Höhere Analysis - Notizen [1]

Jena - Sommersemester 2018

Daniel Lenz

[1] Es handelt sich nicht um ein Skriptum zur Vorlesung. Konstruktive Kommentare sind willkommen.

Inhaltsverzeichnis

KAPITEL 1

Etwas Maß- und Integrationstheorie

Das Konzept des Maßes stellt eine praezise (und sehr allgemeine) Fassung von Volumenmessung zur Verfuegung. Die charakteristische Eigenschaft eines Maßes μ ist die Gleichung

$$\mu(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mu(A_n)$$

fuer paarweise disjunkte A_n, $n \in \mathbb{N}$. Basierend auf einem Maß koennen dann Funktionen integriert werden. Im allgemeinen kann nicht allen Mengen ein Maß zugeordnet werden, und es koennen nicht alle (positiven) Funktionen integriert werden. Die 'guten' Mengen bzw. Funktionen heissen *meßbar*.
Der hier vorgestellten Betrachtungen liefern eine Alternative zu dem in der Vorlesung Analysis III (Lenz) ausgeführten Zugang.
Die wesentliche Rolle dieses Kapitels in der Vorlesung ist es, im naechsten Kapitel die Definition der L^p-Raeume zu ermoeglichen. Diese L^p-Raeume sind wichtige Beispiele fuer die Vorlesung.

1. σ-Algebren und meßbare Funktionen

In diesem Abschnitt lernen wir die 'guten' Mengen und Funktionen kennen.

Definition (σ-Algebra). *Sei X eine Menge und $\mathcal{A}$ eine Familie von Teilmengen von X. Dann heißt $\mathcal{A}$ eine σ-Algebra, wenn folgende Eigenschaften erfuellt sind:*

- *X gehoert zu $\mathcal{A}$.*
- *Mit A gehoert auch $X \setminus A$ zu $\mathcal{A}$.*
- *Mit A_n, $n \in \mathbb{N}$, gehoert auch $\bigcup_n A_n$ zu $\mathcal{A}$.*

Ist $\mathcal{A}$ eine σ-Algebra auf X, so heisst $(X, \mathcal{A})$ ein meßbarer Raum und die Elemente von $\mathcal{A}$ werden meßbar genannt.[1]

Eine σ-Algebra ist also abgeschlossen unter Komplementbildung und abzaehlbaren Vereinigungen und enthaelt jedenfalls den ganzen Raum.

Bemerkung. Es ist $\{X, \emptyset\}$ eine σ-Algebra auf X und in jeder anderen enthalten.

Proposition (Einfache Eigenschaften). *Sei $(X, \mathcal{A})$ ein meßbarer Raum. Dann gilt:*

[1] Im Zuge der sogenannten Rechtschreibereform wurde das Konzept der Meßbarkeit abgeschafft und durch das Konzept der Messbarkeit ersetzt. In diesem Text werden beide Konzepte synonym verwendet.

(a) $A_1, \dots, A_n \in \mathcal{A} \Longrightarrow \bigcup_{j=1}^n A_j \in \mathcal{A}$.
(b) $A_n \in \mathcal{A}$, $n \in \mathbb{N}$, $\Longrightarrow \bigcap_{n=1}^\infty A_n \in \mathcal{A}$.
(c) $A_1, \dots, A_n \in \mathcal{A} \Longrightarrow \bigcap_{j=1}^n A_j \in \mathcal{A}$.
(d) $A, B \in \mathcal{A} \Longrightarrow A \setminus B \in \mathcal{A}$.

Beweis. (a) Waehle $A_j = \emptyset$ fuer $j = n+1, ...$; nutze dritte Eigenschaft.

(b) Das folgt durch doppelte Komplementbildung:

$$\bigcap_n A_n = X \setminus (\bigcup_n (X \setminus A_n)).$$

(c) Das folgt aus (b) nach Wahl von $A_j = X$ fuer $j = n+1, ...$

(d) Das folgt aus (c) und $A \setminus B = A \bigcap (X \setminus B)$. □

Bemerkung / Erinnerung. Zum besseren Verstaendnis mag ein Vergleich mit dem Konzept der Topologie dienen: Eine *Topologie* τ auf X ist eine Familie von Teilmengen von X mit folgenden Eigenschaften:

- $\emptyset, X \in \tau$.
- Mit $U_1, \dots, U_n$ gehoert auch $\bigcap_{j=1}^n U_j$ zu τ.
- Mit U_α, $\alpha \in A$, gehoert auch $\bigcup_\alpha U_\alpha$ zu τ.

Die Elemente von τ heissen dann *offene Mengen*. Das Paar (X, τ) heisst *topologischer Raum*.
Ist (X, d) ein metrischer Raum, so bildet die Familie der Mengen U mit der Eigenschaft, dass zu jedem $x \in U$ ein $r_x > 0$ existiert mit

$$U(x, r_x) := \{y \in X : d(x, y) < r_x\} \subset U$$

eine Topologie (wie man leicht sieht bzw. aus Analysis II weiss). Solche Mengen heissen dann *offen*.
Sind (X_j, τ_j), $j = 1, 2$ Mengen mit Topologien, so heisst eine Funktion $f : X_1 \longrightarrow X_2$ *stetig*, wenn $f^{-1}(U)$ zu τ_1 gehoert fuer jedes U aus τ_2.

Man kann σ-Algebren aus einer gegebene Familie von Teilmengen erzeugen.

Theorem (Erzeugung von σ-Algebren durch $\mathcal{F}$). *Sei $\mathcal{F}$ eine Familie von Teilmengen von X. Dann existiert genau eine σ-Algebra $\mathcal{A}$ auf X mit folgenden beiden Eigenschaften:*

- *$\mathcal{F} \subset \mathcal{A}$.*
- *Gilt $\mathcal{F} \subset \mathcal{B}$ fuer eine σ-Algebra $\mathcal{B}$, so folgt $\mathcal{A} \subset \mathcal{B}$.*

Beweis. Eindeutigkeit. Das ist klar. (Seien $\mathcal{A}_i$, $i = 1, 2$, solche σ-Algebren. Waehle $\mathcal{A}_2 = \mathcal{B}$ und schliesse $\mathcal{A}_1 \subset \mathcal{A}_2$. Vertausche nun die Rollen von $\mathcal{A}_1$ und $\mathcal{A}_2$.)

Existenz. Es existiert ein σ-Algebra, die $\mathcal{F}$ enthaelt (z.B. die Potenzmenge von X). Der Durchschnitt aller solcher σ-Algebren ist wieder ein σ-Algebra (Check!) und enthaelt $\mathcal{F}$ (Check!). Er hat nach Konstruktion die gewuenschten Eigenschaften. □

Definition. *In der Situation des vorigen Theorems nennt man $\mathcal{A}$ die durch $\mathcal{F}$ erzeugte σ-Algebra und schreibt sie auch als $\mathcal{A}(\mathcal{F})$.*

Beispiel. Sei $X = \mathbb{R}$ und seien die Familien $\mathcal{F}_i$ gegeben als:
$\mathcal{F}_0$: Intervalle der Form (a, ∞), $a \in \mathbb{R}$.

$\mathcal{F}_1$: Alle Intervalle.

$\mathcal{F}_2$: Alle offenen Intervalle.

$\mathcal{F}_3$: Alle Intervalle der Form $(-\infty, a)$, $a \in \mathbb{R}$.

$\mathcal{F}_4$: Intervalle der Form $[a, \infty)$, $a \in \mathbb{R}$.

$\mathcal{F}_5$: Intervalle der Form $(-\infty, a]$, $a \in \mathbb{R}$.

$\mathcal{F}_6$: Intervalle der Form (a, b), $a \in \mathbb{R}$.

Dann gilt $\mathcal{A}(\mathcal{F}_0) = \ldots = \mathcal{A}(\mathcal{F}_6)$.
Beweis. Uebung. Nutze Formeln der folgenden Art:

$$[a, b) = \bigcap_n (a - 1/n, b), (a, b) = \bigcup_n [a + 1/n, b), (a, b) = (a, \infty) \setminus [b, \infty)...$$

$\square$

Definition (Borel-σ-Algebra). *Ist (X, τ) ein topologischer Raum, so heisst die von τ erzeugte σ-Algebra die Borel - σ - Algebra auf X.*

Beispiele.
(a) Ist $X = \mathbb{R}$, so ist die Borel-σ-Algebra gerade die oben schon diskutierte von den offenen Intervallen erzeugte σ-Algebra. (Da jede offene Menge eine Vereinigung von abzaehlbar vielen offenen Intervallen ist.)

(b) Ist $X = \mathbb{R}^N$, so wird die Borel-σ-Algebra erzeugt durch die Rechtecke der Form

$$R = (a_1, b_1) \times \ldots \times (a_N, b_N)$$

mit $a_i, b_i \in \mathbb{R}$ mit $a_i \leq b_i$ fuer $i = 1, \ldots, N$.

(c) Ist $X = \mathbb{R} \cup \{-\infty\} \cup \{\infty\}$ die Zweipunktkompaktifizierung von $\mathbb{R}$ mit der von den Intervallen (a, b), $(a, \infty]$ und $[-\infty, b)$, $a, b \in \mathbb{R}$, erzeugten Topologie, so wird die Borel-σ-Algebra erzeugt z.b. durch die Familie

$$(a, \infty], a \in \mathbb{R}.$$

Definition (Meßbarkeit). *Seien $(X_i, \mathcal{A}_i)$, $i = 1, 2$, meßbare Raeume. Eine Funktion $f : X_1 \longrightarrow X_2$ heisst meßbar, wenn $f^{-1}(A) \in \mathcal{A}_1$ fuer alle A aus $\mathcal{A}_2$ gilt.*

Bemerkung. (a) Vergleiche Definition der Stetigkeit.
(b) Offenbar sind konstante Funktionen meßbar (da die auftretenden Urbilder entweder leer oder der gesamte Raum sind).

Aus der Definition folgt sofort, dass die Komposition meßbarer Funktionen wieder meßbar ist:

Proposition (Komposition meßbarer Funktionen ist meßbar). *Seien $(X_i, \mathcal{A}_i)$, $i = 1, 2, 3$ meßbare Räume und $f : X_1 \longrightarrow X_2$ und $g : X_2 \longrightarrow X_3$ meßbar. Dann ist auch $g \circ f : X_1 \longrightarrow X_3$ meßbar.*

Wesentlich ist folgende Vertraeglichkeit von Funktionen mit den charakteristischen Eigenschaften einer σ-Algebra.

Proposition. *Sei $(X, \mathcal{A})$ ein meßbarer Raum und $f : X \longrightarrow Y$ eine Funktion. Dann ist die Menge*

$$\{B \subset Y : f^{-1}(B) \in \mathcal{A}\}$$

eine σ-Algebra.

Beweis. Bezeichne die genannte Menge mit $\mathcal{A}_f$. Dann gilt:
- $Y \in \mathcal{A}_f$: Klar (da $f^{-1}(Y) = X \in \mathcal{A}$).
- $B \in \mathcal{A}_f \Longrightarrow B^c \in \mathcal{A}_f$: $f^{-1}(B^c) = X \setminus f^{-1}(B) \in \mathcal{A}$ (wegen $f^{-1}(B) \in \mathcal{A}$).
- $B_n \in \mathcal{A}_f \Longrightarrow \bigcup B_n \in \mathcal{A}_f$: $f^{-1}(\bigcup_n B_n) = \bigcup_n f^{-1}(B_n) \in \mathcal{A}$ (wegen $f^{-1}(B_n) \in \mathcal{A}$).

□

Uebung. Untersuche entsprechende Aussage fuer Topologien. (Es stellt sich heraus, dass Schnitte, Vereinigung und Komplementbildung mit Urbildbildung vertraeglich ist.)

Lemma (Reicht Urbilder des Erzeuger zu testen). *Seien $(X, \mathcal{A})$ und $(Y, \mathcal{B})$ meßbare Raeume und $\mathcal{B}$ sei von der Familie $\mathcal{F}$ erzeugt. Dann sind fuer $f : X \longrightarrow Y$ aequivalent:*

(i) *Es ist f meßbar.*
(ii) *Es ist $f^{-1}(B)$ meßbar fuer alle $B \in \mathcal{F}$.*

Beweis. (i)$\Longrightarrow$ (ii): Das ist klar.
(ii)$\Longrightarrow$ (i): Betrachte $\mathcal{A}_f := \{B \subset Y : f^{-1}(B) \in \mathcal{A}\}$. Dann gilt:
- $\mathcal{F} \subset \mathcal{A}_f$ (nach (ii)).
- $\mathcal{A}_f$ ist eine σ-Algebra (nach voriger Proposition).

Damit folgt dann nach Konstruktion der erzeugten σ-Algebra

$$\mathcal{B} = \mathcal{A}(\mathcal{F}) \subset \mathcal{A}_f.$$

Das beendet den Beweis. □

Wir ziehen einige Folgerungen aus dem vorigen Lemma.

Folgerung (Stetige Funktionen sind meßbar). *Seien (X_i, τ_i), $i = 1, 2$, topologische Raeume mit den zugehoerigen Borel - σ -Algebren $\mathcal{B}_i$. Ist $f : X_1 \longrightarrow X_2$ stetig, so ist es meßbar.*

Beweis. Da $\mathcal{B}_2$ von τ_2 erzeugt wird, reicht es nach dem vorigen Lemma zu zeigen, dass $f^{-1}(U)$ meßbar ist fuer alle $U \in \tau_2$. Es gilt aber

$$f^{-1}(U) \in \tau_1 \subset \mathcal{B}_1,$$

da f stetig ist und $\mathcal{B}_1$ von τ_1 erzeugt ist. □

Wir ziehen nun Folgerungen fuer meßbare reellwertige Funktionen. Insbesondere werden wir sehen, dass die alle 'ueblichen' Operationen Meßbarkeit erhalten. Insbesondere bleibt Meßbarkeit (ANDERS als Stetigkeit) unter punktweisen Grenzwerten erhalten.

Konvention. Wenn es um Meßbarkeit von Funktionen mit Werten in $\mathbb{R}$ (bzw. $\mathbb{R} \cup \{-\infty\} \cup \{\infty\}$) geht, wird $\mathbb{R}$ (bzw. $\mathbb{R} \cup \{-\infty\} \cup \{\infty\}$) immer mit der Borel - σ - Algebra versehen.

Aus dem vorigen Lemma und den verschiedenen Arten die Borel-σ-Algebra zu erzeugen, erhaelt man sofort:

FOLGERUNG. *Sei $(X, \mathcal{A})$ ein meßbarer Raum und $f : X \longrightarrow \mathbb{R}$ (bzw. $f : X \longrightarrow \mathbb{R} \cup \{\infty\} \cup \{-\infty\}$) eine Funktion. Dann sind aequivaelent:*

(i) *Es ist f Borel-meßbar.*
(ii) *Es ist $f^{-1}(I)$ meßbar fuer alle Intervalle I.*
(iii) *Es ist $f^{-1}(a, \infty]$ meßbar fuer alle $a \in \mathbb{R}$.*
(iv) *Es ist $f^{-1}[-\infty, a)$ meßbar fuer alle $a \in \mathbb{R}$.*

Bemerkung. Aus (ii) in der Folgerung folgt sofort, dass mit f auch $-f$ meßbar ist.

THEOREM. *Sei $(X, \mathcal{A})$ ein meßbarer Raum und seien $f_n : X \longrightarrow \mathbb{R} \bigcup \{-\infty\} \bigcup \{\infty\}$, $n \in \mathbb{N}$, meßbare Funktionen. Dann sind auch die Funktionen*

$$\sup_n f_n, \ \inf_n f_n, \ \limsup_n f_n \ \textit{und} \ \liminf_n f_n$$

meßbar.

Beweis. Setze $g = \sup f_n$. Dann gilt

$$g^{-1}((a, \infty]) = \bigcup_{n=1}^{\infty} f_n^{-1}((a, \infty])$$

und dies ist meßbar als abzaehlbare Vereinigung meßbarer Mengen. Aehnlich kann man $\inf f_n$ behandeln. Damit folgen dann auch die Aussagen fuer

$$\limsup_n f_n = \inf_k \left(\sup_{n \geq k} f_n \right)$$

und

$$\liminf_n f_n = \sup_k \left(\inf_{n \geq k} f_n \right).$$

Das beendet den Beweis. □

Aus dem Theorem erhalten wir die folgende sehr bemerkenswerte Konsequenz (vgl. obige Bemerkung zur Stetigkeit):

FOLGERUNG (Punktweise Grenzwerte sind meßbar). *Sei $(X, \mathcal{A})$ ein meßbarer Raum. Konvergiert die Folge der meßbaren Funktionen*

$$f_n : X \longrightarrow \mathbb{R} \cup \{\infty\} \cup \{-\infty\}$$

punktweise gegen die Funktion f, so ist f meßbar.

Aus dem Theorem erhaelt man auch sofort folgendes Ergebnis.

FOLGERUNG. *Sei $(X, \mathcal{A})$ ein meßbarer Raum. Sind $f, g : X \longrightarrow \mathbb{R}$ meßbar, so sind auch $\max\{f, g\}$ und $\min\{f, g\}$ meßbar. Insbesondere sind auch $f_+ = \max\{f, 0\}$ und $f_- = -\min\{f, 0\} = \max\{-f, 0\}$ meßbar.*

THEOREM (Komponentenweise Meßbarkeit impliziert Meßbarkeit). *Sei $(X, \mathcal{A})$ ein meßbarer Raum und $f_1, \ldots, f_N : X \longrightarrow \mathbb{R}$ meßbar. Dann ist auch die Funktion*

$$F : X \longrightarrow \mathbb{R}^N, F(x) = (f_1(x), \ldots, f_N(x)),$$

meßbar.

Bemerkung. Es ist nicht schwer zu sehen, dass auch die Umkehrung des Theorems gilt: Ist die Funktion $F = (f_1, \ldots, f_n) : X \longrightarrow \mathbb{R}^n$ meßbar, so sind auch die einzelnen Komponenten $f_j = pr_j \circ F$ messbar, wobei $pr_j : \mathbb{R}^n \longrightarrow \mathbb{R}$ die - offenbar - stetige Projektion auf die j-te Komponenten ist.

Beweis. Es reicht zu zeigen, daß die Urbilder von Rechtecken meßbar sind. Sei $R = (a_1, b_1) \times \ldots \times (a_N, b_N)$ ein solches Rechteck. Dann gilt

$$F^{-1}(R) = \bigcap_{j=1}^{N} f_j^{-1}(a_j, b_j)$$

und das ist meßbar als endlicher Schnitt meßbarer Mengen. □

FOLGERUNG. *Sei $(X, \mathcal{A})$ ein meßbarer Raum und seien $f, g : X \longrightarrow \mathbb{R}$ meßbar. Dann sind $f + g$ und fg meßbar.*

Beweis. Wir betrachten nur fg. Der Fall $f + g$ kann analog behandelt werden. Sei

$$\phi : \mathbb{R}^2 \longrightarrow \mathbb{R}, \phi(u, v) = uv.$$

Dann ist ϕ stetig und damit meßbar. Sei weiterhin

$$F : X \longrightarrow \mathbb{R}^2, F(x) = (f(x), g(x)).$$

Dann ist F nach dem vorangehenden Theorem meßbar. Damit ist

$$fg = \phi \circ F$$

als Verknuepfung meßbarer Funktionen wieder meßbar. □

Zum Abschluss dieses Abschnittes führen wir noch eine besonders schoene Klasse von meßbaren Funktionen ein.

DEFINITION (Einfache Funktionen). *Sei $(X, \mathcal{A})$ ein meßbarer Raum. Eine meßbare Funktion $s : X \longrightarrow \mathbb{R}$ heißt einfach, wenn ihr Wertebereich endlich ist.*

Bemerkung. Offenbar bilden die einfachen Funktionen eine Algebra. Tatsaechlich ist dies die kleinste Algebra, die alle characeristischen Funktionen von messbaren Mengen enthaelt.

Fuer uns wird folgende Darstellung einer einfachen Funktion von besonderem Interesse sein: Sind $\alpha_1, \ldots, \alpha_n$ die verschiedenen Werte der einfachen Funktion s, so sind die Urbilder $A_j := \{x \in X : s(x) = \alpha_j\}$ (Niveaumengen) meßbar und disjunkt mit Vereinigung X, und es gilt

$$s = \sum_{j=1}^{n} \alpha_j 1_{A_j}.$$

Das nennen die die *Standarddarstellung* von s.

Bemerkung. Sind umgekehrt $A_1, \ldots, A_n$ meßbar (und nicht notwendig disjunkt) und $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$, so ist $\sum \alpha_j 1_{A_j}$ einfach. (Tatsaechlich ist diese Funktion meßbar als Summe meßbarer Funktionen und nimmt (offenbar) nur endlich viele Werte an.)

Meßbare Funktionen koennen gut durch einfache Funktionen approximiert werden.

THEOREM (Approximation durch einfache Funktionen). *Sei $(X, \mathcal{A})$ ein meßbarer Raum und $f : X \longrightarrow [0, \infty]$ meßbar. Dann existieren einfache Funktionen s_n, $n \in \mathbb{N}$, mit folgenden Eigenschaften:*

- $0 \leq s_1 \leq ...s_n \leq s_{n+1} \leq ... \leq f$.
- $s_n(x) \to f(x)$ *fuer alle* $x \in X$.

Bemerkung. (Uebung) Ist $f : X \longrightarrow \mathbb{R}$ meßbar und beschraenkt, so laesst es sich gleichmaessig durch einfache Funktionen approximieren.

Beweis. Sei fuer $n \in \mathbb{N}$ die Funktion $\phi_n : [0, \infty] \longrightarrow [0, \infty)$ definiert durch

$$\phi_n(t) = n, \text{ falls } n \leq t$$

und

$$\phi_n(t) = \frac{k}{2^n} \text{ falls } k/2^n \leq t < \tfrac{(k+1)}{2^n} \text{ fuer ein } k \in \mathbb{N} \text{ mit } 0 \leq k < 2^n n.$$

Dann nehmen die ϕ_n nur endlich viele Werte an und sind meßbar, und es gilt:

- $\phi_n(x) \leq \phi_{n+1}(x) \leq x$ fuer alle $x \in [0, \infty]$.
- $\phi_n(x) \to x$ fuer alle $x \in [0, \infty]$.

Damit folgt leicht, dass $s_n := \phi_n \circ f$ die gewuenschten Eigenschaften hat. $\square$

2. Maße und Integration positiver Funktionen

In diesem Abschnitt lernen wir Maße kennen. Das Konzept des Maßes verallgemeinert die Volumenmessung.

DEFINITION (Maß). *Sei $(X, \mathcal{A})$ ein meßbarer Raum. Eine Abbildung*

$$\mu : \mathcal{A} \longrightarrow [0, \infty]$$

heißt Maß, wenn gilt:

- $\mu(\emptyset) = 0$.
- $\mu(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mu(A_n)$ *fuer alle meßbaren paarweise diskunkten* A_n, $n \in \mathbb{N}$.

Ist μ ein Maß auf $(X, \mathcal{A})$, so nennt man das Tripel $(X, \mathcal{A}, \mu)$ einen Maßraum. Das Maß μ heißt endlich, wenn $\mu(X) < \infty$ gilt.

Bemerkungen.

- Die zweite Bedingung ist die entscheidenden Bedingung. Sie ist als *σ-Additivitaet* bekannt.
- Gegeben die zweite Bedingung, so ist die Bedingung $\mu(\emptyset) = 0$ aequivalent dazu, dass es eine Menge A in $\mathcal{A}$ gibt mit $\mu(A) < \infty$ (Uebung). Sie dient also nur dazu den trivialen Fall auszuschliessen.

Beispiel - Zaehlmaß Sei X eine beliebige Menge und $\mathcal{A}$ die Potenzmenge von X. Dann ist

$$\mu : \mathcal{A} \longrightarrow [0, \infty], \;\; \mu(E) = \sharp E$$

ein Maß (wobei $\sharp E$ die Anzahl der Element von E bezeichnet). Dieses Maß heisst das *Zaehlmaß* auf X. Auf abzaehlbaren Mengen (z.b. fuer $X = \mathbb{N}$)

ist das Zaehlmaß ein sehr natuerliches Maß. Alle Punkte haben dann das gleiche 'Gewicht'.

Beispiel - Lebesguemass

THEOREM (Lebesgue Maß). *Sei $\mathcal{B}$ die Borel - σ - Algebra auf $\mathbb{R}^N$. Dann gibt es ein eindeutiges Maß λ auf $\mathcal{B}$ mit*

$$\lambda((a_1, b_1) \times \ldots \times (a_N, b_N)) \prod_{j=1}^{N} |b_j - a_j|$$

fuer alle Rechtecke $R = (a_1, b_1) \times \ldots \times (a_N, b_N)$.

Bemerkung. Der Satz besagt, daß das naheliegende Volumen auf den Rechtecken (eindeutig) zu einem Maß fortgesetzt werden kann.

Beweis. Wir koennen hier (aus Zeitgruenden) nicht die Existenz eines Maßes λ mit der gewuenschten Eigenschaft zeigen. Die Aussage ist eigentlich eine Forsetzungsaussage. Man kann sie zum Beispiel mittels des Caratheodoryschen Fortsetzungsatzes zeigen. Wir diskutieren die uebrigen Aussagen.

λ ist eindeutig: Seien λ_1 und λ_2 Maße mit der angegbenen Eigenschaft. Dann ist

$$\{A \in \mathcal{B} : \lambda_1(A) = \lambda_2(A)\}$$

abgeschlossen unter Vereinigung aufsteigender Mengen und dem Schnitt von absteigenden Mengen (mit endlichem Maß). Weiterhin enthaelt das System die schnittstabile Menge aller Rechtecke. Damit gilt dann nach sogenannten Monotonen Klassen Argumenten $\lambda_1 = \lambda_2$ auf der (von den Rechtecken erzeugten) Borel-σ-Algebra. □

DEFINITION. *Es heißt λ das Lebesguemaß und das Integral bzgl. des Lebesguemaßes wird als Lebesgue-Integral bezeichnet.*

Das Lebesgue-Mass ist durch die Translationsinvarianz charakterisiert.

FOLGERUNG. *Sei μ ein Maß auf $\mathbb{R}^N$. Dann sind aequivalent:*

(i) $\mu = \lambda$.

(ii) *Es ist μ translationsinvariant (d.h. es gilt $\mu(t + V) = \mu(V)$ fuer alle $t \in \mathbb{R}^N$ und alle meßbaren V) mit $\mu((0, 1)^N) = 1$.*

Beweis. (i) $\Longrightarrow$ (ii). Wir muessen zeigen, dass λ die beiden angegebenen Eigenschaften hat: Es gilt offenbar $\lambda((0, 1)^N) = 1$. Um die Translationsinvarianz zu zeigen, betrachten wir fuer $t \in \mathbb{R}^N$ das Maß λ_t mit $\lambda_t(A) := \lambda(t+A)$. Dann hat λ_t die charakteristischen Eigenschaften des Lebesguemaß und muss also mit diesem uebereinstimmen aufgrund der schon diskutierten Eindeutigkeit.

(ii) $\Longrightarrow$ (i): Wir betrachten nur $N = 1$ d.h. $\mathbb{R}$. Es gilt $\mu(\{pt\}) = 0$ (da sonst aufgrund der Translationsinvarianz $\mu((0, 1)) = \infty$ gelten muesste.) Damit haengt also das Maß eines Interval nicht davon ab, ob die Randpunkte dazugehoeren oder nicht. Damit folgt aus der Tranlationsinvarianz also

$$\mu([0, 1/n]) = \frac{1}{n}$$

fuer alle $n \in \mathbb{N}$, da

$$[0,1) = \bigcup_{k=0}^{n} (\frac{k}{n} + [0, \frac{1}{n})).$$

Damit hat dann (wieder nach Translationsinvarianz) jedes Intervall der Laenge $1/n$ das Maß $1/n$ und dann jedes Intervall der Laenge k/n das Maß k/n. Durch Ausschopfen eines beliebigen Intervalles durch solche mit rationaler Laenge erhaelt man dann, dass das Maß eines Intervalles gerade sein Laenge ist. Damit folgt (aus dem vorangehenden Theorem) dann $\mu = \lambda$. □

Bemerkung. Die Folgerung ist ein Spezialfall eines allgemeinen Satzes fuer lokalkompakte abelsche Gruppen. Auf einer solchen Gruppe existiert ein (bis auf Normierung) eindeutiges translationsinvariantes Maß auf der Borel-σ-Algebra. Dieses Maß heißt das *Haarmaß* der Gruppe.

Beispiel einer nichtmeßbaren Menge. Betrachte $[0,1]$ mit der Einschraenkung des Lebesguemaß λ und die Aequivalenzrelation

$$x \sim y :\Longleftrightarrow x - y \in \mathbb{Q}.$$

Waehle nun aus jeder Aequivalenzuklasse einen Repraesentanten und nenne die entstehende Menge E. (Dieser Schritt nutzt das Auswahlaxiom!) Wir zeigen, dass dieses E nicht meßbar ist.
Definiere

$$E_r := \{x + r \mod 1 : x \in E\} \subset [0,1)$$

fuer $r \in \mathbb{Q}$. Wie man leicht sieht, gilt dann folgendes:

- $E_r \cap E_s = \emptyset$ fuer $r, s \in \mathbb{Q}$ mit $r \neq s \mod 1$.
- $[0,1) = \bigcup_{r \in \mathbb{Q} \cap [0,1)} E_r$.

Waere E meßbar, so waeren auch alle E_r meßbar und haetten das gleiche Lebesguemaß wie E (aufgrund der Translationsinvarianz). Damit erhielte man in diesem Fall aus den beiden Punkten

$$1 = \lambda([0,1)) = \sum_{r \in \mathbb{Q} \cap [0,1)} \lambda(E_r) = \sum_{r \in \mathbb{Q} \cap [0,1)} \lambda(E).$$

Wir unterscheiden nun zwei Faelle:

$\lambda(E) = 0$*:* Dann gilt also $\lambda([0,1)) = 0$. Das ist ein Widerspruch.
$\lambda(E) > 0$*:* Dann gilt also $\lambda([0,1)) = \infty$. Das ist ein Widerspruch.

Bemerkung (Uebung). Aehnlich kann man zeigen, dass jedes $A \subset \mathbb{R}$ mit positivem Lebesguemass eine nichtmeßbare Teilmenge hat. (Umgekehrt hat natuerlich ein $A \subset \mathbb{R}$ mit Lebesguemaß Null nur meßbare Teilmengen bzgl. der Vervollstaendigung der Borel-σ-Algebra).

Das beendet unsere Diskussion des Beispieles Lebesguemass.

Bemerkung - Warum wir σ-Algebren brauchen. Das obige Beispiel zeigt, dass man auch fuer so 'schoene' Raeume wie den Euklidischen Raum im allgemeinen keine Theorie entwicklen kann, in der jeder Menge in σ-additiver Weise ein translationsinvariantes Volumen zugeordnet wird.
Tatsaechlich laesst sich sogar im Euklidischen Raum der Dimension mindestens drei nicht einmal in additiver Weise ein translations- und drehungsinvariantes Volumen zuordnen (Banach-Tarski Paradoxon). Diese Sachverhalte

sind - in gewissem Sinne - der tiefere Grund dafuer, dass man ueberhaupt eine Theorie der σ-Algebren und messbaren Mengen entwickeln muss.

Proposition (Einfache Eigenschaften des Integrals). *Sei μ ein Maß auf $(X, \mathcal{A})$.*
(a) Sind $A_1, \ldots, A_n$ meßbar und disjunkt, so gilt

$$\mu(\bigcup_{j=1}^{n} A_j) = \sum_{j=1}^{n} \mu(A_j).$$

(b) Sind A, B meßbar mit $A \subset B$, so gilt

$$\mu(B) = \mu(A) + \mu(B \setminus A)$$

und insbesondere $\mu(A) \leq \mu(B)$.

Beweis. (a) Setze $A_{n+1} = A_{n+2} = ... = \emptyset$ und nutze σ-Additivitaet und $\mu(\emptyset) = 0$.
(b) Das folgt aus (a) mit $A_1 = A$ und $A_2 = B \setminus A$. □

Proposition. *Sei μ ein Maß auf $(X, \mathcal{A})$. Sind A_n, $n \in \mathbb{N}$, messbar so gilt*

$$\mu(\bigcup_{n} A_n) \leq \sum \mu(A_n).$$

Beweis. Setze $B_1 := A_1$ und $B_n := A_n \setminus \cup_{k=1}^{n-1} A_k$ fuer $n \geq 2$. Dann sind die B_n paarweise disjunkt mit $B_n \subset A_n$ und $\cup_{k=1}^{n} A_k = \cup_{k=1}^{n} B_k$. Damit folgt die Aussage einfach. □

p

Proposition (Kleiner Grenzwertsatz). *Sei μ ein Maß auf $(X, \mathcal{A})$. Seien $A_1 \subset A_2 \subset ...$ meßbar. Dann gilt fuer $A = \bigcup_n A_n$ die Beziehung*

$$(*) \;\; \mu(A) = \lim_{n} \mu(A_n).$$

Sind $A_1 \supset A_2 \supset ...$ meßbar mit $\mu(A_1) < \infty$. Dann gilt fuer $A = \bigcap_n A_n$

$$\mu(A) = \lim_{n} \mu(A_n).$$

Bemerkung. (a) Ist μ additiv, so ist σ-Additivitaet aequivalent zu $(*)$. In diesem Sinne is $(*)$ eine fundamentale Eigenschaft des Maßes.
(b) Setzt man $A = \lim A_n$ (was in beiden Faellen naeheliegt), so besagt die Proposition gerade

$$\mu(\lim_{n} A_n) = \lim_{n} \mu(A_n).$$

In diesem Sinne handelt es sich um eine Stetigkeitseigenschaft von μ.
(c) Die Voraussetzung $\mu(A_1) < \infty$ in der zweiten Behauptung ist notwendig. Betrachte zum Beispiel $\mathbb{N}$ mit dem Zaehlmaß und den meßbaren Mengen $A_n = [n, \infty)$, $n \in \mathbb{N}$. Dann gilt $\mu(A_n) = \infty$ aber $\mu(A) = 0$ (da $A = \emptyset$).

Beweis. Erste Aussage: Setzte $B_1 := A_1$ und $B_n := A_n \setminus A_{n-1}$ fuer $n \geq 2$. Dann gilt:

- $A_n = \bigcup_{k=1}^{n} B_k$ und damit auch $A = \bigcup_{k=1}^{\infty} B_k$.
- Die B_n sind paarweise disjunkt und meßbar.

Damit folgt dann aus der σ-Additivitaet:

$$\begin{aligned} \mu(A) &= \mu(\bigcup_k B_k) \\ (\sigma - Additivitaet) &= \sum_{k=1}^{\infty} \mu(B_k) \\ &= \lim_n \sum_{k=1}^{n} \mu(B_k) \\ &= \lim_n \mu(A_n). \end{aligned}$$

Zweite Aussage: Das folgt aus der ersten Aussage nach Komplementbildung und Subtraktion des endlichen (!) $\mu(A_1)$: Setze $C_n := A_1 \setminus A_n$. Dann gilt $C_1 \subset C_2 \subset .....$ Sei $C = \bigcup C_n$. Dann gilt nach der ersten Aussage

$$\mu(C) = \lim_n \mu(C_n).$$

Weiterhin gilt

$$\mu(C_n) = \mu(A_1 \setminus A_n) = \mu(A_1) - \mu(A_n)$$

sowie

$$\mu(C) = \mu\left(\bigcup_n C_n\right) = \mu\left(A_1 \setminus \bigcap_n A_n\right) = \mu(A_1) - \mu\left(\bigcap_n A_n\right).$$

Damit folgt die Behauptung. □

Wir werden nun ein Integral auf den einfachen Funktionen einfuehren. Dabei ist die naheliegende Idee natuerlich,

$$\int \sum_{j=1}^{n} \alpha_j 1_{A_j} d\mu = \sum_{j=1}^{n} \alpha_j \mu(A_j)$$

zu setzen. Allerdings stellt sich dann die Frage, ob dies wohldefiniert ist. Daher gehen wir einen geringfuegig anderen Weg. Wir definieren zunaechst das Integral mithilfe der Standarddarstellung einer einfachen Funktion und zeigen dann, dass dies nicht von der Darstellung abhaengt.

Definition (Integral einfacher Funktionen). *Sei μ ein Maß auf $(X, \mathcal{A})$. Sei s eine einfache nichtnegative Funktion mit den verschiedenen Werten $\alpha_1, \ldots, \alpha_n$ und den zugehoerigen Niveaumengen $A_j = \{x \in X : s(x) = \alpha_j\}$, also $s = \sum_{j=1}^n \alpha_j 1_{A_j}$. Dann definiert man*

$$\int s d\mu = \int_X s d\mu = \sum_{j=1}^{n} \alpha_j \mu(A_j).$$

Bemerkung - Rechnen in $[0, \infty]$. In dieser Definition (wie auch an anderen Stellen) brauchen wir Arithmetik d.h. Addition und Multiplikation auf $[0, \infty]$. Diese werden so definiert:

$$a \cdot \infty = 0 \quad \text{fuer } a = 0, \;\; a \cdot \infty = \infty \text{ sonst.}$$

$$a + \infty = \infty \;\text{ fuer alle } a \in [0, \infty].$$

Fuer $a, b \in [0, \infty)$ werden $a + b, a \cdot b$ wie ueblich definiert. Mit diesen Definitionen laeßt sich dann problemlos addieren und multiplizieren. (Probleme treten erst dann auf, wenn man subtrahieren will. Zum Beispiel laesst sich aus $a + c = b + c$ im allgemeinen NICHT schliessen $a = b$ usw.)

PROPOSITION (Wohldefiniert - disjunkte Grundmengen). *Sei μ ein Maß auf $(X, \mathcal{A})$. Ist $s = \sum_{j=1}^{m} \beta_j 1_{B_j}$ eine einfache Funktion mit disjunkten (meßbaren) B_j und nichtnegativen β_j, $j = 1, \ldots, m$. Dann gilt*

$$\int s d\mu = \sum_{j=1}^{n} \beta_j \mu(B_j).$$

Beweis. Ohne Einschraenkung seien alle β_j von Null verschieden. (Ein Term mit $\beta_j = 0$ traegt zur gewuenschten Summe sowieso nicht bei.) Seien $\alpha_1, \ldots, \alpha_n$ die von Null verschiedenen Werte von s und $A_i = \{x \in X : s(x) = \alpha_i\}$ die zugehoerigen Niveaumengen. Dann setzen sich die A_i aus den B_j (disjunkt) zusammen, da die B_j disjunkt sind und das liefert die Behauptung. Hier ist die genaue Buchhaltung: Sei

$$S_k = \{j : \beta_j = \alpha_k\}$$

fuer $k = 1, \ldots, n$. Dann gilt (aufgrund der Disjunktheit der B_j)

$$A_k = \bigcup_{j \in S_k} B_j,$$

wobei die Vereinigung disjunkt ist. Damit folgt also

$$\mu(A_k) = \sum_{j \in B_k} \mu(B_j).$$

Das liefert insgesamt

$$\begin{aligned}
\sum_k \alpha_k \mu(A_k) &= \sum_k \alpha_k \left(\sum_{j \in S_k} \mu(B_j) \right) \\
&= \sum_k \sum_{j \in S_k} \beta_j \mu(B_j) \\
&= \sum_j \beta_j \mu(B_j).
\end{aligned}$$

Das beendet den Beweis. □

PROPOSITION (Einfache Eigenschaften des Integrals einfacher Funktionen). *Sei μ ein Maß auf $(X, \mathcal{A})$. Seien s, t einfache nichtnegative Funktionen.*
(a) Gilt $s \leq t$ so folgt $\int s d\mu \leq \int t d\mu$.
(b) Mit $\lambda > 0$ gilt $\int (s + \lambda t) d\mu = \int s d\mu + \lambda \int t d\mu$.

Beweis. Sei $s = \sum_j^n \alpha_j 1_{A_j}$ und $t = \sum_{k=1}^m \beta_k 1_{B_k}$ mit messbaren paarweise disjunkten A_j bzwl. B_k. Ohne Einschraenkung $n = m$ und $A_j = B_j$ (sonst Betrachten von $A_j \cap B_k$ und nutzen der Proposition zur Wohldefiniertheit). Nun folgen die Aussagen einfach. □

FOLGERUNG (Wohldefiniertheit). *Sei μ ein Maß auf $(X, \mathcal{A})$. Seine $A_1, \ldots, A_n$ messbar und $\alpha_1, \ldots, \alpha_n \geq 0$. Dann gilt fuer die einfache Funktion $h := \sum_{j=1}^n \alpha_j 1_{A_j}$*

$$\int h d\mu = \sum_{j=1}^n \alpha_j \mu(A_j).$$

Beweis. Das folgt mit Induktion nach der Anzahl der Summanden durch Anwenden von Teil (b) der vorigen Proposition. Genauer gilt mit $s = \sum_{j=1}^{n-1} \alpha_j 1_{A_j}$, $t = 1_{A_n}$ und $\lambda = \alpha_n$:

$$\begin{aligned}
\int h d\mu &= \int (s + \lambda t) d\mu \\
&= \int s d\mu + \lambda \int t d\mu \\
&= \int s d\mu + \alpha_n \mu(A_n) \\
\text{(Induktion)} \quad &= \sum_{j=1}^{n-1} \alpha_j \mu(A_j) + \alpha_n \mu(A_n) \\
&= \sum_{j=1}^{n} \alpha_j \mu(A_j).
\end{aligned}$$

Das beendet den Beweis. □

Nachdem wir jetzt das Integral fuer einfache Funktionen definiert haben, werden wir es nun fuer allgemeine messbare nichtnegative Funktionen durch einen Grenzprozess - genauer durch eine Supremumsbildung - definieren.

DEFINITION (Integral nichtnegativer Funktionen). *Sei μ ein Maß auf $(X, \mathcal{A})$. Fuer ein meßbares $f : X \longrightarrow [0, \infty]$ definiert man*

$$\int f d\mu := \int_X f d\mu := \sup_{0 \leq s \leq f,\, s: einfach} \int s d\mu$$

und nennt dies das Integral von f ueber X bzgl. μ. (Hier ist der Wert ∞ fuer das Integral erlaubt.)

Bemerkung. Aus der Definition ergeben sich sofort zwei nuetzliche Beobachtungen:

- Gilt $\int f d\mu < \infty$, so ist $\mu(\{x \in X : f(x) = \infty\}) = 0$. (Andernfalls: Setze $A := \{x \in X : f(x) = \infty\}$ und betrachte $s = n 1_A$ fuer $n \in \mathbb{N}$...)
- Sei $S := \{x \in X : f(x) \neq 0\}$. Dann gilt

$$\int f d\mu = 0 \Longleftrightarrow \mu(S) = 0.$$

  Bew: $\Longleftarrow$: Es gelte $\mu(S) = 0$. Betrachte eine beliebige einfache Funktion s mit $0 \leq s \leq f$. Dann gilt $s = \sum_{j=1}^n \alpha_j 1_{A_j}$ mit disjunkten A_j. Jedes A_j mit $\alpha_j \neq 0$ ist nun eine Teilmenge von S und erfuellt damit $\mu(A_j) \leq \mu(S) = 0$. Damit folgt sofort $\int s d\mu = 0$. Das liefert die gewuenschte Behauptung.

$\Longrightarrow$: Es gelte $\int f d\mu = 0$. Sei $S_n := \{x : f(x) \geq \frac{1}{n}\}$. Dann gilt $f \geq \frac{1}{n} 1_{S_n}$. Damit folgt dann

$$0 \leq \mu(S_n) = \int 1_{S_n} d\mu \leq n \int f d\mu = 0.$$

Damit folgt dann $\mu(S) = \mu(\bigcup_n S_n) = 0$.

Proposition (Monotonie des Integrals). *Sei μ ein Maß auf $(X, \mathcal{A})$. Seien $f, g : X \longrightarrow [0, \infty]$ meßbar mit $f \leq g$. Dann gilt $\int f d\mu \leq \int g d\mu$ und $\int \alpha f d\mu = \alpha \int f d\mu$ fuer alle $\alpha \geq 0$.*

Beweis. Das ist klar. □

Beispiel - Zaehlmass. Sei $\mathbb{N}$ zusammen mit der σ-Algebra der Potenzmenge und dem Zaehlmass ζ gegeben. Dann ist jede Funktion $f : \mathbb{N} \longrightarrow [0, \infty]$ messbar und es gilt

$$\int f d\zeta = \sum_{x \in X} f(x).$$

Bew. Sei s eine beliebige einfache Funktion mit $0 \leq s \leq f$. Seien $\alpha_0, \ldots, \alpha_n$ ihre Werte und $A_0, \ldots, A_n$ die entsprechenden Niveaumengen. Dann gilt wegen $s \leq f$ also

$$\int s d\zeta = \sum_j \alpha_j \zeta(A_j) \leq \sum_j \sum_{x \in A_j} f(x) = \sum_x f(x).$$

Weiterhin gilt fuer $s_n = 1_{[1,n]} f$ offenbar $0 \leq s_n \leq f$ sowie

$$\int s_n d\zeta = \sum_{j=1}^{n} f(j) \to \sum_{j=1}^{n} f(j).$$

Das beendet den Beweis.

Wir lernen nun DAS THEOREM zur Konvergenz von Integralen nichtnegativer Funktionen kennen. Die meisten der anschliessend diskutierten Aussagen sind Folgerungen aus diesem Theorem.

Theorem (Monotone Konvergenz / Satz von Beppo Levi). *Sei μ ein Maß auf $(X, \mathcal{A})$. Seien $f_n : X \longrightarrow [0, \infty]$ meßbar mit $f_n \leq f_{n+1}$ fuer alle $n \in \mathbb{N}$. Sei f der punktweise Grenzwert der f_n. Dann ist f meßbar und nichtnegativ und es gilt*

$$\int f d\mu = \lim_n \int f_n d\mu.$$

Bemerkung. Aufgrund von $f_n \leq f_{n+1}$ existiert der punktweise Grenzwert der f_n (moeglicherweise mit Wert ∞).

Beweis. Es ist f meßbar als punktweiser Grenzwert meßbarer Funktionen. Aufgrund der Monotonie des Integrals gilt

$$\int f_n d\mu \leq \int f_{n+1} d\mu \leq \int f d\mu.$$

Insbesondere existiert also

$$\alpha = \lim_n \int f_n d\mu$$

und es gilt

$$\alpha \leq \int f d\mu.$$

Noch zu zeigen $\alpha \geq \int f d\mu$: Es reicht zu zeigen

$$\alpha \geq \int s d\mu$$

fuer alle einfachen s mit $0 \leq s \leq f$. Tatsaechlich reicht es zu zeigen

$$\alpha \geq \int cs d\mu$$

fuer alle einfachen s mit $0 \leq s \leq f$ und alle $c \in [0,1)$.
Die *Idee* ist nun folgende: Wegen $f_n \nearrow f \geq s$ gilt $f_n(x) \geq cs(x)$ fuer genuegend grosse n fuer jedes feste $x \in X$. Haette man nun diese Aussage gleichmaessig in $x \in X$ waere man fertig. Da wir die Aussage nicht gleichmaessig in n haben, muessen wir etwas sortieren nach den Mengen der x, fuer die die Aussage fuer $n \in \mathbb{N}$ gilt. Das geschieht als naechstes.
Betrachte

$$E_n := \{x \in X : f_n(x) \geq cs(x)\}.$$

Dann gilt

$$E_1 \subset E_2 \subset \ldots \subset E_n \subset ...$$

und

$$X = \bigcup_n E_n.$$

($f(x) = 0$ liefert $x \in E_1$, $f(x) > 0$ liefert $x \in E_n$ fuer genuegend grosses n wegen $f_n(x) \to f(x)$.) Damit folgt aus der Monotonie des Integrals also

$$\begin{aligned} \alpha \leftarrow \int f_n d\mu &\geq \int f_n \cdot 1_{E_n} d\mu \\ &\geq c \int s \cdot 1_{E_n} d\mu \\ &\stackrel{!}{\to} c \int s d\mu. \end{aligned}$$

Nimmt man (fuer den Moment) die letzte Konvergenz als bewiesen an, so kann man schliessen

$$\alpha = \lim_n \int f_n d\mu \geq c \int s d\mu$$

fuer alle einfachen s mit $0 \leq s \leq f$ und alle $c \in (0,1)$. Damit folgt

$$\alpha \geq \int f d\mu.$$

Es bleibt die Aussage (!) zu beweisen: Wegen $X = \bigcup_n E_n$ und der schon (unter dem Namen 'kleiner Grenzwertsatz') bekannten Aussage zur Konvergenz von Maßen, reicht es zu zeigen, dass

$$\phi : \mathcal{A} \longrightarrow [0, \infty],\ \phi(E) = \int s 1_E d\mu$$

ein Mass ist. Letzteres wiederum folgt durch direkte Rechnung fuer einfache Funktionen: Zu zeigen ist

$$\int s \cdot 1_{\cup_n A_n} d\mu = \sum_n \int s 1_{A_n} d\mu,$$

falls A_n paarweise disjunkte messbare Mengen sind. Ohne Einschraenkung $s = 1_B$. Nun folgt die gewuenschte Aussage leicht aus der σ-Additivitaet von μ. $\square$

Bemerkung (Philosophie) Weiter oben haben wir den 'kleinen Grenzwertsatz' kennengelernt:

$$\mu(A) = \lim_n \mu(A_n)$$

fuer $A_1 \subset A_2 \subset ...$ mit $A = \bigcup A_n$. Dieser Satz ist natuerlich ein Spezialfall des vorangehenden Theorems (mit $f_n = 1_{A_n}$ und $f = 1_A$). Umgekehrt liegt er (fuer das Maß ϕ statt μ) dem Beweis des Theorems zugrunde. In diesen Sinne ist das Theorem nichts anderes als eine (extrem nuetzliche) Umformulierung des kleinen Grenzwertsatzes. Dieser Satz wiederum ist, wie oben schon diskutiert, im wesentlichen aequivalent zur σ-Additivitaet. Damit sind σ-Additvitaet, kleiner Grenzwertsatz und vorangehendes Theorem im wesentlichen nichts anderes als verschiedene Perspektiven auf dasselbe Geschehen.

Als Folgerung erhalten wir die Linearitaet des Integrals und viel mehr.

FOLGERUNG (Vertraeglichkeit mit abzaehlbaren Summen). *Sei μ ein Maß auf $(X, \mathcal{A})$. Seien $g_n : X \longrightarrow [0, \infty]$ meßbar. Dann gilt*

$$\int \sum_{n=1}^{\infty} g_n d\mu = \sum_{n=1}^{\infty} \int g_n d\mu.$$

Bemerkung. Auch hier sind (natuerlich) die auftretenden Summen im Sinne der Arithmetik in $[0, \infty]$ zu verstehen.

