Funktionalanalysis I

Prof. Dr. Petra Wittbold
(geTEXt und ergänzt von Frank Gabriel)

Sommersemester 2006

Version vom 8. Juni 2009

Inhaltsverzeichnis

1 Grundlagen **6**

1.1 Metrische Räume . . . 6

1.2 Normierte Vektorräume . . . 10

1.3 Topologische Begriffe . . . 13

1.4 Stetigkeit von Abbildungen . . . 17

1.5 Der Banachsche Fixpunktsatz . . . 19

1.6 Dichtheit und Separabilität . . . 23

1.7 Kompaktheit . . . 28

1.8 Der Satz von Arzelà Ascoli . . . 35

1.9 Der Satz von Baire . . . 39

1.10 Ergänzung: Eigenschaften von normierten Vektorräumen . . . 41

2 Lineare Operatoren und 3 fundamentale Sätze **44**

2.1 Grundbegriffe . . . 44

2.2 Invertierbare Operatoren . . . 52

2.3 Die drei wichtigen Hauptsätze der Funktionalanalysis . . . 60

2.3.1 Der Satz von Banach-Steinhaus . . . 60

2.3.2 Der Satz von der offenen Abbildung . . . 63

2.3.3 Der Satz vom abgeschlossenen Graphen . . . 70

3 Die Sätze von Hahn-Banach und ihre Konsequenzen **72**

3.1 Die Fortsetzungssätze von Hahn-Banach . . . 72

3.2 Geometrische Formen des Satzes von Hahn-Banach . . . 83

3.3 Reflexive Räume . . . 91

3.4 Schwache Konvergenz . . . 96

3.5 Adjungierte und kompakte Operatoren . . . 109

4 Kompakte Operatoren **118**

4.1 Eine Antwort auf $\overline{F(X,Y)} \stackrel{?}{=} K(X,Y)$ . . . 118

4.2 Spektrum und Resolvente beschränkter, linearer Operatoren . . . 129

4.3 Das Spektrum von kompakten Operatoren . . . 138

4.4 Quotientenräume . . . 146

5 Hilberträume, Orthonormalbasen und kompakte Operatoren **153**

5.1 Hilberträume . . . 153

5.2 Orthonormalbasen . . . 166

5.3 Operatoren im Hilbertraum . . . 171

6 Ausblick: Nichtlineare Funktionalanalysis **185**

6.1 Fixpunktsätze . . . 185

6.2 Anwendungen des Schauderschen Fixpunktsatzes . . . 194

6.3 Weitere Fixpunktsätze . . . 198

7 Anhang **209**

7.1 Die allgemeine Version des Satzes von Arzelà-Ascoli . . . 209

7.2 Schwache Konvergenz in l^1 = starke Konvergenz in l^1 . . . 211

7.3 Komplexe Analysis . . . 213

7.3.1 Holomorphe und ganze Funktionen . . . 213

7.3.2 Laurantreihen . . . 214

7.4 Der Satz vom abgeschlossenen Bild . . . 215

7.5 Der Fixpunktsatz von Brouwer . . . 219

7.5.1 Ein Ausflug in die Variationsrechnung . . . 219

7.5.2 Der Beweis des Brouwerschen Fixpunktsatzes . . . 224

7.5.3 Der Satz von Kakutani . . . 226

Literatur

Ein Standardwerk:

D. Werner: Funktionalanalysis, Springer

Klassiker:

Walter Rudin: Functional Analysis, Tata McGraw-Hill

Eberhard Zeidler: Applied Functional Analysis, Serie: Applied Mathematical Sciences Vol. 108 und 109, Springer

Weitere Lehrbücher zur Funktionalanalysis:

Michael Ruzicka : Nichtlineare Funktionalanalysis. Springer

J. Conway: A Course in Functional Anaylsis

M. Reed, B. Simon: Methods of Modern Mathematical

Physics Vol. I: Functional Analysis

Lehrbücher zur Funktionentheorie:

Wolfgang Fischer, Ingo Lieb: Funktionentheorie, vieweg

Walter Rudin: Complex Analysis, Tata McGraw-Hill

Skripten:

Dirk Ferus: Skripten zur Analysis, insbesondere zur komplexen Analysis:

`http://www.math.tu-berlin.de/~ferus/skripten.html`

Abbildung auf der Titelseite (von links nach rechts): *Hans Hahn (1879-1934)* und *Stefan Banach (1892-1945)* [Quelle: `http://www.Wikipedia.de`]

Fehlermeldungen an

- `Frank-Gabriel@arcor.de`
- `wittbold@math.tu-berlin.de`

1 Grundlagen

In diesem Skriptum benutzen wir die folgenden Schreibweisen:

$$\begin{array}{rclrcl} \mathbb{N} & = & \{1,2,3,\ldots\}\,, & \mathbb{N}_0 & = & \{0,1,2,\ldots\}\,, \\ \mathbb{R}^+ & = & [0,+\infty[, & \mathbb{K} & = & \mathbb{R} \text{ oder } \mathbb{C}. \end{array}$$

- Metrische Räume, Konvergenz, Vollständigkeit, (wichtige) Beispiele
- Normierte Räume, Banachräume, wichtige Beispiele, Äquivalenz
- Topologische Begriffe: Kugel, offen, abgeschlossen, Umgebung, innerer Punkt, Topologischer Raum, Vereinigung und Durchschnitt offener sowie abgeschlossener Mengen, Hausdorff-Eigenschaft, das Innere, der Abschluss, der Rand
- Stetigkeit und äquivalente Formulierungen
- Der Banachsche Fixpunktsatz
- Dichtheit und Separabilität sowie nützliche Kriterien, Weierstraßsche Approximationssatz
- Separable Räume: $(C[a,b], \|.\|_\infty)$ und $(l^p(\mathbb{N}), \|.\|_p)$ mit $1 \leq p < \infty$
- Nicht separable Räume: $(l^\infty(\mathbb{N}), \|.\|_\infty)$ und $L^\infty[0,1]$
- (Relative) Kompaktheit und nützliche Kriterien, Satz von Heine-Borel, Präkompaktheit, (relative) Folgenkompaktheit, Häufungspunkt
- Der Satz von Arzelà-Ascoli
- Der sehr wichtige Satz von Baire

Vieles, was ich hier vorstelle, werden Sie natürlich schon aus dem Grundstudium kennen. Sehen Sie diesen Abschnitt dann als schnellen Wiederholungsmöglichkeit an. Ich werde natürlich nicht alles (zumindest zu Beginn) bis ins kleinste Detail beweisen. Wer es genau wissen möchte, sollte dann in seine alten Grundstudiumsunterlagen schauen oder auf die jeweiligen Verweise eingehen.

1.1 Metrische Räume

Definition 1.1 (Metrischer Raum). Sei M eine Menge. Eine Abbildung $d : M \times M \to \mathbb{R}^+$ heißt *Metrik auf M*, wenn für alle $x, y, z \in M$

(i) $d(x,y) = 0 \Longrightarrow x = y$,

(ii) $d(x,y) = d(y,x)$ (Symmetrie),

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

Das Paar (M,d) heißt dann *metrischer Raum.*

Gilt in (i) nur die Implikation „$\Leftarrow$“, so nennt man d eine *Pseudometrik*, *Semimetrik* oder *Halbmetrik.*

Ist sogar $d(x,y) = \infty$ zugelassen, so nennt man d eine *Fastmetrik.*

Bemerkung. Aus den drei Eigenschaften folgt $d(x,y) \geq 0$.

Beispiel 1.1 (**Standardmetrik**). Für $M = \mathbb{R}$ heißt

$$d(x,y) := |x-y|\,, \quad x,y \in \mathbb{R} \tag{1.1}$$

die so genannte *Standardmetrik*.

□

Beispiel 1.2 (**Euklidische Metrik**). Für $M \subset \mathbb{R}^N$, $N \in \mathbb{N}$, heißt

$$d_2(x,y) := \left(\sum_{k=1}^{N} |x_k - y_k|^2\right)^{1/2}, \quad x,y \in \mathbb{R}^N \tag{1.2}$$

die so genannte *Euklidische Metrik*

Die Metrik in (1.2) kann man für $1 \leq p < \infty$ verallgemeinern zu

$$d_p(x,y) := \left(\sum_{k=1}^{N} |x_k - y_k|^p\right)^{1/p}, \quad x,y \in \mathbb{R}^N. \tag{1.3}$$

Man kann zeigen, dass dann für $p = \infty$

$$d_\infty(x,y) = \sup_{k=1,\dots,N} |x_k - y_k|\,, \quad x,y \in \mathbb{R}^N. \tag{1.4}$$

□

Beispiel 1.3 (**Diskrete Metrik**). Sei M eine beliebige Menge. Dann heißt

$$d(x,y) := \begin{cases} 1 & \text{, falls } x = y, \\ 0 & \text{, sonst.} \end{cases} \tag{1.5}$$

die *diskrete Metrik* auf M.

□

Beispiel 1.4. Sei M eine beliebige Menge und d eine Metrik auf M. Dann ist auch

$$\bar{d}(x,y) := \min\{1, d(x,y)\}\,, \quad x,y \in M \tag{1.6}$$

eine Metrik auf M.

□

Beispiel 1.5. Sei $M = C[a,b] := \{\, f : [a,b] \to \mathbb{C} \mid f \text{ stetig} \,\}$. Dann ist

$$d(f,g) := \max_{x\in[a,b]} |f(x) - g(x)|\,, \quad f,g \in C[a,b] \tag{1.7}$$

eine Metrik auf M.

□

Konvergenz:

Definition 1.2 (**Konvergenz, Cauchyfolge**). Sei (M,d) ein metrischer Raum. Seien $(x_n)_n \subset M$ eine Folge und $x \in M$. Man sagt

(i) $(x_n)_n$ *konvergiert gegen* x, falls gilt

$$\forall_{\epsilon>0} \exists_{n_0\in\mathbb{N}} : d(x_n, x) < \epsilon \quad \text{für alle } n \geq n_0.$$

Notation:

$$\lim_{n\to\infty} x_n = x \quad \text{oder} \quad x_n \xrightarrow{d} x.$$

(ii) $(x_n)_n$ *konvergiert in* M, falls ein $x \in M$ existiert, so dass $x_n \xrightarrow{d} x$.

(iii) $(x_n)_n$ heißt *Cauchy-Folge*, falls

$$\forall_{\epsilon>0} \exists_{n_0\in\mathbb{N}} : d(x_n, x_m) < \epsilon \quad \text{für alle } n,m \geq n_0.$$

(iv) Der metrische Raum (M,d) heißt *vollständig*, falls jede Cauchy-Folge in M konvergiert.

Bemerkungen.

1. Direkt aus der Definition folgt

$$\lim_{n\to\infty} x_n = x \text{ in } (M,d) \Longleftrightarrow d(x_n,x) \to x \text{ in } \mathbb{R}.$$

2. Falls $\lim x_n = x$ und $\lim x_n = y$, so folgt $x = y$, d.h. Grenzwerte sind eindeutig: Mit der Dreiecksungleichung folgt nämlich

$$d(x,y) \le d(x,x_n) + d(x_n,y) \overset{n\to\infty}{\longrightarrow} 0.$$

   Also ist $x = y$.

3. Im Allgemeinen sind Cauchy-Folgen in einem metrischen Raum nicht notwendig konvergent (vgl. das folgende Beispiel 1.6).

Beispiel 1.6. Sei $M := C^1[-1,1]$ ausgestattet mit der Metrik (1.7)

$$d(f,g) = \max_{x\in[-1,1]} |f(x) - g(x)|, \quad f,g \in M.$$

Dann ist (M,d) ein metrischer Raum. Betrachte die Funktionsfolge

$$f_n(x) := (x^2 + 1/n)^{1/2}, \quad x \in [-1,1],\ n \in \mathbb{N}.$$

Die Folge $(f_n)_n$ ist stetig differenzierbar, also $(f_n)_n \subset M$. Außerdem konvergiert sie gleichmäßig gegen die nicht(!) differenzierbare Funktion $|.|$. Insbesondere ist daher $(f_n)_n$ eine Cauchy-Folge in M, aber eben nicht konvergent in M, da $f \notin M$.

Wir bemerken noch, dass M einen Untervektorraum[1] von $C[a,b]$ darstellt, aber dieser nicht abgeschlossen bezüglich d ist.

□

Beispiel 1.7. Sei $N \in \mathbb{N}$ und $M := \mathbb{R}^N$ ausgestattet mit der Metrik d_p, $1 \le p \le \infty$, definiert wie in (1.3). Dann gilt

$$\mathbb{R}^N \supset (x_n)_n \overset{d_p}{\to} x$$

genau dann, wenn sämtliche Komponentenfunktionen konvergieren, d.h. wenn

$$\mathbb{R} \supset (x_n^k)_n \to x^k \quad \text{in } \mathbb{R} \text{ für alle } 1 \le k \le N.$$

Das folgt im Grunde aus den folgenden Abschätzungen:

- $p \ne \infty$:

$$|x_k^j - x^j| \le \underbrace{\left(\sum_{i=1}^N |x_k^j - x^j|^p\right)^{1/p}}_{=d_p(x_k,x)} \le N^{1/p} \sup_{1\le j\le n} |x_k^j - x^j|.$$

- $p = \infty$:

$$|x_k^j - x^j| \le \underbrace{\sup_{1\le j\le n} |x_k^j - x^j|}_{d_\infty(x_k,x)} \le \left(\sum_{i=1}^N |x_k^j - x^j|^p\right)^{1/p}.$$

□

[1] Zur Erinnerung: Sei E ein K-Vektorraum und $U \subset E$. Dann heißt U ein *Untervektorraum von* E, wenn gilt: Für alle $x,y \in U$ und $\lambda \in K$ sind

$$x + y \in U \text{ und } \lambda x \in U$$

und

$$U \ne \emptyset.$$

Beispiel 1.8. Sei (M, d) ein metrischer Raum mit der Metrik (1.6)

$$\bar{d}(x, y) := \min\{1, d(x, y)\}, \quad x, y \in M$$

Dann gilt $x_n \xrightarrow{d} x$ genau dann, wenn $x_n \xrightarrow{\bar{d}} x$.

Beispiel 1.9. Sei $C[a, b]$ ausgestattet mit der Metrik (1.7)

$$d(f, g) = \max_{x \in [a,b]} |f(x) - g(x)|, \quad f, g \in M.$$

Dann ist die Konvergenz bezüglich d gleichbedeutend mit der gleichmäßigen Konvergenz auf $C[a, b]$:

$$f_n \xrightarrow{d} f \Longleftrightarrow f_n \overset{\text{glm. auf } [a,\,b]}{\longrightarrow} f.$$

$\square$

1.2 Normierte Vektorräume

Wir beginnen mit einer bekannten

Definition 1.3 (Vektorraum). Sei K ein Körper. Ein *Vektorraum über K* oder ein *K-Vektorraum* ist eine Menge E zusammen mit zwei Abbildungen

$$E \times E \to E, \;\; (v, w) \mapsto v + w,$$
$$K \times E \to E, \;\; (\lambda, v) \mapsto \lambda v,$$

die man *(Vektor-)Addition* und *Skalarmultiplikation* nennt, so dass gilt:

1. E ist zusammen mit der Addition eine abelsche Gruppe. Das neutrale Element bezeichnen wir mit 0 oder mit 0_E. Wir nennen es auch den *Nullvektor von E*.
2. Mit beliebigen $v, w \in E$ und $\lambda, \mu \in K$ ist

$$\lambda(\mu v) = (\lambda\mu)v,$$
$$\lambda(v + w) = \lambda v + \lambda w, \;\; (\lambda + \mu)v = \lambda v + \mu v,$$
$$1v = v.$$

   Dabei ist 1 das neutrale Element der Multiplikation in K. Die Elemente von K bezeichnet man häufig auch als *Skalare*, die Elemente von E als *Vektoren*.

Wohlbekannt ist auch die

Definition 1.4 (Norm). Sei E ein $\mathbb{K}$-Vektorraum. Eine *Norm auf E* ist eine Abbildung $\|.\| : E \to \mathbb{R}^+$ mit den Eigenschaften

(i) $\|x\| = 0 \Longleftrightarrow x = 0$.

(ii) $\|\alpha x\| = |\alpha| \, \|x\|$ für alle $\alpha \in \mathbb{K}$ und $x \in E$.

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

$(E, \|.\|)$ heißt *normierter Vektorraum*. (Falls in (i) nur „$\Leftarrow$“ gilt, so heißt $\|.\|$ eine *Halbnorm*).

Bemerkungen.

1. Ist $(E, \|.\|)$ ein normierter Vektorraum, so definiert

$$d(x, y) := \|x - y\|, \quad x, y \in E \tag{1.8}$$

   eine Metrik auf E. Diese Metrik heißt die von der Norm *induzierte Metrik*. Im Allgemeinen betrachtet man E bzw. Teilmengen von E ausgestattet mit dieser Metrik.

2. Nicht jede Metrik auf einem nicht-trivialen Vektorraum E wird durch eine Norm induziert:

   Betrachte dazu den normierten Vektorraum $(E, \|.\|)$. Dann wird die Metrik

$$d(x, y) := \min\{\|x - y\|, 1\}$$

   von keiner Norm induziert.

   Angenommen, das ist doch der Fall. Dann gibt es also eine Norm $|||\,.\,|||$, so dass

$$|||\, x - y \,||| = d(x, y) = \min\{\|x - y\|, 1\}, \quad x, y \in \mathbb{R}.$$

Daraus folgt

$$||| \, x \, ||| = \min \{\|x\|, 1\} \, .$$

Doch das ist keine Norm, denn für $x = 1$ ist die Eigenschaft (ii) verletzt:

$$2||| \, 1 \, ||| = 2 \neq ||| \, 2 \cdot 1 \, ||| = \|2\| = 1.$$

Auch wenn d von keiner Norm induziert wird, so ist d aber zu einer von einer Norm induzierten Metrik äquivalent, falls E kompakt ist. Setze man nämlich

$$M = 1 \quad \text{und} \quad m = (\max_{x,y \in E} \|x - y\|)^{-1} > 0,$$

so folgt mit $\tilde{d}(x, y) := \|x - y\|$ für alle $x, y \in E$

$$m\tilde{d}(x, y) \leq d(x, y) \leq M\tilde{d}(x, y).$$

Daher sind d und $\tilde{d}$ äquivalent. Warum gilt tatsächlich $m > 0$?

Beispiel 1.10 (**p-Norm**). Auf dem $\mathbb{R}^N$ ist die *p-Norm* gegeben durch

$$\|x\|_p := \begin{cases} \left(\sum_{k=1}^N |x_k|^p\right)^{1/p} & , \text{ für } 1 \leq p < \infty, \\ \sup_{k=1,\ldots,N} |x_k| & , \text{ für } p = \infty \end{cases}, \quad x \in \mathbb{R}^N. \tag{1.9}$$

Die $\|.\|_p$-Norm induziert die p-Metrik d_p (vgl. Beispiel 1.2).

□

Beispiel 1.11. Auf $C[a, b]$ hat man zum Beispiel die Norm

$$\|f\|_\infty := \max_{x \in [a,b]} |f(x)| \, , \quad f \in C[a, b]. \tag{1.10}$$

Es gibt natürlich auch noch viele andere Normen auf $C[a, b]$, so zum Beispiel für $1 \leq p < \infty$ die Norm

$$\|f\|_p := \left(\int_a^b |f(x)|^p \, dt\right)^{1/p}, \quad f \in C[a, b]. \tag{1.11}$$

□

Beispiel 1.12. Betrachte den Vektorraum

$$L^p = \{ \, f :]a, b[\to \mathbb{R} \mid f \text{ Lebesgue-messbar und } \int_a^b |f(t)|^p < \infty \, \}. \tag{1.12}$$

Dann ist $(L^p, \|.\|_p)$ ein normierter Vektorraum.

□

Beispiel 1.13 (**Folgenräume**). Die kommenden Folgenräume sind allesamt normierte Vekttorräume (für nähere Informationen lesen Sie *Werner, p. 3-20*):

a) Der Raum der beschränkten Folgen:

$$l^\infty := l^\infty(\mathbb{N}) = \{ \, (x_n)_n \subset \mathbb{K} \mid (x_n)_n \text{ beschränkte Folge} \, \},$$

ausgestattet mit der Norm

$$\|x\|_\infty = \sup_{n \in \mathbb{N}} |x_n| \, , \quad x = (x_n)_n \in l^\infty.$$

b) Die l^p-Räume:

$$l^p := \{ \, (x_n)_n \subset \mathbb{K} \mid \sum_{k=1}^\infty |x_n|^p < \infty \, \},$$

ausgestattet mit der Norm

$$\|x\|_p = \left(\sum_{n=1}^\infty |x_n|^p\right)^{1/p}, \quad x = (x_n)_n \in l^p.$$

c) Der Raum der konvergenten Folgen, insbesondere der Raum der Nullfolgen:

$$c := \{\, (x_n)_n \subset \mathbb{K} \mid (x_n)_n \text{ konvergent} \,\}$$

bzw.

$$c_0 := \{\, (x_n)_n \subset \mathbb{K} \mid \lim x_n = 0 \,\},$$

ausgestattet mit der Norm

$$\|x\|_\infty = \sup_{n\in\mathbb{N}} |x_n|\,, \quad x = (x_n)_n \in c \text{ bzw. } c_0.$$

□

Definition 1.5 (**Vollständigkeit, Banachraum**). Sei $(E, \|.\|)$ ein normierter Vektorraum. Dann heißt E *vollständig*, falls (E, d) bezüglich der von der Norm induzierten Metrik

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

vollständig ist.

$(E, \|.\|)$ nennt man dann einen *Banachraum*.

Bemerkung. Beachte, dass nicht jeder normierte Vektorraum vollständig ist. Zum Beispiel ist $(C^1[-1, 1], \|.\|_\infty)$ kein Banachraum (vgl. Beispiel 1.6).

Beispiel 1.14 (**für Banachräume**). Die folgenden Räume stellen allesamt Banachräume dar (für nähere Informationen lesen Sie *Werner, p. 3-20*)

a) $(l^p, \|.\|_p)$, $1 \le p \le \infty$.

b) $(C[a, b], \|.\|_\infty)$.

c) $(L^p[a, b], \|.\|_p)$, $1 \le p < \infty$.

d) $L^\infty[a, b], \|.\|_\infty$. Dabei hat man die Norm

$$\|f\|_\infty = \operatorname{esssup}_{x\in[a,b]} |f(x)| = \inf_{N\in\Sigma,\lambda(N)=0} \sup_{t\in I\setminus N} |f(t)|\,, \quad f \in L^\infty[a, b].$$

e) $(\mathbb{R}^n, \|.\|_p)$.

f) $(C^1[a, b], \|.\|)$ mit

$$\|f\| := \max\{\|f\|_\infty, \|f'\|_\infty\}.$$

□

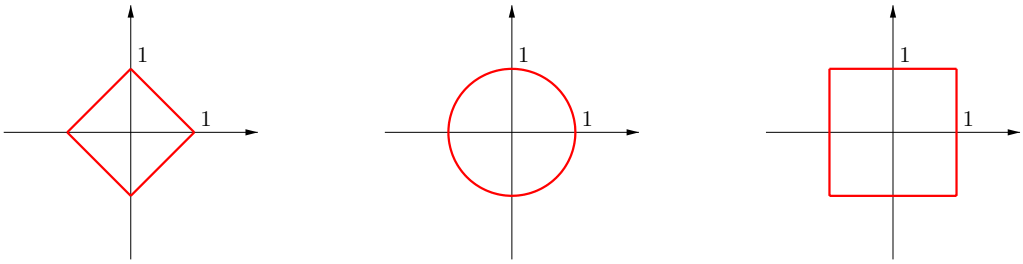
1.3 Topologische Begriffe

Definition 1.6 (Kugel). Sei (M, d) ein metrischer Raum. Seien $x \in M$ und $r > 0$. Dann heißt

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

die *offene Kugel vom Radius r um x*.

Beispiel 1.15. Hier sehen Sie $B(0, 1)$ bezüglich veschiedener Metriken:



$(\mathbb{R}^2, d_1)$ $(\mathbb{R}^2, d_2)$ $(\mathbb{R}^2, d_\infty)$

□

Definition 1.7 (Offen, abgeschlossen, Umgebung, innerer Punkt). Sei (M, d) ein metrischer Raum.

(i) Ein Punkt $x \in M$ heißt *innerer Punkt von M*, falls ein $r > 0$ existiert mit

$$B(x, r) \subset M.$$

(ii) Eine Teilmenge $O \subset M$ heißt *offen*, wenn jeder Punkt in O ein innerer Punkt ist, d.h.

$$\forall_{x \in O} \exists_{r > 0} : B(x, r) \subset O.$$

(iii) Eine Menge $A \subset M$ heißt *abgeschlossen*, falls das Komplement $A^C = M \setminus A$ offen ist.

(iv) Eine Menge $U \subset M$ heißt *Umgebung eines Punktes $x \in M$*, falls ein $r > 0$ exsitiert mit

$$B(x, r) \subset U.$$

Wichtige Fakten:

1. Die offene Kugeln sind nach der Definition tatsächlich offen.
2. (i) $\emptyset$ und M sind offen.
   (ii) $O_1, O_2 \subset M$ offen $\Longrightarrow O_1 \cap O_2$ offen.
   (iii) $(O_i)_{i \in I} \subset M$ offen $\Longrightarrow \bigcup_{i \in I} O_i$ offen.
3. (i) $\emptyset$ und M sind abgeschlossen.
   (ii) $O_1, O_2 \subset M$ abgeschlossen $\Longrightarrow O_1 \cup O_2$ abgeschlossen.
   (iii) $(O_i)_{i \in I} \subset M$ abgeschlossen $\Longrightarrow \bigcap_{i \in I} O_i$ abgeschlossen.

Beweis. Zu 1: Wir müssen zeigen, dass $B(x,r)$ offen ist, d.h. für alle $y \in B(x,r)$ gibt es ein $\epsilon > 0$, so dass $B(y,\epsilon) \subset B(x,r)$.

Sei also $y \in B(x,r)$, d.h. $d(x,y) < r$. Definiere

$$\epsilon := r - d(x,y) > 0.$$

Dann gilt

$$B(y,\epsilon) \subset B(x,r),$$

denn ist $z \in B(y,\epsilon)$, so folgt

$$d(x,z) \leq d(x,y) + d(y,z) \leq d(x,y) + \epsilon = d(x,y) + r - d(x,y) = r,$$

also ist $z \in B(x,r)$.

Zu 2 (i): Es ist trivial, dass M offen ist und es folgt daraus, dass $\emptyset$ abgeschlossen ist. Aber $\emptyset$ ist auch offen, da es keinen Punkt $x \in \emptyset$ gibt, zu dem es eine ϵ-Umgebung $B(x,\epsilon) \subset \emptyset$ geben müsste. Daraus folgt dann, dass X abgeschlossen ist.

Zu 2 (ii): Ist $O_1 \cap O_2 = \emptyset$, so ist nach (i) nichts zu zeigen. Sei also $O_1 \cap O_2 \neq \emptyset$. Sei $x \in O_1 \cap O_2$. Da O_i offen ist, gibt es $\epsilon_i > 0$, so dass

$$B(x,\epsilon_i) \subset O_i, \quad i = 1,2.$$

Setze

$$\epsilon := \min\{\epsilon_1, \epsilon_2\} > 0.$$

Dann gilt

$$B(x,\epsilon) \subset O_1 \cap O_2,$$

also ist $O_1 \cap O_2$ offen.

Zu 2 (iii): Ist die Vereinigung leer, so ist nichts zu zeigen. Sei also $\bigcup O_i \neq \emptyset$. Dann gibt es ein $i \in I$, so dass $O_i \neq \emptyset$. Sei dann $x \in O_i$. Da O_i offen ist, gibt es ein $\epsilon > 0$, so dass

$$B(x,\epsilon) \subset O_i \subset \bigcup_{i \in I} O_i.$$

Das zeigt die Offenheit der Vereinigung.

Zu 3 (i): Das haben wir schon in 2 (i) erledigt.

Zu 3 (ii): Sei also $A_1, A_2 \subset M$ abgeschlossen. Dann sind $M \setminus A_1$ und $M \setminus A_2$ offen, also ist nach 2(ii) auch

$$M \setminus A_1 \cap M \setminus A_2 = M \setminus (A_1 \cup A_2)$$

offen. Daher ist $A_1 \cup A_2$ abgeschlossen.

Zu 3 (iii): Sei also $(A_i)_{i \in I}$ eine abgeschlossene Familie in M. Dann ist für alle $i \in I$ die Menge $M \setminus A_i$ offen. Nach 2 (iii) ist dann auch

$$\bigcup_{i \in I} M \setminus A_i = M \setminus \bigcap_{i \in I} A_i$$

offen. Das zeigt die Abgeschlossenheit von $\bigcap_{i \in I} A_i$. □

Bemerkung. Sei M eine beliebige Menge und sei $\mathcal{T}$ eine Familie von Teilmengen von M mit den Eigenschaften 2. (i), (ii) und (iii), d.h.

(i) $\emptyset, M \in \mathcal{T}$,

(ii) $O_1, O_2 \in \mathcal{T} \Longrightarrow O_1 \cap O_2 \in \mathcal{T}$,

(iii) $(O_i)_{i \in I} \in \mathcal{T} \Longrightarrow \bigcup_{i \in I} O_i \in \mathcal{T}$.

Dann heißt $\mathcal{T}$ eine *Topologie auf* M und $(M, \mathcal{T})$ heißt *topologischer Raum*. Die Elemente in $\mathcal{T}$ nennt man die *offenen Teilmengen von* M.

Achtung. Es gibt Topologien, die nicht von einer Metrik induziert werden (vgl. Beispiel 1.17)!

Es reicht aus, eine Topologie auf einer Menge zu definieren, um Begriffe wie Konvergenz, Stetigkeit usw. erklären zu können.

Definition 1.8 (Hausdorff-Eigenschaft). Ein topologischer Raum $(M, \mathcal{T})$ heißt *hausdorffsch*, wenn man verschiedene Punkte $x, y \in M$ durch offene Kugeln trennen kann, d.h. es gibt $O_1, O_2 \in \mathcal{T}$ mit

$$x \in O_1,\ y \in O_2 \text{ und } O_1 \cap O_2 = \emptyset.$$

Beispiel 1.16. Jeder metrische Raum ist hausdorffsch.

Beweis. Seien also $x, y \in M$ mit $x \neq y$. Setze

$$\epsilon := d(x, y) > 0.$$

Definiere dann

$$O_1 := B(x, \epsilon/3) \quad \text{und} \quad O_2 := B(y, \epsilon/3).$$

Wäre nun $z \in O_1 \cap O_2$, so wäre

$$\epsilon = d(x, y) \leq d(x, z) + d(z, y) < \epsilon/3 + \epsilon/3 = 2/3\epsilon.$$

Das ist aber ein Widerspruch, also gibt es kein $z \in O_1 \cap O_2$: Der Durchschnitt ist leer.

□

Beispiel 1.17. Sei $M := \{1, 2\}$. Wir betrachten die so genannte *Klumpentopologie*

$$\mathcal{T} := \{\emptyset, M\}.$$

Dann ist $(M, \mathcal{T})$ nicht hausdorffsch und kann insbesondere nach Beispiel 1.16 nicht von einer Metrik erzeugt sein.

□

Definition 1.9 (Inneres, Abschluss, Rand). Seien (M, d) ein metrischer Raum und $A \subset M$.

(i) Die Menge

$$\overset{\circ}{A} := A^\circ = \{\, x \in A \mid x \text{ ist innerer Punkt von } A \,\}$$

heißt das *Innere von* A. A° ist die größte offene Teilmenge von A, d.h. $A^\circ \subset A$.

(ii) Die Menge

$$\bar{A} := \{\, x \in M \mid \forall_{r>0} B(x, t) \cap A \neq \emptyset \,\}$$

heißt der *Abschluss von* A. $\bar{A}$ ist die kleineste abgeschlossene Menge M, die A umfasst, also $A \subset \bar{A}$. Man kann zeigen, dass

$$\bar{A} = \{\, x \in M \mid \exists_{(x_n)_n \subset A} : x_n \xrightarrow{d} x \,\}.$$

(iii) Die Menge

$$\partial A := \bar{A} \setminus A^\circ$$

heißt der *Rand von* A und es gilt

$$\partial A = \{\, x \in M \mid \forall_{r>0} : B(x, r) \cap A \neq \emptyset \neq B(x, r) \cap M \setminus A \,\}.$$

Zum Abschluss dieses Abschnittes noch ein nützlicher

Satz 1.1. *Sei M ein vollständiger metrischer Raum und $A \subset M$. Dann gilt*

$$A \text{ ist vollständig} \Longleftrightarrow A \text{ ist abgeschlossen.}$$

Beweis. Zu „$\Longrightarrow$“. Sei (A, d_A) vollständig und sei $(a_k)_{k\in\mathbb{N}}$ eine Folge in A, die gegen $x \in M$ konvergiert. Wir müssen zeigen, dass $x \in A$.

Als konvergente Folge ist (a_k) eine d-Cauchyfolge. Damit ist sie wegen $d(a_k, a_l) = d_A(a_k, a_l)$ auch eine d_A-Cauchyfolge und nach Voraussetzung konvergent gegen ein $a \in A$. Das bedeutet

$$\lim d_A(a_k, a) = \lim d(a_k, a) = 0.$$

Daher konvergiert (a_k) auch in M gegen a. Dann ist aber wegen der Eindeutigkeit des Grenzwertes $x = a \in A$.

Zu „$\Longleftarrow$“. Seien nun A abgeschlossen und (a_k) eine Cauchyfolge in (A, d_A). Das ist dann auch eine Cauchyfolge in (M, d), also konvergent gegen ein $x \in M$. Weil A abgeschlossen ist, liegt x in A und ist der Grenzwert von (a_k) in (A, d_A). $\square$

1.4 Stetigkeit von Abbildungen

Wir kommen nun zum Begriff der *Stetigkeit*, den Sie natürlich schon aus dem Grundstudium kennen. Wir wollen hier auch nur noch einmal schnell die Definition und äquivalente Formulierungen geben:

Definition 1.10 ((Folgen-)Stetigkeit). Seien (M_1, d_1) und (M_2, d_2) zwei metrische Räume sowie $f : M_1 \to M_2$ eine Abbildung.

(i) f heißt *(folgen-)stetig in* $x \in M_1$, wenn für jede Folge $(x_n)_n \subset M_1$ mit $\lim x_n = x$ gilt

$$\lim f(x_n) = f(x).$$

(ii) f heißt *stetig auf* M_1, wenn f stetig in jedem $x \in M_1$.

Manchmal ist es nützlich einige äquivalente Formulierungen zu kennen:

Satz 1.2. *Seien (M_1, d_1) und (M_2, d_2) metrische Räume und $f : M_1 \to M_2$ eine Abbildung. Weiter sei $x \in M_1$. Dann sind die folgenden Aussagen äquivalent:*

(i) f ist stetig in x,

(ii)

$$\forall_{\epsilon>0} \exists_{\delta>0} : (d_1(x, y) < \delta \Longrightarrow d_2(f(x), f(y))) < \epsilon,$$

(iii) Für alle Umgebungen U von $f(x)$ in M_2 gilt

$$f^{-1}(U) \text{ ist Umgebung von } x,$$

(iv) Für alle offenen (abgeschlossenen) Mengen $U \subset M_2$ mit $f(x) \in U$ gilt

$$f^{-1}(U) \text{ ist offen (abgeschlossen) in } M_1 \text{ und enthält } x.$$

Bemerkungen.

1. Man kann mit Hilfe von (iii) die Stetigkeit von Abbildungen

   $$f : (M_1, \mathcal{T}_1) \to (M_2, \mathcal{T}_2)$$

   in topologischen Räumen M_1 und M_2 definieren.

   Natürlich kann man auch die Folgenstetigkeit von Abbildungen in topologischen Räumen definieren. Aber die beiden Begriffe sind i.a. nicht zueinander äquivalent!

2. Die Formulierung (ii) der Stetigkeit von f in x ist äquivalent zu

   $$\forall_{\epsilon>0} \exists_{\delta>0} : f(B(x, \delta)) \subset B(f(x), \epsilon).$$

   Das folgt im Grunde aus

   $$d_1(x, y) < \delta \Longleftrightarrow y \in B(x, \delta) \Longleftrightarrow f(y) \in f(B(x, \delta))$$

   und

   $$d_2(f(x), f(y)) < \epsilon \Longleftrightarrow f(y) \in B(f(x), \epsilon).$$

Beweis des Satzes. Zu „$(i) \Longrightarrow (ii)$“: Sei $\epsilon > 0$. Gäbe es kein $\delta > 0$ wie angegeben, so gäbe es insbesondere zu jedem $k \in \mathbb{N}$ ein $x_k \in G$ mit

$$d_1(x, x_k) < \frac{1}{k+1}, \text{ aber } d_2(f(x), f(x_k)) \geq \epsilon.$$

Dann wäre aber $\lim x_k = x$ und $\lim f(x_k) \neq f(x)$. Widerspruch!

<u>Zu „$(ii) \Longrightarrow (iii)$“:</u> Sei also U eine Umgebung von $f(x)$, d.h. U ist offen mit $f(x) \in U$. Wir müssen zeigen, dass $f^{-1}(U)$ offen ist und $x \in f^{-1}(U)$.

Weil U offen ist, gibt es ein $\epsilon > 0$, so dass

$$B(f(x), \epsilon) \subset U.$$

Nach (ii) gibt es nun ein $\delta > 0$, so dass

$$f(B(x, \delta)) \subset B(f(x), \epsilon) \subset U.$$

Daraus folgt dann

$$B(x, \delta) \subset f^{-1}(U),$$

also ist $f^{-1}(U)$ offen und $x \in f^{-1}(U)$.

<u>Zu „$(iii) \Longrightarrow (iv)$“:</u> Trivial

<u>Zu „$(iv) \Longrightarrow (i)$“:</u> Sei also $(x_k)_k$ eine Folge mit $\lim x_k = x$, d.h. es gilt

$$\forall_{\delta>0} \exists_{k_0 \in \mathbb{N}} : x_k \in B(x, \delta) \quad \text{für alle } k \geq k_0. \tag{1.13}$$

Sei nun $\epsilon > 0$ beliebig. Nach Voraussetzung ist dann $f^{-1}(B(f(x), \epsilon))$ offen und $x \in f^{-1}(B(f(x), \epsilon))$. Daher gibt es ein $\delta > 0$, so dass

$$B(x, \delta) \subset f^{-1}(B(f(x), \epsilon).$$

Wegen (1.13) gibt es dann ein $k_0 \in \mathbb{N}$, so dass

$$x_k \in B(x, \delta) \subset f^{-1}(B(f(x), \epsilon) \quad \text{für alle } k \geq k_0.$$

Aber das heißt gerade

$$f(x_k) \in B(f(x), \epsilon) \quad \text{für alle } k \geq k_0,$$

was zu zeigen war. $\square$

1.5 Der Banachsche Fixpunktsatz

Ein wichtiges Hilfsmittel der Funktionalanalysis und der Theorie von Differentialgleichungen ist der Banachsche Fixpunktsatz. Mit ihm kann man zum Beispiel die Existenz und Eindeutigkeit einer lokalen Lösung eines Anfangswertproblems zeigen.

Definition 1.11 (Strikt kontraktiv). Eine Abbildung $f : X \to Y$ zwischen metrischen Räumen (M_1, d_1) und (M, d_2) heißt *kontrahierend* oder *strikt-kontraktiv*, wenn es ein $k \in]0, 1[$ gibt, so dass für alle $x_1, x_2 \in M_1$

$$d_2(f(x_1), f(x_2)) \leq k d_1(x_1, x_2).$$

Die Zahl k heißt auch ein *Kontraktionsmodul von* f.

Lemma 1.1 (Cauchyfolgen-Kriterium). *Sei $(x_k)_{k\in\mathbb{N}}$ eine Folge im metrischen Raum (M, d). Setze $a_k := d(x_k, x_{k+1})$. Ist $\sum_{k=0}^{\infty} a_k$ konvergent, so ist $(x_k)_{k\in\mathbb{N}}$ eine Cauchyfolge.*

Beweis. Nach der Dreiecksungleichung gilt

$$d(x_k, x_{k+l}) \leq \sum_{j=k}^{k+l-1} d(x_j, x_{j+1}) = \sum_{j=k}^{k+l-1} a_j = |s_{k+l-1} - s_{k-1}|, \tag{1.14}$$

wenn s_k die k-te Partialsumme der Reihe $\sum_{k=0}^{\infty} a_k$ bezeichnet. Die bilden aber nach Voraussetzung eine Cauchyfolge. □

Bemerkung. Die Umkehrung des Lemmas gilt nicht. Wir betrachten dazu $M = \mathbb{R}$ und die Metrik

$$d(x, y) = \begin{cases} \max\{|x|, |y|\} & \text{für } x \neq y, \\ 0 & \text{sonst.} \end{cases}$$

Dann ist $x_k := \frac{1}{k}$ offensichtlich eine Cauchyfolge, aber $a_k := d(x_k, x_{k+1}) = \frac{1}{k}$ und die Reihe aus dem Lemma ist nicht konvergent (Harmonische Reihe).

Satz 1.3 (Banachscher Fixpunktsatz (1922)). *Sei $A \subset X$ eine nichtleere, abgeschlossene Menge eines vollständigen metrischen Raumes (M, d) und sei*

$$f : A \subset X \longrightarrow A$$

eine strikt kontraktive Selbstabbildung. Dann gilt

(i) Die Gleichung

$$f(x) = x, \quad x \in A$$

hat genau eine Lösung $x \in A$, d.h. f hat genau einen Fixpunkt in A.

(ii) Die durch

$$x_0 \in A, \quad x_{n+1} := f(x_n), \quad n = 0, 1, 2, \ldots$$

definierte Folge (x_n) konvergiert gegen die Lösung x für alle Startwerte $x_0 \in A$.

Beweis. Eindeutigkeit. Angenommen, x und y seien zwei Lösungen der Fixpunktgleichung, d.h. $f(y) = y$ und $f(x) = x$. Aufgrund der strikten Kontraktivität von f, gibt es ein $k \in]0,1[$ mit

$$d(x,y) = d(f(x), f(y)) \leq kd(x,y).$$

Wegen $k < 1$ muss also $d(x,y) = 0$, also $x = y$ sein.

Zur Existenz. Seien also $x_0 \in A$ und $(x_n)_n$ definiert wie im Satz. Wir werden zeigen, dass $(x_n)_{n\in\mathbb{N}}$ eine Cauchyfolge ist. Zunächst ist

$$d(x_n, x_{n+1}) \leq kd(x_{n-1}, x_n) \leq k^2 d(x_{n-2}, x_{n-1}) \leq \ldots \leq k^n d(x_0, x_1). \tag{1.15}$$

Wegen $k \in]0,1[$ konvergiert die Reihe $\sum_{n=0}^{\infty} k^n d(x_0, x_1)$ und nach dem Cauchykriterium ist $(x_n)_n$ eine Cauchyfolge. Da X vollständig und M abgeschlossen ist, ist M auch vollständig. Daher konvergiert die Folge $(x_n)_n$ gegen ein $x^* \in M$.

Es bleibt zu zeigen, dass x^* ein Fixpunkt von f ist. Aber nun gilt für alle n

$$\begin{aligned} d(f(x^*), x^*) &\leq d(f(x^*), x_{n+1}) + d(x_{n+1}, x^*) = d(f(x^*), f(x_n)) + d(x_{n+1}, x^*) \\ &\leq k \underbrace{d(x^*, x_n)}_{\to 0} + \underbrace{d(x_{n+1}, x^*)}_{\to 0}. \end{aligned}$$

Es folgt schließlich $d(f(x^*), x^*) = 0$, also $f(x^*) = x^*$. □

Bemerkungen

1. Die Behauptung des Banachschen Fixpunktsatzes gilt auch unter der schwächeren(!) Voraussetzung, dass nicht f, aber für ein $m \in \mathbb{N}$ die m-te Iteration $f^m := f \circ \ldots \circ f$ kontrahierend ist.

   Beweis. Nach dem Banachschen Fixpunktsatz hat f^m genau(!) einen Fixpunkt: Es gibt genau ein $x \in M$ mit $f^m(x) = x$. Wir wenden nun f auf beiden Seiten an und erhalten

   $$f \circ f^m(x) = f(x) \Longleftrightarrow f^m(f(x)) = f(x).$$

   Also ist auch $f(x)$ Fixpunkt von $f^m(x)$. Der Fixpunkt von f^m ist aber eindeutig und daher folgt $f(x) = x$. Da jeder Fixpunkt von f auch Fixpunkt von f^m ist, hat f genau einen Fixpunkt.

   Es bleibt noch zu zeigen, dass die durch $x_0 \in M$ und $x_{n+1} := f(x_n)$ definierte Folge gegen den Fixpunkt konvergiert. Nach dem Banachschen Fixpunktsatz angewendet auf f^m gilt für jedes $y_0 \in M$ und für die Folge $y_{n+1} = f^m(y_n)$ die Tatsache, dass $(y_n)_n$ gegen den Fixpunkt x konvergiert. Setze

   $$y_0^l := x_l, \quad l \in \{0, 1, \ldots, m-1\}.$$

   Nun gilt

   $$\begin{aligned} y_0^l &= x_l \\ y_1^l &= f^m(y_0^l) = f^m(x_l) = f^{m-1}(f(x_l)) = f^{m-1}(x_{l+1}) = \ldots = x_{m+l} \\ &\vdots \\ y_n^l &= \ldots = x_{nm+l}. \end{aligned}$$

   Daher ist y_n^l eine Teilfolge von x_n. Andererseits ist jedes x_n in einer der Teilfolgen y_n^l enthalten und weil diese ja nach Voraussetzung gegen den Fixpunkt x konvergiert, konvergiert damit auch die Folge x_n gegen x. □

2. Die folgenden Gegenbeispiele zeigen, dass keine der Voraussetzungen des Banachschen Fixpunktsatzes weggelassen werden können:

a) $A = \emptyset$ und f beliebig.
Dann kann f keinen Fixpunkt haben.

b) $A = [0,1]$, $\tilde{A} = [2,3]$ und $f : A \to \tilde{A}$.
f bildet nicht in sich selbst ab und kann daher keinen Fixpunkt haben.

c) $A =]0,1[$ und $f(x) = \frac{x}{2}$.
Lösen der Gleichung $f(x) = x$ ergibt $x = 0$. Also hat f keinen Fixpunkt.

d) $A = \mathbb{R}$ und $f(x) = \frac{\pi}{2} + x - \arctan x$.
Es gilt $f'(x) = 1 - \frac{1}{1+x^2}$ und daher folgt mit dem Mittelwertsatz

$$|f(x) - f(y)| = \left|1 - \frac{1}{1+\xi^2}\right| |x-y| < |x-y|\,,$$

d.h. f ist kontraktiv, aber nicht strikt kontraktiv. Und in der Tat hat f in $\mathbb{R}$ auch keinen Fixpunkt.

3. Für die oben definierte iterative Folge gilt für alle $n \in \mathbb{N}$, wenn x der Fixpunkt ist:

$$\begin{aligned} d(x_n, x) &\le \frac{k^n}{1-k} d(x_0, x_1), \\ d(x_{n+1}, x) &\le \frac{k}{1-k} d(x_{n+1}, x_n), \\ d(x_{n+1}, x) &\le k d(x_n, x). \end{aligned}$$

Die erste Ungleichung nennt man *apriori* Fehlerabschätzung und die zweite eine *aposteriori* Fehlerabschätzung. Die letzte Ungleichung zeigt, dass die Methode eine lineare Konvergenzgeschwindigkeit besitzt.

Beweis. Zur apriori-Abschätzung. Aus (1.14) und (1.15) folgt

$$\begin{aligned} d(x_n, x_{n+l}) &\le \sum_{j=n}^{n+l-1} d(x_j, x_{j+1}) \le \sum_{j=n}^{n+l-1} k^n d(x_0, x_1) = \sum_{j=0}^{l-1} k^{j+n} d(x_0, x_1) \\ &= k^n \sum_{j=0}^{l-1} k^j d(x_0, x_1) \le \frac{k^n}{1-k} d(x_0, x_1). \end{aligned}$$

Für $l \to \infty$ folgt daraus dann die behauptete Ungleichung.

Zur aposteriorie-Abschätzung. Wir setzen $y_0 := x_n$. Dann ist

$$y_1 = f(y_0) = f(x_n) = x_{n+1}.$$

Mit der apriori-Abschätzung folgt schließlich die behauptete Ungleichung

$$d(x_{n+1}, x) = d(y_1, x) \le \frac{k}{1-k} d(y_1, y_0) = \frac{k}{1-k} d(x_{n+1}, x_n).$$

Zur letzen Ungleichung. Diese folgt aus

$$d(x_{n+1}, x_{l+1}) = d(f(x_n), f(x_l)) \le k d(x_n, x_l)$$

nach Grenzübergang $l \to \infty$. □

Korollar 1.1. *Sei P eine metrischer Raum, der so genannte Parameterraum. Vorgegeben sei das von $p \in P$ abhängige Fixpunktproblem*

$$f_p x_p = x_p, \quad x_p \in A, \ p \in P, \tag{1.16}$$

wobei die Abbildungen $f_p : A \subset X \to X$ für alle $p \in P$ die Voraussetzungen des Banachschen Fixpunktsatzes erfüllen und k von p abhängig ist. Weiter existiere ein $p_0 \in P$, so dass für alle $x \in M$ gilt:

$$f_p x \to f_{p_0} x \quad \text{in } X \text{ für } p \to p_0.$$

Dann existiert für alle $p \in P$ eine eindeutige Lösung x_p von (1.16) *und es gilt*

$$x_p \to x_{p_0} \quad \text{in } X \text{ für } p \to p_0.$$

Beweis. Sei also x_p die eindeutige Lösung des Problems (1.16), die ja aufgrund des Banachschen Fixpunktsatzes exitiert. Für dieses x_p gilt

$$\begin{aligned} d(x_p, x_{p_0}) &= d(f_p(x_p), f_{p_0}(x_{p_0})) \leq d(f_p(x_p), f_p(x_{p_0})) + d(f_p(x_{p_0}), f_{p_0}(x_{p_0})) \\ &\leq k d(x_p, x_{p_0}) + d(f_p(x_{p_0}), f_{p_0}(x_{p_0})). \end{aligned}$$

Daraus folgt nach Voraussetzung

$$d(x_p, x_{p_0}) \leq \frac{1}{1-k} d(f_p(x_{p_0}), f_{p_0}(x_{p_0})) \to 0 \quad \text{für } p \to p_0.$$

Also konvergiert x_p gegen x_{p_0}. □

1.6 Dichtheit und Separabilität

Definition 1.12 (Dichtheit, Separabilität).

(i) Eine Teilmenge D eines metrischen Raumes (M, d) heißt *dicht in* M, wenn $\bar{D} = M$ oder anders ausgedrückt

$$\forall_{m \in M} \forall_{\epsilon > 0} \exists_{d \in D} : d(m, d) < \epsilon.$$

(ii) Ein metrischer Raum heißt *separabel*, wenn er eine abzählbare dichte Teilmenge besitzt.

Satz 1.4. *Ist (M, d) separabel und $A \subset M$, so ist auch (A, d_A) separabel.*

Beweis. Seien also (M, d) separabel und $A \subset M$. Dann gibt es eine abzählbare Teilmenge D von M, etwa

$$D := \{ x_k \mid k \in \mathbb{N} \}.$$

Wähle zu jedem j und k ein $a_{jk} \in B(x_k, 1/j) \cap A$, sofern diese Menge nicht leer ist. Dann ist

$$Z := \{ a_{jk} \mid j, k \in \mathbb{N} \}$$

abzählbar. Zu jedem $a \in A \subset X$ gibt es ein k mit $d(a, x_k) < 1/j$ und dazu ein j mit $d(x_k, a_{jk}) < 1/j$. Zusammenfassend folgt

$$d(a, a_{jk}) \leq d(a, x_k) + d(x_k, a_{jk}) < \tfrac{2}{j}.$$

d.h. $\bar{Z} = A$. □

Beispiel 1.18. Der $\mathbb{R}^N$, $N \in \mathbb{N}$, ist separabel, denn $\mathbb{Q}^N \subset \mathbb{R}^N$ ist abzählbar und liegt dicht in $\mathbb{R}^N$. Nach Satz 1.4 ist auch jede Telmenge $T \subset \mathbb{R}^N$ separabel.

□

Satz 1.5. *Für einen normierten Raum E sind äquivalent:*

(i) E ist separabel.

(ii) Es gibt eine abzählbare Menge A mit $E = \overline{\mathrm{lin}(A)}$.

Beweis. Zu „$(i) \Longrightarrow (ii)$“: Trivial, denn $E = \bar{A}$ impliziert $E = \overline{\mathrm{lin}(A)}$, weil

$$A \subset \mathrm{lin}(A) \subset E \Longrightarrow E = \bar{A} \subset \overline{\mathrm{lin}(A)} \subset \bar{E} = E.$$

Zu „$(ii) \Longrightarrow (i)$“: Sei also $A \subset E$ abzählbar mit $\overline{\mathrm{lin}(A)} = E$. Definiere

$$Q := \begin{cases} \mathbb{Q} & , \text{ für } \mathbb{K} = \mathbb{R}, \\ \mathbb{Q} + i\mathbb{Q} & , \text{ für } \mathbb{K} = \mathbb{C}. \end{cases}$$

Setze

$$B := \{ \sum_{i=1}^{n} \lambda_i x_i \mid n \in \mathbb{N},\ \lambda_i \in Q,\ x_i \in A \}.$$

Dann ist B abzählbar. Wir werden $\bar{B} = E$ zeigen oder anders ausgedrückt

$$\forall_{x\in E}\forall_{\epsilon>0}\exists_{y\in B} : \|x-y\| < \epsilon.$$

Seien also $x \in E$ und $\epsilon > 0$. Wähle dann $y_0 \in \mathrm{lin}(A)$, also etwa

$$y_0 = \sum_{i=1}^{n} \lambda_i x_i, \quad \lambda_i \in \mathbb{R}, x_i \in A,$$

so dass

$$\|x - y_0\| < \epsilon/2.$$

Wähle dann $\lambda_i' \in Q$ mit

$$|\lambda_i - \lambda_i'| < \frac{\epsilon}{2\sum_{i=1}^{n}\|x_i\|}.$$

Dann gilt für

$$y = \sum_{i=1}^{n} \lambda_i' x_i \in B$$

gerade

$$\|x-y\| \leq \|x-y_0\| + \|y_0 - y\| < \epsilon/2 + \max_i |\lambda_i - \lambda_i'| \sum_{i=1}^{n} \|x_i\| < \epsilon.$$

□

Satz 1.6 (Weierstraßscher Approximationssatz). *Der Unterraum $P[a,b]$ der Polynomfunktionen auf $[a,b]$, $a,b \in \mathbb{R}$, liegt dicht in $(C[a,b], \|.\|_\infty)$.*

Beweis. O.E. sei $a = 0$ und $b = 1$. Andernfalls betrachtet man $g(t) = f(a + t(b-a))$. Ferner reicht es, sich mit reellwertigen Funktionen zu begnügen.

Sei $f \in C[0,1]$ belibig. Betrachte das n-te Bersteinpolynom

$$p_n(s) := B_n(s,f) := \sum_{i=0}^{n} \binom{n}{i} s^i (1-s)^{n-i} f(\frac{i}{n}).$$

Beachte, dass die p_n tatsächlich Polynome sind. Außerdem ist $B_n(s,f)$ linear und monoton in f, d.h.

$$\forall_{x\in[0,1]} : f(x) \leq g(x) \Longrightarrow \forall_{s\in[0,1]} B_n(s,f) \leq B_n(s,g), \tag{1.17}$$

$$\forall_{s\in[0,1]}\forall_{\alpha,\beta\in\mathbb{R}}\forall_{f,g\in C[0,1]} : B_n(s, \alpha f + \beta g) = \alpha B_n(s,f) + \beta B_n(s,g). \tag{1.18}$$

Wir müssen zeigen, dass

$$\forall_{f\in C[0,1]}\exists_{\text{Polynome } P_n} : \|P_n - f\|_\infty \to 0.$$

Die gesuchten Polynome P_n sind gerade die p_n, also müssen wir zeigen, dass

$$\|p_n - f\|_\infty \to 0.$$

Weil f gleichmäßig stetig auf $[0,1]$ ist, gibt es zu $\epsilon > 0$ ein $\delta > 0$ mit

$$|s-t| < \sqrt{\delta} \Longrightarrow |f(s) - f(t)| < \epsilon. \tag{1.19}$$

Daraus ergibt sich mit

$$\alpha := \frac{2\|f\|_\infty}{\delta}$$

die Ungleichung

$$|f(s) - f(t)| < \epsilon + \alpha(t-s)^2, \quad \text{für alle } s,t \in [0,1], \tag{1.20}$$

denn

- Ist $|s-t| < \sqrt{\delta}$, so folgt (1.20) aus (1.19).
- Ist $|s-t| \geq \sqrt{\delta}$, so ergibt sich (1.20) aus

$$\epsilon + \alpha(t-s)^2 \geq \epsilon + \alpha\|f\|_\infty > |f(s)| + |f(t)| \geq |f(s) - f(t)|\,.$$

Setze jetzt

$$g_t(s) := (t-s)^2.$$

Dann kann man (1.20) auch schreiben als

$$-\epsilon - \alpha g_t < f - f(t) < \epsilon + \alpha g_t \quad \text{für alle } t \in [0,1]. \tag{1.21}$$

Wir berechnen jetzt $B_n(.,x_j)$ für $x_j(s) = s^j$, $j = 0,1,2$, mit Hilfe des binomischen Satzes:

$$B_n(s,x_0) = \sum_{i=1}^{n} \binom{n}{i} s^i (1-s)^{n-i} = (s + (1-s))^n = 1.$$

$$\begin{aligned}
B_n(s,x_1) &= \sum_{i=0}^{n} \binom{n}{i} s^i (1-s)^{n-i} \frac{i}{n} = \sum_{i=0}^{n} \binom{n}{i} \frac{i}{n} s^i (1-s)^{n-i} \\
&= \sum_{i=0}^{n} \binom{n-1}{i-1} s^i (1-s)^{n-i} = \sum_{i=1}^{n} \binom{n-1}{i-1} s^i (1-s)^{n-i} \\
&= \sum_{i=0}^{n-1} \binom{n-1}{i} s^{i+1} (1-s)^{n-(i+1)} = s \sum_{i=0}^{n-1} \binom{n-1}{i} s^i (1-s)^{(n-1)-i} \\
&= s(s + (1-s))^{n-1} = s.
\end{aligned}$$

$$\begin{aligned}
B_n(s,x_2) &= \sum_{i=0}^{n} \binom{n}{i} s^i (1-s)^{n-i} \left(\frac{i}{n}\right)^2 \overset{\text{wie oben}}{=} \sum_{i=0}^{n-1} \binom{n-1}{i} s^{i+1} (1-s)^{n-(i+1)} \frac{i+1}{n} \\
&= \frac{1}{n} \sum_{i=0}^{n-1} \binom{n-1}{i} s^{i+1} (1-s)^{n-(i+1)} + \frac{s}{n} \sum_{i=0}^{n-1} \binom{n-1}{i} i s^i (1-s)^{n-(i+1)} \\
&= \frac{B_n(s,x_1)}{n} + \frac{s}{n} \sum_{i=1}^{n-1} \binom{n-1}{i-1} s^i (1-s)^{n-(i+1)} \\
&= \frac{s}{n} + \frac{s}{n} \sum_{i=0}^{n-2} \binom{n-1}{i} s^{i+1} (1-s)^{(n-2)-i} \\
&= \frac{s}{n} + \frac{s^2}{n}(n-1) \sum_{i=0}^{n-2} \binom{n-2}{i} s^i (1-s)^{(n-2)-i} \\
&= \frac{s}{n} + \frac{s^2}{n}(n-1)(s + (1-s))^{n-2} = \frac{s(1-s)}{n} + s^2.
\end{aligned}$$

Wir bilden jetzt die Bernsteinpolynome aus den Funktionen aus (1.21). Wegen der Monotonie (1.17) und der Linearität (1.18) ist dann

$$-B_n(.,\epsilon + \alpha g_t) = B_n(.,-\epsilon - \alpha g_t) < B_n(.,f - f(t)) < B_n(.,\epsilon + \alpha g_t).$$

Wegen

$$B_n(.,f - f(t)) = B_n(.,f) - f(t)\underbrace{B_n(.,1)}_{=1} = B_n(.,f) - f(t).$$

ist also

$$B_n(.,-\epsilon - \alpha g_t) < B_n(.,f) - f(t) < B_n(.,\epsilon + \alpha g_t).$$

Daraus folgt dann für alle $s, t \in [0,1]$

$$\begin{aligned}
|p_n(s) - f(t)| &< B_n(s, \epsilon + \alpha g_t) = \epsilon B_n(s,1) + \alpha B_n(s,(t-s)^2) \\
&= \epsilon + \alpha(B_n(s,t^2) + B_n(s,-2ts) + B_n(s,s^2)) \\
&= \epsilon + \alpha t^2 B_n(s,1) - 2t\alpha B_n(s,s) + \alpha B_n(s,s^2) \\
&= \epsilon + \alpha t^2 - 2t\alpha s + \alpha\left(\frac{s(1-s)}{n} + s^2\right).
\end{aligned}$$

Für $s = t$ ergibt sich somit für alle $t \in [0,1]$

$$\begin{aligned}
|p_n(s) - f(t)| &< \epsilon + \alpha t^2 - 2\alpha t^2 + \alpha\frac{t(1-t)}{n} + \alpha t^2 \\
&= \epsilon + \alpha\frac{t(1-t)}{n} \leq \epsilon + \frac{\alpha}{n}
\end{aligned}$$

Schließlich folgt

$$\|p_n - f\|_\infty \to 0.$$

□

Beispiel 1.19. $(C[a,b], \|.\|_\infty)$ ist separabel.

Zum Nachweis davon, setze

$$D := \operatorname{lin}\{\, t^n \mid n \in \mathbb{N}\,\}.$$

Dann ist $\{\, t^n \mid n \in \mathbb{N}\,\}$ abzählbar und nach dem Weierstraßschen Approximationssatz ist $\bar{D} = C[a,b]$. Mit Satz 1.5 folgt damit die Separabilität.

□

Beispiel 1.20. $(l^p(\mathbb{N}), \|.\|_p)$ ist separabel für $1 \leq p < \infty$.

Sei nämlich e_n der n-te Einheitsvektor

$$e_n := (0, \ldots, 0, 1, 0, \ldots).$$

Weiter sei

$$A := \{\, e_n \mid n \in \mathbb{N}\,\}.$$

Dann ist

$$l^p = \overline{\operatorname{lin}(A)}^{\|.\|_p}.$$

Für $x = (x_n)_n \in l^p$ gilt nämlich

$$\left\| x - \sum_{i=1}^{n} x_i e_i \right\|_p = \|(0, \ldots, 0, x_{n+1}, \ldots)\|_p = \left(\sum_{i=n+1}^{\infty} |x_i| \right)^{1/p} \to 0$$

Bemerke noch, dass man genauso die Separabilität von c_0 zeigen kann.

□

Beispiel 1.21. $(l^\infty(\mathbb{N}), \|.\|_\infty)$ ist nicht separabel.

Wir wissen nämlich, dass $\mathcal{P}(\mathbb{N})$ überabzählbar ist. Betrachte zu jeder Menge $D \in \mathcal{P}(\mathbb{N})$ das Element $\chi_D \in l^\infty(\mathbb{N})$ definiert durch

$$\chi_D(n) := \begin{cases} 1 & , \text{ falls } n \in D, \\ 0 & , \text{ sonst.} \end{cases}$$

Dann gilt für alle $D, C \in \mathcal{P}(\mathbb{N})$ mit $D \neq C$

$$\|\chi_D - \chi_C\|_\infty = \sup_{n \in \mathbb{N}} |\chi_D(n) - \chi_C(n)| = 1. \tag{1.22}$$

Angenommen, $l^\infty(\mathbb{N})$ ist separabel, d.h. es gibt ein <u>abzählbares</u> $A \subset l^\infty(\mathbb{N})$ mit $\bar{A} = l^\infty(\mathbb{N})$. Da A dicht in $l^\infty(\mathbb{N})$ liegt, existiert zu jedem χ_D, $D \in \mathcal{P}(\mathbb{N})$, ein $a_D \in A$ mit

$$\|\chi_D - a_D\|_\infty < 1/4. \tag{1.23}$$

Dann folgt für $C, D \in \mathcal{P}(\mathbb{N})$ mit $D \neq C$

$$\begin{aligned}\|a_D - a_C\|_\infty &= \|a_D - \chi_C + \chi_C - \chi_D + \chi_D - a_C\|_\infty \\ &\geq \underbrace{\|\chi_D - \chi_C\|_\infty}_{\overset{(1.22)}{=} 1} - \underbrace{\|a_D - \chi_D\|_\infty}_{\overset{(1.23)}{<} 1/4} - \underbrace{\|\chi_C - a_C\|_\infty}_{\overset{(1.23)}{<} 1/4} > 1/2.\end{aligned}$$

Also gilt $a_D \neq a_C$, d.h. die Abbildung

$$\mathcal{P}(\mathbb{N}) \to A, \quad D \mapsto a_D$$

ist injektiv. Dann ist aber A überabzählbar. Widerspruch!

Ähnlich kann man zeigen, dass $L^\infty[0,1]$ nicht separabel ist.

$\square$

1.7 Kompaktheit

Definition 1.13 ((Relative) Kompaktheit, (relative) Folgenkompaktheit, Präkompaktheit, Häufungspunkt). Sei (M, d) ein metrischer Raum.

(i) $A \subset M$ heißt *kompakt*, wenn jedes System offener Mengen, das A überdeckt, eine endliche Teilüberdeckung enthält.

(ii) $A \subset M$ heißt *relativ kompakt*, wenn $\bar{A}$ kompakt ist.

(iii) $A \subset M$ heißt *präkompakt* oder *totalbeschränkt*, wenn es zu jedem $\epsilon > 0$ endlich viele offene Kugeln vom Radius ϵ gibt, die A überdecken.

(iv) $A \subset M$ heißt *folgenkompakt*, wenn jede Folge in A eine Teilfolge besitzt, die gegen ein $x \in A$ konvergiert.

(v) $A \subset M$ heißt *relativ folgenkompakt*, falls jede Folge in A eine in A konvergente Teilfolge enthält.

(vi) $x \in M$ heißt *Häufungspunkt* einer Folge $(x_n)_n \subset M$, falls gilt:

$$\forall_{\epsilon>0} \forall_{n\in\mathbb{N}} \exists_{m>n} : x_m \in B(x, \epsilon),$$

d.h. in jeder $\epsilon-$Kugel um x liegen unendlich viele Folgenglieder[2].

Bemerkung. A ist folgenkompakt genau dann, wenn jede Folge in A einen Häufungspunkt besitzt.

Satz 1.7. *Eine kompakte Teilmenge eines metrischen Raumes ist abgeschlossen und beschränkt.*

Beweis. Sei $A \subset M$ kompakt. Zum Beweis müssen wir geeignet offene Überdeckungen von A konstruieren und ausnutzen, dass sie endliche Teilüberdeckungen besitzen.

Zur Beschränktheit. Ist $M = \emptyset$, so ist nichts zu zeigen. Andernfalls sei $x \in M$ und für $k \in \mathbb{N}$ sei

$$U_k := B(x, k+1).$$

Jedes $a \in A$ liegt dann in U_k, sobald $k + 1 > d(a, x)$. Also bildet $(U_n)_{n\in\mathbb{N}}$ eine offene Überdeckung von A und wegen der Kompaktheit von A gibt es ein $n \in \mathbb{N}$ mit

$$A \subset \bigcup_{k=0}^{n} U_k = U_n.$$

Daher ist A beschränkt mit einem Durchmesser $\leq 2(n + 1)$.

Zur Abgeschlossenheit. Sei $x \in M \setminus A$. Zu jedem $a \in A$ sei U_a eine offene Kugel um a mit Radius $\frac{1}{2}d(a, x)$. Offensichtlich bildet $(U_a)_{a\in A}$ eine offene Überdeckung von A und nach Voraussetzung gibt es also ein $n \in \mathbb{N}$ mit $a_0, \ldots, a_n$, so dass

$$A \subset \bigcup_{k=0}^{n} U_{a_k}.$$

Setze

$$\epsilon := \frac{1}{2} \min\{\, d(x, a_k) \mid 0 \leq k \leq n \,\}.$$

[2]Beachten Sie: Beim Grenzwert müssen wirklich alle Folgenglieder ab einem bestimmten $N \in \mathbb{N}$ in der Umgebung liegen

Dann ist $\epsilon > 0$ und für alle $k \in \{0, \ldots, n\}$ gilt

$$B(x, \epsilon) \cap U_{a_k} = \emptyset. \tag{1.24}$$

Daher ist $B(x, \epsilon) \subset M \setminus A$, also $M \setminus A$ offen und A abgeschlossen.

Um (1.24) einzusehen, seien $y \in B(x, \epsilon)$ und $z \in U_{a_k}$ beliebig. Dann gilt

$$\begin{aligned} d(y, z) &\geq d(y, a_k) - d(a_k, z) > d(y, a_k) - \frac{1}{2} d(a_k, x) \geq d(x, a_k) - d(y, x) - \frac{1}{2} d(a_k, x) \\ &= \frac{1}{2} d(a_k, x) - d(y, x) > \frac{1}{2} d(a_k, x) - \epsilon \geq 0. \end{aligned}$$

Also ist $x \neq z$ und damit (1.24) bewiesen. □

Satz 1.8. *Eine abgeschlossene Teilmenge einer kompakten Teilmenge ist kompakt.*

Beweis. Seien $A \subset M$ kompakt und $B \subset A$ eine abgeschlossene Teilmenge. Sei $(U_i)_{i \in I}$ eine offene Überdeckung von B. Wir suchen eine endliche Teilüberdeckung.

Durch Hinzunahme der offenen Menge $U := X \setminus B$ erhält man eine offene Überdeckung von A. Weil A kompakt ist, gibt es eine endliche Teilmenge $K \subset I$ mit

$$A \subset U \cup \bigcup_{k \in K} U_k.$$

Wegen $B \cap U = \emptyset$ ist dann aber

$$B \subset \bigcup_{k \in K} U_k,$$

und wir haben für B eine endliche Teilüberdeckung $(U_i)_{i \in I}$ gefunden. □

Satz 1.9 (**Heine-Borel**). *Eine Teilmenge A des $\mathbb{R}^n$ mit der Standardmetrik ist kompakt genau dann, wenn sie abgeschlossen und beschränkt ist.*

Beweis. Zu „$\Longrightarrow$“. Trivial nach Satz 1.7.

Zu „$\Longleftarrow$“. Sei A beschränkt und abgeschlossen. Dann ist A in einem genügend großen abgeschlossenen Quader Q enthalten. Dieser ist nach Analysis II kompakt. Nach Satz 1.8 ist dann $A \subset Q$ kompakt. □

Bemerkung. In allgemeinen metrischen Räumen bzw. in ∞-dimensionalen normierten Vektorräumen gilt der Satz von Heine-Borel leider nicht mehr. Wir betrachten dazu das folgende

Beispiel 1.22. Sei $(\mathbb{R}, d)$ versehen mit der diskreten Metrik

$$d(x, y) = \begin{cases} 1 & \text{für } x \neq y, \\ 0 & \text{sonst.} \end{cases}$$

Dann ist $A = \mathbb{R}$ abgeschlossen und beschränkt. A ist aber nicht kompakt, denn die Familie $(U_x)_{x \in A}$ mit

$$U_x := B(x, 1/2) = \{x\}$$

ist eine offene Überdeckung von A, die offensichtlich keine echte Teilüberdeckung von A besitzt - und schon gar keine endliche.

□

Beispiel 1.23. Betrachte $(l^\infty, \|.\|_\infty)$ und die Teilmenge

$$A := \{\, e_n \mid n \in \mathbb{N} \,\} \subset l^\infty.$$

Dann ist A beschränkt, denn

$$\|e_1 - e_n\|_\infty \leq 2 \quad \text{für alle } e_n \in A.$$

Weiter ist A abgeschlossen.

A ist also abgeschlossen und beschränkt, aber nicht(!) kompakt, weil die Überdeckung

$$A \subset \bigcup_{n\in\mathbb{N}} B(e_,1/2) = \bigcup_{n\in\mathbb{N}} \{e_n\}$$

keine endliche Teilüberdeckung besitzt.

□

Satz 1.10. *Sei (M, d) ein metrischer Raum. Dann sind äquivalent:*

(i) M ist kompakt.

(ii) M ist folgenkompakt.

(iii) M ist vollständig und präkompakt.

Beweis. Zu „$(i) \Longrightarrow (ii)$“: Angenommen, M ist nicht folgenkompakt.

Dann gibt es eine Folge $(x_n)_n \subset M$, die keinen Häufungspunkt x besitzt, d.h.

$$\forall_{x\in M} \exists_{\epsilon_x > 0} : B(x, \epsilon_x) \text{ enthält nur endlich viele Folgenglieder von } (x_n)_n.$$

Es gilt

$$M \subset \bigcup_{x\in M} B(x, \epsilon_x).$$

Da M kompakt ist, gibt es eine endliche Menge $K \subset M$ mit

$$M \subset \bigcup_{x\in K} B(x, \epsilon_x).$$

Also enthielte M nur endlich viele der x_n. Widerspruch!

Zu „$(ii) \Longrightarrow (iii)$“:

a) Zur Vollständigkeit: Sei $(x_n)_n$ eine Cauchy-Folge in M. Da M folgenkompakt ist, besitzt (x_n) eine konvergente Teilfolge $(x_{n_k})_k$. Also ist auch (x_n) konvergent und (M, d) vollständig.

b) Angenommen, M ist nicht präkompakt.

   Dann gibt es ein $\epsilon > 0$, so dass M nicht durch endlich viele Kugeln mit Radius ϵ überdeckt werden kann:

$$\exists_{\epsilon>0} \forall_{n\in\mathbb{N}} \forall_{t_1,\ldots,t_n\in M} : \bigcup_{k=1}^{n} B(t_k, \epsilon) \subsetneq M.$$

   Wir werden nun induktiv eine Folge ohne konvergente Teilfolge konstruieren. Sei dazu $t_1 \in M$ beliebig. Wegen

$$B(t_1, \epsilon) \neq M$$

gibt es $t_2 \in M$ mit

$$d(t_1, t_2) \geq \epsilon.$$

Nun ist auch

$$B(t_1, \epsilon) \cup B(t_2, \epsilon) \neq M,$$

also existiert ein $t_3 \in M$ mit

$$d(t_3, t_k) \geq \epsilon \text{ für } k = 1, 2.$$

So fortfahrend, erhält man eine Folge $(t_n)_n$ in M, für die

$$d(t_n, t_k) \geq \epsilon \quad \text{für alle } k < n.$$

Es ist klar, dass keine Teilfolge von (t_n) eine Cauchy-Folge sein kann. Daher enthält (t_n) erst recht keine konvergente Folge. Widerspruch!

<u>Zu „$(iii) \Longrightarrow (i)$“:</u> Angenommen, (M, d) ist nicht kompakt.

Dann gibt es eine offene Überdeckung

$$M \subset \bigcup_{i \in I} O_i,$$

die keine endliche Teilüberdeckung von M enthält.

Da M präkompakt ist, existiert zu $\epsilon_1 = 1$ eine endliche Überdeckung von M durch Kugeln mit Radius 1, d.h.

$$M \subset \bigcup_{i=1}^{n_1} B(x_i^{(1)}, \epsilon_1).$$

Aus unserer Annahme folgt, dass mindestens einer dieser Kugeln $B(x_i^{(1)}, \epsilon_1)$ nicht durch endlich viele der Mengen $(O_i)_{i \in I}$ überdeckt werden kann. O.B.d.A. sei das die Kugel $B(x_1^{(1)} \epsilon_1)$.

Nun sei $\epsilon_2 = 1/2$ und seien $(x_1^{(2)}, \ldots, x_{n_2}^{(2)})$ gemäß der Präkompaktheit von M zu ϵ_2 gewählt. Es folgt

$$B(x_1^{(1)}, \epsilon_1) = \bigcup_{k=1}^{n_2} B(x_1^{(1)}, \epsilon_1) \cap B(x_k^{(2)}, \epsilon_2)$$

und eine dieser Mengen, sagen wir $B(x_1^{(1)}, \epsilon_1) \cap B(x_1^{(2)}, \epsilon_2)$, kann nicht endlich überdeckbar durch die O_i sein.

Wir führen jetzt das Verfahren fort mit $\epsilon_3 = 1/4, \epsilon_4 = 1/8, \ldots, \epsilon_n = 2^{-n+1}$ und konstruieren so Punkte s_n, so dass

$$\bigcap_{k=1}^{n} B(s_k, \epsilon_k)$$

für kein $n \in \mathbb{N}$ endlich überdeckbar ist. Insbesondere ist stets

$$B(s_n, \epsilon_n) \cap B(s_{n+1}, \epsilon_{n+1}) \neq \emptyset.$$

Ansonsten würde ja doch eine endliche Überdeckung existieren.

Betrachte nun die entstandene Folge $(s_n)_n$. Dann ist für jedes beliebige $x \in B(s_n, 2^{1-n}) \cap B(s_{n+1}, 2^{-n})$

$$d(s_n, s_{n+1}) \leq d(s_n, x) + d(x, s_{n+1}) \leq 2^{-n+1} + 2^{-n} = 2^{-n}(2+1) = 3 \cdot 2^{-n}.$$

Daher konvergiert $\sum d(s_n, s_{n+1})$ und nach dem Cauchyfolgenkriterium ist $(s_n)_n$ eine Cauchy-Folge.

Weil M vollständig ist, existiert ein $s \in M$ mit

$$s := \lim s_n. \tag{1.25}$$

Da M überdeckt wird von den O_i, gibt es ein $i_0 \in I$, so dass

$$s \in O_{i_0}.$$

Nun ist O_{i_0} offen, d.h. es gibt ein $\epsilon > 0$, so dass

$$B(s, \epsilon) \subset O_{i_0}.$$

Wegen (1.25) gibt es ein $N \in \mathbb{N}$, so dass

$$d(s_n, s) < \epsilon/2 \quad \text{für alle } n \geq N.$$

Außerdem gibt es ein $\tilde{N} \in \mathbb{N}$, so dass

$$2^{1-n} < \epsilon/2 \quad \text{für alle } n \geq \tilde{N}.$$

Für alle $n \geq \max\left\{\tilde{N}, N\right\}$ folgt dann

$$\bigcap_{i=1}^{n} B(s_i, 2^{1-i}) \subset B(s_n, 2^{1-n}) \subset O_{i_0}.$$

In der Tat: Ist $y \in B(s_n, 2^{1-n})$, so gilt

$$d(y, s) \leq d(y, s_n) + d(s_n, s) < \epsilon.$$

Also ist $y \in O_{i_0}$.

Schließlich ist damit

$$\bigcap_{i=1}^{n} B(s_i, 2^{1-i})$$

endlich überdeckbar, nämlich durch ein einziges O_{i_0} im Widerspruch zur Konstruktion der s_n. □

Satz 1.11. *A ist präkompakt genau dann, wenn $\bar{A}$ präkompakt ist.*

Beweis. Zu „$\Longleftarrow$“: Ist $\bar{A}$ präkompakt, so gibt es zu $\epsilon > 0$ endlich viele Punkte $a_i \in \bar{A}$, so dass

$$A \subset \bar{A} \subset \bigcup_i B(a_i, \epsilon/2).$$

Wähle $b_i \in A$ mit

$$d(a_i, b_i) < \epsilon/2.$$

Dann gilt

$$A \subset \bigcup_i B(a_i, \epsilon/2) \subset \bigcup_i B(b_i, \epsilon) \tag{1.26}$$

In der Tat: Ist $s \in \bigcup_i B(a_i, \epsilon/2)$, so gibt es ein i, so dass $s \in B(a_i, \epsilon/2)$. Es folgt

$$d(s, b_i) \leq d(s, a_i) + d(a_i, b_i) < \epsilon,$$

also $s \in \bigcup_i B(b_i, \epsilon)$.

Schließlich heißt (1.26), dass A präkompakt ist.

Zu „$\Longrightarrow$“: Ist $A = \bar{A}$, d.h. $\bar{A} \setminus A = \emptyset$, so ist nichts zu zeigen. Sei also $\bar{A} \setminus A \neq \emptyset$.

Weiter sei A präkompakt. Dann gibt es zu $\epsilon > 0$ endlich viele Punkte $a_i \in A$ mit

$$A \subset \bigcup_i B(a_i, \epsilon/2).$$

Dann folgt

$$\bar{A} \subset \bigcup_i B(a_i, \epsilon). \tag{1.27}$$

In der Tat: Ist $b \in A$, so ist (1.27) klar. Ist $b \in \bar{A} \setminus A$, so gibt es ein $a \in A$ mit

$$d(b, a) < \epsilon/2.$$

Es folgt

$$d(b, a_i) \leq d(b, a) + d(a, a_i) < \epsilon,$$

also $b \in \bigcup_i B(a_i, \epsilon)$.

Schließlich heißt (1.27), dass $\bar{A}$ präkompakt ist. □

Satz 1.12. *Seien (M, d) ein vollständiger metrischer Raum und $A \subset M$ eine Teilmenge. Dann gilt:*

$$A \text{ relativ kompakt} \Longleftrightarrow A \text{ präkompakt.}$$

Beweis. Nach dem letzten Satz ist $\bar{A}$ auch präkompakt und weil M vollständig ist, ist dann nach Satz 1.1 auch $\bar{A}$ vollständig. Nach Satz 1.10 heißt das aber gerade

$$\bar{A} \text{ kompakt } \Longleftrightarrow \bar{A} \text{ präkompakt und vollständig,}$$

was zu zeigen war. □

Satz 1.13. *Jeder kompakte metrische Raum ist separabel.*

Beweis. Seien (M, d) ein kompakter metrischer Raum und $\epsilon > 0$. Dann besitzt die Überdeckung

$$M \subset \bigcup_{s \in M} B(s, \epsilon)$$

eine endliche Teilüberdeckung. Es existieren also zu $n \in \mathbb{N}$ endlich viele $s_1^{(n)}, \ldots, s_{m_n}^{(n)}$ mit

$$M \subset \bigcup_{k=1}^{m_n} B(s_k^{(n)}, 1/n).$$

Dann ist die Menge

$$A := \{\, s_k^{(n)} \mid k = 1, \ldots, m_n,\ n \in \mathbb{N} \,\}$$

offensichtlich abzählbar. A ist aber auch dicht in M. Sei nämlich $\epsilon > 0$ beliebig und $n \in \mathbb{N}$ so gewählt, dass

$$1/n < \epsilon.$$

Für $m \in M$ gibt es ein k, so dass

$$m \in B(s_k^{(n)}, 1/n).$$

Wähle dann $a \in A$ mit $a = s_k^{(n)}$. Dann folgt

$$d(m, s_k^{(n)}) < 1/n < \epsilon,$$

also tatsächlich $\bar{A} = M$. □

Satz 1.14. *Seien (M_1, d_1) und (M_2, d_2) metrische Räume, $f : (M_1, d_2) \to (M_2, d_2)$ stetig und (M_1, d_1) kompakt. Dann gilt*

$$f(M_1) \text{ ist kompakt und } f \text{ ist gleichmäßig stetig.}$$

Beweis. Zur Kompaktheit: Sei also

$$f(M_1) \subset \bigcup_{i \in I} O_i$$

eine beliebige offene Überdekcung in M_2. Dann folgt

$$M_1 \subset \bigcup_{i \in I} f^{-1}(O_i)$$

und weil f stetig ist, ist $f^{-1}(O_i)$ für jedes $i \in I$ offen. Damit haben wir eine offene Überdeckung von M_1 gegeben und weil M_1 kompakt ist, gibt es eine endliche Teilmenge $J \subset I$, so dass

$$M_1 \subset \bigcup_{i \in J} f^{-1}(O_i) \Longrightarrow f(M_1) \subset \bigcup_{i \in J} O_i.$$

Also ist $f(M_1)$ kompakt.

Zur gleichmäßigen Stetigkeit: Sei $\epsilon > 0$. Weil f stetig ist, gibt es zu jedem $x \in A$ ein $\delta_x > 0$ mit

$$f(B(x, \tfrac{1}{2}\delta x) \cap A) \subset B(f(x), \tfrac{1}{2}\epsilon). \tag{1.28}$$

Setze

$$U_x := B(x, \tfrac{1}{2}\delta_x,).$$

Dann ist $(U_x)_{x \in A}$ eine offene Überdeckung von A und wegen der Kompaktheit gibt es $x_1, \ldots, x_n \in A$ mit

$$A \subset U_{x_1} \cup \ldots \cup U_{x_n}.$$

Setze

$$\delta := \min\{\, \tfrac{1}{2}\delta_{x_j} \mid j = 1, \ldots, n \,\}.$$

Wir zeigen, dass dann für alle $x \in A$

$$f(B(x, \delta) \cap A) \subset B(f(x), \epsilon).$$

Seien also $x, y \in A$ und $y \in B(x, \delta)$. Dann gibt es ein $j \in \{1, \ldots, n\}$ mit $x \in U_{x_j}$, also mit

$$d(x, x_j) < \frac{1}{2}\delta_{x_j}.$$

Dann gilt

$$d(y, x_j) \leq d(y, x) + d(x, x_j) < \delta + \tfrac{1}{2}\delta_{x_j} \leq \delta_{x_j}.$$

Daher ist

$$d(f(y), f(x)) \leq d(f(y), f(x_j)) + d(f(x_j), f(x)) \overset{(1.28)}{<} \epsilon/2 + \epsilon/2 = \epsilon.$$

□

1.8 Der Satz von Arzelà Ascoli

Der *Satz von Arzelà Ascoli* ist ein wichtiges Hilfsmittel in der Theorie der Differentialgleichungen, weil man mit ihm z.B. den klassischen Satz von Peano beweisen kann. Dafür reicht im Grunde die erste Version des Satzes, den ich hier vorstellen werde. Arbeitet man jedoch in unendlich dimensionalen Banachräumen, so gilt diese nicht mehr. Dafür gibt es aber dann eine hübsche Verallgemeinerung dieses Satzes.

