# Funktionalanalysis und Integrationstheorie

## Vorlesungsnotizen

**Johannes Kepler Universität Linz**
**Technisch-Naturwissenschaftliche Fakultät**
**Institut für Analysis**

Prof. Dr. Aicke Hinrichs
Wintersemester 2015/16
Version vom 2. Februar 2016

# Inhaltsverzeichnis

# 1 Maß- und Integrationstheorie

Das in den Grundvorlesungen zur Analysis meist zuerst eingeführte bestimmte Integral ist das Riemann-Integral. Seine Einführung ist sehr anschaulich, wichtige Resultate wie der Hauptsatz der Differential- und Integralrechnung lassen sich mit geringem Aufwand beweisen.

Für weite Teile der höheren Analysis ist das Riemann-Integral aber nur bedingt geeignet. Schon die Definition für Funktionen auf dem $\mathbb{R}^d$ ist schwieriger für $d \geq 2$ als für $d = 1$. Noch dramatischer wird es bei Vertauschungssätzen für Grenzwert- und Integral. Unter welchen Bedingungen an die Funktionenfolge $(f_n)$ gilt

$$\lim_{n\to\infty} \int_0^1 f_n(t)\,\mathrm{d}t = \int_0^1 \lim_{n\to\infty} f_n(t)\,\mathrm{d}t?$$

Selbst für stetige $f_n$ kann es sein, dass die Grenzfunktion

$$f(t) = \lim_{n\to\infty} f_n(t)$$

nicht Riemann-integrierbar ist.

Dieses Manko behebt das Lebesgue-Integral. Wichtig wird sein, ein geeignetes System von messbaren Mengen zur Verfügung zu haben. Intervalle reichen nicht. Dies führt zum allgemeinen Begriff der $\sigma$-Algebra und zum Maßbegriff, der für die Analysis wie auch für die Wahrscheinlichkeitsrechnung zentral ist. Aus dem Maß leiten wir dann einen Integralbegriff ab, der im Fall des $\mathbb{R}^d$ das Lebesgue-Integral liefert.

## 1.1 $\sigma$-Algebren

Im folgenden betrachten wir eine Grundmenge $\Omega$ und ihre Potenzmenge $\mathcal{P}(\Omega)$. Für ein $A \in \mathcal{P}(\Omega)$ bezeichnen wir mit $\overline{A} = \Omega \setminus A$ das Komplement von $A$ in $\Omega$. Aus technischen Gründen benötigen wir noch den Begriff eines Rings, bevor wir zum zentralen Begriff der $\sigma$-Algebra kommen.

**Definition 1.1.** Eine Teilmenge $\mathcal{A} \subseteq \mathcal{P}(\Omega)$ heißt *Ring*, falls die folgenden Eigenschaften gelten:

(a) $\emptyset \in \mathcal{A}$

(b) $A, B \in \mathcal{A} \Rightarrow A \setminus B \in \mathcal{A}$

(c) $A, B \in \mathcal{A} \Rightarrow A \cup B \in \mathcal{A}$.

Ein Ring ist also ein differenz- und vereinigungsstabiles Mengensystem, dass zumindest die leere Menge enthält. Wegen $A \cap B = A \setminus (A \setminus B)$ gilt in einem Ring $\mathcal{A}$ auch

(d) $A, B \in \mathcal{A} \Rightarrow A \cap B \in \mathcal{A}$,

ein Ring ist also auch durchschnittsstabil. Die Eigenschaften $(c)$ und $(d)$ lassen sich per Induktion auf endliche Vereinigungen und Durchschnitte erweitern.

**Beispiel.** $\{A \subseteq \Omega \,:\, A \text{ ist endlich}\}$ ist ein Ring.

**Beispiel.** Sei

$$\mathcal{I} = \{(a, b] \,:\, a, b \in \mathbb{R},\, a \leq b\}$$

die Menge aller (links) halboffenen Intervalle. Weiter sei $\mathcal{F}$ die Menge aller endlichen Vereinigungen von Intervallen in $\mathcal{I}$. Dann ist $\mathcal{F}$ ein Ring. Offensichtlich gilt wegen $\emptyset = (a, a]$ die Eigenschaft $(a)$. Auch $(c)$ ist offensichtlich. Überlegen Sie sich die Eigenschaft $(b)$!

**Beispiel.** Analog zum vorhergehenden Beispiel definieren wir im $\mathbb{R}^d$ mittels

$$\mathcal{I}^d = \{(a,b] \,:\, a=(a_j), b=(b_j) \in \mathbb{R}^d,\, a_j \le b_j \text{ für } j=1,\dots,d\}$$

die Menge aller (links) halboffenen Intervalle und mit $\mathcal{F}^d$ die Menge aller endlichen Vereinigungen von Intervallen in $\mathcal{I}^d$. Dann ist $\mathcal{F}^d$ ein Ring.

Beweis: in der Vorlesung. Insbesondere wird gezeigt, dass jede Figur aus $\mathcal{F}^d$ sich als *disjunkte* Vereinigung von Intervallen aus $\mathcal{I}^d$ darstellen läßt.

**Definition 1.2.** Eine Teilmenge $\mathcal{A} \subseteq \mathcal{P}(\Omega)$ heißt *σ-Algebra*, falls die folgenden Eigenschaften gelten:

(a) $\emptyset, \Omega \in \mathcal{A}$

(b) $A \in \mathcal{A} \Rightarrow \overline{A} \in \mathcal{A}$

(c) $A_1, A_2, \dots \in \mathcal{A} \Rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$.

Eine $\sigma$-Algebra ist also abgeschlossen unter der Bildung von Komplementen und abzählbaren Vereinigungen und enthält zumindest die leere und die ganze Menge. Eine $\sigma$-Algebra ist ebenfalls abgeschlossen unter endlichen Vereinigungen und unter endlichen und abzählbaren Durchschnitten (tatsächlich?). Jede $\sigma$-Algebra ist ein Ring. Der Durchschnitt beliebiger $\sigma$-Algebren ist wieder eine $\sigma$-Algebra (Übung!).

**Beispiel** (Kleinste $\sigma$-Algebra)**.** $\{\emptyset, \Omega\}$ ist eine $\sigma$-Algebra.

**Beispiel** (Größte $\sigma$-Algebra)**.** $\mathcal{P}(\Omega)$ ist eine $\sigma$-Algebra.

**Beispiel** (Erzeugte $\sigma$-Algebra)**.** Ist $\mathcal{E} \subseteq \mathcal{P}(\Omega)$ ein beliebiges Mengensystem, dann ist der Durchschnitt aller $\mathcal{E}$ umfassenden $\sigma$-Algebren die sogenannte *von $\mathcal{E}$ erzeugte σ-Algebra*. In diesem Fall heißt $\mathcal{E}$ auch *Erzeuger* von $\sigma(\mathcal{E})$. Offensichtlich gilt für $\mathcal{E}, \mathcal{E}' \subseteq \mathcal{P}(\Omega)$

$$\mathcal{E} \subseteq \mathcal{E}' \Rightarrow \sigma(\mathcal{E}) \subseteq \sigma(\mathcal{E}').$$

Ist $\mathcal{A}$ eine $\sigma$-Algebra und $\mathcal{E} \subseteq \mathcal{A}$, so ist $\sigma(\mathcal{E}) \subseteq \mathcal{A}$.

**Beispiel** (Spur einer $\sigma$-Algebra)**.** Ist $\Omega' \subseteq \Omega$ und ist $\mathcal{A} \subseteq \mathcal{P}(\Omega)$ eine $\sigma$-Algebra, so ist die *Spur* von $\mathcal{A}$ auf $\Omega'$ die durch

$$\mathcal{A}' = \{A \cap \Omega' \,:\, A \in \mathcal{A}\} = \{A' \in \mathcal{P}(\Omega') \,:\, \text{es gibt } A \in \mathcal{A} \text{ mit } A' = A \cap \Omega'\} \subseteq \mathcal{P}(\Omega')$$

gegebene $\sigma$-Algebra auf der Grundmenge $\Omega'$.

Das (zumindest für uns) wichtigste Beispiel ist die von den halboffenen Intervallen $\mathcal{I}^d$ oder von $\mathcal{F}^d$ in $\mathbb{R}^d$ erzeugte $\sigma$-Algebra, die *Borelsche σ-Algebra* genannt wird und mit $\mathcal{B}(\mathbb{R}^d)$ bezeichnet wird. Auch die Spuren $\mathcal{B}(E)$ der Borelschen $\sigma$-Algebra auf Teilmengen $E \subseteq \mathbb{R}^d$ werden wir betrachten. Die Elemente der Borelschen $\sigma$-Algebra heißen *Borelmengen*.

Andere Erzeugendensysteme der Borelschen $\sigma$-Algebra fasst der folgende Satz zusammen.

**Satz 1.3** (Charakterisierung der Borelschen $\sigma$-Algebra). *Die Borelsche $\sigma$-Algebra $\mathcal{B}(\mathbb{R})$ wird erzeugt von jedem der Systeme bestehend aus*

- *allen Intervallen $(-\infty, x]$ mit $x \in \mathbb{R}$.*
- *allen Intervallen $(-\infty, x)$ mit $x \in \mathbb{R}$.*
- *allen Intervallen $(-\infty, x]$ mit $x \in \mathbb{Q}$.*
- *allen Intervallen $(-\infty, x)$ mit $x \in \mathbb{Q}$.*
- *allen offenen Mengen in $\mathbb{R}$.*
- *allen abgeschlossenen Mengen in $\mathbb{R}$.*
- *allen kompakten Mengen in $\mathbb{R}$.*

*Entsprechendes gilt für die Borelsche $\sigma$-Algebra $\mathcal{B}(\mathbb{R}^d)$.*

*Beweis.* Wir führen beispielhaft Nachweise von vier Inklusionen (zwei nur im Fall $d = 1$) durch. Überlegen Sie sich weitere! Wir erinnern zunächst daran, dass per definitionem die Menge $\mathcal{I} = \{(a, b] : a \leq b\}$ von halboffenen Intervallen die Borelsche $\sigma$-Algebra in $\mathcal{R}$ erzeugt.

- Sei $\mathcal{E}$ die Menge aller Intervalle der Form $(-\infty, x]$ mit $x \in \mathbb{Q}$. Dann gilt $\mathcal{E} \subseteq \mathcal{B}(\mathbb{R})$ wegen $(-\infty, x] = \bigcup_{n>|r|}(-n, r)$. Es folgt $\sigma(\mathcal{E}) \subseteq \mathcal{B}(\mathbb{R})$.
- Wegen $(r, s] = (-\infty, s] \setminus (-\infty, r]$ folgt zunächst $\mathcal{I} \subseteq \sigma(\mathcal{E})$ und damit $\mathcal{B}(\mathbb{R}) = \sigma(\mathcal{I}) \subseteq \sigma(\mathcal{E})$. Wir erhalten $\mathcal{B}(\mathbb{R}) = \sigma(\mathcal{E})$.
- Seien $\mathcal{A}$ bzw. $\mathcal{O}$ das System aller abgeschlossenen bzw. offenen Mengen in $\mathbb{R}^d$. Da das Komplement einer abgeschlossenen Menge offen ist, gilt $\mathcal{A} \subseteq \sigma(\mathcal{O})$ und es folgt $\sigma(\mathcal{A}) \subseteq \sigma(\mathcal{O})$. Die Inklusion $\sigma(\mathcal{O}) \subseteq \sigma(\mathcal{A})$ folgt analog. Also ist $\sigma(\mathcal{A}) = \sigma(\mathcal{O})$.
- Sei nun $\mathcal{E}$ die Menge aller Intervalle der Form $(-\infty, x)$ mit $x \in \mathbb{Q}$ und $\mathcal{O}$ das System aller offenen Mengen in $\mathbb{R}$. Wegen $\mathcal{E} \subseteq \mathcal{O}$ folgt $\sigma(\mathcal{E}) \subseteq \sigma(\mathcal{O})$.

□

Leider gibt es keine explizite Beschreibung aller Borelmengen. Allerdings kann man mit folgendem *Prinzip der guten Mengen* viele Eigenschaften beliebiger Borelmengen zeigen:

- Zu zeigen sei, dass jedes Element einer gegebenen $\sigma$-Algebra $\mathcal{A}$ eine gewisse Eigenschaft hat.
- Sei $\mathcal{A}'$ die Menge aller Elemente von $\mathcal{A}$, die diese Eigenschaft haben.
- Man zeigt zunächst, dass jedes Element eines geeigneten Erzeugendensystems $\mathcal{E}$ diese Eigenschaft hat.
- Dann zeigt man, dass $\mathcal{A}'$ selbst eine $\sigma$-Algebra ist.
- Es folgt $\mathcal{A} = \sigma(\mathcal{E}) \subseteq \sigma(\mathcal{A}') \subseteq \mathcal{A}$ und damit wie behauptet $\mathcal{A}' = \mathcal{A}$.

Wir demonstrieren das an folgendem Satz.

**Satz 1.4** (Translationsinvarianz von Borelmengen). *Ist $A \subseteq \mathbb{R}^d$ eine Borelmenge und $x \in \mathbb{R}^d$, so ist auch $x + A = \{x + a : a \in A\}$ eine Borelmenge.*

*Beweis.* Wir setzen
$$\mathcal{A}' = \{A \in \mathcal{B}(\mathbb{R}^d) : x + A \in \mathcal{B}(\mathbb{R}^d)\}.$$
Da Translationen von Intervallen in $\mathcal{I}^d$ wieder Intervalle in $\mathcal{I}^d$ sind, gilt $\mathcal{I}^d \subseteq \mathcal{A}'$. Dass $\mathcal{A}'$ eine $\sigma$-Algebra ist, folgt aus der Vertauschbarkeit von Komplementbildung und Verschiebung sowie Vereinigung und Verschiebung. Da $\mathcal{I}^d$ ein Erzeugendensystem für $\mathcal{B}(\mathbb{R}^d)$ ist, folgt $\mathcal{A}' = \mathcal{B}(\mathbb{R}^d)$. □

## 1.2 Maße

Inhalte und Maße sind Funktionen, die den Mengen aus einer $\sigma$-Algebra positive reelle Zahlen oder $+\infty$ zuordnen. Dabei sollen natürliche Bedingungen wie die Additivität bezüglich disjunkter Mengen gelten. Für das Rechnen mit $+\infty$ benutzen wir die folgenden Vereinbarungen.

- $a + \infty = \infty + a = \infty + \infty = \infty$ für alle $a \in \mathbb{R}$
- Ist $\sum_{i=1}^{\infty} a_i$ eine divergente Reihe mit $a_i \in [0, \infty)$, so ist $\sum_{i=1}^{\infty} a_i = \infty$
- Ist $\sum_{i=1}^{\infty} a_i$ eine Reihe mit $a_i \in [0, \infty]$ und ist mindestens ein $a_i = \infty$, so ist $\sum_{i=1}^{\infty} a_i = \infty$.

**Definition 1.5.** Sei $\mathcal{A} \subseteq$ ein Ring.

1. Ein *Inhalt* oder eine *endlich additive Funktion* auf $\mathcal{A}$ ist eine Funktion $\mu : \mathcal{A} \to [0, \infty]$ mit $\mu(\emptyset) = 0$ und
$$\mu\Big(\bigcup_{i=1}^{n} A_i\Big) = \sum_{i=1}^{n} \mu(A_i)$$
für paarweise disjunkte $A_1, \dots, A_n \in \mathcal{A}$.
2. Ein *Prämaß* oder eine *$\sigma$-additive Funktion* auf $\mathcal{A}$ ist eine Funktion $\mu : \mathcal{A} \to [0, \infty]$ mit $\mu(\emptyset) = 0$ und
$$\mu\Big(\bigcup_{i=1}^{\infty} A_i\Big) = \sum_{i=1}^{\infty} \mu(A_i)$$
für paarweise disjunkte $A_1, A_2, \dots \in \mathcal{A}$ mit $\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$.
3. Ist $\mathcal{A}$ eine $\sigma$-Algebra und $\mu$ ein Prämaß auf $\mathcal{A}$, so heißt $\mu$ *Maß*.
4. Ein Maß auf einer $\sigma$-Algebra $\mathcal{A}$ mit $\mu(\Omega) = 1$ heißt *Wahrscheinlichkeitsmaß*.

Oft spricht man auch von einem Maß auf $\Omega$, falls klar ist, welche $\sigma$-Algebra zugrunde liegt. Insbesondere betrachten wir für Teilmengen $E$ von $\mathbb{R}^d$ meist die Borelsche $\sigma$-Algebra und sprechen von einem Maß auf $E$. Die Bedingung $\mu(\emptyset) = 0$ dient eigentlich nur dazu, den trivialen Fall $\mu(A) = \infty$ für alle $A \in \mathcal{A}$ auszuschließen.

**Beispiel** (Dirac-Maß). Ist $\mathcal{A}$ eine $\sigma$-Algebra und $x \in \Omega$ fixiert, so ist das *Dirac-Maß* $\delta_x$ definiert durch $\delta_x(A) = \chi_A(x)$, ist also 1 für die $x$ enthaltenden $A$ und 0 sonst. Die Eigenschaften eines Maßes prüft man leicht nach.

**Beispiel.** Ein (relativ triviales) Maß, das auf der ganzen Potenzmenge $\mathcal{P}(\Omega)$ definiert werden kann, ist gegeben durch $\mu(A) = \infty$ für $A \neq \emptyset$ (und natürlich $\mu(\emptyset) = 0$).

**Beispiel** (Zählmaß)**.** Ist $\mathcal{A}$ eine $\sigma$-Algebra und $\mu(A)$ gleich der Anzahl der Elemente von $A$ für endliches $A \in \mathcal{A}$ und gleich $\infty$ sonst, so ist $\mu$ ein Maß, das sogenannte *Zählmaß* auf $\mathcal{A}$. Es kann ebenfalls auf der ganzen Potenzmenge $\mathcal{P}(\Omega)$ definiert werden.

**Beispiel.** Wir betrachten jetzt den Ring $\mathcal{F}^d$ der Figuren. Für ein Intervall $I = \prod_{i=1}^d (a_i, b_i] \in \mathcal{I}^d$ sollte man natürlich den Inhalt definieren als

$$\mu(I) = \prod_{i=1}^{d} (b_i - a_i).$$

Im Beweis der Ringeigenschaft in der Vorlesung haben wir insbesondere gezeigt, dass sich jedes $A \in \mathcal{F}^d$ als *disjunkte* Vereinigung von Intervallen $A = \cup_{j=1}^n I_j$ mit $I_j \in \mathcal{I}^d$ schreiben lässt. Dann wird man natürlich den Inhalt additiv fortsetzen durch

$$\mu(A) = \sum_{i=1}^{n} \mu(I_j).$$

Man erhält den *Jordan-Inhalt* $\lambda^d$, der bei der Definition des Riemann-Integrals in $\mathbb{R}^d$ benutzt wird. Die Unabhängigkeit von der konkreten Darstellung als Vereinigung disjunkter Intervalle beweist man dort, sie steckt in der Linearität des Riemann-Integrals. Insgesamt erhalten wir also einen Inhalt (den Jordan-Inhalt) auf dem Ring der Figuren $\mathcal{F}^d$. Diesen werden wir später zu einem Maß auf die $\sigma$-Algebra der Borelmengen fortsetzen, zum Lebesgue-Maß. Dazu müssen wir natürlich zunächst die $\sigma$-Additivität des Jordan-Inhalts auf $\mathcal{F}^d$ zeigen. Das war historisch eine der wichtigen neuen Erkenntnisse von Lebesgue.

Inhalte sind nur additiv. Der nächste Satz zeigt unter anderem, dass zumindest eine Ungleichung (statt einer Gleichung) auch für abzählbare Vereinigungen disjunkter Mengen gilt.

**Satz 1.6.** *Für einen Inhalt $\mu$ auf einem Ring $\mathcal{A}$ gilt:*

*(a) Sind $A, B \in \mathcal{A}$ mit $A \subseteq B$, so ist $\mu(A) \leq \mu(B)$.*

*(b) Ist $A_1, A_2, \dots \in \mathcal{A}$ eine Folge paarweiser disjunkter Mengen, deren Vereinigung wieder in $\mathcal{A}$ ist,so ist*

$$\mu\Big(\bigcup_{i=1}^{\infty} A_i\Big) \geq \sum_{i=1}^{\infty} \mu(A_i).$$

*(c) Ist $A_1, A_2, \dots \in \mathcal{A}$ eine Folge beliebiger Mengen, deren Vereinigung wieder in $\mathcal{A}$ ist, und ist $\mu$ zusätzlich $\sigma$-additiv, so gilt*

$$\mu\Big(\bigcup_{i=1}^{\infty} A_i\Big) \leq \sum_{i=1}^{\infty} \mu(A_i).$$

*Beweis.* (*a*) Schreiben wir $B = A \cup (B \setminus A)$ als disjunkte Vereinigung von Mengen aus $\mathcal{A}$, so folgt $\mu(A) \leq \mu(A) + \mu(B \setminus A) = \mu(B)$.

(*b*) Für jedes $n \in \mathbb{N}$ folgt nach (*a*) $\mu\Big(\bigcup_{i=1}^{\infty} A_i\Big) \geq \mu\Big(\bigcup_{i=1}^{n} A_i\Big) = \sum_{i=1}^{n} \mu(A_i)$ und damit die Behauptung durch Grenzübergang $n \to \infty$.

(c) Mittels des Ansatzes $B_1 = A_1$ und $B_i = A_i \setminus \bigcup_{j=1}^{i-1} A_j$ für $i = 2, 3 \dots$ schreiben wir

$$A = \bigcup_{i=1}^{\infty} A_i = \bigcup_{i=1}^{\infty} B_i$$

als disjunkte Vereinigung von Mengen aus $\mathcal{A}$. Aus $B_i \subseteq A_i$ und (a) folgt die Behauptung

$$\mu\Big(\bigcup_{i=1}^{\infty} A_i\Big) = \mu(A) = \mu\Big(\bigcup_{i=1}^{\infty} B_i\Big) = \sum_{i=1}^{\infty} \mu(B_i) \leq \sum_{i=1}^{\infty} \mu(A_i).$$

□

Mit diesem Satz können wir das folgende praktische Kriterium für $\sigma$-Additivität zeigen.

**Satz 1.7** ($\sigma$-Additivitätskriterium). *Für einen Inhalt $\mu$ auf einem Ring $\mathcal{A}$ sind die folgenden Eigenschaften äquivalent:*

*(a) $\mu$ ist $\sigma$-additiv.*

*(b) Sei $A_1 \subseteq A_2 \subseteq \dots$ eine aufsteigende Folge von Mengen in $\mathcal{A}$ und sei $A = \bigcup_{i=1}^{\infty} A_i$ ebenfalls in $\mathcal{A}$. Dann gilt*

$$\mu(A) = \lim_{i \to \infty} \mu(A_i).$$

*(c) Sei $A_1 \supseteq A_2 \supseteq \dots$ eine absteigende Folge von Mengen in $\mathcal{A}$ und sei $A = \bigcap_{i=1}^{\infty} A_i$ ebenfalls in $\mathcal{A}$. Weiter sei $\mu(A_1) < \infty$. Dann gilt*

$$\mu(A) = \lim_{i \to \infty} \mu(A_i).$$

*(d) Sei $A_1 \supseteq A_2 \supseteq \dots$ eine absteigende Folge von Mengen in $\mathcal{A}$ mit $\emptyset = \bigcap_{i=1}^{\infty} A_i$. Weiter sei $\mu(A_1) < \infty$. Dann gilt*

$$0 = \lim_{i \to \infty} \mu(A_i).$$

*Beweis.* Wir machen einen Ringbeweis $(a) \Rightarrow (b) \Rightarrow (c) \Rightarrow (d) \Rightarrow (a)$.

$(a) \Rightarrow (b)$: Mittels des Ansatzes $B_1 = A_1$ und $B_i = A_i \setminus \bigcup_{j=1}^{i-1} A_j$ für $i = 2, 3, \dots$ schreiben wir wieder $A = \bigcup_{i=1}^{\infty} B_i$ und $A_n = \bigcup_{i=1}^{n} B_i$ als disjunkte Vereinigung von Mengen aus $\mathcal{A}$. Aus der Voraussetzung (a) folgt dann

$$\mu(A) = \sum_{i=1}^{\infty} \mu(B_i) = \lim_{n \to \infty} \sum_{i=1}^{n} \mu(B_i) = \mu(A_n).$$

$(b) \Rightarrow (c)$: Hier setzen wir $B_n = A_1 \setminus A_n$ und benutzen die Voraussetzung (b) für die aufsteigende Folge $B_1 \subseteq B_2 \subseteq \dots$ mit $\bigcup_{i=1}^{\infty} B_i = A_1 \setminus A$. Wir erhalten

$$\mu(A_1) - \mu(A) = \mu(A_1 \setminus A) = \lim_{i \to \infty} \mu(A_1 \setminus A_i) = \lim_{i \to \infty} \big(\mu(A_1) - \mu(A_i)\big) = \mu(A_1) - \lim_{i \to \infty} \mu(A_i).$$

Es folgt die Behauptung. Wo haben wir eigentlich $\mu(A_1) < \infty$ verwendet?

$(c) \Rightarrow (d)$: ist offensichtlich.

$(d) \Rightarrow (a)$: Seien $A_1, A_2, \dots \in \mathcal{A}$ paarweise disjunkt und $A = \bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$. Ist $\sum_{i=1}^{\infty} \mu(A_i) = \infty$, so ist nach Satz 1.6 $(b)$ auch $\mu(A) = \infty$. Ist $\sum_{i=1}^{\infty} \mu(A_i) < \infty$, so ist die Voraussetzung $(d)$ auf die Folge $B_n = \bigcup_{i=n}^{\infty} A_i \in \mathcal{A}$ anwendbar und wir erhalten

$$\mu(A) = \mu(A_1 \cup \dots \cup A_n \cup B_{n+1}) = \sum_{i=1}^{n} \mu(A_i) + \mu(B_{n+1}) \to \sum_{i=1}^{\infty} \mu(A_i),$$

$\mu$ ist also $\sigma$-additiv.

□

Mit diesem Kriterium sind wir jetzt in der Lage, wie angekündigt die $\sigma$-Additivität des Jordan-Inhalts zu zeigen.

**Satz 1.8** ($\sigma$-Additivität des Jordan-Inhalts). *Der Jordan-Inhalt $\lambda^d$ ist $\sigma$-additiv auf dem Ring $\mathcal{F}^d$.*

*Beweis.* In diesem Beweis benutzen wir $\overline{A}$ für den Abschluss der Menge $A \subseteq \mathbb{R}^d$, nicht für das Komplement.

Wir zeigen $(d)$ in Satz 1.7. Sei also $A_1 \supseteq A_2 \supseteq \dots$ eine absteigende Folge von Mengen in $\mathcal{F}^d$ mit $\emptyset = \bigcap_{i=1}^{\infty} A_i$. Wir müssen zeigen: Zu gegebenem $\varepsilon > 0$ existiert ein $n_0$ mit $\lambda^d(A_n) < \varepsilon$ für $n \geq n_0$.

Sei also $\varepsilon > 0$ gegeben. Indem wir jedes $A_n$ ein wenig verkleinern finden wir zunächst Figuren $B_n \in \mathcal{F}^d$ mit Abschluss $\overline{B_n} \subsetneq A_n$ und $\lambda^d(A_n \setminus B_n) \leq 2^{-n}\varepsilon$. Die Mengen $\overline{B_n}$ sind kompakte Teilmengen des Kompaktums $\overline{A_1}$ und haben leeren Durchschnitt. Deshalb gibt es bereits endlich viele der $\overline{B_n}$ mit leerem Durchschnitt, es existiert also ein $n_0$ mit

$$\bigcap_{i=1}^{n} \overline{B_i} = \emptyset \qquad \text{für } n \geq n_0.$$

Für $C_n = \bigcap_{i=1}^{n} B_i$ zeigen wir jetzt induktiv

$$\lambda^d(A_n \setminus C_n) \leq (1 - 2^{-n})\varepsilon.$$

Wegen $C_n = \emptyset$ für $n \geq n_0$ folgt dann die Behauptung. Offensichtlich gilt nach Konstruktion

$$\lambda^d(A_1 \setminus C_1) = \lambda^d(A_1 \setminus B_1) \leq (1 - 2^{-1})\varepsilon = \frac{\varepsilon}{2}.$$

Im Induktionsschritt

$$\begin{aligned}
\lambda^d(A_{n+1} \setminus C_{n+1}) &= \lambda^d\big(A_{n+1} \setminus (C_n \cap B_{n+1})\big) \\
&= \lambda^d\big((A_{n+1} \setminus C_n) \cup (A_{n+1} \setminus B_{n+1})\big) \\
&\leq \lambda^d\big((A_{n+1} \setminus C_n)\big) + \lambda^d\big((A_{n+1} \setminus B_{n+1})\big) \\
&\leq \lambda^d\big((A_n \setminus C_n)\big) + \lambda^d\big((A_{n+1} \setminus B_{n+1})\big) \\
&\leq (1 - 2^{-n})\varepsilon + 2^{-(n+1)}\varepsilon \\
&= (1 - 2^{-(n+1)})\varepsilon
\end{aligned}$$

benutzen wir die Ungleichung $\lambda^d(A \cup B) \leq \lambda^d(A) + \lambda^d(B)$ (wo?), die für beliebige Inhalte gilt (Beweis!).

