

Einführung in die Funktionalanalysis

WiSe 2012/13

J. Baumeister¹

13. Februar 2013

¹Dies sind Aufzeichnungen, die kritisch zu lesen sind, da sie noch nicht endgültig korrigiert sind, und daher auch nicht zitierfähig sind (**Not for quotation without permission of the author**). Hinweise auf Fehler und Verbesserungsvorschläge an baumeister@math.uni-frankfurt.de

Vorwort

Ich habe bemerkt, sagte Herr K., dass wir viele abschrecken von unserer Lehre dadurch, dass wir auf alles eine Antwort wissen. Könnten wir nicht im Interesse der Propaganda eine Liste der Fragen aufstellen, die uns ganz ungelöst erscheinen?

B. Brecht, Geschichten vom Herrn Keuner; gefunden bei [37]

Funktionalanalysis kann in ihrer linearen Theorie als Zusammenführung von Analysis (in mehreren Variablen), Linearer Algebra und topologischen Aspekten angesehen werden. Wesentlich ist, dass die Theorie insbesondere die unendlichdimensionalen Vektorräume im Auge hat. Daraus ergeben sich etwa folgende Fragen:

- Was ist Vollständigkeit?
- Was wird aus dem Satz von Heine-Borel?
- Wann kann man (Folgen-)Kompaktheit haben?
- Welche Basen stehen hier zur Verfügung?
- Wie ist eine Dimensionsformel zu interpretieren?
- Wie ordnen sich Eigenwerte ein?
- Wie sieht eine brauchbare Definition der Stetigkeit aus?
- Wo kann man Orthogonalität definieren?
- Was sind Hyperebenen?
- Wie ordnet sich das Konzept „Konvexität“ ein?

Wir beginnen mit den topologischen Strukturen und grundlegenden funktionalanalytischen Begriffen, die für die Verwendung von Funktionenräumen essentiell sind. Wir beginnen dabei mit zunächst mit sehr wenig Struktur, nämlich mit „Topologischen Räumen“ und stetigen Abbildungen in diesen Räumen. Betrachtet man Topologien auf Vektorräumen, gelangt man zu den topologischen Vektorräumen, wenn man die Stetigkeit der Addition und skalaren Multiplikation fordert. Anschließend betrachten wir spezielle Topologien, die sich aus Metriken ergeben. Erste wichtige Resultate sind hier der Kontraktionssatz und der Satz von Baire.

Normierte Räume sind Vektorräume, in denen den Vektoren eine Länge zugewiesen wird. Hier ordnen sich auch die endlichdimensionalen euklidischen Vektorräume $\mathbb{R}^n$ ein. Wir stellen die wichtigsten Ergebnisse für die normierten Räume zusammen, wobei wir insbesondere auf die vollständigen normierten Räume (Banachräume) eingehen. In diesem Rahmen diskutieren wir dann die Abbildungen, die zur Analysis (Stetigkeit) und Linearer Algebra (Linearität) passen: Stetige lineare Operatoren.

Dualitätsüberlegungen sind ein wichtiges Handlungsprinzip in der Mathematik. In der Funktionalanalysis kommt es hauptsächlich beim Studium der Dualräume zum Tragen. Das Hauptergebnis ist der Satz von Hahn-Banach, der etwas aussagt über die Reichhaltigkeit der Dualräume.

Daran schließen sich Möglichkeiten an, zusätzlich Topologien einzuführen, die sich in Anwendungen nützen lassen, insbesondere, was Kompaktheitsfragen betrifft.

Im Kapitel über Konvexität betrachten wir geometrische Interpretation des Satzes von Hahn-Banach. Hier ordnen sich geometrische und algorithmische Aspekte ein. Von großer Bedeutung in Anwendungen ist die Interpretation der Ergebnisse der linearen Funktionalanalysis im Hilbertraum, insbesondere in der Theorie der linearen partiellen Differentialgleichungen.

Nun folgt das Kapitel mit der Überschrift „Prinzipien der (linearen) Funktionalanalysis“. Hier werden wichtige Abbildungseigenschaften stetiger Abbildungen zusammengestellt. Hierher gehören eigentlich auch der Satz von Baire und der Satz von Hahn-Banach. Wir hielten es aber für zweckmäßig, sie schon früher aufzuführen.

Die Überlegungen zu den Spektren von linearen Operatoren verallgemeinern die Eigenwerttheorie der Linearen Algebra. Besonders einsichtig werden Spektralaussagen bei kompakten Operatoren, Operatoren, die bei Integralgleichungen „das Feld beherrschen“. Hier führen wir auch unbeschränkte Operatoren und deren Spektren ein.

Abschließend gehen wir auf die Fixpunkttheorie und die Anwendung der linearen Theorie für nichtlineare Gleichungen ein.

In unseren Betrachtungen – die oben gestellten Fragen finden darin auch ihre Antworten – wollen wir stets versuchen, den konstruktiven/angewandten Standpunkt einzunehmen. An vielen Stellen deuten wir Anwendungen der Ergebnisse an. Meist beziehen sie sich auf Differential- und Integralgleichungen. Dabei tut sich das Problem auf, dass dazu geeignete Funktionenräume nötig sind, deren Einführung und Nutzung intensiv den Lebesgueschen Integralbegriff nutzen. Darum werden wir uns „herummogeln“, da nicht sicher gestellt ist, dass dieser Integralbegriff allgemein bekannt (und verstanden) ist.

Der Ursprung der Funktionalanalysis liegt am Anfang dieses Jahrhunderts, initiiert durch Arbeiten von Fredholm, Lebesgue und Hilbert. Später wurde sie durch Banach und von Neumann, die die heute geläufigen Begriffe des normierten Raums und des Hilbertraums prägten, kanonisiert. Ihre Beiträge sind eng verbunden mit der Begründung der heutigen Form der Funktionalanalysis. Besonders hervorzuheben ist hier die *Lemberger Schule*, zu der u. a. Banach, Kacmarz, Mazur, Orlicz, Saks, Schauder, Steinhaus, Ulam gehörten; von einigen werden wir noch hören.

Die geschichtliche Entwicklung rund um die „reine“ Funktionalanalysis stellt sich etwa so dar:¹

- Fredholms Artikel über Integral Gleichungen (1900)
- Lebesgues Dissertation über Integration (1902)
- Hilberts Aufsatz über Spektraltheorie (1906)
- Frechets Dissertation über metrische Räume (1906)
- Riesz Artikel über $C[a, b]$ and L_p (1910,1911)
- Hellys Aufsatz (1912,1921)
- Banachs Dissertation über normierte Räume (1922)
- Hahns Artikel (1927) und Banachs Aufsatz über Dualität (1929)
- Banachs und Steinhaus' Artikel (1927)
- Frechets Buch „Les espaces abstraits“ (1928) and Banachs „Theorie des operations lineaires“ (1932)
- Grothendiecks Formulierung des Approximationsproblems (1955)
- Beantwortung des Approxiamtionsproblems durch Enflo (1973)

¹Mehr zur Geschichte findet man bei J. Lindström; siehe [26].

Die Geschichte der „angewandten“ Funktionalanalysis ist eng mit der Entwicklung der Theorie der (partiellen) Differentialgleichungen, insbesondere auch innerhalb der Quantenmechanik, verknüpft:

- Fixpunkttheorie und gewöhnliche Differentialgleichungen im 18. und 19. Jahrhundert
- Untersuchungen zu klassischen Lösungen der Gleichungen der mathematischen Physik, begonnen durch D'Alembert und L. Euler im 18. Jahrhundert.
- Fourier Analysis zur Lösung der Gleichungen der mathematischen Physik um 1800
- Verknüpfung der partiellen Differentialgleichungen mit Integralgleichungen; Cauchy, Dirichlet und andere
- Entwicklung der Potentialtheorie durch Legendre, Laplace, Green und andere
- Formulierung der Fredholm Alternative für Integralgleichungen
- Entwicklung der Integrationstheorie durch Lebesgue
- Studium der kompakten Operatoren und ihrer Spektraldarstellung
- Spektraltheorie bei unbeschränkten Operatoren

Nicht alles Wünschenswerte hat aus Zeitgründen und auf Grund der Vorlieben des Dozenten Platz gefunden in der Vorlesung und im Skript. Darunter sind:

- $L_p(\Omega)$ -Räume und Sobolevräume
- Fouriertransformation
- Kompakte Operatoren
- Anfänge der Theorie linearer partieller Differentialgleichungen

Das Fundament für ein (Eigen-)Studium dieser Themen sollte allerdings gelegt sein.

Funktionalanalytische Kenntnisse sind mittlerweile in vielen Disziplinen der Mathematik wie Differentialgleichungen, Numerik, Wahrscheinlichkeitstheorie oder Approximationstheorie sowie in der theoretischen Physik unabdingbar. Lehrbücher zur Funktionalanalysis sind meist ausgerichtet auf solche spezielle Anwendungen (Struktur von Banachräumen, Partielle Differentialgleichungen, Halbgruppentheorie, Fixpunkttheorie, Integralgleichungen, Numerische Mathematik, ...). Wir listen als Beispiele auf: [1, 2, 3, 4, 5, 10, 20, 19, 22, 23, 25, 35, 37, 39]. Speziellere Literatur führen wir an geeigneten Stellen der Betrachtungen an.

Bei der Erstellung des Skriptums wurde vielfältige Literatur verwendet. Lehrbuchliteratur und Originalartikel sind im Allgemeinen zitiert. Für allgemeine zusätzliche Erkenntnisse aus dem Internet sind meist die Quellen nicht angegeben.

Frankfurt, im Februar 2013

JOHANN BAUMEISTER

Inhaltsverzeichnis

Vorwort	i
1 Topologische Aspekte: Eine kurze Zusammenfassung	1
1.1 Topologische Räume	1
1.2 Folgen in topologischen Räumen	4
1.3 Kompaktheit	5
1.4 Stetigkeit	7
1.5 Übungen	9
2 Metrische Räume	11
2.1 Metrische Räume	11
2.2 Stetigkeit	15
2.3 Vollständigkeit und der Kontraktionssatz	17
2.4 Kompaktheit	19
2.5 Der Satz von Arzela-Ascoli	22
2.6 Der Satz von Baire	26
2.7 Approximationssätze	27
2.8 Übungen	29
3 Normierte Räume	35
3.1 Normen	35
3.2 Banachräume	38
3.3 Stetigkeit	40
3.4 Stetige lineare Funktionale	45
3.5 Basen	50
3.6 Übungen	56
4 Prinzipien/Hauptsätze der Funktionalanalysis	60
4.1 Der Satz von Banach–Steinhaus	60
4.2 Der Satz von der offenen Abbildung	63
4.3 Der Satz vom abgeschlossenen Graphen	65
4.4 Approximationsprinzip: Projektionen	67
4.5 Fixpunktprinzip I–Der Banachsche Fixpunktsatz	73
4.6 Übungen	77
5 Dualität	81
5.1 Der Satz von Hahn–Banach	81
5.2 Reflexivität	87
5.3 Schwache Topologien	89
5.4 Duale Operatoren	94

5.5	Übungen	95
6	Konvexität	99
6.1	Konvexe Mengen	99
6.2	Trennungssätze	103
6.3	Satz vom abgeschlossenen Bild	109
6.4	Der Satz von Krein–Milman	113
6.5	Strikte und gleichmäßige Konvexität	115
6.6	Anhang: Differenzierbare Normen	120
6.7	Anhang: Differenzierbarkeit nichtlinearer Abbildungen	123
6.8	Anhang: Fixpunktprinzip II–Der Schaudersche Fixpunktsatz	128
6.9	Übungen	130
7	Hilberträume	135
7.1	Prähilberträume	135
7.2	Hilberträume: Definition und Beispiele	138
7.3	Der Darstellungssatz von Riesz	140
7.4	Orthonormalbasen	143
7.5	Selbstadjungierte Operatoren	145
7.6	Das Spektrum beschränkter Operatoren	147
7.7	Anhang: Abstandserhaltende Abbildungen	150
7.8	Anhang: Das Lemma von Lax–Milgram	151
7.9	Anhang: Gelfand–Dreier	155
7.10	Übungen	156
	Literatur	159

Kapitel 1

Topologische Aspekte: Eine kurze Zusammenfassung

Wer das erste Knopfloch verfehlt, kommt mit dem
Zuknöpfen nicht zu Rande

J. W. Goethe

In diesem Kapitel bringen wir die Begriffe „Umgebung, Konvergenz, Grenzwert, Kompaktheit“ zusammen. Die „Struktur“, die dies möglich macht, ist die des topologischen Raumes. Für die Betrachtungen wesentlich ist, dass wir mit den topologischen Begriffen nicht nur Nachbarschaften zwischen Punkten, sondern auch zwischen abstrakten Objekten, u. a. auch von Abbildungen erklären können. Eine ausführliche Monographie zum Thema ist [30].

1.1 Topologische Räume

Bei der Analysis von Funktionen einer Variablen haben wir immer wieder mit offenen Intervallen (a, b) hantiert. Dabei fällt auf, dass diese Intervalle gerade als Nachbarschaften aller ihrer Punkte aufgefasst werden können. Diese Eigenschaft geht auch nicht verloren, wenn wir beliebige Vereinigungen solcher Intervalle betrachten. Davon lässt sich das, was in der (mengentheoretischen) Topologie geschieht, leiten.

Definition 1.1 Sei X eine Menge, $X \neq \emptyset$. Eine **Topologie** $\mathcal{T}$ auf X ist eine Familie von Teilmengen von X mit den folgenden Eigenschaften

- a) $\emptyset \in \mathcal{T}, X \in \mathcal{T}$;
- b) $X_1, X_2 \in \mathcal{T} \implies X_1 \cap X_2 \in \mathcal{T}$;
- c) $X_i \in \mathcal{T}, i \in I$ (I Indexmenge) $\implies \bigcup_{i \in I} X_i \in \mathcal{T}$.

Die Elemente in einer Topologie $\mathcal{T}$ auf X heißen **offene Mengen**, das Tupel $(X, \mathcal{T})$ heißt **topologischer Raum**. Die Elemente $x \in X$ bezeichnen wir hier meist als **Punkte**. $\square$

Definition 1.2 Sind $\mathcal{T}_1, \mathcal{T}_2$ Topologien auf X , so heißt $\mathcal{T}_1$ **feiner** als $\mathcal{T}_2$ oder $\mathcal{T}_2$ **gröber** als $\mathcal{T}_1$, wenn $\mathcal{T}_2 \subset \mathcal{T}_1$ gilt. $\square$

Klar, die Relation „feiner“ induziert auf der Familie der Topologien auf X eine Halbordnung wie schon ein flüchtiger Blick lehrt.

Sei X eine Menge, $X \neq \emptyset$. Die grösste Topologie auf X ist durch $\mathcal{T} := \{\emptyset, X\}$ gegeben. Die feinste Topologie auf X erhält man durch $\mathcal{T} := \text{POT}(X)$; sie heisst **diskrete Topologie**. Aussagekräftigere Beispiele topologischer Räume lernen wir in den nächsten Abschnitten kennen.

Definition 1.3 Sei $(X, \mathcal{T})$ ein topologischer Raum. Dann heisst $A \subset X$ **abgeschlossen**, wenn das Komplement $X \setminus A$ offen ist. $\square$

Klar, in einem topologischen Raum $(X, \mathcal{T})$ sind X und $\emptyset$ stets offen und abgeschlossen.

Definition 1.4 Sei $(X, \mathcal{T})$ ein topologischer Raum, $x \in X$. Eine Teilmenge U von X heisst **Umgebung** von x , wenn es $V \in \mathcal{T}$ gibt mit $x \in V \subset U$. Mit $\mathcal{U}(x)$ bezeichnen wir die Familie aller Umgebungen von x . $\square$

Klar, Obermengen von Umgebungen sind selbst wieder Umgebungen. Ebenso folgt mit b) aus Definition 1.1, dass der Schnitt von zwei Umgebungen eines Punktes selbst wieder eine Umgebung dieses Punktes ist.

Satz 1.5 Sei $(X, \mathcal{T})$ ein topologischer Raum, sei $A \subset X$. Dann sind äquivalent:

- a) A ist offen, d.h. $A \in \mathcal{T}$.
- b) Für alle $x \in A$ ist $A \in \mathcal{U}(x)$.

Beweis:

Zu a) $\implies$ b). Klar nach Definition der Umgebung.

Zu b) $\implies$ a). Zu $x \in A$ gibt es $V_x \in \mathcal{T}$ mit $x \in V_x \subset A$. In dieser Bezeichnung haben wir also $\cup_{x \in A} V_x = A$. Nach c) in Definition 1.1 ist $\cup_{x \in A} V_x$ offen, also auch A . $\blacksquare$

Ist nun der schon uns vertraute „Raum“ $\mathbb{R}$ auch als topologischer Raum zu verstehen? Ja, der Satz 1.5 sagt uns, wie das unter Verwendung der offenen Intervalle (a, b) gehen kann:

$$A \subset \mathbb{R} \text{ offen} : \iff \forall x \in A \exists r > 0 ((x - r, x + r) \subset A).$$

Man überzeugt sich leicht, dass damit eine Familie offener Mengen erklärt wird, die eine Topologie auf $\mathbb{R}$ ist. Wir wollen sie die „euklidische Topologie“ nennen, denn wir werden sie später in Beziehung zur „euklidischen Vorstellung“ des Raumes $\mathbb{R}^1$ setzen.

Definition 1.6 Unter der **Produkttopologie** auf dem Produkt $X \times Y$ zweier topologischer Räume $(X, \mathcal{T}_X)$ und $(Y, \mathcal{T}_Y)$ versteht man die grösste Topologie $\mathcal{T}_{X \times Y}$, die alle Mengen $A \times B$ offener Mengen $A \in \mathcal{T}_X$ und $B \in \mathcal{T}_Y$ enthält, also

$$\mathcal{T}_{X \times Y} = \bigcap \{ \mathcal{T}' \mid \mathcal{T}' \text{ Topologie auf } X \times Y \text{ mit } A \times B \in \mathcal{T}' \text{ für alle } A \in \mathcal{T}_X, B \in \mathcal{T}_Y \}.$$

$\square$

Damit sind in offensichtlicher Weise Topologie auf $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}, \mathbb{R}^3 = \mathbb{R} \times \mathbb{R}^2, \dots, \mathbb{R}^n$ erklärt. Später werden wir auch eine Produkttopologie auf einem unendlichen Produkt betrachten.

Definition 1.7 Sei $(X, \mathcal{T})$ ein topologischer Raum und sei $Y \subset X, Y \neq \emptyset$. Dann ist

$$\mathcal{T}_Y := \{ B \subset Y \mid \exists A \in \mathcal{T} \text{ mit } B = Y \cap A \}$$

eine Topologie auf Y , die durch $\mathcal{T}$ auf Y **induzierte Topologie**. $\square$

Der Nachweis, dass in der Tat eine Topologie vorliegt, ist einfach.

Wir wissen, dass ein topologischer Raum $(X, \mathcal{T}_X)$ in jedem Fall offene und abgeschlossene Mengen enthält, nämlich X und $\emptyset$. Damit wird die folgende Begriffsbildung möglich:

$$\overline{A} := \bigcap \{B \subset X \mid A \subset B, B \text{ abgeschlossen}\}, \quad A^\circ := \bigcup \{C \subset X \mid C \subset A, C \text{ offen}\}.$$

$\overline{A}$ heißt die **abgeschlossene Hülle** oder **Abschluss** von A und A° der **offene Kern** von A .

Folgerung 1.8 Sei $(X, \mathcal{T})$ ein topologischer Raum, $A \subset X$. Dann gilt:

a) $\overline{A}$ ist abgeschlossen, $\overline{(\overline{A})} = \overline{A}$.

b) A° ist offen, $(A^\circ)^\circ = A^\circ$.

Beweis:

Die erste Behauptung in a) folgt aus der Beobachtung

$$X \setminus \overline{A} = \bigcup \{X \setminus B \mid A \subset B, B \text{ abgeschlossen}\}.$$

Die zweite Aussage folgt so. Da $\overline{A}$ abgeschlossen ist und $\overline{A} \subset \overline{A}$ ist, gilt nach Definition von $\overline{(\overline{A})}$ schließlich $\overline{(\overline{A})} \subset \overline{A}$. Die Inklusion $\overline{A} \subset \overline{(\overline{A})}$ gilt offensichtlich.

Die erste Aussage in b) folgt aus der Tatsache, dass A° Vereinigung offener Mengen ist, den Beweis der zweiten Behauptung überlassen wir dem Leser. ■

Bei der Hüllenbildung, um aus A den Abschluss $\overline{A}$ zu bekommen, kommen im Allgemeinen Punkte zu A dazu. Wir werden sehen, dass der Abschluss auch als Menge der Berührungspunkte definiert werden könnte.

Definition 1.9 Sei $(X, \mathcal{T})$ ein topologischer Raum, $A \subset X$. $x \in X$ heißt **Berührungspunkt** von A , wenn gilt:

$$\forall U \in \mathcal{U}(x) (U \cap A \neq \emptyset).$$

□

Lemma 1.10 Sei $(X, \mathcal{T})$ ein topologischer Raum, $A \subset X$. Dann gilt

$$\overline{A} = \{x \in X \mid x \text{ Berührungspunkt von } A\}.$$

Beweis:

Sei $\tilde{A} := \{x \in X \mid x \text{ Berührungspunkt von } A\}$. Wir haben also $\overline{A} = \tilde{A}$ zu zeigen.

Sei $x \in \overline{A}$. Annahme $x \notin \tilde{A}$. Dann gibt es $U \in \mathcal{U}(x)$ mit $U \cap A = \emptyset$. O.E. U ist offen. Dann gilt $A \subset X \setminus U$ und $X \setminus U$ ist abgeschlossen. Da wir $x \in \overline{A}$ haben, gilt nach der Definition des Abschlusses $x \in X \setminus U$, was ein Widerspruch ist.

Sei $x \in \tilde{A}$. Sei $B \subset X$, B abgeschlossen und $A \subset B$. Dann ist $X \setminus B$ offen, $X \setminus B \subset X \setminus A$. Annahme $x \in B$. Dann ist $X \setminus B \cap A \neq \emptyset$, was ein Widerspruch ist. Also gilt $x \in B$. Da B mit den obigen Eigenschaften beliebig war, gilt $x \in \overline{A}$. ■

Satz 1.11 Sei $(X, \mathcal{T})$ ein topologischer Raum, $A \subset X$. Dann gilt:

(a) A ist abgeschlossen genau dann, wenn $A = \overline{A}$ gilt.

(b) A ist offen genau dann, wenn $A = A^\circ$ gilt.

Zu a). Die Rückrichtung ist klar, da $\overline{A}$ abgeschlossen ist. Die Hinrichtung folgt so. Offenbar ist $A \subset \overline{A}$. Da $A \subset A$ gilt und A abgeschlossen ist, folgt aus der Definition des Abschlusses $\overline{A} \subset A$.
 Zu b). Die Rückrichtung ist klar, da A° offen ist. Die Hinrichtung folgt so. Offenbar ist $A^\circ \subset A$. Da $A \subset A$ gilt und A offen ist, folgt aus der Definition des offenen Kerns $A \subset A^\circ$.

Definition 1.12 Sei $(X, \mathcal{T})$ ein topologischer Raum. Eine Teilmenge A von X heißt **dicht** (in X), wenn $\overline{A} = X$ gilt.

Enthält X eine abzählbare dichte Teilmenge A , so heißt der Raum $(X, \mathcal{T})$ **separabel**. □

Beispielsweise sind die rationalen Zahlen dicht in den reellen Zahlen.

Definition 1.13 Sei $(X, \mathcal{T})$ ein topologischer Raum, sei $A \subset X$. x heißt **Häufungspunkt** von A , wenn es für alle $U \in \mathcal{U}(x)$ gilt: $U \setminus \{x\} \cap A \neq \emptyset$. Wir setzen $A' := \{x \in X \mid x \text{ Häufungspunkt von } A\}$. □

Es ist klar, dass $A' \subset \overline{A}$ gilt. Ohne Beweis geben wir an:

$$\overline{A} = A \cup A'. \quad (1.1)$$

Definition 1.14 Sei $(X, \mathcal{T})$ ein topologischer Raum, sei $A \subset X$. x heißt **Randpunkt** von A , wenn es ein $U \in \mathcal{U}(x)$ gibt mit $U \cap A \neq \emptyset$ und $U \cap X \setminus A \neq \emptyset$. Die Menge $\partial A := \{x \in X \mid x \text{ Randpunkt von } A\}$ heißt **Rand** von A . □

Klar, der Rand einer Menge ist stets abgeschlossen und es ist leicht einzusehen, dass eine Menge A genau dann abgeschlossen ist, wenn $\partial A \subset A$ gilt.

1.2 Folgen in topologischen Räumen

Die Definition einer Folge in einem topologischen Raum $(X, \mathcal{T})$ ist analog zur Definition in $\mathbb{R}$ bzw. $\mathbb{R}^n$: eine Folge $(x_n)_{n \in \mathbb{N}}$ in X ist eine Abbildung von $\mathbb{N}$ in X gemäß

$$\mathbb{N} \ni n \longmapsto x_n \in X.$$

Eine Teilfolge $(x_{n_k})_{k \in \mathbb{N}}$ einer solchen Folge kommt zustande durch eine injektive Abbildung

$$\mathbb{N} \ni k \longmapsto n_k \in \mathbb{N}.$$

Definition 1.15 Sei $(X, \mathcal{T})$ ein topologischer Raum. Eine Folge $(x_n)_{n \in \mathbb{N}}$ in X **konvergiert** gegen ein x (Grenzwert) in X genau dann, wenn gilt:

$$\forall U \in \mathcal{U}(x) \exists N \in \mathbb{N} \forall n \geq N (x_n \in U).$$

□

Wir wissen aus der Analysis, dass die Abgeschlossenheit einer Menge reeller Zahlen über eine Folgeeigenschaft definierbar ist: Eine Teilmenge A in $\mathbb{R}$ ist abgeschlossen genau dann, wenn jede konvergente Folge in A einen Grenzwert in A besitzt. Dies ist in allgemeinen topologischen Räumen so nicht richtig, genauer gesagt, die Folgeeigenschaft ist nur notwendig. Im allgemeinen Fall benötigt man als Ersatz für die Folge den Begriff des Netzes. Wir gehen nicht darauf ein, sondern benennen topologische Räume, in denen man mit der Folgeeigenschaft wieder auskommt; für unsere Zwecke reicht dies aus.

Definition 1.16 Sei $(X, \mathcal{T})$ ein topologischer Raum. $(X, \mathcal{T})$ erfüllt das **1. Abzählbarkeitsaxiom**, wenn es für alle $x \in X$ eine Mengenfamilie $(U_{n,x})_{n \in \mathbb{N}}$ gibt mit

- (a) $U_{n,x} \in \mathcal{U}(x)$ für alle $n \in \mathbb{N}$,
- (b) zu $U \in \mathcal{U}(x)$ gibt es $n \in \mathbb{N}$ mit $x \in U_{n,x} \subset U$.

Man nennt die Familie $(U_{n,x})_{n \in \mathbb{N}}$ eine **abzählbare Umgebungsbasis von x** . □

Satz 1.17 Sei $(X, \mathcal{T})$ ein topologischer Raum, der das 1. Abzählbarkeitsaxiom erfüllt. Dann gilt:

$$\bar{A} = \{x \in X \mid \exists (x_n)_{n \in \mathbb{N}} \text{ mit } x_n \in A, n \in \mathbb{N}, x = \lim_n x_n\}.$$

Beweis:

Sei $\tilde{A} := \{x \in X \mid \exists (x_n)_{n \in \mathbb{N}} \text{ mit } x_n \in A, n \in \mathbb{N}, x = \lim_n x_n\}$. Wir haben zu zeigen $\bar{A} \subset \tilde{A}$.

Sei $x \in \bar{A}$. Da nach Lemma 1.10 x ein Berührungspunkt ist, gilt $U_{1,x} \cap \dots \cap U_{n,x} \cap A \neq \emptyset$. Wähle $x_n \in U_{1,x} \cap \dots \cap U_{n,x} \cap A$. Damit erhalten wir eine Folge $(x_n)_{n \in \mathbb{N}}$ in A . Wir zeigen $x = \lim_n x_n$. Sei dazu $U \in \mathcal{U}(x)$. Dann gibt es $N \in \mathbb{N}$ mit $U_{N,x} \subset U$. Also $x_n \in U_{N,x} \subset U$ für alle $n \in \mathbb{N}$.

Sei $x \in \tilde{A}$, $x = \lim_n x_n$, $x_n \in A$ für alle $n \in \mathbb{N}$. Annahme: $x \notin \bar{A}$. Dann ist $x \in X \setminus \bar{A}$ und da $X \setminus \bar{A}$ offen ist, gibt es $N \in \mathbb{N}$ mit $x_n \in X \setminus \bar{A}$ für alle $n \geq N$. ■

In der Analysis haben wir benutzt, dass der Grenzwert einer Folge eindeutig bestimmt ist. Dies ist die Konsequenz aus der Tatsache, dass Punkte in $\mathbb{R}$ durch Intervalle „getrennt“ werden können. Dazu die Begrifflichkeit in einem topologischen Raum.

Definition 1.18 Ein topologischer Raum $(X, \mathcal{T})$ heißt **Hausdorffraum** oder **hausdorffsch**, wenn gilt:

$$\forall x, y \in X : x \neq y \exists U \in \mathcal{U}(x) \exists V \in \mathcal{U}(y) (U \cap V = \emptyset)$$

□

Es ist nun sofort klar, dass der Grenzwert einer Folge eindeutig bestimmt ist, wenn der zugrundeliegende topologische Raum hausdorffsch ist.¹

Beispiel 1.19 Ein topologischer Raum $(X, \mathcal{T})$ ist sicher dann nicht hausdorffsch, wenn $\mathcal{T} = \{\emptyset, X\}$ ist mit $\#X \geq 2$. □

1.3 Kompaktheit

Kompaktheit ist (zusammen mit der Stetigkeit) ein überaus wichtiges Konzept der Topologie und damit auch der Funktionalanalysis.

Definition 1.20 Der topologische Raum $(X, \mathcal{T})$ heißt **kompakt**, wenn $(X, \mathcal{T})$ hausdorffsch ist und wenn jede offene Überdeckung von X eine endliche Überdeckung enthält, d. h. wenn aus $X = \cup_{i \in I} X_i$, $X_i \in \mathcal{T}$ für alle $i \in I$, folgt, dass es $i_1, \dots, i_l \in I$ gibt mit $X = \cup_{j=1}^l X_{i_j}$ (endliche Teilüberdeckung). □

Diese Definition kommt etwas „umständlich“ daher, aber bei genauer Betrachtung sieht man die Idee, die dahinter steckt: Übergang von unendlich vielen Objekten X_i , $i \in I$, zu endlich vielen Objekten $X_{i_1}, \dots, X_{i_l}$. Beachte: Eine endliche offene Überdeckung eines topologischen Raumes existiert immer, die Aussage in Definition 1.20 besagt, dass aus jeder offenen Überdeckung eine endliche Teilüberdeckung ausgewählt werden kann.

¹In der Literatur wird ein Hausdorff-Raum $(X, \mathcal{T})$ auch so angesprochen: $(X, \mathcal{T})$ erfüllt das Trennungsaxiom T_2 .

Beispiel 1.21 Sei $\mathbb{R}$ mit der üblichen Topologie ausgestattet. Offensichtlich ist $\mathbb{R}$ nicht kompakt.

Betrachte nun das abgeschlossene Intervall $A := [0, 1]$ in $\mathbb{R}$. Dieses Intervall ist kompakt. Der Beweis, der nur die Eigenschaften von $\mathbb{R}$ und die Resultate, die wir bisher zur Verfügung haben, benutzt, ist nicht ganz trivial. Sei $\mathcal{O} := (U_i)_{i \in I}$ eine offene Überdeckung von A : $A \subset \bigcup_{i \in I} U_i$, U_i offen in $\mathbb{R}$ für alle $i \in I$. Sei

$$L := \{x \in [0, 1] \mid [0, x] \text{ besitzt eine endliche Überdeckung von } \mathcal{O}\}.$$

Wegen $0 \in L$ ist L nicht-leer, und offenbar ist 1 eine obere Schranke für L . Also existiert auf Grund der Vollständigkeit von $\mathbb{R}$ die kleinste obere Schranke $s := \sup_{x \in L} x$; wenn wir $s = 1 \in L$ zeigen können, sind wir fertig. Zunächst wissen wir aber nur, dass $0 \leq s \leq 1$. Da es eine offene Menge $U := U_i$ in $\mathcal{O}$ geben muss, die 0 enthält, gibt es ein $\varepsilon > 0$ mit $[0, 0 + \varepsilon] \subset U$. Das heißt aber, dass das Intervall $[0, 0 + \varepsilon]$ bereits von einer einzigen Menge in $\mathcal{O}$ überdeckt wird. Daher ist $0 + \varepsilon \in U$, also $0 + \varepsilon \leq s$, $0 < s$. Es muss $s \in V := U_j$ für ein U_j in $\mathcal{O}$ gelten. Dann gibt es ein $x \in V$ mit $0 < x < s$, $[x, s] \subset V$, und $x \in L$. Also besitzt $[0, x]$ eine endliche Teilüberdeckung von $\mathcal{O}$. Da zusätzlich $[x, s]$ von V überdeckt wird, besitzt auch $[0, s]$ eine endliche Teilüberdeckung von $\mathcal{O}$, also ist $s \in L$.

Wir nehmen jetzt an, dass $s < 1$ gilt und führen dies zum Widerspruch. Wir finden eine Menge $W := U_k$ in $\mathcal{O}$ mit $s \in W$, also ein Intervall $[s, s + \varepsilon] \subset W$ für ein $\varepsilon > 0$; damit ist $s + \varepsilon < 1$. Da $[s, s + \varepsilon]$ von W und das Intervall $[0, s]$ von endlich vielen Mengen in $\mathcal{O}$ überdeckt wird, gilt dies auch für das Intervall $[0, s + \varepsilon]$, also ist $s + \varepsilon \in L$, und damit ist s keine obere Schranke von L . Widerspruch!

Betrachte nun das halboffene Intervall $A := (0, 1] \subset \mathbb{R}$. A besitzt die offene Überdeckung $\bigcup_{n \in \mathbb{N}} (1/n, 1 + 1/n) \cap (0, 1]$, aber offenbar gibt es davon keine endliche Teilüberdeckung von A . $\square$

Definition 1.22 Sei $(X, \mathcal{T})$ ein Hausdorff-Raum, sei $A \subset X$.

- a) A heißt **kompakt**, wenn $(A, \mathcal{T}_A)$ kompakt ist, wobei $\mathcal{T}_A$ die durch $\mathcal{T}$ auf A induzierte Topologie ist.
- b) A heißt **relativ kompakt**, wenn $\overline{A}$ kompakt ist.

$\square$

Beachte, dass die Spurtopologie die Eigenschaft, Hausdorff-Raum zu sein, erbt.

Satz 1.23 Sei $(X, \mathcal{T})$ ein Hausdorffraum. Dann gilt:

- a) Ist $(X, \mathcal{T})$ kompakt und $A \subset X$ abgeschlossen, dann ist A kompakt.
- b) Ist $A \subset X$ kompakt, dann ist A abgeschlossen.

Beweis:

Zu a). $B := X \setminus A$ ist also offen. Ist nun $A = \bigcup_{i \in I} (U_i \cap A)$ eine offene Überdeckung von A (in der induzierten Topologie), so wird $X = \bigcup_{i \in I} U_i \cup B$ eine offene Überdeckung von X . Also gibt es $i_1, \dots, i_l \in I$ mit $X = \bigcup_{j=1}^l U_{i_j} \cup B$. Dann ist aber $A = \bigcup_{j=1}^l (U_{i_j} \cap A)$ eine endliche Überdeckung von A .

Zu b). Sei $A \subset X$ kompakt. Setze $U := X \setminus A$. Wir wollen mit Satz 1.5 zeigen, dass U offen ist. Sei $x \in U$. Sei $y \in A$. Da X ein Hausdorffraum ist, gibt es offene Mengen V_y, W_y mit

$$x \in V_y, y \in W_y, V_y \cap W_y = \emptyset.$$

Wenn wir nun die Mengen $W_y, y \in A$, betrachten, so bilden sie eine offene Überdeckung von A , da ja $y \in W_y$ für jedes $y \in A$ gilt. Also enthält diese Überdeckung eine endliche Überdeckung $W_{y_1}, \dots, W_{y_m}$:

$$A \subset W := \bigcup_{i=1}^m W_{y_i}.$$

Also gilt nun

$$x \in V := \bigcap_{i=1}^m V_{y_i}, A \subset W, V \cap W = \emptyset.$$

Da V offen ist und $V \subset X \setminus W \subset X \setminus A = U$ gilt, ist U eine Umgebung von x . ■

Satz 1.24 (Cantorscher Durchschnittssatz) Sei $(X, \mathcal{T})$ ein kompakter topologischer Raum und sei $(A_i)_{i \in I}$ eine Familie nicht-leerer abgeschlossener Teilmengen von X . Gilt dann

$$A_{i_1} \cap \dots \cap A_{i_n} \neq \emptyset \text{ für jede Teilmenge } \{i_1, \dots, i_n\} \text{ von } I,$$

so ist

$$\bigcap_{i \in I} A_i \neq \emptyset.$$

Beweis:

Annahme: $\bigcap_{i \in I} A_i = \emptyset$. Dann gilt $X = \bigcup_{i \in I} X \setminus A_i$ und wir haben eine offene Überdeckung von X . Da X kompakt ist, gibt es $i_1, \dots, i_m \in I$ mit $X = \bigcup_{j=1}^m X \setminus A_{i_j}$, also $\emptyset = \bigcup_{j=1}^m A_{i_j}$, was ein Widerspruch zur Voraussetzung ist. ■

Auf die Voraussetzung „ X kompakt“ in Satz 1.24 kann nicht verzichtet werden, wie nachfolgendes Beispiel lehrt.

Beispiel 1.25 Sei $\mathbb{R}$ ausgestattet mit der üblichen Topologie. Betrachte die Mengenfamilie $(A_i)_{i \in \mathbb{N}}$, definiert als $A_i := [i, \infty)$, $i \in \mathbb{N}$. Offenbar ist jede Menge A_i abgeschlossen und je endlich viele Mengen $A_{i_1}, \dots, A_{i_m}$ haben einen nicht-leeren Schnitt. Aber: $\bigcap_{i \in \mathbb{N}} A_i = \emptyset$. □

1.4 Stetigkeit

Die Abbildungen, die zusammen mit den topologischen Räumen eine interessante Struktur ausmachen, sind die stetigen Abbildungen. Die Definition der Stetigkeit, die wir nun liefern werden, wird nicht sehr überraschen, wenn wir uns daran erinnern, wie die Stetigkeit in x bei einer Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ eingeführt wird: zu vorgegebenem ε -Intervall $I := (f(x) - \varepsilon, f(x) + \varepsilon)$ gibt es ein δ -Intervall $J := (x - \delta, x + \delta)$ mit $f(J) \subset I$.

Definition 1.26 Seien $(X, \mathcal{T}_X), (Y, \mathcal{T}_Y)$ topologische Räume und sei $f : X \rightarrow Y$.

- a) f heißt **stetig in** $x \in X$, wenn gilt: $\forall V \in \mathcal{U}(f(x)) \exists U \in \mathcal{U}(x) (f(U) \subset V)$.
- b) f heißt **stetig**, wenn f stetig ist in jedem $x \in X$.

□

Ist $f : X \rightarrow Y$ eine Abbildung, so ist das Urbild von $B \subset Y$ unter dieser Abbildung bezeichnet durch $f^{-1}(B)$. Also

$$f^{-1}(B) = \{x \in X \mid f(x) \in B\}$$

Satz 1.27 Seien $(X, \mathcal{T}_X), (Y, \mathcal{T}_Y)$ topologische Räume und sei $f : X \rightarrow Y$. Es sind äquivalent:

- a) f ist stetig
- b) Für alle $x \in X$ gilt: $f^{-1}(V) \in \mathcal{U}(x)$ für alle $V \in \mathcal{U}(f(x))$.

c) $f^{-1}(V) \in \mathcal{T}_X$ für alle $V \in \mathcal{T}_Y$.

d) $f^{-1}(A)$ ist abgeschlossen für alle abgeschlossenen Teilmengen von Y .

Beweis:

a) und b) sind definitionsgemäß äquivalent; siehe Definition 1.26.

Zu a) $\implies$ c). Sei $V \in \mathcal{T}_Y$. Sei $x \in f^{-1}(V)$. Dann ist $f(x) \in V, V \in \mathcal{U}(f(x))$, und da f stetig ist in x , gibt es $U' \in \mathcal{U}(x)$ mit $f(x) \in f(U') \subset V$, also $U' \subset f^{-1}(V)$. Dazu gibt es $U \in \mathcal{T}_X$ mit $x \in U \subset U'$. Also ist $f^{-1}(V) \in \mathcal{U}(x)$. Mit Satz 1.5 folgt b).

Zu c) $\implies$ d). Sei $A \subset Y$ abgeschlossen, also $V := Y \setminus A$ offen. Also ist $f^{-1}(V) \in \mathcal{T}_X$, d. h. $f^{-1}(A) = X \setminus f^{-1}(V)$ ist abgeschlossen.

Zu d) $\implies$ a). Sei $x \in X$ und sei $V \in \mathcal{U}(f(x))$; o. E. V offen. Dann ist $A := Y \setminus V$ abgeschlossen und daher $U := X \setminus f^{-1}(A)$ offen; ferner $x \in U$ und $f(U) \subset V$. Dies zeigt, dass f stetig in x ist. ■

Bemerkung 1.28 Sei $(X, \mathcal{T})$ ein topologischer Raum. Offensichtlich ist eine Topologie $\mathcal{T}_1$ auf X genau dann feiner als $\mathcal{T}$, wenn die Identität $\text{id} : X \rightarrow X$ stetig ist, wobei im Definitionsbereich die Topologie $\mathcal{T}_1$ und im Wertebereich die Topologie $\mathcal{T}$ gewählt ist.

Die von $\mathcal{T}_X$ induzierte Topologie $\mathcal{T}_Y$ auf $Y \subset X$ ist die grösste Topologie, für die die Einbettung $j_Y : Y \ni y \mapsto y \in X$ stetig ist. □

Satz 1.29 Seien $(X, \mathcal{T}_X), (Y, \mathcal{T}_Y)$ Hausdorffräume und sei $f : X \rightarrow Y$. Es gilt:

a) Ist X kompakt und f stetig, dann ist $f(X)$ kompakt.

b) Ist X kompakt und ist f bijektiv und stetig, dann ist auch die Umkehrabbildung f^{-1} stetig.

Beweis:

Zu a). Ist $\cup_{j \in J} Y_j$ eine offene Überdeckung von $f(X)$, dann ist $\cup_{j \in J} f^{-1}(Y_j)$ eine offene Überdeckung von X , da f stetig ist; siehe Satz 1.27. Da X kompakt ist, gibt es eine endliche Überdeckung $\cup_{l=1}^m f^{-1}(Y_{j_l})$ von X . Daraus resultiert die endliche Überdeckung $\cup_{l=1}^m Y_{j_l}$ von $f(X)$.

Zu b). Sei $g := f^{-1}$. Sei $A \subset X$ abgeschlossen. Dann ist A kompakt (nach Satz 1.23 a)) und $f(A)$ kompakt nach a), also abgeschlossen nach Satz 1.23. Da $g^{-1}(A) = f(A)$ gilt, ist also $g^{-1}(A)$ abgeschlossen. Damit ist mit Satz 1.23 b) die Aussage aus b) gezeigt. ■

Sei $(X_i, \mathcal{T}_i)_{i \in I}$ eine (nicht-leere) Familie topologischer Räume. Wir bilden damit das Mengen-Produkt X in folgendem Sinne:

$$x \in X := \prod_{i \in I} X_i : \iff x = (x_i)_{i \in I} \text{ mit } x_i \in X_i, i \in I\}.$$

Dann sind die **kanonischen Projektionen**

$$\pi_j : X \ni x = (x_i)_{i \in I} \mapsto x_j \in X_j, j \in I,$$

gegeben. Damit gelingt es nun ganz naheliegend eine Topologie auf X anzugeben, die „dazu passt“:

$\mathcal{T} := \prod_{i \in I} \mathcal{T}_i$ sei die grösste Topologie auf X , so dass alle $\pi_j, j \in I$, stetig sind.

Das Paar $(\prod_{i \in I} X_i, \prod_{i \in I} \mathcal{T}_i)$ heißt das **topologische Produkt** der gegebenen Familie topologischer Räume.

Für zwei Faktoren $(X, \mathcal{T}_X), (Y, \mathcal{T}_Y)$ haben wir in Definition 1.6 schon eine Produkttopologie bereitgestellt: $(X \times Y, \mathcal{T}_{X \times Y})$. Wie passen diese Konstruktionen zusammen? Sie fallen zusammen, wie wir nun zeigen wollen. Sei $\mathcal{T}$ die Produkttopologie auf $X \times Y$, die wir eben eingeführt haben.

Sei $U \in \mathcal{T}_X, V \in \mathcal{T}_Y$. Dann sind wegen der Stetigkeit $\pi_X^{-1}(U) = U \times Y \in \mathcal{T}, \pi_Y^{-1}(V) = X \times V \in \mathcal{T}$, also auch $U \times V \in \mathcal{T}$, da $\mathcal{T}$ stabil gegenüber Durchschnittsbildung ist. Dies zeigt, dass $\mathcal{T}_{X \times Y} \subset \mathcal{T}$ gilt.

Die Projektionen π_X, π_Y sind stetig, wenn in $X \times Y$ die Topologie $\mathcal{T}_{X \times Y}$ gewählt ist. Dies sieht man so ein. Ist $U \in \mathcal{T}_X$, so ist $\pi_X^{-1}(U) = U \times Y \in \mathcal{T}_{X \times Y}$. Analog zeigt man dies für π_Y . Damit erhält man $\mathcal{T} \subset \mathcal{T}_{X \times Y}$.

Satz 1.30 Sei $(X_i, \mathcal{T}_i)_{i \in I}$ eine Familie topologischer Räume. Sind alle Räume X_i kompakt, so ist das Produkt $\prod_{i \in I} X_i$ in der Produkttopologie kompakt.

Beweis:

Den Beweis wollen wir hier nicht erbringen; siehe etwa [30].

Wenn wir einen Vektorraum X mit Skalarkörper $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ betrachten, auf dem eine Topologie $\mathcal{T}$ gegeben ist, so können wir $X \times X$ und $X \times \mathbb{K}$ gemäß Definition 1.6 auch als topologische Räume betrachten, da ja $\mathbb{R}$ und $\mathbb{C}$ in bekannter Weise Topologien tragen. Also können wir nach der Stetigkeit der Addition und skalaren Multiplikation als Abbildungen

$$\begin{aligned} + & : X \times X \ni (x, y) \longmapsto x + y \in X, \\ \cdot & : X \times \mathbb{K} \ni (x, \alpha) \longmapsto \alpha \cdot x \in X \end{aligned}$$

fragen. Sind diese Abbildungen stetig, so nennen wir $(X, \mathcal{T})$ einen **topologischen Vektorraum**.

1.5 Übungen

- 1.) Sei $(X, \mathcal{T})$ ein topologischer Raum und sei $A \subset X$. Zeige:
 - (a) $X \setminus \partial A = A^\circ \cup (X \setminus A)^\circ$.
 - (b) $\overline{A} = A \cup \partial A$.
 - (c) $A^\circ = A \setminus \partial A$.
- 2.) Sei $(X, \mathcal{T})$ ein topologischer Raum. Zeige für $A, B \subset X$.
 - (a) $A \subset B \implies A^\circ \subset B^\circ$.
 - (b) $\overline{A \cup B} = \overline{A} \cup \overline{B}$.
 - (c) $\partial(\partial A) \subset \partial(A)$.
- 3.) Seien $(X, \mathcal{T}_X), (Y, \mathcal{T}_Y)$ topologische Räume und sei $f : X \longrightarrow Y$. Zeige die Äquivalenz von
 - (a) f ist stetig.
 - (b) $\overline{f^{-1}(B)} \subset f^{-1}(\overline{B})$ für alle $B \subset Y$.
- 4.) Sind in $\mathbb{R}$ die folgenden Familien $\mathcal{U}_k$ offene Überdeckungen von $A := \{\frac{1}{n} | n \in \mathbb{N}\}$? Falls ja, haben sie eine endliche Teilüberdeckung?

$$\mathcal{U}_1 := \{K_x(2x) | x > 0, x \in \mathbb{R}\}, \mathcal{U}_2 := \{(\frac{1}{n+2}, \frac{1}{n}) | n \in \mathbb{N}\},$$

$$\mathcal{U}_3 := \{(-\infty, n] | n \in \mathbb{N}\}, \mathcal{U}_4 := \{K_1(x), x \in \mathbb{R}\}.$$

5.) Sei $\mathbb{R}^n$ versehen mit der üblichen Topologie. Sind $\mathbb{Z}^n, \mathbb{Q}^n, \mathbb{R}^n \setminus \mathbb{Q}^n$ offen bzw. abgeschlossen in $\mathbb{R}^n$?

6.) Sei $(X, \mathcal{T})$ ein topologischer sei $A \subset X$ kompakt und sei $x \in X \setminus A$. Zeige: Es gibt offene Mengen U, V mit

$$x \in U, A \subset V, U \cap V = \emptyset.$$

7.) Sei $M := \{x \in \mathbb{R} | x = \frac{m}{2^n}, m \in \mathbb{Z}, n \in \mathbb{N}\}$. Zeige:

(a) Weder M noch $\mathbb{R} \setminus M$ ist offen.

(b) Ist $O \subset \mathbb{R}$ offen und $O \subset M$, dann gilt $O = \emptyset$.

(c) $M^\circ = \emptyset, \overline{M} = \mathbb{R}$.

8.) Sei $X := \{x \in \mathbb{R} | x = \frac{1}{m} + \frac{1}{n}, m, n \in \mathbb{N}\}$. Zeige:

(a) X ist nicht offen und nicht abgeschlossen.

(b) Bestimme die Häufungspunkte von X .

9.) Sei $(X_i, \mathcal{T}_i)_{i \in I}$ eine Familie topologischer Räume und sei $(X, \mathcal{T})$ die Produkttopologie. Ist $A_i \subset X_i$ für jedes $i \in I$, so schreiben wir $A := \prod_{i \in I} A_i$ und sagen $x = \{x_i\} \in A$ genau dann, wenn $x_i \in A_i$ gilt für jedes $i \in I$. Zeige: Eine Menge $A := \prod_{i \in I} A_i$ ist offen, wenn jedes A_i offen in $\mathcal{T}_i$ ist und $A_i = X_i$ gilt bis auf endlich viele Ausnahmen.

10.) Sei $(X_i, \mathcal{T}_i)_{i \in I}$ eine Familie topologischer Räume und sei $(X, \mathcal{T})$ die Produkttopologie. Zeige: Jede Projektion $\pi_k : X = \prod_{i \in I} X_i \ni (x_i)_{i \in I} \mapsto x_k \in X_k$ ist eine offene Abbildung, d. h. das Bild einer offenen Menge ist offen.

11.) Sei $(X_i, \mathcal{T}_i)_{i \in I}$ eine Familie topologischer Räume und sei $(X, \mathcal{T})$ die Produkttopologie. Sei $A_i \subset X_i, i \in I$. Zeige: $\prod_{i \in I} \overline{A_i} = \overline{\prod_{i \in I} A_i}$.

12.) Sei $(X_i, \mathcal{T}_i)_{i \in I}$ eine Familie topologischer Räume und sei $(X, \mathcal{T})$ die Produkttopologie. Sei $x := (x_i)_{i \in I} \in X = \prod_{i \in I} X_i$. Zeige: $U \in \mathcal{U}(x)$ genau dann, wenn es für jedes $j \in I$ ein $O_j \in \mathcal{T}_j$ gibt mit $O_j \neq X_j$ nur für endlich viele j und $x \in \prod_{j \in I} O_j \subset U$.

13.) Seien $(X, \mathcal{T}_X), (Y, \mathcal{T}_Y)$ topologische Räume und sei $f : X \rightarrow Y$. Zeige: $G_f := \{(x, f(x)) \in X \times Y | x \in X\}$ (Graph von f) ist abgeschlossen in der Produkttopologie auf $X \times Y$, falls f stetig ist.

Kapitel 2

Metrische Räume

Manche Menschen haben einen Gesichtskreis vom Radius Null und nennen ihn ihren Standpunkt

D. Hilbert

Wir lernen metrische Räume kennen, die die Grundlage für viele Betrachtungen im Folgenden bilden. Es wird manches aus der Analysis schon Bekanntes in neuem Lichte erscheinen lassen. Insbesondere wird der Begriff „vollständig“ auf eine abstrakte Ebene gehoben. Die Nützlichkeit dieses Vorgehens erweist sich hilfreich beim Kontraktionssatz und beim Baireschen Kategoriensatz. Der Kompaktheitsbegriff wird der speziellen Situation angepasst. Die Approximationssätze von Stone und Weierstraß schließen das Kapitel ab. Insgesamt sind nach diesem Kapitel mehr oder minder alle Begriffe, die für die Funktionalanalysis wichtig sind, diskutiert, was in den folgenden Kapiteln hinzukommt, ist die lineare Struktur der metrischen Räume.

2.1 Metrische Räume

Definition 2.1 Sei X eine nichtleere Menge. Eine Abbildung $d : X \times X \rightarrow \mathbb{R}$ heißt **Metrik**, wenn folgende drei Bedingungen erfüllt sind:

- a) $d(x, y) = 0$ genau dann, wenn $x = y$ gilt; (Definitheit)
- b) $d(x, y) = d(y, x)$ für alle $x, y \in X$; (Symmetrie)
- c) $d(x, y) \leq d(x, z) + d(z, y)$ für alle $x, y, z \in X$. (Dreiecksungleichung)

Die Zahl $d(x, y)$ heißt **Abstand** von x, y .

Eine Menge X zusammen mit einer Metrik d wird als **metrischer Raum** (X, d) bezeichnet. $\square$

Entsprechend der Bezeichnungsweise, dass $d(x, y)$ für den Abstand von x, y steht, sollte man erwarten, dass $d(x, y)$ stets nichtnegativ ist. Dies ist aber schon in der Definition 2.1 enthalten, wie folgende Zeile zeigt:

$$0 = d(x, x) \leq d(x, y) + d(y, x) = 2d(x, y) \text{ d. h. } d(x, y) \geq 0.$$

Beispiel 2.2 Jede nichtleere Teilmenge A von $\mathbb{R}$ ist zusammen mit dem Abstand $d(x, y) := |x - y|$ ein metrischer Raum. Die definierenden Eigenschaften sind ohne Mühe nachzuweisen. $\square$

Beispiel 2.3 Sei X eine nichtleere Menge. Dann wird durch

$$d : X \times X \ni (x, y) \mapsto \begin{cases} 1 & , \text{ falls } x \neq y \\ 0 & , \text{ falls } x = y \end{cases} \in \mathbb{R}$$

eine Metrik definiert. Nachrechnen! Interpretiert man den Abstand $d(x, y)$ als Kosten für eine Straßenbahnfahrt in dem Haltestellennetz X , so bedeutet die Wahl der Metrik gerade, dass für jede Fahrt der Einheitspreis 1 (in der gewählten Währung) zu entrichten ist. $\square$

Beispiel 2.4 Der Raum $\mathbb{C}^n$ ist uns als Vektorraum (über dem Skalarkörper $\mathbb{C}$) wohlvertraut. Wir können $\mathbb{C}^n$ in vielfältiger Weise zu einem metrischen Raum machen; hier sind drei Möglichkeiten. Definiere

$$d_\infty(x, y) := \max_{1 \leq i \leq n} |x_i - y_i|, \quad x = (x_1, \dots, x_n), y = (y_1, \dots, y_n) \in \mathbb{C}^n.$$

d_∞ ist eine Metrik, denn: a), b) in Definition 2.1 sind offensichtlich, zu c) wähle man j mit $d_\infty(x, y) = |x_j - y_j|$ und betrachte die Zeile

$$d_\infty(x, y) = |x_j - y_j| \leq |x_j - z_j| + |z_j - y_j| \leq \max_{1 \leq i \leq n} |x_i - z_i| + \max_{1 \leq i \leq n} |z_i - y_i|, \quad 1 \leq j \leq n.$$

Definiere

$$d_1(x, y) := \sum_{i=1}^n |x_i - y_i|, \quad x = (x_1, \dots, x_n), y = (y_1, \dots, y_n) \in \mathbb{C}^n.$$

d_1 ist eine Metrik, denn: a), b) in Definition 2.1 sind offensichtlich, c) liest man aus der Zeile

$$\sum_{i=1}^n |x_i - y_i| \leq \sum_{i=1}^n |x_i - z_i| + \sum_{i=1}^n |z_i - y_i|$$

ab.

Definiere

$$d_2(x, y) := \left(\sum_{i=1}^n |x_i - y_i|^2 \right)^{\frac{1}{2}}, \quad x = (x_1, \dots, x_n), y = (y_1, \dots, y_n) \in \mathbb{C}^n.$$

Auf den Nachweis, dass eine Metrik vorliegt, verzichten wir hier, wir gehen darauf in Abschnitt 3.1 ein. $\square$

Bemerkung 2.5 Die Metrik d_1 heißt Manhattan-Metrik, Cityblock-Metrik oder Taxidriver-Metrik. An der Karte von Manhattan kann nämlich illustriert werden, dass die Luftlinie in den meisten Fällen nicht geeignet ist, um physisch wirklich überwindbare Distanzen anzugeben. Die Manhattan-Metrik folgt derselben Logik wie die Fahrt mit einem Taxi in Manhattan: Das Taxi fährt zum Beispiel zwei Blocks nach Norden und anschließend drei Blocks nach Osten. Es sind also nur Fahrten entlang der vier Haupthimmelsrichtungen möglich, diagonale Fortbewegung durch Häuser ist nicht erlaubt.

Man beachte, dass dieser Abstand nicht invariant gegenüber Drehungen des Koordinatensystems ist. Bei der Metrik d_2 werden wir diese Invarianz später vorführen. $\square$

Aus der Dreiecksungleichung einer Metrik d auf X leitet sich die so genannte **Vierecksungleichung** ab:

$$|d(x, y) - d(u, v)| \leq d(x, u) + d(y, v), \quad x, y, u, v \in X. \quad (2.1)$$

Der Beweis dieser Ungleichung liegt auf der Hand.

Bezeichnungen: Sei (X, d) ein metrischer Raum. Wir definieren **Kugeln**:

$$B_r(x) := \{y \in X | d(x, y) < r\}, \quad \bar{B}_r(x) := \{y \in X | d(x, y) \leq r\} \quad (x \in X, r \geq 0).$$

Damit haben wir unter Verwendung der Dreiecksungleichung für alle $x \in X, r > 0$:

$$x \in \overline{B_{\frac{r}{2}}}(x) \subset B_r(x) \quad (2.2)$$

Ein metrischer (X, d) Raum kann in naheliegender Weise zu einem topologischen Raum gemacht werden. Dazu betrachten wir die d -offenen Mengen: $A \subset X$ nennen wir **d -offen**, wenn gilt:

$$\forall x \in A \exists r > 0 (B_r(x) \subset A)$$

Damit setzen wir $\mathcal{T}_d := \{A \subset X | A \text{ } d\text{-offen}\}$. Nun ist es wirklich einfach nachzurechnen, dass $\mathcal{T}_d$ eine Topologie auf X ist.

Definition 2.6 Sei (X, d) ein metrischer Raum. Die Familie

$$\mathcal{T} := \mathcal{T}_d := \{A \subset X | A \text{ } d\text{-offen}\}$$

heißt die durch d **erzeugte Topologie**. □

Alle Kugeln $B_r(x)$ in einem metrischen Raum sind offen. Die Kugeln $\overline{B_r}(x)$ sind abgeschlossen, wie folgende Überlegung zeigt. Sei $A := X \setminus \overline{B_r}(x)$. Wir zeigen A ist offen. Sei also $y \in A$. Nun ist $\rho := d(x, y) > r$ und wir wählen $0 < s < \rho - r$. Dann ist $B_s(y) \subset A$, wie folgende Zeile für $z \in B_s(y)$ zeigt:

$$\rho = d(y, x) \leq d(y, z) + d(z, x) < s + d(z, x) < \rho - r + d(z, x)$$

Damit ist nun klar, dass $\overline{\overline{B_r}(x)} = \overline{B_r}(x)$ gilt und wir stellen fest, dass die Bezeichnung $\overline{B_r}(x)$ passend gewählt ist.

Zu jeder Umgebung $U \in \mathcal{U}(x)$ können wir eine Kugel $B_r(x)$ finden mit $x \in B_r(x) \subset U$. O. E. kann man den Radius r als $1/n$ mit $n \in \mathbb{N}$ wählen. Diese Beobachtung zeigt uns, dass jeder metrische Raum das 1. Abzählbarkeitsaxiom erfüllt: eine abzählbare Umgebungsbasis ist für x : $(B_{1/n}(x))_{n \in \mathbb{N}}$. Halten wir dies noch fest:

Lemma 2.7 In jedem metrischen Raum gilt das 1. Abzählbarkeitsaxiom.

Definition 2.8 Sei X eine Menge und seien $d, \bar{d}$ Metriken auf X . Diese Metriken heißen **äquivalent**, wenn sie dieselben Topologien erzeugen, d. h. wenn $\mathcal{T}_d = \mathcal{T}_{\bar{d}}$ gilt. □

Klar, Äquivalenz von Metriken ist eine Äquivalenzrelation. Etwa sind alle Metriken d_1, d_2, d_∞ , die wir in Beispiel 2.4 kennen gelernt haben, äquivalent. Wir gehen später darauf ein.

Ist A eine Teilmenge eines metrischen Raumes X , so heißt

$$U_\epsilon(A) := \{y \in X | y \in B_\epsilon(x) \text{ für ein } x \in A\}$$

eine ϵ -**Umgebung** von A . Diese Begriffsbildung stimmt überein mit der Definition der Umgebung eines Punktes, wie wir sie in topologischen Räumen gegeben haben.

Folgerung 2.9 Jeder metrische Raum (X, d) ist in der durch d induzierten Topologie $\mathcal{T}_d$ **hausdorffsch**.

Beweis:

Seien $x, y \in X, x \neq y$. Dann ist $t := d(x, y) > 0$ und mit $r := t/2$ folgt $B_r(x) \cap B_r(y) = \emptyset$ auf Grund der Dreiecksungleichung. ■

Der Abstandsbegriff, der durch eine Metrik eingeführt wird, macht es möglich, die wesentlichen Begriffe der Analysis im Zusammenhang mit Folgen sofort auf metrische Räume zu übertragen. Dies soll nun geschehen.

Definition 2.10 Sei (X, d) ein metrischer Raum und sei $(x_n)_{n \in \mathbb{N}}$ eine Folge in X .

a) Die Folge $(x_n)_{n \in \mathbb{N}}$ heißt **Cauchyfolge**, wenn gilt:

$$\forall \epsilon > 0 \exists N \in \mathbb{N} \forall m, n \geq N (d(x_n, x_m) < \epsilon).$$

a) Die Folge $(x_n)_{n \in \mathbb{N}}$ heißt **konvergent**, wenn es $x \in X$ gibt mit

$$\forall \epsilon > 0 \exists N \in \mathbb{N} \forall n \geq N (d(x_n, x) < \epsilon);$$

x heißt dann **Grenzwert** oder **Limes** der Folge und wir schreiben $x = \lim_n x_n$.

□

Die Tatsache, dass die Topologie, die durch die Metrik eines metrischen Raumes induziert wird, hausdorffsch ist, hat als Konsequenz, dass der Grenzwert einer konvergenten Folge eindeutig bestimmt ist, was wir natürlich auch direkt sehen könnten. Ebenso einfach sieht man ein, dass eine konvergente Folge stets eine Cauchyfolge ist. Die Umkehrung davon gilt nicht immer, wie uns das Beispiel $\mathbb{Q}$ zusammen mit der Abstandsfunktion in $\mathbb{R}$ als Metrik zeigt. Wir kommen auf die damit zusammenhängenden Fragestellungen im nächsten Abschnitt zurück.

Satz 2.11 Sei (X, d) ein metrischer Raum, $A \subset X$. Es sind äquivalent:

a) A ist abgeschlossen.

b) Für jede konvergente Folge $(x_n)_{n \in \mathbb{N}}$ mit $x_n \in A, n \in \mathbb{N}$, gilt: $\lim_n x_n \in A$.

c) $\forall x \in A \forall r > 0 (B_r(x) \cap A \neq \emptyset)$.

Ist A abgeschlossen, dann gilt

$$\overline{A} = \{x \in X | \exists (x_n)_{n \in \mathbb{N}} \text{ mit } x_n \in A \text{ für alle } n \in \mathbb{N} \text{ und } \lim_n x_n = x\}.$$

Beweis:

Eigentlich könnten wir uns beim Beweis mehr oder minder auf die Tatsache berufen, dass das 1. Abzählbarkeitsaxiom gilt. Wir ziehen es aber vor, unabhängig davon den Beweis zu formulieren. Zu a) $\implies$ b). Sei A abgeschlossen. Dann ist $X \setminus A$ offen. Sei $(x_n)_{n \in \mathbb{N}}$ eine konvergente Folge mit $x_n \in A$ für alle $n \in \mathbb{N}$ und sei $x := \lim_n x_n$.

Annahme: $x \in X \setminus A$. Da $X \setminus A$ offen ist, gibt es $\epsilon > 0$ mit $B_\epsilon(x) \subset X \setminus A$. Dazu gibt es $N \in \mathbb{N}$ mit $x_n \in B_\epsilon(x)$ für alle $n \geq N$. Dies ist im Widerspruch zu $x_n \in A$ für alle $n \in \mathbb{N}$.

Zu b) $\implies$ a). Annahme: A sei nicht abgeschlossen, d. h. $X \setminus A$ nicht offen. Dann gibt es $x \in X \setminus A$ derart, dass zu jedem $n \in \mathbb{N}$ ein $x_n \in A$ existiert mit $d(x, x_n) < \frac{1}{n}$. Dann ist aber $(x_n)_{n \in \mathbb{N}}$ eine Folge in A mit $\lim_n x_n = x$. Dies ist ein Widerspruch.

Zu a) $\iff$ c). Ist unmittelbar klar.

Die Darstellung von $\overline{A}$ ergibt sich aus der Äquivalenz a) $\iff$ b). ■

2.2 Stetigkeit

Wir schreiben „Stetigkeit“ neu auf, wobei wir nur die Definition, die wir in Abschnitt 1.1 gegeben haben, an die Situation anpassen. Man vergleiche mit der Definition von Stetigkeit in $\mathbb{R}$.

Definition 2.12 Seien $(X, d_X), (Y, d_Y)$ zwei metrische Räume und sei $f : X \rightarrow Y$ eine Abbildung. f heißt **stetig in** $x \in X$ genau dann, wenn gilt:

$$\forall \epsilon > 0 \exists \delta > 0 \forall y \in X (d_X(x, y) < \delta \implies d_Y(f(x), f(y)) < \epsilon).$$

Ist die Abbildung f stetig in jedem Punkt $x \in X$, so heißt f **stetig**. □

Da ein metrischer Raum stets das 1. Abzählbarkeitsaxiom erfüllt, kann Stetigkeit einer Abbildung $f : X \rightarrow Y$ in einem Punkt z stets auch mit Folgen überprüft werden: f ist stetig in z genau dann, wenn aus $z = \lim_n x_n$ in X folgt $f(z) = \lim_n f(x_n)$ in Y (**Folgenstetigkeit**). Man sieht dies wie in der Analysis von Funktionen auf $\mathbb{R}$ ein.

Definition 2.13 Seien $(X, d_X), (Y, d_Y)$ metrische Räume. Eine Abbildung $f : X \rightarrow Y$ heißt **gleichmäßig stetig**, wenn gilt:

$$\forall \epsilon > 0 \exists \delta > 0 \forall x, x' \in X (d_X(x, x') < \delta \implies d_Y(f(x), f(x')) < \epsilon).$$

□

Beispiel 2.14 Sei (X, d) metrischer Raum und sei $x^0 \in X$. Dann ist die Abbildung

$$f : X \ni x \rightarrow d(x^0, x) \in \mathbb{R}$$

gleichmäßig stetig, insbesondere stetig, wie man an

$$|d(x^0, x) - d(x^0, y)| \leq d(x, y), \quad x, y \in X,$$

abliest. Dabei wird in $\mathbb{R}$ die übliche euklidische Metrik, d. h. die Betragsfunktion, verwendet. □

Bemerkung 2.15 Während die Stetigkeit einer Abbildung $f : X \rightarrow Y$ ($(X, d_X), (Y, d_Y)$ metrische Räume) nur von den jeweiligen Topologien in X, Y abhängt, sind für die gleichmäßige Stetigkeit die Metriken entscheidend, d. h. wir brauchen sie zur Formulierung. Auf der Ebene der topologischen Räume kann gleichmäßige Stetigkeit in **uniformen Räumen** formuliert werden. □

Beispiel 2.16 Seien $(X, d_X), (Y, d_Y)$ metrische Räume und sei $T : X \rightarrow Y$ eine Abbildung mit

$$d_Y(T(x), T(x')) \leq l d_X(x, x') \quad \text{für alle } x, x' \in X,$$

wobei $l \in [0, \infty)$ ist. Dann ist T offenbar stetig, ja sogar gleichmäßig stetig. Man nennt solche Abbildungen **Lipschitzstetig** und die Zahl l eine **Lipschitzkonstante**. Wir werden uns mit solchen Abbildungen noch mehrmals beschäftigen und ihre große Bedeutung kennen lernen. □

Sei D eine nichtleere Teilmenge von $\mathbb{C}$ und seien $f_n : D \rightarrow \mathbb{C}$, $n \in \mathbb{N}$, gegebene Funktionen. Wann dürfen wir die Konvergenz von $(f_n(x))_{n \in \mathbb{N}}$ gegen $f(z)$ für alle $z \in D$ mit einer Funktion (Grenzfunktion) $f : D \rightarrow \mathbb{C}$ erwarten? Beispiele für diese Situation sind die **Monome**

$$m_n : \mathbb{C} \ni z \mapsto z^n \in \mathbb{C}, \quad n \in \mathbb{N}_0.$$

Beispiel 2.17 Für die Folge der Monome $(m_n)_{n \in \mathbb{N}}$, betrachtet auf $D := [-1, 1]$, haben wir offenbar

$$\lim_n m_n(x) = \begin{cases} 0 & , \text{ falls } |x| < 1 \\ 1 & , \text{ falls } x = 1 \end{cases} ,$$

und $\lim_n m_n(x)$ existiert nicht für $x = -1$. Selbst, wenn wir die Folge etwa nur auf $D := [0, 1]$ betrachten, sehen wir, dass die „Grenzfunktion“

$$f : [0, 1] \ni x \mapsto \lim_n m_n(x) \in \mathbb{R}$$

nicht stetig ist, obwohl alle Monome stetig sind. □

Die zu dieser Frage passende Begriffsbildung ist die der gleichmäßigen Konvergenz. Wir formulieren sie im Kontext metrischer Räume.

Definition 2.18 Sei (Y, d) ein metrischer Raum und sei D eine Menge. Eine Funktionenfolge $(f_n)_{n \in \mathbb{N}}$ mit gemeinsamem Definitionsbereich D **konvergiert gleichmäßig** gegen die Funktion $f : D \rightarrow Y$, wenn gilt:

$$\forall \epsilon > 0 \exists N \in \mathbb{N} \forall n \geq N \forall z \in D (d(f_n(z), f(z)) < \epsilon)$$

Wir nennen f dann den **(gleichmäßige) Grenzwert der Funktionenfolge** $(f_n)_{n \in \mathbb{N}}$. □

Klar, der Grenzwert f einer Funktionenfolge $(f_n)_{n \in \mathbb{N}}$ ist eindeutig bestimmt, da der Grenzwert von Folgen in metrischen Räumen eindeutig bestimmt ist. Die gleichmäßige Konvergenz ist offenbar eine Verschärfung der punktweisen Konvergenz $\lim_n f_n(z) = f(z)$ für alle $z \in D$.

Satz 2.19 Seien $(X, d_X), (Y, d_Y)$ metrische Räume, sei $D \subset X$, und seien $f_n : D \rightarrow Y, n \in \mathbb{N}$, stetig in $z \in D$. Ist die Folge $(f_n)_{n \in \mathbb{N}}$ gleichmäßig konvergent gegen f , dann ist f stetig in z .

Beweis:

Sei $\epsilon > 0$. Dazu gibt es $N \in \mathbb{N}$ mit

$$d_Y(f_n(x), f(x)) < \frac{\epsilon}{3}, x \in D, n \geq N.$$

Da f_N stetig in z ist, gibt es $\delta > 0$ mit

$$d_Y(f_N(x), f_N(z)) < \frac{\epsilon}{3}, x \in D \cap B_\delta(z).$$

Sei nun $x \in D \cap B_\delta(z)$. Dann haben wir

$$d_Y(f(x), f(z)) \leq d_Y(f(x), f_N(x)) + d_Y(f_N(x), f_N(z)) + d_Y(f_N(z), f(z)) < \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon.$$

Damit ist die Stetigkeit von f in z gezeigt. ■

Den Inhalt von Satz 2.19 kann man sich auch als Vertauschungsregel merken: Ist $(f_n)_{n \in \mathbb{N}}$ eine gleichmäßig konvergente Folge von stetigen Funktionen mit Grenzwert f , so gilt für eine Folge $(z_k)_{k \in \mathbb{N}}$ mit $\lim_k z_k = z$:

$$\lim_k \lim_n f_n(z_k) = \lim_k f(z_k) = f(z) = \lim_n f_n(z) = \lim_n \lim_k f_n(z_k).$$

2.3 Vollständigkeit und der Kontraktionssatz

Definition 2.20 Ein metrischer Raum (X, d) heißt **vollständig**, wenn jede Cauchyfolge in X konvergiert. $\square$

Wir verzichten hier auf Beispiele vollständiger metrischer Räume, da wir im nächsten Abschnitt in einer spezielleren Situation Beispiele diskutieren werden. Von den metrischen Räumen $\mathbb{R}, \mathbb{R}^d, \mathbb{C}, \mathbb{C}^d$ wissen wir das aber schon.

Zu jedem metrischen Raum (X, d) kann ein vollständiger metrischer Raum $(\hat{X}, \hat{d})$ so konstruiert werden, dass gilt:

$$X \subset \hat{X}, j : X \ni x \mapsto x \in \hat{X} \text{ Isometrie, } j(X) \text{ dicht in } \hat{X}. \quad (2.3)$$

Dabei ist eine Isometrie eine Abbildung, die abstandserhaltend ist. Ein solcher Raum $(\hat{X}, \hat{d})$ heisst **vollständige Hülle** von X . Wir wissen auch aus der Analysis im wesentlichen schon, wie eine solche Hülle zu konstruieren ist: Man fasse die Cauchyfolgen $(x_n)_{n \in \mathbb{N}}$ in X zu neuen Objekten – jedes $x \in X$ kann als konstante Folge in X als eine Cauchyfolge aufgefasst werden – zu X' zusammen, erkläre eine Äquivalenzrelation auf X' durch

$$(x_n)_{n \in \mathbb{N}} \sim (x'_n)_{n \in \mathbb{N}} : \iff \lim_n (x_n - x'_n) = 0,$$

definiere $\hat{X}$ als die Menge der Äquivalenzklassen $[(x_n)_{n \in \mathbb{N}}]$ bezüglich dieser Relation und erkläre die Metrik in $\hat{X}$ auf folgende Weise:

$$\hat{d}([(x_n)_{n \in \mathbb{N}}], [(x'_n)_{n \in \mathbb{N}}]) := \lim_n d(x_n, x'_n).$$

Nun hat man zu verifizieren:

1. $\lim_n d(x_n, x'_n)$ existiert.
2. $\hat{d}([(x_n)_{n \in \mathbb{N}}], [(x'_n)_{n \in \mathbb{N}}])$ ist unabhängig von den Repräsentanten der Äquivalenzklassen.
3. $\hat{d}$ ist eine Metrik.
4. Die oben angeführten Eigenschaften sind erfüllt.

Dies gelingt auf vorgezeichnetem Wege. Wir verfolgen dies nicht weiter, siehe etwa [37].

Definition 2.21 Sei (X, d) ein metrischer Raum. Eine Teilmenge A von X heißt **beschränkt**, wenn der **Durchmesser**

$$\text{diam}(A) := \sup\{d(y, z) | y, z \in A\}$$

von A endlich ist. $\square$

Man sieht sehr schnell, dass die Beschränktheit von A zur Tatsache äquivalent ist, dass es ein $r > 0$ und ein $y \in Y$ gibt mit $A \subset B_r(y)$.

Satz 2.22 (Schachtelungssatz) Sei (X, d) ein vollständiger metrischer Raum. Ist $(A_n)_{n \in \mathbb{N}}$ eine Folge abgeschlossener Teilmengen von X mit

$$A_n \neq \emptyset, A_{n+1} \subset A_n, n \in \mathbb{N}, \lim_n \text{diam}(A_n) = 0,$$

dann gibt es ein $x \in X$ mit

$$\bigcap_{n=1}^{\infty} A_n = \{x\}.$$

Beweis:

Wähle $x_n \in A_n, n \in \mathbb{N}$. Für alle $m, n, N \in \mathbb{N}$ mit $m, n \geq N$ gilt $d(x_m, x_n) \leq \text{diam}(A_N)$. Also ist $(x_n)_{n \in \mathbb{N}}$ eine Cauchyfolge, und daher konvergent. Für $x := \lim_n x_n$ gilt dann $x \in A_n$ für alle n , da $x_m \in A_n$ für alle $m \geq n$ und A_n abgeschlossen ist. Also gilt $x \in \bigcap_{n \in \mathbb{N}} A_n$. Wegen $\lim_n \text{diam}(A_n) = 0$ kann $\bigcap_{n \in \mathbb{N}} A_n$ höchstens ein Element enthalten. ■

In Anwendungen werden Iterationen benutzt, um Folgen zu konstruieren, die einen erwünschten Grenzwert besitzen. Meist interessiert man sich für Punkte, die sich unter einer Abbildung nicht verändern, sogenannten Fixpunkten.

Definition 2.23 Seien (X, d) ein metrischer Raum und sei $T : X \rightarrow X$ eine Abbildung. Ein Punkt $\bar{x} \in X$ heisst **Fixpunkt** von T , wenn $T(\bar{x}) = \bar{x}$ gilt. □

Hier kommt ein Satz, der über die Existenz von Fixpunkten einer Abbildung $T : X \rightarrow X$ und ihrer Berechnung Aufschluss gibt. Er hat ausgeprägte konstruktive Aspekte und er kann bestens zur Berechnung eines Fixpunktes herangezogen werden. Im nächsten Kapitel werden wir im Kontext von normierten Räumen den Satz wieder formulieren und die Anwendung bei gewöhnlichen Differentialgleichungen betrachten. Zentral ist die **Fixpunktiteration**:

$$\text{Gegeben Startwert } x \in X ; x_0 := x, x_{n+1} := T(x_n), n \in \mathbb{N}. \quad (2.4)$$

Satz 2.24 (Kontraktionssatz) Sei (X, d) ein vollständiger metrischer Raum und sei $T : X \rightarrow X$ eine **Kontraktion**, d. h.

$$\exists L \in [0, 1) \forall x, x' \in X (d(T(x), T(x')) \leq Ld(x, x')). \quad (2.5)$$

Dann gilt:

- a) T besitzt einen eindeutig bestimmten Fixpunkt $\bar{x}$.
- b) Für alle Startwerte $x = x_0 \in X$ konvergiert die gemäß (2.4) erzeugte Folge $(x_n)_{n \in \mathbb{N}}$ gegen $\bar{x}$ und wir haben
- c) $d(x_n, \bar{x}) \leq \frac{L^n}{1-L} d(T(x), x), n \in \mathbb{N}$,
- d) $d(x_{n+1}, \bar{x}) \leq \frac{L}{1-L} d(x_{n+1}, x_n), n \in \mathbb{N}$.

Beweis:

Wir zeigen, dass T höchstens einen Fixpunkt besitzt. Seien $\bar{x}, \hat{x} \in X$ Fixpunkte von T . Dann folgt aus der Ungleichung

$$d(\bar{x}, \hat{x}) = d(T(\bar{x}), T(\hat{x})) \leq Ld(\bar{x}, \hat{x})$$

mit der Tatsache $0 \leq L < 1$ sofort $\bar{x} = \hat{x}$.

