Funktionalanalysis

Alexander Grigoryan
Universität Bielefeld

SS 2022

Contents

Chapter 1

Normierter Vektorraum

06.04.22 **Vorlesung 1**

Die Funktionalanalysis ist Analysis in unendlichdimensionalen Vektorräumen, insbesondere in Funktionenräumen, während Analysis I und Analysis II mit endlichdimensionalen Räumen beschäftigt sind.

1.1 Vektorräume und Lineare Operatoren

Wir wiederholen die Begriffe von Vektorraum und linearen Operatoren.

Vektorraum. Ein *Vektorraum* über eine Körper $\mathbb{K}$ ist eine Menge V mit Operationen Addition

$$x, y \in V \mapsto x + y \in V$$

und skalare Multiplikation

$$\alpha \in \mathbb{K},\ x \in V \mapsto \alpha x \in V,$$

die die folgenden Axiomen erfüllen:

1. Nullvektor: es gibt ein $0 \in V$ mit $x + 0 = 0 + x = x$ für alle $x \in V$.
2. Das Negative: für jedes $x \in V$ existiert ein $-x \in V$ mit $x + (-x) = (-x) + x = 0$.
3. Assoziativgesetz für Addition: $(x + y) + z = x + (y + z)$.
4. Kommutativgesetz für Addition: $x + y = y + x$.
5. Skalarmultiplikation mit 1: $1x = x$ für alle x.
6. Assoziativgesetz für Skalarmultiplikation: $(\alpha\beta)\, x = \alpha\, (\beta x)$.
7. Distributivgesetz für Addition von Skalaren: $(\alpha + \beta)\, x = \alpha x + \beta x$.
8. Distributivgesetz für Addition von Vektoren: $\alpha\, (x + y) = \alpha x + \alpha y$.

Der Körper $\mathbb{K}$ wird immer $\mathbb{R}$ oder $\mathbb{C}$ sein.

Operator. Seien X, Y zwei Vektorräume über $\mathbb{K}$. Eine Abbildung $A : X \to Y$ heißt *linear* wenn die folgenden Identitäten gelten:

$$\begin{aligned} A(x_1 + x_2) &= A(x_1) + A(x_2) \\ \alpha A(x) &= A(\alpha x), \end{aligned}$$

für alle $x_1, x_2 \in X$ und $\alpha \in \mathbb{K}$. Lineare Abbildungen heißen auch *Operatoren.* Häufig schreibt man für Operatoren Ax anstatt $A(x)$.

Die Vektorräume X, Y heißen *isomorph* wenn es eine bijektive lineare Abbildung $A : X \to Y$ gibt. Man schreibt in diesem Fall $X \cong Y$.

Das Isomorphismus ist eine Äquivalenzrelation zwischen Vektorräumen, d.h.

1. $X \cong X$ (Id : $X \to X$ ist bijektiv und linear)
2. $X \cong Y \Rightarrow Y \cong X$ (ist $A : X \to Y$ bijektiv und linear so ist $A^{-1} : Y \to X$ bijektiv und linear)
3. $X \cong Y$ und $Y \cong Z \Rightarrow X \cong Z$ (sind $A : X \to Y$ und $B : Y \to Z$ bijektiv und linear so ist $B \circ A : X \to Z$ bijektiv und linear).

Gilt $X \cong Y$, dann sind alle Eigenschaften von Addition und skalar Multiplikation in X und Y identisch.

Unterraum und Faktorraum. Sei V ein Vektorraum. Eine Teilmenge $U \subset V$ heißt *Unterraum* wenn U geschlossen bezüglich Addition und skalar Multiplikation ist, d.h. $x, y \in U \Rightarrow x + y \in U$ und $\alpha x \in U$ $\forall \alpha \in \mathbb{K}$. Dann ist U selbst auch ein Vektorraum.

Gegeben sei ein Unterraum U von V, man definiert eine *Äquivalenzrelation* $\sim$ auf V:

$$x \sim y \text{ genau dann wenn } x - y \in U.$$

Behauptung. *Die Relation $\sim$ besitzt die folgenden Eigenschaften:*

1. $x \sim x$ (Reflexivität)

2. $x \sim y \Rightarrow y \sim x$ (Symmetrie)

3. $x \sim y$ und $y \sim z \Rightarrow x \sim z$ (Transitivität)

Beweis. 1. Wir haben $x \sim x$ da $x - x = 0 \in U$.

2. Gilt $x \sim y$ so gilt $x - y \in U$ und somit auch $y - x = (-1)(x - y) \in U$ und $y \sim x$.

3. Gelten $x \sim y$ und $y \sim z$ dann gilt $x - y \in U$ und $y - z \in U$, woraus folgt

$$x - z = (x - y) + (y - z) \in U$$

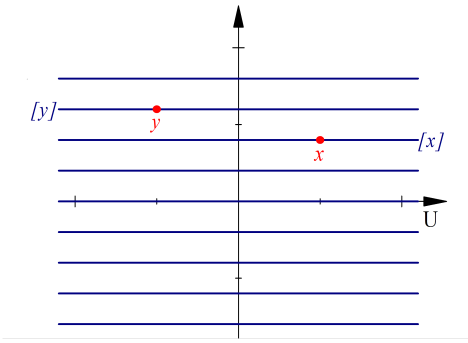
und somit $x \sim z$. ■

Jede Relation $\sim$ mit den o.g. drei Eigenschaften (Reflexivität, Symmetrie und Transitivität) heißt auch eine Äquivalenzrelation.

Für jedes $x \in V$ bezeichnen wir mit $[x]$ die *Äquivalenzklasse* von x, d.h. die folgende Teilmenge von V:

$$[x] = \{z \in V : z \sim x\}.$$

Es ist klar, dass $x \in [x]$.



Äquivalenzklassen von x und y

Behauptung. *Die Äquivalenzklassen erfüllen die folgenden Eigenschaften:*

1. $x \sim y \Leftrightarrow [x] = [y]$
2. $[x]$ *und* $[y]$ *sind entweder gleich oder disjunkt.*

Beweis. 1. Ist $x \sim y$ so gilt für jedes $z \in V$ dass $x \sim z \Leftrightarrow y \sim z$ woraus folgt, dass $z \in [x] \Leftrightarrow z \in [y]$ und somit $[x] = [y]$. Ist $[x] = [y]$ so gilt $x \in [y]$ und somit $x \sim y$ nach Definition von Äquivalenzklassen.

2. Ist der Schnitt $[x] \cap [y]$ nicht leer, so wählen wir ein $z \in [x] \cap [y]$. Dann gelten $z \sim x$ und $z \sim y$ woraus folgt $x \sim y$ und somit $[x] = [y]$. ■

Die Menge von allen Äquivalenzklassen wird mit V/U bezeichnet (manchmal benutzt man auch die Notation $V \operatorname{mod} U$). Man definiert die linearen Operationen in V/U wie folgt:

$$\begin{aligned}[x] + [y] &= [x+y] \\ \alpha [x] &= [\alpha x]\end{aligned}$$

für alle $[x], [y] \in V/U$ und $\alpha \in \mathbb{K}$.

Behauptung. *Diese Operationen sind wohldefiniert, d.h. das Ergebnis ist unabhängig von der Wahl des Vertreters der Klasse.*

Beweis. Zeigen wir, dass für alle $x' \in [x]$ und $y' \in [y]$ gilt

$$[x' + y'] = [x + y]. \tag{1.1}$$

Es gelten $x' \sim x$ und $y' \sim y$, d.h.

$$x' - x \in U \quad \text{und} \quad y' - y \in U,$$

woraus folgt

$$(x' + y') - (x + y) = (x' - x) + (y' - y) \in U.$$

Somit erhalten wir $x' + y' \sim x + y$ was äquivalent zu (1.1). Analog zeigt man, dass $[\alpha x'] = [\alpha x]$. ■

Es ist klar, dass diese Operationen in V/U alle Axiome von Vektorraum erfüllen, so dass V/U ein Vektorraum ist.

Definition. Der Vektorraum V/U heißt der *Faktorraum* (oder *Quotientenraum*) von V, U.

Die Abbildung $x \in V \mapsto [x] \in V/U$ ist eine lineare Abbildung von V nach V/U.

Dimension. Eine Folge $\{x_1, ..., x_n\}$ von Vektoren aus dem Vektorraum V heißt *linear unabhängig* wenn die einzige lineare Kombination von x_k die verschwindet, ist die triviale Kombination, d.h.

$$\alpha_1 x_1 + ... + \alpha_n x_n = 0 \Rightarrow \alpha_1 = ... = \alpha_n = 0,$$

wobei $\alpha_k \in \mathbb{K}$.

Definition. Die *Dimension* $\dim V$ wird wie folgt definiert:

$$\dim V = \sup \{n\text{: es gibt eine Folge von } n \text{ linear unabhängigen Vektoren in } V\}.$$

Insbesondere $\dim V = \infty$ wenn es in V beliebig lange Folgen von linear unabhängigen Vektoren gibt.

Beispiel. Der Raum

$$\mathbb{R}^n = \{(x_1, ..., x_n) : x_i \in \mathbb{R}\}$$

hat Dimension n. Der Raum

$$\mathbb{C}^n = \{(x_1, ..., x_n) : x_i \in \mathbb{C}\}$$

hat Dimension n als ein Vektorraum über $\mathbb{C}$ und Dimension $2n$ als ein Vektorraum über $\mathbb{R}$. Offensichtlich gilt Isomorphismus $\mathbb{C}^n \cong \mathbb{R}^{2n}$ von Vektorräumen über $\mathbb{R}$.

Beispiel. Betrachten wir die Menge

$$\mathbb{R}^\infty = \{\{x_k\}_{k=1}^\infty : x_k \in \mathbb{R}\}$$

von allen unendlichen Folgen von reellen Zahlen. Offensichtlich ist $\mathbb{R}^\infty$ ein Vektorraum über $\mathbb{R}$ bezüglich der Operationen

$$\{x_k\} + \{y_k\} = \{x_k + y_k\} \quad \text{und} \quad \alpha \{x_k\} = \{\alpha x_k\}, \quad \alpha \in \mathbb{R}.$$

Es gilt $\dim \mathbb{R}^\infty = \infty$ da es eine unendliche linear unabhängige Menge von Vektoren in $\mathbb{R}^\infty$ gibt:

$$e_1 = \{1, 0, 0, ...\}, \ e_2 = \{0, 1, 0, 0, ...\}, ..., \ e_k = \{0, ..., 0, \overset{k}{\widehat{1}}, 0, ...\}, ... \tag{1.2}$$

Betrachten wie die folgenden Teilmengen von $\mathbb{R}^\infty$:

$$l^\infty = \left\{ \{x_k\}_{k=1}^\infty : \sup_k |x_k| < \infty \right\} \text{ - die Menge von allen beschränkten Folgen}$$

und, für jedes $p > 0$,

$$l^p = \left\{ \{x_k\}_{k=1}^\infty : \sum_{k=1}^\infty |x_k|^p < \infty \right\}.$$

Es ist offensichtlich, dass l^∞ ein Unterraum ist. Die Menge l^p ist auch ein Unterraum ist (siehe Aufgabe 4). Die beiden Räumen l^∞ und l^p sind ∞-dimensional, da sie die Folge (1.2) enthalten.

Analog definiert man den Raum $\mathbb{C}^\infty$ und die komplexwertigen Versionen von l^p, l^∞.

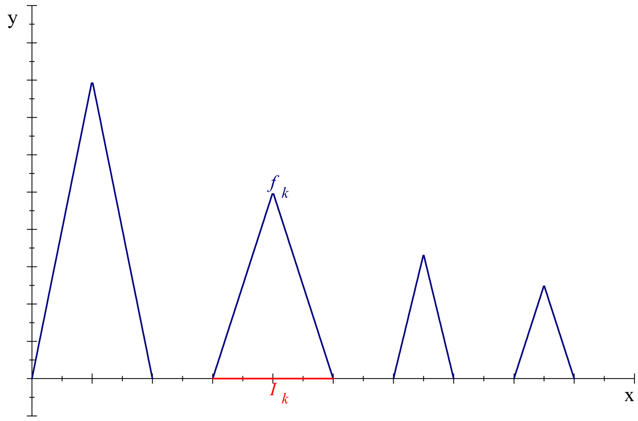
Beispiel. Sei I ein Intervall mit den Grenzpunkten $a < b$, d.h. $I = [a, b]$, oder (a, b), oder $[a, b)$ oder $(a, b]$. Bezeichnen wir mit $C(I)$ die Menge von allen stetigen reellwertigen Funktionen auf I. Die Menge $C(I)$ ist ein Vektorraum über $\mathbb{R}$ mit den folgenden Operationen:

$$\begin{aligned} (f+g)(t) &= f(t)+g(t) \\ (\alpha f)(t) &= \alpha f(t) \end{aligned}$$

für alle $t \in I$. Beweisen wir, dass

$$\dim C(I) = \infty.$$

Dafür wählen wir in I für jedes $n \in \mathbb{N}$ eine Folge $\{I_k\}_{k=1}^n$ von disjunkten Intervallen und betrachten für jedes I_k eine nicht-Null stetige Funktion f_k mit dem Träger in I_k.



Funktionen f_k

Dann ist die Folge $\{f_k\}_{k=1}^n$ linear unabhängig, da die Identität

$$\alpha_1 f_1 + ... + \alpha_n f_n = 0$$

auf I impliziert, dass für jedes I_k und für jedes $t \in I_k$ gilt

$$\alpha_k f_k(t) = \sum_{j=1}^n \alpha_j f_j(t) = 0$$

und somit $\alpha_k = 0$. Da n beliebig gross sein kann, so erhalten wir $\dim C(I) = \infty$.

Beispiel. Für jedes $m \in \mathbb{N}$ bezeichnen wir mit $C^m(I)$ die Menge von m-fach stetig differenzierbaren Funktionen auf I, was offensichtlich ein Unterraum von $C(I)$ ist. Bezeichnen wir mit $C^\infty(I)$ der Durchschnitt von allen $C^m(I)$; d.h. $C^\infty(I)$ ist der Vektorraum von allen unendlich oft differenzierbaren Funktionen auf I. Dieser Raum ist auch ∞-dimensional, d.h.

$$\dim C^\infty(I) = \infty,$$

da die Funktionen f_k im obigen Beispiel sich unendlich oft differenzierbar wählen lassen.

Beispiel. Bezeichnen with mit $\mathcal{P}_n$ die Menge von allen Polynomen des Grades $\leq n$ mit reellen Koeffizienten. D.h., jedes $f \in \mathcal{P}_n$ ist eine Funktion auf $\mathbb{R}$ der Form

$$f(t) = a_0 + a_1 t + ... + a_n t^n,$$

wobei $a_k \in \mathbb{R}$. Es ist klar, dass $\mathcal{P}_n$ ein Unterraum von $C^\infty(\mathbb{R})$ ist. Zeigen wir, dass $\mathcal{P}_n$ isomorph zu $\mathbb{R}^{n+1}$ ist. Dafür betrachten wir die folgende Abbildung:

$$\begin{aligned} \mathbb{R}^{n+1} &\to \mathcal{P}_n \\ (a_0, ..., a_n) &\mapsto a_0 + a_1 t + ... + a_n t^n. \end{aligned}$$

Diese Abbildung ist linear, surjektiv und auch injektiv da

$$a_0 + a_1 t + ... + a_n t^n \equiv 0 \;\Rightarrow\; a_0 = a_1 = ... = a_n = 0.$$

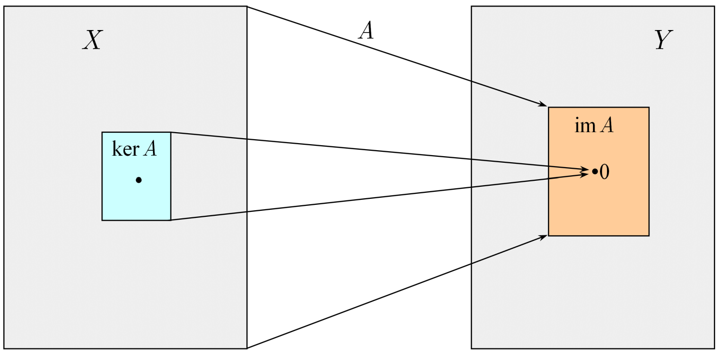
Somit gilt $\mathcal{P}_n \cong \mathbb{R}^{n+1}$ und folglich $\dim \mathcal{P}_n = n + 1$.

Kern und Bildraum. Für einen Operator $A : X \to Y$ zwischen Vektorräumen X und Y definieren wir den *Kern* von A

$$\ker A = \{x \in X : Ax = 0\}$$

und den *Bildraum* von A

$$\operatorname{im} A = \{Ax : x \in X\}.$$



Dann ist $\ker A$ ein Unterraum von X und $\operatorname{im} A$ ist ein Unterraum von Y. Die folgenden zwei Sätze werden in Linearer Algebra bewiesen (siehe auch Aufgabe 2).

Homomorphiesatz. *Es gilt*

$$\operatorname{im} A \cong X / \ker A \tag{1.3}$$

Dimensionssatz *Ist* $\dim X < \infty$, *so gilt*

$$\dim X = \dim \ker A + \dim \operatorname{im} A.$$

Beispiel. Sei M eine beliebige Menge. Bezeichnen wir mit $\mathbb{R}^M$ die Menge von allen Funktionen $f : M \to \mathbb{R}$. Offensichtlich ist $\mathbb{R}^M$ ein Vektorraum über $\mathbb{R}$. Fixieren wir n Elemente $x_1, ..., x_n$ aus M and betrachten die Abbildung $A : \mathbb{R}^M \to \mathbb{R}^n$ wie folgt: für jede Funktion $f : M \to \mathbb{R}$,

$$Af = (f(x_1), f(x_2), ..., f(x_n)) \in \mathbb{R}^n.$$

Es ist klar, dass A linear ist. Der Kern $\ker A$ besteht aus allen Funktionen auf M die in den Punkten $x_1, ..., x_n$ verschwinden, und der Bildraum $\operatorname{im} A$ ist gleich $\mathbb{R}^n$.

Beispiel. In $\mathbb{R}^\infty$ eine Verschiebung von Koordinaten:

$$\begin{aligned} A &: \mathbb{R}^\infty \to \mathbb{R}^\infty \\ A(x_1, x_2, x_3, ...) &= (x_2, x_3, ...), \end{aligned}$$

die offensichtlich ein Operator ist. Der Kern $\ker A$ besteht aus allen Folgen der Form $(x_1, 0, 0, ..)$ mit beliebigem x_1, und $\operatorname{im} A = \mathbb{R}^\infty$.

Beispiel. Betrachten wir in $C[a, b]$ einen *Integraloperator*

$$\begin{aligned} A &: C[a, b] \to \mathbb{R} \\ Af &= \int_a^b f(t)\, dt, \end{aligned}$$

der offensichtlich linear ist. Der Kern $\ker A$ besteht aus allen Funktionen mit dem Integral 0, der Bildraum $\operatorname{im} A$ ist gleich $\mathbb{R}$.

Beispiel. Fixieren wir ein $n \in \mathbb{N}$ und bemerken folgendes: für jede Funktion $f \in C^\infty(\mathbb{R})$ ist die n-te Ableitung $f^{(n)} = \frac{d^n f}{dx^n}$ auch ∞ oft differenzierbar, d.h. liegt in $C^\infty(\mathbb{R})$. Somit erhalten wir eine Abbildung

$$\begin{aligned} A &: C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R}) \\ Af &= \frac{d^n f}{dx^n}, \end{aligned}$$

die offensichtlich linear ist. Der Operator $A = \frac{d^n}{dx^n}$ heißt *Differentialoperator* der Ordnung n. Für diesen Operator gilt

$$\ker A = \mathcal{P}_{n-1} \text{ und } \operatorname{im} A = C^\infty(\mathbb{R}),$$

insbesondere $\dim \ker A = n$. Nach (1.3) erhalten wir den Isomorphismus

$$C^\infty(\mathbb{R}) / \mathcal{P}_{n-1} \cong C^\infty(\mathbb{R}).$$

1.2 Die Norm

Sei V ein Vektorraum über einen Körper $\mathbb{K}$, wobei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$.
Definition. Eine Funktion $N : V \to \mathbb{R}$ heißt *Norm* wenn sie die folgenden Axiome erfüllt:

$(N1)$ $N(x) > 0$ für $x \neq 0$ und $N(0) = 0$ (Definitheit)

$(N2)$ $N(x+y) \leq N(x) + N(y)$ $\forall x, y \in V$ (Dreiecksungleichung)

$(N3)$ $N(\alpha x) = |\alpha| N(x)$ $\forall x \in V, \alpha \in \mathbb{K}$ (absolute Homogenität)

Erfüllt N anstatt $(N1)$ die schwachere Bedingung

$(N1')$ $N(x) \geq 0$ für alle $x \in V$,

so heißt N *Halbnorm* oder *Seminorm.*
Definition. Ein *normierter Vektorraum* ist ein Paar (V, N) wobei V ein Vektorraum über $\mathbb{R}$ oder $\mathbb{C}$ ist und N eine Norm auf V ist.

Normalerweise bezeichnet man die Norm $N(x)$ als $\|x\|$.
Beispiel. In $\mathbb{R}^n$ werden die folgenden Normen häufig benutzt: die 1-Norm

$$\|x\|_1 = |x_1| + ... + |x_n|$$

und die 2-Norm (auch die *Euklidische* Norm genannt):

$$\|x\|_2 = \sqrt{x_1^2 + ... + x_n^2}.$$

Beispiel. In $C[a,b]$ betrachten wir die sup-Norm:

$$\|f\|_{C[a,b]} = \sup_{x\in[a,b]} |f(x)| = \max_{x\in[a,b]} |f(x)|.$$

Im Raum $C^l[a,b]$ mit $l \in \mathbb{N}$ benutzt man die C^l-Norm:

$$\|f\|_{C^l[a,b]} = \max_{0\leq k\leq l} \max_{x\in[a,b]} \left|f^{(k)}(x)\right|$$

(siehe Aufgabe 3).

Beispiel. Betrachten wir verschiedene Normen im Vektorraum $\mathcal{P}_n$ von Polynomen des Grades $\leq n$. Für jedes Polynom

$$f(t) = a_0 + a_1 t + ... + a_n t^n \in \mathcal{P}_n$$

kann man eine Norm aus $\mathbb{R}^{n+1}$ benutzen, zum Beispiel

$$\|f\|_1 = |a_0| + |a_1| + ... + |a_n|$$

oder

$$\|f\|_2 = \sqrt{a_0^2 + a_1^2 + ... + a_n^2}.$$

Andererseits, f ist eine stetige Funktion auf jedem Intervall $[a,b]$ mit reellen $a<b$, und wir können auch die sup-Norm auf $[a,b]$ benutzen:

$$\|f\|_{C[a,b]} = \max_{t\in[a,b]} |f(t)|.$$

Da jedes Polynom unendlich oft differenzierbar ist, so auch $\|f\|_{C^l[a,b]}$ ist wohldefiniert für jedes $l\in\mathbb{N}$ und ist eine Norm in $\mathcal{P}_n$.

Für endlichdimensionale Vektorräume gibt es die folgende Beziehung zwischen verschiedenen Normen.

Satz 1.1 *Seien N_1 und N_2 zwei Normen auf einem endlichdimensionalen Vektorraum V. Dann existieren positive Konstanten c, C mit*

$$cN_2(x) \le N_1(x) \le CN_2(x) \quad \textit{für alle } x\in V. \tag{1.4}$$

Beliebige zwei Normen N_1 und N_2 auf V heißen *äquivalent* wenn sie (1.4) für einige positive c, C erfüllen. Offensichtlich ist die Äquivalenz von Normen eine Äquivalenzrelation d.h. sie ist reflexiv, symmetrisch und transitiv.

Der Satz 1.1 bedeutet folgendes: alle Normen auf einem endlichdimensionalen Vektorraum sind äquivalent.

Beweis. Sei $n=\dim V$. Dann ist V isomorph zu $\mathbb{R}^n$, so reicht es zu beweisen, dass alle Normen auf $\mathbb{R}^n$ äquivalent sind. Darüber hinaus reicht es zu beweisen, dass eine beliebige Norm $N(x)$ auf $\mathbb{R}^n$ zur 1-Norm $\|x\|_1$ äquivalent ist, wobei

$$\|x\|_1 = |x_1| + ... + |x_n|.$$

Betrachten wir die Basisvektoren

$$e_k = (0,...,0,\overset{k}{\widehat{1}},0,...,0),$$

wobei die einzige Eins auf der Position k steht. Für jedes $x=(x_1,...,x_n)\in\mathbb{R}^n$ erhalten wir mit Hilfe von Axiomen $(N2)$ und $(N3)$

$$N(x) = N(x_1e_1+...+x_ne_n) \le |x_1|\,N(e_1)+...+|x_n|\,N(e_n) \le M(|x_1|+...+|x_n|),$$

wobei $M=\max_{1\le k\le n} N(e_k)$. Somit erhalten wir, dass

$$N(x) \le M\|x\|_1. \tag{1.5}$$

Es folgt aus (1.5), dass die Funktion $N:\mathbb{R}^n\to\mathbb{R}$ stetig, weil

$$|N(x)-N(y)| \le N(x-y) \le M\|x-y\|_1.$$

Betrachten die Einheitssphäre

$$S = \{x\in\mathbb{R}^n : \|x\|_1 = 1\},$$

die eine beschränkte abgeschlossene Teilmenge von $\mathbb{R}^n$ ist. Nach dem Extremwertsatz erreicht die stetige Funktion N das Maximum und Minimum auf S. Setzen wir

$$C = \max_S N \quad \text{und} \quad c = \min_S N.$$

Da $0 \notin S$ und somit $N(x) > 0$ auf S, so haben wir

$$C \geq c > 0.$$

Für jedes $x \in \mathbb{R}^n \setminus 0$ gilt $\frac{x}{\|x\|_1} \in S$ woraus folgt

$$c \leq N\left(\frac{x}{\|x\|_1}\right) \leq C$$

und somit

$$c\|x\|_1 \leq N(x) \leq C\|x\|_1. \tag{1.6}$$

Für $x = 0$ gilt diese Ungleichung trivialerweise. Nach (1.6) sind N und die 1-Norm äquivalent, was zu beweisen war. ∎

08.04.22 Vorlesung 2

Für jedes $p \in [1, \infty)$ definieren with die p-Norm von $x \in \mathbb{R}^n$ mit

$$\|x\|_p = (|x_1|^p + ... + |x_n|^p)^{1/p}$$

und für $p = \infty$ die ∞-Norm mit

$$\|x\|_\infty = \max(|x_1|, ..., |x_n|).$$

Für $p = 1$ und $p = 2$ wurden diese Normen oberhalb schon definiert. Der Grund für die Notation $\|x\|_\infty$ ist die Konvergenz

$$\|x\|_p \to \|x\|_\infty \text{ für } p \to \infty$$

(siehe Aufgabe 14). Für jedes $p \in [1, \infty)$ betrachten wir den Vektorraum

$$l^p = \left\{ \{x_k\}_{k=1}^\infty : x_k \in \mathbb{R} \text{ und } \sum_{k=1}^\infty |x_k|^p < \infty \right\}$$

und definieren in l^p die p-Norm mit

$$\|x\|_p = \left(\sum_{k=1}^\infty |x_k|^p \right)^{1/p}.$$

Im Vektorraum

$$l^\infty = \left\{ \{x_k\}_{k=1}^\infty : x_k \in \mathbb{R} \text{ und } \sup_{k \in \mathbb{N}} |x_k| < \infty \right\}$$

definieren wir die ∞-Norm mit

$$\|x\|_\infty = \sup_{k \in \mathbb{N}} |x_k|.$$

Der Raum l^p lässt sich betrachten als eine unendlichdimensionale Erweiterungen von $\mathbb{R}^n$ bezüglich der p-Norm.

Satz 1.2 *Die p-Norm ist eine Norm in $\mathbb{R}^n$ und in l^p für jedes $p \in [1, \infty]$.*

Anfang von dem Beweis. Die Eigenschaften $(N1)$ und $(N3)$ für die p-Norm sind trivialerweise erfüllt. In der Tat, die p-Norm $\|x\|_p$ ist immer nichtnegativ und verschwindet genau dann, wenn alle Komponenten x_k gleich Null sind, d.h. wenn $x = 0$. Die absolute Homogenität beweist man wie folgt: für $\alpha \in \mathbb{R}$, $x \in \mathbb{R}^n$ bzw $x \in l^p$ und für $p \in [1, \infty)$ erhalten wir

$$\|\alpha x\|_p = \left(\sum_k |\alpha x_k|^p\right)^{1/p} = \left(\sum_k |\alpha|^p |x_k|^p\right)^{1/p} = |\alpha| \left(\sum_k |x_k|^p\right)^{1/p} = |\alpha| \, \|x\|_p \, ,$$

während für $p = \infty$ gilt

$$\|\alpha x\|_\infty = \sup_k |\alpha x_k| = |\alpha| \sup_k |x_k| = |\alpha| \, \|x\|_\infty \, .$$

Die Hauptsache ist die Dreiecksungleichung $(N2)$ zu beweisen, d.h.

$$\|x + y\|_p \leq \|x\|_p + \|y\|_p \tag{1.7}$$

für alle x, y aus $\mathbb{R}^n$ bzw l^p. Die Ungleichung (1.7) heißt *Minkowski-Ungleichung*. Wir werden sie mit Hilfe von der Hölder-Ungleichung beweisen. ■

Jedes Paar $p, q \in [1, \infty]$ heißt *konjugierte Hölder-Exponenten* wenn

$$\frac{1}{p} + \frac{1}{q} = 1. \tag{1.8}$$

Z.B. die Paaren $1, \infty$ und $2, 2$ sind konjugierte Hölder-Exponenten.

Lemma 1.3 (Hölder-Ungleichung) *Für alle $x, y \in \mathbb{R}^n$ gilt*

$$|x \cdot y| \leq \|x\|_p \, \|y\|_q \tag{1.9}$$

wobei $x \cdot y = \sum_{k=1}^n x_k y_k$ das Skalarprodukt in $\mathbb{R}^n$ ist, und p, q konjugierte Hölder-Exponenten sind.

Z.B. für $p = q = 2$ erhalten wir die *Cauchy-Schwarz-Ungleichung*

$$|x \cdot y| \leq \|x\|_2 \, \|y\|_2 \, .$$

Beweis. Für $p = 1$ und $q = \infty$ haben wir

$$|x \cdot y| = |x_1 y_1 + ... + x_n y_n| \leq \sum_{k=1}^n |x_k| \, |y_k| \leq \left(\sum_{k=1}^n |x_k|\right) \max_k |y_k| = \|x\|_1 \, \|y\|_\infty \, .$$

Der gleiche Argument gilt für $p = \infty$ und $q = 1$. Jetzt nehmen wir an, dass $1 < p < \infty$ und daher $1 < q < \infty$. Gilt $x = 0$ oder $y = 0$ so ist (1.9) trivial. Nehmen wir an, dass $x, y \neq 0$.

Wenn x in der Ungleichung (1.9) durch αx ersetzt wird (wobei $\alpha \in \mathbb{R}$), so werden die beiden Seiten mit $|\alpha|$ multipliziert, da

$$|(\alpha x) \cdot y| = |\alpha (x \cdot y)| = |\alpha| \, |x \cdot y|$$

und

$$\|\alpha x\|_p \|y\|_q = |\alpha| \|x\|_p \|y\|_q .$$

Somit können wir in (1.9) x durch αx mit einem beliebigen $\alpha \in \mathbb{R} \setminus \{0\}$ ersetzen ohne die Ungleichung zu ändern. Wählen wir $\alpha = \frac{1}{\|x\|_p}$ so dass $\|\alpha x\|_p = 1$. Dann benennen wir αx in x um und somit nehmen ohne Beschränkung der Allgemeinheit an, dass $\|x\|_p = 1$. Analog nehmen wir an, dass $\|y\|_q = 1$.

Weiter benutzen wir die *Young-Ungleichung*

$$\frac{a^p}{p} + \frac{b^q}{q} \geq ab, \tag{1.10}$$

die unter der Bedingung (1.8) für alle $a, b \geq 0$ gilt. Für $a = |x_k|$ und $b = |y_k|$ erhalten wir

$$\frac{|x_k|^p}{p} + \frac{|y_k|^q}{q} \geq |x_k| |y_k|$$

und somit

$$\sum_{k=1}^{n} \left(\frac{|x_k|^p}{p} + \frac{|y_k|^q}{q} \right) \geq \sum_{k=1}^{n} |x_k| |y_k| \geq |x \cdot y| . \tag{1.11}$$

Mit Hilfe von $\|x\|_p = \|y\|_q = 1$ und (1.8) erhalten wir, dass die linke Seite von (1.11) ist

$$\frac{1}{p} \sum_{k=1}^{n} |x_k|^p + \frac{1}{q} \sum_{k=1}^{n} |y_k|^q = \frac{1}{p} \|x\|_p^p + \frac{1}{q} \|y\|_q^q = \frac{1}{p} + \frac{1}{q} = 1 = \|x\|_p \|y\|_q,$$

woraus (1.9) folgt. ∎

Abschluss des Beweises von dem Satz 1.2. Jetzt beweisen wir die Minkowski-Ungleichung (1.7), d.h. für alle $x, y \in \mathbb{R}^n$ bzw $x, y \in l^p$

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

Sei zuerst $p = 1$. Für $x, y \in \mathbb{R}^n$ bzw $x, y \in l^1$ gilt

$$\|x + y\|_1 = \sum_k |x_k + y_k| \leq \sum_k (|x_k| + |y_k|) = \sum_k |x_k| + \sum_k |y_k| = \|x\|_1 + \|y\|_1 .$$

Im Fall $p = \infty$, für $x, y \in \mathbb{R}^n$ bzw $x, y \in l^\infty$ erhalten wir

$$\|x + y\|_\infty = \sup_k |x_k + y_k| \leq \sup_k (|x_k| + |y_k|) \leq \sup_k |x_k| + \sup_k |y_k| = \|x\|_\infty + \|y\|_\infty .$$

Sei jetzt $1 < p < \infty$. Beweisen wir zuerst (1.7) für $x, y \in \mathbb{R}^n$. Da

$$|x_k + y_k| \leq |x_k| + |y_k| ,$$

so benennen wir $|x_k|$ in x_k und $|y_k|$ in y_k um und nehmen ohne Beschränkung der Allgemeinheit an, dass $x_k \geq 0$ und $y_k \geq 0$. Wir haben dann

$$\begin{aligned} \|x + y\|_p^p = \sum_{k=1}^{n} (x_k + y_k)^p &= \sum_{k=1}^{n} (x_k + y_k)(x_k + y_k)^{p-1} \\ &= \sum_{k=1}^{n} x_k (x_k + y_k)^{p-1} + \sum_{k=1}^{n} y_k (x_k + y_k)^{p-1} . \end{aligned} \tag{1.12}$$

Setzen wir $z_k = (x_k + y_k)^{p-1}$ und bemerken, dass nach der Hölder-Ungleichung gilt

$$\sum_{k=1}^{n} x_k (x_k + y_k)^{p-1} = \sum_{k=1}^{n} x_k z_k \leq \|x\|_p \|z\|_q$$

und

$$\sum_{k=1}^{n} y_k (x_k + y_k)^{p-1} = \sum_{k=1}^{n} y_k z_k \leq \|y\|_p \|z\|_q ,$$

wobei $q = \frac{p}{p-1}$ der zu p konjugierte Hölder-Exponent ist. Es gilt

$$\|z\|_q = \left(\sum_{k=1}^{n} (x_k + y_k)^{(p-1)q} \right)^{\frac{1}{q}} = \left(\sum_{k=1}^{n} (x_k + y_k)^{p} \right)^{\frac{1}{p}\frac{p}{q}} = \|x+y\|_p^{p/q} ,$$

wobei wir die Identität $(p-1)\,q = p$ benutzt haben. Somit erhalten wir

$$\sum_{k=1}^{n} x_k (x_k + y_k)^{p-1} \leq \|x\|_p \|x+y\|_p^{p/q}$$

und analog

$$\sum_{k=1}^{n} y_k (x_k + y_k)^{p-1} \leq \|y\|_p \|x+y\|_p^{p/q} .$$

Addieren diese Ungleichungen und Einsetzen in (1.12) ergibt

$$\|x+y\|_p^p \leq \left(\|x\|_p + \|y\|_p \right) \|x+y\|_p^{p/q} .$$

Daraus folgt, dass

$$\|x+y\|_p^{p-p/q} \leq \|x\|_p + \|y\|_p .$$

Da

$$p - \frac{p}{q} = p \left(1 - \frac{1}{q} \right) = p \cdot \frac{1}{p} = 1,$$

so erhalten wir (1.7).

Für $x, y \in l^p$ erhalten wir für jedes $n \in \mathbb{N}$ nach der Minkowski-Ungleichung in $\mathbb{R}^n$

$$\left(\sum_{k=1}^{n} |x_k + y_k|^p \right)^{1/p} \leq \left(\sum_{k=1}^{n} |x_k|^p \right)^{1/p} + \left(\sum_{k=1}^{n} |y_k|^p \right)^{1/p} .$$

Für $n \to \infty$ erhalten die Minkowski-Ungleichung in l^p. ∎

1.3 Metrik und Topologie in normierten Vektorräumen

Jetzt wiederholen wir die Definitionen einer Metrik und eines metrischen Raums.

Definition. Sei X eine Menge. Eine Funktion $d : X \times X \to \mathbb{R}$ heißt *Metrik* (oder *Abstandsfunktion*), wenn sie die folgenden Axiomen von Metrik erfüllt:

$(D1)$ $d(x,y) > 0$ wenn $x \neq y$ und $d(x,x) = 0$ (Definitheit)

$(D2)$ $d(x,y) \leq d(x,z) + d(y,z)$ (Dreiecksungleichung)

$(D3)$ $d(x,y) = d(y,x)$ (Symmetrie).

Das Paar (X,d) heißt *metrischer Raum.* Im metrischen Raum definiert man Konvergenz einer Folge $\{x_n\}_{n\in\mathbb{N}}$ gegen $x \in X$ wie folgt:

$$x_n \to x \text{ für } n \to \infty \ \Leftrightarrow \ d(x_n, x) \to 0 \text{ für } n \to \infty.$$

Man schreibt auch $x_n \xrightarrow{d} x$ und

$$x = d\text{-}\lim_{n\to\infty} x_n = \lim_{n\to\infty} x_n.$$

Mit Hilfe von Konvergenz definiert man abgeschlossene Mengen.

Definition. Eine Menge $F \subset X$ heißt *abgeschlossen* wenn für jede konvergente Folge $\{x_n\}$ aus F gilt $\lim x_n \in F$.

Mit Hilfe von Metrik definiert man auch *offene* Mengen wie folgt. Bezeichnen wir mit $B(x,r)$ eine metrische Kugel in (X,d) mit dem Mittelpunkt x und Radius r, d.h.

$$B(x,r) = \{y \in X : d(x,y) < r\}.$$

Definition. Eine Menge $\Omega \subset X$ heißt *offen*, wenn

$$\forall x \in \Omega \ \ \exists r > 0 \quad \text{mit} \quad B(x,r) \subset \Omega.$$

Behauptung. *Eine Menge $\Omega \subset X$ ist offen genau dann wenn das Komplement $\Omega^c = X \backslash \Omega$ abgeschlossen ist.*

Das Mengensystem von offenen Mengen erfüllt die Axiomen von *Topologie.*

Definition. Eine Topologie auf X ist ein Mengensystem $\mathcal{O} \subset \mathcal{P}(X)$, die die folgenden Axiomen erfüllt:

$(O1)$ $\emptyset$ und X gehören zu $\mathcal{O}$.

$(O2)$ Für beliebige Familie $\{\Omega_\alpha\}_{\alpha\in I}$ von Mengen $\Omega_\alpha \in \mathcal{O}$ auch die Vereinigung $\bigcup_{\alpha\in I} \Omega_\alpha$ gehört zu $\mathcal{O}$.

$(O3)$ Für jede endliche Familie $\{\Omega_k\}_{k=1}^n$ von Mengen $\Omega_k \in \mathcal{O}$ auch der Durchschnitt $\bigcap_{k=1}^n \Omega_k$ gehört zu $\mathcal{O}$.

Die Elemente von $\mathcal{O}$ heißen offene Mengen. Das Paar $(X,\mathcal{O})$ heißt *topologischer Raum.*

Behauptung. *Das Mengensystem von den oberhalb definierten offenen Mengen im metrischen Raum (X,d) erfüllt die Axiome von Topologie.*

Somit is jeder metrische Raum (X,d) auch ein topologischer Raum.

Sei jetzt V ein normierter Vektorraum mit einer Norm $\|\cdot\|$. Die Norm induziert eine Metrik d auf V durch die Festlegung

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

für alle $x, y \in V$, die offensichtlich alle Axiome von Metrik erfüllt. Somit ist (V,d) ein metrischer Raum, sowie auch ein topologischer Raum. Die von der Norm erzeugte Topologie heißt die *Norm-Topologie*. Die Konvergenz in (V,d) sieht wie folgt aus:

$$x_n \to x \iff \|x_n - x\| \to 0.$$

Behauptung. *Seien N_1 und N_2 zwei äquivalente Normen auf V. Dann erzeugen N_1 und N_2 dieselbe Topologie in V.*

Beweis. Die Normen N_1 und N_2 erfüllen *(1.4)* woraus folgt dass

$$x_n \stackrel{N_1}{\to} x \iff x_n \stackrel{N_2}{\to} x.$$

Der Begriff von Konvergenz bestimmt eindeutig abgeschlossene Mengen, und damit auch die offenen Mengen. Somit stimmen die Begriffe von offenen Mengen bezüglich N_1 und N_2 überein, d.h. N_1 und N_2 bestimmen dieselbe Topologie. $\blacksquare$

Es folgt aus dem Satz 1.1, dass in einem endlichdimensionalen Vektorraum V nur *eine* Norm-Topologie gibt. Allerdings ist es *nicht* der Fall für die ∞-dimensionalen Vektorräume.

Beispiel. Betrachten wir den Vektorraum $C[a,b]$ mit der sup-Norm. Dann ist die Konvergenz $f_n \to f$ in diesem Raum stimmt mit der gleichmäßigen Konvergenz $f_n \rightrightarrows f$ auf $[a,b]$ überein.

Beispiel. Betrachten wir im Vektorraum $C^1[0,1]$ zwei Normen:

$$N_1(f) = \|f\|_{C[0,1]} = \max_{t\in[0,1]} |f(t)|$$

und

$$N_2(f) = \|f\|_{C^1[0,1]} = \max\left(\max_{t\in[0,1]} |f(t)|, \max_{t\in[0,1]} |f'(t)|\right).$$

Für die Folge $f_n(t) = \frac{1}{n}\sin nt$ gilt offensichtlich $N_1(f) = \frac{1}{n}$ und somit $f_n \to 0$ bezüglich N_1. Andererseits $f_n' = \cos nt$ und $N_2(f_n) = 1$, woraus folgt, dass $f_n \not\to 0$ bezüglich N_2. Somit erzeugen N_1 und N_2 zwei verschiedene Topologien, insbesondere sind N_1 und N_2 nicht äquivalent.

1.4 Vollständigkeit

Sei (X,d) ein metrischer Raum.

Definition. Eine Folge $\{x_n\}_{n\in\mathbb{N}}$ von Elementen von X heißt *Cauchy-Folge* (oder Fundamentalfolge) wenn $d(x_k, x_m) \to 0$ für $k, m \to \infty$. Der metrische Raum (X,d) heißt *vollständig* wenn jede Cauchy-Folge konvergiert.

Definition. Ein normierter Vektorraum $(V, \|\cdot\|)$ heißt *Banachraum* wenn der metrische Raum (V,d) mit der Metrik $d(x,y) = \|x-y\|$ vollständig ist.

Man benutzt Vollständigkeit von den Räumen um Existenz von Lösungen von bestimmten Gleichungen zu beweisen. Eine mögliche Methode zur Nutzung der Vollständigkeit ist der Fixpunktsatz von Banach aus Analysis II.

Satz 1.4 (Fixpunktsatz von Banach) *Seien (X, d) ein vollständiger metrischer Raum und $T : X \to X$ eine Kontraktionsabbildung, d.h. es gibt ein $q \in (0, 1)$ mit*

$$d\left(T\left(x\right), T\left(y\right)\right) \leq q d\left(x, y\right) \text{ für alle } x, y \in X.$$

Dann besitzt T genau einen Fixpunkt, d.h. es gibt genau ein $x \in X$ mit $T\left(x\right) = x$.

Eine andere Methode ist das folgende *Majorantenkriterium* für Konvergenz von Reihen.

Satz 1.5 *Seien V ein Banachraum und $\{x_n\}$ eine Folge von Elementen von V. Dann gilt die folgende Implikation:*

$$\sum_{n=1}^{\infty} \|x_n\| < \infty \Rightarrow \sum_{n=1}^{\infty} x_n \text{ konvergiert.}$$

Beweis. Betrachten wir die Partialsummen

$$S_k = \sum_{n=1}^{k} x_n.$$

Für $m > k$ haben wir

$$S_m - S_k = \sum_{n=k+1}^{m} x_n$$

und nach der Dreiecksungleichung

$$\|S_m - S_k\| \leq \sum_{n=k+1}^{m} \|x_n\|.$$

Da $\sum_{n=1}^{\infty} \|x_n\| < \infty$, so gilt

$$\sum_{n=k+1}^{m} \|x_n\| \to 0 \text{ für } k, m \to \infty,$$

woraus folgt, dass $\|S_m - S_k\| \to 0$ für $k, m \to \infty$, d.h. $\{S_k\}$ ist eine Cauchy-Folge. Nach der Vollständigkeit von V existiert $\lim S_k$ und somit ist die Reihe $\sum_{k=1}^{\infty} x_k$ konvergent. ■

Jetzt beweisen wir die Vollständigkeit von einigen Räumen.

Satz 1.6 *Die folgenden normierten Vektorräume über $\mathbb{R}$ sind Banachräume.*

(a) *Jeder endlichdimensionale Vektorraum.*

(b) *Der Raum l^p mit der p-Norm, für jedes $p \in [1, \infty]$.*

Beweis. (a) Da jeder endlichdimensionale Vektorraum V isomorph zur $\mathbb{R}^n$ ist, so reicht es zu beweisen, dass $\mathbb{R}^n$ mit einer beliebigen Norm vollständig ist. Es ist bekannt aus Analysis II, dass $\mathbb{R}^n$ mit der 2-Norm vollständig ist. Da jede zwei Normen in $\mathbb{R}^n$ äquivalent sind, so folgt es, dass $\mathbb{R}^n$ mit beliebiger Norm vollständig ist.

13.04.22 Vorlesung 3

(b) Sei $p \in [1, \infty)$. Sei $\left\{x^{(k)}\right\}_{k\in\mathbb{N}}$ eine Folge von Elementen von l^p:

$$\begin{aligned} x^{(1)} &= \left\{x_1^{(1)}, x_2^{(1)}, ..., x_i^{(1)}, ...\right\} \\ x^{(2)} &= \left\{x_1^{(2)}, x_2^{(2)}, ..., x_i^{(2)}, ...\right\} \\ &\ ... \\ x^{(k)} &= \left\{x_1^{(k)}, x_2^{(k)}, ..., x_i^{(k)}, ...\right\} \\ &\ ... \end{aligned}$$

Nehmen wir an, dass $\left\{x^{(k)}\right\}_{k\in\mathbb{N}}$ eine Cauchy-Folge ist, d.h.

$$\left\|x^{(k)} - x^{(m)}\right\|_p \to 0 \ \text{ für } k, m \to \infty.$$

Fixieren wir ein $i \in \mathbb{N}$. Da für jedes $x \in l^p$ gilt $|x_i| \le \|x\|_p$, so erhalten wir

$$\left|x_i^{(k)} - x_i^{(m)}\right| \to 0 \text{ für } k, m \to \infty.$$

Da $\mathbb{R}$ vollständig ist, so existiert der Grenzwert $\lim_{k\to\infty} x_i^{(k)} \in \mathbb{R}$ für jedes i. Bezeichnen wir diesen Grenzwert mit x_i, d.h.

$$x_i := \lim_{k\to\infty} x_i^{(k)}.$$

Somit bekommen wir eine Folge

$$x := \{x_i\}_{i\in\mathbb{N}} = \{x_1, x_2, ..., x_i, ...\}.$$

Es reicht folgendes zu beweisen:

(i) $x \in l^p$

(ii) $x^{(k)} \to x$ in l^p, d.h. $\left\|x^{(k)} - x\right\|_p \to 0$.

Da $\left\{x^{(k)}\right\}_{k=1}^{\infty}$ eine Cauchy-Folge ist, so gilt folgendes:

$$\forall \varepsilon > 0 \ \ \exists N \quad \forall k, m \ge N \quad \left\|x^{(k)} - x^{(m)}\right\|_p \le \varepsilon, \tag{1.13}$$

d.h.

$$\sum_{i=1}^{\infty} \left|x_i^{(k)} - x_i^{(m)}\right|^p \le \varepsilon^p.$$

Insbesondere gilt für jedes n (und für alle $k, m \geq N$)

$$\sum_{i=1}^{n} \left| x_i^{(k)} - x_i^{(m)} \right|^p \leq \varepsilon^p.$$

Für $m \to \infty$ gilt $x_i^{(m)} \to x_i$ und somit

$$\sum_{i=1}^{n} \left| x_i^{(k)} - x_i \right|^p \leq \varepsilon^p.$$

Für $n \to \infty$ erhalten wir

$$\sum_{i=1}^{\infty} \left| x_i^{(k)} - x_i \right|^p \leq \varepsilon^p. \tag{1.14}$$

Es folgt aus (1.14) dass

$$\left\| x^{(k)} - x \right\|_p < \infty$$

und somit $x^{(k)} - x \in l^p$. Da $x^{(k)} \in l^p$, so erhalten wir

$$x = x^{(k)} - \left(x^{(k)} - x \right) \in l^p.$$

Es folgt aus (1.13) und (1.14), dass

$$\forall \varepsilon > 0 \ \ \exists N \quad \forall k \geq N \quad \left\| x^{(k)} - x \right\|_p \leq \varepsilon,$$

und somit $\left\| x^{(k)} - x \right\|_p \to 0$ für $k \to \infty$.

Der Fall $p = \infty$ wird analog behandelt (siehe Aufgabe 11). ■

Satz 1.7 *Der Raum $C^n[a, b]$ mit der $C^n[a, b]$-Norm ist ein Banachraum, für alle reelle $a < b$ und $n \in \mathbb{N} \cup \{0\}$.*

Beweis. Betrachten wir zuerst den Fall $n = 0$ und beweisen, dass jede Cauchy-Folge $\{f_k\}_{k \in \mathbb{N}}$ in $C[a, b]$ konvergiert. Wir haben

$$\sup_{t \in [a,b]} |f_k(t) - f_m(t)| \to 0 \ \text{ für } k, m \to \infty.$$

Folglich gilt für jedes $t \in [a, b]$

$$|f_k(t) - f_m(t)| \to 0 \ \text{ für } k, m \to \infty$$

so dass die numerische Folge $\{f_k(t)\}_{k=1}^{\infty}$ eine Cauchy-Folge ist. Dann existiert für jedes $t \in [a, b]$ der Grenzwert

$$f(t) := \lim_{k \to \infty} f_k(t).$$

Beweisen wir folgendes:

(i) $f \in C[a, b]$

(ii) $\|f_k - f\|_{C[a,b]} \to 0$, d.h. $f_k \rightrightarrows f$ auf $[a, b]$.

Nach der Cauchy-Bedingung haben wir

$$\forall \varepsilon > 0 \; \exists N \quad \forall k, m \geq N \quad \sup_{t \in [a,b]} |f_k(t) - f_m(t)| \leq \varepsilon.$$

Insbesondere gilt für jedes $t \in [a, b]$

$$|f_k(t) - f_m(t)| \leq \varepsilon.$$

Für $m \to \infty$ erhalten wir

$$|f_k(t) - f(t)| \leq \varepsilon,$$

woraus folgt, dass auch

$$\sup_{t \in [a,b]} |f_k(t) - f(t)| \leq \varepsilon.$$

Deshalb erhalten wir die gleichmäßige Konvergenz

$$f_k \rightrightarrows f \quad \text{auf } [a, b].$$

Es ist bekannt, dass der gleichmäßige Grenzwert der Folge von stetigen Funktionen immer stetig ist. Somit gilt $f \in C[a, b]$ und dann auch

$$\|f_k - f\|_{C[a,b]} \to 0.$$

Jetzt beweisen wir, dass $C^n[a, b]$ vollständig für alle $n \in \mathbb{N}$ ist. Sei $\{f_k\}_{k \in \mathbb{N}}$ eine Cauchy-Folge in $C^n[a, b]$, d.h.

$$\|f_k - f_m\|_{C^n[a,b]} \to 0 \quad \text{für } k, m \to \infty.$$

Dann gilt für jedes $i = 0, ..., n$

$$\left\| f_k^{(i)} - f_m^{(i)} \right\|_{C[a,b]} \to 0 \quad \text{für } k, m \to \infty,$$

d.h. die Folge $\left\{ f_k^{(i)} \right\}_{k=1}^{\infty}$ eine Cauchy-Folge in $C[a, b]$ ist. Somit konvergiert diese Folge in $C[a, b]$ für jedes $i = 0, ..., n$. Setzen wir:

$$f_k \rightrightarrows f \quad \text{für } k \to \infty \qquad (\text{Fall } i = 0)$$

und, für jedes $i = 1, ..., n$,

$$f_k^{(i)} \rightrightarrows g_i \quad \text{für } k \to \infty \qquad (\text{Fall } i \geq 1).$$

Benutzen wir die folgende Behauptung.

Behauptung. *Gilt $f_k \rightrightarrows f$ und $f_k' \rightrightarrows g$ auf $[a, b]$ für stetige Funktionen f, g, so ist f differenzierbar und es gilt $f' = g$.*

In der Tat haben wir

$$f_k(t) = f_k(a) + \int_a^t f_k'(s)\, ds$$

woraus nach $k \to \infty$ folgt

$$f(t) = f(a) + \int_a^t g(s)\, ds$$

und somit $f' = g$.

Es folgt aus dieser Behauptung, dass $f' = g_1$, $g_1' = g_2$ usw., und somit $f^{(i)} = g_i$ für alle $i = 1, ..., n$. Folglich erhalten wir, dass

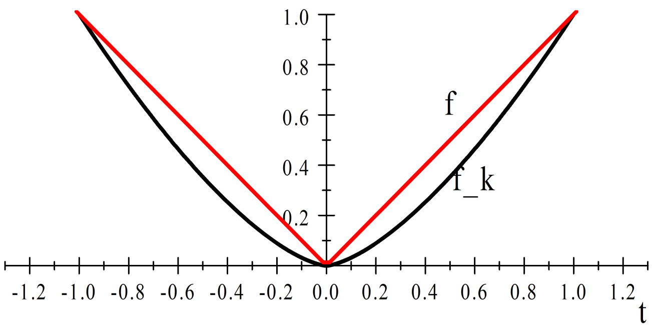
$$f_k^{(i)} \rightrightarrows f^{(i)} \quad \text{für } k \to \infty.$$

Wir beschließen, dass $f \in C^n[a,b]$ und

$$\|f_k - f\|_{C^n[a,b]} \to 0,$$

was zu beweisen war. ∎

Beispiel. Zeigen wir, dass der Raum $C^1[-1,1]$ mit der sup-Norm *nicht* vollständig ist. Dafür betrachten wir die Folge $f_k(t) = |t|^{1+\frac{1}{k}}$:



Konvergenz $f_k \rightrightarrows f$

Diese Folge liegt in $C^1[-1,1]$ und konvergiert gleichmäßig gegen $f(t) = |t|$ für $k \to \infty$. Somit ist $\{f_k\}$ eine Cauchy-Folge bezüglich der sup-Norm, aber der Grenzwert in $C^1[-1,1]$ existiert nicht weil $f \notin C^1[-1,1]$. Allerdings existiert der Grenzwert im größeren Raum $C[-1,1]$.

Beispiel. Als Beispiel von Anwendung des Satzes 1.5 (Majorantenkriterium) untersuchen wir die Summe einer cos-Fourier-Reihe

$$f(t) = \sum_{k=1}^{\infty} a_k \cos kt \tag{1.15}$$

wobei $a_k \in \mathbb{R}$. Es reicht diese Reihe auf der Periode $[0, 2\pi]$ zu betrachten. Wir haben

$$\|\cos kt\|_{C[0,2\pi]} = \sup_{t \in [0,2\pi]} |\cos kt| = 1$$

und somit

$$\|a_k \cos kt\|_{C[0,2\pi]} = |a_k|.$$

Anwendung von dem Majorantenkriterium im Banachraum $C[0,2\pi]$ ergibt folgendes: konvergiert die Reihe $\sum_{k=1}^{\infty} |a_k|$ so konvergiert die Reihe (1.15) im $C[0,2\pi]$ und somit ist ihre Summe f eine stetige Funktion, d.h.

$$\sum_{k=1}^{\infty} |a_k| < \infty \ \Rightarrow \ f \text{ ist stetig.}$$

Betrachten wir jetzt die Reihe (1.15) im $C^1[0,2\pi]$. Wir haben

$$(\cos kt)' = -k \sin kt$$

und somit

$$\sup_{t\in[0,2\pi]} \left|(\cos kt)'\right| = k$$

und

$$\|a_k \cos kt\|_{C^1[0,2\pi]} = k\,|a_k|\,.$$

Anwendung von dem Majorantenkriterium im Banachraum $C^1[0,2\pi]$ ergibt folgendes: konvergiert die Reihe $\sum_{k=1}^{\infty} k\,|a_k|$ so konvergiert die Reihe (1.15) im $C^1[0,2\pi]$ und somit ist die Summe f stetig differenzierbar,

Analog beweist man folgendes: konvergiert die Reihe $\sum_{k=1}^{\infty} k^n\,|a_k|$ so konvergiert die Reihe (1.15) im $C^n[0,2\pi]$ und somit ist die Summe f n-fach stetig differenzierbar, d.h.

$$\sum_{k=1}^{\infty} k^n\,|a_k| < \infty \ \Rightarrow\ f \text{ ist } n\text{-fach stetig differenzierbar.}$$

Diese Bedingung ist erfüllt, z.B. für $a_n = \frac{1}{k^{n+1+\varepsilon}}$ wobei $\varepsilon > 0$.

1.5 Satz von Picard-Lindelöf

Als Beispiel von Anwendung des Begriffes von Vollständigkeit zusammen mit dem Fixpunktsatz von Banach beweisen wir die Existenz von Lösung der gewöhnlichen Differentialgleichung

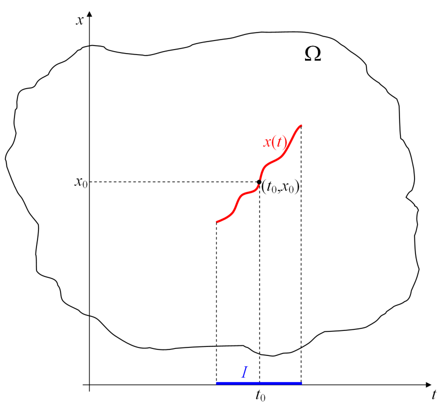
$$x' = f(t,x)\,. \tag{1.16}$$

Here $f(t,x)$ ist eine gegebene Funktion die in einer offenen Teilmenge Ω von $\mathbb{R}^2$ definiert ist, und $x(t)$ – eine unbekannte Funktion.

Definition. Sei $I \subset \mathbb{R}$ ein Intervall. Eine Funktion $x : I \to \mathbb{R}$ heißt eine Lösung von (1.16) wenn $x(t)$ in I differenzierbar ist,

$$(t, x(t)) \in \Omega \ \text{ und } \ x'(t) = f(t, x(t)) \ \text{ für alle } t \in I. \tag{1.17}$$

Die Bedingung $(t,x(t)) \in \Omega$ ist notwendig für die Wohldefiniertheit von $f(t,x(t))$. Sie bedeutet auch, dass der Graph der Funktion x in Ω liegt.



Eine Lösung $x(t)$

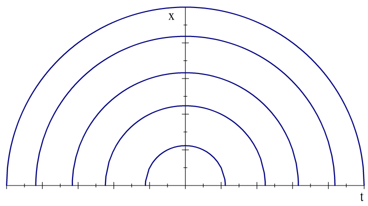
Beispiel. Betrachten wir die Differentialgleichung

$$x' = -\frac{t}{x} \tag{1.18}$$

im Halbebene $\Omega = \{(t, x) \in \mathbb{R}^2 : x > 0\}$. Diese Gleichung lässt sich wie folgt lösen:

$$x' = -\frac{t}{x} \Leftrightarrow xx' = -t \Leftrightarrow \left(\frac{x^2}{2}\right)' = -t \Leftrightarrow \frac{x^2}{2} = -\frac{t^2}{2} + C \Leftrightarrow x(t) = \sqrt{2C - t^2},$$

wobei C eine beliebige positive Konstante ist. Der Definitionsbereich dieser Lösung ist $t \in [-\sqrt{2C}, \sqrt{2C}]$.



Die Lösungen von (1.18)

Die Konstante C lässt sich eindeutig aus der *Anfangsbedingung*

$$x(0) = x_0$$

wie folgt bestimmen: $x(0) = \sqrt{2C} = x_0$ und somit $C = \frac{x_0^2}{2}$.

Betrachten wir das *Anfangswertproblem* für die Gleichung (1.16):

$$\begin{cases} x' = f(t, x), \\ x(t_0) = x_0, \end{cases} \tag{1.19}$$

wobei (t_0, x_0) ein beliebiger Punkt aus Ω ist.

Definition. Eine Funktion $x : I \to \mathbb{R}$ heißt eine Lösung des Anfangswertproblems (1.19) wenn x eine Lösung der Gleichung $x' = f(t, x)$ ist und wenn $t_0 \in I$ und $x(t_0) = x_0$.

Hauptsatz 1.8 (Satz von Picard-Lindelöf) *Sei die Funktion $f : \Omega \to \mathbb{R}$ stetig in (t, x) und stetig differenzierbar in x (d.h. die partielle Ableitung $\partial_x f$ existiert und ist stetig in Ω). Dann für jedes $(t_0, x_0) \in \Omega$ hat das Anfangswertproblem (1.19) eine Lösung $x(t)$.*

Die Eindeutigkeitsaussage gilt auch: gibt es zwei Lösungen x_1 und x_2 von (1.19) die auf den Intervallen I_1 bzw I_2 definiert sind, so gilt $x_1(t) = x_2(t)$ für alle $t \in I_1 \cap I_2$.

Wir fangen den Beweis des Satzes 1.8 mit der folgenden Behauptung an.

Lemma 1.9 *Sei $x : I \to \mathbb{R}$ eine stetige Funktion auf einem Intervall I, so dass $(t, x(t)) \in \Omega$ für alle $t \in I$. Seien $t_0 \in I$ and $(t_0, x_0) \in \Omega$. Die Funktion $x(t)$ löst das Anfangswertproblem (1.19) auf I genau dann wenn sie die folgende Integralgleichung erfüllt:*

$$x(t) = x_0 + \int_{t_0}^{t} f(s, x(s))\, ds \quad \textit{für alle } t \in I. \tag{1.20}$$

Beweis. Bemerken wir zuerst, dass die rechte Seite von (1.20) wohldefiniert ist. Da s zwischen t und t_0 liegt und $t, t_0 \in I$, so liegt s in I und somit ist $f(s, x(s))$ wohldefiniert. Da x und f stetig sind, so ist die Funktion $f(s, x(s))$ stetig in $s \in I$ und somit ist das Integral in (1.20) wohldefiniert.

Sei $x(t)$ eine Lösung von (1.20). Da $x(t)$ stetig ist, so ist die rechte Seite von (1.20) in t differenzierbar. Folglich ist $x(t)$ auch differenzierbar. Ableiten von (1.20) ergibt $x' = f(t, x)$, und die Anfangsbedingung $x(t_0) = x_0$ folgt offensichtlich aus (1.20). Somit löst $x(t)$ das Anfangswertproblem (1.19).

Sei $x(t)$ eine Lösung von (1.19). Es folgt aus der Gleichung $x' = f(t, x(t))$, dass $x'(t)$ stetig ist. Integrieren dieser Gleichung ergibt

$$x(t) - x_0 = \int_{t_0}^{t} x'(s)\, ds = \int_{t_0}^{t} f(s, x(s))\, ds,$$

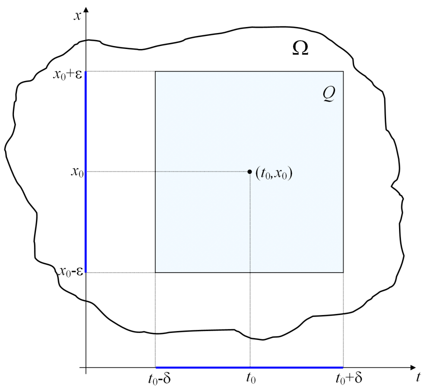
woraus (1.20) folgt. ■

Beweis von dem Satz 1.8. Wir bilden einen vollständigen metrischen Raum (X, d) und eine Selbstabbildung T von X, so dass die Gleichung $Tx = x$ äquivalent zum Anfangswertproblem (1.19) ist. Ist T eine Kontraktionsabbildung so erhalten wir die Existenz der Lösung nach dem Fixpunktsatz von Banach.

Wählen wir positive Konstante ε und δ so klein, dass das Rechteck

$$Q = [t_0 - \delta, t_0 + \delta] \times [x_0 - \varepsilon, x_0 + \varepsilon]$$

eine Teilmenge von Ω ist (was möglich ist da Ω offen ist).



Das Rechteck $Q \subset \Omega$

Da f stetig ist und Q – kompakt, so gilt nach dem Extremwertsatz

$$L := \sup_Q |\partial_x f| < \infty.$$

Es folgt, dass für alle $t \in [t_0 - \delta, t_0 + \delta]$ und $x_1, x_2 \in [x_0 - \varepsilon, x_0 + \varepsilon]$

$$|f(t, x_1) - f(t, x_2)| = \left| \int_{x_2}^{x_1} \partial_x f(t, x)\, dx \right| \leq L |x_1 - x_2|, \tag{1.21}$$

was später benutzt wird.

Fixieren wir ein

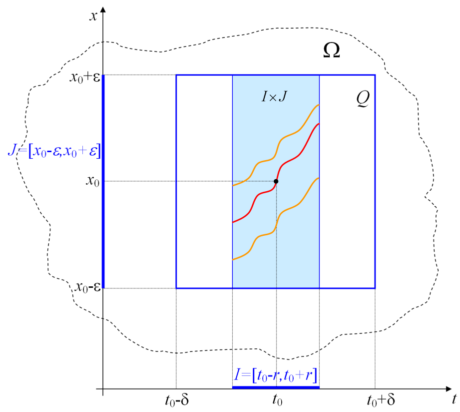
$$r \in (0, \delta], \tag{1.22}$$

das später bestimmt wird, und bezeichnen

$$I = [t_0 - r, t_0 + r] \quad \text{und } J = [x_0 - \varepsilon, x_0 + \varepsilon] \,.$$

Sei X die Menge von allen stetigen Funktionen $x : I \to J$, also

$$X = \{x : I \to J \;\; | \;\; x \text{ ist stetig}\} \,.$$



Funktionen $x \in X$

20.04.22 Vorlesung 4

Definieren wir einen *Integraloperator* T auf Funktionen $x \in X$ wie folgt

$$Tx\,(t) = x_0 + \int_{t_0}^{t} f\,(s, x\,(s))\,ds, \quad t \in I. \tag{1.23}$$

Für jede Funktion $x \in X$ und für alle $s \in I$ liegt der Punkt $(s, x\,(s))$ in $Q \subset \Omega$, so dass das Integral in (1.23) für alle $t \in I$ wohldefiniert ist. Somit ist die Funktion Tx auf dem ganzen Intervall I definiert.

Jetzt müssen wir sicherstellen, dass T eine Selbstabbildung von X ist, d.h.,

$$x \in X \Rightarrow Tx \in X.$$

Die Funktion Tx ist offensichtlich stetig. Es bleibt nur zu zeigen, dass die Werte von $Tx\,(t)$ in J liegen, d.h.

$$|Tx\,(t) - x_0| \le \varepsilon \text{ für alle } t \in I. \tag{1.24}$$

Für jedes $t \in I$ erhalten wir nach (1.23) und LM-Ungleichung für Integrale:

$$\begin{aligned} Tx\,(t) - x_0 \;\; &= \left| \int_{t_0}^{t} f\,(s, x\,(s))\,ds \right| \\ &\le \left| \int_{t_0}^{t} |f\,(s, x\,(s))|\,ds \right| \\ &\le \sup_{(s,x)\in Q} |f\,(s,x)|\,|t - t_0| \\ &\le Mr, \end{aligned}$$

wobei
$$M := \sup_{(s,x)\in Q} |f(s,x)| < \infty.$$
Wir betonen, dass M unabhängig von r definiert ist. Jetzt setzen wir voraus, dass r zusätzlich zu (1.22) noch eine Bedingung erfüllt:
$$r \leq \frac{\varepsilon}{M}. \tag{1.25}$$
Dann ist (1.24) offensichtlich erfüllt und somit erhalten wir, dass $Tx \in X$.

Nach Lemma 1.9, um eine Lösung von (1.19) zu finden, reicht es die Integralgleichung (1.20) zu lösen, was äquivalent zur Gleichung $Tx = x$ ist. D.h., es reicht zu beweisen, dass T einen Fixpunkt besitzt, was mit Hilfe von Fixpunktsatz von Banach bewiesen wird.

Die Menge X ist offensichtlich Teilmenge von $C(I)$ und somit ist ein metrischer Raum mit der Metrik
$$d(x,y) = \|x-y\|_{C(I)} = \sup_{t\in I} |x(t) - y(t)|.$$
Der Raum $C(I)$ ist vollständig, und die Teilmenge X ist offensichtlich abgeschlossen, woraus folgt, dass der metrische Raum (X,d) auch vollständig ist (siehe Aufgabe 15).

Es bleibt zu beweisen, dass $T : X \to X$ eine Kontraktionsabbildung ist, d.h. für ein $q \in (0,1)$ und für alle $x, y \in X$ gilt
$$\|Tx - Ty\| \leq q\,\|x-y\|.$$
Für beliebige Funktionen $x, y \in X$ und für jedes $t \in I$, gilt $x(t), y(t) \in J$, woraus folgt mit Hilfe von (1.21), dass
$$\begin{aligned}
|Tx(t) - Ty(t)| &= \left|\int_{t_0}^{t} f(s,x(s))\,ds - \int_{t_0}^{t} f(s,y(s))\,ds\right| \\
&\leq \left|\int_{t_0}^{t} |f(s,x(s)) - f(s,y(s))|\,ds\right| \\
&\leq \left|\int_{t_0}^{t} L\,|x(s) - y(s)|\,ds\right| \\
&\leq L\,|t-t_0| \sup_{s\in I} |x(s) - y(s)| \\
&\leq Lr\,\|x-y\|
\end{aligned}$$
und somit
$$\|Tx - Ty\| = \sup_{t\in I} |Tx(t) - Ty(t)| \leq Lr\,\|x-y\|.$$
Jetzt wählen wir r so klein, dass r zusätzlich zu (1.22) und (1.25) auch die folgende Bedingung erfüllt:
$$r < \frac{1}{L}.$$
Dann ist T eine Kontraktionsabbildung, und nach dem Fixpunktsatz von Banach hat die Gleichung $Tx = x$ eine Lösung $x \in X$, was zu beweisen war. ∎

1.6 Maß und Lebesgue-Integral

Dieser Abschnitt ist eine Zusammenfassung von Maß- und Integrationstheorie, meistens ohne Beweise.

1.6.1 Motivation

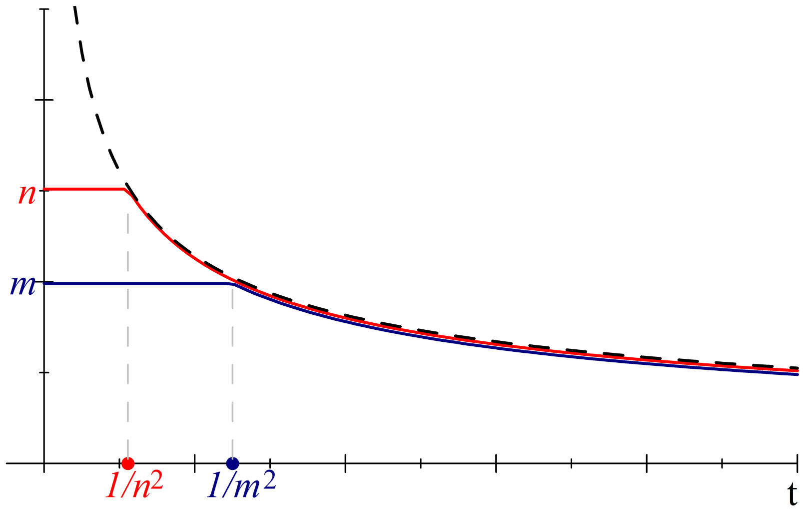
Betrachten wir im Raum $C[a,b]$ die 1-Norm

$$\|f\|_1 = \int_a^b |f(t)|\,dt$$

(siehe Aufgabe 9). Das ist eine Norm in $C[a,b]$ ist, aber der Raum $C[a,b]$ mit dieser Norm ist *nicht* vollständig, was im nächsten Beispiel gezeigt wird.

Beispiel. Betrachten wir die Folge $\{f_n\}_{n=1}^{\infty}$ von Funktionen aus $C[0,1]$:

$$f_n(t) = \begin{cases} n, & t \leq 1/n^2 \\ \frac{1}{\sqrt{t}}, & t \geq 1/n^2 \end{cases} = \min\left(\frac{1}{\sqrt{t}}, n\right)$$



Funktionen f_n (rot) und f_m (blau) und f (schwarz)

Zeigen wir, dass $\{f_n\}$ eine Cauchy-Folge bezüglich der 1-Norm ist aber nicht konvergent. Zuerst bemerken die punktweise Konvergenz

$$f_n(t) \to f(t) := \frac{1}{\sqrt{t}} \text{ für alle } t \in (0,1]$$

und dass

$$\int_0^1 |f(t) - f_n(t)|\,dt = \int_0^{1/n^2} \left(\frac{1}{\sqrt{t}} - n\right) = \left[2\sqrt{t}\right]_0^{1/n^2} - \frac{1}{n} = \frac{2}{n} - \frac{1}{n} = \frac{1}{n} \to 0$$

für $n \to \infty$. Es folgt, dass für $n, m \to \infty$

$$\|f_n - f_m\|_1 = \int_0^1 |f_n - f_m|\,dt \leq \int_0^1 |f - f_n|\,dt + \int_0^1 |f - f_m|\,dt \to 0$$

so that $\{f_n\}$ eine Cauchy-Folge ist. Zeigen jetzt, dass $\{f_n\}$ nicht konvergent in $C[0,1]$ bezüglich der 1-Norm ist. Gibt es einen Grenzwert $g \in C[0,1]$ so gilt

$$\int_0^1 |g - f_n|\, dt = \|g - f_n\|_1 \to 0 \text{ für } n \to \infty.$$

Da auch

$$\int_0^1 |f - f_n|\, dt \to 0,$$

so erhalten wir, dass

$$\int_0^1 |g - f|\, dt \le \int_0^1 |g - f_n|\, dt + \int_0^1 |f - f_n|\, dt \to 0$$

d.h. $\int_0^1 |g - f|\, dt = 0$. Da die beiden Funktionen f und g stetig auf $(0,1]$ sind, so folgt es, dass $f = g$ auf $(0,1]$. Somit ist g unstetig in $t = 0$ und kann nicht in $C[0,1]$ liegen.

Wir sehen, dass die Folge $\{f_n\}$ konvergiert in $C[0,1]$ bezüglich der 1-Norm nicht, aber $\{f_n\}$ konvergiert in einem größeren Raum, wo auch die Funktion $\frac{1}{\sqrt{t}}$ liegt.

Um die Vollständigkeit bezüglich $\|\cdot\|_1$ zu gewinnen, soll man den Raum $C[0,1]$ vervollständigen. Der Raum von allen Riemann-integrierbaren Funktionen reicht nicht (auch wenn uneigentliches Riemann-Integral benutzt wird). Wir brauchen dafür den Begriff von *Lebesgue-integrierbaren* Funktionen, was unterhalb besprochen wird. Im nächsten Abschnitt definieren wir die Lebesgue-Räume L^p mit Integral-p-*Normen.* Insbesondere ist $L^1[a,b]$ eine Vervollständigung von $C[a,b]$ bezüglich der 1-Norm.

1.6.2 Das Maß

Der abstrakte Begriff von Maß ist eine Verallgemeinerung von den Begriffen von Länge, Flächeninhalt, Volumen und sogar Wahrscheinlichkeit.

Definition. Sei X eine Menge. Eine σ-Algebra auf X ist ein Mengensystem S von Teilmengen von X (d.h. $S \subset \mathcal{P}(X)$) mit den folgenden Eigenschaften:

$(A1)$ $\emptyset$ und X sind Elemente von S

$(A2)$ $A \in S \Rightarrow A^c \in S$

$(A3)$ Ist $\{A_k\}_{k\in\mathbb{N}}$ eine Folge von Elementen von S, so gilt $\bigcup_{k\in\mathbb{N}} A_k \in S$.

Es folgt aus $(A1)$ und $(A3)$, dass

$$A, B \in S \Rightarrow A \cup B \in S.$$

Mit Hilfe von $(A2)$ erhalten wir, dass für $A, B \in S$

$$A \cap B = (A^c \cup B^c)^c \in S.$$

Auch für beliebige $A, B \in S$ gilt

$$A \setminus B = A \cap B^c \in S.$$

Man beweist auch, dass für jede Folge $\{A_k\}_{k\in\mathbb{N}}$ aus S gilt $\bigcap_{k=1}^{\infty} A_k \in S$. Somit ist die σ-Algebra S abgeschlossen bezüglich Mengenoperationen $\setminus, \cup, \cap$ wobei die Vereinigung und Durchschnitt sich auf unendlichen (abzählbaren) Folgen durchführen lassen.