Satz 1.15 (von Arzelà-Ascoli: Klassische Version). *Sei (K, d) ein kompakter metrischer Raum und sei $A \subset C(K)$, wobei $C(K)$ wie üblich mit der Supremumsnorm versehen wird. Dann ist A genau dann relativ kompakt, wenn gilt*

(i) A ist punktweise beschränkt, d.h.

$$\forall_{x \in K} \exists_{C_x > 0} : |f(x)| \leq C_x \quad \text{für alle } f \in A.$$

(ii) A ist gleichgradig stetig auf K, d.h.

$$\forall_{\epsilon > 0} \exists_{\delta > 0} : |f(x) - f(y)| < \epsilon \text{ für alle } x, y \in K \text{ mit } d(x, y) < \delta \text{ und alle } f \in A.$$

Bemerkungen.

1. Gilt in dem Satz zusätzlich die Bedingung

   (iii) A ist abgeschlossen,

   so folgt, dass A kompakt ist bzw. wenn A kompakt ist, dass dann (i), (ii) und (iii) folgen.

2. Aus der Kompaktheit von K folgt sofort, dass A *gleichmäßig gleichgradig stetig* ist, d.h.
$$\forall_{\epsilon > 0} \exists_{\delta > 0} \forall_{f \in A} : |f(x) - f(y)| < \epsilon \text{ für alle } x, y \in K \text{ mit } d(x, y) < \delta.$$

Beweis des Satzes. Zu „$\Longrightarrow$“.

- A ist punktweise beschränkt:

  Da A relativ kompakt ist, ist $\bar{A}$ kompakt, also $\bar{A}$ beschränkt und wegen $A \subset \bar{A}$ ist dann auch A beschränkt, d.h. für alle $f \in A$ gilt

  $$\sup_{f \in A} \|f\|_\infty = \sup_{f \in A} \max_{x \in K} |f(x)| < \infty.$$

  Somit ist A insbesondere auch punktweise beschränkt, was zu zeigen war.

- A ist (gleichmäßig) gleichgradig stetig:

  A ist relativ kompakt und demnach A präkompakt. Es existieren also zu $\epsilon > 0$ Funktionen $f_1, \ldots, f_n \in A$ mit

  $$A \subset \bigcup_{i=1}^{n} B(f_i, \tfrac{\epsilon}{3}). \tag{1.29}$$

  Alle $f_1, \ldots, f_n$ sind nach Satz 1.14 gleichmäßig stetig (da K kompakt ist). Deshalb existiert ein $\delta > 0$, so dass

  $$|f_i(x) - f_i(y)| < \tfrac{\epsilon}{3} \text{ für alle } x, y \in K \text{ mit } d(x, y) < \delta \quad \forall_{i=1,\ldots,n}. \tag{1.30}$$

  Ist nun $f \in A$ beliebig, so gibt es wegen (1.29) ein $j \in \{1, \ldots, n\}$, so dass

  $$f \in B(f_j, \tfrac{\epsilon}{3}). \tag{1.31}$$

Dann folgt

$$|f(x)-f(y)| \leq \underbrace{|f(x)-f_j(x)|}_{\overset{(1.31)}{<}\frac{\epsilon}{3}} + \underbrace{|f_j(x)-f_i(y)|}_{\overset{(1.30)}{<}\frac{\epsilon}{3}} + \underbrace{|f_j(y)-f(y)|}_{\overset{(1.31)}{<}\frac{\epsilon}{3}} < \epsilon$$

für alle $x, y \in K$ mit $d(x,y) < \delta$, und somit ist A (gleichmäßig) gleichgradig stetig.

<u>Zu „$\Longleftarrow$“:</u> Für den Beweis benutzen wir ein *Diagonalfolgenargument.*

Nach Satz 1.13 ist die kompakte(!) Menge K separabel. Es gibt also eine abzählbare Menge

$$\{r_1, r_2, \ldots\} \subset K$$

mit $\overline{\{r_1, r_2, \ldots\}} = K$.

Wir müssen zeigen, dass A relativ kompakt ist. Das ist aber äquivalent dazu, dass A relativ folgenkompakt ist:

Sei daher $(f_n)_n \subset A$. Wir müssen zeigen, dass $(f_n)_n$ eine gleichmäßig konvergente Teilfolge besitzt (oder äquivalent dazu: eine konvergente Teilfolge bzgl. der Supremumsnorm).

Weil A punktweise beschränkt ist, ist die Folge $(f_n(r_1))_n$ in $\mathbb{K}$ beschränkt.

Nach dem Satz von Bolzano-Weierstraß enthält jede beschränkte Folge in $\mathbb{K}$ eine konvergente Teilfolge. Dies angewendet auf $(f_n(r_1))_n$ liefert eine konvergente Folge

$$f_{n_1^{(1)}}(r_1), f_{n_2^{(1)}}(r_1), f_{n_3^{(1)}}(r_1), \ldots.$$

Auch die Folge $(f_{n_i^{(1)}}(r_2)_i$ ist beschränkt, und eine geeignete Teilfolge dieser Folge, etwa

$$f_{n_1^{(2)}}(r_2), f_{n_2^{(2)}}(r_2), f_{n_3^{(2)}}(r_2), \ldots$$

konvergiert. Nochmalige Ausdünnung beschert uns eine konvergente Folge

$$f_{n_1^{(3)}}(r_3), f_{n_2^{(3)}}(r_3), f_{n_3^{(3)}}(r_3), \ldots, \quad \textit{etc.}$$

Nun wenden wir das Diagonalfolgenprinzip an:
Die Diagonalfolge $(f_{n_k^{(k)}})_{k\in\mathbb{N}}$ hat die Eigenschaft

$$((f_{n_k^{(k)}}(r_i))_{k\in\mathbb{N}}) \text{ konvergiert in } \mathbb{K} \text{ für alle } i \in \mathbb{N}. \tag{1.32}$$

(Z.B. konvergiert die Folge $(f_{n_1^{(1)}}(r_1)), f_{n_2^{(2)}}(r_1), \ldots$, da die $f_{n_k^{(k)}}(r_1)$, $k \geq 2$, alles Teilfolgen von $f_{n_1^{(1)}}(r_1)$ sind.)

Die Folge $f_k := f_{n_k^{(k)}}$, $k \in \mathbb{N}$, konvergiert also auf $\{r_1, r_2, \ldots\}$.

Wir werden nun die gleichgradige Stetigkeit benutzen, um die gleichmäßige Konvergenz von $(f_k)_{k\in\mathbb{N}}$ zu zeigen. Dazu beweisen wir, dass $(f_k)_k$ bzgl. der Supremumsnorm eine Cauchyfolge bildet:

Sei also $\epsilon > 0$ und wähle $\delta > 0$ gemäß der Voraussetzung der gleichmäßigen gleichgradigen Stetigkeit.

Da K kompakt ist, gibt es endlich viele Kugeln B_i, $i = 1, \ldots, N$, mit Radius $\delta/2$, so dass

$$K \subset \bigcup_{i=1}^{N} B_i.$$

Wegen $\overline{\{r_1, r_2, \ldots\}} = K$ gibt es zu jedem B_i ein $s_i \in \{r_1, r_2, \ldots\}$ mit $s_i \in B_i$ und wegen (1.32) gibt es sodann ein $k_0 = k_0(\epsilon) \in \mathbb{N}$, so dass

$$|f_k(s_i) - f_m(s_i)| < \tfrac{\epsilon}{3} \quad \forall_{k,m\geq k_0} \text{ und } i = 1, \ldots, N. \tag{1.33}$$

Sei jetzt $x \in K$ beliebig. Dieses Element liegt dann in einer der Kugeln B_j, sagen wir $x \in B_i$ und es folgt $d(x, s_i) < \delta$. Aufgrund der gleichgradigen Stetigkeit gilt somit

$$|f_k(x) - f_k(s_i)| < \tfrac{\epsilon}{3} \quad \text{für alle } k \in \mathbb{N}. \tag{1.34}$$

Also implizieren (1.33) und (1.34) für alle $k, m \geq k_0$

$$|f_k(x) - f_m(x)| \leq \underbrace{|f_k(x) - f_k(s_i)|}_{\overset{(1.34)}{<} \epsilon/3} + \underbrace{|f_k(s_i) - f_m(s_i)|}_{\overset{(1.33)}{<} \epsilon/3} + \underbrace{|f_m(s_i) - f_m(x)|}_{\overset{(1.34)}{<} \epsilon/3} < \epsilon$$

Das zeigt $\|f_k - f_m\|_\infty < \epsilon$ für $k, m \geq k_0$ und (f_k) ist in der Tat eine Cauchyfolge. Da $(C(K), \|.\|_\infty)$ vollständig ist und $\bar{A} \subset C(K)$ gilt, ist auch $\bar{A}$ vollständig. Also ist die Cauchyfolge (f_k) konvergent und ihr Limes liegt in $\bar{A}$. Damit ist A relativ folgenkompakt und damit auch relativ kompakt. □

Bemerkung. Der Satz von Arzelà-Ascoli gilt auch für $C(K, \mathbb{K}^n)$. Der Satz gilt aber nicht mehr für $C(K, X)$, wenn $(X, \|.\|)$ ein **unendlich**-dimensionaler Banachraum ist. Es gibt aber einen schönen Ersatz dafür. Man muss dafür die Bedingung der punktweisen Beschränktheit von A ersetzen:

Satz 1.16 (**von Arzelà-Ascoli: Allgemeine Version**). *Sei (K, d) ein kompakter metrischer Raum und sei (X, d) ein vollständiger metrischer Raum. Weiter sei $A \subset C(K, X)$. Dann ist A genau dann relativ kompakt, wenn gilt*

(i) Für alle $k \in K$ ist $A(k) := \{ f(k) \mid f \in A \}$ relativ kompakt in X.

(ii) A ist gleichgradig stetig auf A, d.h.

$$\forall_{\epsilon>0} \exists_{\delta>0} : |f(x) - f(y)| < \epsilon \text{ für alle } x, y \in K \text{ mit } d(x, y) < \delta \quad \forall_{f \in A}.$$

Beweis. Siehe Anhang 7.1. □

Beispiel 1.24. Sei

$$f_n : [0, 1] \to \mathbb{R}, \quad x \mapsto n \arctan(\tfrac{x}{n}).$$

Dann ist

$$F := \{ f_n \mid n \in \mathbb{N} \}$$

relativ kompakt.

a) Beschränktheit: Für $x \in [0, 1]$ gilt

$$f_n(x) = n \arctan(\tfrac{x}{n}) = \begin{cases} 0 & , \quad x = 0, \\ x \frac{\arctan(\frac{x}{n})}{x/n} & , \quad x \in]0, 1]. \end{cases}$$

Mit der Regel von l'Hospital folgt dann

$$\lim_{y \to 0} \frac{\arctan(y)}{y} = \lim_{y \to 0} \frac{\frac{1}{1+y^2}}{1} = 1.$$

Daher gilt für alle $x \in [0, 1]$

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

d.h. insbesondere, dass die Zahlenfolge $(f_n(x))_n$ beschränkt ist für alle $x \in [0, 1]$.

Außerdem konvergiert $(f_n)_n$ punktweise auf $[0, 1]$ gegen die identische Funktion

$$f : [0, 1] \to \mathbb{R}, \quad x \mapsto x.$$

b) Gleichgradige Stetigkeit: Zunächst ist f_n differenzierbar auf $[0,1]$ für alle $n \in \mathbb{N}$:

$$f_n'(x) = \frac{1}{1+(\frac{x}{n})^2} \qquad \text{für alle } x \in [0,1] \text{ und alle } n \in \mathbb{N}.$$

Daraus folgt

$$\|f_n'\|_\infty \leq 1,$$

also nach dem Mittelwertsatz

$$|f_n(x) - f_n(y)| \leq |x-y| \qquad \text{für alle } x, y \in [0,1] \text{ und alle } n \in \mathbb{N}.$$

Daher ist F gleichgradig stetig.

Nach dem klassischen Satz von Arzelà-Ascoli (Satz 1.15) ist F also relativ kompakt. Beachte aber, dass F nicht kompakt ist, da $f \notin F$.

$\square$

1.9 Der Satz von Baire

Der *Satz von Baire* ist ein wirklich wichtiger Satz der Funktionalanalysis, da wir ihn noch oft brauchen werden:

Satz 1.17 (von Baire). *Seien (M, d) ein vollständiger, metrischer Raum, $M \neq \emptyset$ und $(A_k)_k \subset M$ eine Familie abgeschlossener Teilmengen in M mit*

$$M = \bigcup_{k=1}^{\infty} A_k.$$

Dann existiert mindestens ein $k_0 \in \mathbb{N}$ mit $A_{k_0}^\circ \neq \emptyset$, d.h. mindestens ein abgeschlossenes A_k enthält eine offene Kugel.

Beweis. <u>Annahme:</u> $A_k^\circ = \emptyset$ für alle $k \in \mathbb{N}$, d.h. kein A_k enthält eine offene Kugel.

Dann gilt die folgende Tatsache:
Für ein beliebiges $k \in \mathbb{N}$ und eine beliebige offene Menge $O \subset M$ mit $O \neq \emptyset$ und $A_k \subset O$ existiert ein $x \in M$ und ein $\epsilon > 0$, so dass

$$\bar{B}(x, \epsilon) := \{\, y \in M \mid d(x, y) \leq \epsilon \,\} \subset O \setminus A_k. \tag{1.35}$$

Hierbei ist wichtig, dass $O \setminus A_k \neq \emptyset$. Das ist nach Annahme auch der Fall, denn wäre $O \setminus A_k = \emptyset$, so enthielte ja A_k eine offene Kugel, also $A_k^\circ \neq \emptyset$, was ein Widerspruch zur Annahme wäre. Also ist tatsächlich $O \setminus A_k \neq \emptyset$(!).

Wende nun (1.35) sukzessiv an:

- Zunächst finden wir ein $x_1 \in M$ und ein $0 < \epsilon_1 < 1$ mit

  $$\bar{B}_1 := \bar{B}(x_1, \epsilon_1) \subset M \setminus A_1 \neq \emptyset.$$

- Weiter existiert nach (1.35) ein $x_2 \in M$ und ein $0 < \epsilon_2 < \frac{1}{2}$, so dass

  $$\bar{B}_2 := \bar{B}(x_2, \epsilon_2) \subset B(x_1, \epsilon_1) \setminus A_2 \neq \emptyset.$$

  Beachten Sie, dass auch hier wieder $B(x_1, \epsilon_1)$ offen ist und dass

  $$\bar{B}_2 \subset B(x_1, \epsilon_1) \setminus A_2 \subset B(x_1, \epsilon_1) \subset \bar{B}(x_1, \epsilon_1) = \bar{B}_1.$$

- Wir setzen das Verfahren induktiv fort: Es gibt also ein $x_k \in M$ und ein $0 < \epsilon_k < 2^{1-k}$, so dass für alle $k \in \mathbb{N}$

  $$\bar{B}_k := \bar{B}(x_k, \epsilon_k) \subset B(x_{k-1}, \epsilon_{k-1}) \setminus A_k \neq \emptyset$$

  und

  $$\bar{B}_k \subset \bar{B}_{k-1} \tag{1.36}$$

Insbesondere gilt wegen (1.36), dass für alle $k \in \mathbb{N}$

$$x_n \in \bar{B}(x_k, \epsilon_k) = \bar{B}_k \quad \text{für alle } n \geq k,$$

also

$$d(x_n, x_k) < \epsilon_k < 2^{1-k} \quad \text{für alle } n \geq k$$

und daher ist $(x_n)_n$ eine Cauchyfolge in M.

Weil M vollständig ist, gibt es ein $x \in M$ mit

$$x := \lim x_n.$$

Weiter gilt

$$x \in \bar{B}(x_k, \epsilon_k) \subset B(x_{k-1}, \epsilon_{k-1}) \setminus A_k \subset M \setminus A_k \quad \text{für alle } k \in \mathbb{N},$$

denn die Folge $(x_n)_{n \geq k}$ ist eine Folge in der abgeschlossenen Menge $\bar{B}_k$ und deshlab liegt ihr Grenzwert in $\bar{B}_k$.

Damit ist

$$x \in M \setminus A_k \quad \forall_{k \in \mathbb{N}}$$

und so

$$x \in \bigcap_{k \in \mathbb{N}} M \setminus A_k = M \setminus \underbrace{\bigcup_{k \in \mathbb{N}} A_k}_{=M} = \emptyset,$$

doch das ist offensichtlich ein Widerspruch. $\square$

1.10 Ergänzung: Eigenschaften von normierten Vektorräumen

Dieser Abschnitt sei der Äquivalenz von Normen gewidmet. Bevor wir aber dazu kommen, vorerst noch ein kleiner

Satz 1.18. *Sei X ein normierter Vektorraum.*

(i) Aus $x_n \to x$ und $y_n \to y$ folgt $x_n + y_n \to x + y$.

(ii) Aus $\lambda_n \to \lambda$ und $x_n \to x$ folgt $\lambda_n x_n \to \lambda x$.

(iii) Aus $x_n \to x$ folgt $\|x_n\| \to \|x\|$.

Beweis. Zu (i): Das ist einfach wegen

$$\|(x_n + y_n) - (x + y)\| \leq \|x_n - x\| + \|y_n - y\| \to 0.$$

Zu (ii): Das ist auch nicht schwer, denn

$$\|\lambda_n x_n - \lambda x\| \leq \|\lambda_n x - \lambda_n x\| + \|\lambda_n x - \lambda x\| = |\lambda_n| \, \|x_n - x\| + |\lambda_n - \lambda| \, \|x_n\| \to 0.$$

Zu (iii): Das ist genauso einfach, weil

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

Insbesondere ist also eine konvergente Folge (x_n) beschränkt, d.h. die Folge der Normen $(\|x_n\|)$ ist beschränkt. □

Korollar 1.2. *Sei U ein Untervektorraum des normierten Raumes X. Dann ist der Abschluss $\bar{U}$ ebenfalls ein Untervektorraum.*

Beweis. Seien also $x, y \in \bar{U}$. Dann existieren $x_n, y_n \in U$ mit $x_n \to x$ und $y_n \to y$. Es folgt

$$U \ni x_n + y_n \to x + y.$$

Also gilt in der Tat $x + y \in \bar{U}$. Für $\lambda \in \mathbb{K}$ konvergiert (λx_n) gegen (λx). Also gilt auch $\lambda x \in \bar{U}$. □

Wir kommen nun zur angekündigten

Definition 1.14 (Äquivalenz zwischen Normen). Zwei Normen $\|.\|$ und $|||\,.\,|||$ auf einem Vektorraum X heißen *äquivalent*, falls es Zahlen $M \geq m > 0$ gibt, so dass

$$m\|x\| \leq |||\, x \,||| \leq M\|x\| \quad \text{für alle } x \in X.$$

Von Vorteil ist der folgende

Satz 1.19. *Seien $\|.\|$ und $|||\,.\,|||$ zwei Normen auf X. Dann sind folgende Aussagen äquivalent:*

(i) $\|.\|$ und $|||\,.\,|||$ sind äquivalent.

(ii) Eine Folge ist bezüglich $\|.\|$ konvergent genau dann, wenn sie es bezüglich $|||\,.\,|||$ ist. Außerdem stimmen die Grnezwerte überein.

(iii) Eine Folge ist $\|.\|$-Nullfolge genau dann, wenn sie eine $|||\,.\,|||$-Nullfolge ist.

Beweis. Die Implikationskette $(i) \Longrightarrow (ii) \Longrightarrow (iii)$ ist trivial.

<u>Zu $(iii) \Longrightarrow (i)$</u>: Angenommen, für kein $M > 0$ gilt die Ungleichung

$$|||\, x \,||| \leq M\|x\| \quad \text{für alle } x \in X.$$

Für jedes $n \in \mathbb{N}$ gibt es dann ein $x_n \in X$ mit $|||\, x_n \,||| > n\|x_n\|$. Setze

$$y_n := \tfrac{x_n}{n\|x_n\|}.$$

Es folgt

$$\|y_n\| = \tfrac{1}{n} \to 0.$$

Also ist (y_n) eine $\|.\|$-Nullfolge. Nun gilt aber

$$|||\, y_n \,||| > \tfrac{n\|x_n\|}{\|x_n\|} = 1 \quad \text{für alle } n \in \mathbb{N}.$$

Demnach ist (y_n) keine $|||\,.\,|||$-Nullfolge. Widerspruch! □

Bemerkung. Der Satz 1.19 sagt, dass äquivalente Normen dieselben Topologien erzeugen. Weiter folgt, dass $(X, \|.\|)$ und $(X, |||\,.\,|||)$ dieselben Cauchyfolgen besitzen. Daher sind die Räume $(X, \|.\|)$ und $(X, |||\,.\,|||)$ entweder beide vollständig oder eben nicht.

Achtung: Bei Metriken sieht das ganze anders aus: Versieht man nämlich T mit zwei Metriken derart, dass (T, d_1) dieselben konvergenten Folgen wie (T, d_2) besitzt, so brauchen diese Metriken nicht dieselben Cauchyfolgen besitzen:

Beispiel 1.25. Betrachte auf $T = \mathbb{R}$ die Metriken

$$d_1(s,t) := |s - t|\,, \quad d_2 = |\arctan s - \arctan t|\,.$$

Dann ist (n) eine d_2-Cauchyfolge, aber eben keine d_1-Cauchyfolge und $(\mathbb{R}, d_2)$ ist auch nicht vollständig im Gegensatz zu $(\mathbb{R}, d_1)$.

□

Die im Beispiel auftauchenden Metriken sind äquivalent in dem Sinne, dass sie dieselben offenen Mengen erzeugen:

> Seien d_1 und d_2 zwei Metriken auf der nicht-leeren Menge X. Dann heißen d_1 und d_2 äquivalent zueinander, wenn es zu jedem $x \in X$ und $r > 0$ ein $\epsilon > 0$ gibt, so dass
>
> $$B_{d_1}(x, \epsilon) \subset B_{d_2}(x, r) \quad \text{und} \quad B_{d_2}(x, \epsilon) \subset B_{d_1}(x, r).$$

Demzufolge ist es auch nicht verwunderlich, dass der Satz 1.19 i.A. nicht für Metriken zu gelten braucht.

Beispiel 1.26. Die Normen

$$\|f\| = \max\{\|f\|_\infty, \|f'\|_\infty\}, \quad f \in C^1[a,b]$$

und

$$\|f\|_\infty = \sup_{t \in [a,b]} |f(t)|\,, \quad f \in C^1[a,b].$$

sind auf $C^1[a,b]$ nicht äquivalent, da $(C^1[a,b], \|.\|_\infty)$ nach Beispiel 1.6 kein Banachraum ist und nach Beispiel 1.14 f) aber $(C^1[a,b], \|.\|)$ schon einen Banachraum darstellt.

□

Beispiel 1.27. Ebensowenig sind $\|.\|_\infty$ wie in Beispiel 1.26 und

$$\|f\|_1 = \int_0^1 |f(t)|\, dt, \quad f \in C^1[0,1]$$

auf $C^1[0,1]$ äquivalent. Es gilt zwar

$$\|f\|_1 \leq \|f\|_\infty \quad \text{für alle } f \in C^1[0,1],$$

aber $f_n(t) = t^n$ definiert eine $\|.\|_1$-Nullfolge, die keine $\|.\|_\infty$-Nullfolge ist. Nach Satz 1.19 sind deshalb beide Normen nicht äquivalent zueinander.

□

Beispiel 1.28 (**Gewichtete Norm**). Betrachte für $\alpha > 0$ die Normen

$$||| f |||_\alpha = \sup_{0 \le t \le 1} |f(t)| \, e^{-\alpha t}.$$

auf $C[0,1]$. Dann sind $\|.\|_\infty$ und $||| \, . \, |||_\alpha$ äquivalent, denn es gilt

$$||| f |||_\alpha \le \|f\|_\infty \le e^\alpha ||| f |||_\alpha \quad \text{für alle } f \in C[0,1].$$

Auf der einen Seite gilt nämlich für alle $f \in C[0,1]$ und $t \in [0,1]$

$$e^{-\alpha t} \, |f(t)| \le |f(t)| \le \|f\|_\infty$$

und auf der anderen Seite

$$|f(t)| = e^{\alpha t} e^{-\alpha t} \, |f(t)| \le e^{\alpha t} ||| f |||_\alpha \le e^\alpha ||| f |||_\alpha.$$

Die $||| \, . \, |||_\alpha$-Norm wird oft in der Theorie der gewönhlichen Differentialgleichungen benutzt, weil mit dieser oft aus einer nichtkontraktiven Abbildung eine Kontraktion wird, was von Vorteil ist, wenn man den Banachschen Fixpunktsatz anwenden möchte.

□

Abschließend kommmen wir nun zu dem nützlichen

Satz 1.20. *Auf einem endlichdimensionalen Raum sind je zwei Normen äquivalent.*

Beweis. Sei etwa

$$\dim X =: N$$

Weiter sei $\{e_1, \ldots, e_N\}$ eine Basis von X sowie $\|.\|$ eine Norm auf X. Wir wollen zeigen, dass $\|.\|$ äquivalent zur euklidischen Norm

$$\left\| \sum_{i=1}^N \alpha_i e_i \right\|_2 = \left(\sum_{i=1}^N |\alpha_i|^2 \right)^{1/2}$$

ist.

Definiere

$$M := \max\{\|e_1\|, \ldots, \|e_N\|\} > 0.$$

Dann folgt

$$\| \sum_{i=1}^N \alpha_i e_i \| \le \sum_{i=1}^N |\alpha_i| \, \|e_i\| \overset{\text{Hölder}}{\le} \left(\sum_{i=1}^N |\alpha_i|^2 \right)^{1/2} \left(\sum_{i=1}^N \|e_i\|^2 \right)^{1/2}.$$

Also gilt

$$\|x\| \le M\sqrt{N} \|x\|_2 \quad \text{für alle } x \in X.$$

Daraus sieht man, dass $\|.\|$ bezüglich $\|.\|_2$ stetig ist, denn aus $\|x_k - x\|_2 \to 0$ folgt

$$|\|x_k\| - \|x\|| \le \|x_k - x\| \le M\sqrt{N} \|x_k - x\|_2 \to 0.$$

Weiter ist die Menge

$$S := \{ \, x \in X \mid \|x\|_2 = 1 \, \}$$

in $(X, \|.\|_2)$ abgeschlossen, weil

$$S = \|.\|_2^{-1}(\{1\})$$

und $\{1\}$ abgeschlossen sowie $\|.\|_2$ stetig ist. Offensichtlich ist auch S beschränkt bezüglich $\|.\|_2$ und zusammenfassend heißt das, dass S kompakt ist nach dem Satz von Heine Borel (Satz 1.9). Daher nimmt die stetige Funktion $\|.\|$ auf S ihr Minimum $m \ge 0$ an. Es gilt sogar $m > 0$, denn wäre $m = 0$, so folgte $x = 0$ im Widerspruch zu $x \in S$. Schließlich folgt

$$m\|x\|_2 \le \|x\| \quad \text{für alle } x \in X,$$

weil $x/\|x\|_2 \in S$ für $x \neq 0$. Damit ist gezeigt, dass jede Norm zu $\|.\|_2$ äquivalent ist und daraus folgt die Behauptung. □

2 Lineare Operatoren und 3 fundamentale Sätze

- Grundbegriffe: Operatornorm, Dualraum, Isometrie, starke Konvergenz und der Fredholmsche Integraloperator
- Invertierbarkeit und Neumannsche Reihe
- Der Satz von Banach-Steinhaus als Folgerung aus dem Satz von Baire
- Offene und abgeschlossene Abbildungen und der Satz von der offenen Abbildung
- Der Graph eines Operators und der Satz vom abgeschlossenen Graphen als Folgerung aus dem Satz über die offene Abbildung
- Wichtige weitere Folgerungen

2.1 Grundbegriffe

Im Folgenden betrachten wir zwei normierte $\mathbb{K}$-Vektorräume$(X, \|.\|_X)$ und $(Y, \|.\|_Y)$ und eine lineare Abbildung $A : X \to Y$, d.h.

$$A(\alpha x + \beta y) = \alpha A(x) + \beta A(y) \quad \forall_{\alpha,\beta\in\mathbb{K}} \forall_{x,y\in X}.$$

Bezeichnung.
Für einen normierten Raum X setze

$$B_X := \{\, x \in X \mid \|x\| \leq 1 \,\},$$
$$S_X := \{\, x \in X \mid \|x\| = 1 \,\}.$$

B_X wird *die abgeschlossene Einheitskugel in* X und S_X die *Einheitssphäre in* X genannt.

Der folgende Satz stellt nützliche Aussagen über Stetigkeit von linearen Abbildungen bereit:

Satz 2.1. *Seien X und Y normierte Räume und sei $A : X \to Y$ linear. Dann sind folgende Aussagen äquivalent:*

(i) A ist stetig.

(ii) A ist stetig in 0.

(iii) Es gibt ein $C \geq 0$, so dass $\|Ax\|_Y \leq C$ für alle $x \in B_X$.

(iv) Es gibt ein $C \geq 0$, so dass $\|Ax\|_Y \leq C\|x\|_X$ für alle $x \in X$.

(v) A ist gleichmäßig stetig.

Beweis. Die Implikationen $(i) \Longrightarrow (ii)$ und $(v) \Longrightarrow (i)$ sind trivial.

Zu $(ii) \Longrightarrow (iii)$: Da A linear ist, gilt $A0 = 0$. Nun ist A stetig in 0, also gibt es zu $\epsilon = 1$ ein $\delta > 0$, so dass

$$\|Ax\| \leq 1 \quad \text{für alle } x \in X \text{ mit } \|x\| \leq \delta. \tag{2.1}$$

Sei nun $x \in B_X$. Dann gilt $\|\delta x\| = \delta\|x\| \leq \delta$ und daher wegen (2.1)

$$\|A\delta x\| < 1 \Longleftrightarrow \|Ax\| \leq \frac{1}{\delta} =: C.$$

Zu $(iii) \Longrightarrow (iv)$: Sei also $0 \neq x \in X$. Dann ist $\frac{x}{\|x\|} \in B_X$. Daher gibt es ein $C \geq 0$, so dass

$$\|A(x/\|x\|)\| \leq C \Longleftrightarrow \|Ax\| \leq C\|x\|.$$

Das gilt natürlich auch für $x = 0$.

Zu $(iv) \Longrightarrow (v)$: Sei also $\epsilon > 0$. Wähle $\delta := \frac{\epsilon}{C}$ mit dem $C > 0$ aus (iv). Beachte, dass für $C = 0$ nichts zu zeigen wäre. Dann gilt für alle $x, y \in X$ mit $\|x - y\| \leq \delta$

$$\|Ax - Ay\| = \|A(x - y)\| \leq C\|x - y\| \leq C\delta = \epsilon.$$

□

Notationen.

- $L(X, Y)$ ist der $\mathbb{K}$-Vektorraum aller stetigen linearen Abbildungen von X nach Y:

$$\boxed{L(X, Y) := \{\, A : X \to Y \mid A \text{ linear und stetig} \,\}}.$$

  Ein Element $A \in L(X, Y)$ heißt auch *stetiger linearer Operator* oder *beschränkter linearer Operator*. Beachten Sie, dass $L(X, Y) \neq \emptyset$, denn stets ist der *Nulloperator* $x \mapsto 0$ in $L(X, Y)$.

- Ist der Bildraum $Y = X$, so schreiben wir auch

$$\boxed{L(X) := L(X, X)}.$$

- Die Menge

$$\boxed{X' := L(X, \mathbb{K})}$$

  heißt der *topologische Dualraum von* X und die Elemente $A \in X'$ heißen *stetige, lineare Funktionale auf* X.

Bemerkungen.

1. Tatsächlich ist $L(X, Y)$ ein Vektorraum. Das liegt daran, weil Summen und skalare Vielfache von Nullfolgen wieder Nullfolgen sind. Also ist $L(X, Y)$ bezüglich der algebraischen Operationen

$$(A + B)(x) = Ax + Bx$$
$$(\lambda A)(x) = \lambda Ax$$

   ein Vektorraum.

2. Beachten Sie, dass man zwischen dem algebraischen und dem topologischen Dualraum unterscheiden muss. Nicht jedes lineare Funktional ist stetig (vgl. das folgende Beispiel 2.1).

Beispiel 2.1. Sei $(X, \|.\|_\infty)$ ein unendlich-dimensionaler normierter Vektorraum. Dann besitzt X eine Hamelbasis, d.h. eine maximale Menge von linear unabhängigen Elementen

$$\{\, x_i \in X \mid i \in I \,\}$$

aus X, so dass jedes Element $x \in X$ eine eindeutige Darstellung als endliche Linearkombination von Elementen x_i besitzt[3].

O.B.d.A. sei $\|x_i\| = 1$ für alle $i \in I$. Weiter sei

$$\{\, x_{ik} \mid k \in \mathbb{N} \,\} \subset \{\, x_i \mid i \in I \,\}$$

[3] Das kann man z.B. mit dem Lemma von Zorn beweisen

eine abzählbare Teilmenge. Definiere die Abbildung

$$A : X \to \mathbb{K}, \quad x = \sum_{i \in I} \alpha_i x_i \mapsto \sum_{k \in \mathbb{N}} \alpha_{i_k} k.$$

Beachten Sie, dass die rechte Summe endlich ist, da nur endlich viele $\alpha_{i_k} \neq 0$ sind. Daher ist A wohldefiniert. A ist auch linear. Insgesamt ist also A ein Element des algebraischen(!) Dualraums von X.

Nun gilt aber

$$|Ax_{i_k}| = k \quad \forall_{k \in \mathbb{N}},$$

denn $x_{i_k} = \alpha_{i_k} x_{i_k}$, also $\alpha_{i_k} = 1$ und alle anderen $\alpha_{i_l} = 0$. Es folgt

$$\sup_{x \in B_X} |Ax| \overset{\|x_{i_k}\|=1}{\geq} \sup_{k \in \mathbb{N}} |Ax_{i_k}| = \sup_{k \in \mathbb{N}} k = +\infty,$$

d.h. aber, dass A <u>nicht</u> stetig ist.

Wir können also notieren

Der algebraische Dualraum ist i.a. wesentlich größer als der topologische Dualraum.

Ist jedoch $\dim(X) < \infty$, so wissen Sie nach *Analysis 2*, dass jede lineare Abbildung $A : X \to Y$ für jeden beliebigen normierten Vektorraum $(Y, \|.\|_Y)$ stetig ist.
Insbesondere stimmt in diesem Fall $X' = L(X, \mathbb{K})$ mit dem algebraischen Dualraum von X überein.

□

Beispiel 2.2. Betrachte $X = C[a, b]$ ausgestattet mit der $\|.\|_\infty$-Norm. Definiere die Abbildung

$$\delta_{x_0} : C[a, b] \to \mathbb{K}, \quad f \mapsto f(x_0).$$

Offensichtlich ist δ_{x_0} linear und es gilt für alle $f \in C[a, b]$

$$|\delta_{x_0}(f)| = |f(x_0)| \leq \sup_{x \in [a,b]} |f(x)| = \|f\|_\infty < \infty,$$

d.h. δ_{x_0} erfüllt (iv) aus Satz 2.1 mit $C = 1$. Damit ist δ_{x_0} stetig, also $\delta_{x_0} \in X'$ ein stetiges lineares Funktional auf X.

□

Beispiel 2.3 (**Fredholmscher Integraloperator**)**.** Betrachte $X = (C[a, b], \|.\|_\infty)$ und $k \in C([a, b] \times [a, b])$. Definiere die Abbildung

$$A : X \to X, \quad u \mapsto \int_a^b k(x, y)u(y)dy.$$

Offensichtlich ist A linear. Aber A ist auch wohldefiniert, d.h. $Au \in X$: Da k gleichmäßig stetig ist, wähle zu $\epsilon > 0$ ein $\delta = \delta(\epsilon) > 0$ mit

$$\|(x, y) - (x', y')\| < \delta \Longrightarrow |k(x, y) - k(x', y')| < \epsilon,$$

wobei $\|.\|$ die euklidische Norm auf $\mathbb{R}^2$ darstellt.

Für $|x - x'| < \delta$ gilt dann

$$|(Au)(x) - (Au)(x')| \leq \int_a^b |k(x, y) - k(x', y)| \, |u(y)| \, dy < \epsilon \int_a^b |u(y)| \, dy \leq \epsilon \|u\|_\infty.$$

Also ist tatsächlich $Au \in X$ und so A wohldefiniert.

Weiter gilt

$$\|Au\|_\infty = \sup_{x\in[a,b]} \left| \int_a^b k(x,y)u(y)dy \right| \leq \underbrace{\sup_{x\in[a,b]} \int_a^b |k(x,y)|\, dy}_{=:C} \|u\|_\infty,$$

d.h. A erfüllt (iv) von Satz 2.1 und daher ist A stetig, also $A \in L(X,Y)$.

Der hier betrachtete Operator heißt *Fredholmscher Integraloperator* und k sein *Kern*.

□

Die letzten Beispiele motivieren die folgende Norm einzuführen:

Operatornorm.
Für $A \in L(X,Y)$ definiert man

$$\boxed{\|A\| = \|A\|_{L(X,Y)} := \inf\{\, C \geq 0 \mid \|Ax\|_Y \leq C \quad \forall_{x\in B_X}\,\}}.$$

und $\|A\|$ heißt die *Operatornorm von A*.

Es gibt noch eine andere Möglichkeit, die Operatornorm auszudrücken, was viel schöner für die Anwendung ist:

$$\boxed{\|A\| = \sup_{x\in B_X} \|Ax\|}.$$

Beweis. Setze $C_0 := \sup_{x\in B_X} \|Ax\|$. Dann gilt offensichtlich für alle $x \in B_X$

$$\|Ax\| \leq C_0 \Longrightarrow \|A\| \leq C_0.$$

Wählt man andererseits zu $\epsilon > 0$ ein $x_\epsilon \neq 0$ mit

$$\|Ax_\epsilon\| \geq C_0\|1-\epsilon\|, \quad x_\epsilon \in B_X,$$

so ergibt sich

$$\|A\| \geq C_0\|1-\epsilon\|.$$

Da $\epsilon > 0$ beliebig war, folgt $\|A\| = C_0$. □

Eine wichtige Tatsache ist die, dass für alle $x \in X$

$$\boxed{\|Ax\| \leq \|A\|\|x\|}.$$

Beweis. Für $x \in B_X$ gilt

$$\|Ax\| \leq \sup_{x\in B_X} \|Ax\| = \|A\|$$

Für beliebiges $x \in X$ mit $x \neq 0$ ist $\frac{x}{\|x\|} \in B_X$, also

$$\left\| A\frac{x}{\|x\|} \right\| \leq \|A\| \Longleftrightarrow \|Ax\| \leq \|A\|\|x\|$$

und die letzte Ungleichung gilt sogar für $x = 0$. □

Weiter halten wir fest:

$$\boxed{(L(X,Y), \|.\|_{L(X,Y)}) \text{ ist ein normierter Vektorraum}}$$

Beweis. Dass $L(X,Y)$ ein Vektorraum ist, haben wir schon oben gesehen. Die ersten beiden Normeigenschaften sind klar. Um die Dreiecksungleichung zu zeigen, sei $\|x\| \leq 1$. Dann gilt

$$\|(A+B)x\| = \|Ax + Bx\| \leq \|Ax\| + \|Bx\| \leq \|A\| + \|B\|.$$

Der Übergang zum Supremum zeigt, dass $\|A+B\| \leq \|A\| + \|B\|$. □

Es gibt noch viel mehr Möglichkeiten für den Ausdruck $\|A\|$:

$$\boxed{\|A\| = \sup_{x \in B_X} \|Ax\| = \sup_{x \in S_X} \|Ax\| = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|}}.$$

Wir werden aber nur die erste Gleichung in diesem Skriptum benutzen. Daher sparen wir uns die Beweise der beiden anderen Gleichheiten. Sehen Sie es einfach als Übungsaufgabe an!

Beispiel 2.4. Wir wollen jetzt die Operatornorm zu Beispiel 2.2 berechnen: Wir wissen schon, dass $|\delta_{x_0}(f)| \leq \|f\|_\infty$, also

$$\|\delta_{x_0}\|_{X'} \leq \sup_{f \in B_X} \|f\|_\infty = 1.$$

Jetzt benutzen wir einen *Standardtrick*, mit dem wir schließlich $\|\delta_{x_0}\|_{X'} = 1$ zeigen werden. Wir betrachten dazu die Funktion $f = 1 \in C[a,b]$. Dann ist $f \in B_X$, da $\|f\|_\infty = 1$ und daher ist

$$\|\delta_{x_0}\|_{X'} = \sup_{f \in B_X} |\delta_{x_0}(f)| \geq |\delta_{x_0}(1)| = 1.$$

Daraus folgt schließlich $\|\delta_{x_0}\| = 1$.

□

Beispiel 2.5. Wir wollen die Operatornorm zu Beispiel 2.3 bestimmen. Wir wissen schon, dass

$$\|A\|_{L(X)} \leq \sup_{x \in [a,b]} \int_a^b |k(x,y)|\, dy.$$

Nun ist

$$\begin{aligned}\|A\|_{L(x)} &= \sup_{f \in B_X} \|Af\|_\infty = \sup_{f \in B_X} \sup_{x \in [a,b]} \left| \int_a^b k(x,y) f(y) dy \right| \\ &= \sup_{x \in [a,b]} \underbrace{\sup_{f \in B_X} \left| \int_a^b k(x,y) f(y) dy \right|}_{=:M}.\end{aligned}$$

Für $x \in [a,b]$ und $\epsilon > 0$ betrachte $f_\epsilon \in B_X$ definiert durch

$$f_\epsilon(y) := \frac{\overline{k(x,y)}}{|k(x,y)| + \epsilon} \quad \text{für alle } y \in [a,b].$$

Dann folgt

$$M \geq \int_a^b \frac{|k(x,y)|^2}{|k(x,y)| + \epsilon} dy \geq \int_a^b \frac{|k(x,y)| - \epsilon^2}{|k(x,y)| + \epsilon} dy = \int_a^b |k(x,y)|\, dy - \epsilon(b-a).$$

Daraus folgt

$$\|A\| \geq \sup_{x \in [a,b]} \int_a^b |k(x,y)|\, dy - \epsilon(b-a) \quad \text{für alle } \epsilon > 0.$$

Da $\epsilon > 0$ beliebig war, folgt schließlich

$$\boxed{\|A\| = \sup_{x \in [a,b]} \int_a^b |k(x,y)|\, dy}. \tag{2.2}$$

□

Beispiel 2.6. In diesem Beispiel möchte ich zeigen, dass c_0 in c abgeschlossen ist: Betrachte dazu die lineare Abbildung

$$A : c \to \mathbb{K}, \quad x = (t_n)_n \mapsto \lim t_n.$$

Es ist $\|A\| = 1$, denn auf der einen Seite hat man

$$|Ax| = |\lim t_n| = \lim |t_n| \leq \sup |t_n| = \|t_n\|_\infty = \|x\|_\infty \Longrightarrow \|A\| \leq 1$$

und auf der anderen Seite gilt

$$\|A\| \geq |A(1)_n| = 1.$$

Also ist A stetig mit $\|A\| = 1$. Weil A stetig und $\{0\}$ abgeschlossen ist, folgt aus

$$c_0 = A^{-1}(\{0\}) = \{\, x \in c \mid Ax = 0 \,\},$$

dass c_0 abgeschlossen in c ist.

□

Satz 2.2. *Falls $(Y, \|.\|_Y)$ ein Banachraum ist, so ist auch $(L(X,Y), \|.\|_{L(X,Y)})$ ein Banachraum. Insbesondere ist $X' = L(X, \mathbb{K})$ ein Banachraum für jeden beliebigen normierten Vektorraum $(X, \|.\|_X)$.*

Beweis. Sei also $(A_n)_n$ eine Cauchyfolge in $L(X,Y)$, d.h.

$$\forall_{\epsilon>0} \exists_{m_0 \in \mathbb{N}} : ||A_n - A_m||_{L(X,Y)} \leq \epsilon \quad \forall_{n,m \geq m_0}.$$

Da für alle $x \in X$ die Abschätzung

$$\|A_n x - A_m x\|_Y \leq \|A_n - A_m\|_{L(X,Y)} \|x\|_X.$$

besteht, gilt insbesondere für jedes feste $x \in X$

$$\|A_n x - A_m x\| \leq \epsilon \quad \forall_{n,m \geq m_0}, \tag{2.3}$$

d.h. $(A_n x)$ ist eine Cauchyfolge in Y.

Weil Y vollständig ist, existiert also $\lim A_n x$ in Y. Wir bezeichnen den Limes als Ax, d.h.

$$Ax := \lim A_n x.$$

Die so definierte Abbildung $A : X \to Y, \quad x \mapsto Ax$ ist linear, denn

$$\begin{aligned} A(\lambda x_1 + \mu x_2) &= \lim A_n(\lambda x_1 + \mu x_2) = \lim(\lambda A_n x_1 + \mu A_n x_2) \\ &= \lambda \lim A_n x_1 + \mu \lim A_n x_2 = \lambda A x_1 + \mu A x_2. \end{aligned}$$

Wir zeigen jetzt $A \in L(X,Y)$ (also $\|A\| < \infty$) und $\|A_n - A\| \to 0$.

Für alle $x \in X$ gilt

$$\|Ax\| = \lim \|A_n x\| \leq \underbrace{\limsup \|A_n\|}_{=:C} \|x\|.$$

Es ist $C < \infty$, da (A_n) eine Cauchyfolge in $L(X,Y)$ ist und Cauchyfolgen beschränkt sind. Es folgt

$$\|Ax\| \leq C\|x\| \quad \text{für alle } x \in X$$

Also ist A beschränkt und damit stetig: $A \in L(X,Y)$.

Es bleibt zu zeigen, dass $\|A_n - A\| \to 0$: Für $n \to \infty$ in (2.3) erhalten wir für alle $x \in B_X$(!) die Abschätzung

$$\|Ax - A_m x\|_Y \leq \epsilon \quad \forall_{m \geq m_0}.$$

Daraus folgt

$$\|A - A_m\|_{L(X,Y)} = \sup_{x \in B_X} \|Ax - A_m x\|_Y \leq \epsilon \quad \forall_{m \geq m_0},$$

d.h. aber $A_n \to A$ im Sinne der Operatornorm in $L(X,Y)$. □

Satz 2.3. *Sind $\|.\|_a$ und $\|.\|_b$ zwei Normen auf dem Vektorraum X, so sind $\|.\|_a$ und $\|.\|_b$ genau dann äquivalent, wenn*

$$Id : (X, \|.\|_a) \longrightarrow (X, \|.\|_b) \quad \textit{und} \quad Id : (X, \|.\|_b) \longrightarrow (X, \|.\|_a)$$

stetig sind.
Gilt nur die erste Stetigkeit, also $\|x\|_b \leq M\|x\|_a$, so nennt man $\|.\|_a$ feiner *und $\|.\|_b$* gröber.

Beweis. Leicht. □

Komposition (Multiplikation) von Operatoren.

Seien $A \in L(X,Y)$, $B \in L(X,Y)$ und X, Y, Z normierte Vektorräume. Dann definiert man

$$B \circ A : X \to Z, \quad (B \circ A)x := B(Ax) \quad \text{für alle } x \in X.$$

Es ist klar, dass $B \circ A$ wieder linear ist.

Weiter ist $B \circ A$ stetig, denn

$$\|(B \circ A)(x)\| = \|B(Ax)\| \leq \|B\|\|Ax\| \leq \|B\|\|A\|\|x\| \quad \text{für alle } x \in X.$$

Also ist $B \circ A$ stetig und

$$\boxed{\|B \circ A\| \leq \|B\|\|A\|}. \tag{2.4}$$

Auch die Multiplikation ist stetig:
Sind $(A_n), A \in L(X,Y)$ mit $A_n \to A$ und sind $(B_n), B \in L(X,Y)$ mit $B_n \to B$, so gilt auch in $L(X,Y)$

$$B_n \circ A_n \to B \circ A.$$

In der Tat ist

$$\begin{aligned}
\|B_n A_n(x) - B \circ A(x)\| &\leq \|B_n \circ A_n(x) - B_n \circ A(x)\| + \|B_n \circ A(x) - B \circ A(x)\| \\
&= \|B_n(A_n x - Ax)\| + \|(B_n - B)(Ax)\| \\
&\leq \|B_n\|\|A_n x - Ax\| + \|B_n - B\|\|Ax\|
\end{aligned}$$

für alle $x \in X$ und für alle $n \in \mathbb{N}$. Es folgt

$$\begin{aligned}
\|B_n \circ A_n - B \circ A\| &= \sup_{x \in B_X} \|(B_n \circ A_n - B \circ A)(x)\| \\
&\leq \|B_n\| \sup_{x \in B_X} \|A_n x - Ax\| + \|B_n - B\| \sup_{x \in B_X} \|Ax\|. \\
&= \|B_n\|\|A_n - A\| + \|B_n - B\|\|A\|
\end{aligned}$$

Da $B_n \to B$ in $L(X,Y)$, ist $\|B_n\|$ beschränkt. Die rechte Seite geht also für $n \to \infty$ gegen 0.

Wir haben insbeondere gezeigt, dass für alle $A \in L(X)$ und für alle $n \in \mathbb{N}$

$$A^n = \underbrace{A \circ \ldots A}_{n\text{-mal}} \in L(X).$$

Beispiel 2.7. Betrachte auf $\mathbb{R}^2$

$$A : (x,y) \mapsto (x,0) \quad \text{und} \quad B : (x,y) \mapsto (0,y).$$

Dann gilt auf der einen Seite

$$\|A \circ B\| = \sup_{(x,y) \in B_X} \|A \circ B(x,y)\| = 0$$

und auf der anderen Seite

$$\|A\| = \sup_{(x,y)\in B_X} \|A(x,y)\| = 1 = \sup_{(x,y)\in B_X} \|B(x,y)\| = \|B\|.$$

Also ist $\|AB\| < \|A\|\|B\|$.

An diesem Beispiel erkennt man, dass sogar gelegentlich die Striktheit in (2.4) gilt.

$\square$

2.2 Invertierbare Operatoren

Will man eine Gleichung wie z.B. $Ax = b$ nach x auflösen, so würde man das gerne durch Anwendung des Inversen A^{-1} auf beiden Seiten. Dazu wäre es aber vorteilhaft zu wissen, ob A^{-1} überhaupt existieren tut und wenn ja, ob dann auch A^{-1} stetig ist. Um diese Fragen zu klären, zunächst (zur Wiederholung) einige

Begriffe.

* $A \in L(X, Y)$ heißt *invertierbar*, falls A bijektiv und $A^{-1} \in L(Y, X)$ ist[4].
* Ein invertierbarer Operator $A \in L(X, Y)$ heißt auch *Isomorphismus*.
* Die normierten Vektorräume $(X, \|.\|_X)$ und $(Y, \|.\|_Y)$ heißen *isomorph*, falls ein Isomorphismus $A \in L(X, Y)$ existiert[5].

  Man schreibt dann
  $$\boxed{X \simeq Y}$$
  und identifiziert X und Y.
* Eine lineare Abbildung $A : X \to Y$ heißt *isometrisch*, wenn
  $$\|Ax\|_Y = \|x\|_Y \quad \text{für alle } x \in X.$$
  Dann folgt sofort $A \in L(X, Y)$, $\|A\|_{L(X,Y)} = 1$ und A ist injektiv (aber nicht notwendig surjektiv).
* Ist A bijektiv und isometrisch, so ist A invertierbar und heißt *isometrischer Isomorphismus*.
* X und Y heißen *isometrisch isomorph*, wenn ein isometrischer Isomorphismus $A : X \to Y$ existiert. Man schreibt dann
  $$\boxed{X \cong Y}.$$

Lemma 2.1. *Ist $A \in L(X, Y)$ bijektiv, so folgt, dass A^{-1} linear ist.*

Beweis. Seien $y_1, y_2 \in Y$ und $\alpha, \beta \in \mathbb{K}$. Dann gibt es $x_1, x_2 \in X$ mit $Ax_i = y_i$, $i = 1, 2$. Es folgt

$$\begin{aligned} A^{-1}(\alpha y_1 + \beta y_2) &= A^{-1}(\alpha A x_1 + \beta A x_2) = A^{-1}(A(\alpha x_1 + \beta x_2)) = \alpha x_1 + \beta x_2 \\ &= \alpha A^{-1} y_1 + \beta A^{-1} y_2. \end{aligned}$$

□

Beispiel 2.8. Wie wir schon in der Fußnote erwähnt hatten, folgt für bijektives $A \in L(X, Y)$ nicht notwendigerweise $A^{-1} \in L(Y, X)$. Betrachte dazu die lineare Abbildung

$$I : (C^1([0, 1]), \|.\|) \longrightarrow (C^1([0, 1]), \|.\|_\infty)$$

mit der Norm

$$\|f\| = \max\{\|f\|_\infty, \|f'\|_\infty\}, \quad f \in C^1([0, 1]).$$

Dann ist I bijektiv, linear und stetig, weil

$$\|I(f)\|_\infty = \|f\|_\infty \leq \|f\|.$$

Nun ist aber I^{-1} nicht stetig. Dazu müssten nämlich die beiden angegeben Normen nach Satz 2.2 äquivalent sein. Das sind sie aber nicht nach Beispiel 1.25.

□

[4] Achtung: I.A. ist A^{-1}, für $A \in L(X, Y)$ bijektiv, **nicht** notwendig stetig!

[5] Isomorphe normierte Vektorräume besitzen die gleichen topologischen Eigenschaften: Ist z.B. $(X, \|.\|)$ vollständig, so ist auch $(Y, \|.\|)$ vollständig, wenn X und Y isomorph sind

Beispiel 2.9. Wir betrachten den Vektorraum der konvergenten Zahlenfolgen c ausgestattet mit der $\|.\|_\infty$-Norm

$$\|x\|_\infty = \|(t_n)_n\|_\infty = \sup_{n\in\mathbb{N}} |t_n|, \quad x \in c.$$

Dann betrachte den Vektorraum $c_0 \subset c$ der gegen 0 konvergenten Zahlenfolgen, ebenfalls ausgetsattet mit der $\|.\|_\infty$-Norm. Dann gilt

$$\boxed{c \simeq c_0}. \tag{2.5}$$

Beweis. Wir müssen zeigen, dass die lineare (warum?) Abbildung

$$A : c \to c_0, \quad (x_n)_n \mapsto (x_\infty, x_1 - x_\infty, x_2 - x_\infty, \ldots)$$

ein Isomorphismus ist, wobei $x_\infty := \lim x_n$ bedeuten soll.

a) <u>A ist wohldefiniert:</u> Es ist klar, dass $A(x_n)_n \in c_0$.

b) <u>A ist injektiv:</u>

Betrachte

$$A(x_n)_n = (0)_n \Longrightarrow x_\infty = 0 \quad \text{und} \quad x_n = 0 \quad \text{für alle } n.$$

Also folgt aus $Ax = (0)_n$, dass $x = 0$.

c) <u>A ist stetig:</u>

Das folgt aus

$$\|Ax\|_\infty = \sup_{n\in\mathbb{N}}(\underbrace{|x_n - x_\infty|}_{\leq |x_n|+|x_\infty|} \vee |x_\infty|) \leq \sup_{n\in\mathbb{N}} |x_n| + \underbrace{|x_\infty|}_{\lim |x_n| \leq \sup_n |x_n|} \leq 2 \sup_{n\in\mathbb{N}} |x_n| = 2\|x\|_\infty.$$

Also ist $\|A\| \leq 2$ und daher A stetig.

d) <u>A ist surjektiv:</u>

Wenn $y = (y_n)_n \in c_0$, dann ist

$$x = (y_1 + y_2, y_1 + y_3, y_1 + y_4, \ldots) \in c$$

mit

$$\lim x_n = y_1 \quad \text{und} \quad Ax = y.$$

Damit haben wir auch gleich gezeigt, dass A bijektiv ist und

$$A^{-1} : c_0 \to c, \quad A^{-1}y = (y_1 + y_2, y_1 + y_3, y_1 + y_4, \ldots).$$

e) <u>A^{-1} ist linear und wohldefiniert:</u> Klar.

f) <u>A^{-1} ist stetig:</u>

Für alle $y \in c_0$ ist nämlich

$$\|A^{-1}\| \leq 2,$$

denn

$$\|A^{-1}y\|_\infty = \sup_{n\in\mathbb{N}} |y_1 + y_{n+1}| \leq |y_1| + \sup_{n\in\mathbb{N}} |y_{n+1}| \leq 2\|y\|_\infty.$$

Achtung: A ist **nicht** isometrisch: Etwa gilt für $x = (\frac{1}{4} - \frac{1}{n})_n$, dass

$$\tfrac{3}{4} = \|x\|_\infty \neq \|Ax\|_\infty = 1.$$

□

Beispiel 2.10. Wir betrachten die Folgenräume

$$l^1 = \{\, (x_n)_n \mid \sum_{n=1}^{\infty} |x_n| < \infty \,\}, \quad \|x\|_1 = \sum_{n=1}^{\infty} |x_n|\,,$$
$$l^\infty = \{\, (x_n)_n \mid \sup |x_n| < \infty \,\}, \quad \|x\|_\infty = \sup |x_n|\,.$$

Dann sind $(l^1, \|.\|_1)'$ und $(l^\infty, \|.\|_\infty)$ isometrisch isomorph. Kurz gesagt

$$\boxed{(l^1)' \cong l^\infty}. \tag{2.6}$$

Betrachte dazu

$$A : l^\infty \to (l^1)', \quad \alpha = (\alpha_n)_n \mapsto A\alpha,$$

wobei

$$(A\alpha)(x) = (A\alpha)(x_n)_n = \sum_{n=1}^{\infty} \alpha_n x_n \quad \text{für alle } x \in l^1.$$

Dann ist A ein isometrischer Isomorphismus:

a) <u>A ist wohldefiniert</u>, d.h. $A\alpha \in (l^1)'$ für alle $\alpha \in l^\infty$:

Zunächst ist $(A\alpha)(x) = \sum_{n=1}^{\infty} \alpha_n x_n$ konvergent. Weiter gilt

$$|(A\alpha)(x)| \le \sum_{n=1}^{\infty} |\alpha_n|\,|x_n| \le \sup |\alpha_n| \sum_{n=1}^{\infty} |x_n| = \|\alpha\|_\infty \|x\|_1,$$

also ist $(A\alpha)(x)$ sogar absolut konvergent, d.h.

$$A\alpha : l^1 \to \mathbb{R}$$

ist wohldefiniert.

b) <u>A ist linear:</u>

Seien dazu $\alpha = (\alpha_n)_n$, $\beta = (\beta_n)_n \in l^\infty$ sowie $\lambda, \mu \in \mathbb{K}$. Dann gilt für alle $x \in l^1$

$$\begin{aligned} A(\lambda\alpha + \mu\beta)(x) &= \sum_{n=1}^{\infty} (\lambda\alpha_n + \mu\beta_n) x_n = \lambda \sum_{n=1}^{\infty} \alpha_n x_n + \mu \sum_{n=1}^{\infty} \beta_n x_n \\ &= \lambda A(\alpha)(x) + \mu A(\beta)(x) = (\lambda A(\alpha) + \mu A(\beta))(x), \end{aligned}$$

d.h.

$$A(\lambda\alpha + \mu\beta) = \lambda A\alpha + \mu A\beta.$$

c) <u>$A \in L(l^\infty, (l^1)')$:</u>

Trivial, denn A ist nach 2. linear und nach 1. wegen

$$\|A\| \le 1$$

stetig.

d) <u>A ist isometrisch:</u>

Sei dazu $\alpha \in l^\infty$. Wähle

$$x^{(n)} := e_n \in B_{l^1}.$$

Dann ist für alle $n \in \mathbb{N}$

$$\|A\alpha\|_{(l^1)'} = \sup_{x \in B_{l^1}} |(A\alpha)(x)| \ge \left|(A\alpha)x^{(n)}\right| = |\alpha_n|\,.$$

Es folgt für alle $\alpha \in l^\infty$

$$\|A\alpha\|_{(l^1)'} \geq \sup_{n\in\mathbb{N}} |\alpha_n| = \|\alpha\|_{l^\infty}.$$

Zusammen mit a) folgt daher für alle $\alpha \in l^\infty$

$$\|A\alpha\|_{(l^1)'} = \|\alpha\|_{l^\infty},$$

d.h. A ist eine Isometrie.

e) <u>A ist surjektiv:</u>

Sei dazu $\varphi \in (l^1)'$. Setze

$$\alpha_n := \varphi(e_n) \quad \forall_{n\in\mathbb{N}}.$$

Dann gilt $\alpha = (\alpha_n)_n \in l^\infty$, denn

$$|\alpha_n| = |\varphi(e_n)| \leq \|\varphi\|_{(l^1)'}\|e_n\|_{(l^1)'} = \|\varphi\|_{(l^1)'} < \infty \quad \forall_{n\in\mathbb{N}},$$

d.h.

$$\sup|\alpha_n| = \|\alpha\|_{l^\infty} < \infty.$$

Außerdem ist für alle $x = (x_n)_n \in l^1$

$$\begin{aligned}(A\alpha)(x) \quad &= \quad \sum_{n=1}^{\infty} \alpha_n x_n = \lim_{N\to\infty} \sum_{n=1}^{N} \alpha_n x_n = \lim_{N\to\infty} \sum_{n=1}^{N} \varphi(e_n) x_n \\ &\overset{\varphi \text{ linear}}{=} \quad \lim_{N\to\infty} \varphi(\sum_{n=1}^{N} x_n e_n) = \varphi(x).\end{aligned}$$

Das letzte Gleichheitszeichen gilt, weil φ stetig ist und

$$\sum_{n=1}^{N} x_n e_n \overset{N\to\infty}{\longrightarrow} \sum_{n=1}^{\infty} x_n e_n = x \in l^1,$$

d.h. es ist

$$A\alpha = \varphi.$$

f) <u>A ist injektiv:</u> Trivial, da jede Isometrie injektiv ist.

g) <u>A^{-1} ist stetig:</u>

Das folgt aus der Isometrie

$$\left\|A^{-1}y\right\| = \left\|A^{-1}(Ax)\right\| = \|x\| = \|Ax\| < \infty.$$

□

Bemerkungen.

1. Man kann zeigen, dass sogar
$$(l^p)' \cong l^q$$
gilt für $1 \leq p < \infty$ und $\frac{1}{p} + \frac{1}{q} = 1$. q heißt dann der zu p *konjugierte* Exponent.

2. Genauso kann man zeigen, dass
$$(L^p)' \cong L^q$$
gilt für $1 \leq p < \infty$. Im endlichdimensionalen gilt sogar $L^\infty \cong L^1$.

3. Analog wie in 1. kann man zeigen, dass
$$(c_0)' \cong l^1.$$

Die Neumannsche Reihe.

In einem Spezialfall ist es möglich, das Inverse eines Operators zu berechnen. Dies möchte ich hier nun vorführen.

Notation.
Sei $(X, \|.\|_X)$ ein normierter Raum. $I \in L(X)$ sei die *Identität*, d.h. es gilt

$$I \cdot x = x \quad \text{für alle } x \in X.$$

Satz 2.4 (Neumannsche Reihe). *Sei $(X, \|.\|_X)$ ein Banachraum und $A \in L(X)$ mit $\|A\|_{L(X)} < 1$. Dann ist $I - A$ invertierbar und*

$$(I - A)^{-1} = \sum_{n=0}^{\infty} A^n, \quad \left\|(I - A)^{-1}\right\| \leq \frac{1}{1 - \|A\|_{L(X)}}.$$

Bemerkungen.

1. $\sum_{n=0}^{\infty} A^n$ heißt auch die *Neumannsche Reihe.*
2. Der Satz gilt auch in einer etwas allgemeineren Form (wie der Beweis zeigen wird):

   Sei X ein normierter Raum und $A \in L(X)$. Konvergiert $\sum_{n=0}^{\infty} A^n$ in $L(X)$, so ist $I - A$ invertierbar mit

   $$(I - A)^{-1} = \sum_{n=0}^{\infty} A^n.$$

3. Ein Operator $A \in L(X)$ ist invertierbar genau dann, wenn gilt

   $$\exists_{S \in L(X)} : S \cdot A = I = A \cdot S.$$

   In diesem Fall ist $A^{-1} = S$.

   Beweis. Zu „$\Longrightarrow$“: Trivial, denn wähle einfach $S := A^{-1}$.

   Zu „$\Longleftarrow$“: Zunächst ist A injektiv, weil

   $$Ax = 0 \Longrightarrow 0 = S(Ax) = x \Longrightarrow x = 0.$$

   Weiter ist A surjektiv: Sei $y \in X$. Dann setze $x = Sy$. Es folgt $Ax = y$, denn

   $$y = Iy = A(Sy) = Ax.$$

   Damit ist die Existenz von A^{-1} gezeigt. S ist eindeutig bestimmt: Ist nämlich $T \in L(X)$ ein weiterer Operator mit $TA = TA = I$, so folgt

   $$(SA - TA) = 0 \Longrightarrow (S - T)A = 0 \Longrightarrow S - T = 0 \Longrightarrow S = T.$$

   Beachte, dass $A \neq 0$, weil A^{-1} existiert. Aus der Eindeutigkeit folgt $A^{-1} = S$, so dass $A^{-1} \in L(X)$. Also ist A invertierbar. □

Beweis des Satzes. Weil $\|A\| < 1$, ist die geometrische Reihe $\sum_{n=0}^{\infty} \|A\|^n$ konvergent. Da

$$\left\| \sum_{k=m}^{n} A^k \right\| \leq \sum_{k=m}^{n} \|A\|^k \quad \text{für alle } n > m, \quad m, n \in \mathbb{N} \tag{2.7}$$

folgt, dass $(\sum_{k=0}^{N} A^k)_N$ eine Cauchyfolge in $L(X)$ ist. Da X ein Banachraum ist, ist nach Satz 2.2 auch $L(X)$ ein Banachraum, also konvergiert $\sum_{n=0}^{\infty} A^n$ in $L(X)$.

Weiter gilt für alle $N \in \mathbb{N}$

$$(I-A)\sum_{k=0}^{N} A^k = \sum_{k=0}^{N} A^k - \sum_{k=0}^{N} A^{k+1} \overset{\text{Teleskopsumme}}{=} I - A^{N+1} = \sum_{k=0}^{N} A^k (I-A).$$

Wegen $\|A\|^N \to 0$ für $N \to \infty$ und weil die Komposition von Operatoren stetig bzgl. der Operatornorm in $L(X)$ ist, folgt für $N \to \infty$

$$(I-A)\sum_{k=0}^{\infty} A^k = I = \sum_{k=0}^{\infty} A^k (I-A)$$

Aus Punkt 3. der Bemerkung folgt also, dass $(I-A)$ invertierbar ist und

$$(I-A)^{-1} = \sum_{k=0}^{\infty} A^k.$$

Aus (2.7) mit $m = 0$ und $n \to \infty$ folgt schließlich

$$\left\|(I-A)^{-1}\right\| \leq \frac{1}{1-\|A\|}.$$

□

Beispiel 2.11 (Anwendung). Sei $k \in C([0,1]^2)$ mit

$$\sup_{s\in[0,1]} \int_0^1 |k(s,t)|\, dt < 1.$$

Dann ist

$$A_k : C[0,1] \longrightarrow C[0,1], \quad u \mapsto \int_0^1 k(.,t)u(t)dt$$

der Gestalt, dass $\|A_k\| < 1$ (vergleiche dazu Beispiel 2.3) und man kann Satz 2.4 anwenden, um das Problem

$$(*) = \begin{cases} u \in C[0,1] \\ u(s) - \int_0^1 k(s,t)u(t)dt = f(s), \quad s \in [0,1] \end{cases}$$

für beliebiges $f \in C[0,1]$ zu lösen.

Es ist

$$(I-A_k)u = f \Longleftrightarrow u = (I-A_k)^{-1} f = \sum_{n=0}^{\infty} A_k^n f.$$

Definiere nun die *iterierten Kerne* zur Berechnung von A_k^n:

$$k_1(s,t) = k(s,t) \qquad s,t, \in [0,1]$$

$$k_n(s,t) = \int_0^1 k(s,u)k_{n-1}(u,t)du \qquad s,t \in [0,1], \quad n = 2,3,\ldots$$

Dann ist

$$A_k^n = A_{k_n} \quad \text{für alle } n \in \mathbb{N}.$$

Beweis per Induktion:

<u>$n = 1$:</u> Trivial: $A_k = A_{k_1}$.

<u>$n \to n+1$:</u> Es gilt

$$\begin{aligned} A_k^{n+1}u(s) &= A_k \circ A_k^n u(s) \overset{IV}{=} A_k \circ A_{k_n} u(s) = A_k \left(\int_0^1 k_n(.,t)u(t)dt\right)(s) \\ &= \int_0^1 k(s,\sigma) \int_0^1 k_n(\sigma,t)u(t)dtd\sigma \overset{Fubini}{=} \int_0^1 (\int_0^1 k(s,\sigma)k_n(\sigma,t)d\sigma)u(t)dt \\ &= \int_0^1 k_{n+1}(s,t)u(t)dt = A_{k_{n+1}}. \end{aligned}$$

Jetzt gilt die Abschätzung

$$|k_n(s,t)| \leq \|k_{n-1}\|_\infty \int_0^1 |k(s,u)|\, du \quad \text{für alle } s,t \in [0,1].$$

Daraus folgt

$$\|k_n\|_\infty = \sup_{(s,t)\in[0,1]^2} |k_n(s,t)| \leq \|k_{n-1}\|_\infty \underbrace{\sup_{s\in[0,1]} \int_0^1 |k(s,u)|\, du}_{=\|A_k\|<1}.$$

Iterativ erhält man daraus für alle $n \in \mathbb{N}$

$$\|k_n\|_\infty \leq \ldots \leq \|k\|_\infty \underbrace{\|A_k\|^{n-1}}_{<1}.$$

Weil

$$\sum_{n=1}^{\infty} \|A_k\|^{n-1} < \infty,$$

konvergiert auch $\sum_{n=1}^{\infty} k_n$ gleichmäßig auf $[0,1]$ gegen eine Funktion $h \in C([0,1]^2)$. Weiter bemerken wir die Eigenschaft

$$A_g + A_f = A_{g+f}. \tag{2.8}$$

Es folgt

$$\sum_{k=0}^{N} A_k^n = I + \sum_{n=1}^{N} A_k^n = I + \sum_{n=1}^{N} A_{k_n} \overset{(2.8)}{=} I + A_{\sum_{n=1}^{N} k_n} \overset{N\to\infty}{\longrightarrow} I + A_h$$

Weiter erhalten wir

$$\begin{aligned}
\|A_{k_n} - A_h\|_{L(C[0,1])} &= \sup_{u\in B_{C[0,1]}} \left| \int_0^1 (k_n(.,t) - h(.,t))u(t)dt \right| \\
&\leq \sup_{u\in B_{C[0,1]}} \sup_{s\in[0,1]} \int_0^1 |k_n(s,t) - h(s,t)|\, |u(t)|\, dt \\
&\leq \sup_{s\in[0,1]} \int_0^1 |k_n(s,t) - h(s,t)|\, dt \\
&\leq \sup_{s\in[0,1]} \sup_{t\in[0,1]} |k_n(s,t) - h(s,t)| = \underbrace{\|k_n - h\|_\infty}_{\longrightarrow 0}.
\end{aligned}$$

Also gilt

$$A_{k_n} \longrightarrow A_h \quad \text{in } L(C[0,1]).$$

Somit ist

$$(I - A_k)^{-1} = \sum_{n=0}^{\infty} A_k^n = I + A_h,$$

d.h. die gesuchte Lösung u von $(*)$ ist

$$u = (I + A_h)f$$

und u ist eindeutig und hängt stetig von f ab.

Bemerkung. h heißt *auflösender Kern* der Integralgleichung $(*)$ und kann in gewissen Fällen explezit berechnet werden (vgl. Übung).

□

Abschließend möchte ich noch zeigen, dass die Menge der invertierbaren Operatoren offen in $L(X, Y)$ ist:

Korollar 2.1. *Sei X ein Banachraum. Ferner sei der Operator $A \in L(X)$ invertierbar. Gilt dann für $B \in L(X)$ die Abschätzung*

$$\|A - B\|_{L(X)} < \frac{1}{\|A^{-1}\|_{L(X)}},$$

so ist auch B invertierbar und

$$B^{-1} = \sum_{n=0}^{\infty} (A^{-1}(A - B))^n A^{-1}.$$

Bemerkung. Aus Korollar 2.2 wird wieder folgen, dass das obere Korollar auch nur für bijektives A und B gilt.

Beweis. Schreibe B als

$$B = A(I - A^{-1}(A - B)).$$

Nach Voraussetzung ist A invertierbar. Außerdem ist

$$\left\|A^{-1}(A - B)\right\|_{L(X,Y)} \leq \left\|A^{-1}\right\|_{L(X,Y)} \|A - B\|_{L(X,Y)} < 1$$

und aus dem Satz zur Neumannschen Reihe (Satz 2.4) folgt, dass auch $[I - A^{-1}(A - B)]$ invertierbar ist mit

$$(I - A^{-1}(A - B))^{-1} = \sum_{n=0}^{\infty} (A^{-1}(A - B))^n.$$

Somit ist dann aber auch B invertierbar und es gilt

$$B^{-1} = (I - A^{-1}(A - B))^{-1} A^{-1} = \sum_{n=0}^{\infty} (A^{-1}(A - B))^n A^{-1}.$$

□

2.3 Die drei wichtigen Hauptsätze der Funktionalanalysis

Wir kommmen nun zu den Hauptsätzen der Funktionalanalysis:

- Satz von Banach-Steinhaus
- Satz von der offenen Abbildung
- Satz vom abgeschlossenen Graphen

2.3.1 Der Satz von Banach-Steinhaus

Wir beginnen mit dem

Satz 2.5 (**von Banach-Steinhaus: Prinzip der gleichmäßigen Beschränktheit**). *Seien $(X, \|.\|_X)$ ein Banachraum, $(Y, \|.\|_Y)$ ein normierter Vektorraum, I eine beliebige Indexmenge und $A_i \in L(X, Y)$ für alle $i \in I$. Falls*

$$\sup_{i \in I} \|A_i x\| < \infty \quad \text{für alle } x \in X \tag{2.9}$$

gilt, so gilt sogar

$$\sup_{i \in I} \|A_i\| < \infty.$$

Beweis. Setze

$$X_n = \{\, x \in X \mid \sup_{i \in I} \|A_i x\|_Y \leq n \,\}, \quad n \in \mathbb{N}.$$

Wegen (2.9) gilt

$$X = \bigcup_{n \in \mathbb{N}} X_n.$$

Außerdem gilt

$$X_n = \bigcap_{i \in I} \varphi_i^{-1}([0, n]),$$

wobei

$$\varphi_i : X \to \mathbb{K}, \quad x \mapsto \|A_i x\|.$$

Als Komposition von stetigen Abbildungen ist φ_i stetig. Deshalb ist das Urbild der abgeschlossenen Menge $[0, n]$ abgeschlossen für alle $i \in I$. Als beliebige Vereinigung abgeschlossener Teilmengen ist dann auch X_n abgeschlossen.

Aus dem Satz von Baire (Satz 1.17) folgt dann:

$$\exists_{n_0 \in \mathbb{N}} : \overset{\circ}{X}_{n_0} \neq \emptyset.$$

Daher existieren ein $x_0 \in \overset{\circ}{X}_{n_0}$ und ein $\epsilon > 0$ mit

$$\bar{B}(x_0, \epsilon) = \{\, x \in X \mid \|x - x_0\| \leq \epsilon \,\} \subset X_{n_0}.$$

Für $x \in B_X$ ist dann

$$x_0 + \epsilon x \in \bar{B}(x, \epsilon) \subset X_{n_0},$$

denn

$$\|x_0 + \epsilon x - x_0\| = \epsilon \|x\| \leq \epsilon.$$

Es folgt

$$\|A_i(x_0 + \epsilon x)\| \leq n_0 \quad \text{für alle } i \in I.$$

Daher ist

$$\|A_i x\| = \tfrac{1}{\epsilon}\|A_i(\epsilon x)\| \leq \tfrac{1}{\epsilon}\|A_i(\epsilon x + x_0)\| + \tfrac{1}{\epsilon}\|A_i x_0\| \leq \tfrac{2n_0}{\epsilon},$$

d.h.

$$\|A_i\|_{L(X,Y)} = \sup_{x \in B_X} \|A_i x\| \leq \tfrac{2n_0}{\epsilon}.$$

Mit anderen Worten heißt das

$$\sup_{i \in I} \|A_i\| < \infty.$$

□

Definition 2.1 (Starke Konvergenz). Seien $(X, \|.\|_X)$, $Y, \|.\|_Y$ normierte Vektorräume und $(A_n)_n, A \subset L(X,Y)$. Man sagt:

$(A_n)_n$ *konvergiert stark gegen* A, falls

$$Ax = \lim_{n\to\infty} A_n x \quad \text{für alle } x \in X.$$

Notation.

$$A = \text{s-}\lim_{\text{n}\to\infty} A_n.$$

$(A_n)_n$ heißt *stark konvergent in* $L(X,Y)$, falls ein $B \in L(X,Y)$ existiert mit $B = \text{s-}\lim_{\text{n}\to\infty} A_n$.

Bemerkungen.

1. Gilt $A_n \to A$ in der Operatornorm in $L(X,Y)$, so folgt

$$A = \text{s-}\lim_{\text{n}\to\infty} A_n.$$

Beweis. Für alle $x \in X$ gilt

$$\|A_n x - Ax\|_Y \leq \|A_n - A\|_{L(X,Y)} \|x\|_X \overset{n\to\infty}{\longrightarrow} 0.$$

□

2. Die Umkehrung gilt im Allgemeinen nicht, wie das folgende Beispiel zeigen wird.

Beispiel 2.12. Sei $X = c_0$ ausgestattet mit der $\|.\|_\infty$-Norm. Betrachte die Abbildung

$$A_n : c_0 \to c_0, \quad x = (x_n)_n \mapsto x_n \cdot e_n = (0, \ldots, 0, x_n, 0, \ldots).$$

Klar ist, dass $A_n \in L(c_0)$. Nun gilt für jedes $x = (x_n)_n \in c_0$

$$\|A_n x\|_\infty = |x_n| \overset{n\to\infty}{\longrightarrow} 0,$$

d.h.

$$\text{s-}\lim_{\text{n}\to\infty} A_n = A, \quad Ax = 0 \quad \text{für alle } x \in c_0.$$

Aber auf der anderen Seite ist für alle $n \in \mathbb{N}$

$$\|A_n\|_{L(c_0)} = \sup_{x \in B_{c_0}} \|A_n x\|_\infty \overset{x = e_n \in B_{c_0}}{\geq} \|A_n e_n\|_\infty = 1.$$

Somit konvergiert A_n nicht gegen A im Sinne der Operatornorm.

Folgerungen aus dem Satz von Banach-Steinhaus.

1. Seien $(X, \|.\|_X)$ ein Banachraum, $(Y, \|.\|_Y)$ ein normierter Vektorraum und $(A_n)_n \subset L(X,Y)$, so dass

$$Ax := \lim_{n\to\infty} A_n x \quad \text{existiert für alle } x \in X. \tag{2.10}$$

Dann ist $A \in L(X,Y)$ und somit ist (nach Definition)

$$A = \text{s-}\lim_{\text{n}\to\infty} A_n.$$

Beweis. Aus (2.10) folgt, dass $(A_n x)_n$ beschränkt ist in Y für alle $x \in X$. Damit sind die Voraussetzungen des Satzes von Banach-Steinhaus erfüllt und es folgt

$$C := \sup_n \|A_n\|_{L(X,Y)} < \infty.$$

Dann folgt aber für alle $x \in X$

$$\|Ax\|_Y = \lim_{n\to\infty} \|A_n x\|_Y \leq \limsup_{n\to\infty} \|A_n\|_{L(X,Y)} \|x\|_X \leq C\|x\|_X. \tag{2.11}$$

Es ist klar, dass durch (2.10) eine lineare Abbildung $A : X \to Y$ definiert ist. Mit (2.11) folgt $A \in L(X,Y)$. □

Bemerkung. Wenn X nicht vollständig ist, dann gilt die Folgerung 1. (ebenso wie der Satz von Banach-Steinhaus selbst) im Allgemeinen nicht. Vergleiche Beispiel 3.10.

2. Die Komposition von stetigen, linearen Operatoren ist stetig im Sinne der starken Konvergenz von Operatoren:

Seien $(A_n)_n, A \subset L(X,Y)$, $(B_n)_n, B \subset L(X,Y)$, $(Y, \|.\|_Y)$ ein Banachraum, $(X, \|.\|_X)$, $(z, \|.\|_Z)$ normierte Vektorräume und

$$A = \text{s-}\lim_{\text{n}\to\infty} A_n, \quad B = \text{s-}\lim_{\text{n}\to\infty} B_n.$$

Dann gilt

$$B \circ A = \text{s-}\lim_{\text{n}\to\infty} B_n \circ A_n.$$

Beweis. Für $x \in X$ betrachte

$$\begin{aligned}\|B_n \circ A_n x - B \circ Ax\|_Z &\leq \|B_n \circ (A_n x - Ax)\|_Z + \|B(Ax) - B_n(Ax)\|_Z \\ &\leq \|B_n\|_{L(Y,Z)} \|A_n x - Ax\|_Y + \|B(Ax) - B_n(Ax)\|_Z \longrightarrow 0.\end{aligned}$$

Beachte, dass hier nach dem Satz von Banach-Steinhaus der Ausdruck $\|B_n\|$ beschränkt ist und die beiden anderen Ausdrücke wegen der starken Konvergenz gegen 0 gehen. □

2.3.2 Der Satz von der offenen Abbildung

Wir beginnen mit einer

Definition 2.2 (Offene Abbildung). Seien $(X, \|.\|_X)$ und $(Y, \|.\|_Y)$ normierte Vektorräume und $A \in L(X, Y)$. Dann heißt A *offen* oder *offene Abbildung*, wenn A offene Teilmengen von X auf offene Teilmengen von Y abbildet, d.h. $A(U)$ ist eine offene Teilmenge von Y für alle offenen $U \subset X$.

Bemerkung. Beachten Sie, dass offene Abbildungen nicht abgeschlossene Mengen auf abgeschlossene Mengen abbilden müssen. Dazu betrachten wir das folgende

Beispiel 2.13. Betrachte die Projektion

$$P : \mathbb{R}^2 \to \mathbb{R}, \quad (s, t) \mapsto s$$

und die abgeschlossene Menge

$$A := \{ (s, t) \in \mathbb{R}^2 \mid s \geq 0,\ st \geq 1 \}.$$

Dann ist P offen, aber $P(A) =]0, \infty[$ ist nicht abgeschlossen in $\mathbb{R}$.

□

Lemma 2.2. *$A \in L(X, Y)$ ist genau dann offen, wenn*

$$\exists_{r>0} : B_Y(0, r) \subset A(B_X(0, 1)). \tag{2.12}$$

Beweis. Zu „$\Longrightarrow$“: Trivial.

Zu „$\Longleftarrow$“: Sei also U offen in X. Wir müssen zeigen, dass $A(U)$ offen ist. Sei dazu $y_0 \in A(U)$. Wir müssen dann zeigen, dass

$$\exists_{\delta>0} B_Y(y_0, \delta) \subset A(U).$$

Wegen $y_0 \in A(U)$ gibt es ein $x_0 \in X$, so dass

$$Ax_0 = y_0.$$

Nach Voraussetzung ist U offen, also:

$$\exists_{\epsilon>0} : B_X(x_0, \epsilon) \subset U.$$

Aus (2.12) folgt wegen der Linearität von A

$$B_Y(0, \epsilon r) \subset A(B_X(0, \epsilon)) \quad \text{für alle } \epsilon > 0$$

und

$$\underbrace{Ax_0}_{=y_0} + B_Y(0, \epsilon r) \overset{(*)}{=} B_Y(y_0, \underbrace{\epsilon r}_{=:\delta}) \subset Ax_0 + A(B_X(0, \epsilon)) \overset{(**)}{=} A(B_X(x_0, \epsilon)) \subset A(U).$$

Zu $(*)$:

$$y_0 + B_Y(0, \epsilon r) = \{ y_0 + y' \mid \|y'\| < \epsilon r, y' \in Y \} = \{ y \in Y \mid \|y - y_0\| < \epsilon r \} = B_Y(y_0, \epsilon r).$$

Zu $(**)$:

$$Ax_0 + A(B_X(0, \epsilon)) = A(x_0 + B_X(0, \epsilon)) \overset{\text{wie } (*)}{=} A(B_X(x_0, \epsilon)).$$

□

Beispiel 2.14. Die Abbildung

$$A : l^\infty \to c_0, \quad (t_n) \mapsto (\tfrac{1}{n} t_n)$$

ist nicht offen. Es gilt nämlich

$$A(B_{l^\infty}(0,1)) = \{\, (t_n) \in c_0 \mid |t_n| < \tfrac{1}{n} \text{ für alle } n \in \mathbb{N} \,\}$$

Es gibt also kein $r > 0$, so dass $B_{c_0}(0,r) \subset A(B_{l^\infty}(0,1))$.