□

Es bleibt noch, das $\sigma$-additive Prämaß $\lambda^d$ vom Ring der Figuren $\mathcal{F}^d$ auf die gesamte Borelsche $\sigma$-Algebra fortzusetzen. Dass dies in eindeutiger Weise geschehen kann, zeigen wir in allgemeinerer Form mittels eines von Caratheodory gefundenen Konstruktion. Dazu definieren wir zunächst für einen beliebigen Inhalt $\mu$ auf einem Ring $\mathcal{A} \subseteq \mathcal{P}(\Omega)$ und beliebiges $A \subseteq \Omega$

$$\mu^*(A) = \inf \sum_{j=1}^{\infty} \mu(A_j)$$

wobei das Infimum über alle Folgen $A_1, A_2, \ldots$ in $\mathcal{A}$ mit $A \subseteq \bigcup_{j=1}^{\infty} A_j$ genommen wird. Falls es keine solche Folge gibt setzen wir $\mu^*(A) = \infty$.

Wir erhalten eine Mengenfunktion $\mu^* : \mathcal{P}(\Omega) \to [0, \infty]$ mit den Eigenschaften

(a) $\mu^*(\emptyset) = 0$

(b) $A \subseteq B \Rightarrow \mu^*(A) \leq \mu^*(B)$

(c) $A_1, A_2, \dots \in \mathcal{P}(\Omega) \Rightarrow \mu^*\Big(\bigcup_{i=1}^{\infty} A_i\Big) \leq \sum_{i=1}^{\infty} \mu^*(A_i)$.

Eine Mengenfunktion $\mu^* : \mathcal{P}(\Omega) \to [0, \infty]$ mit den Eigenschaften $(a), (b), (c)$ bezeichnet man als *äußeres Maß*. Nur die Eigenschaft $(c)$ bedarf eines Beweises. Ist $\mu^*(A_i) = \infty$ für ein $A_i$, so ist nichts zu zeigen. Also sei $\mu^*(A_i) < \infty$ für alle $i$. Dann wählen wir zu $\varepsilon > 0$ und jedem $i$ Mengen $A_{ij} \in \mathcal{A}$ mit $A_i \subseteq \bigcup_{j=1}^{\infty} A_{ij}$ und

$$\sum_{j=1}^{\infty} \mu(A_{ij}) \leq \mu^*(A_i) + 2^{-i}\varepsilon$$

und erhalten $A = \bigcup_{i=1}^{\infty} A_i \subseteq \bigcup_{i,j=1}^{\infty} A_{ij}$ und

$$\mu^*(A) \leq \sum_{i,j=1}^{\infty} \mu(A_{ij}) \leq \sum_{i=1}^{\infty} \mu^*(A_i) + \varepsilon.$$

Da $\varepsilon > 0$ beliebig war, folgt die Behauptung.

Die Eigenschaft $(c)$ impliziert insbesondere, dass für jedes $A \in \mathcal{P}(\Omega)$

$$\mu^*(Q) \leq \mu^*(Q \cap A) + \mu^*(Q \cap \overline{A}) \qquad \text{für alle } Q \subseteq \Omega$$

gilt. Von einem Maß würden wir natürlich in dieser Ungleichung sogar Gleichheit erwarten. Wir definieren deshalb

**Definition 1.9.** Sei $\mu$ ein Inhalt auf einem Ring $\mathcal{A}$ und sei $\mu^*$ das oben definierte zugehörige äußere Maß. Eine Menge $A \in \mathcal{P}(\Omega)$ heißt $\mu^*$-messbar, falls

$$\mu^*(Q) \geq \mu^*(Q \cap A) + \mu^*(Q \cap \overline{A}) \qquad \text{für alle } Q \subseteq \Omega \tag{1}$$

gilt. Die Menge der $\mu^*$-messbaren Mengen bezeichnen wir mit $\mathcal{M}_{\mu^*}$.

**Bemerkung.** Wie oben bereits bemerkt, kann man (1) ersetzen durch die Gleichheit

$$\mu^*(Q) = \mu^*(Q \cap A) + \mu^*(Q \cap \overline{A}) \qquad \text{für alle } Q \subseteq \Omega. \tag{2}$$

**Satz 1.10.** *Sei $\mu$ ein Inhalt auf einem Ring $\mathcal{A}$ und sei $\mu^*$ das oben definierte zugehörige äußere Maß. Dann ist $\mathcal{M}_{\mu^*}$ eine $\sigma$-Algebra und $\mu^*$ ist ein Maß auf $\mathcal{M}_{\mu^*}$.*

*Beweis.* Zunächst ist klar, dass $\mathcal{M}_{\mu^*}$ die leere Menge und $\Omega$ enthält und mit $A$ auch $\overline{A}$. Den weiteren Beweis teilen wir in einige Schritte auf.

1. Schritt: $\mathcal{M}_{\mu^*}$ ist ein Ring.

Dazu müssen wir noch zeigen, dass mit $A, B \in \mathcal{M}_{\mu^*}$ auch $A \cup B \in \mathcal{M}_{\mu^*}$ gilt. Da $B \in \mathcal{M}_{\mu^*}$ ist, folgt für beliebiges $Q \subseteq \Omega$ zunächst

$$\mu^*(Q \cap \overline{A}) = \mu^*(Q \cap \overline{A} \cap B) + \mu^*(Q \cap \overline{A} \cap \overline{B}).$$

Aus

$$Q \cap (A \cup B) = (Q \cap A) \cup (Q \cap \overline{A} \cap B)$$

folgt mit Eigenschaft $(c)$ von $\mu^*$ auch

$$\mu^*(Q \cap (A \cup B) \leq \mu^*(Q \cap A) + \mu^*(Q \cap \overline{A} \cap B).$$

Zusammen folgt

$$\begin{aligned} \mu^*(Q) &= \mu^*(Q \cap A) + \mu^*(Q \cap \overline{A}) = \mu^*(Q \cap A) + \mu^*(Q \cap \overline{A} \cap B) + \mu^*(Q \cap \overline{A} \cap \overline{B}) \\ &\geq \mu^*(Q \cap (A \cup B) + \mu^*(Q \cap \overline{(A \cup B)})). \end{aligned}$$

2. Schritt: Sind $A_1, \dots, A_n \in \mathcal{M}_{\mu^*}$ paarweise disjunkt, so gilt für jedes $Q \subseteq \Omega$

$$\mu^*\Big(Q \cap \bigcup_{i=1}^n A_i\Big) = \sum_{i=1}^n \mu^*(Q \cap A_i).$$

Sind $A, B \in \mathcal{M}_{\mu^*}$ disjunkt, so ist nach Schritt 1 auch $A \cup B \in \mathcal{M}_{\mu^*}$ und es ergibt sich durch Ersetzen von $Q$ durch $Q \cap (A \cup B)$ in (2) $\mu^*(Q \cap (A \cup B)) \geq \mu^*(Q \cap A) + \mu^*(Q \cap B)$. Der Rest folgt per Induktion.

3. Schritt: Für paarweise disjunkte $A_1, A_2, \dots \in \mathcal{M}_{\mu^*}$ ist auch $A = \bigcup_{i=1}^\infty A_i \in \mathcal{M}_{\mu^*}$ und $\mu^*$ ist $\sigma$-additiv.

Wir setzen $B_n = \bigcup_{i=1}^n A_i$. Wegen Schritt 1 ist $B_n \in \mathcal{M}_{\mu^*}$. Mit Schritt 2 erhalten wir also für $Q \subseteq \Omega$

$$\mu^*(Q) = \mu^*(Q \cap B_n) + \mu^*(Q \cap \overline{B_n}) = \sum_{i=1}^n \mu^*(Q \cap A_i) + \mu^*(Q \cap \overline{B_n}) \geq \sum_{i=1}^n \mu^*(Q \cap A_i) + \mu^*(Q \cap \overline{A}).$$

Es folgt

$$\mu^*(Q) \geq \sum_{i=1}^\infty \mu^*(Q \cap A_i) + \mu^*(Q \cap \overline{A}) \geq \mu^*(Q \cap A) + \mu^*(Q \cap \overline{A})$$

und damit $A \in \mathcal{M}_{\mu^*}$. In der letzten Ungleichungskette gilt also Gleichheit. Dies liefert für $Q = A$ gerade

$$\mu^*(A) = \sum_{i=1}^\infty \mu^*(A_i).$$

4. Schritt: $\mathcal{M}_{\mu^*}$ ist eine $\sigma$-Algebra.

Es bleibt nur noch zu zeigen, dass für beliebige $A_1, A_2, \dots \in \mathcal{M}_{\mu^*}$ auch $A = \bigcup_{i=1}^{\infty} A_i \in \mathcal{M}_{\mu^*}$ ist. Dazu schreiben wir wieder mittels des Ansatzes $B_1 = A_1$ und $B_i = A_i \setminus \bigcup_{j=1}^{i-1} A_j$ für $i = 2, 3 \dots$

$$A = \bigcup_{i=1}^{\infty} A_i = \bigcup_{i=1}^{\infty} B_i$$

als disjunkte Vereinigung von Mengen aus $\mathcal{M}_{\mu^*}$. Dann impliziert Schritt 3 die Behauptung.

$\square$

Starten wir mit dem Jordanschen Inhalt $\mu = \lambda^d$ auf den Figuren $\mathcal{A} = \mathcal{F}^d$, so erhalten wir als $\mathcal{M}_{\mu^*}$ die $\sigma$-Algebra der *Lebesgue-messbaren Mengen* mit dem Lebesgueschen Maß. Weiter ist die Borelsche $\sigma$-Algebra eine Teilalgebra der Lebesgue-messbaren Mengen. Wir habe also nun auch das Lebesgue-Maß für Borelmengen konstruiert.

Ist nun $\mu$ sogar $\sigma$-additiv (wie z.B. der Jordansche Inhalt), also ein Prämaß, so ist $\mu^*$ eine Fortsetzung von $\mu$. Insbesondere stimmt also das Lebesguesche Maß für Figuren mit dem Jordan-Inhalt überein. Das ist der Inhalt des folgenden Satzes.

**Satz 1.11** (Fortsetzungssatz von Caratheodory)**.** *Sei $\mu$ ein Prämaß auf einem Ring $\mathcal{A}$ und sei $\mu^*$ das oben definierte zugehörige äußere Maß. Dann ist $\mathcal{A} \subseteq \mathcal{M}_{\mu^*}$ und $\mu^*(A) = \mu(A)$ für $A \in \mathcal{A}$. Insbesondere ist die Einschränkung von $\mu^*$ auf die von $\mathcal{A}$ erzeugte $\sigma$-Algebra ein $\mu$ fortsetzendes Maß.*

*Beweis.* Wir zeigen zunächst $\mu^*(A) = \mu(A)$ für $A \in \mathcal{A}$. Die Ungleichung $\mu^*(A) \le \mu(A)$ folgt sofort aus der Definition von $\mu^*$ mit der Überdeckung $A = A \cup \emptyset \cup \emptyset \dots$. Die Ungleichung $\mu^*(A) \ge \mu(A)$ ist im Fall $\mu^*(A) = \infty$ trivial. Ist nun $A \subseteq \bigcup_{j=1}^{\infty} A_j$ mit $A_j \in \mathcal{A}$, so ist $A = \bigcup_{j=1}^{\infty} (A \cap A_j)$. Da $A \cap A_j \in \mathcal{A}$ ist, folgt

$$\mu(A) \le \sum_{j=1}^{\infty} \mu(A \cap A_j) \le \sum_{j=1}^{\infty} \mu(A_j).$$

Durch Übergang zum Infimum über alle Überdeckungen $A \subseteq \bigcup_{j=1}^{\infty} A_j$ mit $A_j \in \mathcal{A}$ erhalten wir auch in diesem Fall $\mu^*(A) \ge \mu(A)$ und damit schließlich die Gleichheit $\mu^*(A) = \mu(A)$ für $A \in \mathcal{A}$.

Es bleibt zu zeigen, dass jedes $A \in \mathcal{A}$ auch in $\mathcal{M}_{\mu^*}$ ist, also dass die Ungleichung $\mu^*(Q) \ge \mu^*(Q \cap A) + \mu^*(Q \cap \overline{A})$ für alle $Q \subset \Omega$ erfüllt ist. Im Fall $\mu^*(Q) = \infty$ ist wieder nichts zu zeigen. Ist nun $Q \subseteq \bigcup_{j=1}^{\infty} A_j$ mit $A_j \in \mathcal{A}$, so schreiben wir

$$Q \cap A = \bigcup_{j=1}^{\infty} (A_j \cap A) \qquad \text{und} \qquad Q \cap \overline{A} = \bigcup_{j=1}^{\infty} (A_j \cap \overline{A}),$$

wobei alle $A_j \cap A$ und $A_j \cap \overline{A}$ in $\mathcal{A}$ sind. Es folgt

$$\sum_{j=1}^{\infty} \mu(A_j) = \sum_{j=1}^{\infty} \mu(A_j \cap A) + \sum_{j=1}^{\infty} \mu(A_j \cap \overline{A}) \ge \mu^*(Q \cap A) + \mu^*(Q \cap \overline{A}).$$

Durch Übergang zum Infimum über alle Überdeckungen $A \subseteq \bigcup_{j=1}^{\infty} A_j$ mit $A_j \in \mathcal{A}$ erhalten wir schließlich

$$\mu^*(Q) \ge \mu^*(Q \cap A) + \mu^*(Q \cap \overline{A})$$

wie behauptet.

$\square$

Natürlich ist die Eindeutigkeit des konstruierten Maßes $\mu^*$ von Interesse. Diese liefert der folgende Eindeutigkeitssatz, den wir hier ohne Beweis angeben wollen. Beweise finden sich in der angegebenen Literatur, insbesondere in Dirk Werner: Einführung in die höhere Analysis, Abschitt IV.3. Dazu muss das Prämaß $\mu$ auf dem Ring $\mathcal{A}$ zusätzlich *$\sigma$-endlich* sein, d.h. es muss eine aufsteigende Folge $A_1 \subseteq A_2 \subseteq \dots$ von Mengen in $\mathcal{A}$ geben mit $\mu(A_i) < \infty$ für $i = 1, 2, \dots$ und $\Omega = \cup_{i=1}^{\infty} A_i$. Offensichtlich ist der Jordansche Inhalt auf $\mathbb{R}^d$ ein $\sigma$-endliches Prämaß.

**Satz 1.12** (Eindeutigkeitssatz für Maße). *Seinen $\mu$ und $\nu$ zwei Maße auf einer von einem Ring $\mathcal{A}$ erzeugten $\sigma$-Algebra $\sigma(\mathcal{A})$, die auf $\mathcal{A}$ übereinstimmen und dort $\sigma$-endlich sind. Dann stimmen $\mu$ und $\nu$ auf $\sigma(\mathcal{A})$ überein, es gilt also $\mu = \nu$.*

Im folgenden Satz fassen wir die bisherigen Ergebnisse im Fall des Lebesgue-Maßes zusammen.

**Satz 1.13** (Lebesgue-Maß). *Auf der $\sigma$-Algebra $\mathcal{B}(\mathbb{R}^d)$ der Borelmengen existiert genau ein Maß $\lambda^d$, das Intervallen ihren Jordanschen Inhalt zuordnet. Für $A \in \mathcal{B}(\mathbb{R}^d)$ ist dieses gegeben durch*

$$\lambda^d(A) = \inf \left\{ \sum_{j=1}^{\infty} \lambda^d(I_j) \,:\, I_j \in \mathcal{I}^d,\ A \subseteq \bigcup_{j=1}^{\infty} I_j \right\}.$$

**Bemerkung.** $\lambda^d$ heißt *Lebesgue-Maß*. Durch Einschränkung auf Teilmengen einer Menge $S \subseteq \mathbb{R}^d$ erhält man das Lebesgue-Maß auf $S$. Abgeschlossene und offene Intervallen haben das gleiche Lebesgue-Maß wie das zugehörige halboffene Intervall.

**Bemerkung.** Das Lebesguemaß erfüllt für jedes $A \in \mathcal{B}(\mathbb{R}^d)$ die Gleichungen

$$\lambda^d(A) = \inf\{\lambda^d(O) \,:\, A \subseteq O, O \text{ offen}\} = \sup\{\lambda^d(K) \,:\, K \subseteq A, K \text{ kompakt}\},$$

die *innere* bzw. *äußere Regularität* genannt werden. Für den Beweis dieser Eigenschaften verweisen wir hier auf die Literatur.

**Bemerkung.** Es gibt nicht Borel-messbare Teilmengen von $[0, 1]$ (Übung).

Weiterhin ist das Lebesgue-Maß bis auf eine multiplikative Konstante charakterisiert als einziges translationsinvariantes Maß auf den Borelmengen.

**Satz 1.14** (Translationsinvarianz). *Das Lebesgue-Maß ist translationsinvariant, d.h. für $x \in \mathbb{R}^d$ und $A \in \mathcal{B}(\mathbb{R}^d)$ gilt $\lambda^d(x + A) = \lambda^d(A)$. Ist $\mu$ ein translationsinvariantes Maß auf $\mathcal{B}(\mathbb{R}^d)$, so gibt es eine Konstante $c \geq 0$ mit $\mu = c\lambda^d$.*

*Beweis.* Wir zeigen zunächst die Translationsinvarianz von $\lambda^d$. In Satz 1.4 haben wir gezeigt, dass $\mathcal{B}(\mathbb{R}^d)$ translationsinvariant ist. Setzen wir nun $\mu(A) = \lambda^d(x + A)$, so erhalten wir ein Maß auf $\mathcal{B}(\mathbb{R}^d)$ mit $\mu(I) = \lambda^d(I)$ für $I \in \mathcal{I}^d$. Nach Satz 1.13 stimmt dieses mit $\lambda^d$ überein.

Ist $\mu$ ein translationsinvariantes Maß auf $\mathcal{B}(\mathbb{R}^d)$, so setzen wir $c = \mu\Big((0, 1]^d\Big)$. Durch geeignetes Halbieren, Verdoppeln und Verschieben erhalten wir $\mu(I) = c\lambda^d(I)$ für beliebige Intervalle, deren Seitenlängen dyadische Zahlen der Form $N2^m$ mit $N, m \in \mathbb{Z}$ sind. Durch Approximation erhalten wir $\mu(I) = c\lambda^d(I)$ für beliebige Intervalle und damit auch $\mu(A) = c\lambda^d(A)$ für jede Figur $A \in \mathcal{F}^d$. Die Maße $\mu$ und $c\lambda^d$ stimmen also auf dem Ring $\mathcal{F}^d$ und damit nach Satz 1.12 auf der erzeugten $\sigma$-Algebra $\mathcal{B}(\mathbb{R}^d)$ überein. $\square$

## 1.3 Messbare Funktionen

In diesem Abschnitt wollen wir nun Funktionen zwischen Mengen mit $\sigma$-Algebren betrachten, die die Struktur der $\sigma$-Algebra respektieren, sogenannte *messbare Funktionen*. Für die Integrationstheorie werden das dann Funktionen nach $\mathbb{R}$ mit der Borelschen $\sigma$-Algebra sein. In diesem Abschnitt brauchen wir noch kein Maß und definieren deshalb ein Paar $(\Omega, \mathcal{A})$ einer Menge $\Omega$ mit einer $\sigma$-Algebra $\mathcal{A} \subseteq \mathcal{P}(\Omega)$ als einen *messbaren Raum*.

Haben wir nun zwei messbare Räume $(\Omega_1, \mathcal{A}_1)$ und $(\Omega_2, \mathcal{A}_2)$, dann heißt eine Funktion $f : \Omega_1 \to \Omega_2$ *messbar* oder genauer $(\mathcal{A}_1, \mathcal{A}_2)$-*messbar*, wenn das Urbild jeder Menge aus $\mathcal{A}_2$ in $\mathcal{A}_1$ ist:

$$f^{-1}(B) \in \mathcal{A}_1 \qquad \text{für jedes } B \in \mathcal{A}_2.$$

Vergleichen Sie dies mit der Stetigkeitsdefinition von Abbildungen zwischen topologischen Räumen! Wir vereinbaren nun für die Zukunft, dass wir den $\mathbb{R}^d$ und seine Teilmengen immer mit der Borelschen $\sigma$-Algebra versehen.

Es genügt, die Messbarkeit für die Urbilder eines Erzeugers nachzuprüfen (Übung).

**Beispiel.** $f : \Omega \to \mathbb{R}$ ist genau dann messbar, wenn $f^{-1}\big((-\infty, y]\big) = \{x \in \Omega \,:\, f(x) \le y\}$ in $\mathcal{A}$ ist.

**Beispiel.** Stetige Funktionen $f : \mathbb{R}^m \to \mathbb{R}^d$ sind messbar, da Urbilder von offenen Mengen offen sind und die offenen Mengen die Borelsche $\sigma$-Algebra erzeugen.

**Beispiel.** Kompositionen von messbaren Funktionen sind wieder messbar.

**Beispiel.** Eine Funktion $f : \Omega \to \mathbb{R}^d$ ist messbar genau dann, wenn die Komponentenfunktionen $f_i : \Omega \to \mathbb{R}$, $i = 1, \ldots, d$, messbar sind. Einerseits sind die Koordinatenprojektionen $(x_1, \ldots, x_d) \mapsto x_i$ stetig, ist also $f$ messbar, so auch die Komposition von $f$ mit einer Koordinatenprojektion. Das sind aber die $f_i$. Sind andererseits die $f_i$ messbar und ist $I = I_1 \times \cdots \times I_d$ ein Intervall, so ist $f^{-1}(I) = \bigcap_{j=1}^d f_j^{-1}(I_j)$ ebenfalls messbar. Da die Intervalle einen Erzeuger bilden, folgt die Messbarkeit von $f$.

**Beispiel.** Insbesondere sind konstante Abbildungen und charakteristische Funktionen von messbaren Mengen messbar.

**Beispiel.** Sind $f, g : \Omega \to \mathbb{R}$ messbar, dann sind auch $f \pm g, fg, \max(f, g), \min(f, g), f^+ = \max(f, 0), f^- = -\min(f, 0), |f|, cf$ für $c \in \mathbb{R}$ und $1/f$ (falls $f$ keine Nullstellen hat) messbar. Insbesondere bilden die messbaren Funktionen einen linearen Raum.

**Vereinbarung.** Ist $f : \Omega \to \mathbb{R}$ und $B \subset \mathbb{R}$, so schreiben wir kurz $\{f \in B\} := \{x \in \Omega \,:\, f(x) \in B\}$. Entsprechend benutzen wir die Schreibweisen $\{f \le y\}$ für $y \in \mathbb{R}$ oder $\{f = g\}, \{f \le g\}, \ldots$ für zwei Funktionen $f, g : \Omega \to \mathbb{R}$

Um Folgen und Grenzwerte messbarer Funktionen zu betrachten, ist es wieder sinnvoll, die reellen Zahlen $\mathbb{R}$ durch die Elemente $\pm\infty$ zu $\overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$ zu ergänzen. Die Borelsche $\sigma$-Algebra auf $\overline{\mathbb{R}}$ besteht dann aus allen $B \subseteq \overline{\mathbb{R}}$ mit $B \setminus \{\pm\infty\} \in \mathcal{B}(\mathbb{R})$. Diese wird dann von allen Intervallen der Form $[-\infty, y]$ mit $y \in \mathbb{R}$ erzeugt. Nun wird eine Funktion $f : \Omega \to \overline{\mathbb{R}}$ messbar genau dann, wenn $\{-\infty \le f \le y\} \in \mathcal{A}$ für alle $y \in \mathbb{R}$ ist. Mit diesen Begriffsbildungen ist z.B. die Funktion $\sup_n f_n$ punktweise überall definiert, was für das folgende sehr praktisch ist. Wir vereinbaren noch (außer den natürlichen Rechenregeln für $\pm\infty$), dass $0 \cdot \infty = 0$ ist. Der Ausdruck $\infty - \infty$ muss allerdings undefiniert bleiben, weshalb man auch bei der Bildung von $f - g$ für $f, g : \Omega \to \overline{\mathbb{R}}$ Vorsicht walten lassen muss, er ist nicht unbedingt überall definiert. Mit diesen Vereinbarungen erhalten wir weitere wichtige Beispiele.

**Beispiel.** Ist $(f_n)$ eine Folge messbarer Funktionen $f_n : \Omega \to \overline{\mathbb{R}}$, dann sind auch die punktweise definierten Funktionen $\sup f_n, \inf f_n, \limsup f_n, \liminf f_n$ messbar. Insbesondere ist auch $\lim f_n$ messbar, falls dieser Grenzwert punktweise existiert.

Wir beschließen diesen Abschnitt mit der für die Integrationstheorie fundamentalen Möglichkeit, beliebige messbare Funktionen durch Treppenfunktionen zu approximieren.

**Definition 1.15.** Sei $(\Omega, \mathcal{A})$ ein messbarer Raum. Eine messbare Funktion $f : \Omega \to \mathbb{R}$ heißt *Treppenfunktion*, wenn $f$ nur endlich viele Werte annimmt, also in der Form $f = \sum_{j=1}^{n} \lambda_j \chi_{A_j}$ mit messbaren Mengen $A_j \in \mathcal{A}$ darstellen lässt.

**Bemerkung.** Bis auf die Reihenfolge der Summanden gibt es eine eindeutige Darstellung, wenn man verlangt, dass die $\lambda_j$ paarweise verschieden sind. Diese Darstellung wollen wir kanonische Darstellung nennen.

Wir wissen schon, dass Grenzwerte von Treppenfunktionen (da diese messbar sind) wieder messbar sind. Es gilt auch die Umkehrung!

**Satz 1.16.** *Sei $(\Omega, \mathcal{A})$ ein messbarer Raum.*

*(a) Jede messbare Funktion $f : \Omega \to [0, \infty]$ ist punktweiser Grenzwert einer monoton wachsenden Folge nichtnegativer Treppenfunktionen.*

*(b) Jede messbare Funktion $f : \Omega \to \mathbb{R}$ ist punktweiser Grenzwert einer Folge von Treppenfunktionen.*

*(c) Ist $f$ in (a) oder (b) beschränkt, so können die Folgen von Treppenfunktionen so gewählt werden, dass gleichmäßige Konvergenz gilt.*

*Beweis.* Sei zunächst $f : \Omega \to [0, \infty]$ messbar. Für $n \in \mathbb{N}_0$ und $j = \{0, 1, \ldots, 4^n - 1\}$ setzen wir $f_n$ gleich $j2^{-n}$ auf der (messbaren) Menge $A_{j,n} = \{j2^{-n} \leq f < (j+1)2^{-n}\}$ und gleich $2^n$ auf der (ebenfalls messbaren) Menge $\{f \geq 2^n\}$. Dann ist $f_n$ eine Treppenfunktion, $(f_n)$ ist monoton wachsend, und für jedes $x$ gilt $f_n(x) \to f(x)$. Damit ist $(a)$ und $(c)$ (in diesem Fall) gezeigt. $(b)$ folgt durch Anwendung von $(a)$ auf $f^+ = \max(f, 0), f^- = -\min(f, 0)$ in der Darstellung $f = f_+ - f_-$. □

## 1.4 Integrierbare Funktionen und Integrale

Haben wir einen messbaren Raum $(\Omega, \mathcal{A})$ und ein Maß $\mu$ auf diesem messbaren Raum, so nennen wir $(\Omega, \mathcal{A}, \mu)$ einen *Maßraum*. Wir wollen jetzt messbare Funktionen auf diesem Maßraum integrieren. Dies machen wir in drei Schritten. Zunächst führen wir das kanonische Integral für positive Treppenfunktionen ein, definieren dann durch Grenzübergang das Integral für positive messbare Funktionen und schließlich über Differenzen für (fast) beliebige messbare Funktionen.

Sei also zunächst $f = \sum_{j=1}^{n} \lambda_j \chi_{A_j}$ eine nichtnegative Treppenfunktion in ihrer kanonischen Darstellung mit $\lambda_j \in [0, \infty]$. Dann setzen wir für $E \in \mathcal{A}$

$$\int_E f \, \mathrm{d}\mu = \sum_{j=1}^{n} \lambda_j \mu(A_j \cap E) \in [0, \infty].$$

**Beispiel** (Dirichlet-Funktion)**.** Die Dirichlet-Funktion $f : [0,1] \to \mathbb{R}$ ist die charakteristische Funktion der rationalen Zahlen im Intervall $[0,1]$. Diese ist nicht Riemann-integrierbar, da alle Obersummen 1 und alle Untersummen 0 sind. Da $\mathbb{Q}$ eine abzählbare Vereinigung von Einermengen ist, ist $\mathbb{Q}$ und damit auch $[0,1] \cap \mathbb{Q}$ eine Borel-Menge mit Lebesgue-Maß 0. Die Dirichlet-Funktion ist also eine Treppenfunktion mit der kanonischen Darstellung $f = 1 \cdot \chi_{[0,1]\cap\mathbb{Q}} + 0 \cdot \chi_{[0,1]\setminus\mathbb{Q}}$. Damit existiert also das Integral von $f$ bezüglich des Lebesgue-Maßes auf $[0,1]$ und ist

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

Wir erhalten die folgenden wünschenswerten Eigenschaften für ein Integral. Dazu seien $f, g$ zwei Treppenfunktionen wie oben.

(a) Die Abbildung $\nu : E \to \int_E f \, \mathrm{d}\mu$ definiert ein Maß auf $\mathcal{A}$.

(b) $\int_E (f+g) \, \mathrm{d}\mu = \int_E f \, \mathrm{d}\mu + \int_E g \, \mathrm{d}\mu$.