Sei $x = x_0 \in X$. Die Iterationsfolge $(x_n)_{n \in \mathbb{N}}$ ist gemäß (2.4) wohl definiert. Es gilt:

$$d(x_{k+1}, x_k) \leq L^k d(T(x), x) = L^k d(x_1, x_0), k \in \mathbb{N}; \quad (2.6)$$

$$d(x_n, x_m) \leq \frac{L^n}{1-L} d(x_1, x_0), n, m \in \mathbb{N}, m > n. \quad (2.7)$$

Die Aussage in (2.6) beweist man mit Hilfe der Voraussetzung (2.5) durch Induktion über k . (2.7) folgt mit (2.6) so:

$$\begin{aligned} d(x_n, x_m) &\leq d(x_n, x_{n+1}) + d(x_{n+1}, x_{n+2}) + \cdots + d(x_{m-1}, x_m) \\ &\leq (L^n + L^{n+1} + \cdots + L^{m-1}) d(x_1, x_0) \\ &= L^n (1 + L + \cdots + L^{m-1-n}) d(x_1, x_0) \leq \frac{L^n}{1-L} d(x_1, x_0) \end{aligned}$$

Aus (2.7) folgt, dass in $(x_m)_{m \in \mathbb{N}}$ eine Cauchyfolge vorliegt. Nach Voraussetzung gibt es $\bar{x} \in X$ mit $\bar{x} = \lim_{m \in \mathbb{N}} x_m$. Da T stetig ist, siehe Beispiel 2.16, folgt $T(\bar{x}) = \bar{x}$. Damit ist die Existenz eines Fixpunktes gezeigt. Da der Fixpunkt eindeutig bestimmt ist, konvergiert jede gemäß (2.4) konstruierte Folge unabhängig vom gewählten Startwert $x = x_0$ gegen $\bar{x}$. Damit sind a), b) gezeigt.

Aus (2.7) folgt durch Grenzübergang $m \rightarrow \infty$ die Aussage c). d) folgt aus der Zeile

$$d(x_{n+1}, \bar{x}) \leq Ld(x_n, \bar{x}) \leq Ld(x_n, x_{n+1}) + Ld(x_{n+1}, \bar{x}).$$

■

Hier sind drei Anmerkungen, die die Qualität des Kontraktionssatzes hinsichtlich der Berechenbarkeit des Fixpunktes aufzeigen. Die Aussage b) zeigt, dass die Fixpunktiteration (2.4) beliebig gestartet werden kann. c) ist eine **a priori-Abschätzung**: ohne die Iterationsfolge berechnet zu haben, kann nach Kenntnis von $x_1 := Tx_0$ schon die Anzahl der Iterationsschritte n abgeschätzt werden (mit Hilfe der rechten Seite), die benötigt werden, um eine vorgegebene Genauigkeit von x_n , gemessen in $d(x_n, \bar{x})$ zu erreichen. d) ist eine **a posteriori-Abschätzung**: während der Berechnung der Iterationsfolge kann (mit Hilfe der rechten Seite) entschieden werden, ob x_n schon die gewünschte Genauigkeit, gemessen in $d(x_{n+1}, \bar{x})$, besitzt.

2.4 Kompaktheit

Wir kommen zur Kompaktheit in metrischen Räumen. Die Tatsache, dass in metrischen Räumen stets das 1. Abzählbarkeitsaxiom gilt, hat Konsequenzen für die Beschreibung von Kompaktheit.

Definition 2.25 Sei (X, d) ein metrischer Raum.

- a) (X, d) heißt **folgenkompakt**, wenn jede Folge $(x_n)_{n \in \mathbb{N}}$ in X eine konvergente Teilfolge enthält.
- b) (X, d) heißt **präkompakt**, wenn gilt:

$$\forall r > 0 \exists m \in \mathbb{N} \exists x_1, \dots, x_m \in X \left(X \subset \bigcup_{j=1}^m B_r(x_j) \right)$$

□

Lemma 2.26 Sei (X, d) ein metrischer Raum. Betrachte die folgenden Eigenschaften:

- a) X ist präkompakt.
- b) X enthält eine abzählbare dichte Teilmenge.
- c) Jede offene Überdeckung von X enthält eine abzählbare Überdeckung.

Dann gilt die Implikationskette a) $\implies$ b) $\implies$ c).

Beweis:

Zu a) $\implies$ b). Für alle $n \in \mathbb{N}$ gibt es $x_1^n, \dots, x_{m_n}^n \in X$ mit

$$X = \bigcup_{j=1}^{m_n} B_{\frac{1}{n}}(x_j^n) \text{ für alle } n \in \mathbb{N}.$$

Offenbar ist nun $M := \{x_j^n | 1 \leq j \leq m_n, n \in \mathbb{N}\}$ abzählbar und dicht.

Zu b) $\implies$ c). Sei $M = \{x_j | j \in \mathbb{N}\}$ eine dichte Teilmenge von X . Sei $U \subset X$ offen. Da

$(B_{1/n}(x))_{n \in \mathbb{N}}$ eine abzählbare Umgebungsbasis ist für alle $x \in X$, gibt es zu $x \in U$ ein $n \in \mathbb{N}$ mit $B_{2/n}(x) \subset U$. Da M dicht in X ist, gibt es ein $j \in \mathbb{N}$ mit $x_j \in B_{1/n}(x)$. Daher gilt

$$x \in B_{1/n}(x_j) \subset B_{2/n}(x) \subset U.$$

Daraus folgt, dass die offene Menge U Vereinigung von abzählbar vielen Kugeln ist, nämlich von $(B_{1/n}(x_j))_{n \in \mathbb{N}, j \in \mathbb{N}}$. Hat man nun eine offene Überdeckung von $\cup_{i \in \mathbb{N}} U_i$ von X , so werden zur Überdeckung aller U_i nur die abzählbar vielen Kugeln $B_{1/n}(x_j), n \in \mathbb{N}, j \in \mathbb{N}$, benötigt. ■

Definition 2.27 Ein topologischer Raum heißt abzählbar kompakt, wenn jede offene abzählbare Überdeckung von X eine endliche Überdeckung enthält. □

Satz 2.28 Sei (X, d) ein metrischer Raum. Es sind äquivalent:

- a) X ist kompakt.
- b) X ist abzählbar kompakt.
- c) Ist $(A_n)_{n \in \mathbb{N}}$ eine Folge von abgeschlossenen Teilmengen von X mit $\emptyset \neq A_{n+1} \subset A_n$ für alle $n \in \mathbb{N}$. so gilt

$$\bigcap_{n \in \mathbb{N}} A_n \neq \emptyset.$$

- d) X ist folgenkompakt.
- e) X ist vollständig und präkompakt.

Beweis:

Zu a) $\implies$ b). Klar.

Zu b) $\implies$ c). Ist $\cap_{n \in \mathbb{N}} A_n \neq \emptyset$ nicht erfüllt, erhält man durch Komplementbildung eine offene Überdeckung von X und damit auch eine endliche Überdeckung von X , was nach Komplementbildung sofort einen Widerspruch zur Eigenschaft, dass endliche viele Schnitte der Mengen A_n nicht leer sind.

Zu c) $\implies$ d). Sei $(x_n)_{n \in \mathbb{N}}$ eine Folge. Setze $A_n := \overline{\{x_k | k \geq n\}}$, $n \in \mathbb{N}$. Man folgert mit c), dass es $x \in \cap_{n \in \mathbb{N}} A_n$ gibt. Ein solches x ist Berührungspunkt jeder Menge $\{x_k | k \geq n\}$. Also gibt es zu jedem $l \in \mathbb{N}$ ein x_{n_l} in $B_{1/l}(x) \cap \{x_k | k \geq n\}$. Die Teilfolge $(x_{n_l})_{l \in \mathbb{N}}$ konvergiert offenbar gegen x . Zu d) $\implies$ e). Ist $(x_n)_{n \in \mathbb{N}}$ eine Cauchyfolge, dann besitzt sie eine wegen d) eine konvergente Teilfolge mit einem Grenzwert $x \in X$. Dann konvergiert aber die ganze Folge gegen x . Also ist die Vollständigkeit gezeigt.

Annahme: X ist nicht präkompakt. Dann gilt:

$$\forall r > 0 \forall x_1, \dots, x_n \exists x_{n+1} \in X \forall i = 1, \dots, n (d(x_i, x_{n+1}) \geq r)$$

Klar, $(x_n)_{n \in \mathbb{N}}$ besitzt sicher keine konvergente Teilfolge. Widerspruch!

Zu e) $\implies$ a). Sei $X = \cup_{i \in I} U_i$ eine offene Überdeckung von X . Wegen Lemma 2.26 a) $\implies$ c) können wir $I = \mathbb{N}$ annehmen.

Annahme: $V_k := \cup_{n=1}^k U_n \subsetneq X$ für alle $k \in \mathbb{N}$. Dann ist $X \setminus V_k \neq \emptyset$ für alle $k \in \mathbb{N}$. Wähle für jedes $k \in \mathbb{N}$ ein $x_k \in X \setminus V_k$. Betrachte die so entstandene Folge $(x_k)_{k \in \mathbb{N}}$. Da X präkompakt ist, gibt es ein $\xi_1 \in X$, so dass $B_{2^{-1}}(\xi_1)$ unendlich viele Folgenglieder enthält. Induktiv erhält man so fortfahrend eine Folge $(B_{2^{-k}}(\xi_k))_{k \in \mathbb{N}}$ und eine Teilfolge $(x_{n_k})_{k \in \mathbb{N}}$, so dass $x_{n_k} \in \cap_{m=1}^k B_{2^{-m}}(\xi_m)$ für alle $k \in \mathbb{N}$. Dann gilt $d(x_{n_{k+1}}, x_{n_k}) < 2^{-k+1}$ für alle $k \in \mathbb{N}$. Hieraus folgt sofort, dass $(x_{n_k})_{k \in \mathbb{N}}$ eine Cauchyfolge ist. Diese ist nun konvergent, etwa gegen x . Dann ist

$$x \in \bigcap_{k \in \mathbb{N}} (X \setminus V_k) = X \setminus (\cup_{k \in \mathbb{N}} U_k) = \emptyset,$$

was ein Widerspruch ist. Also $X = \cup_{k=1}^m U_k$ für ein $m \in \mathbb{N}$. ■

Beispiel 2.29 Betrachten wir den metrischen Raum (X, d) , wobei $X := \mathbb{N}$ gewählt und die Metrik d durch den Betrag in $\mathbb{R}$ gegeben ist: $d(x, y) := |x - y|$. An diesem Beispiel kann man die Bedingungen a), ..., e) aus Satz 2.28 betrachten.

Zu a) X ist nicht kompakt, denn X ist keine beschränkte Teilmenge von $\mathbb{R}$.

Zu b) X ist nicht abzählbar kompakt, da etwa die Überdeckung durch die offenen Intervalle $(n - \frac{1}{2}, n + \frac{1}{2})$ keine endliche Überdeckung enthält.

Zu c) Die Mengenfamilie $A_n := \{k \in \mathbb{N} | k \geq n\}$ ist ein Gegenbeispiel.

Zu d) Die Folge $(n)_{n \in \mathbb{N}}$ enthält keine konvergente Teilfolge.

Zu e) X ist zwar vollständig (abgeschlossenen Teilmenge von $\mathbb{R}$), aber nicht präkompakt, da etwa mit dem Radius $r = \frac{1}{2}$ keine Überdeckung von X mit endlich vielen Kugeln möglich ist. $\square$

Der Satz 2.28 beschreibt nun in vollkommener Weise die Kompaktheit in metrischen Räumen. Nun kombinieren wir Kompaktheit (erneut) mit dem zweiten wichtigen Konzept in topologischen Betrachtungen, der Stetigkeit.

Satz 2.30 Seien $(X, d_X), (Y, d_Y)$ metrische Räume und sei $f : X \rightarrow Y$. Ist f stetig und ist X kompakt, dann ist f gleichmäßig stetig.

Beweis:

Sei $\epsilon > 0$. Da f stetig ist in jedem $a \in X$, gibt es zu jedem $a \in X$ ein $\delta(a) > 0$ mit $f(B_{2\delta(a)}) \subset B_{\frac{\epsilon}{2}}(f(a))$. Wegen $X = \cup_{a \in X} B_{\delta(a)}(a)$ und der Kompaktheit folgt, dass es $a_1, \dots, a_n \in X$ gibt mit $X = \cup_{i=1}^n B_{\delta_{a_i}}(a_i)$. Setzt man $\delta := \inf\{\delta(a_1), \dots, \delta(a_n)\}$, so ist $\delta > 0$.

Seien $x, x' \in X$ mit $d_X(x, x') < \delta$. Dann gibt es ein $i \in \{1, \dots, n\}$ mit $x \in B_{\delta(a_i)}(a_i)$. Hieraus folgt $d_X(x', a_i) \leq d_X(x', x) + d_X(x, a_i) < \delta + \delta(a_i) \leq 2\delta(a_i)$. Nach Konstruktion ist $f(x) \in B_{\frac{\epsilon}{2}}(f(a_i))$ und auch $f(x') \in B_{\frac{\epsilon}{2}}(f(a_i))$, insgesamt daher $d_Y(f(x), f(x')) < \epsilon$. $\blacksquare$

Hier ist ein Resultat, das wir aus der Analysis kennen als Existenzsatz für Minima und Maxima von stetigen Funktionen auf beschränkten, abgeschlossenen Intervallen.

Satz 2.31 Sei (X, d) ein kompakter metrischer Raum und sei $f : X \rightarrow \mathbb{R}$ stetig. Dann gibt es $u, w \in X$ mit

$$f(u) = \inf_{x \in X} f(x), \quad f(w) = \sup_{x \in X} f(x).$$

Beweis:

Es ist der Beweis offenbar nur zu einem Fall zu führen.

Aus Satz 1.29 wissen wir, dass $f(X)$ kompakt ist, also insbesondere beschränkt. Dann existiert $\alpha := \inf_{x \in X} f(x)$. Sei $(x_n)_{n \in \mathbb{N}}$ eine Minimalfolge, d. h. $\lim_n f(x_n) = \inf_{x \in X} f(x)$. Da X kompakt ist, enthält diese Folge eine konvergente Teilfolge: $u = \lim_k x_{n_k}$. Da f stetig ist, gilt $f(u) = \lim_k f(x_{n_k}) = \inf_{x \in X} f(x)$. $\blacksquare$

Bemerkung 2.32 Weierstraß¹ klärte die Begriffe „Infimum“ und „Minimum“ im Kern völlig auf und beseitigte damit die vorhandenen Unklarheiten, die aus der Auslassung von Existenzbetrachtungen bei Extremalaufgaben in unendlichdimensionalen Räumen an vielen Stellen entstanden waren; das sogenannte Dirichletproblem² war zentral dabei. $\square$

Wir erinnern an Satz 1.29. Seine Aussagen erhalten keine weitere Qualität, wenn man die topologischen Räume durch metrische Räume ersetzt. Die Interpretation der Aussage, dass das

¹Karl Weierstraß, 1815 – 1897

²P.G.L. Dirichlet, 1805 – 1859

Bild einer kompakten Menge unter einer stetigen Abbildung kompakt ist, kann aber einer neuen Betrachtung unterzogen werden, da Beschränktheit eine Konsequenz von Kompaktheit ist, wie das folgende Lemma zeigt.

Lemma 2.33 *Sei (X, d_X) ein metrischer Raum und $A \subset X$. Ist A kompakt, dann ist A abgeschlossen und beschränkt.*

Beweis:

Die Abgeschlossenheit haben wir in Satz 1.23 gezeigt. Annahme: $\text{diam}(A)$ ist nicht endlich. Dann gibt es Folgen $(x_n)_{n \in \mathbb{N}}, (x'_n)_{n \in \mathbb{N}}$ in A mit $\lim_n d(x_n, x'_n) = \infty$. Da A kompakt ist, ist A folgenkompakt nach Satz 2.28. Daher können wir o. E. annehmen, dass $(x_n)_{n \in \mathbb{N}}, (x'_n)_{n \in \mathbb{N}}$ konvergent sind. Da die Metrik d stetig ist, folgt sofort ein Widerspruch. ■

Kompaktheit und der Kontraktionssatz kommen zusammen bei einer Methode, Bilder von verblüffender Formenvielfalt zu erzeugen, die sogenannten **Fraktale**; Mandelbrot- und Julia-mengen gehören dazu. Der „Sturm“, der die letzten 20 Jahre **Chaos-Mathematik** durch die Literatur gefegt hat, hat diese Bilder mitgebracht. Sie unterscheiden sich als Figuren wesentlich von den klassischen geometrischen Figuren: Flächen- und Längenberechnungen in solchen Bildern sind schwierig, die übliche Darstellung durch einfache Funktionen ist nicht möglich. Das Wort „fraktal“ bedeutet ja gerade „zerbrechlich“, der Bruchrand eines Körpers ist mitunter nicht unähnlich dem Rand von Fraktalen.

Der Kontraktionssatz handelt von Bildabbildungen W , die iterativ angewendet werden mit dem Ergebnis der Figuren

$$W(A), W(W(A)), W(W(W(A))), \dots$$

Der Satz kann nun verwendet werden, um zu beweisen, dass die Folge dieser Figuren gegen einen so genannten **Attraktor** $A^\dagger$ (meist ein Fraktal) konvergiert. Wesentlich ist natürlich eine Kontraktionseigenschaft der Abbildung W . Aber was heißt hier Kontraktion? Hierzu haben wir aber einen Abstandsbegriff für Figuren einführen. Der geeignete Begriff ist der der **Hausdorff-Distanz**.

Seien A, B zwei kompakte Mengen in $\mathbb{R}^n$, versehen mit der Metrik d . Ihre Hausdorff-Distanz ist dann erklärt durch

$$\text{dist}(A, B) := d_H(A, B) := \max\left\{\sup_{a \in A} d(a, B), \sup_{b \in B} d(A, b)\right\}, \quad (2.8)$$

Dabei ist $d(a, B)$ definiert durch $\inf_{b \in B} d(a, b)$; analog $d(b, A)$. Auf Grund der angenommenen Kompaktheit von A, B macht die Definition einen Sinn und d_H ist in der Tat eine Metrik auf dem Raum der kompakten Teilmengen von $\mathbb{R}^n$. Damit ist $(\mathcal{K}, d_H)$ sogar ein vollständiger metrischer Raum. Dies macht es nun möglich den Kontraktionssatz auf Bildabbildungen anzuwenden.

2.5 Der Satz von Arzela-Ascoli

Aus der Analysis wissen wir, dass in $\mathbb{R}^n$ die kompakten Mengen die abgeschlossenen, beschränkten Mengen sind. Die Umkehrung von Lemma 2.33 beweist man hier mit dem Satz von Heine-Borel; siehe etwa [14]. Wir kommen auf die „Lücke“, dass dies in allgemeinen metrischen Räumen nicht gilt, im Kapitel über normierte Räume zurück. Hier wollen wir nun Kompaktheit im unendlich-dimensionalen Raum der stetigen Funktionen untersuchen. Zunächst einige Vorbereitungen.

Definition 2.34 *Sei (Y, d_Y) ein metrischer Raum, D eine Menge. Eine Abbildung $f : D \rightarrow Y$ heißt **beschränkt**, wenn das Bild $f(D)$ eine beschränkte Teilmenge von Y ist.*

□

Seien $(X, d_X), (Y, d_Y)$ metrische Räume, D eine Menge. Wir setzen

$$\begin{aligned}\mathcal{B}(D, Y) &:= \{f : D \rightarrow Y \mid f \text{ beschränkt}\} \\ \mathcal{C}(X, Y) &:= \{f : X \rightarrow Y \mid f \text{ stetig}\} \\ \mathcal{C}_b(X, Y) &:= \mathcal{B}(X, Y) \cap \mathcal{C}(X, Y)\end{aligned}$$

Lemma 2.35 Sei (Y, d_Y) ein vollständiger metrischer Raum, D eine Menge. Dann wird durch

$$d_\infty(f, g) := \sup_{x \in D} d_Y(f(x), g(x)), \quad f, g \in \mathcal{B}(D, Y),$$

eine Metrik definiert und der metrische Raum $\mathcal{B}(D, Y), d_\infty$ ist vollständig.

Beweis:

Das Supremum in der Metrikdefinition existiert, da nach Voraussetzung die beteiligten Abbildungen beschränkt sind. Die Metrikeigenschaften sind einfach nachzurechnen.

Sei $(f_n)_{n \in \mathbb{N}}$ eine Cauchyfolge in $\mathcal{B}(D, Y)$. Dann ist für jedes $x \in D$ $(f_n(x))_{n \in \mathbb{N}}$ eine Cauchyfolge in Y , und daher konvergent. Wir setzen $f(x) := \lim_n f_n(x), x \in D$. Wir haben zu zeigen, dass der punktweise Grenzwert f sogar der Grenzwert bezüglich d_∞ ist. Zuvor ist aber noch zu zeigen, dass $f \in \mathcal{B}(D, Y)$ gilt.

Sei $\varepsilon > 0$. Dann gibt es, da $(f_n)_{n \in \mathbb{N}}$ eine Cauchyfolge ist, $N \in \mathbb{N}$ mit

$$d_Y(f_n(z), f_m(z)) < \varepsilon, \quad n, m \geq N, z \in D.$$

Auf Grund der punktweisen Konvergenz gibt es zu $x \in D$ ein $N_x \in \mathbb{N}$ mit

$$d_Y(f_n(x), f(x)) < \varepsilon, \quad n \geq N_x.$$

O.E. kann man stets $N_x \geq N$ wählen. Da f_N beschränkt ist, gibt es $r > 0$ und $y \in Y$ mit $f_N(D) \subset B_r(y)$. Sei nun $z \in D$. Dann folgt

$$f(z) \in B_\varepsilon(f_N(z)) \subset U_\varepsilon(B_\varepsilon(f_N(z))) \subset B_{2\varepsilon}(f_N(z)) \subset B_{2\varepsilon+r}(y) \quad (2.9)$$

Dies zeigt, dass f beschränkt ist, also $f \in \mathcal{B}(D, Y)$.

Sei $z \in D, n \geq N$. Dann gilt

$$d_Y(f_n(z), f(z)) \leq d_Y(f_n(z), f_N(z)) + d_Y(f_N(z), f(z)) < \varepsilon + \varepsilon.$$

Dies zeigt $f = \lim_n f_n$ bezüglich der Metrik d_∞ . ■

Folgerung 2.36 Seien $(X, d_X), (Y, d_Y)$ vollständige metrische Räume. Dann wird durch

$$d_\infty(f, g) := \sup_{x \in D} d_Y(f(x), g(x)), \quad f, g \in \mathcal{B}(X, Y) \text{ bzw. } f, g \in \mathcal{C}_b(X, Y),$$

eine Metrik definiert und $(\mathcal{B}(X, Y), d_\infty), (\mathcal{C}_b(X, Y), d_\infty)$ sind vollständige metrische Räume.

Beweis:

Es ist nur noch zu zeigen, dass $\mathcal{C}_b(X, Y), d_\infty$ ein vollständiger metrischer Raum ist. Dies folgt aus der Tatsache, dass $\mathcal{C}_b(X, Y)$ abgeschlossen in $\mathcal{B}(X, Y)$ ist. Diese Abgeschlossenheit ist eine Konsequenz aus Lemma 2.19, da Konvergenz bezüglich d_∞ die gleichmäßige Konvergenz ist. ■

Satz 2.37 Sei (X, d_X) ein kompakter metrischer Raum und sei (Y, d_Y) ein vollständiger metrischer Raum. Dann ist $\mathcal{C}(X, Y) = \mathcal{C}_b(X, Y)$ und $(\mathcal{C}(X, Y), d_\infty)$ ist ein vollständiger metrischer Raum.

Beweis:

Wir wissen, dass jede stetige Abbildung $f : X \rightarrow Y$ beschränkt, da X kompakt ist; siehe Lemma 2.33. Aus Lemma 2.36 folgt nun die Aussage. ■

Nun kommen wir zu einem Kompaktheitskriterium im Raum der stetigen Funktionen, benannt nach Arzela und Ascoli.³ Damit wird aufgeklärt, was den notwendigen Bedingungen „Abgeschlossenheit, Beschränktheit“ als Voraussetzung hinzugefügt werden muss, damit hinreichende Bedingungen für Kompaktheit im Raum der stetigen Funktionen entstehen. Eine solche zusätzliche Voraussetzung ist nicht so *universell* wie die Voraussetzungen „Abgeschlossenheit, Beschränktheit“; sie ist von Fall zu Fall verschieden. Hier benötigen wir zur Formulierung einer solchen Zusatzvoraussetzung die folgende Definition.

Definition 2.38 Seien $(X, d_X), (Y, d_Y)$ metrische Räume. Eine Familie $\mathcal{F} \subset \mathcal{C}(X, Y)$ heißt **gleichgradig stetig**, falls gilt:

$$\forall \epsilon > 0 \exists \delta > 0 \forall x, x' \in X \forall f \in \mathcal{F} (d_X(x, x') < \delta \implies d_Y(f(x), f(x')) < \epsilon).$$

□

Besteht die Familie $\mathcal{F}$ nur aus einem Element, dann ist gleichgradig stetig mit gleichmäßig stetig äquivalent.

Satz 2.39 (Satz von Arzela-Ascoli) Sei (X, d_X) kompakter metrischer Raum und sei (Y, d_Y) ein vollständiger metrischer Raum. Sei $\mathcal{F} \subset \mathcal{C}(X, Y)$. Dann sind äquivalent:

- a) $\mathcal{F}$ ist relativ kompakte Teilmenge des vollständigen metrischen Raumes $(\mathcal{C}(X, Y), d_\infty)$ mit der Metrik

$$d_\infty(f, g) := \sup_{x \in X} d_Y(f(x), g(x)).$$

- b) $\mathcal{F}$ ist gleichgradig stetig und die Mengen $\mathcal{F}(x) := \{f(x) | f \in \mathcal{F}\}, x \in X$, sind relativ kompakte Teilmengen in (Y, d_Y) .

Beweis:

a) $\implies$ b) Sei $\epsilon > 0$. Da $\overline{\mathcal{F}}$ kompakt ist, gibt es $f_1, \dots, f_m \in \mathcal{F}$ mit $\mathcal{F} \subset \bigcup_{j=1}^m B_\epsilon(f_j)$. Jedes f_j ist auf der kompakten Menge X auch gleichmäßig stetig. Also existieren $\delta_1 > 0, \dots, \delta_m > 0$ mit $d_Y(f_j(x), f_j(x')) < \epsilon/3$ für alle $x, x' \in X$ mit $d_X(x, x') < \delta_j$ und alle $j = 1, \dots, m$. Setze $\delta := \min\{\delta_1, \dots, \delta_m\}$. Seien nun $x, x' \in X$ mit $d_X(x, x') < \delta$ und $f \in \mathcal{F}$ beliebig gegeben. Dann gilt $f \in B_\epsilon(f_k)$ für ein $k \in \{1, \dots, m\}$. Hieraus folgt

$$d_Y(f(x), f(x')) \leq d_Y(f(x), f_k(x)) + d_Y(f_k(x), f_k(x')) + d_Y(f_k(x'), f(x')) < \epsilon.$$

Also ist $\mathcal{F}$ gleichgradig stetig.

Sei $x \in X$. Wir müssen zeigen, dass $\mathcal{F}(x)$ relativ kompakt in Y ist. Die Abbildung $\omega_x : \mathcal{C}(X, Y) \ni f \mapsto f(x) \in Y$ erfüllt $d_Y(\omega_x(f), \omega_x(g)) \leq d_\infty(f, g)$ für alle $f, g \in \mathcal{C}(X, Y)$. ω_x ist also gleichmäßig stetig. Da $\mathcal{F}$ relativ kompakt ist, ist auch $\mathcal{F}(x) = \omega_x(\mathcal{F})$ relativ kompakt.

b) $\implies$ a) Wir wissen, dass $\mathcal{C}(X, Y)$ vollständig ist. Nach Satz 2.28 ist nur zu zeigen, dass $\mathcal{F}$ präkompakt ist, da diese Eigenschaft äquivalent zur Eigenschaft „relativ kompakt“ ist.

Sei $\epsilon > 0$. Dann existiert ein $\delta > 0$ mit $d_Y(f(x), f(x')) < \epsilon/3$ für alle $x, x' \in X$ mit $d_X(x, x') < \delta$ und für alle $f \in \mathcal{F}$. Da X kompakt, also präkompakt ist, gibt es $x_1, \dots, x_n \in X$ mit $X = \bigcup_{i=1}^n B_\delta(x_i)$. Für jedes i ist $\mathcal{F}(x_i)$ relativ kompakt in Y . Damit ist auch

$$Z := \prod_{i=1}^n \overline{\mathcal{F}(x_i)} \subset \prod_{i=1}^n Y$$

³C. Arzela, 1847-1912, und G. Ascoli, 1843-1896

kompakt; siehe hierzu Satz 1.30 oder nutze die Charakterisierung der Kompaktheit in Satz 2.28. Das Produkt $\prod_{i=1}^n Y$ ist ein vollständiger metrischer Raum mit der l_∞ -Metrik, bezeichnet mit d_Z ; vergleiche mit der Situation $\mathbb{R}^n$. Betrachte die Abbildung $\phi : \mathcal{F} \ni f \mapsto (f(x_1), \dots, f(x_n)) \in Z$. Dann ist das Bild $\phi(\mathcal{F})$ relativ kompakt bzw. präkompakt in Z . Also existieren $f_1, \dots, f_m \in \mathcal{F}$ mit $\phi(\mathcal{F}) \subset \bigcup_{i=1}^m B_{\varepsilon/3}(\phi(f_i))$.

Sei nun $f \in \mathcal{F}$. Dann existiert ein $i \in \{1, \dots, m\}$ mit $d_Z(\phi(f), \phi(f_i)) < \varepsilon/3$, d. h. für alle $j \in \{1, \dots, n\}$ ist $d_Y(f(x_j), f_i(x_j)) < \varepsilon/3$. Sei $x \in X$. Dann existiert ein j mit $d_X(x, x_j) < \delta$ und wir haben

$$d_Y(f(x), f_i(x)) \leq d_Y(f(x), f(x_j))d_Y(f(x_j), f_i(x_j)) + d_Y(f_i(x_j), f_i(x)) < \varepsilon.$$

Da X kompakt ist, ist damit auch $d_\infty(f, f_i) < \varepsilon$. ■

Von besonderem Interesse ist der Satz 2.39 für den Fall $Y = \mathbb{K}$. Hier kann die Voraussetzung „ $\mathcal{F}(x) := \{f(x) | f \in \mathcal{F}\}, x \in X$, sind relativ kompakte Teilmengen in (Y, d_Y) “ ersetzt werden durch „ $\mathcal{F}(x) := \{f(x) | f \in \mathcal{F}\}, x \in X$, sind beschränkte Teilmengen“. Wir schreiben für $\mathbb{K} = \mathbb{R}$ und $b) \implies a)$ einen etwas elementarereren Beweis auf. Also

Beweis von Satz 2.39, $a) \implies a)$ für $\mathbb{K} = \mathbb{R}$:

Sei $(f_n)_{n \in \mathbb{N}}$ eine Folge in $\mathcal{F}$. Wir haben zu zeigen, dass $(f_n)_{n \in \mathbb{N}}$ ein bezüglich d_∞ konvergente Teilfolge enthält.

Schritt 1: Sei $x \in X, n \in \mathbb{N}$. Da $\mathcal{F}$ gleichgradig stetig ist, ist

$$W_x^n := \{y \in X | |f(x) - f(y)| < \frac{1}{n} \text{ für alle } f \in \mathcal{F}\}$$

eine Umgebung von x . Da X kompakt ist und $X = \bigcup_{x \in X} W_x^n$ gilt, gibt es eine endliche Menge $X_n \subset X$ mit $X = \bigcup_{x \in X_n} W_x^n$. Die Menge $X_\infty := \bigcup_{n \in \mathbb{N}} X_n$ ist dann abzählbar und dicht. Wir setzen $X_\infty := \{w_j | j \in \mathbb{N}\}$.

Schritt 2: Nach Voraussetzung ist für jedes $w \in X_\infty$ die Folge $(f_m(w))_{m \in \mathbb{N}}$ beschränkt in $\mathbb{R}$ und enthält somit eine konvergente Teilfolge. Induktiv definieren Teilfolgen $(f_{n,j})_{n \in \mathbb{N}}$ von $(f_{n,0} := f_n)_{n \in \mathbb{N}}$, so dass für alle $j \in \mathbb{N}$ gilt: $(f_{n,j})_{n \in \mathbb{N}}$ ist eine Teilfolge von $(f_{n,j-1})_{n \in \mathbb{N}}$ und $(f_{n,j}(w_j))_{n \in \mathbb{N}}$ konvergiert. Wir betrachten damit die Diagonalfolge $(g_n := f_{n,n})_{n \in \mathbb{N}}$. Nach einem bekannten Diagonalfolgenargument konvergiert $(g_n(w_j))_{n \in \mathbb{N}}$ für alle $j \in \mathbb{N}$.

Schritt 3: Sei $\varepsilon > 0$ und sei dazu $\delta > 0$ gewählt auf Grund der gleichgradigen Stetigkeit der Familie $\mathcal{F}$. Da X kompakt und X_∞ dicht ist, gibt es $w_{i_1}, \dots, w_{i_p} \in X_\infty$ mit $X = \bigcup_{j=1}^p B_\delta(w_{i_j})$.

Schritt 4: Nun wollen wir zeigen, dass $(g_n)_{n \in \mathbb{N}}$ bezüglich d_∞ konvergiert. Sei $x \in X$ beliebig. Wähle $w := w_{i_j} \in \{w_{i_1}, \dots, w_{i_p}\}$ mit $x \in B_\delta(w)$. Dann ist nach Schritt 2 $(g_n(w))_{n \in \mathbb{N}}$ eine Cauchyfolge. Also gilt:

$$\exists N \in \mathbb{N} \forall n, m \geq N (|g_n(w) - g_m(w)| < \varepsilon)$$

Damit erhalten wir mit der Wahl von δ

$$|g_n(x) - g_m(x)| \leq |g_n(x) - g_n(w)| + |g_n(w) - g_m(w)| + |g_m(w) - g_m(x)| < \varepsilon + \varepsilon + \varepsilon$$

für alle $n, m \geq N$. Da diese Abschätzung nicht von x abhängt, ist $(g_n)_{n \in \mathbb{N}}$ eine Cauchyfolge bezüglich d_∞ und, da $\mathcal{C}(X, \mathbb{R})$ vollständig ist, konvergent bezüglich d_∞ . ■

Mit dem Satz von Arzela-Ascoli gelingt eine durchsichtige Absicherung des **Existenzsatzes von Peano** für Anfangswertaufgaben bei gewöhnliche Differentialgleichungen. Das Vorgehen bei diesem Existenzsatz stützt sich auf das **Eulerverfahren** zur Diskretisierung von gewöhnlichen Differentialgleichung. Wir skizzieren das Vorgehen im Zusammenhang mit dem Existenzsatz von **Picard-Lindelöf** im nächsten Kapitel.

2.6 Der Satz von Baire

Der Satz von Baire⁴ lässt sich auf (mindestens) drei unterschiedliche Arten formulieren. Hier sind die drei Versionen:

Satz 2.40 (Satz von Baire/1. Version) *Sei X ein vollständiger metrischer Raum und sei $(A_i)_{i \in \mathbb{N}}$ eine Familie von abgeschlossenen Teilmengen von X mit $\cup_{i \in \mathbb{N}} A_i = X$. Dann existiert ein $i \in \mathbb{N}$, so dass die Menge A_i einen inneren Punkt enthält.*

Satz 2.41 (Satz von Baire/2. Version) *Sei X ein vollständiger metrischer Raum und sei $(A_i)_{i \in \mathbb{N}}$ eine Familie von abgeschlossenen Teilmengen von X , so dass $\cup_{i \in \mathbb{N}} A_i$ einen inneren Punkt enthält. Dann existiert ein $i \in \mathbb{N}$, so dass die Menge A_i einen inneren Punkt enthält.*

Satz 2.42 (Satz von Baire/3. Version) *Sei X ein vollständiger metrischer Raum und sei $(B_i)_{i \in \mathbb{N}}$ eine Familie von offenen, dichten Teilmengen von X . Dann ist $\cap_{i \in \mathbb{N}} B_i$ dicht in X .*

Diese drei Versionen sind äquivalent. Dabei ist nur die Implikation von Version 1 zu Version 2 nicht offensichtlich. Man kann sie auf folgende Weise einsehen. Seien $A_i, i \in \mathbb{N}$, abgeschlossene Mengen in X , so dass $\cup_{i \in \mathbb{N}} A_i$ einen inneren Punkt enthält. Damit ist

$$\tilde{X} := \overline{A} \text{ mit } A := (\cup_{i \in \mathbb{N}} A_i)^\circ$$

nichtleer. Da X vollständig ist und $\tilde{X}$ abgeschlossen ist, ist $\tilde{X}$ mit der Spurtopologie vollständig. Mit $\tilde{A}_0 := \tilde{X} \setminus (\cup_{i \in \mathbb{N}} A_i)^\circ$ und $\tilde{A}_i := \tilde{X} \cap A_i, i \in \mathbb{N}$, ergibt sich

$$\tilde{X} := \cup_{i \in \mathbb{N}} \tilde{A}_i.$$

Die 1. Version des Satzes liefert nun, dass ein $i \in \mathbb{N}_0$ existiert, so dass $\tilde{A}_i$ einen inneren Punkt x enthält. Dieses i muss verschieden von 0 sein, denn $\tilde{X} \setminus \tilde{A}_0$ ist dicht in $\tilde{X}$.

Es bleibt zu beweisen, dass A_i auch einen (bezüglich der von der Metrik auf X erzeugten Topologie) inneren Punkt enthält. Dies folgt so: Wähle (in Übereinstimmung mit der Definition der Spurtopologie) $U \subset X$ offen mit $x \in \tilde{U} := U \cap \tilde{A}$. Da U offen ist, gibt es $r > 0$ mit $B_r(x) \subset U$. Damit ist aber $U \cap A$ nichtleer. Ferner gibt es eine Folge $(a_n)_{n \in \mathbb{N}}$ mit $x = \lim_n a_n$. Also gibt es ein $N \in \mathbb{N}$, so dass $a_n \in B_r(x)$ für alle $n > N$. Insbesondere folgt $a_n \in U \cap A$ für $n > N$. Es gilt nun $U \cap A = U \cap \tilde{X} \cap A = \tilde{U} \cap A \subset \tilde{U}$. Also enthält $\tilde{U}$ die nichtleere offene Menge $V := U \cap A$.

Nun haben wir eine der obigen Versionen zu beweisen. Sie wird in Kapitel 4 benutzt werden, um ein wertvolles Prinzip der linearen Funktionalanalysis zu beweisen, den Satz von Banach–Steinhaus. Wir beweisen die 2. Version. Nach Voraussetzung enthält $\cup_{i \in \mathbb{N}} A_i$ eine offene Kugel $B_{r_0}(x^0)$ ($r_0 > 0$). Wir führen folgende Annahme zum Widerspruch:

Für alle $\varepsilon > 0$ und $x \in X$ und $i \in \mathbb{N}$ gilt $(X \setminus A_i) \cap B_\varepsilon(x) \neq \emptyset$.

Sei $B_0 := B_{r_0}(x^0)$. $(X \setminus A_1) \cap B_0$ ist offen und nach Annahme nicht leer. Also gibt es eine abgeschlossene Kugel $K_1 := \overline{B}_{r_1}(x^1)$, wobei $0 < r_1 < r_0/2$ und $K_1 \subset (X \setminus A_1) \cap B_{r_0}(x^0)$ ist. Dann ist $(X \setminus A_2) \cap B_{r_1}(x^1)$ offen und nach Annahme nicht leer. Also gibt es eine abgeschlossene Kugel $K_2 := \overline{B}_{r_2}(x^2)$, wobei $0 < r_2 < r_0/2$ und $K_2 \subset (X \setminus A_2) \cap B_{r_1}(x^1)$ ist.

Dieses Vorgehen setzen wir induktiv fort: Es gibt eine abgeschlossene Kugel $K_i := \overline{B}_{r_i}(x^i)$, wobei $0 < r_i < r_{i-1}/2$ und $K_i \subset (X \setminus A_i) \cap B_{r_{i-1}}(x^{i-1})$ ist. Die Folge $(K_i)_{i \in \mathbb{N}}$ erfüllt offenbar die Voraussetzungen von Satz 1.24. Es gibt also $x \in X$ mit $\{x\} = \cap_{i \in \mathbb{N}} K_i$. Dann gilt:

$$x \in \bigcap_{i \in \mathbb{N}} (X \setminus A_i) = X \setminus \bigcup_{i \in \mathbb{N}} A_i, \text{ also } x \notin \cup_{i \in \mathbb{N}} A_i.$$

Andererseits ist aber $x \in K_1 \subset B_0 \subset \cup_{i \in \mathbb{N}} A_i$. ■

⁴Baire, Rene Louis, 1874–1932

Bemerkung 2.43 Die 2. Version des Satzes von Baire könnten wir auch so formulieren: Die Vereinigung abzählbar vieler nirgends dichter abgeschlossener Teilmengen eines vollständigen metrischen Raumes hat keine inneren Punkte. Man bezeichnet einen vollständigen metrischen wegen dieser Tatsache als Raum von **2. Kategorie**. $\square$

2.7 Approximationssätze

Wir untersuchen nun, unter welchen Umständen sich eine stetige Funktion durch „einfachere Funktionen“ gut approximieren lassen.

Sei $\mathcal{C}(X) := \mathcal{C}(X, \mathbb{K})$, wobei X stets ein kompakter metrischer Raum ist. Wir nutzen, dass in $\mathcal{C}(X)$ ein Vektorraum über dem Skalkörper $\mathbb{K}$ vorliegt. Dies folgt aus den üblichen Rechenregeln für stetige Funktionen: Summe und skalare Vielfache von stetigen Funktionen sind wieder stetig. Die approximierende Familie $F \subset \mathcal{C}(X)$ wollen wir der Situation anpassen und als linearen Teilraum von $\mathcal{C}(X)$ annehmen.

Beachte, dass eine ϵ -Kugel um $f \in \mathcal{C}(X)$ auf Grund der gegebenen Metrik

$$d_\infty(f, g) := \sup\{|f(x) - g(x)| \mid x \in X\}, \quad f, g \in \mathcal{C}(X),$$

in $\mathcal{C}(X)$ sich als „Streifen“ um den Graphen von f herausstellt. Eine Approximation einer Funktion in dieser Metrik bezeichnen wir auch als **gleichmäßige Approximation**.

Satz 2.44 (Approximationssatz von Stone) Sei X ein kompakter metrischer Raum und sei F ein linearer Teilraum von $\mathcal{C}(X) = \mathcal{C}(X, \mathbb{R})$. Es gelte:⁵

- a) $1 \in F$
- b) $f \in F \implies |f| \in F$
- c) F trennt die Punkte, d. h. zu $x, y \in X$ mit $x \neq y$ gibt es $f \in F$ mit $f(x) \neq f(y)$.

Dann gilt $\overline{F} = \mathcal{C}(X)$.

Beweis:

Wir stellen Vorüberlegungen an: Für $f_1, f_2 \in F$ gilt wegen b)

$$\sup(f_1, f_2) = \frac{1}{2}(f_1 + f_2) + \frac{1}{2}|f_1 - f_2| \in F, \quad \inf(f_1, f_2) = \frac{1}{2}(f_1 + f_2) - \frac{1}{2}|f_1 - f_2| \in F.$$

Zu $x, y \in X$ mit $x \neq y$ und $\alpha, \beta \in \mathbb{R}$ existiert ein $f \in F$ mit $f(x) = \alpha, f(y) = \beta$. Dies folgt so: Nach c) gibt es $g \in F$ mit $g(x) \neq g(y)$. Nun leistet f , definiert durch

$$f(z) := \frac{\alpha - \beta}{g(x) - g(y)}g(z) + \frac{\beta g(x) - \alpha g(y)}{g(x) - g(y)}, \quad z \in X,$$

das Verlangte. Beachte dabei, dass F ein linearer Teilraum ist, und g und 1 zu F gehört.

Sei $f \in \mathcal{C}(X)$, sei $\epsilon > 0$. Wähle zu $x, y \in X$ mit $x \neq y$ gemäß unserer Vorüberlegung $f_{x,y} \in F$ mit

$$f_{x,y}(x) = f(x), \quad f_{x,y}(y) = f(y).$$

Setzt man

$$\mathcal{U}_{x,y} := \{z \in X \mid f_{x,y}(z) < f(z) + \epsilon\},$$

so ist $\mathcal{U}_{x,y}$ auf Grund der Stetigkeit von $f_{x,y}, f$ eine offene Umgebung von x . Damit gilt nun $X = \bigcup_{x \in X} \mathcal{U}_{x,y}$; beachte hierbei, dass offenbar y auf Grund der Stetigkeit der Funktionen $f_{x,y}, f$

⁵Mit 1 bezeichnen wir hier die konstante Funktion $X \ni x \mapsto 1 \in \mathbb{R}$.

zu einem der $U_{x,y}$ gehört. Da X kompakt ist, gibt es $x_1, \dots, x_m \in X$ mit $X = U_{x_1,y} \cup \dots \cup U_{x_m,y}$. Für $f_y := \inf(f_{x_1,y}, \dots, f_{x_m,y})$ gilt nach unserer Vorüberlegung $f \in F$. Ist nun $z \in X$, dann gilt $z \in U_{x_i,y}$ für ein $i \in \{1, \dots, m\}$ und daher $f_y(z) \leq f_{x_i,y}(z) < f(z) + \epsilon$. Beachtet man noch $f_y(y) = f(y)$, so ist $V_y := \{z \in X \mid f_y(z) > f(z) - \epsilon\}$ eine offene Umgebung von y . Wiederum auf Grund der Kompaktheit von X folgt die Existenz von $y_1, \dots, y_n \in X$ mit $X = V_{y_1} \cup \dots \cup V_{y_n}$. Für $g := \sup(f_{y_1}, \dots, f_{y_n})$ gilt nun $g \in F$ und $f(z) - \epsilon < g(z) < f(z) + \epsilon$ für alle $z \in X$, d. h. $d(f, g) < \epsilon$. ■

Die Voraussetzungen im Satz 2.44 können als Reichhaltigkeitsforderungen interpretiert werden: alle Konstanten liegen in F , mit f liegt auch die Betragsfunktion $|f|$ in F . Der Beweis zeigt dann, dass mit f, g sogar $\inf(f, g), \sup(f, g)$ in F liegen.

Satz 2.45 (Approximationssatz von Stone-Weierstraß/Reeller Fall) *Sei X ein kompakter metrischer Raum und sei H ein linearer Teilraum von $\mathcal{C}(X) = \mathcal{C}(X, \mathbb{R})$. Es gelte:*

- a) $1 \in H$
- b) $f, g \in H \implies fg \in H$
- c) H trennt die Punkte von X .

Dann gilt $\overline{H} = \mathcal{C}(X)$.

Beweis:

Setze $F := \overline{H}$. Es genügt zu zeigen, dass F die Voraussetzungen von Satz 2.44 erfüllt, denn dann haben wir

$$\mathcal{C}(X) = \overline{F} = \overline{\overline{H}} = \overline{H}.$$

Da H ein linearer Teilraum von $\mathcal{C}(X)$, ist auch $F = \overline{H}$ ein linearer Teilraum; man sieht dies unter Nutzung von Folgen für den Abschluss schnell ein. Genauso sieht man ein, dass für $f, g \in F$ nach Voraussetzung b) stets auch $fg \in F$ gilt. Klar, $1 \in F$, und F trennt die Punkte. Bleibt noch zu verifizieren, dass b) aus Satz 2.44 erfüllt ist.

Wegen $|f| = \sqrt{f^2}$ genügt es zu zeigen, dass „ $f \in F, 0 \leq f \implies \sqrt{f} \in F$ “ gilt. Sei dazu $f \in F, 0 \leq f$. Da X kompakt ist, ist f beschränkt und wir können daher o. E. $f \leq 1$ annehmen. Da F ein linearer Teilraum von $\mathcal{C}(X)$ ist, ist $g := 1 - f \in F$. Unter Berücksichtigung von $0 \leq g \leq 1$ folgt nun mit der Taylorreihe für $t \mapsto \sqrt{1-t}$ (siehe irgendein Nachschlagewerk über Reihen)

$$\sqrt{f(x)} = \sqrt{1-g(x)} = \sum_{n=1}^{\infty} a_n g(x)^n, \quad a_n = \frac{1}{2n-1} 2^{-2n+1} \binom{2n-1}{n}, \quad n \in \mathbb{N}.$$

Mit der Stirlingschen Formel⁶ folgt

$$0 \leq a_n \leq cn^{-\frac{3}{2}}, \quad n \in \mathbb{N} \quad (c > 0).$$

Also konvergiert die Reihe für $\sqrt{f(\cdot)}$ gleichmäßig, d. h. $\sqrt{f(\cdot)} \in \mathcal{C}(X)$. Wir wissen aber mehr: wegen $g \in F$ ist auch $g^n \in F$ für alle $n \in \mathbb{N}$ (siehe oben) und wir erhalten $\sqrt{f} \in \overline{H} = F$. ■

Beispiel 2.46 *Ein Anwendungsbeispiel von Satz 2.45 ist etwa gegeben durch*

$$X := [0, 1] \times [0, 1], \quad H := \left\{ f \mid f(x, y) = \sum_{i=1}^n f_i(x) g_i(y), \quad f_i, g_i \in \mathcal{C}[0, 1], \quad i = 1, \dots, n \right\}.$$

H ist dicht in $\mathcal{C}(X)$. □

⁶Siehe etwa [18], S. 511.

Satz 2.47 (Approximationssatz von Stone-Weierstraß/Komplexer Fall) Sei X ein kompakter metrischer Raum und sei H ein linearer Teilraum von $\mathcal{C}(X) = \mathcal{C}(X, \mathbb{C})$. Es gelte:

- a) $1 \in H$
- b) $f, g \in H \implies fg \in H$
- c) H trennt die Punkte von X .
- d) Ist f in H , dann ist auch $\bar{f} : X \ni x \mapsto \overline{f(x)} \in \mathbb{C}$ in H .

Dann gilt $\overline{H} = \mathcal{C}(X)$.

Beweis:

Ist $f \in H$, so sind wegen

$$\Re(f) = \frac{1}{2}(f + \bar{f}), \quad \Im(f) = \frac{1}{2i}(f - \bar{f})$$

auch die Funktionen $\Re(f), \Im(f)$ in H . Damit folgt, dass $H_{\mathbb{R}} := \{g \in H \mid g(X) \subset \mathbb{R}\}$ ein linearer Teilraum von $\mathcal{C}(X, \mathbb{R})$ ist und die Eigenschaften a), b), c) in Satz 2.45 erfüllt sind. Also ist $H_{\mathbb{R}}$ dicht in $\mathcal{C}(X, \mathbb{R})$. Ist nun $f \in \mathcal{C}(X, \mathbb{C})$, so sind $\Re(f), \Im(f)$ in $\mathcal{C}(X, \mathbb{R})$ und es gibt zu $\varepsilon > 0$ Funktionen $g_1, g_2 \in H_{\mathbb{R}}$ mit

$$d_{\infty}(\Re(f), g_1) < \varepsilon, \quad d_{\infty}(\Im(f), g_2) < \varepsilon.$$

Es folgt $g := g_1 + ig_2 \in H$ und

$$d_{\infty}(f, g) \leq d_{\infty}(\Re(f), g_1) + d_{\infty}(\Im(f), g_2) < 2\varepsilon.$$

Dies zeigt die Dichtheit von H in $\mathcal{C}(X, \mathbb{C})$. ■

Folgerung 2.48 (Approximationssatz von Weierstraß) Sei $[a, b]$ ein abgeschlossenes Intervall in $\mathbb{R}$. Dann ist der lineare Teilraum der Polynome dicht in $\mathcal{C}[a, b]$.

Beweis:

Die Menge H der Polynome (in einer Variablen), betrachtet auf einem kompakten Intervall $[a, b]$, erfüllt offenbar die Voraussetzungen des Satzes 2.45. Die Trennung von Punkten ist eine Art Interpolationsaufgabe. ■

Der obige Beweis des Approximationssatzes von Weierstraß ist nicht konstruktiv. Mit Hilfe der Bernsteinpolynome lässt sich ein konstruktiver Beweis formulieren; siehe etwa [18].

2.8 Übungen

- 1.) Seien $(X, d_X), (Y, d_Y)$ metrische Räume. Wir definieren auf $X \times Y$:

$$d((x, y), (x', y')) := (d_X(x, x')^2 + d_Y(y, y')^2)^{1/2}.$$

Zeige, d ist Metrik auf $X \times Y$.

- 2.) Sei $\mathcal{F} := \{(z_n)_{n \in \mathbb{N}} \mid z_n \in \mathbb{C} \text{ für alle } n \in \mathbb{N}\}$.

- (a) Setze $l_{\infty} := \{(z_n)_{n \in \mathbb{N}} \in \mathcal{F} \mid \sup_n |z_n| < \infty\}$. Zeige: Durch

$$d_{\infty}(u, v) := \sup_n |u_n - v_n|, \quad u = (u_n)_{n \in \mathbb{N}}, v = (v_n)_{n \in \mathbb{N}},$$

wird auf l_{∞} eine Metrik definiert.

(b) Setze $l_1 := \{(z_n)_{n \in \mathbb{N}} \in \mathcal{F} \mid \sum_{i=1}^{\infty} |z_n| < \infty\}$. Zeige: Durch

$$d_1(u, v) := \sum_{i=1}^{\infty} |u_n - v_n|, \quad u = (u_n)_{n \in \mathbb{N}}, v = (v_n)_{n \in \mathbb{N}},$$

wird auf l_1 eine Metrik d_1 definiert.

- 3.) Sei $(X, \mathcal{T})$ ein kompakter topologischer Raum und sei (Y, d) ein vollständiger metrischer Raum. Dann gilt $\mathcal{C}(X, Y) = \mathcal{C}_b(X, Y)$ und $\mathcal{C}(X, Y)$ wird zusammen mit der Metrik d_∞ zu einem vollständigen metrischen Raum.
- 4.) Sei (X, d) ein metrischer Raum und sei $A \subset X$. Sei $A' := \{x \in \mathbb{R} \mid x \text{ Häufungspunkt von } A\}$. Beweise oder widerlege:
 - (a) $(A \cup B)^\circ = A^\circ \cup B^\circ$.
 - (b) $(A \cup B)' = A' \cup B'$
 - (c) $(A^\circ)' = (A')^\circ$.
- 5.) Seien $A, B \subset \mathbb{R}^n$. Zeige (unter Verwendung der euklidischen Metrik):
 - (a) $(A^\circ)^\circ = A^\circ$.
 - (b) $A + B := \{a + b \mid a \in A, b \in B\}$ ist offen, falls A offen ist.
- 6.) Sei $d(m, n) := |m - n|$, $m, n \in \mathbb{N}$. Zeige:
 - (a) $(\mathbb{N}, d)$ ist ein metrischer Raum.
 - (b) Jede Teilmenge von $\mathbb{N}$ ist offen in der durch d induzierten Topologie.
- 7.) Sei $d(m, n) := \frac{|m-n|}{mn}$, $m, n \in \mathbb{N}$.
 - (a) Zeige: $(\mathbb{N}, d)$ ist ein metrischer Raum.
 - (b) Bestimme alle offenen, abgeschlossenen, kompakten Teilmengen von $\mathbb{N}$.
- 8.) Man entscheide, ob bei d eine Metrik vorliegt.
 - (a) $d(x, y) := \sqrt{|x - y|}$, $x, y \in \mathbb{R}$.
 - (b) $d(x, y) := |x - 2y|$, $x, y \in \mathbb{R}$.
- 9.) Sei (X, d) ein metrischer Raum und seien $A, B \subset X$ abgeschlossene Mengen mit $A \cap B = \emptyset$. Dann gibt es offene Mengen U, V mit

$$A \subset U, B \subset V, U \cap V = \emptyset.$$

- 10.) Zeige das sogenannte **Lebesguesche Lemma**: Sei (X, d_X) ein metrischer Raum, sei $K \subset X$ kompakt, und sei $(U_i)_{i \in I}$ eine Familie offener Teilmengen von X mit $K \subset \bigcup_{i \in I} U_i$. Dann gibt es eine reelle Zahl $\lambda > 0$ mit folgender Eigenschaft: Für jede Teilmenge A von X mit $A \cap K \neq \emptyset$ und $\text{diam}(A) \leq \lambda$ gilt $A \subset U_i$ für ein $i \in I$.
- 11.) Seien $(X, d_X), (Y, d_Y)$ metrische Räume und sei $y \in Y$, $f : X \rightarrow Y$ stetig. Zeige, dass die Abbildung

$$X \ni x \mapsto d_Y(f(x), y) \in \mathbb{R}$$

stetig ist.

- 12.) Sei $X := \mathbb{R}$ und definiere

$$d_+(x, y) := \sqrt{2} \frac{|x - y|}{(1 + |x|)(1 + |y|)}, \quad x, y \in \mathbb{R} \text{ mit } xy \geq 0,$$

und

$$d(x, y) := \begin{cases} d_+(x, y) & , \text{ falls } xy \geq 0 \\ d_+(x, 0) + d_+(0, y) & , \text{ sonst} \end{cases}.$$

(a) Zeige: $(\mathbb{R}, d)$ ist damit ein metrischer Raum.

(b) Versuche eine geometrische Interpretation für $d(x, y)$ zu finden.

- 13.) Seien (X_n, d_n) , $n \in \mathbb{N}$, metrische Räume. Für jedes $x := \{x_n | x_n \in X_n, n \in \mathbb{N}\}$, $y := \{y_n | y_n \in X_n, n \in \mathbb{N}\}$, setze

$$d(x, y) := \sum_{n=1}^{\infty} 2^{-n} \min\{1, d_n(x_n, y_n)\}.$$

Dann wird $(\prod_{n=1}^{\infty} X_n, d)$ zu einem metrischen Raum und die Topologie $\mathcal{T}_d$ auf $\prod_{n=1}^{\infty} X_n$ stimmt mit der Produkttopologie überein.

- 14.) Betrachte $d((x_1, x_2), (y_1, y_2)) := |x_1 - y_1| + |x_2 - y_2|$, $(x_1, x_2), (y_1, y_2) \in \mathbb{R}^2$.
Zeige, d ist eine Metrik auf $\mathbb{R}^2$, und skizziere die Menge

$$K := \{(x_1, x_2) \in \mathbb{R}^2 | d((x_1, x_2), (0, 0))\}.$$

- 15.) Sei $C_b(\mathbb{R}) := \{f : \mathbb{R} \rightarrow \mathbb{R} | f \text{ stetig, beschränkt}\}$ mit der Metrik

$$d_{\infty}(f, g) := \sup_{x \in \mathbb{R}} |f(x) - g(x)|, \quad f, g \in C_b(\mathbb{R}),$$

versehen. Zeige, dass die Abbildung

$$\phi : C_b(\mathbb{R}) \rightarrow C_b(\mathbb{R}), \quad (\phi(f))(x) := \max(f(x), f(x+1)),$$

gleichmäßig stetig ist.

- 16.) Betrachte die Funktion f , definiert durch $x \mapsto x^2$ auf $D_1 := [0, 1]$ bzw. $D_2 := \mathbb{R}$.
Zeige:

- (a) f ist gleichmäßig stetig auf D_1 . Zum Beweis gebe man δ zu vorgegebenem ϵ an.
(b) f ist nicht gleichmäßig stetig auf D_2 .

- 17.) Sei (X, d) ein metrischer Raum, $D \subset X$ und sei $f : D \rightarrow \mathbb{R}$ gleichmäßig stetig. Zeige:

- (a) Es gibt genau eine Funktion $g : \overline{D} \rightarrow \mathbb{R}$ mit $g|_D = f$, g stetig.
(b) g aus (a) ist gleichmäßig stetig.

- 18.) Untersuche, ob die folgenden Funktionenfolgen $(f_n)_{n \in \mathbb{N}}$ auf $[0, a]$ ($a > 0$) gleichmäßig konvergieren.

- (a) $f_n(x) := nxe^{-nx^2}$;
(b) $f_n(x) := (1-x)x^n$;
(c) $f_n(x) := \frac{1}{1+nx^2}$.

- 19.) Betrachte die Funktionen f_n , definiert durch $f_n(x) := \frac{x^{2n}}{1+x^{2n}}$, $x \in \mathbb{R}$, $n \in \mathbb{N}$.

- (a) Bestimme $\lim_n f(x)$ für alle $x \in \mathbb{R}$.
(b) Zeige: $(f_n)_{n \in \mathbb{N}}$ konvergiert gleichmäßig auf $A := \{x \in \mathbb{R} | |x| \leq a\}$ und $B := \{x \in \mathbb{R} | |x| \geq b\}$, wobei $a < 1, b > 1$ erfüllt ist.

20.) Betrachte die Funktionen f_n , definiert durch $f_n(x) := nx(1-x)^n$, $x \in [0, 1]$, $n \in \mathbb{N}$. Zeige:

- (a) $(f_n)_{n \in \mathbb{N}}$ konvergiert punktweise, aber nicht gleichmäßig.
- (b) $\lim_{x \rightarrow 0} \lim_n f_n(x) = \lim_n \lim_{x \rightarrow 0} f_n(x)$.

21.) Sei (X, d) ein kompakter metrischer Raum und sei H ein abgeschlossener linearer Teilraum von $\mathcal{C}(X)$ mit $1 \in H$. Zeige: Es sind äquivalent:

- (a) H ist ein Verband, d. h. $h \in H \implies |h| \in H$.
- (b) H ist eine Algebra, d. h. $h \in H \implies h^2 \in H$.

22.) Zeige, dass auf dem Vektorraum $C[0, 1]$ mit Skalarkörper $\mathbb{R}$ durch

$$d(f, g) := \int_0^1 \frac{|(f-g)(t)|}{1 + |(f-g)(t)|} dt, \quad f, g \in C[0, 1],$$

eine Metrik definiert wird.

Auch durch

$$d_\infty(f, g) := \int_0^1 |(f-g)(t)| dt, \quad f, g \in C[0, 1],$$

wird eine Metrik auf $C[0, 1]$ definiert. Welchen Vorteil/Nachteil könnte man d im Vergleich zu d_∞ zuschreiben?

23.) Sei (X, d) ein metrischer Raum, $A, B \subset X$. Zeige die Äquivalenz von

- (a) $\overline{A} = \overline{B}$.
- (b) $\text{dist}(x, A) = \text{dist}(x, B)$ für alle $x \in X$.

Dabei ist $\text{dist}(x, B) := \inf_{b \in B} d(x, b)$.

24.) Sei (X, d) ein metrischer Raum und sei $A \subset X, A \neq \emptyset$. Zeige:

$$|\text{dist}(x, A) - \text{dist}(y, A)| \leq d(x, y), \quad x, y \in X.$$

25.) Definiere

$$d_1(x, y) := |x - y|, \quad d_2(x, y) := |f(x) - f(y)|, \quad x, y \in \mathbb{R},$$

mit $f(z) := \frac{z}{1+z}$. Zeige:

- (a) d_1, d_2 sind Metriken auf $\mathbb{R}$.
- (b) d_1, d_2 erzeugen dieselbe Topologie auf $\mathbb{R}$.

26.) Seien X und K kompakte metrische Räume, $f : X \times K \rightarrow \mathbb{R}$ eine stetige Abbildung. Zeige, dass die Abbildung: $g : X \rightarrow \mathbb{R}, x \mapsto \sup\{f(x, y) | y \in K\}$ stetig ist.

27.) Betrachte den Münzwurf mit den Ausgängen „Kopf (K), Zahl (Z)“. Ein unendlich wiederholter Münzwurf ist ein Element im Produktraum $\prod_{n \in \mathbb{N}} \{K, Z\}$. Zeige:

- (a) $d(x, y) := \sum_{n \in \mathbb{N}, x_n \neq y_n} 2^{-n}$, $x = (x_n)_{n \in \mathbb{N}}, y = (y_n)_{n \in \mathbb{N}} \in \prod_{n \in \mathbb{N}} \{K, Z\}$ ist eine Metrik.
- (b) Die Projektionen $p_i, i \in \mathbb{N}$, auf dem Produktraum sind stetig. Dabei werde in $X := \{K, Z\}$ die diskrete Topologie und im Produktraum (Folgenraum) $\prod_{n \in \mathbb{N}} X$ die durch die Metrik definierte Topologie verwendet.

28.) Seien $A_1, \dots, A_n$ kompakte Teilmengen des metrischen Raumes (X, d) . Zeige: $\cup_{k \in \mathbb{N}} A_k, \cap_{k \in \mathbb{N}} A_k$ sind kompakt.

- 29.) Eine Teilmenge in einem metrischen Raum heißt perfekt, wenn jedes $x \in X$ Häufungspunkt einer Folge ist. Zeige: $\mathbb{N}, \mathbb{Q}$ sind als Teilmengen von $\mathbb{R}$ perfekt.
- 30.) Sei (X, d) ein vollständiger metrischer Raum, $f : X \rightarrow X$ eine Kontraktion. Betrachte die Mengenfolge

$$X_0 := X, X_{k+1} := f(X_k), k \in \mathbb{N}.$$

- (a) Man gebe eine Abschätzung für den Durchmesser von $A_k, k \in \mathbb{N}$, an.
- (b) Zeige: Es gibt $x \in X$ mit $\bigcap_{k \in \mathbb{N}} A_k = \{x\}$.
- 31.) Sei (X, d) ein folgenkompakter metrischer Raum, sei $A \subset X$. Zeige die Äquivalenz von
- (a) A ist abgeschlossen
- (b) (A, d) ist folgenkompakter metrischer Raum
- 32.) Sei (X, d) ein metrischer Raum, sei $A \subset X$. Zeige die Äquivalenz von
- (a) A ist zusammenhängend, d. h. aus $X = A \cup B, A, B$, offen, $A \cap B = \emptyset$, folgt $A = X$ oder $B = X$.
- (b) Für alle Teilmengen $A \subset X$ mit $A \neq \emptyset, A \neq X$ gilt $\partial A \neq \emptyset$.
- 33.) Sei (X, d) ein zusammenhängender metrischer Raum (siehe vorhergehende Aufgabe) und sei $f : X \rightarrow \mathbb{R}$ stetig, $f(x) \neq 0$ für alle $x \in X$.
Zeige: Gibt es $x^0 \in X$ mit $f(x^0) > 0$, so gilt $f(x) > 0$ für alle $x \in X$.
- 34.) Sei (X, d) ein metrischer Raum, $A \subset X$. Zeige die Äquivalenz von
- (a) $\exists \bar{x} \in X \exists r \geq 0 (A \subset B_r(\bar{x}))$.
- (b) $\forall \tilde{x} \in X \exists s \geq 0 (A \subset B_s(\tilde{x}))$.
- (c) $\text{diam}(A) := \sup_{x, x' \in A} d(x, x') < \infty$.
- 35.) Sei (X, d) ein vollständiger metrischer Raum, sei $A \subset X$. Zeige die Äquivalenz der folgenden Aussagen:
- (a) A ist präkompakt.
- (b) Jede Folge in A besitzt eine in X konvergente Teilfolge.
- (c) A ist relativ kompakt.
- 36.) Betrachte den vollständigen metrischen Raum $(C([0, 1], \mathbb{R}), d_\infty)$.
- (a) Für $m > 0$ sei

$$\mathcal{F}_1 := \{f \in C([0, 1], \mathbb{R}) | f \text{ stetig differenzierbar, } |f'(t)| \leq m \text{ für alle } t \in [0, 1]\}.$$

Zeige $\mathcal{F}_1$ ist gleichgradig stetig.

- (b) Zeige $\mathcal{F}_2 := \{f_\alpha | f_\alpha(t) = t^\alpha, t \in [0, 1], 1 \leq \alpha < 2\}$ ist relativ kompakt in $(C([0, 1], \mathbb{R}), d_\infty)$.
- 37.) Betrachte die Einheitssphäre $S := \{z \in \mathbb{C} | |z| = 1\}$ und dazu die Menge

$$T(S) := \{f \in \mathcal{C}(S, \mathbb{C}) | f(z) = \sum_{k=-n}^n a_k z^k, n \in \mathbb{N}, a_{-n}, \dots, a_n \in \mathbb{C}\}.$$

Zeige:

- (a) S ist eine abgeschlossene, kompakte Teilmenge des metrischen Raumes $\mathbb{C}$. (Man verwende etwa die l_∞ -Metrik in $\mathbb{C}$)
- (b) $T(S)$ ist dicht in $(\mathcal{C}(S, \mathbb{C}), d_\infty)$.

38.) Betrachte die Einheitskugel $S := \{z \in \mathbb{C} \mid |z| = 1\}$ und die Menge

$$T := \{f \in \mathcal{C}(\mathbb{R}, \mathbb{C}) \mid f(t) = \sum_{k=-n}^n a_k e^{ikt}, n \in \mathbb{N}, a_{-n}, \dots, a_n \in \mathbb{C}\}.$$

- (a) Sei $C_{2\pi} := \{f \in \mathcal{C}(\mathbb{R}, \mathbb{C}) \mid f(t + 2\pi) = f(t) \text{ für alle } t \in \mathbb{R}\}$. Zeige:
 $\Phi : \mathcal{C}(\mathbb{R}, \mathbb{C}) \ni f \mapsto \Phi(f) \in C_{2\pi}, \Phi(f)(t) := f(e^{it}), t \in \mathbb{R}$, ist eine Bijektion.
 - (b) Zeige (unter Zuhilfenahme obiger Übungsaufgabe), dass jede stetige 2π -periodische Funktion sich auf $\mathbb{R}$ gleichmäßig durch die trigonometrischen Polynome in T approximieren lässt.
- 39.) Zeige: Die Funktion $\mathbb{R} \ni x \mapsto e^x \in \mathbb{R}$ lässt sich auf $\mathbb{R}$ nicht gleichmäßig durch Polynome (in x) approximieren. (Vergleiche mit dem Approximationssatz von Weierstraß.)

Kapitel 3

Normierte Räume

Die Mathematik ist ein mächtiges und universelles Instrument der Erkenntnis und der Lösung von Aufgaben überall dort, wo sich genügend klar definierte Strukturen abzeichnen

A. Alexandrov

Wir lernen normierte Räume und Banachräume kennen, die in Anwendungen überragende Rollen spielen; Hilberträume betrachten wir in einem speziellen Kapitel. Eine ausführliche Betrachtung widmen wir der Brücke zu den endlichdimensionalen Räumen. Zu der zugrundeliegenden Vektorraumstruktur passen die linearen Abbildungen, deren Stetigkeit wir untersuchen. Abschließend skizzieren wir die Frage nach der Existenz brauchbarer Basen.

3.1 Normen

Nun spezifizieren wir noch mehr: Die Mengen, auf denen wir Topologien betrachten, sind Vektorräume, die Metrik darauf wird nun als Länge eines „Abstandsvektors“ eingeführt.

Beachte: Den **Nullvektor** in Vektorräumen bezeichnen wir stets mit θ ; auch Nullabbildungen sind damit so erfasst.

Definition 3.1 Sei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ und sei X ein $\mathbb{K}$ -Vektorraum. Eine Abbildung $\|\cdot\| : X \rightarrow \mathbb{R}$ heißt **Norm** oder **Abstandsfunktion**, wenn gilt:

- (1) $\|x\| = 0$ genau dann, wenn $x = \theta$.
- (2) $\|ax\| = |a|\|x\|$ für alle $x \in X, a \in \mathbb{K}$.
- (3) $\|x + y\| \leq \|x\| + \|y\|$ für alle $x, y \in X$.

Das Paar $(X, \|\cdot\|)$ heißt dann ein **normierter Raum**. Ist die Bedingung (1) nicht gefordert, spricht man von einer **Halbnorm**. $\square$

Man sieht wieder, dass in einem normierten Raum $(X, \|\cdot\|)$ stets $\|x\| \geq 0$ gilt, wie dies auch erwartet werden kann, wenn $\|x\|$ einen Abstand zum Nullpunkt darstellen soll.

Man sieht sehr schnell, dass ein normierter Raum $(X, \|\cdot\|)$ mit der Metrik

$$d : X \times X \ni (x, y) \mapsto \|x - y\| \in \mathbb{R}$$

zu einem metrischen Raum, und damit auch zu einem topologischen Raum wird; offene, abgeschlossene Kugeln und Umgebungen sind damit wohldefiniert. Wir führen ein:

Schreibweise: $B_r := B_r(\theta)$, $\overline{B}_r := \overline{B}_r(\theta)$, $r \geq 0$.

Beispiel 3.2 Sei $X := \mathbb{C}^n$. Wir haben zu $p \in [1, \infty]$ die Norm

$$\|\cdot\| : X \ni z \mapsto \begin{cases} \left(\sum_{i=1}^n |z_i|^p \right)^{\frac{1}{p}} & , \text{ falls } p \in [1, \infty) \\ \max_{1 \leq i \leq n} |z_i| & , \text{ falls } p = \infty \end{cases} \in \mathbb{R}.$$

Diese Normen bezeichnen wir als p -Normen und schreiben $|\cdot|_p$; für $|\cdot|_2$ schreiben wir kurz $|\cdot|$. Die Einheitskugeln $\bar{B}_1$ für die Fälle $p = 2, 8, \infty$ finden wir in Abbildung 3.1, ebenso für den Fall $p = \frac{1}{2}$ (innerer Stern), der hier keine Bedeutung hat, der aber einen kleinen Ausblick darauf liefert, was für $p \in (0, 1)$ zu erwarten ist. Soviel hier: für $0 < p < 1$ liegt keine Norm vor.

Die Normeigenschaften sind bis auf die Dreiecksungleichung sofort klar. Für $p = 1, p = \infty$ haben wir die Dreiecksungleichung mehr oder minder schon in Beispiel 2.4 nachgerechnet. Sei nun $p \in (1, \infty)$. Zur Verifikation der Dreiecksungleichung ziehen wir das nachfolgende Lemma 3.3 heran.

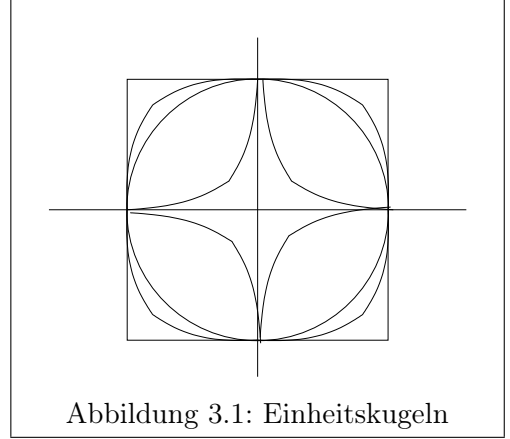


Abbildung 3.1: Einheitskugeln

Sei $q \in (1, \infty)$ mit $q^{-1} + p^{-1} = 1$. (Anmerkung: für $0 < p < 1$ wäre diese Wahl nicht möglich)

Seien $x, y \in \mathbb{C}^n$. Mit Lemma 3.3 erhalten wir:¹

$$\begin{aligned} |x + y|_p^p &= \sum_{i=1}^n |x_i + y_i|^p \\ &\leq \sum_{i=1}^n |x_i| |x_i + y_i|^{p-1} + \sum_{i=1}^n |y_i| |x_i + y_i|^{p-1} \\ &\leq \left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}} \left(\sum_{i=1}^n |x_i + y_i|^{(p-1)q} \right)^{\frac{1}{q}} + \left(\sum_{i=1}^n |y_i|^p \right)^{\frac{1}{p}} \left(\sum_{i=1}^n |x_i + y_i|^{(p-1)q} \right)^{\frac{1}{q}} \\ &\leq \left\{ \left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}} + \left(\sum_{i=1}^n |y_i|^p \right)^{\frac{1}{p}} \right\} \left(\sum_{i=1}^n |x_i + y_i|^p \right)^{1 - \frac{1}{p}} \end{aligned}$$

Daraus liest man nun die Dreiecksungleichung ab. □

Lemma 3.3 (Höldersche Ungleichung) Seien $x_i, y_i \in \mathbb{C}, 1 \leq i \leq n, p \in \mathbb{R}$ mit $1 < p < \infty$. Definiere q durch $p^{-1} + q^{-1} = 1$. Dann gilt

$$\left| \sum_{i=1}^n x_i y_i \right| \leq \sum_{i=1}^n |x_i| |y_i| \leq \left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}} \left(\sum_{i=1}^n |y_i|^q \right)^{\frac{1}{q}}.$$

Beweis:

Als Vorbereitung führen wir für $a, b, r \in \mathbb{R}$ mit $a \geq 0, b \geq 0, r \in (0, 1)$ die folgende Ungleichung an:

$$a^r b^{1-r} \leq ra + (1-r)b \tag{3.1}$$

¹Für $a \geq 0$ sei mit $a^{\frac{1}{p}}$ die nichtnegative Lösung von $x^p = a$ bezeichnet.

Für $a = 0$ oder $b = 0$ ist nichts zu beweisen. Sei nun $0 < a \leq b$. Die stetige Funktion

$$[a, b] \ni t \longmapsto t^{-r} \in \mathbb{R}$$

ist monoton fallend. Daher ist

$$b^{1-r} - a^{1-r} = (1-r) \int_a^b t^{-r} dt \leq (1-r)(b-a)a^{-r}$$

und es folgt

$$a^r b^{1-r} \leq a + (1-r)(b-a) = ra + (1-r)b.$$

Ist $0 < b < a$, dann folgt die Aussage durch Anwendung des eben Bewiesenen nach Vertauschung von r mit $r-1$.

Nun zum eigentlichen Beweis. Setze $r := \frac{1}{p}$. Es ist dann $1-r = \frac{1}{q}$. O. E. $\sum_{j=1}^n |x_j|^p \neq 0$. Sei $a_i := (\sum_{j=1}^n |x_j|^p)^{-1} |x_i|^p$, $b_i := (\sum_{j=1}^n |y_j|^q)^{-1} |y_i|^q$, $i = 1, \dots, n$. Nach Ungleichung (3.1) ist

$$\begin{aligned} |x_i| |y_i| \left(\sum_{j=1}^n |x_j|^p \right)^{-\frac{1}{p}} \left(\sum_{j=1}^n |y_j|^q \right)^{-\frac{1}{q}} &\leq a_i^{\frac{1}{p}} b_i^{\frac{1}{q}} \leq \frac{1}{p} a + \frac{1}{q} b \\ &= \frac{1}{p} \left(\sum_{j=1}^n |x_j|^p \right)^{-1} |x_i|^p + \frac{1}{q} \left(\sum_{j=1}^n |y_j|^q \right)^{-1} |y_i|^q. \end{aligned}$$

Summation über i und Umstellung ergibt die Behauptung. ■

Hinter der Ungleichung (3.1) verbirgt sich die so genannte **Youngsche Ungleichung**:

$$ab \leq \frac{1}{p} a^p + \frac{1}{q} b^q \text{ für } a, b > 0, \frac{1}{p} + \frac{1}{q} = 1. \quad (3.2)$$

Wir geben hier einen alternativen Beweis. Die Funktion $z \longmapsto \log(z)$ ist konkav. Also gilt

$$\log\left(\frac{1}{p} a^p + \frac{1}{q} b^q\right) \geq \frac{1}{p} \log(a^p) + \frac{1}{q} \log(b^q) = \log(ab)$$

und Delogarithmieren führt zum Ergebnis.

Wir formulieren Konvergenz im Kontext von normierten Räumen.

Definition 3.4 Sei $(X, \|\cdot\|)$ ein normierter Raum. Eine Folge $(x_n)_{n \in \mathbb{N}}$ in X **konvergiert** gegen $x \in X$, wenn

$$\forall \varepsilon > 0 \exists N \in \mathbb{N} \forall n \geq N (\|x_n - x\| < \varepsilon)$$

gilt. x heißt dann der (eindeutig bestimmte!) **Grenzwert** oder **Limes** von $(x_n)_{n \in \mathbb{N}}$. □

Lemma 3.5 Sei $(X, \|\cdot\|)$ ein normierter Raum und sei $A \subset X$. Dann sind äquivalent:

- a) A ist abgeschlossen.
- b) Ist $(x_n)_{n \in \mathbb{N}}$ eine konvergente Folge mit $x_n \in A$ für alle $n \in \mathbb{N}$, so gilt $\lim_n x_n \in A$.

Beweis:

Siehe Satz 2.11. ■

3.2 Banachräume

Definition 3.6 Sei $(X, \|\cdot\|)$ ein normierter Raum.

a) Eine Folge $(x_n)_{n \in \mathbb{N}}$ heißt **Cauchyfolge** in X , wenn gilt:

$$\forall \varepsilon > 0 \exists N \in \mathbb{N} \forall n, m \geq N (\|x_n - x_m\| < \varepsilon).$$

b) $(X, \|\cdot\|)$ heißt **vollständig** oder ein **Banachraum**, wenn jede Cauchyfolge in X gegen ein x in X konvergiert.

□

Aus der Analysis I wissen wir, dass $\mathbb{R}$ und $\mathbb{C}$, betrachtet als normierter Raum – die Norm ist der Abstand –, vollständig sind. Daraus schließt man sofort, dass auch $(\mathbb{R}^n, |\cdot|_\infty)$ und $(\mathbb{C}^n, |\cdot|_\infty)$ vollständig sind („Rechnen in Koordinaten“). Mit der Ungleichung (3.5) folgt dann, dass auch $(\mathbb{R}^n, |\cdot|_p)$ und $(\mathbb{C}^n, |\cdot|_p)$, $1 \leq p \leq \infty$, vollständig sind.

Aus der Analysis wissen wir auch, dass jede abgeschlossene Kugel $\overline{B}_r(z)$ folgenkompakt ist. Wir können dies auch aus Satz 2.28 schließen. Als abgeschlossene Teilmenge ist die Einheitskugel

$$S_1 := S_1(\theta) := \{x \in \mathbb{R}^n \mid |x|_2 = 1\}$$

folgenkompakt.

