) \delta(x - z), \quad (4.86)$$

$$F[\phi] = \exp \left[i \int_{-\infty}^{\infty} dx J(x) \phi(x) \right] \Rightarrow \frac{\delta F[\phi]}{\delta \phi(x)} = i J(x) \exp \left[i \int_{-\infty}^{\infty} dy J(y) \phi(y) \right]. \quad (4.87)$$

Dabei haben wir $e^{iz} = 1 + iz + \mathcal{O}(z^2)$ in eine Reihe entwickelt, woraus sich die Funktionalableitung ablesen lässt.

Euler-Lagrange-Gleichungen. Die wohl bekannteste Anwendung der Funktionalableitung in der Physik sind die nach **Leonhard Euler** (1707–1783) und **Joseph-Louis Lagrange** (1736–1813) benannten Euler-Lagrange-Gleichungen. Betrachten wir dazu eine Funktion ϕ auf dem Intervall $[x_1, x_2] \subset \mathbb{R}$, die an den Randpunkten festgehalten wird, $\phi(x_1) = \phi_1$, $\phi(x_2) = \phi_2$. Dadurch werden zwei Punkte $A = (x_1, \phi_1)$ und $B = (x_2, \phi_2)$ in der Ebene definiert. Wir suchen jetzt unter allen Kurven, die A mit B verbinden (d.h. unter allen Funktionen ϕ mit festgehaltenen Randpunkten) jene, für die das nichtlineare Funktional

$$F[\phi] = \int_{x_1}^{x_2} dx f(x, \phi(x), \phi'(x)) \quad (4.88)$$

stationär wird. Offenbar ist das die allgemeine Formulierung des Beispiels (4.76). Bilden wir also die Funktionalableitung:

$$\begin{aligned} F[\phi + h] &= \int_{x_1}^{x_2} dx f(x, \phi(x) + h(x), \phi'(x) + h'(x)) \\ &= \int_{x_1}^{x_2} dx \left[f(x, \phi(x), \phi'(x)) + \frac{\partial f}{\partial \phi(x)} h(x) + \frac{\partial f}{\partial \phi'(x)} h'(x) + \mathcal{O}(h^2) \right] \\ &= F[\phi] + \frac{\partial f}{\partial \phi'(x)} h(x) \Big|_{x_1}^{x_2} + \int_{x_1}^{x_2} dx \left[\frac{\partial f}{\partial \phi(x)} - \frac{d}{dx} \frac{\partial f}{\partial \phi'(x)} \right] h(x) + \mathcal{O}(h^2). \end{aligned} \quad (4.89)$$

Der Randterm verschwindet, da wir die Randpunkte festhalten und dort keine Variation zulassen, sodass $h(x_1) = h(x_2) = 0$. Der Klammerausdruck im Restterm ergibt dann die Funktionalableitung $\delta F[\phi]/\delta \phi(x)$. Damit das Funktional F stationär wird, muss die Variation $\delta F[\phi]$ verschwinden; das ist äquivalent zum Verschwinden des Integranden, d.h. der Funktionalableitung, da h ja eine beliebige glatte Funktion ist, die nur an den Rändern Null sein muss. Somit ist ϕ ein stationärer Punkt des Funktionals F , der die Randbedingungen $\phi(x_1) = \phi_1$ und $\phi(x_2) = \phi_2$ erfüllt, wenn die **Euler-Lagrange-Gleichungen** erfüllt sind:

$$\frac{\delta F[\phi]}{\delta \phi(x)} = \frac{\partial f}{\partial \phi(x)} - \frac{d}{dx} \frac{\partial f}{\partial \phi'(x)} = 0. \quad (4.90)$$

Die Euler-Lagrange-Gleichungen spielen eine wichtige Rolle in vielen Bereichen der Physik. In der klassischen Mechanik hat die **Wirkung** (*action*) eine Gestalt analog zu (4.88). Aus dem **Hamilton-Prinzip** der stationären Wirkung ($\delta S = 0$) folgen die **Newtonschen Bewegungsgleichungen** für die Bahnkurve $\mathbf{q}(t)$ eines Massenpunkts bzw. eines Systems von Massenpunkten:

$$S[\mathbf{q}] = \int_{t_1}^{t_2} dt L(t, \mathbf{q}(t), \dot{\mathbf{q}}(t)), \quad \delta S[\mathbf{q}] = 0 \Leftrightarrow \frac{\partial L}{\partial q_i(t)} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i(t)} = 0. \quad (4.91)$$

Dabei ist L die **Lagrange-Funktion**, die die vollständige Dynamik des Systems beschreibt. Analog verläuft es in der (Quanten-)Feldtheorie, wo die **Lagrangedichte** \mathcal{L} die Dynamik einer Theorie in Form ihrer Felder und deren Ableitungen definiert. Die Euler-Lagrange-Gleichungen sind dann die jeweiligen **Feldgleichungen**, z.B. die Klein-Gordon-Gleichung für Teilchen mit Spin 0, die Dirac-Gleichung (Spin 1/2), die Maxwell-Gleichungen fürs elektromagnetische Feld (Photonen mit Spin 1), die Einstein-Gleichung usw. Die Erwartungswerte der Euler-Lagrange-Gleichungen in der Quantenfeldtheorie heißen **Dyson-Schwinger-Gleichungen**.

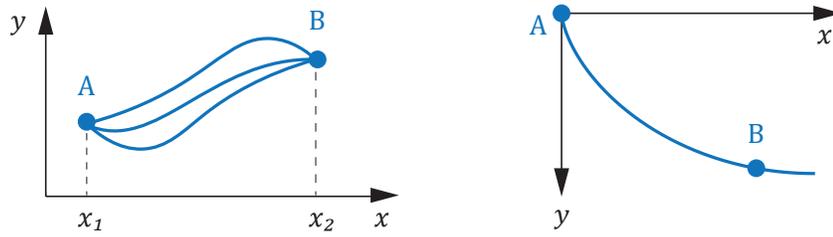


FIG. 4.2: Was ist der kürzeste Weg zwischen A und B? Rechts: Brachistochronenproblem

Kehren wir zu unserem ursprünglichen Problem (4.76) zurück. Was ist der kürzeste Weg in der Ebene, der zwei Punkte A und B verbindet (Fig. 4.2)? Die Bogenlänge $S[y]$ hat die Form (4.88) mit $f(y') = \sqrt{1 + y'^2}$. Bilden wir die Euler-Lagrange-Gleichung:

$$\frac{\partial f}{\partial y} = 0, \quad \frac{\partial f}{\partial y'} = \frac{y'}{\sqrt{1 + y'^2}} \Rightarrow \frac{d}{dt} \frac{\partial f}{\partial y'} = 0 \Rightarrow \frac{\partial f}{\partial y'} = \frac{y'}{\sqrt{1 + y'^2}} = c, \quad (4.92)$$

und somit $y'^2 = c^2(1 + y'^2) \Rightarrow y'^2 = c^2/(1 - c^2) \Rightarrow y' = \pm 1/\sqrt{1 - c^2} =: a$. Wenig überraschend ist die kürzeste Verbindung also eine Gerade $y(x) = ax + b$, wobei sich die Integrationskonstanten a und b aus den Randbedingungen $y(x_1) = y_1$ und $y(x_2) = y_2$ bestimmen.

Ein anspruchsvolleres Beispiel ist das auf **Johann Bernoulli** zurückgehende **Brachistochronenproblem**. Wir suchen die "perfekte Rutsche", bei der ein Massenpunkt entlang einer Kurve $y(x)$ unter dem Einfluss der Schwerkraft reibungslos von einem Punkt A zu einem tiefergelegenen Punkt B gleitet (siehe Fig. 4.2). Wir wollen $y(x)$ so bestimmen, dass die Zeit T , die der Massenpunkt benötigt, minimal wird. Das Funktional lautet daher $T[y] = \int dt = \int ds/v = \int dx \sqrt{1 + y'^2}/v$. Um einen Ausdruck für v zu gewinnen, verwenden wir die Energieerhaltung: $E = mv^2/2 - mgy = 0$ (das Teilchen ruht für $y = 0$). Die Gravitation zeigt hier in die positive y -Richtung, was die Rechnung vereinfacht. Damit ergibt sich $v(y) = \sqrt{2gy}$ und somit

$$T[y] = \int dx \sqrt{\frac{1 + y'^2}{2gy}} = \int dx f(y, y') \Rightarrow \frac{\partial f}{\partial y} - \frac{d}{dx} \frac{\partial f}{\partial y'} = 0 \Rightarrow \frac{\partial f}{\partial y} = \frac{d}{dx} \frac{\partial f}{\partial y'}. \quad (4.93)$$

Da f nicht explizit von x abhängt und somit $\partial f/\partial x = 0$ gilt, lässt sich das Problem einfacher lösen, wenn man die **Beltrami-Identität** verwendet:

$$\frac{df}{dx} = \underbrace{\frac{\partial f}{\partial y} y'}_{=(4.93)} + \frac{\partial f}{\partial y'} y'' + \frac{\partial f}{\partial x} = \frac{d}{dx} \left(y' \frac{\partial f}{\partial y'} \right) + \underbrace{\frac{\partial f}{\partial x}}_{=0} \Rightarrow f - y' \frac{\partial f}{\partial y'} = c. \quad (4.94)$$

Setzen wir die konkreten Werte ein, ergibt sich

$$\frac{\partial f}{\partial y'} = \frac{y'}{\sqrt{2gy}\sqrt{1 + y'^2}} \Rightarrow \sqrt{\frac{1 + y'^2}{2gy}} - \frac{y'^2}{\sqrt{2gy}\sqrt{1 + y'^2}} = \frac{1}{\sqrt{2gy}} \frac{1}{\sqrt{1 + y'^2}} = c \quad (4.95)$$

und somit

$$(1 + y'^2)y = \frac{1}{2gc^2} =: 2a \Rightarrow y' = \sqrt{\frac{2a - y}{y}} \Rightarrow dx = \sqrt{\frac{y}{2a - y}} dy. \quad (4.96)$$

Das lässt sich mit der Variablensubstitution $y = a(1 - \cos u) \Rightarrow dy = a \sin u du$ lösen, denn damit erhält man

$$x = \int dy \sqrt{\frac{y}{2a - y}} = a \int du \sqrt{\frac{1 - \cos u}{1 + \cos u}} \underbrace{\sin u}_{=\sqrt{1 - \cos^2 u}} = a \int du (1 - \cos u) = a(u - \sin u) + const. \quad (4.97)$$

Mit $x(u = 0) = 0$ verschwindet die Integrationskonstante und wir bekommen die Gleichung für eine **Zykloide**:

$$x(u) = a(u - \sin u), \quad y(u) = a(1 - \cos u). \quad (4.98)$$

Einige Bemerkungen:

- Analog zu gewöhnlichen Funktionen lassen sich **Nebenbedingungen** auch bei Variationsproblemen bequem über Lagrange-Multiplikatoren einbauen. Wenn man neben der Suche nach einem Extremwert von $F[\phi]$ auch noch eine Nebenbedingung $G[\phi] = c$ zu erfüllen hat, wendet man das Variationsverfahren auf das Funktional $F[\phi] + \lambda G[\phi]$ an. Dabei ist der Lagrange-Multiplikator λ ein freier Parameter, dessen Wert am Ende so festgelegt wird, dass die Nebenbedingung erfüllt ist.
- Die Funktionalableitung lässt sich auf beliebige Operatoren $A: X \rightarrow Y$ zwischen normierten Räumen verallgemeinern. A heißt dann differenzierbar an der Stelle $f \in X$, wenn es einen stetigen linearen Operator $A'(f)$ gibt, sodass

$$A(f+h) = A(f) + A'(f)h + r(h), \quad \lim_{\|h\|_X \rightarrow 0} \frac{\|r(h)\|_Y}{\|h\|_X} = 0. \quad (4.99)$$

Der Operator $A'(f)$ wird **Fréchet-Ableitung** genannt. Ein noch allgemeinerer Ableitungsbegriff, der auch ohne eine Norm auskommt, ist die **Gâteaux-Ableitung**

$$\delta A(f, h) = \lim_{\epsilon \rightarrow 0} \frac{A(f + \epsilon h) - A(f)}{\epsilon} = \left. \frac{d}{d\epsilon} \right|_{\epsilon=0} A(f + \epsilon h). \quad (4.100)$$

Für Funktionen auf dem \mathbb{R}^n entspricht das der Richtungsableitung $\nabla A \cdot \mathbf{h}$, und für einen Fréchet-differenzierbaren Operator ist $\delta A(f, h) = A'(f)h$.

Numerische Optimierung. Viele Verfahren zur numerischen Optimierung basieren auf dem Minimieren von Funktionalen. Bildet man bei einem nichtlinearen System von (algebraischen oder Differenzial- bzw. Integral-) Gleichungen, die von Variablen $\mathbf{x} = (x_1, \dots, x_N)$ abhängen, die Differenz aus linker und rechter Seite, reduziert sich das Problem auf eine Nullstellensuche: wir suchen jenen Vektor \mathbf{x} , für den $\mathbf{F}(\mathbf{x}) = 0$ erfüllt ist. Linearisiert man $\mathbf{F}(\mathbf{x})$, wobei \mathbf{J} mit $J_{ij} = \partial F_i / \partial x_j$ die Jacobimatrix bezeichnet, bietet sich dazu die **Newton-Methode** an:

$$\mathbf{F}(\mathbf{x}_0 + \delta \mathbf{x}) = \underbrace{\mathbf{F}(\mathbf{x}_0) + \mathbf{J}(\mathbf{x}_0) \delta \mathbf{x}}_{\stackrel{!}{=} 0} + \mathcal{O}(x_i x_j) \Rightarrow \delta \mathbf{x} \approx -\mathbf{J}(\mathbf{x}_0)^{-1} \mathbf{F}(\mathbf{x}_0). \quad (4.101)$$

Hat man $\mathbf{F}(\mathbf{x}_0)$ und $\mathbf{J}(\mathbf{x}_0)$ an einem Startwert \mathbf{x}_0 bestimmt, erhält man so eine neue Suchrichtung $\delta \mathbf{x}$ und kann $\mathbf{F}(\mathbf{x})$ und $\mathbf{J}(\mathbf{x})$ am neuen Punkt $\mathbf{x} = \mathbf{x}_0 + \delta \mathbf{x}$ erneut berechnen. Damit lässt sich das Verfahren iterativ wiederholen, bis das System konvergiert.

Andererseits lässt sich das Problem auch als Variationsproblem interpretieren: wir suchen nach einem stationären Punkt des Funktionals $f(\mathbf{x}) = |\mathbf{F}(\mathbf{x})|^2$. Das entspricht einer Nullstellensuche für den Gradienten ∇f . Auch hier lässt sich die Newton-Methode anwenden,

$$\nabla f(\mathbf{x}_0 + \delta \mathbf{x}) = \nabla f(\mathbf{x}_0) + \mathbf{H}(\mathbf{x}_0) \delta \mathbf{x} + \dots \Rightarrow \delta \mathbf{x} \approx -\mathbf{H}(\mathbf{x}_0)^{-1} \nabla f(\mathbf{x}_0), \quad (4.102)$$

wobei \mathbf{H} die Hesse-Matrix $H_{ij} = \partial^2 f / \partial x_i \partial x_j$ der zweiten Ableitungen bezeichnet. Um die (numerisch oft aufwendige) Berechnung von \mathbf{H} zu vermeiden, bieten sich weitere Zugänge wie der **Gauß-Newton-** oder **Levenberg-Marquardt-Algorithmus** an. Solche Verfahren finden allerdings nur lokale Minima und nicht notwendigerweise das globale Minimum; für diese Aufgabe eignen sich stochastische Optimierungsmethoden.

Kapitel 5

Lineare Operatoren in Hilberträumen

Ein zentrales Thema der Funktionalanalysis ist die Diskussion von **linearen Operatoren** und deren Spektren. Wir haben schon Beispiele solcher Operatoren kennengelernt:

- Matrizen sind lineare Operatoren, die Elemente des \mathbb{R}^n oder \mathbb{C}^n aufeinander abbilden: $\mathbf{A}u = v$. Operatoren auf dem Folgenraum ℓ^2 sind unendlichdimensionale Matrizen.
- Die in Kapitel 3 besprochenen **Integraltransformationen** sind lineare Operatoren, die Funktionen u, v aus bestimmten Funktionenräumen aufeinander abbilden: $Au = v$.
- Die in Kapitel 4 besprochenen **Funktionale** sind Operatoren, die Elemente u eines Vektorraums auf Skalare v abbilden: $A[u] = v$.
- Oft ist man an **Differenzialgleichungen** oder **Integralgleichungen** der allgemeinen Form $(A - z)u = f$ interessiert, wobei f eine Inhomogenität darstellt, z eine komplexe Zahl und A einen linearen Differenzial- oder Integraloperator, z.B.

$$Au(x) = \sum_{k=1}^n a_k(x) \frac{d^k}{dx^k} u(x) \quad \text{oder} \quad Au(x) = \int_a^x dy K(x, y) u(y). \quad (5.1)$$

Der erste Fall beschreibt eine allgemeine lineare Differenzialgleichung n -ter Ordnung. Der zweite Fall ist als **Volterra-Gleichung** bekannt, bzw. wenn man die obere Integralgrenze x durch eine Konstante b ersetzt, als **Fredholm-Gleichung**.

Die Betrachtung von Hilberträumen bringt Vorteile beim Lösen solcher Probleme. Das Ziel dieses Kapitels ist eine Charakterisierung von Operatoren und ihren Eigenschaften, weswegen wir neue Begriffe einführen werden: beschränkte, unbeschränkte, abgeschlossene, kompakte, symmetrische und selbstadjungierte Operatoren, usw. Eine wichtige Frage wird sich um die Lösbarkeit von Gleichungen der Form $(A - z)u = f$ drehen, d.h. unter welchen Bedingungen der inverse Operator $(A - z)^{-1}$ existiert. Offenbar ist der Operator nicht invertierbar, wenn z ein Eigenwert von A ist, denn dann löst es die homogene Gleichung $Au = zu$. Gibt es auch andere Fälle? Mit dieser Verallgemeinerung des Eigenwertproblems auf unendlichdimensionale Räume beschäftigt sich die **Spektraltheorie**.

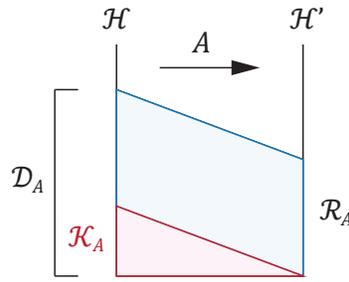


FIG. 5.1: Illustration von Definitionsbereich \mathcal{D}_A , Wertebereich \mathcal{R}_A und Nullraum \mathcal{K}_A

5.1 Grundbegriffe

Wir werden uns im Folgenden ausschließlich auf lineare Operatoren beschränken, denn diese sind mit der linearen Struktur eines Vektorraums verträglich.

DEFINITION: Seien \mathcal{H} und \mathcal{H}' Hilberträume über dem Körper \mathbb{K} der reellen oder komplexen Zahlen. Ein **linearer Operator** $A: \mathcal{D}_A \rightarrow \mathcal{H}'$ ordnet einem Element $u \in \mathcal{D}_A \subset \mathcal{H}$ ein Element $v = Au \in \mathcal{H}'$ mit folgenden Eigenschaften zu:

$$A(u_1 + u_2) = Au_1 + Au_2 \quad \forall u_1, u_2 \in \mathcal{D}_A, \quad (5.2)$$

$$A(\lambda u) = \lambda Au \quad \forall u \in \mathcal{D}_A, \quad \lambda \in \mathbb{K}. \quad (5.3)$$

Dabei bezeichnet $\mathcal{D}_A \subset \mathcal{H}$ den **Definitionsbereich** (*domain*) des Operators A ; dieser ist immer gemeinsam mit dem Operator anzugeben.

Der Definitionsbereich ist somit implizit in der Definition des Operators A . Zwei Operatoren A_1 und A_2 sind nur dann gleich, wenn ihre Definitionsbereiche übereinstimmen und sie auf alle Elemente dieses Definitionsbereichs dieselbe Wirkung haben:

$$A_1 = A_2 \quad \Leftrightarrow \quad \mathcal{D}_{A_1} = \mathcal{D}_{A_2} \quad \text{und} \quad A_1 u = A_2 u \quad \forall u \in \mathcal{D}_{A_1} = \mathcal{D}_{A_2}. \quad (5.4)$$

Insbesondere sind A_1 und A_2 mit $\mathcal{D}_{A_1} \neq \mathcal{D}_{A_2}$ *verschiedene* Operatoren, selbst wenn sie derselben Operatorvorschrift $A_1 u = A_2 u$ genügen. Wir werden uns später konkrete Beispiele dazu ansehen. Es ist nicht immer möglich, einen Operator auf ganz \mathcal{H} zu definieren, z.B. kann ein Differenzialoperator nur auf einer Untermenge von $L^2(a, b)$ definiert werden, eben den differenzierbaren Funktionen. Solche Situationen können zu Problemen führen, denn damit die Summe zweier Operatoren $A + B$ gebildet werden kann, muss u in $\mathcal{D}_A \cap \mathcal{D}_B$ liegen, und damit die Hintereinanderausführung BA wohldefiniert ist, muss u in \mathcal{D}_A und Au in \mathcal{D}_B liegen. Die Multiplikation von A mit einer Zahl $\lambda \in \mathbb{K}$ ändert nichts am Definitionsbereich, d.h. $\mathcal{D}_{\lambda A} = \mathcal{D}_A$.

Das **Bild** bzw. der **Wertebereich** (*image, range*) von A und der **Kern** bzw. **Nullraum** (*kernel, null space*) von A sind definiert durch

$$\mathcal{R}_A := \{Au \in \mathcal{H}' \mid u \in \mathcal{D}_A\}, \quad \mathcal{K}_A := \{u \in \mathcal{D}_A \mid Au = 0\}. \quad (5.5)$$

Gängige alternative Schreibweisen sind auch $\mathcal{D}_A = \text{dom } A$, $\mathcal{R}_A = \text{ran } A$ und $\mathcal{K}_A = \text{ker } A$. Die Verhältnisse sind in Fig. 5.1 skizziert. Aus der Linearität des Operators A folgt sofort, dass \mathcal{K}_A und \mathcal{R}_A Untervektorräume von \mathcal{H} bzw. \mathcal{H}' sind.

Zwei triviale Beispiele für lineare Operatoren $A: \mathcal{H} \rightarrow \mathcal{H}$ sind:

- Der **Nulloperator** $A = 0$ mit $v = Au = 0$ für alle $u \in \mathcal{H}$.
Hier ist $\mathcal{D}_A = \mathcal{K}_A = \mathcal{H}$ und $\mathcal{R}_A = \{0\}$.
- Die **Identität** $A = \mathbb{1}$ mit $v = Au = u$ für alle $u \in \mathcal{H}$.
Es gilt $\mathcal{D}_A = \mathcal{R}_A = \mathcal{H}$ und $\mathcal{K}_A = \{0\}$.

Setzt man in (5.3) $\lambda = 0$, folgt insbesondere $A(0) = 0$ für lineare Operatoren. Beispielsweise haben alle linearen Operatoren $A: \mathbb{R} \rightarrow \mathbb{R}$ die Form $Ax = cx$ mit $c \in \mathbb{R}$; ihre Graphen sind Geraden mit Steigung c , die durch den Ursprung gehen. Eine Gerade, die nicht durch den Ursprung geht ($Ax = cx + d$ mit $d \neq 0$) ist keine lineare Abbildung mehr. Ebenso ist z.B. ein Operator $A: L^2(a, b) \rightarrow L^2(a, b)$ mit $Af(x) = f(x)^2$ kein linearer Operator.

Endlichdimensionale Vektorräume. Betrachten wir zunächst lineare Operatoren auf endlichdimensionalen Vektorräumen. Hier müssen wir uns um die Feinheiten, die uns später begegnen werden, keine Gedanken machen, denn Operatoren auf endlichdimensionalen Vektorräumen sind automatisch beschränkt und somit stetig.

Jeder endlich-dimensionale Hilbertraum \mathcal{H} ist isomorph zu \mathbb{C}^n oder \mathbb{R}^n , deswegen lässt sich auch jede lineare Abbildung $A: \mathcal{H} \rightarrow \mathcal{H}$ in die Gestalt einer Matrixmultiplikation bringen. Zerlegt man einen Vektor $u \in \mathcal{H}$ in eine Orthonormalbasis $\{\varphi_j\}$, dann gilt

$$u = \sum_{j=1}^n \alpha_j \varphi_j \quad \Leftrightarrow \quad \alpha_j = \langle \varphi_j, u \rangle, \quad v = Au = \sum_{j=1}^n \alpha_j A\varphi_j, \quad (5.6)$$

wobei jedes $A\varphi_j$ wieder ein Element von \mathcal{H} ist und eine Komponentendarstellung besitzt:

$$A\varphi_j = \sum_{i=1}^n A_{ij} \varphi_i \quad \Leftrightarrow \quad A_{ij} = \langle \varphi_i, A\varphi_j \rangle. \quad (5.7)$$

Dadurch wird die Matrixdarstellung $\mathbf{A} = (A_{ij})$ des Operators A bezüglich der ONB $\{\varphi_i\}$ definiert. Eingesetzt in (5.6) erhalten wir die Komponentenzerlegung von v ,

$$v = \sum_{j=1}^n \alpha_j \sum_{i=1}^n A_{ij} \varphi_i = \sum_{i=1}^n \left(\sum_{j=1}^n A_{ij} \alpha_j \right) \varphi_i = \sum_{i=1}^n \beta_i \varphi_i, \quad (5.8)$$

und somit für die Komponenten selbst:

$$\beta_i = \sum_{j=1}^n A_{ij} \alpha_j \quad \Leftrightarrow \quad \boldsymbol{\beta} = \mathbf{A} \boldsymbol{\alpha}. \quad (5.9)$$

Die komplexe oder reelle $n \times n$ -Matrix \mathbf{A} charakterisiert den Operator A vollständig: jede lineare Transformation A erzeugt eine eindeutige Matrix \mathbf{A} , und jede Matrix \mathbf{A} erzeugt eine eindeutige lineare Transformation A . Zusammenfassend haben wir also

$$\alpha_i = \langle \varphi_i, u \rangle, \quad \beta_i = \langle \varphi_i, v \rangle, \quad A_{ij} = \langle \varphi_i, A\varphi_j \rangle, \quad v = Au \quad \Leftrightarrow \quad \boldsymbol{\beta} = \mathbf{A} \boldsymbol{\alpha}. \quad (5.10)$$

Damit lässt sich auch der Operator A durch die ONB ausdrücken,

$$Au = \sum_{j=1}^n \alpha_j A\varphi_j = \sum_{ij} A_{ij} \alpha_j \varphi_i = \sum_{ij} A_{ij} \varphi_i \langle \varphi_j, u \rangle \quad \forall u \in \mathcal{H}, \quad (5.11)$$

was sich am einfachsten in der Bra-Ket-Sprache formulieren lässt:

$$A = \left(\sum_{i=1}^n |\varphi_i\rangle\langle\varphi_i| \right) A \left(\sum_{j=1}^n |\varphi_j\rangle\langle\varphi_j| \right) = \sum_{ij} A_{ij} |\varphi_i\rangle\langle\varphi_j|. \quad (5.12)$$

Alles Weitere entspricht dann den Regeln der linearen Algebra:

- Die **adjungierte Matrix** $\mathbf{A}^\dagger = (\mathbf{A}^*)^T \Leftrightarrow (A^\dagger)_{ij} = A_{ji}^*$ ist die transponierte und komplex konjugierte Matrix. Für eine reelle Matrix entspricht die Adjunktion der Transposition: $\mathbf{A}^\dagger = \mathbf{A}^T$.
- Eine Matrix heißt **hermitisch**, wenn $\mathbf{A}^\dagger = \mathbf{A}$ gilt. Hermitesche Matrizen haben reelle Eigenwerte, was besonders wichtig ist, da sie oft physikalische Observablen darstellen. Eine reelle hermitesche Matrix ist symmetrisch: $\mathbf{A}^T = \mathbf{A}$.
- Eine Matrix heißt **unitär**, wenn $\mathbf{A}\mathbf{A}^\dagger = \mathbf{A}^\dagger\mathbf{A} = \mathbb{1}$ gilt, also $\mathbf{A}^{-1} = \mathbf{A}^\dagger$. Für eine reelle Matrix vereinfacht sich die Unitarität zur Orthogonalität: $\mathbf{A}^{-1} = \mathbf{A}^T$.
- Eine Matrix heißt **normal**, wenn $\mathbf{A}\mathbf{A}^\dagger = \mathbf{A}^\dagger\mathbf{A}$ gilt. Normale Matrizen sind unitär diagonalisierbar, d.h. es gibt eine unitäre Matrix \mathbf{U} sodass $\mathbf{A} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^\dagger$, wobei $\mathbf{\Lambda}$ eine Diagonalmatrix ist. Jede hermitesche und jede unitäre Matrix ist normal.

Betrachten wir als Beispiel die lineare Transformation $A: \mathbb{R}^2 \rightarrow \mathbb{R}^2$, die jeden zweidimensionalen Vektor um den Winkel θ gegen den Uhrzeigersinn rotiert. Wie lautet ihre Matrixdarstellung in der kartesischen Basis $\varphi_1 = (1, 0)$, $\varphi_2 = (0, 1)$? Mit Hilfe einer Skizze sieht man

$$\begin{aligned} A\varphi_1 &= (\cos\theta)\varphi_1 + (\sin\theta)\varphi_2, \\ A\varphi_2 &= (-\sin\theta)\varphi_1 + (\cos\theta)\varphi_2 \end{aligned} \quad \Rightarrow \quad \mathbf{A} = (A_{ij}) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}. \quad (5.13)$$

Die Koordinaten des gedrehten Vektors hängen dann mit jenen des ursprünglichen Vektors über $\beta = \mathbf{A}\alpha$ zusammen. \mathbf{A} ist eine unitäre (bzw. orthogonale, da reelle) Matrix, und es ist $\mathcal{K}_A = \{0\}$ und $\mathcal{R}_A = \mathbb{R}^2$. Dieselbe Transformation im Raum \mathbb{C}^2 hat zwar keine geometrische Interpretation, liefert aber dasselbe Ergebnis.

Wie steht es mit der Transformation $A: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ mit der Matrixdarstellung

$$\mathbf{A} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad (5.14)$$

relativ zur kartesischen Basis? Die Koordinaten des transformierten Vektors lauten jetzt $\beta_1 = \alpha_2$, $\beta_2 = 0$, und der zugehörige Operator ist $A\varphi_1 = 0$, $A\varphi_2 = \varphi_1$. Sowohl \mathcal{K}_A als auch \mathcal{R}_A bestehen aus allen Vektoren proportional zu φ_1 und sind somit eindimensionale Untervektorräume.

Betrachten wir den endlichdimensionalen Hilbertraum, der durch die ONB $\{\varphi_n\}$ mit $\varphi_n(x) = \sin nx$ und $n < \infty$ aufgespannt wird. Das ist ein Teilraum des Raums der ungeraden, periodischen Funktionen auf $L^2(-\pi, \pi)$, die am Rand verschwinden (siehe die frühere Diskussion der Fourierreihe in (2.63)). Wie lautet die Matrixdarstellung des Differenzialoperators $A = d^2/dx^2$ in dieser Basis? Es ist

$$A_{kl} = \left\langle \varphi_k, \frac{d^2\varphi_l}{dx^2} \right\rangle = \frac{1}{\pi} \int_{-\pi}^{\pi} dx \sin(kx) \frac{d^2}{dx^2} \sin(lx) = -\frac{l^2}{\pi} \int_{-\pi}^{\pi} dx \sin(kx) \sin(lx) = -k^2 \delta_{kl} \quad (5.15)$$

und somit

$$\mathbf{A} = \begin{pmatrix} -1 & 0 & 0 & \dots \\ 0 & -4 & 0 & \dots \\ 0 & 0 & -9 & \dots \\ \dots & \dots & \dots & \dots \end{pmatrix}. \quad (5.16)$$

5.2 Beschränkte Operatoren

Wie sieht es nun mit Operatoren auf unendlich-dimensionalen Hilberträumen aus? Um die Rechenregeln der linearen Algebra weitgehend übernehmen zu können, schränken wir uns zunächst auf die Klasse der **beschränkten Operatoren** ein:

DEFINITION: Seien \mathcal{H} und \mathcal{H}' Hilberträume. Ein Operator $A: \mathcal{D}_A \rightarrow \mathcal{H}'$ mit $\mathcal{D}_A \subset \mathcal{H}$ heißt **beschränkt**, falls es für alle $u \in \mathcal{D}_A$ eine Konstante $c \geq 0$ gibt, sodass $\|Au\| \leq c\|u\|$, d.h. die Größe $\|Au\|/\|u\|$ ist von oben beschränkt. Die kleinste Zahl c , die diese Ungleichung für alle $u \in \mathcal{D}_A$ erfüllt, heißt **Operatornorm** $\|A\|$ von A :

$$\|A\| = \sup_{\|u\| \neq 0} \frac{\|Au\|}{\|u\|} = \sup_{\|u\|=1} \|Au\|. \quad (5.17)$$

(Da man jeden Vektor $u \neq 0$ auf 1 normieren kann, kann man genauso gut das Supremum über alle Einheitsvektoren betrachten.)