□

Beispiel 2.15 (Quotientenabbildung). Seien X und Y normierte Räume. Eine lineare Abbildung $Q : X \to Y$ heißt *Quotientenabbildung*, wenn Q die offene Kugel $\{\, x \in X \mid \|x\| < 1 \,\}$ auf die offene Kugel $\{\, y \in Y \mid \|y\| < 1 \,\}$ abbildet. Dann ist Q wegen Lemma 2.2 eine offene Abbildung und nach der folgenden Bemerkung 2. auch surjektiv. Ferner ist Q stetig mit $\|Q\| = 1$. Um das einzusehen, seien $\epsilon > 0$ beliebig und $x \in B_X$. Dann ist $\frac{x}{\|x\|+\epsilon} < 1$ und daher

$$\left\| Q(\tfrac{x}{\|x\|+\epsilon}) \right\| < 1 \Longrightarrow \|Q(x)\| < \|x\| + \epsilon.$$

Es folgt

$$\|Q\| = \sup_{x \in B_X} \|Qx\| \leq 1 + \epsilon$$

und weil $\epsilon > 0$ beliebig war schließlich

$$\|Q\| \leq 1.$$

Sei nun $1 \geq \delta > 0$ beliebig. Dann gibt es dazu ein $y \in B_Y$ mit $\|y\| = 1 - \delta < 1$. Weil Q surjektiv ist, gibt es dazu ein x mit $\|x\| < 1$ und $Qx = y$. Es folgt

$$\|Q\| = \sup_{x \in B_X} \|Qx\| \geq \|y\| = 1 - \delta$$

und da δ beliebig war ist

$$\|Q\| \geq 1.$$

□

Bemerkungen.

1. Eine bijektive offene Abbildung $A \in L(X,Y)$ ist invertierbar. Das ist trivial, denn A^{-1} ist stetig.

2. Offene lineare Abbildungen sind surjektiv.

   Beweis. Für $y = 0$ ist $x = 0$. Sei also $y \in Y$ mit $y \neq 0$. Sei $r > 0$. Setze

   $$w := \frac{r}{2} \frac{y}{\|y\|}.$$

   Dann gilt

   $$w \in B_Y(0,r) \overset{\text{Lemma 2.2}}{\subset} A(B_X(0,1)).$$

   Also gibt es ein $\tilde{x} \in B_X(0,1)$, so dass

   $$A\tilde{x} = w = \frac{r}{2} \frac{y}{\|y\|}.$$

   Daraus folgt wegen der Linearität von A

   $$A(\tfrac{2}{r\|y\|\tilde{x}}) = y.$$

   Damit finden wir zu jedem $y \in Y$ ein $x := \frac{2}{r}\|y\|\tilde{x}$, so dass $Ax = y$. □

Für das Resultat aus Bemerkung 2. gilt auch die Umkehrung für Banachräume X und Y:

Satz 2.6 (von der offenen Abbildung). *Seien $(X, \|.\|_X)$, $(Y, \|.\|_Y)$ Banachräume und $A \in L(X, Y)$ surjektiv. Dann ist A offen.*

Beweis. 1. Schritt: Wir wollen zeigen, dass

$$\exists_{r>0} : B_Y(0, 2r) \subset \overline{A(B_X(0,1))}. \tag{2.13}$$

Es ist

$$X = \bigcup_{n\in\mathbb{N}} B_X(0,n) = \bigcup_{n\in\mathbb{N}} n \cdot B_X(0,1).$$

Weil A surjektiv ist, ist dann

$$Y = A(X) = \bigcup_{n\in\mathbb{N}} n \cdot A(B_X(0,1))$$

und somit insbesondere

$$Y = \bar{Y} = \overline{A(X)} = \bigcup_{n\in\mathbb{N}} n \cdot \overline{A(B_X(0,1))}.$$

Da Y vollständig ist, folgt nun aus dem Satz von Baire (Satz 1.17), dass ein $n_0 \in \mathbb{N}$ existiert mit

$$(n_0\overline{A(B_X(0,1))})^\circ \neq \emptyset.$$

Beachte, dass

$$(\lambda M)^\circ = \lambda M^\circ \quad \text{für alle } \lambda > 0.$$

Also folgt

$$n_0(\overline{A(B_X(0,1))})^\circ \neq \emptyset \Longrightarrow (\overline{A(B_X(0,1))})^\circ \neq \emptyset.$$

Also existiert ein $y_0 \in Y$ und ein $r > 0$ mit

$$B_Y(y_0, 2r) \subset (\overline{A(B_X(0,1))})^\circ \subset \overline{A(B_X(0,1))}.$$

Nun ist $A(B_X(0,1))$ symmetrisch, d.h. wenn $z \in A(B_X(0,1))$, dann ist auch $-z \in A(B_X(0,1))$. Daher ist auch $\overline{A(B_X(0,1))}$ symmetrisch. Somit folgt dann aber, dass auch

$$B_Y(-y_0, 2r) \subset \overline{A(B_X(0,1))}.$$

In der Tat: Ist $z \in B_Y(-y_0, 2r)$, so folgt

$$\begin{aligned} \|z + y_0\| < 2r &\Longrightarrow \|-z - y_0\| < 2r \Longrightarrow (-z) \in B_Y(y_0, 2r) \\ &\Longrightarrow (-z) \in \overline{A(B_X(0,1))} \Longrightarrow z \in \overline{A(B_X(0,1))}. \end{aligned}$$

Aus der Linearität des Operators A und der Konvexität von $B_X(0,1)$ folgt außerdem, dass $A(B_X(0,1))$ konvex ist:

- $B_X(0,1)$ konvex: Wir müssen zeigen, dass $\overline{ab} \subset B_X(0,1)$ für alle $a, b \in B_X(0,1)$.

  Sei also $x \in \overline{ab}$. Dann gibt es ein $\lambda \in [0,1]$, so dass

  $$x = \lambda a + (1-\lambda)b.$$

  Es folgt

  $$\|x\| \leq \lambda\|a\| + (1-\lambda)\|b\| < \lambda + (1-\lambda) = 1,$$

  also $x \in B_X(0,1)$.

- $A(B_X(0,1))$ konvex: Analog.

Somit ist demnach auch der Abschluss $\overline{A(B_X(0,1))}$ konvex. Es folgt dann für beliebiges $y \in Y$ mit $\|y\| < 2r$, dass

$$(y_0 + y) \in B_Y(y_0, 2r) \subset \overline{A(B_X(0,1))} \quad \text{und} \quad (-y_0 + y) \in B_X(-y_0, 2r) \subset \overline{A(B_X(0,1))}.$$

Daraus folgt aufgrund der Konvexität von $\overline{A(B_X(0,1))}$ ($\lambda = 1/2$)

$$y = \frac{1}{2}(y_0 + y) + \frac{1}{2}(-y_0 + y) \in \overline{A(B_X(0,1))},$$

d.h. es gilt

$$B_Y(0, 2r) \subset \overline{A(B_X(0,1))}$$

und somit ist (2.13) gezeigt.

<u>2. Schritt:</u> Wir zeigen nun, dass sogar gilt

$$B_Y(0, r) \subset A(B_X(0,1)). \tag{2.14}$$

Aus (2.13) folgt sofort, dass für alle $\epsilon > 0$

$$B_Y(0, 2\epsilon r) \subset \overline{A(B_X(0,\epsilon))}$$

und somit insbesondere ($\epsilon = 2^{-m-1}$)

$$B_Y(0, 2^{-m} r) \subset \overline{A(B_X(0, 2^{-m-1}))}. \tag{2.15}$$

Sei nun $y \in B_Y(0, r)$.

Zu zeigen: Es gibt ein $x \in B_X(0,1)$ mit $y = Ax$.

Wir konstruieren dazu zunächst induktiv eine Folge $(z_n)_n \subset X$ mit

$$\begin{aligned} \|z_n\|_X &< 2^{-n} \quad \text{für alle } n \in \mathbb{N}, \\ \|y - A(z_1 + \ldots + z_n)\|_Y &< 2^{-n} r \quad \text{für alle } n \in \mathbb{N}. \end{aligned} \tag{2.16}$$

Beweis per Induktion.

<u>$n = 1$:</u> $m = 0$ in (2.15) heißt gerade $B_Y(0, r) \subset \overline{A(B_X(0, \frac{1}{2}))}$. Damit existiert ein $z_1 \in B_X(0, \frac{1}{2})$, d.h. $\|z_1\| < \frac{1}{2}$ mit

$$\|y - Az_1\|_Y < \tfrac{r}{2}.$$

<u>$n \to n+1$:</u> Seien für ein $n \in \mathbb{N}$ die ersten Glieder $z_1, \ldots, z_n$ mit der Eigenschaft (2.16) konstruiert. Wegen (2.15) für $m = n$ gilt

$$B_Y(0, 2^{-n} r) \subset \overline{A(B_X(0, 2^{-n-1}))}.$$

Nach Konstruktion ist $(y - A(z_1 + \ldots + z_n)) \in B_Y(0, 2^{-n} r)$. Somit existiert ein $z_{n+1} \in B_X(0, 2^{-n-1})$ mit

$$\|y - A(z_1 + \ldots z_n) - Az_{n+1}\| < 2^{-n-1} r,$$

also

$$\|y - A(z_1 + \ldots + z_{n+1})\| < 2^{-n-1} r.$$

Damit ist $(z_n)_n$ konstruiert.

Wegen $\|z_n\| < 2^{-n}$ für alle $n \in \mathbb{N}$ ist die Folge $(\sum_{k=1}^n z_k)_n$ eine Cauchyfolge in X. Da X vollständig ist, existiert

$$\sum_{k=1}^{\infty} z_k = \lim_{n \to \infty} \sum_{k=1}^{n} z_k =: x \quad \text{in } X,$$

und es gilt

$$\begin{aligned}\|x\| &= \lim_{n\to\infty} \|\sum_{k=1}^{n} z_k\| \leq \limsup_{n\to\infty} \sum_{k=1}^{n} \|z_k\| \overset{(*)}{<} \limsup_{n\to\infty} \sum_{k=1}^{n} 2^{-k} \\ &= \frac{1}{2}\sum_{k=0}^{\infty} 2^{-k} = 1,\end{aligned}$$

also ist $x \in B_X(0,1)$. Beachte, dass die Striktheit in $(*)$ tatsächlich gilt, denn z.B.

$$\sum_{k=1}^{\infty} 2^{-k} - \|z_k\| = \underbrace{(\frac{1}{2} - \|z_1\|)}_{>0} + \underbrace{\sum_{k=2}^{\infty} 2^{-k} - \|z_k\|}_{\geq 0} > 0.$$

Weiter im eigentlichen Beweis: Aus (2.16) und der Stetigkeit von A folgt dann

$$\|y - Ax\|_Y = \lim_{n\to\infty} \|y - A(\sum_{k=1}^{n} z_k)\| \leq 0,$$

d.h. aber gerade $y = Ax$ mit $x \in B_X(0,1)$, was zu zeigen war. □

Bemerkung. Die Aussage des Satzes gilt im Allgemeinen nicht mehr, wenn X oder Y nicht vollständig sind:

Beispiele 2.16.

a) Y nicht vollständig: Betrachte die Identität

$$I : (C^1[0,1], ||| \,.\, |||) \to (C^1[0,1], \|.\|_\infty),$$

wobei für $f \in C^1[0,1]$

$$||| \, f \, ||| := \max\{\|f\|_\infty, \|f'\|_\infty\}.$$

I ist linear, stetig und surjektiv (sogar bijektiv), aber I ist nicht offen, da die Umkehrabbildung

$$I^{-1} : (C^1[0,1], \|.\|_\infty) \to (C^1[0,1], |||.|||)$$

unstetig ist (vgl. Beispiel 2.8).

b) Ein Gegenbeispiel für nicht vollständiges X zu finden ist komplizierter, da ein unvollständiger Raum X stets vervollständigt werden kann, d.h. es existiert ein Banachraum $\hat{X}$, in dem X dicht liegt. Weiter kann ein surjektiver Operator $A \in L(X,Y)$ stets zu einem surjektiven Operator $\hat{A} \in L(\hat{X}, Y)$ in eindeutiger und in operatornormerhaltender Weise fortgesetzt werden.

□

Satz 2.7. *Seien $(X, \|.\|)$ ein normierter Vektorraum, $X_0 \subset X$ ein dichter Untervektorraum, $(Y, \|.\|)$ ein Banachraum und $A_0 \in L(X_0, Y)$. Dann existiert genau ein $A \in L(X,Y)$ mit den Eigenschaften*

1. *$Ax = A_0 x$ für alle $x \in X_0$.*
2. *$\|A\|_{L(X,Y)} = \|A_0\|_{L(X_0,Y)}$.*

Beweis. Sei $x \in X$. Aufgrund der Dichtheit von X_0 existiert eine Folge $(x_n)_n \subset X_0$ mit $x_n \to x$ in X. Dann gilt wegen

$$\|A_0 x_n - A_0 x_m\|_Y \leq \|A_0\|_{L(X_0,Y)} \|x_n - x_m\|_X \quad \text{für alle } n, m \in \mathbb{N},$$

dass $(A_0x_n)_n$ eine Cauchyfolge in Y ist.

Nun ist Y vollständig, also existiert

$$\lim_{n\to\infty} A_0x_n =: Ax.$$

Nachweis der Eigenschaften.

1. <u>Ax ist wohldefiniert:</u> Der Wert Ax hängt nicht von der gewählten Folge $(x_n)_n$ ab, denn ist $(\tilde{x}_n)_n$ eine weitere Folge mit $\tilde{x}_n \to x$ in X, so existiert

$$\lim_{n\to\infty} A_0\tilde{x}_n \text{ in Y}$$

   und aus

$$\|A_0x_n - A_0\tilde{x}_n\|_Y \leq \|A_0\|_{L(X_0,Y)}\|x_n - \tilde{x}_n\|_X \overset{n\to\infty}{\longrightarrow} 0$$

   folgt

$$Ax = \lim_{n\to\infty} A_0x_n = \lim_{n\to\infty} A_0\tilde{x}_n.$$

2. <u>A ist Fortsetzung:</u> Für $x \in X_0$ setze einfach $x_n = x$ für alle $x \in \mathbb{N}$. Wegen der Wohldefiniertheit von A und der Stetigkeit von A_0 folgt dann

$$Ax = \lim A_0x_n = A_0x \Longrightarrow A|_{X_0} = A_0.$$

3. <u>A ist linear:</u> Für $\alpha, \beta \in \mathbb{K}$ und $x, y \in X$ gilt

$$\begin{aligned} A(\alpha x + \beta y) &= \lim A_0(\alpha x_n + \beta y_n) \overset{A_0 \text{ linear}}{=} \lim(\alpha A_0(x_n) + \beta A_0(y_n)) \\ &= \alpha \lim A_0x_n + \beta \lim A_0y_n = \alpha Ax + \beta Ay. \end{aligned}$$

4. <u>A ist stetig und normerhaltend:</u> Auf der einen Seite ist

$$\|A\|_{L(X,Y)} = \sup_{x\in B_X} \|Ax\|_Y \geq \sup_{x\in B_{X_0}} \|A_0x\|_Y = \|A_0\|_{L(X_0,Y)}.$$

   Auf der anderen Seite hat man für beliebiges $x \in X$, d.h. $(x_n)_x \subset X_0$ mit $x_n \to X$ in X die Abschätzung

$$\|Ax\|_{L(X,Y)} = \lim_{n\to\infty} \|A_0x_n\|_Y \leq \lim_{n\to\infty} \|A_0\|_{L(X_0,Y)}\|x_n\|_X = \|A_0\|_{L(X_0,Y)}\|x\|_X.$$

   Nach Supremumsübergang ergibt sich

$$\|A\|_{L(X,Y)} \leq \|A_0\|_{L(X_0,Y)}.$$

   Zusammen ist also

$$\|A\|_{L(X,Y)} = \|A_0\|_{L(X_0,Y)}.$$

   Daraus folgt auch die Stetigkeit von A. □

Bemerkung. Falls X_0 **kein dichter** Untervektorraum von $(X, \|.\|_X)$ ist und $A_0 \in L(X_0, Y)$ sowie $(Y, \|.\|_Y)$ ein Banachraum ist, dann kann im Allgemeinen A_0 <u>nicht</u> zu einem stetigen Operator $A \in L(X, Y)$ fortgesetzt werden.

Ganz anders ist die Situation, wenn $Y = \mathbb{K}$ ist. Das führt dann zu den Fortsetzungssätzen von Hahn-Banach, die wir im Abschnitt 3 kennenlernen werden.

Zum Schluss dieses Abschnitts möchte ich noch einige schöne Folgerungen aus dem Satz über die offene Abbildung vorstellen:

Korollar 2.2. *Seien $(X, \|.\|_X)$, $(Y, \|.\|_Y)$ Banachräume und $A \in L(X, Y)$ bijektiv. Dann ist A invertierbar.*

Beweis. Nach dem Satz von der offenen Abbildung (Satz 2.6) ist A offen. Daher sind die Urbilder offener Teilmengen wieder offen, also A^{-1} stetig. □

Zusatz zum Korollar.
Durch Lemma 2.2 ist es möglich, $\|A^{-1}\|$ abzuschätzen:

Sei $r > 0$, so dass (2.12) erfüllt ist, d.h.

$$B_Y(0,r) \subset A(B_X(0,1)).$$

Dann gilt für $y \in Y$ mit $\|y\| < 1$ gerade $ry \in B_Y(0,r)$ und es gibt ein $x \in X$ mit $\|x\| < 1$, so dass $ry = Ax$, d.h.

$$\|A^{-1}y\| = \frac{\|x\|}{r} < \frac{1}{r} \Longrightarrow \|A^{-1}\| = \sup_{y \in B_Y} \|A^{-1}y\| \leq \frac{1}{r}.$$

Korollar 2.3. *Sei X ein Vektorraum, der bezüglich zweier Normen $\|.\|_1$ und $\|.\|_2$ vollständig ist. Es gebe ein $C > 0$, so dass*

$$\|x\|_2 \leq C\|x\|_1 \quad \textit{für alle } x \in X. \tag{2.17}$$

Dann sind die Normen $\|.\|_1$ und $\|.\|_2$ schon äquivalent, d.h. es existiert ein $D > 0$ mit

$$D\|x\|_1 \leq \|x\|_2 \quad \textit{für alle } x \in X \tag{2.18}$$

Beweis. Betrachte die identische Abbildung

$$I : (X, \|.\|_1) \to (X, \|.\|_2).$$

Es ist klar, dass I bijektiv sowie linear und wegen (2.17) auch stetig ist. Nach Korollar 2.2 ist dann auch die Umkehrabbildung

$$I : (X, \|.\|_2) \to (X, \|.\|_1)$$

stetig, d.h. aber gerade, dass es ein $M > 0$ gibt mit

$$\|Ix\|_1 = \|x\|_1 \leq M\|x\|_2 \quad \text{für alle } x \in X.$$

Also erfüllt $D := \frac{1}{M}$ die Ungleichung (2.18). □

Korollar 2.4. *Seien X und Y Banachräume. Fener sei $A \in L(X,Y)$ injektiv. Dann ist A^{-1} als Operator von* $\operatorname{ran}(A)$ *nach X genau dann stetig, wenn* $\operatorname{ran}(A)$ *abgeschlossen ist.*

Beweis. Zu „$\Longrightarrow$“: Sei also A^{-1} stetig. Dann ist

$$A : X \to \operatorname{ran}(A)$$

ein Isomorphismus. Mit X ist dann auch $\operatorname{ran}(A)$ vollständig und daher ist $\operatorname{ran}(A)$ nach Satz 1.1 auch abgeschlossen.

Zu „$\Longleftarrow$“: Sei $\operatorname{ran}(A)$ abgeschlossen. Dann ist $\operatorname{ran}(A)$ ein Banachraum und die Abbildung

$$A : X \to \operatorname{ran}(A)$$

ist bijektiv. Nach Korollar 2.2 ist daher A^{-1} stetig. □

2.3.3 Der Satz vom abgeschlossenen Graphen

Als weitere Folgerung aus dem Satz über die offene Abbildung erhalten wir den

Satz 2.8 (**vom abgeschlossenen Graphen**). *Seien* $(X, \|.\|_X)$, $(Y, \|.\|_Y)$ *Banachräume und* $A : X \to Y$ *linear mit folgender Eigenschaft*

$$\left.\begin{array}{c} \forall_{(x_n)_n \subset X} : x_n \to x \text{ in } X \\ Ax_n \to y \text{ in } Y \end{array}\right\} \Longrightarrow \begin{array}{l} x \in X = dom(A), \\ y = Ax. \end{array} \tag{2.19}$$

Dann gilt $A \in L(X, Y)$

Bemerkungen.

1. (2.19) ist äquivalent zu:

   A besitzt einen abgeschlossenen Graphen, d.h.

   $$G(A) := \{ (x, Ax) \mid x \in X \} \subset X \times Y$$

   ist eine abgeschlossene Teilmenge von $X \times Y$, wobei der Produktraum ausgestattet sei mit der so genannten *Graphennorm*

   $$\|(x, y)\|_{X \times Y} = \|x\|_X + \|y\|_Y.$$

   Achtung. $(X \times Y, \|.\|_{X \times Y})$ ist wieder ein Banachraum, wenn $(X, \|.\|_X)$ und $(Y, \|.\|_Y)$ Banachräume sind (Vergleiche Beweis von Satz 2.8).

   Beweis der Äquivalenz. Zu „$\Longrightarrow$“: Sei $(x_n) \subset X$ eine Folge mit

   $$(x_n, Ax_n) \overset{\|.\|_{X \times Y}}{\longrightarrow} (x, y).$$

   Wir müssen zeigen, dass $(x, y) \in G(A)$. Nun ist aber

   $$x_n \to x, \quad Ax_n \to y,$$

   d.h. mit (2.19) folgt $y = Ax$, also $(x, y) \in G(A)$.

   Zu „$\Longleftarrow$“: Trivial. □

2. Beachten Sie, den Unterschied von (2.19) zur Stetigkeit von A:

   $$\forall_{(x_n)_n \subset X} : x_n \to x \text{ in } X \Longrightarrow \begin{cases} Ax_n \to y \text{ in } Y, \\ y = Ax. \end{cases}$$

3. Jeder Operator $A \in L(X, Y)$ besitzt einen abgeschlossenen Graphen, denn A ist stetig.

4. Es ist im Allgemeinen leichter zu zeigen, dass ein linearer Operator A einen abgeschlossenen Graphen besitzt als direkt zu zeigen, dass A stetig ist.

Definition 2.3 (**Abgeschlossen**). Ein linearer Operator $A : X \to Y$ mit abgeschlossenem Graphen nennt man kurz *abgeschlossen*.

Beweis des Satzes. Wir wollen Korollar 2.3 anwenden. Statte dazu X mit den folgenden Normen aus:

$$\|x\|_1 = \|x\|_X + \|Ax\|_Y \quad \text{und} \quad \|x\|_2 = \|x\|_X \quad \text{für alle } x \in X$$

Dann ist $\|x\|_2 \leq \|x\|_1$.

Beachte, dass $(X, \|.\|_1)$ vollständig ist:
In der Tat: Sei $(x_n)_n$ eine Cauchyfolge in $(X, \|.\|_1)$:

$$\forall_{\epsilon>0} \exists_{N\in\mathbb{N}} \forall_{n,m\geq N} : \|x_n - x_m\|_1 = \|x_n - x_m\|_X + \|A(x_n - x_m)\|_Y < \epsilon.$$

Dann gilt erst recht

$$\forall_{\epsilon>0} \exists_{N\in\mathbb{N}} \forall_{n,m\geq N} : \|x_n - x_m\|_X, \|A(x_n - x_m)\|_Y < \epsilon.$$

Also ist $(x_n)_n$ Cauchyfolge in $(X, \|.\|_X)$ und $(Ax_n)_n$ eine Cauchyfolge in $(Y, \|.\|_Y)$. Da $(X, \|.\|_X)$ und $(Y, \|.\|_Y)$ Banachräume sind, folgt

$$x_n \to x \text{ in } X, \quad Ax_n \to y \text{ in } Y.$$

Nun ist A abgeschlossen, so dass $y = Ax$ folgt. Daher konvergiert die Cauchyfolge $(x_n)_n$ in $(X, \|.\|_1)$:

$$\|x_n - x\|_1 = \underbrace{\|x_n - x\|_X}_{\to 0} + \|Ax_n - Ax\|_Y \to \|y - Ax\|_Y = 0.$$

Weiter im eigentlichen Beweis: Aus Korollar 2.3 folgt nun, dass es ein $D > 0$ gibt mit

$$D\|x\|_1 = D(\|x\|_X + \|Ax\|_Y) \leq \|x\|_2 = \|x\|_X \quad \text{für alle } x \in X.$$

Das ist äquivalent zu

$$\|x\|_X + \|Ax\|_Y \leq \frac{1}{D}\|x\|_X \quad \text{für alle } x \in X.$$

Also gilt erst Recht

$$\|Ax\|_Y \leq \frac{1}{D}\|x\|_X \quad \text{für alle } x \in X,$$

d.h. A ist stetig. □

Bemerkung. Manchmal ist A nur auf einem Untervektorraum D von X definiert. Man schreibt dafür $D = \mathrm{dom}(A)$ bzw. $A : X \supset \mathrm{dom}(A) \to Y$. An den entsprechenden Stellen ist dann X durch D zu ersetzen.

Beispiel 2.17. Seien $X = Y = C[-1, 1]$ und $D = C^1[-1, 1]$. Betrachte den Ableitungsoperator

$$A : D \to Y, \quad Ax = \dot{x}.$$

Dann ist A abgeschlossen, da man nach *Analysis 1* Grenzwertbildung und Ableitung vertauschen darf.

□

3 Die Sätze von Hahn-Banach und ihre Konsequenzen

- Die verschiedenen Sätze von Hahn-Banach: Version der linearen Algebra: reell und komplex, Fortsetzungsversion für stetige Funktionale
- Wichtige Folgerungen: Trennung von Punkten, Normformel
- Die zwei wichtigen Trennungssätze für konvexe Mengen durch stetige lineare Funktionale
- Reflexive Räume und kanonische Abbildung, insbesondere Identifizierungen
- Reflexive Räume: l^p, L^p für $1 < p < \infty$, Hilberträume, endlich-dimensionale Vektorräume
- Nicht-reflexive Räume: c_0, l^∞, l^1, $L^1(\Omega)$, $L^\infty(\Omega)$
- Schwache- und schwach-* Konvergenz
- Adjungierte und kompakte Operatoren

3.1 Die Fortsetzungssätze von Hahn-Banach

Definition 3.1 (Sublinear). Sei X ein Vektorraum über $\mathbb{K}$. Dann heißt $p : X \to \mathbb{R}$ *sublinear*, wenn

(i) $p(\lambda x) = \lambda p(x)$ für alle $\lambda \geq 0$ und $x \in X$.

(ii) $p(x + y) \leq p(x) + p(y)$ für alle $x, y \in X$.

Beispiele 3.1.

1. Jede Halbnorm ist sublinear.
2. Jede lineare Abbildung auf einem reellen Vektorraum ist sublinear.
3. $(t_n) \mapsto \limsup t_n$ ist sublinear auf dem reellen Vektorraum l^∞; $(t_n) \mapsto \limsup \mathrm{Re}(t_n)$ ist sublinear auf dem komplexen Raum l^∞.

□

Wir kommen nun zum 1. Fortsetzungssatz von Hahn-Banach algebraischer Natur:

Satz 3.1 (Hahn-Banach: Version der linearen Algebra). *Sei X ein Vektorraum über $\mathbb{R}$, $F \subset X$ ein Untervektorraum sowie $p : X \to \mathbb{R}$ sublinear. Sei $\varphi : F \to \mathbb{R}$ linear mit*

$$\varphi(x) \leq p(x) \quad \text{für alle } x \in F.$$

Dann existiert eine lineare Fortsetzung

$$\phi : X \to \mathbb{R}, \quad \phi|_F = \varphi$$

mit

$$\phi(x) \leq p(x) \quad \text{für alle } x \in X. \tag{3.1}$$

Die Eigenschaft (3.1) heißt auch, dass ϕ von p dominiert wird.

Beweis. <u>1. Schritt:</u> Sei $x_0 \in X \setminus F$. Setze zunächst φ fort auf

$$F_1 := F \oplus \mathbb{R}x_0 = F \oplus \{\, \lambda x_0 \mid \lambda \in \mathbb{R} \,\}$$

durch

$$\phi(x + \lambda x_0) := \varphi(x) + \alpha\lambda, \quad x + \lambda x_0 \in F \oplus \mathbb{R}x_0,$$

wobei $\alpha \in \mathbb{R}$ so gewählt werden muss, dass ϕ (3.1) erfüllt, d.h.

$$\phi(y) \leq p(y) \quad \text{für alle } y \in F_1.$$

In der Tat ist ϕ eine Fortsetzung, denn

$$\phi|_F \overset{(\lambda=0)}{=} \varphi.$$

Weiter ist ϕ linear: Seien dazu $a, b \in \mathbb{R}$, $\lambda, \mu \in \mathbb{R}$ und $x, y \in F$. Dann gilt

$$\begin{aligned}
\phi(a(x + \lambda x_0) + b(y + \mu x_0)) &= \phi(ax + by + x_0(a\lambda + b\mu)) \\
&\overset{\text{Def}}{=} \varphi(ax + by) + (a\lambda + b\mu)\alpha \\
&\overset{\text{linear}}{=} a\varphi(x) + b\varphi(y) + a\lambda\alpha + b\mu\alpha \\
&= a(\varphi(x) + \lambda\alpha) + b(\varphi(y) + \mu\alpha) \\
&= a\phi(x + \lambda x_0) + b\phi(y + \mu x_0).
\end{aligned}$$

Nun zur Wahl von α:
Beachte, dass man für $\lambda = 0$ kein α wählen braucht. Für alle $x \in F$ und alle $\lambda \in \mathbb{R} \setminus \{0\}$ soll gelten

$$\begin{aligned}
&\phi(x + \lambda x_0) = \varphi(x) + \lambda\alpha \overset{!}{\leq} p(x + \lambda x_0) \\
\iff &\phi(\tfrac{x}{|\lambda|} + \tfrac{\lambda}{|\lambda|}x_0) = \varphi(\tfrac{x}{|\lambda|}) + \tfrac{\lambda}{|\lambda|}\alpha \leq p(\tfrac{x}{|\lambda|} + \tfrac{\lambda}{|\lambda|}x_0)
\end{aligned}$$

Für $\lambda > 0$ ergibt sich daraus für alle $y = \frac{x}{|\lambda|} \in F$

$$\phi(y + x_0) = \varphi(y) + \alpha \leq p(y + x_0)$$

und für $\lambda < 0$

$$\phi(y - x_0) = \varphi(y) - \alpha \leq p(y - x_0).$$

Beides zusammen heißt gerade

$$\varphi(y) - p(y - x_0) \leq \alpha \leq p(y + x_0) - \varphi(y) \quad \text{für alle } y \in F.$$

Weil für alle $x, y \in F$

$$\begin{aligned}
\varphi(x) + \varphi(y) &\overset{\varphi \text{ linear}}{=} \varphi(x + y) \leq p(x + y) = p(x - x_0 + x_0 + y) \\
&\overset{\text{sublinear}}{\leq} p(x - x_0) + p(y + x_0),
\end{aligned}$$

d.h.

$$\varphi(x) - p(x - x_0) \leq p(y + x_0) - \varphi(y) \quad \text{für alle } x, y \in F,$$

gilt

$$\sup_{x \in F}(\varphi(x) - p(x - x_0)) \leq \inf_{y \in F}(p(y + x_0) - \varphi(y)).$$

Somit erfüllt jedes $\alpha \in \mathbb{R}$ mit

$$\sup_{x \in F}(\varphi(x) - p(x - x_0)) \leq \alpha \leq \inf_{y \in F}(p(y + x_0) - \varphi(y))$$

die gesuchte Eigenschaft.

<u>2. Schritt:</u> Wir wollen das Lemma von Zorn anwenden:

Sei $(M, \leq)$ eine teilweise geordnete nichtleere Menge, in der jede Kette (das ist eine total geordnete Teilmenge, also eine Teilmenge, für deren Elemente stets $a \leq b$ oder $b \leq a$ gilt) eine obere Schranke besitzt. Dann liegt jedes Element von M unter einem maximalen Element von M, also einem Element m mit $m \leq a \Longrightarrow a = m$.

Betrachte die Menge M der Paare (G, ϕ), wobei

- G Untervekttorraum von X,
- $F \subset G$,
- $\phi : G \to \mathbb{R}$ lineare Fortsetzung von φ mit $\phi(x) \leq p(x)$ für alle $x \in G$.

Es ist $M \neq \emptyset$, da $(F, \varphi) \in M$.

Auf M können wir eine teilweise Ordnung $\leq$ definieren durch

$$(G_1, \phi_1) \leq (G_2, \phi_2) :\Longleftrightarrow G_1 \subset G_2, \quad \phi_2|_{G_1} = \phi_1.$$

In der Tat ist (M, ϕ) eine teilweise geordnete Menge, in der jedes Kette eine obere Schranke in M besitzt:
Ist etwa $K = \{ (G_i, \phi_i) \mid i \in I \}$ eine solche Kette, dann ist (G, ϕ) definiert durch

$$G := \bigcup_{i \in I} G_i, \quad \phi(x) := \phi_i(x), \quad \text{falls } x \in G \tag{3.2}$$

für alle $x \in G$ eine obere Schranke, d.h.

$$(G_i, \phi_i) \leq (G, \phi) :\Longleftrightarrow G_i \subset G, \quad \phi|_{G_i} = \phi_i.$$

Beachte, dass ϕ in (3.2) wohldefiniert ist, da K total geordnet ist.

Nach dem Lemma von Zorn besitzt die Menge M ein maximales Element (G, ϕ).

Behauptung.

$$G = X \text{ (und die gesuchte Fortsetzung von } \varphi \text{ ist gerade } \phi).$$

Beweis davon: Angenenommen, es ist $G \subsetneq X$.

Dann können wir durch den ersten Schritt eine echte lineare Fortsetzung ϕ_1 von ϕ auf

$$G_1 := G \oplus \mathbb{R}x_0, \quad x_0 \notin G,$$

definieren, die durch p dominiert wird. Dann wäre aber

$$(G, \phi) \leq (G_1, \phi_1)$$

und das wäre ein Widerspruch zur Maximalität von (G, ϕ). $\square$

Bemerkung. Im Allgemeinen ist der Parameter α im 1. Schritt nicht eindeutig bestimmt. Es kann daher durchaus unendlich-viele lineare, von p dominierte Fortsetzungen von φ geben.

Als nächstes wollen wir eine komplexe Version des Satzes von Hahn-Banach formulieren. Zum Beweis ist das folgende Lemma sehr hilfreich:

Lemma 3.1. *Sei X ein $\mathbb{C}$-Vektorraum.*

(i) Ist $f : X \to \mathbb{R}$ eine $\mathbb{R}$-lineares Abbildung, so ist

$$\varphi_f : X \to \mathbb{C}, \quad \varphi_f(x) = f(x) - if(ix), \quad x \in X$$

eine $\mathbb{C}$-lineare Abbildung und $\operatorname{Re}(\varphi_f) = f$.

(ii) Ist $\varphi : X \to \mathbb{C}$ eine $\mathbb{C}$-lineare Abbildung, so ist

$$f = \operatorname{Re}(\varphi) : X \to \mathbb{R}$$

eine $\mathbb{R}$-lineare Abbildung und $\varphi = \varphi_f$.

(iii) Ist $p : X \to \mathbb{R}$ eine Halbnorm und $\phi : X \to \mathbb{C}$ eine $\mathbb{C}$-lineare Abbildung, so gilt die Äquivalenz

$$|\phi(x)| \leq p(x) \ \ \forall_{x \in X} \quad \Longleftrightarrow \quad |\operatorname{Re}(\phi(x))| \leq p(x) \ \ \forall_{x \in X}.$$

(iv) Ist X ein normierter Raum und $\phi : X \to \mathbb{C}$ eine $\mathbb{C}$-lineare, stetige Abbildung, so ist

$$\|\phi\| = \|\operatorname{Re}\phi\|.$$

Mit anderen Worten: $\phi \mapsto \operatorname{Re}\phi$ ist eine bijektive $\mathbb{R}$-lineare Abbildung zwischen dem Raum der $\mathbb{C}$-linearen und dem der $\mathbb{R}$-wertigen $\mathbb{R}$-linearen Funktionale. Im normierten Fall ist sie isometrisch.

Beweis. Zu (i): Es ist klar, dass

$$\varphi_f \ \mathbb{R}\text{-linear}, \quad \operatorname{Re}(\varphi_f) = f.$$

Es reicht daher zu zeigen, dass für alle $x \in X$

$$\varphi_f(ix) = i\varphi_f(x).$$

Nun gilt für alle $x \in X$

$$\begin{aligned} \varphi_f(ix) &= f(ix) - if(-x) \overset{\mathbb{R}\text{-linear}}{=} -i^2(f(ix) + if(x)) \\ &= i(f(x) - if(ix)) = i\varphi_f(x). \end{aligned}$$

Zu (ii): Da φ $\mathbb{C}$-linear ist, ist φ erst recht $\mathbb{R}$-linear. Dann ist aber auch $f = \operatorname{Re}(\varphi)$ $\mathbb{R}$-linear. Außerdem gilt für alle $x \in X$

$$\begin{aligned} \operatorname{Re}(\varphi(ix)) + i\operatorname{Im}(\varphi(ix)) &\overset{\text{Def}}{=} \varphi(ix) \overset{\mathbb{C}\text{-linear}}{=} i\varphi(x) \\ &= i(\operatorname{Re}(\varphi(x)) + i\operatorname{Im}(\varphi(x))) = i\operatorname{Re}(\varphi(x)) - \operatorname{Im}(\varphi(x)). \end{aligned}$$

Daraus folgt

$$\operatorname{Im}(\varphi(x)) = -\operatorname{Re}(\varphi(ix)) \quad \text{für alle } x \in X.$$

Daher gilt für alle $x \in X$

$$\varphi(x) = \operatorname{Re}(\varphi(x)) + i\operatorname{Im}(\varphi(x)) = \operatorname{Re}(\varphi(x)) - i\operatorname{Re}(\varphi(x)) = f(x) - if(ix) = \varphi_f(x).$$

Zu (iii): Die Richtung „$\Longrightarrow$“ ist trivial, da $\operatorname{Re}(\phi) \leq |\phi|$ gilt. Für die „$\Longleftarrow$“-Richtung benutze die Aussage

$$\forall_{z \in \mathbb{C}} \exists_{\theta \in \mathbb{R}} : \mathrm{e}^{i\theta} z = |z|.$$

Dann gilt nämlich

$$\begin{aligned} |\phi(x)| = \mathrm{e}^{i\theta}\phi(x) &\overset{\mathbb{C}\text{-linear}}{=} \phi(\mathrm{e}^{i\theta}x) \overset{\text{wegen } |\cdot|}{=} \operatorname{Re}(\phi(\mathrm{e}^{i\theta})) \\ &\leq p(\mathrm{e}^{i\theta}x) = \left|\mathrm{e}^{i\theta}\right| p(x) = p(x). \end{aligned}$$

Zu (iv): Zunächst gilt

$$\|\phi\| = \sup_{x \in B_X} |\phi(x)| \geq \sup_{x \in B_X} |\mathrm{Re}(\phi(x))| = \|\mathrm{Re}\,\phi\|.$$

Wähle nun

$$p(x) := |\mathrm{Re}(\phi(x))|\,.$$

Dann gilt trivialerweise $|\mathrm{Re}(\phi(x))| \leq p(x)$ und aus (iii) folgt somit

$$|\phi(x)| \leq p(x) = |\mathrm{Re}(\phi(x))|$$

also

$$|\phi(x)| \leq |\mathrm{Re}(\phi(x))|\,.$$

Zusammen folgt

$$\|\phi\| = \|\mathrm{Re}(\phi)\|.$$

□

Satz 3.2 (**Hahn-Banach: Version der linearen Algebra - Komplexe Fassung**). *Sei X ein $\mathbb{C}$-Vektorraum, $p : X \to \mathbb{R}$ sublinear, F ein Untervektorraum von X und $\varphi : F \to \mathbb{C}$ eine $\mathbb{C}$-lineare Abbildung mit*

$$\mathrm{Re}(\varphi(x)) \leq p(x) \quad \textit{für alle } x \in F.$$

Dann existiert eine $\mathbb{C}$-lineare Abbildung $\phi : X \to \mathbb{C}$, $\phi|_F = \varphi$, mit

$$\mathrm{Re}(\phi(x)) \leq p(x) \quad \textit{für alle } x \in X.$$

Beweis. Wende Satz 3.1 an auf $\mathrm{Re}(\varphi) : F \to \mathbb{R}$. Es folgt, dass $\mathrm{Re}(\varphi)$ eine $\mathbb{R}$-lineare Fortsetzung $f : X \to \mathbb{R}$ besitzt mit

$$f(x) \leq p(x) \quad \text{für alle } x \in X.$$

Nach Lemma 3.1 (i) ist

$$\phi : X \to \mathbb{C}, \quad \phi(x) := f(x) - if(ix), \quad x \in \mathbb{C}$$

$\mathbb{C}$-linear mit $\mathrm{Re}(\phi) = f \leq p$ und das zeigt die letzte Abschätzung im Satz.

Wir betrachten nun die $\mathbb{C}$-lineare Abbildung $\varphi : F \to \mathbb{C}$. Nach Lemma 3.1 (ii) ist für $x \in F$

$$\varphi(x) = \varphi_{\mathrm{Re}(\varphi)}(x) = \mathrm{Re}(\varphi(x)) - i\,\mathrm{Re}(\varphi(ix)).$$

Wegen $f|_F = \mathrm{Re}(\varphi)$ folgt daraus für $x \in F$

$$\varphi(x) = \mathrm{Re}(\varphi(x)) - i\,\mathrm{Re}(\varphi(ix)) = f(x) - if(ix) = \phi(x)$$

und daher wie gewünscht $\phi|_F = \varphi$. □

Bemerkung. Falls p im Satz 3.2 sogar eine Halbnorm ist, d.h. es gilt sogar

$$p(\lambda x) = |\lambda|\,p(x) \quad \text{für alle } x \in X, \lambda \in \mathbb{C},$$

dann ist die Abschätzung

$$\mathrm{Re}(\phi(x)) \leq p(x) \quad \text{für alle } x \in X$$

wegen Lemma 3.1 (iii) äquivalent zu

$$|\phi(x)| \leq p(x) \quad \text{für alle } x \in X.$$

Endlich ist es soweit und wir können die algebraischen Hahn-Banach-Sätze auf normierte Räume und stetige lineare Funktionale anwenden:

Satz 3.3 (**Hahn-Banach: Fortsetzungsversion**). *Sei* $(X, \|.\|_X)$ *ein normierter* $\mathbb{K}$-*Vektorraum,* F *ein Untervektorraum von* X *und* $\varphi \in F' = L(F, \mathbb{K})$. *Dann existiert eine Fortsetzung* $\phi \in X'$ *von* φ *mit*

$$\|\phi\|_{X'} = \|\varphi\|_{F'}.$$

Jedes stetige Funktional kann also normgleich fortgesetzt werden.

Beweis. Wende Satz 3.2 und seine nachfolgende Bemerkung an auf die Halbnorm

$$p : X \to \mathbb{R}, \quad p(x) = \|\varphi\|_{F'} \|x\|_X, \quad x \in X.$$

Wir erhalten so eine $\mathbb{K}$-lineare Fortsetzung $\phi : X \to \mathbb{K}$ von φ mit

$$|\phi(x)| \leq p(x) = \|\varphi\|_{F'} \|x\|_X \quad \text{für alle } x \in X.$$

Insbesondere sieht man daran, dass ϕ stetig ist, d.h. $\phi \in X'$. Weiter gilt

$$\|\phi\|_{X'} = \sup_{x \in B_X} |\phi(x)| \leq \|\varphi\|_{F'}.$$

Außerdem ist

$$\|\phi\|_{X'} = \sup_{x \in B_X} \|\phi(x)\| \geq \sup_{x \in B_F} \|\phi(x)\| = \sup_{x \in B_F} \|\varphi(x)\| = \|\varphi\|_{F'},$$

d.h. $\|\phi\|_{X'} = \|\varphi\|_{F'}$. □

Bemerkung. Eine solche Fortsetzung ϕ ist im Allgemeinen nicht eindeutig bestimmt und der Satz 3.3 gilt für allgemeine Operatoren nicht, wie wir später sehen werden.

Wir kommen nun zu wichtigen Folgerungen für Funktionale. In manchen Lehrbüchern finden sie oft auch $x' \in X'$ als Funktionalelement statt $\varphi \in X'$:

Korollar 3.1. *Seien* $(X, \|.\|_X)$ *ein normierter* $\mathbb{K}$-*Vektorraum und* $x \in X$ *mit* $x \neq 0$. *Dann existiert ein Funktional* $\varphi \in X'$ *mit*

$$\varphi(x) = \|x\|_X \quad \textit{und} \quad \|\varphi\|_{X'} = 1.$$

Beweis. Betrachte

$$F := \mathbb{K} \cdot x = \{ \lambda x \mid \lambda \in \mathbb{K} \}.$$

Definiere dann

$$\varphi : F \to \mathbb{K}, \quad \varphi(\lambda x) = \lambda \|x\|_X \quad \text{für alle } \lambda \in \mathbb{K}.$$

Dann ist φ eine $\mathbb{K}$-lineare Abbildung: Seien $\alpha, \beta \in \mathbb{K}$ und $y_1, y_2 \in F$. Es folgt

$$\begin{aligned} \varphi(\alpha y_1 + \beta y_2) &= \varphi(\alpha\lambda_1 x + \beta\lambda_2 x) = \varphi(x(\alpha\lambda_1 + \beta\lambda_2)) \\ &= (\alpha\lambda_1 + \beta\lambda_2)\|x\| = \alpha\lambda_1\|x\| + \beta\lambda_2\|x\| \\ &= \alpha\varphi(\lambda_1 x) + \beta\varphi(\lambda_2 x) = \alpha\varphi(y_1) + \beta\varphi(y_2). \end{aligned}$$

Weiter ist

$$\|\varphi\|_{F'} = \sup_{\substack{y = \lambda x \in F \\ \|\lambda x\| \leq 1}} |\varphi(y)| = \sup_{\substack{y = \lambda x \in F \\ \|\lambda x\| \leq 1}} \|\lambda x\|_X \leq 1.$$

Auf der anderen Seite ist

$$1 = \frac{1}{\|x\|}\|x\| = \varphi(\underbrace{\frac{x}{\|x\|}}_{\in B_F}) \leq \sup_{y \in B_F} |\varphi(y)| = \|\varphi\|_{F'}.$$

Demnach ist

$$\|\varphi\|_{F'} = 1.$$

Also ist $\varphi \in F'$. Nach Satz 3.3 gibt es eine Fortsetzung $\phi \in X'$ von φ mit

$$\|\phi\|_{X'} = \|\varphi\|_{F'} = 1.$$

Außerdem ist

$$\phi(x) = \|x\|_X,$$

denn für $\lambda \in \mathbb{K}$ mit $\lambda \neq 0$ gilt

$$\phi(x) = \frac{\lambda}{\lambda}\phi(x) = \frac{1}{\lambda}\phi(\lambda x) = \frac{1}{\lambda}\varphi(\lambda x) = \|x\|$$

und für $x = 0$ ist das sowieso klar. □

Aus Korollar 3.1 folgt unmittelbar, dass die Elemente aus X' die Punkte von X trennen:

Korollar 3.2. *Sei $(X, \|.\|_X)$ ein $\mathbb{K}$-Vektorraum und $x_1, x_2 \in X$ mit $x_1 \neq x_2$. Dann existiert ein Funktional $\varphi \in X'$ mit*

$$\varphi(x_1) \neq \varphi(x_2) \quad \textit{und} \quad \|\varphi\|_{X'} = 1.$$

Beweis. Wende Korollar 3.1 an auf

$$x = x_1 - x_2 \neq 0.$$

Dann existiert ein $\varphi \in X'$ mit $\|\varphi\|_{X'} = 1$ und

$$\varphi(x) = \underbrace{\varphi(x_1 - x_2)}_{=\varphi(x_1)-\varphi(x_2)} = \|x\|_X = \|x_1 - x_2\|_X \overset{x_1 \neq x_2}{\neq} 0.$$

Es folgt

$$\varphi(x_1) \neq \varphi(x_2).$$

□

Bemerkung. Das Element $\varphi \in X'$ mit

$$\|\varphi\|_{x'} = 1 \quad \text{und} \quad \varphi(x) = \|x\|_X$$

aus Korollar 3.1 ist im Allgemeinen nicht eindeutig bestimmt. Aber es es gibt eine Möglichkeit, damit es eindeutig wird:

$$\boxed{\varphi \in X' \text{ ist eindeutig, wenn } X' \text{ strikt konvex ist.}} \tag{3.3}$$

Erinnerung. Ein normierter Vektorraum $(Y, \|.\|_Y)$ heißt *strikt-konvex*, wenn gilt
Für alle $y_1, y_2 \in Y$ mit $\|y_1\|_Y = \|y_2\|_Y = 1$ und $y_1 \neq y_2$ ist

$$\|\lambda y_1 + (1-\lambda)y_2\|_Y < 1 \quad \text{für alle } \lambda \in]0,1[.$$

Interpretation. Die Einheitskugel in einem strikt konvexen normierten Vektorraum ist „rund“.

Beispiele 3.2.

1. Die Einheitskugel in $(X, \|.\|) = (\mathbb{R}^2, \|.\|_2)$ ist strikt konvex:

   Seien $x = (x_1, x_2)$ und $y = (y_1, y_2)$ mit $x \neq y$ und $\|x\|_2 = \|y\|_2 = 1$. Dann gilt

$$\begin{aligned}\|\lambda y + (1-\lambda)x\|_2^2 &= (\lambda y_1 + (1-\lambda)x_1)^2 + (\lambda y_2 + (1-\lambda)x_2)^2 \\ &= \lambda^2(y_1^2 + y_2^2) + 2\lambda(1-\lambda)(x_1y_1 + x_2y_2) + (1-\lambda)^2(x_1^2 + x_2^2) \\ &= \lambda^2\|y\|_2^2 + 2\lambda(1-\lambda)\langle x, y\rangle + (1 - 2\lambda + \lambda^2)\|x\|_2^2 \\ &= 2\lambda^2 - 2\lambda + 1 + 2\lambda(1-\lambda)\langle x, y\rangle \\ &= 1 + \underbrace{(1-\lambda)2\lambda}_{>0}(\langle x, y\rangle - 1).\end{aligned}$$

   Wir müssen also zeigen, dass $\langle x, y\rangle - 1 < 0$, d.h.

$$\langle x, y\rangle < 1 = \|x\|_2^2\|y\|_2^2. \tag{3.4}$$

   Das ist aber (fast) die Cauchy-Schwarzsche Ungleichung. Die Striktheit gilt, wenn x und y linear unabhängig sind.

   Ist also $x = \lambda y$ für ein $\lambda \in \mathbb{R}$, so gilt

$$x_1 = \lambda y_1 \quad \text{und} \quad x_2 = \lambda y_2.$$

   Daraus folgt

$$1 = \|x\|_2^2 = x_1^2 + x_2^2 = \lambda^2(y_1^2 + y_2^2) = \lambda^2\|y\|_2^2 = \lambda^2,$$

   also ist $\lambda = 1$ oder $\lambda = -1$. Doch $\lambda = 1$ können wir ausschließen, da sonst $x = y$ wäre. Für $\lambda = -1$ gilt

$$\langle x, y\rangle = \langle x, -x\rangle = -\langle x, x\rangle = -1 < 1.$$

   Damit ist (3.4) gezeigt und die Einheitskugel strikt konvex.

2. Die Einheitskugel in $(X, \|.\|) = (\mathbb{R}^2, \|.\|_1)$ ist bezüglich der Norm

$$\|(x, y)\|_1 = |x| + |y|$$

   nicht strikt konvex:

   Betrachte etwa $y_1 = (1, 0)$ und $y_2 = (0, 1)$. Dann ist $y_1 \neq y_2$, $\|y_1\|_1 = \|y_2\| = 1$, aber

$$\|\lambda y_1 + (1-\lambda)y_2\|_1 = |\lambda| + |1-\lambda| \overset{\lambda=1/2}{=} 1.$$

Vergleichen Sie auch Beispiel 1.15, um sich von der „Rundheit“ zu überzeugen.

□

Beweis von (3.3). Angenommen, es existieren $\varphi_1, \varphi_2 \in X'$ mit $\varphi_1 \neq \varphi_2$ sowie $\|\varphi_1\|_{X'} = \|\varphi_2\|_{X'} = 1$ und

$$\varphi_1(x) = \varphi_2(x) = \|x\|.$$

Dann folgt für alle $\lambda \in]0, 1[$

$$\begin{aligned}1 &= |\lambda \underbrace{\varphi_1(x/\|x\|)}_{=1} + (1-\lambda)\underbrace{\varphi_2(x/\|x\|)}_{=1}| \leq \sup_{y \in B_X} |\lambda\varphi_1(y) + (1-\lambda)\varphi_2(y)| \\ &= \|\lambda\varphi_1 + (1-\lambda)\varphi_2\|_{X'} \overset{\text{strikt konvex}}{<} 1\end{aligned}$$

Aber $1 < 1$ ist ein Widerspruch. □

Korollar 3.3 (**Normformel**). *In jedem normierten Raum gilt*

$$\|x\|_X = \sup_{\varphi \in B_{X'}} |\varphi(x)| = \max_{\varphi \in B_{X'}} |\varphi(x)| \quad \textit{für alle } x \in X.$$

Beweis. Für $x = 0$ ist die Behauptung trivial. Sei also $x \neq 0$.

Zu „$\geq$“: Es gilt

$$\sup_{\varphi \in B_{X'}} |\varphi(x)| \leq \sup_{\varphi \in B_{X'}} \|\varphi\|_{X'} \|x\|_X \leq \|x\|_X.$$

Zu „$\leq$“: Nach Korollar 3.1 existiert zu $x \in X \setminus \{0\}$ ein $\varphi \in X'$ mit

$$\|\varphi\|_{X'} = 1 \quad \text{und} \quad \varphi(x) = \|x\|.$$

Daraus folgt

$$\|x\| = \varphi(x) \overset{\|\varphi\|=1}{\leq} \sup_{x \in B_{X'}} |\varphi(x)|.$$

□

Bemerke die Symmetrie in der Normformel zur Definition

$$\|\varphi\|_{X'} = \sup_{x \in B_X} |\varphi(x)| \quad \text{für alle } x \in X'.$$

Im Gegensatz hierzu wird das Supremum in der Normformel stets angenommen.

Lemma 3.2. *Seien (X, d) ein metrischer Raum, $U \subset X$ eine abgeschlossene, nichtleere Menge und $x \notin U$. Dann gilt*

$$d(x, U) = \inf\{\, d(x, u) \mid u \in U \,\} > 0.$$

Beweis. Angenommen, es gibt eine Folge $(u_n) \subset U$ mit

$$\lim_{n \to \infty} d(x, u_n) = 0,$$

d.h. $u_n \to x$. Weil U abgeschlossen ist, folgt $x \in U$. Widerspruch! □

Korollar 3.4. *Seien $(X, \|.\|_X)$ ein normierter Raum, U ein abgeschlossener Untervektorraum von X und $x \in X \setminus U$. Dann gilt:*

$$\exists_{\varphi \in X'} : \varphi|_U = 0, \quad \|\varphi\| = 1 \quad \textit{und} \quad \varphi(x) = d(x, U) > 0.$$

Beweis. Setze

$$U_0 := \operatorname{lin}(U \cup \{x\}) = U \oplus \operatorname{lin}\{x\}.$$

Definiere auf U_0 die Abbildung

$$u_0'(u + \alpha x) := \alpha d(x, U) = \alpha \inf\{\, \|x - u\| \mid u \in U \,\} \quad \text{für alle } u \in U \text{ und } \alpha \in \mathbb{K}.$$

Dann ist $u_0' : U_0 \to \mathbb{K}$ linear, weil für alle $u_1, u_2 \in U$ und $\lambda, \mu, \alpha_1, \alpha_2 \in \mathbb{K}$

$$\begin{aligned} u_0'(\lambda(u_1 + \alpha x) + \mu(u_2 + \alpha_2 x)) &= u_0'((\lambda u_1 + \mu u_2) + x(\lambda\alpha_1 + \mu\alpha_2)) \\ &= (\lambda\alpha_1 + \mu\alpha_2) d(x, U) = \lambda\alpha_1 d(x, U) + \mu\alpha_2 d(x, U) \\ &= \lambda u_0'(y_1 + \alpha_1 x) + \mu u_0'(y_2 + \alpha_2 x). \end{aligned}$$

Offensichtlich ist auch $u_0' = 0$ auf U. Ferner gilt für $\alpha \neq 0$

$$d(x, U) \leq \left\| x - \tfrac{-u}{\alpha} \right\|,$$

also

$$|u_0'(u + \alpha x)| \leq |\alpha| \left\| x - \tfrac{-u}{\alpha} \right\| = \|\alpha x + u\|.$$

Daher ist $\|u_0'\| \leq 1$.

Sei nun $\epsilon > 0$. Wir können dann wegen $d(x, U) > 0$ (Lemma 3.2) ein u_ϵ wählen, so dass

$$\|x - u_\epsilon\| \leq (1 + \epsilon) d(x, U).$$

Daraus folgt

$$u_0'(x - u_\epsilon) = 1 \cdot d(x, U) \geq \tfrac{1}{1+\epsilon} \|x - u_\epsilon\|.$$

Wegen $x - u_\epsilon \neq 0$ ist also

$$\|u_0'\| \geq \tfrac{1}{1+\epsilon} \overset{\epsilon \to 0^+}{\longrightarrow} 1.$$

Damit ist insgesamt $\|u_0'\| = 1$ gezeigt und daher ist $u_0' \in L(U_0)$. Nach dem Satz von Hahn-Banach (Satz 3.3) gibt es dann ein $\varphi \in X'$ mit $\varphi|_{U_0} = u_0'|_{U_0}$ und $\|\varphi\| = \|u_0'\|$. Es gilt dann wie gewünscht

$$\varphi|_U \overset{\alpha=0}{=} 0, \quad \|\varphi\| = \|u_0'\| = 1 \quad \text{und} \quad \varphi(x) = u_0'(0 + 1 \cdot x) = d(x, U) > 0.$$

□

Lemma 3.3. *Sei X ein normierter Raum und S ein Untervektorraum mit* $\dim S < \infty$. *Dann ist S abgeschlossen und es gibt einen komplementären abgeschlossenen Untervektorraum T, d.h.*

$$X = S \oplus T.$$

Beweis. Jeder endlich-dimensionale normierte Raum ist vollständig, also ist S abgeschlossen.

Sei $x_1, \ldots, x_n$ eine Basis von S und $e_1, \ldots, e_n$ die duale Basis von S', d.h.

$$e_i(x_j) = \delta_{ij}.$$

Nach dem Satz von Hahn-Banach gibt es Fortsetzungen $f_1, \ldots, f_n \in X'$ von $e_1, \ldots, e_n$, d.h.

$$f_j|_S = e_j.$$

Definiere $P : X \to S$ durch

$$Px = \sum_{i=1}^n f_i(x) x_i. \tag{3.5}$$

Wir bemerken, dass diese Abbildung wohldefiniert ist und offensichtlich ist P linear und stetig. Weiter gilt $P \circ P = P$, denn

$$\begin{aligned}(P \circ P)(x) &= P(\sum_{i=1}^n f_i(x) x_i) = \sum_{j=1}^n f_j(\sum_{i=1}^n f_i(x) x_i) x_j \\ &= \sum_{j=1}^n \sum_{i=1}^n f_i(x) \underbrace{f_j(x_i)}_{=\delta_{ij}} x_j = \sum_{i=1}^n f_i(x) x_i = Px.\end{aligned}$$

Außerdem ist $P|_S = I$, denn für $x \in S$ gilt

$$\begin{aligned}P(x) &= \sum_{i=1}^n f_i(x) x_i = \sum_{i=1}^n f_i(\sum_{j=1}^n \lambda_j x_j) x_i \\ &= \sum_{i=1}^n \sum_{j=1}^n \lambda_j \underbrace{f_i(x_j)}_{=\delta_{ij}} x_i = \sum_{i=1}^n \lambda_j x_j = x.\end{aligned}$$

Setze nun

$$T := \ker P.$$

Dann ist T tatsächlich ein abgeschlossener Untervektorraum von X, denn $T = P^{-1}(\{0\})$ ist abgeschlossen als Urbild der abgeschlossenen Menge $\{0\}$ unter der stetigen Abbildung P. Außerdem gilt $T \cap S = \{0\}$, denn angenommen, es gibt ein $s \in S \cap T$ mit $s \neq 0$. Dann ist $s \in T = \ker P$, d.h. $P(s) = 0$, aber wegen $s \in S$ gilt auch

$$0 = P(s) = s$$

und das ist ein Widerspruch zu $s \neq 0$.

Die gesuchte Zerlegung ist nun

$$x = (x - P(x)) + P(x).$$

$P(x) \in S$ ist dabei sowieso klar. Weiter gilt $x - P(x) \in T$, denn

$$P(x - P(x)) = P(x) - P \circ P(x) = P(x) - P(x) = 0.$$

$\square$

3.2 Geometrische Formen des Satzes von Hahn-Banach

In diesem Abschnitt ist unser Ziel, zwei konvexe Mengen durch eine Hyperebene zu trennen, die durch ein lineares stetiges Funktional beschrieben wird. Wir haben also das folgende

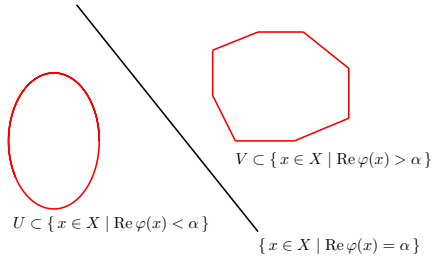
Trennungsproblem:
Existiert zu konvexem $U \subset X$ und $V \subset X$ ein Funktional $\varphi \in X'$ mit $\varphi \neq 0$ und

$$\sup_{x \in U} \varphi(x) \leq \inf_{x \in V} \varphi(x)$$

im Falle $\mathbb{K} = \mathbb{R}$ bzw.

$$\sup_{x \in U} \operatorname{Re} \varphi(x) \leq \inf_{x \in V} \operatorname{Re} \varphi(x)$$

im Falle $\mathbb{K} = \mathbb{C}$?



Ein wichtiges Hilfmittel zum Beweis der folgenden Trennungssätze ist die folgende

Definition 3.2 (Minkowskifunktional). Sei X ein $\mathbb{K}$-Vektorraum und $C \subset X$ eine Teilmenge. Das *Minkowskifunktional*

$$p_C : X \to [0, \infty]$$

wird definiert durch

$$p_C(x) := \inf\{ \alpha > 0 \mid x \in \alpha C \}.$$

C heißt *absorbierend*, falls $p_C(x) < \infty$ für alle $x \in X$.

Beispiel 3.3. Sei $(X, \|.\|_X)$ ein $\mathbb{K}$-Vektorraum und

$$C := \{ x \in X \mid \|x\| < 1 \}.$$

Dann ist

$$p_C(x) = \inf\{ \alpha > 0 \mid x \in \alpha C \} = \inf\{ \alpha > 0 \mid \|x\| < \alpha \} = \|x\|.$$

□

Lemma 3.4. *Sei $(X, \|.\|_X)$ ein normierter $\mathbb{K}$-Vektorraum und $C \subset X$ eine offene, konvexe Teilmenge mit $0 \in C$. Dann gilt*

(i) p_C ist sublinear.

(ii) C ist absorbierend, genauer:

$$\exists_{M \geq 0} \forall_{x \in X} : 0 \leq p_C(x) \leq M\|x\|.$$

(iii) $C = \{ x \in X \mid p_C(x) < 1 \}$.

Beweis. Zu (iii):

- $\subset$: Sei $x \in C$. Da C offen ist, gibt es ein $\epsilon > 0$, so dass

  $$(1 + \epsilon)x \in C.$$

  Also ist

  $$x \in \tfrac{1}{1+\epsilon} C$$

  und damit

  $$p_C(x) \leq \tfrac{1}{1+\epsilon} < 1.$$

- $\supset$: Sei nun $x \in X$ mit $p_C(x) < 1$. Dann gibt es ein $\alpha \in]0,1[$ mit $x \in \alpha C$, also

$$\tfrac{1}{\alpha} x \in C.$$

Nach Voraussetzung ist auch $0 \in C$. Da C konvex ist, folgt

$$x = \alpha \tfrac{1}{\alpha} x + (1-\alpha) \cdot 0 \in C$$

Zu (ii): Es ist klar, dass $p_C(x) \geq 0$, da $\alpha > 0$ in der Infbildung.

Zur Bestimmung von M: Da C offen ist und $0 \in C$, gibt es ein $r > 0$, so dass

$$B_X(0, 2r) \subset C.$$

Für alle $x \in X$ mit $x \neq 0$ ist $r \frac{x}{\|x\|} \in B_X(0, 2r)$, also auch

$$r \tfrac{x}{\|x\|} \in C \Longleftrightarrow x \in \tfrac{\|x\|}{r} C.$$

Daraus folgt

$$p_C(x) \leq \tfrac{\|x\|}{r} \quad \text{für alle } x \in X \tag{3.6}$$

Also ist $M := \frac{1}{r}$ das Gesuchte. Beachte, dass die Ungleichung (3.6) auch für $x = 0$ gilt.

Zu (i): Um die Homogenität nachzuweisen, seien $\lambda > 0$ und $x \in X$. Dann ist

$$\begin{aligned} p_C(\lambda x) &= \inf\{\, \alpha > 0 \mid \lambda x \in \alpha C \,\} = \inf\{\, \alpha > 0 \mid x \in \tfrac{\alpha}{\lambda} C \,\} \\ &= \inf\{\, \lambda\beta > 0 \mid x \in \beta C \,\} = \lambda \inf\{\, \beta > 0 \mid x \in \beta C \,\} = \lambda p_C(x). \end{aligned}$$

Für $\lambda = 0$ ist $p_C(0) = 0$.

Zur Subadditivität: Seien $x, y \in X$ und $\epsilon > 0$. Dann gilt

$$p_C(x) + \epsilon > 0 \quad \text{und} \quad p_C(y) + \epsilon > 0.$$

Dann folgt

$$p_C\left(\tfrac{x}{p_C(x)+\epsilon}\right) = \tfrac{p_C(x)}{p_C(x)+\epsilon} \overset{\epsilon>0}{<} 1$$

und daher wegen (iii)

$$\tfrac{x}{p_C(x)+\epsilon} \in C.$$

Analog folgert man

$$\tfrac{y}{p_C(y)+\epsilon} \in C.$$

Nun ist C konvex und daher

$$\theta \tfrac{x}{p_C(x)+\epsilon} + (1-\theta) \tfrac{y}{p_C(y)+\epsilon} \in C \quad \text{für alle } x, y \in X \text{ und } \theta \in [0,1].$$

Wähle

$$\theta := \tfrac{p_C(x)+\epsilon}{p_C(x)+p_C(y)+2\epsilon} (< 1).$$

Dann ist

$$1 - \theta = \tfrac{p_C(y)+\epsilon}{p_C(x)+p_C(y)+2\epsilon}$$

Einsetzen liefert

$$\tfrac{x}{p_C(x)+p_C(y)+2\epsilon} + \tfrac{y}{p_C(x)+p_C(y)+2\epsilon} \in C.$$

Damit ist

$$x + y \in (p_C(x) + p_C(y) + 2\epsilon) C,$$

d.h.

$$p_C(x+y) \leq p_C(x) + p_C(y) + 2\epsilon.$$

Da $\epsilon > 0$ beliebig war, folgt daraus dann die Subadditivität. $\square$

Konvention.
Seien $A, B \subset X$. Dann definiere

$$A \pm B := \{\, a \pm b \mid a \in A, b \in B \,\}.$$

Ist A sowie B konvex, so ist auch $A \pm B$ komvex.

Beweis. Seien $z_1, z_2 \in A \pm B$, also gibt es $a_1, a_2 \in A$ und $b_1, b_2 \in B$, so dass

$$z_1 = a_1 \pm b_1, \quad z_2 = a_2 \pm b_2.$$

Für $\lambda \in [0,1]$ gilt dann

$$\begin{aligned} \lambda z_1 + (1-\lambda) z_2 &= \lambda(a_1 \pm b_1) + (1-\lambda)(a_2 \pm b_2) \\ &= \underbrace{[\lambda a_1 + (1-\lambda) a_2]}_{\in A,\text{ da konvex}} \pm \underbrace{[\lambda b_1 + (1-\lambda) b_2]}_{\in B,\text{ da konvex}} \in A \pm B. \end{aligned}$$

Also ist $A \pm B$ konvex. □

Nun zum ersten Trennungssatz:

Satz 3.4 (1. Trennungssatz, Trennung im weiteren Sinne). *Sei $(X, \|.\|_X)$ ein normierter $\mathbb{K}$-Vektorraum und $U, V \neq \emptyset$ disjunkte, konvexe Teilmengen von X. Weiter sei U offen. Dann gibt es ein $\varphi \in X'$ mit*

$$\operatorname{Re}\varphi(u) < \operatorname{Re}\varphi(v) \quad \textit{für alle } u \in U \textit{ und } v \in V. \tag{3.7}$$

Bemerkungen.

1. Geometrische Interpretation: Die Hyperebene $\{\, x \in X \mid \operatorname{Re}(\varphi(x)) = \alpha \,\}$ trennt U und V.
2. An dem Satz sieht man auch, dass stets $\varphi \neq 0$ gilt. Denn wäre $\varphi = 0$, so ergäbe sich ein Widerspruch zur Striktheit in (3.7).
3. Dass U offen ist, ist wesentliche Voraussetzung im Falle $\dim X = \infty$, wie das folgende Beispiel zeigen wird:

Beispiel 3.4. Sei $X = c_0$ sowie

$$U := \{\, (x_n)_n : \mathbb{N} \to \mathbb{R} \mid x_n \neq 0 \text{ nur für endliche viele } n \,\}.$$

U ist nicht offen in c_0, denn um $(0, 0, \ldots)$ gibt es keine offene Kugel, die ganz in U liegt, d.h. es gibt wieder und wieder Elemente aus c_0, die nicht in U liegen.

Der Dualraum von c_0 ist $c_0' \cong l^1$ (vgl. die Bemerkung zum Beispiel 2.10) und

$$\varphi(x) = \sum_{n=1}^{\infty} \varphi_n x_n, \quad \varphi \in l^1, x \in c_0.$$

Definiere

$$V := \{\, (1/2)^n \mid n \in \mathbb{N} \,\}.$$

Behauptung: Es gibt kein $\varphi \in X'$, so dass für alle $x \in U$ und $v_n \in V$

$$\varphi(x) = \sum_{n=1}^{\infty} \varphi_n x_n < \sum_{n=1}^{\infty} \varphi_n 2^{-n} = \sum_{n=1}^{\infty} \varphi_n v_n. \tag{3.8}$$

Angenommen, (3.8) gilt doch. Dann betrachte

$$x^{(k)} := (\operatorname{sgn}\varphi_1, \operatorname{sgn}\varphi_2, \ldots, \operatorname{sgn}\varphi_k, 0, \ldots) \in U$$

Dann gilt nämlich

$$\varphi(x^{(k)}) = \sum_{n=1}^{k} |\varphi_n| < \sum_{n=1}^{k} \varphi_n 2^{-k}.$$

Für $k \to \infty$ folgt

$$\sum_{n=1}^{\infty} |\varphi_n| \leq \sum_{n=1}^{\infty} 2^{-n} = \frac{1}{2} \sum_{n=1}^{\infty} \varphi_n 2^{-n+1} \leq \frac{1}{2} \sum_{n=1}^{\infty} |\varphi_n| \,.$$

Das geht aber nur für $\varphi = 0$. Doch das wäre ein Widerspruch zur Striktheit in (3.8).

□

Beweis von Satz 3.4.

<u>1. Schritt:</u> $V = \{x_0\}$.

1. Fall: $\mathbb{K} = \mathbb{R}$.
Wir müssen zeigen, dass es ein $\varphi \in X'$ gibt mit

$$\varphi(x) < \varphi(x_0) \quad \text{für alle } x \in U. \tag{3.9}$$

O.B.d.A. sei $0 \in U$ (denn wir wollen Lemma 3.4 anwenden). Andernfalls betrachte für $u_0 \in U$ beliebig

$$V' := \{x_0\} - \{u_0\} = \{x_0 - u_0\}\,, \quad U' := U - \{u_0\}\,.$$

Dann sind U' und V' wieder konvex (siehe Konvention oben) und U' ist offen, da U' eine Verschiebung der offenen Menge U um u_0 ist. Weiter ist

$$V' \cap U' = \emptyset,$$

denn wäre das nicht so, so gebe es ein $u \in U$ mit

$$x_0 - u_0 = u - u_0 \Longleftrightarrow x_0 = u,$$

doch ein solches u gibt es nicht, da $U \cap V = \emptyset$.

Außerdem gilt auch (3.9):

$$\varphi(x - u_0) < \varphi(x_0 - u_0) \Longleftrightarrow \varphi(x) - \varphi(u_0) < \varphi(x_0) - \varphi(u_0),$$

also

$$\varphi(x) < \varphi(x_0) \quad \text{für alle } x \in U.$$

Wir können also tatsächlich $0 \in U$ annehmen.

Wir betrachten jetzt das Minkowski-Funktional p_U. Für dieses zeigen wir jetzt

$$p_U(x_0) = \inf\{\, \lambda > 0 \mid x_0 \in \lambda U \,\} \geq 1. \tag{3.10}$$

Es ist klar, dass $x_0 \notin 0 \cdot U$ und $x_0 \notin 1 \cdot U$, da $U \cap V = \emptyset$.

Annahme: Es gibt ein $\lambda \in]0, 1[$ mit $x_0 \in \lambda U$, d.h. es gibt dazu ein $u \in U$ mit $x_0 = \lambda u$.

Dann folgt aus der Konvexität von U und da u und $0 \in U$:

$$\lambda u + (1 - \lambda) \cdot 0 = \lambda u = x_0 \in U,$$

aber das ist ein Widerspruch zu $U \cap V = \emptyset$. Damit ist (3.10) gezeigt.

Sei nun

$$F = \mathbb{R} x_0$$

und sei

$$\phi : F \to \mathbb{R}, \quad \phi(\lambda x_0) := \lambda \quad \text{für alle } \lambda \in \mathbb{R}.$$

Wir haben früher schon einmal gesehen, dass ϕ linear ist (vgl. den Beweis zu Korollar 3.1).

Nun gilt

$$\phi(y) \leq p_U(y) \quad \text{für alle } y \in F,$$

denn sei $y := \lambda x_0$. Dann gilt

- Für $\lambda \leq 0$ ist

$$\phi(y) = \phi(\lambda x_0) = \lambda \leq 0 \overset{\textit{Lemma 3.4 (ii)}}{\leq} p_U(\lambda x_0) = p_U(y).$$

- Für $\lambda > 0$ ist

$$\phi(y) = \phi(\lambda x_0) = \lambda \overset{(3.10)}{\leq} \lambda p_U(x_0) \overset{\text{sublinear}}{=} p_U(\lambda x_0) = p_U(y).$$

Nach dem Fortsetzungsatz 3.1 von Hahn-Banach gibt es eine lineare Fortsetzung φ von ϕ, so dass

$$\varphi(y) = \phi(y) \quad \text{für alle } y \in F, \quad \varphi(x) \leq p_U(x) \quad \text{für alle } x \in X.$$

Weiter gilt

$$\varphi(x_0) \overset{x_0 \in F}{=} \phi(x_0) \overset{\text{Def.},\lambda=1}{=} 1.$$

Wir zeigen nun, dass $\varphi \in X'$, d.h. dass auch noch zusätzlich φ stetig ist. Nach Lemma 3.4 (ii) gibt es nämlich ein $M \geq 0$, so dass

$$p_U(x) \leq M\|x\| \quad \text{für alle } x \in X$$

und daraus folgt

$$\pm\varphi(x) = \varphi(\pm x) \leq p_U(\pm x) \leq M\|x\| \quad \text{für alle } x \in X,$$

also

$$|\varphi(x)| \leq M\|x\| \quad \text{für alle } x \in X$$

und damit $\varphi \in X'$.

Wegen Lemma 3.4 (iii) gilt schließlich

$$\varphi(x) \leq p_U(x) < 1 \overset{(??)}{=} \varphi(x_0) \quad \text{für alle } x \in U.$$

2. Fall: $\mathbb{K} = \mathbb{C}$.
Wir müssen zeigen, dass es ein $\varphi \in X'$ gibt, so dass

$$\operatorname{Re}\varphi(x) < \operatorname{Re}\varphi(x_0) \quad \text{für alle } x \in U \tag{3.11}$$

Nach Fall 1. gibt es ein $\mathbb{R}$-wertiges lineares stetiges Funktional $\phi : X \to \mathbb{R}$ mit

$$\phi(x) < \phi(x_0) \quad \text{für alle } x \in U.$$

Setze

$$\varphi(x) := \phi(x) - i\phi(ix).$$

Dann ist φ als Verknüpfung stetiger Abbildung selbst wieder stetig.

Nach Lemma 3.1 (i) ist $\varphi : X \to \mathbb{C}$ also ein $\mathbb{C}$-lineares stetiges Funktional, also $\varphi \in X'$ und

$$\phi(x) = \operatorname{Re}\varphi(x).$$

Das zeigt (3.11).

<u>2. Schritt:</u> V beliebig.

Betrachte

$$M := U - V.$$

Dann ist M konvex, denn U und V sind konvex. Weiter ist $M \neq \emptyset$, weil $U, V \neq \emptyset$. Des Weiteren ist M offen, da

$$M = \bigcup_{v \in V} \underbrace{U - \{v\}}_{\text{offen}}$$

und beliebige Vereinigungen offener Mengen offen sind.

Wende nun den ersten Schritt an auf

$$M \quad \text{und} \quad \{0\}\,.$$

Es ist $M \cap \{0\} = \emptyset$ gesichert, da wegen $U \cap V = \emptyset$ in der Tat $0 \notin M$ gilt.

Es gibt also ein $\varphi \in X'$, so dass für alle $u \in U$ und $v \in V$

$$\operatorname{Re}(\varphi(u - v)) < \operatorname{Re}\varphi(0) = 0.$$

Daraus folgt

$$\operatorname{Re}\varphi(u) < \operatorname{Re}\varphi(v) \quad \text{für alle } u \in U \text{ und } v \in V.$$

□

Lemma 3.5. *Sei (X, d) ein metrischer Raum und seien $A, B \subset X$ mit $A, B \neq \emptyset$ und $A \cap B = \emptyset$. Ferner sei A kompakt und B abgeschlossen. Dann gilt*

$$d(A, B) = \inf\{\, d(a, b) \mid a \in A, b \in B \,\} > 0. \tag{3.12}$$

Beweis. Seien dazu $(b_n)_n \subset B$ und $(a_n)_n \subset A$ zwei Folgen mit der Annahme

$$\lim_{n\to\infty} d(b_n, a_n) = 0.$$

Dann gibt es wegen der Kompaktheit von A eine konvergente Teilfolge mit

$$\lim_{k\to\infty} a_{n_k} = a \in A.$$

Daraus folgt

$$\lim_{k\to\infty} d(b_{n_k}, a) \leq \lim_{k\to\infty} d(b_{n_k}, a_{n_k}) + \lim_{k\to\infty} d(a_{n_k}, a) = 0.$$

Da B abgeschlossen ist, folgt dann $a \in B$ sowie

$$a \in A \cap B.$$

Doch das ist ein Widerspruch zu $A \cap B = \emptyset$. Also gilt tasächlich (3.12). □

Satz 3.5 (2. Trennungssatz, Trennung im strikten Sinne). *Sei $(X, \|.\|_X)$ ein normierter Vektorraum und seien $U, V \neq \emptyset$ disjunkte, konvexe Teilmengen von X. Weiter seien U abgeschlossen und V kompakt. Dann gibt es ein lineares stetiges Funktional $\varphi \in X'$ und $\alpha_1, \alpha_2 \in \mathbb{R}$, so dass*

$$\operatorname{Re}\varphi(u) \leq \alpha_1 < \alpha_2 \leq \operatorname{Re}\varphi(v) \quad \textit{für alle } u \in U \textit{ und } v \in V.$$

Beweis. Zunächst ist nach Lemma 3.5

$$d := \operatorname{dist}(U, V) > 0.$$

gesichtert. Sei nun $0 < r < d$. Betrachte

$$U_r := \underbrace{U}_{\text{konvex}} + \underbrace{B_X(0, r)}_{\text{konvex}} = \bigcup_{u\in U} \underbrace{(u + B_X(0, r))}_{\text{offen}}.$$

Dann ist U_r offen und konvex. Weiter ist

$$U_r \cap V = \emptyset,$$

da für alle $u_r := u + x \in U_r$ und $v \in V$

$$\|(u+x) - v\| = \|u - v + x\| \geq \|u - v\| - \|x\| \geq d - \|x\| > d - r > 0.$$

Nach dem ersten Trennungssatz 3.4 gibt es ein $\varphi \in X'$ mit

$$\operatorname{Re}\varphi(u+x) < \operatorname{Re}\varphi(v) \quad \text{für alle } u \in U, v \in V \text{ und } x \in B_X(0,r). \tag{3.13}$$

Sei nun $\tilde{x} \in B_X$. Dann gilt

$$-\tilde{x} \in B_X \quad \text{und} \quad \pm \tfrac{r}{2}\tilde{x} \in B_X(0,r)$$

und daher

$$\begin{aligned}\operatorname{Re}\varphi(u \pm \tfrac{r}{2}\tilde{x}) &= \operatorname{Re}\varphi(u) \pm \tfrac{r}{2}\operatorname{Re}\varphi(\tilde{x}) < \operatorname{Re}\varphi(v)\\ &= \operatorname{Re}\varphi(u) + \tfrac{r}{2}|\operatorname{Re}\varphi(\tilde{x})| < \operatorname{Re}\varphi(v).\end{aligned}$$

Nach Supremumsübergang folgt

$$\operatorname{Re}\varphi(u) + \tfrac{r}{2}\sup_{x\in B_X}|\operatorname{Re}\varphi(x)| \leq \operatorname{Re}\varphi(v),$$

also

$$\operatorname{Re}\varphi(u) + \underbrace{\frac{r}{2}\|\varphi\|_{X'}}_{=:\epsilon} < \operatorname{Re}\varphi(v).$$

Dann ist nach der Bemerkung Punkt 2. $\epsilon > 0$ und daher

$$\operatorname{Re}\varphi(u) < \operatorname{Re}\varphi(u) + \tfrac{\epsilon}{2} < \operatorname{Re}\varphi(u) + \epsilon \leq \operatorname{Re}\varphi(v). \tag{3.14}$$

Setze

$$\alpha_1 := \sup_{u\in U}\operatorname{Re}\varphi(u), \quad \alpha_2 := \inf_{v\in V}\operatorname{Re}\varphi(v).$$

Dann folgt aus (3.14)

$$\operatorname{Re}\varphi(u) \leq \alpha_1 < \alpha_2 \leq \operatorname{Re}\varphi(v),$$

was zu zeigen war. □

Korollar 3.5. *Seien $(X, \|.\|_X)$ ein normierter $\mathbb{K}$-Vektorraum und $U \subset X$ ein Untervektorraum. Dann gilt*

$$U \text{ dicht in } X \iff (\forall_{\varphi\in X'} : \varphi|_U = 0 \Longrightarrow \varphi = 0).$$

Beweis. Zu „$\Longrightarrow$“: Sei also

$$\varphi(u) = 0 \quad \text{für alle } u \in U. \tag{3.15}$$

Weiter liegt U dicht in X, d.h. zu jedem $x \in X$ gibt es eine Folge $(u_n)_n \subset U$ mit

$$\lim_{n\to\infty}\|x - u_n\|_X = 0.$$

Betrachte nun

$$0 \leq |\varphi(x)| \leq |\varphi(x-u_n)| + \underbrace{|\varphi(u_n)|}_{=0 \text{ wegen } (3.15)} \leq \|\varphi\|_{X'}\|x-u_n\|_X \longrightarrow 0.$$

Also ist $\varphi = 0$.

Zu „$\Longleftarrow$“: Angenommen $\bar{U} \neq X$. Dann gibt es ein $x_0 \in X \setminus \bar{U}$ und nach dem 2. Trennungssatz 3.5 angewendet auf $\bar{U}$ und $V := \{x_0\}$ gibt es ein $\phi \in X'$, so dass

$$\operatorname{Re}\phi(u) < \operatorname{Re}\phi(x_0) \quad \text{für alle } u \in \bar{U}, \tag{3.16}$$

denn $\bar{U} \cap V = \emptyset$, U ist abgeschlossen, V ist kompakt und U sowie V sind konvex.