(c) $\int_E \alpha f \, \mathrm{d}\mu = \alpha \int_E f \, \mathrm{d}\mu$ für $\alpha \geq 0$.

(d) $f \geq g$ impliziert $\int_E f \, \mathrm{d}\mu \geq \int_E g \, \mathrm{d}\mu$.

Bei (a) ist nur die $\sigma$-Additivität von $\nu$ zu zeigen. Sei also $(E_i)$ eine Folge disjunkter Mengen in $\mathcal{A}$ und $E = \bigcup_{i=1}^{\infty} E_i$. Dann gilt

$$\nu(E) = \sum_{j=1}^{n} \lambda_j \mu(A_j \cap E) = \sum_{j=1}^{n} \lambda_j \sum_{i=1}^{\infty} \mu(A_j \cap E_i) = \sum_{i=1}^{\infty} \sum_{j=1}^{n} \lambda_j \mu(A_j \cap E_i) = \sum_{i=1}^{\infty} \nu(E_i).$$

Für den Nachweis von (b) stellen wir auch $g = \sum_{k=1}^{m} \mu_k \chi_{B_k}$ kanonisch dar und erhalten per Definition auf den Mengen $E_{jk} = E \cap A_j \cap B_k$

$$\int_{E_{jk}} f \, \mathrm{d}\mu = \lambda_j \mu(E_{jk}), \quad \int_{E_{jk}} g \, \mathrm{d}\mu = \mu_k \mu(E_{jk}), \quad \int_{E_{jk}} (f+g) \, \mathrm{d}\mu = (\lambda_j + \mu_k)\mu(E_{jk}).$$

Summation über alle $j, k$ liefert nun wegen $E = \bigcup_{j,k} E_{jk}$ die Behauptung. Nebenbei haben wir erhalten, dass bei *jeder* Darstellung von $f$ als Treppenfunktion $f = \sum_{j=1}^{n} \lambda_j \chi_{A_j}$ die Formel

$$\int_E f \, \mathrm{d}\mu = \sum_{j=1}^{n} \lambda_j \mu(A_j \cap E)$$

gilt, nicht nur bei der kanonischen.

Eigenschaft (c) ist offensichtlich und Eigenschaft (d) folgt, da $f - g$ eine positive Treppenfunktion ist und deshalb nichtnegatives Integral hat, zusammen mit (b) und $f = (f - g) + g$.

Ist nun $f : \Omega \to [0, \infty]$ eine messbare Funktion und $E \in \mathcal{A}$, dann definieren wir

$$\int_E f \, \mathrm{d}\mu = \sup \left\{ \int_E g \, \mathrm{d}\mu \,:\, g \text{ Treppenfunktion mit } 0 \leq g \leq f \right\} \in [0, \infty].$$

Die Monotonie des Integrals für Treppenfunktionen impliziert, dass wir hier im Fall einer Treppenfunktion das gleiche Integral erhalten wie vorher. Wieder sammeln wir ein paar offensichtliche und natürliche Eigenschaften für messbare $f, g : \Omega \to [0, \infty]$.

(a) Gilt $0 \leq f \leq g$ auf $E$, so ist $\int_E f \, \mathrm{d}\mu \leq \int_E g \, \mathrm{d}\mu$.

(b) Sind $E_1, E_2 \in \mathcal{A}$ mit $E_1 \subseteq E_2$, so gilt $\int_{E_1} f \, \mathrm{d}\mu \leq \int_{E_2} f \, \mathrm{d}\mu$.

(c) Verschwindet $f$ auf $E$, so ist $\int_E f \, \mathrm{d}\mu = 0$.

(d) Ist $\mu(E) = 0$, so gilt $\int_E f \, \mathrm{d}\mu = 0$ (auch falls $f$ auf $E$ unendlich ist).

An dieser Stelle können wir bereits einen wichtigen Konvergenzsatz für Integrale - den Satz von Beppo Levi von der monotonen Konvergenz - zeigen, weitere folgen im nächsten Abschnitt.

**Satz 1.17.** *Seien $f_n : \Omega \to [0, \infty]$ für $n \in \mathbb{N}$ messbare Funktionen und sei $f_1 \leq f_2 \leq$ monoton wachsend mit $\lim_{n\to\infty} f_n = f$ punktweise. Dann ist $f$ messbar und es gilt*

$$\lim_{n\to\infty} \int_E f_n \, \mathrm{d}\mu = \int_E f \, \mathrm{d}\mu \qquad \text{für alle } E \in \mathcal{A}.$$

*Beweis.* Die Messbarkeit des punktweisen Grenzwerts $f$ haben wir bereits im letzten Abschnitt eingesehen. Die Monotonie des Integrals zeigt, dass die Zahlenfolge $(\int_E f_n \, \mathrm{d}\mu)$ monoton wächst und durch $\int_E f \, \mathrm{d}\mu$ nach oben beschränkt ist. Deshalb hat sie einen Grenzwert

$$a = \lim_{n\to\infty} \int_E f_n \, \mathrm{d}\mu = \sup_{n\to\infty} \int_E f_n \, \mathrm{d}\mu \in [0, \infty]$$

mit $a \leq \int_E f \, \mathrm{d}\mu$. Wir müssen also noch $a \geq \int_E f \, \mathrm{d}\mu$ zeigen. Nach Definition des Integrals reicht es also $\int_E g \, \mathrm{d}\mu \leq a$ für alle Treppenfunktionen $g$ mit $0 \leq g \leq f$ zu zeigen.
Dazu betrachten wir für $c \in [0, 1)$ die Mengen $E_n = \{f_n \geq cg\} \cap E$. Diese sind messbar. Die Monotonie der $f_n$ impliziert $E_1 \subseteq E_2 \subseteq \ldots$. Außerdem ist wegen $f_n \nearrow f$ und $g \leq f$ auch $\lim E_n = \bigcup E_n = E$. Aus den Eigenschaften des Integrals folgt

$$a \geq \int_E f_n \, \mathrm{d}\mu \geq \int_{E_n} f_n \, \mathrm{d}\mu \geq \int_{E_n} cg \, \mathrm{d}\mu = c \int_{E_n} g \, \mathrm{d}\mu$$

für jedes $n \in \mathbb{N}$. Durch Grenzübergang $c \to 1$ erhalten wir $a \geq \int_{E_n} g \, \mathrm{d}\mu$ und schließlich durch die $\sigma$-Additivität der Abbildung $A \to \int_A g \, \mathrm{d}\mu$ die Behauptung $a \geq \int_E g \, \mathrm{d}\mu$. □

Dieser Konvergenzsatz legt ein Beweisverfahren nahe, wie es typisch für die Integrationstheorie ist. Wollen wir eine Eigenschaft des Integrals für beliebige messbare Funktionen zeigen, dann zeigen wir sie zunächst für Indikatorfunktionen (das ist meist einfach), dann übertragen wir sie auf Treppenfunktionen (durch Additivität) und zeigen sie schließlich per Grenzübergang für allgemeine Funktionen. Wir wollen dies anwenden, um die folgenden Eigenschaften des Integrals für messbare Funktionen $f, g : \Omega \to [0, \infty]$ einzusehen.

(a) $\int_E f \, \mathrm{d}\mu = \int_\Omega \chi_E f \, \mathrm{d}\mu$ für alle $E \in \mathcal{A}$.

(b) $\int_E (f + g) \, \mathrm{d}\mu = \int_E f \, \mathrm{d}\mu + \int_E g \, \mathrm{d}\mu$ für alle $E \in \mathcal{A}$.

(c) $\int_E cf \, \mathrm{d}\mu = c \int_E f \, \mathrm{d}\mu$ für alle $E \in \mathcal{A}$ und $c \geq 0$.

*Beweis.* (a): Für Treppenfunktionen folgt die Aussage aus der Definition. Nach Satz 1.16 kann $f$ monoton approximiert werden durch Treppenfunktionen $f_n$, also $f_n \nearrow f$. Also gilt auch $\chi_E f_n \nearrow \chi_E f$ und die Behauptung folgt mit Grenzübergang aus dem Satz von der monotonen Konvergenz.

(b): Wegen (a) genügt es, die Behauptung für $E = \Omega$ zu zeigen. Für Treppenfunktionen $f, g$ haben wir die Behauptung bereits eingesehen. Wähle nun Treppenfunktionen $f_n \nearrow f$ und $g_n \nearrow g$. Dann gilt $f_n + g_n \nearrow f + g$ und die Behauptung folgt durch Grenzübergang aus dem Satz von der monotonen Konvergenz.

(c): geht analog. □

Die Additivität des Integrals (b) liefert die folgende alternative Formulierung des Satzes von der monotonen Konvergenz.

**Satz 1.18.** *Seien $f_n : \Omega \to [0, \infty]$ für $n \in \mathbb{N}$ messbare Funktionen und sei $f = \sum_{n=1}^{\infty} f_n$. Dann ist $f$ messbar und es gilt*

$$\int_E f \,\mathrm{d}\mu = \sum_{n=1}^{\infty} \int_E f_n \,\mathrm{d}\mu \qquad \text{für alle } E \in \mathcal{A}.$$

Schließlich führen wir noch das Integral für nicht notwendig positive Funktionen ein, indem wir diese in positiven und negativen Anteil aufteilen.

**Definition 1.19.** Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum, sei $f : \Omega \to [-\infty, \infty]$ messbar und seien $f_+, f_- : \Omega \to [0, \infty]$ positiver und negativer Teil von $f$ mit $f = f_+ - f_-$. Sind $\int_\Omega f_+ \,\mathrm{d}\mu$ und $\int_\Omega f_- \,\mathrm{d}\mu$ endlich, so heißt $f$ *integrierbar* (genauer *$\mu$-integrierbar*) und wir setzen

$$\int_\Omega f \,\mathrm{d}\mu = \int_\Omega f_+ \,\mathrm{d}\mu - \int_\Omega f_- \,\mathrm{d}\mu.$$

Für nichtnegative Funktionen mit endlichem Integral stimmt natürlich die neue Definition des Integrals mit der alten überein. Aus $0 \le f_+, f_- \le |f| \le f_+ + f_-$ folgt

$$f \text{ ist integrierbar} \qquad \Longleftrightarrow \qquad \int_\Omega |f| \,\mathrm{d}\mu < \infty.$$

Integrierbarkeit über eine messbare Teilmenge $A \in \mathcal{A}$ kann man analog einführen oder alternativ über

$$\int_A f \,\mathrm{d}\mu = \int_\Omega f \,\chi_A \,\mathrm{d}\mu.$$

Es gibt in der Literatur verschiedene Schreibweisen für das Integral:

$$\int_A f \,\mathrm{d}\mu = \int_A f(x) \,\mathrm{d}\mu(x) = \int_A f(x) \mu(\,\mathrm{d}x).$$

Für Lebesgue-Integrale, also Integrale bezüglich des Lebesgue-Maßes auf $\mathbb{R}^d$, schreiben wir einfach $\int_A f(x) \,\mathrm{d}x$.

Für die folgende Formulierung wichtiger Eigenschaften des Integrals seien $f, g$ integrierbar und $\alpha, \beta$ reelle Zahlen.

(a) **Linearität**: $\alpha f + \beta g$ ist integrierbar mit $\int_\Omega (\alpha f + \beta g) \,\mathrm{d}\mu = \alpha \int_\Omega f \,\mathrm{d}\mu + \beta \int_\Omega g \,\mathrm{d}\mu$. Insbesondere bilden die integrierbaren Funktionen einen linearen Raum und $f \to \int_\Omega f \,\mathrm{d}\mu$ ist ein lineares Funktional auf diesem Raum.

(b) **Monotonie**: $f \ge g$ impliziert $\int_\Omega f \,\mathrm{d}\mu \ge \int_\Omega g \,\mathrm{d}\mu$

(c) **Dreiecksungleichung**: $\left|\int_\Omega f \, \mathrm{d}\mu\right| \leq \int_\Omega |f| \, \mathrm{d}\mu$

(d) Sind $A, B \in \mathcal{A}$ disjunkt, so gilt $\int_{A\cup B} f \, \mathrm{d}\mu = \int_A f \, \mathrm{d}\mu + \int_B f \, \mathrm{d}\mu$.

*Beweis.* Für die Monotonie in (b) reicht es, den Fall $g = 0$ zu betrachten. Dieser wurde bereits weiter oben erledigt. Die Dreiecksungleichung (c) folgt dann wegen $\pm f \leq |f|$. Die Eigenschaft (d) folgt aus der Linearität in (a) und $\int_A f \, \mathrm{d}\mu = \int_\Omega f \, \chi_A \, \mathrm{d}\mu$.

Zum Beweis von (a) bemerken wir zunächst, dass aus der Integrierbarkeit von $f, g$ die Messbarkeit von $f, g$ folgt. Damit ist auch $\alpha f + \beta g$ messbar. Weiter sind $\int_\Omega |f| \, \mathrm{d}\mu, \int_\Omega |g| \, \mathrm{d}\mu < \infty$ und die Integrierbarkeit von $\alpha f + \beta g$ folgt aus

$$\int_\Omega \left|\alpha f + \beta g\right| \mathrm{d}\mu \leq |\alpha| \int_\Omega |f| \, \mathrm{d}\mu + |\beta| \int_\Omega |f| \, \mathrm{d}\mu < \infty.$$

Für den Beweis der Linearitätsformel können wir uns auf die Fälle $\alpha = \beta = 1$ und $\alpha \neq 0, \beta = 0$ beschränken. Zur Abkürzung schreiben wir $\int f$ für $\int_\Omega f \, \mathrm{d}\mu$.

Im Fall $\alpha = \beta = 1$ benutzen wir die Identität $f_+ + g_+ + (f+g)_- = f_- + g_- + (f+g)_+$ um mit der Linearität des Integrals für positive Funktionen

$$\int f_+ + \int g_+ + \int (f+g)_- = \int f_- + \int g_- + \int (f+g)_+$$

zu folgern. Umstellen und Definition des Integrals liefert dann $\int f + \int g = \int (f+g)$.

Den Fall $\alpha \geq 0, f \geq 0$ haben wir schon erledigt. Für den allgemeinen Fall $\alpha \geq 0$ benutzen wir die Identität $f_+ = f + f_-$ und erhalten mit der Linearität des Integrals für positive Funktionen

$$\alpha \int f_+ = \int \alpha f_+ = \int \alpha(f + f_-) = \int \alpha f + \alpha \int f_-,$$

woraus dann

$$\alpha \int f = \alpha \int f_+ - \alpha \int f_- = \int \alpha f$$

folgt.

Ist nun $\alpha = -1$, so folgt die Behauptung aus dem bereits gezeigten wegen

$$0 = \int \Big(f + (-f)\Big) = \int f + \int (-f).$$

Ist schließlich $\alpha < 0$ beliebig, erhalten wir die Behauptung aus

$$\int \alpha f = \int (-\alpha)(-f) = -\alpha \int (-f) = \alpha \int f.$$

□

**Beispiel** (Dirac-Maß)**.** Sei $\mathcal{A}$ eine $\sigma$-Algebra, $x \in \Omega$ fixiert und $\delta_x$ das Dirac-Maß definiert durch $\delta_x(A) = \chi_A(x)$. Dann gilt $\int_\Omega f \, \mathrm{d}\delta_x = f(x)$. Dies zeigt man nacheinander für Indikatorfunktionen, Treppenfunktionen, nichtnegative messbare Funktionen und allgemeine messbare Funktionen.

**Beispiel** (Zählmaß - Summen als Integrale)**.** Wir betrachten nun den Maßraum $(\mathbb{N}, \mathcal{P}(\mathbb{N}), \mu)$ mit dem Zählmaß $\mu$. Also ist $\mu(A)$ die Anzahl der Elemente von $A$, wenn $A$ endlich ist und $\mu(A) = \infty$ für unendliches $A$. Eine Funktion $a : \mathbb{N} \to \mathbb{R}$ identifizieren wir wie üblich mit einer Folge $(a_n)$. Offensichtlich ist jede solche Funktion messbar. Für die spezielle Treppenfunktion $\chi_{\{n\}}$ ergibt sich aus der Definition für Integrale von Treppenfunktionen $\int_{\mathbb{N}} \chi_{\{n\}}\,\mathrm{d}\mu = 1$. Für nichtnegative Folgen $a = (a_n)$ erhalten wir also

$$\int_{\mathbb{N}} a\,\mathrm{d}\mu = \sum_{n=1}^{\infty} a_n.$$

Eine beliebige Folge $a = (a_n)$ reeller Zahlen ist integrierbar genau dann, wenn

$$\int_{\mathbb{N}} |a|\,\mathrm{d}\mu = \sum_{n=1}^{\infty} |a_n| < \infty$$

ist, also die Reihe $\sum_{n=1}^{\infty} a_n$ absolut konvergiert. In diesem Fall ist ebenfalls $\int_{\mathbb{N}} a\,\mathrm{d}\mu = \sum_{n=1}^{\infty} a_n$.

**Beispiel** (Lebesgue-Integral und Riemann-Integral)**.** Für stetige Funktionen $f : [a,b] \to \mathbb{R}$ stimmen Riemann- und Lebesgue-Integral überein. Um dies einzusehen, können wir uns auf nichtnegative $f$ beschränken. Da $f$ gleichmäßig stetig ist, gibt es nichtnegative Treppenfunktionen $f_n$ mit $f_n \nearrow f$ gleichmäßig. Hier können wir für die Stufen einfach Intervalle nehmen. Dann stimmen Riemann- und Lebesgue-Integral für jedes $f_n$ überein. Aus der gleichmäßigen Konvergenz $f_n \to f$ folgt nun, dass das Riemann-Integral von $f$ Grenzwert der Riemann-Integrale der $f_n$ ist. Aus dem Satz von der monotonen Konvergenz folgt, dass das Lebesgue-Integral von $f$ Grenzwert der Lebesgue-Integrale der $f_n$ ist. Da auch diese Grenzwerte übereinstimmen, sind Riemann- und Lebesgue-Integral von $f$ gleich.

Es gibt eine Charakterisierung von Lebesgue für beschränkte Riemann-integrierbare Funktionen, aus der folgt, dass jede solche Funktion auch Lebesgue-integrierbar ist und die Integrale übereinstimmen. Diese Charakterisierung lautet: $f$ ist Riemann-integrierbar genau dann, wenn die Menge der Unstetigkeitsstellen von $f$ eine Lebesguesche Nullmenge ist. Wir wollen diese Charakterisierung hier nicht beweisen.

Im Falle von uneigentlichen Integralen ist allerdings die Lebesguesche Integrationstheorie schwerfälliger als die Riemannsche. So existiert z.B. das uneigentliche Riemann-Integral $\int_0^\infty \frac{\sin x}{x}\,\mathrm{d}x$, das Lebesgue-Integral hingegen nicht, da $\int_0^\infty \left|\frac{\sin x}{x}\right|\mathrm{d}x = \infty$ ist. Natürlich kann man sich auch hier behelfen, indem man $\lim_{b\to\infty} \int_0^b \frac{\sin x}{x}\,\mathrm{d}x$ betrachtet.

**Beispiel** (Integration bezüglich Dichten)**.** Sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum. Wir haben schon eingesehen, dass für jede positive Treppenfunktion $g : \Omega \to \mathbb{R}$ die Abbildung $\nu : A \to \int_A g\,\mathrm{d}\mu$ ein Maß auf $\mathcal{A}$ definiert. Dies ist ebenfalls richtig für beliebiges positives messbares $g$, da dann für paarweise disjunkte Mengen $A_j$ und $A = \bigcup_j A_j$ aus der punktweisen Konvergenz der Reihe $\chi_A = \sum_j \chi_{A_j}$ und Satz 1.18 die $\sigma$-Additivität $\nu(A) = \sum_j \nu(A_j)$ folgt. Eine messbare Funktion $f$ ist genau dann $\nu$-integrierbar, wenn $\int_\Omega |f| g\,\mathrm{d}\mu < \infty$ ist. Man zeigt wieder nacheinander für Indikatorfunktionen, Treppenfunktionen, positive messbare Funktionen und beliebige messbare Funktionen (mit $\int_\Omega |f| g\,\mathrm{d}\mu < \infty$) die Gleichheit

$$\int_\Omega f\,\mathrm{d}\nu = \int_\Omega f g\,\mathrm{d}\mu.$$

Man spricht hier von einer *Integrationsdichte* $g$ und der Integration bezüglich des Maßes $g\,\mathrm{d}\mu$ mit der Dichte $g$. Der Satz von Radon-Nikodým, den wir später zeigen, charakterisiert diese Maße mit Dichten.

## 1.5 Konvergenzsätze

Neben dem Satz von Beppo-Levi über die monotone Konvergenz sind zwei weitere Konvergenzsätze von großer Bedeutung für die Integrationstheorie, das Lemma von Fatou und der Satz von Lebesgue von der dominierten Konvergenz. In diesem Abschnitt sei wieder ein Maßraum $(\Omega, \mathcal{A}, \mu)$ fixiert.

**Satz 1.20** (Lemma von Fatou). *Für messbare Funktionen $f_n : \Omega \to [0, \infty]$ gilt*

$$\int_\Omega \liminf_{n\to\infty} f_n \,\mathrm{d}\mu \le \liminf_{n\to\infty} \int_\Omega f_n \,\mathrm{d}\mu.$$

**Satz 1.21** (Satz von der dominierten Konvergenz). *Seien $f_n : \Omega \to \mathbb{R}$ messbar mit punktweisem Grenzwert $\lim_{n\to\infty} f_n = f$. Falls es eine integrierbare Funktion $g : \Omega \to [0, \infty]$ mit $|f_n| \le g$ für alle $n \in \mathbb{N}$ gibt, dann sind auch die Funktionen $f_n$ und $f$ integrierbar und es gilt*

$$\int_\Omega f \,\mathrm{d}\mu = \lim_{n\to\infty} \int_\Omega f_n \,\mathrm{d}\mu \qquad \textit{und} \qquad \lim_{n\to\infty} \int_\Omega |f_n - f| \,\mathrm{d}\mu = 0.$$

*Beweis von Satz 1.20.* Die Messbarkeit von $\liminf_{n\to\infty} f_n$ haben wir bereits eingesehen. Die Behauptung folgt dann mit dem Satz von der monotonen Konvergenz aus

$$\int_\Omega \liminf_{n\to\infty} f_n \,\mathrm{d}\mu = \int_\Omega \sup_k \inf_{n\ge k} f_n \,\mathrm{d}\mu = \sup_k \int_\Omega \inf_{n\ge k} f_n \,\mathrm{d}\mu \le \sup_k \inf_{n\ge k} \int_\Omega f_n \,\mathrm{d}\mu = \liminf_{n\to\infty} \int_\Omega f_n \,\mathrm{d}\mu.$$

□

*Beweis von Satz 1.21.* Die Messbarkeit von $f$ haben wir bereits eingesehen, die Integrierbarkeit von $f_n$ und $f$ folgt dann aus $|f_n| \le g$ und $|f| \le g$. Die Funktion $h = g + |f|$ ist dann ebenfalls intergrierbar. Die Dreiecksungleichung zeigt $|f - f_n| \le h$. Das Lemma von Fatou liefert dann

$$\int_\Omega h \,\mathrm{d}\mu = \int_\Omega \lim_{n\to\infty} \big(h - |f_n - f|\big) \,\mathrm{d}\mu \le \liminf_{n\to\infty} \int_\Omega \big(h - |f_n - f|\big) \,\mathrm{d}\mu = \int_\Omega h \,\mathrm{d}\mu - \limsup_{n\to\infty} \int_\Omega |f_n - f| \,\mathrm{d}\mu$$

und damit $\limsup_{n\to\infty} \int_\Omega |f_n - f| \,\mathrm{d}\mu \le 0$. Da $\liminf_{n\to\infty} \int_\Omega |f_n - f| \,\mathrm{d}\mu \ge 0$ offensichtlich ist, folgt $\lim_{n\to\infty} \int_\Omega |f_n - f| \,\mathrm{d}\mu = 0$. Schließlich folgt

$$\left| \int_\Omega f \,\mathrm{d}\mu - \int_\Omega f_n \,\mathrm{d}\mu \right| \le \int_\Omega \big|f - f_n\big| \,\mathrm{d}\mu \to 0.$$

□

Man nennt den Satz von Lebesgue auch Satz von der majorisierten Konvergenz wegen der integrierbaren *Majorante* $g$ in der Voraussetzung. Auf die Existenz der integrierbaren Majorante kann man nicht verzichten (Übung). Für einen endlichen Maßraum $\Omega$, also $\mu(\Omega) < \infty$, ist die Voraussetzung des Satzes von Lebesgue insbesondere erfüllt, wenn die Funktionen $f_n$ gleichmäßig beschränkt sind. In diesem Fall kann man einfach $g = M$ mit einem $M \ge |f_n|$ für alle $n \in \mathbb{N}$ wählen.

Eine schöne Folgerung ist der folgende Satz. Ein Beweis mittels Riemannscher Integrationstheorie wäre sehr aufwendig.

**Korollar 1.22** (Arzelà-Osgood). *Sei $f_n : [0, 1] \to \mathbb{R}$ für $n \in \mathbb{N}$. Ist die Folge $(f_n)$ gleichmäßig beschränkt und konvergiert punktweise gegen 0, dann konvergiert auch die Folge der Integrale $\int_0^1 f_n(x) \,\mathrm{d}x$ gegen 0.*

Eine weitere Folgerung ist das folgende allgemeine Kriterium zur Vertauschung von Differentiation und Integral für Parameterintegrale.

**Satz 1.23.** *Die Funktion $f : \mathbb{R}^{d+1} \to \mathbb{R}$ sei stetig differenzierbar und für alle $t \in \mathbb{R}$ sei die Funktion $\mathbb{R}^d \to \mathbb{R}, x \mapsto f(t,x)$ integrierbar. Außerdem existiere für jedes $t_0 \in \mathbb{R}$ eine Umgebung $U$ von $t_0$ und eine integrierbare Funktion $g : \mathbb{R}^d \to \mathbb{R}$, so dass die Abschätzung*

$$\left|\frac{\partial f}{\partial t}(t,x)\right| \leq g(x) \qquad \text{für } t \in U \text{ und } x \in \mathbb{R}^d$$

*gilt. Dann gilt für alle $t \in \mathbb{R}$*

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^d} f(t,x)\,\mathrm{d}x = \int_{\mathbb{R}^d} \frac{\partial f}{\partial t}(t,x)\,\mathrm{d}x.$$

*Beweis.* Für gegebenes $t_0$ sei $(t_n)$ eine Folge in der Umgebung $U$ von $t_0$ aus der Voraussetzung mit $t_n \to t_0$. Dann gilt für den Differenzenquotienten

$$\frac{\int_{\mathbb{R}^d} f(t_n,x)\,\mathrm{d}x - \int_{\mathbb{R}^d} f(t_0,x)\,\mathrm{d}x}{t_n - t_0} = \int_{\mathbb{R}^d} \frac{f(t_n,x) - f(t_0,x)}{t_n - t_0}\,\mathrm{d}x.$$

Die Integranden $h_n(x) = \frac{f(t_n,x)-f(t_0,x)}{t_n-t_0}$ im rechten Integral konvergieren punktweise gegen $h(x) = \frac{\partial f}{\partial t}(t_0,x)$. Außerdem gilt nach dem Mittelwertsatz $h_n(x) = \frac{\partial f}{\partial t}(t,x)$ für ein $t$ zwischen $t_0$ und $t_n$. Dann ist $t \in U$ und nach Voraussetzung gilt $|h_n(x)| \leq g(x)$. Es gibt also eine integrierbare Majorante für die Funktionen $h_n$. Aus dem Satz von der dominierten Konvergenz folgt dann

$$\lim_{n\to\infty} \frac{\int_{\mathbb{R}^d} f(t_n,x)\,\mathrm{d}x - \int_{\mathbb{R}^d} f(t_0,x)\,\mathrm{d}x}{t_n - t_0} = \int_{\mathbb{R}^d} h(x)\,\mathrm{d}x = \int_{\mathbb{R}^d} \frac{\partial f}{\partial t}(t_0,x)\,\mathrm{d}x.$$

Da dies für jede Folge $t_n \to t_0$ (mit $t_n \in U$) gilt, folgt die Behauptung.

□

Eine *Nullmenge* bezüglich eines Maßes $\mu$ ist eine Menge die Teilmenge einer Menge mit Maß 0 ist. Man sagt dann, dass eine Eigenschaft für Elemente eines Maßraumes *fast überall*, kurz *f.ü.* oder *$\mu$-f.ü.* erfüllt ist, wenn die Ausnahmemenge, auf der sie nicht erfüllt ist, eine Nullmenge ist. Man beachte, dass die Ausnahmemenge selbst nicht messbar sein muss. Zum Beispiel ist $\lambda$-fast jede Zahl irrational, die Funktionenfolge $(t^n)$ konvergiert $\lambda$-f.ü. gegen 0.

**Beispiel** (Teufelstreppe)**.** Cantorfunktion, f.ü. konstant, surjektiv auf [0,1], in Vorlesung

Wir beschließen den Abschnitt mit einer letzten Anwendung über den Zusammenhang zwischen punktweiser und gleichmäßiger Konvergenz.