Beweis. Dies folgt offenbar aus dem vorangehenden Theorem (mit $f_N := \sum_{n=1}^{N} g_n$), wenn die Vertraeglichkeit des Integrals mit endlichen Summen gilt. (Umgekehrt liefert die Aussage natuerlich die Vertraeglichkeit des Integrals mit endlichen Summen.) Entsprechend geht es also genau darum, diese Vertaeglichkeit zu zeigen.

Entsprechend zeigen wir daher zunaechst $\int (f+g) d\mu = \int f d\mu + \int g d\mu$ fuer meßbare nichtnegative f, g. Nach dem Theorem zur Approximation meßbarer Funktionen existieren monoton wachsende Folgen einfacher nichtnegativer Funktionen (s_n) und (t_n) mit $s_n \to f$ und $t_n \to g$. Dann konvergiert also $(t_n + s_n)$ monoton wachsend gegen $f + g$. Damit folgt mit zweimaliger Anwendung des vorigen Theorem und der Linearitaet des Integrals fuer einfache Funktionen dann

$$\begin{aligned}
\int (f+g) d\mu &= \lim_n \int (s_n + t_n) d\mu \\
&= \lim_n (\int s_n d\mu + \int t_n d\mu) \\
&= \int f d\mu + \int g d\mu.
\end{aligned}$$

Damit koennen wir nun zur eigentlichen Aussage kommen: Wir setzen

$$f_N := \sum_{k=1}^{N} g_k.$$

Dann konvergieren die f_N monoton wachsend gegen $f = \sum_{k=1}^{\infty} f_k$. Damit folgt aus dem vorigen Theorem und der Vorueberlegung dann

$$\int f d\mu = \lim_N \int f_N d\mu = \lim_N \sum_{k=1}^{N} \int g_k d\mu = \sum_{k=1}^{\infty} \int g_k d\mu.$$

Das beendet den Beweis. □

Beispiel - Anwendung. Betrachte $\mathbb{N}$ versehen mit dem Zaehlmaß und $a : \mathbb{N} \times \mathbb{N} \longrightarrow [0, \infty]$. Dann gilt

$$\sum_i \sum_j a_{ij} = \sum_j \sum_i a_{ij}.$$

(Beweis: $f_i : \mathbb{N} \longrightarrow [0, \infty], f_i(j) = a_{ij} ...$)

THEOREM (Lemma von Fatou). *Sei μ ein Maß auf $(X, \mathcal{A})$. Seien $f_n : X \longrightarrow [0, \infty]$ meßbar. Dann gilt*

$$\int \liminf_n f_n d\mu \leq \liminf_n \int f_n d\mu.$$

Bemerkung. Im allgemeinen ist die Ungleichung strikt, sogar wenn die (f_n) punktweise konvergieren. Betrachte zum Beispiel $\mathbb{N}$ mit dem Zaehlmaß und $f_n = 1_{\{k\}}$ (fuehrt auf $0 \leq 1$) oder $f_n = 1_{\{k \geq n\}}$ (fuehrt auf $0 \leq \infty$). Insbesondere gibt es *keinen* Grenzwertsatz, der nur die punktweise Konvergenz der entsprechenden Funktionen voraussetzt. Stattdessen ist immer eine gewisse Form von Gleichmäßigkeit der Konvergenz noetig.

Beweis. Setze $g_k := \inf_{n \geq k} f_n$. Dann ist g_k meßbar und nichtnegativ. Nach Konstruktion konvergiert (g_k) monoton wachsend gegen $\liminf_n f_n$. Damit folgt aus dem Theorem ueber monotone Konvergenz

$$\int (\liminf_n f_n) d\mu = \lim_k \int g_k d\mu$$

fuer alle $k \in \mathbb{N}$. Offenbar gilt $g_k \leq f_k$ und wir erhalten also

$$\int \liminf_n f_n d\mu \leq \int f_k d\mu$$

fuer alle $k \in \mathbb{N}$. Bildet man auf der rechten (und der linken ;-) Seite den $\liminf$ ueber k, so folgt die gewuenschte Aussage. □

Wir kommen nun noch Erweiterung des im Beweis des Theorems zur monotonen Konvergenz Gedankens.

PROPOSITION (Maße mit Dichte). *Sei μ ein Maß auf $(X, \mathcal{A})$ und $f : X \longrightarrow [0, \infty]$ meßbar. Dann ist*

$$\phi : \mathcal{A} \longrightarrow [0, \infty], \phi(E) := \int f 1_E d\mu$$

ein Maß, und es gilt

$$\int g d\phi = \int g f d\mu$$

fuer alle meßbaren $g \geq 0$. Insbesondere gilt $\phi(E) = 0$ fuer alle E mit $\mu(E) = 0$.

Beweis. Φ *ist Maß.* Es ist die σ-Additivitaet zu zeigen. Seien also A_n, $n \in \mathbb{N}$, paarweise disjunkte messbare Mengen. Offenbar gilt

$$f 1_{\bigcup_n A_n} = \sum_{n=1}^{\infty} f 1_{A_n}.$$

Damit folgt die gewuenschte Aussage (mit $g_n = f 1_{A_n}$) leicht aus der Folgerung zum Theorem ueber monotone Konvergenz.

Integralformel '$\int g d\phi = \int g f d\mu$'. Diese folgt fuer $g = 1_A$ direkt aus der Definition und damit dann fuer einfache Funktionen aus der Linearitaet und dann fuer allgemeine meßbare Funktionen nach Grenzuebergang unter Nutzen des Satzes ueber Monotone Konvergenz.

Die letzte Aussage folgt einfach aus dem vorangehenden. $\square$

Bemerkung. Ein Maß ν mit $\nu(E) = 0$ falls $\mu(E) = 0$ heißt absolut stetig bzgl. μ. Die Proposition besagt also unter anderem, dass ϕ absolut stetig bzgl. μ ist. Tatsaechlich gilt auch die Umkehrung 'Satz von Radon-Nikodym' (falls μ σ-endlich ist, d.h. es meßbare Mengen A_n gibt mit $\mu(A_n) < \infty$ und $\bigcup A_n = X$). Ist in diesem Fall ν absolut stetig bzgl. μ, so existiert eine meßbare Funktion $f : X \longrightarrow [0, \infty]$ mit $\nu(E) = \int 1_E f d\mu$. Das werden wir spaeter beweisen.

Definition (Das Maß $f\mu$). *Sei μ ein Maß auf $(X, \mathcal{A})$ und $f : X \longrightarrow [0, \infty]$ meßbar und ϕ das Maß aus der vorigen Proposition. Dann heißt f die Dichte oder auch Ableitung des Maßes ϕ bzgl. μ und man definiert auch $f\mu := \phi$ und $\frac{d\phi}{d\mu} := f$.*

Fuer gewisse Situationen erweist es sich als praktisch auch **Einschraenkungen von σ-Algebren und Maßen auf meßbare Teilmengen** zur Verfuegung zu haben. Ist $(X, \mathcal{A}, \mu)$ ein Maßraum und $E \subset X$ meßbar, so ist - wie man leicht sieht -

$$\mathcal{A}_E := \{A \cap E : A \in \mathcal{A}\}$$

eine σ-Algebra auf E und

$$\mu_E : \mathcal{A}_E \longrightarrow [0, \infty], \;\; \mu_E(B) := \mu(B)$$

ein Maß. Es heisst $\mathcal{A}_E$ die Einschrankung von $\mathcal{A}$ auf E und μ_E die Einschraenkung von μ auf E.
Fuer meßbare $f : X \longrightarrow [0, \infty]$ gilt dann

$$\int f \cdot 1_E d\mu = \int f|_E d\mu_E.$$

(Das ist klar fuer $f = 1_A$ mit $A \in \mathcal{A}$ und folgt dann aufgrund der Linearitaet fuer allgemeine einfache Funktionen und dann aus monotoner Konvergenz fuer beliebige meßbare Funktionen.) Wir schreiben auch

$$\int_E f d\mu \text{ statt } \int f d\mu_E.$$

3. Integration komplexer Funktionen und $\mathcal{L}^1(X,\mu)$

Im vorangehenden Abschnitt haben wir die Integration nichtnegativer Funktionen kennengelernt. In diesem Abschnitt lernen wir die Integration von komplexwertigen Funktionen mit einer gewissen Beschraenktheitseigenschaft kennen. Damit haben wir dann die beiden gaengigen Varianten von Integrationstheorie kennengelernt. Eine aehnliche Situation ist uns auch in Analysis I schon begegnet bei der Summation von Folgen: Man hat eine ueberzeugende Theorie fuer Summation von Folgen mit nichtnegativen Gliedern und eine weitere ueberzeugende Theorie fuer absolut konvergente Folgen. Tatsaechlich sind entsprechende Betrachtungen ein Spezialfall unserer Erwaegungen hier (mit dem Raum der natuerlichen Zahlen versehen mit dem Zaehlmaß).

Definition. *Sei $(X,\mathcal{A},\mu)$ ein Maßraum. Dann definieren wir*

$$\mathcal{L}^1(X,\mathcal{A},\mu) = \mathcal{L}^1(X,\mu) = \mathcal{L}^1(X)$$

als die Menge der meßbaren Funktionen $f : X \longrightarrow \mathbb{C}$ mit

$$\int |f| d\mu < \infty.$$

Fuer ein solches f mit Realteil u und Imaginaerteil v (also $f = u + iv$) definieren wir dann

$$\int f d\mu := \int u_+ d\mu - \int u_- d\mu + i\int v_+ d\mu - i\int v_- d\mu.$$

Bemerkungen.

- Hier ist $\mathbb{C}$ mit der Borel-σ-Algebra versehen. Dann bedeutet Meßbarkeit von $f : X \longrightarrow \mathbb{C}$ gerade, dass die entsprechende Funktion
$$\widetilde{f} : X \longrightarrow \mathbb{R}^2, \widetilde{f}(x) = (\Re f(x), \Im f(x))$$
meßbar ist.
- Da f meßbar ist, sind dann auch $|f| = |\cdot| \circ f$ meßbar, sowie $\Re f, \Im f$ meßbar. Damit sind insbesondere die auftretenden Terme $u_\pm$ und $v_\pm$ meßbar und nichtnegativ. Daher existieren die entsprechenden Integrale. Tatsaechlich sind alle Integrale auf der rechten Seite der angegebenen Gleichungen endlich wegen $|u_\pm| \leq |u| \leq |f|$ und $|v_\pm| \leq |v| \leq |f|$ und der vorausgesetzten Endlichkeit des Integrals ueber $|f|$.
- Wenn das Integral linear sein soll, muss es so definiert werden, da offenbar gilt
$$f = u_+ - u_- + iv_+ - iv_-.$$

Beispiel. Sei $\mathbb{N}$ ausgestattet mit der σ-Algebra $\mathcal{P}$ aller Teilmengen und dem Zaehlmass ζ. Dann gilt

$$\mathcal{L}^1(\mathbb{N}, \mathcal{P}, \zeta) = \{x : \mathbb{N} \longrightarrow \mathbb{C} : \sum |x(n)| < \infty\}.$$

Es handelt sich also genau um die absolute konvergenten Reihen. Man setzt

$$\ell^1 := \ell^1(\mathbb{N}) := \mathcal{L}^1(\mathbb{N}, \mathcal{P}, \zeta).$$

Theorem ($\mathcal{L}^1$ ist Vektorraum). *Sei $(X, \mathcal{A}, \mu)$ ein Maßraum. Seien $f, g \in \mathcal{L}^1(X, \mathcal{A}, \mu)$ und $\alpha, \beta \in \mathbb{C}$ gegeben. Dann gehoert auch $\alpha f + \beta g$ zu $\mathcal{L}^1(X, \mathcal{A}, \mu)$ und es gilt*

$$\int (\alpha f + \beta g) d\mu = \alpha \int f d\mu + \beta \int g d\mu.$$

Weiterhin gilt

$$\int \Re f d\mu = \Re \int f d\mu \text{ und } \int \Im f d\mu = \Im \int f d\mu.$$

Sind f und g reell mit $f \leq g$, so folgt auch noch

$$\int f d\mu \leq \int g d\mu.$$

Beweis. *$\alpha f + \beta g \in \mathcal{L}^1(X, \mathcal{A}, \mu)$:* Es ist $\alpha f + \beta g$ meßbar als Linearkombination meßbarer Funktionen. Weiterhin gilt

$$|\alpha f + \beta g| \leq |\alpha||f| + \beta|g|.$$

Damit folgt die erste Aussage leicht.
Zur Gleichung: Uebung. Hinweis: Man zeige zunaechst die folgenden drei Aussagen:

- $\int f d\mu = \int \Re f d\mu + i \int \Im f d\mu$ fuer alle $f \in \mathcal{L}^1$.
- $\int (f + g) = \int f + \int g$ fuer reellwertige f, g in $\mathcal{L}^1$.
- $\int i f d\mu = i \int f d\mu$.

Damit ergibt sich dann die gewuenschte Aussage.

Zu den letzten beiden Aussagen: Vertraeglichkeit mit Imaginaer- und Realteil wurde mitbeweisen. Die Ungleichung folgt leicht aus der Linearitaet. □

Proposition (Dreiecksungleichung). *Sei $(X, \mathcal{A}, \mu)$ ein Maßraum. Dann gilt*

$$\left| \int f d\mu \right| \leq \int |f| d\mu$$

fuer alle $f \in \mathcal{L}^1(X, \mathcal{A}, \mu)$.

Beweis. Waehle $\alpha \in \mathbb{C}$ mit $|\alpha| = 1$ und $|\int f d\mu| = \alpha \int f d\mu$. Dann gilt

$$\begin{aligned}
\left|\int f d\mu\right| &= \alpha \int f d\mu \\
\text{(Linearitaet)} &= \int \alpha f d\mu \\
\text{(Linke Seite reell)} &= \int \Re(\alpha f) d\mu \\
(\Re(\alpha f) \leq |\alpha f| = |f|) &\leq \int |f| d\mu.
\end{aligned}$$

Das beendet den Beweis. □

Der wesentliche Grenzwertsatz zur Integration nichtnegativer Funktionen ist das Monotone - Konvergenz - Theorem. Der entsprechende Grenzwertsatz zur Integration in $\mathcal{L}^1$ ist das folgende Theorem. Es ist (natuerlich ;-) eine Folge aus dem Theorem ueber monotone Konvergenz in Form des Lemma von Fatou.

THEOREM (Dominierte Konvergenz / Satz von Lebesgue). *Sei $(X, \mathcal{A}, \mu)$ ein Maßraum. Seien $f_n : X \longrightarrow \mathbb{C}$, $n \in \mathbb{N}$, und $f : X \longrightarrow \mathbb{C}$ meßbare Funktionen mit $f_n(x) \to f(x)$, $n \to \infty$, fuer alle $x \in X$. Gibt es ein $g \in \mathcal{L}^1(X, \mathcal{A}, \mu)$ mit $|f_n| \leq g$ fuer alle $n \in \mathbb{N}$, so gehoert auch f zu $\mathcal{L}^1(X, \mathcal{A}, \mu)$ und es gilt*

$$\int |f_n - f| d\mu \to 0, n \to \infty$$

und damit insbesondere

$$\int f d\mu = \lim_n \int f_n d\mu.$$

Beweis. Es ist f meßbar als Grenzwert meßbarer Funktionen und es gilt $|f| \leq g$ (da $|f_n| \leq g$ und f der punktweise Grenzwert der f_n ist). Damit gehoert f zu $\mathcal{L}^1$.

Wir zeigen nun die Konvergenz: Die entscheidende **Idee** ist es, die Funktion

$$h_n := 2g - |f_n - f| \geq 0$$

zu betrachten.

Dann gilt $h_n \geq 0$ und $h_n \to 2g$ punktweise. Damit folgt aus dem Lemma von Fatou also

$$\begin{aligned}
\int 2g d\mu &= \int \lim h_n d\mu \\
\text{(Fatou)} &\leq \liminf_n \int h_n d\mu \\
&= \liminf_n \left(\int 2g d\mu - \int |f_n - f| d\mu \right) \\
\text{(Rechenregeln lim inf)} &= \int 2g d\mu - \limsup_n \int |f_n - f| d\mu.
\end{aligned}$$

Wegen $g \in \mathcal{L}^1$ ist $\int 2g d\mu$ endlich und wir koennen es auf beiden Seiten subtrahieren und erhalten

$$\limsup_n \int |f_n - f| d\mu \leq 0.$$

Da die auftretenden Integranden nichtnegativ sind, folgt

$$0 \leq \liminf_n \int |f_n - f| d\mu \leq \limsup_n \int |f_n - f| d\mu \leq 0$$

und wir erhalten

$$0 = \lim_n \int |f_n - f| d\mu.$$

Das ist gerade die erste Aussage. Aus der Dreiecksungleichung folgt dann die letzte Aussage. □

Wir **erinnern** nun kurz an das Konzept von (Halb)norm: Sei V ein Vektorraum über $\mathbb{C}$. Eine Abbildung

$$\| \cdot \| : V \longrightarrow [0, \infty)$$

heisst Halbnorm, wenn gilt:

- $\|\alpha v + \beta w\| \leq |\alpha| \|v\| + |\beta| \|w\|$ fuer alle $v, w \in V$, $\alpha, \beta \in \mathbb{C}$.
- $\|\alpha v\| = |\alpha| \|v\|$ fuer alle $v \in V$ und $\alpha \in \mathbb{C}$.

Strukturell ist eine Halbnorm ist also eine Abbildung nach $[0, \infty)$, die 'moeglichst viel' Vektorraumstruktur respektiert. Anschaulich kann man es sich als eine Art Laengenmessung vorstellen. Fuer eine Halbnorm gilt jedenfalls $\|0\| = 0$. Gilt auch die Umkehrung (d.h $\|v\| = 0 \Longrightarrow v = 0$), so spricht man von einer Norm.

Proposition. *Sei $(X, \mathcal{A}, \mu)$ ein Maßraum. Dann wird durch*

$$\| \cdot \|_1 : \mathcal{L}^1(X, \mathcal{A}, \mu) \longrightarrow [0, \infty), \;\; \|f\|_1 := \int |f| d\mu,$$

eine Halbnorm definiert.

Beweis. Das ist einfach. □

Bemerkungen. (a) Es ist $\| \cdot \|_1$ im allgemeinen keine Norm. Tatsaechlich gilt (Uebung): $\| \cdot \|_1$ ist eine Norm genau dann, wenn es keine meßbare N mit $\mu(N) = 0$ und $N \neq \emptyset$ gibt.
Insbesondere ist also $\| \cdot \|_1$ eine Norm, wenn μ das Zaehlmass ist.
(b) (Halbnorm auf $\mathcal{L}^1$ versus Norm auf L^1). Auf dem Raum $\mathcal{L}^1(X, \mathcal{A}, \mu)$ ist $\| \cdot \|_1$ nur eine Halbnorm. Das laesst sich durch Herausfaktorisieren der fast ueberall verschwindenden Funktionen beheben und dem werden wir im folgenden noch genauer nachgehen.

4. Nullmengen und fast überall gültige Eigenschaften

Wir gehen nun auf die (verschwindende) Rolle von Nullmengen in der Theorie ein.

Definition (Nullmengen und Gültigkeit μ - fast überall). *Sei* $(X, \mathcal{A}, \mu)$ *ein Maßraum.*
(a) Eine meßbare Menge N *heisst Nullemge, wenn* $\mu(N) = 0$ *gilt. Die Komplemente von Nullmengen heissen Mengen von vollem Maß.*
(b) Sei $\mathcal{P}$ *eine Eigenschaft, die ein Element von* X *haben kann. Dann gilt* $\mathcal{P}$ μ*-fast-überall (*μ*-f.ü. oder auch nur f.ü.), wenn es eine Nullmenge* N *gibt, soda3* $\mathcal{P}$ *fuer alle* $x \in X \setminus N$ *gilt.*

Bemerkungen.

- Natuerlich haengt das Konzept der Nullmenge bzw. der Gueltigkeit fast überall stark vom gegebenen Maß μ ab. Betrachtet man z. B. $X = \mathbb{N}$ mit dem Zaehlmaß, so gilt ein Eigenschaft fast-überall genau dann, wenn sie fuer alle Punkte gilt.
- Man beachte, daß nicht gefordert wird, daß die Menge
$$\{x : \ \mathcal{P} \text{ gilt nicht fuer } x\}$$
eine Nullmenge ist. Tatsaechlich kann es sein, daß diese Menge gar nicht meßbar ist (s.u.).

Beispiele fuer gaengige Eigenschaften $\mathcal{P}$ eines $x \in X$:

- $f(x) > 0$ (fuer gegebenes f).
- $f(x) = g(x)$ (fuer gegebene f und g).
- $f_n(x)$ konvergiert bzw. $f_n(x)$ konvergiert nicht (fuer gegebenen Folge (f_n)).

Lemma. *Sei* $(X, \mathcal{A}, \mu)$ *ein Maßraum und* $f : X \longrightarrow \mathbb{C}$ *messbar und* $S := \{x : f(x) \neq 0\}$*. Dann sind aequivalent:*

(i) *Es gilt* $\mu(S) =$*. verschwindet ausserhalb einer Nullmenge).*
(ii) *Es gilt* $\|f\|_1 = 0$.

In diesem Fall gilt natuerlich $\int f d\mu = 0$*. Insbesondere gilt fuer jedes messbare Funktion* f *und jede Nullmengen* N *auch*
$$\int |f| 1_N d\mu = 0.$$

Beweis. Es gilt offenbar $S = \{x : |f(x)| \neq 0\}$. Damit folgt aus der Bemerkung nach der Definition des Integrals ueber nichtnegative Funktionen dann
$$\mu(S) = 0 \Longleftrightarrow 0 = \int |f| d\mu = \|f\|_1.$$
Zu $\int f d\mu = 0$: Das folgt aus der Monotonie des Integrals und
$$0 \leq u_\pm, v_\pm \leq |f|$$
fuer $f = u + iv$.
Zur letzten Aussage: Sei $g := |f| 1_N$. Dann gilt fuer $S_g := \{x \in X : g(x) \neq 0\}$ natuerlich $S_g \subset N$. Damit gilt $\mu(S_g) = 0$ und es folgt aus der ersten Aussage des Lemma (mit S_g statt S) dann $\int |g| d\mu = 0$. □

Anwendung - Integration fast überall gleicher Funktionen. Sei $(X, \mathcal{A}, \mu)$ ein Maßraum. Dann wird durch
$$f \sim g \Longleftrightarrow \quad f = g \ \mu \text{ -f.ü.}$$

eine Aequivalenzrelation auf den meßbaren Funktionen definiert (wie man leicht sieht). Nach dem obigen Lemma gilt auch

$$f \sim g \Longleftrightarrow \int |f-g| d\mu = 0.$$

Insbesondere gehoert fuer f, g mit $f \sim g$ die Funktion f zu $\mathcal{L}^1$ genau dann, wenn g zu $\mathcal{L}^1$ gehoert (da $f = g + (f-g)$ bzw. $g = f + (g-f)$ und $(f-g)$ wegen $\int |f-g| d\mu = 0 < \infty$ zu $\mathcal{L}^1$ gehoert). In diesem Fall gilt

$$\int |f| d\mu = \int |g| d\mu$$

wegen

$$|\int |f| d\mu - \int |g| d\mu| \leq \int ||f| - |g|| d\mu \leq \int |f-g| d\mu = 0.$$

Aehnlich folgt mit der Dreiecksungleichung

$$\int f 1_E d\mu = \int g 1_E d\mu$$

fuer alle meßbaren E aufgrund von

$$|\int f 1_E d\mu - \int g 1_E d\mu| \leq \int |f-g| 1_E d\mu \leq \int |f-g| d\mu = 0.$$

Anwendung - Behandlung fast überall definierter Funktionen. Sei $(X, \mathcal{A}, \mu)$ ein Maßraum. Sei f f.ü. definiert d.h. es gebe eine Nullmenge N so daß f auf $X \setminus N$ definiert ist. Dann heißt f meßbar, wenn

$$f^{-1}(V) \cap (X \setminus N)$$

meßbar ist fuer alle meßbaren V. Dann ist (Uebung) f genau dann meßbar, wenn die Fortsetzung $\widetilde{f}$ von f auf ganz X durch Setze von Null auf N meßbar ist. Man ueberlege sich, dass diese Definition nicht davon abhaengt, welche der (unter Umstaenden mehreren) möglichen Nullmengen N man waehlt. Dann definiert man

$$\int f d\mu := \int_{X \setminus N} f d\mu.$$

Ebenso sagt man - etwas lax, dass ein fast ueberall definierte messbare Funktion f zu $\mathcal{L}^1(X, \mathcal{A}, \mu)$ gehoert, wenn fuer die fast ueberall definierte Funktion $|f|$ gilt

$$\int |f| d\mu < \infty,$$

(wobei das Integral im gerade definierten Sinne zu verstehen ist). Mit diesen Definitionen erhalten wir folgenden Satz.

THEOREM. *Sei $(X, \mathcal{A}, \mu)$ ein Maßraum. Seien $f_n : X \longrightarrow \mathbb{C}$ f.ü. definierte meßbare Funktionen mit*

$$\sum_{n=1}^{\infty} \int |f_n| d\mu < \infty.$$

Dann existiert $f(x) = \sum_{n=1}^\infty f_n(x)$ f.ü. (als absolut konvergente Reihe), und es gehoert f zu $\mathcal{L}^1(X,\mu)$, und es gilt

$$\int f d\mu = \sum_{n=1}^\infty \int f_n d\mu.$$

Beweis. Wir finden eine Nullmenge, ausserhalb derer 'alles funktioniert'. Seien S_n meßbare Mengen mit $\mu(X \setminus S_n) = 0$, so dass f_n auf S_n definiert ist fuer jedes $n \in \mathbb{N}$. Die Mengen S_n stellen dann also so eine Art Traeger (support) fuer f_n dar. Sei $S := \bigcap_n S_n$. Dann gilt

$$\mu(X \setminus S) = \mu(\bigcup_n X \setminus S_n) = 0.$$

Sei

$$\varphi : X \longrightarrow [0,\infty]$$

auf $X \setminus S$ durch 0 definiert und auf S durch

$$\varphi(x) = \sum_{n=1}^\infty |f_n(x)|.$$

Dann gilt nach dem Satz ueber monotone Konvergenz

$$\int \varphi d\mu = \int_S \varphi d\mu = \sum_{n=1}^\infty \int_S |f_n| d\mu = \sum_{n=1}^\infty \int |f_n| d\mu < \infty.$$

Aus der Endlichkeit folgt, dass

$$E := \{x \in X : \varphi(x) = \infty\}$$

das Maß Null hat. Dann ist also

$$G := S \cap (X \setminus E)$$

meßbar mit

$$\mu(X \setminus G) = \mu((X \setminus S) \cup E) = 0.$$

Auf G sind alle vorkommenden Funktionen definiert und es konvergiert φ absolut mit endlichen Werten. Insbesondere existiert also f auf G und es gilt $|f| \leq \varphi$ auf G und damit gehoert f zu $\mathcal{L}^1$ und nach dem Satz ueber dominierte Konvergenz folgt

$$\int f d\mu = \int_G f d\mu = \sum_{n=1}^\infty \int_G f_n d\mu = \sum_{n=1}^\infty \int f_n d\mu.$$

Das beendet den Beweis. □

Bemerkung. Ganz analog kann man 'fast-ueberall'-Versionen des Satzes von Lebesgue und des Satzes von Beppo-Levi beweisen.

Gilt $\mu(N) = 0$ so erwartet man natuerlich $\mu(E) = 0$ fuer alle $E \subset N$. Tatsaechlich kann man diese Beziehung zeigen fuer diejenigen $E \subset N$, die zu $\mathcal{A}$ gehoeren:

$$0 \leq \mu(E) \leq \mu(E) + \mu(N \setminus E) = \mu(N) = 0.$$

Fuer diejenigen $N \subset E$, die nicht zu $\mathcal{A}$ gehoeren, kann man es nicht zeigen, da ja $\mu(E)$ gar nicht definiert ist. Das ist unter Umstaenden unpraktisch.

Gluecklicherweise kann man aber Teilmengen von Nullmengen zur σ-Algebra hinzufuegen. Das ist der Inhalt des folgenden Theorem.

THEOREM (Vervollstaendigung der σ-Algebra bzgl. eines Maßes). *Sei* $(X, \mathcal{A}, \mu)$ *ein Maßraum. Sei* $\overline{\mathcal{A}}^{\mu}$ *die Familie aller Teilmengen* E *von* X *fuer die* $A, B \in \mathcal{A}$ *existieren mit*

- $A \subset E \subset B$,
- $\mu(B \setminus A) = 0$.

Dann ist $\overline{\mathcal{A}}^{\mu}$ *die kleinste* σ*-Algebra, die* $\mathcal{A}$ *enthaelt und jede Teilmenge einer* μ*-Nullmenge. Weiterhin wird durch*

$$\overline{\mu} : \overline{\mathcal{A}}^{\mu} \longrightarrow [0, \infty], \overline{\mu}(E) = \mu(A) = \mu(B)$$

ein Maß auf $\overline{\mathcal{A}}^{\mu}$*, das* μ *fortsetzt.*

Bemerkungen.

- Hat der Maßraum $(X, \mathcal{A}, \mu)$ die Eigenschaft, dass jede Teilmenge einer Nullmenge wieder meßbar ist, so heißt er *vollstaendig*. Es heißt $\overline{\mathcal{A}}^{\mu}$ die *Vervollstaendigung* von $\mathcal{A}$ bzgl. μ und $\overline{\mu}$ die Vervollstaendigung von μ.
- Es ist $(X, \mathcal{A}, \mu)$ vollstaendig genau dann, wenn $\mathcal{A} = \overline{\mathcal{A}}^{\mu}$ gilt (aufgrund der Minimalitaet).
- Das Theorem besagt, dass wir ohne Einschraenkung annehmen koennen, daß der Maßraum des Definitionsbereiches der Funktionen vollstaendig ist.
- **!!! Vorsicht** bei der Vervollstaendigung des Bildraumes der Funktionen: Es kann pasieren, dass eine meßbare Funktion nach Vervollstaendigen des Bildraumes nicht mehr meßbar ist (da man nun mit viel mehr Mengen testen muss).

Beweis. (Nicht in der Vorlesung gegeben) Wir zeigen eine Reihe von Behauptungen, die zusammengenommen das Theorem liefern.

Es ist $\overline{\mu}$ *wohldefiniert:* Seien $A \subset E \subset B$ und $A' \subset E \subset B'$ mit A, A', B, B' aus $\mathcal{A}$ und $\mu(B \setminus A) = 0 = \mu(B' \setminus A')$ gegeben. Dann gilt

$$A \setminus A' \subset E \setminus A' \subset B' \setminus A'.$$

Damit folgt

$$\mu(A \setminus A') \leq \mu(B' \setminus A') = 0,$$

und es ergibt sich

$$\mu(A) = \mu(A \cap A') + \mu(A \setminus A') = \mu(A \cap A').$$

sowie ganz analog durch Vertauschen von A und A'

$$\mu(A') = \mu(A \cap A').$$

Somit erhalt man insgesamt

$$\mu(A) = \mu(A').$$

Es ist $\overline{\mathcal{A}}^{\mu}$ *eine* σ*-Algebra:* Wir muessen drei Eigenschaften nachweisen.

- X gehoert zu $\overline{\mathcal{A}}^{\mu}$: Waehle $A = X = B$.

- Mit E gehoert auch $X \setminus E$ zu $\overline{\mathcal{A}}^{\mu}$: Das folgt direkt durch Komplementbildung an allen Stellen: Seien A, B aus $\mathcal{A}$ mit $A \subset E \subset B$ und $\mu(B \setminus A) = 0$. Dann gehoeren B^c, A^c zu $\mathcal{A}$ und es gilt $B^c \subset E^c \subset A^c$ sowie $\mu(A^c \setminus B^c) = \mu(B \setminus A) = 0$ (da $A^c \setminus B^c = B \setminus A$).
- Mite E_n, $n \in \mathbb{N}$,gehoert auch $\bigcup_n E_n$ zu $\overline{\mathcal{A}}^{\mu}$: Das ist einfach.

Es enthaelt $\overline{\mathcal{A}}^{\mu}$ sowohl $\mathcal{A}$ als auch alle Teilmengen von μ-Nullmengen. Fuer E aus $\mathcal{A}$ kann man waehlen $A = E = B$. Fuer eine Teilmenge E einer Nullmenge N kann man waehlen $A = \emptyset$ und $B = N$.

Enthaelt die σ-Algebra $\mathcal{B}$ sowohl $\mathcal{A}$ als auch alle Teilmengen von μ-Nullmengen, so enthaelt sie $\overline{\mathcal{A}}^{\mu}$: Nach Konstruktion kann man jedes Element E aus $\overline{\mathcal{A}}^{\mu}$ darstellen als

$$E = A \bigcup S$$

mit A aus $\mathcal{A}$ und $S = E \setminus A \subset B \setminus A$ Teilmenge einer Nullmenge.

Es ist $\overline{\mu}$ σ-additiv: Das ist einfach. □

5. Der Satz von Fubini-Tonelli

Seien X und Y Mengen und $\mathcal{A}_X$ eine σ-Algebra auf X und $\mathcal{A}_Y$ eine σ-Algebra auf Y. Dann erzeugen die Mengen der Form $A \times B$ mit $A \in \mathcal{A}_X$ und $B \in \mathcal{A}_Y$ eine σ-Algebra auf $X \times Y$, die wir als die *Produkt-σ-Algebra* $\mathcal{A}_X \times \mathcal{A}_Y$ bezeichnen. Sind weiterhin μ_X und μ_Y Maße auf $\mathcal{A}_X$ bzw. $\mathcal{A}_Y$, so gibt es ein eindeutiges Mass μ auf Produkt-σ-Algebra mit

$$\mu(A \times B) := \mu_X(A)\mu_Y(B)$$

fuer beliebige $A \in \mathcal{A}_X$ und $B \in \mathcal{A}_Y$. Dieses Mass heisst das *Produktmass.*

THEOREM. *Seien X, Y, $\mathcal{A}_X, \mathcal{A}_Y$ und μ_X und μ_Y wie oben. Sei $f : X \times Y \longrightarrow \mathbb{C}$ gegeben. Dann sind aequivalent:*

(i) *Es gehoert f zu $\mathcal{L}^1(X \times Y, \mu)$.*
(ii) *Es ist f μ-messbar und $|f(x, \cdot)|$ gehoert fuer μ_X fast alle $x \in X$ zu $\mathcal{L}^1(Y, \mu_Y)$ und die Funktion $F : X \longrightarrow \mathbb{C}$ mit $F(x) = \int_Y |f(x, y)| \, d\mu(y)$, falls $|f(x, \cdot)| \in \mathcal{L}^1(Y, \mu_Y)$, und $F(x) = 0$ sonst, gehoert zu $\mathcal{L}^1(X, \mu_X)$.*

In diesem Fall gilt

$$\int_{X \times Y} f(x, y) \, d\mu(x, y) = \int_X \left(\int_Y f(x, y) \, d\mu_Y(y) \right) \, d\mu_X(x).$$

Bemerkung.

- Natuerlich gilt entsprechendes, wenn die Rollen von X und Y vertauscht werden.
- Ist $f \geq 0$ messbar, so liefert der Satz, dass auf jeden Fall gilt

$$\int f(x, y) d\mu(x, y) = \int_X \int_Y f(x, y) d\mu_Y(x) d\mu_X(x),$$

  wobei allerdings beide Seiten ∞ sein koennen.

Der *Beweis* kann mit den zur Verfügung stehenden Methoden geführt werden, wuerde aber recht viel Zeit kosten (ohne über die Aussage hinausgehenden wesentlichen Erkentnisgewinn zu liefern). Daher geben wir ihn hier nicht. Wir bemerken stattdessen, dass (i)$\Longrightarrow$ (ii) als *Satz von Fubini* bekannt ist und (ii)$\Longrightarrow$ (i) als *Satz von Tonelli*.

KAPITEL 2

Die L^p-Raeume

Soweit es um Integration geht, koennen Funktionen f und g, die fast ueberall uebereinstimmen, nicht unterschieden werden. Das legt es nahe, eine Integrationstheorie fuer die Klassen zu entwickeln. Das passt auch zu unserer Beobachtung im vorigen Abschnitt, dass man erst nach Identifizieren eine Norm auf den entsprechenden Raeumen von Funktionen erhaelt. Hier untersuchen wir das systematisch. Die vorgestellten L^p-Raeume sind wesentliche Beispiele fuer normierte Räume.

Auf dem Vektorraum der meßbaren Funktionen wird durch

$$f \sim g :\Longleftrightarrow \ f = g \ \mu \text{ fast ueberall}$$

eine Aequivalenzrelation definiert. Man definiert fuer ein meßbares f die Klasse von f als die Menge der messbaren g die fast ueberall mit f uebereinstimmen. Man schreibt $[f]$ fuer diese Klasse. Die Mengen der Klassen bilden einen Vektorraum mit

$$[f] + [g] = [f + g], \ \alpha[f] = [\alpha f]$$

fuer $\alpha \in \mathbb{C}$ und f, g messbar. Dabei ist natuerlich nachzupruefen, dass diese Abbildungen wohldefiniert sind und tatsaechlich die Vektorraumaxiome erfuellen.
Alternativ laesst sich die Aequivalenzrelation $\sim$ und die zugehoerigen Klassen auch mit dem Unterraum

$$\mathcal{N} := \{f : \ f \text{ messbar mit } f = 0 \ \mu \text{ fast ueberall}\}$$

beschreiben. Es gilt naemlich

$$f \sim g \Longleftrightarrow f - g \in \mathcal{N}$$

bzw.

$$[f] = f + \mathcal{N}.$$

Das liefert aufgrund von allgemeiner Theorie ueber das Faktorisieren nach Unterraeumen dann direkt, dass die Klassen einen Vektorraum bilden bzgl. der oben angegebenen Verknuepfungen.

Es gehoert f zu $\mathcal{L}^1$ genau dann, wenn ein / alle Elemente seiner Klasse zu $\mathcal{L}^1$ gehoeren. Man definiert

$$L^1(X) := L^1(X, \mu) := L^1(X, \mathcal{A}, \mu) := \mathcal{L}^1(X, \mathcal{A}, \mu)/\sim = \mathcal{L}^1(X, \mathcal{A}, \mu)/\mathcal{N}$$

sowie

$$\int : L^1(X, \mu) \longrightarrow \mathbb{C}, \int [f] d\mu := \int f d\mu.$$

Nach den vorhergehenden Ueberlegungen ist das wohldefiniert. Weiterhin gilt offenbar

$$\mathcal{N} = \{h \in \mathcal{L}^1(X, \mathcal{A}, \mu) : \|h\|_1 = 0\}.$$

Daher wird $L^1(X, \mu)$ mit

$$\|[f]\|_1 := \int |f| d\mu$$

zu einem normierten Raum.
Wir werden sehen, dass dieser Raum vollstaendig ist. Hier ist das entscheidende Lemma.

LEMMA. *Sei $(X, \mathcal{A}, \mu)$ ein Maßraum und (f_n) eine Cauchy-Folge in $\mathcal{L}^1(X, \mathcal{A}, \mu)$. Dann gibt es ein $f \in \mathcal{L}^1(X, \mathcal{A}, \mu)$ mit folgenden beiden Eigenschaften:*

- *Die Folge (f_n) hat eine Teilfolge, die μ fast ueberall gegen f konvergiert.*
- $\|f_n - f\|_1 \to 0$, $n \to \infty$.

Bemerkung. (a) Das im Lemma gegebene f ist bis auf eine Nullmenge eindeutig - und mehr laesst sich auch nicht erwarten.
(Bew: Sei g eine weitere solche Funktion. Dann gilt:

$$\|f - g\|_1 \leq \|f - f_n\|_1 + \|f_n - g\|_1 \to 0, n \to \infty.$$

Damit stimmen also f und g bis auf eine Nullmenge ueberein.)
(b) Ist h eine beliebige Funktion, fuer die eine Teilfolge (f_{n_k}) existiert mit $f_{n_k} \to h$ punktweise fast ueberall, so folgt $h = f$ fast ueberall. Damit ist also der Grenzwert eindeutig dadurch bestimmt, dass man eine fast ueberall dagegen konvergente Teilfolge findet.
(Bew: Anwenden des Lemma auf (f_{n_k}) liefert eine Funktion f' mit $\|f_{n_k} - f'\|_1 \to 0$ und eine Teilfolge $f_{n_{k_l}}$, die punktweise gegen f' konvergiert. Da die Teilfolge $f_{n_{k_l}}$ natuerlich weiterhin punktweise fast sicher gegen h konvergiert, folgt $h = f'$ fast ueberall. Da f_{n_k} weiterhin natuerlich bgzl. $\|\cdot\|_1$ gegen f konvergiert folgt (wie in (a)) dann auch $f' = f$ fast ueberall. Insgesamt ergibt sich also $f = f' = h$ fast ueberall.)
(c) Die Folge (f_n) selber wird im allgemeinen nicht fast ueberall konvergieren, wie man leicht an Beispielen sieht. Konkret etwa auf $[0, 1]$ mit Lebesguemass die charakteristische Funktionen von Intervallen der Form $[\frac{k}{2^n}, \frac{k+1}{2^n}]$ fuer $n = 1, 2, ...$ und $k = 0, \ldots, 2^n$, betrachten. Dann gilt fuer alle $x \in [0, 1]$ also $\liminf f_n(x) = 0$ und $\limsup f_n(x) = 1$ sowie offenbar $\|f_n\|_1 \to 0$.
Beweis. Sei (f_n) eine Cauchy-Folge bzgl. $\|\cdot\|_1$. Die wesentliche **Idee** ist es eine Teilfolge (f_{n_k}) zu waehlen fuer die $\|f_{n_k} - f_{n_{k+1}}\|_1$ sehr klein ist. Dann wird also $|f_{n_k}(x) - f_{n_{k+1}}(x)|$ fuer die meisten x klein sein. Damit ist dann $\sum_k |f_k(x) - f_{k+1}|(x) < \infty$ und es existiert entsprechend

$$f(x) = \lim_k f_{n_k} = \lim_k (f_{n_1} + \sum_{l=1}^{k-1} (f_{n_{l+1}} - f_{n_l}).$$

Genauer waehlen wir eine Teilfolge (f_{n_k}) mit

$$\|f_{n_k} - f_{n_{k+1}}\|_1 \leq \frac{1}{3^{k+1}}.$$

Setze

$$g_k := |f_{n_k} - f_{n_{k+1}}|$$

und

$$M_k := \{x : g_k(x) > \frac{1}{2^{k+1}}\},$$

Dann gilt

$$\mu(M_k) = \int 1_{M_k} d\mu \leq \int 2^{k+1} g_k d\mu \leq \left(\frac{2}{3}\right)^{k+1}.$$

Fuer $x \in X$, die zu keinem der M_l gehoeren, ist dann also

$$\sum_{l=1}^{\infty} |f_{n_l}(x) - f_{n_{l+1}}(x)| \leq \sum_l \frac{1}{2^{l+1}} \infty.$$

Da es bei absoluter Konvergenz der Reihe nicht auf endlich viele Terme ankommt, reicht es natuerlich, dass die x nicht zu M_l gehoeren fuer $l \geq k$. Entsprechend setzen wir

$$N_k := \bigcup_{l \geq k} M_k$$

und halten fest, dass

$$\sum_{l=1}^{\infty} |f_{n_l}(x) - f_{n_{l+1}}(x)| < \infty$$

fuer $x \notin N_k$ fuer ein $k \in \mathbb{N}$. Entsprechend erhalten wir absolute Konvergenz fuer alle x, die nicht zu

$$N_\infty := \bigcap_k N_k$$

gehoeren. Wir zeigen nun, dass die Masse der N_k gegen 0 konvergieren: Es gilt

$$\mu(N_k) \leq \sum_{l \geq k} \mu(M_l) \leq \sum_{l \geq k} \left(\frac{2}{3}\right)^l \to 0, k \to \infty.$$

Insbesondere folgt aus $\mu(N_\infty) \leq \mu(N_k)$ also

$$\mu(N_\infty) = 0.$$

Wir setzen

$$X' := X \setminus N_\infty = \bigcup_k X \setminus N_k.$$

Aufgrund der vorangegangenen Diskussion ist also

$$\sum_{l \geq 1} (f_{n_{l+1}} - f_{n_l})$$

auf X' absolut konvergent. Wegen

$$f_{n_k}(x) := f_{n_1}(x) + \sum_{l=1}^{k-1} (f_{n_{l+1}} - f_{n_l})$$

existiert dann also

$$\lim_k f_{n_k}(x)$$

fuer alle $x \in X'$. Definiere $f : X \longrightarrow \mathbb{C}$ durch $f(x) = \lim_k f_{n_k}(x)$ fuer $x \in X'$ und $f(x) = 0$ sonst.
Damit konvergiert dann also die Teilfolge (f_{n_k}) punktweise fast ueberall gegen f. Weiterhin gilt

$$\begin{aligned}
\|f - f_{n_k}\|_1 &= \int |f - f_{n_k}| d\mu \\
(X \setminus X' \text{ Nullmenge}) &= \int 1_{X'} |f - f_{n_k}| d\mu \\
(f_{n_k} \to f \text{ f. u.}) &= \int (\lim_l 1_{X'} |f_{n_l} - f_{n_k}| d\mu \\
(Fatou) &\leq \liminf_l \int 1_{X'} |f_{n_l} - f_{n_k}| d\mu \\
&= \liminf_l \|f_{n_l} - f_{n_k}\| \\
&\to 0, k \to \infty.
\end{aligned}$$

Da (f_n) eine Cauchy-Folge ist, folgt dann aus

$$\|f - f_n\| \leq \|f - f_{n_k}\| + \|f_{n_k} - f_n\|$$

die gewuenschte Konvergenz $\|f - f_n\| \to 0, n \to \infty$. □

Bemerkung. Der Beweis zeigt sogar die μ-fast gleichmaessige Konvergenz der Teilfolge (d.h. fuer jedes $\varepsilon > 0$ existiert eine messbare Ausnahmemenge A_ε mit $\mu(A_\varepsilon) \leq \varepsilon$, so dass die Teilfolge auf dem Komplement von A_ε gleichmaessig konvergiert). Wenn es nur um die Ausage des Lemma gilt, kann man auch einen etwas kuerzeren Beweis wie folgt geben: Waehle (f_{n_k}) mit

$$\sum_k \|f_{n_k} - f_{n_{k+1}}\|_1 < \infty.$$

Setze $g := \sum_k |f_{n_k} - f_{n_{k+1}}|$. Dann gilt nach monotoner Konvergenz

$$\int g d\mu = \sum_k \int |f_{n_k} - f_{n_{k+1}}| d\mu = \sum_k \|f_{n_k} - f_{n_{k+1}}\|_1 < \infty.$$

Damit folgt dann $g(x) < \infty$ fuer fast alle $x \in X$. Nun kann man wie im obigen Beweis der Schluss zu Ende gefuehrt werden.

THEOREM (Vollstaendigkeit von L^1). *Sei $(X, \mathcal{A}, \mu)$ ein Maßraum. Dann wird $L^1(X, \mathcal{A}, \mu)$ mit $\|\cdot\|_1$ zu einem vollstaendigen normierten Raum.*

Beweis. Das folgt sofort aus dem vorigen Lemma. □

Notation. Ein vollstaendiger normierter Vektorraum wird auch *Banachraum* genannt.

Wir kommen nun zu den L^p Raeumen: Sei $(X, \mathcal{A}, \mu)$ ein Massraum.

Sei $p \in [1, \infty)$ gegeben. Dann definiert man

$$\mathcal{L}^p(X, \mathcal{A}, \mu) := \{f : X \longrightarrow \mathbb{C} : \text{messbar mit } \int |f|^p d\mu < \infty\}.$$

Wir definieren weiterhin

$$\|\cdot\|_p : \mathcal{L}^p(X, \mathcal{A}, \mu) \longrightarrow [0, \infty), \ \|f\|_p := \left(\int |f|^p d\mu\right)^{1/p}.$$

Offenbar ist $\mathcal{L}^p(X, \mathcal{A}, \mu)$ ein Vektorraum.

Sei $p = \infty$ gegeben. Dann definiert man

$\mathcal{L}^\infty := \{f : X \longrightarrow \mathbb{C} : \text{messbar, es existiert ein } C \geq 0 \text{ mit } |f(x)| \leq C \text{ fast ueberall}\}.$

Offenbar ist $\mathcal{L}^p(X, \mathcal{A}, \mu)$ fuer jedes $p \in [1, \infty]$ ein Vektorraum. Wir definieren weiterhin

$$\|f\|_\infty := \inf\{C \geq 0 : |f(x)| \leq C \text{ f.u.}\}.$$

(Man mache sich klar, dass man das Infimum auch durch ein Minimum ersetzen kann.) Es heisst $\|\cdot\|_\infty$ das *wesentliche Supremum* von f.

Bemerkung. Man kann auch den Vektorraum aller beschraenkten messbaren Funktionen auf X mit der Supremum Norm ausstatten. Das liefert einen vollstaendigen Vektorraum (der unabhaengig vom gewaehlten Maß ist). Das ist aber fuer Anwendungen nicht so interessant, da dann Funktionen die fast ueberall uebereinstimmen als verschieden betrachtet werden.

Bemerkung. (Uebung) Gehoert f zu allen $\mathcal{L}^p(X, \mathcal{A}, \mu)$, $p \in [1, \infty]$, so gilt

$$\|f\|_\infty = \lim_{p \to \infty} \|f\|_p.$$

Das 'erklaert' die Bezeichnung $\|\cdot\|_\infty$ fuer diese Halbnorm.

Wir werden zeigen, dass $\|\cdot\|_p$ eine Halbnorm ist. Dazu bedarf es noch einer Vorbereitung, die auch fuer sich schon von grossem Interese ist. Es zeigt sich naemlich, dass L^p und L^q fuer $p, q \in [1, \infty]$ mit

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

zusammengehoeren. (Fuer $p = 1$ setzt man hier $q = \infty$ und fuer $p = \infty$ setzt man $q = 1$.) Es ist hilfreich, sich folgende aequivalente Beziehungen fuer p, q klarzumachen:

$$p + q = pq$$

bzw. weniger symmetrisch aber auch nuetzlich

$$q = p(q-1), \quad p = q(p-1),$$

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

$$q = \frac{p}{p-1}, \quad p = \frac{q}{q-1}.$$

Bemerkung. Der Fall $p = 2 = q$ wird spaeter noch von besonderem Interesse ein. (Hilbertraumtheorie).