Da U ein Untervektorraum ist, ist auch $\bar{U}$ ein Untervektorraum, d.h.

$$\lambda u \in \bar{U} \quad \text{für alle } \lambda \in \mathbb{K} \text{ und } u \in \bar{U}.$$

Dann gilt für alle $\lambda \in \mathbb{R}$ mit $\lambda \neq 1$

$$\operatorname{Re}\phi(\lambda u) = \lambda \operatorname{Re}\phi(u) < \operatorname{Re}\phi(x_0)$$

und das geht nur, wenn ϕ auf ganz U konstant Null ist. Nach Voraussetzung folgt dann aber, dass ϕ das Nullfunktional ist, und dies steht im Widerspruch zu der strikten(!) Ungleichung in (3.16). $\square$

Bemerkung. Die Korollare 3.1 bis 3.4 aus dem letzen Abschnitt können auch aus den Trennungssätzen hergeleitet werden. Genauso kann das obere Korollar mit dem Hahn-Banach-Satz hergeleitet werden.

3.3 Reflexive Räume

Wir haben gesehen, dass wenn $(X, \|.\|_X)$ ein beliebiger normierter Raum ist, dann stets sein Dualraum $(X', \|.\|_{X'})$ vollständig ist (vgl. Satz 2.2).

Wir betrachten nun den so genannten *Bidualraum* $X'' := (X')'$ ausgestattet mit der Norm

$$\|x''\|_{X''} := \sup_{\varphi \in B_{X'}} |x''(\varphi)|, \quad x'' \in X''.$$

Dann ist auch wieder $(X'', \|.\|_{X''})$ vollständig.

Jedes $x \in X$ definiert über die so genannte *kanonische Abbildung*

$$j : X \to X'', \quad x \mapsto j(x)(\varphi) := \varphi(x), \quad \varphi \in X'$$

ein lineares stetiges Funktional.

In der Tat ist j wohldefiniert, d.h. $j(x) \in X''$: Seien dazu $\varphi, \psi \in X'$ und $\lambda, \mu \in \mathbb{K}$. Dann gilt

$$j(x)(\lambda\varphi + \mu\psi) = (\lambda\varphi + \mu\psi)(x) = \lambda\varphi(x) + \mu\psi(x) = \lambda j(x)(\varphi) + \mu j(x)(\psi)$$

und damit ist die Linearität von $j(x)$ gezeigt. Die Stetigkeit folgt aus

$$|j(x)(\varphi)| = |\varphi(x)| \leq \|\varphi\|_{X'} \|x\|_X.$$

Somit ist $j(x) \in X''$ gezeigt.

Eigenschaften von $j : X \to X''$.

1. <u>j ist linear:</u> Für $x, y \in X$ und $\lambda, \mu \in \mathbb{K}$ gilt nämlich mit $\varphi \in X'$

   $$j(\lambda x + \mu y)(\varphi) = \varphi(\lambda x + \mu y) = \lambda\varphi(x) + \mu\varphi(y) = \lambda j(x)(\varphi) + \mu j(y)(\varphi),$$

   also

   $$j(\lambda x + \mu y) = \lambda j(x) + \mu j(y).$$

2. <u>j ist isometrisch und daher insbesondere beschränkt und injektiv:</u> Denn betrachte

   $$\|j(x)\|_{X''} = \sup_{\varphi \in B_{X'}} |j(x)(\varphi)| = \sup_{\varphi \in B_{X'}} |\varphi(x)| \overset{\text{Normformel}}{=} \|x\|_X.$$

3. <u>j ist im Allgemeinen nicht surjektiv:</u> Betrachte zum Beispiel[6]

   $$X = c_0 \Longrightarrow X' = (c_0)' \cong l^1 \Longrightarrow X'' = (l^1)' \cong l^\infty.$$

   Dann gilt

   $$j(x) = x.$$

   Denn identifiziert man $\varphi = (\varphi_n)_n \in l^1$ mit dem Funktional

   $$\varphi : x = (x_n)_n \mapsto \sum_n x_n \varphi_n,$$

   so sieht man

   $$j(x)(\varphi) = \varphi(x) = \sum x_n \varphi_n = x''(\varphi),$$

   wo $x'' \in l^\infty$ das Funktional

   $$x'' : (\varphi_n)_n \mapsto \sum x_n \varphi_n$$

   darstellt. Also folgt

   $$x'' = x = j(x)$$

   und hieran sieht man, dass $j(x)$ nicht surjektiv ist, denn nicht jede beschränkte Folge (= Element aus l^∞) ist eine Nullfolge (= Element aus c_0).

[6] Vergleichen Sie dazu auch das Beispiel 2.10 und seine Bemerkung

Wir fassen nocheinmal zusammen:

Satz 3.6. *Die Abbildung*

$$j : X \to X'', \quad j(x)(\varphi) = \varphi(x)$$

ist eine (im Allgemeinen nicht surjektive) lineare Isometrie.

Insgesamt folgt also, dass X mit dem Unterraum $j(X) \subset X''$ identifiziert werden kann, d.h.

$$\boxed{X \cong j(X) \subset X''}.$$

Daher ist auch die symmetrische Schreibweise

$$\langle x, \varphi \rangle_{X,X'} = \langle j(x), \varphi \rangle_{X'',X'} = \langle \varphi, x \rangle_{X',X} = \varphi(x)$$

mit der so genannten *Dualitätsklammer* $\langle ., . \rangle$ gerechtfertigt.

Bemerkungen.

1. Ist X vollständig, so ist $j(X)$ abgeschlossen in X''

   Anders ausgedrückt: Ein Banachraum X wird mit einem abgeschlossenen Unterraum von X'' identifiziert.

   Beweis. Da X vollständig ist, ist wegen der Identifikation mit $j(X)$ auch $j(X)$ vollständig. Somit ist $j(X)$ abgeschlossen. □

2. Ist X nicht vollständig, so kann $\overline{j(X)}$ als Vervollständigung von X angesehen werden.

   Beweis. $\overline{j(X)}$ ist im Banachraum X'' abgeschlossen, also ist $\overline{j(X)}$ nach Satz 1.1 vollständig. □

Satz 3.7 (und Definition). *Ist* $(X, \|.\|_X)$ *ein normierter* $\mathbb{K}$*-Vektorraum, so gibt es einen Banachraum* $(\hat{X}, \|.\|_{\hat{X}})$*, in dem* X *dicht liegt.* $\hat{X}$ *heißt* Vervollständigung von X.

Beweis. Benutze dazu die Bemerkung 2. Dann erfüllt nämlich

$$\hat{X} := \overline{j(X)}$$

das Gewünschte, denn X liegt dicht in $\hat{X}$, weil

$$X \cong j(X) \quad \text{sowie} \quad \overline{j(X)} = \hat{X},$$

und $\hat{X}$ ist vollständig, d.h. modulo Identifikation mit $j(X)$ kann $\hat{X}$ als Vervollständigung von X angesehen werden. □

Definition 3.3 (Reflexiv). Ein normierter $\mathbb{K}$-Vektorraum $(X, \|.\|_X)$ heißt *reflexiv*, wenn die kanonische Abbildung surjektiv ist, d.h. wenn

$$j(X) = X''.$$

Beispiele 3.5 (**für reflexive Räume**).

- Hilberträume sind reflexiv (siehe Korollar 5.2).
- Die Sobolevräume $W^{k,p}(\Omega)$ für $1 < p < \infty$ sind reflexiv. Das werden wir aber erst im nächsten Semester zeigen, wenn wir uns mit solchen Räumen beschäftigen.
- Jeder endlich-dimensionale Vektorraum $(X, \|.\|_X)$ ist reflexiv. Das folgt nämlich aus Linearer Algebra (siehe Fischer: Satz 6.1.6).
- $L^p(\Omega)$ und l^p sind für $1 < p < \infty$ reflexiv.

  Beweis. Für $1 < p < \infty$ stimmt die kanonische Abbildung $j : l^p \to (l^p)''$ mit dem identischen Operator
  $$\mathrm{I} : l^p \to l^p$$
  überein und ist deswegen surjektiv. Das zeigt man wie oben im „nicht-surjektiv-Beispiel“, denn
  $$(l^p)'' \cong (l^q)' \cong l^p, \quad \tfrac{1}{p} + \tfrac{1}{q} = 1, \quad 1 < p < \infty$$
  nach der Bemerkung zu Beispiel 2.10.

  Die gleichen Überlegungen gelten für $L^p(\Omega)$

  □

Beispiele 3.6 (**für nicht reflexive Räume**).

- c_0 ist nicht reflexiv, wie wir oben gesehen hatten.
- l^1 ist nicht reflexiv: $(l^1)'' = ((l^1)')' \cong (l^\infty)' \ncong l^1$.
- l^∞ ist nicht reflexiv: $(l^\infty)'' = ((l^\infty)')' \ncong (l^1)' \cong l^\infty$.
- $L^1(\Omega)$ und $L^\infty(\Omega)$ sind nicht reflexiv.
- $C(K)$ für kompaktes K ist nicht reflexiv.

□

Bemerkungen.

1. Ist X reflexiv, so kann X mit X'' identifiziert werden, d.h.
   $$X \cong X''.$$
   Die Umkehrung gilt aber im Allgemeinen nicht, d.h.
   $$X \cong X'' \not\Longrightarrow j \text{ surjektiv.}$$
   Es ist demnach möglich, dass es eine lineare bijektive isometrische Abbildung $i : X \to X''$ gibt, somit also X und X'' identifiziert werden können, aber die kanonische Abbildung $j : X \to X''$ nicht surjektiv ist! Ein solches Beispiel findet man etwa in: *R.C. James, A non reflexive Banach space isometric with second conjugate space, Proc. Nat. Acad. Sc. USA 37 (1951), 174-177.*
2. Da X'' als Dualraum vollständig ist, ist ein reflexiver Raum wegen $X \cong X''$ auch vollständig.

   Anders gesagt: Ein unvollständiger Raum hat keine Chance, reflexiv zu sein.

Satz 3.8. *Ein abgeschlossener Teilraum eines reflexiven Banachraumes ist reflexiv.*

Beweis. Sei also $(X, \|.\|_X)$ ein reflexiver Banachraum und sei $F \subset X$ ein abgeschlossener Teilraum.

Betrachte $\varphi \in F''$ und definiere

$$\phi(x') = \varphi(x'|_F), \quad x' \in X.$$

Dann ist $\phi \in X''$, denn

$$|\phi(x')| = |\varphi(x'|_F)| \leq \|\varphi\|_{F''} \|x'|_F\|_{F'} \leq \|\varphi\|_{F''} \|x'\|_{X'}$$

und die Linearität ist sowieso klar.

Weil X reflexiv ist, gibt es ein $x_0 \in X$, so dass

$$\phi(x') = x'(x_0) \quad \text{für alle } x' \in X',$$

denn ϕ ist ja surjektiv. Es gilt also

$$\varphi(x'|_F) = x'(x_0). \tag{3.17}$$

Wir behaupten: $x_0 \in F$.

Angenommen, $x_0 \notin F$. Dann existiert, da F abgeschlossen ist, nach Korollar 3.4 ein $x' \in X'$ mit

$$x'|_F = 0 \quad \text{und} \quad x'(x_0) \neq 0,$$

d.h. $\varphi(x'|_F) = 0$ und das ist ein Widerspruch zu (3.17).

Weiter im eigentlichen Beweis: Sei nun $y' \in F'$. Dann gibt es nach dem Satz von Hahn-Banach (Satz 3.3) eine Fortsetzung $x' \in X'$ mit

$$x'|_F = y' \quad \text{und} \quad \|x'\|_{X'} = \|y'\|_{F'}.$$

Daraus folgt

$$\varphi(y') = \varphi(x'|_F) \overset{(3.17)}{=} x'(x_0) = y'(x_0,)$$

also

$$\varphi = j(x_0)$$

und damit ist $j : F \to F''$ surjektiv und F reflexiv. □

Satz 3.9. *Ein Banachraum X ist genau dann reflexiv, wenn X' reflexiv ist.*

Beweis. <u>Zu „$\Longrightarrow$“:</u> Wir müssen zeigen, dass

$$j_{X'} : X' \to (X')'', \quad j_{X'}(x')(x'') = x''(x')$$

surjektiv ist. Sei dazu also $\phi \in X''' \cong (X'')'$.

Zu zeigen: $\exists x_0' \in X' : \phi(x'') = x''(x_0')$ für alle $x'' \in X''$.

Wir setzen

$$x_0' := \phi \circ j_X \tag{3.18}$$

mit der kanonischen Abbildung $j_X : X \to X''$, also ist

$$x_0'(x) = \phi(j_X(x)).$$

Es ist $x_0' \in X'$, denn x_0' ist eine Verkettung linearer stetiger Funktionen.

Sei nun $x'' \in X''$. Da X reflexiv ist, gibt es ein $x_0 \in X$, so dass

$$x'' = j_X(x_0) \Longrightarrow x''(x') = x'(x_0) \tag{3.19}$$

Daraus folgt

$$\phi(x'') = \phi(j_X(x_0)) = (\phi \circ j_X)(x_0) \overset{(3.18)}{=} x_0'(x_0) \overset{(3.19)}{=} x''(x_0'),$$

was zu zeigen war.

Zu „$\Longleftarrow$“: Sei nun X' reflexiv. Wegen dem eben Bewiesenden ist dann auch X'' reflexiv. Wegen Punkt 1. in der oberen Bemerkung gilt

$$X \text{ vollständig } \Longrightarrow j(X) \text{ abgeschlossen in } X''.$$

Also ist nach Satz 3.8 der Teilraum $j(X) \subset X''$ reflexiv und wegen $X \cong j(X)$ auch X selbst. □

Zum Schluss notieren wir noch:

Satz 3.10. *Sei M eine Teilmenge eines normierten Raumes X. Dann ist M genau dann beschränkt, wenn für alle $\varphi \in X'$ die Menge $\varphi(M) \subset \mathbb{K}$ beschränkt ist.*

Beweis. Zu „$\Longrightarrow$“: Trivial.

Zu „$\Longleftarrow$“: Betrachte die Funktionale $i_X(x)$ für $x \in M$, die auf dem Banachraum X' definiert sind. Dann gilt nach Voraussetzung

$$\sup_{x \in M} |\varphi(x)| = \sup_{x \in M} |i_X(x)(\varphi)| < \infty \quad \text{für alle } \varphi \in X'.$$

Aus dem Satz von Banach-Steinhaus (Satz 2.5) folgt nun

$$\sup_{x \in M} \|x\| = \sup_{x \in M} \|i_X(x)\| < \infty.$$

□

Für die duale Version des Satzes muss die Vollständigkeit gefordert werden:

Satz 3.11. *Sei X ein Banachraum und $M' \subset X'$. Dann ist M' genau dann beschränkt, wenn für alle $x \in X$ die Menge $\{\, \varphi(x) \mid \varphi \in M' \,\}$ beschränkt ist.*

Beweis. Zu „$\Longrightarrow$“: Trivial.

Zu „$\Longleftarrow$“: Nach Voraussetzung gilt also

$$\sup_{\varphi \in M'} \|\varphi(x)\| < \infty \quad \text{für alle } x \in X.$$

Aus dem Satz von Banach-Steinhaus (Satz 2.5) folgt auch schon

$$\sup_{\varphi \in M'} \|\varphi\| < \infty,$$

was zu zeigen war. □

3.4 Schwache Konvergenz

Definition 3.4 (Starke und schwache Konvergenz). Sei $(X, \|.\|_X)$ ein Banachraum.

(i) Die Folge $(x_n)_n \subset X$ konvergiert *stark gegen* $x \in X$, wenn

$$\lim_{n\to\infty} \|x_n - x\|_X = 0.$$

Notation. $x_n \to x$ in X.

(ii) Die Folge $(x_n)_n \subset X$ konvergiert *schwach gegen* $x \in X$, wenn

$$\forall_{\varphi\in X'} : \lim_{n\to\infty} \langle \varphi, x_n - x\rangle = 0.$$

Notation. $x_n \rightharpoonup x$ in X oder $\sigma(x, x') - \lim x_n = x$.

Bemerkungen und Eigenschaften.

1. Der schwache Grenzwert ist eindeutig bestimmt.

   Sei nämlich $\hat{x} \neq x$ ein weiterer schwacher Grenzwert der Folge $(x_n)_n \subset X$. Dann gilt für alle $\varphi \in X'$
   $$\langle \varphi, x\rangle = \lim \langle \varphi, x_n\rangle = \langle \varphi, \hat{x}\rangle.$$
   Daraus folgt $x = \hat{x}$ im Widerspruch zu $x \neq \hat{x}$.

2. Aus $x_n \rightharpoonup x$ folgt für jede Teilfolge $x_{n'} \subset x_n$ ebenfalls $x_{n'} \rightharpoonup x$.

   Denn liegen in einer vorgegebenen Umgebung von X fast alle Glieder, so gilt das erst recht für die Teilfolgenglieder.

3. Ist $x_n \to x$, so folgt $x_n \rightharpoonup x$.

   Für alle $\varphi \in X'$ gilt nämlich
   $$|\langle \varphi, x_n - x\rangle| \leq \|\varphi\|_{X'} \underbrace{\|x_n - x\|_X}_{\to 0} \to 0.$$
   Die Umkehrung gilt aber im Allgemeinen nicht. Vergleiche das folgende Beispiel 3.7 unten.

4. Gilt $x_n \rightharpoonup x$, so folgt
   $$(\|x_n\|_X)_n \text{ beschränkt} \quad \text{und} \quad \|x\|_X \leq \liminf_{n\to\infty} \|x_n\|_X.$$
   Die letzte Ungleichung drückt auch die so genannte *schwache Unterhalbstetigkeit* aus.

   Beweis. Betrachte die kanonische Abbildung
   $$j : X \to X''.$$
   Dann ist $(j(x_n))_n$ punktweise beschränkt, denn
   $$\langle j(x_n), \varphi\rangle = j_{x_n}(\varphi) = \langle \varphi, x_n\rangle$$
   ist konvergent und damit beschränkt, weil $x_n \rightharpoonup x$.

   Nach dem Satz von Banach-Steinhaus gilt dann für alle $n \in \mathbb{N}$:
   $$\exists_{M\geq 0} : \|j(x_n)\|_{X''} \overset{\text{Isometrie}}{=} \|x_n\|_X \leq M.$$

Nun ergibt sich aus

$$\langle \varphi, x_n \rangle \leq \|\varphi\|_{X'} \|x_n\|_X \quad \text{für alle } \varphi \in X' \text{ und } n \in \mathbb{N}$$

nach Grenzübergang

$$\langle \varphi, x \rangle \leq \|\varphi\|_{X'} \liminf_{n\to\infty} \|x_n\|_X \quad \text{für alle } \varphi \in X'$$

Wegen der Normformel heißt das

$$\|x\|_X \leq \liminf_{n\to\infty} \|x_n\|.$$

□

5. Es gilt genau dann $x_n \rightharpoonup x$, wenn
   (i) $\|x_n\|_X$ beschränkt,
   (ii) $\lim \langle \varphi, x_n - x \rangle = 0$ für alle $\varphi \in M \subset X'$ mit $\overline{\operatorname{lin}(M)} = X'$.

   Dabei bezeichnet $\operatorname{lin}(M)$ die lineare Hülle von M, d.h. die Menge aller endlichen Linearkombinationen von Elementen aus M.

Beweis. Zu „$\Longrightarrow$“: Trivial mit 4.

Zu „$\Longleftarrow$“: Sei also $\psi \in X'$ beliebig. Sei weiter $(x_n)_n \subset X$ eine Folge nach Vorausetzung. Dann gibt es ein $K \geq 0$, so dass

$$\|x\|_X \leq \tfrac{K}{2} \quad \text{und} \quad \|x_n\|_X \leq \tfrac{K}{2}.$$

Sei $\phi_k \in \operatorname{lin}(M)$. Dann gibt es $\varphi_i^{(k)} \in M$ und $\alpha_i^{(k)} \in \mathbb{K}$ mit

$$\phi_k = \sum_{i=1}^{m_k} \alpha_i^{(k)} \varphi_i^{(k)}.$$

Setze

$$C := (\max_n m_n + 1) \cdot \max_{k,i} \left|\alpha_i^{(k)}\right|.$$

Sei $\epsilon > 0$. Wegen der Dichtheit von $\operatorname{lin}(M)$ in X' gibt es ein $K_0 \in \mathbb{N}$, so dass

$$\|\phi_k - \psi\|_{X'} < \tfrac{\epsilon}{2(K+1)} \quad \text{für alle } k \geq K_0.$$

Da $\varphi_i^{(k)} \in M$, gibt es nach (ii) ein $N_0 \in \mathbb{N}$, so dass für alle i und k

$$\left|\left\langle \varphi_i^{(k)}, x_n - x \right\rangle\right| < \tfrac{\epsilon}{2C} \quad \text{für alle } n \geq N_0.$$

Setze

$$N := \max(N_0, K_0).$$

Dann gilt für alle $n \geq N$

$$\begin{aligned}
|\langle \psi, x_n - x \rangle| &= |\langle \psi, x_n \rangle - \langle \psi, x \rangle| \\
&\leq |\langle \psi, x_n \rangle - \langle \phi_N, x_n \rangle| + |\langle \phi_N, x_n \rangle - \langle \phi_N, x \rangle| + |\langle \phi_N, x \rangle - \langle \psi, x \rangle| \\
&= |\langle \psi - \phi_N, x_n \rangle| + |\langle \phi_N, x_n - x \rangle| + |\langle \phi_N - \psi, x \rangle| \\
&\leq \|\psi - \phi_N\|_{X'} \|x_n\|_X + |\langle \phi_N, x_n - x \rangle| + \|\phi_N - \psi\|_{X'} \|x\|_X \\
&\leq K \|\psi - \phi_N\|_{X'} + \left|\left\langle \sum_{i=1}^{m_N} \alpha_i^{(N)} \varphi_i^{(N)}, x_n - x \right\rangle\right| \\
&< \tfrac{\epsilon}{2} + \left|\sum_{i=1}^{m_N} \alpha_i^{(N)} \left\langle \varphi_i^N, x_n - x \right\rangle\right| \leq \tfrac{\epsilon}{2} + \sum_{i=1}^{m_N} \left|\alpha_i^{(N)}\right| \left|\left\langle \varphi_i^N, x_n - x \right\rangle\right| \\
&< \tfrac{\epsilon}{2} + \sum_{i=1}^{m_N} |\alpha_i^{(N)}| \tfrac{\epsilon}{2C} < \epsilon.
\end{aligned}$$

□

6. Im Falle $\dim X < \infty$ gilt
$$x_n \rightharpoonup x \Longleftrightarrow x_n \to x.$$
Dieselbe Äquivalenz gilt auch für $X = l^1$.

Beweis. <u>Zu „$\Longleftarrow$“:</u> Trivial nach 3.

<u>Zu „$\Longrightarrow$“:</u> Sei also $\dim X = k < \infty$. Dann gibt es nach Linearer Algebra eine Basis $\{e_i\}_{i=1,\dots,k}$, so dass
$$x = \sum_{i=1}^{k} \alpha_i e_i, \quad x_n = \sum_{i=1}^{k} \beta_i^{(n)} e_i, \quad \alpha_i, \beta_i^{(n)} \in \mathbb{K}.$$
Des Weiteren gibt es nach linearer Algebra genau ein $\varphi_j \in X'$ mit
$$\varphi_j(e_i) = \delta_{ij}.$$
Dann gilt
$$\begin{aligned} |\langle \varphi_j, x_n - x\rangle| &= \left|\left\langle \varphi_j, \sum_{i=1}^{k} (\beta_i^{(n)} - \alpha_i) e_i \right\rangle\right| = \left|\sum_{i=1}^{k} (\beta_i^{(n)} - \alpha_i)\right| \underbrace{|\langle \varphi_j, e_i\rangle|}_{=\delta_{ji}} \\ &= \left|\beta_j^n - \alpha_j\right|. \end{aligned}$$
Nach Voraussetzung gilt $x_n \rightharpoonup x$, also
$$\lim_{n\to\infty} \left|\beta_j^{(n)} - \alpha_j\right| = 0 \quad \text{für alle } j = 1, \dots, k. \tag{3.20}$$
Es folgt
$$\begin{aligned} \|x - x_n\|_X &= \left\|\sum_{i=1}^{k} (\beta_i^{(n)} - \alpha_i) e_i\right\|_X = \left|\sum_{i=1}^{k} (\beta_i^{(n)} - \alpha_i)\right| \|e_i\|_X \\ &\le \sum_{i=1}^{k} \left|\beta_i^{(n)} - \alpha_i\right| \underbrace{\|e_i\|_X}_{\neq 0,\text{ da Basis}} \overset{n\to\infty}{\longrightarrow} 0 \text{ nach (3.20).} \end{aligned}$$

<u>$X = l^1$:</u> Siehe Anhang 7.2.

□

7. Gilt $x_n \rightharpoonup x$ in X und $\varphi_n \to \varphi$ in X', so folgt
$$\langle \varphi_n, x_n\rangle \to \langle \varphi, x\rangle.$$

Beweis. Nach 4. ist x_n beschränkt, also gibt es ein $M \ge 0$, so dass
$$\|x_n\|_X \le M.$$
Es folgt
$$\begin{aligned} |\langle \varphi_n, x_n\rangle - \langle \varphi, x\rangle| &= |\langle \varphi_n - \varphi, x_n\rangle + \langle \varphi, x_n\rangle - \langle \varphi, x\rangle| \\ &\le |\langle \varphi_n - \varphi, x_n\rangle| + |\langle \varphi, x_n - x\rangle| \\ &\le \|\varphi_n - \varphi\|_{X'} \|x_n\|_X + |\langle \varphi, x_n - x\rangle| \\ &\le M \underbrace{\|\varphi_n - \varphi\|_{X'}}_{\to 0} + \underbrace{|\langle \varphi, x_n - x\rangle|}_{\to 0} \to 0. \end{aligned}$$

□

Beispiel 3.7. Betrachte

$$X = l^2, \quad x = (x_1, x_2, \ldots), \quad \|x\| = \left(\sum_{n=1}^{\infty} |x_n|^2\right)^{1/2}.$$

Weiter betrachte die Folge

$$(e^{(k)})_{k\in\mathbb{N}}, \quad e^{(k)} = (0, 0, \ldots, 0, 1, 0, \ldots).$$

Es folgt

$$\left\|e^{(k)}\right\| = 1 \quad \text{und} \quad \left\|e^{(k)} - e^{(m)}\right\| = \begin{cases} 0, & \text{falls } k = m, \\ \sqrt{2} & \text{sonst.} \end{cases}$$

Also ist $(e^{(k)})_{k\in\mathbb{N}}$ keine Cauchy-Folge und somit erst recht <u>nicht</u> stark konvergent.

Nun ist $X' \cong l^2$ nach der Bemerkung zu Beispiel 2.10. Sei $\phi \in X'$. Dann gibt es nach dem Darstellungssatz von Riesz-Fréchet (Satz 5.4) genau ein $\varphi \in l^2$, so dass

$$\phi(x) = \langle \phi, x\rangle = \sum_{n=1}^{\infty} \varphi_n x_n.$$

Es folgt für alle $\phi \in X'$

$$\phi(e^{(k)}) = \sum_{n=1}^{\infty} \varphi_n e_n^{(k)} = \varphi_k \overset{k\to\infty}{\longrightarrow} 0,$$

denn, da $\varphi \in l^2$, muss φ_k eine Nullfolge sein (notwendige Bedingung für die Konvergenz von Reihen).

Wir haben also gezeigt, dass $e^{(k)} \rightharpoonup (0, 0, \ldots)$ und damit, dass die schwache Konvergenz nicht die starke Konvergenz impliziert.

□

Hier noch ein Beispiel, welches zeigt, dass die schwache Konvergenz nicht die starke impliziert:

Beispiel 3.8. Sei $X = L^2(\mathbb{R})$ ausgestattet mit der Norm

$$\|v\|_{L^2} := \left(\int_{\mathbb{R}} |v(x)|^2\, dx\right)^{1/2}.$$

Weiter bezeichne mit $(.,.)$ Skalarprodukt in $L^2(\mathbb{R})$, welches die $\|.\|_{L^2}$ erzeugt. Betrachte die Funktionenfolge $(u_n)_n$ definiert durch

$$u_n := \frac{1}{\sqrt{n}} \chi_{]n,2n[}(x), \quad x \in \mathbb{R},\ n \in \mathbb{N}.$$

(Hierbei ist für eine Menge $A \subset \mathbb{R}$ die Funktion χ_A die so genannte *Indikatorfunktion von* A definiert durch $\chi_A(x) = 1$, wenn $x \in A$ und $\chi_A(x) = 0$ für $x \notin A$).
Dann gilt

* $\|u_n\|_{L^2} = \left(\int_n^{2n} \frac{1}{n}\right)^{1/2} = 1,$
* $(u_n)_n$ ist keine Cauchy-Folge in $L^2(\mathbb{R})$, denn für alle geraden $m \in \mathbb{N}$ und $n = \frac{3}{2}m$ gilt

$$\begin{aligned}\|u_m - u_n\|_{L^2}^2 &= \left\|\frac{1}{\sqrt{m}}\chi_{]m,2m[} - \frac{1}{\sqrt{n}}\chi_{]n,2n[}\right\|_{L^2}^2 = \left\|\frac{1}{\sqrt{m}}\chi_{]m,2m[} - \sqrt{\frac{2}{3m}}\chi_{]\frac{3}{2}m,3m[}\right\|_{L^2}^2 \\ &= \int_m^{2m} \frac{1}{m} dx - 2\int_{\frac{3}{2}m}^{2m} \frac{1}{\sqrt{m}}\sqrt{\frac{2}{3m}} dx + \int_{\frac{3}{2}m}^{3m} \frac{2}{3m} dx = 1 - \sqrt{\frac{2}{3}} + 1.\end{aligned}$$

Allerdings gilt

* $u_n \rightharpoonup u \equiv 0$ in $L^2(\mathbb{R})$, denn für alle $f \in L^2(\mathbb{R}) \cong (L^2(\mathbb{R}))'$ gilt

$$\begin{aligned}
\langle f, u_n \rangle_{(L^2(\mathbb{R}))', L^2(\mathbb{R})} &= (f, u_n)_{L^2(\mathbb{R})} = \int_{\mathbb{R}} f(x) u_n(x) dx = \int_n^{2n} \frac{1}{\sqrt{n}} f(x) dx \\
&\overset{\text{Hölder}}{\leq} \left(\int_n^{2n} \frac{1}{n} dx \right)^{1/2} \left(\int_n^{2n} |f(x)|^2 \, dx \right)^{1/2} \\
&= 1 \cdot \left(\int_n^{2n} |f(x)|^2 \, dx \right)^{1/2} \overset{f \in L^2(\mathbb{R})}{\leq} \left(\int_n^{+\infty} |f(x)|^2 \, dx \right)^{1/2} \\
&\to 0 \quad \text{für } n \to \infty.
\end{aligned}$$

□

Satz 3.12. *Seien $(X, \|.\|_X)$ ein Banachraum und $M \subset X$ konvex und abgeschlossen. Weiter sei $(x_n)_n \subset M$ eine Folge mit $x_n \rightharpoonup x$. Dann folgt $x \in M$.*

Beweis. Angenommen, $x \notin M$. Nach dem 2. Trennungssatz von Hahn-Banach (Satz 3.2) angewandt auf die abgeschlossene konvexe Menge M und die kompakte konvexe Menge $\{x\}$ gibt es ein $\varphi \in X'$ und $\alpha_1, \alpha_2 \in \mathbb{R}$ mit

$$\operatorname{Re} \varphi(y) \leq \alpha_1 < \alpha_2 \leq \operatorname{Re} \varphi(x) \quad \text{für alle } y \in M.$$

Insbesondere gilt also

$$\underbrace{\operatorname{Re} \varphi(x_n)}_{\to \operatorname{Re} \varphi(x)} \leq \alpha_1 < \alpha_2 \leq \operatorname{Re} \varphi(x) \quad \text{für alle } n \in \mathbb{N}.$$

Widerspruch. □

Bemerkungen.

1. Die schwache Abgeschlossenheit impliziert stets die Normabgeschlossenheit.

   Ist nämlich $(x_n)_n$ eine Folge in $M \subset X$ mit $x_n \to x$, so folgt erst recht $x_n \rightharpoonup x$ und wegen der schwachen Abgeschlossenheit folgt dann $x \in M$.

2. Für nicht konvexe Mengen gilt im Allgemeinen nicht:

   $$\text{normabgeschlossen} = \text{schwach abgeschlossen.}$$

   Betrachte dazu das folgende

Beispiel 3.9. Sei $X = c_0$ ausgestattet mit der $\|.\|_\infty$-Norm. Nach der Bemerkung zu Beispiel 2.10 ist $X' \cong l^1$. Betrachte die Menge

$$S_X := \{ x = (x_n)_n \in c_0 \mid \|x\|_\infty = \sup_n |x_n| = 1 \}.$$

Dann ist S_X normabgeschlossen, denn wir hatten in der Tat in Beispiel 2.6 schon einmal gesehen, dass für $y^{(n)} \in c_0$ gilt

$$||y^{(n)} - y||_\infty \to 0 \Longrightarrow y \in c_0.$$

Ist also $y^{(n)} \subset S_X$, so folgt weiter

$$|\underbrace{||y^{(n)}||_\infty}_{=1} - ||y||_\infty| \leq ||y^{(n)} - y||_\infty \to 0,$$

also $\|y\|_\infty = 1$ und daher $y \in S_X$, was die Abgeschlossenheit zeigt.

Betrachte nun die Folge $(e^{(n)})_n$, die an der n-ten Stelle eine 1 und sonst nur 0 als Einträge hat. Offensichtlich ist $e^{(n)} \in S_X$ für alle $n \in \mathbb{N}$ und

$$e^{(n)} \rightharpoonup 0 \quad \text{in } c_0,$$

denn für alle $x = (x_n) \in l^1$ gilt

$$\left\langle x, e^{(n)} \right\rangle = \sum_{k=1}^{\infty} x_k e_k^{(n)} = x_n,$$

d.h. speziell $\sum |x_n| < \infty$, also

$$x_n \to x.$$

Nun ist aber $0 \notin S_X$, d.h. S_X ist nicht schwach abgeschlossen.

Aber es gilt eben auch, dass S_X nicht konvex ist, denn betrachte

$$x = (1, 1/2, 1/3, 1/4, \ldots) \in S_X \quad \text{und} \quad y = (1/2, 1, 1/3, 1/4, \ldots) \in S_X,$$

dann folgt

$$\frac{1}{2}x + \frac{1}{2}y = (3/4, 3/4, 1/3, 1/4, \ldots) \notin S_X.$$

□

Eine nützliche Konsequenz aus Satz 3.12 ist das

Korollar 3.6 (Satz von Mazur). *Sei $(X, \|.\|_X)$ ein normierter Vektorraum und $(x_n)_n \subset X$ eine Folge mit $x_n \rightharpoonup x$ in X. Dann gibt es eine Folge von Konvexkombinationen der Elemente $(x_n)_n$, d.h. genauer eine Folge $(y_n)_n$ von der Form*

$$y_n = \sum_{k=1}^{N_m} \alpha_k^{(n)} x_k, \quad \sum_{k=1}^{N_m} \alpha_k^{(n)} = 1 \quad \forall_n,$$

die stakt in X gegen x konvergiert, d.h.

$$y_n \to x \quad \text{in } \in X.$$

Beweis. Betrachte den Abschluss der Menge aller endlichen Konvexkombinationen von Elementen x_n, $n \in \mathbb{N}$ bezüglich $\|.\|_X$[7]:

$$C := \overline{\text{co}}(\{\, x_n \mid n \in \mathbb{N} \,\}).$$

Die Menge ist trivialerweise normabgeschlossen und konvex, also nach Satz 3.12 schwach abgeschlossen, d.h.

$$(x_n)_n \subset C, \quad x_n \rightharpoonup x \Longrightarrow x \in C.$$

Nach Konstruktion von C existiert dann aber eine Folge von Konvexkombinationen von Elementen aus $\{\, x_n \mid n \in \mathbb{N} \,\}$, die stark gegen $x \in X$ konvergieren. □

Aus dem Dualraum X' eines normierten Vektorraumes $(X, \|.\|_X)$ kann analog eine schwache Konvergenz definiert werden durch

$$\lim x''(\varphi_n) = \lim \langle x'', \varphi_n \rangle = \langle x'', \varphi \rangle \quad \text{für alle } x'' \in X''.$$

[7] C ist die *abgeschlossene konvexe Hülle* der Menge $\{\, x_n \mid n \in \mathbb{N} \,\}$, d.h. der Abschluss des Schnittes aller konvexen Obermenge oder anders ausgedrückt:

$$\text{co}(M) = \{\, \sum_{k=1}^{N_m} \alpha_k^{(n)} x_k \mid x_k \in M, \sum_{i=1}^{N_m} \alpha_i^{(n)} = 1, \alpha_i^{(n)} \geq 0 \,\}$$

Auf X' kann aber auch noch ein weiterer Konvergenzbegriff definiert werden, die so genannte *schwach*-Konvergenz*, die von dem Primalraum X erzeugt wird:

Definition 3.5 (Schwach*-Konvergenz). Eine Folge $(\varphi_n)_n \subset X'$ *konvergiert schwach* gegen* $\varphi \in X'$, wenn

$$\lim_{n\to\infty} \langle \varphi_n, x\rangle = \langle \varphi, x\rangle \quad \text{für alle } x \in X.$$

Das ist sozusagen die punktweise Konvergenz.

Notation.

$$\varphi_n \rightharpoonup_* \varphi.$$

Bemerkungen.

1. Die schwach*-Konvergenz ist eindeutig, denn

$$\langle \varphi, x\rangle = \lim \langle \varphi_n, x\rangle = \langle \tilde{\varphi}, x\rangle \Longrightarrow \varphi = \tilde{\varphi}.$$

2. Wenn X vollständig ist, dann sind schwach*-konvergente Folgen auch normbeschränkt in X', denn gilt

$$(\varphi_n)_n \subset X', \quad \varphi_n \rightharpoonup_* \varphi,$$

dann ist die Familie $(\varphi_n)_n \subset X'$ insbesondere punktweise beschränkt und nach dem Satz von Banach-Steinhaus ist dann $(\varphi_n)_n$ normbeschränkt in X'.

3. Falls X nicht vollständig ist, dann gilt 2. im Allgemeinen nicht (Vergleiche das folgende Beispiel 3.10 unten)

4. Aus der schwachen Konvergenz in X' folgt die schwach*-Konvergenz, denn ist $(\varphi_n) \subset X'$, so gilt

$$\langle x'', \varphi_n\rangle \to \langle x'', \varphi\rangle \quad \text{für alle } x'' \in X''$$

und daraus folgt insbesondere mit der kanonischen Abbildung j

$$\langle j(x), \varphi_n\rangle_{X'',X'} \to \langle j(x), \varphi\rangle_{X'',X'} \quad \text{für alle } x \in X.$$

D.h. aber gerade

$$\langle \varphi_n, x\rangle_{X',X} \to \langle \varphi, x\rangle_{X',X} \quad \text{für alle } x \in X,$$

also

$$\varphi_n \rightharpoonup_* \varphi.$$

5. Falls X reflexiv ist, d.h. $j(X) = X''$ gilt, dann fallen die beiden Konvergenzbegriffe der schwachen und schwach*-Konvergenz offensichtlich nach 4. zusammen.

6. Für beliebige normierte Vektorräume gilt 5. im Allgemeinen nicht mehr (Vergleiche das folgende Beispiel 3.11 unten).

Beispiel 3.10. Betrachte $X = d$ ausgestattet mit der $\|.\|_\infty$-Norm. Dann ist $(d, \|.\|_\infty)$ nicht vollständig (vgl. *Werner, p. 8*). Betrachte weiterhin

$$\varphi_n : d \to \mathbb{K}, \quad (s_m)_{m\in\mathbb{N}} \mapsto n \cdot s_n.$$

Dann ist $\varphi_n \in d'$: Die Linearität ist klar. Die Stetigkeit folgt aus

$$\|\varphi_n\|_{X'} = \sup_{x\in B_d} |\varphi_n(x)| = \sup_{x_n\in B_d} |n \cdot x_n| \leq n.$$

Sei nun $e^{(n)}$ das Element mit einer 1 an der Stelle n und sonst nur Nullen an den anderen Stellen. Dann ist $e^{(n)} \in B_d$ und

$$\|\varphi_n\|_{X'} \geq |n \cdot e^n| = n.$$

Wir haben also sogar gezeigt, dass

$$\|\varphi_n\|_{X'} = n.$$

Hieran sieht man auch schon, dass φ_n nicht normbeschränkt ist. Aber es gilt

$$\varphi_n \rightharpoonup_* 0,$$

denn für alle $(s_m)_m \in d$ gilt

$$\langle \varphi_n, s_m \rangle = n \cdot s_n \to 0,$$

denn nur endlich viele s_m sind $\neq 0$.

□

Beispiel 3.11. Sei $X = c_0$. Dann gilt nach der Bemerkung zu Beispiel 2.10

$$X' \cong l^1 \text{ und } X'' \cong l^\infty$$

und wir wissen nach Beispiel 3.6, dass c_0 nicht reflexiv ist.

Betrachte nun die bekannte Folge $(e^{(n)})_{n\in\mathbb{N}} \in l^1$. Dann gilt

$$e^{(n)} \rightharpoonup_* 0 \quad \text{in } l^1,$$

denn für alle $x = (x_n)_n \in c_0$ gilt

$$\left\langle e^{(n)}, x \right\rangle = \sum_{k=1}^{\infty} e_k^{(n)} x_k = x_n \to 0.$$

Aber für die konstante 1-Folge

$$x'' = (1, 1, 1, \ldots) \in l^\infty \cong X''$$

gilt für alle $n \in \mathbb{N}$

$$\left\langle x'', e^{(n)} \right\rangle = \sum_{k=1}^{\infty} x''_k e_k^{(n)} = 1,$$

d.h.

$$\left\langle x'', e^{(n)} \right\rangle \not\to 0$$

und somit ist $(e^{(n)})_n$ nicht(!) schwach konvergent.

□

Das Auswahlprinzip von Bolzano-Weierstraß[8] angewandt auf die schwache*-Konvergenz gilt genau dann und nur dann, wenn $\dim(X) < \infty$ (vgl. Satz 4.11).

Einen Ausweg aus diesem Dilemma, der sehr schön in den Anwendungen ist, bietet der folgende

Satz 3.13. *Sei $(X, \|.\|_X)$ ein separabler normierter Vektorraum. Dann besitzt jede in X' beschränkte Folge eine schwach*-konvergente Teilfolge.*

Beweis. Sei also $(\varphi_n)_n \subset X'$ eine Folge mit

$$\|\varphi_n\|_{X'} \leq C \quad \text{für alle } n \in \mathbb{N}.$$

Weiter sei

$$\{\, x_n \mid n \in \mathbb{N} \,\} \subset X$$

eine abzählbar dichte Teilmenge von X.

[8] Jede beschränkte Folge besitzt eine konvergente Teilfolge

Dann ist $(\varphi_n(x_1))_n \subset \mathbb{K}$ beschränkt, also gibt es nach dem Satz von Bolzano-Weierstraß eine Teilfolge $(\varphi_{n_k^{(1)}}(x_1))_k$, so dass $(\varphi_{n_k^{(1)}}(x_1))_k$ in $\mathbb{K}$ konvergiert.

Weiter ist dann $(\varphi_{n_k^{(1)}})(x_2)$ beschränkt in $\mathbb{K}$. Also existiert wieder eine Teilfolge $(\varphi_{n_k^{(2)}}(x_2))_k$ von $(\varphi_{n_k^{(1)}}(x_2))_k$, so dass $\varphi_{n_k^{(2)}}(x_2))_k$ konvergiert in $\mathbb{K}$.

Nach dem Diagonalfolgenprinzip finden wir so eine Teilfolge $(\varphi_{n_k})_k$, so dass

$$\lim_{k\to\infty} \varphi_{n_k}(x_i) =: \varphi(x) \quad \text{existiert für alle } i \in \mathbb{N}. \tag{3.21}$$

Sei $x \in X$ beliebig und sei $\epsilon > 0$. Dann existiert wegen der Dichtheit von $\{\, x_n \mid n \in \mathbb{N} \,\}$ ein x_i, $i \in \mathbb{N}$, mit

$$\|x - x_i\|_X < \tfrac{\epsilon}{3(C+1)}.$$

Wegen (3.21) existiert ein $n_0 \in \mathbb{N}$ mit

$$|\varphi_{n_k}(x_i) - \varphi_{n_l}(x_i)| < \epsilon/3 \quad \text{für alle } n_k, n_l \geq n_0.$$

Dann gilt aber für alle $n_k, n_l \geq n_0$

$$\begin{aligned}
|\varphi_{n_k}(x) - \varphi_{n_l}(x)| &\leq |\varphi_{n_k}(x) - \varphi_{n_k}(x_i)| + |\varphi_{n_k}(x_i) - \varphi_{n_l}(x_i)| + |\varphi_{n_l}(x_i) - \varphi_{n_l}(x)| \\
&< \|\varphi_{n_k}\|_{X'} \|x - x_i\|_X + \epsilon/3 + \|\varphi_{n_l}\|_{X'} \|x_i - x\|_X \\
&< C\tfrac{\epsilon}{3(C+1)} + \epsilon/3 + C\tfrac{\epsilon}{3(C+1)} < \epsilon
\end{aligned}$$

d.h. $(\varphi_{n_k}(x))_k$ ist eine Cauchy-Folge in $\mathbb{K}$ und somit existiert

$$\lim \varphi_{n_k}(x) =: \varphi(x) \quad \text{für alle } x \in X \tag{3.22}$$

Es ist klar, dass φ linear ist und wegen

$$\|\varphi(x)\| = \lim \|\varphi_{n_k}(x)\| \leq \lim \|\varphi_{n_k}\| \|x\| \leq C\|x\| \quad \text{für alle } x \in X$$

ist dann $\varphi \in X'$. (3.22) besagt schließlich gerade

$$\varphi_{n_k} \rightharpoonup_* \varphi \quad \text{in } X'.$$

□

Beispiel 3.12. Sei $X = l^\infty$. Dann ist X nach Beispiel 1.21 nicht separabel. Betrachte $\varphi_n \in (l^\infty)'$ definiert durch

$$\varphi_n(x) = x_n \quad \text{für alle } x = (x_k)_k \in l^\infty.$$

Dann ist $\varphi_n : l^\infty \to \mathbb{K}$ linear und für alle $n \in \mathbb{N}$ gilt

$$\|\varphi_n\|_{(l^\infty)'} = \sup_{x \in B_{l^\infty}} |\varphi_n(x)| = \sup_{x \in B_{l^\infty}} |x_n| \leq 1,$$

d.h. $\varphi_n \in (l^\infty)'$ und insbesondere ist $(\varphi_n)_n$ auch beschränkt in $(l^\infty)'$.

Aber $(\varphi_n)_n$ besitzt keine schwach*-konvergente Teilfolge.

In der Tat: Ist $(\varphi_{n_k})_k$ eine beliebige Teilfolge und $x = (x_n)_n \in l^\infty$ definiert durch

$$x_n := \begin{cases} (-1)^k & , \text{ wenn } n = k, \\ 0 & , \text{ sonst,} \end{cases}$$

so gilt

$$\langle \varphi_{n_k}, x \rangle_{(l^\infty)', l^\infty} = x_{n_k} = (-1)^k$$

und dies konvergiert offensichtlich nicht für $k \to \infty$.

Bemerkungen.

1. Auf die Voraussetzung, dass X separabel ist, kann nicht verzichtet werden. Das haben wir eben gesehen.

2. Die schwach*-Konvergenz kann über die so genannte schwach*-Topologie eingeführt werden. Es gilt dann stets, dass $B_{X'}$ schwach*-kompakt ist (*Satz von Banach-Alaoglu-Bourbaki*).

   Allerdings ist Kompaktheit eben nur in metrischen Räumen äquivalent zur Folgenkompaktheit. Es gilt aber, dass die schwach*-Topologie auf $B_{X'}$ metrisierbar ist genau dann, wenn X separabel ist. In diesem Fall gilt:

   Ist $\{ x_n \mid n \in \mathbb{N} \}$ dicht in B_X, dann definiert

   $$d(\varphi, \psi) = \sum_{n=1}^{\infty} 2^{-n} \left|(\varphi - \psi)(x_n)\right|, \quad \varphi, \psi \in B_{X'}$$

   eine Metrik auf $B_{X'}$, die den schwach*-Konvergenzbegriff erzeugt:

   $$\varphi \rightharpoonup_* \varphi \Longleftrightarrow d(\varphi_n, \varphi) \to 0.$$

Kommen wir nun zur schwachen Konvergenz. Für sie gilt ein analoges Resultat nur in reflexiven Räumen. Dafür vorbereitend ist der folgende

Satz 3.14. *Sei $(X, \|.\|_X)$ ein normierter Vektorraum.*

(i) Ist X' separabel, so ist auch X separabel.

(ii) Sepziell gilt, wenn X reflexiv ist:

$$X \text{ separabel} \Longleftrightarrow X' \text{ separabel.}$$

Bemerkung. Die Umkehrung von (i) gilt im Allgemeinen nicht: Zum Beispiel ist $X = l^1$ nach Beispiel 1.20 separabel, aber $X' \cong l^\infty$ ist nicht separabel nach Beispiel 1.21. Vergleiche auch das Beispiel 3.13.

Beweis. Zu (ii): Die Richtung „$\Longleftarrow$“ ist gerade (i) und „$\Longrightarrow$“ folgt aus $X \cong X''$ und ebenfalls aus (i), denn wenn X separabel ist, dann auch X'' und wegen (i) dann auch X'.

Zu (i): Sei $(\varphi_n)_n$ eine abzählbar dichte Teilmenge in X'. Wir wissen

$$\|\varphi_n\|_{X'} = \sup_{x \in B_X} |\varphi_n(x)|.$$

Dann existiert zu jedem $n \in \mathbb{N}$ ein $x_n \in B_X$, so dass

$$\frac{1}{2}\|\varphi_n\|_{X'} \leq |\varphi_n(x_n)| \tag{3.23}$$

Setze

$$F := \operatorname{lin}\{ x_n \mid n \in \mathbb{N} \}.$$

Behauptung: F ist dicht in X.

Zum Nachweis dazu sei $\varphi \in X'$ mit

$$\varphi|_F = 0.$$

Sei $\epsilon > 0$. Dann existiert wegen der Dichtheitsvoraussetzung an φ_n ein $n \in \mathbb{N}$, so dass

$$\|\varphi_n - \varphi\| < \epsilon/3.$$

Dann gilt

$$\begin{aligned}\|\varphi\|_{X'} &\leq \|\varphi - \varphi_n\|_{X'} + \|\varphi_n\|_{X'} \overset{(3.23)}{<} \epsilon/3 + 2\,|\varphi_n(x_n)| \\ &\overset{(*)}{=} \epsilon/3 + 2\,|(\varphi_n - \varphi)(x_n)| \leq \epsilon/3 + 2\|\varphi_n - \varphi\|_{X'} \underbrace{\|x_n\|_X}_{\leq 1} < \epsilon,\end{aligned}$$

also $\varphi = 0$. An der Stelle $(*)$ geht ein, dass $x_n \in F$ und deswegen $\varphi|_F = 0$ gilt.

Nach Korollar 3.5 zum 2. Trennungssatz ist nun $\bar{F} = X$, also X separabel. □

Beispiel 3.13. Auch wenn l^1 und l^∞ nicht isometrisch isomorph zueinander sind, so ist die Abbildung

$$A : l^1 \to (l^\infty)', \quad (Ax)(y) = \sum_{n=1}^{\infty} x_n y_n$$

für $x = (x_n)$ und $y = (y_n)$ trotzdem isometrisch, aber nicht surjektiv.

Zur Isometrie: Zunächst ist nach der Hölderschen Ungleichung $xy \in l^1$ und es folgt

$$|(Ax)(y)| = \left|\sum_{n=1}^{\infty} x_n y_n\right| \leq \sum_{n=1}^{\infty} |x_n y_n| \leq \sup_k |y_k| \sum_{n=1}^{\infty} |x_n| = \|y\|_{l^\infty} \|x\|_{l^1},$$

also

$$\|Ax\| \leq \|x\|.$$

Wir setzen nun $x \neq 0$ voraus, weil die Behauptung für $x = 0$ sowieso gilt. Sei dann für alle $n \in \mathbb{N}$

$$y_n := \begin{cases} \frac{\bar{x}_n}{|x_n|} & \text{für } x_n \neq 0 \\ 0 & \text{für } x_n = 0. \end{cases}$$

Es folgt $\|y_n\|_{l^\infty} = 1$ und daher

$$\|Ax\| \geq \left|\sum_{n=1}^{\infty} |x_n|\right| = \|x\|_{l^1}.$$

Zusammenfassend ist also $\|Ax\| = \|x\|$ gezeigt.

Wir wollen nun zeigen, dass A nicht surjektiv ist. Betrachte dazu das Funktional

$$\lim : c \to \mathbb{K}.$$

Setze dieses mit dem Satz von Hahn-Banach zu einem stetigen Funktional

$$\varphi : l^\infty \to \mathbb{K}$$

fort. Angenommen, φ hätte eine Darstellung

$$\varphi(y) = \sum_{n=1}^{\infty} x_n y_n,$$

so wäre

$$x_k = \varphi(e_k) = \lim e_k = 0 \quad \text{für alle } k \in \mathbb{N} \Longrightarrow \varphi = 0.$$

Widerspruch!

□

Satz 3.15. *Sei $(X, \|.\|_X)$ ein reflexiver Banachraum. Dann besitzt jede in X beschränkte Folge eine schwach konvergente Teilfolge.*

Bemerkungen.

1. In nicht-reflexiven Räumen gilt der Satz im Allgemeinen nicht. Vergleiche das Beispiel 3.12 mit der beschränkten Folge $(\varphi_n)_n$ in $(l^\infty)'$. Wir hatten gesehen, dass $(\varphi_n)_n$ keine schwach*-konvergente Teilfolge besitzt. Damit besitzt sie erst recht keine schwach konvergente Teilfolge.

2. Es gibt eine genauere Version des oberen Satzes:

   Satz von Banach-Eberlein-Shmulyan.
   Sei $(X, \|.\|_X)$ ein Banachraum. Dann ist X reflexiv genau dann, wenn jede beschränkte Folge in X eine schwach konvergente Teilfolge besitzt.

   Wir haben leider nicht die Mittel, um „$\Longleftarrow$“ zu zeigen.

Beweis von Satz 3.15. Sei $(x_n)_n$ eine beschränkte Folge in X. Betrachte

$$F := \overline{\text{lin}}\{\, x_n \mid n \in \mathbb{N} \,\}$$

Als abgeschlossener linearer Teilraum von X ist F nach Satz 3.8 reflexiv. Wegen Satz 1.5 ist F aus Konstruktionsgründen separabel. Also ist nach Satz 3.14 (i) auch F' separabel. F' ist ein Banachraum, also nach Satz 3.9 auch reflexiv. Also sind wir in folgender Situation:

$$F' \text{ reflexiv} + \text{separabel} \overset{\text{Satz 3.14 (ii)}}{\Longrightarrow} F'' \text{ separabel.}$$

Da $F'' \cong F$, betrachten wir nun

$$(j_{x_n})_n \subset F'' \cong F.$$

Diese sind beschränkt, da die kanonische Abbildung j isometrisch ist. Somit folgt jetzt aus Satz 3.13 (denn F'' ist ja separabel), dass $(j_{x_n})_n$ eine schwach*-konvergente Teilfolge besitzt, d.h.

$$\left\langle j_{x_{n_k}}, \varphi \right\rangle \to \langle j_x, \varphi \rangle \quad \text{für alle } \varphi \in F'$$

für eine geeignete Teilfolge $(x_{n_k})_k$ und $x \in F$ (hier geht ein: $F'' = j(F)$). Das heißt aber gerade

$$\langle \varphi, x_{n_k} \rangle \to \langle \varphi, x \rangle \quad \text{für alle } \varphi \in F'.$$

Da für jedes $\varphi \in X'$ gilt, dass $\varphi|_F \in F' = L(F, \mathbb{K})$ und alle $x_{n_k}, x \in F$, folgt somit für alle $\varphi \in X'$

$$\lim_{k\to\infty} \langle \varphi, x_{n_k} \rangle = \lim_{k\to\infty} \langle \varphi|_F, x_{n_k} \rangle = \langle \varphi|_F, x \rangle = \langle \varphi, x \rangle$$

d.h. $x_{n_k} \rightharpoonup x$ in X. □

Nützlich ist oft folgendes Teilfolgenprinzip

Satz 3.16. *Sei X ein reflexiver Banachraum. Besitzen alle schwach konvergenten Teilfolgen einer beschränkten Folge $(x_n)_n \subset X$ ein und denselben Limes $x \in X$, dann konvergiert die gesamte Folge $(x_n)_n$ schwach gegen $x \in X$.*

Beweis. Angenommen, $(x_n)_n$ konvergiert nicht schwach gegen x in X. Dann existiert ein $\epsilon > 0$, eine aufsteigende Folge $(n_k)_k$ mit $n_k \to \infty$ und ein $\varphi \in X'$, so dass

$$|\, \langle \varphi, x_{n_k} \rangle_{X',X} \,| \geq \epsilon \quad \text{für alle } k \in \mathbb{N}. \tag{3.24}$$

Nach Voraussetzung ist aber $(x_n)_n$ beschränkt und somit auch die Teilfolge $(x_{n_k})_k$. Nach dem Satz von Eberlein-Smulian besitzt daher $(x_{n_k})_k$ eine schwach konvergente Teilfolge, da X reflexiv ist. Nach Voraussetzung konvergiert diese Teilfolge, nennen wir sie $(x_{n'_k})_k$, schwach gegen denselben Limes x. Andererseits gilt für die Teilfolge aber auch die Ungleichung (3.24). Dies ist ein Widerspruch! □

Bemerkung. Das Teilfolgenprinzip gilt i.A. nicht in nicht-reflexiven Räumen.

In Dualräumen gilt ein analoges Teilfolgenprinzip für die schwach$*$-Konvergenz:

Satz 3.17. *Sei X ein separabler normierter Raum. Besitzen alle schwach$*$-konvergenten Teilfolgen einer beschränkten Folge $(\varphi_n)_n \subset X'$ ein und denselben Limes $\varphi \in X'$, dann konvergiert die gesamte Folge (φ_n) schwach$*$ gegen $\varphi \in X'$.*

Bemerkung. Auf die Voraussetzung, dass X separabel ist, kann im Satz 3.17 nicht verzichtet werden.

3.5 Adjungierte und kompakte Operatoren

Seien X und Y Vektorräume. Ähnlich wie man jedem X den Dualraum zuordnet, kann man jedem Operator $A \in L(X,Y)$ einen „dualen“ Operator A' zuordnen:

Definition 3.6 (Adjungierte Operator). Seien X, Y Vektorräume und $A \in L(X,Y)$. Der Operator[9] $A' : Y' \to X'$ definiert durch

$$(A'y')(x) := y'(Ax) \quad \text{für alle } x \in X \text{ und } y' \in Y'$$

heißt der zu A *adjungierte Operator*.

Wir zeigen zunächst, dass A' wohldefiniert ist, d.h. es ist $A' \in L(Y', X')$:

In der Tat ist $A'y' \in X'$. Auf der einen Seite ist nämlich $A'y' : X \to \mathbb{K}$ linear: Für alle $x_1, x_2 \in X$ und $\alpha, \beta \in \mathbb{K}$ gilt

$$\begin{aligned}(A'y')(\alpha x_1 + \beta x_2) &= y'(A(\alpha x_1 + \beta x_2)) = \alpha y'(Ax_1) + \beta y'(Ax_2) \\ &= \alpha (A'y')(x_1) + \beta (A'y')(x_2).\end{aligned}$$

Auf der anderen Seite ist $A'y'$ stetig wegen

$$\begin{aligned}\|A'y'\|_{X'} &= \sup_{x \in B_X} |(A'y')(x)| = \sup_{x \in B_X} |y'(Ax)| \le \sup_{x \in B_X} \|y'\|_{Y'} \|Ax\|_Y \\ &= \|y'\|_{Y'} \sup_{x \in B_X} \|Ax\|_Y = \|y\|_{Y'} \|A\|_{L(X,Y)}.\end{aligned} \tag{3.25}$$

Insgesamt folgt also $A'y' \in X'$.

Es bleibt zu zeigen, dass $A' \in L(Y', X')$: A' ist linear, denn für alle $\alpha, \beta \in \mathbb{K}$ und $y_1', y_2' \in Y'$ gilt

$$\begin{aligned}A'(\alpha y_1' + \beta y_2')(x) &= (\alpha y_1' + \beta y_2')(Ax) = \alpha y_1'(Ax) + \beta y_2'(Ax) \\ &= \alpha A'(y_1')(x) + \beta A'(y_2')(x),\end{aligned}$$

und das für alle $x \in X$, also

$$A'(\alpha y_1' + \beta y_2') = \alpha A'(y_1') + \beta A'(y_2').$$

Außerdem ist A' stetig wegen

$$\|A'\|_{L(X',Y')} = \sup_{y' \in B_{Y'}} \|A'y'\| \overset{(3.25)}{\le} \sup_{y' \in B_{Y'}} \|y'\|_{Y'} \|A\|_{L(X,Y)} \le \|A\|_{L(X,Y)}.$$

Also ist $A' \in L(Y', X')$ und schließlich die Wohldefiniertheit von A' gezeigt.

Beispiel 3.14. Seien X, Y endlich-dimensionale normierte Vektorräume und $A \in L(X,Y)$ repräsentiert durch eine Matrix $[A]$. Dann ist nach *Linearer Algebra* $A' \in L(Y', X')$ repräsentiert durch die transponierte Matrix $[A]^*$.

Konkret sei dazu

$$A : \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 3 \end{pmatrix} : \mathbb{R}^3 \to \mathbb{R}^2$$

gegeben. Die Dualräume sind $(\mathbb{R}^3)' = \mathbb{R}^3$ und $(\mathbb{R}^2)' = \mathbb{R}^2$. Seien

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \quad y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \quad \text{und} \quad y' = \begin{pmatrix} y_1' \\ x_2' \end{pmatrix}.$$

[9] In der amerikanischen Literatur schreibt man für den adjungierten Operator oft A^*

Dann gilt nach *Linearer Algebra*

$$y'(y) = y_1' y_1 + y_2' y_2.$$

Weiter ist

$$A' : \mathbb{R}^3 \to \mathbb{R}^2,$$

also

$$\begin{aligned}(A'y')(x) = y'(Ax) = y'\left(\begin{pmatrix} x_1 + 2x_2 \\ x_2 + 3x_3 \end{pmatrix}\right) &= y_1'(x_1 + 2x_2) + y_2'(x_2 + 3x_3) \\ &= x_1(y_1') + x_2(2y_1' + y_2') + x_3(3y_2'),\end{aligned}$$

d.h.

$$A'y' = A'\begin{pmatrix} y_1' \\ y_2' \end{pmatrix} = \begin{pmatrix} y_1' \\ 2y_1' + y_2' \\ 3y_2' \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 2 & 1 \\ 0 & 3 \end{pmatrix}\begin{pmatrix} y_1' \\ y_2' \end{pmatrix}.$$

Schließlich folgt daraus

$$A' = \begin{pmatrix} 1 & 0 \\ 2 & 1 \\ 0 & 3 \end{pmatrix} = A^T.$$

□

Beispiel 3.15 (**Linksshift-Operator**). Seien $1 \leq p < \infty$ und $X = Y = l^p$. Es gilt (vgl. Bemerkung zu Beispiel 2.10):

$$X' = Y' \cong l^q \quad \text{mit } \tfrac{1}{p} + \tfrac{1}{q} = 1.$$

Betrachte den *Linksshift-Operator*

$$A : l^p \to l^p, \quad (x_1, x_2, x_3, \ldots) \mapsto (x_2, x_3, x_4, \ldots,).$$

Offensichtlich ist $A \in L(l^p)$. Wegen

$$Ax = A(x_n)_n = (x_{n+1})_n$$

gilt dann

$$(A'y')(x) = y'(Ax) = \sum_{n=1}^{\infty} y_n \cdot x_{n+1} = \sum_{n=2}^{\infty} y_{n-1} x_n,$$

d.h.

$$A'y' = (0, y_1, y_2, y_3, \ldots) \quad \text{für alle } y' \cong (y_n)_n \in l^q.$$

Demnach ist A' der so genannte *Rechtsshift-Operator* auf l^q.

□

Beispiel 3.16 (**Kernoperator**). Seien $X = Y = L^2(0,1)$ und $k \in C([0,1]^2)$ (oder allgemeiner $k \in L^2((0,1)^2)$. Betrachte den *Kern-Operator*

$$A : L^2(0,1) \to L^2(0,1), \quad u \mapsto \int_0^1 k(.,t)u(t)dt.$$

Wir haben in Beispiel 2.5 gesehen, dass $A \in L(L^2(0,1))$. Außerdem ist nach der Bemerkung zu Beispiel 2.10

$$X' = Y' \cong L^2(0,1)$$

und der isometrische Isomorphismus $\phi : L^2(0,1) \to (L^2(0,1))'$ ist gegeben durch

$$\phi(f)(g) = \int_0^1 f(s)g(s)ds \quad \text{für alle } f, g \in L^2(0,1),$$

wie der Darstellungssatz von Riesz-Frechét (Satz 5.4) noch zeigen wird. Es folgt

$$\begin{aligned}(A'y')(u) &= y'(Au) = \int_0^1 y'(s)Au(s)ds = \int_0^1 y'(s)\int_0^1 k(s,t)u(t)dtds \\ &\overset{\text{Fubini}}{=} \int_0^1 u(t)\left(\int_0^1 k(s,t)y'(s)ds\right)dt,\end{aligned}$$

d.h. aber gerade

$$A'y' = \int_0^1 k(s,.)y'(s)ds \quad \text{für alle } y' \in Y \cong L^2.$$

Wir stellen fest, dass A' und A ziemlich ähnlich aussehen. Er vertauschen anscheinend nur die Variablen in k.

□

Satz 3.18. *Seien X, Y, Z normierte Vektorräume.*

(i) Die Abbildung $\phi : A \in L(X,Y) \to A' \in L(Y',X')$ ist linear und isometrisch, d.h.

$$\|A\|_{L(X,Y)} = \|A'\|_{L(Y',X')}.$$

(ii) Für alle $B \in L(X,Y)$ und $A \in L(Y,Z)$ gilt

$$(A \circ B)' = B' \circ A'.$$

Bemerkung. Die Abbildung $A \to A'$ ist im Allgemeinen nicht surjektiv. Vergleiche Beispiel **??**.

Beweis. Zu (i): Die Abbildung ϕ ist in der Tat linear: Sind nämlich $A, B \in L(X,Y)$ und $\alpha, \beta \in \mathbb{K}$, so gilt

$$\phi(\alpha A + \beta B) = (\alpha A + \beta B)' = \alpha A' + \beta B' = \alpha\phi(A) + \beta\phi(B).$$

Die Isometrie folgt aus

$$\begin{aligned}\|A\|_{L(X,Y)} &= \sup_{x\in B_X} \|Ax\|_Y \overset{\text{Normformel}}{=} \sup_{x\in B_X}\sup_{y'\in B_{Y'}} |y'(Ax)| \\ &= \sup_{y'\in B_{Y'}}\sup_{x\in B_X} |y'(Ax)| = \sup_{y'\in B_{Y'}}\sup_{x\in B_X} |(A'y')(x)| \\ &= \sup_{y'\in B_{Y'}} \|A'y'\|_{X'} = \|A'\|_{L(Y',X')}.\end{aligned}$$

Zu (ii): Seien also $A \in L(Y,Z)$ und $B \in L(X,Y)$, also $A' \in L(Z',Y')$ und $B' \in L(Y',X')$ und somit

$$(A \circ B)' \in L(Z',X') \quad \text{sowie} \quad B' \circ A' \in L(Z',X').$$

Dann gilt für alle $x \in X$ und $z \in Z$

$$(B' \circ A')(z')(x) = B'(A'(z'))(x) = A'(z')(Bx) = z'(ABx) = (AB)'(z')(x),$$

also

$$(A \circ B)' = B' \circ A'.$$

□

Wir kommen nun zu den kompakten Operatoren:

Definition 3.7 (Kompakter Operator). Seien X, Y normierte Vektorräume. Ein Operator $A \in L(X,Y)$ heißt *kompakt*, wenn $A(B_X)$ relativ kompakt in Y ist.

Beispiel 3.17. Seien X, Y normierte Vektorräume mit $\dim X = n < \infty$. Dann ist jede lineare stetige Abbildung $A : X \to Y$ kompakt. Nach Satz 4.11 ist nämlich B_X kompakt und daher ist dann auch $A(B_X)$ kompakt, weil A stetig ist.

□

Bemerkungen.

1. A ist genau dann kompakt, wenn A beschränkte Mengen auf relativ-kompakte Mengen abbildet.
2. A ist genau dann kompakt, wenn für jede beschränkte Folge $(x_n)_n \subset X$ eine Teilfolge $(x_{n_k})_k$ existiert, so dass $(Ax_{n_k})_k$ in Y konvergiert.

Beweis. **Zu 1.** Zu „$\Longrightarrow$“: Sei also $M \subset X$ beschränkt. Dann gibt es ein $r > 0$, so dass $M \subset r \cdot B_X$. Wegen der Linearität von A folgt dann

$$A(M) \subset A(r \cdot B_X) = r \cdot A(B_X)$$

und somit

$$\overline{A(M)} \subset r\overline{A(B_X)}.$$

Weil $\overline{A(B_X)}$ kompakt ist, ist damit auch $\overline{A(M)}$ kompakt.

Zu „$\Longleftarrow$“: Trivial.

Zu 2. Zu „$\Longrightarrow$“: Sei also A kompakt, d.h. $\overline{A(B_X)}$ ist kompakt. Sei weiter $(x_n)_n$ eine beschränkte Folge, d.h. es gibt ein $M > 0$, so dass

$$\|x_n\|_X \leq M \quad \text{für alle } n \in \mathbb{N}.$$

Beachte, dass der Fall $M = 0$ trivial wäre. Es folgt $\frac{x_n}{M} \in B_X$ für alle $n \in \mathbb{N}$ und so

$$A(\tfrac{x_n}{M}) \in \overline{B_X} \quad \text{für alle } n \in \mathbb{N}.$$

Daher gibt es eine Teilfolge $(x_{n_k})_k$, so dass $(A(\frac{x_{n_k}}{M}))_k$ konvergiert. Wegen

$$A(\tfrac{x_{n_k}}{M}) = \tfrac{1}{M} A(x_{n_k})$$

ist dann auch $A(x_{n_k})$ konvergent.

Zu „$\Longleftarrow$“: Wir müssen zeigen, dass $\overline{A(B_X)}$ kompakt ist, d.h. jede Folge aus $\overline{A(B_X)}$ besitzt eine konvergente Teilfolge.

Seien also $y_n \in \overline{A(B_X)}$ beliebig und $z_n \in \overline{A(B_X)}$ etwa mit $z_n = Ax_n$ und $\|x_n\|_X \leq 1$, so dass

$$\|y_n - z_n\|_Y \leq 1/n.$$

Da $(x_n)_n$ beschränkt ist, gibt es nach Voraussetzung eine Teilfolge $(x_{n_k})_k$ mit

$$\lim_{k \to \infty} A(x_{n_k}) = z.$$

Daraus folgt

$$\|y_{n_k} - z\|_Y \leq \|y_{n_k} - z_{n_k}\|_Y + \|z_{n_k} - z\|_Y \to 0.$$

Also ist $\overline{A(B_X)}$ kompakt.

□

Beispiel 3.18. Seien $k \in C([0,1]^2)$ und

$$A : C[0,1] \to C[0,1], \quad u \mapsto \int_0^1 k(.,s)u(s)ds.$$

Wir haben in Beispiel 2.5 gesehen, dass $A \in L(C[0,1])$.

Behauptung: A ist kompakt.

Dazu müssen wir zeigen, dass

$$\{\, Au \mid u \in B_{C[0,1]} \,\}$$

relativ kompakt in $(C[0,1], \|.\|_\infty)$ ist. Das folgt aber aus dem Satz von Arzelà-Ascoli (Satz 1.15), denn

- $\{ Au \mid u \in B_{C[0,1]} \}$ ist punktweise beschränkt wegen

$$|Au(t)| = \left| \int_0^1 k(s,t)u(s)ds \right| \leq \int_0^1 |k(s,t)|\,|u(s)|\,ds \overset{\|u\|\leq 1}{\leq} \int_0^1 |k(s,t)|\,ds < \infty,$$

  weil $k \in C([0,1]^2)$.

- $\{ Au \mid u \in B_{C[0,1]} \}$ ist gleichgradig stetig: Für alle $u \in B_{C[0,1]}$ gilt nämlich

$$\left|Au(t) - Au(\tilde{t})\right| \overset{\|u\|\leq 1}{\leq} \int_0^1 \left|k(t,s) - k(\tilde{t},s)\right| ds < \epsilon$$

  für $\left|t - \tilde{t}\right| < \delta$, wobei $\delta > 0$ gemäß der gleichmäßigen Stetigkeit von k auf $[0,1]^2$ gewählt ist. D.h. aber gerade, dass $\{ Au \mid u \in B_{C[0,1]} \}$ gleichgradig stetig ist.

□

Eigenschaften kompakter Operatoren.

1. Seien X und Y Banachräume. Dann ist

$$K(X,Y) := \{ A \in L(X,Y) \mid A \text{ kompakt} \}$$

   ein abgeschlossener linearer Untervektorraum von $L(X,Y)$. Das heißt insbesondere, dass $(K(X,Y), \|.\|_{L(X,Y)})$ ein Banachraum ist.

2. Seien X, Y, Z und W normierte Vektorräume und $A \in K(X,Y)$, $S \in L(Y,Z)$ und $T \in L(W,X)$. Dann gilt:

$$S \circ A \in K(X,Z) \quad \text{und} \quad A \circ T \in K(W,Y). \tag{3.26}$$

3. Betrachte die so genannten *finite rank Operators*

$$F(X,Y) := \{ A \in L(X,Y) \mid \text{Bild}(A) \text{ ist endlich-dimensionaler UVR von } Y \}.$$

   Dann gilt $F(X,Y) \subset K(X,Y)$.

   Insbesondere gilt $\overline{F(X,Y)} \subset K(X,Y)$: Ist also $(A_n)_n \subset F(X,Y)$ eine Folge mit $\|A_n - A\| \to 0$, so folgt $A \in K(X,Y)$.

Beweis. <u>Zu 1:</u> Zunächst ist $K(X,Y)$ tatsächlich ein Untervektorraum, denn

- Mit A ist auch λA kompakt für $\lambda \in \mathbb{K}$.
- Seien $A, B \in K(X,Y)$ und sei $(x_n)_n \subset X$ eine beschränkte Folge. Wähle dann eine Teilfolge $(x_{n_k})_k$, so dass $(Ax_{n_k})_k$ konvergiert. Da auch B kompakt ist, können wir eine Teilfolge $(x_{n_{k_l}})_l$ wählen, so dass $(B(x_{n_{k_l}}))_l$ konvergiert. Dann konvergiert auch

$$(A(x_{n_{k_l}}) + B(x_{n_{k_l}}))_l$$

  und $A + B$ ist kompakt.

Wir haben damit gezeigt, dass $K(X,Y)$ ein Untervektorraum ist. Um die Abgeschlossenheit von $K(X,Y)$ zu zeigen, sei $A \in \overline{K(X,Y)}$. Weiter sei $\epsilon > 0$. Dann existiert ein Operator $A_\epsilon \in K(X,Y)$ mit

$$\|A - A_\epsilon\|_{L(X,Y)} = \sup_{x \in B_X} \|Ax - A_\epsilon x\|_Y \leq \epsilon/2 \tag{3.27}$$

$A_\epsilon B_X$ ist relativ kompakt in Y, d.h. insbesondere präkompakt und so existieren $y_1, \ldots, y_n$ mit

$$A_\epsilon B_X \subset \bigcup_{i=1}^n B_Y(y_i, \epsilon/2). \tag{3.28}$$

Sei nun $x \in B_X$. Dann existiert wegen (3.28) ein $i_0 \in \{1, \dots, n\}$, so dass

$$\|A_\epsilon x - y_{i_0}\| \leq \epsilon/2. \tag{3.29}$$

Somit folgt dann aus (3.27) und (3.29)

$$\|Ax - y_{i_0}\| \leq \|Ax - A_\epsilon x\| + \|A_\epsilon x - y_{i_0}\| \leq \epsilon,$$

d.h.

$$AB_X \subset \bigcup_{i=1}^{n} B_Y(y_i, \epsilon),$$

also ist AB_X präkompakt und somit, da Y vollständig ist, relativ kompakt. Demnach ist $A \in \overline{K(X,Y)}$ und $K(X,Y)$ ist abgeschlossen.

Zu 2:

- $S \circ A$: Sei $(x_n)_n$ eine beschränkte Folge in X. Da A kompakt ist, exsitiert eine Teilfolge $(x_{n_k})_k$, so dass $(Ax_{n_k})_k$ in Y konvergiert. Da S stetig ist, ist dann auch $(S \circ A)(x_{n_k})_k$ konvergent in Z.
- $A \circ T$: Analog.

Zu 3: Sei $A \in F(X,Y)$. Da A stetig ist, ist AB_X eine beschränkte Teilmenge des endlichdimensionalen Vektorraums $A(X)$. Wir haben also aufgrund des Satzes von Heine-Borel

$$\overline{AB_X} \text{ abgeschlossen und beschränkt} \Longrightarrow \overline{AB_X} \text{ kompakt.}$$

Der Zusatz ist trivial wegen

$$\overline{F(X,Y)} \subset \overline{K(X,Y)} \overset{1.}{=} K(X,Y).$$

□

Bemerkung. Es war bist 1974 ein offenes Problem, ob für beliebige Banachräume Y gilt, dass

$$\overline{F(X,Y)} = K(X,Y).$$

Enflo gab dazu die negative Antwort (vergleiche Abschnitt 4.1).

Satz 3.19 (von Schauder). *Seien X und Y Banachräume und $A \in L(X,Y)$. Dann gilt*

$$A \in K(X,Y) \Longleftrightarrow A' \in K(Y', X').$$

Beweis. Zu „$\Longrightarrow$“: Sei also $(y'_n)_n$ eine beschränkte Folge in Y'. Wir müssen zeigen, dass $(A'y'_n)_n \subset \overline{L(Y,\mathbb{K})}$ eine konvergente Teilfolge in Y' besitzt.

Weil A kompakt ist, ist

$$K := \overline{AB_X}$$

eine kompakte Teilmenge von Y. Betrachte nun

$$f_n := y'_n|_K, \quad n \in \mathbb{N}. \tag{3.30}$$

Da $(y'_n)_n$ beschränkt in Y' ist, ist $(f_n)_n$ eine punktweise beschränkte Familie in $C(K,\mathbb{K})$. Die Familie ist außerdem gleichgradig stetig, denn für alle $n \in \mathbb{N}$ und $k, l \in K$ gilt

$$|f_n(k) - f_n(l)| = |y'_n(k) - y'_n(l)| = |y'_n(k-l)| \leq \|y'_n\|_{Y'} \|k-l\|_X \leq M\|k-l\|_X.$$

Da K ein kompakter metrischer Raum ist, folgt mit dem Satz von Arzelà-Ascoli (Satz 1.15), dass eine Teilfolge $(f_{n_k})_k$ existiert, die gleichmäßig auf dem kompakten(!) K konvergiert. Weil

$$\begin{aligned}
\left\|A'y'_{n_k} - A'y'_{n_l}\right\|_{X'} &= \sup_{x\in B_X} \left|A'y'_{n_k}(x) - A'y'_{n_l}(x)\right| \overset{\text{Def}}{=} \sup_{x\in B_X} \left|y'_{n_k}(Ax) - y'_{n_l}(Ax)\right| \\
&= \sup_{z\in AB_X} \left|y'_{n_k}(z) - y'_{n_l}(z)\right| \overset{(3.30)}{=} \sup_{z\in AB_X} |f_{n_k}(z) - f_{n_l}(z)| \\
&\overset{\overline{AB_X}=K}{\leq} \sup_{z\in K} |f_{n_k}(z) - f_{n_l}(z)| = \|f_{n_k} - f_{n_l}\|_\infty
\end{aligned}$$

und $(f_{n_k})_k$ eine Cauchy-Folge in $(C(K,\mathbb{K}), \|.\|_\infty)$ ist, ist $(A'y'_{n_k})_k$ eine Cauchy-Folge in X' und konvergiert deswegen in X'.

Damit ist die Existenz der konvergenten Teilfolge gezeigt.

<u>Zu „$\Longleftarrow$“</u>: Sei also A' kompakt. Nach Teil 1. des Beweises („$\Leftarrow$“) ist dann

$$A'' \in K(X'', Y'').$$

Da $j_X : X \to X''$ stetig (und sogar isometrisch) ist, ist dann nach Eigenschaft 2.

$$A'' \circ j_X \in K(X, Y'').$$

Es gilt außerdem

$$\begin{array}{ccccc} & X & \xrightarrow{A} & Y & \\ j_X & \downarrow & & \downarrow & j_Y \\ & X'' & \xrightarrow{A''} & Y'' & \end{array}, \qquad A'' \circ j_X = j_Y \circ A. \tag{3.31}$$

Beweis von (3.31). Einfaches Nachrechnen zeigt

$$[A''(j_X(x))](y') = j_X(x)(A'y') = (A'y')(x) = y'(Ax) = j_Y(Ax)(y')$$

für alle $y' \in Y'$ und alle $x \in X$. Daraus folgt

$$A''(j_X(x)) = j_Y(Ax) \quad \text{in } Y'' \text{ für alle } x \in X,$$

also

$$A'' \circ j_X = j_Y \circ A$$

und das zeigt (3.31).

Zurück zum eigentlichen Beweis: Weil $A'' \circ j_X \in K(X, Y'')$, ist also nach (3.31)

$$j_Y \circ A \in K(X, Y'').$$

Weil j_Y isometrisch ist, folgt somit $A \in K(X, Y)$.

In der Tat: Sei $(x_n)_n$ eine beschränkte Folge in X. Da $j_Y \circ A$ kompakt ist, gibt es eine Teilfolge $(x_{n_k})_k$, so dass $(j_Y(Ax_{n_k}))_k$ in Y'' konvergiert. Weil j_Y eine Isometrie ist, ist auch $(Ax_{n_k})_k$ eine Cauchy-Folge in Y. Da Y vollständig ist, konvergiert $(Ax_{n_k})_k$ in Y. □

Zammenhänge zwischen A und A'.

Sei $A : X \to Y$ ein linearer Operator aus normierten Vektorräumen X und Y. Dann bezeichnet man

$$\ker(A) := \{\, x \in X \mid Ax = 0 \,\}$$

als den *Kern von* A und

$$R(A) := \{\, Ax \mid x \in X \,\}$$

heißt das *Bild von* A. Weiter bezeichnet man für Untervektorräume $U \subset X$ und $V \subset X'$

$$U^\perp := \{\, x' \in X' \mid x'(x) = 0 \text{ für alle } x \in U \,\}$$

als den *Annihilator von U in X'* und

$$V_\perp := \{\, x \in X \mid x'(x) = 0 \text{ für alle } x' \in V \,\}$$

als den *Annihilator von V in X.*

$U^\perp$ (sprich: U polare oben) bzw. $V_\perp$ (sprich: V polare unten) sind stets abgeschlossene Untervektorräume von X' und X.

Satz 3.20. *Seien X, Y normierte Vektorräume und $A \in L(X, Y)$. Dann gilt*

(i) $\ker A' = R(A)^\perp$.

(ii) $\ker A = R(A')_\perp$.

(iii) $\overline{R(A)} = \ker(A')_\perp$.

(iv) $\overline{R(A')} \subset \ker(A)^\perp$.

Bemerkung. Sind X, Y Banachräume, so gilt sogar Gleichheit in (iv). Das ist eine Aussage des Satzes vom abgeschlossenen Bild. Siehe Anhang 7.4.

Beweis. **Zu (i).** „$\subset$“: Sei also $y' \in \ker(A')$, d.h. $A'y' = 0$. Dann gilt für alle $x \in X$

$$y'(Ax) = (A'y')(x) = 0$$

und somit ist $y' \in R(A)^\perp$.

„$\supset$“: Sei $y' \in R(A)^\perp$. Dann ist für alle $x \in X$

$$A'y'(x) = y'(Ax) = 0,$$

d.h. $A'y' = 0$ in X', also $y' \in \ker(A')$.

Zu (ii). „$\subset$“: Sei $x \in \ker(A)$, d.h. $Ax = 0$. Dann gilt für $y' \in Y'$

$$(A'y')(x) = y'(Ax) = 0$$

und somit $x \in R(A')_\perp$.

„$\supset$“: Sei $x \in R(A')_\perp$. Dann gilt für alle $y' \in Y'$

$$y'(Ax) = (A'y')(x) = 0,$$

also $Ax = 0$ in Y und somit $x \in \ker(A)$.

Zu (iii). „$\subset$“: Da $\ker(A)_\perp$ ein abgeschlossener Untervektorraum ist, genügt es zu zeigen, dass

$$R(A) \subset \ker(A')_\perp.$$

Sei also $y = Ax \in R(A)$. Dann ist für alle $y' \in \ker(A')$

$$y'(y) = y'(Ax) = (A'y')(x) = 0,$$

also $y' \in \ker(A')_\perp$.