Da \mathcal{H} und \mathcal{H}' im Prinzip verschiedene Hilberträume sein können, bezieht sich die Norm des Vektors u auf den Hilbertraum \mathcal{H} und jene von Au auf \mathcal{H}' , daher sollten wir eigentlich präziser schreiben: $\|Au\|_{\mathcal{H}'} \leq c\|u\|_{\mathcal{H}}$.

Beschränkte Operatoren haben also eine endliche Norm. Die Bedeutung der Operatornorm wird aus Fig. 5.2 klar. Das Supremum $\|A\|$ wird in der Regel nicht für jedes Element u erreicht werden, bzw. muss es auch für gar kein u erreicht werden (deswegen auch "sup" statt "max"). Um zu zeigen, dass $\|A\| =: c$ die Operatornorm ist, muss man zunächst beweisen, dass c tatsächlich eine obere Schranke für $\|Au\|/\|u\|$ darstellt (d.h. $\|Au\| \leq c\|u\|$ für alle $u \in \mathcal{D}_A$). Zweitens muss man zeigen: entweder es gibt ein $u_0 \in \mathcal{D}_A$ mit $u_0 \neq 0$, für das diese obere Schranke tatsächlich erreicht wird, also $\|Au_0\|/\|u_0\| = c$; oder es gibt eine Folge $(u_n) \subset \mathcal{D}_A$ mit $\|u_n\| \neq 0$, sodass $\|Au_n\|/\|u_n\| \rightarrow c$ für $n \rightarrow \infty$.

Betrachten wir als Beispiel den Hilbertraum $L^2(0, 1)$ und den Operator A , der u mit x multipliziert:

$$v(x) = (Au)(x) = xu(x). \quad (5.18)$$

Der Operator ist linear und die Multiplikation für alle $u \in L^2(0, 1)$ wohldefiniert. Wegen

$$\|Au\|^2 = \int_0^1 dx x^2 |u(x)|^2 \leq \int_0^1 dx |u(x)|^2 = \|u\|^2 \quad (5.19)$$

ist A beschränkt und wir können mit Sicherheit sagen, dass $\|A\| \leq 1$ gilt. Ist die Operatornorm dann $\|A\| = 1$? Dazu müssten wir ein u_0 finden, sodass $\|Au_0\| = \|u_0\|$, was wegen

$$\|u_0\|^2 - \|Au_0\|^2 = \int_0^1 dx |u_0(x)|^2 (1 - x^2) > 0 \quad (5.20)$$

offenbar aber nicht möglich ist, da beide Faktoren im Integranden positiv sind. Versuchen wir es daher mit einer Folge, z.B. $u_n(x) = x^n$ mit $n > -\frac{1}{2}$ (ansonsten wäre u_n nicht quadratintegrierbar). Dann gilt

$$\|u_n\|^2 = \int_0^1 dx x^{2n} = \frac{1}{2n+1}, \quad \|Au_n\|^2 = \int_0^1 dx x^{2+2n} = \frac{1}{2n+3} \Rightarrow \frac{\|Au_n\|}{\|u_n\|} = \sqrt{\frac{2n+1}{2n+3}} \xrightarrow{n \rightarrow \infty} 1.$$

Diese Folge erreicht im Limes $n \rightarrow \infty$ also den Maximalwert 1 (siehe Fig. 5.2 links), und somit lautet auch die Operatornorm $\|A\| = 1$.

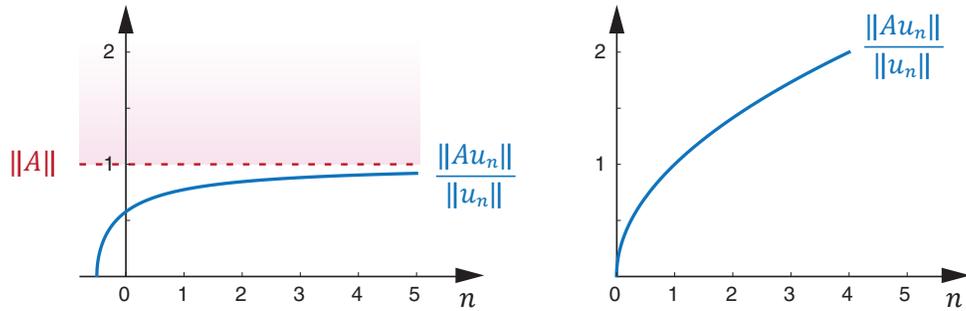


FIG. 5.2: Bei beschränkten Operatoren ist $\|Au\|/\|u\|$ von oben beschränkt, bei unbeschränkten Operatoren nicht. Die kleinste obere Schranke ist die Operatornorm $\|A\|$

Um zu zeigen, dass A **unbeschränkt** ist, muss man lediglich *eine* Folge (u_n) mit $\|u_n\| \neq 0$ finden, sodass $\|Au_n\|/\|u_n\| \rightarrow \infty$.

Untersuchen wir den Operator A im Hilbertraum $L^2(0, 1)$, der u mit $1/x$ multipliziert, also $(Au)(x) = u(x)/x$. Das ist nicht mehr für alle $u \in L^2(0, 1)$ wohldefiniert, deswegen müssen wir den Definitionsbereich \mathcal{D}_A einschränken. Wenn $u(x)$ für $x \rightarrow 0$ hinreichend schnell verschwindet, dann ist $u/x \in L^2(0, 1)$, daher könnten wir verlangen:

$$\mathcal{D}_A = \left\{ u \in L^2(0, 1) \mid \frac{u}{x} \in L^2(0, 1) \right\}. \quad (5.21)$$

Das Analogon zu (5.19) liefert keine obere Schranke, im Gegenteil: es sieht so aus, als würde die Norm $\|Au\|$ bereits als Limes einfacher Funktionen divergieren. Betrachten wir z.B. die Folge

$$u_n(x) = \begin{cases} 0 & 0 < x < \frac{1}{n} \\ 1 & \frac{1}{n} < x < 1 \end{cases} \quad \text{mit } n > 1, \quad (5.22)$$

dann sehen wir, dass A ein unbeschränkter Operator ist (siehe Fig. 5.2 rechts):

$$\|u_n\|^2 = \int_{1/n}^1 dx = \frac{n-1}{n}, \quad \|Au_n\|^2 = \int_{1/n}^1 \frac{dx}{x^2} = n-1 \quad \Rightarrow \quad \frac{\|Au_n\|}{\|u_n\|} = \sqrt{n} \xrightarrow{n \rightarrow \infty} \infty.$$

Bei unbeschränkten Operatoren gibt es Probleme mit dem Definitionsbereich, was zu Schwierigkeiten führt, da man nicht mehr bedenkenlos Summen und Produkte von Operatoren bilden kann. Auch Ableitungsoperatoren sind von dieser Art, wie das folgende Beispiel zeigt:

Für den Ableitungsoperator $A : C^1(0, 1) \rightarrow C^0(0, 1)$ mit $Au = u'$ gilt $\mathcal{D}_A = C^1(0, 1)$ und $\mathcal{R}_A = C^0(0, 1)$, während der Nullraum \mathcal{K}_A die Menge aller konstanten Funktionen $u = \text{const.}$ ist. Ob dieser Operator beschränkt ist, hängt stark von der Wahl der Normen ab. Wählt man die L^2 -Hilbertraumnorm, dann ist der Ableitungsoperator nicht beschränkt, wie man mit Hilfe der Folge $u_n(x) = \sin(n\pi x)$ sehen kann:

$$\|u_n\|^2 = \int_0^1 dx \sin^2(n\pi x) = \frac{1}{2}, \quad \|Au_n\|^2 = \int_0^1 dx (n\pi)^2 \cos^2(n\pi x) = \frac{(n\pi)^2}{2} \quad \Rightarrow \quad \frac{\|Au_n\|}{\|u_n\|} = |n\pi|. \quad (5.23)$$

Mit der folgenden Wahl der Normen kann man allerdings zeigen, dass der Ableitungsoperator beschränkt ist:

$$\|u\|_{C^1} = \|u\|_{C^0} + \|u'\|_{C^0}, \quad \|u\|_{C^0} = \|u\|_{C^1}. \quad (5.24)$$

Wir werden auf unbeschränkte Operatoren in Abschnitt 5.3 zurückkommen. Bei beschränkten Operatoren kann man dagegen immer davon ausgehen, dass sie auf ganz \mathcal{H} definiert sind, was vieles erleichtert. Insbesondere sind in endlich-dimensionalen Räumen alle linearen Operatoren beschränkt. Wir werden die Menge aller linearen und beschränkten Operatoren $A : \mathcal{H} \rightarrow \mathcal{H}'$ mit $\mathcal{B}(\mathcal{H}, \mathcal{H}')$ bezeichnen.

Stetigkeit. Eine sehr nützliche Eigenschaft linearer Operatoren ist, dass für sie Beschränktheit synonym mit **Stetigkeit** ist.

DEFINITION: Ein linearer Operator A heißt **stetig**, falls für jede Nullfolge $(u_n) \subset \mathcal{D}_A$ gilt:

$$\lim_{n \rightarrow \infty} u_n = 0 \quad \Rightarrow \quad \lim_{n \rightarrow \infty} Au_n = 0. \quad (5.25)$$

Für lineare Operatoren ist die Stetigkeit am Ursprung äquivalent mit der Stetigkeit auf dem gesamten Definitionsbereich \mathcal{D}_A : für eine Folge (u_n) , die gegen einen Grenzwert $u \in \mathcal{D}_A$ konvergiert, bildet $(u_n - u) \rightarrow 0$ eine Nullfolge, und aus $A(u_n - u) \rightarrow 0$ folgt dann wegen der Linearität des Operators $Au_n \rightarrow Au$. Ist ein linearer Operator stetig am Ursprung, ist er also automatisch überall stetig:

$$\lim_{n \rightarrow \infty} u_n = u \quad \Rightarrow \quad \lim_{n \rightarrow \infty} Au_n = Au. \quad (5.26)$$

Die Stetigkeit von A bedeutet, dass man den Operator mit dem Grenzwert vertauschen kann, also $Au = A(\lim_{n \rightarrow \infty} u_n) = \lim_{n \rightarrow \infty} Au_n$. Die entscheidende Aussage ist aber folgende:

SATZ: Ein linearer Operator A ist genau dann stetig, wenn er beschränkt ist.

Dazu müssen wir beide Richtungen beweisen. Falls A beschränkt ist, gilt $\|Au_n\| \leq \|A\| \|u_n\|$; für eine Nullfolge mit $u_n \rightarrow 0$ folgt dann auch $Au_n \rightarrow 0$, also ist A stetig. Falls A unbeschränkt ist, dann gibt es eine Folge (u_n) mit $\|u_n\| \neq 0$ sodass $\|Au_n\| \geq n \|u_n\|$. Dann ist $v_n = u_n / (n \|u_n\|) \rightarrow 0$ eine Nullfolge, für die $\|Av_n\| \geq 1$ für alle n gilt; das verletzt (5.25) und somit ist A nicht stetig.

Matrixdarstellung. Die Äquivalenz von Beschränktheit und Stetigkeit erlaubt es uns, mit beschränkten Operatoren in separablen Hilberträumen wie mit unendlich-dimensionalen Matrizen zu rechnen. Betrachten wir einen linearen beschränkten Operator A , der auf ganz \mathcal{H} definiert ist, und sei $\{\varphi_j\}$ eine Orthonormalbasis in \mathcal{H} . Dann ist ein Hilbertraumvektor $u \in \mathcal{H}$ durch den Grenzwert der Folge seiner Partialsummen definiert:

$$u = \sum_{j=1}^{\infty} \alpha_j \varphi_j = \lim_{n \rightarrow \infty} \sum_{j=1}^n \alpha_j \varphi_j = \lim_{n \rightarrow \infty} u_n. \quad (5.27)$$

Wegen der Stetigkeit und Linearität von A können wir dann schreiben:

$$v = Au = \lim_{n \rightarrow \infty} Au_n = \lim_{n \rightarrow \infty} A \left(\sum_{j=1}^n \alpha_j \varphi_j \right) = \lim_{n \rightarrow \infty} \sum_{j=1}^n \alpha_j A\varphi_j = \sum_{j=1}^{\infty} \alpha_j A\varphi_j. \quad (5.28)$$

Wie im endlichdimensionalen Fall können die Vektoren $A\varphi_j$ nach der ONB $\{\varphi_j\}$ zerlegt werden, was die Matrixdarstellung \mathbf{A} des Operators A bezüglich der ONB $\{\varphi_i\}$ definiert:

$$A\varphi_j = \sum_{i=1}^{\infty} A_{ij} \varphi_i, \quad A_{ij} = \langle \varphi_i, A\varphi_j \rangle, \quad \mathbf{A} = (A_{ij}). \quad (5.29)$$

Eingesetzt in (5.28) erhalten wir die Komponentenzerlegung von v ,

$$v = \sum_{j=1}^{\infty} \alpha_j \sum_{i=1}^{\infty} A_{ij} \varphi_i = \sum_{i=1}^{\infty} \left(\sum_{j=1}^{\infty} A_{ij} \alpha_j \right) \varphi_i = \sum_{i=1}^{\infty} \beta_i \varphi_i, \quad (5.30)$$

und somit für die Komponenten selbst:

$$\beta_i = \sum_{j=1}^{\infty} A_{ij} \alpha_j \quad \Leftrightarrow \quad \boldsymbol{\beta} = \mathbf{A} \boldsymbol{\alpha}. \quad (5.31)$$

Ferner gilt, dass die Spalten der Matrix \mathbf{A} quadratsummierbar sind:

$$\sum_{i=1}^{\infty} |A_{ij}|^2 = \sum_{i=1}^{\infty} |\langle \varphi_i, A\varphi_j \rangle|^2 = \sum_{i=1}^{\infty} \langle A\varphi_j, \varphi_i \rangle \langle \varphi_i, A\varphi_j \rangle = \langle A\varphi_j, A\varphi_j \rangle = \|A\varphi_j\|^2 < \infty. \quad (5.32)$$

Dasselbe lässt sich auch für die Zeilen zeigen. Somit erzeugt jeder lineare beschränkte Operator A bezüglich einer gegebenen ONB eine unendlich-dimensionale Matrix \mathbf{A} , deren Spalten und Zeilen quadratsummierbar sind. Allerdings erzeugt nicht jede solche Matrix auch einen linearen und beschränkten Operator! Ein Gegenbeispiel ist eine Diagonalmatrix, deren Elemente entlang der Diagonalen unbeschränkt ansteigen.

Shift-Operatoren. Als Beispiel betrachten wir einen separablen Hilbertraum \mathcal{H} mit einer ONB $\{\varphi_i\}$. Dann ist der **Rechtshift-Operator** definiert durch $A\varphi_i = \varphi_{i+1}$ mit $\mathcal{D}_A = \mathcal{H}$. Wegen

$$u = \sum_{i=1}^{\infty} \alpha_i \varphi_i \quad \Rightarrow \quad Au = \sum_{i=1}^{\infty} \alpha_i \varphi_{i+1}, \quad \|Au\|^2 = \sum_{i=1}^{\infty} |\alpha_i|^2 = \|u\|^2 \quad (5.33)$$

ist A beschränkt mit der Operatornorm $\|A\| = 1$. Die Matrixdarstellung lautet

$$\begin{aligned} A_{ij} &= \langle \varphi_i, A\varphi_j \rangle = \langle \varphi_i, \varphi_{j+1} \rangle = \delta_{i,j+1}, \\ \beta_i &= \langle \varphi_i, Au \rangle = \sum_{j=1}^{\infty} \alpha_j \langle \varphi_i, \varphi_{j+1} \rangle = \begin{cases} 0 & i = 1, \\ \alpha_{i-1} & i > 1, \end{cases} \\ \boldsymbol{\beta} = \mathbf{A} \boldsymbol{\alpha} &= \begin{pmatrix} 0 & & & & \\ 1 & 0 & & & \\ & 1 & 0 & & \\ & & 1 & 0 & \\ & & & \ddots & \ddots \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \\ \vdots \end{pmatrix} = \begin{pmatrix} 0 \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \vdots \end{pmatrix}. \end{aligned} \quad (5.34)$$

Der Nullraum ist $\mathcal{K}_A = \{0\}$ und der Wertebereich \mathcal{R}_A ist das orthogonale Komplement $\{\varphi_1\}^\perp$, d.h. alle Vektoren mit $\beta_1 = 0$. Der **Linkshift-Operator** mit $A\varphi_1 = 0$ und $A\varphi_i = \varphi_{i-1}$ für $i > 1$ bewirkt analog

$$\boldsymbol{\beta} = \mathbf{A} \boldsymbol{\alpha} = \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & 0 & 1 & \\ & & & \ddots & \ddots \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \\ \vdots \end{pmatrix} = \begin{pmatrix} \alpha_2 \\ \alpha_3 \\ \alpha_4 \\ \alpha_5 \\ \vdots \end{pmatrix}. \quad (5.35)$$

Hier gilt $\mathcal{D}_A = \mathcal{R}_A = \mathcal{H}$, und der Nullraum \mathcal{K}_A ist die lineare Hülle von φ_1 .

Hilbert-Schmidt-Operatoren. Lässt sich die Operatornorm $\|A\|$ allgemein durch die Matrixelemente a_{ij} ausdrücken? Für endlich-dimensionale Matrizen ist das einfach, z.B. für die Drehmatrix in (5.13):

$$\|A\varphi_1\| = \|A\varphi_2\| = \sqrt{\cos^2 \theta + \sin^2 \theta} = 1 \quad \Rightarrow \quad \|A\| = 1. \quad (5.36)$$

Die Frage lässt sich auch bejahen, falls es eine ONB gibt, die den Operator A diagonalisiert. Dann gilt $A\varphi_i = \lambda_i \varphi_i$ und somit $\|A\varphi_i\| = |\lambda_i|$, sodass die Operatornorm (als Supremum über alle möglichen Einheitsvektoren φ_i) der **größte Absolutbetrag der Eigenwerte** ist: $\|A\| = \max |\lambda_i|$. Die Drehmatrix (5.13) hat z.B. Eigenwerte $e^{\pm i\theta} \Rightarrow \|A\| = 1$.

Im allgemeinen Fall ist die Angabe der Operatornorm schwieriger, wir können aber zumindest eine obere Schranke angeben. Zu diesem Zweck definieren wir für ein festes i einen Vektor $\mathbf{A}_i \in \ell^2$ mit Komponenten $(\mathbf{A}_i)_j = A_{ij}$, sodass $\|\mathbf{A}_i\|^2 = \sum_j |A_{ij}|^2$. Damit können wir nämlich (5.31) als Skalarprodukt im ℓ^2 schreiben und die Cauchy-Schwarz-Ungleichung (1.36) ausnützen:

$$\beta_i = \sum_j A_{ij} \alpha_j = \langle \mathbf{A}_i^*, \boldsymbol{\alpha} \rangle \quad \Rightarrow \quad |\beta_i| = |\langle \mathbf{A}_i^*, \boldsymbol{\alpha} \rangle| \leq \|\mathbf{A}_i\| \|\boldsymbol{\alpha}\| = \|\mathbf{A}_i\| \|u\|, \quad (5.37)$$

denn es ist ja $\|u\|^2 = \sum_{j=1} |\alpha_j|^2$ bezüglich der Norm in \mathcal{H} dasselbe wie $\|\boldsymbol{\alpha}\|^2$ bezüglich der Norm in ℓ^2 . Daraus erhalten wir eine obere Schranke für die Operatornorm:

$$\|Au\|^2 = \sum_i |\beta_i|^2 \leq \|u\|^2 \sum_i \|\mathbf{A}_i\|^2 = \|u\|^2 \sum_{ij} |A_{ij}|^2 \quad \Rightarrow \quad \|A\| \leq \sqrt{\sum_{ij} |A_{ij}|^2}. \quad (5.38)$$

Falls diese Größe endlich bleibt, ist der Operator sicher beschränkt. Operatoren dieser Art, bei denen die Matrix \mathbf{A} nicht nur in ihren Zeilen und Spalten quadratsummierbar ist wie in (5.32), sondern unter Summation *aller* Einträge konvergiert,

$$\text{Tr}(\mathbf{A}^\dagger \mathbf{A}) = \sum_{ij} |A_{ij}|^2 = \sum_{ij} |\langle \varphi_i, A\varphi_j \rangle|^2 = \sum_j \langle A\varphi_j, A\varphi_j \rangle = \sum_j \|A\varphi_j\|^2 < \infty, \quad (5.39)$$

heißen **Hilbert-Schmidt-Operatoren**. In vielen Fällen wird diese Grenze sehr großzügig gewählt sein, z.B. bekommen wir für die Drehmatrix (5.13)

$$\sum_{i=1}^2 \sum_{j=1}^2 |A_{ij}|^2 = 2(\cos^2 \theta + \sin^2 \theta) = 2 \quad \Rightarrow \quad \|A\| \leq \sqrt{2}, \quad (5.40)$$

aber wir wissen ja bereits, dass in diesem Fall $\|A\| = 1$ gilt. Andererseits erfüllen die vorhin betrachteten Shift-Operatoren die Eigenschaft (5.39) offenbar nicht, da man unendlich oft die 1 aufsummieren würde.

Hilbert-Schmidt-Operatoren liegen oft als **Integraloperatoren** $A: L^2(\Omega) \rightarrow L^2(\Omega)$ vor. Dabei ist $\Omega = (a, b)$ ein reelles Intervall und der Integralkern $K \in L^2(\Omega \times \Omega)$ eine reelle Funktion in den zwei Variablen $x, y \in \Omega$,

$$v(x) = Au(x) = \int_{\Omega} dy K(x, y) u(y) \quad \text{mit} \quad \mathcal{D}_A = L^2(\Omega), \quad (5.41)$$

die folgende Eigenschaft erfüllt:

$$\int_{\Omega} dx \int_{\Omega} dy |K(x, y)|^2 < \infty. \quad (5.42)$$

Hier verläuft der Beweis der Beschränktheit gleich wie vorher: (5.41) lässt sich für ein festes $x \in L^2(\Omega)$ als Skalarprodukt schreiben, womit wir die Cauchy-Schwarz-Ungleichung anwenden können:

$$v(x) = Au(x) = \langle K(x, \cdot)^*, u \rangle \quad \Rightarrow \quad |v(x)| = |\langle K(x, \cdot)^*, u \rangle| \leq \|K(x, \cdot)\| \|u\|. \quad (5.43)$$

Dann folgt

$$\|Au\|^2 = \int_{\Omega} dx |v(x)|^2 \leq \|u\|^2 \int_{\Omega} dx \|K(x, \cdot)\|^2 = \|u\|^2 \int_{\Omega} dx \int_{\Omega} dy |K(x, y)|^2 < \infty, \quad (5.44)$$

wobei das Integral auf der rechten Seite laut Voraussetzung (5.42) endlich ist. Damit ist der Operator A beschränkt und für die Operatornorm gilt $\|A\|^2 \leq \int_{\Omega} dx \int_{\Omega} dy |K(x, y)|^2$.

Insbesondere kann man zeigen, dass die obere Schranke der Summe über alle Einheitsvektoren $u = \varphi_j$ mit $v = A\varphi_j$ entspricht:

$$\begin{aligned} \sum_j \|A\varphi_j\|^2 &= \sum_j \int_{\Omega} dx |\langle K(x, \cdot)^*, \varphi_j \rangle|^2 = \int_{\Omega} dx \sum_j \langle K(x, \cdot)^*, \varphi_j \rangle \langle \varphi_j, K(x, \cdot)^* \rangle \\ &= \int_{\Omega} dx \langle K(x, \cdot)^*, K(x, \cdot)^* \rangle = \int_{\Omega} dx \|K(x, \cdot)\|^2 = \int_{\Omega} dx \int_{\Omega} dy |K(x, y)|^2. \end{aligned} \quad (5.45)$$

Aus (5.39) und (5.45) folgt dann die Gleichheit der beiden Ausdrücke in der Matrix- und Integraldarstellung:

$$\|A\|^2 \leq \sum_{ij} |A_{ij}|^2 = \sum_j \|A\varphi_j\|^2 = \int_{\Omega} dx \int_{\Omega} dy |K(x, y)|^2 < \infty. \quad (5.46)$$

Die Matrixelemente von \mathbf{A} lauten

$$A_{ij} = \langle \varphi_i, A\varphi_j \rangle = \int_{\Omega} dx \int_{\Omega} dy \varphi_i^*(x) K(x, y) \varphi_j(y). \quad (5.47)$$

Ein Spezialfall eines Hilbert-Schmidt-Operators ist der folgende Integraloperator mit der Stufenfunktion als Integralkern:

$$v(x) = Au(x) = \int_0^x dy u(y) = \int_0^1 dy \Theta(x - y) u(y) \quad (5.48)$$

Er ist für alle $u \in L^2(0, 1)$ definiert, da die Quadratintegrierbarkeit von u auf dem Intervall $(0, 1)$ auch die Integrierbarkeit impliziert. Es gilt

$$\|A\|^2 \leq \int_0^1 dx \int_0^1 dy \Theta(x - y)^2 = \int_0^1 dx \int_0^x dy = \int_0^1 dx x = \frac{1}{2}. \quad (5.49)$$

Funktionen beschränkter Operatoren. Die Menge aller linearen, beschränkten Operatoren auf einem Hilbertraum \mathcal{H} wird mit

$$\mathcal{B}(\mathcal{H}) := \mathcal{B}(\mathcal{H}, \mathcal{H}) = \{A: \mathcal{H} \rightarrow \mathcal{H} \mid A \text{ ist linear und beschränkt}\} \quad (5.50)$$

bezeichnet und bildet selbst wieder einen Banachraum mit der Operatornorm $\|A\|$. Er erfüllt alle Voraussetzungen eines Vektorraums, z.B. sind λA und $A_1 + A_2$ wieder Elemente von $\mathcal{B}(\mathcal{H})$. Die Operatornorm erfüllt alle Eigenschaften einer Norm: es gilt $\|A\| \geq 0$, $\|A\| = 0 \Leftrightarrow A = 0$, $\|\lambda A\| = |\lambda| \|A\|$ sowie die Dreiecksungleichung

$$\|A_1 + A_2\| = \sup_{\|u\|=1} \|A_1 u + A_2 u\| \leq \sup_{\|u\|=1} (\|A_1 u\| + \|A_2 u\|) = \|A_1\| + \|A_2\|. \quad (5.51)$$

Daraus sieht man, dass $A_1 + A_2$ wieder ein beschränkter Operator ist. Man kann auch zeigen, dass der Raum $\mathcal{B}(\mathcal{H})$ bezüglich der Operatornorm vollständig ist.

$\mathcal{B}(\mathcal{H})$ besitzt aber noch weitere Eigenschaften, denn wir können Operatoren ja nicht nur addieren und mit Skalaren multiplizieren, sondern auch *miteinander* multiplizieren. Das Produkt AB mit $A, B \in \mathcal{B}(\mathcal{H})$ ist wieder ein beschränkter Operator aus $\mathcal{B}(\mathcal{H})$, denn es gilt

$$\|ABu\| \leq \|A\| \|Bu\| \leq \|A\| \|B\| \|u\| \quad \Rightarrow \quad \|AB\| \leq \|A\| \|B\|, \quad (5.52)$$

und somit insbesondere $\|A^n\| \leq \|A\|^n$. Damit wird $\mathcal{B}(\mathcal{H})$ zu einer Algebra; mit der Definition des adjungierten Operators bildet er eine **C*-Algebra**, siehe z.B. [Bla93, Wer05].

Mit der Operatornorm ist die Konvergenz von Folgen erklärt und man kann **Funktionen** von Operatoren definieren wie z.B. Polynome und Potenzreihen. Eine für $|z| < R$ analytische Funktion $f(z)$ lässt sich durch eine Reihe $f(z) = \sum_{n=0}^{\infty} c_n z^n$ darstellen, z.B.

$$\frac{1}{1-z} = \sum_{n=0}^{\infty} z^n \quad (|z| < 1), \quad e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!} \quad (|z| < \infty). \quad (5.53)$$

Dann konvergiert auch die entsprechende Operatorreihe, hier die **Neumannsche Reihe** und die **Operator-Exponentialfunktion**:

$$(\mathbf{1} - A)^{-1} = \sum_{n=0}^{\infty} A^n \quad (\|A\| < 1), \quad e^A = \sum_{n=0}^{\infty} \frac{A^n}{n!} \quad (\|A\| < \infty). \quad (5.54)$$

Man überprüft leicht, dass das wieder beschränkte Operatoren aus $\mathcal{B}(\mathcal{H})$ sind, z.B. ist

$$\left\| \sum_{n=0}^{\infty} A^n \right\| \leq \sum_{n=0}^{\infty} \|A^n\| \leq \sum_{n=0}^{\infty} \|A\|^n = \frac{1}{1 - \|A\|} < \infty, \quad (5.55)$$

wobei wir im ersten Schritt die Dreiecksungleichung verwendet haben und im zweiten Schritt $\|A^n\| \leq \|A\|^n$. Analog zur geometrischen Reihe zeigt man, dass der Operator $(\mathbf{1} - A)$ invertierbar ist und sein Inverses die Neumannsche Reihe ergibt:

$$(\mathbf{1} - A) \sum_{n=0}^{\infty} A^n = \sum_{n=0}^{\infty} A^n (\mathbf{1} - A) = \lim_{N \rightarrow \infty} \left(\sum_{n=0}^N A^n - \sum_{n=1}^{N+1} A^n \right) = \mathbf{1} - \lim_{N \rightarrow \infty} A^{N+1} = \mathbf{1}. \quad (5.56)$$

Dabei haben wir $\|A^{N+1}\| \leq \|A\|^{N+1} \rightarrow 0$ für $N \rightarrow \infty$ benützt. Falls die Operatoren A und B kommutieren ($AB = BA$), gilt für die Exponentialfunktion auch $e^{A+B} = e^A e^B$.

Dualraum. Wir haben uns ja schon in Abschnitt 4.1 mit **Funktionalen** beschäftigt. Aus der Perspektive der Operatorsprache sind Funktionale nichts anderes als Operatoren, die Elemente eines Vektorraums auf den Körper der reellen oder komplexen Zahlen abbilden. Insbesondere sind **stetige lineare Funktionale** $F : \mathcal{H} \rightarrow \mathbb{C}$ auf einem Hilbertraum \mathcal{H} Spezialfälle von beschränkten Operatoren, womit wir alle Aussagen des letzten Abschnitts sofort übertragen können; z.B. lautet die Operatornorm

$$\|F\| = \sup_{\|u\| \neq 0} \frac{|F[u]|}{\|u\|} = \sup_{\|u\|=1} |F[u]| \quad \forall u \in \mathcal{H}. \quad (5.57)$$

Die Menge $\mathcal{H}^* := \mathcal{B}(\mathcal{H}, \mathbb{C})$ aller stetigen linearen Funktionale auf \mathcal{H} bildet den topologischen **Dualraum** von \mathcal{H} . In diesem Fall gibt es eine sehr nützliche Aussage:

SATZ (Darstellungssatz von Riesz und Fréchet): Zu jedem stetigen linearen Funktional F , das auf dem ganzen Hilbertraum \mathcal{H} definiert ist, existiert genau ein Vektor $f \in \mathcal{H}$, sodass sich F als Skalarprodukt

$$F[u] = \langle f, u \rangle \quad \forall u \in \mathcal{H} \quad (5.58)$$

darstellen lässt. Außerdem ist die Norm von f gleich der Operatornorm des Funktionals F , d.h. $\|f\| = \|F\|$.

Das besagt, dass \mathcal{H} isomorph zu seinem Dualraum \mathcal{H}^* ist, wobei der Isomorphismus in der Zuordnung $F \in \mathcal{H}^* \longleftrightarrow f \in \mathcal{H}$ besteht. Beispielsweise ist im Hilbertraum $\mathcal{H} = L^2(\mathbb{R})$ für ein festes $f \in \mathcal{H}$ die Zuordnung

$$F[u] := \langle f, u \rangle = \int_{-\infty}^{\infty} dx f(x)^* u(x) \quad (5.59)$$

ein stetiges lineares Funktional auf $L^2(\mathbb{R})$, und jedes stetige lineare Funktional auf $L^2(\mathbb{R})$ lässt sich auf diese Art schreiben. Darauf basiert letztlich auch die Bra-Ket-Formulierung, da jeder Ket-Vektor $|f\rangle \in \mathcal{H}$ isomorph zu einem Bra-Vektor $\langle f|$ aus dem Dualraum von \mathcal{H} ist.