Beispiel 3.7 Sei $C[a, b] := \{f : [a, b] \rightarrow \mathbb{R} \mid f \text{ stetig}\}$. Wir machen $X := C[a, b]$ zu einem normierten Raum durch

$$\|f\|_\infty := \max_{t \in [a, b]} |f(t)|, \quad f \in C[a, b]. \quad (3.3)$$

Diese Norm wird von der Metrik

$$d_\infty(f, g) := \sup_{t \in [a, b]} |f(t) - g(t)|$$

„induziert“. Wir wissen deshalb nach Satz 2.30 schon, dass $(C[a, b], \|\cdot\|_\infty)$ ein Banachraum ist.

□

Nun wollen wir die **reellen Folgenräume** einführen. Sie sind unendlichdimensionale Banachräume, in denen einfach zu „rechnen“ ist und mit denen viele Eigenschaften zu leicht illustrieren sind.

Wir setzen

$$l := \{(x_n)_{n \in \mathbb{N}} \mid x_n \in \mathbb{R} \text{ für alle } n \in \mathbb{N}\}.$$

l ist ein Vektorraum mit Skalarkörper $\mathbb{R}$; Addition und skalare Multiplikation erfolgt folgenreihweise. Spezielle Elemente in l sind:

- $e^n := (e_k^n)_{k \in \mathbb{N}}$ mit $e_k^n := \delta_{nk}$, $n, k \in \mathbb{N}$.
- $f^n := (f_k^n)_{k \in \mathbb{N}}$ mit $f_k^n := 0$ für $k \geq n$, $f_k^n = 1$ für $k < n$, $n \in \mathbb{N}$.

Die Elemente des Teilraums $f := \{(x_n)_{n \in \mathbb{N}} \in l \mid x_n = 0 \text{ für fast alle } k \in \mathbb{N}\}$ lassen sich durch Linearkombinationen der Elemente e^n , $n \in \mathbb{N}$, darstellen.

Nun setzen wir mit $p \in [1, \infty]$ formal auf l

$$\|(x_n)_{n \in \mathbb{N}}\|_p := \begin{cases} \left(\sum_{i=1}^{\infty} |x_i|^p \right)^{\frac{1}{p}}, & \text{falls } p \in [1, \infty) \\ \max_{1 \leq i < \infty} |x_i| & \text{falls } p = \infty \end{cases}.$$

und nutzen dies zur Definition

$$l_p := \{x = (x_n)_{n \in \mathbb{N}} \in l \mid \|x\|_p < \infty\}. \quad (3.4)$$

Damit wird $(l_p, \|\cdot\|_p)$ zu einem normierten Raum. Der Nachweis der Norm-Eigenschaften „Definitheit, Homogenität“ ist einfach. Der Beweis der Dreiecksungleichung für $p \in [1, \infty)$ stützt sich mit $q := 1 - \frac{1}{p}$ in bekannter Weise erneut auf die Hölder-Ungleichung, die auch für „Reihen“ gilt, wie man aus der Zeile

$$\left| \sum_{k=1}^n x_k y_k \right| \leq \left(\sum_{k=1}^n |x_k|^p \right)^{\frac{1}{p}} \left(\sum_{k=1}^n |y_k|^q \right)^{\frac{1}{q}} \leq \left(\sum_{k=1}^{\infty} |x_k|^p \right)^{\frac{1}{p}} \left(\sum_{k=1}^{\infty} |y_k|^q \right)^{\frac{1}{q}}$$

mit Grenzübergang $n \rightarrow \infty$ abliest.

Lemma 3.8 Die Räume $(l_p, \|\cdot\|_p)$, $1 \leq p \leq \infty$, sind Banachräume.

Beweis:

Den Nachweis führen wir nur für $1 \leq p < \infty$, der Beweis für $p = \infty$ ist sogar noch elementarer.

Sei also $(Z^n)_{n \in \mathbb{N}}$ eine Cauchyfolge in l_p ; $Z^n := (x_k^n)_{k \in \mathbb{N}}$, $n \in \mathbb{N}$. Sei $\varepsilon > 0$. Nun gibt es $N \in \mathbb{N}$ mit

$$\|Z^m - Z^n\|_p \leq \varepsilon, \quad m, n \geq N.$$

Für alle $k \in \mathbb{N}$ gilt

$$|x_k^m - x_k^n|^p \leq \sum_{l=1}^{\infty} |x_l^m - x_l^n|^p = \|Z^m - Z^n\|_p^p, \text{ also } |x_k^m - x_k^n| \leq \|Z^m - Z^n\|_p \leq \varepsilon, \quad m, n \geq N.$$

Damit ist für festes k die Folge $(x_k^m)_{m \in \mathbb{N}}$ eine Cauchy-Folge in $\mathbb{K}$ und daher konvergent: $x_k := \lim_m x_k^m$, $k \in \mathbb{N}$, $Z := (x_k)_{k \in \mathbb{N}}$. Wegen

$$\left(\sum_{k=1}^M |x_k^m - x_k^n|^p \right)^{\frac{1}{p}} \leq \|Z^m - Z^n\|_p \leq \varepsilon, \quad M \in \mathbb{N}, m, n \geq N,$$

erhalten wir durch Grenzübergang $n \rightarrow \infty$

$$\left(\sum_{k=1}^M |x_k^m - x_k|^p \right)^{\frac{1}{p}} \leq \varepsilon, \quad M \in \mathbb{N}, m \geq N,$$

Da M beliebig war, gilt

$$\|Z^m - Z\|_p \leq \varepsilon, \quad m \geq N.$$

Daraus folgt zum Einen $Z = Z^N + (Z - Z^N) \in l_p$ und zum Anderen $Z = \lim_m Z^m$ in l_p . ■

Lemma 3.9 Wir setzen

$$\begin{aligned} c &:= \{x = (x_n)_{n \in \mathbb{N}} \in l \mid \lim_n x_n \text{ existiert}\} \\ c_0 &:= \{x = (x_n)_{n \in \mathbb{N}} \in l \mid \lim_n x_n = 0\} \end{aligned}$$

- (a) c, c_0 sind abgeschlossene Teilräume von l_∞ mit $c_0 \subset c \subset l_\infty$.
- (b) Die Räume $(c, \|\cdot\|_\infty)$, $(c_0, \|\cdot\|_\infty)$ sind Banachräume.
- (c) $l_p \subset l_q \subset c_0 \subset c \subset l_\infty$, $1 \leq p \leq q < \infty$.
- (d) $\bigcup_{1 \leq p < q} l_p \subset l_q \subset \bigcap_{q < p < \infty} l_p$.

Beweis:

$a)$ ist offenbar. Da $(l_\infty, \|\cdot\|_\infty)$ ein Banachraum ist, sind wegen $a)$ auch $(c, \|\cdot\|_\infty), (c_0, \|\cdot\|_\infty)$ vollständig.

Zu $c)$: Man hat für $l_p \subset l_q$ zu zeigen $\|(x_n)_{n \in \mathbb{N}}\|_p \leq \|(x_n)_{n \in \mathbb{N}}\|_q$. O.E. kann man $\|(x_n)_{n \in \mathbb{N}}\|_q = 1$ annehmen. Dann ist aber die Ungleichung sofort zu sehen, denn dann ist $|x_n| \leq 1, n \in \mathbb{N}$. Die Aussage $l_q \subset c_0 \subset c \subset l_\infty$ ist nahezu trivial.

Den Beweis zu $d)$ überlassen wir dem Leser. ■

Die Inklusion $l_p \subset l_q$ ist für $1 \leq p < q < \infty$ strikt, denn die Folge $(n^{-s})_{n \in \mathbb{N}}$ mit $ps \leq 1 < qs$ liegt in $l_q \setminus l_p$.

3.3 Stetigkeit

Wir formulieren die Definition der Stetigkeit, wie wir sie in topologischen Räumen gegeben haben, um auf den Fall normierter Räume:

Definition 3.10 Seien $(X, \|\cdot\|_X), (Y, \|\cdot\|_Y)$ normierte Räume. Eine Abbildung $f : D \rightarrow Y, D \subset X$ heißt stetig in $z \in D$, wenn gilt:

$$\forall \varepsilon > 0 \exists \delta > 0 \forall x \in D (\|x - z\|_X < \delta \implies \|f(x) - f(z)\|_Y < \varepsilon).$$

□

Beispiel 3.11 Sei $(X, \|\cdot\|)$ ein normierter Raum. Die Normabbildung

$$f : X \ni x \mapsto \|x\| \in \mathbb{R}$$

ist stetig. Dies folgt aus

$$|f(x) - f(y)| = ||x| - |y|| \leq \|x - y\|.$$

□

Ist $(X, \|\cdot\|_X)$ ein normierter Raum, so ist X ein topologischer Raum. Dann ist auch $X \times X$ ein topologischer Raum, versehen mit der Produkttopologie. Der Produktraum $X \times X$ ist sogar ein normierter Raum in einer Weise, dass die Produkttopologie von der Norm erzeugt wird. Diese Norm kann etwa so gewählt werden:

$$\|(x, x')\| := (\|x\|_X^2 + \|x'\|_X^2)^{\frac{1}{2}}.$$

Dann werden die (Struktur-)Abbildungen

$$+ : X \times X \ni (x, x') \mapsto x + x' \in X, \cdot : \mathbb{K} \times X \ni (a, x) \mapsto ax \in X$$

im Vektorraum X als stetig erkannt und X kann damit als topologischer Vektorraum angesehen werden.

Definition 3.12 Seien $\|\cdot\|_1, \|\cdot\|_2$ zwei Normen auf dem Vektorraum X . Diese Normen heißen **äquivalent**, wenn es Konstanten $c, d > 0$ gibt mit

$$c\|x\|_1 \leq \|x\|_2 \leq d\|x\|_1 \text{ für alle } x \in X.$$

□

Offenbar ist „äquivalent“ eine Äquivalenzrelation.

Man rechnet für $p \in [1, \infty)$ sehr einfach die folgende Ungleichung für die l_∞ -Normen auf $\mathbb{C}^n$ nach:

$$|x|_\infty \leq |x|_p \leq n^{\frac{1}{p}} |x|_\infty, \quad x \in \mathbb{C}^n. \quad (3.5)$$

Man sieht hier in der rechten Schranke die Abhängigkeit von der Dimension. Sie ist auch unvermeidbar und deutet schon an, dass solche Abschätzungen möglicherweise im unendlichdimensionalen Kontext nicht möglich sind.

Die Ungleichung (3.5) besagt, dass in $\mathbb{C}^n$ die Konvergenz, betrachtet in unterschiedlichen Normen $|\cdot|_p$, zu keinen unterschiedlichen Ergebnissen führt. Man fasst dies damit zusammen, dass die Normen-Familie $|\cdot|_p, p \in [1, \infty]$, eine Familie äquivalenter Normen ist.

Bemerkung 3.13 Betrachte die normierten Räume $(\mathbb{C}^n, |\cdot|_p), 1 \leq p \leq \infty$. Sei $D \subset \mathbb{C}^n$ und sei $f : D \rightarrow \mathbb{C}^m$ eine Abbildung. Die Ungleichung (3.5) besagt, dass die Stetigkeit von f auch überprüft werden kann, wenn man die Norm $|\cdot|_p$ gegen die Norm $|\cdot|_\infty$ austauscht. Damit fällt das Rechnen meist einfacher, da koordinatenweise gerechnet werden kann. $\square$

Satz 3.14 Sei X ein endlichdimensionaler Vektorraum (über $\mathbb{R}$ oder $\mathbb{C}$). Dann sind alle Normen auf X äquivalent.

Beweis:

Schritt 1:

Wir zeigen zunächst das Resultat, dass es zu jeder Norm $\|\cdot\|$ auf $\mathbb{C}^n$ Konstanten $d_1, d_2 > 0$ gibt mit

$$d_1 |a|_\infty \leq \|a\| \leq d_2 |a|_\infty, \quad a \in \mathbb{C}^n. \quad (3.6)$$

Wähle dazu die Basis der kanonischen Einheitsvektoren $e^1, \dots, e^n$ in $\mathbb{C}^n$. Damit gilt

$$\|a\| = \left\| \sum_{j=1}^n a_j e^j \right\| \leq \max_{1 \leq j \leq n} \|e^j\| \cdot |a|_\infty, \quad a = (a_1, \dots, a_n) \in \mathbb{C}^n.$$

Damit ist die eine Seite von (3.6) gezeigt und wir lesen ab, dass die Abbildung

$$f : \mathbb{C}^n \ni a = (a_1, \dots, a_n) \mapsto \|a\| \in \mathbb{R}$$

stetig ist, wobei in $\mathbb{C}^n$ die Norm $|\cdot|_\infty$ Verwendung findet. Betrachte nun die Einheitssphäre

$$S := \{a \in \mathbb{C}^n \mid |a|_\infty = 1\}.$$

Da die Normabbildung stetig ist, ist S eine abgeschlossene Teilmenge der folgenkompakten Menge $\overline{B}_1$, also selbst folgenkompakt. Daher nimmt die Abbildung f auf S das Minimum an; siehe 2.31. Also gibt es $a^* \in S$ mit

$$\|a\| \geq \|a^*\| =: d_1 \text{ für alle } a \in S.$$

Da $a^* \in S$ ist, kann $\|a^*\|$ nicht verschwinden und es ist daher $d_1 > 0$.

Sei nun $a \in \mathbb{C}^n, a \neq \theta$. Dann haben wir

$$\|a\| = |a|_\infty \left\| \frac{a}{|a|_\infty} \right\| \geq |a|_\infty d_1.$$

Damit ist die linke Abschätzung von (3.6) für alle $a \neq \theta$ gezeigt; für $a = \theta$ ist die Abschätzung aber trivialerweise richtig.

Schritt 2:

Seien $\|\cdot\|, \|\cdot\|_\sim$ zwei Normen in X und sei $x^1, \dots, x^n$ eine Basis in X . Wir definieren in $\mathbb{C}^n$ zwei Normen durch

$$\|a\|_* := \left\| \sum_{i=1}^n a_i x^i \right\|, \quad \|a\|_{*,\sim} := \left\| \sum_{i=1}^n a_i x^i \right\|_\sim, \quad a = (a_1, \dots, a_n) \in \mathbb{C}^n.$$

Nach Schritt 1 sind beide Normen $\|\cdot\|_*, \|\cdot\|_{*,\sim}$ äquivalent zu $|\cdot|_\infty$. Dann sind aber auch $\|\cdot\|_*, \|\cdot\|_{*,\sim}$ äquivalent, und es gibt Konstanten $c_1, c_2 > 0$ mit $c_1 \|a\|_* \leq \|a\|_{*,\sim} \leq c_2 \|a\|_*, a \in \mathbb{C}^n$. Damit folgt für $x = \sum_{i=1}^n a_i x^i \in X, a = (a_1, \dots, a_n) \in \mathbb{C}^n$,

$$\|x\| = \|a\|_* \leq c_1^{-1} \|a\|_{*,\sim} = c_1^{-1} \|x\|_\sim, \quad \|x\|_\sim = \|a\|_{*,\sim} \leq c_2 \|a\|_* = c_2 \|x\|.$$

Damit ist alles gezeigt. $\square$

Satz 3.14 besagt, dass in einem endlichdimensionalen Vektorraum über $(\mathbb{R} \text{ oder } \mathbb{C})$ die induzierte Topologie und die metrischen Eigenschaften von der gewählten Norm unabhängig sind. Anders verhält es sich mit den „geometrischen Eigenschaften“, wie wir noch sehen werden.

Satz 3.15 *Sei $(X, \|\cdot\|)$ ein normierter Raum. Dann gilt:*

- a) *Ist $\dim X < \infty$, so ist $(X, \|\cdot\|)$ vollständig.*
- b) *Jeder lineare Teilraum U von X mit $\dim U < \infty$ ist eine abgeschlossene Teilmenge von X .*
- c) *Die Einheitskugel $\overline{B}_1$ ist folgenkompakt genau dann, wenn $\dim X < \infty$ gilt..*

Beweis:

Zu a). Sei $\{x^1, \dots, x^n\}$ eine Basis von X . Sei $\|\cdot\|_\sim$ die folgende Norm auf X .

$$\|\cdot\|_\sim : X \ni x = \sum_{i=1}^n a_i x^i \mapsto |(a_1, \dots, a_n)|_\infty \in \mathbb{R}. \quad (3.7)$$

Da $(\mathbb{C}^n, |\cdot|_\infty)$ vollständig ist, ist auch $(X, \|\cdot\|_\sim)$ vollständig und da nach Satz 3.14 die Norm $\|\cdot\|_\sim$ zu $\|\cdot\|$ äquivalent ist, ist auch $(X, \|\cdot\|)$ vollständig.

Zu b). Sei $U \subset X$ ein linearer Teilraum mit $\dim U < \infty$. Nach a) ist $(U, \|\cdot\|)$ vollständig. Daraus folgt sofort, dass U auch abgeschlossen ist (siehe Lemma 3.5).

Zu c). Sei $\dim X < \infty$. Betrachte erneut die Norm $\|\cdot\|_\sim$ aus 3.7. Der Satz von Bolzano–Weierstraß besagt, dass die Einheitskugel in dieser Norm folgenkompakt ist. Nach Satz 3.14 ist auch die Einheitskugel in der gegebenen Norm folgenkompakt. Sei $\overline{B}_1$ folgenkompakt, also kompakt; siehe Satz 2.28. Nach Satz 2.28 (Präkompaktheit) gibt es Punkte $x^1, \dots, x^m \in \overline{B}_1$ mit

$$\overline{B}_1 \subset B_{\frac{1}{2}}(x_1) \cup \dots \cup B_{\frac{1}{2}}(x_m).$$

Sei U der von $\{x_1, \dots, x_m\}$ erzeugte lineare Teilraum von X . Annahme: $U \neq X$.

Wähle $x \in X \setminus U$. Da U abgeschlossen ist nach b), gilt

$$a := \text{dist}(x, U) := \inf\{\|x - u\| \mid u \in U\} > 0.$$

Also existiert $y \in U$ mit

$$a \leq \|x - y\| \leq \frac{3}{2}a. \quad (3.8)$$

Für $z := \frac{1}{\|x - y\|}(x - y)$ gilt $\|z\| = 1$, d. h. $z \in \overline{B}_1$. Also $\|z - x_i\| \leq \frac{1}{2}$ für ein $i \in \{1, \dots, m\}$.

Ferner gilt

$$x = y + \|x - y\|z = y + \|x - y\|x_i + \|x - y\|(z - x_i)$$

mit $y + \|x - y\|x_i \in U$. Nach Definitionen von $\text{dist}(x, U)$ impliziert dies

$$a \leq \| \|x - y\|(z - x_i) \| = \|x - y\| \|z - x_i\|, \text{ also } \|x - y\| \geq \frac{a}{\|z - x_i\|} \geq 2a.$$

Dies ist im Widerspruch zu (3.8) ■

Die interessanten Abbildungen in der linearen Algebra sind die linearen Abbildungen. Diesen wollen wir uns nun zuwenden. Als ein Resultat erhalten wir, dass alle linearen Abbildungen auf endlichdimensionalen Räumen stetig sind.

Satz 3.16 *Seien $(X, \|\cdot\|_X), (Y, \|\cdot\|_Y)$ normierte Räume und sei $L : X \rightarrow Y$ linear. Dann sind äquivalent:*

- (a) *L ist beschränkt, d. h. es gibt $c > 0$ mit $\|L(x)\|_Y \leq c\|x\|_X$ für alle $x \in X$.*
- (b) *L ist stetig in jedem $x^0 \in X$.*
- (c) *L ist stetig in $x^0 := \theta$.*

Beweis:

(a) $\implies$ (b). Sei $x^0 \in X$. Sei $\varepsilon > 0$. Wähle $\delta := \frac{\varepsilon}{c}$. Für $\|x - x^0\|_X < \delta$ gilt dann

$$\|L(x) - L(x^0)\|_Y = \|L(x - x^0)\|_Y \leq c\|x - x^0\|_X < \varepsilon.$$

(b) $\implies$ (c). Klar.

(c) $\implies$ (a). Wähle $\varepsilon := 1$ und dazu $\delta > 0$ mit

$$\|x - \theta\|_X < \delta \implies \|L(x) - L(\theta)\|_Y < 1.$$

Für $x \in X \setminus \{\theta\}$ gilt dann mit $z := x\|x\|_X^{-1} \frac{\delta}{2}$

$$\|z\|_X = \frac{\delta}{2}, \|L(z)\|_Y \leq 1, \text{ d. h. } \|L(x)\|_Y \leq \frac{2}{\delta}\|x\|_X$$

Setze nun $c := \frac{2}{\delta}$. ■

Beachte, dass Stetigkeit von L in einem Punkt x auch mit der Folgenstetigkeit geprüft werden kann.

Lineare Abbildungen zwischen normierten Räumen werden, wenn sie stetig sind, auch meist als **beschränkt** bezeichnet. Dies hat seinen Grund darin, dass nach Satz 3.16 bei stetigen Abbildungen das Bild der Einheitskugel beschränkt ist.

Beispiel 3.17 *Seien $(X, \|\cdot\|_X), (Y, \|\cdot\|_Y)$ normierte Räume und sei $L : X \rightarrow Y$ eine lineare Abbildung. Dann ist L stetig genau dann, wenn L gleichmäßig stetig ist. Dies folgt aus der Tatsache, dass bei einer linearen Abbildung Stetigkeit äquivalent ist mit einer Abschätzung*

$$\|L(x) - L(x')\|_Y \leq c\|x - x'\|_X, \quad x, x' \in X,$$

mit einem $c \in \mathbb{R}$; siehe Satz 3.16 □

Seien $(X, \|\cdot\|_X), (Y, \|\cdot\|_Y)$ normierte Räume mit Skalarkörper $\mathbb{K}$.

Bezeichnung: $\mathcal{B}(X, Y) := \{L : X \rightarrow Y \mid L \text{ linear und stetig}\}$, $\mathcal{B}(X) := \mathcal{B}(X, X)$

Klar, $\mathcal{B}(X, Y)$ ist ein Vektorraum mit Skalarkörper $\mathbb{K}$.

Definition 3.18 Seien $(X, \|\cdot\|_X), (Y, \|\cdot\|_Y)$ normierte Räume und sei $L : X \rightarrow Y$ linear und stetig. Dann heißt $\|L\|_{X \rightarrow Y} := \sup\{\|L(x)\|_Y \mid \|x\|_X \leq 1\}$ die **Operatornorm** von L . $\square$

Aus den Äquivalenzen von Satz 3.16 entnimmt man, dass die Definition der Operatornorm Sinn macht.

Satz 3.19 Seien $(X, \|\cdot\|_X), (Y, \|\cdot\|_Y)$ normierte Räume. Dann ist $(\mathcal{B}(X, Y), \|\cdot\|_{X \rightarrow Y})$ ein normierter Raum. Er ist vollständig, wenn $(Y, \|\cdot\|_Y)$ vollständig ist.

Beweis:

Es ist einfach einzusehen, dass die Operatornorm in der Tat eine Norm ist. Die Vollständigkeitsaussage folgt so: Sei $(L_k)_{k \in \mathbb{N}}$ eine Cauchyfolge in $(\mathcal{B}(X, Y), \|\cdot\|_{X \rightarrow Y})$. Wegen

$$\|L_k x - L_l x\|_Y \leq \|L_k - L_l\|_{X \rightarrow Y} \|x\|_X$$

ist für jedes $x \in X$ die Folge $(L_k x)_{k \in \mathbb{N}}$ eine Cauchy-Folge in Y . Da Y vollständig ist, konvergiert $(L_k x)_{k \in \mathbb{N}}$ gegen einen Grenzwert, den wir Lx nennen. Nun ist einfach nachzurechnen, dass hierdurch eine lineare Abbildung $L : X \rightarrow Y, Lx := \lim_k L_k x$, definiert wird. Sei nun $\varepsilon > 0$ beliebig. Dann gibt es $N \in \mathbb{N}$ mit

$$\|L_k - L_l\|_{X \rightarrow Y} \leq \varepsilon \text{ für alle } k, l \geq N.$$

Daraus folgt

$$\|L_k x - L_l x\|_Y \leq \|L_k - L_l\|_{X \rightarrow Y} \|x\|_X \leq \varepsilon \|x\|_X \text{ für alle } k, l \geq N.$$

Grenzübergang $l \rightarrow \infty$ liefert

$$\|L_k x - Lx\|_Y \leq \varepsilon \|x\|_X \text{ für alle } k \geq N.$$

Nun bilden wir das Supremum über $\overline{B}_1$ und erhalten damit

$$\|L_k - L\|_{X \rightarrow Y} \leq \varepsilon \text{ für alle } k \geq N.$$

Wegen $L = L - L_k + L_k$ ist $L \in \mathcal{B}(X, Y)$. Ferner haben wir gezeigt $L = \lim_k L_k$ in $\mathcal{B}(X, Y)$. $\blacksquare$

Nun schreiben wir statt $\|L\|_{X \rightarrow Y}$ meist kurz $\|L\|$. Aus dem Kontext wird schon klar werden, auf welche Räume sich die Operatornorm bezieht.

Häufig wird die Stetigkeit eines Operators ausgenutzt in folgender Weise:

$$\|L(x)\|_Y \leq \|L\| \|x\|_X, x \in X. \quad (3.9)$$

Folgerung 3.20 Seien $(X, \|\cdot\|_X), (Y, \|\cdot\|_Y), (Z, \|\cdot\|_Z)$ normierte Räume und seien $L : X \rightarrow Y, K : Y \rightarrow Z$ stetig. Dann ist auch $K \circ L$ stetig und es gilt

$$\|K \circ L\| \leq \|K\| \|L\|.$$

Beweis:

Einfach zu verifizieren unter Heranziehung von (3.9) oder der Folgenstetigkeit. $\blacksquare$

Im Spezialfall $X = Y = Z$ ergibt sich aus Folgerung 3.20, dass $\mathcal{B}(X)$ abgeschlossen ist gegenüber Addition, skalarer Multiplikation, Hintereinanderausführung (Multiplikation). $\mathcal{B}(X)$ wird daher auch als eine **Algebra** bezeichnet.

Satz 3.21 Seien $(X, \|\cdot\|_X), (Y, \|\cdot\|_Y)$ normierte Räume und sei X endlichdimensional. Dann ist jede lineare Abbildung $L : X \rightarrow Y$ stetig.

Beweis:

Wähle eine Basis $x^1, \dots, x^n$ in X und definiere eine Norm $\|\cdot\|_\sim$ in X durch

$$\|\cdot\|_\sim : X \ni x = \sum_{i=1}^n a_i x^i \mapsto \max_{1 \leq i \leq n} |a_i| \in \mathbb{R}.$$

Dann gilt für $x = \sum_{i=1}^n a_i x^i$:

$$\|L(x)\|_Y = \left\| \sum_{i=1}^n a_i L(x^i) \right\|_Y \leq n \cdot \max_{1 \leq i \leq n} \|L(x^i)\|_Y \|x\|_\sim \leq n \cdot d \cdot \max_{1 \leq i \leq n} \|L(x^i)\|_Y \|x\|_X$$

Dabei haben wir Satz 3.14 verwendet; die Konstante d ist daraus abgeleitet. Mit Satz 3.16 folgt nun die Behauptung. ■

Beispiel 3.22 Sei $X := C_\infty[0, 2\pi] := \{f : [0, 2\pi] \rightarrow \mathbb{R} \mid f \text{ unendlich oft differenzierbar}\}$. Auf X betrachte die Norm

$$\|\cdot\| : X \ni f \mapsto \int_0^{2\pi} |f(t)| dt \in \mathbb{R}$$

und betrachte dazu die lineare Abbildung (Ableitung)

$$D : X \ni f \mapsto f' \in X.$$

Setze

$$f_n(t) := \frac{1}{n} \cos(nt), t \in [0, 2\pi], n \in \mathbb{N}.$$

Es gilt $\|f_n\|_1 = \frac{2\pi}{n}, n \in \mathbb{N}$. Also ist $(f_n)_{n \in \mathbb{N}}$ eine Nullfolge in X . Aber es ist

$$\|Df_n\|_1 = 4, n \in \mathbb{N}.$$

Somit ist D nicht stetig. □

3.4 Stetige lineare Funktionale

Ist X ein Vektorraum über dem Skalkörper $\mathbb{K}$, so können wir den **algebraischen Dualraum**

$$X' := \{\lambda : X \rightarrow \mathbb{K} \mid \lambda \text{ linear}\}$$

eingeführen. Ist $(X, \|\cdot\|)$ ein normierter Raum mit Skalkörper $\mathbb{K}$, so können wir auch den Vektorraum X^* der stetigen linearen Abbildungen von X in $\mathbb{K}$ betrachten:

$$X^* := \{\lambda : X \rightarrow \mathbb{K} \mid \lambda \text{ linear und stetig}\}.$$

Für die Anwendung eines Funktionals $\lambda \in X^*$ auf ein Element $x \in X$ benutzen wir folgende

Schreibweise: Für $\lambda(x)$ schreiben wir meist: $\lambda(x) = \langle \lambda, x \rangle_{X^*, X} = \langle \lambda, x \rangle$.

Wir werden später im Kapitel über Hilberträume sehen, dass diese Wahl der Schreibweise einen „tieferen Sinn“ ergibt. Die Abbildung

$$\langle \cdot, \cdot \rangle : X^* \times X \ni (\lambda, x) \mapsto \lambda(x) \in \mathbb{K}$$

nennen wir die **duale Paarung**.

Folgerung 3.23 Sei X ein normierter Raum über dem Skalkörper $\mathbb{K}$. Dann gilt:

(a) X^* ist, versehen mit der Norm

$$\|\lambda\|_* := \sup\{|\langle \lambda, x \rangle| \mid x \in X, \|x\| \leq 1\}, \lambda \in X^*,$$

ein Banachraum.

(b) Die duale Paarung $\langle \cdot, \cdot \rangle : X^* \times X \longrightarrow \mathbb{K}$ ist linear in jedem Argument, also bilinear, und stetig.

Beweis:

Da $\mathbb{K}$ ein vollständiger normierter Raum ist, folgt die Aussage in a) aus Satz 3.19. Die Linearitätsaussagen sind offensichtlich. Zur Stetigkeit: Seien $\lambda, \lambda' \in X^*, x, x' \in X$. Dann liest man die Stetigkeit aus

$$\begin{aligned} |\langle \lambda, x \rangle - \langle \lambda', x' \rangle| &= |\langle \lambda, x - x' \rangle + \langle \lambda - \lambda', x' \rangle| \\ &\leq |\langle \lambda, x - x' \rangle| + |\langle \lambda - \lambda', x' \rangle| \\ &\leq \|\lambda\|_* \|x - x'\| + \|\lambda - \lambda'\|_* \|x'\| \end{aligned}$$

ab. ■

Definition 3.24 Sei $(X, \|\cdot\|)$ ein normierter Raum über dem Skalkörper $\mathbb{K}$. Der Banachraum $(X^*, \|\cdot\|_*)$ heißt der (topologische) **Dualraum** von $(X, \|\cdot\|)$. □

Mit X^* werden wir uns im Kapitel 5 genauer beschäftigen. Dort werden wir insbesondere die Reichhaltigkeit von X^* untersuchen. Hier bestimmen wir für einige konkrete Beispiele den Dualraum, die Reichhaltigkeit ist daraus ablesbar. Zur Erinnerung: In der Linearen Algebra hat man die Reichhaltigkeit von X' im Fall $\dim X = n$ schon konstruktiv erledigt: Es gibt zu einer Basis $x^1, \dots, x^n$ von X eine **duale Basis** $\lambda_1, \dots, \lambda_n$ von X' :

$$\langle \lambda_i, x^j \rangle = \delta_{i,j}, \quad i, j = 1, \dots, n.$$

Wir wissen aus der Linearen Algebra auch, dass im Fall $\dim X < \infty$ eine lineare bijektive Abbildung von $X \longrightarrow X'$ existiert. Da in diesem Fall $X' = X^*$ gilt (siehe Satz 3.21), ist also X^* vermöge einer linearen, bijektiven Abbildung gleichgesetzt mit X . Diese Abbildung ist wiederum stetig, ebenso wie seine Inverse. Für den unendlichdimensionalen Fall benötigen wir:

Definition 3.25 Zwei Banachräume $(X, \|\cdot\|_X), (Y, \|\cdot\|_Y)$ heißen **isomorph** (in Zeichen $\cong$), wenn es eine lineare bijektive Abbildung $T : X \longrightarrow Y$ und Konstanten m, M gibt mit

$$m\|x\|_X \leq \|Tx\|_Y \leq M\|x\|_X, \quad x \in X. \quad (3.10)$$

□

Die linke Ungleichung in (3.10) beschreibt die Stetigkeit von T , die rechte die Stetigkeit von T^{-1} . Die obige Bemerkung über den Dualraum von endlichdimensionalen Räumen besagt, dass $(\mathbb{R}^n)^* \cong \mathbb{R}^n$ gilt. Ein Isomorphismus zwischen $\mathbb{R}^n$ und $(\mathbb{R}^n)^*$ ist gegeben durch

$$S : \mathbb{R}^n \ni x = (x_1, \dots, x_n) \longmapsto Sx \in (\mathbb{R}^n)^*, \quad \langle Sx, y \rangle = \sum_{i=1}^n x_i y_i \quad (y = (y_1, \dots, y_n)).$$

Die duale Paarung ist daran schon zu erkennen und sie ist gegeben durch

$$\langle \lambda, x \rangle = \sum_{i=1}^n \lambda_i x_i \in \mathbb{R}, \quad \lambda = (\lambda_1, \dots, \lambda_n), \quad x = (x_1, \dots, x_n) \in \mathbb{R}^n.$$

Wir wollen nun zeigen, dass der Dualraum der reellen Folgenräume $(l_p, \|\cdot\|_p)$ für $1 \leq p < \infty$ durch $(l_q, \|\cdot\|_q)$ mit $1/p + 1/q = 1$ dargestellt wird, in dem Sinne, dass es einen Isomorphismus von l_q auf l_p^* gibt.

Sei also $1 \leq p < \infty$ und $1/p + 1/q = 1$. Wir definieren die Abbildung (duale Paarung)

$$T : l_q \longrightarrow l_p^*, \langle T((y_k)_{k \in \mathbb{N}}), (x_k)_{k \in \mathbb{N}} \rangle := \sum_{k=1}^{\infty} y_k x_k.$$

Offensichtlich ist T linear. Mit der Hölderschen Ungleichung folgt

$$|\langle T((y_k)_{k \in \mathbb{N}}), (x_k)_{k \in \mathbb{N}} \rangle| = \left| \sum_{k=1}^{\infty} y_k x_k \right| \leq \|(y_k)_{k \in \mathbb{N}}\|_q \|(x_k)_{k \in \mathbb{N}}\|_p,$$

woraus wir

$$\|T((y_k)_{k \in \mathbb{N}})\|_{l_p^*} \leq \|(y_k)_{k \in \mathbb{N}}\|_q \quad (3.11)$$

ablesen. Es gilt in (3.11) sogar Gleichheit. Wir zeigen dies in einer Fallunterscheidung.

1. Fall: $q = \infty$

Es gilt

$$|\langle T((y_k)_{k \in \mathbb{N}}), e^l \rangle| = |y_l| \leq \|(y_k)_{k \in \mathbb{N}}\|_{\infty},$$

also $\|(y_k)_{k \in \mathbb{N}}\|_{\infty} \leq \|T((y_k)_{k \in \mathbb{N}})\|_{l_p^*}$.

2. Fall: $1 < q < \infty$

Betrachte die Folge $(x_k)_{k \in \mathbb{N}}$ mit

$$x_k := \begin{cases} \frac{|y_k|}{y_k} & \text{falls } y_k \neq 0 \\ 0 & \text{sonst} \end{cases}$$

Dann gilt

$$\|(x_k)_{k \in \mathbb{N}}\|_p^p = \sum_{k=1}^{\infty} |x_k|^p = \sum_{k=1}^{\infty} |y_k|^{(q-1)p} = \sum_{k=1}^{\infty} |y_k|^q = \|(y_k)_{k \in \mathbb{N}}\|_q^q,$$

also $(x_k)_{k \in \mathbb{N}} \in l_p$. Ferner gilt

$$\begin{aligned} |\langle T((y_k)_{k \in \mathbb{N}}), (x_k)_{k \in \mathbb{N}} \rangle| &= \sum_{k=1}^{\infty} |y_k|^q = \|(y_k)_{k \in \mathbb{N}}\|_q \|(y_k)_{k \in \mathbb{N}}\|_q^{q-1} \\ &= \|(y_k)_{k \in \mathbb{N}}\|_q \|(x_k)_{k \in \mathbb{N}}\|_p^{p(q-1)/q} = \|(y_k)_{k \in \mathbb{N}}\|_q \|(x_k)_{k \in \mathbb{N}}\|_p \end{aligned}$$

Daraus folgt $\|(y_k)_{k \in \mathbb{N}}\|_{\infty} \leq \|T((y_k)_{k \in \mathbb{N}})\|_{l_p^*}$.

Damit haben wir nun gezeigt, dass T stetig ist mit $\|T\| = 1$. T ist also eine Isometrie. Wir zeigen nun durch erneute Fallunterscheidung, dass T sogar surjektiv ist. Sei $\lambda \in l_p^*$. Wir definieren $(y_k)_{k \in \mathbb{N}}$ durch $y_k := \langle \lambda, e^k \rangle, k \in \mathbb{N}$.

1. Fall: $q = \infty$

Es gilt

$$|y_l| = |\langle \lambda, e^l \rangle| \leq \|\lambda\|_{l_1^*}, l \in \mathbb{N}.$$

Dies zeigt $(y_k)_{k \in \mathbb{N}} \in l_{\infty}$, $\|(y_k)_{k \in \mathbb{N}}\|_{\infty} \leq \|\lambda\|_{l_1^*}$.

2. Fall: $1 < q < \infty$

Betrachte die Folge $(x_k)_{k \in \mathbb{N}}$ mit

$$x_k := \begin{cases} \frac{|y_k|}{y_k} & \text{falls } y_k \neq 0 \\ 0 & \text{sonst} \end{cases}$$

und damit die Folgen $X^m, Y^m \in \mathbb{N}$, mit

$$X_k^m := \begin{cases} x_k & \text{falls } k \leq m \\ 0 & \text{sonst} \end{cases}, Y_k^m := \begin{cases} y_k & \text{falls } k \leq m \\ 0 & \text{sonst} \end{cases}, m \in \mathbb{N}.$$

Dann rechnet man nach

$$\|Z^m\|_p^p = \|Y^m\|_q^q, \|Y^m\|_q^q \leq \|\lambda\|_{l_p^*} \|X^m\|_p = \|\lambda\|_{l_p^*} \|Y^m\|_p^{q/p}, m \in \mathbb{N}.$$

Dies zeigt $\|Y^m\|_q \leq \|\lambda\|_{l_p^*}$ und Grenzübergang $m \rightarrow \infty$ liefert $(y_k)_{k \in \mathbb{N}} \in l_q, \|(y_k)_{k \in \mathbb{N}}\|_q \leq \|\lambda\|_{l_p^*}$. In beiden Fällen haben wir

$$\langle T((y_k)_{k \in \mathbb{N}}), e^l \rangle = y_l = \langle \lambda, e^l \rangle, l \in \mathbb{N}.$$

Damit stimmen die Funktionale $T((y_k)_{k \in \mathbb{N}})$ und λ auf der in l_q dichten Teilmenge $\{e^l | l \in \mathbb{N}\}$ überein und es folgt wegen der Stetigkeit der beiden Funktionale $\lambda = T((y_k)_{k \in \mathbb{N}})$.

Lemma 3.26 *Es gilt:*

(a) $l_p^* \cong l_q, 1 \leq p < \infty, 1/p + 1/q = 1$. Die duale Paarung in $l_q \times l_p$ ist gegeben durch

$$\langle (y_n)_{n \in \mathbb{N}}, (x_n)_{n \in \mathbb{N}} \rangle = \sum_{n \in \mathbb{N}} y_n x_n, (y_n)_{n \in \mathbb{N}} \in l_q, (x_n)_{n \in \mathbb{N}} \in l_p.$$

(b) $l_1 \subset l_\infty^*, l_\infty^* \not\cong l_1$.

(c) $l_1 \cong c_0^*$. Die duale Paarung in $l_1 \times c_0$ ist gegeben durch

$$\langle (y_n)_{n \in \mathbb{N}}, (x_n)_{n \in \mathbb{N}} \rangle = \sum_{n \in \mathbb{N}} y_n x_n, (y_n)_{n \in \mathbb{N}} \in l_1, (x_n)_{n \in \mathbb{N}} \in c_0.$$

(d) l_p ist separabel für $1 \leq p < \infty$, l_∞ ist nicht separabel.

Beweis:

a) haben wir oben abgeleitet, b) und c) geben wir ohne Beweis an. Die Separabilität von $l_p, 1 \leq p < \infty$, zeigen wir im nächsten Abschnitt. Hier ist der Beweis zur Tatsache, dass l_∞ nicht separabel ist.

Zu $A \subset \mathbb{N}$ definiere $(x^A = (x_k^A)_{k \in \mathbb{N}})$ durch $x_k^A = 1$, falls $k \in A$, $x_k^A = 0$, sonst. Für $A, B \subset \mathbb{N}, A \neq B$ unterscheiden sich x^A, x^B in mindestens einem Folgenglied; also $\|x^A - x^B\|_\infty = 1$. Da die Potenzmenge von $\mathbb{N}$ überabzählbar ist, ist $(B_{\frac{1}{2}}(x^A)_{A \subset \mathbb{N}})$ eine Familie von überabzählbaren disjunkten Kugeln in l_∞ . Eine dichte Teilmenge $\bar{D}$ von l_∞ muss mit jeder dieser Kugeln einen nicht-leeren Schnitt besitzen, kann also nicht abzählbar sein. ■

Wir führen nun kursorisch die so genannten **Lebesgue-Räume** ein. Dazu ist zuerst der Begriff der Messbarkeit in $\mathbb{R}^n$ zu erklären. Eine Menge $A \subset \mathbb{R}^n$ heißt **messbar**, wenn sie zur σ -Algebra der **Borel-messbaren** Mengen gehört. Diese σ -Algebra ist die kleinste σ -Algebra, die alle Quader in $\mathbb{R}^n$ enthält. Dabei ist eine σ -Algebra eine Teilmenge der Potenzmenge von $\mathbb{R}^n$, die $\mathbb{R}^n$ enthält und abgeschlossen ist bezüglich Komplementbildung und abzählbarer disjunkter Vereinigung. Allen diesen Borel-messbaren Mengen ordnet man nun ein **Maß**, das so genannte **Borel-Maß**, zu, das den Inhalt der Quader erweitert. Dieses Maß ist additiv auf disjunkten Vereinigungen.

Zu diesen Borel-messbaren Mengen nimmt man nun alle Mengen A hinzu, die „zwischen zwei Borel-messbaren Mengen liegen“ ($B_1 \subset A \subset B_2$), welche dasselbe Borel-Maß haben; ihr **Lebesgue-Maß** ist dann das Borel-Maß von B_2 . Die Mengen, für die das Lebesgue-Maß auf diese Weise definiert ist, heißen **Lebesgue-messbar**. Durch „Produktbildung“ erhält man Messbarkeit und das Lebesguemaß auch auf $\mathbb{K} = \mathbb{C}$.

Sei nun $\Omega \subset \mathbb{R}^n$ eine Lebesgue-messbare Menge mit positivem Lebesgue-Maß. Jede Funktion $f: \Omega \rightarrow \mathbb{K}$, deren Urbilder messbar sind für alle messbaren Mengen in $\mathbb{K}$. Für jede solche messbare Funktion f wird für jedes $1 \leq p < \infty$ durch

$$\|f\|_p := \left(\int_{\Omega} |f(x)|^p dx \right)^{1/p}$$

ein Kandidat für eine Halbnorm erklärt. Den Wert $\|f\|_p = \infty$ „schließen“ wir aus, indem wir setzen:

$$\mathcal{L}_p(\Omega) := \{f : \Omega \rightarrow \mathbb{K} \mid f \text{ messbar}, \|f\|_p < \infty\}.$$

Aus der Theorie der messbaren Mengen folgt, dass $\|f\|_p = 0$ genau dann gilt, wenn $\{x \in \Omega \mid f(x) = \theta\}$ eine Menge mit Lebesgue-Maß Null ist. $\|\cdot\|_p$ ist eine Halbnorm auf $\mathcal{L}_p(\Omega)$. Die Dreiecksungleichung folgt wiederum aus einer entsprechenden **Hölder-Ungleichung**:

$$\|fg\|_1 \leq \|f\|_p \|g\|_q, \quad f \in \mathcal{L}_p(\Omega), g \in \mathcal{L}_q(\Omega).$$

Dabei ist wieder $1/p + 1/q = 1$. Um zu einer Norm zu kommen, bilden wir Äquivalenzklassen und identifizieren die Funktionen, die sich nur auf einer Menge vom Lebesgue-Maß unterscheiden, also

$$f \sim g \iff \|f - g\|_p = 0.$$

Nach dieser Identifikation fassen wir die „neuen Objekte“ zusammen in $L_p(\Omega)$ und verwenden $\|f\|_p$ auch für diese Objekte. Die Höldersche Ungleichung überträgt sich auch auf die Objekte in $L_p(\Omega)$. Damit ist $(L_p(\Omega), \|\cdot\|_p)$ ein Banachraum. Ist $\Omega \subset \mathbb{R}^n$ offen, dann ist der Raum

$$C_0^\infty(\Omega) := \{f : \Omega \rightarrow \mathbb{R} \mid f \text{ stetig}, \exists K \subset \Omega (K \text{ kompakt}, f(x) = 0, x \notin K)\}$$

dicht in $L_p(\Omega)$.

Kommen wir nun zur Einführung von $L_\infty(\Omega)$. Zur Messbarkeit ist nichts Neues anzumerken. Für jede messbare Funktion f wird durch

$$\|f\|_\infty := \text{ess sup } |f(x)| := \inf_N \sup_{x \in \Omega \setminus N} |f(x)|$$

ein Kandidat für eine Halbnorm erklärt. Den Wert $\|f\|_\infty = \infty$ „schließen“ wir aus, indem wir setzen:

$$\mathcal{L}_\infty(\Omega) := \{f : \Omega \rightarrow \mathbb{K} \mid f \text{ messbar}, \|f\|_\infty < \infty\}.$$

$\|\cdot\|_\infty$ ist eine Halbnorm auf $\mathcal{L}_\infty(\Omega)$. Die Dreiecksungleichung ist hier einfach zu beweisen. Um zu einer Norm zu kommen, bilden wir wieder Äquivalenzklassen und identifizieren die Funktionen, die sich nur auf einer Menge vom Lebesgue-Maß unterscheiden. Nach dieser Identifikation fassen wir die „neuen Objekte“ zusammen in $L_\infty(\Omega)$ und verwenden $\|f\|_\infty$ auch für diese Objekte. Damit ist $(L_\infty(\Omega), \|\cdot\|_\infty)$ ein Banachraum. Achtung: Der Raum enthält keine dichte abzählbare Teilmenge, ist also nicht separabel.

Lemma 3.27 *Es gilt:*

- (a) $(L_p(\Omega), \|\cdot\|_p)$ sind Banachräume für alle $1 \leq p \leq \infty$.
- (b) $L_p(\Omega) \subset L_q(\Omega), 1 \leq p \leq \infty$.
- (c) $L_p(\Omega)^* = L_q(\Omega), 1 \leq p < \infty, 1/p + 1/q = 1$. Die Dualitätsabbildung ist gegeben durch

$$\langle \cdot, \cdot \rangle : L_q(\Omega) \times L_p(\Omega) \ni (\lambda, f) \mapsto \int_\Omega \langle \lambda(x), f(x) \rangle_2 dx \in \mathbb{R}.$$

- (d) $L_1(\Omega) \subset L_\infty(\Omega)^*, L_1(\Omega) \neq L_\infty(\Omega)^*$.

Beweis:

Zum Beweis verweisen wir auf die obigen Überlegungen und die Literatur. ■

3.5 Basen

Der einfachste Basisbegriff in Banachräumen ist der vom Vektorraum geerbte Begriff der Hamelschen oder algebraischen Basis. In endlichdimensionalen Räumen ist die algebraische Basis bekanntlich ein fundamentaler und fruchtbarer Begriff. Mit Hilfe des Zornschen Lemmas² beweist man, dass jeder Vektorraum eine solche Basis besitzt. Solche Basen sind sehr „groß“, was ja nicht überraschend ist, da immer mit endlich vielen Elementen aus der Hamelbasis jedes Element erzeugbar sein muss. Ferner beziehen Hamelbasen keine Konvergenzeigenschaften aus dem zugrundeliegenden Banachraum ein.

Satz 3.28 *Jeder Banachraum hat eine Hamelbasis, die entweder endlich oder überabzählbar unendlich ist*

Beweis:

Sei X ein Banachraum und $\{e^i | i \in \mathbb{N}\}$ eine Familie mit $X = \cup_{n \in \mathbb{N}} U_n$, wobei für $n \in \mathbb{N}$ mit U_n die lineare Hülle von $e^1, e^2, \dots$ bezeichnet sei; wenn $\{e^i | i \in \mathbb{N}\}$ eine Hamelbasis ist, trifft dies zu. Dann ist nach Satz 3.15 jeder Teilraum U_n abgeschlossen. Nach dem Satz von Baire (siehe Satz 2.40) gibt es ein $N \in \mathbb{N}$, so dass das Innere von U_N nicht-leer ist. Also gibt es ein $r > 0, x^0 \in U_N$ mit $B_r(x^0) \subset U_N$. Da U_N ein linearer Raum ist, gilt auch $B_r \subset U_N$. Dann gilt aber sogar $X \subset U_N$, wie man mit skalarer Multiplikation (Stauchung) einsehen kann, und die Hamelbasis muss endlich sein. ■

Im Fall von unendlichdimensionalen Banach- und insbesondere Hilberträumen (siehe Kapitel 7) erweisen sich andere Basisbegriffe als natürlicher und tragfähiger. Der quasi natürliche Basisbegriff in unendlichdimensionalen Banachräumen ist der der Schauderbasis. Im Unterschied zur Hamelschen Basis, die die Darstellbarkeit jedes Vektors durch eine Linearkombination zusichert, werden beim Konzept der Schauderbasis unendliche Linearkombinationen im Sinn von Darstellungen in unendlichen Reihen zugelassen. Eine bewährte Quelle für Resultate über Basen sind [13, 25, 31]. Eine schöne Zusammenstellung ist auch [6], die im nächsten Kapitel insbesondere auch benutzt wird im Zusammenhang mit **greedy-Algorithmen**.

Zunächst zu unterschiedlichen Konvergenzbegriffen bei Reihen.

Definition 3.29 *Sei $(X, \|\cdot\|)$ ein normierter Raum und $(x^i)_{i \in \mathbb{N}}$ eine Folge.*

- (a) *Die Reihe $\sum_{i=1}^{\infty} x^i$ heißt **konvergent**, wenn die Folge $(s^n)_{n \in \mathbb{N}}$ der Partialsummen ($s^n := \sum_{i=1}^n x^i, n \in \mathbb{N}$) konvergiert. Der Wert s der Reihe ist der Grenzwert der Folge der Partialsummen, also $s := \lim_n s^n$; wir schreiben kurz $\sum_{i=1}^{\infty} x^i$ auch für den Reihenwert.*
- (b) *Die Reihe $\sum_{i=1}^{\infty} x^i$ heißt **absolut konvergent**, wenn die Reihe $\sum_{i=1}^{\infty} \|x^i\|$ konvergiert.*
- (c) *Die konvergente Reihe $\sum_{i=1}^{\infty} x^i$ heißt **unbedingt konvergent**, wenn die Reihe für jede Permutation σ der natürlichen Zahlen ($\sigma : \mathbb{N} \rightarrow \mathbb{N}$ bijektiv) die (umgeordnete) Reihe $\sum_{i=1}^{\infty} x^{\sigma(i)}$ konvergiert.*

□

Die Vollständigkeit eines normierten Raumes $(X, \|\cdot\|)$ kann man mit Reihen beschreiben. Es sind nämlich äquivalent:

²Lemma von Zorn: Sei $(A, \leq)$ eine nichtleere partiell geordnete Menge, so dass jede total geordnete Teilmenge $\mathcal{N}$ (d. h. für alle $n_1, n_2 \in \mathcal{N}$ gilt $n_1 \leq n_2$ oder $n_2 \leq n_1$) eine obere Schranke besitzt (d. h. es gibt ein $m \in M$ mit $n \leq m$ für alle $n \in \mathcal{N}$). Dann besitzt M ein maximales Element, d. h. es gibt ein $m_0 \in M$, so dass für alle $m \in M$ gilt: $m_0 \leq m \implies m \leq m_0$

- (a) $(X, \|\cdot\|)$ ist ein Banachraum.
- (b) Jede absolut konvergente Reihe in X konvergiert.

Die Implikation $a) \implies b)$ ist ganz einfach. Zum Beweis der Implikation $b) \implies a)$ konstruiert man aus einer Cauchyfolge $(x_k)_{k \in \mathbb{N}}$ in X eine Teilfolge $(y^j := x^{j+1} - x^j)_{j \in \mathbb{N}}$, so dass die Reihe $\sum_{j=1}^{\infty} \|y^j\|$ konvergiert.

Wir wissen aus der Analysis, dass nicht jede konvergente Reihe auch absolut konvergiert (harmonische und alternierende harmonische Reihe!). Aus der Analysis wissen wir auch, dass Umordnungen einer Reihe, wenn sie nicht absolut konvergiert, zu einem desaströsen Ergebnis führen kann: jede reelle Zahl kann als Reihenwert erreicht werden. Dank des Riemannschen Umordnungssatzes weiß man, dass in $\mathbb{R}$ absolute Konvergenz und unbedingte Konvergenz äquivalent sind. In unendlichdimensionalen Räumen gilt dieser Satz nicht: unbedingte und absolute Konvergenz fallen auseinander, denn es gibt stets eine Reihe, die unbedingt, aber nicht absolut konvergiert; siehe [11].

Definition 3.30 Sei X ein Banachraum. Eine Folge $(e^i)_{i \in \mathbb{N}}$ in X heißt **Schauderbasis** in X , wenn es zu jedem $x \in X$ eine eindeutig bestimmte Folge von Skalaren $(\alpha_i)_{i \in \mathbb{N}}$ gibt mit

$$x = \sum_{i=1}^{\infty} \alpha_i e^i.$$

□

Beachte, dass jedes Element e^i einer Schauderbasis nicht der Nullvektor sein kann, da Eindeutigkeit einer Darstellung gefordert ist. Ferner sind je endlich viele Elemente $e^{i_1}, \dots, e^{i_n}$ in einer Schauderbasis linear unabhängig sind, wiederum wegen der Eindeutigkeit einer Darstellung. Damit ist klar, dass der Begriff der Schauderbasis eine Weiterentwicklung der algebraischen Basis in unendlichdimensionalen Räumen ist. In der Definition über eine Reihenentwicklung ist die Reihenfolge der Basiselemente wichtig: die unbedingte Konvergenz der Reihenentwicklung ist hier noch nicht verlangt; eine Permutation einer Schauderbasis muss also keine Schauderbasis sein.

Definition 3.31 Ist $B = (e^i)_{i \in \mathbb{N}}$ eine Schauderbasis des Banachraums X , so heißt die Folge von Funktionalen $(\lambda_n)_{n \in \mathbb{N}}$, die gegeben ist durch

$$\lambda_n : X \ni x = \sum_{i=1}^{\infty} a_i e^i \longmapsto a_n \in \mathbb{K}, n \in \mathbb{N},$$

die zu B gehörende Folge von **Koeffizientenfunktionalen**.

□

Die Koeffizientenfunktionale sind offensichtlich linear. Ferner gilt: $\lambda_n(e^k) = \delta_{n,k}, n, k \in \mathbb{N}$.

Jede Schauderbasis $B = (e^i)_{i \in \mathbb{N}}$ in einem Banachraum $(X, \|\cdot\|)$ gibt Anlass zu einer Familie von **kanonischen Projektionen**:

$$S_n : X \ni x \longmapsto \sum_{k=1}^n \lambda_k(x) e^k \in X, n \in \mathbb{N}.$$

Offensichtlich sind die Abbildungen P_n linear. Ihre Stetigkeit ist nicht elementar zu beweisen, ebenso wie die Stetigkeit der Koeffizientenfunktionale λ_n . Wir werden die Stetigkeit beider Objekte im nächsten Kapitel beweisen, wenn wir einige starke Hilfsmittel zur Verfügung haben.

Ihre Bedeutung finden die Projektionen u. a. beim Entwurf von *Greedy-Algorithmen*. Dies sind Algorithmen, die gierig die größten Werte der Koeffizientenfunktionale nutzen wollen. Dabei ist

die Umordnung von Darstellungen durch eine Schauderbasis nötig. Die Werte der Koeffizientenfunktionale können als vergleichbar zu den Fourierkoeffizienten angesehen werden.

Die Projektionen, ihre Bildbereiche sind ja endlichdimensional, können als Indiz genommen werden, dass Banachräume mit Schauderbasen als „strukturell einfach“ angesehen werden können; wir können sie uns ja sogar aus eindimensionalen Teilräumen zusammengesetzt vorstellen. Die Tatsache, dass es Banachräume gibt – wir lernen solche kennen – die keine Schauderbasis besitzen – Hilberträume besitzen eine solche immer – lassen vermuten, dass Banachräume auch strukturell sehr komplex sein können. Dazu später.

Folgerung 3.32 *Sei X ein Banachraum mit einer Schauderbasis in X . Dann ist X separabel, d. h. es gibt eine abzählbare dichte Teilmenge in X .*

Beweis:

Sei $(e^i)_{i \in \mathbb{N}}$ eine Schauderbasis in X . Sei $\mathbb{C}_{\mathbb{Q}} := \{a = u + iv \in \mathbb{C} \mid u, v \in \mathbb{Q}\}$ und sei $A := \bigcup_{n \in \mathbb{N}} A_n$ mit $A_n := \{x = \sum_{i=1}^n a_i e^i \mid a_1, \dots, a_n \in \mathbb{C}_{\mathbb{Q}}\}$. Offenbar ist A abzählbar. Wir zeigen, dass A auch dicht in X ist. Sei $x \in X$, $x = \sum_{i=1}^{\infty} a_i e^i$, und sei $\varepsilon > 0$. Dann gibt es $N \in \mathbb{N}$ mit

$$\left\| \sum_{i=1}^n a_i e^i - x \right\| < \varepsilon \text{ für alle } n \geq N.$$

Wähle zu $a_1, \dots, a_N \in \mathbb{C}$ $b_1, \dots, b_N \in \mathbb{C}_{\mathbb{Q}}$ mit

$$\left\| \sum_{i=1}^N a_i e^i - \sum_{i=1}^N b_i e^i \right\| < \varepsilon,$$

was wegen der Dichte von $\mathbb{Q}$ in $\mathbb{R}$ möglich ist. Mit der Dreiecksungleichung folgt für das Element $z := \sum_{i=1}^N b_i e^i \in A_N$ die Abschätzung $\|z - x\| < 2\varepsilon$. ■

Beispiel 3.33 *Der Folgenraum l_p hat die Basis*

$$e^1 := (1, 0, 0, 0, \dots), e^2 := (0, 1, 0, 0, \dots), e^3 := (0, 0, 1, 0, \dots), \dots,$$

unabhängig von der Wahl $p \in [1, \infty)$. Auch der Raum c_0 der Nullfolgen hat diese Basis. Die Räume als lineare Hülle sind verschieden, die Basis nicht!

Betrachte die Reihe $\sum_{k=1}^{\infty} \frac{1}{k} e^k$ in l_2 . Diese Reihe ist offensichtlich nicht absolut konvergent. Sie ist aber unbedingt konvergent. Dazu ziehe man c) aus Lemma 3.36 heran.

Der Raum l_{∞} besitzt keine Schauderbasis, denn er ist nicht separabel; siehe Lemma 3.26. Dass

$$e^1 := (1, 0, 0, 0, \dots), e^2 := (0, 1, 0, 0, \dots), e^3 := (0, 0, 1, 0, \dots), \dots,$$

keine Schauderbasis von l_{∞} darstellt, kann man aber auch direkt sehen: Das Element $x = (1, 1, \dots) \in l_{\infty}$ kann durch $e^1, e^2, \dots$ nur so hingeschrieben werden:

$$x = \sum_{i=1}^{\infty} e^i$$

Diese Reihe ist aber offenbar in l_{∞} nicht konvergent, da alle „Reste“ der Partialsummen Norm 1 besitzen.

Für den Raum c der konvergenten Folgen benötigen wir noch einen zusätzlichen Vektor, um die Grenzwerte der Elemente in c zu berücksichtigen. Man kann hierfür irgendein Element in $c \setminus c_0$ auswählen, etwa $e^0 := (1, 1, 1, \dots)$. Stellt man nun ein $x \in c$ in der Gestalt

$$x = \sum_{i=0}^{\infty} a_i e^i$$

dar, so ist der zu e^0 gehörende Koeffizient a_0 gerade der Grenzwert der Folge $x = (a_1, a_2, a_3, \dots)$. □

Beispiel 3.34 Wir wissen aus dem Approximationssatz von Weierstraß, dass die Polynome dicht im Raum $C[a, b]$ sind. Dennoch bilden sie hier keine Schauderbasis. Man sieht dies mit einem Argument aus dem Bereich der Potenzreihen ein. $\square$

Beispiel 3.35 Ein herausragendes Beispiel einer Schauderbasis ist das Haar-System. Es ist eine Basis im Raum der Lebesgue-integrierbaren Funktion auf dem kompakten Intervall $[0, 1]$, also in $L_1[0, 1]$. Diese Basis dient als Ausgangspunkt für die vielfältigen Basen der **wavelets**; siehe etwa [7]. $\square$

Lemma 3.36 Sei $(X, \|\cdot\|)$ ein Banachraum und sei $(x^i)_{i \in \mathbb{N}}$ eine Folge in X . Es sind äquivalent:

- (a) Die Reihe $\sum_{i=1}^{\infty} x^{\sigma(i)}$ konvergiert für jede Permutation σ von $\mathbb{N}$.
- (b) Die Reihe $\sum_{i=1}^{\infty} e^{n_i}$ konvergiert für jede Folge $(n_i)_{i \in \mathbb{N}}$ natürlicher Zahlen $n_1 < n_2 < \dots$.
- (c) Die Reihe $\sum_{i=1}^{\infty} \vartheta_i x^i$ konvergiert für jede Folge $(\vartheta_i)_{i \in \mathbb{N}}$ in $\{-1, +1\}$.
- (d) Für jedes $\varepsilon > 0$ gibt es ein $N \in \mathbb{N}$, so dass für jede endliche Menge $\eta \subset \mathbb{N}$ mit $\min_{i \in \eta} i > N$ stets $\|\sum_{i \in \eta} x^i\| < \varepsilon$ gilt.

Beweis:

b) $\implies$ c)

Sei $(\vartheta_i)_{i \in \mathbb{N}}$ eine Folge in $\{-1, +1\}$. Für $i \in \mathbb{N}$ setze $\vartheta_i^+ := \max\{\vartheta_i, 0\}$ und $\vartheta_i^- := \max\{-\vartheta_i, 0\}$; also $\vartheta_i = \vartheta_i^+ - \vartheta_i^-$. Betrachte die Reihen $\sum_{i=1}^{\infty} \vartheta_i^+ x^i$ und $\sum_{i=1}^{\infty} \vartheta_i^- x^i$. Die Indizes der nicht-verschwindenden Reihenglieder steigen streng monoton, weshalb beide Reihen nach b) konvergieren; daher konvergiert auch ihre Differenz und daher auch $\sum_{i=1}^{\infty} \vartheta_i x^i$.

c) $\implies$ a)

Sei $n_1 < n_2 < \dots$. Für $i \in \mathbb{N}$ setze $\vartheta_i := +1$, falls $i \in \{n_k | k \in \mathbb{N}\}$, -1 , sonst. Nach c) konvergiert die Reihe $\sum_{i=1}^{\infty} \vartheta_i x^i$ ebenso wie auch die Reihe $\sum_{i=1}^{\infty} x^i$, und damit auch die Reihe

$$\sum_{i=1}^{\infty} x^{n_i} = \frac{1}{2} \left(\sum_{i=1}^{\infty} \vartheta_i x^i + \sum_{i=1}^{\infty} x^i \right).$$

d) $\implies$ b)

Sei zu $\varepsilon > 0$ ein $N \in \mathbb{N}$ nach d) gewählt und seien $n_1 < n_2 < \dots$. Sei $m \geq N$ und $k \in \mathbb{N}$. Setze $\eta := \{n_m, n_{m+1}, \dots, n_{m+k}\}$. Wegen $N \leq m \leq n_m$ ist $\min_{i \in \eta} i \geq N$ und somit

$$\left\| \sum_{i \in \eta} x^i \right\| = \left\| \sum_{i=m}^{m+k} x^{n_i} \right\| < \varepsilon.$$

Die Reihe $\sum_{i=1}^{\infty} x^{n_i}$ ist also eine Reihe, deren Partialsummen eine Cauchyfolge bilden. Sie ist daher konvergent.

d) $\implies$ a)

Sei zu $\varepsilon > 0$ ein $N \in \mathbb{N}$ nach d) gewählt und sei σ eine Permutation von $\mathbb{N}$. Sei $M \in \mathbb{N}$ so gewählt, dass $\sigma(i) > N$ für $i \geq M$. Dann erfüllt für $m \geq \max\{M, N\}$ und $k \in \mathbb{N}$ die Menge $\eta := \{\sigma(m), \dots, \sigma(m+k)\}$ die Bedingung in d). Daher gilt

$$\left\| \sum_{i=m}^{m+k} x^{\sigma(i)} \right\| = \left\| \sum_{i \in \eta} x^i \right\| < \varepsilon,$$

d. h. die Reihe $\sum_{i=1}^{\infty} x^{\sigma(i)}$ ist eine Reihe, deren Partialsummen eine Cauchyfolge ist, und daher konvergent.

b) $\implies$ d)

Angenommen, d) gelte nicht. Dann gibt es also ein $\varepsilon > 0$, so dass für alle $N \in \mathbb{N}$ eine endliche Menge η_N mit $\min_{i \in \eta_N} i > N$ existiert, so dass $\|\sum_{i \in \eta_N} x^i\| \geq \varepsilon$. Die Mengen η_N können so gewählt werden, dass

$$q_N := \max_{i \in \eta_N} i < \min_{i \in \eta_{N+1}} i =: p_{N+1}$$

gilt, weshalb $\eta := \cup_{N \in \mathbb{N}} \eta_N$ eine unendliche Teilmenge von $\mathbb{N}$ bildet, die wir in natürlicher Reihenfolge geordnet als Folge $(n_i)_{i \in \mathbb{N}}$ betrachten. Nach Konstruktion gibt es zu jedem $N \in \mathbb{N}$ natürliche Zahlen $m_0 > N$ und k_0 , so dass $\eta_N = \{n_{m_0}, \dots, n_{m_0+k_0}\}$ und daher

$$\left\| \sum_{i=m_0}^{m_0+k_0} x^{n_i} \right\| = \left\| \sum_{i \in \eta_N} x^i \right\| \geq \varepsilon.$$

Also kann die Reihe $\sum_{i=1}^{\infty} x^{n_i}$ keine konvergente Reihe sein.

a) $\implies$ d)

Sei σ eine Permutation von $\mathbb{N}$, die für jedes $N \in \mathbb{N}$ die Menge $\{p_{N+1}, p_{N+2}, \dots, q_N\}$ auf sich selbst abbildet, so dass $\sigma^{-1}(\eta_N) = \{p_N, p_{N+1}, \dots, p_N + k_N\}$, wobei $k_N = \#\eta_N$ ist. Dann ist

$$\left\| \sum_{i=p_N}^{p_N+k_N} x^{\sigma(i)} \right\| = \left\| \sum_{i \in \eta_N} x^i \right\| \geq \varepsilon.$$

Also kann die Reihe $\sum_{i=1}^{\infty} x^{n_i}$ keine konvergente Reihe sein. ■

Satz 3.37 *Konvergiert eine Reihe $\sum_{i=1}^{\infty} x^i$ in einem Banachraum X unbedingt, so ist der Wert der Reihe $\sum_{i=1}^{\infty} x^{\sigma(i)}$ unabhängig von der Permutation σ der natürlichen Zahlen $\mathbb{N}$.*

Beweis:

Sei die Reihe $\sum_{i=1}^{\infty} x^i$ unbedingt konvergent in X mit Reihenwert $s \in X$. Annahme: Es gibt eine Permutation σ der natürlichen Zahlen $\mathbb{N}$, so dass der Wert der Reihe $s' := \sum_{i=1}^{\infty} x^{\sigma(i)}$ verschieden von s ist; also $s \neq s'$. Der Satz von Hahn-Banach 5.5 – wir beweisen ihn im Kapitel 5 über Dualität – erlaubt die Wahl eines Funktional $\lambda \in X^*$ mit $\langle \lambda, s \rangle \neq \langle \lambda, s' \rangle$. Dann konvergiert die Reihe $\sum_{i=1}^{\infty} \langle \lambda, x^{\sigma(i)} \rangle$ nicht unbedingt, denn die Permutation ändert ja den Reihenwert. Der Riemannsche Umordnungssatz impliziert die Existenz einer Permutation σ^* , für die die Reihe $\sum_{i=1}^{\infty} \langle \lambda, x^{\sigma^*(i)} \rangle$ divergiert. Dann divergiert auch die Reihe $\sum_{i=1}^{\infty} x^{\sigma^*(i)}$ in Widerspruch zur vorausgesetzten unbedingten Konvergenz. ■

Wir wissen schon, dass nicht jeder Banachraum eine Schauderbasis besitzt. Das Beispiel, das wir kennen, ist l_{∞} , ein nicht separabler Raum. Nach [12] gibt es aber sogar separable Banachräume, die keine Schauderbasis besitzen. In dieser „Lücke“ ist ein Ergebnis von Mazur angesiedelt, dass eine erste Antwort auf die ursprünglich von Banach (~ 1930) formulierten Fragen zu Schauderbasisen darstellt.

Satz 3.38 (Mazur) *Jeder unendlichdimensionale Banachraum enthält einen abgeschlossenen unendlichdimensionalen Teilraum, der eine Schauderbasis besitzt.*

Beweis:

Für den Beweis siehe etwa [31]. ■

Im Allgemeinen konvergiert die Entwicklung eines Elements bezüglich einer Schauderbasis nicht unabhängig von der Summationsreihenfolge. Diese Unabhängigkeit von der Reihenfolge ist aber für viele Anwendungen von großer Bedeutung. Es lohnt sich daher, Basen zu betrachten, die eine unbedingte Konvergenz jeder Reihenentwicklung zulassen.

Definition 3.39 Sei X ein Banachraum und sei $B = (e^i)_{i \in \mathbb{N}}$ eine Schauderbasis. Diese Basis heißt **unbedingte Basis**, wenn für jedes $x \in X$ die Entwicklung $\sum_{i \in \mathbb{N}} a_i(x) e^i$ bezüglich der Basis unbedingt konvergiert, d. h. für jede Umordnung gegen x konvergiert. $\square$

Der Begriff der unbedingten Basis ist offenbar stärker als der der Schauderbasis. Es ist daher klar, dass nicht jeder Banachraum, der eine Schauderbasis besitzt, eine unbedingte Basis besitzt. Ein Beispiel für dieses Phänomen ist der Raum $L_1[0, 1]$ der Lebesgue-integrierbaren Funktionen auf $[0, 1]$. Im Gegensatz zum Satz von Mazur 3.38 hat die Frage, ob in jedem unendlichdimensionalen Banachraum ein unendlichdimensionaler Teilraum existiert, der eine unbedingte Schauderbasis besitzt, eine negative Antwort; siehe dazu [16, 17].

Definition 3.40 Sei X ein Vektorraum.

- (a) Zwei lineare Teilräume X_0, X_1 heißen **algebraisch komplementär**, falls gilt: $X = X_0 + X_1, X_0 \cap X_1 = \{\theta\}$. Ist dies der Fall, dann heißt X_1 ein **algebraisches Komplement** von X_0 und wir schreiben $X = X_0 \oplus X_1$.
- (b) Ist X ein normierter Raum und sind X_0, X_1 algebraisch komplementäre Teilräume von X , so heißen sie **topologisch komplementär**, wenn die (zugehörige) Projektion von X auf X_0 entlang X_1

$$P : X = X_0 \oplus X_1 \ni x_0 + x_1 \longmapsto x_0 \in X_0$$

stetig ist. $\square$

$\square$

In einem Hilbertraum besitzt jeder abgeschlossene Unterraum ein (sogar orthogonales) Komplement; siehe Kapitel 7. Die bereits von Banach aufgeworfene Frage, ob jeder abgeschlossene Unterraum in einem Banachraum ein topologisches Komplement besitzt, konnte erst 1973 beantwortet werden, nämlich negativ; siehe [12]. Inzwischen sind viele Gegenbeispiele bekannt, etwa besitzt der Unterraum c_0 der Nullfolgen des Raumes l_∞ kein Komplement; siehe [28]. Nach einem Resultat von J. Lindenstrauß und L. Tzafriri muss ein Banachraum, in dem alle abgeschlossenen Unterräume komplementiert sind, zu einem Hilbertraum isomorph sein; siehe [25]. Nach T. Gowers und B. Maurey ([17]) gibt es einen unendlichdimensionalen Banachraum X , in dem aus $X = V \oplus W$ stets $\dim V < \infty$ oder $\dim W < \infty$ folgt. Insgesamt hat sich gezeigt, dass Banachräume also eine sehr von der Einfachheit der Hilberträume abweichende Komplexität haben können.

Als kleine Kuriosität sei folgendes Resultat angefügt. Der Satz von Bernstein-Schröder aus der Mengenlehre besagt:

Ist eine Menge M zu einer Teilmenge der Menge N gleichmächtig und ist N zu einer Teilmenge von M gleichmächtig, so besitzen M und N dieselbe Mächtigkeit.

Eine Übertragung auf die Banachräume könnte auf das folgende Problem hinauslaufen:

Sind X und Y Banachräume und ist X zu einem Teilraum von Y isomorph sowie Y zu einem Teilraum von X isomorph, sind dann X und Y isomorph?

Die Antwort ist NEIN! Siehe [36] hierzu und zu sonstigen „Kuriositäten“.

3.6 Übungen

- 1.) Sei $X := \mathbb{R}$ versehen mit der euklidischen Norm $|\cdot|$. Zeige: Sind $A, B \subset X$ abgeschlossen und disjunkt ($A \cap B = \emptyset$), so gibt es offene Mengen $\mathcal{U}, V \subset X$ mit

$$A \subset \mathcal{U}, B \subset V, \mathcal{U} \cap V = \emptyset.$$

(Bemerkung: Wir haben damit gezeigt, dass der euklidische Raum $(\mathbb{R}, |\cdot|)$ normal ist.)

- 2.) Sei $(X, \|\cdot\|)$ ein Banachraum und sei $F \subset X$ ein linearer Teilraum. Dann ist auch $(\overline{F}, \|\cdot\|)$ ein Banachraum.
- 3.) Sei $(X, \|\cdot\|)$ ein normierter Raum und seien $a, b \in [0, \infty)$. Zeige:

$$\{x + y \in X \mid x, y \in X, \|x\| = a, \|y\| = b\} = \{z \in X \mid |a - b| \leq \|z\| \leq a + b\}.$$

- 4.) Sei $(X, \|\cdot\|)$ ein normierter Raum und sei $F \subset X$ ein abgeschlossener linearer Teilraum.
- (a) Zeige, dass durch

$$x \sim y : \Longleftrightarrow x - y \in F$$

auf X eine Äquivalenzrelation erklärt wird.

- (b) Durch $\|\cdot\|_{\sim} : X/\sim \ni [X] \longmapsto \inf\{\|x + z\| \mid z \in F\} \in \mathbb{R}$ wird eine Norm auf $X/\sim$ definiert.

- 5.) Eine abgeschlossene Teilmenge M von $\mathbb{R}^n$ (versehen mit der euklidischen Norm) heißt diskret, wenn gilt:

$$\forall x \in M \exists r > 0 (B_r(x) \cap M = \{x\}).$$

- (a) Zeige: Ist M diskret und beschränkt, so ist M eine endliche Menge.
- (b) Gilt die Behauptung aus a) auch ohne die Voraussetzung „beschränkt“?

- 6.) Für m sei $\mathbb{R}^{m,m}$ der Vektorraum der $m \times m$ Matrizen $A = (a_{ij})_{1 \leq i \leq m, 1 \leq j \leq m}$. Zeige: Durch

$$\|A\| := \left(\sum_{i=1}^m \sum_{j=1}^m |a_{ij}|^2 \right)^{1/2}$$

wird auf $\mathbb{R}^{m,m}$ eine Norm definiert, die folgende multiplikative Eigenschaft besitzt:

$$\|AB\| \leq \|A\| \cdot \|B\|$$

- 7.) Sei $(X, \|\cdot\|)$ ein reeller normierter Raum. Es gelte die Parallelogrammidentität, d. h. es gelte

$$\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2) \text{ für alle } x, y \in X.$$

Zeige: Durch

$$\langle x, y \rangle := \frac{1}{4}(\|x + y\|^2 - \|x - y\|^2)$$

wird ein Skalarprodukt auf X erklärt.

- 8.) Sei $(X, \|\cdot\|)$ ein normierter Raum und sei $(\alpha_n)_{n \in \mathbb{N}}$ irgendeine Folge reeller Zahlen mit $\alpha_n \geq 1, n \in \mathbb{N}$. Zeige die Äquivalenz von

- (a) X ist Banachraum
- (b) Ist $(x_n)_{n \in \mathbb{N}}$ eine Folge in X , für die die Reihe $\sum_{n=1}^{\infty} \alpha_n \|x_n\|$ konvergiert, so konvergiert auch die Reihe $\sum_{n=1}^{\infty} x_n$.

- 9.) Sei $m_n : \mathbb{R} \ni t \mapsto t^n \in \mathbb{R}$ das Monom n -ten Grades, sei $\mathcal{P}_n$ der reelle Vektorraum, der von den Monomen $m_0, \dots, m_n$ aufgespannt wird. Setze $X := \cup_{n \in \mathbb{N}_0} \mathcal{P}_n$. Zeige: X ist ein reeller Vektorraum, in dem durch

$$\|p\| := \max_{0 \leq i \leq n} |a_i| \text{ falls } p(t) = \sum_{i=0}^n a_i t^i$$

eine Norm erklärt ist.

- 10.) Die Bezeichnungen seien wie in der vorangehenden Aufgabe gewählt. Zeige:

- (a) Die Abbildung

$$T : X \longrightarrow X, (Tp)(t) := a_0 + \sum_{i=1}^n \frac{a_i}{i} t^i \text{ falls } p(t) = \sum_{i=0}^n a_i t^i$$

ist linear und stetig.

- (b) T ist bijektiv.

- (c) T^{-1} ist linear aber nicht stetig.

- 11.) Sei $\|\cdot\|$ eine der Normen $|\cdot|_1, |\cdot|_2, |\cdot|_\infty$ auf $\mathbb{R}^2$. Sei $K := \{(x, y) \in \mathbb{R}^2 \mid x - y = 4\}$. Zeige:

- (a) $a := \inf\{\|v\| \mid v \in K\} \leq 4$.

- (b) Es gibt ein $u \in K$ mit $\|u\| = a$.

- (c) Berechne ein $u \in K$ mit $\|u\| = a$.

- 12.) Sei $(X, \|\cdot\|)$ ein normierter Raum und sei $U \subset X$ ein linearer Teilraum. Zeige: $\overline{U}$ ist ein linearer Teilraum von X .

- 13.) Sei $(X, \|\cdot\|)$ ein normierter Raum und sei $U \subset X$ ein linearer Teilraum. Sei $T : U \longrightarrow \mathbb{R}$ stetig und linear. Zeige: Die stetige Fortsetzung $\overline{T}$ von T auf $\overline{U}$ ist ebenfalls linear.

$$(\overline{T}(x) := \lim_{u \rightarrow x, u \in U} T(u), x \in \overline{U}).$$

- 14.) Sei $E := \{(x, y) \in \mathbb{R}^2 \mid x^2 + xy + 4y^2 \leq 1\}$. Finde eine Norm $\|\cdot\|$ in $\mathbb{R}^2$, so dass E in dieser Norm die Kugel um den Nullpunkt mit Radius 2 ist.

- 15.) Sei $1 \leq p < \infty$. Man betrachte die Shiftoperatoren $L, R : l_p(\mathbb{C}) \longrightarrow l_p(\mathbb{C})$. Sind definiert durch $L((x_1, x_2, \dots)) = (x_2, x_3, \dots)$ und $R((x_1, x_2, \dots)) = (0, x_1, x_2, \dots)$. Zeige, dass L, R stetig sind und berechne jeweils die Operatornorm.

- 16.) Betrachte den Banachraum $(C[0, 1], \|\cdot\|_\infty)$. Zeige, dass die Polynome keine Schauderbasis bilden. Hinweis betrachte etwa die Funktion $x \mapsto 1$.