Die Elemente der σ-Algebra S heißen auch *messbare Mengen* (oder S-messbare Mengen).

Bemerkung. Der Buchstabe "σ" in "σ-Algebra" bezieht sich auf die abzählbare Vereinigung in $(A3)$. Gilt $(A3)$ nur für die endlichen Vereinigungen, so heißt S eine Algebra (ohne "σ").

Definition. Ein Maß μ auf S ist eine Funktion $\mu : S \to \mathbb{R}$ mit den folgenden Eigenschaften:

$(M1)$ (*Positivität*) $\mu(A) \geq 0$ für alle $A \in S$.

$(M2)$ (σ-*Additivität*) für jede Folge $\{A_k\}_{k\in\mathbb{N}}$ von disjunkten messbaren Mengen gilt

$$\mu\left(\bigsqcup_{k\in\mathbb{N}} A_k\right) = \sum_{k\in\mathbb{N}} \mu(A_k). \tag{1.26}$$

Man nennt den Wert $\mu(A)$ das Maß von A.

Es folgt aus (1.26) mit $A_k = \emptyset$ dass $\mu(\emptyset) = 0$. Es folgt auch, dass für beliebige disjunkte Mengen $A, B \in S$ gilt

$$\mu(A \sqcup B) = \mu(A) + \mu(B), \tag{1.27}$$

da $A \sqcup B = \bigsqcup_{k\in\mathbb{N}} A_k$ mit $A_1 = A$, $A_2 = B$ und $A_k = \emptyset$ für alle $k > 2$.

Bemerkung. Der Buchstabe "σ" in "σ-Additivität" bezieht sich auf die abzählbare Vereinigung/Summe in (1.26). Die ähnliche Eigenschaft (1.26) heißt die *Additivität* (ohne "σ").

Definition. Ein *Maßraum* ist das Dreifache (X, S, μ), wobei S eine σ-Algebra auf X ist und μ ein Maß auf S.

Eine besondere Rolle in der Maßtheorie spielen die Nullmengen.

Definition. Jede messbare Menge mit Maß 0 heißt *Nullmenge*.

Es folgt aus (1.26), dass $\mu(\emptyset) = 0$ so dass $\emptyset$ eine Nullmenge ist. Aber es kann auch nichtleere Nullmengen geben. Es folgt aus (1.26), dass abzählbare Vereinigung von Nullmengen wieder eine Nullmenge ist.

Definition. Ein Maß (μ, S) heißt *vollständig*, wenn jede Teilmenge einer Nullmenge auch Nullmenge ist.

Jedes Maß (μ, S) lässt sich vervollständigen wie folgt. Definieren wir die Erweiterung S' der σ-Algebra S mit

$$S' = \{A \cup B : A \in S \text{ und } B \text{ ist Teilmenge einer Nullmenge}\}.$$

Dann ist S' auch eine σ-Algebra, und μ lässt sich auf S' wie folgt erweitern:

$$\mu(A \cup B) = \mu(A).$$

Man kann beweisen, dass (μ, S') ein vollständiges Maß ist, was die *Vervollständigung* von (μ, S) heißt.

1.6.3 Lebesgue-Maß in $\mathbb{R}^n$

Sei X ein beschränktes Intervall in $\mathbb{R}$. Für jedes Intervall $I \subset X$ definieren wir die Länge $\ell(I)$ mit

$$\ell(I) = b - a$$

wobei a, b die Endpunkten von I sind. Wir möchten die Länge ℓ wie ein Maß betrachten.

Die Menge von allen Intervallen auf X ist keine σ-Algebra, da die Vereinigung zweier Intervallen kein Interval sein kann.

Sei $\mathcal{B}(X)$ die minimale σ-Algebra auf X die alle Intervalle in X enthält. Solche σ-Algebra existiert als Durchschnitt von allen σ-Algebren die alle Intervalle enthalten, und ist offensichtlich eindeutig bestimmt. Die σ-Algebra $\mathcal{B}(X)$ heißt die *Borel-σ-Algebra* and die Elemente von $\mathcal{B}(X)$ heißen Borel(-messbare) Mengen in X.

Satz 1.10 *Die Länge ℓ lässt sich auf $\mathcal{B}(X)$ als ein Maß λ eindeutig erweitern.*

Das Maß λ heißt das *Lebesgue-Maß* in X. Betrachten wir die Vervollständigung von $(\lambda, \mathcal{B}(X))$. Der Definitionsbereich der Vervollständigung von $(\lambda, \mathcal{B}(X))$ heißt die *Lebesgue-σ-Algebra* und wird mit $\mathcal{L}(X)$ bezeichnet. Elemente von $\mathcal{L}(X)$ heißen *Lebesgue-messbare* Mengen.

Man kann zeigen, dass λ auf der ganze Potenzmenge $\mathcal{P}(X)$ nicht als Maß fortsetzbar. Das Lebesgue-Maß lässt sich auch auf $X = \mathbb{R}$ definieren, aber mit Werten in $[0, \infty]$.

Es ist klar aus der Konstruktion, dass jede Menge $\{x\}$ die aus einem Punkt $x \in X$ besteht eine Nullmenge ist. Es folgt, dass die Menge von allen rationalen Zahlen in X auch Nullmenge ist als abzählbare Vereinigung von Nullmengen.

Ähnliche Konstruktion gilt in $\mathbb{R}^n$. Fixieren wir eine beschränkte offene Teilmenge X von $\mathbb{R}^n$. Man definiert erst das Volumen ℓ_n von n-dimensionalen Quadern

$$Q = I_1 \times ... \times I_n \subset X$$

mit

$$\ell_n(Q) = \ell(I_1) ... \ell(I_n).$$

Sei $\mathcal{B}(X)$ die Borel-σ-Algebra in X, d.h. $\mathcal{B}(X)$ ist die minimale σ-Algebra auf X, die alle Quadern enthält. Man kann zeigen, dass $\mathcal{B}(X)$ die minimale σ-Algebra ist die all offene Teilmengen von X enthält.

Satz 1.11 *Das Volumen ℓ_n lässt sich auf $\mathcal{B}(X)$ als ein Maß λ_n eindeutig erweitern.*

Das Maß λ_n heißt das n-dimensionale Lebesgue-Maß auf X. Die Vervollständigung von $(\lambda_n, \mathcal{B}(X))$ wird mit $(\lambda_n, \mathcal{L}(X))$ bezeichnet. Die Elemente der α-Algebra $\mathcal{L}(X)$ heißen Lebesgue-messbare Mengen in X.

1.6.4 Messbare Funktionen und Lebesgue-Integration

Sei (X, S, μ) ein Maßraum.

Definition. Eine Funktion $f : X \to \mathbb{R}$ heißt *messbar* (oder S-messbar) wenn die Menge

$$\{f < c\} = \{x \in X : f(x) < c\}$$

messbar für jedes $c \in \mathbb{R}$ ist, d.h. in S liegt.

Man kann beweisen, dass Summe und Produkt von messbaren Funktionen wieder messbar sind.

Beispiel. Die Indikatorfunktion 1_A einer messbaren Menge A ist messbare Funktion. Auch lineare Kombination

$$\sum_{k=1}^{N} c_k \mathbf{1}_{A_k}$$

von Indikatorfunktionen messbar ist. Zum Beispiel, die Dirichlet-Funktion

$$f(t) = \mathbf{1}_{\mathbb{Q}} = \begin{cases} 1, & t \in \mathbb{Q} \\ 0, & t \in \mathbb{R} \setminus \mathbb{Q} \end{cases}$$

ist messbar bezüglich Lebesgue-Maßes, da $\mathbb{Q}$ messbar als eine abzählbare Menge ist.

Beispiel. Jede stetige Funktion f auf einem Intervall $X \subset \mathbb{R}$ ist messbar bezüglich Lebesgue-Maßes. In der Tat, die Menge $\{f < c\}$ ist offen und somit Borel-messbar. Auch jede stetige Funktion auf einer offenen Teilmenge $X \subset \mathbb{R}^n$ ist Borel-messbar.

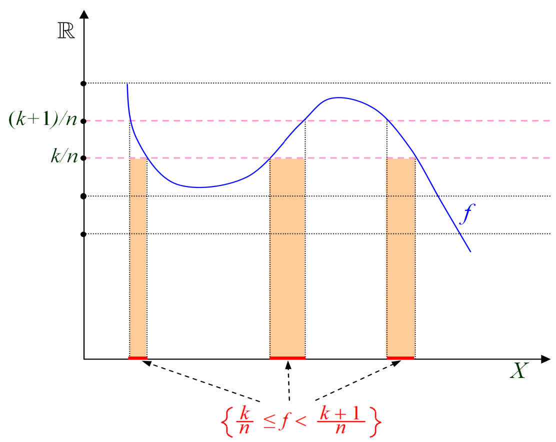
Definition. Für jede nicht-negative messbare Funktion f auf einem Maßraum (X, S, μ) definiert man das *Lebesgue-Integral* $\int_X f d\mu$ als einen Wert aus $[0, \infty]$, wie folgt:

$$\int_X f d\mu = \lim_{n\to\infty} \sum_{k=0}^{\infty} \frac{k}{n} \mu \left\{ x \in X : \frac{k}{n} \le f(x) < \frac{k+1}{n} \right\}. \tag{1.28}$$

Bemerken wir, dass die Menge

$$\left\{ \frac{k}{n} \le f < \frac{k+1}{n} \right\} = \left\{ f < \frac{k+1}{n} \right\} \setminus \left\{ f < \frac{k}{n} \right\}$$

messbar ist. Deshalb ist die Summe in (1.28) wohldefiniert und heißt *Lebesgue-Summe*.



Man kann zeigen, dass der Grenzwert (1.28) immer in $[0, \infty]$ existiert. Somit ist das Integral $\int_X f d\mu$ für jede nichtnegative messbare Funktion f immer definiert und nimmt den Wert in $[0, \infty]$.

Beispiel. Für jede messbare Menge $A \subset X$ gilt

$$\int_X \mathbf{1}_A d\mu = \mu(A).$$

Zum Beispiel, für die Dirichlet-Funktion $f = \mathbf{1}_{\mathbb{Q}\cap[0,1]}$ auf dem Intervall $[0,1]$ erhalten wir

$$\int_{[0,1]} f \, d\lambda = \lambda(\mathbb{Q}\cap[0,1]) = 0.$$

Bemerken wir, dass die Dirichlet-Funktion f nicht Riemann-integrierbar, aber das Lebesgue-Integral von f existiert.

Definition. Eine messbare Funktion f heißt *(Lebesgue-)integrierbar* bezüglich μ wenn

$$\int_X |f| \, d\mu < \infty. \tag{1.29}$$

In diesem Fall definiert man das Lebesgue-Integral von f mit

$$\int_X f d\mu = \int_X f_+ d\mu - \int_X f_- d\mu \tag{1.30}$$

wobei

$$f_+ = \max(f,0) \quad \text{und} \quad f_- = \max(-f,0).$$

Es ist leicht zu sehen, dass

$$f_+ - f_- = f \quad \text{und} \quad f_+ + f_- = |f|$$

woraus folgt $f_+ = \frac{f+|f|}{2}$ und $f_- = \frac{|f|-f}{2}$. Man kann beweisen, dass die Funktion $|f|$ messbar ist, woraus folgt, dass auch f_+ und f_- messbar sind. Diese Funktionen sind nichtnegativ, und es folgt aus (1.29), dass auch

$$\int_X f_+ d\mu < \infty \quad \text{und} \quad \int_X f_- d\mu < \infty.$$

Somit ist die rechte Seite in (1.30) eine Differenz von reellen Zahlen, so dass das Integral $\int_X f d\mu$ als eine reelle Zahl definiert ist.

Beispiel. Zeigen wir, dass jede stetige Funktion f auf einem beschränkten und abgeschlossenen Intervall $[a,b]$ Lebesgue-integrierbar ist (d.h., integrierbar bezüglich des Lebesgue-Maßes λ). Wir wissen schon, dass f messbar ist. Die Stetigkeit impliziert, dass

$$M := \sup_{[a,b]} |f| < \infty,$$

woraus folgt, dass für die Lebesgue-Summe für $\int_{[a,b]} |f| \, d\lambda$ die folgende Ungleichung gilt:

$$\begin{aligned}
\sum_{k=0}^{\infty} \frac{k}{n} \mu\left\{\frac{k}{n} \le |f| < \frac{k+1}{n}\right\} &= \sum_{\left\{k:\frac{k}{n}\le M\right\}} \frac{k}{n}\left\{\frac{k}{n} \le |f| < \frac{k+1}{n}\right\} \\
&\le M \sum_{k=0}^{\infty} \mu\left\{\frac{k}{n} \le |f| < \frac{k+1}{n}\right\} \\
&\le M(b-a),
\end{aligned}$$

wobei wir benutzt haben, dass die Mengen $\left\{\frac{k}{n} \le |f| < \frac{k+1}{n}\right\}$ disjunkt sind und ihre Vereinigung eine Teilmenge von $[a,b]$ ist. Es folgt, dass

$$\int_{[a,b]} |f| \, d\lambda < \infty,$$

was zu beweisen war. Man kann auch beweisen, dass die Lebesgue-Integral und Riemann-Integral in diesem Fall übereinstimmen:

$$\int_{[a,b]} f d\lambda = \int_a^b f(t) \, dt. \tag{1.31}$$

Man kann auch beweisen, dass jede Riemann-integrierbare Funktion f auch Lebesgue-integrierbar ist und die Identität (1.31) gilt. Die Umkehrung gilt nicht. Zum Beispiel, die Dirichlet-Funktion $\mathbf{1}_{\mathbb{Q}}$ ist Lebesgue-integrierbar aber nicht Riemann-integrierbar.

22.04.22 Vorlesung 5

Das Lebesgue-Integral hat die folgenden Eigenschaften, wobei f und g entweder nichtnegative messbare Funktionen oder integrierbare Funktionen sind:

1. $\int_X (f+g) \, d\mu = \int_X f d\mu + \int_X g d\mu$
2. $\int_X (cf) \, d\mu = c \int_X f d\mu$

   Zum Beispiel, für die Funktion $f = \sum_{k=1}^N c_k \mathbf{1}_{A_k}$ mit messbaren Mengen A_k gilt

$$\int_X f d\mu = \sum_{k=1}^N c_k \mu(A_k).$$

3. $f \le g \Rightarrow \int_X f d\mu \le \int_X g d\mu$.

   Daraus folgt, dass

$$\left|\int_X f d\mu\right| \le \int_X |f| \, d\mu.$$

4. Für eine messbare Funktion $f \ge 0$ gilt die Äquivalenz

$$\int_X f d\mu = 0 \ \Leftrightarrow \ \mu\{f > 0\} = 0.$$

Gilt eine Eigenschaft außerhalb einer Nullmenge, so sagt man, dass diese Eigenschaft *fast überall* gilt, kurz f.ü. (oder für fast alle Elemente von X). Z.B. für nichtnegative Funktion f

$$\{f = 0\} = X \setminus \{f > 0\}$$

und die Bedingung $\mu\{f > 0\} = 0$ bedeutet, dass $f = 0$ gilt fast überall. Somit erhalten wir die Äquivalenz

$$\int_X f d\mu = 0 \quad \Leftrightarrow \quad f = 0 \text{ f.ü.} \tag{1.32}$$

Jede messbare Teilmenge Y von X definiert man

$$\int_Y f d\mu := \int_X f \mathbf{1}_Y d\mu.$$

Betrachten wir in der Menge Y die σ-Algebra $S_Y = \{A \in S : A \subset Y\}$ und Maß $\mu|_{S_Y}$ so dass $(Y, S_Y, \mu|_{S_Y})$ ein Maßraum ist. Man kann zeigen, dass

$$\int_Y f d\mu|_{S_Y} = \int_Y f d\mu.$$

Für zwei disjunkte messbare Teilmenge Y_1 und Y_2 von X gilt

$$\int_{Y_1 \sqcup Y_2} f d\mu = \int_{Y_1} f d\mu + \int_{Y_2} f d\mu.$$

1.6.5 Konvergenz fast überall und Fatou-Lemma

Sei (X, S, μ) ein Maßraum.

Definition. Man sagt, dass eine Folge $\{f_k\}_{k=1}^{\infty}$ von Funktionen auf X gegen eine Funktion f auf X *fast überall* konvergiert und schreibt

$$f_k \overset{\text{f.ü.}}{\to} f \quad \text{oder} \quad f_k \to f \text{ f.ü.}$$

wenn die Bedingung $f_k(x) \to f(x)$ für fast alle $x \in X$ gilt, d.h.

$$\{x \in X : f_k(x) \to f(x) \quad \text{für } k \to \infty\} = X \setminus N,$$

wobei N eine Nullmenge ist.

Wir nehmen ohne Beweis die folgenden Aussagen an.

Lemma 1.12 *Sei $\{f_k\}$ eine Folge von messbaren Funktionen auf X. Gilt*

$$f_k \overset{\text{f.ü.}}{\to} f$$

so ist f aus messbar.

Lemma 1.13 (Fatou-Lemma) *Sei $\{f_k\}$ eine Folge von nichtnegativen messbaren Funktionen auf X mit*

$$f_k \overset{\text{f.ü.}}{\to} f.$$

Gilt für alle k und eine Konstante C

$$\int_X f_k \, d\mu \le C,$$

so gilt auch

$$\int_X f \, d\mu \le C.$$

Eine äquivalente Version von Fatou-Lemma besagt folgendes: für jede Folge von nichtnegativen messbaren Funktionen $\{f_k\}$ auf X gilt

$$\int_X \left(\liminf_{k\to\infty} f_k\right) d\mu \le \liminf_{k\to\infty} \int_X f_k d\mu.$$

1.7 Lebesgue-Räume

1.7.1 Die p-Norm

Sei (X, S, μ) ein Maßraum. Für jedes $p \in [1, +\infty)$ und für jede messbare Funktion $f : X \to \mathbb{R}$ definieren wir die p-Norm von f mit

$$\|f\|_p := \left(\int_X |f|^p \, d\mu\right)^{1/p}.$$

Da $|f|^p$ eine nichtnegative messbare Funktion ist, so ist $\|f\|_p$ immer wohldefiniert als Element von $[0, \infty]$. Zum Beispiel, wir haben

$$\|f\|_1 = \int_X |f| \, d\mu \text{ and } \|f\|_2 = \left(\int_X f^2 \, d\mu\right)^{1/2}.$$

Trotz des Namens ist die p-Norm *nicht unbedingt* eine Norm. Z.B. $\|f\|_p$ can gleich ∞ sein oder $\|f\|_p$ kann gleich 0 für $f \not\equiv 0$ sein (z.B. für Dirichlet-Funktion f).

Wir besprechen später eine Möglichkeit die p-Norm zu einer Norm zu machen. Jetzt betrachten wir die Eigenschaften der p-Norm.

Bemerken wir, dass die absolut Homogenität für die p-Norm offensichtlich erfüllt ist: für jedes $\alpha \in \mathbb{R}$ haben wir

$$\|\alpha f\|_p = \left(\int_X |\alpha|^p |f|^p \, d\mu\right)^{1/p} = |\alpha| \left(\int_X |f|^p \, d\mu\right)^{1/p} = |\alpha| \, \|f\|_p \, .$$

Jetzt beweisen wir die Dreiecksungleichung.

Lemma 1.14 (Hölder-Ungleichung) *Seien $p, q > 1$ die konjugierten Hölder-Exponenten, d.h.*

$$\frac{1}{p} + \frac{1}{q} = 1. \tag{1.33}$$

Dann für alle messbare Funktionen f, g auf X gilt

$$\int_X |fg| \, d\mu \le \|f\|_p \, \|g\|_q \tag{1.34}$$

(wobei das Product $0 \cdot \infty$ als 0 definiert ist).

Beweis von Lemma 1.14. Umbenennen wir $|f|$ in f und $|g|$ in g, so dass f und g nicht-negativ sind. Gilt $\|f\|_p = 0$, so haben wir $f = 0$ f.ü.und somit

$$\int_X |fg| \, d\mu = 0$$

so dass (1.34) gilt. Nehmen wir an, dass $\|f\|_p > 0$ und $\|g\|_q > 0$. Gilt $\|f\|_p = \infty$, so ist (1.34) wieder offensichtlich. Wir können jetzt annehmen, dass die beiden Werte $\|f\|_p$ und $\|g\|_q$ Elemente von $(0, +\infty)$ sind.

Wenn f in der Ungleichung (1.34) durch αf ersetzt wird (wobei $\alpha > 0$), so werden die beiden Seiten von (1.34) mit α multipliziert, und die Gültigkeit von (1.34) ändert sich nicht. Deshalb können wir annehmen, dass $\|f\|_p = 1$ und analog auch $\|g\|_q = 1$. Nach der Young-Ungleichung (1.10) erhalten wir

$$f(x)\,g(x) \le \frac{f^p(x)}{p} + \frac{g^q(x)}{q} \quad \text{für jedes } x \in X,$$

woraus folgt

$$\begin{aligned} \int_X fg d\mu &\le \frac{1}{p}\int_X f^p\,d\mu + \frac{1}{q}\int_X g^q\,d\mu \\ &= \frac{1}{p} + \frac{1}{q} = 1 = \|f\|_p\,\|g\|_q\,, \end{aligned}$$

was zu beweisen war. ∎

Satz 1.15 (Minkowski-Ungleichung) *Für alle messbare Funktionen f, g auf X and für alle $p \in [1,\infty)$ gilt*

$$\|f+g\|_p \le \|f\|_p + \|g\|_p\,. \tag{1.35}$$

Beweis. Für $p = 1$ haben wir

$$\|f+g\|_1 = \int_X |f+g|\,d\mu \le \int_X |f|\,d\mu + \int_X |g|\,d\mu = \|f\|_1 + \|g\|_1\,.$$

Sei $p > 1$. Ist $\|f\|_p$ oder $\|g\|_p$ gleich ∞, so gilt (1.35) trivialerweise. Nehmen wir an, dass die beiden Werte $\|f\|_p$ und $\|g\|_p$ in $[0,\infty)$ sind. Da $|f+g| \le |f| + |g|$, so können wir ohne Beschränkung der Allgemeinheit annehmen, dass f und g nichtnegativ sind.

Die Ungleichung

$$(a+b)^p \le 2^p\,(a^p + b^p)$$

für nichtnegative a, b ergibt

$$\int_X (f+g)^p\;d\mu \le 2^p \left(\int_X f^p\,d\mu + \int_X g^p\,d\mu\right) < \infty,$$

woraus folgt, dass $\|f+g\|_p < \infty$. Im Fall $\|f+g\|_p = 0$ gilt (1.35) trivialerweise, so wir können weiter annehmen, dass

$$\|f+g\|_p \in (0,+\infty)\,.$$

Wir haben

$$\begin{aligned} \|f+g\|_p^p &= \int_X (f+g)^p\;d\mu = \int_X (f+g)\,(f+g)^{p-1}\,d\mu \\ &= \int_X f\,(f+g)^{p-1}\;d\mu + \int_X g\,(f+g)^{p-1}\;d\mu. \end{aligned} \tag{1.36}$$

Sei $q = \frac{p}{p-1}$ der zu p konjugierte Hölder-Exponent, d.h. $\frac{1}{p} + \frac{1}{q} = 1$. Nach der Hölder-Ungleichung erhalten wir

$$\begin{aligned} \int_X f\,(f+g)^{p-1}\;d\mu &\le \|f\|_p \left\|(f+g)^{p-1}\right\|_q \\ &= \|f\|_p \left(\int_X (f+g)^{(p-1)q}\;d\mu\right)^{1/q} = \|f\|_p\,\|f+g\|_p^{p/q} \end{aligned}$$

und analog
$$\int_X g\,(f+g)^{p-1}\,d\mu \le \|g\|_p\,\|f+g\|_p^{p/q}.$$
Addieren diese Ungleichungen ergibt nach (1.36)
$$\|f+g\|_p^p \le \left(\|f\|_p + \|g\|_q\right)\|f+g\|_p^{p/q}.$$
Dividieren durch $\|f+g\|_p^{p/q}$ ergibt (1.35) da $p-\frac{p}{q}=p\left(1-\frac{1}{q}\right)=1$. ■

Beispiel. Betrachten wir die Menge $X=\{1,2,...,n\}$ und das *Zählmaß* μ auf der σ-Algebra $S=\mathcal{P}(X)$; d.h.
$$\mu(A) = \text{die Anzahl der Elemente von } A.$$
Insbesondere gilt $\mu(\{i\})=1$ für jedes $i=1,...,n$ und $\mu(X)=n$. Dann ist (X,S,μ) ein Maßraum. Jede Funktion $f:X\to\mathbb{R}$ lässt sich mit der Folge $(f(1),...,f(n))\in\mathbb{R}^n$ identifizieren. Es gilt
$$\int_X f d\mu = \sum_{i=1}^n f(i)\,\mu(\{i\}) = \sum_{i=1}^n f(i).$$
Die p-Norm der Funktion $f:X\to\mathbb{R}$ ist
$$\|f\|_p = \left(\int_X |f|^p\,d\mu\right)^{1/p} = \left(\sum_{i=1}^n |f(i)|^p\right)^{1/p},$$
was mit der p-Norm des Vektors $(f(1),...,f(n))$ in $\mathbb{R}^n$ übereinstimmt.

Die Ungleichung (1.34) bedeutet folgendes:
$$\sum_{i=1}^n |f(i)\,g(i)| \le \|f\|_p\,\|g\|_q,$$
was mit der Hölder-Ungleichung von Lemma 1.3 übereinstimmt.

Die Minkowski-Ungleichung (1.35) ergibt in diesem Fall die Minkowski-Ungleichung des Satzes 1.2 in $\mathbb{R}^n$.

Beispiel. Betrachten wir die Menge $X=\mathbb{N}=\{1,2,...\}$ und das Zählmaß μ auf der σ-Algebra $S=\mathcal{P}(X)$. Für jede Funktion $f:X\to[0,\infty)$ haben wir
$$\int_X f d\mu = \sum_{i=1}^\infty f(i).$$
Die p-Norm einer Funktion $f:X\to\mathbb{R}$ ist
$$\|f\|_p = \left(\int_X |f|^p\,d\mu\right)^{1/p} = \left(\sum_{i=1}^\infty |f(i)|^p\right)^{1/p},$$
was mit der p-Norm von Folgen übereinstimmt. Die Minkowski-Ungleichung (1.35) stimmt in diesem Fall mit der Minkowski-Ungleichung des Satzes 1.2 in l^p.

1.7.2 Definition von L^p

We haben schon gesehen, dass die p-Norm $\|f\|_p$ von messbaren Funktionen zwei Bedingungen von der Definition einer Norm erfüllt: die absolute Homogenität und die Dreiecksungleichung. Noch fehlen zwei andere Bedingungen: die Endlichkeit (die Norm muss nur die Werte in $[0,\infty)$ annehmen) und die Definitheit. Im allgemeinen Fall gelten diese Bedingungen nicht: die p-Norm kann den Wert ∞ annehmen (z.B. für die Funktion $f(t) = \frac{1}{t}$ auf $(0,1)$ gilt $\|f\|_1 = \int_0^1 \frac{dt}{t} = \infty$) und die p-Norm von f kann 0 für eine nicht-Null Funktion f sein (z.B. für die Dirichlet-Funktion).

Unser Zweck ist jetzt einen Vektorraum zu erstellen wo die p-Norm eine Norm ist. Dafür betrachten wir für jedes $p \in [1,\infty)$ die Menge

$$\mathcal{L}^p = \left\{ f : X \to \mathbb{R} : \ f \text{ ist messbar und } \ \|f\|_p < \infty \right\}.$$

Es folgt aus der Minkowski-Ungleichung, dass die Menge $\mathcal{L}^p$ ein Vektorraum über $\mathbb{R}$ ist. Nach Definition ist die p-Norm endlich auf $\mathcal{L}^p$ aber noch nicht unbedingt positiv definit: obwohl $\|f\|_p \geq 0$ für alle $f \in \mathcal{L}^p$ gilt, es kann sein $\|f\|_p = 0$ für nicht-Null Funktion f. Somit ist die p-Norm eine Halbnorm auf $\mathcal{L}^p$.

Um die p-Norm zu einer Norm machen, man muss alle f mit $\|f\|_p = 0$ als ein Nullelement betrachten.

Betrachten wir die folgende Teilmenge von $\mathcal{L}^p$:

$$\mathcal{N} = \left\{ f : X \to \mathbb{R} : \ f \text{ ist messbar und } \ \|f\|_p = 0 \right\}.$$

Lemma 1.16 *Die Menge $\mathcal{N}$ hat die folgenden Eigenschaften.*

(*a*) *Es gilt $f \in \mathcal{N}$ genau dann wenn $f = 0$* f.ü. *Insbesondere ist $\mathcal{N}$ unabhängig von p.*

(*b*) *$\mathcal{N}$ ein Unterraum von $\mathcal{L}^p$.*

Beweis. (*a*) Wir haben

$$\|f\|_p = 0 \Leftrightarrow \int_X |f|^p \, d\mu = 0 \Leftrightarrow |f|^p = 0 \text{ f.ü.} \Leftrightarrow f = 0 \text{ f.ü.}$$

(*b*) Offensichtlich ist $\mathcal{N}$ eine Teilmenge von $\mathcal{L}^p$ da $\|f\|_p = 0$ ergibt $\|f\|_p < \infty$. Für $f, g \in \mathcal{N}$ gilt

$$\|f+g\|_p \leq \|f\|_p + \|g\|_p = 0$$

so dass $\|f+g\|_p = 0$ und somit $f + g \in \mathcal{N}$. Auch für jede Konstante $\alpha \in \mathbb{R}$ gilt $\|\alpha f\|_p = |\alpha| \, \|f\|_p = 0$ und somit $\alpha f \in \mathcal{N}$. Es folgt, dass $\mathcal{N}$ ein Unterraum von $\mathcal{L}^p$ ist. ■

Definition. *(Lebesgue-Raum)* Für jedes $p \in [1,\infty)$ bezeichnen wir mit $L^p = L^p(X,\mu)$ den Faktorraum $\mathcal{L}^p/\mathcal{N}$:

$$L^p = \mathcal{L}^p/\mathcal{N}.$$

Nach der Definition besteht der Faktorraum $\mathcal{L}^p/\mathcal{N}$ aus den Äquivalenzklassen der Äquivalenzrelation $f - g \in \mathcal{N}$, was nach Lemma 1.16 äquivalent zu $f = g$ f.ü. ist. Für jede Funktion $f \in \mathcal{L}^p$ bezeichnen wir mit $[f]$ die Äquivalenzklasse von f, d.h.

$$[f] = \{g \in \mathcal{L}^p : f = g \text{ f.ü.}\} = \{g : X \to \mathbb{R} : f = g \text{ f.ü.}\}.$$

Dann gilt

$$L^p = \{[f] : f \in \mathcal{L}^p\}.$$

Nach der Definition des Faktorraums werden die lineare Operationen in L^p wie folgt definiert:

$$\begin{aligned} [f] + [g] &= [f+g] \\ \alpha\,[f] &= [\alpha f] \end{aligned} \tag{1.37}$$

für alle $f, g \in \mathcal{L}^p$ und $\alpha \in \mathbb{R}$. Definieren wir die p-Norm von Äquivalenzklassen mit

$$\|[f]\|_p = \|f\|_p \,. \tag{1.38}$$

Lemma 1.17 *Die p-Norm (1.38) ist wohldefiniert und ist eine Norm in L^p.*

Beweis. Zeigen wir, dass die Definition (1.38) unabhängig von der Wahl des Vertreters f der Klasse $[f]$ ist. In der Tat, im Fall $g \in [f]$ gilt $f - g \in \mathcal{N}$ und somit

$$\|f\|_p = \|f - g + g\|_p \le \|f - g\|_p + \|g\|_p = \|g\|_p$$

und analog $\|g\|_p \le \|f\|_p$, woraus folgt $\|f\|_p = \|g\|_p$. Somit ist $\|[f]\|_p$ wohldefiniert.

Die p-Norm auf L^p ist endlich und erfüllt die absolute Homogenität und die Dreiecksungleichung. Die Definitheit lässt sich wie folgt beweisen:

$$\|[f]\|_p = 0 \Leftrightarrow \|f\|_p = 0 \Leftrightarrow f \in \mathcal{N} \Leftrightarrow [f] = 0.$$

■

Beispiel. Betrachten wir wieder die Dirichlet-Funktion $f = \mathbf{1}_{\mathbb{Q}\cap[0,1]}$ auf $X = [0,1]$ mit Lebesgue-Maß λ. Es gilt $f = 0$ f.ü. da

$$\lambda\{x \in X : f(x) \neq 0\} = \lambda(\mathbb{Q} \cap [0,1]) = 0.$$

Somit liegt f in $\mathcal{N}$ und $[f]$ ist ein Nullelement des Raums $L^p[0,1]$.

Konvention. Man nennt die Elemente von L^p auch Funktionen (obwohl sie keine Funktionen sind) und schreibt $f \in L^p$ anstatt $[f] \in L^p$ oder $f \in \mathcal{L}^p$. Es ist so gemacht um die Terminology und Notation zu vereinfachen. Allerdings muss man immer daran erinnern, dass die Aussage “f ist eine Funktion aus L^p” folgendes bedeutet: “f ist eine Funktion, die ein Vertreter eines Elements aus L^p ist” oder “f ist eine Funktion aus $\mathcal{L}^p$” . Allerdings wird der Raum $\mathcal{L}^p$ selten explizit benutzt.

1.7.3 Vollständigkeit von L^P

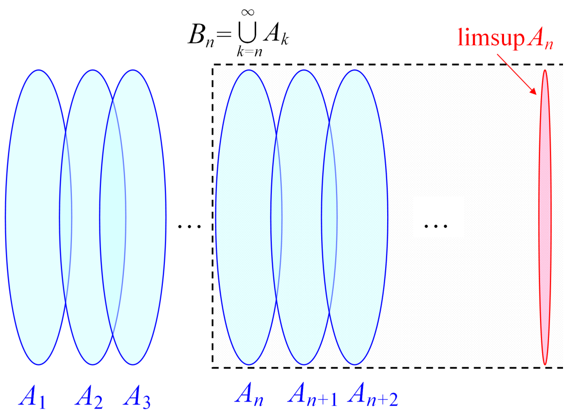
Hauptsatz 1.18 *Für jeden Maßraum (X, S, μ) und für jedes $p \in [1, \infty)$ ist $L^p(X, \mu)$ ein Banachraum.*

Da L^p ein normierter Vektorraum ist, so brauchen wir nur die Vollständigkeit davon beweisen. Wir fangen mit zwei Lemmas an.

Definition. Für eine Folge $\{A_n\}_{n\in\mathbb{N}}$ von Teilmengen von X definieren wir $\limsup A_n$ mit

$$\limsup A_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k.$$

Bemerken wir, dass die Folge $\{\bigcup_{k=n}^{\infty} A_k\}_{n\in\mathbb{N}}$ monoton fallend bezüglich Inklusion ist.



Lemma 1.19 (Borel-Cantelli-Lemma) *Sei $\{A_n\}_{n\in\mathbb{N}}$ eine Folge von messbaren Teilmengen von X mit*

$$\sum_{n\in\mathbb{N}} \mu(A_n) < \infty.$$

Dann gilt $\mu(\limsup A_n) = 0$.

Beweis. Bezeichnen wir $B_n = \bigcup_{k=n}^{\infty} A_k$ und

$$A := \limsup A_n = \bigcap_{n=1}^{\infty} B_n.$$

Offensichtlich sind B_n und A messbar. Nach der σ-Subadditivität von Maß erhalten wir

$$\mu(B_n) \leq \sum_{k=n}^{\infty} \mu(A_k) \to 0 \text{ für } n \to \infty.$$

Da $A \subset B_n$ für alle n, so folgt es $\mu(A) \leq \mu(B_n)$ und somit $\mu(A) = 0$. ∎

Lemma 1.20 (Tschebyschow-Ungleichung) *For jede messbare Funktion f auf X und für alle $t > 0$, $p \in [1,\infty)$ gilt*

$$\mu\{f \geq t\} \leq \frac{1}{t^p} \|f\|_p^p.$$

Beweis. Setzen wir $A = \{f \geq t\}$. Diese Menge ist messbar, da A das Komplement der messbaren Menge $\{f < t\}$ ist. Es gilt

$$\|f\|_p^p = \int_X |f|^p \, d\mu = \int_A |f|^p \, d\mu + \int_{A^c} |f|^p \, d\mu \geq \int_A |f|^p \, d\mu \geq \int_A t^p \, d\mu = t^p \mu(A),$$

woraus die Behauptung folgt. ∎

Beweis von Hauptsatz 1.18. Wir müssen beweisen, dass jede Cauchy-Folge in L^p konvergiert. Eine Folge $\{[f_n]\}$ in L^p ist eine Cauchy-Folge, wenn

$$\|f_n - f_m\|_p = \|[f_n] - [f_m]\|_p \to 0 \quad \text{für } n, m \to \infty,$$

d.h. wenn $\{f_n\}$ eine Cauchy-Folge in $\mathcal{L}^p$ ist. Die Konvergenz der Folge $\{[f_n]\}$ in L^p gegen ein $[f] \in L^p$ bedeutet, dass

$$\|[f_n] - [f]\|_p \to 0 \quad \text{für } n \to \infty$$

was äquivalent zu
$$\|f_n - f\|_p \to 0 \text{ für } n \to \infty,$$
d.h. $\{f_n\}$ konvergiert gegen f bezüglich der Halbnorm $\|\cdot\|_p$. In diesem Fall schreiben wir, dass
$$f_n \xrightarrow{L^p} f \text{ für } n \to \infty.$$
Nach Aufgabe 16 (a), die Cauchy-Folge $\{f_n\}$ konvergiert wenn eine Teilfolge $\{f_{n_k}\}$ dieser Folge konvergiert. Nach Aufgabe 16 (b) gibt es eine Teilfolge $\{f_{n_k}\}$ so dass für alle $k \geq 1$ gilt
$$\left\|f_{n_{k+1}} - f_{n_k}\right\|_p \leq 4^{-k}.$$
Um die Notation zu vereinfachen, umbenennen wir die Folge $\{f_{n_k}\}$ zurück in $\{f_n\}$. Dann $\{f_n\}$ ist eine Cauchy-Folge in $\mathcal{L}^p$ mit
$$\|f_{n+1} - f_n\|_p \leq 4^{-n} \text{ für alle } n \in \mathbb{N}, \tag{1.39}$$
und wir müssen beweisen, dass es eine Funktion $f \in \mathcal{L}^p$ gibt mit
$$\|f_n - f\|_p \to 0 \text{ für } n \to \infty.$$

27.04.22 Vorlesung 6

Beweisen wir zunächst, dass die Folge $\{f_n\}$ gegen eine Funktion f f.ü. konvergiert. Dafür betrachten wir für jedes $n \in \mathbb{N}$ die Menge
$$A_n = \left\{x \in X : |f_{n+1}(x) - f_n(x)| \geq 2^{-n}\right\}.$$
Nach der Tschebyschow-Ungleichung und (1.39) erhalten wir
$$\mu(A_n) \leq 2^{np} \|f_{n+1} - f_n\|_p^p \leq 2^{np} 4^{-np} = 2^{-np}.$$
Es folgt, dass
$$\sum_{n \in \mathbb{N}} \mu(A_n) < \infty.$$
Nach dem Borel-Cantelli-Lemma hat die Menge
$$A = \limsup A_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k$$
das Maß 0. Betrachten wir das Komplement von A:
$$\begin{aligned} A^c &= \left(\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k\right)^c = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k^c \\ &= \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} \left\{x \in X : |f_{k+1}(x) - f_k(x)| < 2^{-k}\right\}, \end{aligned}$$
woraus folgt, dass
$$x \in A^c \quad \Leftrightarrow \quad \exists n \ \ \forall k \geq n \quad |f_{k+1}(x) - f_k(x)| < 2^{-k}.$$

Jetzt verwenden wir die folgende Eigenschaft der numerischen Folgen: gilt für eine reellwertige Folge $\{a_k\}$, dass für ein n

$$|a_{k+1} - a_k| < 2^{-k} \text{ für alle } k \geq n$$

so ist $\{a_k\}$ eine Cauchy-Folge, da für alle $m > k \geq n$ gilt

$$\begin{aligned} |a_k - a_m| &\leq |a_k - a_{k+1}| + |a_{k+1} - a_{k+2}| + ... + |a_{m-1} - a_m| \\ &\leq 2^{-k} + 2^{-(k+1)} + ... \leq 2^{-(k-1)} \to 0 \text{ für } k \to \infty. \end{aligned}$$

Somit ist die numerische Folge $\{f_k(x)\}_{k\in\mathbb{N}}$ eine Cauchy-Schwarz für jedes $x \in A^c$. Es folgt, dass die Folge $\{f_k(x)\}_{k\in\mathbb{N}}$ für alle $x \in A^c$ konvergiert.

Setzen wir

$$f(x) = \begin{cases} \lim\limits_{k\to\infty} f_k(x), & x \in A^c \\ 0, & x \in A \end{cases}$$

Dann gilt

$$f_k \overset{\text{f.ü.}}{\to} f \quad \text{für} \quad k \to \infty,$$

da $f_k(x) \to f(x)$ für alle x außerhalb der Nullmenge A. Nach dem Lemma 1.12 ist die Funktion f messbar.

Beweisen wir jetzt, dass $\|f_n - f\|_p \to 0$. Nach Definition der Cauchy-Folge gibt es für jedes $\varepsilon > 0$ ein $N \in \mathbb{N}$ so dass

$$\forall n, m \geq N \text{ gilt } \|f_n - f_m\|_p \leq \varepsilon,$$

d.h.

$$\int_X |f_n - f_m|^p \, d\mu \leq \varepsilon^p. \tag{1.40}$$

Da

$$f_m \overset{\text{f.ü.}}{\to} f \quad \text{für} \quad m \to \infty,$$

so gilt auch

$$|f_n - f_m|^p \overset{\text{f.ü.}}{\to} |f_n - f|^p \quad \text{für } m \to \infty.$$

Nach dem Fatou-Lemma und (1.40) erhalten wir

$$\int_X |f_n - f|^p \, d\mu \leq \varepsilon^p.$$

Es folgt, dass

$$\forall n \geq N \text{ gilt } \|f_n - f\|_p \leq \varepsilon$$

und somit

$$\|f_n - f\|_p \to 0 \text{ für } n \to \infty.$$

Es bleibt zu bemerken, dass $\|f\|_p$ endlich ist, da

$$\|f\|_p \leq \|f_n\|_p + \|f - f_n\|_p < \infty.$$

■

Es gibt auch den Raum L^∞ der auch vollständig ist – siehe Aufgaben 30, 31.

1.7.4 Der Raum $L^p_{\mathbb{C}}$

Betrachten wir jetzt komplexwertige messbare Funktionen und komplexwertige Lebesgue-Räume. Sei (X, S, μ) ein Maßraum wie zuvor.

Definition. Eine komplexwertige Funktion $f : X \to \mathbb{C}$ heißt messbar wenn $\operatorname{Re} f$ und $\operatorname{Im} f$ messbar sind. Eine messbare Funktion $f : X \to \mathbb{C}$ heißt integrierbar, wenn

$$\int_X |f| \, d\mu < \infty.$$

Da $|\operatorname{Re} f| \leq |f|$ und $|\operatorname{Im} f| \leq |f|$, so sind $\operatorname{Re} f$ und $\operatorname{Im} f$ auch integrierbar.

Definition. Für eine integrierbare Funktion $f : X \to \mathbb{C}$ definieren wir das Lebesgue-Integral mit

$$\int_X f d\mu = \int_X (\operatorname{Re} f) \, d\mu + i \int_X (\operatorname{Im} f) \, d\mu.$$

Man zeigt, das Integral linear ist, d.h.

$$\int_X (f+g) \, d\mu = \int_X f d\mu + \int_X g d\mu$$

und

$$\int_X cf \ d\mu = c \int_X f d\mu$$

für alle integrierbare Funktion f, g und alle $c \in \mathbb{C}$. Darüber hinaus gelten

$$\left| \int_X f d\mu \right| \leq \int_X |f| \, d\mu$$

und

$$\overline{\int_X f d\mu} = \int_X \overline{f} \ d\mu$$

wobei $\overline{w}$ die komplexe Konjugierte von w ist (siehe Aufgabe 32).

Definition. Für eine komplexwertige messbare Funktion f auch X definieren wir die p-Norm mit

$$\|f\|_p := \left(\int_X |f|^p \, d\mu \right)^{1/p} = \| |f| \|_p .$$

Die p-Norm ist offensichtlich nichtnegativ und erfüllt die absolute Homogenität über $\mathbb{C}$: für jedes $c \in \mathbb{C}$ gilt

$$\|cf\|_p = \left(\int_X |cf|^p \, d\mu \right)^{1/p} = \left(|c|^p \int_X |f|^p \, d\mu \right)^{1/p} = |c| \, \|f\|_p .$$

Die Hölder-Ungleichung und die Minkowski-Ungleichung gelten über $\mathbb{C}$ auch. Zum Beispiel, für alle messbare Funktionen $f, g : X \to \mathbb{C}$ gilt die Hölder-Ungleichung da

$$\int_X |fg| \, d\mu = \int_X |f| \, |g| \, d\mu \leq \| |f| \|_p \, \| |g| \|_q = \|f\|_p \, \|g\|_q .$$

Im Fall $p = q = 2$ erhalten wir die Cauchy-Schwarz-Ungleichung

$$\int_X |fg|\, d\mu \leq \|f\|_2 \|g\|_2 .$$

Die Minkowski-Ungleichung wird wie folgt bewiesen:

$$\|f + g\|_p = \||f + g|\|_p \leq \||f| + |g|\|_p \leq \||f|\|_p + \||g|\|_p = \|f\|_p + \|g\|_p .$$

Betrachten wir jetzt die Konvergenz bezüglich der p-Norm.

Definition. Eine Folge $\{f_n\}$ von komplexwertigen messbaren Funktionen auf X konvergiert gegen $f : X \to C$ bezüglich der p-Norm wenn

$$\|f_n - f\|_p \to 0 \text{ für } n \to \infty.$$

In diesen Fall schreibt man $f_n \overset{L^p}{\to} f$.

Lemma 1.21 *Seien $u = \operatorname{Re} f$ und $v = \operatorname{Im} f$. Dann gilt*

$$\frac{\|u\|_p + \|v\|_p}{2} \leq \|f\|_p \leq \|u\|_p + \|v\|_p . \tag{1.41}$$

Folglich ist die Konvergenz $f_n \overset{L^p}{\to} f$ äquivalent zu

$$\operatorname{Re} f_n \overset{L^p}{\to} \operatorname{Re} f \quad \textit{und} \quad \operatorname{Im} f_n \overset{L^p}{\to} \operatorname{Im} f.$$

Beweis. Da $f = u + iv$ so erhalten wir

$$\|f\|_p = \|u + iv\|_p \leq \|u\|_p + \|iv\|_p = \|u\|_p + \|v\|_p .$$

Da $|f| = \sqrt{u^2 + v^2}$ so erhalten wir $|f| \geq |u|$ und $|f| \geq |v|$ woraus folgt $\|f\|_p \geq \|u\|_p$ und $\|f\|_p \geq \|v\|_p$ und somit

$$\|f\|_p \geq \frac{\|u\|_p + \|v\|_p}{2}.$$

Die zweite Aussage folgt aus (1.41), da wir mit der Notation $f_n = u_n + iv_n$ erhalten

$$\frac{\|u_n - u\|_p + \|v_n - v\|_p}{2} \leq \|f_n - f\|_p \leq \|u_n - u\|_p + \|v_n - v\|_p$$

und somit

$$\begin{aligned} \|f_n - f\|_p \to 0 \quad &\Leftrightarrow \quad \|u_n - u\|_p + \|v_n - v\|_p \to 0 \\ &\Leftrightarrow \quad \|u_n - u\|_p \to 0 \text{ und } \|v_n - v\|_p \to 0. \end{aligned}$$

■

Bezeichnen wir mit $\mathcal{L}^p_{\mathbb{C}}$ die komplexwertige Version von $\mathcal{L}^p$, d.h.

$$\mathcal{L}^p_{\mathbb{C}} = \left\{ f : X \to \mathbb{C} : f \text{ is messbar und } \|f\|_p < \infty \right\}.$$

Die p-Norm ist eine Halbnorm auf $\mathcal{L}^p_{\mathbb{C}}$. Die Menge

$$\mathcal{N}_{\mathbb{C}} = \left\{ f : X \to \mathbb{C}: \ \ \|f\|_p = 0 \right\} = \{ f : X \to \mathbb{C}: \ \ f = 0 \text{ f.ü.} \}$$

ist offensichtlich ein Unterraum von $\mathcal{L}^p_{\mathbb{C}}$. Jetzt definieren wir den komplexwertigen Lebesgue-Raum als den Faktorraum

$$L^p_{\mathbb{C}} := \mathcal{L}^p_{\mathbb{C}} / \mathcal{N}_{\mathbb{C}}.$$

Wie im reellwertigen Fall, die p-Norm ist eine Norm in $L^p_{\mathbb{C}}$.Der Hauptsatz 1.18 gilt auch für $L^p_{\mathbb{C}}$.

Korollar 1.22 *Der Raum $L^p_{\mathbb{C}}$ mit der p-Norm ist ein Banachraum.*

Beweis. Sei $\{f_n\}$ eine Cauchy-Folge in $L^p_{\mathbb{C}}$. Nach dem Lemma 1.21 sind auch $\{\operatorname{Re} f_n\}$ und $\{\operatorname{Im} f_n\}$ Cauchy-Folgen in L^p. Somit konvergieren die Folgen $\{\operatorname{Re} f_n\}$ und $\{\operatorname{Im} f_n\}$ in L^p:

$$\operatorname{Re} f_n \overset{L^p}{\to} u, \quad \operatorname{Im} f_n \overset{L^p}{\to} v.$$

Nach dem Lemma 1.21 erhalten wir, dass $f_n \overset{L^p_{\mathbb{C}}}{\to} f := u + iv$. ■

1.7.5 $C[a,b]$ als Unterraum von $L^p[a,b]$

Fixieren wir reelle Zahlen $a < b$ und ein $p \in [1, \infty)$. Der Raum $L^p([a,b], \lambda)$ wird kurz mit $L^p[a,b]$ bezeichnet. Für jede stetige Funktion f auf $[a,b]$ ist $|f|^p$ auch stetig und somit integrierbar, woraus folgt, dass $f \in \mathcal{L}^p[a,b]$. Somit liegt die Äquivalenzklasse $[f]$ von f in $L^p[a,b]$. Weiter brauchen wir die folgende Aussage.

Lemma 1.23 *Für zwei stetige Funktionen f, g auf $[a,b]$ gilt*

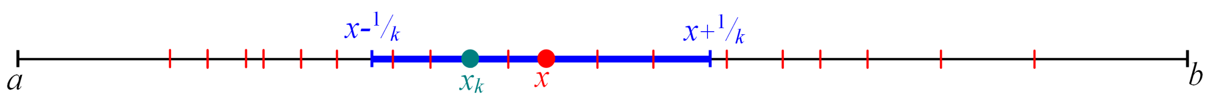
$$f = g \ \text{ f.ü.} \ \Leftrightarrow \ f(x) = g(x) \ \ \forall x \in [a,b].$$

Beweis. Sei

$$N = \{x \in [a,b] : f(x) \neq g(x)\}.$$

Die Bedingung $f = g$ f.ü. bedeutet, dass $\lambda(N) = 0$. Beweisen wir, dass $N = \emptyset$. Sei N nichtleer, dann wählen wir ein $x \in N$. Für jedes $k \in N$ betrachten wir das Intervall

$$J_k = \left[x - \tfrac{1}{k}, x + \tfrac{1}{k}\right] \cap [a,b].$$



Da

$$\lambda(J_k) > 0 = \lambda(N),$$

so ist J_k keine Teilmenge von N. Somit existiert $x_k \in J_k \setminus N$. Da $x_k \notin N$, so gilt

$$f(x_k) = g(x_k).$$

Da $x_k \to x$ für $k \to \infty$ und die Funktionen f und g stetig sind, daraus folgt, dass auch

$$f(x) = g(x),$$

was im Widerspruch zur Annahme $x \in N$ steht. Somit ist N leer, und $f(x) = g(x)$ für alle $x \in [a,b]$. ∎

Es folgt, dass für $f, g \in C[a,b]$

$$[f] = [g] \Leftrightarrow f = g$$

so dass die Abbildung

$$C[a,b] \ni f \mapsto [f] \in L^p[a,b]$$

injektiv ist. Somit lässt sich $C[a,b]$ als einen Unterraum von $L^p[a,b]$ betrachten.

Wir beweisen später, dass $C[a,b]$ in $L^p[a,b]$ dicht liegt. Der Banachraum $L^p[a,b]$ ist somit eine *Vervollständigung* von $C[a,b]$ bezüglich der p-Norm.

Chapter 2

Hilbertraum

2.1 Skalarprodukt

Sei V ein Vektorraum über $\mathbb{R}$. Sei $\sigma(x, y)$ eine Funktion auf $V \times V$ mit Werten in $\mathbb{R}$.

Definition. Die Funktion σ heißt *Skalarprodukt* (oder inneres Produkt) auf V wenn sie die folgenden Axiome erfüllt:

$(S1)$ positive Definitheit: $\sigma(x, x) \geq 0$ und $\sigma(x, x) = 0 \Leftrightarrow x = 0$.

$(S2)$ Linearität in x:

$$\sigma(x_1 + x_2, y) = \sigma(x_1, y) + \sigma(x_2, y) \quad \text{und} \quad \sigma(\alpha x, y) = \alpha\sigma(x, y)$$

für alle $x_1, x_2, x, y \in V$ und $\alpha \in \mathbb{R}$.

$(S3)$ Symmetrie: $\sigma(y, x) = \sigma(x, y)$.

Ein Vektorraum mit Skalarprodukt heißt *Skalarproduktraum* (oder *Prähilbertraum*)

Es folgt aus $(S2)$ und $(S3)$ dass $\sigma(x, y)$ auch in y linear ist. Eine Funktion $\sigma(x, y)$ die linear in x und y ist heißt eine *Bilinearform*. Ein Skalarprodukt ist somit eine symmetrische positiv definite Bilinearform.

Normalerweise bezeichnet man das Skalarprodukt mit (x, y) statt $\sigma(x, y)$.

Beispiel. In $\mathbb{R}^n$ definiert man das Skalarprodukt mit

$$(x, y) = \sum_{k=1}^{n} x_k y_k,$$

und es wird auch mit $x \cdot y$ bezeichnet.

Beispiel. In l^2 kann man analog das Skalarprodukt definieren:

$$(x, y) = \sum_{k=1}^{\infty} x_k y_k.$$

Da $\sum_{k=1}^{\infty} x_k^2 < \infty$ und $\sum_{k=1}^{\infty} y_k^2 < \infty$, so konvergiert die Reihe $\sum_{k=1}^{\infty} x_k y_k$ absolut nach der Cauchy-Schwarz Ungleichung. Offensichtlich sind alle Axiome von Skalarprodukt erfüllt.

Beispiel. In $C[a,b]$ mit reellen $a < b$ definiert man das Skalarprodukt mit

$$(f,g) = \int_a^b f(t)\, g(t)\, dt. \tag{2.1}$$

Beispiel. Betrachten wir den Lebesgue-Raum L^2 auf einem Maßraum (X, S, μ), und setzen wir für $f, g \in L^2$

$$(f,g) = \int_X fg d\mu.$$

Nach der Cauchy-Schwarz Ungleichung gilt

$$\int_X |fg|\, d\mu \leq \|f\|_2 \|g\|_2 < \infty$$

so dass die Funktion fg integrierbar ist und somit (f,g) wohldefiniert ist. Die Symmetrie und Linearität sind offensichtlich. Auch haben wir

$$(f,f) = \int_X f^2 d\mu = \|f\|_2^2 \geq 0.$$

Gilt $(f,f) = 0$, so erhalten wir $\|f\|_2 = 0$ und somit $f = 0$ als Element von L^2. Daraus folgt die positive Definitheit. Deshalb is (f,g) ein Skalarprodukt in L^2.

Sei jetzt V ein Vektorraum über $\mathbb{C}$.

Definition. Eine Funktion $\sigma : V \times V \to \mathbb{C}$ heißt Skalarprodukt (oder *Hermitesche Form*) auf V wenn sie die folgenden Axiome erfüllt:

$(H1)$ positive Definitheit: für alle $x \in V$ ist $\sigma(x,x)$ reel, $\sigma(x,x) \geq 0$ und $\sigma(x,x) = 0 \Leftrightarrow x = 0$.

$(H2)$ Linearität in x:

$$\sigma(x_1 + x_2, y) = \sigma(x_1, y) + \sigma(x_2, y) \quad \text{und} \quad \sigma(\alpha x, y) = \alpha\sigma(x,y) \tag{2.2}$$

für alle $x_1, x_2, x, y \in V$ und $\alpha \in \mathbb{C}$.

$(H3)$ Hermitesche Symmetrie: $\sigma(y,x) = \overline{\sigma(x,y)}$ (wobei $\overline{w}$ die komplexe Konjugierte von w ist)

Es folgt aus $(H2)$ und $(H3)$ dass

$$\sigma(x, y_1 + y_2) = \sigma(x, y_1) + \sigma(x, y_2) \quad \text{und} \quad \sigma(x, \alpha y) = \overline{\alpha}\sigma(x,y), \tag{2.3}$$

da

$$\sigma(x, y_1 + y_2) = \overline{\sigma(y_1 + y_2, x)} = \overline{\sigma(y_1, x)} + \overline{\sigma(y_2, x)} = \sigma(x, y_1) + \sigma(x, y_2)$$

und

$$\sigma(x, \alpha y) = \overline{\sigma(\alpha y, x)} = \overline{\alpha\sigma(y,x)} = \overline{\alpha}\,\overline{\sigma(y,x)} = \overline{\alpha}\sigma(x,y).$$

Man sagt, dass $\sigma(x,y)$ *semilinear* in y ist. Eine Funktion $\sigma(x,y)$ die (2.2) und (2.3) erfüllt heißt eine *Sesquilinearform.* Ein Skalarprodukt ist somit eine Hermit-symmetrische positiv definite Sesquilinearform.

Beispiel. In $\mathbb{C}^n$ definiert man das Skalarprodukt mit

$$(x, y) = \sum_{k=1}^{n} x_k \overline{y_k}.$$

Z.B. $(H1)$ gilt da

$$(x, x) = \sum_{k=1}^{n} x_k \overline{x_k} = \sum_{k=1}^{n} |x_k|^2 \geq 0$$

und $(x, x) = 0 \Leftrightarrow$alle $x_k = 0 \Leftrightarrow x = 0$.

Beispiel. Betrachten wir den komplexwertigen Raum l^2:

$$l^2_{\mathbb{C}} = \left\{ \{x_k\}_{k=1}^{\infty} : x_k \in \mathbb{C} \text{ und } \sum_{k=1}^{\infty} |x_k|^2 < \infty \right\}.$$

Das Skalarprodukt in $l^2_{\mathbb{C}}$ ist ähnlich definiert:

$$(x, y) = \sum_{k=1}^{\infty} x_k \overline{y_k}.$$

Beispiel. Betrachten wir den Lebesgue-Raum $L^2_{\mathbb{C}}$ auf einem Maßraum (X, S, μ) und definieren wir das Skalarprodukt mit

$$(f, g) = \int_X f \overline{g} \, d\mu.$$

Nach der Cauchy-Schwarz -Ungleichung ist die Funktion $f\overline{g}$ integrierbar, da

$$\int_X |f \overline{g}| \, d\mu \leq \|f\|_2 \, \|\overline{g}\|_2 = \|f\|_2 \, \|g\|_2 < \infty.$$

Beweisen wir, dass (f, g) ein Skalarprodukt ist. Die positive Definitheit gilt, weil

$$(f, f) = \int_X f \ \overline{f} \ d\mu = \int_X |f|^2 \, d\mu = \|f\|_2^2 \geq 0,$$

und

$$(f, f) = 0 \Leftrightarrow \|f\|_2 = 0 \Leftrightarrow f = 0 \text{ als Element von } L^2.$$

Die Linearität in f ist offensichtlich. Die Hermitesche Symmetrie gilt, da

$$(g, f) = \int_X g \overline{f} \, d\mu = \int_X \overline{\overline{g} f} \, d\mu = \overline{\int_X f \overline{g} \, d\mu} = \overline{(f, g)}.$$

2.2 Skalarproduktnorm

Wir haben gesehen, dass in den Räumen L^2 und $L^2_{\mathbb{C}}$ die Cauchy-Schwarz-Ungleichung gilt

$$|(f,g)| \leq \|f\|_2 \|g\|_2 ,$$

und die folgende Identität gilt:

$$\|f\|_2 = \sqrt{(f,f)}.$$

Die ähnlichen Eigenschaften gelten in jedem Skalarproduktraum.

Satz 2.1 *In jedem Skalarproduktraum V über $\mathbb{R}$ oder $\mathbb{C}$ ist die Funktion*

$$\|x\| := \sqrt{(x,x)}$$

immer eine Norm. Darüber hinaus gilt es für alle $x, y \in V$

$$|(x,y)| \leq \|x\| \|y\| \tag{2.4}$$

(die allgemeine Cauchy-Schwarz-Ungleichung).

Die Norm $x \mapsto \sqrt{(x,x)}$ heißt die *Skalarproduktnorm.* Z.B. die 2-Norm in L^2 ist eine Skalarproduktnorm. Man kann beweisen, dass die p-Norm in L^p mit $p \neq 2$ keine Skalarproduktnorm ist (siehe Aufgabe 35).

29.04.22 Vorlesung 7

Beweis. Die Funktion $\|x\| = \sqrt{(x,x)}$ ist reel, nicht-negativ und erfüllt die absolute Homogenität, da für jedes $\alpha \in \mathbb{C}$

$$\|\alpha x\| = \sqrt{(\alpha x, \alpha x)} = \sqrt{\alpha \overline{\alpha}\,(x,x)} = |\alpha| \|x\| .$$

Die Definitheit gilt auch, da

$$\|x\| = 0 \Leftrightarrow (x,x) = 0 \Leftrightarrow x = 0.$$

Es bleibt nur die Dreiecksungleichung (= die Minkowski-Ungleichung)

$$\|x+y\| \leq \|x\| + \|y\|$$

zu beweisen, was machen wir nach dem Beweis von (2.4).

Nach Eigenschaften von Skalarprodukt haben wir für jedes $t \in \mathbb{R}$

$$(x+ty, x+ty) \geq 0$$

und

$$(x+ty, x+ty) = (x,x) + t\,(x,y) + t\,(y,x) + t^2\,(y,y) .$$

Da

$$(x,y) + (y,x) = (x,y) + \overline{(x,y)} = 2\operatorname{Re}(x,y) ,$$

so erhalten wir

$$\boxed{(x+ty, x+ty) = \|x\|^2 + 2t\operatorname{Re}(x,y) + t^2 \|y\|^2}$$

woraus folgt
$$\|x\|^2 + 2t\,\mathrm{Re}\,(x,y) + t^2\,\|y\|^2 \geq 0 \;\; \text{für alle } t \in \mathbb{R}.$$
Da die quadratische Funktion
$$t \mapsto \|x\|^2 + 2t\,\mathrm{Re}\,(x,y) + t^2\,\|y\|^2$$
für alle $t \in \mathbb{R}$ nicht-negativ ist, so ist ihre Diskriminante nichtpositiv, d.h.
$$\mathrm{Re}\,(x,y)^2 - \|x\|^2\,\|y\|^2 \leq 0,$$
woraus folgt
$$|\mathrm{Re}\,(x,y)| \leq \|x\|\,\|y\|\,. \tag{2.5}$$
Für einen reellwertigen Raum V ist diese Ungleichung äquivalent zu (2.4). Im Fall von komplexwertigen Raum finden wir eine komplexe Zahl α mit $|\alpha| = 1$ und so dass $\alpha\,(x,y)$ reel ist. Dann erhalten wir mit Hilfe von (2.5)
$$|(x,y)| = |\alpha\,(x,y)| = |\mathrm{Re}\,(\alpha\,(x,y))| = |\mathrm{Re}\,(\alpha x, y)| \leq \|\alpha x\|\,\|y\| = \|x\|\,\|y\|\,,$$
was (2.4) beweist.

Die Dreiecksungleichung folgt aus (2.5) wie folgt:
$$\begin{aligned}\|x+y\|^2 &= (x+y, x+y) = \|x\|^2 + 2\,\mathrm{Re}\,(x,y) + \|y\|^2 \\ &\leq \|x\|^2 + 2\,\|x\|\,\|y\| + \|y\|^2 \\ &= (\|x\| + \|y\|)^2\,,\end{aligned}$$
was äquivalent zu $\|x+y\| \leq \|x\| + \|y\|$ ist. ∎

Wir haben oberhalb die folgende Identität bewiesen:
$$\boxed{\|x+y\|^2 = \|x\|^2 + 2\,\mathrm{Re}\,(x,y) + \|y\|^2}, \tag{2.6}$$
die auch später häufig verwendet wird.

Zweiter Beweis. In diesem Beweis nehmen wir an, dass V ein Skalarproduktraum über $\mathbb{R}$ ist. Fixieren wir $x, y \in V$. Sei U ein 2-dimensionaler Unterraum von V der x und y enthält. Dann U ist auch ein Skalarproduktraum. Jetzt benutzen wir den Satz aus der Linearen Algebra: alle n-dimensionale Skalarprodukträume sind isometrisch. Somit ist U isometrisch zu $\mathbb{R}^2$. Deshalb können wir annehmen, dass $x, y \in \mathbb{R}^2$ und somit auch, dass x und y komplexe Zahlen sind. Für $x = x_1 + ix_2$ gilt
$$\|x\|^2 = x_1^2 + x_2^2 = |x|^2\,,$$
wobei $\|x\|$ die Skalarproduktnorm ist und $|x|$ der Betrag der komplexen Zahl x ist. Für Betrag gilt die Dreiecksungleichung
$$|x+y| \leq |x| + |y|\,,$$
was äquivalent zur
$$\|x+y\| \leq \|x\| + \|y\|\,.$$
Da
$$x\overline{y} = (x_1 + ix_2)\,(y_1 - iy_2) = (x_1y_1 + x_2y_2) + i\,(x_2y_1 - x_1y_2)\,,$$
so gilt
$$|(x,y)| = |x_1y_1 + x_2y_2| = |\mathrm{Re}\,x\overline{y}| \leq |x\overline{y}| = |x|\,|y| = \|x\|\,\|y\|\,,$$
was (2.4) beweist. ∎

2.3 Geometrische Eigenschaften

Sei V ein Skalarproduktraum über $\mathbb{R}$ oder $\mathbb{C}$.

Definition. Zwei Vektoren $x, y \in V$ heißen *orthogonal*, wenn $(x, y) = 0$. Man schreibt in diesem Fall $x \perp y$.

Für orthogonale Vektoren gilt der Satz des Pythagoras:

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

Beweis ist trivial, da in diesem Fall

$$\|x+y\|^2 = \|x\|^2 + \|y\|^2 + 2\,\mathrm{Re}\,(x, y) = \|x\|^2 + \|y\|^2 .$$

Dieser Satz erlaubt eine Verallgemeinerung für n orthogonale Vektoren: sind $x_1, ..., x_n$ zueinander orthogonale Vektoren in V so gilt

$$\|x_1 + ... + x_n\|^2 = \|x_1\|^2 + ... + \|x_n\|^2$$

(siehe Aufgabe 33).

Für beliebige Vektoren $x, y \in V$ gilt die Parallelogrammgleichung:

$$\|x+y\|^2 + \|x-y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right), \tag{2.7}$$

die man analog beweist (siehe Aufgabe 34).

Es gibt auch die folgende Umkehrung der Parallelogrammgleichung: gilt (2.7) für eine Norm in einem normierten Raum, so ist sie eine Skalarproduktnorm, d.h. es existiert ein Skalarprodukt $(\cdot, \cdot)$ mit $\|x\| = \sqrt{(x, x)}$ (siehe Aufgabe 36). In anderen Wörtern heißt das, dass die Parallelogrammgleichung in einem normierten Raum genau dann gilt, wenn dieser Raum ein Skalarproduktraum ist.

2.4 Definition von Hilbertraum

Definition. Ein Hilbertraum ist ein vollständiger Skalarproduktraum; d.h. ein Hilbertraum ist ein Vektorraum mit Skalarprodukt, der vollständig bezüglich der Skalarproduktnorm ist.

Beispiel. Jeder endlichdimensionaler Skalarproduktraum ist Hilbertraum, da endlichdimensionaler normierter Raum immer vollständig ist (Satz 1.6). Insbesondere sind $\mathbb{R}^n$ und $\mathbb{C}^n$ mit standarten Skalarprodukten die Hilberträume.

Der Raum l^2 (und auch $l^2_{\mathbb{C}}$) ist ein Hilbertraum nach dem Satz 1.6.

Der Raum $L^2(X, \mu)$ ist ein Hilbertraum nach dem Satz 1.18, und so ist $L^2_{\mathbb{C}}(X, \mu)$.

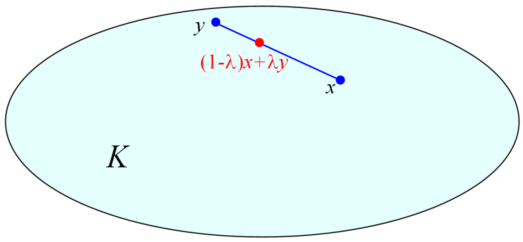
Der Raum $C[a, b]$ mit dem Skalarprodukt (2.1) ist ein Skalarproduktraum aber kein Hilbertraum, da $C[a, b]$ bezüglich der 2-Norm nicht vollständig ist.

2.5 Konvexe Mengen im Hilbertraum

Sei V ein Vektorraum über $\mathbb{R}$ oder $\mathbb{C}$.

Definition. Eine Menge $K \subset V$ heißt konvex, wenn für alle $x, y \in K$ und $\lambda \in [0,1]$ gilt

$$\lambda x + (1-\lambda)\, y \in K.$$



Die Menge

$$[x,y] := \{\lambda x + (1-\lambda)\, y : 0 \le \lambda \le 1\}$$

heißt die *Verbindungsstrecke* zwischen x und y. So ist die Menge K konvex wenn sie mit jeden zwei Punkten x, y auch die Verbindungsstrecke $[x,y]$ enthält.

Beispiel. Jeder Unterraum U von V ist konvex, da für alle $x, y \in U$ auch beliebige lineare Kombination von x, y in U liegt.

Beispiel. Sei $(V, \|\cdot\|)$ ein normierter Raum. Für jedes $R > 0$ betrachten wir eine Kugel

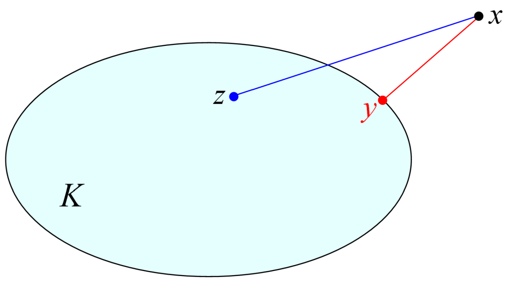
$$B_R = \{z \in V : \|z\| \le R\}.$$

Dann ist B_R konvex (siehe Aufgabe 37).

Die konvexen Mengen im Hilbertraum haben eine besondere Eigenschaft.

Hauptsatz 2.2 *Seien H ein Hilbertraum (über $\mathbb{R}$ oder $\mathbb{C}$) und $K \subset H$ eine abgeschlossene konvexe Menge. Dann für jedes $x \in H$ gibt es genau ein $y \in K$ mit dem minimalen Abstand $\|x-y\|$, d.h. mit*

$$\|x-y\| = \min\{\|x-z\| : z \in K\}. \tag{2.8}$$



Jedes $y \in K$ mit (2.8) heißt die *beste Approximation* von x aus K. Der Satz 2.2 sichert die Existenz und Eindeutigkeit der besten Approximation wenn K konvex und abgeschlossen ist.

Für jede Teilmenge $K \subset H$ und jedes $x \in H$ heißt der Wert $\inf_{z\in K} \|x-z\|$ der *Abstand* von x zu K. Für allgemeine Teilmenge K (die nicht abgeschlossen oder konvex ist) soll weder die Existenz einer Minimumstelle noch die Eindeutigkeit der Fall sein.

Beispiel. In Banachräumen ohne Skalarprodukt fehlt die Eindeutigkeit von y sogar für abgeschlossene konvexe Mengen. Zum Beispiel, betrachten wir $\mathbb{R}^2$ mit der 1-Norm, die abgeschlossene Kugel

$$K = \left\{z \in \mathbb{R}^2 : \|z\|_1 \leq 1\right\},$$

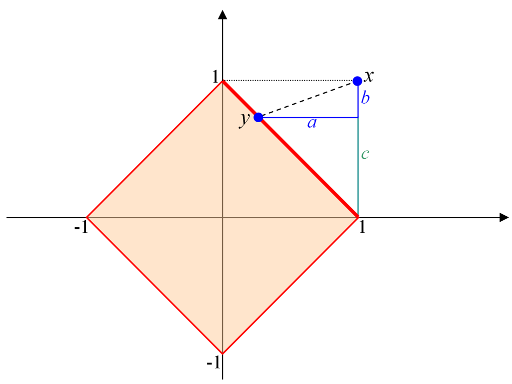
die auch konvex ist, und den Punkt $x = (1,1)$. Dann jeder Punkt $y = (y_1, y_2)$ mit $y_1 + y_2 = 1$ und nichtnegativen y_1, y_2 ist die beste Approximation von x, da

$$\|x-y\|_1 = |1-y_1| + |1-y_2| = 1 - y_1 + 1 - y_2 = 2 - (y_1+y_2) = 1,$$

während für alle $z \in K$ gilt

$$\|x-z\|_1 \geq \|x\|_1 - \|z\|_1 \geq 2 - 1 = 1.$$

Einen anderen Beweis sieht man auf dem Bild:



Es gilt $\|x-y\|_1 = a + b = c + b = 1$.

Beweis von dem Satz 2.2. Ohne Beschränkung der Allgemeinheit nehmen wir $x = 0$ an. Setzen wir

$$a = \inf\{\|z\| : z \in K\}. \tag{2.9}$$

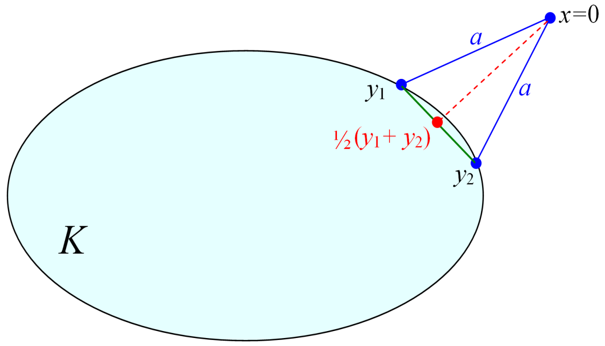
Wir müssen die Existenz und Eindeutigkeit von $y \in K$ mit $\|y\| = a$ beweisen.

Zunächst beweisen wir die Eindeutigkeit von y. Gilt $\|y_1\| = \|y_2\| = a$ für zwei verschiedene Punkte $y_1, y_2 \in K$, so erhalten wir nach der Parallelogrammgleichung

$$\|y_1+y_2\|^2 = 2\|y_1\|^2 + 2\|y_2\|^2 - \|y_1-y_2\|^2 < 4a^2,$$

woraus folgt

$$\left\|\frac{y_1+y_2}{2}\right\| < a. \tag{2.10}$$



Allerdings haben wir nach der Konvexität von K, dass

$$\frac{y_1+y_2}{2} \in [y_1, y_2] \subset K,$$

und somit nach (2.9)

$$\left\|\frac{y_1+y_2}{2}\right\| \geq a$$

was im Widerspruch zu (2.10) steht.

Jetzt beweisen wir die Existenz von y. Nach der Definition (2.9) von a gibt es eine Folge $\{z_n\}_{n=1}^{\infty}$ aus K mit $\|z_n\| \to a$ für $n \to \infty$. Die Folge $\{z_n\}$ heißt die *minimierende Folge*. Wir beweisen, dass diese Folge konvergiert und der Grenzwert $y = \lim z_n$ in K liegt und erfüllt $\|y\| = a$.

Zunächst zeigen wir, dass die Folge $\{z_n\}$ eine Cauchy-Folge ist. Nach der Parallelogrammgleichung haben wir für alle $n, m \geq 1$

$$\|z_n - z_m\|^2 = 2\|z_n\|^2 + 2\|z_m\|^2 - 4\left\|\frac{z_n+z_m}{2}\right\|^2.$$

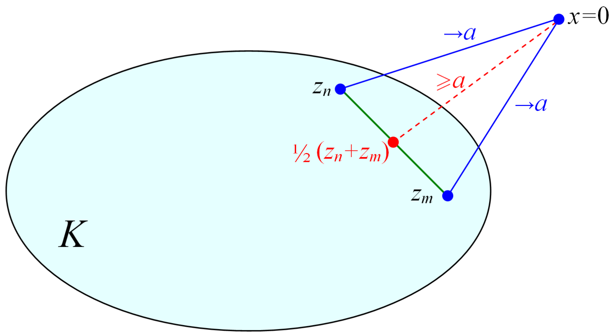
Da $\frac{z_n+z_m}{2} \in K$, so gilt $\left\|\frac{z_n+z_m}{2}\right\| \geq a$ und somit

$$\|z_n - z_m\|^2 \leq 2\|z_n\|^2 + 2\|z_m\|^2 - 4a^2.$$

Für $n, m \to \infty$ erhalten wir $\|z_n\| \to a$ und $\|z_m\| \to a$ woraus folgt

$$\limsup_{n,m\to\infty} \|z_n - z_m\|^2 \leq 2a^2 + 2a^2 - 4a^2 = 0$$

und $\lim_{n,m\to\infty} \|z_n - z_m\| = 0$.