**Satz 1.24** (Satz von Egorov)**.** *Ist $(\Omega, \mathcal{A}, \mu)$ ein endlicher Maßraum und $(f_n)$ eine Folge messbarer Funktionen, die f.ü. gegen 0 konvergiert, dann existiert zu jedem $\varepsilon > 0$ eine Menge $A \in \mathcal{A}$ mit $\mu(\overline{A}) < \varepsilon$, auf der $(f_n)$ gleichmäßig gegen 0 konvergiert.*

*Beweis.* Nach Voraussetzung gilt für jedes $\delta > 0$

$$\lim_{m\to\infty} \mu\left(\bigcap_{n\geq m} \{|f_n| \leq \delta\}\right) = \mu\left(\bigcup_{m=1}^{\infty}\bigcap_{n\geq m} \{|f_n| \leq \delta\}\right) = \mu(\Omega).$$

Insbesondere existiert für jedes $k \in \mathbb{N}$ ein $m_k \in \mathbb{N}$ mit

$$\mu(A_k) \geq \mu(\Omega) - \varepsilon 2^{-k} \qquad \text{für die Menge } A_k := \bigcap_{n \geq m_k} \big\{|f_n| \leq 1/k\big\}.$$

Setzen wir nun $A = \bigcap_{k=1}^{\infty} A_k$, so erhalten wir

$$\mu(\overline{A}) \leq \sum_{k=1}^{\infty} \mu(\overline{A_k}) \leq \varepsilon \sum_{k=1}^{\infty} 2^{-k} = \varepsilon.$$

Außerdem gilt für $x \in A$ und $n \geq m_k$ die Ungleichung $|f_n(x)| \leq 1/k$, die Konvergenz $f_n \to 0$ ist also gleichmäßig auf $A$. □

## 1.6 Produktmaße und Satz von Fubini

In diesem Abschnitt soll es um die Vertauschbarkeit der Integration bei iterierten Integralen gehen, Ziel ist der Satz von Fubini. Um diesen ebenfalls in voller Allgemeinheit abhandeln zu können, benötigen wir zunächst Produkt-$\sigma$-Algebren und Produktmaße.

**Definition 1.25.** Sind $(\Omega_i, \mathcal{A}_i)$ für $i = 1, 2$ messbare Räume, so ist die *Produkt-$\sigma$-Algebra* $\mathcal{A}_1 \otimes \mathcal{A}_2$ die von den Rechtecken $A_1 \times A_2$ mit $A_i \in \mathcal{A}_i$ erzeugte $\sigma$-Algebra auf $\Omega_1 \times \Omega_2$.

Sei $(\Omega, \mathcal{A})$ ein dritter messbarer Raum. Wann ist eine Funktion $f : \Omega \to \Omega_1 \times \Omega_2$ messbar , genauer $(\mathcal{A}, \mathcal{A}_1 \otimes \mathcal{A}_2)$-messbar? Dazu betrachten wir die Koordinatenprojektionen $p_i : \Omega_1 \times \Omega_2 \to \Omega_i$ gegeben durch $p_i(x_1, x_2) = x_i$ und die Koordinatenfunktionen $f_i = p_i \circ f$. Dann gilt

$$f \text{ ist messbar} \quad \iff \quad f_i \text{ ist messbar für } i = 1, 2. \tag{3}$$

*Beweis.* Da $p_1^{-1}(A_1) = A_1 \times \Omega_2$ für $A_1 \in \mathcal{A}_1$ messbar ist, ist $p_1$ messbar. Ebenso ist $p_2$ messbar. Ist also $f$ messbar, so sind auch die $f_i = p_i \circ f$ als Komposition messbarer Funktionen messbar. Ist umgekehrt $f_i$ messbar und $A_i \in \mathcal{A}_i$ für $i = 1, 2$, so erhalten wir

$$f^{-1}(A_1 \times A_2) = f_1^{-1}(A_1) \cap f_2^{-1}(A_2) \in \mathcal{A}.$$

Da die $A_1 \times A_2$ per definitionem einen Erzeuger bilden, ist $f$ messbar. □

**Beispiel.** Die Produkt-$\sigma$-Algebra $\mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R})$ ist die Borelsche $\sigma$-Algebra $\mathcal{B}(\mathbb{R}^2)$ auf $\mathbb{R}^2$.

Nach Definition von $\mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R})$ sind nämlich die halboffenen Intervalle $I = I_1 \times I_2$ mit $I_1, I_2 \in \mathcal{I}$ in $\mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R})$. Da diese $\mathcal{B}(\mathbb{R}^2)$ erzeugen, folgt $\mathcal{B}(\mathbb{R}^2) \subseteq \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R})$. Andererseits folgt aus (3) und der Stetigkeit (und damit $(\mathcal{B}(\mathbb{R}^2), \mathcal{B}(\mathbb{R}))$-Messbarkeit) der Koordinatenprojektionen $p_i : \mathbb{R}^2 \to \mathbb{R}$ die $(\mathcal{B}(\mathbb{R}^2), \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R}))$-Messbarkeit der Identität auf $\mathbb{R}^2$. Das bedeutet aber nichts anderes als $\mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R}) \subseteq \mathcal{B}(\mathbb{R}^2)$.

Analog zeigt man

$$\mathcal{B}(\mathbb{R}^{d_1}) \otimes \mathcal{B}(\mathbb{R}^{d_2}) = \mathcal{B}(\mathbb{R}^{d_1 + d_2}).$$

Zu einer gegebenen Menge $A \subseteq \Omega_1 \times \Omega_2$ betrachten wir für festes $x_1 \in \Omega_1$ den Schnitt

$$A_{x_1} = \{x_2 \in \Omega_2 \, : \, (x_1, x_2) \in A\} \subseteq \Omega_2.$$

Nach (3) ist die Funktion $\varphi : x_2 \mapsto (x_1, x_2)$ messbar. Ist dann $A \in \mathcal{A}_1 \otimes \mathcal{A}_2$, so ist auch jeder Schnitt $A_{x_1} = \varphi^{-1}(A)$ messbar, also $A_{x_1} \in \mathcal{A}_2$. Analog definieren wir für festes $x_2 \in \Omega_2$ den Schnitt

$$A^{x_2} = \{x_1 \in \Omega_1 \,:\, (x_1, x_2) \in A\} \subseteq \Omega_1.$$

Es folgt $A^{x_2} \in \mathcal{A}_1$ für $A \in \mathcal{A}_1 \otimes \mathcal{A}_2$.

Unser nächstes Ziel ist nun, für gegebene Maßräume $(\Omega_i, \mathcal{A}_i, \mu_i)$ für $i = 1, 2$ ein *Produktmaß* $\mu$ auf $\mathcal{A}_1 \otimes \mathcal{A}_2$ zu finden mit

$$\mu(A_1 \times A_2) = \mu_1(A_1)\mu_2(A_2) \qquad \text{für } A_i \in \mathcal{A}_i, i = 1, 2.$$

Als Vorbereitung benötigen wir das

**Lemma 1.26.** *Sind $(\Omega_i, \mathcal{A}_i, \mu_i)$ für $i = 1, 2$ $\sigma$-endliche Maßräume, dann sind die Funktionen $x_1 \mapsto \mu_2(A_{x_1})$ und $x_2 \mapsto \mu_1(A^{x_2})$ für jedes $A \in A_1 \times A_2$ messbar (bezüglich $\mathcal{A}_1$ bzw. $\mathcal{A}_2$).*

*Beweisskizze.* Es genügt, die Behauptung für $x_1 \mapsto \mu_2(A_{x_1})$ zu zeigen. Sei dazu zunächst $\mu_2(\Omega_2) < \infty$. Dann betrachten wir das Mengensystem

$$\mathcal{A} := \{A \in A_1 \times A_2 \,:\, x_1 \mapsto \mu_2(A_{x_1}) \text{ ist } \mathcal{A}_1 - \text{messbar}\}\,.$$

Offensichtlich sind die Rechtecke $A_1 \times A_2$ mit $A_i \in \mathcal{A}_i$ in $\mathcal{A}$.

Wir zeigen, dass $\mathcal{A}$ eine $\sigma$-Algebra ist. Dann folgt wie behauptet $\mathcal{A} = A_1 \otimes A_2$. Für $A = \emptyset$ und $\mathcal{A} = \Omega_1 \times \Omega_2$ ist $x_1 \mapsto \mu_2(A_{x_1})$ eine konstante Funktion und damit messbar. Also haben wir $\emptyset, \Omega_1 \times \Omega_2 \in \mathcal{A}$. Wegen $\mu_2\big(\overline{A_{x_1}}\big) = \mu_2(\Omega_2) - \mu_2\big(A_{x_1}\big)$ ist $x_1 \mapsto \mu_2\big(\overline{A_{x_1}}\big)$ messbar, falls $x_1 \mapsto \mu_2(A_{x_1})$ messbar ist. Aus $A \in \mathcal{A}$ folgt also $\overline{A} \in \mathcal{A}$. Seien nun $A_1, A_2, \dots \in \mathcal{A}$ disjunkt und sei $A = \bigcup_{i=1}^{\infty}$. Dann ist auch

$$x_1 \mapsto \mu_2(A_{x_1}) = \sum_{i=1}^{\infty} \mu_2\big((A_i)_{x_1}\big)$$

messbar als punktweiser Grenzwert. Also ist $\mathcal{A}$ ein Mengensystem mit den Eigenschaften einer $\sigma$-Algebra, wobei allerdings zunächst die Vereinigungsbedingung nur für disjunkte Mengen gilt. Man zeigt schließlich mittels der Durchschnittsstabilität des erzeugenden Systems der Rechtecke, dass die Disjunktheit nicht notwendig ist, also $\mathcal{A}$ tatsächlich eine $\sigma$-Algebra ist.

Ist nun $\mu_2$ nicht endlich, aber $\sigma$-endlich, wählen wir eine Folge $B_1 \subseteq B_2 \subseteq \dots$ in $\mathcal{A}_2$ mit $\mu_2(B_n) < \infty$ und $\bigcup B_n = \Omega_2$ und betrachten die endlichen Masse $\nu_n(A_2) = \mu(A_2 \cap B_n)$. Aus der schon gezeigten Messbarkeit von $x_1 \mapsto \nu_n(A_{x_1})$ folgt dann die Messbarkeit von $x_1 \mapsto \mu_2(A_{x_1})$ wegen $\mu_2(A_{x_1}) = \sup_n \nu_n(A_{x_1})$. $\square$

**Satz 1.27.** *Sind $(\Omega_i, \mathcal{A}_i, \mu_i)$ für $i = 1, 2$ $\sigma$-endliche Maßräume, dann gibt es ein eindeutig bestimmtes Maß $\mu_1 \otimes \mu_2$ auf $A_1 \otimes A_2$, das* Produktmaß *von $\mu_1$ und $\mu_2$, mit*

$$\mu_1 \otimes \mu_2(A_1 \times A_2) = \mu_1(A_1)\mu_2(A_2) \qquad \textit{für } A_i \in \mathcal{A}_i, i = 1, 2.$$

*Weiter gilt*

$$\mu_1 \otimes \mu_2(A) = \int_{\Omega_1} \mu_2(A_{x_1})\,\mathrm{d}\mu_1(x_1) = \int_{\Omega_2} \mu_1(A^{x_2})\,\mathrm{d}\mu_2(x_2)$$

*für alle $A \in A_1 \times A_2$.*

*Beweis.* Die Bedingung

$$\mu(A_1 \times A_2) = \mu_1(A_1)\mu_2(A_2) \qquad \text{für } A_i \in \mathcal{A}_i, i = 1,2$$

bestimmt $\mu$ eindeutig auf dem Ring der endlichen Vereinigungen von Rechtecken $A_1 \times A_2$ mit $A_i \in \mathcal{A}_i, i = 1,2$. Nach Satz 1.12 ist $\mu_1 \otimes \mu_2$ damit eindeutig bestimmt.

Für die Existenz definieren wir

$$\mu(A) = \int_{\Omega_1} \mu_2(A_{x_1})\,\mathrm{d}\mu_1(x_1) \qquad \text{für } A \in A_1 \times A_2$$

und zeigen, dass $\mu$ ein Maß ist und

$$\mu(A_1 \times A_2) = \mu_1(A_1)\mu_2(A_2) \qquad \text{für } A_i \in \mathcal{A}_i, i = 1,2$$

erfüllt. Die Messbarkeit des Integranden haben wir in Lemma 1.26 gezeigt. $\mu(\emptyset) = 0$ ist offensichtlich. Die $\sigma$-Additivität folgt mit dem Satz von der monotonen Konvergenz. Tatsächlich gilt für eine disjunkte Vereinigung $A = \bigcup_n A_n$ von Mengen $A_n \in A_1 \otimes A_2$

$$\begin{aligned} \mu(A) &= \int_{\Omega_1} \mu_2\left(\bigcup_n (A_n)_{x_1}\right) \mathrm{d}\mu_1(x_1) = \int_{\Omega_1} \sum_n \mu_2\big((A_n)_{x_1}\big)\,\mathrm{d}\mu_1(x_1) \\ &= \sum_n \int_{\Omega_1} \mu_2\big((A_n)_{x_1}\big)\,\mathrm{d}\mu_1(x_1) = \sum_n \mu(A_n). \end{aligned}$$

Schließlich haben wir auch

$$\mu(A_1 \times A_2) = \int_{\Omega_1} \mu_2(A_2)\chi_{A_1}(x_1)\,\mathrm{d}\mu_1(x_1) = \mu_1(A_1)\mu_2(A_2)$$

für $A_i \in \mathcal{A}_i, i = 1,2$. □

**Beispiel.** Das Produktmaß $\lambda \otimes \lambda$ auf der Produkt-$\sigma$-Algebra $\mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R}) = \mathcal{B}(\mathbb{R}^2)$ ist das Lebesguemaß $\lambda^2$. Das folgt daraus, dass das Produktmaß auf den Intervallen in $\mathcal{I}^2$ und damit auch den Figuren in $\mathcal{F}^2$ nach Definition mit dem Jordan-Inhalt übereinstimmt. Analog sieht man $\lambda^{d_1} \otimes \lambda^{d_2} = \lambda^{d_1+d_2}$. Die Integralformel

$$\lambda^{d_1+d_2}(A) = \int_{\mathbb{R}^{d_1}} \lambda^{d_2}(A_{x_1})\,\mathrm{d}\lambda^{d_1}(x_1)$$

ist eine Verallgemeinerung des *Prinzips von Cavalieri.*

Wir kommen jetzt zur Vertauschbarkeit von iterierten Integralen, dem Satz von Fubini. Ist $f : \Omega_1 \times \Omega_2 \to [-\infty,\infty]$, so betrachten wir für festes $x_1 \in \Omega_1$ die durch Festhalten von $x_1$ induzierte Funktion $f_{x_1} : \Omega_2 \to [-\infty,\infty]$ gegeben durch $f_{x_1}(x_2) = f(x_1,x_2)$. Dann ist $f_{x_1} = f \circ \varphi$ mit $\varphi : x_2 \mapsto (x_1,x_2)$ und $f_{x_1}$ ist messbar sofern $f$ messbar ist. Analog definieren wir für festes $x_2 \in \Omega_2$ die durch Festhalten von $x_2$ induzierte Funktion $f^{x_2} : \Omega_1 \to [-\infty,\infty]$ gegeben durch $f^{x_2}(x_1) = f(x_1,x_2)$. Wieder ist $f^{x_2}$ messbar, sofern $f$ messbar ist.

Wir wollen unter geeigneten Voraussetzungen die Formel

$$\int_{\Omega_1\times\Omega_2} f\,\mathrm{d}\mu_1 \otimes \mu_2 = \int_{\Omega_1} \left(\int_{\Omega_2} f_{x_1}\,\mathrm{d}\mu_2(x_2)\right) \mathrm{d}\mu_1(x_1) = \int_{\Omega_2} \left(\int_{\Omega_1} f^{x_2}\,\mathrm{d}\mu_1(x_1)\right) \mathrm{d}\mu_2(x_2) \tag{4}$$

oder einprägsamer

$$\int_{\Omega_1\times\Omega_2} f\,\mathrm{d}\mu_1 \otimes \mu_2 = \int_{\Omega_1} \int_{\Omega_2} f(x_1,x_2)\,\mathrm{d}\mu_2(x_2)\,\mathrm{d}\mu_1(x_1) = \int_{\Omega_2} \int_{\Omega_1} f(x_1,x_2)\,\mathrm{d}\mu_1(x_1)\,\mathrm{d}\mu_2(x_2)$$

zeigen.

**Satz 1.28** (Satz von Fubini). *Seien $(\Omega_i, \mathcal{A}_i, \mu_i)$ für $i = 1, 2$ $\sigma$-endliche Maßräume.*

*(a) Ist $f : \Omega_1 \times \Omega_2 \to [0, \infty]$ messbar, dann gilt (4).*

*(b) Ist $f : \Omega_1 \times \Omega_2 \to \mathbb{R}$ integrierbar, dann sind die Funktionen $f_{x_1}$ und $f^{x_2}$ fast überall integrierbar und es gilt (4).*

*(c) Ist $f : \Omega_1 \times \Omega_2 \to \mathbb{R}$ messbar und ist*

$$\int_{\Omega_1} \int_{\Omega_2} |f(x_1, x_2)| \,\mathrm{d}\mu_2(x_2) \,\mathrm{d}\mu_1(x_1) < \infty,$$

*dann ist auch*

$$\int_{\Omega_2} \int_{\Omega_1} |f(x_1, x_2)| \,\mathrm{d}\mu_1(x_1) \,\mathrm{d}\mu_2(x_2) < \infty,$$

*$f$ ist integrierbar und es gilt (4).*

**Bemerkung.** Den Teil (a) nennt man auch den Satz von Tonelli.

*Beweis.* Wir zeigen nur (a), die Aussagen in (b) und (c) folgen daraus. Die Aussage gilt nach Satz 1.27 für Indikatorfunktionen messbarer Mengen wegen

$$\mu_1 \otimes \mu_2(A) = \int_{\Omega_1} \mu_2(A_{x_1}) \,\mathrm{d}\mu_1(x_1) = \int_{\Omega_1} \int_{\Omega_2} \chi_A(x_1, x_2) \,\mathrm{d}\mu_2(x_2) \,\mathrm{d}\mu_1(x_1).$$

Mit dem üblichen Verfahren überträgt sie sich auf Treppenfunktionen und monotone Grenzwerte von Treppenfunktionen. □

# 2 Funktionalanalysis

Die Funktionalanalysis ist am Anfang des 20. Jahrhunderts aus dem Bedürfnis entstanden, den gemeinsamen Kern von Problemen aus verschiedenen Gebieten wie Schwingungs- und Elektrizitätslehre (Fourierreihen), Potentialtheorie, Quantentheorie und Lösung von Differential- und Integralgleichungen herauszuarbeiten. Grundidee dabei ist, eine Klasse von Funktionen als Elemente eines linearen Raumes aufzufassen und diesen mit einer geeigneten Metrik oder Norm zu versehen. Aus den Ergebnissen einer Theorie solcher normierter Räume und den zwischen ihnen wirkenden linearen Operatoren ergeben sich dann vielfältige Anwendungsmöglichkeiten.

Wir betrachten als motivierendes Beispiel ein Anfangswertproblem für eine einfache lineare Differentialgleichung

$$x'(t) - a(t)x(t) = b(t), x(0) = 0.$$

Hier sind $a, b$ gegebene stetige Funktionen. Gesucht ist eine differenzierbare Funktion $x : [0,1] \to \mathbb{R}$, die die Gleichung und die Anfangsbedingung erfüllt. Durch Integration erhält man die im wesentlichen äquivalente Integralgleichung

$$x(s) = f(s) + \int_0^s x(t)a(t)\,\mathrm{d}t.$$

Diese kann man im Sinne der linearen Algebra als ein lineares Gleichungssystem mit unendlich vielen Variablen $x(t)$ und unendlich vielen Gleichungen (für jedes $s \in [0,1]$ eine) auffassen. Betrachtet man nun $x$ als Element eines linearen Raumes von Funktionen und den Operator $T$ gegeben durch $(Tx)(s) = \int_0^s x(t)a(t)\,\mathrm{d}t$ als Operator, der stetige Funktionen auf stetige Funktionen abbildet, so erhält man die Gleichung

$$x - Tx = f,$$

die es zu lösen gilt. der Operator $T$ ist ein linearer Operator.

Allgemeiner betrachtet man *Fredholmsche Integralgleichungen* dieser Form mit einem durch einen Kern $k(s,t)$ gegebenen Integraloperator

$$(Tx)(s) = \int_0^1 k(s,t)x(t)\,\mathrm{d}t.$$

Die Lösungstheorie solcher Operatorgleichungen ist eine der vielfältigen Anwendungen der funktionalanalytischen Methode.

## 2.1 Normierte Räume und Banachräume

Wir beginnen zunächst mit dem Studium der zugrundeliegenden normierten Räume. Grundlage ist ein linearer Raum über dem Skalarkörper $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$. Für die meisten Betrachtungen spielt die Wahl des Skalarkörpers keine Rolle.

**Definition 2.1.** Eine Funktion $\|\cdot\| : X \to [0, \infty)$ auf einem linearen Raum mit den Eigenschaften

(a) $\|\lambda x\| = |\lambda|\|x\|$ für $\lambda \in \mathbb{K}$ und $x \in X$

(b) $\|x + y\| \leq \|x\| + \|y\|$ für $x, y \in X$

heißt *Halbnorm* auf $X$. Eine Halbnorm $\|\cdot\| : X \to [0, \infty)$ auf $X$ heißt *Norm*, wenn zusätzlich

(c) $\|x\| = 0 \iff x = 0$

gilt. Dann heißt $X$ *halbnormierter bzw. normierter Raum.*

Die Ungleichung (b) heißt Dreiecksungleichung. Um verschiedene Normen zu unterscheiden, schreiben wir auch gelegentlich $\|x\|_X$ oder $\|x|X\|$ für $\|x\|$.

Ist $X$ ein normierter Raum, dann definiert

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

eine Metrik auf $X$, die *induzierte* Metrik. Jeder normierte Raum ist also ein metrischer Raum und wir können topologische Begriffe benutzen. Wichtig werden sein:

- *Konvergenz:* Eine Folge $(x_n)$ in $X$ konvergiert gegen $x \in X$ (also $x_n \to x$), wenn $d(x_n, x) = \|x_n - x\| \to 0$ gilt.
- *Cauchyfolgen:* Eine Folge $(x_n)$ in $X$ ist eine Cauchyfolge, wenn zu jedem $\varepsilon > 0$ ein $n_0$ existiert mit $d(x_n, x_m) = \|x_n - x_m\| < \varepsilon$ für $n, m \geq n_0$ gilt.
- *offene Mengen:* Eine Menge $O \subset X$ ist offen, wenn es zu jedem $x \in O$ ein $\varepsilon > 0$ gibt mit $U(x, \varepsilon) = \{y \in X \,:\, \|y - x\| < \varepsilon\} \subseteq O$.
- *abgeschlossene Mengen:* Eine Menge $A \subseteq X$ ist abgeschlossen, wenn $X \setminus A$ offen ist.
- *Kompaktheit:* Eine Menge $K \subseteq X$ ist kompakt, wenn jede Folge in $K$ eine konvergente Teilfolge mit Grenzwert in $K$ hat. Dies ist genau dann der Fall, wenn jede offene Überdeckung von $K$ eine endliche Teilüberdeckung besitzt.
- *Vollständigkeit:* $X$ ist vollständig, wenn jede Cauchyfolge in $X$ konvergent ist.

**Definition 2.2.** Ein vollständiger normierter Raum heißt *Banachraum.*

**Beispiel.** Wir betrachten den linearen Raum $X = \mathbb{K}^d$ mit verschiedenen Normen. Dazu sei $x = (\xi_1, \dots, \xi_d) \in \mathbb{K}^d$. Dann sind durch

$$\begin{aligned}
\|x\|_1 &= \sum_{k=1}^{d} |\xi_k| \\
\|x\|_2 &= \left( \sum_{k=1}^{d} |\xi_k|^2 \right)^{1/2} \\
\|x\|_\infty &= \max_{k=1,\dots,d} |\xi_k|
\end{aligned}$$

Normen auf $X$ definiert. Die entsprechenden normierten Räume bezeichnet man mit $\ell_1^d, \ell_2^d, \ell_\infty^d$. Die induzierte Metrik von $\ell_2^d$ ist die bekannte euklidische Metrik

$$d(x,y) = \sqrt{\sum_{k=1}^{d} |\xi_k - \eta_k|^2}$$

für $x = (\xi_1, \dots, \xi_d), y = (\eta_1, \dots, \eta_d) \in \mathbb{K}^d$.

**Beispiel.** Sei $M$ eine beliebige Menge. Dann ist der Raum $\ell_\infty(M)$ aller beschränkten Funktionen $x : M \to \mathbb{K}$ mit der Supremumsnorm

$$\|x\|_\infty = \sup_{t \in M} |x(t)|$$

ein Banachraum. Dass $\ell_\infty(M)$ ein linearer Raum ist, ist dabei offensichtlich. Die punktweise Dreiecksungleichung liefert

$$|(x+y)(t)| = |x(t) + y(t)| \leq |x(t)| + |y(t)| \leq \|x\|_\infty + \|y\|_\infty.$$

Übergang zum Supremum auf der linken Seite zeigt dann die Dreiecksungleichung

$$\|x+y\|_\infty \leq \|x\|_\infty + \|y\|_\infty.$$

Die anderen Normeigenschaften sind offensichtlich.

Es bleibt, die Vollständigkeit von $\ell_\infty(M)$ nachzuweisen. Dazu sei $(x_n)$ eine Cauchyfolge in $\ell_\infty(M)$. Wegen

$$|x_n(t) - x_m(t)| \leq \|x_n - x_m\|_\infty$$

ist dann jede Folge $\big(x_n(t)\big)$ eine Cauchyfolge in $\mathbb{K}$ und damit konvergent. Also hat die Funktionenfolge $(x_n)$ einen punktweisen Grenzwert $x$. Es ist noch zu zeigen, dass $x$ selbst beschränkt ist und die Konvergenz $x_n \to x$ auch in der Supremumsnorm gilt.

Da $(x_n)$ eine Cauchyfolge ist, gibt es zu jedem $\varepsilon > 0$ ein $n_0$, so dass $\|x_n - x_m\|_\infty < \varepsilon$ für $n, m \geq n_0$ gilt. Damit gilt für jedes $t \in M$ $|x_n(t) - x_m(t)| < \varepsilon$ für $n, m \geq n_0$. Grenzübergang $m \to \infty$ liefert $|x_n(t) - x(t)| \leq \varepsilon$ für $n \geq n_0$ und alle $t \in M$ und damit $\|x_n - x\|_\infty \leq \varepsilon$ für $n \geq n_0$. Damit ist zunächst $x_n - x$ beschränkt, also ist auch $x$ beschränkt. Außerdem ist nun $x_n \to x$ in $\ell_\infty(M)$ gezeigt.

Die Konvergenz in $\ell_\infty(M)$ ist gerade die gleichmäßige Konvergenz.

Das folgende einfache Kriterium erleichtert oft den Nachweis der Vollständigkeit: Ist $X$ ein Banachraum und $U \subseteq X$ abgeschlossen, dann ist $U$ vollständig.

**Beispiel.** Für einen metrischen Raum $M$ bezeichnet man mit $C(M)$ den linearen Raum aller stetigen Funktionen $x : M \to \mathbb{K}$ und mit $C_b(M)$ den linearen Teilraum aller beschränkten stetigen Funktionen. Für kompaktes $M$ ist $C(M) = C_b(M)$. Der Raum $C_b(M)$ mit der Supremumsnorm (und damit für kompaktes $M$ der Raum $C(M)$ mit der Supremumsnorm) ist ein Banachraum. Mit dem obigen Kriterium ist nur zu zeigen, dass $C_b(M)$ in $\ell_\infty(M)$ abgeschlossen ist, also dass gleichmäßige Grenzwerte stetiger Funktionen wieder stetig sind. Dies beweist man wie in der Grundvorlesung Analysis.

**Beispiel.** Als weiteres Beispiel wollen wir Räume von Folgen $x = (\xi_k) = (\xi_k)_{k \in \mathbb{N}}$ in $\mathbb{K}$ betrachten. Zunächst ist $\ell_\infty = \ell_\infty(\mathbb{N})$ der Raum aller beschränkten Folgen mit der Supremumsnorm $\|x\|_\infty =$

$\sup_n |\xi_k|$. Nun sei $c$ der lineare Teilraum aller konvergenten Folgen, $c_0$ der lineare Teilraum aller Nullfolgen und $d$ der lineare Teilraum aller abbrechenden Folgen, die also $\xi_k = 0$ für genügend großes $k$ erfüllen. Offensichtlich ist $d \subseteq c_0 \subseteq c \subseteq \ell_\infty$ und diese Inklusionen sind echt. Man zeigt (Vorlesung), dass $c$ und $c_0$ abgeschlossene Teilräume von $\ell_\infty$ und damit selbst Banachräume sind. Der Raum $d$ ist nicht abgeschlossen, wie z.B. die Cauchyfolge $(x_n)$ mit

$$x_n = (1, 1/2, 1/3, \dots, 1/n, 0, 0, \dots)$$

zeigt, deren Grenzwert $x = (1/n)$ in $c_0$ existiert, aber nicht in $d$ liegt. Also ist $d$ nicht vollständig.

**Beispiel.** Betrachten wir den Raum $C^1[0,1]$ aller stetig differenzierbaren Funktionen als linearen Teilraum von $C[0,1]$ (mit der Supremumsnorm), dann ist $C^1[0,1]$ nicht abgeschlossen in $C[0,1]$ und damit kein Banachraum. Versieht man hingegen $C^1[0,1]$ mit der Norm

$$\|x\| = \|x\|_\infty + \|x'\|_\infty,$$

so erhalten wir einen Banachraum (Beweise in Vorlesung).