- 17.) Betrachte die Folge $(\ln(n+1)^{-1})_{n \in \mathbb{N}}$ liegt in c_0 , aber in keinem $l_p, 1 \leq p < \infty$.

- 18.) Betrachte auf $C[0, 1]$ die Norm $\|f\|_\infty := \sup_{t \in [0, 1]} |f(t)|$ der gleichmäßigen Konvergenz. Damit definieren wir auf $C^1[0, 1] := \{f : [0, 1] \longrightarrow \mathbb{R} \mid f \text{ stetig differenzierbar}\}$ die Abbildung

$$C^1[0, 1] \ni f \mapsto \|f\|_{1, \infty} := |f(0)| + \|f'\|_\infty \in \mathbb{R}.$$

Zeige:

- (a) $(C^1[0, 1], \|\cdot\|_{1, \infty})$ ist ein Banachraum.

- (b) Die Abbildung

$$T : C[0, 1] \ni g \mapsto Tg \in C^1[0, 1], Tg(x) := \int_0^x g(s) ds,$$

ist linear, stetig und injektiv, aber nicht surjektiv.

- 19.) Sei $(X, \|\cdot\|)$ ein normierter Raum und sei $(\alpha_n)_{n \in \mathbb{N}}$ irgendeine Folge reeller Zahlen mit $\alpha_n \geq 1, n \in \mathbb{N}$. Zeige die Äquivalenz von
- $(X, \|\cdot\|)$ ist Banachraum
 - Ist $(x_n)_{n \in \mathbb{N}}$ eine Folge in X , für die die Reihe $\sum_{n=1}^{\infty} \alpha_n \|x_n\|$ konvergiert, so konvergiert auch die Reihe $\sum_{n=1}^{\infty} x_n$.
- 20.) Betrachte normierte Räume $(X, \|\cdot\|_X), (Y, \|\cdot\|_Y)$, dazu eine Folge $(T_n)_{n \in \mathbb{N}}$ in $\mathcal{B}(X, Y)$, $T \in \mathcal{B}(X, Y)$ und eine Folge $(x_n)_{n \in \mathbb{N}}$ in X .
- Zeige:
Konvergiert $(x_n)_{n \in \mathbb{N}}$ in X gegen z und konvergiert $(T_n)_{n \in \mathbb{N}}$ in $\mathcal{B}(X, Y)$ gegen T (Operatorkonvergenz), so konvergiert $(T_n x_n)_{n \in \mathbb{N}}$ in Y gegen Tz .
 - Zeige: Die folgende Abänderung von a) gilt nicht!
Konvergiert $(x_n)_{n \in \mathbb{N}}$ in X gegen z und konvergiert $(T_n x)_{n \in \mathbb{N}}$ in Y für alle $x \in X$ (starke Konvergenz), so konvergiert $(T_n x_n)_{n \in \mathbb{N}}$ in Y gegen Tz .
- 21.) Betrachte $c := \{x = (x_k)_{k \in \mathbb{N}} \in l_{\infty} \mid \lim_k x_k \text{ existiert}\}$ als Teilraum des reellen Banachraumes $(l_{\infty}, \|\cdot\|)$ und darin die Elemente
- $$f^0 := (1, 1, 1, \dots), f^1 := (0, 1, 1, \dots), f^2 := (0, 0, 1, 1, \dots), f^3 := (0, 0, 0, 1, 1, \dots), \dots$$
- Zeige: $(f^i)_{i \in \mathbb{N}_0}$ ist eine Schauderbasis in c .
 - Berechne $\|\sum_{i=0}^N f^i\|_{\infty}, \|\sum_{i=0}^N (-1)^i f^i\|_{\infty}$.
 - Ist die Basis eine unbedingte Basis?
- 22.) Illustration einiger ungewohnter Effekte in unendlichdimensionalen Räumen.
- Finde eine lineare, surjektive Abbildung $T \in \mathcal{B}(X)$ mit einem Banachraum X , so dass $\ker(T) := \{x \in X \mid Tx = \theta\} \neq \{\theta\}$.
 - Finde eine lineare, injektive Abbildung $T \in \mathcal{B}(X)$ mit einem Banachraum X , so dass $\text{ran}(T) := \{Tx \in X \mid x \in X\} \neq X$.
 - Sei $\mathcal{P}$ der Raum der Polynome, betrachtet als Teilraum von $C[-2, 2]$. Seien $m_n := x^n, n \in \mathbb{N}$, die Monome in $\mathcal{P}$. Ist die lineare Abbildung T , definiert auf der Basis der Monome durch $m_n \mapsto 3^n m_n, n \in \mathbb{N}$, stetig?
- 23.) Sei $1 \leq p < q < \infty$. Zeige:
- $\|x\|_q \leq \|x\|_p$ für alle $x \in l_p$;
 - $l_p \subset l_q$.
- 24.) Sei $l_b := \{(x_n)_{n \in \mathbb{N}} \mid \sum_{n \in \mathbb{N}} |x_{n+1} - x_n| < \infty\}$. Zeige: Durch

$$\|x\|_b := |x_1| + \sum_{n \in \mathbb{N}} |x_{n+1} - x_n|, x = (x_n)_{n \in \mathbb{N}} \in l_b,$$

wird eine Norm auf l_b erzeugt. Liegt in $(l_b, \|\cdot\|_b)$ sogar ein Banachraum vor?

- 25.) Sei $(X, \|\cdot\|)$ ein normierter Raum. Zeige die Äquivalenz von:
- Es gibt eine endliche oder abzählbare Menge $A \subset X$ mit $\overline{A} = X$.
 - Es gibt eine abzählbare Menge B mit $\overline{\text{span}(B)} = X$.
- 26.) Für $f \in C^1[0, 1]$ setze:

$$\|f\|_{(1)} := |f(0)| + \sup_{t \in [0, 1]} |f'(t)|, \|f\|_{(2)} := \max \left\{ \left| \int_0^1 f(t) dt \right|, \sup_{t \in [0, 1]} |f'(t)| \right\}.$$

Zeige: $\|\cdot\|_{(i)}$ ist jeweils eine Norm. Sind die Normen äquivalent?

27.) Sei $(x_n)_{n \in \mathbb{N}}$ eine Folge in einem Banachraum X . Es gelte: $\sum_{k \in \mathbb{N}} \|x_k\| \leq M < \infty$. Zeige, dass die Reihe $\sum_{k \in \mathbb{N}} x_k$ gegen ein x mit $\|x\| \leq M$ konvergiert.

28.) Betrachte die Abbildung

$$T : c \longrightarrow c_0, (x_n)_{n \in \mathbb{N}} \longmapsto (y_n)_{n \in \mathbb{N}} \text{ mit } y_n := \begin{cases} \lim_k x_k & , \text{ falls } n = 1 \\ x_n - \lim_k x_k & , \text{ sonst} \end{cases}.$$

(a) Zeige: T ist bijektiv und stetig.

(b) Berechne $\|T\|, \|T^{-1}\|$.

Kapitel 4

Prinzipien/Hauptsätze der Funktionalanalysis

Wenn ich weiter geblickt habe, so deshalb, weil ich auf den Schultern von Riesen stehe

I. Newton

Hier bauen wir das Studium stetiger linearer Operatoren aus. Wir beweisen Hauptsätze der (linearen) Funktionalanalysis, die in Anwendungen von überragender Bedeutung sind. Hierher gehören eigentlich auch der Satz von Baire und der Satz von Hahn-Banach. Wir hielten es aber für zweckmäßig, den ersten Satz schon früher und den zweiten Satz anschließend aufzuführen.

4.1 Der Satz von Banach–Steinhaus

Satz 4.1 (Banach–Steinhaus (1927)/Hahn (1927)) Sei $(X, \|\cdot\|_X)$ ein Banachraum, sei $(Y, \|\cdot\|_Y)$ ein normierter Raum, und sei $(T_i)_{i \in I}$ eine Familie stetiger Operatoren von X nach Y . Es gelte:

$$\forall x \in X \exists c_x \geq 0 \forall i \in I (\|T_i x\| \leq c_x) \quad (4.1)$$

Dann gibt es $c \geq 0$ mit

$$\|T_i\|_{X \rightarrow Y} \leq c \text{ für alle } i \in I. \quad (4.2)$$

Beweis:

Wir beweisen zunächst:

$$\exists z \in X \exists r > 0 \exists d > 0 \forall x \in B_r(z) \forall i \in I (\|T_i x\|_Y \leq d) \quad (4.3)$$

Sei dazu für $k \in \mathbb{N}$

$$A_k := \{x \in X \mid \|T_i x\|_Y \leq k \text{ für alle } i \in I\} = \bigcap_{i \in I} \{x \in X \mid \|T_i x\|_Y \leq k\}.$$

Jedes A_k ist abgeschlossen, da jedes T_i stetig ist. Nach Voraussetzung liegt jedes $x \in X$ in einem A_k , also

$$X = \bigcup_{k \in \mathbb{N}} A_k.$$

Nach dem Satz von Baire – siehe Satz 2.41 – gilt:

$$\exists k_0 \in \mathbb{N} \exists z \in X \exists r > 0 (B_r(z) \subset A_{k_0}).$$

(4.3) gilt also mit $d := k_0$. Wegen (4.3) gilt für alle $x \in B_r, i \in I$

$$\|T_i x\|_Y = \|T_i(x+z) - T_i(z)\|_Y \leq \|T_i(x+z)\|_Y + \|T_i(z)\|_Y \leq d + c_z.$$

Hieraus folgt für alle $x \in B_1$

$$\|T_i x\|_Y \leq r^{-1}(d + c_z) =: c, \text{ d. h. } \|T_i\|_{X \rightarrow Y} \leq c, i \in I.$$

■

Hier ist ein **Beweis** des obigen Satzes von Banach-Steinhaus, der eine „Elementarisierung“ des ursprünglichen Beweises von Hahn und Banach darstellt; siehe [32]. Er ist indirekt.

Annahme: $\sup_{i \in I} \|T_i\|_{X \rightarrow Y} = \infty$. Dann gibt es eine Folge $(T_n)_{n \in \mathbb{N}}$ in der Familie $(T_i)_{i \in I}$ mit $\|T_n\|_{X \rightarrow Y} \geq 4^n, n \in \mathbb{N}$. Ausgehend von $x^0 = \theta$ konstruieren wir mit dem Hilfsresultat aus Lemma 4.2 induktiv eine Folge $(x^n)_{n \in \mathbb{N}}$ mit

$$\|x^n - x^{n-1}\|_X \leq 3^{-n}, \frac{3}{2} \|T_n x^n\|_Y \geq 3^{-n} \|T_n\|_{X \rightarrow Y}.$$

Dann folgt für $m, n \in \mathbb{N}, m > n$, mit der geometrischen Reihe

$$\|x^m - x^n\|_X \leq \|x^m - x^{m-1}\|_X + \dots + \|x^{n+1} - x^n\|_X \leq (3^{-m} + \dots + 3^{-(n+1)}) \leq 3^{-(n+1)} \frac{3}{2}. \quad (4.4)$$

Also ist $(x^n)_{n \in \mathbb{N}}$ eine Cauchyfolge und daher konvergent. Sei $x = \lim_m x^m$. Dann folgt durch Grenzübergang $m \rightarrow \infty$ in (4.4)

$$\|x - x^n\|_X \leq \frac{1}{2} 3^{-n}, n \in \mathbb{N}.$$

Nun gilt für alle $n \in \mathbb{N}$

$$\begin{aligned} \|T_n x\|_Y &\geq \|T_n x^n\|_Y - \|T_n x - T_n x^n\|_Y \geq \frac{2}{3} 3^{-n} \|T_n\|_{X \rightarrow Y} - \|T_n\|_{X \rightarrow Y} \|x - x^n\|_X \\ &\geq \frac{2}{3} 3^{-n} \|T_n\|_{X \rightarrow Y} - \frac{1}{2} 3^{-n} \|T_n\|_{X \rightarrow Y} = \frac{1}{6} 3^{-n} \|T_n\|_{X \rightarrow Y} \geq \frac{1}{6} \left(\frac{3}{4}\right)^n, \end{aligned}$$

also $\lim_n \|T_n x\|_Y = \infty$, was ein Widerspruch zur Voraussetzung (4.1) ist.

Lemma 4.2 Seien $(X, \|\cdot\|_X), (Y, \|\cdot\|_Y)$ normierte Räume, $T \in \mathcal{B}(X, Y)$. Dann gilt:

$$\forall x \in X \forall r > 0 \left(\sup_{z \in \overline{B}_r(x)} \|Tz\|_Y \geq r \|T\|_{X \rightarrow Y} \right) \quad (4.5)$$

Beweis:

Sei $x \in X, r > 0, h \in X$. Dann gilt unter Verwendung der Dreiecksungleichung

$$\|Th\|_Y \leq \frac{1}{2} (\|T(x+h)\|_Y + \|T(x-h)h\|_Y) \leq \max(\|T(x+h)\|_Y, \|T(x-h)h\|_Y).$$

Also folgt durch Supremumsbildung $h \in \overline{B}_1$

$$r \|T\|_{X \rightarrow Y} = \sup_{h \in \overline{B}_r} \|Th\|_Y \leq \sup_{z \in \overline{B}_r(x)} \|Tz\|_Y.$$

■

Definition 4.3 Seien X, Y normierte Räume und sei $(T_n)_{n \in \mathbb{N}}$ eine Folge stetiger linearer Operatoren von X nach Y . Wir sagen, $(T_n)_{n \in \mathbb{N}}$ **konvergiert stark** gegen T , wenn $\lim_n T_n x = Tx$ für alle $x \in X$ gilt. Wir schreiben diese Konvergenz als $T_n \xrightarrow{s} T$. $\square$

Beachte, dass die starke Konvergenz etwas anderes ist als die Konvergenz der Operatoren der Norm nach im Raum der stetigen linearen Operatoren. Klar, starke Konvergenz folgt aus der Konvergenz der Operatornorm nach.

Folgerung 4.4 Seien X, Y normierte Räume, sei $(T_n)_{n \in \mathbb{N}}$ eine Folge stetiger linearer Operatoren von X nach Y sei $T : X \rightarrow Y$ stetig und linear, und es gelte $T_n \xrightarrow{s} T$. Dann gilt $\|T\| \leq \liminf_n \|T_n\|$.

Beweis:

Sei $(T_{n_k})_{k \in \mathbb{N}}$ eine Teilfolge mit $\liminf_n \|T_n\| = \lim_k \|T_{n_k}\|$. Für $x \in X$ gilt dann $\|Tx\| = \lim_k \|T_{n_k}x\|$, $\|T_{n_k}x\| \leq \|T_{n_k}\| \|x\|$ für alle $k \in \mathbb{N}$, Nun folgt $\|Tx\| \leq \liminf_k (\|T_{n_k}\| \|x\|) = \lim_k \|T_{n_k}\| \|x\| = (\liminf_n \|T_n\|) \|x\|$. $\blacksquare$

Man kann die Folgerung 4.4 auch so lesen: Die Abbildung

$$f : \mathcal{B}(X, Y) \ni T \mapsto \|T\| \in \mathbb{R}$$

ist nach unten beschränkt und unterhalbstetig.

Folgerung 4.5 Seien X, Y, Z Banachräume, sei $(T_n)_{n \in \mathbb{N}}$ eine Folge stetiger linearer Operatoren von X nach Y , sei $(S_n)_{n \in \mathbb{N}}$ eine Folge stetiger linearer Operatoren von X nach Y und sei $T : X \rightarrow Y, S : Y \rightarrow Z$ stetig und linear. Gilt $T_n \xrightarrow{s} T, S_n \xrightarrow{s} S$, dann folgt $S_n \circ T_n \xrightarrow{s} S \circ T$.

Beweis:

Für jedes $y \in Y$ ist $(S_n y)_{n \in \mathbb{N}}$ konvergent also beschränkt. Nach Satz 4.1 gibt es $c \geq 0$ mit $\|S_n\| \leq c$ für alle $n \in \mathbb{N}$. Sei $x \in X$. Aus

$$\|S_n \circ T_n x - S \circ Tx\| \leq \|S_n \circ (T_n - T)x\| + \|(S_n - S) \circ Tx\| \leq c \|T_n x - Tx\| + \|S_n(Tx) - S(Tx)\|$$

liest man die Behauptung ab. $\blacksquare$

Satz 4.6 Seien X, Y Banachräume, sei $(T_n)_{n \in \mathbb{N}}$ eine Folge stetiger linearer Operatoren von X nach Y . Dann sind äquivalent:

- (a) Es gibt einen stetigen linearen Operator $T : X \rightarrow Y$ mit $T_n \xrightarrow{s} T$.
- (b) Es gilt $\sup_{n \in \mathbb{N}} \|T_n\| < \infty$ und es gibt eine in X dichte Teilmenge A , so dass $(T_n x)_{n \in \mathbb{N}}$ eine Cauchyfolge ist für alle $x \in A$.

Beweis:

(a) $\implies$ (b).

Wir können $A := X$ setzen und haben nur die Beschränktheit der Folge $(T_n)_{n \in \mathbb{N}}$ in $\mathcal{B}(X, Y)$ zu zeigen. Diese folgt wie für $(S_n)_{n \in \mathbb{N}}$ im Beweis zu Folgerung 4.5.

(b) $\implies$ (a).

Sei $c := \sup_{n \in \mathbb{N}} \|T_n\|$, sei $x \in X$ und $\varepsilon > 0$. Es gibt $y \in A$ mit $\|x - y\| < \frac{\varepsilon}{c+1}$ und es gibt $N \in \mathbb{N}$ mit $\|T_n y - T_m y\| < \varepsilon$ für alle $m, n \geq N$. Damit folgt für $m, n \geq N$

$$\|T_n x - T_m x\| \leq \|T_n(x - y)\| + \|(T_n - T_m)y\| + \|T_m(y - x)\| \leq c \|x - y\| + \|T_n y - T_m y\| + c \|x - y\| < 3\varepsilon.$$

Also ist $(T_n x)_{n \in \mathbb{N}}$ eine Cauchyfolge in Y und somit konvergent. Es ist daher sinnvoll, zu definieren:

$$T : X \longrightarrow Y, Tx := \lim_n T_n x, x \in X.$$

Offenbar ist T linear, ja sogar stetig, denn wir haben ja

$$\|Tx\| = \lim_n \|T_n x\| \leq c\|x\|, x \in X.$$

Nach Konstruktion gilt $T_n \xrightarrow{s} T$. ■

Beispiel 4.7 Als Anwendung für den Satz 4.6 skizzieren wir einen Konvergenzsatz für Quadraturformeln. Wir setzen $X := C[a, b]$, versehen mit der Supremumsnorm $\|\cdot\|_\infty$, und betrachten eine Folge $(Q_n)_{n \in \mathbb{N}}$ von **Quadraturformeln**:

$$Q_n x := \sum_{j=1}^{m_n} g_{n,j} x(t_{n,j}), x \in X, n \in \mathbb{N};$$

dabei sind $a \leq t_{n,1} < \dots < t_{n,m_n} \leq b$ **Stützstellen** und $g_{n,1}, \dots, g_{n,m_n}$ **Gewichte**. Jedes $Q_n x$ soll eine Approximation für $\int_a^b x(s) ds$ darstellen. Offenbar ist Q_n ein linearer stetiger Operator von X nach $\mathbb{R}$ für alle $n \in \mathbb{N}$. Uns interessiert die Frage nach der Konvergenz der näherungsweisen Integration mittels $(Q_n)_{n \in \mathbb{N}}$ also: Gilt $Q_n \xrightarrow{s} Q$ für $Qx := \int_a^b x(s) ds, x \in X$? Da nach dem Satz von Weierstraß gilt, dass die Menge $\mathcal{P}$ der Polynome auf $[a, b]$ in X dicht ist, d. h. dass $\overline{\mathcal{P}} = X$ gilt, ist nach Satz 4.6 hinreichend, dass wir nachweisen:

$$\|Q_n\| \leq c \text{ für alle } n \in \mathbb{N}, \lim_n Q_n p = Qp \text{ für alle } p \in \mathcal{P}.$$

Da $\|Q_n\| \leq \sum_{j=1}^{m_n} |g_{n,j}|$ für alle $n \in \mathbb{N}$ gilt, benötigen wir also eine Schranke für die l_1 -Norm der Gewichte. Wir erhalten diese, wenn gilt:

$$Q_n x = Qx \text{ für alle konstanten Funktionen } x \text{ und } g_{n,j} \geq 0, j = 1, \dots, m_n, n \in \mathbb{N},$$

denn dann haben wir

$$\|Q_n\| \leq \sum_{j=1}^{m_n} g_{n,j} = Q_n 1 = b - a.$$

Damit lässt sich etwa die Konvergenz aller **Gaußschen Quadraturformeln** absichern, bei denen die Gewichte nichtnegativ sind und die im „Grenzwert“ exakt für die Polynome sind. Erfasst wird etwa auch die **Simpsonsche Regel**:

$$Q_{2n} x := \frac{h}{3} (x(a) + 4x(t_1) + 2x(t_2) + \dots + 4x(t_{2n-1}) + x(b)),$$

wobei $t_{i+1} - t_i = h, 0 \leq i \leq 2k-1, h := \frac{b-a}{2n}, t_0 := a$. Hier ist allerdings die starke Konvergenz von $(Q_{2n})_{n \in \mathbb{N}}$ auf den Polynomen noch nachzurechnen; sie folgt aus der Tatsache, dass Konvergenz für alle $f \in C^4[a, b]$ vorliegt. Der Vorteil ist, dass wir damit sogar auf Konvergenz in $C[a, b]$ geschlossen werden kann. □

4.2 Der Satz von der offenen Abbildung

Definition 4.8 Seien X, Y Banachräume. $T \in \mathcal{B}(X, Y)$ heißt **offen**, falls $T(U)$ offen in Y ist für jede offene Menge in X . □

Lemma 4.9 Seien X, Y normierte Räume, sei $T : X \longrightarrow Y$ linear, stetig und surjektiv. Dann ist T eine offene Abbildung.

Beweis:

Abkürzung: $X_\varrho := B_\varrho$ in $X, Y_\varrho := B_\varrho$ in $Y; \varrho \geq 0$. Der Beweis besteht aus drei Schritten:

- (1) $\forall \varrho > 0 \exists \tau > 0 (Y_\tau \subset \overline{T(X_\varrho)})$
- (2) $\forall \varrho > 0 \exists \sigma > 0 (Y_\sigma \subset T(X_\varrho))$
- (3) Beweis der Behauptung.

Zu (1). Sei $\varrho > 0; \delta := \frac{1}{2}\varrho$. Es gilt $X = \cup_{n \in \mathbb{N}} X_{n\delta}$. Da T surjektiv ist, folgt:

$$Y = T(X) = \cup_{n \in \mathbb{N}} T(X_{n\delta}) = \subset \cup_{n \in \mathbb{N}} \overline{nT(X_\delta)}$$

Nach dem Satz von Baire (Satz 2.41) gibt es $n \in \mathbb{N}, y' \in Y, \tau' > 0$, mit: $B_{\tau'}(y') \subset \overline{nT(X_\delta)}$. Wegen $\overline{nT(X_\delta)} = n\overline{T(X_\delta)}$ gibt es $y \in Y, \tau'' > 0$, mit $B_{\tau''}(y) \subset \overline{T(X_\delta)}$. Aus $X_\delta - X_\delta \subset X_\varrho$ folgt $T(X_\delta) - T(X_\delta) \subset T(X_\varrho)$ und somit

$$U := \cup_{z \in B_{\tau''}(y)} (z - B_{\tau''}(y)) = B_{\tau''}(y) - B_{\tau''}(y) \subset \overline{T(X_\delta)} - \overline{T(X_\delta)} \subset \overline{T(X_\delta) - T(X_\delta)} \subset \overline{T(X_\varrho)}.$$

Da U offen ist und θ enthält, gibt es $\tau > 0$ mit $Y_\tau \subset U \subset \overline{T(X_\varrho)}$.

Zu (2): Sei $\varrho > 0; r_0 := \frac{\varrho}{2}$. Wähle $(r_n)_{n \in \mathbb{N}}$ so, dass gilt: $\sum_{n=1}^{\infty} r_n < r_0$; man kann etwa $r_n := r_0 3^{-n}$ wählen. Nach (1) haben wir:

$$\forall n \in \mathbb{N}_0 \exists \tau_n > 0 (Y_{\tau_n} \subset \overline{T(X_{r_n})}).$$

O.E. $\lim_n \tau_n = 0$. Sei $\sigma := \tau_0$ und $y \in Y_\sigma$. Dann ist $y \in \overline{T(X_{r_0})}$. Also gibt es $x_0 \in X_{r_0}$ mit $\|y - Tx_0\| < \tau_1$, d.h. $y - Tx_0 \in Y_{\tau_1}$. Dann ist $y - Tx_0 \in \overline{T(X_{r_1})}$ und somit existiert ein $x_1 \in X_{r_1}$ mit

$$\|y - Tx_0 - Tx_1\| < \tau_2, \text{ d.h. } y - Tx_0 - Tx_1 \in Y_{\tau_2}.$$

Dieser Prozess liefert eine Folge $(x_n)_{n \in \mathbb{N}}$ in X mit

$$x_n \in X_{r_n}, \|y - T(\sum_{l=0}^n x_l)\| < \tau_{n+1}, n \in \mathbb{N}_0.$$

Wegen $\|x_n\| < r_n, n \in \mathbb{N}, \sum_{n=0}^{\infty} r_n < \infty$, ist die Reihe $\sum_{l=0}^{\infty} x_l$ konvergent; sei x ihr Reihenwert. Dann gilt:

$$\|x\| \leq \sum_{n=0}^{\infty} \|x_n\| \leq r_0 + \sum_{n=1}^{\infty} r_n < 2r_0 = \varrho, \|y - Tx\| = \lim_n \|y - T(\sum_{l=0}^n x_l)\| \leq \lim_n \tau_n = 0.$$

Dies bedeutet $x \in X_\varrho, y = Tx \in T(X_\varrho)$.

Zu (3): Sei $U \subset X$ offen. Sei $y \in T(U)$, etwa $y = Tx, x \in U$. Da U offen ist, gibt es $\varrho > 0$ mit $x + X_\varrho \subset U$. Nach (2) gibt es $\sigma > 0$ mit $Y_\sigma \subset T(X_\varrho)$. Nun folgt

$$y + Y_\sigma = Tx + Y_\sigma \subset Tx + T(X_\varrho) = T(x + X_\varrho) \subset T(U).$$

Dies zeigt, dass $T(U)$ offen ist. ■

Satz 4.10 (Satz von der offenen Abbildung) Seien X, Y Banachräume und sei $T : X \longrightarrow Y$ linear und stetig. Dann sind äquivalent:

- (a) T ist offen.

(b) Es gibt $\delta > 0$ mit $B_\delta \subset T(B_1)$.

(c) T ist surjektiv.

Beweis:

(a) $\implies$ (b).

Da $\theta \in T(B_1)$ und da $T(B_1)$ nach (a) offen ist, folgt offensichtlich (b).

(b) $\implies$ (a).

Sei $U \subset X$ offen. Sei $y \in T(U)$, etwa $y = Tx, x \in U$. Da U offen ist, gibt es $\varrho > 0$ mit $x + B_\varrho \subset U$.

Mit (b) erhalten wir $\sigma > 0$ mit $B_\sigma \subset T(B_\varrho)$. Nun folgt

$$y + B_\sigma = Tx + B_\sigma \subset Tx + T(B_\varrho) = T(x + B_\varrho) \subset T(U).$$

Dies zeigt, dass $T(U)$ offen ist.

(a) $\implies$ (c). Klar bei Verwendung von (b).

(c) $\implies$ (a). Siehe Lemma 4.9. ■

Satz 4.11 (Satz von der stetigen Inversen (Banach)) Seien X, Y Banachräume und sei $T : X \longrightarrow Y$ linear, stetig und bijektiv. Dann ist T^{-1} stetig.

Beweis:

Nach Satz 4.10 ist $T = (T^{-1})^{-1}$ eine offene Abbildung, d. h. $(T^{-1})^{-1}(U)$ ist offen, falls U offen ist. Dies ist aber gerade die Stetigkeit von T^{-1} (siehe Satz 1.27). ■

Folgerung 4.12 Seien $(X, \|\cdot\|_X), (Y, \|\cdot\|_Y)$ Banachräume und sei $T : X \longrightarrow Y$ linear und stetig. Dann sind äquivalent:

(a) T ist injektiv und $\text{ran}(T)$ ist abgeschlossen.

(b) Es gibt $c > 0$ mit $\|Tx\| \geq c\|x\|$ für alle $x \in X$.

Beweis:

(a) $\implies$ (b). Sei $Z := \text{ran}(T)$. $(Z, \|\cdot\|_Y)$ ist ein Banachraum. Da $T : X \longrightarrow Z$ linear, stetig und bijektiv ist, ist T^{-1} stetig. Also gibt es $d > 0$ mit $\|T^{-1}z\|_X \leq d\|z\|_Y$ für alle $z \in Z = \text{ran}(T)$. Ersetze z durch Tx .

(b) $\implies$ (a). Wir haben

$$c\|x\| \leq \|Tx\| \leq d\|x\|, x \in X, \text{ mit } d := \|T\|.$$

Aus der linken Ungleichung folgt die Injektivität von T . Ist $(Tx^n)_{n \in \mathbb{N}}$ eine konvergente Folge in $\text{ran}(T)$, $y := \lim_n Tx^n$, so ist, wiederum nach der linken Ungleichung, $(x^n)_{n \in \mathbb{N}}$ eine Cauchyfolge in X , also konvergent. Sei $x = \lim_n x^n$. Aus der Stetigkeit von T folgt $y = \lim_n Tx^n = Tx$, also $y \in \text{ran}(T)$. ■

4.3 Der Satz vom abgeschlossenen Graphen

Sind $(X, \|\cdot\|_X), (Y, \|\cdot\|_Y)$ normierte Räume, so ist auch $X \times Y$ ein normierter Raum bzgl. der Norm

$$\|(x, y)\|_{X \times Y} := \|x\|_X + \|y\|_Y, (x, y) \in X \times Y.$$

Die Produkttopologie auf $X \times Y$ stimmt offenbar mit der durch die Norm $\|\cdot\|_{X \times Y}$ induzierten Topologie überein. Sind X, Y Banachräume, so ist offenbar auch $X \times Y$ ein Banachraum. In Anlehnung an reelle Funktionen geben wir die folgende Definition.

Definition 4.13 Seien X, Y normierte Räume, sei D ein linearer Teilraum in X und sei $T : D \rightarrow Y$ linear (D Definitionsmenge).

a) Die Menge $G(T) := \{(x, Tx) | x \in D\}$ heißt der **Graph von T** .

b) T heißt **abgeschlossen**, wenn $G(T)$ als Teilmenge von $X \times Y$ abgeschlossen ist.

□

Mit der Berücksichtigung eines Definitionsbereiches D als (möglicherweise echte) Teilmenge von X bereiten wir die Betrachtung unbeschränkter/nichtstetiger Operatoren vor.

Mit Folgen ausgedrückt bedeutet die Graphen-Abgeschlossenheit:

Aus $\lim_n x_n = x, \lim_n Tx_n = y$ folgt $x \in D$ und $y = Tx$.

Man beachte den Unterschied zur Stetigkeit von T :

Aus $\lim_n x_n = x \in D$ folgt $\lim_n T(x_n) = y$ und $y = Tx$.

Lemma 4.14 Seien X, Y normierte Räume, D linearer Teilraum von X und $T : D \rightarrow Y$ linear. Dann sind äquivalent:

i) T ist abgeschlossen.

ii) Ist $(x_n)_{n \in \mathbb{N}}$ eine Folge in D und ist $x = \lim_n x_n, y = \lim_n Tx_n$, so gilt $x \in D, y = Tx$.

Zusatz: Sind X, Y sogar Banachräume, so sind die obigen Aussagen äquivalent zu

iii) $(D, \|\cdot\|_T)$ ist ein Banachraum, wobei $\|\cdot\|_T : D \ni x \mapsto \|x\|_X + \|Tx\| \in \mathbb{R}$ die so genannte **Graphennorm** ist.

Beweis:

i) $\iff$ ii) ist klar, wir beweisen den Zusatz.

i) $\implies$ iii).

Klar, durch $\|\cdot\|_T$ ist eine Norm definiert. Sei $(x_n)_{n \in \mathbb{N}}$ eine Cauchyfolge in D bezüglich der Graphennorm. Dann ist $((x_n, Tx_n))_{n \in \mathbb{N}}$ eine Cauchyfolge in $X \times Y$, also $x_n \rightarrow x, Tx_n \rightarrow y$. Mit i) $x \in D, y = Tx$.

iii) $\implies$ i).

Sei $((x_n, Tx_n))_{n \in \mathbb{N}}$ mit $x_n \rightarrow x, Tx_n \rightarrow y$. Dann $((x_n, Tx_n))_{n \in \mathbb{N}}$ eine Cauchyfolge in $Z := G(T)$. Mit der Vollständigkeit von Z folgt $\lim_n x_n = x, \lim_n Tx_n = y \in T(X)$. Also $x \in D, y = Tx$. ■

Satz 4.15 (Satz vom abgeschlossenen Graphen) Seien X, Y Banachräume und sei $T : X \rightarrow Y$ linear. Dann sind äquivalent:

(a) T ist abgeschlossen.

(b) T ist stetig.

Beweis:

(a) $\implies$ (b).

Sei T abgeschlossen, also $G(T) := \{(x, Tx) | x \in X\}$ abgeschlossen in $X \times Y$; somit ist $G(T)$, versehen mit der Norm in $X \times Y$ ein Banachraum. Definiere

$$P : G(T) \ni (x, Tx) \mapsto x \in X.$$

Offenbar ist P linear. P ist stetig, denn $\|P(x, Tx)\| = \|x\| \leq \|x\| + \|Tx\| = \|(x, Tx)\|, x \in X$. P ist injektiv, denn aus $P((x, Tx)) = \theta$ folgt $x = \theta$ und daher $(x, Tx) = (\theta, \theta)$. P ist offenbar surjektiv. Nach Satz 4.10 ist P^{-1} stetig, also auch $T = QP^{-1}$, wobei $Q : X \times Y \ni (x, y) \mapsto y \in Y$.

(b) $\implies$ (a). Klar. ■

Beispiel 4.16 Wir geben ein Beispiel eines abgeschlossenen, linearen, nicht stetigen Operators. Sei $X := C[0, 1]$, versehen mit der Supremumsnorm $\|\cdot\|_\infty$. Sei $D := C^1[0, 1]$ und sei $T : X \supset D \ni x \mapsto x' \in X$. Offenbar ist T linear. T ist nicht stetig, da für $x_n(t) := t^n$ gilt;

$$\|x_n\|_\infty = 1, \|Tx_n\|_\infty = n, n \in \mathbb{N}.$$

T ist abgeschlossen, denn ist $(x_n)_{n \in \mathbb{N}}$ eine Folge in D mit $x_n \rightarrow x \in X, Tx_n \rightarrow y \in X$, so gilt $x_n \rightarrow x$ gleichmäßig, $x'_n \rightarrow y$ gleichmäßig. Daraus folgt: x ist differenzierbar, $x' = y, x \in D$. □

4.4 Approximationsprinzip: Projektionen

Unendlichdimensionale Banachräume dienen dazu, Probleme in der mathematischen Physik, in der Biomathematik, in der Finanzmathematik, ... zu modellieren und Existenz- und Eindeutigkeitsfragen von Lösungen zu studieren. In der Angewandten Mathematik sucht man dann in diesen Banachräumen Ansätze, diese Lösungen einer Berechnung zuzuführen. Dabei spielt die Approximationstechnik die entscheidende Rolle: komplexe, unendlichdimensionale Objekte sollen durch einfache, endlichdimensionale Objekte ersetzt werden und dann auf die Approximationsgüte untersucht werden. Das Maß, das dafür in erster Linie zur Verfügung steht, ist die Norm im zugrundeliegenden Banachraum. Der Übergang vom unendlichdimensionalen Objekt zum endlichdimensionalen Objekt geschieht meist mit Projektionen auf endlichdimensionale Räume.

In diesem Abschnitt beschäftigen wir uns überwiegend mit den obigen Fragen in normierten Räumen/Banachräumen, die eine (unbedingte) Schauderbasis besitzen. Wir wiederholen (siehe Definition 3.40).

Definition 4.17 Sei X ein Vektorraum.

- (a) Eine lineare Abbildung $P : X \rightarrow X$ heißt **Projektion** in X , falls $P \circ P = P$ gilt.
- (b) Zwei lineare Teilräume X_0, X_1 heißen **algebraisch komplementär**, falls gilt: $X = X_0 + X_1, X_0 \cap X_1 = \{\theta\}$. Ist dies der Fall, dann heißt X_1 ein **algebraisches Komplement** von X_0 und wir schreiben $X = X_0 \oplus X_1$.
- (c) Ist X ein normierter Raum und sind X_0, X_1 algebraisch komplementäre Teilräume von X , so heißen sie **topologisch komplementär**, wenn die Abbildung

$$X = X_0 \oplus X_1 \ni x_0 + x_1 \mapsto x_0 \in X$$

stetig ist. Diese Abbildung heißt **Projektion von X auf X_0 entlang X_1** .

□

Offenbar ist in der Definition 4.17 die Abbildung $\text{id} - P$ die Projektion von X auf X_1 entlang X_0 .

Folgerung 4.18 Sei X ein normierter Raum. Ist $P : X \rightarrow X$ eine Projektion, so ist auch $Q := \text{id} - P$ eine Projektion und es gelten die folgenden Aussagen.

$$P \circ Q = Q \circ P = \theta, \text{ran}(P) = \ker(Q), \ker(P) = \text{ran}(Q), X = \ker(P) \oplus \text{ran}(P), X = \ker(Q) \oplus \text{ran}(Q).$$

Beweis:

Alle Aussagen sind einfach einzusehen. Wir beweisen etwa $X = \ker(P) \oplus \operatorname{ran}(P)$. Wegen $\operatorname{id} = P + Q$ und $\operatorname{ran}(Q) = \ker(P)$ ist die Aussage $X = \ker(P) + \operatorname{ran}(P)$ klar. Sei nun $x \in \ker(P) \cap \operatorname{ran}(P)$. Dann gilt $x = Pz, Px = \theta$, und damit $x = Pz = P(P(z)) = P(x) = \theta$. ■

Folgerung 4.19 Sei X ein normierter Raum und seien X_0, X_1 komplementäre Teilräume in X . Dann gilt:

- (a) Die Abbildung $P : X = X_0 \oplus X_1 \ni (x_0 + x_1) \mapsto x_0 \in X$ ist eine Projektion P mit $\operatorname{ran}(P) = X_0, \ker(P) = X_1$.
- (b) Die Abbildung $Q : X = X_0 \oplus X_1 \ni (x_0 + x_1) \mapsto x_1 \in X$ ist eine Projektion Q mit $\operatorname{ran}(Q) = X_1, \ker(Q) = X_0$ und es gilt $P + Q = \operatorname{id}$.

Beweis:

Die Aussagen sind leicht einzusehen; siehe auch Folgerung 4.18. ■

Satz 4.20 Sei X ein Banachraum und sei X zerlegt gemäß $X = X_0 \oplus X_1$ mit komplementären Teilräumen X_0, X_1 . Dann sind äquivalent:

- (a) Die Zerlegung ist topologisch.
- (b) X_0, X_1 sind abgeschlossen.

Beweis:

- (a) $\implies$ (b) Sei die Zerlegung topologisch komplementär, also $P : X = X_0 \oplus X_1 \ni x = x_0 + x_1 \mapsto x_0 \in X$ stetig. Betrachte eine konvergente Folge $(x_0^n)_{n \in \mathbb{N}}$ in X_0 ; sei $\lim_n x_0^n = z = z_0 + z_1 \in X_0 \oplus X_1$. Dann gilt mit $x_0^n = Px_0^n, n \in \mathbb{N}$, wegen der Stetigkeit von P $z_0 = Pz = \lim_n Px_0^n = \lim_n x_0^n = z$, also $z \in X_0$. Analog zeigt man die Abgeschlossenheit von X_1 .
- (b) $\implies$ (a) Seien X_0, X_1 abgeschlossene Teilräume. Es ist zu zeigen, dass die Projektion

$$P : X = X_0 \oplus X_1 \ni x_0 + x_1 \mapsto x_0 \in X$$

stetig ist. Wir beweisen die Stetigkeit von P mit Satz 4.15. Also haben wir die Abgeschlossenheit von P zu zeigen. Sei $(x_n)_{n \in \mathbb{N}}$ eine Folge in X mit $x = \lim_n x_n, y = \lim_n Px_n$. Da $Px_n \in X_0$ für alle $n \in \mathbb{N}$ und da X_0 abgeschlossen ist, gilt $y \in X_0$. Ferner gilt $x - y = \lim_n (x_n - Px_n) \in X_1$. Also haben wir $x = y + (x - y) \in X_0 + X_1$ und mit der Eindeutigkeit der Darstellung $Px = y$, wie erforderlich. ■

Folgerung 4.21 Sei X ein Banachraum und sei X_0 ein abgeschlossener Teilraum von X . Dann sind äquivalent:

- (a) Es gibt eine stetige Projektion $P : X \rightarrow X$ mit $\operatorname{ran}(P) = X_0$.
- (b) Es gibt einen abgeschlossenen Teilraum X_1 von X mit $X = X_0 \oplus X_1$. (Die Zerlegung ist also topologisch.)

Beweis:

- (a) $\implies$ (b) Wir wissen $X = X_0 \oplus X_1$ mit $X_1 = \ker(P)$; siehe Folgerung 4.18. Da P stetig ist, ist $\ker(P)$ abgeschlossen.
- (b) $\implies$ (a) Siehe Satz 4.20. ■

Der Raum X_1 in obiger Folgerung ist isomorph zu X/X_0 . Der Isomorphismus j ist gegeben durch

$$j : X \ni x \mapsto [x] \in X/X_0.$$

Bekanntlich ist dank der Abgeschlossenheit von X_0 der Äquivalenzklassenraum X/X_0 ein Banachraum. Die Norm ist gegeben durch

$$X/X_0 \ni [x] \longmapsto \|[x]\| := \inf\{\|x - u\| \mid u \in X_0\} \in \mathbb{R}$$

Bekanntlich gibt es Banachräume, in dem es einen abgeschlossenen Teilraum gibt, der kein topologisches Komplement zulässt. Das „bekannteste“ Beispiel: Banachraum l_∞ , abgeschlossener Teilraum c_0 ; siehe [37].

Nach diesen mehr allgemeineren Anmerkungen kommen wir zurück zu den Betrachtungen über Schauderbasisen und wollen nun die Frage klären, wie solche Basen genutzt werden können zur Approximation. Hier ist eine Wiederholung der Definition.

Definition 4.22 Sei X ein normierter Raum. Eine Folge $(e^i)_{i \in \mathbb{N}}$ in X heißt **(unbedingte) Schauderbasis** in X , wenn es zu jedem $x \in X$ eine eindeutig bestimmte Folge von Skalaren $(a_i)_{i \in \mathbb{N}}$ gibt, so dass die Reihe $\sum_{i=1}^{\infty} a_i e^i$ (unbedingt) gegen x konvergiert. $\square$

Ein handliches Kriterium, mit dem man überprüfen kann, ob eine gegebene Folge eine Schauderbasis bildet, ist in folgendem Lemma aufgeführt.

Lemma 4.23 Sei $(e^n)_{n \in \mathbb{N}}$ eine Folge im normierten Raum $(X, \|\cdot\|)$, $e^n \neq \theta$ für alle $n \in \mathbb{N}$. Die lineare Hülle $\text{span}(e_1, e_2, e_3, \dots)$ sei eine dichte Teilmenge von X . Dann ist $(e^n)_{n \in \mathbb{N}}$ eine Schauderbasis von X , falls es ein $c > 0$ gibt, so dass für jede Folge $(a_n)_{n \in \mathbb{N}}$ von Skalaren in $\mathbb{K}$

$$\left\| \sum_{i=1}^k a_i e^i \right\| \leq c \left\| \sum_{i=1}^m a_i e^i \right\|, \quad k, m \in \mathbb{N}, k < m, \quad (4.6)$$

gilt.

Beweis:

Aus den Voraussetzungen folgt, dass jedes $x \in X$ eine Darstellung der Form $x = \sum_{i=1}^{\infty} a_i e^i$ besitzt. Um die Eindeutigkeit der Entwicklung zu beweisen, zeigen wir induktiv, dass aus $\sum_{i=1}^{\infty} a_i e^i = \theta$ folgt, dass alle a_i verschwinden. Sei also $\sum_{i=1}^{\infty} a_i e^i = \theta$. Dann gilt wegen (4.6) für $n = 1$ mit Grenzübergang $m \rightarrow \infty$, dass $|a_1| \|e^1\| = 0$ gelten muss, und daher wegen $e^1 \neq \theta$ schließlich $a_1 = 0$. Gilt nun $a_1 = \dots = a_n = 0$, so haben wir für $m > n + 1$

$$\left\| \sum_{i=1}^{n+1} a_i e^i \right\| = |a_{n+1}| \|e^{n+1}\| \leq c \left\| \sum_{i=1}^m a_i e^i \right\|,$$

woraus wie im Induktionsanfang folgt, dass $a_{n+1} = 0$ gelten muss. $\blacksquare$

Besitzt der Banachraum X eine Schauderbasis $(e^i)_{i \in \mathbb{N}}$, so sind die Koeffizientenfunktionale

$$\lambda_n : X \ni x = \sum_{i=1}^{\infty} a_i e^i \longmapsto a_n \in K$$

wohldefiniert. Wir wissen schon, dass sie lineare Funktionale sind. Wir wollen nun auch ihre Stetigkeit beweisen.

Definition 4.24 Sei $(e^i)_{i \in \mathbb{N}}$ eine Schauderbasis im Banachraum X mit Koeffizientenfunktionalen $(\lambda_i)_{i \in \mathbb{N}}$. Dann heißen die Abbildungen

$$P_n : X \ni x = \sum_{i=1}^{\infty} \lambda_i(x) e^i \longmapsto \sum_{i=1}^n \lambda_i(x) e^i \in X$$

die (zugehörigen) **kanonischen Projektionen**. $\square$

Satz 4.25 Sei $(e^i)_{i \in \mathbb{N}}$ eine Schauderbasis im Banachraum $(X, \|\cdot\|)$ mit kanonischen Projektionen $(P_n)_{n \in \mathbb{N}}$. Dann gilt:

- (a) Jede Projektion P_n ist linear und $\dim \operatorname{ran}(P_n) = n$ für alle $n \in \mathbb{N}$.
- (b) $P_n P_m = P_m P_n = P_{\min(m,n)}$ für alle $m, n \in \mathbb{N}$.
- (c) $\lim_n P_n x = x$ für alle $x \in X$.
- (d) Durch $\|\cdot\| : X \ni x \mapsto \sup_n \|P_n x\| \in \mathbb{R}$ wird auf X eine Norm erklärt.
- (e) $(X, \|\cdot\|)$ ist vollständig und die Norm $\|\cdot\|$ ist äquivalent zur Norm $\|\cdot\|$.
- (f) Alle kanonischen Projektionen sind stetig und es gibt eine Konstante $pc \geq 1$ mit $\|P_n\|_{X \rightarrow X} \leq pc, n \in \mathbb{N}$.
- (g) Alle Koeffizientenfunktionale sind stetig und es gilt

$$1 \leq \|e^k\| \|\lambda_k\|_* \leq bc, k \in \mathbb{N}, \text{ mit einer Konstanten } bc \geq 1. \quad (4.7)$$

Beweis:

Zu (a): Die Linearität folgt aus der Eindeutigkeit einer Darstellung, die Dimensionsaussage aus der Tatsache, dass $e^1, \dots, e^n$ linear unabhängig sind.

Zu (b): Folgt mit der Tatsache $\lambda_k(e^l) = \delta_{kl}$.

Zu (c): Dies ist Teil der Definition einer Schauderbasis.

Zu (d): Wegen (c) ist für jedes $x \in X$ die Folge $(P_n x)_{n \in \mathbb{N}}$ konvergent und damit auch beschränkt. Daher ist $d := \sup_{n \in \mathbb{N}} \|P_n\|_{X \rightarrow X}$ endlich. Die Normeigenschaften sind offensichtlich.

Zu (e): Wir zeigen die Vollständigkeit von $(X, \|\cdot\|)$. Sei dazu $(x^n)_{n \in \mathbb{N}}$ eine Cauchyfolge in X bezüglich der Norm $\|\cdot\|$; $x^n = \sum_{i=1}^{\infty} a_i^n e^i, n \in \mathbb{N}$. Sei $\varepsilon > 0$. Dann gibt es $N \in \mathbb{N}$ mit

$$\|x^r - x^s\| = \sup_{m \in \mathbb{N}} \left\| \sum_{i=1}^m (a_i^r - a_i^s) e^i \right\| < \varepsilon, r, s \geq N.$$

Wir haben wegen $\|x\| = \lim_n \|P_n x\|$ für alle $x \in X$ die Abschätzung $\|x\| \leq \|x\|$ für alle $x \in X$. Damit ist $(x^n)_{n \in \mathbb{N}}$ auch eine Cauchyfolge in X bezüglich der Norm $\|\cdot\|$. Sei $y := \lim_n x^n$ im Raum $(X, \|\cdot\|)$. Wir haben

$$\begin{aligned} |a_k^n| &= \|a_k^n e^k\| \|e^k\|^{-1} = \|e^k\|^{-1} \left\| \sum_{i=1}^k a_i^n e^i - \sum_{i=1}^{k-1} a_i^n e^i \right\| \\ &\leq \|e^k\|^{-1} \left(\left\| \sum_{i=1}^k a_i^n e^i \right\| + \left\| \sum_{i=1}^{k-1} a_i^n e^i \right\| \right) \leq 2 \|e^k\|^{-1} \|x^n\|. \end{aligned}$$

Daraus lesen wir ab, dass für jedes $k = 1, 2, \dots$ nun $(a_k^n)_{n \in \mathbb{N}}$ eine Cauchyfolge in $\mathbb{K}$ bildet und daher konvergent ist; wir setzen $y_k := \lim_n a_k^n, k \in \mathbb{N}$. Wegen

$$\sum_{i=1}^m (y_i - x_i^s) e^i = \lim_r \sum_{i=1}^m (a_i^r - x_i^s) e^i$$

gilt

$$\left\| \sum_{i=1}^m (y_i - x_i^s) e^i \right\| = \lim_r \left\| \sum_{i=1}^m (a_i^r - x_i^s) e^i \right\| \leq \varepsilon$$

für alle $s \geq N$ und alle $m \in \mathbb{N}$. Dies zeigt dass $\sum_{i=1}^{\infty} y_i e^i$ existiert und dass wegen $y = \lim_n x^n$ im Raum $(X, \|\cdot\|)$ gilt: $y = \sum_{i=1}^{\infty} y_i e^i$. Dies zeigt auch, dass $(\sum_{i=1}^{\infty} a_i^n e^i)_{n \in \mathbb{N}}$ in $(X, \|\cdot\|)$ konvergiert. Also gibt es $M \in \mathbb{N}$, so dass

$$\left\| \sum_{i=r}^s a_i^n e^i \right\| < \varepsilon \text{ für alle } r, s \geq M.$$

Nun haben wir

$$\begin{aligned} \left\| \sum_{i=r}^s y_i e^i \right\| &= \sup_{t \geq r} \left\| \sum_{i=r}^{\min(t,s)} y_i e^i \right\| \\ &\leq \sup_{r \leq t \leq s} \left(\left\| \sum_{i=r}^t (y_i - a_i^n) e^i \right\| + \left\| \sum_{i=r}^t a_i^n e^i \right\| \right) \\ &\leq \sup_{r \leq t \leq s} \left(\left\| \sum_{i=r}^t (y_i - a_i^n) e^i \right\| + \left\| \sum_{i=r}^{r-1} (y_i - a_i^n) e^i \right\| + \varepsilon \right) \leq \varepsilon \end{aligned}$$

Dies zeigt, dass $\sum_{i=1}^{\infty} y_i e^i$ konvergent im Raum $(X, \|\cdot\|)$ ist.

Sei $\tilde{X} := X$, versehen mit der Norm $\|\cdot\|$. Wir wissen, dass $(\tilde{X}, \|\cdot\|)$ vollständig ist. Die Abbildung $j : \tilde{X} \ni x \mapsto x \in X$ ist offenbar bijektiv und wegen $\|x\| \leq \|\tilde{x}\|$ für alle $x \in X$ ist j auch stetig. Nach dem Satz von der offenen Abbildung ist auch j^{-1} stetig. Dann gibt es eine Konstante $c > 0$ mit $c\|\tilde{x}\| \leq \|x\|$ für alle $x \in X$.

Zu (f): Mit der erzielten Ungleichung $c\|\tilde{x}\| \leq \|x\| \leq \|\tilde{x}\|$ für alle $x \in X$ folgt

$$\|P_n x\| = \left\| \sum_{i=1}^n \lambda_i(x) e^i \right\| \leq \sup_{n \in \mathbb{N}} \left\| \sum_{i=1}^n \lambda_i(x) e^i \right\| = \|\tilde{x}\| \leq c^{-1} \|x\|, \quad x \in X.$$

Damit ist die Stetigkeit gezeigt. Die gleichmäßige Beschränktheit haben wir schon mitbewiesen, wir könnten sie nun auch mit Banach-Steinhaus auf Grund von c) beweisen.

Zu (g): Unter f) haben wir schon $|\lambda_k(x)| \leq 2\|e^k\|^{-1}\|\tilde{x}\|, x \in X$, gezeigt. Mit der Äquivalenz der Normen $\|\cdot\|, \|\cdot\|$ folgt mit einer Konstanten $c > 0$

$$|\lambda_k(x)| \leq 2\|e^k\|^{-1}c^{-1}\|x\|, \quad x \in X.$$

Wegen $\lambda_k(e^k) = 1$, ist die behauptete Abschätzung bewiesen. ■

Die kleinste Konstante pc mit $\|P_n\|_{X \rightarrow X} \leq pc, n \in \mathbb{N}$, bezeichnen wir mit $BC(B)$, also $BC(B) := \sup_n \|P_n\|_{X \rightarrow X}$. $BC(B)$ heißt die **Basiskonstante** der Schauderbasis B .

Kommen wir nun zu den Möglichkeiten der Approximation in einem Banachraum X mit (unbedingter) Schauderbasis $B := (e^i)_{i \in \mathbb{N}}$. Wir skizzieren drei Möglichkeiten.

Lineare Approximation

Für gegebenes $n \in \mathbb{N}$ verwendet man als Approximation für ein $x \in X$ die die n -te Partialsumme der Reihe $\sum_{i=1}^{\infty} \lambda_i(x) e^i$, also $P_n x$ mit der zugehörigen Projektion. Wir setzen:

$$E_n(x, B) := \inf_{a_1, \dots, a_n \in \mathbb{K}} \left\| x - \sum_{i=1}^n a_i e^i \right\| \quad (4.8)$$

Man könnte erwarten, dass der beste Approximationsfehler $E_n(x, B)$ durch $P_n x$ realisiert wird; die Fourierreihen sind dafür Vorbild. Allerdings haben wir dort Orthogonalität zur Verfügung, die hier fehlt, im Kapitel über Hilberträume kommen wir darauf zurück. Hier ist eine gewisse Bestätigung dieser Erwartung.

Satz 4.26 Sei $B = (e^i)_{i \in \mathbb{N}}$ eine Schauderbasis im Banachraum $(X, \|\cdot\|)$ mit kanonischen Projektionen $(P_n)_{n \in \mathbb{N}}$. Dann gilt mit einer Konstanten $C \geq 1$:

$$\|x - P_n x\| \leq C E_n(x, B), \quad n \in \mathbb{N}, \text{ für alle } x \in X. \quad (4.9)$$

Beweis:

Nach Definition des linearen Bestapproximationsfehlers $E_n(x, B)$ gibt es zu jedem $\varepsilon > 0$ Koeffizienten $a_1, \dots, a_n \in \mathbb{K}$ mit

$$\|x - y\| \leq E_n(x, B) + \varepsilon, \quad \text{wobei } y := \sum_{i=1}^n a_i e^i \text{ ist.}$$

Damit haben wir die Abschätzung

$$\|x - P_n x\| \leq \|x - y\| + \|y - P_n x\| \leq E_n(x, B) + \varepsilon + \|P_n y - P_n x\|,$$

wobei wir genutzt haben, dass $P_n|_{\text{span}(e^1, \dots, e^n)} = \text{id}$ ist. Nun folgt

$$\begin{aligned} \|x - P_n x\| &\leq E_n(x, B) + \varepsilon + \|P_n(y - x)\| \\ &\leq E_n(x, B) + \varepsilon + \|P_n\|_{X \rightarrow X} \|y - x\| \\ &\leq E_n(x, B) + \varepsilon + \text{BC}(E_n(x, B) + \varepsilon) \\ &\leq (\text{BC} + 1)E_n(x, B) + (\text{BC} + 1)\varepsilon, \end{aligned}$$

woraus die Behauptung folgt, da $\varepsilon > 0$ beliebig ist. ■

Damit erweist sich die lineare Projektion bezüglich einer Schauderbasis nahezu als bestmögliche Approximation. Zu vergleichen ist diese Art der Approximation mit

n-Term – Approximation

Für gegebenes $n \in \mathbb{N}$ verwendet man als Approximation für ein $x \in X$ die die „beste Partialsumme“ der Reihe $\sum_{i=1}^{\infty} \lambda_i(x) e^i$, die aus n Termen besteht. Also sollten wir

$$F_n(x, B) := \inf_{\Lambda \subset \mathbb{N}, \#\Lambda = n} \inf_{a_1, \dots, a_n \in \mathbb{K}} \|x - \sum_{i \in \Lambda} a_i e^i\| \quad (4.10)$$

betrachten. Die Ungleichung $F_n(x, B) \leq E_n(x, B)$ ist offensichtlich. Die Vorgehensweise ist nun aber eine nichtlineare, denn Approximationen zu x und x' lassen sich im Allgemeinen nicht aus den Approximationen zu x bzw. x' additiv erhalten.

Greedy – Approximation

Diese Art der Approximation nutzt eine absteigende Anordnung der Koeffizienten einer Darstellung durch eine Schauderbasis.

Definition 4.27 Sei $(e^i)_{i \in \mathbb{N}}$ eine Schauderbasis im Banachraum X mit Koeffizientenfunktionen $(\lambda_i)_{i \in \mathbb{N}}$. Eine Permutation $\rho: \mathbb{N} \rightarrow \mathbb{N}$ heißt **absteigend** bezüglich $x \in X$, wenn gilt:

$$|\lambda_{\rho(i)}(x)| \geq |\lambda_{\rho(i+1)}(x)| \text{ für alle } i \in \mathbb{N}.$$

$D(x)$ sei die Menge aller bezüglich x absteigenden Permutationen. □

Damit setzen wir für $x \in X$ mit $\rho \in D(x)$:

$$G_n(x, B, \rho) := \sum_{i=1}^n \lambda_{\rho(i)}(x) e^{\rho(i)}.$$

Diese n -Greedy-Approximation eines $x \in X$ ist also eine Partialsumme einer umgeordneten Reihenentwicklung von x . Die Vorgehensweise ist wiederum eine nichtlineare, denn Approximationen zu x und x' lassen sich im Allgemeinen nicht aus den Approximationen zu x bzw. x' additiv erhalten.

Konvergenz von $G_n(x, B, \rho)$ für $n \rightarrow \infty$ gegen x für alle $x \in X$ wird man daher jedenfalls nur dann erwarten können, wenn jede Umordnung von B eine Schauderbasis von X ist, es sich bei B also um eine unbedingte Basis handelt. Im Fall einer bedingten Basis wird dagegen im allgemeinen keine Konvergenz vorliegen. Daran schließen sich nun reizvolle Fragen an, die die Konvergenz auch unter schwächeren Bedingungen (Greedy-Basis, Quasi-Greedy-Basis, demokratische Basis) als der Unbedingtheit der Basis sicherstellen; siehe [6, 34]. Die effektive Approximation in Anwendungen (Signal- und Bildverarbeitung) ist vielfältig erprobt.

4.5 Fixpunktprinzip I–Der Banachsche Fixpunktsatz

Im Kapitel über metrische Räume haben wir den Kontraktionssatz (Satz 2.24) in vollständigen metrischen Räumen kennengelernt. Nun formulieren wir den Kontraktionssatz in metrischen Räumen um auf die Situation normierter Räume. Im Kapitel über Konvexität werden wir den Schauderschen Fixpunktsatz vorstellen. Zunächst aber einige allgemeine Anmerkungen über die Anwendung von Fixpunktgleichungen.

In Anwendungen, insbesondere in der mathematischen Physik, modellieren überwiegend Gleichungen die Fragestellungen. Im Allgemeinen entstehen diese Gleichungen durch eine Betrachtung der Beziehungen zwischen den Ursachen und Gegebenheiten (Kraft, Druck, Temperatur, Spannung, Dichte, Volatilität, ...) und den Wirkungen (Geschwindigkeit, Beschleunigung, Ausdehnung, Aktienkurse, ...). In diesen Gleichungen kann man dann – systemtheoretisch gesprochen – die Quantitäten

Eingangsgrößen, Systemparameter, Ausgangsgrößen

unterscheiden. Die Gleichungen stellen dann diese Größen in eine Beziehung. Mathematisch resultiert meist eine (lineare/nichtlineare) Gleichung

$$Ax = y, \quad (4.11)$$

wobei der „Operator“ A die Systemparameter enthält und die Umsetzung der Eingangsgrößen y in die gesuchten Wirkungen x beschreibt. Formal ist dies eine **Gleichung 1. Art**. Manchmal kommt diese Gleichung in der Form

$$(\text{id} - K)x = y \quad (4.12)$$

daher. Eine solche Gleichung bezeichnet man meist als **Gleichung 2. Art**. Man spricht dann von **Integralgleichungen 1. Art** bzw. **2. Art**, wenn der Operator A bzw. K ein Integraloperator ist. In der „einfachsten Form“ sind solche Integraloperatoren B wie folgt definiert:

$$B : C[a, b] \ni x \longmapsto Bx \in C[a, b], \quad Bx(t) := \int_a^b \kappa(t, s)x(s) ds, \quad t \in [a, b]. \quad (4.13)$$

Dabei heißt κ **Integalkern**. Bei der Aufgabe, gewöhnliche Differentialgleichungen mit dem Fixpunktsatz zu lösen, werden wir eine solche (nichtlineare) Integralgleichung 2. Art ableiten. In (4.12) ist die Fixpunktgleichung und damit die Fixpunktiteration schon abzulesen:

$$x^{k+1} = Kx^k + y, \quad x^0 \text{ gegeben.} \quad (4.14)$$

In vielen Fällen kann man aus der Gleichung (4.11) eine Gleichung der Form (4.12) herstellen. Ein üblicher „Trick“ ist, eine **Zerlegungsstrategie** zu verfolgen:

$$A = \text{id} - B, \text{ mit } B := \text{id} - A. \quad (4.15)$$

Damit dies einen Vorteil bringt, müssen im Allgemeinen schon besondere, vorteilhafte Eigenschaften vorliegen. Eine allgemeine Klasse, die hier einzuordnen sind, sind die **Regularisierungsverfahren** bei Gleichungen 1. Art, die wir später noch skizzieren werden. Das „Zerlegungsverfahren“ findet man auch in der Numerischen Mathematik bei der Lösung von Gleichungssystemen. Man geht aus von dem Gleichungssystem

$$Ax = y \text{ mit } A \in \mathbb{R}^{n,n}. \quad (4.16)$$

Die Matrix wird dann zerlegt in eine obere Dreiecksmatrix U , eine Diagonalmatrix D und eine untere Dreiecksmatrix L :

$$A = L + D + U. \quad (4.17)$$

Dann lassen sich unter geeigneten Voraussetzungen die Iterationsverfahren

$$x^{k+1} = -D^{-1}(L + U)x^k + D^{-1}y, \quad x^0 \text{ gegeben}, \quad (4.18)$$

bzw.

$$x^{k+1} = -(L + D)^{-1}x^k + (L + D)^{-1}y, \quad x^0 \text{ gegeben}, \quad (4.19)$$

betrachten. Das Verfahren in (4.18) nennt man das **Jakobi- bzw. Gesamtschrittverfahren**, das Vorgehen in (4.19) das **Gauß-Seidel- bzw. Einzelschrittverfahren**. In der numerischen Realisierung kann man auf die explizite Invertierung von D bzw. $L + D$ verzichten auf Kosten von Gleichungslösern.

Satz 4.28 (Banachscher Fixpunktsatz) Sei $(X, \|\cdot\|)$ ein Banachraum, sei V ein abgeschlossener Teilraum von X und sei $T : V \rightarrow V$ eine Kontraktion, d. h.

$$\exists L \in [0, 1) \forall x, y \in V (\|Tx - Ty\| \leq L\|x - y\|). \quad (4.20)$$

Dann haben wir folgende Aussagen:

a) Es gibt in V genau einen Fixpunkt $\bar{x}$ von T .

b) Die Iteration

$$x = x^0 \in V, \quad x^{n+1} := Tx^n, \quad n \in \mathbb{N}_0, \quad (4.21)$$

knvergiert gegen $\bar{x}$ für alle $x \in V$, und wir haben

$$c) \|x^n - \bar{x}\| \leq \frac{L^n}{1-L} \|Tx - x\|, \quad n \in \mathbb{N}.$$

$$(d) \|x^{n+1} - \bar{x}\| \leq \frac{L}{1-L} \|x^{n+1} - x^n\|, \quad n \in \mathbb{N}.$$

Beweis:

Betrachte den metrischen Raum (V, d) mit $d(x, y) := \|x - y\|$, $x, y \in V$. Dieser metrische Raum ist vollständig, da V abgeschlossen ist. Wende nun Satz 2.24 an. ■

Beispiel 4.29 Sei $a < b$, $F : [a, b] \rightarrow [a, b]$ stetig und sei F differenzierbar in (a, b) . Es gelte mit $q \in [0, 1)$:

$$|F'(\xi)| \leq q \text{ für alle } \xi \in (a, b). \quad (4.22)$$

Dann besitzt F genau einen Fixpunkt in $[a, b]$, denn nach dem Mittelwertsatz der Differentialrechnung ist F eine Kontraktion.

Beachte: Die Voraussetzung (4.22) ist keine notwendige Voraussetzung für die Existenz eines Fixpunktes, denn es reicht ja schon die Stetigkeit von F aus, um einen Fixpunkt zu garantieren ($F(a) - a \geq 0, F(b) - b \leq 0$). □

Beispiel 4.30 $X = \mathbb{R}$, $\|\cdot\| := |\cdot|$, $M := [0, \infty)$, $F : M \ni t \mapsto t + \frac{1}{1+t} \in M$. Wir haben

$$|F(t) - F(s)| < |t - s|.$$

F besitzt keinen Fixpunkt! Beachte, dass

$$|F(t) - F(s)| \leq q|t - s|, \quad t, s \in M,$$

für kein $q \in [0, 1)$ erreichbar ist. □

Beispiel 4.31 $X := \mathbb{R}$, $\|\cdot\| := |\cdot|$, $M := (0, \infty)$, $F : M \ni t \mapsto qt \in M$ mit $q \in [0, 1)$. Offenbar gilt

$$|F(t) - F(s)| \leq q|t - s|, \quad t, s \in M,$$

aber F besitzt keinen Fixpunkt. Beachte: M ist nicht abgeschlossen! □

Zahlreiche Naturgesetze und mathematische Modelle aus den verschiedensten Anwendungsgebieten führen in ihrer einfachsten Form auf gewöhnliche Differentialgleichungen, also auf Gleichungen, die Ableitungen einer gesuchten Funktion einer Variablen – meist der Zeit – enthalten. Mit unseren Mitteln können wir nun einen Existenzsatz dazu beweisen.

Wir betrachten das Existenzproblem von **Anfangswertaufgaben bei gewöhnlichen Differentialgleichungen**. Damit meinen wir eine Aufgabe der folgenden Art:

Gegeben: $f : [t_0, t_1] \times D \rightarrow \mathbb{R}^n$, $D \subset \mathbb{R}^n$, $y^0 \in D$ ($t_0, t_1 \in \mathbb{R}$, $t_0 < t_1$).

Gesucht: **Lösung** φ der **Anfangswertaufgabe**
(AWA) $y' = f(t, y)$, $y(t_0) = y^0$.

In obiger Aufgabenstellung nennt man f **rechte Seite** und y^0 den **Anfangswert** zum **Anfangszeitpunkt** t_0 und $y(t_0) = y^0$ die **Anfangsbedingung**. Mit **Lösung** ist dabei eine differenzierbare Funktion $\varphi : [t_0, t_1] \rightarrow D$ gemeint, für die gilt:

$$\varphi'(t) = f(t, \varphi(t)), \quad t \in [t_0, t_1], \quad \varphi(t_0) = y^0. \quad (4.23)$$

Sei die rechte Seite f stetig. Der Hauptsatz der Differential- und Integralrechnung besagt, dass die Suche nach einer Lösung der Anfangswertaufgabe (AWA) verknüpft werden kann mit der Suche nach einer Lösung der folgenden **Integralgleichung**:

$$(IG) \quad \varphi(t) = y^0 + \int_{t_0}^t f(s, \varphi(s)) ds, \quad t \in [t_0, t_1].$$

Denn wird (IG) gelöst durch φ , so ist φ differenzierbar und löst die Anfangswertaufgabe. Umgekehrt, ist φ eine Lösung der Anfangswertaufgabe, so sieht man durch Integration sofort, dass φ auch (IG) löst. Die Gleichung (IG) hat gegenüber der Aufgabe (AWA) den Vorzug, dass die Anfangsbedingung eingearbeitet ist, darin nur φ vorkommt und nicht φ und φ' . Die Gleichung (IG) kann daher im Raum $C([t_0, t_1]; \mathbb{R}^n)$ untersucht werden.

Satz 4.32 (Satz von Picard–Lindelöf) Sei $f : [t_0, t_1] \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ stetig, $y^0 \in \mathbb{R}^n$ beliebig, und es gebe $L \geq 0$ mit

$$|f(t, u) - f(t, v)| \leq L|u - v| \text{ für alle } u, v \in \mathbb{R}^n, t \in [t_0, t_1].$$

Dann gibt es genau eine stetig differenzierbare Funktion $\varphi : [t_0, t_1] \rightarrow \mathbb{R}^n$ mit

$$\varphi'(t) = f(t, \varphi(t)), \quad t \in [t_0, t_1], \quad \varphi(t_0) = y^0.$$

Beweis:

Wir haben zu zeigen, dass es genau eine stetige Abbildung $\varphi : [t_0, t_1] \longrightarrow \mathbb{R}^n$ gibt, die die Integralgleichung (IG) erfüllt.

Betrachte

$$X := C([t_0, t_1]; \mathbb{R}^n), \quad \|g\|_\infty := \max_{t \in [t_0, t_1]} e^{-2Lt} |f(g)|, \quad g \in X.$$

(L ist die Lipschitzkonstante aus den Voraussetzungen des Satzes.) Damit ist $(X, \|\cdot\|_\infty)$ ein Banachraum, denn das „Gewicht“ e^{-2Lt} ändert nichts an der Aussage von Beispiel 3.7.

Wir wollen Satz 2.24 auf die Fixpunktgleichung anwenden, als die sich (IG) interpretieren lässt. Dazu definieren wir:

$$T : X \longrightarrow X, \quad T(\varphi)(t) := y^0 + \int_{t_0}^t f(s, \varphi(s)) ds, \quad t \in [t_0, t_1].$$

Offenbar gilt für $\varphi, \psi \in X, t \in [t_0, t_1]$:

$$\begin{aligned} |T(\varphi)(t) - T(\psi)(t)| &\leq \int_{t_0}^t |f(s, \varphi(s)) - f(s, \psi(s))| ds \leq L \int_{t_0}^t |\varphi(s) - \psi(s)| ds \\ &= L \int_{t_0}^t |\varphi(s) - \psi(s)| e^{-2Ls} e^{2Ls} ds \leq L \|\varphi - \psi\|_\infty \int_{t_0}^t e^{2Ls} ds \\ &\leq \frac{1}{2} e^{2Lt} \|\varphi - \psi\|_\infty, \end{aligned}$$

d. h.

$$\|T(\varphi) - T(\psi)\| \leq \frac{1}{2} \|\varphi - \psi\|_\infty.$$

Damit sind die Voraussetzungen des Satzes 2.24 erfüllt und T besitzt genau einen Fixpunkt. ■

Folgerung 4.33 Sei $f : [t_0, t_1] \times \overline{B}_r(y^0) \longrightarrow \mathbb{R}^n$ stetig und es gebe $L \geq 0$ mit

$$|f(t, x) - f(t, x')| \leq L|x - x'| \text{ für alle } x, x' \in \overline{B}_r(y^0), t \in [t_0, t_1].$$

Dann gibt es genau eine stetig differenzierbare Funktion $\varphi : [t_0, t_0 + a] \longrightarrow \mathbb{R}^n$ mit

$$\varphi'(t) = f(t, \varphi(t)), \quad t \in [t_0, t_0 + a], \quad \varphi(t_0) = y^0;$$

hierbei sind $a := \min(t_1 - t_0, rm^{-1})$, $m := \max\{|f(s, x)| : s \in [t_0, t_1], x \in \overline{B}_r(y^0)\}$.

Beweis:

Wir setzen f vom „Rechteck“ $R := [t_0, t_1] \times \overline{B}_r(y^0)$ zu $\tilde{f}$ auf $[t_0, t_1] \times \mathbb{R}^n$ fort durch

$$\tilde{f}(t, x) := \begin{cases} f(t, x) & , \text{ falls } x \in \overline{B}_r(y^0) \\ f(t, y^0 + r|x - y^0|^{-1}(x - y^0)) & , \text{ sonst} \end{cases}.$$

Nun erfüllt $\tilde{f}$ die Voraussetzungen des Satzes 4.32: die Stetigkeit ist klar, die Lipschitzstetigkeit ist erfüllt mit derselben Lipschitzkonstante. Es gibt also $\tilde{\varphi} : [t_0, t_1] \longrightarrow \mathbb{R}^n$ mit

$$\tilde{\varphi}'(t) = \tilde{f}(t, \tilde{\varphi}(t)), \quad t \in [t_0, t_1], \quad \tilde{\varphi}(t_0) = y^0.$$

Solange nun $\tilde{\varphi}(t) \in \overline{B}_r(y^0)$ gilt, ist $\tilde{f}(t, \tilde{\varphi}(t)) = f(t, \tilde{\varphi}(t))$. $\tilde{\varphi}$ verläßt das Rechteck R frühestens zum Zeitpunkt $t^* = t_0 + a$, da

$$|\tilde{\varphi}(t) - y^0| \leq \int_{t_0}^t |\tilde{f}(s, \tilde{\varphi}(s))| ds \leq (t - t_0)m$$

gilt. Also ist $\varphi|_{[t_0, t_0+a]}$ eine Lösung der vorgegebenen Aufgabe. Sie ist eindeutig bestimmt, da die Lösung der zugeordneten Integralgleichung ($f = \tilde{f}$) auf $[t_0, t_0 + a]$ eindeutig bestimmt ist. ■

Beispiel 4.34 Hier ist ein Modell für das Wachstum einer Population.

$$(AWA) \quad y' = y(a - by), \quad y(0) = y^0; \quad (a, b > 0).$$

Die rechte Seite – eine explizite Abhängigkeit von der Zeit t liegt nicht vor, man sagt, die Differentialgleichung sei autonom – heißt **logistisches Gesetz** des Populationswachstums.

Mit $f(t, y) := y(a - by)$ gilt für $y \in \overline{B}_r(y^0)$ mit $r > 0$

$$|f(t, x) - f(t, x')| \leq L|x - x'|, \quad \text{wobei } L := \max\left\{\left|\frac{\partial f}{\partial y}(t, \xi)\right| \mid t \in \mathbb{R}, \xi \in \overline{B}_r(y^0)\right\}.$$

Wir erhalten eine eindeutige Lösung zumindest für ein „kleines Stück“ in die Zukunft (lokale Lösung). Aber die Lösung läßt sich explizit ausrechnen. Wir geben sie an, die Verifikation, dass eine Lösung vorliegt, ist einfach:

$$\varphi(t) := \frac{ay^0}{by^0 + (a - by^0)\exp(-at)}, \quad t \in [0, \infty).$$

Es fällt auf, dass $\lim_{t \rightarrow \infty} \varphi(t) = \frac{a}{b}$ gilt: über „lange Zeit“ stellt sich das „Gleichgewicht“ $\frac{a}{b}$ ein ($\frac{a}{b}$ ist eine Nullstelle der rechten Seite!). Erweiterte Modelle sind mit dem Namen **Volterra-Lotka**-Modelle verbunden. □

Beispiel 4.35 Betrachte

$$(AWA) \quad y' = 1 + y^2, \quad y(0) = 0.$$

Die Lösung der Anfangswertaufgabe von (AWA) ist der Tangens ($\tan'(y) = 1 + y^2$). Diese Lösung zeigt uns, dass wir nicht immer erwarten dürfen, dass die Lösung für die ganze Zukunft existiert, obwohl die rechte Seite für alle Zeiten definiert ist und hinreichend oft differenzierbar ist. Der Grund für die eingeschränkte Existenz liegt darin begründet, dass hier die rechte Seite „großes Wachstum“ aufweist. □

4.6 Übungen

- 1.) Sei $(C([0, 1]), \|\cdot\|_\infty)$ der bekannte normierte Raum und für jedes $n \in \mathbb{N}$ sei

$$F_n := \{f \in C([0, 1]) \mid \exists t \in [0, 1] \text{ mit } |f(t) - f(s)| \leq n|t - s| \text{ für alle } s \in [0, 1]\}.$$

Man zeige, dass die Mengen F_n in $C([0, 1])$ abgeschlossen sind und keine inneren Punkte besitzen. Hieraus folgere man, dass es eine dichte Menge von Funktionen in $C([0, 1])$ gibt, die an keinem Punkt in $[0, 1]$ differenzierbar sind.

- 2.) Sei $(x_n)_{n \in \mathbb{N}}$ eine reelle Folge. Betrachte die Zuordnung T mit

$$T((x_n)_{n \in \mathbb{N}}) := \begin{cases} (1 - \sum_{n=1}^{\infty} |x_n|^2)^{\frac{1}{2}} & , \text{ falls } n = 1 \\ x_{n-1} & , \text{ sonst} \end{cases}.$$

Zeige:

- (a) T ist eine wohldefinierte Abbildung von $\overline{B}_1 \rightarrow \overline{B}_1$, wobei $\overline{B}_1$ die abgeschlossene Einheitskugel in l_2 ist.
- (b) $T : l_2 \supset \overline{B}_1 \rightarrow l_2$ ist stetig.
- (c) T besitzt keinen Fixpunkt.

- 3.) Betrachte

$$T : C[a, b] \ni f \mapsto \int_0^b f(t) dt \in \mathbb{R}.$$

- (a) Zeige: T ist linear und stetig.
 - (b) Bestimme T^{-1} .
 - (c) Zeige, es gibt eine Folge $(g_n)_{n \in \mathbb{N}}$ in $C[0, 1]$ und g in $C[0, 1]$ mit $\lim_n g_n = g$ in $C[0, 1]$, $\lim_n \|T^{-1}g_n - T^{-1}g\| = 1$.
- 4.) Seien X, Y, Z Banachräume, A ein linearer, stetiger Operator von X nach Y und T ein abgeschlossener, linearer Operator von Y nach Z mit Definitionsbereich $D(T)$ und $\text{ran}(A) \subset D(T)$. Zeige, dass $T \circ A : X \rightarrow Z$ stetig ist.
- 5.) Seien X, Y, Z normierte Räume, wobei X oder Y ein Banachraum sei. Sei $B : X \times Y \rightarrow Z$ eine bilineare Abbildung. Es gelte: T ist partiell stetig, d. h. $T(\cdot, y) : X \rightarrow Z, T(x, \cdot) : Y \rightarrow Z$ sind stetig für alle $x \in X, y \in Y$. Zeige: T ist stetig.
Hinweise (unter der Annahme, dass X ein Banachraum ist): Für alle $x \in X$ gibt es eine Konstante K_x mit $\|T(x, y)\|_Z \leq K_x \|y\|_Y$ für alle $y \in Y$. Wende den Satz von Banach Steinhaus auf die Familie $\{T(\cdot, y) \mid \|y\|_Y \leq 1\}$ an.
- 6.) Betrachte die Anfangswertaufgabe

$$y' = -e^{-y} \cos(t), \quad y(0) = 1.$$

- (a) Zeige: Die Anfangswertaufgabe hat eine eindeutig bestimmte Lösung in $[0, 1]$.
 - (b) Berechne diese Lösung.
- 7.) Betrachte

$$F(x, y) := \frac{1}{2} \left(1 + \frac{\sin(x)}{4} + y, 1 + \sin(y) + x \right), \quad (x, y) \in \mathbb{R}^2.$$

- (a) Untersuche die Kontraktionseigenschaften von F jeweils bezüglich der Norm $\|\cdot\| := \|\cdot\|_{\infty}$ und $\|\cdot\| := \|\cdot\|_2$ in $\mathbb{R}^2$.
- (b) Wie oft ist bei Verwendung der Fixpunktiteration

$$(x^{n+1}, y^{n+1}) := F(x^n, y^n),$$

ausgehend von $(0, 0)$, zu iterieren, bis der Fixpunkt mit Genauigkeit 10^{-2} erreicht ist.