Somit ist $\{z_n\}$ eine Cauchy-Folge in K. Nach der Vollständigkeit von H hat diese Folge den Grenzwert $y = \lim_{n\to\infty} z_n$. Da K abgeschlossen ist, gilt $y \in K$. Die Bedingung $z_n \to y$ impliziert nach der Dreiecksungleichung, dass

$$|\|y\| - \|z_n\|| \leq \|y - z_n\| \to 0 \text{ für } n \to \infty$$

und somit $\|z_n\| \to \|y\|$, woraus $\|y\| = a$ folgt. ∎

Bemerkung. Der Beweis funktioniert auch im Fall, wenn H ein Skalarproduktraum ist und K eine konvexe Teilmenge von H ist, die vollständig als metrischer Raum ist.

Definition. Seien V ein Skalarproduktraum und M eine Teilmenge von V. Ein Vektor $x \in V$ heißt orthogonal zu M wenn $x \perp z$ für alle $z \in M$. Man schreibt in diesem Fall $x \perp M$, d.h.

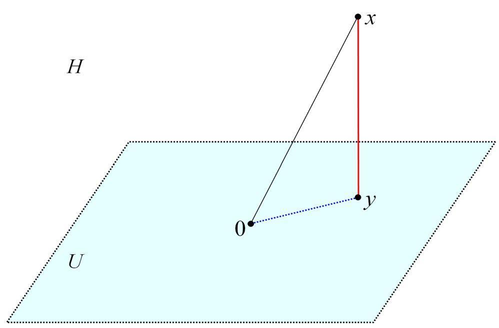
$$x \perp M \Leftrightarrow x \perp z \quad \forall z \in M.$$

Satz 2.3 *Seien H ein Hilbertraum (über $\mathbb{R}$ oder $\mathbb{C}$) und U ein abgeschlossener Unterraum von H. Seien $x \in H$ und $y \in U$. Dann sind die folgenden Eigenschaften äquivalent:*

(a) *y ist die beste Approximation von x in U;*

(b) *$x - y \perp U$.*

Folglich, für jedes $x \in H$ existiert genau ein $y \in U$ mit $x - y \perp U$.



Definition. Der Punkt $y \in U$ mit $x - y \perp U$ heißt die *orthogonale Projektion* von x auf U.

Mit Hilfe von diesem Begriff kann der Satz 2.3 wie folgt umformuliert werden:

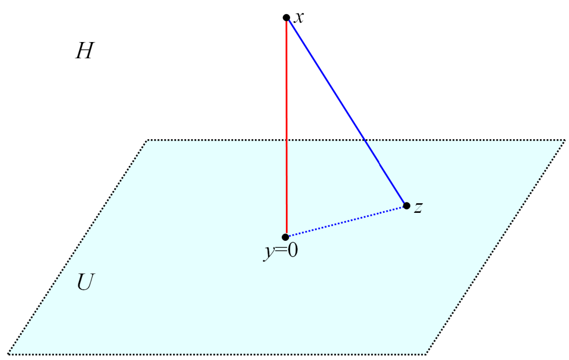
- ein Punkt $y \in U$ ist die beste Approximation von x in U genau dann, wenn y die orthogonale Projektion von x auf U ist;

- die orthogonale Projektion auf U existiert immer und ist eindeutig bestimmt.

Beweis. Die Existenz und Eindeutigkeit der orthogonalen Projektion folgt aus dem Satz 2.2 und der Äquivalenz $(a) \Leftrightarrow (b)$. Wir brauchen nur diese Äquivalenz zu beweisen. Ohne Beschränkung der Allgemeinheit nehmen wir an, dass $y = 0$.

$(b) \Rightarrow (a)$ Sei x orthogonal zu U, beweisen wir, dass 0 die beste Approximation von x in U ist. Für alle $z \in U$ haben wir $x \perp z$ und somit

$$\|x - z\|^2 = \|x\|^2 + \|z\|^2 .$$



Es ist klar, dass $\|x\|^2 + \|z\|^2$ minimal genau dann ist, when $z = 0$.

$(a) \Rightarrow (b)$ Sei 0 die beste Approximation von x in U. Beweisen wir, dass $x \perp U$, d.h. $(x, z) = 0$ für alle $z \in U$. Fixieren wir in $z \in U$ und betrachten die folgende Funktion von $t \in \mathbb{R}$:

$$f(t) = \|x - tz\|^2 = \|x\|^2 - 2t \operatorname{Re}(x, z) + t^2 \|z\|^2 .$$

Diese Funktion nimmt das Minimum an $t = 0$ an, da $f(t)$ der Abstand zwischen x und $tz \in U$ ist. Somit gilt $f'(0) = 0$. Da

$$f'(t) = -2 \operatorname{Re}(x, z) + 2t \|z\|^2$$

und $f'(0) = -2 \operatorname{Re}(x, z)$, so folgt es $\operatorname{Re}(x, z) = 0$.

Im reellwertigen Hilbertraum ist es äquivalent zu $(x, z) = 0$, d.h. $x \perp z$. Im komplexwertigen Fall wählen wir eine Zahl $\alpha \in \mathbb{C}$ mit $|\alpha| = 1$ und so dass $(x, \alpha z) = \overline{\alpha}(x, z)$ reell ist. Da $\alpha z \in U$, so gilt

$$(x, \alpha z) = \operatorname{Re}(x, \alpha z) = 0$$

woraus $(x, z) = 0$ folgt. ■

04.05.22 Vorlesung 8

2.6 Orthogonale Projektoren

Seien H ein Hilbertraum über den Körper $\mathbb{K}$, wobei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$.

Definition. Für jede Teilmenge $M \subset H$ definieren wir die Menge $M^\perp$ wie folgt:

$$M^\perp = \{x \in H : x \perp M\} = \{x \in H : (x, y) = 0 \ \forall y \in M\} .$$

Die Menge $M^\perp$ heißt das *orthogonale Komplement* von M.

Es gilt immer

$$M \subset \left(M^\perp\right)^\perp,$$

da jedes $y \in M$ orthogonal zu $M^\perp$ nach Definition ist.

Lemma 2.4 *Das orthogonale Komplement $M^\perp$ ist immer ein abgeschlossener Unterraum.*

Beweis. Sind $x_1, x_2 \in M^\perp$, so es folgt, dass

$$(x_1 + x_2, y) = (x_1, y) + (x_2, y) = 0$$

für alle $y \in M$, woraus folgt $x_1 + x_2 \in M^\perp$. Analog gilt $\alpha x \in M^\perp$ für alle $x \in M^\perp$ und Skalare $\alpha \in \mathbb{K}$. Somit ist $M^\perp$ ein Unterraum.

Sei $\{x_n\}$ eine Folge aus $M^\perp$, die einen Grenzwert $x = \lim x_n$ in H hat. Dann gilt für alle $y \in H$

$$(x, y) = \lim_{n\to\infty} (x_n, y)$$

da

$$|(x, y) - (x_n, y)| = |(x - x_n, y)| \leq \|x - x_n\| \, \|y\| \to 0.$$

Da $(x_n, y) = 0$ für alle $y \in M$, so folgt es, dass auch $(x, y) = 0$ und somit $x \in M^\perp$. Daher ist $M^\perp$ abgeschlossen. ■

Bemerkung. Im Beweis haben wir die folgende Eigenschaft von Skalarprodukt gezeigt: die Konvergenz $x_n \to x$ in V ergibt die Konvergenz $(x_n, y) \to (x, y)$ für alle $y \in V$.

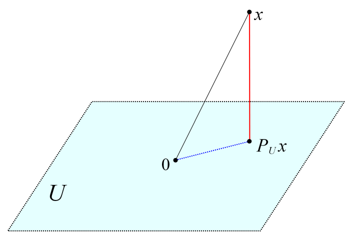
Sei U ein Unterraum von H. Bemerken wir, dass U abgeschlossen genau dann ist, wenn U vollständig ist (und somit auch ein Hilbertraum ist). Insbesondere ist eine endlichdimensionaler Unterraum U immer abgeschlossen. Unterhalb betrachten wir die Eigenschaften der Projektion auf den abgeschlossenen Unterräumen von H.

Definition. Sei U ein abgeschlossener Unterraum von einem Hilbertraum H.

Der (orthogonale) Projektor P_U auf U ist die Abbildung

$$P_U : H \to H$$

so dass $P_U x$ die orthogonale Projektion von x auf U ist.



D.h. $P_U x$ wird mit den folgenden Bedingungen eindeutig bestimmt:

$$\boxed{P_U x \in U \quad \text{und} \quad x - P_U x \perp U}.$$

Nach dem Satz 2.3 ist P_U wohldefiniert.

Satz 2.5 *Der Projektor $P = P_U$ hat die folgenden Eigenschaften.*

(a) *P ist eine lineare Abbildung.*

(b) $\operatorname{im} P = U$ *und* $\ker P = U^\perp$

(c) (Idempotenz) $P^2 = P$

(d) $\|Px\| \leq \|x\|$ $\forall x \in H$.

(e) (Symmetrie) $(Px, y) = (x, Py)$

(f) *Es gilt die Identität* $P_U + P_{U^\perp} = \mathrm{Id}$.

Beweis. Für jedes x haben wir nach Definition

$$x - Px \perp U,$$

was äquivalent zu

$$x - Px \in U^\perp. \tag{2.11}$$

(a) Beweisen wir, dass für alle $x, y \in H$

$$P(x+y) = Px + Py.$$

Da $x - Px \in U^\perp$ und $y - Py \in U^\perp$, so erhalten wir

$$x - Px + y - Py \in U^\perp,$$

d.h.

$$(x+y) - (Px + Py) \in U^\perp.$$

Da $P(x+y)$ ein eindeutiger Vektor mit

$$(x+y) - P(x+y) \in U^\perp$$

ist daraus folgt, dass $Px + Py = P(x+y)$. Analog beweist man, dass $P(\alpha x) = \alpha Px$ für alle $\alpha \in \mathbb{K}$.

(b) Nach der Definition von P haben wir im $P \subset U$. Für jedes $x \in U$ gilt offensichtlich $Px = x$, woraus folgt im $P = U$. Für den Kern haben wir nach (2.11)

$$\ker P = \{x \in H : Px = 0\} = \{x \in H : x \perp U\} = U^\perp.$$

(c) Da $Px \in U$, so erhalten wir $P^2 x = P(Px) = Px$, woraus folgt $P^2 = P$.

(d) Da

$$x = Px + (x - Px)$$

und $Px \in U$, $x - Px \in U^\perp$, so erhalten wir nach dem Satz von Pythagoras

$$\|x\|^2 = \|Px\|^2 + \|x - Px\|^2$$

woraus $\|Px\| \leq \|x\|$ folgt.

(e) Da $Px \in U$ und $y - Py \in U^\perp$, so haben wir

$$(Px, y - Py) = 0$$

und somit

$$(Px, y) = (Px, Py).$$

Analog haben wir

$$(Py, x) = (Py, Px)$$

woraus folgt

$$(x, Py) = (Px, Py)$$

und somit $(Px, y) = (x, Py)$.

(f) Die Identität $P_U + P_{U^\perp} = \mathrm{Id}$ bedeutet, dass für alle $x \in H$

$$P_U x + P_{U^\perp} x = x. \tag{2.12}$$

Setzen wir $z = x - P_U x$ und bemerken, dass

$$z \in U^\perp \text{ und } \; x - z = P_U x \perp U^\perp,$$

da $P_U x \subset U$. Nach Definition ist z die orthogonale Projektion von x auf $U^\perp$, d.h. $z = P_{U^\perp} x$, woraus folgt

$$x - P_U x = P_{U^\perp} x$$

und somit (2.12). ■

Bemerkung. Nach der Definition der Operatornorm folgt es aus (d) dass

$$\|P_U\| := \sup_{x \neq 0} \frac{\|P_U x\|}{\|x\|} \leq 1,$$

und wenn $U \neq \{0\}$ dann $\|P_U\| = 1$.

Korollar 2.6 *Für jeden abgeschlossenen Unterraum U gilt $\left(U^\perp\right)^\perp = U$.*

Beweis. Da

$$P_U + P_{U^\perp} = \mathrm{Id} = P_{U^\perp} + P_{\left(U^\perp\right)^\perp},$$

es folgt, dass $P_U = P_{\left(U^\perp\right)^\perp}$ und somit

$$U = \operatorname{im} P_U = \operatorname{im} P_{\left(U^\perp\right)^\perp} = \left(U^\perp\right)^\perp.$$

■

2.7 Begriff von Winkel

Definition. Für zwei Vektoren $x, y \neq 0$ im reellwertigen Hilbertraum H definiert man den *Winkel* $\varphi = \measuredangle xy$ zwischen x, y wie folgt:

$$\varphi = \arccos \frac{(x, y)}{\|x\| \, \|y\|},$$

was äquivalent zu

$$\cos \varphi = \frac{(x, y)}{\|x\| \, \|y\|}, \quad \varphi \in [0, \pi]. \tag{2.13}$$

Da nach der Cauchy-Schwarz-Ungleichung gilt $\frac{(x,y)}{\|x\|\|y\|} \in [-1,1]$, so ist φ wohldefiniert.

Beispiel. Sei $U \subset H$ ein abgeschlossener Unterraum und setzen wir $P = P_U$, $x' = Px$ und

$$\alpha = \measuredangle xx'.$$

Dann gilt

$$(x,x') = (x,Px) = \left(x,P^2x\right) = (Px,Px) = \|x'\|^2$$

und somit

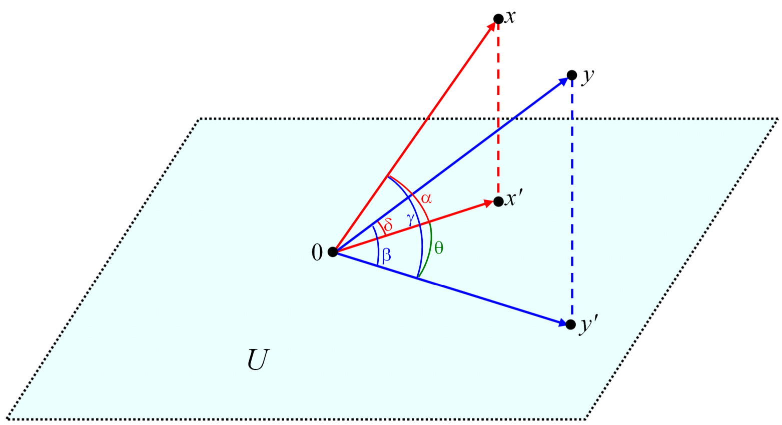
$$\cos\alpha = \frac{(x,x')}{\|x\|\,\|x'\|} = \frac{\|x'\|^2}{\|x\|\,\|x'\|} = \frac{\|x'\|}{\|x\|}$$

und

$$\|x'\| = \|x\|\cos\alpha.$$

Diese Identität stimmt mit der geometrische Definition von Kosinus überein als der Quotient von Ankathete und Hypotenuse.

Beispiel. Wählen wir noch einen Punkt $y \in H$, setzen $y' = Py$ und betrachten neben α auch die folgenden Winkel: $\beta = \measuredangle yy', \quad \gamma = \measuredangle xy', \quad \delta = \measuredangle x'y, \quad \theta = \measuredangle x'y'$.



Dann gilt

$$(Px,y) = (x',y) = \|x'\|\,\|y\|\cos\delta = \|x\|\,\|y\|\cos\alpha\cos\delta$$

und analog

$$(x,Py) = \|x\|\,\|y\|\cos\beta\cos\gamma.$$

Da $(Px,y) = (x,Py)$, so erhalten wir die Identität

$$\cos\alpha\cos\delta = \cos\beta\cos\gamma,$$

die sonst nicht so offensichtlich ist.

Analog haben wir

$$(x',y') = \|x'\|\,\|y'\|\cos\theta = \|x\|\,\|y'\|\cos\alpha\cos\theta$$

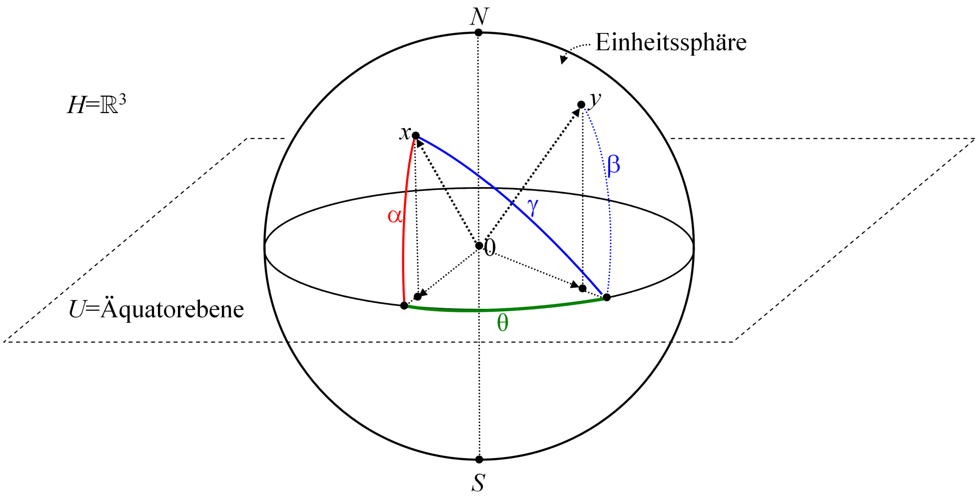
und

$$(x',y') = (Px,Py) = \left(x,P^2y\right) = (x,Py) = (x,y') = \|x\|\,\|y'\|\cos\gamma$$

woraus folgt

$$\cos\gamma = \cos\alpha\cos\theta.$$

Diese Identität ist ein Analog von dem Satz von Pythagoras in der *sphärischen* Geometrie.



In der Tat, für die kleinen Werte von α, γ, θ erhalten wir

$$1 - \frac{\gamma^2}{2} \approx \left(1 - \frac{\alpha^2}{2}\right)\left(1 - \frac{\theta^2}{2}\right) \approx 1 - \frac{\alpha^2 + \theta^2}{2}$$

und somit

$$\gamma^2 \approx \alpha^2 + \theta^2.$$

Weitere Eigenschaften von den Winkeln befinden sich in Aufgaben $38, 39$.

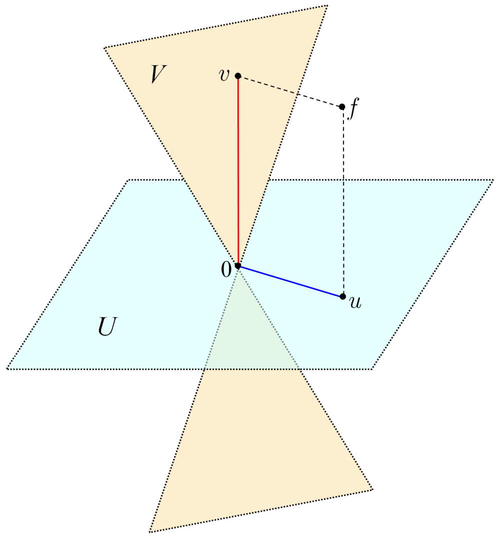
2.8 Approximation von Funktionen

Sei H ein Hilbertraum über den Körper $\mathbb{K} = \mathbb{C}$ oder $\mathbb{R}$. Gegeben seien n linear unabhängige Vektoren $u_1, ..., u_n \in H$, betrachten wir den Unterraum

$$U = \operatorname{span}\{u_1, ..., u_n\} = \{c_1 u_1 + ... + c_n u_n : c_k \in \mathbb{K}\}.$$

Der Unterraum U heißt die *lineare Hülle* von $u_1, ..., u_n$. Sei V das orthogonale Komplement von U:

$$V = U^{\perp} = \{v \in H : (v, u_1) = (v, u_2) = ... = (v, u_n) = 0\}.$$



Betrachten wir zwei folgende Optimierungsprobleme: gegeben sei ein Vektor $f \in H$, man bestimme

1. die beste Approximation u von f aus U;

2. die beste Approximation v von f aus V.

Da $V = U^{\perp}$, so ist V abgeschlossen. Da U endlichdimensional ist, so ist U auch abgeschlossen. Nach dem Satz 2.3 erhalten wir folgendes: die Lösung u des Optimierungsproblems 1 ist die orthogonale Projektion $u = P_U f$ von f auf U, und die Lösung v des Optimierungsproblems 2 ist die orthogonale Projektion $v = P_V f$ von f auf V. Nach dem Satz 2.5 haben wir $P_U + P_V = \mathrm{Id}$ und somit

$$f = u + v,$$

so reicht es u zu bestimmen. Wir haben

$$u = c_1 u_1 + ... + c_n u_n = \sum_{j=1}^{n} c_j u_j,$$

wobei die Koeffizienten $c_1, ..., c_n \in \mathbb{K}$ noch bestimmt werden müssen. Multiplizieren mit u_i für $i = 1, ..., n$ ergibt

$$(\sum_{j=1}^{n} c_j u_j, u_i) = (u, u_i) = (f, u_i) - (v, u_i) = (f, u_i)\,,$$

und wir erhalten das folgende System von n Gleichungen für Bestimmung von den Koeffizienten c_j:

$$\sum_{j=1}^{n} (u_j, u_i)\, c_j = (f, u_i) \quad \text{für alle } i = 1, ..., n,$$

d.h.

$$\begin{cases} c_1\left(u_1,u_1\right)+c_2\left(u_2,u_1\right)+...+c_n\left(u_n,u_1\right)=\left(f,u_1\right) \\ c_1\left(u_1,u_2\right)+c_2\left(u_2,u_2\right)+...+c_n\left(u_n,u_2\right)=\left(f,u_2\right) \\ ... \\ c_1\left(u_1,u_n\right)+c_2\left(u_2,u_n\right)+...+c_n\left(u_n,u_n\right)=\left(f,u_n\right) \end{cases} \tag{2.14}$$

Die Matrix $M = \left(\left(u_j,u_i\right)\right)_{i,j=1}^n$ des Systems (2.14) ist die Gramsche Matrix von den Vektoren $u_1,...,u_n$, und nach der linearen Unabhängigkeit von $u_1,...,u_n$ gilt $\det M \neq 0$. Damit erhalten wir die Lösung von (2.14)

$$\begin{pmatrix} c_1 \\ c_2 \\ ... \\ c_n \end{pmatrix} = M^{-1}\begin{pmatrix} (f,u_1) \\ (f,u_2) \\ ... \\ (f,u_n) \end{pmatrix}.$$

Die Lösung des Optimierungsproblems 1 ist $u = c_1u_1 + ... + c_nu_n$, und die Lösung des Optimierungsproblems 2 ist $v = f - u$.

Betrachten wir einige Beispiele.

Beispiel. Sei $H = L^2[a,b]$ wobei $a < b$ reell sind. Seien $n = 1$ und $u_1 = 1$, d.h. U ist der Unterraum, der aus Konstanten besteht, und $V = U^\perp$ besteht aus den Funktionen $v \in L^2[a,b]$ mit

$$(v,1) = \int_{[a,b]} v d\lambda = 0.$$

Das System (2.14) für $c = c_1$ reduziert sich zur Gleichung

$$c\,(1,1) = (f,1)\,,$$

woraus folgt

$$c = \frac{1}{b-a}\int_{[a,b]} f d\lambda,$$

und $u = c$, $v = f - c$.

Beispiel. Sei $H = L^2[0,1]$. Betrachten wir für $n = 5$ die Funktionen

$$u_1 = 1, \quad u_2(t) = t, \quad u_3(t) = t^2, \quad u_4(t) = t^3, \quad u_5(t) = t^4$$

so dass

$$U = \operatorname{span}\{u_1,...,u_5\} = \mathcal{P}_4.$$

Dann haben wir

$$(u_j,u_i) = \int_0^1 t^{i+j-2}dt = \frac{1}{i+j-1}$$

und somit

$$M = \begin{pmatrix} 1 & \frac12 & \frac13 & \frac14 & \frac15 \\ \frac12 & \frac13 & \frac14 & \frac15 & \frac16 \\ \frac13 & \frac14 & \frac15 & \frac16 & \frac17 \\ \frac14 & \frac15 & \frac16 & \frac17 & \frac18 \\ \frac15 & \frac16 & \frac17 & \frac18 & \frac19 \end{pmatrix}.$$

Für $f(t) = e^t$ erhalten wir

$$\begin{aligned}
(f,u_1) &= \int_0^1 e^t dt = e-1, \qquad (f,u_2) = \int_0^1 te^t dt = 1 \\
(f,u_3) &= \int_0^1 t^2e^t dt = e-2, \qquad (f,u_4) = \int_0^1 t^3e^t dt = 6-2e \\
(f,u_5) &= \int_0^1 t^4e^t dt = 9e-24
\end{aligned}$$

woraus folgt

$$\begin{aligned}
\begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \end{pmatrix} &= M^{-1} \begin{pmatrix} e-1 \\ 1 \\ e-2 \\ 6-2e \\ 9e-24 \end{pmatrix} \\
&= \begin{pmatrix} 25 & -300 & 1050 & -1400 & 630 \\ -300 & 4800 & -18\,900 & 26\,880 & -12\,600 \\ 1050 & -18\,900 & 79\,380 & -117\,600 & 56\,700 \\ -1400 & 26\,880 & -117\,600 & 179\,200 & -88\,200 \\ 630 & -12\,600 & 56\,700 & -88\,200 & 44\,100 \end{pmatrix} \begin{pmatrix} e-1 \\ 1 \\ e-2 \\ 6-2e \\ 9e-24 \end{pmatrix} \\
&= \begin{pmatrix} 9545e-25945 \\ 506580-186360e \\ 825930e-2245110 \\ 3455480-1271200e \\ 630630e-1714230 \end{pmatrix} \approx \begin{pmatrix} 1.00005 \\ 0.998\,45 \\ 0.510\,58 \\ 0.13966 \\ 0.06\,948 \end{pmatrix}
\end{aligned}$$

Somit ist die Funktion

$$\begin{aligned}
u(t) = {} & (9545e - 25\,945) + (506\,580 - 186\,360e)\,t + (825\,930e - 2245\,110)\,t^2 \\
& + (3455\,480 - 1271\,200e)\,t^3 + (630\,630e - 1714\,230)\,t^4 \\
& \approx 1.00005 + 0.99845\,t + 0.51058\,t^2 + 0.13966\,t^3 + 0.06948\,t^4
\end{aligned}$$

die beste Approximation von e^t aus $\mathcal{P}_4$ auf dem Intervall $[0,1]$ bezüglich der 2-Norm.

Die Taylor-Formel ergibt eine andere Approximation von e^t aus $\mathcal{P}_4$: das Taylor-Polynom von e^t des Grades 4 ist

$$v(t) = 1 + t + \frac{1}{2}t^2 + \frac{1}{6}t^3 + \frac{1}{24}t^4.$$

Vergleichen wir die Abstände zwischen e^t und u bzw v in $L^2[0,1]$:

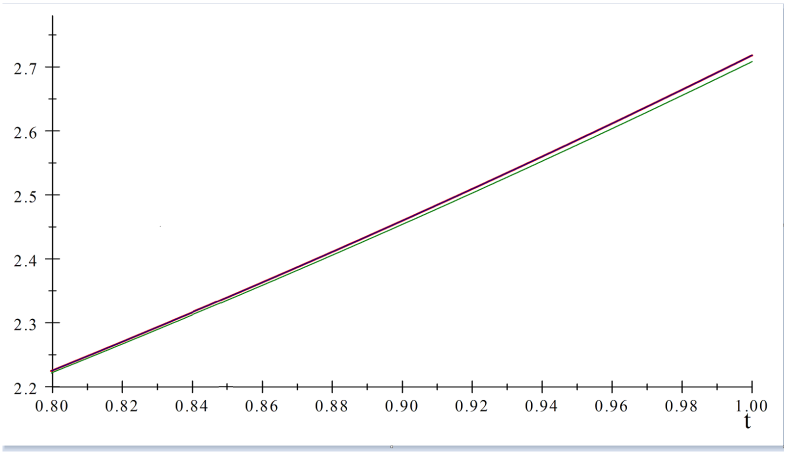
$$\left\|e^t - u\right\|_2^2 = \int_0^1 \left(e^t - u(t)\right)^2 dt \approx 2.759\,575\,738\,205\,1 \times 10^{-10}$$

und

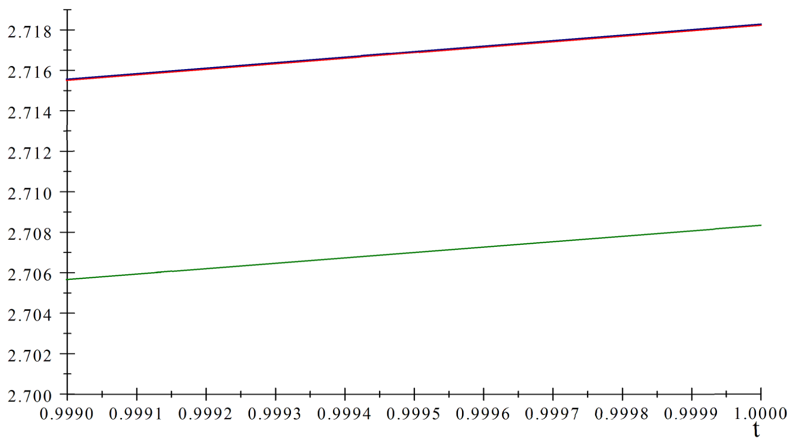
$$\left\|e^t - v\right\|_2^2 = \int_0^1 \left(e^t - v(t)\right)^2 dt \approx 8.725\,649\,582\,904\,23 \times 10^{-6}\ .$$

Die Graphen von diesen Funktionen sehen wie folgt aus (e^t blau, u rot, v grün):

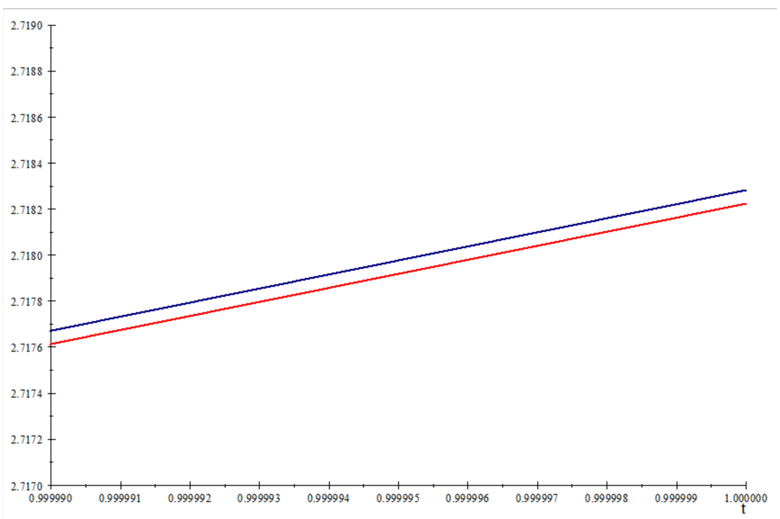
auf dem Intervall $(0.8, 1)$ (wo e^t und u praktisch identisch sind):



auf dem Intervall $(0.999, 1)$ (wo e^t und u auch sehr nahe sind):



und auf dem Intervall $(0.99999, 1)$ (wo nur e^t und u gezeigt wird):



Die Funktion u ergibt auch eine bessere punktweise Approximation von e^t auf dem Intervall $[0,1]$ als das Taylor-Polynom v. Zum Beispiel, man vergleiche

$e^{1/2} = 1.648\,721\,270...$
$u\,(1/2) = 1.648\,722\,050...$
$v\,(1/2) = 1.648\,437\,5$

06.05.22 **Vorlesung 9**

2.9 Stetige lineare Funktionale im Hilbertraum

Sei V ein Vektorraum über $\mathbb{K}$ (wobei $K = \mathbb{R}$ oder $\mathbb{C}$).

Definition. Ein lineares *Funktional* auf V ist eine lineare Abbildung $F : V \to \mathbb{K}$.

Ist V ein normierter Vektorraum, so kann man über die Stetigkeit von F sprechen.

Beispiel. Sei H ein Hilbertraum. Für jedes $a \in H$ betrachten wir das folgende lineare Funktional

$$F_a : H \to \mathbb{K}$$
$$F_a\,(x) = (x,a)\,.$$

Dieses Funktional ist stetig, da $x_n \to x$ in H impliziert $(x_n, a) \to (x, a)\,.$

Z.B. in $\mathbb{R}^n$ bekommen wir die folgenden Funktionale:

$$F_a\,(x) = (x,a) = \sum_{k=1}^{n} a_k x_k,$$

wobei x_k und a_k die Komponenten von x bzw a in der Standardbasis $\{e_k\}$ sind. In $\mathbb{R}^n$ gibt es kein anderes lineares Funktional. In der Tat, für jedes lineares Funktional $F : \mathbb{R}^n \to \mathbb{R}$ und jedes $x \in \mathbb{R}^n$ erhalten wir nach der Linearität

$$F\,(x) = F\left(\sum_{k=1}^{n} x_k e_k\right) = \sum_{k=1}^{n} x_k F\,(e_k) = \sum_{k=1}^{n} a_k x_k = F_a\,(x)\,,$$

wobei $a_k = F\,(e_k)\,.$ Eine ähnliche Aussage gilt in allen Hilberträumen.

Satz 2.7 (Rieszscher Darstellungssatz) *Zu jedem stetigen linearen Funktional F auf H gibt es genau ein $a \in H$ mit $F = F_a$, d.h.*

$$F\,(x) = (x,a) \qquad \forall x \in H.$$

In anderen Wörtern, die Menge von allen stetigen linearen Funktionalen in H stimmt mit der Menge $\{F_a\}_{a\in H}$ überein.

Beweis. Beweisen wir zunächst die Existenz von a. Dafür betrachten wir den Kern von F:

$$U = \ker F = F^{-1}\,(\{0\}) = \{x \in H : F\,(x) = 0\}\,.$$

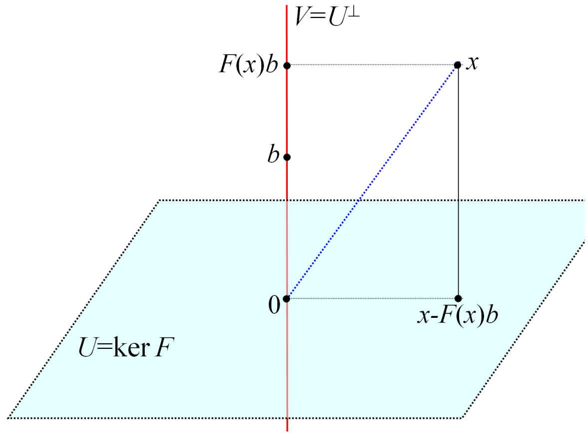
Da F linear und stetig ist, so ist U ein abgeschlossener Unterraum von H.

Setzen wir $V = U^\perp$. Gilt $V = \{0\}$, so gilt auch

$$U = \left(U^\perp\right)^\perp = V^\perp = H$$

und somit $F \equiv 0$. In diesem Fall setzen wir $a = 0$ so dass $F(x) = (x, a)$ für alle $x \in H$.

Sei $V \neq \{0\}$. Für jedes $b \in V \setminus \{0\}$ haben wir $b \notin U$ und somit $F(b) \neq 0$. Da V ein Unterraum ist, so existiert ein $b \in V \setminus \{0\}$ mit $F(b) = 1$.



Dann haben wir für jedes $x \in H$

$$F(F(x)b) = F(x)F(b) = F(x)$$

und somit

$$F(x - F(x)b) = 0,$$

woraus folgt

$$x - F(x)b \in U.$$

Da $b \perp U$, so folgt es daraus, dass

$$(x - F(x)b, b) = 0$$

und somit

$$(x, b) - F(x)(b, b) = 0$$

und

$$F(x) = \frac{(x, b)}{(b, b)} = (x, a)$$

wobei $a = \frac{b}{(b,b)}$.

Beweisen wir jetzt die Eindeutigkeit von a. Gilt $F(x) = (x, a_1) = (x, a_2)$ für alle $x \in H$, so erhalten wir

$$(x, a_1 - a_2) = 0.$$

Einsetzen $x = a_1 - a_2$ ergibt $\|a_1 - a_2\|^2 = 0$ und somit $a_1 = a_2$. ■

Bemerkung. Es folgt aus diesem Beweis, dass im Fall $V \neq \{0\}$

$$x - F(x)b \perp V$$

woraus folgt $P_V x = F(x)\, b$ und somit $V = \operatorname{im} P_V = \operatorname{span}\{b\}$ und $\dim V = 1$.

Bemerkung. Sei V ein normierter Vektorraum. Die Menge von allen stetigen linearen Funktionalen auf V ist ein Vektorraum, der heißt der *Dualraum* von V und wird mit V' bezeichnet. Satz 2.7 etabliert eine bijektive (und offensichtlich lineare) Abbildung

$$H' \ni F \mapsto a \in H$$

zwischen H' und H. Somit sind H' und H linear isomorph. Später in dieser Vorlesung studieren wir die Dualräume von anderen Banachräumen.

2.10 Orthogonale Folgen

Definition. Eine Folge $\{v_n\}$ von Vektoren im Skalarproduktraum heißt *orthogonal*, wenn $v_n \perp v_m$ für alle $n \neq m$ und $v_n \neq 0$ für alle n. Die Folge $\{v_n\}$ heißt *orthonormal*, wenn sie orthogonal ist und $\|v_n\| = 1$ für alle n.

Jede orthogonale Folge ist linear unabhängig (Aufgabe 33). Ist $\{v_n\}$ eine orthogonale Folge, so ist die Folge $u_n = \frac{v_n}{\|v_n\|}$ orthonormal.

Beispiel. 1. In $\mathbb{R}^n$ (und $\mathbb{C}^n$) ist die Folge $e_1, ..., e_n$ orthonormal, wobei

$$e_k = (0, ..., \overset{k}{\widehat{1}}, ...0).$$

2. In l^2 ist die Folge $\{e_k\}_{k=1}^{\infty}$ orthonormal, wobei

$$e_k = (0, ..., \overset{k}{\widehat{1}}, ..., 0, ...).$$

3. In $L^2[0, 2\pi]$ ist die folgende Folge orthogonal:

$$1, \cos x, \sin x, \cos 2x, \sin 2x, ..., \cos nx, \sin nx, ...$$

(Aufgabe 41).

4. In $L^2_{\mathbb{C}}[0, 2\pi]$ ist die folgende Folge orthogonal:

$$\left\{e^{inx}\right\}_{n\in\mathbb{Z}} = ..., e^{-3ix}, e^{-2ix}, e^{-ix}, 1, e^{ix}, e^{2ix}, e^{3ix}, ...$$

(Aufgabe 41).

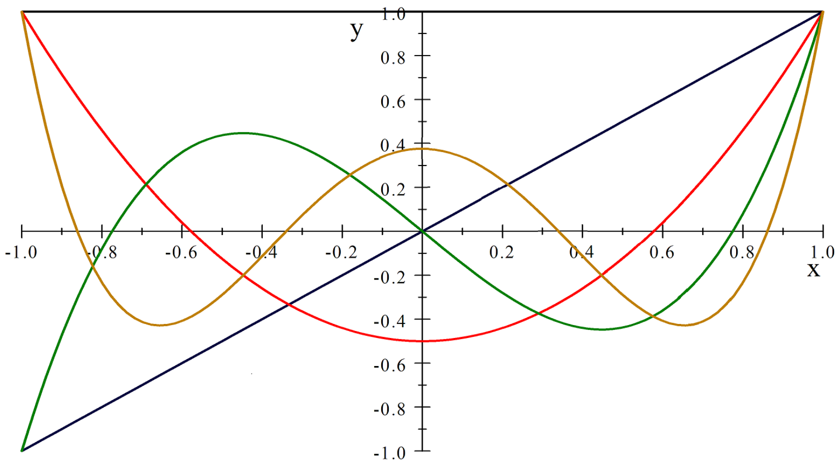
5. In $L^2[-1, 1]$ ist die Folge $\{L_n(x)\}_{n=0}^{\infty}$ von *Legendre-Polynomen* eine orthogonale Folge, wobei

$$L_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n} \left(x^2 - 1\right)^n$$

ein Polynom des Grades n ist (siehe Aufgabe 45). Z.B. wir haben

$$\begin{aligned}
L_0(x) &= 1, \\
L_1(x) &= x, \\
L_2(x) &= \frac{1}{2}\left(3x^2 - 1\right), \\
L_3(x) &= \frac{1}{2}\left(5x^3 - 3x\right), \\
L_4(x) &= \frac{1}{8}\left(35x^4 - 30x^2 + 3\right),
\end{aligned}$$

usw. Diese Polynome sind auf dem nächsten Bild gezeichnet:



6. Sei D die Einheitskreisscheibe in $\mathbb{C}$ mit dem Radius 1, d.h.

$$D = \{z \in \mathbb{C} : |z| \leq 1\}.$$

Dann ist die folgende Funktionenfolge orthogonal in $L^2_{\mathbb{C}}(D, \lambda_2)$:

$$1, \, z, \, z^2, ..., \, z^n, ...$$

(siehe Aufgabe 46).

Satz 2.8 *Sei* $\{e_n\}_{n=1}^N$ *eine orthonormale Folge im Hilbertraum* H*, wobei* N *entweder natürliche Zahl oder* ∞ *ist.*

(a) (Besselsche Ungleichung) *Für jedes* $x \in H$ *gilt*

$$\|x\|^2 \geq \sum_{n=1}^{N} |(x, e_n)|^2 \tag{2.15}$$

(b) (Parsevalsche Gleichung) *Für*

$$x = \sum_{n=1}^{N} \alpha_n e_n$$

mit $\alpha_n \in \mathbb{K}$ *gelten die Identitäten*

$$\alpha_n = (x, e_n) \quad \text{für alle } n \tag{2.16}$$

und

$$\|x\|^2 = \sum_{n=1}^{N} |\alpha_n|^2 = \sum_{n=1}^{N} |(x, e_n)|^2 . \tag{2.17}$$

Beweis. (a) Es reicht zu beweisen, dass für jede natürliche Zahl $m \leq N$ gilt

$$\|x\|^2 \geq \sum_{n=1}^{m} |(x, e_n)|^2 .$$

Dann für endliches N erhalten wir (2.15) für $m = N$, und für $N = \infty$ lassen wir $m \to \infty$.

Betrachten wir den Unterraum

$$U = \operatorname{span}\{e_1, ..., e_m\}$$

der endlichdimensional und somit abgeschlossen ist. Dann ist die Projektion $P_U x$ eine lineare Kombination von $e_1, ..., e_m$, sei

$$P_U x = \alpha_1 e_1 + ... + \alpha_m e_m. \tag{2.18}$$

Wir haben nach dem Satz der Pythagoras

$$\|x\|^2 \geq \|P_U x\|^2 = \sum_{n=1}^{m} \|\alpha_n e_n\|^2 = \sum_{n=1}^{m} |\alpha_n|^2 . \tag{2.19}$$

Bestimmen wir jetzt α_n für $n = 1, ..., m$. Es folgt aus (2.18)

$$(x, e_n) = (x, P_U e_n) = (P_U x, e_n) = \alpha_n.$$

Einsetzen in (2.19) ergibt (2.15).

(b) Für jede natürliche Zahl $m \leq N$ setzen wir

$$x_m = \sum_{n=1}^{m} \alpha_n e_n.$$

Wie oberhalb gilt $(x_m, e_n) = \alpha_n$ für $m \geq n$ und

$$\|x_m\|^2 = \sum_{n=1}^{m} |\alpha_n|^2 .$$

Ist N endlich, so setzen wir $m = N$ und erhalten (2.16) und (2.17). Sei $N = \infty$. Lassen wir $m \to \infty$ und erhalten, dass $x_m \to x$. Für jedes n erhalten wir

$$(x_m, e_n) \to (x, e_n) \text{ für } m \to \infty,$$

woraus (2.16) folgt. Auch gilt $\|x_m\|^2 \to \|x\|^2$ für $m \to \infty$, woraus (2.17) folgt. ■

2.11 Orthonormalbasis

Sei $\{e_n\}_{n=1}^N$ eine Folge im Hilbertraum, wobei N entweder endlich oder unendlich ist (d.h. $N \in \mathbb{N} \cup \{\infty\}$).

Definition. Die Folge $\{e_n\}_{n=1}^N$ heißt eine *Basis* von H wenn es für jedes $x \in H$ genau eine Folge $\{\alpha_n\}_{n=1}^N$ von Elementen von $\mathbb{K}$ mit

$$x = \sum_{n=1}^{N} \alpha_n e_n$$

gibt. Die Werte α_n heißen die Koordinaten von x in der Basis $\{e_n\}$.

Definition. Die Folge $\{e_n\}_{n=1}^N$ heißt ein *Erzeugendensystem* von H wenn $\operatorname{span}\{e_n\}$ dicht in H liegt.

Erinnern wir uns, dass span $\{e_n\}$ die lineare Hülle von $\{e_n\}$ ist, d.h. die Menge von allen *endlichen* linearen Kombinationen von $\{e_n\}$ mit den Koeffizienten aus $\mathbb{K}$. Die lineare Hülle ist immer ein Unterraum. Eine Teilmenge $M \subset H$ liegt dicht im H wenn der Abschluss $\overline{M}$ von M mit H übereinstimmt (d.h. für jedes Element $x \in H$ existiert eine Folge $\{y_n\} \subset M$ mit $y_n \to x$). Der Abschluss von span $\{e_n\}$ heißt die *abgeschlossene lineare Hülle* von $\{e_n\}$ und wird mit $\overline{\operatorname{span}}\,\{e_n\}$ bezeichnet (offensichtlich ist $\overline{\operatorname{span}}\,\{e_n\}$ ein abgeschlossener Unterraum). Somit ist die Folge $\{e_n\}$ ein Erzeugendensystem genau dann, wenn $\overline{\operatorname{span}}\,\{e_n\} = H$.

Definition. Die Folge $\{e_n\}_{n=1}^N$ heißt *maximal* in H wenn die folgende Implikation für alle $x \in H$ gilt:

$$x \perp e_n \text{ für alle } n \ \Rightarrow x = 0.$$

Satz 2.9 *Sei $\{e_n\}_{n=1}^\infty$ eine orthonormale Folge. Dann sind die folgenden Eigenschaften äquivalent:*

1. *$\{e_n\}$ ist eine Basis;*
2. *$\{e_n\}$ ist ein Erzeugendensystem;*
3. *$\{e_n\}$ ist maximal;*
4. *für jedes $x \in H$ gilt*

$$x = \sum_{n=1}^\infty (x, e_n)\, e_n \ ;$$

5. (Parsevalsche Gleichung) *für jedes $x \in H$ gilt*

$$\|x\|^2 = \sum_{n=1}^\infty |(x, e_n)|^2 \,.$$

Zunächst beweisen wir ein Lemma.

Lemma 2.10 *Sei $\{e_n\}_{n=1}^\infty$ eine orthonormale Folge und $\{\alpha_n\}_{n=1}^\infty$ eine Folge von Skalaren mit*

$$\sum_{n=1}^\infty |\alpha_n|^2 < \infty.$$

Dann ist die Reihe

$$\sum_{n=1}^\infty \alpha_n e_n$$

konvergent.

Beweis. Betrachten wir die partiellen Summen

$$x_n = \sum_{k=1}^n \alpha_k e_k.$$

Für $n > m$ gilt

$$\|x_n - x_m\|^2 = \left\| \sum_{k=m+1}^n \alpha_k e_k \right\|^2 = \sum_{k=m+1}^n |\alpha_k|^2 \to 0$$

für $n, m \to \infty$. Folglich ist $\{x_n\}$ eine Cauchy-Folge und somit konvergiert, so dass auch die Reihe $\sum_{k=1}^{\infty} \alpha_k e_k$ konvergiert. $\blacksquare$

Beweis von Satz 2.9. Wir beweisen die folgenden Implikationen:

$$1 \Rightarrow 2 \Rightarrow 3 \Rightarrow 4 \Rightarrow 1 \text{ und } 1 \Rightarrow 5 \Rightarrow 3.$$

$1 \Rightarrow 2$. Die endlichen Summen $\sum_{n=1}^{N} \alpha_n e_n$ liegen in $\operatorname{span}\{e_n\}$, woraus folgt, dass

$$\sum_{n=1}^{\infty} \alpha_n e_n \in \overline{\operatorname{span}}\{e_n\}.$$

Da jedes x sich in der Form $\sum_{n=1}^{\infty} \alpha_n e_n$ darstellen lässt, so erhalten wir, dass

$$\overline{\operatorname{span}}\{e_n\} = H$$

und somit ist $\{e_n\}$ ein Erzeugendensystem.

$2 \Rightarrow 3$. Wir müssen beweisen, dass

$$x \perp e_n \text{ für alle } n \;\Rightarrow\; x = 0.$$

Die Bedingung $x \perp e_n$ impliziert, dass x orthogonal zu $\operatorname{span}\{e_n\}$ ist und dann auch orthogonal zu $\overline{\operatorname{span}}\{e_n\}$. Da $\overline{\operatorname{span}}\{e_n\} = H$, so erhalten wir $x \perp H$, insbesondere $x \perp x$, woraus folgt $x = 0$.

$3 \Rightarrow 4$. We müssen beweisen, dass für jedes $x \in H$ gilt

$$x = \sum_{n=1}^{\infty} \alpha_n e_n$$

wobei $\alpha_n = (x, e_n)$. Nach der Besselschen Ungleichung gilt

$$\|x\|^2 \geq \sum_{n=1}^{\infty} |\alpha_n|^2$$

so dass $\sum_{n=1}^{\infty} |\alpha_n|^2 < \infty$. Nach dem Lemma 2.10 ist die Reihe $\sum_{n=1}^{\infty} \alpha_n e_n$ konvergent, sei

$$z = \sum_{n=1}^{\infty} \alpha_n e_n.$$

Es bleibt zu beweisen, dass $x = z$. Nach dem Satz 2.8 gilt $\alpha_n = (z, e_n)$ und somit

$$(z, e_n) = (x, e_n).$$

Daraus folgt, dass

$$z - x \perp e_n \text{ für alle } n.$$

Da die Folge $\{e_n\}$ maximal ist, so erhalten wir $z - x = 0$, was zu beweisen war.

11.05.22 Vorlesung 10

$4 \Rightarrow 1$. Da für jedes $x \in H$ die Identität

$$x = \sum_{n=1}^{\infty} (x, e_n)\, e_n$$

gilt, so ist jedes x in der Form $x = \sum_{n=1}^{\infty} \alpha_n e_n$ darstellbar (und zwar mit $\alpha_n = (x, e_n)$). Darüber hinaus sind die Koeffizienten α_n nach dem Satz 2.8 eindeutig bestimmt. Somit ist $\{e_n\}$ eine Basis.

1 $\Rightarrow$ 5. Da $\{e_n\}$ eine Basis ist, erhalten wir für jedes x eine Entwicklung

$$x = \sum_{n=1}^{\infty} \alpha_n e_n.$$

Es folgt nach Satz 2.8, dass $\alpha_n = (x, e_n)$ und

$$\|x\|^2 = \sum_{n=1}^{\infty} |(x, e_n)|^2 . \tag{2.20}$$

5 $\Rightarrow$ 3. Gilt $x \perp e_n$ für alle n, so erhalten wir $(x, e_n) = 0$ und somit nach (2.20) $\|x\| = 0$ und $x = 0$. Somit ist die Folge $\{e_n\}$ maximal. ∎

Korollar 2.11 *Sei $\{e_n\}_{n=1}^{\infty}$ eine Orthonormalbasis im H. Dann gilt für alle $x, y \in H$*

$$(x, y) = \sum_{n=1}^{\infty} \alpha_n \overline{\beta_n} \tag{2.21}$$

wobei $\alpha_n = (x, e_n)$ und $\beta_n = (y, e_n)$.

Für Beweis brauchen wir die folgende Behauptung.

Behauptung. *In jedem Skalarproduktraum*

$$x_n \to x \quad \text{und} \quad y_n \to y \ \Rightarrow \ (x_n, y_n) \to (x, y) .$$

Beweis. Wir haben nach der Cauchy-Schwarz-Ungleichung:

$$\begin{aligned} |(x_n, y_n) - (x, y)| &= |(x_n - x, y_n) + (x, y_n - y)| \\ &\leq |(x_n - x, y_n)| + |(x, y_n - y)| \\ &\leq \|x_n - x\| \, \|y_n\| + \|x\| \, \|y_n - y\| . \end{aligned} \tag{2.22}$$

Die Folge $\{\|y_n\|\}$ ist beschränkt wegen $\|y_n\| \to \|y\|$, während

$$\|x_n - x\| \to 0 \quad \text{und} \quad \|y_n - y\| \to 0 \quad \text{für } n \to \infty.$$

Somit erhalten wir, dass die rechte Seite von (2.22) gegen 0 für $n \to \infty$ konvergiert, woraus $(x_n, y_n) \to (x, y)$ folgt. ∎

Beweis von Korollar 2.11. Setzen wir

$$x_n = \sum_{k=1}^{n} \alpha_k e_k \quad \text{und} \quad y_n = \sum_{l=1}^{n} \beta_l e_l,$$

so dass $x_n \to x$ und $y_n \to y$ für $n \to \infty$. Dann gilt

$$\begin{aligned} (x_n, y_n) &= \left(\sum_{k=1}^{n} \alpha_k e_k, \sum_{l=1}^{n} \beta_l e_l \right) = \sum_{k=1}^{n} \sum_{l=1}^{n} (\alpha_k e_k, \beta_l e_l) \\ &= \sum_{k=1}^{n} \sum_{l=1}^{n} \alpha_k \overline{\beta_l} \, (e_k, e_l) = \sum_{k=1}^{n} \alpha_k \overline{\beta_k}, \end{aligned}$$

da $(e_k, e_l) = 1$ für $k = l$ ist und sonst $(e_k, e_l) = 0$. Für $n \to \infty$ erhalten wir

$$(x_n, y_n) \to (x, y)$$

und somit

$$\sum_{k=1}^{n} \alpha_k \overline{\beta_k} \to \sum_{k=1}^{\infty} \alpha_k \overline{\beta_k}$$

woraus (2.21) folgt. ∎

2.12 Existenz orthogonaler Basis im separablen Hilbertraum

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

Da jeder normierter Vektorraum auch ein metrischer Raum ist, so kann man diesen Begriff für normierten Räumen anwenden.

Beispiel. $\mathbb{R}$ ist separabel da die abzählbare Menge $\mathbb{Q}$ in $\mathbb{R}$ dicht liegt. Analog auch $\mathbb{R}^n$ separabel ist, da die abzählbare Mengen $\mathbb{Q}^n$ von Vektoren $(x_1, ..., x_n)$ mit rationalen Komponenten in $\mathbb{R}^n$ dicht liegt.

Beispiel. Der Raum l^p ist separabel für alle $p \in [1, \infty)$, da die Menge M von Elementen von l^p der Form $(x_1, ..., x_n, 0, 0, ...)$ mit beliebigen $n \in \mathbb{N}$ und $x_1, ..., x_n \in \mathbb{Q}$ abzählbar ist und in l^p dicht liegt. Der Raum l^∞ ist im Gegenteil *nicht* separabel (siehe Aufgabe 47).

Wir werden später beweisen, dass die Räume $C[a, b]$ und $L^p[a, b]$ mit $p \in [1, \infty)$ auch separabel sind. Insbesondere ist $L^2[a, b]$ ein separabler Hilbertraum.

Satz 2.12 *Sei H ein Hilbertraum (über $\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$) mit $\dim H = \infty$. Dann sind die folgenden Eigenschaften äquivalent:*

1. *Es gibt eine Orthonormalbasis $\{e_n\}_{n=1}^{\infty}$.*
2. *Es gibt ein Erzeugendensystem $\{g_n\}_{n=1}^{\infty}$.*
3. *H ist separabel.*

Beweis. $1 \Rightarrow 2$. Eine Basis ist auch ein Erzeugendensystem nach Satz 2.9.

$2 \Rightarrow 3$. Betrachten wir die rationale lineare Hülle von $\{g_n\}$:

$$\operatorname{span}_{\mathbb{Q}} \{g_n\} = \{c_1 g_1 + ... + c_n g_n : n \in \mathbb{N},\ \operatorname{Re} c_k, \operatorname{Im} c_k \in \mathbb{Q}\}.$$

Die Menge $\operatorname{span}_{\mathbb{Q}} \{g_n\}$ ist offensichtlich abzählbar, und

$$\overline{\operatorname{span}_{\mathbb{Q}}} \{g_n\} = \overline{\operatorname{span}_{\mathbb{K}}} \{g_n\} = H.$$

Damit ist H separabel.

$3 \Rightarrow 2$. Trivial: ist $M = \{g_n\}_{n=1}^{\infty}$ eine abzählbare dichte Menge, so ist die Folge $\{g_n\}$ Erzeugendensystem, da

$$\overline{\operatorname{span}} \{g_n\} \supset \overline{M} = H.$$

$2 \Rightarrow 1$. Gegeben sei ein Erzeugendensystem $\{g_n\}_{n=1}^{\infty}$, so bilden wir eine Orthogonalbasis $\{e_n\}$ mit Hilfe von dem *Gram-Schmidtschen Orthogonalisierungsverfahren*. Danach kann man die Vektoren e_n normieren, d.h. e_n durch $\frac{e_n}{\|e_n\|}$ ersetzen und somit eine Orthonormalbasis erhalten. Wir definieren die Folge $\{e_n\}_{n=1}^{\infty}$ per Induktion nach n so dass für jedes n gilt:

- die Folge $\{e_1, ..., e_n\}$ ist orthogonal;
- $g_1, ..., g_n \in \operatorname{span}\{e_1, ..., e_n\}$.

Ohne Beschränkung der Allgemeinheit nehmen wir an, dass $g_n \neq 0$ für alle n.

Für $n = 1$ setzen wir $e_1 = g_1$. Angenommen, dass die Vektoren $e_1, ..., e_n$ schon definiert sind, bemerken wir, dass der Unterraum $E_n := \operatorname{span}\{e_1, ..., e_n\}$ endlichdimensional und abgeschlossen ist. Dann existiert ein $g_m \notin E_n$ da sonst erhalten wir

$$H = \overline{\operatorname{span}}\{g_m\} \subset E_n$$

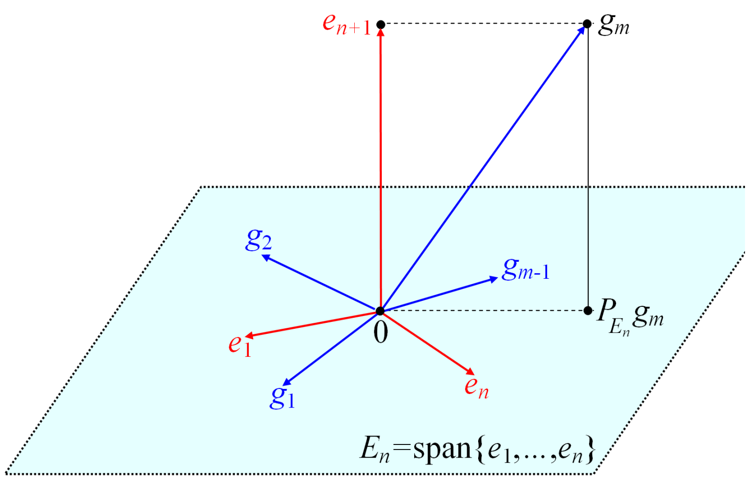
was im Widerspruch zu $\dim H = \infty$ steht. Wählen wir ein $g_m \notin E_n$ mit dem minimalen möglichen Wert von m, d.h.

$$g_1, ..., g_{m-1} \in E_n \quad \text{und} \quad g_m \notin E_n.$$

Da E_n alle $g_1, ..., g_n$ enthält, so haben wir $m \geq n + 1$.

Setzen wir

$$e_{n+1} = g_m - P_{E_n} g_m. \tag{2.23}$$



Dann ist e_{n+1} orthogonal zu E_n und somit zu allen $e_1, ..., e_n$, so dass $\{e_1, ..., e_{n+1}\}$ eine orthogonale Folge ist. Da

$$g_m = P_{E_n} g_m + e_{n+1}$$

und $P_{E_n} g_m \in E_n$, so erhalten wir, dass

$$g_m \in \operatorname{span}(E_n, e_{n+1}) = E_{n+1}.$$

Nach der Minimalität von m gilt $g_1, ..., g_{m-1} \in E_n$, woraus folgt $g_1, ..., g_m \in E_{n+1}$ und somit auch $g_1, ..., g_{n+1} \in E_{n+1}$ (da $m \geq n+1$).

Nach dem Induktionsprinzip erhalten wir eine orthogonale Folge $\{e_n\}_{n=1}^{\infty}$ mit der Eigenschaft, dass $\operatorname{span}\{e_n\}$ alle g_k enthält. Daraus folgt, dass

$$\overline{\operatorname{span}}\{e_n\} \supset \overline{\operatorname{span}}\{g_n\} = H.$$

Folglich ist $\{e_n\}$ ein Erzeugendensystem und somit auch eine Basis nach dem Satz 2.9. ■

Definition. Zwei Skalarprodukträume V, U heißen *isomorph* (oder isometrisch) wenn es eine bijektive lineare Abbildung $A : U \to V$ gibt, die das Skalarprodukt bewahrt, d.h.

$$(x, y)_V = (Ax, Ay)_U \text{. für alle } x, y \in V.$$

Korollar 2.13 *Jeder separable Hilbertraum H über $\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$ ist isomorph (als Skalarproduktraum) zu einem der folgenden Räumen: $\mathbb{R}^n$, l^2, $\mathbb{C}^n$, $l^2_{\mathbb{C}}$.*

Beweis. Betrachten wir den Fall von $\mathbb{K} = \mathbb{C}$. Gilt $n := \dim H < \infty$, so ist H isomorph zu $\mathbb{C}^n$. Sei $\dim H = \infty$. Nach dem Satz 2.12 existiert in H eine Orthonormalbasis $\{e_n\}_{n=1}^{\infty}$. Für jedes $x \in H$ existiert eine eindeutige Entwicklung

$$x = \sum_{k=1}^{\infty} x_k e_k$$

mit $x_k \in \mathbb{C}$, und

$$\sum_{k=1}^{\infty} |x_k|^2 = \|x\|^2 < \infty.$$

Daraus folgt, dass die Folge $(x_1, x_2, ...)$ ein Element von $l^2_{\mathbb{C}}$ ist. Somit erhalten wir eine Abbildung

$$A : H \to l^2_{\mathbb{C}}$$
$$Ax = (x_1, x_2, ...) \, .$$

Die Abbildung A ist linear und bijektiv. In der Tat, es gilt $\ker A = \{0\}$ und $\operatorname{im} A = l^2_{\mathbb{C}}$, da für jede Folge $(x_1, x_2, ...) \in l^2_{\mathbb{C}}$ die Reihe $\sum_{k=1}^{\infty} x_k e_k$ konvergiert (siehe Lemma 2.10).

Es bleibt noch zu zeigen, dass die Abbildung A das Skalarprodukt bewahrt, d.h.

$$(x, y)_H = (Ax, Ay)_{l^2_{\mathbb{C}}} \quad \forall x, y \in H. \tag{2.24}$$

Nach dem Korollar 2.11, $x = \sum_{k=1}^{\infty} x_k e_k$ und $y = \sum_{k=1}^{\infty} y_k e_k$ ergeben

$$(x, y)_H = \sum_{k=1}^{\infty} x_k \overline{y_k},$$

und nach der Definition von Skalarprodukt in $l^2_{\mathbb{C}}$ gilt

$$(Ax, Ay)_{l^2_{\mathbb{C}}} = \sum_{k=1}^{\infty} x_k \overline{y_k},$$

woraus (2.24) folgt. ■

2.13 Separabilität von Lebesgue-Räumen

Definition. Seien A, B zwei Teilmengen von einem metrischen Raum X. We sagen dass

$$A \textit{ approximiert } B$$

und schreiben

$$A \precapprox B,$$

wenn $B \subset \overline{A}$ wobei $\overline{A}$ der Abschluss von A ist.; d.h.

$$A \precapprox B \quad \Leftrightarrow \quad B \subset \overline{A}.$$

Die Bedingung $A \precapprox B$ lässt sich auch wie folgt verstehen: für jedes $b \in B$ gibt es eine Folge $\{a_n\}_{n \in \mathbb{N}}$ von Elementen aus A mit $a_n \to b$, d.h. jedes Element von B mit Elementen von A approximiert werden kann. Z.B.

$$\mathbb{Q} \precapprox \mathbb{R}.$$

Die Menge A muss nicht unbedingt eine Teilmenge von B sein obwohl in Anwendungen dies häufig der Fall ist. Ist A eine Teilmenge von B so bedeutet $A \precapprox B$ dass A dicht in B liegt.

Für den Begriff "A approximiert B" gibt es keine gebräuchliche Bezeichnung; das Zeichen $\precapprox$ wird nur in diesem Kurs gebraucht.$\lesssim$

Lemma 2.14 *Die folgenden Aussagen gelten für Teilmengen A, B, C von einem normierten Vektorraum V.*

(a) $A \precapprox B \precapprox C$ *impliziert* $A \precapprox C$.

(b) $A \precapprox B$ *impliziert* $\operatorname{span} A \precapprox \operatorname{span} B$.

Bemerken wir, dass die Relation $A \precapprox B$ nicht symmetrisch ist.

Beweis. (a) Wir haben $C \subset \overline{B}$ und $B \subset \overline{A}$, woraus folgt $\overline{B} \subset \overline{A}$ und somit $C \subset \overline{A}$, was äquivalent zu $A \precapprox C$ ist.

(b) Da jedes Element von B mit Elementen von A approximiert wird, so wird jede endliche lineare Kombination von Elementen von B mit endlichen linearen Kombinationen von Elementen von A approximiert. ■

Definieren wir den Unterraum $C_0[a, b]$ von $C[a, b]$ wie folgt:

$$C_0[a, b] = \{f \in C[a, b] : f(a) = f(b) = 0\}.$$

Satz 2.15 *Für alle reelle $a < b$ gilt folgendes.*

(a) $L^1[a, b]$ *ist separabel.*

(b) $C_0[a, b]$ *liegt dicht in* $L^1[a, b]$.

Beweis. Die Aussagen (a) und (b) lassen sich wie folgt umformulieren:

$$M \precapprox L^1[a,b], \tag{2.25}$$

wobei im Fall (a) M eine abzählbare Teilmenge $M \subset L^1[a,b]$ ist und im Fall (b) $M = C_0[a,b]$.

Wir beweisen (2.25) mit Hilfe von den folgenden Schritten:

$$M \precapprox \operatorname{span} T_0 \precapprox \operatorname{span} T_1 \precapprox \operatorname{span} T_2 \precapprox \operatorname{span} T_3 \precapprox L^1[a,b], \tag{2.26}$$

wobei

$$\begin{aligned} T_0 &= \{\mathbf{1}_J : J \text{ ist Teilintervall von } [a,b]\} \\ T_1 &= \{\mathbf{1}_U : U \text{ ist abzählbare Vereinigung von Teilintervallen von } [a,b]\} \\ T_2 &= \{\mathbf{1}_E : E \text{ ist messbare Teilmenge von } [a,b]\} \end{aligned}$$

und

$$T_3 = \left\{f \in L^1[a,b] : f \geq 0\right\}.$$

Offensichtlich gelten die Inklusionen:

$$T_0 \subset T_1 \subset T_2 \subset T_3 \subset L^1[a,b].$$

Da jede integrierbare Funktion eine Differenz zweier nichtnegativen integrierbaren Funktion ist, so erhalten wir

$$\operatorname{span} T_3 = L^1[a,b]$$

so dass die letzte Relation $\precapprox$ in (2.26) gilt.

Beweisen wir, dass $\operatorname{span} T_2 \precapprox T_3$ (und somit $\operatorname{span} T_2 \precapprox \operatorname{span} T_3$). Nach der Definition von Lebesgue-Integral gilt für jede Funktion $f \in T_3$

$$\int_{[a,b]} f d\lambda = \lim_{n\to\infty} \sum_{k=0}^{\infty} \frac{k}{n} \lambda\left\{\frac{k}{n} \leq f(x) < \frac{k+1}{n}\right\} = \lim_{n\to\infty} \int_{[a,b]} f_n d\lambda, \tag{2.27}$$

wobei

$$f_n = \sum_{k=0}^{\infty} \frac{k}{n} \mathbf{1}_{\left\{\frac{k}{n} \leq f < \frac{k+1}{n}\right\}} \in \overline{\operatorname{span}} T_2,$$

da die Mengen $\left\{\frac{k}{n} \leq f < \frac{k+1}{n}\right\}$ messbar sind. Da $f \geq f_n$, so folgt es aus (2.27), dass

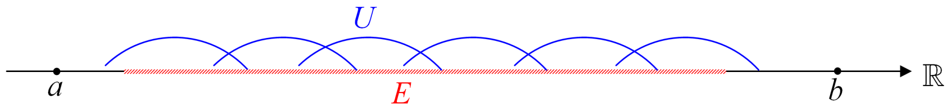
$$\|f - f_n\|_1 = \int_{[a,b]} (f - f_n)\, d\lambda \to 0 \text{ für } n \to \infty,$$

und somit $f \in \overline{\operatorname{span}} T_2$ und $\operatorname{span} T_2 \precapprox T_3$.

13.05.22 Vorlesung 11

Beweisen wir jetzt, dass $T_1 \precapprox T_2$ und somit $\operatorname{span} T_1 \precapprox \operatorname{span} T_2$. Dafür verwenden wir die folgende Eigenschaft von Lebesgue-Maß, die aus der Konstruktion von messbaren Mengen folgt: für jede messbare Menge $E \subset [a,b]$ und jedes $\varepsilon > 0$ existiert eine Teilmenge U von $[a,b]$ so dass die folgenden Eigenschaften erfüllt sind:

- U ist eine abzählbare Vereinigung von Intervallen;
- $U \supset E$;
- $\lambda(U \setminus E) < \varepsilon$.



Für jedes $n \in \mathbb{N}$ setzen wir $\varepsilon = \frac{1}{n}$ und erhalten die entsprechende Menge $U = U_n$ so dass $\lambda(U_n \setminus E) < \frac{1}{n}$ und somit $\lambda(U_n \setminus E) \to 0$ für $n \to \infty$. Daraus folgt, dass

$$\|\mathbf{1}_{U_n} - \mathbf{1}_E\|_1 = \int_{[a,b]} \mathbf{1}_{U_n \setminus E} d\lambda = \lambda(U_n \setminus E) \to 0.$$

Da $\mathbf{1}_E$ ein beliebiges Element von T_2 ist und $\mathbf{1}_{U_n} \in T_1$, so erhalten wir $T_1 \precapprox T_2$.

Beweisen wir, dass $\operatorname{span} T_0 \precapprox T_1$ (und somit $\operatorname{span} T_0 \precapprox \operatorname{span} T_1$). Jede Menge $U \in T_1$ ist eine abzählbare Vereinigung von Intervallen, und sie lässt sich auch als eine *disjunkte* Vereinigung von Intervallen darstellen:

$$U = \bigsqcup_{k=1}^{\infty} J_k$$

(wobei J_k Teilintervalle von $[a,b]$ sind). Daraus folgt, dass

$$\mathbf{1}_U = \sum_{k=1}^{\infty} \mathbf{1}_{J_k}$$

und

$$\|\mathbf{1}_U - \sum_{k=1}^{n} \mathbf{1}_{J_k}\|_1 = \int_{[a,b]} \left(\sum_{k=n+1}^{\infty} \mathbf{1}_{J_k}\right) d\lambda = \lambda\left(\bigsqcup_{k=n+1}^{\infty} J_k\right) = \sum_{k=n+1}^{\infty} \lambda(J_k) \to 0 \text{ für } n \to \infty.$$

Da $\mathbf{1}_U$ ein beliebiges Element von T_1 ist und $\sum_{k=1}^{n} \mathbf{1}_{J_k} \in \operatorname{span} T_0$, so beschließen wir, dass $\operatorname{span} T_0 \precapprox T_1$.

Es bleibt zu beweisen, dass $M \precapprox \operatorname{span} T_0$. Im Fall (a) setzen wir

$$M = \operatorname{span}_{\mathbb{Q}} \{\mathbf{1}_J : J \text{ ist ein Teilintervall von } [a,b] \text{ mit den Grenzpunkten aus } \mathbb{Q}\},$$

d.h. M ist die Menge von allen endlichen linearen Kombinationen mit rationalen Koeffizienten von Indikatorfunktionen von Intervallen mit rationalen Grenzpunkten. Es ist klar, dass M abzählbar ist und

$$M \precapprox \operatorname{span} T_0.$$

Es folgt aus (2.26), dass auch

$$M \precapprox L^1[a,b],$$

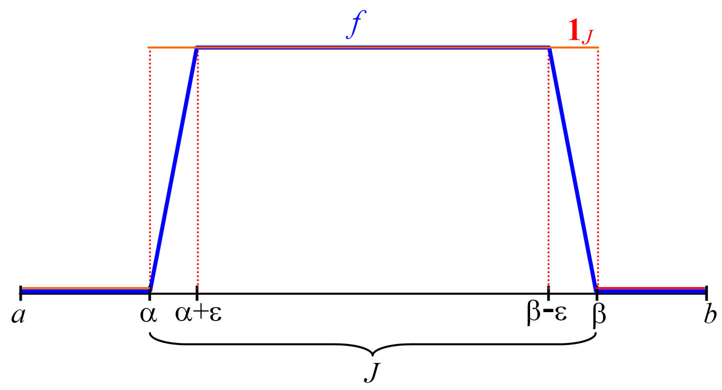
d.h. $L^1[a,b]$ separabel ist.

Im Fall (b) gilt $M = C_0[a,b]$ und we müssen beweisen, dass

$$C_0[a,b] \precapprox T_0,$$

d.h. die Indikatorfunktionen von Intervallen müssen von stetigen Funktionen approximiert werden. Für jedes Intervall $J \in T_0$ mit den Grenzen $\alpha < \beta$ und jedes hinreichend kleines $\varepsilon > 0$ betrachten wir die folgende stetige Funktion auf $[a, b]$

$$f(x) = \begin{cases} 1, & x \in [\alpha + \varepsilon, \beta - \varepsilon] \\ 0, & x \in [a, \alpha] \cup [\beta, b] \\ \text{linear auf } [\alpha, \alpha + \varepsilon] \text{ und } [\beta - \varepsilon, \beta] . \end{cases} \tag{2.28}$$



Dann gilt $f \in C_0[a, b]$ und

$$\|\mathbf{1}_J - f\|_1 \leq 2\varepsilon,$$

woraus folgt $C_0[a, b] \precapprox T_0$. ■

Im nächsten Beweis verwenden wir die Signum-Funktion: für alle $t \in \mathbb{R}$

$$\operatorname{sgn} t = \begin{cases} 1, & t \geq 0, \\ -1, & t < 0. \end{cases}$$

Offensichtlich gilt die Identität $t = |t| \operatorname{sgn} t$.

Satz 2.16 *Für alle reelle $a < b$ und $p \in [1, \infty)$ gilt folgendes.*

(a) $L^p[a, b]$ ist separabel.

(b) $C_0[a, b]$ liegt dicht in $L^p[a, b]$.

Beweis. Der Fall $p = 1$ wurde im Satz 2.15 bewiesen. Sei $p > 1$. Definieren wir eine Abbildung

$$\begin{aligned} \Phi &: L^1[a, b] \to L^p[a, b] \\ \Phi(f) &= |f|^{1/p} \operatorname{sgn} f. \end{aligned}$$

Da $|\Phi(f)|^p = |f|$ und $f \in L^1[a, b]$, so ist $|\Phi(f)|^p$ integrierbar und somit $\Phi(f) \in L^p[a, b]$.

Die Abbildung Φ hat die inverse Abbildung

$$\begin{aligned} \Psi &: L^p[a, b] \to L^1[a, b] \\ \Psi(g) &= |g|^p \operatorname{sgn} g, \end{aligned}$$

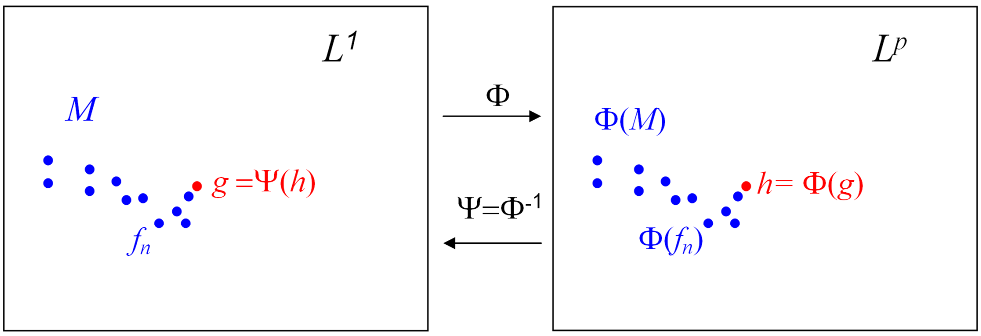
da
$$\Phi \circ \Psi(g) = \Phi\left(|g|^p \operatorname{sgn} g\right) = \left(|g|^p\right)^{1/p} \operatorname{sgn} g = |g| \operatorname{sgn} g = g$$
und analog
$$\Psi \circ \Phi(f) = \Psi\left(|f|^{1/p} \operatorname{sgn} f\right) = \left(|f|^{1/p}\right)^p \operatorname{sgn} f = |f| \operatorname{sgn} f = f.$$

Die Abbildung Φ hat die folgende Eigenschaft: für beliebige zwei Funktion $f, g \in L^1[a,b]$ gilt
$$|\Phi(f) - \Phi(g)| \le 2|f-g|^{1/p}, \tag{2.29}$$
(was unterhalb bewiesen wird), woraus folgt
$$\|\Phi(f) - \Phi(g)\|_p^p = \int_{[a,b]} |\Phi(f) - \Phi(g)|^p \, d\lambda \le 2^p \int_{[a,b]} |f-g| \, d\lambda = 2^p \|f-g\|_1 . \tag{2.30}$$

Jetzt beweisen wir, dass
$$M \precapprox L^1[a,b] \quad \Rightarrow \quad \Phi(M) \precapprox L^p[a,b] . \tag{2.31}$$

Im Fall (a) ist $\Phi(M)$ abzählbar, woraus folgt, dass $L^p[a,b]$ separabel ist. Im Fall (b) bemerken wir, dass
$$\Phi(C_0[a,b]) = C_0[a,b],$$
woraus folgt, dass $C_0[a,b]$ dicht in $L^p[a,b]$ liegt.

