CARL VON OSSIETZKY
universität OLDENBURG
FAKULTÄT V
MATHEMATIK- UND NATURWISSENSCHAFTEN

Prof. Dr. Dietmar Pfeifer
Institut für Mathematik

Analytische Prinzipien der Stochastik

Stand: April 2016

Inhalt

Vorbemerkung

Moderne stochastische Methoden und Theorien kommen ohne Hilfsmittel aus der Analysis, speziell der Funktionalanalysis, nicht aus. Hierzu zählt insbesondere der Bereich der *Stochastischen Prozesse*, die u.a. für die Modellierung von Finanzmärkten (z.B. der geometrischen Brown'schen Bewegung im Black-Scholes-Modell) in stetiger Zeit das wesentliche Hilfsmittel bilden (vgl. etwa [11], [14], [17], [19], [20] und [28]). Aber auch in der Versicherungsmathematik, insbesondere in der Risiko- und Ruintheorie, spielen Stochastische Prozesse eine wichtige Rolle, so z.B. beim klassischen Risikoprozess, der auf dem Poisson-Prozess beruht, oder bei den Phasentyp-Verteilungen, die sich aus zeitstetigen Markoff-Prozessen mit diskretem Zustandraum ableiten. Der vorliegende Text soll deshalb helfen, das Verständnis Stochastischer Prozesse und der *Stochastischen Analysis* durch eine systematische Behandlung ihrer funktionalanalytischen Grundlagen und Prinzipien zu erleichtern.

Der Text orientiert sich dabei stark an einer Sichtweise auf die Funktionalanalysis, die ich während meiner 16-jährigen Zeit an der RWTH Aachen bei meinem akademischen Lehrer Paul-Leo Butzer kennen und schätzen gelernt habe. Gut verständliche Monographien zur Vertiefung dieses Stoffs sind die Bücher von BACHMAN / NARICI (2000) [1], ROYDEN (1988) [25], TAYLOR (1963) [29] und WERNER [30]; vgl. aber auch die Monographien [5], [9], [18], [26], [27] und [31].

I Grundlagen der Funktionalanalysis

In der Funktionalanalysis vereinen sich in sehr fruchtbarer Weise Aspekte der „klassischen" Analysis für Funktionen reeller Veränderlicher mit solchen der Linearen Algebra. Insbesondere sind häufig – in der Regel unendlich-dimensionale – Vektorräume geeigneter Funktionen (z.B. stetiger oder differenzierbarer Funktionen) Gegenstand des Interesses. Durch den Begriff der *Metrik* oder *Norm* lassen sich in solchen algebraischen Strukturen topologische Eigenschaften und damit auch interessante Konvergenzbegriffe einführen. Solche Konvergenzen kommen z.B. in der Stochastischen Finanzmathematik schon beim Übergang vom zeitdiskreten Cox-Ross-Rubinstein- zum Black-Scholes-Modell vor (vgl. das Skript DISKRETE STOCHASTISCHE FINANZMATHEMATIK, Kapitel 8). Weil Funktionen in der Funktionalanalysis eine so bedeutsame Rolle spielen, bezeichnen wir im Folgenden die zu Grunde liegenden Mengen mit dem Buchstaben $\mathcal{X}$ und ihre Elemente mit den Buchstaben $f, g, h, \ldots$

I. 1. Metrische Räume

Definition 1 (metrischer Raum)**.** Für eine nicht-leere Menge $\mathcal{X}$ sei eine Abbildung $\rho:\ \mathcal{X}\times\mathcal{X}\to\mathbb{R}$ gegeben mit den Eigenschaften

$$\rho(f,g)\geq 0 \text{ für alle } f,g\in\mathcal{X}$$

$$\rho(f,g)=0 \Leftrightarrow f=g \text{ für alle } f,g\in\mathcal{X}$$

$$\rho(f,g)=\rho(g,f) \text{ für alle } f,g\in\mathcal{X} \quad \text{[Symmetrie]}$$

$$\rho(f,h)\leq\rho(f,g)+\rho(g,h) \text{ für alle } f,g,h\in\mathcal{X} \quad \text{[Dreiecksungleichung]}$$

Dann heißt ρ eine *Metrik* oder *Abstandsfunktion* auf $\mathcal{X}$, $\rho(f,g)$ heißt Abstand zwischen f und $g \in \mathcal{X}$ und das Paar $(\mathcal{X},\rho)$ heißt *metrischer Raum.*

Bemerkung: Aus der Dreiecksungleichung ergibt sich sofort auch noch die folgende äquivalente Ungleichung:

$$|\rho(f,h)-\rho(g,h)| \le \rho(f,g) \text{ für alle } f,g,h \in \mathcal{X}$$

sowie die sog. *Vierecksungleichung*

$$\left|\rho(f_1,g_1)-\rho(f_2,g_2)\right| \le \rho(f_1,f_2)+\rho(g_1,g_2) \text{ für alle } f_1,f_2,g_1,g_2 \in \mathcal{X}.$$

Beispiel 1. Die folgenden Strukturen bilden metrische Räume:

a) $\mathcal{X} = \mathbb{R}^1$ mit der Metrik ρ, gegeben durch

$$\rho(f,g) := |f-g|, \quad f,g \in \mathcal{X}$$

(euklidischer Abstand in $\mathbb{R}^1$)

b) $\mathcal{X} = \mathbb{R}^d$ mit der Metrik ρ, gegeben durch

$$\rho(f,g) := \sqrt{\sum_{i=1}^{d}(f_i-g_i)^2}, \quad f=(f_1,\cdots,f_d),\ g=(g_1,\cdots,g_d)\in\mathcal{X}$$

(euklidischer Abstand in $\mathbb{R}^d$ mit $d \in \mathbb{N}$)

c) $\mathcal{X} = \mathbb{R}^d$ mit der Metrik ρ, gegeben durch

$$\rho(f,g) := \max_{1\le i\le d}\{|f_i-g_i|\}, \quad f=(f_1,\cdots,f_d),\ g=(g_1,\cdots,g_d)\in\mathcal{X}$$

d) $\mathcal{X} = \ell^1 = \left\{ f \in \mathbb{R}^{\mathbb{N}} \,\middle|\, \sum_{i=1}^{\infty} |f_i| < \infty \right\}$ mit der Metrik ρ, gegeben durch

$$\rho(f,g) := \sum_{i=1}^{\infty}|f_i-g_i|, \quad f=\{f_i\}_{i\in\mathbb{N}},\ g=\{g_i\}_{i\in\mathbb{N}}\in\mathcal{X}$$

[dieser metrische Raum wird z.B. bei der *diskreten Fourier-Transformation* verwendet, vgl. das Skript RISIKOTHEORIE, Kapitel II.3]

e) $\mathcal{X} = \ell^2 = \left\{ f \in \mathbb{R}^{\mathbb{N}} \,\middle|\, \sum_{i=1}^{\infty} f_i^2 < \infty \right\}$ mit der Metrik ρ, gegeben durch

$$\rho(f,g) := \sqrt{\sum_{i=1}^{\infty}\left(f_i - g_i\right)^2}, \quad f = \{f_i\}_{i\in\mathbb{N}},\ g = \{g_i\}_{i\in\mathbb{N}} \in \mathcal{X}$$

f) $\mathcal{X} = C[a,b] = \left\{ f : [a,b] \to \mathbb{R} \mid f \text{ stetig} \right\}$ mit $a < b,\ a,b \in \mathbb{R}$ mit der Metrik $\rho,$ gegeben durch

$$\rho(f,g) := \max_{a\le x\le b}\left\{| f(x) - g(x)|\right\}, \quad f,g \in \mathcal{X}$$

g) $\mathcal{X} = C[a,b] = \left\{ f : [a,b] \to \mathbb{R} \mid f \text{ stetig} \right\}$ mit der Metrik $\rho,$ gegeben durch

$$\rho(f,g) := \int_a^b | f(x) - g(x)|\,dx, \quad f,g \in \mathcal{X}$$

[im Riemann'schen oder Lebesgue'schen Sinn, vgl. das Skript STOCHASTIK, Kapitel I.7]

Beispiel 2. Es sei $(\Omega,\mathcal{A})$ ein Messraum. Dann lassen sich auf verschiedene Weise Metriken auf dem Raum $\mathcal{X}$ der Wahrscheinlichkeitsmaße über $\mathcal{A}$ definieren[1]:

a) $\rho(P,Q) := \sup_{A\in\mathcal{A}}\left|P(A) - Q(A)\right|, \quad P,Q \in \mathcal{X}$ [Totalvariation]

b) $\rho(P,Q) := \sqrt{\int\left(\sqrt{f_P} - \sqrt{f_Q}\right)^2 d\mu}, \quad P,Q \in \mathcal{X}$ [Hellinger-Distanz],

wobei μ ein beliebiges σ-endliches dominierendes Maß für P und Q und f_P bzw. f_Q die zugehörigen Dichten bezüglich μ seien.[2] $\rho(P,Q)$ ist dabei unabhängig von der speziellen Wahl sowohl des dominierenden Maßes als auch der zugehörigen μ-Dichten.

c) $\rho(P,Q) := \inf_{X,Y\in\mathcal{Z}}\left\{E\left(|X-Y|\right) \,\middle|\, \mathcal{L}(X) = P, \mathcal{L}(Y) = Q\right\}, \quad P,Q \in \mathcal{X}$ [Wasserstein-Distanz],

wobei $\mathcal{Z}$ die Menge aller Zufallsvariablen auf einem (prinzipiell beliebigen, aber genügend "großen", festen) Messraum mit Werten in Ω und $\mathcal{L}(X), \mathcal{L}(Y)$ die Verteilungen von X und Y bezeichnen.[3]

---

[1] Aus Gründen der Konvention schreiben wir hier für die Elemente von $\mathcal{X}$ nicht *f,g*, … , sondern wie in der Stochastik üblich *P,Q*, …

[2] Ein Maß μ heißt ein Maß ν dominierend, wenn jede μ-Nullmenge auch eine ν-Nullmenge ist. Nach dem Satz von Radon und Nikodym (vgl. das Skript STOCHASTIK, Satz 21) existiert dann eine Dichte *f* von ν bezüglich μ. Beispielsweise kann immer $\mu = P + Q$ als gemeinsam dominierendes Maß für *P* und *Q* gewählt werden.