Zum Beweis betrachten wir den Nullraum \mathcal{K}_F , d.h. die Menge aller Vektoren u , für die $F[u] = 0$ gilt. Für $\mathcal{K}_F = \mathcal{H}$ ist (5.58) trivial erfüllt, denn falls $F[u] = \langle f, u \rangle = 0 \forall u \in \mathcal{H}$, muss auch $f = 0$ sein. Falls \mathcal{K}_F ein Untervektorraum von \mathcal{H} ist, ist er wegen der Stetigkeit von F selbst wieder ein (abgeschlossener) Hilbertraum, da jede konvergente Folge in \mathcal{K}_F einen Grenzwert in \mathcal{K}_F hat. Damit können wir laut Projektionssatz (1.46) $\mathcal{H} = \mathcal{K}_F \oplus (\mathcal{K}_F)^\perp$ schreiben. Wählen wir jetzt irgendein $v \in (\mathcal{K}_F)^\perp$, dann liegt der Vektor $w = F[u]v - F[v]u$ wieder in \mathcal{K}_F , da wegen der Linearität $F[w] = F[u]F[v] - F[v]F[u] = 0$ gilt, d.h. w ist orthogonal zu v :

$$\langle v, w \rangle = F[u]\|v\|^2 - F[v]\langle v, u \rangle = 0 \quad \Rightarrow \quad F[u] = \left\langle \frac{F[v]^* v}{\|v\|^2}, u \right\rangle. \quad (5.60)$$

Das hat aber genau die gewünschte Form in (5.58) mit $f = F[v]^* v / \|v\|^2$. Um auch die Eindeutigkeit zu beweisen, nehmen wir an, es gäbe zwei f_1 und f_2 , die (5.58) für alle u erfüllen. Dann ist $\langle f_1 - f_2, u \rangle = 0 \forall u \in \mathcal{H}$, aber der einzige Vektor orthogonal zu allen $u \in \mathcal{H}$ ist der Nullvektor, also ist $f_1 = f_2$.

Zum Beweis von $\|F\| = \|f\|$ verwenden wir die Cauchy-Schwarz-Ungleichung: $|F[u]| = |\langle f, u \rangle| \leq \|f\| \|u\|$, also gilt $|F[u]| / \|u\| \leq \|f\|$ und somit bildet $\|f\|$ eine obere Schranke für die Operatornorm $\|F\|$. Wenn wir ein u finden, für das diese Schranke erreicht wird, gilt auch $\|F\| = \|f\|$. Ein solches u gibt es, nämlich $u = f$:

$$\frac{|F[u]|}{\|u\|} = \frac{|\langle f, f \rangle|}{\|f\|} = \frac{\|f\|^2}{\|f\|} = \|f\|. \quad (5.61)$$

Adjungierter Operator. Der Satz von Riesz und Fréchet erlaubt es uns, den früher für endlich-dimensionale Vektorräume besprochenen adjungierten Operator A^\dagger auch für unendlich-dimensionale Hilberträume zu definieren, und zwar über das Skalarprodukt:

DEFINITION: Seien \mathcal{H} und \mathcal{H}' Hilberträume und sei $A \in \mathcal{B}(\mathcal{H}, \mathcal{H}')$ ein linearer und beschränkter Operator. Dann ist der **adjungierte Operator** $A^\dagger \in \mathcal{B}(\mathcal{H}', \mathcal{H})$ ein linearer, beschränkter Operator, der eindeutig definiert ist durch

$$\langle A^\dagger v, u \rangle = \langle v, Au \rangle \quad \forall u \in \mathcal{H}, v \in \mathcal{H}'. \quad (5.62)$$

In der mathematischen Literatur wird der adjungierte Operator meist mit A^* bezeichnet; wir verwenden das Symbol A^\dagger , da es in der Physik gebräuchlicher ist und der Stern schon für die komplexe Konjugation vergeben ist. Genaugenommen sollten wir auch $\langle A^\dagger v, u \rangle_{\mathcal{H}} = \langle v, Au \rangle_{\mathcal{H}'}$ schreiben, aber zur übersichtlicheren Notation lassen wir die Indizes weg. Der adjungierte Operator A^\dagger erfüllt die folgenden Eigenschaften:

- Eindeutigkeit: dazu betrachten wir für ein gegebenes $v \in \mathcal{H}'$ das lineare Funktional $F[u] := \langle v, Au \rangle \forall u \in \mathcal{H}$. Weil A beschränkt ist, ist auch F ein beschränktes (und somit stetiges) lineares Funktional, denn aus der Cauchy-Schwarz-Ungleichung folgt

$$|F[u]| = |\langle v, Au \rangle| \leq \|v\| \|Au\| \leq \|A\| \|u\| \|v\|. \quad (5.63)$$

Nach dem Satz von Riesz und Fréchet gibt es dann ein eindeutiges Element $f \in \mathcal{H}$ sodass $F[u] = \langle f, u \rangle$. Dieses Element bezeichnen wir mit $f = A^\dagger v$.

- A^\dagger ist linear: $A^\dagger(v_1 + v_2) = A^\dagger v_1 + A^\dagger v_2$, $A^\dagger(\lambda v) = \lambda A^\dagger v$ für $v, v_1, v_2 \in \mathcal{H}'$ und $\lambda \in \mathbb{K}$. Das sieht man einfach durch Einsetzen:

$$\begin{aligned} \langle A^\dagger(v_1 + v_2), u \rangle &= \langle v_1 + v_2, Au \rangle = \langle v_1, Au \rangle + \langle v_2, Au \rangle \\ &= \langle A^\dagger v_1, u \rangle + \langle A^\dagger v_2, u \rangle = \langle A^\dagger v_1 + A^\dagger v_2, u \rangle \quad \forall u \in \mathcal{H}, \\ \langle A^\dagger(\lambda v), u \rangle &= \langle \lambda v, Au \rangle = \lambda^* \langle v, Au \rangle = \lambda^* \langle A^\dagger v, u \rangle = \langle \lambda A^\dagger v, u \rangle \quad \forall u \in \mathcal{H}. \end{aligned}$$

- $(A^\dagger)^\dagger = A$: Es gilt für alle $u \in \mathcal{H}$, $v \in \mathcal{H}'$

$$\langle v, (A^\dagger)^\dagger u \rangle = \langle (A^\dagger)^\dagger u, v \rangle^* = \langle u, A^\dagger v \rangle^* = \langle A^\dagger v, u \rangle = \langle v, Au \rangle. \quad (5.64)$$

- Auf die gleiche Weise beweist man folgende Relationen:

$$(A_1 + A_2)^\dagger = A_1^\dagger + A_2^\dagger, \quad (\lambda A)^\dagger = \lambda^* A^\dagger, \quad (AB)^\dagger = B^\dagger A^\dagger. \quad (5.65)$$

- A^\dagger ist ein beschränkter Operator mit $\|A^\dagger\| = \|A\|$: Es gilt

$$\|A^\dagger v\|^2 = \langle A^\dagger v, A^\dagger v \rangle = \langle v, AA^\dagger v \rangle \leq \|v\| \|AA^\dagger v\| \leq \|v\| \|A\| \|A^\dagger v\| \quad (5.66)$$

und daher $\|A^\dagger v\| \leq \|v\| \|A\|$ für alle $v \in \mathcal{H}'$, also $\|A^\dagger\| \leq \|A\|$. Umgekehrt gilt dann

$$\|Au\|^2 = \langle Au, Au \rangle = \langle u, A^\dagger Au \rangle \leq \|u\| \|A^\dagger\| \|Au\| \quad (5.67)$$

und somit $\|Au\| \leq \|u\| \|A^\dagger\|$ für alle $u \in \mathcal{H}$, also $\|A\| \leq \|A^\dagger\|$. Zusammengenommen ergibt sich daraus $\|A^\dagger\| = \|A\|$.

$$\blacksquare \|AA^\dagger\| = \|A^\dagger A\| = \|A\|^2: \quad (5.68)$$

Aus (5.52) folgt $\|AA^\dagger\| \leq \|A\| \|A^\dagger\| = \|A\|^2$ und $\|A^\dagger A\| \leq \|A^\dagger\| \|A\| = \|A\|^2$, sodass AA^\dagger und $A^\dagger A$ beides beschränkte Operatoren sind. Aus (5.66) und (5.67) folgt

$$\begin{aligned} \|A^\dagger v\|^2 &\leq \|v\|^2 \|AA^\dagger\| &\Rightarrow \|A^\dagger\|^2 &= \|A\|^2 \leq \|AA^\dagger\|, \\ \|Au\|^2 &\leq \|u\|^2 \|A^\dagger A\| &\Rightarrow \|A\|^2 &\leq \|A^\dagger A\|, \end{aligned} \quad (5.69)$$

womit sich in Summe die Aussage ergibt.

$$\blacksquare \mathcal{K}_A = (\mathcal{R}_{A^\dagger})^\perp \text{ und } \mathcal{K}_{A^\dagger} = (\mathcal{R}_A)^\perp: \quad (5.70)$$

Wir beweisen zuerst die Inklusion $\mathcal{K}_A \subset (\mathcal{R}_{A^\dagger})^\perp$: Sei $u \in \mathcal{K}_A$, sodass $Au = 0$. Dann gilt auch $\langle A^\dagger v, u \rangle = \langle v, Au \rangle = 0$ und somit $A^\dagger v \perp u$ für alle $v \in \mathcal{H}'$, also $u \in (\mathcal{R}_{A^\dagger})^\perp$. Für die umgekehrte Inklusion sei $u \in (\mathcal{R}_{A^\dagger})^\perp$, d.h. $0 = \langle u, A^\dagger v \rangle = \langle Au, v \rangle$ und somit $Au \perp v$ für alle $v \in \mathcal{H}'$. Dann gilt $Au \in (\mathcal{H}')^\perp = \{0\}$, denn das orthogonale Komplement eines gesamten Hilbertraums ist der Nullvektor, also $Au = 0$ und somit $u \in \mathcal{K}_A$. Aus $\mathcal{K}_A \subset (\mathcal{R}_{A^\dagger})^\perp$ und $(\mathcal{R}_{A^\dagger})^\perp \subset \mathcal{K}_A$ folgt die Mengengleichheit $\mathcal{K}_A = (\mathcal{R}_{A^\dagger})^\perp$. Zum Beweis der zweiten Beziehung $\mathcal{K}_{A^\dagger} = (\mathcal{R}_A)^\perp$ müssen wir in der ersten Relation nur A durch A^\dagger ersetzen.

Mit diesen Betrachtungen können wir jetzt auch den Begriff des selbstadjungierten Operators für beschränkte Operatoren definieren (den allgemeinen Fall unbeschränkter Operatoren werden wir uns in Abschnitt 5.3 ansehen):

DEFINITION: Sei \mathcal{H} ein Hilbertraum. Ein linearer, beschränkter Operator $A \in \mathcal{B}(\mathcal{H})$ heißt **selbstadjungiert**, wenn $A^\dagger = A$. Mit Gl. (5.62) ist das äquivalent zu:

$$\langle Av, u \rangle = \langle v, Au \rangle \quad \forall u, v \in \mathcal{H}. \quad (5.71)$$

Ein oft einfacheres Kriterium der Selbstadjungiertheit ist folgendes: Sei \mathcal{H} ein Hilbertraum über \mathbb{C} und $A \in \mathcal{B}(\mathcal{H})$. Dann ist A selbstadjungiert genau dann, wenn der **Erwartungswert** $\langle u, Au \rangle$ des Operators A bezüglich des Vektors u für alle $u \in \mathcal{H}$ reell ist. In der Quantenmechanik ist der Erwartungswert der erwartete Messwert der physikalischen Observablen A im Zustand u .

Der Beweis in eine Richtung ist klar: falls A selbstadjungiert ist, dann gilt $\langle u, Au \rangle = \langle Au, u \rangle = \langle u, Au \rangle^*$ für alle $u \in \mathcal{H}$. Der Beweis in die andere Richtung verläuft so: falls $\langle u, Au \rangle \in \mathbb{R} \quad \forall u \in \mathcal{H}$, dann gilt auch

$$\langle u + v, A(u + v) \rangle = \langle A(u + v), u + v \rangle, \quad (5.72)$$

$$\langle u - iv, A(u - iv) \rangle = \langle A(u - iv), u - iv \rangle \quad (5.73)$$

für alle $u + v$ und $u - iv \in \mathcal{H}$. Wenn wir diese Gleichungen termweise auswerten, ergibt sich

$$\begin{aligned} \langle u, Au \rangle + \langle v, Av \rangle + \langle v, Au \rangle + \langle u, Av \rangle &= \langle Au, u \rangle + \langle Av, v \rangle + \langle Au, v \rangle + \langle Av, u \rangle, \\ \langle u, Au \rangle + \langle v, Av \rangle + i\langle v, Au \rangle - i\langle u, Av \rangle &= \langle Au, u \rangle + \langle Av, v \rangle - i\langle Au, v \rangle + i\langle Av, u \rangle. \end{aligned} \quad (5.74)$$

Verwenden wir wieder $\langle u, Au \rangle = \langle Au, u \rangle \quad \forall u \in \mathcal{H}$, folgt

$$\begin{aligned} \langle v, Au \rangle + \langle u, Av \rangle &= \langle Au, v \rangle + \langle Av, u \rangle, \\ \langle v, Au \rangle - \langle u, Av \rangle &= -\langle Au, v \rangle + \langle Av, u \rangle \end{aligned} \quad (5.75)$$

und in Summe $\langle Av, u \rangle = \langle v, Au \rangle \quad \forall u, v \in \mathcal{H}$.

Für einen selbstadjungierten Operator lässt sich die Operatornorm auch mit Hilfe des Erwartungswerts bestimmen:

$$\|A\| = \sup_{\|u\| \leq 1} |\langle u, Au \rangle|. \quad (5.76)$$

Wie sieht der adjungierte Operator A^\dagger in der Matrixdarstellung aus? Mit der ONB $\{\varphi_i\}$ schreiben wir

$$(A^\dagger)_{ij} = \langle \varphi_i, A^\dagger \varphi_j \rangle = \langle A \varphi_i, \varphi_j \rangle = \langle \varphi_j, A \varphi_i \rangle^* = A_{ji}^* \Rightarrow \mathbf{A}^\dagger = (\mathbf{A}^*)^T, \quad (5.77)$$

d.h. wie erwartet ist der adjungierte Operator die adjungierte Matrix. Der Erwartungswert $\langle u, Au \rangle$ lautet in der Matrixdarstellung

$$\langle u, Au \rangle = \sum_{ij} \alpha_i^* \alpha_j \langle \varphi_i, A \varphi_j \rangle = \boldsymbol{\alpha}^\dagger \mathbf{A} \boldsymbol{\alpha}. \quad (5.78)$$

Damit sind die Rechts- und Linksshift-Operatoren aus (5.34–5.35) gerade adjungiert zueinander, denn die Operation $\mathbf{A} \rightarrow \mathbf{A}^\dagger$ entspricht ja der Spiegelung an der Diagonalen und komplexen Konjugation. Für den Hilbert-Schmidt-Integraloperator (5.41) lautet die Bestimmungsgleichung des adjungierten Operators

$$\langle g, Au \rangle = \int_{\Omega} dx g(x)^* \int_{\Omega} dy K(x, y) u(y) = \int_{\Omega} dy \left[\int_{\Omega} dx K(x, y)^* g(x) \right]^* u(y) = \langle A^\dagger g, u \rangle,$$

und somit ist der adjungierte Operator gegeben durch

$$A^\dagger g(x) = \int_{\Omega} dy K(y, x)^* g(y). \quad (5.79)$$

Der Operator ist selbstadjungiert, wenn $K(x, y) = K(y, x)^*$.

SATZ (Hellinger-Toeplitz): Sei $A: \mathcal{H} \rightarrow \mathcal{H}$ ein linearer Operator, der auf dem ganzen Hilbertraum definiert ist. Falls A selbstadjungiert ist ($A = A^\dagger$), dann ist er auch beschränkt, d.h. $A \in \mathcal{B}(\mathcal{H})$.

Im Umkehrschluss bedeutet das, dass unbeschränkte selbstadjungierte Operatoren nicht auf dem ganzen Hilbertraum definiert werden können (darauf werden wir in Abschnitt 5.3 zurückkommen). Leider kommt man in der Quantenmechanik ohne unbeschränkte Operatoren nicht aus, denn die dort als Observablen auftretenden Operatoren wie der Orts-, Impuls- und Hamiltonoperator sind unbeschränkte selbstadjungierte Operatoren. Insbesondere erfüllen der Ortsoperator X und Impulsoperator P die Vertauschungsrelationen $[X, P] = i\mathbb{1}$. Der folgende Satz zeigt, dass diese Relation nicht durch lineare beschränkte Operatoren erfüllt werden kann, sodass mindestens einer der Operatoren X und P unbeschränkt sein muss:

SATZ (Wielandt): Seien X und P lineare Operatoren in einem Hilbertraum \mathcal{H} . Falls X und P beschränkt sind, erfüllen sie die Vertauschungsrelation $[X, P] = i\mathbb{1}$ nicht.

Der Beweis wird durch Widerspruch geführt. Nehmen wir an, X und P seien beschränkte Operatoren, die die Vertauschungsrelation erfüllen. Dann gilt $[P^{n+1}, X] = P^n[P, X] + [P^n, X]P$, wie man sofort sieht, wenn man die Kommutatoren ausschreibt, und daher

$$[P^{n+1}, X] = -iP^n + [P^n, X]P = -i(n+1)P^n, \quad (5.80)$$

wobei der letzte Schritt durch Induktion erfolgt ist. Daraus erhalten wir

$$(n+1)\|P^n\| = \|[P^{n+1}, X]\| \leq \|P^{n+1}X\| + \|X P^{n+1}\| \leq 2\|X\|\|P^n\| = c\|P^n\|. \quad (5.81)$$

Wenn X und P beschränkt sind, ist c endlich. Da c nicht von n abhängt, muss für hinreichend großes n somit $P^n = 0$ sein. Dann folgt aber aus (5.80): $P^n = 0 \Rightarrow P^{n-1} = 0 \Rightarrow P^{n-2} = 0 \Rightarrow \dots \Rightarrow P = 0$, und somit $[X, P] = 0$.

Projektionsoperatoren. Eine wichtige Klasse von beschränkten Operatoren sind orthogonale Projektoren, die auf abgeschlossene Unterräume eines Hilbertraums projizieren. Betrachten wir einen solchen abgeschlossenen Untervektorraum $\mathcal{M} \subset \mathcal{H}$. Der Projektionssatz (1.46) besagt dann, dass wir den Hilbertraum in $\mathcal{H} = \mathcal{M} \oplus \mathcal{M}^\perp$ zerlegen können. In anderen Worten, jedes Element $u \in \mathcal{H}$ kann eindeutig als $u = v + w$ geschrieben werden, wobei $v \in \mathcal{M}$, $w \in \mathcal{M}^\perp$ und $v \perp w$. Die orthogonale Projektion bzw. der **Projektor** auf \mathcal{M} ist dann der lineare Operator $P: \mathcal{H} \rightarrow \mathcal{H}$ mit $Pu = v$. Allein aus dieser Definition lässt sich bereits eine Reihe weiterer Aussagen treffen:

- Der Operator P ist beschränkt, d.h. es ist $P \in \mathcal{B}(\mathcal{H})$. Es ist ja $u = v + w$ und $v \perp w$, daher $\|u\|^2 = \|v\|^2 + \|w\|^2$ und somit $\|Pu\| = \|v\| \leq \|u\|$, also $\|P\| \leq 1$. Die obere Schranke wird für $u \in \mathcal{M}$ erreicht, denn in diesem Fall ist $\|Pu\| = \|u\|$. Solange P nicht der Nulloperator ist, gilt also $\|P\| = 1$.
- Wegen $Pu = v$ und $Pw = 0$ gilt $\mathcal{R}_P = \mathcal{M}$ und $\mathcal{K}_P = \mathcal{M}^\perp$.
- Aus $P^2 u = Pv = v$ folgt die **Idempotenz**: $P^2 = P$.
- Der Projektor ist selbstadjungiert: $P = P^\dagger$, denn es gilt für alle $u_i = v_i + w_i$ mit $v_i \in \mathcal{M}$ und $w_i \in \mathcal{M}^\perp$:

$$\langle u_2, Pu_1 \rangle = \langle v_2 + w_2, v_1 \rangle = \langle v_2, v_1 \rangle = \langle v_2, v_1 + w_1 \rangle = \langle Pu_2, u_1 \rangle. \quad (5.82)$$

- Es gilt $\langle u, Pu \rangle \geq 0$ für alle $u \in \mathcal{H}$, denn

$$\langle u, Pu \rangle = \langle v + w, P(v + w) \rangle = \langle v + w, v \rangle = \langle v, v \rangle = \|v\|^2 \geq 0. \quad (5.83)$$

- Wegen $(\mathbb{1} - P)u = u - v = w \in \mathcal{M}^\perp$ ist $(\mathbb{1} - P)$ der orthogonale Projektor auf \mathcal{M}^\perp , und es gilt $P(\mathbb{1} - P) = P - P^2 = P - P = 0$. Daraus folgt

$$\mathcal{R}_{\mathbb{1}-P} = \mathcal{M}^\perp = \mathcal{K}_P, \quad \mathcal{K}_{\mathbb{1}-P} = \mathcal{M} = \mathcal{R}_P. \quad (5.84)$$

- Wegen $Pv = v$ und $Pw = 0$ wirkt ein Projektor P wie die Identität auf \mathcal{R}_P und wie der Nulloperator auf $\mathcal{R}_{\mathbb{1}-P}$:

$$P = \begin{cases} \mathbb{1} & \text{auf } \mathcal{R}_P, \\ 0 & \text{auf } \mathcal{R}_{\mathbb{1}-P}. \end{cases} \quad (5.85)$$

SATZ: Ein **Projektor** $P \in \mathcal{B}(\mathcal{H})$ ist durch folgende äquivalente Aussagen charakterisiert:

- (a) P bildet jedes $u \in \mathcal{H}$ auf einen abgeschlossenen Teilraum \mathcal{M} von \mathcal{H} ab.
- (b) $P^2 = P$ und $P = P^\dagger$.
- (c) $P^2 = P$ und $\|P\| = 1$.
- (d) $P^2 = P$ und $\mathcal{R}_P = \mathcal{K}_P^\perp$.
- (e) $P^2 = P$ und $\langle u, Pu \rangle \geq 0 \quad \forall u \in \mathcal{H}$.

Die Beweise für $(a \rightarrow b, c, d, e)$ haben wir schon geliefert. Man kann auch alle anderen Richtungen beweisen, z.B. folgt $(b \rightarrow c)$ aus (5.68): $\|P\|^2 = \|PP^\dagger\| = \|P^2\| = \|P\| \Rightarrow \|P\| = 1$; $(b \rightarrow d)$ ist eine unmittelbare Konsequenz aus (5.70): $\mathcal{K}_P = (\mathcal{R}_{P^\dagger})^\perp = (\mathcal{R}_P)^\perp$, und $(b \rightarrow e)$ folgt aus $\langle u, Pu \rangle = \langle u, P^2 u \rangle = \langle Pu, Pu \rangle = \|Pu\|^2 \geq 0 \quad \forall u \in \mathcal{H}$.

Einige Beispiele für Projektoren:

- Der Null- und Einsoperator sind (triviale) Projektoren.
- Der Multiplikationsoperator $P : L^2(\mathbb{R}) \rightarrow L^2(\mathbb{R})$ mit $Pu(x) = I_\Omega(x)u(x)$, wobei I_Ω die charakteristische Funktion aus (4.22) mit $\Omega \subset \mathbb{R}$ bezeichnet, ist ein Projektor. Es gilt $P^2u(x) = I_\Omega(x)^2u(x) = I_\Omega(x)u(x)$ und $P^\dagger = P$, da die Funktion I_Ω reell ist. P projiziert $L^2(\mathbb{R})$ auf den Unterraum aller Funktionen in $L^2(\mathbb{R})$ mit Träger in Ω .
- Für einen festen Einheitsvektor φ mit $\|\varphi\| = 1$ projiziert

$$P_\varphi u = \varphi \langle \varphi, u \rangle \quad \Leftrightarrow \quad P_\varphi = |\varphi\rangle\langle\varphi| \quad (5.86)$$

aus beliebigen Vektoren $u \in \mathcal{H}$ den longitudinalen Anteil parallel zu φ heraus, d.h. er projiziert auf den Unterraum \mathcal{M} , der durch den Einheitsvektor φ aufgespannt wird. Die Projekteureigenschaften sind erfüllt, denn es gilt

$$\begin{aligned} P_\varphi^2 u &= \varphi \langle \varphi, \varphi \rangle \langle \varphi, u \rangle = \varphi \langle \varphi, u \rangle, \\ \langle u, P_\varphi u \rangle &= \langle u, \varphi \rangle \langle \varphi, u \rangle = |\langle \varphi, u \rangle|^2 \geq 0. \end{aligned} \quad (5.87)$$

Die transversale Komponente $(1 - P_\varphi)u$ steht dann normal auf φ :

$$\langle \varphi, P_\varphi u \rangle = \langle \varphi, u \rangle, \quad \langle \varphi, (1 - P_\varphi)u \rangle = 0. \quad (5.88)$$

Die Matrixelemente in einer ONB $\{\varphi_j\}$ lauten $P_{ij} = \langle \varphi_i, P_\varphi \varphi_j \rangle = \langle \varphi_i, \varphi \rangle \langle \varphi, \varphi_j \rangle$. Ist φ eines der Basiselemente, dann ist \mathbf{P} eine Matrix, bei der alle Elemente außer dem Diagonalelement in der entsprechenden Zeile und Spalte Null sind. Ist φ eine Kombination aus Basiselementen, dann gibt es entsprechend viele nichtverschwindende Einträge.

- Im Raum \mathbb{R}^2 sind beispielsweise die Matrizen

$$\mathbf{P}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{und} \quad \mathbf{P}_2 = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \quad (5.89)$$

Projektoren, denn sie erfüllen $\mathbf{P}_i^2 = \mathbf{P}_i$ und $\mathbf{P}_i^\dagger = \mathbf{P}_i$. Die erste Matrix projiziert immer die obere Komponente eines Vektors $\mathbf{x} = (x_1, x_2)$ heraus: $\mathbf{P}_1 \mathbf{x} = (x_1, 0)$. Die zweite Matrix \mathbf{P}_2 projiziert ihn in Diagonalrichtung: $\mathbf{P}_2 \mathbf{x} = \frac{1}{2}(x_1 + x_2)(1, 1)$.

- Der Operator $P : \ell^2 \rightarrow \ell^2$ mit $P(x_1, \dots, x_{k-1}, x_k, x_{k+1}, \dots, x_n) = (0, \dots, 0, x_k, 0, \dots, 0)$ ist ein Projektor: Es gilt $P^2 = P$, da $P(0, \dots, 0, x_k, 0, \dots, 0) = (0, \dots, 0, x_k, 0, \dots, 0)$, und $P = P^\dagger$, weil $\langle x, Px \rangle = |x_k|^2 \geq 0$.
- Für ein ONS $\{\varphi_i\}$, d.h. eine Menge zueinander orthogonaler Einheitsvektoren, mit zugehörigen Projektoren P_{φ_i} ist

$$P = \sum_i P_{\varphi_i} = \sum_i |\varphi_i\rangle\langle\varphi_i| \quad (5.90)$$

wieder ein Projektor, der auf den von den φ_i aufgespannten Unterraum projiziert:

$$\begin{aligned} P^2 &= \sum_{ij} |\varphi_i\rangle\langle\varphi_i, \varphi_j\rangle\langle\varphi_j| = \sum_i |\varphi_i\rangle\langle\varphi_i| = P, \\ \langle u, Pu \rangle &= \sum_i \langle u, \varphi_i \rangle \langle \varphi_i, u \rangle = \sum_i |\langle \varphi_i, u \rangle|^2 \geq 0. \end{aligned} \quad (5.91)$$

Unitäre Operatoren. Eine weitere wichtige Klasse von linearen, beschränkten Operatoren sind die unitären Operatoren, die als Isomorphismen sowohl die lineare als auch die metrische Struktur eines Hilbertraums erhalten. Unitäre Operatoren lassen die Länge eines Vektors (seine Norm) und den Winkel zwischen zwei Vektoren (d.h. ihr Skalarprodukt) invariant und entsprechen somit einer Drehung im Hilbertraum.

DEFINITION: Ein beschränkter linearer Operator $U \in \mathcal{B}(\mathcal{H})$ auf einem Hilbertraum \mathcal{H} heißt **unitär** genau dann, wenn

- U **isometrisch** ist, d.h. $\|Ux\| = \|x\| \quad \forall x \in \mathcal{H}$,
- U **surjektiv** ist, d.h. $\mathcal{R}_U = \mathcal{H}$.

Die Aussage, dass U isometrisch ist, ist gleichbedeutend mit $U^\dagger U = \mathbf{1}$, denn es gilt

$$\|Ux\| = \|x\| \quad \Rightarrow \quad \langle Ux, Ux \rangle = \langle x, U^\dagger Ux \rangle = \langle x, x \rangle \quad \forall x \in \mathcal{H}, \quad (5.92)$$

was sich mit Hilfe der Polarisationsidentität (1.29) zu $\langle x, U^\dagger Uy \rangle = \langle x, y \rangle \quad \forall x, y \in \mathcal{H}$ erweitern lässt. Eine Isometrie ist trivialerweise injektiv, denn es gilt $Ux = 0 \Leftrightarrow x = 0$, d.h. $\mathcal{K}_U = \{0\}$. Damit besitzt sie auf ihrem Wertebereich einen inversen Operator $U^{-1}: \mathcal{R}_U \rightarrow \mathcal{H}$. Aus $U^\dagger U = \mathbf{1}$ folgt dann $U^{-1} = U^\dagger|_{\mathcal{R}_U}$. Welche Bedeutung hat dann UU^\dagger ? Es gilt für alle $x \in (\mathcal{R}_U)^\perp$ und $y \in \mathcal{H}$ (und somit $Uy \in \mathcal{R}_U$):

$$\langle U^\dagger x, y \rangle = \langle x, Uy \rangle = 0 \quad \forall y \in \mathcal{H} \quad \Rightarrow \quad U^\dagger x = 0 \quad \forall x \in (\mathcal{R}_U)^\perp. \quad (5.93)$$

Zusammengefasst ergibt das:

$$U^\dagger = \begin{cases} U^{-1} & \text{auf } \mathcal{R}_U, \\ 0 & \text{auf } (\mathcal{R}_U)^\perp, \end{cases} \quad \Rightarrow \quad UU^\dagger = \begin{cases} 1 & \text{auf } \mathcal{R}_U, \\ 0 & \text{auf } (\mathcal{R}_U)^\perp. \end{cases} \quad (5.94)$$

Damit ist UU^\dagger der Projektor auf \mathcal{R}_U ! Erst wenn U auch surjektiv (und somit unitär) ist, gilt $UU^\dagger = U^\dagger U = \mathbf{1}$ und infolgedessen $U^{-1} = U^\dagger$.

In einem endlichdimensionalen Hilbertraum $\mathcal{H} = \mathbb{R}^n$ oder \mathbb{C}^n ist jede Isometrie auch unitär: aus $U^\dagger U = \mathbf{1}$ folgt $\det(U^\dagger U) = |\det U|^2 = 1$, daher ist U invertierbar und es folgt $\mathcal{R}_U = \mathcal{H}$. In unendlichdimensionalen Hilberträumen gibt es allerdings Isometrien, die nicht unitär sind. Ein Beispiel ist der Shift-Operator in (5.33): es gilt zwar $\|A\varphi_i\| = \|\varphi_{i+1}\| = 1$, doch A ist nicht surjektiv, da $\varphi_1 \notin \mathcal{R}_A$.

SATZ: Für einen Operator $U \in \mathcal{B}(\mathcal{H})$ sind folgende Aussagen äquivalent:

- U ist unitär,
- $U^\dagger U = UU^\dagger = \mathbf{1} \quad \Leftrightarrow \quad U^{-1} = U^\dagger$,
- $\mathcal{R}_U = \mathcal{H}$ und $\langle Ux, Uy \rangle = \langle x, y \rangle \quad \forall x, y \in \mathcal{H}$.

Der **Translationsoperator** $U_a: L^2(\mathbb{R}) \rightarrow L^2(\mathbb{R})$ mit $U_a f(x) = f(x - a)$ ist ein unitärer Operator. Er ist isometrisch, denn es gilt

$$\langle U_a f, U_a g \rangle = \int_{-\infty}^{\infty} dx f(x - a)^* g(x - a) = \int_{-\infty}^{\infty} dx f(x)^* g(x) = \langle f, g \rangle \quad \forall f, g \in L^2(\mathbb{R}). \quad (5.95)$$

Aus $U_0 = \mathbf{1}$ und $U_a U_b = U_b U_a = U_{a+b}$ folgt außerdem $U_a^{-1} = U_{-a}$ für alle $a \in \mathbb{R}$, daher ist $\mathcal{R}_{U_a} = L^2(\mathbb{R})$ und der Operator U_a ist unitär.