- 8.) Für eine invertierbare Matrix $A \in \mathbb{R}^{n,n}$ ist die Matrix $X := A^{-1}$ offensichtlich eine Lösung der (nichtlinearen) Gleichung

$$X^{-1} - A = \Theta. \quad (4.24)$$

Zeige:

- (a) Die Iteration

$$X^{n+1} := X^n + X^n(E - AX^n), \quad n = 0, 1, \dots, \quad (4.25)$$

stellt gerade das Newton-Verfahren zur Lösung von (4.24) dar.

- (b) Für jede Startmatrix X^0 mit $\|E - AX^0\| \leq q < 1$ ($\|\cdot\|$ ist eine Matrixnorm) konvergiert die in (4.25) erklärte Folge $(X^n)_{n \in \mathbb{N}}$ gegen die Matrix A^{-1} und es gelten die Abschätzungen

$$\|X^n - A^{-1}\| \leq (1 - q)^{-1} \|X^0\| \|E - AX^0\| \leq (1 - q)^{-1} \|X^0\| q^{2^n}, \quad n = 0, 1, \dots$$

- 9.) Betrachte $D_q := \{f : [a, b] \rightarrow \mathbb{R} \mid f \text{ stetig differenzierbar}, |f'(\xi)| \leq q, \xi \in [a, b]\} \subset C[a, b]$. Zeige, dass der Abschluss von D_q in $C[a, b]$ (bezüglich der gleichmäßigen Konvergenz) die Menge $Lip_q := \{f : [a, b] \rightarrow \mathbb{R} \mid f \text{ stetig}, |f(t) - f(s)| \leq q|t - s|, t, s \in [a, b]\}$ ist.
- 10.) Seien U, V abgeschlossene Teilräume eines Banachraumes $(X, \|\cdot\|)$ mit $U \cap V = \{\theta\}$. Dann sind äquivalent:
- (a) $U \oplus V$ ist abgeschlossen.
 - (b) $P : U \oplus V \ni (u, v) \mapsto u \in U$ ist stetig.
 - (c) $\|\cdot\|$ mit $\|u + v\| := \|u\| + \|v\|, u \in U, v \in V$, ist eine zu $\|\cdot\|$ äquivalente Norm auf $U \oplus V$.
- 11.) Seien U, V abgeschlossene Teilräume eines Banachraumes $(X, \|\cdot\|)$ mit $U \cap V = \{\theta\}$. Dann sind äquivalent:
- (a) $U \oplus V$ ist abgeschlossen.
 - (b) Es gibt eine Konstante $c > 0$ mit $\|u\| \leq c\|u + v\|$ für alle $u \in U, v \in V$.
 - (c) $c := \inf\{\|u - v\| \mid u \in U, v \in V, \|u\| = \|v\| = 1\} > 0$.
- 12.) Betrachte $c_f := \{x = (x_k)_{k \in \mathbb{N}} \in c_0 \mid x_k \neq 0 \text{ für fast alle } k \in \mathbb{N}\}$ als Teilraum des Banachraumes $(c_0, \|\cdot\|)$ und damit die Abbildungen

$$A_k : c_f \ni x = (x_k)_{k \in \mathbb{N}} \mapsto A_k x \in \mathbb{R} \text{ mit } A_k x := \sum_{j=1}^k x_j, \quad k \in \mathbb{N}.$$

Zeige:

- (a) $A_k \in \mathcal{B}(c_f, \mathbb{R})$ für alle $k \in \mathbb{N}$.
- (b) $\sup_k |A_k x| < \infty$ für alle $x \in c_f$.
- (c) $\sup_k \|A_k\|_{c_f \rightarrow \mathbb{R}} = \infty$.

Die obigen Beobachtungen zeigen, dass der Satz von Banach-Steinhaus nicht zutrifft. Wieso?

- 13.) Illustration einiger ungewohnter Effekte in unendlichdimensionalen Räumen.

- (a) Finde eine lineare, surjektive Abbildung $T \in \mathcal{B}(X)$ mit einem Banachraum X , so dass $\ker(T) := \{x \in X \mid Tx = \theta\} \neq \{\theta\}$.
- (b) Finde eine lineare, injektive Abbildung $T \in \mathcal{B}(X)$ mit einem Banachraum X , so dass $\text{ran}(T) := \{Tx \in X \mid x \in X\} \neq X$.
- (c) Sei $\mathcal{P}$ der Raum der Polynome, betrachtet als Teilraum von $C[-2, 2]$. Seien $m_n := x^n, n \in \mathbb{N}$, die Monome in $\mathcal{P}$. Ist die lineare Abbildung T , definiert auf der Basis der Monome durch $m_n \mapsto 3^n m_n, n \in \mathbb{N}$, stetig?

14.) Betrachte den Banachraum l_2 die Abbildung $T : l_2 \ni x = (x_k)_{k \in \mathbb{N}} \mapsto ((1 - 1/k)x_k)_{k \in \mathbb{N}} \in l_2$.

- (a) Zeige T ist stetig und $\|T\| = 1$.
- (b) Zeige: $U := \{x \in X \mid \|Tx\| \geq 1\}$ ist abgeschlossen.
- (c) Gibt es ein $u \in U$ mit $\|u\| = \inf_{z \in U} \|z\|$?

15.) Seien X, Y normierte Räume, sei $T : X \rightarrow Y$ linear und es gelte $\dim Y < \infty$. Zeige die Äquivalenz von

- (a) T ist stetig.
- (b) $\ker(T) := \{x \in X \mid Tx = \theta\}$ ist abgeschlossen.

Folgere, dass T offen ist, wenn T stetig ist.

Hinweis: Man darf verwenden, dass X/U ein normierter Raum ist, wenn U ein abgeschlossener Teilraum von X ist; die passende Norm in X/U ist $\|[x]\| := \inf\{\|x - u\| \mid u \in U\}$, $[x] \in X/U$.

16.) Betrachte den Banachraum $(l_\infty, \|\cdot\|_\infty)$ und darin den abgeschlossenen Teilraum c_0 . Für $x = (x^k)_{k \in \mathbb{N}} \in l_\infty$ setze

$$S_x := \{y = (y^k)_{k \in \mathbb{N}} \mid |y^k| \leq |x^k|, k \in \mathbb{N}\}.$$

Sei $A \subset l_\infty$. Zeige die Äquivalenz von

- (a) A ist kompakt.
- (b) $A \subset S_x$ für ein $x \in c_0$.

Hinweis: Welches x würde das „kleinste“ in (b) sein?

17.) Betrachte den Raum $X := C[-1, 1]$ der stetigen Funktionen auf $[-1, 1]$, versehen mit der Supremumsnorm $\|\cdot\| := \|\cdot\|_\infty$. Bekanntlich ist $(X, \|\cdot\|)$ ein Banachraum. Zeige:

- (a) $U := \{f \in C[-1, 1] \mid f(t) = f(-t), t \in [-1, 1]\}$ ist ein abgeschlossener Teilraum.
- (b) Es gibt einen abgeschlossenen Teilraum V mit $X = U \oplus V$.
- (c) Die Projektion P auf den Teilraum U ist stetig.

Kapitel 5

Dualität

Mathematics is the most beautiful and most powerful
creation of human spirit

S. Banach

Das zentrale Ergebnis ist der Satz von Hahn–Banach, der Aufschluss über die Existenz von stetigen linearen Funktionalen liefert. Viele reizvolle Fragestellungen finden damit auch einen Lösungsansatz. Aus den Resultaten ziehen wir im nächsten Kapitel (überwiegend) geometrische Folgerungen im Rahmen des Konzepts „Konvexität“. Stetige lineare Funktionale erlauben es, in einem Banachraum eine „duale Topologie“ einzuführen, die in manchen Fragen (Kompaktheit, Existenz von Minima, ...) einfache Antworten gestattet. Abschließend betrachten wir duale Operatoren und ihre Eigenschaften.

5.1 Der Satz von Hahn–Banach

Kurz eine Wiederholung von Bezeichnungen und Begriffen, die wir im Kapitel 3 schon benutzt haben.

Definition 5.1 Sei X ein Vektorraum über dem Skalkörper $\mathbb{K}$. Wir nennen

$$X' := \{\lambda : X \rightarrow \mathbb{K} \mid \lambda \text{ linear}\}$$

den **algebraischen Dualraum** von X . Die Elemente von X' werden auch **lineare Funktionale** genannt. $\square$

Schreibweise: Sind $\lambda \in X', x \in X$, so schreiben wir für $\lambda(x)$ meist $\langle \lambda, x \rangle$. Die Abbildung

$$\langle \cdot, \cdot \rangle : X' \times X \ni (\lambda, x) \mapsto \langle \lambda, x \rangle \in \mathbb{K}$$

heißt die (**kanonische**) **duale Paarung**.

Definition 5.2 Sei X ein topologischer Vektorraum über dem Skalkörper $\mathbb{K}$. Wir nennen

$$X^* := \{\lambda : X \rightarrow \mathbb{K} \mid \lambda \text{ linear, stetig}\}$$

den **topologischen Dualraum** von X . $\square$

Klar, stets gilt $\theta \in X^* \subset X'$. Ferner, X', X^* sind selbst wieder Vektorräume über dem Skalkörper $\mathbb{K}$. Wir wissen schon, dass X^* ein vollständiger normierter Raum ist, wobei die Norm durch

$$\|\lambda\|_* := \sup\{|\langle \lambda, x \rangle| \mid \|x\| \leq 1\}, \lambda \in X^*,$$

definiert ist.

Bemerkung 5.3 Sei X ein normierter Raum. Dann ist auch $X \times \mathbb{K}$ (in der Produkttopologie) ein normierter Raum. Wir können $(X \times \mathbb{K})'$, $(X \times \mathbb{K})^*$ mit $X' \times \mathbb{K}$ bzw. $X^* \times \mathbb{K}$ identifizieren vermöge der Abbildungen

$$J : (X \times \mathbb{K})' \longrightarrow X' \times \mathbb{K}, J(\lambda) := (\lambda_1, k) \text{ wobei } \langle \lambda_1, x \rangle := \langle \lambda, (x, 0) \rangle, k := \langle \lambda, (\theta, 1) \rangle,$$

bzw.

$$J : (X \times \mathbb{K})^* \longrightarrow X^* \times \mathbb{K}, J(\lambda) := (\lambda_1, k) \text{ wobei } \langle \lambda_1, x \rangle := \langle \lambda, (x, 0) \rangle, k := \langle \lambda, (\theta, 1) \rangle.$$

Dies sieht man im Fall $(X \times \mathbb{K})^*$ so ein:

J ist linear. J ist auch injektiv, denn aus

$$J(\lambda) = \theta, \text{ d. h. } \langle \lambda, (x, 0) \rangle = 0 \text{ für alle } x \in X, \langle \lambda, (\theta, 1) \rangle = 0,$$

folgt

$$\langle \lambda, (x, r) \rangle = \langle \lambda, (x, 0) \rangle + \langle \lambda, (\theta, r) \rangle = \langle \lambda_1, x \rangle + r \langle \lambda, (\theta, 1) \rangle = 0 \text{ für alle } x \in X, r \in \mathbb{K},$$

und daher $\lambda = \theta$.

J ist surjektiv, denn zu $\langle \lambda_1, x \rangle \in X^* \times \mathbb{K}$ definiert man ein Urbild $\lambda \in (X \times \mathbb{K})^*$ durch $\langle \lambda, (x, r) \rangle := \langle \lambda_1, x \rangle + rk$, $x \in X, k \in \mathbb{K}$.

J ist eine Isometrie, denn

$$\|J(\lambda)\|_{X^* \times \mathbb{K}} = \|(\lambda_1, k)\|_{X^* \times \mathbb{K}} = \|\lambda_1\|_{X^*} + |k|, \|\lambda\|_* = \|\lambda_1\|_{X^*} + |k|.$$

Damit „setzt“ J den Raum $(X \times \mathbb{K})^*$ normerhaltend gleich $X^* \times \mathbb{K}$. □

Beispiel 5.4 Sei $X := C([a, b])$, versehen mit der Maximumsnorm $\|\cdot\|_\infty$. Wir betrachten die Abbildung

$$J : X \longrightarrow X^*, \langle J(x), y \rangle := \int_a^b x(t)y(t)dt, y \in X.$$

Offenbar ist diese Abbildung linear. Sie ist auch stetig, denn

$$\|J(x)\|_* \leq \|x\|_\infty(b-a), x \in X.$$

Aber sie ist nicht surjektiv, denn die Punktauswertungsfunktionale

$$\delta_s : X \ni y \longmapsto y(s) \in \mathbb{R} (s \in [a, b])$$

liegen nicht im Bild von J , wohl aber in X^* . Dies sieht man daran, dass aus der Annahme $\delta_s = J(x)$ für ein $x \in X$ $x(t) = 0$ für $t \neq s$ folgt, im Gegensatz zu $\delta_s \neq \theta$. □

Das obige Beispiel zeigt, dass X^* durchaus sehr reichhaltig sein kann. Diese Reichhaltigkeit haben wir auch schon bei den Folgenräumen $l_p, 1 \leq p \leq \infty$ gesehen. In Banachräumen mit Schauderbasen ergibt sich diese Reichhaltigkeit implizit aus der Stetigkeit der Koeffizientenfunktionale; siehe Satz 4.25. Nun gehen wir der Frage nach, ob dies eine allgemeine Beobachtung ist.

Das Hauptergebnis ist der folgende Fortsetzungssatz, der besagt, dass sich stetige lineare Funktionale von einem beliebigen linearen Teilraum normerhaltend auf den ganzen Raum fortsetzen lassen. Zunächst aber ein Satz über die Fortsetzbarkeit linearer Funktionale, also Stetigkeit der Funktionale ist noch nicht involviert.

Satz 5.5 (Hahn–Banach/reelle Version) Sei X ein Vektorraum über dem Skalkörper $\mathbb{R}$, sei Y ein linearer Teilraum von X , sei $\lambda_0 \in Y'$ und sei $p : X \longrightarrow \mathbb{R}$. Es gelte:

i) $p(x + x') \leq p(x) + p(x')$ für alle $x, x' \in X$.

ii) $p(rx) = rp(x)$ für alle $r \geq 0, x \in X$.

iii) $\langle \lambda_0, y \rangle \leq p(y)$ für alle $y \in Y$.

Dann gibt es $\lambda \in X'$ mit:

(a) $\langle \lambda, y \rangle = \langle \lambda_0, y \rangle$ für alle $y \in Y$, d. h. $\lambda|_Y = \lambda_0$

(b) $\langle \lambda, x \rangle \leq p(x)$ für alle $x \in X$.

Beweis:

Ist $Y = X$, dann ist nichts zu beweisen. Sei also $Y \neq X$.

1. Teil des Beweises.

Sei $x_0 \in X \setminus Y$. Setze $Y_1 := \text{span}(Y \cup \{x_0\})$.

Offenbar gibt es zu jedem $z \in Y_1$ genau ein $y \in Y, r \in \mathbb{R}$ mit $z = rx_0 + y$. Eine Fortsetzung λ_1 von λ_0 auf Y_1 muss also wegen der Forderung der Linearität folgendermaßen operieren:

$$\langle \lambda_1, rx_0 + y \rangle = rc + \langle \lambda_0, y \rangle, r \in \mathbb{R}, y \in Y,$$

wobei $c \in \mathbb{R}$ noch so festzulegen ist, dass (b) gilt. Aus der Voraussetzung über λ_0 folgt für $y_1, y_2 \in Y$:

$$\langle \lambda_0, y_1 \rangle - \langle \lambda_0, y_2 \rangle = \langle \lambda_0, y_1 - y_2 \rangle \leq p(y_1 - y_2) \leq p(y_1 + x_0) + p(-y_2 - x_0), \quad (5.1)$$

d. h.

$$-\langle \lambda_0, y_2 \rangle - p(-y_2 - x_0) \leq -\langle \lambda_0, y_1 \rangle + p(y_1 + x_0).$$

Also

$$m := \sup\{-\langle \lambda_0, y_2 \rangle - p(-y_2 - x_0) | y_2 \in Y\} \leq \inf\{-\langle \lambda_0, y_1 \rangle + p(y_1 + x_0) | y_1 \in Y\} =: M. \quad (5.2)$$

Nun können wir c wählen. Wähle $c \in [m, M]$. Sei $z := rx_0 + y, r \in \mathbb{R}, y \in Y$.

Ist $r > 0$, dann haben wir

$$\begin{aligned} \langle \lambda_1, z \rangle &= rc + \langle \lambda_0, y \rangle = r(c + \langle \lambda_0, r^{-1}y \rangle) \\ &\leq r(-\langle \lambda_0, r^{-1}y \rangle + p(r^{-1}y + x_0) + \langle \lambda_0, r^{-1}y \rangle) \\ &= rp(r^{-1}y + x_0) = p(y + rx_0) = p(z). \end{aligned}$$

Ist $r = 0$, dann haben wir

$$\langle \lambda_1, z \rangle = \langle \lambda_0, y \rangle \leq p(y) = p(z).$$

Ist $r < 0$, dann haben wir mit $\rho := -r$

$$-p(\rho^{-1}y - x_0) - \langle \lambda_0, -\rho^{-1}y \rangle \leq c, \quad -\rho p(\rho^{-1}y - x_0) + \langle \lambda_0, y \rangle \leq \rho c,$$

also

$$\langle \lambda_1, z \rangle = rc + \langle \lambda_0, y \rangle \leq \rho p(\rho^{-1}y - x_0) = p(y + rx_0) = p(z).$$

Wir haben also gezeigt:

$$\langle \lambda_1, z \rangle \leq p(z) \text{ für alle } z \in Y_1.$$

2. Teil. Sei

$$\mathcal{M} := \{(Y_1, \lambda_1) | Y_1 \text{ linearer Teilraum von } X, Y \subset Y_1, \lambda_1 \in Y_1', \lambda_1|_Y = \lambda_0, \langle \lambda_1, z \rangle \leq p(z) \forall z \in Y_1\}.$$

Offenbar ist $(Y, \lambda_0) \in \mathcal{M}$. Eine Halbordnung auf $\mathcal{M}$ wird definiert durch

$$(Y_1, \lambda_1) \preceq (Y_2, \lambda_2) : \Longleftrightarrow Y_1 \subset Y_2, \lambda_2|_{Y_1} = \lambda_1.$$

Ist nun $\hat{\mathcal{M}}$ eine total geordnete Teilmenge von $\mathcal{M}$ bezüglich dieser Halbordnung, so ist offenbar durch

$$\hat{Y} := \cup_{(Y_1, \lambda_1) \in \hat{\mathcal{M}}} Y_1, \hat{\lambda} : \hat{Y} \longrightarrow \mathbb{R}, \hat{\lambda}(y) := \lambda_{Y_1}(y), \text{ falls } y \in Y_1,$$

ein Element $(\hat{Y}, \hat{\lambda}) \in \mathcal{M}$ gegeben, das obere Schranke von $\hat{\mathcal{M}}$ ist. Aus dem Lemma von Zorn folgt: $\mathcal{M}$ enthält ein maximales Element $(\bar{Y}, \bar{\lambda})$. Aus dem 1. Teil folgt $\bar{Y} = X$. ■

Die Bedingungen *i), ii)* in Satz 5.5 werden zusammengefasst unter der Aussage **p ist sublinear**. Für eine sublineare Abbildung p gilt $p(\theta) = 0$, denn mit einem $x \in X$ haben wir $0 = 0 \cdot p(x) = p(0 \cdot x) = p(\theta)$. Wichtig ist auch die Eigenschaft $p(-x) \geq -p(x), x \in X$, für eine sublineare Abbildung; sie folgt aus $0 = p(\theta) = p(x - x) \leq p(x) + p(-x)$. Ein Beispiel für eine sublineare Abbildung ist eine Halbnorm, also eine Normabbildung ohne die Gültigkeit der Definitheitseigenschaft.

Satz 5.6 (Hahn–Banach/komplexe Version) *Sei X ein Vektorraum über dem Skalkörper $\mathbb{C}$, sei Y linearer Teilraum von X , sei $\lambda_0 \in Y'$, und sei $p : X \longrightarrow \mathbb{R}$. Es gelte:*

$$i) \quad p(x + x') \leq p(x) + p(x') \text{ für alle } x, x' \in X.$$

$$ii) \quad p(\alpha x) = |\alpha|p(x) \text{ für alle } \alpha \in \mathbb{C}, x \in X,$$

$$iii) \quad |\langle \lambda_0, y \rangle| \leq p(y) \text{ für alle } y \in Y.$$

Dann gibt es $\lambda \in X'$ mit:

$$(a) \quad \langle \lambda, y \rangle = \langle \lambda_0, y \rangle \text{ für alle } y \in Y, \text{ d. h. } \lambda|_Y = \lambda_0.$$

$$(b) \quad |\langle \lambda, x \rangle| \leq p(x) \text{ für alle } x \in X.$$

Beweis:

X kann als Vektorraum $X_{\mathbb{R}}$ über dem Skalkörper $\mathbb{R}$ aufgefasst werden:

$$X_{\mathbb{R}} := X, \text{ Addition wie in } X, \text{ Multiplikation nur mit reellen Zahlen.}$$

Ebenso verfahren wir mit Y , also haben wir $Y = Y_{\mathbb{R}}$. Sei $\lambda_{0, \mathbb{R}} := \Re \lambda_0$, d. h. $\langle \lambda_{0, \mathbb{R}}, y \rangle := \Re \langle \lambda_0, y \rangle, y \in Y$. Offenbar gilt $\lambda_{0, \mathbb{R}} : Y_{\mathbb{R}} \longrightarrow \mathbb{R}$ linear, $\langle \lambda_{0, \mathbb{R}}, y \rangle \leq p(y)$ für alle $y \in Y_{\mathbb{R}}$. Aus dem Satz von Hahn–Banach in der reellen Version (siehe Satz 5.5) folgt: Es gibt $\lambda_{\mathbb{R}} \in X'_{\mathbb{R}}$ mit

$$\lambda_{\mathbb{R}}|_{Y_{\mathbb{R}}} = \lambda_{0, \mathbb{R}}, \langle \lambda_{\mathbb{R}}, x \rangle \leq p(x) \text{ für alle } x \in X = X_{\mathbb{R}}.$$

Setze $\lambda := \lambda_{\mathbb{R}} - i\lambda_{\mathbb{R}}(i\cdot)$, d. h. $\langle \lambda, x \rangle := \langle \lambda_{\mathbb{R}}, x \rangle - i\langle \lambda_{\mathbb{R}}, ix \rangle, x \in X$. Sei $y \in Y$. Wir haben

$$\begin{aligned} \langle \lambda, y \rangle &= \langle \lambda_{\mathbb{R}}, y \rangle - i\langle \lambda_{\mathbb{R}}, iy \rangle = \langle \lambda_{0, \mathbb{R}}, y \rangle - i\langle \lambda_{0, \mathbb{R}}, iy \rangle \\ &= \Re \langle \lambda_0, y \rangle - i\Re \langle \lambda_0, iy \rangle = \Re \langle \lambda_0, y \rangle - i\Re i \langle \lambda_0, y \rangle \\ &= \Re \langle \lambda_0, y \rangle + i\Im \langle \lambda_0, y \rangle = \langle \lambda_0, y \rangle. \end{aligned}$$

Also gilt $\lambda|_Y = \lambda_0$.

Wir zeigen die Linearität von λ . Seien $x, z \in X$.

$$\begin{aligned} \langle \lambda, x + z \rangle &= \langle \lambda_{\mathbb{R}}, x + z \rangle - i\langle \lambda_{\mathbb{R}}, i(x + z) \rangle \\ &= \langle \lambda_{\mathbb{R}}, x \rangle + \langle \lambda_{\mathbb{R}}, z \rangle - i\langle \lambda_{\mathbb{R}}, ix \rangle - i\langle \lambda_{\mathbb{R}}, iz \rangle \\ &= \langle \lambda, x \rangle + \langle \lambda, z \rangle \end{aligned}$$

Seien $x \in X, \xi = \alpha + i\beta \in \mathbb{C}$:

$$\begin{aligned}
\langle \lambda, \xi x \rangle &= \langle \lambda, \alpha x + i\beta x \rangle = \langle \lambda, \alpha x \rangle + \langle \lambda, i\beta x \rangle \\
&= \langle \lambda, \alpha x \rangle - i\langle \lambda_{\mathbb{R}}, i\alpha x \rangle + \langle \lambda_{\mathbb{R}}, i\beta x \rangle - i\langle \lambda_{\mathbb{R}}, -\beta x \rangle = \alpha(\langle \lambda_{\mathbb{R}}, x \rangle - i\langle \lambda_{\mathbb{R}}, ix \rangle) \\
&= \beta(\langle \lambda_{\mathbb{R}}, ix \rangle + \beta i\langle \lambda_{\mathbb{R}}, x \rangle) \\
&= \alpha(\langle \lambda_{\mathbb{R}}, x \rangle - i\langle \lambda_{\mathbb{R}}, ix \rangle) - i\beta(i\langle \lambda_{\mathbb{R}}, ix \rangle - i\langle \lambda_{\mathbb{R}}, x \rangle) \\
&= \xi \langle \lambda, x \rangle.
\end{aligned}$$

Damit ist nun insgesamt die Linearität von $\lambda: X \rightarrow \mathbb{C}$ gezeigt.

Sei $x \in X$. Wähle $\rho \in \mathbb{R}$ mit $|\langle \lambda, x \rangle| = e^{i\rho} \langle \lambda, x \rangle$. Es folgt:

$$|\langle \lambda, x \rangle| = e^{i\rho} \langle \lambda, x \rangle = \langle \lambda, e^{i\rho} x \rangle = \Re \langle \lambda, e^{i\rho} x \rangle = \langle \lambda_{\mathbb{R}}, e^{i\rho} x \rangle \leq p(e^{i\rho} x) = |e^{i\rho}| p(x) = p(x).$$

Damit erfüllt λ alle geforderten Bedingungen. ■

Die Formulierung des Satzes von Hahn–Banach in der Fassung 5.5 bzw. 5.6 nennt man die **analytische**. Im Zusammenhang mit den Trennungssätzen im Abschnitt 6.2 lernen wir die **geometrischen** Fassungen kennen. Die geometrische Fassung knüpft an an der Beweistechnik zu Satz 5.5.

Bisher haben wir nur Funktionale in X' betrachtet; eine Topologie in X war nicht nötig. Nun wollen wir die Frage nach der Existenz von stetigen linearen Funktionalen, die nicht verschwinden, betrachten. Dabei stellt sich die Frage, ob die Räume X' und X^* wirklich auseinanderfallen. Dass dies im Allgemeinen so ist, zeigt folgendes Beispiel.

Beispiel 5.7 Betrachte $c_f := \{x = (x_k)_{k \in \mathbb{N}} \mid x_k = 0 \text{ für fast alle } k \in \mathbb{N}\}$ als linearen Teilraum des Banachraumes l_2 . Dann gibt es ein algebraisches Komplement (Ergänzung der Hamelbasis in c_f zu einer Hamelbasis von l_2) U von c_f , also $l_2 = c_f \oplus U$. Betrachte damit

$$\lambda: c_f \oplus U \ni x = (x_k)_{k \in \mathbb{N}} + u \mapsto \sum_{k=1}^{\infty} x_k \in \mathbb{K}.$$

Offenbar ist λ linear, aber λ ist nicht stetig, denn $\lim_n \langle \lambda, x^n \rangle = \infty$ und $\|x^n\|_2 \leq 2, n \in \mathbb{N}$, für $x^n := (1, 1/2, 1/3, \dots, 1/n, 0, 0, \dots), n \in \mathbb{N}$. □

Satz 5.8 (Fortsetzungssatz für normierte Räume) Sei $(X, \|\cdot\|)$ ein normierter Raum, sei U ein linearer Teilraum von X und sei $\lambda_0: U \rightarrow \mathbb{K}$ linear und stetig. Dann gibt es $\lambda \in X^*$ mit

$$\lambda|_U = \lambda_0, \|\lambda\|_* = \|\lambda_0\|_*.$$

Beweis:

U ist ein normierter Raum und $\|\lambda_0\|_* = \sup_{u \in U, \|u\| \leq 1} |\langle \lambda_0, u \rangle|$. Definiere eine Halbnorm und damit eine sublineare Abbildung p durch $p(x) := \|\lambda_0\|_* \|x\|, x \in X$. Dann ist offenbar $|\langle \lambda_0, u \rangle| \leq p(u), u \in U$. Nach den obigen Fortsetzungssätzen gibt es ein $\lambda \in X'$ mit $|\langle \lambda, x \rangle| \leq p(x) = \|\lambda_0\|_* \|x\|, x \in X$. λ ist also stetig und $\|\lambda\|_* \leq \|\lambda_0\|_*$. Andererseits ist $|\langle \lambda_0, u \rangle| = |\langle \lambda, u \rangle| \leq \|\lambda\|_* \|u\|, u \in U$, also $\|\lambda_0\|_* \leq \|\lambda\|_*$. Somit gilt $\|\lambda\|_* = \|\lambda_0\|_*$. ■

Folgerung 5.9 Sei $(X, \|\cdot\|)$ ein normierter Raum. Dann gilt:

- (a) Ist $x_0 \in X, x_0 \neq \theta$, dann gibt es $\lambda \in X^*$ mit $\langle \lambda, x_0 \rangle = \|x_0\|, \|\lambda\|_* = 1$.
- (b) Ist U ein linearer abgeschlossener Teilraum von $X, x_0 \in X \setminus U$, dann gibt es ein $\lambda \in X^*$ mit $\langle \lambda, x \rangle = 0, x \in U, \langle \lambda, x_0 \rangle = \text{dist}(x_0, U) := \inf_{u \in U} \|u - x_0\|, \|\lambda\|_* = 1$.

(c) Für alle $x \in X$ ist $\|x\| = \sup_{\lambda \in X^*, \|\lambda\|_* \leq 1} |\langle \lambda, x \rangle|$.

Beweis:

Zu (a)

Sei die sublineare Abbildung p als Norm $\|\cdot\|$ gewählt. Sei $U := \text{span}(\{x_0\})$ und definiere

$$\lambda_0 : U \ni u = ax_0 \mapsto a\|x_0\| \in \mathbb{K}.$$

Offenbar ist λ_0 linear und erfüllt die Abschätzung

$$|\langle \lambda_0, u \rangle| \leq \|u\| = p(u), \quad u \in U.$$

Mit dem Satz von Hahn–Banach (reelle oder komplexe Version) erhalten wir die Existenz von $\lambda \in X'$ mit

$$\langle \lambda, u \rangle = \langle \lambda_0, u \rangle, \quad u \in U, \quad |\langle \lambda, x \rangle| \leq \|x\|, \quad x \in X.$$

Da die Norm stetig ist, ist λ stetig in θ und daher überall stetig. Wegen $\langle \lambda, x_0 \rangle = \|x_0\| \neq 0$ gilt $\|\lambda\|_* = 1$.

Zu (b)

Da U abgeschlossen ist, ist $\delta := \text{dist}(x_0, U) > 0$. Sei $V := \text{span}(\{x_0\} \cup U)$. Wir definieren $\lambda_0 \in U'$ durch

$$\langle \lambda_0, ax_0 + u \rangle := a\delta, \quad a \in \mathbb{K}, \quad u \in U.$$

Offenbar ist $\langle \lambda_0, u \rangle = 0, u \in U$, und $\langle \lambda_0, x_0 \rangle = \delta$. λ ist stetig, denn für $y := ax_0 - u, a \in \mathbb{K} \setminus \{0\}, u \in U$, haben wir

$$\|y\| = \|a(x_0 - a^{-1}u)\| = |a|\|x_0 - a^{-1}u\| \geq |a|\delta = |\langle \lambda_0, y \rangle|,$$

also $|\langle \lambda_0, y \rangle| \leq \|y\|$. Da diese Abschätzung auch trivialerweise für den Fall $a = 0$ gilt, ist $\|\lambda_0\|_* \leq 1$.

Zu $\varepsilon > 0$ gibt es ein $u \in U$ mit $\|x_0 - u\| \leq \delta + \varepsilon$; setze $z := \|x_0 - u\|^{-1}(x_0 - u)$. Wir erhalten

$$\|\lambda_0\|_* \geq |\langle \lambda_0, z \rangle| = \|x_0 - u\|^{-1}\delta \geq (\delta + \varepsilon)^{-1}\delta.$$

Im Grenzübergang $\varepsilon \rightarrow 0$ folgt $\|\lambda_0\|_* \geq 1$. Also ist $\|\lambda_0\|_* = 1$. Nun setzen wir λ_0 normerhaltend von V auf X fort; siehe Satz 5.8.

Zu (c)

Sei $x \in X$; o. E. $x \neq \theta$. Offenbar gilt $\sup_{\lambda \in X^*, \|\lambda\|_* \leq 1} |\langle \lambda, x \rangle| \leq \|x\|$. Aus (a) folgt $\sup_{\lambda \in X^*, \|\lambda\|_* \leq 1} |\langle \lambda, x \rangle| \geq \|x\|$. ■

Wir können (a) in Folgerung 5.9 so interpretieren, dass X^* die Punkte in X trennt, denn ist $x \neq x'$, dann gibt es $\lambda \in X^*$ mit $\langle \lambda, x \rangle \neq \langle \lambda, x' \rangle$.

Folgerung 5.10 Sei $(X, \|\cdot\|)$ ein normierter Raum. Dann ist X separabel, wenn X^* separabel ist.

Beweis:

Es gibt nach Voraussetzung eine abzählbare Menge $A = \{\alpha_n | n \in \mathbb{N}\} \subset X^*$, die dicht in $S := \{\lambda \in X^* | \|\lambda\|_* = 1\}$ ist. Zu jedem $n \in \mathbb{N}$ gibt es $x_n \in X$ mit $\|x_n\| = 1$ und $|\langle \alpha_n, x_n \rangle| \geq \frac{1}{2}\|\alpha_n\|_*$. Annahme: $X \neq U := \overline{\text{span}(x_1, x_2, \dots)}$. Sei etwa $x_0 \in X \setminus U$. Dann gibt es nun $\lambda \in S$ mit $\langle \lambda, x \rangle = 0$ für alle $x \in U$; siehe Folgerung 5.9 (b). Wähle α_n mit $\|\alpha_n - \lambda\|_* < \frac{1}{3}\|\lambda\|_*$, was wegen der Dichtheit von A möglich ist. Dann ergibt

$$\frac{1}{2}\|\lambda\|_* - \frac{1}{2}\|\alpha_n - \lambda\|_* \leq \frac{1}{2}\|\alpha_n\|_* \leq |\langle \alpha_n, x_n \rangle| = |\langle \alpha_n - \lambda, x_n \rangle| \leq \|\alpha_n - \lambda\|$$

einen Widerspruch. Also gilt $X = U$ und X ist separabel. ■

Lineare stetige Funktionale sind ein nützliches Hilfsmittel, eine Fragestellung in einem normierten Raum X in den zugehörigen Skalkörper $\mathbb{K}$ zu übersetzen. Ein Beispiel dafür haben wir in Satz 3.37 kennengelernt im Zusammenhang mit der Konvergenz unbedingter Reihen. Hier skizzieren wir ein weiteres Beispiel.

Definition 5.11 Sei X ein normierter Raum mit Skalkörper $\mathbb{C}$, sei D eine offene Teilmenge von $\mathbb{C}$ und sei $f : D \rightarrow X$ eine (vektorwertige) Funktion.

- (a) f heißt **schwach holomorph**, falls $\lambda \circ f : D \rightarrow \mathbb{C}$ holomorph für alle $\lambda \in X^*$ ist.
- (b) f heißt **stark holomorph**, falls

$$\lim_{h \rightarrow 0} \frac{f(z+h) - f(z)}{h} \text{ existiert für alle } z \in D$$

gilt, wobei der Grenzwert genommen wird bezüglich der Norm-Konvergenz.

Satz 5.12 (Liouville) Sei X ein normierter Raum mit Skalkörper $\mathbb{C}$ und sei $f : \mathbb{C} \rightarrow X$ schwach holomorph. Ist f beschränkt, dann ist f konstant.

Beweis:

Die Abbildung $\lambda \circ f : \mathbb{C} \rightarrow \mathbb{C}$ ist beschränkt für jedes $\lambda \in X^*$. Nach dem klassischen Satz von Liouville ist $\lambda \circ f : \mathbb{C} \rightarrow \mathbb{C}$ konstant für jedes $\lambda \in X^*$. Dann gilt also

$$\langle \lambda, f(z) \rangle = \langle \lambda, f(0) \rangle \text{ für alle } z \in \mathbb{C}, \lambda \in X^*.$$

Da X^* die Punkte trennt, muss $f(z) = f(0)$ gelten für alle $z \in \mathbb{C}$. ■

5.2 Reflexivität

Nun können wir für normierte Räume nach dem Dualraum des Banachraumes X^* fragen.

Definition 5.13 Sei X ein normierter Raum. Der Dualraum $(X^*)^*$ von X^* heisst **Bidualraum** und wird mit X^{**} bezeichnet. □

Sei $(X, \|\cdot\|)$ ein normierter Raum. Die folgende Überlegung zeigt, dass man X stets als einen linearen Teilraum von X^{**} auffassen kann. Für $x \in X$ ist nämlich durch

$$J_X(x)(\lambda) := \langle \lambda, x \rangle, \lambda \in X^*,$$

ein lineares Funktional $J_X(x)$ in $(X^*)'$ gegeben, welches wegen

$$|J_X(x)(\lambda)| = |\langle \lambda, x \rangle| \leq \|\lambda\|_* \|x\|, x \in X,$$

stetig ist. Also ist $J_X(x) \in X^{**}$ für alle $x \in X$. Die hierdurch definierte Abbildung $J_X : X \rightarrow X^{**}$ ist offensichtlich linear. Aus der Definition der Operatornorm folgt

$$\|J_X(x)\|_{X^* \rightarrow \mathbb{K}} = \sup\{|\langle \lambda, x \rangle| : \|\lambda\|_* \leq 1\} \leq \|x\|, x \in X. \quad (5.3)$$

In (5.3) besteht sogar Gleichheit nach Folgerung 5.9 (c). Also haben wir folgenden Satz bewiesen:

Satz 5.14 Sei $(X, \|\cdot\|)$ ein normierter Raum. Dann ist die Abbildung

$$J_X : X \rightarrow X^{**}, J_X(x)(\lambda) := \langle \lambda, x \rangle, \lambda \in X^*, x \in X,$$

eine lineare, normerhaltende (isometrische), injektive Abbildung.

Beweis:

Es ist schon alles nachgerechnet. ■

Man bezeichnet J_X aus Satz 5.14 die **kanonische Einbettung** von X in seinen Bidualraum X^{**} . Sie ist im Allgemeinen nicht surjektiv.

Definition 5.15 Sei $(X, \|\cdot\|)$ ein normierter Raum. X heisst **reflexiv**, wenn die kanonische Einbettung J_X surjektiv ist. □

Da der Bidualraum X^{**} stets vollständig ist, ist eine notwendige Bedingung für Reflexivität die Vollständigkeit von X . Klar, endlichdimensionale normierte Räume sind reflexiv, da wir ja schon aus der linearen Algebra wissen, dass der (algebraische) Dualraum mit sich selbst identifiziert werden kann. Diese Aussage werden wir später auf Hilberträume ausweiten.

Beispiel 5.16 Die Räume $l_p, 1 < p < \infty$ sind reflexiv. Dies folgt aus der Tatsache, dass wir wissen, dass l_p^* isometrisch isomorph zu l_q ist mit $1/p + 1/q = 1$. Dieselbe Aussage gilt auch für die Räume $L_p(\Omega), 1 < p < \infty$. $l_1, L_1(\Omega)$ und $l_\infty, L_\infty(\Omega)$ sind nicht reflexiv, denn ihre Dualräume sind $l_\infty, L_\infty(\Omega)$ bzw. $l_\infty^* \neq l_1, L_\infty(\Omega)^* \neq L_1(\Omega)$. $c_0, c, C[a, b]$ sind nicht reflexiv (betrachtet in ihrer üblichen Norm). □

Satz 5.17 Jeder abgeschlossene Teilraum eines reflexiven Banachraumes ist ein reflexiver Banachraum.

Beweis:

Sei U ein abgeschlossener Unterraum des reflexiven Raums X . Offenbar ist U selbst wieder ein Banachraum. Zu $\mu \in U^{**}$ definiere $\mu_U : X^* \rightarrow \mathbb{K}$ durch

$$\langle \mu_U, \lambda \rangle = \langle \mu, \lambda|_U \rangle, \lambda \in X^*.$$

μ_U ist offenbar linear mit $\|\mu_U\|_{X^{**}} \leq \|\mu\|_{U^{**}}$. Da X reflexiv ist, gibt es eine $x \in X$ mit

$$\langle \lambda, x \rangle = \langle \mu_U, \lambda \rangle = \langle \mu, \lambda|_U \rangle, \lambda \in X^*.$$

Annahme: $x \notin U$.

Dann können wir nach 5.9 (b) ein $\lambda \in X^*$ finden mit $\lambda|_U = \theta$ und $\langle \lambda, x \rangle \neq 0$. Dies ist ein Widerspruch.

Jedes $\rho \in U^*$ kann mit dem Fortsetzungssatz 5.8 zu einem $\lambda \in X^*$ fortgesetzt werden. Dann haben wir wegen $x \in U$ $\langle \lambda, x \rangle = \langle \rho, x \rangle$ und daher

$$\langle \mu_U, \rho \rangle = \langle \mu, \lambda|_U \rangle = \langle \mu|_U, \lambda \rangle = \langle \lambda, x \rangle = \langle \rho, x \rangle.$$

Damit ist $J_U : U \rightarrow U^{**}$ surjektiv. ■

Satz 5.18 Ein Banachraum X ist reflexiv genau dann, wenn X^* reflexiv ist.

Beweis:

Sei X reflexiv. Dann ist die kanonische Einbettung $J_X : X \rightarrow X^{**}$ ein isometrischer Isomorphismus. Sei $\rho \in X^{***}$. Wir haben $\mu \in X^*$ zu finden mit

$$\langle \rho, \lambda \rangle = \langle \lambda, \mu \rangle \text{ für alle } \lambda \in X^{**}.$$

Betrachte $\mu := \rho \circ J_X$. Dann ist $\mu \in X^*$. Zu $\lambda \in X^{**}$ gibt es auf Grund der Reflexivität von X ein $x \in X$ mit $\lambda = J_X(x)$. Dann haben wir

$$\langle \rho, \lambda \rangle = \langle \rho, J_X(x) \rangle = \langle \rho \circ J_X, x \rangle = \langle \mu, x \rangle = \langle J_X(x), \mu \rangle = \langle \lambda, \mu \rangle.$$

Sei X^* reflexiv. Annahme: X ist nicht reflexiv.

Dann ist $U := J_X(X)$ ist ein echter Teilraum von X^{**} , der auch abgeschlossen ist, da X ein Banachraum und J_X eine isometrische Abbildung ist. Dann gibt es nach Folgerung 5.9 (b) ein $\rho \in X^{***}$ mit $\rho|_U = \theta, \rho \neq \theta$. Da X^* reflexiv ist, gibt es $\mu \in X^*$ mit $\rho = J_{X^*}(\mu)$. Dann haben wir

$$0 = \langle \rho, J_X(x) \rangle = \langle J_{X^*}(\mu), J_X(x) \rangle = \langle J_X(x), \mu \rangle = \langle \mu, x \rangle \text{ für alle } x \in X.$$

Also ist $\mu = \theta$ und damit auch ρ , was im Widerspruch zu $\rho \neq \theta$ ist. ■

5.3 Schwache Topologien

Sei X ein normierter Vektorraum. Neben der damit vorliegenden Topologie können wir mit Hilfe des Dualraums eine weitere Topologie definieren. Dies machen wir, indem wir die Familie

$$(U_{\varepsilon, M})_{\varepsilon > 0, M \subset X^*: \#M < \infty} \text{ mit } U_{\varepsilon, M} := \{x \in X \mid |\langle \lambda, x \rangle| < \varepsilon \text{ für alle } \lambda \in M\}$$

betrachten und sie als **Nullumgebungsbasis** verwenden; durch Verschiebung haben wir dann insgesamt eine Umgebungsbasis. Die offenen Mengen ergeben sich dann so:

$$U \subset X \text{ offen} : \iff \forall u \in U \exists U_{\varepsilon, M} \text{ mit } u + U_{\varepsilon, M} \subset U.$$

Es liegt offenbar ein System offener Mengen vor, die eine Topologie auf X bilden. Die damit definierte Topologie auf X nennen wir die **schwache Topologie** und bezeichnen sie auch als $\sigma(X, X^*)$ -Topologie.

Dieses Konstruktionsprinzip könnten wir nun auf dem Dualraum X^* analog wiederholen (zumindest im Fall, dass X ein normierter Raum ist) und somit die $\sigma(X^*, X^{**})$ -Topologie auf X^* erhalten. Eine größere Bedeutung spielt aber eine durch X auf X^* induzierte Topologie. Wir betrachten dazu die Familie

$$(V_{\varepsilon, M})_{\varepsilon > 0, M \subset X: \#M < \infty} \text{ mit } V_{\varepsilon, M} := \{\lambda \in X^* \mid |\langle \lambda, x \rangle| < \varepsilon \text{ für alle } x \in M\}.$$

Diese Familie verwenden wir als **Nullumgebungsbasis** in X^* ; durch Verschiebung haben wir dann insgesamt eine Umgebungsbasis. Die offenen Mengen ergeben sich dann so:

$$V \subset X^* \text{ offen} : \iff \forall \lambda \in V \exists V_{\varepsilon, M} \text{ mit } \lambda + V_{\varepsilon, M} \subset V.$$

Es liegt offenbar ein System offener Mengen vor, die eine Topologie auf X^* bilden. Die damit definierte Topologie auf X^* nennen wir die **schwach*-Topologie** und bezeichnen sie auch als $\sigma(X^*, X)$ -Topologie. Klar, wenn X reflexiver normierter Raum ist, dann stimmt die $\sigma(X^*, X)$ -Topologie mit der $\sigma(X^*, X^{**})$ -Topologie überein.

Bemerkung 5.19 Sei X ein topologischer Vektorraum. Die $\sigma(X, X^*)$ -Topologie ist die größte Topologie auf X , so dass alle $\lambda \in X^*$ stetig sind.

Die $\sigma(X^*, X)$ -Topologie ist die größte Topologie auf X^* , so dass die Abbildungen $\lambda \mapsto \langle \lambda, x \rangle$ für alle $x \in X$ stetig sind. □

Von den Bezügen dieser Topologien untereinander handelt dieser Abschnitt im Folgenden. Die Begriffe *offen*, *abgeschlossen*, *(folgen-)kompakt*, *konvergent*, ... können sich nun auf diese verschiedenen Topologien beziehen.

Die $\sigma(X, X^*)$ -Topologie ist schwächer als die Norm-Topologie/starke Topologie in X , die $\sigma(X^*, X)$ -Topologie ist schwächer als die Norm-Topologie in X^* . Das 1. Abzählbarkeitsaxiom gilt für die $\sigma(X, X^*)$ -Topologie genau dann, wenn X endlichdimensional ist, also, wenn $\sigma(X, X^*)$ -Topologie und Norm-Topologie in X übereinstimmen. Dies hat zur Konsequenz, dass

im Allgemeinen (Überdeckungs-)Kompaktheit und Folgenkompaktheit nicht zusammenfallen. Eine damit verwandte Fragestellung ist, ob sich die abgeschlossene Einheitskugel in X bzw. X^* in der $\sigma(X, X^*)$ - bzw. $\sigma(X, X^*)$ -Topologie durch eine Metrik erzeugen lässt. Dies ist in einem separablen Banachraum X möglich in X^* und in einen Banachraum X , wenn X^* separabel ist. Alle diese Anmerkungen lassen sich mit Beispielen in den Folgenräumen belegen oder widerlegen.

Beispiel 5.20 Sei X ein unendlichdimensionaler Banachraum. Die Beweise der folgenden Fakten überlassen wir dem Leser.

Die Einheitssphäre $S_1 := \{x \in X \mid \|x\| = 1\}$ ist zwar stark abgeschlossen, aber nicht schwach abgeschlossen, denn der schwache Abschluss von S_1 ist die abgeschlossene Einheitskugel $\overline{B}_1$.

B_1 ist nicht offen in der schwachen Topologie, denn die Menge der inneren Punkte von B_1 in der schwachen Topologie ist leer. $\square$

Definition 5.21 Sei X ein topologischer Vektorraum.

(a) Eine Folge $(x^n)_{n \in \mathbb{N}}$ **konvergiert schwach** gegen $x \in X$, wenn es zu jedem $\varepsilon > 0$ und zu jedem $M \subset X^*$ mit $\#M < \infty$ ein $N \in \mathbb{N}$ gibt mit $x^n - x \in U_{\varepsilon, M}$ für alle $n \geq N$.

Wir schreiben $x = w - \lim_n x^n$ (oder $x^n \rightharpoonup x$).

(b) Eine Folge $(\lambda^n)_{n \in \mathbb{N}}$ **konvergiert schwach***¹ gegen $\lambda \in X$, wenn es zu jedem $\varepsilon > 0$ und zu jedem $M \subset X$ mit $\#M < \infty$ ein $N \in \mathbb{N}$ gibt mit $\lambda^n - \lambda \in V_{\varepsilon, M}$ für alle $n \geq N$.

Wir schreiben $\lambda = w^* - \lim_n \lambda^n$ (oder $\lambda^n \xrightarrow{*} \lambda$).

$\square$

Aus der Definition der Mengen $U_{\varepsilon, M}, V_{\varepsilon, M}$, insbesondere der Tatsache, dass hierbei M endlich ist, folgt, dass wir die Konvergenz etwas einfacher beschreiben können.

Folgerung 5.22 Sei X ein normierter Vektorraum, $(x^n)_{n \in \mathbb{N}}$ eine Folge in X , $(\lambda^n)_{n \in \mathbb{N}}$ eine Folge in X^* . Dann haben wir für $x \in X, \lambda \in X^*$:

(a) $x = w - \lim_n x^n$ genau dann, wenn $\langle \lambda, x \rangle = \lim_n \langle \lambda, x^n \rangle$ für alle $\lambda \in X^*$.

(b) $\lambda = w^* - \lim_n \lambda^n$ genau dann, wenn $\langle \lambda, x \rangle = \lim_n \langle \lambda^n, x \rangle$ für alle $x \in X$.

Beweis:

Schauen wir uns etwa (a) an. Es gelte $x^n \rightharpoonup x$. Wähle zu $\varepsilon > 0, \lambda \in X^*$, die Umgebung $x + U_{\varepsilon, \{\lambda\}}$ von x . Man erhält $N \in \mathbb{N}$ mit $x^n \in x + U_{\varepsilon, \{\lambda\}}$ für alle $n \geq N$. Dies zeigt $\langle \lambda, x \rangle = \lim_n \langle \lambda, x^n \rangle$. Alle anderen Aussagen beweist man genauso leicht. $\blacksquare$

Bemerkung 5.23 Die schwache Abgeschlossenheit oder die schwache Vollständigkeit eines normierten Raumes ist im Allgemeinen nicht klar und sogar ein Banachraum muss nicht schwach vollständig sein. Ein solches Beispiel ist $X := c_0$. Betrachte dazu in c_0 die Folge $(x^n)_{n \in \mathbb{N}}$ mit

$$x^n := (1, \dots, 1, 0, 0, \dots), n \in \mathbb{N}.$$

Für jedes $\lambda \in l_1 = c_0^*, \lambda = (y_k)_{k \in \mathbb{N}}$, gilt

$$|\langle \lambda, x^n - x^m \rangle| \leq \sum_{k=\min(n,m)+1}^{\max(n,m)} |y_k| \leq \sum_{k=\min(n,m)+1}^{\infty} |y_k|, m, n \in \mathbb{N},$$

und dies zeigt, dass $(x^n)_{n \in \mathbb{N}}$ eine schwache Cauchyfolge ist. Deren Kandidat für den Grenzwert ist die konstante Folge $(1, \dots, 1, \dots)$, die aber nicht in c_0 sondern in $l_\infty = l_1^* = (c_0)^{**}$ liegt.

Gegen die konstante Folge $(1, \dots, 1, \dots)$ konvergiert $(x^n)_{n \in \mathbb{N}}$ schwach* in l_∞ . Daher konvergiert auch jede Teilfolge davon gegen die konstante Folge schwach* in l_∞ . Die in c_0 beschränkte Folge $(x^n)_{n \in \mathbb{N}}$ hat also offenbar keine schwach konvergente Teilfolge. $\square$

¹Sprich „schwach-Stern“

Bemerkung 5.24 Eine „Besonderheit“ bezüglich schwacher Konvergenz liegt in l_1 vor. Das Lemma von Schur besagt nämlich, dass die Norm-Konvergenz äquivalent mit der schwachen Konvergenz. Wir verweisen für den Beweis auf [37]. $\square$

Folgerung 5.25 Sei X ein normierter Vektorraum. Dann sind sowohl die $\sigma(X, X^*)$ - als auch die $\sigma(X^*, X)$ -Topologie eine Hausdorff-Topologie.

Beweis:

Wir beweisen nur die erste Aussage, die zweite Aussage beweist man analog. Seien $x, y \in X, x \neq y$. Nach Folgerung 5.9 (a) gibt es ein $\lambda \in X^*$ mit $\langle \lambda, x - y \rangle = \|x - y\| \neq 0$, d. h. $\langle \lambda, x \rangle \neq \langle \lambda, y \rangle$. Sei $\varepsilon := \frac{1}{2} |\langle \lambda, x \rangle - \langle \lambda, y \rangle|$. Dann sind

$$\frac{-1}{\lambda} ((\langle \lambda, x \rangle - \varepsilon, \langle \lambda, x \rangle + \varepsilon)), \quad \frac{-1}{\lambda} ((\langle \lambda, y \rangle - \varepsilon, \langle \lambda, y \rangle + \varepsilon)),$$

offene Mengen bezüglich der $\sigma(X, X^*)$ -Topologie, die die Punkte x, y trennen. $\blacksquare$

Aus Lemma 5.25 wissen wir, dass die Limiten bezüglich der $\sigma(X, X^*)$ -Topologie eindeutig bestimmt sind. Um den Unterschied zwischen der schwach mit dem Fortsetzungssatz 5.8 zu einem $\lambda \in X^*$ fortgesetzt werden. Dann haben wir und der schwach*-Konvergenz herauszuarbeiten, betrachten wir den Raum l_1 , der Dualraum von c_0 ist und l_∞ zum Dualraum hat. Für die Elemente $(e^i)_{i \in \mathbb{N}}$ der Schauderbasis in c_0 erhalten wir für jedes $x = (x_k)_{k \in \mathbb{N}}$ mit den Koeffizientenfunktionalen $(\lambda_k)_{k \in \mathbb{N}}$

$$\langle \lambda_k, x \rangle = x_k, k \in \mathbb{N},$$

und daher, dass $(\lambda_k)_{k \in \mathbb{N}}$ in $l_1 = c_0^*$ schwach* gegen Null konvergiert. Aber für $x = (1, 1, 1, \dots) \in l_\infty$ gilt $\langle \lambda_k, x \rangle = 1, k \in \mathbb{N}$. Da der schwache und der schwach*-Grenzwert übereinstimmen müssen, sofern sie existieren, ist die Folge $(\lambda_k)_{k \in \mathbb{N}}$ nicht schwach konvergent in l_1 .

Folgerung 5.26 Sei X ein normierter Vektorraum. Dann gilt:

- (a) Eine Menge $M \subset X$ ist beschränkt genau dann, wenn $\sup_{x \in M} |\langle \lambda, x \rangle| < \infty$ (schwache Beschränktheit) gilt für alle $\lambda \in X^*$.
- (b) Konvergiert die Folge $(x^k)_{k \in \mathbb{N}}$ schwach gegen $x \in X$, dann gilt $\|x\| \leq \inf_k \|x^k\|$.
- (c) $\overline{B}_r(x)$ ist $\sigma(X, X^*)$ -folgenabgeschlossen für alle $x \in X, r \geq 0$.
- (d) Eine Menge $M' \subset X^*$ ist beschränkt genau dann, wenn $\sup_{\lambda \in M'} |\langle \lambda, x \rangle| < \infty$ gilt für alle $x \in X$.

Beweis:

Zu (a) Dass aus der Beschränktheit die schwache Beschränktheit folgt, ist offensichtlich. Sei M nun schwach beschränkt, also $\sup_{x \in M} |\langle \lambda, x \rangle| < \infty$ für alle $\lambda \in X^*$. Wir interpretieren diese Bedingung in X^{**} : also $\sup_{x \in M} |J_X(x)(\lambda)| < \infty$ für alle $\lambda \in X^*$. Nach dem Prinzip der gleichmäßigen Beschränktheit ist $\{J_X(x) | x \in M\}$ in X^{**} beschränkt. Da J_X eine Isometrie ist nach Satz 5.14, ist auch M beschränkt.

Zu (b) Folgt aus (a), denn eine schwach konvergente Folge ist schwach beschränkt.

Zu (c) Folgt aus (b).

Zu (d) Folgt direkt aus dem Prinzip der gleichmäßigen Beschränktheit. $\blacksquare$

Satz 5.27 Sei X ein Banachraum und sei $(\lambda_k)_{k \in \mathbb{N}}$ eine Folge in X^* . Dann konvergiert die Folge $(\lambda_k)_{k \in \mathbb{N}}$ schwach* genau dann, wenn folgende Bedingungen erfüllt sind:

- (a) $\sup_{k \in \mathbb{N}} \|\lambda_k\|_* < \infty$;

(b) $(\langle \lambda_k, x \rangle)_{k \in \mathbb{N}}$ ist eine Cauchyfolge für alle $x \in A \subset X$ mit einer in X dichten Teilmenge A .

Beweis:

Dies ist eine Anwendung von Satz 4.6. ■

Satz 5.28 *Ist X ein endlichdimensionaler Banachraum, dann stimmen starke und schwache Topologie überein.*

Beweis:

Man führe eine Basis ein. Bezüglich der schwachen Topologie sind die Koeffizientenfunktionale bezüglich dieser Basis stetig, daher auch bezüglich der Norm-Topologie. ■

Wir wissen, dass die abgeschlossene Einheitskugel eines normierten Raums X genau dann kompakt ist in der Norm-Topologie, wenn die Dimension von X endlich ist; siehe Satz 3.15. Im Fall $\dim X = \infty$ ist auch $\dim X^* = \infty$. Daher ist in unendlich-dimensionalen normierten Räumen auch die Einheitskugel B_1 in X^* niemals kompakt in der Norm-Topologie. Die folgenden Sätze liefern uns jedoch die Folgenkompaktheit in der schwach*-Topologie in Banachräumen, der erste Satz unter der Voraussetzung der Separabilität von X , der zweite Satz ohne diese Voraussetzung.

Satz 5.29 *Sei X ein separabler Banachraum. Dann ist die abgeschlossene Einheitskugel $\overline{B}_1$ in X^* schwach*-folgenkompakt.*

Beweis:

Sei $(\lambda_k)_{k \in \mathbb{N}}$ eine norm-beschränkte Folge in X^* . Dann ist auch $(\langle \lambda_k, x \rangle)_{k \in \mathbb{N}}$ eine beschränkte Folge in $\mathbb{K}$ für alle $x \in X$. Da X separabel ist, gibt es eine abzählbare in X dichte Teilmenge A . Durch Auswahl einer Diagonalfolge (siehe Anmerkungen nach Satz 2.39) kann erreicht werden, dass – ohne die Notation zu ändern – $(\langle \lambda_k, x \rangle)_{k \in \mathbb{N}}$ eine Cauchyfolge ist für alle $x \in A$. Nun kann Satz 5.27 angewendet werden. ■

Eine Erweiterung ist

Satz 5.30 (Satz von Alaoglu-Bourbaki) *Sei $(X, \|\cdot\|)$ ein Banachraum. Dann ist die abgeschlossene Einheitskugel $\overline{B}_1$ in X^* in der $\sigma(X^*, X)$ -Topologie kompakt.*

Beweis:

Wir betrachten mit dem Produktraum $Y := \mathbb{K}^X = \prod_{x \in X} \mathbb{K}$ die Abbildung

$$\Phi : X^* \longrightarrow Y, \lambda \longmapsto (\langle \lambda, x \rangle)_{x \in X}.$$

Wir statteten hierbei X^* mit der schwach*-Topologie $\sigma(X^*, X)$ und Y mit der Produkttopologie aus und zeigen, dass gilt:

- (a) Φ ist linear und bijektiv; Φ und Φ^{-1} sind stetig.
- (b) $\Phi(B_1)$ ist kompakt in Y .

Zu (a). Für jedes $x \in X$ ist $\lambda \longmapsto \langle \lambda, x \rangle$ stetig. Hieraus folgt die Stetigkeit von Φ auf Grund der Definition der Produkt-Topologie. Ferner ist Φ offensichtlich linear und injektiv. Also existiert $\Phi^{-1} : \Phi(X^*) \longrightarrow X^*$. Es bleibt zu zeigen, dass Φ^{-1} ebenfalls stetig ist². Dies folgt aber unmittelbar aus der Tatsache, dass für jedes feste $x \in X$ die Abbildung $y = (y_x)_{x \in X} \longmapsto$

²Beachte, dass wir hier den Satz über die stetige Inverse nicht anwenden können, da wir nicht im Kontext von Banachräumen sind.

$\langle \Phi^{-1}(y), x \rangle = y_x$ auf $\Phi(X^*)$ stetig ist.

Zu (b). Es gilt $\Phi(B_1) = K := K_1 \cap K_2$ mit

$$K_1 := \{y = (y_x)_{x \in X} \in Y \mid |y_x| \leq \|x\|\} = \prod_{x \in X} [-\|x\|, \|x\|]$$

$$K_2 := \{y = (y_x)_{x \in X} \in Y \mid y_{x+z} = y_x + y_z, y_{ax} = ay_x \text{ für alle } x, z \in X, a \in \mathbb{K}\}$$

K_1 ist als Produkt kompakter Mengen auch kompakt. K_2 ist abgeschlossen, denn K_2 lässt sich darstellen in der Gestalt

$$K_2 = \left(\bigcap_{x, z \in X} A_{x, z} \right) \cap \left(\bigcap_{x \in X, a \in \mathbb{K}} C_{x, a} \right)$$

mit den abgeschlossenen Mengen

$$A_{x, z} := \{y = (y_x)_{x \in X} \in Y \mid y_{x+z} = y_x + y_z\}, \quad C_{x, a} := \{y = (y_x)_{x \in X} \in Y \mid y_{ax} = ay_x\}$$

für $x, z \in X, a \in \mathbb{K}$. Die Abgeschlossenheit ist hier eine einfache Konsequenz aus der Stetigkeit der Abbildungen $y = (y_x)_{x \in X} \mapsto y_{x+z} - y_x - y_z$ und $y = (y_x)_{x \in X} \mapsto y_{ax} - ay_x$. Also ist K als abgeschlossene Teilmenge der kompakten Menge K_1 kompakt. ■

Da kompakte Mengen in unendlichdimensionalen normierten Räumen eher rar sind, der Wechsel zu der schwächeren schwach*-Topologie aber in vielen Situationen keine große Einschränkung bedeutet bzw. diese Topologie auf natürlichem Wege ins Spiel kommt, gibt uns der Satz 5.30 eine Fülle „neuer“ kompakter Mengen an die Hand.

Eine Anwendung, bei der der Satz von Alaoglu-Borbaki hilfreich ist, ist etwa in folgender Situation gegeben:

$$X = l_1, X^* = l_\infty.$$

Wir wissen ja, dass X, X^* nicht reflexiv sind, die Kompaktheit ihrer Einheitskugeln (und damit der beschränkten Mengen) im Zweifel ist. Mit dem Resultat in Satz 5.30 haben wir aber ein Kompaktheitsinstrument in l_∞ zur Hand. Beachte, für l_1 können wir so nicht schließen, da ja l_1 nicht Dualraum von l_∞ ist.

Satz 5.31 *Ist X ein reflexiver Banachraum, so ist die abgeschlossene Einheitskugel $\overline{B}_1$ in $\sigma(X, X^*)$ -folgenkompakt.*

Beweis:

Sei $(x_k)_{k \in \mathbb{N}}$ eine Folge in der abgeschlossenen Einheitskugel $\overline{B}_1$ von X . Der Teilraum $U := \text{span}(x_1, x_2, \dots)$ ist dann ein separabler reflexiver Banachraum nach Satz 5.17. Es gilt also $J_U(U) = U^{**}$ und U^{**} ist separabel. Mit U^{**} ist nach Folgerung 5.10 auch U^* separabel. Also enthält die Folge $(J_U(x_k))_{k \in \mathbb{N}}$ in U^{**} eine gegen $z \in U^{**}$ $\sigma(U^{**}, U^*)$ -konvergente Teilfolge; o. E. sei diese nicht neu bezeichnet. Da J_U ein isometrischer Isomorphismus ist, folgt $\lim_k \langle \mu, x_k \rangle = \langle \mu, z \rangle$ für alle $\mu \in U^*$. Da $x_k \in U, k \in \mathbb{N}, J_U^{-1}(z) \in U$, folgt $\lim_k \langle \lambda, x_k \rangle = \langle \lambda, J_U^{-1}(z) \rangle$ für alle $\lambda \in X^*$. Also ist $J_U^{-1}(z) = w - \lim_k x_k$. ■

Der Satz von Eberlein–Smulian (siehe etwa [37]) ist eine Verschärfung des obigen Satzes 5.31. Er besagt, dass die Reflexivität auch notwendig ist für die schwache Folgenkompaktheit der abgeschlossenen Einheitskugel. Wir verzichten auf dieses Resultat.

Das Ergebnis von Satz 5.31 ist enorm hilfreich, ersetzt es doch die Folgenkompaktheit der abgeschlossenen Einheitskugel in endlichdimensionalen Räumen (Satz von Heine-Borel) in unendlichdimensionalen Räumen.

Die starke Topologie in einem Banachraum X mit Norm $\|\cdot\|$ ist metrisierbar, d.h. die Topologie wird durch eine Metrik erzeugt, nämlich durch $d := \|\cdot - \cdot\|$. Wie sieht es mit dieser Eigenschaft bezüglich der schwachen Topologien aus? Hier sind Antworten für unendlichdimensionale Banachräume X ohne Belege. $(X, \sigma(X, X^*))$, $(X^*, \sigma(X^*, X))$ sind nie metrisierbar. Die (Relativ-)Topologie $\sigma(X, X^*)$ auf der abgeschlossenen Einheitskugel $\overline{B}_1$ in X kann metrisierbar sein oder nicht. Die (Relativ-)Topologie $\sigma(X^*, X)$ auf der abgeschlossenen Einheitskugel $\overline{B}_1$ in X^* ist metrisierbar genau dann, wenn X separabel ist.

5.4 Duale Operatoren

Seien X, Y normierte Räume und sei $T : X \rightarrow Y$ linear und stetig. Ist $\lambda \in Y^*$, so wird durch $T'\lambda := \lambda \circ T$ offenbar ein stetiges lineares Funktional in X^* erzeugt. Die so eingeführte Abbildung

$$T^* : Y^* \rightarrow X^*, T^*\lambda := \lambda \circ T, \lambda \in Y^*, \quad (5.4)$$

ist offensichtlich linear.

Definition 5.32 Seien X, Y normierte Räume und sei $T : X \rightarrow Y$ linear und stetig. Die durch (5.4) definierte Abbildung T^* heisst der zu T **duale Operator**. $\square$

Folgerung 5.33 Seien X, Y normierte Räume und sei $T : X \rightarrow Y$ linear und stetig. Dann ist T^* stetig und es gilt $\|T^*\| = \|T\|$.

Beweis:

Es gilt

$$\sup_{\|\lambda\|_* \leq 1} \|T^*\lambda\| = \sup_{\|\lambda\|_* \leq 1} \sup_{\|x\| \leq 1} |(T^*\lambda)(x)| = \sup_{\|x\| \leq 1} \sup_{\|\lambda\|_* \leq 1} |\lambda(Tx)| = \sup_{\|x\| \leq 1} \|Tx\| = \|T\|.$$

Für die vorletzte Gleichheit siehe Folgerung 5.9 (c). $\blacksquare$

Folgerung 5.34 Seien X, Y normierte Räume. Die Zuordnung $T \rightarrow T^*$ definiert eine isometrische lineare Abbildung von $\mathcal{B}(X, Y)$ nach $\mathcal{B}(Y^*, X^*)$.

Beweis:

Die Linearität folgt aus der Definition, die Isometrie-eigenschaft aus Folgerung 5.33. $\blacksquare$

Seien X, Y normierte Räume. Zweimaliges Anwenden der Dualisierung liefert zu gegebenem $T \in \mathcal{B}(X, Y)$ einen Operator $T^{**} \in \mathcal{B}(X^{**}, Y^{**})$. Man rechnet einfach nach, dass

$$T^{**} \circ J_X = J_Y \circ T \quad (5.5)$$

gilt; dabei sind J_X, J_Y die kanonischen Einbettungen in die Bidualräume.

Folgerung 5.35

- (a) Sind X, Y, Z normierte Räume und ist $T \in \mathcal{B}(X, Y), S \in \mathcal{B}(Y, Z)$, so ist $(S \circ T)^* = T^* \circ S^*$.
- (b) Sind X, Y normierte Räume und ist $T \in \mathcal{B}(X, Y)$ bijektiv, so ist auch T^* bijektiv und $(T^*)^{-1} \in \mathcal{B}(Y^*, X^*), (T^*)^{-1} = (T^{-1})^*$.

Beweis:

Zu (a) Folgt mit $\lambda \in Z^*$ aus

$$\langle \lambda, S \circ Tx \rangle = \langle S^* \lambda, Tx \rangle = \langle T^* \circ S^* \lambda, x \rangle.$$

Zu (b) Ist T stetig und bijektiv, dann ist auch T^{-1} stetig und es gilt $\text{id}_X = T^{-1} \circ T$, $\text{id}_Y = T \circ T^{-1}$. Offenbar ist $\text{id}_X^* = \text{id}_{X^*}$. Nun folgt aus (a)

$$\text{id}_{X^*} = T^* \circ (T^{-1})^*, \text{id}_{Y^*} = (T^{-1})^* \circ T^*.$$

Aus der ersten Gleichung folgt $\text{ran}(T^*) = X^*$, aus der zweiten Gleichung folgt $\ker(T^*) = \{\theta\}$. Damit ist T^* bijektiv und es gilt $(T^*)^{-1} = (T^{-1})^*$. ■

Lemma 5.36 Seien X, Y Banachräume, sei $T : X \longrightarrow Y$ linear und stetig. Dann ist T^* schwach folgenstetig.

Beweis:

Sei die Folge $(x^n)_{n \in \mathbb{N}}$ eine gegen x schwach konvergente Folge in X . Sei $\lambda \in Y^*$. Dann gilt

$$\langle \lambda, Tx \rangle = \langle T^* \lambda, x \rangle = \lim_n \langle T^* \lambda, x^n \rangle = \lim_n \langle \lambda, \lambda T x^n \rangle.$$

Daraus liest man die Behauptung ab. ■

Beispiel 5.37 Betrachte auf dem Banachraum c_0 die Abbildung

$$T : c_0 \ni x = (x_k)_{k \in \mathbb{N}} \longmapsto y := \left(\frac{1}{k} x_k\right)_{k \in \mathbb{N}} \in c_0.$$

Offenbar ist T linear und stetig. Wir wissen, dass $c_0 = l_1$ gilt. Also ist für $\lambda = (y_k)_{k \in \mathbb{N}} \in l_1$ wegen

$$\langle T^* \lambda, x \rangle = \langle \lambda, Tx \rangle = \sum_{k=1}^{\infty} y_k \frac{1}{k} x_k$$

$$T^* \lambda = \left(\frac{1}{k} y_k\right)_{k \in \mathbb{N}}.$$

□

5.5 Übungen

- 1.) Seien X, Y Mengen. Eine Abbildung $\Delta : \text{POT}(X) \longrightarrow \text{POT}(Y)$ heisst eine **Dualität**, wenn

$$\Delta(\cup_{i \in I} A_i) = \cap_{i \in I} \Delta(A_i)$$

gilt für alle Familien $(A_i)_{i \in I}$ in $\text{POT}(X)$. Zeige für eine Dualität:

- (a) $A \subset B \implies \Delta(A) \supset \Delta(B)$;
- (b) $\Delta(A) = \cap \{\Delta(\{x\}) \mid x \in A\}$;
- (c) $A \subset \cup \{D \mid \Delta(D) \subset \Delta(A)\}$.

- 2.) Seien X, Y Mengen und sei $\Delta : \text{POT}(X) \longrightarrow \text{POT}(Y)$ eine Dualität; siehe oben. Definiere

$$\Delta^* : \text{POT}(Y) \longrightarrow \text{POT}(X), \Delta^*(A) := \cup \{D \mid D \subset \Delta(A)\}.$$

- (a) Zeige: Δ^* ist eine Dualität mit $\Delta^{**} := (\Delta^*)^* = \Delta$.
- (b) Berechne Δ^*, Δ^{**} für $\Delta(A) := X \setminus A$.

- 3.) Sei X eine Menge. Eine Abbildung $H : \text{POT}(X) \rightarrow \text{POT}(X)$ heisst **Hüllenoperator**, falls gilt:

$$A \subset H(A), H(A) = H(H(A)), A \subset B \implies H(A) \subset H(B).$$

- (a) Finde mindestens zwei Beispiele für Hüllenoperatoren.
 (b) Ist $\Delta : \text{POT}(X) \rightarrow \text{POT}(Y)$ eine Dualität, dann ist $\Delta^* \circ \Delta$ ein Hüllenoperator.
- 4.) Sei $X := \{f \in C[0, 1] \mid f(1) = 0\}$ versehen mit der Maximumsnorm $\|\cdot\|_\infty$.
- (a) Zeige: X ist damit ein Banachraum.
 (b) Betrachte $\lambda : X \ni f \mapsto \int_0^1 t f(t) dt \in \mathbb{R}$. Zeige: $\lambda \in X^*$, $\|\lambda\| = \frac{1}{2}$, und es gibt kein $f \in X$ mit $\|f\|_\infty = 1, \lambda(f) = \frac{1}{2}$.
- 5.) Sei $X := l_2$ und sei $T : X \rightarrow X$ linear und stetig. Zeige: Es gibt genau eine „Matrix“ $A = (a_{ij})_{i,j \in \mathbb{N}}$ mit

$$Tx = \left(\sum_{k \in \mathbb{N}} a_{ik} x_k \right)_{i \in \mathbb{N}}, x = (x_n)_{n \in \mathbb{N}}.$$

- 6.) Seien X, Y normierte Räume und T eine lineare Abbildung von X nach Y . Man zeige, dass T genau dann stetig ist, wenn $T \circ g \in X^*$ für alle $g \in Y^*$ gilt.
- 7.) Sei X ein normierter Raum über dem Skalkörper $\mathbb{C}$ und sei $X_{\mathbb{R}}$ derselbe Raum X aufgefasst als reeller normierter Raum. Dann ist

$$H : (X_{\mathbb{R}})^* \rightarrow X^*, \lambda \mapsto \lambda(\cdot) - i\lambda(i\cdot),$$

ein $\mathbb{R}$ -linearer isometrischer Isomorphismus.

- 8.) Gib Isometrien S, T an mit

$$S : c^* \rightarrow l_1, T : c_0 \rightarrow l_1.$$

- 9.) Seien X, Y normierte Räume und sei $L : X \rightarrow Y$ linear. Beweise die Äquivalenz von
- (a) L ist stetig (in der Normtopologie).
 (b) Für jede Folge in X , die schwach gegen θ konvergiert, konvergiert auch die Bildfolge schwach gegen θ .
- 10.) Sei $(X, \|\cdot\|)$ ein Banachraum und sei Y ein Teilraum von X . Zeige: Ist X reflexiv und Y abgeschlossen, dann ist auch $(Y, \|\cdot\|)$ reflexiv.
- 11.) Sei $(X, \|\cdot\|)$ ein separabler Banachraum. Dann ist die abgeschlossene Einheitskugel $\overline{B}_1$ in X^* schwach*-folgenkompakt.
- 12.) Sei X ein reeller normierter Raum, sei $(x^k)_{k \in \mathbb{N}}$ eine Folge in X , sei $(a_k)_{k \in \mathbb{N}}$ eine Folge in $\mathbb{R}$ und sei $c > 0$. Zeige die Äquivalenz von
- (a) Es gibt ein $\lambda \in X^*$ mit $\|\lambda\|_* \leq c, \langle \lambda, x^i \rangle = a_i, i \in I$.
 (b) $|\sum_{i=1}^n \beta_i a_i| \leq c \|\sum_{i=1}^n \beta_i x^i\|$ für alle $n \in \mathbb{N}, \beta_1, \dots, \beta_n \in \mathbb{R}$.
- 13.) Betrachte auf $(l_\infty, \|\cdot\|_\infty)$ die Abbildungen

$$p_1(x) := \limsup_n x_n, x = (x_n)_{n \in \mathbb{N}}$$

$$p_2(x) := \limsup_n \sigma(x) \text{ wobei } \sigma(x) = 1/n \sum_{j=1}^n x_j, x = (x_n)_{n \in \mathbb{N}}.$$

Zeige: p_1, p_2 sind sublinear.

- 14.) Sei X ein reeller Vektorraum. Setze $\Gamma := \{p|p \text{ sublinear}\}$ und definiere

$$p \preceq q : \Longleftrightarrow p(x) \leq q(x), x \in X, p, q \in \Gamma.$$

Zeige für $p \in \Gamma$ die Äquivalenz von

- (a) p ist linear, d. h. $p \in X'$.
 - (b) p ist minimales Element von Γ , d. h. ist $q \preceq p, q \in \Gamma$, dann gilt $q = p$.
- 15.) Sei X ein reeller Vektorraum und sei $p : X \rightarrow \mathbb{R}$ sublinear. Dann existiert ein $\lambda \in X'$ mit $\langle \lambda, x \rangle \leq p(x)$ für alle $x \in X$.
Hinweis: Verwende obige Übungsaufgabe.

- 16.) Sei X ein Banachraum. Beweise oder widerlege durch ein Gegenbeispiel

- (a) Ist $\lambda = \lim_n \lambda_n$ in X^* und $x = \lim_n x_n$ in X , so gilt $\langle \lambda, x \rangle = \lim_n \langle \lambda_n, x_n \rangle$.
 - (b) Ist $\lambda = \lim_n \lambda_n$ in X^* und $x = w - \lim_n x_n$ in X , so gilt $\langle \lambda, x \rangle = \lim_n \langle \lambda_n, x_n \rangle$.
 - (c) Ist $\lambda = w^* - \lim_n \lambda_n$ in X^* und $x = \lim_n x_n$ in X , so gilt $\langle \lambda, x \rangle = \lim_n \langle \lambda_n, x_n \rangle$.
 - (d) Ist $\lambda = w^* - \lim_n \lambda_n$ in X^* und $x = w - \lim_n x_n$ in X , so gilt $\langle \lambda, x \rangle = \lim_n \langle \lambda_n, x_n \rangle$.
- 17.) Betrachte den Banachraum l_2 mit der Schauder-Basis $(e^i)_{i \in \mathbb{N}}$; $e_k^i = \delta_{ik}$, $k \in \mathbb{N}$. Sei $M := \{x^{k,l} := e^k + (k-1)e^l | 1 \leq k < l\}$.
- (a) Bestimme $M_w := \{x | x = w - \lim_k x_k, (x_k)_{k \in \mathbb{N}} \in l_2, x_k \in M \text{ für alle } k \in M\}$.
 - (b) Zeige $\theta \in M$.

- 18.) Betrachte den Banachraum l_∞ und sei $M := c_f := \{x = (x_k)_{k \in \mathbb{N}} | x_k = 0 \text{ für fast alle } k \in \mathbb{N}\}$. Zeige:

- (a) $\overline{M} = l_\infty$ in der Norm-Topologie von l_∞ .
 - (b) Bestimme $M_w := \{x | x = w^* - \lim_k x_k, (x_k)_{k \in \mathbb{N}} \in l_2, x_k \in M \text{ für alle } k \in M\}$, wobei l_∞ als Dualraum von l_1 zu verstehen ist.
- 19.) Seien X, Y Banachräume und seien $T : X \rightarrow Y, S : Y^* \rightarrow X^*$ lineare Operatoren mit

$$\langle \lambda, Tx \rangle = \langle S\lambda, x \rangle, x \in X, \lambda \in Y^*.$$

Dann sind T, S stetig und $S = T^*$.

- 20.) Sei $X := c_0$. dann haben wir $X^* = l_1, X^{**} = l_\infty$. Zeige, dass

$$\lambda : l_1 \ni x = (x_k)_{k \in \mathbb{N}} \mapsto \sum_{k=1}^{\infty} x_k \in \mathbb{K}$$

stetig ist, aber λ nicht stetig ist bezüglich der w^* -Topologie, induziert durch l_∞ .

- 21.) Sei X ein Banachraum und sei $T : X \rightarrow X$ stetig. Zeige: Aus $x = w - \lim_n x^n$ folgt $Tx = w - \lim_n Tx$.