[3] Das Symbol $\mathcal{L}(X)$ für die Verteilung der Zufallsvariablen *X* ist vor allem in der englischsprachigen Literatur gebräuchlich ("Law of *X*"). Wir benutzen es hier, weil das Symbol *P* bereits für die Bezeichnung eines Elements von $\mathcal{X}$ vergeben ist.

Definition 2 (topologische Eigenschaften)**.** In einem metrischen Raum $(\mathcal{X}, \rho)$ heißt

$$K_r(f) := \{g \in \mathcal{X} \mid \rho(f,g) < r\}, \quad f \in \mathcal{X},\ r > 0$$

offene Kugel mit Mittelpunkt *f* und Radius *r*. Die Menge

$$\bar{K}_r(f) := \{g \in \mathcal{X} \mid \rho(f,g) \le r\}, \quad f \in \mathcal{X},\ r > 0$$

heißt *abgeschlossene Kugel*[4] mit Mittelpunkt *f* und Radius *r*.

Ein Element *f* einer Teilmenge $A \subseteq \mathcal{X}$ heißt *innerer Punkt von A*, wenn eine reelle Zahl $r > 0$ derart existiert, dass $K_r(f) \subseteq A$ gilt. Die Menge aller inneren Punkte von *A* heißt *Inneres von A* und wird mit dem Symbol A^o bezeichnet.

Eine Teilmenge $A \subseteq \mathcal{X}$ heißt *offen*, falls jeder Punkt von *A* ein innerer Punkt von *A* ist.

Eine *offene Umgebung* $U(f)$ eines Punktes $f \in \mathcal{X}$ ist eine offene Menge, die *f* als inneren Punkt enthält.

Ein Element *f* einer Teilmenge $A \subseteq \mathcal{X}$ heißt *Häufungspunkt* von *A*, wenn jede offene Umgebung $U(f)$ mindestens einen Punkt $g \in A$ mit $g \ne f$ enthält. Die Menge der Häufungspunkte von *A* wird mit dem Symbol A^+ bezeichnet.

Eine Teilmenge $A \subseteq \mathcal{X}$ heißt *abgeschlossen*, falls die Menge *A* alle ihre Häufungspunkte enthält, d.h. wenn $A^+ \subseteq A$ gilt. Die Menge $\bar{A} := A \cup A^+$ heißt *abgeschlossene Hülle* von *A*, $\partial A := \bar{A} \setminus A^o$ heißt *Rand* von *A*.

Satz 1 (topologische Eigenschaften)**.** Es sei $(\mathcal{X}, \rho)$ ein metrischer Raum. Dann gilt:

a) Ist $A \subseteq \mathcal{X}$ offen, so ist das Komplement A^c abgeschlossen; ist umgekehrt $A \subseteq \mathcal{X}$ abgeschlossen, so ist das Komplement A^c offen.

b) Ist $\{A_i \mid i \in I\}$ mit einer nicht-leeren Indexmenge *I* eine Familie von Teilmengen von $\mathcal{X}$, so gilt: Sind alle A_i offen, so auch $\bigcup_{i \in I} A_i$; sind alle A_i abgeschlossen, so auch $\bigcap_{i \in I} A_i$. Ist *I* zusätzlich *endlich*, so gilt auch: Sind alle A_i offen, so auch $\bigcap_{i \in I} A_i$; sind alle A_i abgeschlossen, so auch $\bigcup_{i \in I} A_i$.

c) Jede abgeschlossene Kugel $\bar{K}_r(f)$ mit $f \in \mathcal{X},\ r > 0$ ist eine abgeschlossene Menge, ebenso jede abgeschlossene Hülle $\bar{A}$ für jede Teilmenge $A \subseteq \mathcal{X}$. Das Innere A^o jeder Teilmenge $A \subseteq \mathcal{X}$ ist eine offene Menge.

---

[4] Dies ist zu unterscheiden von der abgeschlossenen Hülle $\overline{K_r(f)}$; i. Allg. gilt nur $\bar{K}_r(f) \supseteq \overline{K_r(f)}$, obwohl beide Mengen abgeschlossen sind. Die nachfolgende Bemerkung zeigt, dass diese Mengen tatsächlich unterschiedlich sein können.

Beweis:

a) Sei $A \subseteq \mathcal{X}$ offen und *f* ein Häufungspunkt von A^c. Wir nehmen zunächst an, dass $f \in A$ gilt. Dann ist f innerer Punkt von *A*, also existiert ein $r > 0$ mit $U(f) := K_r(f) \subseteq A$: Widerspruch zur Voraussetzung, dass *f* ein Häufungspunkt von A^c ist, denn $U(f)$ enthält *keinen* Punkt $g \in A^c$! Damit ist notwendig $f \notin A$, also $f \in A^c$. Also ist A^c abgeschlossen.
Sei nun $A \subseteq \mathcal{X}$ abgeschlossen und $f \in A^c$. Wir zeigen, dass f ein innerer Punkt von A^c ist. Wegen $f \in A^c$ ist jedenfalls f kein Häufungspunkt von *A*, da *A* abgeschlossen ist. Also gibt es eine offene Umgebung $U(f)$, die *keinen* Punkt $g \in A$ mit $g \neq f$ enthält. Wegen $f \in A^c$ ist damit $U(f) \cap A = \varnothing$ und somit $U(f) \subseteq A^c$. Nach Definition der Umgebung ist also *f* ein innerer Punkt von A^c.

b) Seien zunächst alle A_i offen und $f \in \bigcup_{i \in I} A_i$; wir zeigen, dass f ein innerer Punkt von $\bigcup_{i \in I} A_i$ ist. Nach Definition der Vereinigung existiert ein $j \in I$ mit $f \in A_j \subseteq \bigcup_{i \in I} A_i$; also ist *f* ein innerer Punkt von A_j, so dass ein $r > 0$ existiert mit $K_r(f) \subseteq A_j \subseteq \bigcup_{i \in I} A_i$. Also ist f auch ein innerer Punkt von $\bigcup_{i \in I} A_i$.
Seien nun alle A_i abgeschlossen. Dann sind die A_i^c sämtlich offen, also auch $\bigcup_{i \in I} A_i^c$. Folglich ist nach Teil a) des Satzes die Menge $\bigcap_{i \in I} A_i = \left(\bigcup_{i \in I} A_i^c \right)^c$ abgeschlossen.
Im Fall einer *endlichen* Indexmenge *I* argumentiert man für den ersten Teil der Aussage folgendermaßen: ist $f \in \bigcap_{i \in I} A_i$, so existieren reelle Zahlen $r_i > 0$, $i \in I$ mit $K_{r_i}(f) \subseteq A_i$ für alle $i \in I$. Dann ist aber $r := \min_{i \in I} \{r_i\} > 0$ mit $K_r(f) \subseteq K_{r_i}(f) \subseteq A_i$, also auch $K_r(f) \subseteq \bigcap_{i \in I} A_i$, d.h. $\bigcap_{i \in I} A_i$ ist offen. Der zweite Teil ergibt sich analog oben durch Komplementbildung.

c) Es sei g ein Häufungspunkt der abgeschlossenen Kugel $\bar{K}_r(f)$. Dann existiert eine Folge $\{f_n\}_{n \in \mathbb{N}} \subseteq \bar{K}_r(f)$ mit $f_n \in U_n(g) := K_{1/n}(g)$, also $\rho\left(f_n, g\right) < \dfrac{1}{n}$ für alle $n \in \mathbb{N}$. Mit der Dreiecksungleichung folgt nun

$$\rho(f, g) \leq \rho(f_n, g) + \rho(f_n, f) < r + \frac{1}{n} \text{ für alle } n \in \mathbb{N}$$

und somit $\rho(f, g) \leq r$, also $g \in \bar{K}_r(f)$. Also ist $\bar{K}_r(f)$ abgeschlossen.