Auch die früher besprochene **Fouriertransformation** ist ein unitärer Operator. Aus der Formel von Plancherel (3.43) sieht man, dass die Fouriertransformation $\mathcal{F} : S(\mathbb{R}) \rightarrow S(\mathbb{R})$ ein isometrischer Isomorphismus auf dem Schwartz-Raum ist. Da $S(\mathbb{R})$ dicht im $L^2(\mathbb{R})$ liegt, lässt sich dieser Isomorphismus auf ganz $L^2(\mathbb{R})$ erweitern, sodass die Fouriertransformation zu einem unitären Operator auf $L^2(\mathbb{R})$ wird.

Der **Basiswechsel** $\psi_i = U\varphi_i$ zwischen zwei Orthonormalbasen $\{\varphi_i\}$ und $\{\psi_i\}$ wird durch einen unitären Operator vermittelt, da ja $\|\psi_i\| = \|U\varphi_i\| = \|\varphi_i\| = 1$ gilt. Ein Vektor $f \in \mathcal{H}$ hat in verschiedenen Basen unterschiedliche Komponenten,

$$f = \sum_i \alpha_i \varphi_i = \sum_k \beta_k \psi_k, \quad (5.96)$$

sodass der Übergang von $\{\varphi_i\}$ zu $\{\psi_i\}$ einer Drehung $\beta = \mathbf{U}^\dagger \alpha$ entspricht:

$$\beta_k = \langle \psi_k, f \rangle = \sum_i \alpha_i \langle \psi_k, \varphi_i \rangle = \sum_i \alpha_i \langle U\varphi_k, \varphi_i \rangle = \sum_i \alpha_i \langle \varphi_k, U^\dagger \varphi_i \rangle = \sum_i (U^\dagger)_{ki} \alpha_i. \quad (5.97)$$

Das entspricht einer Transformation $\mathbf{A}' = \mathbf{U}^\dagger \mathbf{A} \mathbf{U}$ für die Matrixdarstellung eines Operators, wobei \mathbf{A} die Matrix in der ONB $\{\varphi_i\}$ bezeichnet und \mathbf{A}' jene in $\{\psi_i\}$:

$$A'_{ij} = \langle \psi_i, A\psi_j \rangle = \langle U\varphi_i, AU\varphi_j \rangle = \langle \varphi_i, U^\dagger A U \varphi_j \rangle = (U^\dagger A U)_{ij}. \quad (5.98)$$

Betrachten wir als Beispiel den \mathbb{R}^2 mit den beiden Orthonormalbasen

$$\varphi_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \varphi_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \psi_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \psi_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}. \quad (5.99)$$

Die Transformationsmatrix \mathbf{U} zwischen den beiden Systemen ist unitär:

$$U_{ij} = \langle \varphi_i, U\varphi_j \rangle = \langle \varphi_i, \psi_j \rangle \Rightarrow \mathbf{U} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad \mathbf{U}\mathbf{U}^\dagger = \mathbf{U}^\dagger\mathbf{U} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \quad (5.100)$$

Um z.B. den Vektor $(1, -1)$ vom $\{\varphi_i\}$ ins $\{\psi_i\}$ -System zu transformieren, müssen wir ihn mit \mathbf{U}^\dagger multiplizieren: $\mathbf{U}^\dagger (1, -1) = (0, \sqrt{2}) = \sqrt{2} \psi_2$. Die Länge des Vektors hat sich nicht verändert.

Wie wir oben am Beispiel des Translationsoperators gesehen haben, bilden die unitären Operatoren eines Hilbertraums eine **Gruppe**: die Hintereinanderausführung zweier unitärer Operatoren ist wieder ein unitärer Operator, es gilt ein Assoziativgesetz, und es gibt ein eindeutiges Einselement sowie ein eindeutiges inverses Element. Unitäre Gruppen spielen in vielen Bereichen der Physik eine wichtige Rolle:

- Die Gruppe $SO(3)$ besteht aus orthogonalen 3×3 -Matrizen $\mathbf{A}^T = \mathbf{A}^{-1}$ mit $\det \mathbf{A} = 1$ und bewirkt Raumdrehungen im \mathbb{R}^3 .
- Die Gruppe $U(1)$ ist die Menge der komplexen Zahlen $e^{i\theta}$ mit $\theta \in \mathbb{R}$. Sie ist die Symmetriegruppe der Quantenmechanik bzw. der Quantenelektrodynamik.
- Die Gruppe $SU(n)$ besteht aus unitären $n \times n$ -Matrizen $\mathbf{U}^\dagger = \mathbf{U}^{-1}$ mit $\det \mathbf{U} = 1$ und ist eine der wichtigsten Gruppen in der Teilchenphysik. Die Gruppe $SU(2)$ beschreibt den Spin und Isospin; die $SU(3)$ beschreibt die Farbe der Quarks in der Quantenchromodynamik sowie den Flavor für up-, down- und strange Quarks; $SU(2) \times SU(2)$ hängt eng mit der Lorentzgruppe der speziellen Relativitätstheorie zusammen, usw. Das Standardmodell der Teilchenphysik wird durch die Symmetriegruppe $U(1) \times SU(2) \times SU(3)$ beschrieben.

5.3 Unbeschränkte Operatoren

Im Fall eines unendlich-dimensionalen Hilbertraums sind lineare Operatoren zwar mit der linearen, aber nicht notwendigerweise mit der topologischen Struktur verträglich. Typische in der Quantenmechanik auftretende Operatoren wie der Ortsoperator, Impulsoperator und Hamiltonoperator sind unbeschränkt und damit unstetig. Wir haben schon früher gesehen, dass bei unbeschränkten Operatoren Probleme mit dem Definitionsbereich auftreten und man sie nicht auf dem ganzen Hilbertraum definieren kann. Die Definitionsbereiche sind dann meist auch keine vollständigen Räume mehr. Sehen wir uns zunächst einige Beispiele für unbeschränkte Operatoren an:

- Betrachten wir den **Multiplikationsoperator** $A: L^2(\Omega) \rightarrow L^2(\Omega)$, $Au(x) = h(x)u(x)$ auf dem Raum $L^2(\Omega)$ mit $\Omega \subset \mathbb{R}^n$, wobei $h: \Omega \rightarrow \mathbb{C}$ eine komplexwertige Funktion ist. Da für beliebige Funktionen h das Produkt hu nicht mehr in $L^2(\Omega)$ sein wird (setze z.B. $h(x) = 1/x$ wie in (5.21)), müssen wir den Definitionsbereich des Operators auf jene Funktionen u einschränken, für die das gilt:

$$\mathcal{D}_A = \{u \in L^2(\Omega) \mid hu \in L^2(\Omega)\}. \quad (5.101)$$

Falls h eine beschränkte Funktion ist, d.h. falls es ein $c \geq 0$ gibt sodass $|h(x)| \leq c$ für alle $x \in \Omega$, ist auch der Operator A beschränkt:

$$\|y\|^2 = \|Au\|^2 = \int_{\Omega} dx |h(x)u(x)|^2 \leq c^2 \int_{\Omega} dx |u(x)|^2 = c^2 \|u\|^2, \quad (5.102)$$

und der Definitionsbereich ist der gesamte Hilbertraum: $\mathcal{D}_A = L^2(\Omega)$.

- Bezeichnen wir ein allgemeines Element des Folgenraums ℓ^2 mit $\mathbf{x} = (x_n)_{n=0}^{\infty}$, dann bildet der Multiplikationsoperator $A: \ell^2 \rightarrow \ell^2$ mit $A\mathbf{x} = \mathbf{y}$ jedes \mathbf{x} auf $\mathbf{y} = (a_n x_n)_{n=0}^{\infty}$ ab, wobei $(a_n)_{n=0}^{\infty} \subset \mathbb{C}$ eine beliebige Folge ist. Im Allgemeinen muss \mathbf{y} nicht mehr im ℓ^2 liegen, deswegen schränken wir den Definitionsbereich ein: $\mathcal{D}_A = \{\mathbf{x} \in \ell^2 \mid \mathbf{y} \in \ell^2\}$. Ist die Folge $(a_n)_{n=0}^{\infty} \subset \mathbb{C}$ beschränkt, d.h. $|a_n| \leq c \forall n \in \mathbb{N}$, dann ist auch $\mathbf{y} \in \ell^2$,

$$\|\mathbf{y}\|^2 = \|A\mathbf{x}\|^2 = \sum_{n=1}^{\infty} |a_n x_n|^2 \leq c^2 \sum_{n=1}^{\infty} |x_n|^2 = c^2 \|\mathbf{x}\|^2, \quad (5.103)$$

und der Definitionsbereich ist der gesamte Hilbertraum: $\mathcal{D}_A = \ell^2$.

- Der **Impulsoperator** in der Quantenmechanik hat die Operatorvorschrift $P = -id/dx$, d.h. er ist proportional zum Ableitungsoperator. Wir haben schon in (5.23) gesehen, dass dieser Operator auf dem Hilbertraum L^2 unbeschränkt ist.
- In drei Raumdimensionen ist der **Schrödingeroperator** $H: L^2(\mathbb{R}^3) \rightarrow L^2(\mathbb{R}^3)$ durch die Operatorvorschrift $H = -\Delta + V$ gegeben, wobei Δ den Laplace-Operator und $V: \mathbb{R}^3 \rightarrow \mathbb{R}$ das Potential bezeichnet. In der Quantenmechanik ist H der Hamiltonoperator, der die Energie eines quantenmechanischen Teilchens beschreibt. Auch dieser Operator ist unbeschränkt. Ein möglicher Definitionsbereich wäre

$$\mathcal{D}_H = \{u \in L^2(\mathbb{R}^3) \mid u \in C^2(\mathbb{R}^3) \text{ und } (-\Delta + V)u \in L^2(\mathbb{R}^3)\}. \quad (5.104)$$

Abgeschlossene Operatoren. Das Problem mit unbeschränkten Operatoren ist der Verlust der Stetigkeit (5.26). Damit gehen die Vertauschbarkeit von Limes und Operatoranwendung und somit viele aus endlichen-dimensionalen Räumen bekannte Rechenregeln verloren. Man kann aber den Spielraum in der Wahl der Definitionsbereiche ausnützen, um die Stetigkeit durch das Konzept der **Abgeschlossenheit** zu ersetzen, das dann wieder zu ähnlich angenehmen Eigenschaften wie bei beschränkten Operatoren führt.

Wir werden im Folgenden immer voraussetzen, dass der Definitionsbereich \mathcal{D}_A dicht im Hilbertraum \mathcal{H} liegt, d.h. $\overline{\mathcal{D}_A} = \mathcal{H}$. Man sagt dann auch: der Operator A ist **dicht definiert**. Das impliziert insbesondere, dass das orthogonale Komplement von \mathcal{D}_A der Nullvektor ist:

$$\overline{\mathcal{D}_A} = \mathcal{H} \quad \Leftrightarrow \quad (\mathcal{D}_A)^\perp = \{0\}. \quad (5.105)$$

Diese Forderung legt den Definitionsbereich des Operators aber noch nicht eindeutig fest. Wie wir in Abschnitt 5.4 sehen werden, hat die Wahl eines passenden Definitionsbereiches auch physikalische Auswirkungen und kann zu verschiedenen Spektren führen.

DEFINITION: Ein linearer Operator B auf \mathcal{D}_B heißt **Erweiterung** bzw. **Fortsetzung** eines anderen Operators A auf \mathcal{D}_A , wenn $\mathcal{D}_A \subset \mathcal{D}_B$ und $Bu = Au$ für alle $u \in \mathcal{D}_A$. Man schreibt dann $A \subset B$. Umgekehrt nennt man $A = B|_{\mathcal{D}_A}$ eine **Einschränkung** von B auf \mathcal{D}_A .

Die Freiheit in der Wahl des Definitionsbereichs lässt sich an folgendem Beispiel illustrieren. Betrachten wir die Ableitungsoperatoren $A_n: L^2(\mathbb{R}) \rightarrow L^2(\mathbb{R})$ mit derselben Operatorvorschrift $u \mapsto u'$, aber verschiedenen Definitionsbereichen:

$$A_n u = \frac{du}{dx}, \quad \mathcal{D}_{A_n} = \{u \in L^2(\mathbb{R}) \cap C^n(\mathbb{R}) \mid u' \in L^2(\mathbb{R})\}, \quad n \in \mathbb{N}. \quad (5.106)$$

Diese Operatorvorschrift ist auf allen $u \in \mathcal{D}_{A_n}$ ausführbar und liefert Elemente $u' \in L^2(\mathbb{R})$. Die $(n+1)$ -fach stetig differenzierbaren Funktionen bilden eine Untermenge der n -fach stetig differenzierbaren Funktionen, $C^{n+1}(\mathbb{R}) \subset C^n(\mathbb{R})$, und somit gilt $\mathcal{D}_{A_{n+1}} \subset \mathcal{D}_{A_n}$ und $A_{n+1} \subset A_n$, d.h. A_n ist eine Erweiterung von A_{n+1} . Deswegen hat z.B. A_4 mindestens drei verschiedene Erweiterungen, nämlich A_3 , A_2 und A_1 . Alle Definitionsbereiche \mathcal{D}_{A_n} liegen dicht im Hilbertraum $L^2(\mathbb{R})$, d.h. $\overline{\mathcal{D}_{A_n}} = L^2(\mathbb{R}) \forall n \in \mathbb{N}$.

Beschränkte Operatoren können auf diese Weise immer auf den ganzen Hilbertraum erweitert werden. Nehmen wir zu \mathcal{D}_A den Grenzwert u jeder Cauchyfolge $(u_n) \in \mathcal{D}_A$ hinzu, bilden wir den Abschluss $\overline{\mathcal{D}_A}$. Die Stetigkeit (5.25) garantiert dann, dass die Bildfolge (Au_n) ebenfalls eine Cauchyfolge ist, und die Fortsetzung des Operators A erklärt Au auch für alle Grenzwerte u . Da diese Prozedur immer möglich ist, kann man bei beschränkten Operatoren von vornherein immer von $\mathcal{D}_A = \mathcal{H}$ ausgehen – deswegen mussten wir uns in Abschnitt 5.2 auch keine Gedanken über Definitionsbereiche machen.

Unbeschränkte Operatoren sind gleichzeitig unstetig, deswegen kann die Eigenschaft (5.26) für sie nicht mehr gelten. Es wird also Cauchyfolgen $(u_n) \in \mathcal{D}_A$ geben, deren Bilder (Au_n) keine Cauchyfolgen mehr sind und divergieren. Die Idee ist dann, nur die Grenzwerte *jener* Cauchyfolgen (u_n) in \mathcal{D}_A aufzunehmen, deren Bilder ebenfalls Cauchyfolgen sind, und die entsprechenden Grenzwerte als Au zu definieren. Dadurch erweitert man die Definitionsmenge \mathcal{D}_A ; den neuen Operator nennt man den **Abschluss** \bar{A} von A . Vorsicht: dieser Prozess ist nicht dasselbe wie die Vervollständigung von \mathcal{D}_A , d.h. $\mathcal{D}_{\bar{A}} \neq \overline{\mathcal{D}_A}$, denn wir nehmen ja nicht die Grenzwerte *aller* Cauchyfolgen in den Definitionsbereich auf. Das ist nur bei beschränkten Operatoren möglich, wo der so erweiterte Definitionsbereich (ganz \mathcal{H}) dann auch abgeschlossen ist.

DEFINITION: Seien X und Y zwei metrische Räume. Ein Operator $A: X \rightarrow Y$ auf $\mathcal{D}_A \subset X$ heißt **abgeschlossen** genau dann, wenn für jede konvergente Folge $(u_n) \subset \mathcal{D}_A$ mit Grenzwert u , für die auch $(Au_n) \subset \mathcal{R}_A$ eine konvergente Folge mit Grenzwert v ist, gilt:

$$u \in \mathcal{D}_A, \quad Au = v. \quad (5.107)$$

Ein Operator heißt **abschließbar** genau dann, wenn A eine abgeschlossene Erweiterung $B = \bar{B}$ besitzt, d.h. $A \subset B$. Der **Abschluss** \bar{A} ist die kleinste abgeschlossene Erweiterung eines abschließbaren Operators A .

Der Begriff der Abgeschlossenheit ähnelt jenem der Stetigkeit in (5.26), ist aber schwächer: ein abgeschlossener Operator erfüllt wieder $Au = A(\lim_{n \rightarrow \infty} u_n) = \lim_{n \rightarrow \infty} Au_n$, aber eben nur für jene konvergenten Folgen (u_n) , für die auch (Au_n) konvergent ist. Man kann zeigen, dass aus jeweils zwei der Eigenschaften: (i) A ist abgeschlossen, (ii) \mathcal{D}_A ist abgeschlossen und (iii) A ist beschränkt, die dritte folgt, d.h.:

- A ist abgeschlossen und \mathcal{D}_A ist abgeschlossen $\Rightarrow A$ ist beschränkt (diese Aussage ist auch als **Satz vom abgeschlossenen Graphen** bekannt);
- A ist beschränkt und \mathcal{D}_A ist abgeschlossen $\Rightarrow A$ ist abgeschlossen;
- A ist abgeschlossen und beschränkt $\Rightarrow \mathcal{D}_A$ ist abgeschlossen.

Daraus folgt insbesondere, dass ein abgeschlossener, unbeschränkter Operator keinen abgeschlossenen Definitionsbereich haben kann – ein solcher Operator kann daher nie auf ganz \mathcal{H} definiert werden!

Betrachten wir als Beispiel den Operator $A: L^2(0,1) \rightarrow L^2(0,1)$ mit

$$Au(x) = \frac{u(x)}{x^\alpha}, \quad \alpha > 0, \quad \mathcal{D}_A = \{u \in L^2(0,1) \mid u(x) = 0 \text{ in einer Umgebung von } x = 0\}. \quad (5.108)$$

Der Definitionsbereich \mathcal{D}_A liegt zwar dicht in $L^2(0,1)$, aber der Operator ist nicht abgeschlossen. Um das zu sehen, wählen wir z.B. $u(x) = x^\alpha$. Diese Funktion liegt offenbar nicht in \mathcal{D}_A , wird aber approximiert durch die Folge

$$u_n(x) = \begin{cases} x^\alpha & \frac{1}{n} < x \leq 1 \\ 0 & \text{sonst} \end{cases} \quad \Rightarrow \quad Au_n(x) = \begin{cases} 1 & \frac{1}{n} < x \leq 1 \\ 0 & \text{sonst} \end{cases} \quad (5.109)$$

Es gilt $u_n \in \mathcal{D}_A$, $u_n \rightarrow u$ und $Au_n \rightarrow 1$, doch wegen $u \notin \mathcal{D}_A$ ist A nicht abgeschlossen. Später werden wir mit Hilfe des adjungierten Operators sehen, dass der folgende Operator abgeschlossen ist:

$$A_1 u(x) = \frac{u(x)}{x^\alpha}, \quad \alpha > 0, \quad \mathcal{D}_{A_1} = \left\{ u \in L^2(0,1) \mid \frac{u}{x^\alpha} \in L^2(0,1) \right\}. \quad (5.110)$$

Kehren wir zum **Ableitungsoperator** $A: X \rightarrow Y$ mit der Vorschrift $Au = du/dx = u'$ zurück, wobei X und Y Funktionenräume über dem Intervall $\Omega = [a, b] \subset \mathbb{R}$ sind. Welchen Definitionsbereich \mathcal{D}_A wählt man sinnvollerweise, d.h. welche Bedingungen soll u erfüllen? Damit wir eine Funktion u differenzieren können, sollte $u \in C^1(\Omega)$ sein. Außerdem soll $Au = u'$ wieder in Y liegen. Wir könnten also fordern: $\mathcal{D}_A = \{u \in X \mid u \in C^1(\Omega), u' \in Y\}$. Beschränken wir uns auf den Banachraum $X = Y = C(\Omega)$ der stetigen Funktionen, sodass $A: C(\Omega) \rightarrow C(\Omega)$ mit $\mathcal{D}_A = C^1(\Omega)$, kann man zeigen, dass A abgeschlossen ist: der Raum $C^1(\Omega)$ ist selbst wieder ein Banachraum, d.h. jede Cauchyfolge $(u_n) \subset C^1(\Omega)$ konvergiert gegen eine Funktion $u \in C^1(\Omega)$, und die Ableitungen differenzierbarer Funktionen sind gerade die stetigen Funktionen.

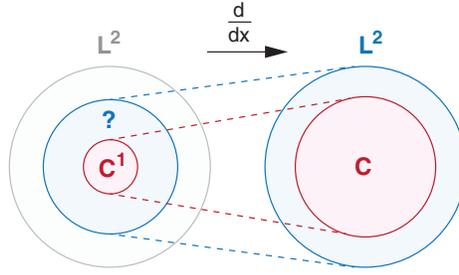


FIG. 5.3: Abschluss des Ableitungsoperators: Was ist der kleinstmögliche sinnvolle Definitionsbereich, sodass die Bilder L^2 -Funktionen sind?

Andererseits sind wir an Hilberträumen interessiert, doch die stetigen Funktionen bilden keinen Hilbertraum. Gehen wir daher zum Raum der quadratintegriblen Funktionen $L^2(\Omega)$ und sehen uns den Operator $A : L^2(\Omega) \rightarrow L^2(\Omega)$ an. Die Situation ist in Fig. 5.3 skizziert. Setzen wir jetzt $\mathcal{D}_A = C^1(\Omega)$, dann ist A nicht abgeschlossen: Wenn wir Cauchyfolgen von differenzierbaren Funktionen in der L^2 -Norm bilden, wird das aus $C^1(\Omega)$ herausführen. Ebenso wird es im Bildraum Cauchyfolgen stetiger Funktionen geben, die in der L^2 -Norm gegen unstetige (aber integrable Funktionen) konvergieren – denken Sie an das Beispiel der Stufenfunktion in (1.20) zurück. Solche unstetigen Funktionen werden aber nicht durch Ableitungen differenzierbarer Funktionen erzeugt. Was ist dann die kleinstmögliche konsistente Erweiterung von \mathcal{D}_A , d.h. wie lautet der Abschluss \bar{A} des Operators?

Betrachten wir dazu eine Folge $(u_n) \in C^1(\Omega)$, die gegen ein Element $u \in \mathcal{D}_{\bar{A}}$ konvergiert (wobei $\mathcal{D}_{\bar{A}}$ noch zu bestimmen bleibt), sowie die zugehörige Folge $(u'_n) \in C(\Omega)$, die gegen ein Element $v = \bar{A}u \in L^2(\Omega)$ konvergiert, also $\|u_n - u\| \rightarrow 0$ und $\|u'_n - v\| \rightarrow 0$ in der L^2 -Norm. Welche Bedingung muss u dann erfüllen? Dazu stellen wir fest, dass wir jede differenzierbare Funktion $u_n \in C^1[a, b]$ folgendermaßen schreiben können:

$$u_n(x) = u_n(a) + \int_a^x dy u'_n(y) \quad \forall x \in [a, b]. \quad (5.111)$$

Was passiert im Limes $n \rightarrow \infty$? Laut Voraussetzung gilt $\|u_n - u\| \rightarrow 0$, und aus $\|u'_n - v\| \rightarrow 0$ kann man zeigen, dass auch $\int_a^x dy u'_n(y)$ in der L^2 -Norm gegen $\int_a^x dy v(y)$ konvergiert. Dann muss auch $u_n(0)$ eine konvergente Folge sein. Der Grenzwert lautet also

$$u(x) = u(a) + \int_a^x dy v(y). \quad (5.112)$$

Das ist genau die Eigenschaft, nach der wir gesucht haben. Funktionen dieser Form heißen **absolut stetig**. Hierbei muss v keine stetige Funktion mehr sein, sondern nur integrabel. Falls v stetig ist, dann ist u überall differenzierbar und $u' = v$; ansonsten stellt (5.112) eine nützliche Verallgemeinerung des Begriffs der Differenzierbarkeit dar, denn es gilt $u' = v$ fast überall und z.B. sind stückweise differenzierbare Funktionen jetzt auch inkludiert. Somit ist der Abschluss von A gegeben durch $\bar{A} : L^2(\Omega) \rightarrow L^2(\Omega)$ mit

$$\bar{A}u = u', \quad \mathcal{D}_{\bar{A}} = \{u \in L^2(\Omega) \mid u \text{ absolut stetig, } u' \in L^2(\Omega)\}. \quad (5.113)$$

Dieser Definitionsbereich ist ein Spezialfall der **Sobolev-Räume** $W^{k,p}(\Omega)$ für allgemeine L^p -Räume und (schwache) Ableitungen der Ordnung k ; im obigen Fall ist $p = 2$ und $k = 1$.

Adjungierter Operator. Die Probleme mit unbeschränkten Operatoren setzen sich in der Definition des adjungierten Operators (5.62) fort. Im unbeschränkten Fall ist diese nicht mehr trivial, weil dafür mehrere Voraussetzungen fehlen. Betrachten wir für ein fixes $v \in \mathcal{H}$ das Funktional $F[u] = \langle v, Au \rangle$. Wenn der Operator A unbeschränkt ist, ist dieses Funktional möglicherweise auch unbeschränkt und somit nicht stetig, und dann garantiert der Satz von Riesz und Fréchet nicht mehr, dass jedes $F[u]$ als $\langle f, u \rangle$ geschrieben werden kann, was uns ja erlaubt hat, den adjungierten Operator A^\dagger mittels $f = A^\dagger v$ zu definieren.

Versuchen wir, einen adjungierten Operator ohne diese Voraussetzungen zu definieren. Nehmen wir zunächst an, es gäbe irgendwelche festen Vektoren $v \in \mathcal{H}$, für die die Relation

$$\langle f, u \rangle = \langle v, Au \rangle \quad \forall u \in \mathcal{D}_A \quad (5.114)$$

nach wie vor gilt, mit irgendwelchen f , die dann von v abhängen. Wir bezeichnen alle Paare (v, f) , für die die Zuordnung (5.114) gelingt, als "zulässig". Zumindest der Nullvektor $v = f = 0$ liegt sicher in dieser Menge. Solange \mathcal{D}_A dicht in \mathcal{H} liegt, ist f auch eindeutig durch v bestimmt: wäre nämlich (5.114) für verschiedene f_1, f_2 erfüllt, dann gilt $\langle f_1 - f_2, u \rangle = 0 \quad \forall u \in \mathcal{D}_A$; dann muss aber $f_1 - f_2 \in (\mathcal{D}_A)^\perp = \{0\}$ sein und somit $f_1 = f_2$. (Wäre \mathcal{D}_A nicht dicht in \mathcal{H} , könnten wir ein beliebiges Element aus $(\mathcal{D}_A)^\perp$ addieren, ohne die Gültigkeit der Gleichung zu verletzen.) Da f von v abhängt, können wir $f = A^\dagger v$ schreiben, wodurch der adjungierte Operator definiert wird. Dieser hat jetzt eine wohldefinierte Definitionsmenge \mathcal{D}_{A^\dagger} , nämlich eben jene v , für die (5.114) gilt. Damit können wir definieren:

DEFINITION: Seien \mathcal{H} und \mathcal{H}' Hilberträume und $A: \mathcal{H} \rightarrow \mathcal{H}'$ ein linearer Operator mit dichtem Definitionsbereich $\overline{\mathcal{D}_A} = \mathcal{H}$. Dann ist der **adjungierte Operator** definiert durch

$$\langle A^\dagger v, u \rangle = \langle v, Au \rangle \quad \forall u \in \mathcal{D}_A, v \in \mathcal{D}_{A^\dagger}. \quad (5.115)$$

Betrachten wir als Beispiel wieder den Ableitungsoperator $Au = u'$ auf dem Hilbertraum $\mathcal{H} = L^2(\Omega)$ mit $\Omega = [a, b] \subset \mathbb{R}$. Wir setzen ihn als abgeschlossen voraus, d.h. sein Definitionsbereich ist durch (5.113) gegeben. Um den adjungierten Operator zu bestimmen, müssen wir jene zulässigen Paare (v, f) finden, die $\langle f, u \rangle = \langle v, Au \rangle$ für alle $u \in \mathcal{D}_A$ erfüllen, also

$$\int_a^b dx f^*(x) u(x) \stackrel{!}{=} \int_a^b dx v^*(x) u'(x) \quad \forall u \in \mathcal{D}_A. \quad (5.116)$$

Durch partielle Integration der rechten Seite bekommen wir:

$$\int_a^b dx u(x) (f(x) + v'(x))^* = u(b) v^*(b) - u(a) v^*(a) \quad \forall u \in \mathcal{D}_A. \quad (5.117)$$

Sofern die rechte Seite verschwindet, lesen wir aus der linken Seite $f = A^\dagger v = -v'$ ab. Da A^\dagger dann wieder ein Ableitungsoperator ist, muss dessen Definitionsbereich ebenfalls (5.113) genügen. Ohne weitere Einschränkungen an A lautet der adjungierte Operator also

$$A^\dagger v = -\frac{dv}{dx}, \quad \mathcal{D}_{A^\dagger} = \{v \in \mathcal{H} \mid v \text{ absolut stetig, } v' \in L^2, v(a) = v(b) = 0\}. \quad (5.118)$$

Das ist offenbar ein engerer Definitionsbereich als \mathcal{D}_A .

Was passiert, wenn wir von vornherein weitere Randbedingungen an A stellen? Betrachten wir dazu folgende Operatoren:

$$\begin{aligned}
A u &= \frac{du}{dx}, & \mathcal{D}_A &= \{u \in \mathcal{H} \mid u \text{ absolut stetig, } u' \in \mathcal{H}\}, \\
A_a u &= \frac{du}{dx}, & \mathcal{D}_{A_a} &= \{u \in \mathcal{H} \mid u \text{ absolut stetig, } u' \in \mathcal{H}, u(a) = 0\}, \\
A_b u &= \frac{du}{dx}, & \mathcal{D}_{A_b} &= \{u \in \mathcal{H} \mid u \text{ absolut stetig, } u' \in \mathcal{H}, u(b) = 0\}, \\
A_{ab} u &= \frac{du}{dx}, & \mathcal{D}_{A_{ab}} &= \{u \in \mathcal{H} \mid u \text{ absolut stetig, } u' \in \mathcal{H}, u(a) = u(b) = 0\}, \\
A_z u &= \frac{du}{dx}, & \mathcal{D}_{A_k} &= \{u \in \mathcal{H} \mid u \text{ absolut stetig, } u' \in \mathcal{H}, u(a) = zu(b), z \in \mathbb{C} \setminus \{0\}\}.
\end{aligned} \tag{5.119}$$

Ohne weitere Einschränkungen an A ist \mathcal{D}_{A^\dagger} durch (5.118) gegeben; das entspricht offenbar aber gerade $\mathcal{D}_{A_{ab}}$. Für den Operator A_a mit der Randbedingung $u(a) = 0$ muss dagegen $v(b) = 0$ gelten, damit die rechte Seite in (5.117) verschwindet. Insgesamt finden wir:

$$\mathcal{D}_{A^\dagger} = \mathcal{D}_{A_{ab}}, \quad \mathcal{D}_{A_a^\dagger} = \mathcal{D}_{A_b}, \quad \mathcal{D}_{A_b^\dagger} = \mathcal{D}_{A_a}, \quad \mathcal{D}_{A_{ab}^\dagger} = \mathcal{D}_A, \quad \mathcal{D}_{A_z^\dagger} = \mathcal{D}_{A_{1/z^*}}. \tag{5.120}$$

Die letzte Relation ergibt sich daraus, dass für $u(a) = zu(b)$ die rechte Seite in (5.117) verschwindet, wenn $v^*(b) = zv^*(a)$ erfüllt ist, also $v(a) = v(b)/z^*$. Allgemein kann man zeigen, dass für einen dicht definierten Operator $A \subset B \Rightarrow B^\dagger \subset A^\dagger$ gilt, was hier offenbar erfüllt ist: z.B. ist $\mathcal{D}_{A_{ab}} \subset \mathcal{D}_A$ und damit $A_{ab} \subset A$, und umgekehrt ist $A^\dagger \subset A_{ab}^\dagger$.

Der adjungierte Operator bietet eine bequeme Methode, um den Abschluss eines Operators A zu bilden bzw. Aussagen darüber zu treffen, wann ein Operator abschließbar ist. Für einen dicht liegenden linearen Operator gilt nämlich [Bla93, Gro14]:

- Der zu A adjungierte Operator A^\dagger ist automatisch abgeschlossen: $A^\dagger = \overline{A^\dagger}$. (5.121)

Das folgt aus der Stetigkeit des Skalarprodukts: Sei $(v_n) \subset \mathcal{D}_{A^\dagger}$ eine Folge mit Grenzwert v , für die auch die Folge $(A^\dagger v_n)$ gegen einen Grenzwert f konvergiert. Dann folgt aus $\langle A^\dagger v_n, u \rangle = \langle v_n, Au \rangle$ für alle $u \in \mathcal{D}_A$ und $n \in \mathbb{N}$ im Limes $n \rightarrow \infty$: $\langle f, u \rangle = \langle v, Au \rangle$, und somit $v \in \mathcal{D}_{A^\dagger}$ und $A^\dagger v = f$.