**Beispiel** (Die Räume $\mathcal{L}_p(\mu)$ und $L_p(\mu)$)**.** Sei $(\Omega, \mathcal{A}, \mu)$ ein fester Maßraum und sei $0 < p < \infty$. Dann definieren wir

$$\mathcal{L}_p(\mu) = \left\{ x : \Omega \to \mathbb{R} \,:\, x \text{ ist messbar und } \int_\Omega |x|^p \, \mathrm{d}\mu < \infty \right\}.$$

Ist $A \subset \mathbb{R}^d$ und $\mu$ das Lebesguemaß auf der Borelschen $\sigma$-Algebra auf $A$, so schreibt man auch $\mathcal{L}_p(A)$, bei z.B. $\mathcal{L}_p[0,1]$ lässt man oft noch die Klammern um $A = [0,1]$ weg. Die Räume $\mathcal{L}_p(\mu)$ sind reelle lineare Räume. Nichttrivial ist nur die Eigenschaft, dass aus $x, y \in \mathcal{L}_p(\mu)$ auch $x + y \in \mathcal{L}_p(\mu)$ folgt. Dabei ist die Messbarkeit von $x + y$ klar, die Endlichkeit des Integrals folgt aus der punktweisen Ungleichung $|x + y| \le 2 \max\{|x|, |y|\}$.

Wir setzen nun

$$\|x\|_p = \left( \int_\Omega |x|^p \, \mathrm{d}\mu \right)^{1/p} \qquad \text{für } x \in \mathcal{L}_p(\mu).$$

Wir wollen zeigen, dass dies eine Halbnorm auf $\mathcal{L}_p(\mu)$ definiert. Im allgemeinen ist es keine Norm, da ein $x$, dass fast überall, aber nicht überall, gleich 0 ist, die Norm 0 hat. Die Homogenität $\|\lambda x\|_p = |\lambda| \|x\|_p$ ist dabei klar. Für die Dreiecksungleichung $\|x + y\|_p \le \|x\|_p + \|y\|_p$ im Fall $1 < p < \infty$ benötigen wir als wichtige Vorbereitung die *Hölder-Ungleichung*, der Fall $p = 1$ ist klar.

**Satz 2.3** (Hölder-Ungleichung)**.** *Sei $1 < p < \infty$ und sei $q \in (1, \infty)$ der zu $p$* konjugierte Index *gegeben durch die Gleichung $\frac{1}{p} + \frac{1}{q} = 1$. Sind dann $x \in \mathcal{L}_p(\mu)$ und $y \in \mathcal{L}_q(\mu)$, dann ist $xy \in \mathcal{L}_1(\mu)$ und es gilt*

$$\|xy\|_1 \le \|x\|_p \|y\|_q.$$

*Beweis.* Da die Logarithmusfunktion konkav ist, gilt

$$\log(\xi^r \eta^{1-r}) = r \log \xi + (1 - r) \log \eta \le \log(r\xi + (1 - r)\eta) \quad \text{für } 0 \le r \le 1, \xi, \eta > 0.$$

Aus der Monotonie der Logarithmusfunktion folgt dann

$$\xi^r \eta^{1-r} \le r\xi + (1 - r)\eta \quad \text{für } 0 \le r \le 1, \xi, \eta > 0.$$

Ist nun $a = \|x\|_p^p, b = \|y\|_q^q$ und nehmen o.B.d.A $a, b > 0$ an und setzen wir

$$r = \frac{1}{p}, 1 - r = \frac{1}{q}, \xi = \frac{|x|^p}{a}, \eta = \frac{|y|^q}{b}$$

und integrieren die Ungleichung über $\Omega$, so erhalten wir

$$\frac{\|xy\|_1}{a^{1/p}b^{1/q}} \leq \frac{1}{p}\frac{\|x\|_p^p}{a} + \frac{1}{q}\frac{\|y\|_q^q}{b} = \frac{1}{p} + \frac{1}{q} = 1$$

und schließlich

$$\|xy\|_1 \leq a^{1/p}b^{1/q} = \|x\|_p\|y\|_q.$$

□

Aus der Hölder-Ungleichung leiten wir nun die Dreiecksungleichung ab, die einen eigenen Namen trägt.

**Satz 2.4** (Minkowski-Ungleichung). *Sei $1 \leq p < \infty$ und seien $x, y \in \mathcal{L}_p(\mu)$. Dann gilt*

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

*Beweis.* Für $p = 1$ ist die Ungleichung trivial, also sei $p > 1$ und $q$ der zu $p$ konjugierte Index. Dann ist

$$\|x + y\|_p^p = \big\||x + y|^p\big\|_1 \leq \big\||x| \cdot |x + y|^{p-1}\big\|_1 + \big\||y| \cdot |x + y|^{p-1}\big\|_1.$$

Wegen $(p - 1)q = p$ ist $|x + y|^{p-1} \in \mathcal{L}_q(\mu)$ und wir erhalten aus der Hölder-Ungleichung

$$\big\||x| \cdot |x + y|^{p-1}\big\|_1 \leq \|x\|_p\big\||x + y|^{p-1}\big\|_q \qquad \text{und} \qquad \big\||y| \cdot |x + y|^{p-1}\big\|_1 \leq \|y\|_p\big\||x + y|^{p-1}\big\|_q$$

und somit

$$\|x + y\|_p^p \leq \big(\|x\|_p + \|y\|_p\big)\big\||x + y|^{p-1}\big\|_q = \big(\|x\|_p + \|y\|_p\big)\|x + y\|_p^{p-1}.$$

Division durch $\|x + y\|_p^{p-1}$ liefert die Behauptung. □

Die Hölder-Ungleichung benutzt man manchmal auch in der folgenden verallgemeinerten Form

**Satz 2.5** (Hölder-Ungleichung II). *Seien $0 < p, q, r < \infty$ und $\frac{1}{r} = \frac{1}{p} + \frac{1}{q}$. Sind dann $x \in \mathcal{L}_p(\mu)$ und $y \in \mathcal{L}_q(\mu)$, dann ist $xy \in \mathcal{L}_r(\mu)$ und es gilt*

$$\|xy\|_r \leq \|x\|_p\|y\|_q.$$

Diese Ungleichung folgt sofort aus der ursprünglichen Hölder-Ungleichung durch setzen von

$$\tilde{p} = \frac{r}{p}, \tilde{q} = \frac{r}{q} \qquad \text{und} \qquad \tilde{x} = |x|^r, \tilde{y} = |y|^r.$$

**Beispiel.** Sei $\Omega = \{1, 2, \ldots, d\}$ und sei $\mu$ das Zählmaß auf $\Omega$. Eine Funktion $x : \Omega \to \mathbb{R}$ identifizieren wir dann mit einem Vektor $x = (\xi_1, \ldots, \xi_d) \in \mathbb{R}^d$. Jede solche Funktion ist messbar und wir haben $\int_\Omega x \, \mathrm{d}\mu = \sum_{k=1}^d \xi_k$. Wir erhalten für alle $p > 0$ also $\mathcal{L}_p(\mu) = \mathbb{R}^d$ und

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

Für $p \geq 1$ ist dies eine Norm auf $\mathbb{R}^d$, denn in unserem konkreten Fall bedeutet fast überall sogar überall. Den normierten Raum $(\mathbb{R}^d, \|\cdot\|_p)$ bezeichnet man mit $\ell_p^d$. Dies stimmt für $p = 1, 2$ mit unserer bisherigen Bezeichnung überein. Die Hölder-Ungleichung ist dann die Ungleichung

$$\sum_{k=1}^{d} |\xi_k \eta_k| \leq \left(\sum_{k=1}^{d} |\xi_k|^p\right)^{1/p} \left(\sum_{k=1}^{d} |\eta_k|^q\right)^{1/q}.$$

Den Spezialfall $p = q = 2$ kann man auch als

$$|\langle x, y\rangle| \leq \|x\|_2 \|y\|_2$$

für $x, y \in \mathbb{R}^d$ schreiben, das ist die Cauchy-Schwarz-Ungleichung im $\mathbb{R}^d$. Die Minkowski-Ungleichung lautet

$$\left(\sum_{k=1}^{d} |\xi_k + \eta_k|^p\right)^{1/p} \leq \left(\sum_{k=1}^{d} |\xi_k|^p\right)^{1/p} + \left(\sum_{k=1}^{d} |\eta_k|^p\right)^{1/p}$$

und gilt ebenfalls für $p \geq 1$.

Wir zeigen noch die folgenden Ungleichungen zwischen verschiedenen $p$-Normen: Ist $0 < p_2 < p_1 \leq \infty$, dann gilt für alle $x \in \mathbb{R}^d$

$$\|x\|_{p_1} \leq \|x\|_{p_2} \leq d^{1/p_2 - 1/p_1} \|x\|_{p_1}.$$

Hier setzen wir $1/\infty = 0$ für $p_1 = \infty$. In diesem Fall folgt die zweite Ungleichung direkt aus $|\xi_k|^{p_2} \leq \|x\|_\infty^{p_2}$ für $k = 1, \ldots, d$. Im Fall $p_2 < \infty$ benutzen wir die allgemeine Form der Hölder-Ungleichung mit $1/r = 1/p_2 - 1/p_1$ und erhalten

$$\|x\|_{p_2} = \left(\sum_{k=1}^{d} |\xi_k|^{p_2}\right)^{1/p_2} = \left(\sum_{k=1}^{d} 1^{p_2} \cdot |\xi_k|^{p_2}\right)^{1/p_2} \leq \left(\sum_{k=1}^{d} 1^r\right)^{1/r} \left(\sum_{k=1}^{d} |\xi_k|^{p_1}\right)^{1/p_1} = d^{1/r} \|x\|_{p_1}.$$

Für die erste Ungleichung benutzen wir die Monotonie der Potenzfunktion $t^{p_1} \leq t^{p_2}$ für $0 \leq t \leq 1$ und erhalten

$$\|x\|_{p_1}^{p_1} = \|x\|_{p_2}^{p_1} \sum_{k=1}^{d} \left|\frac{\xi_k}{\|x\|_{p_2}}\right|^{p_1} \leq \|x\|_{p_2}^{p_1} \sum_{k=1}^{d} \left|\frac{\xi_k}{\|x\|_{p_2}}\right|^{p_2} = \|x\|_{p_2}^{p_1}.$$

**Beispiel.** Sei nun $\Omega = \mathbb{N}$ und sei wieder $\mu$ das Zählmaß auf $\mathbb{N}$. Eine Funktion $x : \mathbb{N} \to \mathbb{R}$ identifizieren wir dann mit einer Folge $x = (\xi_k)_{k \in \mathbb{N}}$. Jede solche Funktion ist messbar und wir haben $\int_\Omega x \, \mathrm{d}\mu = \sum_{k=1}^\infty \xi_k$ falls die Reihe absolut konvergiert. Für $p > 0$ erhalten wir

$$\mathcal{L}_p(\mu) = \left\{ x = (\xi_k) \in \mathbb{R}^{\mathbb{N}} \,:\, \sum_{k=1}^{\infty} |\xi_k|^p < \infty \right\}$$

und

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

Für $p \geq 1$ ist dies eine Norm, den resultierenden normierten Raum bezeichnet man mit $\ell_p$. Die Hölder-Ungleichung ist dann die Ungleichung

$$\sum_{k=1}^{\infty} |\xi_k \eta_k| \leq \left(\sum_{k=1}^{\infty} |\xi_k|^p\right)^{1/p} \left(\sum_{k=1}^{\infty} |\eta_k|^q\right)^{1/q}.$$

Den Spezialfall $p = q = 2$ kann man auch als

$$|\langle x, y\rangle| \leq \|x\|_2 \|y\|_2$$

für $x, y \in \ell_2$ schreiben, das ist die Cauchy-Schwarz-Ungleichung in $\ell_2$. Die Minkowski-Ungleichung lautet

$$\left(\sum_{k=1}^{\infty} |\xi_k + \eta_k|^p\right)^{1/p} \leq \left(\sum_{k=1}^{\infty} |\xi_k|^p\right)^{1/p} + \left(\sum_{k=1}^{\infty} |\eta_k|^p\right)^{1/p}$$

und gilt ebenfalls für $p \geq 1$.

In den letzten beiden Beispielen sind die $\mathcal{L}_p$-Halbnormen sogar Normen. Da dies im Allgemeinen nicht der Fall ist, geht man von $\mathcal{L}_p(\mu)$ zu einem normierten Raum $L_p(\mu)$ über, indem man Äquivalenzklassen von Funktionen betrachtet, die sich nur in einer Nullmenge unterscheiden. Wir definieren also eine Äquivalenzrelation auf $\mathcal{L}_p(\mu)$ durch

$$x \sim y \iff x = y \text{ fast überall } \iff \|x - y\|_p = 0.$$

Der Raum $L_p(\mu)$ ist dann der lineare Raum aller Äquivalenzklassen $[x]$ mit der Norm $\|[x]\|_p = \|x\|_p$. Algebraisch ist das also der Quotientenraum von $\mathcal{L}_p(\mu)$ nach dem Nullraum

$$N = \{x \in \mathcal{L}_p(\mu) \,:\, x = 0 \text{ fast überall}\} = \{x \in \mathcal{L}_p(\mu) \,:\, \|x\|_p = 0\}.$$

Da für $x \sim y$ auch $\|x\|_p = \|y\|_p$ gilt, ist $\|[x]\|_p$ wohldefiniert. Üblicherweise spricht man im Umgang mit Elementen von $L_p(\mu)$ wieder über Funktionen, obwohl eigentlich Äquivalenzklassen von Funktionen gemeint sind. Man schreibt also $x \in L_p(\mu)$ statt $[x] \in L_p(\mu)$ für eine Funktion $x : \Omega \to \mathbb{R}$, sollte sich aber immer im klaren sein, dass $x$ damit nur bis auf eine Nullmenge eindeutig gegeben ist. In diesem Sinne sagt man auch, dass Hölder- und Minkowski-Ungleichung für Funktionen $x \in L_p(\mu)$ gelten.

**Satz 2.6.** *Der Raum $L_p(\mu)$ ist ein Banachraum.*

*Beweis.* Es bleibt noch die Vollständigkeit von $L_p(\mu)$ nachzuweisen. Sei also eine Cauchyfolge in $L_p(\mu)$ gegeben, genauer: eine Folge von Repräsentanten einer Cauchyfolge. Dann finden wir eine Teilfolge $(x_n)$ dieser Cauchyfolge mit

$$\|x_{n+1} - x_n\|_p \leq 2^{-n}.$$

Wir werden die Konvergenz dieser Teilfolge zeigen. Eine Cauchyfolge mit einer konvergenten Teilfolge ist selbst konvergent (Analysis 1, wissen Sie noch wie der Beweis geht?). Damit wäre der Satz gezeigt.

Wir betrachten die Funktion $y = \sum_{n=1}^{\infty} |x_{n+1} - x_n|$ und ihre Partialsummen $y_m = \sum_{n=1}^{m} |x_{n+1} - x_n|$, wobei $y$ Werte in $[0, \infty]$ annehmen kann. Da $\mathcal{L}_p(\mu)$ ein Vektorraum ist, ist $y_m \in \mathcal{L}_p(\mu)$ und die Minkowski-Ungleichung liefert

$$\|y_m\|_p \leq \sum_{n=1}^{m} \|x_{n+1} - x_n\|_p \leq \sum_{n=1}^{m} 2^{-n} \leq 1.$$

Da die Folge $(y_m)$ monoton gegen $y$ konvergiert, erhalten wir aus dem Satz von der monotonen Konvergenz

$$\int_\Omega y^p \,\mathrm{d}\mu = \lim_{m\to\infty} \int_\Omega y_m^p \,\mathrm{d}\mu = \lim_{m\to\infty} \|y_m\|_p^p \leq 1,$$

also ist $y$ fast überall endlich und damit ist die Reihe $x = \sum_{n=1}^{\infty}(x_{n+1} - x_n)$ fast überall absolut konvergent. Auf der übrigbleibenden Nullmenge setzen wir $x = 0$. Als punktweiser Limes messbarer Funktionen ist $x$ messbar und es gilt $|x| \leq y$ und damit

$$\int_\Omega |x|^p \,\mathrm{d}\mu \leq \int_\Omega y^p \,\mathrm{d}\mu \leq 1.$$

Also ist $x \in \mathcal{L}_p(\mu)$.

Wir zeigen nun $x = \sum_{n=1}^{\infty}(x_{n+1} - x_n)$ mit Konvergenz in $\mathcal{L}_p(\mu)$, wir haben schon die fast überall-Konvergenz. Nun konvergiert

$$h_m = \left| x - \sum_{n=1}^{m}(x_{n+1} - x_n) \right|^p$$

punktweise f.ü. gegen 0. Außerdem gilt

$$h_m \leq \left( \sum_{n=m+1}^{\infty} |x_{n+1} - x_n| \right)^p \leq y^p,$$

$h_m$ hat also eine integrierbare Majorante. Nach dem Satz von der dominierten Konvergenz folgt $\int_\Omega h_m \,\mathrm{d}\mu \to 0$ und die Konvergenz der Teleskopreihe in $\mathcal{L}_p(\mu)$ ist gezeigt.

Die Partialsummen dieser Teleskopreihe sind aber $x_n - x_1$, also finden wir, dass die Folge $(x_n - x_1)$ in $\mathcal{L}_p(\mu)$ gegen $x \in \mathcal{L}_p(\mu)$ konvergiert. Damit ist auch die Folge $(x_n)$ in $\mathcal{L}_p(\mu)$ konvergent (gegen $x + x_1$) und der Beweis ist vollständig. $\square$

**Beispiel** (Die Räume $\mathcal{L}_\infty(\mu)$ und $L_\infty(\mu)$)**.** Sei wieder $(\Omega, \mathcal{A}, \mu)$ ein fester Maßraum und sei $p = \infty$. Eine messbare Funktion $x : \Omega \to \mathbb{R}$ heißt *essentiell beschränkt*, wenn es ein $K > 0$ gibt mit $\mu(\{|f| > K\}) = 0$. Der Raum aller essentiell beschränkten Funktionen wird mit $\mathcal{L}_\infty(\mu)$ bezeichnet. Für $x \in \mathcal{L}_\infty(\mu)$ ist $\|x\|_\infty$ das Infimum über alle solchen Konstanten $K$, das man auch als

$$\|x\|_\infty = \inf_{N \text{ Nullmenge}} \sup_{t \in \Omega \setminus N} |x(t)|$$

schreiben kann. Beachten Sie den Unterschied zur Supremumsnorm auf dem Raum der beschränkten Funktionen auf einer beliebigen Menge!

Dieses Infimum wird angenommen. Ist nämlich $(N_k)$ eine Folge von Nullmengen mit $\sup_{t \in \Omega \setminus N_k} |x(t)| < \|x\|_\infty + \frac{1}{k}$, dann ist $N = \bigcup_k N_k$ eine Nullmenge mit $\sup_{t \in \Omega \setminus N} |x(t)| = \|x\|_\infty$.

Wieder ist $\|\cdot\|_\infty$ eine Halbnorm auf $\mathcal{L}_\infty(\mu)$ und es gilt

$$\|x\|_\infty = 0 \iff x = 0 \text{ fast überall}$$

und durch Bildung der Äquivalenzklassen bezüglich fast-überall-Gleichheit erhalten wir den normierten Raum $L_\infty(\mu)$. Betrachtet man $p = 1$ und $q = \infty$ mittels $\frac{1}{1} + \frac{1}{\infty} = 1$ als konjugierte Indizes, gilt die Höldersche Ungleichung: Für $f \in L_1(\mu)$ und $g \in L_\infty(\mu)$ ist $fg \in L_1(\mu)$ und es ist $\|fg\|_1 \leq \|f\|_1 \|g\|_\infty$.

Ist $\mu$ das Zählmaß auf $\Omega = \{1, \ldots, d\}$, so ist $\mathcal{L}_\infty(\mu) = L_\infty(\mu) = \ell_\infty^d$. Ist $\mu$ das Zählmaß auf $\Omega = \mathbb{N}$, so ist $\mathcal{L}_\infty(\mu) = L_\infty(\mu) = \ell_\infty$.

Wir zeigen noch die Vollständigkeit von $L_\infty(\mu)$. Dazu sei $(x_n)$ eine Cauchyfolge in $L_\infty(\mu)$ und seien $N_{n,m}$ Nullmengen mit

$$\|x_n - x_m\|_\infty = \sup_{t \notin N_{n,m}} |x_n(t) - x_m(t)|.$$

Dann erhalten wir für die Nullmenge $N = \bigcup_{n,m} N_{n,m}$ sogar

$$\|x_n - x_m\|_\infty = \sup_{t \notin N} |x_n(t) - x_m(t)| \qquad \text{für alle } n, m \in \mathbb{N}.$$

Auf $\Omega \setminus N$ ist also die Folge $(x_n)$ eine Cauchyfolge in der Supremumsnorm, also eine gleichmäßige Cauchyfolge mit einem beschränkten punktweisen Grenzwert $x$. Setzen wir noch $x = 0$ auf $N$, so erhalten wir eine messbare beschränkte Funktion $x$

$$\|x_n - x\|_\infty \leq \sup_{t \notin N} |x_n(t) - x(t)| \to 0,$$

es ist also $x_n \to x$ in $L_\infty(\mu)$.

Wir beschließen zunächst die Behandlung der $L_p$-Räume mit folgendem nützlichen Dichtheitssatz.

**Satz 2.7.** *Ist $1 \leq p \leq \infty$ und $x \in L_p(\mu)$, dann gibt es eine Folgen von Treppenfunktionen $(x_n)$, die gegen $x$ in $L_p(\mu)$ konvergiert, d.h für die* $\lim \|x_n - x\|_p = 0$ *ist.*

*Beweis.* Für $p = \infty$ ist $x$ beschränkt außerhalb einer Nullmenge und dort gleichmäßiger Grenzwert einer Folge von Treppenfunktionen. Hier ist also nichts zu zeigen.

Ist $1 \leq p < \infty$ und $x \geq 0$, dann können wir $x$ monoton durch Treppenfunktionen $(x_n)$ approximieren. Wegen

$$|x - x_n|^p = (x - x_n)^p \leq x^p$$

ist $x - x_n \in L_p(\mu)$ und damit auch $x_n \in L_p(\mu)$. Die Konvergenz $\|x - x_n\|_p \to 0$ in folgt nun aus dem Satz von der dominierten Konvergenz.

Ist schließlich $x \in L_p(\mu)$ beliebig, dann sind $x_+, x_- \leq |x|$ in $L_p(\mu)$ und wir können $x_+, x_-$ und damit auch $x = x_+ - x_-$ in $L_p$ durch Treppenfunktionen approximieren. □

Wir haben Beispiele von Normen auf dem gleichen linearen Raum kennengelernt, die verschiedene Topologien erzeugen. So war der Raum $C^1[0,1]$ mit der Norm $\|x\|_\infty + \|x'\|_\infty$ ein Banachraum, mit der Norm $\|x\|_\infty$ hingegen nicht. Wir wollen uns mit Normen beschäftigen, die die gleiche Topologie, also insbesondere die gleichen konvergenten Folgen erzeugen - das sind *äquivalente* Normen. Vorher tragen wir kurz einige einfache allgemeine Eigenschaften einer Norm zusammen:

(a) Die Addition ist stetig, aus $x_n \to x$ und $y_n \to y$ folgt $x_n + y_n \to x + y$.

(b) Die skalare Multiplikation ist stetig, aus $\lambda_n \to \lambda$ und $x_n \to x$ folgt $\lambda_n x_n \to \lambda x$.

(c) Die Norm ist stetig, aus $x_n \to x$ folgt $\|x_n\| \to \|x\|$.

Diese Eigenschaften beweist man analog zu den entsprechenden Eigenschaften in $\mathbb{K}$. Wegen

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

folgt (a). Wegen

$$\|\lambda_n x_n - \lambda x\| \leq \|\lambda_n x_n - \lambda_n x\| + \|\lambda_n x - \lambda x\| = |\lambda_n|\|x_n - x\| + |\lambda_n - \lambda|\|x_n\|$$

folgt (b). Für (c) leitet man zunächst die Ungleichung

$$\big|\|x\| - \|y\|\big| \leq \|x - y\|$$

aus der Dreiecksungleichung ab und erhält die Stetigkeit der Norm aus

$$\big|\|x\| - \|x_n\|\big| \leq \|x - x_n\|,$$

Die Normabbildung $\|\cdot\| : X \to \mathbb{R}$ ist also sogar eine Lipschitzabbildung mit Lipschitzkonstante 1.

**Satz 2.8** (Äquivalente Normen). *Seien $\|\cdot\|_1$ und $\|\cdot\|_2$ zwei Normen auf einem linearen Raum $X$. Dann sind die folgenden Aussagen äquivalent:*

*(a) Es gibt zwei Konstanten $0 < a \leq b < \infty$ mit $a\|x\|_1 \leq \|x\|_2 \leq b\|x\|_1$ für alle $x \in X$.*

*(b) Eine Folge ist bezüglich $\|\cdot\|_1$ konvergent genau dann, wenn sie bezüglich $\|\cdot\|_2$ konvergent ist.*

*(c) Eine Folge ist bezüglich $\|\cdot\|_1$ eine Nullfolge genau dann, wenn sie bezüglich $\|\cdot\|_2$ eine Nullfolge ist.*

*In diesem Fall heißen die beiden Normen* äquivalent.

*Beweis.* Da die Implikationen $(a) \Rightarrow (b) \Rightarrow (c)$ offensichtlich sind, müssen wir nur $(c) \Rightarrow (a)$ zeigen. Nehmen wir an, dass es keine solche Konstante $b$ gibt wie in $(a)$ gefordert. Dann finden wir eine Folge $(x_n)$ in $X$ mit $\|x_n\|_2 > n\|x_n\|_1$. Dann ist die Folge $(y_n) = (x_n)/(n\|x_n\|_1)$ eine Nullfolge bezüglich $\|\cdot\|_1$, aber wegen $\|y_n\|_2 > 1$ keine Nullfolge bezüglich $\|\cdot\|_1$. Die Existenz von $a$ zeigt man analog. □

Wir haben schon eingesehen, dass auf $\mathbb{K}^d$ die $\ell_p$-Normen $\|\cdot\|_p$ alle äquivalent sind. Es gilt sogar

**Satz 2.9.** *Auf einem endlichdimensionalen linearen Raum sind alle Normen äquivalent.*

*Beweis.* Beweis in Vorlesung. □

Endlichdimensionale normierte Räume sind also automatisch Banachräume, lineare Teilräume von endlichdimensionalen Räumen sind abgeschlossen. Außerdem ist eine Menge in einem endlichdimensionalen Raum kompakt genau dann, wenn sie beschränkt und abgeschlossen ist. Diesen Satz kennen Sie als Satz von Heine-Borel aus der Analysisvorlesung. Wir werden bald zeigen, dass dies nur in endlichdimensionalen Räumen gilt. Dazu benötigen wir das folgende Lemma von der "Fastsenkrechten".

**Satz 2.10** (Lemma von Riesz). *Ist $X$ ein normierter Raum und ist $U$ ein abgeschlossener echter (d.h. $U \subsetneq X$) Teilraum, dann existiert zu jedem $\delta \in (0,1)$ ein $x_\delta \in X$ mit $\|x_\delta\| = 1$ und $\|x_\delta - u\| > 1 - \delta$ für alle $u \in U$.*

*Beweis.* Wir wählen ein $x \in X \setminus U$ und betrachten den Abstand $d = d(x, U) = \inf_{u \in U} \|x - u\|$ von $x$ zu $U$. Wäre $d = 0$, dann gäbe es eine Folge $(u_n)$ in $U$ mit $u_n \to x$, was aber der Abgeschlossenheit von $U$ widerspricht. Also ist $d > 0$ und damit gibt es ein $u_\delta \in U$ mit $\|x - u_\delta\| > \frac{d}{1-\delta}$. Wir setzen

$$x_\delta = \frac{x - u_\delta}{\|x - u_\delta\|}$$

und erhalten für beliebiges $u \in U$

$$\|x_\delta - u\| = \frac{\big\|x - (u_\delta + \|x - u_\delta\| u)\big\|}{\|x - u_\delta\|} \geq \frac{d}{d/(1-\delta)} = 1 - \delta.$$

□

Es folgt die angekündigte Charakterisierung der endlichdimensionalen normierten Räume mittels des Satzes von Heine-Borel.

**Satz 2.11.** *Ein normierter Raum $X$ ist endlichdimensional genau dann, wenn jede beschränkte und abgeschlossene Menge kompakt ist. Weiter ist dies genau dann der Fall, wenn die abgeschlossene Einheitskugel $B_X = \{x \in X : \|x\| \leq 1\}$ kompakt ist.*

*Beweis.* Wir haben nur noch zu zeigen, dass aus der Kompaktheit von $B_X$ die Endlichdimensionalität von $X$ folgt. Nehmen wir dazu an, dass $X$ nicht endlichdimensional ist. Wir werden dann eine Folge $(x_n)$ konstruieren mit $\|x_n\| = 1$ für alle $n \in \mathbb{N}$ und $\|x_n - x_m\| > \frac{1}{2}$ für alle $n \neq m$. Offensichtlich besitzt diese Folge keine konvergente Teilfolge, $B_X$ ist also nicht kompakt.