Beweisen wir jetzt (2.31), d.h.
$$\Phi(M) \precapprox L^p[a,b],$$
vorausgesetzt dass $M \precapprox L^1[a,b]$. Für jede Funktion $h \in L^p[a,b]$ liegt $g := \Psi(h)$ im $L^1[a,b]$. Somit existiert eine Folge $\{f_n\} \subset M$ mit $f_n \overset{L^1}{\to} g$, d.h.
$$\|f_n - g\|_1 \to 0 \text{ für } n \to \infty.$$



Nach (2.30) gilt
$$\|\Phi(f_n) - h\|_p^p = \|\Phi(f_n) - \Phi(g)\|_p^p \le 2^p \|f_n - g\|_1 \to 0 \quad \text{für } n \to \infty$$

d.h.

$$\Phi\left(f_n\right) \stackrel{L^p}{\to} h\ .$$

Da $\Phi(f_n) \in \Phi\left(M\right)$, so folgt es daraus dass $\Phi\left(M\right) \precapprox L^p\left[a,b\right]$.

Es bleibt noch die Ungleichung in (2.29) zu begründen. Dafür reicht es zu beweisen, dass die Funktion

$$\varphi\left(t\right) = \left|t\right|^{1/p} \operatorname{sgn} t$$

die folgende Ungleichung für alle $t, s \in \mathbb{R}$ erfüllt:

$$\left|\varphi\left(t\right) - \varphi\left(s\right)\right| \le 2\left|t-s\right|^{1/p}. \tag{2.32}$$

Ohne Beschränkung der Allgemeinheit können wir annehmen, dass mindestens eine von t, s nichtnegativ ist da sonst t und s durch $-t$ und $-s$ ersetzt werden können. Sei $t \ge s$ und somit $t \ge 0$.

Im Fall $s \ge 0$ gilt

$$\left|\varphi\left(t\right) - \varphi\left(s\right)\right| = t^{1/p} - s^{1/p}.$$

Zeigen wir dass

$$t^{1/p} - s^{1/p} \le \left(t-s\right)^{1/p}, \tag{2.33}$$

woraus (2.32) folgen wird. In der Tat, es gilt

$$\begin{aligned}
\left(t-s\right)^{1/p} + s^{1/p} &= t^{1/p}\left[\left(1-\frac{s}{t}\right)^{1/p} + \left(\frac{s}{t}\right)^{1/p}\right] \\
&\ge t^{1/p}\left[\left(1-\frac{s}{t}\right) + \left(\frac{s}{t}\right)\right] \\
&= t^{1/p},
\end{aligned}$$

wobei wir benutzt haben, dass $\frac{s}{t}$ und $1-\frac{s}{t}$ im Intervall $[0,1]$ liegen und $\frac{1}{p} \le 1$, und woraus (2.33) folgt.

Im Fall $s < 0$ gilt

$$\begin{aligned}
\left|\varphi\left(t\right) - \varphi\left(s\right)\right| &= t^{1/p} + \left|s\right|^{1/p} \le 2\max\left(t^{1/p}, \left|s\right|^{1/p}\right) \\
&= 2\max\left(t, \left|s\right|\right)^{1/p} \le 2\left(t+\left|s\right|\right)^{1/p} = 2\left(t-s\right)^{1/p}.
\end{aligned}$$

∎

Korollar 2.17 *$L^p_{\mathbb{C}}\left[a,b\right]$ ist separabel für alle $p \in [1,\infty)$.*

Beweis. Es folgt nach der Konstruktion von $L^p_{\mathbb{C}}$ dass

$$L^p_{\mathbb{C}}\left[a,b\right] = \left\{f = u + iv : u, v \in L^p\left[a,b\right]\right\}. \tag{2.34}$$

Sei M eine abzählbare dicht liegende Teilmenge von $L^p\left[a,b\right]$. Dann ist die Menge

$$M_{\mathbb{C}} = \left\{u + iv : u, v \in M\right\}$$

auch abzählbar. Zeigen wir, dass $M_{\mathbb{C}}$ dicht in $L^p_{\mathbb{C}}\left[a,b\right]$ liegt. In der Tat, für jede Funktion $f \in L^p_{\mathbb{C}}\left[a,b\right]$ gibt es die Folgen $\{u_n\}$ und $\{v_n\}$ aus M mit

$$u_n \stackrel{L^p}{\to} \operatorname{Re} f \quad \text{und} \quad v_n \stackrel{L^p}{\to} \operatorname{Im} f,$$

und daraus folgt, dass $f_n := u_n + iv_n \in L^p_{\mathbb{C}}$ und

$$f_n \overset{L^p_{\mathbb{C}}}{\to} f,$$

was zu beweisen war. ∎

Korollar 2.18 *Die Hilberträume $L^2[a,b]$ und $L^2_{\mathbb{C}}[a,b]$ besitzen abzählbare Orthogonalbasen. Darüber hinaus ist $L^2[a,b]$ isomorph zu l^2 und $L^2_{\mathbb{C}}[a,b]$ isomorph zu $l^2_{\mathbb{C}}$.*

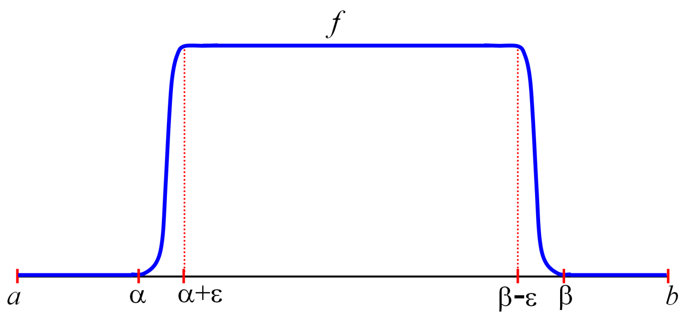
Beweis. Die erste Aussage folgt aus dem Satz 2.12, da die Hilberträume $L^2[a,b]$ und $L^2_{\mathbb{C}}[a,b]$ separabel sind. Die zweite Aussage folgt aus dem Korollar 2.13, da diese Räume unendlichdimensional sind. ∎

18.05.22 Vorlesung 12

Bemerkung. Betrachten wir für jedes $n \in \mathbb{N}$ den Raum

$$C_0^n[a,b] = \{f \in C^n[a,b] : f(a) = f(b) = 0\}.$$

Eine kleine Modifizierung von den obigen Beweisen ergibt, dass auch $C_0^n[a,b]$ dicht in $L^p[a,b]$ für alle $p \in [1,\infty)$ liegt. In der Tat, es reicht im Beweis von dem Satz 2.15 anstatt der Funktion $f \in C_0[a,b]$ aus (2.28) eine Funktion $f \in C_0^n[a,b]$ wählen, wie auf dem folgenden Bild:



2.14 Orthogonalbasen in $L^2[a,b]$

2.14.1 Polynome

Bezeichnen wir mit $\mathcal{P}$ die Menge von allen Polynomen von reeller Variable x mit reellen Koeffizienten, d.h.

$$\mathcal{P} = \bigcup_{n=0}^{\infty} \mathcal{P}_n.$$

Offensichtlich ist $\mathcal{P}$ ein Unterraum von $C[a,b]$ und von $L^p[a,b]$ für alle $p \in [1,\infty)$ und $a, b \in \mathbb{R}$, $a < b$.

Satz 2.19 (1. Approximationssatz von Weierstraß) *$\mathcal{P}$ liegt dicht in $C[a,b]$ bezüglich der* sup-*Norm; d.h. für jede $f \in C[a,b]$ gibt es eine Folge $\{P_n\}_{n\in\mathbb{N}}$ von Polynomen mit $P_n \rightrightarrows f$ auf $[a,b]$.*

Beweis in Analysis 2.

Satz 2.20 *$\mathcal{P}$ liegt dicht in $L^p[a,b]$ für alle $p \in [1,\infty)$; d.h. für jede $f \in L^p[a,b]$ gibt es eine Folge $\{P_n\}_{n\in\mathbb{N}}$ von Polynomen mit $P_n \xrightarrow{L^p} f$ auf $[a,b]$.*

Beweis. Nach dem Satz 2.19 haben wir

$$\mathcal{P} \precapprox C[a,b] \quad \text{bezüglich der sup-Norm.}$$

Da

$$\|f\|_p = \left(\int_{[a,b]} |f|^p \, d\lambda\right)^{1/p} \le \lambda([a,b])^{1/p} \sup|f|,$$

so folgt as daraus dass

$$\mathcal{P} \precapprox C[a,b] \quad \text{bezüglich der } p\text{-Norm.}$$

Da nach Satz 2.16 haben wir

$$C[a,b] \precapprox L^p[a,b],$$

woraus $\mathcal{P} \precapprox L^p[a,b]$ folgt. ■

Satz 2.21 *Seien $a < b$ reell.*

(a) Die Folge $\{1, x, x^2, ...\}$ ist ein Erzeugendensystem in $L^2[a,b]$.

(b) Die Folge $\{L_n\}_{n=0}^{\infty}$ von Legendre-Polynomen ist eine Orthogonalbasis in $L^2[-1,1]$.

Beweis. (a) Wir haben span $\{1, x, x^2, ...\} = \mathcal{P}$ und nach dem Satz 2.20 $\mathcal{P} \precapprox L^2[a,b]$. Es folgt, dass $\overline{\mathcal{P}} = L^2[a,b]$,

$$\overline{\operatorname{span}}\{1, x, x^2, ...\} = L^2[a,b],$$

so dass $\{1, x, x^2, ...\}$ ein Erzeugendensystem in $L^2[a,b]$ ist.

(b) Die Legendre Polynome werden wie folgt definiert:

$$L_n(x) = \frac{1}{2^n n!}\frac{d^n}{dx^n}\left(x^2-1\right)^n.$$

Die Funktion $L_n(x)$ ist ein Polynom des Grades n und die Folge $\{L_n\}$ ist orthogonal in $L^2[-1,1]$ (siehe Aufgabe 45). Nach dem Satz 2.9 reicht es zu zeigen, dass $\{L_n\}_{n=0}^{\infty}$ ein Erzeugendensystem ist. Dafür reicht es zu beweisen, dass

$$\operatorname{span}\{1, x, x^2, ...\} = \operatorname{span}\{L_0, L_1, L_2, ...\}.$$

Eigentlich gilt diese Identität für jede Folge $\{L_n\}_{n=0}^{n}$ von Polynomen vorausgesetzt, dass $\deg L_n = n$.

Die Inklusion "$\supset$" ist offensichtlich, da span $\{1, x, x^2, ...\}$ alle Polynome enthält. Für die Inklusion "$\subset$" reicht es zu zeigen, dass für alle $n \ge 0$

$$x^n \in \operatorname{span}\{L_0, ..., L_n\}. \tag{2.35}$$

Für $n = 0$ ist es offensichtlich da $L_0 = \text{const}$. Für den Induktionsschritt "von $< n$ nach n" bemerken wir, dass

$$L_n(x) = cx^n + Q(x),$$

wobei $c \neq 0$ und Q ein Polynom mit $\deg Q \leq n - 1$. Nach der Induktionsvoraussetzung gilt

$$Q \in \operatorname{span}\{L_0, ..., L_{n-1}\}.$$

woraus folgt

$$x^n = c^{-1}(L_n - Q) \in \operatorname{span}\{L_0, ..., L_{n-1}, L_n\},$$

was zu beweisen war. ■

2.14.2 Trigonometrische Polynome

Die elementaren trigonometrischen Polynome sind die Funktionen

$$1, \cos x, \sin x, \cos 2x, \sin 2x, ..., \cos nx, \sin nx, ... \tag{2.36}$$

wobei x eine reelle Variable ist und $n \in \mathbb{N}$. Ein trigonometrisches Polynom ist eine beliebige endliche lineare Kombination der Funktionen (2.36):

$$a_0 + a_1 \cos x + b_1 \sin x + ... + a_n \cos nx + b_n \sin nx,$$

wobei $n \in \mathbb{N}$ und $a_k, b_k \in \mathbb{R}$. Offensichtlich ist jedes trigonometrisches Polynom 2π-periodisch. Die Menge von allen trigonometrischen Polynomen wird mit $\mathcal{T}$ bezeichnet, d.h.

$$\mathcal{T} = \operatorname{span}\{1, \cos x, \sin x, \cos 2x, \sin 2x, ..., \cos nx, \sin nx, ...\}.$$

Bezeichnen wir mit $C_{2\pi}$ die Menge von reellwertigen stetigen 2π-periodischen Funktionen auf $\mathbb{R}$. Man kann $C_{2\pi}$ als Unterraum von $C[0, 2\pi]$ betrachten. Offensichtlich ist $\mathcal{T}$ ein Unterraum von $C_{2\pi}$.

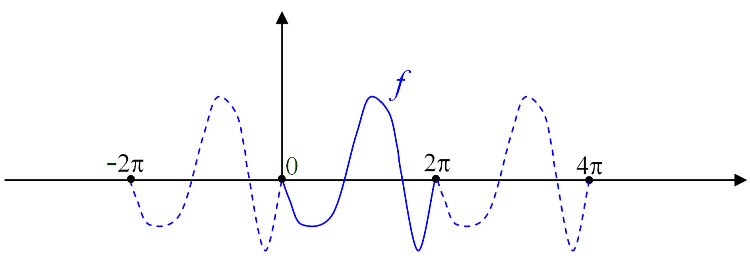
Satz 2.22 (2. Approximationssatz von Weierstraß) *$\mathcal{T}$ liegt dicht in $C_{2\pi}$ bezüglich der* sup-*Norm.*

Der Beweis erfolgt in einem späteren Kapitel.

Satz 2.23 *Die Folge (2.36) ist eine Orthogonalbasis in $L^2[0, 2\pi]$.*

Beweis. Die Folge (2.36) ist orthogonal in $L^2[0, 2\pi]$ (Aufgabe 41). Somit reicht es zu beweisen, dass sie ein Erzeugendensystem in $L^2[0, 2\pi]$ ist.

Jede Funktion $f \in C_0[0, 2\pi]$ lässt sich als eine 2π-periodische Funktion auf $\mathbb{R}$ fortsetzen, und diese Fortsetzung ist stetig da $f(0) = f(2\pi)$.



Somit haben wir $C_0[0,2\pi] \subset C_{2\pi}$. Nach dem Satz 2.22 gilt $\mathcal{T} \precapprox C_{2\pi}$ und somit auch

$$\mathcal{T} \precapprox C_0[0,2\pi]. \tag{2.37}$$

(Bemerken wir, dass der Raum $C[0,2\pi]$ sich nicht als Teilmenge von $C_{2\pi}$ betrachten lässt, da die 2π-periodische Fortsetzung von $f \in C[0,2\pi]$ nicht unbedingt stetig ist; deshalb verwenden wir hier $C_0[0,2\pi]$ anstatt $C[a,b]$.)

Die Relation (2.37) gilt bezüglich der sup-Norm und somit auch bezüglich der 2-Norm. Da nach dem Satz 2.16 gilt

$$C_0[0,2\pi] \precapprox L^2[0,2\pi],$$

so erhalten wir

$$\mathcal{T} \precapprox L^2[0,2\pi],$$

woraus folgt, dass die Folge

$$1, \cos x, \sin x, \cos 2x, \sin 2x, ..., \cos nx, \sin nx, ...$$

ein Erzeugendensystem ist. ■

In einem späteren Kapitel geben wir noch einen Beweis von dem Satz 2.23 ohne den 2.Approximationssatz von Weierstraß.

Satz 2.24 *Die Folge $\{e^{inx}\}_{n\in\mathbb{Z}}$ ist eine Orthogonalbasis in $L^2_{\mathbb{C}}[0,2\pi]$.*

Beweis. Die Folge $\{e^{inx}\}_{n\in\mathbb{Z}}$ ist orthogonal in $L^2_{\mathbb{C}}[0,2\pi]$ ist (Aufgabe 41). Beweisen wir, dass $\{e^{inx}\}_{n\in\mathbb{Z}}$ ein Erzeugendensystem in $L^2_{\mathbb{C}}[0,2\pi]$ ist. Nach dem Satz 2.23 gilt

$$\operatorname{span}_{\mathbb{R}}\{1, \cos x, \sin x, \cos 2x, \sin 2x, ..., \cos nx, \sin nx, ...\} \precapprox L^2_{\mathbb{R}}[0,2\pi]$$

Da

$$\cos nx = \frac{e^{inx} + e^{-inx}}{2} \quad \text{und} \quad \sin nx = \frac{e^{inx} - e^{-inx}}{2i},$$

so erhalten wir, dass alle Funktionen $\cos nx$ und $\sin nx$ in $\operatorname{span}_{\mathbb{C}}\{e^{inx}\}$ liegen, woraus folgt

$$\operatorname{span}_{\mathbb{C}}\left\{e^{inx}\right\} \precapprox L^2_{\mathbb{R}}[0,2\pi].$$

Jede Funktion aus $L^2_{\mathbb{C}}[0,2\pi]$ hat der Form $u+iv$ wobei $u,v \in L^2_{\mathbb{R}}[0,2\pi]$ woraus folgt, dass

$$\operatorname{span}_{\mathbb{C}} L^2_{\mathbb{R}}[0,2\pi] = L^2_{\mathbb{C}}[0,2\pi].$$

Es folgt, dass

$$\operatorname{span}_{\mathbb{C}}\left\{e^{inx}\right\} \precapprox L^2_{\mathbb{C}}[0,2\pi],$$

so dass $\{e^{inx}\}$ ein Erzeugendensystem ist. ■

Es folgt aus dem Satz 2.23, dass jede Funktion $f \in L^2[0,2\pi]$ eine eindeutige Darstellung in der Form

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty}(a_n \cos nx + b_n \sin nx) \tag{2.38}$$

mit reellwertigen Koeffizienten a_n, b_n hat, wobei die Reihe in $L^2[0, 2\pi]$ konvergiert. Diese Reihe heißt die *Fourierreihe* der Funktion f, und die Zahlen a_n, b_n heißen die Fourier-Koeffizienten von f.

Für beliebige Orthogonalbasis $\{v_n\}$ in einem Hilbertraum H gilt für jedes $u \in H$ die Darstellung

$$u = \sum_n c_n v_n$$

mit den Koeffizienten

$$c_n = \frac{(u, v_n)}{\|v_n\|^2}.$$

Ist $\{v_n\}$ eine Orthonormalbasis so gilt diese Darstellung nach dem Satz 2.9; für den allgemeinen Fall siehe Aufgabe 49. Die Zahlen c_n heißen auch die Fourier-Koeffizienten von u in der Basis $\{v_n\}$.

Da

$$\|\cos nx\|^2_{L^2[0,2\pi]} = \|\sin nx\|^2_{L^2[0,2\pi]} = \pi \text{ und } \quad \|1\|^2_{L^2[0,2\pi]} = 2\pi,$$

so erhalten wir die folgenden Formeln für die Koeffizienten der Fourierreihe (2.38): für $n \geq 1$

$$a_n = \frac{1}{\pi}(f, \cos nx) \quad \text{und} \quad b_n = \frac{1}{\pi}(f, \sin nx)$$

und für $n = 0$

$$\frac{a_0}{2} = \frac{1}{2\pi}(f, 1),$$

d.h. für alle $n \geq 0$

$$a_n = \frac{1}{\pi}\int_{[0,2\pi]} f(x)\cos nx\, d\lambda \ \text{ und } \ b_n = \frac{1}{\pi}\int_{[0,2\pi]} f(x)\sin nx\, d\lambda.$$

Ebenfalls hat jede Funktion $f \in L^2_{\mathbb{C}}[0, 2\pi]$ eine eindeutige Darstellung in der Form

$$f(x) = \sum_{n\in\mathbb{Z}} c_n e^{inx} \tag{2.39}$$

mit komplexwertigen Koeffizienten c_n, wobei die Reihe in $L^2_{\mathbb{C}}[0, 2\pi]$ konvergiert. Diese Reihe heißt auch die Fourierreihe der Funktion f. Für die Koeffizienten c_n erhält man

$$c_n = \frac{(f, e^{inx})}{\|e^{inx}\|^2_{L^2_{\mathbb{C}}[0,2\pi]}} = \frac{1}{2\pi}\int_{[0,2\pi]} f(x)\, e^{-inx} d\lambda$$

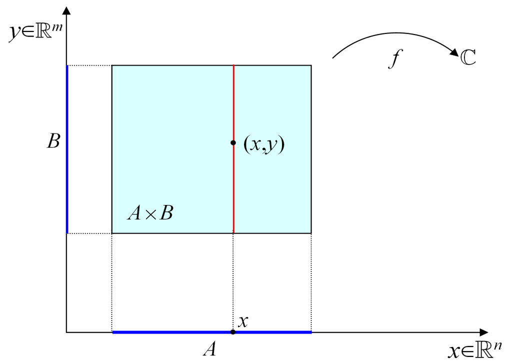
(siehe auch Aufgabe 51).

2.15 Orthogonalbasen in $L^2(A \times B)$

Wir bezeichnen mit λ_n das Lebesgue-Maß in $\mathbb{R}^n$. Für jede beschränkte messbare Teilmenge $A \subset \mathbb{R}^n$ betrachten wir den Hilbertraum $L^2(A, \lambda_n)$ (über $\mathbb{R}$ oder $\mathbb{C}$).

Satz 2.25 *Seien $A \subset \mathbb{R}^n$ und $B \subset \mathbb{R}^m$ beschränkte messbare Mengen. Sei $\{u_k(x)\}_{k\in\mathbb{N}}$ eine Orthonormalbasis in $L^2(A, \lambda_n)$ und $\{v_l(y)\}_{l\in\mathbb{N}}$ – eine Orthonormalbasis in $L^2(B, \lambda_m)$. Dann ist die Folge $\{u_k(x)\, v_l(y)\}_{k,l\in\mathbb{N}}$ eine Orthonormalbasis in $L^2(A \times B, \lambda_{n+m})$.*

Die Menge $A \times B$ ist eine Teilmenge von $\mathbb{R}^{n+m}$. Ist $u(x)$ eine Funktion auf A und $v(y)$ eine Funktion auf B, so heißt die Funktion $u(x)\,v(y)$ das *Tensorprodukt* von u und v. Diese Funktion hat offensichtlich den Definitionsbereich $A \times B$. Sind u und v messbar so ist $u(x)\,v(y)$ auch messbar.



Im Beweis verwenden wir den folgenden Satz aus der Maß- und Integrationstheorie.

Satz von Fubini *Ist f eine (komplexwertige) integrierbare Funktion auf $A \times B$, so gilt*

$$\int_{A\times B} f d\lambda_{n+m} = \int_A \left(\int_B f(x,y)\, d\lambda_m(y) \right) d\lambda_n(x), \tag{2.40}$$

wobei die Funktion $y \mapsto f(x,y)$ für fast alle $x \in A$ auf B integrierbar ist und die Funktion $x \mapsto \int_B f(x,y)\, d\lambda_m(y)$ auf A integrierbar ist.

Die gleiche Aussage gilt wenn f eine nichtnegative messbare Funktion ist (in diesem Fall soll "integrierbar" durch "nichtnegativ messbar" überall ersetzt werden).

Beweis von dem Satz 2.25. Setzen wir

$$w_{k,l}(x,y) = u_k(x)\, v_l(y)$$

und bemerken zuerst, dass $w_{k,l}$ messbar auf $A \times B$ ist und in $L^2(A \times B)$ liegt, da nach (2.40)

$$\begin{aligned}
\int_{A\times B} |w_{k,l}|^2 d\lambda_{n+m} &= \int_A \left(\int_B |u_k(x)|^2 |v_l(y)|^2 d\lambda_m(y) \right) d\lambda_n(x) \\
&= \int_A \left(\int_B |v_l(y)|^2 d\lambda_m(y) \right) |u_k(x)|^2 d\lambda_n(x) \\
&= \|v_l\|^2_{L^2(B)} \|u_k\|^2_{L^2(A)} = 1 < \infty.
\end{aligned}$$

Insbesondere gilt

$$\|w_{k,l}\|_{L^2(A\times B)} = 1.$$

20.05.22 Vorlesung 13

Jetzt beweisen wir, dass die Doppelfolge $\{w_{k,l}\}$ orthogonal ist. Es gilt nach (2.40)

$$\begin{aligned}
(w_{k,l}, w_{k',l'}) &= \int_{A\times B} w_{k,l}\overline{w_{k'l'}}d\lambda_{n+m} \\
&= \int_{A\times B} u_k(x)\, v_l(y)\,\overline{u_{k'}(x)\, v_{l'}(y)}d\lambda_{n+m} \\
&= \int_A u_k(x)\,\overline{u_{k'}(x)}d\lambda_n(x) \int_B v_l(y)\,\overline{v_{l'}(y)}d\lambda_m(y) \\
&= (u_k, u_{k'})_{L^2(A)}\,(v_l, v_l')_{L^2(B)}\,.
\end{aligned}$$

Im Fall $(k,l) \neq (k',l')$ gilt entweder $k \neq k'$ oder $l \neq l'$ so dass entweder $(u_k, u_k') = 0$ oder $(v_l, v_{l'}) = 0$, und somit immer $(w_{k,l}, w_{k',l'}) = 0$ so dass die Folge $\{w_{k,l}\}$ orthogonal ist.

Jetzt beweisen wir, dass die Folge $\{w_{k,l}\}$ eine Basis ist. Dafür reicht zu beweisen, dass diese Folge maximal ist, d.h. für jede Funktion $f \in L^2(A\times B)$ gilt

$$\forall k,l \in \mathbb{N} \quad (f, w_{k,l}) = 0 \ \Rightarrow f = 0.$$

Nach (2.40) gilt

$$\begin{aligned}
(f, w_{k,l}) &= \int_{A\times B} f(x,y)\,\overline{u_k(x)\, v_l(y)}d\lambda_{n+m} \\
&= \int_A \left(\int_B f(x,y)\,\overline{v_l(y)}d\lambda_m(y)\right)\overline{u_k(x)}d\lambda_n(x) \\
&= \int_A g_l(x)\,\overline{u_k(x)}d\lambda_n(x)\,, \qquad (2.41)
\end{aligned}$$

wobei

$$g_l(x) = \int_B f(x,y)\,\overline{v_l(y)}d\lambda_m(y)\,. \qquad (2.42)$$

Nach dem Satz von Fubini, das Integral (2.42) ist für fast alle $x \in A$ definiert. Wir setzen $g_l(x) = 0$ wenn das Integral (2.42) nicht definiert ist. Zeigen wir, dass $g_l \in L^2(A)$. Nach der Cauchy-Schwarz-Ungleichung gilt

$$|g_l(x)|^2 \leq \int_B |f(x,y)|^2\, d\lambda_m(y) \int_B |v_l(y)|^2\, d\lambda_m(y) = \int_B |f(x,y)|^2\, d\lambda_m(y)$$

und nach dem Satz von Fubini

$$\int_A |g_l(x)|^2\, d\lambda_n(x) \leq \int_A \left(\int_B |f(x,y)|^2\, d\lambda_m(y)\right) d\lambda_n(x) = \int_{A\times B} |f(x,y)|^2\, d\lambda_{n+m} < \infty$$

so dass $\|g_l\|_{L^2(A)} < \infty$ und somit $g_l \in L^2(A)$.

Es folgt aus (2.41), dass für alle $k \in \mathbb{N}$

$$(g_l, u_k) = \int_A g_l(x)\,\overline{u_k(x)}d\lambda_n(x) = (f, w_{k,l}) = 0.$$

Da $\{u_k\}$ eine Basis in $L^2(A)$ ist, so folgt es, dass $g_l = 0$ als Element von $L^2(A)$ d.h. $g_l(x) = 0$ für fast alle $x \in A$. Somit gibt es eine Nullmenge $N_l \subset A$ mit

$$\int_B f(x,y)\overline{v_l(y)}d\lambda_m(y) = 0 \quad \text{für alle } x \in A \setminus N_l. \tag{2.43}$$

Nach dem Satz von Fubini ist Funktion $f^2(x,\cdot)$ auf B integrierbar für fast alle $x \in A$. Somit existiert eine Nullmenge $N' \subset A$ so dass

$$f(x,\cdot) \in L^2(B) \quad \text{für alle } x \in A \setminus N'. \tag{2.44}$$

Die Menge

$$N = \bigcup_{l \in \mathbb{N}} N_l \cup N'$$

ist auch Nullmenge, und für alle $x \in A \setminus N$ gelten (2.43) und (2.44), woraus folgt

$$(f(x,\cdot), v_l)_{L^2(B)} = 0.$$

Da $\{v_l\}$ eine Basis in $L^2(B)$ ist, so folgt es, dass $f(x,\cdot) = 0$ als Element von $L^2(B)$, und zwar für alle $x \in A \setminus N$. Es folgt, dass

$$\begin{aligned}
\int_{A\times B} |f(x,y)|^2 d\lambda_{n+m} &= \int_A \left(\int_B |f(x,y)|^2 d\lambda_m(y) \right) d\lambda_n(x) \\
&= \int_{A\setminus N} \|f(x,\cdot)\|^2_{L^2(B)} d\lambda_n(x) \\
&= \int_{A\setminus N} 0 d\lambda_n(x) = 0
\end{aligned}$$

und somit $f = 0$ as Element von $L^2(A \times B)$, was zu beweisen war. ■

Beispiel. Da die Folge $\{L_k\}_{k=0}^{\infty}$ von Legendre-Polynomen eine Orthogonalbasis in $L^2[-1,1]$ ist, so folgt es aus dem Satz 2.25, dass die Folge

$$\{L_k(x)\,L_m(y)\}_{k,m=0}^{\infty}$$

eine Orthogonalbasis in $L^2\left([-1,1]^2\right)$ ist. Zum Beispiel,

$$L_1(x)\,L_2(y) = \text{const}\, x \left(3y^2 - 1\right).$$

Per Induktion erhält man, dass die Folge

$$\{L_{k_1}(x_1)\,L_{k_2}(x_2)\,...L_{k_n}(x_n)\}_{k_1,...,k_n=0}^{\infty}$$

eine Orthogonalbasis in $L^2([-1,1]^n)$ ist, wobei $x = (x_1,...,x_n) \in \mathbb{R}^n$.

Beispiel. Betrachten wir die folgende Folge von komplexwertigen Funktionen auf $\mathbb{R}^n$:

$$\left\{e^{i\mathbf{k}\cdot x}\right\}_{\mathbf{k}\in\mathbb{Z}^n}, \tag{2.45}$$

wobei $\mathbf{k} = (k_1,...,k_n) \in \mathbb{Z}^n$, $x = (x_1,...,x_n) \in \mathbb{R}^n$ und

$$\mathbf{k} \cdot x = \sum_{j=1}^{n} k_j x_j$$

(i ist die imaginäre Einheit). Es folgt aus den Sätzen 2.24 und 2.25, dass die Folge (2.45) eine Orthogonalbasis in $L^2_{\mathbb{C}}([0,2\pi]^n)$ ist. Jede Funktion $f \in L^2_{\mathbb{C}}([0,2\pi]^n)$ lässt sich wie folgt darstellen:

$$f = \sum_{\mathbf{k}\in\mathbb{Z}^n} c_{\mathbf{k}} e^{i\mathbf{k}\cdot x}, \tag{2.46}$$

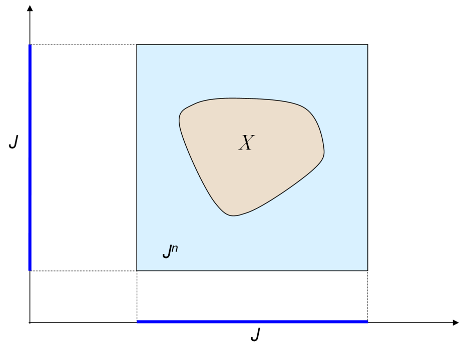
wobei

$$c_{\mathbf{k}} = \frac{1}{(2\pi)^n} \int_{[0,2\pi]^n} f(x)\, e^{-i\mathbf{k}\cdot x} d\lambda_n$$

und die Reihe (2.46) in $L^2([0,2\pi]^n)$ konvergiert (siehe auch Aufgabe 51). Die Reihe (2.46) heißt *mehrdimensionale Fourierreihe.*

Korollar 2.26 *Für beliebige beschränkte messbare Menge $X \subset \mathbb{R}^n$, der Hilbertraum $L^2(X,\lambda_n)$ ist separabel. Folglich besitzt $L^2(X,\lambda_n)$ eine Orthonormalbasis.*

Beweis. Da X beschränkt ist, so gibt es ein Intervall $J \subset \mathbb{R}$ so dass $X \subset J^n$. Nach den Korollar 2.18 besitzt $L^2(J,\lambda)$ eine Orthonormalbasis. Nach dem Satz 2.25 besitzt auch $L^2(J^n,\lambda_n)$ eine Orthonormalbasis. Nach dem Satz 2.12 ist $L^2(J^n,\lambda_n)$ separabel. Beweisen wir, dass auch $L^2(X,\lambda_n)$ separabel ist, woraus nach dem Satz 2.12 folgt, dass $L^2(X,\lambda_n)$ eine Orthonormalbasis besitzt.



Jede Funktion $f \in L^2(X,\lambda_n)$ lässt sich auf J^n fortsetzen wie folgt:

$$f = 0 \text{ auf } J^n \setminus X.$$

Die erweiterte Funktion f liegt in $L^2(J^n,\lambda_n)$ da

$$\int_{J^n} |f|^2 d\lambda_n = \int_X |f|^2 d\lambda_n + \int_{J^n\setminus X} |f|^2 d\lambda_n = \int_X |f|^2 d\lambda_n < \infty.$$

Unter diese Fortsetzung wird das Skalarprodukt bewahrt: für $f,g \in L^2(X,\lambda)$ gilt

$$(f,g)_{L^2(J^n,\lambda_n)} = \int_{J^n} f\overline{g} d\lambda_n = \int_X f\overline{g} d\lambda_n + \int_{J^n\setminus X} f\overline{g} d\lambda_n = \int_X f\overline{g} d\lambda_n = (f,g)_{L^2(X,\lambda_n)}.$$

Somit lässt $L^2(X,\lambda_n)$ sich als einen Unterraum von dem Skalarproduktraum $L^2(J^n,\lambda_n)$ betrachten. Dieser Unterraum ist abgeschlossen da $L^2(X,\lambda_n)$ vollständig ist (Aufgabe 15). Da der Raum $L^2(J^n,\lambda_n)$ separabel ist, so ist jeder abgeschlossene Unterraum davon auch separabel (siehe Aufgabe 50). Folglich ist $L^2(X,\lambda_n)$ separabel. ∎

2.16 Weitere Beispiele von Orthogonalbasen

Sei μ ein beliebiges endliches Borel-Maß auf $[a,b]$. Alle stetige Funktionen auf $[a,b]$ sind messbar bezüglich μ und liegen in $L^2([a,b],\mu)$. Man beweist wie im Satz 2.16, dass

$$C[a,b] \text{ liegt dicht in } L^2([a,b],\mu),$$

und wie im Satz 2.20, dass

$$\mathcal{P} \text{ liegt dicht in } L^2([a,b],\mu).$$

Sei $\{Q_n\}_{n=0}^{\infty}$ eine Folge von Polynomen mit den folgenden zwei Eigenschaften:

(i) $\deg Q_n = n$;

(ii) $\{Q_n\}$ ist orthogonal in $L^2([a,b],\mu)$.

Dann erhält man wie im Beweis von dem Satz 2.21, dass

$$\operatorname{span}\{Q_0,Q_1,Q_2,...\} = \operatorname{span}\left\{1,x,x^2,...\right\} = \mathcal{P},$$

woraus folgt, dass $\{Q_n\}$ eine Orthogonalbasis in $L^2([a,b],\mu)$ ist.

Solche Folge $\{Q_n\}$ existiert immer: man erhält sie aus $\{1,x,x^2,...\}$ mit Hilfe von dem Orthogonalisierungsverfahren.

Normalerweise erhält man ein neues Maß μ auf $[a,b]$ wie folgt. Wählen wir eine positive Funktion $w(x)$ auf $[a,b]$ die bezüglich λ integrierbar ist, und definieren $\mu(A)$ für jede messbare Menge $A\subset[a,b]$ mit

$$\mu(A) = \int_A w(x)\,d\lambda.$$

Man schreibt in diesem Fall

$$d\mu = w(x)\,d\lambda.$$

Die Funktion $w(x)$ heißt das Gewicht oder die Dichte von μ bezüglich λ. Das Skalarprodukt in $L^2([a,b],\mu)$ ist somit

$$(f,g)_{L^2([a,b],\mu)} = \int_{[a,b]} f(x)\,g(x)\,w(x)\,d\lambda.$$

Beispiel. Für jede nichtnegative ganze Zahl n betrachten wir die folgende Funktion auf dem Intervall $[-1,1]$:

$$T_n(x) = \cos(n\arccos x)$$

Man kann zeigen dass die Funktion T_n ein Polynom von x des Grades n ist. Mit Substitution $y=\arccos x$ erhalten wir eine äquivalente Definition

$$\cos ny = T_n(\cos y).$$

Zum Beispiel, es gilt

$$T_0(x)=1,\quad T_1(x)=x,\quad T_2(x)=2x^2-1,\quad T_3(x)=4x^3-3x,\quad \text{usw.}$$

da

$$\cos 2y = 2\cos^2 y - 1, \quad \cos 3y = 4\cos^3 y - 3\cos y, \quad \text{usw.}$$

Die Funktionen T_n heißen *Tschebyschow-Polynome erster Art.* Man kann auch beweisen, dass die Folge $\{T_n\}$ orthogonal in $L^2([-1,1],\mu)$ ist, wobei

$$d\mu = \frac{d\lambda}{\sqrt{1-x^2}}$$

(siehe Aufgabe 57).

Beispiel. Die Tschebyschow-Polynome *zweiter Art* definiert man auf dem Intervall $(-1,1)$ mit

$$U_{n-1}(x) = \frac{\sin(n \arccos x)}{\sqrt{1-x^2}},$$

für jedes $n \in \mathbb{N}$. Mit Substitution $y = \arccos x$ erhalten wir eine äquivalente Definition

$$\sin ny = U_{n-1}(\cos y)\sin y$$

Zum Beispiel,

$$U_0(x) = 1, \quad U_1(x) = 2x, \quad U_2(x) = 4x^2 - 1, \quad U_3(x) = 8x^3 - 4x, \text{ usw.}$$

da

$$\sin 2y = 2\sin y \cos y, \quad \sin 3y = \left(4\cos^2 y - 1\right)\sin y, \quad \text{usw.}$$

Man kann zeigen, dass U_k wirklich ein Polynom des Grades k ist und die Folge $\{U_k\}_{k=0}^{\infty}$ orthogonal in $L^2([-1,1],\mu)$ mit dem Maß

$$d\mu = \sqrt{1-x^2}\,d\lambda$$

ist (siehe Aufgabe 57).

Beispiel. Betrachten wir auf $[-1,1]$ das Gewicht

$$w(x) = (1-x)^{\alpha}(1+x)^{\beta},$$

wobei $\alpha, \beta > -1$ (unter diese Bedingung ist w integrierbar bezüglich λ). Die orthogonale Polynome bezüglich dieser Gewichtfunktion heißen *Jacobi-Polynome* und werden mit $P_n^{(\alpha,\beta)}$ bezeichnen.

Zum Beispiel, $P_n^{(0,0)}$ sind Legendre-Polynome, $P_n^{\left(-\frac{1}{2},-\frac{1}{2}\right)}$ sind Tschebyschow-Polynome erster Art, $P_n^{\left(\frac{1}{2},\frac{1}{2}\right)}$ sind Tschebyschow-Polynome zweiter Art, und $P_n^{(\alpha,\alpha)}$ heißen auch *Gegenbauer-Polynome.*

Beispiel. Betrachten wir in $\mathbb{R}$ das *Gauß-Maß* γ:

$$d\gamma = e^{-x^2}d\lambda,$$

und den Hilbertraum $L^2(\mathbb{R},\gamma)$. Dieses Maß ist endlich, und alle Polynome liegen in $L^2(\mathbb{R},\gamma)$. Man kann beweisen, dass die Folge $\{1, x, x^2, ...\}$ ein Erzeugendensystem in $L^2(\mathbb{R},\gamma)$ ist. Mit Hilfe von dem Gram-Schmidtschen Orthogonalisierungsverfahren erhält man eine Orthogonalbasis in $L^2(\mathbb{R},\gamma)$, die aus Polynomen besteht. Diese Polynome sind *Hermite-Polynome*

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n} e^{-x^2}$$

mit $n = 0, 1, 2, ...$ (siehe Aufgabe 58). Zum Beispiel, es gilt

$$H_0(x) = 1, \quad H_1(x) = 2x, \quad H_2(x) = 4x^2 - 2, \quad H_3(x) = 8x^3 - 12x, \quad \text{usw.}$$

Beispiel. Betrachten wir in $[0,\infty)$ das Maß

$$d\mu = e^{-x} d\lambda.$$

Im Hilbertraum $L^2([0,\infty),\mu)$ erhält man eine Orthogonalbasis aus den *Laguerre-Polynomen*:

$$L_n(x) = e^x \frac{d^n}{dx^n}\left(x^n e^{-x}\right).$$

Zum Beispiel,

$$L_0(x) = 1, \quad L_1(x) = -x+1, \quad L_2(x) = x^2 - 4x + 2, \quad L_3(x) = -x^3 + 9x^2 - 18x + 6, \quad \text{usw.}$$

Chapter 3

Lineare Operatoren im Hilbertraum

3.1 Operatornorm und die Operatorenräume

Seien X, Y zwei normierte Vektorräume über $\mathbb{K}$ (wobei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{C}$) mit den Normen $\|\cdot\|_X$ und $\|\cdot\|_Y$. Eine lineare Abbildung $A : X \to Y$ heißt auch ein (linearer) *Operator*. Die Menge von allen linearen Operatoren von X nach Y wird mit $\mathcal{L}(X, Y)$ bezeichnet. Diese Menge ist ein Vektorraum mit den folgenden linearen Operationen:

$$\begin{aligned} (A + B)\, x &= Ax + Bx \\ (\alpha A)\, x &= \alpha\,(Ax)\,, \end{aligned}$$

wobei $x \in X$ und $\alpha \in \mathbb{R}$. Das Nullelement in $\mathcal{L}(X, Y)$ ist der Nulloperator $Ax \equiv 0$.

Zum Beispiel, $\mathcal{L}(\mathbb{R}^m, \mathbb{R}^n)$ besteht aus allen linearen Abbildungen von $\mathbb{R}^m$ nach $\mathbb{R}^n$, d.h. aus allen Matrizen $n \times m$.

Definition. Für jeden Operator $A \in \mathcal{L}(X, Y)$ definieren wir die *Operatornorm* $\|A\|_{X\to Y}$ mit

$$\|A\|_{X\to Y} = \sup_{x\in X\setminus\{0\}} \frac{\|Ax\|_Y}{\|x\|_X}. \tag{3.1}$$

Die Tiefstellungen X, Y, $X \to Y$ in der Bezeichnung der Norm kann man weglassen wenn es klar in welchem Raum die Norm benutzt wird. Z.B. in (3.1) ist es eindeutig klar ohne Indizes, dass $\|x\|$ die Norm in X ist, $\|Ax\|$ die Norm in Y und $\|A\|$ – die Operatornorm. Somit schreiben wir (3.1) um wie folgt:

$$\|A\| = \sup_{x\in X\setminus\{0\}} \frac{\|Ax\|}{\|x\|}. \tag{3.2}$$

Bemerken wir, dass $\|A\| \in [0, \infty]$. Es folgt aus (3.2) dass

$$\|Ax\| \leq \|A\|\, \|x\| \quad \text{für alle } x \in X. \tag{3.3}$$

Nach (3.2) ist $\|A\|$ die kleinste obere Schranke von dem Quotient $\frac{\|Ax\|}{\|x\|}$, was auch bedeutet, dass $\|A\|$ gleich die minimale Zahl C mit

$$\|Ax\| \leq C\, \|x\| \quad \text{für alle } x \in X$$

ist oder $\|A\|$ gleich ∞ ist wenn solches C nicht existiert.

25.05.22 Vorlesung 14

Da für alle $\lambda > 0$

$$\frac{\|A(\lambda x)\|}{\|\lambda x\|} = \frac{\|Ax\|}{\|x\|},$$

so folgt es aus aus (3.2) dass

$$\|A\| = \sup_{x \in X,\ \|x\|=1} \|Ax\|. \tag{3.4}$$

Die Operatornorm hat die folgenden Eigenschaften:

(a) $\|A + B\| \leq \|A\| + \|B\|$

da

$$\|A + B\| = \sup_{\|x\|=1} \|(A + B)x\| \leq \sup_{\|x\|=1} \|Ax\| + \sup_{\|x\|=1} \|Bx\| = \|A\| + \|B\|.$$

(b) $\|\alpha A\| = |\alpha| \|A\|$ für alle $\alpha \in \mathbb{K}$

da

$$\|\alpha A\| = \sup_{\|x\|=1} \|(\alpha A)x\| = \sup_{\|x\|=1} \|\alpha (Ax)\| = |\alpha| \sup_{\|x\|=1} \|Ax\| = |\alpha| \|A\|.$$

(c) $\|A\| = 0 \Leftrightarrow A = 0$

da nach (3.2)

$$\|A\| = 0 \Leftrightarrow \|Ax\| = 0\ \forall x \in X \Leftrightarrow Ax = 0\ \forall x \in X \Leftrightarrow A = 0.$$

Um die Operatornorm zu einer richtigen Norm zu machen, fehlt noch die Endlichkeit, da es Operatoren mit $\|A\| = \infty$ gibt.

Definition. Ein Operator $A : X \to Y$ heißt *beschränkt* wenn $\|A\| < \infty$. Die Menge von allen beschränkten Operatoren von X nach Y wird mit $\mathcal{B}(X, Y)$ bezeichnet.

Es folgt aus (a) und (b), dass $\mathcal{B}(X, Y)$ ein Unterraum von $\mathcal{L}(X, Y)$ ist und somit ein Vektorraum. Darüber hinaus ist die Operatornorm eine Norm in $\mathcal{B}(X, Y)$, so dass $\mathcal{B}(X, Y)$ ein normierter Vektorraum ist.

Satz 3.1 *Ein Operator $A \in \mathcal{L}(X, Y)$ ist beschränkt genau dann, wenn A stetig ist.*

Beweis. Ist A beschränkt, so gilt

$$\|Ax - Ay\| = \|A(x - y)\| \leq \|A\| \|x - y\|$$

woraus folgt, dass $Ay \to Ax$ für $y \to x$. Somit ist A stetig und sogar gleichmäßig stetig.

Ist A stetig, so ergibt die Stetigkeit an $0 \in X$ folgendes: für jedes $\varepsilon > 0$ existiert δ mit

$$\|x\| \leq \delta \Rightarrow \|Ax\| \leq \varepsilon.$$

Da für jedes $x \in X$ mit $\|x\| = 1$ gilt

$$\|\delta x\| = \delta,$$

so erhalten wir

$$\|A(\delta x)\| \leq \varepsilon$$

woraus folgt

$$\|Ax\| \leq \frac{\varepsilon}{\delta}$$

und somit nach (3.4) $\|A\| \leq \frac{\varepsilon}{\delta} < \infty$. ■

Beispiel. Alle Operatoren $A : \mathbb{R}^m \to \mathbb{R}^n$ sind beschränkt, da die Komponenten von A lineare Funktionen auf $\mathbb{R}^n$ sind, die immer stetig sind. Es gilt auch folgendes: ist X endlichdimensional (wobei Y beliebig ist), so sind alle Operatoren aus $\mathcal{L}(X,Y)$ beschränkt.

Beispiel. Sei P ein Projektor im Hilbertraum H. Wir wissen, dass $\|Px\| \leq \|x\|$ für alle $x \in H$, woraus folgt $\|P\| \leq 1$. Ist $P \neq 0$, so gibt es nicht-Null Vektor $x \in \operatorname{im} P$. Da $Px = x$, so erhalten wir in diesem Fall, dass $\|P\| = 1$.

Beispiel. Betrachten wir die normierten Vektorräume $X = C^1[0,2\pi]$ und $Y = C[0,2\pi]$, die beiden mit der sup-Norm. Betrachten wie die Abbildung

$$\begin{aligned} D &: X \to Y \\ Df &= f'. \end{aligned}$$

Diese Abbildung ist linear, aber unbeschränkt, d.h. $\|D\| = \infty$. In der Tat, für die Funktion $f(x) = \sin nx \in X$ mit $n \in \mathbb{N}$ haben wir

$$\|f\| = \sup_{[0,2\pi]} |f| = 1 \quad \text{und} \quad \|Df\| = \sup_{[0,2\pi]} |f'| = \sup_{[0,2\pi]} |n \cos nx| = n,$$

so dass der Quotient $\frac{\|Df\|}{\|f\|} = n$ unbeschränkt ist und somit $\|D\| = \infty$.

Satz 3.2 *Ist Y ein Banachraum, so ist $\mathcal{B}(X,Y)$ auch Banachraum.*

Beweis. Die Konvergenz in $\mathcal{B}(X,Y)$ wird als die Konvergenz bezüglich der Operatornorm definiert, d.h.

$$A_n \to A \Leftrightarrow \|A_n - A\| \to 0.$$

Diese Konvergenz heißt auch *gleichmäßige* Konvergenz von Operatoren.

Wir beweisen, dass jede Cauchy-Folge $\{A_n\}$ in $\mathcal{B}(X,Y)$ einen Grenzwert $A \in \mathcal{B}(X,Y)$ besitzt. Für jedes $x \in X$ ist die Folge $\{A_n x\}$ eine Cauchy-Folge in Y da

$$\|A_n x - A_m x\| \leq \|A_n - A_m\| \, \|x\|$$

und $\|A_n - A_m\| \to 0$ für $n, m \to \infty$. Somit konvergiert $A_n x$ in Y und wir definieren einen Operator $A : X \to Y$ mit

$$Ax = \lim_{n\to\infty} A_n x \quad \forall x \in X. \tag{3.5}$$

Offensichtlich ist A ist linear. Die Konvergenz in (3.5) ist eine punktweise Konvergenz, die auch *starke Konvergenz* von Operatoren heißt. Trotz des Names ist die starke Konvergenz schwacher als gleichmäßige Konvergenz, so wir müssen noch beweisen, dass

$$\|A_n - A\| \to 0.$$

Daraus wird es nach der Dreiecksungleichung folgen, dass $\|A\| < \infty$ so dass $A \in \mathcal{B}(X, Y)$.

Da $\{A_n\}$ eine Cauchy-Folge ist, so reicht es zu zeigen, dass eine Teilfolge von $\{A_n\}$ gegen A gleichmäßig konvergiert (siehe Aufgabe 16). Wählen wir eine Teilfolge $\{A_{n_k}\}$ mit $\|A_{n_k} - A_{n_{k+1}}\| < 2^{-k}$ und benennen sie in $\{A_k\}$ um. Dann gilt für alle k

$$\|A_k - A_{k+1}\| < 2^{-k}.$$

Es folgt aus (3.5) dass für alle $x \in X$

$$\begin{aligned}(A - A_k)\, x &= Ax - A_k x \\ &= \lim_{n\to\infty} (A_n x - A_k x) \\ &= \lim_{n\to\infty} \sum_{j=k}^{n-1} (A_{j+1}x - A_j x) \\ &= \sum_{j=k}^{\infty} (A_{j+1}x - A_j x)\end{aligned}$$

und somit

$$\begin{aligned}\|(A - A_k)\, x\| &\le \sum_{j=k}^{\infty} \|A_{j+1}x - A_j x\| \\ &\le \sum_{j=k}^{\infty} \|A_{j+1} - A_j\| \, \|x\| \\ &\le \sum_{j=k}^{\infty} 2^{-j} \, \|x\| \\ &= 2^{-k+1} \, \|x\|,\end{aligned}$$

woraus folgt

$$\|A - A_k\| = \sup_{x \in X \setminus \{0\}} \frac{\|(A - A_k)\, x\|}{\|x\|} \le 2^{-k+1} \to 0 \quad \text{für} \quad k \to \infty.$$

■

Im Fall $Y = X$ bezeichnen wir $\mathcal{L}(X) = \mathcal{L}(X, X)$ und $\mathcal{B}(X) = \mathcal{B}(X, X)$.

Definition. Im Vektorraum $\mathcal{L}(X)$ definieren wir die Multiplikation von Operatoren als ihre Komposition, d.h. für Operatoren $A, B \in \mathcal{L}(X)$ definieren wir das Produkt $AB : X \to X$ mit

$$(AB)\, x = A(Bx) \quad \forall x \in X.$$

Es ist klar, dass $AB \in \mathcal{L}(X)$. Die Multiplikation ist nicht kommutativ, aber assoziativ und besitzt ein Einheitselement: die identische Abbildung $I : X \to X$, die auch die Identität heißt. Die Multiplikation ist offensichtlich bilinear bezüglich der linearen Operationen in $\mathcal{L}(X)$, d.h.

$$(A + B)\, C = AC + BC \quad \text{und} \quad C(A + B) = CA + CB. \tag{3.6}$$

Zum Beispiel, für jedes $x \in X$

$$((A + B)\, C)\, x = (A + B)(Cx) = A(Cx) + B(Cx) = AC(x) + BC(x) = (AC + BC)\, x,$$

was die erste Identität in (3.6) beweist. In anderen Wörtern ist die Multiplikation von Operatoren distributiv bezüglich der Addition.

Die Multiplikation von Operatoren ist *submultiplikativ* bezüglich der Operatornorm, d.h. für alle $A, B \in \mathcal{L}(X)$,

$$\|AB\| \leq \|A\| \, \|B\|, \tag{3.7}$$

da für jedes $x \in X$

$$\|(AB)\, x\| = \|A\,(Bx)\| \leq \|A\| \, \|Bx\| \leq \|A\| \, \|B\| \, \|x\|.$$

Daraus folgt, dass das Produkt von Operatoren aus $\mathcal{B}(X)$ auch in $\mathcal{B}(X)$ liegt. Somit ist $\mathcal{B}(X)$ *eine normierte Algebra*, d.h. ein Vektorraum mit Multiplikation die assoziativ und distributiv ist, besitzt ein Einheitselement und erfüllt (3.7). Ist X ein Banachraum, so ist $\mathcal{B}(X)$ eine *Banachalgebra*.

Für jeden Operator $A \in \mathcal{B}(X)$ (bzw $\mathcal{L}(X)$) ist auch die Potenz A^n für alle $n \in \mathbb{N}$ definiert:

$$A^n = \underbrace{A...A}_{n} \text{ und } A^0 = I.$$

Für jedes Polynom $P(z) = \sum_{k=0}^{n} c_k z^k$ mit $c_k \in \mathbb{K}$ definieren wir den Operator $P(A)$ mit

$$P(A) = \sum_{k=0}^{n} c_k A^k.$$

Es ist klar, dass für zwei PolynomeP, Q die folgenden Identitäten gelten:

$$\begin{aligned} (P+Q)(A) &= P(A) + Q(A) \\ (PQ)(A) &= P(A)\, Q(A). \end{aligned}$$

3.2 Integraloperatoren

Sei X eine beschränkte messbare Teilmenge von $\mathbb{R}^n$. Betrachten wir den Hilbertraum

$$H = L^2(X, \lambda_n)$$

(über $\mathbb{R}$ oder $\mathbb{C}$) und den folgenden *Integraloperator* K auf Funktionen $f : X \to \mathbb{K}$:

$$(Kf)(x) = \int_X k(x,y)\, f(y)\, d\lambda_n(y), \tag{3.8}$$

wobei $k(x,y)$ eine gegebene messbare Funktion auf $X \times X$ ist (reellwertig bzw komplexwertig). Die Funktion $k(x,y)$ heißt der *Kern* des Operators K. Unter bestimmten Bedingungen über k beweist man, dass K wohldefiniert ist und sogar ein beschränkter Operator in H ist.

Satz 3.3 *Nehmen wir an, dass der Kern k die folgende Bedingung erfüllt:*

$$\int_{X \times X} |k(x,y)|^2 \, d\lambda_{2n}(x,y) < \infty, \tag{3.9}$$

d.h. $k \in L^2(X \times X, \lambda_{2n})$. Dann gilt $Kf \in H$ für jedes $f \in H$. Darüber hinaus ist der Operator $K : H \to H$ linear, beschränkt und es gilt

$$\|K\| \leq C := \left(\int_{X \times X} |k(x,y)|^2 \, d\lambda_{2n}(x,y) \right)^{1/2}.$$

Erfüllt der Kern k eines Integraloperators K die Bedingung (3.9) so heißt K ein *Hilbert-Schmidt Operator*.

Beweis. Zuerst zeigen wir, dass $Kf(x)$ für fast alle x definiert ist, d.h. die Funktion $k(x,y)f(y)$ für fast alle x integrierbar bezüglich y ist. Diese Funktion ist messbar auf $X \times X$. Für jedes $x \in X$ erhalten wir nach dem Cauchy-Schwarz Ungleichung, dass

$$\left(\int_X |k(x,y) f(y)| \, d\lambda_n(y)\right)^2 \leq \int_X |k(x,y)|^2 \, d\lambda_n(y) \int_X |f|^2 \, d\lambda_n.$$

Für fast alle x gilt

$$\int_X |k(x,y)|^2 \, d\lambda_n(y) < \infty$$

nach der Voraussetzung (3.9) und dem Satz von Fubini. Es folgt dass die Funktion

$$y \mapsto k(x,y) f(y)$$

integrierbar für fast alle x ist und somit $Kf(x)$ für fast alle x definiert ist.

Es folgt aus (3.8) dass

$$|(Kf)(x)|^2 \leq \int_X |k(x,y)|^2 \, d\lambda_n(y) \int_X |f|^2 \, d\lambda_n = \|f\|_2^2 \int_X |k(x,y)|^2 \, d\lambda_n(y),$$

und somit

$$\begin{aligned} \|Kf\|_2^2 &= \int_X |Kf(x)|^2 \, d\lambda_n(x) \leq \int_X \left(\|f\|_2^2 \int_X |k(x,y)|^2 \, d\lambda_n(y)\right) d\lambda_n(x) \\ &= \|f\|_2^2 \int_{X \times X} |k(x,y)|^2 \, d\lambda_{2n} = C^2 \|f\|_2^2. \end{aligned} \tag{3.10}$$

Somit gilt $\|Kf\|_2 < \infty$ und $Kf \in L^2(X, \lambda_n)$. Es folgt aus (3.10), dass $\|K\| \leq C < \infty$ so dass K beschränkt ist. ■

Noch eine Bedingung für die Beschränktheit von K ist in der Aufgabe 25 gegeben: gilt

$$M := \max\left(\sup_{x \in X} \int_X |k(x,y)| \, d\lambda_n(y), \ \sup_{y \in X} \int_X |k(x,y)| \, d\lambda_n(x)\right) < \infty$$

so ist der Operator $K : H \to H$ wohldefiniert und

$$\|K\| \leq M. \tag{3.11}$$

3.3 Adjungierter Operator

Sei H ein Hilbertraum.

Definition. Seien $A, B \in \mathcal{B}(H)$. Die Operatoren A, B heißen *adjungiert* wenn für alle $x, y \in H$ gilt

$$(Ax, y) = (x, By). \tag{3.12}$$

Diese Definition ist symmetrisch bezüglich A und B da (3.12) äquivalent zu

$$(Bx, y) = (x, Ay)$$

ist. Der Operator B heißt der *adjungierte Operator* von A (und A heißt der adjungierte Operator von B).

Satz 3.4 *Für jeden Operator $A \in \mathcal{B}(H)$ gibt es genau einen adjungierten Operator $B \in \mathcal{B}(H)$, und es gilt $\|B\| = \|A\|$.*

Beweis. Beweisen wir zuerst die folgende Aussage: gilt für $u, v \in H$ die Identität

$$(x, u) = (x, v) \quad \text{für alle } x \in H, \tag{3.13}$$

so gilt $u = v$. In der Tat, es folgt aus (3.13), dass $(x, u - v) = 0$ für all $x \in H$. Insbesondere für $x = u - v$ erhalten wir $\|u - v\| = 0$ und somit $u = v$.

Es folgt, dass der adjungierte Operator eindeutig bestimmt ist, da für zwei Operatoren B_1 und B_2, die (3.12) erfüllen, gilt

$$(x, B_1 y) = (x, B_2 y) \quad \forall x, y \in H,$$

woraus folgt $B_1 y = B_2 y$ und somit $B_1 = B_2$.

Jetzt beweisen wir die Existenz von einem Operator B, der (3.12) erfüllt. Für jedes $y \in H$ betrachten wir das Funktional

$$\begin{aligned} F_y &: H \to \mathbb{K} \\ F_y(x) &= (Ax, y). \end{aligned}$$

Offensichtlich ist F_y linear, und auch beschränkt, da für alle $x \in H$

$$|F_y(x)| = |(Ax, y)| \leq \|Ax\| \|y\| \leq (\|A\| \|x\|) \|y\| = (\|A\| \|y\|) \|x\|. \tag{3.14}$$

Somit ist F_y auch stetig nach dem Satz 3.1. Nach dem Satz 2.7 (Rieszscher Darstellungssatz) gibt es ein eindeutiges Element $v \in H$ mit

$$F_y(x) = (x, v) \quad \forall x \in H.$$

Dann definieren wir die Abbildung $B : H \to H$ mit $By = v$ (da v nur von y abhängig ist). Somit erhalten wir die Identität

$$(Ax, y) = F_y(x) = (x, By) \quad \forall x, y \in H,$$

d.h. die Identität (3.12). Es bleibt noch zu zeigen, dass B linear und beschränkt ist.

Wir haben für alle $x, y_1, y_2 \in H$

$$(x, B(y_1 + y_2)) = (Ax, y_1 + y_2) = (Ax, y_1) + (Ax, y_2) = (x, By_1) + (x, By_2) = (x, By_1 + By_2),$$

woraus folgt

$$B(y_1 + y_2) = By_1 + By_2.$$

Analog beweist man, dass $B(\alpha y) = \alpha By$, so dass B linear ist.

Es folgt aus (3.12) und (3.14), dass für alle $x, y \in H$

$$|(x, By)| = |(Ax, y)| \leq \|A\| \|x\| \|y\|$$

woraus folgt, dass für $x = By$

$$\|By\|^2 \leq \|A\| \|By\| \|y\|$$

und somit

$$\|By\| \leq \|A\| \|y\|$$

und $\|B\| \leq \|A\|$. Insbesondere ist B beschränkt.

Analog erhalten wir $\|A\| \leq \|B\|$, woraus folgt $\|A\| = \|B\|$. ∎

01.06.22 Vorlesung 15

Der adjungierte Operator von A wird mit A^* bezeichnet. Dann wird A^* mit der Identität

$$(Ax, y) = (x, A^*y) \quad \text{für alle } x, y \in H$$

eindeutig bestimmt.

Beispiel. Für $A = I$ gilt

$$(Ix, y) = (x, y) = (x, Iy)$$

so dass $I^* = I$.

Beispiel. Jede Matrix $A = (a_{ij})_{i,j=1}^n$ mit $a_{ij} \in \mathbb{R}$ bestimmt einen Operator

$$\begin{aligned} A : \mathbb{R}^n &\to \mathbb{R}^n \\ x &\mapsto Ax \end{aligned}$$

wobei Ax das Produkt der Matrix A und des Spaltenvektors x ist. Wir haben

$$(Ax, y) = \sum_{i=1}^n (Ax)_i \, y_i = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_j y_i.$$

Austauschen von i und j ergibt

$$(Ax, y) = \sum_{i=1}^n \sum_{j=1}^n x_i a_{ji} y_j = \sum_{i=1}^n x_i \left(A^T y\right)_i = \left(x, A^T y\right),$$

woraus folgt $A^* = A^T$.

Im Fall $a_{ij} \in \mathbb{C}$ bestimmt die Matrix A einen Operator in $\mathbb{C}^n$. Wir benutzen die Definition von dem Skalarprodukt in $\mathbb{C}^n$ und erhalten analog

$$\begin{aligned} (Ax, y) &= \sum_{i=1}^n (Ax)_i \, \overline{y_i} = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_j \overline{y}_i \\ &= \sum_{i=1}^n \sum_{j=1}^n x_i a_{ji} \overline{y}_j = \sum_{i=1}^n x_i \left(A^T \overline{y}\right)_i = \sum_{i=1}^n x_i \overline{(\overline{A}^T y)_i} = (x, \overline{A}^T y) \end{aligned}$$

und somit $A^* = \overline{A}^T$.

Beispiel. Sei X eine beschränkte messbare Teilmenge von $\mathbb{R}^n$ (z.B. ein Intervall in $\mathbb{R}$). Sei K ein Integraloperator in $H = L^2(X, \lambda_n)$ mit dem Kern $k(x,y)$:

$$Kf(x) = \int_X k(x,y) f(y) \, d\lambda_n(y).$$

und nehmen wir an, dass $k(x,y)$ die Bedingung (3.9) erfüllt, d.h.

$$\int_{X\times X} |k(x,y)|^2 \, d\lambda_{2n} < \infty$$

Nach dem Satz 3.3 ist K beschränkt. Bestimmen wir K^*. Zuerst bemerken wir, dass alle $f, g \in L^2(X, \lambda_n)$ nach der Cauchy-Schwarz Ungleichung gilt

$$\begin{aligned}\left(\int_{X\times X} \left|k(x,y) f(y) \overline{g(x)}\right| d\lambda_{2n}\right)^2 &\leq \int_{X\times X} |k(x,y)|^2 \, d\lambda_{2n} \int_{X\times X} \left|f(y) \overline{g(x)}\right|^2 d\lambda_{2n} \\ &= \left(\int_{X\times X} |k(x,y)|^2 \, d\lambda_{2n}\right) \|f\|_2^2 \|g\|_2^2 < \infty,\end{aligned}$$

so dass die Funktion $k(x,y) f(y) \overline{g(x)}$ auf $X \times X$ integrierbar ist. Somit erhalten wir nach dem Satz von Fubini

$$\begin{aligned}(Kf, g) &= \int_X \left(\int_X k(x,y) f(y) \, d\lambda_n(y)\right) \overline{g(x)} d\lambda_n(x) \\ &= \int_{X\times X} k(x,y) f(y) \overline{g(x)} d\lambda_{2n}(x,y) \\ &= \int_X \left(\int_X k(x,y) \overline{g(x)} d\lambda_n(x)\right) f(y) \, d\lambda_n(y) \\ &= \int_X f(x) \overline{\left(\int_X \overline{k(y,x)} g(y) \, d\lambda_n(y)\right)} d\lambda_n(x) \\ &= (f, K^* g),\end{aligned}$$

wobei

$$K^* g(x) = \int_X \overline{k(y,x)} g(y) \, d\lambda_n(y).$$

Somit ist K^* auch ein Integraloperator und zwar mit dem Kern

$$k^*(x,y) = \overline{k(y,x)}.$$

Im nächsten Satz werden die wichtigsten Eigenschaften der Adjunktion aufgelistet.

Satz 3.5 *Für Operatoren $A, B \in \mathcal{B}(H)$ gelten die folgenden Eigenschaften.*

1. $(A+B)^* = A^* + B^*$
2. $(\alpha A)^* = \overline{\alpha} A^*$ *für* $\alpha \in \mathbb{K}$
3. $(A^*)^* = A$

4. $(AB)^* = B^*A^*$

5. $\|A^*\| = \|A\|$.

Beweis. 1. Für alle $x, y \in H$ haben wir nach Definition von adjungiertem Operator

$$((A+B)\,x, y) = (Ax, y) + (Bx, y) = (x, A^*y) + (x, B^*y) = (x, (A^* + B^*)\,y)\,.$$

Da auch

$$((A+B)\,x, y) = (x, (A+B)^*\,y)\,,$$

so erhalten wir $(A+B)^* = A^* + B^*$ nach der Eindeutigkeit von adjungiertem Operator.

2. Wir haben

$$((\alpha A)\,x, y) = \alpha\,(Ax, y) = \alpha\,(x, A^*y) = (x, \overline{\alpha}A^*y)\,.$$

Da auch

$$((\alpha A)\,x, y) = (x, (\alpha A)^*\,y)\,,$$

so erhalten wir $(\alpha A)^* = \overline{\alpha}A^*$.

3. As der Identität

$$(Ay, x) = (y, A^*x)$$

erhalten wir mit Hilfe vom komplexer Konjugation, dass

$$(A^*x, y) = (x, Ay)\,, \tag{3.15}$$

Da auch

$$(A^*x, y) = (x, (A^*)^*\,y)$$

so erhalten wir $(A^*)^* = A$.

4. Für alle x, y gilt nach (3.15)

$$((AB)^*\,x, y) = (x, ABy) = (x, A\,(By)) = (A^*x, By) = (B^*A^*x, y)$$

woraus $(AB)^* = B^*A^*$ folgt.

5. Die Identität $\|A^*\| = \|A\|$ haben wir schon im Satz 3.4 bewiesen. ■

Satz 3.6 *Für jeden Operator $A \in \mathcal{B}(H)$ gelten die Identitäten:*

$$\ker A = (\operatorname{im} A^*)^\perp \qquad \textit{und} \qquad \overline{\operatorname{im} A} = (\ker A^*)^\perp$$

wobei $\overline{U}$ den Abschluss des Unterraums U bezeichnet.

Beweis. Wir haben

$$\begin{aligned} x \in (\operatorname{im} A^*)^\perp &\Leftrightarrow x \perp A^*y \quad \forall y \in H \\ &\Leftrightarrow (x, A^*y) = 0 \quad \forall y \in H \\ &\Leftrightarrow (Ax, y) = 0 \quad \forall y \in H \\ &\Leftrightarrow Ax = 0 \\ &\Leftrightarrow x \in \ker A, \end{aligned}$$

und somit $\ker A = (\operatorname{im} A^*)^\perp$. Anwendung von dieser Identität auf A^* anstatt A ergibt

$$\ker A^* = (\operatorname{im} A)^\perp ,$$

woraus folgt

$$(\ker A^*)^\perp = (\operatorname{im} A)^{\perp\perp} .$$

Nach Korollar 2.6 gilt für abgeschlossenen Unterraum U

$$U^{\perp\perp} = U.$$

When U nicht unbedingt abgeschlossen ist, so folgt es daraus, dass

$$U^{\perp\perp} = \overline{U}^{\perp\perp} = \overline{U}.$$

Für $U = \operatorname{im} A$ erhalten wir

$$(\operatorname{im} A)^{\perp\perp} = \overline{\operatorname{im} A}$$

und somit $(\ker A^*)^\perp = \overline{\operatorname{im} A}$. ■

3.4 Inverser Operator

Seien X ein normierter Vektorraum und $A \in \mathcal{L}(X)$. Der Operator A hat den inversen Operator $A^{-1} \in \mathcal{L}(X)$ genau dann, wenn A injektiv und surjektiv ist, d.h. wenn

$$\ker A = \{0\} \quad \text{und} \quad \operatorname{im} A = X. \tag{3.16}$$

Dann ist A^{-1} auch linear. Ist der Operator A^{-1} beschränkt, so sagen wir, dass Operator A einen *beschränkten inversen Operator* hat und schreiben $A^{-1} \in \mathcal{B}(X)$. Diese Schreibweise bedeutet sowohl die Existenz als auch die Beschränktheit des inversen Operators.

Der nächste Hilfssatz liefert die bequemen Bedingungen für die Existenz und Beschränktheit von A^{-1}, die häufig benutzt wird.

Satz 3.7 *Ein Operator $A \in \mathcal{B}(H)$ hat einen beschränkten inversen Operator genau dann, wenn die folgenden zwei Bedingungen erfüllt sind:*

(a) $\ker A^* = \{0\}$

(b) *und* $\|Ax\| \geq c\,\|x\|$ *für ein* $c > 0$ *und alle* $x \in H$.

Existiert $A^{-1} \in \mathcal{B}(H)$, so existiert auch $(A^*)^{-1} \in \mathcal{B}(H)$ und es gilt $(A^{-1})^* = (A^*)^{-1}$.

Beweis. Beweisen wir zuerst, dass

$$A^{-1} \in \mathcal{B}(H) \Rightarrow (a) \text{ und } (b).$$

Wir haben nach dem Satz 3.6 und (3.16)

$$\ker A^* = (\operatorname{im} A)^\perp = H^\perp = \{0\} ,$$

d.h. die Bedingung (a).

Da A^{-1} beschränkt ist, so setzen wir $C = \|A^{-1}\|$ so dass für alle $y \in H$

$$\left\|A^{-1}y\right\| \leq C \left\|y\right\|.$$

Für $y = Ax$ erhalten wir

$$\|x\| \leq C \|Ax\|$$

woraus (b) mit $c = C^{-1}$ folgt.

Jetzt beweisen wir, dass

$$(a) \text{ und } (b) \Rightarrow A^{-1} \in \mathcal{B}(H).$$

Für die Existenz von A^{-1} überprüfen wir die Bedingungen (3.16), d.h.

$$\ker A = \{0\} \quad \text{und} \quad \operatorname{im} A = H.$$

Für jedes $x \in \ker A$ haben wir $Ax = 0$ und nach (b) auch $x = 0$, so dass $\ker A = \{0\}$. Nach (a) und Satz 3.6 haben wir

$$\overline{\operatorname{im} A} = (\ker A^*)^{\perp} = H.$$

Um zu beweisen, dass $\operatorname{im} A = H$, reicht es zu zeigen, dass $\operatorname{im} A$ abgeschlossen ist. Dafür betrachten wir eine Folge $\{y_n\} \subset \operatorname{im} A$ die gegen ein $y \in H$ konvergiert, und beweisen, dass $y \in \operatorname{im} A$. Da $y_n \in \operatorname{im} A$ so gilt $y_n = Ax_n$ für ein $x_n \in H$. Nach (b) gilt

$$\|x_n - x_m\| \leq c^{-1} \|Ax_n - Ax_m\| = c^{-1} \|y_n - y_m\| \to 0 \text{ für } n, m \to \infty,$$

so dass $\{x_n\}$ eine Cauchy-Folge ist. Somit konvergiert auch die Folge $\{x_n\}$. Setzen wir $x = \lim x_n$ und erhalten nach der Stetigkeit von A dass

$$Ax = \lim Ax_n = \lim y_n = y,$$

woraus folgt, dass $y \in \operatorname{im} A$ und dass $\operatorname{im} A$ abgeschlossen ist.

Wir beschließen, dass die Bedingungen (3.16) erfüllt sind und der inverse Operator A^{-1} existiert. Es bleibt noch zu zeigen, dass $\|A^{-1}\| < \infty$. In (b) setzen wir $y = Ax$ und erhalten

$$\|y\| \geq c \left\|A^{-1}y\right\|.$$

Da $\operatorname{im} A = H$, so ist y hier beliebig, woraus folgt $\|A^{-1}\| \leq c^{-1} < \infty$.

Jetzt beweisen wir die zweite Aussage:

$$A^{-1} \in \mathcal{B}(H) \Rightarrow (A^*)^{-1} \in \mathcal{B}(H) \quad \text{und} \quad \left(A^{-1}\right)^* = (A^*)^{-1}.$$

Aus der Identität

$$AA^{-1} = A^{-1}A = I$$

erhalten wir nach dem Satz 3.5

$$I = I^* = \left(AA^{-1}\right)^* = \left(A^{-1}\right)^* A^*$$

und analog

$$I = A^* \left(A^{-1}\right)^*.$$

Es folgt, dass der inverse von A^* existiert und gleich $(A^{-1})^*$ ist, d.h. $(A^*)^{-1} = (A^{-1})^*$. Daraus folgt, dass $(A^*)^{-1} \in \mathcal{B}(H)$. ■

3.5 Spektrum eines Operators

Sei X ein normierter Vektorraum über $\mathbb{C}$.

Definition. Für jeden Operator $A \in \mathcal{L}(X)$ definieren wir die *Resolventenmenge* $\rho(A)$ mit

$$\rho(A) = \left\{\lambda \in \mathbb{C} : (A - \lambda I)^{-1} \in \mathcal{B}(X)\right\}.$$

Der Operator $(A - \lambda I)^{-1}$ heißt die *Resolvente* von A.

Das *Spektrum* $\sigma(A)$ von A ist das Komplement von $\rho(A)$ in $\mathbb{C}$:

$$\begin{aligned}\sigma(A) &= \mathbb{C} \setminus \rho(A) = \\ &\left\{\lambda \in \mathbb{C} : (A - \lambda I)^{-1} \text{ entweder existiert nicht oder existiert aber ist unbeschränkt}\right\}.\end{aligned}$$

Sei X ein normierter Vektorraum über $\mathbb{R}$. Um das Spektrum von $A \in \mathcal{L}(X)$ in diesem Fall zu definieren, betrachten wir die Komplexifizierung von X:

$$X_{\mathbb{C}} = \{x + iy : x, y \in X\},$$

die ein normierter Vektorraum über $\mathbb{C}$ ist, und den komplexifizierten Operator

$$A_{\mathbb{C}}(x + iy) = Ax + iAy \ \text{ für alle } x, y \in X,$$

so dass $A_C \in \mathcal{L}(X_{\mathbb{C}})$ (siehe Aufgaben 6, 40, 67). Dann setzen wir

$$\sigma(A) := \sigma(A_{\mathbb{C}}).$$

Sei X wieder ein normierter Vektorraum über $\mathbb{C}$. Es folgt aus der Definition, dass $\lambda \in \rho(A)$ genau dann gilt, wenn

- $\ker(A - \lambda I) = \{0\}$
- und $\operatorname{im}(A - \lambda I) = X$
- und $(A - \lambda I)^{-1}$ ist beschränkt.

Entsprechend gilt $\lambda \in \sigma(A)$ genau dann, wenn

- entweder $\ker(A - \lambda I) \neq \{0\}$
- oder $\operatorname{im}(A - \lambda I) \neq X$
- oder $(A - \lambda I)^{-1}$ existiert aber ist unbeschränkt.