THEOREM (Hoelder Ungleichung). *Seien $p, q \in [1, \infty]$ mit $1/p + 1/q = 1$ gegeben. Dann gilt fuer beliebige $f \in \mathcal{L}^p(X, \mathcal{A}, \mu)$ und $g \in \mathcal{L}^q(X, \mathcal{A}, \mu)$, dass fg zu $\mathcal{L}^1(X, \mathcal{A}, \mu)$ gehoert und*

$$\int |fg| d\mu \leq \|f\|_p \|g\|_q$$

erfuellt.

Beweis. Der Fall $p = 1, q = \infty$ oder $p = \infty, q = 1$ sind einfach. Wir betrachten daher nur die Situation $1 < p, q < \infty$.
Der Fall $\|f\|_p = 0$ oder $\|g\|_q = 0$ ist einfach. Wir setzen daher nun $\|f\|_p \neq 0$ und $\|g\|_q \neq 0$ voraus. Wir verwenden die bekannte Ungleichung

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}$$

fuer $a, b \geq 0$ und p, q mit $1/p + 1/q = 1$. (Ein Beweis dieser Ungleichung wird im Anschluss gegeben.) Anwenden auf

$$a = \frac{|f(x)|}{\|f\|_p}, b = \frac{|g(x)|}{\|g\|_q}$$

liefert

$$\frac{|f(x)g(x)|}{\|f\|_p\|g\|_q} \leq \frac{1}{p}\frac{|f(x)|^p}{\|f\|_p^p} + \frac{1}{q}\frac{|g(x)|^q}{\|g\|_q^q}.$$

Integration über X liefert dann

$$\int \frac{|f(x)g(x)|}{\|f\|_p\|g\|_q} d\mu \leq \frac{1}{p}\int_X \frac{|f(x)|^p}{\|f\|_p^p} d\mu + \frac{1}{q}\int_X \frac{|g(x)|^q}{\|g\|_q^q} d\mu = \frac{1}{p} + \frac{1}{q} = 1.$$

Damit folgt die gewuenschte Ungleichung leicht. □

Geometrischer Beweis der bekannten Ungleichung. Betrachte die Funktion

$$r : [0, \infty) \longrightarrow [0, \infty),\ r(x) = x^{p-1}$$

und ihre Stammfunktion

$$R(x) = \frac{1}{p}x^p.$$

Die Umkehrfunktion von r ist gegeben durch die Funktion

$$s : [0, \infty) \longrightarrow [0, \infty),\ s(y) = y^{\frac{1}{p-1}}$$

mit Stammfunktion

$$S(y) = \frac{1}{q}y^q.$$

(Hier nutzt man $\frac{1}{p-1} + 1 = \frac{1+p-1}{p-1} = \frac{p}{p-1} = q$.) Nach diesen Vorbereitungen folgt die gewuenschte Aussage einfach geometrisch.

Proposition. *Es ist $\|\cdot\|_p$ ein Halbnorm.*

Beweis. Der Fall $p = \infty$ bzw. $p = 1$ ist einfach. Wir betrachten nun $1 < p < \infty$.
Offenbar gilt $\|\alpha f\|_p = |\alpha|\|f\|_p$.
Wir zeigen nun die Dreiecksungleichung. Der Fall $\|f + g\|_p = 0$ ist klar. Wir setzen also $\|f + g\| \neq 0$ voraus. Wir nutzen die Hoelderungleichung:

$$\begin{aligned}
\|f+g\|_p^p &= \int_X |f+g||f+g|^{p-1} d\mu \\
&\leq \int_X |f||f+g|^{p-1} d\mu + \int_X |g||f+g|^{p-1} d\mu \\
\text{(Hoelder)} \quad &\leq \|f\|_p\|f+g\|_p^{p/q} + \|g\|_p\|f+g\|_p^{p/q} \\
&= (\|f\|_p + \|g\|_p)\|f+g\|_p^{p/q}.
\end{aligned}$$

Damit folgt die Aussage nach Multiplikation mit $\|f+g\|_p^{-p/q}$ (unter Nutzen von $p - p/q = 1$). Das beendet den Beweis. $\square$

Wir kommen nun zur Definition der eigentlichen L^p Raeume. Wir erinnern daran, dass wir auf dem Vektorraum aller messbaren Funktionen durch

$$f \sim q \Longleftrightarrow \ f = g \ \mu \text{ fast ueberall}$$

eine Aequivalenzrelation eingefuehrt hatten und fuer die Klasse $[f]$ eines messbaren f gerade gilt

$$[f] = f + \mathcal{N}$$

mit

$$\mathcal{N} := \{f : f \text{ messbar mit } f = 0 \ \mu \text{ fast ueberall } \}.$$

Dann gehoert f zu $\mathcal{L}^p$ genau dann, wenn ein / alle Elemente seiner Klasse zu $\mathcal{L}^p$ gehoeren. Man definiert

$$L^p(X) := L^p(X,\mu) := L^p(X,\mathcal{A},\mu) := \mathcal{L}^p(X,\mathcal{A},\mu)/\sim = \mathcal{L}^1(X,\mathcal{A},\mu)/\mathcal{N}.$$

Es gilt

$$\mathcal{N} = \{f : \|f\|_p = 0\}.$$

Daher wird $L^p(X,\mu)$ mit

$$\|\cdot\|_p : L^p(X,\mu) \longrightarrow [0,\infty), \ \|[f]\|_p := \|f\|_p,$$

zu einem normierten Raum.

Lemma (Vollstaendigkeit der $\mathcal{L}^p$). *Sei $(X,\mathcal{A},\mu)$ ein Massraum und $p \in [1,\infty]$ gegeben. Ist (f_n) eine Cauchy-Folge in $\mathcal{L}^p(X,\mathcal{A},\mu)$ so gibt es ein $f \in \mathcal{L}^p(X,\mathcal{A},\mu)$ und eine Teilfolge (f_{n_k}) mit*

- *$f_{n_k} \to f$ punktweise fast ueberall.*
- *$\|f - f_n\|_p \to 0$, $n \to \infty$.*

Beweis. Den Fall $p = \infty$ ueberlassen wir dem Leser zur Uebung. Sei nun $1 \le p < \infty$. Sei die Teilfolge (f_{n_k}) so gewaehlt, dass

$$\sum_{k=1}^{\infty} \|f_{n_k} - f_{n_{k+1}}\|_p < \infty$$

gilt. Setze

$$g_N := \sum_{k=1}^{N-1} |f_{n_k} - f_{n_{k+1}}|.$$

Dann ist die Folge (g_N^p) monoton wachsend mit punktweisem Grenzwert

$$\widetilde{g} : X \longrightarrow [0,\infty], \widetilde{g}(x) := \sum_{k=1}^{\infty} |f_{n_k}(x) - f_{n_{k+1}}(x)|.$$

(Hier ist der Wert ∞ fuer $\widetilde{g}$ ausdruecklich moeglich.) Weiterhin gilt nach der schon bewiesenen Dreiecksungleichung fuer alle natuerlichen Zahlen N dann

$$\begin{aligned}
\int g_N^p d\mu &= \|g_N\|_p^p \\
&\leq \left(\sum_{k=1}^{N} \|f_{n_k} - f_{n_{k+1}}\|_p^p\right)^p \\
&\leq \left(\sum_{k=1}^{\infty} \|f_{n_k} - f_{n_{k+1}}\|_p\right)^p \\
&=: C \\
&< \infty.
\end{aligned}$$

Damit ist dann nach dem Satz von der Monotonen Konvergenz die Funktion $\widetilde{g}^p$ integrierbar. Insbesondere ist also $\widetilde{g}$ fast ueberall endlich. Da $\widetilde{g}$ fast ueberall endlich ist, ist dann auch

$$\sum_{k\geq 1}(f_{n_k} - f_{n_{k+1}})$$

μ-fast ueberall absolut konvergent. Daher existiert dann also der Grenzwert der Folge

$$f_{n_N} = f_{n_1} + \sum_{k=1}^{N-1}(f_{n_{k+1}} - f_{n_k})$$

μ-fast ueberall. Die zugehoerige Funktion werde mit f bezeichnet. Nun koennen wir den Beweis wie im Falle von $\mathcal{L}^1$ abschliessen: Es gilt nach dem Lemma von Fatou dann

$$\begin{aligned}
\|f - f_{n_k}\|_p^p &= \int |f - f_{n_k}|^p d\mu \\
(f_{n_k} \to f \text{ f. u.}) \quad &= \int (\lim_l |f_{n_l} - f_{n_k}|^p d\mu \\
(Fatou) \quad &\leq \liminf_l \int |f_{n_l} - f_{n_k}|^p d\mu \\
&= \liminf_l \|f_{n_l} - f_{n_k}\|_p^p \\
&\to 0, k \to \infty.
\end{aligned}$$

Es gilt also $\|f - f_{n_k}\|_p \to 0, k \to \infty$. Da (f_n) eine Cauchy-Folge ist, folgt dann auch $\|f - f_n\|_p \to 0, n \to \infty$. □

THEOREM (Vollstaendigkeit der L^p). *Sei $(X, \mathcal{A}, \mu)$ ein Massraum und $p \in [1, \infty]$ gegeben. Dann ist $L^p(X, \mathcal{A}, \mu)$ vollstaendig.*

Beweis. Das folgt sofort aus dem vorigen Lemma. □

Die Hoeldersche Ungleichung liefert sofort noch die Reichhaltigkeit der stetigen linearen Abbildungen von L^p in den zugrundeliegenden Koerper. Diese linearen Abbildungen werden spaeter eine sehr grosse Rolle spielen. Tatsaechlich ist es eine Grundidee hoeherer Analysis, dass man einen Vektorraum studiert, indem man stetige lineare Abbildungen von dem Vektorraum in den Koerper betrachtet.

FOLGERUNG. *Sei $(X, \mathcal{A}, \mu)$ ein Massraum und $p \in [1, \infty]$ gegeben. Sei $q \in [1, \infty]$ mit $1/p + 1/q = 1$ gewaehlt. Dann definiert jedes $[g] \in L^q(X, \mathcal{A}, \mu)$ eine lineare Abbildung*

$$j_{[g]} : L^p(X, \mathcal{A}, \mu) \longrightarrow \mathbb{C}, \ \ j_g([f]) = \int_X g f d\mu$$

mit

$$|j_{[g]}(f)| \leq \|g\|_q \|[f]\|_p.$$

Bemerkung. Spaeter werden wir sehen, dass lineare Abbildungen mit der angegebenen Ungleichung *stetig* sind und wir werden zeigen, dass fuer $1 \leq p < \infty$ alle stetigen linearen Abbildungen in der angegebenen Weise entstehen.

Bemerkung - Schachtelung der L^p, siehe Uebung. Es liegt nahe eine Schachtelung der L^p Raeume zu erwarten. Schon am Beispiel $\mathbb{R}$ mit dem Lebesguemass kann man aber ablesen, dass im allgemeinen fuer $p \neq q$ keine der beiden Inklusionen $L^p \subset L^q$ oder $L^q \subset L^p$ gilt (Uebung). In zwei speziellen Situationen gilt aber eine Schachtelung:

- Sei $\mathbb{N}$ mit der Potenzmengen und dem Zaehlmass ζ ausgestattet und $\ell^p := \mathcal{L}(\mathbb{N}, \mathcal{P}, \zeta)$ fuer $p \in [1, \infty]$. Dann gilt
$$\ell^p \subset \ell^q \quad \text{fuer } p \leq q.$$
Insbesondere gilt $\ell^p \subset \ell^\infty$ fuer alle $p \in [1, \infty]$.
(Bew. Gilt $\sum_n |x(n)|^p < \infty$, so muss $(x(n))$ eine Nullfolge sein und $\sum_N |x(n)|^q < \infty$ folgt fuer $q \geq p$.)
- Ist $(X, \mathcal{A}, \mu)$ ein beliebiger Massraum mit $\mu(X) < \infty$, so gilt
$$\mathcal{L}^q(X, \mathcal{A}, \mu) \subset \mathcal{L}^p(X, \mathcal{A}, \mu) \quad \text{fuer } p \leq q..$$
(Bew. Es gilt $|f|^p \leq 1 + |f|^q$ fuer alle messbaren $f : X \longrightarrow \mathbb{C}$. Gilt $f \in \mathcal{L}^q$, so gehoert $|f|^q$ zu $\mathcal{L}^1$ und wegen $\mu(X) < \infty$ gehoert auch 1 zu $\mathcal{L}^1$ und es folgt $f \in \mathcal{L}^p$.)
Ein Massraum mit $\mu(X) = 1$ heisst auch *Wahrscheinlichkeitsraum*. Wahrscheinlichkeitsraeume sind die Grundobjekte der Stochastik.

Bemerkung zum Fall $0 < p < 1$. Man koennte versucht sein, eine aehnliche Theorie wie oben auch fuer $0 < p < 1$ zu entwickeln. Es stellen sich dann aber zwei Probleme:

- Die zugehoerigen Abbildungen $\| \cdot \|_p$ erfuellen i.a. nicht die Dreiecksungleichung (Uebung).
- Im allgemeinen gibt es ausser der Nullabbbildung KEINE linearen Abbildung
$$\Phi : L^p \longrightarrow \mathbb{C}$$
mit $|\Phi(f)| \leq C\|f\|_p$ fuer alle $f \in L^p$.

Notation. Auch wenn es streng mathematisch gesehen nicht ganz korrekt ist, ist es voellig ueblich bei Betrachtungen von L^p-Raeumen das Bilden von Klassen in der Notation wegzulassen. Man spricht dann zum Beispiel von einer Funktion f aus L^p oder eine Folge von Funktionen (f_n) in L^p.

KAPITEL 3

Etwas Hilbertraumtheorie

In diesem Abschnitt studieren wir Vektorraume mit einem Skalarprodukt. Das Skalarprodukt erlaubt es Laengen UND Winkel zu messen. Damit spielen Hilbertraeume eine fundamentale Rolle fuer weitere Untersuchungen:

- ihre Geometrie ist in gewisser Weise recht nahe an der im endlichdimensionalen Fall bekannten Geometrie;
- anhand von Hilbertraeumen koennen beispielhaft Grundphaenomene in unendlichdimensionalen Raeumen diskutiert werden;
- sie sind wesentlich fuer die mathematische Behandlung der Quantenmechanik.

Das wichtigste (und in gewissem Sinne auch einzige) Beispiel eines Hilbertraumes ist $L^2(X, \mu)$. In diesem Sinne geht es also in diesem Abschnitt um ein Studium der L^p-Raeume fuer $p = 2$.

1. Vektorräume mit Semiskalarprodukt

Wir beginnen mit einer Untersuchung von (Semi)skalarprodukten.

Definition. *Sei V ein Vektorraum über $\mathbb{K}$ mit $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$. Eine Abbildung $s : V \times V \longrightarrow \mathbb{K}$ heisst (semi) Skalarprodukt auf V, wenn gilt*

- *$s(x, \lambda y + \mu z)) = \lambda s(x, y) + \mu s(x, z)$ ('s ist linear im zweiten Argument')*
- *$s(x, y) = \overline{s(y, x)}$*
- *$s(x, x) \geq 0$*

fuer alle $x, y, z \in V$ und $\lambda, \mu \in \mathbb{K}$. Gilt darueberhinaus noch $s(x, x) > 0$ fuer $x \neq 0$, so heisst s ein Skalarprodukt. Dann schreibt man meist $\langle \cdot, \cdot \rangle$ (statt $s(\cdot, \cdot)$).

Bemerkungen.

- Ist s ein Semiskalarprodukt, so gilt

$$s(\lambda x + \mu y, z) = \overline{\lambda} s(x, z) + \overline{\mu} s(y, z)$$

  (d.h. s ist antilinear im ersten Argument).
- Ist s ein Semiskalarprodukt, so gilt

$$s(0, 0) = 0$$

  (Denn $0 = 0x$ also $s(0, 0) = s(0, 0x) = 0s(0, x) = 0$.)
- Manche Autoren definieren (Semi)skalarprodukte linear im ersten Argument und antilinear im zweiten Argument. Das aendert strukturell nichts, führt aber zu (leichten) Veraenderungen in manchen Formeln.

Notation. Ist s ein (Semi)Skalarprodukt auf V, so setzen wir

$$q : V \longrightarrow [0,\infty), q(x) = s(x,x),$$

und

$$\|\cdot\| : V \longrightarrow [0,\infty), \|x\| := q(x)^{1/2} = s(x,x)^{1/2}.$$

Beispiele (Uebung).

- $\mathbb{K}^N$ mit dem Euklidischen / Standard Skalarprodukt

$$\langle x, y\rangle := \sum_{j=1}^{N} \overline{x_j} y_j.$$

- $C[0,1]$ = stetige Funktionen auf [0,1] mit dem Skalarprodukt

$$\langle f, g\rangle := \int_0^1 \overline{f(t)} g(t) dt.$$

  Beachte: $\langle f, f\rangle = 0$ ist nur fuer $f \equiv 0$ moeglich, da f stetig ist.
- Sei $(X, \mathcal{A}, \mu)$ ein Massraum. Dann definiert

$$\langle f, g\rangle := \int f g d\mu$$

  ein Skalarprodukt. (Integral existiert aufgrund der Hoelder Ungleichung.) Es gilt

$$\langle f, f\rangle = \|f\|_2^2.$$

  Insbesondere ist also $L^2(X, \mathcal{A}, \mu)$ vollstaendig bzgl. der durch das Skalarpodukt erzeugten Norm. Solche Raeume werden wir spaeter Hilbertraeume nennen.
  - Spezialfall:

$$\ell^2 := L^2(\mathbb{N}, \text{Zaehlmass}) := \{c : \mathbb{N} \longrightarrow \mathbb{C} : \sum_{k=1}^{\infty} |c(n)|^2 < \infty$$

    mit dem Skalarprodukt

$$\langle c, d\rangle := \sum_{k=1}^{\infty} \overline{c(k)} d(k).$$

  - Spezialfall: $\ell^2(\{1,\ldots,N\}) := L^2(\{1,\ldots,N\}, \text{Zaehlmass}) = \mathbb{K}^N$

Fuer Semiskalarprodukte gelten drei fundamentale Sachverhalte:

- Cauchy-Schwarz Ungleichung,
- Polarisierung,
- Parallelogrammidentitaet.

Das untersuchen wir nun:

PROPOSITION (Cauchy-Schwarz-Bunyakowski Ungleichung). *Sei $s(\cdot,\cdot)$ ein Semiskalarprodukt auf V. Dann gilt*

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

fuer alle $x, y \in V$.

Beweis. Sei

$$F(t) := s(x+ty, x+ty) = \|x\|^2 + ts(x,y) + ts(y,x) + t^2\|y\|^2.$$

Nach Voraussetzung gilt $F \geq 0$. Das ist nur moeglich, wenn die gewuenschte Ungleichung gilt. Hier sind die Details: Ohne Einschraenkung $s(x,y) \geq 0$ (sonst Multiplizieren mit $e^{i\alpha}$). Ohne Einschraenkung $s(x,y) > 0$ (sonst ist die Aussage sowieso klar). Damit gilt also

$$F(t) = \|x\|^2 + 2ts(x,y) + t^2\|y\|^2.$$

Wegen $F \geq 0$ und $s(x,y) \neq 0$ folgt $\|y\| > 0$. Dann ist auch

$$\frac{1}{\|y\|^2}F = \frac{\|x\|^2}{\|y\|^2} + 2t\frac{s(x,y)}{\|y\|^2} + t^2 = \left(t + \frac{s(x,y)}{\|y\|^2}\right)^2 + \frac{\|x\|^2}{\|y\|^2} - \left(\frac{s(f,g)}{\|g\|^2}\right)^2$$

ein nichtnegatives Polynom. Damit folgt

$$\frac{\|x\|^2}{\|y\|^2} - \left(\frac{s(x,y)}{\|y\|^2}\right)^2 \geq 0.$$

Das liefert die Aussage. $\square$

Bemerkungen.

- Ist s ein Skalarprodukt und gilt $|s(x,y)| = s(x,x)^{1/2}s(y,y)^{1/2}$ fuer $x, y \in V$, so sind x und y linear abhaengig. (Uebung: O.E. $s(x,x) = s(y,y) = 1$. Betrachte $z = x - s(y,x)y$. Dann gilt

$$s(z,z) = s(x,x) - s(y,x)s(x,y) - \overline{s(y,x)}s(y,x) + |s(y,x)|^2 s(y,y) = 1-1-1+1 = 0.)$$

- Fuer den Fall des Skalarproduktes auf einem L^2-Raum ist die Cauchy-Schwarz-Ungleichung eine direkte Folgerung der Hoelder Ungleichung.

Bemerkung - Skalarprodukt und Winkelmessung (Uebung). Sei V ein Vektorraum ueber $\mathbb{R}$ mit Skalarprodukt. Ausgehend von der Cauchy-Schwarz Ungleichung kann man den (Cosinus des) Winkel zwischen zwei Vektoren x, y (die nicht der Nullvektor sind) definieren durch

$$\widetilde{\cos}(x,y) := \frac{s(x,y)}{\|x\|\|y\|}.$$

Mit dieser Definition gelten dann aus der Geometrie des $\mathbb{R}^2$ bekannte Formeln fuer den cosinus wie etwa:

$$s(x,y) = \|x\|\|y\|\widetilde{\cos}(x,y)$$

sowie

$$\widetilde{\cos}(x,y) = 0$$

falls x und y linear abhaengig sind. Das ist kein Zufall. Tatsaechlich kann man (wie?) zu gegebenen $x, y \in V$ eine bijektive lineare Abbildung U zwischen $\mathbb{R}^2$ und dem durch x, y erzeugten Unterraum $\mathrm{Lin}(x,y) := \{\alpha x + \beta y : \alpha, \beta \in \mathbb{R}\}$ angeben

$$U : \mathbb{R}^2 \longrightarrow \mathrm{Lin}(x,y)$$

mit $\langle v, w\rangle = s(Uv, Uw)$ (mit den ueblichen Skalarprodukt $\langle\cdot,\cdot\rangle$ auf $\mathbb{R}^2$). Damit gilt dann (Warum?) $\cos(v,w) = \widetilde{\cos}(v,w)$, wobei $\cos(v,w)$ den Cosinus des von den Vektoren v, w eingeschlossenen Winkels bezeichnet.

Folgerung. *Ist s ein Semiskalarprodukt auf V, so ist*

$$\|x\| := s(x,x)^{1/2}$$

eine Halbnorm, d.h. es gilt

- $\|x+y\| \leq \|x\| + \|y\|$,
- $\|\lambda x\| = |\lambda|\|x\|$,
- $\|x\| \geq 0$,

fuer alle $x \in V$ und $\lambda \neq 0$. Ist s sogar ein Skalarpodukt, so ist $\|\cdot\|$ sogar eine Norm, d.h. es gilt zusaetzlich noch $\|x\| > 0$ fuer alle $x \neq 0$.

Beweis. Bis auf die erste Eigenschaft ist alles klar. Wir zeigen die erste Eigenschaft:

$$\begin{aligned}
\|f+g\|^2 &= s(f+g, f+g) \\
&= s(f,f) + s(f,g) + s(g,f) + s(g,g) \\
&\leq \|f\|^2 + 2\|f\|\|g\| + \|g\|^2 \\
&= (\|f\| + \|g\|)^2.
\end{aligned}$$

Das beendet den Beweis. □

Die folgenden beiden Aussage befassen sich mit der 'quadratischen Form'

$$q(x) = s(x,x)$$

eines Semiskalarproduktes. Mathematisch stellen sie nichts anders dar als Umsetzungen der Formeln

$$(a+b)^2 - (a-b)^2 = 4ab \text{ und } (a+b)^2 + (a-b)^2 = 2a^2 + 2b^2.$$

Es ist moeglich, die Werte von s auszurechnen, wenn man nur die Werte $\|x\| := s(x,x)^{1/2}$ fuer $x \in V$ kennt. Das ist unter dem Namen Polarisierung bekannt.

Proposition (Polarisierung). *Ist s ein Semiskalarprodukt auf V, so gilt mit $q(x) := s(x,x) = \|x\|^2$*

$$s(x,y) = \frac{1}{4}\left(q(x+y) - q(x-y) + iq(x-iy) - iq(x+iy)\right)$$

(falls $\mathbb{K} = \mathbb{C}$) und

$$s(x,y) = \frac{1}{4}\left(q(x+y) - q(x-y)\right)$$

(falls $\mathbb{K} = \mathbb{R}$).

Beweis. Das folgt direkt durch Einsetzen. □

Wir kommen nun zur Parallelogrammidentitaet.

Proposition (Parallelogrammidentitaet). *Sei s eine Semiskalarprodukt auf V und $q(x) = s(x,x)$. Dann gilt fuer alle $x, y \in V$*

$$q(x+y) + q(x-y) = 2(q(x) + q(y)).$$

(Zeichnung)

Beweis. Das folgt durch direkte Rechnung. □

Bemerkungen.

- (Jordan/von Neumann) Eine (Halb)Norm auf einem Vektorraum wird genau dann durch ein (Semi)Skalarprodukt induziert, wenn die Parallelogrammidentitat gilt. Die (Halb)Norm ist dann durch die Polarisierungsidentitaet gegeben. In diesem Sinne ist die Parallelogrammindentitaet die fundamentale Eigenschaft eines Raumes mit innerem Produkt.
- Die Norm von ℓ^p wird genau fuer $p = 2$ durch ein Skalarprodukt induziert, wie man mittels der Parallelogrammidentitaet sehen kann (Uebung).

Wir kommen nun zu grundlegenden geometrischen Begriffen, die ein Skalarprodukt auf einem Vektorraum ermoeglicht. Wie schon erwaehnt, koennen in einem Vektorraum mit Skalarprodukt sowohl Laengen und Winkel gemessen werden. Das spiegelt sich in folgenden Konzepten wieder:

- Laengenmessung wird im Konzept der Norm gefasst und das erlaubt es dann insbesondere eine Metrik (und damit eine Topologie) einzufuehren.
- Was Winkelmessung betrifft, so wird fuer uns vor allem das Konzept des rechten Winkels d.h. der Orthogonalitaet eine Rolle spielen. Damit einher geht dann das Konzept der orthogonalen Projektion.

Ein Vektorraum V mit Skalarprodukt $\langle\cdot,\cdot\rangle$ traegt die Norm (s.o.)

$$\|\cdot\| : V \longrightarrow [0,\infty), \|x\| = \langle x, x\rangle^{1/2}.$$

Damit wird dann eine Metrik $d = d_{\langle\cdot,\cdot\rangle}$ durch

$$d(x,y) := \|x - y\|$$

auf V induziert. Wann immer im folgenden im Kontext eines Raumes mit Skalarprodukt von metrischen Eigenschaften (Konvergenz von Folgen, Cauchy-Folgen, Abgeschlossenheit, Stetigkeit....) die Rede ist, wird die eben definierte Metrik zugrunde gelegt.

Proposition. *Sei V ein Vektorraum mit Skalarprodukt mit induzierter Norm $\|\cdot\|$ und induzierter Metrik d. Dann sind $\|\cdot\| : V \longrightarrow \mathbb{R}$ und $\langle\cdot,\cdot\rangle : V \times V \longrightarrow \mathbb{K}$ stetig.*

Beweis. Das folgt einfach aus der Dreiecksungleichung und Cauchy-Schwarz.
Stetigkeit von $\|\cdot\|$: Das folgt leicht aus der folgenden Konsequenz der Dreiecksungleichung

$$|\|x\| - \|y\|| \leq \|x - y\|.$$

Stetigkeit von $\langle\cdot,\cdot\rangle$: Aus Dreiecksunglgeichung und Cauchy-Schwarz-Ungleichung folgt

$$|\langle x,y\rangle - \langle x',y'\rangle| \leq |\langle x - x', y\rangle| + |\langle x', y - y'| \leq \|x - x'\|\|y\| + \|x'\|\|y - y'\|.$$

Nun ergibt sich die gewuenschte Aussage leicht aus der schon gezeigten Stetigkeit der Norm. $\square$

Definition (Hilbertraum). *Ein Vektorraum V mit Skalarprodukt $\langle \cdot, \cdot \rangle$ und induzierter Metrik d heisst Hilbertraum, wenn er bzgl. d vollstaendig ist (d.h. jede Cauchy Folge bzgl. d einen Grenzwert hat).*

Notation. Wir schreiben (oft) $(H, \langle \cdot, \cdot \rangle)$ fuer einen Hilbertraum mit seinem Skalaprodukt.

Wie schon Zum Abschluss des Abschnittes kommen wir noch zum Konzept der Orthogonalitaet.

Definition (Orthogonal). *Sei V Vektorraum mit Skalarprodukt. Dann heissen $x, y \in V$ orthogonal (senkrecht), wenn gilt*

$$s(x, y) = 0.$$

Man schreibt dann $x \perp y$. Gilt $A \subset V$, so definiert man das orthogonale Komplement von A durch $A^\perp := \{u \in V : s(u, a) = 0$ fuer alle $a \in A\}$.

Beispiele. V mit Skalarprodukt. Dann $V^\perp = \{0\}$, $\{0\}^\perp = V$.

Wichtige Deutung. Ist $s(y, y) = 1$, so ist

$$z = x - s(y, x)y$$

senkrecht auf y d.h. es gilt $s(z, y) = 0$. Es gibt also $s(y, x)$ die Laenge der Komponente von x in Richtung y an. (Zeichnung.) Bew: Nachrechnen:

$$s(z, y) = s(x - s(y, x)y, y) = s(x, y) - \overline{s(y, x)}s(y, y) = 0.$$

Proposition. *Sei $(V, \langle \cdot, \cdot \rangle)$ ein Vektorraum mit Skalarprodukt und $A \subset V$ eine beliebige Teilmenge. Dann ist $A^\perp$ ein abgeschlossener Unterraum.*

Beweis. (Uebung). □

2. Hilbertraeume und Approximationssatz

In diesem Abschnitt lernen wir eine fundamentale Eigenschaft eines Hilbertraumes kennen.

Theorem (Approximationssatz). *Sei $(H, \langle \cdot, \cdot \rangle)$ ein Hilbertraum. Ist C eine abgeschlossene konvexe Teilmenge von H, so gibt es zu jedem $x \in V$ genau ein $y \in C$ mit*

$$\|x - y\| = d(x, C) := \inf\{\|x - z\| : z \in C\}.$$

Damit existiert also die beste Approximation an x in C. (Zeichnung.)

Beweis. Bei dem Satz handelt es sich um eine fundamentale Eigenschaft eines Hilbertraum. Entsprechend spielen die fundamentalen Eigenschaften des Hilbertraum naemlich Vollstaendigkeit und Parallelogrammidentitaet

$$\|v + w\|^2 + \|v - w\|^2 = 2\|v\|^2 + 2\|w\|^2$$

eine Rolle.

Wir setzen $d := d(x, C)$. Existenz und Eindeutigkeit werden ganz aehnlich gezeigt.

Eindeutigkeit. Seien y_1, y_2 solche Punkte. Anwenden der Parallelogrammidentitaet mit $v = x - y_1$ und $w = x - y_2$ liefert

$$\|2x - y_1 - y_2\|^2 + \|y_1 - y_2\|^2 = 2\|x - y_1\|^2 + 2\|x - y_2\|^2.$$

Damit folgt

$$4\|x - \frac{1}{2}(y_1 + y_2)\|^2 + \|y_1 - y_2\|^2 = 4d^2.$$

Aufgrund der Konvexitaet von C und der Definition von d laesst sich der linkeste Term durch $4d^2$ nach unten abschaetzen. Damit folgt

$$\|y_1 - y_2\|^2 = 0.$$

Existenz. Sei (y_n) eine Folge in C mit

$$\|x - y_n\|^2 \to d^2.$$

Einsetzen in die Parallelogrammidentitaet mit $v = x - y_n$ und $w = x - y_m$ liefert (wie im Eindeutigkeitsteil)

$$4\|x - \frac{1}{2}(y_n - y_m)\|^2 + \|y_n - y_m\|^2 = 2\|x - y_n\|^2 + 2\|x - y_m\|^2.$$

Damit folgt (Details, Zeichnung Parallelogramm)

$$\|y_n - y_m\| \to 0.$$

Daher ist (y_n) eine Cauchy Folge. Aufgrund der Hilbertraumeigenschaft konvergiert dann (y_n). Da C abgeschlossen ist, gehoert der Grenzwert y wieder zu C. Weiterhin gilt nach Definition

$$\|x - y\| = \lim \|x - y_n\| = d.$$

Das beendet den Beweis. $\square$

Bemerkung. (a) Sowohl die Voraussetzung der Konvexitaet als auch der Abgeschlossenheit sind noetig. (Uebung): Ist die Menge konvex, aber nicht abgeschlossen, so muss es keine beste Approximation geben (sie koenne ja gerade fehlen). Ist die Menge abgeschlossen und nicht konvex, kann es z.b. mehrere beste Appoximationen geben (klar). Es kann dann auch keine beste Approximation geben: $x = e_1$. $A = \{(1 + \frac{1}{j})e_j : j > 1\}$. (Im endlichdimensionalen Hilbertraum gibt es natuerlich beste Approximationen fuer beliebige abgeschlossenen Mengen. Warum? Kompaktheit!)
(b) In Raeumen, die keine Hilbertraeume sind, muss die Aussage nicht gelten. (Uebung).

Theorem (Projektionssatz). *Sei $(H, \langle\cdot,\cdot\rangle)$ ein Hilbertraum und U ein abgeschlossener Unterraum. Dann laesst sich jedes $x \in H$ eindeutig schreiben als*

$$x = y + z \text{ mit } y \in U \text{ und } z \in U^\perp$$

und es gilt

$$\|x - y\| = \min\{\|x - u\| : u \in U\}.$$

Beweis. Eindeutigkeit. Sei $x = y + z = y' + z'$ mit $y, y \in U$ und $z, z' \in U^\perp$. Dann gilt

$$y - y' = z' - z \in U \cap U^\perp = \{0\}.$$

Damit folgt die Eindeutigkeit.

Existenz: Es ist U abgeschlossen und konvex. Daher existiert nach dem vorigen Satz ein (eindeutiges) $y \in U$ mit

$$\|x - y\| = d(x, U) = \inf\{\|x - z\| : z \in U\}.$$

Sei $z = x - y$. Dann gilt also

$$x = y + z$$

mit $y \in U$.
Noch zu zeigen $z \perp U$: Sei $u \in U$ beliebig. Dann hat die Funktion

$$F = F_u : \mathbb{R} \longrightarrow [0, \infty), F(t) = \|(x - y) + tu\|^2$$

ein Minimum bei $t = 0$ nach Konstruktion von $z = x - y$. Es gilt

$$F(t) = \|x - y\|^2 + t\langle x - y, u\rangle + t\langle u, (x - y)\rangle + t^2\|u\|^2,$$

also

$$F(t) = \|x - y\|^2 + 2t\Re\langle (x - y), u\rangle + t^2\|u\|^2.$$

Da F diffbar ist und ein Minimum in $t = 0$ hat folgt

$$0 = F^{'}(0) = \Re\langle (x - y), u\rangle$$

fuer jedes beliebige $u \in U$. Damit folgt

$$0 = \Re\langle (x - y), \lambda u\rangle$$

fuer alle $u \in U$, $\lambda \in \mathbb{K}$. Mit $\lambda = \overline{\langle (x - y), u\rangle}$ folgt

$$0 = |\langle (x - y), u\rangle|^2.$$

Das liefert die gewuenschte Orthogonalitaet. □

Bemerkung.

- Im Beweis wird eine Beziehung zwischen Minimalitaet (des Abstands) und Orthogonalitaet (des Abstandsvektors) hergestellt. Das Argument ist typisch fuer *Variationsrechnung*: Minimalitaet liefert Verschwinden gewisser Ableitungen und das bedeutet im konkreten Falle dann gerade eine Orthogonalitaet. Aehnliche Schluesse werden in der Physik z.b. beim Hamiltonprinzip der Bewegung durchgefuehrt und fuehren auf die Euler-Lagrangeschen Differentialgleichungen.
- Auch im Approximationssatz laesst sich eine Orthogonalitaet entdecken. Der Differenzvektor minimaler Laenge ist senkrecht auf einer geeignet definierten Tangentialebene an die Menge C. Genauer gilt (Uebung): Es ist y das eindeutige Element aus C mit

  $$\Re\langle x - y, z - y\rangle \leq 0$$

  fuer alle $z \in C$.

Beispiel. Seien $e_1, \ldots, e_N$ normiert und paarweise orthogonal in einem Hilbertraum H und $U := Lin\{e_1, \ldots, e_N\}$. Dann gilt fuer $x \in H$ die Gleichung

$$x = y + z$$

mit $y = \sum_{j=1}^N \langle e_j, x\rangle e_j \in U$ und $z = x - y \in U^\perp$. Damit handelt es sich um die in dem vorigen Theorem beschriebene eindeutige Darstellung von x. Insbesondere gilt also fuer jedes $c_1, \ldots, c_N \in \mathbb{K}$

$$\|x - \sum_{j=1}^N c_j e_j\|^2 \geq \|x - \sum_{j=1}^N \langle e_j, x\rangle e_j\|^2.$$

(Es ist instruktiv auch direkt die beste Approximationseigenschaft nachzurechnen: Es gilt

$$\begin{aligned} \|x - \sum_{j=1}^N c_j e_j\|^2 &= \|(x - \sum_{j=1}^N \langle e_j, x\rangle e_j) + \sum_{j=1}^N \langle e_j, x\rangle e_j - \sum_{j=1}^N c_j e_j\|^2 \\ &= \|(x - \sum_{j=1}^N \langle e_j, x\rangle e_j) + \sum_{j=1}^N (\langle e_j, x\rangle - c_j) e_j\|^2 \\ (Pythagoras) \quad &= \|(x - \sum_{j=1}^N \langle e_j, x\rangle e_j)\|^2 + \|\sum_{j=1}^N (\langle e_j, x\rangle - c_j) e_j\|^2 \\ &\geq \|(x - \sum_{j=1}^N \langle e_j, x\rangle e_j)\|^2. \end{aligned}$$

Es wird also die Differenz minimal fuer $c_j = \langle e_j, x\rangle$.)

Der vorige Satz legt die folgende Definition nahe.

Definition. *Ist U ein abgeschlossener Unterraum eines Hilberteraum H, so heisst die Abbildung*

$$P_U : H \longrightarrow H, x \mapsto y,$$

(mit $x = y + z$ mit $y \in U$ und $z \in U^\perp$) die orthogonale Projektion auf U.

Proposition (Charakteristikum Projektionen). *Sei U ein abgeschlossener Unterraum eines Hilberteraum H. Dann gilt fuer $P = P_U$ folgendes:*

- *$P(\alpha x + \beta y) = \alpha Px + \beta Py$ fuer alle $x, y \in H$ und $\alpha, \beta \in \mathbb{K}$. 'P ist linear'*
- *$P = P^2$. ' P ist idempotent'*
- *$\langle Pw, v\rangle = \langle w, Pv\rangle$. fuer alle $v, w \in H$. 'P ist selbstadjungiert'*

Beweis. Zum ersten Punkt: Es gilt

$$\alpha x + \beta y = \alpha Px + \beta Py + (\alpha x - \alpha Px) + (\beta y - \beta Py).$$

Damit folgt die gewuenschte Aussage leicht aus der Eindeutigkeit der Zerlegung.

Zum zweiten Punkt: Es gilt $Px = Px + 0$ mit $Px \in U$ und $0 \in U^\perp$. Damit folgt die gewuenschte Aussage leicht aus der Eindeutigkeit der Zerlegung.

Zum dritten Punkt: Wir schreiben $w = y + z$, $v = y' + z'$ mit $y, y' \in U$ und $z, z' \in U^\perp$. Nun laesst sich die Aussage direkt nachrechnen. □

Bemerkung. Es laesst sich zeigen (Uebung), dass jede lineare Abbildung, die selbstadjungiert und idempotent ist eine orthogonale Projektion ist (wobei der zugehoerige Unterraum gerade das Bild der Abbildung ist)

FOLGERUNG. *(a) Ist U ein abgeschlossener Unterraum eines Hilbertraumes, so gilt $(U^{\perp})^{\perp} = U$.*
(b) Ist A eine beliebige Teilmenge eines Hilbertraumes, so gilt $\overline{Lin(A)} = A^{\perp\perp}$.

Bemerkung. Es ist in (a) die Abgeschlossenheit von U noetig, da orthogonale Komplemente automatisch abgeschlossen sind, siehe Uebung. Weiterhin ist offenbar (a) eine Folge von (b). Wie unser Beweis zeigt, kann man auch (b) aus (a) herleiten.

Beweis. (a) $U \subset U^{\perp\perp}$: Fuer $x \in U$ und $z \in U^{\perp}$ gilt nach Definiton von $U^{\perp}$

$$\langle x, z\rangle = 0.$$

Damit gilt $x \in U^{\perp\perp}$

$U^{\perp\perp} \subset U$: Sei $x \in U^{\perp\perp}$. Dann gilt nach dem vorigen Satz $x = y + z$ mit $y \in U$ und $z \in U^{\perp}$. Damit folgt (wegen $x \in U^{\perp\perp}$ und $y \in U \subset U^{\perp\perp}$)

$$z = x - y \in U^{\perp} \cap U^{\perp\perp} = \{0\}.$$

Damit folgt $x = y \in U$.

(b) Offenbar gilt $A^{\perp} = (LinA)^{\perp} = (\overline{LinA})^{\perp}$. Damit folgt (b) sofort aus (a) (mit $U = \overline{LinA}$). □

Damit koennen wir den Projektionssatz wie folgt umformulieren.

FOLGERUNG. *Ist U ein abschgeschlossener Unterraum eine Hilbertraums. Dann gilt $I = P_U + P_{U^{\perp}}$.*

Beweis. Sei x ein beliebiges Element des Hilbertraumes. Dann gilt $x = y+z$ mit $y \in U = U^{\perp\perp}$ und $z \in U^{\perp}$. Damit ist also $y = P_U x$ (da $y \in U$ und $z \in U^{\perp}$ und es ist $z = P_{U^{\perp}} x$ (da $z \in U^{\perp}$ und $y \in U^{\perp\perp}$). □

Fuer spaetere Anwendung notieren wir noch folgendes Lemma.

LEMMA. *Sei A eine Teilmenge eines Hilbertraumes. Dann sind aequivalent:*
(i) $\overline{LinA} = V$
(ii) $A^{\perp} = \{0\}$.

Beweis. (ii)$\Longrightarrow$ (i): Das folgt sofort aus der vorangegangenen Folgerung

$$\overline{LinA} = A^{\perp\perp} = \{0\}^{\perp} = V.$$

(i)$\Longrightarrow$ (ii): Wendet man zunaechst (a) und dann (b) der Folgerung und schliesslich (i) an, so folgt

$$A^{\perp} = (A^{\perp})^{\perp\perp} = \overline{LinA}^{\perp} = \{0\}.$$

Das ist aber gerade (ii). □

Bemerkung. Mengen A wie im Lemma heissen *total* oder auch *Erzeugendensystem im Hilbertraum.* Das wird uns spaeter noch beschaeftigen.

3. Entwicklung nach Orthonormalbasen

Wir untersuchen Entwicklungen der Form

$$x = \sum c_k e_k$$

mit e_k Orthonormalsystem und $\sum |c_k|^2 < \infty$. Im Hilbertraum gilt:

- Jede solche Summe stellt ein x aus dem Raum dar.
- Hat (e_k) eine gewisse Vollstaendigkeitseigenschaft (Orthonormal*basis*), so kann jedes x aus dem Raum kann eindeutig so dargestellt werden mit $c_k = \langle e_k, x\rangle$.

Das erlaubt es dann, mit Koordinaten wie im Euklidischen Raum zu rechnen. Darum geht es in diesem Abschnitt.

DEFINITION. *Sei V ein Vektorraum mit Semiskalarprodukt s. Sei I eine Indexmenge und e_j, $j \in I$, Element von V. Dann heissen die e_j ein Orthonormalsystem (ONS), wenn gilt*

$$s(e_i, e_j) = \delta_{i,j}$$

fuer alle $i, j \in I$.

Wir sammeln einige wesentliche Eigenschaften orthogonaler Vektoren:

PROPOSITION (Eigenschaften orthogonaler Vektoren). *Sei V ein Vektorraum mit (Semi)Skalarprodukt s.*
(a) (Pythagoras) Gilt $x \perp y$ so folgt

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2.$$

(b) Ist e_j, $j = 1, \ldots, N$ ein endliches Orthonormalsystem so gilt fuer jedes $x \in V$

$$x - \sum_{j=1}^{N} s(e_k, x)e_k \perp e_j, \;\; j = 1, \ldots, N.$$

(c) (Besselsche Ungleichung) Ist $\{e_j : j \in I\}$ ein beliebiges Orthonormalsystem, so gilt fuer jedes $x \in V$

$$\|x\|^2 \geq \sum_{j\in I} |s(e_j, x)|^2.$$

Definition von Summen mit ueberabzaehlbar vielen Summanden: Zur Definition der Summe in (c): Ist I eine Indexmenge und sind $c_j, j \in I$, nichtnegative Zahlen, so definiert man

$$\sum_{j\in I} c_j := \sup\{\sum_{j\in A} c_j : A \subset I \text{ endlich}\}.$$

Aus

$$\sum_{j\in I} c_j < \infty$$

folgt dann, dass hoechstens abzaehlbar viele der c_j nicht verschwinden. (Da fuer jedes $n \in \mathbb{N}$ die Menge

$$I_n := \{j \in I : c_j \geq 1/n\}$$

endlich sein muss und

$$\{j \in I : c_j \neq 0\} = \bigcup_n I_n$$

gilt.)

Beweis. (a) Direkte Rechnung:

$$\|x+y\|^2 = s(x+y, x+y) = s(x,x) + s(x,y) + s(y,x) + s(y,y) = \|y\|^2 + \|x\|^2.$$

(b) Direkte Rechnung (s.o. fuer den Fall eines Vektors e_1):

$$s(x - \sum_{k=1}^{N} s(e_k, x)e_k, e_j) = s(x, e_j) - \sum_{k=1}^{N} \overline{s(e_k, x)} s(e_k, e_j) = s(x, e_j) - s(x, e_j) = 0.$$

(c) Seien $e_{j_1}, \ldots, e_{j_N}$ ein beliebiges endliches Teilsystem von (e_j). Dann gilt:

$$\begin{aligned}
\|x\|^2 &= \|x - \sum_{k=1}^{N} s(e_{j_k}, x)e_{j_k} + \sum_{k=1}^{N} s(e_{j_k}, x)e_{j_k}\|^2 \\
&\overset{(a,b)}{=} \|x - \sum_{k=1}^{N} s(e_{j_k}, x)e_{j_k}\|^2 + \|\sum_{k=1}^{N} s(e_{j_k}, x)e_{j_k}\|^2 \\
&\geq \|\sum_{k=1}^{N} s(e_{j_k}, x)e_{j_k}\|^2 \\
(a) \quad &= \sum_{k=1}^{N} |s(e_{j_k}, x)|^2.
\end{aligned}$$

Da die Aussage fuer beliebige endliche Teilsysteme gilt, folgt sie fuer die Ursprungsmenge. □

Die vorangehende Proposition zeigt die Nuetzlichkeit eines Orthonormalsystems. Aus jedem abzaehlbaren System linear unabhaengiger Vektoren kann man ein Orthonormalsystem gewinnen.

Proposition. *(Gram/Schmidtsches Orthogonalisierungsverfahren) Sei V ein Vektorraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$. Seien die Vektoren $v_1, \ldots, v_N$ (bzw. v_j, $j \in \mathbb{N}$) linear unabhaengig. Dann bilden die induktiv definierten Vektoren*

$$e_1 := \frac{1}{\|v_1\|} v_1, \; e_{k+1} := \frac{1}{\|v_{k+1} - \sum_{j=1}^{k} s(e_j, v_{k+1})e_j\|} (v_{k+1} - \sum_{j=1}^{k} s(e_j, v_{k+1})e_j)$$

ein Orthonormalsystem mit

$$Lin\{e_1, \ldots, e_k\} = Lin\{v_1, \ldots, v_k\}$$

fuer alle k.

Beweis. Der Beweis wird durch Induktion nach k gefuehrt ueber die Aussage: Es bilden $e_1, \ldots, e_k$ ein Orthonormalsystem mit

$$\mathrm{Lin}\{e_1, \ldots, e_k\} = \mathrm{Lin}\{v_1, \ldots, v_k\}.$$

$k = 1$: klar.

$k \Longrightarrow (k+1)$: Es ist $w_{k+1} := v_{k+1} - \sum_{j=1}^{k} s(e_j, v_{k+1})e_j$ senkrecht auf e_l, $l = 1, \ldots, e_k$ (nach voriger Proposition) und verschwindet nicht (aufgrund der linearen Unabhaengigkeit der v_j und der Bedingung and die Huellen). Durch Normieren erhalten wir e_{k+1} und $e_1, \ldots, e_{k+1}$ bilden ein Orthonormalsystem. Weiterhin gilt

$$\begin{aligned} Lin\{e_1, \ldots, e_k, e_{k+1}\} &= Lin\{e_1, \ldots, e_k, w_{k+1}\} \\ &= Lin\{e_1, \ldots, e_k, v_{k+1}\} \\ &= Lin\{v_1, \ldots, v_{k+1}\}. \end{aligned}$$

Dabei haben wir im zweiten Schritt die Definition der $w's$ genutzt und im letzten Schritt die Induktionsannahme fuer k. $\square$

Bemerkung. Auch wenn die (v_j) nicht linear unabhaengig sind, kann man das Gram/Schmidtsche Verfahren in einer einfachen Modifikation anwenden. Dazu streicht man sukzessive alle diejenigen N bei denen $w_N = 0$ gilt. (Hier wird das im Beweis eingefuehrte w_N verwendet.)

Wir wenden uns nun Entwicklungen nach Orthonormalsystemen zu. Da wir apriori keine Abzaehlbarkeitsforderungen an die Indexmengen unserer Orthonormalsystem stellen, erweist sich folgende **Notation** als sinnvoll: Sei J eine Menge. Sei F eine Funktion auf den endlichen Teilmengen von J mit Werten in einem metrischen Raum M (mit Metrik d). Dann definieren wir

$$\lim_{A \to J} F(A) = x,$$

falls fuer jedes $\varepsilon > 0$ eine endliches $A \subset J$ existiert mit

$$d(F(B), x) \leq \varepsilon$$

fuer jedes endliche $B \supset A$. Diese Definition von Konvergenz ist gut mit Stetigkeit vertraeglich: Ist (N, e) ein weiterer metrischer Raum und $h : M \longrightarrow N$ stetig, so folgt aus $\lim_{A \to J} F(A) = x$ sofort

$$\lim_{A \to J} h(F(A)) = h(x).$$

(Bew. Da h stetig ist, existiert zu $\varepsilon > 0$ ein $\delta > 0$ mit $e(h(y), h(x)) \leq \varepsilon$ fuer alle $y \in X$ mit $d(y, x) \leq \delta$. Wegen $\lim_{A \to J} F(A) = x$ existiert zu $\delta > 0$ ein endliches $A_\delta \subset J$ mit $d(F(A), x) \leq \delta$ fuer all $A \supset A_\delta$. Damit folgt dann fuer solche A sofort

$$e(h(F(A)), h(x)) \leq \varepsilon.$$

Das ist aber gerade die gewuenschte Aussage.) Wir werden diese Stetigkeit fuer Funktionen wie $\|\cdot\|$ und $\langle \cdot, \cdot \rangle$ anwenden.