- Falls A einen Abschluss \overline{A} besitzt, gilt auch $A^\dagger = (\overline{A})^\dagger$. (5.122)

- A ist abschließbar genau dann, wenn \mathcal{D}_{A^\dagger} dicht in \mathcal{H} liegt; dann gilt $\overline{A} = A^{\dagger\dagger}$. (5.123)

Kommen wir auf das Beispiel (5.108) mit $Au(x) = u(x)/x^\alpha$, $\alpha \in \mathbb{R}$ zurück. Wir haben gesehen, dass dieser Operator auf dem Definitionsbereich \mathcal{D}_A nicht abgeschlossen war. Wie lautet der adjungierte Operator? Aus $\langle A^\dagger v, u \rangle = \langle v, Au \rangle$ folgt

$$\int_0^1 dx \left(A^\dagger v(x) \right)^* u(x) \stackrel{?}{=} \int_0^1 dx v^*(x) \frac{u(x)}{x^\alpha} = \int_0^1 dx \left(\frac{v(x)}{x^\alpha} \right)^* u(x) \quad \forall u \in L^2(0, 1), \tag{5.124}$$

und somit

$$A^\dagger v(x) = \frac{v(x)}{x^\alpha}, \quad \mathcal{D}_{A^\dagger} = \left\{ v \in L^2(0, 1) \mid \frac{v}{x^\alpha} \in L^2(0, 1) \right\} \supset \mathcal{D}_A. \tag{5.125}$$

Es gilt $\mathcal{D}_A \subset \mathcal{D}_{A^\dagger}$, denn die Menge der Funktionen $u \in L^2(0, 1)$ mit $u(x) = 0$ in einer Umgebung von $x = 0$ ist eine Untermenge von \mathcal{D}_{A^\dagger} . Weil die Operatorvorschrift dieselbe ist, finden wir $A \subset A^\dagger$; wie wir gleich sehen werden, ist das die definierende Eigenschaft eines *symmetrischen* Operators. Wir wissen inzwischen aber auch, dass A^\dagger abgeschlossen ist; somit ist auch $A_1 = A^\dagger$ in (5.110) abgeschlossen.

Symmetrisch und selbstadjungiert. Die Beispiele in (5.119) machen klar, dass der Definitionsbereich \mathcal{D}_{A^\dagger} des adjungierten Operators sowohl größer als auch kleiner als \mathcal{D}_A sein kann. Um einen selbstadjungierten Operator zu konstruieren, müssen wir daher auch die Definitionsbereiche aneinander anpassen.

DEFINITION:

- Ein linearer, dicht definierter Operator A auf \mathcal{D}_A heißt **symmetrisch** genau dann, wenn

$$\langle Av, u \rangle = \langle v, Au \rangle \quad \forall u, v \in \mathcal{D}_A. \quad (5.126)$$

Das ist gleichbedeutend mit der Forderung $A \subset A^\dagger$, oder ausgeschrieben:

$$A^\dagger u = Au \quad \forall u \in \mathcal{D}_A \quad \text{und} \quad \mathcal{D}_A \subset \mathcal{D}_{A^\dagger}. \quad (5.127)$$

- Ein Operator heißt **selbstadjungiert** genau dann, wenn $A^\dagger = A$, oder ausgeschrieben:

$$A^\dagger u = Au \quad \forall u \in \mathcal{D}_A \quad \text{und} \quad \mathcal{D}_A = \mathcal{D}_{A^\dagger}. \quad (5.128)$$

- Ein Operator heißt **wesentlich selbstadjungiert** genau dann, wenn er symmetrisch ist und sein Abschluss \bar{A} selbstadjungiert ist.

Der Ableitungsoperator $A = d/dx$ erfüllt offenbar keine dieser Bedingungen, denn es mangelt ja schon an der Operatorvorschrift ($A^\dagger = -A$). Starten wir stattdessen aber mit dem Operator $P = -id/dx$ und führen die Schritte (5.116–5.117) aus, sehen wir, dass der adjungierte Operator P^\dagger dieselbe Operatorvorschrift besitzt wie P selbst: $P^\dagger = -id/dx$. Damit dieser Operator symmetrisch wird, muss aber zusätzlich $\mathcal{D}_P \subset \mathcal{D}_{P^\dagger}$ gelten. Sehen wir uns dazu die Operatoren P , P_a , P_b und P_{ab} analog zu (5.119) an. Die ersten drei sind offenbar nicht symmetrisch, denn wegen $\mathcal{D}_{P^\dagger} = \mathcal{D}_{P_{ab}}$ und $\mathcal{D}_{P_a^\dagger} = \mathcal{D}_{P_b}$ gilt weder $\mathcal{D}_P \subset \mathcal{D}_{P^\dagger}$ noch $\mathcal{D}_{P_a} \subset \mathcal{D}_{P_a^\dagger}$. Der vierte Operator P_{ab} ist symmetrisch, aber nicht selbstadjungiert, denn es gilt $\mathcal{D}_{P_{ab}^\dagger} = \mathcal{D}_P$ und somit $\mathcal{D}_{P_{ab}} \subset \mathcal{D}_{P_{ab}^\dagger}$.

Um einen selbstadjungierten Operator zu erhalten, müssen wir die Definitionsbereiche von P und P^\dagger offenbar so abstimmen, dass die rechte Seite $u(b)v^*(b) - u(a)v^*(a)$ in (5.117) Null wird, und zwar genau so, dass *dieselbe* Bedingung für u und v resultiert. Versuchen wir es daher mit dem fünften Operator P_z in (5.119):

$$u(a) = z u(b) \quad \Rightarrow \quad z v^*(b) = v^*(a) \quad \Rightarrow \quad v(a) = \frac{1}{z^*} v(b). \quad (5.129)$$

Das liefert nur dann dieselbe Bedingung, wenn $zz^* = |z|^2 = 1$ erfüllt ist, also $z = e^{i\gamma}$. In diesem Fall ist der Operator selbstadjungiert. Es ergibt sich sogar eine unendliche Schar selbstadjungierter Impulsoperatoren auf dem Hilbertraum $\mathcal{H} = L^2(a, b)$, die durch $\gamma \in \mathbb{R}$ charakterisiert sind:

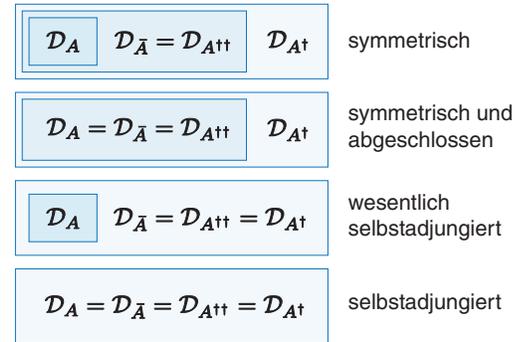
$$P_\gamma u = -i \frac{du}{dx}, \quad \mathcal{D}_{P_\gamma} = \left\{ u \in \mathcal{H} \mid u \text{ absolut stetig, } u' \in \mathcal{H}, u(a) = e^{i\gamma} u(b) \right\}. \quad (5.130)$$

Bei beschränkten Operatoren spielen diese Unterschiede keine Rolle, da sie immer auf ganz \mathcal{H} definiert werden können; d.h. dort bedeuten *symmetrisch*, *selbstadjungiert* und *wesentlich selbstadjungiert* dasselbe. Dasselbe gilt natürlich auch für Operatoren auf endlich-dimensionalen Hilberträumen, die ja immer beschränkt sind. Es sind aber gerade die unbeschränkten selbstadjungierten Operatoren, die in der Quantenmechanik wichtig sind.

Dazu einige Bemerkungen:

- Weil für einen symmetrischen Operator $\mathcal{D}_A \subset \mathcal{D}_{A^\dagger}$ gilt und \mathcal{D}_A dicht in \mathcal{H} liegt, liegt auch \mathcal{D}_{A^\dagger} dicht in \mathcal{H} . Aus (5.123) folgt dann, dass ein symmetrischer Operator A immer abschließbar ist, und sein Abschluss ist $\bar{A} = A^{\dagger\dagger}$.
- Für einen symmetrischen Operator nimmt der Erwartungswert $\langle u, Au \rangle$ reelle Werte an, da $\langle u, Au \rangle = \langle Au, u \rangle = \langle u, Au \rangle^*$ für alle $u \in \mathcal{D}_A$.
- Ein selbstadjungierter Operator ist immer abgeschlossen, da $A = A^\dagger$ gilt und A^\dagger abgeschlossen ist.

- Ein selbstadjungierter Operator ist auch wesentlich selbstadjungiert, und letzterer ist auch symmetrisch. Die Umkehrungen gelten im Allgemeinen nicht. Die Situation ist in der nebenstehenden Figur skizziert, z.B. erfüllt ein symmetrischer Operator nur



$$A \subset \bar{A} = A^{\dagger\dagger} \subset A^\dagger, \quad (5.131)$$

und erst für einen selbstadjungierten gilt die Gleichheit $A = \bar{A} = A^{\dagger\dagger} = A^\dagger$.

- Der Nachweis der Selbstadjungiertheit ist oft nicht einfach, da man dazu nicht nur die Rechenvorschrift “ $A = A^\dagger$ ” beweisen muss, sondern auch einen Definitionsbereich finden muss, der dicht in \mathcal{H} liegt und für den $\mathcal{D}_A = \mathcal{D}_{A^\dagger}$ gilt. Deswegen sind die folgenden weiteren Kriterien zur Selbstadjungiertheit nützlich. Für einen symmetrischen Operator A in einem Hilbertraum \mathcal{H} sind folgende Aussagen äquivalent [Bla93]:

$$A = A^\dagger \Leftrightarrow \bar{A} = A \text{ und } \mathcal{K}_{A^\dagger \pm i \mathbb{1}} = \{0\}, \Leftrightarrow \mathcal{R}_{A \pm i \mathbb{1}} = \mathcal{H}. \quad (5.132)$$

Betrachten wir den Multiplikationsoperator aus Gl. (5.101) mit $A: L^2(\Omega) \rightarrow L^2(\Omega)$, $Au(x) = h(x)u(x)$ und $\mathcal{D}_A = \{u \in L^2(\Omega) \mid hu \in L^2(\Omega)\}$, wobei h eine reellwertige Funktion ist. Dieser Operator ist selbstadjungiert. Zunächst kann man sich überlegen, dass \mathcal{D}_A tatsächlich dicht in $L^2(\Omega)$ liegt. Sei nun $v \in \mathcal{D}_{A^\dagger}$, dann gibt es ein $f = A^\dagger v \in L^2(\Omega)$ sodass $\langle v, Au \rangle = \langle f, u \rangle$, d.h.

$$\int_{\Omega} dx (h(x)v(x) - f(x))^* u(x) = 0 \quad \forall u \in \mathcal{D}_A. \quad (5.133)$$

Laut Operatorvorschrift ist $A^\dagger v(x) = h(x)v(x)$, d.h. der Operator A ist symmetrisch. Für einen dicht definierten symmetrischen Operator folgt die Eigenschaft $A \subset A^\dagger$ automatisch. Um zu zeigen, dass A selbstadjungiert ist, muss man die Gegenrichtung $A \supset A^\dagger$ beweisen, d.h. man muss zeigen, dass jedes mögliche $v \in \mathcal{D}_{A^\dagger}$ in \mathcal{D}_A liegt. Das ist etwas mühsamer [Bla93] und geht so: Die obige Gleichung gilt für alle $u \in \mathcal{D}_A$. Wählt man $u(x) = \chi_R(x)w(x)$ mit $w \in L^2(\Omega)$, wobei χ_R eine geeignete Abschneidefunktion mit einem Abschneideparameter R ist, sodass $\chi_R(x) = 1$ für $R \rightarrow \infty$, dann muss auch der Integrand verschwinden, d.h. $\chi_R(x)f(x) = \chi_R(x)h(x)v(x)$. Im Limes $R \rightarrow \infty$ folgt dann $f(x) = h(x)v(x)$, und wegen $f \in L^2(\Omega)$ muss $v \in \mathcal{D}_A$ gelten.

5.4 Das Spektrum linearer Operatoren

Die Bestimmung des **Spektrums** eines Operators ist die Verallgemeinerung des Eigenwertproblems in der linearen Algebra auf unendlichdimensionale Hilberträume. Dazu sammeln wir zuerst die wichtigsten Aussagen für lineare Operatoren auf endlichdimensionalen Hilberträumen, die man ja immer durch komplexe $n \times n$ -Matrizen darstellen kann.

Gilt für eine komplexe $n \times n$ -Matrix \mathbf{A} die Relation $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$ für ein $\mathbf{v} \neq 0$, mit $\lambda \in \mathbb{C}$, dann heißt λ **Eigenwert** der Matrix \mathbf{A} zum **Eigenvektor** \mathbf{v} . Die Menge aller Eigenwerte von \mathbf{A} heißt **Spektrum** von \mathbf{A} . Ist λ ein Eigenwert, dann sind die folgenden Aussagen äquivalent:

- $(\mathbf{A} - \lambda)\mathbf{v} = 0$ für ein $\mathbf{v} \neq 0$;
- Die Matrix $\mathbf{A} - \lambda$ ist nicht invertierbar, d.h. das Inverse $(\mathbf{A} - \lambda)^{-1}$ existiert nicht;
- Es gibt ein $\mathbf{v} \neq 0$ aus dem Nullraum $\mathcal{K}_{\mathbf{A}-\lambda}$, d.h. der Nullraum $\mathcal{K}_{\mathbf{A}-\lambda}$ ist nichttrivial;
- Es gilt $\det(\mathbf{A} - \lambda) = 0$.

Die Determinante liefert ein bequemes Kriterium, um die Eigenwerte λ von \mathbf{A} zu bestimmen, denn $\det(\mathbf{A} - \lambda)$ ist ein Polynom vom Grad n in λ mit genau n Nullstellen $\lambda_1, \dots, \lambda_n \in \mathbb{C}$. Diese müssen nicht alle verschieden sein: die **algebraische Vielfachheit** bezeichnet die Vielfachheit eines Eigenwerts als Nullstelle des Polynoms; ist sie 1, heißt der Eigenwert **einfach**. Somit besitzt jede komplexe $n \times n$ -Matrix genau n Eigenwerte, wenn man sie mit ihren Vielfachheiten zählt. Die Eigenvektoren zu verschiedenen Eigenwerten sind linear unabhängig.

Der **Eigenraum** zu einem gegebenen Eigenwert λ ist der Nullraum $\mathcal{K}_{\mathbf{A}-\lambda}$. Seine Dimension $\dim \mathcal{K}_{\mathbf{A}-\lambda}$ heißt **geometrische Vielfachheit** von λ und gibt die Anzahl der Eigenvektoren für dieses λ an (sie ist kleiner oder gleich der algebraischen Vielfachheit). Ist sie größer als 1, heißt der Eigenwert **entartet**. Es gilt $\dim \mathcal{K}_{\mathbf{A}-\lambda} + \dim \mathcal{R}_{\mathbf{A}-\lambda} = n$, wobei $\dim \mathcal{R}_{\mathbf{A}-\lambda}$ den Rang der Matrix $\mathbf{A} - \lambda$ bezeichnet.

Für eine **hermitesche Matrix** $\mathbf{A} = \mathbf{A}^\dagger$ wird das Eigenwertproblem $\mathbf{A}\mathbf{v}_i = \lambda_i\mathbf{v}_i$ besonders einfach, denn ihre Eigenwerte $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$ sind reell und ihre Eigenvektoren $\mathbf{v}_1, \dots, \mathbf{v}_n$ bilden eine ONB von \mathbb{C}^n . (Das werden wir später allgemeiner beweisen.) Eine hermitesche Matrix kann immer auf eine **Diagonalform** gebracht werden, indem man eine unitäre Matrix $\mathbf{U} = (\mathbf{v}_1, \dots, \mathbf{v}_n)$ definiert, deren Spalten die Eigenvektoren sind. Dann ist nämlich

$$\mathbf{U}^\dagger \mathbf{A} \mathbf{U} = \mathbf{\Lambda} = \text{diag} \{ \lambda_1, \dots, \lambda_n \} \quad (5.134)$$

die Diagonalmatrix der Eigenwerte, oder umgekehrt: $\mathbf{A} = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^\dagger$. Das ist äquivalent zur **Spektraldarstellung** der Matrix \mathbf{A} , wobei $\langle \cdot, \cdot \rangle$ das Skalarprodukt im \mathbb{C}^n bezeichnet:

$$\mathbf{A}\mathbf{x} = \sum_{i=1}^n \mathbf{v}_i \underbrace{\langle \mathbf{v}_i, \mathbf{A}\mathbf{x} \rangle}_{=\langle \mathbf{A}\mathbf{v}_i, \mathbf{x} \rangle} = \sum_{i=1}^n \lambda_i \mathbf{v}_i \langle \mathbf{v}_i, \mathbf{x} \rangle \quad \forall \mathbf{x} \in \mathbb{C}^n \quad \Leftrightarrow \quad \mathbf{A} = \sum_{i=1}^n \lambda_i |\mathbf{v}_i\rangle \langle \mathbf{v}_i|. \quad (5.135)$$

Im Folgenden werden wir versuchen, diese Aussagen auf unendlichdimensionale Hilberträume zu übertragen. Nachdem man die Determinante nicht ohne Weiteres verallgemeinern kann, werden wir als "Spektrum" des Operators die Menge aller λ bezeichnen, für die $(\mathbf{A} - \lambda)^{-1}$ nicht existiert. Während das Spektrum einer Matrix \mathbf{A} nur aus ihren Eigenwerten besteht, werden wir sehen, dass der unendlichdimensionale Fall wesentlich komplizierter ist. Die Äquivalenz der Aussagen ganz oben ist dann im Allgemeinen *falsch*, z.B. gibt es selbstadjungierte Operatoren, die überhaupt keine Eigenwerte besitzen und trotzdem für bestimmte Werte $\lambda \in \mathbb{C}$ nicht invertierbar sind.

Inverser Operator. Bevor wir uns mit dem Eigenwertproblem beschäftigen, stellen wir uns die Frage, wann ein Operator $A : \mathcal{H} \rightarrow \mathcal{H}$ ein Inverses besitzt. Das zentrale Problem der Operatortheorie ist die Lösung (d.h. die Inversion) der inhomogenen Gleichung

$$Au = f, \quad f \in \mathcal{H}. \quad (5.136)$$

Wir wollen gleich den allgemeineren Fall unbeschränkter Operatoren miteinschließen, bei dem A nicht mehr auf ganz \mathcal{H} definiert werden kann, sondern nur auf $\mathcal{D}_A \subset \mathcal{H}$. Wir werden aber nach wie vor davon ausgehen, dass \mathcal{D}_A dicht in \mathcal{H} liegt, sodass $(\mathcal{D}_A)^\perp = \{0\}$. Der Operator A lässt sich dann als Abbildung $A : \mathcal{D}_A \rightarrow \mathcal{R}_A$ interpretieren, sodass jedes Element $u \in \mathcal{D}_A$ auf genau ein Element $f \in \mathcal{R}_A$ abgebildet wird.

Wann existiert der inverse Operator? Offenbar kann die Abbildung rückgängig gemacht werden, wenn A injektiv ist, d.h. wenn A je zwei verschiedenen Elementen aus \mathcal{D}_A verschiedene Elemente aus \mathcal{R}_A zuordnet. Dann ordnet der inverse Operator A^{-1} mit $u = A^{-1}f$ jedem $f \in \mathcal{R}_A$ genau ein Element aus $u \in \mathcal{D}_A$ zu, sodass $\mathcal{D}_{A^{-1}} = \mathcal{R}_A$ und $\mathcal{R}_{A^{-1}} = \mathcal{D}_A$.

SATZ: Der Operator A hat ein **Inverses** A^{-1} (d.h. A ist injektiv) genau dann, wenn die homogene Gleichung $Au = 0$ nur die Lösung $u = 0$ erlaubt, d.h. wenn der Nullraum \mathcal{K}_A nur das Nullelement enthält: $\mathcal{K}_A = \{0\}$.

Zum Beweis “ \Rightarrow ”: Falls A injektiv ist, kann $Au = 0$ maximal eine Lösung haben, und da $u = 0$ eine Lösung ist, ist sie die einzige. Zum Beweis “ \Leftarrow ”: Falls $Au = 0$ nur die Lösung $u = 0$ hat, müssen wir die Injektivität beweisen, d.h. wir müssen zeigen, dass $Av = f$ maximal eine Lösung hat. Für $Av_1 = f$ und $Av_2 = f$ folgt aus der Linearität von A und \mathcal{D}_A , dass $A(v_1 - v_2) = 0$, sodass $v_1 = v_2$.

Falls A^{-1} existiert, lassen sich weitere Aussagen treffen. Im Folgenden werden wir die Abgeschlossenheit des Operators A voraussetzen, denn dann ist auch das Inverse abgeschlossen:

SATZ: Falls A abgeschlossen ist und A^{-1} existiert, dann ist auch A^{-1} abgeschlossen.

Beweis: Sei f_n eine Folge in \mathcal{R}_A mit $f_n \rightarrow f$ und $u_n = A^{-1}f_n \rightarrow u$. Wir müssen zeigen, dass $f \in \mathcal{R}_A$ und $A^{-1}f = u$; dann ist A^{-1} abgeschlossen. Wegen $Au_n = f_n$, $f_n \rightarrow f$ und $u_n \rightarrow u$ impliziert die Abgeschlossenheit von A , dass $u \in \mathcal{D}_A$ und $Au = f$. Somit ist $f \in \mathcal{R}_A$ und $u = A^{-1}f$.

Wann ist der inverse Operator A^{-1} beschränkt? Ein beschränkter Operator ist nämlich auch stetig, sodass die Lösung $u = A^{-1}f$ stetig von f abhängt. Auf einem endlichdimensionalen Raum ist der inverse Operator (so wie alle Operatoren) automatisch beschränkt, aber auf unendlichdimensionalen Räumen ist das Inverse eines beschränkten Operators *nicht* notwendigerweise beschränkt. Offenbar muss folgendes gelten:

$$\|A^{-1}\| \leq c \quad \Leftrightarrow \quad \|A^{-1}f\| \leq c\|f\| \quad \Leftrightarrow \quad \|u\| \leq c\|Au\| \quad \Leftrightarrow \quad \|Au\| \geq \frac{\|u\|}{c}. \quad (5.137)$$

A muss also *von unten* durch die Null beschränkt sein (“**nullbeschränkt**”), d.h. Au darf für kein $u \in \mathcal{D}_A$ mit $u \neq 0$ Null werden, auch nicht als Limes einer Folge. Der Fall $Au = 0$ für irgendein $u \neq 0$ führt wieder auf den Fall zurück, wo die Inverse nicht existiert. Aber auch wenn eine Folge (u_n) mit $\|u_n\| \neq 0$ und $\|Au_n\| \rightarrow 0$ existiert, ist A^{-1} unbeschränkt.

		A^{-1} beschränkt	A^{-1} unbeschränkt	A^{-1} existiert nicht
$\mathcal{R}_A = \mathcal{H}$	$\overline{\mathcal{R}_A} = \mathcal{H}$	1A Regulär		3A
	$\overline{\mathcal{R}_A} \neq \mathcal{H}$		2B	3B
$\mathcal{R}_A \neq \mathcal{H}$	$\overline{\mathcal{R}_A} = \mathcal{H}$			3C
	$\overline{\mathcal{R}_A} \neq \mathcal{H}$	1C	2C	3C

FIG. 5.4: Klassifikation der Inversen eines abgeschlossenen Operators

Damit gelangen wir zur Klassifikation in Fig. 5.4. Die drei Spalten entsprechen dem bisher Gesagten: in der ersten Spalte ist A^{-1} beschränkt, in der zweiten unbeschränkt, und in der dritten existiert A^{-1} nicht, weil A einen nichttrivialen Nullraum besitzt. Für die Zeilen treffen wir eine weitere Unterscheidung:

- (A) $\mathcal{R}_A = \mathcal{H}$, d.h. der Wertebereich ist der gesamte Hilbertraum.
- (B) $\mathcal{R}_A \neq \mathcal{H}$, aber der Wertebereich liegt dicht in \mathcal{H} , d.h. $\overline{\mathcal{R}_A} = \mathcal{H}$ bzw. $\mathcal{R}_A^\perp = \{0\}$.
- (C) $\mathcal{R}_A \neq \mathcal{H}$ und $\overline{\mathcal{R}_A} \neq \mathcal{H}$, d.h. $\mathcal{R}_A^\perp \neq \{0\}$ und somit $\dim \mathcal{R}_A^\perp \neq 0$.

Für abgeschlossene Operatoren erlaubt der folgende Satz weitere Einschränkungen:

SATZ: Falls A abgeschlossen ist und A^{-1} existiert, dann ist \mathcal{R}_A genau dann abgeschlossen, wenn A^{-1} beschränkt ist.

- (A) $\mathcal{R}_A = \mathcal{H}$: Da \mathcal{H} vollständig (und somit abgeschlossen) ist, muss A^{-1} , sofern es existiert, beschränkt sein. Das führt auf die **regulären Operatoren**, die ein Inverses besitzen (1A). Alle anderen Zustände deuten auf irgendwelche Probleme hin, bei denen der inverse Operator entweder nicht auf ganz \mathcal{H} definiert werden kann oder gar nicht existiert. Insbesondere ist der Fall (2A) ausgeschlossen.
- (B) $\mathcal{R}_A \neq \mathcal{H}$, aber $\overline{\mathcal{R}_A} = \mathcal{H}$: Wenn \mathcal{R}_A nur dicht in \mathcal{H} liegt, ist \mathcal{R}_A nicht abgeschlossen. Falls A^{-1} existiert, muss es unbeschränkt sein (2B). Das schließt den Fall (1B) aus.
- (C) $\mathcal{R}_A \neq \mathcal{H}$ und $\overline{\mathcal{R}_A} \neq \mathcal{H}$: \mathcal{R}_A liegt nicht dicht in \mathcal{H} , d.h. sein orthogonales Komplement \mathcal{R}_A^\perp besteht nicht nur aus dem Nullvektor. Hier folgt nur, dass für beschränktes A^{-1} der Wertebereich \mathcal{R}_A abgeschlossen ist (1C) und für unbeschränktes A^{-1} nicht (2C).

Für den Spezialfall **beschränkter Operatoren** lässt sich der Operator A auf dem ganzen Hilbertraum definieren und ist dann auch abgeschlossen. Als Abbildung $A: \mathcal{H} \rightarrow \mathcal{H}$ liest sich Fig. 5.4 dann so: In den Fällen (1) und (2) ist A injektiv, im Fall (3) nicht. Im Fall (A) ist A surjektiv (da $\mathcal{R}_A = \mathcal{H}$), in den Fällen (B) und (C) nicht. Nur im Fall (1A) ist A bijektiv und besitzt ein Inverses. Für alle anderen Zustände ist A^{-1} nicht wohldefiniert, entweder weil der Operator nicht injektiv ist oder weil er nicht auf den ganzen Hilbertraum abbildet. Im Spezialfall **endlichdimensionaler Matrizen** treten nur die Zustände (1A) und (3C) auf; das ist eine Konsequenz der Fredholmschen Alternative, die wir später besprechen werden.

Untersuchen wir als Beispiel die **Ableitungsoperatoren** A, A_a, A_b, A_{ab}, A_z in (5.119). Sie führen alle auf die Differentialgleichung $u' = f$, allerdings mit verschiedenen Randbedingungen. Der inverse Operator (sofern er existiert) ist ein Integraloperator der Form

$$u(x) = \int_a^x dy f(y) + c = \int_a^b dy \Theta(x-y) f(y) + c, \quad (5.138)$$

wobei die Konstante c durch die jeweiligen Randbedingungen spezifiziert wird. Alle obigen Ableitungsoperatoren sind unbeschränkt, abgeschlossen und bilden zunächst auf den gesamten Hilbertraum $L^2(a, b)$ ab. Daran ändern auch die Randbedingungen nichts, denn dabei handelt es sich ja nur um Punkte vom Maß Null. Sehen wir uns an, was verschiedene Randbedingungen für A bewirken:

- Der “maximale” Operator A , der ohne Randbedingungen auskommt, erlaubt eine Lösung der homogenen Gleichung $Au(x) = u'(x) = 0$ mit $u(x) = \text{const}$. Er ist somit nicht invertierbar und gehört zum Typ (3A).
- Die Operatoren A_a und A_b mit Randbedingungen $u(a) = 0$ bzw. $u(b) = 0$ sind invertierbar, z.B. ist A_a^{-1} durch (5.138) mit $c = 0$ gegeben. Das ist ein Hilbert-Schmidt-Operator der Form (5.41) mit $K(x, y) = \Theta(x - y)$ und somit beschränkt. A_a und A_b sind damit reguläre Operatoren (1A).
- Für den Operator A_{ab} mit der Randbedingung $u(a) = u(b) = 0$ ergibt (5.138)

$$u(x) = \int_a^x dy f(y), \quad \int_a^b dy f(y) = 0. \quad (5.139)$$

A_{ab}^{-1} ist zwar beschränkt, aber \mathcal{R}_A deckt wegen der zusätzlichen Lösungsbedingung nicht mehr den gesamten Hilbertraum ab. Somit ist A_{ab} vom Typ (1C).

- Für den Operator A_z mit Randbedingung $u(a) = zu(b)$ liefert (5.138) für $z \neq 1$:

$$u(x) = \int_a^b dy G(x, y) f(y), \quad G(x, y) = \Theta(x - y) + \frac{z}{1 - z}. \quad (5.140)$$

Das ist wieder ein beschränkter Operator vom Typ (1A). Im Fall $z = 1$ resultiert dagegen keine Einschränkung an c , daher bleibt die allgemeine Lösungsformel (5.138) unverändert, allerdings ergibt sich wieder die Zusatzbedingung $\int_a^b dy f(y) = 0$; somit ist der Operator nicht invertierbar und gehört zum Typ (3C).

Weitere Beispiele:

- Der Multiplikationsoperator $Au(x) = xu(x)$ auf $L^2(0, 1)$ aus (5.18) ist beschränkt mit der Operatornorm $\|A\| = 1$. Allerdings gibt es kein u , für das dieser Wert tatsächlich erreicht wird, deswegen ist $\mathcal{R}(A)$ nicht abgeschlossen, liegt aber dicht in $L^2(0, 1)$. Der inverse Operator $u(x) = v(x)/x$ ist unbeschränkt, somit ist A vom Typ (2B).
- Für den Rechtsshift-Operator (5.34) ist $\|Au\|/\|u\| = 1$, $\mathcal{K}_A = \{0\}$ und $\mathcal{R}_A^\perp \neq \{0\}$, deshalb ist er vom Typ (1C). Der Linksshift-Operator erfüllt $\mathcal{K}_A \neq \{0\}$ und $\mathcal{R}_A = \mathcal{H}$ und ist vom Typ (3A).

Spektrum und Resolventenmenge. Wir kommen jetzt zur Bestimmung des Spektrums eines Operators. Dazu betrachten wir einen linearen, abgeschlossenen Operator A , der auf einer dichten Teilmenge \mathcal{D}_A eines Hilbertraums \mathcal{H} definiert ist. Statt $Au = f$ betrachten wir die inhomogene Operatorgleichung

$$(A - z\mathbb{1})u = f \quad \text{oder kurz} \quad (A - z)u = f, \quad z \in \mathbb{C}, \quad (5.141)$$

wobei z eine komplexe Zahl ist. Für jene $z \in \mathbb{C}$, für die der Operator $A - z$ invertierbar ist, heißt sein Inverses die **Resolvente** $R(z) = (A - z)^{-1}$ des Operators A an der Stelle z . Mit ihrer Hilfe lässt sich die Lösung der Gleichung angeben:

$$u = (A - z)^{-1}f = R(z)f. \quad (5.142)$$

Wir sind insbesondere an jenen Punkten z in der komplexen Ebene interessiert, für die die Resolvente *nicht* existiert. Das sind die **singulären Punkte**, bei denen es irgendein Problem mit der Inversion gab. Dieser Fall tritt sicher ein, wenn z ein Eigenwert von A ist. Im endlichdimensionalen Fall ist das auch die einzige Möglichkeit für singuläre Punkte. Wie wir gleich sehen werden, gibt es im unendlichdimensionalen Fall noch weitere Möglichkeiten, also singuläre Punkte, die keine Eigenwerte sind. Die **regulären Punkte** in z sind dagegen jene, für die die Resolvente $R(z)$ existiert.

Zur Klassifikation der regulären und singulären Punkte wenden wir unser früheres Schema aus Fig. 5.4 auf den Operator $A - z$ an. Um die Figur zu vereinfachen, kontrahieren wir die ersten beiden Zeilen, denn diese erfüllen beide die Bedingung $\overline{\mathcal{R}_{A-z}} = \mathcal{H}$. Je nach dem Wert von $z \in \mathbb{C}$ fällt der Operator $A - z$ dann in eine der Klassen in Fig. 5.5. Wir werden außerdem unseren Sprachgebrauch dahingehend anpassen, dass wir den Wertebereich \mathcal{R}_{A-z} immer als eingebettet in den ganzen Hilbertraum betrachten, dann kommt nämlich auch der Begriff der Surjektivität ins Spiel. Das ändert nichts an der Klassifikation in Fig. 5.5, nur an ihrer Nomenklatur: der Fall, dass $A - z$ nicht injektiv ist, entspricht wie früher (3ABC); der Fall, dass $A - z$ bijektiv ist und $R(z)$ existiert, entspricht (1A); und in allen anderen Fällen ist $A - z$ zwar injektiv, aber nicht surjektiv. Das ermöglicht einen einfacheren Vergleich mit der Literatur, wo die folgenden Begriffe oft nur für beschränkte Operatoren $A: \mathcal{H} \rightarrow \mathcal{H}$ (die auf dem ganzen Hilbertraum definiert sind) erklärt werden.