Für die Konstruktion der Folge $(x_n)$ starten wir mit einem beliebigen $x_1$ mit $\|x_1\| = 1$ und setzen $U = \operatorname{span}\{x_1\}$ als lineare Hülle von $x_1$ an. Da $X$ unendlichdimensional ist, ist $U \subsetneq X$ und wir finden mit dem Lemma von Riesz für $\delta = \frac{1}{2}$ ein $x_2$ mit $\|x_2\| = 1$ und $\|x_2 - u\| > \frac{1}{2}$ für $u \in U$. Insbesondere ist $\|x_1 - x_2\| > \frac{1}{2}$. Nun setzen wir $U = \operatorname{span}\{x_1, x_2\}$ und finden mit dem Lemma von Riesz ein $x_3$ mit $\|x_3\| = 1$ und $\|x_3 - u\| > \frac{1}{2}$ für $u \in U$. Insbesondere ist $\|x_1 - x_3\|, \|x_2 - x_3\| > \frac{1}{2}$. Dieses Verfahren setzen wir induktiv fort. Im $n$-ten Schritt ist $U = \operatorname{span}\{x_1, x_2, \dots, x_n\}$ und wir finden mit dem Lemma von Riesz ein $x_{n+1}$ mit $\|x_{n+1}\| = 1$ und $\|x_{n+1} - u\| > \frac{1}{2}$ für $u \in U$. Insbesondere ist $\|x_k - x_{n+1}\| > \frac{1}{2}$ für $k = 1, \dots, n$. □

## 2.2 Lineare und stetige Operatoren

Während wir im letzten Abschnitt normierte Räume und Banachräume eingeführt und studiert und einige wichtige Beispiele kennengelernt haben, ist dieser Abschnitt den stetigen linearen Abbildungen zwischen normierten Räumen gewidmet. Wir nennen eine linearen Abbildung $T : X \to Y$ zwischen zwei normierten Räumen $X, Y$ einen Operator, ein stetiger Operator ist also eine stetige lineare Abbildung. Ist $Y = \mathbb{K}$ der Skalarkörper, so spricht man statt von einem Operator auch von einem *Funktional*. Der Raum aller stetigen Operatoren von einem normierten Raum $X$ in einen normierten Raum $Y$ bezeichnet man mit $L(X, Y)$. Der Raum $X' = L(X, \mathbb{K})$ aller stetigen Funktionale heißt *Dualraum* von $X$. Wir schreiben übrigens oft $Tx$ für das Bild $T(x)$.

Das folgende Kriterium zeigt insbesondere, dass eine lineare Abbildung schon stetig ist, wenn sie stetig bei 0 ist.

**Satz 2.12.** *Seien $X, Y$ normierte Räume und sei $T : X \to Y$ linear. Dann sind äquivalent:*

*(a) $T$ ist stetig.*

*(b) $T$ ist stetig in $0 \in X$.*

*(c) $T$ ist* beschränkt*, d.h. es gibt ein $C \geq 0$ mit $\|Tx\|_Y \leq C\|x\|_X$.*

*(d) $T$ ist Lipschitz-stetig.*

*(e) $T$ ist gleichmäßig stetig.*

*Beweis.* Die Implikationen $(c) \Rightarrow (d) \Rightarrow (e) \Rightarrow (a) \Rightarrow (b)$ sind offensichtlich. Wir haben also nur die Implikation $(b) \Rightarrow (c)$ zu zeigen. Sei $T$ also stetig bei 0. Würde $(c)$ nicht gelten, so gäbe es eine Folge $(x_n)$ in $X$ mit $\|Tx_n\|_Y > n\|x_n\|_X$. Setzen wir nun $x'_n = \frac{x_n}{n\|x_n\|_X}$, so erhalten wir $\|x'_n\|_X = \frac{1}{n} \to 0$ und $\|Tx'_n\| > 1$ für alle $n$. Dann wäre $(x'_n)$ eine Nullfolge in $X$, aber $(Tx'_n)$ konvergiert nicht gegen $0 = T(0) \in Y$, also wäre $T$ nicht stetig in 0. □

Wir werden im weiteren oft beide Normen in $X$ und $Y$ einfach mit $\|\cdot\|$ bezeichnen (statt dem eindeutigeren $\|\cdot\|_X$ bzw. $\|\cdot\|_Y$). Machen Sie sich immer klar, in welchem normierten Raum das Element ist, dessen Norm Sie berechnen.

Geht man in $X$ und/oder $Y$ zu einer äquivalenten Norm über, so bleibt die Stetigkeit eines Operators erhalten, allerdings ändert sich im Allgemeinen die Operatornorm.

Aus dieser Charakterisierung ergibt sich, dass $L(X,Y)$ selbst ein linearer Raum ist, da Summen und skalare Vielfache von Nullfolgen wieder Nullfolgen sind. Hierbei werden wie üblich die Summe $S+T$ und das Vielfache $\lambda T$ für $S,T \in L(X,Y)$ und $\lambda \in \mathbb{K}$ punktweise definiert, also

$$(S+T)x = Sx + Tx \qquad \text{und} \qquad (\lambda T)x = \lambda(Tx).$$

Der Nulloperator $0 \in L(X,Y)$ gegeben durch $0(x) = 0 \in Y$ für alle $x \in X$ ist natürlich stetig. Definieren wir nun die *Operatornorm* von $T$ oder einfach Norm $\|T\|$ als kleinste mögliche Konstante in $(c)$, also

$$\|T\| = \inf\{C \geq 0 \,:\, \|Tx\| \leq C\|x\| \text{ für alle } x \in X\} = \sup_{x \neq 0} \frac{\|Tx\|}{\|x\|} = \sup_{\|x\|=1} \|Tx\| = \sup_{\|x\| \leq 1} \|Tx\|.$$

Oft benutzt man diese Definition einfach zur Abschätzung

$$\|Tx\| \leq \|T\|\,\|x\| \qquad \text{für alle } x \in X.$$

Wir zeigen nun, dass die Operatornorm tatsächlich eine Norm ist, also dass $L(X,Y)$ mit der Operatornorm selbst ein normierter Raum ist. Offensichtlich ist die Norm des Nulloperators gleich 0. Ist andererseits $\|T\| = 0$, so folgt $\|Tx\| \leq \|T\|\,\|x\| = 0$, also $\|Tx\| = 0$ für alle $x \in X$. Da $Y$ ein normierter Raum ist, ist dann $Tx = 0$ für alle $x \in X$ und $T$ ist der Nulloperator. Die Homogenität $\|\lambda T\| = |\lambda|\|T\|$ ist fast offensichtlich wegen

$$\|\lambda T\| = \sup_{x \neq 0} \frac{\|(\lambda T)x\|}{\|x\|} = \sup_{x \neq 0} \frac{\|\lambda(Tx)\|}{\|x\|} = \sup_{x \neq 0} \frac{|\lambda|\|Tx\|}{\|x\|} = |\lambda|\|T\|.$$

Ebenso leicht folgt die Dreiecksungleichung aus

$$\begin{aligned}\|S+T\| &= \sup_{\|x\| \leq 1} \|(S+T)x\| = \sup_{\|x\| \leq 1} \|Sx + Tx\| \leq \sup_{\|x\| \leq 1} \big(\|Sx\| + \|Tx\|\big) \\ &\leq \sup_{\|x\| \leq 1} \|Sx\| + \sup_{\|x\| \leq 1} \|Tx\| = \|S\| + \|T\|.\end{aligned}$$

Die Vollständigkeit klärt der folgende Satz.

**Satz 2.13.** *Ist $X$ ein beliebiger normierter Raum und ist $Y$ ein Banachraum, dass ist $L(X,Y)$ ein Banachraum. Insbesondere ist der Dualraum $X' = L(X, \mathbb{K})$ ein Banachraum.*

*Beweis.* Ist $(T_n)$ eine Cauchyfolge in $L(X,Y)$, dann ist wegen $\|T_n x - T_m x\| \leq \|T_n - T_m\|\,\|x\|$ die Folge $(T_n x)$ für jedes $x \in X$ eine Cauchyfolge in $Y$. Nach Voraussetzung ist $Y$ vollständig, also konvergiert $(T_n x)$ in $Y$, sagen wir $T_n x \to Tx \in Y$. Damit ist also eine Abbildung $T : X \to Y$

definiert. Wir zeigen nun, dass $T$ linear und beschränkt ist und dass $\|T_n - T\| \to 0$ gilt. Dann ist $T_n \to T$ in $L(X,Y)$ gezeigt und damit die Vollständigkeit von $L(X,Y)$.

Zunächst ist $T$ linear wegen

$$T(\lambda x + \mu y) = \lim T_n(\lambda x + \mu y) = \lim \big(\lambda T_n x + \mu T_n y\big) = \lambda \lim T_n x + \mu \lim T_n y = \lambda T x + \mu T y.$$

Weiterhin ist $T$ beschränkt wegen

$$\|Tx\| = \|\lim T_n x\| = \lim \|T_n x\| \leq \sup_n \|T_n x\| \leq \sup_n \|T_n\| \|x\|.$$

Als Cauchyfolge ist $(T_n)$ beschränkt, d.h. $\sup \|T_n\| < \infty$, womit die Beschränktheit (und damit Stetigkeit) von $T$ gezeigt ist. Nun wählen wir zu $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ mit $\|T_n - T_m\| < \varepsilon$ für $n, m \geq n_0$. Für jedes $x \in X$ gilt also $\|T_n x - T_m x\| < \varepsilon \|x\|$ und die Stetigkeit der Norm in $Y$ impliziert durch Grenzübergang $m \to \infty$ die Abschätzung $\|T_n x - T x\| < \varepsilon \|x\|$. Also ist $\|T_n - T\| < \varepsilon$ für $n \geq n_0$ und die Konvergenz $\|T_n - T\| \to 0$ ist gezeigt. □

Die Komposition stetiger Abbildungen $T : X \to Y$ und $S : Y \to Z$ ist wieder stetig. Die Ungleichung

$$\|ST\| \leq \|S\| \, \|T\|$$

ist eine quantitative Abschätzung der Operatornorm einer Komposition stetiger Operatoren. Sie gilt wegen

$$\|(ST)x\| = \|S(Tx)\| \leq \|S\| \, \|Tx\| \leq \|S\| \, \|T\| \, \|x\| \qquad \text{für alle } x \in X.$$

Ist $X$ endlichdimensional, so ist *jede* lineare Abbildung $T : X \to Y$ in einen normierten Raum $Y$ stetig. Um dies einzusehen, wählen wir in $X$ eine Basis $\{e_1, \ldots, e_n\}$. Dann ist

$$x = \sum_{k=1}^{n} \lambda_k e_k \mapsto \sum_{k=1}^{n} |\lambda_k|$$

eine Norm auf $X$ und damit äquivalent zur Ausgangsnorm $\|\cdot\|$ von $X$. Es gibt also eine Konstante $C > 0$ mit

$$\sum_{k=1}^{n} |\lambda_k| \leq C \left\| \sum_{k=1}^{n} \lambda_k e_k \right\|$$

und wir erhalten die Stetigkeit von $T$ aus

$$\|Tx\| = \left\| \sum_{k=1}^{n} \lambda_k T e_k \right\| \leq \sum_{k=1}^{n} |\lambda_k| \|Te_k\| \leq C \max_{1 \leq k \leq n} \|Te_k\| \left\| \sum_{k=1}^{n} \lambda_k e_k \right\| = C \max_{1 \leq k \leq n} \|Te_k\| \|x\|.$$

Wir führen nun einige Beispiele für stetige Funktionale an. Die Linearität der Funktional ist jeweils offensichtlich.

**Beispiel** (Punktauswertungen)**.** Sei $X = C[0,1]$ und sei $t_0 \in [0,1]$. Dann ist die Punktauswertung $a_{t_0} : x \mapsto x(t_0)$ ein stetiges Funktional auf $C[0,1]$ mit Norm 1.

Tatsächlich gilt

$$|a_{t_0}(x)| = |x(t_0)| \leq \sup_{t \in [0,1]} |x(t)| = \|x\|_\infty.$$

Für die konstante Funktion $x = 1$ gilt Gleichheit.

**Beispiel** (Grenzwertfunktional auf $c$)**.** Sei $X = c$ der Raum der konvergenten Folgen. Dann ist das Funktional $\lim : x = (\xi_k) \mapsto \lim \xi_k$ ein stetiges Funktional auf $c$ mit Norm 1.

Tatsächlich gilt

$$|\lim(x)| = |\lim \xi_k| \leq \sup_k |\xi_k| = \|x\|_\infty.$$

Für die konstante Folge $x = (1)$ gilt Gleichheit.

Hieraus erhält man sehr schön die Abgeschlossenheit von $c_0 = \lim^{-1}\big(\{0\}\big)$ in $c$.

**Beispiel** (Lineare Funktionale auf $\ell_p$)**.** Sei $X = \ell_p$ mit $1 \leq p \leq \infty$ und sei $a = (\alpha_k) \in \ell_q$, wobei $q$ der zu $p$ konjugierte Index ist. Dann ist das Funktional $a : x = (\xi_k) \mapsto \sum_{k=1}^\infty \alpha_k \xi_k$ ein stetiges Funktional auf $\ell_p$ mit Norm $\|a\|_q$. Hier identifizieren wir nicht ganz koscher aber sinnvoll die Folge $a$ und das lineare Funktional $a$.

Tatsächlich folgt aus der Hölderungleichung

$$|a(x)| \leq \|a\|_q \|x\|_p.$$

Ist $1 < p < \infty$ und wählt man $\xi_k = |\alpha_k|^{q/p}$, so gilt Gleichheit. Was muss man dazu in den Fällen $p = 1, \infty$ wählen?

**Beispiel** (Integrationsfunktional)**.** Sei $X = C[0,1]$ und $y \in C[0,1]$. Dann ist durch

$$I_y(x) = \int_0^1 x(t)y(t)\,\mathrm{d}t$$

ein stetiges Funktional $I_y$ auf $C[0,1]$ definiert.

Tatsächlich ist

$$|I_y(x)| \leq \int_0^1 |x(t)|\,|y(t)|\,\mathrm{d}t \leq \int_0^1 |y(t)|\,\mathrm{d}t\,\|x\|_\infty.$$

Dies zeigt auch $\|I_y\| \leq \int_0^1 |y(t)|\,\mathrm{d}t$. Hier gilt sogar Gleichheit, was man einsieht, indem man zu $\varepsilon > 0$ die Funktion $x = \frac{\overline{y}}{|y|+\varepsilon}$ betrachtet. Dann ist nämlich $x$ stetig mit $\|x\|_\infty \leq 1$ und

$$|I_y(x)| = \int_0^1 \frac{|y(t)|^2}{|y(t)|+\varepsilon}\,\mathrm{d}t \geq \int_0^1 \frac{|y(t)|^2 - \varepsilon^2}{|y(t)|+\varepsilon}\,\mathrm{d}t = \int_0^1 \big(|y(t)| - \varepsilon\big)\,\mathrm{d}t = \int_0^1 |y(t)|\,\mathrm{d}t - \varepsilon.$$

Es folgt $\|I_y\| \geq \int_0^1 |y(t)|\,\mathrm{d}t - \varepsilon$, und zwar für jedes $\varepsilon > 0$.

Wir setzen fort mit einigen Beispielen von stetigen Operatoren. Wieder ist die Linearität in allen Beispielen offensichtlich.

**Beispiel** (Matrizen)**.** Jede $m \times n$-Matrix $A = (\alpha_{hk})$ mit Einträgen in $\mathbb{K}$ beschreibt eine lineare und damit stetige Abbildung $A : \mathbb{K}^n \to \mathbb{K}^m$. Dann ist $\big\|A : \ell_1^n \to \ell_\infty^m\big\| = \max_{h,k} |\alpha_{hk}|$. Tatsächlich ist für $x = (\xi_k) \in \mathbb{K}^n$

$$\|Ax\|_\infty = \max_{1\leq h\leq m} \left|\sum_{k=1}^n \alpha_{hk}\xi_k\right| \leq \max_{1\leq h\leq m} \left(\max_{1\leq k\leq n} |\alpha_{hk}|\right) \sum_{k=1}^n |\xi_k| = \max_{h,k} |\alpha_{hk}| \|x\|_1.$$

Für welches $x$ gilt hier Gleichheit?

**Beispiel** (Äquivalente Normen)**.** Seien $\|\cdot\|_1$ und $\|\cdot\|_2$ zwei Normen auf einem linearen Raum $X$. Dann sind die beiden Normen äquivalent genau dann, wenn die Identität $I$ von $X$ stetig ist als Abbildung $I : (X, \|\cdot\|_1) \to (X, \|\cdot\|_2)$ und als Abbildung $I : (X, \|\cdot\|_2) \to (X, \|\cdot\|_1)$. Tatsächlich ist dann $b = \|I : (X, \|\cdot\|_1) \to (X, \|\cdot\|_2)\|$ die kleinste Konstante, für die $\|x\|_2 \le b\|x\|_1$ und $a = \|I : (X, \|\cdot\|_2) \to (X, \|\cdot\|_1)\|^{-1}$ ist die größte Konstante, für die $a\|x\|_1 \le \|x\|_2$ für alle $x \in X$ gilt.

**Beispiel** (Integraloperatoren I)**.** Wir betrachten einen Kern $k \in C\big([0,1]^2\big)$ und definieren den zugehörigen *Fredholmschen Integraloperator* $T_k : C[0,1] \to C[0,1]$ durch

$$(T_k x)(s) = \int_0^1 k(s,t)x(t)\,\mathrm{d}t \qquad \text{für } x \in C[0,1].$$

Dann zeigen wir, dass $T_k$ ein stetiger Operator mit $\|T_k\| = \sup_{s\in[0,1]} \int_0^1 |k(s,t)|\,\mathrm{d}t$ ist (in Vorlesung).

**Beispiel** (Integraloperatoren II)**.** Ist $k \in C\big([0,1]^2\big)$ ein stetiger Kern, so zeigen ähnliche Argumente wie oben, dass der zugehörigen *Fredholmschen Integraloperator* $T_k$ auch wohldefiniert und stetig zwischen anderen Funktionenräumen ist, z.B. als Operator in $L_p[0,1]$ oder als Operator von $L_1[0,1]$ nach $C[0,1]$. Auch die Bedingungen an den Kern kann man ändern. Wir wollen jetzt einen Kern $k \in L_2\big([0,1]^2\big)$ betrachten und zeigen, dass $T_k : L_2[0,1] \to L_2[0,1]$ wohldefiniert und stetig ist. Dazu benötigen wir den Satz von Fubini und die Hölder-Ungleichung im Fall $p = q = 2$. (Beweis in Vorlesung).

**Beispiel** (Differentialoperatoren)**.** Eine weitere große Klasse von linearen Operatoren sind Differentialoperatoren. Wir betrachten hier den Ableitungsoperator $D : C^1[0,1] \to C[0,1], x \mapsto x'$. Versehen wir $C^1[0,1]$ mit der Norm $\|x\|_{C^1} = \|x\|_\infty + \|x'\|_\infty$, die $C^1[0,1]$ zu einem Banachraum macht, dann ist $D$ stetig wegen $\|x'\|_\infty \le \|x\|_{C^1}$. Versehen wir dagegen $C^1[0,1]$ mit der Norm $\|x\|_\infty$, dann ist $D$ nicht stetig. Das zeigen die Polynome $p_n(t) = t^n$, für die $\|p_n\|_\infty = 1$ und $\|p_n'\|_\infty = n$ ist.

## 2.3 Hauptsätze über stetige Operatoren

In diesem Abschnitt wollen wir einige der Hauptresultate über stetige Operatoren kennenlernen, die alle auf dem Satz von Baire beruhen. Diesen Satz (auch bekannt als Bairescher Kategoriensatz) kann man allgemeiner für metrische und auch topologische Räume formulieren und beweisen. Wir benötigen hier die folgende vereinfachte Version für Banachräume.

**Satz 2.14** (Satz von Baire)**.** *Ist $X$ ein Banachraum und ist $(A_n)$ eine Folge von abgeschlossenen Mengen, deren Vereinigung ganz $X$ ist, dann besitzt mindestens eine der Mengen $A_n$ einen inneren Punkt.*

*Beweis.* Nehmen wir an, dass die Behauptung nicht gilt, also keine der abgeschlossenen Mengen $A_n$ einen inneren Punkt besitzt. Dann schneidet jede der offenen Mengen $O_n = X \setminus A_n$ jede Kugel in $X$, diese Mengen sind *dicht* in $X$. Wir werden jetzt einen Punkt $x \in X$ konstruieren, der in allen $O_n$ liegt. Dann ist der Durchschnitt der $O_n$ nichtleer und damit die Vereinigung der $A_n$ nicht ganz $X$, im Widerspruch zur Voraussetzung.

Dazu bezeichnen wir mit $U_\varepsilon(x)$ bzw. $B_\varepsilon(x)$ die offene bzw. abgeschlossene Kugel mit Radius $\varepsilon$ und Mittelpunkt $x$. Da $X$ ein normierter Raum ist, können wir zu jeder offenen Menge $O$ und jedem $x \in O$ ein $\varepsilon$ finden, so dass sogar $B_\varepsilon(x) \subseteq O$ gilt. Dieses Argument werden wir jetzt mehrfach anwenden.

Wir wählen zunächst ein $x_1 \in O_1$ und ein $\varepsilon_1 > 0$, so dass die abgeschlossenen Kugel $B_{\varepsilon_1}(x_1) \subseteq O_1$ ist. Das ist möglich, da $O_1$ offen ist. Nun ist $V_1 = U_{\varepsilon_1}(x_1) \cap O_2$ offen und nach Voraussetzung nichtleer, wir können also $x_2 \in V_1$ und $\varepsilon_2 < \varepsilon_1/2$ wählen, so dass die abgeschlossenen Kugel $B_{\varepsilon_2}(x_2) \subseteq V_1$ ist. Nun ist $V_2 = U_{\varepsilon_2}(x_2) \cap O_3$ offen und nach Voraussetzung nichtleer, wir können also $x_3 \in V_2$ und $\varepsilon_3 < \varepsilon_2/2$ wählen, so dass die abgeschlossenen Kugel $B_{\varepsilon_3}(x_3) \subseteq V_2$ ist. Dieses Verfahren setzen wir induktiv fort und erhalten eine Folge $(x_n)$ in $X$ und eine Nullfolge $\varepsilon_n$, so dass $B_{\varepsilon_n}(x_n) \subseteq U_{\varepsilon_{n-1}}(x_{n-1})$ und damit

$$\cdots \subseteq U_{\varepsilon_n}(x_n) \subseteq U_{\varepsilon_{n-1}}(x_{n-1}) \subseteq \cdots \subseteq U_{\varepsilon_1}(x_1)$$

gilt. Die Folge $(x_n)$ ist also eine Cauchyfolge und damit konvergent, da $X$ vollständig ist. Sei $x$ der Grenzwert. Nach Konstruktion ist dann $x \in B_{\varepsilon_n}(x_n) \subseteq O_n$ für alle $n$ und der Beweis ist beendet. □

Aus dem Satz von Baire folgt direkt eines der Hauptresultate für stetige Operatoren, das Prinzip der gleichmäßigen Beschränktheit. Erstaunlicherweise kann man aus der punktweisen Beschränktheit einer Operatorenmenge eine gleichmäßige Schranke für die Norm der Operatoren ableiten.

**Satz 2.15** (Prinzip der gleichmäßigen Beschränktheit)**.** *Ist $X$ ein Banachraum, $Y$ ein normierter Raum und $\mathcal{M} \subseteq L(X,Y)$ eine Familie von stetigen Operatoren, die punktweise beschränkt sind, d.h. für jedes $x \in X$ ist $\{\|Tx\| : T \in \mathcal{M}\}$ eine beschränkte Menge in $\mathbb{R}$. Dann ist $\{\|T\| : T \in \mathcal{M}\}$ beschränkt.*

*Beweis.* Wir setzen

$$A_n = \{x \in X : \|Tx\| \leq n \text{ für alle } T \in \mathcal{M}\}.$$

Da die Operatoren $T \in \mathcal{M}$ stetig sind, ist $A_n$ abgeschlossen. Die Voraussetzung besagt gerade, dass die Vereinigung der $A_n$ ganz $X$ ist. Aus dem Satz von Baire folgt, dass ein $A_n$ einen inneren Punkt enthält. Es gibt also $n \in \mathbb{N}$, $x_0 \in X$ und $\varepsilon > 0$, so dass für alle $x \in X$ mit $\|x - x_0\| < \varepsilon$ die Ungleichung $\|Tx\| \leq n$ für alle $T \in \mathcal{M}$ gilt. Insbesondere folgt für alle $x \in X$ mit $\|x\| < \varepsilon$ und alle $T \in \mathcal{M}$ die Ungleichung

$$\|Tx\| \leq \|T(x + x_0)\| + \|Tx_0\| \leq n + n = 2n$$

und damit für alle $x \in X$ mit $\|x\| < 1$

$$\|Tx\| \leq \frac{2n}{\varepsilon}.$$

Das bedeutet aber

$$\|T\| \leq \frac{2n}{\varepsilon} \qquad \text{für alle } T \in \mathcal{M}.$$

□

Eine oft benutzte Variante für Folgen von Operatoren ist der

**Satz 2.16** (Satz von Banach-Steinhaus)**.** *Ist $X$ ein Banachraum, $Y$ ein normierter Raum und $T_n \in L(X,Y)$ eine Folge von stetigen Operatoren, die punktweise konvergiert, d.h. für jedes $x \in X$ ist existiert $\lim T_n x =: Tx$. Dann ist $T$ ein stetiger Operator mit $\|T\| \leq \sup_n \|T_n\| < \infty$.*

*Beweis.* Die Linearität von $T$ ist offensichtlich. Wir setzen $\mathcal{M} = \{T_n : n \in \mathbb{N}\}$. Da $(T_n x)_n$ für jedes $x$ konvergiert, ist $\{\|Tx\| : T \in \mathcal{M}\} = \{\|T_n x\| : n \in \mathbb{N}\}$ für jedes $x$ beschränkt. Das Prinzip von der gleichmäßigen Beschränktheit liefert nun $M = \sup_n \|T_n\| < \infty$. Für jedes $x \in X$ folgt

$$\|Tx\| = \lim_n \|T_n x\| \le \sup_n \|T_n\| \, \|x\| = M\|x\|$$

und damit die Stetigkeit von $T$ samt Ungleichung $\|T\| \le M$. □

Die Vollständigkeit von $X$ ist im Satz von Banach-Steinhaus und damit auch im Satz von der gleichmäßigen Beschränktheit eine wesentliche Voraussetzung. Ist z.B. $X = d$ der Raum der abbrechenden Folgen mit der Supremumsnorm, dann sind die Funktionale $T_n : d \to \mathbb{K}$ gegeben durch $T_n x = n\xi_n$ für $x = (\xi_k)$ stetig mit $\|T_n\| = n$, die Folge der Normen $(\|T_n\|)$ ist also unbeschränkt. Für jedes $x \in d$ ist aber $\lim T_n x = \lim \xi_n = 0$, also konvergiert $T_n$ punktweise gegen $T = 0$.

Der zweite Hauptsatz über lineare stetige Operatoren ist der Satz von der offenen Abbildung. Dazu benötigen wir zunächst die

**Definition 2.17.** Eine Abbildung zwischen metrischen (oder topologischen) Räumen, die offene Mengen auf offene Mengen abbildet, heißt *offen*.

Diese Definition ist natürlich maßgeschneidert, um die Stetigkeit der inversen Abbildung zu überprüfen, da eine Abbildung ja stetig ist, wenn Urbilder offener Mengen offen sind. In dieser Definition der Stetigkeit kann man natürlich offene Mengen durch abgeschlossene Mengen ersetzen, da das Komplement von des Urbilds einer Menge das Urbild des Komplements der Menge ist. Dies gilt nicht mehr für die Offenheit einer Abbildung. Ein Gegenbeispiel ist die Koordinatenabbildung $p : (x, y) \mapsto x$ auf $\mathbb{R}^2$. Diese ist offen, aber es gibt abgeschlossene Mengen $A \subseteq \mathbb{R}^2$, für die $p(A)$ nicht abgeschlossen ist (Übung).

Ist $T : X \to Y$ eine lineare Abbildung zwischen normierten Räumen, dann ist die Offenheit $T$ äquivalent dazu, dass $T$ offene Kugeln um 0 (oder auch nur die offene Einheitskugel) in Nullumgebungen abbildet. Bezeichnen wir wieder mit $U_\varepsilon$ die offene Kugel mit Radius $\varepsilon$ und Mittelpunkt 0, so haben wir also die äquivalenten Bedingungen

(a) $T$ ist offen.

(b) Zu jedem $\delta > 0$ existiert ein $\varepsilon > 0$ mit $U_\varepsilon \subseteq T(U_\delta)$.

(c) Es existiert ein $\varepsilon > 0$ mit $U_\varepsilon \subseteq T(U_1)$.

Hier sind die Implikationen $(a) \Rightarrow (b)$ und $(b) \Rightarrow (c)$ klar und der Beweis von $(c) \Rightarrow (a)$ ist einfach (Vorlesung).

Eine offene lineare Abbildung $T : X \to Y$ ist surjektiv, da $T(X)$ ein linearer Raum ist und die lineare Hülle einer Nullumgebung in $Y$ ganz $Y$ ist. Der Satz von der offenen Abbildung liefert die Umkehrung dieser Aussage im Fall von Banachräumen.

Für den Beweis benötigen wir das folgende auch sonst oft nützliche Vollständigkeitskriterium. Dazu nennen wir eine Reihe $\sum x_k$ mit $x_k \in X$ absolut konvergent, wenn $\sum \|x_k\| < \infty$ ist.