Anwendung. Ist (e_j), $j \in J$, ein Orthonormalsystem und sind c_j, $j \in J$, Elemente aus $\mathbb{K}$, so definiert man F auf den endlichen Teilmengen von J durch

$$F(A) = \sum_{j \in A} c_j e_j.$$

Falls existiert, so schreibet man in diesem Fall

$$\sum_{j \in J} c_j e_j := \lim_{A \to J} F(A) = \lim_{A \to J} \sum_{j \in A} c_j e_j.$$

THEOREM (Darstellung mit Koeffizienten). *Sei $(H, \langle\cdot,\cdot\rangle)$ ein Hilbertraum und (e_j) ein Orthonormalsystem. Seien $c_j \in \mathbb{K}$ mit $\sum |c_j|^2 < \infty$ gegeben. Dann existiert*

$$x = \sum_{j\in J} c_j e_j = \lim_{A\to J} \sum_{j\in A} c_j e_j,$$

und es gilt

$$\|x\|^2 = \sum |c_j|^2.$$

Beweis. Es koennen nur abzaehlbar viele c_j nicht verschwinden. Daher koennen wir ohne Einschraenkung annehmen, dass die Indexmenge abzaehlbar ist. Wir zeigen zunaechst, dass

$$S_N := \sum_{j=1}^{N} c_j e_j$$

eine Cauchy Folge ist. Es gilt

$$\|S_N - S_M\|^2 = \sum_{j=N+1}^{M} |c_j|^2 \to 0, N, M \to \infty.$$

Daher konvergiert (S_N) gegen ein $x \in V$ (Hilbertraum). Wegen $\sum |c_j|^2 < \infty$ existiert zu jedem $\varepsilon > 0$ ein endliches A mit

$$\sum_{j\notin B} |c_j|^2 < \varepsilon$$

fuer alle $B \supset A$. Damit gilt dann fuer solche B

$$\|x - \sum_{j\in B} c_j e_j\|^2 = \lim_{N\to\infty} \|S_N - \sum_{j\in B} c_j e_j\|^2 \leq \sum_{j\notin A} |c_j|^2 < \varepsilon.$$

Damit ist $x = \lim_{A\to J} \sum_{j\in A} c_j e_j$ gezeigt.
Aufgrund der Stetigkeit der Norm folgt dann

$$\|x\|^2 = \lim_{A\to J} \| \sum_{j\in A} c_j e_j\|^2 = \lim_{A\to J} \sum_{j\in A} |c_j|^2 = \sum_{j\in J} |c_j|^2.$$

Das beendet den Beweis. □

Bemerkung. Der Satz liefert insbesondere, dass man die Reihe umsortieren kann. (Denn es kommt nur darauf an die 'wesentlichen' c_j beruecksichtigt zu haben.)

FOLGERUNG (Allgemeine Besselsche Ungleichung). *Sei H ein Hilbertraum mit Orthonormalsystem (e_j). Dann existiert fuer jedes $x \in H$ der Vektor*

$$y := \sum_{j\in J} \langle e_j, x\rangle e_j = \lim_{A\to J} \sum_{j\in A} \langle e_j, x\rangle e_j$$

und es gilt

$$x - y \perp e_j, j \in J, \text{ sowie } x - y \perp y.$$

Insbesondere gilt

$$\|x\|^2 = \|x - \sum \langle e_j, x\rangle e_j\|^2 + \sum |\langle e_j, x\rangle|^2.$$

Beweis. Nach dem vorigen Theorem und der Besselsche Ungleichung existiert

$$y := \sum \langle e_j, x\rangle e_j = \lim_{A\to J} y_A$$

mit

$$y_A := \sum_{j\in A} \langle e_j, x\rangle e_j$$

fuer $A \subset J$ endlich. Wir zeigen zunaechst $x - y \perp e_k$ unter Verwendung der Stetigkeit des Skalarproduktes:

$$\langle x-y, e_k\rangle = \lim_{A\to J} \langle x - y_A, e_k\rangle = \lim_{A\to J} (\langle x, e_k\rangle - \langle y_A, e_k\rangle) = \lim_{A\to J} 0 = 0.$$

Damit folgt $\langle x-y, y_A\rangle = 0$ fuer jede endliche Teilmenge A von J und damit

$$\langle x-y, y\rangle = 0.$$

Das liefert

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

Das beendet den Beweis. □

LEMMA (Charakterisierung Basis). *Sei $(H, \langle\cdot,\cdot\rangle)$ ein Hilbertraum und (e_j) ein Orthonormalsystem. Dann sind aequivalent:*

(i) *$\{e_j : j \in I\}$ ist maximal (d.h. jedes Orthonormalsystem $\{e'_\alpha : \alpha \in A\}$, das (e_j) enthaelt stimmt mit diesem ueberein).*
(ii) *Es gilt $\{e_j : j \in I\}^\perp = \{0\}$.*
(iii) *Es gilt $\overline{Lin\{e_j\}} = H$.*
(iv) *Fuer jedes $x \in H$ gilt $x = \sum\langle e_j, x\rangle e_j$.*
(v) *Fuer jedes $x \in H$ gilt $\|x\|^2 = \sum_j |\langle e_j, x\rangle|^2$. (Parsevalsche Gleichung)*

Beweis.
(i)$\Longrightarrow$ (ii): Waere die Aussage $\{e_j\}^\perp = \{0\}$ falsch, so gaebe es ein $x \perp \{e_j\}$ mit $x \neq 0$ und das widerspraeche der Maximalitaet von (e_j).

(ii)$\Longleftrightarrow$ (iii): Das wurde oben schon gezeigt.

(ii) / (iii)$\Longrightarrow$ (iv): Nach Besselscher Ungleichung gilt $\sum |\langle e_j, x\rangle|^2 < \infty$. Damit existiert nach dem Darstellungssatz $y = \sum\langle e_j, x\rangle e_j$. Es ist nach Konstruktion $x - y \in \{e_j\}^\perp$. Mit (ii) folgt dann $x - y = 0$ und damit

$$x = \sum\langle e_j, x\rangle e_j.$$

(iv)$\Longleftrightarrow$ (v): Das folgt sofort aus der (in der vorigen Folgerung gezeigten) Gleichung

$$\|x\|^2 = \|x-y\|^2 + \sum_j |\langle e_j, x\rangle|^2$$

mit $y = \sum_j \langle e_j, x\rangle e_j$.

(iv)$\Longrightarrow$ (i): Sei $x \perp e_j$ fuer alle $j \in I$. Nach (iv) kann man x darstellen als

$$x = \sum\langle e_j, x\rangle e_j = \sum 0 e_j = 0.$$

Das liefert (i). □

Definition (Orthonormalbasis). *Ein Orthonormalsystem in einem Hilbertraum heisst (Orthonormal)basis (ONB), wenn es eine der Eigenschaften des vorangehenden Lemmas erfuellt.*

Bemerkung. Im folgenden wird manchmal verkuerzt von Basis statt von Orthonormalbasis gesprochen. Im allgemeinen bedeutet aber Basis in einem Vektorraum etwas anderes als Orthonormalbasis, naemlich ein Menge von linear unabhaengigen Vektoren, so dass jeder Vektor als *endliche* Linearkombination aus diesen Vektoren dargestellt werden kann (*Hamelsche Basis*).

Notation. Ist $e_j, j \in I$, eine Basis im Hilbertraum, so kann man also nach dem vorigen Lemma jedes x aus dem Hilbertraum darstellen als

$$x = \sum_{j \in I} \langle e_j, x \rangle e_j.$$

Dies Darstellung heisst *Entwicklung von x* nach der Basis e_j, $j \in I$. Die Zahlen $\langle e_j, x \rangle$ heissen *Koeffizienten* der Entwicklung. Man kann sie als Koordinaten deuten. Tatsaechlich handelt es sich, falls der Hilbertraum der Euklidische Raum $\mathbb{K}^N$ ist und $e_j, j = 1, \ldots, N$ die Standardorthonormalbasis, genau um die Koordinaten. Denn fuer jedes Element $x \in \mathbb{K}^N$ gilt offenbar die Gleichung

$$x = \sum_{j=1}^{N} x_j e_j = \sum_{j=1}^{N} \langle e_j, x \rangle e_j.$$

Man kann nach Auswahl einer Basis in jedem Hilbertraum (fast) genauso mit Koordinaten 'rechnen' wie im Euklidischen Raum und das ist einer der grossen Vorteile von Hilbertraeumen. Das werden wir nun diskutieren:

Folgerung. *Sei $e_j, j \in I$, eine Orthonormalbasis in einem Hilbertraum H. Dann gilt fuer x und y aus H:*

- $x = \sum \langle e_j, x \rangle e_j$.
- $\|x\|^2 = \sum |\langle e_j, x \rangle|^2$.
- $\langle x, y \rangle = \sum_{j \in I} \langle x, e_j \rangle \langle e_j, y \rangle$.

Beweis. Die erste und zweite Aussage folgen sofort aus der Charakterisierung einer Basis im vorangehenden Lemma. Die letzte Eigenschaft folge einfach durch Grenzwertbildung. $\square$

Bemerkung. Die Eigenschaft (iv) eines Orthonormalsystem in obigem Lemma wird (in der Physik) auch als Vollstaendigkeitsrelation bezeichnet und als

$$I = \sum |e_j\rangle \langle e_j|$$

geschrieben.

Beispiel $\ell^2(X)$. (Dieses Beispiel wurde bereits behandelt, um ein uebarbzaehlbares ONS anzugeben; hier zeigen wir, dass dieses ONS sogar eine

ONB ist). Sei X eine beliebige Menge versehen mit der Potenzmenge $\mathcal{P}$ und dem Zaehlmass ζ. Dann ist

$$L^2(X, \mathcal{P}, \zeta) =: \ell^2(X)$$

ein Hilbertraum mit Skalarprodukt

$$\langle f, g \rangle = \sum_{x \in X} \overline{f(x)} g(x).$$

Eine Orthonormalbasis ist gegeben durch die e_x, $x \in X$,

$$e_x(y) = \delta_{x,y}.$$

Beweis. Offenbar bilden die $e_x, x \in X$ ein Orthonormalsystem. Es reicht also zu zeigen, dass ein $f \in \ell^2(X)$ mit $f \perp e_x$ fuer alle $x \in X$ verschwinden muss. Das ist aber klar. Das beendet den Beweis.

Speziallfaelle sind folgende:

Der Hilbertraum $\mathbb{K}^N$. Es gilt (offenbar)

$$\mathbb{K}^N = \ell^2(\{1, \ldots, N\}).$$

In diesem Fall handelt es sich bei den e_x, $x \in X$, gerade um die Standartnormalbasis $e_1, \ldots, e_N$.

ℓ^2. Es ist (der oben eingefuehrte Vektorraum mit Skalarprodukt) $\ell^2 := \ell^2(\mathbb{N})$ ein Hilbertraum. Die oben eingefuehrt Orthonormalbasis ist dann gegeben durch die $e_j, j \in \mathbb{N}$ mit $e_j(k) = \delta_{j,k}$.

Beispiel $L^2(\mathbb{T}^N)$. Sei $\mathbb{T}^N := \mathbb{R}^N / \mathbb{Z}^N$ mit dem Lebesguemass ausgestattet. Dann bilden die Funktionen

$$e_k : \mathbb{T}^N \longrightarrow \mathbb{C}, x + \mathbb{Z}^N \mapsto e^{2\pi i k x},$$

$k \in \mathbb{Z}^N$, eine Orthonormalbasis. Die Entwicklung

$$f = \sum_{k \in \mathbb{Z}^N} \langle e_k, f \rangle e_k$$

nach dieser Orthonormalbasis ist als *Fourierreihe* der Funktion f bekannt.
Beweis. Offenbar bilden die angegebenen Funktionen ein Orthonormalsystem. Der Beweis der Basiseigenschaft ist etwas aufwendiger. Wir werden das im kommenden Semester als einfache Folgerung erhalten.

Mittels Basen kann man die orthogonale Projektion auf einen Unterraum explizit ausrechnen:

FOLGERUNG. *Sei H Hilbertraum und U ein abgeschlossener Unterraum. Sei (e_j) eine Basis von U. Dann ist die orthogonale Projektion von H auf U gegeben durch*

$$P_U x := \sum_j \langle e_j, x \rangle e_j.$$

Beweis. Es gilt $x = y + z$ mit $y \in U$ und $z \perp U$. Wegen $y \in U$, der Konstruktion der (e_j) und wegen $z \perp U$ gilt

$$P_U x = y = \sum \langle e_j, y \rangle e_j = \sum \langle e_j, y + z \rangle e_j.$$

Das liefert die Aussage. $\square$

Die vorangehenden Betrachtungen zeigen den Nutzen von Basen. Als naechstes geht es darum, Existenz einer Basis zu zeigen.

Erinnerung. Eine Teilmenge A eines Hilbertraum heisst total, wenn $A^{\perp} = \{0\}$ gilt (d.h. wenn $Lin(A)$ dicht ist).

THEOREM. *Jeder Hilbertraum besitzt eine Basis. Diese Basis kann genau dann mit abzaehlbarer Indexmenge gewaehlt werden, wenn der Hilbertraum eine abzaehlbare totale Menge besitzt.*

Beweis. Wir unterscheiden zwei Faelle:

Fall 1: Der Hilbertraum hat eine abzaehlbar totale Menge. Anwenden des Gram/Schmidtschen Orthogonalisierungsverfahren auf diese Menge liefert ein abzaehlbares totales Orthonormalsystem. Dieses ist nach der Charakterisierung eine Basis.

Fall 2: Der Hilbertraum hat keine abzaehlbar totale Menge. Dann gibt es insbesondere auch keine abzaehlbare Orthonormalbasis (denn eine Orthonormalbasis ist immer total). Mit dem Zornschen Lemma kann man aber die Existenz eines maximalen Orthonormalsystem zeigen. Dieses ist nach der vorangegangenen Charakterisierung eine Orthonormalbasis. $\square$

Exkurs - Das Lemma von Zorn Wir diskutieren wir das Lemma von Zorn. Es liefert die Existenz maximaler Element in halbgeordneten Mengen (unter geeigneten Vorausetzungen).

Eine Menge M heisst *halbgeordnet* bzgl. einer Ordnungsrelation $\prec$, wenn gilt

(01) $a \prec a$ fuer alle $a \in M$
(02) $a \prec b$, $b \prec c$ impliziert $a \prec c$.
(03) $a \prec b$, $b \prec a$ impliziert $a = b$.

Nicht fuer jedes Paar (a, b) muss eine der Relationen $a \prec b$ oder $b \prec a$ gelten! Eine Teilmenge N von M heisst total geordnet, wenn fuer jedes Paar $(a, b) \in N \times N$ eine der Beziehungen $a \prec b$ oder $b \prec a$ gilt.
Ein Element $s \in M$ heisst obere Schranke einer Teilmenge R von M, wenn fuer jedes $r \in R$ gilt $r \prec s$.
Ein Element $m \in M$ heisst maximales Element in M, wenn aus $m \prec a$ fuer ein $a \in M$ folgt $m = a$ (d.h. wenn es kein echt groesseres Element in M gibt).

Beispiel. $\mathbb{R}$ mit $\leq$ oder $\mathbb{R}$ mit $\geq$. In diesem Fall gibt es kein maximales Element. Aber es hat jede beschraenkte abgeschlossene Menge ein maximales Element.

Beispiel. (M, d) metrischer Raum, $x \in M$.

- $K(x) := \{U_r(x) : r > 0\}$ mit $U_r(x) \prec U_s(x)$ falls $s < r$. Totalgeordnet.
- $U :=$ Umgebungen von x mit $U \prec V$ falls $V \subset U$. Nicht totalgeordnet.

In beiden Systemen gibt es (im allgemeinen) kein maximales Element. (Warum?)

Beispiel. Man kann leicht (endliche) halbgeordnete Mengen angeben, in denen kein maximales Element existiert und nicht alle Elemente gegenseitig vergleichbar sind.

LEMMA. *(von Zorn) Besitzt in einer halbgeordneten Menge jede geordnete Teilmenge eine obere Schranke, so existiert (mindestens) ein maximales Element.*

Bemerkung. Das Zornsche Lemma ist aequivalent zum *Auswahlaxiom*: Seien A_i, $i \in I$, nichtleere Mengen. Dann gibt es eine Funktion F auf I mit $F(i) \in A_i$ fuer alle $i \in I$, d.h. $\prod_{i \in I} A_i \neq \emptyset$. Das Auswahlaxiom fuehrt zu verblueffenden Konsequenzen (Banach/Tarski). Es ist mit den ueblichen Grundlagen der Mengenlehre / Logik vertraeglich. **Ende des Exkurses.**

Bemerkung. (a) Es sind aequivalent:

- Existenz einer abzaehlbaren Orthonormalbasis.
- Existenz einer abzaehlbaren totalen Menge.
- Existenz einer abzaehlbaren dichten Menge.

(Bew. Die Aequivalenz der ersten beiden Punkte folgt aus dem Theorem. Zur Aequivalenz der letzten beiden Punkte: Eine Richtung ist klar. Die andere Richtung folgt durch Bilden von Linearkombinationen mit rationalen Koeffizienten.)
(b) Wenn der Hilbertraum eine abzaehlbare Orthonormalbasis besitzt, so sind alle Orthonormalbasen abzaehlbar. (Uebung.)

DEFINITION. *Ein Hilbertraum heisst separabel, wenn er eine abzaehlbares totale Menge besitzt.*

Beispiel. $\ell^2 = \ell^2(\mathbb{N})$ Offenbar ist ℓ^2 ein separabler Hilbertraum.

Das ist in gewisser Weise das allgemeinste Beispiel eines separablen Hilbertraumes.

THEOREM (Separable Hilbertraeume sind ℓ^2). *Sei H ein beliebiger separabler Hilbertraum und (e_j), $j \in \mathbb{N}$, eine Orthonormalbasis. Dann gibt es genau eine lineare stetige Abbildung*

$$J : \ell^2 \longrightarrow H \ \textit{ mit } \ J(\delta_n) = e_n$$

fuer alle $n \in \mathbb{N}$. Diese Abbildung ist bijektiv und es gilt $\langle Jx, Jy \rangle = \langle x, y \rangle$.

Beweis. Existenz. Fuer $x \in \ell^2$ ist $\sum |x(k)|^2$ endlich und damit existiert in H auch $\sum x(k)e_k$. Wir definieren $J(x) := \sum x(k)e_k$. Dann ist (offenbar) J linear und es gilt aufgrund der oben gezeigten Saetze

$$\langle J(x), J(x) \rangle = \|J(x)\|^2 = \sum |x(k)|^2 = \|x\|^2 = \langle x, x \rangle.$$

Damit ist J eine Isometrie. Offenbar gilt $J(\delta_n) = e_n$.

Eindeutigkeit. Da $Lin\{\delta_n\}$ dicht ist, folgt die Eindeutigkeit.

Zur letzten Aussage: Da die Abbildung eine Isometrie ist, ist sie injektiv. Jedes $x \in H$ laesst sich als $\sum c_j e_j$ mit $\sum |c_j|^2 < \infty$ darstellen und erfuellt also $x = J((c_j))$. Damit ist J surjektiv.

Mit Polarisation folgt aus Normerhaltung leicht, dass

$$\langle J(x), J(y)\rangle = \langle x, y\rangle$$

fuer alle $x, y \in \ell^2$. □

Bemerkung. Der vorangehende Satz ist von grosser struktureller Relevanz. Er besagt, dass man - sofern es die Geometrie des Hilbertraumes betrifft - sich im separablen Fall auf Betrachtung von ℓ^2 einschraenken kann. Das spielt auch bei der Entwicklung der Quantenmechank von eine Rolle: Die auf Heisenberg zurueckgehende 'Matrizenmechanik' basiert auf dem Hilbertraum ℓ^2. Die von Schroedinger entwickelte 'Wellenmachanik' basiert auf dem Hilbertraum $L^2(\mathbb{R}^N, \lambda)$. Die Aequivalenz der beiden Zugaenge beruht maßgeblich auf dem vorangehenden Satz.

Fuer beliebige (nicht notwendig separable) Hilbertraeume gilt eine ganz analoge Aussage mit ganz analogem Beweis.

Theorem (Jeder Hilbertraum ist $\ell^2(S)$ fuer ein geeignetes S). *Sei H ein beliebiger Hilbertraum und (e_s), $s \in S$, eine Orthonormalbasis. Dann gibt es genau eine lineare stetige Abbildung*

$$J : \ell^2(S) \longrightarrow H \;\; mit \;\; J(1_s) = e_s$$

fuer alle $s \in S$. Diese Abbildung ist bijektiv und es gilt $\langle Jx, Jy\rangle = \langle x, y\rangle$.

Es laesst sich die *Dimension* eines Hilbertraumes definieren als Mächtigkeit einer ONB. Dazu muss man natuerlich zeigen, dass die Mächtigkeit nicht von der Wahl der ONB abhaengt. Im Falle von endlichen bzw. abzaehlbaren ONB wissen wir das (s.o.).

4. Der Rieszsche Darstellungssatz

Eine wesentliche Methode zum Studium mathematischer Objekte ist die Untersuchung (gewisser) Abbildungen auf ihnen. Im Falle von normierten Raeumen sind die naheliegenden Abbildungen die stetigen linearen Abbildungen in den Koerper. Die Menge dieser Abbildungen heisst Dualraum. Hier bestimmen wir noch den Dualraum eines Hilbertraum.

Wir beginnen mit einigen Definitionen. Sei $(H, \langle\cdot,\cdot\rangle)$ eine Hilbertraum. Eine Abbildung $\varphi : H \longrightarrow \mathbb{K}$ hiesst *linear*, wenn gilt $\varphi(\lambda x + y) = \lambda\varphi(x) + \varphi(y)$ fuer alle $x, y \in H$ und $\lambda \in \mathbb{K}$. Eine solche Abbildung heisst dann auch *lineares Funktional* auf H. Die Menge aller stetigen linearen Funktionale heisst der *Dualraum* von H und wird mit H' bezeichnet. (Die Menge aller linearen Funktionale wird auch als *algebraischer Dualraum* bezeichnet). Der Dualraum wird offenbar durch

$$\lambda\varphi + \psi : H \longrightarrow \mathbb{K},\; x \mapsto \lambda\varphi(x) + \psi(x),$$

zu einem Vektorraum. Auf dem Dualraum von H wird durch

$$\|\varphi\| := \sup\{|\varphi(x)| : \|x\| \le 1\}$$

eine Norm definiert (wie man leicht sieht). Tatsaechlich hat der Dualraum eines Hilbertraumes wesentlich mehr Struktur als nur die eines normierten Raumes. Das ist der Inhalt des naechsten Satzes.

THEOREM (Rieszscher Darstellungssatz). *Sei* $(H, \langle\cdot,\cdot\rangle)$ *ein Hilbertraum. Dann erzeugt jedes* $y \in H$ *durch*

$$F_y : H \longrightarrow \mathbb{K},\ y \mapsto \langle y, \cdot\rangle,$$

ein stetiges lineares Funktional auf H *mit* $\|F_y\| = \|y\|$. *Die Abbildung*

$$H \longrightarrow H', y \mapsto F_y,$$

ist antilinear (d.h. $F_{\lambda x+\mu z} = \overline{\lambda} F_x + \overline{\mu} F_z$*) und bijektiv. Insbesondere wird also jedes stetige lineare Funktional auf* H *eindeutig durch ein* F_y *dargestellt.*

Bemerkung.

- Die entscheidende Aussage ist die Surjektivitaet von F.
- Der Satz gibt eine mathematische Behandlung der in der Physik ueblichen 'Bracket' Notation, in der die Vektor $|x\rangle$ mit dem 'Vektor' (eigentlich dem durch den Vektor y induzierten Element des Dualraumes) $\langle y|$ zum Sklalar $\langle y, x\rangle$ zusammen gebracht wird.

Beweis. Fuer jedes $y \in H$ ist F_y offenbar ein lineares Funktional mit

$$|F_y(x)| = |\langle y, x\rangle| \leq \|y\|\|x\|.$$

Also ist F_y stetig mit $\|F_y\| \leq \|y\|$. Wegen $F_y(y) = \|y\|^2$ gilt sogar $\|F_y\| = \|y\|$. Die Antilinearitaet ist einfach zu zeigen. Da $y \mapsto F_y$ isometrisch ist, folgt aus der (Anti)linearitaet, dass $y \mapsto F_y$ injektiv ist.

Noch z.z. F ist surjektiv: Sei $\varphi \in H'$. Wir unterscheiden zwei Faelle:

$\varphi \equiv 0$. Dann koennen (muessen ;-) wir $y = 0$ waehlen.

Es gilt nicht $\varphi \equiv 0$. Dann ist $N := Ker\varphi$ ein abgeschlossener echter Unterraum von H. Damit ist $N^\perp \neq \{0\}$. (Sonst: $N = N^{\perp\perp} = H$, also $\varphi \equiv 0$). Sei $z \in N^\perp$ mit $z \neq 0$ beliebig. Dann gehoert fuer jedes $x \in H$ das Element

$$\varphi(x)z - \varphi(z)x$$

zu $Ker\varphi = N$ (Nachrechnen!). Also gilt wegen $z \in N^\perp$ dann

$$0 = \langle z, \varphi(x)z - \varphi(z)x\rangle = \varphi(x)\langle z, z\rangle - \varphi(z)\langle z, x\rangle.$$

Damit folgt

$$\varphi(x) = \frac{\varphi(z)}{\|z\|^2}\langle z, x\rangle = \langle \overline{\frac{\varphi(z)}{\|z\|^2}} z, x\rangle.$$

Das liefert die Behauptung mit $y = \overline{\frac{\varphi(z)}{\|z\|^2}} z$. □

Bemerkung. Es mag (zunaechst) verblueffen, dass wir irgendein $z \in N^\perp$ waehlen konnten. Aber nachtraeglich wird klar, das $N = \{y\}^\perp = \{z\}^\perp$ gilt und damit

$$N^\perp = (Lin\{z\})^{\perp\perp} = Lin\{z\}$$

eindimensional ist. (Zeichnung: Kern als Hyperebenen, Niveauflaechen.) Damit kann man auch einen alternativen Beweis wie folgt geben: Sei $N := Ker\varphi$.
Schritt 1: Es ist $N^\perp$ eindimensional.
(Bew. $x \in N^\perp$, $z \in N^\perp$. Dann ist

$$x - \frac{\varphi(x)}{\varphi(z)} z \in N \cap N^\perp = \{0\}.$$

Damit sind x und z linear abhaengig.)
Schritt 2: Sei nun $z \in N^{\perp}$ mit $\varphi(z) = 1$. Dann kann man jedes $x \in H$ eindeutig schreiben als

$$x = y + \lambda z$$

mit $y \in N$ und

$$\lambda = \frac{\langle z, x\rangle}{\langle z, z\rangle}.$$

(Bew. Da N eindimensional ist, gibt es auf jeden Fall eine Darstellung von x der Form $x = y + \widetilde{z}$ mit $y \in N$ und $\widetilde{z} = \lambda z$. Die Formel fuer λ folgt dann sofort durch Bilden des Skalarproduktes mit z.
Schritt 3: Es gilt $\varphi(x) = \lambda = \frac{\langle z,x\rangle}{\langle z,z\rangle}$.
(Bew. Das folgt aus Schritt 2 wegen $\varphi(z) = 1$.)

5. Etwas zu schwacher Konvergenz*

Schwache Konvergenz im Hilbertraum ist ein nützliches Hilfsmittel fuer weitergehende Untersuchungen. In dieser Vorlesung spielt das keine groessere Rolle. Allerdings gibt es eine Reihe von Uebungsaufgaben, in denen schwache Konvergenz eine Rolle spielt. Die entsprechenden Ergebnisse werden in diesem Abschnitt kurz zusammengestellt.

Sei H ein Hilbertraum. Eine Folge (x_n) in H heisst schwach konvergent (gegen $x \in H$), wenn gilt

$$\lim_{n\to\infty} \langle y, x_n\rangle = \langle y, x\rangle$$

fuer alle $y \in H$.

Jedes Orthonormalsystem ist - offenbar - ein Beispiel fuer eine schwach (gegen 0) konvergente Folge. Insbesondere ist also schwache Konvergenz tatsaechlich schwaecher als die uebliche (Norm)Konvergenz im Hilbertraum. Das genauer Verhaeltnis der beiden Konzepte von Konvergenz ist in folgender Proposition gegeben.

Proposition. *Sei (x_n) eine Folge im Hilbertraum H. Dann konvergiert x_n gegen x genau dann, wenn es schwach gegen x konvergiert und $\|x_n\| \to \|x\|$ gilt.*

Schwache Konvergenz kann als starke Form der Beschraenktheit verstanden werden:

Proposition. *Jede schwach konvergente Folge ist beschraenkt.*

Beweis. □

Tatsaechlich gilt auch eine Art Umkehrung.

Proposition. *Jede beschraenkte Folge im Hilbertraum hat eine schwach konvergente Teilfolge.*

Als Folgerung ergibt sich sofort:

Folgerung. *Ist S eine beschraenkte Menge im Hilbertraum, so hat jede Folge aus S eine schwach konvergente Teilfolge.*

KAPITEL 4

Beschraenkte Operatoren im Hilbertraum

In diesem Abschnitt werfen wir einen ersten Blick auf lineare Operatoren zwischen Hilbertraeumen. Uns wird dabei besonders die Situation interessieren, dass die Raeume unendlichdimensional sind. Denn im endlichdimensionalen Fall gelten automatisch eine ganze Reihe von Eigenschaften:
Sei H ein endlichdimensionaler Vektorraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$. Dann gilt (Uebung):

(a) Es ist $(H, \langle \cdot, \cdot \rangle)$ ein Hilbertraum (d.h. es ist H vollständig bzgl. der Norm $\|x\| = \langle x, x \rangle^{1/2}$).
(b) Jede lineare Abbildung von H nach H ist stetig.
(c) Eine Abbildung $A : H \longrightarrow H$ ist genau dann injektiv, wenn sie surjektiv ist.

1. Grundlegendes

Definition. *Seien H und K Hilbertraeume. Eine Abbildung $A : H \longrightarrow K$ heisst linear, wenn*

$$A(\lambda x + \mu y) = \lambda A(x) + \mu A(y)$$

fuer alle $x, y \in H$, $\lambda, \mu \in \mathbb{K}$ gilt.

Notation. Eine lineare Abbildung wird auch als *linearer Operator* oder auch nur *Operator* bezeichnet. Oft laesst man dann die Klammern um das Argument weg und schreibt also Ax statt $A(x)$ (wie es auch bei Matrizen ueblich ist). Fuer spaeter fuehren wir zu einem linearen Operator A noch das Bild von A,

$$\mathrm{Ran}(A) := \{Ax : x \in H\}$$

und den Kern von A,

$$\mathrm{Ker}(A) := \{x \in H : Ax = 0\}$$

ein.

Bemerkung. Den Fall $K = \mathbb{K}$ haben wir im vorigen Abschnitt schon behandelt.

Uns wird es (zunaechst) vor allem um stetige lineare Abbildungen zwischen Hilbertraeumen gehen. Offenbar bilden die stetigen linearen Abbildungen zwischen den Hilbertraeumen H und K einen Vektorraum. Dieser wird mit $L(H, K)$ bezeichnet.

Proposition (Charakterisierung Stetigkeit). *Seien H und K Hilbertraeume mit zugehoerigen Normen $\| \cdot \|_H$ und $\| \cdot \|_K$ und $A : H \longrightarrow K$ linear. Dann sind die folgenden Aussagen aequivalent:*

(i) *Es ist A stetig.*
(ii) *Es ist A stetig in* 0.
(iii) *Es gibt ein $C \geq 0$ mit*
$$\|Ax\|_K \leq C\|x\|_H$$
fuer alle $x \in H$.

Bemerkung. Lineare Operatoren, die (iii) erfuellen, werden als *beschraenkt* bezeichnet. Die Proposition besagt also, dass ein Operator genau dann beschraenkt ist, wenn er stetig ist.

Beweis. (i)$\Longrightarrow$ (ii): Klar.
(ii)$\Longrightarrow$ (iii): Da A stetig in 0 ist, gibt es ein $\delta > 0$ mit $\|Ax\|_H \leq 1$ fuer alle $x \in H$ mit $\|x\|_H \leq \delta$. Damit gilt dann also fuer alle $x \in H$ mit $x \neq 0$
$$\|Ax\|_K = \frac{\|x\|}{\delta}\|A(\frac{\delta}{\|x\|}x)\|_K \leq \frac{\|x\|}{\delta}.$$
(iii)$\Longrightarrow$ (i): Es gilt
$$\|Ax - Ay\|_K = \|A(x-y)\|_K \leq C\|x-y\|_K$$
und die Behauptung folgt. □

Das C in (iii) ist natuerlich nicht eindeutig. Tatsaechlich ist mit jedem solchen C auch jedes groessere C' moeglich. Damit stellt sich die Frage nach dem kleinsten solchen C.

Proposition. *Ist $A : H \longrightarrow K$ ein linearer Operator zwischen Hilbertraeumen, so gilt*
$$\inf\{C \geq 0 : \|Ax\| \leq C\|x\| \textit{fuer alle } x \in H\} = \sup\{\|Ax\| : \|x\| \leq 1\} =: S$$
sowie
$$\|Ax\| \leq S\|x\|$$
fuer alle $x \in H$. (Hier ist der Wert ∞ fuer S erlaubt.)

Beweis. Bezeichne das Infimum in der Aussage mit I und das Supremum mit S.
$S \leq I$: Sei $C \geq 0$ mit $\|Ax\| \leq C\|x\|$ fuer alle $x \in H$. Dann folgt $S \leq C$. Da dies fuer alle solchen C gilt, folgt $S \leq I$.
$I \leq S$: Gilt $S = \infty$, so ist die Aussage klar. Sei also nun $S < \infty$. Weiterhin folgt aus der Definition von S fuer $x \neq 0$ sofort
$$\|Ax\| = \|x\|\|A\frac{1}{\|x\|}x\| \leq S\|x\|.$$
Das liefert $I \leq S$ und auch die letzte Aussage. □

Bemerkung. Fuer stetige Operatoren liefert die Proposition insbesondere dass das Infimum angenommen wird (also ein Minimum ist). Es ist erst einmal nicht klar, ob das Supremum angenommen wird. Tatsaechlich ist das im allgemeinen falsch (Uebung).

Definition. *Ist $A : H \longrightarrow K$ ein linearer Operator zwischen Hilbertraeumen, so definiert man*
$$\|A\| := \sup\{\|Ax\| : \|x\| \leq 1\}.$$

Dann gilt also nach dem schon gezeigte $\|A\| < \infty$ genau dann, wenn A ein stetiger Operator ist. Tatsaechlich hat $\|\cdot\|$ auf dem Vektorraum der stetigen Operatoren viele schoene Eigenschaften:

THEOREM (Vollstaendigkeit von $L(H, K)$). *Seien H und K Hilbertraeume. Dann definiert $\|\cdot\|$ eine Norm auf $L(H, K)$. Bezueglich der durch diese Norm induzierten Metrik $d(A, B) := \|A - B\|$ wird $L(H, K)$ zu einem vollstaendigen Raum.*

Beweis. Es ist einfach zu sehen, dass $\|\cdot\|$ eine Norm definiert. Wir zeigen nun die Vollstaendigkeit. Sei (A_n) eine Cauchy-Folge in $L(H, K)$. Wegen

$$\|A_n x - A_m x\| \leq \|A_n - A_m\| \|x\|$$

ist dann insbesondere also $(A_n x)$ eine Cauchy-Folge in K fuer jedes $x \in H$. Da K vollstaendig ist, konvergiert diese Cauchy-Folge. Wir koennen also die Abbildung

$$A : H \longrightarrow K \text{ definieren durch } A(x) := \lim A_n x.$$

Da jedes A_n linear ist folgt

$$A(\lambda x + \mu y) = \lim_n A_n(\lambda x + \mu y) = \lim_n (\lambda A_n x + \mu A_n y) = \lambda A(x) + \mu A(y)$$

fuer alle $x, y \in H$ und $\lambda, \mu \in \mathbb{K}$. Damit ist also A linear.
Es bleibt zu zeigen, dass A stetig ist und (A_n) im Sinne von $\|\cdot\|$ gegen A konvergiert.
Sei dazu nun zu $\varepsilon > 0$ eine natuerliche Zaehl N gewaehlt mit

$$\|A_n - A_m\| \leq \varepsilon$$

fuer alle $n, m \geq N$. Dann gilt fuer $n \geq N$ und jedes $x \in H$ mit $\|x\| \leq 1$

$$\|(A - A_n)x\| = \lim_m \|A_m x - A_n x\| \leq \varepsilon \|x\|.$$

Damit ist also $A - A_n$ ein beschraenkter linearer Operator. Damit ist $A = A - A_N + A_N$ ein beschraenkter Operator und es folgt

$$\|A - A_n\| \to 0, n \to \infty.$$

Das beendet den Beweis. □

Bemerkung. Der Beweis nutzt die Vollstaendigkeit von K. Die Vollstaendigkeit von H spielt keine Rolle.

PROPOSITION (Submultiplikativitaet der Norm). *Sind $A : H \longrightarrow K$ und $B : K \longrightarrow L$ lineare stetige Operatoren zwischen Hilbertraeumen, so gilt*

$$\|BA\| \leq \|B\| \|A\|.$$

Beweis. Das ist einfach. (Uebung) □

Wir betrachten nun einige *Beispiele.*

Seien dazu zunaechst $(X, \mathcal{A}, \mu)$ und $(Y, \mathcal{B}, \nu)$ Massraeume und $H = L^2(X, \mu)$ und $K = L^2(Y, \nu)$.

Multiplikationsoperatoren. Sei $\varphi : X \longrightarrow \mathbb{K}$ messbar und beschraenkt. Dann ist

$$M_\varphi : L^2(X, \mu) \longrightarrow L^2(X, \mu), \ M_\varphi f = \varphi f,$$

ein linearer beschraenkter Operator mit

$$\|M_\varphi\| \leq \|\varphi\|_\infty.$$

Er wird der *Operator der Multiplikation* mit φ genannt. Ist $(X, \mathcal{A}, \mu)$ σ-endlich, so gilt sogar Gleicheit in obiger Formel fuer die Norm (Uebung).
(Bew. Linear: klar. Normabschaetzung: klar.)

Ist $X = \{1, \ldots, N\}$ mit dem Zaehlmass, so wird M_φ durch eine Diagonalmatrix mit Eintraegen $(\varphi(n))$ dargestellt. Aehnlich verhaelt es sich fuer $X = \mathbb{N}$.

Hilbert-Schmidt Operatoren. Sei $k : X \times Y \longrightarrow \mathbb{K}$ messbar mit

$$\|k\|_2^2 := \int_{X \times Y} |k(x,y)|^2 d\mu \times \nu(x,y) < \infty.$$

Dann ist

$$K : L^2(Y, \nu) \longrightarrow L^2(X, \mu), Kf(x) = \int k(x,y) f(y) d\nu(y),$$

ein linearer beschraenkter Operator mit

$$\|K\| \leq \|k\|_2.$$

Ein solcher Operator heisst *Hilbert-Schmidt Operator mit Kern k.*
(Bew. Es ist $(x,y) \mapsto k(x,y)f(y)$ messbar. Nach dem Satz von Fubini folgt weiterhin

$$\begin{aligned} \int_X \left(\int_X |k(x,y) f(y)| d\nu(y) \right)^2 d\mu(x) &\leq \int_X \int_Y |k(x,y)|^2 d\nu(y) \int_Y |f(z)|^2 d\nu(z) d\mu(x) \\ &= \|f\|^2 \cdot \int_{X \times Y} |k(x,y)|^2 d\mu(x) \nu(y). \end{aligned}$$

Damit ist $\int k(x,y) f(y) d\nu(y)$ fuer μ fast alle $x \in X$ endlich und

$$x \mapsto \int k(x,y) f(y) d\nu(y)$$

gehoert zu $L^2(X, \mu)$. Weiterhin zeigt dies auch einfach die gewueschte Abschaetzung.)
Eine wesentliche Eigenschaft von Hilbert-Schmidt Operatoren ist folgende.

PROPOSITION. *Ist K ein Hilbert-Schmidt Operator von $L^2(Y, \nu)$ nach $L^2(X, \mu)$ mit Kern k, so gilt fuer jede Orthonormalbasis (e_j) von $L^2(Y, \nu)$*

$$\|k\|_2^2 = \sum_j \|K e_j\|^2.$$

(Insbesondere ist also die rechte Seite endlich und unabhaengig von der Orthonormalbasis.)

Bemerkung. Es sind Hilbert-Schmidt Operatoren durch die Endlichkeit der Summe charakerisiert (Uebung) und so lassen sich dann auch Hilbert-Schmidt Operatoren in allgemeinen Hilbertraeumen definieren.

Beweis. Wir rechnen

$$\begin{aligned}
\sum_j \|Ke_j\|^2 &= \sum_j \int_X |Ke_j(x)|^2 d\mu(x) \\
&= \int_X \sum_j \left| \int_Y k(x,y) e_j(y) d\nu(y) \right|^2 d\mu(x) \\
&= \int_X \sum_j |\langle \overline{k(x,\cdot)}, e_j \rangle|^2 d\mu(x) \\
\text{(Vollstaendigkeit)} \quad &= \int_X \|k(x,\cdot)\|^2_{L^2(Y,\nu)} d\mu(x) \\
&= \int_X \int_Y |k(x,y)|^2 d\nu(y) d\mu(x) \\
&= \|k\|_2^2.
\end{aligned}$$

Das beendet den Beweis. □

Operatoren auf $\ell^2(X)$. Sei $A : \ell^2(X) \longrightarrow \ell^2(X)$ ein beschraenkter Operator. Dann ist fuer jedes $x \in X$ also die Abbildung

$$\ell^2(X) \longrightarrow \mathbb{K}, f \mapsto Af(x),$$

stetig und linear (als Verknuepfung stetiger linearer Abbildungen). Damit existiert also nach dem Rieszschen Darstellungssatz fuer jedes $x \in X$ ein $a_x \in \ell^2(X)$ mit

$$Af(x) = \langle a_x, f \rangle = \sum_{x \in X} \overline{a_x(y)} f(y).$$

Definiert man

$$a : X \times X \longrightarrow \mathbb{K},\ a(x,y) = \overline{a_x(y)},$$

so gilt also

$$Af(x) = \langle a_x, f \rangle = \sum_{y \in Y} a(x,y) f(y).$$

Es heisst a die *Matrix* von A.
Nutzt man, dass jeder Hilbertraum die Form $\ell^2(X)$ hat (nach Wahl einer ONB), so kann man fuer Operatoren in beliebigen Hilbertraeumen von zugeorigen Matrizen und Matrixelementen sprechen (nachdem man eine ONB gewaehlt hat).

Bemerkung. Nach Konstruktion hat die Matrix a von A quadratsummierbare Zeilen (wegen $a_x \in \ell^2(X)$ fuer alle $x \in X$). Tatsaechlich hat die Matrix auch quadrastsummierbare Spalten (wegen $a(\cdot, y) = A1_y \in \ell^2(X)$) fuer alle $y \in X$). Es ist aber im allgemeinen nicht der Fall, dass a selber quadratsummierbar ist. Tatsaechlich gilt das genau dann, wenn A ein Hilbert-Schmidt Operator ist (s.o.). Eine effektive Charakterisierung der beschraenkten Operatoren durch Summierbarkeitsbedingungen an die Matrix gibt es (im unendlich dimensionalen Hilbertraum) nicht.

2. Ein erster Blick auf Spektraltheorie

Wir werden uns hier kurz mit der Frage der stetigen Invertierbarkeit von $A-\lambda I$ befassen. Dazu zunaechst eine kurze **Erinnerung**: Ein linearer Operator T im Hilbertraum $\mathcal{H}$ ist genau dann bijektiv, wenn ein linearer Operator B existiert mit $AB = I = BA$. Dann heisst A auch *invertierbar* und B die *Inverse* zu A. Ist B stetig, so heisst A *stetig invertierbar*.

Fuer einen Operator T im Hilbertraum nennt man

$$\varrho(T) := \{z \in \mathbb{C} : T - zI \text{ bijektiv mit stetiger Inverser}\}$$

die *Resolventenmenge* von T und

$$\sigma(T) := \mathbb{C} \setminus \varrho(T)$$

das *Spektrum* von T. Man beachte, dass die Inverse einer bijektiven linearen Abbildung automatisch wieder eine lineare Abbildung ist.

Bemerkung. Tatsaechlich ist die Inverse einer beschraenkten linearen Abbildung automatisch wieder beschraenkt. Das werden wir aber erst spaeter beweisen (Satz von der stetigen Inversen).

Beispiel. Ist $X = \{1, \ldots, N\}$ mit dem Zaehlmass, so gilt $L^2(X, \zeta) = \mathbb{K}^N$ und es ist jede lineare Abbildung A in $L^2(X, \zeta)$ stetig und durch eine Matrix a gegeben. Dann gilt

$$\sigma(A) = \text{Eigenwerte der Matrix } a.$$

(Beweis. Denn es ist (wegen der Endlichdimensionalitaet) A invertierbar genau dann, wenn es stetig invertierbar ist und das gilt genau dann, wenn A injektiv ist.)

Tatsaechlich lassen sich nicht nur in endlichdimensionalen Hilbertraeumen, sondern in beliebigen Hilbertraeumen Eigenwerte und Eigenvektoren in der naheliegenden Weise definieren: Ist T ein beschraenkter Operator im Hilbertraum H und gilt $(T - z)x = 0$ fuer ein $x \in H$ mit $x \neq 0$ und $z \in \mathbb{K}$, so heißt λ ein *Eigenwert von* T und x ein Eigenvektor (zu λ). Offenbar gehoert jeder Eigenwert zum Spektrum. Die Umkehrung gilt im unendlichdimensionalen Hilbertraum jedoch nicht, wie man sich leicht an Beispielen klarmachen kann (Uebung).

Grundlegend fuer die naehere Untersuchung von Spektrum und Resolvente ist die folgende Operatorvariante der geometrischen Reihe:

Proposition. *Ist T ein beschraenkter Operator mit $\|T\| < 1$, so ist $I - T$ invertierbar und die Inverse S ist gegeben durch die normkonvergente Reihe $\sum_{k=0}^{\infty} T^k$.*

Beispiel. Beispiel : Multiplikationsoperatoren. Sei $(X, \mathcal{A}, \mu)$ ein Massraum mit σ-endlichem Mass μ. Sei $\psi : X \longrightarrow \mathbb{C}$ messbar und beschraenkt. Dann definert man den *Operator der Multiplikation mit f auf* $L^2(X, \mu)$ durch

$$M_\psi : L^2(X, \mu) \longrightarrow L^2(X, \mu), M_\psi f = \psi f.$$

Dann gilt

$$\sigma(M_\psi) = \{\lambda \in \mathbb{C} : \mu(\psi^{-1}(U_\varepsilon(\lambda))) > 0 \text{fuer alle } \varepsilon > 0\}$$

(Wesentlicher Wertebereich von ψ). (Uebung)

3. Adjungierte Operatoren

In den bisherigen Ueberlegungen zu Operatoren haben wir das Skalarprodukt noch nicht wirklich genutzt.

PROPOSITION (Adjungierte Operator). *Seien H und K Hilbertraeume und $A : H \longrightarrow K$ ein beschraenkter Operator. Dann existiert ein eindeutiger linearer Operator $A^* : K \longrightarrow H$ mit*

$$\langle g, Af\rangle_K = \langle A^* g, f\rangle_H$$

fuer alle $f \in H$ und $g \in K$. Es gilt

$$\|A\| = \|A^*\|.$$

Beweis. Eindeutigkeit. Das ist klar (da alle Skalarprodukte von A^*g festgelegt sind).
Existenz. Betrachte (bei festem $g \in K$) die Abbildung

$$H \longrightarrow \mathbb{K}, f \mapsto \langle g, Af\rangle_K.$$

Diese Abbildung ist offenbar linear und wegen

$$|\langle g, Af\rangle| \leq \|g\|\|A\|\|f\|$$

auch stetig. Damit gibt es nach dem Rieszschen Darstellungssatz ein eindeutiges $\widetilde{g} \in H$ mit

$$\langle g, Af\rangle_K = \langle \widetilde{g}, f\rangle_H$$

fuer alle $f \in H$. Offenbar ist die Abbildung

$$K \longrightarrow H, g \mapsto \widetilde{g},$$

linear. Definiere nun $A^*g := \widetilde{g}$.
Zur letzten Aussage: Direkte Rechnung liefert

$$\begin{aligned} \|A\| &= \sup\{\|Ax\| : \|x\| \leq 1\} \\ &= \sup\{|\langle Ax, y\rangle| : \|x\|, \|y\| \leq 1\} \\ &= \sup\{|\langle x, A^*y\rangle| : \|x\|, \|y\| \leq 1\} \\ &= \sup\{\|A^*y\| : \|y\| \leq 1\} \\ &= \|A^*\|. \end{aligned}$$

Das beendet den Beweis. □

DEFINITION. *Es heisst A^* der zu A adjungierte Operator.*

PROPOSITION (Die Abbildung $*$ als Involution). *Seien A und B beschraenkte lineare Operatoren vom Hilbertraum H in den Hilbertaum K. Dann gilt:*
(a) $(A + \lambda B)^* = A^* + \overline{\lambda} B^*$
(b) $(A^*)^* = A$.
(c) Ist I die Identitaet auf H, so gilt $I^* = I$.