Dementsprechend trifft man die folgenden Definitionen:

- Die **Resolventenmenge** $\rho(A)$ ist die Menge aller regulären Punkte in $z \in \mathbb{C}$, an denen der Operator $A - z$ invertierbar ist:

$$\rho(A) := \{z \in \mathbb{C} \mid \mathcal{K}_{A-z} = \{0\}, \mathcal{R}_{A-z} = \mathcal{H}\}. \quad (5.143)$$

Das entspricht dem Fall (1A). Laut unseren früheren Überlegungen gilt in diesem Fall $\mathcal{R}_{A-z} = \mathcal{H}$, und der Operator ist nullbeschränkt: $\|(A - z)u\|/\|u\| > 0$ für alle $u \neq 0$, weswegen auch die Resolvente $R(z)$ beschränkt ist.

- Alle anderen Punkte z sind singulär und gehören zum **Spektrum** $\sigma(A)$. Das ist die Komplementärmenge der Resolventenmenge, an deren Stellen es mit der Resolvente irgendein Problem gibt:

$$\sigma(A) = \mathbb{C} \setminus \rho(A). \quad (5.144)$$

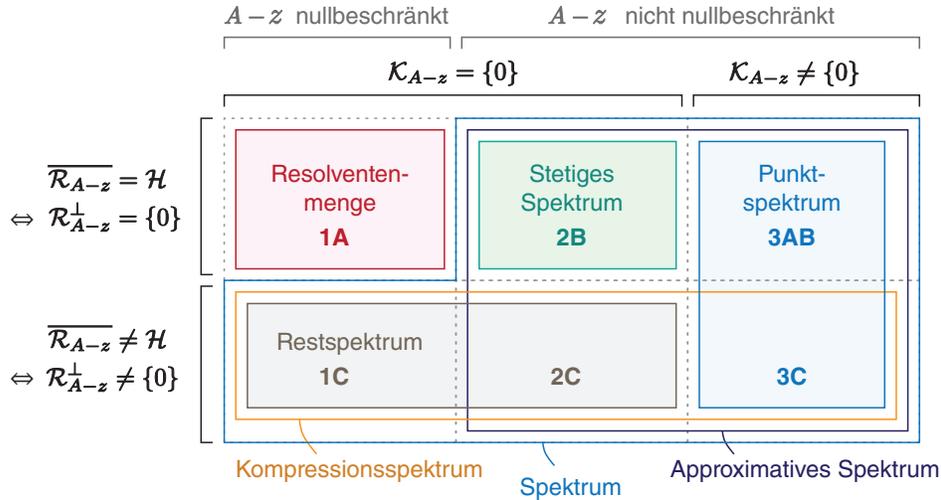


FIG. 5.5: Klassifikation des Spektrums eines Operators

Das Spektrum unterteilt man in drei disjunkte Teilbereiche $\sigma(A) = \sigma_p(A) \cup \sigma_c(A) \cup \sigma_r(A)$:

- Das **Punktspektrum** $\sigma_p(A)$ besteht aus den **Eigenwerten** $z = \lambda_i$ des Operators A . An diesen Stellen ist der Operator $A - z$ nicht injektiv, d.h. er hat einen nichttrivialen Nullraum $\mathcal{K}_{A-z} \neq \{0\}$:

$$\sigma_p(A) := \{z \in \mathbb{C} \mid \mathcal{K}_{A-z} \neq \{0\}\}. \quad (5.145)$$

Das entspricht den Fällen (3ABC). Jede nichttriviale Lösung der Gleichung $(A - z)u = 0$ heißt **Eigenvektor** von A . Die Dimension des Eigenraums \mathcal{K}_{A-z} für ein gegebenes z ist die **geometrische Vielfachheit** (*multiplicity*). Ist diese Dimension größer als 1, wird der Eigenwert **entartet** genannt, ansonsten nichtentartet.

- Für die restlichen Teilbereiche ist $A - z$ zwar injektiv, aber nicht surjektiv, d.h. der Wertebereich deckt nicht den gesamten Hilbertraum ab: $\mathcal{R}_{A-z} \neq \mathcal{H}$. Hier trifft man folgende Unterscheidungen: Das **stetige** oder **kontinuierliche Spektrum** $\sigma_c(A)$ besteht aus jenen Punkten $z \in \mathbb{C}$, für die der Wertebereich von $A - z$ dicht in \mathcal{H} liegt:

$$\sigma_c(A) := \{z \in \mathbb{C} \mid \mathcal{K}_{A-z} = \{0\}, \mathcal{R}_{A-z} \neq \mathcal{H}, \overline{\mathcal{R}_{A-z}} = \mathcal{H}\}. \quad (5.146)$$

Das entspricht dem Fall (2B). Daneben wird auch oft das **approximative Spektrum** als Vereinigung der Fälle (2) und (3) definiert; hier ist das gemeinsame Charakteristikum, dass $A - z$ nicht nullbeschränkt ist, d.h. es gibt eine Folge (u_n) mit $u_n \neq 0$, sodass $\|(A - z)u_n\| \rightarrow 0$ (dann ist z ein "approximativer" Eigenwert).

- Das **Restspektrum** $\sigma_r(A)$ besteht aus jenen Punkten $z \in \mathbb{C}$, für die Wertebereich von $A - z$ nicht dicht in \mathcal{H} liegt (er wurde "komprimiert"):

$$\sigma_r(A) := \{z \in \mathbb{C} \mid \mathcal{K}_{A-z} = \{0\}, \overline{\mathcal{R}_{A-z}} \neq \mathcal{H}\}. \quad (5.147)$$

Das entspricht den Fällen (1C, 2C). In diesem Fall ist die Dimension von \mathcal{R}_{A-z}^\perp ungleich Null und wird als **Defekt** (*deficiency*) bezeichnet. Das **Kompressionsspektrum** ist durch $\overline{\mathcal{R}_{A-z}} \neq \mathcal{H}$ ohne Ausschluss des Punktspektrums charakterisiert.

Für **selbstadjungierte Operatoren** lassen sich nun einige wichtige Aussagen treffen:

SATZ:

- Falls A **symmetrisch** ist, ist das approximative Spektrum reell und Eigenvektoren zu verschiedenen Eigenwerten sind orthogonal.
- Falls A **selbstadjungiert** ist, dann liegt jedes z mit $\operatorname{Im} z \neq 0$ in der Resolventenmenge. Das (reelle) Punktspektrum fällt mit dem Kompressionsspektrum zusammen, und die Vielfachheit jedes Eigenwerts ist identisch mit seinem Defekt.

Wir beweisen diese Aussagen nacheinander:

- Das Punktspektrum eines symmetrischen Operator ist reell, denn wir wissen ja bereits, dass für einen symmetrischen Operator der Erwartungswert $\langle u, Au \rangle$ reell ist. Mit $Au = \lambda u$ gilt $\langle u, Au \rangle = \lambda \|u\|^2 \in \mathbb{R}$ und somit $\lambda \in \mathbb{R}$ für $\|u\| \neq 0$.

- Allgemeiner gilt das auch fürs approximative Spektrum: Setze $z = x + iy$ mit $x, y \in \mathbb{R}$, dann ist

$$\|(A-z)u\|^2 = \langle (A-x)u - iyu, (A-x)u - iyu \rangle = \|(A-x)u\|^2 + y^2 \|u\|^2 + iy \langle u, (A-x)u \rangle - iy \langle (A-x)u, u \rangle.$$

Verwenden wir im letzten Term die Symmetrie von A und den Umstand, dass x reell ist, heben sich die Kreuzterme weg und wir finden $\|(A-z)u\|^2 \geq y^2 \|u\|^2$. Falls also $\operatorname{Im} z = y \neq 0$, dann ist der Operator $A - z$ nullbeschränkt, d.h. er muss vom Typ (1A) oder (1C) sein. Im Umkehrschluss: Falls z im approximativen Spektrum liegt, muss $\operatorname{Im} z = 0$ gelten, also $z \in \mathbb{R}$.

- Um zu zeigen, dass Eigenvektoren zu verschiedenen Eigenwerten orthogonal sind, setzen wir $Au_1 = \lambda_1 u_1$ und $Au_2 = \lambda_2 u_2$ mit $\lambda_1 \neq \lambda_2$. Dann ist $\lambda_1, \lambda_2 \in \mathbb{R}$ und es gilt

$$0 = \langle u_1, Au_2 \rangle - \langle u_1, Au_2 \rangle = \langle u_1, Au_2 \rangle - \langle Au_1, u_2 \rangle = (\lambda_2 - \lambda_1^*) \langle u_1, u_2 \rangle = (\lambda_2 - \lambda_1) \langle u_1, u_2 \rangle. \quad (5.148)$$

Wegen $\lambda_1 \neq \lambda_2$ muss $\langle u_1, u_2 \rangle = 0$ sein, d.h. die Eigenvektoren u_1 und u_2 sind orthogonal.

- Ein symmetrischer Operator $A - z$ mit $\operatorname{Im} z \neq 0$ ist also vom Typ (1A) oder (1C). Ist der Operator zusätzlich selbstadjungiert, gilt $A = A^\dagger$. Wir wenden jetzt die frühere Aussage (5.70) auf $A - z$ an:

$$\mathcal{K}_{A-z^*} = \mathcal{R}_{A-z}^\perp. \quad (5.149)$$

Wir haben sie zwar nur im Kontext beschränkter Operatoren bewiesen, sie gilt aber allgemein, solange A dicht definiert ist. Falls $A - z$ vom Typ (1C) ist, dann gilt $\mathcal{R}_{A-z}^\perp \neq \{0\}$, und somit muss z^* mit $\operatorname{Im} z \neq 0$ ein Eigenwert von A sein. Das ist aber ein Widerspruch, denn ein selbstadjungierter Operator hat ja nur reelle Eigenwerte. Folglich muss jedes z mit $\operatorname{Im} z \neq 0$ in der Resolventenmenge liegen.

- Für reelle z gilt $\mathcal{K}_{A-z} = \mathcal{R}_{A-z}^\perp$. Daraus folgt: Falls $\mathcal{R}_{A-z}^\perp = \{0\}$ (Typ AB), muss auch $\mathcal{K}_{A-z} = \{0\}$ sein (Typ 1 oder 2); das schließt den Fall (3AB) aus. Falls $\mathcal{R}_{A-z}^\perp \neq \{0\}$ (Typ C), muss auch $\mathcal{K}_{A-z} \neq \{0\}$ sein (Typ 3); das schließt die Fälle (1C) und (2C) aus. Im Fall (3C) ist jedes $v \in \mathcal{R}_{A-z}^\perp$ gleichzeitig ein Eigenvektor von A mit Eigenwert z , und die Vielfachheit ist identisch mit dem Defekt:

$$\dim \mathcal{K}_{A-z} = \dim \mathcal{R}_{A-z}^\perp. \quad (5.150)$$

Während für einen symmetrischen Operator nur das approximative Spektrum reell ist, ist für einen selbstadjungierten Operator also das gesamte Spektrum reell: $\sigma(A) \subset \mathbb{R}$. Das ist in der Physik von enormer Bedeutung, denn physikalische Observablen werden durch selbstadjungierte Operatoren beschrieben und deren Spektren entsprechen den Messwerten, die immer reell sein müssen. Für einen selbstadjungierten Operator liegt z entweder im Punktspektrum (3C), das den gebundenen Zuständen entspricht, im stetigen Spektrum (2B), das den freien Zuständen entspricht, oder in der Resolventenmenge (1A). Das Restspektrum, für das es keine physikalische Interpretation gibt, ist dagegen leer: $\sigma_r(A) = \emptyset$.

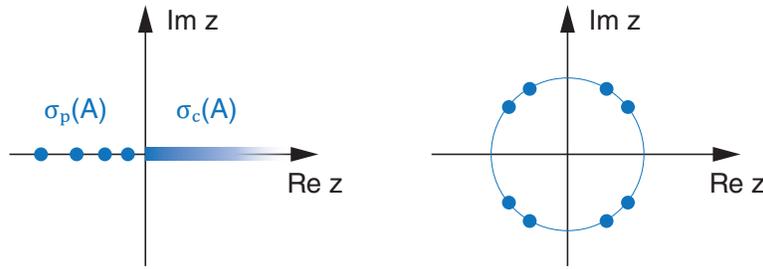


FIG. 5.6: Beispiele für das rein reelle Spektrum eines selbstadjungierten Operators (links) und das Spektrum eines unitären Operators auf dem Einheitskreis (rechts)

Ein Spezialfall beschränkter Operatoren sind **unitäre Operatoren**, die wir schon früher besprochen haben. Eine unitäre Transformation

$$B = UAU^{-1}, \quad \mathcal{D}_B = \{x \in \mathcal{H} \mid U^{-1}x \in \mathcal{D}_A\} \quad (5.151)$$

lässt das Spektrum des Operators A invariant: Für $z \in \rho(A)$ ist $A - z$ invertierbar, doch wegen $B - z = U(A - z)U^{-1}$ und der Invertierbarkeit von U ist dann auch $B - z$ invertierbar. Somit bleibt die Resolventenmenge und folglich auch das Spektrum unter der unitären Transformation invariant. Man kann zeigen, dass die unitäre Transformation auch die Teilspektren $\sigma_p(A)$, $\sigma_c(A)$ und $\sigma_r(A)$ getrennt invariant lässt. Für einen unitären Operator U selbst gilt:

SATZ: Das Spektrum eines unitären Operators U liegt auf dem Einheitskreis $|z| = 1$.

Dazu müssen wir nur zeigen, dass $U - z$ für $|z| < 1$ oder $|z| > 1$ invertierbar ist, sodass z in der Resolventenmenge liegt. Das gelingt am schnellsten über die Neumannsche Reihe. Für $|z| < 1$ ist $U - z = U(\mathbb{1} - zU^{-1})$ laut (5.56) invertierbar, da U selbst invertierbar ist und $\|zU^{-1}\| = |z|\|U^{-1}\| = |z| < 1$ gilt. Für $|z| > 1$ ist $U - z = -z(\mathbb{1} - U/z)$ invertierbar, da $\|U/z\| = 1/|z| < 1$. Daraus folgt, dass das Spektrum auf dem Einheitskreis $|z| = 1$ liegen muss.

Die Resolvente ist für $z \in \rho(A)$ ein linearer beschränkter Operator. Darüberhinaus ist sie auch eine holomorphe Funktion von z in der komplexen Ebene, daher kann man den Apparat der Funktionentheorie darauf anwenden. Beispielsweise folgt aus $A - z = A - z_0 - (z - z_0)$ die **1. Resolventenidentität**:

$$R(z)^{-1} = R(z_0)^{-1} - (z - z_0) \Leftrightarrow R(z) = R(z_0) + (z - z_0)R(z_0)R(z). \quad (5.152)$$

Hier haben wir von links mit $R(z_0)$ multipliziert und von rechts mit $R(z)$. Durch Iteration erhält man $R(z) = R(z_0) + (z - z_0)R(z_0)^2 + (z - z_0)^2R(z_0)^3 + \dots$. Falls $|z - z_0|\|R(z_0)\| < 1$, erhält man daraus die entsprechende Neumannsche Reihe:

$$R(z) = \sum_{n=0}^{\infty} (z - z_0)^n R(z_0)^{n+1} \quad \forall z, z_0 \in \rho(A). \quad (5.153)$$

Auf solche Identitäten stößt man in der Streutheorie, der Statistischen Physik oder im Elektromagnetismus häufig. Für $H = H_0 + V$ gilt auch $H - z = H_0 - z + V$ und somit $R^{-1} = R_0^{-1} + V$. Multipliziert man mit R_0 von links und R von rechts, erhält man die **2. Resolventenidentität** $R = R_0 - R_0 V R$. Solche Gleichungen sind in verschiedenen Kontexten auch als **Lippmann-Schwinger-** und **Bethe-Salpeter-Gleichungen** bekannt.

Beispiele. Wir wollen uns nun eine Reihe von Beispielen für Operatoren und deren Spektren ansehen, um das bisher Gesagte zu konkretisieren. Wir beginnen mit solchen Operatoren, die ein reines **Punktspektrum** besitzen:

- Der **Nulloperator** $A = 0$ hat als einzigen Punkt im Spektrum $z = 0$. In diesem Fall gilt $\mathcal{K}_A = \mathcal{H}$ und $\mathcal{R}_A = \{0\}$, also $\mathcal{R}_A^\perp = \mathcal{H}$. Damit ist $A - z$ vom Typ (3C). Wegen $\dim \mathcal{R}_A^\perp = \infty$ ist der Eigenwert $z = 0$ unendlichfach entartet. Für alle anderen Werte $z \neq 0$ ist $A - z$ invertierbar und die Resolvente lautet $R(z) = -1/z$.
- Der **Einsoperator** $A = \mathbb{1}$ hat als einzigen Punkt im Spektrum $z = 1$. Dieser ist unendlichfach entartet, denn jeder Vektor ist ein Eigenvektor. Auch hier ist $\mathcal{K}_{A-1} = \mathcal{H}$, $\mathcal{R}_{A-1} = \{0\}$, $\mathcal{R}_A^\perp = \mathcal{H}$, und der Eigenwert $z = 1$ ist unendlichfach entartet. Für alle anderen Werte $z \neq 1$ ist $A - z$ invertierbar und die Resolvente lautet $R(z) = 1/(1 - z)$.
- Betrachten wir den **unitären Operator** $A: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ aus (5.13), dessen Matrixdarstellung einer zweidimensionalen Drehmatrix entspricht:

$$\begin{aligned} A\varphi_1 &= (\cos \theta) \varphi_1 + (\sin \theta) \varphi_2, \\ A\varphi_2 &= (-\sin \theta) \varphi_1 + (\cos \theta) \varphi_2, \end{aligned} \quad \mathbf{A} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}. \quad (5.154)$$

Die Eigenwertgleichung $(\mathbf{A} - \lambda) \mathbf{u} = 0$ ergibt

$$(\cos \theta - \lambda)^2 + \sin^2 \theta = 0 \quad \Rightarrow \quad \lambda^2 - 2\lambda \cos \theta + 1 = (\lambda - e^{i\theta})(\lambda - e^{-i\theta}) = 0 \quad (5.155)$$

und somit $\lambda_\pm = e^{\pm i\theta}$, mit den beiden Eigenvektoren $(1, i)$ und $(1, -i)$. Wie erwartet liegen die Eigenwerte auf dem Einheitskreis, da der Operator unitär ist. Für alle anderen Punkte $z \neq \lambda_\pm$ ist die Gleichung $(\mathbf{A} - z) \mathbf{u} = \mathbf{f}$ invertierbar und liefert

$$\mathbf{u} = R(z) \mathbf{f}, \quad R(z) = (\mathbf{A} - z)^{-1}. \quad (5.156)$$

- Sei P der orthogonale **Projektor** auf einen abgeschlossenen Untervektorraum $\mathcal{M} \subset \mathcal{H}$, sodass $P^2 = P = P^\dagger$. Aus der Eigenwertgleichung $Pu = \lambda u$ folgt $\lambda^2 = \lambda$ mit den zwei Eigenwerten $\lambda = 0$ und $\lambda = 1$. Für den Eigenwert $\lambda = 0$ besteht der Nullraum \mathcal{K}_P der Gleichung $Pu = 0$ aus allen Vektoren $u \in \mathcal{M}^\perp$, d.h. der Eigenwert ist $(\dim \mathcal{M}^\perp)$ -fach entartet. Für $\lambda = 1$ setzt sich der Nullraum \mathcal{K}_{P-1} von $Pu = u$ aus allen Vektoren $u \in \mathcal{M}$ zusammen und die Vielfachheit des Eigenwertes ist $\dim \mathcal{M}$. Alle anderen Werte $z \neq \lambda$ liegen in der Resolventenmenge, d.h. die Gleichung $(P - z)u = f$ lässt sich invertieren. Wie lautet $(P - z)^{-1}$? Da $(\mathbb{1} - P)$ ebenfalls ein Projektor ist, gilt

$$H := \alpha P + \beta (\mathbb{1} - P) \quad \Leftrightarrow \quad H^{-1} = \frac{P}{\alpha} + \frac{\mathbb{1} - P}{\beta} \quad \alpha \neq 0, \quad \beta \neq 0, \quad (5.157)$$

wie man durch Einsetzen in $HH^{-1} = \mathbb{1}$ mittels $P^2 = P$, $(\mathbb{1} - P)^2 = \mathbb{1} - P$ und $P(\mathbb{1} - P) = 0$ sieht. Für $H = P - z$ findet man durch Koeffizientenvergleich $\alpha = 1 - z$ und $\beta = -z$, und somit lautet die Lösung

$$u = R(z) f, \quad R(z) = (P - z)^{-1} = \frac{P}{1 - z} - \frac{\mathbb{1} - P}{z}. \quad (5.158)$$

Interessanter wird es, wenn wir uns auch Operatoren mit einem **stetigen Spektrum** ansehen:

- Betrachten wir den **Ortsoperator** A mit $Au(x) = xu(x)$ auf dem Raum $L^2(0, 1)$. Dieser Operator ist selbstadjungiert: $A^\dagger = A$ (was auch aus der Diskussion von Multiplikationsoperatoren um (5.133) folgt) und kann somit bestenfalls ein Punktspektrum und ein stetiges Spektrum besitzen. Das Punktspektrum ergibt sich aus der Eigenwertgleichung $xu(x) = \lambda u(x)$, die für alle x erfüllt sein muss. Das erlaubt aber nur die Nulllösung $u = 0$, d.h. der Operator hat keine Eigenwerte. Liegen dann alle $z \in \mathbb{C}$ in der Resolventenmenge? Die Lösung

$$u = R(z) f, \quad R(z) = \frac{1}{x - z} \quad (5.159)$$

ist nur für $z \notin [0, 1]$ wohldefiniert. Für $z \in [0, 1]$ liegt u nur dann in $L^2(0, 1)$, wenn $f(x)$ für $x = z$ hinreichend schnell verschwindet. Oder anders gesagt: der Wertebereich von $x - z$ ist nicht mehr der gesamte $L^2(0, 1)$, liegt aber dicht. Alle Punkte $z \in [0, 1]$ liegen daher im stetigen Spektrum, und alle anderen Punkte in der Resolventenmenge.

- Sehen wir uns den **Hilbert-Schmidt-Operator** $A: L^2(0, 1) \rightarrow L^2(0, 1)$ aus (5.48) an:

$$Au(x) = \int_0^x dy u(y) = \int_0^1 dy \Theta(x - y) u(y). \quad (5.160)$$

Die Eigenwertgleichung $Au = \lambda u$ lautet $\int_0^x dy u(y) = \lambda u(x)$. Für $\lambda = 0$ erlaubt das nur die Nulllösung $u = 0$. Für $\lambda \neq 0$ liefert die Ableitung $u(x) = \lambda u'(x) \Rightarrow u(x) = c e^{x/\lambda}$, doch aus der Eigenwertgleichung folgt die Bedingung $u(0) = 0$, was wieder auf die Nulllösung zurückführt. Der Operator A hat also keine Eigenwerte. Wie lautet dann das Inverse von $A - z$? Wir haben

$$(A - z)u = f \Leftrightarrow \int_0^x dy u(y) = \underbrace{zu(x) + f(x)}_{=: w(x)} \Rightarrow u(x) = zu'(x) + f'(x) = w'(x),$$

was wegen $u = (w - f)/z$ für $z \neq 0$ auf eine inhomogene lineare Differenzialgleichung 1. Ordnung für w führt, die gelöst werden kann:

$$w'(x) = \frac{w(x) - f(x)}{z}, \quad w(0) = 0 \Rightarrow w(x) = -\frac{1}{z} \int_0^x dy f(y) e^{(x-y)/z}. \quad (5.161)$$

Die Lösung für u ist dann für alle $z \neq 0$ wohldefiniert und lautet

$$u(x) = R(z) f(x) = \frac{w(x) - f(x)}{z} = -\frac{f(x)}{z} - \frac{1}{z^2} \int_0^x dy f(y) e^{(x-y)/z}, \quad (5.162)$$

und somit liegen alle $z \neq 0$ in der Resolventenmenge. Was passiert für $z = 0$? In diesem Fall hat $Au = f$ die Lösung $u(x) = f'(x)$, sofern f differenzierbar ist und $f(0) = 0$ gilt. Diese Menge liegt dicht in $L^2(0, 1)$, d.h. es ist $\mathcal{R}_A^\perp = \{0\}$. Somit gehört der Punkt $z = 0$ zum stetigen Spektrum.

Noch interessanter wird es für Operatoren, die auch ein **Restspektrum** aufweisen:

- Betrachten wir den **Rechtsshift-Operator** aus (5.34) mit $A\varphi_i = \varphi_{i+1}$, wobei $\{\varphi_i\}$ eine ONB bildet. Die Eigenwertgleichung lautet

$$(\mathbf{A} - \lambda) \mathbf{u} = \begin{pmatrix} -\lambda & & & \\ 1 & -\lambda & & \\ & 1 & -\lambda & \\ & & 1 & -\lambda \\ & & & \ddots \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \\ \vdots \end{pmatrix} = 0 \quad \Rightarrow \quad \begin{array}{l} \lambda u_1 = 0, \\ \lambda u_2 = u_1, \\ \lambda u_3 = u_2, \\ \vdots \end{array} \quad (5.163)$$

Für $\lambda \neq 0$ folgt $u_k = 0$ für alle k , und für $\lambda = 0$ ebenso. A hat also keine Eigenwerte. Wie lautet die Lösung der inhomogenen Gleichung $(A - z)u = f$? Es ist

$$-zu_1 = f_1, \quad u_1 - zu_2 = f_2, \quad u_2 - zu_3 = f_3, \dots \quad (5.164)$$

Falls $z = 0$ und $f_1 = 0$, lautet die Lösung $u_1 = f_2, u_2 = f_3$, usw. Somit gilt $\mathcal{R}_A^\perp \neq \{0\}$, und weil der Punkt $z = 0$ kein Eigenwert ist, muss er im Restspektrum liegen. Der Defekt ist $\dim \mathcal{R}_A^\perp = 1$. Für $z \neq 0$ ist die Lösung gegeben durch

$$u_1 = -\frac{f_1}{z}, \quad u_2 = \frac{u_1 - f_2}{z}, \quad \dots, \quad u_n = \frac{u_{n-1} - f_n}{z}, \dots \quad (5.165)$$

Das bedeutet aber noch nicht, dass jeder Punkt $z \neq 0$ in der Resolventenmenge liegt! Denn es könnte ja irgendwelche Werte von z geben, für die der Hilbertraumvektor \mathbf{u} nicht konvergiert, d.h. $\sum_k |u_k|^2 = \infty$. Dazu behilft man sich mit dem adjungierten Operator A^\dagger , was hier der **Linksshift-Operator** $A^\dagger \varphi_i = \varphi_{i-1}$ ist. Laut (5.70) gilt

$$\mathcal{K}_{A-z} = (\mathcal{R}_{A^\dagger - z^*})^\perp \quad \text{und} \quad \mathcal{K}_{A^\dagger - z^*} = (\mathcal{R}_{A-z})^\perp. \quad (5.166)$$

Nachdem A keine Eigenwerte hat, ist $\mathcal{K}_{A-z} = \{0\}$ und somit kann A^\dagger kein Restspektrum besitzen. Die Eigenwertgleichung für A^\dagger lautet

$$(\mathbf{A}^\dagger - \lambda) \mathbf{u} = \begin{pmatrix} -\lambda & 1 & & \\ & -\lambda & 1 & \\ & & -\lambda & 1 \\ & & & -\lambda \\ & & & & \ddots \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \\ \vdots \end{pmatrix} = 0 \quad \Rightarrow \quad \begin{array}{l} u_2 = \lambda u_1, \\ u_3 = \lambda u_2, \\ u_4 = \lambda u_3, \\ \vdots \end{array} \quad (5.167)$$

Der resultierende Eigenvektor $\mathbf{u} \sim (1, \lambda, \lambda^2, \lambda^3, \dots)$ ist ein Hilbertraumvektor genau dann, wenn $|\lambda| < 1$. Das Punktspektrum von A^\dagger besteht also aus allen Werten von z innerhalb des Einheitskreises, was laut (5.166) identisch mit dem Restspektrum von A ist. Man kann zeigen, dass das Spektrum $\sigma(A)$ eine abgeschlossene Menge ist, weswegen auch der Einheitskreis $|z| = 1$ selbst im Spektrum liegen muss – da er aber nicht im Punktspektrum von A^\dagger liegt (und somit auch nicht im Restspektrum von A), muss er ein Teil des stetigen Spektrums sein. In Summe hat $\sigma(A)$ also eine reichhaltige Struktur: alle Punkte $|z| < 1$ liegen im Restspektrum und alle Punkte $|z| = 1$ im stetigen Spektrum. Hier sieht man auch deutlich den Unterschied zum endlichdimensionalen Fall: dort hat jede der Matrizen \mathbf{A} und \mathbf{A}^\dagger den Eigenwert $\lambda = 0$ als einzigen Punkt im Spektrum, und somit ist $\mathbf{A} - z$ bis auf $z = 0$ immer invertierbar.

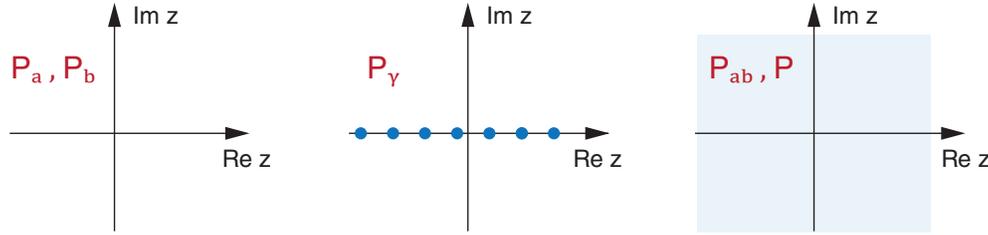


FIG. 5.7: Spektren verschiedener Impulsoperatoren

Sehen wir uns abschließend noch die aus (5.119) konstruierten **Impulsoperatoren** P , P_a , P_b , P_{ab} und P_γ im Hilbertraum $\mathcal{H} = L^2(a, b)$ an. Die Operatorvorschrift ist überall dieselbe, $u \mapsto -i du/dx$, aber die Definitionsbereiche sind verschieden:

$$\begin{aligned}
 \mathcal{D}_P &= \{u \in \mathcal{H} \mid u \text{ absolut stetig, } u' \in \mathcal{H}\}, \\
 \mathcal{D}_{P_a} &= \{u \in \mathcal{H} \mid u \text{ absolut stetig, } u' \in \mathcal{H}, u(a) = 0\}, \\
 \mathcal{D}_{P_b} &= \{u \in \mathcal{H} \mid u \text{ absolut stetig, } u' \in \mathcal{H}, u(b) = 0\}, \\
 \mathcal{D}_{P_{ab}} &= \{u \in \mathcal{H} \mid u \text{ absolut stetig, } u' \in \mathcal{H}, u(a) = u(b) = 0\}, \\
 \mathcal{D}_{P_\gamma} &= \{u \in \mathcal{H} \mid u \text{ absolut stetig, } u' \in \mathcal{H}, u(a) = e^{i\gamma} u(b), \gamma \in \mathbb{R}\}.
 \end{aligned} \tag{5.168}$$

Wir haben früher gesehen, dass P_{ab} symmetrisch ist und P_γ selbstadjungiert, und es gilt $P^\dagger = P_{ab}$ und $P_a^\dagger = P_b$. Welchen Einfluss haben die verschiedenen Definitionsbereiche aufs Spektrum? In allen Fällen lautet die Eigenwertgleichung $u'(x) = i\lambda u(x) \Rightarrow u(x) = c e^{i\lambda x}$, und die inhomogene Gleichung hat die formale Lösung

$$u'(x) = i\lambda u(x) + if(x) \quad \Rightarrow \quad u(x) = u(a) e^{iz(x-a)} + i \int_a^x dy f(y) e^{iz(x-y)}. \tag{5.169}$$

- Für den Operator P_a hat die Eigenwertgleichung wegen der Randbedingung $u(a) = 0$ nur die Nulllösung, daher hat P_a keine Eigenwerte und das Punktspektrum ist leer. In der inhomogenen Lösung bleibt wegen $u(a) = 0$ nur der Integralterm übrig, der die Resolvente $R(z) = (P_a - z)^{-1}$ über $u(x) = R(z) f(x)$ definiert. $R(z)$ ist ein Hilbert-Schmidt-Operator der Form (5.41) mit dem Integralkern $K(x, y) = i \Theta(x - y) e^{iz(x-y)}$, und $R(z)$ ist für alle $z \in \mathbb{C}$ beschränkt. Folglich ist das Spektrum *leer* und alle $z \in \mathbb{C}$ liegen in der Resolventenmenge.
- Die Eigenwertgleichung des Operators P_γ führt auf $e^{i\lambda a} = e^{i(\gamma + \lambda b)}$, somit lauten die Eigenwerte $\lambda = (\gamma + 2n\pi)/(a - b)$. Für alle anderen z ist die Resolvente wohldefiniert.
- Der Operator P_{ab} hat wieder keine Eigenwerte, doch hier führen die Randbedingungen $u(a) = u(b) = 0$ zu einer Einschränkung für f , nämlich $\int_a^b dy f(y) e^{iz(b-y)} = 0$ für jedes $z \in \mathbb{C}$. Somit ist $(\mathcal{R}_{P_{ab}-z})^\perp \neq \{0\}$ und diese Punkte liegen im Restspektrum. Das Spektrum von P_{ab} ist daher die gesamte komplexe Ebene! Für den adjungierten Operator $P_{ab}^\dagger = P$ ist das Punktspektrum ebenfalls ganz \mathbb{C} , da die Eigenwertgleichung keine Bedingung an λ liefert.