Sei nun g ein Häufungspunkt der abgeschlossenen Hülle $\bar{A} = A \cup A^+$ einer Teilmenge $A \subseteq \mathcal{X}$. Die Aussage ist trivial, wenn schon *g* Häufungspunkt von *A* oder $g \in A$ ist. Sei also $g \notin A$ und g Häufungspunkt von $A^+ \setminus A$. Dann gibt es ähnlich dem Beweis zu Teil c) eine Folge $\{g_n\}_{n \in \mathbb{N}} \subseteq A^+$ von Häufungspunkten von *A* mit $g_n \neq g$ und $\rho\left(g_n, g\right) < \dfrac{1}{n}$ für alle $n \in \mathbb{N}$. Jedes

g_n ist Häufungspunkt von *A*, also gibt es ein $f_n \in A,\ f_n \neq g$ mit $f_n \in U_n(g_n) := K_{1/n}(g_n)$, d.h. $\rho(f_n, g_n) < \frac{1}{n}$ für alle $n \in \mathbb{N}$. Mit der Dreiecksungleichung folgt nun

$$\rho(f_n, g) \leq \rho(f_n, g_n) + \rho(g_n, g) < \frac{2}{n} \text{ für alle } n \in \mathbb{N}.$$

Sei nun $U(g)$ eine beliebige Umgebung von *g*. Dann enthält $U(g)$ als offene Menge eine geeignete Kugel $K_\varepsilon(g) \subseteq U(g)$ mit $\varepsilon > 0$ und $f_n \in K_\varepsilon(g)$ für $n > \frac{2}{\varepsilon}$. Jede Umgebung von *g* enthält also einen von *g* verschiedenen Punkt aus *A*, womit *g* auch Häufungspunkt von *A* ist und somit zu $A^+ \subseteq \overline{A}$ gehört. Damit ist aber $\overline{A}$ abgeschlossen.

Für das letzte Resultat benutzen wird das Hilfsergebnis, dass jeder Punkt $g \in K_r(f)$ für ein $f \in \mathcal{X}$ und $r > 0$ innerer Punkt von $K_r(f)$ ist. Sei dazu $s := \frac{1}{2}\{r - \rho(f,g)\} > 0$. Dann ist nämlich $K_s(g) \subseteq K_r(f)$ wegen

$$\rho(h, f) \leq \rho(h, g) + \rho(g, f) < \frac{1}{2}\{r - \rho(f,g)\} + \rho(f,g) < \frac{1}{2}\{r + \rho(f,g)\} < r.$$

Sei nun $f \in A^o$. Dann gibt es ein $r > 0$ mit $K_r(f) \subseteq A$. Nach vorigem existiert aber zu jedem $g \in K_r(f)$ ein $s > 0$ mit $K_s(g) \subseteq K_r(f) \subseteq A$, d.h. jedes $g \in K_r(f)$ ist auch innerer Punkt von *A*, so dass damit $K_r(f) \subseteq A^o$ gilt, was den Beweis abschließt. ■

Bemerkung: Auf die Endlichkeit der Indexmenge *I* in Teil b) von Satz 1 kann nicht verzichtet werden, weil ansonsten $\inf_{i \in I}\{r_i\} = 0$ sein kann. Beispiel: im metrischen Raum aus Beispiel 1 a) sind etwa die Mengen $A_i := \left(-\frac{1}{i}, 1 + \frac{1}{i}\right),\ i \in I := \mathbb{N}$ sämtlich offen, aber das abgeschlossene Intervall $\bigcap_{i \in I} A_i = [0,1]$ ist nicht offen.

Wenn wir in Beispiel 1, Teil a) den metrischen Raum $(\mathcal{X}, \rho) = (\mathbb{Z}, \rho)$ mit der euklidischen Metrik wählen, so gilt offenbar

$$K_r(f) = \overline{K_r(f)} = \{f\} \text{ für alle } f \in \mathcal{X} \text{ und } 0 < r \leq 1, \text{ aber } \overline{K}_1(f) = \{f - 1, f, f + 1\}.$$

Dies zeigt, dass hier $\overline{K_1(f)} \subset \overline{K}_1(f)$ mit $\overline{K_1(f)} \neq \overline{K}_1(f)$ ist für alle $f \in \mathcal{X}$, was die Aussage in der Fußnote zu Definition 2 bestätigt.

Definition 3 (Grenzwert und Cauchy-Folge)**.** Es sei $(\mathcal{X},\rho)$ ein metrischer Raum. Eine Folge $\{f_n\}_{n\in\mathbb{N}}$ von Elementen aus $\mathcal{X}$ heißt *konvergent in* $\mathcal{X}$ *gegen einen Grenzpunkt* $f\in\mathcal{X}$, wenn zu jedem $\varepsilon>0$ ein $N\in\mathbb{N}$ derart existiert, dass gilt:

$$\rho(f_n,f)<\varepsilon \text{ für alle } n\geq N.$$

Man schreibt dann auch: $f=\rho\text{-}\lim\limits_{n\to\infty} f_n$ oder kurz nur $f=\lim\limits_{n\to\infty} f_n$.

Eine Folge $\{f_n\}_{n\in\mathbb{N}}$ in $\mathcal{X}$ heißt *Cauchy-Folge* (oder *Fundamentalfolge*) in $\mathcal{X}$, wenn zu jedem $\varepsilon>0$ ein $N\in\mathbb{N}$ existiert mit:

$$\rho(f_n,f_m)<\varepsilon \text{ für alle } n,m\geq N.$$

Satz 2 (Eigenschaften von Folgen)**.** Es sei $(\mathcal{X},\rho)$ ein metrischer Raum, $A\subseteq\mathcal{X}$ und $\{f_n\}_{n\in\mathbb{N}}$ eine Folge von Elementen aus $\mathcal{X}$. Dann gilt:

a) Die Konvergenz von $\{f_n\}_{n\in\mathbb{N}}$ gegen $f\in\mathcal{X}$ ist äquivalent zu der Aussage: zu jeder offenen Umgebung $U(f)$ existiert ein $N\in\mathbb{N}$ derart, dass gilt:

$$f_n\in U(f) \text{ für alle } n\geq N.$$

b) Der Grenzwert $f=\lim\limits_{n\to\infty} f_n$ ist im Falle der Existenz eindeutig bestimmt.
c) Ein Punkt $f\in\mathcal{X}$ ist genau dann Häufungspunkt von *A*, wenn gilt: es existiert eine Folge $\{f_n\}_{n\in\mathbb{N}}$ von *paarweise verschiedenen* Punkten aus *A* mit $f=\lim\limits_{n\to\infty} f_n$.
d) *A* ist genau dann abgeschlossen, wenn jeder Grenzpunkt $f\in\mathcal{X}$ einer Folge $\{f_n\}_{n\in\mathbb{N}}\subseteq A$ zu *A* gehört.
e) Jede in $\mathcal{X}$ konvergente Folge $\{f_n\}_{n\in\mathbb{N}}$ ist eine Cauchy-Folge.
f) Ist $\{f_n\}_{n\in\mathbb{N}}$ in $\mathcal{X}$ konvergent gegen $f\in\mathcal{X}$ und $\{g_n\}_{n\in\mathbb{N}}$ eine weitere Folge von Punkten in $\mathcal{X}$ mit Grenzwert $g\in\mathcal{X}$, so gilt:

$$\lim_{n\to\infty}\rho(f_n,g_n)=\rho(f,g).$$

g) Sind $\{f_n\}_{n\in\mathbb{N}}$ und $\{g_n\}_{n\in\mathbb{N}}$ Cauchy-Folgen in $\mathcal{X}$, so ist $\{\rho(f_n,g_n)\}_{n\in\mathbb{N}}$ Cauchy-Folge in $\mathbb{R}^1$.

Beweis:

a) Es ist klar, dass die angegebene Bedingung die Konvergenz von $\{f_n\}_{n\in\mathbb{N}}$ gegen $f\in\mathcal{X}$ impliziert, weil man $U(f):=K_\varepsilon(f)$ wählen kann. Sei nun umgekehrt $\{f_n\}_{n\in\mathbb{N}}$ gegen $f\in\mathcal{X}$ konvergent und $U(f)$ eine beliebige offene Umgebung von f. Dann ist f innerer Punkt von $U(f)$, d.h. es existiert ein $\varepsilon>0$ mit $f\in K_\varepsilon(f)\subseteq U(f)$. Wegen der Konvergenz von $\{f_n\}_{n\in\mathbb{N}}$ gegen f existiert also ein $N\in\mathbb{N}$ derart, dass $f_n\in K_\varepsilon(f)\subseteq U(f)$ für alle $n\geq N$ gilt, was zu zeigen war.

b) Seien $f, g \in \mathcal{X}$ mit der Eigenschaft: zu jedem $\varepsilon > 0$ existiert ein $N \in \mathbb{N}$ derart, dass

$$\rho(f_n, f) < \varepsilon \text{ und } \rho(f_n, g) < \varepsilon \text{ für alle } n \geq N.$$

Mit der Dreiecksungleichung ergibt sich dann:

$$\rho(f, g) \leq \rho(f_n, g) + \rho(f_n, f) < 2\varepsilon \text{ für alle } \varepsilon > 0, \text{ also } \rho(f, g) = 0 \text{ und damit } f = g.$$

c) Sei zunächst $f \in \mathcal{X}$ ein Häufungspunkt von *A*. Wir definieren eine streng monoton fallende Folge $\{\varepsilon_n\}_{n\in\mathbb{N}}$ und die gewünschte Folge $\{f_n\}_{n\in\mathbb{N}}$ in *A* rekursiv wie folgt: für ε_1 wählen wir eine beliebige positive reelle Zahl. Dann existiert nach Definition des Häufungspunkts in der offenen Umgebung $K_{\varepsilon_1}(f)$ ein von *f* verschiedenes Element $f_1 \in A$. Sind nun $\varepsilon_1, \cdots, \varepsilon_n$ und $f_1, \cdots, f_n \in A$ für $n \in \mathbb{N}$ bereits konstruiert, so wähle $\varepsilon_{n+1} := \frac{1}{2} \min_{1\leq k\leq n} \{\varepsilon_k, \rho(f_k, f)\}$. Nach Definition des Häufungspunkts existiert dann in der offenen Umgebung $K_{\varepsilon_{n+1}}(f)$ ein von f verschiedenes Element $f_{n+1} \in A$. Dieses ist nach Wahl von ε_{n+1} auch von allen früheren $f_1, \cdots, f_n \in A$ verschieden. Die Folge $\{f_n\}_{n\in\mathbb{N}}$ konvergiert nun in *A* gegen *f*, wie gewünscht.