Definition. Gilt $\ker(A - \lambda I) \neq \{0\}$, so heißt λ ein *Eigenwert* von A. Der Unterraum $\ker(A - \lambda I)$ von X heißt der *Eigenraum* von A mit dem Eigenwert λ. Jeder nicht-Null Vektor $v \in \ker(A - \lambda I)$ heißt ein *Eigenvektor* von A mit dem Eigenwert λ. Die Bedingung $v \in \ker(A - \lambda I)$ ist offensichtlich äquivalent zu $Av = \lambda v$. Der Wert $m = \dim\ker(A - \lambda I)$ heißt die *Vielfachheit* des Eigenwertes λ.

03.06.22 Vorlesung 16

Die Menge von allen Eigenwerten von A wird mit $\sigma_p(A)$ bezeichnet und heißt das *Punktspektrum* von A. Offensichtlich ist $\sigma_p(A)$ eine Teilmenge von $\sigma(A)$.

Beispiel. Sei A eine $n \times n$ Matrix mit komplexwertigen Koeffizienten. Wir betrachten A als einen Operator in $\mathbb{C}^n$. Der Operator $A - \lambda I$ ist invertierbar (und dann ist $(A - \lambda I)^{-1}$ automatisch beschränkt) genau dann, wenn $\det(A - \lambda I) \neq 0$, woraus folgt

$$\lambda \in \sigma(A) \Leftrightarrow \det(A - \lambda I) = 0.$$

Die Funktion $\lambda \mapsto \det(A - \lambda I)$ ist ein Polynom von Grad n, das heißt das *charakteristische Polynom* von A. Somit stimmt das Spektrum $\sigma(A)$ mit der Menge von den Nullstellen des Polynoms P überein, und diese Menge ist nichtleer nach dem Fundamentalsatz der Algebra.

Andererseits ist die Bedingung $\det(A - \lambda I) = 0$ äquivalent zur Existenz eines Spaltenvektors $v \neq 0$ mit $(A - \lambda I)v = 0$, was bedeutet, dass λ ein Eigenwert der Matrix A ist. Insbesondere erhalten wir $\sigma_p(A) = \sigma(A)$. Wir werden später sehen, dass diese Identität für die unendlichdimensionalen Operatoren nicht unbedingt der Fall ist, d.h. $\sigma_p(A)$ eine echte Teilmenge von $\sigma(A)$ sein kann.

Beispiel. Sei K ein kompakter metrischer Raum. Betrachten wir den Banachraum $X = C(K)$ von stetigen komplexwertigen Funktionen auf K. Definieren wir einen Operator $A : X \to X$ mit

$$Af(t) = a(t) f(t) \quad \forall t \in K \tag{3.17}$$

wobei $a(t)$ eine gegebene stetige Funktion auf K ist. Es gilt offensichtlich

$$\|Af\| = \sup_{t \in K} |a(t) f(t)| \leq \sup |a| \sup |f| = C \|f\|$$

wobei $C = \sup |a| < \infty$. Somit ist der Operator A beschränkt. Wir zeigen, dass

$$\sigma(A) = a(K) := \{a(t) : t \in K\}. \tag{3.18}$$

Für jedes $\lambda \notin a(K)$ ist die Funktion $b(t) = \frac{1}{a(t) - \lambda}$ wohldefiniert und stetig auf K. Definieren wir den Operator $B \in \mathcal{B}(X)$ mit

$$Bf(t) = b(t) f(t).$$

Dann gilt für alle $f \in X$

$$(A - \lambda I) Bf = (a - \lambda) bf = f \quad \text{und} \quad B(A - \lambda I) f = f.$$

Wir beschließen dass

$$(A - \lambda I) B = B(A - \lambda I) = I$$

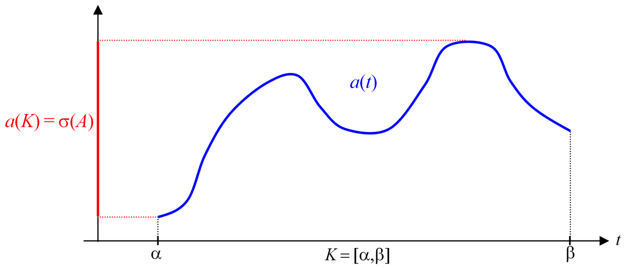
und somit $B = (A - \lambda I)^{-1}$, woraus folgt $\lambda \notin \sigma(A)$.

Sei $\lambda \in a(K)$, z.B. $\lambda = a(t_0)$ für ein $t_0 \in K$. Dann gilt für jede Funktion $f \in C(K)$

$$(A - \lambda I) f(t_0) = (a(t_0) - \lambda) f(t_0) = 0.$$

Deshalb verschwinden an der Stelle t_0 alle Funktionen aus dem Bildraum $\operatorname{im}(A - \lambda I)$, woraus folgt, dass $1 \notin \operatorname{im}(A - \lambda I)$ und $\operatorname{im}(A - \lambda I) \neq X$. Somit ist $A - \lambda I$ nicht invertierbar ist, d.h. $\lambda \in \sigma(A)$.

Seien $K = [\alpha, \beta]$ und a eine stetige Funktion auf $[\alpha, \beta]$ die auf keinem offenen Teilintervall von $[\alpha, \beta]$ Konstante ist.



In diesem Fall gibt es keinen Eigenwert von A, d.h. $\sigma_p(A) = \emptyset$, obwohl $\sigma(A)$ nichtleer ist. In der Tat, ist λ ein Eigenwert, so gilt für seine Eigenfunktion $f \in X \setminus \{0\}$, dass $a(t) f(t) = \lambda f(t)$, woraus folgt

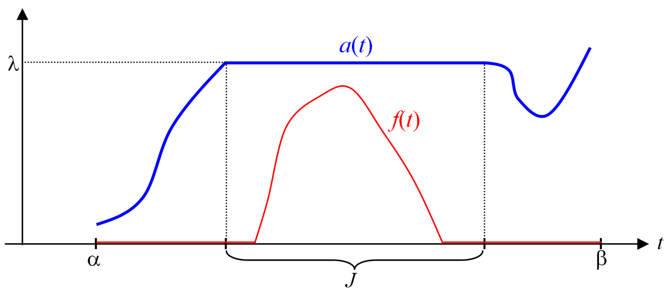
$$f(t) \neq 0 \Rightarrow a(t) = \lambda.$$

Somit ist $a(t)$ gleich Konstante auf der Menge $\{f \neq 0\}$, die nicht-leer und offen ist, und insbesondere auf einem offenen Intervall, was im Widerspruch zur Voraussetzung steht.

Ist $a(t)$ gleich eine Konstante λ auf einem nicht-leeren offenen Intervall $J \subset [\alpha, \beta]$ so gilt

$$a(t) f(t) = \lambda f(t)$$

für jede Funktion f mit $\operatorname{supp} f \subset J$ so dass alle solche Funktionen die Eigenfunktionen von A mit dem Eigenwert λ sind.



Satz 3.8 *Seien X ein Banachraum und $A \in \mathcal{B}(X)$. Dann ist das Spektrum $\sigma(A)$ eine nichtleere abgeschlossene beschränkte Teilmenge von $\mathbb{C}$.*

Dass $\sigma(A)$ abgeschlossen und beschränkt ist, beweist man in Aufgabe 69. Dass $\sigma(A)$ nichtleer ist, nehmen wir ohne Beweis an. Wir beweisen unterhalb, dass $\sigma(A) \neq \emptyset$ im speziellen Fall wenn A ein *selbstadjungierter* Operator in Hilbertraum ist (siehe Satz 3.10).

Satz 3.9 *Seien H ein Hilbertraum und $A \in \mathcal{B}(H)$. Dann gilt $\lambda \in \rho(A)$ genau dann, wenn die folgenden zwei Bedingungen erfüllt sind:*

(a) $\ker\left(A^* - \overline{\lambda} I\right) = \{0\}$

(b) $\|Ax - \lambda x\| \geq c\|x\|$ *für ein $c > 0$ und alle $x \in H$.*

Beweis. Anwendung von dem Satz 3.7 mit dem Operator $B = A - \lambda I$ anstatt A ergibt

$$\begin{aligned}\lambda \in \rho(A) &\Leftrightarrow (A - \lambda I)^{-1} \in \mathcal{B}(H) \\ &\Leftrightarrow B^{-1} \in \mathcal{B}(H) \\ &\Leftrightarrow \ker B^* = \{0\} \text{ und } \|Bx\| \geq c\|x\| \text{ für ein } c > 0.\end{aligned}$$

Da $B^* = (A - \lambda I)^* = A^* - \overline{\lambda} I$ so erhalten wir dass $\lambda \in \rho(A) \Leftrightarrow (a)$ und (b). ∎

3.6 Selbstadjungierte Operatoren

Sei H ein Hilbertraum über $\mathbb{C}$.

Definition. Ein Operator $A \in \mathcal{B}(H)$ heißt *selbstadjungiert* wenn $A^* = A$.

Äquivalente Definition: ein Operator $A \in \mathcal{B}(H)$ ist selbstadjungiert genau dann, wenn

$$(Ax, y) = (x, Ay) \tag{3.19}$$

für alle $x, y \in H$. Die Operatoren, die (3.19) erfüllen, heißen auch symmetrisch oder Hermitesch.

Es folgt aus der Definition, dass die Menge von selbstadjungierten Operatoren ein Unterraum von $\mathcal{B}(H)$ ist. Man kann zeigen, dass dieser Unterraum abgeschlossen ist und somit auch ein Banachraum.

Beispiel. Eine $n \times n$ Matrix $A = (a_{ij})$ mit reellwertigen Koeffizienten ist selbstadjungiert als ein Operator in $\mathbb{R}^n$ genau dann, wenn $A^T = A$, d.h. wenn $a_{ij} = a_{ji}$, d.h. wenn A symmetrisch ist. Eine $n \times n$ Matrix $A = (a_{ij})$ mit komplexwertigen Koeffizienten ist selbstadjungiert als Operator in $\mathbb{C}^n$ genau dann, wenn $\overline{A^T} = A$, d.h. wenn $\overline{a_{ji}} = a_{ij}$. Solche Matrizen heißen *Hermitesch*.

Beispiel. Jeder orthogonale Projektor P im Hilbertraum ist selbstadjungiert, was aus dem Satz 2.5 folgt.

Beispiel. Sei X eine beschränkte messbare Teilmenge von $\mathbb{R}^n$. Betrachten wir den Hilbertraum $H = L^2(X, \lambda_n)$ den folgenden Integraloperator $K : H \to H$:

$$Kf(x) = \int_X k(x, y) f(y) \, d\lambda_n(y),$$

wobei der Kern $k(x, y)$ eine der Bedingungen (3.9), (3.11) erfüllt. Unter diesen Bedingungen ist K beschränkt, und K^* ist auch ein Integraloperator mit dem Kern $\overline{k(y, x)}$. Offensichtlich ist K selbstadjungiert genau dann, wenn

$$\overline{k(y, x)} = k(x, y). \tag{3.20}$$

Ein Kern $k(x, y)$ mit der Bedingung (3.20) heißt Hermitesch. Im Fall $\mathbb{K} = \mathbb{R}$ bedeutet (3.20) die Symmetry:

$$k(x, y) = k(y, x).$$

Im nächsten Satz beweisen wir die allgemeinen Eigenschaften von selbstadjungierten Operatoren.

Hauptsatz 3.10 *Sei A ein beschränkter selbstadjungierter Operator im Hilbertraum $H \neq \{0\}$. Dann gilt folgendes.*

1. $\sigma(A) \subset \mathbb{R}$.
2. $\sigma(A) \subset [-\|A\|, \|A\|]$.
3. *Mindestens einer von den Werten $\|A\|, -\|A\|$ liegt in $\sigma(A)$, insbesondere gilt*
$$\|A\| = \max |\sigma(A)|. \tag{3.21}$$
4. *Sind λ, μ zwei verschiedene Eigenwerte von A, so sind die Eigenräume* $\ker(A - \lambda I)$ *und* $\ker(A - \mu I)$ *orthogonal.*

Beweis. 1. Sei $\lambda = \alpha + i\beta$ eine komplexe Zahl mit $\beta \neq 0$. Wir beweisen, dass $\lambda \notin \sigma(A)$, d.h. $\lambda \in \rho(A)$, was äquivalent zu
$$(A - \lambda I)^{-1} \in \mathcal{B}(H).$$
Wir haben
$$A - \lambda I = A - \alpha I - i\beta I = \beta \left(\frac{A - \alpha I}{\beta} - iI\right).$$
Somit reicht es zu zeigen, dass $i \in \rho\left(\frac{A-\alpha I}{\beta}\right)$. Da der Operator $\frac{A-\alpha I}{\beta}$ selbstadjungiert ist, so können wir ihn in A umbenennen. Dann reicht es zu zeigen, dass $i \in \rho(A)$.

Nach dem Satz 3.9 reicht es die folgenden zwei Bedingungen zu überprüfen:

(*a*) $\ker\left(A^* - \bar{i}I\right) = \{0\}$, d.h. $\ker(A + iI) = \{0\}$

(*b*) $\|Ax - ix\| \geq c\|x\|$ für ein $c > 0$.

Beweis von (*a*). Sei $x \in \ker(A + iI)$, d.h. $Ax + ix = 0$. Daraus folgt
$$(Ax, x) + i(x, x) = 0. \tag{3.22}$$
Bemerken wir, dass (x, x) reell ist und (Ax, x) auch reell ist, was aus der folgenden Identität folgt:
$$(Ax, x) = (x, Ax) = \overline{(Ax, x)}.$$
Es folgt aus (3.22), dass $(x, x) = 0$ und somit $x = 0$, was beweist, dass
$$\ker(A + iI) = \{0\}.$$

Beweis von (*b*). Wir haben nach (2.6)
$$\|Ax - ix\|^2 = \|Ax\|^2 - 2\,\mathrm{Re}\,(Ax, ix) + \|ix\|^2.$$
Da (Ax, x) reell ist, so haben wir $\mathrm{Re}\,(Ax, ix) = 0$, woraus folgt
$$\|Ax - ix\|^2 = \|Ax\|^2 + \|x\|^2 \geq \|x\|^2,$$
so dass (*b*) mit $c = 1$ erfüllt ist.

2. Beweisen wir, dass jedes $\lambda \in \mathbb{R} \setminus [-\|A\|, \|A\|]$ in der Resolventenmenge $\rho(A)$ liegt. Dafür überprüfen wir die Bedingungen (*a*) und (*b*) von dem Satz 3.9, die für reellwertiges λ wie folgt aussehen:

(a) $\ker(A - \lambda I) = \{0\}$

(b) $\|Ax - \lambda x\| \geq c\,\|x\|$.

Es ist klar, dass $(b) \Rightarrow (a)$, da für jedes $x \in \ker(A - \lambda I)$ gilt $Ax - \lambda x = 0$ und dann nach (b) auch $x = 0$. Die Bedingung (b) folgt aus $|\lambda| > \|A\|$ wie folgt:

$$\|Ax - \lambda x\| \geq \|\lambda x\| - \|Ax\| \geq |\lambda|\,\|x\| - \|A\|\,\|x\| = (|\lambda| - \|A\|)\,\|x\|,$$

so dass (b) mit der Konstante $c = |\lambda| - \|A\| > 0$ erfüllt ist.

3. Gilt $\|A\| = 0$, so haben wir $A = 0$, und $\lambda = 0$ ist ein Eigenwert. Sei $\|A\| > 0$. Ohne Beschränkung der Allgemeinheit können wir annehmen, dass $\|A\| = 1$. Nehmen wir das Gegenteil an, dass die beiden Werte $\|A\| = 1$ und $-\|A\| = -1$ in $\rho(A)$ liegen, d.h.

$$(A - I)^{-1} \in \mathcal{B}(H) \quad \text{und} \quad (A + I)^{-1} \in \mathcal{B}(H).$$

Da

$$A^2 - I = (A + I)(A - I),$$

daraus folgt, dass

$$\left(A^2 - I\right)^{-1} = (A - I)^{-1}(A + I)^{-1} \in \mathcal{B}(H).$$

Somit $1 \in \rho(A^2)$ und nach dem Satz 3.9 gilt den Operator A^2 die Bedingung (b): es gibt ein $c > 0$ so dass für alle $x \in H$

$$\left\|A^2 x - x\right\| \geq c\,\|x\|. \tag{3.23}$$

Andererseits haben wir

$$\left\|A^2 x - x\right\|^2 = \left\|A^2 x\right\|^2 - 2\operatorname{Re}\left(A^2 x, x\right) + \|x\|^2. \tag{3.24}$$

Da $A = A^*$, so gilt

$$\left(A^2 x, x\right) = (Ax, Ax) = \|Ax\|^2.$$

Da $\|A\| = 1$, so haben wir $\|A^2\| \leq 1$ und somit

$$\left\|A^2 x\right\| \leq \|x\|. \tag{3.25}$$

Es folgt aus (3.23), (3.24), (3.25), dass

$$\begin{aligned} c^2\,\|x\|^2 &= \left\|A^2 x - x\right\|^2 \\ &= \left\|A^2 x\right\|^2 - 2\,\|Ax\|^2 + \|x\|^2 \\ &\leq \|x\|^2 - 2\,\|Ax\|^2 + \|x\|^2 \\ &= 2\,\|x\|^2 - 2\,\|Ax\|^2. \end{aligned}$$

Somit erhalten wir

$$2\,\|Ax\|^2 \leq \left(2 - c^2\right)\|x\|^2$$

und

$$\|Ax\|^2 \leq \left(1 - c^2/2\right)\|x\|^2,$$

woraus folgt

$$\|A\| = \sup_{x\in H\setminus\{0\}} \frac{\|Ax\|}{\|x\|} \leq \sqrt{1-c^2/2} < 1,$$

was im Widerspruch zu $\|A\| = 1$ steht.

4. Wir wissen schon, dass λ und μ reell sind. Wir müssen beweisen, dass für alle $x \in \ker(A-\lambda I)$ und $y \in \ker(A-\mu I)$ gilt $x \perp y$, vorausgesetzt $\lambda \neq \mu$. Wir haben

$$(Ax, y) = (\lambda x, y) = \lambda (x, y)$$

und

$$(Ax, y) = (x, Ay) = (x, \mu y) = \mu (x, y),$$

da $\mu \in \mathbb{R}$. Somit erhalten wir

$$\lambda (x, y) = \mu (x, y),$$

woraus $(x, y) = 0$ folgt, da $\lambda \neq \mu$. ■

Bemerkung. Es folgt aus dem Satz 3.10, dass das Spektrum $\sigma(A)$ nicht leer ist (vorausgesetzt $H \neq \{0\}$). Erinnern wir uns, dass sogar die Existenz eines Eigenwertes einer Matrix nicht-trivial ist und aus dem Fundamentalsatz der Algebra folgt.

Beweisen wir mit Hilfe von dem Satz 3.10 die folgende Aussage, die aus Linearer Algebra bekannt ist:

für jeden selbstadjungierten Operator A im Hilbertraum H einer endlichen Dimension $n \in \mathbb{N}$ gibt es eine Orthogonalbasis in H, die aus den Eigenvektoren von A besteht.

Der Beweis erfolgt per Induktion nach n. Der Induktionsanfang für $n = 1$ ist trivial. Für Induktionsschritt von $< n$ nach n, wählen wir einen Eigenwert λ von A, z.B., $\lambda = \|A\|$ oder $\lambda = -\|A\|$, und betrachten den Eigenraum $K = \ker(A - \lambda I)$. Der Unterraum $K^\perp$ ist A-invariant (siehe Aufgabe 70). Da $\dim K^\perp < n$ so beschließen wir nach der Induktionsvoraussetzung, dass es in $K^\perp$ eine Orthogonalbasis aus den Eigenvektoren von A gibt. In K gibt es auch eine Orthogonalbasis aus den Eigenvektoren von A, da alle nicht-Null Elemente von K Eigenvektoren von A sind. Die Vereinigung zweier Basen in K und in $K^\perp$ liefert eine Orthogonalbasis in H, die aus den Eigenvektoren von A besteht.

08.06.22 **Vorlesung 17**

3.7 Kompakte Operatoren

Sei X ein Banachraum über $\mathbb{K}$.

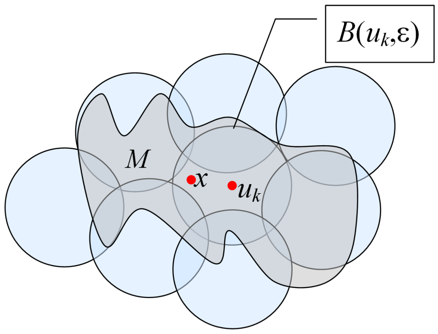
Definition. Eine Teilmenge M von X heißt *präkompakt* (oder relativ kompakt), wenn der Abschluss $\overline{M}$ kompakt ist.

Es gibt verschiedene äquivalente Definitionen von Präkompaktheit:

1. M ist präkompakt genau dann, wenn jede Folge von Elementen aus M eine konvergente Teilfolge enthält (aber der Grenzwert soll nicht unbedingt in M liegen).

2. M ist präkompakt genau dann, wenn M *totalbeschränkt* ist, d.h. für jedes $\varepsilon > 0$ existiert eine endliche Folge $\{u_k\}_{k=1}^N$ in X so dass
$$\forall x \in M \quad \exists k = 1, ..., N \quad \text{mit } \|x - u_k\| < \varepsilon. \tag{3.26}$$
Die Folge $u_1, ..., u_N$ mit dieser Eigenschaft heißt ein *ε-Netz* von M.

Die Bedingung (3.26) bedeutet, dass die Vereinigung von den Kugeln $B(u_k, \varepsilon)$ die Menge M überdeckt.



Es folgt aus der Definition, dass jede Teilmenge der präkompakten Menge ist präkompakt. Jede präkompakte Teilmenge M ist immer beschränkt, da sonst es eine Folge $\{x_n\} \subset M$ mit $\|x_n\| \to \infty$ gibt, die somit keine konvergente Teilfolge enthält.

In endlichdimensionalen Räumen gilt auch die Umkehrung: jede beschränkte Teilmenge ist präkompakt (Satz von Bolzano-Weierstraß), aber im unendlichdimensionalen Raum gilt diese Aussage nicht.

Beispiel. Sei X ein unendlichdimensionaler Hilbertraum. Beweisen wir, dass die Einheitskugel
$$Q = \{x \in X : \|x\| \leq 1\}$$
nicht präkompakt ist, obwohl Q beschränkt ist. Da $\dim X = \infty$, so gibt es eine orthonormale Folge $\{x_n\}_{n=1}^\infty$ in X (man bildet diese Folge per Induktion). Dann liegen alle x_n in Q. Allerdings enthält die Folge $\{x_n\}$ keine konvergente Teilfolge, da für alle $n \neq m$ gilt
$$\|x_n - x_m\|^2 = \|x_n\|^2 + \|x_m\|^2 = 2,$$
woraus folgt, dass keine Teilfolge von $\{x_n\}$ eine Cauchy-Folge ist.

Erinnern wir uns daran, dass ein Operator $A \in \mathcal{L}(X)$ beschränkt genau dann ist, wenn für ein $C > 0$ und alle $x \in X$ gilt
$$\|Ax\| \leq C \|x\|.$$
Die Beschränktheit von A ist äquivalent zu jeder der folgenden Bedingungen:

1. die Bildmenge $A(Q)$ ist beschränkt;

2. die Bildmenge $A(M)$ ist beschränkt für beliebige beschränkte Teilmenge $M \subset X$.

Definition. Ein Operator $A \in \mathcal{L}(X)$ heißt *kompakt*, wenn die Bildmenge $A(M)$ für beliebige beschränkte Teilmenge $M \subset X$ präkompakt ist.

Satz 3.11 *Sei $A \in \mathcal{L}(X)$.*

(i) *A ist kompakt genau dann, wenn die Bildmenge $A(Q)$ der Kugel Q präkompakt ist.*

(ii) *A ist kompakt genau dann, wenn für jede beschränkte Folge $\{x_n\}$ aus X die Folge $\{Ax_n\}$ eine konvergente Teilfolge hat.*

(iii) *Ist A kompakt so ist A auch beschränkt.*

Beweis. (i) Ist A kompakt dann ist $A(Q)$ präkompakt da Q beschränkt ist. Beweisen wir die Umkehrung $\Leftarrow$. Für jedes $r > 0$ bezeichnen wir

$$Q_r = \{x \in X : \|x\| \leq r\}.$$

Ist $A(Q)$ präkompakt so ist auch $A(Q_r) = rA(Q)$ präkompakt für jedes $r > 0$. Da jede beschränkte Menge M in einer Kugel Q_r liegt, so ist $A(M)$ eine Teilmenge von $A(Q_r)$ und somit präkompakt, woraus folgt, dass A kompakt ist.

(ii) Sei A kompakt. Ist $\{x_n\}$ eine beschränkte Folge so ist $M = \{x_n\}$ eine beschränkt Menge und somit $A(M) = \{Ax_n\}$ präkompakt ist. Somit hat die Folge $\{Ax_n\}$ eine konvergente Teilfolge.

Für die Umkehrung $\Leftarrow$ beweisen wir, dass $A(M)$ für jede beschränkte Teilmenge M präkompakt ist, Sei $\{y_n\}$ eine Folge aus $A(M)$. Dann gibt es $x_n \in M$ mit $Ax_n = y_n$. Die Folge $\{x_n\}$ ist beschränkt, woraus folgt dass die Folge $\{y_n\}$ eine konvergente Teilfolge hat. Somit ist $A(M)$ präkompakt und A ist kompakt.

(iii) Da $A(Q)$ präkompakt ist, so ist $A(Q)$ beschränkt, woraus folgt, dass A beschränkt ist. ■

Beispiel. Ein Operator $A \in \mathcal{B}(X)$ heißt *endlichdimensional*, wenn $\dim \operatorname{im} A < \infty$. Für endlichdimensionalen Operator A ist $A(Q)$ eine beschränkte Teilmenge vom endlichdimensionalen Unterraum $\operatorname{im} A$ und somit ist $A(Q)$ präkompakt. Daraus folgt, dass A kompakt ist.

Beispiel. Der identische Operator I im unendlichdimensionalen Hilbertraum H ist nie kompakt, da $I(Q) = Q$ nicht präkompakt ist.

Beispiel. Sei P ein orthogonaler Projektor im Hilbertraum H auf den abgeschlossenen Unterraum $U = \operatorname{im} P$. Ist U endlichdimensional, so wissen wir schon, dass P kompakt ist. Ist U unendlichdimensional, so ist P nicht kompakt, da für die Einheitskugel Q von U gilt $P(Q) = Q$ und Q nicht präkompakt ist.

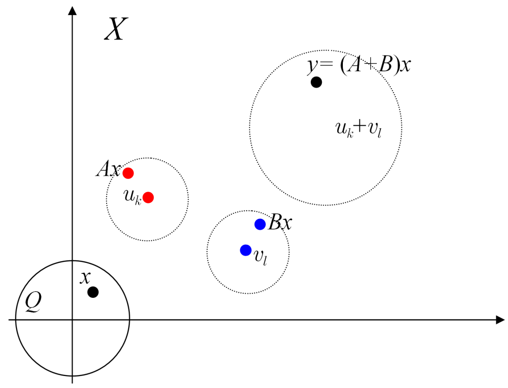
Die Menge von kompakten Operatoren wird mit $\mathcal{K}(X)$ bezeichnet, so dass $\mathcal{K}(X) \subset \mathcal{B}(X)$.

Satz 3.12 *Ist X ein Banachraum so ist $\mathcal{K}(X)$ ein abgeschlossener Unterraum von $\mathcal{B}(X)$.*

Darüber hinaus ist $\mathcal{K}(X)$ ein *Ideal* der Algebra $\mathcal{B}(X)$, d.h. für alle $A \in \mathcal{B}(X)$ und $C \in \mathcal{K}(X)$ sind die Operatoren AC und CA kompakt (siehe Aufgabe 71). Man kann auch zeigen, dass im Hilbertraum die Menge von endlichdimensionalen Operatoren dicht in $\mathcal{K}(X)$ liegt (siehe Aufgabe 72).

Beweis. Seien A und B kompakte Operatoren in X. Um zu beweisen, dass $A+B$ auch kompakt ist, zeigen wir, dass $(A+B)(Q)$ präkompakt ist, d.h. totalbeschränkt, vorausgesetzt, dass $A(Q)$ und $B(Q)$ totalbeschränkt sind. Seien $\{u_k\}$ ein ε-Netz von $A(Q)$ und $\{v_l\}$ ein ε-Netz von $B(Q)$.

Zeigen wir, dass die Doppelfolge $\{u_k+v_l\}$ ein 2ε-Netz in $(A+B)(Q)$ ist.



Für jedes $y \in (A+B)(Q)$ gilt

$$y=(A+B)x=Ax+Bx \quad \text{für ein } x \in Q.$$

Dann gilt für einige k und l

$$\|Ax-u_k\|<\varepsilon \quad \text{und} \quad \|Bx-v_l\|<\varepsilon,$$

woraus folgt

$$\|y-(u_k+v_l)\|=\|Ax+Bx-u_k-v_l\| \leq \|Ax-u_k\|+\|Bx-v_l\|<2\varepsilon.$$

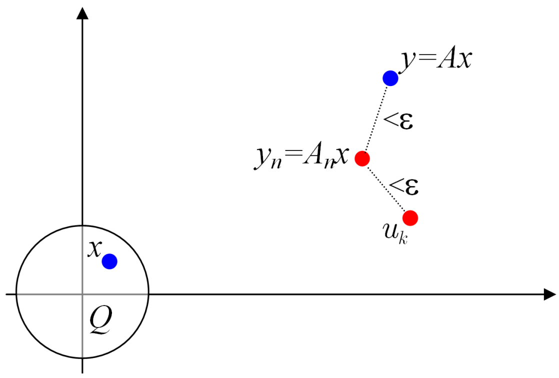
Somit ist die Folge $\{u_k+v_l\}$ ein 2ε-Netz in $(A+B)(Q)$, woraus folgt, dass $(A+B)(Q)$ totalbeschränkt ist.

Analog zeigt man, dass αA für jedes $\alpha \in \mathbb{K}$ kompakt ist. Somit ist $\mathcal{K}(X)$ ein Unterraum von $\mathcal{B}(X)$.

Jetzt beweisen wir, dass $\mathcal{K}(X)$ abgeschlossen ist. Sei $\{A_n\}_{n=1}^{\infty}$ eine Folge von kompakten Operatoren in X. Nehmen wir an, dass $A_n \to A \in \mathcal{B}(X)$ bezüglich der Operatornorm, und beweisen dass A auch kompakt ist. Dafür zeigen wir, dass $A(Q)$ totalbeschränkt ist. Für jedes $\varepsilon>0$ gibt es einen Operator A_n mit

$$\|A-A_n\|<\varepsilon.$$

Sei $\{u_k\}$ ein ε-Netz von $A_n(Q)$. Zeigen wir, dass $\{u_k\}$ ein 2ε-Netz von $A(Q)$ ist. Für jedes $y \in A(Q)$ gilt $y=Ax$ für ein $x \in Q$.



Da $y_n := A_n x$ in $A_n(Q)$ liegt, so gibt es ein u_k mit $\|y_n - u_k\| < \varepsilon$, woraus folgt

$$\begin{aligned}\|y - u_k\| &\le \|y - y_n\| + \|y_n - u_k\| \\ &\le \|(A - A_n)x\| + \varepsilon \\ &\le \|A - A_n\| \, \|x\| + \varepsilon < 2\varepsilon.\end{aligned}$$

Somit ist $\{u_k\}$ ein 2ε-Netz von $A(Q)$. ∎

Beispiel. Gegeben sei eine beschränkte Folge $\{\alpha_k\}_{k=1}^{\infty}$ von reellen (bzw komplexen) Zahlen. Betrachten wir den folgenden Operator A in l^2 (bzw $l^2_{\mathbb{C}}$): für $x = \{x_k\}_{k\in\mathbb{N}}$ setzen wir

$$Ax = \{\alpha_k x_k\}_{k\in\mathbb{N}} = \{\alpha_1 x_1, \alpha_2 x_2, ..., \alpha_k x_k, ...\}.$$

Da

$$\|Ax\|_2^2 = \sum_{k\in\mathbb{N}} |\alpha_k x_k|^2 \le \sup_k |\alpha_k|^2 \|x\|_2^2 < \infty,$$

so liegt Ax in l^2, A ist ein Operator in l^2 und es gilt

$$\|A\| \le \sup_k |\alpha_k| < \infty. \tag{3.27}$$

Somit haben wir $A \in \mathcal{B}(l^2)$.

Bemerken wir, dass die Elemente

$$e_n = \{0, 0, ..., \overset{n}{\widehat{1}}, 0, 0, ...\}$$

von Standardbasis in l^2 die Eigenvektoren von A mit den Eigenwerten α_n sind:

$$Ae_n = \alpha_n e_n.$$

Beweisen wir folgendes: gilt $\alpha_n \to 0$ so ist A kompakt, d.h. $A \in \mathcal{K}(l^2)$. Dafür betrachten wir für jedes $n \in \mathbb{N}$ den folgenden Operator A_n in l^2:

$$A_n x = \{\alpha_1 x_1, ..., \alpha_n x_n, 0, 0, ...\}.$$

Da im $A_n \subset \operatorname{span}\{e_1, ..., e_n\}$, so ist der Operator A_n endlichdimensional und somit kompakt. Wir haben

$$(A - A_n)x = \{0, ...0, \alpha_{n+1}x_{n+1}, ..., \alpha_k x_k, ...\},$$

so dass nach (3.27) gilt

$$\|A - A_n\| \leq \sup_{k \geq n+1} |\alpha_k| \,.$$

Die Voraussetzung $\alpha_n \to 0$ impliziert $\sup_{k \geq n+1} |\alpha_k| \to 0$ für $n \to \infty$, woraus folgt, dass $\|A_n - A\| \to 0$. Somit ist A auch kompakt.

Die Bedingung $\alpha_n \to 0$ ist nicht nur hinreichend für die Kompaktheit von A aber auch notwendig, was man aus dem folgenden Lemma sieht.

Lemma 3.13 *Sei H ein Hilbertraum und $A \in \mathcal{K}(H)$. Sei $\{v_n\}_{n=1}^{\infty}$ eine orthonormale Folge von Eigenvektoren von A mit den Eigenwerten α_n, d.h. $Av_n = \alpha_n v_n$. Dann gilt $\alpha_n \to 0$ für $n \to \infty$.*

Beweis. Nehmen wir das Gegenteil an, dass $\alpha_n \not\to 0$. Dann existiert ein $\varepsilon > 0$ und eine Teilfolge $\{\alpha_{n_k}\}$ mit $|\alpha_{n_k}| \geq \varepsilon$ für alle k. Wir können $\{\alpha_{n_k}\}$ in $\{\alpha_n\}$ umbenennen und somit annehmen, dass $|\alpha_n| \geq \varepsilon$ für alle n. Da A kompakt ist und die Folge $\{v_n\}$ beschränkt ist, so hat die Folge $\{Av_n\}$ eine konvergente Teilfolge.

Andererseits, da $v_n \perp v_m$ für $n \neq m$, so erhalten wir

$$\|Av_n - Av_m\|^2 = \|\alpha_n v_n - \alpha_m v_m\|^2 = |\alpha_n|^2 \|v_n\|^2 + |\alpha_m|^2 \|v_m\|^2 \geq 2\varepsilon^2.$$

Somit kann die Folge $\{Av_n\}$ keine Cauchy Teilfolge enthalten, was im Widerspruch zur Voraussetzung steht. ■

Der folgende Satz liefert ein wichtiges Beispiel von kompakten Operatoren.

Satz 3.14 *Sei X eine beschränkte messbare Teilmenge von $\mathbb{R}^n$. Sei $k(x,y)$ eine messbare Funktion auf $X \times X$ mit*

$$\int_{X \times X} |k(x,y)|^2 \, d\lambda_{2n}(x,y) < \infty. \tag{3.28}$$

Dann ist der Integraloperator

$$Kf(x) = \int_X k(x,y) f(y) d\lambda_n(y)$$

ein kompakter Operator in $L^2(X, \lambda_n)$.

Beweis. Nach dem Satz 3.3 ist K ein beschränkter Operator in $L^2(X, \lambda_n)$. Um zu zeigen, dass K kompakt ist, werden wir K mit endlichdimensionalen Operatoren approximieren. Dafür wählen wir zuerst eine Orthonormalbasis $\{u_i\}_{i \in \mathbb{N}}$ im Hilbertraum $L^2(X, \lambda_n)$ die nach dem Korollar 2.26 existiert. Nach dem Satz 2.25 erhalten wir die Orthonormalbasis

$$\{u_i(x) \, u_j(y)\}_{i,j \in \mathbb{N}}$$

im Hilbertraum $L^2(X \times X, \lambda_{2n})$. Wir stellen diese Doppelfolge dar als eine Folge

$$\{w_m(x,y)\}_{m \in \mathbb{N}},$$

wobei

$$w_m(x,y) = u_i(x) \, u_j(y) \tag{3.29}$$

für einige i, j. Da $k \in L^2(X \times X, \lambda_{2n})$, so gilt eine Darstellung von k in der Basis $\{w_m\}$:

$$k(x,y) = \sum_{m=1}^{\infty} c_m w_m(x,y)$$

mit Koeffizienten c_m. Nach der Parsevalschen Gleichung haben wir

$$\|k\|^2_{L^2(X\times X,\lambda_{2n})} = \sum_{m=1}^{\infty} |c_m|^2 .$$

Da nach dem Satz 3.3 gilt $\|K\| \leq \|k\|_{L^2(X\times X,\lambda_{2n})}$, so erhalten wir

$$\|K\|^2 \leq \sum_{m=1}^{\infty} |c_m|^2 . \tag{3.30}$$

Betrachten wir für jedes $N \in \mathbb{N}$ den Integraloperator K_N mit dem Kern

$$k_N(x,y) = \sum_{m=1}^{N} c_m w_m(x,y), \tag{3.31}$$

d.h.

$$K_N f(x) = \int_X k_N(x,y) f(y) d\lambda_n(y) .$$

Der Operator $K - K_N$ hat den Kern

$$k(x,y) - k_N(x,y) = \sum_{m=N+1}^{\infty} c_m w_m(x,y),$$

und es folgt aus (3.30) dass

$$\|K - K_N\|^2 \leq \sum_{m=N+1}^{\infty} |c_m|^2 \to 0 \text{ für } N \to \infty,$$

so dass $K_N \to K$ bezüglich der Operatornorm.

Nach dem Satz 3.12 reicht es zu beweisen, dass K_N kompakt ist. Dafür betrachten wir die Integraloperatoren mit den Kernen $w_m(x,y)$:

$$W_m f(x) = \int_X w_m(x,y) f(y) d\lambda_n(y),$$

und bemerken, dass K_N nach (3.31) eine endliche lineare Kombination von W_m ist:

$$K_N = \sum_{m=1}^{N} c_m W_m .$$

Nach dem Satz 3.12 reicht es zu beweisen, dass jeder Operator W_m kompakt ist. Aus (3.29) erhalten wir, dass für einige i, j,

$$W_m f(x) = \int_X u_i(x) u_j(y) f(y) d\lambda_n(y) = c\, u_i(x),$$

wobei $c = \int_X u_j(y) f(y) d\lambda_n(y)$. Wir sehen, dass $W_m f$ immer ein Vielfache von u_i ist, so dass $\dim \operatorname{im} W_m = 1$. Somit ist der Operator W_m kompakt, was zu beweisen war. ∎

Ist der Kern $k(x,y)$ zusätzlich Hermitesch, so erhalten wir, dass der Integraloperator K selbstadjungiert und kompakt ist.

3.8 Satz von Hilbert-Schmidt

We haben oberhalb schon bemerkt, dass es für jeden selbstadjungierten Operator A im endlichdimensionalen Hilbertraum H eine Orthogonalbasis in H gibt die aus den Eigenvektoren von A besteht. Diese Aussage heißt Diagonalisierung von A da die Matrix von A in dieser Basis diagonal ist.

Der nächste Satz besagt die Diagonalisierung von selbstadjungierten *kompakten* Operatoren in allgemeinen Hilberträumen.

Hauptsatz 3.15 (Satz von Hilbert-Schmidt) *Sei A ein kompakter selbstadjungierter Operator im Hilbertraum H. Dann gilt folgendes.*

1. *Jedes $\lambda \in \sigma(A) \setminus \{0\}$ ist ein Eigenwert von A mit endlicher Vielfachheit.*

2. *Die Menge von allen unterschiedlichen nicht-Null Eigenwerten von A ist höchstens abzählbar und somit lässt sich als eine Folge $\{\lambda_n\}_{n=1}^N$ nummerieren wobei $N \in \mathbb{N} \cup \{\infty\}$. Im Fall $N = \infty$ gilt $\lambda_n \to 0$ für $n \to \infty$.*

3. (Spektralzerlegung) *Bezeichnen wir*
$$K_n = \ker(A - \lambda_n I) \quad \text{für} \quad n \geq 1 \quad \text{und} \quad K_0 = \ker A.$$
Dann sind die Unterräume $\{K_n\}_{n=0}^\infty$ zueinander orthogonal, und es gilt für alle $x \in H$
$$x = \sum_{n=0}^N P_{K_n} x \tag{3.32}$$
und
$$Ax = \sum_{n=1}^N \lambda_n P_{K_n} x. \tag{3.33}$$

4. (Diagonalisierung) *Ist K_0 (oder H) separabel, so gibt es eine Orthogonalbasis in H die aus den Eigenvektoren von A besteht.*

10.06.22 **Vorlesung 18**

Die Identitäten (3.32) und (3.33) lassen sich wie folgt umschreiben:
$$I = \sum_{n=0}^N P_{K_n}$$
und
$$A = \sum_{n=1}^N \lambda_n P_{K_n}, \tag{3.34}$$
wobei im Fall $N = \infty$ die Konvergenz der Reihen *punktweise* ist, d.h. die starke Konvergenz von Operatoren. Allerdings gilt in (3.34) sogar die Konvergenz in der Operatornorm – siehe Aufgabe 65.

Die Identität (3.34) bedeutet folgendes: jeder kompakte selbstadjungierte Operator ist eine (endliche oder abzählbare) lineare Kombination von endlichdimensionalen Projektoren P_{K_n} mit zueinander orthogonalen K_n und mit den Koeffizienten $\lambda_n \to 0$.

Die Umkehrung gilt auch: für jede Folge $\{K_n\}_{n=1}^{\infty}$ von zueinander orthogonalen endlichdimensionalen Unterräumen und für jede Nullfolge $\{\lambda_n\}_{n=1}^{\infty}$ von reellen Zahlen, konvergiert die rechte Seite von (3.34) in der Operatornorm, und die Summe A ist ein kompakter selbstadjungierter Operator (siehe Aufgaben 73, 74). Die Darstellung (3.34) von kompakten selbstadjungierten Operatoren heißt *Spektralzerlegung.*

Für den Beweis des Hauptsatzes 3.15 brauchen wir das folgende Lemma.

Lemma 3.16 *Sei $\{K_n\}_{n=1}^{N}$ eine (endliche oder abzählbare) Folge von abgeschlossenen zueinander orthogonalen Unterräumen von H. Betrachten wir den Unterraum*

$$U = \bigoplus_{n=1}^{N} K_n := \overline{\operatorname{span}}\,\{K_n\}_{n=1}^{N}.$$

Dann für jedes $x \in U$ gilt

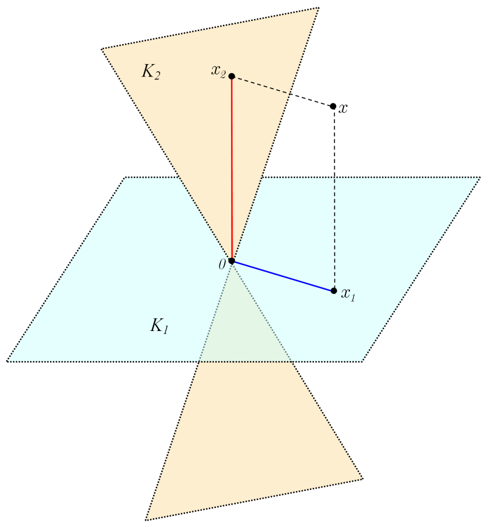
$$x = \sum_{n=1}^{N} P_{K_n} x. \tag{3.35}$$

Der Ausdruck $\bigoplus_{n=1}^{N} K_n$ heißt die *direkte Summe* der Unterräume K_n.

Beweis. Fixieren wir ein $x \in H$ and setzen

$$x_n = P_{K_n} x,$$

so dass die Folge $\{x_n\}$ orthogonal ist.



Projektionen x_n von x auf K_n

Beweisen wir zunächst, dass im Fall $N = \infty$ die Reihe $\sum_{n=1}^{\infty} x_n$ konvergent ist. Ohne Beschränkung der Allgemeinheit können wir annehmen, dass alle $x_n \neq 0$. Setzen wir

$$e_n = \frac{x_n}{\|x_n\|}$$

so dass $\{e_n\}$ eine orthonormale Folge ist. Nach der Besselschen Ungleichung gilt

$$\sum_{n=1}^{\infty} |(x, e_n)|^2 \leq \|x\|^2 .$$

Da

$$(x, e_n) = \frac{1}{\|x_n\|}(x, x_n) = \frac{1}{\|x_n\|}(x, P_{K_n}x_n) = \frac{1}{\|x_n\|}(P_{K_n}x, x_n) = \frac{1}{\|x_n\|}\|x_n\|^2 = \|x_n\|$$

so erhalten wir

$$\sum_{n=1}^{\infty} \|x_n\|^2 \leq \|x\|^2 < \infty \tag{3.36}$$

(siehe auch die Aufgabe 54). Nach dem Lemma 2.10 ist die Reihe $\sum_{n=1}^{\infty} x_n = \sum_{n=1}^{\infty} \|x_n\| e_n$ konvergent.

Jetzt beweisen wir (3.35) für alle $x \in U$, was äquivalent zu

$$x = \sum_{n=1}^{N} x_n \tag{3.37}$$

ist. Dafür setzen wir

$$y = x - \sum_{n=1}^{N} x_n$$

und erhalten für jedes $n \leq N$

$$\begin{aligned} P_{K_n}y &= P_{K_n}x - P_{K_n}\left(\sum_{m=1}^{N} x_m\right) \\ &= x_n - \sum_{m=1}^{N} P_{K_n}x_m \\ &= x_n - x_n = 0, \end{aligned}$$

da

$$P_{K_n}x_m = \begin{cases} x_n, & m = n, \\ 0, & m \neq n. \end{cases}$$

Somit gilt $y \perp K_n$ für alle n, woraus folgt, dass $y \perp U$. Da auch $y \in U$, so erhalten wir $y = 0$, woraus (3.37) folgt. ■

Beweis von dem Satz 3.15. Der Beweis besteht aus 4 Teilen.

Teil 1. Sei $\lambda \in \mathbb{R}$. Nach dem Satz 3.9 ist die Bedingung $\lambda \in \rho(A)$ äquivalent zu den folgenden zwei Bedingungen:

(a) $\ker\left(A^* - \overline{\lambda}I\right) = \{0\}$

(b) $\|Ax - \lambda x\| \geq c\|x\|$ für ein $c > 0$ und alle $x \in H$.

In unserem Fall haben wir $A^* = A$ und $\overline{\lambda} = \lambda$, so dass (a) äquivalent zu

$$\ker(A - \lambda I) = \{0\}$$

ist, was offensichtlich aus (b) folgt. Somit gilt

$$\lambda \in \rho(A) \Leftrightarrow (b) .$$

Da die Bedingung (b) äquivalent zu

$$\inf_{\|x\|=1} \|A - \lambda x\| > 0$$

ist, so folgt es daraus, dass

$$\lambda \in \sigma(A) \Leftrightarrow \inf_{\|x\|=1} \|Ax - \lambda x\| = 0. \tag{3.38}$$

Jetzt beweisen wir, dass jedes $\lambda \in \sigma(A) \setminus \{0\}$ ein Eigenwert von A ist. Nach dem Satz 3.10 ist λ reell. Nach (3.38) existiert eine Folge $\{x_n\}_{n=1}^{\infty}$ von Vektoren mit $\|x_n\| = 1$ und

$$\|Ax_n - \lambda x_n\| \to 0 \text{ für } n \to \infty.$$

Da A kompakt ist und die Folge $\{x_n\}$ beschränkt ist, so hat die Folge $\{Ax_n\}$ eine konvergente Teilfolge. So, wir können annehmen, dass die ganze Folge $\{Ax_n\}$ konvergiert, und setzen

$$y = \lim_{n\to\infty} Ax_n.$$

Da

$$\lim_{n\to\infty} (Ax_n - \lambda x_n) = 0$$

und $\lambda \neq 0$. so folgt es daraus, dass $\{x_n\}$ auch konvergent ist und

$$\lim_{n\to\infty} \lambda x_n = y.$$

Insbesondere gilt $\|y\| = |\lambda| \neq 0$. Es folgt auch

$$Ay = \lim_{n\to\infty} A(\lambda x_n) = \lambda \lim_{n\to\infty} (Ax_n) = \lambda y$$

so dass y ein Eigenvektor mit dem Eigenwert λ ist. Insbesondere ist λ ein Eigenwert von A.

Beweisen wir jetzt, dass die Vielfachheit von λ endlich ist. Gilt

$$\dim\ker(A - \lambda I) = \infty,$$

so existiert im Unterraum $\ker(A - \lambda I)$ eine orthonormale Folge $\{v_n\}_{n=1}^{\infty}$. Für alle n haben wir $Av_n = \alpha_n v_n$ mit $\alpha_n = \lambda$. Nach dem Lemma 3.13 muss die Folge $\{\alpha_n\}$ gegen 0 konvergieren, was nicht der Fall ist. Somit gilt

$$\dim\ker(A - \lambda I) < \infty,$$

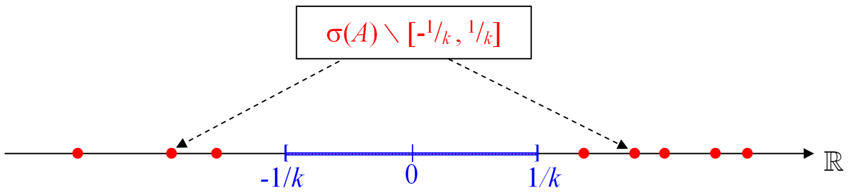
was zu beweisen war.

Teil 2. Beweisen wir, dass für jedes $\varepsilon > 0$ die Menge $\sigma(A) \setminus [-\varepsilon, \varepsilon]$ endlich ist. Ist diese Menge unendlich, so existiert eine Folge $\{\alpha_n\}_{n=1}^{\infty}$ von verschiedenen Eigenwerten von A mit $|\alpha_n| > \varepsilon$. Sei v_n ein Eigenvektor von α_n mit $\|v_n\| = 1$. Nach dem Satz 3.10 ist die Folge $\{v_n\}$ orthogonal. Nach Lemma 3.13 beschließen wir, dass $\alpha_n \to 0$, was im Widerspruch zu $|\alpha_n| > \varepsilon$ steht.

Da jede Menge $\sigma(A) \setminus \left[-\frac{1}{k}, \frac{1}{k}\right]$ für jedes $k \in \mathbb{N}$ endlich ist und

$$\sigma(A) \setminus \{0\} = \bigcup_{k=1}^{\infty} \left(\sigma(A) \setminus \left[-\tfrac{1}{k}, \tfrac{1}{k}\right]\right)$$

so erhalten wir, dass die Menge $\sigma(A) \setminus \{0\}$ höchstens abzählbar ist und somit sich nummerieren lässt.



Bezeichnen wir mit $\{\lambda_n\}_{n=1}^N$ die Folge von allen Elementen von $\sigma(A) \setminus \{0\}$, d.h. von allen nicht-Null Eigenwerten von A. Im Fall $N = \infty$ beschließen wir nach dem Lemma 3.13, dass $\lambda_n \to 0$ für $n \to \infty$.

Teil 3. Sei $\{\lambda_n\}_{n=1}^N$ die Folge von allen nicht-Null Eigenwerten von A, wie oberhalb, und bezeichnen wir

$$K_n = \ker(A - \lambda_n I).$$

Setzen wir auch $\lambda_0 = 0$ und $K_0 = \ker A$ (ist $\lambda_0 = 0$ ein Eigenwert von A so ist K_0 der Eigenraum von λ_0, sonst $K_0 = \{0\}$). Nach dem Satz 3.10 sind die Unterräume K_n zueinander orthogonal. Setzen wir

$$U = \bigoplus_{n=0}^N K_n = \overline{\operatorname{span}} \{K_n\}_{k=0}^\infty. \tag{3.39}$$

Nach dem Lemma 3.16 gilt für alle $x \in U$

$$x = \sum_{n=0}^N P_{K_n} x \tag{3.40}$$

woraus folgt

$$Ax = \sum_{n=0}^N A(P_{K_n} x) = \sum_{n=0}^N \lambda_n P_{K_n} x$$

und somit

$$Ax = \sum_{n=1}^N \lambda_n P_{K_n} x, \tag{3.41}$$

wobei wir benutzt haben, dass $\lambda_0 = 0$. Aus (3.40) und (3.41) erhalten wir die Identitäten (3.32) und (3.33) für alle $x \in U$. Es bleibt zu beweisen, dass $U = H$.

Da $P_{K_n} x \in K_n$, so folgt aus (3.41), dass

$$x \in U \Rightarrow Ax \in U.$$

Somit ist der Unterraum U *invariant* bezüglich A, d.h. $AU \subset U$. Da U ein abgeschlossener Unterraum ist, so ist die Aussage $U = H$ äquivalent zu $U^\perp = \{0\}$. Nehmen wir das Gegenteil an, dass $U^\perp \neq \{0\}$, und führen die Annahme zum Widerspruch. Nach

Zeigen wir zuerst, dass der Unterraum $U^\perp$ auch invariant bezüglich A ist, d.h.

$$y \in U^\perp \Rightarrow Ay \in U^\perp. \tag{3.42}$$

In der Tat erhalten wir für $y \in U^\perp$, dass $y \perp AU$ (da $AU \subset U$) d.h.

$$\forall x \in U \quad (y, Ax) = 0$$

woraus folgt

$$\forall x \in U \quad (Ay, x) = 0;$$

daher $Ay \perp x$ für alle $x \in U$ und $Ay \in U^{\perp}$, was beweist (3.42) (siehe auch Aufgabe 70).

Betrachten wir den Operator $B = A|_{U^{\perp}}$ im Hilbertraum $U^{\perp}$. Offensichtlich B ist auch selbstadjungiert und kompakt (die Eigenschaften Selbstadjungiertheit und Kompaktheit bewahren sich unter Einschränkung auf einen Unterraum).

Gilt $B = 0$, so ist $\lambda = 0$ ein Eigenwert mit dem Eigenraum $U^{\perp}$ (da nach Voraussetzung $U^{\perp} \neq \{0\}$).

Gilt $B \neq 0$, so enthält $\sigma(B)$ nach dem Satz 3.10 ein nicht-Null Element ($\|B\|$ oder $-\|B\|$), und nach dem Teil 1 ist dieses Element von $\sigma(B)$ ein Eigenwert. Somit gibt es immer einen Eigenvektor $v \in U^{\perp}$ von B. Dann ist v auch ein Eigenvektor von A, was ein Widerspruch ist, da nach der Definition von U alle Eigenvektoren von A in U liegen, während $v \in U^{\perp}$.

Damit beschließen wir, dass $U = H$, und die Identität (3.40) gilt für alle $x \in H$. Daraus folgt, dass auch 3.41 für alle $x \in H$ gilt, was zu beweisen war.

Teil 4. Jeder Unterraum $K_n = \ker(A - \lambda_n I)$, $n \geq 0$, ist abgeschlossen, so dass jeder K_n selbst ein Hilbertraum ist.

Der Raum K_0 ist separabel nach der Voraussetzung, und nach dem Satz 2.12 gibt es in K_0 eine Orthonormalbasis, endlich oder abzählbar.

Jeder Raum K_n mit $n \geq 1$ ist endlichdimensional, so dass es in K_n auch eine Orthonormalbasis gibt.

Die Vereinigung von den Orthonormalbasen aus allen K_n, $n \geq 0$, ist höchstens abzählbar. Bezeichnen wir diese Vereinigung mit $\{e_k\}$. Dann ist $\{e_k\}$ eine orthonormale Folge (da die Räume K_n zueinander orthogonal sind), und nach dem Teil 3 gilt

$$\overline{\operatorname{span}}\{e_k\} = \overline{\operatorname{span}}\{K_n\} = \bigoplus_{n=0}^{N} K_n = H.$$

Somit ist $\{e_k\}$ eine Orthonormalbasis in H. Es bleibt nur zu bemerken, dass jedes e_k in einem Eigenraum von A liegt, so dass e_k ein Eigenvektor von A ist. ∎

3.9 Randwertproblem und Greenscher Operator

Betrachten wir auf einem Intervall $[a, b]$ den Differentialoperator

$$Lu = -(pu')' + qu, \tag{3.43}$$

zweiter Ordnung, wobei $p \in C^1[a, b]$ und $q \in C[a, b]$ gegebene Funktionen sind. Wir werden immer annehmen dass $p > 0$ und $q \geq 0$ auf $[a, b]$. Zum Beispiel, im Fall $p \equiv 1$ und $q \equiv 0$ gilt $Lu = -u''$.

Betrachten wir die Differentialgleichung $Lu = f$ wobei $f \in C[a, b]$ eine gegebene Funktion ist und $u \in C^2[a, b]$ unbekannt ist. Allerdings bestimmt diese Gleichung die Funktion u nicht eindeutig: um die Eindeutigkeit von u zu sichern braucht man zusätzliche *Anfangs*- oder *Randbedingungen*.

Betrachten wir das folgende *Randwertproblem*

$$\begin{cases} -(pu')' + qu = f \\ u(a) = u(b) = 0. \end{cases} \tag{3.44}$$

15.06.22 Vorlesung 19

Satz 3.17 *Unter den Bedingungen $p > 0$ und $q \geq 0$ hat das Randwertproblem (3.44) für jedes $f \in C[a,b]$ genau eine eindeutige Lösung $u \in C^2[a,b]$.*

Beweis. Zuerst beweisen wir die folgende Aussage: für alle Funktionen $u \in C^2[a,b]$ und $v \in C_0^1[a,b]$ gilt die Identität

$$\boxed{(Lu, v) = \int_a^b (pu'v' + quv)\, dx}, \tag{3.45}$$

wobei $(\cdot,\cdot)$ das Skalarprodukt in $L^2[a,b]$ bezeichnet. In der Tat ergibt die partielle Integration

$$(Lu, v) = \int_a^b \left(-(pu')' v + quv\right) dx = [-pu'v]_a^b + \int_a^b (pu'v' + quv)\, dx,$$

woraus (3.45) folgt, da $[-pu'v]_a^b = 0$. Die Identität (3.45) impliziert insbesondere die Symmetrie des Ausdrucks (Lu, v): für alle $u, v \in C_0^2[a,b]$ gilt

$$(Lu, v) = (u, Lv). \tag{3.46}$$

Mit Hilfe von (3.45) beweisen wir die Eindeutigkeit der Lösung von (3.44). Gibt es zwei Lösungen u_1 und u_2 von (3.44), so setzen wir $v = u_1 - u_2$ so dass v das folgende Randwertproblem erfüllt:

$$\begin{cases} Lv = 0 \\ v(a) = v(b) = 0 \end{cases}. \tag{3.47}$$

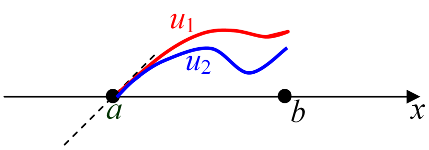
Nach (3.45) gilt

$$0 = (Lv, v) = \int_a^b \left(p(v')^2 + qv^2\right) dx. \tag{3.48}$$

Da $p > 0$ und $q \geq 0$, so folgt es aus (3.48), dass $v' = 0$ und somit $v = 0$, was zu beweisen war.

Jetzt beweisen wir die Existenz der Lösung von (3.44). Dafür betrachten wir zwei *Anfangswertprobleme*:

$$\begin{cases} Lu_1 = f \\ u_1(a) = 0,\ u_1'(a) = 1 \end{cases} \quad \text{und} \quad \begin{cases} Lu_2 = 0 \\ u_2(a) = 0,\ u_2'(a) = 1 \end{cases}$$



Nach dem Satz von Picard-Lindelöf für lineare DGLen haben diese Probleme die Lösingen u_1 und u_2 auf dem ganzen Intervall $[a,b]$. Für jedes $c \in \mathbb{R}$ erfüllt die Funktion

$$u = u_1 + cu_2$$

die Gleichung $Lu = f$ und die Bedingung $u(a) = 0$. Wählen wir c so dass auch $u(b) = 0$, d.h.

$$c = -\frac{u_1(b)}{u_2(b)},$$

was möglich ist, da $u_2(b) \neq 0$. In der Tat, im Fall $u_2(b) = 0$ löst u_2 das Randwertproblem (3.47), und somit erhalten wir nach der Eindeutigkeit, dass $u_2 \equiv 0$ was im Widerspruch zu $u_2'(a) = 1$ steht. Somit erhalten wir eine Lösung u von (3.44). ■

Bemerkung. Im Fall $q < 0$ gilt die Aussage des Satzes 3.17 nicht. Zum Beispiel, seien $q \equiv -1$, $p \equiv 1$ und $[a, b] = [0, \pi]$. Dann ist das Randwertproblem (3.44) äquivalent zu

$$\begin{cases} u'' + u = f \\ u(0) = u(\pi) = 0. \end{cases} \tag{3.49}$$

Für $f = 0$ hat dieses Problem eine nicht-Null Lösung $u(x) = \sin x$ so dass die Eindeutigkeit nicht erfüllt ist. Die Existenz ist auch nicht immer der Fall. Z.B. zeigen wir, dass es für die Funktion $f(x) = \sin x$ keine Lösung von (3.49) gibt. In der Tat, ist u eine Lösung so erhalten wir nach (3.46) dass

$$\int_0^\pi (u'' + u) f \, dx = (Lu, f) = (u, Lf) = \int_0^\pi (f'' + f) u \, dx = 0,$$

während nach (3.49)

$$\int_0^\pi (u'' + u) f \, dx = \int_0^\pi f^2 dx > 0.$$

Im Gegenteil dazu hat das Anfangswertproblem eine eindeutige Lösung unabhängig vom Vorzeichen von q.

Unter den Voraussetzungen des Satzes 3.17 erhalten wir einen Operator

$$\begin{gathered} K : C[a, b] \to C[a, b] \\ Kf = u, \end{gathered}$$

wobei u die eindeutige Lösung von (3.44) ist. Offensichtlich ist der Operator K linear.

Definition. Der Operator K heißt der *Greensche Operator* des Randwertproblems (3.44).

Wie wir unterhalb beweisen, der Greensche Operator K lässt sich zu einen kompakten selbstadjungierten Operator in $L^2[a, b]$ erweitern.

Beispiel. Bestimmen wir den Greenschen Operator K für den Fall $Lu = -u''$ (d.h. $p \equiv 1$ und $q \equiv 0$) und $[a, b] = [0, 1]$, d.h. für das Randwertproblem

$$\begin{cases} -u'' = f \\ u(0) = u(1) = 0 \end{cases}. \tag{3.50}$$

Wir haben für jedes $x \in [0, 1]$

$$u'(x) = u'(0) + \int_0^x u''(y) \, dy = -\int_0^x f(y) \, dy + C$$

wobei $C = u'(0)$, und

$$u(x) = u(0) + \int_0^x u'(z)\,dz = -\int_0^x \left(\int_0^z f(y)\,dy\right) dz + Cx.$$

Nach dem Satz von Fubini gilt

$$\begin{aligned}
\int_0^x \left(\int_0^z f(y)\,dy\right) dz &= \int_0^x \left(\int_0^x \mathbf{1}_{\{y\le z\}} f(y)\,dy\right) dz \\
&= \int_{[0,x]\times[0,x]} \mathbf{1}_{\{y\le z\}} f(y)\,d\lambda_2(y,z) \\
&= \int_0^x \left(\int_0^x \mathbf{1}_{\{y\le z\}} f(y)\,dz\right) dy \\
&= \int_0^x f(y) \left(\int_y^x dz\right) dy \\
&= \int_0^x (x-y)\, f(y)\,dy,
\end{aligned}$$

woraus folgt

$$u(x) = \int_0^x (y-x)\, f(y)\,dy + Cx.$$

Wir wählen jetzt C um die Randbedingung $u(1) = 0$ zu erfüllen:

$$0 = u(1) = \int_0^1 (y-1)\, f(y)\,dy + C$$

so dass

$$C = \int_0^1 (1-y)\, f(y)\,dy.$$

Es folgt, dass

$$\begin{aligned}
u(x) &= \int_0^x (y-x)\, f(y)\,dy + x\int_0^1 (1-y)\, f(y)\,dy \\
&= \int_0^1 \big(\mathbf{1}_{\{y\le x\}}(y-x) + x(1-y)\big) f(y)\,dy \\
&= \int_0^1 k(x,y)\, f(y)\,dy,
\end{aligned}$$

wobei

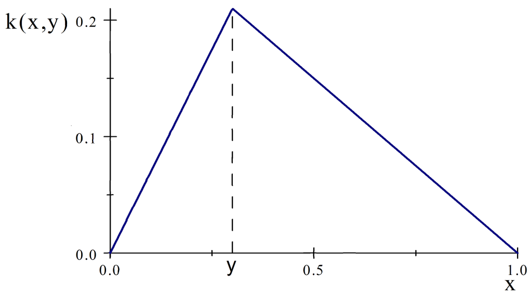
$$\begin{aligned}
k(x,y) &= \mathbf{1}_{\{y\le x\}}(y-x) + x - xy = \begin{cases} y-xy, & x\ge y \\ x-xy, & x<y \end{cases} \\
&= \min(x,y) - xy.
\end{aligned}$$

Somit erhalten wir, dass

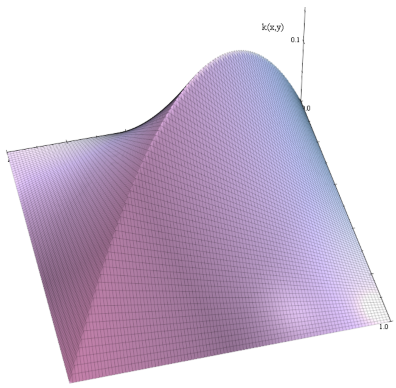
$$Kf(x) = u(x) = \int_0^1 k(x,y)\, f(y)\,dy,$$

d.h. der Greensche Operator K ein Integraloperator mit dem Kern $k(x,y)$ ist. Der Operator K mit dem Kern k lässt sich als ein Operator in $L^2[0,1]$ betrachten. Nach dem Satz 3.14 ist der Operator K kompakt. Da der Kern $k(x,y)$ symmetrisch ist, so ist K auch selbstadjungiert.

Die Funktion $k(x,y)$ heißt die *Greensche Funktion* des Randwertproblems (3.50). Diese Funktion wird auf den Bildern gezeigt:



Die Greensche Funktion $k(x,y) = \min(x,y) - xy$ als eine Funktion von x



Die Greensche Funktion $k(x,y) = \min(x,y) - xy$

Beispiel. Im Fall

$$Lu = -(pu')'$$

(und somit $q \equiv 0$) lässt sich die Greensche Funktion auf $[a,b]$ analog bestimmen:

$$k(x,y) = \min(v(x), v(y)) - \frac{v(x)v(y)}{v(b)}, \tag{3.51}$$

wobei

$$v(x) = \int_a^x \frac{dz}{p(z)}.$$

Für den Beweis von Kompaktheit und Selbstadjungiertheit von K im allgemeinen Fall brauchen wir zwei Lemmas.

Lemma 3.18 (A-priori-Abschätzung) *Für jede Funktion $f \in C[a,b]$ und für die Lösung u von (3.44) gilt die Ungleichung*

$$\sup_{[a,b]} |u| \leq C \|f\|_2 \,, \tag{3.52}$$

mit einer Konstante C die von p, a, b bestimmt wird.

Man kann (3.52) in Bezug auf den Greenschen Operator wie folgt umschreiben:

$$\sup_{[a,b]} |Kf| \leq C \|f\|_2 \,. \tag{3.53}$$

Beweis. Beweisen wir zuerst, dass

$$\sup_{x\in[a,b]} |u(x)|^2 \leq c \|f\|_2 \|u\|_2 \tag{3.54}$$

mit einer Konstante c. Es folgt aus (3.44) und (3.45) dass

$$(f,u) = (Lu,u) = \int_a^b \left(p(u')^2 + qu^2\right) dy \geq \int_a^b p(u')^2 \, dy, \tag{3.55}$$

wobei wir benutzt haben dass $q \geq 0$. Andererseits erhalten wir nach der Cauchy-Schwarz-Ungleichung, dass für alle $x \in [a,b]$

$$\begin{aligned} u(x)^2 &= \left(\int_a^x u' dy\right)^2 = \left(\int_a^x \frac{1}{\sqrt{p}}(\sqrt{p}u')dy\right)^2 \\ &\leq \left(\int_a^x \frac{1}{p} dy\right)\left(\int_a^x p(u')^2 \, dy\right) \\ &\leq \left(\int_a^b \frac{1}{p} dy\right)\left(\int_a^b p(u')^2 \, dy\right) = c \int_a^b p(u')^2 \, dy, \end{aligned} \tag{3.56}$$

wobei

$$c = \int_a^b \frac{dy}{p(y)}.$$

Es folgt aus (3.55) und (3.56) dass

$$u(x)^2 \leq c(f,u) \leq c \|f\|_2 \|u\|_2 \,, \tag{3.57}$$

woraus (3.54) folgt.

Es folgt aus (3.57), dass

$$\|u\|_2^2 = \int_a^b u(x)^2 \, dx \leq \int_a^b c \|f\|_2 \|u\|_2 \, dx = (b-a) c \|f\|_2 \|u\|_2 \,,$$

und somit

$$\|u\|_2 \leq (b-a) c \|f\|_2 \,. \tag{3.58}$$

Einsetzen von (3.58) in (3.54) ergibt

$$\sup_{x\in[a,b]} |u(x)|^2 \leq c \|f\|_2 \|u\|_2 \leq c^2 (b-a) \|f\|_2^2 \,,$$

woraus (3.52) mit $C = c\sqrt{b-a}$ folgt. ∎

Lemma 3.19 *Sei $K : C[a,b] \to C[a,b]$ ein linearer Operator mit der folgenden Eigenschaft: es gibt eine Konstante C so dass für jedes $f \in C[a,b]$ gilt*

$$\sup_{[a,b]} |Kf| \le C \|f\|_2 . \tag{3.59}$$

Dann gelten die folgenden Aussagen.

(*a*) *Der Operator K lässt sich auf $L^2[a,b]$ erweitern so dass für jedes $f \in L^2[a,b]$ die Funktion Kf stetig ist und (3.59) erfüllt.*

(*b*) *Der erweiterte Operator $K : L^2[a,b] \to L^2[a,b]$ ist kompakt.*

Beweis. (*a*) Betrachten wir zwei Banachräume: $X = L^2[a,b]$ und $Y = C[a,b]$. Im X betrachten wir den Unterraum $U = C[a,b]$. Obwohl die Vektorräume U und Y identisch sind, die Normen in U und Y unterschiedlich: in U benutzen wir die 2-Norm wie in X, und in Y – die sup-Norm.

Nach (3.59) hat der Operator $K : U \to Y$ die Norm $\|K\| \le C$, insbesondere ist K beschränkt. Da U dicht in X liegt (Satz 2.16) und Y vollständig ist (Satz 1.7), so gibt es eine eindeutige Erweiterung von K auf X mit Erhaltung der Norm (siehe Aufgabe 61).

Der erweiterte Operator $X \to Y$ wird auch mit K bezeichnet, d.h.

$$K : L^2[a,b] \to C[a,b],$$

und (3.59) gilt für alle $f \in L^2[a,b]$.

(*b*) Für jedes $x \in [a,b]$ betrachten wir die lineare Abbildung

$$\begin{aligned} K_x : L^2[a,b] &\to \mathbb{R} \\ f &\mapsto Kf(x). \end{aligned}$$

Die Bedingung (3.59) ergibt, dass für jedes $x \in [a,b]$

$$|K_x f| \le C \|f\|_2 ,$$

d.h. K_x ein stetiges lineares Funktional auf $L^2[a,b]$ ist. Nach dem Rieszschen Darstellungssatz 2.7 gibt es ein $k_x \in L^2[a,b]$ mit

$$K_x f = (f, k_x) \quad \forall f \in L^2[a,b], \tag{3.60}$$

d.h.

$$Kf(x) = K_x f = \int_{[a,b]} k_x(y) f(y)\, d\lambda(y) = \int_{[a,b]} k(x,y) f(y)\, d\lambda(y),$$

wobei $k(x,y) = k_x(y)$. Somit ist K ein Integraloperator mit dem Kern $k(x,y)$.

Die Bedingung (3.59) lässt sich wie folgt umschreiben:

$$|(f, k_x)| = |Kf(x)| \le C \|f\|_2$$

für alle $f \in L^2[a,b]$ und $x \in [a,b]$. Einsetzen hier $f = k_x$ ergibt

$$\|k_x\|_2^2 \le C \|k_x\|_2$$

und somit

$$\|k_x\|_2 \le C.$$

Es folgt, dass

$$\begin{aligned}\int_{[a,b]\times[a,b]} k^2(x,y)\,d\lambda_2(x,y) &= \int_{[a,b]}\left(\int_{[a,b]} k_x^2(y)\,d\lambda(y)\right)d\lambda(x)\\ &= \int_{[a,b]} \|k_x\|^2\,d\lambda(x) \le C^2(b-a) < \infty.\end{aligned}$$

Nach dem Satz 3.14 ist der Integraloperator K mit dem Kern $k(x,y)$ kompakt. ■

22.06.22 Vorlesung 20

Wir betrachten wieder das Randwertproblem (3.44) auf einem Intervall $[a,b]$:

$$\begin{cases} Lu = f \\ u(a) = u(b) = 0 \end{cases} \tag{3.61}$$

wobei $Lu = -(pu')' + qu$ und

$$0 < p \in C^1[a,b] \text{ und } 0 \le q \in C[a,b]. \tag{3.62}$$

Nach den Satz 3.17, für jedes $f \in C[a,b]$ gibt es eine eindeutige Lösung $u \in C^2[a,b]$ von (3.61). Somit erhalten wir den Greenschen Operator

$$\begin{aligned} K &: C[a,b] \to C[a,b] \\ Kf &= u \end{aligned}$$

Nach dem Lemma 3.18 (und (3.53)) gilt für jedes $f \in C[a,b]$ die folgende Ungleichung:

$$\sup_{[a,b]} |Kf| \le C\,\|f\|_2, \tag{3.63}$$

mit einer Konstante C.

Satz 3.20 *Unter den Voraussetzungen (3.62) gelten die folgenden Aussagen.*

(a) *Der Greensche Operator K von (3.61) lässt sich auf $L^2[a,b]$ erweitern, so dass für jedes $f \in L^2[a,b]$ die Funktion Kf stetig ist. Darüber hinaus ist der erweiterte Operator kompakt in $L^2[a,b]$.*

(b) *K ist selbstadjungiert in $L^2[a,b]$.*

(c) *$(Kf,f)_{L^2} > 0$ für alle $f \in C[a,b]\setminus\{0\}$.*

(d) *$\ker K = \{0\}$ in $L^2[a,b]$.*

Beweis. (a) Nach (3.63) erfüllt K die Voraussetzungen von Lemma 3.19, woraus alle Aussagen folgen. Weiter betrachten wir K als einen Operator in $L^2[a,b]$.

(b) Wir müssen beweisen, dass für alle $f, g \in L^2[a,b]$ gilt

$$(Kf, g) = (f, Kg)\,. \tag{3.64}$$

Da $C[a,b]$ dicht in $L^2[a,b]$ liegt, so reicht es (3.64) für $f, g \in C[a,b]$ zu beweisen. Setzen wir

$$u = Kf \quad \text{und} \quad v = Kg.$$

Dann sind u und v die Lösungen des Randwertproblems, insbesondere,

$$Lu = f \quad \text{und} \quad Lv = g,$$

und es gelten nach (3.45) die Identitäten

$$(f, Kg) = (Lu, v) = \int_a^b (pu'v' + quv)\, dx \tag{3.65}$$

und

$$(Kf, g) = (u, Lv) = (Lv, u) = \int_a^b (pv'u' + qvu)\, dx. \tag{3.66}$$

Offensichtlich sind die rechten Seiten von (3.66) und (3.65) gleich, woraus (3.64) folgt.

(c) Setzen wir $u = Kf$. Nach (3.66) gilt

$$(Kf, f) = \int_a^b \left(p\,(u')^2 + qu^2\right) dx,$$

woraus folgt $(Kf, f) \geq 0$. Gilt $(Kf, f) = 0$, so muss u' identisch 0 sein, woraus folgt $u \equiv 0$ und somit $f = Lu \equiv 0$, was im Widerspruch zu $f \not\equiv 0$ steht. Somit gilt $(Kf, f) > 0$.

(d) Nach dem Satz 3.6 gilt

$$(\ker K)^\perp = (\ker K^*)^\perp = \overline{\operatorname{im} K}.$$

Für jedes $u \in C_0^2[a,b]$ setzen wir $f = Lu \in C[a,b]$ so dass u die Lösung des Randwertproblems (3.44) mit dieser Funktion f ist, d.h. $u = Kf$. Somit erhalten wir, dass $u \in \operatorname{im} K$ und folglich

$$C_0^2[a,b] \subset \operatorname{im} K. \tag{3.67}$$

Wir wissen, dass $C_0^2[a,b]$ dicht in $L^2[a,b]$ liegt (siehe Satz 2.16 und Bemerkung danach). Es folgt aus (3.67), dass auch $\operatorname{im} K$ dicht in $L^2[a,b]$ liegt, d.h.

$$\overline{\operatorname{im} K} = L^2[a,b]\,,$$

und somit

$$(\ker K)^\perp = L^2[a,b]$$

und $\ker K = \{0\}$. ■

3.10 Sturm-Liouville-Problem

Für den Operator

$$L = -\left(pu'\right)' + qu$$

auf $[a, b]$ wie oberhalb betrachten wir das *Sturm-Liouville-Problem*

$$\begin{cases} Lv = \lambda v, \\ v(a) = v(b) = 0, \end{cases} \tag{3.68}$$

wobei $v \in C^2[a, b]$ eine unbekannte Funktion ist und λ eine unbekannte Konstante, die *spektraler Parameter* heißt. Offensichtlich ist $v \equiv 0$ eine Lösung von (3.68) für alle λ.