**Satz 2.18.** *Ein normierter Raum $X$ ist vollständig, wenn jede absolut konvergente Reihe in $X$ auch in $X$ konvergiert.*

*Beweis.* Sei zunächst jede absolut konvergente Reihe in $X$ konvergent und sei eine Cauchyfolge gegeben. Wir zeigen die Konvergenz einer Teilfolge und damit ist, wie schon früher beim Beweis der Vollständigkeit von $L_p$ benutzt, die gesamte Folge konvergent. Dazu wählen wir eine Teilfolge $(x_n)$ der Ausgangsfolge mit $\|x_{n+1} - x_n\| < 2^{-n}$. Dann ist wegen $\sum_{\|} x_{k+1} - x_k\| < \infty$ die Reihe $\sum_{k=1}^{\infty}(x_{k+1} - x_k)$ absolut konvergent und damit nach Voraussetzung konvergent. Die Partialsummen dieser Reihe sind aber $(x_n - x_1)$, also ist diese Folge und damit auch die Folge $(x_n)$ konvergent.

Sei nun umgekehrt $X$ vollständig und $\sum x_k$ absolut konvergent. Dann bilden die Partialsummen wegen

$$\left\|\sum_{k=n+1}^{m} x_k\right\| \leq \sum_{k=n+1}^{m} \|x_k\|$$

eine Cauchyfolge und konvergieren damit in $X$. □

**Satz 2.19** (Satz von der offenen Abbildung)**.** *Sind $X$ und $Y$ Banachräume und ist $T \in L(X,Y)$ surjektiv, dann ist $T$ offen.*

*Beweis.* Wir zeigen (c) oben, also die Existenz eines $\varepsilon > 0$ mit $U_\varepsilon \subseteq T(U_1)$. Der Beweis erfolgt in zwei Schritten. Zunächst zeigen wir mit dem Satz von Baire und der Vollständigkeit von $Y$ die Existenz eines $\varepsilon > 0$ mit $U_\varepsilon \subseteq \overline{T(U_1)}$. Anschließend folgern wir mit der Vollständigkeit von $X$ daraus $U_\varepsilon \subseteq T(U_1)$.

Für den ersten Schritt benutzen wir die abgeschlossenen Mengen $A_n = \overline{T(U_n)}$, deren Vereinigung ganz $Y$ ist, da $T$ ja surjektiv ist. Also enthält ein $A_n$ eine offene Kugel $U_\varepsilon(y_0)$. Da $A_n$ symmetrisch ist, d.h. $-A_n = \{-y \,:\, y \in A_n\} = A_n$ gilt, enthält $A_n$ auch die offene Kugel $U_\varepsilon(-y_0)$. Da $A_n$ auch konvex ist, enthält $A_n$ mit den beiden Kugeln $U_\varepsilon(y_0)$ und $U_\varepsilon(-y_0)$ auch $U_\varepsilon$. Also ist $U_\varepsilon \subseteq \overline{T(U_n)}$ und damit $U_{\varepsilon/N} \subseteq \overline{T(U_1)}$. Damit ist der erste Schritt gezeigt.

Für den zweiten Schritt betrachten wir o.B.d.A. $\varepsilon = 1$, indem wir $T$ durch $T/\varepsilon$ ersetzen. Unsere Voraussetzung ist nun also $U_1 \subseteq \overline{T(U_1)}$, zeigen wollen wir $U_1 \subseteq T(U_1)$. Sei also $y \in U_1$, d.h. $\|y\| < 1$. Wir werden $y$ induktiv durch Elemente aus $T(U_1)$ approximieren. Wir wählen zunächst $\varepsilon < 1$ mit $\|y\| < \varepsilon$ und $0 < \alpha < 1$ mit $\varepsilon < 1 - \alpha$ und setzen $z = y/\varepsilon$. Wegen $\|z\| < \|y\|/\varepsilon < 1$ ist $z \in U_1 \subseteq \overline{T(U_1)}$ und wir finden ein $y_0 = Tx_0 \in T(U_1)$ mit

$$\|z - y_0\| < \alpha.$$

Wegen $\|(z - y_0)/\alpha\| < 1$ ist finden wir ein $y_1 = Tx_1 \in T(U_1)$ mit

$$\left\|\frac{z - y_0}{\alpha} - y_1\right\| < \alpha,$$

also

$$\|z - y_0 - \alpha y_1\| < \alpha^2.$$

Anschließend finden wir $y_2 = Tx_2 \in T(U_1)$ mit

$$\|z - y_0 - \alpha y_1 - \alpha^2 y_2\| < \alpha^3.$$

Dieses Verfahren setzen wir induktiv fort und finden eine Folge $(x_n)$ mit $x_n \in U_1)$ und

$$\left\|z - T\sum_{k=0}^{n} \alpha^k x_k\right\| < \alpha^{n+1}.$$

Wegen $\alpha < 1$ konvergiert $\sum_{k=0}^{\infty} \|\alpha^k x_k\| \le \sum_{k=0}^{\infty} \alpha^k = \frac{1}{1-\alpha}$. Aus der Vollständigkeit von $X$ folgt die Konvergenz der Reihe

$$x = \sum_{k=0}^{\infty} \alpha^k x_k$$

in $X$ und $\|x_0\| \le \frac{1}{1-\alpha}$ und $z = Tx$. Damit ist schließlich $y = \varepsilon z = T(\varepsilon x)$ und $\|\varepsilon x\| \le \frac{\varepsilon}{1-\alpha} < 1$ und $y \in T(U_1)$ gezeigt. $\square$

Die beiden nächsten Sätze sind wichtige direkte Folgerungen aus dem Satz von der offenen Abbildung.

**Satz 2.20** (Satz von der stetigen Inversen)**.** *Sind $X$ und $Y$ Banachräume und ist $T \in L(X,Y)$ bijektiv, dann ist $T^{-1}$ stetig.*

**Satz 2.21** (Äquivalente Normen)**.** *Sind auf einem linearen Raum $X$ zwei Normen $\|\cdot\|_1$ und $\|\cdot\|_2$ gegeben, die beide $X$ zu einem Banachraum machen und für die es eine Konstante $C > 0$ gibt mit*

$$\|x\|_2 \le C\|x\|_1 \quad \textit{für alle } x \in X,$$

*dann sind die beiden Normen äquivalent.*

## 2.4 Hilberträume

Hilberträume sind Banachräume, deren Norm aus einem Skalarprodukt erzeugt wird. Neben der Abstandsmessung im normierten Raum kann man mit einem Skalarprodukt auch Winkel messen.

**Definition 2.22.** Sei $H$ ein linearer Raum über dem Körper $\mathbb{K}$. Eine Funktion $\langle \cdot, \cdot \rangle : X \times X \to \mathbb{K}$ mit den Eigenschaften

(a) *Linearität:* $\langle \lambda x + \mu y, z\rangle = \lambda\langle x, z\rangle + \mu\langle y, z\rangle$ für $x, y, z \in H$ und $\lambda, \mu \in K$

(b) *Symmetrie:* $\langle x, y\rangle = \overline{\langle y, x\rangle}$ für $x, y \in H$

(c) *positive Definitheit:* $\langle x, x\rangle \ge 0$ für $x \in H$ und $\langle x, x\rangle = 0 \Leftrightarrow x = 0$

heißt *Skalarprodukt* oder *inneres Produkt* auf $H$. Dann heißt $H$ mit diesem Skalarprodukt *Prä-Hilbertraum.*

Aus den Eigenschaften (a) und (b) folgt direkt

(a') $\langle x, \lambda y + \mu z\rangle = \overline{\lambda}\langle x, y\rangle + \overline{\mu}\langle x, z\rangle$ für $x, y, z \in H$ und $\lambda, \mu \in K$

Im reellen Fall ist ein Skalarprodukt also eine positiv definite Bilinearform, im komplexen Fall muss man Skalare aus dem zweiten Faktor als konjugiert komplexe Zahl herausziehen. Hier spricht man von einer *Sesquilinearform.*

**Beispiel.** Aus der linearen Algebra ist das Skalarprodukt

$$\langle x, y\rangle = \sum_{k=1}^{n} \xi_k \overline{\eta_k} \quad \text{für } x = (\xi_k), y = (\eta_k) \in \mathbb{K}^n$$

auf $\mathbb{K}^n$ bekannt. Die euklidische Norm erhält man als $\|x\|_2 = \sqrt{\langle x, x\rangle}$. Das war der Raum $\ell_2^n$.

**Beispiel.** Auf dem Folgenraum $\ell_2$ ist durch

$$\langle x, y\rangle = \sum_{k=1}^{\infty} \xi_k \overline{\eta_k} \quad \text{für } x = (\xi_k), y = (\eta_k) \in \ell_2$$

ein Skalarprodukt definiert. Die Konvergenz der Reihe folgt aus der Hölder-Ungleichung im Fall $p = q = 2$. Die Norm in $\ell_2$ erhält man als $\|x\|_2 = \sqrt{\langle x, x\rangle}$.

**Beispiel.** Allgemeiner sei $(\Omega, \mathcal{A}, \mu)$ ein fester Maßraum. Dann ist auf $L_2(\mu)$ ein Skalarprodukt definiert durch

$$\langle x, y\rangle = \int_\Omega x(t)\overline{y(t)}\, \mathrm{d}\mu(t)$$

ein Skalarprodukt definiert. Die Endlichkeit des Integrals folgt wieder aus der Hölder-Ungleichung im Fall $p = q = 2$. Die Norm in $L_2(\mu)$ erhält man als $\|x\|_2 = \sqrt{\langle x, x\rangle}$.

Ist nun $H$ ein Prä-Hilbertraum, so setzen wir in Analogie zu diesen Beispielen $\|x\| = \sqrt{\langle x, x\rangle}$. Offensichtlich ist $\|\lambda x\| = |\lambda|\|x\|$ für $x \in H$ und $\lambda \in \mathbb{K}$ und $\|x\| \geq 0$ und $\|x\| = 0 \iff x = 0$. Wir werden gleich einsehen, dass auch die Dreiecksungleichung erfüllt ist und dies somit eine Norm auf $H$ definiert. Dazu brauchen wir die

**Satz 2.23** (Cauchy-Schwarz-Ungleichung). *In jedem Prä-Hilbertraum $H$ gilt*

$$\big|\langle x, y\rangle\big| \leq \|x\|\,\|y\| \quad \textit{für alle } x, y \in H.$$

*Beweis.* Für $y = 0$ ist nichts zu zeigen. Für den Fall $y \neq 0$ können wir uns wegen der Homogenität auf den Fall $\|y\| = 1$ beschränken. Dann folgt

$$0 \leq \big\langle x - \langle x, y\rangle y, x - \langle x, y\rangle y\big\rangle = \langle x, x\rangle - |\langle x, y\rangle|^2 - |\langle x, y\rangle|^2 + |\langle x, y\rangle|^2 = \|x\|^2 - |\langle x, y\rangle|^2$$

und damit die Behauptung $|\langle x, y\rangle| \leq \|x\|^2$. □

**Satz 2.24** (Dreiecksungleichung). *In jedem Prä-Hilbertraum $H$ gilt*

$$\|x + y\| \leq \|x\| + \|y\| \quad \textit{für alle } x, y \in H.$$

*Beweis.* Aus der Cauchy-Schwarz-Ungleichung folgt

$$\|x + y\|^2 = \langle x + y, x + y\rangle = \|x\|^2 + 2\mathrm{Re}\,\langle x, y\rangle + \|y\|^2 \leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 = \big(\|x\| + \|y\|\big)^2$$

und damit die Behauptung. □

Ein Prä-Hilbertraum ist also ein normierter Raum. Einen vollständigen Prä-Hilbertraum nennt man *Hilbertraum*.

**Beispiel.** $\ell_2^n$, $\ell_2$, $L_2(\mu)$ aus den obigen Beispielen sind Hilberträume. Betrachtet man das $L_2[0,1]$-Skalarprodukt eingeschränkt auf $C[0,1]$, so erhält man einen Prä-Hilbertraum, aber keinen Hilbertraum.

**Satz 2.25** (Parallelogrammgleichung). *Ist $H$ ein Prä-Hilbertraum, dann gilt*

$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2 \quad \textit{für alle } x, y \in H.$$

*Beweis.* Dies rechnet man einfach nach. □

Man kann auch zeigen, dass ein normierter Raum, in dem die Parallelogrammgleichung gilt, ein Prä-Hilbertraum ist, die Norm also aus einem Skalarprodukt erzeugt wird.

**Satz 2.26** (Stetigkeit des Skalarprodukts)**.** *In jedem Prä-Hilbertraum $H$ ist für jedes $x \in H$ das lineare Funktional $a_x : y \mapsto \langle x, y\rangle$ stetig auf $H$ mit $\|a_x\| = \|x\|$. Ebenso ist für jedes $y \in H$ das lineare Funktional $b_y : x \mapsto \langle x, y\rangle$ stetig auf $H$ mit $\|b_y\| = \|y\|$..*

*Beweis.* Wir zeigen die Aussage für das Funktional $b_y$ für festes $y \in H$. Für $y = 0$ ist nichts zu zeigen, also sei $y \neq 0$. Die Linearität von $b_y$ ist klar, die Stetigkeit folgt aus der Cauchy-Schwarz-Ungleichung

$$|b_y(x)| = \big|\langle x, y\rangle\big| \leq \|y\|\|x\|.$$

Dies zeigt auch $\|b_y\| \leq \|y\|$. Setzt man $x = y/\|y\|$, so erhält man auch $\|b_y\| \geq \langle x, y\rangle = \|y\|$. □

Das folgende Lemma zeigt eine Verschärfung der Definitheit des Skalarprodukts. Wir erinnern dazu an den Begriff der Dichtheit in metrischen Räumen. Eine Teilmenge $M$ eines metrischen Raumes $X$ heißt *dicht* in $X$, wenn jede Kugel in $X$ einen Punkt von $M$ enthält. Offenbar ist $M$ dicht in $X$ genau dann, wenn $\overline{M} = X$ gilt. Ist $M$ dicht und abgeschlossen, dann ist $M = X$.

**Lemma 2.27.** *Sei $H$ ein Prä-Hilbertraum, sei $U \subseteq H$ dicht und sei $x \in H$. Ist dann $\langle x, u\rangle = 0$ für alle $u \in U$, so ist $x = 0$.*

*Beweis.* Der lineare Teilraum $Y = \{y \in H : \langle x, y\rangle = 0\}$ ist als Urbild $a_x^{-1}(\{0\})$ des stetigen linearen Funktionals $a_x : y \mapsto \langle x, y\rangle$ abgeschlossen und enthält die Menge $U$. Da $U$ dicht ist, ist $Y$ ebenfalls dicht. Also ist $Y = H$ und die Behauptung $x = 0$ folgt wegen $x \in Y$ aus der Definitheit des Skalarprodukts. □

Wichtig am Skalarprodukt ist, dass man damit einen sinnvollen Begriff der Orthogonalität definieren kann. Dazu erinnern wir daran, dass im $\mathbb{K}^n$ zwei Vektoren orthogonal sind, wenn ihr Skalarprodukt 0 ist.

**Definition 2.28** (Orthogonalität, orthogonales Komplement)**.** Sei $H$ ein Prä-Hilbertraum. Dann heißen $x, y \in H$ *orthogonal* und wir schreiben $x \perp y$, wenn $\langle x, y\rangle = 0$ ist. Zwei Teilmengen $A, B \subseteq H$ heißen *orthogonal* und wir schreiben $A \perp B$, wenn $x \perp y$ für alle $x \in A$ und $y \in B$ ist. Die Menge

$$A^\perp := \{y \in H : x \perp y = 0 \text{ für alle } x \in A\}$$

heißt *orthogonales Komplement* von $A$.

Die folgenden Eigenschaften sind (fast) offensichtlich:

(a) *Satz des Pythagoras:* Sind $x, y$ orthogonal, dann gilt $\|x + y\|^2 = \|x\|^2 + \|y\|^2$.

(b) *Monotonie:* $A \subseteq B \Leftrightarrow B^\perp \subseteq A^\perp$

(c) Das orthogonale Komplement $A^\perp$ ist immer ein abgeschlossener linearer Teilraum von $H$.

(d) $A \subseteq (A^\perp)^\perp$

(e) Das orthogonale Komplement der abgeschlossenen linearen Hülle von $A$ stimmt mit dem orthogonalen Komplement von $A$ überein: $A^\perp = \big(\overline{\operatorname{span} A}\big)^\perp$.

In der Approximationstheorie ist der folgende Satz wichtig, in dem die Vollständigkeit eine zentrale Voraussetzung ist. Wir benötigen diesen Satz für den Beweis des anschließenden Projektionssatzes.

**Satz 2.29** (Existenz und Eindeutigkeit der Bestapproximation). *Sei $H$ ein Hilbertraum und $U$ ein abgeschlossener linearer Teilraum von $H$. Dann existiert zu jedem $x \in H$ ein eindeutig bestimmtes Element $u \in U$ mit*

$$\|x-u\| = \inf_{v \in U} \|x-v\| =: \operatorname{dist}(x,U).$$

*Außerdem ist $x-u \perp U$. Man nennt $u$ die* Bestapproximation *zu $x$ in $U$.*

*Beweis.* Zum Beweis der Existenz setzen wir $d = \operatorname{dist}(x,U) = \inf_{v\in U} \|x-v\|$ und wählen eine Folge $(u_n)$ in $U$ mit $\|x-u_n\| \to d$. Nun gilt für $\varepsilon > 0$ wegen der Parallelogrammgleichung

$$\begin{aligned} 4d^2 + \|u_n-u_m\|^2 &\le 4\|(u_n+u_m)/2 - x\|^2 + \|u_n-u_m\|^2 = \|u_n+u_m-2x\|^2 + \|u_n-u_m\|^2 \\ &= 2\|u_n-x\|^2 + 2\|u_m-x\|^2 \le 4d^2 + 4\varepsilon^2, \end{aligned}$$

wobei die letzte Ungleichung für $m,n \ge n_0$ gilt. Es folgt $\|u_n-u_m\| \le 2\varepsilon$ für für $m,n \ge n_0$ gilt, also ist $(u_n)$ eine Cauchyfolge. Da $H$ vollständig ist, konvergiert $(u_n)$, sagen wir $u_n \to u \in H$. Da $U$ abgeschlossen ist, ist $u \in U$. Nach Konstruktion ist $d = \lim \|x-u_n\| = \|x-u\|$.

Nun folgt die Eindeutigkeit der Bestapproximation folgendermaßen. Sind $u, v$ Bestapproximationen, so ist nach dem gerade gezeigten die Folge $(u,v,u,v,u,v,\dots)$ eine Cauchyfolge, also muss $u = v$ sein.

Wir zeigen schließlich noch $x-u \perp U$. Dazu wählen wir $v \in U$ beliebig und setzen $y = x-u$. Dann ist $\|y\| \le \|y-tv\|$ für jedes $t \in \mathbb{R}$ wegen der Bestapproximationseigenschaft. Quadrieren und Zusammenfassen liefert die Ungleichung $2t\operatorname{Re}\langle z,u\rangle \le t^2\|u\|^2$ für alle $t \in \mathbb{R}$. Dies kann aber nur im Fall $\operatorname{Re}\langle z,u\rangle = 0$ gelten. Benutzt man nun (im komplexen Fall) $it$ statt $t$ in diesem Argument, dann erhält man auch $\operatorname{Im}\langle z,u\rangle = 0$ und damit wie behauptet $\langle z,u\rangle = 0$. □

Die Abbildung $P_U : H \to U$, die $x \in H$ die Bestapproximation $u = P_U(x) = P_U x$ zuordnet, heißt *Orthogonalprojektion* von $H$ auf $U$. Es ist eine Projektion, da für $x \in U$ offensichtlich $P_U x = x$ gilt und deshalb $P_U^2 = P_U$ ist. Außerdem ist $P_U$ eine lineare Abbildung. Dabei ist $P_U(\lambda x) = \lambda P_U(x)$ klar. Für die Linearität müssen wir noch $P_U(x+y) = P_U x + P_U y$ zeigen. Für jedes $v \in U$ ist nun wegen der Orthogonalität $x - P_U x \perp v$ und $y - P_U y \perp v$ auch $(x+y) - (P_U x + P_U y) \perp v$ und der Pythagoras impliziert

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

Also ist $P_U x + P_U y$ tatsächlich die Bestapproximation $P_U(x+y)$ zu $x+y$ in $U$.

**Satz 2.30** (Projektionssatz). *Sei $H$ ein Hilbertraum und $U$ ein abgeschlossener linearer Teilraum von $H$. Dann ist $H$ die direkte (und orthogonale) Summe $H = U \oplus U^\perp$, d.h zu jedem $x \in H$ existieren eindeutig bestimmte Elemente $u \in U$ und $u^\perp \in U^\perp$ mit $x = u + u^\perp$. Hierbei ist $u = P_U x$ und $u^\perp = P_{U^\perp} x$ und $\|x\|^2 = \|u\|^2 + \|u^\perp\|^2$.*

*Beweis.* Die Direktheit der Summe $U \oplus U^\perp$ und damit die Eindeutigkeit der Zerlegung $x = u + u^\perp$ folgt, da aus $u \in U \cap U^\perp$ sofort $\langle u, u\rangle = 0$ und damit $u = 0$ folgt. Für die Existenz setzen wir natürlich $u = P_U x$ und $u^\perp = x - u$ und erhalten $x = u + u^\perp$ und $u^\perp \in U^\perp$. Die Gleichung $\|x\|^2 = \|u\|^2 + \|u^\perp\|^2$ ist dann nichts anderes als der Pythagoras. Dass auch $u^\perp = P_{U^\perp} x$ gilt, folgt wegen $\big(U^\perp\big)^\perp = U$, was wir im folgenden Satz etwas allgemeiner zeigen. □

**Satz 2.31.** *Sei $H$ ein Hilbertraum und $U$ ein linearer Teilraum von $H$. Dann gilt $\big(U^\perp\big)^\perp = \overline{U}$. Insbesondere ist $U$ dicht, wenn $U^\perp = \{0\}$ ist.*

*Beweis.* Die Inklusion $\big(U^\perp\big)^\perp \supseteq \overline{U}$ ist klar, da $\big(U^\perp\big)^\perp$ abgeschlossen ist und $\big(U^\perp\big)^\perp \supseteq U$ gilt. Zum Beweis der Inklusion $\big(U^\perp\big)^\perp \subseteq \overline{U}$ schreiben wir $x \in \big(U^\perp\big)^\perp = \big(\overline{U}^\perp\big)^\perp \subseteq H$ als $x = v + v^\perp$ mit $v \in \overline{U}$ und $v^\perp \in \overline{U}^\perp$ und erhalten nach Definition $\langle x, v^\perp\rangle = 0 = \langle v, v^\perp\rangle$, woraus schließlich $\langle v^\perp, v^\perp\rangle = 0$ und damit $v^\perp = 0$ und $x = v \in \overline{U}$ folgen. □

Mittels des Projektionssatzes charakterisieren wir jetzt den Dualraum $H'$ der stetigen linearen Funktionale auf $H$. Dieser sogenannte Darstellungssatz von Fréchet-Riesz zeigt, dass man $H'$ mit $H$ identifizieren kann, dass jedes stetige lineare Funktional von der Form $x \mapsto \langle x, y\rangle$ ist.

**Satz 2.32** (Darstellungssatz von Fréchet-Riesz)**.** *Sei $H$ ein Hilbertraum und $a \in H'$ ein stetiges lineares Funktional auf $H$. Dann gibt es ein eindeutig bestimmtes $y \in H$ mit $a(x) = \langle x, y\rangle$ für alle $x \in H$. Weiter gilt $\|a\| = \|y\|$.*

*Beweis.* Wir setzen o.B.d.A. $\|a\| = 1$ voraus. Sei $U = a^{-1}\big(\{0\}\big)$ der Kern von $a$. Dann hat $U$ die Kodimension 1 und damit ist $U^\perp$ eindimensional, sagen wir $U^\perp = \operatorname{span}\big(\{z\}\big)$ für ein $y \in H$ mit $a(z) = 1$. Dann lässt sich nach dem Projektionssatz jedes $x$ eindeutig schreiben als $x = u + \lambda z$ mit $u \in U$. Dann ist $a(x) = \lambda a(z) = \lambda$ und $\langle x, z\rangle = \lambda\|z\|^2$. Mit $y = z/\|z\|^2$ erhalten wir also tatsächlich $a(x) = \langle x, y\rangle$. Die Normgleichung $\|a\| = \|y\|$ haben wir schon weiter oben gezeigt. Die Eindeutigkeit folgt, da für $y_1, y_2 \in H$ mit $a(x) = \langle x, y_1\rangle = \langle x, y_2\rangle$ für alle $x \in H$ sofort $\langle x, y_1 - y_2\rangle = 0$ für alle $x \in H$ und damit $y_1 = y_2$ ist. □

Als erste prominente Anwendung des Darstellungssatzes von Fréchet-Riesz präsentieren wir den Satz von Radon-Nikodým, der Maße mit Dichten als absolutstetige Maße charakterisiert. Wir erinnern zunächst an Maße mit Dichten aus der Maßtheorie. Ist $\mu$ ein Maß auf einem messbaren Raum $(\Omega, \mathcal{A})$ und ist $g \geq 0$ eine messbare Funktion auf $\Omega$, dann ist durch den Ansatz $\nu : A \mapsto \int_A g \,\mathrm{d}\mu$ ein Maß (bezüglich derselben $\sigma$-Algebra $\mathcal{A}$) gegeben. Dann gilt offenbar die Implikation

$$\mu(A) = 0 \Rightarrow \nu(A) = 0 \quad \text{für jedes } A \in \mathcal{A}.$$

Ein Maß $\nu$ mit dieser Eigenschaft heißt *absolutstetig* bezüglich des Maßes $\mu$ und man schreibt $\nu \ll \mu$. Für $\sigma$-endliches $\mu$ und $\nu$ sind die Maße $\nu$ mit Dichten die einzigen absolutstetigen Maße bezüglich $\mu$.

**Satz 2.33** (Satz von Radon-Nikodým)**.** *Seien $\mu$ und $\nu$ zwei $\sigma$-endliche Maße auf einem messbaren Raum $(\Omega, \mathcal{A})$ und sei $\nu \ll \mu$. Dann existiert eine messbare Funktion $g : \Omega \to [0, \infty)$ mit $\nu(A) = \int_A g \,\mathrm{d}\mu$ für $A \in \mathcal{A}$, d.h. $\nu$ besitzt eine Dichte bezüglich $\mu$.*

*Beweis.* Es genügt, die Behauptung für endliches $\mu$ und $\nu$ zu zeigen. Mittels disjunkter Zerlegungen in Mengen mit endlichem Maß überträgt man das Ergebnis auf $\sigma$-endliche Maße.

Seien also $\mu$ und $\nu$ endliche Maße mit $\nu \ll \mu$ und sei $\sigma = \mu + \nu$ das (endliche) Summenmaß. Auf dem Hilbertraum $L_2(\sigma)$ betrachten wir das lineare Funktional

$$a(x) = \int_\Omega x \, \mathrm{d}\mu \qquad \text{für } x \in L_2(\sigma).$$

Wir benutzen nun die Inklusionen und Ungleichungen für $L_p$-Räume, die in Aufgabe 43 in Übungsserie 7 gezeigt wurden:

Ist $\mu$ ein endliches Maß auf einer Menge $\Omega$ und sind $1 \leq p < q \leq \infty$, dann gilt $\|x\|_p \leq \mu(\Omega)^{1/p-1/q}\|x\|_q$ für $x \in L_q(\mu)$ und damit die Inklusion $L_q(\mu) \subseteq L_p(\mu)$.

Es folgt $L_2(\sigma) \subseteq L_2(\mu) \subseteq L_1(\mu)$ und

$$|a(x)| \leq \|x\|_{L_1(\mu)} \leq \mu(\Omega)^{1/2}\|x\|_{L_2(\mu)} \leq \mu(\Omega)^{1/2}\|x\|_{L_2(\sigma)}.$$

Also ist $a$ ein stetiges lineares Funktional auf $L_2(\sigma)$. Der Darstellungssatz von Fréchet-Riesz liefert nun eine Funktion $h \in L_2(\sigma)$ mit

$$a(x) = \int_\Omega x \, \mathrm{d}\mu = \int_\Omega xh \, \mathrm{d}\sigma \qquad \text{für alle } x \in L_2(\sigma).$$

Wegen der Endlichkeit von $\sigma$ sind beschränkte messbare Funktionen in $L_2(\sigma)$ und wir erhalten für $A \in \mathcal{A}$

$$\mu(A) = \int_\Omega \chi_A \, \mathrm{d}\mu = \int_\Omega \chi_A h \, \mathrm{d}\sigma = \int_A h \, \mathrm{d}\sigma \tag{5}$$

sowie

$$\nu(A) = \sigma(A) - \mu(A) = \int_A (1-h) \, \mathrm{d}\sigma.$$

Wegen $\mu(A), \nu(A) \geq 0$ für messbares $A$ folgt $0 \leq h \leq 1$ $\sigma$-fast überall und damit auch $\mu$-fast überall. Für $B = \{h = 0\}$ folgt $\mu(B) = 0$ und deshalb auch $\nu(B) = 0$ wegen $\nu \ll \mu$. Also ist auch $\sigma(B) = 0$ und damit $0 < h \leq 1$ $\sigma$-fast überall. Die Umkehrfunktion $h^{-1}$ existiert also $\sigma$-fast überall.