Ist umgekehrt $\{f_n\}_{n\in\mathbb{N}}$ eine derartige Folge in *A* mit Grenzwert $f \in A$, so ist f Häufungspunkt von *A*, weil nach Definition der Konvergenz in jeder Umgebung $K_\varepsilon(f)$ mit beliebigem $\varepsilon > 0$ unendlich viele, paarweise (und damit auch von *f*) verschiedene Folgenglieder liegen.

d) Sei zunächst *A* abgeschlossen und f Grenzpunkt von $\{f_n\}_{n\in\mathbb{N}}$. Falls $\{f_n\}_{n\in\mathbb{N}}$ eine Teilfolge $\{f_{k_n}\}_{n\in\mathbb{N}}$ aus paarweise verschiedenen Elementen enthält, so ist *f* auch Häufungspunkt von *A* und gehört daher wegen der Abgeschlossenheit von *A* zu *A*. Falls keine solche Teilfolge existiert, folgt $f_n = f$ für genügend große $n \geq N \in \mathbb{N}$, so dass auch hier $f \in A$ folgt. Für die Umkehrung nehmen wir an, dass *g* ein Häufungspunkt von *A* ist. Nach Teil c) dieses Satzes existiert damit eine Folge $\{f_n\}_{n\in\mathbb{N}}$ von (paarweise verschiedenen) Punkten aus *A* mit Grenzwert *g*. Nach Voraussetzung gehört aber *g* zu *A*, also ist *A* abgeschlossen.

e) Dies folgt mit der Dreiecksungleichung aus

$$\rho(f_n, f_m) \leq \rho(f_n, f) + \rho(f_m, f) < 2\varepsilon \text{ für alle } n, m \geq N.$$

f) Nach zweimaliger Anwendung der Dreiecksungleichung erhält man:

$$\rho(f_n, g_n) \leq \rho(f_n, g) + \rho(g_n, g) \leq \rho(f, g) + \rho(f_n, f) + \rho(g_n, g) \quad \text{und}$$
$$\rho(f, g) \leq \rho(f_n, g) + \rho(f_n, f) \leq \rho(f_n, g_n) + \rho(g_n, g) + \rho(f_n, f),$$

woraus

$$\begin{aligned}\rho(f, g) &\leq \lim_{n\to\infty} \rho(f_n, g_n) + \lim_{n\to\infty} \rho(g_n, g) + \lim_{n\to\infty} \rho(f_n, f) \\ &= \lim_{n\to\infty} \rho(f_n, g_n) \leq \rho(f, g) + \lim_{n\to\infty} \rho(f_n, f) + \lim_{n\to\infty} \rho(g_n, g) = \rho(f, g)\end{aligned}$$

und damit Gleichheit folgt.

g) Dies folgt aus den für alle $n, m \in \mathbb{N}$ gültigen Ungleichung

$$\begin{aligned}|\rho(f_n,g_n)-\rho(f_m,g_m)| &\leq |\rho(f_n,g_n)-\rho(f_n,g_m)|+|\rho(f_n,g_m)-\rho(f_m,g_m)| \\ &\leq \rho(f_n,f_m)+\rho(g_n,g_m)\end{aligned}$$

nach weiterer Anwendung der Vierecksungleichung. ■

Definition 4 (Vollständigkeit)**.** Ein metrischer Raum $(\mathcal{X},\rho)$ heißt *vollständig*, wenn jede Cauchy-Folge $\{f_n\}_{n\in\mathbb{N}}$ in $\mathcal{X}$ gegen ein Element $f\in\mathcal{X}$ konvergiert.

Beispiel 3. Die metrischen Räume aus Beispiel 1, Teile a) bis f) sind sämtlich vollständig. Der metrische Raum aus Beispiel 1, Teil g) ist dagegen nicht vollständig:

Wähle etwa $\mathcal{X}=C[0,2]$ und die Funktionenfolge $\{f_n\}_{n\in\mathbb{N}}$ definiert durch $f_n(x):=\min\{x^n,1\}$, $x\in[0,2]$, $n\in\mathbb{N}$. Dann ist $\{f_n\}_{n\in\mathbb{N}}$ Cauchy-Folge, denn es ist

$$\rho(f_n,f_m)=\int_0^2 |f_n(x)-f_m(x)|\,dx=\int_0^1 |x^n-x^m|\,dx=\left|\frac{1}{n+1}-\frac{1}{m+1}\right|\to 0 \text{ für } n,m\to\infty,$$

aber es liegt keine Konvergenz gegen ein Element $f\in\mathcal{X}$ in der Metrik ρ vor, da mit $f:=\mathbb{1}_{[1,2]}$[5] zwar

$$\int_0^2 |f_n(x)-f(x)|\,dx=\int_0^1 x^n\,dx=\frac{x^{n+1}}{n+1}\Bigg|_0^1=\frac{1}{n+1}\to 0 \text{ für } n\to\infty$$

gilt; jedoch ist f kein Element von $\mathcal{X}$ und damit dieser metrische Raum $(\mathcal{X},\rho)$ nicht vollständig.

Für den Rest dieses Abschnitts wollen wir noch die Stetigkeit von Abbildungen zwischen metrischen Räumen betrachten.

Definition 5 (Stetigkeit)**.** Es seien $(\mathcal{X},\rho_{\mathcal{X}})$ und $(\mathcal{Y},\rho_{\mathcal{Y}})$ metrische Räume. Eine Abbildung $T:\mathcal{X}\to\mathcal{Y}$ heißt *stetig* im Punkt $f\in\mathcal{X}$, wenn zu jeder (offenen) Umgebung $\mathcal{V}$ von Tf[6] in $\mathcal{Y}$ eine (offene) Umgebung $\mathcal{U}$ von f in $\mathcal{X}$ existiert mit $T(\mathcal{U})\subseteq\mathcal{V}$. T heißt *stetig auf* $\mathcal{X}$, wenn T stetig in jedem Punkt $f\in\mathcal{X}$ ist.

Lemma 1 (Äquivalenzen zur Stetigkeit I)**.** Unter den Voraussetzungen von Definition 5 sind folgende Aussagen äquivalent zur Stetigkeit von T in $f\in\mathcal{X}$:

a) Für jede Umgebung $\mathcal{V}$ von Tf enthält $T^{-1}(\mathcal{V})$ eine Umgebung $\mathcal{U}$ von f in $\mathcal{X}$.
b) Zu jedem $\varepsilon>0$ existiert ein $\delta>0$ mit der Eigenschaft:

$$\rho_{\mathcal{X}}(f,g)<\delta \;\Rightarrow\; \rho_{\mathcal{Y}}(Tf,Tg)<\varepsilon \text{ für alle } g\in\mathcal{X}.$$

c) Für jede Folge $\{f_n\}_{n\in\mathbb{N}}$ in $\mathcal{X}$ mit $\lim_{n\to\infty}\rho_{\mathcal{X}}(f_n,f)=0$ gilt $\lim_{n\to\infty}\rho_{\mathcal{Y}}(Tf_n,Tf)=0$.

---

[5] Hier bezeichnet $\mathbb{1}_A$ wie üblich die Indikatorfunktion zur Menge A.

[6] Wir benutzen hier abkürzend die Schreibweise Tf statt $T(f)$.