Das führt uns vor Augen, welchen drastischen Einfluss der Definitionsbereich eines Operators aufs Spektrum haben kann. Abbildungen mit derselben Operatorvorschrift, aber verschiedenen Definitionsbereichen sind auch verschiedene Operatoren!

Sturm-Liouville-Problem. Als Anwendungsbeispiele wollen wir jetzt allgemeine gewöhnliche **lineare Differenzialgleichungen 2. Ordnung** betrachten. Unter bestimmten Randbedingungen kann der entsprechende Operator immer selbstadjungiert gemacht werden:

SATZ: Seien A , B und C glatte, reellwertige Funktionen auf dem Intervall $[a, b] \subset \mathbb{R}$ und der Differenzialoperator L auf $C^2(a, b)$ gegeben durch

$$L = -A(x) \frac{d^2}{dx^2} - B(x) \frac{d}{dx} + C(x). \quad (5.170)$$

Dann ist L bezüglich des Skalarprodukts

$$\langle v, u \rangle_w = \int_a^b dx w(x) v^*(x) u(x) \quad (5.171)$$

mit einer Gewichtsfunktion w symmetrisch, sofern folgende Bedingungen erfüllt sind:

$$w(x) = \exp \left[\int^x dy \frac{B(y) - A'(y)}{A(y)} \right], \quad A(x) w(x) (v(x)^* u'(x) - v'(x)^* u(x)) \Big|_a^b = 0. \quad (5.172)$$

Zum Beweis betrachten wir den Integranden der Relation $\langle v, Lu \rangle - \langle Lv, u \rangle$:

$$\begin{aligned} w [v^* Lu - (Lv)^* u] &= w [v^* (-Au'' - Bu' + Cu) - (-Av'' - Bv' + Cv)^* u] \\ &= -Aw \underbrace{(v^* u'' - v''^* u)}_{=(v^* u' - v'^* u)'} - Bw (v^* u' - v'^* u) = -\frac{d}{dx} [Aw (v^* u - v'^* u)]. \end{aligned} \quad (5.173)$$

Im letzten Schritt haben wir (5.172) verwendet:

$$w'(x) = \frac{B(x) - A'(x)}{A(x)} w(x) \quad \Rightarrow \quad Bw = Aw' + A'w = (Aw)'. \quad (5.174)$$

Damit der Operator L symmetrisch wird, muss also die Randbedingung (5.172) erfüllt sein:

$$\langle v, Lu \rangle - \langle Lv, u \rangle = \int_a^b dx w [v^* Lu - (Lv)^* u] = -Aw (v^* u - v'^* u) \Big|_a^b. \quad (5.175)$$

Daraus folgt insbesondere, dass jeder Differenzialoperator L der obigen Form durch Multiplikation mit dem Gewicht w in die Standardform eines **Sturm-Liouville-Operators** gebracht werden kann, wenn wir $p = Aw \Rightarrow p' = Bw$ und $q = Cw$ setzen:

$$wL = -\underbrace{Aw}_p \frac{d^2}{dx^2} - \underbrace{Bw}_{p'} \frac{d}{dx} + \underbrace{wC}_q \quad \Rightarrow \quad L = \frac{1}{w} \left(-\frac{d}{dx} p \frac{d}{dx} + q \right), \quad (5.176)$$

wobei p , q und w reellwertige Funktionen mit $p \in C^1(a, b)$, $q, w \in C(a, b)$ und $p, w > 0$ sind. Die resultierende Eigenwertgleichung für L lautet

$$(L - \lambda) u = 0 \quad \Leftrightarrow \quad (pu')' + (-q + \lambda w) u = 0 \quad (5.177)$$

und muss folgende Randbedingung erfüllen:

$$p(x) (v(x)^* u'(x) - v'(x)^* u(x)) \Big|_a^b = 0. \quad (5.178)$$

Damit L zusätzlich **selbstadjungiert** wird, muss auch $\mathcal{D}_L = \mathcal{D}_{L^\dagger}$ erfüllt sein. Das lässt sich auf verschiedene Arten realisieren: gilt z.B. $p(a) = p(b) = 0$, dann verschwindet (5.178), solange u und v an den Randpunkten regulär sind. Ansonsten ist eine hinreichende Bedingung für die Selbstadjungiertheit

$$\alpha_1 u(a) + \alpha_2 u'(a) = 0, \quad \beta_1 u(b) + \beta_2 u'(b) = 0, \quad \alpha_i, \beta_i \in \mathbb{R}, \quad (5.179)$$

wobei weder α_1 und α_2 noch β_1 und β_2 gleichzeitig verschwinden dürfen. Falls $p(x)$ an den Rändern nämlich ungleich Null ist, lässt sich die Randbedingung (5.178) so schreiben:

$$\det \mathbf{M}(a) = \det \mathbf{M}(b), \quad \mathbf{M}(x) = \begin{pmatrix} v(x)^* & v'(x)^* \\ u(x) & u'(x) \end{pmatrix}, \quad (5.180)$$

wobei $\det \mathbf{M}$ die von u und v^* erzeugte **Wronski-Determinante** ist. Das ist sicher erfüllt, wenn $\det \mathbf{M}(a) = \det \mathbf{M}(b) = 0$ gilt. In diesem Fall ist $\mathbf{M}(a) \boldsymbol{\alpha} = \mathbf{M}(b) \boldsymbol{\beta} = 0$, was für $\boldsymbol{\alpha} \neq 0$ und $\boldsymbol{\beta} \neq 0$ das Resultat in (5.179) liefert. Somit genügen $u(a)$, $v(a)$ bzw. $u(b)$, $v(b)$ denselben Einschränkungen und es gilt $\mathcal{D}_L = \mathcal{D}_{L^\dagger}$. Zwei Spezialfälle von (5.179) sind:

- **Dirichlet-Randbedingungen:** $\alpha_2 = \beta_2 = 0$, d.h. $u(a) = u(b) = 0$,
- **Neumann-Randbedingungen:** $\alpha_1 = \beta_1 = 0$, d.h. $u'(a) = u'(b) = 0$.

Anstelle von (5.179) sind z.B. aber auch **periodische Randbedingungen** $u(a) = u(b)$, $u'(a) = u'(b)$ für die Selbstadjungiertheit von L hinreichend. Bevor wir uns konkrete Beispiele für Sturm-Liouville-Probleme ansehen, folgen noch einige Bemerkungen:

- Falls L bezüglich des Skalarprodukts $\langle \cdot, \cdot \rangle_w$ selbstadjungiert ist, liefert die Eigenwertgleichung $(L - \lambda)u = 0$ ein reelles Punktspektrum mit $\lambda \in \mathbb{R}$, und die Eigenvektoren u zu verschiedenen Eigenwerten sind orthogonal bezüglich $\langle \cdot, \cdot \rangle_w$.
- Ein **reguläres** Sturm-Liouville-Problem ist durch (5.177) auf einem beschränkten Intervall $[a, b]$ mit Randbedingungen (5.179) und reellwertigen Funktionen p, q, w mit $p, w > 0$ und $p, p', q, w \in C[a, b]$ gegeben. Die meisten der in Anwendungen auftretenden Fälle sind allerdings **singuläre** Probleme, die durch mindestens eine der folgenden Eigenschaften charakterisiert sind: das Intervall ist unbeschränkt; $p(x) > 0$ für alle $x \in (a, b)$ und p verschwindet an einem oder beiden der Randpunkte; das Gewicht $w(x)$ verschwindet an einigen Punkten; oder man ersetzt eine oder beide der Randbedingungen durch die Forderung, dass die Grenzwerte $u(a)$ und $u(b)$ bloß existieren.
- Als selbstadjungierter Operator kann L auch ein stetiges Spektrum haben. Die Lösung der Gleichung $Lu = \lambda u$ lässt sich laut (4.62) durch eine Greenfunktion darstellen, $u(x) = \lambda \int dy K(x - y) u(y)$, womit $1/\lambda$ gleichzeitig zum Eigenwert des Operators K wird. Für reguläre Probleme ist dieser Operator **kompakt**. Wie wir später sehen werden, ist das Spektrum eines kompakten Operators eine abzählbare Menge mit Null als einzigem möglichen Häufungspunkt. Deswegen muss auch das Spektrum von L abzählbar sein, mit einzigem Häufungspunkt $\lambda \rightarrow \infty$. Die im Folgenden besprochenen Probleme sind zwar größtenteils singulär, für sie gelten diese Aussagen aber noch immer.
- Die Eigenwerte λ und Eigenvektoren u der Gleichung $Lu = \lambda u$ lassen sich durch einen Potenzreihenansatz und anschließenden Koeffizientenvergleich bestimmen, siehe z.B. [Lan16]. Als Differenzialgleichung 2. Ordnung kann $Lu = \lambda u$ für jedes λ maximal zwei unabhängige Lösungen haben, sodass die Vielfachheit jedes Eigenwerts ≤ 2 ist.

	$p(x)$	$q(x)$	$w(x)$	a	b	Eigenwerte λ
Schwingungsgleichung	1	0	1			
Legendre-DG	$1 - x^2$	0	1	-1	1	$l(l+1), l \in \mathbb{N}_0$
verallg. Legendre-DG	$1 - x^2$	$m^2/(1 - x^2)$	1	-1	1	$l(l+1), m \leq l$
Hermite-DG	e^{-x^2}	0	e^{-x^2}	$-\infty$	∞	$2n, n \in \mathbb{N}_0$
Laguerre-DG	$x e^{-x}$	0	e^{-x}	0	∞	$n, n \in \mathbb{N}_0$
Tschebyschev-DG	$\sqrt{1 - x^2}$	0	$1/\sqrt{1 - x^2}$	-1	1	$n^2, n \in \mathbb{N}_0$
Bessel-DG	x	ν^2/x	x	0	1	$\omega_{\nu k}^2, k \in \mathbb{N}$

In der obigen Tabelle sind einige häufig auftretende Sturm-Liouville-Probleme zu Gl. (5.177) angeführt. Wir wollen sie etwas näher diskutieren:

- Die **Schwingungsgleichung** $u'' + \lambda u = 0$ entspricht dem Sturm-Liouville-Operator für $p(x) = w(x) = 1$ und $q(x) = 0$. Ihre Eigenfunktionen sind die in Kapitel 2 besprochenen Fourierpolynome, z.B. erhält man mit den Randbedingungen $u(0) = u(1) = 0$ die Eigenfunktionen $u_n(x) = \sin(n\pi x)$ mit $\lambda_n = (n\pi)^2$ und $n \in \mathbb{N}$.
- Die **Legendresche Differenzialgleichung** $(1 - x^2)u'' - 2xu' + \lambda u = 0$ entspricht dem Fall $p(x) = 1 - x^2$, $q(x) = 0$ und $w(x) = 1$. Wegen $p(\pm 1) = 0$ sind die Randbedingungen für $a = -1$ und $b = 1$ erfüllt, solange $u(\pm 1)$ existiert. Mit $w(x) = 1$ ergibt (5.171) dann das übliche Skalarprodukt in $L^2(-1, 1)$. Die Eigenwerte lauten $\lambda = l(l+1)$ mit $l \in \mathbb{N}_0$, und die Eigenfunktionen sind die **Legendre-Polynome** $P_l(x)$:

$$P_l(x) = \frac{1}{2^l l!} \frac{d^l}{dx^l} (x^2 - 1)^l \Rightarrow \begin{aligned} P_0(x) &= 1, \\ P_1(x) &= x, \\ P_2(x) &= \frac{1}{2} (3x^2 - 1), \\ P_3(x) &= \frac{1}{2} (5x^3 - 3x), \dots \end{aligned} \quad (5.181)$$

P_l ist ein Polynom vom Grad l , sodass die n -te Ableitung $(P_l)^{(n)}$ für $n > l$ verschwindet. Es gilt außerdem $P_l(-x) = (-1)^l P_l(x)$, $P_l(1) = 1$ und $P_l(-1) = (-1)^l$.

Um die **Rodrigues-Formel** (5.181) zu beweisen, zeigen wir, dass $u = [(x^2 - 1)^n]^{(n)}$ die Legendre-DG löst. Dazu machen wir Gebrauch von der Leibnizregel für höhere Ableitungen:

$$(fg)^{(n)} = \sum_{k=0}^n \binom{n}{k} f^{(k)} g^{(n-k)}. \quad (5.182)$$

Wir starten von der Relation $d/dx (x^2 - 1)^n = 2nx(x^2 - 1)^{n-1}$, also $(x^2 - 1) d/dx (x^2 - 1)^n = 2nx(x^2 - 1)^n$. Davon bilden wir die $(n+1)$ -te Ableitung und setzen beide Seiten in die Leibnizregel ein:

$$\sum_{k=0}^{n+1} \binom{n+1}{k} (x^2 - 1)^{(k)} \underbrace{[(x^2 - 1)^n]^{(n+2-k)}}_{=u^{(2-k)}} = 2n \sum_{k=0}^{n+1} \binom{n+1}{k} x^{(k)} \underbrace{[(x^2 - 1)^n]^{(n+1-k)}}_{=u^{(1-k)}}. \quad (5.183)$$

Auf der linken Seite gibt es nur Terme für $k \leq 2$ und auf der rechten Seite nur für $k = 0, 1$:

$$(x^2 - 1)u'' + 2x(n+1)u' + n(n+1)u = 2nxu' + 2n(n+1)u, \quad (5.184)$$

woraus sich die Legendre-DG $(x^2 - 1)u'' + 2xu' = n(n+1)u$ ergibt.

Die Polynome $\sqrt{l + \frac{1}{2}} P_l$ bilden eine Orthonormalbasis in $L^2(-1, 1)$:

$$\int_{-1}^1 dx P_l(x) P_{l'}(x) = \frac{2 \delta_{ll'}}{2l + 1}. \quad (5.185)$$

Da die lineare Hülle der Legendre-Polynome $\{P_0, \dots, P_n\}$ gleich jener der Monome $\{1, \dots, x^n\}$ ist, kann man erstere mittels Gram-Schmidt-Verfahren aus letzteren erzeugen, wie wir es früher in (1.72) auch schon getan haben. Die Entwicklung einer Funktion in Legendre-Polynome wird meist so geschrieben:

$$f(x) = \sum_{l=0}^{\infty} (2l + 1) c_l P_l(x) \quad \Leftrightarrow \quad c_l = \frac{1}{2} \int_{-1}^1 dx P_l(x) f(x). \quad (5.186)$$

Die Legendre-Polynome lassen sich durch eine **erzeugende Funktion** $\Phi(x, h)$ generieren. Dazu schreibt man

$$\Phi(x, h) = \sum_{l=0}^{\infty} h^l P_l(x) \quad \Leftrightarrow \quad P_n(x) = \frac{1}{n!} \frac{\partial^n}{\partial h^n} \Phi(x, h) \Big|_{h=0}, \quad (5.187)$$

womit sie zunächst einfach die Koeffizienten der Taylorreihe für $\Phi(x, h)$ sind. Die erzeugende Funktion lässt sich aber explizit angeben:

$$\Phi(x, h) = \frac{1}{\sqrt{1 - 2xh + h^2}}, \quad |h| < 1. \quad (5.188)$$

Aus den Eigenschaften von $\Phi(x, h)$ lassen sich durch Einsetzen der Taylorreihe und Koeffizientenvergleich verschiedene nützliche Rekursionsbeziehungen für die Legendre-Polynome herleiten, z.B.

$$\begin{aligned} (1 - 2xh + h^2) \frac{\partial \Phi}{\partial h} &= (x - h) \Phi \quad \Rightarrow \quad (n + 1) P_{n+1} = (2n + 1) x P_n - n P_{n-1}, \\ (x - h) \frac{\partial \Phi}{\partial x} &= h \frac{\partial \Phi}{\partial h} \quad \Rightarrow \quad x P'_n - P'_{n-1} = n P_n. \end{aligned}$$

- Die **verallgemeinerte Legendresche Differenzialgleichung** entspricht dem Fall $q(x) = m^2/(1 - x^2)$:

$$\frac{d}{dx} [(1 - x^2) u'] + \left(-\frac{m^2}{1 - x^2} + \lambda \right) u = 0. \quad (5.189)$$

Die Lösung $\lambda = l(l + 1)$ mit $l \in \mathbb{N}_0$ ist analog wie vorher. Darüberhinaus gibt es im Intervall $[-1, 1]$ nur dann nichtsinguläre Lösungen, wenn auch die Konstante m nur diskrete Werte annimmt, sodass $m \in \mathbb{Z}$ und $m = -l \dots l$. Die Eigenfunktionen sind die **zugeordneten Legendre-Polynome** P_l^m :

$$\begin{aligned} P_l^m(x) &= \frac{(-1)^m (1 - x^2)^{m/2}}{2^l l!} \frac{d^{l+m}}{dx^{l+m}} (x^2 - 1)^l, \\ \int_{-1}^1 dx P_l^m(x) P_{l'}^m(x) &= \frac{2}{2l + 1} \frac{(l + m)!}{(l - m)!} \delta_{ll'}. \end{aligned} \quad (5.190)$$

Die Legendre-Differenzialgleichung ergibt sich bei Problemen mit sphärischer Symmetrie, z.B. bei der Lösung der **Laplace-Gleichung** $\Delta u(r, \theta, \phi) = 0$ in Kugelkoordinaten. Diese lautet

$$\Delta u = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial u}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial u}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2 u}{\partial \phi^2} = 0. \quad (5.191)$$

Mit einem Separationsansatz $u(r, \theta, \phi) = R(r) S(\theta) T(\phi)$ wird das zu

$$\frac{RST}{r^2} \left[\underbrace{\frac{1}{R} \frac{\partial}{\partial r} \left(r^2 \frac{\partial R}{\partial r} \right)}_{=\alpha} + \underbrace{\frac{1}{\sin^2 \theta} \left(\frac{\sin \theta}{S} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial S}{\partial \theta} \right) + \frac{1}{T} \frac{\partial^2 T}{\partial \phi^2} \right)}_{=-\alpha} \right] = 0. \quad (5.192)$$

Für $u = RST \neq 0$ muss die Klammer verschwinden. Da der erste Teil nur von r abhängt und der zweite nur von θ und ϕ , müssen beide Teile gleich einer Konstanten sein und sich gegenseitig wegheben. Setzt man $R(r) = U(r)/r$, vereinfacht sich die Radialgleichung zu $r^2 U'' = \alpha U$, was durch einen Potenzreihenansatz gelöst werden kann. Der Winkelanteil lässt sich wiederum in zwei Teile separieren, die jeweils nur von θ bzw. ϕ abhängen und deswegen gleich einer Konstanten sein müssen:

$$\underbrace{\frac{\sin \theta}{S} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial S}{\partial \theta} \right)}_{=\beta} + \alpha \sin^2 \theta + \underbrace{\frac{1}{T} \frac{\partial^2 T}{\partial \phi^2}}_{=-\beta} = 0. \quad (5.193)$$

Mit der Substitution $x = \cos \theta$ bekommt man

$$\sin \theta \frac{\partial}{\partial \theta} = \sin \theta \frac{\partial x}{\partial \theta} \frac{\partial}{\partial x} = -\sin^2 \theta \frac{\partial}{\partial x} = -(1-x^2) \frac{\partial}{\partial x} \quad (5.194)$$

und somit wird die Gleichung für θ zu einer verallgemeinerten Legendre-DG:

$$\frac{d}{dx} [(1-x^2) S'] + \left(-\frac{\beta}{1-x^2} + \alpha \right) S = 0. \quad (5.195)$$

Von dieser wissen wir bereits, dass sie die Lösungen $\alpha = l(l+1)$ mit $l \in \mathbb{N}_0$ und $\beta = m^2$ mit $m = -l \dots l$ hat, und die Eigenlösungen $S(\theta)$ sind die verallgemeinerten Legendre-Polynome $P_l^m(\cos \theta)$. Damit lässt sich auch die verbleibende Gleichung für ϕ lösen: $T'' = -m^2 T \Rightarrow T(\phi) \sim e^{\pm im\phi}$. Der gesamte Winkelanteil $S(\theta)T(\phi)$ wird dann durch die **Kugelflächenfunktionen** (*spherical harmonics*) beschrieben:

$$Y_{lm}(\theta, \phi) = \sqrt{\frac{2l+1}{4\pi} \frac{(l-m)!}{(l+m)!}} P_l^m(\cos \theta) e^{im\phi}. \quad (5.196)$$

Diese bilden eine ONB auf der Kugelschale $d\Omega = d\theta \sin \theta d\phi$:

$$\int d\Omega Y_{l'm'}^*(\theta, \phi) Y_{lm}(\theta, \phi) = \delta_{ll'} \delta_{mm'}, \quad \int d\Omega = \int_0^\pi \underbrace{d\theta \sin \theta}_{-d \cos \theta} \int_0^{2\pi} d\phi = \int_{-1}^1 dx \int_0^{2\pi} d\phi = 4\pi. \quad (5.197)$$

- Die **Hermiteische Differenzialgleichung** $u'' - 2xu' + \lambda u = 0$ entspricht dem Sturm-Liouville-Problem mit $p(x) = w(x) = e^{-x^2}$ und $q(x) = 0$. Hier verschwindet $p(x)$ für $|x| \rightarrow \infty$, weswegen der Operator auf ganz \mathbb{R} selbstadjungiert ist, sofern $u(\pm\infty)$ existiert. Die Gleichung wird gelöst durch $\lambda = 2n$ mit $n \in \mathbb{N}_0$, und die Eigenfunktionen sind die **Hermite-Polynome** $H_n(x)$:

$$\begin{aligned} H_0(x) &= 1, \\ H_1(x) &= 2x, \\ H_2(x) &= 4x^2 - 2, \dots \end{aligned} \quad \int_{-\infty}^{\infty} dx e^{-x^2} H_n(x) H_m(x) = \sqrt{\pi} 2^n n! \delta_{nm}. \quad (5.198)$$

Sie treten bei der Lösung des quantenmechanischen harmonischen Oszillators auf. Auch hier gibt es eine erzeugende Funktion $\Phi(x, h) = \sum_{n=0}^{\infty} H_n(x) h^n / n! = e^{-h^2 + 2hx}$, woraus sich Rekursionsbeziehungen herleiten lassen:

$$\frac{\partial \Phi}{\partial x} = 2h\Phi \Rightarrow H'_n = 2nH_{n-1}, \quad \frac{\partial \Phi}{\partial h} = 2(x-h)\Phi \Rightarrow H_{n+1} = 2xH_n - H'_n.$$

- Die **Laguerresche Differenzialgleichung** $xu'' + (1-x)u' + \lambda u = 0$ entspricht dem Fall $p(x) = xe^{-x}$, $q(x) = 0$ und $w(x) = e^{-x}$. Hier verschwindet $p(x)$ für $x = 0$ und $x \rightarrow \infty$, woraus sich die Intervallgrenzen $a = 0$ und $b = \infty$ ergeben. Die Eigenfunktionen zu den Eigenwerten $\lambda = n$ mit $n \in \mathbb{N}_0$ sind die **Laguerre-Polynome** $L_n(x)$:

$$\begin{aligned} L_0(x) &= 1, \\ L_1(x) &= 1 - x, \\ L_2(x) &= 1 - 2x + \frac{x^2}{2}, \dots \end{aligned} \quad \int_0^{\infty} dx e^{-x} L_n(x) L_m(x) = \delta_{nm}. \quad (5.199)$$

Auch hier gibt es eine erzeugende Funktion, aus der sich Rekursionsrelationen herleiten lassen, siehe z.B. [Lan16].

- Die **Besselsche Differenzialgleichung** $x^2 u'' + xu' + (\lambda x^2 - \nu^2) u = 0$ entspricht dem Fall $p(x) = x$, $q(x) = \nu^2/x$, $w(x) = x$ mit einem Parameter $\nu \geq 0$. Für $\lambda = 1$ ist die bei $x = 0$ reguläre Lösung die **Besselfunktion 1. Art der Ordnung ν** :

$$J_\nu(x) = \sum_{n=0}^{\infty} \frac{(-1)^n}{\Gamma(n+1)\Gamma(n+\nu+1)} \left(\frac{x}{2}\right)^{2n+\nu}. \quad (5.200)$$

Setzt man $u(x) = J_\nu(x \omega_{\nu k})$ wieder in die Besselgleichung ein, wobei $\omega_{\nu k}$ die k -te Nullstelle von J_ν bezeichnet, erhält man wieder dieselbe Gleichung zurück, sofern $\lambda = \omega_{\nu k}^2$ erfüllt ist; somit ist das die Lösung für den Eigenwert. Außerdem folgt automatisch die Randbedingung $u(1) = J_\nu(\omega_{\nu k}) = 0$. Die Orthogonalitätsrelation lautet

$$\int_0^1 dx x J_\nu(x \omega_{\nu k}) J_\nu(x \omega_{\nu k'}) = \frac{1}{2} (J_{\nu+1}(\omega_{\nu k}))^2 \delta_{kk'}. \quad (5.201)$$

Besselfunktionen treten bei zylindrischen Problemen auf: löst man die Laplace-Gleichung in Zylinderkoordinaten mit dem Separationsansatz $u(\rho, z, \varphi) = R(\rho) Z(z) \Phi(\varphi)$, ergibt sich für $R(\rho)$ die Besselgleichung. Die zweite linear unabhängige Lösung ist durch die **Besselfunktionen 2. Art** gegeben, die auch **Neumann-** oder **Weber-Funktionen** genannt werden:

$$Y_\nu(x) = \lim_{n \rightarrow \nu} \frac{\cos(n\pi) J_n(x) - J_{-n}(x)}{\sin(n\pi)}. \quad (5.202)$$

Diese haben bei $x = 0$ eine logarithmische Singularität und einen Pol ν -ter Ordnung. Falls ν nicht ganzzahlig ist, dann stellt bereits $J_{-\nu}(x)$ die zweite unabhängige Lösung dar, während das für ganzzahliges ν wegen $J_{-\nu} = (-1)^\nu J_\nu$ nicht möglich ist. Anstatt der J_ν und Y_ν gibt man auch oft deren Linearkombinationen an, die **Hankelfunktionen** $H_\nu^\pm = J_\nu \pm iY_\nu$.

Zusammenfassend sehen wir also, dass man statt der Fourierreihe als Eigensystem der Schwingungsgleichung auch die Eigensysteme anderer Operatoren zur Konstruktion von Orthonormalbasen heranziehen kann. Das bringt Vorteile bezüglich der Konvergenz, die sich verbessert, wenn man das Basissystem des jeweiligen Operators wählt. Erfüllt eine Funktion f die Randbedingungen der Differenzialgleichung und ist sie auf dem Definitionsbereich $[a, b]$ stetig und stückweise stetig differenzierbar, dann konvergiert die Summe $f = \sum_n c_n \varphi_n$ mit $c_n = \int_a^b dx w(x) \varphi_n^*(x) f(x)$ sowohl punktweise als auch gleichmäßig. Erfüllt f die Randbedingungen nicht, dann konvergiert die Reihe noch immer im quadratischen Mittel bzw. an den Unstetigkeitsstellen gegen den Mittelwert der Funktionswerte.

Kompakte Operatoren. Eine spezielle Klasse beschränkter Operatoren sind kompakte Operatoren. Diese sind “fast endlichdimensional” in dem Sinn, dass sie sich in der Operatornorm durch endlichdimensionale Operatoren approximieren lassen. Die Untersuchung kompakter Operatoren war für die mathematische Entwicklung der Theorie der linearen Operatoren in unendlichdimensionalen Hilberträumen von großer Bedeutung. Kompakte Operatoren spiegeln viele Eigenschaften aus der endlichdimensionalen linearen Algebra wider, was sich in der einfachen Gestalt ihres Spektrums und in der einfachen Herleitung des Spektralsatzes für selbstadjungierte kompakte Operatoren zeigt.

Kompakte Operatoren sind dadurch definiert, dass sie *beschränkte* Mengen auf *relativ kompakte* Mengen abbilden. Zur Begriffsklärung wiederholen wir einige Konzepte aus der Analysis (von denen wir bisher ja auch schon reichlich Gebrauch gemacht haben):

- Eine Menge \mathcal{M} heißt **offen**, wenn für jedes $x \in \mathcal{M}$ jede hinreichend kleine Umgebung von x wieder in \mathcal{M} liegt. Das Standardbeispiel einer offenen Menge ist das Intervall $(0, 1) \subset \mathbb{R}$: jede reelle Zahl mit der Eigenschaft $0 < x < 1$ ist nur von Zahlen mit derselben Eigenschaft umgeben. Der Begriff der offenen Menge lässt sich auf verschiedenen Abstraktionsstufen definieren: im allgemeinsten Fall topologischer Räume mit Hilfe offener Kugeln, in metrischen Räumen mit Hilfe der Metrik oder in normierten Räumen durch die Norm. In einem normierten Raum ist eine offene Menge dadurch definiert, dass es für jedes $x \in \mathcal{M}$ ein $c > 0$ gibt, sodass alle Elemente $y \in \mathcal{M}$ mit $\|x - y\| < c$ wieder in \mathcal{M} liegen. Wir werden die folgenden Aussagen im Kontext eines Banachraums X treffen, d.h. eines vollständigen normierten Raums.
- Der **Abschluss** einer Menge $\mathcal{M} \subset X$ ist die Menge $\overline{\mathcal{M}}$, die aus den Grenzwerten aller Folgen besteht, die aus \mathcal{M} konstruiert werden können. $\overline{\mathcal{M}}$ enthält alle Elemente aus \mathcal{M} (da der Grenzwert der Folge x, x, x, \dots wieder x ist), kann aber auch Elemente aus X enthalten, die nicht in \mathcal{M} liegen. Eine Menge \mathcal{M} heißt **abgeschlossen** in X , falls $\overline{\mathcal{M}} = \mathcal{M}$; d.h. für jede Folge von Elementen aus \mathcal{M} , die in X konvergiert, liegt der Grenzwert in \mathcal{M} . Jede Menge aus endlich vielen Elementen ist abgeschlossen. Für $\mathcal{M} = (0, 1) \subset \mathbb{R}$ ist der Abschluss das Intervall $[0, 1]$. Eine Menge ist offen (abgeschlossen) genau dann, wenn ihr Komplement abgeschlossen (offen) ist. Es gibt Mengen, die weder offen noch abgeschlossen sind (z.B. $(0, 1] \subset \mathbb{R}$) und solche, die gleichzeitig offen und abgeschlossen sind (z.B. \mathbb{R} und die leere Menge \emptyset).
- Seien $\mathcal{A} \subset \mathcal{B}$ zwei Teilmengen eines Banachraums X . \mathcal{A} liegt **dicht** in \mathcal{B} , falls $\overline{\mathcal{A}} = \mathcal{B}$. Jedes Element aus \mathcal{B} kann dann beliebig genau durch Elemente aus \mathcal{A} approximiert werden, d.h. für jedes $b \in \mathcal{B}$ und jedes $\epsilon > 0$ gibt es ein $a \in \mathcal{A}$ sodass $\|a - b\| < \epsilon$. Falls \mathcal{A} dicht in \mathcal{B} und \mathcal{B} dicht in X liegt, dann liegt auch \mathcal{A} dicht in X . Beispielsweise liegen die rationalen Zahlen dicht in den reellen Zahlen; die Menge aller Polynome $P(a, b)$ liegt dicht im Raum der stetigen Funktionen $C(a, b)$ bezüglich der Maximumsnorm (Weierstraßscher Approximationssatz) sowie der L^2 -Norm; und die Menge der stetigen Funktionen liegt dicht in $L^2(a, b)$, weswegen auch die Polynome dicht in $L^2(a, b)$ liegen. Die Menge \mathbb{N} der natürlichen Zahlen liegt nicht dicht in den rationalen Zahlen.
- Eine Menge \mathcal{M} ist **beschränkt**, falls der Abstand zwischen allen Elementen der Menge von oben beschränkt ist, d.h. falls es eine Zahl c gibt, sodass $\|x - y\| \leq c \forall x, y \in \mathcal{M}$. Daraus folgt insbesondere auch $\|x\| \leq c < \infty \forall x, y \in \mathcal{M}$.