- 22.) Betrachte $S_n : X \ni x = (x_1, x_2, \dots, x_{n+1}, x_n, \dots) \mapsto (x_{n+1}, x_{n+2}, \dots) \in X, R_n : X \ni x = (x_1, x_2, \dots, x_{n+1}, x_n, \dots) \mapsto (\underbrace{0, \dots, 0}_{n\text{-mal}}, x_1, x_2, \dots) \in X$. Zeige für $x \in l_2$:

- (a) $w - \lim_n S_n x = \theta$, aber $(S_n x)_{n \in \mathbb{N}}$ konvergiert nicht der Norm nach.
 - (b) $w - \lim_n R_n x = \theta$, aber $(R_n x)_{n \in \mathbb{N}}$ konvergiert nicht der Norm nach.
- 23.) Sei X ein Banachraum. Dann sind äquivalent:
- (a) X separabel und reflexiv.

- (b) X^* separabel und reflexiv.
- 24.) Sei X ein reflexiver Banachraum und sei $(e^n)_{n \in \mathbb{N}}$ eine Schauderbasis in X . Zeige: Die Koeffizientenfunktionale $(\lambda_n)_{n \in \mathbb{N}}$ bilden eine Schauderbasis in X^* .
- 25.) Sei X ein normierter Raum und sei $U \subset X$ ein linearer Teilraum. Betrachte die Einbettung $j : U \ni u \mapsto u \in X$. Zeige die Äquivalenz von
- (a) $j^* : X^* \rightarrow U^*$ ist injektiv.
 - (b) $j^* : X^* \rightarrow U^*$ ist isometrischer Isomorphismus.
 - (c) U ist dicht in X .
- 26.) Sei X ein Banachraum, sei $(x^n)_{n \in \mathbb{N}}$ eine dichte Teilmenge in X und sei $B := \overline{B}_1$ die abgeschlossene Einheitskugel in X^* . Betrachte die Abbildung

$$d : B \times B \ni (\lambda, \mu) \mapsto \sum_{n=1}^{\infty} 2^{-n} \frac{|\langle \lambda, x^n \rangle - \langle \mu, x^n \rangle|}{1 + |\langle \lambda, x^n \rangle - \langle \mu, x^n \rangle|} \in \mathbb{R}.$$

Zeige:

- (a) d ist eine Metrik auf B .
- (b) d erzeugt die $\sigma(X^*, X)$ -(Relativ-)Topologie auf B .

Kapitel 6

Konvexität

Was ist rund?

Titel eines Manuskripts von B. Kawohl

Aus dem Satz von Hahn–Banach ziehen wir nun (überwiegend) geometrische Folgerungen unter Heranziehung des Konzepts der Konvexität. Konvexität kann als Bindeglied zwischen linearer und nichtlinearer Funktionalanalysis betrachtet werden. In der Optimierungstheorie ist Konvexität von großer Bedeutung hinsichtlich Existenz und Charakterisierung von Lösungen und algorithmischen Betrachtungen. Eine Kurzzusammenfassung erschließt (zum Teil im Anhang) die geometrischen Eigenschaften der Norm. Die Prähilberträume stellen diesbezüglich eine interessante Klasse von normierten Räumen dar. Wir betrachten in diesem Kapitel fast ausschließlich reelle Vektorräume.

6.1 Konvexe Mengen

Definition 6.1 Sei X ein Vektorraum.

- (a) Jedes $z \in X$ mit $z = y + a(x - y) = ax + (1 - a)y$, $a \in [0, 1]$, heißt **Konvexkombination** von $x, y \in X$.
- (b) $M \subset X$ heißt **konvex**, falls für alle $x, y \in M$ und $a \in [0, 1]$ gilt: $ax + (1 - a)y \in M$.

Klar, jeder lineare Teilraum und jeder affine Unterraum eines Vektorraums ist konvex.

Beispiel 6.2 Sei $(X, \|\cdot\|)$ ein normierter Raum. Die offenen Kugeln $B_r(x^0)$ und die abgeschlossenen Kugeln $\overline{B}_r(x^0)$ sind konvex. Dies folgt mit der Dreiecksungleichung. Hier ist der Beweis zu $\overline{B}_r(x^0)$.

Seien $x, y \in \overline{B}_r(x^0)$, d. h. $\|x - x^0\| \leq r$, $\|y - x^0\| \leq r$. Sei $a \in [0, 1]$.

$$\begin{aligned}\|ax + (1 - a)y - x^0\| &= \|a(x - x^0) + (1 - a)(y - x^0)\| \leq \|a(x - x^0)\| + \|(1 - a)(y - x^0)\| \\ &= |a|\|x - x^0\| + |1 - a|\|y - x^0\| \leq ar + (1 - a)r = r.\end{aligned}$$

Dies zeigt: $ax + (1 - a)y \in \overline{B}_r(x^0)$. □

Lemma 6.3 Seien X, Y Vektorräume, sei $T : X \rightarrow Y$ linear, seien $K_1, K_2 \subset X, K_3 \subset Y$ konvex und seien $a, b \in \mathbb{R}$. Dann gilt

- (a) $K_1 \cap K_2$ ist konvex.

(b) $K_1 \times T(K_2) := \{(u, v) \in X \times Y | u \in K_1, v \in T(K_1)\}$ ist konvex (in $X \times Y$).

(c) $\overline{T}^{-1}(K_3) = \{u \in X | T(u) \in K_3\}$ ist konvex.

(d) $aK_1 + bK_2 := \{au + bv | u \in K_1, v \in K_2\}$ ist konvex.

Beweis:

Einfach nachzurechnen. ■

Lemma 6.4 Sei $(X, \|\cdot\|_X)$ ein normierter Raum und sei $K \subset X$ konvex. Dann ist auch $K(\epsilon) := K + \overline{B}_\epsilon := \{u \in X | \exists x \in K \text{ mit } \|u - x\|_X \leq \epsilon\}$, der Abschluss von K und das Innere von K konvex.

Beweis:

Trivial ■

Beispiel 6.5 Sei X ein normierter Raum mit topologischem Dualraum X^* . Sei $\lambda \in X^*, r \in \mathbb{R}$. Dann sind

$$H_{\lambda,r}^+ := \{x \in X | \langle \lambda, x \rangle \geq r\}, H_{\lambda,r}^- := \{x \in X | \langle \lambda, x \rangle \leq r\},$$

$$H_{\lambda,r} := \{x \in X | \langle \lambda, x \rangle = r\}, H_{\lambda,r}^+ \setminus H_{\lambda,r}, H_{\lambda,r}^- \setminus H_{\lambda,r}$$

konvex. Dies folgt im wesentlichen aus der Linearität von $\langle \cdot, \cdot \rangle$ im 2. Argument.

Die Mengen

$$H_{\lambda,r}^+, H_{\lambda,r}^- \text{ bzw. } H_{\lambda,r}^+ \setminus H_{\lambda,r}, H_{\lambda,r}^- \setminus H_{\lambda,r}$$

werden **Halbräume** bzw. **offene Halbräume**, $H_{\lambda,r}$ wird **Hyperebene** genannt.

Wegen

$$H_{\lambda,r} = \overline{\lambda}^{-1}(\{r\}), H_{\lambda,r}^- = \overline{\lambda}^{-1}((-\infty, r]), H_{\lambda,r}^+ = \overline{\lambda}^{-1}([r, \infty))$$

sind $H_{\lambda,r}, H_{\lambda,r}^-, H_{\lambda,r}^+$ $\sigma(X, X^*)$ -abgeschlossen. □

Beispiel 6.6 Sei X ein Vektorraum und sei $A \subset X$. Setze

$$A' := \{\lambda \in X' | \langle \lambda, u \rangle \leq 1 \text{ für alle } u \in A\}.$$

Dann ist A' eine konvexe Teilmenge von X' . A' heißt die zu A **polare Menge**. □

Ist eine Teilmenge A eines Vektorraumes X nicht konvex, so kann man versuchen, eine „kleinste“ (im Sinne der Inklusion) konvexe Menge K zu finden, die A enthält. Da der Raum X sicherlich konvex ist, macht dies Sinn.

Definition 6.7 Sei X ein Vektorraum und sei $A \subset X$. Die Menge

$$K := \bigcap \{C | C \supset A, C \text{ konvex}\}$$

heißt **konvexe Hülle** von A . Wir schreiben

$$K = \text{co}(A).$$

Eine naheliegende Verallgemeinerung der Regel (a) aus Lemma 6.3 für beliebige Durchschnitte belegt, dass $\text{co}(A)$ stets konvex ist. Offenbar gilt:

$$\text{co}(A_1) \subset \text{co}(A_2), \text{ falls } A_1 \subset A_2 \subset X.$$

Folgerung 6.8 Sei X ein Vektorraum und sei $A \subset X$. Dann gilt

$$\text{co}(A) = \{x \in X \mid x = \sum_{i=1}^n a_i u^i, u^1, \dots, u^n \in A, a_1, \dots, a_n \in [0, 1], \sum_{i=1}^n a_i = 1, n \in \mathbb{N}\}$$

Beweis:

Sei $K := \{x \in X \mid x = \sum_{i=1}^n a_i u^i, u^1, \dots, u^n \in A, a_1, \dots, a_n \in [0, 1], \sum_{i=1}^n a_i = 1, n \in \mathbb{N}\}$.

K ist offenbar konvex und A ist in K enthalten. Daraus folgt sofort $\text{co}(A) \subset K$. Ist $x \in K$, dann ist offenbar x in jeder konvexen Menge $C \subset X$ mit $A \subset C$; daher $x \in \text{co}(A)$. Also gilt $\text{co}(A) = K$. ■

Beispiel 6.9 Wir betrachten die Menge $S := \{(e^{\sigma(1)} \mid \dots \mid e^{\sigma(n)}) \in \mathbb{R}^{n,n} \mid \sigma \in \mathcal{S}_n\}$ ($\mathcal{S}_n$ ist die Menge der Permutationen von $1, 2, \dots, n$)

Wir wissen $\#S = n!$. Setze

$$\mathcal{D} := \text{co}(S).$$

Man rechnet nach, dass für eine Matrix $D = (d_{ij}) \in \mathcal{D}$ gilt:

$$\sum_{k=1}^n d_{kj} = \sum_{k=1}^n d_{ik} = 1, \quad d_{ij} \geq 0, \quad 1 \leq i, j \leq n.$$

Die Matrizen in $\mathcal{D}$ nennt man **doppeltstochastische Matrizen**. □

Satz 6.10 (Satz von Carathéodory (1911)/Steinitz (1914)) Sei X ein n -dimensionaler Vektorraum und sei $A \subset X, A \neq \emptyset$. Dann gilt

$$\text{co}(A) = \left\{ x = \sum_{i=1}^m a_i x^i \mid \sum_{i=1}^m a_i = 1, x^i \in A, a_i \in [0, 1], 1 \leq i \leq m, m \leq n+1 \right\}$$

Beweis:

O. E. sei $X = \mathbb{R}^n$.

Sei $K := \{x = \sum_{i=1}^m a_i x^i \mid x^i \in A, a_i \in [0, 1], 1 \leq i \leq m, \sum_{i=1}^m a_i = 1, m \leq n+1\}$.

Sei $x \in \text{co}(A)$, d. h. $x = \sum_{i=1}^k a_i x^i, x^i \in A, a_i \in [0, 1], 1 \leq i \leq k$, und $\sum_{i=1}^k a_i = 1$. Ist $k \leq n+1$, ist $x \in K$ und nichts ist mehr zu zeigen. Sei also nun $k > n+1$. Da $\dim X = n$ und $k > n+1$ ist, ist $\text{rang}(M) \leq n+1$, wobei $M := \begin{pmatrix} x^1 & \dots & x^k \\ 1 & \dots & 1 \end{pmatrix} \in \mathbb{R}^{(n+1) \times k}$ ist. Also gibt es $b_1, \dots, b_k \in \mathbb{R}$ mit

$$\sum_{i=1}^k b_i x^i = \theta, \quad \sum_{i=1}^k b_i = 0, \quad \sum_{i=1}^k b_i^2 \neq 0.$$

Setze

$$Q := \{q \in \mathbb{R} \mid qb_i + a_i \geq 0, 1 \leq i \leq k\}$$

Q ist abgeschlossen, nicht-leer, da $0 \in Q$, und verschieden von $\mathbb{R}$, da nicht alle b_i verschwinden. Sei $q \in Q$, so dass $qb_j + a_j = 0$ für mindestens ein $j \in \{1, \dots, k\}$ ist. Ein solches q existiert, da $Q \neq \mathbb{R}$ und Q abgeschlossen ist. Dann ist

$$x = \sum_{i=1}^k a_i x^i + q \sum_{i=1}^k b_i x^i = \sum_{i=1}^k (a_i + qb_i) x^i = \sum_{i \neq j} (a_i + qb_i) x^i$$

und wir haben eine Darstellung von x als Konvexkombination durch $k-1$ Vektoren, da $\sum_{i \neq j} (a_i + qb_i) = 1$ und daher $a_i + qb_i \in [0, 1], 1 \leq i \leq k$, gilt. ■

Den obigen Satz von Carathéodory kann man sich an einem Dreieck in der Ebene gut veranschaulichen. Etwa erhält man den Schwerpunkt $\bar{x}$ eines Dreiecks mit den Endpunkten

$$P_1 : a^1, P_2 : a^2, P_3 : a^3$$

(Koordinaten a^i bzgl. der Standardbasis) als

$$\bar{x} = \frac{1}{3}a^1 + \frac{1}{3}a^2 + \frac{1}{3}a^3$$

(Je nach Betrachtungsweise, kann man dies auch als definierende Gleichung für den Schwerpunkt ansehen.)

Aus dem Satz von Carathéodory schließt man sehr einfach, dass $\text{co}(A)$ beschränkt (kompakt) ist, falls A beschränkt (kompakt) ist.

Definition 6.11 Sei X ein Vektorraum und seien $x^0, \dots, x^k \in X$ affin linear unabhängig, d. h. $x^1 - x^0, \dots, x^k - x^0$ linear unabhängig. Dann nennen wir

$$\Sigma(x^0, \dots, x^k) := \text{co}(\{x^0, \dots, x^k\})$$

einen k – Simplex. □

Bemerkung 6.12 Mit dem Konvexitätsbegriff für Mengen können wir schnell auch die Konvexität von Funktionen erklären. Sei X ein reeller Vektorraum und sei $f : D \rightarrow \mathbb{R}$ eine Funktion, $D \subset X$ konvex. Wir sagen f ist **konvex**, falls

$$\text{epi}(f) := \{(x, r) \in D \times \mathbb{R} \mid f(x) \leq r\}$$

konvex ist. Die Menge $\text{epi}(f)$ heißt der **Epigraph** von X .

Wir schreiben die Konvexität von f äquivalent auf: f ist konvex genau dann, wenn für alle $x, y \in D$ und für alle $a \in [0, 1]$ gilt:

$$f(ax + (1 - a)y) \leq af(x) + (1 - a)f(y).$$

Beispielsweise ist die Funktion $\mathbb{R}^n \ni x \mapsto \langle x, Ax \rangle \in \mathbb{R}$ konvex genau dann, wenn $A \in \mathbb{R}^{n,n}$ positiv semidefinit ist. □

Beispiel 6.13 Sei $D \subset \mathbb{R}$ konvex und sei $f : D \rightarrow \mathbb{R}$ konvex. Dann ist f stetig. Seien $c, a, b, d \in D$ mit $c < a < b < d$. Dann ist f auf dem Intervall $[a, b]$ Lipschitzstetig, also insbesondere stetig, mit Lipschitzkonstante

$$L := \max\left\{\frac{f(a) - f(c)}{a - c}, \frac{f(d) - f(b)}{d - b}\right\}.$$

Dies folgt aus der Tatsache, dass die Abbildung

$$(x, y) \mapsto \frac{f(x) - f(y)}{x - y}, \quad x, y \in D, x \neq y,$$

monoton nichtfallend ist in jedem Argument. Dies nachzurechnen überlassen wir dem Leser. □

Betrachte folgende Aufgabe:

Minimiere $\langle c, x \rangle$
unter den Nebenbedingungen $x \in \mathbb{R}^n, Ax = b, x \geq \theta,$

Diese Aufgabe wird als „Standardaufgabe der Linearen Optimierung“ bezeichnet; andere Aufgaben der Linearen Optimierung lassen sich in diese Standardaufgabe umschreiben. Lineare Optimierung ist Kern des Fachgebiets *Operations Research*. Hier sind typische Anwendungen der linearen Optimierung: **Transportprobleme** (Kostenminimierung beim Transport von Gütern zu den Märkten), **Rundreiseprobleme** (Kostenminimierung bei der Ausarbeitung einer Tour durch eine vorgegebene Anzahl von Städten). Letzteres Problem ist ein Problem der **ganzzahligen Optimierung**. Das *Simplexverfahren* ist ein effizienter Algorithmus zur Lösung solcher Aufgaben. Es basiert auf der Tatsache, dass die zulässige Menge

$$\{x \in \mathbb{R}^n | Ax = b, x \geq \theta\}$$

„Ecken“ besitzt – im Zusammenhang mit Extrempunkten kommen wir darauf zurück – und dass eine Lösung der Aufgabe stets auch in einer Ecke angenommen wird.

6.2 Trennungssätze

Mit Hilfe des algebraischen Dualraums haben wir oben Halbräume und Hyperebenen erklärt. Verwenden wir dabei stetige lineare Funktionale in einem topologischen Vektorraum erhalten wir abgeschlossene Halbräume. Mit solchen Halbräumen wollen wir konvexe Mengen trennen. Wir streben dabei nicht die allgemeinsten Resultate an. Zunächst eine Vorbereitung.

Lemma 6.14 *Sei X ein normierter Raum und sei $U \subset X$ eine konvexe Nullumgebung. Definiere damit das so genannte **Minkowski-Funktional**:*

$$p(x) := p_U(x) := \inf\{r \geq 0 | x \in rU\}, \quad x \in X. \quad (6.1)$$

Dann gilt:

- (a) $p(x)$ ist definiert für alle $x \in X$.
- (b) $p(x) \geq 0, x \in X, p(\theta) = 0$.
- (c) $p(ax) = ap(x), x \in X, a \in [0, \infty)$.
- (d) $p(x + x') \leq p(x) + p(x'), x, x' \in X$.
- (e) $\{x \in X | p(x) < 1\} \subset U \subset \{x \in X | p(x) \leq 1\}$.
- (f) $U^\circ = \{x \in X | p(x) < 1\}$.
- (g) $X \setminus \overline{U} = \{x \in X | p(x) > 1\}$.

Beweis:

(a), (b) sind unmittelbar klar. Zu (c)

$$\begin{aligned} p(ax) &= \inf\{r \geq 0 | ax \in rU\} = \inf\{ar' \geq 0 | x \in r'U\} \\ &= a \inf\{r' \geq 0 | x \in r'U\} = ap(x). \end{aligned}$$

Zu (d) Sei $\varepsilon > 0$. Sei $s := p(x) + \varepsilon, s' := p(x') + \varepsilon$. Dann ist $x \in sU, x' \in s'U$, und da U konvex ist, folgt

$$(s + s')^{-1}(x + x') = (s + s')^{-1}(ss^{-1}x + ss'^{-1}x') \in U.$$

Also gilt $p((s + s')^{-1}(x + x')) \leq 1$, und wir erhalten

$$p(x + x') = (s + s')p((s + s')^{-1}(x + x')) \leq s + s' = p(x) + p(x') + 2\varepsilon.$$

Da $\varepsilon > 0$ beliebig ist, ist die Behauptung bewiesen.

Zu (e)

Sei $p(x) < 1$. Dann gibt es ein $r \in (0, 1)$ mit $x \in rU$. Wegen $\theta \in U$, $x = rr^{-1}x + (1-r)\theta$ und da U konvex ist, ist $x \in U$. Wenn $x \in U$ ist, dann ist $p(x) \leq 1$ und damit folgt $U \subset \{x \in X | p(x) \leq 1\}$.

Zu (f) Die Aussage $U^\circ \subset \{x \in X | p(x) \leq 1\}$ ist wegen (e) klar.

Annahme: $x \in U^\circ$ und $p(x) = 1$. Dann gibt es, da x ein innerer Punkt von U ist, $\delta > 0$ mit $x + \delta x \in U$. Also $x \in (1 + \delta)^{-1}U$, und damit $p(x) \leq (1 + \delta)^{-1} < 1$, also ein Widerspruch.

Zu (g) Die Aussage $\bar{U} \subset \{x \in X | p(x) \geq 1\}$ ist wegen (e) klar.

Annahme: $x \in X \setminus \bar{U}$ und $p(x) = 1$. Dann gibt es eine monoton fallende Folge $s_n)_{n \in \mathbb{N}}$ in $(1, \infty)$ mit $\lim_n s_n = 1$, $s_n^{-1}x \in U$. Dann ist $x \in \bar{U}$, ein Widerspruch! ■

Satz 6.15 (Eidelheit, 1936) Sei X ein normierter Raum über $\mathbb{R}$. Seien A, B konvexe Teilmengen von X mit

$$A^\circ \neq \emptyset, A^\circ \cap B = \emptyset, B \neq \emptyset.$$

Dann gibt es $\lambda \in X^* \setminus \{\theta\}$, $r \in \mathbb{R}$, mit

$$A \subset H_{\lambda, r}^-, A^\circ \cap H_{\lambda, r} = \emptyset, B \subset H_{\lambda, r}^+. \quad (6.2)$$

Beweis:

Sei $a \in A^\circ$. $A' := A - a$, $B' := B - a$. Dann gilt: A', B' konvex, $\theta \in A'^\circ$, $A'^\circ \cap B' \neq \emptyset$, $B' \neq \emptyset$. Also können wir o. E. von vorneherein annehmen, dass $\theta \in A^\circ$ ist. Sei $y \in B$. Sei $Z := y + A - B$. Dann ist Z konvex.

Sei U eine Nullumgebung mit $\theta \in U \subset A$. Dann ist $V := \cup_{b \in B} (y - b + U)$ ist eine Nullumgebung mit $\theta \in V \subset Z$. Also ist $\theta \in Z^\circ$.

Annahme: $y \in Z^\circ$. Sei U eine Nullumgebung mit $y + U \subset Z$. Dann ist $U \subset A - B$ und damit $\theta \in (^\circ A - B) \subset (^\circ A) - B$, was einen Widerspruch zur Voraussetzung ergibt.

Also wissen wir nun

$$Z \text{ konvex, } \theta \in Z^\circ, y \notin Z^\circ.$$

Wir wollen nun den Fortsetzungssatz 5.5 nutzen. Dazu einige Vorbereitungen. Wir betrachten das Minkowski-Funktional zu Z :

$$p(x) := \inf\{r \geq 0 | x \in rZ\}, x \in X.$$

Da Z eine konvexe Nullumgebung ist, gilt: $p : X \rightarrow \mathbb{R}$ ist wohldefiniert. p ist sublinear, denn

$$p(x + x') \leq p(x) + p(x'), p(rx) = rp(x) \text{ für alle } x, x' \in X, r \geq 0.$$

Sei $Y := \text{span}(\{y\})$. Wir definieren das lineare Funktional $\lambda_0 : Y \rightarrow \mathbb{R}$ durch:

$$\langle \lambda_0, \alpha y \rangle := \alpha p(y), \alpha \in \mathbb{R}.$$

Dazu gibt es nach Satz 5.5 $\lambda \in X'$ mit

$$\lambda|_Y = \lambda_0, \langle \lambda, x \rangle \leq p(x) \text{ für alle } x \in X.$$

Mit dem Lemma 6.14 folgt:

$$p(z) > 1, z \in X \setminus \bar{Z}, p(z) = 1, z \in \bar{Z} \setminus Z^\circ, p(y) \geq 1.$$

Dies zeigt

$$Z \subset H_{\lambda, 1}^-, y \in H_{\lambda, 1}^+, \lambda \neq \theta \text{ da } \langle \lambda, y \rangle \geq 1. \quad (6.3)$$

Wir wollen nun zeigen, dass λ stetig ist, d. h. dass $\lambda \in X^*$ gilt.

Wegen $\theta \in Z^\circ \subset \overset{-1}{\lambda}((-\infty, 1])$ ist $\overset{-1}{\lambda}((-\infty, 1])$ eine Nullumgebung. Da λ linear ist, ist auch $\overset{-1}{\lambda}([-1, \infty)) = -\overset{-1}{\lambda}((-\infty, 1])$ eine Nullumgebung, und daher auch $\overset{-1}{\lambda}([-1, 1])$. Nun ist $\overset{-1}{\lambda}([- \varepsilon, \varepsilon])$ eine Nullumgebung für alle $\varepsilon > 0$. Dies zeigt, dass λ stetig in θ ist, also stetig überall ist.

Aus (6.3) folgt mit der eingeführten Darstellung $Z = y + A - B$

$$\langle \lambda, y \rangle \geq 1, \langle \lambda, a - b \rangle \leq 1 - \langle \lambda, y \rangle \leq 0 \text{ für alle } a \in A, b \in B,$$

und daher

$$\alpha := \sup_{a \in A} \langle \lambda, a \rangle \leq \beta := \inf_{b \in B} \langle \lambda, b \rangle.$$

Wähle $r \in [\alpha, \beta]$. Dann gilt also

$$A \subset H_{\lambda, r}^-, B \subset H_{\lambda, r}^+. \quad (6.4)$$

Annahme: Es gibt z in $A^\circ \cap H_{\lambda, r}$.

Dann gibt es $\delta > 0$ mit $z + \delta y \in A$. Also ist $\langle \lambda, z + \delta y \rangle = r + \delta \langle \lambda, y \rangle \geq r + \delta$. Dies bedeutet $z + \delta y \in H_{\lambda, r}^+ \setminus H_{\lambda, r}$ im Widerspruch zu $z + \delta y \in A \subset H_{\lambda, r}^-$. ■

Bemerkung 6.16 Wir beschäftigen uns hier nur mit der Trennung im reellen Fall. Im Fall, dass ein Vektorraum über $\mathbb{C}$ vorliegt, trennt man zunächst reell ($X := X_{\mathbb{R}}$) und geht dann zum komplex-linearen Funktional über wie im Beweis zu Satz 5.6; Terme $\langle \lambda, x \rangle$ in den Resultaten sind dann zu ersetzen durch $\Re \langle \lambda, x \rangle$. □

Bemerkung 6.17 Das Minkowski-Funktional nehmen wir zum Anlass, die lokalkonvexen Räume zu skizzieren. Sie sind in gewisser Weise die topologischen Vektorräume, in denen sich ziemlich vollkommen die Ergebnisse der normierten Räume einschließlich ihrer schwachen Topologien wiederfinden lassen. Konvexität spielt eine entscheidende Rolle.

In einem topologischen Vektorraum sind die Vektorraumverknüpfungen „Addition, skalare Multiplikation“ (bezüglich der induzierten Produkttopologien) stetig, was zur Konsequenz hat, dass es reicht, lokale topologische Fragestellungen immer in der Umgebung des Nullpunktes zu studieren.

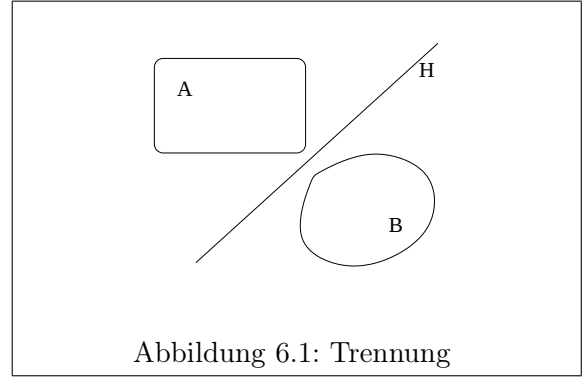
Ein topologischer Vektorraum $(X, \mathcal{T})$ heißt **lokalkonvex**, falls es darin eine Nullumgebungsbasis aus konvexen Mengen gibt. Damit äquivalent ist, dass die Topologie $\mathcal{T}$ von einer Familie $(p_i)_{i \in I}$ von Halbnormen (Normen, denen die Definitheit fehlt) gemäß

$$U \in \mathcal{U}(\theta) \iff \exists \varepsilon > 0 \exists i_1, \dots, i_m \in I (\theta \in V := \{x \in X \mid p_{i_j}(x) < \varepsilon, j = 1, \dots, m\} \subset U$$

erzeugt wird. Mit der Familie der Halbnormen $x \mapsto |\langle \lambda, x \rangle|, \lambda \in X^*$, ordnet sich die schwache Topologie auf einem normierten Raum X ein. Mit Hilfe eines angepassten Minkowskifunktional gelingt es den Satz von Hahn-Banach und Trennungssätze zu beweisen und damit wieder die Reichhaltigkeit von X^* nachzuweisen.

Sei $C^\infty(\Omega)$ der Vektorraum der unendlich oft stetig partiell differenzierbaren (also glatten) Funktionen auf der nichtleeren offenen Teilmenge $\Omega \subset \mathbb{R}^n$. Damit setzen wird

$$C_0^\infty(\Omega) := \{f \in C^\infty(\Omega) \mid \exists K \subset \Omega \text{ mit } f(x) = 0, x \in \Omega \setminus K, K \text{ kompakt}\}.$$



$C_0^\infty(\Omega)$ kann so topologisiert werden, dass ein lokalkonvexer Raum mit abzählbarer Nullumgebungsbasis entsteht. Man nimmt dazu eine abzählbare „Ausschöpfung“ von Ω mit einer Familie $(K_j)_{j \in \mathbb{N}}$ von kompakten Teilmengen von Ω und definiert Halbnormen $p_{\alpha,j}, \alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}_0^n, j \in \mathbb{N}$, durch

$$p_{\alpha,j}(f) := \max_{x \in K_j} \left| \frac{\partial^{\alpha_1 + \dots + \alpha_n} f}{\partial^{\alpha_1} x_1 \dots \partial^{\alpha_n} x_n}(x) \right|, f \in C^\infty(\Omega).$$

Damit ist nun auch der Dualraum X^* , bezeichnet durch $\mathcal{D}(\Omega)$, der stetigen linearen Funktionalen auf $X := C^\infty(\Omega)$ erklärt. Dieses Paar $C^\infty(\Omega), \mathcal{D}(\Omega)$ spielt eine große Rolle beim Studium der partiellen Differentialgleichungen: $C^\infty(\Omega)$ als Raum der **Testfunktionen**, $\mathcal{D}(\Omega)$ als Raum der **Distributionen** (verallgemeinerte Funktionen). $\square$

Folgerung 6.18 Sei X ein reeller normierter Raum. Sei B eine konvexe, abgeschlossene, nicht-leere Teilmenge von X und sei $z \in X \setminus B$. Dann gibt es $\lambda \in X^* \setminus \{\theta\}, r \in \mathbb{R}$, mit

$$\langle \lambda, z \rangle < r \leq \langle \lambda, x \rangle \text{ für alle } x \in B. \quad (6.5)$$

Beweis:

Sei $d := \text{dist}(z, B) := \inf_{b \in B} \|z - b\|$. Da B abgeschlossen ist, ist $d > 0$. Also gilt $B_d(z) \cap B = \emptyset$. Da $B_d(z)$ und B konvex sind, folgt mit Satz 6.15 die Behauptung ($A := B_d(z)$). $\blacksquare$

Bemerkung 6.19 Die obige Folgerung macht uns im Beweis klar, wo die Schwierigkeit liegt, denselben Trennungssatz in einem allgemeinen topologischen Vektorraum über $\mathbb{R}$ zu beweisen: Statt $B_d(z)$ könnten wir sofort eine offene Umgebung von z erhalten, aber wir haben ohne weitere Voraussetzung nicht, dass diese Umgebung konvex ist, wie wir es von $B_d(z)$ wissen. Die Folgerung wäre also auch richtig in einem topologischen Vektorraum über $\mathbb{R}$, der zu jedem Punkt mindestens eine konvexe offene Umgebung zulässt. Solche Räume sind die lokalkonvexen Räume. $\square$

Die Forderung nach inneren Punkten in den Voraussetzungen des obigen Trennungssatzes und implizit in der Folgerung ist wesentlich, ohne sie kann man – zumindest im unendlichdimensionalen Fall – sofort Gegenbeispiele für die Trennbarkeit angeben.

Beispiel 6.20 Ist X ein reeller normierter Raum und seien A, B konvexe Mengen, die dicht in X sind. Dann reicht die Voraussetzung $A \cap B = \emptyset$ nicht für die Trennbarkeit aus. Eine konkrete Situation ist etwa in l_2 gegeben durch

$$A := c_f, B := x + c_f \text{ mit } x = (2^{-k+1})_{k \in \mathbb{N}}.$$

Die beiden Mengen A, B kann man nur trivial trennen. $\square$

Beispiel 6.21 Im obigen Beispiel sind die Mengen A, B nicht abgeschlossen. Hierzu einen Verschärfung.

Sei $X := l_1, A_0 := \{x = (x^k)_{k \in \mathbb{N}} \mid x^{2k} = 0, k \in \mathbb{N}\}, B := \{x = (x^k)_{k \in \mathbb{N}} \mid x^{2k} = 2^{-k} x^{2k-1}, k \in \mathbb{N}\}$. Sei $z = (z^k)_{k \in \mathbb{N}}$ mit $z^{2k-1} = 0, z^{2k} = 2^{-k}, k \in \mathbb{N}$. Damit definieren wir $A := A_0 - z$ und haben, dass A, B konvex und abgeschlossen sind und $A \cap B = \emptyset$ gilt. A, B sind nicht trennbar. $\square$

Definition 6.22 Sei X ein normierter Raum. Eine beschränkte konvexe Teilmenge B von X besitzt die **Trennungseigenschaft** genau dann, wenn für jede konvexe Teilmenge A von X mit $A \cap B = \emptyset$ ein $\lambda \in X^*$ existiert mit

$$\sup_{x \in A} \langle \lambda, x \rangle < \inf_{y \in B} \langle \lambda, y \rangle.$$

$\square$

Klar, in der Situation von Definition 6.22 gilt für jedes $\alpha \in [\sup_{x \in A} \langle \lambda, x \rangle, \inf_{y \in B} \langle \lambda, y \rangle]$

$$A \subset H_{\lambda, \alpha}^+, B \subset H_{\lambda, \alpha}^-.$$

Jedes $H_{\lambda, \alpha}$ nennen wir eine **trennende Hyperebene**.

Ein wichtiges Aussage hierzu ist das Resultat von Klee (siehe [21]), dass jede schwach kompakte konvexe Menge die Trennungseigenschaft besitzt. Es ist sogar mehr bekannt. Eine beschränkte Menge B mit nicht-leerem Inneren ist schwach kompakt genau dann, wenn B die Trennungseigenschaft besitzt; siehe [15]. Ohne Beweis geben wir für die abgeschlossene Einheitskugel $B := \overline{B}_1$ in einem normierten Raum X an (siehe [15]): Es sind äquivalent:

- (a) B ist schwach kompakt.
- (b) B ist schwach abzählbar kompakt.
- (c) B ist schwach folgenkompakt.
- (d) Ist $S \subset X$ schwach abgeschlossen und gilt $B \cap S = \emptyset$, so gilt $\text{dist}(B, S) = \inf\{\|x - y\| \mid x \in B, y \in S\} > 0$.
- (e) B besitzt die Trennungseigenschaft.

Satz 6.23 Sei X ein normierter Raum und sei $A \subset X$ konvex. Dann sind äquivalent:

- (a) A ist abgeschlossen (in der Ausgangstopologie/starken Topologie).
- (b) A ist $\sigma(X, X^*)$ -abgeschlossen.

Zusatz: Aus (a) folgt

$$A = \bigcap_{A \subset H_{\lambda, r}^+} H_{\lambda, r}^+. \quad (6.6)$$

Beweis:

(a) $\implies$ (b)

Setze $C := \bigcap_{A \subset H_{\lambda, r}^+} H_{\lambda, r}^+$. Klar, $A \subset C$. Sei $x \notin A$. Nach Folgerung 6.18 gibt es $\lambda \in X^* \setminus \{\theta\}, r \in \mathbb{R}$, mit

$$A \subset H_{\lambda, r}^+, x \notin H_{\lambda, r}^+.$$

Also $x \notin C$. Dies zeigt auch $C \subset A$.

Damit $A = C$ $\sigma(X, X^*)$ -abgeschlossen, denn: Sei $x \in X \setminus C$, d. h. $x \in H_{\mu, s}^- \setminus H_{\mu, s}$ für ein $\mu \in X^*, s \in \mathbb{R}$ mit $A \subset H_{\mu, s}^+$. Setze $\varepsilon := \frac{1}{2}(s - \langle \mu, x \rangle)$, $M := \{\mu\}$. Dann haben wir

$$x + U_{\varepsilon, M} \subset H_{\mu, s}^- \setminus H_{\mu, s} \subset X \setminus C.$$

(b) $\implies$ (a)

Jede $\sigma(X, X^*)$ -abgeschlossene Menge ist offenbar abgeschlossen (in der Ausgangstopologie). ■

Beispiel 6.24 Die Aussage von Satz 6.23 wird falsch, wenn die Voraussetzung „ A konvex“ weggelassen wird. Das Gegenbeispiel sollten wir in unendlichdimensionalen Räumen suchen, da in endlichdimensionalen Räumen die schwache und die Ausgangstopologie „äquivalent“ sind.

Wir betrachten $A := \{x \in X \mid \|x\| = 1\}$ im Raum $X := l_2$. A ist offenbar abgeschlossen und nicht konvex. A ist nicht $\sigma(X, X^*)$ -abgeschlossen, denn für die Folge $(e^k)_{k \in \mathbb{N}}$ der „Einheitsvektoren“ gilt, dass θ $\sigma(X, X^*)$ -Grenzwert der Folge ist. Dies sieht man mit Folgerung 5.22 so: Sei $\lambda \in l_2^*$. Wir erhalten (siehe Abschnitt 3.4) ein Element $(z_n)_{n \in \mathbb{N}} \in l_2$ mit

$$\langle \lambda, (x_n)_{n \in \mathbb{N}} \rangle = \sum_{j=1}^{\infty} z_j x_j \text{ für alle } (x_n)_{n \in \mathbb{N}} \in l_2.$$

Daraus folgt

$$\langle \lambda, e^k \rangle = \sum_{j=1}^{\infty} z_j e_j^k = z_k, \quad k \in \mathbb{N}.$$

Da $(z_n)_{n \in \mathbb{N}} \in l_2$ ist, folgt $\lim_k z_k = 0$, und es ist gezeigt $\lim_k \langle \lambda, e^k \rangle = 0$. $\square$

Definition 6.25 Sei X ein normierter Raum und sei $A \subset X$. Wir setzen

$$\overline{\text{co}}(A) := \bigcap_{A \subset K, K \text{ konvex}, K = \overline{K}} K,$$

und nennen $\overline{\text{co}}(A)$ die **abgeschlossene konvexe Hülle** von A . $\square$

Folgerung 6.26 Sei X ein normierter Raum und sei $A \subset X$. Es gilt:

- (a) $\overline{\text{co}}(A)$ ist konvex und abgeschlossen.
- (b) $\overline{\text{co}}(A) = \overline{\text{co}(\overline{A})} = \overline{\text{co}(A)}$.

Beweis:

(a) ist sofort einzusehen, (b) überlassen wir dem Leser als Übung. $\blacksquare$

Satz 6.27 (Mazur) Sei X ein normierter Raum. Konvergiert die Folge $(x^k)_{k \in \mathbb{N}}$ gegen x bezüglich der $\sigma(X, X^*)$ -Topologie, so gilt

$$x \in \overline{\text{co}}(\{x^k | k \in \mathbb{N}\}),$$

was bedeutet: für alle $\varepsilon > 0$ gibt es $x^1, \dots, x^m$ und $a_1, \dots, a_m \in \mathbb{R}, a_1, \dots, a_m \geq 0$, mit

$$\sum_{j=1}^m a_j = 1, \quad \|x - \sum_{j=1}^m a_j x^j\| < \varepsilon.$$

Beweis:

Sei $B := \{x^k | k \in \mathbb{N}\}, A := \overline{\text{co}}(B)$. A ist konvex und abgeschlossen, also nach Satz 6.23 auch abgeschlossen bezüglich der $\sigma(X, X^*)$ -Topologie. Also gilt wegen $B \subset A$ offenbar $x \in \overline{\text{co}}(B)$. Der Rest der Behauptung leitet sich daraus ab. $\blacksquare$

Satz 6.28 (Helly) Sei X ein normierter Raum und seien $\lambda_1, \dots, \lambda_n \in X^*, a_1, \dots, a_n \in \mathbb{K}$. Dann sind folgende Aussagen äquivalent:

- (a) $\forall \varepsilon > 0 \exists x \in X (\|x\| \leq 1 + \varepsilon, \langle \lambda_i, x \rangle = a_i, i = 1, \dots, n)$
- (b) Es gilt $|\sum_{i=1}^n \alpha_i a_i| \leq \|\sum_{i=1}^n \alpha_i a_i\|$ für alle $\alpha_1, \dots, \alpha_n \in \mathbb{K}$.

Beweis:

(a) $\implies$ (b) Seien $\alpha_1, \dots, \alpha_n \in \mathbb{R}, \varepsilon > 0$. Wähle $x \in X$ gemäß (a). Dann ist

$$|\sum_{i=1}^n \alpha_i a_i| = |\sum_{i=1}^n \alpha_i \langle \lambda_i, x \rangle| \leq \|\sum_{i=1}^n \alpha_i \lambda_i\| \|x\| \leq \|\sum_{i=1}^n \alpha_i \lambda_i\| (1 + \varepsilon).$$

Da $\varepsilon > 0$ beliebig ist, ist b) bewiesen.

(b) $\implies$ (a) Wir können o. E. annehmen, dass $\lambda_1, \dots, \lambda_n$ linear unabhängig sind. Dann ist die Abbildung

$$T : X \ni x \longmapsto (\langle \lambda_1, x \rangle, \dots, \langle \lambda_n, x \rangle) \in \mathbb{R}^n$$

linear, stetig. T ist auch surjektiv, denn ist T nicht surjektiv, dann gibt es $\mu = (\mu_1, \dots, \mu_n) \in \mathbb{R}^n \cong \mathcal{S}(R^n)^* \neq \theta$ mit

$$0 = \langle \mu, Tx \rangle = \sum_{i=1}^n \mu_i \langle Tx \rangle = \left\langle \sum_{i=1}^n \mu_i \lambda_i, x \right\rangle \text{ für alle } x \in X,$$

also $\mu_1 = \dots = \mu_n = 0$, was mit der linearen Unabhängigkeit von $\lambda_1, \dots, \lambda_n$ ein Widerspruch zu $\mu \neq \theta$ ist. Damit ist T nun eine offene Abbildung und es gibt $\delta > 0$ mit $B_\delta \subset T(B_1)$; siehe Satz 4.10. Dann folgt $B_{\delta(1+\varepsilon)} \subset T(B_1)$ und daraus liest man (a) ab. ■

Bemerkung 6.29 *Im reflexiven Fall kann man den Satz von Helly einfacher zeigen. Ferner kann man die Bedingung (a) mit $\varepsilon = 0$ aufschreiben.* □

Satz 6.30 *Sei X ein reflexiver Banachraum und sei $K \subset X$ konvex, abgeschlossen und nicht-leer. Dann gibt es zu jedem $x \in X$ ein $y \in K$ mit*

$$\|x - y\| = \inf_{z \in K} \|x - z\|. \quad (6.7)$$

Beweis:

Sei $x \in X$ und sei dazu $(z^n)_{n \in \mathbb{N}}$ eine Folge in K mit

$$\lim_n \|x - z^n\| = \inf_{z \in K} \|x - z\|.$$

Wegen $\|z^n\| \leq \|z^n - x\| + \|x\|$ ist die Folge $(z^n)_{n \in \mathbb{N}}$ beschränkt und besitzt daher nach Satz 5.31 eine schwach konvergente Teilfolge $(z^{n_k})_{k \in \mathbb{N}}$; $z^{n_k} \rightharpoonup y$. Da K abgeschlossen und konvex ist, ist K auch schwach abgeschlossen; also $y \in K$. Da auch $x - z^{n_k} \rightharpoonup x - y$ gilt, folgt $\|x - y\| \leq \liminf_k \|x - z^{n_k}\|$; siehe Folgerung 5.26. ■

Beachte, dass wir hier keine Eindeutigkeit behauptet haben, sie kann allgemein auch nicht nachgewiesen werden. Die Eindeutigkeit hängt an der geometrischen Qualität der Kugel in einem normierten Raum. Ein Gegenbeispiel zur Eindeutigkeit überlegt man sich leicht in der Situation $X = \mathbb{R}^2$, versehen mit der l_∞ - oder l_1 -Norm. Die Einheitskugeln $K = B_1$ sind nicht „rund“ und es ist leicht ein $x \in \mathbb{R}^2 \setminus K$ anzugeben, so dass die Aufgabe (6.7) mindestens zwei Lösungen besitzt.

6.3 Satz vom abgeschlossenen Bild

Kommen wir zurück zu den dualen Operatoren aus dem letzten Kapitel.

Definition 6.31 *Sei X ein normierter Raum, $U \subset X$ ein linearer Teilraum und V ein linearer Teilraum von X^* . Die **Annihilatoren** von U und V sind*

$$U^\perp = \{\lambda \in X^* \mid \langle \lambda, u \rangle = 0 \text{ für alle } u \in U\}, \quad {}^\perp V = \{x \in X \mid \langle \lambda, x \rangle = 0 \text{ für alle } \lambda \in V\}.$$

□

Die Annihilatoren sind offenbar Unterräume von X^* bzw. X . $U^\perp$ ist abgeschlossen bezüglich der Norm-Topologie von X^* . ${}^\perp V$ ist als Durchschnitt von Nullräumen stetiger Funktionale abgeschlossen in X . $U^\perp$ lässt sich als Durchschnitt der Nullräume der $J_X(x)$, $x \in U$, schreiben. Da jedes $J_X(x)$ stetig bezüglich der $\sigma(X^*, X)$ -Topologie ist, ist $U^\perp$ sogar $\sigma(X^*, X)$ -abgeschlossen in X^* .

Seien X, Y Banachräume und sei $T \in \mathcal{B}(X, Y)$. Man zeigt mit einem Trennungssatz oder direkt mit dem Satz von Hahn-Banach

$$\ker(T^*) = {}^\perp \text{ran}(T), \quad \ker(T) = \text{ran}(T^*)^\perp,$$

Identitäten, die schon aus der Linearen Algebra bekannt sind. Entsprechend erwartet man etwa $\text{ran}(T) = {}^\perp \ker(T^*)$, was im Allgemeinen aber nicht gilt. Hier ist der Satz, der die Sachlage klärt.

Satz 6.32 (Banachs Satz vom abgeschlossenen Bild) *Seien X, Y Banachräume, sei $T : X \rightarrow Y$ linear und stetig. Dann sind äquivalent:*

- (a) $\text{ran}(T)$ ist abgeschlossen
- (b) $\text{ran}(T^*)$ ist abgeschlossen
- (c) $\text{ran}(T) = {}^\perp \ker(T^*)$
- (d) $\text{ran}(T^*) = \ker(T)^\perp$

Beweis:

Zu (a) $\implies$ (b).

Sei $Z := \text{ran}(T)$. Z ist als Teilmenge von Y ein Banachraum. Betrachte $S : X \ni x \mapsto Tx \in Z$. S ist linear und stetig. Wir zeigen $\text{ran}(S^*) = \text{ran}(T^*)$.

Sei $\mu \in Z^*$. Mit dem Satz von Hahn-Banach kann μ fortgesetzt werden zu $\lambda \in Y^*$ mit

$$\langle \mu, Tx \rangle = \langle \lambda, Tx \rangle, \quad x \in X.$$

Daher gilt $S^*\mu = T^*\mu$. Dies zeigt $\text{ran}(S^*) \subset \text{ran}(T^*)$.

Sei $\lambda \in Y^*$ und damit $\mu := \lambda|_Z$. Dann folgt mit

$$\langle T^*\lambda, x \rangle = \langle \lambda, Tx \rangle = \langle \mu, Sx \rangle = \langle S^*\mu, x \rangle, \quad x \in X.$$

$T^*\lambda = S^*\mu$. Dies zeigt $\text{ran}(T^*) \subset \text{ran}(S^*)$.

Also genügt es anzunehmen, dass $\text{ran}(T) = Y$ gilt, d. h. dass T surjektiv ist. Aus dem Satz von der offenen Abbildung folgt

$$\exists c > 0 \forall y \in Y \exists x \in X (Tx = y, \|x\| \leq c\|y\|).$$

Also haben wir für jedes $\lambda \in Y^*$ und $y = Tx$

$$|\langle \lambda, y \rangle| = |\langle \lambda, Tx \rangle| = |\langle T^*\lambda, x \rangle| \leq \|x\| \|T^*\lambda\| \leq c\|y\| \|T^*\lambda\|.$$

Daher

$$\|\lambda\|_* = \sup_{\|y\| \leq 1} |\langle \lambda, y \rangle| \leq c\|T^*\lambda\|.$$

Daraus liest man ab, dass $\text{ran}(T^*)$ abgeschlossen ist.

Zu (b) $\implies$ (a).

Betrachte $S : X \ni x \mapsto Tx \in Z := \overline{\text{ran}(T)}$. S^* ist injektiv, denn: $S^*\lambda = \theta$ impliziert

$$\langle \lambda, Sx \rangle = \langle \lambda, Tx \rangle = \langle S^*\lambda, x \rangle = 0, \quad x \in X;$$

da $\text{ran}(S) = \text{ran}(T)$ dicht in Z ist, muss $\lambda = \theta$ gelten.

Wie oben zeigt man $\text{ran}(T^*) = \text{ran}(S^*)$. Da daher $\text{ran}(S^*)$ abgeschlossen ist, ist S^* ein linearer, bijektiver, stetiger Operator von Z^* in den Banachraum $\text{ran}(S^*)$. Nach dem Satz von der offenen Abbildung ist $(S^*)^{-1}$ stetig.

Betrachte die Eigenschaft

$$\forall \rho > 0 \exists \tau > 0 (B_\tau \cap Z \subset \overline{S(B_\rho)}) \quad (6.8)$$

In Lemma 4.9 haben wir im Beweis gezeigt, dass aus (6.8)

$$\forall \rho > 0 \exists \sigma > 0 (B_\sigma \cap Z \subset S(B_\rho)) \quad (6.9)$$

folgt. Trifft (6.8), also auch (6.9) zu, dann folgt aus Satz 4.10, dass S surjektiv ist. Dies bedeutet $\text{ran}(S) = \text{ran}(T) = Z$ und damit ist $\text{ran}(T)$ abgeschlossen.

Trifft (6.8) nicht zu, dann gibt es eine Folge $(z_n)_{n \in \mathbb{N}}$ in Z mit

$$\lim_n z_n = \theta, z_n \notin \overline{S(B_\rho)}, n \in \mathbb{N}.$$

Da $\overline{S(B_\rho)}$ abgeschlossen und konvex ist, gibt es zu jedem $n \in \mathbb{N}$ ein $\lambda_n \in Z^*$, $\|\lambda_n\|_* = 1$, mit

$$\langle \lambda_n, z_n \rangle > \sup_{\|x\| \leq \rho} |\langle \lambda_n, Sx \rangle|;$$

siehe 6.18 und beachte dabei, dass mit $\|x\| \leq \rho$ auch $\| -x \| \leq \rho$ gilt. Daraus folgt $\|S^* \lambda_n\| < \rho^{-1} \|\lambda_n\|_* \|z_n\|$, $n \in \mathbb{N}$ und daher $\lim_n S^* \lambda_n = \theta$. Dies ist im Widerspruch zur Tatsache, dass S^* eine stetige Inverse hat.

Zu (a) $\implies$ (c).

Aus

$$\langle \lambda, Tx \rangle = \langle T^* \lambda, x \rangle, x \in X, \lambda \in Y^*,$$

folgt $\text{ran}(T) \subset N := \{y \in Y \mid \langle \lambda, y \rangle = 0 \text{ für alle } \lambda \in \ker(T^*)\}$. Wir zeigen, dass (a) auch $N \subset \text{ran}(T)$ impliziert.

Annahme: $y \in N \setminus \text{ran}(T)$. Mit dem Satz von Hahn–Banach (siehe Folgerung 5.9 (b)) erhalten wir $\lambda \in Y^*$ mit

$$\langle \lambda, y \rangle \neq 0, \langle \lambda, Tx \rangle = 0 \text{ für alle } x \in X.$$

Daraus folgt $\langle T^* \lambda, x \rangle = 0$, $x \in X$, und daher $T^* \lambda = \theta$, $y \notin N$. Dies ist ein Widerspruch.

Zu (c) $\implies$ (a).

Dies ist klar, da ${}^\perp \ker(T^*)$ abgeschlossen ist.

Zu (b) $\implies$ (d).

Beachte, dass wir wegen (a) $\iff$ (b) schon wissen, dass $\text{ran}(T)$ abgeschlossen ist.

Die Inklusion $\text{ran}(T^*) \subset \ker(T)^\perp$ ist klar. Wir zeigen, dass (b) auch $\ker(T)^\perp \subset \text{ran}(T^*)$ impliziert.

Betrachte mit dem Graphen $G(T)$ von T den Operator

$$S : G(T) \ni (x, Tx) \longmapsto Tx \in Z := \text{ran}(T).$$

Da T stetig ist, ist $G(T)$ abgeschlossen, also ein Banachraum. S ist linear, stetig, surjektiv. Dann hat die Faktorabbildung

$$\hat{S} : G(T)/(\ker(T) \times Y) \ni ([x, Tx]) \longmapsto Tx \in Z$$

eine stetige Inverse $\hat{S}^{-1}$.

Sei nun $\lambda \in \ker(T)^\perp$. Definiere ein Funktional μ auf $Z = \text{ran}(T)$ durch

$$\langle \mu, y \rangle := \langle \lambda, x \rangle, \text{ falls } \hat{S}^{-1}(y) = ([x, Tx]).$$

μ ist wohldefiniert, denn $\hat{S}^{-1}(y) = ([x, Tx]) = ([z, Tz])$ impliziert $x - z \in \ker(T)$ und daher $\langle \lambda, x - z \rangle = 0$. Offenbar ist μ linear und stetig, da $\hat{S}^{-1}$ stetig ist. Damit ist $\mu \in Z^*$. Sei $\rho \in Y^*$ die Fortsetzung von μ nach dem Satz von Hahn–Banach. Dann gilt

$$\langle \rho, Tx \rangle = \langle \lambda, Tx \rangle = \langle \mu, x \rangle, \quad x \in X.$$

Dies zeigt $T^*\rho = \lambda$.

Zu (d) $\implies$ (b).

Dies ist wieder offensichtlich, da $\ker(T)^\perp$ abgeschlossen ist. ■

Folgerung 6.33 Seien X, Y Banachräume, sei $T : X \longrightarrow Y$ linear und stetig. Dann gilt:

- (a) $\operatorname{ran}(T) = Y$ genau dann, wenn T^* eine stetige Inverse auf $\operatorname{ran}(T^*)$ besitzt.
- (b) $\operatorname{ran}(T^*) = X^*$ genau dann, wenn T eine stetige Inverse auf $\operatorname{ran}(T)$ besitzt.

Beweis:

Wir zeigen (a). Sei $\operatorname{ran}(T) = Y$. Aus

$$T^*\lambda = \theta, \quad \langle \lambda, Tx \rangle = \langle T^*\lambda, x \rangle \quad \text{für alle } x \in X,$$

folgt $\lambda = \theta$. Also ist T^* injektiv und da nach Satz 6.32 $\operatorname{ran}(T^*)$ abgeschlossen ist, hat T^* eine stetige Inverse auf $\operatorname{ran}(T^*)$.

T^* besitze eine stetige Inverse auf $\operatorname{ran}(T^*)$. Dann ist $\ker(T^*) = \{\theta\}$ und wir erhalten aus (c) in Satz 6.32 $\operatorname{ran}(T) = Y$.

(b) zeigt man ziemlich analog. ■

Beispiel 6.34 Betrachte auf dem Banachraum c_0 die Abbildung

$$T : c_0 \ni x = (x_k)_{k \in \mathbb{N}} \longmapsto y := \left(\frac{1}{k}x_k\right)_{k \in \mathbb{N}} \in c_0.$$

Offenbar ist T linear und stetig. Da alle „Einheitsfolgen“ $e^n := (\delta_{k,n})_{k \in \mathbb{N}}$ in $\operatorname{ran}(T)$ liegen, ist $\operatorname{ran}(T)$ dicht in c_0 . Da die Folge $(\frac{1}{k})_{k \in \mathbb{N}}$ in c_0 liegt, aber nicht in $\operatorname{ran}(T)$ liegen kann, ist $\operatorname{ran}(T)$ nicht abgeschlossen. □

Der obige Satz 6.32 hat bietet Einsicht in die Probleme bei der Lösung einer Gleichung

$$Tx = y \tag{6.10}$$

mit einem linearen, injektiven (o. E.) Operator $T : X \longrightarrow Y$, wenn $\dim(X) = \infty$ gilt. Ist $\operatorname{ran}(T)$ abgeschlossen, dann haben wir eine stetige Inverse S^{-1} von $S : X \ni x \longmapsto Tx \in Y$. Für $y, y' \in \operatorname{ran}(T)$ haben wir dann Lösungen $x, x' \in X$ der Gleichung mit

$$\|x - x'\| = \|S^{-1}y - S^{-1}y'\| \leq \|S^{-1}\| \|y - y'\|$$

Dies zeigt, dass der „Fehler“ $\varepsilon := \|y - y'\|$ in den „Fehler“ $\eta := \|x - x'\|$ umgesetzt wird, verstärkt mit der Konstanten $\|S^{-1}\|$, eine recht befriedigende Situation. Was tun, wenn wir y, y' nicht beide in $\operatorname{ran}(T)$ liegen. Hier sind wir dann in der günstigen Situation, dass wir dann auf $\operatorname{ran}(T)$ projizieren können; siehe Approximationssatz 6.56. Da Projektionen – wir gehen nicht genauer darauf ein – stetige Abbildungen sind, ist die Lösung der Gleichung wieder eine „gutartige“ Operation.

Ist $\operatorname{ran}(T)$ nicht abgeschlossen, aber dicht in Y , dann stehen beide obigen Vorgehensweisen (Verwendung von S^{-1} , Projektion) nicht ohne Vorsichtsmaßnahmen zur Verfügung: S^{-1} ist nicht

stetig, Projektionen auf nicht abgeschlossenen Mengen sind nicht abgesichert. Jede kleine Störung einer rechten Seite y kann zu einer großen Störung in der Lösung x führen.

Die eben beschriebene Situation findet man in der Regel bei allen Problemen, bei denen die Gleichung (6.10) ein Modell für einen Vorgang darstellt, der entgegen der natürlichen Kausalkette gerichtet ist, etwa: Schluß von Intensitätsschwankungen bei Bildern eines Computertomographen auf Dichteschwankungen. Da in Zeiten des Computers gerade diese **inversen Fragestellungen** interessant wurden, hat sich in den letzten 25 Jahren das Gebiet der „Inversen Probleme“ etabliert.

6.4 Der Satz von Krein–Milman

Definition 6.35 Sei X ein Vektorraum. Sei $B \subset A \subset X$, $B \neq \emptyset$. B heisst **extremal** in A , falls gilt:

$$x, y \in A, a \in (0, 1) \subset \mathbb{R}, ax + (1 - a)y \in B \implies x, y \in B.$$

□

Offensichtlich ist eine Teilmenge stets extremal in sich. In $\mathbb{R}^3$ sind extremal die Ecken eines Würfels im Würfel, die Kugeloberfläche in der Kugel, die Kanten in einem Vieleck. Offenbar ist die Eigenschaft „extremal in“ transitiv.

Definition 6.36 Sei X ein Vektorraum und sei $A \subset X$, $x \in A$. x heisst **Extremalpunkt** von A , falls $\{x\}$ extremal in A ist. Wir schreiben dann $x \in \text{ext}(A)$. □

Lemma 6.37 Sei X ein normierter Raum und sei $A \subset X$ konvex. Dann sind für $x \in A$ äquivalent:

- (a) $x \in \text{ext}(A)$.
- (b) $A \setminus \{x\}$ ist konvex.

Beweis:

(a) $\implies$ (b)

Seien $y, z \in A \setminus \{x\}$, $a \in (0, 1)$. Da A konvex ist, ist $ay + (1 - a)z \in A$. Da $x \in \text{ext}(A)$ ist, ist $ay + (1 - a)z \neq x$. Also ist $ay + (1 - a)z \in A \setminus \{x\}$.

(b) $\implies$ (a)

Seien $y, z \in A$, sei $a \in (0, 1)$ und sei $x := ay + (1 - a)z$. Wegen (b) kann $y \neq x$ und $z \neq x$ nicht gleichzeitig gelten. Sei etwa $y = x$. Dann ist $x = y = z$, also $x \in \text{ext}(A)$. ■

Es ist leicht einzusehen, dass $\text{ext}(A) = \emptyset$ ist, falls A offen ist. Etwas genauer grenzen die folgenden Beispiele die Möglichkeiten ab.

Beispiel 6.38 $\text{ext}(A)$ ist im allgemeinen nicht endlich, selbst wenn A kompakt ist. Dies sieht man schon in folgender einfachen Situation ein: $X := \mathbb{R}^n$, $A := \overline{B}_1$. Hier haben wir nämlich $\text{ext}(A) = \overline{B}_1 \setminus B_1 =: \mathcal{S}_1$. Beweis dazu:

Seien $x, y \in A$, $a \in [0, 1]$, $ax + (1 - a)y \in \mathcal{S}_1$. Dann haben wir

$$1 = a^2|x|^2 + (1 - a)^2|y|^2 + 2a(1 - a)(x - y) \leq |y|^2 + a^2(|x| - |y|)^2 + 2a|y|(|x| - |y|) \leq \max(|x|, |y|).$$

O. E. $|y| \leq |x|$. Also ist $|x| = 1$ und

$$1 < a^2 + 2a(a - 1) + (1 - a)^2 = (a + (1 - a))^2 = 1, \text{ falls } |y| < 1.$$

Somit $|y| = |x| = 1$, d. h. $x, y \in \mathcal{S}_1$. Dies zeigt $\mathcal{S}_1 \subset \text{ext}(A)$. Wegen Lemma 6.37 ist $A^\circ \cap \text{ext}(A) = B_1 \cap \text{ext}(A) = \emptyset$. Also $\mathcal{S}_1 = \text{ext}(A)$. □

Beispiel 6.39 $\text{ext}(A)$ ist nicht notwendigerweise abgeschlossen, selbst wenn A kompakt ist. Hier ist das passende Beispiel:

$$X := \mathbb{R}^3, A_1 := \{(0, 0, 1)\}, A_2 := \{(0, 0, -1)\}, A_3 := \{x \in \mathbb{R}^3 | x = (a, b, 0), a^2 + (b-1)^2 = 1\}, .$$

Setze $A := \text{co}(A_1 \cup A_2 \cup A_3)$. Es ist $\text{ext}(A) = (A_1 \cup A_2 \cup A_3) \setminus \{\theta\}$, also $\text{ext}(A)$ nicht abgeschlossen. $\square$

Lemma 6.40 Sei X ein normierter Raum. Sei $A \subset X$ kompakt, $A \neq \emptyset, \lambda \in X^*$. Dann ist

$$A_0 := \{x \in X | \langle \lambda, x \rangle = \inf_{a \in A} \langle \lambda, a \rangle\}$$

eine kompakte extreme Teilmenge von A .

Beweis:

A_0 ist nicht leer, da λ stetig und A kompakt ist. A_0 ist abgeschlossen, da $\lambda \in X^*$. Sei $\alpha := \inf_{a \in A} \langle \lambda, a \rangle$. Seien $x, y \in A, a \in (0, 1), ax + (1-a)y \in A_0$. Dann gilt

$$\langle \lambda, x \rangle \geq \alpha, \langle \lambda, y \rangle \geq \alpha, a\langle \lambda, x \rangle + (1-a)\langle \lambda, y \rangle = \langle \lambda, ax + (1-a)y \rangle = \alpha,$$

und daher $\langle \lambda, x \rangle = \alpha, \langle \lambda, y \rangle = \alpha$, d. h. $x, y \in A_0$. $\blacksquare$

Satz 6.41 Sei X ein normierter Raum und sei $A \subset X$ kompakt, nicht-leer. Dann ist $\text{ext}(A) \neq \emptyset$.

Beweis:

$$\mathcal{F} := \{S \subset A | S \text{ extremal in } A, S \text{ abgeschlossen}, S \neq \emptyset\}.$$

Wegen $A \in \mathcal{F}$ ist $\mathcal{F} \neq \emptyset$. Durch die Inklusion ist $\mathcal{F}$ halbgeordnet. Sei $\mathcal{F}_0$ eine totalgeordnete Teilmenge von $\mathcal{F}$; wir setzen

$$S_u := \bigcap_{S \in \mathcal{F}_0} S.$$

Da A kompakt ist und die endliche Durchschnittseigenschaft gilt, ist nach dem Cantorsche Durchschnittssatz (siehe Satz 1.24) S_u nichtleer. Ferner ist offenbar S_u abgeschlossen und extremal in A , da die Eigenschaft „extremal in“ transitiv ist. Also ist $S_u \in \mathcal{F}$ und daher eine untere Schranke für $\mathcal{F}_0$. Nach dem Lemma von Zorn gibt es ein minimales Element in $\mathcal{F}$; sei S_m dieses minimale Element.

Annahme: S_m enthält mindestens zwei Elemente $u, v, u \neq v$. Nach Folgerung 5.9 gibt es $\lambda \in X^*$ mit $\langle \lambda, u \rangle < \langle \lambda, v \rangle$. Sei

$$A_0 := \{x \in S_m | \langle \lambda, x \rangle = \inf_{y \in S_m} \langle \lambda, y \rangle\}.$$

Da A kompakt ist, ist A_0 eine extreme Teilmenge von S_m , also auch eine extreme Teilmenge von A , also $A_0 \in \mathcal{F}$ und $A_0 \subset S_m$. Da S_m minimal in $\mathcal{F}$ ist, gilt $A_0 = S_m$. Dies ist ein Widerspruch, da $v \notin A_0$.

Also ist S_m eine einelementige Menge: $S_m := \{z\}$. Dann ist nach Konstruktion $z \in \text{ext}(A)$. $\blacksquare$

Lemma 6.42 Sei X ein normierter Raum. Sei A nicht-leer. Dann gilt $\text{ext}(\text{co}(A)) \subset A$.

Beweis:

Sei $x \in \text{ext}(\text{co}(A)) \subset \text{co}(A)$; also

$$x = \sum_{i=1}^m a_i x^i \text{ mit } \sum_{i=1}^m a_i = 1, a_i \in [0, 1], x^i \in A, i = 1, \dots, m.$$

O.E. $a_i > 0, i = 1, \dots, m$. Ist $m = 1$, so gilt offensichtlich $x = x^1 \in A$. Sei $m \geq 2$. Dann gilt $a_1 < 1$ und

$$x = a_1 x^1 + \sum_{i=2}^m a_i x^i = a_1 x^1 + (1 - a_1) \sum_{i=2}^m (1 - a_1)^{-1} a_i x^i.$$

Da $x \in \text{ext}(\text{co}(A)), x^1 \in \text{co}(A), \sum_{i=2}^m (1 - a_1)^{-1} a_i x^i \in \text{co}(A)$, folgt mit der Definition der Extremalität $x = x^1 = \sum_{i=2}^m (1 - a_1)^{-1} a_i x^i \in A$. ■

Satz 6.43 (Krein–Milman, 1940) *Sei X ein normierter Raum. Sei $A \subset X$ nicht-leer, konvex und kompakt. Dann gilt $A = \overline{\text{co}}(\text{ext}(A))$.*

Beweis:

Nach Lemma 6.42 gilt: $\overline{\text{co}}(\text{ext}(A)) = \overline{\text{co}}(\text{ext}(\text{co}(A))) \subset \overline{\text{co}}(A) = A$.

Annahme: Es gibt $x \in A \setminus \overline{\text{co}}(\text{ext}(A))$. Dann gibt es nach Folgerung 6.18 $\lambda \in X^* \setminus \{\theta\}$ mit

$$\langle \lambda, x \rangle < \inf_{w \in \overline{\text{co}}(\text{ext}(A))} \langle \lambda, w \rangle.$$

Sei $A_0 := \{y \in A \mid \langle \lambda, y \rangle = \inf_{v \in A} \langle \lambda, v \rangle\}$; $\inf_{v \in A} \langle \lambda, v \rangle$ ist wohldefiniert, da A kompakt ist. Nach Lemma 6.40 ist A_0 kompakt und extremale Teilmenge von A . Nach Satz 6.41 gilt $\text{ext}(A_0) \neq \emptyset$. Sei $y_0 \in \text{ext}(A_0)$. Dann gilt $y_0 \in \text{ext}(A) \subset \overline{\text{co}}(\text{ext}(A))$. Daraus folgt ein Widerspruch, den man aus

$$\langle \lambda, y_0 \rangle \leq \langle \lambda, x \rangle < \inf_{w \in \overline{\text{co}}(\text{ext}(A))} \langle \lambda, w \rangle \leq \langle \lambda, y_0 \rangle.$$

abliest. Also gilt $A \subset \overline{\text{co}}(\text{ext}(A))$. ■

Folgerung 6.44 *Sei X ein reeller normierter Raum. Sei $A \subset X$ konvex, kompakt, $A \neq \emptyset, \lambda \in X^*$. Dann gibt es $z \in \text{ext}(A)$ mit $\langle \lambda, z \rangle = \inf_{a \in A} \langle \lambda, a \rangle$.*

Beweis:

Es ist $Q := \{u \in A \mid \langle \lambda, u \rangle = \inf_{a \in A} \langle \lambda, a \rangle\}$ kompakt und nichtleer. Ferner ist Q nach Lemma 6.40 extremal in A . Nach Satz 6.41 gilt $\text{ext}(Q) \neq \emptyset$. Sei $z \in \text{ext}(Q)$. Dann ist z Extrempunkt von A und es gilt $\langle \lambda, z \rangle = \inf_{x \in A} \langle \lambda, x \rangle$. ■

Folgerung 6.44 ist in der Linearen Optimierung (X ist ein $\mathbb{R}^n$) die Grundlage des Simplexverfahrens. Hier ist A ein beschränktes Polyeder P . Dessen Extrempunkte sind nur endlich viele und leicht zu bestimmen/charakterisieren: sie sind die Ecken des Polyeders. Der Satz besagt dann, dass „optimale Punkte“ in A auch in den Ecken zu finden sind.

Beispiel 6.45 *Die Einheitskugel $\overline{B}_1$ in c_0 enthält keine Extrempunkte. Der Grund für diese Aussage ist, dass $\overline{B}_1$ in c_0 nicht kompakt ist. Man sieht dies so ein.*

Sei $x \in \overline{B}_1, x = (x^k)_{k \in \mathbb{N}}$; also $\lim_k x^k = 0$. Wähle $N \in \mathbb{N}$ mit $|x^N| < \frac{1}{2}$. Definiere $y, z \in c_0$ durch

$$y^k := z^k := x^k, k \neq N, y^N := x^N - \frac{1}{2}, z^N := x^N + \frac{1}{2}.$$

Dann gilt $y, z \in \overline{B}_1, x = \frac{1}{2}(y + z)$ und daher $x \notin \text{ext}(\overline{B}_1)$. □

6.5 Strikte und gleichmäßige Konvexität

Eine Eigenschaft, die einige geometrische Aussagen aus der Hilbertraumtheorie erlaubt, ist die strikte Konvexität in normierten Räumen. Sie besagt im wesentlichen, dass der Rand der Einheitskugel keine Gerade enthält.

Definition 6.46 Ein normierter Raum $(X, \|\cdot\|)$ heißt **strikt konvex**, falls aus $\|x\| = \|y\| = 1, \|x - y\| > 0$, stets $\|x + y\| < 2$ folgt. $\square$

Lemma 6.47 Sei $(X, \|\cdot\|)$ ein normierter Raum. Es sind äquivalent:

- (a) X ist strikt konvex.
- (b) Ist $\|x\| = \|y\| = 1, \|x - y\| > 0$, dann gilt für jedes $a \in (0, 1)$ $\|ax + (1 - a)y\| < 1$.
- (c) Ist $\|x + y\| = \|x\| + \|y\|, x \neq \theta, y \neq \theta$, dann gibt es $c > 0$ mit $x = cy$.

Beweis:

Zu (a) $\iff$ (b).

Seien $x, y \in X$ mit $\|x\| = \|y\| = 1, \|x - y\| > 0$, und $u := ax + (1 - a)y \in X$ mit $a \in (0, 1)$. O.E. $a \leq \frac{1}{2}$. Dann gilt offenbar $u = 2az + (1 - 2a)y$ mit $z := \frac{1}{2}(x + y)$ und daher folgt $\|u\| \leq 2a\|z\| + (1 - 2a)\|y\| < 2a + (1 - 2a) = 1$.

Damit ist (a) $\implies$ (b) gezeigt. Die Rückrichtung ist trivial.

Zu (a) $\implies$ (c).

Seien $x, y \in X$ mit $\|x + y\| = \|x\| + \|y\|, x \neq \theta, y \neq \theta$. Setze $x^1 := \|x\|^{-1}x, y^1 := \|y\|^{-1}y, a := \|x\|(\|x\| + \|y\|)^{-1}$. Dann ist $\|x^1\| = \|y^1\| = 1$ und

$$1 = \frac{\|x + y\|}{\|x\| + \|y\|} = \left\| \frac{1}{\|x\| + \|y\|}x + \frac{1}{\|x\| + \|y\|}y \right\| = \|ax^1 + (1 - a)y^1\|.$$

Wegen (b), was ja mit (a) äquivalent ist, folgt $x^1 = y^1$. Damit gilt $x = cy$ mit $c := \|x\|\|y\|^{-1}$.

Zu (c) $\implies$ (a).

Seien $x, y \in X$ mit $\|x\| = \|y\| = 1, \|x - y\| > 0$. Offenbar ist $\|x + y\| \leq 2$.

Annahme: $\|x + y\| = 2$. Dann ist $\|x + y\| = \|x\| + \|y\|$ und daher $x = cy$ für ein $c > 0$. Dann folgt aber $c = 1$ und $x = y$, was im Widerspruch zu $\|x - y\| > 0$ steht. $\blacksquare$

Offenbar sind die Räume $l_p, p = 1, \infty$ nicht strikt konvex, da der Rand der Einheitskugel „Kanten“ besitzen. Die Räume $l_p, 1 < p < \infty$ sind strikt konvex, sie sind sogar gleichmäßig konvex, besitzen also eine Eigenschaft, die strenger ist, das eigentlich wirklich geometrische Konzept ist und die wir unten einführen.

Beispiel 6.48 Der Banachraum $(C[a, b], \|\cdot\|)$ ist nicht strikt konvex. Dies folgt so:

O. E. $[a, b] = [0, 1]$. Wähle $x(t) := 1, y(t) := t, t \in [0, 1]$. Dann ist $x \neq y$ und $\|x\|_\infty = \|y\|_\infty = 1$. Aber es gilt $\|x + y\|_\infty = 2$.

Definition 6.49 Ein normierter Raum $(X, \|\cdot\|)$ heißt **gleichmäßig konvex**, falls gilt:

$$\forall \varepsilon > 0 \exists \delta > 0 \forall x, y \in X (\|x\| = \|y\| = 1, \|x - y\| \geq \varepsilon \implies \|\frac{1}{2}(x + y)\| \leq 1 - \delta)$$

$\square$

Bemerkung 6.50 Die beiden Definitionen 6.46, 6.49 bedürfen einer Kommentierung: Statt von einem strikt konvexen bzw. einem gleichmäßig konvexen Raum zu sprechen, sollten wir eigentlich von einer strikt konvexen bzw. gleichmäßig konvexen Norm sprechen. Diese Vorsicht ist umso mehr geboten, da es ja in einem Vektorraum zwei äquivalente Normen geben mag, von der eine die Eigenschaft haben mag, die andere nicht; es gibt solche Fälle; meist erreicht man dies durch „Renormierung“ einer gegebenen Norm. Also muss in jeder Situation klar sein, welche Norm gemeint ist. $\square$

Folgerung 6.51 Jeder gleichmäßig konvexe normierte Raum X ist strikt konvex.

Beweis:

Seien $x, y \in X, x \neq y$ mit $\|x + y\| = \|x\| + \|y\|$. O.E. $\|x\| = 1$. Definiere

$$x' := (x + y)\|x + y\|^{-1}, z := \frac{1}{2}(x + x'), a := \|z\|, b := \|x + y - z\|$$

Es gilt $\|x + y - x'\| = \|y\|$, also $b \leq \|y\|$. Ferner ist $a \leq 1, a + b \geq 1 + \|y\|$, also $a = \|\frac{1}{2}(x + x')\| = 1$. Aus der gleichmäßigen Konvexität folgt $x = x'$, also $x\|y\| = y$. ■

Lemma 6.52 Sei X ein normierter Raum. Es sind äquivalent:

(a) X ist gleichmäßig konvex.

(b) Für je zwei Folgen $(x^n)_{n \in \mathbb{N}}, (y^n)_{n \in \mathbb{N}}$ in X mit $\|x^n\| = \|y^n\| = 1, n \in \mathbb{N}$, gilt:

$$\text{Aus } \lim_n \|\frac{1}{2}(x^n + y^n)\| = 1 \text{ folgt } \lim_n (x^n - y^n) = \theta.$$

(c) Für je zwei Folgen $(x^n)_{n \in \mathbb{N}}, (y^n)_{n \in \mathbb{N}}$ in X mit $\limsup_n \|x^n\| \leq 1, \limsup_n \|y^n\| \leq 1, n \in \mathbb{N}$, gilt:

$$\text{Aus } \lim_n (\|\frac{1}{2}(x^n + y^n)\|) = 1 \text{ folgt } \lim_n (x^n - y^n) = \theta.$$

Beweis:

(a) $\implies$ (b) Dies ist klar.

(b) $\implies$ (c) Wir zeigen $\lim_n \|x^n\| = 1$ indirekt. Annahme: $\alpha := \liminf_k \|x^{n_k}\| < 1$. Dann haben wir eine Teilfolge $(x^{n_k})_{k \in \mathbb{N}}$ mit $\lim_k \|x^{n_k}\| = \alpha$. Dann folgt

$$\begin{aligned} 1 &= \lim_k \|\frac{1}{2}(x^{n_k} + y^{n_k})\| = \liminf_k \|\frac{1}{2}(x^{n_k} + y^{n_k})\| \\ &\leq \frac{1}{2} \liminf_k \|x^{n_k}\| + \frac{1}{2} \liminf_k \|y^{n_k}\| < 1, \end{aligned}$$

also ein Widerspruch. Damit wissen wir $\lim_n \|x^n\| = 1$ und aus Symmetriegründen $\lim_n \|y^n\| = 1$. O.E. können wir nun annehmen: $x^n \neq \theta, y^n \neq \theta, n \in \mathbb{N}$. Betrachte nun die Folgen

$$(x^n \|x^n\|^{-1})_{n \in \mathbb{N}}, (y^n \|y^n\|^{-1})_{n \in \mathbb{N}}.$$

Es gilt $\lim \|u^n + v^n - (x^n + y^n)\| = 0$, also $\lim_n \|u^n + v^n\| = \lim_n \|x^n + y^n\| = 2$. Nach (b) folgt $\lim_n \|u^n - v^n\| = 0$ und daher auch $\lim_n \|x^n - y^n\| = 0$.

(c) $\implies$ (a) Annahme: X ist nicht gleichmäßig konvex. Dann gibt es $\varepsilon > 0$ und zu jedem $\delta_n := 1/n$ Elemente $x^n, y^n \in X$ mit $\|x^n\|, \|y^n\| = 1, \|x^n - y^n\| \geq \varepsilon$, aber $\|\frac{1}{2}(x^n + y^n)\| > 1 - 1/n, n \in \mathbb{N}$. Die Existenz solcher Folgen $(x^n)_{n \in \mathbb{N}}, (y^n)_{n \in \mathbb{N}}$ steht dann aber im Widerspruch zu (c). ■

Lemma 6.53 Sei X gleichmäßig konvexer Banachraum. Ist $(x^n)_{n \in \mathbb{N}}$ eine Folge in X mit

$$\alpha := \lim_n \|x^n\| = \lim_{n,m} \|\frac{1}{2}(x^n + x^m)\|,$$

Dann konvergiert diese Folge.

Beweis:

Ist $\alpha = 0$, ist nichts zu beweisen. Sei nun $\alpha > 0$ und wir können o.E. annehmen $\alpha = 1$. Wegen der Vollständigkeit von X reicht es zu zeigen, dass $(x^n)_{n \in \mathbb{N}}$ eine Cauchyfolge ist. Andernfalls gibt es $\varepsilon > 0$ und $n_1 < m_1 < n_2 < m_2 < n_3 < \dots$ mit $\|x^{n_k} - x^{m_k}\| \geq \varepsilon, k \in \mathbb{N}$. Aus der Voraussetzung folgt, dass

$$\lim_k \|\frac{1}{2}(x^{n_k} + x^{m_k})\| = \lim_k \|x^{n_k}\| = \lim_k \|x^{m_k}\|.$$

Dies widerspricht (c) in Lemma 6.52. ■

Lemma 6.54 Sei X gleichmäßig konvexer Banachraum. Dann sind für eine Folge $(x^n)_{n \in \mathbb{N}}$ in X äquivalent:

- (a) $\lim_n x^n = x$.
(b) $w - \lim_n x^n = x$, $\lim_n \|x^n\| = \|x\|$.