Beweis:

a) Folgt aus der Äquivalenz $T(\mathcal{U}) \subseteq \mathcal{V} \;\Leftrightarrow\; \mathcal{U} \subseteq T^{-1}(\mathcal{V})$.

b) Sei T stetig in f. Wähle $\mathcal{V} = K_\varepsilon(Tf;\mathcal{Y})$ in $\mathcal{Y}$. Dann existiert eine Umgebung $\mathcal{U}$ von f in $\mathcal{X}$ mit $T(\mathcal{U}) \subseteq \mathcal{V}$. Nach Definition der Umgebung existiert dann ein $\delta > 0$ mit $K_\delta(f;\mathcal{X}) \subseteq \mathcal{U}$, so dass $T(K_\delta(f;\mathcal{X})) \subseteq T(\mathcal{U}) \subseteq \mathcal{V}$ ist, woraus die Beziehung "$\rho_\mathcal{X}(f,g) < \delta \;\Rightarrow\; \rho_\mathcal{Y}(Tf,Tg) < \varepsilon$ für alle $g \in \mathcal{X}$" folgt. Sei umgekehrt diese Beziehung vorausgesetzt und $\mathcal{V}$ eine Umgebung von Tf in $\mathcal{Y}$. Dann existiert ein $\varepsilon > 0$ so, dass $K_\varepsilon(Tf;\mathcal{Y}) \subseteq \mathcal{V}$ ist. Nach Voraussetzung existiert also ein $\delta > 0$ mit "$\rho_\mathcal{X}(f,g) < \delta \;\Rightarrow\; \rho_\mathcal{Y}(Tf,Tg) < \varepsilon$ für alle $g \in \mathcal{X}$". Mit der Wahl $\mathcal{U} = K_\delta(f;\mathcal{X})$ erhält man nun die gewünschte Umgebung $\mathcal{U}$ von f in $\mathcal{X}$ mit $T(\mathcal{U}) \subseteq \mathcal{V}$.

c) Folgt unmittelbar aus Teil b). ■

Satz 3 (Äquivalenzen zur Stetigkeit II)**.** Es seien $(\mathcal{X},\rho_\mathcal{X})$ und $(\mathcal{Y},\rho_\mathcal{Y})$ metrische Räume und $T:\mathcal{X} \to \mathcal{Y}$ eine Abbildung. Folgende Aussagen sind äquivalent:

a) T ist stetig auf $\mathcal{X}$.
b) Für jede offene Menge $B \subseteq \mathcal{Y}$ ist $T^{-1}(B)$ offen in $\mathcal{X}$.
c) Für jede abgeschlossene Menge $C \subseteq \mathcal{Y}$ ist $T^{-1}(C)$ abgeschlossen in $\mathcal{X}$.
d) Es ist $T(\overline{A}) \subseteq \overline{T(A)}$ für alle Teilmengen $A \subseteq \mathcal{X}$.

Beweis: Wir führen einen Ringschluss:

a) $\Rightarrow$ b): Sei $B \subseteq \mathcal{Y}$ offen. Falls $T^{-1}(B) = \varnothing$ gilt, ist die Aussage trivial. Anderenfalls existiert ein $f \in \mathcal{X}$ mit $Tf \in B$. Da B offen ist, kann als Umgebung $\mathcal{V}$ von Tf B selbst gewählt werden. T ist nun insbesondere stetig in f, also existiert eine Umgebung $\mathcal{U}$ von f mit $T(\mathcal{U}) \subseteq \mathcal{V}$ bzw. äquivalent $\mathcal{U} \subseteq T^{-1}(\mathcal{V}) \subseteq T^{-1}(B)$, was zu zeigen war.

b) $\Rightarrow$ c): Folgt durch Übergang zu Mengenkomplementen mit Satz 1, Teil a).

c) $\Rightarrow$ d): Die Menge $B := T^{-1}\left(\overline{T(A)}\right)$ ist abgeschlossen, mit $A \subseteq B$ wegen $T(A) \subseteq \overline{T(A)}$. Es folgt $\overline{A} \subseteq B$ (da B abgeschlossen ist) und somit $T(\overline{A}) \subseteq T(B) \subseteq \overline{T(A)}$, was zu zeigen war.

d) $\Rightarrow$ a): Wir zeigen zunächst, dass aus Teil d) dieses Satzes der Teil c) aus Lemma 1 folgt. Sei dazu $\{f_n\}_{n\in\mathbb{N}}$ eine Folge in $\mathcal{X}$ mit $\lim_{n\to\infty} \rho_\mathcal{X}(f_n, f) = 0$ für ein $f \in \mathcal{X}$. Mit der Wahl $A := \{f_n \mid n \in \mathbb{N}\}$ (aufgefasst als Menge!) ergibt sich $\overline{A} = A \cup \{f\}$ sowie $T(A) = \{Tf_n \mid n \in \mathbb{N}\}$. Nach Voraussetzung gilt nun $T(\overline{A}) = \{Tf_n \mid n \in \mathbb{N}\} \cup \{Tf\} \subseteq \overline{T(A)}$, so dass entweder schon $Tf \in T(A)$ gilt oder Tf ein Häufungspunkt von $T(A) = \{Tf_n \mid n \in \mathbb{N}\}$ ist, also in jedem Fall $\lim_{n\to\infty} \rho_\mathcal{Y}(Tf_n, Tf) = 0$ folgt. Da $\{f_n\}_{n\in\mathbb{N}}$ mit Grenzwert f beliebig war, folgt die Stetigkeit von T in jedem Punkt $f \in \mathcal{X}$, also a). ■

Bemerkung: Ist $(\mathcal{X}, \rho)$ ein metrischer Raum, so ist die Metrik $\rho\colon\ \mathcal{X}\times\mathcal{X}\to\mathbb{R}$ aufgrund der Vierecksungleichung selbst auch eine stetige Abbildung, d.h. es ist

$$\lim_{n\to\infty}\rho(f_n, g_n) = \rho\left(\lim_{n\to\infty} f_n, \lim_{n\to\infty} g_n\right) \text{ für alle konvergenten Folgen } \{f_n\}_{n\in\mathbb{N}}, \{g_n\}_{n\in\mathbb{N}} \text{ in } \mathcal{X},$$

denn mit $f := \lim\limits_{n\to\infty} f_n,\ g := \lim\limits_{n\to\infty} g_n$ gilt ja

$$0 \le |\rho(f_n, g_n) - \rho(f, g)| \le \rho(f_n, f) + \rho(g_n, g) \to 0 \text{ für } n\to\infty.$$

I. 1.1. Vervollständigung metrischer Räume

In diesem Abschnitt wollen wir als Hauptresultat zeigen, dass jeder (ggf. unvollständige) metrische Raum in geeigneter Weise vervollständigt werden kann.

Definition 6 (Isometrie)**.** Es seien $(\mathcal{X}, \rho_{\mathcal{X}})$ und $(\mathcal{Y}, \rho_{\mathcal{Y}})$ metrische Räume. Eine bijektive Abbildung $T:\mathcal{X}\to\mathcal{Y}$ heißt *isometrisch*, wenn gilt:

$$\rho_{\mathcal{X}}(f, g) = \rho_{\mathcal{Y}}(Tf, Tg) \quad \text{für alle } f, g\in\mathcal{X}.$$

T heißt dann auch *Isometrie* (zwischen den metrischen Räumen $(\mathcal{X}, \rho_{\mathcal{X}})$ und $(\mathcal{Y}, \rho_{\mathcal{Y}})$). Zwei metrische Räume heißen entsprechend *isometrisch*, falls eine Isometrie zwischen ihnen existiert.

Bemerkung: Mit *T* ist natürlich auch die Umkehrabbildung T^{-1} eine Isometrie (zwischen den metrischen Räumen $(\mathcal{Y}, \rho_{\mathcal{Y}})$ und $(\mathcal{X}, \rho_{\mathcal{X}})$).

Definition 7 (Homöomorphismus)**.** Es seien $(\mathcal{X}, \rho_{\mathcal{X}})$ und $(\mathcal{Y}, \rho_{\mathcal{Y}})$ metrische Räume. Eine bijektive Abbildung $T:\mathcal{X}\to\mathcal{Y}$ heißt *Homöomorphismus*, wenn sowohl *T* als auch T^{-1} stetig sind. Die Räume $(\mathcal{X}, \rho_{\mathcal{X}})$ und $(\mathcal{Y}, \rho_{\mathcal{Y}})$ heißen entsprechend *homöomorph*, falls ein Homöomorphismus zwischen ihnen existiert.

Bemerkung: Trivialerweise ist jede Isometrie ein Homöomorphismus, nicht jedoch umgekehrt.

Definition 8 (Dichtheit)**.** Es sei $(\mathcal{X}, \rho)$ ein metrischer Raum. Eine Teilmenge $A\subseteq\mathcal{X}$ heißt *dicht in* $\mathcal{X}$, wenn $\overline{A} = \mathcal{X}$ gilt.

Bemerkung: Äquivalent dazu ist: zu jedem $\varepsilon > 0$ und jedem $f\in\mathcal{X}$ existiert ein $g\in A$ mit $\rho(f, g) < \varepsilon$.

Definition 9 (Vervollständigung)**.** Es sei $(\mathcal{X}, \rho)$ ein metrischer Raum. Ein metrischer Raum $(\tilde{\mathcal{X}}, \tilde{\rho})$ heißt *Vervollständigung* von $(\mathcal{X}, \rho)$, wenn folgendes gilt:

a) $(\mathcal{X}, \rho)$ ist isometrisch zu $(\tilde{\mathcal{Y}}, \tilde{\rho})$, wobei $\tilde{\mathcal{Y}}$ dicht in $\tilde{\mathcal{X}}$ ist;

b) $(\tilde{\mathcal{X}}, \tilde{\rho})$ ist vollständig.