Definition. Jede nicht-Null Lösung $v \in C^2[a, b]$ von (3.68) heißt *Eigenfunktion* von (3.68) und der entsprechende Wert λ heißt *Eigenwert* von (3.68).

Die *Vielfachheit* von λ is die maximale Anzahl von linear unabhängigen Eigenfunktionen mit dem Eigenwert λ. Ein Eigenwert λ heißt einfach wenn die Vielfachheit von λ gleich 1 ist.

Das Problem ist alle Eigenwerte und Eigenfunktionen von (3.68) zu bestimmen. Dieses Problem heißt auch das *Eigenwertproblem.*

Hauptsatz 3.21 (Satz von Sturm-Liouville) *Seien* $0 < p \in C^1[a, b]$ *und* $0 \le q \in C[a, b]$. *Dann gelten die folgenden Aussagen.*

(a) *Alle Eigenwerte von (3.68) sind einfach. Die Menge von allen Eigenwerten von (3.68) lässt sich als eine Folge* $\{\lambda_n\}_{n=1}^{\infty}$ *nummerieren so dass diese Folge reell, positiv, streng monoton steigend ist und es gilt* $\lambda_n \to \infty$ *für* $n \to \infty$.

(b) *Die entsprechende Folge* $\{v_n\}_{n=1}^{\infty}$ *von Eigenfunktionen ist eine Orthogonalbasis in* $L^2[a, b]$.

Beispiel. Betrachten wir das Sturm-Liouville-Problem für den Operator $Lu = -u''$ auf $[0, \pi]$:

$$\begin{cases} v'' + \lambda v = 0, \\ v(0) = v(\pi) = 0 \end{cases} \tag{3.69}$$

Für $\lambda > 0$ die allgemeine Lösung der Gleichung $v'' + \lambda v = 0$ ist

$$v(x) = C_1 \cos \sqrt{\lambda} x + C_2 \sin \sqrt{\lambda} x.$$

Die Bedingung $v(0) = 0$ ergibt $C_1 = 0$, und die Bedingung $v(\pi) = 0$ ergibt $\sin \sqrt{\lambda}\pi = 0$, was äquivalent zu $\sqrt{\lambda} \in \mathbb{Z}$ ist. Somit erhalten wir alle Eigenwerte

$$\lambda_n = n^2, \quad n = 1, 2, ...,$$

und die entsprechenden Eigenfunktionen sind

$$v_n(x) = \sin nx.$$

Nach dem Satz 3.21 beschließen wir, dass die Folge $\{\sin nx\}_{n=1}^{\infty}$ eine Orthogonalbasis in $L^2[0, \pi]$ ist.

Bemerkung. Im Satz 3.21 wird es vorausgesetzt wie zuvor, dass $p > 0$ und $q \geq 0$ auf $[a,b]$. Allerdings lässt der Satz 3.21 sich auf beliebige (signierte) Funktionen $q \in C[a,b]$ erweitern (wobei p positiv bleibt). In der Tat ist die Gleichung

$$-(pv')' + qv = \lambda v$$

äquivalent zu

$$-(pv')' + (q+c)v = (\lambda + c)v$$

für beliebige Konstante $c \in \mathbb{R}$, d.h. v ist eine Eigenfunktion von L mit dem Eigenwert λ genau dann, wenn v eine Eigenfunktion von dem Operator

$$L_c = -(pu')' + (q+c)u$$

mit dem Eigenwert $\lambda + c$ ist. Für hinreichend grosses $c > 0$ gilt $q + c > 0$, so dass der Satz 3.21 sich auf den Operator L_c anwenden lässt. Somit erhalten wir, dass die Eigenwerte und Eigenfunktionen von L die gleichen Eigenschaften haben, außer Positivität: die Aussage $\lambda_n > 0$ soll in diesem Fall durch $\lambda_n + c > 0$ ersetzt werden.

Bemerkung. Es gibt noch weitere Informationen über v_n und λ_n: jede Eigenfunktion v_n hat genau $n-1$ Nullstellen in (a,b), und die Eigenwerte erfüllen die folgende *Weyl-Asymptotik*:

$$\lambda_n \sim cn^2 \text{ für } n \to \infty, \tag{3.70}$$

wobei $c = \pi^2 \left(\int_a^b \frac{dx}{\sqrt{p(x)}} \right)^{-2}$.

Für das obige Beispiel mit $v_n = \sin nx$ auf dem Intervall $[0,\pi]$ und $\lambda_n = n^2$ sind diese Eigenschaften offensichtlich: die Funktion $\sin nx$ hat in $(0,\pi)$ genau $n-1$ Nullstellen $x = \frac{\pi k}{n}$, $k = 1, 2, ..., n-1$, und λ_n erfüllt (3.70) mit $c = 1$.

Beweis von dem Satz 3.21. Wir fangen mit dem Beweis von (b) an.

(b) Wir beweisen zuerst die Existenz einer Orthogonalbasis in $L^2[a,b]$ die aus Eigenfunktionen besteht. Dafür verwenden wir den Greenschen Operator $K : L^2[a,b] \to L^2[a,b]$: für jede Funktion $f \in L^2[a,b]$ ist die Funktion $u = Kf$ stetig, und wenn $f \in C[a,b]$ so liegt u in $C^2[a,b]$ und löst das Randwertproblem (3.44) d.h.

$$\begin{cases} Lu = f \\ u(a) = u(b) = 0 \end{cases}. \tag{3.71}$$

Sei $v \in C^2[a,b]$ eine Lösung von (3.68), d.h.

$$\begin{cases} Lv = \lambda v \\ v(a) = v(b) = 0 \end{cases}. \tag{3.72}$$

Dann lösst v das Randwertproblem (3.71) mit $f = \lambda v$, so dass

$$v = \lambda K v. \tag{3.73}$$

Beweisen wir die Umkehrung davon: jede Funktion $v \in L^2[a,b]$, die (3.73) erfüllt, liegt in $C^2[a,b]$ und löst (3.72). In der Tat ist die Funktion Kv stetig (Satz 3.20), und es folgt aus (3.73), dass auch v stetig ist; daraus folgt, dass $Kv \in C^2[a,b]$ und nach (3.73) auch

$v \in C^2[a,b]$ (das *Bootstrap-Argument*). Folglich löst v (3.71) mit $f = \lambda v$, was äquivalent zu (3.72) ist.

Da K kompakt und selbstadjungiert ist (Satz 3.20) und $L^2[a,b]$ separabel ist, so gibt es nach dem Satz von Hilbert-Schmidt (Satz 3.15) eine Orthogonalbasis $\{v_n\}_{n=1}^{\infty}$ in $L^2[a,b]$, die aus den Eigenfunktionen von K besteht.

Sei α_n der Eigenwert von v_n, d.h.

$$Kv_n = \alpha_n v_n.$$

Da $\ker K = \{0\}$ (Satz 3.20), so gilt $\alpha_n \neq 0$, woraus folgt

$$v_n = \frac{1}{\alpha_n} K v_n.$$

Folglich liegt v_n in $C^2[a,b]$ und v_n ist eine Eigenfunktion von (3.72) mit dem Eigenwert $\lambda_n = \frac{1}{\alpha_n}$.

(*a*) Zeigen wir zuerst, dass die Folge $\{\lambda_n\}_{n=1}^{\infty}$ alle Eigenwerte von (3.72) enthält. Sei λ ein Eigenwert von (3.72), der nicht in der Folge $\{\lambda_n\}$ liegt, und sei v eine Eigenfunktion von λ. Da $\lambda \neq \lambda_n$, so gilt $v \perp v_n$ in $L^2[a,b]$ (siehe Aufgabe 79). Somit ist v orthogonal zu allen Elementen v_n der Basis, woraus folgt dass $v = 0$ und λ kein Eigenwert ist.

Da K selbstadjungiert ist, so sind alle α_n und somit auch $\boxed{\lambda_n \text{ reell}}$.

Zeigen wir, dass $\boxed{\lambda_n > 0}$. Nach dem Satz 3.20 gilt $(Kf, f) > 0$ für alle nicht-Null stetige Funktionen f. Insbesondere gilt

$$(Kv_n, v_n) > 0$$

was äquivalent zu

$$\alpha_n (v_n, v_n) > 0$$

und somit $\alpha_n > 0$, woraus folgt $\lambda_n > 0$.

Da $\alpha_n \to 0$ (Satz 3.15), so erhalten wir $\boxed{\lambda_n \to \infty}$.

Jetzt beweisen wir, dass jeder Eigenwert λ von (3.72) $\boxed{\text{einfach}}$ ist. Seien v_1 und v_2 zwei Eigenfunktionen von λ. Zeigen wir, dass v_1 und v_2 linear abhängig sind. Da $v_1(a) = 0$, so muss $v_1'(a) \neq 0$ erfüllt werden, da sonst nach dem Satz von Picard-Lindelöf gilt $v_1 \equiv 0$. Somit existiert eine Konstante c mit $v_2'(a) = cv_1'(a)$. Dann erfüllt die Funktion $v = v_1 - cv_2$ die Differentialgleichung $Lv = \lambda v$ mit $v(a) = v'(a) = 0$, woraus folgt dass $v \equiv 0$. Somit sind die Lösungen v_1 und v_2 linear abhängig, was zu beweisen war.

Insbesondere sind alle Elemente der Folge $\{\lambda_n\}$ verschieden. Da $\lambda_n \to \infty$, so lässt die Folge $\{\lambda_n\}$ (und somit auch $\{v_n\}$) sich so umnummerieren, dass die neue Folge $\{\lambda_n\}$ streng monoton steigend ist. Dafür bezeichnen wir mit Λ die Menge von allen Eigenwerten $\{\lambda_n\}$ und bemerken dass Λ ein minimales Element hat, was aus $\lambda_n \to \infty$ folgt. Wählen wir das minimale Element in Λ und benennen es in λ_1 um. Die Menge $\Lambda \setminus \{\lambda_1\}$ hat auch das minimale Element, benennen wir es in λ_2 um, usw. Somit erhalten wir eine neue Nummerierung $\{\lambda_n\}$ der Folge von Eigenwerten so dass die neue Folge $\boxed{\text{streng monoton steigend}}$ ist. ■

24.06.22 Vorlesung 21

Mit Hilfe von dem Satz 3.21 geben wir einen alternativen Beweis von dem Satz 2.23 ohne den 2.Approximationssatz von Weierstraß zu benutzen.

2.Beweis von dem Satz 2.23. Wir beweisen, dass die Folge von trigonometrischen Monomen

$$1, \cos x, \sin x, \cos 2x, \sin 2x, ..., \cos nx, \sin nx, ... \tag{3.74}$$

eine Orthogonalbasis in $L^2[-\pi,\pi]$ ist. Wie zuvor, bezeichnen wir mit $\mathcal{T}$ die Menge von allen trigonometrischen Polynomen. Da die Folge (3.74) orthogonal ist (siehe Aufgabe 41), so reicht es zu beweisen, dass die Folge (3.74) ein Erzeugendensystem in $L^2[-\pi,\pi]$ ist, d.h.

$$\mathcal{T} \precapprox L^2[-\pi,\pi].$$

Da

$$C^1[-\pi,\pi] \precapprox L^2[-\pi,\pi]$$

so reicht es zu beweisen, dass

$$\mathcal{T} \precapprox C^1[-\pi,\pi] \tag{3.75}$$

bezüglich der L^2-Norm. Wir beweisen, dass jede Funktion $f \in C^1[-\pi,\pi]$ eine Darstellung durch eine Fourier-Reihe zulässt:

$$f(x) = a_0 + \sum_{n=1}^{\infty}(a_n \cos nx + b_n \sin nx),$$

wobei die Reihe in $L^2[-\pi,\pi]$ konvergiert, woraus (3.75) folgt.

Sei f ungerade. Nach dem obigen Beispiel ist die Folge $\{\sin nx\}_{n=1}^{\infty}$ eine Orthogonalbasis in $L^2[0,\pi]$ und somit gibt es eine Darstellung von f in dieser Basis:

$$f(x) = \sum_{n=1}^{\infty} b_n \sin nx,$$

wobei die Reihe in $L^2[0,\pi]$ konvergiert. Da $\sin nx$ auch ungerade ist, so gilt diese Darstellung auch in $L^2[-\pi,\pi]$.

Sei f gerade. Dann ist f' ungerade und somit gibt es eine Darstellung

$$f'(x) = \sum_{n=1}^{\infty} c_n \sin nx,$$

wobei die Reihe in $L^2[-\pi,\pi]$ konvergiert. Integrieren von dieser Reihe ergibt

$$f(x) = f(0) + \int_0^x f'(t)\,dt = f(0) + \sum_{n=1}^{\infty} c_n \left[-\frac{\cos nt}{n}\right]_0^x,$$

wobei die Reihe wieder in $L^2[-\pi,\pi]$ konvergiert (nach Aufgabe 53 ist Integrieren mit L^2-Limes vertauschbar). We haben

$$c_n \left[-\frac{\cos nt}{n}\right]_0^x = \frac{c_n}{n} - c_n \frac{\cos nx}{n}$$

Die Reihe $\sum_{n=1}^{\infty} \frac{c_n}{n}$ ist konvergent da die Folgen $\{c_n\}$ und $\left\{\frac{1}{n}\right\}$ in l^2 liegen, woraus folgt

$$f(x) = a_0 + \sum_{n=1}^{\infty} a_n \cos nx \tag{3.76}$$

wobei $a_0 = f(0) + \sum_{n=1}^{\infty} \frac{c_n}{n}$ und $a_n = -\frac{c_n}{n}$ für $n \geq 1$, und die Reihe (3.76) in $L^2[-\pi, \pi]$ konvergiert.

Sei jetzt f eine beliebige Funktion aus $C^1[-\pi, \pi]$. Dann gilt eine Zerlegung

$$f = f_0 + f_1$$

wobei $f_0, f_1 \in C^1[-\pi, \pi]$, f_0 gerade und f_1 ungerade ist, z.B.

$$f_0(x) = \frac{f(x) + f(-x)}{2} \quad \text{und} \quad f_1(x) = \frac{f(x) - f(-x)}{2}.$$

Somit erhalten wir

$$f_0(x) = a_0 + \sum_{n=1}^{\infty} a_n \cos nx, \qquad f_1(x) = \sum_{n=1}^{\infty} b_n \sin nx$$

und

$$f(x) = a_0 + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx),$$

wobei alle Reihen in $L^2[-\pi, \pi]$ konvergieren. ■

3.11 Schrödinger-Gleichung

In Quantenmechanik beschreibt man die Bewegung von einem Elementarteilchen in einem elektrischen Feld mit Hilfe von der Schrödinger-Gleichung. Die eindimensionale Schrödinger-Gleichung sieht wie folgt aus:

$$i\frac{\partial \psi}{\partial t} = -\frac{\partial^2 \psi}{\partial x^2} + V(x)\psi, \tag{3.77}$$

wobei $V(x)$ eine gegebene reellwertige stetige Funktion auf einem Intervall $[a, b]$ ist (elektrisches Potential des elektrischen Felds), und $\psi(x, t)$ eine unbekannte komplexwertige Funktion ist, die von der Position $x \in [a, b]$ und von der Zeit $t \in \mathbb{R}$ abhängt. Die Funktion ψ heißt die *Wellenfunktion* des Teilchens.

Die Wellenfunktion hat die folgende physikalische Bedeutung: $|\psi(x, t)|^2$ ist die Wahrscheinlichkeitsdichte von dem Teilchen an Position x zum Zeitpunkt t zu sein; d.h. die Position X_t des Teilchens zum Zeitpunkt t ist eine Zufallsvariable, und es gilt für jedes Intervall $J \subset [a, b]$

$$\mathbb{P}(X_t \in J) = \int_J |\psi(x, t)|^2 \, dx.$$

Die Schrödinger-Gleichung erhält man mit Hilfe von *Quantisierung* der Identität für die Gesamtenergie

$$E = \frac{p^2}{2m} + V,$$

wobei V die potentielle Energie und $\frac{p^2}{2m}$ die kinetische Energie sind, und p der Impuls ist. Nach Einsetzen

$$E = i\frac{\partial}{\partial t} \quad \text{und} \quad p = i\frac{\partial}{\partial x}$$

und Anwendung auf ψ erhält man (3.77) (im Fall $m = 1/2$).

Erlegen wir die folgende Randbedingung und Anfangsbedingung auf:

$$\psi(a,t) = \psi(b,t) = 0 \quad \text{für alle } t \in \mathbb{R}, \tag{3.78}$$

und

$$\psi(x,0) = \psi_0(x), \tag{3.79}$$

wobei ψ_0 eine gegebene Funktion ist. Die Randbedingung (3.78) bedeutet, dass das Teilchen im Intervall $[a,b]$ eingesperrt ist. Wir betrachten die Anfangsfunktion als bekannt, und die Aufgabe ist die Wellenfunktion $\psi(x,t)$ für alle $t > 0$ zu bestimmen.

Um das Problem (3.77)-(3.78)-(3.79) zu lösen, verwenden wir die Methode der *Trennung der Variablen.* Dafür betrachten wir zuerst der Differentialoperator

$$Lv = -v'' + V(x)\,v,$$

der auch der *Hamiltonoperator* des physikalischen Systems heißt, und das entsprechende Strum-Liouville-Problem auf $[a,b]$, d.h.

$$\begin{cases} Lv = \lambda v \\ v(a) = v(b) = 0. \end{cases}$$

Nach dem Satz 3.21 existiert eine Orthonormalbasis $\{v_n\}_{n=1}^{\infty}$ in $L^2[a,b]$, die aus den Eigenfunktionen von L besteht. Sei λ_n der Eigenwert von v_n. Dann erfüllt die Funktion

$$\psi_n(x,t) = v_n(x)\,e^{-i\lambda_n t}$$

die Randbedingung (3.78) und die Schrödinger-Gleichung (3.77) da

$$i\frac{\partial \psi_n}{\partial t} = i(-i\lambda_n)\,v_n(x)\,e^{-i\lambda_n t} = (Lv_n)\,e^{-i\lambda_n t} = L\psi_n = -\frac{\partial^2 \psi}{\partial x^2} + V(x)\,\psi.$$

Die Lösung des Problems (3.77)-(3.78)-(3.79) suchen wir in der Form

$$\psi(x,t) = \sum_{n=1}^{\infty} c_n \psi_n(x,t) = \sum_{n=1}^{\infty} c_n v_n(x)\,e^{-i\lambda_n t}, \tag{3.80}$$

wobei die Koeffizienten $c_n \in \mathbb{C}$ noch unbekannt sind. Für $t = 0$ erhalten wir die Gleichung

$$\psi_0(x) = \sum_{n=1}^{\infty} c_n v_n(x), \tag{3.81}$$

woraus c_n bestimmt werden können:

$$c_n = (\psi_0, v_n).$$

Die Wellenfunktion als eine Wahrscheinlichkeitsdichte muss noch die folgende Bedingung erfüllen:

$$\int_{[a,b]} |\psi(x,t)|^2 \, dx = 1, \tag{3.82}$$

da das Integral gleich die Gesamtwahrscheinlichkeit $\mathbb{P}(X_t \in [a,b])$ ist. So nehmen wir an, dass

$$\int_{[a,b]} |\psi_0|^2 \, dx = 1,$$

insbesondere $\psi_0 \in L^2[a,b]$. Nach der Parsevalschen Gleichung gilt

$$\sum_{n=1}^{\infty} |c_n|^2 = \|\psi_0\|_2^2 = 1$$

und somit auch für alle $t \in \mathbb{R}$

$$\sum_{n=1}^{\infty} \left| c_n e^{-i\lambda_n t} \right|^2 = \sum_{n=1}^{\infty} |c_n|^2 = 1.$$

Daraus folgt, dass die Reihe (3.80) für jedes $t \in \mathbb{R}$ in $L^2[a,b]$ konvergiert, und für die Summe $\psi(x,t)$ gilt

$$\|\psi(\cdot,t)\|_2^2 = 1,$$

d.h. (3.82).

Man kann beweisen dass unter bestimmten weiteren Voraussetzungen über ψ_0 die Funktion ψ aus (3.80) 2-fach differenzierbar ist und die Gleichung (3.77) erfüllt.

Analog löst man andere partielle Differentialgleichungen, z.B. die *Wärmeleitungsgleichung*

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}$$

und die *Wellengleichung*

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2}$$

3.12 Weitere Beispiele

Beispiel. Der *harmonische Oszillator* in der Quantenmechanik hat den Hamiltonoperator

$$Lv = -v'' + x^2 v$$

auf $\mathbb{R}$. In diesem Fall gibt es keine Randbedingungen, aber man fordert $v \in L^2(\mathbb{R})$ an. Das entsprechende Eigenwertproblem ist wie folgt:

$$v'' - x^2 v + \lambda v = 0, \quad v \in L^2(\mathbb{R}).$$

Man kann zeigen, dass die Eigenfunktionen dieses Problems sind

$$v_n(x) = e^{-x^2/2} H_n(x), \quad n = 0, 1, 2, ...$$

wobei H_n die Hermite-Polynome sind, und die entsprechenden Eigenwerte sind

$$\lambda_n = 2n + 1$$

(siehe auch die Aufgabe 57). Man kann auch beweisen, dass die Folge $\{v_n\}$ eine Orthogonalbasis in $L^2(\mathbb{R})$ ist (eine Verallgemeinerung des Hauptsatzes 3.21).

Beispiel. Sei Ω eine beschränkte offene Teilmenge von $\mathbb{R}^n$. Betrachten wir den *Laplace-Operator* Δ der auf Funktionen $f \in C^2(\Omega)$ wie folgt wirkt:

$$\Delta = \sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_j^2}.$$

Für das entsprechende Eigenwertproblem

$$\begin{cases} \Delta v + \lambda v = 0 \text{ in } \Omega \\ v|_{\partial\Omega} = 0 \end{cases} \tag{3.83}$$

kann man folgendes beweisen: es gibt eine Orthogonalbasis $\{v_k\}_{k=1}^{\infty}$ in $L^2(\Omega)$ die aus den Eigenfunktionen von (3.83) besteht, die Folge $\{\lambda_k\}_{k=1}^{\infty}$ von den entsprechenden Eigenwerten ist monoton steigend, und es gelten $\lambda_k > 0$ und $\lambda_k \to +\infty$ für $k \to \infty$ (die Eigenwerte müssen jedoch nicht unbedingt einfach sein). Für den Beweis konstruiert man den entsprechenden Greenschen Operator und zeigt, dass dieser Operator selbstadjungiert und kompakt in $L^2(\Omega)$ ist.

Die Schrödinger-Gleichung für ein Teilchen in Ω sieht wie folgt aus:

$$i\frac{\partial\psi}{\partial t} = -\Delta\psi + V(x)\,\psi,$$

wobei $V(x)$ eine gegebene Funktion in Ω ist und $\psi = \psi(x,t)$ die (unbekannte) Wellenfunktion ist. Diese Gleichung mit den Rand- und Anfangsbedingungen

$$\psi(x,t)\,|_{x\in\partial\Omega} = 0 \quad \text{und} \quad \psi(x,0) = \psi_0(x)$$

lässt sich auch mit Hilfe von Methode der Trennung der Variable lösen, und dafür verwendet man eine Verallgemeinerung des Satzes von Sturm-Liouville, dass das Eigenwertproblem

$$\begin{cases} \Delta v - V(x)\,v + \lambda v = 0 \text{ in } \Omega \\ v|_{\partial\Omega} = 0 \end{cases}$$

eine Orthogonalbasis in $L^2(\Omega)$ aus den Eigenfunktionen liefert.

Chapter 4

Funktionalkalkül von selbstadjungierten Operatoren

In diesem Kapitel definieren wir den Begriff $f(A)$ wobei A ein selbstadjungierter Operator ist und f eine stetige Funktion auf $\mathbb{R}$ ist.

4.1 Polynome von Operatoren

Sei A ein Operator in einem Vektorraum V über $\mathbb{C}$ und $f(z)$ ein Polynom mit komplexwertigen Koeffizienten:

$$f(z) = c_0 + c_1 z + ... + c_n z^n.$$

Definieren wir $f(A)$ als einen Operator in V mit

$$f(A) = c_0 I + c_1 A + ... + c_n A^n. \tag{4.1}$$

Es ist offensichtlich, dass für zwei Polynome f, g die folgenden Identitäten gelten:

(i) $(f+g)(A) = f(A) + g(A)$

(ii) $(fg)(A) = f(A)\, g(A)$.

Ist V ein normierter Vektorraum, dann $A \in \mathcal{B}(V) \Rightarrow f(A) \in \mathcal{B}(V)$.

Seien A ein selbstadjungierter Operator im Hilbertraum H und $f(z)$ ein Polynom mit reellwertigen Koeffizienten. Dann ist $f(A)$ auch selbstadjungiert.

In diesem Kapitel definieren wir $f(A)$ für selbstadjungierte Operatoren auch wenn f eine beliebige stetige Funktion ist. Diese Theorie heißt das *Funktionalkalkül* von selbstadjungierten Operatoren.

Wir fangen mit dem folgenden Satz an.

Satz 4.1 (Der spektrale Abbildungssatz für Polynome) *Seien A ein beschränkter Operator im normierten Vektorraum V und $f(z)$ ein Polynom. Dann gilt*

$$\sigma(f(A)) = f(\sigma(A)) := \{f(z) : z \in \sigma(A)\}. \tag{4.2}$$

Der Beweis ist unterhalb.

Korollar 4.2 *Seien A ein selbstadjungierter Operator im Hilbertraum H und $f(z)$ ein Polynom mit reellwertigen Koeffizienten. Dann gilt*

$$\|f(A)\| = \sup_{z\in\sigma(A)} |f(z)|. \tag{4.3}$$

Beweis. Nach dem Satz 3.10 gilt für jeden selbstadjungierten Operator B in H

$$\|B\| = \sup |\sigma(B)|. \tag{4.4}$$

Da $f(A)$ selbstadjungiert ist, setzen wir hier $B = f(A)$ und erhalten mit Hilfe von (4.2), dass

$$\|f(A)\| = \sup |\sigma(f(A))| = \sup |f(\sigma(A))| = \sup_{z\in\sigma(A)} |f(z)|.$$

■

Beispiel. Angenommen, dass $\sigma(A) \subset [0,1]$, schätzen wir die Norm $\|f(A)\|$ für das Polynom $f(z) = z - z^2$ ab. Nach (4.3) gilt

$$\left\|A - A^2\right\| = \sup_{z\in\sigma(A)} \left|z - z^2\right| \le \sup_{z\in[0,1]} \left|z - z^2\right| = \frac{1}{4},$$

d.h.

$$\left\|A - A^2\right\| \le \frac{1}{4}.$$

Diese Ungleichung ist hoch nicht-trivial sogar wenn A eine symmetrische 2×2 Matrix!

29.06.22 Vorlesung 22

Für den Beweis von Satz 4.1 brauchen wir das folgende Lemma.

Lemma 4.3 *Seien A, B zwei beschränkten Operatoren im normierten Vektorraum V mit $AB = BA$. Dann gilt die Äquivalenz*

$$(AB)^{-1} \in \mathcal{B}(V) \quad \Leftrightarrow \quad A^{-1} \in \mathcal{B}(V) \ \text{ und } \ B^{-1} \in \mathcal{B}(V).$$

Beweis. Sind A^{-1} und B^{-1} in $\mathcal{B}(V)$ so gilt auch $B^{-1}A^{-1} \in \mathcal{B}(V)$ und

$$(AB)\left(B^{-1}A^{-1}\right) = A\left(BB^{-1}\right)A^{-1} = AA^{-1} = I.$$

Analog beweist man $(B^{-1}A^{-1})(AB) = I$. Es folgt, dass

$$(AB)^{-1} = B^{-1}A^{-1} \in \mathcal{B}(V).$$

Setzen wir $C = AB$ und beweisen die Umkehrung: ist $C^{-1} \in \mathcal{B}(V)$ so sind A^{-1} und B^{-1} in $\mathcal{B}(V)$. Die Idee ist wie folgt: wenn wir schon wissen, dass A^{-1} und B in $\mathcal{B}(V)$ sind, dann

$$C^{-1}B = (AB)^{-1}B = (BA)^{-1}B = A^{-1}B^{-1}B = A^{-1}.$$

Dies schlägt die folgende Beweismethode vor: angenommen, dass $C^{-1} \in \mathcal{B}(V)$, wir zeigen, dass $C^{-1}B$ die inverse von A ist, d.h.

$$\left(C^{-1}B\right)A = I \quad \text{und} \quad A\left(C^{-1}B\right) = I. \tag{4.5}$$

Die erste Identität in (4.5) folgt direkt aus $AB = BA$:

$$\left(C^{-1}B\right)A = C^{-1}(BA) = C^{-1}(AB) = C^{-1}C = I.$$

Für den Beweis von der zweiten Identität in (4.5) zeigen wir zunächst, dass C mit A (und B) kommutiert:

$$CA = (AB)A = A(BA) = A(AB) = AC.$$

Daraus folgt, dass auch C^{-1} mit A (und B) kommutiert, da $AC = CA$ ergibt

$$C^{-1}(AC)C^{-1} = C^{-1}(CA)C^{-1}$$

d.h.

$$C^{-1}A = AC^{-1}.$$

Jetzt können wir die zweite Identität in (4.5) beweisen:

$$A\left(C^{-1}B\right) = \left(AC^{-1}\right)B = \left(C^{-1}A\right)B = C^{-1}(AB) = I.$$

Aus $A^{-1} = C^{-1}B$ folgt es, dass $A^{-1} \in \mathcal{B}(V)$ und analog gilt $B^{-1} \in \mathcal{B}(V)$. ■

Korollar 4.4 *Seien $A_1, ..., A_n$ beschränkte Operatoren in V die miteinander kommutieren. Dann gilt die Äquivalenz*

$$(A_1...A_n)^{-1} \in \mathcal{B}(V) \Leftrightarrow A_k^{-1} \in \mathcal{B}(V) \;\; \forall k = 1, .., n.$$

Beweis. Beweis per Induktion nach n. Der Fall $n = 1$ ist trivial. Für den Induktionsschritt von n nach $n+1$ erhalten wir nach dem Lemma 4.3 und Induktionsvoraussetzung, dass

$$\begin{aligned}(\underbrace{A_1...A_n}_{A}\underbrace{A_{n+1}}_{B})^{-1} \in \mathcal{B}(V) &\Leftrightarrow (A_1...A_n)^{-1} \in \mathcal{B}(V) \;\text{ und } A_{n+1}^{-1} \in \mathcal{B}(V) \\ &\Leftrightarrow A_1^{-1}, ..., A_n^{-1} \in \mathcal{B}(V) \quad \text{und } A_{n+1}^{-1} \in \mathcal{B}(V),\end{aligned}$$

was zu beweisen war. ■

Beweis von dem Satz 4.1. Für jedes $\lambda \in \mathbb{C}$ beweisen wir, dass

$$\lambda \in \sigma(f(A)) \Leftrightarrow \lambda \in f(\sigma(A)).$$

Nach dem Fundamentalsatz der Algebra lässt sich das Polynom $f(z) - \lambda$ in ein Produkt von Linearfaktoren zerlegen:

$$f(z) - \lambda = c(z - z_1)...(z - z_n),$$

wobei $z_1, ..., z_n$ die (komplexwertigen) Nullstellen von $f(z) - \lambda$ sind. Dann auch gilt

$$f(A) - \lambda I = c(A - z_1 I)...(A - z_n I).$$

Nach dem Korollar 4.4 haben wir

$$\begin{aligned}\lambda \in \rho(f(A)) &\Leftrightarrow (f(A) - \lambda I)^{-1} \in \mathcal{B}(H) \\ &\Leftrightarrow ((A - z_1 I)...(A - z_n I))^{-1} \in \mathcal{B}(H) \\ &\Leftrightarrow (A - z_k I)^{-1} \in \mathcal{B}(H) \;\; \forall k = 1, ..., n \\ &\Leftrightarrow z_k \in \rho(A) \;\; \forall k = 1, ..., n\end{aligned}$$

Daraus folgt

$$\begin{aligned}\lambda \in \sigma(f(A)) &\Leftrightarrow \exists k \quad z_k \in \sigma(A)\\ &\Leftrightarrow \exists z \in \sigma(A) \text{ mit } f(z) - \lambda = 0\\ &\Leftrightarrow \lambda = f(z) \text{ für ein } z \in \sigma(A)\\ &\Leftrightarrow \lambda \in f(\sigma(A)),\end{aligned}$$

was zu beweisen war. ■

4.2 Stetige Funktionen von selbstadjungierten Operatoren

Hauptsatz 4.5 *Sei A ein selbstadjungierter Operator im Hilbertraum H. Dann existiert genau eine Abbildung*

$$\mathcal{A} : C(\sigma(A)) \to \mathcal{B}(H)$$

mit den folgenden Eigenschaften:

1. (Selbstadjungiertheit) *Für jede (reellwertige) Funktion $f \in C(\sigma(A))$ ist $\mathcal{A}(f)$ ein selbstadjungierter Operator.*

2. (Erhaltung der Norm) $\|\mathcal{A}(f)\| = \sup_{\sigma(A)} |f|$ *für alle $f \in C(\sigma(A))$.*

3. (Erhaltung von algebraischen Operationen) *Für alle $f, g \in C(\sigma(A))$ gelten die Identitäten*

$$\begin{aligned}\mathcal{A}(f+g) &= \mathcal{A}(f) + \mathcal{A}(g)\\ \mathcal{A}(cf) &= c\mathcal{A}(f) \quad \forall c \in \mathbb{R}\\ \mathcal{A}(fg) &= \mathcal{A}(f)\,\mathcal{A}(g).\end{aligned}$$

4. (Beziehung zu A) *Für $f \equiv 1$ gilt $\mathcal{A}(f) = I$ und für $f(z) = z$ gilt $\mathcal{A}(f) = A$.*

Der Beweis des Satzes 4.5 befindet sich unterhalb. Die Eigenschaften 2 und 3 bedeuten folgendes: die Abbildung$\mathcal{A}$ ist ein Homomorphismus von den normierten Algebras $C(\sigma(A))$ und $\mathcal{B}(H)$.

Definition. Die Abbildung $\mathcal{A}$ heißt das *Funktionalkalkül* des Operators A. Für jede Funktion $f \in C(\sigma(A))$ setzen wir

$$f(A) := \mathcal{A}(f).$$

Nach dem Satz 4.5 hat $f(A)$ die folgenden Eigenschaften:

1. $f(A)$ ist selbstadjungiert.

2. $\|f(A)\| = \sup_{\sigma(A)} |f|$.

3. Für zwei Funktionen $f, g \in C(\sigma(A))$ gelten

$$\begin{aligned}(f+g)(A) &= f(A) + g(A) \\ (cf)(A) &= c(f(A)) \\ (fg)(A) &= f(A)\,g(A).\end{aligned} \tag{4.6}$$

4. Für $f \equiv 1$ gilt $f(A) = I$ und für $f(z) = z$ gilt $f(A) = A$.

Es folgt aus (4.6), dass für $f(z) = z^k$ gilt $f(A) = A^k$ (wobei $k \in \mathbb{N}$) und somit für ein Polynom

$$f(z) = c_0 + c_1 z + ... + c_n z^n$$

gilt

$$f(A) = c_0 I + c_1 A + ... + c_n A^n,$$

was mit der Definition (4.1) von einem Polynom von A übereinstimmt.

Der Satz 4.5 erweitert somit die Definition $f(A)$ von Polynomen f auf alle stetige Funktionen f auf $\sigma(A)$.

Beispiel. Sei A ein selbstadjungierter Operator mit $\sigma(A) \geq 0$. Zeigen wir, dass es einen selbstadjungierten Operator X mit $X^2 = A$ gibt. Dafür betrachten wir die Funktion $f(z) = \sqrt{z}$ die auf $[0, \infty)$ und somit auch auf $\sigma(A)$ definiert und stetig ist. Nach (4.6) erfüllt der Operator $X = f(A)$ die Identität

$$X^2 = f(A)\,f(A) = f^2(A) = z(A) = A.$$

Somit ist $\sqrt{A}$ wohldefiniert als $f(A)$.

Beispiel. Zeigen wir, dass für jeden selbstadjungierten Operator A der Operator $I + A^2$ beschränkt invertierbar ist. Die Funktion $f(z) = \frac{1}{1+z^2}$ ist auf $\mathbb{R}$ und somit auch auf $\sigma(A)$ stetig so dass der Operator $f(A)$ wohldefiniert ist. Zeigen wir, dass

$$\left(I + A^2\right)^{-1} = f(A).$$

Für die Funktion $g(z) = 1 + z^2$ haben wir $I + A^2 = g(A)$, und es gilt nach (4.6)

$$f(A)\left(I + A^2\right) = f(A)\,g(A) = (fg)(A) = \left(\frac{1}{1+z^2}\left(1+z^2\right)\right)(A) = 1(A) = I$$

und analog

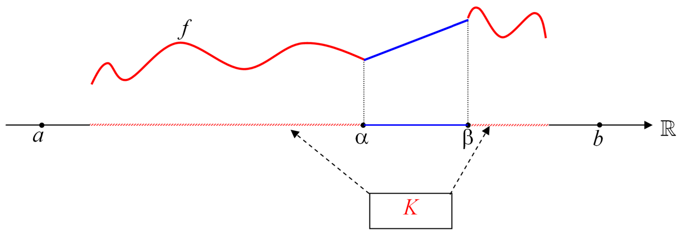
$$\left(I + A^2\right) f(A) = I.$$

Somit ist $f(A)$ der inverse Operator von $I + A^2$. Da $f(A)$ beschränkt ist, so beschließen wir, dass $I + A^2$ beschränkt invertierbar ist.

Für den Beweis des Satzes 4.5 brauchen wir das folgende Lemma.

Lemma 4.6 *Sei K eine kompakte Teilmenge von $\mathbb{R}$. Dann für jede Funktion $f \in C(K)$ existiert eine stetige Fortsetzung von f auf $\mathbb{R}$. Folglich gibt es eine Folge $\{f_n\}_{n=1}^{\infty}$ von Polynomen mit $f_n \rightrightarrows f$ auf K für $n \to \infty$.*

Beweis. Es existiert ein beschränktes offenes Intervall (a,b) das K überdeckt. Definieren wir die Funktion f in $\mathbb{R}\setminus(a,b)$ mit $f\equiv 0$. Die Menge $(a,b)\setminus K$ ist eine offene Menge und deshalb $(a,b)\setminus K$ ist eine disjunkte Vereinigung von offenen Intervallen. Sei (α,β) eines von diesen Intervallen.



Die Endpunkte α,β entweder liegen in K oder stimmen mit a oder b überein, so dass $f(\alpha)$ und $f(\beta)$ schon definiert sind. Dann erweitern wir f auf das ganze Intervall $[\alpha,\beta]$ linear, so dass f stetig auf $\mathbb{R}$ ist.

Nach dem Approximationssatz von Weierstraß (Satz 2.19) gibt es eine Folge $\{f_n\}$ von Polynomen, die gegen f auf $[a,b]$ gleichmäßig konvergiert, woraus auch $f_n\rightrightarrows f$ auf K folgt. ■

Beweis von Hauptsatz 4.5. Beweisen wir zunächst die Eindeutigkeit des Funktionalkalküls $\mathcal{A}$. Da $\mathcal{A}(1)=I$, $\mathcal{A}(z)=A$ und folglich $\mathcal{A}\left(z^k\right)=A^k$ für alle $k\in\mathbb{N}$, so erhalten wir, dass für jedes Polynom

$$f(z)=c_0+c_1z+...+c_nz^n$$

gilt

$$\mathcal{A}(f)=c_0I+c_1A+...+c_nA^n. \tag{4.7}$$

Insbesondere ist $\mathcal{A}(f)$ eindeutig bestimmt für Polynome f. Sei f eine beliebige Funktion aus $C(\sigma(A))$. Da $\sigma(A)$ kompakt (beschränkt und abgeschlossen) ist, so gibt es nach dem Lemma 4.6 eine Folge $\{f_n\}$ von Polynomen mit $f_n\rightrightarrows f$ auf $\sigma(A)$.

Dann gilt

$$\|\mathcal{A}(f_n)-\mathcal{A}(f)\|=\|\mathcal{A}(f_n-f)\|=\sup_{\sigma(A)}|f_n-f|\to 0$$

d.h. $\mathcal{A}(f_n)\to\mathcal{A}(f)$ in $\mathcal{B}(H)$. Somit ist $\mathcal{A}(f)$ eindeutig bestimmt für alle $f\in C(\sigma(A))$.

Jetzt beweisen wir die Existenz des Funktionalkalküls $\mathcal{A}$. Für Polynome f definieren wir $\mathcal{A}(f)$ mit (4.7), d.h. $\mathcal{A}(f)=f(A)$. Offensichtlich erfüllt diese Definition alle Eigenschaften 1-4 (insbesondere die Eigenschaft 2 gilt nach dem Korollar 4.2).

Für beliebige Funktion $f\in C(\sigma(A))$ finden wir wieder nach Lemma 4.6 eine Folge $\{f_n\}$ von Polynomen mit

$$f_n\rightrightarrows f\quad\text{auf } \sigma(A). \tag{4.8}$$

Die Folge $\{\mathcal{A}(f_n)\}$ von Operatoren ist dann eine Cauchy-Folge, da nach dem Korollar 4.2

$$\|\mathcal{A}(f_n)-\mathcal{A}(f_m)\|=\|f_n(A)-f_m(A)\|=\|(f_n-f_m)(A)\|=\sup_{\sigma(A)}|f_n-f_m|\to 0$$

für $n,m\to\infty$. Da $\mathcal{B}(H)$ ein Banachraum ist, so existiert der Grenzwert $\lim\mathcal{A}(f_n)$, und wir setzen

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

Allerdings brauchen wir noch zu beweisen, dass $\lim \mathcal{A}(f_n)$ unabhängig von der Wahl der Folge $\{f_n\}$ ist. Gegeben sei noch eine Folge $\{g_n\}$ von Polynomen mit

$$g_n \rightrightarrows f \quad \text{auf } \sigma(A)$$

Dann gilt

$$f_n - g_n \rightrightarrows 0 \quad \text{auf} \quad \sigma(A)$$

und somit

$$\|\mathcal{A}(f_n) - \mathcal{A}(g_n)\| = \|(f_n - g_n)(A)\| = \sup_{\sigma(A)} |f_n - g_n| \to 0,$$

woraus folgt $\lim \mathcal{A}(f_n) = \lim \mathcal{A}(g_n)$.

Somit ist $\mathcal{A}(f)$ wohldefiniert. Beweisen wir alle Eigenschaften 1-4.

1. $\mathcal{A}(f)$ ist selbstadjungiert als ein Grenzwert in $\mathcal{B}(H)$ der Folge von selbstadjungierten Operatoren (Aufgabe 75).

2. Für die Folge $\{f_n\}$ von Polynomen wie in (4.8) erhalten wir nach dem Korollar 4.2

$$\|\mathcal{A}(f)\| = \lim_{n\to\infty} \|\mathcal{A}(f_n)\| = \lim_{n\to\infty} \|f_n(A)\| = \lim_{n\to\infty} \sup_{\sigma(A)} |f_n| = \sup_{\sigma(A)} |f|.$$

3. Beweisen wir z.B. die Identität $\mathcal{A}(fg) = \mathcal{A}(f)\mathcal{A}(g)$ für alle $f, g \in C(\sigma(A))$. Seien $\{f_n\}$ und $\{g_n\}$ zwei Folgen von Polynomen mit

$$f_n \rightrightarrows f \quad \text{und } g_n \rightrightarrows g \quad \text{auf } C(\sigma(A)).$$

Daraus folgt, dass

$$f_n g_n \rightrightarrows fg \quad \text{auf} \quad C(\sigma(A)).$$

Da $f_n g_n$ ein Polynom ist, so erhalten wir

$$\mathcal{A}(fg) = \lim_{n\to\infty} \mathcal{A}(f_n g_n) = \lim_{n\to\infty} \mathcal{A}(f_n)\mathcal{A}(g_n) = \mathcal{A}(f)\mathcal{A}(g).$$

Die Identitäten $\mathcal{A}(f+g) = \mathcal{A}(f) + \mathcal{A}(g)$ und $\mathcal{A}(cf) = c\mathcal{A}(f)$ werden ähnlich bewiesen.

4. Da $f \equiv 1$ und $f = z$ die Polynome sind, so erhalten wir $\mathcal{A}(1) = 1(A) = I$ und $\mathcal{A}(z) = z(A) = A$. ■

Wie es schon gesagt wurde, wir benutzen weiter die Bezeichnung

$$f(A) \equiv \mathcal{A}(f).$$

01.07.22 Vorlesung 23

4.3 Der spektrale Abbildungssatz

Jetzt erweitern wir den spektralen Abbildungssatz 4.1 von Polynomen auf stetige Funktionen.

Satz 4.7 (Der spektrale Abbildungssatz für stetige Funktionen) *Sei A ein selbstadjungierter Operator in H. Dann gilt für jede Funktion $f \in C(\sigma(A))$*

$$\sigma(f(A)) = f(\sigma(A)).$$

Beweis. Beweisen wir für jedes $\lambda \in \mathbb{R}$ die Äquivalenz

$$\lambda \in f(\sigma(A)) \Leftrightarrow \lambda \in \sigma(f(A)).$$

Für ein $\lambda \notin f(\sigma(A))$ betrachten wir die Funktion

$$g(z) = \frac{1}{f(z) - \lambda} \in C(\sigma(A)).$$

Offensichtlich gilt

$$(f(z) - \lambda) g(z) = g(z)(f(z) - 1) = 1,$$

woraus folgt, dass

$$(f(A) - \lambda I) g(A) = g(A)(f(A) - \lambda I) = I.$$

Somit erhalten wir

$$(f(A) - \lambda I)^{-1} = g(A) \in \mathcal{B}(H)$$

und $\lambda \notin \sigma(f(A))$.

Sei $\lambda \in f(\sigma(A))$, so beweisen wir, dass $\lambda \in \sigma(f(A))$. Das Letzte ist äquivalent zu

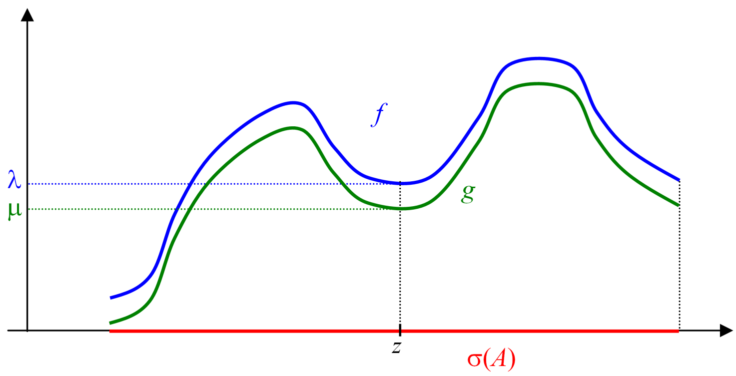
$$\inf_{\|x\|=1} \|f(A)x - \lambda x\| = 0 \tag{4.9}$$

(siehe (3.38) im Beweis von dem Satz 3.15). Da $\lambda \in f(\sigma(A))$, so existiert ein $z \in \sigma(A)$ mit $\lambda = f(z)$.

Fixieren wir ein $\varepsilon > 0$ und wählen ein Polynom g mit

$$\sup_{\sigma(A)} |f - g| < \varepsilon. \tag{4.10}$$

Bezeichnen wir $\mu = g(z)$. Es folgt aus (4.10) dass $|\lambda - \mu| < \varepsilon$.



Nach dem Satz 4.1 gilt

$$g(\sigma(A)) = \sigma(g(A)).$$

Da $\mu \in g(\sigma(A))$, so erhalten wir, dass auch $\mu \in \sigma(g(A))$ und somit

$$\inf_{\|x\|=1} \|g(A)x - \mu x\| = 0.$$

Insbesondere gibt es ein $x \in H$ mit $\|x\| = 1$ und

$$\|g(A)x - \mu x\| < \varepsilon.$$

Es folgt, dass

$$\begin{aligned}
\|f(A)x - \lambda x\| &\le \|f(A)x - g(A)x\| + \|g(A)x - \mu x\| + \|\mu x - \lambda x\| \\
&\le \|f(A) - g(A)\| + \varepsilon + |\mu - \lambda| \\
&\le \sup_{\sigma(A)} |f - g| + \varepsilon + |\mu - \lambda| \\
&< 3\varepsilon.
\end{aligned}$$

Da $\varepsilon > 0$ beliebig ist, so (4.9) folgt. ∎

4.4 Weitere Eigenschaften von Funktionalkalkül

Im nächsten Satz beweisen wir die weiteren Eigenschaften von Funktionalkalkül des Hauptsatzes 4.5.

Satz 4.8 *Das Funktionalkalkül vom selbstadjungierten Operator A erfüllt die folgenden Eigenschaften.*

5. (Erhaltung der Konvergenz) *Für jede Folge $\{f_n\}$ von Funktionen aus $C(\sigma(A))$ gilt die Implikation*

$$f_n \rightrightarrows f \text{ auf } \sigma(A) \ \Rightarrow \ f_n(A) \to f(A) \text{ in } \mathcal{B}(H).$$

6. (Erhaltung der Eigenvektoren) *Für $\lambda \in \mathbb{R}$, $x \in H$ and $f \in C(\sigma(A))$ gilt die Implikation*

$$Ax = \lambda x \Rightarrow f(A)x = f(\lambda)x.$$

Folglich gilt $f(\sigma_p(A)) \subset \sigma_p(f(A))$.

7. (Erhaltung der Komposition) *Für alle $f \in C(\sigma(A))$ und $g \in C(f(\sigma(A)))$ gilt*

$$(g \circ f)(A) = g(f(A)). \tag{4.11}$$

Bemerkung. Für beliebige Funktion $g \in C(f(\sigma(A)))$ ist die Komposition $g \circ f$ auf $\sigma(A)$ definiert, so dass der Operator $(g \circ f)(A)$ wohldefiniert ist. Die rechte Seite von (4.11) ist auch wohldefiniert da $\sigma(f(A)) = f(\sigma(A))$ und somit $g \in C(\sigma(f(A)))$.

Beweis. 5. Nach dem Satz 4.5 gilt

$$\|f_n(A) - f(A)\| = \|(f_n - f)(A)\| = \sup_{\sigma(A)} |f_n - f| \to 0.$$

6. Die Bedingung $Ax = \lambda x$ impliziert $A^n x = \lambda^n x$ für alle nichtnegative ganze Zahlen n, woraus folgt, dass für jedes Polynom f gilt

$$f(A)x = f(\lambda)x. \tag{4.12}$$

Für beliebige Funktion $f \in C(\sigma(A))$ wählen wir eine Folge $\{f_n\}$ von Polynomen mit

$$f_n \rightrightarrows f \quad \text{auf} \quad \sigma(A).$$

Nach dem obigen Argument gilt es

$$f_n(A)\,x = f_n(\lambda)\,x. \tag{4.13}$$

Nach 5 gilt $f_n(A) \to f(A)$ und somit $f_n(A)\,x \to f(A)\,x$. Da $\lambda \in \sigma(A)$ so gilt auch $f_n(\lambda) \to f(\lambda)$. Für $n \to \infty$ erhalten wir aus (4.13) die Identität (4.12).

Die Inklusion $f(\sigma_p(A)) \subset \sigma_p(f(A))$ wird wie folgt bewiesen: für jedes $\lambda \in \sigma_p(A)$ gibt es einen Eigenvektor x, und es folgt aus (4.12) dass $f(\lambda)$ ein Eigenwert von $f(A)$ ist, d.h. $f(\lambda) \in \sigma_p(f(A))$.

7. Sei zuerst g ein Polynom:

$$g(t) = c_0 + c_1 t + ... + c_n t^n.$$

Dann gilt dann

$$(g \circ f)(z) = c_0 + c_1 f(z) + ... + c_n f(z)^n$$

und somit

$$(g \circ f)(A) = c_0 + c_1 f(A) + ... + c_n f(A)^n = g(f(A)).$$

Für beliebige Funktion $g \in C(f(\sigma(A)))$ gibt es eine Folge $\{g_n\}$ von Polynomen mit

$$g_n \rightrightarrows g \text{ auf } f(\sigma(A)).$$

Nach dem obigen Argument gilt die Identität

$$(g_n \circ f)(A) = g_n(f(A)). \tag{4.14}$$

Für $n \to \infty$ erhalten wir folgendes. In der linken Seite haben wir

$$\sup_{\sigma(A)} |g_n \circ f - g \circ f| = \sup_{f(\sigma(A))} |g_n - g| \to 0$$

und somit nach 5

$$(g_n \circ f)(A) \to (g \circ f)(A). \tag{4.15}$$

In der rechten Seite von (4.14) haben wir nach dem Satz 4.7

$$\|g_n(f(A)) - g(f(A))\| = \sup_{\sigma(f(A))} |g_n - g| = \sup_{f(\sigma(A))} |g_n - g| \to 0,$$

und somit

$$g_n(f(A)) \to g(f(A)) \quad \text{für } n \to \infty. \tag{4.16}$$

Aus (4.14), (4.15) und (4.16) erhalten wir $(g \circ f)(A) = g(f(A))$, was zu beweisen war. ∎

Beispiel. Es ist bekannt, dass

$$\left(1 + \frac{z}{n}\right)^n \to \exp(z) \quad \text{für } n \to \infty$$

und die Konvergenz gleichmäßig auf jedem beschränkten Intervall ist. Da $\sigma(A)$ beschränkt ist, so erhalten wir

$$\left(I + \frac{1}{n}A\right)^n \to \exp(A).$$

Analog gilt die Identität

$$\exp(A) = \sum_{n=0}^{\infty} \frac{A^n}{n!}.$$

Beispiel. Betrachten wir eine Gleichung $\exp(X) = A$ wobei A ein gegebener selbstadjungierter Operator ist und X ein unbekannter Operator. Gilt $\sigma(A) > 0$, so ist die Funktion $f(z) = \ln z$ auf $\sigma(A)$ definiert, und wir erhalten eine Lösung $X = \ln A$, da nach dem Satz 4.8 gilt

$$\exp(\ln A) = (\exp \circ \ln)(A) = A.$$

Beispiel. Sei A ein kompakter selbstadjungierter Operator im Hilbertraum H, und sei $\{\lambda_n\}$ die Folge von allen verschiedenen Eigenwerten von A (siehe den Satz 3.15). Setzen wir $K_n = \ker(A - \lambda_n I)$ und beweisen, dass für jede Funktion $f \in C(\sigma(A))$ und alle $x \in H$ gilt

$$f(A)x = \sum_n f(\lambda_n) P_{K_n} x. \tag{4.17}$$

Nach dem Satz 3.15 haben wir für jedes $x \in H$

$$x = \sum_n x_n$$

wobei $x_n = P_{K_n} x \in K_n$. Da $Ax_n = \lambda_n x_n$, so erhalten wir nach dem Teil 6 des Satzes 4.8, dass

$$f(A)x_n = f(\lambda_n) x_n,$$

woraus folgt

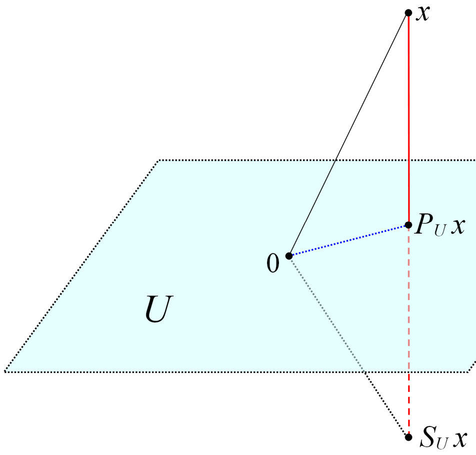
$$f(A)x = f(A) \sum_n x_n = \sum_n f(A)x_n = \sum_n f(\lambda_n) x_n,$$

was zu beweisen war.

Beispiel. (Aufgabe 88) Setzen wir $|A| = f(A)$ wobei $f(z) = |z|$. Sei U ein abgeschlossener Unterraum von H. Dann gilt für den Projektor P_U

$$|P_U| = P_U.$$

Sei S_U die *Spiegelung* an U, d.h. $S_U = 2P_U - I$.



Dann gilt

$$|S_U| = I.$$

Chapter 5

Dualräume

5.1 Definition von Dualraum

Sei V ein normierter Vektorraum über $\mathbb{K}$ (mit $\mathbb{K} = \mathbb{C}$ oder $\mathbb{R}$). Wir betrachten den Raum $\mathcal{B}(V, \mathbb{K})$ von allen beschränkten linearen Abbildungen $f : V \to \mathbb{K}$, d.h. den Raum von allen beschränkten *Funktionalen* auf V. Nach dem Satz 3.1 ist die Beschränktheit von f äquivalent zur Stetigkeit.

Erinnern wir uns, dass $\mathcal{B}(V, \mathbb{K})$ ein normierter Vektorraum ist und zwar mit der Operatornorm

$$\|f\| = \sup_{x \in V \setminus \{0\}} \frac{|f(x)|}{\|x\|} = \sup_{x \in V,\ \|x\|=1} |f(x)| \tag{5.1}$$

Da $\mathbb{K}$ vollständig ist, so nach dem Satz 3.2 ist $\mathcal{B}(V, \mathbb{K})$ ein Banachraum.

Definition. Der Banachraum $\mathcal{B}(V, \mathbb{K})$ heißt der *Dualraum* von V und wird mit V^* bezeichnet.

Hier besprechen wir die Aufgabe, den Dualraum eines gegebenen normierten Vektorraums V zu bestimmen.

Definition. Zwei normierte Vektorräume U_1 und U_2 heißen *linear isometrisch* (=*isomorph*), wenn es eine lineare Bijektion $\varphi : U_1 \to U_2$ gibt, die die Norm bewahrt, d.h. $\|\varphi(x)\| = \|x\|$ für alle $x \in U_1$. Man schreibt in diesem Fall $U_1 \cong U_2$. Die Abbildung φ heißt eine *lineare Isometrie* (auch Isomorphismus).

Zum Beispiel, jeder n-dimensionale Skalarproduktraum ist linear isometrisch zu $\mathbb{R}^n$. Nach dem Korollar 2.13, jeder ∞-dimensionale separable Hilbertraum ist linear isometrisch zu l^2.

In vielen Fällen kann man den Dualraum V^* bis auf lineare Isometrie bestimmen.

Beispiel. Sei $V = H$ ein Hilbertraum über $\mathbb{R}$. Jedes $a \in H$ bestimmt ein beschränktes (=stetiges) lineares Funktional $f_a \in H^*$ mit

$$f_a(x) = (x, a). \tag{5.2}$$

Somit erhalten wir die Abbildung

$$\begin{aligned} \varphi &: H \to H^* \\ \varphi(a) &= f_a. \end{aligned}$$

Nach dem Rieszschen Darstellungssatz (Satz 2.7) ist die Abbildung φ surjektiv, d.h. jedes $f \in H^*$ stimmt mit einem f_a überein. Zeigen wir, dass die Abbildung φ isometrisch ist, d.h. $\|f_a\| = \|a\|$ (daraus auch folgt, dass φ injektiv ist). In der Tat folgt es aus (5.2) und (5.1) dass

$$\|f_a\| = \sup_{x\in H\setminus\{0\}} \frac{|(x,a)|}{\|x\|} \le \sup_{x\in H\setminus\{0\}} \frac{\|x\|\,\|a\|}{\|x\|} = \|a\|$$

und

$$\|f_a\| \ge \frac{|(a,a)|}{\|a\|} = \frac{\|a\|^2}{\|a\|} = \|a\|\,,$$

woraus $\|f_a\| = \|a\|$ folgt. Somit ist φ eine lineare Isometrie, und es gilt

$$H^* \cong H. \tag{5.3}$$

Man sagt, dass H *selbstdual* ist.

5.2 Dualraum von l^p

Satz 5.1 *Seien $p, q \in (1, \infty)$ die konjugierten Hölder-Exponenten (d.h. $\frac{1}{p} + \frac{1}{q} = 1$). Dann gilt*

$$(l^p)^* \cong l^q. \tag{5.4}$$

Zum Beispiel, für $p = q = 2$ erhalten wir $(l^2)^* \cong l^2$ was auch aus (5.3) folgt, da l^2 ein Hilbertraum ist.

Bemerkung. Es gilt auch $L^p(X, \mu)^* \cong L^q(X, \mu)$, aber dies werden wir nicht beweisen.

06.07.22 Vorlesung 24

Beweis. Für jedes $a = \{a_k\}_{k=1}^\infty \in l^q$ definieren wir ein lineares Funktional f_a auf l^p wie folgt:

$$f_a(x) = \sum_{k=1}^\infty a_k x_k, \quad x \in l^p. \tag{5.5}$$

Nach der Hölder-Ungleichung des Lemmas 1.3 haben wir für alle $n \in \mathbb{N}$

$$\sum_{k=1}^n |a_k x_k| \le \left(\sum_{k=1}^n |a_k|^q\right)^{1/q} \left(\sum_{k=1}^n |x_k|^p\right)^{1/p} \le \|a\|_q \|x\|_p\,, \tag{5.6}$$

woraus folgt, dass die Reihe $\sum_{k=1}^\infty a_k x_k$ absolut konvergiert und es gilt

$$|f_a(x)| \le \|a\|_q \|x\|_p\,.$$

Es folgt dass

$$\|f_a\| \le \|a\|_q\,, \tag{5.7}$$

so dass f_a ein beschränktes lineares Funktional auf l^p ist, d.h. $f_a \in (l^p)^*$. Wir erhalten eine lineare Abbildung

$$\begin{aligned} \varphi : l^q &\to (l^p)^* \\ \varphi(a) &= f_a. \end{aligned} \tag{5.8}$$

Beweisen wir, dass diese Abbildung surjektiv ist, d.h. jedes $f \in (l^p)^*$ mit einem f_a übereinstimmt. Bezeichnen wir

$$e_k = \{0, , ..., 0, 1, 0, 0, ...\}$$

wobei 1 auf der Stelle k steht und alle andere Komponenten gleich 0 sind. Offensichtlich liegt e_k in l^p. Es gilt für jedes $x \in l^p$

$$x = \sum_{k=1}^{\infty} x_k e_k$$

wobei die Reihe in l^p konvergiert, da

$$\begin{aligned} \left\| x - \sum_{k=1}^{n} x_k e_k \right\|_p &= \|(0, ...0, x_{n+1}, x_{n+2}, ...)\|_p \\ &= \left(\sum_{k=n+1}^{\infty} |x_k|^p \right)^{1/p} \to 0 \end{aligned}$$

für $n \to \infty$. Da f stetig ist, so folgt es

$$f(x) = \sum_{k=1}^{\infty} f(x_k e_k) = \sum_{k=1}^{\infty} f(e_k) x_k = \sum_{k=1}^{\infty} a_k x_k$$

mit

$$a_k = f(e_k).$$

Es bleibt zu beweisen, dass die Folge $a := \{a_k\}$ in l^q liegt. Fixieren wir ein $n \in$N und betrachten wir das folgende Element $x \in l^p$:

$$\begin{aligned} &x_k = |a_k|^{q-1} \operatorname{sgn} a_k \ \text{ für } k \le n, \\ &x_k = 0 \ \text{ für } k > n. \end{aligned}$$

Dann gilt

$$\|x\|_p^p = \sum_{k=1}^{n} |a_k|^{p(q-1)} = \sum_{k=1}^{n} |a_k|^q$$

und

$$f(x) = f\left(\sum_{k=1}^{n} x_k e_k \right) = \sum_{k=1}^{n} a_k x_k = \sum_{k=1}^{n} a_k \left(|a_k|^{q-1} \operatorname{sgn} a_k \right) = \sum_{k=1}^{n} |a_k|^q,$$

so dass

$$\|x\|_p^p = f(x) = \sum_{k=1}^{n} |a_k|^q.$$

Es folgt, dass

$$f(x) \le \|f\| \, \|x\|_p = \|f\| \, f(x)^{1/p}$$

und somit

$$f(x)^{1/q} = f(x)^{1-1/p} \le \|f\|$$

und

$$\left(\sum_{k=1}^{n} |a_k|^q\right)^{1/q} \leq \|f\| .$$

Für $n \to \infty$ erhalten wir

$$\|a\|_q \leq \|f\| \tag{5.9}$$

und somit $a \in l^q$. Folglich gilt $f = f_a$. Der Vergleich von (5.7) und (5.9) ergibt

$$\|f_a\| = \|a\|_q .$$

Somit ist die Abbildung φ eine lineare Isometrie, woraus $(l^p)^* \cong l^q$ folgt. ■

Bemerkung. Analog kann man beweisen dass

$$\left(l^1\right)^* \cong l^\infty$$

was zu $p = 1$ und $q = \infty$ entspricht. Der obige Beweis muss wie folgt modifiziert werden. In diesem Fall $x \in l^1$, $a \in l^\infty$ und anstatt der Hölder-Ungleichung (5.6) verwendet man die offensichtliche Ungleichung

$$\sum_{k=1}^{\infty} |a_k x_k| \leq \sup_{k\in\mathbb{N}} |a_k| \sum_{k=1}^{\infty} |x_k| = \|a\|_\infty \|x\|_1 ,$$

woraus folgt

$$\|f_a\| \leq \|a\|_\infty ,$$

was analog zu (5.7) ist. Im zweiten Teil von dem Beweis (die Surjektivität der Abbildung φ_a aus (5.8)) muss man zeigen, dass die Folge $a_k = f(e_k)$ in l^∞ liegt, und das ist offensichtlich da

$$|f(e_k)| \leq \|f\| \, \|e_k\|_1 = \|f\| \quad \text{für alle } k \in \mathbb{N},$$

woraus folgt

$$\|a\|_\infty \leq \|f\| ,$$

was analog zu (5.9) ist.

Allerdings gilt $(l^\infty)^* \cong l^1$ nicht! Für jedes $a \in l^1$ kann man ein $f_a \in (l^\infty)^*$ wie zuvor mit (5.5) definieren, aber die Abbildung φ_a aus (5.8 nicht mehr surjektiv ist. Um die neuen Elemente von $(l^\infty)^*$ zu verstehen, müssen wir zuerst die Theorie weiter entwickeln.

5.3 Satz von Hahn-Banach

Um die weiteren Eigenschaften von Dualräumen untersuchen zu können, brauchen wir den folgenden Satz.

Hauptsatz 5.2 *Sei X ein normierter Vektorraum über $\mathbb{R}$ und sei X_0 ein Unterraum von X. Sei f_0 ein beschränktes lineares Funktional auf X_0. Dann lässt f_0 sich auf X mit Erhaltung der Norm fortsetzen, d.h. es gibt ein beschränktes lineares Funktional f auf X mit*

1. $f = f_0$ *auf* X_0
2. $\|f\| = \|f_0\|$.

Beweis. Im Fall wenn X_0 dicht in X liegt, die Fortsetzung f existiert und ist eindeutig bestimmt nach Aufgabe 61.

In der allgemeinen Situation muss X_0 nicht in X dicht liegen, so gilt die Eindeutigkeit nicht mehr, und der Beweis von Existenz wird komplizierter.

Im ersten Schritt zeigen wir die folgende Aussage: jedes beschränktes lineares Funktional g auf einem Unterraum $Y \subsetneqq X$ lässt such auf einen größeren Unterraum Z mit Erhaltung der Norm fortsetzen. Dafür wählen wir ein $z \in X \setminus Y$ und setzen

$$Z = \operatorname{span}\{Y, z\} = \{y + \alpha z : y \in Y, \alpha \in \mathbb{R}\}. \tag{5.10}$$

Ohne Beschränkung der Allgemeinheit nehmen wir an, dass

$$\|g\| = 1,$$

und beweisen, dass es ein Funktional h auf Z gibt mit

1. $h = g$ auf Y,
2. $\|h\| = 1$,

d.h. h ist eine Fortsetzung von g auf Z mit Erhaltung der Norm.

Um $\|h\| = 1$ zu beweisen, reicht es zu zeigen, dass

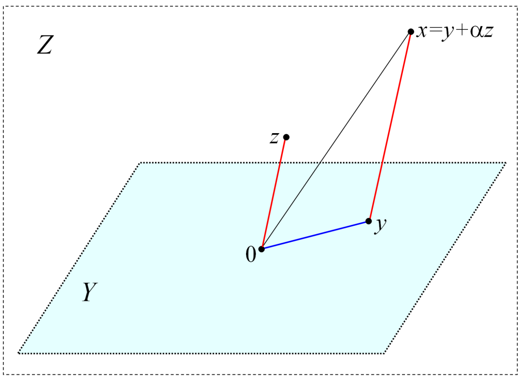
$$h(x) \leq \|x\| \quad \forall x \in Z. \tag{5.11}$$

In der Tat impliziert (5.11), dass auch $|h(x)| \leq \|x\|$ und somit $\|h\| \leq 1$, woraus folgt $\|h\| = 1$, da $\|h\| \geq \|g\| = 1$.

Nach (5.10) gilt für jedes $x \in Z$ die eindeutige Darstellung in der Form

$$x = y + \alpha z,$$

für ein $y \in Y$ und ein $\alpha \in \mathbb{R}$.



Für jede lineare Fortsetzung h von g auf Z muss die folgende Identität gelten:

$$h(x) = h(y) + ah(z) = g(y) + \alpha h(z) = g(y) + \alpha c,$$

wobei $c = h(z)$. Somit wird h eindeutig von dem Wert c bestimmt. Die Zahl $c \in \mathbb{R}$ muss so gewählt werden, um (5.11) zu sichern, was äquivalent zur folgenden Bedingung ist:

$$g(y) + \alpha c \leq \|y + \alpha z\| \qquad \forall y \in Y \quad \forall \alpha \in \mathbb{R}. \tag{5.12}$$

Für $\alpha = 0$ gilt diese Ungleichung trivialerweise da

$$g(y) \leq \|g\| \|y\| = \|y\|.$$

Für $\alpha > 0$ dividieren wir (5.12) durch α und schreiben (5.12) um wie folgt:

$$c \leq \left\| \frac{1}{\alpha} y + z \right\| - g\left(\frac{1}{\alpha} y\right).$$

Setzen wir $y_1 = y/\alpha$ und schreiben diese Ungleichung um wie folgt:

$$c \leq \|y_1 + z\| - g(y_1) \text{ für alle } y_1 \in Y. \tag{5.13}$$

Für $\alpha < 0$ dividieren wir wieder (5.12) durch α und erhalten

$$c \geq \frac{1}{\alpha} \|y + \alpha z\| - \frac{1}{\alpha} g(y) = -\left\| \frac{1}{\alpha} y + z \right\| - g\left(\frac{1}{\alpha} y\right).$$

Die Substitution $y_2 = y/\alpha$ ergibt die äquivalente Bedingung

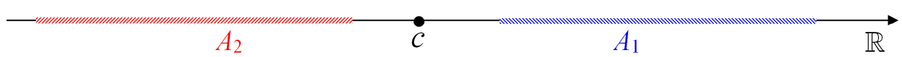
$$c \geq -\|y_2 + z\| - g(y_2) \text{ für alle } y_2 \in Y. \tag{5.14}$$

Somit muss die Zahl c die folgenden zwei Teilmengen von $\mathbb{R}$ trennen:

$$A_1 = \{\|y_1 + z\| - g(y_1) : y_1 \in Y\}$$

und

$$A_2 = \{-\|y_2 + z\| - g(y_2) : y_2 \in Y\}.$$



Nach dem Vollständigkeitsaxiom von $\mathbb{R}$ gibt es ein solches $c \in \mathbb{R}$ genau dann wenn

$$a_1 \geq a_2 \text{ für alle } a_1 \in A_1 \text{ und } a_2 \in A_2,$$

d.h. wenn für alle $y_1, y_2 \in Y$ gilt

$$\|y_1 + z\| - g(y_1) \geq -\|y_2 + z\| - g(y_2).$$

Diese Bedingung ist äquivalent zu

$$g(y_1 - y_2) \leq \|y_1 + z\| + \|y_2 + z\|,$$

und diese Ungleichung wirklich gilt, da

$$g(y_1 - y_2) \leq \|y_1 - y_2\| = \|(y_1 + z) - (y_2 + z)\| \leq \|y_1 + z\| + \|y_2 + z\|.$$

Somit ist die Existenz von h bewiesen.

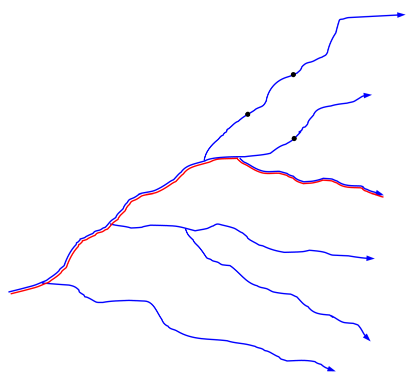
Im zweiten Schritt betrachten wir alle Paare (Y, g) von Unterräumen $Y \subset X$ und beschränkten Funktionalen g auf Y, und bezeichnen mit M die Menge von allen solchen Paaren (Y, g). Auf M definieren wir eine partielle Ordnung $\preceq$ wir folgt: für zwei Elemente (Y, g) und (Z, h) von M schreiben wir

$$(Y, g) \preceq (Z, h)$$

wenn $Y \subset Z$ und h eine Fortsetzung von g mit Erhaltung der Norm ist. Offensichtlich ist $\preceq$ eine partielle Ordnung auf M, d.h. für alle $a, b, c \in M$ gelten

1. $a \preceq a$,
2. $a \preceq b \preceq c \Rightarrow a \preceq c$,
3. $a \preceq b \preceq a \Rightarrow a = b$.

Eine Teilmenge $K \subset M$ heißt eine *Kette* wenn K total geordnet ist, d.h. für alle $a, b \in K$ gilt $a \preceq b$ oder $b \preceq a$. Die Kette K heißt *maximal*, wenn sie keine echte Teilmenge einer anderen Kette ist.



Ein Beispiel von einer partiell geordneten Menge (blau). Eine maximale Kette ist mit rot gezeigt.

Der *Maximalkettensatz* von Hausdorff besagt folgendes: jedes Element $a \in M$ einer partiell geordneten Menge M liegt in einer maximalen Kette K. Im Fall wenn alle Ketten endlich sind und die Längen von allen Ketten beschränkt sind, ist diese Aussage offensichtlich, da es für jedes $a \in M$ immer eine Kette $K \ni a$ mit maximaler Länge gibt, die somit maximal ist. Aber für beliebige partiell geordnete Mengen ist die Existenz der maximalen Kette nicht offensichtlich und ist äquivalent zum Auswahlaxiom der Mengenlehre (und auch zur transfiniten Induktion). Man kann den Maximalkettensatz als ein Axiom der Mengenlehre anstatt des Auswahlaxioms annehmen, was wir jetzt tun.

Nach dem Maximalkettensatz liegt (X_0, f_0) in einer maximalen Kette

$$K = \{(Y_i, g_i)\}_{i \in I}$$

wobei I eine Indexmenge ist. Diese Kette ist total geordnet, d.h. für alle $i, j \in I$ gilt $Y_i \subset Y_j$ oder $Y_j \subset Y_i$, und falls $Y_i \subset Y_j$ dann g_j eine Fortsetzung von g_i mit Erhaltung der Norm ist. Es folgt, dass die Menge

$$U = \bigcup_i Y_i$$

ein Unterraum von X, und dass das folgende Funktional f auf U

$$f(x) = g_i(x) \quad \text{für } x \in Y_i$$

wohldefiniert ist. Dann ist (U, f) ein Element von M und es gilt $(Y_i, g_i) \preceq (U, f)$ für alle $i \in I$. Da $K \cup \{(U, f)\}$ eine Kette ist und K maximal ist, so beschließen wir, dass $(U, f) \in K$. Insbesondere ist (U, f) ein maximales Element von K.

Zeigen wir, dass $U = X$. Ist U eine echte Teilmenge von X, so kann man nach dem ersten Schritt das Funktional f weiter fortsetzen, was wegen der Maximalität von (U, f) nicht möglich ist. Somit gilt $U = X$. Dann ist f eine Fortsetzung von f_0 auf X mit Erhaltung der Norm. ■

08.07.22 Vorlesung 25

5.4 Bidualraum

Sei X ein normierter Raum über $\mathbb{R}$. Betrachten wir den *Bidualraum*

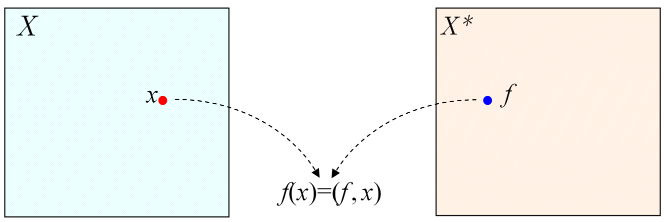
$$X^{**} := (X^*)^*.$$

Wir zeigen hier, dass der Raum X sich immer betrachten lässt als Unterraum von X^{**}.

Für jedes $x \in X$ und $f \in X^*$ definieren wir die *Paarung*

$$(f, x) := f(x)$$

und betrachten die Paarung als eine reellwertige Funktion auf $X \times X^*$.