„$\supset$“: Wir müssen zeigen, dass

$$\overline{R(A)} \supset \ker(A')_\perp \iff \overline{R(A)} \subset (\ker(A')_\perp)^C.$$

Sei also $y \in \overline{R(A)}^C$. Dann existiert nach Korollar 3.4 des Satzes von Hahn-Banach ein $y' \in Y'$ mit

$$y'(y) \neq 0 \quad \text{und} \quad y'|_{\overline{R(A)}} = 0.$$

Damit gilt dann:

$$(A'y')(x) = y'(Ax) = 0 \quad \text{für alle } x \in X,$$

d.h. $A'y' = 0$, also $y' \in \ker(A')$. Somit folgt dann, dass $y \in (\ker(A')_\perp)^C$, denn sonst wäre ja $y'(y) = 0$.

Zu (iv). Wieder genügt es zu zeigen, dass

$$R(A') \subset \ker(A)^\perp.$$

Sei also $A'y' \in R(A')$. Dann gilt für alle $x \in \ker(A)$

$$(A'y')(x) = y'(Ax) = 0.$$

Also gilt $A'y' \in \ker(A)^\perp$. □

Mit Hilfe der obigen „Orthogonalitätsbeziehungen“ zwischen dem Kern und dem Bild eines Operators bzw. des adjungierten Operators erhält man zum Beispiel

Korollar 3.7. *Sei $A \in L(X, Y)$ ein Operator mit abgeschlossenem Bild $R(A)$ in Y. Dann gilt*

$$A \text{ ist surjektiv } (R(A)=Y) \Longleftrightarrow A' \text{ ist injektiv } (\ker(A') = \{0\}).$$

Genauer: Die Operatorgleichung $Ax = y$ ist für $y \in Y$ lösbar genau dann, wenn

$$A'y' = 0 \Longrightarrow y'(y) = 0,$$

d.h. $y \in \ker(A')_\perp$

Beweis. Nach dem letzten Satz gilt

$$\overline{R(A)} = R(A) = \ker(A')_\perp.$$

Gilt also $Y = R(A)$, so folgt $\ker(A') = \{0\}$ und umgekehrt folgt aus $\ker(A') = \{0\}$ sofort $Y = R(A)$. □

4 Kompakte Operatoren

- Eine Antwort auf die Frage $\overline{F(X,Y)} = K(X,Y)$
- Spektraltheorie: Resolvente, Resolventenmenge, Spektrum, Resolventengleichung
- Bestandteile des Spektrums: Punkt-, kontinuierliches- und Restspektrum, Berechnung derselben
- Lemma und Kompaktheitssatz von Riesz
- Fredholm'sche Alternative und Anwendung auf Integralgleichung
- Spektralsatz für kompakte Operatoren
- Quotientenräume und Dimensionsformel für kompakte Operatoren

4.1 Eine Antwort auf $\overline{F(X,Y)} \stackrel{?}{=} K(X,Y)$

In diesem Abschnitt seien - wenn nichts anderes gesagt wird - X und Y Banachräume. Wir wollen der Frage nachgehen, wann eigentlich

$$\overline{F(X,Y)} = K(X,Y)$$

gilt. Bevor wir dazu kommen, wollen wir zunächst einmal klären, wie die *finite rank Operatoren* aussehen:

Sei $A \in F(X,Y)$. Dann existiert eine Basis $y_1, \ldots, y_n$ von $A(X)$, d.h. es gibt zu jedem $x \in X$ genau ein Element

$$(\varphi_1(x), \varphi_2(x), \ldots, \varphi_n(x)) \in \mathbb{K}^n$$

mit

$$Ax = \sum_{i=1}^{n} \varphi_i(x) y_i.$$

Weil A linear und die $(y_i)_{i=1,\ldots,n}$ linear unabhängig sind, ist dann auch

$$\varphi_i : X \to \mathbb{K}$$

linear für alle $i = 1, \ldots, n$.

Beweis. Seien also $x, y \in X$ und $\lambda, \mu \in \mathbb{K}$. Dann ist auf der einen Seite

$$A(\lambda x + \mu y) = \sum_{i=1}^{n} \varphi_i(\lambda x + \mu y) y_i$$

und auf der anderen Seite wegen der Linearität von A

$$A(\lambda x + \mu y) = \lambda Ax + \mu Ay = \lambda \sum_{i=1}^{n} \varphi_i(x) y_i + \mu \sum_{i=1}^{n} \varphi_i(y) y_i.$$

Daraus folgt

$$\sum_{i=1}^{n} (\varphi_i(\lambda x + \mu y) - \lambda \varphi_i(x) - \mu \varphi_i(y)) y_i = 0.$$

Da die $(y_i)_{i=1,\ldots,n}$ linear unabhängig sind, folgt schließlich

$$\varphi_i(\lambda x + \mu y) = \lambda \varphi_i(x) + \mu \varphi_i(y) \quad \text{für alle } i = 1, \ldots, n.$$

□

Wir zeigen nun sogar mit Hilfe der folgenden Lemmata, dass

$$\varphi_i \in X' \quad \text{für alle } i = 1, \dots, n.$$

Lemma 4.1. *Sei $\phi : X \to \mathbb{K}$ linear. Dann gilt*

$$\phi : X \to \mathbb{K} \text{ stetig} \Longleftrightarrow \ker(\phi) \text{ abgeschlossener UVR von } X.$$

Beweis. Zu „$\Longrightarrow$“: Trivial, da

$$\ker(\phi) = \phi^{-1}(\{0\})$$

und Urbilder abgeschlossener Mengen (hier: $\{0\}$) unter stetigen Funktionen abgeschlossen sind.

Zu „$\Longleftarrow$“: Angenommen, ϕ ist nicht stetig, d.h.

$$\sup_{x \in B_X} |\phi(x)| = +\infty \tag{4.1}$$

Dann gibt es für alle $M > 0$ ein $x \in B_X$, so dass

$$|\phi(x)| > M.$$

Sei $x_0 \in X$ und sei $\epsilon > 0$. Nach (4.1) existiert ein $x \in B_X$ mit

$$|\varphi(x)| > \tfrac{|\phi(x_0)|}{\epsilon}.$$

Somit gilt für $y = \epsilon x$

$$\|y\| \leq \epsilon \quad \text{und} \quad |\phi(y)| > |\phi(x_0)|. \tag{4.2}$$

Zunächst ist

$$x_0 - \frac{\phi(x_0)}{\phi(y)} \cdot y \in \ker(\phi),$$

denn

$$\phi(x_0 - \tfrac{\phi(x_0)}{\phi(y)} \cdot y) \overset{\phi \text{ linear}}{=} \phi(x_0) - \tfrac{\phi(x_0)}{\phi(y)} \cdot \phi(y) = 0.$$

Weiter gilt

$$\left\| x_0 - x_0 - \tfrac{\phi(x_0)}{\phi(y)} \cdot y \right\| = \left| \tfrac{\phi(x_0)}{\phi(y)} \right| \|y\| \overset{(4.2)}{<} \|y\| \overset{(4.2)}{\leq} \epsilon,$$

d.h.

$$x_0 - \frac{\phi(x_0)}{\phi(y)} \cdot y \in B(x_0, \epsilon).$$

Damit haben wir gezeigt, dass

$$x_0 - \tfrac{\phi(x_0)}{\phi(y)} \cdot y \in \ker(\phi) \cap B(x_0, \epsilon),$$

d.h. $\ker(\phi)$ ist dicht in X: $\overline{\ker(\phi)} = X$. Da aber nach Voraussetzung $\ker(\phi)$ abgeschlossen ist, folgt

$$\ker(\phi) = X \Longrightarrow \phi = 0.$$

Doch das ist ein Widerspruch zur Unstetigkeit von ϕ. □

Lemma 4.2. *Sei A ein abgeschlossener Untervektorraum von X und B ein endlich-dimensionaler linearer Teilraum von X. Dann ist auch $A + B$ abgeschlossen.*

Beweis. Wir führen den Beweis induktiv: Es reicht daher zu zeigen:

$$b \notin A \Longrightarrow A + \operatorname{lin}\{b\} \text{ ist abgeschlossen.}$$

Ist nämlich $b \in A$, so ist $A + \operatorname{lin}\{b\} = A$, da $A \subset X$ ein Untervektorraum ist. Also ist $A + \operatorname{lin}\{b\}$ trivialerweise abgeschlossen.

Betrachte jetzt die Folge $(a_n + \lambda_n b)_n \subset A + \operatorname{lin}\{b\}$ mit

$$\lim_{n\to\infty} a_n + \lambda_n b = y \quad \text{in } X.$$

Angenommen, es gibt eine Teilfolge mit

$$|\lambda_{n_k}| \to +\infty, \quad (k \to \infty).$$

Dann gilt

$$\tfrac{a_{n_k}+\lambda_{n_k} b}{\lambda_{n_k}} \overset{k\to\infty}{\longrightarrow} \tfrac{y}{\infty} = 0,$$

d.h.

$$-\tfrac{a_{n_k}}{\lambda_{n_k}} \overset{k\to\infty}{\longrightarrow} b.$$

Folglich ist $b \in \bar{A} = A$. Widerspruch!

Somit ist $(\lambda_n)_n$ beschränkt in $\mathbb{K}$. Nach dem Satz von Bolzano-Weierstraß (Satz 4.11) existiert eine Teilfolge $(\lambda_{n_k})_k$ mit

$$\lambda_{n_k} \overset{k\to\infty}{\longrightarrow} \lambda \quad \text{in } \mathbb{K}.$$

Damit ist aber

$$a_{n_k} = \underbrace{(a_{n_k} + \lambda_{n_k} b)}_{\to y} - \lambda_{n_k} b \overset{k\to\infty}{\longrightarrow} y - \lambda b \in \bar{A} = A.$$

Es folgt

$$y = (y - \lambda b) + \lambda b \in A + \operatorname{lin}\{b\}.$$

□

Die Stetigkeit von φ_i folgt nun mit Lemma 4.1 und 4.2 leicht aus

$$\underbrace{\ker(A)}_{\text{abgeschlossen, da } A \text{ stetig}} + \underbrace{\operatorname{lin}(\{\, x_j \mid j = 1, \dots, n, j \neq i \,\})}_{\text{endlich-dimensional}} = \ker(\varphi_i), \quad i = 1, \dots, n, \tag{4.3}$$

wobei $x_1, \dots, x_n \in X$ mit $Ax_i = y_i$ für alle $i = 1, \dots, n$.

Beweis von (4.3). Zu „$\subset$“: Sei also $a \in \ker(A) + \operatorname{lin}(\{\, x_j \mid j = 1, \dots, n, j \neq i \,\})$. Dann gibt es $x \in \ker(A)$ und $\tilde{x} \in \operatorname{lin}(\{\, x_j \mid j = 1, \dots, n, j \neq i \,\})$, so dass

$$a = x + \tilde{x}.$$

Wir müssen zeigen, dass

$$\varphi_i(a) \overset{\varphi_i \text{ linear}}{=} \varphi_i(x) + \varphi_i(\tilde{x}) = 0 \quad \text{für alle } i = 1, \dots, n.$$

a) Ist $x \in \ker(A)$, so folgt

$$0 = Ax = \sum_{i=1}^{n} \varphi_i(x) y_i$$

und da die $(y_i)_i$ linear unabhängig sind schließlich

$$\varphi_i(x) = 0 \quad \text{für alle } i = 1, \dots, n.$$

b) Zunächst gilt für alle $j = 1, \dots, n$

$$\underline{y_j} = Ax_j = \sum_{i=1}^{n} \varphi_i(x_j)\underline{y_i}.$$

Das geht aber nur, wenn

$$\varphi_i(x_j) = \delta_{ij}.$$

Ist nun $\tilde{x} \in \operatorname{lin}(\{\, x_j \mid j = 1, \dots, n, j \neq i \,\})$, so gibt es $\alpha_i \in \mathbb{K}$, so dass

$$\tilde{x} = \sum_{j \neq i} \alpha_j x_j.$$

Dann ist also für alle $i = 1, \dots, n$

$$\varphi_i(\tilde{x}) = \sum_{j \neq i}^{n} \alpha_j \underbrace{\varphi_i(x_j)}_{=\delta_{ij}} = 0.$$

Damit ist $\varphi_i(a) = 0$, also $a \in \ker(\varphi_i)$ für alle $i = 1, \dots, n$ gezeigt.

<u>Zu „$\supset$“:</u> Sei also $x \in \ker(\varphi_i)$, d.h. $\varphi_i(x) = 0$. Dann gilt

$$Ax = \sum_{j=1}^{n} \varphi_j(x) y_j = \sum_{j \neq i} \varphi_j(x) y_j + \underbrace{\varphi_i(x)}_{=0} y_i = \sum_{j \neq i} \varphi_j(x) y_j. \tag{4.4}$$

Sei nun

$$\tilde{x} = \sum_{j \neq i} \varphi_j(x) x_j \in \operatorname{lin}(\{\, x_j \mid j = 1, \dots, n, j \neq i \,\}).$$

Dann folgt

$$A\tilde{x} \overset{A \text{ linear}}{=} \sum_{j \neq i} \varphi_j(x) \underbrace{Ax_j}_{=y_j} = \sum_{j \neq i} \varphi_j(x) y_j \overset{(1.27)}{=} Ax,$$

also

$$A(x - \tilde{x}) = 0 \Longrightarrow x - \tilde{x} \in \ker(A)$$

und somit

$$x = (x - \tilde{x}) + \tilde{x} \in \ker(A) + \operatorname{lin}(\{\, x_j \mid j = 1, \dots, n, j \neq i \,\}).$$

□

Wir fassen nun unser Resultat nochmal als Satz zusammen:

Satz 4.1. *Jedes $A \in F(X, Y)$ ist von der Form*

$$Ax = \sum_{i=1}^{n} \varphi_i(x) y_i, \quad \varphi_i \in X',\ y_i \in Y,\ i = 1, \dots, n =: \dim(A(x)). \tag{4.5}$$

Umgekehrt ist klar, dass jeder Operator $A : X \to Y$ von der Form (4.5) in $F(X, Y)$ liegt.

WIr kommen jetzt zu einem neuen Basisbegriff:

Definition 4.1 (Schauderbasis). Eine Folge von Elementen $\{ y_n \mid n \in \mathbb{N} \} \subset Y$ in einem normierten Vektorraum $(Y, \|.\|_Y)$ heißt *Schauderbasis*, falls gilt

$$\forall_{y \in Y} \exists!_{(\alpha_n)_n \subset \mathbb{K}} : y = \sum_{n=1}^{\infty} \alpha_n y_n.$$

Die Summe ist dabei im folgenden (Konvergenz-)Sinne zu verstehen:

$$\lim_{N \to \infty} \|y - \sum_{n=1}^{N} \alpha_n y_n\| = 0.$$

Bemerkungen.

1. Aus der Eindeutigkeitsbedingung an die Koeffizienten folgt sofort, dass jede endliche Familie von Elementen einer Schauderbasis linear unabhängig in Y ist; insbesondere sind alle $y_n \neq 0$.

   Beweis. Sei etwa
   $$\lambda_1 y_1 + \ldots + \lambda_n y_n = 0.$$
   Angenommen, es gibt ein $i = 1, \ldots, n$ mit $\lambda_i \neq 0$. O.E. sei $i = 1$. Dann gilt
   $$y_1 = \frac{\lambda_2}{\lambda_1} y_2 + \ldots + \frac{\lambda_n}{\lambda_1} y_n. \tag{4.6}$$
   Nach Definition gilt
   $$y_1 = \sum_{i=1}^{\infty} \alpha_i y_i = 1 \cdot y_1, \quad \alpha_i = 0 \text{ für alle } i \geq 2.$$
   Nach Vergleich mit (4.6) folgt wegen der Eindeutigkeit der Koeffizienten $\lambda_2 = \ldots = \lambda_n = 0$, also $y_1 = 0$. Dann ist aber $\alpha_1 = 1$ nicht eindeutig. Widerspruch! □

2. Ist $\{ y_n \mid n \in \mathbb{N} \}$ eine Schauderbasis, so ist auch jede Familie der Form $\{ \alpha_n y_n \mid n \in \mathbb{N} \}$ mit
   $$(\alpha_n)_n \subset \mathbb{K} \setminus \{0\}$$
   eine Schauderbasis. Jede Schauderbasis kann daher „normalisiert“ werden:
   $$\{y_n\} \longrightarrow \{ \tfrac{y_n}{|y_n|} \mid n \in \mathbb{N} \}.$$
   Beachte, dass nach 1. tatsächlich alle $y_n \neq 0$ sind.

3. Besitzt Y eine Schauderbasis $\{ y_n \mid n \in \mathbb{N} \}$, so ist Y nach Satz 1.5 separabel, denn $\operatorname{lin}\{ y_n \mid n \in \mathbb{N} \}$ ist dicht in Y.

4. Schauderbasis $\neq$ Hamelbasis.

5. Jede Schauderbasis ist gleichzeitig auch eine Galerkinbasis[10], aber eine Galerkinbasis ist im Allgemeinen keine Schauderbasis (vergleichen Sie dazu das Beispiel 4.1 unten).

Beispiel 4.1. Sei $X = c_0$ (oder $X = l^1$). Betrachte

$$\begin{aligned} y_1 &:= \tfrac{1}{2} e^{(1)} = (\tfrac{1}{2}, 0, 0, \ldots), \\ y_n &:= \tfrac{1}{2} e^{(n)} - e^{(n-1)} = (0, \ldots, 0, -1, \tfrac{1}{2}, 0, \ldots). \end{aligned}$$

Dann gilt

[10] Eine Familie von linear unabhängigen Elementen $\{ y_n \mid n \in \mathbb{N} \} \subset Y$ heißt *Galerkinbasis*, falls
$$\operatorname{dist}(y, \operatorname{lin}(\{y_1, \ldots, y_n\})) \overset{n \to \infty}{\longrightarrow} 0 \quad \text{für alle } y \in Y.$$

- $\{ y_n \mid n \in \mathbb{N} \}$ ist linear unabhängig:

$$0 = \sum_{i=1}^{\infty} \lambda_i y_i = (\tfrac{1}{2}\lambda_1 - \lambda_2, \tfrac{1}{2}\lambda_2 - \lambda_3, \ldots, \tfrac{1}{2}\lambda_{n-1} - \lambda_n, \ldots).$$

  Es folgt

$$\lambda_2 = \tfrac{1}{2}\lambda_1, \ldots, \lambda_n = \tfrac{1}{2}\lambda_{n-1}, \ldots,$$

  also

$$\lambda_i = (\tfrac{1}{2})^{i-1}\lambda_1 \quad \text{für alle } i \in \mathbb{N} \tag{4.7}$$

  Betrachten eine beliebige (endliche) Teilmenge $\{y_{n_1}, \ldots, y_{n_N}\} \subset \{ y_n \mid n \in \mathbb{N} \}$. Dann ist der letzte von Null verschiedene Eintrag in der oberen Summe gerade $\frac{1}{2}\lambda_{n_N}$, also ist $\lambda_{n_N} = 0$ und wegen (4.7) folgt schließlich $\lambda_i = 0$ für alle anderen $i = n_1, \ldots, n_{N-1}$.

- $\operatorname{lin}\{y_1, \ldots, y_n\} = \operatorname{lin}\{e^{(1)}, \ldots, e^{(n)}\}$:

  Zum Beispiel ist

$$e^{(1)} = 2y_1, \quad e^{(n)} = 2y_n + 2^2 y_{n-1} + \ldots + 2^n y_1, \ n \geq 2.$$

  Dabei kann man den letzten Ausdruck leicht durch Induktion zeigen.

Das zeigt, dass $\{ y_n \mid n \in \mathbb{N} \}$ eine Gelerkinbasis in X ist. Nun gilt aber

$$\begin{aligned}\sum_{n=1}^{N} 2^{-n} y_n &= \tfrac{1}{4}e^{(1)} + \tfrac{1}{4}(\tfrac{1}{2}e^{(2)} - e^{(1)}) + \ldots + 2^{-N}(\tfrac{1}{2}e^{(n)} - e^{(N-1)}) \\ &= 2^{-N-1}e^{(N)}.\end{aligned}$$

Somit folgt

$$\sum_{n=1}^{\infty} 2^{-n} y_n = \lim_{N\to\infty} \sum_{n=1}^{N} 2^{-n} y_n = \lim_{N\to\infty} 2^{-N-1} e^{(N)} = 0,$$

also

$$0 = \sum_{n=1}^{\infty} 0 \cdot y_n = \sum_{n=1}^{\infty} 2^{-n} y_n$$

und damit hat man keine eindeutige Darstellung. Daher ist $\{ y_n \mid n \in \mathbb{N} \}$ keine Schauderbasis.

□

Beispiel 4.2. Sei $X = c_0$ oder $X = l^p$, $1 \leq p < \infty$. Dann ist $\{ e^{(n)} \mid n \in \mathbb{N} \}$ eine Schauderbasis für X.

□

Beispiel 4.3. $X = C[0,1]$ besitzt eine Schauderbasis. (Übung: Schaudersche Funktionensystem).

□

Wenn $(Y, \|.\|_Y)$ eine Schauderbasis $\{ y_n \mid n \in \mathbb{N} \}$ besitzt, dann definiert man für alle $n \in \mathbb{N}$ die zugehörigen *Koeffizientenfunktionale*

$$\phi_n : Y \to \mathbb{K}, \quad y \mapsto \lambda_n,$$

wobei λ_n der zugehörige n-te Koeffizient in

$$y = \sum_{n=1}^{\infty} \lambda_n y_n$$

ist.

Wegen der Eindeutigkeit der Darstellung gilt

a) ϕ_i ist linear für alle $i \in \mathbb{N}$.

b) Es gilt folgende Relation

$$\phi_i(y_j) = \delta_{ij} \quad \text{für alle } i, j \in \mathbb{N}.$$

Deswegen werden die Funktionale ϕ_i auch *biorthogonale Funktionale* genannt.

Weiter werden für alle $n \in \mathbb{N}$ folgende Operatoren definiert

$$P_n : Y \to Y, \quad y \mapsto \sum_{k=1}^{n} \lambda_k y_k,$$

wobei die Summe die n-te Partialsumme von

$$y = \sum_{n=1}^{\infty} \lambda_n y_n$$

darstellt.

Eigenschaften von P_n.

1. P_n ist linear für alle $n \in \mathbb{N}$.
2. P_n ist eine Projektion auf $\operatorname{lin}\{y_1, \ldots, y_n\}$, d.h.

$$P_n \circ P_n = P_n,$$

denn

$$P_n \circ P_n(y) = P_n(\sum_{k=1}^{n} \lambda_k y_k) \overset{P_n \text{ linear}}{=} \sum_{k=1}^{n} \lambda_k \underbrace{P_n(y_k)}_{=y_k} = P_n(y).$$

3. $\lim P_n y \to y$ in Y für $n \to \infty$ für alle $y \in Y$.

Damit bleibt nur noch die Frage nach der Stetigkeit von ϕ_n und P_n zu klären. Eine Antwort bietet der

Satz 4.2. *Sei $(Y, \|.\|_Y)$ ein <u>Banachraum</u> mit Schauderbasis $\{ y_n \mid n \in \mathbb{N} \}$ und zugehörige Koeffizientenfunktionale $(\phi_n)_n$ und Projektionen $(P_n)_n$. Dann gilt*

$$\phi_n \in X' \quad \text{und} \quad P_n \in L(Y) \quad \text{für alle } n \in \mathbb{N}$$

und es existiert ein $M \geq 1$, so dass

$$\|P_n\|_{L(Y)} \leq M \quad \text{für alle } n \in \mathbb{N}.$$

Bemerkung. Die Aussage des Satzes ist im Allgemeinen falsch, wenn Y nicht vollständig ist. Zum Beispiel ist $(Y, \|.\|_Y) = (d, \|.\|_\infty)$ kein Banachraum und $(b^{(n)})_n$ mit

$$b^{(1)} := e^{(1)}, \quad b^{(n)} := e^{(n)} - e^{(n-1)}, \ n \geq 2$$

eine Schauderbasis. Aber man kann zeigen, dass die ϕ_n und P_n nicht stetig sind. (Übung).

Um den Satz zu beweisen, brauchen wir noch ein Lemma aus der *Analysis 1*:

Lemma 4.3. *Sei $(a_{km})_{km}$ eine reelle Folge. Dann gilt*

$$\sup_k \lim_{m \to \infty} a_{km} \leq \limsup_{m \to \infty} \sup_k a_{km}.$$

Beweis. Selbst. □

Beweis von Satz 4.2. Wir definieren auf Y eine neue Norm wie folgt

$$||| \, y \, ||| := \sup_{k \in \mathbb{N}} \|P_k(y)\| \quad \text{für alle } y \in Y.$$

Da

$$\lim_{k \to \infty} P_k(y) = y \quad \text{in } (Y, \|.\|) \text{ für alle } y \in Y,$$

gilt offensichtlich

$$\|y\| \leq ||| \, y \, ||| \quad \text{für alle } y \in Y. \tag{4.8}$$

Zwischenbehauptung. $(Y, ||| \, . \, |||)$ ist vollständig.

Aus der Zwischenbehauptung und Korollar 2.3 folgt dann, dass es ein $M \geq 0$ gibt, so dass

$$||| \, y \, ||| \leq M\|y\| \quad \text{für alle } y \in Y,$$

d.h. aber gerade, dass

$$\|P_k y\| \leq M\|y\| \quad \text{für alle } k \in \mathbb{N} \text{ und alle } y \in Y.$$

Damit ist $P_k \in L(Y)$ für alle $k \in \mathbb{N}$ und

$$\|P_k\|_{L(Y)} \leq M \quad \text{für alle } k \in \mathbb{N}.$$

Bemerke, dass $M \geq 1$: Für $P_k = 0$ ist das sowieso klar und für $P_k \neq 0$ gilt

$$\|P_k\|_{L(Y)} = \left\|P_k^2\right\|_{L(Y)} \leq \|P_k\|_{L(Y)}^2 \Longrightarrow \|P_k\|_{L(Y)} \geq 1.$$

Weiter im Beweis: Weil

$$\phi_n(y) y_n = P_n(y) - P_{n-1}(y) \quad \text{für alle } n \in \mathbb{N} \text{ und } y \in Y,$$

folgt

$$|\phi_n(y)| \, \|y_n\|_Y = \|\phi_n(y) y_n\|_Y \leq (\|P_n\|_{L(Y)} + \|P_{n-1}\|_{L(Y)})\|y\|_Y \leq 2M\|y\|_Y.$$

Somit ist

$$|\phi_n(y)| \leq \tfrac{2M}{\|y_n\|_Y}\|y\|_Y \quad \text{für alle } n \in \mathbb{N} \text{ und } y \in Y,$$

d.h.

$$\phi_n \in Y' \quad \text{und} \quad \|\phi_n\|_{Y'} \leq \tfrac{2M}{\|y_n\|_Y} \quad \text{für alle } n \in \mathbb{N}.$$

Beweis der Zwischenbehauptung. Sei $(x_n)_n$ eine Cauchy-Folge in $(Y, ||| \, . \, |||)$. Dann ist $(x_n)_n$ wegen (4.8) auch eine Cauchy-Folge in $(Y, \|.\|)$. Weil $(Y, \|.\|)$ ein Banachraum ist, gilt also

$$x_n \to x \quad \text{in } (Y, \|.\|).$$

Nach Definition der $||| \, . \, |||$-Norm ist auch für jedes feste k die Folge $(P_k(x_n))_n$ eine Cauchyfolge in $(Y, \|.\|)$. Also gilt

$$P_k(x_n) \overset{n \to \infty}{\longrightarrow} z_k \quad \text{in } (Y, \|.\|) \text{ für alle } k \in \mathbb{N}.$$

Wir zeigen jetzt, dass

$$z_k \overset{k \to \infty}{\longrightarrow} x \quad \text{in } (Y, \|.\|).$$

Sei dazu $\epsilon > 0$. Da $(x_n)_n$ eine Cauchy-Folge in $(Y, \|.\|)$ ist, existiert ein $n_0 \in \mathbb{N}$ mit

$$\|x_{n_0} - x_n\|_Y \leq \tfrac{\epsilon}{3} \quad \text{für alle } n \geq n_0. \tag{4.9}$$

Da $P_k y \overset{k \to \infty}{\longrightarrow} y$ für alle $y \in Y$, gibt es ein $k_0 \in \mathbb{N}$ mit

$$\|P_k x_{n_0} - x_{n_0}\|_Y \leq \tfrac{\epsilon}{3M} \quad \text{für alle } k \geq k_0. \tag{4.10}$$

Es folgt dann

$$\begin{aligned}
\|z_k - x\|_Y &= \lim_{n\to\infty} \|P_k(x_n) - x_n\|_Y \\
&\le \lim_{n\to\infty} (\|P_k(x_n) - P_k(x_{n_0})\|_Y + \underbrace{\|P_k(x_{n_0}) - x_{n_0}\|_Y}_{\overset{(4.10)}{\le} \epsilon/3} + \underbrace{\|x_{n_0} - x_n\|_Y}_{\overset{(4.11)}{\le} \epsilon/3}) \\
&\le \lim_{n\to\infty} \underbrace{\|P_k\|_{L(Y)}}_{\le M} \underbrace{\|x_n - x_{n_0}\|_Y}_{\overset{(4.10)}{\le} \epsilon/(3M)} + \tfrac{2}{3}\epsilon \le \epsilon.
\end{aligned}$$

Wir wissen jetzt also, dass

$$z_k \to x \quad \text{in } (Y, \|.\|).$$

Da ϕ_j auf $\operatorname{lin}\{y_1, \dots, y_k\}$ stetig ist (denn lineare Funktionale auf endlich-dimensionalen Vektorräumen sind stets stetig) und $P_k(x_n) \in \operatorname{lin}\{y_1, \dots, y_n\}$, gilt für alle $1 \le j \le k$

$$\phi_j(x_n) = \phi_j(P_k(x_n)) \overset{n\to\infty}{\longrightarrow} \phi_j(z_k) =: \alpha_j.$$

Außerdem ist für alle $j = 1, \dots, n$ und für alle $k \in \mathbb{N}$

$$\phi_j(z_{k+1}) = \lim_{n\to\infty} \phi_j(P_{k+1}(x_n)) = \lim_{n\to\infty} \phi_j(P_k(x_n)) = \phi_j(z_k) = \alpha_j,$$

d.h.

$$z_k = \sum_{j=1}^{k} \alpha_j y_j \quad \text{für alle } k \in \mathbb{N}.$$

Da $z_k \to x$, folgt weiter

$$x = \sum_{j=1}^{\infty} \alpha_j y_j.$$

Nach Voraussetzung ist $\{\, y_n \mid n \in \mathbb{N} \,\}$ eine Schauderbasis. Daher ist die Darstellung von x eindeutig. Somit folgt

$$P_k(x) = \sum_{j=1}^{k} \alpha_j y_j = z_k \quad \text{für alle } k \in \mathbb{N}.$$

Somit ist

$$\begin{aligned}
\|x_n - x\| &= \sup_k \|P_k(x_n) - P_k(x)\|_Y = \sup_k \lim_{m\to\infty} \|P_k(x_n) - P_k(x_m)\|_Y \\
&\overset{\text{Lemma 4.3}}{\le} \limsup_{m\to\infty} \sup_k \|P_k(x_n - x_m)\|_Y = \limsup_{m\to\infty} \|x_n - x_m\|
\end{aligned}$$

und da $(x_n)_x$ eine Cauchyfolge in $(Y, |||\,.\,|||)$ ist, folgt schließlich

$$x_n \to x \quad \text{in } (Y, |||\,.\,|||).$$

□

Mit Satz 4.2 folgt nun leicht der

Satz 4.3. *Sei $(Y, \|.\|_Y)$ ein Banachraum mit Schauderbasis und $(X, \|.\|_X)$ ein beliebiger Banachraum. Dann gilt*

$$\overline{F(X,Y)} = K(X,Y).$$

Insbesondere gilt der Satz für die separablen Banachräume c_0, l^p, $C[0,1]$ und $L^p[0,1]$, wobei $(1 \le p < \infty)$.

Beweis. Nach Satz 4.2 sind die zur Schauderbasis $\{ y_n \mid n \in \mathbb{N} \}$ gehörigen Projektionen P_n stetig und es gibt ein $M \geq 1$, so dass

$$\|P_n\|_{L(Y)} \leq M \quad \text{für alle } n \in \mathbb{N} \tag{4.11}$$

Sei nun $A \in K(X,Y)$. Dann ist wegen (3.26)

$$P_n \circ A \in F(X,Y) \quad \text{für alle } n \in \mathbb{N}.$$

Weil A kompakt ist, ist AB_X präkompakt. Sei $\epsilon > 0$. Dann existieren wegen der Präkompaktheit $y_1, \ldots, y_k \in Y$ mit

$$AB_X \subset \bigcup_{i=1}^{k} B(y_i, \tfrac{\epsilon}{3M}).$$

Weiter gibt es wegen $P_n y_i \to y_i$ ein $n_0 \in \mathbb{N}$, so dass

$$\|P_n y_i - y_i\|_Y \leq \epsilon/3 \quad \text{für alle } n \geq n_0 \text{ und } i = 1, \ldots, k.$$

Wenn $y \in AB_x$, dann existiert ein $i_0 \in \{1, \ldots, k\}$, so dass

$$\|y - y_i\|_Y \leq \tfrac{\epsilon}{3m}.$$

Es folgt

$$\begin{aligned} \|P_n y - y\|_Y &\leq \|P_n y - P_n y_{i_0}\|_Y + \|P_n y_{i_0} - y_{i_0}\|_Y + \|y_{i_0} - y\|_Y \\ &\leq \|P_n\|_{L(Y)} \|y - y_i\|_Y + \tfrac{\epsilon}{3} + \tfrac{\epsilon}{3M} \leq M \tfrac{\epsilon}{3M} + \tfrac{2\epsilon}{3} = \epsilon. \end{aligned}$$

Wir haben damit gezeigt, dass

$$P_n \to I \quad \text{auf } AB_X \text{ gleichmäßig.}$$

Nun folgt schließlich

$$\|P_n \circ A - A\|_{L(X,Y)} = \sup_{x \in B_X} \|P_n(Ax) - Ax\|_Y = \sup_{y \in AB_X} \|P_n y - y\|_Y \overset{n\to\infty}{\longrightarrow} 0,$$

also ist $\overline{F(X,Y)} = K(X,Y)$. □

Bemerkung. Erstaunlicherweise kann man zeigen, dass $\overline{F(X,Y)} = K(X;Y)$ auch für die nicht-separablen(!) Banachräume $Y = l^\infty$ und $Y = L^\infty$ gilt, obwohl diese keine Schauderbasis besitzen.

Besitzt der Raum Y keine Schauderbasis, aber ist noch separabel, so kann man den Satz durch eine kleine Zusatzvoraussetzung retten:

Satz 4.4. *Sei X ein beliebiger Banachraum und Y ein (separabler) Banachraum mit der Eigenschaft:*

Es gibt eine Folge (S_n) in $F(Y)$ mit

$$\lim_{n\to\infty} S_n y = y \quad \textit{für alle } y \in Y. \tag{4.12}$$

Dann gilt $\overline{F(X,Y)} = K(X,Y)$.

Beweis. Sei also $T \in K(X,Y)$. Dann ist $S_n T \in F(X,Y)$ und es genügt zu zeigen, dass

$$\|S_n T - T\| \to 0.$$

Sei $\epsilon > 0$. Nach dem Satz von Banach-Steinhaus ist

$$K := \sup \|S_n\| < \infty.$$

Nun ist T kompakt und daher gibt es endlich viele $y_1, \ldots, y_r$ mit

$$\overline{T(B_X)} \subset \bigcup_{i=1}^{r} \{\, y \in Y \mid \|y - y_i\| < \epsilon \,\}.$$

Wegen (4.12) gibt es ein $N \in \mathbb{N}$, so dass

$$\|S_n y_i - y_i\| \leq \epsilon \quad \text{für alle } n \geq N,\ i = 1, \ldots, r.$$

Wähle jetzt für $x \in B_X$ ein $j \in \{1, \ldots, r\}$ mit $\|Tx - y_j\| < \epsilon$. Für alle $n \geq N$ folgt dann

$$\begin{aligned} \|S_n Tx - Tx\| &\leq \|S_n(Tx - y_j)\| + \|S_n y_j - y_j\| + \|y_j - Tx\| \\ &\leq \|S_n\| \|Tx - y_j + \epsilon + \epsilon\| \leq K\epsilon + 2\epsilon = (K+2)\epsilon, \end{aligned}$$

was zu zeigen war. □

Bemerkung.

1. Offensichtlich ist Y wegen der Eigenschaft (4.12) tatsächlich separabel.
2. Die Eigenschaft (4.12) ist schwächer als $\|S_n - I\| \to 0$, denn sonst würde hieraus $\dim Y < \infty$ folgen.

4.2 Spektrum und Resolvente beschränkter, linearer Operatoren

Sei $(X, \|.\|_X)$ ein Banachraum und sei $A \in L(X)$. Dann heißt

$$\boxed{\rho(A) := \{\, \lambda \in \mathbb{K} \mid (\lambda I - A) \text{ ist invertierbar in } L(X) \,\}} \tag{4.13}$$

die *Resolventenmenge*[11] *von* A.

Für $\lambda \in \rho(A)$ heißt

$$\boxed{R(\lambda, A) := (\lambda I - A)^{-1}} \tag{4.14}$$

die *Resolvente von* A *in* λ.

Schließlich nennt man

$$\boxed{\sigma(A) := \mathbb{K} \setminus \rho(A)} \tag{4.15}$$

das *Spektrum von* A.

Satz 4.5. *Seien* $\lambda_0 \in \rho(A)$ *und* $\lambda \in \mathbb{K}$ *mit*

$$|\lambda - \lambda_0| < \|R(\lambda_0, A)\|^{-1}.$$

Dann ist $\lambda \in \rho(A)$ *und*

$$R(\lambda, A) = \sum_{n=0}^{\infty} (\lambda_0 - \lambda)^n R(\lambda_0, A)^{n+1}.$$

Beweis. Schreibe

$$\lambda I - A = (\lambda_0 I - A) + (\lambda - \lambda_0) I = [I - (\lambda_0 - \lambda) R(\lambda_0, A)](\lambda_0 I - A).$$

Aus der Voraussetzung folgt

$$\|(\lambda_0 - \lambda) R(\lambda_0, A)\| = |\lambda_0 - \lambda| \, \|R(\lambda_0, A)\| < 1.$$

Somit folgt aus dem Satz zur Neumannschen Reihe (Satz 2.4), dass

$$I - (\lambda_0 - \lambda) R(\lambda_0, A)$$

invertierbar ist und

$$(I - (\lambda_0 - \lambda) R(\lambda_0, A))^{-1} = \sum_{n=0}^{\infty} (\lambda_0 - \lambda)^n R(\lambda_0, A)^n.$$

Als Komposition invertierbarer Operatoren ist $\lambda I - A$ invertierbar und

$$(\lambda I - A)^{-1} = R(\lambda_0, A) \cdot \sum_{n=0}^{\infty} (\lambda_0 - \lambda)^n R(\lambda_0, A)^n.$$

Also ist

$$R(\lambda, A) = \sum_{n=0}^{\infty} (\lambda_0 - \lambda)^n R(\lambda_0, A)^{n+1}.$$

□

[11] Wir erinnern uns, dass nach Korollar 2.2 gilt:

$$(\lambda I - A) \text{ invertierbar in } L(X) \iff \lambda I - A \text{ bijektiv}.$$

Dementsprechend kann man also in (4.13) „invertierbar“ durch „bijektiv“ ersetzen.

Korollar 4.1. *Sei $A \in L(X)$. Dann ist $\rho(A)$ offen und $\sigma(A)$ abgeschlossen.*

Beweis. Wir zeigen zunächst die Offenheit: Nach dem letzten Satz gilt

$$|\lambda - \lambda_0| < \|R(\lambda_0, A)\|^{-1} \Longleftrightarrow \lambda \in B(\lambda_0, \|R(\lambda_0, A)\|^{-1}) \Longrightarrow \lambda \in \rho(A).$$

Also ist in der Tat

$$B(\lambda_0, \left\|R(\lambda_0, A)^{-1}\right\|) \subset \rho(A)$$

und daher $\rho(A)$ offen.

Die Abgeschlossenheit von $\sigma(A) = \mathbb{K} \setminus \rho(A)$ ist klar. □

Satz 4.6. *Sei $A \in L(X)$. Dann gilt*

(i) *$\sigma(A)$ ist kompakt.*

(ii) *Für alle $\lambda \in \sigma(A)$ gilt $|\lambda| \leq \|A\|_{L(X)}$.*

(iii) *Falls $\mathbb{K} = \mathbb{C}$, dann ist $\sigma(A) \neq \emptyset$.*

Beweis. Zu (i) und (ii): Wir wolllen den Satz von Heine-Borel (Satz 1.9) anwenden:

$$\sigma(A) \text{ abgeschlossen und beschränkt} \Longrightarrow \sigma(A) \text{ kompakt.}$$

Im Korollar 4.1 haben wir schon gesehen, dass $\sigma(A)$ abgeschlossen ist. Um die Beschränktheit, d.h. (ii) zu zeigen, sei also $\lambda \in \mathbb{K}$ mit

$$|\lambda| > \|A\|.$$

Dann ist

$$\lambda I - A = \lambda(I - \tfrac{1}{\lambda}A)$$

und

$$\left\|\tfrac{A}{\lambda}\right\| = \frac{1}{|\lambda|}\|A\| < 1.$$

Also gilt wegen dem Satz über die Neumannsche Reihe (Satz 2.4)

$$(\lambda I - A)^{-1} = \lambda^{-1}(I - \tfrac{1}{\lambda}A)^{-1} = \lambda^{-1}\sum_{n=0}^{\infty}\lambda^{-n}A^n = \sum_{n=0}^{\infty}\lambda^{-n-1}A^n, \tag{4.16}$$

d.h.

$$\lambda \in \rho(A) \Longleftrightarrow \lambda \notin \sigma(A).$$

Somit folgt (ii) und mit Heine-Borel dann (i).

Zu (iii): Angenommen, $\sigma(A) = \emptyset$, d.h. $\rho(A) = \mathbb{C}$. Dann ist für alle $\varphi \in (L(X))'$ die Abbildung

$$\lambda \in \mathbb{C} \mapsto \varphi \circ R(\lambda, A) \in \mathbb{C}$$

eine ganze Funktion, d.h. eine auf ganz $\mathbb{C}$ holomorphe Funktion:

In jedem Punkt $\lambda_0 \in \mathbb{C}$ ist nämlich $\varphi \circ R(\lambda, A)$ in eine Potenzreihe entwickelbar, denn mit Satz 4.5 ist

$$\begin{aligned}\varphi \circ R(\lambda, A) =& \varphi \circ \sum_{n=0}^{\infty}(-1)^n(\lambda - \lambda_0)^n R(\lambda_0, A)^{n+1} \\ =& \sum_{n=0}^{\infty}(-1)^n(\lambda - \lambda_0)^n \underbrace{\varphi \circ R(\lambda_0, A)^{n+1}}_{\in\mathbb{C}}.\end{aligned}$$

Damit ist $\varphi \circ R(\lambda, A)$ holomorph auf ganz $\mathbb{C}$, also ganz.

Wir wollen jetzt die Annahme zum Widerspruch führen: Weil

$$\begin{aligned} |\varphi \circ R(\lambda, A)| &\overset{(4.16)}{\leq} \frac{\|\varphi\|}{|\lambda|} \sum_{n=0}^{\infty} \left\| \tfrac{A}{\lambda} \right\|^n = \frac{\|\varphi\|}{|\lambda|} \cdot \frac{1}{1 - \frac{\|A\|}{|\lambda|}} \\ &= \frac{\|\varphi\|}{|\lambda| - \|A\|} \end{aligned} \tag{4.17}$$

für alle $|\lambda| > \|A\|$, folgt, dass die Abbildung

$$\lambda \in \mathbb{C} \mapsto \varphi \circ R(\lambda, A)$$

beschränkt ist. Mit dem Satz von Liouville

Jede beschränkte ganze Funktion ist konstant.

folgt dann, dass

$$\varphi \circ R(., A) = const \text{ auf } \mathbb{C}.$$

Wegen (4.17) gilt aber

$$\lim_{|\lambda| \to \infty} |\varphi \circ R(\lambda, A)| = 0,$$

d.h.

$$\varphi \circ R(\lambda, A) = 0 \quad \text{für alle } \lambda \in \mathbb{C} \text{ und alle } \varphi \in L(X)'. \tag{4.18}$$

Dann folgt aber

$$R(\lambda, A) = 0, \tag{4.19}$$

und das ist ein Widerspruch zur Invertierbarkeit.

Nachweis von (4.19): Angenommen, $R(\lambda, A) \neq 0$. Dann gibt es nach Korollar 3.1 ein $\varphi \in L(X)'$, so dass

$$\varphi \circ R(\lambda, A) = \|R(\lambda, A)\|.$$

Also nach (4.18)

$$\|R(\lambda, A)\| = 0 \Longrightarrow R(\lambda, A) = 0.$$

Widerspruch! □

Als nützlich wird sich der so genannte Spektralradius erweisen, den wir jetzt definieren: Sei $A \in L(X)$ und X ein Banachraum. Dann heißt

$$\boxed{r(A) := \sup\{\, |\lambda| \mid \lambda \in \sigma(A) \,\}} \tag{4.20}$$

der *Spektralradius von A*.

Nach Satz 4.6(ii) gilt

$$r(A) \leq \|A\|.$$

Satz 4.7. *Sei* $\mathbb{K} = \mathbb{C}$. *Dann gilt:*

$$\exists_{\lambda \in \sigma(A)} : r(A) = |\lambda|\,.$$

Beweis. Aus Satz 4.6 folgt

$$\sigma(A) \neq \emptyset \quad \text{und} \quad \sigma(A) \text{ kompakt.}$$

Also nimmt die stetige Funktion $|.|$ auf $\sigma(A)$ das Maximum an. □

Satz 4.8. *Sei wieder $\mathbb{K} = \mathbb{C}$. Dann gilt*

$$r(A) = \lim_{n\to\infty} \|A^n\|^{1/n} = \inf_n \|A^n\|^{1/n}.$$

Beweis. 1. Schritt: Verallgemeinerung des Satzes zur Neumannschen Reihe. Angenommen, es existiert ein $n_0 \in \mathbb{N}$ mit

$$\|A^{n_0}\| < 1.$$

Dann folgt

$$\begin{aligned}\sum_{m=0}^{\infty} \|A^m\| &= (\|A^0\| + \|A^1\| + \ldots + \|A^{n_0-1}\|) + (\|A^{n_0}\| + \ldots + \|A^{2n_0-1}\|) \\ &\quad + \ldots + (\|A^{m_0\cdot n_0}\| + \ldots + \|A^{m_0\cdot n_0+n_0-1}\|) + \ldots \\ &= \sum_{m=0}^{\infty}\sum_{p=0}^{n_0-1} \|A^{m\cdot n_0+p}\| \le \sum_{m=0}^{\infty}\sum_{p=0}^{n_0-1} \|A^{n_0}\|^m \|A\|^p \\ &= \underbrace{\sum_{p=0}^{n_0-1} \|A\|^p}_{=:C<\infty} \cdot \sum_{m=0}^{\infty} \|A^{n_0}\|^m < \infty,\end{aligned}$$

weil $\|A^{n_0}\| < 1$. Somit konvergiert die Reihe $\sum_{m=0}^{\infty} A^m$ in $L(X)$.

Jetzt folgt genauso wie im Beweis des Satzes zur Neumannschen Reihe (Satz 2.4), dass $I - A$ invertierbar ist, d.h.

$$1 \in \sigma(A) \quad \text{und} \quad (I-A)^{-1} = \sum_{n=0}^{\infty} A^n.$$

2. Schritt: Sei nun $n \in \mathbb{N}$ und $\lambda \in \mathbb{C}$ mit

$$|\lambda| > \|A^n\|^{1/n}, \tag{4.21}$$

d.h. es gilt

$$\|(\tfrac{1}{\lambda}A)^n\| < 1.$$

Also folgt nach dem 1. Schritt

$$1 \in \rho(\tfrac{1}{\lambda}A) \iff I - \tfrac{1}{\lambda}A \text{ invertierbar} \iff \lambda \in \rho(A).$$

Es folgt demnach mit (4.21) wegen $\lambda \notin \sigma(A)$, dass

$$r(A) \le \inf_n \|A^n\|^{1/n} \le \liminf_n \|A^n\|^{1/n} \le \limsup_n \|A^n\|^{1/n} \overset{(*)}{\le} r(A).$$

Bevor wir nun $(*)$ beweisen, möchte ich noch ein wichtiges Resulat aus der komplexen Analysis rekapitulieren (vergleichen Sie dazu auch den Anhang 7.3):

3. Schritt: In der komplexen Analysis betrachtet man so genannte *Laurantreihen*:

$$\sum_{n=0}^{\infty} a_n\lambda^n + \sum_{n=1}^{\infty} a_{-n}\lambda^{-n}, \quad a_n, a_{-n} \in \mathbb{C}.$$

Die erste Summe wird *Nebenteil* und die zweite *Hauptteil* genannt.

Besitzt die auf dem Kreisring

$$\{\, \lambda \in \mathbb{C} \mid |\lambda| > r \,\}$$

holomorphe Funktion auf dem kleineren Kreisring

$$\{\, \lambda \in \mathbb{C} \mid |\lambda| > R \,\}, \quad (R > r)$$

die Laurententwicklung

$$\sum_{n=-\infty}^{\infty} a_n \lambda^n,$$

dann gilt diese Darstellung auch auf dem größeren Kreisring (und damit auch für den Außenbereich, wenn man solche Ringe zulässt).

<u>4. Schritt:</u> Zeige nun $(*)$: Sei $\rho > r(A)$. Wir hatten gesehen (vgl. (4.16)):

Für $|\lambda| > \|A\|$ gilt

$$R(\lambda, A) = \sum_{n=0}^{\infty} \lambda^{-n-1} A^n. \tag{4.22}$$

Da (4.22) konvergent ist, gilt dann für alle $\varphi \in L(X)'$

$$\varphi \circ R(\lambda, A) = \sum_{n=0}^{\infty} \lambda^{-n-1} \varphi \circ A^n \quad \text{für alle } |\lambda| > \|A\|.$$

Auf der anderen Seite ist $\varphi \circ R(., A)$ für $\varphi \in L(X)'$ eine holomorphe Funktion auf der Resolventenmenge $\rho(A)$, d.h. speziell auf

$$\{\, \lambda \in \mathbb{C} \mid |\lambda| > r(A) \,\}.$$

Wir wenden nun den 3. Schritt an auf $\lambda = \rho$: Dann ist

$$\varphi \circ R(\rho, A) = \sum_{n=0}^{\infty} \rho^{-n-1} \varphi \circ A^n = \sum_{n=0}^{\infty} \varphi(\rho^{-n-1} A^n)$$

und die Reihe konvergiert absolut. Somit ist die Folge

$$(\varphi(\rho^{-n-1} A^n))_n$$

beschränkt in $\mathbb{C}$ für alle $\varphi \in L(X)'$. Damit ist die Folge $(\rho^{-n-1} A^n)_n$ schwach beschränkt in $L(X)$. Nach Korollar 3.1 gibt es ein $\varphi \in L(X)'$ mit

$$\varphi(\rho^{-n-1} A^n) = \left\|\rho^{-n-1} A^n\right\|.$$

Damit ist die Folge $(\rho^{-n-1} A^n)_n$ auch normbeschränkt in $L(X)$, d.h. es gibt ein $C > 0$ mit

$$\|A^n\| < C \rho^{n+1}.$$

Daraus folgt

$$\|A^n\|^{1/n} \leq \sqrt[n]{C} \rho \sqrt[n]{\rho} \quad \text{für alle } n \in \mathbb{N}.$$

Es ergibt sich

$$\limsup_{n \to \infty} \|A^n\|^{1/n} \leq \rho \leq r(A),$$

d.h. $(*)$ ist bewiesen. □

Satz 4.9 (**Resolventengleichung**). *Seien $A \in L(X)$ und $\lambda, \mu \in \rho(A)$. Dann gilt*

$$R(\lambda, A) - R(\mu, A) = (\mu - \lambda) R(\lambda, A) R(\mu, A).$$

Beweis. Einfaches Nachrechnen ergibt

$$\begin{aligned} R(\lambda, A) - R(\mu, A) &= R(\lambda, A)(I - (\lambda I - A) R(\mu, A)) \\ &= R(\lambda, A)((\mu I - A) - (\lambda I - A)) R(\mu, A) = (\mu - \lambda) R(\lambda, A) R(\mu, A). \end{aligned}$$

□

Bestandteile des Spektrum.

1. Die Menge

$$\boxed{\sigma_p(A) := \{\, \lambda \in \mathbb{C} \mid \lambda I - A \text{ nicht injektiv} \,\}} \tag{4.23}$$

   heißt das *Punktspektrum von A*. Der Buchstabe „p“ steht für „point“.

   Ein Element $\lambda \in \sigma_p(A)$ heißt *Eigenwert von A*.

   Ein Element $x \in X \setminus \{0\}$ mit $Ax = \lambda x$ heißt *Eigenvektor von A zum Eigenwert* λ.

   $\ker(\lambda I - A)$ heißt der *Eigenraum von A zum Eigenwert* λ.

   $\dim(\ker(\lambda I - A))$ heißt die *Vielfachheit des Eigenwerts* λ.

2. Die Menge

$$\boxed{\sigma_c(A) := \{\lambda \in \mathbb{C} | \begin{array}{c} \lambda I - A \text{ injektiv, aber nicht surjektiv} \\ \text{und } \overline{R(\lambda I - A)} = X \end{array} \}} \tag{4.24}$$

   heißt das *kontinuierliche Spektrum von A*. Das „c“ steht für „continous“.

3. Die Menge

$$\boxed{\sigma_r(A) := \{\, \lambda \in \mathbb{C} \mid \lambda I - A \text{ injektiv und } \overline{R(\lambda I - A)} \subsetneq X \,\}} \tag{4.25}$$

   heißt das *Restspektrum von A*. Das „r“ steht für „rest“.

Man kann sich leicht überzeugen, dass

$$\sigma(A) = \sigma_p(A) \dot\cup \sigma_c(A) \dot\cup \sigma_r(A).$$

Beispiel 4.4. Sei $\dim(X) < \infty$. Dann wird $A \in L(X)$ nach linearer Algebra durch eine Matrix $[A]$ dargestellt. Es ist

$$\sigma(A) = \{\text{Eigenwerte von [A]}\}.$$

□

Beispiel 4.5. Seien $X = l^2$, $\mathbb{K} = \mathbb{C}$ und $A \in L(l^2)$ der Rechtsshift-Operator

$$A((x_n)_n) = (0, x_1, x_2, x_3, \ldots), \quad x = (x_n)_n \in l^2.$$

- $\underline{\lambda \in \mathbb{C}}$: Die Gleichung $Ax = \lambda x$ für $x \in l^2$ ist äquivalent zu

$$(0, x_1, x_2, \ldots) = (\lambda x_1, \lambda x_2, \lambda x_3, \ldots) \tag{4.26}$$

  Das heißt aber gerade $(x_n) = 0 = (0, 0, 0, \ldots)$, also $\sigma_p(A) = \emptyset$, denn ist $\lambda \neq 0$, dann folgt aus (4.26)

$$\lambda x_1 = 0 \Longrightarrow x_1 = 0 \Longrightarrow x_i = 0 \quad \text{für alle } i.$$

  und für $\lambda = 0$ ist das sowieso klar.

- $\underline{\lambda = 0}$: Dann ist

$$R(0I - A) = R(-A) = R(A) = \{\, (y_n)_n \in l^2 \mid y_1 = 0 \,\}$$

  und somit

$$\overline{R(A)} \subsetneq l^2,$$

  d.h.

$$0 \in \sigma_r(A).$$

- Wie sieht $r(A)$ aus? Nach n-mal Shiften hat man

$$\|A^n x\|_{l^2} = \left(\sum_{k=1}^{\infty} |x_{k+n}|^2 \right)^{1/2} \leq \|x\|_{l^2}.$$

Damit folgt

$$\|A^n\| = \sup_{x \in B_{l^2}} \|A^n x\| \leq 1 \quad \text{für alle } n \in \mathbb{N}.$$

Wegen $\|A^n e_{n+1}\|_{l^2} = 1$ und $\|e_{n+1}\|_{l^2} = 1$ folgt schließlich

$$\|A^n\| = 1 \quad \text{für alle } n \in \mathbb{N}.$$

Daher ist

$$\|A^n\|^{1/n} = 1 \quad \text{für alle } n \in \mathbb{N}.$$

Daraus folgt mit Satz 4.8

$$r(A) = 1.$$

□

Versucht man nun das Bild

$$R(\lambda I - A) = \{ (\lambda x_1, \lambda x_2 - \lambda x_1, \lambda x_3 - \lambda x_2, \ldots) \mid (x_n)_n \in l^2 \}$$

aus dem letzten Beispiel zu untersuchen, so scheint das problematisch. Abhilfe schafft gelegentlich das folgende

Lemma 4.4. *Sei $A \in L(X)$, also $A' \in L(X')$. Dann gilt*

$$\lambda \in \sigma_p(A') \Longleftrightarrow \overline{R(\lambda I - A)} \subsetneq X.$$

Speziell folgt

$$\sigma_r(A) \subset \sigma_p(A') \subset \sigma_p(A) \dot{\cup} \sigma_r(A).$$

Beweis. Zu „$\Longleftarrow$“: Sei also $\overline{R(\lambda I - A)} \neq X$. Da $\overline{R(\lambda I - A)}$ ein echter abgeschlossener Untervektorraum von X ist, gibt es nach Korollar 3.4 ein $\varphi \in X' \setminus \{0\}$ mit

$$\varphi(\lambda x - Ax) = 0 \quad \text{für alle } x \in X.$$

Daraus folgt für alle $x \in X$

$$\varphi(Ax) = \lambda \varphi(x) \Longrightarrow (A'\varphi)(x) = \lambda \varphi(x),$$

also

$$A'\varphi = \lambda \varphi.$$

Somit ist

$$\lambda \in \sigma_p(A').$$

Zu „$\Longrightarrow$“: Angenommen, $\overline{R(\lambda I - A)} = X$ für $\lambda \in \sigma_p(A')$. Dann existiert ein Eigenvektor $\varphi \in X' \setminus \{0\}$ zum Eigenwert λ von A, d.h. für alle $x \in X$ gilt

$$(A'\varphi)(x) = \lambda \varphi(x) \Longrightarrow \varphi(Ax) = \lambda \varphi(x).$$

Es folgt

$$\varphi(\lambda x - Ax) = 0 \quad \text{für alle } x \in X,$$

also

$$\varphi|_{R(\lambda I - A)} \equiv 0.$$

Da aber $\overline{R(\lambda I - A)} = X$, folgt $\varphi \equiv 0$. Widerspruch! □

Beispiel 4.6 (**Fortsetzung von Beispiel 4.5**). In Beispiel 3.15 haben wir gesehen, dass $A' \in L(X') \cong L(l^2)$ gegeben ist durch den Linksshift

$$A'(x_n)_n = (x_2, x_3, x_4, \ldots).$$

Bestimmung von $\sigma_p(A')$: Sei $\lambda \in \mathbb{C}$. Wir werden sehen (vgl. (4.27)), dass $|\lambda| < 1$.

Es gilt

$$A'((x_n)_n) = \lambda(x_n)_n \Longleftrightarrow (x_2, x_3, \ldots) = (\lambda x_1, \lambda x_2, \lambda x_3, \ldots).$$

Wir können $\lambda = 0$ wegen $(x_2, x_3, \ldots) \neq 0$ ausschließen. Sei nun x_1 beliebig. Dann folgt

$$\begin{array}{rcl} x_2 & = & \lambda x_1 \\ x_3 & = & \lambda x_2 = \lambda^2 x_1 \\ \vdots & & \\ x_n & = & \lambda^{n-1} x_1 \quad \text{für alle } n \in \mathbb{N}. \end{array}$$

Also ist

$$(x_n)_n = x_1(\lambda^{n-1})_n \in l^2,$$

d.h.

$$x_1 = 0 \quad \text{oder} \quad \sum_{n=1}^{\infty} \left|\lambda^{n-1}\right|^2 < \infty.$$

Wir können aber $x_1 = 0$ ausschließen, da sonst $x_i = 0$ für alle $i \in \mathbb{N}$ folgen würde. Die andere Lösung ergibt

$$\sum_{n=0}^{\infty} (|\lambda|^2)^n < \infty \Longleftrightarrow |\lambda| < 1, \tag{4.27}$$

d.h.

$$\sigma_p(A') = \{\, \lambda \in \mathbb{C} \mid |\lambda| < 1 \,\}.$$

Für A folgt aufgrund des Lemmas 4.4

$$\sigma_r(A) \subset \sigma_p(A') \subset \underbrace{\sigma_p(A)}_{=\emptyset} \dot{\cup} \sigma_r(A),$$

d.h.

$$\sigma_r(A) = \sigma_p(A') = \{\, \lambda \in \mathbb{C} \mid |\lambda| < 1 \,\}.$$

Weil $r(A) = 1$ gilt, ist

$$\sigma(A) \subset \{\, \lambda \in \mathbb{C} \mid |\lambda \leq 1| \,\}$$

und damit folgt zusammen wegen der Kompaktheit von $\sigma(A)$ (vgl. Satz 4.6), dass

$$\sigma(A) = \{\, \lambda \in \mathbb{C} \mid |\lambda| < 1 \,\}$$

und

$$\sigma_c(A) = \{\, \lambda \in \mathbb{C} \mid |\lambda| = 1 \,\}.$$

□

Satz 4.10. *Seien X ein Banachraum und $A \in L(X)$. Dann gilt*

(i) A ist invertierbar genau dann, wenn A' invertierbar ist.

(ii) $\sigma(A) = \sigma(A')$.

Beweis. Zu (ii): Das folgt aus (i), denn

$$\lambda \in \rho(A) \Longleftrightarrow \underbrace{\lambda I - A}_{\in L(X)} \text{ invertierbar} \overset{(i)}{\Longleftrightarrow} \lambda I' - A' \text{ invertierbar} \Longleftrightarrow \lambda \in \rho(A').$$

Zu (i): Zunächst zeigen wir „$\Longrightarrow$“: Sei also A invertierbar. Dann ist $A^{-1} \in L(X)$, also $(A^{-1})' \in L(X')$ und

$$(A^{-1})' \circ A' = (A \circ A^{-1})' = I' = (A^{-1} \circ A)' = A' \circ (A^{-1})'.$$

Somit ist also A' invertierbar und

$$(A')^{-1} = (A^{-1})'.$$

Zu „$\Longleftarrow$“: Sei also A' invertierbar. Dann ist anch dem eben gezeigten auch A'' invertierbar in $L(X'')$.

Wir erinnern uns an den Beweis des Satzes von Schauder (Satz 3.19). Dort hatten wir folgendes gezeigt:

$$\begin{array}{ccccc} & X & \xrightarrow{A} & X & \\ j_X & \downarrow & & \downarrow & j_X \\ & X'' & \xrightarrow{A''} & X'' & \end{array}, \qquad A'' \circ j_X = j_X \circ A. \tag{4.28}$$

Zunächst ist $C := \left\|(A'')^{-1}\right\| > 0$, da A'' invertierbar ist. Also gilt die Abschätzung

$$\|j_X(x)\|_{X''} = \left\|(A'')^{-1}(A''(j_X(x)))\right\|_{X''} \leq \underbrace{\left\|(A'')^{-1}\right\|_{L(X'')}}_{=C>0} \|A'' j_X(x)\|_{X''} \tag{4.29}$$

Es folgt für alle $x \in X$

$$\|Ax\|_X \overset{(*)}{=} \|j_X(Ax)\|_{X''} \overset{(4.28)}{=} \|A''(j_X(x))\|_{X''} \overset{(4.29)}{\geq} \tfrac{1}{C}\|j_X(x)\|_{X''} \overset{(*)}{=} \tfrac{1}{C}\|x\|_X. \tag{4.30}$$

An den Stellen $(*)$ fließt ein, dass j_X eine Isometrie ist. Damit ist A injektiv, denn aus $Ax = 0$ folgt

$$\|Ax\| = 0 \overset{(4.30)}{\geq} \tfrac{1}{C}\|x\| \Longrightarrow \|x\| = 0 \Longrightarrow x = 0.$$

Nach Korollar 3.7 ist A surjektiv, da A' injektiv ist.

Damit haben wir gezeigt, dass A bijektiv ist. Nach Korollar 2.2 zum Satz von der offenen Abbildung ist somit A invertierbar. $\square$

4.3 Das Spektrum von kompakten Operatoren

Vorbereitend zunächst ein sehr wichtiges

Lemma 4.5 (von Riesz). *Seien $(X, \|.\|_X)$ ein normierter Vektorraum und U ein echter abgeschlossener Untervektorraum von X. Dann existiert zu jedem $0 < \delta < 1$ ein $x \in X$ mit*

$$\|x\| = 1 \quad \text{und} \quad d(x, U) := \inf\{\, \|x - u\| \mid u \in U \,\} \geq 1 - \delta.$$

Bemerkung. Das Riezsche Lemma gilt im Allgemeinen nicht für $\delta = 0$. Vergleiche dazu das Beispiel 4.7

Beweis des Lemmas. Sei also $1 > \delta > 0$ und $z \notin U$, d.h. $z \in X \setminus U$. Setze

$$d := d(z, U) > 0.$$

Beachte, dass tatsächlich $d > 0$, denn andernfalls gäbe es eine Folge $(u_n)_n$ mit $\|u_n - z\| \to 0$ und z läge in $\bar{U} = U$. Doch z liegt nicht in U.

Weiter gilt dann

$$d < \tfrac{d}{1-\delta}$$

und es existiert somit ein $u \in U$ mit

$$0 < d \leq \|z - u\| < \tfrac{d}{1-\delta} \tag{4.31}$$

Setze dann

$$x := \tfrac{u-z}{\|u-z\|}.$$

Dann ist $x \notin U$, denn wäre $x \in U$, so folgte

$$z = u - \|u - z\| x \in U,$$

und das ist ein Widerspruch. Weiter ist $\|x\| = 1$ und für alle $y \in U$ gilt

$$\begin{aligned} \|x - y\| &= \left\| \tfrac{u-z}{\|u-z\|} - y \right\| = \tfrac{1}{\|u-z\|} \|u - z - \|u - z\| y\| \\ &= \tfrac{1}{\|u-z\|} ||z - \underbrace{(u - \|u - z\| y)}_{\in U}|| \geq \tfrac{d}{\|u-z\|} \overset{(4.31)}{>} 1 - \delta. \end{aligned}$$

□

Mit dem Riezschen Lemma können wir nämlich nun die Umkehrung des Satzes von Bolzano-Weierstraß beweisen:

Satz 4.11 (Kompaktheitssatz von F. Riesz). *Für einen normierten Raum X sind äquivalent:*

(i) $\dim X < \infty$.

(ii) $B_X = \{\, x \in X \mid \|x\| \leq 1 \,\}$ *ist kompakt.*

(iii) Jede bschränkte Folge in X besitzt eine konvergente Teilfolge.

Beweis. Zu $(i) \Longrightarrow (ii)$: Nach dem Satz von Heine-Borel (Satz 1.9) ist $B_X \subset X$ genau dann kompakt, wenn B_X beschränkt und abgeschlossen ist. Fertig.

Zu $(ii) \Longrightarrow (iii)$: In einem kompakten metrischen Raum besitzt nach Satz 1.10 jede Folge eine konvergente Teilfolge.

Zu $(iii) \Longrightarrow (i)$: Angenommen, $\dim X = \infty$. Sei $x_1 \in X$ mit $\|x_1\| = 1$. Setze

$$U_1 := \operatorname{lin}\{x_1\}.$$

Dann ist U_1 endlich-dimensional, also abgeschlossen und von X verschieden.

Nach dem Rieszschen Lemma (Lemma 4.5) angewendet auf $\delta = 1/2$, existiert ein $x_2 \in X$ mit

$$\|x_2\| = 1 \quad \text{und} \quad \|x_2 - x_1\| \geq 1/2.$$

Betrachte jetzt

$$U_2 := \operatorname{lin}\{x_1, x_2\}.$$

Wende das Lemma erneut an: Dann gibt es $x_3 \in X$ mit

$$\|x_3\| = 1, \quad \|x_3 - x_1\| \geq 1/2 \quad \text{und} \quad \|x_3 - x_2\| \geq 1/2.$$

Setze das Verfahren fort, so dass induktiv eine Folge $(x_n)_n$ mit $\|x_n\| = 1$ und $\|x_n - x_m\| \geq 1/2$ für alle $m, n \in \mathbb{N}$, $m \neq n$, definiert wird.

Die Folge $(x_n)_n$ ist offenbar beschränkt, ist aber keine Cauchy-Folge, also enthält erst recht keine konvergente Teilfolge. Widerspruch! $\square$

Beispiel 4.7. Betrachte

$$X := \{\, f \in C[0,1] \mid f(1) = 0 \,\} \quad \text{und} \quad U := \{\, f \in X \mid \int_0^1 f(t)dt = 0 \,\}.$$

Versehe dann X mit der $\|.\|_\infty$-Norm. Offensichtlich ist U ein echter(!) Untervektorraum von X, denn beispielsweise ist $g(t) := \sin(t-1) \in X$, aber $g(t) \notin U$. Weiter ist U abgeschlossen: Ist nämlich $(f_n)_n \subset U$ eine Folge mit $f_n \to f \in X$, so folgt aufgrund der gleichmäßigen Konvergenz gerade

$$0 = \lim \int_0^1 f_n dt = \int_0^1 \lim f_n dt = \int_0^1 f dt,$$

also ist auch $f \in U$.

Gäbe es nun ein Element $f \in X$ mit

$$\|f - u\|_\infty \geq \|f\|_\infty = 1 \quad \text{für alle } u \in U, \tag{4.32}$$

so bekäme man den folgenden Widerspruch:

Setze

$$f_n(t) := 1 - t^n.$$

Dann ist $f_n \in X$ mit

$$\|f_n\|_\infty = 1 \quad \text{für alle } n \in \mathbb{N} \tag{4.33}$$

und

$$\int_0^1 f_n(t)dt = 1 - \tfrac{1}{n+1} \quad \text{für alle } n \in \mathbb{N}. \tag{4.34}$$

Definiere

$$\lambda_n := \frac{\int_0^1 f(t)dt}{1 - \frac{1}{n+1}} \quad \text{und} \quad u_n := f - \lambda_n f_n.$$

Dann ist $u_n \in U$ für alle $n \in \mathbb{N}$, weil

$$\int_0^1 u_n(t)dt = \int_0^1 f(t)dt - \lambda_n \int_0^1 f_n(t)dt = (1 - \tfrac{1}{n+1})\lambda_n - \lambda_n \int_0^1 f_n(t)dt \overset{(4.34)}{=} 0.$$

Wegen der Annahme (4.32) gilt dann für alle $n \in \mathbb{N}$

$$\|f - u_n\|_\infty = |\lambda_n| \, \|f_n\|_\infty \overset{(4.33)}{=} |\lambda_n| \geq 1,$$

also

$$\left|\int_0^1 f(t)dt\right| \geq 1.$$

Auf der anderen Seite gilt nach Annahme

$$\left|\int_0^1 f(t)dt\right| \leq \int_0^1 |f(t)|\, dt \leq \|f\|_\infty = 1.$$

Damit ist also

$$\left|\int_0^1 f(t)dt\right| = 1.$$

Zunächst ist wegen $\|f\|_\infty \leq 1$ immer $f(t) \leq 1$ für alle $t \in [0,1]$. Da $f(1) = 0$ und f stetig ist, gibt es ein $\epsilon > 0$, so dass

$$1 > |f(t)| > 0 \quad \text{für alle } t \in]1-\epsilon, 1[.$$

Daraus folgt

$$1 = \left|\int_0^1 f(t)dt\right| \leq \int_0^{1-\epsilon} |f(t)|\, dt + \int_{1-\epsilon}^1 |f(t)|\, dt < 1 - \epsilon + 1 - (1-\epsilon) = 1.$$

Widerspruch.

□

Lemma 4.6. *Seien $A \in K(X)$, X ein Banachraum und $S = I - A$. Dann gilt*

(i) $\ker(S)$ *ist ein endlich-dimensionaler Untervektorraum von X.*

(ii) $R(S)$ *ist ein abgeschlossener Untervektorraum von X.*

Beweis. Zu (i): Zunächst ist klar, dass

$$\ker(S) = S^{-1}(\{0\})$$

ein abgeschlossener Untervektorraum von X ist, da S stetig ist.

Sei jetzt $(x_n)_n$ eine beschränkte Folge in $\ker(S)$. Da A kompakt ist, existiert eine Teilfolge $(x_{n_k})_k$, so dass

$$Ax_{n_k} \to y \quad \text{in } X \tag{4.35}$$

für ein gewisses $y \in X$ (vgl. Bemerkung, Punkt 2 zur Definition 3.7 eines kompakten Operators).

Weiter gilt wegen $x_{n_k} \in \ker(S)$

$$Sx_{n_k} = 0 \Longleftrightarrow x_{n_k} - Ax_{n_k} = 0 \Longleftrightarrow x_{n_k} = Ax_{n_k} \quad \text{für alle } k.$$

Somit ist

$$x_{n_k} \overset{(4.35)}{\longrightarrow} y \quad \text{in } \ker(S),$$

da $\ker(S)$ abgeschlossen ist.

Damit ist gezeigt, dass jede beschränkte Folge in $\ker(S)$ eine konvergente Teilfolge besitzt.

Nach Satz 4.11 ist schließlich

$$\dim(\ker(S)) < \infty.$$

Zu (ii): Da $\dim(\ker(S)) < \infty$, existiert nach Lemma 3.3 ein abgeschlossener Untervektorraum V von X mit

$$X = \ker(S) \oplus V.$$

Es gilt also insbesondere

$$\ker(S) \cap V = \{0\}. \tag{4.36}$$

Betrachte nun die Einschränkung von S auf V, d.h.

$$S_V := S|_V.$$

Dann gilt

- S_V ist injektiv auf V: Aus $S_V x = 0$ für $x \in V$ folgt nämlich wegen (4.36) gerade $x = 0$.
- Es gilt $R(S_V) = R(S)$:

  Es ist trivial, dass $R(S_V) \subset R(S)$, denn S_V ist Einschränkung von S.

  Die andere Implikation $R(S) \subset R(S_V)$ folgt für $x = u + v$ mit $u \in \ker(S)$ und $v \in V$ aus

  $$Sx = S(u + v) = Su + Sv = Sv.$$

Es genügt daher zu zeigen, dass $R(S_V)$ abgeschlossen ist.

Sei dazu $y \in \overline{R(S_V)}$. Dann existiert eine Folge $(x_n)_n \subset V$ mit

$$\underbrace{S_V x_n}_{x_n - Ax_n} = Sx_n \to y \quad \text{in } X. \tag{4.37}$$

Angenommen, $(x_n)_n$ ist unbeschränkt. O.B.d.A. können wir annehmen, dass

$$\|x_n\| \to +\infty. \tag{4.38}$$

(Sonst gehen wir zu einer Teilfolge über: Bsp.: $a_n = (1 + (-1)^n)n$).

Betrachte dann

$$v_n := \tfrac{x_n}{\|x_n\|}, \quad n \in \mathbb{N}.$$

Wegen (4.37) ist $S_v(x_n)$ beschränkt und mit (4.38) folgt dann

$$\underbrace{S_V(v_n)}_{v_n - Av_n} = \tfrac{1}{\|x_n\|} S_V(x_n) \to 0 \quad \text{in } X. \tag{4.39}$$

Da $(v_n)_n$ beschränkt (denn $\|v_n\| = 1$ für alle n) und A kompakt ist, existiert eine Teilfolge $(v_{n_k})_k$ mit

$$Av_{n_k} \to z \quad \text{in } X \tag{4.40}$$

Mit (4.39) folgt dann

$$v_{n_k} \to z \in V,$$

da V abgeschlossen ist und

$$\|v_{n_k} - z\| \leq \underbrace{\|v_{n_k} - Av_{n_k}\|}_{\overset{(4.39)}{\to} 0} + \underbrace{\|Av_{n_k} - z\|}_{\overset{(4.40)}{\to} 0} \to 0.$$

Weil $\|v_n\| = 1$ für alle $n \in \mathbb{N}$, folgt auch

$$\|z\| = 1,$$

also insbesondere $z \neq 0$.

Nochmaliges Ausnutzen von (4.39) ergibt aber

$$S_V(z) = 0,$$

obwohl $z \neq 0$, d.h. S_V ist nicht injektiv. Widerpsurch!

Die Annahme der Unbeschränktheit von $(x_n)_n$ ist also falsch. Demnach ist $(x_n)_n$ beschränkt und da A kompakt ist, existiert eine Teilfolge $(x_{n_k})_k$ mit

$$Ax_{n_k} \to z \quad \text{in } X.$$

Weil nach (4.37)

$$S_V x_{n_k} = x_{n_k} - Ax_{n_k} \to y,$$

folgt so, dass

$$x_{n_k} = S_V x_{n_k} + Ax_{n_k} \to y + z =: x \quad \text{in } V.$$

Da S_V stetig ist, folgt somit

$$S_V(x_{n_k}) \to S_V(x) \overset{(4.37)}{=} y \in R(S_V).$$

□

Satz 4.12 (**Fredholm'sche Alternative**). *Seien X ein Banachraum, $A \in K(X)$ und $S := I - A$. Dann sind äquivalent:*

(i) S bijektiv.

(ii) S injektiv.

(iii) S surjektiv.

Beweis. Zu *(ii)* $\Longrightarrow$ *(iii)*: Sei also S injektiv. Angenommen, S ist nicht surjektiv, d.h.

$$X_1 := R(S) \neq X.$$

Dann ist nach Lemma 4.6 X_1 ein echter, abgeschlossener Untervektorraum von X, d.h. $X_1 \subsetneq X$.

Definiere nun induktiv die Folge $(X_n)_n$ von Untervektorräumen mit

$$X_{n+1} := S(X_n) \quad \text{für alle } n \in \mathbb{N}.$$

Dann gilt:

- X_n ist invariant unter S, d.h. $S(X_n) \subset X_n$ und dann auch unter A:

  a) Aus der Invarianz unter S folgt die Invarianz unter A, denn

  $$x \in X_n \Longrightarrow Ax = \underbrace{x}_{\in X_n} - \underbrace{Sx}_{\in X_n} \in X_n.$$

  b) Invarianz unter S: Induktiv.

  $\underline{n = 1:}$ Aus $x \in X_1 = S(X)$ folgt $Sx \in S(X) = X_1$.

  $\underline{n \to n+1:}$ Sei $x \in X_{n+1} = S(X_n)$. Dann gibt es ein $y \in X_n$ mit $x = Sy$. Nach Induktionsvoraussetzung ist X_n invariant unter S, also

  $$x = Sy \in X_n \Longrightarrow Sx \in X_{n+1} = S(X_n).$$

- X_n ist ein abgeschlossener Untervektorraum von X für alle $n \in \mathbb{N}$ (Das folgt wieder mit Lemma 4.6).

- $X_{n+1} \subsetneq X_n$ für alle $n \in \mathbb{N}$.

  a) $X_{n+1} \subset X_n$ folgt sofort aus der Invarianz von X_n unter S, denn

  $$X_{n+1} = S(X_n) \subset X_n.$$

b) $X_{n+1} \neq X_n$: Da $R(S) \neq X$, existiert ein $y \in X$ mit

$$Sx \neq y \quad \text{für alle } x \in X.$$

Da S injektiv ist (denn $S(Sx) \neq S(y)$ für $Sx \neq y$), folgt induktiv

$$\underbrace{S^{n+1}x}_{\in S^{n+1}(X)=X_{n+1}} \neq \underbrace{S^n y}_{\in S^n(X)=X_n} \quad \text{für alle } x \in X,$$

also $X_{n+1} \neq X_n$.

Nach dem Lemma von Riesz (Lemma 4.5) existiert für alle $n \in \mathbb{N}$ ein $u_n \in X_n$ mit

$$\|u_n\| = 1 \quad \text{und} \quad d(u_n, X_{n+1}) \geq 1/2. \tag{4.41}$$

Sei $m > n$. Dann ist $X_m \subset X_{n+1}$ und daher

$$\underbrace{Su_n}_{\in X_{n+1}} + \underbrace{u_m - Su_m}_{=Au_m \in X_m \subset X_{n+1}} \in X_{n+1}$$

und

$$\|Au_n - Au_m\| = \|u_n - (Su_n + u_m - Su_m)\| \overset{(4.41)}{\geq} 1/2 \quad \text{für alle } m > n.$$

Somit ist $(Au_n)_n$ keine Cauchy-Folge und enthält daher auch keine konvergente Teilfolge. Widerspruch, da $(u_n)_n$ beschränkt und A kompakt ist und deshalb $(Au_n)_n$ eine konvergente Teilfolge besitzt. Also folgt die Surjektivität von S.

Zu $(iii) \Longrightarrow (ii)$: Nun sei $S = I - A$ surjektiv. Dann ist nach Korollar 3.7 S' injektiv.

Nach dem Satz von Schauder (Satz 3.19) ist mit A auch A' kompakt. Die schon bewiesende Implikation angewendet auf

$$S' = I' - A'$$

liefert somit, dass S' surjektiv ist. Also ist S' bijektiv und somit invertierbar. Mit Satz 4.10 folgt dann, dass S invertierbar ist, also ist S injektiv. □

Beispiel 4.8 (**Anwendung der Fredholm'schen Alternative**). Sei $k \in C([a,b]^2)$. Dann ist

$$A : C[a,b] \to C[a,b], \quad u \mapsto \int_a^b k(.,t)u(t)dt$$

kompakt (vgl. Beispiel 3.18). Daher gilt:

Die Integralgleichung

$$u(s) - \int_a^b k(s,t)u(t)dt = f(s), \quad s \in [a,b] \tag{4.42}$$

besitzt eine eindeutige Lösung $u \in C[a,b]$ für alle $f \in C[a,b]$ genau dann, wenn (4.42) für jedes $f \in C[a,b]$ mindestens eine Lösung $u \in C[a,b]$ besitzt (Surjektivität) bzw. genau dann, wenn die homogene Gleichung

$$u(s) - \int_a^b k(s,t)u(t)dt = 0, \quad s \in [a,b]$$

nur die triviale Lösung $u \equiv 0$ besitzt (Injektivität).

□

Satz 4.13 (**Spektralsatz für kompakte Operatoren**). *Sei X ein unendlichdimensionaler Banachraum und sei $A \in K(X)$. Dann gilt*

(i) $0 \in \sigma(A)$.

(ii) $\sigma(A) \setminus \{0\} \subset \sigma_p(A)$.

(iii) $\sigma(A)$ *ist endlich oder es gibt eine Nullfolge* $(\lambda_n)_n \subset \mathbb{C} \setminus \{0\}$ *mit*

$$\sigma(A) = \{\, \lambda_n \mid n \in \mathbb{N} \,\} \cup \{0\}.$$

Bemerkungen.

1. Im endlichdimensionalen Fall braucht (i) nicht gelten.
2. Es gilt nicht immer $\sigma(A) \setminus \{0\} = \sigma_p(A)$, da eventuell $0 \in \sigma_p(A)$.

Beweis. Zu (i): Angenommen $0 \notin \sigma(A)$. Dann ist (vgl. Eigenschaften von kompakten Operatoren (3.26)):

$$I = \underbrace{A^{-1}}_{\in L(X)} \circ \underbrace{A}_{\in K(X)} \in K(X).$$

Somit besitzt dann jede beschränkte Folge in X eine konvergente Teilfolge. Mit Satz 4.11 folgt $\dim X < \infty$. Widerspruch.

Zu (ii): Sei $\lambda \in \mathbb{C} \setminus \{0\}$ mit $\lambda \notin \sigma_p(A)$. Dann ist $(\lambda I - A)$ injektiv. Aber

$$\lambda I - A = \lambda(I - \underbrace{\tfrac{1}{\lambda}A}_{\in K(X)})$$

und aus dem Satz der Fredholm'schen Alternative (Satz 4.12) folgt somit, dass $I - \frac{1}{\lambda}A$ invertierbar ist, d.h. aber auch, dass $\lambda I - A$ invertierbar ist. Also ist $\lambda \notin \sigma(A)$.

Zu (iii): Sei $\epsilon > 0$. Angenommen,

$$\lambda \in \{\, \sigma(A) \mid |\lambda| > \epsilon \,\}$$

ist unendlich. Dann existiert eine Folge von Eigenwerten $(\lambda_n)_n$ von A und zugehörige Eigenvektoren $(x_n)_n$ mit

$$\lambda_n \neq \lambda_m \quad \text{für alle } m \neq n \quad \text{und} \quad \|x_n\| = 1 \quad \text{für alle } n \in \mathbb{N}.$$

Zwischenbehauptung: Die Eigenvektoren zu verschiedenen Eigenwerten sind linear unabhängig. Das heißt hier konkret:

$$\{\, x_n \mid n \in \mathbb{N} \,\} \text{ ist linear unabhängig.}$$

Wir beweisen das induktiv:

$n = 1$: Trivial, da $x_1 \neq 0$, denn x_1 ist Eigenvektor.