Satz 4 (Vervollständigungssatz I)**.** Jeder (unvollständige) metrische Raum $(\mathcal{X},\rho)$ besitzt (mindestens) eine Vervollständigung $(\tilde{\mathcal{X}},\tilde{\rho})$. Ist $(\underset{\sim}{\mathcal{X}},\underset{\sim}{\rho})$ eine weitere Vervollständigung von $(\mathcal{X},\rho)$, so sind $(\tilde{\mathcal{X}},\tilde{\rho})$ und $(\underset{\sim}{\mathcal{X}},\underset{\sim}{\rho})$ isometrisch, d.h. "die" Vervollständigung ist eindeutig bis auf Isometrie.

Beweis: Sei $\hat{\mathcal{X}}$ die Menge aller Cauchy-Folgen $\mathbf{f}=\{f_n\}_{n\in\mathbb{N}}$ in $\mathcal{X}$. Wir definieren eine Äquivalenzrelation auf $\hat{\mathcal{X}}\times\hat{\mathcal{X}}$ vermöge

$$\mathbf{f}\sim\mathbf{g} \;:\Leftrightarrow\; \lim_{n\to\infty}\rho(f_n,g_n)=0.$$

Für $\mathbf{f}\in\hat{\mathcal{X}}$ bezeichne $\tilde{\mathbf{f}}$ die zugehörige Äquivalenzklasse. Sei nun $\tilde{\mathcal{X}}$ die Menge aller so gebildeten Äquivalenzklassen von Cauchy-Folgen. Wir definieren eine Metrik $\tilde{\rho}$ auf $\tilde{\mathcal{X}}$ vermöge

$$\tilde{\rho}(\tilde{\mathbf{f}},\tilde{\mathbf{g}}):=\lim_{n\to\infty}\rho(f_n,g_n) \text{ für alle } \tilde{\mathbf{f}},\tilde{\mathbf{g}}\in\tilde{\mathcal{X}}.$$

Dieser Ausdruck ist wohldefiniert, da nach Satz 2, Teil g) mit $\mathbf{f}$ und $\mathbf{g}$ ebenfalls $\{\rho(f_n,g_n)\}_{n\in\mathbb{N}}$ Cauchy-Folge (in $\mathbb{R}^1$) ist und jede Cauchy-Folge wegen der Vollständigkeit von $\mathbb{R}^1$ dort auch konvergiert. Ferner ist $\tilde{\rho}(\tilde{\mathbf{f}},\tilde{\mathbf{g}})$ unabhängig vom gewählten Repräsentanten der jeweiligen Äquivalenzklasse, denn für $\tilde{\mathbf{f}}_1=\tilde{\mathbf{f}}_2$, $\tilde{\mathbf{g}}_1=\tilde{\mathbf{g}}_2$ folgt $\lim\limits_{n\to\infty}\rho(f_{1n},g_{1n})=\lim\limits_{n\to\infty}\rho(f_{2n},g_{2n})$ wegen

$$\left|\rho(f_{1n},g_{1n})-\rho(f_{2n},g_{2n})\right|\le\rho(f_{1n},f_{2n})+\rho(g_{1n},g_{2n})\to 0 \text{ für } n\to\infty$$

nach der Vierecksungleichung. Die Eigenschaften einer Metrik lassen sich leicht durch direktes Nachrechnen zeigen. Die ursprüngliche Menge $\mathcal{X}$ lässt sich nunmehr isometrisch in $\tilde{\mathcal{X}}$ einbetten, indem jedem $f\in\mathcal{X}$ zunächst die konstante (Cauchy-)Folge $\mathbf{f}^\circ$ mit $f_n^\circ:=f$ für alle $n\in\mathbb{N}$ und darüber die Äquivalenzklasse $\tilde{\mathbf{f}}^\circ$ zugeordnet wird. $\tilde{\mathbf{f}}^\circ$ besteht damit aus allen Folgen, die gegen f konvergieren. Insbesondere gilt

$$\tilde{\rho}(\tilde{\mathbf{f}}^\circ,\tilde{\mathbf{g}}^\circ)=\rho(f,g) \text{ für alle } f,g\in\mathcal{X}.$$

Wir zeigen zunächst, dass die Menge $\tilde{\mathcal{Y}}:=\{\mathbf{f}^\circ \mid f\in\mathcal{X}\}$ dicht in $\tilde{\mathcal{X}}$ ist. Sei dazu $\varepsilon>0$ und $\tilde{\mathbf{g}}\in\tilde{\mathcal{X}}$ vorgegeben. Für die Repräsentanten-Cauchy-Folge $\mathbf{g}$ existiert also ein $N\in\mathbb{N}$ mit

$$\rho(g_n,g_m)<\varepsilon \text{ für alle } n,m\ge N.$$

Mit der Setzung $\tilde{\mathbf{f}}_n:=\tilde{\mathbf{g}}_n^\circ$ mit $\mathbf{g}_n^\circ=\{g_n,g_n,g_n,\cdots\}$ für alle $n\in\mathbb{N}$ erhält man nun eine Folge $\{\tilde{\mathbf{f}}_n\}_{n\in\mathbb{N}}$ in $\tilde{\mathcal{Y}}$ mit $\tilde{\mathbf{g}}=\tilde{\rho}\text{-}\lim\limits_{n\to\infty}\tilde{\mathbf{f}}_n$, denn es ist

$$\tilde{\rho}(\tilde{\mathbf{f}}_n,\tilde{\mathbf{g}})=\lim_{m\to\infty}\rho(f_{nm},g_m)=\lim_{m\to\infty}\rho(g_n,g_m)<\varepsilon \text{ für alle } n\ge N.$$

Es bleibt zu zeigen, dass in $(\tilde{\mathcal{X}},\tilde{\rho})$ alle Cauchy-Folgen konvergieren. Sei dazu $\{\tilde{\mathbf{g}}_n\}_{n\in\mathbb{N}}$ eine beliebige Cauchy-Folge in $\tilde{\mathcal{X}}$. Zu $\varepsilon>0$ existiert also ein $N\in\mathbb{N}$ mit $\tilde{\rho}(\tilde{\mathbf{g}}_n,\tilde{\mathbf{g}}_m)<\varepsilon$ für alle $n,m\ge N$. Zu jedem $\tilde{\mathbf{g}}_n$ existiert ferner wegen der Dichtheit von $\tilde{\mathcal{Y}}$ in $\mathcal{X}$ ein $\tilde{\mathbf{f}}_n=\{f_n,f_n,f_n,\cdots\}\in\tilde{\mathcal{Y}}$ mit

$\tilde{\rho}\left(\tilde{\mathbf{g}}_n,\tilde{\mathbf{f}}_n\right)<\dfrac{1}{n}$ für alle $n\in\mathbb{N}$. Den gewünschten Grenzwert $\tilde{\mathbf{h}}$ von $\{\tilde{\mathbf{g}}_n\}_{n\in\mathbb{N}}$ erhält man nun durch die Setzung

$$h_n := f_n \text{ für alle } n\in\mathbb{N}.$$

Zunächst ist $\mathbf{h}=\{f_n\}_{n\in\mathbb{N}}$ Cauchy-Folge in $\mathcal{X}$, denn es gilt

$$\rho(f_n,f_m)=\tilde{\rho}\left(\tilde{\mathbf{f}}_n,\tilde{\mathbf{f}}_m\right)\le\tilde{\rho}\left(\tilde{\mathbf{f}}_n,\tilde{\mathbf{g}}_n\right)+\tilde{\rho}\left(\tilde{\mathbf{g}}_n,\tilde{\mathbf{g}}_m\right)+\tilde{\rho}\left(\tilde{\mathbf{g}}_m,\tilde{\mathbf{f}}_m\right)<\frac{1}{n}+\varepsilon+\frac{1}{m}\le 3\varepsilon$$

für alle $n,m\ge M:=\max\left\{N,\dfrac{1}{\varepsilon}\right\}$. Also ist $\tilde{\mathbf{h}}\in\tilde{\mathcal{X}}$. Es ist aber auch $\tilde{\mathbf{h}}$ Grenzwert von $\{\tilde{\mathbf{g}}_n\}_{n\in\mathbb{N}}$, denn es gilt

$$\tilde{\rho}\left(\tilde{\mathbf{g}}_n,\tilde{\mathbf{h}}\right)\le\tilde{\rho}\left(\tilde{\mathbf{f}}_n,\tilde{\mathbf{g}}_n\right)+\tilde{\rho}\left(\tilde{\mathbf{f}}_n,\tilde{\mathbf{h}}\right)<\frac{1}{n}+\lim_{m\to\infty}\rho(f_n,f_m)\le 4\varepsilon$$

für alle $n\ge M=\max\left\{N,\dfrac{1}{\varepsilon}\right\}$. Also ist $\left(\tilde{\mathcal{X}},\tilde{\rho}\right)$ eine Vervollständigung von $(\mathcal{X},\rho)$, womit der erste Teil des Satzes bewiesen ist. Zum Beweis des zweiten Teils betrachten wir zwei Vervollständigungen $\left(\tilde{\mathcal{X}},\tilde{\rho}\right)$ und $\left(\underset{\sim}{\mathcal{X}},\underset{\sim}{\rho}\right)$ von $(\mathcal{X},\rho)$, mit den zu $\mathcal{X}$ isometrischen dichten Teilmengen $\tilde{\mathcal{Y}}$ bzw. $\underset{\sim}{\mathcal{Y}}$ sowie das Abbildungsdiagramm