Beweis. Das ist einfach. □

PROPOSITION (Die Abbildung $*$ und Multiplikation). *Sind $A : H \longrightarrow K$ und $B : K \longrightarrow L$ beschraenkte Operatoren zwischen Hilbertraeumen, so gilt*

$$(BA)^* = A^* B^*.$$

Beweis. Das ist einfach. □

FOLGERUNG (Die Abbildung $*$ und das Spektrum). *Ist $A : H \longrightarrow H$ ein beschraenkter Operator im Hilbertraum H, so gilt $\sigma(A^*) = \overline{\sigma(A)}$.*

Beweis. Nach den beiden vorigen Propositionen gilt

$$(A^* - \overline{\lambda})B^* = (A - \lambda)^* B^* = (B(A - \lambda))^*$$

und entsprechend

$$B^*(A^* - \overline{\lambda})B^* = B^*(A - \lambda)^* = ((A - \lambda)B)^*$$

sowie $I^* = I$. □

PROPOSITION. *Ist $A : H \longrightarrow K$ ein beschraenkter Operator zwischen Hilbertraeumen, so gilt*

$$\|A\|^2 = \|A^*A\| = \|AA^*\|.$$

Beweis. Es reicht die erste Gleichung zu zeigen. Wir zeigen zwei Ungleichungen: Es gilt

$$\|A^*A\| \leq \|A^*\| \| A\| = \|A\|^2.$$

(Hier nutzen wir im ersten Schritt die Submultiplikativitaet der Norm und im zweiten Schritt die Gleichheit $\|A\| = \|A^*\|$.)
Weiterhin gilt

$$\begin{aligned} \|A^*A\| &= \sup\{\|A^*Ax\| : \|x\| \leq 1\} \\ &\geq \sup\{|\langle A^*Ax, x\rangle| : \|x\| \leq 1\} \\ &= \sup\{|\langle Ax, Ax\rangle : \|x\| \leq 1\} \\ &= \sup\{\|Ax\|^2 : \|x\| \leq 1\} \\ &= \|A\|^2. \end{aligned}$$

Das beendet den Beweis. □

Bemerkung. Die vorangehende Gleichung ist von grosser struktureller Relevanz. Betrachtet man den Vektorraum, $L(H)$, aller stetigen linearen Operatoren vom Hilbertraum H in sich selber, so gilt:

- $L(H)$ ist eine Algebra mit einer Norm $\| \cdot \|$ und einer Involution $*$.
- Die Norm $\| \cdot \|$ ist submultiplikativ und es ist $L(H)$ vollstaendig.
- Es gilt $\|A^*A\| = \|AA^*\|$.

Eine Algebra mit diesen Eigenschaften heisst C^*-Algebra. Es ist also $L(H)$ eine C^*-Algebra. Tatsaechlich laesst sich zeigen, dass jede C^*-Algebra eine Unteralgebra von $L(H)$ ist (mit einem unter Umstaenden sehr grossen H). Es spielen C^*-Algebren eine wesentliche Rolle in Betrachtungen zur Physik

PROPOSITION. *Sei A ein beschraenkter Operator vom Hilbertraum H in den Hilbertraum K. Dann gilt*

$$Ker(A) = Ran(A^*)^{\perp} \text{ und } Ker(A^*) = Ran(A)^{\perp}.$$

Beweis. Wegen $(A^*)^* = A$ reicht es die erste Gleichung zu zeigen. Dann folgt die zweiten Aussage durch Betrachten von A^*. Direkte Rechnung liefert

$$\begin{aligned} Ker(A) &= \{f : Af = 0\} \\ &= \{f : \langle Af, g\rangle = 0 \text{ fuer alle } g \in H\} \\ &= \{f : \langle f, A^* g\rangle = 0 \text{ fuer alle } g \in H\ \} \\ &= \{A^* g : g \in H\}^{\perp}. \end{aligned}$$

Das beendet den Beweis. $\square$

Notation. Wir werden uns nun (oft) auf die Situation $H = K$ einschraenken. Ein linearer Operator $A : H \longrightarrow H$ wird dann als *linearer Operator im Hilbertraum* H bezeichnet.

Wir betrachten nun die *Beispiele* von oben.
Seien dazu zunaechst $(X, \mathcal{A}, \mu)$ und $(Y, \mathcal{B}, \nu)$ Massraeume und $H = L^2(X, \mu)$ und $K = L^2(Y, \nu)$.

Multiplikationsoperatoren. Sei $\varphi : X \longrightarrow \mathbb{K}$ messbar und beschraenkt und

$$M_\varphi : L^2(X, \mu) \longrightarrow L^2(X, \mu),\ M_\varphi f = \varphi f,$$

der zugehoerige Operator der Multiplikation mit φ. Dann gilt $M_\varphi^* = M_{\overline{\varphi}}$. (Bew. direkte Rechnung.)

Hilbert-Schmidt Operatoren. Sei $k : X \times Y \longrightarrow \mathbb{K}$ messbar mit

$$\|k\|_2^2 := \int_{X\times Y} |k(x, y)|^2 d\mu(x)\nu(y) < \infty$$

und

$$K : L^2(Y, \nu) \longrightarrow L^2(X, \mu), Kf(x) = \int k(x, y)f(y)d\nu(y),$$

der zugehoerige Hilbert-Schmidt Operator mit Kern k. Dann ist K^* ein Hilbert-Schmidt Operator mit Kern $\widetilde{k}$ gegeben durch

$$\widetilde{k} : Y \times X \longrightarrow \mathbb{K},\ \widetilde{k}(y, x) = \overline{k(x, y)}.$$

(Bew. Direkte Rechnung.)

Operatoren auf $\ell^2(X)$. Sei $A : \ell^2(X) \longrightarrow \ell^2(X)$ ein beschraenkter linearer Operator mit Matrix

$$a : X \times X \longrightarrow \mathbb{K}$$

d.h. mit

$$Af(x) = \sum_{y\in Y} a(x, y)f(y)$$

fuer alle $f \in \ell^2(X)$. Dann ist die Matrix des adjungierten Operator gegeben durch die adjungierte Matrix

$$a^* : X \times X \longrightarrow \mathbb{K},\ a^*(x, y) = \overline{a(y, x)}.$$

(Bew. Direkte Rechnung liefert fuer die Matrix b des Adjungierten

$$\begin{aligned} b(x,y) &= \langle e_x, A^* e_y \rangle \\ &= \langle A e_x, e_y \rangle \\ &= \overline{\langle e_y, A e_x \rangle} \\ &= \overline{a(y,x)}. \end{aligned}$$

Das beendet den Beweis.)

4. Isometrien im Hilbertraum

Wesentliche Klassen von linearen Operatoren lassen sich mittels ihrer Adjungierten charakterisieren. So haben wir zum Beispiel schon gesehen, dass orthogonale Projektionen gerade diejenigen linearen Operatoren P mit $P = P^* = P^2$ sind (s.o.). Wir lernen nun einige weitere solche Klassen kennen.

Ein Operator $U : H \longrightarrow H$ heißt *Isometrie*, wenn eine der folgenden drei aequivalenten Eigenschaften gilt:

- $\|Uf\| = \|f\|$ fuer alle $f \in H$.
- $\langle Uf, Ug \rangle = \langle f, g \rangle$ fuer alle $f, g \in H$.
- $U^* U = id$.

(Beweis der Aequivalenz: (1) $\Longleftrightarrow$ (2): Polarisation.
(2) $\Longleftrightarrow$ (3) Einfach nach Definition des Adjungierten.)

Ein Operator $U : H \longrightarrow H$ heißt *unitaer*, wenn eine der folgenden drei aequivalenten Eigenschaften gilt:

- U ist eine surjektive Isometrie.
- $U^* U = id = U U^*$.
- $U^* = U^{-1}$.

(Beweis der Aequivalenz: (2)$\Longleftrightarrow$ (3) Einfach nach Definition des Inversen.
(2)$\Longrightarrow$ (1): Unter Verwendung der Charakterisierung von Isometrien folgt aus (2), dass U eine surjektive Isometrie ist.
(1)$\Longrightarrow$ (2): Da U eine Isometrie ist gilt $U^* U = I$ und U ist injektiv. Ist U noch surjektiv, so ist das Linksinverse auch das Rechtsinverse (und damit das Inverse) und die Aussage (2) folgt.)

Fuer unitaere Abbildungen koennen wir die Lage des Spektrums bestimmen.

Proposition (Spektrum unitaerer Abbildungen ist in S^1 enthalten). *Fuer unitaeres $U : H \longrightarrow H$ gilt $\sigma(U) \subset \{z \in \mathbb{C} : |z| = 1\}$.*

Beweis. Das folgt aus der geometrischen Reihe fuer Operatoren und Ausklammern:
Fuer $\lambda \in \mathbb{C}$ mit $|\lambda| > 1$ ist

$$U - \lambda = -\lambda(I - \frac{1}{\lambda} U)$$

invertierbar (wegen $\|\frac{1}{\lambda} U\| < 1$).
Fuer $\lambda \in \mathbb{C}$ mit $|\lambda| < 1$ ist

$$U - \lambda = U^*(I - \lambda U)$$

invertierbar (wegen $\|\lambda U\| < 1$). $\square$

PROPOSITION (Struktur Kern und Bild von Isometrien). *Sei $U : H \longrightarrow H$ eine Isometrie. Dann gilt:*

- $Ker(U - id) = Ker(U^* - id)$.
- $Ker(U - id)^\perp = \overline{Ran(U - id)}$.

Insbesondere gilt dann also

$$H = Ker(U - id) \oplus \overline{Ran(U - id)}.$$

Beweis. Zum ersten Punkt: Es sind zwei Inklusionen zu zeign.
$\subset$: $(U - id)f = 0$ impliziert $0 = U^*(U - id)f = U^*Uf - U^*f = f - U^*f$.
$\supset$: Sei $(U^* - id)f = 0$. Dann gilt also $f = U^*f$. Damit folgt

$$\begin{aligned} \|(U - id)f\|^2 &= \langle Uf, Uf\rangle - \langle Uf, f\rangle - \langle f, Uf\rangle + \langle f, f\rangle = 0 \\ &= \langle Uf, Uf\rangle - \langle f, U^*f\rangle - \langle U^*f, f\rangle + \langle f, f\rangle \\ &= 0. \end{aligned}$$

(Hier haben wir im letzten Schritt $U^*f = f$ und die Isometrie von U benutzt, welche liefern, dass alle Terme gerade $\|f\|^2$ sind.)
Zum zweiten Punkt: Unter Nutzen des ersten Punktes und der allgemeinen Struktur folgt:

$$\begin{aligned} Ker(U - id)^\perp &= Ker(U^* - id)^\perp \\ (KerA^* = RanA^\perp) &= (Ran(U - id))^{\perp\perp} \\ (Projektionssatz) &= \overline{Ran(U - id)}. \end{aligned}$$

Das beendet den Beweis. $\square$

Die folgende Aussage ist unter dem Namen 'von Neumanscher Ergodensatz' bekannt. Sie wurde unabhaengig voneinander durch Carleman und durch von Neuman um 1931 herum beweisen.

THEOREM (Abstrakter von Neumannscher Ergodensatz - Isometrie im Hilbertraum, von Neumman, Carleman '31). *Sei U eine Isometrie auf dem Hilbertraum H. Sei P die orthogonale Projektion auf $Ker(U - id)$. Dann gilt*

$$\frac{1}{N}\sum_{k=0}^{N-1} U^k\xi \to P\xi, N \to \infty,$$

fuer jedes $\xi \in H$.

Bemerkung. Die Behauptung gilt auch wenn $\frac{1}{N}\sum_{k=0}^{N-1}$ durch $\frac{1}{N}\sum_{k=p}^{N+q}$ mit beliebigen aber festen $p, q \in \mathbb{N}$ ersetzt werden (Aenderung der Summe in endlich vielen Termen).

Beweis. Definiere fuer $N \in \mathbb{N}_0$ den Mittelungsoperator $A_N := \frac{1}{N}\sum_{k=0}^{N-1}$. Dann gilt - wie man leicht sieht - $\|A_N\| \leq 1$ fuer alle $N \in \mathbb{N}_0$. Zu zeigen:

$$A_N\xi \to P\xi, N \to \infty,$$

fuer alle $\xi \in H$. Nach dem Strukurproposition zu Bild und Kern von Isometrien aus dem vorigen Abschnitt gilt

$$H = Ker(U - id) \oplus \overline{Ran(U - id)}.$$

Wir zeigen die entsprechende Konvergenz auf $Ker(U-id)$ und auf $\overline{Ran(U-id)}$.

Es gilt $A_N\xi \to P\xi$, $N \to \infty$, *fuer* $\xi \in Ker(U-id)$*:* Sei $\xi \in Ker(U-id)$. Dann gilt also $U\xi = \xi$ und damit $U^k\xi = \xi$ fuer alle $k \in \mathbb{N}_0$. Damit folgt die gewuenschte Aussage sofort.

Es gilt $A_N\xi \to P\xi$, $N \to \infty$, *fuer* $\xi \in \overline{Ran(U-id)}$*:* Aufgrund der Zerlegung des Hilbertraumes gilt

$$P\xi = 0$$

fuer alle $\xi \in \overline{Ran(U-id)}$. Wegen $\|A_N\| \leq 1$ fuer alle $N \in \mathbb{N}_0$ reicht es $\xi \in Ran(U-id)$ zu betrachten. Sei also

$$\xi = (U-id)\eta.$$

Dann wird $A_N\xi$ eine Teleskopsumme und die gewuenschte Aussage folgt einfach:

$$\begin{aligned} A_N\xi &= A_N(U-id)\eta \\ &= A_N U\eta - A_N\eta \\ &= \frac{1}{N}\sum_{k=1}^{N} U^k\eta - \frac{1}{N}\sum_{k=0}^{N-1} U^k\eta \\ &= \frac{1}{N}(U^N\eta - \eta) \\ &\to 0, N \to \infty. \end{aligned}$$

Das ist die gewuenschte Aussage. □

Anwendung des Ergodensatzes. Sei $(X, \mathcal{A}, \mu)$ ein Massraum und $\mu(X) = 1$. Sei $T : X \longrightarrow X$ *maßerhaltend* d.h. es ist T messbar mit

$$\mu(T^{-1}A) = \mu(A)$$

fuer alle $A \in \mathcal{A}$. Dann ist

$$U_T : L^2(X,\mu) \longrightarrow L^2(X,\mu), U_T f = f \circ T,$$

eine Isometrie. Ist $\mathcal{I}$ der Unterraum der T-invarianten Funktionen aus L^2 (d.h. der Funktionen f mit $f = f \circ T$ in L^2) und P die Projektion auf $\mathcal{I}$, so gilt

$$\frac{1}{N}\sum_{n=0}^{N-1} f \circ T^k \to Pf, N \to \infty.$$

Ist T *ergodisch* (d.h. gilt $\mathcal{I} = Lin\{1\}$), so gilt $Pf = \langle 1, f\rangle 1$ (da 1 eine Orthonormalbasis von $\mathcal{I}$ ist). Dann vereinfacht sich die Aussage zu

$$\frac{1}{N}\sum_{n=0}^{N-1} f \circ T^k \to \langle 1, f\rangle 1, N \to \infty.$$

Die linke Seite kann als eine Art 'Zeitmittel' gesehen werden die rechte als ein 'Raummittel' ueber f. Die Untersuchung solcher Raum- und Zeitmittel ist ein wesentliches Thema der Ergodentheorie.

Bew. Es ist zu zeigen, dass U_T eine Isometrie ist, dann folgen die uebrigen Aussagen aus dem vorangehenden Ergodensatz. Betrachte

$$\mathcal{F} := \{f : X \longrightarrow [0,\infty] : \int f d\mu = \inf f \circ T d\mu\}.$$

Dann enthaelt $\mathcal{F}$ alle charakteristischen Funktionen von Mengen (da T masserhaltend ist). Des weiteren ist $\mathcal{F}$ offenbar abgeschlossen unter endlichen Linearkombinationen mit positiven Koeffizienten und unter monotoner Konvergenz. Damit sieht man leicht

$$\mathcal{F} = \{f : X \longrightarrow [0,\infty] : f \text{ messbar}\}.$$

Dann folgt die gewuenschte Isometrie von U_T sofort

$$\langle Uf, Uf\rangle = \int_X |f \circ T|^2 d\mu = \int_X |f|^2 \circ T d\mu = \int_X |f|^2 d\mu = \langle f, f\rangle.$$

5. Selbstadjungierte Operatoren

In diesem Abschnitt geht es um eine besonders wichtige Klasse von Operatoren. Es sind dies die Operatoren, die in der mathematischen Beschreibung der Quantenmechanik gerade fuer physikalische Observablen stehen.

Ein Operator A im Hilbertraum H heisst *selbstadjungiert*, wenn $A = A^*$ gilt.

Es gibt viele selbstadjungierte Operatoren. Tatsaechlich sind fuer jeden Operator B im Hilbertraum H die beiden Operatoren

$$B + B^* \text{ und } i(B - B^*)$$

selbstadjungiert. Insbesondere laesst sich im komplexen Hilbertraum jeder beschraenkte Operator B schreiben als

$$B = \Re B + i\Im B$$

mit den selbstadjungierten Operatoren

$$\Re B = \frac{1}{2}(B + B^*), \ \Im B := \frac{1}{2i}(B - B^*).$$

Das entspricht gerade der Zerlegung einer komplexen Zahl in Real- und Imaginaerteil.

Bemerkung. Es ist fuer jeden Operator B auch BB^* selbstadjungiert. Damit laesst sich auch zu jedem B ein selbstadjungierter Operator $|B| = (B^*B)^{1/2}$ und ein isometrischer Operator U definieren mit

$$B = U|B|.$$

Das entspricht der Polarzerlegung einer komplexen Zahl. Wir geben hier nicht die Details zur Definition von $(B^*B)^{1/2}$.

Fuer selbadjungierte Operatoren laesst sich die Norm ueber die Form ausrechnen.

Proposition. *Sei A ein beschraenkter selbstadjungierter Operator im Hilbertraum H. Dann gilt*

$$\|A\| = \sup\{|\langle x, Ax\rangle| : \|x\| \leq 1\}.$$

Bemerkung. Offenbar gilt fuer jedes beliebige A jedenfalls

$$\|A\| = \sup\{|\langle x, Ay\rangle| : \|x\|, \|y\| \leq 1\}.$$

Beweis. Wir bezeichnen das Supremum mit S. Die Ungleichung $\geq$ folgt einfach mit Cauchy-Schwarz Ungleichung und der Definition von $\|\cdot\|$. Wir zeigen nun $\leq$. Gemäß der vorangehenden Bemerkung ist die Aufgabe 'aus $\langle x, Ax\rangle$ den Term '$\langle x, Ay\rangle$ zu machen'. Seien x, y mit $\|x\|, \|y\| \leq 1$ gegeben. Wir rechnen:

$$\begin{array}{rcl} S & \geq & \frac{S}{2}(\|x\|^2 + \|y\|^2) \\ \text{(Parallelogrammidentitaet)} & = & \frac{S}{4}(\|x+y\|^2 + \|x-y\|^2) \\ \text{(Definition } S\text{, }\Delta\text{-Ugl)} & \geq & \frac{1}{4}|\langle x+y, A(x+y)\rangle - \langle x-y, A(x-y)\rangle| \\ \text{(Ausrechnen)} & = & \frac{1}{2}|\langle x, Ay\rangle + \langle y, Ax\rangle| \\ \text{(A selbstadjungiert)} & = & |\Re\langle x, Ay\rangle|. \end{array}$$

Nach Ersetzen von x durch αx mit $|\alpha| = 1$ und

$$|\Re\langle \alpha X, Ay\rangle| = |\langle x, Ay\rangle|$$

folgt nun

$$|\langle x, Ay\rangle| \leq |\Re\langle \alpha X, Ay\rangle| \leq S.$$

Da dies fuer alle x, y mit $\|x\|, \|y\| \leq 1$ gilt, folgt die gewuenschte Aussage nun sofort. $\square$

Proposition. *Ist A ein selbstadjungierter Operator im komplexen Hilbertraum H, so gilt*

$$\|(A - zI)x\|^2 = \|(A - Rez)x\|^2 + |\Im z|^2 \|x\|^2$$

fuer alle $z \in \mathbb{C}$ und $x \in H$. Insbesondere gilt

$$\|(A - zI)x\| \geq |\Im z| \|x\|$$

fuer alle $z \in \mathbb{C}$ und $x \in H$.

Beweis. Die erste Gleichung folgt durch direkte Rechnung unter Nutzen der Selbstadjungiertheit. Aus dieser Gleichung folgt sofort die Ungleichung

$$\|(A - z)x\| \geq \|\Im z| \|x\|$$

fuer alle $z \in \mathbb{C}$ und $x \in H$. $\square$

Bemerkung. Die zweite Ungleichung kann man auch direkt herleiten aus

$$\|(A - z)x\| \|x\| \geq |\langle (A - z)x, x\rangle| = |\langle (A - \Re z)x, x\rangle + i\Im z\langle x, x\rangle| \geq |\Im z| \|x\|^2.$$

Theorem. *Sei A ein selbstadjungierter Operator im komplexen Hilbertraum H. Dann gilt $\sigma(A) \subset \mathbb{R}$. Genauer ist $A - zI$ fuer jedes $z \in \mathbb{C} \setminus \mathbb{R}$ bijektiv und die Inverse $(A - zI)^{-1}$ ist beschraenkt durch $\frac{1}{|\Im z|}$.*

Bemerkung. Fuer die Quantenmechanik ist es fundamental, dass das Spektrum eines selbstadjungierten Operator reell ist, denn das erlaubt es das Spektrum als die moeglichen Ergebnisse einer Messung zu interpretieren.

Beweis. Wir nutzen die Ungleichung

$$(*) \quad \|(A-z)x\| \geq |\Im z| \|x\|$$

fuer alle $z \in \mathbb{C}$ und $x \in H$. Diese Ungleichung ist fundamental fuer die folgenden Betrachtungen und wird in jedem einzelnen Schritt genutzt.

$A - zI$ ist injektiv. Das folgt sofort aus $(*)$.

$A - zI$ ist surjektiv. Wir zeigen, dass $A - zI$ dichtes Bild hat und dieses Bild abgeschlossen ist. (Dann folgt Surjektivitaet.)
Bild von $(A - zI)$ ist abgeschlossen: Sei $y_n = (A - zI)x_n$ eine Folge im Bild, die gegen y konvergiert. Dann ist (y_n) eine Cauchy-Folge. Damit ist nach $(*)$ auch (x_n) eine Cauchy-Folge. Sei x ihr Grenwert. Dann gilt

$$(A - zI)x = \lim(A - zI)x_n = \lim y_n = y.$$

Bild von $A - zI$ ist dicht: Aus $(*)$ (mit $\overline{z}$) folgt

$$Ran(A - z)^{\perp} = Ker(A^* - \overline{z}) = Ker(A - \overline{z}) = \{0\}.$$

Damit ist dann also $Ran(A - zI)$ dicht und abgeschlossen, also stimmt es mit H ueberein.

Zur Schranke an die Norm. Die Schranke an die Inverse folgt aus $(*)$. Denn es gilt fuer $x = (A - zI)^{-1}y$

$$\|y\| = \|(A - zI)x\| \geq |Imz| \|x\|$$

und damit

$$\|x\| \leq \frac{1}{|\Im z|} \|y\|.$$

Das beendet den Beweis. □

Im endlichdimensionalen Fall wissen wir, dass das Spektrum aus Eigenwerten (d.h. $\lambda \in \mathbb{R}$ mit $(A - \lambda)x = 0$ fuer ein geeignetes $x \neq 0$) besteht. Im unendlichdimensionalen Hilbertraum ist das falsch (s.o.). Mittels einer Variante des Beweises der vorigen Proposition laesst sich allerdings das Spektrum selbstadjungierter Operatoren durch 'Fasteigenfunktionen' charakterisieren.

PROPOSITION. *Sei A ein selbstadjungierter Operator im Hilbertraum.*
(a) Es gehoert $\lambda \in \mathbb{R}$ genau dann zum Spektrum $\sigma(A)$, wenn fuer alle $\varepsilon > 0$ ein $x_\varepsilon \in H$ existiert mit $\|x_\varepsilon\| = 1$ und

$$\|(A - \lambda I)x_\varepsilon)\| \leq \varepsilon.$$

(b) Es gehoert $\lambda \in \mathbb{R}$ genau dann zur Resolventenmenge, wenn ein $c > 0$ existiert mit

$$\|(A - \lambda I)x\| \geq c\|x\|$$

fuer alle $x \in H$.

Beweis. Offenbar sind die Aussagen (a) und (b) aequivalent. Wir zeigen nur (b): Gehoert λ zur Resolventemenge und ist B die beschraenkte Inverse von $(A - \lambda I)$, so gilt

$$\|x\| = \|B(A - \lambda I)x\| \leq \|B\| \|(A - \lambda I)x\|$$

fuer alle $x \in H$. Gilt umgekehrt die Ungleichung $\|(A - \lambda I)x\| \geq c\|x\|$ fuer alle $x \in H$, so ist $(A - \lambda I)$ also injektiv. Weiterhin sieht man wie im Beweis der vorigen Proposition, dass $(A - \lambda I)$ dichtes Bild hat. Weiterhin folgt

$$Ran(A - \lambda I)^{\perp} = Ker(A - \lambda I) = \{0\}.$$

Damit ist also $(A - \lambda I)$ surjektiv. Schliesslich folgt aus der Ungleichung noch die Beschraenktheit des inversen Operator. □

Bemerkung. Ein Operator T im Hilbertraum H heisst *normal*, wenn eine der folgenden aequivalenten Aussagen gilt:

- Es gilt $TT^* = T^*T$.
- Es gilt $\|Tx\| = \|T^*x\|$ fuer alle $x \in H$.

(Beweis der Aequivalenz kann leicht mittels Polarisation gefuehrt werden.) Beispiele normaler Operatoren sind offenbar unitaere Operatoren und selbstadjungierte Operatoren. Eine Theorie normaler Operatoren umfasst also beide Klassen von Operatoren. Tatsaechlich sind selbstadjungierte Operatoren gerade diejenigen normalen Operatoren, deren Spektrum in $\mathbb{R}$ enthalten ist und unitaere Operatoren sind diejenigen normalen Operatoren, deren Spektrum in der Einheitskreislinie enthalten ist. Entsprechendes koennen wir aber in dieser Vorlesung nicht beweisen.

Beispiel : Multiplikationsoperatoren mit reellen Funktionen. Sei $(X, \mathcal{A}, \mu)$ ein Massraum mit σ-endlichem Mass μ. Sei $\psi : X \longrightarrow \mathbb{R}$ messbar und beschraenkt. Dann definert man den *Operator der Multiplikation mit f auf* $L^2(X, \mu)$ durch

$$M_\psi : L^2(X, \mu) \longrightarrow L^2(X, \mu), M_\psi f = \psi f.$$

Aus dem im Laufe der bisherigen Abschnitte bewiesenen erhalten wir:

- $\|M_\psi\| = \|\psi\|_\infty$.
- $M_\psi^* = M_{\overline{\psi}} = M_\psi$.
- $\sigma(M_\psi) = \{\lambda \in \mathbb{R} : \mu(\psi^{-1}(U_\varepsilon(\lambda))) > 0 \text{fuer alle } \varepsilon > 0\}$ (Wesentlicher Wertebereich von ψ).

Bemerkung. Es laesst sich zeigen, dass jeder normale Operator unitaer equivalent zu einem Multiplikatonsoperator ist. Das ist eine fundamentale Eigenschaft normaler Operatoren. Die entsprechende Aussage heisst *Spektralsatz.* Dieser Satz verallgemeinert die Diagonalisierbarkeit von normalen Matrizen im Endlichdimensionalen.

6. Kompakte Operatoren

In diesem Abschnitt lernen wir noch die wichtige Klasse der kompakten linearen Operatoren im Hilbertraum kennen. Diese verhalten sich besonders ‘aehnlich’ den Operatoren in endlichdimensionalen Raeumen. Dabei wird - etwas lax gesprochen - Endlichdimensionalitaet des Raumes wird durch die Forderung einer ‘Fast-Eindlichdimensionalitaet’ des Operators ersetzt. Es gilt

$$\text{Hilbert Schmidt Operatoren} \subset \text{ kompakte Operatoren} \subset \text{ beschraenkte Operatoren},$$

wobei die Inklusionen im unendlichdimensoinalen Hilbertraum echt sind (waehrend im endlichdimensionalen Fall alle drei Klassen zusammenfallen).

Exkurs. Sei (X, d) ein metrischer Raum. Eine Teilmenge S von X heisst *total beschraenkt*, wenn zu jedem $\varepsilon > 0$ Elemente $x_1, \ldots, x_n \in X$ existieren mit

$$S \subset \cup_{j=1}^{n} U_\varepsilon(x_j),$$

(wobei $U_s(x)$ die offene s-Kugel um x bezeichnet. Offenbar ist S genau dann total beschraenkt, wenn $\overline{S}$ total beschraenkt ist.
Es gilt:

- Es ist S total beschraenkt genau dann, wenn jede Folge in S eine Teilfolge enthaelt, die eine Cauchy-Folge ist.
- Es ist S total beschraenkt und vollstaendig genau dann, wenn jede Folge in S eine in S konvergente Teilfolge enthaelt.

Eine Teilmenge S heisst *kompakt*, wenn sie total beschraenkt und vollstaendig ist.
Ist (X, d) vollstaendig, so ist S offenbar vollstaendig genau dann, wenn S abgeschlossen ist. Damit sind die kompakten Mengen in diesem Fall also die abgeschlossenen total beschraenkten Teilmengen von X.
Ein $S \subset X$ heisst *relativ kompakt*, wenn $\overline{S}$ kompakt ist. Ist (X, d) vollstaendig, so ist also S relativ kompakt genau dann, wenn es total beschraenkt ist.
Das beendet den Exkurs.

DEFINITION. *Ein linearer Operator A vom Hilbertraum H in den Hilbertraum K heisst kompakt, wenn das Bild der Einheitskugel von H unter A relativ kompakt in K ist.*

Meist wird in Anwendungen folgende Charakterisierung von Kompaktheit eines Operator genutzt.

FOLGERUNG. *Ein Operator A von Hilbertraum H in den Hilbertraum K ist genau dann kompakt, wenn fuer jede beschraenkte Folge (x_n) in H die Folge (Ax_n) eine konvergente Teilfolge enthaelt.*

Beweis. Das ist klar (nach den Definitionen und der Diskussion im Exkurs). $\square$

Bemerkung. Wie schon erwaehnt haben kompakte Operatoren starke Aehnlichkeiten mit Operatoren in endlichdimensionalen Raeumen. Eine Anhaltspunkt dafuer gibt auch folgende Beobachtung: Enthaelt das Bild der Einheitskugel unter einem kompakten A eine Orthogonalsystem e_j, so muss notwendig die Indexmengen abzaehlbar sein und gelten $\|e_j\| \to 0$ (da es zu jedem $c > 0$ hoechstens endlich viele j mit $|e_j| \geq c > 0$ geben kann).

PROPOSITION. *Jeder kompakte Operator ist beschraenkt.*

Beweis. Sei A ein kompakter Operator. Dann ist nach Definition also $\{Ax : \|x\| \leq 1\}$ relativ kompakt und damit ist die stetige Funktion $\|\cdot\|$ darauf beschraenkt. Damit folgt sofort die Beschraenktheit von A. $\square$

Bemerkung. Die Umkehrung gilt im unendilchdimensionalen Fall nicht. Betrachte etwa die Identitaet I auf dem unendlichdimensionalen H und sei

e_n, $n \in \mathbb{N}$, eine Orthonormalsystems. Dann hat Ie_n offenbar keine konvergente Teilfolge.

Die folgenden Propositionen zeigen, dass Kompaktheit mit den 'ueblichen' Operationen gut verträglich ist.

PROPOSITION. *Seien H, K, L Hilbertraeume.*
(a) Sind $A : H \longrightarrow K$ und $B : H \longrightarrow K$ kompakt, so ist auch $A + \lambda B$ kompakt.
(b) Sei $A : H \longrightarrow K$ kompakt und $B : K \longrightarrow L$ und $C : L \longrightarrow H$ beschraenkt. Dann sind auch BA und AC kompakt.

Beweis. (a) Das ist einfach (da die Summe zweier kompakter Mengen wieder kompakt ist).
(b) Zu BA: Es ist das Bild kompakter Mengen unter dem stetigen B kompakt.
Zu AC: Es ist das Bild der Einheitskugel unter dem stetigen C beschraenkt und damit in einer skalierten Version der Einheitskugel enthalten. □

PROPOSITION. *Seien H und K Hilbertraeume und $A : H \longrightarrow K$ linear und beschraenkt. Dann sind aequivaelent:*
(i) *A ist kompakt.*
(ii) *A^*A ist kompakt.*
(iii) *AA^* ist kompakt.*
(iv) *A^* ist kompakt.*

Beweis. (i)$\Longrightarrow$ (ii): Das folgt aus dem weiter oben gezeigten.
(ii)$\Longrightarrow$ (i): Sei (x_n) eine beschraenkte Folge. Sei ohne Einschraenkung A^*Ax_n konvergent. Dann gilt

$$\|A(x_n - x_m)\|^2 = \langle A(x_n - x_m), A(x_n - x_m)\rangle = \langle A^*A(x_n - x_m), x_n - x_m\rangle \to 0.$$

(i)$\Longrightarrow$ (iii): Das folgt aus der vorigen Proposition.
(iii)$\Longrightarrow$ (iv): Das folgt wie (ii)$\Longrightarrow$ (i) nach Ersetzen von A durch A^*.
(iv)$\Longrightarrow$ (ii): Das folgt aus der vorigen Proposition. □

PROPOSITION. *Seien H, K Hilbertraeume und $A_n : H \longrightarrow K$ kompakte Operatoren mit $A_n \to A$ bzgl. $\|\cdot\|$. Dann ist auch A kompakt.*

Beweis. Wir zeigen Totalbeschraenktheit von AB_1, wobei B_1 die Einheiskugel in H ist. Sei $\varepsilon > 0$ gegeben. Waehle $N \in \mathbb{N}$ mit $\|A_N - A\| < \frac{\varepsilon}{2}$. Da A_N kompakt ist, gibt es dann $x_1, \ldots, x_k \in K$ mit

$$A_N B_1 \subset \bigcup_{j=1}^{k} U_{\varepsilon/2}(x_j).$$

Wegen $\|A - A_N\| < \varepsilon/2$ gilt dann

$$AB_1 \subset \bigcup_{j=1}^{k} U_{\varepsilon}(x_j)$$

und der Beweis ist beendet. □

Bemerkung. Die kompakten Operatoren im Hilbertraum H bilden nach den vorangehenden Ueberlegungen eine vollstaendige normierte Algebra mit Involution. Weiterhin gilt $\|AA^*\| = \|A^*A\|$. Es handelt sich also um eine C^*-Algebra.

Wir diskutieren nun einige *Beispiele.*

Operatoren mit endlichdimensionalem Bild. Offenbar sind Operatoren mit endlichdimensionalem Bild kompakt. Solche Operatoren werden auch oft als endlichdimensionale Operatoren bezeichnet. Nach unseren obigen Betrachtungen ist dann auch jeder Normgrenzwert endlichdimensionaler Operatoren kompakt.

Bemerkung. Es gilt auch die Umkehrung (Uebung): Ein Operator ist genau dann kompakt, wenn er ein Normgrenzwert endlichdimensionaler Operatoren ist.

Hilbert-Schmidt Operatoren. Seien H, K Hilbertraeume und (e_j) eine Orthonormalbasis von H und Sei $A : H \longrightarrow K$ beschraenkt mit

$$\sum_j \|Ae_j\|^2 < \infty.$$

Dann ist A kompakt.
(Bew. Ohne Einschraenkung sei die Indexmenge abzaehlbar und durch $\mathbb{N}$ gegeben. (Andernfalls koennen wir uns auf die (notwendig) abzaehlbare Menge der j mit $Ae_j \neq 0$ beschraenken.) Sei P_N die orthogonale Projektion auf $Lin\{e_1, \ldots, e_N\}$. Dann ist $A_N := AP_N$ kompakt als endlichdimensionaler Operator. Weiterhin gilt fuer jedes $x = \sum c_j e_j$ in H.

$$\begin{aligned} \|(A - A_N)x\|^2 &= \|\sum_{j>N} c_j Ae_j\|^2 \\ \text{(Dreiecksungl.)} \quad &\leq \left(\sum_{j>N} |c_j|\|Ae_j\|\right)^2 \\ \text{(CS)} \quad &\leq \|x\|^2(\sum_{j>N} \|Ae_j\|^2)^2 \end{aligned}$$

Mit

$$\sum_{j>N} \|Ae_j\|^2 \to 0, N \to \infty,$$

folgt die gewuenschte Normkonvergenz.)

Grenzwerte von Hilbert-Schmidt Operatoren. Ist A ein Normgrenzwert von Hilbert-Schmidt Operatoren, so ist A aufgrund der bisherigen Ueberlegungen kompakt. - Das ist eine der Hauptmethoden, um Kompaktheit von konkreten Operatoren zu zeigen.

Wir wenden uns nun dem Spektralsatz fuer kompakte selbstadjungierte Operatoren zu. Aus der linearen Algebra wissen folgendes: Ist H ein endlichdimensionaler Hilbertraum und $A : H \longrightarrow H$ selbstadjungiert, so existiert eine Orthonormalbasis $e_1, \ldots, e_N$ von H aus Eigenvektoren zu A und mit den zugehoerigen Eigenwerten λ_j gilt dann

$$Ax = \sum_{j=1}^{N} \lambda_j \langle e_j, x\rangle e_j.$$

Im unendlichdimensionalen Hilbertraum gilt ein entsprechendes Ergebnis, wenn man zusaetzlich die Kompaktheit von A fordert. Das laesst sich wieder so verstehen, dass die kompakten Operatoren nahe an den endlichdimensionalen Operatoren sind.
Im Beweis ist im wesentlichen die Existenz von 'vielen' Eigenwerten zu zeigen. Das wird induktiv geschehen basierend auf der folgenden Proposition, die die Existenz *eines* Eigenwertes zeigt.

PROPOSITION. *Sei A ein kompakter selbstadjungierter Operator im Hilbertraum H. Dann ist $\|A\|$ oder $-\|A\|$ ein Eigenwert von A.*

Beweis. Der Fall $\|A\| = 0$ ist klar. Sei nun $\|A\| \neq 0$. Wir wissen schon $\|A\| = \sup\{|\langle x, Ax\rangle| : \|x\| \leq 1\}$. Damit existiert also eine Folge (x_n) mit $\|x_n\| = 1$ und

$$\langle x_n, Ax_n\rangle \to \lambda$$

mit $|\lambda| = \|A\|$. Damit folgt also

$$0 \leq \|Ax_n - \lambda x_n\|^2 = \|Ax_n\|^2 - 2\lambda\langle x_n, Ax_n\rangle + \lambda^2 \leq 2\lambda^2 - 2\lambda\langle x_n, Ax_n\rangle \to 0.$$

Das liefert

$$\|Ax_n - \lambda x_n\| \to 0.$$

Aufgrund der Kompaktheit von A koennen wir ohne Einschraekung annehmen, dass Ax_n gegen y konvergiert. Dann gilt aber auch $\lambda x_n \to y$ und damit $x_n \to \frac{1}{\lambda} y =: x$. Insgesamt folgt

$$Ax = \lim_N Ax_n = y = \lambda x.$$

Das beendet den Beweis. $\square$

THEOREM (Spektralsatz fuer kompakte selbstadjungierte Operatoren). *Sei $A \neq 0$ ein kompakter selbstadjungierter Operator im Hilbertraum H. Dann existiert eine endliche oder abzaehlbare Menge N und ein Orthonormalsystem $(e_j)_{j\in N}$ von Eigenvektoren von A mit zugehoerigen Eigenwerten $\lambda_j \in \mathbb{R} \setminus \{0\}$ mit monoton fallenden $|\lambda_j|$ und, falls N unendlich ist, $\lambda_j \to 0$, so dass fuer alle $x \in H$ gilt*

$$Ax = \sum_{j\in N} \lambda_j \langle e_j, x\rangle e_j.$$

Beweis. **Plan:** Wir konstruieren induktiv unter Nutzen der vorigen Proposition ein Orthonormalsystem $e_1, \ldots, ..$ aus Eigenvektoren (e_j) zu Eigenwerten $\lambda_1, \ldots$, so dass mit den abgeschlossenen Unterraeumen

$$H_1 := H, \;\; H_{n+1} := Lin\{e_1, \ldots, e_n\}^\perp, \; n \geq 1,$$

gilt

$$|\lambda_n| = \|A|_{H_n}\|$$

fuer alle $n \in N$. Dann folgen die uebrigen Aussagen automatisch (siehe unten).

Sei $n = 1$. Es gilt $H_1 = H$. Die vorige Proposition liefert ein e_1 mit $Ae_1 = \lambda_1 e_1$ mit $|\lambda_1| = \|A|_{H_1}\|$.

Der Schluss von n auf $n + 1$. Seien $e_1, \ldots, e_n$ und zughoerige $\lambda_1, ..., \lambda_n$ und $H_1, \ldots, H_n$ schon konstruiert. Sei H_{n+1} das orthogonale Komplement der lineare Huelle von $e_1, \ldots, e_n$. Dann gilt $A(H_{n+1}) \subset H_{n+1}$. (Denn fuer $y = Ax$ mit $x \in H_{n+1}$ sieht man sofort

$$\langle y, e_j \rangle = \langle Ax, e_j \rangle = \langle x, Ae_j \rangle = 0.)$$

Die Einschraenkung von A auf H_{n+1} ist also ein Operator in H_{n+1}. Da A kompakt ist, ist auch die Einschraenkung kompakt. Offenbar ist sie selbstadjungiert. Anwenden der vorigen Proposition liefert dann die gewuenschte Ausage fuer $n + 1$.

Gilt $A|_{H_n} = 0$ fuer ein n, so bricht die Konstruktion ab und N ist endlich. Andernfalls erhaelt man eine abzaehlbare Menge N. In diesem Fall gilt

$$|\lambda_j| = \|\lambda_j e_j\| = \|Ae_j\| \to 0, j \to \infty.$$

(Zur letzten Konvergenz: Es gilt:

- Aufgrund der Kompaktheit hat (Ae_j) und jede seiner Teilfolgen eine konvergente Teilfolge.
- Jede konvergente Teilfolge von (Ae_j) konvergiert gegen 0. (Denn $Ae_j \to y$ impliziert

$$\langle y, z \rangle = \lim \langle Ae_j, z \rangle = \lim_n \langle e_j, Az \rangle = 0.)$$

Damit konvergiert dann (Ae_j) gegen 0.)

Es bleibt die Aussage zur Darstellung von Ax zu beweisen:
Sei

$$K := \bigcap_n H_n.$$

Dann gilt offenbar $K = \mathrm{Lin}\{e_j : j = 1, \ldots, \}^\perp$ und damit

$$K^\perp = \overline{\mathrm{Lin}\{e_j : j \in N\}}.$$

Wir untersuchen nun zunaechst A auf K und auf $K^\perp$ (und setzen das dann zusammen).
Es gilt $Ax = 0$ fuer alle $x \in K$ (da $\|Ax\| \leq |\lambda_n| \|x\|$ fuer $x \in H_n$).
Fuer $x \in K^\perp$ gilt

$$x = \sum_j \langle e_j, x \rangle e_j$$

und damit dann

$$Ax = \sum_j \lambda_j \langle e_j, x \rangle e_j.$$

Fuer ein allgemeines $x = x^* + n$ mit $x^* \in K^\perp$ und $n \in N$ gilt dann

$$Ax = Ax^* = \sum_j \lambda_j \langle e_j, x^* \rangle e_j = \sum_j \lambda_j \langle e_j, x \rangle e_j.$$

Das beendet den Beweis. □

Bemerkung - Umformulierung im separablen Hilbertraum. Sei H ein separabler Hilbertraum und A ein selbstadjungierter kompakter Operator in H. Nach dem vorigen Satz gibt es ein ONS (e_j), $j \in N$, von Eigenvektoren mit nichtverschwindenden Eigenwerten und auf $K := \mathrm{Lin}\{e_j : j \in N\}^{\perp}$ verschwindet A. Damit sind die Elemente von K also alle Eigenvektoren von A zum Eigenwert 0. Ergaenzt man das ONS e_j durch Wahl einer ONB von K zu einer ONB von ganz H und nennt das entstehende ONB $(\widetilde{e}_j)$ mit zugehoerigen Eigenwerten $\widetilde{\lambda}_j$, so hat man eine ONB aus Eigenvektoren gefunden und es gilt $\widetilde{\lambda}_j \to 0, j \to \infty$. Mit der eindeutigen unitaeren Abbildung $U : H \longrightarrow \ell^2(\mathbb{N})$ mit $U\widetilde{e}_j = 1_j$ gilt dann also

$$UAU^{-1} = M_{\varphi}$$

fuer $\varphi : \mathbb{N} \longrightarrow \mathbb{R}, \varphi(j) = \widetilde{\lambda}_j$ und φ hat die Eigenschaft

$$\varphi(j) \to 0, j \to \infty.$$

Damit kann man sich also kompakte selbstadjungierte Operatoren vorstellen, als Multiplikationsoperatoren mit Funktionen, die gegen Null konvergieren.

Folgerung. *(Uebung) Das Spektrum eines kompakten selbstadjungierten Operator in einem unendlichdimensionalen Hilbertraum besteht aus nichtverschwindenden Eigenwerten sowie der Null.*

KAPITEL 5

Etwas zu normierten Räumen

In diesem Abschnitt lernen wir sehr allgemeine Vektorraume kennen. Das gibt manchen Betrachtungen, die wir vorher schon durchgefuehrt haben, eine systematische Grundlage. Auf der Ebene der Beispiele kann man es so sehen, dass in den vorigen beiden Kapiteln $L^2(X,\mu)$ betrachtet wurde, waehrend es nun um $L^p(X,\mu)$ fuer allgemeine $p \in [1,\infty)$ (und noch viele weitere Raeume) geht.

1. Normen

Wir beginnen, in dem wir das zentrale Konzept zur 'Laengenmessung' in einem Vektorraum noch einmal vorstellen.

Definition. *Sei E ein Vektorraum ueber $\mathbb{K}$. Eine Abbildung*

$$\|\cdot\| : E \longrightarrow [0,\infty)$$

heisst Norm auf E, wenn fuer alle $x,y \in E$ und $\lambda \in \mathbb{K}$ gilt

(N1) $\|x\| = 0$ *genau dann wenn* $x = 0$,
(N2) $\|\lambda x\| = |\lambda| \|x\|$,
(N3) $\|x+y\| \leq \|x\| + \|y\|$.

Dann heisst $(E, \|\cdot\|)$ ein normierter Raum. Eine Abbildung, die $(N2)$ und $(N3)$ erfuellt, heisst Halbnorm.

Beispiele fuer normierte Raeume:

- Hilbertraeume sind normierte Raeume (mit der durch das Skalarprodukt induzierten Norm).
- Ist $(X,\mathcal{A},\mu)$ ein Massraum und $p \in [1,\infty]$ beliebigt, so ist $L^p(X,\mathcal{A},\mu)$ mit $\|\cdot\|_p$ eine normiertrer Raum.
- Ist (X,d) ein metrischer Raum, so definiert man fuer eine beschraenkte Funktion $f : X \longrightarrow \mathbb{C}$ dann

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

  Damit wird dann der Vektorraum $C_b(X)$ der beschaenkten stetigen Funktionen auf X zu einem normierten Raum. Sei $C_0(X)$ der Vektorraum der stetigen $f : X \longrightarrow \mathbb{C}$, fuer die fuer jedes $\varepsilon > 0$ ein kompakte $K_\varepsilon \subset X$ existiert mit

  $$|f(x)| \leq \varepsilon$$

  fuer $x \notin K_\varepsilon$. Dann ist $(C_0(X), \|\cdot\|_\infty)$ ein normierter Raum. Sonderfaelle sind (Uebung) der Raum

  $$C_0(\mathbb{R}) = \{f : \mathbb{R} \longrightarrow \mathbb{C} : f \text{ stetig mit } \lim_{x\to\pm\infty} f(x) = 0\}$$

und

$$C_0(\mathbb{N}) = \{f : \mathbb{N} \longrightarrow \mathbb{C} : \lim_{n\to\infty} f(n) = 0.\}.$$

PROPOSITION. *Sei E ein Vektorraum mit Norm $\|\cdot\|$. Dann induziert $\|\cdot\|$ eine Metrik auf E mittels*

$$d = d_{\|\cdot\|} : E \times E \longrightarrow [0,\infty),\ d(x,y) = \|x-y\|.$$

Die Abbildungen

$$A : E \times E \longrightarrow E, (x,y) \mapsto x+y$$

und

$$M : \mathbb{K} \times E \longrightarrow E, (\lambda, x) \mapsto \lambda x)$$

sind stetig.

Beweis. d ist Metrik. Einfach.

Stetigkeit der Addition A und der Multiplikation M.
Zu A: Es gilt

$$\begin{aligned} d(A(x,y), A(x_0,y_0)) &= \|(x+y) - (x_0+y_0)\| \\ &\leq \|x - x_0| + \|y - y_0\| \\ &= d(x,x_0) + d(y,y_0). \end{aligned}$$

Das liefert leicht die Stetigkeit von A.

Zu M: Es gilt

$$\begin{aligned} d(M(\lambda,x), M(\lambda_0,x_0)) &= \|\lambda x - \lambda_0 x_0\| \\ &= \|\lambda x - \lambda x_0 + \lambda x_0 - \lambda_0 x_0\| \\ &\leq |\lambda|\|x - x_0\| + |\lambda - \lambda_0|\|x_0\| \\ &\leq (|\lambda_0| + |\lambda - \lambda_0|) d(x,x_0) + |\lambda - \lambda_0|\|x_0\|. \end{aligned}$$

Das liefert leicht die Stetigkeit von M. □

Bemerkung.

- Metriken erzeugen Topologien. Topologien auf einem Vektorraum, fuer die A und M stetig sind, heissen auch *Vektorraumtopologien.*
- Die Nichtausgeartetheit der Norm spielt bei den Definitionen (eigentlich) keine Rolle. Die enstprechenden Aussagen gelten auch fuer Halbnormen.
- Die Nichtausgeartetheit ist wesentlich fuer die Hausdorffeigenschaft.

Notation. Im Zusammenhang mit metrischen Raeumen und damit auch mit normierten Raeumen verwenden wir folgende Bezeichnungen:

- Wir bezeichnen die offene bzw. abgeschlossene Kugeln mit U bzw B gemaess

$$U_\varepsilon(x) = \{y \in E : \|y - x\| < \varepsilon\}$$

  und

$$B_\varepsilon(x) = \{y \in E : \|y - x\| \leq \varepsilon\}.$$

- Wir setzen $x_n \xrightarrow{\|\cdot\|} x$ falls $x_n \to x$ bzgl d.
- (x_n) heisst Cauchy Folge bzgl. $\|\dot{\|}$ falls (x_n) Cauchy-Folge bzgl. d.
- $(E, \|\cdot\|)$ heisst vollstaendig, wenn (E,d) vollstaendig ist.

PROPOSITION (Variante Dreiecksungleichung). *Sei $(E, \|\cdot\|)$ eine normierter Raum. Dann gilt*

$$|\|x\| - \|y\|| \leq \|x - y\|$$

fuer alle $x, y \in E$. Insbesondere ist $\|\cdot\| : E \longrightarrow [0, \infty)$ stetig.