$n \to n+1$: Angenommen,

$$x_{n+1} = \sum_{i=1}^{n} \alpha_i x_i, \quad \alpha_i \in \mathbb{C},\ i = 1, \ldots, n.$$

Auf der einen Seite ist

$$A(x_{n+1}) = \lambda_{n+1} x_{n+1} = \lambda_{n+1} \sum_{i=1}^{n} \alpha_i x_i$$

und auf der anderen Seite

$$A(x_{n+1}) = A(\sum_{i=1}^{n} \alpha_i x_i) = \sum_{i=1}^{n} \alpha_i \lambda_i x_i.$$

Zusammenfassend heißt das

$$\sum_{i=1}^{n} \alpha_i (\lambda_{n+1} - \lambda_i) x_i = 0.$$

Die x_i sind nach Induktionsvorausetzung linear unabhängig. Also folgt

$$\alpha_i \underbrace{(\lambda_{n+1} - \lambda_i)}_{\neq 0} = 0 \quad \text{für alle } i$$

und so

$$\alpha_i = 0 \quad \text{für alle } i \Longrightarrow x_{n+1} = 0.$$

Widerspruch, da x_{n+1} Eigenvektor und Eigenvektoren $\neq 0$ sind.

Weiter im eigentlichen Beweis: Sei nun

$$X_n := \operatorname{lin}\{x_1, \ldots, x_n\}, \quad n \in \mathbb{N}.$$

Dann ist $X_{n-1} \subset X_n$. Nach dem Lemma von Riesz (Lemma 4.5) existiert zu jedem $n \in \mathbb{N}$ ein $u_n \in X_n$ mit

$$\|u_n\| = 1 \quad \text{und} \quad d(u_n, X_{n-1}) \geq 1/2. \tag{4.43}$$

Für die $(u_n)_n$ gilt:

- $Au_n \in X_n$, denn
$$u_n = \sum_{i=1}^{n} \alpha_i x_i \Longrightarrow Au_n = \sum_{i=1}^{n} \alpha_i \lambda_i x_i \in X_n.$$
- $\lambda_n u_n - Au_n \in X_{n-1}$, denn:
$$\lambda_n u_n - Au_n = \lambda_n \sum_{i=1}^{n} \alpha_i x_i - \sum_{i=1}^{n} \alpha_i \lambda_i x_i = \sum_{i=1}^{n-1} (\lambda_n - \lambda_i)\alpha_i x_i \in X_{n-1}.$$

Damit folgt für alle $m < n$:

$$\begin{aligned} \|Au_n - Au_m\| &= \|\lambda_n u_n - (Au_m + \lambda_n u_n - Au_n)\| \\ &\overset{\lambda_n \neq 0}{=} |\lambda_n| \, ||u_n - \tfrac{1}{\lambda_n}(Au_m + \lambda_n u_n - Au_n)||. \end{aligned}$$

Wir bemerken, dass $Au_n \in X_{n-1}$ und $Au_m \in X_m \subset X_{n-1}$, da $m < n$. Also folgt

$$\|Au_n - Au_m\| \geq \underbrace{|\lambda_n|}_{\geq \epsilon} \underbrace{d(u_n, X_{n-1})}_{\overset{(4.43)}{\geq} 1/2} \geq \epsilon/2,$$

d.h. (Au_n) enthält keine Cauchy-Folge und somit keine konvergente Teilfolge. Widerpsurch, da $(u_n)_n$ beschränkt und A kompakt. □

Rückblick. Sei $A \in K(X)$ und sei $\lambda \in \mathbb{C} \setminus \{0\}$. Dann bedeutet die Fredholm-Alternative:

$$(\lambda I - A)u = f \text{ ist eindeutig lösbar in } X \text{ für alle } f \in X \tag{4.44}$$

oder äquivalent dazu:

$$(\lambda I - A)u = 0 \text{ besitzt nur die triviale Lösung } u \equiv 0. \tag{4.45}$$

Wenn

$$\ker(\lambda I - A) \neq \{0\},$$

d.h. (4.45) nicht-triviale Lösungen besitzt, dann gilt

$$\text{(4.44) ist lösbar (aber nicht eindeutig) für ein } f \in X$$

oder äquivalent dazu

$$f \in R(\lambda I - A) = \overline{R(\lambda I - A)} \overset{\text{Satz 3.20}}{=} Ker(\lambda I' - A')_{\perp}.$$

Das wiederum ist gleichbedeutend damit, dass

$$\varphi(f) = 0 \quad \text{für alle } \varphi \in \ker(\lambda I' - A')$$

bzw. dass für alle Lösungen der homogenen adjungierten Gleichung

$$(\lambda I' - A')\varphi = 0$$

gerade

$$\varphi(f) = 0$$

gilt.

4.4 Quotientenräume

Aus dem endlichdimensionalen ist Ihnen die Dimensionsformel

$$\dim(X) = \dim(\ker(S)) + \dim(R(S))$$

für $S \in L(X)$ wohlbekannt. Umstellen dieser Gleichung ergibt

$$\dim(X) - \dim(R(S)) = \dim(\ker(S)).$$

Oder anders ausgedrückt

$$\dim(X/R(S)) = \dim(\ker(S))$$

Der linke Ausdruckt stellt dabei ein Maß für die Nicht-Surjektivität von S dar. Dagegen ist $\dim(\ker(S))$, die so genannte *Codimension von* $R(S)$, ein Maß für die nicht-Injektivität. Wir werden nun den unendlichdimensionalen Fall untersuchen:

Sei jetzt also $\dim X = \infty$ und sei $(X, \|.\|)$ ein normierter Vektorraum. Weiter sei U ein abgeschlossener(!) Untervekltorraum von X. Dann definiert man eine Äquivalenzrelation: Für $x, y \in X$ definiere

$$x \sim_U y :\Longleftrightarrow x - y \in U.$$

Die Menge

$$X/U := \{\, x + U \mid x \in X \,\}$$

ist die *Menge aller Äquivalenzklassen modulo* U.

In der linearen Algebra lernt man, dass ein Quotientenraum eine Vektorraumstruktur besitzt:

$$(x + U) + (y + U) = (x + y + U),$$
$$\lambda(x + U) = \lambda x + U.$$

Außerdem definiert man die *Quotientenabbildung*

$$q : X \to X/U, \quad x \mapsto x + U.$$

Satz 4.14 (**Eigenschaften.**).

1. *q ist linear.*
2. *q ist surjektiv.*
3. $\ker(q) = U$.

Beweis. Zu 1.: Für $x, y \in X$ und $\lambda, \mu \in \mathbb{K}$ gilt

$$q(\lambda x + \mu y) = (\lambda x + \mu y) + U = \lambda(x + U) + \mu(y + U) = \lambda q(x) + \mu q(y).$$

Zu 2.: Sei $[b] \in X/U$. Dann gibt es ein $x \in X$, so dass

$$b = x + U,$$

also ist q surjektiv.

Zu 3.: Die Behauptung folgt aus

$$q(u) = u + U = 0 \Longleftrightarrow u \sim 0 \Longleftrightarrow u \in U.$$

□

Wir definieren nun eine Norm auf X/U durch

$$\|q(x)\|_{X/U} := d(x, U) = \inf\{\, \|x - u\| \mid u \in U \,\}.$$

Beweis. Zu (i): Ist $q(x) = 0$, d.h. $x \in U$, so folgt $\|q(x)\| = d(x, U) = 0$. Auf der anderen Seite folgt aus $\|q(x)\| = d(x, U) = 0$ gerade $x \in \bar{U} = U$, also $q(x) = 0$, denn angenommen, $x \notin U$. Wir zeigen, dass x Randpunkt von U ist, also

$$\forall_{\epsilon>0} \exists_{y \in U} : y \in B(\epsilon, x),$$

und erhalten so einen Widerspruch.

Wenn also x kein Randpunkt ist, so gibt es ein $\epsilon' > 0$, so dass für alle $y \in U$ gilt $y \notin B(\epsilon', x)$. Doch dann ist $d(x, y) \geq \epsilon'$ für alle $y \in U$ und somit

$$d(x, U) = \inf\{\, \|x - u\| \mid u \in U \,\} \geq \epsilon' > 0.$$

Widerspruch, da $d(x, U) = 0$ nach Voraussetzung. Also ist x doch Randpunkt und wegen $U = \bar{U}$ ist $x \in U$.

Zu (ii): Wir müssen zeigen, dass $\|q(\lambda x)\| = |\lambda| \, \|q(x)\|$.

1. Fall: $\lambda = 0$. Dann ist

$$\|q(0 \cdot x) = \|q(0)\|\| = d(0, U) = \inf\{\, \|u\| \mid u \in U \,\} = 0,$$

da U Untervektorraum ist und so $0 \in U$ gilt.

2. Fall: $\lambda \neq 0$. Mit $\tilde{u}$ ist auch $\frac{\tilde{u}}{\lambda} \in U$, da U Untervektorraum ist. Daraus folgt

$$\begin{aligned}
\|q(\lambda x)\| &= d(\lambda x, U) = \inf\{\, \|\lambda x - \tilde{u}\| \mid \tilde{u} \in U \,\} \\
&= |\lambda| \inf\{\, \left\|x - \tfrac{\tilde{u}}{\lambda}\right\| \mid \tilde{u} \in U \,\} = |\lambda| \inf\{\, \|x - u\| \mid u \in U \,\} \\
&= |\lambda| \, d(x, U) = |\lambda| \, \|q(x)\|.
\end{aligned}$$

Zu (iii): Wir müssen die Dreiecksungleichung

$$\|q(x) + q(y)\| \leq \|q(x)\| + \|q(y)\|$$

zeigen. Seien dazu $x, y \in X$. Wähle zu $\epsilon > 0$ Elemente $u_1, u_2 \in U$ mit

$$\|x - u_1\| \leq \|q(x)\| + \epsilon, \quad \|y - u_2\| \leq \|q(y)\| + \epsilon.$$

Dann folgt

$$\begin{aligned}
\|q(x) + q(y)\| = \|q(x + y)\| &\leq ||(x + y) - \underbrace{(u_1 + u_2)}_{\in U}|| \\
&\leq \|x - u_1\| + \|y - u_2\| \leq \|q(x)\| + \|q(y)\| + 2\epsilon.
\end{aligned}$$

Da $\epsilon > 0$ beliebig war, folgt so die Dreiecksungleichung.

Zu (iv): Wir müssen noch die Wohldefiniertheit zeigen. Für $x_1, x_2 \in X$ sei also $q(x_1) = q(x_2)$. Dann gibt es ein $u \in U$ mit

$$x_1 = x_2 + u.$$

Also folgt

$$\begin{aligned}
\|q(x_1)\| &= d(x, U) = \inf\{\, \|x_1 - \tilde{u}\| \mid \tilde{u} \in U \,\} \\
&= \inf\{\, \|x_2 + u - \tilde{u}\| \mid \tilde{u} \in U \,\} = \inf\{\, \|x_2 - v\| \mid v \in U \,\} \\
&= d(x_2, U) = \|q(x_2)\|.
\end{aligned}$$

□

Satz 4.15.

1. *Bezüglich der Norm* $\|.\|_{X/U}$ *ist* $q \in L(X, X/U)$.
2. *Ist* X *ein Banachraum, so ist auch* X/U *ein Banachraum.*
3. $$U' \cong X'/U^{\perp} \tag{4.46}$$
$$U^{\perp} \cong (X/U)' \tag{4.47}$$
4. q *ist eine Quotientenabbildung im Sinne der Definition (siehe Beispiel 2.15).*

Beweis. Zu 1.: Dass q linear ist, haben wir schon oben gesehen. q ist aber auch stetig, denn

$$\|q\| = \sup_{x \in B_X} \|q(x)\| = \sup_{x \in B_X} \inf\{\, \|x - u\| \mid u \in U \,\} \overset{u=0}{\leq} \sup_{x \in B_X} \|x\| \leq 1.$$

Wir zeigen sogar

$$\|q\| = 1. \tag{4.48}$$

Nach dem Lemma von Riesz (Lemma 4.5) gibt es nämlich zu jedem $0 < \delta < 1$ ein $x_\delta \in X$ mit

$$\|x_\delta\| = 1 \quad \text{und} \quad \|q(x_\delta)\| = d(x_\delta, U) \geq 1 - \delta.$$

Damit folgt

$$\|q\| = \sup_{x \in B_X} \|q(x)\| \geq \|q(x_\delta)\| \geq 1 - \delta.$$

Da δ beliebig war, folgt schließlich $\|q\| \geq 1$.

Zu 2.: Sei also $(q(x_n))_n$ eine Cauchy-Folge in X/U. Dann existiert eine Teilfolge $q(x_{n_k})_k$ mit

$$\left\|q(x_{n_k}) - q(x_{n_{k+1}})\right\| < 2^{-k} \quad \text{für alle } k \in \mathbb{N}. \tag{4.49}$$

Definiere

$$y_{n_k} := x_{n_k} - x_{n_{k+1}}.$$

Wegen der Linearität von q heißt (4.49) gerade

$$\left\|q(x_{n_k}) - q(x_{n_{k+1}})\right\| = \|q(y_{n_k})\| < 2^{-k} \quad \text{für alle } k \in \mathbb{N}. \tag{4.50}$$

Da

$$\|q(y_{n_k})\| = d(y_{n_k}, U),$$

existiert zu jedem k ein $u_{n_k} \in U$ mit

$$\|y_{n_k} - u_{n_k}\| < \|q(y_{n_k})\| + 2^{-k}. \tag{4.51}$$

Also gilt wegen (4.50) und (4.51)

$$\sum_{k=1}^{\infty} \|y_{n_k} - u_{n-k}\| \overset{(4.51)}{<} \sum_{k=1}^{\infty} (\|q(y_{n_k})\| + 2^{-k}) \overset{(4.50)}{<} \sum_{k=1}^{\infty} 2^{-k+1} < \infty.$$

Also ist $(\sum_{k=1}^{N} (y_{n_k} - u_{n_k}))_N$ eine Cauchy-Folge in X und da X ein Banachraum ist, konvergiert die obere Reihe, d.h.

$$\sum_{k=1}^{\infty} (y_{n_k} - u_{n_k}) < \infty.$$

Weil q stetig ist, folgt

$$q(\sum_{k=1}^{\infty}(y_{n_k}-u_{n_k})) = \lim_{N\to\infty} q(\sum_{k=1}^{N}(y_{n_k}-u_{n_k})) \overset{q \text{ linear}}{=} \lim_{N\to\infty}\sum_{k=1}^{N} q(y_{n_k}-\underbrace{u_{n_k}}_{\in U}) = \lim_{N\to\infty}\sum_{k=1}^{N} q(y_{n_k})$$

$$= \lim_{N\to\infty}\sum_{k=1}^{N} q(x_{n_k}) - q(x_{n_{k+1}}) \overset{\text{Teleskopsumme}}{=} \lim_{N\to\infty}(q(x_{n_1}) - q(x_{n_{N+1}})),$$

d.h. aber gerade, dass

$$\lim_{k\to\infty} q(x_{n_k})$$

in X/U existiert. Da $(q(x_n))_n$ eine Cauchy-Folge in X/U ist, konvergiert die gesamte Folge in X/U.

<u>Zu 3. (4.46):</u> Betrachte die Abbildung

$$\sigma : U' \to X'/U^{\perp}, \quad \varphi \mapsto \tilde{\varphi} + U^{\perp}, \tag{4.52}$$

wobei $\tilde{\varphi}$ eine beliebige Fortsetzung von φ auf ganz X sei (diese existiert nach dem Satz von Hahn-Banach immer).

Wir müssen zeigen, dass (4.52) ein isometrischer Isomorphismus ist:

a) <u>Wohldefiniertheit:</u> Sind $\tilde{\varphi}_1, \tilde{\varphi}_2 \in X'$ zwei Fortsetzungen von φ, so gilt

$$(\tilde{\varphi}_1 - \tilde{\varphi}_2)|_U = 0,$$

denn

$$\tilde{\varphi}_1|_U = \varphi = \tilde{\varphi}_2|_U.$$

Es folgt daraus

$$\tilde{\varphi}_1 - \tilde{\varphi}_2 \in U^{\perp} \Longrightarrow \tilde{\varphi}_1 + U^{\perp} = \tilde{\varphi}_2 + U^{\perp}.$$

b) <u>Linearität und Stetigkeit:</u> Klar.

c) <u>Surjektivität:</u> σ ist surjektiv, denn für alle $\psi \in X'$ (d.h. $\psi + U^{\perp} \in X'/U^{\perp}$) ist $\psi|_U \in U'$ und

$$\sigma(\psi|_U) = \psi + U^{\perp}.$$

d) <u>Isometrie:</u> Zunächst gilt für $\psi \in U^{\perp}$ und $\varphi \in U'$

$$\|\varphi - \psi\|_{X'} \geq \|\varphi - \psi\|_{U'} \geq \|\varphi\|_{U'} - \underbrace{\|\psi\|_{U'}}_{=0} = \|\varphi\|_{U'}.$$

Daraus folgt

$$\|\sigma(\varphi)\|_{X'/U^{\perp}} = \inf\{\, \|\varphi - \psi\|_{X'} \mid \psi \in U^{\perp} \,\} \geq \|\varphi\|_{U'},$$

also

$$\|\varphi\|_{U'} \leq \|\sigma(\varphi)\|_{X'/U^{\perp}}.$$

Auf der anderen Seite existiert nach dem Satz von Hahn-Banach (Satz 3.3) eine Fortsetzung $\tilde{\varphi} \in X'$ von φ mit

$$\|\varphi\|_{U'} = \|\tilde{\varphi}\|_{X'}.$$

Weiter gilt

$$\begin{aligned}\|\sigma(\varphi)\|_{X'/U^{\perp}} &= \|\tilde{\varphi} + U\|_{X'/U^{\perp}} = \|q(\tilde{\varphi})\|_{X'/U^{\perp}} \leq \|q\|_{(X/U)'}\|\tilde{\varphi}\|_{X'} \\ &= \|\tilde{\varphi}\|_{X'} = \|\varphi\|_{U'}.\end{aligned}$$

Insgesamt folgt also

$$\|\varphi\|_{U'} = \|\sigma(\varphi)\|_{X'/U^{\perp}}.$$

Zu 3. (4.47): Betrachte dazu die Abbildung

$$\tau : (X/U)' \to U^\perp, \quad \varphi \mapsto \varphi \circ q.$$

Wir müssen wieder zeigen, dass τ ein isometrischer Isomorphismus ist.

a) τ ist linear und stetig: Um die Linearität nachzuweisen, seien $\varphi, \tilde{\varphi} \in (X/U)'$ und $\mu, \lambda \in \mathbb{K}$. Dann gilt

$$\tau(\lambda\varphi + \mu\tilde{\varphi}) = (\lambda\varphi + \mu\tilde{\varphi}) \circ q = \lambda(\varphi \circ q) + \mu(\tilde{\varphi} \circ q) = \lambda\tau\varphi + \mu\tau(\tilde{\varphi}).$$

τ ist auch stetig, denn

$$\begin{aligned}\|\tau\|_{L((X/U)',U^\perp)} &= \sup_{\varphi \in B_{(X/U)'}} \|\tau(\varphi)\|_{X'} = \sup_{\varphi \in B_{(X/U)'}} \|\varphi \circ q\|_{X'} \\ &\leq \sup_{\varphi \in B_{(X/U)'}} \|\varphi\|_{(X/U)'} \|q\|_{(X/U)'} \leq \|q\|_{(X/U)'} = 1.\end{aligned}$$

b) τ ist wohldefiniert: Es ist $\tau(\varphi) \in U^\perp$, denn für $u \in U$ gilt

$$\tau(\varphi)(u) = (\varphi \circ q)(u) = \varphi(u + U) = \varphi(0) \overset{\text{linear}}{=} 0.$$

c) Surjektivität: Sei $x \in X$. Für $\psi \in U^\perp$ setze

$$\varphi(x + U) := \psi(x), \quad \varphi \in (X/U)'.$$

Dann gilt

$$\varphi(x + U) \overset{q \text{ surjektiv}}{=} \varphi(q(x)) = \tau(\varphi)(x) = \psi(x).$$

d) Isometrie: Es gilt für alle $u \in U$

$$|\varphi(x + U)| = |\varphi(q(x))| = |\tau(\varphi)(x)| = |\tau(\varphi)(x) + \underbrace{\tau(\varphi)(u)}_{=0}|$$

$$|\tau(\varphi)(x + u)| \leq \|\tau(\varphi)\|_{X'} \|x + u\|_{(X/U)'}$$

und daher

$$\|\varphi\|_{(X/U)'} \leq \|\tau(\varphi)\|_{X'}.$$

Auf der anderen Seite gilt

$$\|\tau(\varphi)\|_{X'} = \|\varphi \circ q\|_{X'} \leq \|\varphi\|_{(X/U)'} \|q\|_{(X/U)'} = \|\varphi\|_{(X/U)'}.$$

Insgesamt folgt so

$$\|\tau(\varphi)\|_{X'} = \|\varphi\|_{(X/U)'}.$$

Zu 4: Sei dazu $x \in X$ mit $\|x\| < 1$. Dann folgt

$$\|q(x)\| = d(x, U) \leq \|x\| < 1.$$

□

Nach dieser getanen Arbeit nun endlich der entscheidene

Satz 4.16 (Dimensionsformel für kompakte Operatoren). *Seien $A \in K(X)$ und $\lambda \in \mathbb{C} \setminus \{0\}$. Dann gilt*

$$\dim(\ker(\lambda I - A)) = \dim(X/R(\lambda I - A)) = \dim(\ker(\lambda I' - A')) = \dim(X'/R(\lambda I' - A')).$$

Insbesondere sind alle Dimensionen $< \infty$, da nach Lemma 4.6 $\dim(\ker(\lambda I - A)) < \infty$ gilt.

Beweis. Um Schreibarbeit zu sparen, setze

$$\begin{array}{rclrcl} \alpha & := & \dim(\ker(\lambda I - A)), & \beta & := & \dim(X/R(\lambda I - A)), \\ \alpha' & := & \dim(\ker(\lambda I' - A')), & \beta' & := & \dim(X'/R(\lambda I' - A')). \end{array}$$

Wir müssen also zeigen:

$$\alpha = \beta = \alpha' = \beta'.$$

Nach (4.46) mit

$$U := \ker(\lambda I - A)$$

gilt

$$\begin{array}{rcl} [\ker(\lambda I - A)]' & \overset{(4.46)}{\cong} & X'/[\ker(\lambda I - A)]^{\perp} \overset{\text{Satz 3.20}(iv)}{=} X'/\overline{R(\lambda I' - A')} \\ & \overset{\text{Lemma 4.6}}{=} & X'/R(\lambda I' - A'), \end{array}$$

wobei zur Anwendung des Lemmas auch noch der Satz von Schauder (Satz 3.19) eingeht, wonach mit $\lambda I - A$ auch $\lambda I' - A'$ kompakt ist. Es folgt

$$\alpha = \dim \ker(\lambda I - A) = \dim[\ker(\lambda I - A)]' = \dim X'/R(\lambda I' - A') = \beta'.$$

Analog folgt mit (4.47) angewendet auf $U = R(\lambda I - A)$

$$\ker(\lambda I' - A') \overset{\text{Satz 3.20}(i)}{=} R(\lambda I - A)^{\perp} \overset{(4.47)}{\cong} (X/R(\lambda I - A))'.$$

Also

$$\alpha' = \dim(\ker(\lambda I' - A')) = \dim(X/R(\lambda I - A))' = \dim(X/R(\lambda I - A)) = \beta.$$

Wir zeigen jetzt

$$\alpha \leq \beta \tag{4.53}$$

Mit den gleichen Argumenten folgt dann auch $\alpha' \leq \beta'$, da mit A auch A' kompakt ist nach dem Satz von Schauder (Satz 3.19). Somit wäre dann gezeigt

$$\alpha \leq \beta = \alpha' \leq \beta' = \alpha \quad (< \infty)$$

und wir sind fertig.

Nun zum Nachweis von (4.53): Angenommen,

$$\beta < \alpha \quad (< \infty).$$

Dann existiert nach Lemma 3.3 abgeschlossene Untervektorräume $U, V \subset X$ mit

$$X = \ker(\lambda I - A) \oplus U = R(\lambda I - A) \oplus V$$

und $\dim V = \beta < \infty$. Wegen $\beta < \alpha$ gibt es eine stetige, lineare, surjektive Abbildung

$$\phi : \ker(\lambda I - A) \to V,$$

die nicht injektiv ist, d.h. es gibt ein $x_0 \in \ker(\lambda I - A) \setminus \{0\}$ mit

$$\phi(x_0) = 0.$$

Betrachte die stetige, lineare Projektion P auf $\ker(\lambda I - A)$ aus dem Beweis von Lemma 3.3. Es ist

$$\ker(\lambda I - A) = R(P) \quad \text{und} \quad U = \ker(P).$$

Betrachte nun den Operator

$$T := A + \phi \circ P \in K(X).$$

Tatsächlich ist $T \in K(X)$, denn $A \in K(X)$ und $P \in L(X)$, also $\phi \circ P \in F(X) \subset K(X)$. Außerdem ist wegen $x_0 \in \ker(\lambda I - A)$

$$(\lambda I - T)(x_0) = \underbrace{(\lambda I - A)(x_0)}_{=0} - \phi \circ P(x_0) = \phi(x_0) = 0,$$

d.h. $\lambda I - T$ ist nicht injektiv. Nach dem Satz von der Fredholm'schen Alternative (Satz 4.12) ist $\lambda I - T$ auch nicht surjektiv.

Auf der anderen Seite gilt aber für alle $x \in U$ wegen $U = \ker P$

$$Px = 0.$$

Damit folgt

$$(\lambda I - T)(U) = (\lambda I - A)U - \underbrace{(\phi \circ P)(U)}_{=0} = (\lambda I - A)U = R(\lambda I - A),$$

weil $X = \ker(\lambda I - A) \oplus U$. Weiter gilt für $x \in \ker(\lambda I - A)$ nach dem Beweis des Lemmas 3.3

$$Px = x$$

und somit

$$\begin{aligned}(\lambda I - T)(\ker(\lambda I - A)) &= \underbrace{(\lambda I - A)(\ker(\lambda I - A))}_{=0} - \phi \circ P(\ker(\lambda I - A)) \\ &= \phi(\ker(\lambda I - A)) \overset{\phi \text{ surjektiv}}{=} V.\end{aligned}$$

Also ist

$$(\lambda I - T)X = (\lambda I - T)(\ker(\lambda I - A) \oplus U) = R(\lambda I - A) \oplus V = X.$$

Doch das ist ein Widerspruch zur Nicht-Surjektivität von $\lambda I - T$. Also gilt (4.53). $\square$

5 Hilberträume, Orthonormalbasen und kompakte Operatoren

- Hilberträume und ihre Eigenschaften: (Stetiges) Skalarprodukt, Prähilbertraum, induzierte Norm, Parallelogrammgleichung und Polarisation, Dreiecksungleichung, Phytagoras, Cauchy-Schwarzsche Ungleichung
- Projektionssatz
- Orthogonalprojektion auf abgeschlossene Untervektorräume
- Darstellungssatz von Riesz-Fréchet
- Jeder Hilbertraum ist reflexiv
- Orthonormalsysteme- und Basen: Gram-Schmidt Orthonormalisierungsverfahren, Besselsche Ungleichung, Parsevalsche Gleichung, Fourierentwicklung sowie unbedingte und absolute Konvergenz
- Jeder separable Hilbertraum besitzt eine Orthonormalbasis
- Hilbert-Adjungierte, unitäre, selbstadjungierte und normale Operatoren
- Spektralzerlegungssatz für komnpakte, normale Operatoren
- Wurzeln aus Operatoren

5.1 Hilberträume

Definition 5.1 (Skalarprodukt, Prähilbertraum, orthogonal). Sei X ein $\mathbb{K}$-Vektorraum. Eine Abbildung

$$(.,.) : X \times X \to \mathbb{K}$$

heißt *Skalarprodukt* (oder *inneres Produkt*), falls

(i) Linearität: Für alle $x, y, z \in X$ und $\lambda \in \mathbb{K}$ gilt

$$(x + y, z) = (x, z) + (y, z), \quad (\lambda x, y) = \lambda(x, y).$$

(ii) Symmetrie: $(x, y) = \overline{(y, x)}$ für alle $x, y \in X$.

(iii) Positive Definitheit: $(x, x) > 0$ für alle $x \neq 0$.

Das Paar $(X, (.,.))$ nennt man dann einen *Prähilbertraum*.

Zwei Vektoren $x, y \in X$ heißen *orthogonal*, falls $(x, y) = 0$. Notation: $x \perp y$.

Wenn $U, V \subset X$, dann schreib man

$$U \perp V \Longrightarrow u \perp v \quad \text{für alle } u \in U \text{ und } v \in V.$$

Man bezeichnet

$$U^\perp := \{\, x \in X \mid x \perp u \text{ für alle } u \in U \,\}$$

als den *Orthogonalraum von* U.

Auf natürliche Weise definiert für $x \in X$

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

eine Norm auf X.

Die ersten beiden Eigenschaften einer Norm sind leicht einzusehen. Die Dreiecksungleichung zeigen wir in Satz 5.1.

Wenn $(X, |.|)$ vollständig ist, dann heißt X *Hilbertraum.*

Beispiel 5.1. Aus der linearen Algebra sind die Hilberträume $\mathbb{K}^N$ mit dem Skalarprodukt

$$(x, y) := \sum_{i=1}^{N} x_i \bar{y}_i, \quad x, y \in \mathbb{K}^N$$

bekannt.

□

Beispiel 5.2. Sei $X := C([a, b], \mathbb{K})$ und sei

$$(f, g) := \int_a^b f(t)\bar{g}(t)dt, \quad f, g \in X.$$

Dann ist $(X, (., .))$ ein Prähilbertraum, aber kein Hilbertraum unter der natürlichen assozierten Norm

$$\|f\|_2 = \sqrt{\int_a^b |f(t)|^2 \, dt}.$$

Um das einzusehen, sei $a = -1$ und $b = 1$. Definiere $f_n : [-1, 1] \to \mathbb{R}$ durch

$$f_n(t) := \begin{cases} 0, & -1 \leq t \leq -\frac{1}{n} \\ \frac{1}{2}(nt + 1), & -\frac{1}{n} \leq t \leq \frac{1}{n} \\ 1, & \frac{1}{n} \leq t \leq 1. \end{cases}$$

Dann ist $f_n \in C([-1, 1], \mathbb{R})$ sowie stückweise affin und monton, wobei die Werte von f_n zwischen 0 und 1 liegen.

Nun gilt für $n \leq m$

$$f_n(t) - f_m(t) = 0, \quad \text{wenn } |t| > \tfrac{1}{n}.$$

Folglich ist

$$\begin{aligned} \|f_n - f_m\|^2 &= \int_{-1}^{1} (f_n - f_m)^2(t)dt = \int_{-1/n}^{1/n} (f_n - f_m)^2(t)dt \\ &= \int_{-1/n}^{1/n} f_n^2(t)dt + \int_{-1/n}^{1/n} f_m^2(t)dt - 2\int_{-1/n}^{1/n} f_n(t)f_m(t)dt \end{aligned}$$

Da $f_n, f_m \geq 0$ und da

$$\int_{-1/n}^{1/n} f_m(t)dt, \int_{-1/n}^{1/n} f_n(t)dt \leq \int_{-1/n}^{1/n} \tfrac{1}{2}dt = \tfrac{1}{n}$$

ist

$$\|f_n - f_m\|^2 \leq \tfrac{4}{n}, \quad \text{wenn } m \geq n.$$

Daher bilden die f_n eine Cauchy-Folge in $(C[-1, 1], \mathbb{R}, (., .))$. Wenn es nun eine stetige(!) Funktion $f : [-1, 1] \to \mathbb{R}$ gäbe, so dass

$$\lim_{n\to\infty} \|f - f_n\| = 0,$$

so müsste wegen der Eigenschaft von f_n der Ausdruck

$$\int_{-1}^{-1/n} f^2(t)dt + \int_{-1/n}^{1/n} (f - f_n)^2(t)dt + \int_{1/n}^{1} (f - 1)^2(t)dt$$

für $n \to \infty$ gegen 0 gehen. Weil alle Integranden positiv sind, folgt bei gegebenem $0 < \epsilon < 1$ und für $n \geq \frac{1}{\epsilon}$, dass

$$\int_{-1}^{-\epsilon} f^2(t)dt \leq \int_{-1}^{-1/n} f^2(t)dt, \quad \int_{\epsilon}^{1} (f-1)^2(t)dt \leq \int_{1/n}^{1} (f-1)^2(t)dt.$$

Die rechten Seiten dieser Ungleichungen streben aber gegen 0. Folglich gilt für alle $0 < \epsilon < 1$, dass

$$\int_{1}^{-\epsilon} f^2(t)dt = \int_{1}^{\epsilon} (f-1)^2(t)dt = 0.$$

Daher ist (in den Räumen $C([-1,-\epsilon],\mathbb{R})$ bzw. $C([\epsilon,1],\mathbb{R})$

$$f|_{[-1,-\epsilon]} = 0 \quad \text{und} \quad (1-f)|_{[\epsilon,1]} = 0.$$

Für $\epsilon \to 0$ folgt dann aus der angenommenen Stetigkeit von f der Widerspruch

$$0 = f(0) \quad \text{aber} \quad 1 - f(0) = 0.$$

□

Beispiel 5.3. Sei $X := L^2(]a,b[,\mathbb{K})$ und wie in Beispiel 5.2

$$(f,g) := \int_a^b f(t)\bar{g}(t)dt, \quad f,g \in X.$$

Dann ist $(L^2(a,b), \|.\|_2)$ ein Hilbertraum. Die Vollständigkeit folgt aus dem Satz von Fischer-Riesz (vgl. Analysis 3).

□

Beispiel 5.4. Sei $X := l^2$ ausgestattet mit dem Skalarprodukt

$$(x,y) = \sum_{i=1}^{\infty} x_i \bar{y}_i, \quad x,y \in X. \tag{5.1}$$

Dann ist X ausgestattet mit der Norm

$$\|x\|_{l^2} = \sqrt{\sum_{i=1}^{\infty} |x_i|^2}, \quad x \in X$$

ein Hilbertraum.

Beachte, dass die Konvergenz der Reihe (5.1) aus der Hölderschen Ungleichung folgt.

□

Satz 5.1. *Seien $(X,(.,.))$ ein Prähilbertraum und $x,y \in X$. Dann gilt:*

(i) Parallelogrammgleichung:

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

(ii) Cauchy-Schwarzsche Ungleichung:

$$|(x,y)| \leq |x|\,|y|.$$

(iii) Dreiecksungleichung:

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

(iv) Satz von Phytagoras:

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

(v) Es gilt für alle $\lambda \in \mathbb{K}$

$$x \perp y \Longleftrightarrow |y| \leq |\lambda x + y|.$$

Beweis. Zu (i): Nachrechnen ergibt

$$\begin{aligned}|x+y|^2+|x-y|^2 &= (x+y,x+y)+(x-y,x-y)\\ &= (x,x)+(x,y)+(y,x)+(y,y)+(x,y)-(x,y)-(y,x)+(y,y)\\ &= 2\,|x|^2+|y|^2\,.\end{aligned}$$

Zu (ii) bis (v): Es gilt für $\lambda \in \mathbb{K}$

$$\begin{aligned}0 \le |\lambda x+y|^2 = (\lambda x+y,\lambda x+y) &= |\lambda|^2\,|x|^2+\lambda(x,y)+\bar{\lambda}(y,x)+|y|^2\\ &= |\lambda|^2\,|x|^2+\lambda(x,y)+\bar{\lambda}\overline{(x,y)}+|y|^2\,.\end{aligned}$$

Also ist

$$0 \le |\lambda x+y|^2 = |\lambda|^2\,|x|^2+2\,\mathrm{Re}(\lambda(x,y))+|y|^2\,. \tag{5.2}$$

Mit $\lambda = 1$ folgt aus (5.2) sofort (iv).

Bemerke, dass für $x = 0$ sämtliche Behauptungen trivial sind. Sei also im Folgenden $x \neq 0$. Setze dann

$$\lambda := -\frac{(y,x)}{|x|^2}$$

Einsetzen in (5.2) ergibt

$$0 \le \frac{|(x,y)|^2}{|x|^2} - 2\frac{|(x,y)|^2}{|x|^2} + |y|^2 = |y|^2 - \frac{|(x,y)|^2}{|x|^2} \tag{5.3}$$

Damit folgt dann

$$\frac{|(x,y)|^2}{|x|^2} \le |y|^2 \Longrightarrow |(x,y)|^2 \le |x|^2\,|y|^2$$

und damit ist (ii) gezeigt.

Wir zeigen nun die Rückrichtung von (v): Angenommen, $(x,y) \neq 0$. Dann wähle das oben definierte λ, so dass mit (5.3)

$$|\lambda x+y|^2 = |y|^2 - \frac{|(x,y)|^2}{|x|^2} < |y|^2\,.$$

Das ist aber ein Widerspruch zu $|\lambda x+y|^2 \ge |y|^2$. Also gilt $(x,y) = 0$.

Aus (5.2) folgt aber auch die Hinrichtung von (v), denn aus $(x,y) = 0$ folgt

$$|\lambda x+y|^2 = |\lambda|^2\,|x|^2+|y|^2 \ge |y|^2\,.$$

(iii) folgt schließlich aus (5.2), denn für $\lambda = 1$ ist

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

□

Bemerkung. Die Paralellogrammgleichung charakterisiert genau die Normen, die von einem Skalarprodukt induziert werden, d.h. ist $\|.\|$ eine Norm auf einem Vektorraum X, dann existiert ein Skalarprodukt $((.,.))$ auf $X \times X$ mit

$$\|x\| = ((x,x))^{1/2}, \quad x \in X$$

genau dann, wenn $\|.\|$ die Parallelogrammgleichung erfüllt. Anders gesagt: Ein normierter Raum X ist genau dann ein Prähilbertraum, wenn die Paralelogrammgleichung gilt.

Für einen Beweis brauchen wir die Tatsache, dass in einem Prähilbertraum nicht nur die Norm durch das Skalarprodukt ausgedrückt werden kann, sondern auch das Skalarprodukt durch dir Norm:

1. Für $\mathbb{K} = \mathbb{R}$ gilt

$$(x, y) = \tfrac{1}{4}(|x + y|^2 - |x - y|^2). \tag{5.4}$$

2. Für $\mathbb{K} = \mathbb{C}$ gilt

$$(x, y) = \tfrac{1}{4}(|x + y|^2 - |x - y|^2 + i\,|x + iy|^2 - i\,|x - iy|^2). \tag{5.5}$$

Diese *Polarisationsformeln* zeigt man durch leichtes Nachrechnen.

Beweis der Bemerkung. Wir behandeln nur den Fall $\mathbb{K} = \mathbb{R}$. Für $\mathbb{K} = \mathbb{C}$ geht das genauso.

Die Richtung „$\Longrightarrow$“ haben wir schon mit Satz 5.1 gezeigt. Um „$\Longleftarrow$“ zu zeigen, setze wie in (5.4)

$$(x, y) := \tfrac{1}{4}(|x + y|^2 - |x - y|^2).$$

<u>Zu (i):</u> Für $x_1, x_2, y \in X$ gilt nach Parallelogrammgleichung

$$\begin{aligned} |x_1 + x_2 + y|^2 &= 2\,|x_1 + y|^2 + |x_2| - |x_1 - x_2 + y|^2 =: \alpha, \\ |x_1 + x_2 + y|^2 &= 2\,|x_2 + y|^2 + |x_1| - |-x_1 + x_2 + y|^2 =: \beta. \end{aligned}$$

Folglich ist

$$\begin{aligned} |x_1 + x_2 + y|^2 &= \frac{\alpha + \beta}{2} \\ &= |x_1 + y|^2 + |x_1|^2 + |x_2 + y|^2 + |x_2|^2 - \tfrac{1}{2}(|x_1 - x_2 + y| + |-x_1 + x_2 + y|^2). \end{aligned}$$

Analog ergibt sich

$$|x_1 + x_2 - y|^2 = |x_1 - y|^2 + |x_1|^2 + |x_2 - y|^2 + |x_2|^2 - \tfrac{1}{2}(|x_1 - x_2 - y| + |-x_1 + x_2 - y|^2).$$

Es folgt

$$\begin{aligned} (x_1 + x_2, y) &= \tfrac{1}{4}(|x_1 + x_2 + y|^2 - |x_1 + x_2 - y|^2) \\ &= \tfrac{1}{4}(|x_1 + y|^2 + |x_2 + y|^2 - |x_1 - y|^2 - |x_2 - y|^2) \\ &= (x_1, y) + (x_2, y). \end{aligned}$$

Zeige nun

$$(\lambda x, y) = \lambda(x, y). \tag{5.6}$$

Eben haben wir schon gesehen, dass dies für $\lambda \in \mathbb{N}$ gilt. Nach konstruktion gilt (5.6) auch für $\lambda = 0$ und $\lambda = -1$, also auch für $\lambda \in \mathbb{Z}$. Weiter gilt (5.6) auch für $\lambda = \frac{m}{n} \in \mathbb{Q}$, denn

$$n(\lambda x, y) = n(m\tfrac{x}{n}, y) = m(x, y) = n\lambda(x, y).$$

Die (wegen der Stetigkeit von $|.|$ nach Lemma 5.1) stetigen Funktionen (von $\mathbb{R}$ nach $\mathbb{R}$)

$$\lambda \mapsto (\lambda x, y), \quad \lambda \to \lambda(x, y)$$

stimmen auf $\mathbb{Q}$ daher überein und sind deshalb auf $\mathbb{R}$ gleich.

<u>Zu (ii):</u> Das ist klar.

<u>Zu (iii):</u> Das folgt aus $(x, x) = |x|^2$. □

Lemma 5.1. *Das Skalarprodukt eines Prähilbertraums X ist eine stetige Abbildung von $X \times X$ nach $\mathbb{K}$.*

Beweis. Seien $(x_n)_n, (y_n)_n \subset X$ Folgen mit $x_n \to x$ und $y_n \to y$. Dann folgt

$$|(x_n, y_n) - (x, y)| = |(x_n - x, y_n) + (x, y_n - y)| \overset{\text{CSU}}{\leq} |x_n - x|\,|y_n| + |x|\,|y_n - y| \to 0.$$

□

Bemerkungen.

1. Genauso zeigt man, dass auf einem Prähilbertraum die Abbildungen $x \mapsto (x, y)$ und $y \mapsto (x, y)$ stetig sind.
2. Ist $(X, |.|)$ ein Prähilbertraum und $\|.\|$ ein äquivalente Norm auf X, so braucht $(X, \|.\|)$ kein Prähilbertraum sein.
3. Ist X ein Prähilbertraum, $U \subset X$ ein dichter Unterraum und $x \in X$ mit $(x, u) = 0$ für alle $u \in U$, so ist $x = 0$.

   Beweis. Sei $x \in X$ beliebig. Betrachte die Menge
   $$Y := \{\, y \in X \mid (x, y) = 0 \,\}.$$
   Dann ist Y aufgrund der Stetigkeit von $y \mapsto (x, y)$ abgeschlossen und es gilt $U \subset Y \subset X$. Daraus folgt
   $$X = \bar{U} \supset \bar{Y} = Y \subset X \Longrightarrow Y = X.$$
   Speziell ist also $x \in Y$ und das heißt
   $$|x|^2 = (x, x) = 0 \Longrightarrow x = 0.$$

   □

4. Unterräume von Prähilberträumen sind Prähilberträume. Das folgt, indem man das Skalarprodukt einschränkt.
5. Sei $\tilde{X}$ die Vervollständigung des Prähilbertraumes X. Da die Parallelogrammgleichung in X gilt, überträgt sie sich aus Stetigkeitsgründen auf $\tilde{X}$. Mit der oberen Bemerkung folgt dann, dass $\tilde{X}$ ein Hilbertraum ist.

Bevor wir zum großen Projektionssatz kommen, noch einige nützliche Eigenschaften:

1. $A^\perp$ ist stets ein abgeschlossener Unterraum von X.
2. $A \subset (A^\perp)^\perp$.
3. $A^\perp = (\overline{\text{lin}} A)^\perp$.

Beweis. <u>Zu 1:</u> Die Unterraum-Eigenschaft ist klar. Um die Abgeschlossenheit zu zeigen, sei $(a_n)_n \subset A^\perp$ eine Folge mit $a_n \to a \in X$. Dann folgt wegen der Steteigkeit von $x \mapsto (x, y)$

$$0 = (a_n, x) \to (a, x) = 0 \quad \text{für alle } x \in A,$$

also $a \in A^\perp$.

<u>Zu 2:</u> Sei also $a \in A$. Dann gilt für alle $x \in A^\perp$ gerade $(x, a) = 0$, also

$$a \in (A^\perp)^\perp = \{\, y \in X \mid (y, x) = 0 \quad \text{für alle } x \in A^\perp \,\}.$$

<u>Zu 3:</u> Um $\supset$ zu zeigen, sei $x \in (\overline{\text{lin}} A)^\perp$, d.h. $(x, \tilde{a}) = 0$ für alle $\tilde{a} \in \overline{\text{lin}} A$. Dann gilt für alle $a \in A \subset \overline{\text{lin}}(A)$ erst recht

$$(x, a) = 0,$$

also $x \in A^\perp$.

Jetzt zur Implikation $\subset$: Sei dazu $x \in A^\perp$. Sei $a \in \overline{\text{lin}\, A}$. Dann gibt es eine Folge $(a_n)_n \subset \text{lin}\, A$, so dass $|a_n - a| \to 0$. Es folgt

$$(x, a) = \lim(x, a_n) = \lim(x, \sum_{i=1}^m \lambda_{n_i} a_{n_i}) = \lim \sum_{i=1}^m \bar{\lambda}_{n_i} (x, a_{n_i}) = 0.$$

□

Satz 5.2 (Projektionssatz). *Sei $(H, (.,.))$ ein Hilbertraum und $K \subset H$ eine nicht-leere abgeschlossene konvexe Menge. Dann gilt*

$$\forall_{f \in H} \exists!_{u \in K} : |f - u| = d(f, K) = \inf\{\, |f - v| \mid v \in K \,\}.$$

Das Element u ist eindeutig charakterisiert durch die Familie von Ungleichungen

$$\mathrm{Re}(f - u, v - u) \leq 0 \quad \text{für alle } v \in K.$$

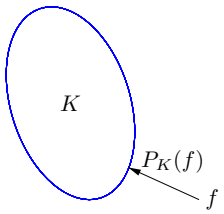
Das Element u bezeichnen wir mit $P_k(f)$ und meinen damit die „Projektion von f auf K“.

Bemerkung. Achtung: Die Abbildung

$$P_K : H \to K, \quad f \mapsto P_K(f)$$

ist im Allgemeinen nicht linear. Es gilt aber stets

$$P_k \circ P_K = P_K.$$



Beweis. <u>Zur Existenz:</u> Definiere

$$d := \inf\{\, |f - v| \mid v \in K \,\}.$$

Weiter sei $(v_n) \subset K$ eine Folge mit

$$d_n := |f - v_n| \to d.$$

Setze

$$x := \frac{f - v_n}{2}, \quad y = \frac{f - v_m}{2}, \quad m, n \in \mathbb{N}$$

Mit der Parallelogrammgleichung folgt dann

$$|f - \tfrac{v_n + v_m}{2}|^2 + \left|\tfrac{v_n - v_m}{2}\right|^2 = \tfrac{1}{2} d_n^2 + \tfrac{1}{2} d_m^2.$$

Bemerke, dass $\frac{v_n + v_m}{2} \in K$, da K konvex ist. Also folgt

$$\left|f - \tfrac{v_n + v_m}{2}\right|^2 \geq d^2$$

und daraus

$$\left|\tfrac{v_n - v_m}{2}\right|^2 \leq \tfrac{d_n^2 + d_m^2}{2} - d^2 \overset{m,n \to \infty}{\longrightarrow} 0.$$

Somit ist $(v_n)_n$ eine Cauchy-Folge in H. Nun ist H vollständig, d.h. es gibt ein $u \in H$ mit

$$v_n \to u \quad \text{in } H.$$

Da K abgeschlossen ist, folgt $u \in K$. Außerdem gilt

$$d = \lim_{n \to \infty} d_n = \lim_{n \to \infty} |f - v_n| = |f - u|.$$

<u>Zur Familie der Ungleichungen:</u> Zeige nun, dass u die Familie der Ungleichungen erfüllt. Sei dazu $w \in K$ und $t \in]0, 1[$. Da K konvex ist, gilt

$$v = (1 - t)u + tw \in K.$$

Daher gilt

$$|f - u| \leq |f - [(1 - t)u + tw]| = |(f - u) - t(w - u)|.$$

Daraus folgt

$$|f - u|^2 \leq |f - u|^2 - 2t\,\mathrm{Re}(f - u, w - u) + t^2 |w - u|^2.$$

Umstellen und teilen durch t liefert

$$0 \leq -2\operatorname{Re}(f-u, w-u) + t\,|w-u|^2 .$$

Also haben wir für $t \to 0^+$

$$0 \leq -2\operatorname{Re}(f-u, w-u) \Longrightarrow \operatorname{Re}(f-u, w-u) \leq 0.$$

Das zeigt „$\Longrightarrow$“.

Um die umgekehrte Richtung zu zeigen, sei nun $z \in K$ und es gelte für alle $v \in K$

$$\operatorname{Re}(f-z, v-z) \leq 0.$$

Dann folgt aber

$$\begin{aligned}
|f-z|^2 - |f-v|^2 &= |f-z|^2 - |f-z+z-v|^2 \\
&= |f-z|^2 - ((f-z)+(z-v), (f-z)+(z-v)) \\
&= |f-z|^2 - |f-z|^2 + 2\operatorname{Re}(f-z, v-z) - |v-z|^2 \\
&= -|v-z|^2 + 2\underbrace{\operatorname{Re}(f-z, v-z)}_{\leq 0} \leq 0.
\end{aligned}$$

Somit gilt für alle $v \in K$

$$|f-z| \leq |v-f| .$$

<u>Zur Eindeutigkeit:</u> Dazu wähle $u_1, u_2 \in K$, so dass diese die Familie der Ungleichung erfüllen. Dann ist

$$\operatorname{Re}(f-u_1, v-u_1) \leq 0 \quad \text{für alle } v \in K.$$

Also speziell auch für $v = u_2$:

$$\operatorname{Re}(f-u_1, u_2-u_1) \leq 0. \tag{5.7}$$

Analog gilt

$$\operatorname{Re}(f-u_2, u_1-u_2) = \operatorname{Re}(u_2-f, u_2-u_1) \leq 0. \tag{5.8}$$

Nun berechne (5.7) + (5.8):

$$\operatorname{Re}(u_2-u_1, u_2-u_1) \leq 0,$$

d.h.

$$|u_2-u_1|^2 \leq 0 \Longrightarrow u_2 = u_1.$$

□

Bemerkung. P_K ist Lipschitz-stetig mit Lipschitz-Konstante $L = 1$.

Beweis. Seien $f_1, f_2 \in H$. Weiter setze

$$u_1 := P_K(f_1) \quad \text{und} \quad u_2 := P_K(f_2).$$

Dann gilt für alle $v \in K$

$$\operatorname{Re}(f_1-u_1, v-u_1) \leq 0.$$

Speziell gilt wieder, da $u_2 \in K$

$$\operatorname{Re}(f_1-u_1, u_2-u_1) \leq 0$$

bzw. analog

$$\operatorname{Re}(f_2-u_2, u_1-u_2) \leq 0.$$

Addition ergibt dann

$$\operatorname{Re}(f_1-f_2+(u_2-u_1), u_2-u_1) \leq 0.$$

Daraus folgt

$$|u_2-u_1|^2 \leq \operatorname{Re}(f_1-f_2, u_1-u_2) \overset{\text{CSU}}{\leq} |f_1-f_2|\,|u_1-u_2| .$$

Also folgt schließlich, (da $u_1 \neq u_2$ wegen der Eindeutigkeit)

$$|u_1-u_2| \leq 1 \cdot |f_1-f_2| .$$

□

Satz 5.3 (**Orthogonalprojektion auf abgeschlossene Untervektorräume**). *Seien* $(H, (.,.))$ *ein Hilbertraum,* $\{0\} \subsetneq U \subseteq H$ *ein abgeschlossener Untervektorraum sowei* P_U *die Bestapproximation auf* U *aus Satz 5.2. Dann gilt*

$$P_U \in L(H), \quad \|P_U\|_{L(H)} = 1, \quad R(P_U) = U,$$
$$\ker(P_U) = U^\perp, \quad H = U \oplus U^\perp.$$

Das Element $P_U(f)$, $f \in H$, *ist eindeutig charakterisiert durch*

$$P_U(f) \in U, \quad (f - P_U(f), w) = 0 \quad \textit{für alle } w \in U. \tag{5.9}$$

d.h. $P_U(f)$ *ist das eindeutige Element aus* U, *so dass*

$$f - P_U(f) \in U^\perp.$$

P_U *wird in diesem Fall* Orthogonalprojektion auf U *genannt.*

Beweis. Nach Satz 5.2 ist $P_U(f)$, $f \in H$, charakterisiert durch

$$\mathrm{Re}(f - P_U(f), v - P_U(f)) \leq 0 \quad \text{für alle } v \in U.$$

Das ist gleichbedeutend mit

$$\mathrm{Re}(f - P_U(f), w) \leq 0 \text{ für alle } w \in U.$$

Das gilt aber auch für alle $-w \in U$, so dass

$$\mathrm{Re}(f - P_U(f), w) \geq 0 \text{ für alle } w \in U,$$

also folgt

$$\mathrm{Re}(f - P_U(f), w) = 0 \text{ für alle } w \in U, \tag{5.10}$$

Ist $U \subset \mathbb{C}$, so gilt (5.10) auch für ein rein imaginäres w, also gilt insgesamt

$$(f - P_U(f), w) = 0 \quad \text{für alle } w \in U$$

und damit ist (5.9) gezeigt.

a) <u>Linearität:</u> Seien $f_1, f_2 \in H$ und $\lambda_1, \lambda_2 \in \mathbb{K}$. Dann gilt, weil $U^\perp$ ein Untervektorraum ist

$$\lambda_1(f_1 - P_U(f_1)) + \lambda_2(f_2 - P_U(f_2)) \in U^\perp.$$

Das ist gleichbedeutend mit

$$(\lambda_1 f_1 + \lambda_2 f_2) - (\lambda_1 P_U(f_1) + \lambda_2 P_U(f_2)) \in U^\perp \tag{5.11}$$

Wegen der Charakterisierung der Bestapproximation ist aber

$$w := P_U(\lambda_1 f_1 + \lambda_2 f_2)$$

das eindeutige Element $w \in U$ mit

$$\lambda_1 f_1 + \lambda_2 f_2 - w \in U^\perp.$$

Mit (5.11) folgt damit

$$(\lambda_1 f_1 + \lambda_2 f_2) - (\lambda_1 P_U(f_1) + \lambda_2 P_U(f_2)) = (\lambda_1 f_1 + \lambda_2 f_2) - w,$$

also

$$\lambda_1 P_U(f_1) + \lambda_2 P_U(f_2) = P_U(\lambda_1 f_1 + \lambda_2 f_2),$$

d.h. P_U ist linear.

b) Stetigkeit und Norm: Nach der obigen Bemerkung ist P_U Lipschitz-stetig mit Konstante 1, d.h.

$$\|P_U\| \leq 1.$$

Für Projektionen gilt stets $\|P_U\| \geq 1$. Damit ist die Stetigkeit und $\|P_U\| = 1$ gezeigt. Es ist also $P_U \in L(H)$.

c) $R(P_U) = U$: Das ist trivial.

d) $\ker(P_U) = U^\perp$: Es gilt in der Tat

$$f \in \ker(P_U) \Longleftrightarrow P_U(f) = 0 \Longleftrightarrow f - \underbrace{0}_{=P_U(f)} \in U^\perp.$$

e) $X = U \oplus_{top} U^\perp$:

* Zunächst ist $U \cap U^\perp = \{0\}$: Klar dabei ist $\{0\} \subset U \cap U^\perp$. Zeige jetzt $\{0\} \supset U \cap U^\perp$. Angenommen, es gibt ein $u \in U \cap U^\perp$ mit $u \neq 0$. Dann gilt

$$u \in U^\perp \Longleftrightarrow (u, \tilde{u}) = 0 \quad \text{für alle } \tilde{u} \in U.$$

Also insbesondere auch für $\tilde{u} = u$:

$$(u, u) = 0 \Longrightarrow u = 0.$$

Widerspruch!

* Für $x \in H$ gilt

$$x = P_U x + (I - P_U)x,$$

denn $P_U \in U$ und $(I - P_U) \in \ker(P_U) = U^\perp$, da $P_U - P_U \circ P_U = 0$.

Die Zerlegung

$$X = U \oplus U^\perp$$

ist eine topologische direkte Zerlegung in dem Sinne, dass

$$x_n = u_n + v_n \to x = u + v \Longleftrightarrow \begin{cases} u_n \to u & \text{in } U \\ v_n \to v & \text{in } U^\perp. \end{cases}$$

Beachte, dass U und $U^\perp$ abgeschlossen(!) sind.

□

Bemerkung. Ist X ein Banachraum, U ein abgeschlossener Untervektorraum und existiert eine stetige, lineare Projektion P auf U, so gilt auch in diesem Fall

$$X = U \oplus_{top} \ker(P).$$

Im Allgemeinen existiert eine solche Projektion nicht in jedem abgeschlossenen Untervektorraum U eines Banachraumes X. Man sagt in diesem Fall

$$U \text{ ist nicht projezierbar.}$$

Lindenstrauss-Tzafiri haben gezeigt, dass ein Banachraum X, in dem jeder abgeschlossene Untervektorraum projezierbar ist, eine äquivalente Norm besitzt, die von einem Skalarprodukt induziert wird.

Korollar 5.1. *Sei U ein Unterraum des Hilbertraumes H. Dann gilt*

$$\bar{U} = (U^\perp)^\perp.$$

Beweis. Sei V ein abgeschlossener Unterraum von H. Aus Satz 5.3 folgt zunächst

$$I - P_V = P_{V^\perp}.$$

Setze $V := \bar{U}$. Dann folgt

$$U^\perp = V^\perp \quad \text{und} \quad I - P_{V^\perp} = P_{V^{\perp\perp}},$$

also

$$P_V = P_{V^{\perp\perp}} \Longrightarrow \bar{U} = U^{\perp\perp}.$$

□

Satz 5.4 (**Darstellungssatz von Riesz-Fréchet**). *Sei $(H, (.,.))$ ein Hilbertraum. Dann gilt*

$$\forall_{\varphi \in H'} \exists!_{f \in H} : \varphi(v) = (v, f) \quad \textit{für alle } v \in H$$

und es gilt

$$|f|_H = |\varphi|_{H'}.$$

Die Abbildung

$$\phi : H \to H', \quad f \mapsto \phi(f) = \phi_f = (., f)$$

ist ein konjugiert linearer isometrischer Isomorphismus.

Beweis.

a) <u>ϕ ist wohldefiniert:</u> In der Tat ist $(., f)$ linear und stetig:

$$|\phi_f(x)| = |(x, f)| \overset{CSU}{\leq} |x|\,|f| \Longrightarrow |\phi_f|_{H'} \leq |f|_H. \tag{5.12}$$

b) <u>ϕ ist konjugiert linear:</u> Für alle $f, f_1, f_2 \in H$ und alle $\lambda \in \mathbb{K}$ gilt

$$\phi(f_1 + f_2) = \phi(f_1) + \phi(f_2), \quad \phi(\lambda f) = \bar{\lambda}\phi(f).$$

c) <u>Isometrie:</u> Nach (5.12) gilt

$$|\phi_f|_{H'} \leq |f|_H.$$

Auf der anderen Seite definiere für $f \neq 0$

$$g := \frac{f}{|f|}.$$

Dann ist $|g| = 1$ und

$$\phi_f(g) = (g, f) = \frac{1}{|f|}(f, f) = |f|,$$

also insgesamt

$$|\phi_f|_{H'} = |f|_H.$$

d) <u>Surjektivität:</u> Sei $\varphi \in H'$. Wenn $\varphi = 0$, dann ist $f = 0$ der gesuchte Kandidat, denn

$$\phi(0) = 0 = \varphi.$$

Sei also $\varphi \neq 0$. Dann ist

$$U := \ker(\varphi)$$

ein echter, abgeschlossener Untervektorraum von H (denn φ stetig) und somit

$$U^\perp \neq \{0\}.$$

Sei dann $w \in U^\perp$ mit $\varphi(w) = 1$, also insbesondere $w \neq 0$. Wir werden zeigen, dass

$$\varphi = \phi_f = \phi(f) \quad \text{mit } f = \frac{w}{|w|^2}. \tag{5.13}$$

Sei dazu $v \in H$. Dann ist

$$v - \varphi(v) \cdot w \in \ker(\varphi) = U,$$

denn

$$\varphi(v - \varphi(v) \cdot \varphi(w)) = \varphi(v) - \varphi(v) \underbrace{\varphi(w)}_{=1} = 0.$$

Da $w \in U^\perp$, folgt somit

$$(v - \varphi(v) \cdot w, w) = 0 \Longleftrightarrow (v, w) - \varphi(v) |w|^2 = 0,$$

also

$$\varphi(v) = \frac{1}{|w|^2}(v, w) = (v, f) = \phi_f(v) \quad \text{für alle } v \in H.$$

Damit ist (5.13) gezeigt.

□

Bemerkung. Die Räume H und H' können mittels ϕ miteinander identifiziert werden.

Korollar 5.2. *Jeder Hilbertraum ist reflexiv.*

Beweis.

- Zunächst ist H' ein Hilbertraum unter dem Skalarprodukt

  $$\langle \phi(x), \phi(y) \rangle := (y, x) \quad \text{für alle } x, y \in H.$$

  Zu (i): Für $x, \tilde{x}, y \in H$ gilt

  $$\begin{aligned} \langle \phi(x) + \phi(\tilde{x}), \phi(y) \rangle &= \langle \phi(x + \tilde{x}), \phi(y) \rangle = (y, x + \tilde{x}) = (y, x) + (y, \tilde{x}) \\ &= \langle \phi(x), \phi(y) \rangle + \langle \phi(\tilde{x}, \phi(y) \rangle . \end{aligned}$$

  Weiter gilt für $\lambda \in \mathbb{K}$

  $$\langle \lambda \phi(x), \phi(y) \rangle = \left\langle \phi(\bar{\lambda} x), \phi(y) \right\rangle = (y, \bar{\lambda} x) = \lambda (y, x) = \lambda \langle \phi(x), \phi(y) \rangle .$$

  Zu (ii): Für $x, y \in H$ gilt

  $$\overline{\langle \phi(x), \phi(y) \rangle} = \overline{(y, x)} = (x, y) = \langle \phi(y), \phi(x) \rangle .$$

  Zu (iii): Für alle $x \neq 0$ gilt

  $$\langle \phi(x), \phi(x) \rangle = (x, x) > 0.$$

  Zu (iv): Dass $(H', (., .))$ vollständig ist, ist klar.

- Es gilt

  $$|\varphi|_{H'}^2 \overset{\text{Satz 5.4}}{=} |\phi^{-1}(\varphi)|_H^2 = (\phi^{-1}(\varphi), \phi^{-1}(\varphi)) = \langle \varphi, \varphi \rangle \quad \text{für alle } \varphi \in H',$$

  d.h. die duale Norm $|.|_{H'}$ wird von $\langle ., . \rangle$ induziert.

- Also kann nach Satz 5.4 H' mittels des kanonischen konjugierten linearen iosometrischen Isomorphismus ϕ' mit H'' identifiziert werden:

$$\begin{array}{ccc} H & & \\ \phi \downarrow & \overset{j_H}{\searrow} & \\ H' & \overset{\phi'}{\to} & H'' \end{array} \qquad \text{Kanonische Abbildung: } j_H = \phi' \circ \phi.$$

Zunächst ist klar, dass j_H tatsächlich die kanonische Abbildung ist. Wir müssen noch zeigen, dass $j_H : H \to H''$ surjektiv ist, d.h.

$$\forall_{\psi \in H''} \exists_{x_\psi \in H} : \psi(\varphi) = j_H(x_\psi)(\varphi) = \varphi(x_\psi) \quad \text{für alle } \varphi \in H'.$$

Sei also $\psi \in H''$. Dann existiert nach Satz 5.4 ein $x_\psi \in H$ mit

$$\psi = \phi'(\phi(x_\psi)).$$

Für alle $\varphi \in H'$ gilt dann zunächst

$$\psi(\varphi) = \phi'(\phi(x_\psi))(\varphi) = \langle \varphi, \phi(x_\psi) \rangle .$$

Nun gibt es nach Satz 5.4 zu φ ein x_φ, so dass $\varphi = \phi(x_\varphi)$ und somit folgt weiter

$$\begin{aligned} \psi(\varphi) &= \langle \phi(x_\varphi), \phi(x_\psi) \rangle = (x_\psi, x_\varphi) \\ &= (x_\psi, \phi^{-1}(\varphi)) = \phi(\phi^{-1}\varphi(x_\psi)) = \varphi(x_\psi). \end{aligned}$$

□

Bemerkungen

1. Eine Folge $(x_n)_n \subset H$ konvergiert genau dann schwach gegen $x \in H$, wenn

$$(x_n - x, y) \to 0 \quad \text{für alle } y \in H.$$

Das folgt leicht aus dem Satz von Riesz-Fréchet (Satz 5.4).

2. In Hilberträumen gilt

$$x_n \to x \iff \begin{cases} x_n \rightharpoonup x, \\ |x_n| \to |x| . \end{cases}$$

Die Richtung „$\Longrightarrow$“ ist klar. Zeige nun $\Longleftarrow$: Aus 1. folgt insbesondere für $y = x$

$$(x_n - x, x) \to 0 \iff (x_n, x) \to (x, x).$$

Also

$$\begin{aligned} (x_n - x)^2 = (x_n - x, x_n - x) &= |x_n|^2 - (x, x_n) - (x_n, x) + |x|^2 \\ &= |x_n|^2 - 2\,\mathrm{Re}(x_n, x) + |x|^2 \to |x|^2 - 2\underbrace{\mathrm{Re}(x, x)}_{=|x|^2} + |x|^2 = 0. \end{aligned}$$

3. Das Symbol „$\perp$“ ist nicht zufällig so gewählt für den Orthogonalraum, denn es gilt

$$\text{Othogonalraum im Hilbertraum} \;\cong\; \text{Annihilator im Banachraum.}$$

Ist nämlich $U \subset H$, so ist

$$U^\perp = \{\, v \in H \mid (u, v) = 0 \quad \text{für alle } u \in U \,\}.$$

Hier ist das der Orthogonalraum. Auf der anderen Seite ist

$$U^\perp \overset{\text{Riesz}}{\cong} \{\, \phi(v) \in H' \mid \phi(v)(u) = 0 \quad \text{für alle } u \in U \,\} = \{\, \varphi \in H' \mid \varphi|_U = 0 \,\}$$

und das ist der wohlbekannte Annihilator.

4. Jede beschränkte Folge in H enthält eine schwach konvergente Teilfolge, denn H ist reflexiv und die Behauptung folgt so aus Satz 3.15.

5.2 Orthonormalbasen

Eine Familie $(e_i)_{i\in I}$ heißt *Orthonormalsystem* (kurz ONS) in $(H, (.,.))$, falls

$$(e_i, e_j) = \delta_{ij} \quad \text{für alle } i, j \in I.$$

Beispiel 5.5. Im $\mathbb{K}^N$ ist $(e_1, \ldots, e_N)$ ein Orthonormalsystem.

□

Beispiel 5.6. In $(l^2(\mathbb{N}), (.,.)_2)$ mit dem Skalarprodukt

$$((x_n)_n, (y_n)_n)_2 = \sum_{n=1}^{\infty} x_n \bar{y}_n, \quad (x_n), (y_n) \in l^2$$

ist $(e_n)_{n\in\mathbb{N}}$ ein Orthonormalsystem.

□

Beispiel 5.7. In $(L^2_{\mathbb{C}}(0, T; dx), (.,.)_2)$ mit dem Skalarprodukt

$$(f, g)_2 = \int_0^T f(x)\bar{g}(x)dx, \quad f, g \in L^2_{\mathbb{C}}(0, T; dx)$$

ist $(e_n)_{n\in\mathbb{N}}$ mit

$$e_n(x) = \frac{1}{\sqrt{T}} e^{in\frac{2\pi}{T}x}, \quad x \in]0, T[,\ n \in \mathbb{Z}$$

ein Orthonormalsystem.

□

Eigenschaften.

1. Ist $(e_i)_{i\in I}$ ein Orthonormalsystem, so ist $(e_i)_{i\in I}$ linear unabhängig.
2. Seien $\dim U = n < \infty$ und $\{e_1, \ldots, e_n\}$ ein Orthonormalsystem. Dann ist die orthogonale Projektion auf
$$U_n := \operatorname{lin}\{e_1, \ldots, e_n\}$$
gegeben durch
$$P_U x = \sum_{k=1}^{n} (x, e_k) e_k.$$
3. Gram-Schmidt-Orthonormalisierungsverfahren: Seien $\{u_1, \ldots, u_n\}$ linear unabhängige Vektoren in H. Dann kann induktiv daraus ein Orthonormalsystem $\{e_1, \ldots, e_n\}$ mit
$$\operatorname{lin}\{u_1, \ldots, u_n\} = \operatorname{lin}\{e_1, \ldots, e_n\}$$
kostruiert werden:
$$e_1 := \frac{u_1}{|u_1|}.$$
Wenn $e_1, \ldots, e_k$ schon konstruiert sind, dann setze
$$f_{k+1} := u_{k+1} - \sum_{i=1}^{k} (u_{k+1}, e_i) e_i$$
$$e_{k+1} := \frac{f_{k+1}}{|f_{k+1}|}.$$
Insbesondere hat jeder endlichdimensionale Hilbertraum ein Orthonormalsystem.

Beweis. Zu 1: Sei *o.E.*

$$x = \lambda_1 e_1 + \ldots + \lambda_n e_n.$$

Dann folgt

$$(e_i, x) = \lambda_i = 0 \quad \text{für alle } i \in \{1, \ldots, n\}.$$

Zu 2: Zunächst ist

$$\sum_{k=1}^{n} (x, e_k) e_k \in U$$

und

$$(x - P_U x, e_i) = (x, e_i) - (\sum_{k=1}^{n} (x, e_k) e_k, e_i) = (x, e_i) - \sum_{k=1}^{n} (x, e_k) \underbrace{(e_k, e_i)}_{=\delta_{ik}} = 0$$

für alle $i = 1, \ldots, n$. Daraus folgt

$$(x - P_U x, u) = 0 \quad \text{für alle } u \in U_n.$$

Zu 3: Die Behauptungen sind alle klar und in der Tat gilt $(e_i, e_j) = 0$, denn

$$(f_i, e_l) = (u_i, e_l) - \sum_{j=1}^{i-1} (u_i, e_j) \underbrace{(e_i, e_l)}_{=\delta_{il}} = 0.$$

□

Definition 5.2. Ein Orthonormalsystem $\{ e_i \mid i \in I \}$ in einem unendlichdimensionalen Hilbertraum $(H, (.,.))$ heißt *Orthonormalbasis* (kurz: ONB) in H, wenn

$$\overline{\text{lin}\{ e_i \mid i \in I \}} = H.$$

Satz 5.5. *Jeder separable Hilbertraum besitzt eine Orthonormalbasis.*

Beweis. Da H separabel ist, gibt es eine Folge $(x_n)_n \subset H$ mit

$$\overline{\text{lin}\{ x_n \mid n \in \mathbb{N} \}} = H.$$

Extrahiere aus $(x_n)_n$ eine Folge linear unabhängiger Vektoren $(u_n)_n$:

$$u_1 := x_1$$

Wenn x_2 linear unabhängig von x_1, dann setze

$$u_2 := x_2,$$

andernfalls fahre mit x_3 fort usw.

Wende nun das Gram-Schmidt Orthonormalisierungsverfahren an. □

Lemma 5.2 (**Bessel'sche Ungleichung**). *Sei* $\{ e_n \mid n \in \mathbb{N} \}$ *ein Orthonormalsystem. Dann gilt*

$$\sum_{n=1}^{\infty} |(x, e_n)|^2 \leq |x|^2 \quad \textit{für alle } x \in H.$$

Beweis. Sei $N \in N$ und

$$U_N := \operatorname{lin}\{e_1, \ldots, e_N\}.$$

Dann gilt

$$x_N = x - \sum_{n=1}^{N} (x, e_n) e_n \in U_N^{\perp},$$

denn

$$(x_N, e_k) = (x, e_k) - \sum_{i=1}^{N} (x, e_n) \underbrace{(e_n, e_k)}_{=\delta_{nk}} = 0.$$

Mit dem Satz von Phytagoras folgt

$$\begin{aligned} |x|^2 &= |\underbrace{x_N}_{\in U_N^{\perp}}|^2 + |\underbrace{\sum_{n=1}^{N} (x, e_n) e_n}_{\in U_N}|^2 = |x_N|^2 + |(x, e_1)e_1 + \ldots + (x, e_N)e_N|^2 \\ &\overset{\text{Phytagoras}}{=} |x_N|^2 + \sum_{n=1}^{N} |(x, e_n)|^2 \underbrace{|e_n|^2}_{=1}. \end{aligned}$$

Es folgt

$$|x|^2 \geq \sum_{n=1}^{N} |(x, e_n)|^2 \qquad \text{für alle } N \in \mathbb{N}.$$

Für $N \to \infty$ folgt schließlich die Behauptung. □

Satz 5.6 (**Fourierentwicklung**). *Sei $\{ e_n \mid n \in \mathbb{N} \}$ eine Orthonormalbasis des Hilbertraumes $(H, (.,.))$. Dann gilt für alle $x, y \in H$*

(i)

$$x = \sum_{n=1}^{\infty} (x, e_n) e_n.$$

Das ist die Fourierreihe *von x bezüglich der Basis $\{ e_n \mid n \in \mathbb{N} \}$ und (x, e_n) ist der n-te* Fourierkoeffizient *von x.*

(ii)

$$(x, y) = \sum_{n=0}^{\infty} (x, e_n)(e_n, y).$$

(iii) Parsevalsche Gleichung:

$$|x|^2 = \sum_{n=1}^{\infty} |(x, e_n)|^2 .$$

Beweis. Zu (i): Sei

$$U_n := \operatorname{lin}\{e_1, \ldots, e_n\}, \quad n \in \mathbb{N}.$$

Es gilt

$$H = \overline{\bigcup_{n \in \mathbb{N}} U_n}. \tag{5.14}$$

Sei n fest. Dann gilt für alle $x \in U_n$ und $N \geq n$

$$P_N(x) = \sum_{i=1}^{N} (x, e_j) e_i = \sum_{i=1}^{N} (\sum_{j=1}^{n} x_j e_j, e_i) e_i = \sum_{i=1}^{N} \sum_{j=1}^{n} x_j \underbrace{(e_j, e_i)}_{=\delta_{ij}} e_i = \sum_{j=1}^{n} x_j e_j = x.$$

Das heißt also

$$\lim_{N\to\infty} P_N(x) = \lim_{N\to\infty} \sum_{n=1}^{N} (x, e_n)e_n = x.$$

Somit gilt folglich

$$\lim_{N\to\infty} P_N(x) = x \quad \text{für alle } x \in \bigcup_{n\in\mathbb{N}} U_n. \tag{5.15}$$

Da $\bigcup U_n$ dicht in H liegt und die Operatoren $(P_N)_N$ in $L(H)$ beschränkt sind, folgt

$$\lim_{N\to\infty} P_N(x) = x \quad \text{für alle } x \in H. \tag{5.16}$$

Sei nämlich $\epsilon > 0$. Wegen (5.14) gibt es zu $x \in H$ ein $u \in \bigcup U_n$, so dass

$$|x - u| \leq \tfrac{\epsilon}{2}. \tag{5.17}$$

Wegen (5.15) gibt es dann ein $N_0 \in \mathbb{N}$, so dass für alle $N \geq N_0$

$$|P_N(x) - u| \leq \tfrac{\epsilon}{2}. \tag{5.18}$$

Es folgt schließlich für alle $N \geq N_0$

$$|P_N(x) - x| \leq \underbrace{|P_N(x) - u|}_{\overset{(5.18)}{\leq}\ \epsilon/2} + \underbrace{|u - x|}_{\overset{(5.17)}{\leq}\ \epsilon/2} \leq \epsilon.$$

Damit ist (5.16) gezeigt und somit auch (i).

Zu (ii): Nach Lemma 5.1 ist $(.,.)$ stetig auf $H \times H$. Daher

$$\begin{aligned}(x, y) &= \lim_{N\to\infty} (P_N x, P_N y) = \lim_{N\to\infty} (\sum_{k=1}^{N} (x, e_k)e_k, \sum_{i=1}^{N} (y, e_i)e_i) \\ &= \lim_{N\to\infty} \sum_{k=1}^{N} (x, e_k) \sum_{i=1}^{N} (y, e_i) \underbrace{(e_k, e_i)}_{=\delta_{ik}} = \lim_{N\to\infty} \sum_{k=1}^{N} (x, e_k)(e_k, y).\end{aligned}$$

Zu (iii): Folgt aus (ii) mit $x = y$. □

Die Parsevalsche Gleichung besagt, dass die Fourierkoeffizientenfolge

$$((x, e_n))_n \in l^2_{\mathbb{K}}(\mathbb{N}).$$

Davon gilt auch die Umkehrung:

Lemma 5.3. *Sei $(\alpha_n)_n \in l^2_{\mathbb{K}}(\mathbb{N})$. Dann konvergiert*

$$x := \sum_{n=1}^{\infty} \alpha_n e_n$$

in H und es gilt

$$\alpha_n = (x, e_n) \quad \textit{für alle } n \in \mathbb{N}.$$

Beweis. Für $m \leq n$ betrachte

$$\left| \sum_{k=m}^{n} \alpha_k e_k \right|^2 = (\sum_{k=m}^{n} \alpha_k e_k, \sum_{i=m}^{n} \alpha_i e_i) = \sum_{k=m}^{n} |\alpha_k|^2 . \tag{5.19}$$

Da $(\alpha_k)_k \in l^2_{\mathbb{K}}(\mathbb{N})$, ist

$$(\sum_{k=1}^{n} |\alpha_k|^2)_n$$

eine Cauchy-Folge in $\mathbb{K}$. Aus (5.19) folgt, dass

$$(\sum_{k=1}^{n} \alpha_k e_k)_n$$

eine Cauchy-Folge in H ist. Nun ist H vollständig, also konvergiert

$$\sum_{k=1}^{\infty} \alpha_k e_k =: x$$

in H. Weiter gilt für alle $i \in \mathbb{N}$

$$(x, e_i) = \lim_{N\to\infty} (\sum_{k=1}^{N} \alpha_k e_k, e_i) = \alpha_i.$$

□

Bemerkungen.

1. Nach der Parsevalschen Gleichung und Lemma 5.3 ist die Abbildung

$$U : H \to l^2_{\mathbb{K}}(\mathbb{N}), \quad x \mapsto ((x, e_n))_n$$

linear, bijektiv und isometrisch. Das heißt aber gerade

$$\boxed{H \cong l^2_{\mathbb{K}}(\mathbb{N})}.$$

2. Die Fourierreihe

$$x = \sum_{k=1}^{\infty} (x, e_k) e_k$$

konvergiert *unbedingt*, d.h.

$$\sum_{k=1}^{\infty} (x, e_{\pi(k)}) e_{\pi(k)}$$

konvergiert gegen x für jede beliebige Permutation

$$\pi : \mathbb{N} \to \mathbb{N}.$$

Im Skalarenkörper $\mathbb{K}$ ist nach *Analysis 1* die unbedingte Konvergenz von Reihen äquivalent zu ihrer absoluten Konvergenz.

Im unendlich-dimensionalen Fall ist das nicht richtig (Vergleiche das Beispiel 5.8 unten):

Absolute Konvergenz impliziert unbedingte Konvergenz, aber nicht umgekehrt.

Das ist die Aussage des Satzes von *Dvoretzky-Rogers*.

Beispiel 5.8. Sei $(\alpha_k)_k \in l^2_{\mathbb{K}}(\mathbb{N}) \setminus l^1_{\mathbb{K}}(\mathbb{N})$. Zum Beispiel liegt $(\alpha_k = \frac{1}{k})_k$ in diesem Raum. Weiter sei $\{ e_k \mid k \in \mathbb{N} \}$ eine Orthonormalbasis. Nach Lemma 5.2 konvergiert die Reihe

$$\sum_{k=1}^{\infty} \alpha_k e_k$$

unbedingt in H, aber eben nicht absolut, da

$$\sum_{k=1}^{\infty} |\alpha_k e_k| = \sum_{k=1}^{\infty} |\alpha_k| \underbrace{|e_k|}_{=1} = +\infty,$$

weil $\alpha_k \notin l^1_{\mathbb{K}}(\mathbb{N})$.

□

5.3 Operatoren im Hilbertraum

Definition 5.3 (Adjungierte, unitär, normal, selbstadjungiert). Sei $A \in L(H)$ und $\phi : H \to H$ der kanonische Isomorphismus aus dem Rieszschen Darstellungssatz (Satz 5.4).

(i) Dann heißt

$$A^* = \phi^{-1} \circ A' \circ \phi \qquad \begin{array}{ccccc} & H' & \xrightarrow{A'} & H' & \\ \phi & \uparrow & & \downarrow & \phi^{-1} \\ & H & \xrightarrow{A^*} & H & \end{array}$$

die *Hilbertraum-Adjungierte zu A.*

A^* ist charakterisiert durch

$$A'(\phi)(x) = \phi(y)(Ax) = (Ax, y) = (x, A^*y) = \phi(A^*y)(x) \quad \text{für alle } x, y \in H.$$

(ii) A heißt *selbstadjungiert*, wenn $A = A^*$.

(iii) A heißt *unitär*, wenn $AA^* = A^*A = I$.

(iv) A heißt *normal*, wenn $AA^* = A^*A$.

Bemerkungen.

1. Unitäre und selbstadjungierte Operatoren sind offensichtlich normal.
2. Selbstadjungierte Operatoren sind charakterisiert durch $(Ax, y) = (x, Ay)$ für alle $x, y \in H$.
3. Unitäre Operatoren sind charakterisiert durch $(Ax, Ay) = (x, y)$ für alle $x, y \in H$.
4. Normale Operatoren sind charakterisiert durch $(Ax, Ay) = (A^*x, A^*y)$ für alle $x, y \in H$.

Beispiel 5.9. Seien $H = L^2_{\mathbb{C}}(a, b)$, $k \in L^2_{\mathbb{C}}((a, b)^2)$ und

$$(Au)(s) = \int_a^b k(s, t)u(t)dt, \quad s \in]a, b[\text{ f.ü. und für alle } u \in L^2_{\mathbb{C}}(a, b).$$

Dann ist $A \in L(H)$ und A^* ist definiert durch

$$(A^*v)(t) = \int_a^b \overline{k(s, t)}v(s)ds, \quad t \in]a, b[\text{ f.ü. und für alle } v \in L^2_{\mathbb{C}}(a, b),$$

denn

$$\begin{aligned} (A^*u, v) &= \int_a^b Au(s)\overline{v(s)}ds = \int_a^b \left(\int_a^b k(s, t)u(t)dt \right) \overline{v(s)}ds \\ &\overset{\text{Fubini}}{=} \int_a^b u(t) \left(\int_a^b k(s, t)\overline{u(s)}ds \right) dt \\ &= \int_a^b u(t) \underbrace{\overline{\left(\int_a^b \overline{k(s, t)}u(s)ds \right)}}_{=A^*v} dt. \end{aligned}$$

Wir sehen hieran auch, dass A selbstadjungiert ist, wenn

$$k(s, t) = \overline{k(t, s)} \quad \text{f.ü. auf }]a, b[\times]a, b[.$$

Solche Kerne heißen *symmetrisch*. Ein Beispiel für einen solchen Kern ist

$$k(s,t) = h(t-s)$$

mit rellwertigen und geradem h.

□

Beispiel 5.10. A^*A und AA^* sind stets selbstadjungiert.

□

Eigenschaften. Seien $A, B \in L(H)$ und $\lambda \in \mathbb{K}$. Dann gilt

1. $(A+B)^* = A^* + B^*$.
2. $(\lambda A)^* = \bar{\lambda} A^*$.
3. $(A \circ B)^* = B^* \circ A^*$.
4. $\|A^*\|_{L(H)} = \|A'\|_{L(H)} = \|A\|_{L(H)}$., also insbesondere $A^* \in L(H)$.
5. $A^{**} = A$.
6. $\|AA^*\|_{L(H)} = \|A^*A\|_{L(H)} = \|A\|^2_{L(H)}$.
7. Ist A selbstadjungiert oder unitär, so ist A normal.
8. Für $A \in L(H)$ ist äquivalent:
   (i) A unitär.
   (ii) A surjektiv und isometrisch.
   (iii) A surjektiv und A ist winkelerhaltend, d.h.
   $$(Ax, Ay) = (x, y) \quad \text{für alle } x, y \in H.$$
9. Ist $A \in L(H)$ normal, so gilt
   $$|Ax| = |A^*x| \quad \text{für alle } x \in H,$$
   d.h. insbesondere
   $$\ker(A) = \ker(A^*).$$
10. Ist $A \in L(H)$ normal, so ist auch A^n normal für alle $n \in \mathbb{N}$.
11. $\ker A = R(A^*)^\perp$ und $\ker A^* = R(A)^\perp$. Insbesondere ist A genau dann injektiv, wenn $R(A^*)$ dicht liegt.

Beweis. <u>Zu 1 bis 3:</u> Folgen direkt aus der Definition.