$$\begin{tikzcd}
\hat{\mathcal{X}} \arrow[r, "T_1"] \arrow[d, "T_2"] & \hat{\tilde{\mathcal{Y}}} \arrow[r, "S_1"] & \tilde{\mathcal{X}} \arrow[ddll, shift left, "Z^{-1}"] \\
\hat{\underset{\sim}{\mathcal{Y}}} \arrow[d, "S_2"] & & \\
\underset{\sim}{\mathcal{X}} \arrow[uurr, shift left, "Z"] & &
\end{tikzcd}$$

wobei hier $\hat{\mathcal{X}}$ die Menge der Äquivalenzklassen von Cauchy-Folgen in $\mathcal{X}$ und entsprechend $\hat{\tilde{\mathcal{Y}}}$ und $\hat{\underset{\sim}{\mathcal{Y}}}$ die Mengen der Äquivalenzklassen von Cauchy-Folgen in $\tilde{\mathcal{Y}}$ bzw. in $\underset{\sim}{\mathcal{Y}}$ bezeichnen. Die Abbildungen T_1 und T_2 seien folgendermaßen definiert: sind I_1 und I_2 die Isometrien zwischen $(\mathcal{X},\rho)$ und $\left(\tilde{\mathcal{Y}},\tilde{\rho}\right)$ bzw. zwischen $(\mathcal{X},\rho)$ und $\left(\underset{\sim}{\mathcal{Y}},\underset{\sim}{\rho}\right)$, so sei

$$T_i\tilde{\mathbf{f}}:=\widetilde{I_i\mathbf{f}} \text{ für } i=1,2 \text{ und alle } \tilde{\mathbf{f}}\in\hat{\mathcal{X}}.$$

Mit ~ sei wie oben der Übergang zu Äquivalenzklassen bezeichnet. T_1 und T_2 sind dann Isometrien zwischen $\hat{\mathcal{X}}$ und $\hat{\tilde{\mathcal{Y}}}$ bzw. zwischen $\hat{\mathcal{X}}$ und $\hat{\underset{\sim}{\mathcal{Y}}}$ bezüglich der Metriken $\hat{\tilde{\rho}}$ und $\hat{\underset{\sim}{\rho}}$, definiert durch

$$\hat{\tilde{\rho}}\left(T_1\tilde{\mathbf{f}},T_1\tilde{\mathbf{g}}\right):=\hat{\underset{\sim}{\rho}}\left(T_2\tilde{\mathbf{f}},T_2\tilde{\mathbf{g}}\right):=\rho(\mathbf{f},\mathbf{g}) \text{ für alle } \mathbf{f},\mathbf{g}\in\hat{\mathcal{X}}.$$

Dabei ist wie oben die rechte Seite unabhängig vom speziellen Repräsentanten der Äquivalenzklasse. Die Abbildungen S_1 und S_2 seien ferner definiert durch

$$S_1\left(\tilde{\mathbf{h}}\right):=\tilde{\rho}\text{-}\lim_{n\to\infty}h_n \text{ für alle } \tilde{\mathbf{h}}\in\hat{\tilde{\mathcal{Y}}} \quad\text{sowie}\quad S_2\left(\tilde{\mathbf{k}}\right):=\underset{\sim}{\rho}\text{-}\lim_{n\to\infty}k_n \text{ für alle } \tilde{\mathbf{k}}\in\hat{\underset{\sim}{\mathcal{Y}}}.$$

Diese Abbildungen sind wohldefiniert, da $\tilde{\mathcal{X}}$ und $\underset{\sim}{\mathcal{X}}$ vollständig und die jeweiligen rechten Seiten wieder unabhängig von der speziellen Wahl der Repräsentanten sind. Wir wollen zeigen, dass S_1 und S_2 Isometrien sind. Dann existiert nämlich auch eine Isometrie *Z* zwischen $\left(\underset{\sim}{\mathcal{X}},\underset{\sim}{\rho}\right)$ und $\left(\tilde{\mathcal{X}},\tilde{\rho}\right)$ im Sinne des obigen Diagramms, womit der Satz bewiesen wäre. Die Abbildungen S_1 und S_2 sind jedenfalls surjektiv, da wegen der jeweiligen Dichtheitseigenschaften jedes Element aus $\tilde{\mathcal{X}}$ bzw. aus $\underset{\sim}{\mathcal{X}}$ Grenzwert einer geeigneten Cauchy-Folge in $\tilde{\mathcal{Y}}$ bzw. in $\underset{\sim}{\mathcal{Y}}$ ist. Sie sind auch injektiv (und damit bijektiv), denn zwei Cauchy-Folgen besitzen genau dann denselben Grenzwert, wenn sie äquivalent sind. Ferner ist (vgl. die Bemerkung auf S. 11 unten)

$$\widehat{\tilde{\rho}}\left(\tilde{\mathbf{h}}_1,\tilde{\mathbf{h}}_2\right)=\lim_{n\to\infty}\tilde{\rho}\left(h_{1n},h_{2n}\right)=\tilde{\rho}\left(\lim_{n\to\infty}h_{1n},\lim_{n\to\infty}h_{2n}\right) \text{ für alle } \tilde{\mathbf{h}}_1,\tilde{\mathbf{h}}_2\in\widehat{\tilde{\mathcal{Y}}} \text{ sowie}$$

$$\widehat{\underset{\sim}{\rho}}\left(\tilde{\mathbf{h}}_1,\tilde{\mathbf{h}}_2\right)=\lim_{n\to\infty}\underset{\sim}{\rho}\left(h_{1n},h_{2n}\right)=\underset{\sim}{\rho}\left(\lim_{n\to\infty}h_{1n},\lim_{n\to\infty}h_{2n}\right) \text{ für alle } \tilde{\mathbf{h}}_1,\tilde{\mathbf{h}}_2\in\widehat{\underset{\sim}{\mathcal{Y}}}.$$

Damit sind also die Abbildungen S_1 und S_2 Isometrien und daher auch $Z=S_1\circ T_1\circ T_2^{-1}\circ S_2^{-1}$ bzw. $Z^{-1}=S_2\circ T_2\circ T_1^{-1}\circ S_1^{-1}$, wie zu zeigen war. ∎

Beispiel 4 (Vervollständigungen metrischer Räume)**.**

a) Wählen wir in Beispiel 1, Teil a) als Grundmenge $\mathcal{X}=\mathbb{Q}$, so ist $(\mathbb{Q},\rho)$ mit der euklidischen Metrik nicht vollständig, weil etwa die Folge **f** mit $f_n:=\left(1+\frac{1}{n}\right)^n$, $n\in\mathbb{N}$ zwar eine Cauchy-Folge in $\mathbb{Q}$ ist, aber dort nicht bezüglich der Metrik ρ konvergiert. Eine geeignete Vervollständigung ist hier offensichtlich der metrische Raum $(\mathbb{R},\rho)$ mit dem Grenzwert $\lim_{n\to\infty}f_n=e$ [Euler'sche Zahl].

b) Wählen wir in Beispiel 1, Teil f) als Grundmenge $\mathcal{X}$ die Menge aller Polynome $\mathcal{P}[a,b]$ über dem Intervall $[a,b]$, so ist $(\mathcal{P}[a,b],\rho)$ mit der Supremumsmetrik nicht vollständig, weil etwa die Folge **f** mit $f_n(x):=\left(1+\frac{x}{n}\right)^n$, $n\in\mathbb{N}$; $x\in[a,b]$ zwar eine Cauchy-Folge in $\mathcal{P}[a,b]$ ist, aber dort nicht bezüglich der Metrik ρ konvergiert. Eine Vervollständigung ist hier der metrische Raum $(C[a,b],\rho)$ mit dem (nicht-polynomialen) Grenzwert $\lim_{n\to\infty}f_n(x)=e^x$, $x\in[a,b]$ [Exponentialfunktion]. Ein konstruktiver Beweis für diese Aussage (*Weierstraß'scher Approximationssatz*) kann mit Hilfe der so genannten *Bernstein-Polynome* $B_n f\in\mathcal{P}[a,b]$ für jedes $f\in C[a,b]$ geführt werden, denn es gilt:

$$\lim_{n\to\infty}B_n f(x):=\frac{1}{(b-a)^n}\sum_{k=0}^{n}\binom{n}{k}f\left(a+(b-a)\frac{k}{n}\right)(x-a)^k(b-x)^{n-k}=f(x)$$

gleichmäßig in $x\in[a,b]$ für alle $f\in C[a,b]$. (Dies folgt übrigens aus dem Gesetz der großen Zahlen für Binomialverteilungen.)