Beweis. Es gilt

$$\|x\| = \|x - y + y\| \leq \|x - y\| + \|y\|$$

also

$$\|x\| - \|y\| \leq \|x - y\|$$

und analog (nach Vertauschen von x und y)

$$\|y\| - \|x\| \leq \|y - x\| = \|x - y\|.$$

Damit folgt

$$|\|x\| - \|y\|| \leq \|x - y\|.$$

Stetigkeit von $\|\cdot\|$. Nach dem schon gezeigten gilt

$$|\|x\| - \|x_0\|| \leq d(x, x_0)$$

und die Stetigkeit von $\|\cdot\|$ folgt sofort.

□

DEFINITION (Banachraum). *Ein vollstaendiger normierter Vektorraum heisst Banachraum.*

Beispiele für Banachraeume.

- Hilbertraeume. (Diese sind nach Definition vollstaendig.)
- $L^p(X, \mathcal{A}, \mu)$ mit $\|\cdot\|_p$ fuer $p \in [1, \infty]$ und $(X, \mathcal{A}, \mu)$ Massraum
- $(C_0(X), \|\cdot\|_\infty)$ ist vollstaendig (Uebung).
- (Uebung) Sei $C_c = C_c(\mathbb{N})$ der Vektorraum der Funktionen auf $\mathbb{N}$ die fuer alle bis auf endlich viele Punkte verschwinden. Dann ist $C_c(X)$ nicht vollstaendig als Teilraum von ℓ^2. Ebenso ist $C_c(X)$ nicht vollstaendig als Teilraum von $C_0(X)$. Tastsaechlich gilt:
  - $\overline{C_c(\mathbb{N})} = C_0(\mathbb{N})$.
  - $C_c(\mathbb{N}) \neq C_0(\mathbb{N})$.

Wir diskutieren nun, wie die Bildung von Quotienten mit der Norm vertraeglich ist: Sei dazu E ein normierter Vektorraum und F ein abgeschlossener Unterraum von E und E/F der Quotient mit der kanonischen Projektion

$$\pi : E \longrightarrow E/F, x \mapsto x + F.$$

THEOREM. *Sei $(E, \|\cdot\|)$ ein normierter Raum und F ein abgeschlossener Unterraum von E. Dann wird durch*

$$|||\cdot||| : E/F \longrightarrow [0, \infty),\ t \mapsto \inf\{\|y\| : y \in t\} =: |||t|||$$

eine Norm auf E/F eingefuehrt. Es ist $U \subset E/F$ offen genau dann, wenn $\pi^{-1}(U)$ offen ist. Insbesondere ist also π stetig und offen. Ist $(E, \|\cdot\|)$ vollstaendig, so ist auch $(E/F, |||\cdot|||)$ vollstaendig.

Bemerkung.

- Die Aussage zu π liefert gerade, dass die Definition der Norm genau mit der Stetigkeit von π abgestimmt ist.

- Der Beweis der Vollstaendigkeit ist aehnlich wie ein entsprechender Beweis fuer die L^p Raeume. Das ist kein Zufall, denn auch fuer die L^p Raume ging es darum die Vollstaendigkeit einer Quotientennorm zu zeigen.
- Der Quotient E/F kann manchmal auch als Ersatz fuer das (in nicht Hilbertraeumen nicht existierende) orthogonale Komplement von F gesehen werden, in dem Sinne, dass E in gewisser Weise gleich $E/F + F$ sein sollte.

Beweis. Es ist nicht schwer zu zeigen, dass $|||\cdot|||$ eine Norm ist. Die Abgeschlossenheit von F liefert die Nichtausgeartetheit.

Aussage zur π. Die Definition von $|||\cdot|||$ zeigt

$$U_\delta(\pi(x)) = \pi(U_\delta(x)).$$

Damit sieht man leicht, dass $U \subset E/F$ offen ist genau dann, wenn $\pi^{-1}(U)$ offen ist. Damit ist π stetig. Die Offenheit von π folgt nun so: Sei U offen in E. Dann ist auch

$$\pi^{-1}(\pi(U)) = \cup_{f\in F}(f + U)$$

offen. Damit ist $\pi(U)$ also offen, nach dem schon gezeigten.

Vollstaendigkeit: Sei (t_n) eine Cauchy Folge in E/F. Ohne Einschraenkung (sonst Teilfolge) koennen wir

$$|||t_n - t_m||| < \frac{1}{2^{n+1}}$$

fuer alle $m \geq n$ annehmen. Daher koennen wir induktiv $x_n \in E$ mit $\pi(x_n) = t_n$ konstruieren mit

$$\|x_n - x_{n+1}\| < \frac{1}{2^n} :$$

$n = 1$: Waehle x_1 mit $\pi(x_1)$ beliebig.
$n \Longrightarrow (n + 1)$: Seien $x_1, \ldots, x_n$ wie gewuenscht schon konstruiert. Wegen

$$|||\pi(x_n) - t_{n+1}||| < \frac{1}{2^{n+1}}$$

und der Definition von $|||\cdot|||$ gibt es also ein x_{n+1} mit

$$\|x_n - x_{n+1}\| < \frac{1}{2^n}.$$

Die (x_n) sind dann offenbar eine Cauchy Folge in E (Dreiecksungleichung). Daher konvergiert (x_n) gegen ein $x \in E$. Dann gilt aufgrund der Stetigkeit von π

$$t_n = \pi(x_n) \to \pi(x) =: t.$$

Das liefert die gewuenschte Vollstaendigkeit. $\square$

Bemerkung. (a) Ein Vektorraum kann natuerlich mehrere Normen erlauben. Damit stellt sich dann die Frage, welche Normen man als aequivalent betrachten will. Zwei naheliegende Ansaetze sind es Normen $\|\cdot\|_1$ und $\|\cdot\|_2$, als aequivalent zu betrachten, wenn sie

- dieselbe Topologie erzeugen, d.h. $id : (E, \|\cdot\|_1) \longrightarrow (E, \|dot\|_2)$ und $id : (E, \|\cdot\|_2) \longrightarrow (E, \|dot\|_1)$ stetig sind,
- sie Konstanten $c, C > 0$ zulassen mit $\|\|_1 \leq c\|\cdot\|_2$ und $\|\cdot\|_2 \leq c\|\cdot\|_1$.

Tatsaechlich sind beide Forderungen aequivalent. Es gilt naemlich folgendes: Sei E ein Vektorraum und $\|\cdot\|_1$, $\|\cdot\|_2$ Normen auf E. Dann sind aequivalent:

(i) Es ist $Id : (X, \|\cdot\|_1) \longrightarrow (X, \|\cdot\|_2)$ stetig.
(ii) Es existiert ein $C > 0$ mit $\|x\|_2 \leq C\|x\|_1$ fuer alle $x \in E$.
(iii) Aus $x_n \to 0$ bzgl. $\|\cdot\|_1$ folgt $x_n \to 0$ bzgl. $\|\cdot\|_2$.

Beweis. (Uebung). Wir werden gleich eine noch allgemeiner Aussage kennen lernen. □

(b) Ist E ein endlichdimensionale Vektorraum, so sind alle Normen aequivalent (Uebung). Fuer unendlichdimensonale Vektorraeume gilt das nicht.

2. Stetige Abbildungen zwischen normierten Raeumen

In diesem Abschnitt untersuchen wir die Stetigkeit von linearen Abbildungen zwischen normierten Vektorraeumen. Fuer den Spezialfall der linearen Abbildungen zwischen Hilbertraumes haben wir das schon frueher untersucht. Es zeigt sich, dass die damals gemachten Ueberlegungen leicht verallgemeinert werden koennen.

Theorem (Charakterisierung Stetigkeit von linearen Abbildungen zwischen normierten Raeumen). *Seien $(E, \|\cdot\|_E)$ und $(F, \|\cdot\|_F)$ normierte Raeume und $T : E \longrightarrow F$ linear. Dann sind aequivalent:*

(i) *T ist stetig.*
(ii *T ist stetig bei* 0.
(iii) *Es existiert ein $C > 0$ mit $\|Tx\|_F \leq C\|x\|_E$ fuer alle $x \in E$.*
(iv) *Es gilt* $\sup_{\|x\|_E \leq 1} \|Tx\|_F < \infty$.

Beweis. (i)$\Longrightarrow$ (ii): Das ist klar.

(ii)$\Longrightarrow$ (iii): Aufgrund von (ii) existiert ein $\delta > 0$ mit $\|Tx\|_F \leq 1$ fuer alle $x \in E$ mit $\|x\|_E \leq \delta$. Damit erhaelt man fuer beliebige $x \neq 0$ dann also

$$\|Tx\|_F = \frac{\|x\|_E}{\delta}\|T(\frac{\delta}{\|x\|_E}x)\|_F \leq \frac{\|x\|_E}{\delta}.$$

Das zeigt (iii).

(iii)$\Longrightarrow$ (iv): Das ist klar.

(iv)$\Longrightarrow$ (i): Sei $S := \sup_{\|x\|_E \leq 1} \|Tx\|_F$. Sei $x \in E$ beliebig. Dann gilt fuer alle $y \in E$ mit $y \neq x$ also

$$\|Tx - Ty\|_F = \|T(x-y)\| = \|x-y\|_E \|T(\frac{1}{\|x-y\|_E}(x-y))\|_F \leq S\|x-y\|_E.$$

Damit folgt sofort die Stetigkeit von T.

Weil es so schoen ist, noch ein Beweise fuer weitere Implikationen:

(ii)$\Longrightarrow$ (i): Es gelte $x_n \to x$. Dann folgt $x - x_n \to 0$. Damit folgt aus (ii) dann $T(x_n - x) \to 0$. Aufgrund der Linearitaet von T ergibt sich $Tx_n \to Tx$. Das ist (i).

(ii)$\Longrightarrow$ (iii): Angnommen nein! Dann gibt es zu jedem $n \in \mathbb{N}$ ein $x_n \in E$ mit

$$\|Tx_n\|_F > n\|x_n\|_E.$$

Es gilt dann $x \neq 0$ und wir koennen ohne Einschraenkung (sonst Skalieren) annehmen $\|x_n\|_E = \frac{1}{n}$. Dann folgt also $x_n \to 0$ aber $\|Tx_n\|_F \geq 1$ fuer alle $n \in \mathbb{N}$. Das ist ein Widerspruch zu (ii).

(iii)$\Longrightarrow$ (ii): Das ist klar.

(iii)$\Longrightarrow$ (iv): Das ist klar.

(iv)$\Longrightarrow$ (iii): Setze

$$C := \sup\{\|Tx\|_F : \|x\|_E \leq 1\}.$$

Dann folgt fuer $x \neq 0$ sofort

$$\|Tx\|_F = \|x\|_E \|T\frac{1}{\|x\|_E} x\|_F \leq \|x\|_E C.$$

Natuerlich gilt das auch fuer $x = 0$ (da dann beide Seite 0 sind). Insgesamt folgt also (iii):

(ii)$\Longrightarrow$ (iv): Da T stetig stetig in 0 ist, gibt es ein $\delta > 0$ mit $TB_\delta^E(0) \subset B_1^F(0)$. Dann gilt also

$$\sup\{\|Tx\|_F : \|x\|_E \leq \delta\} \leq 1.$$

Damit folgt (nach Skalieren mit $1/\delta$)

$$\sup\{\|Tx\|_F : \|x\|_E \leq 1\} \leq \frac{1}{\delta}.$$

Das ist gerade (iv).

(iv)$\Longrightarrow$ (ii): Angenommen nein! Dann existiert zu jedem $n \in \mathbb{N}$ ein $x_n \in E$ mit $\|x_n\| \leq 1$ und $\|Tx_n\|_F \geq n$. Setze $y_n := \frac{1}{n}x_n$. Dann gilt $y_n \to 0$ aber $\|Ty_n\|_F \geq 1$. Das ist ein Widerspruch. $\square$

Bemerkung.

- Nach (iii) bedeutet die Stetigkeit von T gerade, dass die Halbnorm $x \mapsto \|Tx\|_F$ auf E durch die Halbnorm $\|\cdot\|_E$ abgeschaetzt werden kann.
- Den Fall von Hilbertraeumen haben wir schon behandelt.
- Im unendlichdimensionalen Fall ist es leicht moeglich nichtstetige lineare Abbildungen zu erzeugen: Mit dem Zornschen Lemma zeigt man leicht die Existenz von Basen (x_α), $\alpha \in A$, von E und (y_β), $\beta \in B$, von F (d.h. jedes Element von E bzw. F kann eindeutig als *endliche* Linearkombination der (x_α) bzw. (y_β) geschrieben werden. Ohne Einschraenkung enthaelte die Indexmengen A und B jeweils eine Kopie von $\mathbb{N} \cup \{0\}$. Dann gibt es fuer jede Folge c_n die eine eindeutige lineare Abbildung

  $$T : E \longrightarrow F$$

  mit $Tx_n = y_n$ und $Tx_\alpha = 0$ fuer $\alpha \notin \mathbb{N} \cup \{0\}$. Offenbar ist T nicht stetig.

Proposition (Norm auf dem Raum der linearen Abbildungen). *Seien $(E, \|\cdot\|_E)$ und $(F, \|\cdot\|_F)$ normierte Raeume. Dann gilt fuer jedes lineare $T : E \longrightarrow F$*

$$\begin{aligned} &\inf \quad \{C > 0 : \|Tx\|_F \leq C\|x\|_E \quad \text{alle } x \in E\} \\ = \ &\sup \quad \{\|Tx\|_F : \|x\|_E \leq 1\} = S. \end{aligned}$$

sowie

$$\|Tx\| \leq \|T\|\|x\|$$

fuer alle $x \in E$.

Beweis. Das folgt wie im Hilbertraum. Hier sind die Details:
Wir zeigen zunaechst Gleichheit von

$$I := \inf\{C > 0 : \|Tx\|_F \leq C\|x\|_E \text{ alle } x \in E\}$$

und

$$S := \sup\{\|Tx\|_F : \|x\|_E \leq 1\}.$$

$I \leq S$: Nach Definition von S gilt fuer $x \neq 0$

$$\|Tx\|_F = \|x\|_E \|T\frac{1}{\|x\|_E} x\|_F \leq S\|x\|_E.$$

Damit folgt $I \leq S$.
$S \leq I$: Sei $\|Tx\|_F \leq C\|x\|_E$ fuer alle $x \in E$. Dann gilt also insbesondere

$$\|Tx\|_F \leq C$$

fuer alle $\|x\|_E \leq 1$. Damit folgt $S \leq C$ und somit $S \leq I$.
Es gilt die Ungleichung $\|Tx\| \leq \|T\|\|x\|$. Im Beweis von $I \leq S$ wurde das gezeigt fuer $x \neq 0$. Offenbar gilt es auch fuer $x = 0$. Die Gueltigkeit dieser Ungleichung zeigt auch, dass man das Infimum in der Definition von I durch ein Minimum ersetzen kann. □

Es ist $\|\cdot\|$ eine Norm. Das ist einfach.

DEFINITION. *Sind E und F normierte Raeume, so bezeichnen wir den Vektorraum aller stetigen linearen Abbildungen von E nach F mit $L(E, F)$.*

Nach der Prosposition zur Stetigkeit von linearen Abbildungen gehoert eine lineare Abbildung $T : E \longrightarrow F$ zu $L(E, F)$ genau, dann wenn

$$\|T\| := \sup\{\|Tx\|_F : \|x\|_E \leq 1\} < \infty.$$

Nach der vorigen Proposition gilt dann

$$\|Tx\|_F \leq \|T\|\|x\|_E$$

fuer alle $x \in E$.
Wie im Hilbertraum beweist man die folgende Aussage.

PROPOSITION. *Sind $(E, \|\cdot\|_E)$ und $(F, \|\cdot\|_F)$ normierte Raeume, so wird durch das eben definierte $\|\cdot\|$ eine Norm auf $L(E, F)$ gegeben.*

THEOREM. *Seien $(E, \|\cdot\|_E)$ und $(F, \|\cdot\|_F)$ normierte Raeume. Ist $(F, \|\cdot\|_F)$ vollstaendig. So ist $(L(E, F), \|\cdot\|)$ vollstaendig.*

Beweis. Das folgt wie im Hilbertaum. Hier sind die Details:
Sei (T_n) eine Cauchy Folge, d.h. fuer alle $\varepsilon > 0$ existiere ein $N_\varepsilon \in \mathbb{N}$ mit

$$\|T_n - T_m\| \leq \varepsilon$$

fuer alle $n, m \geq N_\varepsilon$. Dann ist $(T_n x)$ eine Cauchy Folge fuer jedes $x \in E$. Da F vollstaendig ist, existiert

$$Tx := \lim_{n\to\infty} T_n x$$

fuer jedes $x \in E$.
T ist linear. $T(x+\lambda x) = \lim_{n\to\infty} T_n(x+\lambda y) = ...$
T ist stetig. Sei $N = N_1$. Da $\|\cdot\|$ stetig ist, gilt dann

$$\|Tx\| = \|\lim_{n\to\infty} T_n x\| = \lim_{n\to\infty} \|T_n x\| \leq \limsup_{n\to\infty}(\|T_n x - T_N x\| + \|T_N x\|) \leq (1+\|T_N\|)\|x\|.$$

T_n konvergiert gegen T. Fuer $n \geq N_\varepsilon$ gilt

$$\|(T-T_n)x\| = \lim_{k\to\infty} \|(T_k - T_n)x\| \leq \varepsilon \|x\|$$

fuer alle $x \in E$. □

Wir werden das alles bald anwenden fuer den Fall, dass $(F, \|\cdot\|_F)$ gerade der zugrundeliegende Koerper ist. Das fuehrt auf das Konzept des Dualraumes. Wir werden zeigen, dass der Dualraum sehr gross ist. Dazu dient der folgende Abschnitt.

3. Der Satz von Hahn / Banach

Der Satz von Hahn-Banach liefert die Existenz 'vieler' schoener Funktionale. Genauer besagt er die Fortsetzbarkeit von Funktionalen auf Teilraeumen auf den gesamten Raum unter Erhalt einer Abschaetzung. Diese Abschaetzung kann (unter anderem) dazu verwendet werden die Stetigkeit der entsprechenden Funktionale zu garantieren.
Der Beweis des Satzes von Hahn-Banach hat zwei Bestandteile. Der eine Bestandteile liefert eine echte Fortsetzbarkeit eines noch nicht auf dem gesamten Raum definieren Funktionals. Dies ist ein konstruktiver Beweis. Der andere Bestandteil besagt die Existenz einer maximalen Fortsetzung. Dabei wird das sogenannte Zornsche Lemma benutzt. Zusammengenommen liefern beide Bestandteile die Aussage.
Wir werden den Satz von Hahn-Banach zur Untersuchung normierter Raeume verwenden. Das ist aber nicht die einzige Anwendung. Der Satz ist allgemeiner.

Definition (Sublineare Funktionale und Halbnnormen). *Sei E ein Vektorraum ueber $\mathbb{K}$. Ein sublineares Funktional p auf E ist eine Abbildung $p : E \longrightarrow \mathbb{R}$ mit*

- *$p(\lambda x) = \lambda p(x)$ fuer alle $x \in E$ und $\lambda \geq 0$.*
- *$p(x+y) \leq p(x) + p(y)$ fuer alle $x, y \in E$.*

Bemerkung. (a) Ein sublineares Funktional p mit $p(\lambda x) = |\lambda| p(x)$ fuer alle $x \in E$ und $\lambda \in \mathbb{K}$ ist gerade eine Halbnorm. Eine Halbnorm p mit $p(x) \neq 0$ fuer $x \neq 0$ ist eine Norm.
(b) Nicht jedes sublineare Funktional kommt von einer Halbnorm. Betrachte zum Beispiel auf dem Raum M_n der (reellen) $n \times n$ Matrizen das Funktional

$$p(x) := \max \sigma(\frac{1}{2}(x + x^*)).$$

Dann ist p sublinear, kann aber auch negative Werte annehmen (Uebung). Statt M_n kann man auch die Algebra aller beschraenkten Operatoren in einem beliebigen Hilbertraum betrachten.

Notation. Seien E und F Vektorraume (ueber demselben Koerper). Ist G ein Unterraum von E und $z : G \longrightarrow F$ linear und H eine Unterraum von E mit $G \subset H$, so nennt man $Z : H \longrightarrow F$ eine Forsetzung von z (auf H), wenn $Z(x) = z(x)$ fuer alle $x \in G$ gilt.

LEMMA (Fortsetzungslemma). *Sei E ein Vektorraum ueber $\mathbb{R}$ und $p : E \longrightarrow \mathbb{R}$ sublinear. Sei $z : G \longrightarrow \mathbb{R}$ linear auf dem Unterraum G von E mit $z \leq p$ (d.h. $z(x) \leq p(x)$ fuer $x \in G$). Sei $x_0 \in E \setminus G$. Dann existiert eine Fortsetzung Z von z auf*

$$H = Lin\{x_0, G\} = \{\lambda x_0 + y : \lambda \in \mathbb{R}, y \in G\} \longrightarrow \mathbb{R}$$

mit $Z \leq p$.

Beweis. Ist Z eine Fortsetzung von z auf H, so gibt es offenbar ein a (naemlich $a = Z(x_0)$ mit

$$Z(\lambda x_0 + y) = \lambda a + z(y)$$

fuer alle $\lambda \in \mathbb{R}$ und $y \in G$. Umgekehrt wird nach Wahl eines a durch

$$Z(\lambda x_0 + y) = \lambda a + z(y)$$

fuer alle $\lambda \in \mathbb{R}$ und $y \in G$ offenbar eine Fortsetzung von z definiert. Die Frage ist also, ob man ein solches a finden kann unter Erhalt der Ungleichung

$$Z(\lambda x_0 + y) \leq p(\lambda x_0 + y)$$

fuer alle $\lambda \in \mathbb{R}$ und $y \in G$. Da beiden Seiten unter Skalieren mit $1/\lambda$ (fuer $\lambda \neq 0$) invariant sind, reicht es also ein $a \in \mathbb{R}$ zu finden mit

$$Z(x_0 + y) \leq p(x_0 + y) \text{ also } a \leq p(x_0 + y) - z(y)$$

fuer alle $y \in G$ und

$$Z(-x_0 + y) \leq p(-x_0 + y) \text{ also } a \geq -p(-x_0 + y) + z(y)$$

fuer alle $z \in G$. Wir bestimmen nun ein geeignetes a:
Fuer alle $x, y \in G$ gilt

$$z(x) + z(y) = z(x + y) \leq p(x + x_0 + (y - x_0)) \leq p(x + x_0) + p(y - x_0).$$

Also folgt

$$z(y) - p(y - x_0) \leq p(x + x_0) - z(x)$$

fuer alle $x, y \in G$. Damit folgt

$$m := \sup_{y \in G}\{z(y) - p(y - x_0)\} \leq \inf_{x \in G}\{p(x + x_0) - z(x)\} =: M.$$

Waehle nun $a \in [m, M]$. Dann gilt also nach Konstruktion

$$-p(-x_0 + y) + z(y) \leq a \leq p(x_0 + y) - z(y)$$

fuer alle $x, y \in G$ und die beiden obigen Ungleichungen sind erfuellt. $\square$

THEOREM (Hahn - Banach: reeller Fall). *Sei E ein Vektorraum ueber $\mathbb{R}$ und p ein sublineares Funktional auf E. Sei F ein Unterraum von E und $\varphi : F \longrightarrow \mathbb{R}$ linear mit $\varphi(x) \leq p(x)$ fuer alle $x \in F$. Dann existiert ein lineares Funktional Φ auf E mit $\Phi|_F = \varphi$ und $\Phi \leq p$.*

Ende der Vorlesung

Beweis. Sei $\mathcal{Z}$ die Menge der Fortsetzungen von φ, die die Ungleichung erfuellen. Es besteht also $\mathcal{Z}$ aus den Paaren (G, ψ) mit folgenden Eigenschaften:

- G Unterraum von E mit $G \subset E$,
- $\psi : G \longrightarrow \mathbb{R}$ linear mit $\psi \leq p$
- $\psi_F = \varphi$.

Dann ist $\mathcal{Z}$ nicht leer (da $(F, \varphi) \in \mathcal{Z}$). Durch $(G_1, \psi_1) \prec (G_2, \psi_2)$ falls $G_1 \subset G_2$ und $\psi_2|_{G_1} = \psi_1$ wird eine Halbordnung auf $\mathcal{Z}$ definiert. Ist $\mathcal{C}$ eine total geordnete Menge in $\mathcal{Z}$, so ist offenbar (G_m, ψ_m) mit

$$G_m := \cup_{(G,\psi)\in\mathcal{C}} G,\ \psi_m(x) = \psi(x);\text{ fuer } x \in G \text{ mit } (G, \psi) \in \mathcal{C}$$

eine obere Schranke von $\mathcal{C}$. Daher gibt es nach dem Lemma von Zorn also ein maximales Element (G, ψ) in $\mathcal{Z}$. Dann muss aber nach dem vorigen Lemma gelten $G = E$. (Sonst gaebe es $x_0 \in E \setminus G$ und wir koennten, nach dem vorigen Lemma, ψ noch auf $Lin\{x_0, G\}$ fortsetzen. Widerspruch). $\square$

THEOREM (Hahn-Banach: Halbnorm). *Sei E ein Vektorraum ueber $\mathbb{K}$, p eine Halbnorm auf E. Sei F ein Unterraum von E und $\varphi : E \longrightarrow \mathbb{K}$ linear mit $|\varphi| \leq p$. Dann existiert ein lineares $\Phi : E \longrightarrow \mathbb{K}$ mit $\Phi|_F = \varphi$ und $|\Phi| \leq p$.*

Beweis. $\mathbb{K} = \mathbb{R}$. Nach Voraussetzung gilt $\varphi \leq p$ (denn $\varphi \leq |\varphi| \leq p$). Daher existiert nach dem vorigen Satz eine Fortsetzung $\Phi : E \longrightarrow \mathbb{R}$ linear mit $\Phi \leq p$. Dann gilt also auch $-\Phi(x) = \Phi(-x) \leq p(-x) = p(x)$ und $|\Phi| \leq p$ folgt.

$K = \mathbb{C}$. Wir beginnen mit einer Vorueberlegung zum 'Komplexifizieren' von reellen Funktionalen.

- Ist $\varphi : E \longrightarrow \mathbb{C}$ linear (ueber $\mathbb{C}$), so ist $u := \Re\varphi : E \longrightarrow \mathbb{R}$ linear ueber $\mathbb{R}$ mit $\varphi(x) = u(x) - iu(ix)$. (Kleine Rechnung: $\Im\varphi(x) = -\Re\varphi(ix)$...)
- Ist $u : E \longrightarrow \mathbb{R}$ linear (ueber $\mathbb{R}$), so ist $\varphi(x) := u(x) - iu(ix)$ linear ueber $\mathbb{C}$ mit $\Re\varphi = u$. (Kleine Rechnung: Offenbar linear ueber $\mathbb{R}$; reicht zu zeigen $\varphi(ix) = i\varphi(x)$; Einsetzen...)

Damit bestimmt also der Realteil des Funktionals schon das Funktional. Entsprechend reicht es, den Realteil fortzusetzen und wir koennen damit den komplexen Fall wie folgt auf den reelen Fall zurueckfuehren:
Betrachte $u := \Re\varphi$. Dann gilt $|u| \leq p$ und u ist linear ueber $\mathbb{R}$. Damit folgt aus dem schon untersuchten reellen Fall, dass ein lineares (ueber $\mathbb{R}$) $U : E \longrightarrow \mathbb{R}$ existiert mit $|U| \leq p$ und $U|_F = u$. Dann ist nach der Vorueberlegung

$$\Phi : E \longrightarrow \mathbb{C},\ \Phi(x) = U(x) - iU(ix)$$

linear ueber $\mathbb{C}$ und eine Fortsetzung von φ auf E. Es bleibt die Ungleichung $|\Phi| \leq p$ zu zeigen: Es gilt mit geeignetem $\alpha \in \mathbb{R}$

$$\begin{array}{rcl} \|\Phi(x)| & = & e^{i\alpha}\Phi(x) \\ & = & \Phi(e^{i\alpha}x) \\ (reell) & = & \Re\Phi(e^{i\alpha}x) \\ (Definition\ U) & = & U(e^{i\alpha}x) \\ & \leq & p(e^{i\alpha}x) \\ (p\ Halbnorm) & = & p(x). \end{array}$$

Das beendet den Beweis. □

4. Dualraeume normierter Raeume

In diesem Abschnitt lernen wir grundlegende Eigenschaften der Dualraeume normierter Raeume kennen. Es zeigt sich, dass diese Raeume immer vollstaendig sind und genuegend viele Funktionale enthalten, um (nicht nur) die Punkte des Raumes von einander zu trennen. Die Aussagen dieses Abschnittes sind Folgerungen aus den beiden vorherigen Abschnitten.

DEFINITION. *Sei $(E, \|\cdot\|)$ ein normierter Raum. Dann heisst der Vektorraum der stetigen linearen Abbildungen von E in den zugrundeliegenden Köper der Dualraum von E und wird mit E' bezeichnet.*

Offenbar gilt also fuer einen normierten Raum $(E, \|\cdot\|)$ dann

$$E' = L(E, \mathbb{K}),$$

wobei der zugrundeliegende Koerper mit der Norm $|\cdot|$ versehen ist. Damit erhalten wir aus den Aussgen zu $L(E, F)$ von Abschnitt 2 dieses Kapitels sofort die folgenden Aussagen.

LEMMA. *Sei $(E, \|\cdot\|)$ ein normierter Raum und $\varphi : E \longrightarrow \mathbb{K}$ linear. Dann sind aquivalent:*

(i) *Es gehoert φ zu E'.*
(ii) *Es gibt ein $C \geq 0$ mit $|\varphi(x)| \leq C\|x\|$ fuer alle $x \in E$.*
(iii) *Es gilt $\sup\{|\varphi(x)| : \|x\| \leq 1\} < \infty$.*

THEOREM. *Ist $(E, \|\cdot\|)$ normierter Raum, so ist der Dualraum $(E', \|\cdot\|)$ mit*

$$\|\varphi\| := \sup\{|\varphi(x)| : \|x\| \leq 1\}$$

ein vollstaendiger normierter Raum.

Beweis. Das folgt da $(\mathbb{K}, |\cdot|)$ vollstaendig ist. □

THEOREM (Hahn-Banach fuer normierte Raeume). *Sei $(E, \|\cdot\|)$ ein normierter Vektorraum ueber $\mathbb{K}$. Ist U ein Unterraum von E und $\Psi : U \longrightarrow \mathbb{K}$ ein stetiges Funktional auf U (mit $|\Psi| \leq C\|\cdot\|$ fuer ein $C \geq 0$), so laesst sich Ψ zu einem stetigen Funktional ψ auf E fortsetzen (mit $|\psi| \leq C\|\cdot\|$).*

Beweis. Das folgt sofort aus dem Satz von Hahn-Banach. □

Damit kommen nun zu entscheidenden Konsequenzen des Satzes von Hahn-Banach fuer normierte Raeume: Auf normierten Raeumen gibt es viele Funktionale; genauer gesagt, genuegend viele, um die Punkte (und mehr) voneinander zu trennen.

Folgerung (Hahn-Banach Konsequenz: Ausrechnen der Norm). *Sei $(E, \|\cdot\|)$ ein normierter Vektorraum ueber $\mathbb{K}$ und $x \in E$. Dann existiert ein stetiges $\varphi : E \longrightarrow \mathbb{K}$ mit $\varphi(x) = \|x\|$ und $\|\varphi\| \leq 1$. Insbesondere gilt*

$$\|z\| = \max\{|\varphi(z)| : \varphi \in E' : \|\varphi\| \leq 1\}$$

fuer alle $z \in E$.

Beweis. Zur ersten Aussage: Sei U der von x aufgespannte Unterraum und

$$\Phi : U \longrightarrow \mathbb{K},\ \alpha x \mapsto \alpha\|x\|.$$

Dann gilt offenbar $|\Phi| \leq \|\cdot\|$. Nach dem vorangehenden Satz von Hahn-Banach fuer normierte Raeume existiert dann eine Fortsetzung φ von Φ mit $\|\varphi\| \leq 1$ und $\varphi(x) = \|x\|$.
Zum 'Insbesondere': Offenbar gilt

$$\sup\{|\varphi(z)| : \varphi \in E' : \|\varphi\| \leq 1\} \leq \|z\|.$$

Mit der ersten schon bewiesenen Aussage folgt dann die gewuenschte Gleichheit. □

Aus der vorangehenden Folgerung ziehen wir nun weitere Folgerungen. Zunaechst sagt die Folgerung, dass es zu $x \neq 0$ ein lineares stetiges φ gibt mit $\varphi(x) \neq 0$. Damit kann man also die 0 von $x \neq 0$ durch φ unterschweiden. Tatsaechlich kann man beliebige Punkte voneinander unterscheiden:

Folgerung (Hahn-Banach Konsequenz: Trennung der Punkte). *Sei $(E, \|\cdot\|)$ ein normierter Vektorraum ueber $\mathbb{K}$ und $x \in E$. Dann trennt $E^{'}$ die Punkte von E (d.h. zu $x, y \in E$ mit $x \neq y$ existiert ein $\varphi \in E^{'}$ mit $\varphi(x) \neq \varphi(y)$).*

Beweis. Anwenden der vorangehenden Folgerung auf $z = x - y \neq 0$ liefert ein $\varphi \in E'$ mit $0 \neq \varphi(z) = \varphi(x) - \varphi(y)$. □

Auch kann man nicht nur die Norm von Vektoren, sondern auch von Abbildungen mittels des Dualraumes ausrechnen.

Folgerung. *Sei $T : E \longrightarrow F$ eine stetige Abbildung zwischen normierten Raeumen. Dann gilt*

$$\|T\| := \sup\{|(\varphi, Tx)| : \varphi \in E^{'} \textit{ mit } \|\varphi\| \leq 1,\ x \in E,\ \textit{mit } \|x\| \leq 1\}.$$

Beweis. Das folgt leicht durch Anwenden der obigen Folgerung zum Ausrechnen der Norm. □

Wir wollen nun noch eine Aussage zur Trennung von Punkten von abgeschlossenen Unterraeumen beweisen. Dazu brauchen wir noch folgendes Hilfslemma, das auch fuer sich schon von Interesse ist.

Lemma. *Sei $(E, \|\cdot\|)$ ein normierter Raum und $\varphi : E \longrightarrow \mathbb{K}$ linear. Dann sind aequivalent:*

(i) *Es ist* φ *stetig.*
(ii) *Es ist* $Ker(\varphi)$ *abgeschlossen.*

Bemerkung. (a) Es ist ausgesprochen bemerkenswert, dass zur Feststellung der Stetigkeit die Abgeschlossenheit eines einzigen Urbildes ausreicht.
(b) Die Aussage gilt nicht fuer allgemeine lineare Abbildungen $T : E \longrightarrow F$ wenn F unendlichdimensional ist. (Waehle eine Basis (x_α), $\alpha \in A$. Sei ohne Einschraenkung $\mathbb{N} \cup \{0\} \subset A$. Dann gibt es eine eindeutige lineare Abbildung

$$T : E \longrightarrow E$$

mit $Tx_0 = 0$ und $Tx_n = nx_n$ fuer $n \in \mathbb{N}$ und $Tx_\alpha = x_\alpha$ fuer $\alpha \in A \setminus (\mathbb{N} \cup \{0\})$. Dann ist $Ker(T) = Lin\{x_{\alpha_0}\}$ abgeschlossen, aber T ist nicht stetig.)

Beweis. Es reicht Stetigkeit bei 0 zu zeigen. Sei (x_n) eine Folge in E mit $x_n \to 0$. Zu zeigen $\varphi(x_n) \to 0$. *Angenommen* $\varphi(x_n)$ konvergiert nicht gegen 0. Dann koennen wir ohne Einschraenkung (sonst Teilfolge) annehmen $|\varphi(x_n)| \geq C > 0$ fuer alle n. Weiterhin gibt es dann offenbar ein $x \in E$ mit $\varphi(x) = 1$. Betrachte nun

$$y_n := x - \frac{1}{\varphi(x_n)} x_n.$$

Dann gilt $\varphi(y_n) = 0$, d.h. es gehoert jedes y_n zu $Ker(\varphi)$. Weiterhin konvergiert aufgrund der Voraussetzungen offenbar (y_n) gegen x. Da $Ker(\varphi)$ abgeschlossen ist gehoert damit also x zu $Ker(\varphi)$. Das ist ein Widerspruch zu $\varphi(x) = 1$. □

Damit koennen wir nun folgende Anwendung des Satzes von Hahn-Banach geben.

FOLGERUNG. *Sei* $(E, \|\cdot\|)$ *ein normierter Raum und* U *ein abgeschlossener Unterraum und* $x \notin U$. *Dann existiert ein* $\varphi \in E'$ *mit* $\varphi(x) = 1$ *und* $\varphi = 0$ *auf* U.

Beweis. Wir betrachten das Funktional $\psi : Lin(x, U) \longrightarrow \mathbb{K}, \alpha x + u \mapsto \alpha$. Dann ist der Kern von ψ (d.h. U) abgeschlossen und damit ist ψ stetig. Nun laesst sich Hahn Banach anwenden. □

Interpretation der Folgerungen - Trennung durch Hyperebenen. Sei E ein Vektorraum. Eine Teilmenge H von H der Form $H = x + U$ mit einem Unterraum U heisst, *Hyperebene* wenn $Lin\{s, U\} = E$ fuer jedes $s \notin U$. (Dann heisst U auch 1-codimensional.) Ist E ein normierter Raum, so stehen abgeschlossene Hyperebenen und stetige Funktionale in einer engen Beziehung:
- Ist φ eine stetiges Funktional mit $\varphi \neq 0$, so ist

$$H_\alpha := \{x \in E : \varphi(x) = \alpha\}$$

eine Hyperebene fuer jedes $\alpha \in \mathbb{K}$. In der Tat gilt $H_\alpha = x + Ker(\varphi)$ fuer jedes $x \in H_\alpha$ und $Ker(\varphi)$ ist offenbar 1-codimensional.
- Ist umgekehrt $H = x + U$ eine abgeschlossene Hyperebene, so ist U abgeschlossen und damit gibt es dann (nach der vorigen Folgerung aus Hahn-Banach) ein stetiges φ mit $H = H_1$.

Eine Grundproblem in verschiedenen Anwendungen ist die sogenannte *Trennung konvexer Mengen durch abgeschlossene Hyperebenen.* Dabei geht es darum, zu konvexen disjunkten Teilmengen C_1 und C_2 eines Vektorraumes eine abgeschlossene Hyperebene H zu finden, so dass C_1 'auf der einen Seite von H' liegt und C_2 auf der anderen Seite (Zeichnung.) Nach den vorhergegangenen Betrachtungen laesst sich das auch verstehen als die Frage, ob es eine lineares stetiges φ und $\alpha \in \mathbb{R}$ gibt mit

$$C_1 \subset \{x : \varphi(x) \leq \alpha\} \text{ und } C_2 \subset \{x : \varphi(x) > \alpha\}.$$

Dieses Problem laesst sich nun mit dem Satz von Hahn-Banach unter geeigneter Wahl von p angehen. Tatsaechlich lassen sich die obigen Folgerungen so versehen, dass wir es fuer die beiden Faelle

- $C_1 = \{x\}$, $C_2 = \{y\}$.
- $C_1 = \{x\}$ und $C_2 = U$.

Ende der Vorlesung

Bemerkung. Im allgemeinen gibt es fuer normierte Vektorraeume (anders als fuer Hilbertraeume) keine kanonische Beziehung zwischen dem Raum und seinem Dualraum. Allerdings gibt es immer eine Einbettung des Raumes in den Dualraum des Dualraum. Genauer gilt folgendes: Ist $(X, \|\cdot\|)$ ein normierter Raum, so ist die Abbildung

$$J : X \longrightarrow (X')', J(x)(\varphi) := \varphi(x),$$

isometrisch. Denn es gilt

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

(Dabei nutzen wir eine der Folgerungen aus dem Satz von Hahn-Banach im letzten Schritt.) Ist J surjektiv, so heisst X *reflexiv.* Offenbar ist jeder reflexive normierte Raum vollstaendig (Dualraum ist Banachraum). Offenbar ist jeder endlichdimensionale normierte Raum reflexiv. (Denn: Dualraum hat selbe Dimension wie Ursprungsraum...).

5. Die Dualraeume der L^p-Raeume

In diesem Abschnitt bestimmen wir die Dualraeume der L^p-Raeume (im σ-endlichen Fall).

Sei $(X, \mathcal{A}, \mu)$ ein Massraum und seien $p \in [1, \infty]$ und $q \in [1, \infty]$ mit $1/p + 1/q = 1$ gewaehlt. (Dabei ist wieder der Fall $q = 1, p = \infty$ und $q = \infty, p = 1$ eingeschlossen.) Dann wissen wir schon aus der Hoelder Ungleichung, dass jedes $f \in L^q(X, \mu)$ durch

$$j(f) : L^p(X, \mu) \longrightarrow \mathbb{K},\ j(f)(g) := \int_X fg d\mu,$$

ein lineares Funktional mit $\|j(f)\| \leq \|f\|_q$ definiert. Wir werden nun ein Umkehrung dieses Sachverhaltes zeigen. Diese Umkehrung gilt allgemein. Wir werden sie hier aber nur unter eine zusaetzlichen Voraussetzung zeigen. Genauer werden wir noch voraussetzen, dass der Massraum σ-endlich ist, d.h. dass messbare Mengen $X_n \subset X$ existieren mit $X_n \subset X_{n+1}$ und $\mu(X_n) < \infty$ fuer alle $n \in \mathbb{N}$. Dann wird unser Beweis der Umkehrung aus zwei Schritten bestehen:

- Zu $\varphi \in (L^p)'$ existiert ein messbares f mit
$$(*)\ \ \varphi(g) = \int fgd\mu$$
fuer 'viele' $g \in L^p$.
- Ein messbares f mit $(*)$ gehoert zu L^q und es gilt $(*)$ fuer alle $g \in L^p$.

Die Menge der 'vielen' f wird mit folgendem Raum L_0 gefasst:

$L_0 := \{g : X \longrightarrow \mathbb{C} : g$ messbar, beschraenkt, es gibt $n \in \mathbb{N}$ mit $g(x) = 0$ fuer $x \notin X_n\}$.

(Beachte, dass L_0 von der Auswahl (X_n) abhaengt, ohne dass dies in der Notation sich widerspiegelt.) Der Raum L_0 stellt, wenn man so will, in unserem Kontext ein Analogon dar zum Raum der stetigen Funktionen mit kompaktem Traeger. Offenbar gilt fuer L_0 folgendes:

- Es ist L_0 ein Vektorraum. (Das ist klar.)
- Fuer $1 \leq p \leq \infty$ gilt $L_0 \subset L^p(X, \mu)$. (Das folgt, da die Funktionen aus L_0 beschraenkt sind und auf einer Menge endlichen Masses leben.)
- Fuer $1 \leq p < \infty$ ist L_0 sogar dicht in $L^p(X, m)$. (Wie man sich leicht ueberlegt, ist die lineare Huelle der charakteristischen Funktionen von Mengen mit endlichem Mass dicht in $L^p(X, m)$. Es reicht also zu zeigen, dass jede charakteristische Funktion einer Menge von endlichem Mass approximiert werden kann. Dazu reicht es zu zeigen, dass die charakteristischen Funktionen von Mengen von endlichem Mass in X_n approximiert werden koennen. Das ist aber klar, da diese Funktionen zu L_0 gehoeren.)

Bemerkung. Man beachte, dass L_0 im allgemeinen nicht dicht in L^∞ ist. Das ist ein erstes Signal dafuer, dass fuer $p = \infty$ manches anders ist.

Damit koennen wir jetzt den zweiten Schritt unsere allgemeinen Beweisschemas machen.

LEMMA. *Wir betrachten die vorstehend geschilderte Situation. Sei* $1 \leq p, q \leq \infty$ *mit* $1/p + 1/q = 1$ *gegeben. Sei* $f : X \longrightarrow \mathbb{C}$ *messbar mit*
$$fg \in \mathcal{L}^1(X, \mu) \ \ \text{und} \ \ \left| \int_X f(x)g(x)d\mu(x) \right| \leq C\|g\|_p$$
fuer alle $g \in L_0$. *Dann gehoert* f *zu* $\mathcal{L}^q(X, \mu)$ *mit* $\|f\|_q \leq C$.

Beweis. *Sei* $p = 1$. Angenommen es gilt ess-sup$|f| > C$. Dann existiert also ein $\delta > 0$ so, dass fuer
$$X_\delta := \{x \in X : |f(x)| \geq C + \delta\}$$
gilt $\mu(X_\delta) > 0$. Also existiert dann ein $N \in \mathbb{N}$ mit
$$\mu(X_\delta \cap X_N) > 0.$$
Setze
$$g := \overline{\operatorname{sgn} f} \frac{1}{\mu(X_\delta \cap X_N)} 1_{X_\delta \cap X_N}.$$

Dann gehoert g zu L_0, erfuellt $\|g\|_1 = 1$ und es gilt

$$\left| \int f(x)g(x)d\mu(x) \right| \geq C + \delta = (C+\delta)\|g\|_1 > C\|g\|_1.$$

Das ist ein Widerspruch.

<u>*Sei*</u> $1 < p < \infty$. Wir beschreiben kurz die **Idee** des Beweis: Waehle $g = \overline{\mathrm{sgn} f}|f|^{q-1}$. Dann gilt

$$\int_X fgd\mu = \int_X |f|^q d\mu$$

und das liefert dann die Behauptung. Wir werden das mit einer Modifikationen durchfuehren. Genauer brauchen wir ein $g \in L_0$. Das fuehrt auf eine Approximation. Hier sind die Details: Ohne Einschraenkung sei f nicht identisch 0. Wir definieren fuer $n \in \mathbb{N}$ die 'doppelt abgeschnittene Funktion' durch

$$f_n : X \longrightarrow \mathbb{C};\ f_n(x) := \begin{cases} f(x) & \text{fuer } x \in X_n \text{ mit } |f(x)| \leq n \\ 0 & \textit{sonst.} \end{cases}$$

Dann ist f_n also beschraenkt und verschwindet ausserhalb der Menge X_n, die endliches Mass hat. Damit gehoert f_n also zu allen L^r. Offenbar gilt

- $f_n \in L^q(X,\mu)$,
- $\|f_n\|_q \neq 0$ fuer grosse n.
- $\|f_n\|_q \to \left(\int |f|^q\right)^{1/q} \in [0,\infty]$,

Setze

$$g_n := \overline{\mathrm{sgn} f_n}|f_n|^{q-1}.$$

Dann folgt $g_n \in L_0$. Das liefert dann

$$\begin{aligned} C\|g_n\|_p &\geq \left| \int_X fg_n d\mu \right| \\ &= \left| \int_X f_n g_n d\mu \right| \\ &= \int |f_n|^q d\mu \\ &= \|f_n\|_q^q \end{aligned}$$

fuer alle $n \in \mathbb{N}$. Division durch $\|g_n\|_p$ liefert dann

$$C \geq \frac{\|f_n\|_q^q}{\|g_n\|_p} = \|f_n\|_q$$

fuer alle $n \in \mathbb{N}$. (Zum Beweis der letzten Gleichheit nutzen wir

$$\|g_n\|_p = \left(\int |f_n|^{p(q-1)} d\mu \right)^{1/p} = \|f_n\|_q^{q/p}$$

sowie $q - q/p = 1$.) Aufgrund der oben diskutierten Konvergenz der $\|f_n\|_q$ gegen $\left(\int |f|^q\right)^{1/q}$ folgt dann $f \in \mathcal{L}^q(X,\mu)$ mit $\|f\|_q \leq C$.

Sei $p = \infty$. Wir lassen diesen Fall als Uebungsaufgabe. Wir werden die entsprechende Aussage nicht benoetigen. $\square$

THEOREM. *Sei $(X, \mathcal{A}, \mu)$ ein σ-endlicher Massraum. Sei $1 \leq p < \infty$ und $q \in [1, \infty]$ mit $1/p + 1/q = 1$ gewaehlt (d.h. $q = \infty$ falls $p = 1$). Dann ist die Abbildung*

$$j : L^q(X, \mu) \longrightarrow L^p(X, \mu)',\ f \mapsto j(f),$$

linear, bijektiv und isometrisch (d.h. es existiert zu jedem $\varphi \in L^p(X, \mu)'$ genau ein $f \in L^q(X, \mu)$ mit $\varphi = j(f)$ und es gilt $\|f\|_q = \|\varphi\|$).

Bemerkungen.

- Die Aussage des Satzes gilt genauso auch im nicht σ-endlichen Fall. Der Beweis ist allerdings aufwendiger. Wir fuehren ihn hier nicht.
- Der Satz gilt nicht fuer $p = \infty$. Tatsaechlich ist der Dualraum von L^∞ der Raum der sogenannten *endlich additiven Masse* auf X (vgl. folgender Beweis)
- Hat X endliches Mass und gilt $1 \leq p \leq 2$, so lasst sich die Aussage leicht aus dem Rieszschen Lemma und dem vorigen Lemma folgern. In diesem Fall ist naemlich $\kappa : L^2 \longrightarrow L^p, f \mapsto f$, stetig. Damit ist dann fuer jedes $\varphi \in (L^p)'$ auch $\varphi \circ \kappa$ stetig. Damit gibt es dann nach dem Riesschen Lemma ein $f \in L^2$ mit

  $$\varphi(g) = \varphi(\kappa(g)) = \int fg d\mu$$

  fuer alle $g \in L^2 \subset L^p$. Mit dem vorigen Lemma folgt dann die gewuenschte Aussage.

Beweis. Wir zeigen eine Reihe von Aussagen. Der Knackpunkt ist aber die Surjektivitaet der Abbildung j.

Es ist j eine Isometrie (d.h. es gilt $\|f\|_q = \|j(f)\|$). Nach Hoelder Ungleichung gilt

$$|j(f)(g)| \leq \|f\|_q \|g\|_p$$

fuer alle $g \in L^p(X, \mu)$. Damit folgt also $\|j(f)\| \leq \|f\|_q$. Umgekehrt gilt natuerlich

$$|\int_X fg d\mu| = |j(f)(g)| \leq \|j(f)\| \|g\|_p$$

fuer alle $g \in L_0 \subset L^p(X, \mu)$. Damit folgt aus dem vorangehenden Lemma dann $\|f\|_q \leq \|j(f)\|$. Insgesamt zeigt dies die gewuenschte Isometrie.

Es ist j injektiv. Da j eine Isometrie ist, ist es injektiv.

Surjektivitaet von j.
Seien messbare Mengen $X_n \subset X$ gewaehlt mit $X_n \subset X_{n+1}$ und $\mu(X_n) < \infty$ fuer alle $n \in \mathbb{N}$.