Beweis:

(a) $\implies$ (b) Dies ist klar.

(a) $\implies$ (b) Für $x = \theta$ ist nichts zu beweisen. Sei $x \neq \theta$. O.E. $\|x\| = 1$.

Nach Folgerung 5.9 gibt es $\lambda \in X^*$ mit $\|\lambda\|_* = 1$, $\langle \lambda, x \rangle = \|x\| = 1$. Wegen $\|\frac{1}{2}(x^n + x^m)\| \leq \frac{\|x^n\|}{2} + \frac{\|x^m\|}{2}$, $m, n \in \mathbb{N}$, erhalten wir

$$\begin{aligned} 1 &\geq \limsup_{m,n} \|\frac{1}{2}(x^n + x^m)\| \geq \liminf_{m,n} \|\frac{1}{2}(x^n + x^m)\| \\ &\geq \liminf_{m,n} |\langle \lambda, \frac{1}{2}(x^n + x^m) \rangle| = |\langle \lambda, x \rangle| \\ &= \|x\| = 1 \end{aligned}$$

Also gilt $\lim_{m,n} \|\frac{1}{2}(x^n + x^m)\| = \lim_n \|x^n\| = 1$ und $(x^n)_{n \in \mathbb{N}}$ konvergiert nach Lemma 6.53. Wegen (b) ist $x = \lim_n x^n$. ■

Ein Hauptergebnis zur geometrischen Struktur von Banachräumen ist

Satz 6.55 (Milman (1938)-Pettis (1939)) Jeder gleichmäßig konvexe Banachraum ist reflexiv

Beweis:

Sei $\mu \in X^{**}$. O.E. $\|\mu\| = \|\mu\|_{X^* \rightarrow \mathbb{K}} = 1$. Wähle eine Folge $(\lambda_n)_{n \in \mathbb{N}}$ in X^* mit $\|\lambda_n\| \leq 1$, $n \in \mathbb{N}$, $\lim_n |\langle \mu, \lambda_n \rangle| = \|\mu\| = 1$. Nach dem Satz von Helly 6.28 gibt es zu jedem $m \in \mathbb{N}$ ein $x^m \in X$ mit $\|x^m\| \leq 1 + \frac{1}{m}$ mit

$$\langle \lambda_i, x^m \rangle = \langle \mu, \lambda_i \rangle, i = 1, \dots, m. \quad (6.11)$$

Sei $n \leq m$. Dann ist

$$\|x^n + x^m\| \geq |\langle \lambda_n, x^n \rangle + \langle \lambda_n, x^m \rangle| = 2|\langle \mu, \lambda_n \rangle|.$$

Nun folgt $\liminf_{n,m} \|\frac{1}{2}(x^n + x^m)\| \geq 1$, $\liminf \|x^n\| \geq 1$. Da $\limsup_{m,n} \|\frac{1}{2}(x^n + x^m)\| \leq 1$, $\limsup_n \|x^n\| \leq 1$, folgt hieraus $\lim_n \|x^n\| = \lim_{n,m} \|\frac{1}{2}(x^n + x^m)\| = 1$. Aus Lemma 6.53 folgt, dass $x := \lim_n x^n$ existiert. Nun folgt durch Grenzübergang $m \rightarrow \infty$ in (6.11)

$$\|x\| = 1, \langle \lambda_n, x \rangle = \langle \mu, \lambda_n \rangle, n \in \mathbb{N}.$$

Sei nun $\lambda_0 \in X^*$, $\|\lambda_0\| = 1$. Ersetze die Folge $(\lambda_n)_{n \in \mathbb{N}}$ durch die Folge $(\lambda_n)_{n \in \mathbb{N}_0}$. Dann finden wir wie oben ein $\hat{x} \in X$ mit

$$\|\hat{x}\| = 1, \langle \lambda_n, \hat{x} \rangle = \langle \mu, \lambda_n \rangle, n \in \mathbb{N}_0,$$

und es gilt

$$\|\frac{1}{2}(x + \hat{x})\| \geq |\langle \lambda_n, \frac{1}{2}(x + \hat{x}) \rangle| = |\langle \mu, \lambda_n \rangle| \text{ für alle } n \in \mathbb{N}_0,$$

und daher $\|\frac{1}{2}(x + \hat{x})\| = 1$. Dies impliziert mit der gleichmäßigen Konvexität $x = \hat{x}$. Insbesondere gilt $\langle \lambda_0, x \rangle = \langle \mu, \lambda_0 \rangle$. Da λ_0 als ein beliebig gewähltes Funktional in X^* , das nicht verschwindet, angesehen werden kann, haben wir $x \in X$ gefunden mit $\mu = J_X(x)$. ■

Welche Räume sind gleichmäßig konvex? Wir werden im nächsten Abschnitt die Prähilberträume kennenlernen als Beispiele gleichmäßig konvexer Räume. Die Reflexivität der vollständigen Prähilberträume ergibt sich (im nächsten Kapitel) auf direktem Weg, unabhängig von der gleichmäßigen Konvexität. In den Anhängen dieses Kapitels werden wir in einem „geometrischen“ Zusammenhang noch genauer klären können, welche Räume gleichmäßig konvex sind.

Wir wissen, dass $\mathbb{R}^2$ reflexiv ist, unabhängig von der gewählten Norm. Da die l_1 -Norm in $\mathbb{R}^2$ nicht gleichmäßig konvex ist, haben wir also ein Beispiel eines reflexiven Raumes, in dem es äquivalente Normen gibt, von denen eine nicht gleichmäßig konvex ist. Diese Aussage lässt sich noch verschärfen: Nach einem Resultat von M.M. Day gibt es einen separablen, reflexiven, strikt konvexen Banachraum, der zu keinem gleichmäßigem Banachraum isomorph ist; siehe [9].

Satz 6.56 (Approximationssatz) *Sei X ein Banachraum, sei $z \in X$ und C eine konvexe abgeschlossene Menge. Betrachte damit die Approximationsaufgabe*

$$\text{Gesucht } x \in C \text{ mit } \|z - x\| = \inf_{y \in C} \|z - y\|.$$

Dann gilt:

- (a) *Ist X strikt konvex, so besitzt diese Aufgabe höchstens eine Lösung.*
- (b) *Ist X gleichmäßig konvex, so besitzt diese Aufgabe eine eindeutig bestimmte Lösung.*

Beweis:

Sei $a := \inf_{y \in C} \|z - y\|$.

Zu (a) Seien z, z' Lösungen der Aufgabe. Dann gilt $\|z - x\| = \|z' - x\| = a$. Wegen $u := \frac{1}{2}(z + z') \in C$ ist $\|u - x\| \geq a$. Mit der Dreiecksungleichung folgt $\|u - x\| \leq a$. Also gilt $\|u - x\| = a$. Da X strikt konvex ist, folgt $u = z = z'$.

Zu (b) Da X nach Satz 6.55 reflexiv ist, existiert eine Lösung nach Satz 6.30. Die Eindeutigkeit folgt aus (a) mit Folgerung 6.51. ■

Definition 6.57 *Sei X ein normierter Raum. Die Abbildung*

$$\delta := \delta_X : [0, 2] \ni \varepsilon \mapsto \inf \left\{ 1 - \frac{1}{2}(\|x + y\|) \mid \|x\| = \|y\| = 1, \|x - y\| = \varepsilon \right\} \text{ in } \mathbb{R}$$

heißt **Konvexitätsmodul**. □

Sicherlich gilt $\delta_X(0) = 0$. Man kann einsehen, dass $\delta_X(\cdot)$ monoton nichtfallend ist. Als „Episode“ sei angefügt, dass für jeden Banachraum X gilt:

$$\delta_X(\varepsilon) \leq \delta_{l_2}(\varepsilon), \varepsilon \geq 0.$$

l_2 wird hier benutzt als ein Beispiel eines Prähilbertraumes; Prähilberträume haben alle denselben Konvexitätsmodul, wie wir im nächsten Kapitel sehen werden.

Die Räume $l_p, 1 < p < \infty$, sind gleichmäßig konvex. Die Berechnung der Konvexitätsmoduli für die Räume $l_p, 1 < p < \infty$ ergibt sich aus den Abschätzungen von Clarkson.

Lemma 6.58 *Sei $1 < p < \infty$ und sei $1 < q < \infty$ mit $\frac{1}{p} + \frac{1}{q} = 1$. Dann gilt für $a, b \in \mathbb{R}$:*

- (a) $|a + b|^p + |a - b|^p \leq 2^{p-1}(|a|^p + |b|^p)$ falls $2 \leq p < \infty$.
- (b) $|a + b|^q + |a - b|^q \leq 2(|a|^p + |b|^p)^{q-1}$ falls $1 < p \leq 2$.

Beweis:

Zu (a)

$$\begin{aligned}
|a+b|^p + |a-b|^p &\leq (|a+b|^2 + |a-b|^2)^{\frac{1}{2}} = \sqrt{2}(|a|^2 + |b|^2)^{\frac{1}{2}} \\
&\leq \sqrt{2} 2^{\frac{1}{2}-\frac{1}{p}}(|a|^p + |b|^p)^{\frac{1}{p}} = 2^{1-\frac{1}{p}}(|a|^p + |b|^p)^{\frac{1}{p}}
\end{aligned}$$

Zu b)

Diese Behauptung folgt aus der Gültigkeit der Ungleichung

$$(1+x)^q + (1-x)^q \leq 2(1+x^p)^{q-1}, \quad 0 \leq x \leq 1. \quad (6.12)$$

Beweis dazu:

Setze $f(\alpha, x) := (1 + \alpha^{1-q}x)(1 + \alpha x)^{q-1} + (1 + \alpha^{1-q}x)(1 - \alpha x)^{q-1}$, $0 \leq \alpha \leq 1, 0 \leq x \leq 1$. Dann ist $f(1, x)$ die linke Seite in (6.12), $f(x^{p-1}, x)$ die rechte Seite, da $(p-1)(q-1) = 1$. Man rechnet nach:

$$\frac{\partial f}{\partial \alpha}(\alpha, x) = (q-1)x + (1 - \alpha^{-p})((1 + \alpha x)^{q-2} - (1 - \alpha x)^{q-2}).$$

Es ist $(1 - \alpha^{-p}) \leq 0$, da $\alpha \leq 1$, $(1 + \alpha x)^{q-2} - (1 - \alpha x)^{q-2} \geq 0$, da $q \geq 2$. ■

Satz 6.59 Die Räume $l_p, 1 < p < \infty$, sind gleichmäßig konvex. Der Modul der Konvexität ist

$$\delta(\varepsilon) := 1 - (1 - (\frac{1}{2}\varepsilon)^q)^{\frac{1}{q}}, \quad 0 < \varepsilon \leq 2.$$

Beweis:Nutze Lemma 6.58 ■

Der Satz 6.59 ist auch richtig, wenn man l_p durch $L_p(\Omega)$ ersetzt.

6.6 Anhang: Differenzierbare Normen

Definition 6.60 Ein Banachraum X heißt **glatt**, wenn für alle $x \in X \setminus \{\theta\}$ es genau ein $\lambda \in X^*$ mit $\|\lambda\|_* = 1$ und $\langle \lambda, x \rangle = \|x\|$ gibt. □

Diese Definition ist unabhängig davon, ob der Raum reell oder komplex ist. Man kann zeigen, dass X glatt ist genau dann, wenn die „reelle Kopie“ $X_{\mathbb{R}}$ glatt ist. Wir können uns also wieder o. E. auf den reellen Fall konzentrieren.

Aus der Definition geht nicht hervor, weshalb Glattheit von X – sie drückt ja wohl etwas in Richtung von Differenzierbarkeit der Norm aus – etwas mit der Differenzierbarkeit der Norm in X zu tun haben sollte. Dies wird im Anhang 6.7 dahingehend aufgeklärt, dass Glattheit äquivalent mit der Gateaux-Differenzierbarkeit der Norm in $X \setminus \{\theta\}$ ist. Das Gateaux-Differential (Ableitung) $Df(x) \in X^*$ in x ist dann das nach Definition 6.60 existierende eindeutig bestimmte $\lambda \in X^*$; siehe Anhang 6.7.

Die geometrische Interpretation der Definition 6.6 ist aufschlussreich: λ beschreibt eine Hyperebene $H_{\lambda, \|x\|}$

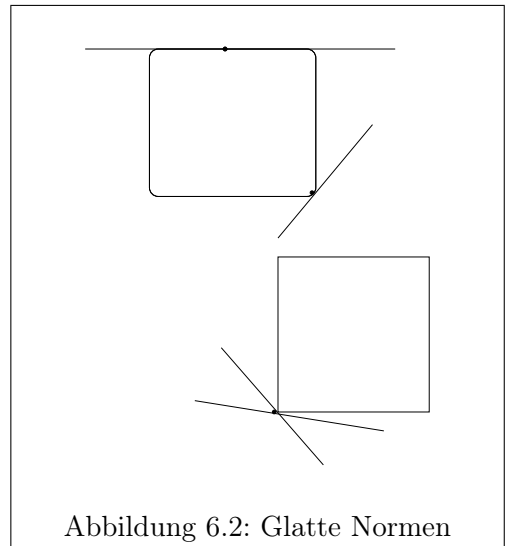


Abbildung 6.2: Glatte Normen

(in Normalform ($\|\lambda\|_* = 1!$)) durch den Punkt x mit $\theta \in H_{\lambda, \|x\|}^-$; wir nennen sie eine **Stützhyperebene** in x an die Kugel $\bar{B}_{\|x\|}$. Die Eindeutigkeit von λ mit dieser Eigenschaft schließt die Fälle aus, dass diese Kugel nicht „rund“, also Ecken hat; siehe Abbildung 6.2¹. Zur Illustration:

Beispiel 6.61 Betrachte die Normabbildung $f : l_1 \ni x = (x_k)_{k \in \mathbb{N}} \mapsto \|x\|_1 = \sum_{k=1}^{\infty} |x_k| \in \mathbb{R}$. Sei e^k der k -te Einheitsvektor in l_1 . Wir betrachten die Differenzierbarkeit in $x^0 = (x_k^0)_{k \in \mathbb{N}}$ „in Richtung“ $v = e^k$. Bestimme dazu

$$\lim_{h \rightarrow 0} \frac{f(x^0 + he^k) - f(x^0)}{h} = \lim_{h \rightarrow 0} \frac{|x_k^0 + h| - |x_k^0|}{h}.$$

Man erkennt, dass der Grenzwert genau dann existiert, wenn $x_k^0 \neq 0$ ist, und dass er dann gleich $\text{sign}(x_k^0)$ ist. Also ist sicherlich notwendig für irgendeine Differenzierbarkeit, dass $x_k^0 \neq 0$ für alle $k \in \mathbb{N}$ gilt. Als Kandidat für eine Ableitung kommt dann die Folge $(\text{sign}(x_k^0))_{k \in \mathbb{N}}$ als Element von $l_1^* = l_{\infty}$ in Frage. Man kann nun zeigen, dass damit die Gâteaux-Ableitung gemäß

$$Df(x^0)(v) = \sum_{k=1}^{\infty} \text{sign}(x_k^0) v_k, \quad v = (v_k)_{k \in \mathbb{N}},$$

gefunden ist. Dies geht so: Wir haben zu zeigen

$$\lim_{h \rightarrow 0} \sum_{k=1}^{\infty} \frac{|x_k^0 + hv_k| - |x_k^0|}{h} = \sum_{k=1}^{\infty} \text{sign}(x_k^0) v_k.$$

Dazu die Hilfsbetrachtung

$$\left| \frac{|x_k^0 + hv_k| - |x_k^0| - h \text{sign}(x_k^0) v_k}{h} \right| = \left| \frac{|1 + hv_k/x_k^0| - 1 - hv_k/x_k^0}{h/|v_k^0|} \right| \leq 2|v_k|, \quad k \in \mathbb{N},$$

da $|1 + a| - (1 + a) = 0$ für $a \geq -1$, und $= -2 - 2a$ für $a < -1$. Wählt man nun zu $\varepsilon > 0$ eine natürliche Zahl N so, dass $\sum_{k > N} |v_k| \leq \varepsilon$ gilt, erhält man

$$\left| \sum_{k=1}^{\infty} \frac{|x_k^0 + hv_k| - |x_k^0| - h \text{sign}(x_k^0) v_k}{h} \right| \leq \sum_{k=1}^N |\dots| + \sum_{N+1}^{\infty} |\dots|,$$

wobei die erste, endliche Summe mit $h \rightarrow 0$ gegen 0 strebt und die zweite Summe unabhängig von h durch 2ε majorisiert wird. Wegen

$$\frac{\|x^0 + v\|_1 - \|x^0\|_1 - Df(x^0)(v)}{\|v\|_1} = 1 \quad \text{für } v := x_k^0 e^k$$

liegt nach Lemma 6.71 (siehe Anhang 6.7) keine Fréchet-Differenzierbarkeit vor. □

Definition 6.62 Sei X ein normierter Raum. Die Abbildung

$$\rho := \rho_X : [0, \infty) \ni \tau \mapsto \frac{1}{2} \sup_{\|x\|=\|y\|=1} (\|x + \tau y\| + \|x - \tau y\| - 2) \in \mathbb{R}$$

heißt **Modul der Glattheit** von X . □

Man kann ziemlich schnell sehen, dass $\rho(\tau) \geq 0$ gilt für alle $\tau \geq 0$ und $\rho(0) = 0$.

¹Damit ist nun auch die Frage aus dem Epigraph zu diesem Kapitel aufgeklärt.

Lemma 6.63 *Sei X ein normierter Raum. Dann ist der Modul der Glattheit eine konvexe und stetige Funktion.*

Beweis:

Seien $x, y \in X$ mit $\|x\| = \|y\| = 1$. Betrachte

$$f_{x,y}(t) := \frac{\|x + ty\| + \|x - ty\|}{2} - 1.$$

Dann gilt für $a \in [0, 1]$

$$\begin{aligned} f_{x,y}(at + (1-a)s) &= \frac{\|x + (at + (1-a)s)y\| + \|x - (at + (1-a)s)y\|}{2} - 1 \\ &\leq \frac{a\|x + ty\| + (1-a)\|x + sy\| + a\|x - ty\| + (1-a)\|x - sy\|}{2} - 1 \\ &= af_{x,y}(t) + (1-a)f_{x,y}(s) \end{aligned}$$

Also ist $f_{x,y}$ konvex für alle x, y .

Sei $\varepsilon > 0$. Dann existieren $x, y \in X$ mit

$$\rho(at + (1-a)s) - \varepsilon \leq f_{x,y}(at + (1-a)s) \leq af_{x,y}(t) + (1-a)f_{x,y}(s) \leq a\rho(t) + (1-a)\rho(s).$$

Also ist ρ konvex, da $\varepsilon > 0$ beliebig war. Die Stetigkeit ist einfach nachzurechnen. ■

Definition 6.64 *Der Banachraum X heißt **gleichmäßig glatt**, falls mit dem Modul der Glattheit $\rho = \rho_X$ gilt:*

$$\lim_{\tau \rightarrow 0} \frac{\rho(\tau)}{\tau} = 0.$$

□

Es ist einfach einzusehen, dass X gleichmäßig glatt ist genau dann, wenn gilt:

$$\forall \varepsilon > 0 \exists \delta > 0 \forall x, y \in X (\|x\| = 1, \|y\| \leq 1 \implies \|x + y\| + \|x - y\| - 2 < \varepsilon\|y\|). \quad (6.13)$$

Es ist leicht einzusehen, dass ein Prähilbertraum gleichmäßig glatt ist.

Satz 6.65 *Jeder gleichmäßig glatte normierte Raum ist glatt.*

Beweis:

Annahme: X ist nicht glatt.

Dann gibt es $x^0 \in X$ und $\lambda, \mu \in X^*$, so dass $\lambda \neq \mu$, $\|\lambda\|_* = \|\mu\|_* = 1$ und $\langle \lambda, x^0 \rangle = \|x^0\| = \langle \mu, x^0 \rangle$. Wir können o. E. annehmen, dass $\|x^0\| = 1$ gilt. Sei $y^0 \in X$ mit $\|y^0\| = 1$ und $\langle \lambda - \mu, y^0 \rangle > 0$. Nun haben wir für $t > 0$

$$0 < t\langle \lambda - \mu, y^0 \rangle = \frac{1}{2}(\langle \lambda, x^0 + ty^0 \rangle + \langle \mu, x^0 - ty^0 \rangle) - 1 \leq \frac{1}{2}(\|x^0 + ty^0\| + \|x^0 - ty^0\|) - 1, \quad (6.14)$$

weshalb $0 < \langle \lambda - \mu, y^0 \rangle \leq \rho(t)/t$. Da $t > 0$ beliebig war, kann X nicht gleichmäßig glatt sein. ■

Nun kommen wir zum Zusammenhang von gleichmäßiger Konvexität und gleichmäßiger Glattheit. Ausgangspunkt sind die folgenden „Dualitätsformeln“.

Lemma 6.66 (Lindenstrauss-Tzafriri, 1978) *Sei X ein Banachraum. Dann haben wir für jedes x in X mit $\|x\| = 1$ und jedes $\lambda \in X^*$ mit $\|\lambda\|_* = 1$:*

$$\rho_{X^*}(\tau) = \sup\left\{\frac{1}{2}\tau\varepsilon - \delta_X(\varepsilon) \mid 0 < \varepsilon \leq 2\right\}, \tau > 0 \quad (6.15)$$

$$\rho_X(\tau) = \sup\left\{\frac{1}{2}\tau\varepsilon - \delta_{X^*}(\varepsilon) \mid 0 < \varepsilon \leq 2\right\}, \tau > 0 \quad (6.16)$$

Beweis:

Den Beweis verschieben wir in den Anhang 6.7. ■

Folgerung 6.67 Für jeden Banachraum X ist die Funktion $t \mapsto \rho_X(t)/t$ monoton nichtfallend.

Beweis:

Folgt aus der Formel (6.16). ■

Satz 6.68 Sei X ein Banachraum. Dann gilt:

- (a) X ist gleichmäßig glatt genau dann, wenn X^* gleichmäßig konvex ist.
- (b) X ist gleichmäßig konvex genau dann, wenn X^* gleichmäßig glatt ist.

Beweis:

Zu (a)

Sei X gleichmäßig glatt. Annahme: X^* ist nicht gleichmäßig konvex.

Dann gibt es $\varepsilon_0 \in (0, 2]$ mit $\delta_{X^*}(\varepsilon_0) = 0$ und wir erhalten aus 6.66 für alle $\tau > 0$ $0 < \frac{1}{2}\varepsilon_0 \leq \rho_X(\tau)/\tau$, was bedeutet, dass X nicht gleichmäßig glatt ist.

Sei X^* gleichmäßig konvex. Annahme: X ist nicht gleichmäßig glatt.

Dann gilt $\lim_{t \downarrow 0} \rho_X(t)/t \neq 0$, was bedeutet, dass ein $\varepsilon > 0$ existiert, so dass für alle $\delta > 0$ ein t existiert mit $0 < t < \delta$ und $t\varepsilon \leq \rho_X(t)$. Dann gibt es eine Folge $(\tau_n)_{n \in \mathbb{N}}$ mit $0 < \tau_n < 1$, $\rho_X(\tau_n) > \frac{1}{2}\varepsilon\tau_n$, $n \in \mathbb{N}$, $\lim_n \tau_n = 0$. Gemäß 6.66 erhalten wir zu jedem $n \in \mathbb{N}$ ein $\varepsilon_n \in (0, 2]$, so dass $\frac{1}{2}\varepsilon\tau_n \leq \frac{1}{2}\tau_n\varepsilon_n - \delta_{X^*}(\varepsilon_n)$. Dies impliziert $0 < \delta_{X^*}(\varepsilon_n) \leq \frac{1}{2}\tau_n(\varepsilon_n - \varepsilon)$, insbesondere $\varepsilon < \varepsilon_n$ und $\lim_n \delta_{X^*}(\varepsilon_n) = 0$. Da δ_{X^*} monoton nichtfallend ist, folgt $\delta_{X^*}(\varepsilon) \leq \delta_{X^*}(\varepsilon_n)$, $n \in \mathbb{N}$. Dies zeigt aber $\delta_{X^*}(\varepsilon) \leq 0$, was ein Widerspruch zur gleichmäßigen Konvexität ist.

Zu (b)

Der Beweis erfolgt entlang von (a) durch Vertauschung der Rollen von X und X^* . ■

Folgerung 6.69 Jeder gleichmäßig glatte Banachraum X ist reflexiv.

Beweis:

Ist X gleichmäßig glatt, so ist X^* gleichmäßig konvex und daher nach Satz 6.55 reflexiv. Nach Satz 5.18 ist X reflexiv. ■

6.7 Anhang: Differenzierbarkeit nichtlinearer Abbildungen

Kurz ein Abstecher in die nichtlineare Funktionalanalysis; siehe z. B. [8].

Definition 6.70 Seien X, Y normierte Räume, $U \subset X$ offen und $f : U \rightarrow Y$ eine Abbildung.

- (a) f heißt **Gâteaux-differenzierbar** in $x^0 \in U$, falls ein stetiger Operator $T : X \rightarrow Y$ existiert mit

$$\lim_{h \rightarrow 0} \frac{f(x^0 + hv) - f(x^0)}{h} = Tv \text{ für alle } v \in X. \quad (6.17)$$

- (b) f heißt **Fréchet-differenzierbar** in $x^0 \in U$, falls die Konvergenz in (6.17) gleichmäßig in $v \in B_1$ ist.
- (c) f heißt **Gâteaux- bzw. Fréchet-differenzierbar** in U , falls f in jedem $x^0 \in U$ Gâteaux- bzw. Fréchet-differenzierbar ist.

Der Grenzwert in (6.17) hängt von x^0 ab; wir schreiben $Df(x^0)$ statt T und nennen $Df(x^0)$ die Gâteaux- bzw. Fréchet-Ableitung von f in x^0 .

Die Fréchet-Ableitung reflektiert die Idee der linearen Approximation einer nichtlinearen Abbildung.

Lemma 6.71 *Seien X, Y normierte Räume, $U \subset X$ offen und $f : U \rightarrow Y$ eine Abbildung. Dann sind äquivalent für $x^0 \in U$:*

- (a) f ist Fréchet-differenzierbar in x^0 .
- (b) $f(x^0 + u) = f(x^0) + Df(x^0)u + r(u)$ mit $\lim_{\|u\| \rightarrow 0} r(u)\|u\|^{-1} = \theta$.

Beweis:

Den Beweis überlassen wir dem Leser, er ist ja aus der Analysis schon bekannt. ■

Uns interessiert hier insbesondere die Differenzierbarkeit der Norm in einem normierten Raum. An der Betragsfunktion sehen wir sofort die zu erwartenden Schwierigkeiten: der Nullvektor θ in einem normierten Raum sollte diesbezüglich eine Sonderrolle spielen; siehe Beispiel 6.61.

Lemma 6.72 *Sei X ein normierter Raum. Sei $x^0 \in X, \|x^0\| = 1$. Die Normfunktion $f : X \ni x \mapsto \|x\| \in \mathbb{R}$ ist Gâteaux-differenzierbar in x^0 genau dann, wenn es genau ein $\lambda \in X^*$ mit $\|\lambda\|_* = \langle \lambda, x^0 \rangle = 1$ gibt; in diesem Fall ist $Df(x^0) = \lambda$.*

Beweis:

Sei f Gâteaux-differenzierbar in $x^0 \in X \setminus \{\theta\}$ mit $\lambda := Df(x^0) \in X^*$. Aus der Definition der Gâteaux-Differenzierbarkeit folgt sofort $\langle \lambda, x^0 \rangle = 1$ und $\|\lambda\|_* \leq 1$, ja sogar $\|\lambda\|_* = 1$. Sei $\mu \in X^*$ ein weiteres Funktional mit diesen Eigenschaften. Für $v \in X$ und $h > 0$ gilt dann

$$\begin{aligned} \langle \lambda, v \rangle - \langle \mu, v \rangle &= \frac{\langle \lambda, x^0 + hv \rangle + \langle \mu, x^0 - hv \rangle - 2}{h} \\ &\leq \frac{\|x^0 + hv\| - 1}{h} + \frac{\|x^0 - hv\| - 1}{h} \end{aligned}$$

was mit $h \rightarrow 0$ gegen $\langle \lambda, v \rangle + \langle \mu, -v \rangle$ konvergiert. Daher ist $\langle \lambda, v \rangle - \langle \mu, v \rangle \leq 0$ für alle $v \in X$. Deshalb gilt (ersetze v durch $-v$) $\lambda = \mu$.

Wir beweisen nun die Umkehrung. Die einseitigen Grenzwerte

$$p^+(v) := \lim_{h \downarrow 0} \frac{\|x^0 + hv\| - 1}{h}, \quad p^-(v) := \lim_{h \uparrow 0} \frac{\|x^0 + hv\| - 1}{h},$$

existieren, denn die Funktion $g : \mathbb{R} \setminus \{0\} \ni h \mapsto \frac{\|x^0 + hv\| - 1}{h}$ ist monoton nicht fallend. Dazu der Beleg: Für $0 < h_1 < h_2$ folgt unter Nutzung der Dreiecksungleichung

$$\|x^0 + h_1 v\| - 1 \leq \frac{h_1}{h_2} \|x^0 + h_2 v\| + (1 - \frac{h_1}{h_2}) - 1 = h_1 \frac{\|x^0 + h_2 v\| - 1}{h_2}. \quad (6.18)$$

Analog sieht man die Monotonie in $(\infty, 0)$. Wegen $g(-h) \leq g(h)$, $h > 0$, ist alles gezeigt.

Nun beobachtet man $p^+(v) = -p^-(-v)$, $p^-(v) \leq p^+(v)$, $v \in X$, und dass p^+ sublinear ist. Um Gâteaux-Differenzierbarkeit zu zeigen, muss man nur noch $p^- = p^+$ zeigen. Sei dazu $v^0 \in X$ und sei $\alpha \in [p^-(v^0), p^+(v^0)]$. Definiere auf $U := \text{span}(\{v^0\})$ das lineare Funktional μ_0 durch $\langle \mu_0, av^0 \rangle := a\alpha$. Klar, $\langle \mu_0, av^0 \rangle \leq p^+(av^0)$, $a \in \mathbb{R}$. Mit dem Satz von Hahn-Banach 5.5 erhalten wir $\mu \in X^*$ mit $\mu|_U = \mu_0$, $\langle \mu, x \rangle \leq p^+(x)$, $x \in X$. Wegen

$$\langle \mu, x \rangle \leq p^+(x) \leq |p^+(x)| \leq \|x\|, \quad x \in X,$$

ist $\mu \in X^*$ und $\|\mu\|_* \leq 1$. Schließlich gilt

$$-\langle \mu, x^0 \rangle = \langle \mu, -x^0 \rangle \leq p^+(-x^0) = -1.$$

Damit ergibt sich $\|\mu\|_* = 1$. Da es nur ein Funktional $\lambda \in X^*$ mit diesen Eigenschaften gibt, muss $\mu = \lambda$ sein und wir erhalten $p^-(v^0) = \alpha = p^+(v^0)$. Da v^0 beliebig war, ist $p^- = p^+$. ■

Lemma 6.73 Sei X ein normierter Raum. Sei $x^0 \in X, \|x^0\| = 1$. Die Normfunktion $f : X \ni x \mapsto \|x\| \in \mathbb{R}$ ist Fréchet-differenzierbar in x^0 genau dann, wenn es genau ein $\lambda \in X^*$ mit folgenden Eigenschaften gibt:

- (a) $\|\lambda\|_* = \langle \lambda, x^0 \rangle = 1$;
- (b) ist $(\lambda_n)_{n \in \mathbb{N}}$ eine Folge in X^* mit $\|\lambda_n\|_* \leq 1, n \in \mathbb{N}$, und $\lim_n \langle \lambda_n, x^0 \rangle = 1$, so gilt $\lim_n \lambda_n = \lambda$ in X^* .

In diesem Fall ist $Df(x^0) = \lambda$.

Beweis:

Ist f Fréchet-differenzierbar in x^0 , dann ist f Gâteaux-differenzierbar in x^0 und die Existenz von λ gemäß (a) ist wegen Lemma 6.72 gesichert. Sei nun $(\lambda_n)_{n \in \mathbb{N}}$ eine Folge in X^* wie in (b). Sei $v \in X$. Dann gilt unter Verwendung von Lemma 6.71

$$\begin{aligned} \langle \lambda, v \rangle - \langle \lambda_n, v \rangle &= \langle \lambda, x^0 + v \rangle + \langle \lambda_n, x^0 - v \rangle - 1 - \langle \lambda_n, x^0 \rangle \\ &\leq \|x^0 + v\| - \|x^0\| + \|x^0 - v\| - \|x^0\| + 1 - \langle \lambda_n, x^0 \rangle \\ &= \langle \lambda, v \rangle + r(v) + \langle \lambda, -v \rangle + r(-v) + 1 - \langle \lambda_n, x^0 \rangle \\ &= r(v) + r(-v) + 1 - \langle \lambda_n, x^0 \rangle. \end{aligned}$$

Wegen $\lim_{\|v\| \rightarrow 0} r(v)/\|v\| = \theta$ erhält man zu $\varepsilon > 0$ ein $\delta > 0$, so dass

$$\|r(v)\| \leq \varepsilon \|v\| \text{ für alle } v \text{ mit } \|v\| \leq \delta.$$

Wählt man $N \in \mathbb{N}$ so, dass $1 - \langle \lambda_n, x^0 \rangle \leq \delta \varepsilon, n \geq N$, so folgt $\langle \lambda, v \rangle - \langle \lambda_n, v \rangle \leq 3\delta \varepsilon, n \geq N, \|v\| = \delta$. Daher erhalten wir

$$\|\lambda_n - \lambda\| = \delta^{-1} \sup_{\|v\|=1} (\langle \lambda, \delta v \rangle - \langle \lambda_n, \delta v \rangle) \leq 3\varepsilon, n \geq N.$$

Damit ist die notwendige Richtung gezeigt.

Wir beweisen die Rückrichtung. Sei $(v_n)_{n \in \mathbb{N}}$ eine Folge mit $\lim_n v_n = \theta$. Dann ist wegen Lemma ??

$$\lim_n \frac{\|x^0 + v_n\| - 1 - \langle \lambda, v_n \rangle}{\|v_n\|} = 0$$

zu zeigen. Da $\langle \lambda, x^0 \rangle = 1 = \|\lambda\|_*$, ist der Ausdruck ≥ 0 . Wähle Funktionale $\lambda_n \in \overline{B}_1 \subset X^*$ mit $\langle \lambda_n, x^0 + v_n \rangle = \|x^0 + v_n\|$. Wegen

$$\langle \lambda_n, x^0 \rangle = \langle \lambda_n, x^0 + v_n \rangle - \langle \lambda_n, v_n \rangle = \|x^0 + v_n\| - \langle \lambda_n, v_n \rangle$$

und $\lim_n \|v_n\| = 0$ ist $\lim_n \langle \lambda_n, x^0 \rangle = 1$. Damit folgt mit der Voraussetzung $\lim_n \|\lambda_n - \lambda\| = 0$. Deshalb

$$\begin{aligned} 0 &\leq \frac{\|x^0 + v_n\| - 1 - \langle \lambda, v_n \rangle}{\|v_n\|} = \frac{\langle \lambda_n - \lambda, v_n \rangle + \langle \lambda_n, x^0 \rangle - 1}{\|v_n\|} \\ &\leq \frac{\langle \lambda_n - \lambda, v_n \rangle}{\|v_n\|} \leq \|\lambda_n - \lambda\|_*. \end{aligned}$$

Also folgt mit $\lim_n \|\lambda_n - \lambda\|_* = 0$ das Gewünschte. ■

Beweis zu Lemma 6.66.

Wir beweisen die erste Formel.

Sei $\tau > 0, 0 < \varepsilon \leq 2, x, y \in X$ mit $\|x\| = \|y\| = 1$. Mit dem Satz von Hahn-Banach erhalten wir $\lambda, \mu \in X^*$ mit $\|\lambda\| = \|\mu\| = 1$, so dass

$$\langle \lambda, x + y \rangle = \|x + y\| \text{ und } \langle \mu, x - y \rangle = \|x - y\|.$$

Dann

$$\begin{aligned} \|x + y\| + \tau\|x - y\| - 2 &= \langle \lambda, x + y \rangle + \tau\langle \mu, x - y \rangle = \|x - y\| - 2 \\ &= \langle \lambda + \tau\mu, x \rangle + \langle \lambda - \tau\mu, y \rangle - 2 \\ &\leq \|\lambda + \tau\mu\|_* + \|\lambda - \tau\mu\|_* - 2 \\ &\leq \sup\{\|\lambda + \tau\mu\|_* + \|\lambda - \tau\mu\|_* - 2 \mid \|x\| = \|y\| = 1\} \\ &= 2\rho_X(\tau) \end{aligned}$$

Ist zusätzlich $\|x - y\| \geq \varepsilon$, dann erhalten wir

$$\frac{1}{2}\tau\varepsilon - \rho_{X^*} \leq 1 - \frac{1}{2}\|x + y\|.$$

Also ist $\frac{1}{2}\tau\varepsilon - \rho_{X^*}(\tau) \leq \delta_X(\varepsilon)$ und da $0 < \varepsilon \leq 2$ beliebig ist, folgt

$$\sup\{\frac{1}{2}\tau\varepsilon - \delta_X(\varepsilon) \mid 0 < \varepsilon \leq 2\} \leq \rho_{X^*}(\tau).$$

Seien nun $\lambda, \mu \in X^*$ mit $\|\lambda\|_* = \|\mu\|_* = 1$ und sei $\delta > 0$. Für $\tau > 0$ gibt es $x^0, y^0 \in X$ mit $x^0 \neq y^0$ und $\|x^0\| = \|y^0\| = 1$, so dass

$$\|\lambda + \tau\mu\|_* \leq \langle \lambda + \tau\mu, x^0 \rangle + \delta, \quad \|\lambda - \tau\mu\|_* \leq \langle \lambda - \tau\mu, y^0 \rangle + \delta.$$

Mit diesen Ungleichungen erhält man

$$\begin{aligned} \|\lambda + \tau\mu\|_* + \|\lambda - \tau\mu\|_* - 2 &\leq \langle \lambda + \tau\mu, x^0 \rangle + \langle \lambda - \tau\mu, y^0 \rangle - 2 + 2\delta \\ &= \langle \lambda, x^0 + y^0 \rangle + \langle \mu, x^0 - \tau y^0 \rangle - 2 + 2\delta \\ &\leq \|x^0 + y^0\| - 2 + \tau|\langle \mu, x^0 - y^0 \rangle| + 2\delta. \end{aligned}$$

Wir setzen $\varepsilon_0 := |\langle \mu, x^0 - y^0 \rangle|$. Dann ist $0 < \varepsilon_0 \leq \|x^0 + y^0\| \leq 2$ und

$$\begin{aligned} \frac{1}{2}(\|x^0 + \tau y^0\|_* + \|x^0 - \tau y^0\|_*) - 1 &\leq \frac{1}{2}\tau\varepsilon_0 + \delta - \delta_X(\varepsilon_0) \\ &\leq \delta + \sup\{\frac{1}{2}\tau\varepsilon_0 - \delta_X(\varepsilon) \mid 0 < \varepsilon \leq 2\}. \end{aligned}$$

Da $\delta > 0$ beliebig ist, folgt

$$\rho_{X^*}(\tau) \leq \sup\{\frac{1}{2}\tau\varepsilon - \delta_X(\varepsilon) \mid 0 < \varepsilon \leq 2\}.$$

Damit ist die erste Formel bewiesen.

Wir beweisen die zweite Formel.

Sei $\tau > 0$ und seien $\lambda, \mu \in X^*$ mit $\|\lambda\|_* = \|\mu\|_* = 1$. Für $\eta > 0$ existieren dann x^0, y^0 in X mit $\|x^0\| = \|y^0\| = 1$, so dass

$$\|\lambda + \mu\|_* - \eta \leq \langle \lambda + \mu, x^0 \rangle, \quad \|\lambda - \mu\|_* - \eta \leq \langle \lambda - \mu, y^0 \rangle.$$

Dann

$$\begin{aligned}
\|\lambda + \mu\|_* + \tau\|\lambda - \mu\|_* - 2 &\leq \langle \lambda + \mu, x^0 \rangle + \tau\langle \lambda - \mu, y^0 \rangle - 2\eta(1 + \tau) \\
&\leq \sup\{\|x + \tau y\| + \|x - \tau y\| - 2 \mid \|x\| = \|y\| = 1\} + \eta(1 + \tau) \\
&= 2\rho_X(\tau) + \eta(1 + \tau).
\end{aligned}$$

Falls $0 < \varepsilon \leq \|\lambda - \mu\| \leq 2$, dann haben wir

$$\frac{1}{2}\tau\varepsilon - \rho_X(\tau) - \eta(1 + \tau) \leq 1 - \frac{1}{2}\|\lambda + \mu\|_*,$$

was

$$\frac{1}{2}\tau\varepsilon - \rho_X(\tau) - \eta(1 + \tau) \leq \delta_{X^*}(\varepsilon).$$

Da $\eta > 0$ beliebig ist, erhalten wir

$$\frac{1}{2}\tau\varepsilon - \rho_X(\tau) \leq \delta_{X^*}(\varepsilon)$$

für alle $\varepsilon \in (0, 2]$ und schließlich

$$\sup\{\frac{1}{2}\tau\varepsilon - \rho_X(\tau) \mid 0 < \varepsilon \leq 2\} \leq \rho_X(\tau).$$

Seien $x, y \in X$ mit $\|x\| = \|y\| = 1$ und sei $\tau > 0$. Mit dem Satz von Hahn-Banach erhalten wir $\lambda, \mu \in X^*$ mit $\|\lambda\| = \|\mu\| = 1$, so dass

$$\langle \lambda, x + \tau y \rangle = \|x + \tau y\| \text{ und } \langle \mu, x - \tau y \rangle = \|x - \tau y\|.$$

Dann

$$\begin{aligned}
\|x + \tau y\| + \|x - \tau y\| - 2 &= \langle \lambda, x + \tau y \rangle + \langle \mu, x - \tau y \rangle - 2 \\
&= \langle \lambda + \mu, x \rangle + \tau\langle \lambda - \mu, y \rangle - 2 \\
&\leq \|\lambda + \mu\|_* + \|\lambda - \mu\|_* - 2 \\
&\leq \sup\{\|\lambda + \mu\|_* + \tau|\langle \lambda - \mu, y \rangle|\} - 2.
\end{aligned}$$

Wir definieren $\varepsilon_0 = |\langle \lambda - \mu, y \rangle|$. Dann ist $0 < \varepsilon_0 \leq \|x^0 - y^0\| \leq 2$ und

$$\begin{aligned}
\frac{1}{2}(\|x^0 + \tau y^0\| + \|x^0 - \tau y^0\|) - 1 &\leq \frac{1}{2}(\|\lambda + \mu\|_* + \tau|\langle \lambda - \mu, y \rangle|) - 1 \\
&= \frac{1}{2}\tau\varepsilon_0 - (1 - \frac{1}{2}\|\lambda + \mu\|_*) \\
&\leq \frac{1}{2}\tau\varepsilon_0 - \delta_{X^*}(\varepsilon_0) \\
&\leq \sup\{\frac{1}{2}\tau\varepsilon - \delta_{X^*}(\varepsilon) \mid 0 < \varepsilon \leq 2\}.
\end{aligned}$$

Dies impliziert

$$\sup\{\frac{1}{2}\tau\varepsilon - \rho_X(\tau) \mid 0 < \varepsilon \leq 2\} \geq \rho_X(\tau).$$

■

6.8 Anhang: Fixpunktprinzip II–Der Schaudersche Fixpunktsatz

Der Fixpunktsatz von Brouwer, den wir nun anführen, ist schwieriger zu beweisen als der Banachsche Fixpunktsatz. Wir müssen hier auf den Beweis verzichten.

Satz 6.74 (Fixpunktsatz von Brouwer, 1912) *Sei $M \subset \mathbb{R}^n$ konvex, abgeschlossen, beschränkt und nicht-leer. Sei $f : M \rightarrow M$ stetig. Dann besitzt f einen Fixpunkt d. h. es gibt $x \in M$ mit*

$$f(x) = x.$$

Hier ist eine Skizze der Beweisschritte.

Für $n = 1$ ist der Satz trivial, der Zwischenwertsatz erledigt schon alles. Der Beweis für $n \geq 2$ kann in mehreren Etappen erbracht werden:

1. $M = \overline{B}_r$ (o. E. $r = 1$). Siehe etwa [27].
2. $\text{int}(M) \neq \emptyset$ (o. E. $\theta \in \text{int}(M)$). Sei $M \subset \overline{B}_r$. Ziehe $\overline{B}_r$ auf M zusammen mittels $R : \overline{B}_r \rightarrow M$ stetig, $R(x) = x$ für alle $x \in M$ (Durchstoßpunkte sind wichtig). Ein Fixpunkt von $f \circ R : \overline{B}_r \rightarrow \overline{B}_r$ ist auch ein Fixpunkt von f , d. h. $x = (f \circ R)(x)$ genau dann, wenn $x = f(x)$.
3. M allgemein. Steige in der Dimension n ab und wende 2. an. Beachte: $M \neq \emptyset$ relativ zur affinen Hülle von M .

Die Übertragung von Satz 6.74 auf unendlich-dimensionale Räume ist ohne zusätzliche Voraussetzungen nicht möglich.

Beispiel 6.75 *Sei $X := l_2$. Wir setzen:*

$$Tx =: y, y_k := \begin{cases} (1 - \|x\|^2)^{1/2} & , \text{ falls } k = 1 \\ x_{k-1} & , \text{ falls } k > 1 \end{cases}, \text{ für } x = (x_k)_{k \in \mathbb{N}}.$$

Wegen $\|Tx\|^2 = (1 - \|x\|^2) + \sum_{k=1}^{\infty} |x_k|^2 = 1, x \in X$, gilt $T : \overline{B}_1 \rightarrow \overline{B}_1$. T ist stetig, denn $\|Tx - Ty\|^2 = \|\|x\|^2 - \|y\|^2\| + \|x - y\|^2, x, y \in X$. Aber T besitzt keinen Fixpunkt, denn aus $Tx = x, x = (x_k)_{k \in \mathbb{N}}$, folgt:

$$x_1 = (1 - \|x\|^2)^{1/2}, x_1 = x_2 = x_3 = \dots$$

Wegen $\|x\|^2 = \sum_{k=1}^{\infty} |x_k|^2 < \infty$, folgt $x_1 = x_2 = x_3 = \dots = 0$, d. h. $\|x\| = 0$. Dies ist im Widerspruch zu $x_1 = (1 - \|x\|^2)^{1/2} = 1$. ■

Die Verallgemeinerung von Satz 6.74 auf unendlich dimensionale Räume gelingt, wenn wir die Voraussetzungen

„ M abgeschlossen und beschränkt“ als „ M kompakt“

lesen.

Satz 6.76 (Fixpunktsatz von Schauder, 1930) *Sei X ein Banachraum und sei $M \subset X$ konvex, kompakt, nicht-leer, sei $T : M \rightarrow M$ stetig. Dann besitzt T einen Fixpunkt.*

Beweis:

$\overline{T(M)}$ ist eine abgeschlossene Teilmenge der kompakten Menge M kompakt und damit kompakt. Sei $n \in \mathbb{N}$. Offensichtlich gilt $\overline{T(M)} \subset \cup_{y \in T(M)} B_{\frac{1}{n}}(y)$. Da $\overline{T(M)}$ aber kompakt ist, gibt es endlich viele $y_1^n, \dots, y_m^n \in T(M)$ mit

$$T(M) \subset \overline{T(M)} \subset \cup_{i=1}^m B_{\frac{1}{n}}(y_i^n);$$

siehe 2.28. Also gilt

$$\min_{1 \leq i \leq m} \|T(x) - y_i^n\| < \frac{1}{n} \text{ für alle } x \in M. \quad (6.19)$$

Sei $M_n := \overline{\text{co}}(y_1^n, \dots, y_m^n)$. Dann ist M_n konvex, abgeschlossen, kompakt, $M_n \subset \overline{\text{co}}(T(M)) \subset \overline{M} = M$. Wir definieren:

$$a_i^n : M \ni x \mapsto \max\{0, \frac{1}{n} - \|T(x) - y_i^n\|\} \in \mathbb{R}, 1 \leq i \leq m.$$

Jedes a_i^n ist stetig und es gilt wegen (6.19) $\sum_{i=1}^m a_i^n(x) \neq 0$ für alle $x \in M$. Wir definieren den so genannten **Schauderoperator**:

$$P_n(x) := \left(\sum_{i=1}^m a_i^n(x) y_i^n \right) \left(\sum_{i=1}^m a_i^n(x) \right)^{-1}, x \in M.$$

Wir haben $P_n : M_n \rightarrow M_n$, da $y_1^n, \dots, y_m^n \in M_n$ und jedes $P_n(x)$ eine Konvexkombination der $y_1^n, \dots, y_m^n$ ist. P_n ist stetig, da $a_1^n, \dots, a_m^n$ stetig sind. Anwendung von Satz 6.74 ergibt die Existenz eines Punktes $x_n \in M_n$ mit $P_n x_n = x_n$. Da M kompakt ist, gibt es eine Teilfolge $(x_{n_k})_{k \in \mathbb{N}}$ und $x \in M$ mit $x = \lim_k x_{n_k}$. x ist Fixpunkt von T , denn:

$$\begin{aligned} \|x_{n_k} - Tx\| &= \|P_{n_k} x_{n_k} - Tx\| \leq \|P_{n_k} x_{n_k} - Tx_{n_k}\| + \|Tx_{n_k} - Tx\| \\ &= \left\| \left(\sum_{i=1}^m a_i^{n_k}(x) (y_i^{n_k} - Tx) \right) \left(\sum_{i=1}^m a_i^{n_k}(x) \right)^{-1} \right\| + \|Tx_{n_k} - Tx\| \\ &\leq \left(\sum_{i=1}^m a_i^{n_k}(x) \|y_i^{n_k} - Tx\| \right) \left(\sum_{i=1}^m a_i^{n_k}(x) \right)^{-1} + \|Tx_{n_k} - Tx\| \\ &\leq \left(\sum_{i=1}^m a_i^{n_k}(x) \frac{1}{n_k} \right) \left(\sum_{i=1}^m a_i^{n_k}(x) \right)^{-1} + \|Tx_{n_k} - Tx\| \quad (\text{beachte die Definition von } a_i) \\ &= \frac{1}{n_k} + \|Tx_{n_k} - Tx\| \end{aligned}$$

Daraus folgt $\lim_k \|x_{n_k} - Tx\| = 0$, $\|x - Tx\| = 0$, d. h. $x = Tx$. ■

Bemerkung 6.77 Der Banachsche Fixpunktsatz bedient sich einer nahezu rein analytischen Argumentation, die Beweise des Brouwerschen und in der Konsequenz des Schauderschen Fixpunktsatz verwenden daneben auch topologische Argumente. □

Folgerung 6.78 Sei X ein Banachraum und sei $M \subset X$ konvex, nicht-leer, sei $T : M \rightarrow M$ stetig, und sei $\overline{T(M)}$ kompakt. Dann besitzt T einen Fixpunkt.

Beweis:

Sei $N := \overline{\text{co}}(T(M))$. N ist kompakt und konvex; $N \subset M$. Die Einschränkung $T|_N : N \rightarrow N$ besitzt einen Fixpunkt nach Satz 6.76. Dieser Fixpunkt ist auch Fixpunkt von T . ■

Bemerkung 6.79 Die Fixpunktsätze von Brouwer und Schauder sagen nichts über die Eindeutigkeit eines Fixpunktes; sie ist auch ohne weitere Voraussetzungen und weitere Argumentation nicht zu erwarten. Dementsprechend ist auch zu erwarten, dass das so durchsichtige Berechnungsverfahren der sukzessiven Approximation durch Iteration im allgemeinen nicht zum Ziel führt. Es gibt aber durchaus auch Iterationsverfahren, die in der Lage sind, Fixpunkte einzuschließen; was man dazu braucht ist eine Halbordnung auf X und spezielle Voraussetzungen für die Abbildung, deren Fixpunkt man sucht.² $\square$

Bemerkung 6.80 Die Reichweite des Schauderschen Fixpunktsatzes ist außerordentlich groß: der Existenzsatz von Peano für Differentialgleichungen lässt sich damit beweisen, Existenzsätze für periodische Lösungen bei Differentialgleichungen basieren auf diesem Fixpunktsatz, ... $\square$

Das **Schaudersche Prinzip** leitet sich aus dem folgenden Satz ab.

Satz 6.81 Sei X ein Banachraum, sei $T : M \rightarrow M$ kompakt und es gebe $r > 0$ mit

$$\|x\| < r \text{ für alle } x \in X, \sigma \in [0, 1] \text{ mit } x = \sigma T(x). \quad (6.20)$$

Dann besitzt T einen Fixpunkt.

Beweis:

O. E. $r = 1$. Definiere

$$S : X \rightarrow X, S(x) := \begin{cases} T(x) & , \text{ falls } \|T(x)\| \leq 1 \\ \frac{T(x)}{\|T(x)\|} & , \text{ falls } \|T(x)\| > 1 \end{cases}.$$

Dann ist $S : \overline{B}_1 \rightarrow \overline{B}_1$ stetig. Da $\overline{T(\overline{B}_1)}$ kompakt ist, ist auch $\overline{S(\overline{B}_1)}$ kompakt. Nach Folgerung 6.78 hat S einen Fixpunkt z .

Annahme: $\|T(z)\| \geq 1$. Dann ist $z = S(z) = \sigma T(z)$, falls $\sigma = \|T(z)\|^{-1}$, und $\|z\| = \|S(z)\| = 1$, was ein Widerspruch zur Voraussetzung ist. Also ist $\|T(z)\| < 1$ und folglich $z = S(z) = T(z)$. $\blacksquare$

Die Bedingung (6.20) stellt eine a priori-Abschätzung für die Lösungen von skalierten Fixpunktgleichungen dar.

6.9 Übungen

- 1.) Der **Trennungssatz von Stone** lautet: Seien A, B konvexe Teilmengen im Vektorraum X über $\mathbb{R}$. Es gelte $A \cap B = \emptyset$. Dann gibt es konvexe Mengen C, D mit

$$A \subset C, B \subset D, C \cap D = \emptyset, D = X \setminus C.$$

Beweise diesen Satz. Hinweis: Betrachte die Mengen

$$\mathcal{C} := \{K \subset X \mid K \text{ konvex}, K \cap B = \emptyset, A \subset K\}, \mathcal{D} := \{K \subset X \mid K \text{ konvex}, K \cap C = \emptyset, B \subset K\}$$

und verwende das Zornsche Lemma.

- 2.) Sei X ein reeller Banachraum, $\lambda \in X^* \setminus \{\theta\}$ und sei $A := \{y \in X \mid \lambda(y) = 1\}$. Zeige:

²Siehe: Collatz, L. Funktionalanalysis und numerische Mathematik, Springer-Verlag, 1968

- (a) A ist nichtleer, konvex, abgeschlossen.
 (b) $\inf_{y \in A} \|y\| = \|\lambda\|_*^{-1}$.
 (c) Falls λ seine Norm nicht annimmt, d. h. falls es kein $x_0 \in X$ gibt mit $|\langle \lambda, x_0 \rangle| = \|\lambda\|_*$, so gibt es kein $y_0 \in A$ mit $\|y_0\| = \inf_{y \in A} \|y\|$.
- 3.) Sei
- $$A := \{(1, 0, 1), (1, 0, -1)\} \cup \{(\cos \theta, \sin \theta, 0) | \theta \in [0, 2\pi]\} \subset \mathbb{R}^3$$
- und setze $B := \text{co}(A)$. Zeige:
- (a) B kompakt.
 (b) $\text{ext}(B)$ ist nicht kompakt.
- Existiert ein solches Beispiel auch in $\mathbb{R}^2$?
- 4.) Sei $f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ unterhalbstetig, nach unten beschränkt und eigentlich. Sei $\varepsilon > 0, x^* \in X$ mit $f(x^*) \leq \inf\{f(x) | x \in X\} + \varepsilon$. Dann gibt es $\bar{x} \in X$ mit
- $f(\bar{x}) \leq f(x^*)$;
 - $|x^* - \bar{x}| \leq \gamma$;
 - $f(\bar{x}) \leq f(x) + \frac{\varepsilon}{\gamma}|x - \bar{x}|$ für alle $x \in \mathbb{R}^n$.
- Hinweis: Betrachte $g : x \mapsto f(x) + \frac{\varepsilon}{\gamma}|x - x^*|$.
- 5.) Sei $(X, \|\cdot\|)$ ein normierter Raum und sei $K \subset X$ konvex.
- (a) Zeige; Die offenen und abgeschlossenen Kugeln in X sind konvex.
 (b) Zeige: Sind $K_1, \dots, K_m$ konvexe Teilmengen von X , so ist auch $\cap_{j=1}^m K_j$ konvex.
- 6.) Sei X ein reeller Banachraum, $\lambda \in X^* \setminus \{\theta\}$ und sei $A := \{y \in X | \langle \lambda, y \rangle = 1\}$. Zeige:
- (a) A ist nichtleer, konvex, abgeschlossen.
 (b) $\inf_{y \in A} \|y\| = \|\lambda\|_*^{-1}$.
 (c) Falls λ seine Norm nicht annimmt, d. h. falls es kein $x_0 \in X$ gibt mit $|\langle \lambda, x_0 \rangle| = \|\lambda\|_*$, so gibt es kein $y_0 \in A$ mit $\|y_0\| = \inf_{y \in A} \|y\|$.
- 7.) Sei $(X, \|\cdot\|)$ ein normierter Raum und sei $S_1 := \{x \in X | \|x\| = 1\}$. Beweise: X ist strikt konvex genau dann, wenn $\text{ext}(\overline{B}_1) = S_1$ gilt.
- 8.) Sei
- $$A := \{(1, 0, 1), (1, 0, -1)\} \cup \{(\cos \theta, \sin \theta, 0) | \theta \in [0, 2\pi]\} \subset \mathbb{R}^3$$
- und setze $B := \text{co}(A)$. Zeige:
- (a) B kompakt.
 (b) $\text{ext}(B)$ ist nicht kompakt.
- Existiert ein solches Beispiel auch in $\mathbb{R}^2$?
- 9.) Sei $f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ unterhalbstetig, nach unten beschränkt und eigentlich. Sei $\varepsilon > 0, x^* \in X$ mit $f(x^*) \leq \inf\{f(x) | x \in X\} + \varepsilon$. Dann gibt es $\bar{x} \in X$ mit
- $f(\bar{x}) \leq f(x^*)$;
 - $|x^* - \bar{x}| \leq \gamma$;
 - $f(\bar{x}) \leq f(x) + \frac{\varepsilon}{\gamma}|x - \bar{x}|$ für alle $x \in \mathbb{R}^n$.
- Hinweis: Betrachte $g : x \mapsto f(x) + \frac{\varepsilon}{\gamma}|x - x^*|$.
- 10.) Seien X, Y ein Vektorraum über dem Skalkörper $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$. Zeige:

- (a) Ist $A \subset X$ konvex und $\alpha \in \mathbb{K}$, dann ist $\alpha K := \{\alpha x | x \in K\}$ konvex.
 - (b) Sind $A, B \subset X$ konvex, dann ist auch $A + B := \{x + y | x \in A, y \in B\}$ konvex.
 - (c) Ist $A \subset X$ konvex und $T : X \rightarrow Y$ linear, dann ist auch $T(A)$ konvex.
 - (d) Ist $B \subset Y$ konvex und $T : X \rightarrow Y$ linear, dann ist auch $T^{-1}(B)$ konvex.
- 11.) Seien X ein Vektorraum über dem Skalkörper $\mathbb{R}$. Eine Teilmenge K von X heisst konvexer Kegel, wenn K konvex ist und $\alpha K \subset K$ für alle $\alpha \in [0, \infty)$ gilt. Zeige: Ein konvexer Kegel K mit der Eigenschaft

$$K \cap (-K) = \{\theta\} \quad (\text{Positivität})$$

definiert eine Halbordnung auf X vermöge

$$x \geq y : \iff x - y \in K.$$

- 12.) Sei $(X, \|\cdot\|)$ ein normierter Raum und Y ein endlichdimensionaler Teilraum von X . Das **Tschebyscheff-Approximationsproblem** lautet:
Gegeben $x \in X$, gesucht $y \in Y$ mit $\|x - y\| = \inf_{z \in Y} \|x - z\|$.
- (a) Zeige, dass das Tschebyscheff-Approximationsproblem stets lösbar ist.
 - (b) Finde eine Situation, in der das Tschebyscheff-Approximationsproblem nicht eindeutig lösbar ist.
- 13.) Sei X ein Vektorraum über dem Skalkörper $\mathbb{R}$. Wir nennen eine Teilmenge A **affin**, wenn gilt:

$$x, y \in A \implies rx + (1 - r)y \in A \text{ für alle } r \in \mathbb{R}.$$

Sei $\mathcal{A} := \{A \subset X | A \text{ affin}\}$. Zeige, dass durch

$$A \parallel B : \iff \exists x \in X (A = x + B)$$

eine Äquivalenzrelation auf $\mathcal{A}$ definiert wird.

- 14.) Sei X ein Banachraum, versehen mit einer Halbordnung $\preceq$, sei $T : M \rightarrow M$, wobei M konvex sei. Sei $T = T_1 + T_2$ mit

$$T_1 u \preceq T_1 v, T_2 v \preceq T_2 u, \text{ falls } u \preceq v.$$

und seien T_1, T_2 stetig. Ausgehend von $v_0, w_0 \in M$ betrachte die Iteration

$$v_{n+1} := T_1 v_n + T_2 w_n, w_{n+1} := T_1 w_n + T_2 v_n, n \in \mathbb{N}_0.$$

Damit gelte $v_0 \preceq v_1 \preceq w_1 \preceq w_0$. Dann gilt:

- (a) $T(M_n) \subset M_n$, wobei $M_n := \{u \in M | v_n \preceq u \preceq w_n\}$, $n \in \mathbb{N}_0$.
 - (b) Ist M_n kompakt für ein $n \in \mathbb{N}$, so besitzt T einen Fixpunkt z und es gilt $z \in M_n$ für alle $n \in \mathbb{N}_0$.
- 15.) Sei X ein reeller Vektorraum, $C \subset X$ konvex und $p : X \rightarrow \mathbb{R}$ sublinear. Dann existiert $\lambda \in X'$ mit

$$\langle \lambda, x \rangle \leq p(x), x \in X, \inf_{u \in C} p(x) = \inf_{u \in C} \langle \lambda, u \rangle.$$

- 16.) Sei X ein Banachraum. X heisst **schwach gleichmäßig konvex** in $\lambda \in X^*$, $\|\lambda\|_* = 1$, falls für alle Folgen $(x^n)_{n \in \mathbb{N}}, (y^n)_{n \in \mathbb{N}}$ mit $\|x^n\| = \|y^n\| = 1, n \in \mathbb{N}$, und $\lim_n \langle \lambda, x^n + y^n \rangle = 2$ folgt: $\lim_n \|x^n - y^n\| = 0$. Zeige: X ist schwach gleichmäßig konvex in $\lambda \in X^*$, $\|\lambda\|_* = 1$, genau dann, wenn es für alle $\varepsilon > 0$ ein $\delta > 0$ gibt, so dass $\|x\| = \|y\| = 1, \|x - y\| \geq \varepsilon$ impliziert $\langle \lambda, x + y \rangle \leq 2(1 - \delta)$.

17.) Sei X ein Banachraum und sei $B := \overline{B}_1$. Zeige die Äquivalenz von

- (a) B schwach kompakt.
- (b) B schwach abzählbar kompakt.
- (c) B schwach folgenkompakt.
- (d) Ist $S \subset X$ schwach abgeschlossen, $B \cap S = \emptyset$, dann gilt

$$\text{dist}(B, S) = \inf\{\|b - x\| \mid b \in B, x \in S\} > 0.$$

- (e) B hat die Trennungseigenschaft.

18.) Sei X ein normierter Raum. Zeige:

- (a) Ist X^* strikt konvex, dann ist X strikt konvex.
- (b) Ist X reflexiv und strikt konvex, dann ist X^* strikt konvex.

19.) Sei X ein Banachraum und sei U ein linearer Teilraum von X . Zeige: ${}^\perp(U^\perp) = \overline{U}$.

20.) Seien X, Y Banachräume und sei $T \in \mathcal{B}(X, Y)$. Zeige: $\ker(T^*) = \text{ran}(T)^\perp$, $\ker(T) = {}^\perp\text{ran}(T^*)$, $\overline{\text{ran}(T)} = {}^\perp\ker(T^*)$.

21.) Sei U ein abgeschlossener linearer Teilraum des normierten Raumes X . Dann existieren (kanonische) isometrische Isomorphismen:

$$U^\perp = (X/U)^*, \quad U^* = X^*/U^\perp$$

22.) Sei X ein reflexiver Banachraum und sei V ein linearer Teilraum von X^* . Dann gilt:

$$({}^\perp V)^\perp = \overline{V}$$

23.) Sei X ein Banachraum und seien A, B lineare Teilräume von X . Dann gilt:

$$A \cap B = {}^\perp(A^\perp + B^\perp), \quad A^\perp \cap B^\perp = (A + B)^\perp$$

24.) Sei X ein Banachraum und seien A, B lineare Teilräume von X . Dann ist $A + B$ abgeschlossen in X genau dann, wenn $A + B = {}^\perp(A^\perp + B^\perp)$ gilt.

25.) Sei X ein normierter Raum und sei $C \subset X$ konvex und sei $C^\circ \neq \emptyset$. Zeige: $\overline{C}$ konvex, $(\overline{C})^\circ = C^\circ$, $\overline{(C^\circ)} = C$.

26.) Seien X, Y Banachräume und seien $S, T \in \mathcal{B}(X, Y)$. Es gebe $c_1, c_2 > 0$ mit

$$c_1 |\langle \lambda, Sx \rangle| \leq |\langle \lambda, Tx \rangle| \leq c_2 |\langle \lambda, Sx \rangle| \text{ für alle } \lambda \in X^*, x \in X.$$

Zeige: $\text{ran}(T)$ abgeschlossen genau dann, wenn $\text{ran}(S)$ abgeschlossen ist.

27.) Seien X, Y Banachräume und sei $T \in \mathcal{B}(X, Y)$ mit $\|Tx\| = \|x\|$ für alle $x \in X$. Seien X_1, Y_1 die abgeschlossenen Einheitskugeln in X bzw. Y . Zeige: $z \in \text{ext}(X_1)$ genau dann, wenn $Tz \in \text{ext}(Y_1)$.

28.) Seien X, Y Banachräume und seien $S, T \in \mathcal{B}(X, Y)$. S, T heißen norm-äquivalent, wenn es $c_1, c_2 > 0$ gibt mit

$$c_1 \|Tx\| \leq \|Tx\| \leq c_2 \|Sx\| \text{ für alle } x \in X.$$

Zeige: $\text{ran}(T)$ abgeschlossen genau dann, wenn $\text{ran}(S)$ abgeschlossen ist.

29.) Seien X, Y Banachräume und seien $S, T \in \mathcal{B}(X, Y)$. Es gebe $c_1, c_2 > 0$ mit

$$c_1 |\langle \lambda, Tx \rangle| \leq |\langle \lambda, Sx \rangle| \leq c_2 |\langle \lambda, Tx \rangle| \text{ für alle } \lambda \in Y^*, x \in X.$$

Zeige: Es gibt $a \in \mathbb{R}$ mit $S = aT$.

Hinweis: Sind y, y' linear unabhängig, dann gibt es zu $n \in \mathbb{N}$ ein $\lambda \in Y^*$ mit $\langle \lambda, y \rangle = n, \langle \lambda, y' \rangle = 0$; vergleiche nun mit der Voraussetzung.

30.) Sei X ein normierter Raum und C eine konvexe Teilmenge. Man finde einen Ansatz für die Beschreibung einer Eigenschaft „ A ist gleichmäßig konvex“.

31.) Betrachte die abgeschlossene Einheitskugel $\overline{B}_1$ in l_∞ . Zeige:

$$\text{ext}(\overline{B}_1) = \{x = (x_k)_{k \in \mathbb{N}} \mid |x_k| = 1 \text{ für alle } k \in \mathbb{N}\}.$$

Kapitel 7

Hilberträume

Zu Beginn des 19. Jahrhunderts wurde ein selbstzerstörendes demokratisches Prinzip in die Mathematik eingeführt (vor allem durch D. Hilbert), nachdem alle Axiomensysteme das gleiche Recht auf Analyse haben und der Wert einer Leistung nicht durch seine Bedeutung und seinen Nutzen für andere Disziplinen bestimmt wird, sondern allein durch seine Schwierigkeit, wie beim Bergsteigen. . . . Merkwürdige Fragen wie Fermats Problem oder Summen von Primzahlen wurden zu angeblich zentralen Problemen

V.I. Arnold

Hilberträume sind bei der Modellierung und Lösung von Problemen der theoretischen und angewandten Naturwissenschaften von überragender Bedeutung. Ihre Theorie ist nicht zuletzt entstanden durch das Bedürfnis der Quantenmechanik ein sicheres mathematisches Fundament zu geben. Wir ordnen die Hilberträume in die Banachraumtheorie ein, betrachten die spezifischen zusätzlichen geometrischen Möglichkeiten und werfen einen Blick auf das Spektrum linearer beschränkter Operatoren.

7.1 Prähilberträume

Definition 7.1 Sei X ein Vektorraum über $\mathbb{K}$. Eine Abbildung $\langle \cdot | \cdot \rangle : X \times X \longrightarrow \mathbb{K}$ heißt ein **inneres Produkt** oder **Skalarprodukt** auf X , wenn folgende Bedingungen erfüllt sind:

a) $\langle x | x \rangle > 0$ für alle $x \in X \setminus \{\theta\}$; (Definitheit)

b) $\langle x | y \rangle = \overline{\langle y | x \rangle}$ für alle $x, y \in X$; (Antisymmetrie)

c) $\langle ax + by | z \rangle = a \langle x | z \rangle + b \langle y | z \rangle$ für alle $x, y, z \in X, a, b \in \mathbb{K}$. (Linearität)

Einen Vektorraum X zusammen mit einem Skalarprodukt $\langle \cdot, \cdot \rangle$ bezeichnet man als **euklidischen Raum**, falls $\mathbb{K} = \mathbb{R}$ ist und als **unitären Raum**, falls $\mathbb{K} = \mathbb{C}$ ist. Allgemein sprechen wir bei $(X, \langle \cdot | \cdot \rangle)$ von einem **Prähilbertraum**¹. $\square$

Wegen a) ist $\langle x | x \rangle \in \mathbb{R}$ für alle $x \in X$. Aus b) und c) folgt, dass ein inneres Produkt eine so genannte **Sesquilinearform** ist: linear im ersten Argument, *konjugiert linear* im 2. Argument. Im

¹Die Bezeichnungsweise $\langle \cdot | \cdot \rangle$ für ein Skalarprodukt ist an eine Notationsgewohnheit der Physiker angelehnt.

Falle des Skalkörpers $\mathbb{R}$ liegt auch Linearität im 2. Argument vor und statt von Antisymmetrie sprechen wir bei b) von Symmetrie. Aus

$$\langle \theta | z \rangle = \langle z - z | z \rangle = \langle z | z \rangle - \langle z | z \rangle = 0, \quad z \in X,$$

lesen wir ab, dass $\langle z, z \rangle = 0$ genau dann ist, wenn $x = \theta$ ist.

Beispiel 7.2 Auf dem Vektorraum $\mathbb{R}^n$ haben wir das Skalarprodukt

$$\langle \cdot | \cdot \rangle_2 : \mathbb{R}^n \times \mathbb{R}^n \ni (x, y) \mapsto \sum_{i=1}^n x_i y_i \in \mathbb{R}.$$

Es heißt **euklidisches Skalarprodukt**. Dass in der Tat ein Skalarprodukt vorliegt, verifiziert man leicht. $\square$

Lemma 7.3 (Cauchy–Schwarzsche Ungleichung) Sei $(X, \langle \cdot | \cdot \rangle)$ ein Prähilbertraum. Dann gilt für alle $x, y \in X$

$$|\langle x | y \rangle| \leq \sqrt{\langle x | x \rangle} \sqrt{\langle y | y \rangle},$$

wobei Gleichheit genau dann gilt, wenn x, y linear abhängig sind.

Beweis:

Seien $x, y \in X$. Ist $y = \theta$, dann ist die Aussage schon klar. Sei also nun $y \neq \theta$. Offenbar gilt $0 \leq \langle x - ay | x - ay \rangle$ für alle $a \in \mathbb{K}$, also

$$0 \leq \langle x | x \rangle - 2\Re a \langle x | y \rangle + |a|^2 \langle y | y \rangle.$$

Setze $a := -\overline{\langle x | y \rangle} \langle y | y \rangle^{-1}$. Dann folgt

$$0 \leq \langle x | x \rangle - \frac{|\langle x | y \rangle|^2}{\langle y | y \rangle},$$

woraus wir die Aussage nun ablesen.

Die Aussage über Gleichheit folgt aus der Tatsache, dass $0 = \langle x - ay | x - ay \rangle$ genau dann gilt, wenn $x = ay$ ist. $\blacksquare$

Folgerung 7.4 Ist $(X, \langle \cdot | \cdot \rangle_X)$ ein Prähilbertraum, dann wird auf X eine Norm definiert durch

$$\|x\|_X := \sqrt{\langle x | x \rangle_X}, \quad x \in X.$$

Beweis:

Die Normeigenschaften sind bis auf die Dreiecksungleichung sofort einsichtig. Die Dreiecksungleichung folgt mit der Cauchy–Schwarzschen Ungleichung so:

$$\|x + y\|_X^2 = \|x\|_X^2 + 2\Re \langle x | y \rangle_X + \|y\|_X^2 \leq \|x\|_X^2 + 2\|x\|_X \|y\|_X + \|y\|_X^2 = (\|x\|_X + \|y\|_X)^2.$$

$\blacksquare$

Sei $(X, \langle \cdot | \cdot \rangle_X)$ ein euklidischer Raum. Wir wissen aus der Cauchy–Schwarzschen Ungleichung $|\langle x | y \rangle| \leq \|x\|_X \|y\|_X$ für alle $x, y \in X$, also

$$-1 \leq \frac{\langle x | y \rangle_X}{\|x\|_X \|y\|_X} \leq 1 \quad \text{für alle } x, y \in X \setminus \{\theta\}.$$

Nun gibt es zu $x, y \in X \setminus \{\theta\}$ also einen eindeutig bestimmten Winkel $\vartheta = \vartheta(x, y)$ mit

$$\frac{\langle x | y \rangle_X}{\|x\|_X \|y\|_X} = \cos(\vartheta(x, y)), \quad \vartheta(x, y) \in [0, \pi].$$

Wir nennen $\vartheta(x, y)$ den **Winkel** zwischen x und y .

Wir haben hier einen euklidischen Raum betrachtet, damit die Winkeldefinition gelingt. Ein inhaltliches Analogon zum Winkelbegriff im unitären Raum gibt es nicht.

Definition 7.5 Sei $(X, \langle \cdot | \cdot \rangle)$ ein euklidischer Raum. $x, y \in X \setminus \{\theta\}$ heißen **orthogonal**, wenn $\langle x | y \rangle = 0$ gilt. $\square$

Wir sehen also, dass Orthogonalität von x, y in einem euklidischen Raum gerade bedeutet, dass der Winkel zwischen x, y ein Rechter, also $\frac{\pi}{2}$ ist.

Im $\mathbb{R}^2$, einem „Ausschnitt“ unseres Anschauungsraumes $\mathbb{R}^3$, mit dem üblichen euklidischen Skalarprodukts haben wir den sogenannten Kosinussatz. Dieser besagt, dass in einem Dreieck, beschrieben durch die Ortsvektoren $x, y, x - y$, gilt:

$$\textbf{Kosinussatz: } |x - y|^2 = |x|^2 + |y|^2 - 2 \cos(\vartheta(x, y)) |x| |y|$$

Das Skalarprodukt macht also geometrische Anschauung möglich. Hier sind noch zwei Identitäten, die auch etwas mit Elementargeometrie zu tun haben. In einem euklidischen Raum $(X, \langle \cdot | \cdot \rangle_X)$ haben wir:

$$\textbf{Parallelogrammidentität: } \|x + y\|_X^2 + \|x - y\|_X^2 = 2\|y\|_X^2 + 2\|x\|_X^2, \quad x, y \in X.$$

$$\textbf{Satz des Pythagoras: } \|x + y\|_X^2 = \|y\|_X^2 + \|x\|_X^2 \text{ für } x, y \in X \text{ mit } \langle x | y \rangle_X = 0.$$

Diese Identitäten ergeben sich durch einfaches Rechnen mit dem Skalarprodukt. Daraus ergibt sich der Konvexitätsmodul eines Prähilbertraumes X in einfacher Weise:

$$\delta_X(\varepsilon) = 1 - \sqrt{1 - \frac{\varepsilon^2}{4}}, \quad 0 \leq \varepsilon \leq 2. \quad (7.1)$$

Prähilberträume sind also gleichmäßig konvex. Achtung: sie sind nicht notwendigerweise reflexiv, da sie ja nicht notwendigerweise vollständig sind.

Bemerkung 7.6 Sei $(X, \|\cdot\|)$ ein normierter Raum. Die Norm $\|\cdot\|$ wird durch ein Skalarprodukt $\langle \cdot | \cdot \rangle$ gemäß $\|x\| := \sqrt{\langle x, x \rangle}$ induziert genau dann, wenn in X die Parallelogrammidentität gilt. Das Skalarprodukt ist dann definiert durch

$$\langle x | y \rangle := \frac{1}{4}(\|x + y\|^2 - \|x - y\|^2) + \frac{i}{4}(\|x + iy\|^2 - \|x - iy\|^2), \quad x, y \in X.$$

$\square$

Ist X ein Vektorraum über dem Skalarkörper $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, so können wir nach Abbildungen suchen, die aus einem Paar von Vektoren aus X einen Skalar erzeugen, also im Sinne der Physik „gerichteten Größen“ eine „ungerichtete Größe“ zuordnet.

Erinnern wir noch an das **Kreuzprodukt**. Im Unterschied zum Skalarprodukt erzeugt es aus zwei Vektoren wieder einen Vektor. Wir beschränken uns auf den Anschauungsraum $\mathbb{R}^3$.

Seien $u = (u_1, u_2, u_3), v = (v_1, v_2, v_3) \in \mathbb{R}^3$. Wir definieren $w := u \times v$ durch

$$w_1 := u_2 v_3 - u_3 v_2, \quad w_2 := u_3 v_1 - u_1 v_3, \quad w_3 := u_1 v_2 - u_2 v_1.$$

Anschaulich ist $w = u \times v$ ein Vektor, der auf der Ebene, die durch u, v aufgespannt wird, senkrecht steht. Dies bestätigt man durch Nachrechnen von

$$\langle u, u \times v \rangle = \langle v, u \times v \rangle = 0.$$

Nun bleiben im $\mathbb{R}^3$ „zwei Seiten“ für den orthogonalen Vektor $u \times v$. Es ist nun gerade so eingerichtet, dass die **Dreifingerregel der rechten Hand** die drei Vektoren $u, v, u \times v$ wiedergibt. Weitere Regeln, die man leicht bestätigt, sind $(x, y, z \in \mathbb{R}^3)$:

$$z \times (x \times y) = \langle z | y \rangle x - \langle z | x \rangle y \quad (7.2)$$

$$\langle x | y \rangle^2 + |x \times y|^2 = |x|^2 |y|^2 \quad (7.3)$$

$$|x \times y| = |x| |y| \sin(\vartheta(x, y)) \quad (7.4)$$

Die Identität (7.3) ist eine Verschärfung der Cauchy–Schwarzschen Ungleichung. Die Identität in (7.4) ist der sogenannte **Sinussatz**; dabei ist $\vartheta(x, y)$ der Winkel zwischen x und y . Da die rechte Seite von (7.4) gerade die Fläche des von x, y aufgespannten Parallelogramms ist, erhält $|x \times y|$ eine neue Interpretation.

Ist eine Ebene $\mathcal{E}$ im $\mathbb{R}^3$ gegeben (in Parameterdarstellung) durch

$$\mathcal{E} := \{p + su + tv | s, t \in \mathbb{R}\} \text{ (} p \text{ ist Aufpunkt, } u, v \text{ sind Richtungsvektoren)},$$

so ist der Abstand eines Punktes q von der Ebene $\mathcal{E}$ gegeben durch

$$\text{dist}(q, \mathcal{E}) = \frac{|\langle u \times v, q - p \rangle|}{|u \times v|}. \quad (7.5)$$

In der Physik ist der Unterschied „Skalar/Vektor“ ganz bedeutend – wir haben ihn hier anhand der verschiedenen Produkte kennengelernt. Die Verabredung ist:

Skalares Feld: Jedem $x \in \mathbb{R}^3$ wird ein Skalar $U(x) \in \mathbb{R}$ zugeordnet.