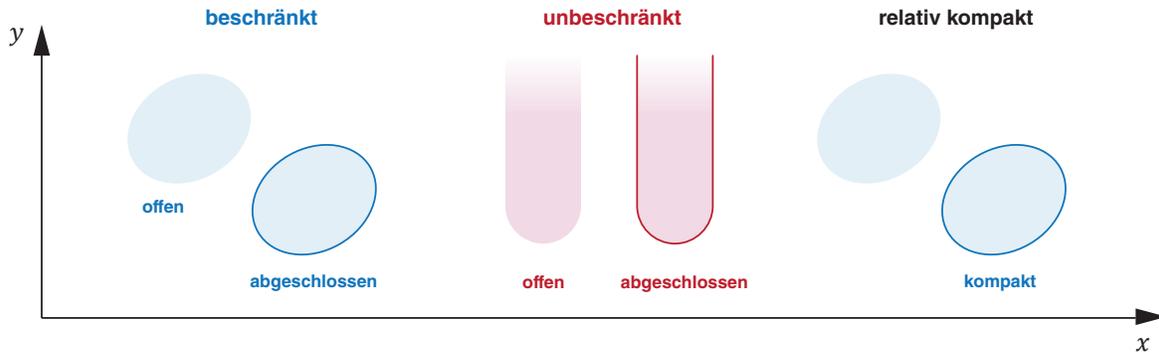


FIG. 5.8: Illustration der Begriffe offen/abgeschlossen, beschränkt/unbeschränkt, relativ kompakt und kompakt im \mathbb{R}^2 . Jede beschränkte Teilmenge des \mathbb{R}^n ist relativ kompakt.

- Eine Menge $\mathcal{M} \subset X$ heißt **relativ kompakt**, wenn jede unendliche Teilmenge aus \mathcal{M} eine Cauchyfolge enthält. Cauchyfolgen haben natürlich Grenzwerte in X , da wir X als vollständig vorausgesetzt haben, doch diese Grenzwerte müssen nicht in \mathcal{M} liegen. Falls \mathcal{M} endlich ist, gibt es keine unendliche Teilmenge verschiedener Elemente, deswegen nennen wir \mathcal{M} in diesem Fall ebenfalls relativ kompakt.
- Eine Menge $\mathcal{M} \subset X$ heißt **kompakt**, wenn \mathcal{M} relativ kompakt und vollständig ist, d.h. wenn jede unendliche Teilmenge aus \mathcal{M} eine Cauchyfolge enthält, deren Grenzwert wieder in \mathcal{M} liegt. In anderen Worten, \mathcal{M} ist relativ kompakt, wenn $\overline{\mathcal{M}}$ kompakt ist, d.h. wenn \mathcal{M} einen kompakten Abschluss besitzt.

Der **Satz von Heine-Borel** besagt, dass eine Teilmenge des \mathbb{R}^n genau dann kompakt ist, wenn sie beschränkt und abgeschlossen ist. Eine Konsequenz daraus ist, dass jede beschränkte Teilmenge des \mathbb{R}^n einen kompakten Abschluss besitzt, d.h. jede beschränkte Teilmenge des \mathbb{R}^n ist relativ kompakt. Das ist gleichbedeutend mit dem **Satz von Bolzano-Weierstraß**: jede beschränkte Folge in \mathbb{R}^n enthält mindestens eine konvergente Teilfolge, also mindestens einen Häufungspunkt. Beispielsweise ist das Intervall $[0, 1] \subset \mathbb{R}$ kompakt, aber nicht die Mengen \mathbb{N} (nicht beschränkt) oder $(0, 1]$ (nicht abgeschlossen). Im \mathbb{R}^n ist die Einheitskugel $\|x\| \leq 1$ beschränkt und abgeschlossen und bildet somit eine kompakte Menge. Alle diese Aussagen lassen sich auf allgemeine endlichdimensionale Vektorräume erweitern: dort sind die Begriffe *beschränkt* und *relativ kompakt* synonym (Fig. 5.8).

In unendlichdimensionalen Banachräumen gilt noch immer, dass eine kompakte Teilmenge beschränkt und abgeschlossen ist, aber die Umkehrung muss nicht mehr stimmen:

- Der Prototyp einer beschränkten und abgeschlossenen Menge ist die **Einheitskugel** $\|u\| \leq 1$. In einem unendlichdimensionalen Raum ist diese aber nicht mehr kompakt, denn sie enthält die unendliche Teilmenge der Einheitsvektoren $\{\varphi_j\}$ mit $\|\varphi_j\| = 1$, die wegen $\|\varphi_i - \varphi_j\| = \sqrt{2}$ für $i \neq j$ alle denselben Abstand voneinander haben – diese Teilmenge enthält daher keine Cauchyfolge.
- Die früher diskutierte Folge $(f_n)_{n=1}^\infty$ aus (1.20) bildet eine unendliche Teilmenge des Banachraums $C[-1, 1]$ der stetigen Funktionen bezüglich der Maximumsnorm $\|\cdot\|_\infty$. Diese Teilmenge ist beschränkt wegen $\|f_n\|_\infty = 1$, enthält aber keine Cauchyfolge und ist somit nicht kompakt.

Übertragen wir das Gesagte nun auf lineare beschränkte Operatoren in $\mathcal{B}(\mathcal{H})$. Diese sind gleichzeitig stetig, und als solches bilden sie beschränkte Teilmengen wieder auf beschränkte Teilmengen ab. Insbesondere ist das Bild einer kompakten Menge unter einer stetigen Funktion ebenfalls wieder kompakt (deswegen nimmt z.B. eine reellwertige stetige Funktion auf einer nichtleeren kompakten Definitionsmenge ein globales Minimum und ein globales Maximum an). Im unendlichdimensionalen Fall sind die Begriffe *beschränkt* und *kompakt* allerdings nicht mehr synonym, deswegen verschärft man die Stetigkeit durch folgende Definition:

SATZ: Ein Operator $K \in \mathcal{B}(\mathcal{H})$ heißt **kompakt** oder **vollstetig** genau dann, wenn irgendeine der folgenden äquivalenten Aussagen zutrifft:

- (a) K bildet beschränkte Teilmengen von \mathcal{H} auf relativ kompakte Teilmengen in \mathcal{H} ab.
- (b) Das Bild der offenen Einheitskugel $\|u\| < 1$ ist relativ kompakt.
- (c) Für jede beschränkte Folge $(u_n) \subset \mathcal{H}$ (d.h. $\|u_n\| \leq c$) enthält die Bildfolge (Ku_n) eine Teilfolge, die gegen ein Element aus \mathcal{H} konvergiert.
- (d) Wenn K_n beschränkte, endlichdimensionale Operatoren sind (sodass $\dim \mathcal{R}_{K_n} < \infty$), dann ist K der Limes der Operatornorm $\|K - K_n\| \rightarrow 0$. Folglich kann jeder kompakte Operator durch beschränkte, endlichdimensionale Operatoren approximiert werden.
- (e) Für jede orthonormale (unendliche) Folge $(\varphi_n) \subset \mathcal{H}$ gilt $K\varphi_n \rightarrow 0$.

Man kann zeigen, dass kompakte Operatoren einen Untervektorraum und auch eine Unter-*algebra* von $\mathcal{B}(\mathcal{H})$ bilden. Umgekehrt ist nicht jeder beschränkte Operator kompakt: z.B. ist der Einsoperator auf einem unendlichdimensionalen Hilbertraum nicht kompakt, denn er bildet die Einheitskugel wieder auf die Einheitskugel ab. Dasselbe gilt für unitäre Operatoren. Jeder lineare Operator auf einem endlichdimensionalen Hilbertraum ist dagegen kompakt, z.B. wegen der Eigenschaft (d): die konstante Folge $K_n = K$ konvergiert gegen K .

Die Prototypen kompakter Operatoren sind die **Hilbert-Schmidt-Operatoren**, die durch die Eigenschaften (5.39) bzw. (5.42) charakterisiert sind:

$$\int_{\Omega} dx \int_{\Omega} dy |K(x, y)|^2 = \sum_{ij} |K_{ij}|^2 = \sum_j \|K\varphi_j\|^2 < \infty, \quad (5.203)$$

wobei $\{\varphi_j\}$ eine ONB bezeichnet. Damit diese Summe konvergiert, muss $(K\varphi_j)$ eine Nullfolge sein, was genau der Eigenschaft (e) entspricht. Ein Beispiel ist der Operator $K\varphi_n = \varphi_n/n$; angewandt auf die Einheitskugel u mit $\|u\| \leq 1$ ergibt er

$$u = \sum_{n=1}^{\infty} u_n \varphi_n, \quad \|u\|^2 = \sum_{n=1}^{\infty} |u_n|^2 \leq 1 \quad \Rightarrow \quad Au = \sum_{n=1}^{\infty} \frac{u_n}{n} \varphi_n = \sum_{n=1}^{\infty} \beta_n \varphi_n. \quad (5.204)$$

Wegen $\|u\|^2 \leq 1$ muss sicher auch $|u_n| \leq 1$ gelten und somit $|\beta_n| \leq 1/n$. Die entsprechende Menge im Raum $\ell^2(\mathbb{C})$ wird auch **Hilbert-Würfel** genannt:

$$\left\{ x \in \ell^2(\mathbb{C}) \mid x = (x_n)_{n \in \mathbb{N}}, |x_n| \leq \frac{1}{n} \quad \forall n \in \mathbb{N} \right\}. \quad (5.205)$$

Spektrum eines kompakten Operators. Wie sieht das **Spektrum** eines kompakten Operators aus? Sei K ein kompakter Operator auf \mathcal{H} und $z \in \mathbb{C}$. Dann lassen sich folgende Aussagen treffen:

- Jeder kompakte Operator ist auch beschränkt und kann somit auf dem ganzen Hilbertraum definiert werden. Dann ist er auch abgeschlossen.
- Man kann zeigen, dass der Operator $K - z$ genau dann injektiv ist, wenn er auch surjektiv ist; außerdem ist sein Wertebereich \mathcal{R}_{K-z} für jedes $z \in \mathbb{C}$ abgeschlossen [Nee22]. Aus der Schnittmenge beider Bedingungen folgt, dass $K - z$ nur vom Typ (1A) oder (3C) sein kann, d.h. jedes $z \in \mathbb{C}$ liegt entweder im Punktspektrum (es ist ein Eigenwert) oder in der Resolventenmenge. Das stetige Spektrum und Restspektrum sind leer.
- Kompakte Operatoren haben zwar die angenehme Eigenschaft, dass sie beschränkte Teilmengen auf (relativ) kompakte Teilmengen “komprimieren”, das führt aber zwangsläufig auf Probleme mit der Inversion: ein kompakter Operator auf einem unendlichdimensionalen Hilbertraum ist nicht invertierbar. Die Eigenschaft (e) besagt, dass ein kompakter Operator jede orthonormale Folge mit $\|\varphi_n\| = 1$ auf $K\varphi_n \rightarrow 0$ abbildet, somit ist er nicht nullbeschränkt. Sein Inverses muss daher unbeschränkt sein, aber das schließt den Fall (1A) aus – somit existiert es nicht. Dann ist auch der Operator $K - z$ für $z = 0$ nicht invertierbar, weswegen der Punkt $z = 0$ nicht in der Resolventenmenge liegen kann und daher Teil des Spektrums sein muss.
- Für $z \neq 0$ gilt $\dim \mathcal{K}_{K-z} < \infty$: Auf seinen Nullraum eingeschränkt ist $K - z$ Null und damit gilt $z\mathbb{1}|_{\mathcal{K}_{K-z}} = K|_{\mathcal{K}_{K-z}}$; doch da $K|_{\mathcal{K}_{K-z}}$ noch immer ein kompakter Operator ist, muss auch $z\mathbb{1}|_{\mathcal{K}_{K-z}}$ kompakt sein. Wir wissen aber, dass der Einsoperator nicht kompakt ist – es sei denn, \mathcal{K}_{K-z} ist endlich. Daraus folgt, dass für einen kompakten Operator die Vielfachheit jedes Eigenwerts $z \neq 0$ höchstens endlich sein kann.

SATZ (Riesz-Schauder): Das Spektrum eines kompakten Operators K auf einem Hilbertraum \mathcal{H} hat folgende Eigenschaften:

- (a) Jedes $z \in \sigma(K)$ mit $z \neq 0$ liegt in $\sigma_p(K)$ und ist somit ein Eigenwert. Der Eigenraum \mathcal{K}_{K-z} für jeden Eigenwert z ist endlichdimensional.
- (b) Falls $\dim H = \infty$, dann ist $z = 0$ Teil des Spektrums.
- (c) Für jedes $c > 0$ ist die Anzahl der Eigenwerte mit $|z| \geq c$ endlich.

Die Aussagen (a) und (b) entsprechen dem bereits Gesagten; den Beweis für (c) findet man z.B. in [Nee22]. Somit ist das Spektrum eines kompakten Operators rein diskret und besteht entweder aus endlich vielen Eigenwerten, oder die Eigenwerte bilden eine Nullfolge, d.h. der einzige mögliche Häufungspunkt der Eigenwerte ist $z = 0$. Alle $z \in \mathbb{C}$, die nicht Eigenwerte sind, gehören zur Resolventenmenge $\rho(K)$. Das ist schon sehr nahe am Spektrum einer endlichdimensionalen Matrix! Es gibt allerdings keine Aussage darüber, ob die Eigenwerte reell oder komplex sind.

Fredholmsche Alternative. Wir können nun Aussagen über die Lösungen von Gleichungen der Art $(K - z)u = f$ treffen, z.B. von **Fredholm-Integralgleichungen** der Form

$$\int_a^b dy K(x, y) u(y) - zu(x) = f(x). \quad (5.206)$$

Dazu stellen wir fest:

- Sei $A = K - z$, sodass $f \in \mathcal{R}_A$. Wie wir gesehen haben, ist \mathcal{R}_A abgeschlossen und $z = 0$ liegt entweder in der Resolventenmenge oder im Punktspektrum. Falls $z \in \rho(A)$, muss $\mathcal{R}_A = \mathcal{H}$ gelten und $\mathcal{K}_A = \{0\}$ (siehe Fig. 5.4). Dann ist $(K - z)u = f$ für alle $f \in \mathcal{H}$ eindeutig lösbar. Liegt z im Punktspektrum, dann spannen die Eigenvektoren von K den Untervektorraum \mathcal{K}_A mit endlicher Dimension $\dim \mathcal{K}_A$ auf.
- Man kann zeigen [z.B. Sta11, Wei00], dass $\dim \mathcal{K}_A = \dim \mathcal{K}_{A^\dagger}$ gilt. In anderen Worten: wenn z ein Eigenwert von K ist, dann ist z^* ein Eigenwert von K^\dagger . Das bedeutet: falls z in der Resolventenmenge liegt, gilt $\dim \mathcal{K}_A = \dim \mathcal{K}_{A^\dagger} = 0$, und auch A^\dagger erlaubt nur die triviale Lösung. Falls z im Punktspektrum liegt, ist $\dim \mathcal{K}_A = \dim \mathcal{K}_{A^\dagger} \neq 0$, sodass auch $A^\dagger v$ nichttriviale Lösungen haben muss, nämlich die Eigenvektoren von K^\dagger .
- Aus (5.70) folgt $\mathcal{K}_{A^\dagger} = (\mathcal{R}_A)^\perp$, und weil \mathcal{R}_A abgeschlossen ist, auch $\mathcal{K}_{A^\dagger}^\perp = \overline{\mathcal{R}_A} = \mathcal{R}_A$. Die Lösung $f \in \mathcal{R}_A$ muss daher orthogonal auf alle v stehen, die die Gleichung $A^\dagger v = 0$ lösen, nämlich die Eigenvektoren von K^\dagger . Für $z \in \rho(A)$ ist das trivial erfüllt, denn mit $\mathcal{R}_A = \mathcal{H}$ gilt $\mathcal{K}_{A^\dagger} = \{0\}$, sodass $A^\dagger v = 0$ nur die Nulllösung hat und K^\dagger keine Eigenwerte besitzt. Somit existiert eine Lösung f genau dann, wenn $f \perp \mathcal{K}_{A^\dagger}$, also wenn $\langle v, f \rangle = 0$ für alle Eigenvektoren v von K^\dagger gilt.

Daraus ergibt sich die auf [Erik Ivar Fredholm](#) (1866–1927) zurückgehende Lösbarkeitsbedingung der Gleichung $(K - z)u = f$:

SATZ (Fredholmsche Alternative): Sei K ein kompakter Operator auf dem Hilbertraum \mathcal{H} , mit $z \in \mathbb{C}$ und $A = K - z$. Dann gilt eine der folgenden Alternativen:

- Entweder ist $\dim \mathcal{K}_A = \dim \mathcal{K}_{A^\dagger} = 0$; dann erlauben sowohl $Au = 0$ als auch $A^\dagger v = 0$ nur die Nulllösung, und die Gleichung $Au = f$ hat genau eine Lösung $\forall f \in \mathcal{H}$.
- Oder es gilt $\dim \mathcal{K}_A = \dim \mathcal{K}_{A^\dagger} = n \neq 0$; dann hat die homogene Gleichung $Au = 0$ nichttriviale Lösungen $\{u_1, \dots, u_n\}$ und $A^\dagger v = 0$ nichttriviale Lösungen $\{v_1, \dots, v_n\}$. Die Gleichung $Au = f$ hat Lösungen genau dann, wenn f im orthogonalen Komplement von \mathcal{K}_{A^\dagger} liegt, d.h. wenn $\langle v_i, f \rangle = 0$ für alle $v_i \in \mathcal{K}_{A^\dagger}$. Die allgemeine Lösung von $Au = f$ hat dann die Form

$$u = u_p + \sum_{i=1}^n c_i u_i, \quad (5.207)$$

wobei u_p eine partikuläre Lösung der Gleichung $Au = f$ ist.

Die Relation $\dim \mathcal{K}_A = \dim \mathcal{K}_{A^\dagger} = \dim \mathcal{R}_A^\perp$ ist die Verallgemeinerung aus der linearen Algebra, wo die Dimension einer $N \times N$ -Matrix A durch $N = \dim \mathcal{K}_A + \dim \mathcal{R}_A$ gegeben ist, wobei $\dim \mathcal{R}_A$ der Rang der Matrix ist.

Spektralsatz für kompakte selbstadjungierte Operatoren. Ein kompakter Operator hat ein reines Punktspektrum, wobei die Eigenwerte entweder eine endliche Menge oder eine Nullfolge bilden. Für einen selbstadjungierten Operator ist das Punktspektrum reell und die Eigenvektoren zu verschiedenen Eigenwerten sind orthogonal. Kombiniert man beide Eigenschaften, lässt sich die **Spektraldarstellung** der linearen Algebra auf kompakte selbstadjungierte Operatoren verallgemeinern. Dazu benötigen wir folgenden Hilfssatz:

LEMMA: Ein kompakter, selbstadjungierter Operator A auf einem Hilbertraum \mathcal{H} hat mindestens einen Eigenwert, nämlich $+||A||$ oder $-||A||$, und damit auch einen Eigenvektor.

Beweis: Wenn A beschränkt ist, gilt $||A|| < \infty$. Falls A zusätzlich selbstadjungiert ist, lässt sich die Operatornorm mittels (5.76) auch über den Erwartungswert angeben:

$$||A|| = \sup_{||u|| \leq 1} |\langle u, Au \rangle|. \quad (5.208)$$

Nachdem die Operatornorm die kleinste obere Schranke über alle Vektoren u darstellt, kann man zwei Fälle unterscheiden: (i) Entweder es gibt ein u mit $||u|| = 1$, für welches der Erwartungswert $\langle u, Au \rangle =: \lambda$ das Supremum erreicht. Da der Erwartungswert für einen selbstadjungierten Operator reell ist, muss in diesem Fall $\lambda = +||A||$ oder $\lambda = -||A||$ gelten. Dann ist aber

$$||Au - \lambda u||^2 = ||Au||^2 - 2\lambda \underbrace{\langle u, Au \rangle}_{=\lambda} + \lambda^2 \underbrace{||u||^2}_{=1} = ||Au||^2 - \lambda^2 \leq \underbrace{||A||^2}_{=\lambda^2} \underbrace{||u||^2}_{=1} - \lambda^2 = 0 \quad (5.209)$$

und somit $Au = \lambda u$, d.h. λ ist ein Eigenwert von A und u der zugehörige Eigenvektor. (ii) Oder es gibt eine Folge (u_n) mit $||u_n|| = 1$, sodass $\langle u_n, Au_n \rangle \rightarrow \pm ||A|| =: \lambda$. Dann haben wir wieder

$$||Au_n - \lambda u_n||^2 = \underbrace{||Au_n||^2}_{\leq ||A||^2 = \lambda^2} - 2\lambda \langle u_n, Au_n \rangle + \lambda^2 \underbrace{||u_n||^2}_{=1} \leq 2\lambda^2 - 2\lambda \underbrace{\langle u_n, Au_n \rangle}_{\rightarrow \lambda} \rightarrow 0, \quad (5.210)$$

und somit konvergiert $Au_n - \lambda u_n$. Woher wissen wir aber, dass Au_n und u_n selbst konvergieren? Hier kommt die Kompaktheit von A ins Spiel: Die Menge der u_n ist normiert und somit beschränkt, also ist die Folge (Au_n) kompakt und somit auch konvergent (oder sie hat zumindest eine konvergente Teilfolge). Wegen der Konvergenz von Au_n und $Au_n - \lambda u_n$ muss auch u_n konvergieren. Weil A stetig (da beschränkt) ist, ist der Grenzwert $Au = \lambda u$, also ist λ ein Eigenwert von A und u der zugehörige Eigenvektor.

Dieses Prinzip lässt sich fortsetzen. Angenommen, wir haben auf diese Weise den Eigenwert λ_1 mit $|\lambda_1| = ||A||$ gefunden. Dann muss $||Au||/||u|| \leq ||A|| = \lambda_1$ für jeden weiteren Vektor $u \in \mathcal{H}$ gelten. Sei v_1 der zu λ_1 zugehörige normierte Eigenvektor und $\mathcal{M}_1 = \text{span}\{v_1\}$ seine lineare Hülle. Dann sind alle Vektoren aus $\mathcal{H}_2 := \mathcal{M}_1^\perp$ orthogonal zu v_1 . Der Operator A lässt die beiden Untervektorräume \mathcal{M}_1 und \mathcal{H}_2 invariant, denn angewandt auf \mathcal{M}_1 liefert er ein Vielfaches von v_1 , und angewandt auf $w \in \mathcal{H}_2$ ergibt er wieder ein Element aus \mathcal{H}_2 :

$$\langle v_1, Aw \rangle = \langle Av_1, w \rangle = \lambda_1 \langle v_1, w \rangle = 0, \quad \text{d.h.} \quad w \perp v_1 \Rightarrow Aw \perp v_1. \quad (5.211)$$

Wir bezeichnen die Einschränkung von A auf \mathcal{H}_2 mit $A_2 = A|_{\mathcal{H}_2}$. Dann ist A_2 in \mathcal{H}_2 wieder ein selbstadjungierter und kompakter Operator. Da $||A||$ die kleinste obere Schranke ist, gilt

$$||A_2|| = \sup_{w \in \mathcal{H}_2} \frac{||A_2 w||}{||w||} = \sup_{w \in \mathcal{H}_2} \frac{||Aw||}{||w||} \leq ||A||. \quad (5.212)$$

Damit können wir das obige Lemma erneut auf \mathcal{H}_2 anwenden, woraus wir den nächsten Eigenwert λ_2 mit $|\lambda_2| = ||A_2|| \leq ||A|| = |\lambda_1|$ erhalten. Der zugehörige normierte Eigenvektor v_2 ist automatisch orthogonal zu v_1 , da er ja in \mathcal{H}_2 lebt.

Im nächsten Schritt definieren wir $\mathcal{M}_2 = \text{span}\{v_1, v_2\}$, $\mathcal{H}_3 = \mathcal{M}_2^\perp$ und $A_3 = A|_{\mathcal{H}_3}$ und erhalten den Eigenwert $|\lambda_3| \leq |\lambda_2| \leq |\lambda_1|$ mit zugehörigem Eigenvektor v_3 , der orthogonal zu v_1 und v_2 ist. Das lässt sich beliebig fortsetzen. Auf diese Weise erhält man Eigenwerte $|\lambda_1| \geq |\lambda_2| \geq |\lambda_3| \geq \dots$, deren Eigenvektoren $\{v_1, v_2, v_3, \dots\}$ ein Orthonormalsystem bilden. Das Verfahren kann auf zwei Arten ausgehen:

- Entweder ist nach $n + 1$ (also endlich vielen) Schritten das Supremum erstmals Null und somit $|\lambda_n| > 0$, womit alle weiteren Suprema ebenfalls verschwinden. Dann haben wir n nichtverschwindende Eigenwerte gefunden (wobei alle Multiplizitäten mitgezählt wurden). In einem unendlichdimensionalen Hilbertraum hat der Eigenwert $\lambda = 0$ dann eine unendliche Multiplizität, da jedes Element aus \mathcal{M}_n^\perp ein Eigenvektor von A ist.
- Oder das Supremum verschwindet in keinem endlichen Schritt, sodass die Eigenwerte eine Nullfolge bilden: $|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_n| \geq \dots \geq 0$, mit orthonormalen Eigenvektoren $\{v_1, v_2, \dots, v_n, \dots\}$. Dass der einzige mögliche Häufungspunkt der Eigenwerte Null ist, sieht man so: andernfalls wäre die Folge (v_n/λ_n) nämlich beschränkt, und wegen $Av_n = \lambda v_n \Rightarrow A(v_n/\lambda_n) = v_n$ müsste die Folge (v_n) wegen der Kompaktheit von A eine konvergente Teilfolge haben, was nicht der Fall ist, da es sich um Einheitsvektoren handelt. Daher muss $\lambda_n \rightarrow 0$ gelten.

Die Eigenfunktionen $\{v_n\}$ für $\lambda_n \neq 0$ bilden nicht notwendigerweise eine ONB für den gesamten Hilbertraum, aber zumindest für den Wertebereich \mathcal{R}_A . Sei nämlich u ein beliebiger Hilbertraumvektor, dann ist

$$w := u - \sum_{i=1}^n v_i \langle v_i, u \rangle \quad (5.213)$$

ein Element aus $\mathcal{M}_n^\perp = \mathcal{H}_{n+1}$, da alle $\langle v_k, w \rangle = 0$ verschwinden, und es gilt

$$Aw = Au - \sum_{i=1}^n \underbrace{\lambda_i v_i \langle v_i, u \rangle}_{=v_i \langle \lambda_i v_i, u \rangle} = Au - \sum_{i=1}^n v_i \langle Av_i, u \rangle = Au - \sum_{i=1}^n v_i \langle v_i, Au \rangle. \quad (5.214)$$

Wenn die Eigenvektoren eine ONB bilden sollen, sodass jedes $Au \in \mathcal{R}_A$ durch sie aufgespannt werden kann, muss also $\lim_{n \rightarrow \infty} \|Aw\| = 0$ erfüllt sein. Wir wissen, dass $\|Aw\| \leq |\lambda_{n+1}| \|w\|$ für jeden Vektor $w \in \mathcal{H}_{n+1}$ gelten muss, da $|\lambda_{n+1}| = \|A_{n+1}\|$ das Supremum darstellt. Weil die Eigenwerte eine Nullfolge bilden, gilt $\|Aw\| \leq |\lambda_{n+1}| \|w\| \leq |\lambda_{n+1}| \|u\| \rightarrow 0$, also $Aw \rightarrow 0$. Zusammenfassend ergibt sich folgender Satz:

SATZ (Spektralsatz): Sei A ein kompakter, selbstadjungierter Operator auf dem Hilbertraum \mathcal{H} . Dann hat A abzählbar (endlich oder unendlich) viele reelle Eigenwerte $\lambda_i \neq 0$ mit $|\lambda_1| \geq |\lambda_2| \geq \dots \geq 0$. Die Vielfachheit jedes Eigenwerts $\lambda_i \neq 0$ ist endlich, und Null ist der einzige mögliche Häufungspunkt der Eigenwerte: $\lim_{i \rightarrow \infty} \lambda_i = 0$. Jeder Vektor $Au \in \mathcal{R}_A$ lässt sich als

$$Au = \sum_{i=1}^{\infty} v_i \langle v_i, Au \rangle = \sum_{i=1}^{\infty} \lambda_i v_i \langle v_i, u \rangle \Leftrightarrow A = \sum_{i=1}^{\infty} \lambda_i |v_i\rangle \langle v_i| \quad (5.215)$$

darstellen, wobei die v_i die Eigenvektoren zum Eigenwert λ_i sind.

Einige Bemerkungen:

- Der Spektralsatz ist äußerst nützlich, denn mit seiner Hilfe kann man für kompakte selbstadjungierte Operatoren beliebige **stetige Funktionen** $f: \mathbb{R} \rightarrow \mathbb{C}$ definieren:

$$f(A) = \sum_{i=1}^{\infty} f(\lambda_i) |v_i\rangle\langle v_i|. \quad (5.216)$$

- Die Herleitung des Spektralsatzes liefert nicht nur die Aussage (5.215), sondern mit dem **Ritzschen Variationsverfahren** auch eine konkrete Vorschrift zur *Bestimmung* der Eigenwerte, nämlich: um den Eigenwert $\lambda_i \neq 0$ und dessen Eigenvektor v_i zu gewinnen, finde das absolute Maximum der Funktion $u \mapsto \langle u, Au \rangle$ unter der Nebenbedingung $\|u\| = 1$ für $u \perp v_1, v_2, \dots, v_{i-1}$.
- Bezieht man auch die Eigenvektoren aus dem Nullraum von A mit ein, d.h. jene für $\lambda_i = 0$, dann bildet die Menge *aller* Eigenvektoren von A eine ONB für \mathcal{H} . Aus der Herleitung der Fredholmschen Alternative sieht man nämlich, dass für einen kompakten, selbstadjungierten ($A = A^\dagger$) Operator $\mathcal{K}_A^\perp = \mathcal{R}_A$ und $\mathcal{R}_A^\perp = \mathcal{K}_A$ gelten muss, also $\mathcal{H} = \mathcal{K}_A \oplus \mathcal{R}_A$. Somit lässt sich jedes Element $f \in \mathcal{H}$ schreiben als

$$f = h + \sum_i v_i \langle v_i, f \rangle \quad \text{mit} \quad h \in \mathcal{K}_A \Leftrightarrow Ah = 0, \quad (5.217)$$

wobei h orthogonal zu allen Eigenvektoren v_i für $\lambda_i \neq 0$ ist:

$$0 = \langle v_i, Ah \rangle = \langle Av_i, h \rangle = \lambda_i \langle v_i, h \rangle \Rightarrow h \perp v_i. \quad (5.218)$$

Falls $f \in \mathcal{R}_A$, muss $h = 0$ sein; falls $f \in \mathcal{R}_A^\perp = \mathcal{K}_A$, verschwindet die Summe wegen $\langle v_i, f \rangle = 0$. Die Menge aller Eigenvektoren v_i für $\lambda_i \neq 0$ bildet daher eine ONB für \mathcal{H} genau dann, wenn $\lambda = 0$ kein Eigenwert von A ist, d.h. wenn $\dim \mathcal{K}_A = 0$.

Ausblick: Wegen der einfachen Gestalt ihres Spektrums haben wir den Spektralsatz nur für *kompakte* selbstadjungierte Operatoren hergeleitet. Man kann ihn auch für allgemeine (sowohl beschränkte als auch unbeschränkte) selbstadjungierte Operatoren formulieren, obwohl die Herleitung deutlich aufwendiger wird. In (5.215) gehen dann nicht nur die Eigenwerte ein, sondern auch das stetige Spektrum, weswegen die Spektralsumme zu einem Spektralintegral wird. Dazu definiert man eine **Spektralschar** (*projection-valued measure*); Näheres findet man z.B. in [Gro14, Bla93, Wei00]. Darauf basiert auch die in der Quantenmechanik gebräuchliche Schreibweise

$$A = \int a_i |a_i\rangle\langle a_i|, \quad \text{z.B.} \quad X = \int dx x |x\rangle\langle x|, \quad \langle x, x' \rangle = \delta(x - x') \quad (5.219)$$

für den Ortsoperator, der laut unseren früheren Überlegungen ja keine Eigenwerte, sondern nur ein stetiges Spektrum besitzt. Die entsprechenden “Eigenvektoren” $|x\rangle$ des Operators X sind dann **uneigentliche Eigenvektoren** in einem distributionellen Sinn, und die Operatoren werden zu operatorwertigen Distributionen. Solche Objekte spielen insbesondere in der Quantenfeldtheorie eine wichtige Rolle.