Fixiere zunaechst ein $n \in \mathbb{N}$: Dann gehoert fuer jedes messbare $A \subset X_n$ die Funktion 1_A zu L^p. Damit kann man dann definieren

$$\nu(A) := \varphi(1_A).$$

Behauptung. Es ist ν ein komplexes Mass auf X_n, d.h. es gilt $\nu(\cup_n A_n) = \sum_n \nu(A_n)$ mit absolut konvergenter Summe fuer eine Folge paarweise disjunkter A_n.
Beweis der Behauptung: Sei $A = \cup_j A_j$ mit disjunkten messbaren Teilmengen von X_n, so konvergiert nach dem Satz ueber die monotone Konvergenz oder

dem Satz ueber dominierte Konvergenz $\sum_{j=1}^N 1_{A_j}$ gegen 1_A in L^p. (Hier verwenden wir $1 \leq p < \infty$.) Damit folgt aufgrund der Stetigkeit von φ dann

$$\begin{aligned} \nu(A) &= \varphi(1_A) \\ &= \varphi(\sum_j 1_{A_j}) \\ \text{(Konvergenz, Stetigkeit } \varphi) \quad &= \sum_j \varphi(1_{A_j}) \\ &= \sum_j \nu(A_j). \end{aligned}$$

Das beendet den Beweis der Behauptung.

Behauptung. Es ist ν absolut stetig bzgl. μ, d.h. aus $\mu(A) = 0$ folgt $\nu(A) = 0$. Beweis der Behauptung. Gilt $\mu(N) = 0$ fuer ein messbares $N \subset X_n$, so gilt $1_N = 0$ in L^p.

Aus dem komplexen Satz von Radon Nikodym ergibt sich nun aus den beiden Behauptungen sofort die Existenz eines $f_n \in \mathcal{L}^1(X_n, \mu)$ mit

$$\varphi(1_A) = \nu(A) = \int 1_A f_n d\mu$$

fuer jede messbare Teilmenge A von X_n. Damit folgt dann

$$\varphi(g) = \int f_n g d\mu$$

fuer jedes $g \in \mathcal{L}^\infty(X_n, \mu)$, da man solche g gleichmaessig, also auch in L^p, durch Linearkombinationen von charakteristischen Funktionen approximieren kann. (Unter einer solchen Approximation konvergiert die Linke Seite, da φ stetig ist und die Approximation in L^p erfolgt und es konvergert die rechte Seite da f_n zu L^1 gehoert und die Approximation gleichmaessig erfolgt.) Daraus ergibt sich, dass fuer $n \leq m$ gilt

$$f_n = f_m \quad \mu \text{ fast ueberall auf } X_n.$$

Damit koennen wir dann die f_n zusammensetzen zu einer Funktion

$$f : X \longrightarrow \mathbb{C} \text{ mit } f = f_n \text{ auf } X_n \ \mu \text{ fast ueberall.}$$

Dann gilt fuer $g \in L_0$ nach Konstruktion also

$$\int f g d\mu = \varphi(g).$$

Das liefert

$$\left| \int f g d\mu \right| = |\varphi(g)| \leq \|\varphi\| \|g\|_p$$

fuer alle $g \in L_0$. Nach dem vorangehenden Lemma gilt dann $f \in L^p(X, \mu)$ mit $\|f\|_p \leq \|\varphi\|$. Damit existiert dann $j(f)$ und stimmt mit φ auf L_0 ueberein. Da L_0 dicht in L^p ist, folgt

$$j(f) = \varphi$$

und der Satz ist bewiesen. □

Ende der Vorlesung

KAPITEL 6

Der Satz von Baire

In diesem Abschnitt lernen wir ein grundlegendes Resultat fuer vollstaendige metrische Raeume kennen. Wir werden es hauptsaechlich auf geeigneten Vektoräumen anwenden. Aber das Resultat verlangt keine Vektorraumstruktur! Es gibt zwei (aequivalente) Varianten des Satzes und wir werden beide diskutieren.

Erinnerung. Eine Menge A in einem metrischen Raum (M, d) ist dicht, wenn $\overline{A} = M$ gilt. Das ist genau dann der Fall, wenn A jede offene Menge in M schneidet.

THEOREM (Satz von Baire). *Sei (M, d) ein vollstaendiger metrischer Raum. Seien U_n, $n \in \mathbb{N}$, offene dichte Mengen in M. Dann ist $\bigcap_n U_n$ dicht.*

Bemerkung. Der Satz liefert eine sehr bemerkenswerte Aussage. Zum Beispiel ist $\mathbb{Q} \cap (\mathbb{Q} + r) = \emptyset$, wenn r irrational ist. Der Schnitt von dichten Mengen ist also im allgemeinen leer und damit gar nicht dicht. Wahlt man aber umgekehrt $\mathbb{Q} = \{q_j : j \in \mathbb{N}\}$ eine Abzaehlung von $\mathbb{Q}$ und definiert fuer $\varepsilon > 0$

$$Q^\varepsilon := \bigcup_j (q_j - \frac{\varepsilon}{2^j}, q_j + \frac{\varepsilon}{2^j})$$

so ist Q^ε offen und dicht (und hat Lebesguemass kleiner als 2ε). Aber es ist (trotz des kleinen Lebesguemass) nach dem Satz von Baire

$$\bigcap_n (\mathbb{Q}^{\varepsilon_n} + r_n)$$

dicht fuer alle Folgen (r_n), (ε_n) in $\mathbb{R}$ mit $\varepsilon_n > 0$ fuer alle $n \in \mathbb{N}$.

Beweis. Sei $D := \cap U_n$. Wir zeigen, dass D jede offene nichtleere Menge schneidet. Sei $V = V_0$ eine beliebige offene nichtleere Menge. Da U_1 dicht ist, gilt dann

$$V \cap U_1 \neq \emptyset.$$

Damit koennen wir ein $x_1 \in M$ und $\varepsilon_1 > 0$ finden mit

$$B_{\varepsilon_1}(x_1) \subset V \cap U_1.$$

Da U_2 dicht ist, gilt natuerlich

$$U_{\varepsilon_1}(x_1) \cap U_2 \neq \emptyset.$$

Induktiv koennen wir dann $x_n \in M$ und $0 < \varepsilon_n < \varepsilon_{n-1}/2$ finden mit

$$B_{\varepsilon_n}(x_n) \subset U_{\varepsilon_{n-1}}(x_{n-1}) \cap U_n$$

fuer alle $n \in \mathbb{N}$. Nach Konstruktion und wegen der Vollstaendigkeit von (M, d) ziehen die Kugelen $B_{\varepsilon_n}(x_n)$ sich dann zu einem Punkt x zusammen.

Dieser Punkt gehoert sowohl zu V als auch zu jedem U_n. Hier sind die Detaisl: Es gilt nach Konstruktion

- $\varepsilon_{n+1} \leq \frac{\varepsilon_1}{2^n}$ fuer alle $n \in \mathbb{N}$.
- $B_{\varepsilon_{n+1}}(x_{n+1}) \subset B_{\varepsilon_n}(x_n) \subset ... \subset B_1(x_1) \subset V$ fuer alle $n \in \mathbb{N}$.
- $B_{\varepsilon_n}(x_n) \subset U_n$ fuer alle $k, n \in \mathbb{N}$.

Aus den ersten beiden Punkten folgt sofort, dass die (x_n) eine Cauchy Folge bilden, deren Grenzwert (nach dem zweiten Punkt) dann zu allen $B_{\varepsilon_n}(x_n)$ und damit auch zu V gehoert. Weiterhin gehoert der Grenzwert nach dem dritten Punkt aber auch zu jedem U_n damit zu D gehoert. Damit haben also D und V einen nichtleere Schnittmenge. □

Wir lernen nun noch eine Variante des Satz von Baire kennen. Dabei handelt es sich um 'durch Komplementbildung' entstehende Aussagen. Dabei sind die folgenden Zusammenhaenge zwischen einer Menge U und ihrem Komplement F grundlegend:

- Es ist U offen genau dann wenn F abgeschlossen ist.
- Es ist U dicht genau dann, wenn F leeres Inneres hat.

(*Erinnerung.* In einem metrischen Raum (M, d) ist das Innere A° einer Menge A definiert durch

$$A^\circ = \{x \in A : \text{es existiert ein } r > 0 \text{ mit } U_r(x) \subset A\}.$$

Damit folgt dann der erste Punkt direkt aus der Definition von Offenheit und Abgeschlossenheit. Zweiter Punkt U dicht $\Longleftrightarrow$ U schneidet jede offene Kugel $\Longleftrightarrow$ keine offene Kugel ist in F enthalten $\Longleftrightarrow$ F hat leeres Inneres.)

Theorem (Variante Satz von Baire). *Sei (M, d) ein nichtleerer vollstaendiger metrischer Raum. Seien F_n, $n \in \mathbb{N}$, abgeschlossene Teilmengen von M. Gilt*

$$M = \bigcup_n F_n,$$

so existiert ein $N \in \mathbb{N}$ mit $F_N^\circ \neq \emptyset$.

Beweis. Angenommen nein: Dann ist als F_n^c dicht fuer jedes $n \in \mathbb{N}$ (da F_n leeres Inneres hat). Weiterhin ist F_n^c natuerlich offen fuer jedes $n \in \mathbb{N}$ (da F_n abgeschlossen ist). Damit folgt aus dem schon bewiesenen Satz von Baire dann, dass

$$\emptyset = M^c = \left(\bigcup_n F_n\right)^c = \bigcap_n F_n^c$$

dicht in M ist. Das ist ein Widerspruch. □

Bemerkung. (a) Wir haben die Variante aus der zunaechst gegebenen Form hergeleitet. Mit ein wenig Aufwand kann man auch (wieder mit Komplementbildung) umgekehrt vorgehen.
(Bew: Sei $G := \cap U_n$. Zu zeigen G ist dicht, d.h. $M \setminus G$ enthaelt keine offene Menge.
Angenommen: $U_r(x) \subset M \setminus G$ fuer ein $x \in M$ und $r > 0$. Dann gelte ohne Einschraenkung (sonst r verkleinern) $\overline{U_r(x)} \subset M \setminus G$.

Sei $\widetilde{M} := \overline{U_r(x)}$. Dann ist $(\widetilde{M}, d)$ ein vollstaendiger metrischer Raum und es gilt

$$\widetilde{M} = (M \setminus G) \cap \widetilde{M} = \bigcup_n (M \setminus U_n) \cap \widetilde{M}.$$

Als Schnitt zweier abgeschlossener Teilmengen von M ist $(M \setminus U_n) \cap \widetilde{M}$ abgeschlossen in M und damit auch in $\widetilde{M}$. Also existiert nach dem Satz von Baire ein $\delta > 0$ und $y \in \widetilde{M}$ und $N \in \mathbb{N}$ mit

$$U_\delta^{\widetilde{M}}(y) \subset (M \setminus U_N) \cap \widetilde{M}.$$

(Hier bezeichnet $U_\delta^{\widetilde{M}}(y)$ die Kugel in $\widetilde{M}$.) Dann erfuellt die offene und nichtleere (da $y \in \overline{U_r(x)}$) Menge $U_\delta(y) \cap U_r(x)$ die Inklusion

$$U_\delta(y) \cap U_r(x) \subset U_\delta(y) \cap \widetilde{M} = U_\delta^{\widetilde{M}}(y) \subset M \setminus U_N.$$

Damit enthaelt also $M \setminus U_N$ eine offene Menge. Das ist ein Widerspruch zur Dichtheit von U_N.)

(b) Die Variante kann auch mit einem aehnlichen Schluss wie die zunaechst gegebene Version hergeleitet werden.
(Widerspruchsbeweis: Angenommen $F_n^\circ = \emptyset$ fuer alle $n \in \mathbb{N}$. Dann gilt fuer alle $x \in M$ und alle $\varepsilon > 0$ und $n \in \mathbb{N}$

$$U_\varepsilon(x) \cap F_n^c \neq \emptyset. \ (*)$$

Wir konstruieren nun eine Cauchy Folge ohne Grenzwert. Dazu werden wir induktiv $x_n \in M$ und $\varepsilon_n > 0$ konstruieren mit

- $B_{\varepsilon_n}(x_n) \subset B_{\varepsilon_m}(x_m)$ fuer alle $n \geq m$,
- $\varepsilon_n \to 0$, $n \to \infty$,
- $B_{\varepsilon_n}(x_n) \subset F_n^c$.

Dann bilden die (x_n) eine Cauchy-Folge aufgrund der ersten beiden Punkte. Aufgrund der Vollstaeendigkeit des Raumes konvergiert diese Cauchy-Folge. Der Grenzwert x muss dann in jedem $B_{\varepsilon_n}(x_n)$ liegen und gehoert damit zu keinem F_n. Das ist ein Widerspruch zu $M = \cup_n F_n$.

Hier sind die Details: Sei $x_0 \in M$ beliebig und $\varepsilon_0 = 1$. Nach $(*)$ existiert ein

$$x_1 \in F_1^c \cap U_{\varepsilon_0}(x_0).$$

Da F_1^c und U_{ε_0} offen sind, gibt es $0 < \varepsilon_1$ mit

$$U_{\varepsilon_1}(x_1) \subset B_{\varepsilon_1}(x_1) \subset F_1^c \cap U_{\varepsilon_0}(x_0).$$

und $\varepsilon_1 \leq \varepsilon_0/2$. Mit diesem Schluss koennen wir dann induktiv koennen unter Verwendung von $(*)$ eine Folge (x_n) in M und $\varepsilon_n > 0$ finden mit

$$U_{\varepsilon_n}(x_n) \subset B_{\varepsilon_n}(x_n) \subset F_n^c \cap U_{\varepsilon_{n-1}}(x_{n-1})$$

und $\varepsilon_n \leq \varepsilon_{n-1}/2$.
(Induktionsschritt: Seien $x_0, \ldots, x_n$ mit diesen Eigenschaften schon konstruiert. Dann gilt aufgrund von $(*)$ natuerlich

$$U_{\varepsilon_n}(x_n) \cap F_{n+1}^c \neq \emptyset.$$

Da $U_{\varepsilon_n}(x_n)$ und F_{n+1}^c offen sind, gibt es dann also x_{n+1} und $0 < \varepsilon_{n+1} < \varepsilon_n/2$ mit

$$U_{\varepsilon_{n+1}}(x_{n+1} \subset B_{\varepsilon_{n+1}}(x_{n+1}) \subset F_{n+1}^c \cap U_{\varepsilon_n}(x_n)$$

Das beendet den Induktionsbeweis.)

Zeichnung.

Also folgt fuer $k \leq m$

$$B_{\varepsilon_m}(x_m) \subset B_{\varepsilon_{m-1}}(x_{m-1}) \subset \ldots \subset B_{\varepsilon_{k+1}}(x_{k+1}) \subset B_{\varepsilon_k}(x_k).$$

Insbesondere folgt

$$d(x_k, x_m) \leq \varepsilon_k \leq \frac{1}{2^k}.$$

Damit ist (x_k) eine Cauchy Folge. Aufgrund der Vollstaendigkeit (!) existiert dann

$$x = \lim_{n\to\infty} x_n.$$

Da die Folge (x_n) fuer alle grossen Indices vollstaendig in der abgeschlossenen Kugel $B_{\varepsilon_k}(x_k)$ enthalten ist, gilt dies auch fuer den Grenzwert. (Denn: $d(x, x_k) = \lim_{n\to\infty} d(x_k, x_n) \leq \varepsilon_k$.) Es folgt also

$$x \in B_{\varepsilon_k}(x_k) \subset F_k^c$$

fuer alle $k \in \mathbb{N}$. Das liefert

$$x \notin \bigcup F_n = M.$$

Dieser Widerspruch beweist die Behauptung.)

Bemerkungen. Die beiden Vorausetzungen im Satz sind noetig.

- Die Vollstaendigkeit ist eine wesentliche Voraussetzung im Satz. So kann man etwa den metrischen Raum $\mathbb{Q}$ mit der Euklidischen Metrik via
$$\mathbb{Q} = \bigcup_{q\in\mathbb{Q}} \{q\}$$
als abzaehlbare Vereinigung abgeschlossener Mengen mit leerem Inneren darstellen bzw. es ist $G_q := \{q\}^c$ offen und dicht in $\mathbb{Q}$ mit
$$\bigcap_q G_q = \emptyset.$$
- Die Voraussetzung der Abzaehlbarkeit der F_n ist ebenfalls noetig. Denn man hat natuerlich immer $M = \cup_{m\in M}\{m\}$ bzw. $\cap\{m\}^c = \emptyset$. Fuer $M = \mathbb{R}$ mit der Euklidischen Metrik, liefert das die Notwendigkeit der Abzaehlbarkeitsannahme.

Der Satz von Baire legt es nahe, abzaehlbare Schnitte offener Mengen zu untersuchen.

Definition. *Sei (M, d) ein vollstaendiger metrischer Raum. Eine Teilmenge B von M heisst G_δ-Menge, wenn sie ein abzaehlbarer Schnitt von offenen Mengen ist. Eine Teilmenge D heisst F_σ-Menge, wenn sie eine abzaehlbare Vereinigung von abgeschlossenen Mengen ist.*

Bemerkung. Offenheit ist mit endlichen Schnitten vertraeglich und Abgeschlossenheit mit endlichen Vereinigungen. In diesem Sinne sind die abzaehlbaren Operationen das naechstbeste.

Damit erhalten wir als Folgerung aus dem Satz von Baire noch folgende Aussage (aus der offenbar auch direkt der Satz von Baire folgt).

FOLGERUNG (Abzaehlbare Schnitte dichter G_δ-Mengen ist dichte G_δ-Menge). *Sei (M,d) ein vollstaendiger metrischer Raum. Dann ist ein abzaehlbarer Schnitt von dichten G_δ-Mengen wieder ein dichte G_δ-Menge.*

Beweis. Offenbar ist ein abzaehlbarer Schnitt von G_δ Mengen wieder eine G_δ Menge. Er ist dicht nach einer Folgerung aus dem Satz von Baire. □

Interpretation. Es gibt verschiedene Arten die Groesse einer Menge zu messen. Im topologischen Kontext gelten oft dichte G_δ-Mengen als gross. Ihre Elemente werden dann als *typisch* oder *generisch* bezeichnet. Damit spielen dichte G_δ-Mengen in der Topologie eine aehnliche Rolle wie Mengen von vollem Mass in der Masstheorie. So sind ja auch Mengen von vollem Mass abgeschlossen unter abzaehlbaren Schnitten. Es stellt sich dann allerdings heraus, dass typische Elemente oft im Sinne der naheligenden Beispiele recht untypische Eigenschaften haben :-)

Wir diskutieren nun einige prominente G_δ-Mengen und allgemeine Eigenschaften von G_δ-Mengen.

Beispiele. (a) Die irrationalen Zahlen sind eine dichte G_δ-Menge in $\mathbb{R}$ (mit der Euklidischen Metrik).

Bew. Sei q_n, $n \in \mathbb{N}$ eine Abzaehlung von $\mathbb{Q}$. Sei zu jedem $n \in \mathbb{N}$ die Menge

$$U_n := \mathbb{R} \setminus \{q_n\}.$$

Dann ist jedes U_n dicht und offen und die Menge der irrationalen Zahlen gerade der Schnitt ueber alle U_n, $n \in \mathbb{N}$.

(b) Die Menge der nirgends differenzierbaren stetigen Funktionen auf $[0,1]$ enthaelt eine dichte G_δ-Menge (im Raum der stetigen Funktionen auf $[0,1]$ mit der Supremumnorm).

Bew. Hier geben wir nur eine Skizze: Eine wichtige Idee ist es, geeignete Funktionen mit vielen 'Zacken' zu konstruieren: Sei I ein Intervall mit nichtleerem Inneren in $[0,1]$ und $C > 0$ und $U_{I,C}$ die Menge der f fuer die fuer KEIN $x \in I$ gilt

$$|f(x) - f(y)| \leq C|x-y| \text{ fuer alle } y \in I.$$

Dann ist $U_{I,C}$ eine dichte offene Menge. (Offen:klar. Dicht: Wahle Funktionen mit vielen kleinen 'Zacken'.) Die Menge der nirgends diffbaren Funktionen ist dann gegeben durch

$$\bigcap_{I \text{mit rationalen Endpunkten}, C \in \mathbb{N}} U_{I,C}.$$

Das beendet den Beweis.

(c) Sei (f_n) eine Folge von stetigen Funktionen auf dem vollstaendigen metrischen Raum (M,d), die punktweise gegen die Funktion f konvergieren. Dann ist die Menge der Stetigkeitspunkte von f eine dichte G_δ-Menge.

Bew. Wir behandeln das in der Uebung. Hier geben wir eine Skizze. Definiere zu $\varepsilon > 0$ die Menge U_ε als die Menge der $x \in M$ fuer die ein $\delta > 0$ und ein $N \in \mathbb{N}$ existieren mit

$$|f_N(y) - f(y)| \leq \varepsilon$$

fuer alle $y \in U_\delta(x)$. Nach Konstruktion ist U_ε offen (Warum?). Weiterhin ist U_ε dicht. (Um das zu sehen, wenden wir den Satz von Baire auf eine beliebige abgeschlossene Kugel B mit nichtleerem Inneren und die abgeschlossenen Mengen

$$F_n := \{x \in B : |f_n(x) - f_m(x)| \leq \varepsilon \text{ fuer alle } m \geq n\}$$

an.) Die Menge der Stetigkeitspunkte von M ist aber $\bigcap_n U_{1/n}$.

(d) Die oben diskutierte 'Aufdickung' von $\mathbb{Q}$. ($\mathbb{Q}$ selber ist keine G_δ-Menge, denn es ist dicht und Komplement ist G_δ-Menge).

Bemerkung. Die Beispiele zeigen, dass es oft sehr schwer sein kann, Elemente von G_δ-Mengen konkret anzugeben (auch wenn diese Elemente gerade die typische Eigenschaften haben).

Wir kommen nun zu allgemeinen Eigenschaften von G_δ-Mengen.

- Das Urbild einer G_δ-Menge unter einer stetigen Funktion ist wieder eine G_δ-Menge.
- Der Schnitt einer G_δ-Menge mit einer offenen Menge ist eine G_δ-Menge.
- Ist G eine dichte G_δ-Menge, so ist ihr Komplement keine dichte G_δ-Menge (wenn der Raum vollstaendig ist). Bew. ;-)
- Sei (M, d) vollstaendiger metrischer Raum.
  - Ist M abzaehlbar, so liegt mindestens ein Punkt aus M diskret. Bew. $M = \cup_{m\in M}\{m\}$. Nun Anwenden des Satz von Baire.
  - Hat M keine diskreten Punkte, so ist M ueberabzaehlbar. Bew. vorige Aussage.
  - Hat M keine diskreten Punkte, so ist jede dichte G_δ-Menge in M ueberabzaehlbar.
    Bew. Angenommen $G = \{g_k : k \in \mathbb{N}\}$ G_δ-Menge. Sei
    $$G_k := G \setminus \{g_k\} := G \cap \{g_k\}^c.$$
    Dann ist G_k dicht (da M keine diskreten Punkte hat) und eine G_δ Menge (als Schnitt einer offenen und einer G_δ-Menge). Daher ist dann auch
    $$\cap G_k = \emptyset$$
    eine dichte Menge. Widerspruch.

Wir beenden den Abschnitt mit einer Diskussion einer (frueher) ueblichen Formulierung des Satzes von Baire. Eine Menge A des metrischen Raumes (M, d) heisst nirgends dicht, wenn $(\overline{A})^\circ = \emptyset$. Eine Menge D heisst von erster Kategorie, wenn sie eine abzaehlbare Vereinigung nirgends dichter Mengen ist. Eine Menge heisst von zweiter Kategorie, wenn sie nicht von erster Kategorie ist. Damit laesst sich der Satz so umformulieren:

Bairescher Kategoriensatz. Jeder vollstaendige metrische Raum ist von zweiter Kategorie in sich selbst.

Bemerkung. Fuer lokalkompakte hausdorff Raeume gilt der Satz von Baire ebenfalls mit (im wesentlichen) dem selben Beweis, den wir oben gegeben haben. Wir verzichten auf weitere Diskussion (da wir nicht einmal das Konzept des lokalkompakten Raumes hier eingefuehrt haben).

Ende der Vorlesung

KAPITEL 7

Anwendungen des Satz von Baire auf beschraenkte Operatoren

Es gibt vier grosse Anwendungen des Satz von Baire in der Operatortheorie. Diese sind

- Der Satz von Banach / Steinhaus auch bekannt als Prinzip der gleichmaessigen Beschraenktheit.
- Der Satz von der offenen Abbildung.
- Der Satz von der stetigen Inversen.
- Der Satz vom abgeschlossenen Graphen.

Die ersten drei Anwendungen setzen Stetigkeit der involvierten Operatoren voraus. Sie werden in diesem Kapitel behandelt. Die vierte Anwendung wird im kommenden Kapitel behandelt.

THEOREM (Satz von Banach Steinhaus / Prinzip der gleichmaessigen Beschraenktheit). *Sei $(E, \|\cdot\|)$ eine Banachraum und $(F, \|\cdot\|_F)$ ein normierter Raum. Seien $T_\alpha, \alpha \in A$, beschraenkte Operatoren von E nach F. Ist T_α, $\alpha \in A$, punktweise beschraenkt (d.h. es gilt fuer jedes $x \in E$ $\sup_{\alpha\in A} \|T_\alpha x\|_E < \infty$), so ist T_α gleichmaessig beschraenkt (d.h. es gilt $\sup_{\alpha\in A} \|T_\alpha\| < \infty$).*

Beweis. Wir beweisen den Satz in zwei Schritten.

Schritt 1. Es existiert eine Kugel $U_r(p)$ mit $\|T_\alpha x\| \le C_0$ fuer alle $x \in U_r(p)$.
Bew. Sei

$$F_k := \{x \in E : \|T_\alpha x\| \le k \text{ fuer alle } \alpha \in A\} = \bigcap_{\alpha\in A} \{x \in E : \|T_\alpha x\| \le k\}.$$

Dann ist F_k abgeschlossen als Schnitt abgeschlossener Mengen (da jedes T_α stetig ist). Wegen der punktweisen Beschraenktheit gilt weiterhin

$$E = \bigcup_k F_k.$$

Damit folgt nach dem Satz von Baire die Aussage.

Schritt 2. Es existiert ein $C > 0$ mit $\|T_\alpha\| \le C$ fuer alle $\alpha \in A$.
Bew. Fuer $\|x\| < r$ gilt

$$\|T_\alpha x\| \le \|T_\alpha(x + p)\| + \|T_\alpha p\| \le C_0 + C_p.$$

Damit folgt die Aussage mit $C := \frac{C_0+C_p}{r}$. □

Anwendung. Sei H ein Hilbertraum und (x_n) eine Folge in H, die schwach gegen $x \in H$ konvergiert, d.h.

$$\langle x_n, y\rangle \to \langle x, y\rangle$$

fuer alle $y \in H$ erfuellt. Dann gilt $\sup_n \|x_n\| < \infty$.

Bew. Betrachte die durch x_n induzierten Funktionale

$$\varphi_n : H \longrightarrow \mathbb{K},\ \varphi_n(y) = \langle x_n, y\rangle.$$

Dann sind die φ_n, $n \in \mathbb{N}$, punktweise beschraenkt (da konvergent) und da H vollstaendig ist, folgt aus dem vorigen Satz dann sofort

$$\sup_n \|\varphi\| < \infty.$$

Mit $\|\varphi_n\| = \|x_n\|$ (s.o.) folgt dann die gewuenschte Aussage. Das beendet den Beweis.

Theorem (Satz von der offenen Abbildung). *Seien E, F Banachraeume und $T : E \longrightarrow F$ eine linearer beschraenkter Operator. Ist T surjektiv, so ist T offen (d.h. fuer jede offenen Menge $U \subset E$ ist auch TU offen in F).*

Beweis. Seien U_r bzw. V_r die offenen Kugeln um den Ursprung vom Radius r in E bzw. F.

Schritt 1. Fuer jedes $r > 0$ existiert ein $s > 0$ mit $V_s \subset \overline{TU_r}$.
Bew. Sei $\delta := r/2$. Dann gilt

$$E = \cup_n U_{n\delta} = \cup_n n U_\delta.$$

Damit folgt aufgrund der Surjektivitaet

$$F = TE = \cup_n nTU_\delta = \cup_N n\overline{TU_\delta}.$$

Nach dem Satz von Baire enthaelt dann eine der Mengen $n\overline{TU_\delta}$ eine offene Kugel. Dann gibt es also (Skalieren) eine offenen Kugel V mit

$$V \subset \overline{TU_\delta}.$$

(Diese Kugel liegt unter Umstaenden noch an der falschen Stelle). Wegen $\delta = r/2$ gilt $U_r \supset U_\delta - U_\delta$, also

$$TU_r \supset TU_\delta - TU_\delta.$$

Damit folgt

$$\overline{TU_r} \supset \overline{TU_\delta - TU_\delta} \supset \overline{TU_\delta} - \overline{TU_\delta} \supset V - V.$$

Die letzte Menge ist offen (als Vereinigung offener Mengen) und enthaelt den Ursprung.

Schritt 2. Fuer jedes $r > 0$ enthaelt TU_r eine Kugel um den Ursprung.
Bew. Sei $r_0 := r/2$. Sei s_0 nach Schritt 1 mit$V_{s_0} \subset \overline{TU_{r_0}}$ gewaehlt. Wir zeigen

$$V_{s_0} \subset TU_{2r_0} = TU_r.$$

Waehle dazu fuer jedes $n \in \mathbb{N}$ e3in $r_n > 0$ sodass gilt

$$\sum_n r_n < r_0$$

und waehle nach Schritt 1 nun $s_n > 0$ mit

$$V_{s_n} \subset \overline{TU_{r_n}}.$$

Ohne Einschraenkung

$$s_n \to 0$$

(sonst Verkleinern). Sei nun $y \in V_{s_0}$ beliebig. Dann gehoert also y zu $\overline{TU_{r_0}}$. Daher gibt es $x_0 \in U_{r_0}$ mit

$$\|y - Tx_0\| < s_1$$

also

$$y - Tx_0 \in V_{s_1}.$$

Wegen $V_{s_1} \subset \overline{TU_{r_1}}$ existiert dann $x_1 \in U_{r_1}$ mit

$$\|(y - Tx_0) - Tx_1\| < s_2,$$

also

$$y - Tx_0 - Tx_1 \in V_{s_2}.$$

Induktiv finden wir dann eine Folge (x_n) mit

- $x_n \in U_{r_n}$,
- $\|y - T(\sum_{k=0}^{n} x_k)\| < s_{n+1}$.

Wegen $\|x_k\| \leq r_k$ und

$$\sum_{k=0}^{\infty} r_k < 2r_0 < \infty$$

und da E ein Banachraum ist, existiert

$$x := \lim_{N\to\infty} \sum_{k=0}^{N} x_k$$

und erfuellt

$$\|x\| < r_0 + \sum_{k=1}^{\infty} r_k = 2r_0 = r$$

sowie

$$\|y - Tx\| = \lim_{N} \|y - T\sum_{k=0}^{N} x_k\| \leq \limsup s_{N+1} = 0.$$

Es ist also $y = Tx$ mit $x \in U_r$.

Schritt 3: Es ist T offen. Sei $W \subset E$ offen und $x \in W$ und $y = Tx$. Dann existiert eine offene Kugel U um den Ursprung in E mit $x + U \subset W$. Nach Schritt 2 existiert eine offene Kugel V um 0 in F mit $V \subset TU$. Damit gilt

$$y + V \subset y + TU = Tx + TU = T(x + U) \subset TW.$$

Das beendet den Beweis. □

Theorem (Satz von der stetigen Inversen). *Seien E, F Banachraeume und $T : E \longrightarrow F$ linear, beschraenkt und bijektiv. Dann ist T^{-1} stetig.*

Beweis. Sei $S := T^{-1}$. Nach dem vorigen Satz ist $T = S^{-1}$ offen. Damit ist fuer jede offene Menge U in E auch $S^{-1}U$ offen und die Stetigkeit von S folgt. □

Bemerkung. (Uebung) Aus dem Satz von der stetigen Inversen kann man auch den Satz von der offenen Abbildung herleiten, indem man zu Quotienteraeumen uebergeht. (Evtl. Details).

Anwendung. Sei H ein Hilbertraum und $T : H \longrightarrow H$ beschraenkt und linear. Dann gilt fuer die Resolventenmenge $\varrho(H)$ dann

$$\varrho(H) = \{z \in \mathbb{K} : T - zI \text{ bijektiv}\}.$$

(Die in der Definition urspruenglich noch geforderte Stetigkeit der Inversen folgt aus dem Satz von der stetigen Inversen und ist also automatisch erfuellt.
Bemerkung. Eine weitere Anwendung des Satz von Baire betrifft die Anzahl der Elemente einer algebraischen Basis eines normierten vollstaendigen Raumes E unendlicher Dimension. Eine solche Basis muss ueberabzaehlbar viele Elemente haben. (Beweis: Angenommen es gibt eine Basis $\{e_1, \ldots, \}$ mit abzaehlbar vielen Elementen. Dann sind die Unterraume $F_n := \mathrm{Lin}\{e_1, \ldots, e_n\}$ endlichdimensional und damit abgeschlossen. Weiterhin gilt (Basis!) dann $\cup_n F_n = E$. Damit muss nach dem Satz von Baire dann eines der F_n eine offene Menge enthalten. Daraus folgt dann (wie?), dass dieses F_n dann schon der ganze Raum ist.

Ende der Vorlesung

KAPITEL 8

Elemente der Theorie unbeschraenkter Operatoren

Ein wichtiger Operator ist der Laplace

$$\Delta f = \sum_{j=1}^{N} \partial_j^2 f$$

fuer $f \in C_c^\infty(\mathbb{R}^N)$. Will man bei seiner Untersuchung auf die starken Hilbertraummethoden zurueckgreifen, stellen sich zwei Probleme:

- der Operator ist nicht auf dem ganzen Hilbertraum definiert;
- der Operator ist nicht beschraenkt.

Diese beiden Probleme haengen (wie wir sehen werden) eng miteinander zusammen. Um mit ihnen umzugehen, fuehren wir ein neues Konzept ein, das Konzept der 'Abgeschlossenheit eines Operators'. Auf abgeschlossene Operatoren kann man wesentliche Teile der Theorie der beschraenkten, ueberall definierte Operatoren verallgemeineren. Ein geignet definierter Laplaceoperator ist abgeschlossen (und kann dann gut behandelt werden).

1. Abgeschlossene Operatoren

In diesem Abschnitt stellen wir das Konzept der Abgeschlossenheit eines Operators vor und beweisen einen grundlegenden Satz dazu, den Satz vom abgeschlossenen Graphen.

Definition (Operator). *Seien E, F Vektorraeume. Ein linearer Operator von E nach F ist eine lineare Abbildung $T : D(T) \longrightarrow F$, wobei $D(T)$ ein Unterraum von E ist. Es heisst $D(T)$ der Definitionsbereich von T und*

$$G(T) := \{(x, Tx) : x \in D(T)\}$$

der Graph von T.

Fuer die weiteren Untersuchungen ist es nuetzlich, sich klar zu machen, dass Operatoren von E nach F via Bildung des Graphen gerade Unterraeumen G von $E \times F$ mit der Eigenschaft

$$(0, y) \in G \Longrightarrow y = 0$$

entsprechen. (Genauer gilt folgendes: Sind E und F Vektorraeume, so ist fuer jeden Operator T von E nach F der Graph $G(T)$ ein Unterraum G von $E \times F$ mit dieser Eigenschaft. Umgekehrt ist jeder Unterraum G von $E \times F$ mit dieser Eigenschaft der Graph eines eindeutigen Operator T. Dieser ist gegeben durch

$$D(T) : \{x \in E : \text{es existiert } y \in F \text{ mit } (x, y) \in G\ \}$$

$$Tx = y.$$

Hierbei ist T wohldefiniert aufgrund der Vorrausetzung, wie man leicht nachrechnen kann.)

Beispiele allgemeiner Operatoren.

- Laplaceoperator auf $L^2(\mathbb{R}^N, \lambda)$:

$$\Delta : C_c^\infty(\mathbb{R}^N) \subset L^2(\mathbb{R}^N, \lambda) \longrightarrow L^2(\mathbb{R}^N, \lambda), \;\; \Delta f = \sum_{j=1}^N \partial_j^2 f.$$

- Multiplikation mit x^2 auf $L^2(\mathbb{R}^N, \lambda)$:

$$M_{|\cdot|^2} : C_c(\mathbb{R}^N) \subset L^2(\mathbb{R}^N, \lambda) \longrightarrow L^2(\mathbb{R}^N, \lambda), \;\; M_{|\cdot|^2} f = (x \mapsto |x|^2 f(x)).$$

- Harmonischer Oszillator in der Quantenmechanik: $-\Delta + M_{|\cdot|^2}$.

Seien E, F normierte Raeume, so ist auch $E \times F$ mit

$$\|(x,y)\| := \|x\| + \|y\|$$

ein normierter Raum. Er ist vollstaendig genau dann, wenn E und F vollstaendig sind. (Nachrechnen!).

Bemerkung. Man kann auch

$$\|(x,y)\|_p := (\|x\|^p + \|y\|^p)^{1/p}$$

fuer $1 \leq p < \infty$ bzw.

$$\|(x,y)\|_\infty := \max\{\|x\|, \|y\|\}$$

einfuehren. Diese Normen sind alle aequivalent. Denn es sind alle Normen auf endlichdimensionalen Raeumen aequivalent und es gilt mit

$$J : E \times F \longrightarrow \mathbb{R}^2, J(x,y) := (\|x\|_E, \|y\|_F)$$

offenbar $\|(x,y)\|_p = \|J(x,y)\|_{p,\mathbb{R}^2}$.

Lemma. *Seien E, F normierte Raeume. Sei T ein linearer Operator von E nach F. Dann sind aequivalent.*

(i) *Es ist $G(T)$ abgeschlossen in $E \times F$.*
(ii) *Ist (x_n) eine Folge in $D(T)$ mit $x_n \to x$ und $Tx_n \to y$, so gilt $x \in D(T)$ und $Tx = y$.*

Sind E, F Banachraeume so ist dies aequivalent zu

(iii) *$(D(T), \|\cdot\|_T)$ ist vollstaendig, wobei $\|x\|_T := \|x\| + \|Tx\|$.*

Beweis. Die Aequivalenz von (*i*) und (*ii*) ist klar, denn (*ii*) ist gerade die Folgencharakterisierung von Abgeschlossenheit. Die Aequivalenz von (i) und (iii) ist klar, denn $G(T)$ ist abgeschlossen, genau dann wenn $G(T)$ vollstaendig ist, was wiederum zur Vollstaendigkeit von $D(T)$ in $\|\cdot\|_T$ aequivaelent ist. □

Uns wird es eigentlich immer um Operatoren gehen, die die Bedingungen des Lemma erfuellen:

Definition (Abgeschlossene Operatoren). *Seien E, F normierte Raeume und T ein linearer Operator von E nach F. Dann heisst T abgeschlossen, wenn er eine der Eigenschaften des vorigen Lemma erfuellt.*

Bemerkung. Stetigkeit eines Operators bedeutet

$$x_n \to x \Longrightarrow Tx_n \to Tx.$$

Damit ist Stetigkeit eine Forderung der *Konvergenz*.
Abgeschlossenheit eines Operators bedeutet

$$x_n \to x \text{ und } Tx_n \to y \Longrightarrow Tx_n \to Tx.$$

Damit ist Abgeschlossenheit lediglich eine Forderung der *Konsistenz*.

Die vorangehende Bemerkung laesst erwarten, dass Abgeschlossenheit eine schwaechere Eigenschaft als Stetigkeit ist. Die naechste Proposition besagt, dass Stetigkeit in der Tat eine staerkere Eigenschaft als Abgeschlosseheit ist.

PROPOSITION (Beschraenkte Operatoren sind abgeschlossen). *Seien E, F normierte Raeume und $T \in L(E, F)$. Dann ist T abgeschlossen.*

Beweis. Sei $(x_n, Tx_n) \to (x, y)$. Dann gilt $x_n \to x$ und daher (T stetig) $Tx_n \to Tx$. Weiterhin gilt $Tx_n \to y$. Insgesamt folgt $Tx = \lim Tx_n = y$. □

PROPOSITION (Abgeschlossenheit vertraeglich mit Inversenbildung). *Sind E, F normierte Raeume und T ein injektiver Operator von E nach F und $T^{-1} : TE \longrightarrow E$ der inverse Operator. Dann ist T abgeschlossen, genau dann wenn T^{-1} abgeschlossen ist.*

Beweis. Das ist klar, da die Graphen von T und T^{-1} durch Vertauschen der Komponenten auseinander hervorgehen. □

Bemerkung. Proposition zeigt die Flexibilitaet unseres Konzeptes von nur auf einem Teilraum definierten Operators und von Abgeschlossenheit. Ohne weitere Voraussetzungen an den Raum und (fast) ohne Beweis erhalten wir folgendes:

- Die Inverse eines injektiven Operators ist wieder ein Operator.
- Die Inverse eines injektiven abeschlossenen Operator ist wieder abgeschlossen.

Zum Vergleich mag der Satz von der stetigen Inversen dienen. Dieser enhaelt eine entsprechende Aussage fuer stetige Operatoren. Allerdings sind die Voraussetzungen wesentlich staerker, denn die Raeume muessen Banachraeume sein und der Operator muss bijektiv sein, und der Beweis ist wesentlich aufwaendiger.

THEOREM (Satz vom abgeschlossenen Graphen). *Seien E, F Banachraeume und T ein abgeschlossener linearer Operator von E nach F. Ist $D(T)$ abgeschlossen, so ist T beschraenkt.*

Bemerkungen.

- Die Vorraussetzung der Abgeschlossenheit des Definitionsbereiches ist erfuellt z.B. fuer $D(T) = E$. Tatsaechlich kann man das ohne Einschraenkung nach Verkleinern auch gleich so annehmen.
- Man kann den Satz als gute und als schlechte Nachricht lesen:

  Die gute Nachricht betrifft die Inversen von bijektiven abgeschlossenen Operatoren. Diese sind automatisch stetig (siehe Folgerung unten).

Die schlechte Nachricht betrifft die uns interessierenden unbeschraenkten Operatoren. Diese sind abgeschlossen. Damit koennen sie *unter keinen Umstaenden* auf dem ganzen Banachraum definiert werden. (Denn dann muessten sie beschraenkt sein).

- Es gilt auch die Umkehrung: Ist T beschraenkt und $D(T)$ abgeschlossen, so ist T abgeschlossen. Das folgt aus einer vorangehenden Proposition mit $E = D(T)$.
- Der Satz besagt sinngemaess, dass die beiden 'Probleme', die sich im Zusammenhang mit dem Operator T stellen:
  – T ist nicht ueberall definiert,
  – T ist unbeschraenkt,

  eigentlich nur zwei Seiten derselben Medaille sind.

Beweis. Ohne Einschraenkung sei $D(T) = E$. (Da E ein Banachraum ist, ist auch das abgeschlossene $D(T)$ ein Banachraum. Wir konnen uns auf $D(T)$ einschraenken). Nach Voraussetzung ist $G(T)$ als abgeschlossener Teilraum des Banachraum (!) $E \times F$ ebenfalls ein Banachraum. Betrachte nun (*Zeichnung*)

$$P : G(T) \longrightarrow E,\ P(x, Tx) = x,$$

$$Q : G(T) \longrightarrow F,\ Q(x, Tx) = Tx.$$

Dann gilt:

Es sind P, Q linear. klar.

P, Q sind stetig. (Nur P) $\|P(x, Tx)\| = \|x\| \leq \|x\| + \|Tx\| = \|(x, Tx)\|$.

Es ist P bijektiv. Injektiv: $0 = P(x, Tx) = x$. Dann $Tx = 0$, also $(x, Tx) = (0, 0)$.
Surjektiv: $x = P(x, Tx)$ fuer jedes $x \in E$.

Damit sind alle Voraussetzungen des Satzes von der stetigen Inversen fuer P erfuellt und es ist P^{-1} stetig. Damit ist dann auch $QP^{-1} = T$ stetig. $\square$

Ein wichtiges Thema (nicht nur) im Zusammenhang mit unbeschraenkten Operatoren ist Invertierbarkeit. Wir **erinnern** noch einmal an die entsprechenden Begriffe: Ein Operator T von E nach F heisst *bijektiv* oder *invertierbar*, wenn $T : D(T) \longrightarrow F$ bijektiv ist. Es ist T genau dann bijektiv, wenn ein lineaerer Operator $R : F \longrightarrow E$ existiert mit

$$RT = I_{D(T)} \text{ und } TR = I_F.$$

Der Operator R heisst dann der *Inverse* von T und wird mit T^{-1} bezeichnet. Ist T^{-1} stetig, so heisst T auch *stetig invertierbar.*
Eine entscheidende Folgerung des Satzes vom abgeschlossenen Graphen ist folgende.

FOLGERUNG (Stetigkeit der Inversen von abgeschlossenen Operatoren). *Seien E, F Banachraeume und der Operator T von E nach F sei linearer, abgeschlossen und bijektiv. Dann ist T^{-1} stetig.*

Beweis. Da T abgeschlossen ist, ist auch T^{-1} abgeschlossen. Da T surjektiv ist, gilt $D(T^{-1}) = F$. Damit folgt aus dem Satz vom abgeschlossenen Graphen also die gewuenschte Stetigkeit. $\square$

Bemerkung.

- Fuer beschraenkte ueberall definierte Operatoren kennen wir den Satz schon (als Satz von der stetigen Inversen). In der Folgerung wird 'beschaenkt und ueberall definiert' zu 'abgeschlossen' erweitert.
- Fuer die Inverse eines bijektiven Operators zwischen Banachraeumen gibt es also folgende beide Moeglicheiten:
  - Wenn der Operator abgeschlossen ist, so besagt die vorangehende Folgerung, dass die Inverse 'von alleine stetig ist' ist.
  - Wenn der Operator nicht abgeschlossen ist, so ist die Inverse unter keinen Umstaenden stetig. (Bew. Ang. Inverse stetig. Dann ist die Inverse stetig und ueberall definiert und damit (siehe Bemerkung nach Satz vom abgeschlossenen Graphen) abgeschlossen. Damit ist auch der Operator abgeschlossen. Widerspruch.)

  Damit kann man eine schoene Theorie mit stetigen Inversen nur fuer abgeschlossene Operatoren machen (und dort ist Stetigkeit automatisdch). Wir werden uns eigentlich nur mit abgeschlossenen Operatoren befassen.

Bemerkung - magisches Dreieck. Seien E, F Banachraeume und T ein linearer Operator von E nach F. Dann gilt fuer die drei Eigenschaften

- T beschraenkt,
- T abgeschlossen,
- $D(T)$ abgeschlossen,

dass je zwei dieser Eigenschaften die dritte implizieren. Insbesondere sind also je zwei der Eigenschaften aequivalent, wenn die dritte gilt. (Beweis: Die meisten dieser Aussagen sind einfach. die einzige wirklich schwere Aussage ist der Satz vom abgeschlossenen Graphen. Hier sind die Details:
T beschraenkt und $D(T)$ abgeschlossen. Dann ist T abgeschlossen: Einfach $((x_n, Tx_n) \to (x, y)$. Dann $x_n \to x$ also $((D(T)$ abg), $x \in D(T)$, also (T stetig) $Tx = \lim Tx_n = y$).

T beschraenkt, T abgeschlossen. Dann ist $D(T)$ abgeschlossen: Sei (x_n) Folge in $D(T)$ mit $x_n \to x$. Da T beschraenkt ist, ist (Tx_n) eine Cauchy Folge. Da F Banachraum ist, ist dann Tx_n konvergent gegen ein y. Damit konvergiert (x_n, Tx_n) gegen (x, y). Da T abgeschlossen ist, folgt $x \in D(T)$ (und $Tx = y$).

$D(T)$ abgeschlossen und T abgeschlossen. Dann ist T beschraenkt. Satz vom abgeschlossenen Graphen.)

Nach diesen allgemeinen Untersuchungen kommen wir nun auf die Beispiele vom Anfang des Abschnittes zurueck.

Beispiele.

Maximaler Operator der Multiplikation. Sei $(X, \mathcal{A}, \mu)$ ein Maßraum und $u : X \longrightarrow \mathbb{C}$ meßbar und $p \in [1, \infty]$. Dann definiert man -den *maximalen Operator der Multiplikation mit u in L^p* via

$$D(M_u) := \{f \in L^p(X, \mu) : uf \in L^p(X, \mu)$$

$$M_u f = uf.$$

Sicher ist diese Definition keine grosse Ueberraschung. Wie wir schon wissen, ist M_u beschraenkt, wenn u beschraenkt ist und man kann - im σ-endlichen Fall - auch leicht die Umkehrung sehen, vgl. Uebung. Hier zeigen wir nun, dass fuer beliebiges u der Operator M_u abgeschlossen ist. (Das liefert dann nach dem Satz vom abgeschlossenen Graphen noch einen weiteren Beweis der Stetigkeit fuer den Fall des beschraenkten u, weil dann offenbar M_u ueberall definiert ist).

Beweis der Abgeschlossenheit: Sei (f_n) eine Folge in L^p mit $f_n \to f$ in L^p und $M_u f_n \to g$ (in L^p). Die Idee ist, dass Konvergenz in L^p punkteweise Konvergenz gibt und damit alles leicht folgt. Hier sind die Details:
Wegen $f_n \to f$ gibt es eine Teilfolge f_{n_k}von (f_n), die punktweise fast sicher gegen f konvergiert. Wegen $M_u f_{n_k} = u f_{n_k} \to g$ gibt es eine Teilfolge $f_{n_{k_l}}$ mit $u f_{n_{k_l}} \to g$ fast sicher. Damit gilt dann also $f_{n_{k_l}} \to f$ fast sicher und damit auch

$$u f_{n_{k_l}} \to u f \text{ fast sicher.}$$

Weiterhin gilt nach Konstruktion

$$u f_{n_{k_l}} \to g \text{ fast sicher.}$$

Damit folgt $uf = g \in L^p$ und die Abgeschlossenheit folgt.