Mit dem Satz von der monotonen Konvergenz erhält man aus (5) für jede messbare Funktion $x \geq 0$ und jedes $A \in \mathcal{A}$

$$\int_A x \, \mathrm{d}\mu = \int_A xh \, \mathrm{d}\mu.$$

Setzen wir hier $x = h^{-1}$, so ergibt sich

$$\int_A h^{-1} \, \mathrm{d}\mu = \int_A 1 \, \mathrm{d}\sigma = \sigma(A) = \mu(A) + \nu(A)$$

und damit mit $g = h^{-1} - 1$ wie behauptet

$$\nu(A) = \int_A h^{-1} \, \mathrm{d}\mu - \mu(A) = \int_A g \, \mathrm{d}\mu.$$

□

Will man nicht nur positive Dichten im Satz von Radon-Nikodým betrachten, so benötigt man (reellwertige) signierte Maße.

**Definition 2.34.** Eine $\sigma$-additive Mengenfunktion $\mu : \mathcal{A} \to \mathbb{R}$ auf einer $\sigma$-Algebra $\mathcal{A}$ heißt *signiertes Maß*.

Sind $\mu_+$ und $\mu_-$ endliche Maße, so ist $\mu_+ - \mu_-$ ein signiertes Maß. Ein signiertes Maß kann man immer auf diese Weise in einen negativen und einen positiven Anteil zerlegen, die dann selbst Maße sind.

**Satz 2.35** (Hahn-Jordan-Zerlegung). *Sei $\mu$ ein signiertes Maß auf einem messbaren Raum $(\Omega, \mathcal{A})$. Dann existieren Maße $\mu_+$ und $\mu_-$ auf $\mathcal{A}$ und messbare Mengen $P$ und $N$*

*(a)* $\mu = \mu_+ - \mu_-$

*(b)* $P \cup N = \Omega$ *und* $P \cap N = \emptyset$

*(c)* $\mu(A) = \mu_+(A)$ *für messbares* $A \subseteq P$ *und* $\mu(B) = -\mu_-(B)$ *für messbares* $B \subseteq N$.

*Beweis.* in Vorlesung □

Ist $\mu$ ein Maß und $\nu$ ein signiertes Maß (bezüglich derselben $\sigma$-Algebra $\mathcal{A}$), so heißt $\nu$ wieder *absolutstetig* bezüglich des Maßes $\mu$ (und wir schreiben $\nu \ll \mu$), wenn

$$\mu(A) = 0 \Rightarrow \nu(A) = 0 \quad \text{für jedes } A \in \mathcal{A}.$$

gilt. Ist nun $\nu = \nu_+ - \nu_-$ die Hahn-Jordan-Zerlegung mit den zugehörigen messbaren Mengen $P, N$, so folgt aus $\nu \ll \mu$ auch $\nu_+ \ll \mu$ und $\nu_- \ll \mu$. Ist nämlich $\mu(A) = 0$, so ist auch $\mu(A \cup P) = \mu(A \cup N) = 0$ und damit $\nu_+(A) = \nu(A \cup P) = 0$ und $\nu_-(A) = \nu(A \cup N) = 0$.

**Satz 2.36** (Satz von Radon-Nikodým, Version für signierte Maße). *Seien $\mu$ ein $\sigma$-endliches Maße auf einem messbaren Raum $(\Omega, \mathcal{A})$ und sei $\nu$ ein signiertes Maß auf $\mathcal{A}$ mit $\nu \ll \mu$. Dann existiert eine integrierbare Funktion $g : \Omega \to \mathbb{R}$ mit $\nu(A) = \int_A g \, d\mu$ für $A \in \mathcal{A}$.*

*Beweis.* Wir benutzen die Hahn-Jordan-Zerlegung $\nu = \nu_+ - \nu_-$ und die zugehörigen messbaren Mengen $P, N$. Wie wir oben eingesehen haben, sind dann $\nu_+, \nu_-$ absolutstetige Maße bezüglich $\mu$ und der Satz von Radon-Nikodým für Maße liefert Dichten $g_+, g_-$ für $\nu_+, \nu_-$. Dann ist $g = g_+ - g_-$ die gewünschte Funktion. □

## 2.5 Orthonormalsysteme und Orthonormalbasen in Hilberträumen

Wir betrachten in diesem Abschnitt einen Hilbertraum $H$.

**Definition 2.37.** Eine Teilmenge von $S \subset H$ heißt *Orthonormalsystem*, falls gilt:

(a) $\|e\| = 1$ für jedes $e \in S$

(b) $e \perp f$ für alle $e, f \in S$ mit $e \neq f$.

Eine *Orthonormalbasis* oder ein *vollständiges Orthonormalsystem* ist ein Orthonormalsystem $S$ mit $\overline{\operatorname{span} S} = H$.

Ein Hilbertraum $H$ (oder auch ein normierter Raum) heißt *separabel*, falls es eine abzählbare dichte Teilmenge in $H$ gibt. Wir werden in diesem Abschnitt hauptsächlich separable Hilberträume betrachten, dann ist nämlich jedes Orthonormalsystem höchstens abzählbar, da verschiedene Elemente $e, f$ eines Orthonormalsystems nach dem Pythagoras einen Abstand $\|e - f\| = \sqrt{2}$ haben. In einem Hilbertraum mit einem überabzählbaren Orthonormalsystem kann also keine abzählbare dichte Menge existieren.

**Beispiel.** In $\ell_2^n$ bilden die kanonischen Einheitsvektoren $e_k = (0, \dots, 0, 1, 0, \dots, 0)$ mit der 1 an der $k$-ten Stelle eine Orthonormalbasis $S = \{e_1, \dots, e_n\}$.

**Beispiel.** In $\ell_2$ bilden die kanonischen Einheitsvektoren $e_k = (0, \dots, 0, 1, 0, \dots)$ mit der 1 an der $k$-ten Stelle eine (abzählbare) Orthonormalbasis $S = \{e_k : k \in \mathbb{N}\}$.

**Beispiel.** Im Raum $L_2[-\pi, \pi]$ ist das System

$$S = \left\{ \frac{1}{\sqrt{2\pi}} \right\} \cup \left\{ \frac{1}{\sqrt{\pi}} \cos kt \,:\, k \in \mathbb{N} \right\} \cup \left\{ \frac{1}{\sqrt{\pi}} \sin kt \,:\, k \in \mathbb{N} \right\}$$

eine Orthonormalbasis, die sogenannte *trigonometrische Orthonormalbasis*. Überlegen Sie sich, dass dies tatsächlich ein Orthonormalsystem ist. Die Basiseigenschaft ist nicht so leicht zu zeigen.

Orthonormalsysteme sind immer linear unabhängige Mengen. Ist nämlich $\{e_1, \dots, e_n\}$ ein Orthonormalsystem und sind $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ mit $\sum_{k=1}^n \lambda_k e_k = 0$, so folgt durch skalare Multiplikation mit einem $e_\ell$ aus der Orthonormalität sofort

$$0 = \left\langle \sum_{k=1}^n \lambda_k e_k, e_\ell \right\rangle = \sum_{k=1}^n \lambda_k \langle e_k, e_\ell \rangle = \lambda_\ell$$

für alle $\ell = 1, \dots, n$.

Mit dem im Beweis des folgenden Satzes verwendeten Verfahrens kann man eine linear unabhängige Menge schrittweise *orthonormalisieren.*

**Satz 2.38** (Gram-Schmidt-Verfahren)**.** *Ist $\{x_n : n \in \mathbb{N}\}$ eine linear unabhängige Menge in $H$, dann existiert ein Orthonormalsystem $\{e_n : n \in \mathbb{N}\}$ mit* $\operatorname{span}\{x_1, \dots, x_n\} = \operatorname{span}\{e_1, \dots, e_n\}$ *für alle $n \in \mathbb{N}$.*

*Beweis.* Wir setzen $e_1 = \frac{x_1}{\|x_1\|}$. Sind $e_1, \dots, e_{n-1}$ bereits konstruiert, so setzen wir $f_n = x_n - \sum_{k=1}^{n-1} \langle x_n, e_k \rangle e_k \neq 0$ und $e_n = \frac{f_n}{\|f_n\|}$. □

In jedem separablen Hilbertraum $H$ gibt es eine abzählbare oder endliche linear unabhängige Menge, deren abgeschlossene lineare Hülle $H$ ist. Mit dem Gram-Schmidt-Verfahren erhalten wir dann eine abzählbares oder endliches Orthonormalsystem, dessen abgeschlossene lineare Hülle $H$ ist, also eine Orthonormalbasis. In jedem separablen Hilbertraum gibt es also eine Orthonormalbasis.

Der folgende Satz gibt ein Kriterium für die Konvergenz von Reihen der Form $\sum_k \alpha e_k$ für ein Orthonormalsystem $(e_n)$.

**Satz 2.39.** *Sei $(e_n)_{n \in \mathbb{N}}$ ein Orthonormalsystem in $H$ und sei $a = (\alpha_n)_{n \in \mathbb{N}}$ eine Folge in $\mathbb{K}$. Dann konvergiert die Reihe $\sum_{k=1}^\infty \alpha_k e_k$ genau dann, wenn $\sum_{k=1}^\infty |\alpha_k|^2 < \infty$, also $a \in \ell_2$ ist.*

*Beweis.* Ist $s_n = \sum_{k=1}^n \alpha_k e_k$, dann ist

$$\|s_n - s_m\|^2 = \sum_{k,\ell=n+1}^m \alpha_k \overline{\alpha_\ell} \langle e_k, e_\ell \rangle = \sum_{k=n+1}^m |\alpha_k|^2.$$

Also ist $(s_n)$ eine Cauchyfolge (und damit konvergent) genau dann, wenn $a \in \ell_2$ ist. □

In diesem Fall konvergiert die Reihe $\sum_{k=1}^{\infty} \alpha_k e_k$ auch unbedingt, das heißt $\sum_{k=1}^{\infty} \alpha_{\pi(k)} e_{\pi(k)}$ konvergiert für jede Umordnung $\pi$ von $\mathbb{N}$ gegen den gleichen Grenzwert. Die Konvergenz folgt dabei aus der Bedingung $a \in \ell_2$, die natürlich umordnungsinvariant ist. Ist nun $x = \sum_{k=1}^{\infty} \alpha_k e_k$ und $y = \sum_{k=1}^{\infty} \alpha_{\pi(k)} e_{\pi(k)}$, so folgt $\langle x, e_k \rangle = \langle y, e_k \rangle = \alpha_k$ für alle $k \in \mathbb{N}$ und damit $\langle x, z \rangle = \langle y, z \rangle$ für alle $z \in U = \overline{\operatorname{span}(e_k)}$. Offenbar gilt diese Gleichung ebenfalls für $z \in U^\perp$ und damit für alle $z \in H$. Es folgt $x = y$.

Wir sammeln jetzt einige Eigenschaften von Orthogonalreihen der Form $\sum_{k=1}^{\infty} \alpha_k e_k$ insbesondere mit $\alpha_k = \langle x, e_k \rangle$ für ein $x \in H$.

**Satz 2.40.** *Sei $\{e_1, e_2, \dots\}$ ein Orthonormalsystem in $H$ und seien $x, y \in H$. Dann gilt:*

*(a)* Bessel-Ungleichung*:* $\sum_{k=1}^{\infty} \left|\langle x, e_k \rangle\right|^2 \leq \|x\|^2$

*(b)* $\sum_{k=1}^{\infty} \left|\langle x, e_k \rangle \langle e_k, y \rangle\right| < \infty$

*(c)* $\sum_{k=1}^{\infty} \langle x, e_k \rangle e_k$ *konvergiert unbedingt.*

*(d) Die Abbildung* $P : x \mapsto \sum_{k=1}^{\infty} \langle x, e_k \rangle e_k$ *ist die Orthogonalprojektion auf* $\overline{\operatorname{span}(e_k)}$.

*Beweis.* (a) Für jedes $n \in \mathbb{N}$ ist $x_n = x - \sum_{k=1}^{n} \langle x, e_k \rangle e_k$ orthogonal zu $e_1, \dots, e_n$. Aus dem Satz von Pythagoras folgt

$$\|x\|^2 = \|x_n\|^2 + \left\| \sum_{k=1}^{n} \langle x, e_k \rangle e_k \right\|^2 = \|x_n\|^2 + \sum_{k=1}^{n} |\langle x, e_k \rangle|^2 \geq \sum_{k=1}^{n} |\langle x, e_k \rangle|^2.$$

Die Bessel-Ungleichung ergibt sich für $n \to \infty$.

(b) Dies folgt aus der Cauchy-Schwarz-Ungleichung in $\ell_2$ und der Bessel-Ungleichung.

(c) Dies folgt aus Satz 2.39 und der Bessel-Ungleichung.

(d) Wegen

$$\left\langle x - \sum_{k=1}^{\infty} \langle x, e_k \rangle e_k, e_n \right\rangle = 0$$

ist $x - Px \perp S =\perp \overline{\operatorname{span}(e_k)}$ ist $x - Px \in \overline{\operatorname{span}(e_k)}^\perp$.

□

Für Orthonormalbasen gelten noch stärkere Aussagen, die wir in folgendem Satz zusammenfassen. Die entsprechenden Aussagen für einen endlichdimensionalen Hilbertraum kennen Sie schon aus der linearen Algebra.

**Satz 2.41.** *Für ein Orthonormalsystem $S = \{e_1, e_2, \dots\}$ in einem separablen unendlichdimensionalen Hilbertraum sind die folgenden Aussagen äquivalent:*

*(a) $S$ ist eine Orthonormalbasis.*

*(b) Ist $x \in H$ mit $x \perp S$, dann ist $x = 0$.*

*(c)* $x = \sum_{k=1}^{\infty} \langle x, e_k \rangle e_k$ *für alle $x \in H$.*

*(d)* $\langle x, y \rangle = \sum_{k=1}^{\infty} \langle x, e_k \rangle \langle e_k, y \rangle$ *für alle $x, y \in H$.*

*(e)* Parseval-Gleichung*:* $\|x\|^2 = \sum_{k=1}^{\infty} \left|\langle x, e_k \rangle\right|^2$ *für alle $x \in H$.*

*Beweis.* $(b) \Rightarrow (a)$: Aus $(b)$ folgt $(\operatorname{span} S)^\perp = \{0\}$ und damit aus Satz 2.31 die Dichtheit von $\operatorname{span} S$ in $H$. Also ist $S$ eine Orthonormalbasis.

$(a) \Rightarrow (c)$ folgt aus dem vorhergehenden Satz.

$(c) \Rightarrow (d)$ folgt durch skalare Multiplikation von $(c)$ mit $y$.

$(d) \Rightarrow (e)$ folgt durch Setzen von $y = x$.

$(e) \Rightarrow (b)$: Wenn $(b)$ nicht gilt, dann gibt es ein $x \in H$ mit $\|x\| = 1$ und $x \perp S$. Also ist $S \cup \{x\}$ und es folgt $\sum_{k=1}^\infty \left|\langle x, e_k\rangle\right|^2 = 0$, d.h $(e)$ gilt dann für $x$ nicht. □

Aus dieser Charakterisierung folgt, dass alle separablen unendlichdimensionalen Hilberträume als normierte Räume nicht zu unterscheiden sind. Dazu sagen wir, dass zwei normierte Räume $X$ und $Y$ *isometrisch isomorph* sind, wenn es einen bijektiven linearen Operator $T : X \to Y$ gibt mit $\|Tx\| = \|x\|$. Wir schreiben dann $X \cong Y$.Dann nennt man $T$ einen *isometrischen Isomorphismus*. Ein solcher ist offenbar stetig und hat eine stetige Inverse mit $\|T\| = \|T^{-1}\| = 1$.

**Satz 2.42** (Satz von Fischer-Riesz). *Ist $H$ ein separabler Hilbertraum, dann ist $H$ isometrisch isomorph zu $\ell_2$, also $H \cong \ell_2$. Insbesondere ist $L_2[0,1] \cong \ell_2$.*

*Beweis.* Wählen wir eine Orthonormalbasis $\{e_1, e_2, \dots\}$ in $H$, so definieren wir $T : H \to \ell_2$ durch $Tx = \big(\langle x, e_k\rangle\big)$. Die Parseval-Gleichung zeigt, dass $T$ wohldefiniert ist und $\|Tx\| = \|x\|$ für alle $x \in H$ gilt. Umgekehrt ist für $a = (\alpha_k) \in \ell_2$ die Reihe $x = \sum_{k=1}^\infty \alpha_k e_k$ in $H$ absolut konvergent und wir erhalten $Tx = a$. Also ist $T$ surjektiv. Aus $Tx = 0$ folgt wegen $0 = \|Tx\| = \|x\|$ auch $x = 0$, also ist $T$ auch injektiv und damit bijektiv. Insgesamt ist $T$ ein isometrischer Isomorphismus. □

Der konstruierte isometrische Isomorphismus $T : H \to \ell_2$ erfüllt sogar $\langle Tx, Ty\rangle = \langle x, y\rangle$ für alle $x, y \in H$, erhält also Skalarprodukte, ist „winkeltreu".

**Beispiel.** Wir betrachten noch einmal das trigonometrische System in $L_2[-\pi, \pi]$. Dazu bezeichnen wir die Funktionen des trigonometrischen Systems mit $c_0(t) = \frac{1}{\sqrt{2\pi}}$ sowie $c_k(t) = \frac{1}{\sqrt{\pi}} \cos kt$ und $s_k(t) = \frac{1}{\sqrt{\pi}} \sin kt$. Dann erhalten wir für jedes $x \in L_2[-\pi, \pi]$ die Darstellung

$$x = \langle x, c_0\rangle c_0 + \sum_{k=1}^\infty \langle x, c_k\rangle c_k + \langle x, s_k\rangle s_k$$

die *Fourierreihe* von $x$ mit den *Fourierkoeffizienten*

$$\langle x, c_0\rangle = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} x(t)\,\mathrm{d}t, \quad \langle x, c_k\rangle = \frac{1}{\sqrt{\pi}} \int_{-\pi}^{\pi} x(t) \cos kt\,\mathrm{d}t, \quad \langle x, s_k\rangle = \frac{1}{\sqrt{\pi}} \int_{-\pi}^{\pi} x(t) \sin kt\,\mathrm{d}t.$$

Diese Fourierreihe ist also gegen $x$ konvergent in $L_2$, insbesondere also fast überall. Unter welchen Voraussetzungen sie z.B. für eine stetige Funktion überall gegen $x$ konvergiert ist eine delikate Frage, die in der Fourieranalysis oder harmonischen Analysis eingehend untersucht wird. Die Parsevalsche Gleichung liefert

$$\int_{-\pi}^{\pi} |x(t)|^2\,\mathrm{d}t = \left|\langle x, c_0\rangle\right|^2 + \sum_{k=1}^\infty \left(\left|\langle x, c_k\rangle\right|^2 + \left|\langle x, s_k\rangle\right|^2\right).$$

## 2.6 Der Satz von Hahn-Banach

In diesem Abschnitt wollen wir uns näher mit dem Dualraum $X'$ der stetigen linearen Funktionale auf einem Banachraum $X$ beschäftigen. Der Darstellungssatz von Frechét-Riesz liefert eine sehr befriedigende Antwort für Hilberträume $H$, der Dualraum $H'$ kann isometrisch isomorph mit $H$ identifiziert werden. In diesem Abschnitt wollen wir auf ähnliche Weise die Dualräume von $L_p$-Räumen charakterisieren. Außerdem wollen wir den Satz von Hahn-Banach beweisen, der zeigt, dass der Dualraum eines beliebigen normierten Raumes sehr reichhaltig ist.

Wir beginnen mit den Folgenräumen $\ell_p$.

**Satz 2.43.** *Ist $1 \leq p < \infty$ und $\frac{1}{p} + \frac{1}{q} = 1$, dann ist $\ell_p' \cong \ell_q$. Die Abbildung $T : \ell_q \to \ell_p'$ gegeben durch $(Tx)(y) = \sum_{k=1}^{\infty} \xi_k \eta_k$ für $x = (\xi_k) \in \ell_q$ und $y = (\eta_k) \in \ell_p$ ist ein isometrischer Isomorphismus. Dieselbe Abbildung ist ein isometrischer Isomorphismus zwischen $\ell_1$ und $c_0'$. Wir haben also auch $c_0' \cong \ell_1$.*

*Beweis.* Wir betrachten nur den Fall $1 < p < \infty$. Die anderen Behauptungen ergeben sich ähnlich.

Die Hölder-Ungleichung zeigt, dass für $x = (\xi_k) \in \ell_q$ und $y = (\eta_k) \in \ell_p$ die Reihe $\sum_{k=1}^{\infty} \xi_k \eta_k$ absolut konvergiert und $\|(Tx)(y)\| \leq \|x\|_q \|y\|_p$ gilt. Die Linearität von $Tx$ und $T$ sowie die Ungleichung $\|Tx\| \leq \|x\|_q$ sind dann klar. Ist $Tx = 0$, so ist $\xi_k = (Tx)(e_k) = 0$ für alle $k \in \mathbb{N}$ und damit $x = 0$. Also ist $T$ injektiv.

Wir zeigen noch zusammen, dass $T$ surjektiv ist und auch $\|Tx\| \geq \|x\|_q$ ist. Sei also $y' \in \ell_p'$. Dann definieren wir $x = (\xi_k) = (y'(e_k))$. Dazu setzen wir noch $\eta_k = 0$ für $\xi_k = 0$ und $\eta_k = |\xi_k|^q / \xi_k$ sonst. Dann ist $|\eta_k|^p = |\xi_k|^q$ und wir erhalten für jedes $n \in \mathbb{N}$

$$\sum_{k=1}^{n} |\xi_k|^q = \sum_{k=1}^{n} \xi_k \eta_k = y' \left( \sum_{k=1}^{n} \eta_k e_k \right) \leq \|y'\| \left( \sum_{k=1}^{n} |\eta_k|^p \right)^{1/p} = \|y'\| \left( \sum_{k=1}^{n} |\xi_k|^q \right)^{1/p} .$$

Es folgt

$$\left( \sum_{k=1}^{n} |\xi_k|^q \right)^{1/q} \leq \|y'\|$$

für jedes $n \in \mathbb{N}$ und damit $x \in \ell_q$ und $\|x\|_q \leq \|y'\|$. Weiter ist nun $y'(e_k) = (Tx)(e_k)$ für alle $k \in \mathbb{N}$. Somit stimmen $y'$ und $Tx$ auf dem dichten Teilraum der abbrechenden Folgen überein und es folgt aus ihrer Stetigkeit schließlich $y' = Tx$. □

Der Dualraum von $\ell_\infty$ ist nicht $\ell_1$. Man kann zeigen, dass $\ell_\infty'$ nicht separabel ist. In diesem Fall ist die Abbildung $T$ im Satz zwar isometrisch und damit injektiv, aber nicht surjektiv. Für die endlichdimensionalen Räume zeigt der Beweis ohne Einschränkung an $p$ die Dualität $\left(\ell_p^n\right)' = \ell_q^n$.

Im Fall allgemeiner $L_p$-Räume kennzeichnet man die Dualräume ganz analog.

**Satz 2.44.** *Sei $(\Omega, \mathcal{A}, \mu)$ ein $\sigma$-endlicher Maßraum. Ist $1 \leq p < \infty$ und $\frac{1}{p} + \frac{1}{q} = 1$, dann ist $L_p(\mu)' \cong L_q(\mu)$. Die Abbildung $T : L_q(\mu) \to L_p(\mu)'$ gegeben durch $(Tx)(y) = \int xy \, \mathrm{d}\mu$ für $x \in L_q(\mu)$ und $y \in L_p(\mu)$ ist ein isometrischer Isomorphismus.*

*Beweis.* in Vorlesung □

Wieder ist im allgemeinen $L_\infty' \neq L_1$. Wichtig ist auch die Charakterisierung des Dualraums des Raums $C(K)$ der stetigen Funktionen auf einem kompakten metrischen (oder topologischen) Raum $K$. Der sogenannte Darstellungssatz von Riesz (nicht zu verwechseln mit dem Darstellungssatz von Frechét-Riesz) beschreibt die linearen Funktionale auf $C(K)$ gerade als Integrale bezüglich signierter Maße: zu jedem $a \in C(K)'$ existiert ein signiertes Maß $\mu$ mit $a(x) = \int_K x \,\mathrm{d}\mu$.

Wir kommen schließlich zum Satz von Hahn-Banach, der insbesondere zeigt, dass der Dualraum eines Banachraums sehr reichhaltig ist. Dies benötigt man, um die Struktur eines Banachraums $X$ mittels der stetigen Funtkionale auf $X$ studieren zu können.

**Satz 2.45** (Satz von Hahn-Banach). *Sei $X$ ein normierter Raum und $U$ ein linearer Teilraum von $X$. Dann existiert zu jedem linearen stetigen Funktional $u' \in U'$ eine lineare stetige und normerhaltende Fortsetzung $x' \in X'$, d.h. $x'(u) = u'(u)$ für $u \in U$ und $\|x'\| = \|u'\|$.*

*Beweis.* in Vorlesung □

**Korollar 2.46.** *Ist $X$ ein normierter Raum und $x \in X$, dann existiert ein $x' \in X'$ mit $\|x'\| = 1$ und $x'(x) = \|x\|$. Insbesondere gilt für jedes $x \in X$*

$$\|x\| = \sup_{x' \in B_{X'}} |x'(x)| = \max_{x' \in B_{X'}} |x'(x)|.$$

*Beweis.* Für $x = 0$ ist nichts zu zeigen, also nehmen wir o.B.d.A. $\|x\| = 1$ an. Auf $U = \operatorname{span}\{x\}$ ist durch $u' : \lambda x \mapsto \lambda$ ein lineares stetiges Funktional mit $\|u'\| = 1$ und $u'(x) = 1$ definiert, welches nach dem Satz von Hahn-Banach eine Fortsetzung $x' \in X'$ mit $\|x'\| = 1$ und $x'(x) = u'(x) = 1 = \|x\|$ hat. □

Vergleichen Sie die mit der Normdefinition von $x' \in X'$

$$\|x'\| = \sup_{x \in B_X} |x'(x)|,$$

in der das Supremum im Allgemeinen nicht angenommen wird.

**Korollar 2.47.** *Sei $X$ ein normierter Raum, $U$ ein abgeschlossener linearer Teilraum und $x \in X \setminus U$. Dann existiert ein $x' \in X'$ mit $x'|_U = 0$ und $x'(x) \neq 0$.*

*Beweis.* Wir definieren auf $\operatorname{span}\big(U \cup \{x\}\big)$ das lineare Funktional $u' : u + \lambda x \mapsto \lambda$. Der Kern von $u'$ ist der abgeschlossene Teilraum $U$. Das folgende Lemma zeigt, dass $u'$ stetig ist. Jede Hahn-Banach-Fortsetzung $x'$ von $u'$ zeigt die Behauptung. □

**Lemma 2.48.** *Ein lineares Funktional auf einem normierten Raum ist stetig genau dann, wenn sein Kern abgeschlossen ist.*

*Beweis.* Ist das lineare Funktional $a : X \to \mathbb{K}$ stetig, so ist $U = \ker(a)$ als Urbild der abgeschlossenen Menge $\{0\}$ abgeschlossen. Ist umgekehrt $U$ abgeschlossen und $a \neq 0$, so wählen wir ein $x \in X$ mit $a(x) = 1$. Da $x \notin U$ ist, gibt es ein $\varepsilon > 0$ mit $B_\varepsilon(x) \cap U = \emptyset$. Sei nun $(x_i)$ eine Nullfolge in $X$. Wir haben zu zeigen, dass $a(x_i) \to 0$ gilt. Ist $a(x_i) \neq 0$, so folgt aus $a(x - x_i/a(x_i)) = 0$ sofort $\|\frac{x_i}{a(x_i)}\| > \varepsilon$ und damit $|a(x_i)| < \frac{\|x_i\|}{\varepsilon} \to 0$. □

**Korollar 2.49.** *Sei $X$ ein normierter Raum und $U$ ein linearer Teilraum. Dann ist $U$ dicht in $X$ genau dann, wenn für jedes $x' \in X'$ mit $x'|_U = 0$ sogar $x' = 0$ gilt.*

*Beweis.* Dies folgt sofort aus dem vorhergehenden Korollar. □

**Korollar 2.50.** *Die Abbildung $T : \ell_1 \to \ell_\infty'$ gegeben durch $(Tx)(y) = \sum_{k=1}^\infty \xi_k \eta_k$ für $x = (\xi_k) \in \ell_1$ und $y = (\eta_k) \in \ell_\infty$ ist isometrisch, aber nicht surjektiv.*

*Beweis.* Der Beweis von $\|Tx\| = \|x\|$ ist einfach. Wir betrachten nun das stetige Grenzwertfunktional $\lim : c \to \mathbb{K}$ auf dem abgeschlossenen Teilraum $c \subset \ell_\infty$ der konvergenten Folgen. Dieses Funktional hat Norm 1. Jede Hahn-Banach-Fortsetzung $x' \in \ell_\infty'$ von lim heißt Banach-Limes und ist nicht im Bild von $T$. Wäre nämlich

$$x'(y) = \sum_{k=1}^\infty \xi_k \eta_k$$

für ein $x \in \ell_1$ und alle $y \in \ell_\infty$, so würde für $y = e_k \in c$ folgen, dass

$$\xi_k = x'(e_k) = \lim(e_k) = 0$$

für alle $k \in \mathbb{N}$ und damit $x' = x = 0$ folgen, ein Widerspruch. □

# Index