<u>Zu 4:</u> Es gilt

$$\|A^*\| = \left\|\phi^{-1} \circ A' \circ \phi\right\| \overset{\phi \text{ Isometrie}}{=} \|A'\| \overset{\text{Satz 3.18}}{=} \|A\|.$$

<u>Zu 5:</u> Es gilt für alle $x, y \in H$

$$(Ax, y) = (x, A^*y) = \overline{(A^*x, y)} = \overline{(y, A^{**}x)} = (A^{**}x, y).$$

<u>Zu 6:</u> Zunächst gilt für alle $x \in H$

$$|Ax|^2 = (Ax, Ax) = (x, A^*Ax) \overset{\text{CSU}}{\leq} |x|\,|A^*Ax|\,.$$

Nach Supremumsübergang folgt

$$\|A\|^2 \leq \|A^*A\| \leq \|A^*\|\|A\| \overset{4.}{=} \|A\|^2,$$

also

$$\|A\|^2 = \|A^*A\|.$$

Schließlich ist

$$\|A\|^2 \overset{4.}{=} \|A^*\|^2 \overset{\text{eben gezeigt}}{=} \|A^{**}A^*\| \overset{5.}{=} \|AA^*\|.$$

<u>Zu 7:</u> Trivial.

<u>Zu 8:</u> **Zu „$(i) \Rightarrow (ii)$“:** Sei A unitär, d.h.

$$AA^* = A^*A = I.$$

Dann gilt

$$1 = \|I\| = \|A^*A\| \overset{6.}{=} \|A\|^2 \Longrightarrow \|A\| \overset{4.}{=} \|A^*\| = 1.$$

Daraus folgt

$$\|Ax\| \leq \|A\|\|x\| = \|x\| = \|A^*Ax\| \leq \|A^*\|\|Ax\| = \|Ax\|.$$

Also ist

$$\|Ax\| = \|x\|.$$

d.h. A ist isometrisch.

Sei nun $y \in H$. Dann setze

$$x := A^*y.$$

Es folgt

$$Ax = AA^*y = y,$$

also ist A surjektiv.

Zu „$(ii) \Rightarrow (iii)$“: Zeige nur den Fall $\mathbb{K} = \mathbb{R}$. Der Fall $\mathbb{K} = \mathbb{C}$ geht analog mit der entsprechenden Polarisationsformel. Für alle $x, y \in H$ gilt

$$\begin{aligned}(Ax, Ay) &= \tfrac{1}{4}(|Ax + Ay|^2 - |Ax - Ay|^2) = \tfrac{1}{4}(|A(x+y)|^2 - |A(x-y)|^2) \\ &\overset{A \text{ isometrisch}}{=} \tfrac{1}{4}(|x+y|^2 - |x-y|^2) = (x, y).\end{aligned}$$

Zu „$(iii) \Rightarrow (i)$“: Zunächst gilt für alle $x, y \in H$

$$(Ax, Ay) = (x, A^*Ay) = (x, y),$$

woraus

$$A^*A = I$$

folgt. Weiter gilt für alle $x, y \in H$

$$(A^*x, A^*y) = (x, A^{**}A^*y) \overset{5.}{=} (x, AA^*y) = (x, y),$$

also

$$AA^* = I.$$

<u>Zu 9:</u> Für alle $x \in H$ gilt

$$|Ax|^2 = (Ax, Ax) = (x, A^*Ax) \overset{A \text{ normal}}{=} (x, AA^*x) = (A^*x, A^*x) = |A^*x|^2.$$

<u>Zu 10:</u> Per Induktion:

$n = 1$: Trivial.

$n \to n+1$: Betrachte

$$\begin{aligned}A^{n+1}(A^{n+1})^* &= A \circ A^n \circ (A \circ A^n)^* = A \circ A^n \circ (A^n)^* \circ A^* \\ &\overset{\text{IV}}{=} A \circ (A^n)^* \circ A^n \circ A^* \overset{\text{IV}}{=} (A^n)^* \circ A \circ A^* \circ A^n \\ &\overset{\text{IV}}{=} (A^n)^* \circ A^* \circ A \circ A^n = (A^{n+1})^*(A^{n+1}).\end{aligned}$$

Zu 11: Zunächst gilt $\ker A = (R(A^*))^\perp$, weil

$$\begin{aligned} Ax = 0 &\iff (Ax, y) = 0 \quad \text{für alle } y \in H \\ &\iff (x, A^*y) = 0 \quad \text{für alle } y \in H \\ &\iff x \in R(A^*)^\perp. \end{aligned}$$

Daher gilt auch

$$\ker A^* = (R(A)^{**})^\perp = (R(A))^\perp.$$

Zum Zusatz: Sei zunächst $\overline{R(A^*)} = H$. Aus Korollar 5.1 folgt dann

$$H = \overline{R(A^*)} = (R(A)^\perp)^\perp,$$

also

$$R(A)^\perp = \{0\} \Longrightarrow \ker(A) = \{0\}$$

und daher ist A injektiv.

Ist umgekehrt A injektiv, also $ker(A) = R(A^*)^\perp = \{0\}$, so folgt erneut mit Korollar 5.1 schließlich

$$\overline{R(A^*)} = (R(A^*)^\perp)^\perp = (\{0\})^\perp = H.$$

Also liegt $R(A^*)$ dicht in H. □

Satz 5.7 (von Hellinger-Toeplitz). *Seien H ein Hilbertraum und $A \in L(H)$ mit der Eigenschaft*

$$(Ax, y) = (x, Ay) \quad \textit{für alle } x, y \in H.$$

Dann ist A stetig und damit insbesondere selbstadjungiert.

Beweis. Aufgrund des Satzes vom abgeschlossenen Graphen (Satz 2.8) müssen wir zeigen:

$$x_n \to 0,\ Ax_n \to z \Longrightarrow z = 0.$$

Und tatsächlich gilt

$$(z, z) = (\lim Ax_n, z) = \lim(Ax_n, z) = \lim(x_n Az) = 0.$$

□

Satz 5.8. *Seien $\mathbb{K} = \mathbb{C}$ und $A \in L(H)$. Dann ist A genau dann selbstadjungiert, wenn $(Ax, x) \in \mathbb{R}$ für alle $x \in H$.*

Beweis. Zu „$\Longrightarrow$“: Für $x \in H$ gilt

$$(Ax, x) = (x, A^*x) = (x, Ax) = \overline{(Ax, x)}.$$

Zu „$\Longleftarrow$“: Sei $\lambda \in \mathbb{C}$. Dann betrachte die reelle(!) Zahl

$$(A(x + \lambda y), x + \lambda y) = (Ax, x) + \bar{\lambda}(y, Ax) + \lambda(x, Ay) + |\lambda|^2 (Ay, y).$$

Die komplex Konjugierte davon ist

$$(A(x + \lambda y), x + \lambda y) = (Ax, x) + \lambda(y, Ax) + \bar{\lambda}(x, Ay) + |\lambda|^2 (Ay, y).$$

Setze $\lambda = 1$ und $\lambda = -i$ ein. Dann folgt

$$\begin{aligned} (Ax, y) + (Ay, x) &= (y, Ax) + (x, Ay) \\ (Ax, y) - (Ay, x) &= -(y, Ax) + (x, Ay). \end{aligned}$$

Also ist tatsächlich $(Ax, y) = (x, Ay)$. □

Lemma 5.4. *Seien $A \in L(H)$ selbstadjungiert und $\mathbb{K} = \mathbb{R}$. Dann gilt*

$$\|A\| = \sup_{|x|=1} |(Ax, x)| .$$

Beweis. Zu „$\geq$“: Es gilt für alle $x \in H$

$$|(Ax, x)| \overset{\text{CSU}}{\leq} |Ax|\,|x| \leq \|A\|\,|x|^2 .$$

Für $|x| = 1$ gilt also stets

$$|(Ax, x)| \leq \|A\|,$$

d.h.

$$\sup_{|x|=1} |(Ax, x)| \leq \|A\|.$$

Zu „$\leq$“: Für alle $x, y \in H$ gilt

$$\begin{aligned}
(A(x+y), x+y) - (A(x-y), x-y) &= (Ax, x) + (Ax, y) + (Ay, x) + (Ay, y) \\
&\quad -(Ax, x) + (Ax, y) + (Ay, x) - (Ay, y) \\
&= 2(Ax, y) + 2(Ay, x) = 2(Ax, y) + 2(y, A^*x) \\
&\overset{A=A^*}{=} 2(Ax, y) + 2(y, Ax) \overset{\mathbb{K}=\mathbb{R}}{=} 4(Ax, y).
\end{aligned}$$

Daraus folgt

$$\begin{aligned}
4\,|(Ax, y)| &\leq |(A(x+y), x+y)| + |(A(x-y), x-y)| \\
&= |x+y|^2\,|(A\tfrac{x+y}{|x+y|}, \tfrac{x+y}{|x+y|})| + |x-y|^2\,|(A\tfrac{x-y}{|x-y|}, \tfrac{x-y}{|x-y|}| \\
&\leq (|x+y|^2 + |x-y|^2) \max\left\{|(A\tfrac{x+y}{|x+y|}, \tfrac{x+y}{|x+y|})|, |(A\tfrac{x-y}{|x-y|}, \tfrac{x-y}{|x-y|})|\right\}
\end{aligned}$$

Nach der Parallelogrammgleichung bedeutet dies

$$4\,|(Ax, y)| \leq 2(|x|^2 + |y|^2) \cdot \sup_{|z|=1} |(Az, z)| .$$

Es folgt schließlich

$$\|A\| = \sup_{|x|,|y|\leq 1} |(Ax, y)| \leq \sup_{|z|=1} |(Az, z)| .$$

□

Daraus folgt nun leicht

Korollar 5.3. *Seien $A \in L(H)$ selbstadjungiert und $\mathbb{K} = \mathbb{R}$. Außerdem gelte $(Ax, x) = 0$ für alle $x \in H$. Dann ist $A = 0$.*

Beweis. Nach Lemma 5.4 ist $\|A\| = 0$, also $A = 0$. □

Lemma 5.5. *Sei $A \in L(H)$. Dann gilt:*

(i) $\sigma(A^*) = \{\, \bar{\lambda} \mid \lambda \in \sigma(A) \,\}$.

(ii) Ist A selbstadjungiert, so gilt $\sigma_p(A) \subset \mathbb{R}$.

(iii) Ist A normal und x Eigenvektor von A zum Eigenwert λ, so ist x Eigenvektor von A^* zum Eigenwert $\bar{\lambda}$.

(iv) Ist A normal und λ, μ Eigenwerte von A mit $\lambda \neq \mu$, so gilt

$$\ker(\lambda I - A) \perp \ker(\mu I - A).$$

(v) Ist $\mathbb{K} = \mathbb{C}$ und A normal, so gilt

$$\exists_{\lambda \in \sigma(A)} : |\lambda| = \|A\|.$$

(vi) Ist $\mathbb{K} = \mathbb{R}$ und A selbstadjungiert und kompakt, so ist $\|A\|$ oder $-\|A\| \in \sigma(A)$.

Bemerkungen.

1. Man kann in (ii) sogar $\sigma(A) \subset \mathbb{R}$ zeigen, doch das ist viel schwieriger.
2. Wegen (v) und Satz 4.7 gilt speziell für den Spektralradius

$$r(A) = \max_{\lambda \in \sigma(A)} |\lambda| = \|A\|.$$

Beweis des Lemmas. Zu (i): Das folgt aus

$$\begin{aligned} \lambda \in \rho(A) &\iff (\lambda I - A) \text{ bijektiv} \iff (\lambda I - A)' \text{ bijektiv} \overset{\phi \text{ Isomorphismus}}{\iff} (\lambda I - A)^* \text{ bijektiv} \\ &\iff \bar{\lambda} I^* - A^* \text{ bijektiv} \iff \bar{\lambda} \in \rho(A^*). \end{aligned}$$

Zu (ii): Ist $\lambda \in \sigma_p(A)$, so gibt es ein $x \in H \setminus \{0\}$ mit

$$Ax = \lambda x,$$

d.h.

$$\begin{aligned} \lambda |x|^2 &= \lambda(x,x) = (\lambda x, x) = (Ax, x) = (x, A^* x) \overset{A = A^*}{=} (x, Ax) \\ &= (x, \lambda x) = \bar{\lambda}(x,x) = \bar{\lambda} |x|^2 . \end{aligned}$$

Nun ist $x \neq 0$, da x Eigenvektor ist. Daher ist auch $|x| \neq 0$ und wir können beruhigt dadurch teilen, so dass

$$\lambda = \bar{\lambda},$$

also $\lambda \in \mathbb{R}$ folgt.

Zu (iii): Ist A normal, so ist auch $\lambda I - A$ normal, denn

$$\begin{aligned} (\lambda I - A)(\lambda I - A)^* &= \lambda I (\lambda I)^* - \lambda I A^* - A(\lambda I)^* + AA^* = \lambda \bar{\lambda} I^* - \lambda A^* - \bar{\lambda} A I^* + A^* A \\ &= (\lambda I)^* \lambda I - A^* \lambda I - (\lambda I)^* A + A^* A = (\lambda I - A)^* (\lambda I - A). \end{aligned}$$

Nach Eigenschaft 9. ist daher

$$\ker(\lambda I - A) = \ker(\lambda I - A)^* = \ker(\bar{\lambda} I^* - A^*).$$

Zu (iv): Seien $x \in \ker(\lambda I - A)$ und $y \in \ker(\mu I - A)$ mit $x \neq 0 \neq y$. Betrachte

$$\lambda(x,y) = (\lambda x, y) = (Ax, y) = (x, A^* y) \overset{(i)}{=} (x, \bar{\mu} y) = \mu(x,y).$$

Wegen $\lambda \neq \mu$ folgt $(x, y) = 0$, also $x \perp y$.

<u>Zu (v):</u> Zunächst existiert nach Satz 4.7 ein $\lambda \in \sigma(A)$, so dass

$$r(A) = |\lambda| \,.$$

Zeigen wir also

$$r(A) = \|A\|, \tag{5.20}$$

so sind wir fertig.

Zunächst gilt

$$\begin{aligned} \left\|A^2\right\|^2 &\overset{\text{Eigenschaft 6}}{=} \left\|A^2(A^2)^*\right\| \overset{\text{Eigenschaft 3}}{=} \left\|A^2(A^*)^2\right\| = \left\|(AA^*)^2\right\| = \|(AA^*)(AA^*)\| \\ &\overset{\text{Eigenschaft 6}}{=} \|AA^*\|^2 \overset{\text{Eigenschaft 6}}{=} (\|A\|^2)^2, \end{aligned}$$

also

$$\left\|A^2\right\| = \|A\|^2 \quad \text{für } A \text{ normal.}$$

Da nach Eigenschaft 10. mit A auch A^n normal ist für alle $n \in \mathbb{N}$, gilt für alle $k \in \mathbb{N}$

$$\|A\|^{2^k} = \| \, A^{2^k} \, \| \,. \tag{5.21}$$

Mit Satz 4.8 folgt

$$r(A) = \lim_{n\to\infty} \|A^n\|^{1/n} = \lim_{k\to\infty} \| \, A^{2^k} \, \|^{1/2^k} \overset{(5.21)}{=} \lim_{k\to\infty} \|A\|^{2^k \cdot 1/2^k} = \|A\|.$$

Das zeigt (5.20).

<u>Zu (vi):</u> Nach *Analysis 1* gibt es eine Folge $(x_{n'})_{n'} \subset B_H$ mit

$$|(Ax_{n'}, x_{n'})| \to \sup_{|z|=1} |(Az, z)| \overset{\text{Lemma 5.4}}{=} \|A\|.$$

Da A kompakt ist, gibt es eine Teilfolge $(x_n)_n \subset (x_{n'})_{n'}$, so dass die Grenzwerte

$$\lambda := \lim_{n\to\infty} (Ax_n, x_n), \quad y := \lim_{n\to\infty} Ax_n \tag{5.22}$$

existieren. Es gilt

$$\begin{aligned} |Ax_n - \lambda x_n|^2 &= (Ax_n - \lambda x_n, Ax_n - \lambda x_n) \\ &\overset{\mathbb{K}=\mathbb{R}}{=} |Ax_n|^2 - 2\lambda(Ax_n, x_n) + \lambda^2 \underbrace{|x_n|^2}_{=1} \\ &\leq \|A\|^2 \underbrace{|x_n|^2}_{=1} - 2\lambda(Ax_n, x_n) + \lambda^2 \\ &= 2\lambda^2 - 2\lambda \underbrace{(Ax_n, x_n)}_{\to\lambda} \to 0. \end{aligned} \tag{5.23}$$

Also ist

$$y = \lim_{n\to\infty} \lambda x_n,$$

denn

$$|\lambda x_n - y| \leq \underbrace{|\lambda x_n - Ax_n|}_{\overset{(5.23)}{\to} 0} + \underbrace{|Ax_n - y|}_{\overset{(5.22)}{\to} 0} \to 0.$$

Weiter folgt wegen der Stetigkeit von A

$$Ay = A(\lim \lambda x_n) = \lambda \lim Ax_n = \lambda y.$$

Ist also $y \neq 0$, so ist λ Eigenwert von A und die Behauptung wegen $|\lambda| = \|A\|$ gezeigt.

Ist $y = 0$, so gilt $Ax_n \to 0$. Daraus folgt

$$\|A\| = \lim(Ax_n, x_n) \overset{\text{CSU}}{\leq} \lim |Ax_n|^2 \underbrace{\|x_n\|^2}_{=1} = 0,$$

also $\|A\| = 0$ und so $A = 0$, woraus ebenfalls die Behauptung folgt. □

Satz 5.9 (**Spektralzerlegungssatz für kompakte, normale Operatoren**). *Sei $A : H \to H$ linear im unendlichdimensionalen, separablen Hilbertraum H. Dann sind äquivalent:*

(i) $A \in K(H)$ und A ist normal (falls $\mathbb{K} = \mathbb{C}$) bzw. selbstadjungiert (falls $\mathbb{K} = \mathbb{R}$).

(ii) Es existiert ein Orthonormalsystem aus Eigenvektoren von A

$$\{ e_n \mid n \in I \}, \quad I = \mathbb{N} \text{ oder } I \text{ endlich}$$

mit Eigenwerten $\{ \lambda_n \mid n \in I \} \subset \mathbb{C} \setminus \{0\}$ bzw. $\mathbb{R} \setminus \{0\}$ mit $\lambda_n \to 0$, falls $I = \mathbb{N}$ und so dass

$$Ax = \sum_{n \in I} \lambda_n (x, e_n) e_n \quad \text{für alle } x \in H$$

und für $x \in H$ ein $h_x \in \ker(A)$ exsitiert mit

$$x = h_x + \sum_{n \in I} (x, e_n) e_n. \tag{5.24}$$

Beweis. Zu „$(ii) \Longrightarrow (i)$“: Nach der Besselschen Ungleichung (Lemma 5.2) gilt für alle $x \in H$

$$((x, e_n))_n \in l^2(I).$$

Somit ist auch

$$(\lambda_n (x, e_n))_n \in l^2(I) \quad \text{für alle } x \in H,$$

denn $\lambda_n \to 0$, falls $I = \mathbb{N}$. Weiter gilt für alle $x \in H$ und $m, n \in I$ mit $m \leq n$

$$\begin{aligned} |\sum_{k=m}^{n} \lambda_k (x, e_k) e_k|^2 &= \left(\sum_{k=m}^{n} \lambda_k (x, e_k) e_k, \sum_{l=m}^{n} \lambda_l (x, e_l) e_l \right) \\ &= \sum_{k=m}^{n} \lambda_k (x, e_k) \sum_{l=m}^{n} \bar{\lambda}_l \overline{(x, e_l)} \underbrace{(e_k, e_l)}_{=\delta_{kl}} \\ &= \sum_{k=m}^{n} \lambda_k \bar{\lambda}_k (x, e_k) \overline{(x, e_k)} = \sum_{k=m}^{n} |\lambda_k|^2 |(x, e_k)|^2 . \end{aligned}$$

Deswegen folgt daraus auch die Konvergenz der Reihe

$$\sum_{n \in I} \lambda_n (x, e_n) e_n.$$

Wegen

$$|Ax|^2 = |\sum_{n \in I} (x, e_n) e_n|^2 < \infty$$

ist A beschränkt. Offensichtlich ist A auch linear, also $A \in L(H)$.

Wir zeigen jetzt

$$\ker(A) \subset \ker(A^*) \tag{5.25}$$

Sei dazu $x \in \ker(A)$, also

$$Ax = 0 \Longrightarrow (Ax, e_m) = 0 \quad \text{für alle } m \in I.$$

Daraus erhalten wir

$$(Ax, e_m) = \sum_{n \in I} \lambda_n (x, e_n) \underbrace{(e_n, e_m)}_{=\delta_{nm}} = 0,$$

also

$$\underbrace{\lambda_m}_{\neq 0} (x, e_m) = 0 \Longrightarrow (x, e_m) = 0 \quad \text{für alle } m \in I.$$

Es folgt

$$(A^* x, e_m) = (x, Ae_m) = (x, \lambda_m e_m) = \bar{\lambda}_m (x, e_m) = 0$$

und schließlich

$$(A^* x, y) = 0 \quad \text{für alle } y \in H.$$

Das heißt aber gerade

$$A^* x = 0$$

und demnach ist $x \in \ker(A^*)$ und somit (5.25) gezeigt.

Wir berechnen jetzt die Darstellung von A^*: Für alle $x \in H$ gilt

$$\begin{aligned} A^* x &= A^*(h_x + \sum_{n \in I} (x, e_n) e_n) = \underbrace{A^* h_x}_{\overset{(5.25)}{=} 0} + \sum_{n \in I} A^* e_n \\ &= \sum_{n \in I} (x, e_n) \bar{\lambda}_n e_n, \end{aligned}$$

also

$$A^* x = \sum_{n \in I} \bar{\lambda}_n (x, e_n) e_n \quad \text{für alle } x \in H. \tag{5.26}$$

Wir zeigen jetzt, dass A normal bzw. selbstadjungiert ist:

a) Falls $\mathbb{K} = \mathbb{R}$, dann ist nach Voraussetzung $\bar{\lambda}_n = \lambda_n$ für alle $n \in I$ und so wegen (5.26)

$$A^* x = Ax \quad \text{für alle } x \in H.$$

Demnach ist $A^* = A$, also A selbstadjungiert.

b) Falls $\mathbb{K} = \mathbb{C}$, dann gilt

$$\begin{aligned} AA^* x &= A(\sum_{n \in I} \bar{\lambda}_n (x, e_n) e_n) = \sum_{n \in I} \bar{\lambda}_n (x, e_n) A e_n \\ &= \sum_{n \in I} \bar{\lambda}_n \lambda_n = A^* A x \quad \text{für alle } x \in H, \end{aligned}$$

also

$$AA^* = A^* A.$$

Demnach ist A normal.

Es bleibt damit zu zeigen, dass $A \in K(H)$:

a) Ist I endlich, so ist $A \in F(H) \subset K(H)$.

b) Ist $I = \mathbb{N}$, so gilt für

$$A_N x := \sum_{k=1}^{N} \lambda_k (x, e_k) e_k \in F(H) \subset K(H)$$

die Abschätzung

$$|Ax - A_N x|^2 = |\sum_{k=N+1}^{\infty} \lambda_k (x, e_k) e_k|^2 \overset{\text{wie oben}}{=} \sum_{k=N+1}^{\infty} |\lambda_k|^2 \, |(x, e_k))|^2$$
$$\leq \sup_{k \geq N+1} \underbrace{|\lambda_k|^2}_{\to 0} |x|^2 \to 0.$$

Daraus folgt

$$\|A - A_N\| \overset{N \to \infty}{\longrightarrow} 0.$$

Somit ist

$$A \in \overline{F(H)} \subset K(H).$$

<u>Zu „$(i) \Longrightarrow (ii)$“</u>: Für $A = 0$ ist die Behauptung trivial. Sei daher $A \neq 0$.

<u>Konstruktionsschema:</u>

1) $n = 1$: Nach Lemma 5.5 existiert ein $\lambda_1 \in \sigma(A) \setminus \{0\}$ mit

$$\lambda_1 = +\|A\| \quad \text{oder} \quad \lambda_1 = -\|A\|,$$

wobei in der Tat $\|A\| \neq 0$ wegen $A \neq 0$.

Nach dem Spektralsatz für kompakte Operatoren (Satz 4.13) gilt

$$\sigma(A) \setminus \{0\} \subset \sigma_p(A),$$

also ist λ_1 Eigenwert von A und es gibt ein $e_1 \in H$ mit

$$|e_1| = 1 \quad \text{und} \quad Ae_1 = \lambda_1 e_1.$$

Setze

$$H_1 := (\text{lin}\{e_1\})^{\perp}.$$

Dann gilt

$$A, A^* : H_1 \to H_1,$$

denn für alle $h_1 \in H_1$ gilt

$$(Ah_1, e_1) = (h_1, A^* e_1) = (h_1, \bar{\lambda}_1 e_1) = \lambda_1 (h_1, e_1) = 0.$$

Analog zeigt man das für A^*.

Betrachte jetzt die Abbildung

$$A_1 := A|_{H_1} : H_1 \to H_1.$$

Dann ist $A_1 \in K(H)$ und A_1 ist normal (falls $\mathbb{K} = \mathbb{C}$) bzw. selbstadjungiert (falls $\mathbb{K} = \mathbb{R}$), da

$$(A_1)^* = (A|_{H_1})^* = A^*|_{H_1}.$$

Jetzt gibt es zwei Möglichkeiten:

a) $A_1 = 0$: Die Konstruktion bricht ab. Es gilt nämlich $H_1 \subset \ker(A)$, ja, es gilt sogar

$$H_1 = \ker(A), \tag{5.27}$$

denn für $x \in \ker(A)$ ist

$$0 = (Ax, e_1) = (x, A^* e_1) = (x, \bar{\lambda}_1 e_1) = \lambda_1 (x, e_1)$$

und wegen $\lambda_1 \neq 0$ folgt

$$(x, e_1) = 0 \quad \text{für alle } x \in H_1,$$

also $\ker(A) \subset H_1$, was (5.27) zeigt.

Setze nun für $x \in H_1$

$$h_x := x - (x, e_1)e_1.$$

Dann ist $h_x \in H_1 = \ker(A)$ und

$$x = h_x + (x, e_1)e_1$$

sowie

$$Ax = \underbrace{Ah_x}_{=0} + (x, e_1)Ae_1 = \lambda_1(x, e_1)e_1.$$

Damit sind wir fertig.

b) $A_1 \neq 0$: Die Konstruktion wird fortgesetzt. Weiter bei 2).

2) $n = 2$: Es existiert wieder ein $\lambda_2 \in \sigma(A_1) \setminus \{0\}$ mit

$$|\lambda_2| = \|A_1\| \neq 0.$$

Wieder folgt mit dem Spektralsatz (Satz 4.13), dass λ_2 Eigenwert von A_1 ist, da A_1 kompakt ist. Also gibt es ein $e_2 \in H_1$ mit

$$|e_2| = 1 \quad \text{und} \quad A_1 e_2 = \lambda_2 e_2.$$

Nach Konstruktion gilt

$$e_1 \perp e_2.$$

Außerdem ist

$$|\lambda_2| = \|A_1\| = \|A|_{H_1}\| \leq \|A\| = |\lambda_1|. \tag{5.28}$$

Setze

$$H_2 := (\operatorname{lin}\{e_1, e_2\})^\perp, \quad A_2 := A|_{H_2} : H_2 \to H_2.$$

Wieder folgt die Wohldefiniertheit von A_2 und $A_2 \in K(H_2)$. Außerdem ist A wieder normal (falls $\mathbb{K} = \mathbb{C}$) bzw. selbstadjungiert (falls $\mathbb{K} = \mathbb{R}$). Erneut gibt es zwei Möglichkeiten:

a) $A_2 = 0$: Die Konstruktion bricht ab. Wieder ist $H_2 = \ker(A)$. Setze für $x \in H$

$$h_x := x - (x, e_1)e_1 - (x, e_2)e_2.$$

Dann gilt $h_x \in \ker(A) = H_2$ und damit auch

$$Ax = \underbrace{Ah_x}_{=0} + \lambda_1(x, e_1)e_1 + \lambda_2(x, e_2)e_2.$$

b) $A_2 \neq 0$: Die Konstruktion wird fortgesetzt...

Wir bemerken, dass es für das Konstruktionschema zwei möglich Ausgänge gibt:

a) Die Konstruktion bricht nach endlich vielen Schritten ab, d.h. es gibt ein $N \in \mathbb{N}$, so dass

$$I := \{1, 2, 3, \ldots, N\}$$

und

$$x = h_x + \sum_{k=1}^{N} (x, e_k)e_k, \quad h_x \in \ker(A) = H_N$$

sowie

$$Ax = \sum_{k=1}^{N} \lambda(x, e_k)e_k.$$

b) Die Konstruktion bricht nie ab: Dann erhalten wir eine Folge von Eigenwerten $(\lambda_n)_n \subset \mathbb{K} \setminus \{0\}$ mit

$$\begin{aligned} \|A\| = |\lambda_1| \geq |\lambda_2| \geq |\lambda_2| \geq \dots, \\ |\lambda_n| = \|A_{n-1}\| \quad \text{für alle } n \in \mathbb{N} \end{aligned}$$

und eine Folge von Eigenvektoren $e_n \in H$, so dass für alle $n \in \mathbb{N}$.

$$\begin{aligned} Ae_n = \lambda_n e_n, \quad |e_n| = 1, \\ A_n = A|_{H_n} \neq 0, \quad H_n \subsetneq \ker(A) \end{aligned}$$

Nun ist A kompakt. Also gilt $\lambda_n \to 0$ nach Satz 4.11.

Für $x \in H$ setze

$$x_n := x - \sum_{k=1}^{n} (x, e_k) e_k \subset H_N = (\operatorname{lin}\{e_1, \dots, e_n\})^{\perp},$$

denn

$$(x_n, e_j) = (x, e_j) - (x, e_j) = 0 \quad \text{für alle } j = 1, \dots, n.$$

Weiter gilt

$$|Ax_n| \overset{A_n : H_n \to H_n}{=} |A_n x_n| \leq \|A_n\| \, |x_n| = |\lambda_{n+1}| \, |x_n|$$

und

$$\begin{aligned} |x_n|^2 &= |x|^2 - 2\operatorname{Re}(x, \sum_{k=1}^{n} (x, e_k) e_k) + \sum_{k=1}^{n} |(x, e_k)|^2 \\ &= |x|^2 - 2\operatorname{Re} \sum_{k=1}^{n} \underbrace{\overline{(x, e_k)}(x, e_k)}_{=|(x,e_k)|^2} + \sum_{k=1}^{n} |(x, e_k)|^2 \\ &= |x|^2 - \underbrace{\sum_{k=1}^{n} |(x, e_k)|^2}_{\geq 0} \leq |x|^2 . \end{aligned}$$

Daraus folgt

$$|Ax_n| \to 0,$$

denn

$$|Ax_n| \leq |\lambda_{n+1}| \, |x_n| \leq \underbrace{|\lambda_{n+1}|}_{\to 0} |x| \to 0.$$

Schließlich folgt

$$Ax = \underbrace{Ax_n}_{\to 0} + \sum_{k=1}^{n} \lambda_k (x, e_k) e_k,$$

also nach Grenzübergang

$$Ax = \sum_{k=1}^{\infty} \lambda_k (x, e_k) e_k.$$

Setze nun für $x \in H$

$$h_x := x - \sum_{k=1}^{\infty} (x, e_k) e_k.$$

Dann ist wie gewünscht $h_x \in \ker(A)$, denn

$$Ah_x = \lim Ax_n = 0.$$

$\square$

Bemerkungen.

1. Die Folge der $(\lambda_n)_n$ kann so geordnet werden, dass

$$(\lambda_n) = (\underbrace{\mu_1, \ldots, \mu_1}_{d_1\text{-viele}}, \underbrace{\mu_2, \ldots, \mu_2}_{d_2\text{-viele}}, \underbrace{\mu_3, \ldots, \mu_3}_{d_3\text{-viele}}, \ldots),$$

wobei (μ_k) die paarweise veschiedenen Eigenvektoren von A sind und die d_i die Dimensionen des jeweiligen Eigenraums (geometrische Vielfachheiten) darstellen, d.h.

$$\begin{aligned} d_1 &:= \dim(\ker(\mu_1 I - A)), \\ d_2 &:= \dim(\ker(\mu_2 I - A)), \\ &\vdots \end{aligned}$$

2. Sei

$$E_n := P_{\ker(\mu_n I - A)}$$

die orthogonale Projektion auf den Eigenraum $\ker(\mu_n I - A)$, d.h.

$$E_n x = \sum_{i=1}^{d_n} (x, e_i^n) e_i^n$$

Dann gilt

$$\begin{aligned} Ax &= \sum_{n=1}^{\infty} \mu_n E_n x \quad \text{falls } I = \mathbb{N}, \qquad (5.29) \\ Ax &= \sum_{n=1}^{N} \mu_n E_n x \quad \text{falls } I = \{1, \ldots, N\}. \end{aligned}$$

Die Reihe (5.29) konvergiert punktweise, d.h.

$$|Ax - \sum_{n=1}^{N} \mu_n E_n x| \overset{N\to\infty}{\longrightarrow} 0.$$

Sie konvergiert sogar in der Operatornorm, denn

$$|| \, A - \sum_{k=1}^{N} \mu_k E_k \, || = || \sum_{k>N} \mu_k E_k \, || \overset{\text{Bemerkung 2}}{=} \max_{k>N} |\mu_k| \to 0.$$

Daher ist auch die Schreibweise

$$A = \sum_{n=1}^{\infty} \mu_n E_n$$

gerechtfertigt.

Satz 5.10. *Sei $A \in K(H)$ selbstadjungiert und postitiv, d.h.*

$$(Ax, x) \geq 0 \quad \textit{für alle } x \in H.$$

Dann existiert genau ein positiv, selbstadjungierter Operator $B : H \to H$ mit $B^2 = A$.
Notation: $B := A^{1/2}$.

Beweis. Zur Existenz: Nach Satz 5.9 gilt

$$Ax = \sum_{n\in I} \lambda_n (x, e_n) e_n \quad \text{für alle } x \in H.$$

Daraus folgt

$$(Ax, x) = \sum_{n\in I} \lambda_n \left|(x, e_n)\right|^2 \geq 0 \quad \text{für alle } x \in H$$

und insbesondere

$$(Ae_k, e_k) = \lambda_k \geq 0 \quad \text{für alle } k \in I.$$

Definiere $B : H \to H$ durch

$$Bx := \sum_{n\in I} \sqrt{\lambda_n} (x, e_n) e_n, \quad x \in H.$$

B ist wohldefiniert, da $\lambda_n \geq 0$ und B ist auch wieder positiv, denn

$$(Bx, x) = \sum_{n\in I} \sqrt{\lambda_n} \left|(x, e_n)\right|^2 \geq 0 \quad \text{für alle } x \in H.$$

Außerdem ist nach Satz 5.9 $B \in K(H)$ und selbstadjungiert, denn $(\sqrt{\lambda_n})_n \subset \mathbb{R}^+ \setminus \{0\}$ und

$$Be_k = \sum_{n\in I} \sqrt{\lambda_n} \underbrace{(e_k, e_n)}_{=\delta_{kn}} e_n = \sqrt{\lambda_k} e_k,$$

also ist $\sqrt{\lambda_k}$ Eigenwert von B zum Eigenwert e_k.

Weiter gilt $B^2 = A$, denn

$$\begin{aligned} B^2 x &= \sum_{m\in I} \sqrt{\lambda_m} (\sum_{n\in I} \sqrt{\lambda_n} (x, e_n) e_n, e_m) e_m \\ &= \sum_{m\in I} \sum_{n\in I} \sqrt{\lambda_m} \sqrt{\lambda_n} (x, e_n) \underbrace{(e_n, e_m)}_{=\delta_{mn}} e_m \\ &= \sum_{m\in I} \lambda_m (x, e_m) e_m = Ax. \end{aligned}$$

Zur Eindeutigkeit: Angenommen, es existiert ein weiterer positiver und selbstadjungierter Operator $C \in K(H)$ mit

$$C^2 = A.$$

Dann gibt es nach Satz 5.9 eine Darstellung

$$C = \sum_{k\in I} \mu_k (x, e_k) e_k.$$

Weil C positiv ist, gilt

$$\mu_k = (Ce_k, e_k) \geq 0 \quad \text{für alle } k \in I.$$

Weiter folgt

$$\begin{aligned} Ax = C^2 x = C(\sum_{k\in I} \mu_k (x, e_k) e_k) &= \sum_{l\in I} \mu_l (\sum_{k\in I} \mu_k (x, e_k) e_k, e_l) e_l \\ = \sum_{l\in I} \sum_{k\in I} \mu_l \mu_k (x, e_k) \underbrace{(e_k, e_l)}_{=\delta_{kl}} e_l &= \sum_{k\in I} \mu_k^2 (x, e_k) e_k. \end{aligned}$$

Nun ist $(\mu_l^2)_l$ die Folge der Eigenwerte von A, denn

$$Ae_l = \sum_{k\in I} \mu_k^2 \underbrace{(e_l, e_k)}_{=\delta_{kl}} e_k = \mu_l^2 e_l.$$

Daher gilt

$$\mu_k^2 = \lambda_{\sigma(k)} \quad \text{und} \quad e_k = e^{\lambda}_{\sigma(k)}$$

für eine bestimmte Permutation $\sigma(k)$. Aus der unbedingten punktweisen Konvergenz der beiden Reihen für B und C folgt daher $B = C$ (vergleiche die Bemerkung 2. zu Lemma 5.3). $\square$

6 Ausblick: Nichtlineare Funktionalanalysis

6.1 Fixpunktsätze

Im ersten Abschnitt hatten wir den Banachschen Fixpunktsatz rekapituliert. Das ist natürlich nicht der einzige Fixpunktsatz in der Welt der Mathematik. Es gibt noch viele mehr, so wie zum Beispiel

- der Fixpunktsatz von Brouwer oder
- der Fixpunktsatz von Schauder,

die wir in diesem Abschnitt studieren wollen. Der wohl einfachste „Fixpunktsatz“ ist der Folgende:

> Sei $f : [a, b] \to [a, b]$ stetig. Dann besitzt f in $[a, b]$ mindestens einen Fixpunkt.

Der Beweis nutzt den Zwischenwertsatz aus der reellen Analysis aus: Setze

$$g(x) := f(x) - x.$$

Da f eine Selbstabbildung von $[a, b]$ nach $[a, b]$ ist, gilt

$$f(a) \geq a \quad \text{und} \quad f(b) \leq b.$$

Damit folgt

$$g(a) = f(a) - a \geq 0 \quad \text{und} \quad g(b) = f(b) - b \leq 0.$$

g ist auch stetig. Also folgt aus dem Zwischenwertsatz die Exsitenz eines $x_0 \in [a, b]$ mit

$$g(x_0) = 0 = f(x_0) - x_0 \Longrightarrow f(x_0) = x_0.$$

Damit ist x_0 der gesuchte Fixpunkt.

Wir bemerken, dass dieser Satz insbesondere für das Einheitsintervall $[-1, 1]$ gilt. In dieser Situation gilt ein analoges Resulatat in beliebig endlich-dimensionalen Räumen:

Satz 6.1 (Fixpunktsatz von Brouwer, 1912). *Jede stetige Abbildung*

$$\phi : \overline{B_1(0)} \subset \mathbb{R}^N \to \overline{B_1(0)}$$

besitzt mindestens einen Fixpunkt.

Hierbei bezeichnet

$$\overline{B_1(0)} := \{\, x \in \mathbb{R}^N \mid \|x\| \leq 1 \,\}$$

die abgeschlossene Einheitskugel im $\mathbb{R}^N$, $N \geq 1$.

Beweis. Siehe Anhang. □

Bemerkungen.

1. Ein analoges Resulatat gilt in endlich-dimensionalen Banachräumen für Mengen M, die homöomorph[12] zur Einheitskugel sind.

   Dies sind zum Beispiel alle nicht-leeren konvexen kompakten Mengen des $\mathbb{R}^N$ sowie alle nicht-leeren kompakten Mengen im $\mathbb{R}^N$, die sternförmig bzgl. einer abgeschlossenen Kugel sind.

[12] Ein Homöomorphismus ist eine stetige bijektive Abbildung mit stetiger Umkehrabbildung

Beweis. Sei $T : M \to M$ eine stetige Abbildung und

$$h : \overline{B_1(0)} \to M$$

ein Homöomorphismus. Setze

$$\phi := h^{-1} \circ T \circ h : \overline{B_1(0)} \to \overline{B_1(0)}.$$

Dann ist ϕ stetig und mit dem Satz von Brouwer folgt die Existenz eines Fixpunktes x_0 von ϕ, d.h.

$$\phi(x_0) = x_0$$

bzw.

$$h^{-1}(T(h(x_0))) = x_0.$$

Wende h auf beiden Seiten an:

$$T(h(x_0)) = h(x_0).$$

Also ist $h(x_0)$ der gesuchte Fixpunkt von T. □

2. Der Fixpunktsatz gilt **nicht** in unendlich-dimensionalen Räumen (vergleiche das nächste Beispiel 6.1)!

Beispiel 6.1. Betrachte den unendlich-dimensionalen(!) Banachraum $X = l^2(\mathbb{N})$ ausgestattet mit der Norm

$$\|x\|_{l^2} = \sqrt{\sum_{n=0}^{\infty} |x_n|^2}, \quad x = (x_n)_n \in l^2(\mathbb{N}).$$

Betrachte die Abbildung

$$\phi : \overline{B_1(0)} \to \overline{B_1(0)}$$

definiert durch

$$\phi(x) := \phi((x_n)_n) = (\sqrt{1 - \|x\|_{l^2}^2}, x_0, x_1, x_2, \ldots)$$

für alle $x \in \overline{B_1(0)}$. Bemerke, dass für alle $x \in \overline{B_1(0)}$

$$\|\phi(x)\|_{l^2}^2 = (1 - \|x\|_{l^2}^2) + \underbrace{\sum_{n=0}^{\infty} |x_n|^2}_{=\|x\|_{l^2}^2} = 1. \tag{6.1}$$

Dies zeigt, dass ϕ in der Tat wohldefiniert und eine Selbstabbildung der abgeschlossen Einheitskugel von l^2 in sich selbst ist (tatsächlich bildet ϕ sogar in die Sphäre $\{\, x \in l^1(\mathbb{N}) \mid \|x\|_{l^2} = 1 \,\}$ ab).

ϕ ist außerdem stetig, denn ist $(x^{(n)})_n, x \subset l^2(\mathbb{N})$ mit

$$|| \, x^{(n)} - x \, ||_{l^2} \to 0,$$

so folgt

$$|| \, \phi(x^{(n)}) - \phi(x) \, ||_{l^2}^2 = (\sqrt{1 - || \, x^{(n)} \, ||_{l^2}^2} - \sqrt{1 - \|x\|_{l^2}^2})^2 + || \, x^{(n)} - x \, ||_{l^2}^2 \overset{n\to\infty}{\to} 0.$$

Die Abbildung ϕ besitzt aber keinen Fixpunkt. Angenommen, das ist doch der Fall, also gibt es $x \in \overline{B_1(0)}$ mit

$$x = \phi(x).$$

Wegen (6.1) ist also

$$\|x\|_{l^2} = 1. \tag{6.2}$$

Folglich erhalten wir für die Komponenten von x die Gleichungen

$$\begin{aligned} x_0 &= (\phi(x))_0 = \sqrt{1 - \|x\|_{l^2}^1} = 0, \\ x_1 &= (\phi(x))_1 = x_0 = 0, \\ x_2 &= (\phi(x))_2 = x_1 = 0, \\ &\vdots \end{aligned}$$

Sukzessive erhalten wir so $x_n = 0$ für alle $n \in \mathbb{N}$, d.h. $\|x\|_{l^2} = 0$ und somit einen Widerspruch zu (6.2).

□

Im weiteren Verlauf wird sich das folgende Lemma als sehr nützlich erweisen:

Lemma 6.1 (**von Mazur, 1930**). *Sei X ein Banachraum und $M \subset X$ relativ kompakt. Dann ist auch $\operatorname{co} M$ relativ kompakt.*

Beweis. Nach Satz 1.9 ist die relative Kompaktheit von M gleichbedeutend mit der Präkompaktheit von M.

Sei also $\epsilon > 0$. M ist präkompakt und daher gibt es $z_1, \ldots, z_n \in M$ mit

$$M \subset \bigcup_{i=1}^{n} B(z_i, \tfrac{\epsilon}{2}),$$

d.h. für alle $x \in M$ existiert ein $j \in \{1, \ldots, n\}$ mit

$$\|x - z_j\| < \tfrac{\epsilon}{2}. \tag{6.3}$$

Definiere jetzt die Funktion

$$v : M \to \{1, \ldots, n\}, \quad v(x) := j,$$

wobei j der kleinste Index ist, der (6.3) erfüllt.

Ist $y \in \operatorname{co} M$, so gibt es eine Darstellung

$$y = \sum_{i=1}^{m} \alpha_i y_i$$

mit $m = m(y) \in \mathbb{N}$, $\alpha_i \in [0, 1]$, $y_i \in M$, $i = 1, \ldots, m$ und $\sum_{i=1}^{m} \alpha_i = 1$. Es folgt

$$|| \, y - \sum_{i=1}^{m} \alpha_i z_{v(y_i)} \, || = || \sum_{i=1}^{m} \alpha_i (y_i - z_{v(y_i)}) \, || \leq \sum_{i=1}^{m} \alpha_i \, || \, y_i - z_{v(y_i)} \, || < \tfrac{\epsilon}{2}.$$

Wegen

$$\sum_{i=1}^{m} \alpha_i z_{v(y_i)} \in K := \operatorname{co}(z_1, \ldots, z_n)$$

gilt daher

$$\operatorname{co} M \subset \bigcup_{x \in K} B(x, \tfrac{\epsilon}{2}). \tag{6.4}$$

Betrachte jetzt die Funktion

$$\psi : [0, 1]^n \to X, \quad (\alpha_1, \ldots, \alpha_n) \mapsto \sum_{i=1}^{n} \alpha_i z_i.$$

Offensichtlich ist ψ stetig. Setze

$$A := \{ (\alpha_1, \ldots, \alpha_n \in [0,1]^n) \mid \sum_{i=1}^n \alpha_i = 1 \} \subset [0,1]^n.$$

Dann ist A abgeschlossen, denn sei $(\alpha^k)_k \subset A$ eine Folge mit $\alpha^k \to \alpha \in [0,1]^n$. Es folgt

$$|| \sum_{i=1}^n \alpha_i - 1 || \leq || \sum_{i=1}^n \alpha_i - \sum_{i=1}^n \alpha_i^k || + \underbrace{|| \sum_{i=1}^n \alpha_i^k - 1 ||}_{=0} \to 0,$$

also

$$\sum_{i=1}^n \alpha_i = 1$$

und damit ist das Grenzelement wieder in A. Die abgeschlossene Menge A ist auch beschränkt und daher ist A kompakt. Weiter ist $\psi(A) = K$ und weil ψ stetig ist, ist auch K kompakt, d.h. es gibt $k_1, \ldots, k_N \in K$ mit

$$K \subset \bigcup_{i=1}^N B(k_i, \tfrac{\epsilon}{2}),$$

d.h. für alle $x \in K$ existiert ein $i \in \{1, \ldots, N\}$ mit

$$\|x - k_i\| < \tfrac{\epsilon}{2}. \tag{6.5}$$

Sei nun $y \in \operatorname{co} M$ beliebig. Wähle dazu $x \in K$, so dass nach (6.4) gerade $y \in B(x, \frac{\epsilon}{2})$ gilt. Weiter wähle ein $i \in \{1, \ldots, N\}$, so dass (6.5) gilt. Es folgt

$$\|y - k_i\| \leq \|y - x\| + \|x - k_i\| < \epsilon,$$

also ist

$$\operatorname{co} M \subset \bigcup_{i=1}^N B(k_i, \epsilon)$$

und somit $\operatorname{co} M$ präkompakt bzw. relativ kompakt. □

Als nächstes wollen wir geeignete Stetigkeitsbegriffe für nicht-lineare Operatoren einführen:

Definition 6.1 (**Beschränkt, stetig, kompakt**). Seien X, Y Banachräume, $M \subset X$ eine nicht-leere Teilmenge und $A : M \to Y$ ein Operator. Dann heißt A

(i) *beschränkt*, wenn A beschränkte Teilmengen von M in beschränkte Teilmengen von Y abbildet.

(ii) *stetig*, wenn

$$\left\{ \begin{array}{c} (x_n)_n, x \subset M \\ x_n \to x \text{ in } X \end{array} \right\} \Longrightarrow Ax_n \to Ax \text{ in } Y.$$

(iii) *kompakt*, wenn A stetig ist und außerdem A beschränkte Teilmengen von M in relativ kompakte Teilmengen von Y abbildet.

Bemerkungen.

1. Falls $A : X \to Y$ linear ist, dann gilt nach Satz 2.1

$$A \text{ ist beschränkt genau dann, wenn } A \text{ stetig ist.}$$

Für nicht lineare Operatoren gilt diese Äquivalenz nicht: Beschränkte nicht-lineare Operatoren sind natürlich nicht notwendigerweise stetig.

Aber auch die umgekehrte Implikation gilt im Allgemeinen nicht (vergleiche Sie dazu das Beispiel 6.2 unten).

2. Falls $\dim(Y) < \infty$, dann gilt:

$$A \text{ ist stetig und beschränkt genau dann, wenn } A \text{ kompakt ist.}$$

Beweis. Zu „$\Longrightarrow$“: Diese Richtung gilt immer.

Zu „$\Longleftarrow$“: A ist stetig nach Voraussetzung. Es bleibt zu zeigen, dass A beschränkte Mengen in relativ kompakte Mengen abbildet. Sei dazu E eine beschränkte Teilmenge von M. Da A beschränkt ist, ist $A(E)$ eine beschränkte Teilmenge von Y. Dann ist aber $\overline{A(E)}$ beschränkt und abgeschlossen in Y. Da Y endlich-dimensional ist, ist $\overline{A(E)}$ nach dem Satz von Heine-Borel somit kompakt. □

3. Falls $\dim(X) < \infty$ und M eine abgeschlossene Teilmenge von X ist, dann gilt:

$$A \text{ ist stetig genau dann, wenn } A \text{ kompakt ist.}$$

Beweis. Zu „$\Longleftarrow$“: Diese Richtung ist trivial.

Zu „$\Longrightarrow$“: Wir müssen nur zeigen, dass A beschränkte Teilmengen in relativ kompakte Mengen abbildet, d.h. dass aus der Bildfolge jeder beschränkten Folge eine konvergente Teilfolge ausgewählt werden kann. Sei dazu $N \subset M$ beschränkt. Fener sei $(y_n)_n$ eine beliebige beschränkte Folge in $A(N)$. Dann gibt es eine Folge $(x_n)_n \subset N$, so dass

$$y_n = Ax_n \quad \text{für alle } n \in \mathbb{N}.$$

Weil N beschränkt ist, ist auch die Folge $(x_n)_n$ beschränkt und wegen $\dim(X) < \infty$ enthält $(x_n)_n$ daher nach dem Auswahlprinzip von Bolzano-Weierstrass eine konvergente Teilfolge $(x_{n_k})_k$. Sei x der Grenzwert dieser Teilfolge. Da M abgeschlossen ist, gilt $x \in M$. Weil A stetig ist, folgt $Ax_{n_k} \to Ax$, und somit enthält die Bildfolge $(y_n = Ax_n)_n$ eine konvergente Teilfolge. □

4. Falls $\dim(X) = \dim(Y) = \infty$ gilt, dann gibt es stetige beschränkte Abbildungen, die nicht kompakt sind.

Beispiel 6.2. Sei $X = l^2$ ausgestattet mit der Norm

$$\|u\|_{l^2} := \sqrt{\sum_{k=1}^{\infty} |u_k|^2}, \quad u = (u_k)_k \in l^2.$$

Definiere die Abbildung $f : \overline{B_1(0)} \subset l^2 \to l^2$ durch

$$f(u) := (\sum_{k=1}^{\infty} \frac{2^{-k}}{1 + 3^{-k} - u_k}, 0, 0, \ldots).$$

Wir zeigen zunächst, dass

$$\sum_{k=1}^{\infty} \frac{2^{-k}}{1 + 3^{-k} - u_k} < \infty. \tag{6.6}$$

Es ist nämlich

$$\left| \frac{2^{-k}}{1 + 3^{-(k+1)} - u_{k+1}} : \frac{2^{-k}}{1 + 3^{-k-u_k}} \right| = \tfrac{1}{2} \left| \frac{1 + 3^{-k} - u_k}{1 + 3^{-k-1} - u_{k+1}} \right| \to \tfrac{1}{2} < 1,$$

denn[13] $u_k \to 0$, da $u_k \in l^2$. Aus dem Quotientenkriterium folgt also in der Tat (6.6). Außerdem gilt

$$\|f(u)\|_{l^2} = \left| \sum_{k=1}^{\infty} \frac{2^{-k}}{1+3^{-k}-u_k} \right| \overset{(6.6)}{<} \infty,$$

also $f(u) \in l^2$. Insgesamt haben wir damit gezeigt, dass die Abbildung f wohldefiniert ist. f ist auch stetig, aber f ist unbeschränkt. Betrachtet man nämlich die beschränkte Menge $\{ e_n \mid n \in \mathbb{N} \}$ von Elementen

$$e_n := (0, \dots, 0, 1, 0 \dots,) \in \overline{B_1(0)} \subset l^2,$$

so gilt

$$\|f(e_n)\|_{l^2} = \sum_{k=1}^{\infty} \frac{2^{-k}}{1+3^{-k}-\delta_{n_k}} \geq \frac{2^{-n}}{1+3^{-n}-1} = \left(\frac{3}{2}\right)^n \to \infty \quad \text{für} \quad n \to \infty.$$

Also ist das Bild $f((e_n))$ unbeschränkt.

□

Anhand von Beispiel 6.1 haben wir gesehen, dass der Brouwersche Fixpunktsatz nur in endlich-dimensionalen Räumen funktioniert. In unendlich-dimensionalen Räumen X kann die Existenz eines Fixpunktes einer stetigen Selbstabbildung $\phi : M \subset X \to X$ nur unter einer zusätzlichen Kompaktheitsbedingung nachgewiesen werden:

Satz 6.2 (**Fixpunktsatz von Schauder, 1. Version (1930)**). *Es seien $(X, \|.\|_X)$ ein Banachraum sowie $M \subset X$ eine nicht-leere konvexe und kompakte Menge. Dann besitzt jede stetige Abbildung $\phi : M \to M$ mindestens einen Fixpunkt in M.*

Bemerkung. Der Satz von Brouwer ist quasi ein Spezialfall des Satzes von Schauder, da die abgeschlossene Einheitskugel in einem endlich-dimensionalen Banachraum stets nicht-leer, konvex und kompakt ist.

Alternativ kann die Kompaktheitsbedingung nicht in die Menge M, sondern in den „Operator gesteckt werden“:

Satz 6.3 (**Fixpunktsatz von Schauder, 2. Version**). *Es seien $(X, \|.\|_X)$ ein Banachraum sowie $M \subset X$ eine nicht-leere konvexe, beschränkte und abgeschlossene Menge. Dann besitzt jede stetige kompakte Abbildung $\phi : M \to M$ mindestens einen Fixpunkt in M.*

Bemerkungen.

1. Falls $\dim(X) = \infty$, dann sind abgeschlossene und beschränkte Mengen nach Beispiel 1.23 nicht notwendigerweise kompakt.

2. Beide Versionen des Schauderschen Fixpunktsatzes sind äquivalent!

   Beweis. Zu „$\Longrightarrow$“: Das ist leicht, denn eine stetige Abbildung $\phi : M \to M$ mit kompakten $\overline{M}$ ist eine kompakte Abbildung.

   Zu „$\Longleftarrow$“: Zunächst ist M beschränkt und ϕ kompakt, so dass $\overline{\phi(M)}$ kompakt ist. Nach dem Lemma von Mazur ist dann auch

   $$N := \overline{\operatorname{co} \phi(M)}$$

[13] Ist $u_k \in l^2$, so ist $\sum_{k=1}^{\infty} |u_k|^2 < \infty$ und wegen der absoluten Konvergenz dann auch $\sum_{k=1}^{\infty} u_k^2 < \infty$. Daher muss notwendigerweise $u_k^2 \to 0$ und so auch $u_k \to 0$

kompakt.

Weiter ist $\operatorname{co}\phi(M) \subset M$. Das zeigen wir per Induktion:

<u>$n = 1$:</u> Dann ist $x_1 \in \phi(M) \subset M$.

<u>$n \to n+1$:</u> Seien $x_1, \ldots, x_{n+1} \in \phi(M) \subset M$ und $\alpha_1, \ldots, \alpha_{n+1} \in [0,1]$ mit $\sum \alpha_i = 1$. Dann gilt

$$\alpha_1 x_1 + \ldots \alpha_{n+1} x_{n+1} = \alpha_1 x_1 + (1-\alpha_1) \underbrace{\left[\tfrac{\alpha_2}{1-\alpha_1} x_2 + \ldots + \tfrac{\alpha_{n+1}}{1-\alpha_1} x_{n+1}\right]}_{=:p}.$$

Da M konvex ist, reicht es also zu zeigen, dass $p \in M$: Es ist

$$p = \underbrace{\tfrac{\alpha_2}{\alpha_2+\ldots+\alpha_{n+1}}}_{\in[0,1]} x_2 + \ldots + \underbrace{\tfrac{\alpha_{n+1}}{\alpha_2+\ldots+\alpha_{n+1}}}_{\in[0,1]} x_{n+1}$$

und

$$\tfrac{\alpha_2}{\alpha_2+\ldots+\alpha_{n+1}} + \ldots + \tfrac{\alpha_{n+1}}{\alpha_2+\ldots+\alpha_{n+1}} = 1.$$

Nach Induktionsvoraussetzung ist also $p \in M$.

Damit haben wir $\operatorname{co}\phi(M) \subset M$ gezeigt, aber es gilt sogar $N \subset M$, denn

$$N = \overline{\operatorname{co}\phi(M)} \subset \bar{M} = M.$$

Weiter bildet ϕ die Menge N in sich ab, denn

$$N \subset M \Longrightarrow \phi(N) \subset \phi(M) \Longrightarrow \phi(N) \subset \overline{\operatorname{co}\phi(N)} \subset \overline{\operatorname{co}\phi(M)} = N.$$

Die Abbildung $\phi : N \to N$ ist stetig, denn Einschränkungen stetiger Abbildungen sind immer stetig. Weiter ist N kompakt, konvex und nicht-leer, denn $M \neq \emptyset$.

Nach Satz 6.2 besitzt $\phi|_N$ mindestens einen Fixpunkt $\hat{x} \in N$. Damit ist $\hat{x}$ auch Fixpunkt von ϕ. □

Um die 1. Version des Fixpunktsatzes zu beweisen, benötigen wir die Tatsache, dass man kompakte Operatoren durch „endlich-dimensionale“ Operatoren approximieren kann:

Lemma 6.2 (Schauder-Operatoren). *Seien $(X, \|.\|_X)$ ein Banachraum, $M \subset X$ nicht-leer und beschränkt sowie $\phi : M \to X$ kompakt. Dann existiert für alle $n \in \mathbb{N}$ ein kompakter Operator $\phi_n : M \to X$ mit folgenden Eigenschaften:*

(i) $\|\phi(x) - \phi_n(x)\| \leq \frac{1}{n}$ *für alle $x \in M$.*

(ii) $\dim(\operatorname{lin}(\phi_n(M))) < +\infty$.

(iii) $\phi_n(M) \subset \operatorname{co}\phi(M)$.

Die Operatoren ϕ_n heißen Schauder-Operatoren.

Beweis. Da M beschränkt und ϕ kompakt ist, ist $\phi(M)$ relativ kompakt, also nach Satz 1.10 präkompakt, weil X ein Banachraum ist. Folglich enthält, für beliebiges $n \in \mathbb{N}$, die Überdeckung von $\phi(M)$ durch die Familie der offenen Mengen

$$\bigcup_{x \in M} B(\phi(x), \tfrac{1}{n})$$

eine endliche Teilüberdeckung, d.h. es existieren $x_1, \ldots, x_{m_n} \in M$ mit

$$\phi(M) \subset \bigcup_{i=1}^{m_n} B(\phi(x_i), \tfrac{1}{n}). \tag{6.7}$$

Definiere nun für alle $n \in \mathbb{N}$ die Abbildung $\phi_n : M \to X$ durch

$$\phi_n(x) := \frac{\sum_{i=1}^{m_n} \alpha_i(x)\phi(x_i)}{\sum_{i=1}^{m_n} \alpha_i(x)}, \quad x \in M,$$

wobei

$$\alpha_i(x) := \max\{\tfrac{1}{n} - \|\phi(x) - \phi(x_i)\|, 0\} \tag{6.8}$$

für alle $i = 1, \ldots, m_n$.

Dann gilt:

- ϕ_n ist wohldefiniert, da wegen (6.7) mindestens ein $\alpha_i(x) > 0$ und somit

$$\sum_{i=1}^{m_n} \alpha_i(x) > 0 \quad \text{für alle } x \in M. \tag{6.9}$$

- ϕ_n ist stetig, da ϕ und alle α_i stetig sind und (6.9) gilt.
- Offensichtlich gilt

$$\phi(M) \subset \operatorname{lin}\{\phi(x_1), \ldots, \phi(x_{m_n})\}.$$

  Daraus folgt

$$\dim \operatorname{lin} \phi_n(M) \leq m_n < \infty$$

  für alle $n \in \mathbb{N}$ und damit ist (ii) gezeigt.

- Für alle $x \in M$ gilt

$$\begin{aligned}
\|\phi_n(x) - \phi(x)\| &= \| \phi_n(x) - \frac{\sum^{m_n} \alpha_i(x)\phi(x_i)}{\sum_{i=1}^{m_n} \alpha_i(x)} \| \\
&= \frac{1}{\sum_{i=1}^{m_n} \alpha_i(x)} \| \sum_{i=1}^{m_n} \alpha_i(x)(\phi(x_i) - \phi(x)) \| \\
&\leq \frac{\sum_{i=1}^{m_n} \alpha_i(x)\|\phi(x_i) - \phi(x)\|}{\sum_{i=1}^{m_n} \alpha_i(x)} \overset{(6.7)}{\leq} \frac{\sum_{i=1}^{m_n} \alpha_i(x)\frac{1}{n}}{\sum_{i=1}^{m_n} \alpha_i(x)} = \frac{1}{n}.
\end{aligned}$$

  Damit ist auch (i) gezeigt.

- Für alle $x \in M$ gilt nach dem eben Gezeigten:

$$\|\phi_n(x)\| \leq \|\phi_n(x) - \phi(x)\| + \|\phi(x)\| \leq \tfrac{1}{n} + \|\phi(x)\|.$$

  Weil $\phi(M)$ realtiv kompakt, d.h. insbesondere beschränkt ist, folgt sofort, dass auch $\phi_n(M)$ beschränkt ist.

  Da $\phi_n(M)$ eine Teilmenge des endlich-dimensionalen Raumes $\operatorname{lin}\{\phi(x_1), \ldots, \phi(x_{m_n})\}$ ist, folgt sofort, dass $\phi_n(M)$ relativ kompakt ist. Damit ist ϕ_n kompakt.

- Weil

$$\phi_n(x) = \sum_{i=1}^{m_n} \lambda_i(x)\phi(x_i)$$

  mit

$$\lambda_i(x) := \frac{\alpha_i(x)}{\sum_{i=1}^{m_n} \alpha_i(x)}$$

und $\lambda_i(x) \in [0,1]$ sowie

$$\sum_{i=1}^{m_n} \lambda_i = 1,$$

folgt sofort

$$\phi_n(M) \subset \operatorname{co} \phi(M).$$

Damit ist schließlich auch (iii) gezeigt.

□

Beweis von Satz 6.2. Weil $\phi : M \to M$ stetig und M kompakt ist, ist ϕ kompakt, denn abgeschlossene Mengen von kompakten Mengen sind wieder kompakt.

Nach Lemma 6.2 kann ϕ also durch endlich-dimensionale kompakte Operatoren ϕ_n approximiert werden. Diese seien wir im Beweis des Lemmas definiert. Für $n \in \mathbb{N}$ sei nun

$$M_n := \overline{\operatorname{co}\{\phi(x_1), \ldots, \phi(x_{m_n})\}}.$$

Wieder, da M konvex und abgeschlossen und $\phi : M \to M$ eine Selbstabbildung ist, folgt

$$M_n \subset M \quad \text{für alle } n \in \mathbb{N}.$$

Da M beschränkt ist, ist dann auch M_n für alle $n \in \mathbb{N}$ beschränkt. Weiter ist M_n konvex.

Weil $M_n \subset \operatorname{lin}\{\phi(x_1), \ldots, \phi(x_{m_n})\}$, also M_n in einem endlich-dimensionalen Teilraum von X liegt, folgt weiter, dass M_n kompakt ist.

Wieder ist $\phi_n(M_n) \subset M_n$ nach Lemma 6.2 (iii).

Zusammenfassend ist M_n also konvex, nicht-leer, kompakt und endlich-dimensional. Nach der Bemerkung zum Satz von Brouwer ist daher M_n homöomorph zur abgeschlossenen Kugel $\overline{B_1(0)}$.

Es folgt mit dem Satz von Brouwer, dass die stetige Abbildung $\phi_n : M_n \to M_n$ mindestens einen Fixpunkt $x_n \in M_n \subset M$ besitzt, d.h.

$$\phi(x_n) = x_n.$$

Da M beschränkt ist, ist die Folge $(x_n)_n \subset M$ beschränkt. Nun ist ϕ kompakt und daher existiert eine Teilfolge $(x_{n_k})_k$, so dass

$$\phi(x_{n_k}) \overset{k\to\infty}{\longrightarrow} y \quad \text{in } M.$$

Daher

$$\begin{aligned} || \phi_{n_k}(x_{n_k}) - y || &\leq || \phi_{n_k}(x_{n_k})\phi(x_{n_k}) || + || \phi(x_{n_k}) - y || \\ &\overset{\text{Lemma 6.2}}{\leq} \tfrac{1}{n_k} + || \phi(x_{n_k}) - y || \overset{k\to\infty}{\longrightarrow} 0, \end{aligned}$$

d.h.

$$x_{n_k} = \phi_{n_k}(x_{n_k}) \overset{k\to\infty}{\longrightarrow} y.$$

Somit folgt aus der Stetigkeit von ϕ

$$\phi(y) = \lim_{k\to\infty} \phi(x_{n_k}) = y,$$

also ist y Fixpunkt von ϕ. □

6.2 Anwendungen des Schauderschen Fixpunktsatzes

- Anwendungen des Fixpunktsatzes auf Integralgleichungen und Differentialgleichungen
- Leray-Schauder-Prinzip

Beispiel 6.3. Seien $k \in C([a,b]^2 \times \mathbb{R})$, $\lambda \in \mathbb{R}^+$ und $f \in C[a,b]$. Betrachte die nicht-lineare Integralgleichung

$$u(x) - \lambda \int_a^b k(x,y,u(y))dy = f(x), \quad x \in [a,b] \tag{6.10}$$

Für alle $R > 0$ existiert dann mindestens eine Lösung $u \in C[a,b]$ von (6.10) mit

$$\|u\|_\infty = \sup_{x\in[a,b]} |u(x)| \leq R + \|f\|_\infty,$$

falls

$$\lambda(b-a) \max_{\substack{(x,y)\in[a,b]^2 \\ r\leq R+\|f\|_\infty}} |k(x,y,r)| \leq R, \tag{6.11}$$

d.h. $\lambda \geq 0$ muss klein genug sein.

Beweis. Schreibe das Problem als Fixpunkgleichung:

$$u = \phi(u), \quad u \in C[a,b],$$

wobei

$$\phi(u)(x) := f(x) + \lambda \int_a^b k(x,y,u(y))dy, \quad x \in [a,b],\ u \in C[a,b].$$

Sei $R > 0$ und λ gemäß (6.11) gewählt. Definiere

$$M := \{\, u \in C[a,b] \mid \|u\|_\infty \leq R + \|f\|_\infty \,\}.$$

Zunächst ist $M \neq \emptyset$, denn $f \in M$. Offensichtlich ist M auch beschränkt. Aber M ist auch abgeschlossen, denn ist $(u_k)_k \subset M$ mit $u_k \to u$, so folgt nämlich

$$\|u\|_\infty \leq \|u_k\|_\infty + \|u_k - u\|_\infty \leq R + \|f\|_\infty + \underbrace{\|u_k - u\|_\infty}_{\to 0},$$

also $u \in M$.

Weiter ist M konvex: Seien dazu $u, v \in M$ und $\lambda \in [0,1]$. Dann gilt

$$\begin{aligned}\|\lambda u + (1-\lambda)v\|_\infty &\leq \lambda\|u\|_\infty + (1-\lambda)\|v\|_\infty \\ &\leq \lambda(R + \|f\|_\infty) + (1-\lambda)(R + \|f\|_\infty) = R + \|f\|_\infty,\end{aligned}$$

also $\overline{uv} \subset M$.

Die Abbildung $\phi|_M : M \to M$ ist in der Tat eine Selbstabbildung, denn auf der einen Seite gilt

$$\begin{aligned}\|\phi(u)\|_\infty &= \|\, f + \lambda \int_a^b k(.,y,u(y))dy \,\|_\infty \\ &\leq \|f\|_\infty + \lambda(b-a) \max_{\substack{(x,y)\in[a,b]^2 \\ r\leq R+\|f\|_\infty}} |k(x,y,r)| \leq \|f\|_\infty + R\end{aligned}$$

und auf der anderen Seite ist $\phi(u) \in C[a,b]$. k ist nämlich auf der kompakten Menge $[a,b]^2 \times [-R - \|f\|_\infty, R + \|f\|_\infty]$ gleichmäßig stetig. Gleiches gilt für f auf $[a,b]$. Daher kann man zu $\epsilon > 0$ ein $\delta = \delta(\epsilon, R) > 0$ wählen, so dass

$$\begin{aligned}&\max\{|x_2 - x_1|, |y_2 - y_1|, |v_2 - v_1| \leq \delta\},\ |v_1|, |v_2| \leq R + \|f\|_\infty \\ \Longrightarrow\ &|k(x_1,y_1,v_1) - k(x_2,y_2,v_2)| \leq \tfrac{\epsilon}{2\lambda(b-a)} \text{ und } |f(x_1) - f(x_2)| \leq \tfrac{\epsilon}{2}.\end{aligned} \tag{6.12}$$

Ist also $u \in M$ und $\|x_1 - x_2\| \leq \delta$, so folgt sofort

$$|(\phi u)(x_1) - (\phi u)(x_2)| \leq |f(x_1) - f(x_2)| + \int_a^b \lambda \, |k(x_1, y, u(y)) - k(x_2, y, u(y))| \, dt \leq \epsilon.$$

Damit ist $\phi(u)|_M$ stetig. Dieses Argument zeigt uns außerdem, dass für eine beschränkte Teilmenge $K \subset M$ die Teilmenge $\phi(K)$ gleichradig stetig ist. Wegen $\phi(K) \subset M$ ist klar, dass $\phi(K)$ punktweise beschränkt ist. Nach dem Satz von Arzelà-Ascoli ist also $\phi(K)$ relativ kompakt.

Um die Kompaktheit von $\phi|_M$ nachzuweisen, ist nur noch die Stetigkeit von ϕ zu zeigen. Es genügt dafür zu zeigen, dass die gesamte Abbildung stetig[14] ist. Dazu seien $\epsilon > 0$ und $u \in C[a, b]$. Wähle $R \geq 0$ so, dass $\|u\|_\infty \leq \|f\|_\infty + R - 1$ und wähle $\delta = \delta(\epsilon, R)$ wie oben. Sei

$$\delta' := \min\{\delta, 1\} \quad \text{und} \quad \|u - v\|_\infty \leq \delta'.$$

Dann ergibt sich für alle $x \in [a, b]$

$$|(\phi u)(x) - (\phi v)(x)| \leq \int_a^b |k(x, y, u(y)) - k(x, y, v(y))| \, dt \leq \tfrac{\epsilon}{2} \leq \epsilon,$$

denn

$$\|u - v\|_\infty \leq \delta' \leq \delta$$

sowie $\|u\|_\infty \leq R + \|f\|_\infty$ und

$$\|v\|_\infty \leq \|v - u\|_\infty + \|u\|_\infty \leq \delta' + R + \|f\|_\infty - 1 \leq 1 + R + \|f\|_\infty - 1 = R + \|f\|_\infty.$$

Da x beliebig war, folgt schließlich

$$\|\phi u - \phi v\|_\infty \leq \epsilon$$

und das zeigt die Stetigkeit von ϕ.

Mit der 2. Version des Fixpunktsatzes von Schauder folgt nun, dass $\phi|_M$ einen Fixpunkt $u \in M$ besitzt. Damit ist (6.10) gelöst.

Bemerkungen.

1. Falls k beschränkt auf $[a, b]^2 \times \mathbb{R}$ ist, so folgt nach dem eben gezeigten, dass für alle $R > 0$ die Gleichung (6.10) eine Lösung $u \in C[a, b]$ besitzt, falls nur

   $$\lambda \leq \frac{R}{(b - a)\|k\|_\infty}.$$

   Gilt jetzt sogar $R \to \infty$, so folgt, dass (6.10) eine Lösung $u \in C[a, b]$ für beliebiges $\lambda \in \mathbb{R}^+$ besitzt. Speziell ist also

   $$\begin{cases} u \in C[a, b] \\ u(x) - \lambda \int_a^b \sin(u(y)) dy = f(x), \quad x \in [a, b] \end{cases}$$

   lösbar für alle $f \in C[a, b]$ und $\lambda \in \mathbb{R}^+$.

2. Der *Hammerstein-Operator* ist gegeben durch

   $$\phi(u)(x) = \int_a^b k(x, y) f(y, u(y)) dt.$$

   Sind $k \in C([a, b]^2 \times \mathbb{R})$ und $f \in C([0, 1] \times \mathbb{R})$, so handelt es sich um einen Spezialfall von Beispiel 6.3.

□

[14] Wir zeigen hier gleich die Stetigkeit der ganzen Abbildung. Eigentlich folgt die Stetigkeit von $\phi|_M$ schon aus der oberen $\epsilon - \delta$-Geschichte

Beispiel 6.4. Eine weitere Anwendung des Schauderschen Fixpunktsatzes lässt sich in der Lehre der Differentialgleichungen finden. So wird der Fixpunktsatz etwa dazu benutzt, um die Existenz von geöhnlichen Differentialgleichungen $y' = f(x, y)$ bzw. Anfangswertproblemen

$$\begin{cases} y' = f(x, y) \\ y(x_0) = y_0 \end{cases}$$

mit stetiger rechter Seite f nachzuweisen. Man kommt so zum *Satz von Peano*. Wer mehr darüber erfahren möchte, sollte den Kurs *Differentialgleichungen 1* besuchen.

□

Korollar 6.1 (Leray-Schauder-Prinzip). *Seien $(X, \|.\|_X)$ ein Banachraum und $\phi : X \to X$ kompakt mit der Eigenschaft, dass es ein $R > 0$ gibt, so dass für alle $t \in [0, 1]$ und für alle Lösungen $u \in X$ von*

$$u = t\phi(u)$$

die Abschätzung

$$\|u\|_X \leq R$$

gilt. Dann besitzt ϕ einen Fixpunkt.

Bemerkungen.

1. Die Eigenschaft an die Abbildung ϕ bedeutet, dass man nur zeigen muss, dass eine etwaige Lösung von $u = t\phi(u)$ der Abschätzung $\|u\| \leq R$ genügt, falls eine solche existiert. Man muss nicht zeigen, dass eine solche existiert.

2. Dieses Beweis-Prinzip heißt *a-priorie-Abschätzung*.

Beweis. Definiere

$$M := \{\, u \in X \mid \|u\|_X \leq 2R \,\} \neq \emptyset.$$

M ist abgeschlossen, beschränkt und konvex. Setze

$$\tilde{\phi}(u) := \begin{cases} \phi(u) & \text{falls } \|\phi(u)\| \leq 2R, \\ 2R \cdot \frac{\phi(u)}{\|\phi(u)\|} & \text{sonst.} \end{cases}$$

Dann ist $\tilde{\phi} : M \to M$ wohldefiniert, stetig und beschränkt. Ja, $\tilde{\phi}$ ist sogar kompakt, denn ist $(u_n)_n \subset M$ eine beschränkte Folge, so existiert entweder eine Teilfolge $(u_{n_k})_k$ mit

a) $\|\phi(u_{n_k})\| \leq 2R$ für alle $k \in \mathbb{N}$. Wegen der Kompaktheit von ϕ folgt die Existenz einer Teilfolge $(u_{n_{k_l}})_l$ mit

$$\phi(u_{n_{k_l}}) = \tilde{\phi}(u_{n_{k_l}}) \to y$$

Oder aber

b) $\|\phi(u_{n_k})\| > 2R$ für alle $k \in \mathbb{N}$. Weil ϕ kompakt ist, existiert wieder eine Teilfolge $(u_{n_{k_l}})_l$ mit

$$\phi(u_{n_{k_l}}) \to y \neq 0.$$

Beachten Sie, das in der Tat $y \neq 0$, weil $\|\phi(u_{n_k})\| > 2R > 0$. Es folgt dann

$$\tilde{\phi}(u_{n_{k_l}}) = \frac{2R\phi(u_{n_{k_l}})}{\|\,\phi(u_{n_{k_l}})\,\|} \to \frac{2Ry}{\|y\|}.$$

Zusammenfassend besitzt $\tilde{\phi}$ nach dem Schauderschen Fixpunktsatz einen Fixpunkt $u \in M$. 2 Fälle sind möglich:

<u>1. Fall:</u> $\|\phi(u)\| \leq 2R$. Dann ist

$$u = \tilde{\phi}(u) = \phi(u)$$

und wir sind fertig.

<u>2. Fall:</u> $\|\phi(u)\| \geq 2R$. Dann folgt

$$u = \tilde{\phi}(u) = \frac{2R\phi(u)}{\|\phi(u)\|} = \underbrace{\frac{2R}{\|\phi(u)\|}}_{=:t\in[0,1]} \cdot \phi(u).$$

Es folgt

$$\|u\| = 2R. \tag{6.13}$$

Auf der anderen Seite folgt aus

$$u = t\phi(u), \quad t \in [0,1]$$

gemäß der a-priori-Abschätzung gerade $\|u\| \leq R$. Doch das ist ein Widerspruch zu (6.13), d.h. der 2. Fall taucht gar nicht erst auf. $\square$

6.3 Weitere Fixpunktsätze

- Kuratowski-Maß der Nicht-Kompaktheit einer Menge
- Verdichtende Abbildungen
- Uniform konvexe Räume
- Fixpunktsatz von Darbo-Sadovskiĭ
- Fixpunktsatz von Browder-Göhde-Kirk
- Fixpunktsatz von Krasnoselskiĭ

Der *Durchmesser einer Menge* $S \subset X$ ist gegeben durch

$$\operatorname{diam}(S) = \sup\{\, \|x - y\| \mid x, y \in S \,\}.$$

Lemma 6.3. *Sei $S \subset X$ beschränkt. Dann gilt*

(i) $\operatorname{diam}(S) = \operatorname{diam}(\bar{S})$.

(ii) $\operatorname{diam}(\operatorname{co} S) = \operatorname{diam}(S)$.

Beweis. Zu (i): Wegen $S \subset \bar{S}$ gilt zunächst $\operatorname{diam}(S) \leq \operatorname{diam}(\bar{S})$. Wir müssen also nur noch $\operatorname{diam}(\bar{S}) \leq \operatorname{diam}(S)$ zeigen. Angenommen, $\operatorname{diam}(S) < \operatorname{diam}(\bar{S})$. Dann gibt es $\bar{x}, \bar{y} \in \bar{S}$, so dass

$$\epsilon := \|\bar{x} - \bar{y}\| - \operatorname{diam}(S) > 0.$$

Seien nun $x, y \in S$ mit

$$\|\bar{x} - x\| < \tfrac{\epsilon}{2} \quad \text{und} \quad \|\bar{y} - y\| < \tfrac{\epsilon}{2}.$$

Dann folgt

$$\begin{aligned} \|x - y\| &= \|x - \bar{x} + \bar{x} - \bar{y} + \bar{y} - y\| \geq \|\bar{x} - \bar{y}\| - \|x - \bar{x}\| - \|y - \bar{y}\| \\ &> \epsilon + \operatorname{diam}(S) - \epsilon = \operatorname{diam}(S). \end{aligned}$$

Doch das ist ein Widerspruch.

Zu (ii): Wegen $S \subset \operatorname{co} S$ gilt zunächst $\operatorname{diam}(S) \leq \operatorname{diam}(\operatorname{co} S)$. Wir müssen also nur noch $\operatorname{diam}(\operatorname{co} S) \leq \operatorname{diam}(S)$ zeigen. Dazu seien $x, y \in \operatorname{co} S$. Dann gibt es folgende Darstellungen

$$x = \sum_{i=1}^{p} \lambda_i x_i, \quad \text{und} \quad y = \sum_{k=1}^{r} \mu_k y_k.$$

Damit folgt

$$\begin{aligned} \|x - y\| &= \left\| \sum_{i=1}^{p} \lambda_i - y \right\| = \left\| \sum_{i=1}^{p} \lambda_i - y \sum_{i=1}^{p} \lambda_i \right\| \\ &= \left\| \sum_{i=1}^{p} \lambda_i (x_i - y) \right\| = \left\| \sum_{i=1}^{p} \lambda_i \sum_{k=1}^{r} \mu_k (x_i - y_k) \right\| \\ &\leq \sum_{i=1}^{p} \lambda_i \sum_{k=1}^{r} \mu_k \|x_i - y_i\| \leq \sum_{i=1}^{p} \lambda_i \sum_{k=1}^{r} \mu_k \operatorname{diam}(S) = \operatorname{diam}(S). \end{aligned}$$

□

Definition 6.2 (Kuratowskisches Nichtkompaktheitsmaß). Seien $(X, \|.\|_X)$ ein Banachraum und $M \subset X$ eine beschränkte Menge. Dann heißt

$$\gamma(M) := \inf\{\, \epsilon > 0 \mid M \subset \bigcup_{i=1}^{n} B_i,\ n \in \mathbb{N},\ \operatorname{diam}(B_i) \leq \epsilon,\ i = 1, \ldots, n \,\}$$

das *(Kuratowski-)Maß der Nichtkompaktheit von M.*

Bemerkungen.

1. Ohne Einschränkung kann man annehmen, dass die überdeckenden Mengen aus der Definition abgeschlossen sind, denn es gilt nach Lemma 6.3 stets $\operatorname{diam} A = \operatorname{diam} \bar{A}$.
2. Für kompaktes M gilt $\gamma(M) = 0$. Die offene Überdeckung $\bigcup_{x \in M} B(x, \frac{\epsilon}{2})$ besitzt nämlich eine endliche Teilüberdeckung. Eine Menge M ist also umso weniger kompakt, je größer $\gamma(M)$ ist. Das gerechtfertigt auch den Namen „Nichtkompaktheitsmaß“ für γ.

Beispiel 6.5. Sei X ein unendlichen-dimensionaler(!) Banachraum. Nach Satz 4.11 ist dann die abgeschlossene Einheitskugel B_X nicht kompakt. Es gilt $\gamma(B_X) = 2$.

Beweis. Da sich jede Menge selbst überdeckt, gilt

$$\gamma(B_X) \leq \operatorname{diam} B_X = 2.$$

Um die Gleichheit zu zeigen, brauchen wir ein Resultat aus der Topologie, nämlich den Satz von Lyusternik-Shnirelman-Borsuk: Angenommen, es gilt $\gamma(M) < 2$. Dann gibt es endlich viele abgeschlossene Mengen $A_1, \ldots, A_n$ mit $\operatorname{diam} A_i < 2$, so dass

$$B_x \subset A_1 \cup \ldots \cup A_n.$$

Sei jetzt $E \subset X$ mit $\dim E = n < \infty$. Setze

$$\tilde{A}_j := A_j \cap S_E.$$

Offensichtlich sind die $\tilde{A}_j$ abgeschlossene Teilmengen der Sphäre S_E und überdecken dieselbe. Der Satz von Lyusternik-Shnirelman-Borsuk impliziert nun, dass eine dieser Mengen, sagen wir $\tilde{A}_{j_0}$, einen Punkt x_0 samt $-x_0$ (*Antipode*) enthält. Daraus erhalten wir den Widerspruch

$$2 = \|x_0 - (-x_0)\| \leq \operatorname{diam}(\tilde{A}_{j_0}) \leq \operatorname{diam}(A_{j_0}) < 2.$$

□

Eigenschaften des Maßes.

1. $\gamma(\emptyset) = 0$.
2. $\gamma(M) = 0$ genau dann, wenn M relativ kompakt (=präkompakt) ist.
3. $\gamma(\beta M) = |\beta|\, \gamma(M)$ für alle $\beta \in \mathbb{K}$.
4. $\gamma(M_1 + M_2) \leq \gamma(M_1) + \gamma(M_2)$.
5. Ist $N \subset M$, so folgt $\gamma(N) \subset \gamma(M)$.
6. $\gamma(M_1 \cup M_2) = \max\{\gamma(M_1), \gamma(M_2)\}$.
7. $\gamma(M) = \gamma(\bar{M})$.
8. $\gamma(M) = \gamma(\operatorname{co}(M)) = \gamma(\overline{\operatorname{co}(M)})$
9. Für eine absteigende Folge abgeschlossener Mengen $M_1 \supset M_2 \supset \ldots \neq \emptyset$ mit $\gamma(M_n) \to 0$ ist $\bigcap_{n=1}^{\infty} M_n$ kompakt und nicht-leer.