Der Laplaceoperator auf $L^2(\mathbb{R}^N, \lambda)$. Wir erinnern zunaechst an zwei Aussagen ueber den Schwartzraum $\mathcal{S}(\mathbb{R}^N)$ und die Fouriertransformation F (die wir aus Analysis III kennnen):
(1) Es ist $F : \mathcal{S}(\mathbb{R}^N) \longrightarrow \mathcal{S}(\mathbb{R}^N)$ bijektiv mit

$$\int_{\mathbb{R}^N} |f(x)|^2 d\lambda(x) = \int_{\mathbb{R}^N} |Ff(k)|^2 d\lambda(k)$$

fuer alle $f \in \mathcal{S}(\mathbb{R}^N)$.
(2) Es gilt

$$\Delta f = F^{-1}(|\cdot|^2 F f)$$

fuer alle $f \in \mathcal{S}(\mathbb{R}^N)$.
Nach Aussage (1) laesst sich, F zu einer unitaeren Abbildung von $L^2(\mathbb{R}^N, \lambda)$ nach $L^2(\mathbb{R}^N, \lambda)$ fortsetzen, die wir wieder mit F bezeichnen. Da $M_{|\cdot|^2}$ nach dem vorigen Beispiel ein abgeschlossener Operator ist, ist dann jedenfalls auch

$$F^{-1} M_{|\cdot|^2} F$$

ein abgeschlossener Operator. Nach Aussage (1) ist aber $F^{-1} M_{|\cdot|^2} F$ jedenfalls eine Fortsetzung des 'ueblichen' Laplacoperator auf $\mathcal{S}(\mathbb{R}^N)$. Damit koennen wir also *definieren*

$$\Delta := F^{-1} M_{|\cdot|^2} F.$$

Mit dieser Definition wird Δ ein abgeschlossener Operator, der den ueblichen Lapolacoperator auf $\mathcal{S}(\mathbb{R}^N)$ (und damit inbesondere auf $C_c^\infty(\mathbb{R}^N)$) fortsetzt.

2. Etwas Stabilitaetstheorie

In diesem Abschnitt untersuchen wir Abgeschlossenheit und Invertierbarkeit auf Stabilitaet unter Stoerungen. Genauer werden wir den Operator $T+S$ untersuchen, wobei T abgeschlossen bzw. stetig invertierbar ist und S klein ist. Dabei wird es um 'klein bzgl. T' gehen. Was das genau bedeuten soll, wird Teil der Untersuchungen sein.

Ende der Vorlesung

Definition. *Seien E, F normierte Raeume und S, T lineare Operatoren von E nach F. Der Operator S heisst T-beschraenkt, wenn gilt*

- $D(S) \supset D(T)$
- *es gibt $a, b \geq 0$ mit*
$$\|Sx\| \leq a\|x\| + b\|Tx\|$$
fuer alle $x \in D(T)$.

Das Infimum aller dieser b heisst die T-Schranke von S.

Bemerkung.

- Im allgemeinen ist die T Schranke kein moegliches b.
- Der Operator S ist T-beschraenkt, wenn $\|x\|_S \leq C\|x\|_T$ gilt.
- Ist $S \in L(E, F)$, so gilt fuer jeden Operator T, dass S T-beschraenkt mit Schranke 0 ist. (Denn $\|Sx\| \leq C\|x\| + 0\|Tx\|$.)

Fuer den naechsten Satz muessen wir die Summe von Operatoren von E nach F definieren. Sind T und S Operatoren von E nach F, so definiert man den Operator $T+S$ durch

$$D(T+S) := D(T) \cap D(S)$$
$$(T+S)x := Tx + Sx.$$

Theorem (Stabilitaet der Abgeschlossenheit). *Seien E, F Banachraeume und S, T lineare Operatoren von E nach F. Der Operator S sei T-beschraenkt mit T-Schranke kleiner als Eins. Dann ist $T+S$ mit $D(T+S) = D(T)$ genau dann abgeschlossen, wenn T abgeschlossen ist.*

Beweis. Da S die T-Schranke kleiner als Eins hat, gibt es b mit $0 \leq b < 1$ und $a \geq 0$ mit
$$\|Sx\| \leq a\|x\| + b\|Tx\|$$
fuer alle $x \in D(T)$. Wir zeigen nun, dass $\|\cdot\|_T$ und $\|\cdot\|_{T+S}$ aequivalent sind. Es gilt fuer alle $x \in D(T)$ zweierlei:

Erstens:
$$\|Tx\| \leq \|(T+S)x\| + \|Sx\| \leq \|(T+S)x\| + a\|x\| + b\|Tx\|$$
also
$$(1-b)\|Tx\| \leq a\|x\| + \|(T+S)x\|$$
also
$$\|x\|_T = \|x\| + \|Tx\| \leq C(\|x\| + \|(T+S)x\|) = C\|x\|_{T+S}.$$
Zweitens:

$$\begin{aligned}
\|x\|_{T+S} &= \|x\| + \|(T+S)x\| \\
&\leq \|x\| + \|Tx\| + \|Sx\| \\
&\leq (1+a)\|x\| + (1+b)\|Tx\| \\
&= D\|x\|_T.
\end{aligned}$$

Damit sind die Normen $\|\cdot\|_T$ und $\|\cdot\|_{T+S}$ auf $D(T) = D(T+S)$ aequivalent. Es ist dann also $(D(T), \|\cdot\|_T)$ genau dann vollstaendig, wenn $(D(T+S), \|\cdot\|_{T+S})$ vollstaendig ist und die Aussage folgt aus der Charakterisierung der Abgeschlossenheit. $\square$

Die folgende Proposition ist der Ursprung aller Stabilitaetsresultate fuer Invertierbarkeit.

PROPOSITION (Neumannsche Reihe). *Sei E ein Banachraum und $S \in L(E)$ mit* $\limsup_{N\to\infty} \|S^N\|^{1/N} < 1$. *Dann ist $I - S$ stetig invertierbar mit*

$$(I - S)^{-1} = \sum_{n=0}^{\infty} S^n,$$

wobei die Summe absolut konvergiert.

Bemerkungen.

- Man kann das Resultat als ein Stueck Stoerungstheorie verstehen: Die stetig invertierbare Identitaet wird durch das 'kleine' S gestoert und das Ergebnis ist weiterhin stetig invertierbar.
- Der Spezialfall $E = F = \mathbb{C}$ liefert die geometrische Reihe:
$$\frac{1}{1-q} = \sum_{n=0}^{\infty} q^n.$$
- Absolute Konvergenz von $\sum_{n=0}^{\infty} S^n$ heisst gerade $\sum_{n=0}^{\infty} \|S^n\| < \infty$. Da E ein Banachraum ist, impliziert dies die Existenz von
$$\sum_{n=0}^{\infty} S^n = \lim_{N\to\infty} \sum_{n=0}^{N} S^n.$$
(Tatsaechlich kann man das fuer beliebige Banachraeume E mit Norm $\|\cdot\|$ genauso beweisen, wie man es fuer $E = \mathbb{C}$ mit Norm $|\cdot|$ in Analysis I beweist.)

Beweis. Nach dem Wurzelkriterium gilt

$$\sum_{n=0}^{\infty} \|S^n\| < \infty.$$

Das liefert die absolute Konvergenz der Reihe. Diese absolute Konvergenz liefrt dann die Existenz von $T := \sum_{n=0}^{\infty} S^n$. Damit erhaelt man dann nach kurzer Rechnung

$$\begin{aligned}
(I-S)T &= (I-S)\lim_{N\to\infty}\sum_{n=0}^{N} S^n \\
&= \lim_{N\to\infty}(I-S)\sum_{n=0}^{N} S^n \\
\text{(Teleskop-Summe)} \quad &= \lim_{N\to\infty}(I-S^{N+1}) \\
&= I.
\end{aligned}$$

Dabei haben wir im letzten Schritt $\|S^{N+1}\| \to 0$ verwendet.
Analog zeigt man

$$T(I-S) = ... = I.$$

Damit ist $(I-S)$ also stetig invertierbar mit dem angegebenen Inversen. □

Bemerkung. Gilt $\|S\| < 1$, so ist $\limsup_N \|S^N\|^{1/N} < 1$ automatisch erfuellt. Dieser Fall spielt in den Anwendungen eine grosse Rolle.

FOLGERUNG. *Sei E ein Banachraum und $T \in L(E)$ stetig invertierbar. Sei $S \in L(E)$ mit $\|ST^{-1}\| < 1$ gegeben. Dann ist $T+S$ stetig invertierbar mit*

$$(T+S)^{-1} = \sum_{n=0}^{\infty}(-1)^n T^{-1}(ST^{-1})^n,$$

wobei die Summe absolut konvergiert.

Beweis. Es gilt

$$T+S = (I+ST^{-1})T = (I-(-ST^{-1}))T.$$

Nach Voraussetzung hat $-ST^{-1}$ die Norm kleiner als Eins. Damit ist es nach der Proposition ueber die Neumannsche Reihe invertierbar. Weiterhin ist T invertiebar. Nun folgt die gewuenscht Aussage einfach. □

Bemerkungen.

- Auch dieses Resultat kann man als ein Stueck Stoerungstheorie verstehen: Die stetig invertierbare T wird durch das (bzgl. T) 'kleine' S gestoert und das Ergebnis ist weiterhin stetig invertierbar. Die Folgerung besagt gerade, dass um jedes stetig invertierbare T noch Platz ist.
- Die Menge der stetig invertierbaren Elemente in $L(E)$ wird auch mit $GL(E)$ bezeichnet. Man macht sich leicht klar, dass diese Menge ein Gruppe ist bzgl. der Multiplikation (mit der Identitaet I als neutralem Element). Die Folgerung besagt, dass $GL(E)$ offen in $L(E)$ ist.
- Die Terme in der Summe moegen etwas unsymmetrisch aussehen. Tatsaechlich ist aber der Term an der Stelle n gegeben durch

  $$T^{-1}ST^{-1}\cdots T^{-1}ST^{-1},$$

  wobei n-mal der Faktor S auftritt.

Fuer den naechsten Satz muessen wir das Produkt TS von Operatoren definieren. Fuer Operatoren S von E nach F und T von F nach G definiert man den Operator TS von E nach G durch

$$D(TS) = \{x \in D(S) : Sx \in D(T)\},\ (TS)x = T(Sx).$$

Ist also T ein bijektiver Operator von E nach F (d.h. $T : D(T) \longrightarrow F$ ist bijektiv) und S ein Operator von E nach F mit $D(S) \supset D(T)$, so ist der Operator ST^{-1} auf ganz F definiert (da fuer alle $y \in F$ $T^{-1}y \in D(T) \subset D(S)$ gilt).

Theorem (Stabilitaet der stetigen Invertierbarkeit). *Seien E, F Banachraeume und S, T lineare Operatoren von E nach F. Es sei T abgeschlossen und bijektiv (also $T^{-1} \in L(F, E)$). Gilt $D(S) \supset D(T)$ und $\|ST^{-1}\| < 1$, so ist auch $T + S$ abgeschlossen und bijektiv und es gilt*

$$(T + S)^{-1} = \sum_{n=0}^{\infty} (-1)^n T^{-1} (ST^{-1})^n,$$

wobei die Reihe absolut konvergiert (also insbesondere auch bzgl. der Norm in $L(E, F)$).

Bemerkung. Formal folgt das sofort wie bei den Betrachtungen zu beschraenkten Operatoren zu Beginn des Kapitels. Das Problem ist, dass $T+S$ nicht stetig ist und man darum einen zweiten Blick werfen muss.

Beweis. Wir zeigen eine Reihe von Aussagen:
Es gilt $T + S = (I + ST^{-1})T$. In der Tat sind beide Seiten genau auf $D(T)$ definiert und liefern dort den gleichen Wert.

$T + S$ ist bijektiv mit der angegebenen Inversen. Es ist $T : D(T) \longrightarrow F$ bijektiv mit stetiger Inverser $T^{-1} : F \longrightarrow D(T)$. Weiterhin ist nach Voraussetzung $A = ST^{-1}$ beschraenkt mit $\|A\| < 1$ und damit ist nach Neumannscher Reihe $(I + A) : F \longrightarrow F$ bijektiv mit stetiger Inverser

$$(I + A)^{-1} = \sum_{n=0}^{\infty} (-1)^n A^n.$$

Damit ist

$$T + S = (I + ST^{-1})T = (I + A)T$$

bijektiv mit stetiger Inverser gegeben durch

$$T^{-1} \sum_{n=0}^{\infty} (-1)^n (ST^{-1})^n = \sum_{n=0}^{\infty} (-1)^n T^{-1} (ST^{-1})^n.$$

Es ist $T + S$ abgeschlossen. Da $(T + S)^{-1}$ stetig und ueberall definiert ist, ist es abgeschlossen. Damit ist dann auch $T + S$ abgeschlossen. □

3. Spektraltheorie abgeschlossener Operatoren

Fuer abgeschlossene Operatoren T in einem normierten Raum E entwickeln wir eine Loesungstheorie fuer Gleichungen der Form

$$(T - zI)x = y$$

bei gegebenem $y \in E$ und $z \in \mathbb{C}$. Wir wollen

- Loesbarkeit der Gleichung fuer alle y (d.h. Surjektivitaet von $T - zI$).
- Eindeutigkeit der Loesung (d.h. Injektivitaet von $T - zI$).
- Stetige Abhaengigkeit der Loesung x von y (d.h. Stetigkeit von $(T - zI)^{-1}$.
- Stetige Abhaengigkeit von z.

Die ersten drei Punkte bedeuten gerade, dass $T - zI$ stetig invertierbar ist. Entsprechende Betrachtungen sind als *Spektraltheorie* bekannt. Im Falle von Operatoren in endlichdimensionalen Raeumen (Matrizen) geht es gerade um Eigenwerte.

Wir *definieren* zunaechst einmal die Menge der λ, auf denen 'alles gutgeht'.

Definition. *Sei E ein normierter Raum und T ein linearer Operator von E nach E. Dann definiert man die Resolventenmenge $\varrho(T)$ von T als*

$$\varrho(T) := \{\lambda \in \mathbb{K} : T - \lambda I : D(T) \longrightarrow E \text{ bijektiv mit stetiger Inverser}\}.$$

Das Komplement $\mathbb{K} \setminus \varrho(T)$ heisst Spektrum von T und wird mit $\sigma(T)$ bezeichnet. Die Abbildung

$$R = R_T : \varrho(T) \longrightarrow L(E),\ \lambda \mapsto R_T(\lambda) = (T - \lambda I)^{-1},$$

heisst die Resolvente von T.

Notation. Wir schreiben auch $T - \lambda$ statt $T - \lambda I$. Oft wird auch die Resolventenmenge als Resolvente bezeichnet.

Bemerkung - Zur Abgeschlossenheit von T. Die Definition von Spektrum und Resolvente setzt nicht die Abgeschlosseheit von T voraus. Tatsaechlich sind aber die Begriffsbildungen nur fuer abgeschlossene Operatoren interessant. Es gilt naemlich folgendes: Ist T nicht abgeschlossen, so gilt

$$\varrho(T) = \emptyset.$$

(Bew: Sei $z \in \mathbb{C}$ beliebig.
Fall 1: Es ist $T - z$ nicht bijektiv: Dann gehoert z nicht zu $\varrho(T)$.

Fall 2: Es ist $T - z$ bijektiv: Dann kann $(T - z)^{-1}$ nicht stetig sein. Denn sonst waere $(T - z)^{-1}$ abgeschlossen (s.o.) und damit dann auch $T - z$ abgeschlossen und damit dann T abgeschlossen.)
Ist E ein Banachraum, so folgt aus der Abgeschlossenheit von T automatisch die Stetigkeit von $(T - \lambda)^{-1}$.
Wir werden es im folgenden daher (fast) ausschliessich mit abgeschlossenen Operatoren in Banachraeumen zu tun haben.

Grundlegende Eigenschaften von Spektrum und Resolvente liefert der folgende Satz. Er besagt, dass die Resolventenmenge eine schoene Menge ist und die Resolvente eine schoene Funktion.

Theorem (Grundlegende Eigenschaften von Resolvente und Spektrum). *Sei E ein Banachraum und T ein Operator von E nach E. Dann ist $\varrho(T)$ eine offene Teilmenge von $\mathbb{K}$, also $\sigma(T)$ eine abgeschlossene Teilmenge von $\mathbb{K}$, und die Resolvente*

$$R : \varrho \longrightarrow L(E),\ \lambda \mapsto (T - \lambda)^{-1},$$

ist analytisch, d.h. um jeden Punkt in eine absolut konvergente (also inbesondere normkonvergente) Potenzreihe mit positivem Konvergenzradius entwickelbar.

Beweis. Das folgt letzten Endes aus der Neumannschen Reihe. Sei $z_0 \in \varrho(T)$. Dann ist $(T - z_0)$ invertierbar. Damit ist dann nach dem Satz ueber die Stabilitaet der Invertierbarkeit auch

$$T - z = (T - z_0) + (z_0 - z) =' T + S'$$

invertierbar falls

$$1 > \|(z - z_0)(T - z_0)^{-1}\| = |z - z_0|\|(T - z_0)^{-1}\|$$

also

$$|z - z_0| < \frac{1}{\|(T - z_0)^{-1}\|}.$$

Das zeigt die Offenheit der Resolventenmenge. Weiterhin folgt aus dem Satz ueber die Stabilitaet der Invertierbarkeit auch noch

$$\begin{aligned}(T - z)^{-1} &= ((T - z_0) + (z_0 - z))^{-1} \\ &= \sum_{n=0}^{\infty} (T - z_0)^{-1}((z - z_0)(T - z_0)^{-1})^n \\ &= \sum_{n=0}^{\infty} (z - z_0)^n (T - z_0)^{-n-1}\end{aligned}$$

mit absolut konvergenten Potenzreihe. Damit folgt die Analytizitaet. □

FOLGERUNG. *Sei die Situation wie im Theorem. Dann ist*

$$R : \varrho(T) \longrightarrow L(E),\ \lambda \mapsto (T - \lambda)^{-1},$$

stetig.

Beweis. Es gilt allgemein, dass eine analytische Funktion stetig ist. Konkret ist

$$(T - z)^{-1} - (T - z_0)^{-1} = \sum_{k=1}^{\infty} (z - z_0)^n (T - z_0)^{-n-1}$$

absolut konvergent und

$$\| \ldots \| \leq |z - z_0|\|(T - z_0)^{-2}\| \sum_{k=0}^{\infty} \|((z - z_0)(T - z_0)^{-1}\|^k \leq |z - z_0| C$$

fuer z nahe z_0 (da dann $\|((z - z_0)(T - z_0)^{-1})\| \leq q < 1$ gleichmaessig). □

Wir wenden uns nun Beispielen zu.

Im Rahmen der Betrachtungen zu Hilbertraeumen haben wir schon einige Beispiele gesehen. Insbesondere wissen wir, dass das Spektrum im Falle von linearen Operatoren im endlichdimensionalen Raeumen (die dann durch Matrizen gegeben werden) gerade die Eigenwerte sind.

Beispiel - Linksshift auf ℓ^p: Sei $p \in [1, \infty]$ gegeben und $E = \ell^p$ mit $\| \cdot \|_p$. Dabei ist

$$\ell^p := \ell^p(\mathbb{N}) := \begin{cases} \{x : \mathbb{N} \longrightarrow \mathbb{C} : \sum_{n=1}^{\infty} |x(n)|^p < \infty\} & : \ 1 \leq p < \infty \\ \{x : \mathbb{N} \longrightarrow \mathbb{C} : x \text{ beschraenkt}\} & : \ p = \infty \end{cases}$$

und

$$\|x\|_p := \begin{cases} \left(\sum_{n=1}^{\infty} |x(n)|^p\right)^{1/p} & : \quad 1 \leq p < \infty \\ \sup\{|x(n)| : n \in \mathbb{N}\} & : \quad p = \infty. \end{cases}$$

Sei T der Shift-Operator auf ℓ^p, d.h.

$$T : \ell^p \longrightarrow \ell^p, (Tx)(n) = x(n+1).$$

Dann gilt offenbar $\|Tx\| \leq \|x\|$ und damit $\|T\| \leq 1$. Wir berechnen nun das Spektrum:

Es gilt $\sigma(T) \subset B_1(0)$. Das folgt sofort aus $\|T\| \leq 1$. In der Tat ist fuer $\lambda \in \mathbb{C}$ mit $|\lambda| > 1$ der Operator

$$T - \lambda = -\lambda(I - \frac{1}{\lambda}T)$$

nach dem Lemma zur Neumanschen Reihe stetig invertierbar wegen

$$\|\frac{1}{\lambda}T\| = \frac{\|T\|}{|\lambda|} < 1.$$

Es gilt $U_1(0) \setminus \{0\} \subset \sigma(T)$*:* Sei $\lambda \in U_1(0) \setminus \{0\}$. Dann gehoert

$$x : \mathbb{N} \longrightarrow \mathbb{C}, x(n) = \lambda^n,$$

zu ℓ^p (wegen $|\lambda| < 1$) und verschwindet nicht (wegen $\lambda \neq 0$). Ausserdem gilt offenbar $Tx = \lambda x$, d.h.

$$(T - \lambda)x = 0.$$

Damit ist also $T - \lambda$ nicht injektiv.

Nimmt man die vorangehenden Aussagen mit der Abgeschlossenheit de Spektrums zusammen so folgt

$$\sigma(T) = B_1(0).$$

Ende der Vorlesung

Bemerkung. Fuer $1 \leq p < \infty$ sind die $\lambda \in B_1(0) \setminus U_1(0)$ **keine** Eigenwerte (wie man leicht sieht). Das zeigt, dass das Spektrum eines Operators in einem unendlich dimensionalen Raum im allgemeinen nicht (nur) aus Eigenwerten besteht.

Beispiel - Rechtsshift auf ℓ^2: Sei $S : \ell^2 \longrightarrow \ell^2$ der Rechtsshift d.h.

$$(Sx)(n) = \begin{cases} x(n-1) & : \quad n \geq 2 \\ 0 & : \quad n = 1. \end{cases}$$

Dann ist S beschraenkt, und es gilt - wie man direkt sieht -

$$\langle Sx, y \rangle = \langle x, Ty \rangle$$

fuer alle $x, y \in \ell^2$. Damit folgt also $S = T^*$ und $T = S^*$. Damit folgt

$$\sigma(S) \stackrel{!}{=} \overline{\sigma(T)} = B_1(0).$$

Es bleibt ! zu zeigen: Wir zeigen die entsprechende Aussage fuer die Resolvente, d.h.

$$\varrho(S) = \overline{\varrho(T)}.$$

Wir zeigen $\supset$: Sei $\lambda \in \varrho(T)$. Dann existiert also ein beschraenkter Operator R mit

$$(T - \lambda)R = I = R(T - \lambda).$$

Durch Adjungieren folgt dann

$$R^*(T^* - \overline{\lambda}) = I = (T^* - \overline{\lambda})R^*.$$

Damit gehoert dann also $\overline{\lambda}$ zu $\varrho(T^*) = \varrho(S)$.
Wir zeigen $\subset$: Das folgt analog unter Nutzen von $S^* = T$.

Beispiel - Multiplikationsoperatoren. Sei $(X, \mathcal{A}, \mu)$ ein σ-endlicher Maßraum und $u : X \longrightarrow \mathbb{C}$ meßbar. Sei $1 \leq p \leq \infty$ und M_u der maximale Operator der Multiplikation mit u auf $L^p(X, \mu)$ d.h.

$$D(M_u) := \{f \in L^p(X, \mu) : uf \in L^p(X, m)\}$$
$$M_u f = uf.$$

Dann gilt:

- Fuer $1 \leq p < \infty$ ist $D(M_u)$ dicht in $L^p(X, \mu)$. Gilt $p = \infty$, so ist $D(M_u)$ genau dann dicht, wenn u wesentlich beschraenkt ist. In diesem Fall gilt $D(M_u) = L^\infty(X, \mu)$.
- $\sigma(M_u) =$ wesentlicher Wertebereich von $u := \{\lambda \in C : \mu(u^{-1}(U_\lambda(\varepsilon))) > 0 \text{ for all } \varepsilon > 0\}$.
- $\varrho(M_u) = \{z : \text{existiert } \varepsilon > 0 \text{ mit } |u(x) - z| \geq \varepsilon \text{ fuer fast alle } x \in X\}$.
- $(M_u - z)^{-1} = M_{\frac{1}{u-z}}$ fuer alle $z \in \varrho(M_u)$.
- Es ist λ ein Eigenwert von $D(M_u)$ genau dann, wenn $\mu(u^{-1}(\{\lambda\})) > 0$ gilt.

Beweis. Wir lassen den Beweis des ersten und vierten Punkt als Uebungsaufgabe. Wir zeigen nun die mittleren beiden Punkte zusammen. Sei dazu

$$\varrho' := \{z : \text{existiert } \varepsilon > 0 \text{ mit } |u(x) - z| \geq \varepsilon \text{ fuer fast alle } k \in X\}.$$

Es gilt $\varrho' = \varrho(T)$:
Fuer $z \in \varrho'$ ist $g := \frac{1}{u-z}$ eine beschraenkte Funktion und es gilt offenbar

$$(M_u - z)M_g = I, \;\; M_g(M_u - z) = I_{D(M_u)}.$$

Es ist also $z \in \varrho(T)$ und die angegebene Formel fuer die Inverse von $(M_u - z)$ gilt.
Fuer $z \notin \varrho'$ gilt fuer jedes $\varepsilon > 0$

$$\mu(\{x : |u(x) - z| \leq \varepsilon\}) > 0.$$

Da μ σ-endlich ist, gibt es dann eine meßbare Menge A mit $0 < \mu(A) < \infty$ und

$$A \subset \{x : |u(x) - z| \leq \varepsilon\}.$$

Dann gehoert die charakteristische Funktion 1_A von A zu $L^p(X, \mu)$ und es gilt

$$\|(M_u - z)1_A\| \leq \varepsilon \|1_A\|.$$

Da $\varepsilon > 0$ beliebig ist, gehoert dann also z nicht zu ϱ (s.o.).
Damit sind alles gewuenschten Aussagen gezeigt.

Bemerkung.

- Ein wichtiger Spezialfall ergibt sich fuer $X = \mathbb{N}$, $\mathcal{A} = \mathcal{P}(X)$ und μ Zaehlmass. Dann ist das Spektrum gerade gegeben durch

$$\overline{\{u(k) : k \in \mathbb{N}\}}.$$

- Jede abgeschlossene nichtleere Menge in $\mathbb{K}$ kann Spektrum (eines Multiplikationsoperator) sein. (Uebung...)
- (Uebung) Ist man in einer topologischen Situation, so kann man das Spektrum durch das Bild von u beschreiben. Genauer gilt folgendes: Sei X sogar ein topologischer Raum und es gelte fuer jede offene nichtleere Menge U noch $\mu(U) > 0$ und es sei $u : X \longrightarrow \mathbb{C}$ stetig. Dann gilt

$$\text{wesentliches Bild von } u = \overline{\text{Bild von } u}.$$

KAPITEL 9

Nachtrag - der Satz von Radon-Nikodym und die Lebesgue-Zerlegung

In diesem Abschnitt lernen wir eine weitere Klasse von Maßen kennen: die komplexen Maße. Sie stehen zu den positiven Maßen in einem aehnlichen Verhaeltnis wie $\mathcal{L}^1$ zu den nichtnegativen meßbaren Funktionen. Wir werden untersuchen, wie man ein komplexes Maß in Bezug auf ein gegebenes positives Maß zerlegen kann.

Definition (Komplexes Maß). *Sei $(X, \mathcal{A})$ ein meßbarer Raum. Eine Funktion $\mu : \mathcal{A} \longrightarrow \mathbb{C}$ heißt komplexes Maß, wenn*

$$\mu(E) = \sum_{n=1}^{\infty} \mu(E_n)$$

mit absolut konvergenter Summe gilt fuer jede Zerlegung (E_n) von E. Hier heißt (E_n) ein Zerlegung von E, wenn die E_n, $n \in \mathbb{N}$, meßbar und paarweise disjunkt sind mit $E = \bigcup_n E_n$.

Notation. Wir werden die bisher betrachteten Maße in diesem Abschnitt (meist) als positive Maße bezeichnen, um sie von den komplexen Maßen zu unterscheiden. Mit 'Maß' werden wir sowohl komplexe als auch positive Maße meinen.

Bemerkungen.

- Statt absoluter Konvergenz kann man auch nur die Konvergenz der Summe fordern. Dann wird die Summe naemlich unbedingt konvergent sein (d.h. fuer jede Umordnung ebenfalls konvergieren) und damit dann absolut konvergent (nach einem bekannten Satz aus der Analysis).
- Offenbar ist jedes komplexe Maß mit reellen nichtnegativen Werten auch ein positives Maß. Die Umkehrung gilt aber nicht (da ein komplexes Maß nicht den Wert ∞ annehmen darf). Tatsaechlich sind die komplexen Maße mit Werten in $[0, \infty)$ genau die positiven endlichen Maße, wie man sich leicht ueberlegt.
- Ein komlexes Maß mit reellen Werten heißt *signiertes Maß*.

Unser naechstes Ziel ist es, zu einem komplexen Maß μ ein positives Maß ν zu finden mit

$$|\mu(E)| \leq \nu(E)$$

fuer alle meßbaren E. Falls es ein solches ν gibt, muß gelten

$$\nu(E) = \sum_n \nu(E_n) \geq \sum_n |\mu(E_n)|$$

fuer jede Zerlegung (E_n) von E. Das legt es nahe $|\mu|$ zu definieren als

$$|\mu|(E) := \sup_{(E_n) \text{ Zerlegung von } E} \sum_n |\mu(E_n)|.$$

Wir bezeichnen $|\mu|$ als den *Betrag* von μ.

Theorem. *Sei μ ein komplexes Maß. Dann ist $|\mu|$ ein positives Maß mit $|\mu|(X) < \infty$. Weiterhin ist $|\mu|$ das kleinste positive Maß ν mit $|\mu(E)| \leq \nu(E)$ fuer alle meßbaren Mengen E.*

Bemerkungen. Der Endlichkeit von $|\mu|(X)|$ liefert, dass μ tatsaechlich nur Werte annimmt in der Kugel um 0 mit dem endlichen Radius $|\mu|(X)$.

Beweis. Wenn $|\mu|$ ein Maß ist, folgt die Aussage zum 'Weiterhin' sofort aus der Konstruktion.

Wir zeigen zunaechst die σ-Additivitaet von $|\mu|$. Sei (E_n) ein Zerlegung von E. Zu zeigen:

$$|\mu|(E) = \sum_n |\mu|(E_n).$$

Die Ungleichung '$\geq$'. Sei zu jedem n ein beliebiges t_n mit $t_n < |\mu|(E_n)$ gewaehlt. Dann gibt es zu jedem E_n eine Zerlegung $A_{n,m}$, $m \in \mathbb{N}$, mit

$$t_n \leq \sum_m |\mu(A_{n,m})|.$$

Dann ist aber $(A_{n,m})_{n,m}$ eine Zerlegung von E und damit folgt

$$|\mu|(E) \geq \sum_{n,m} |\mu(A_{n,m})| = \sum_n \sum_m |\mu(A_{n,m})| \geq \sum_n t_n.$$

Da die t_n beliebig (mit $t_n < |\mu|(E_n)$) sind, folgt die gewuenschte Ungleichung.

Die Ungleichung '$\leq$'. Sei (A_m) ein Zerlegung von E. Setze

$$A_{n,m} := A_m \cap E_n$$

fuer $n, m \in \mathbb{N}$. Dann ist $(A_{n,m})_m$ fuer jedes n ein Zerlegung von E_n und es ist $(A_{n,m})_n$ fuer jedes m eine Zerlegung von A_m. Damit folgt dann

$$\begin{aligned}
\sum_m |\mu(A_m)| &= \sum_m \left| \sum_n \mu(A_{n,m}) \right| \\
&\leq \sum_{m,n} |\mu(A_{n,m})| \\
(Fubini) &= \sum_n \sum_m |\mu(A_{n,m})| \\
&\leq \sum_n |\mu|(E_n).
\end{aligned}$$

Hier wird in der ersten Zeile genutzt, daß μ ein komplexes Maß ist und $(A_{n,m})_n$ eine Zerlegung von A_m und in der letzten Zeile wird die Definition von $|\mu|$ genutzt und daß $(A_{n,m})_m$ eine Zerlegung von E_n ist. Obige Ungleichungskette liefert die gewuenschte Ungleichung.

Es gilt $|\mu|(X) < \infty$. Wir zeigen zunaechst eine Hilfsaussage.

Zwischenbehauptung. Gilt $|\mu|(E) = \infty$, so existieren disjunkte A, B mit $A \cup B = E$ und $|\mu(A)| \geq 1$ und $|\mu|(B) = \infty$.
Beweis der Zwischenbehauptung. Nach Voraussetzung existiert eine Zerlegung (A_n) von E mit

$$\sum_n |\mu(A_n)| \geq 8(1 + |\mu(E)|).$$

Betrachten von Imaginaer- und Realteil liefert o.E.

$$\sum_n |\Re(\mu(A_n))| \geq 4(1 + |\mu(E)|).$$

Zerlegen in Summanden mit positivem und negativem Realteil liefert dann eine Teilfolge A_{n_k} mit o.E.

$$\sum_k \Re(\mu(A_{n_k})) \geq 2(1 + |\mu(E)|).$$

Insgesamt koennen wir dann also ohne Einschraenkung annehmen, dass

$$|\sum_{n=1}^N \mu(A_n)| \geq 1 + |\mu(E)|$$

gilt. Setze

$$A := \bigcup_{n=1}^N A_n, \quad B := E \setminus A.$$

Dann gilt

$$|\mu(A)| = |\sum_{n=1}^N \mu(A_n)| \geq 1 + |\mu(E)| \geq 1$$

und

$$|\mu(B)| = |\mu(E) - \mu(A)| \geq |\mu(A)| - |\mu(E)| \geq 1.$$

Da $|\mu|$ ein Maß ist, muss außerdem gelten

$$\infty = |\mu|(E)| = |\mu|(A) + |\mu|(B).$$

Damit hat also (mindestens) eine der Mengen A, B unendliches $|\mu|$ Maß. Ohne Einschraenkung sei dies B. Das beendet den Beweis der Zwischenbehauptung.

Gilt $|\mu|(X) = \infty$, so kann man mit der Zwischenbehauptung induktiv eine Folge von paarweise disjunkten Mengen A_n finden mit $|\mu(A_n)| \geq 1$. Dann konvergiert

$$\mu\left(\bigcup_n A_n\right) = \sum_n \mu(A_n)$$

nicht. Das ist ein Widerspruch. $\square$

Eine wesentliche Folgerung aus der Moeglichkeit, den Betrag eines Maßes zu bilden, ist die Zerlegung eines Masses in eine Summe von nichtnegativen endlichen Massen. Das ist der Inhalt der folgenden Proposition.

PROPOSITION. *Sei μ ein komplexes Maß. Dann existieren endliche positive Maße $\mu_1, \mu_2, \mu_3, \mu_4$ mit $\mu_j \le |\mu|$ und*

$$\mu = \mu_1 - \mu_2 + i\mu_3 - i\mu_4$$

(d.h. $\mu(E) = \mu_1(E) - \mu_2(E) + i\mu_3(E) - i\mu_4(E)$ fuer alle messbaren E).

Bemerkung. Die von uns im Beweis gegebene Zerlegung hat eine Minimalitaetseigenschaft, die sie eindeutig macht (d.h. jede weitere Zerlegung $\mu = \mu_1^* - \mu_2^* + i\mu_3^* - i\mu_4^*$ gilt $\mu_j^* \ge \mu_j$). Wir werden das hier nicht vertiefen.

PROOF. Das zeigt man aehnlich wie die Zerlegung einer komplexen Zahl als Linearkombination aus positiven Zahlen. Hier sind die Details: Wir definieren zunaecht $\Re\mu$ und $\Im\mu$ durch

$$(\Re\mu)(E) := \Re(\mu(E)) \text{ und } (\Im\mu)(E) := \Im(\mu(E)).$$

Dann sind offenbar $\nu_1 := \Re\mu$ und $\nu_2 := \Im\mu$ komplexe Maße mit reellen Werten und $|\nu_1|, |\nu_2| \le |\mu|$. Nun definieren wir

$$\mu_1 := \frac{1}{2}(|\nu_1| + \nu_1) \text{ und } \mu_2 := \frac{1}{2}(|\nu_1| - \nu_2)$$

sowie

$$\mu_3 := \frac{1}{2}(|\nu_2| + \nu_2) \text{ und } \mu_4 := \frac{1}{2}(|\nu_2| - \nu_2).$$

Dann sind die μ_j, $j = 1, \ldots, 4$, nach Konstruktion positive Maße mit

$$\mu_j(X) \le |\mu|(X)$$

fuer alle j sowie

$$\mu = \mu_1 - \mu_2 + i\mu_3 - i\mu_4.$$

Das beendet den Beweis. □

Bemerkung. Die komplexen Maße bilden eine Vektorraum mit

$$(\alpha\mu + \beta\nu)(E) := \alpha\mu(E) + \beta\nu(E).$$

Mittels $\|\mu\| := |\mu|(X)$ (s.u.) wird dies zu einem normierten Vektorraum. Dieser Raum ist vollstaendig (Uebung).

DEFINITION (Absolut stetig). *Sei ν ein positives Maß und μ ein beliebiges Maß. Es heißt μ absolut stetig bzgl. ν, wenn $\mu(E) = 0$ fuer jedes E mit $\nu(E) = 0$ gilt. Dann schreibt man $\mu \ll \nu$.*

Bemerkung. Es 'erklaert' sich die Bezeichnung *absolut stetig* wie folgt: Sei ν ein positives Maß und μ ein komplexes Maß. Dann sind aequivalent:

(i) Es ist μ absolut stetig bzgl. ν.
(ii) Zu jedem $\varepsilon > 0$ existiert ein $\delta > 0$ mit $|\mu(E)| < \varepsilon$ falls $\nu(E) < \delta$.

Beweis. (ii)$\Longrightarrow$ (i): Das ist einfach.

(i)$\Longrightarrow$ (ii): Sei (ii) falsch. Dann existiert ein $\varepsilon > 0$ und Mengen E_n mit $\nu(E_n) \le 2^{-n}$ aber $|\mu(E_n)| \ge \varepsilon$. Insbesondere gilt also $|\mu|(E_n) \ge \varepsilon$. Setze

$$A_n := \bigcup_{i \ge n} E_i, \quad A := \bigcap_n A_n.$$

Dann gilt also $\nu(A_n) \le \sum_{i \ge n} 2^{-i} = 2^{-n+1}$ und $A_n \supset A_{n+1}$.

Weiterhin gilt $\nu(A_1) \leq 2 < \infty$ und $|\mu|(A_1) \leq |\mu|(X) < \infty$ Damit folgt

$$\nu(A) = \lim_n \nu(A_n) = 0$$

und (wegen $A_n \supset E_n$)

$$|\mu|(A) = \lim_n |\mu|(A_n) \geq \varepsilon.$$

Damit ist also $|\mu|$ nicht absolut stetig bezueglich ν. Damit kann aber μ nicht absolut stetig sein bezueglich ν. (Sonst: Sei (E_n) eine Zerlegung einer ν Nullmenge. Dann gilt $\nu(E_n) = 0$ fuer jedes n. Damit folgt also $\mu(E_n) = 0$ fuer jedes n. Damit folgt $\sum_n |\mu(E_n)| = 0$ fuer die Zerlegung (E_n). Da dies fuer jede Zerlegung gilt, folgt $|\mu|(E) = 0$.) $\square$

Bemerkung. Eine entsprechende Aussage gilt im allgemeinen nicht, wenn μ ein positives (aber nicht endliches) Maß ist: Sei ν die Einschraenkung des Lebesguemaß auf $[0,1]$ und $\mu(E) = \int_E \frac{1}{t} d\nu(t)$. Dann gilt die Aussage der Proposition nicht, wie man leicht durch Betrachten von Intervallen der Form $I = [0,t]$ (mit $0 < t \leq 1$) sehen kann.

Wir brauchen noch einige weitere Konzepte.

Definition. *Sei $(X, \mathcal{A})$ ein Maßraum.*
(a) Sei μ ein Maß auf X und A eine Teilmenge von X. Dann heißt μ auf A konzentriert, wenn

$$\mu(E) = \mu(A \cap E)$$

gilt fuer alle meßbaren E.
(b) Die Maße μ_1 und μ_2 auf X heißen gegenseitig singulaer,

$$\mu_1 \perp \mu_2,$$

wenn es disjunkte meßbare Menge A, B gibt, so daß μ_1 auf A konzentriert ist und μ_2 auf B konzentriert ist.

Bemerkung. Es ist μ auf A konzentiert genau dann wenn gilt $\mu(E) = 0$ fuer alle E mit $E \cap A = \emptyset$. Ist μ positiv, so ist dies aequivalent zu $\mu(A^c) = 0$. (Beweis: Zweite Aequivalenz ist klar. Zur ersten Aequivalenz: Ist μ auf A konzentriert, so gilt die Aussage sicher. Umgekehrt folgt aus der Aussage auch

$$\mu(E) = \mu(E \cap A) + \mu(E \cap A^c) = \mu(E \cap A).$$

Theorem (Lebesgue-Radon-Nikodym). *$(X, \mathcal{A})$ ein meßbarer Raum. Sei ν ein positives σ-endliches Maß und sei μ ein komplexes Maß.*
(a) Es existiert ein eindeutiges Paar (μ_{ac}, μ_{sing}) von komplexen Maßen mit

$$\mu = \mu_{ac} + \mu_{sing}, \;\; \mu_{ac} \ll \nu, \;\; \textit{und} \;\; \mu_{sing} \perp \nu.$$

Ist μ positive (und endlich), so sind es auch μ_{ac} und μ_{sing}.
(b) Es gibt ein (bis auf Nullmengen) eindeutig bestimmtes $h \in \mathcal{L}^1(X, \nu)$ mit

$$\mu_{ac}(E) = \int_E h d\nu$$

fuer alle meßbaren E.

Bemerkungen.

- Die Aussage von Teil (a) des Satzes ist als *Lebesgue-Zerlegung von μ bezueglich ν* bekannt. Es heißt μ_{ac} der (bzgl. ν) *absolut stetig Anteil von μ* und μ_{sing} der (bzgl. ν) *singulaere Anteil von μ.*
- Teil (b) des Satze ist als *Satz von Radon-Nikodym bekannt.* Wir werden noch eine Fassung fuer positive Maße kennenlenernen. Es heißt h die *Radon-Nikodym Ableitung* von μ_{ac} bezueglich ν. Man schreibt auch $h = d\mu_{ac}/d\nu$ und (wie schon oben eingefuehrt) $\mu_{ac} = hd\nu$.
- Teil (b) des Satzes ist die Umkehrung zu von uns weiter oben besprochenen Aussagen.

Beweis. Wir zeigen zunaechst die *Eindeutigkeit* der Lebesgue-Zerlegung. Dabei verwenden wir die einfachen Eigenschaften aus der vorangehenden Proposition: Sei $\mu_{ac}^{'}, \mu_{sing}^{'}$ eine weitere Zerlegung. Dann gilt

$$\mu_{ac}^{'} - \mu_{ac} = \mu_{sing} - \mu_{sing}^{'}$$

sowie

$$\mu_{ac}^{'} - \mu_{ac} \ll \nu, \;\; \mu_{sing} - \mu_{sing}^{'} \perp \nu.$$

Damit erfuellt $\kappa := \mu_{ac}^{'} - \mu_{ac} = \mu_{sing} - \mu_{sing}^{'}$ also sowohl $\kappa \ll \nu$ als auch $\kappa \perp \nu$. Das liefert leicht $\kappa = 0$ und die gewuenschte Eindeutigkeit folgt.

Wir zeigen nun die *Existenz* der Lebesgue-Zerlegung und von h. (Damit beweisen wir die verbliebenen Teile von (a) und (b) in einem.)

Wir beginnen mit zwei leichten Vereinfachungen.

Ohne Einschraenkung sei μ ein positives beschraenktes Maß. (Andernfalls koennen wir μ schreiben als $\mu = \mu_1 - \mu_2 + i\mu_3 - i\mu_4$ und jedes dieser Teile einzeln behandeln.)
Ohne Einschraenkung sei ν endlich. (Andernfalls koennen wir ν ersetzen durch $w\nu$ mit einer strikt positiven meßbaren Funktion w mit $\int wd\nu = 1$. Dann haben ν und $w\nu$ dieselben Nullmengen und wir koennen mit $w\nu$ statt w weiterarbeiten.)

Sei das Maß ϕ definiert durch $\phi = \mu + \nu$, d.h.

$$\phi(E) = \mu(E) + \int w 1_E d\nu.$$

Dann ist ϕ ein endliches Maß und es gilt

$$(+) \quad \int f d\phi = \int f d\mu + \int f w d\nu$$

fuer alle nichtnegativen meßbaren f sowie fuer alle $f \in \mathcal{L}^1(X, \phi)$. (Aussage folgt zunaechst fuer $f = 1_E$, anschliessend fuer einfache Funktionen und dann nach Grenzuebergang fuer nichtnegative Funktionen und dann nach entsprechender Zerlegung fuer $\mathcal{L}^1$ Funktionen). Wir betrachten nun das lineare Funktional

$$L^2(X, \phi) \longrightarrow \mathbb{C}, \;\; f \mapsto \int f d\mu.$$

Wir machen uns zunaechst klar, daß dieses Funktional wohldefiniert ist: Wegen $\mu \leq \phi$ ist f tatsaechlich μ fast ueberall eindeutig bestimmt. Wegen $|f| \leq 1 + |f|^2$ und der Endlichkeit des Maßes μ und $f \in \mathrm{L}^2(X, \phi)$ gehoert f

dann zu $\mathcal{L}^1(X,\mu)$. Damit existiert $\int f d\mu$. Wir zeigen nun, daß dieses Funktional stetig ist: Fuer $f \in L^2(X,\phi)$ erhalten wir

$$\left|\int f d\mu\right| \leq \int |f| d\mu \leq \int |f| d\phi \leq \left(\int |f|^2 d\phi\right)^{1/2} \phi(X)^{1/2}.$$

Wegen $\phi(X) < \infty$ ist also die Abbildung

$$L^2(\phi) \longrightarrow \mathbb{C}, f \mapsto \int f d\mu,$$

ein stetiges Funktional. Damit existiert also nach dem Lemma von Riesz ein $g \in L^2(\phi)$ mit

$$(*) \quad \int f d\mu = \int f g d\phi.$$

Mit der Definition von ϕ erhalten wir aus $(*)$ und $(+)$ dann

$$\int f d\mu = \int f g d\mu + \int f g w d\nu.$$

(Beachte, daß fg als Produkt von zwei Funktionen aus $L^2(X,\phi)$ zu $\mathcal{L}^1(X,\phi)$ gehoert.) Durch Substraktion ergibt sich dann

$$\int f(1-g) d\mu = \int f g w d\nu$$

fuer alle $f \in L^2(X,\phi)$. Aus dieser Gleichung kann man nun schließen, dass ϕ-fast ueberall die Funktion g nur Werte in $[0,1]$ annimmt. (Gibt es etwa eine Menge A von positivem μ-Maß mit $g < 0$ auf A, so haette man nach Wahl von $f = 1_A$ dann $LS > \mu(A) > 0$ aber $RS \leq 0$. Die uebrigen Aussagen folgen analog.) Damit gilt also

$$(**) \quad (1-g)\mu = g w \nu$$

als Gleichheit von (aufgrund von $0 \leq g \leq 1$) positiven Maßen. Das ist (fast schon) die gewuenschte Zerlegung. Hier sind die Details: Wir setzen

$$A := \{x : 0 \leq g(x) < 1\}, \; B := \{x : g(x) = 1\}$$

und definieren

$$\mu_{ac}(E) := \mu(A \cap E), \;\; \mu_{sing}(E) := \mu(B \cap E).$$

Dann gilt (wegen $0 \leq g \leq 1$) also

$$\mu = \mu_{ac} + \mu_{sing}.$$

Weiterhin folgt nach Einsetzen von $f = 1_B$ und Nutzen von $g = 1$ auf B aus $(**)$ sofort

$$0 = \int_B w d\nu.$$

Wegen $w > 0$ folgt also $\nu(B) = 0$. Da μ_{sing} offenbar auf B konzentriert ist gilt also

$$\mu_{sing} \perp \nu.$$

Außerdem folgt aus $(**)$ mit $h = 1_A \frac{gw}{1-g} \geq 0$ sofort

$$\mu_{ac} = 1_A \mu = \frac{1_A}{1-g}(1-g)\mu \overset{(**)}{=} h\nu.$$

Setzt man $E = X$, so folgt $h \in \mathcal{L}^1$. Damit ist (b) beweisen. Damit folgt dann auch der noch verbliebene Teil von (a). □

Bemerkung. Die Vollstaendigkeit von $L^2(\phi)$ ist wesentlich fuer den Beweis. Sie liefert die Existenz von g.

FOLGERUNG (Radon-Nikodym). *Sei $(X, \mathcal{A})$ ein Maßraum und seien ν und μ positive σ-endliche Maße mit $\mu \ll \nu$. Dann existiert ein nichtnegatives meßbares h mit $\mu = h\nu$.*

Bemerkung. Das angegebene h gehoert (offenbar) genau dann zu $\mathcal{L}^1$, wenn μ endlich ist.

Beweis. Seien meßbare paarweise disjunkte Mengen X_n mit $\nu(X_n) < \infty$ und $\mu(X_n) < \infty$ und $X = \bigcup_n X_n$ gewaehlt. (Die Existenz solcher Mengen ist nicht schwer zu sehen: Waehle Y_n zu ν und Z_n zu μ. Ohne Einschraenkung seien Y_n und Z_n aufsteigend. Setze nun $X_n^{'} := Y_n \cap Z_n$ und betrachte $X_n := X_n^{'} \setminus X_{n-1}^{'}$.)
Dann kann man auf jedem einzelnen X_n das vorangehende Theorem anwenden: Sei μ_n die Einschraenkung von μ auf X_n d.h.

$$\mu_n(E) = \mu(X_n \cap E).$$

Dann liefert Teil (a) des vorigen Theorems $\mu_n = \mu_{ac} + \mu_{sing}$ und nach Voraussetzung ist $\mu_{sing} = 0$. Damit liefert dann Teil (b) eine Funktion $h_n \geq 0$ auf X_n mit

$$\mu_n = h_n \nu.$$

Zusammensetzen der h_n liefert dann mit $h = \sum_n h_n$ die gewuenschte Aussage gemaeß

$$\begin{aligned}
\mu(E) &= \mu(\bigcup_{n=1}^{\infty} X_n \cap E) \\
(X_n \text{ paarweise disjunkt}) \quad &= \sum_n \mu(X_n \cap E) \\
&= \sum \mu_n(E) \\
&= \sum_n \int 1_E h_n d\nu \\
(\text{Monotone Konvergenz}) \quad &= \int h 1_E d\nu.
\end{aligned}$$

Das ist die gewuenschte Aussage. □

Bemerkung. Die Voraussetzung der σ-Endlichkeit von ν ist notwendig, wie folgendes Beispiel zeigt: $\mu = \lambda$ Lebesguemaß auf $[0, 1]$ und ν das Zaehlmaß auf $[0, 1]$ (das also jedem Punkt aus $[0, 1]$ das Maß 1 zuordnet). Dann gilt $\mu \ll \nu$, aber es gibt kein h mit $\mu = h\nu$.