Für jedes fixiertes $x \in X$ bestimmt die Paarung ein Funktional auf X^* wie folgt:

$$\Phi_x : X^* \to \mathbb{R}$$
$$f \mapsto (f, x)$$

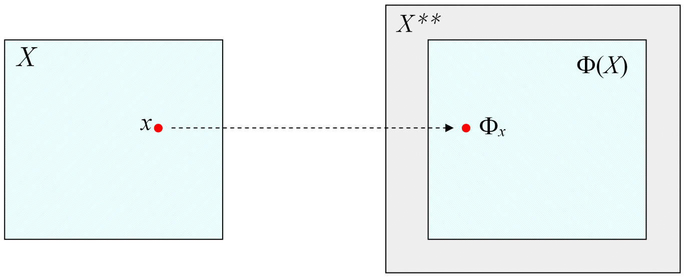
d.h.

$$\Phi_x(f) = f(x) \quad \text{für alle } f \in X^*.$$

Satz 5.3 *Das Funktional Φ_x auf X^* ist linear, beschränkt und somit gehört zu X^{**}. Die Abbildung*

$$\begin{aligned}\Phi : X &\to X^{**} \\ x &\mapsto \Phi_x\end{aligned}$$

ist eine lineare Isometrie zwischen X und dem Unterraum $\Phi(X) \subset X^{**}$.



Beweis. Das Funktional Φ_x ist offensichtlich linear:

$$\Phi_x(f+g) = (f+g)(x) = f(x) + g(x) = \Phi_x(f) + \Phi_x(g)$$

und analog $\Phi_x(cf) = c\Phi_x(f)$ für $c \in \mathbb{R}$. Das Funktional Φ_x ist beschränkt, da

$$|\Phi_x(f)| = |f(x)| \leq \|f\| \|x\|,$$

woraus folgt, dass

$$\begin{aligned}\|\Phi_x\|_{X^{**}} &= \sup_{f \in X^* \setminus \{0\}} \frac{\Phi_x(f)}{\|f\|} = \sup_{f \in X^* \setminus \{0\}} \frac{f(x)}{\|f\|} \\ &\leq \sup_{f \in X^* \setminus \{0\}} \frac{\|f\| \|x\|}{\|f\|} = \|x\| < \infty. \end{aligned} \tag{5.15}$$

Somit liegt Φ_x in X^{**} und Φ ist wirklich eine Abbildung von X nach X^{**}.

Offensichtlich ist die Abbildung $\Phi : X \to X^{**}$ linear: für alle $x, y \in X$

$$\Phi_{x+y}(f) = f(x+y) = f(x) + f(y) = \Phi_x(f) + \Phi_y(f)$$

so dass $\Phi_{x+y} = \Phi_x + \Phi_y$ und analog $\Phi_{cx} = c\Phi_x$ für $c \in \mathbb{R}$. Beweisen wir jetzt, dass Φ normerhaltend ist, d.h. für alle $x \in X$

$$\|\Phi_x\|_{X^{**}} = \|x\|_X,$$

was auch implizieren wird, dass Φ_x injektiv ist. Nach (5.15) gilt

$$\|\Phi_x\|_{X^{**}} \leq \|x\|_X.$$

Um die umgekehrte Ungleichung

$$\|\Phi_x\|_{X^{**}} \geq \|x\|_X$$

zu erhalten, reicht es die folgende Aussage zu beweisen:

$$\forall x \in X \quad \exists f \in X^* \text{ mit } \|f\| = 1 \text{ und } f(x) = \|x\|, \tag{5.16}$$

da (5.16) implizieren wird, dass

$$\|\Phi_x\|_{X^{**}} \geq \frac{\Phi_x(f)}{\|f\|} = \frac{f(x)}{\|f\|} = \|x\|.$$

Fixieren wir ein $x \in X$ und betrachten einen eindimensionalen Unterraum X_0 von X wie folgt:

$$X_0 = \operatorname{span}\{x\} = \{\alpha x : \alpha \in \mathbb{R}\}.$$

Definieren wir f auf X_0 mit

$$f(\alpha x) = \alpha \|x\| \quad \text{für alle } \alpha \in \mathbb{R}.$$

Dann ist f ein lineares Funktional auf X_0 und es gilt

$$\|f\| = \sup_{\alpha \in \mathbb{R} \setminus \{0\}} \frac{|f(\alpha x)|}{\|\alpha x\|} = \sup_{\alpha \in \mathbb{R} \setminus \{0\}} \frac{|\alpha \|x\||}{\|\alpha x\|} = 1 \text{ und } f(x) = \|x\|,$$

Nach dem Satz von Hahn-Banach gibt es eine Fortsetzung von f auf X mit der Norm 1, woraus (5.16) folgt.

Somit ist $\Phi : X \to X^{**}$ eine injektive lineare Abbildung mit Erhaltung der Norm. Es folgt, dass Φ eine lineare Isometrie zwischen X und $\Phi(X)$ ist. ■

Wir identifizieren jedes $x \in X$ mit Φ_x und somit X mit $\Phi(X)$, und betrachten X immer als Unterraum von X^{**}, d.h.

$$X \subset X^{**}.$$

Definition. Ein Banachraum X heißt *reflexiv* wenn

$$X = X^{**}.$$

In anderen Wörtern, X ist reflexiv wenn die Abbildung $\Phi : X \to X^{**}$ surjektiv ist, d.h. $\Phi(X) = X^{**}$.

Beispiel. Beweisen wir, dass der Raum $X = l^p$ für jedes $p \in (1, \infty)$ reflexiv ist. Nach dem Satz 5.1 gilt

$$(l^p)^* \cong l^q \tag{5.17}$$

wobei $q = \frac{p}{p-1}$, und jedes $f \in (l^p)^*$ hat der Form

$$f(x) = \sum_{k=1}^{\infty} f_k x_k \quad \text{für alle } x = \{x_k\} \in l^p,$$

wobei $\{f_k\} \in l^q$. Es gilt auch

$$\|f\|_{(l^p)^*} = \|\{f_k\}\|_{l^q}.$$

Wir identifizieren jedes $f \in (l^p)^*$ mit der Folge $\{f_k\} \in l^q$ und schreiben

$$(l^p)^* = l^q.$$

Nach dem gleichen Argument gilt $(l^q)^* = l^p$, woraus folgt $(l^p)^{**} = l^p$.

5.5 Banachlimes

Der Isomorphismus (5.17) gilt auch für $p = 1$, so dass

$$\left(l^1\right)^* = l^\infty,$$

woraus folgt

$$\left(l^1\right)^{**} = \left(l^\infty\right)^*.$$

Hier besprechen wir kurz $(l^\infty)^*$. Nach dem Satz 5.3 lässt sich l^1 mit einem Unterraum von $(l^\infty)^*$ identifizieren:

$$l^1 \subset \left(l^\infty\right)^*.$$

In diesem Fall funktioniert diese Identifizierung wie folgt. Jedes $a \in l^1$ bestimmt ein Funktional Φ_a auf l^∞ mit

$$\Phi_a(x) = \sum_{k=1}^{\infty} a_k x_k, \quad \text{für alle } x = \{x_k\} \in l^\infty,$$

und es gilt $\|\Phi_a\|_{(l^\infty)^*} = \|a\|_{l^1}$. Wir identifizieren jedes $a \in l^1$ mit $\Phi_a \in (l^\infty)^*$ und somit l^1 mit einem Unterraum von $(l^\infty)^*$.

Satz 5.4 *Der Raum l^1 ist eine echte Teilmenge von $(l^\infty)^*$, d.h. $(l^\infty)^* \setminus l^1 \neq \emptyset$. Folglich ist l^1 eine echte Teilmenge von $(l^1)^{**}$, und l^1 ist nicht reflexiv.*

Beweis. Wir finden ein Funktional $f \in (l^\infty)^*$ das mit keinem Φ_a übereinstimmt. Dafür betrachten wir den folgenden Unterraum $U \subset l^\infty$:

$$U = \left\{\{x_k\} \in l^\infty : \lim_{k\to\infty} x_k \text{ existiert}\right\} = \{\{x_k\} : \ \{x_k\} \text{ ist konvergent}\}.$$

Für jedes $x = \{x_k\} \in U$ definieren wir das Funktional

$$f(x) = \lim_{k\to\infty} x_k.$$

Offensichtlich ist f linear und es gilt

$$|f(x)| \leq \sup |x_k| = \|x\|_\infty,$$

so dass f beschränkt ist und zwar mit der Norm $\|f\| \leq 1$. Tatsächlich gilt $\|f\| = 1$ da für die konstante Folge $x = \{1, 1, ...\}$ gilt $f(x) = 1 = \|x\|_\infty$.

Nach dem Satz von Hahn-Banach (Hauptsatz 5.2) lässt f sich auf den ganzen Raum l^∞ mit der Norm 1 fortsetzen. Diese Fortsetzung ist nicht eindeutig, aber jede solche Fortsetzung heißt *Banachlimes* und wird mit $\operatorname{LIM} x_k$ bezeichnet. Somit ist $\operatorname{LIM} x_k$ für alle beschränkte Folgen $\{x_k\}$ definiert und bestimmt ein Element von $(l^\infty)^*$.

Zeigen wir, dass der Banachlimes mit keinem Φ_a übereinstimmt. Gilt für ein $a \in l^1$ und für alle $x \in l^\infty$ die Identität

$$\operatorname{LIM} x_k = \Phi_a(x) = \sum_{k=1}^{\infty} a_k x_k,$$

so erhalten wir, dass für jede konvergente Folge $\{x_k\}$

$$\lim x_k = \sum_{k=1}^{\infty} a_k x_k$$

was nicht möglich ist, da $\lim x_k$ von den einzelnen Komponenten x_k unabhängig ist. ■

5.6 Schwache Topologie

Sei X ein normierter Vektorraum über $\mathbb{R}$. Der Dualraum X^* ist immer ein Banachraum. Insbesondere gibt es Begriff von *Norm-Konvergenz* in X^* bezüglich der Norm in X^*: ist $\{f_n\}$ eine Folge von Elementen von X^*, so gilt $f_n \to f$ wenn $\|f_n - f\| \to 0$ für $n \to \infty$. Die Norm-Konvergenz in X^* heißt auch die *starke* Konvergenz.

Definition. Eine Folge $\{f_n\}$ von Elementen von X^* konvergiert *schwach* (punktweise) gegen $f \in X^*$ wenn

$$f_n(x) \to f(x) \text{ für alle } x \in X.$$

Man schreibt in diesem Fall

$$f_n \rightharpoonup f \quad \text{oder} \quad f_n \stackrel{w}{\to} f \quad \text{oder} \quad f = w\text{-}\lim f_n,$$

wobei "w" für "weak" steht.

Die Norm-Konvergenz impliziert offensichtlich die schwache Konvergenz, aber die Umkehrung gilt im allgemeinen Fall nicht (im Fall $\dim X < \infty$ stimmen die schwache Konvergenz und die Norm-Konvergenz überein).

Man kann zeigen, dass die schwache Konvergenz eine Konvergenz in einer Topologie in X^* ist, die man die *schwache Topologie* nennt (im Gegenteil zur Norm-Topologie). Wir brauchen die schwache Topologie explizit nicht, und definieren alle dazugehörigen Begriffe (z.B. Kompaktheit) direkt, mit Hilfe von Folgen.

Die schwache Konvergenz (bzw Topologie) in X^* heißt auch schwache* Konvergenz (bzw Topologie), da dieser Begriff sich auf den Dualraum bezieht.

Beispiel. Sei $X = C[a,b]$. Wir bezeichnen die Funktionen aus X mit $x(t)$ und die Funktionale auf X mit $f(x)$. Für jedes $t \in [a,b]$ definieren wir die *Dirac-Funktion* $\delta_t \in C[a,b]^*$ wie folgt:

$$\delta_t(x) = x(t), \quad x \in C[a,b].$$

Offensichtlich ist δ_t ein lineares beschränktes Funktional auf $C[a,b]$ da

$$\|\delta_t\| = \sup_{x \in X \setminus \{0\}} \frac{|\delta_t(x)|}{\|x\|} = \sup_{x \in X \setminus \{0\}} \frac{|x(t)|}{\sup_{[a,b]} |x|} \leq 1.$$

Sei $\{t_n\}$ eine konvergente Folge in $[a,b]$ mit $t = \lim t_n$. Dann gilt für jedes $x \in C[a,b]$

$$\delta_{t_n}(x) = x(t_n) \to x(t) = \delta_t(x)$$

woraus folgt

$$\delta_{t_n} \rightharpoonup \delta_t$$

für $n \to \infty$. Zeigen wir, dass

$$\delta_{t_n} \not\to \delta_t,$$

vorausgesetzt $t_n \neq t$. Wir haben

$$\begin{aligned} \|\delta_{t_n} - \delta_t\| &= \sup_{x \in C[a,b], \|x\|=1} |\delta_{t_n}(x) - \delta_t(x)| \\ &= \sup_{x \in C[a,b], \|x\|=1} |x(t_n) - x(t)| = 2, \end{aligned}$$

da es immer eine Funktion $x \in C[a,b]$ gibt mit

$$\|x\| = 1, \quad x(t_n) = 1 \text{ und } x(t) = -1.$$

Wir beschließen, dass

$$\|\delta_{t_n} - \delta_t\| \not\to 0.$$

Wir besprechen unterhalb die Vollständigkeit und Kompaktheit bezüglich der schwachen Konvergenz.

Definition. Eine Folge $\{f_n\}$ von Elementen von X^* heißt *schwache Cauchy-Folge* wenn für jedes $x \in X$ die numerische Folge $\{f_n(x)\}$ eine Cauchy-Folge ist.

Definition. Der Raum X^* heißt *schwach vollständig*, wenn jede schwache Cauchy-Folge schwach konvergent ist.

Definition. Eine Menge $M \subset X^*$ heißt *schwach präkompakt* wenn jede Folge $\{f_n\}$ von Elementen von M eine schwach konvergente Teilfolge enthält.

Definition. Eine Menge $M \subset X^*$ heißt *schwach beschränkt*, wenn für jedes $x \in X$ die Menge $\{f(x)\}_{f \in M}$ beschränkt ist (als Teilmenge von $\mathbb{R}$).

Offensichtlich gelten die folgenden Implikationen:

$$\begin{array}{ccc} \textit{Norm-präkompakt} & \Rightarrow & \textit{Norm-beschränkt} \\ \Downarrow & & \Downarrow \\ \textit{schwach präkompakt} & \Rightarrow & \textit{schwach beschränkt} \end{array}$$

Wir werden beweisen, dass

$$\begin{array}{ccc} & & \textit{Norm-beschränkt} \\ & & \Updownarrow \\ \textit{schwach präkompakt} & \Leftrightarrow & \textit{schwach beschränkt} \end{array}$$

5.7 Prinzip der gleichmäßigen Beschränktheit

Hauptsatz 5.5 (Satz von Banach-Steinhaus, das Prinzip der gleichmäßigen Beschränktheit) *Sei X ein Banachraum. Dann jede schwach beschränkte Menge M von X^* ist Norm-beschränkt.*

Da jede beschränkte Menge auch schwach beschränkt ist, sind die Begriffe "beschränkt" und "schwach beschränkt" äquivalent.

Für den Beweis benutzen wir den folgenden Satz.

Satz 5.6 (Satz von Baire) *Sei (X,d) ein vollständiger metrischer Raum, und sei $\{S_n\}_{n=1}^{\infty}$ eine Folge von abgeschlossen Teilmengen von X mit*

$$X = \bigcup_{n=1}^{\infty} S_n.$$

Dann enthält mindestens eine Menge S_n eine offene Kugel.

Bemerkung. Die offene Kugel $B(z,r)$ mit Zentrum $z \in X$ and Radius r ist die Menge

$$B(z,r) = \{x \in X : d(x,z) < r\}.$$

Wir nehmen immer an, dass $r > 0$.

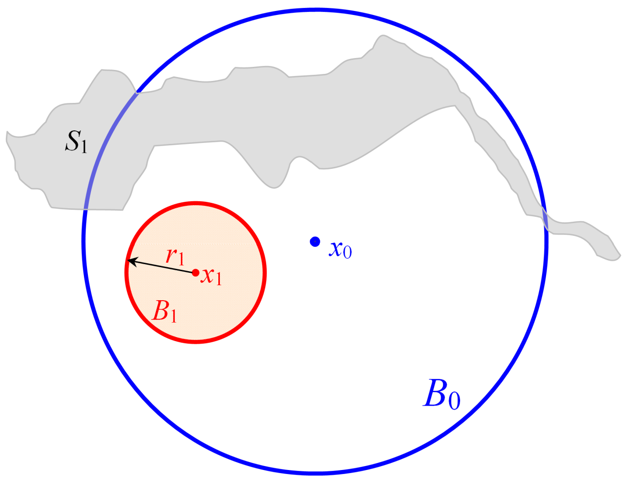
Bemerkung. Ohne Vollständigkeit gilt diese Aussage nicht. Z.B. $X = \mathbb{Q}$ ist abzählbar und somit ist X eine abzählbare Vereinigung von einzelnen Punkten, die kein offenes Intervall enthalten.

13.07.22 Vorlesung 26

Beweis. Nehmen wir das Gegenteil an, dass keine Menge S_n eine offene Kugel enthält. Wählen wir ein beliebiges $x_0 \in X$ und setzen $B_0 = B(x_0, 1)$. Die Menge $B_0 \setminus S_1$ ist nicht leer, so wählen wir einen Punkt

$$x_1 \in B_0 \setminus S_1.$$

Da $B_0 \setminus S_1$ eine offene Menge ist, so enthält diese Menge eine Kugel $B_1 = B(x_1, r_1)$ mit $r_1 > 0$.



Der Radius r_1 kann immer so klein gewählt werden, dass die folgenden Bedingungen erfüllt werden:

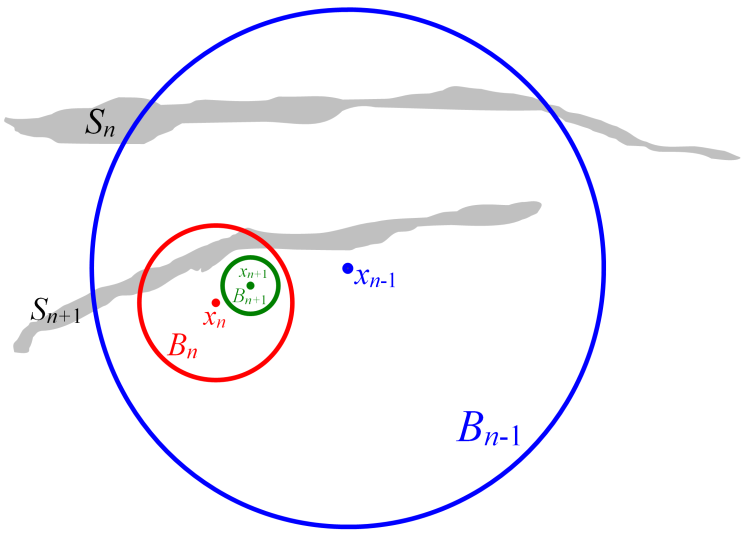
$$r_1 < \frac{1}{2} \quad \text{und} \quad \overline{B}_1 \subset B_0 \setminus S_1$$

Da die Menge $B_1 \setminus S_2$ nicht leer und offen ist, so existiert analog eine Kugel $B_2 = B(x_2, r_2)$ mit

$$r_2 < \frac{1}{2^2} \quad \text{und} \quad \overline{B}_2 \subset B_1 \setminus S_2.$$

Per Induktion erhalten wir für jedes $n \geq 1$ eine Kugel $B_n = B(x_n, r_n)$ mit

$$r_n < \frac{1}{2^n} \quad \text{und} \quad \overline{B}_n \subset B_{n-1} \setminus S_n. \tag{5.18}$$



Da $x_n \in B_n \subset B_{n-1}$, so erhalten wir

$$d\left(x_n, x_{n-1}\right) < r_{n-1} < \frac{1}{2^{n-1}},$$

woraus folgt, dass die Folge $\{x_n\}$ eine Cauchy-Folge ist. Somit existiert der Grenzwert

$$x = \lim_{n \to \infty} x_n.$$

Für alle $m > n$ gilt

$$x_m \in B_m \subset B_{m-1} \subset ... \subset B_n$$

d.h.

$$x_m \in B_n,$$

woraus folgt für $m \to \infty$, dass

$$x \in \overline{B}_n \text{ für alle } n.$$

Da nach (5.18) gilt $\overline{B}_n \subset B_{n-1} \setminus S_n$ so folgt es, dass

$$x \in B_{n-1} \setminus S_n \text{ für alle } n.$$

Wir erhalten, dass $x \notin S_n$ für alle n, was im Widerspruch zur Voraussetzung $X = \bigcup_{n=1}^{\infty} S_n$ steht. ■

Definition. Eine Teilmenge S von X heißt *nirgends dicht* wenn der Abschluss $\overline{S}$ keine Kugel enthält.

Seien S_n nirgends dichte Teilmengen von X für alle $n \in \mathbb{N}$. Anwendung von dem Satz 5.6 auf die Mengen $\overline{S}_n$ ergibt dass

$$\bigcup_{n=1}^{\infty} S_n \neq X,$$

d.h. das Komplement $K = X \setminus \bigcup_{n=1}^{\infty} S_n$ nicht leer ist. Darüber hinaus folgt es aus dem Beweis, dass *die Menge K dicht in X liegt.* In der Tat haben für im Beweis gezeigt, dass die Menge $K \cap B_0$ nicht leer ist wobei $B_0 = B(x_0, 1)$. Genauso beweist man dass $K \cap B_0 \neq \emptyset$ für *beliebige* Kugel B_0, was bedeutet, dass K dicht in X liegt.

Definition. Eine Teilmenge $M \subset X$ heißt *mager* wenn $M = \bigcup_{n=1}^{\infty} S_n$ für eine Folge $\{S_n\}$ von nirgends dichten Teilmengen S_n.

Der Satz von Baire lässt sich wie folgt umformulieren: wenn X vollständig ist und $M \subset X$ eine magere Teilmenge von X ist, so liegt das Komplement $X \setminus M$ dicht in X.

Beweis von Satz 5.5. Sei M eine schwach beschränkte Teilmenge von X^*. Das bedeutet, dass die Menge $\{f(x)\}_{f \in M} \subset \mathbb{R}$ beschränkt für jedes $x \in X$ ist, d.h.

$$\forall x \in X \quad \sup_{f \in M} |f(x)| < \infty. \tag{5.19}$$

Wir beweisen, dass die Menge M Norm-beschränkt ist, d.h.

$$\sup_{f \in M} \|f\| < \infty.$$

Dafür betrachten wir für jedes $n \in \mathbb{N}$ die Menge

$$S_n = \left\{ x \in X : \sup_{f \in M} |f(x)| \leq n \right\}.$$

Es folgt aus (5.19) dass

$$X = \bigcup_{n=1}^{\infty} S_n.$$

Da f stetig ist und somit die Menge

$$\{x \in X : |f(x)| \leq n\} = f^{-1}([-n, n])$$

abgeschlossen ist, so ist auch die Menge

$$\begin{aligned} S_n &= \{x \in X : |f(x)| \leq n \quad \forall f \in M\} \\ &= \bigcap_{f \in M} \{x \in X : |f(x)| \leq n\} \end{aligned}$$

abgeschlossen als der Durchschnitt von abgeschlossenen Mengen.

Nach dem Satz 5.6 erhalten wir, dass eine von S_n eine Kugel enthält, d.h.

$$S_n \supset B(z, r) \quad \text{für einige } n \in \mathbb{N}, \ z \in X \text{ und } r > 0.$$

Nach Definition von S_n erhalten wir, dass

$$\forall x \in B(z, r) \quad \sup_{f \in M} |f(x)| \leq n$$

und somit

$$\sup_{x \in B(z,r)} \sup_{f \in M} |f(x)| \leq n.$$

Für jedes $y \in B(0, 1)$ liegt der Punkt $x = z + ry$ in $B(z, r)$, da

$$\|x - z\| = \|ry\| = r\|y\| < r.$$

Dann gilt

$$f(x) = f(z) + rf(y)$$

und

$$f(y) = \frac{1}{r}f(x) - \frac{1}{r}f(z).$$

Daraus folgt, dass

$$\sup_{y\in B(0,1)} \sup_{f\in M} |f(y)| \leq \frac{1}{r} \sup_{x\in B(z,r)} \sup_{f\in M} |f(x)| + \frac{1}{r} \sup_{f\in M} |f(z)| \leq \frac{1}{r}n + \frac{1}{r}m,$$

wobei

$$m = \sup_{f\in M} |f(z)| < \infty.$$

Somit erhalten wir

$$\sup_{f\in M} \sup_{y\in B(0,1)} |f(y)| < \infty,$$

was äquivalent zu

$$\sup_{f\in M} \|f\| < \infty,$$

ist. Somit ist M Norm-beschränkt, was zu beweisen war. ■

5.8 Schwache Vollständigkeit des Dualraums

Satz 5.7 *Sei X ein Banachraum über $\mathbb{R}$. Dann ist der Dualraum X^* schwach vollständig.*

Beweis. Sei $\{f_n\}$ eine schwache Cauchy-Folge in X^*, d.h. für jedes $x \in X$ die Folge $\{f_n(x)\}$ eine Cauchy-Folge in $\mathbb{R}$ ist. Wir müssen beweisen, dass die Folge $\{f_n\}$ schwach konvergiert. Für jedes $x \in X$ ist die Folge $\{f_n(x)\}$ konvergent in $\mathbb{R}$. Setzen wir

$$f(x) := \lim_{n\to\infty} f_n(x). \tag{5.20}$$

Offensichtlich ist f ein lineares Funktional auf X. Beweisen wir, dass f stetig ist. Für jedes $x \in X$ ist die Folge $\{f_n(x)\}$ beschränkt, so dass die Folge $\{f_n\}$ schwach beschränkt ist. Nach dem Hauptsatz 5.5 ist die Folge $\{f_n\}$ auch Norm-beschränkt, d.h. existiert ein $C \in \mathbb{R}$ mit

$$\|f_n\| \leq C \text{ für alle } n.$$

Dann gilt für jedes $x \in X$

$$|f_n(x)| \leq C\|x\|$$

und somit nach (5.20)

$$|f(x)| = \lim_{n\to\infty} |f_n(x)| \leq C\|x\|,$$

woraus folgt

$$\|f\| \leq C.$$

Deshalb ist f beschränkt und somit stetig, insbesondere $f \in X^*$. Es folgt aus (5.20), dass $f_n \rightharpoonup f$ was zu beweisen war. ■

5.9 Schwache Präkompaktheit im Dualraum

Satz 5.8 (Satz von Banach-Alaoglu) *Sei X ein separabler Banachraum über $\mathbb{R}$. Dann ist jede beschränkte Teilmenge von X^* schwach präkompakt.*

Somit erhalten wir die folgenden Implikationen

$$\text{schwach präkompakt} \overset{\text{trivial}}{\Rightarrow} \text{schwach beschränkt} \overset{\text{S.5.5}}{\Rightarrow} \text{beschränkt} \overset{\text{S.5.8}}{\Rightarrow} \text{schwach präkompakt}$$

woraus folgt

$$\text{schwach präkompakt} \Leftrightarrow \text{schwach beschränkt} \Leftrightarrow \text{beschränkt}\ .$$

Beweis. Es reicht folgendes zu beweisen: jede beschränkte Folge $\{f_n\}$ von Elementen von X^* besitzt eine schwach konvergente Teilfolge (analog zum Satz von Bolzano-Weierstraß). Nach der Separabilität von X gibt es in X eine dicht liegende Folge $\{x_k\}_{k\in\mathbb{N}}$. Die Folge von Werten

$$f_1(x_1),\ f_2(x_1), ...,\ f_n(x_1), ...$$

ist beschränkt in $\mathbb{R}$ und somit hat eine konvergente Teilfolge:

$$f_{n_1}(x_1),\ f_{n_2}(x_1), ...,\ f_{n_k}(x_1), ...$$

Benennen wir diese Folge mit Hilfe von Hochstellung um wie folgt: $f_k^{(1)} := f_{n_k}$. Dann ist die Folge $\{f_n^{(1)}\}$ eine Teilfolge von $\{f_n\}$ und

$$f_1^{(1)}(x_1),\ f_2^{(1)}(x_1), ...,\ f_n^{(1)}(x_1), ... \text{ konvergiert.}$$

Analog ist die Folge

$$f_1^{(1)}(x_2),\ f_2^{(1)}(x_2), ...,\ f_n^{(1)}(x_2), ...$$

beschränkt und somit hat eine konvergente Teilfolge:

$$f_1^{(2)}(x_2),\ f_2^{(2)}(x_2), ...,\ f_n^{(2)}(x_2), ... \text{ konvergiert.}$$

Per Induktion nach k erhalten wir die Folge $\{f_n^{(k)}\}_{\ n\in\mathbb{N}}$ so dass für jedes $k\in\mathbb{N}$

1. $\{f_n^{(k+1)}\}_{n\in\mathbb{N}}$ ist eine Teilfolge von $\{f_n^{(k)}\}_{n\in\mathbb{N}}$;
2. die Folge $f_n^{(k)}(x_k)$ konvergiert für $n\to\infty$.

Somit erhalten wir eine Tabelle von Folgen

$$\begin{array}{l}
f_1^{(1)},\quad f_2^{(1)}, \dots\dots\dots\dots\dots\dots,\ f_n^{(1)}, ...\\
f_1^{(2)},\quad f_2^{(2)}, \dots\dots\dots\dots\dots\dots,\ f_n^{(2)}, ...\\
...\\
f_1^{(k)},\quad f_2^{(k)}, ...,\ f_k^{(k)}, \dots\dots\dots,\ f_n^{(k)}, ...\\
f_1^{(k+1)},\quad f_2^{(k+1)}, ...,\ f_{k+1}^{(k+1)}, ...,\ f_n^{(k+1)}, ...\\
...
\end{array}$$

mit der folgenden Eigenschaften:

1. die $(k+1)$-te Zeile ist eine Teilfolge der k-ten Zeile;
2. die k-te Zeile konvergiert für $n \to \infty$ an der Stelle x_k.

Betrachten wir die *Diagonalfolge*

$$f_1^{(1)},\ f_2^{(2)}, ..., f_n^{(n)}, ...,$$

die eine Teilfolge von $\{f_n\}$ ist. Diese Folge konvergiert all allen Stellen x_k, da $\{f_n^{(n)}\}_{n \geq k}$ eine Teilfolge von $\{f_n^{(k)}\}_{n \geq k}$ ist, und die letzte Folge an x_k konvergiert.

Setzen wir

$$g_n = f_n^{(n)},$$

so dass $\{g_n\}$ eine Teilfolge von $\{f_n\}$ ist und $\{g_n\}_{n \in \mathbb{N}}$ an allen Stellen x_k konvergiert. Beweisen wir, dass $\{g_n\}$ schwach konvergent ist. Nach dem Satz 5.7 reicht es zu beweisen, dass $\{g_n\}$ ein schwache Cauchy-Folge ist, d.h. für jedes $x \in X$

$$\lim_{n,m \to \infty} |g_n(x) - g_m(x)| = 0.$$

Sei auch C eine obere Schranke für die Folge $\{\|f_n\|\}$, die nach Voraussetzung existiert. Dann ist C auch eine obere Schranke für $\{\|g_n\|\}$. Fixieren wir ein $x \in X$ und $\varepsilon > 0$ und wählen x_k so dass $\|x - x_k\| < \varepsilon$. Dann erhalten wir für alle $n, m \in \mathbb{N}$

$$\begin{aligned}
|g_n(x) - g_m(x)| &\leq |g_n(x) - g_n(x_k)| + |g_n(x_k) - g_m(x_k)| + |g_m(x_k) - g_m(x)| \\
&\leq \|g_n\| \|x - x_k\| + |g_n(x_k) - g_m(x_k)| + \|g_m\| \|x_k - x\| \\
&\leq 2C\varepsilon + |g_n(x_k) - g_m(x_k)|.
\end{aligned}$$

Da $\{g_n(x_k)\}_{n \in \mathbb{N}}$ konvergent und somit eine Cauchy-Folge ist, so erhalten wir

$$\limsup_{n,m \to \infty} |g_n(x) - g_m(x)| \leq 2C\varepsilon + \limsup_{n,m \to \infty} |g_n(x_k) - g_m(x_k)| = 2C\varepsilon.$$

Da ε eine beliebige positive Zahl ist, so folgt es, dass

$$\lim_{n,m \to \infty} |g_n(x) - g_m(x)| = 0,$$

was zu beweisen war. ∎

Beispiel. Sei H ein Hilbertraum über $\mathbb{R}$. Nach dem Satz 2.7 gilt $H^* \cong H$, und jedes Element $a \in H$ bestimmt ein Element von H^* mit

$$a(x) = (x, a).$$

Die schwache Konvergenz in H^* kann man als schwache Konvergenz in H wie folgt verstehen: für Elemente a_n, a aus H gilt $a_n \rightharpoonup a$ wenn $(x, a_n) \to (x, a)$ für alle $x \in H$. Nach dem Satz 5.7 erhalten wir, dass H schwach vollständig ist. Ist H separabel, so erhalten wir nach dem Satz 5.8, dass jede beschränkte Teilmenge von H (z.B. jede Kugel) schwach präkompakt ist.

Erinnern wir uns, dass jede orthonormale Folge $\{e_n\}_{n \in \mathbb{N}}$ in H nicht Norm-präkompakt ist da $\|e_n - e_m\|^2 = 2$ für alle $n \neq m$. Allerdings ist die Folge $\{e_n\}$ beschränkt und somit *schwach* präkompakt nach dem Satz 5.8. In der Tat konvergiert die ganze Folge schwach gegen 0, da für jedes $x \in X$

$$\sum_{n=1}^{\infty} |(x, e_n)|^2 \leq \|x\|^2 < \infty$$

und somit $(x, e_n) \to 0$, was genau bedeutet, dass $e_n \rightharpoonup 0$.

Beispiel. Seien $p, q \in (1, \infty)$ die konjugierten Hölder-Exponenten. Da $l^q \cong (l^p)^*$, so ist der Begriff von schwache Konvergenz in l^q wohldefiniert ist. Dann ist l^q schwach vollständig, und jede beschränkte Teilmenge von l^q ist schwach präkompakt.

Chapter 6

Sonstiges

15.07.22 **Vorlesung 27**

6.1 Satz von Stone-Weierstraß

Sei X ein kompakter metrischer Raum. Bezeichnen wir mit $C(X)$ die Menge von allen stetigen reellwertigen Funktionen auf X. Offensichtlich ist $C(X)$ ein Vektorraum über $\mathbb{R}$, und die sup-Norm

$$\|f\| = \sup_X |f|$$

eine Norm in $C(X)$ ist. Wie im Satz 1.7 beweist man, dass $C(X)$ ein Banachraum ist. Bemerken wir, dass in $C(X)$ nicht nur die linearen Operationen definiert sind, sondern auch die Multiplikation:

$$(fg)(x) = f(x)\,g(x) \quad \forall x \in X.$$

Die Multiplikation ist bilinear bezüglich der linearen Operationen, und es gilt

$$\|fg\| \le \|f\| \, \|g\| \, .$$

Ein Banachraum mit solcher Multiplikation heißt eine *Banachalgebra*. Somit ist $C(X)$ eine Banachalgebra. Diese Banachalgebra ist zusätzlich kommutativ, assoziativ, und hat die Einheit – die konstante Funktion 1.

Eine Teilmenge $\mathcal{A}$ von $C(X)$ heißt eine *Algebra* (oder Unteralgebra von $C(X)$), wenn $\mathcal{A}$ ein Unterraum ist und

$$f, g \in \mathcal{A} \Rightarrow fg \in \mathcal{A}.$$

Eine Algebra $\mathcal{A}$ heißt *unitär* wenn $1 \in \mathcal{A}$.

Hauptsatz 6.1 (Satz von Stone-Weierstraß) *Sei X ein kompakter metrischer Raum. Sei $\mathcal{A}$ eine Teilmenge von $C(X)$ mit den folgenden Eigenschaften:*

1. *$\mathcal{A}$ ist eine unitäre Algebra;*
2. *$\mathcal{A}$ trennt die Punkte von X, d.h. für jedes Paar $x, y \in X$ mit $x \neq y$ existiert eine Funktion $f \in \mathcal{A}$ mit $f(x) \neq f(y)$.*

Dann liegt $\mathcal{A}$ dicht in $C(X)$.

Beispiel. Seien $X = [a, b]$ und $\mathcal{P}$ die Menge von allen Polynomen mit reellen Koeffizienten. Offensichtlich ist $\mathcal{P}$ eine unitäre Algebra und $\mathcal{P}$ trennt die Punkte (z.B. das Polynom $f(x) = x$ trennt alle Punkte). Somit liegt $\mathcal{P}$ dicht in $C[a, b]$, was den 1. Approximationssatz von Weierstraß (Satz 2.19) beweist.

Beispiel. Sei X der Einheitskreis in $\mathbb{C}$ d.h.

$$X = \left\{ e^{ix} : x \in \mathbb{R} \right\}.$$

Jede 2π-periodische Funktion f auf $\mathbb{R}$ lässt sich als eine Funktion auf X betrachten, und umgekehrt, jede Funktion f auf X lässt sich auf eine 2π-periodische Funktion auf $\mathbb{R}$ erweitern. Insbesondere lässt sich der Raum $C(X)$ mit dem Raum $C_{2\pi}$ von stetigen 2π-periodischen Funktionen identifizieren.

Bezeichnen wir mit $\mathcal{T}$ die Menge von trigonometrischen Polynomen

$$f(x) = a_0 + \sum_{k=1}^{n} (a_k \cos kx + b_k \sin kx)$$

und betrachten die Elemente von $\mathcal{T}$ als Funktionen auf X d.h. $\mathcal{T} \subset C(X)$. Die Menge $\mathcal{T}$ ist eine Algebra da die Produkte der From $\varphi(kx)\psi(mx)$ wobei jede von Funktionen φ und ψ eine von Funktionen cos und sin ist, sich als lineare Kombinationen von $\sin(k \pm m)x$ und $\cos(k \pm m)x$ darstellen lassen. Die Algebra $\mathcal{T}$ ist unitäre da $1 \in \mathcal{T}$, und $\mathcal{T}$ trennt die Punkte von X, da für $x, y \in X$ mit $x \neq y$ gilt entweder $\sin x \neq \sin y$ oder $\cos x \neq \cos x$. Nach dem Satz 6.1 liegt $\mathcal{T}$ dicht in $C(X)$. Folglich liegt $\mathcal{T}$ dicht auch in $C_{2\pi}$, was den 2. Approximationssatz von Weierstraß (Satz 2.22) beweist.

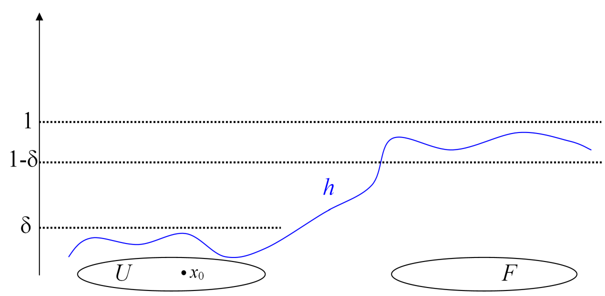
Beweis von Hauptsatz 6.1. Beweisen wir zunächst zwei Behauptungen.

Behauptung 1. *Für jedes $x_0 \in X$ und jede abgeschlossene Menge $F \subset X$ mit $x_0 \notin F$, gibt es eine offene Umgebung U von x_0, die von F disjunkt ist und die folgende Eigenschaft erfüllt: für jedes $\delta > 0$ existiert eine Funktion $h \in \mathcal{A}$ mit*

$$0 \leq h \leq 1 \tag{6.1}$$

und

$$\sup_U h \leq \delta, \qquad \inf_F h \geq 1 - \delta. \tag{6.2}$$



Da $\mathcal{A}$ die Punkte trennt, für jedes $x \neq x_0$ existiert eine Funktion $f_x \in \mathcal{A}$ mit

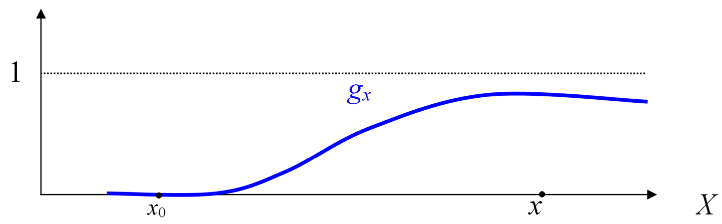
$$f_x(x_0) \neq f_x(x).$$

Wir können annehmen, dass $f_x(x_0) = 0$ indem wir eine Konstante aus f_x subtrahieren. Die Funktion

$$g_x = \frac{f_x^2}{2\|f_x^2\|}$$

gehört auch zu $\mathcal{A}$ und erfüllt

$$0 \leq g_x < 1, \quad g_x(x_0) = 0 \ \text{ und } \ g_x(x) > 0.$$



Betrachten wir die offene Menge $G_x = \{g_x > 0\}$. Dann ist $\{G_x\}_{x \in F}$ eine offene Überdeckung von F. Da F eine kompakte Menge ist, es gibt eine endliche Teilüberdeckung $G_{x_1}, ..., G_{x_m}$. Betrachten wir die Funktion

$$g = \frac{1}{m}\left(g_{x_1} + ... + g_{x_m}\right),$$

die zu $\mathcal{A}$ gehört und die folgenden Eigenschaften erfüllt:

$$0 \leq g < 1, \quad g(x_0) = 0 \ \text{ und } \ g(x) > 0 \text{ für alle } x \in F.$$

Da F kompakt ist, so haben wir

$$\gamma := \inf_F g > 0$$

und $\gamma < 1$. Setzen wir

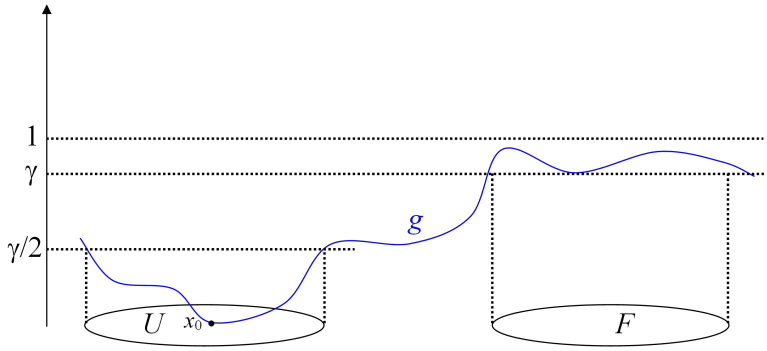
$$U = \{g < \gamma/2\}$$

so dass U eine offene Menge ist, $x_0 \in U$ und $U \cap F = \emptyset$. Somit haben wir

$$0 \leq g < 1$$

und

$$\sup_U g \leq \frac{\gamma}{2}, \quad \inf_F g = \gamma.$$



Um eine Funktion h zu bekommen die (6.2) mit beliebig kleinem δ erfüllt, benutzen wir das folgende Verfahren. Wählen wir zuerst eine ganze Zahl k so dass

$$\frac{1}{\gamma} < k < \frac{2}{\gamma},$$

und betrachten für alle $n \in \mathbb{N}$ die Funktionen

$$h_n = 1 - (1 - g^n)^{k^n}.$$

Offensichtlich haben wir $h_n \in \mathcal{A}$ und $0 \le h_n \le 1$. Zeigen wir, dass

$$\inf_F h_n \to 1 \text{ für } n \to \infty. \tag{6.3}$$

und

$$\sup_U h_n \to 0 \text{ für } n \to \infty. \tag{6.4}$$

Da $g \ge \gamma$ auf F, so haben wir

$$\inf_F h_n \ge 1 - (1 - \gamma^n)^{k^n}.$$

Da $g \le \gamma/2$ auf U, so haben wir

$$\sup_U h_n \le 1 - \left(1 - \left(\frac{\gamma}{2}\right)^n\right)^{k^n}.$$

Bezeichnen wir γ bzw $\gamma/2$ mit q, so dass $0 < q < 1$. Da $q^n \to 0$ für $n \to \infty$, wir haben

$$\ln\left((1 - q^n)^{k^n}\right) = k^n \ln(1 - q^n) \sim -k^n q^n = -(kq)^n.$$

Daraus folgt

$$\lim_{n\to\infty} \ln\left((1 - q^n)^{k^n}\right) = \begin{cases} -\infty, & kq > 1, \\ 0, & kq < 1 \end{cases}$$

und somit

$$\lim_{n\to\infty} \left(1 - (1 - q^n)^{k^n}\right) = \begin{cases} 1, & kq > 1, \\ 0, & kq < 1. \end{cases} \tag{6.5}$$

Da $k\gamma > 1$, so erhalten wir aus (6.5) mit $q = \gamma$ dass

$$1 - (1 - \gamma^n)^{k^n} \to 1 \text{ für } n \to \infty$$

und somit (6.3). Da $k\gamma/2 < 1$, so erhalten wir aus (6.5) mit $q = \gamma/2$ dass

$$1 - \left(1 - \left(\frac{\gamma}{2}\right)^n\right)^{k^n} \to 0 \text{ für } n \to \infty,$$

und somit (6.4).

Gegeben sei ein $\delta > 0$, so gibt es nach (6.3) und (6.4) ein $n \in \mathbb{N}$ mit

$$\inf_F h_n > 1 - \delta, \quad \sup_U h_n < \delta,$$

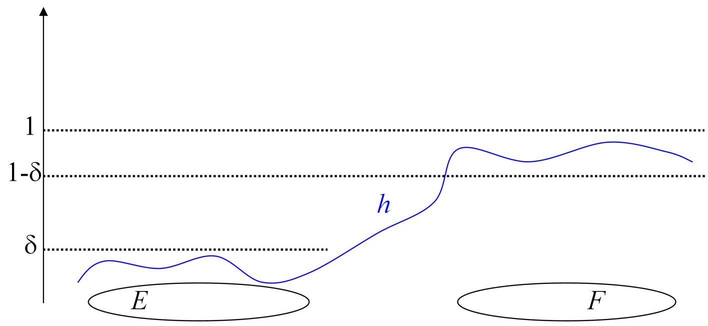
so dass die Funktion $h = h_n$ (6.2) erfüllt.

Behauptung 2. *Seien E, F zwei abgeschlossene disjunkte Teilmengen von X. Dann für jedes $\delta > 0$ existiert eine Funktion $h \in \mathcal{A}$ mit*

$$0 \le h \le 1$$

und

$$\sup_E h \le \delta, \qquad \inf_F h \ge 1 - \delta.$$



Für jedes $x \notin F$ sei U_x die Umgebung von x aus Behauptung 1: für jedes $\delta' > 0$ existiert eine Funktion $h_x \in \mathcal{A}$ mit

$$0 \le h_x \le 1$$

und

$$\sup_{U_x} h_x \le \delta', \qquad \inf_F h_x \ge 1 - \delta'.$$

Dann ist $\{U_x\}_{x \in E}$ eine offene Überdeckung von E. Somit existiert eine endliche Teilüberdeckung $U_{x_1}, ..., U_{x_n}$. Wählen wir $\delta' = \delta/n$ und setzen $h_k := h_{x_k}$ so dass $h_k \in \mathcal{A}$ und

$$0 \le h_k \le 1,$$

$$\sup_{U_{x_k}} h_k \le \delta/n \ , \quad \inf_F h_k \ge 1 - \delta/n.$$

Betrachten wir die Funktion

$$h = h_1 h_2 ... h_n.$$

Offensichtlich haben wir $h \in \mathcal{A}$ und $0 \leq h \leq 1$. Jedes $x \in E$ liegt in einem U_{x_k} woraus folgt

$$h_k(x) \leq \delta/n$$

und somit

$$h(x) \leq h_k(x) \leq \delta/n \leq \delta,$$

woraus folgt

$$\sup_E h \leq \delta.$$

Da für jedes $x \in F$ und für alle k gilt

$$h_k(x) \geq 1 - \delta/n,$$

so erhalten wir nach der Bernoullischen Ungleichung

$$h(x) \geq \left(1 - \frac{\delta}{n}\right)^n \geq 1 - \delta,$$

woraus folgt

$$\inf_F h \geq 1 - \delta,$$

was zu beweisen war.

Jetzt beweisen wir, dass $\mathcal{A}$ dicht in $C(X)$ liegt. Gegeben seien $f \in C(X)$ und $\varepsilon > 0$, beweisen wir, dass es eine Funktion $g \in \mathcal{A}$ gibt mit

$$\sup |f - g| \leq 2\varepsilon.$$

Wir können annehmen, dass $f > 0$ indem wir zu f eine große Konstante addieren. Betrachten wir für jedes $n \in \mathbb{N}$ die Mengen

$$E_n = \{f \leq (n-1)\varepsilon\}, \qquad F_n = \{f \geq n\varepsilon\}. \tag{6.6}$$

Nach der Behauptung 2 gibt es eine Funktion $h_n \in \mathcal{A}$ mit

$$0 \leq h_n \leq 1$$

und

$$\sup_{E_n} h_n \leq \delta, \quad \inf_{F_n} h_n \geq 1 - \delta.$$

Wählen wir eine ganze Zahl $N > \frac{1}{\varepsilon} \sup f$ und setzen $\delta = \frac{1}{N}$.

Bemerken wir, dass

$$\emptyset = E_1 \subset E_2 \subset ... \subset E_N \subset E_{N+1} = X$$

und

$$X = F_0 \supset F_1 \supset ... \supset F_N = \emptyset.$$

Betrachten wir die Funktion

$$g = \varepsilon \sum_{n=1}^{N} h_n$$

und schätzen $\sup |f-g|$ ab. Da $E_1 = \emptyset$ und $E_{N+1} = X$, so liegt jedes $x \in X$ in einer der Mengen $E_{k+1} \setminus E_k$, $k = 1, ..., N$.

Sei
$$x \in E_{k+1} \setminus E_k,$$
so dass nach (6.6)
$$(k-1)\,\varepsilon < f(x) \leq k\varepsilon.$$
Für $n \geq k+1$ haben wir $x \in E_{k+1} \subset E_n$ und somit $h_n(x) \leq \delta$. Da $h_n(x) \leq 1$ für alle n, so erhalten wir

$$\begin{aligned} g(x) &= \varepsilon \sum_{n=1}^{k} h_n(x) + \varepsilon \sum_{n=k+1}^{N} h_n(x) \\ &\leq \varepsilon k + \varepsilon \delta N \\ &= \varepsilon (k+1), \end{aligned}$$

und somit
$$g(x) - f(x) \leq \varepsilon (k+1) - (k-1)\,\varepsilon = 2\varepsilon. \tag{6.7}$$
Für $n \leq k-1$ haben wir $x \in E_k^c \subset F_{k-1} \subset F_n$ und somit $h_n(x) \geq 1-\delta$, woraus folgt

$$\begin{aligned} g(x) &\geq \varepsilon \sum_{n=1}^{k-1} h_n(x) \\ &\geq \varepsilon (1-\delta)(k-1) \\ &= \varepsilon (k-1) - \varepsilon\delta (k-1) \\ &\geq \varepsilon (k-1) - \varepsilon \delta N \\ &= \varepsilon (k-2) \end{aligned}$$

und somit
$$g(x) - f(x) \geq \varepsilon (k-2) - \varepsilon k = -2\varepsilon. \tag{6.8}$$
Da (6.7) und (6.8) für jedes $x \in X$ gelten, so erhalten wir

$$\sup_{x\in X} |f(x) - g(x)| \leq 2\varepsilon,$$

was zu beweisen war. ■

<u>Ende</u>

6.2 * Satz über die offene Abbildung

Definition. Eine Abbildung $A : X \to Y$ zwischen zwei metrischen Räumen heißt *offen*, wenn für jede offene Teilmenge $\Omega \subset X$ das Bild $A(\Omega)$ auch offen ist.

Beispiel. Sei $B : Y \to X$ eine stetige invertierbare Abbildung. Dann ist die inverse Abbildung $A = B^{-1}$ offen, da für jede offene Teilmenge $\Omega \subset X$ das Bild $A(\Omega) = B^{-1}(\Omega)$ nach der Stetigkeit von B offen ist.

Hauptsatz 6.2 (Satz über die offene Abbildung oder Satz von Banach-Schauder) *Seien X und Y Banachräume und sei $A : X \to Y$ eine stetige lineare surjektive Abbildung. Dann ist die Abbildung A offen.*

Bemerkung. Die Bedingung von Surjektivität ist wichtig, was das Gegenbeispiel $A = 0$ zeigt.

Beweis. Bezeichnen wir mit $U(x,r)$ die Kugel in X und mit $V(y,r)$ die Kugel in Y. Auch schreiben wir kurz $U(r) = U(0,r)$ und $V(r) = V(0,r)$. Beweisen wir zunächst die folgende Behauptung: es gibt ein $\varepsilon > 0$ mit

$$\overline{A(U(1))} \supset V(\varepsilon). \tag{6.9}$$

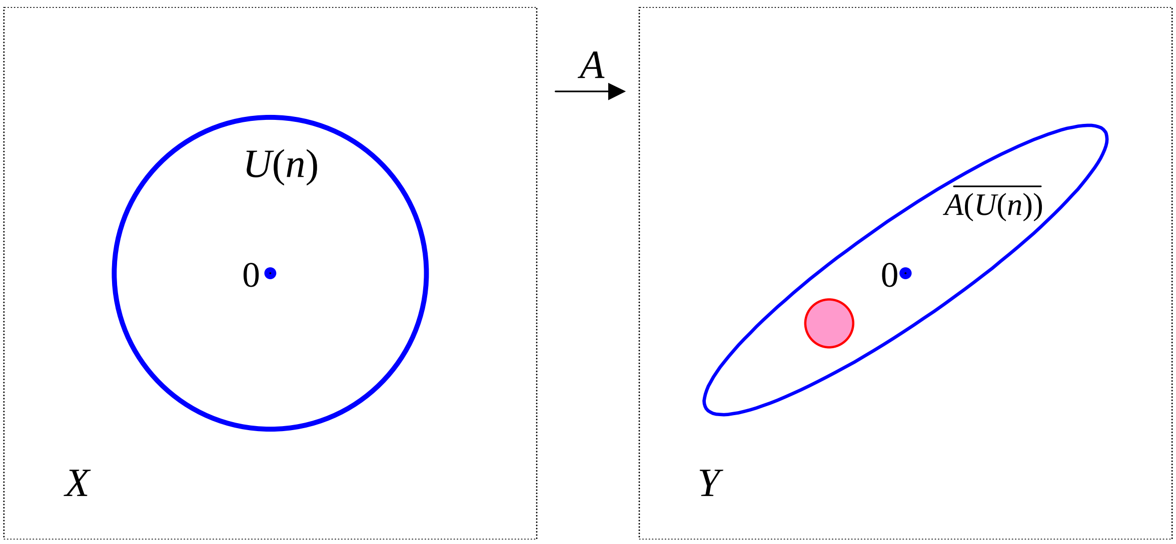
Da

$$X = \bigcup_{n=1}^{\infty} U(n)$$

und $A(X) = Y$, so erhalten wir, dass

$$Y = \bigcup_{n=1}^{\infty} A(U(n)).$$

Nach dem Satz 5.6 beschließen wir, dass eine von den Mengen $\overline{A(U(n))}$ eine offene Kugel enthält.



Nach der Linearität von A haben wir

$$A(U(n)) = 2nA\left(U\left({}^1/_2\right)\right),$$

woraus folgt, dass auch $\overline{A(U({}^1/_2))}$ eine offene Kugel enthält. Wählen wir in dieser Kugel einen Punkt y_0, der auch in der Bildmenge $A(U({}^1/_2))$ liegt. Dann für ein hinreichend kleines $\varepsilon > 0$ gilt

$$\overline{A(U({}^1/_2))} \supset V(y_0, \varepsilon).$$

Nach der Wahl von y_0 existiert ein $x_0 \in U({}^1/_2)$ mit $Ax_0 = y_0$. Da für jedes $x \in U({}^1/_2)$ gilt $x - x_0 \in U(1)$, so erhalten wir

$$U(1) \supset U\left({}^1/_2\right) - x_0,$$

$$A(U(1)) \supset A\left(U\left({}^1/_2\right) - x_0\right) = A\left(U\left({}^1/_2\right)\right) - y_0$$

und somit

$$\overline{A(U(1))} \supset \overline{A(U({}^1/_2))} - y_0 \supset V(y_0, \varepsilon) - y_0 = V(\varepsilon),$$

was ergibt (6.9).

Die Behauptung (6.9) impliziert nach der Linearität von A das folgende: für jedes r gilt

$$\overline{A(U(r))} \supset V(\varepsilon r). \tag{6.10}$$

Beweisen wir, dass

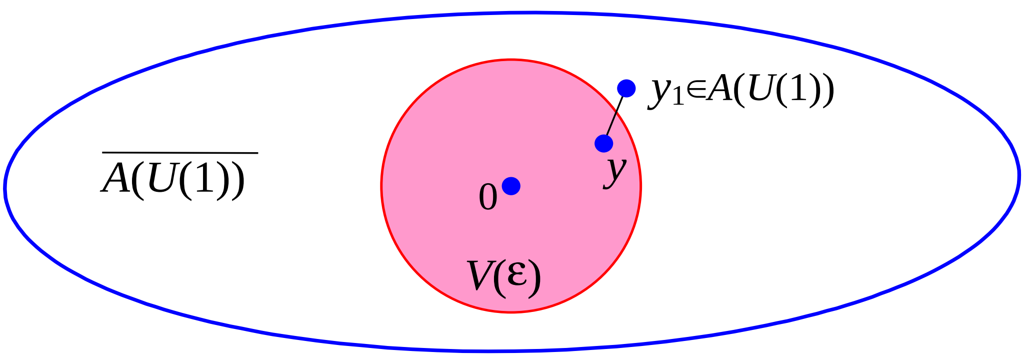
$$A(U(2)) \supset V(\varepsilon). \tag{6.11}$$

Nehmen wir ein $y \in V(\varepsilon)$ und beweisen, dass $y = Ax$ für ein $x \in U(2)$ gilt. Nach (6.9) oder (6.10) gilt

$$y \in \overline{A(U(1))}$$

und somit existiert ein $y_1 \in A(U(1))$ mit beliebig kleinem $\|y - y_1\|$, insbesondere mit

$$\|y - y_1\| < \varepsilon/2.$$

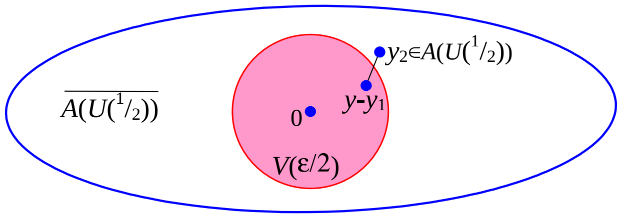


Da $y - y_1 \in V(\varepsilon/2)$ und nach (6.10)

$$\overline{A(U({}^1/_2))} \supset V(\varepsilon/2),$$

so existiert ein $y_2 \in A\left(U\left({}^1/_2\right)\right)$ mit beliebig kleinem $\|(y-y_1)-y_2\|$, insbesondere mit

$$\|y-y_1-y_2\| < \varepsilon/4.$$



Da $y-y_1-y_2 \in V(\varepsilon/4)$ und somit

$$y-y_1-y_2 \in \overline{A\left(U\left({}^1/_4\right)\right)},$$

so existiert ein $y_3 \in A\left(U\left({}^1/_4\right)\right)$ mit beliebig kleinem $\|(y-y_1-y_2)-y_3\|$, insbesondere mit

$$\|y-y_1-y_2-y_3\| < \varepsilon/8.$$

Per Induktion erhalten wir zwei Folgen $\{x_n\}$ und $\{y_n\}$ mit den folgenden Eigenschaften:

1. $y_n \in A\left(U\left(2^{-(n-1)}\right)\right)$
2. $\|y-y_1-...-y_n\| < \varepsilon 2^{-n}$.

Dann existiert $x_n \in U\left(2^{-(n-1)}\right)$ mit $y_n = Ax_n$. Da $\|x_n\| < 2^{-(n-1)}$ und somit

$$\sum_{n=1}^{\infty}\|x_n\| < \sum_{n=1}^{\infty}2^{-(n-1)} = 2 < \infty,$$

so ist die Reihe $\sum_{n=1}^{\infty} x_n$ konvergent nach dem Satz 1.5. Setzen wir $x = \sum_{n=1}^{\infty} x_n$ und bemerken, dass

$$\|x\| \le \sum_{n=1}^{\infty}\|x_n\| < 2,$$

so dass $x \in U(2)$. Da A stetig ist, so gilt

$$Ax = \sum_{n=1}^{\infty} Ax_n = \sum_{n=1}^{\infty} y_n = y,$$

woraus folgt, dass $y \in A(U(2))$, was beweist (6.11).

Es folgt aus (6.11), dass für alle $x \in X$ und $r > 0$ gilt

$$A(U(x,r)) \supset V\left(y, \frac{\varepsilon}{2}r\right) \tag{6.12}$$

mit $y = Ax$.

Sei Ω eine beliebige offene Teilmenge von X. Beweisen wir, dass $A(\Omega)$ offen ist. Für jedes $y \in A(\Omega)$ existiert $x \in \Omega$ mit $y = Ax$. Dann existiert ein $r > 0$ so dass $U(x,r) \subset \Omega$. Nach (6.12) erhalten wir

$$A(\Omega) \supset A(U(x,r)) \supset V\left(y, \frac{\varepsilon}{2}r\right),$$

woraus folgt, dass $A(\Omega)$ offen ist. ■

6.3 * Satz von der inversen Abbildung und Zerlegung von Spektrum

Korollar 6.3 (Satz von der inversen Abbildung oder Satz vom stetigen Inversen) *Sei A eine stetige lineare bijektive Abbildung zwischen Banachräumen X und Y. Dann ist die inverse Abbildung A^{-1} auch stetig.*

Beweis. Die inverse Abbildung $B = A^{-1}$ existiert weil A bijektiv ist. Die Abbildung $B : Y \to X$ ist genau dann stetig, wenn für jede offene Menge $\Omega \subset X$ das Urbild $B^{-1}(\Omega)$ offen ist. Dies ist wirklich der Fall, da $B^{-1}(\Omega) = A(\Omega)$ offen nach dem Satz 6.2 ist. ■

Erinnern wir die Definition des Spektrums. Für einen Operator A im Banachraum X über $\mathbb{C}$ gilt $\lambda \in \sigma(A)$ genau dann, wenn

1. entweder $\ker(A - \lambda I) \neq \{0\}$
2. oder $\operatorname{im}(A - \lambda I) \neq X$
3. oder $(A - \lambda I)^{-1}$ existiert aber ist nicht beschränkt.

Nehmen wir an, dass die Möglichkeiten 1 und 2 gelten nicht, d.h. $\ker(A - \lambda I) = \{0\}$ und $\operatorname{im}(A - \lambda I) = X$. Dann ist die Abbildung $A - \lambda I$ eine Bijektion, so dass die inverse Abbildung $(A - \lambda I)^{-1}$ existiert. Ist A beschränkt, so ist auch $A - \lambda I$ beschränkt, und nach Korollar 6.3 ist $(A - \lambda I)^{-1}$ beschränkt. Somit kann diese Möglichkeit nicht geschehen.

Erinnern wir auch, dass jedes λ mit $\ker(A - \lambda I) \neq \{0\}$ heißt ein Eigenwert von A, und die Menge von allen Eigenwerten von A heißt das Punktspektrum von A und wird mit $\sigma_p(A)$ bezeichnet. Somit erhalten wir die folgende Aussage.

Korollar 6.4 *Sei A ein beschränkter Operator im Banachraum X. Dann für jedes $\lambda \in \sigma(A) \setminus \sigma_p(A)$ gilt $\operatorname{im}(A - \lambda I) \neq X$.*

Für selbstadjungierten Operatoren gilt mehr.

Korollar 6.5 *Sei A ein selbstadjungierter Operator im Hilbertraum H. Dann gilt die Äquivalenz*

$$\lambda \in \sigma(A) \setminus \sigma_p(A) \Leftrightarrow \operatorname{im}(A - \lambda I) \neq H \text{ und } \overline{\operatorname{im}(A - \lambda I)} = H.$$

Beweis. Falls $\lambda \notin \sigma(A)$, so gilt $\operatorname{im}(A - \lambda I) = H$. Falls $\lambda \in \sigma_p(A)$ so gilt $\ker(A - \lambda I) \neq \{0\}$ und somit nach dem Satz 3.6

$$\overline{\operatorname{im}(A - \lambda I)} = \ker(A - \lambda I)^{\perp} \neq H.$$

Nehmen wir jetzt an, dass $\lambda \in \sigma(A) \setminus \sigma_p(A)$. Dann nach Korollar 6.4 gilt $\operatorname{im}(A - \lambda I) \neq H$, und nach dem Satz 3.6

$$\overline{\operatorname{im}(A - \lambda I)} = \ker(A - \lambda I)^{\perp} = H,$$

da $\ker(A - \lambda I) = \{0\}$. ■

Definition. Die Menge von $\lambda \in \mathbb{C}$ mit

$$\operatorname{im}(A - \lambda I) \neq H \text{ und } \overline{\operatorname{im}(A - \lambda I)} = H.$$

heißt das *stetige Spektrum* des Operators A und wird mit $\sigma_c(A)$ bezeichnet.

Dann ergibt Korollar 6.5 die folgende Zerlegung des Spektrums von selbstadjungierten Operatoren:

$$\sigma(A) = \sigma_p(A) \sqcup \sigma_c(A). \tag{6.13}$$

Beispiel. Betrachten wir den Operator

$$A = \sum_{k=1}^{\infty} \lambda_k P_{U_k},$$

wobei $\{\lambda_k\}_{k=1}^{\infty}$ eine beschränkt Folge von reellen Zahlen ist, und $\{U_k\}$ eine Folge von zueinander orthogonalen abgeschlossenen Unterräumen von H mit $\bigoplus_{k=1}^{\infty} U_k = H$ und $U_k \neq \{0\}$ ist. Dann ist A selbstadjungiert und

$$\sigma(A) = \overline{\{\lambda_k\}}.$$

Die Eigenwerte von A sind genau die Zahlen λ_k, d.h.

$$\sigma_p(A) = \{\lambda_k\}.$$

Nach (6.13) erhalten wir

$$\sigma_c(A) = \overline{\{\lambda_k\}} \setminus \{\lambda_k\}.$$

Z.B. ist $\{\lambda_k\}$ eine Folge von allen rationalen Zahlen in einem Intervall $[a, b]$, so ist $\sigma_c(A)$ die Menge von allen irrationalen Zahlen in $[a, b]$.

Bemerkung. Sei H der Hamilton-Operator von Schrödinger-Gleichung (obwohl dieser Operator unbeschränkt ist, viele von Theorie von beschränkten selbstadjungierten Operatoren gilt für H auch). Das Spektrum $\sigma(H)$ hat die folgende physikalische Bedeutung: $\sigma(H)$ ist genau die Menge von Werten von Energie im gegebenen System. Die Werte von dem Punktspektrum von H entsprechen zu gebundenen Zuständen des System, die mit den Quanteneffekten von Interaktion von den Teilchen zu tun haben, wobei die einzelnen Teilchen das stetige Spektrum haben.

6.4 * Kompaktheit in $C[a, b]$ und Satz von Arzela-Ascoli

Satz 6.6 (Satz von Arzela-Ascoli) *Eine Teilmenge M von $C[a, b]$ ist präkompakt genau dann, wenn die folgenden zwei Bedingungen erfüllt sind:*

- *die Menge M ist beschränkt, d.h. es gibt eine Konstante C mit $\sup|f| < C$ für alle $f \in M$;*
- *die Menge ist* gleichgradig stetig, *d.h. für jedes $\varepsilon > 0$ existiert δ mit*

  $$|f(x) - f(y)| < \varepsilon$$

  für alle $f \in M$ und alle $x, y \in [a, b]$ mit $|x - y| < \delta$.

Beweis. Beweisen wir, dass jede Folge $F = \{f_n\}_{n=1}^{\infty} \subset M$ eine konvergente Teilfolge enthält. Wir sagen, dass die Folge F an der Stelle $x \in [a,b]$ konvergiert, fall die numerische Folge $\{f_n(x)\}$ konvergiert. Sei $\{r_k\}_{k=1}^{\infty}$ die Folge von rationalen Zahlen auf $[a,b]$. Da die Folge F an der Stelle r_1 beschränkt ist, so hat F eine Teilfolge F_1 die an r_1 konvergiert. Analog hat F_1 eine Teilfolge F_2 die an r_2 konvergiert, usw. Für jedes n ist F_n eine Teilfolge von F_{n-1} die an r_n konvergiert. Sei g_n der n-te Glied von F_n. Dann ist die Folge $G = \{g_n\}$ eine Teilfolge von F die an allen Stellen r_k konvergiert.

Beweisen wir, dass die Folge G gleichmäßig auf $[a,b]$ konvergiert. Es reicht zu zeigen, dass G eine Cauchy-Folge in $C[a,b]$ ist. Gegeben ist ein $\varepsilon > 0$, finden wir δ so dass

$$|x-y| < \delta \Rightarrow f(x) - f(y) < \varepsilon \quad \text{für alle } f \in M.$$

Dann wählen wir eine Folge $q_1 < q_2 < ... < q_l$ von rationalen Zahlen auf $[a,b]$ so dass

$$[a,b] \subset \bigcup_{k=1}^{l} (q_k - \delta, q_k + \delta).$$

Da die Folgen $\{g_n(q_k)\}_{n=1}^{\infty}$ konvergieren, so existiert ein N mit

$$\forall n,m \geq N \quad \forall k = 1,..,l \ \text{ gilt } \ |g_n(q_k) - g_m(q_k)| < \varepsilon.$$

Jedes $x \in [a,b]$ liegt in einem Intervall $(q_k - d, q_k + \delta)$, so dass $|x - q_k| < \delta$. Dann gilt für alle g_n

$$|g_n(x) - g_n(q_k)| < \varepsilon$$

woraus folgt

$$\begin{aligned}|g_n(x) - g_m(x)| &\leq |g_n(x) - g_n(q_k)| + |g_n(q_k) - g_m(q_k)| + |g_m(q_k) - g_m(x)| \\ &\leq 3\varepsilon.\end{aligned}$$

Somit ist die Folge $\{g_n\}$ eine Cauchy-Folge in $C[a,b]$.

Zweiter Beweis. Sei M totalbeschränkt und sei $f_1, ..., f_m$ ein ε-Netz von M, d.h. für jede Funktion $f \in M$ existiert f_k mit $\|f - f_k\| < \varepsilon$ (wobei $\|f\| = \sup|f|$). Setzen wir $C = \max_k \|f_k\|$ und erhalten für alle $f \in M$, dass

$$\|f\| \leq C + \varepsilon.$$

Somit ist die Menge M beschränkt. Jede Funktion f_k ist gleichmäßig stetig, so dass existiert δ mit

$$|f_k(x) - f_k(y)| < \varepsilon$$

für alle f_k und alle $x, y \in [a,b]$ mit $|x-y| < \delta$. Daraus folgt für alle $f \in M$

$$\begin{aligned}|f(x) - f(y)| &\leq |f(x) - f_k(x)| + |f_k(x) - f_k(y)| + |f_k(y) - f(y)| \\ &< 3\varepsilon.\end{aligned}$$

Somit ist M gleichgradig stetig.

Sei jetzt M beschränkt und gleichgradig stetig, und beweisen wir, dass M totalbeschränkt ist. Sei $c = \sup_{f \in M} \|f\|$. Gegeben sind ε und δ wie in (b), betrachten wir eine Zerlegung $x_0 < x_1 < .... < x_n = 0$ des Intervalls $[a,b]$ mit $x_{k+1} - x_k < \delta$ und eine Zerlegung $y_0 < y_1 < .... < y_m$ des Intervalles $[-c,c]$ mit $y_{l+1} - y_l < \varepsilon$. Bezeichnen wir mit N die

Menge von stetigen Funktionen auf $[a, b]$ die an den Stellen x_k die Werte y_l annehmen, und zwischen x_k und x_{k+1} immer linear sind. Die Menge N ist offensichtlich endlich. Beweisen wir, dass N ein 5ε-Netz von M ist. Für jede Funktion $f \in M$ definieren wir eine Funktion $h \in N$ wie folgt. Für jedes x_k liegt $f(x_k)$ in einem Intervall $[y_l, y_{l+1}]$; dann setzen wir $h(x_k) = y_l$, was die Funktion h eindeutig bestimmt. Daraus folgt, dass für alle k

$$|h(x_k) - f(x_k)| < \varepsilon.$$

Dann für jedes $x \in [x_k, x_{k+1}]$ gilt

$$\begin{aligned}
|h(x) - f(x)| &\leq |h(x) - h(x_k)| + |h(x_k) - f(x_k)| + |f(x_k) - f(x)| \\
&\leq |h(x_{k+1}) - h(x_k)| + 2\varepsilon \\
&\leq |h(x_{k+1}) - f(x_{k+1})| + |f(x_{k+1}) - f(x_k)| + |f(x_k) - h(x_k)| + 2\varepsilon \\
&\leq 5\varepsilon.
\end{aligned}$$

was zu beweisen war. ■

Beispiel. Sei $\alpha \in (0, 1)$. Dann jede beschränkte Teilmenge von $C^\alpha[a, b]$ ist präkompakt in $C[a, b]$.

6.5 * Dualraum von $C[a, b]$

Geben wir zwei Beispiele von Elementen von $C[a, b]^*$ an.

Beispiel. Die Abbildung

$$C[a, b] \ni f \mapsto \Phi(f) := \int_a^b f(t)\, dt \in \mathbb{R}$$

ist offensichtlich linear und beschränkt, da

$$|\Phi(f)| \leq (b - a) \sup |f|.$$

Somit ist Φ ein Element von $C[a, b]^*$.

Beispiel. Für jedes $t \in [a, b]$ definieren wir die Abbildung

$$C[a, b] \ni f \mapsto \delta_t(f) := f(t) \in \mathbb{R}.$$

Die Abbildung δ_t ist linear und beschränkt, da

$$|\delta_t(f)| = |f(t)| \leq \sup |f|.$$

Somit ist δ_t ein Element von $C[a, b]^*$. Das Funktional δ_t heißt die *Dirac-Funktion* an der Stelle t.

Bemerken wir, dass $\delta_t(f)$ sich als Riemann-Stieltjes-Integral[1] darstellen lässt:

$$\delta_t(f) = \int_a^b f(s)\, d\theta_t(s)$$

[1] Siehe Abschnitt 7.5.

wobei

$$\theta_t(s) = \begin{cases} 1, & s \geq t, \\ 0, & s < t. \end{cases}$$

Es stellt sich heraus, dass alle Elemente von $C[a,b]^*$ mit Hilfe von Riemann-Stieltjes-Integral beschrieben werden können. Für jede monoton steigende Funktion F definieren wir das Funktional Φ_F auf $C[a,b]$ mit

$$\Phi_F(f) = \int_a^b f dF,$$

wobei die rechte Seite das Riemann-Stieltjes-Integral ist. Das Funktional Φ_F ist offensichtlich linear, aber auch beschränkt, da

$$|\Phi_F(f)| \leq (F(b) - F(a)) \sup |f|.$$

Somit ist Φ_F ein Element von $C[a,b]^*$. Diese Konstruktion kann weiter verallgemeinert werden indem man statt einer monoton steigenden Funktion F die Differenz von zwei monoton steigenden Funktionen benutzt.

Für jede Funktion F auf dem Intervall $[a,b]$ definieren wir die Variation von F mit

$$\operatorname{Var}_{[a,b]} F = \sup_Z \sum_{k=1}^n |F(\lambda_k) - F(\lambda_{k-1})|,$$

wobei

$$Z = \{a = \lambda_0 < \lambda_1 < ... < \lambda_n = b\} \tag{6.14}$$

eine beliebige Zerlegung von $[a,b]$ ist. Die Variation nimmt die Werte in $[0,\infty]$ und erfüllt die folgenden Eigenschaften:

1. $\operatorname{Var}_{[a,b]}(F+G) \leq \operatorname{Var}_{[a,b]} F + \operatorname{Var}_{[a,b]} G$
2. $\operatorname{Var}_{[a,b]}(\alpha F) = |\alpha| \operatorname{Var}_{[a,b]} F$ für alle $\alpha \in \mathbb{R}$
3. $\operatorname{Var}_{[a,b]} F = 0 \Leftrightarrow F = \text{const auf } [a,b]$.

Gilt $\operatorname{Var}_{[a,b]} F < \infty$ so heißt F eine Funktion *von beschränkter Variation*. Die Menge von Funktionen F mit $\operatorname{Var}_{[a,b]} F < \infty$ wird mit $BV[a,b]$ bezeichnet. Es folgt aus den obigen Eigenschaften, dass $BV[a,b]$ ein Vektorraum ist, und $\operatorname{Var}_{[a,b]} F$ eine Halbnorm in $BV[a,b]$ ist.

Zum Beispiel, jede monotone Funktion F hat beschränkte Variation, da in diesem Fall

$$\operatorname{Var}_{[a,b]} F = |F(b) - F(a)|.$$

Dann auch die Differenz $F - G$ von zwei monotone Funktionen gehört zu $BV[a,b]$. Man kann zeigen, dass jede Funktion in $BV[a,b]$ eine Differenz zweier monotonen Funktionen ist[2].

[2] Man kann beweisen, dass die Funktion $G(x) = \operatorname{Var}_{[a,x]} F$ monoton steigend ist, und auch $H(x) := G(x) - F(x)$ monoton steigend. Da gilt $F = G - H$, so ist F eine Differenz zweier monoton steigenden Funktionen.

Für Funktionen $f \in C[a,b]$ und $F \in BV[a,b]$ definiert man das Riemann-Stieltjes-Integral

$$\int_a^b f dF = \lim_{\delta_Z \to 0} S(f, F, Z, T) \tag{6.15}$$

wobei Z eine Zerlegung von $[a,b]$ ist, wie in (6.14), $T = \{t_k\}_{k=1}^n$ die Folge von Zwischenstellen, und $S(f, F, Z, T)$ die entsprechende Riemann-Stieltjes Summe, d.h.

$$S(f, F, Z, T) = \sum_{k=1}^n f(t_k)\left(F(\lambda_k) - F(\lambda_{k-1})\right). \tag{6.16}$$

Man kann zeigen, dass der Grenzwert in (6.15) existiert für alle $f \in C[a,b]$ und $F \in BV[a,b]$, und somit ist das Integral $\int_a^b f dF$ wohldefiniert. Das ist eine Verallgemeinerung von Definition von Riemann-Stieltjes-Integral bezüglich monotoner Funktion F, die wir im Abschnitt 7.5 besprochen haben.