Beweis. Zu 1: Trivial.

Zu 2: Zunächst gilt wegen 7. $\gamma(M) = \gamma(\bar{M})$ und mit der oberen Bemerkung folgt die Behauptung.

Zu 3-6: Das überlass ich Ihnen.

Zu 7: Klar nach Lemma 6.3 (i).

Zu 8: Wegen 7. und 5. müssen wir nur $\gamma(\operatorname{co} M) \leq \gamma(M)$ zeigen. Sei also $\epsilon_0 > \gamma(M)$. Dann gibt es eine endliche Überdeckung

$$M \subset \bigcup_{j=1}^{n} A_j$$

durch Mengen A_j vom Durchmesser

$$\operatorname{diam} A_j \leq \epsilon_0.$$

Nach Lemma 6.3 (ii) können wir sogar die A_i als konvex annehmen. Nun bemerken wir, dass

$$\begin{aligned}\operatorname{co} M &\subset \operatorname{co}\left(A_1 \cup \bigcup_{j=2}^{n} A_j\right) \subset \operatorname{co}\left(A_1 \cup \operatorname{co}\bigcup_{j=2}^{n} A_j\right)\\ &\subset \operatorname{co}\left(A_1 \cup \operatorname{co}\left(A_2 \cup \operatorname{co}\bigcup_{j=3}^{n} A_j\right)\right) \subset \text{ usw.}\end{aligned}$$

Zwischenbehauptung: Für konvexe Mengen C_1 und C_2 und deren konvexe Hülle $C = \operatorname{co}(C_1 \cup C_2)$ gilt

$$\gamma(C) \leq \max\{\gamma(C_1), \gamma(C_2)\}. \tag{6.14}$$

Dann folgt nämlich aus der obigen Implikationskette gerade $\gamma(\operatorname{co} M) \leq \epsilon_0$ und daher $\gamma(\operatorname{co} M) \leq \gamma(M)$.

Beweis von (6.14): Weil C_1 und C_2 beschränkt sind, gibt es ein $R > 0$ mit

$$\|x\| \leq R \quad \text{für alle } x \in C_1 \cup C_2.$$

Jedes $x \in C$ kann als

$$x = \lambda x_1 + (1-\lambda)x_2$$

mit geeigneten $x_1 \in C_1$, $x_2 \in C_2$ und $0 \leq \lambda \leq 1$ dargestellt werden.

Sei jetzt $N \in \mathbb{N}$ beliebig und $\lambda_k := \frac{k}{N}$. Wähle dann zu λ ein λ_k mit

$$|\lambda_k - \lambda| \leq \tfrac{1}{N}.$$

Setze

$$z := (\lambda - \lambda_k)x_1 + (\lambda_k - \lambda)x_2.$$

Dann kann x als

$$x = \lambda_k x_1 + (1-\lambda_k)x_2 + z$$

geschrieben werden, wobei $\|z\| \leq \frac{2R}{N}$ gilt. Daraus folgt weiter

$$C \subset \bigcup_{k=0}^{N} (\lambda_k C_1 + (1-\lambda_k)C_k) + \frac{2R}{N} B_X.$$

Schließlich gilt

$$\begin{aligned}\gamma(C) &\overset{4.,5.}{\leq} \gamma\Big(\bigcup_{k=0}^{N} (\lambda_k C_1 + (1-\lambda_k)C_k)\Big) + \gamma\Big(\frac{2R}{N} B_X\Big)\\ &\overset{3.}{=} \gamma\Big(\bigcup_{k=0}^{N} (\lambda_k C_1 + (1-\lambda_k)C_k)\Big) + \frac{2R}{N}\underbrace{\gamma(B_X)}_{\overset{\text{Bsp. 6.5}}{=} 2}\\ &\overset{6.,3.,4.}{\leq} \max_{0\leq k\leq n} (\lambda_k \gamma(C_1) + (1-\lambda_k)\gamma(C_2)) + \frac{4R}{N} \leq \max\{\gamma(C_1), \gamma(C_2)\} + \frac{4R}{N}.\end{aligned}$$

Nun war $N \in \mathbb{N}$ beliebig und damit ist (6.14) gezeigt.

<u>Zu 9:</u> Wegen $M_n \subset \bigcap_{k=1}^{\infty} M_k$ gilt nach 5.

$$\gamma(\bigcap_{k=1}^{\infty} M_k) \leq \gamma(M_n) \to 0$$

und das heißt nach 2., dass $\bigcap_{k=1}^{\infty} M_k$ kompakt ist, denn ein beliebiger Durchschnitt von abgeschlossenen Mengen ist wieder abgeschlossen.

Wir müssen noch zeigen, dass der Durchschnitt nicht leer ist. Wähle dafür zu $n \in \mathbb{N}$ Punkte $x_n \in M_n$. Da

$$\{\, x_n \mid n \geq m \,\} \subset K_m,$$

folgt

$$\begin{aligned}\gamma(\{\, x_n \mid n \geq 1 \,\}) &\overset{6.}{=} \max\{\gamma(\{x_1\}), \{\, x_n \mid n \geq 2 \,\}\} = \gamma(\{\, x_n \mid n \geq 2 \,\}) = \dots \\ &= \gamma(\{\, x_n \mid n \geq m \,\}) \leq \gamma(K_m) \to 0.\end{aligned}$$

Daher ist die Menge $\{\, x_n \mid n \geq 1 \,\}$ relativkompakt und es gibt eine Teilfolge $(x_{n_k})_k \subset (x_n)_n$ mit $x_{n_k} \to x$. Da der Durchschnitt abgeschlossen ist, liegt der Grenzwert x im Durchschnitt und deswegen ist dieser auch nicht leer. □

Definition 6.3 (Verdichtend). Sei $(X, \|.\|_X)$ ein Banachraum. Eine stetige Abbildung $\phi : M \subset X \to X$ heißt *verdichtend* oder *kondensierend*, wenn

$$\gamma(\phi(A)) < \gamma(A)$$

für alle beschränkten Teilmengen $A \subset X$ mit $\gamma(A) > 0$.

Beispiel 6.6. Kompakte Abildungen ϕ_1 sind verdichtend, denn es gilt stets $\gamma(\phi_1(A)) = 0$ für jede beschränkte Menge A.

□

Beispiel 6.7. Strikte k-Kontraktionen sind verdichtend, denn für jede beschränkte Menge A ist

$$\gamma(\phi_2(A)) \leq k\gamma(A) < \gamma(A).$$

□

Beispiel 6.8. Summen von kompakten und strikten k-Kontraktiven Abbildungen sind verdichtend, denn ist ϕ_1 kompakt und ϕ_2 kondensierend, so folgt für jede beschränkte Menge A:

$$\gamma(\phi_1(A) + \phi(A_2)) \overset{4.}{\leq} \underbrace{\gamma(\phi_1(A))}_{=0} + \gamma(\phi_2(A)) \overset{\text{Bsp 6.7}}{\leq} k\gamma(A) < \gamma(A).$$

□

Satz 6.4 (Fixpunktsatz von Darbo-Sadovskiĭ, 1955). *Seien $(X, \|.\|_X)$ ein Banachraum und $M \subset X$ eine abgeschlossene, nicht-leere, konvexe, beschränkte Teilmenge. Dann besitzt jede stetige verdichtende Abbildung $\phi : M \to M$ mindestens einen Fixpunkt.*

Beweis. Die Idee des Beweises besteht darin, eine konvexe, kompakte Teilmenge $\tilde{M} \subset M$ zu finden, die invariant unter ϕ ist, so dass der Fixpunktsatz von Schauder in der ersten Version angewendet werden kann.

Sei also $y \in M$. Betrachte die Menge

$$\mathcal{K} := \{\, K \subset M \mid K \text{ ist abgeschlossen und konvex, } y \in K \text{ und } \phi(K) \subset K \,\}.$$

Offensichtlich ist $M \in \mathcal{K}$, also $\mathcal{K} \neq \emptyset$.

Definiere jetzt

$$\tilde{M} := \bigcap_{K \in \mathcal{K}} K.$$

Zunächst sehen wir, dass $\tilde{M}$ als Durchschnitt abgeschlossener konvexer Mengen wieder abgeschlossen und konvex ist. Weiter ist $\tilde{M} \neq \emptyset$, denn $\mathcal{K} \neq \emptyset$ und $y \in K$. Außerdem sehen wir leicht, dass für alle $K \in \mathcal{K}$ die Implikationskette

$$\phi(\tilde{M}) = \phi(\bigcap_{K \in \mathcal{K}} K) \subset \phi(K) \subset K$$

gilt. Also ist

$$\phi(\tilde{M}) \subset \bigcap_{K \in \mathcal{K}} K = \tilde{M}.$$

Damit müssen wir nur noch zeigen, dass $\tilde{M}$ kompakt ist: Dazu bemerken wir zunächst, dass

$$N := \overline{\text{co}}(\phi(\tilde{M}) \cup \{y\}) \overset{\phi(\tilde{M}) \subset \tilde{M}}{\subset} \overline{co}(\tilde{M} \cup \{y\}) \overset{y \in \tilde{M}}{\subset} = \overline{co}(\tilde{M}) = \tilde{M}. \tag{6.15}$$

Auf der anderen Seite ist N abgeschlossen, konvex, $y \in N$ und

$$\phi(N) \overset{(6.15)}{\subset} \phi(\tilde{M}) \subset \overline{\text{co}}(\phi(\tilde{M}) \cup \{y\}) = N.$$

Zusammenfassend ist also $N \in \mathcal{K}$ und daher ist

$$\tilde{M} = \bigcap_{K \in \mathcal{K}} K \subset N.$$

Wir haben damit gezeigt, dass

$$\tilde{M} = N = \overline{\text{co}}(\phi(\tilde{M}) \cup \{y\}).$$

Es folgt

$$\begin{aligned} \gamma(\tilde{M}) &= \gamma(\overline{\text{co}}(\phi(\tilde{M}) \cup \{y\})) \overset{8.}{=} \gamma(\phi(\tilde{M}) \cup \{y\}) \\ &\overset{6.}{\leq} \max\{\gamma(\phi(\tilde{M})), \underbrace{\gamma(\{y\})}_{=0}\} = \gamma(\phi(\tilde{M})). \end{aligned}$$

Angenommen, $\gamma(\tilde{M}) > 0$. Dann folgt, weil ϕ verdichtend ist:

$$\gamma(\tilde{M}) \overset{5.}{\leq} \gamma(\phi(\tilde{M})) < \gamma(\tilde{M}).$$

Widerspruch, also gilt doch $\gamma(\tilde{M}) = 0$, d.h. $\tilde{M}$ ist relativ kompakt und da $\tilde{M}$ abgeschlossen ist, somit kompakt.

Schließlich können wir jetzt die 1. Version des Schauderschen Fixpunktsatzes auf

$$\phi|_{\tilde{M}} : \tilde{M} \to \tilde{M}$$

anwenden und sind fertig. □

Bemerkung. Der Darbo-Sadovskiïsche Fixpunktsatz enthält als Spezialfall sowohl den Fixpunktsatz von Banach als auch den von Schauder. Beachten Sie, dass letzterer in den Beweis einging.

Korollar 6.2 (**Fixpunktsatz von Krasnoselskiï**). *Seien X ein Banachraum und $M \subset X$ abgeschlossen, konvex, beschränkt und nicht-leer. Weiter seien $\phi_1 : M \to X$ kompakt, $\phi_2 : M \to X$ eine strikte Kontraktion und $\phi := \phi_1 + \phi_2 : M \to M$ eine Selbstabbildung. Dann besitzt ϕ einen Fixpunkt.*

Beweis. In Beispiel 6.8 wurde gezeigt, dass ϕ verdichtend ist. Fertig. □

Wir wollen noch weitere Fixpunktsätze studieren und betrachten zunächst *nicht-expansive* Abbildungen. Das sind Lipschitz-stetige Abbildungen mit Lipschitz-Konstante 1. Unser Ziel ist es, die Kompaktheitsvoraussetzung an ϕ fallen zu lassen. Wir machen das wieder gut, indem wir zusätzliche geometrische Eigenschaften des Banachraumes voraussetzen. Bevor wir dazu kommen, vorerst noch ein kleines

Lemma 6.4. *Seien $(X, \|.\|_X)$ ein Banachraum, M eine abgeschlossene, konvexe, nicht-leere Teilmenge von X und $\phi : M \to M$ nicht expansiv. Dann existiert eine Folge $(x_n)_n \subset M$ mit*

$$\|\phi(x_n) - x_n\|_X \overset{n\to\infty}{\longrightarrow} 0.$$

Beweis. Sei $y \in M$ und $0 < \lambda \leq 1$. Betrachte die Abbildung

$$\phi_\lambda : M \to M, \quad \phi_\lambda(x) := \lambda y + (1-\lambda)\phi(x).$$

Diese Abbildung ist wohldefiniert, d.h. $\phi_\lambda(x) \in M$, weil $y, \phi(x) \in M$ und M konvex ist. Offensichtlich ist ϕ_λ stetig. Weiter gilt für alle $x, \tilde{x} \in M$ die Abschätzung

$$\|\phi_\lambda(x) - \phi_\lambda(\tilde{x})\| = (1-\lambda)\|\phi(x) - \phi(\tilde{x})\| \leq \underbrace{(1-\lambda)}_{<1}\|x - \tilde{x}\|,$$

d.h. ϕ_λ ist eine strikte $(1-\lambda)$-Kontraktion. Nach dem Banachschen Fixpunktsatz gilt dann:

$$\forall_{\lambda \in]0,1[} \exists!_{x_\lambda \in M} : \phi_\lambda(x_\lambda) = x_\lambda.$$

Weiter ist

$$\|\phi(x_\lambda) - x_\lambda\| = \| \tfrac{1}{1-\lambda}(\underbrace{\phi_\lambda(x_\lambda)}_{=x_\lambda} - \lambda y) - x_\lambda \| = \tfrac{\lambda}{1-\lambda}\|x_\lambda - y\| \overset{\lambda\to 0}{\longrightarrow} 0.$$

Beachten Sie beim Grenzwertübergang, dass x_λ tatsächlich beschränkt ist, weil $x_\lambda \in M$.

Wählen wir nun $\lambda := \frac{1}{n}$, so sind wir fertig. □

Nun zum versprochenen Fixpunktsatz:

Satz 6.5. *Seien $(H, (.,.))$ ein Hilbertraum sowie $M \subset H$ abgeschlossen, beschränkt, konvex und nicht-leer. Dann besitzt jede nicht-expansive Abbildung $\phi : M \to M$ mindestens einen Fixpunkt.*

Beweis. Nach Lemma 6.4 existiert eine Folge $(x_n)_n \subset M$ mit

$$|\phi(x_n) - x_n| \to 0. \tag{6.16}$$

Da M beschränkt ist, ist auch die Folge (x_n) beschränkt. Außerdem ist H reflexiv und daher gibt es nach Satz 3.15 eine Teilfolge $(x_{n_k})_k$ mit

$$x_{n_k} \rightharpoonup x \quad \text{in } H.$$

Nun ist M konvex und abgechlossen und deswegen ist M nach Satz 3.12 schwach abgeschlossen, d.h. $x \in M$. Ferner gilt

$$\begin{aligned} |x_{n_k} - \phi(x)|^2 &= |(x_{n_k} - x) + (x - \phi(x))|^2 \\ &= |x_{n_k} - x|^2 + 2\,\mathrm{Re}(x_{n_k} - x, x - \phi(x)) + |x - \phi(x)|^2 . \end{aligned}$$

Nach Voraussetzung ist $(x_{n_k} - x)$ eine schwache Nullfolge, so dass der mittlere Term gegen 0 strebt. Es folgt

$$|x_{n_k} - \phi(x)|^2 - |x_{n_k} - x|^2 \to |x - \phi(x)|^2 . \tag{6.17}$$

Auf der anderen Seite gilt

$$\begin{aligned} |x_{n_k} - \phi(x)| &\leq |x_{n_k} - \phi(x_{n_k})| + |\phi(x_{n_k}) - \phi(x_{n_k})| \\ &\overset{(*)}{\leq} |x_{n_k} - \phi(x_{n_k})| + |x_{n_k} - x| , \end{aligned}$$

wobei an der Stelle $(*)$ eingeht, dass ϕ nicht-expansiv ist. Wir bemerken, dass nach (6.16) der vorletzte Summand gegen 0 geht. Daher folgt

$$\limsup_{k\to\infty}(|x_{n_k} - \phi(x)| - |x_{n_k} - x|) \leq 0.$$

Bemerke, dass der Limes wegen der Beschränktheit von x_{n_k} tatsächlich $< \infty$ ist. Daher gilt ebenso

$$0 \leq \limsup_{k\to\infty}(|x_{n_k} - \phi(x)| + |x_{n_k} - x|) < \infty.$$

Wir können damit die obere Ungleichung multiplizieren und erhalten

$$\limsup_{k\to\infty}(|x_{n_k} - \phi(x)| - |x_{n_k} - x|) \cdot \limsup_{k\to\infty}(|x_{n_k} - \phi(x)| + |x_{n_k} - x|) \leq 0.$$

Ein Satz aus der Analysis sagt, dass

$$\limsup_n f_n \cdot g_n \leq \limsup_n f_n \cdot \limsup_n g_n, \qquad \text{falls } f_n \geq 0 \text{ f.ü.}$$

Daher folgt

$$\limsup_{k\to\infty}(|x_{n_k} - \phi(x)|^2 - |x_{n_k} - x|^2) \leq 0. \tag{6.18}$$

Nach Vergleich von (6.18) mit (6.17) sieht man schließlich $\phi(x) = x$. □

Definition 6.4 (Gleichmäßige Konvexität, Konvexitätsmodul). Ein Banachraum $(X, \|.\|_X)$ heißt *gleichmäßig konvex* oder *uniform konvex*, wenn es zu jedem $\epsilon > 0$ ein $\delta > 0$ mit der Eigenschaft

$$\|x\|, \|y\| \leq 1,\ \|x - y\| \geq \epsilon \implies \left\| \frac{x+y}{2} \right\| \leq 1 - \delta$$

gibt. Man nennt

$$\delta_X : \epsilon \mapsto \inf\{\, 1 - \left\| \frac{x+y}{2} \right\| \mid \|x\|, \|y\| \leq 1,\ \|x - y\| \geq \epsilon \,\}$$

das *Konvexitätsmodul von X*.

Bemerkungen.

1. Uniform konvexe Räume sind strikt konvex.

2. Die Einheitskugel in einem uniformen Raum ist „rund“. Die „Rundheit“ kann sogar quantifiziert werden.

3. Hilberträume sind uniform konvex. Nach der Parallelogrammgleichung gilt nämlich

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

   Daraus folgt

$$\left|\tfrac{x+y}{2}\right|^2 = \tfrac{1}{2}(|x|^2 + |y|^2) - \tfrac{1}{4}\,|x-y|^2\,.$$

   Sei $0 < \epsilon < 2$ beliebig. Setze dann $\delta := 1 - \sqrt{1 - \frac{\epsilon^2}{4}}$. Für $x, y \in H$ mit $\|x\|, \|y\| \leq 1$ und $\|x - y\| \geq \epsilon$ gilt dann nach der oben umgeformten Parallelogrammgleichung

$$\left|\tfrac{x+y}{2}\right| \leq \sqrt{1 - \tfrac{\epsilon^2}{4}} = 1 - \delta.$$

   Beachte, dass ϵ höchstens den Wert 2 annehmen kann, denn

$$\epsilon \leq \|x - y\| \leq \|x\| + \|y\| \leq 2.$$

   Für $\epsilon = 2$ wählt man einfach $\delta \in]0, 1[$ beliebig.

4. Die Räume l^p und L^p sind uniform konvex für $1 < p < \infty$. Dagegen sind l^1, l^∞, L^1 und L^∞ nicht uniform konvex, denn sie nicht mal strikt-konvex.

5. Der *Satz von Milman* sagt, dass jeder uniforme konvexe Banachraum reflexiv ist. Der Beweis ist sehr aufwendig und ist bei *Yosida, p. 127* zu finden. Der Beweis wird aber einfacher, sobald man Mittel der *Funktionalanylyis 2* zur Verfügung hat. Daher tragen wir den Beweis im nächsten Semester nach!

6. Die uniforme Konvexität bleibt beim Übergang zu einer äquivalenten Norm i.A. nicht erhalten.

7. X ist genau dann uniform konfex, wenn für alle $\epsilon > 0$ auch $\delta_X(\epsilon) > 0$ ist.

8. Die Abbildung δ_X ist monoton wachsend. Außerdem gilt stets $\delta_X \in [0, 1]$, denn

$$0 \leq \left\|\tfrac{x+y}{2}\right\| \leq \tfrac{1}{2}(\|x\| + \|y\|) \leq 1.$$

   Außerdem ist $\epsilon \in]0, 2]$, denn

$$0 < \epsilon \leq \|x - y\| \leq \|x\| + \|y\| \leq 2.$$

   Mit der Monotonie folgt daher weiter

$$\delta_X(2) \leq 1.$$

Satz 6.6 (**Fixpunktsatz von Browder-Göhde-Kirk, 1965**). *Seien $(X, \|.\|_X)$ ein uniform konvexer Banachraum sowie $M \subset X$ eine abgeschlossene, beschränkte, konvexe und nicht-leere Menge. Dann besitzt jede stetige nicht-expansive Selbstabbildung $\phi : M \to M$ mindestens einen Fixpunkt.*

Beweis. <u>1. Schritt:</u> Existenz einer Funktion φ mit $\varphi(\epsilon) \to 0$ für $\epsilon \to 0$, so dass für $x, y \in M$:

$$\|x - \phi(x)\| \leq \epsilon, \|y - \phi(y)\| \leq \epsilon \Longrightarrow \left\| \frac{x+y}{2} - \phi\left(\frac{x+y}{2}\right) \right\| \leq \varphi(\epsilon). \tag{6.19}$$

Wir bemerken, dass

$$z := \frac{x+y}{2} \in M,$$

weil M konvex ist. Weiter sehen wir, dass

$$\|\phi(x) - x\| \leq \|\phi(z) - \phi(x)\| + \|\phi(x) - x\| \leq \|z - x\| + \epsilon, \tag{6.20}$$

wobei wir in der letzten Abschätzung ausgenutzt haben, dass ϕ nicht-expansiv ist.

Definiere jetzt

$$\rho := \rho(x, y, \epsilon) := \frac{1}{2}\|x - y\| + \epsilon = \|z - x\| + \epsilon = \|z - y\| + \epsilon. \tag{6.21}$$

Dann wird aus (6.20)

$$\|\phi(z) - x\| \leq \rho$$

und analog erhält man

$$\|\phi(z) - y\| \leq \rho.$$

Setze

$$x_1 := \frac{\phi(z) - x}{\rho}, \quad y_1 := \frac{\phi(z) - y}{\rho}.$$

Dann gilt:

$$\|x_1\| \leq \rho, \quad \|y_1\| \leq 1, \quad \|x_1 - y_1\| = \frac{1}{\rho}\|x - y\|.$$

Daraus folgt wegen der uniformen Konvexität

$$\frac{\|\phi(z) - z\|}{\rho} = \left\| \frac{x_1 + y_1}{2} \right\| \leq 1 - \delta_X\left(\frac{\|x - y\|}{\rho}\right).$$

Nach (6.21) gilt nun

$$\|x - y\| = 2\rho - 2\epsilon$$

und daher folgt

$$\|\phi(z) - z\| \leq \rho\left(1 - \delta_X\left(2 - 2\frac{\epsilon}{\rho}\right)\right).$$

<u>1. Fall:</u> $\rho \leq \sqrt{\epsilon}$. Dann gilt wegen $\delta_X \in [0, 1]$

$$\|\phi(z) - z\| \leq \sqrt{\epsilon}.$$

<u>2. Fall:</u> $\rho > \sqrt{\epsilon}$. Dann gilt $2 - 2\frac{\epsilon}{\rho} > 2 - 2\sqrt{\epsilon}$ und wegen der Monotonie von δ_X folgt

$$\delta_X(2 - 2\tfrac{\epsilon}{\rho}) \geq \delta(2 - 2\sqrt{\epsilon}).$$

Damit können wir nun abschätzen:

$$\|\phi(z) - z\| \leq \rho(1 - \delta_X(2 - 2\sqrt{\epsilon})) \overset{(6.21)}{\leq} (\tfrac{1}{2}\operatorname{diam}(M) + \epsilon)(1 - \delta_X(2 - 2\sqrt{\epsilon}))$$

Setzen wir also zusammenfassend

$$\varphi(\epsilon) := \sqrt{\epsilon} + (\tfrac{1}{2}\operatorname{diam}(M) + \epsilon)(1 - \delta_M(2 - 2\sqrt{\epsilon})),$$

so ist die Abschätzung in (6.21) gezeigt.

Es bleibt zu zeigen, dass $\lim_{\epsilon \to 0} \varphi(\epsilon) = 0$. Dazu reicht es

$$\lim_{\alpha \to 2} \delta_X(\alpha) \overset{a)}{=} \sup_{\alpha < 2} \delta_X(\alpha) \overset{b)}{=} 1.$$

einzusehen.

Zu b: Wegen $\delta_X(\alpha) \in [0,1]$, ist $\sup_{\alpha<2} \delta_X(\alpha) \leq 1$ klar. Um „$\geq$“ zu zeigen, machen wir die Annahme, dass

$$\sup_{\alpha<2} \delta_X(\alpha) < 1.$$

Dann gibt es ein $\beta > 0$ mit

$$\delta_X(\alpha) < 1 - \beta \quad \text{für alle } \alpha < 2.$$

Außerdem existieren zu jedem $n \in \mathbb{N}$ Punkte $x_n, y_n \in B_X$ mit

$$\|x_n - y_n\| \geq 2 - \tfrac{1}{n}, \quad 1 - \left\|\tfrac{x_n+y_n}{2}\right\| \leq 1 - \beta.$$

Setze $\tilde{y}_n = -y_n$. Dann folgt

$$1 - \left\|\tfrac{x_n+\tilde{y}_n}{2}\right\| \leq 1 - (1 - \tfrac{1}{2n}) = \tfrac{1}{2n} \tag{6.22}$$

und

$$\|x_n - \tilde{y}_n\| \geq 2\beta. \tag{6.23}$$

Aus (6.22) erhält man

$$1 - \left\|\tfrac{x_n-\tilde{y}_n}{2}\right\| \geq \delta_X(2\beta) > 0,$$

wobei die letzte strikte Ungleichung gilt, weil X uniform konvex ist. Das ist aber ein Widerspruch zu (6.23) für große n.

Zu a: Folgt aus der Monotonie von δ_X.

2. Schritt: Existenz des Fixpunktes. Setze dazu

$$\begin{aligned} \eta(r) &= \inf\{\, \|\phi(x) - x\| \mid x \in M,\ \|x\| \leq r \,\}, \\ s_0 &= \inf\{\, r \geq 0 \mid \eta(r) = 0 \,\}. \end{aligned}$$

Nun ist M beschränkt. Für hinreichend großes r ist daher nach Lemma 6.4 $\eta(r) = 0$, also $s_0 < \infty$.

Sei $(x_n)_n \subset M$ eine Folge mit

$$\|x_n\| \to s_0 \quad \text{und} \quad \|F(x_n) - x_n\| = 0.$$

Dann gilt

$$x := \lim_{n\to\infty} x_n \tag{6.24}$$

und x ist der gesuchte Fixpunkt, denn

$$0 = \lim_{n\to\infty} \phi(x_n) - x_n = \phi(x) - x \Longrightarrow \phi(x) = x.$$

Angenommen, (6.24) gilt nicht. Notwendigerweise ist dann $s_0 > 0$, denn sonst wäre (x_n) eine Nullfolge. Weiter gibt es eine Teilfolge (x_{n_k}) mit

$$\left\|x_{n_{k+1}} - x_{n_k}\right\| \geq \tau$$

für ein geeignetes $\tau > 0$. Wir wählen jetzt

$$s_0 < s_1 < 2s_0$$

mit

$$s_2 := s_1(1 - \delta_X(\tfrac{\tau}{2s_0})) < s_0.$$

Ist k groß genug, so gilt $\|x_{n_k}\| \leq s_1$ und deswegen

$$\left\|\tfrac{x_{n_k}}{s_1}\right\| \leq 1, \quad \left\|\tfrac{x_{n_{k+1}}}{s_1}\right\| \leq 1, \quad \left\|\tfrac{x_{n_{k+1}}}{s_1} - \tfrac{y_{n_k}}{s_1}\right\| \geq \tfrac{\tau}{s_1}.$$

Wegen der gleichmäßigen Konvexität folgt dann

$$\left\| \frac{x_{n_k} + x_{n_{k+1}}}{2s_1} \right\| \leq 1 - \delta_X(\tfrac{\tau}{s_1}),$$

und daher wegen der Monotonie von δ_X

$$\left\| \frac{x_{n_k} + x_{n_{k+1}}}{2} \right\| \leq s_1(1 - \delta_X(\tfrac{\tau}{s_1})) \leq s_2 < s_0$$

für große k. Nun gilt aber $\left\| \phi(x_{n_k}) - x_{n_{k+1}} \right\| \to 0$ und Anwendung des Resulats aus dem ersten Schritt ergibt

$$\left\| \phi(\tfrac{x_{n_k} + x_{n_{k+1}}}{2}) - \tfrac{x_{n_k} + x_{n_{k+1}}}{2} \right\| \to 0.$$

Also ist $\eta(s_2) = 0$. Widerspruch zur Minimalität von s_0, denn $s_2 < s_0$. $\square$

7 Anhang

7.1 Die allgemeine Version des Satzes von Arzelà-Ascoli

In diesem Abschnitt wollen wir die allgemeine Version des folgenden Satzes beweisen:

Satz 7.1 (**von Arzelà-Ascoli: Allgemeine Version**). *Sei (K, d_K) ein kompakter metrischer Raum und sei (X, d_X) ein vollständiger metrischer Raum. Weiter sei $A \subset C(K, X)$. Dann ist A genau dann relativ kompakt, wenn gilt*

(i) Für alle $k \in K$ ist $A(k) := \{ f(k) \mid f \in A \}$ relativ kompakt in X.

(ii) A ist gleichgradig stetig auf A, d.h.

$$\forall_{\epsilon>0} \exists_{\delta>0} : |f(x) - f(y)| < \epsilon \text{ für alle } x, y \in K \text{ mit } d(x, y) < \delta \quad \forall_{f \in A}.$$

Beweis. Zu „$\Longrightarrow$“: Die gleichradige Stetigkeit zeigt man wie im Beweis des klassischen Satzes (Satz 1.15). Wir müssen also nur noch die relative Kompaktheit der Menge $A(k)$ für jedes $k \in K$ zeigen:

Da A relativ kompakt ist, ist A insbesondere nach Satz 1.11 präkompakt. Es gibt also zu $\epsilon > 0$ endlich viele $f_1, \ldots, f_n \in A$, so dass

$$A \subset \bigcup_{i=1}^{n} B(f_i, \epsilon/3).$$

Insbesondere gilt dann für alle $k \in K$

$$A(k) \subset \bigcup_{i=1}^{n} B(f_i(k), \epsilon/3).$$

Damit ist $A(k)$ für jedes $k \in K$ präkompakt und nach Satz 1.11 ist dann auch $\overline{A(k)}$ präkompakt. Weil $\overline{A(k)} \subset X$ und X vollständig ist, folgt aus Satz 1.8 sofort, dass $\overline{A(k)}$ vollständig ist. Insgesamt folgt so aus Satz 1.10, dass $\overline{A(k)}$ kompakt, also $A(k)$ relativ kompakt ist.

Zu „$\Longleftarrow$“: Zunächst ist $C(K, X)$ vollständig, da X vollständig ist. Nach Satz 1.8 ist dann auch $\bar{A} \subset C(K, X)$ vollständig. Es genügt daher die Präkompaktheit von $\bar{A}$ zu zeigen, denn aus Satz 1.10 folgt dann sofort, dass $\bar{A}$ kompakt ist.

Seien $\epsilon > 0$ und $x \in K$ beliebig. Da A gleichgradig stetig ist, gibt es zu beliebigem $f \in A$ ein $\delta > 0$, so dass

$$d_X(f(x), f(y)) < \frac{\epsilon}{4} \quad \text{für alle } y \in K \text{ mit } d(x, y) < \delta.$$

Nun ist K kompakt und daher gibt es endlich viele Punkte x_i, $i = 1, \ldots I$, und zugehörige offene Kugeln $B(x_i, \delta_i)$, so dass

$$X \subset \bigcup_{i=1}^{I} B(x_i, \delta_i).$$

Da nach Voraussetzung jedes $A(x_i)$, $i = 1, \ldots, I$, relativ kompakt ist, ist auch die Vereinigung

$$V := \bigcup_{i=1}^{I} A(x_i)$$

relaitv kompakt in X. Es gibt also endlich viele $y_j \in V$, $j = 1, \ldots, J$, so dass

$$V \subset \bigcup_{j=1}^{J} B(y_j, \epsilon/4),$$

d.h.

$$\forall_{f\in A}\forall_{i\in\{1,\dots,I\}}\exists_{j\in\{1,\dots,J\}} : d_X(f(x_i)-y_j) < \epsilon/4.$$

Definiere ϕ als die Menge der Abbildungen

$$\varphi : \{1,\dots,I\} \to \{1,\dots,J\}.$$

Dann ist ϕ endlich. Weiter setze

$$L_\varphi := \{\, f \in A \mid \max_{1\le i\le I} d_X(f(x_i)-y_{\varphi(i)}) < \epsilon/4 \,\}, \quad \varphi\in\phi.$$

Offensichtlich ist nach Konstruktion

$$A = \bigcup_{\varphi\in\phi} L_\varphi.$$

Zu beliebigen $\varphi\in\phi$, $f,g\in L_\varphi$ und $z\in K$ gibt es $i\in\{1,\dots,I\}$, so dass $z\in B(x_i,\delta_i)$ und

$$d_X(f(z),f(x_i)) < \epsilon/2 \quad \text{sowie} \quad d_X(g(z),g(x_i)) < \epsilon/2.$$

Nach Definition ist

$$d_X(f(x_i),g(x_i)) \le d_X(f(x_i),y_{\varphi(i)}) + d_X(y_{\varphi(i)},g(x_i)) < \epsilon/2.$$

Zusammenfassend folgt für alle $z\in K$

$$d_X(f(z),g(z)) \le d_X(f(z),f(x_i)) + d_X(f(x_i),g(x_i)) + d_X(g(x_i),g(z)) < \epsilon,$$

also

$$\sup_{f,g\in L_\varphi} d_X^{sup}(f,g) < \epsilon \quad \text{für alle } \varphi\in\phi.$$

Zu jedem $f\in A$ gibt es dann ein $\varphi\in\phi$, so dass $f\in L_\varphi$. Außerdem existiert ein beliebiges, aber festes $g_\varphi\in L_\varphi$ mit

$$d_X^{\sup}(f,g_\varphi) \le \sup_{f,g\in L_\varphi} d^{\sup}(f,g) < \epsilon.$$

Daher gilt

$$A \subset \bigcup_{\varphi\in\phi} B(g_\varphi,\epsilon)$$

und das zeigt die Präkompaktheit von A. $\square$

7.2 Schwache Konvergenz in l^1 = starke Konvergenz in l^1

In Abschnitt 3.4 haben wir gesehen, dass aus der starken Konvergenz stets die schwache folgt. In l^1 gilt sogar die Umkehrung:

Satz 7.2 (**Lemma von Schur**). *Eine Folge konvergiert genau dann stark in l^1, wenn sie schwach konvergiert.*

Beweis. Zu „$\Longrightarrow$“: Klar.

Zu „$\Longleftarrow$“: Sei also o.E. $(x_n)_n$ eine schwache Nullfolge aus l^1. Konvergiert die Folge dann stark, so konvergiert sie ebenfalls gegen Null. Angenommen, die Folge $(x_n)_n$ konvergiert nicht stark. Dann gibt es ein $\epsilon > 0$ und eine Teilfolge - nennen wir sie wieder $(x_n)_n$, um Doppelindizes zu vermeiden - so dass

$$\|x_n\|_{l^1} \geq \epsilon > 0 \quad \text{für alle } n \in \mathbb{N}.$$

Wir konstruieren nun induktiv eine Teilfolge $(y_n)_n \subset (x_n)_n$ und eine Folge $(a(n))_n \subset \mathbb{N}$ mit $a(0) = 0$, $a(n-1) < a(n)$ und

$$\sum_{k=1}^{a(n-1)} |y_n(k)| \leq \frac{1}{n} \quad \text{und} \quad \sum_{k=a(n-1)-1}^{a(n)} |y_n(k)| \geq \|y_n\|_{l^1} - \frac{2}{n}. \tag{7.1}$$

Beweis per Induktion:

Induktionsanfang: Für $n = 1$ wähle $y_1 = x_1$. Wegen

$$\lim_{N\to\infty} \sum_{k=1}^{N} |y_1| = \|y_1\|_{l^1}$$

gibt es ein $a(1) \in \mathbb{N}$ mit

$$\sum_{k=1}^{a(1)} |y_1(k)| \geq \|y_1\|_{l^1} - \frac{2}{1}.$$

Induktionsschritt: Seien $y_1, \ldots, y_{n-1}$ und $a(1), \ldots, a(n-1)$ gewählt. Betrachte die stetigen, linearen Funktionale

$$x'_k : l^1 \to \mathbb{K}, \quad x \mapsto x(k), \quad k \in \mathbb{N}.$$

Nach Voraussetzung gilt $x'_k(x_n) \to 0$ für $n \to \infty$. Ist etwa $y_{n-1} = x_l$ gewählt, so gibt es ein $m > l$ derart, dass mit $y_n := x_m$

$$\sum_{k=1}^{a(n-1)} |y_n(k)| = \sum_{k=1}^{a(n-1)} |x'_k(x_m)| \leq \frac{1}{n}.$$

Wie im Fall $n = 1$ kann ein $a(n) > a(n-1)$ gewählt werden, so dass

$$\sum_{k=1}^{a(n)} |y_n(k)| \geq \|y_n\|_{l^1} - \frac{1}{n}. \tag{7.2}$$

Daraus erhalten wir nun zusammen

$$\sum_{k=a(n-1)+1}^{a(n)} |y_n(k)| = \sum_{k=1}^{a(n)} |y_n(k)| - \sum_{k=1}^{a(n-1)} |y_n(k)| \geq \|y_n\|_{l^1} - \frac{2}{n}.$$

Somit ist der Induktionsbeweis geführt.

Um den Widerspruchsbeweis zu beenden, definiere nun

$$y(k) := \frac{\overline{y_n(k)}}{|y_n(k)|}, \quad \text{falls } a(n-1) < k < a(n)$$

und $y(k) = 0$, falls $y_n(k) = 0$ für ein k mit $a(n-1) < k < a(n)$. Dann ist $y \in l^\infty$ und die Abbildung

$$x' : l^1 \to \mathbb{K}, \quad x \mapsto \sum_{k=1}^{\infty} x(k)y(k)$$

stellt ein stetiges Funktional auf l^1 dar. Für dieses gilt nun

$$\begin{aligned}
|x'(y_n)| &= \left|\sum_{k=1}^{\infty} y_n(k)y(k)\right| = \left|\sum_{k=1}^{a(n-1)} y_n(k)y(k) + \sum_{k=a(n-1)+1}^{a(n)} y_n(k)y(k) + \sum_{k=a(n)+1}^{\infty} y_n(k)y(k)\right| \\
&\geq \left|\sum_{k=a(n-1)+1}^{a(n)} y_n(k)y(k)\right| - \sum_{k=1}^{a(n-1)} |y_n(k)y(k)| - \sum_{k=a(n)+1}^{\infty} |y_n(k)y(k)| \\
&= \sum_{k=a(n-1)+1}^{a(n)} |y_n(k)| - \sum_{k=1}^{a(n-1)} |y_n(k)| - \sum_{k=a(n)+1}^{\infty} |y_n(k)| \\
&\overset{(7.1)}{\geq} \|y_n\|_{l^1} - \frac{2}{n} - \frac{1}{n} - \sum_{k=a(n)+1}^{\infty} |y_n(k)y(k)| \,.
\end{aligned}$$

Weiterhin gilt wegen (7.2)

$$\sum_{k=a(n)+1}^{\infty} |y_n(k)| = \|y_n\|_{l^1} - \sum_{k=1}^{a(n)} |y_n(k)| \leq \frac{1}{n}.$$

und daher folgt insgesamt

$$|x'(y_n)| = \left|\sum_{k=1}^{\infty} y_n(k)y(k)\right| \geq \epsilon - \frac{4}{n}.$$

Doch das ist ein Widerspruch zu $y_n \rightharpoonup 0$! $\square$

7.3 Komplexe Analysis

Auch in der Funktionalanalysis braucht man ab und zu Hilfsmittel aus der Funktionentheorie. Ich stelle hier die wichtigsten Sachen zusammen und folge dabei dem Fischer-Lieb. Für detailierte Auskünfte verweise ich auf dieses Werk.

7.3.1 Holomorphe und ganze Funktionen

Definition 7.1 (Holomorph). Seien $G \subset \mathbb{C}$ offen, $f : G \to \mathbb{C}$ und $z_0 \in G$.

(i) f heißt in z_0 *komplex differenzierbar*, wenn

$$\lim_{z \to z_0} \frac{f(z) - f(z_0)}{z - z_0}$$

existiert. Wir nennen dann $f'(z_0) := \lim_{z \to z_0} \frac{f(z)-f(z_0)}{z-z_0}$ die *Ableitung von f in z_0*.

(ii) f heißt *holomorph* oder *regulär analytisch auf G*, wenn f in jedem Punkt von G komplex differenzierbar ist.

(iii) f heißt *holomorph* oder *regulär analytisch in $z_0 \in G$*, wenn f in einer Umgebung von z_0 komplex differenzierbar ist.

(iv) f heißt *ganz*, wenn f auf ganz $\mathbb{C}$ holomorph ist.

Beispiele 7.1.

1. Konstante Funktionen sind komplex differenzierbar mit $f'(z) = 0$.
2. $f(z) = z^n$ ist komplex differenzierbar mit $f'(z) = nz^{n-1}$.
3. $f(z) = \overline{z}$ ist nirgends komplex differenzierbar, jedoch stetig. In der Analysis musste man solch eine Funktion mühevoll konstruieren: Erinnern Sie sich an die Takagi-Funktion?

□

Satz 7.3 (Liouville). *Jede beschränkte ganze Funktion ist konstant.*

Satz 7.4. *Eine auf einer offenen Menge U holomorphe Funktion ist um jeden Punkt $z_0 \in U$ in eine Potenzreihe*

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n$$

entwickelbar.

7.3.2 Laurantreihen

Kreisring: Sei $0 \leq r < R \leq \infty$ und sei $a \in \mathbb{C}$. Defniere

$$\boxed{K_{r,R}(a) := \{\, z \in \mathbb{C} \mid r < |z-a| < R \,\}}.$$

Definition 7.2 (**Laurant-Reihe**).

(i) Eine *Laurent-Reihe* ist eine Reihe der Form

$$\sum_{n=-\infty}^{\infty} a_n (z-a)^n$$

auf $K_{r,R}(a)$. Die Reihe konvergiert in $z_0 \in K_{r,R}(a)$ genau dann, wenn die Reihen

$$\sum_{n=-\infty}^{-1} a_n (z_0-a)^n \quad \text{und} \quad \sum_{n=0}^{\infty} a_n (z_0-a)^n$$

konvergieren.

(ii) Die Reihe $\sum_{n=-\infty}^{-1} a_n (z_0-a)^n$ heißt *Hauptteil* der Laurent-Reihe. Dagegen heißt die Reihe $\sum_{n=0}^{\infty} a_n (z_0-a)^n$ *Nebenteil* der Laurent-Reihe.

Beispiel 7.2. Betrachte $f(z) = e^z + e^{\frac{1}{z}}$. Dann gilt in $K_{0,\infty}(0)$:

$$f(z) = \sum_{n=-\infty}^{-1} \frac{1}{|n|!} z^n + 1 + \sum_{n=0}^{\infty} \frac{1}{n!} z^n.$$

□

7.4 Der Satz vom abgeschlossenen Bild

Die folgenden Ausführen stammen von Prof. Dr. Karl-Heinz Förster.

Bevor wir den großen Satz vom abgeschlossenen Bild beweisen können, brauchen wir einige Hilfssätze. Nützlich wird folgende Menge sein:

$$R_{q,n} := \{\, y \in Y \mid \exists_{x \in X} : \|y - Ax\| \leq q\|y\|,\ \|x\| \leq n\|y\| \,\}.$$

Lemma 7.1. *Seien X, Y Banachräume, $A \in L(X,Y)$ und $0 \leq q < 1$ sowie $n \in \mathbb{N}$. Weiter sei $M \subset Y$ ein Unterraum mit $R(A) \subset M \subset R_{q,n}$. Dann gilt $R(A) = M$.*

Beweis. Sei also $y \in M$. Wir müssen zeigen, dass $y \in R(A)$. Dazu sei $w \in R(A)$. Dann folgt

$$y - w \in M - R(A) \subset M - M = M \subset R_{q,n}. \tag{7.3}$$

Nach Voraussetzung ist $M \subset R_{q,n}$ und daher gibt es zu $y \in R_{q,n}$ ein $x_1 \in X$ mit

$$|\, y - \underbrace{Ax_1}_{=w} \,| \leq q\|y\| \quad \text{und} \quad \|x_1\| \leq n.\|y\|.$$

Wegen (7.3) ist $y - Ax_1 \in R_{q,n}$, d.h. es gibt ein x_2 mit

$$|\, \underbrace{(y - Ax_1) - Ax_2}_{=y - A(x_1+x_2) \in R_{q,n}} \,| \leq q\|y - Ax_1\| \leq q^2\|y\|$$

und

$$\|x_2\| \leq n\|x - Ax_1\| \leq nq\|y\|.$$

Führe das Verfahren induktiv fort. Dann gibt es also eine Folge (x_k) mit

$$\|y - A(x_1 + \ldots x_k)\| \leq q^k\|y\|$$

und

$$\|x_k\| \leq q^{k-1}\|y\|, \quad q < 1.$$

Es folgt, dass der Grenzwert (geometrische Reihe)

$$x := \sum_{k=1}^{\infty} x_k$$

existiert. Also ist $Ax = y$ und so $y \in R(A)$ gezeigt. □

Lemma 7.2. *Seien X, Y Banachräume, $A \in L(X,Y)$, $q \geq 0$ und $M \subset X$ ein Unterraum mit $M \not\subset R_{q,n}$ für alle $n \in \mathbb{N}$. Dann existiert $(y_n) \subset M$ mit $\|y_n\| = 1$, so dass für alle $x \in X$*

$$\|y_n - Ax\| > q \quad \textit{oder} \quad \|x\| > n.$$

Beweis. Nach Voraussetzung gibt es für alle $n \in \mathbb{N}$ ein $\tilde{y}_n \in M$ mit $\tilde{y}_n \notin R_{q,n}$ und $\tilde{y} \neq 0$.

Angenommen, es ist $y := \frac{\tilde{y}}{\|\tilde{y}\|} \in R_{q,n}$. Dann ist $y_n \in M$ und $\|y_n\| = 1$. Dann gibt es ein $x \in X$ mit

$$\|y - Ax\| \leq q\|y_n\| = q \quad \text{und} \quad \|x\| \leq n\|y_n\| = n.$$

Ferner gilt

$$\|\tilde{y}_n - A(\|\tilde{y}_n\|x)\| \leq n\|\tilde{y}_n\|$$

und

$$\|\|\tilde{y}_n\|x\| \leq n\|\tilde{y}_n\|,$$

also $\tilde{y}_n \in R_{q,n}$. Widerspruch. Damit ist $y_n \notin R_{q,n}$ mit $y_n \in M$ und $\|y_n\| = 1$. Für alle $x \in X$ gilt

$$\|y_n - A\| > q\|y_n\| = q \quad \text{oder} \quad \|x\| > n\|y_n\| = n.$$

Damit ist das Lemma bewiesen. □

Lemma 7.3. *Seien X, Y Banachräume, $A \in L(X,Y)$ und $0 < q < 1$. Gilt $\overline{R(A)} \not\subset R_{q,n}$ für alle $n \in \mathbb{N}$, so folgt $\overline{R(A')} \neq R(A')$.*

Beweis. Zu $n \in \mathbb{N}$ definiere

$$V_n := \{\, (\tfrac{1}{n}x, Ax) \mid x \in X \,\} \subset X \times Y.$$

Dann ist $V_n \subset X \times Y =: Z$ abgeschlossen bezüglich der Norm

$$\|(x,y)\|_Z = \|x\| + \|y\|, \quad x \in x, y \in Y,$$

weil der Graph $G(nA) = V_n$ abgeschlossen ist.

Setze nun

$$M := \overline{R(A)} \not\subset R_{q,n} \quad \text{für alle } n \in \mathbb{N}.$$

Wir wollen das Lemma 7.2 anwenden. Sei dazu $(y_n) \subset \overline{R(A)}$ mit $\|y_n\| = 1$ und

$$\|y_n - A\| > q \quad \text{oder} \quad \|x\| > n. \tag{7.4}$$

für alle $x \in X$. Daher gilt

$$\big\|(0, y_n) - (\tfrac{1}{n}x, Ax)\big\|_Z = \tfrac{1}{n}\|x\| + \|y_n - Ax\| \overset{(7.4)}{\geq} \min(1, q).$$

Es folgt

$$\operatorname{dist}((0, y_n), V_n) \geq q > 0.$$

Nach Korollar 3.4 gibt es nun ein $z_n' \in Z'$ mit

$$z_n'|_{V_n} = 0, \quad \|z_n'\| = 1 \quad \text{und} \quad z_n'((0, y_n)) = \|[0, y_n]\| = \operatorname{dist}((0, y_n), V_n) \geq q.$$

Daraus folgt

$$\|z_n'((0, y_n))\| \leq \|(0, y_n)\| \leq 1$$

Betrachte die Einbettung

$$P_Y : Y \to Z, \quad y \mapsto (0, y)$$

und setze

$$y_n' := z_n' \circ P_Y \in Y'$$

Es ist

$$y_n'(y) = z_n' \circ P_Y(y) = z_n'((0, y)). \tag{7.5}$$

Sei $y' \in \ker(A')$. Nach Satz 3.20 gilt

$$y_n \in \overline{R(A)} = (\ker A')_\perp.$$

Ferner gilt

$$y_n'(y_n) - \underbrace{y'(y_n)}_{=0} = y_n'(y_n) \overset{(7.5)}{=} z_n'((0, y_n)) \geq q$$

und daher

$$\|y_n' - y\| \geq q \quad \text{für alle } y' \in \ker A'.$$

Daraus folgt

$$\operatorname{dist}(y_n', \ker A') \geq q. \tag{7.6}$$

Nebenrechung: Aufgrund von $z_n'|_{V_n} = 0$ gilt

$$z_n'((\tfrac{1}{n}x, Ax)) = 0 \Longrightarrow z_n'((\tfrac{1}{n}x, 0) + (0, Ax)) = 0,$$

also

$$z_n'((0, Ax)) = -z_n'((\tfrac{1}{n}x, 0)). \tag{7.7}$$

Zurück zum eigentlich Beweis. Für alle $x \in X$ gilt jetzt

$$A'y_n'(x) = y_n'(Ax) \overset{(7.5)}{=} z_n'((0, Ax)) \overset{(7.7)}{=} -z_n'((\tfrac{1}{n}x, 0)),$$

also

$$|A'y_n'(x)| \leq \underbrace{\|z_n'\|}_{\leq 1} \left\|\tfrac{1}{n}x\right\| \leq \tfrac{1}{n}\|x\|.$$

Damit folgt schließlich

$$\|A'y_n'\| \leq \tfrac{1}{n} \to 0. \tag{7.8}$$

Angenommen, $R(A')$ ist abgeschlossen. Betrachte dann die Abbildungen $A' : Y' \to X'$ und

$$\tilde{A}' : Y'/\ker A' \to R(A'), \quad \tilde{A}([y']) = A'y'.$$

Beide Abbildungen sind stetig und bijektiv. Nach Satz Korollar 2.2 ist dann

$$(\tilde{A}')^{-1} : R(A') \to Y'/\ker(A')$$

stetig. Betrachte

$$(\tilde{A}')^{-1} \underbrace{A'y_n'}_{\to 0} = (\tilde{A}')^{-1}\tilde{A}'([y_n']) = [y_n'].$$

Aufgrund der Stetigkeit von $(\tilde{A}')^{-1}$ folgt dann schließlich

$$\|[y_n]\| \to 0,$$

obwohl nach (7.6)

$$\|[y_n]\| = \operatorname{dist}(y_n', \ker A') \geq q.$$

Widerspruch! Also ist $\overline{R(A')}$ nicht abgeschlossen und $\overline{R(A')} \not\subset R(A')$ gezeigt. □

Satz 7.5 (vom abgeschlossenen Bild). *Seien X, Y Banachräume und $A \in L(X, Y)$. Dann sind äquivalent:*

(i) $R(A)$ *ist abgeschlossen.*

(ii) $R(A')$ *ist abgeschlossen.*

(iii) $R(A) = (\ker A')_\perp$.

(iv) $R(A') = (\ker A)^\perp$.

Beweis. Zu „$(i) \Longleftrightarrow (iii)$“: Das ist Satz 3.20 (iii).

Zu „$(iv) \Longrightarrow (ii)$“: Trivial.

Zu „$(i) \Longrightarrow (iv)$“: Wir beginnen mit „$\supset$“. Sei also $R(A) = \overline{R(A)}$. Ferner sei $x' \in (\ker A)^{\perp}$. Wir müssen zeigen, dass $x' \in R(A')$, d.h. $x' = T'y'$ und das das heißt $x'(\ker A) = 0$.

Wegen $x' \in R(A')$ gibt es ein $u' \in (X/\ker A)'$ mit

$$x' = u' \circ q,$$

wobei $q : X \to X/U$ die Quotientenabbildung darstellt. Betrachte nun

$$\tilde{A} : X/\ker A \to R(A) = \overline{R(A)}, \quad \tilde{A}([x]) = Ax.$$

Dann ist $\tilde{A}$ bijektiv und stetig und aus Korollar 2.2 folgt die Stetigkeit von $\tilde{A}^{-1}$. Setze

$$v' := u' \circ (\tilde{A})^{-1} \in R(A)'.$$

Nach Satz 3.3 gibt es ein $y' \in Y'$ mit

$$y'|_{R(A)} = v'.$$

Für $x \in X$ folgt

$$(A'y')(x) = y'(\underbrace{Ax}_{\in R(A)}) = y'(\tilde{A}([x])) = v'(\tilde{A}([x])) = u' \circ (\tilde{A})^{-1} \circ (\tilde{A}([x])) = u'([x]) = x'(x),$$

also

$$A'y' = x',$$

was zu zeigen war.

Nun zu „$\subset$“: Das folgt aus Satz 3.20 (iv).

Zu „$(ii) \Longrightarrow (i)$“: Jetzt kommen die Lemmata zum Einsatz. Betrachte also

$$R_{q,n} := \{\, y \in Y \mid \exists_{x \in X} : \|y - Ax\| \leq q\|y\|,\ \|x\| \leq n\|y\| \,\}.$$

Da $R(A')$ abgeschlossen ist, gibt es nach Lemma 7.3 für alle $q \in]0,1[$ ein $n_q \in \mathbb{N}$ mit $\overline{R(A)} \subset R_{q,n_q}$. Daraus folgt

$$R(A) \subset \underbrace{\overline{(R(A))}}_{=:M} \subset R_{q,n_q}.$$

Aus Lemma 7.1 folgt schließlich

$$R(A) = \overline{R(A)}.$$

$\square$

7.5 Der Fixpunktsatz von Brouwer

In diesem Abschnitt möchte ich einen Beweis des Brouwerschen Fixpunktsatz mittels Hilfsmitteln aus der Variationsrechnung geben, mit deren Methoden wir uns jetzt vertraut machen wollen.

7.5.1 Ein Ausflug in die Variationsrechnung

Notation. Seien $m, d \in \mathbb{N}$. Wir schreiben für

1. Matrizen: $P = (p^i_j) \in \mathbb{R}^{m\times d}$, wobei obere Indizes Zeilenindizes bezeichnen.
2. Vektoren: $z = (z^i) \in \mathbb{R}^m$ und $x = x_j \in \mathbb{R}^d$.
3. Gradient von $w = (w^1, \dots, w^m) : \mathbb{R}^d \to \mathbb{R}^m$: $\nabla w = (\partial_j w^i)_{i,j}$.
4. Ableitung von einer Funktion $L(P, z)$ nach den einzelnen Komponenten:

$$D_p L = (L_{p^1_1}, \dots, L_{p^m_d}) \quad \text{bzw.} \quad D_z L = (L_{z^1}, \dots, L_{z^m}).$$

Definition 7.3 (**Energiefunktional, Lagrangefunktion**). Seien $\Omega \subset \mathbb{R}^d$ ein glattes Gebiet und

$$L : \mathbb{R}^{m\times d} \times \mathbb{R}^m \times \bar{\Omega} \to \mathbb{R}, \quad L = L(P, z, x) = L(p^1_1, \dots, p^m_d, z^1, \dots, z^m, x_1, \dots, x_d)$$

eine glatte Funktion. Eine Abbildung der Form

$$E : \mathbb{R}^m \to \mathbb{R}, \quad E(w) := \int_\Omega L(\nabla w(x), w(x), x)dx \tag{7.9}$$

heißt *Energiefunktional* und L heißt die *Lagrangefunktion* des Energiefunktionals $E(.)$.

Problem. Sei $g : \partial\Omega \to \mathbb{R}^m$ eine glatte Funktion. Suche das Minimum des Energiefunktionals (7.9) unter allen glatten Funktionen $w : \Omega \subset \mathbb{R}^d \to \mathbb{R}^m$ mit

$$w = g \quad \text{auf} \quad \partial\Omega. \tag{7.10}$$

Lemma 7.4 (**Euler-Lagrange-Gleichungen**). *Sei $u = (u^1, \dots, u^m)$ ein glattes Minimum von* (7.9) *unter allen glatten Funktionen, die* (7.10) *erfüllen. Dann ist u Lösung der Euler-Lagrange-Gleichungen*

$$L_{z^k}(\nabla u, u, x) = \sum_{j=1}^d \partial_j (L_{p^k_j}(\nabla u, u, x)) \quad \text{in } \Omega,\ k = 1, \dots, m. \tag{7.11}$$

Beweis. Sei $v = (v^1, \dots, v^m) \in C^\infty_0(\Omega)$. Definiere

$$i : \mathbb{R} \to \mathbb{R}, \quad i(\tau) := E(u + \tau v).$$

Weil v auf dem Rand verschwindet, erfüllt auch $u + \tau v$ die Randbedingung (7.10). Da u Minimum ist, muss $i'(0) = 0$ sein. Nach Definition gilt

$$i(\tau) = \int_\Omega L(\nabla u + \tau\nabla v, u + \tau v, x)dx$$

und damit könnnen wir nun die 1. Variation berechnen:

$$
\begin{aligned}
i'(\tau) &= \frac{d}{d\tau}\int_\Omega L(\nabla u + \tau\nabla v, u + \tau v, x)dx = \int_\Omega \frac{d}{d\tau} L(\nabla u + \tau\nabla v, u + \tau v, x)dx \\
&= \int_\Omega \left(\sum_{j=1}^d \sum_{k=1}^m L_{p_j^k}(\nabla u + \tau\nabla v, u + \tau v, x)\partial_j v^k + \sum_{k=1}^m L_{z^k}(\nabla u + \tau\nabla v, u + \tau v, x)v^k \right) dx.
\end{aligned}
$$

Aus $i'(0) = 0$ folgt dann

$$
0 = \int_\Omega \left(\sum_{j=1}^d \sum_{k=1}^m L_{p_j^k}(\nabla u, u, x)\partial_j v^k + \sum_{k=1}^m L_{z^k}(\nabla u, u, x)v^k \right) dx.
$$

Partielle Integration ergibt

$$
\begin{aligned}
0 &= \int_\Omega \left(\sum_{j=1}^d \sum_{k=1}^m -\partial_j(L_{p_j^k}(\nabla u, u, x))v^k + \sum_{k=1}^m L_{z^k}(\nabla u, u, x)v^k \right) dx \\
&= \int_\Omega \sum_{k=1}^m v^k \left(\sum_{j=1}^d -\partial_j(L_{p_j^k}(\nabla u, u, x)) + L_{z^k}(\nabla u, u, x) \right) dx.
\end{aligned}
$$

Diese Gleichung muss für alle $v \in C_0^\infty(\Omega)$ gelten. Das geht aber nur, wenn

$$
-\sum_{j=1}^d \partial_j(L_{p_j^k}(\nabla u, u, x)) + L_{z^k}(\nabla u, u, x)v^k = 0 \quad \text{in } \Omega,\ k = 1, \dots, m.
$$

□

Definition 7.4 (**Null-Lagrange-Funktion**). Die Funktion L heißt *Null-Lagrangefunktion*, wenn die zugehörigen Euler-Lagrange-Gleichungen (7.11) für alle glatten Funktionen erfüllt sind.

Wir werden jetzt zeigen, dass der Wert des Energiefunktionals $E(w)$ nur von den Randwerten der Funktion w abhängt:

Satz 7.6. *Sei L eine Null-Lagrangefunktion und seien $u, \tilde{u}$ zwei glatte Funktionen mit*

$$
u = \tilde{u} \quad \text{auf } \partial\Omega.
$$

Dann gilt schon

$$
E(u) = E(\tilde{u})
$$

Beweis. Definiere $j : [0, 1] \to \mathbb{R}$ durch

$$
j(\tau) := E(\tau u + (1 - \tau)\tilde{u}).
$$

Dann ist

$$
\begin{aligned}
j'(\tau) = &\int_\Omega \sum_{i=1}^d \sum_{k=1}^m L_{p_i^k}(\tau\nabla u + (1 - \tau)\nabla\tilde{u}, \tau u + (1 - \tau)\tilde{u}, x)(\partial_i u^k - \partial_i \tilde{u}^k)dx \\
&+ \int_\Omega \sum_{k=1}^m L_{z^k}(\tau\nabla u + (1 - \tau)\nabla\tilde{u}, \tau u + (1 - \tau)\tilde{u}, x)(u^k - \tilde{u}^k)dx
\end{aligned}
$$

Mit partieller Integration und wegen der Tatsache $u = \tilde{u}$ auf $\partial\Omega$ folgt

$$j'(\tau) = \sum_{k=1}^{m} \int_\Omega (-\sum_{i=1}^{d} \partial_i (L_{p_i^k}(\tau\nabla u + (1-\tau)\nabla\tilde{u}, \tau u + (1-\tau)\tilde{u}, x)) \\ + L_{z^k}(\tau\nabla u + (1-\tau)\nabla\tilde{u}, \tau u + (1-\tau)\tilde{u}, x))(u^k - \tilde{u}^k)dx.$$

Nun ist $\tau u + (1-\tau)\tilde{u}$ eine Lösung der Euler-Lagrange-Gleichungen, d.h. es gilt $j'(\tau) = 0$ und somit ist j konstant. Es folgt sofort $E(u) = E(\tilde{u})$, wenn man $\tau = 0$ und $\tau = 1$ setzt. □

Definition 7.5 (Kofaktormatrix). Sei $A \in \mathbb{R}^{d\times d}$ eine Matrix. Dann heißt

$$\text{cof } A$$

die *Kofaktormatrix*, deren (k,i)-ter Eintrag aus der Determinante der $(d-1)\times(d-1)$ Matrix A_i^k besteht, die man durch Streichen der k-ten Zeile und der i-ten Spalte erhält, d.h.

$$(\text{cof } A)_i^k := (-1)^{i+k} \det A_i^k.$$

Beispiel 7.3.

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \Longrightarrow \text{cof } A = \begin{pmatrix} 4 & -3 \\ -2 & 1 \end{pmatrix}$$

Weiter gilt

$$(\det A)I = -2I = \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix}.$$

Auf der anderen Seite ist

$$A^T \text{cof } A = \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix},$$

also gilt

$$(\det A)I = A^T(\text{cof } A). \tag{7.12}$$

Dieser Sachverhalt gilt ganz allgemein für Matrizen $A \in \mathbb{R}^{d\times d}$ (vgl. lineare Algebra).

□

Wir bereiten mit einem Lemma den Satz von *Landers-Ball* vor:

Lemma 7.5. *Sei $u : \mathbb{R}^d \to \mathbb{R}^d$ eine glatte Funktion. Dann gilt*

$$\sum_{i=1}^{d} \partial_i (\mathit{cof}\, \nabla u)_i^k = 0, \quad k = 1, \ldots, d.$$

Beweis. Nach (7.14) gilt für $P \in \mathbb{R}^{d\times d}$ und $i, j = 1, \ldots, d$

$$(\det P)\delta_{ij} = \sum_{k=1}^{d} p_i^k (\text{cof } P)_j^k. \tag{7.13}$$

Daraus folgt für $r, s = 1, \ldots, d$, wobei klar ist, dass $i = j = s$ gewählt werden kann:

$$\frac{\partial \det P}{\partial p_s^r} = \sum_{k=1}^{d} \delta_{kr} (\text{cof } P)_s^k + p_s^k \frac{\partial(\text{ cof } P)_s^k}{\partial p_s^r} = (\text{cof } P)_s^r + \underbrace{\sum_{k=1}^{d} p_s^k (-1)^{k+j} \frac{\partial(\det P)_s^k}{\partial p_s^r}}_{=0}.$$

Also ist

$$\frac{\partial \det P}{\partial p_s^r} = (\operatorname{cof} P)_s^r. \tag{7.14}$$

Setze nun $P = \nabla u = (\nabla_j u^k)_{k,j}$. Dann haben wir also

$$(\det \nabla u)\delta_{ij} = \sum_{k=1}^{d} (\partial_i u^k)(\operatorname{cof} \nabla u)_j^k \quad \text{und} \quad \frac{\partial \det \nabla u}{\partial_s u^r} = (\operatorname{cof} \nabla u)_s^r. \tag{7.15}$$

Differenziere nun den linkes Ausdruck von (7.15) nach x_j. Auf der einen Seite erhalten wir so mit Zuhilfenahme des rechten Ausdrucks von (7.15)

$$\partial_j (\det \nabla u)\delta_{ij} = \sum_{r,s=1}^{d} \delta_{ij} \frac{\partial \det \nabla u}{\partial_s u^r} \partial_j \partial_s u^r = \sum_{r,s=1}^{d} \delta_{ij} (\operatorname{cof} \nabla u)_s^r \partial_j \partial_s u^r$$

und auf der anderen Seite

$$\partial_j \sum_{k=1}^{d} (\partial_i u^k)(\operatorname{cof} \nabla u)_j^k = \sum_{k=1}^{d} \partial_j \partial_i u^k (\operatorname{cof} \nabla u)_j^k + \partial_i u^k \partial_j (\operatorname{cof} \nabla u)_j^k.$$

Summieren über $j = 1, \dots, d$ ergibt zusammenfassend

$$\sum_{j,r,s=1}^{d} \delta_{ij} (\operatorname{cof} \nabla u)_s^r \partial_j \partial_s u^r = \sum_{k,j=1}^{d} \partial_j \partial_i u^k (\operatorname{cof} \nabla u)_j^k + \partial_i u^k \partial_j (\operatorname{cof} \nabla u)_j^k$$

Umstellen ergibt für $i = 1, \dots, d$

$$\begin{aligned} \sum_{k,j=1}^{d} \partial_i u^k \partial_j (\operatorname{cof} \nabla u)_j^k &= \sum_{j,r,s=1}^{d} \delta_{ij} (\operatorname{cof} \nabla u)_s^r \partial_j \partial_s u^r - \sum_{k,j=1}^{d} \partial_j \partial_i u^k (\operatorname{cof} \nabla u)_j^k \\ &= \sum_{r,s=1}^{d} (\operatorname{cof} \nabla u)_s^r \partial_i \partial_s u^r - \sum_{k,j=1}^{d} \partial_j \partial_i u^k (\operatorname{cof} \nabla u)_j^k = 0 \end{aligned}$$

Folglich erhalten wir für $i = 1, \dots, d$

$$\sum_{k=1}^{d} \partial_i u^k \left(\sum_{j=1}^{d} \partial_j (\operatorname{cof} \nabla u)_j^k \right) = 0.$$

Also ist der Vektor $(\sum_{j=1}^{d} \partial_j (\operatorname{cof} \nabla u)_j^k)_{k=1,\dots,d}$ eine Lösung des linearen Gleichungssystems $A^T y = 0$ mit $A = \nabla u$. Sei $x_0 \in \mathbb{R}^d$.

<u>1. Fall:</u> $\det \nabla u(x_0) \neq 0$. Dann folgt sofort die Behauptung

$$\sum_{j=1}^{d} \partial_j (\operatorname{cof} \nabla u(x_0))_j^k = 0, \quad k = 1, \dots, d.$$

<u>2. Fall:</u> $\det \nabla u(x_0) = 0$. Wähle dann $\epsilon_0 > 0$, so dass $\det(\nabla u(x_0) + \epsilon I) \neq 0$ für alle $0 < \epsilon \leq \epsilon_0$. Führe obige Rechnung mit $\tilde{u} := u + \epsilon x$ aus. Nach Grenzwertübergang $\epsilon \to 0$ folgt die Behauptung. □

Satz 7.7 (**Landers (1942), Ball (1976)**). *Die Determinante*

$$L(P) = \det P, \quad P \in \mathbb{R}^{d \times d}$$

ist eine Null-Lagrangefunktion.

Beweis. Wegen den Euler-Lagrange-Gleichungen (Lemma 7.4) müssen wir nur zeigen, dass für jede glatte Funktion $u : \Omega \to \mathbb{R}^d$ gilt:

$$\sum_{i=1}^{d} \partial_i (L_{p_i^k}(\nabla u)) = 0, \quad k = 1, \dots, d.$$

Aufgrund von (7.14) wissen wir aber, dass

$$L_{p_i^k}(\nabla u) = (\operatorname{cof} \nabla u)_i^k, \quad i, k = 1, \dots, d.$$

Die Behauptung folgt somit aus Lemma 7.5. □

Für den nächsten Abschnitt benötigen wir noch einige Begriffe:

Definition 7.6 (Glättungsoperator, Faltung). Gegeben sei eine nichtnegative Funktion $J \in C_0^\infty(\Omega)$ mit den Eigenschaften

(i) $J(x) = 0$, falls $|x| \geq 1$,

(ii) $\int_{\mathbb{R}^d} J(x)dx = 1$.

Für $\epsilon > 0$ ist der *Glättungsoperator* definiert durch

$$J_\epsilon(x) := \epsilon^{-d} J(\tfrac{x}{\epsilon}).$$

J_ϵ besitzt die Eigenschaft (ii) und es gilt $\operatorname{supp}(J_\epsilon) \subset \overline{B_\epsilon(0)}$.

Die Konvolution

$$(J_\epsilon * f)(x) := \int_{\mathbb{R}^d} J_\epsilon(x-y) f(y) dy,$$

welche für Funktionen f definiert ist, für die die rechte Seite Sinn macht, heißt *Regularisierung* von f.

Beispiel 7.4. Ein typisches Beispiel für J ist

$$J(x) = \begin{cases} k \exp(-\frac{1}{1-|x|^2}), & \text{falls } |x| < 1 \\ 0, & \text{falls } |x| \geq 1, \end{cases}$$

wobei k eine Normierungskonstante ist

□

Wir werden im nächsten Abschnitt auch noch den folgenden Satz benötigen:

Satz 7.8. *Sei Ω ein gebiet des R^N und sei $w : \mathbb{R}^N \to \mathbb{R}$ außerhalb von Ω identisch Null. Weiter sei $\epsilon > 0$. Dann gilt für $w \in C(\Omega)$ und realtiv kompaktes $K \subset \Omega$:*

$$\lim_{\epsilon \to 0^+} J_\epsilon w(x) = w(x) \quad \textit{gleichmäßig auf } K.$$

Den Beweis verschieben wir ins nächste Semester, da wir diesen Satz noch im Zusammenhang mit Sobolevräume benötigen werden und diese Räume sind gerade Bestandteil des Moduls Funktionalanalysis II.

7.5.2 Der Beweis des Brouwerschen Fixpunktsatzes

Mit den Resultaten aus dem letzten Abschnitt haben wir nun endlich alle Werkzeuge zusammen, um den entscheidenen Satz zu beweisen:

Satz 7.9 (**Fixpunktsatz von Brouwer, 1912**). *Jede stetige Abbildung*

$$\phi : \overline{B_1(0)} \subset \mathbb{R}^N \to \overline{B_1(0)}$$

besitzt mindestens einen Fixpunkt.

Beweis. 1. Schritt: Es gibt keine glatte Funktion

$$w : \overline{B_1(0)} \to \partial\overline{B_1(0)},$$

so dass für alle $x \in \partial\overline{B_1(0)}$ gilt

$$w(x) = x.$$

Angenommen, solch eine Funktion w existiert doch. Dann sei $\tilde{w}$ die identische Funktion auf $\overline{B_1(0)}$, d.h.

$$\tilde{w}(x) = x \quad \text{für alle } x \in \overline{B_1(0)}.$$

Dann gilt auch

$$\tilde{w}(x) = w(x) \quad \text{für alle } x \in \partial\overline{B_1(0)}.$$

Nun ist die Determinante nach Satz 7.7 eine Null-Lagrangefunktion und daher liefert der Satz 7.6

$$\int_B \det \nabla w dx = \int_B \underbrace{\det \nabla \tilde{w}}_{=1} dx = \mathrm{vol}(B) \neq 0. \tag{7.16}$$

Wir erinnern uns daran, dass

$$w(x) \in \partial\overline{B_1(0)} = \{\, y \in \mathbb{R}^N \mid |y| = 1 \,\}$$

und daher ist $|w(x)|^2 = 1$. Differentiation dieses Ausdrucks ergibt

$$(\nabla w)^T w = 0,$$

d.h. 0 ist ein Eigenwert von $(\nabla w(x))^T$ für alle $x \in \overline{B_1(0)}$, denn w ist Eigenvektor, da $|w| = 1$, also $w \neq 0$. Daher folgt

$$\det(\nabla w(x) - 0\tilde{w}(x)) = \det(\nabla w(x)) = 0.$$

Widerspruch zu (7.16).

2. Schritt: Es gibt keine stetige Funktion

$$w : \overline{B_1(0)} \to \partial\overline{B_1(0)},$$

so dass für alle $x \in \partial\overline{B_1(0)}$ gilt

$$w(x) = x.$$

Angenommen, solch ein w existiert doch. Dann setzen wir w durch $w(x) = x$ für $x \in \mathbb{R}^n \setminus \overline{B_1(0)}$ auf ganz $\mathbb{R}^N$ fort. Daraus und wegen $w(x) \in \partial\overline{B_1(0)}$ folgt $|w(x)| \geq 1$, also insbesondere $w(x) \neq 0$. Sei $\epsilon > 0$. Definiere

$$w_\epsilon(x) := J_\epsilon * w.$$

Wir können sogar w_ϵ so wählen, dass $w_\epsilon(x) \neq 0$ auf ganz $\mathbb{R}^n$:

- Für $x \in \mathbb{R}^N \setminus B_2(0)$ und $\epsilon > 0$ klein genug gilt

$$\begin{aligned} w_\epsilon(x) &= \int_{B_\epsilon(0)} J_\epsilon(|y|) w(x-y) dy = \int_{B_\epsilon(0)} J_\epsilon(|y|)(x-y) dy \\ &= x \underbrace{\int_{B_\epsilon(0)} J_\epsilon(|y|) dy}_{=1} + \underbrace{\int_{B_\epsilon(0)} J_\epsilon(|y|) y dy}_{=0} = x. \end{aligned} \tag{7.17}$$

  Das eine Integral ist $= 0$, weil $J_\epsilon(|y|)$ eine radialsymmetrische Funktion und y eine antisymmetrische Funktion ist.

  Wegen

$$w_\epsilon(x) = x \quad \text{für alle } x \in \mathbb{R}^n \setminus B_2(0) \tag{7.18}$$

  gilt demnach für diese x die Abschätzung $|w_\epsilon(x)| > 2$.

- Nach Satz 7.8 gilt

$$\lim_{\epsilon \to 0} w_\epsilon(x) = w(x) \quad \text{für alle } x \in \overline{B_2(0)}.$$

  Wähle dann $\delta = \frac{1}{2}|w(x)| > 0$, da $w(x) \neq 0$. Dann gilt für alle $x \in \overline{B_2(0)}$

$$|w(x)| - |w_\epsilon(x)| \leq |w(x) - w_\epsilon(x)| < \delta \Longrightarrow |w_\epsilon(x)| \geq |w(x)| - \delta = \frac{1}{2}|w(x)| \geq \frac{1}{2}.$$

Es gibt also in der Tat ein genügend kleines $\epsilon > 0$, so dass

$$|w_\epsilon(x)| \geq \tfrac{1}{2} \quad \text{für alle } x \in \mathbb{R}^n,$$

also insbesondere $w_\epsilon(x) \neq 0$.

Betrachte nun die glatte(!) Abbildung

$$\tilde{w} : \overline{B_2(0)} \to \partial\overline{B_2(0)}, \quad \tilde{w}(x) = \frac{2w_\epsilon(x)}{|w_\epsilon(x)|}.$$

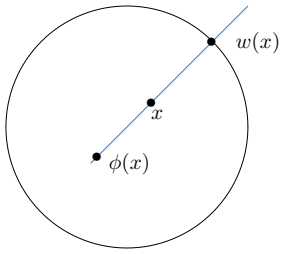
Für alle $x \in \partial\overline{B_2(0)}$ gilt dann

$$\tilde{w}(x) \overset{(7.18)}{=} \frac{2x}{|x|} \overset{x \in \partial\overline{B_2(0)}}{=} \frac{2x}{2} = x.$$

Widerspruch zu dem Resultat aus Schritt 1!

<u>3. Schritt:</u>

Angenommen, die stetige Funktion $\phi : \overline{B_1(0)} \to \overline{B_1(0)}$ besitzt keinen Fixpunkt. Definiere die Abbildung $w : \overline{B_1(0)} \to \partial\overline{B_1(0)}$ dadurch, dass $w(x)$ der Punkt auf $\partial\overline{B_1(0)}$ ist, der vom Strahl, der aus $\phi(x)$ startet und durch x geht, getroffen wird. Wir bemerken, dass w wohldefiniert ist, da stets $\phi(x) \neq x$ für alle $x \in \overline{B_1(0)}$ gilt. Offensichtlich ist



$$w(x) = x \quad \text{für alle } x \in \partial\overline{B_1(0)}.$$

Weiter ist w stetig. Widerspruch zu dem Resultat aus Schritt 2! Fertig. □

7.5.3 Der Satz von Kakutani

In Kapitel 6 bzw. in Beispiel 6.1 haben wir gesehen, dass der Brouwersche Fixpunktsatz nicht in unendlich-dimensionalen Räumen funktioniert. Selbst mit zusätzlichen Voraussetzungen an den Raum bekommt man keine Verbesserung:

Satz 7.10 (Kakutani, 1943). *Sei H ein unendlich-dimensionaler separabler Hilbertraum. Dann gibt es eine stetige Abbildung $f : H \to H$, die die abgeschlossene Einheitskugel in sich selbst abbildet und keinen Fixpunkt besitzt.*

Beweis. Da H separabel ist, gibt es nach Satz 5.5 eine Orthonormalbasis $(y_n)_{n\in\mathbb{Z}}$ von H, d.h. es gilt

$$(y_n, y_m) = \delta_{nm} \quad \text{für alle } n, m \in \mathbb{Z}. \tag{7.19}$$

Definiere für alle $n \in \mathbb{Z}$ die Abbildung

$$\tilde{U} : H \to H, \quad y_n \to y_{n+1}.$$

Nach Satz 5.6 hat man für jedes $x \in H$ die Darstellung

$$x = \sum_{i=-\infty}^{\infty} \alpha_i y_i,$$

wobei

$$\sum_{i=-\infty}^{\infty} |\alpha_i|^2 = \sum_{i=-\infty}^{\infty} |(x, y_i)|^2 \overset{\text{Parseval}}{=} |x|^2 < \infty. \tag{7.20}$$

Wir können daher die Abbildung $\tilde{U}$ erweitern auf beliebige Elemente x durch

$$U(x) = \sum_{i=-\infty}^{\infty} \alpha_i y_{i+1}. \tag{7.21}$$

Zunächst ist U linear, denn für $x, y \in H$ und $\lambda, \mu \in \mathbb{K}$ gilt

$$\begin{aligned} U(\lambda x + \mu y) &= U(\lambda \sum \alpha_i y_i + \mu \sum \beta_i y_i) = U(\sum (\lambda\alpha_i + \mu\beta_i) y_i) \\ &= \sum (\lambda\alpha_i + \mu\beta_i) y_{i+1} = \lambda \sum \alpha_i y_{i+1} + \mu \sum \beta_i y_{i+1} = \lambda U(x) + \mu U(y). \end{aligned}$$

Weiter ist U beschränkt, weil

$$\begin{aligned} |U(x)|^2 &= (U(x), U(x)) = (\sum_i \alpha_i y_{i+1}, \sum_j \alpha_j y_{j+1}) \\ &= \sum_i \alpha_i \sum_j \bar{\alpha}_j \underbrace{(y_{i+1}, y_{j+1})}_{=\delta_{ij}} = \sum_i |\alpha_i|^2 \overset{(7.20)}{<} \infty. \end{aligned}$$

Insgesamt haben wir damit gezeigt, dass U stetig ist. Außerdem bildet U die Menge

$$S_r := \{\, x \in H \mid |x| = r \,\}, \quad 0 < r < \infty,$$

in sich selbst ab, weil für $x \in S_r$ gilt gerade

$$|U(x)|^2 = \sum_i |\alpha_i|^2 \overset{\text{Parseval}}{=} |x|^2 = r^2.$$

Definiere jetzt

$$f(x) := \tfrac{1}{2}(1 - |x|) y_0 + U(x).$$

Offensichtlich ist f stetig und bildet die Kugel $\overline{B_1(0)}$ in sich selbst ab, denn für alle $x \in H$ mit $|x| \leq 1$ gilt

$$|f(x)| \leq \tfrac{1}{2}(1-|x|)\underbrace{|y_0|}_{=1} + \underbrace{|U(x)|}_{=|x|} = \tfrac{1}{2}(1-|x|) + |x| \leq 1.$$

Angenommen f hat auf $\overline{B_1(0)}$ einen Fixpunkt, d.h. es gibt ein $x_0 \in \overline{B_1(0)}$ mit $f(x_0) = x_0$. Das kann man schreiben als

$$x_0 - U(x_0) = \tfrac{1}{2}(1-|x_0|)y_0. \tag{7.22}$$

Wir untersuchen drei Fälle:

<u>1. Fall:</u> $x_0 = 0$. Dann folgt aus (7.22) der Widerspruch

$$0 = \tfrac{1}{2}y_0,$$

denn $y_0 \neq 0$ als Basisvektor.

<u>2. Fall:</u> $|x_0| = 1$. Aus (7.22) ergibt sich dann

$$x_0 = U(x_0) \tag{7.23}$$

Für $x_0 = \sum \alpha_i y_i$ gilt außerdem $\sum |\alpha_i|^2 = 1$, d.h. $\alpha_i \neq 0$ für mindestens ein $i \in \mathbb{Z}$. Mithilfe von (7.21) ergibt sich (7.23) zu

$$\sum_{i=-\infty}^{\infty} \alpha_i y_i = \sum_{i=-\infty}^{\infty} \alpha_i y_{i+1}.$$

Jetzt bilden wir das Skalarprodukt mit y_j, $j \in \mathbb{Z}$, und erhalten

$$\alpha_j = \alpha_{j-1} \quad \text{für alle } j \in \mathbb{Z}.$$

Damit sind alle α_j gleich, also etwa $\alpha_j \equiv: \alpha \neq 0$ und daher

$$\sum_{j\in\mathbb{Z}} |\alpha_j|^2 = \sum_{j\in\mathbb{Z}} |\alpha|^2 = \infty \neq 1.$$

Widerspruch!

<u>3. Fall:</u> $0 < |x_0| < 1$. Sei dazu $x_0 = \sum \alpha_i y_i$. Dann muss wieder $\sum |\alpha_i|^2 < 1$ gelten. Wir setzen nun den Ausdruck für x_0 in (7.22) ein und kriegen

$$\sum_{i\in\mathbb{Z}} (\alpha_i - \alpha_{i-1}) y_i = \tfrac{1}{2}(1-|x_0|)y_0.$$

Daraus folgt

$$\begin{aligned} \alpha_0 - a_1 &= \tfrac{1}{2}(1-|x_0|) > 0, \\ \alpha_i &= \alpha_{i-1}, \quad i \neq 0. \end{aligned}$$

Insgesamt ergibt sich damit

$$\ldots = \alpha_{-2} = \alpha_{-1} < \alpha_0 = \alpha_1 = \ldots$$

und daher der Widerspruch

$$\sum_{i\in\mathbb{Z}} |\alpha_i|^2 = \infty.$$

Wir stellen fest, dass wir in allen drei Fällen einen Widerspruch erhalten. Damit muss die Annahme falsch sein und f hat doch keinen Fixpunkt. $\square$