Vektorielltes Feld: Jedem $x \in \mathbb{R}^3$ wird ein Vektor $V(x) \in \mathbb{R}^3$ zugeordnet.

Beispiele für skalare Felder sind:

- Temperaturverteilung im Raum;
- Druckwert in der Atmosphäre;
- Potential einer elektrischen Punktladung;
- Zentralfeld der Gravitation.

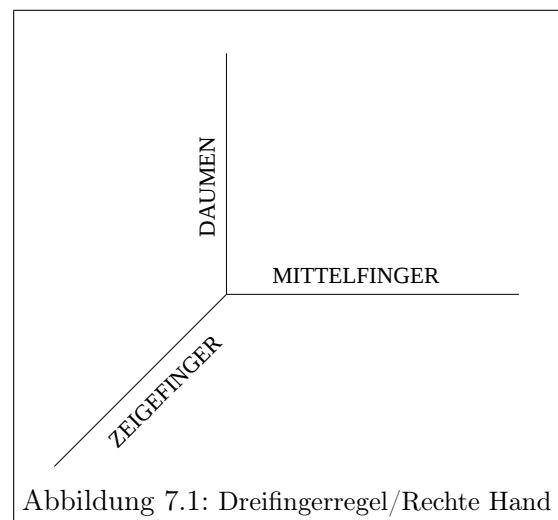


Abbildung 7.1: Dreifingerregel/Rechte Hand

Beispiele für vektorielle Felder sind:

- Gravitationsfeld eines Massenpunktes;
- Geschwindigkeitsverteilung in einer strömenden Flüssigkeit;
- Magnetfeld der Erde.

7.2 Hilberträume: Definition und Beispiele

Sei $(X, \langle \cdot | \cdot \rangle)$ ein Prähilbertraum. Wir haben unter Nutzung der Cauchy–Schwartzschen Ungleichung gesehen, dass X zu einem normierten Raum $(X, \| \cdot \|)$ wird, wobei die Norm definiert ist durch

$$\|x\| := \sqrt{\langle x | x \rangle}, \quad x \in X.$$

Diesen normierten Raum können wir nun auf Vollständigkeit hinterfragen.

Definition 7.7 Der Prähilbertraum $(X, \langle \cdot | \cdot \rangle_X)$ heißt **Hilbertraum**, wenn der damit definierte normierte Raum $(X, \| \cdot \|_X)$ vollständig ist. $\square$

Die Bezeichnung **Prähilbertraum** für einen Raum mit Skalarprodukt erklärt sich nun aus der Definition 7.7.

Wir wissen schon, dass der euklidische Raum $(\mathbb{R}^n, |\cdot|_2)$ bezüglich der euklidischen Norm $|\cdot|_2$, die ja durch ein Skalarprodukt induziert wird, vollständig ist.

Beispiel 7.8 Die Prähilberträume $(l_2, \langle \cdot | \cdot \rangle_2)$, $(L_2(\Omega), \|\cdot\|_2)$ sind Hilberträume. Das innere Produkt in $L_2(\Omega)$ ist gegeben durch

$$\langle f | g \rangle_p := \int_{\Omega} \langle f(x) | g(x) \rangle_2 dx, \quad f, g \in L_2(\Omega).$$

□

Beispiel 7.9 Wir wissen schon, dass der euklidische Raum $(\mathbb{R}^n, |\cdot|_2)$ bezüglich der euklidischen Norm $|\cdot|_2$, die ja durch ein Skalarprodukt induziert wird, vollständig ist.

Das euklidische Skalarprodukt auf $\mathbb{R}^n$ ist nur ein Spezialfall einer großen Familie von Skalarprodukten: Mit einer Matrix $A \in \mathbb{R}^{n,n}$ machen wir den Ansatz

$$\langle x | y \rangle_A := \langle Ax | y \rangle_2, \quad x, y \in \mathbb{R}^n. \quad (7.6)$$

Damit diese Abbildung $\langle \cdot | \cdot \rangle_A$ symmetrisch wird, muss A symmetrisch sein, d. h. $A = A^t$ gelten, wobei A^t die transponierte Matrix ist. Damit $\langle \cdot | \cdot \rangle_A$ die Definitheitsbedingung erfüllt, muss A positiv definit sein, d. h. $\langle Ax | x \rangle_2 > 0$ für alle $x \in \mathbb{R}^n \setminus \{0\}$ gelten. Sind nun diese Bedingungen erfüllt, dann liegt in $\langle \cdot | \cdot \rangle_A$ in der Tat ein Skalarprodukt vor. □

Beispiel 7.10 $C[a, b]$ kann durch

$$\langle f | g \rangle := \int_a^b f(t)g(t) dt, \quad f, g \in C[a, b],$$

zu einem Prähilbertraum gemacht werden. Hier liegt aber kein Hilbertraum vor, wie wir an der Folge $(f_n)_{n \in \mathbb{N}}$ in $C[-1, 1]$ definiert als

$$f_n(t) := \begin{cases} 0 & , \text{ falls } t \leq 0 \\ nt & , \text{ falls } 0 < t \leq \frac{1}{n}, \quad n \in \mathbb{N}, \\ 1 & , \text{ falls } \frac{1}{n} < t \leq 1 \end{cases}$$

sehen können. Es liegt zwar eine Cauchyfolge vor, wie man aus

$$\int_{-1}^1 |f_n(t) - f_m(t)|^2 dt \leq \frac{1}{3} \left(\frac{1}{n^2} + \frac{1}{m} \right) \quad \text{für } n \geq m$$

abliest. Aber diese Folge konvergiert in jedem Punkt gegen $f : [-1, 1] \rightarrow \mathbb{R}$ mit

$$f(t) := \begin{cases} 0 & , \text{ falls } t \leq 0 \\ 1 & , \text{ falls } 0 < t \end{cases}$$

d. h. als Grenzwert im Sinne der Norm $\|\cdot\|_{\sigma}$ kommt nur f in Frage. f ist aber nicht stetig. Übrigens, $C[a, b]$ versehen mit der Norm

$$\|f\|_1 := \int_a^b |f(t)| dt, \quad f, g \in C[a, b],$$

ist auch nicht vollständig, wie wir analog sehen können. ■

Folgerung 7.11 *Jeder Hilbertraum ist gleichmäßig konvex und daher auch reflexiv.*

Beweis:

Wir haben dies schon vorbereitet: siehe (7.1) und Satz 6.55. ■

7.3 Der Darstellungssatz von Riesz

Wir werfen erneut einen Blick auf das so genannte Approximationsproblem.

Satz 7.12 *Sei $(X, \langle \cdot | \cdot \rangle)$ ein Hilbertraum und sei $K \subset X$ konvex, abgeschlossen und nichtleer. Dann gibt es zu jedem $x \in X$ ein eindeutig bestimmtes $y \in K$ mit*

$$\|x - y\| = \text{dist}(x, K) = \inf_{z \in K} \|x - z\|. \quad (7.7)$$

Zusatz: Es gilt

$$\Re(\langle x - y | z - y \rangle) \leq 0 \text{ für alle } z \in K. \quad (7.8)$$

Beweis:

Der Satz 6.56 erledigt die Existenz- und Eindeutigkeitsfrage, wir zeigen den Zusatz.

Sei $z \in K$. Dann ist $z_t := tz + (1 - t)y = y + t(z - y)$ in K für alle $t \in (0, 1)$. Daher ist

$$\|x - y\|^2 \leq \|x - z_t\|^2 = \|x - y\|^2 - 2t\langle x - y, z - y \rangle + t^2\|z - y\|^2, \quad t \in (0, 1),$$

und daher

$$0 \leq -2\Re(t\langle x - y | z - y \rangle) + t^2\|z - y\|^2, \quad t \in (0, 1),$$

was nach Kürzen mit t und Grenzübergang $t \rightarrow 0$

$$0 \leq -2\Re(\langle x - y | z - y \rangle)$$

ergibt. ■

Bemerkung 7.13 *Die Existenz einer Lösung der Approximationsaufgabe 7.7 kann man direkt auf die Parallelogrammidentität stützen, ohne den Umweg über die Reflexivität zu nehmen. Man zeigt dazu, dass eine Minimalfolge, also eine Folge $(y^n)_{n \in \mathbb{N}}$ mit $\lim_n \|x - y^n\| = a := \text{dist}(x, K)$, eine Cauchyfolge ist. Man schließt dabei mit der Parallelogrammidentität und der Konvexität von K folgendermaßen.*

$$\begin{aligned} \|y^n - y^m\|^2 &= \|(x - y^m) - (x - y^n)\|^2 \\ &= -\|(x - y^m) + (x - y^n)\|^2 + 2\|x - y^m\|^2 + 2\|x - y^n\|^2 \\ &= -4\left\|\frac{1}{2}(y^m + y^n) - x\right\|^2 + 2\|x - y^m\|^2 + 2\|x - y^n\|^2 \\ &\leq -4a^2 + 2\|x - y^m\|^2 + 2\|x - y^n\|^2 \end{aligned}$$

Also ist $\lim_{m,n} \|y^n - y^m\| = 0$. □

Die Lösung y in Satz 7.12 wird **Bestapproximation** von x bezüglich K oder die **Projektion** von y auf K genannt. Wir schreiben dafür $y = P_K(x)$ und haben damit den **Projektionsoperator**

$$P_K \ni X \mapsto P_K(x) \in K$$

eingeführt. Es ist klar, dass $P_K = P_K \circ P_K$ gilt. P_K ist im Allgemeinen nicht linear, denn das Bild von P_K ist ja im Allgemeinen kein linearer Teilraum von X .

Die Ungleichung 7.8 wird **Variationsungleichung** genannt. Sie stellt die notwendige Bedingung dafür dar, dass y eine Lösung der Approximationsaufgabe ist. Man kann zeigen, dass sie auch hinreichend ist; wir übergehen dies. Die hier zu beobachtende Variationsungleichungen ist Spezialfall einer allgemeineren „Familie“ von Variationsungleichungen, die im Kontext von partiellen Differentialgleichungen etwa Kontakt- und Hindernisprobleme beschreiben.

Im Zweidimensionalen kann man die Variationsungleichung 7.8 geometrisch deuten. Dann bedeutet $\langle x - y | z - y \rangle \leq 0$ für alle $z \in K$, dass der Winkel zwischen $x - y$ und $z - y$ für alle $z \in K$ ein stumpfer ist. Wir können dies auch so ausdrücken:

$$y - K \subset H_{x-y,0}^+.$$

Sei $(X, \langle \cdot | \cdot \rangle)$ ein Prähilbertraum und U ein linearer Teilraum von X . Dann definieren wir mit

$$U^\perp := \{x \in X | \langle x | y \rangle = 0 \text{ für alle } y \in U\}$$

das orthogonale Komplement von U : in der Abbildung 7.2 ist die zweidimensionale Situation festgehalten. $U^\perp$ ist selbst wieder ein linearer Teilraum von X . Im folgenden Satz wird die Bedeutung dieser Begriffsbildung klar.

Satz 7.14 (Zerlegungssatz) *Sei $(X, \langle \cdot | \cdot \rangle)$ ein Hilbertraum und U ein linearer abgeschlossener Teilraum von X . Dann gibt es zu jedem $x \in X$ eindeutig bestimmte $u \in U, v \in U^\perp$ mit $x = u + v$, d. h.*

$$X = U \oplus U^\perp.$$

Beweis:

Wir wenden Satz 7.12 mit $K := U$ an. Damit ist schon klar, dass es zu jedem $x \in X$ ein eindeutig bestimmtes $u \in U$ gibt mit

$$\langle x - u | v \rangle = 0 \text{ für alle } v \in U;$$

beachte, dass wir nun in K einen linearen Teilraum haben. Da mit u auch $-u$ in U liegt, folgt $x - u \in U^\perp$. Also haben wir die Zerlegung $x = u + (x - u) \in U + U^\perp$ gezeigt. Die Eindeutigkeit der Zerlegung ist sofort klar, da stets $U \cap U^\perp = \{\theta\}$ gilt. ■

Ist $(X, \langle \cdot | \cdot \rangle)$ ein Hilbertraum und U ein linearer abgeschlossener Teilraum von X , dann kann man nach Satz 7.14 Abbildungen $p_U, p_{U^\perp}$ mit folgenden Eigenschaften definieren:

- a) $x = p_U(x) + p_{U^\perp}(x)$ für alle $x \in X$,
- b) $p_U \circ p_U = p_U, p_{U^\perp} \circ p_{U^\perp} = p_{U^\perp}$,

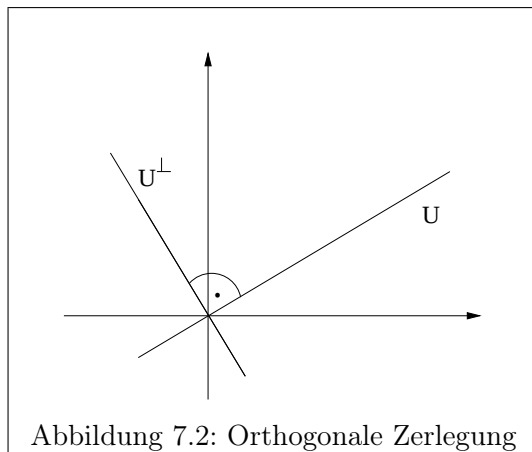


Abbildung 7.2: Orthogonale Zerlegung

$$c) \quad p_U \circ p_{U^\perp} = p_{U^\perp} \circ p_U = \theta.$$

Die Abbildungen $p_U, p_{U^\perp}$ werden **orthogonale Projektionen** genannt.

Der folgende Satz von Riesz² liefert nun eine wichtige Erkenntnis zum Dualraum von Hilberträumen.

Satz 7.15 (Darstellungssatz von Riesz) Sei $(X, \langle \cdot, \cdot \rangle)$ ein Hilbertraum mit induzierter Norm $\| \cdot \|$. Dann ist die Abbildung

$$R_X : X \ni x \longmapsto \langle \cdot | x \rangle_X \in X^* \quad (7.9)$$

bijektiv, isometrisch und konjugiert linear, d. h.

$$R_X(ax + by) = \bar{a}R_X(x) + \bar{b}R_X(y) \text{ für alle } x, y \in X, a, b \in \mathbb{K}. \quad (7.10)$$

Beweis:

Offenbar wird für jedes $x \in X$ durch $x \longmapsto \langle \cdot | x \rangle_X$ ein Funktional $\lambda_x \in X'$ erklärt. Wegen

$$|\langle \lambda_x, z \rangle| \leq \|z\| \|x\|, \quad |\langle \lambda_x, x \rangle| = \|x\|^2,$$

ist $\lambda_x \in X^*$. Also ist R_X wohldefiniert, offenbar konjugiert linear, isometrisch und stetig. Die Injektivität von R_X folgt aus der Isometrieeigenschaft. Es bleibt die Surjektivität von R_X zu zeigen.

Sei $\lambda \in X^*$; o. E. $\lambda \neq \theta$. Wir setzen $U := \ker(\lambda)$. Wegen $U \neq X$ ist $U^\perp \neq \{\theta\}$. Wir wählen $u \in U^\perp$ mit $\langle \lambda, u \rangle = 1$. Für $z \in X$ gilt

$$z = z - \langle \lambda, z \rangle u + \langle \lambda, z \rangle u, \quad z - \langle \lambda, z \rangle u \in U,$$

also

$$\langle z | u \rangle = \langle z - \langle \lambda, z \rangle u | u \rangle + \langle \langle \lambda, z \rangle u | u \rangle = \langle \lambda, z \rangle \langle u | u \rangle,$$

und somit

$$\lambda = \lambda_x = R_X(x) \text{ für } x = \frac{u}{\|u\|^2}.$$

■

Sei X ein Hilbertraum. Mit Hilfe des Riesz'schen Darstellungssatz kann man X^* selbst zu einem Hilbertraum machen, indem man das Skalarprodukt $\langle \cdot | \cdot \rangle_{X^*}$ in X^* wie folgt definiert:

$$\langle \lambda | \mu \rangle_{X^*} := \langle R_X^{-1}(\lambda) | R_X^{-1}(\mu) \rangle_X, \quad \lambda, \mu \in X^*; \quad (7.11)$$

dabei ist R_X die Riesz–Abbildung aus Satz 7.15. Dies können wir nun ausformulieren zu

Folgerung 7.16 Jeder Hilbertraum ist reflexiv.

Beweis:

Sei $(X, \langle \cdot | \cdot \rangle_X)$ ein Hilbertraum, sei R_X die zugehörige Riesz–Abbildung und sei J_X die kanonischen Einbettung von X in X^{**} . Sei $\mu \in X^{**}$. Mit der Riesz–Abbildung R_X definieren wir ein lineares Funktional ρ durch

$$\langle \rho, y \rangle := \overline{\langle \mu, R_X(y) \rangle}, \quad y \in X.$$

Offenbar ist ρ in X^* . Sei $x := R_X^{-1}(\rho)$. Sei nun $\lambda \in X^*$ beliebig. Da R_X surjektiv ist, gibt es $y \in X$ mit $\lambda = R_X(y)$. Dann haben wir

$$\langle \mu, \lambda \rangle = \langle \mu, R_X(y) \rangle = \overline{\langle \rho, y \rangle} = \overline{\langle y | R_X^{-1}(\rho) \rangle} = \langle R_X^{-1}(\rho) | y \rangle = \langle x | y \rangle = \langle R_X(y), x \rangle = \langle \lambda, x \rangle.$$

Dies zeigt $\mu = J_X(x)$. ■

²Der Satz ist benannt nach F. Riesz (1880–1956). Es gibt noch den Mathematiker M. Riesz.

Bemerkung 7.17 Die Reflexivität eines Hilbertraums haben wir in Folgerung 7.11 schon aus der gleichmäßigen Konvergenz gefolgert. $\square$

Für den Hilbertraum l_2 bedeutet der Rieszsche Darstellungssatz, dass zu jedem Funktional $\lambda \in l_2^*$ eine Folge $x = (x_n)_{n \in \mathbb{N}} \in l_2$ gibt mit

$$\langle \lambda, z \rangle = \sum_{n \in \mathbb{N}} z_n x_n, \quad z = (z_n)_{n \in \mathbb{N}} \in l_2.$$

Dies hatten wir bereits in Abschnitt 3.4 auf „direktem Weg“ gesehen.

Im Prähilbertraum $C([0, 1])$, versehen mit dem üblichen Skalarprodukt, gilt der Rieszsche Darstellungssatz übrigens nicht; man mache sich dies an dem linearen Funktional $f \mapsto f(0)$ klar.

Noch eine Bemerkung zur Bezeichnung von dualer Paarabbildung $\langle \cdot, \cdot \rangle : X^* \times X \rightarrow \mathbb{K}$ und Skalarprodukt $\langle \cdot, \cdot \rangle$ in einem Hilbertraum X . Dank des Rieszschen Darstellungssatzes ist die Ähnlichkeit der Bezeichnung gerechtfertigt.

7.4 Orthonormalbasen

Wir betrachten nun das Basisproblem in Hilberträumen unter Einbeziehung der Orthogonalität.

Definition 7.18 Sei $(X, \langle \cdot, \cdot \rangle_H)$ ein Prähilbertraum. Sei I eine Indexmenge und sei $\{x_i | i \in I\} \subset X \setminus \{\theta\}$.

(a) $\{x_i | i \in I\}$ heißt ein **Orthogonalsystem**, falls gilt:³

$$\langle x_i | x_j \rangle_X = \delta_{ij} \text{ für alle } i, j \in I, i \neq j.$$

(b) $\{x_i | i \in I\}$ heißt ein **Orthonormalsystem**, falls gilt:

$$\langle x_i | x_j \rangle_X = \delta_{ij} \text{ für alle } i, j \in I.$$

$\square$

Ist $\{x_i | i \in I\}$ ein **Orthogonalsystem**, dann ist $\{x_i | i \in I\}$ linear unabhängig. Ist das Orthogonalsystem abzählbar, also $I = \mathbb{N}$ oder I endlich, dann kann man mit dem **Gram-Schmidtschen Orthonormalisierungsverfahren** das Orthogonalsystem in algorithmisch durchsichtiger Weise zu einem Orthonormalsystem machen.

Satz 7.19 Sei $(X, \langle \cdot, \cdot \rangle_X)$ ein Prähilbertraum mit induzierter Norm $\|\cdot\|_X$ und sei $\{x_i | i \in I\} \subset X$ ein **Orthonormalsystem**. Sei J eine endliche Teilmenge von I und seien $\alpha_j \in \mathbb{K}, j \in J$. Sei $y \in X$. Dann gilt:

- (a) $\|y - \sum_{j \in J} \alpha_j x_j\|_X^2 = \|y\|^2 - \sum_{j \in J} |\langle y | x_j \rangle_X|^2 + \sum_{j \in J} |\alpha_j - \langle y | x_j \rangle_X|^2$
- (b) Genau dann ist $\|y - \sum_{j \in J} \alpha_j x_j\|_X^2$ in der Wahl von $(\alpha_j)_{j \in J}$ minimal, wenn $\alpha_j = \langle y | x_j \rangle_X$ für alle $j \in J$.
- (c) $\|y - \sum_{j \in J} \langle y | x_j \rangle_X x_j\|_X^2 = \|y\|_X^2 - \sum_{j \in J} |\langle y | x_j \rangle_X|^2$
- (d) $\|\sum_{j \in J} \alpha_j x_j\|_X^2 = \sum_{j \in J} |\alpha_j|^2$

³Mit δ_{ij} bezeichnen wir das Kroneckersymbol: $\delta_{ij} = 1$, falls $i = j$, $= 0$ sonst

Beweis:

b), c), d) folgen unmittelbar aus a). Also haben wir nur a) zu beweisen.

$$\begin{aligned}
\|y - \sum_{j \in J} \alpha_j x_j\|_X^2 &= \langle y - \sum_{j \in J} \alpha_j x_j | y - \sum_{j \in J} \alpha_j x_j \rangle_X \\
&= \|y\|_X^2 - \sum_{j \in J} \overline{\alpha_j} \langle y | x_j \rangle_X - \sum_{j \in J} \alpha_j \langle x_j | y \rangle_X + \sum_{j \in J} \alpha_j \overline{\alpha_j} \\
&= \|y\|_X^2 - \sum_{j \in J} |\langle y | x_j \rangle_X|^2 + \sum_{j \in J} |\alpha_j - \langle y | x_j \rangle_X|^2
\end{aligned}$$

Daraus ist a) ablesbar. ■

Folgerung 7.20 Sei $(X, \langle \cdot | \cdot \rangle_X)$ ein Prähilbertraum mit induzierter Norm $\|\cdot\|_H$ und sei $\{x_i | i \in I\} \subset X$ ein **Orthonormalsystem**. Sei $y \in X$. Dann gilt die **Besselsche Ungleichung**:

$$\sum_{i \in I} |\langle y | x_i \rangle_X|^2 \leq \|y\|_X^2 \quad (7.12)$$

Beweis:

„Grenzübergang“ von J nach I in c) von Satz 7.19. ■

Folgerung 7.21 Sei $(X, \langle \cdot | \cdot \rangle_X)$ ein Prähilbertraum mit induzierter Norm $\|\cdot\|_H$ und sei $\{x_i | i \in I\} \subset X$ ein **Orthonormalsystem**. Sei $y \in X$. Dann ist die Menge $\{i \in I | \langle y | x_i \rangle_X \neq 0\}$ abzählbar.

Beweis:

Aus der Bessel-Ungleichung folgt, dass die Menge $\{i \in I | |\langle y | x_i \rangle_X| > \frac{1}{n}\}$ endlich ist für alle $n \in \mathbb{N}$. ■

Definition 7.22 Sei $(X, \langle \cdot | \cdot \rangle_X)$ ein Prähilbertraum. Sei I eine Indexmenge und sei $\{x_i | i \in I\}$ ein **Orthonormalsystem**. $\{x_i | i \in I\}$ heißt eine **Orthonormalbasis**, falls $\{x_i | i \in I\}$ eine bezüglich der Inklusion $\subset$ maximales Element in der Menge der Orthonormalsysteme ist. □

Mit dem Lemma von Zorn kann man einsehen, dass jeder Hilbertraum eine Orthonormalbasis besitzt, die ein vorgegebenes Orthonormalsystem enthält. Beachte, dass eine Orthonormalbasis keine Vektorraumbasis ist.

Satz 7.23 Sei $(X, \langle \cdot | \cdot \rangle_X)$ ein Hilbertraum mit induzierter Norm $\|\cdot\|_X$ und sei $\{x_i | i \in I\} \subset X$ ein **Orthonormalsystem**. Dann sind äquivalent:

- (a) $\{x_i | i \in I\}$ ist eine Orthonormalbasis.
- (b) $y = \sum_{i \in I} \langle y | x_i \rangle_X x_i$ für alle $y \in X$ und die Reihe konvergiert unbedingt.
- (c) $\langle x | y \rangle_X = \sum_{i \in I} \langle x | x_i \rangle_X \overline{\langle y | x_i \rangle_X}$ für alle $x, y \in X$.
- (d) $\|y\|_X^2 = \sum_{i \in I} |\langle y | x_i \rangle_X|^2$
- (e) $\{x_i | i \in I\}$ ist eine dichte Teilmenge von X .

Beweis:

Die Implikationen (b) $\implies$ (e) und (c) $\implies$ (d) sind klar.

(a) $\implies$ (b) Zunächst ist zu klären, ob die Reihe in (b) konvergiert. Nach Satz 7.19 (d) ist $\sum_{j \in J} \alpha_j x_j \|_X^2 = \sum_{j \in J} |\alpha_j|^2$ für eine endliche Teilmenge J von I . Die Bessel-Ungleichung liefert

$$\sum_{j \in J} |\langle y | x_j \rangle_X|^2 \leq \|y\|_X^2$$

und da X vollständig ist, konvergiert die Reihe $\|\sum_{i \in I} \langle y|x_i \rangle x_i\|^2$. Setze $x := y - \sum_{i \in I} \langle y|x_i \rangle x_i$, dann gilt für alle $i \in I$ wegen der Stetigkeit des Skalarprodukts

$$\begin{aligned} \langle x|x_i \rangle_X &= \langle y - \sum_{j \in I} \langle y|x_j \rangle x_j | x_i \rangle_X = \langle y|x_i \rangle_X - \langle \sum_{j \in I} \langle y|x_j \rangle x_j | x_i \rangle_X \\ &= \langle y|x_i \rangle_X - \langle y|x_i \rangle_X \langle x_i|x_i \rangle_X = 0 . \end{aligned}$$

Da $\{x_i | i \in I\}$ maximal ist, muss $x = \theta$, also $y = \sum_{i \in I} \langle y|x_i \rangle x_i$ gelten.

(e) $\implies$ (a) Sei $x \in \{x_i | i \in I\}^\perp$, dann ist $x \in \overline{\{x_i | i \in I\}}^\perp = H^\perp = \{\theta\}$.

(b) $\implies$ (c) Seien $x, y \in H$. Dann gilt

$$\begin{aligned} \langle x|y \rangle_X &= \langle \sum_{i \in I} \langle x|x_i \rangle x_i | \sum_{i \in I} \langle y|x_i \rangle x_i \rangle_X \\ &= \sum_{i \in I} \langle \langle x|x_i \rangle x_i | \langle y|x_i \rangle x_i \rangle_X = \sum_{i \in I} \langle x|x_i \rangle_X \langle x_i|y \rangle_X . \end{aligned}$$

(d) $\implies$ (b) Sei $J \subset I$ endlich, dann gilt nach Satz 7.19 (c)

$$\|y - \sum_{j \in J} \langle y|x_j \rangle x_j\|_X^2 = \|y\|_X^2 - \sum_{j \in J} |\langle y|x_j \rangle_X|^2 ,$$

also $y = \sum_{j \in I} \langle y|x_j \rangle x_j$. ■

Die Skalare $\langle y|x_i \rangle_X x_i, i \in I$, in (b) des obigen Satzes 7.23 heißen in Anlehnung an Fourierreihen die **Fourierkoeffizienten** von y .

In Abschnitt 3.5 haben wir den Begriff der Schauder-Basis in einem Banachraum kennengelernt. Wir stellen nun fest, dass jede abzählbare Orthonormalbasis in einem Hilbertraum eine unbedingte Schauder-Basis darstellt. Klar, der zugrundeliegende Hilbertraum ist dann separabel. Ohne Beweis führen wir an, dass jeder separable Hilbertraum $(X, \langle \cdot | \cdot \rangle)$ Orthonormalbasen $(x^n)_{n \in \mathbb{N}}, (y^n)_{n \in \mathbb{N}}$, besitzt, so dass gilt:

$$\langle x^n | y^m \rangle_X = \delta_{mn} , m, n \in \mathbb{N} .$$

Dieses Paar von Basen heisst dann **biorthogonal**.

7.5 Selbstadjungierte Operatoren

Dank des Rieszschen Darstellungssatzes finden wir im Rahmen der Hilbertraumtheorie einen anderen Zugang zur dualen Abbildung.

Seien X, Y Hilberträume und sei $T : X \longrightarrow Y$ stetig. Dann ist für jedes $y \in Y$ das Funktional $x \longmapsto \langle Tx|y \rangle$ ein stetiges Funktional auf X . Nach Satz 7.15 gibt es daher genau ein Element $z_y \in X$ mit

$$\langle Tx|y \rangle = \langle x|z_y \rangle \text{ für alle } x \in X . \quad (7.13)$$

Definition 7.24 Seien X, Y Hilberträume und sei $T : X \longrightarrow Y$ stetig. Dann heisst die durch (7.13) definierte Zuordnung $Y \ni y \longmapsto z_y \in X$ die **adjungierte Abbildung** T^* von T ; also

$$\langle Tx|y \rangle = \langle x|T^*y \rangle \text{ für alle } x \in X . \quad (7.14)$$

□

Wir haben in Definition 7.24 im Kontext von Hilberträumen X eine adjungierte Abbildung T^* eingeführt, die in der Wirkung wie die duale Abbildung T^* bei normierten Räumen daher kommt. Der Zusammenhang zwischen beiden ist mit Hilfe der Riesz-Abbildungen ziemlich einfach herzustellen.

Seien X, Y Hilberträume und sei $T : X \rightarrow Y$ stetig. Seien $R_X : X \rightarrow X^*$, $R_Y : Y \rightarrow Y^*$ die Riesz-Abbildungen von X bzw. Y . Dann rechnet man einfach nach:

$$T^* = R_X^{-1} \circ T^* \circ R_Y \quad (7.15)$$

Die Unterscheidung zwischen T^* und T^* geben wir nun wieder auf und schreiben für T^* nun T^* . Die Unterscheidung ergibt sich etwa durch die Bezeichnungen „dual, adjungiert“.

Satz 7.25 *Seien X, Y Hilberträume. Dann gelten:*

- (a) *Die Zuordnung $\mathcal{B}(X, Y) \ni T \mapsto T^* \in \mathcal{B}(Y, X)$ ist eine Isometrie.*
- (b) *$(T^*)^* = T$ für alle $T \in \mathcal{B}(X, Y)$.*
- (c) *$\|T^*T\| = \|T\|^2$ für alle $T \in \mathcal{B}(X, Y)$.*

Beweis:

Zu (a). Folgt unmittelbar aus (7.15).

Zu (b). Seien $x \in X, y \in Y$. dann haben wir

$$\langle y | T^{**}x \rangle = \langle T^*y | x \rangle = \overline{\langle x | T^*y \rangle} = \overline{\langle Tx | y \rangle} = \langle y | Tx \rangle.$$

Zu (c).

$$\|T\|^2 = \sup_{\|x\| \leq 1} \|Tx\|^2 = \sup_{\|x\| \leq 1} \langle Tx | Tx \rangle = \sup_{\|x\| \leq 1} \langle x | T^*Tx \rangle \leq \|T^*T\| \leq \|T^*\| \|T\| = \|T^*\|^2.$$

■

Satz 7.26 *Seien X, Y Hilberträume. Für $T \in \mathcal{B}(X, Y)$ sind äquivalent:*

- (a) *$\text{ran}(T)$ ist abgeschlossen.*
- (b) *$\text{ran}(T) = \ker(T^*)^\perp$.*
- (c) *$\text{ran}(T^*)$ ist abgeschlossen.*
- (d) *$\text{ran}(T^*) = \ker(T)^\perp$.*

Beweis:

Vorbemerkung: Wir haben

$$\begin{aligned} \text{ran}(T)^\perp &= \{y \in Y | \langle Tx | y \rangle = 0 \text{ für alle } x \in X\} \\ &= \{y \in Y | \langle x | T^*y \rangle = 0 \text{ für alle } x \in X\} = \ker(T^*). \end{aligned}$$

(a) $\iff$ (b), (c) $\iff$ (d). Siehe Vorbemerkung.

(a) $\iff$ (c). Folgt aus Satz 6.32 und aus (7.15).

■

Folgerung 7.27 *Seien X ein Hilbertraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$. Es gelte mit einem $c > 0$:*

$$\Re \langle Tx | x \rangle \geq c \|x\|^2, \quad x \in X.$$

Dann gilt $\text{ran}(T^) = X$.*

Beweis:

Aus der Cauchy-Schwarzschen Ungleichung folgt

$$\|Tx\| \|x\| \geq \Re \langle Tx | x \rangle \geq c \|x\|^2, \quad x \in X.$$

Also gilt $\|Tx\| \geq c \|x\|, x \in X$, und damit hat T eine stetige Inverse. Mit Folgerung 6.33 folgt die Behauptung.

■

Definition 7.28 Seien X, Y Hilberträume, sei $T : X \rightarrow Y$ linear und stetig. Dann heisst T **selbstadjungiert**, falls $T = T^*$ gilt. $\square$

Lemma 7.29 Seien X, Y Hilberträume und sei $T \in \mathcal{B}(X, Y)$ selbstadjungiert. Dann gelten:

- (a) $\langle Tx|x \rangle \in \mathbb{R}$ für alle $x \in X$.
- (b) $\|T\| = \sup\{|\langle Tx|x \rangle| : \|x\| = 1\}$.

Beweis:

Zu (a). Wie in der Linearen Algebra: Da T selbstadjungiert ist, gilt für $x \in X$:

$$\overline{\langle Tx|x \rangle} = \langle x|Tx \rangle = \langle x|T^*x \rangle = \langle Tx|x \rangle.$$

Zu (b). Offenbar gilt

$$c := \sup_{\|x\|=1} |\langle Tx|x \rangle| \leq \sup_{\|x\|=1} \|Tx\| \|x\| = \|T\|.$$

Um $\|T\| \leq c$ zu zeigen, beachten wir, dass wegen $T^* = T$ für alle $x, z \in X$

$$\langle T(x \pm z), x \pm z \rangle = \langle Tx|x \rangle \pm 2\Re\langle Tx|z \rangle + \langle Tz|z \rangle$$

gilt. Hieraus folgt

$$\Re\langle Tx|z \rangle = \frac{1}{4}(\langle T(x+z)|x+z \rangle - \langle T(x-z)|x-z \rangle).$$

Für festes x, z mit $\|x\| \leq 1, \|z\| \leq 1$ gehen wir nun so vor: Wir wählen $a \in \mathbb{K}$ mit $|a| = 1$ und $|\langle Tx|z \rangle| = \Re\langle Tx|az \rangle$ und erhalten

$$|\langle Tx|z \rangle| = \Re\langle Tx|az \rangle \leq \frac{c}{4}(\|x+az\|^2 + \|x-az\|^2) = \frac{c}{4}(2\|x\|^2 + 2\|z\|^2) \leq c.$$

Hieraus folgt die Behauptung nun so:

$$\|T\|^2 = \sup_{\|x\|=1} \langle Tx|Tx \rangle \leq \sup_{\|x\|=1, \|z\|=1} \langle Tx|z \rangle \|Tx\| \leq c \sup_{\|x\|=1} \|Tx\| = c\|T\|.$$

■

7.6 Das Spektrum beschränkter Operatoren

Definition 7.30 Sei X ein Banachraum über dem Skalkörper $\mathbb{K}$. Sei $T \in \mathcal{B}(X)$. Die Teilmenge

$$\rho(T) := \{\lambda \in \mathbb{K} \mid \lambda - T \text{ ist bijektiv}\}$$

heisst **Resolventenmenge** von T . Die Abbildung

$$R : \rho(T) \ni \lambda \mapsto R_\lambda := (\lambda - T)^{-1} \in \mathcal{B}(X) \quad (7.16)$$

heisst die **Resolvente** von T . $\square$

Wir benötigen die so genannte **geometrische Reihe** im Banachraum. Als Vorbereitung betrachten wir „kleine Störungen“ der Identität, d.h. wir wollen $T := \text{id} - H$ betrachten für ein $H \in \mathcal{B}(X)$, X Banachraum. Es ist zu erwarten, dass T invertierbar ist, wenn H nur klein genug ist. Hier soll Kleinheit mit der Norm erklärt werden. Wir können diese Frage klären durch einen Vergleich mit der geometrischen Reihe. Wir betrachten formal die Identität

$$(\text{id} - H)(\text{id} + H + H^2 + \cdots + H^k) = \text{id} - H^{k+1}.$$

Wenn wir diese Identität mit $(\text{id} - H)^{-1}$ multiplizieren und $(\text{id} - H)^{-1}$ auf beiden Seiten abziehen, erhalten wir

$$(\text{id} + H + H^2 + \cdots + H^k) - (\text{id} - H)^{-1} = -(\text{id} - H)^{-1} H^{k+1}. \quad (7.17)$$

Also sollte $\text{id} - H$ invertierbar sein, falls $(H^k)_{k \in \mathbb{N}}$ gegen Null konvergiert der Norm nach.

Satz 7.31 (Neumann-Reihe) Sei X ein Banachraum und sei $H \in \mathcal{B}(X)$ und es gelte $\|H\|_{X \rightarrow X} < 1$. Dann ist $\text{id} - H$ invertierbar und es gelten

$$(a) \quad (\text{id} - H)^{-1} = \sum_{k=0}^{\infty} H^k.$$

$$(b) \quad \|(\text{id} + H + H^2 + \cdots + H^k) - (\text{id} - H)^{-1}\| \leq \frac{\|H\|^{k+1}}{1 - \|H\|}.$$

Beweis:

$\sum_{n=1}^{\infty} H^k$ ist eine in $\mathcal{B}(X)$ konvergente Reihe ist, denn die geometrische Reihe $\sum_{i=0}^{\infty} \|H\|_{X \rightarrow X}^k$ ist eine Majorante. ■

Satz 7.32 Sei X ein Banachraum und sei $T \in \mathcal{B}(X)$. Dann gelten:

(a) $\rho(T)$ ist offen.

(b) Ist $\lambda_0 \in \rho(T)$, dann haben wir

$$R_{\lambda} = (\lambda - T)^{-1} = \sum_{k=0}^{\infty} (\lambda - \lambda_0)^k ((\lambda_0 - T)^{-1})^{k+1} = \sum_{k=0}^{\infty} (\lambda - \lambda_0)^k R_{\lambda_0}^{k+1} \quad (7.18)$$

falls $|\lambda - \lambda_0| < \|R_{\lambda_0}\|^{-1}$ gilt.

Beweis:

Es gilt

$$\lambda - T = (\lambda_0 - T) \circ (1 - (\lambda_0 - T)^{-1}(\lambda_0 - \lambda)), \quad \|(\lambda_0 - T)^{-1}(\lambda_0 - \lambda)\| = \|R_{\lambda_0}\| |\lambda_0 - \lambda|.$$

Eine Anwendung der Neumannschen Reihe (siehe 7.31) ergibt die Behauptung. ■

Definition 7.33 Sei X ein Banachraum und sei $T \in \mathcal{B}(X)$. Dann heisst $\sigma := \mathbb{K} \setminus \rho(T)$ das **Spektrum** von T und jedes $\lambda \in \sigma(T)$ heisst **Spektralwert** von T . □

Definition 7.34 Sei X ein Banachraum über dem Skalkörper $\mathbb{K}$ und sei $T \in \mathcal{B}(X)$. Sei $\lambda \in \sigma(T)$.

(a) λ heisst **Eigenwert** von T , falls $\lambda - T$ nicht injektiv ist; ein $x \in \ker(\lambda - T) \setminus \{\theta\}$ heisst **Eigenvektor** von T . Der Raum $\ker(\lambda - T)$ heisst **Eigenraum** und $\dim(\ker(\lambda - T))$ (geometrische) **Vielfachheit** von λ .

(b) Wir nennen $\sigma_p(T) := \{\lambda \in \mathbb{K} \mid \lambda \text{ Eigenwert von } T\}$ das **Punktspektrum** von T .

(c) Die Menge

$$\sigma_c(T) := \{\lambda \in \mathbb{K} \mid \lambda - T \text{ injektiv, nicht surjektiv, } \overline{\text{ran}(\lambda - T)} = X\}$$

heisst das **kontinuierliche Spektrum** von T .

(d) Die Menge

$$\sigma_r(T) := \{\lambda \in \mathbb{K} \mid \lambda - T \text{ injektiv, nicht surjektiv, } \overline{\text{ran}(\lambda - T)} \neq X\}$$

heisst das **Restspektrum** von T .

□

Das Spektrum zerfällt also in disjunkte Teilmengen:

$$\sigma(T) = \sigma_p(T) \cup \sigma_c(T) \cup \sigma_r(T).$$

Man sieht sofort, dass im Falle $\lambda \in \sigma_c(T)$ die Abbildung $\lambda - T : X \rightarrow \text{ran}(\lambda - T)$ eine nicht stetige Inverse besitzt.

Beispiel 7.35 Sei

$$X := l_2, T(x)_k := \begin{cases} 0 & , \text{ falls } k = 1 \\ x_{k-1} & , \text{ falls } k > 1 \end{cases}, \text{ wenn } x = (x_k)_{k \in \mathbb{N}}.$$

Nun ist $0 \in \sigma_r(T)$, denn T ist zwar injektiv, aber $\text{ran}(T) = \{(x_k)_{k \in \mathbb{N}} | x_1 = 0\}$ ist ein abgeschlossener Teilraum von l_2 , der verschieden von X ist. Beachte, dass diese Konstruktion mit endlichen Tupeln so nicht möglich ist. □

Beispiel 7.36 Sei $X := l_2, T(x)_k := \frac{1}{k}x_k, k \in \mathbb{N}$, falls $x = (x_k)_{k \in \mathbb{N}}$. Nun ist $0 \in \sigma_c(T)$, da T zwar injektiv, T wegen $(\frac{1}{k})_{k \in \mathbb{N}} \notin \text{ran}(T)$ aber nicht surjektiv und $c_0(\mathbb{K}) \subset \text{ran}(T)$ ein dichter Teilraum von X ist. Beachte, dass diese Konstruktion mit endlichen Tupeln so nicht möglich ist. □

Satz 7.37 Sei X ein Banachraum über dem Skalkörper $\mathbb{K}$ und sei $T \in \mathcal{B}(X)$. Dann ist $\sigma(T)$ kompakt und es gilt $\sigma(T) \subset \overline{B}_{\|T\|}$.

Zusatz: Ist $\mathbb{K} = \mathbb{C}$, dann folgt $\sigma(T) \neq \emptyset$.

Beweis:

Ist $T = \theta$, dann ist nicht zu beweisen. Sei nun $T \neq \theta$.

Ist $|\lambda| > \|T\|$, dann liest man aus

$$\lambda - T = \lambda(1 - \frac{1}{\lambda}T)$$

ab, dass $\lambda \in \rho(T)$ gilt. Also ist $\sigma(T) \subset \overline{B}_{\|T\|}$ schon klar. Da $\overline{B}_{\|T\|}$ kompakt und $\sigma(T)$ abgeschlossen ist, ist $\sigma(T)$ kompakt.

Wir kommen zum Zusatz.

$\mathcal{B}(X)$ ist ein Banachraum. Sei $\mu \in \mathcal{B}(X)^*$. Damit definieren wir

$$f_\mu := f : \rho(T) \ni \lambda \mapsto \mu(R_\lambda) \in \mathbb{C}.$$

Seien $\lambda_0 \in \rho(T)$ und sei $\lambda \in \mathbb{C}$ mit $|\lambda - \lambda_0| < \|R_{\lambda_0}\|^{-1}$. Damit haben wir

$$f(\lambda) = \mu(R_\lambda) = \sum_{k=0}^{\infty} (\lambda - \lambda_0)^k \mu(R_{\lambda_0}^{k+1})$$

nach (b) in Satz 7.32. Also hat f eine konvergente Potenzreihenentwicklung (siehe unten) in einer Umgebung von λ_0 . Dies bedeutet, dass f holomorph in $\rho(T)$ ist.

Annahme: $\sigma(T) = \emptyset$, d. h. $\rho(T) = \mathbb{C}$.

f ist als stetige Funktion auf $\overline{B}_{2\|T\|}$ beschränkt. Für $|\lambda| > 2\|T\|$ gilt mit der Neumannschen Reihe

$$|f(\lambda)| = |\mu(R_\lambda)| = \left| \frac{1}{\lambda} \sum_{k=0}^{\infty} \mu(T^k) \lambda^{-k} \right| \leq \frac{1}{|\lambda|} \|\mu\|_* \sum_{k=0}^{\infty} \|T\|^k \|\lambda\|^k \leq 2\|\mu\|_* \frac{1}{|\lambda|}.$$

Also ist die ganze Funktion f beschränkt auf $\mathbb{C}$. Nach dem Satz von Liouville ist f konstant. Aus der obigen Abschätzung folgt $f = f_\mu = \theta$. Also ist $\mu(R_\lambda) = 0$ für alle $\mu \in \mathcal{B}(X)$. Nach dem Satz von Hahn–Banach folgt $R_\lambda = \theta$, was ein Widerspruch ist, da R_λ ein bijektiver Operator ist. ■

Bemerkung 7.38 Zum Ergebnis von Satz 7.37 gibt es die Verschärfung $\sigma(T) \subset \overline{B}_r$ mit $r := r(T) := \inf_n \sqrt[n]{\|T^n\|}$. Dabei heisst $r(T)$ **Spektralradius** von T . Beachte, dass wegen $\|T^n\| \leq \|T\|^n$ wirklich eine Verschärfung vorliegt. Mit Mitteln der Funktionentheorie beweist man im Fall $\mathbb{K} = \mathbb{C}$ auch $r(T) = \max_{\lambda \in \sigma(T)} |\lambda|$. $\square$

7.7 Anhang: Abstandserhaltende Abbildungen

Welche Abbildungen passen zur geometrischen Sicht des euklidischen Raums $\mathbb{R}^n$? Es sind dies die Abbildungen, die den euklidischen Abstand erhalten. Zunächst zu den **Symmetriegruppen**. Dazu folgende **Bezeichnungen**:

$$\begin{aligned}\mathcal{S}(n) &:= \{A \in \mathbb{R}^{n,n} \mid \det(A) \neq 0\}, \\ \mathcal{O}(n) &:= \{A \in \mathbb{R}^{n,n} \mid A^t A = A A^t = E\}, \\ \mathcal{SO}(n) &:= \{A \in \mathcal{O}(n) \mid \det(A) = 1\}.\end{aligned}$$

Mit dem Determinantensatz $\det(AB) = \det(A)\det(B)$ folgt, dass $\mathcal{S}(n)$, $\mathcal{O}(n)$ und $\mathcal{SO}(n)$ Gruppen sind bezüglich der Operation „Matrixmultiplikation“. $\mathcal{S}(n)$ heisst **Symmetriegruppe** des $\mathbb{R}^n$, $\mathcal{SO}(n)$ heisst **spezielle orthogonale Gruppe**; sie steht für die Drehungen in der euklidischen Ebene. Die **euklidische Gruppe** $\mathcal{E}(n)$ wird definiert als diejenige Untergruppe von $\mathcal{S}(n)$, welche von $\mathcal{SO}(n)$ und der Gruppe der **Translationen**

$$\mathcal{T} := \{T_b \mid b \in \mathbb{R}^n\} \quad (T_b(x) := b + x, b \in \mathbb{R}^n, x \in \mathbb{R}^n)$$

erzeugt wird.

Klar, eine Transformation $A \in \mathcal{O}(n)$ erhält wegen

$$\langle Ax \mid Ay \rangle = \langle A^t Ax \mid y \rangle = \langle x \mid y \rangle, \quad x, y \in \mathbb{R}^n,$$

Winkel und euklidischen Abstand. Der folgende Satz belegt, dass dies schon fast der allgemeine Fall winkel- und längenerhaltender Abbildungen ist.

Satz 7.39 Ist $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ eine Abbildung, die den euklidischen Abstand invariant lässt, d. h. für die

$$|f(x) - f(y)| = |x - y| \quad \text{für alle } x, y \in \mathbb{R}^n,$$

gilt, dann ist f von der Form $f = T_b \circ A$ mit $A \in \mathcal{O}(n)$ und $b \in \mathbb{R}^n$; b und A sind dabei eindeutig bestimmt.

Beweis:

Sei zunächst $f(\theta) = \theta$. Wir zeigen, dass f linear ist. Nach Voraussetzung gilt

$$|f(x)| = |f(x) - \theta| = |x - \theta| = |x|$$

für alle $x \in \mathbb{R}^n$. Nun folgt für alle $x, y \in \mathbb{R}^n$:

$$2\langle f(x), f(y) \rangle = |f(x)|^2 + |f(y)|^2 - |f(x) - f(y)|^2 = |x|^2 + |y|^2 - |x - y|^2 = 2\langle x \mid y \rangle.$$

Also lässt f auch das euklidische Skalarprodukt invariant. Mit der Standardbasis $e^1, \dots, e^n \in \mathbb{R}^n$ folgt damit nun $\langle f(e^i) \mid f(e^j) \rangle = \delta_{ij}$, $1 \leq j \leq n$, und $f(e^1), \dots, f(e^n)$ ist eine Basis von $\mathbb{R}^n$. Sei $x \in \mathbb{R}^n$, $x = \sum_{i=1}^n x_i e^i$. Damit haben wir

$$x^i = \langle x \mid e^i \rangle = \langle f(x) \mid f(e^i) \rangle, \quad 1 \leq i \leq n, \quad f(x) = \sum_{i=1}^n x_i f(e^i).$$

Daraus lesen wir ab, dass f eine surjektive Abbildung von $\mathbb{R}^n$ nach $\mathbb{R}^n$ ist, die das Skalarprodukt invariant lässt. Nun können wir ohne Einschränkungen $\mathbb{R}^n$ mit $\mathbb{R}^{n,1}$ identifizieren und annehmen (nach Wahl einer Basis), dass mit $A \in \mathbb{R}^{n,n}$ gilt:

$$f(x) = Ax, \quad x \in \mathbb{R}^{n,1}.$$

Aus $\langle Ax|Ay \rangle = \langle x|y \rangle$, $x, y \in \mathbb{R}^{n,1}$, folgt

$$\langle x|y \rangle = \langle A^t Ax|y \rangle = \langle x|A^t Ay \rangle, \quad x, y \in \mathbb{R}^{n,1}.$$

Daraus folgt $A^t A = E$ und schließlich auch $AA^t = E$.

Sei nun $b := f(\theta)$ beliebig. Setze $g(x) := f(x) - b$, $x \in \mathbb{R}^n$. Man sieht, dass auch g den Abstand invariant lässt. Außerdem gilt $g(\theta) = 0$. Aus obigem Spezialfall folgt die Existenz von $A \in \mathcal{O}(n)$ mit

$$f(x) = b + Ax, \quad x \in \mathbb{R}^{n,1}, \quad \text{also } f = T_b \circ A.$$

Für jede andere solche Darstellung

$$f = T_c \circ B$$

mit $c \in \mathbb{R}^{n,1}$ und $B \in \mathcal{O}(n)$ folgt zunächst $b = f(\theta) = c$ und dann $Ax = Bx$ für alle $x \in \mathbb{R}^{n,1}$, also $A = B$. ■

Beachte: Jede abstandserhaltende Abbildung ist stetig.

Folgerung 7.40 *Ist $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ eine Abbildung, die den euklidischen Abstand invariant lässt, dann lässt sie auch die Winkel invariant, d. h. es gilt:*

$$\langle f(x), f(y) \rangle = \langle x, y \rangle \quad \text{für alle } x, y \in \mathbb{R}^n.$$

Beweis:

Folgt aus der obigen Beweisführung. ■

7.8 Anhang: Das Lemma von Lax-Milgram

Nun beweisen wir eine Verallgemeinerung des Rieszschen Darstellungssatzes⁴ auf unsymmetrische Sesquilinearformen. Den Ansatz für diese Verallgemeinerung kann man an der Beobachtung ablesen, dass in jedem Hilbertraum $(X, \cdot, \cdot)_X$ die Gleichung

$$\langle v, u \rangle_X = \lambda(v) \tag{7.19}$$

für jedes $\lambda \in X^*$ eine eindeutige Lösung u hat, nämlich $u = R_X^{-1}(\lambda)$.

Definition 7.41 *Sei $(X, \cdot, \cdot)_X$ ein Hilbertraum über dem Skalarkörper $\mathbb{K}$ mit induzierter Norm $\|\cdot\|_X$. Eine Sesquilinearform $b : X \times X \rightarrow \mathbb{K}$ heißt **beschränkt**, wenn es eine Konstante c_b gibt mit*

$$|b(x, y)| \leq c_b \|x\|_X \|y\|_X \quad \text{für alle } x, y \in X.$$

*Sie heißt **koerziv**, wenn es eine Konstante $c_e > 0$ gibt mit*

$$|b(x, x)| \geq c_e \|x\|_X^2 \quad \text{für alle } x \in X.$$

□

⁴Die Bezeichnung „Lemma“ für das Resultat ist historisch begründet; hier wird es zu einem Satz. Für den Originalartikel siehe [24].

Satz 7.42 (Lax-Milgram) Sei $(X, \langle \cdot | \cdot \rangle_X)$ ein Hilbertraum über dem Skalkörper $\mathbb{K}$ mit induzierter Norm $\| \cdot \|_X$ und sei $b : X \times X \rightarrow \mathbb{K}$ eine beschränkte, koerzive Sesquilinearform. Dann gibt es einen bijektiven Operator $A \in \mathcal{B}(X)$ mit $A^{-1} \in \mathcal{B}(X)$, so dass gilt:

$$b(x, Aw) = \langle x | w \rangle_X \text{ für alle } x, w \in X. \quad (7.20)$$

Ferner gilt $\|A\| \leq c_e^{-1}$, $\|A^{-1}\| \leq c_b$.

Beweis:

Das Funktional $b(\cdot, y)$ ist linear und stetig mit $\|b(\cdot, y)\|_{X^*} \leq c_b \|y\|$. Nach dem Darstellungssatz von Riesz existiert daher genau ein Ty mit

$$b(x, y) = \langle x, Ty \rangle_X \text{ für alle } x \in X. \quad (7.21)$$

Da beide Formen in (7.21) sesquilinear sind, ist $T : X \rightarrow X$ linear und wegen

$$\|Ty\|^2 = \langle Ty, Ty \rangle_X = b(Ty, y) \leq c_b \|Ty\|_X \|y\|_X \quad (7.22)$$

auch stetig. Aus

$$c_e \|y\|_X^2 \leq |b(y, y)| = \langle y | Ty \rangle_X \leq \|Ty\|_X \|y\|_X, \quad (7.23)$$

folgt $\ker(T) = \{\theta\}$. Weiter folgt aus dieser Abschätzung, dass der Bildraum $\text{ran}(T)$ abgeschlossen ist. Denn wenn für eine Folge $(y_n)_{n \in \mathbb{N}}$ gilt $w = \lim_n Ty_n$, so liefert (7.23)

$$\|y_n - y_m\| \leq c_e^{-1} \|Ty_n - Ty_m\|.$$

Damit ist $(y_n)_{n \in \mathbb{N}}$ eine Cauchyfolge in X und es gibt $y \in X$ mit $y = \lim_n y_n$. Wegen der Stetigkeit von T gilt $Ty = \lim_n Ty_n = w$.

Sei $z \in \text{ran}(T)^\perp$. Aus $\langle z | Ty \rangle_X = 0$ für alle $y \in X$ folgt $0 = \langle z | Tz \rangle_X = b(z, z) \geq c_e \|z\|_X^2$ und damit $z = \theta$. Da $\text{ran}(T)$ abgeschlossen ist, erhalten wir $\text{ran}(T) = X$. Damit ist T bijektiv und $A := T^{-1}$ leistet das Gewünschte, allerdings ist die Stetigkeit von A noch offen. Wir verzichten hier auf den Beweis, denn der Satz 4.11, den wir später beweisen werden, liefert die Stetigkeit. Die Normabschätzungen für A folgen aus (7.22) und (7.23), indem man dort x durch $x := T^{-1}x$ ersetzt. ■

Dieser Satz wird häufig angewendet auf Lösungen von Gleichungen, die mittels Bilinearformen gegeben sind. Solche Gleichungen sind bei den elliptischen partiellen Differentialgleichungen vorzufinden. Hier ist eine Einordnung.

Sei $(V, \langle \cdot | \cdot \rangle_V)$ ein reeller Hilbertraum mit induzierter Norm $\| \cdot \|_V$ und sei $a : V \times V \rightarrow \mathbb{R}$ eine Bilinearform, die beschränkt und koerziv ist:

$$|a(u, v)| \leq c_b \|u\|_V \|v\|_V, \quad a(u, u) \geq c_e \|u\|_V^2 \text{ für alle } u, v \in V. \quad (7.24)$$

Wir betrachten damit folgende **variationell** formulierte Gleichung:

Gegeben $\lambda \in V^*$

Gesucht $u \in V$ mit

$$a(v, u) = \langle \lambda, v \rangle \text{ für alle } v \in V. \quad (7.25)$$

Betrachte dazu die Aufgabe

Gegeben $\lambda \in V^*$

Gesucht $u \in V$ mit

$$\frac{1}{2}a(u, u) - \langle \lambda, u \rangle = \min_{v \in V} \left(\frac{1}{2}a(v, v) - \langle \lambda, v \rangle \right) \quad (7.26)$$

Wir stellen den Zusammenhang zwischen den obigen Aufgaben her.

Lemma 7.43 Sei $(V, \langle \cdot | \cdot \rangle_V)$ ein reeller Hilbertraum mit induzierter Norm $\| \cdot \|_V$ und sei $a : V \times V \rightarrow \mathbb{R}$ eine Bilinearform, die beschränkt, koerziv ist. Ferner sei sie symmetrisch, d. h. $a(w, v) = a(v, w)$ für alle $v, w \in V$. Dann sind für $u \in V$ äquivalent:

(a) u ist eine Lösung von (7.26).

(b) $a(u, v) = \langle \lambda, v \rangle$ für alle $v \in V$.

Beweis:

a) $\implies$ b). Sei $v \in V$. Dann gilt auf Grund der Optimalität von u für alle $t > 0$

$$\frac{1}{2}a(u + tv, u + tv) - \langle \lambda, u + tv \rangle \geq \frac{1}{2}a(u, u) - \langle \lambda, u \rangle, \text{ also } ta(u, v) + \frac{1}{2}t^2a(v, v) - t\langle \lambda, v \rangle \geq 0.$$

Division mit t und Grenzübergang $t \rightarrow 0$ ergibt $a(u, v) - \langle \lambda, v \rangle \geq 0$. Da dies auch für $-v$ gilt folgt $a(u, v) - \langle \lambda, v \rangle = 0$.

b) $\implies$ a). Sei $v \in V$. Eine einfache Rechnung zeigt

$$\frac{1}{2}a(v, v) - \langle \lambda, v \rangle = \frac{1}{2}a(v - u, v - u) + a(u, v - u) - \langle \lambda, v - u \rangle + \frac{1}{2}a(u, u) - \langle \lambda, u \rangle,$$

woraus mit b) folgt

$$\frac{1}{2}a(v, v) - \langle \lambda, v \rangle = \frac{1}{2}a(v - u, v - u) + \frac{1}{2}a(u, u) - \langle \lambda, u \rangle.$$

Da $a(v - u, v - u) \geq 0$ folgt, dass u (7.26) löst. ■

Folgerung 7.44 Sei $(V, \langle \cdot | \cdot \rangle_V)$ ein reeller Hilbertraum mit induzierter Norm $\| \cdot \|_V$ und sei $a : V \times V \rightarrow \mathbb{R}$ eine Bilinearform, die beschränkt und koerziv ist. Dann besitzt die Aufgabe 7.25 genau eine Lösung $u = u_\lambda$. Damit gilt:

$$\|u_\lambda - u_\mu\|_V \leq c_e^{-1} \|\lambda - \mu\|_{V^*}, \quad \lambda, \mu \in V^*. \quad (7.27)$$

Zusatz: Ist die Bilinearform auch symmetrisch, d. h. $a(w, v) = a(v, w)$ für alle $v, w \in V$, dann ist diese Lösung u charakterisiert als Lösung der Aufgabe (7.26).

Beweis:

Zunächst zur Eindeutigkeit. Sind $u, \hat{u}$ zwei Lösungen, dann gilt offenbar $a(u - \hat{u}, u - \hat{u}) = 0$. Die Koerzivität impliziert $u = \hat{u}$.

Nun zur Existenz. Die Gleichung lautet mit der Riesz-Abbildung R_V :

$$a(v, u) = \langle v | R_X^{-1} \lambda \rangle_V, \quad u, v \in V.$$

Mit der Abbildung A aus Satz 7.42 erhalten wir

$$Au = R^{-1} \lambda, \quad u = A^{-1} R^{-1} \lambda.$$

Daraus folgt

$$\|u_\lambda - u_\mu\|_V = \|A^{-1} R^{-1} \lambda - A^{-1} R^{-1} \mu\|_V \leq c_e^{-1} \|R^{-1}(\lambda - \mu)\|_V = c_e^{-1} \|\lambda - \mu\|_{V^*}$$

Der Zusatz folgt mit der Symmetrie aus Lemma 7.43. ■

Mit dem Lemma von Lax-Milgram kann man die Historie des **Dirichlet-Problems** gut nachvollziehen und damit auch ein gutes Stück „Entwicklung der Analysis“ einsehen. Dieses Problem hat folgende (zweidimensionale) Beschreibung:

Sei Ω eine beschränkte offene Menge in $\mathbb{R}^2$, sei $f : \Omega \rightarrow \mathbb{R}$ und $c \geq 0$. Gesucht ist eine Lösung u des **Randwertproblems**

$$\begin{cases} -\Delta u + cu = f & \text{in } \Omega \\ u = 0 & \text{in } \partial\Omega \end{cases} \quad (7.28)$$

Eine **klassische Lösung** der Aufgabe wird eine Funktion $u \in C^2(\Omega) \cap C(\bar{\Omega})$ genannt, die zusätzlich $u|_{\partial\Omega} = \theta$ erfüllt. Im allgemeinen gibt es keine solche klassische Lösung! Dies kann etwa folgende Gründe haben:

- Die rechte Seite ist keine stetige Funktion.
- Das Gebiet Ω ist nicht „schön“ genug, um eine solche Lösung zuzulassen (einspringende Ecken, ...).

Mit der folgenden Überlegung wollen wir nun mindestens eine Differenzierbarkeitsordnung für die Lösung loswerden. Zunächst nehmen wir aber eine genügend glatte Lösung der Aufgabe an und gehen dann folgendermaßen vor. Wir multiplizieren in der ersten Gleichung in (7.28) mit einer so genannten **Testfunktion** v , die ebenfalls hinreichend glatt angenommen werde, und integrieren dann partiell. Dann erhalten wir in formaler Rechnung

$$\int_{\mathbb{R}^2} (\nabla u(\xi) \nabla v(\xi) + cu(\xi)v(\xi)) d\xi = \int_{\mathbb{R}^2} f(\xi)v(\xi) d\xi. \quad (7.29)$$

Nun setzen wir

$$a(v, u) := \int_{\mathbb{R}^2} (\nabla u(\xi) \nabla v(\xi) + cu(\xi)v(\xi)) d\xi, \quad \lambda(v) := \int_{\mathbb{R}^2} f(\xi)v(\xi) d\xi,$$

Damit haben wir die Aufgabe (7.28) umgeschrieben in eine Aufgabe der Form (7.25), allerdings fehlt noch der Hilbertraum V , in dem alles „zueinander passt“. Dieser Raum V ergibt sich als homogener Sobolevraum $H_0^1(\Omega)$ als Vervollständigung von $C_0^\infty(\Omega)$ unter der Norm

$$\|v\| := \left(\int_{\mathbb{R}^2} |\nabla v(\xi)|^2 + |v(\xi)|^2 d\xi \right)^{\frac{1}{2}},$$

die ja als von einem Skalarprodukt induziert angesehen werden kann. Dabei ist

$$C_0^\infty(\Omega) := \{w \in C^\infty(\Omega) | w = \theta \text{ außerhalb einer kompakten Menge } K \subset \Omega\},$$

$$C^\infty(\Omega) := \{w : \Omega \rightarrow \mathbb{R} | w \text{ unendlich oft differenzierbar}\}.$$

Auf die rigorose Einordnung, um das Lemma von Lax-Milgram anwenden zu können, müssen wir hier verzichten, da für die sinnvolle Betrachtung der Sobolev-Räume das Lebesgue-Integral vonnöten ist. Es ist möglich! Siehe hierzu etwa [33, 38]. Eine Lösung auf diesem Weg nennt man eine **schwache Lösung**.

Riemanns Ansatz, das Dirichlet-Problem zu lösen, bestand in der Nutzung des Zusammenhangs mit der Aufgabe (7.26), die Bilinearform a ist ja symmetrisch. Dieser Ansatz wird **Dirichlet-Prinzip** genannt. Im Raum klassischer Lösungen besitzt diese Aufgabe im Allgemeinen keine Lösung! Dass eine Optimierungsaufgabe, die nach unten beschränkt ist, also ein Infimum besitzt, kein Minimum haben muss, wurde hierbei nicht in Betracht gezogen. Erst Weierstraß klärte hierzu mit Kompaktheitsbetrachtungen die Sachlage. Man kommt auch auf diesem Weg zum Sobolev-Raum $H_0^1(\Omega)$ als sinnvollerweise zu nutzender Hilbertraum.

Beide Ansätze, das Dirichletproblem zu lösen, führen auch zu guten Verfahren, die Lösung gut zu approximieren. Wir machen dies an den Aufgaben (7.25), (7.26) deutlich.

Die Approximation, die für numerisches Rechnen (mit numerischen auf einem Computer umgesetzten Algorithmen) zugänglich ist, muss sich auf endlichdimensionale Betrachtungen stützen. Eine Lösung ist also in einem endlichdimensionalen Raum zu suchen. Wir wählen daher einen **Ansatzraum** $V_h \subset V$ und schreiben als Ersatz folgende Aufgaben auf:

Gegeben $\lambda \in V^*$

Gesucht $u_h \in V_h$ mit

$$a(v, u_h) = \langle \lambda, v \rangle \text{ für alle } v \in V_h. \quad (7.30)$$

Betrachte dazu die Aufgabe

Gegeben $\lambda \in V^*$

Gesucht $u_h \in V_h$ mit

$$\frac{1}{2}a(u_h, u_h) - \langle \lambda, u_h \rangle = \min_{v \in V_h} \left(\frac{1}{2}a(v, v) - \langle \lambda, v \rangle \right) \quad (7.31)$$

Der Parameter $h > 0$ kann als Diskretisierungsparameter aufgefasst werden. In der Theorie der **Finiten Elemente** steht er für die Feinheit einer Triangulierung eines Gebiets (Ω im Dirichlet-Problem). Nach Wahl einer Basis im linearen Raum V_h kommt man nach geeigneten Umformungen und Betrachtungen zu einem linearen Gleichungssystem, zu dessen Lösung es hinreichend gute Software gibt. Dazu siehe etwa [29].

7.9 Anhang: Gelfand-Dreier

Beginnen wir mit einem Zusammenspiel von Räumen. Dazu benötigen wir die Begriffsbildung der Einbettung. Sind X, Y topologische Räume mit $X \subset Y$, so heißt die Abbildung $\iota : X \ni x \mapsto x \in Y$ die **(kanonische) Einbettung**. Ist diese Einbettung stetig, so nennen wir X **stetig eingebettet** in Y . Ist X zusätzlich noch dicht in Y , so nennen wir X **stetig und dicht eingebettet** in Y .

Satz 7.45 *Seien X, Y Banachräume. Ist X reflexiv und ist X stetig und dicht eingebettet in Y , so ist Y^* stetig eingebettet in X^* . Ist X reflexiv, so ist Y^* sogar dicht eingebettet.*

Beweis:

Ist X dicht in Y eingebettet, dann ist die Abbildung $\iota^* : Y^* \ni \mu \mapsto \mu|_X \in X^*$ linear und injektiv, denn: Die Linearität ist klar, die Injektivität folgt so:

$$\mu_X = \mu'_X \implies \langle \mu, x \rangle = \langle \mu', x \rangle \text{ für alle } x \in X,$$

woraus wegen der Dichtheit von X in Y und der Stetigkeit der Einbettung von X in Y die Gleichheit $\mu = \mu'$ folgt.

Beweisen wir nun den Zusatz. Gilt für ein $x \in X$, dass $\langle \lambda, x \rangle = 0$ ist für alle $\lambda \in Y^*$, so ist

$$\|x\|_Y = \sup_{\lambda \in Y^* \setminus \{\theta\}} \frac{\langle \lambda, x \rangle}{\|\lambda\|_{Y^*}} = 0,$$

also $x = \theta$ in Y . Wegen der Reflexivität von X bedeutet dies, dass ein auf Y^* verschwindendes Funktional $J_X(x) \in X^{**}$ das Nullfunktional ist. Nach dem Satz von Hahn-Banach ist dies nur möglich, wenn $\overline{Y^*} = X^*$ ist. ■

Sei $(V, \|\cdot\|)$ ein reflexiver Banach-Raum und sei $(H, \langle \cdot, \cdot \rangle)$ ein Hilbert-Raum. Nach dem Rieszschen Darstellungssatz können wir H mit seinem Dualraum H^* identifizieren. Ist nun V

dicht in H eingebettet, folgt daraus wie gesehen, dass H dicht in V^* eingebettet ist. Schreiben wir für eine dichte Einbettung kurz $\xhookrightarrow{d}$, so gilt:

$$V \xhookrightarrow{d} H \equiv H^* \xhookrightarrow{d} V^*$$

Ein Tripel (V, H, V^*) , das diese Eigenschaften erfüllt, nennen wir einen **Gelfand-Dreier (Evolutionstripel)**.

Beispiel 7.46 Das „einfachste“ Standardbeispiel für einen Gelfand-Dreier liefert

$$V := H_0^1(a, b) \xhookrightarrow{d} H := L_2(a, b) \xhookrightarrow{d} H^{-1}(a, b).$$

Dabei ist $H_0^1(a, b)$ der Banachraum, der entsteht, wenn man

$$C_0^1(a, b) := \{f : |a, b| \longrightarrow \mathbb{R} \mid f(a) = f(b) = 0, f \text{ stetig differenzierbar in } (a, b)\}$$

bezüglich der Norm

$$f \longmapsto \|f\| := \left(\int_{(a,b)} |f'(t)|^2 dt \right)^{\frac{1}{2}}$$

vervollständigt. $H^{-1}(a, b)$ bezeichnet dann den Dualraum von $H_0^1(a, b)$. □

Weitere Beispiele findet man bei der Analyse von partiellen Differentialgleichungen, insbesondere bei Evolutionsgleichungen.

7.10 Übungen

- 1.) Betrachte in l_2 den so genannten **Hilbertwürfel**

$$W := \{(x_n)_{n \in \mathbb{N}} \mid |x_n| \leq \frac{1}{n}, n \in \mathbb{N}\}.$$

Zeige:

- (a) W ist beschränkt.
 - (b) W ist konvex.
 - (c) W ist kompakt.
 - (d) $(\frac{1}{2n})_{n \in \mathbb{N}}$ ist ein Randpunkt von W .
- 2.) Sei X ein reeller Hilbertraum und $a : X \times X \longrightarrow \mathbb{R}$ eine stetige, symmetrische Bilinearform, für die gilt:

$$a(x, x) \geq \|x\|^2, x \in X. \quad (\text{Elliptizität})$$

Zu gegebenem $\lambda \in X^*$ betrachte das Problem

$$\text{Minimiere } I(X) := \frac{1}{2}a(x, x) - \langle \lambda, x \rangle, x \in X.$$

Zeige: Es gibt genau eine Lösung und diese Lösung ist zugleich Lösung von

$$\text{Gesucht } x \in X \text{ mit } a(x, y) = \langle \lambda, y \rangle \text{ für alle } y \in X.$$

- 3.) Sei X ein Hilbertraum und sei $(x_n)_{n \in \mathbb{N}}$ eine Folge in X . Zeige für $x \in X$ die Äquivalenz von

- (a) $\lim_n x_n = x$ (in der Normtopologie).
 (b) $(x_n)_{n \in \mathbb{N}}$ konvergiert schwach gegen x und $\lim_n \|x_n\| = \|x\|$.
 4.) Sei $(X, \langle \cdot | \cdot \rangle_X)$ ein Hilbertraum mit Skalarkörper $\mathbb{C}$ und induzierter Norm $\|\cdot\|_X$. Seien $x, y \in X$. Zeige die so genannte **Polarisationsidentität**:

$$\langle x | y \rangle_X = \frac{1}{4} \|x + y\|^2 - \|x - y\|_H^2 + i \|x + iy\|_H^2 - i \|x - iy\|_H^2$$

- 5.) Sei X ein reeller Hilbertraum und sei $P : X \rightarrow X$ eine Projektion, d.h. $P \circ P = P$. Zeige die Äquivalenz von
 (a) P ist orthogonale Projektion, d.h. $\ker(P) = (\text{ran}(P))^\perp$;
 (b) P ist selbstadjungiert, d.h. $P = P^*$.
 (c) P ist normal, d.h. $P^*P = PP^*$.
 (d) $\langle Px, x \rangle \geq 0$ für alle $x \in X$.
 6.) Sei X ein reeller Hilbertraum und sei $(T_n)_{n \in \mathbb{N}}$ eine Folge stetiger linearer Operatoren von X nach X . Es gelte: $(T_n)_{n \in \mathbb{N}}$ konvergiert punktweise schwach gegen T , $(T_n^*)_{n \in \mathbb{N}}$ konvergiert punktweise schwach gegen T^* . Zeige: $(T_n \circ T_n^*)_{n \in \mathbb{N}}$ konvergiert punktweise schwach gegen $T \circ T^*$.
 7.) Sei X ein komplexer Hilbertraum. Dann sind für einen stetigen linearen Operator $L : X \rightarrow X$ äquivalent:
 (a) L ist normal.
 (b) Es gibt stetige lineare Operatoren $L_1, L_2 : X \rightarrow X$ mit

$$L = L_1 + L_2, \quad L_1, L_2 \text{ sind selbstadjungiert, } L_1 \circ L_2 = L_2 \circ L_1.$$

- 8.) Seien $(X_1, \langle \cdot, \cdot \rangle_1)$, $(X_2, \langle \cdot, \cdot \rangle_2)$ Hilberträume. Dann ist $X := X_1 \times X_2$, versehen mit dem Skalarprodukt $\langle (x_1, x_2), (y_1, y_2) \rangle_t := t \langle x_1, y_1 \rangle + (1 - t) \langle x_2, y_2 \rangle$, für jedes $t \in (0, 1)$ ein Hilbertraum.
 9.) Sei X ein reeller Hilbertraum und $T : X \rightarrow X$ habe die Darstellung

$$Tx = \sum_{n \in \mathbb{N}} \lambda_n \langle x, e^n \rangle e^n, \quad x \in X,$$

wobei $(e^n)_{n \in \mathbb{N}}$ eine Orthonormalbasis und λ_n eine Nullfolge ist.

- (a) Sei $\psi : \sigma(T) \rightarrow \mathbb{R}$ stetig. Man mache einen Versuch, einen Operator $\psi(T)$ zu definieren.
 (b) Versuche eine Darstellung von $(T^*T)^{1/2}$.
 10.) Sei X ein Hilbertraum mit Orthonormalbasis $(e^n)_{n \in \mathbb{N}}$ und sei $(\lambda_n)_{n \in \mathbb{N}}$ eine reelle Nullfolge.
 (a) Wie sieht das Spektrum von

$$T : X \rightarrow X, \quad Tx := \sum_{n=1}^{\infty} \lambda_n \langle x | e^n \rangle e^n,$$

aus?

- (b) Berechne den adjungierten Operator T^* .

Literaturverzeichnis

- [1] H.W. Alt. *Lineare Funktionalanalysis*. Springer, 2002.
- [2] J. Appell and M. Văth. *Elemente der Funktionalanalysis*. Vieweg, Braunschweig, 2005.
- [3] W. Arendt. *Funktionalanalysis*, 2006.
- [4] K. Atkinson and W. Han. *Theoretical Numerical Analysis. A Functional Analysis Framework*. Texts in applied mathematics, vol. 39. Springer, New York, 2001.
- [5] J.-P. Aubin. *Applied Functional Analysis*. Wiley, Chichester, second edition, 1999.
- [6] D. Bengs. Greedy approximation im raum der funktionen von beschränkter variation und anwendungen auf die bildverarbeitung. Master's thesis, Fachbereich Informatik und Mathematik, 2011.
- [7] M. Blatter. *Wavelets—Eine Einführung*. Vieweg, Braunschweig, 1998.
- [8] C. Chidume. *Geometric properties of Banach spaces and nonlinear iterations*. Springer, New York, 2009.
- [9] M. Day. Reflexive Banach spaces not isomorphic to uniformly convex spaces. *Bull. Amer. Math. Soc.*, pages 313–317, 1941.
- [10] M. Dobrowolski. *Angewandte Funktionalanalysis*. Springer, 2006.
- [11] A. Dvoretzky and C.A. Rogers. Absolute and unconditional convergence in normed linear spaces. volume 36 of *Proceedings of the National Academy of Science of the United States of America*, pages 192–197. Academy of Science of the United States of America, 1950.
- [12] P. Enflo. A counterexample to the approximation problem in Banach spaces. *Acta Mathematica*, 130:309–317, 1973.
- [13] M. Fabian, V. Montesinos, P. Habala, V. Zizler, and P. Hajek. *Banach space theory: the basis for linear and nonlinear analysis*. Springer, New York, 2011.
- [14] O. Forster. *Analysis 1,2*. Vieweg, Wiesbaden, 1996.
- [15] H.-L. Gau and N.-C. Wong. Some converses of the strong separation theorems. *Proc. Amer. Math. Soc.*, 124:2443–2449, 1996.
- [16] T. Gowers. A new dichotomy for Banach spaces. *Geometric Functional Analysis*, pages 1083–1093, 1996.
- [17] T. Gowers and B. Maurey. The unconditional basic sequence problem. *J. Amer. Math. Soc.*, pages 851–874, 1993.

- [18] H. Heuser. *Lehrbuch der Analysis, Teil 1*. Teubner, 1990.
- [19] H. Heuser. *Funktionalanalysis*. Teubner, 1992.
- [20] F. Hirzebruch and W. Scharlau. *Einführung in die Funktionalanalysis*. Bibliographisches Institut, 1970.
- [21] V. Klee. Convex sets in linear spaces. *Duke Math. J.*, 18:443–466, 1951.
- [22] R. Kress. *Linear Integral Equations*. Springer, Berlin, 1999.
- [23] Kreyszig. *Introductory Functional Analysis with Applications*. John Wiley & Sons, New York, 1978.
- [24] P.D. Lax and A.N. Milgram. Parabolic equations. *Annals of Mathematics Studies*, 33:167–190, 1964.
- [25] J. Lindenstrauss and L. Tzafriri. *Classical Banach spaces I and II*. Springer, Berlin, 1977.
- [26] J. Lindström. On the origin and early history of functional analysis. Technical report, Uppsala University, 2008.
- [27] J.T. Marti. *Konvexe Analysis*. Birhäuser, Basel, 1977.
- [28] R. Meise and D. Vogt. *Einführung in die Funktionalanalysis*. Vieweg, Braunschweig, 1992.
- [29] A. Quarterioni, R. Sacco and F. Saleri. *Numerical Mathematics*. Springer, New York, 2000.
- [30] H. Schubert. *Topologie*. Teubner, Stuttgart, 1971.
- [31] I. Singer. *Bases in Banach spaces I*, volume 154 of *Grundlehren der Mathematischen Wissenschaften*. Springer, Berlin, 1970.
- [32] A.D. Sokal. A really simple elementary proof of the uniform boundedness theorem. *Amer. Math. Monthly*, 118:450–452, 2011.
- [33] W.A. Strauss. *Partielle Differentialgleichungen*. Vieweg, 1992.
- [34] V.N. Temlyakov. Greedy approximation. *Acta Mathematica*, 17.
- [35] J. Weidmann. *Linear operators in Hilbert spaces*. Springer, 1980.
- [36] D. Werner. Unzerlegbare Banachräume. Über die Arbeiten von W.T. Gowers. Freie Universität, Fachbereich Mathematik und Informatik, 14 Seiten, 2000.
- [37] D. Werner. *Funktionalanalysis*. Springer, 2002.
- [38] J. Wloka. *Partielle Differentialgleichungen*. Teubner, Stuttgart, 1982.
- [39] K. Yosida. *Functional analysis*. Springer, 1980.