Für jedes $F \in BV[a,b]$ betrachten wir die Abbildung

$$\begin{aligned} \Phi_F &: C[a,b] \to \mathbb{R} \\ \Phi_F(f) &= \int_a^b f dF. \end{aligned}$$

Die Abbildung Φ_F ist offensichtlich linear. Sie ist auch beschränkt, da (6.16) impliziert

$$|S(f, F, Z, T)| \le \sup|f| \sum_{k=1}^n |F(\lambda_k) - F(\lambda_{k-1})| \le \sup|f| \operatorname{Var}_{[a,b]} F$$

woraus folgt

$$\left|\int_a^b f dF\right| \le \sup|f| \operatorname{Var}_{[a,b]} F$$

und somit

$$\|\Phi_F\| \le \operatorname{Var}_{[a,b]} F. \tag{6.17}$$

Somit bestimmt jede Funktion $F \in BV[a,b]$ ein beschränktes Funktional $\Phi_F \in C[a,b]^*$. Die Umkehrung gilt auch wie folgt.

Hauptsatz 6.7 (2. Darstellungssatz von Riesz) *Für jedes $\Phi \in C[a,b]^*$ gibt es eine Funktion $F \in BV[a,b]$ mit*

$$\Phi = \Phi_F$$

und

$$\|\Phi\| = \operatorname{Var}_{[a,b]} F. \tag{6.18}$$

Beweis. Bezeichnen wir mit $B[a,b]$ die Menge von allen beschränkten Funktionen auf $[a,b]$. Offensichtlich ist $B[a,b]$ ein Vektorraum. Darüber hinaus ist $B[a,b]$ ein normierter Vektorraum mit der sup-Norm, und $C[a,b]$ ist ein Unterraum von $B[a,b]$. Sei Φ ein Element von $C[a,b]^*$. Da Φ ein beschränktes Funktional auf $C[a,b]$ ist, so existiert nach dem Satz von Hahn-Banach eine Fortsetzung von Φ (auch mit Φ bezeichnet) auf $B[0,1]$ mit Erhaltung der Norm. Dann ist Φ auf allen beschränkten Funktionen auf $[a,b]$ definiert. Für jedes $x \in [a,b]$ setzen wir

$$F(x) = \Phi\left(\mathbf{1}_{[a,x]}\right)$$

und zeigen, dass $F \in BV\,[a,b]$. Gegeben sei eine Zerlegung (6.14) von $[a,b]$, setzen wir für $k=1,...,n$

$$s_k = \operatorname{sgn}\left(F\left(\lambda_k\right) - F\left(\lambda_{k-1}\right)\right)$$

und betrachten die Funktion

$$f = s_1 \mathbf{1}_{[\lambda_0,\lambda_1]} + \sum_{k=2}^{n} s_k \mathbf{1}_{(\lambda_{k-1},\lambda_k]}.$$

Diese Funktion ist beschränkt (und sogar $\|f\| = \sup|f| = 1$) und somit $\Phi(f)$ ist definiert. Dann erhalten wir

$$\begin{aligned}
\Phi(f) &= s_1 \Phi\left(\mathbf{1}_{[\lambda_0,\lambda_1]}\right) + \sum_{k=2}^{n} s_k \Phi\left(\mathbf{1}_{(\lambda_{k-1},\lambda_k]}\right) \\
&= s_1 \Phi\left(\mathbf{1}_{[\lambda_0,\lambda_1]}\right) + \sum_{k=2}^{n} s_k \left(\Phi\left(\mathbf{1}_{[a,\lambda_k]}\right) - \Phi\left(\mathbf{1}_{[a,\lambda_{k-1}]}\right)\right) \\
&= \sum_{k=1}^{n} s_k \left(F\left(\lambda_k\right) - F\left(\lambda_{k-1}\right)\right) \\
&= \sum_{k=1}^{n} \left|F\left(\lambda_k\right) - F\left(\lambda_{k-1}\right)\right|,
\end{aligned}$$

woraus folgt

$$\sum_{k=1}^{n} \left|F\left(\lambda_k\right) - F\left(\lambda_{k-1}\right)\right| = \Phi(f) \le \|\Phi\| \, \|f\| = \|\Phi\|$$

und somit

$$\operatorname{Var}_{[a,b]} F \le \|\Phi\| < \infty. \tag{6.19}$$

Beweisen wir jetzt, dass

$$\Phi(f) = \Phi_F(f) \ \text{ für alle } f \in C\,[a,b]\,. \tag{6.20}$$

Für jede Zerlegung $Z = \{\lambda_k\}_{k=0}^{n}$ von $[a,b]$ definieren wir die Treppenfunktion

$$f_Z = f\left(\lambda_1\right) \mathbf{1}_{[\lambda_0,\lambda_1]} + \sum_{k=2}^{n} f\left(\lambda_k\right) \mathbf{1}_{(\lambda_{k-1},\lambda_k]}.$$

Diese Funktion liegt in $B\,[0,1]$, und es gilt wie oberhalb

$$\Phi\left(f_Z\right) = f\left(\lambda_1\right) F\left(\lambda_1\right) + \sum_{k=2}^{n} f\left(\lambda_k\right) \left(F\left(\lambda_k\right) - F\left(\lambda_{k-1}\right)\right) = S\left(f, F, Z\right),$$

wobei $S(f,F,Z)$ die Riemann-Stieltjes Summe mit Zwischenstellen $t_k = \lambda_k$ ist. Da f gleichmäßig stetig auf $[a,b]$ ist, so erhalten wir

$$f_Z \rightrightarrows f \quad \text{für } \delta_Z \to 0.$$

Da Φ stetig ist, so folgt es daraus, dass

$$\Phi\left(f_Z\right) \to \Phi(f)\,.$$

Andererseits, es gilt nach Definition des Riemann-Stieltjes Integrals, dass

$$\Phi\left(f_Z\right) = S\left(f, F, Z\right) \to \int_a^b f dF = \Phi_F\left(f\right),$$

woraus (6.20) folgt.

Der Vergleich von (6.17) und (6.19) ergibt (6.18). ■

Die Abbildung $F \mapsto \Phi_F$ ist allerdings nicht injektiv. Zum Beispiel, $F - G = \text{const}$ impliziert, dass $\Phi_F = \Phi_G$.

Für jede Funktion $F \in BV\left[a, b\right]$ und $x \in (a, b)$ definieren wir

$$F\left(x+\right) = \lim_{\varepsilon \to 0+} F\left(x + \varepsilon\right).$$

Der Grenzwert existiert da F eine Differenz zweier monoton steigenden Funktionen ist. Dann definieren wir die *rechtsstetige Version* F_r von F mit

$$F_r\left(x\right) = \begin{cases} F\left(x\right), & x = a \text{ oder } x = b, \\ F\left(x+\right), & a < x < b. \end{cases}$$

Dann F_r ist wieder eine Differenz von monoton steigenden Funktionen und somit $F_r \in BV\left[a, b\right]$. Die Menge U von Punkten wo F unstetig ist, ist höchstens abzählbar, woraus folgt, dass F und F_r außer U übereinstimmen. Es folgt auch dass für alle $x \in (a, b)$

$$F_r\left(x+\right) = F\left(x+\right) = F_r\left(x\right),$$

so dass F_r rechtsstetig in (a, b) ist.

Lemma 6.8 *Für jede Funktion $F \in BV\left[a, b\right]$ gelten*

$$\Phi_{F_r} = \Phi_F$$

und

$$\operatorname{Var}_{[a,b]} F_r \le \operatorname{Var}_{[a,b]} F. \tag{6.21}$$

Beweis. In der Tat, die Zerlegung $Z = \{\lambda_k\}_{k=0}^n$ in der Definition des Riemann-Stieltjes Integral lässt sich mit beliebig kleiner Feinheit δ_Z so wählen dass alle λ_k außer U liegen, woraus folgt, dass für beliebige Funktion $f \in C\left[a, b\right]$

$$S\left(f, F, Z, T\right) = S\left(f, F_r, Z, T\right).$$

Für $\delta_Z \to 0$ erhalten wir

$$\Phi_{F_r}\left(f\right) = \int_a^b f dF_r = \int_a^b f dF = \Phi_F\left(f\right).$$

Nach Definition haben wir

$$\operatorname{Var}_{[a,b]} F_r = \sup_Z \sum_{k=1}^n \left|F_r\left(\lambda_k\right) - F_r\left(\lambda_{k-1}\right)\right|,$$

wobei $Z = \{\lambda_k\}_{k=0}^n$ eine beliebige Zerlegung von $[a,b]$ ist. Es folgt aus der Definition von F_r dass für jedes $\varepsilon > 0$ es eine Zerlegung $Z' = \{\lambda'_k\}_{n=0}^n$ gilt so dass $\lambda'_k \geq \lambda_k$ und

$$\sum_{k=1}^n |F(\lambda'_k) - F_r(\lambda_k)| < \varepsilon.$$

Daraus folgt, dass

$$\sum_{k=1}^n |F_r(\lambda_k) - F_r(\lambda_{k-1})| \leq \varepsilon + \sum_{k=1}^n \left|F(\lambda'_k) - F_r\left(\lambda'_{k-1}\right)\right| \leq \varepsilon + \operatorname{Var}_{[a,b]} F.$$

Da ε beliebig ist, so erhalten wir (6.21). ■

Lemma 6.9 *Seien F und G Funktionen aus $BV[a,b]$ die auf (a,b) rechtsstetig sind. Gilt $\Phi_F = \Phi_G$ so gilt*

$$F(x) - F(a) = G(x) - G(a) \quad \textit{für alle } x \in [a,b]. \tag{6.22}$$

Beweis. Seien F und G zuerst monoton steigend. Fixieren wir ein $s \in (a,b)$ und betrachten die Folge $\{f_n\}$ von stetigen Funktionen wie folgt:

$$f_n(x) = \begin{cases} 1, & a \leq x \leq s \\ 0, & x > s + \frac{1}{n} \\ \text{linear}, & s \leq x \leq s + \frac{1}{n}. \end{cases}$$

Wir haben

$$\int_a^b f_n dF = \int_a^s f_n dF + \int_s^{s+\frac{1}{n}} f_n dF.$$

Es gilt

$$\int_a^s f_n dF = \int_a^s 1 dF = F(s) - F(a).$$

Da $f_n \leq 1$, so gilt

$$0 \leq \int_s^{s+\frac{1}{n}} f_n dF \leq F\left(s + \frac{1}{n}\right) - F(s).$$

Da F an s rechtsstetig ist, so gilt

$$F\left(s + \frac{1}{n}\right) - F(s) \to 0 \ \text{ für } n \to \infty,$$

woraus folgt

$$\int_s^{s+\frac{1}{n}} f_n dF \to 0$$

und somit

$$\lim_{n\to\infty} \Phi_F(f_n) = \lim_{n\to\infty} \int_a^b f_n dF = F(s) - F(a).$$

Analog gilt

$$\lim_{n\to\infty} \Phi_G(f_n) = G(s) - G(a).$$

Da $\Phi_F(f_n) = \Phi_G(f_n)$, so erhalten wir (6.22).

Für beliebige $F \in BV[a,b]$ gilt eine Darstellung

$$F = F_1 - F_2 \tag{6.23}$$

wobei F_1 und F_2 monoton steigend sind. Ist F rechtsstetig in (a,b) so erhalten wir für alle $x \in (a,b)$

$$\begin{aligned} F(x) &= \lim_{\varepsilon\to 0} F(x+\varepsilon) = \lim_{\varepsilon\to 0} F_1(x+\varepsilon) - \lim_{\varepsilon\to 0} F_2(x+\varepsilon) \\ &= (F_1)_r(x) - (F_2)_r(x). \end{aligned}$$

Somit können wir ohne Beschränkung der Allgemeinheit annehmen, dass F_1 und F_2 in (6.23) rechtsstetig sind. Analog gilt die Darstellung

$$G = G_1 - G_2,$$

wobei G_1 und G_2 auch monoton steigend und rechtsstetig sind. Weiter nehmen wir ohne Beschränkung der Allgemeinheit an, dass

$$F_1(a) = F_2(a) = G_1(a) = G_2(a) = 0.$$

Da $\Phi_F = \Phi_G$ so erhalten wir

$$\Phi_{F_1} - \Phi_{F_2} = \Phi_{F_1 - F_2} = \Phi_{G_1 - G_2} = \Phi_{G_1} - \Phi_{G_2}$$

und somit

$$\Phi_{F_1+G_2} = \Phi_{F_1} + \Phi_{G_2} = \Phi_{F_2} + \Phi_{G_1} = \Phi_{F_2+G_1}.$$

Da $F_1 + G_2$ und $F_2 + G_1$ monoton steigend und rechtsstetig sind, so erhalten wir nach dem ersten Teil des Beweises dass

$$F_1 + G_2 = F_2 + G_1$$

woraus folgt

$$F = F_1 - F_2 = G_1 - G_2 = G.$$

■

Betrachten wir jetzt einen Unterraum von $BV[a,b]$:

$$V[a,b] := \{F \in BV[a,b] : F(a) = 0 \text{ und } F \text{ ist rechtsseitig stetig in } (a,b)\}.$$

Dann ist $\mathrm{Var}_{[a,b]} F$ eine Norm in $V[a,b]$, so dass $V[a,b]$ ein normierter Vektorraum ist.

Hauptsatz 6.10 *Es gilt Isomorphismus von normierten Vektorräumen*

$$C[a,b]^* \cong V[a,b].$$

Folglich ist $V[a,b]$ ein Banachraum.

Beweis. Betrachten wir die lineare Abbildung

$$\begin{aligned} \mathcal{A} : V[a,b] &\to C[a,b]^* \\ \mathcal{A}(F) &= \Phi_F. \end{aligned}$$

Nach dem Lemma 6.9 ist $\mathcal{A}$ injektiv. Nach dem Satz 6.7, für jedes $\Phi \in C[a,b]^*$ gibt es eine Funktion $F \in BV[a,b]$ mit

$$\Phi = \Phi_F \quad \text{und} \quad \|\Phi\| = \operatorname{Var}_{[a,b]} F.$$

Nach dem Lemma 6.8 erhalten wir

$$\Phi = \Phi_{F_r} \text{ und } \operatorname{Var}_{[a,b]} F_r \le \operatorname{Var}_{[a,b]} F = \|\Phi\|. \tag{6.24}$$

Ohne Beschränkung der Allgemeinheit nehmen wir an, dass $F(a) = 0$. Dann liegt Φ_{F_r} in $V[a,b]$ und es gilt

$$\mathcal{A}(F_r) = \Phi,$$

so dass $\mathcal{A}$ surjektiv und somit auch bijektiv ist. Es bleibt zu zeigen, dass $\mathcal{A}$ eine Isometrie ist. In der Tat, nach (6.17) haben wir

$$\|\Phi\| = \|\Phi_{F_r}\| \le \operatorname{Var}_{[a,b]} F_r, \tag{6.25}$$

was zusammen mit (6.24) ergibt

$$\operatorname{Var}_{[a,b]} F_r = \|\Phi\|.$$

Somit ist $\mathcal{A}$ eine Isometrie. ■

6.6 * Satz von Picard-Lindelöf für lineare Differentialgleichungen

In diesem Abschnitt betrachten wir ein lineares Normalsystem von DGLen, das die folgende Form hat:

$$x' = A(t)\,x + B(t), \tag{6.26}$$

wobei $x = x(t)$ eine unbekannte Funktion mit den Werten in $\mathbb{R}^n$ ist und $A(t)$ und $B(t)$ gegebene Funktionen auf einem Intervall $I \subset \mathbb{R}$ mit Werten jeweils in $\mathbb{R}^{n\times n}$ und $\mathbb{R}^n$.

Insbesondere ist $A(t)\,x$ ein Vektor in $\mathbb{R}^n$, wie die anderen Terme in (6.26). Koordinateweise lautet (6.26) wie folgt:

$$x_i' = \sum_{l=1}^{n} A_{ij}(t)\,x_j + B_i(t), \quad i = 1, ..., n,$$

wobei A_{ij} und B_i die Komponenten jeweils von A und B sind. Wir nehmen immer an, dass die Abbildungen $A : I \to \mathbb{R}^{n\times n}$ und $B : I \to \mathbb{R}^n$ stetig auf I sind, also alle Komponenten $A_{ij}(t)$ und $B_i(t)$ stetige Funktionen von $t \in I$ sind. Wir betonen, dass der Definitionsbereich von (6.26) ist $I \times \mathbb{R}^n$, so dass jede Lösung von (6.26) muss auf einem Teilintervall von I definiert sein.

Hauptsatz 6.11 (Satz von Picard-Lindelöf für lineare Normalsystemen) *Seien $A(t)$ und $B(t)$ stetig auf I.*

(*a*) (Existenz) *Für alle $t_0 \in I$ und $x_0 \in \mathbb{R}^n$, existiert eine Lösung $x(t)$ des Anfangswertproblems*

$$\begin{cases} x' = A(t)x + B(t) \\ x(t_0) = x_0 \end{cases} \tag{6.27}$$

die auf dem Intervall I definiert ist.

(*b*) (Eindeutigkeit) *Sind $x(t)$ und $y(t)$ zwei Lösungen von (6.27) auf einem Teilintervall $I' \subset I$, so gilt es $x(t) \equiv y(t)$ auf I'.*

Für den Beweis von Satz 6.11 brauchen wir das folgende Lemma.

Lemma 6.12 (Grönwall-Lemma) *Sei $z(t)$ eine nicht-negative stetige Funktion auf einem Intervall $[a,b]$ und $t_0 \in [a,b]$. Gilt die folgende Ungleichung für alle $t \in [a,b]$:*

$$z(t) \le C + L\left|\int_{t_0}^{t} z(s)\,ds\right|, \tag{6.28}$$

mit beliebigen Konstanten $C, L \ge 0$, dann gilt für alle $t \in [a,b]$ die Ungleichung

$$z(t) \le Ce^{L|t-t_0|}. \tag{6.29}$$

Bemerkung. Normalerweise formuliert man Grönwall-Lemma für den Fall $t_0 = a$. In diesem Fall entfallen die Betragzeichen in (6.28) und (6.29), weil $t \ge t_0$.

Beweis. Es reicht die Behauptung im Fall $C > 0$ zu beweisen, da der Fall $C = 0$ daraus folgt indem man $C \to 0$ lässt. In der Tat, ist (6.28) mit $C = 0$ erfüllt, so ist (6.28) auch mit jedem $C > 0$ erfüllt. Es folgt, dass (6.29) mit jedem $C > 0$ erfüllt ist, und das ergibt (6.29) mit $C = 0$.

Also, nehmen wir an, dass $C > 0$ und definieren eine Funktion F auf dem Interval $[t_0, b]$ wie folgt

$$F(t) = C + L\int_{t_0}^{t} z(s)\,ds.$$

Bemerken wir, dass die Funktion F positive und differenzierbar ist, und $F' = Lz$ gilt. Die Bedingung (6.28) ergibt für $t \in [t_0, b]$ dass $z \le F$ und somit

$$F' = Lz \le LF.$$

Diese Differentialungleichung lässt sich wie folgt lösen. Dividieren durch F ergibt

$$\frac{F'}{F} \le L,$$

und durch Integration über $[t_0, t]$ erhalten wir, dass

$$\ln\frac{F(t)}{F(t_0)} = \int_{t_0}^{t} \frac{F'(s)}{F(s)}ds \le \int_{t_0}^{t} L ds = L(t - t_0),$$

für alle $t \in [t_0, b]$. Daraus folgt

$$F(t) \le F(t_0)e^{L(t-t_0)} = Ce^{L(t-t_0)}.$$

Da $z \leq F$, erhalten wir (6.29) für alle $t \in [t_0, b]$.

Auf dem Intervall $[a, t_0]$ betrachten wir analog die Funktion

$$F(t) = C + L \int_t^{t_0} z(s)\, ds,$$

die positive und differenzierbar ist. Da $F' = -Lz$ und nach (6.28) $z \leq F$, erhalten wir die Differentialungleichung

$$F' \geq -LF,$$

die ergibt für $t \in [a, t_0]$

$$\ln \frac{F(t_0)}{F(t)} = \int_t^{t_0} \frac{F'(s)}{F(s)} ds \geq -\int_t^{t_0} L ds = -L(t_0 - t) = -L|t - t_0|$$

und

$$z(t) \leq F(t) \leq F(t_0)\, e^{L|t-t_0|} = Ce^{L|t-t_0|}.$$

■

Beweis von Satz 6.11. Wählen wir ein beschränktes abgeschlossenes Interval $[a, b] \subset I$, derart, dass $t_0 \in [a, b]$. Wir beweisen die folgenden zwei Behauptungen:

1. die Eindeutigkeit der Lösung auf $[a, b]$, sowie auf jedem Teilintervall $I' \subset I$;

2. die Existenz einer Lösung auf $[a, b]$, sowie auf dem ganzen Intervall I.

Bemerken wir, dass für jede Lösung $x(t)$ von (6.27) auf $[a, b]$ und für alle $t \in [a, b]$ gilt

$$\begin{aligned} x(t) &= x_0 + \int_{t_0}^t x'(s)\, ds \\ &= x_0 + \int_{t_0}^t (A(s)\, x(s) + B(s))\, ds. \end{aligned} \tag{6.30}$$

Sei $x(t)$ und $y(t)$ zwei Lösungen von (6.27) auf the interval $[a, b]$, dann die beiden Funktionen erfüllen die Integralgleichung (6.30). Daraus folgt, dass

$$x(t) - y(t) = \int_{t_0}^t A(s)(x(s) - y(s))\, ds,$$

für alle $t \in [a, b]$. Sei $\|\cdot\|$ eine Norm in $\mathbb{R}^n$. Wir benutzen die Ungleichung

$$\left\| \int_{t_0}^t f(s)\, ds \right\| \leq \left| \int_{t_0}^t \|f(s)\|\, ds \right|, \tag{6.31}$$

die für jede stetige Funktion $f(s)$ mit Werten in $\mathbb{R}^n$ gilt. Mit Hilfe von (6.31) erhalten wir

$$\begin{aligned} \|x(t) - y(t)\| &\leq \left| \int_{t_0}^t \|A(s)(x(s) - y(s))\|\, ds \right| \\ &\leq \left| \int_{t_0}^t \|A(s)\| \|(x(s) - y(s))\|\, ds \right| \\ &\leq L \left| \int_{t_0}^t \|(x(s) - y(s))\|\, ds \right|, \end{aligned} \tag{6.32}$$

wobei
$$L = \sup_{s\in[a,b]} \|A(s)\|. \tag{6.33}$$
Da die Verkettung $s \mapsto A(s) \mapsto \|A(s)\|$ stetig ist, ist die Funktion $\|A(s)\|$ beschränkt auf dem Intervall $[a,b]$, so dass $L < \infty$. Nach (6.32) gilt für die Funktion
$$z(t) = \|x(t) - y(t)\|$$
die folgende Ungleichung:
$$z(t) \le L \left|\int_{t_0}^{t} z(s)\, ds\right|.$$
Lemma 6.12 mit $C = 0$ ergibt $z(t) \le 0$ und somit $z(t) = 0$, $x(t) \equiv y(t)$ auf $[a,b]$.

Sei $x(t), y(t)$ zwei Lösungen von (6.27) auf einem Intervall $I' \subset I$. Für jedes Intervall $[a,b] \subset I'$ mit $t_0 \in [a,b]$ haben wir $x(t) \equiv y(t)$ auf $[a,b]$. Da jedes Intervall als eine Vereinigung von beschränkten geschlossenen Intervallen dargestellt werden kann, gewinnen wir die Identität $x(t) \equiv y(t)$ auf I'.

Jetzt beweisen wir die Existenz einer Lösung von (6.27) auf $[a,b]$ mit Hilfe von Annäherung durch eine Funktionenfolge $\{x_k(t)\}_{k=0}^{\infty}$ auf $[a,b]$. Die Näherungslösungen werden induktiv definiert wie folgt:
$$x_0(t) \equiv x_0$$
und
$$x_k(t) = x_0 + \int_{t_0}^{t} \left(A(s)\, x_{k-1}(s) + B(s)\right) ds, \quad k \ge 1. \tag{6.34}$$
Es ist klar, dass alle Funktionen $x_k(t)$ stetig auf $[a,b]$ sind. Wir beweisen, dass die Folge $\{x_k\}_{k=0}^{\infty}$ auf $[a,b]$ gegen eine Lösung von (6.27) konvergiert für $k \to \infty$. Mit Hilfe von (6.34) und
$$x_{k-1}(t) = x_0 + \int_{t_0}^{t} \left(A(s)\, x_{k-2}(s) + B(s)\right) ds$$
erhalten wir, für jedes $k \ge 2$ und $t \in [a,b]$,
$$\begin{aligned} \|x_k(t) - x_{k-1}(t)\| &\le \left|\int_{t_0}^{t} \|A(s)\| \, \|x_{k-1}(s) - x_{k-2}(s)\|\, ds\right| \\ &\le L \left|\int_{t_0}^{t} \|x_{k-1}(s) - x_{k-2}(s)\|\, ds\right|, \end{aligned} \tag{6.35}$$
wobei L wie früher nach (6.33) definiert ist. Bezeichnen wir
$$z_k(t) = \|x_k(t) - x_{k-1}(t)\|,$$
und schreiben (6.35) in der folgenden Form um:
$$z_k(t) \le L \left|\int_{t_0}^{t} z_{k-1}(s)\, ds\right|. \tag{6.36}$$
Erst schätzen wir die Funktion $z_1(t) = \|x_1(t) - x_0(t)\|$ für $t \in [t_0, b]$ ab, wie folgt:
$$z_1(t) = \left\|\int_{t_0}^{t} \left(A(s)\, x_0 + B(s)\right) ds\right\| \le M(t - t_0),$$

wobei
$$M = \sup_{s\in[a,b]} \|A(s)\,x_0 + B(s)\| < \infty.$$
Es folgt aus (6.36), dass für $t \in [t_0, b]$
$$\begin{aligned}
z_2(t) &\le LM \int_{t_0}^{t} (s-t_0)\,ds = LM\frac{(t-t_0)^2}{2},\\
z_3(t) &\le L^2M \int_{t_0}^{t} \frac{(s-t_0)^2}{2}ds = L^2M\frac{(t-t_0)^3}{2\cdot 3},\\
&\dots\\
z_k(t) &\le L^{k-1}M\frac{(t-t_0)^k}{k!} \le \frac{(c\,(t-t_0))^k}{k!},
\end{aligned}$$
wobei $c = \max(L, M)$.

Mit dem gleichen Argument behandeln wir den Fall $t \in [a, t_0]$ und gewinnen die folgende Ungleichung
$$\|x_k(t) - x_{k-1}(t)\| \le \frac{(c\,|t-t_0|)^k}{k!},$$
für alle $t \in [a, b]$. Da die Potenzreihe
$$\sum_{k=0}^{\infty} \frac{(c\,|t-t_0|)^k}{k!}$$
konvergiert für alle reellen t, und zwar gleichmässig auf jedem beschränkten Intervall von t, auch die Funktionenreihe
$$\sum_{k=1}^{\infty} \|x_k(t) - x_{k-1}(t)\|$$
konvergiert gleichmässig auf $[a, b]$. Daraus folgt, dass die Funktionenreihe mit Vektor-Werten
$$\sum_{k=1}^{\infty} (x_k(t) - x_{k-1}(t))$$
konvergiert gleichmässig auf $[a, b]$. Da die Partialsummen dieser Reihe gleich $x_k(t) - x_0$ sind, gewinnen wir, dass die Funktionenfolge $\{x_k(t)\}$ gleichmäßig auf $[a, b]$ konvergiert. Bezeichnen wir
$$x(t) = \lim_{k\to\infty} x_k(t)\,.$$
Die Funktion $x(t)$ ist stetig auf $[a, b]$ als der Limes von einer gleichmäßig konvergierten Funktionenfolge von stetigen Funktionen. In der Identität (6.34), also
$$x_k(t) = x_0 + \int_{t_0}^{t} (A(s)\,x_{k-1}(s) + B(s))\,ds,$$
lassen wir k gegen unendlich gehen, und erhalten, dass der Grenzwert $x(t)$ die folgende Integralgleichung erfüllt:
$$x(t) = x_0 + \int_{t_0}^{t} (A(s)\,x(s) + B(s))\,ds \tag{6.37}$$

(das Integralzeichen und der Limes sind vertauschbar auf jedem kompakten Intervall). Wir behaupten, dass $x(t)$ das Anfangswertproblem (6.27) auf $[a,b]$ löst. Da die rechte Seite von (6.37) eine differenzierbare Funktion von t ist, so ist $x(t)$ auch differenzierbar, und

$$x' = \frac{d}{dt}\left(x_0 + \int_{t_0}^{t} (A(s)\,x(s) + B(s))\,ds.\right) = A(t)\,x(t) + B(t).$$

Schließlich, es ist klar von (6.37), dass $x(t_0) = x_0$. Deshalb löst $x(t)$ das Anfangswertproblem (6.27) auf $[a,b]$.

Jetzt definieren wir eine Lösung auf ganzem Intervall I. Es gibt eine wachsende Folge von beschränkten geschlossenen Intervallen $\{[a_i,b_i]\}_{i=1}^{\infty}$, derart, dass ihre Vereinigung gleich I ist; wir nehmen auch an, dass $t_0 \in [a_i,b_i]$ für alle i. Bezeichnen mit $x_i(t)$ eine Lösung von (6.27) auf $[a_i,b_i]$. Dann ist $x_{i+1}(t)$ auch eine Lösung von (6.27) auf $[a_i,b_i]$, und nach Eindeutigkeit des 1. Teils gewinnen wir, dass $x_{i+1}(t) = x_i(t)$ auf $[a_i,b_i]$; also, in der Folge $\{x_i(t)\}$ ist jede Funktion eine Fortsetzung der vorangehenden Funktion. Daraus folgt, dass the Funktion

$$x(t) := x_i(t) \text{ für } t \in [a_i,b_i]$$

wohldefiniert für alle $t \in I$ ist, und deshalb $x(t)$ eine Lösung von Anfangswertproblem (6.27) auf I ist. ■

Der Beweis der Existenz der Lösung führt zur folgenden Methode für Bestimmung der Lösung $x(t)$ des Anfangswertproblems (6.27). Man definiert eine Folge von *Näherungslösungen* $x_k(t)$ nach den Regeln

$$\begin{aligned} x_0(t) &\equiv x_0, \\ x_{k+1}(t) &= x_0 + \int_{t_0}^{t} (A(s)\,x_k(s) + B(s))\,ds. \end{aligned}$$

Diese Folge $\{x_k(t)\}$ heißt die *Picarditeration*. Nach dem Beweis von Satz 6.11 konvergiert die Folge $\{x_k(t)\}$ gegen die Lösung $x(t)$ für alle $t \in I$, und zwar gleichmäßig auf jedem kompakten Teilintervall $[a,b] \subset I$.

Wir verwenden den Satz 6.11 in der folgenden Situation. Betrachten wir eine lineare Differentialgleichung 2-er Ordnung:

$$u'' + a(t)\,u' + b(t)\,u = f(t) \tag{6.38}$$

wobei $a(t), b(t)$ und $f(t)$ gegebene stetige Funktionen auf einem Intervall I und u unbekannt ist. Zusammen mit (6.38) betrachten wir auch die Anfangsbedingungen für u:

$$u(t_0) = \alpha, \quad u'(t_0) = \beta \tag{6.39}$$

wobei $t_0 \in [a,b]$ und $\alpha, \beta \in R$ gegeben sind.

Korollar 6.13 *Das Anfangswertproblem (6.38)-(6.39) hat eine eindeutige Lösung $u(t) \in C^2(I)$ auf I.*

Beweis. Betrachten wir den Vektor-Funktion

$$\mathbf{x}(t) = \begin{pmatrix} \mathbf{x}_1(t) \\ \mathbf{x}_2(t) \end{pmatrix} := \begin{pmatrix} u(t) \\ u'(t) \end{pmatrix}.$$

Dann gilt

$$\begin{aligned}\mathbf{x}' &= \begin{pmatrix} u' \\ u'' \end{pmatrix} = \begin{pmatrix} u' \\ -au' - bu + f \end{pmatrix} = \begin{pmatrix} \mathbf{x}_2 \\ -b\mathbf{x}_1 - a\mathbf{x}_2 + f \end{pmatrix} \\ &= \begin{pmatrix} 0 & 1 \\ -b & -a \end{pmatrix} \begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{pmatrix} + \begin{pmatrix} 0 \\ f \end{pmatrix}\end{aligned}$$

d.h.

$$\mathbf{x}' = A(t)\,\mathbf{x} + B(t) \tag{6.40}$$

wobei

$$A(t) = \begin{pmatrix} 0 & 1 \\ -b(t) & -a(t) \end{pmatrix} \text{ und } B(t) = \begin{pmatrix} 0 \\ f(t) \end{pmatrix}.$$

Offensichtlich ist die Gleichung (6.38) äquivalent zu (6.40) wobei $\mathbf{x}(t)$ stetig differenzierbar sein soll. Die Anfangsbedingungen (6.39) sind äquivalent zu

$$\mathbf{x}(t_0) = \mathbf{x}_0 := \begin{pmatrix} \alpha \\ \beta \end{pmatrix}. \tag{6.41}$$

Nach dem Satz 6.11 hat die Gleichung (6.40) mit der Anfangsbedingung (6.41) eine eindeutige Lösung $\mathbf{x}(t)$ auf I, woraus die Existenz und Eindeutigkeit von u folgen. ■

Chapter 7

* Spektralsatz für selbstadjungierte Operatoren

7.1 Operator-Ungleichungen

Sei A ein selbstadjungierter Operator im Hilbertraum H über $\mathbb{C}$. Dann ist (Ax, x) reell für alle $x \in H$, da nach der Selbstadjungiertheit

$$(Ax, x) = (x, Ax)$$

und nach der Hermiteschen Symmetrie

$$(Ax, x) = \overline{(x, Ax)},$$

woraus folgt $(x, Ax) = \overline{(x, Ax)}$ und somit $(x, Ax) \in \mathbb{R}$.

Definition. Ein selbstadjungierter Operator A heißt *nichtnegativ definit* wenn $(Ax, x) \geq 0$ für alle $x \in H$. Man schreibt in diesem Fall $A \geq 0$.

Für zwei selbstadjungierte Operatoren A und B schreiben wir $A \geq B$ wenn $A - B \geq 0$, was äquivalent zu $(Ax, x) \geq (Bx, x)$ für alle $x \in H$.

Beispiel. Wir haben immer $A^2 \geq 0$, da

$$\left(A^2 x, x\right) = (Ax, Ax) \geq 0.$$

Beispiel. Jeder orthogonale Projektor P_U auf einen abgeschlossenen Unterraum $U \subset H$ ist nichtnegativ definit, da

$$(P_U x, x) = (P_U x, P_U x + P_{U^\perp} x) = \|P_U x\|^2 \geq 0.$$

Bemerken wir auch, dass die Operatoren P_U und $P_{U^\perp}$ nicht vergleichbar sind (vorausgesetzt, dass U und $U^\perp$ nicht trivial sind), d.h.

$$P_U \not\leq P_{U^\perp} \text{ und } P_{U^\perp} \not\leq P_U.$$

Beispiel. Der Greensche Operator K (Satz 3.20) ist nichtnegativ definit.

Im nächsten Lemma alle Operatoren sind selbstadjungiert im H.

Lemma 7.1 *Die folgende Eigenschaften von $\geq$ gelten.*

1. $A \geq B$ *und* $B \geq C \Rightarrow A \geq C$
2. $A \geq 0$ *und* $B \geq 0 \Rightarrow A + B \geq 0.$
3. $A \geq 0$ *und* $\alpha \in [0, \infty) \Rightarrow \alpha A \geq 0,$ *während für* $\alpha \in (-\infty, 0]$ *gilt* $\alpha A \leq 0.$
4. $A_n \geq 0$ *und* $A_n x \to Ax$ *für alle* $x \in H \Rightarrow A \geq 0.$
5. $-\|A\| I \leq A \leq \|A\| I$

Beweis. 1. $(Ax, x) \geq (Bx, x)$ und $(Bx, x) \geq (Cx, x)$ ergeben $(Ax, x) \geq (Cx, x)$.
2. $(Ax, x) \geq 0$ und $(Bx, x) \geq 0$ ergeben $((A + B)x, x) \geq 0.$
3. $(Ax, x) \geq 0$ ergibt $((\alpha A)x, x) \geq 0$ für $\alpha \geq 0$ und $((\alpha A)x, x) \leq 0$ für $\alpha \leq 0.$
4. $(A_n x, x) \geq 0$ und $A_n x \to Ax$ ergeben $(Ax, x) \geq 0.$
5. $(Ax, x) \leq \|Ax\| \|x\| \leq \|A\| \|x\|^2 = \|A\| (Ix, x)$ ergibt $A \leq \|A\| I$. Analog gilt $A \geq -\|A\| I$. ■

Satz 7.2 *Setzen wir*

$$a = \inf_{\|x\|=1} (Ax, x) \quad \text{und} \quad b = \sup_{\|x\|=1} (Ax, x).$$

Dann gelten die Identitäten

$$\min \sigma(A) = a \quad \text{und} \quad \max \sigma(A) = b. \tag{7.1}$$

Insbesondere gilt die Äquivalenz

$$A \geq 0 \quad \Leftrightarrow \quad \sigma(A) \geq 0. \tag{7.2}$$

Beweis. Für (7.1) siehe Übungen. Dann gilt

$$A \geq 0 \Leftrightarrow a \geq 0 \Leftrightarrow \sigma(A) \geq 0,$$

woraus (7.2) folgt. ■

Beispiel. Für eine symmetrische Matrix A bedeutet die Äquivalenz (7.2), dass A genau dann nichtnegativ definit ist, wenn alle Eigenwerte von A nichtnegativ sind.

Korollar 7.3 *Es gilt*

$$A \geq B \text{ und } B \geq A \Rightarrow A = B.$$

Beweis. Wir haben $\sigma(A - B) \geq 0$ und $\sigma(A - B) \leq 0$ woraus folgt $\sigma(A - B) = \{0\}$. Dann gilt $\|A - B\| = \sup |\sigma(A - B)| = 0$ und somit $A = B$. ■

Bemerkung. Somit ist $\geq$ eine partielle Ordnung auf der Menge von selbstadjungierten Operatoren.

Satz 7.4 *Für alle Funktion $f, g \in C(\sigma(A))$ gilt die Äquivalenz:*

$$f \geq g \text{ auf } \sigma(A) \quad \Leftrightarrow \quad f(A) \geq g(A).$$

Somit bewahrt das Funktionalkalkül $\mathcal{A} : C(\sigma(A)) \to \mathcal{B}(H)$ nicht nur die Struktur von normierten Algebras, sondern auch die partielle Ordnung.

Beweis. Wir haben nach den Sätzen 7.2 und 4.7

$$\begin{aligned} f(A) \geq g(A) &\Leftrightarrow f(A) - g(A) \geq 0 \Leftrightarrow \sigma(f(A) - g(A)) \geq 0 \\ &\Leftrightarrow \sigma((f-g)(A)) \geq 0 \Leftrightarrow (f-g)(\sigma(A)) \geq 0 \Leftrightarrow \ f \geq g \text{ auf } \sigma(A). \end{aligned}$$

■

7.2 Starke Konvergenz von Operatoren

Definition. Sei $\{A_n\}$ eine Folge von Operatoren im Hilbertraum. Man sagt, dass die Folge $\{A_n\}$ gegen den Operator A *stark konvergiert*, wenn

$$A_n x \to Ax \text{ für alle } x \in H.$$

Man schreibt in diesem Fall $A_n \xrightarrow{s} A$ oder $A = s\text{-}\lim A_n$.

Früher haben wir auch die Konvergenz $A_n \to A$ bezüglich der Operatornorm betrachten. Wir schreiben in diesem Fall $A = \lim A_n$. Die Konvergenz bezüglich der Operatornorm ist offensichtlich stärker als die starke Konvergenz, d.h.

$$A = \lim A_n \Rightarrow A = s\text{-}\lim A_n.$$

Es gibt auch die schwache Konvergenz: $A_n \xrightarrow{w} A$ oder $A = w\text{-}\lim A_n$ wenn

$$(A_n x, y) \to (Ax, y)$$

für alle $x, y \in H$. Die starke Konvergenz ist stärker als die schwache Konvergenz.

Die Konvergenz bezüglich der Operatornorm heißt auch *gleichmäßige* Konvergenz, die starke Konvergenz heißt auch *punktweise* Konvergenz.

Nach dem Lemma 7.1 bewahrt die starke Konvergenz die Operator-Ungleichungen zwischen selbstadjungierten Operatoren:

$$A_n \geq B_n \Rightarrow s\text{-}\lim A_n \geq s\text{-}\lim B_n.$$

Eine Folge $\{A_n\}$ von Operatoren heißt beschränkt wenn die Folge $\{\|A_n\|\}$ von den Normen beschränkt ist.

Satz 7.5 *Sei $\{A_n\}$ eine beschränkte monoton steigende (bzw fallende) Folge von selbstadjungierten Operatoren. Dann s-$\lim A_n$ existiert und ist ein selbstadjungierter Operator.*

Beweis. Wir benutzen die folgende Ungleichung, die für alle selbstadjungierte Operatoren B mit $B \geq 0$ gilt:

$$\|Bx\|^2 \leq \|B\| (Bx, x)$$

(siehe Aufgabe 80). Daraus folgt, dass für $n \geq m$ und alle $x \in H$

$$\|(A_n - A_m) x\|^2 \leq \|A_n - A_m\| ((A_n - A_m) x, x).$$

Die Folge $(A_n x, x)$ ist monotone steigend und beschränkt, somit ist sie eine Cauchy-Folge, d.h.

$$((A_n - A_m) x, x) \to 0 \text{ für } n, m \to \infty.$$

Da die Normen $\|A_n - A_m\|$ beschränkt sind, so erhalten wir, dass

$$\|(A_n - A_m) x\| \to 0 \text{ für } n, m \to \infty.$$

Dann ist die Folge $\{A_n x\}$ eine Cauchy-Folge und somit hat den Grenzwert $\lim A_n x$. Setzen wir für jedes x

$$Ax := \lim A_n x.$$

Dann ist A ein linearer Operator in H, und $A = s\text{-}\lim A_n$. Da

$$\|Ax\| = \lim \|A_n x\| \leq \sup_n \|A_n\| \|x\|,$$

so ist der Operator A beschränkt. Da

$$(Ax, y) = \lim (A_n x, y) = \lim (x, A_n y) = (x, Ay),$$

so ist A selbstadjungiert. ■

7.3 Monotone Grenzwerte der stetigen Funktionen und Funktionalkalkül

Satz 7.6 *Let A ein selbstadjungierter Operator und $\{f_n\}$ eine monoton fallende Folge von nichtnegativen stetigen Funktionen auf $\sigma(A)$. Dann gilt folgendes.*

1. $s\text{-}\lim f_n(A)$ *existiert (und ist selbstadjungiert).*

2. *Sei g_n eine andere solche Folge mit $\lim f_n \leq \lim g_n$ auf $\sigma(A)$. Dann gilt*

$$s\text{-}\lim f_n(A) \leq s\text{-}\lim g_n(A).$$

Insbesondere

$$\lim f_n = \lim g_n \ \Rightarrow s\text{-}\lim f_n(A) = s\text{-}\lim g_n(A).$$

Für den Beweis von zweitem Teil brauchen wir ein Lemma.

Lemma 7.7 (Satz von Dini) *Sei $\{h_n\}$ eine monoton fallende Folge von nichtnegativen stetigen Funktionen auf einem kompakten metrischen Raum X. Gilt $h_n(x) \to 0$ für alle $x \in X$, so gilt auch*

$$\sup h_n \to 0 \quad \textit{für } n \to \infty.$$

Beweis. Für ein vorgegebenes $\varepsilon > 0$ setzen wir

$$U_n = \{x \in X : h_n(x) < \varepsilon\}.$$

Da $h_n(x) \to 0$ für $n \to \infty$, liegt jedes x in einem U_n. Da die Mengen U_n offen sind, erhalten wir eine offene Überdeckung $\{U_n\}$ von X. Sei $U_{n_1}, ...U_{n_k}$ eine endliche Teilüberdeckung, und setzen wir $N = \max(n_1, ..., n_k)$. Da jedes $x \in X$ in einem von U_{n_i} liegt, erhalten wir für alle $n \geq N$

$$h_n(x) \leq h_{n_i}(x) < \varepsilon,$$

woraus folgt $\sup_{x \in K} |h_n| < \varepsilon$, was zu beweisen war. ■

Beweis von dem Satz 7.6. 1. Die Bedingung $f_{n+1} \leq f_n$ auf $\sigma(A)$ impliziert $f_{n+1}(A) \leq f_n(A)$, so dass die Folge $\{f_n(A)\}$ monoton fallend ist. Da

$$\|f_n(A)\| = \sup_{\sigma(A)} f_n \leq \sup_{\sigma(A)} f_1,$$

so ist die Folge $\{\|f_n(A)\|\}$ beschränkt. Dann s-$\lim f_n(A)$ existiert und ist selbstadjungiert nach dem Satz 7.5.

2. Wir werden beweisen, dass für jedes k und jedes $\varepsilon > 0$ existiert ein N so dass für alle $n \geq N$ gilt

$$\sup_{\sigma(A)} (f_n - g_k) \leq \varepsilon. \tag{7.3}$$

Nach dem Satz 7.4 folgt es aus (7.3), dass

$$f_n(A) - g_k(A) \leq \varepsilon I.$$

Für $n \to \infty$, $k \to \infty$ und $\varepsilon \to 0$ erhalten wir nach Lemma 7.1

$$s\text{-}\lim f_n(A) - s\text{-}\lim g_k(A) \leq 0,$$

was zu beweisen war.

Um (7.3) zu beweisen, setzen wir

$$h_n = (f_n - g_k)_+.$$

Dann ist $\{h_k\}$ eine monoton fallende Folge von stetigen Funktionen auf X. Da

$$\lim f_n \leq \lim g_n \leq g_k,$$

erhalten wir, dass $\lim h_n = 0$. Nach Lemma 7.7 beschließen wir, dass $\sup h_n \to 0$, woraus (7.3) folgt. ■

Definition. Bezeichnen wir mit M_+ die Menge von reellwertigen Funktionen f auf $\mathbb{R}$ mit der folgenden Eigenschaft: es gibt eine monoton fallende Folge $\{f_n\}$ von nichtnegativen stetigen Funktionen auf $\mathbb{R}$ mit $f_n(z) \to f(z)$ $\forall z \in \mathbb{R}$. Bezeichnen wir mit M die Menge von Differenzen $f - g$ mit $f, g \in M_+$.

Offensichtlich ist jede Funktion $f \in M$ beschränkt auf jeder kompakten Teilmenge von $\mathbb{R}$. Es folgt aus der Definition, dass $C_+(\mathbb{R}) \subset M_+$ und $C(\mathbb{R}) \subset M$, aber M enthält auch unstetige Funktionen. Z.B. für jedes $a \in \mathbb{R}$ liegt die Funktion $\mathbf{1}_{(-\infty,a]}$ in M_+, und für alle $a < b$ liegt die Funktion

$$\mathbf{1}_{(a,b]} = \mathbf{1}_{(-\infty,b]} - \mathbf{1}_{(-\infty,a]}$$

in M.

Definition. Sei A ein selbstadjungierter Operator. Für jede Funktion $f \in M_+$ definieren wir den Operator $f(A)$ mit

$$f(A) = s\text{-}\lim f_n(A),$$

wobei f_n eine monoton fallende Folge $\{f_n\}$ von nichtnegativen stetigen Funktionen auf $\mathbb{R}$ mit $f_n(z) \to f(z)\ \forall z \in \mathbb{R}$.

Nach dem Satz 7.6 der Grenzwert s-lim $f_n(A)$ existiert, ist selbstadjungiert und unabhängig von der Wahl der Folge $\{f_n\}$.

Satz 7.8 *Für alle Funktionen $f, g \in M_+$ gelten die folgenden Eigenschaften.*

1. *$f \leq g$ auf $\sigma(A)$ impliziert $f(A) \leq g(A)$.*
2. *$\|f(A)\| \leq \sup_{\sigma(A)} f$*
3. *$f + g$, cf mit $c > 0$ und fg liegen in M_+ und*

$$\begin{aligned}(f+g)(A) &= f(A) + g(A)\\ (cf)(A) &= c(f(A))\\ (fg)(A) &= f(A)g(A).\end{aligned}$$

Beweis. Seien $\{f_n\}$ und $\{g_n\}$ zwei monoton fallende Folgen von nichtnegativen stetigen Funktionen auf $\mathbb{R}$, die gegen f bzw g punktweise konvergieren.

1. Dies folgt aus dem Satz 7.6, da

$$f = \lim f_n \leq \lim g_n = g \quad \text{auf } \sigma(A)$$

impliziert $f(A) \leq g(A)$.

2. Da $0 \leq f \leq \sup f$, so erhalten wir die Operator-Ungleichungen

$$0 \leq f(A) \leq (\sup f)\, I.$$

Daraus folgt

$$0 \leq \sigma(f(A)) \leq \sup f$$

and

$$\|f(A)\| = \sup |\sigma(f(A))| \leq \sup f.$$

3. Die Folge $\{f_n g_n\}$ ist monoton fallend, und $f_n g_n \to fg$ punktweise, so dass $fg \in M_+$. Wir haben nach der Definition von $(fg)(A)$

$$(f_n g_n)(A) \to (fg)(A).$$

Andererseits gilt

$$(f_n g_n)(A) = f_n(A)\, g_n(A) \xrightarrow{s} f(A)g(A),$$

da $f_n(A) \xrightarrow{s} f(A)$ und $g_n(A) \xrightarrow{s} g(A)$ (siehe Aufgaben), woraus folgt $(fg)(A) = f(A)g(A)$.

Die Aussagen über cf und $f + g$ werden analog bewiesen. ■

Für jede Funktion $f \in M$ existieren Funktionen $g, h \in M_+$ mit $f = g - h$. Dann setzen wir

$$f(A) = g(A) - h(A).$$

Das nächste Lemma impliziert, dass diese Definition unabhängig von der Wahl der Funktionen g, h ist.

Lemma 7.9 *Seien $g_1, g_2, h_1, h_2 \in M_+$. Dann gilt*

$$g_1 - h_1 \leq g_2 - h_2 \Rightarrow g_1(A) - h_1(A) \leq g_2(A) - h_2(A).$$

Insbesondere

$$g_1 - h_1 = g_2 - h_2 \Rightarrow g_1(A) - h_1(A) = g_2(A) - h_2(A)$$

Beweis. Wir haben

$$g_1 + h_2 \leq g_2 + h_1$$

woraus folgt nach Satz 7.8

$$g_1(A) + h_2(A) \leq g_2(A) + h_1(A),$$

und somit

$$g_1(A) - h_1(A) \leq g_2(A) - h_2(A).$$

■

Der folgende Satz erweitert den Satz 7.8 für Funktionen aus M.

Satz 7.10 *Für alle Funktionen $f, g \in M$ gelten die folgenden Eigenschaften.*

1. *$f \leq g$ auf $\sigma(A)$ impliziert $f(A) \leq g(A)$.*
2. *$\|f(A)\| \leq \sup_{\sigma(A)} |f|$.*
3. *$f + g$, cf mit $c \in \mathbb{R}$ und fg liegen in M und*

$$\begin{aligned}(f+g)(A) &= f(A) + g(A)\\ (cf)(A) &= c(f(A))\\ (fg)(A) &= f(A)g(A).\end{aligned}$$

Beweis. 1. Dies folgt aus Lemma 7.9.

2. Setzen wir $c = \sup |f|$. Dann gilt $-c \leq f \leq c$ und somit

$$-cI \leq f(A) \leq cI,$$

woraus folgt $\sigma(f(A)) \subset [-c, c]$ und $\|f(A)\| \leq c$.

3. Sei $f = f_1 - f_2$ und $g = g_1 - g_2$ mit $f_1, f_2, g_1, g_2 \in M_+$. Dann gilt

$$fg = (f_1 - f_2)(g_1 - g_2) = (f_1 g_1 + f_2 g_2) - (f_1 g_2 + f_2 g_1)$$

woraus folgt $fg \in M$ und

$$(fg)(A) = (f_1 g_1 + f_2 g_2)(A) - (f_1 g_2 + f_2 g_1)(A).$$

Dann erhalten wir

$$\begin{aligned}f(A)g(A) &= (f_1(A) - f_2(A))(g_1(A) - g_2(A))\\ &= (f_1(A)g_1(A) + f_2(A)g_2(A)) - (f_1(A)g_2(A) + f_2(A)g_1(A))\\ &= (f_1 g_1 + f_2 g_2)(A) - (f_1 g_2 + f_2 g_1)(A)\\ &= (fg)(A).\end{aligned}$$

Die Aussagen über cf und $f + g$ werden analog bewiesen. ■

7.4 Spektralschar

Gegeben sei ein selbstadjungierter Operator A in H. Da für jedes $\lambda \in \mathbb{R}$ die Funktion $\mathbf{1}_{(-\infty,\lambda]}$ in M_+ liegt, so können wir setzen

$$E_\lambda = \mathbf{1}_{(-\infty,\lambda]}(A).$$

Dann ist E_λ ein selbstadjungiert Operator.

Definition. Die Familie $\{E_\lambda\}_{\lambda\in\mathbb{R}}$ von Operatoren heißt die *Spektralschar* von A.

Im nächsten Satz werden die Haupteigenschaften von der Spektralschar bewiesen.

Satz 7.11 *Die Spektralschar von einem selbstadjungierten Operator A hat die folgenden Eigenschaften.*

1. *E_λ ist ein orthogonaler Projektor in H.*
2. *Die Spektralschar $\{E_\lambda\}_{\lambda\in\mathbb{R}}$ ist monoton steigend bezüglich λ, d.h. $E_\lambda \le E_\mu$ für $\lambda \le \mu$.*
3. *$E_\lambda = 0$ für $\lambda < \inf\sigma(A)$ und $E_\lambda = I$ für $\lambda \ge \sup\sigma(A)$.*
4. *Die Funktion $\lambda \to E_\lambda$ ist rechtsseitig stetig bezüglich der starken Konvergenz von Operatoren, d.h.*
$$E_{\lambda+h} \xrightarrow{s} E_\lambda \;\; \text{für } h \to 0+.$$

Jede Familie $\{E_\lambda\}_{\lambda\in\mathbb{R}}$ mit diesen Eigenschaften heißt eine *Spektralschar*.

Beweis. 1. Die Funktion $f = \mathbf{1}_U$ wobei U eine Teilmenge von $\mathbb{R}$ ist, erfüllt $f^2 = f$. Daraus folgt, dass $E_\lambda^2 = E_\lambda$. Da E_λ selbstadjungiert ist, so beschließen wir, dass E_λ ein orthogonaler Projektor ist.

2. Da $\mathbf{1}_{(-\infty,\lambda]} \le \mathbf{1}_{(-\infty,\mu]}$, so erhalten wir $E_\lambda \le E_\mu$.

3. Für $\lambda < \inf\sigma(A)$ gilt $\mathbf{1}_{(-\infty,\lambda]} = 0$ auf $\sigma(A)$, woraus $E_\lambda = 0$. Für $\lambda \ge \sup\sigma(A)$ gilt $\mathbf{1}_{(-\infty,\lambda]} = 1$ auf $\sigma(A)$, woraus folgt $E_\lambda = 1(A) = I$.

4. Wir beweisen, dass für jede monoton fallende Folge $\{h_n\}_{n=1}^\infty$ mit $h_n \to 0+$ gilt

$$E_{\lambda+h_n} \xrightarrow{s} E_\lambda \text{ für } n \to \infty.$$

Für jedes n definieren wir eine stetige Funktion f_n auf $\mathbb{R}$ wie folgt:

$$f_n(z) = \begin{cases} 1, & z \le \lambda + h_n \\ 0, & z \ge \lambda + 2h_n \\ \text{linear}, & \lambda + h_n \le z \le \lambda + 2h_n. \end{cases}$$

Da

$$\mathbf{1}_{(-\infty,\lambda]} \le \mathbf{1}_{(-\infty,\lambda+h_n]} \le f_n,$$

so erhalten wir

$$E_\lambda \le E_{\lambda+h_n} \le f_n(A). \tag{7.4}$$

Da die Folge $\{f_n\}$ monoton fallend ist und $f_n \to \mathbf{1}_{(-\infty,\lambda]}$, erhalten wir

$$s\text{-}\lim f_n(A) = E_\lambda.$$

Da die Folge $\{E_{\lambda+h_n}\}$ von Operatoren monoton fallend und beschränkt ist, existiert nach dem Satz 7.5 der Grenzwert s-lim $E_{\lambda+h_n}$. Es folgt aus (7.4) und Lemma 7.1, dass

$$E_\lambda \leq s\text{-}\lim E_{\lambda+h_n} \leq E_\lambda$$

woraus s-lim $E_{\lambda+h_n} = E_\lambda$ folgt. ■

Beispiel. Sei $\{K_n\}_{n=1}^\infty$ eine Folge von zueinander orthogonalen abgeschlossenen Unterräumen $K_n \neq \{0\}$ vom Hilbertraum H mit $\bigoplus_{n=1}^\infty K_n = H$. Sei $\{\lambda_n\}_{n=1}^\infty$ eine beschränkte Folge von reellen Zahlen. Dann konvergiert die Reihe $\sum_{n=1}^\infty \lambda_n P_{K_n}$ stark, und wir setzen

$$A = s\text{-}\sum_{n=1}^\infty \lambda_n P_{K_n},$$

so dass A ein selbstadjungierter Operator ist. Für diesen Operator erhalten wir

$$E_\lambda = \mathbf{1}_{(-\infty,\lambda]}(A) = s\text{-}\sum_{n=1}^\infty \mathbf{1}_{(-\infty,\lambda]}(\lambda_n) P_{K_n} = s\text{-}\sum_{\{n:\lambda_n\leq\lambda\}} P_{K_n},$$

woraus folgt $E_\lambda = P_{U_\lambda}$ wobei

$$U_\lambda = \bigoplus_{\{n:\lambda_n\leq\lambda\}} K_n.$$

7.5 Riemann-Stieltjes-Integral

Sei f eine stetige Funktion auf einem Intervall $[a,b]$ und m eine monoton wachsende Funktion auf $[a,b]$. Definieren wir das Riemann-Stieltjes-Integral

$$\int_a^b f(t)\, dm(t)$$

wie folgt. Für jede Zerlegung

$$Z = \{a = \lambda_0 < \lambda_1 < ... < \lambda_n = b\}$$

von $[a,b]$ mit den Zwischenstellen $T = \{t_k\}_{k=1}^n$, wobei $t_k \in [\lambda_{k-1}, \lambda_k]$, definieren wir die Riemann-Stieltjes-Summe

$$S(f,m,Z,T) = \sum_{k=1}^n f(t_k)(m(\lambda_k) - m(\lambda_{k-1})).$$

Z.B. für $m(\lambda) = \lambda$ stimmt $S(f,m,Z,T)$ mit der Riemann-Summe

$$S(f,Z,T) = \sum_{k=1}^n f(t_k)(\lambda_k - \lambda_{k-1})$$

für das Riemann-Integral

$$\int_a^b f(t)\, dt$$

überein.

Setzen wir

$$\delta_Z = \max_{1\leq k\leq n} \left(\lambda_k - \lambda_{k-1}\right).$$

Man kann beweisen, dass der Grenzwert

$$\lim_{\delta_Z\to 0} S\left(f, m, Z, T\right)$$

existiert. Dann definiert man das Riemann-Stieltjes-Integral mit

$$\int_a^b f\left(t\right) dm\left(t\right) = \lim_{\delta_Z\to 0} S\left(f, m, Z, T\right).$$

Beispiel. Ist m stetig differenzierbar, so erhält man die Identität

$$\int_a^b f\left(t\right) dm\left(t\right) = \int_a^b f\left(t\right) m'\left(t\right) dt.$$

Beispiel. Sei m eine Treppenfunktion, d.h. es gibt eine Zerlegung $\{a = t_0 < ... < t_n = b\}$ mit

$$m\left(t\right) = m_k \quad \text{für } t \in \left(t_k, t_{k+1}\right).$$

Dann gilt

$$\int_a^b f\left(t\right) dm\left(t\right) = \sum_{k=1}^n f\left(t_k\right)\left(m_k - m_{k-1}\right).$$

Sei A ein selbstadjungierter Operator in H und sei E_λ seine Spektralschar. Wählen wir ein Intervall $[a, b]$ das $\sigma\left(A\right)$ überdeckt. Für jede Funktion $f \in C\left[a, b\right]$ definieren wir das Riemann-Stieltjes-Integral bezüglich E_λ wie folgt. Für jede Zerlegung

$$Z = \{a = \lambda_0 < \lambda_1 < ... < \lambda_n = b\}$$

von $[a, b]$ mit den Zwischenstellen $T = \{t_k\}_{k=1}^n$, definieren wir die Riemann-Stieltjes-Summe

$$S\left(f, E, Z, T\right) = \sum_{k=1}^n f\left(t_k\right)\left(E_{\lambda_k} - E_{\lambda_{k-1}}\right), \tag{7.5}$$

die ein selbstadjungierter Operator ist. Dann definiert man

$$\int_a^b f\left(\lambda\right) dE_\lambda = \lim_{\delta_Z\to 0} S\left(f, E, Z, T\right),$$

vorausgesetzt, dass der Grenzwert existiert (bezüglich der gleichmäßigen oder starken Konvergenz von Operatoren).

7.6 Spektralsatz

Hauptsatz 7.12 (Spektralsatz) *Für jeden selbstadjungierten Operator A und jede stetige Funktion f auf $[a,b] \supset \sigma(A)$ gilt*

$$f(A) = \int_a^b f(\lambda)\, dE_\lambda, \tag{7.6}$$

wobei die Riemann-Stieltjes-Summen gegen $f(A)$ in der Operatornorm konvergieren.

Darüber hinaus gilt für jedes $x \in H$

$$\|f(A)x\|^2 = \int_a^b f(\lambda)^2\, d\,\|E_\lambda x\|^2, \tag{7.7}$$

wobei das Integral in (7.7) ein Riemann-Stieltjes-Integral bezüglich der monoton steigenden Funktion $\lambda \mapsto \|E_\lambda x\|^2$ ist.

Insbesondere erhalten wir für $f(z) = z$

$$A = \int_a^b \lambda dE_\lambda. \tag{7.8}$$

Die Identität (7.8) heißt die *spektrale Zerlegung* von A. Sie ist eine Verallgemeinerung der Identität (3.33) von dem Hilbert-Schmidt-Satz (Hauptsatz 3.15).

Beweis. Betrachten wir die Funktion

$$f_{Z,T}(\lambda) = \sum_{k=1}^n f(t_k)\, \mathbf{1}_{(\lambda_{k-1},\lambda_k]}(\lambda),$$

d.h. $f_{Z,T}(\lambda) = f(t_k)$ für $\lambda \in (\lambda_{k-1}, \lambda_k]$. Nach der gleichmäßigen Stetigkeit von f erhalten wir

$$\sup_{[a,b]} |f_{Z,T} - f| \to 0 \quad \text{für} \quad \delta_Z \to 0.$$

Da $f_{Z,T} \in M$, so ist $f_{Z,T}(A)$ wohldefiniert, und

$$f_{Z,T}(A) = \sum_{k=1}^n f(t_k)\, \mathbf{1}_{(\lambda_{k-1},\lambda_k]}(A) = \sum_{k=1}^n f(t_k) \left(E_{\lambda_k} - E_{\lambda_{k-1}}\right) = S(f,E,Z,T). \tag{7.9}$$

Somit gilt für $\delta_Z \to 0$

$$\|S(f,E,Z,T) - f(A)\| = \|f_{Z,T}(A) - f(A)\| \le \sup_{\sigma(A)} |f_{Z,T} - f| \to 0,$$

woraus (7.6) folgt.

Da für $k \neq m$ gilt

$$\mathbf{1}_{(\lambda_{k-1},\lambda_k]} \mathbf{1}_{(\lambda_{m-1},\lambda_m]} = 0,$$

es folgt, dass

$$\left(E_{\lambda_k} - E_{\lambda_{k-1}}\right)\left(E_{\lambda_m} - E_{\lambda_{m-1}}\right) = 0$$

und somit für jedes $x \in H$

$$\left(\left(E_{\lambda_k} - E_{\lambda_{k-1}}\right) x, \left(E_{\lambda_m} - E_{\lambda_{m-1}}\right) x\right) = 0.$$

Es folgt aus (7.9), dass

$$\left\| f_{Z,T}(A)(x) \right\|^2 = \sum_{k=1}^{n} f(t_k)^2 \left\| \left(E_{\lambda_k} - E_{\lambda_{k-1}}\right) x \right\|^2 . \tag{7.10}$$

Beweisen wir, dass für alle $\mu < \lambda$ gilt

$$\left\| (E_\lambda - E_\mu) x \right\|^2 = \left\| E_\lambda x \right\|^2 - \left\| E_\mu x \right\|^2 . \tag{7.11}$$

Da

$$\mathbf{1}_{(-\infty,\ \mu]} \mathbf{1}_{(\ \mu,\lambda]} = 0,$$

es folgt, dass

$$(E_\mu x, (E_\lambda - E_\mu) x) = 0,$$

und nach dem Satz von Pythagoras

$$\left\| E_\lambda x \right\|^2 = \left\| E_\mu x \right\|^2 + \left\| (E_\lambda - E_\mu) x \right\|^2 ,$$

woraus (7.11) folgt. Offensichtlich impliziert (7.11), dass die Funktion $\lambda \mapsto \left\| E_\lambda x \right\|^2$ monoton steigend ist.

Einsetzen von (7.11) in (7.10) ergibt

$$\begin{aligned}
\left\| f_{Z,T}(A)(x) \right\|^2 &= \sum_{k=1}^{n} f(t_k)^2 \left(\left\| E_{\lambda_k} x \right\|^2 - \left\| E_{\lambda_{k-1}} x \right\|^2 \right) \\
&= S\left(f, \left\| E_\lambda x \right\|^2, Z, T\right)
\end{aligned}$$

Für $\delta_T \to 0$ erhalten wir

$$\begin{aligned}
\left\| f(A) x \right\|^2 &= \lim_{\delta_Z \to 0} \left\| f_{Z,T}(A)(x) \right\|^2 \\
&= \lim_{\delta_Z \to 0} S\left(f, \left\| E_\lambda x \right\|^2, Z, T\right) \\
&= \int_a^b f(\lambda)^2 \, d \left\| E_\lambda x \right\|^2 .
\end{aligned}$$

■

Bemerkung. Für jedes Intervall $I = (a, b]$ kann man $\mathbf{1}_I(A) = E_b - E_a$ als ein operatorwertiges Maß von I betrachten. Genau so, wie das Lebesgue-Maß von Intervallen nach borelschen Mengen fortgesetzt wird, kann auch das operatorwertiges Maß auf alle borelsche Mengen fortgesetzt werden. Bezeichnen wir dieses Maß mit $E(U)$ für jede borelsche Menge $U \subset \mathbb{R}$. Weiter definiert man das Lebesgue-Integral bezüglich des Maßes E von jeder borelschen Funktion f auch dem Spektrum (die Lebesgue-Summen konvergieren schwach gegen das Integral). Dann setzt man für alle borelsche Funktionen f auf $\sigma(A)$

$$f(A) := \int_{\sigma(A)} f dE.$$

Insbesondere gilt diese Definition für unbeschränkte Funktionen f. In diesem Fall ist $f(A)$ ein *unbeschränkter* Operator, deren Definitionsbereich ein dicht liegender Unterraum von H ist. Die Theorie von unbeschränkten selbstadjungierten Operatoren gehört zu dieser Vorlesung nicht, aber die Hauptergebnisse davon sind ähnlich wie für beschränkten Operatoren.

7.7 Multiplikationsoperatoren

Sei (X, S, μ) ein Maßraum. Betrachten wir den Hilbertraum $L^2 = L^2(X, \mu)$ über $\mathbb{C}$. Jede Funktion $a \in L^\infty(X)$ bestimmt einen Operator T_a in L^2 mit

$$T_a f(x) = a(x) f(x).$$

Offensichtlich ist T_a beschränkt da

$$\|T_a f\|_{L^2} \leq \|a\|_{L^\infty} \|f\|_{L^2}.$$

Der Operator T_a heißt *Multiplikationsoperator.* Ist a reellwertig so ist T_a selbstadjungiert da

$$(T_a f, g) = \int_X a f \overline{g} d\mu = \int_X f \overline{a g} d\mu = (f, T_a g).$$

Wir beweisen here dass jeder beschränkte selbstadjungierte Operator A eine Darstellung als Multiplikationsoperator hat.

Gegeben seien zwei Hilberträume H_1 und H_2. Eine Abbildung

$$U : H_1 \to H_2$$

heißt *unitär* wenn U linear ist und das Skalarprodukt unter U behalten wird, d.h. für alle $x, y \in H_1$

$$(Ux, Ux)_{H_2} = (x, y)_{H_1}.$$

Insbesondere gilt

$$\|Ux\|_{H_2} = \|x\|_{H_1}.$$

Offensichtlich ist jede unitäre Abbildung U injektiv. Ist U auch bijektiv, so ist U eine lineare Isometrie von H_1 und H_2. Zwei Hilberträume die linear isometrisch sind haben gleiche Eigenschaften.

Der folgende Satz ist eine Version des Spektralsatzes für selbstadjungierte Operatoren.

Hauptsatz 7.13 *Sei H ein separabler Hilbertraum über $\mathbb{C}$ und A ein beschränkter selbstadjungierter Operator in H. Dann existiert ein Maßraum (X, S, μ), eine lineare Isometrie $U : L^2(X, \mu) \to H$ und eine reellwertige Funktion $a \in L^\infty(X)$ so dass*

$$A = U T_a U^{-1}. \tag{7.12}$$

Die Identität (7.12) ist äquivalent zu

$$AU = UT_a,$$

d.h. das folgende Diagram ist kommutativ:

$$\begin{array}{ccc} L^2 & \xrightarrow{U} & H \\ {\scriptstyle T_a}\downarrow & & \downarrow{\scriptstyle A} \\ L^2 & \xrightarrow{U} & H \end{array}$$

Beweis. Das Spektrum $\sigma(A)$ liegt im Intervall $J = [-\|A\|, \|A\|]$. Fixieren wir ein $\xi \in H$ und betrachten die Abbildung

$$\begin{aligned} \Phi_\xi &: C(J) \to \mathbb{R} \\ \Phi_\xi(f) &= (f(A)\xi, \xi) \end{aligned}$$

wobei $C(J)$ der Raum von reellwertigen stetigen Funktionen auf J ist. Da der Operator $f(A)$ selbstadjungiert ist, so ist die Zahl $(f(A)\xi, \xi)$ reell.

Offensichtlich ist Φ_ξ ein lineares Funktional auf $C(J)$. Das Funktional Φ_ξ ist beschränkt bezüglich der sup-Norm da

$$|(f(A)\xi, \xi)| \le \|f(A)\| \|\xi\|^2 \le \|\xi\|^2 \sup_J |f|.$$

Das Funktional Φ_ξ ist positiv im folgenden Sinn: if die Funktion $f \in C(J)$ nichtnegativ so gilt $\Phi_\xi(f) \ge 0$. In der Tat, für die Funktion $g = \sqrt{f} \in C(J)$ erhalten wir

$$\Phi_\xi(f) = (f(A)\xi, \xi) = (g(A)g(A)\xi, \xi) = (g(A)\xi, g(A)\xi) = \|g(A)\xi\|^2 \ge 0.$$

Es folgt aus dem Darstellungssatz von Riesz (Satz 6.7): jedes beschränktes positives Funktional Φ auf $C(J)$ hat der Form

$$\Phi(f) = \int_J f d\mu$$

für ein Borel-Maß μ auf J. Somit erhalten wir für jedes $\xi \in H$ ein Borel-Maß μ_ξ auf J so dass für alle $f \in C(J)$ gilt

$$(f(A)\xi, \xi) = \Phi_\xi(f) = \int_J f d\mu_\xi.$$

Betrachten wir jetzt die Abbildung

$$\begin{aligned} U_\xi &: C(J) \to H \\ U_\xi f &= f(A)\xi. \end{aligned}$$

Zeigen wir, dass unter dieser Abbildung der L^2-Norm behalten wird: für alle $f, g \in C(J)$ gilt

$$(U_\xi f, U_\xi g)_H = (f(A)\xi, g(A)\xi) = ((fg)(A)\xi, \xi) = \int_H fg d\mu_\xi = (f, g)_{L^2}.$$

Insbesondere ist die Abbildung U_ξ beschränkt bezüglich der L^2-Norm in $C(J)$ und der Norm in H. Da $C(J)$ in $L^2(J, \mu_\xi)$ dicht liegt, so lässt U_ξ sich auf $L^2(J, \mu_\xi)$ eindeutig erweitern. Somit erhalten wir einen unitären Operator

$$U_\xi : L^2(J, \mu_\xi) \to H.$$

Bemerken wir, dass

$$AU_\xi f = Af(A)\xi = f(A)A\xi = U_\xi g$$

wobei

$$g(t) = f(t)t = T_a f(t)$$

wobei $a(t) = t$. Somit erhalten wir die Identität

$$AU_\xi = U_\xi T_a. \tag{7.13}$$

Das einziges Problem dass noch bleibt ist ob die Abbildung U_ξ surjektiv ist. Ist es der Fall, so ist U_ξ eine lineare Isometrie, und alles bewiesen ist.

In allgemeinem Fall ist U_ξ nicht unbedingt surjektiv. Bezeichnen wir mit H_ξ das Bild von U_ξ so dass H_ξ ein abgeschlossener Unterraum von H ist. Dann ist U_ξ eine lineare Isometrie zwischen $L^2\left(J, \mu_\xi\right)$ und H_ξ, die (7.13) erfüllt.

Wir zeigen unterhalb dass H sich in der Form von direkter Summe

$$H = \bigoplus_{k=1}^{N} H_k \tag{7.14}$$

darstellen lässt, wobei $H_k = H_{\xi_k}$ und $N \in \mathbb{N} \cup \{\infty\}$. Für jedes ξ_k erhalten wir ein Maß $\mu_k = \mu_{\xi_k}$ auf $J_k = J$ und ein lineare Isometrie

$$U_k = U_{\xi_k} : L^2(J_k, \mu_k) \to H_k$$

so dass

$$AU_k = U_k T_{a_k}$$

für die Funktion $a_k(t) = t$ auf J_k. Dann betrachten wir die Menge

$$X = \bigsqcup_{k=1}^{N} J_k$$

und definieren wir ein Maß μ auf X mit

$$\mu|_{J_k} = \mu_k.$$

Für jede Funktion $f \in L^2(X, \mu)$ setzen wir

$$f_k = f\mathbf{1}_{J_k} \in L^2(J_k, \mu_k)$$

so dass $f = \sum_k f_k$. Jetzt definieren wir die Abbildung

$$\begin{aligned} U &: L^2(X, \mu) \to H \\ Uf &= \sum_k U_k f_k. \end{aligned}$$

Die Abbildung U ist unitär da

$$\|Uf\|^2 = \sum_k \|U_k f_k\|^2 = \sum_k \|f_k\|_{L^2}^2 = \|f\|_{L^2}^2 .$$

Die Abbildung U ist surjektiv, da für jedes $\eta \in H$ eine Darstellung gilt

$$\eta = \sum_k \eta_k \quad \text{mit } \eta_k \in H_k$$

und wir setzen

$$f_k = U_k^{-1}\eta_k \in L^2(J_k, \mu_k)$$

und $f = \sum_k f_k$ so dass $Uf = \eta$. Somit ist U eine lineare Isometrie. Für die Abbildung U gilt

$$AUf = \sum_k AU_k f_k = \sum_k U_k T_{a_k} f_k = UT_a f$$

wobei die Funktion a auf X wird wie folgt definiert:

$$a|_{J_k} = a_k.$$

Jetzt beweisen wir die Darstellung (7.14). Der Beweis ist analog zum Gram-Schmidt-Verfahren. Sei $\{\zeta_k\}$ eine dicht liegende Folge von Elementen von H. Setzen wir $\xi_1 = \zeta_1$ und bestimmen den Unterraum $H_1 = H_{\xi_1}$. Dann wählen wir ζ_k mit minimalen Wert von k so dass $\zeta_k \notin H_1$ und definieren

$$\xi_2 = \zeta_k - P_{H_1}\zeta_k$$

so dass $\xi_2 \perp H_1$ und $\zeta_k - \xi_2 \in H_1$. Dann setzen wir $H_2 = H_{\xi_2}$ und bemerken, dass $H_1 \perp H_2$. In der Tat, wir müssen beweisen, dass

$$(f(A)\xi_1, g(A)\xi_2) = 0$$

für beliebige Funktionen $f, g \in C(J)$. Wir haben

$$(f(A)\xi_1, g(A)\xi_2) = (fg(A)\xi_1, \xi_2).$$

Da $fg(A)\xi_1 \in H_1$ und $\xi_2 \perp H_1$, so erhalten wir die Orthogonalität von H_1 und H_2.

Weiter wählen wir wieder ζ_k mit minimalen Wert von k so dass $\zeta_k \notin H_1 \oplus H_2$ und definieren

$$\xi_3 = \zeta_k - P_{H_1 \oplus H_2}\zeta_k$$

so dass $\xi_3 \perp H_1 \oplus H_2$ und $\zeta_k - \xi_3 \in H_1 \oplus H_2$, und setzen $H_3 = H_{\xi_3.}$und so weiter. Unter diesem Verfahren haben wir immer

$$\begin{aligned} \zeta_1 &= \xi_1 \in H_1 \\ \zeta_2 &\in \operatorname{span}\{\xi_1, \xi_2\} \in H_1 \oplus H_2 \\ \zeta_3 &\in \operatorname{span}\{\xi_1, \xi_2, \xi_2\} \in H_1 \oplus H_2 \oplus H_3 \end{aligned}$$

usw. so dass alle ζ_k in $\bigoplus_{k=1}^N H_k$ liegen, woraus (7.14) folgt. ■