c) Eine Vervollständigung des metrischen Raumes aus Beipiel 1, Teil g) ist gegeben durch den metrischen Raum der bezüglich des Lebesgue-Maßes $\mathfrak{m}^1$ über $[a,b]$ integrierbaren Funktionen[7], mit derselben formalen Definition der Metrik ρ. Dies folgt aus dem *Satz von Lusin* (vgl. BAUER (1992), Satz 26.7, ELSTRODT (1996), Satz 1.18 und Korollar 1.19 in Kapitel VIII, oder RUDIN (1987), Satz 2.24). Danach gibt es zu jeder Lebesgue-messbaren Funktion f auf $[a,b]$ und zu jedem $\varepsilon > 0$ eine Lebesgue-messbare Menge $A_\varepsilon \subseteq [a,b]$ mit $\mathfrak{m}^1(A_\varepsilon) < \varepsilon$ sowie eine stetige Funktion $f_\varepsilon \in C[a,b]$ mit $f(x) = f_\varepsilon(x)$ für alle $x \in A_\varepsilon^c$ (man sagt auch, f sei "fast stetig" [nicht zu verwechseln mit: "fast überall stetig"]). Insbesondere folgt

$$\begin{aligned}\rho(f, f_\varepsilon) &:= \int_a^b |f(x) - f_\varepsilon(x)|\,\mathfrak{m}^1(dx) = \int_{A_\varepsilon} |f(x) - f_\varepsilon(x)|\,\mathfrak{m}^1(dx) \\ &\leq \int_{A_\varepsilon} |f(x)|\,\mathfrak{m}^1(dx) + \varepsilon \cdot \sup_{x\in[a,b]} \left\{ |f_\varepsilon(x)| \right\}\end{aligned}$$

mit $\lim_{\varepsilon \downarrow 0} \int_{A_\varepsilon} |f(x)|\,\mathfrak{m}^1(dx) = 0$ nach dem Lebesgue'schen *Satz von der majorisierten Konvergenz* (vgl. das Skript STOCHASTIK, Satz 19), da $|f| \cdot \mathbb{1}_{A_\varepsilon}$ für $\varepsilon \downarrow 0$ $\mathfrak{m}^1$-fast überall punktweise gegen Null konvergiert und durch $|f|$ majorisiert ist.

I. 1.2. Kategorien von Mengen, Kompaktheit, Separabilität

Definition 10 (Durchmesser)**.** Es sei $(\mathcal{X}, \rho)$ ein metrischer Raum und $A \subseteq \mathcal{X}$ nicht-leer. Die (ggf. unendliche) Zahl

$$\delta(A) := \sup_{f,g\in A} \rho(f, g)$$

heißt *Durchmesser von A*. Ist $\delta(A)$ endlich, so heißt *A beschränkt*.

Satz 5 (Cantor'scher Durchschnittssatz)**.** Ein metrischer Raum $(\mathcal{X}, \rho)$ ist genau dann vollständig, wenn für jede monoton fallende Folge $\{A_n\}_{n\in\mathbb{N}}$ nicht-leerer abgeschlossener Mengen in $\mathcal{X}$ mit $\lim_{n\to\infty} \delta(A_n) = 0$ gilt:

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

In diesem Fall gilt genauer: $\bigcap_{n=1}^{\infty} A_n \neq \varnothing$ besteht aus genau einem Punkt.

[7] Genauer gesagt, der Raums der Äquivalenzklassen von fast überall gleichen integrierbaren Funktionen. Anderenfalls ergibt sich bei ρ lediglich eine so genannte *Pseudo-Metrik*, bei der nur $f = g \Rightarrow \rho(f, g) = 0$ für alle $f, g \in \mathcal{X}$ gilt. Hierauf kommen wir in Abschnitt I.1.5 noch einmal genauer zurück.

Beweis: "$\Rightarrow$": Für $n \in \mathbb{N}$ wähle $f_n \in A_n$. Wegen der Antitonie der Mengenfolge ist dann auch $f_m \in A_n$ für alle $m \geq n$, woraus $\rho(f_n, f_m) \leq \delta(A_n)$ folgt für $m \geq n$. Damit ist $\{f_n\}_{n\in\mathbb{N}}$ aber Cauchy-Folge in $\mathcal{X}$, also existiert ein Grenzwert $f \in \mathcal{X}$ mit $\lim\limits_{n\to\infty} \rho(f_n, f) = 0$. Wir zeigen, dass $f \in A_n$ gilt für alle $n \in \mathbb{N}$; dann ist auch $f \in \bigcap\limits_{n=1}^{\infty} A_n$ und damit $\bigcap\limits_{n=1}^{\infty} A_n \neq \varnothing$. Sei dazu $n \in \mathbb{N}$ fest gewählt. Die Folge $\{f_{n+k}\}_{k\in\mathbb{N}}$ liegt nach obigem ganz in A_n mit $\lim\limits_{k\to\infty} \rho(f_{n+k}, f) = 0$. Nach Satz 2, Teil d) ist also aufgrund der Abgeschlossenheit $f \in A_n$, was zu zeigen war.

$\bigcap\limits_{n=1}^{\infty} A_n \neq \varnothing$ besteht darüber hinaus aus genau einem Punkt, denn wenn wir annehmen, dass ein weiterer, von f verschiedener Punkt g in $\bigcap\limits_{n=1}^{\infty} A_n$ liegt, erhalten wir einen Widerspruch, weil aus $f \neq g$ folgt: $\delta\left(\bigcap\limits_{n=1}^{\infty} A_n\right) \geq \rho(f, g) > 0$ und zugleich $\delta\left(\bigcap\limits_{n=1}^{\infty} A_n\right) \leq \delta(A_m)$ für alle $m \in \mathbb{N}$ mit $\lim\limits_{n\to\infty} \delta(A_n) = 0$.

"$\Leftarrow$": Sei $\{f_n\}_{n\in\mathbb{N}}$ Cauchy-Folge in $\mathcal{X}$. Wir konstruieren rekursiv eine monoton fallende Folge $\{A_n\}_{n\in\mathbb{N}}$ nicht-leerer abgeschlossener Mengen in $\mathcal{X}$ mit $\lim\limits_{n\to\infty} \delta(A_n) = 0$, deren Durchschnitt gerade den Grenzwert von $\{f_n\}_{n\in\mathbb{N}}$ als (einziges) Element enthält, womit die Aussage gezeigt ist. Zunächst existiert ein $n_1 \in \mathbb{N}$ mit $\rho\left(f_n, f_{n_1}\right) < \frac{1}{2}$ für $n \geq n_1$. Wähle die abgeschlossene Kugel $A_1 := \bar{K}_1\left(f_{n_1}\right)$. Ist der Index n_k für $k \in \mathbb{N}$ bereits konstruiert, so existiert ein $n_{k+1} > n_k$ mit $\rho\left(f_n, f_{n_{k+1}}\right) < \frac{1}{2^{k+1}}$ für $n \geq n_{k+1}$. Wähle die abgeschlossene Kugel $A_{k+1} := \bar{K}_{2^{-k}}\left(f_{n_{k+1}}\right)$. Die so konstruierte Folge $\{A_n\}_{n\in\mathbb{N}}$ ist monoton fallend mit $\lim\limits_{n\to\infty} \delta(A_n) = \lim\limits_{n\to\infty} \frac{1}{2^{n-1}} = 0$, so dass nach Voraussetzung $\bigcap\limits_{n=1}^{\infty} A_n \neq \varnothing$ ist. Sei nun $f \in \bigcap\limits_{n=1}^{\infty} A_n$. Dann gilt also $\lim\limits_{k\to\infty} \rho\left(f_{n_k}, f\right) = 0$, d.h. die Cauchy-Folge $\{f_n\}_{n\in\mathbb{N}}$ enthält die konvergente Teilfolge $\left\{f_{k_n}\right\}_{n\in\mathbb{N}}$. Damit ist aber $\{f_n\}_{n\in\mathbb{N}}$ selbst konvergent gegen f, denn zu jedem $\varepsilon > 0$ existieren $N, M \in \mathbb{N}$ mit $\rho\left(f_n, f_{n_k}\right) < \varepsilon$ für $n, n_k \geq N$ und $\rho\left(f, f_{n_k}\right) < \varepsilon$ für $n_k \geq M$, so dass man für $n \geq N$ und $n_k \geq \max\{N, M\}$ erhält: $\rho(f_n, f) \leq \rho\left(f_n, f_{n_k}\right) + \rho\left(f, f_{n_k}\right) < 2\varepsilon$. Also ist $\mathcal{X}$ vollständig. ■

Definition 11 (nirgends dicht)**.** Eine Teilmenge A eines metrischen Raumes $(\mathcal{X}, \rho)$ heißt *nirgends dicht in* $\mathcal{X}$, falls $\left(\bar{A}\right)^{\circ} = \varnothing$ gilt.

Bemerkung: Äquivalent hierzu ist:

a) $\overline{\left(\overline{A}\right)^c} = \mathcal{X}$

b) Jede abgeschlossene Kugel enthält eine weitere abgeschlossene Kugel, die frei von Punkten von A ist.

Ein typisches Beispiel hierfür ist $A = \mathbb{N}$ im metrischen Raum $(\mathcal{X}, \rho)$ aus Beispiel 1, Teil a).

Definition 12 (von erster und zweiter Kategorie)**.** Eine Teilmenge A eines metrischen Raumes $(\mathcal{X}, \rho)$ heißt *von der ersten Kategorie* (oder *mager*), wenn A darstellbar ist als abzählbar unendliche Vereinigung nirgends dichter Mengen. Eine Menge, die nicht von erster Kategorie ist, ***heißt*** *von zweiter Kategorie*.

Satz 6 (Baire'scher Kategoriensatz)**.** Jeder vollständige metrische Raum $(\mathcal{X}, \rho)$ ist von zweiter Kategorie.

Beweis: Wir führen einen Widerspruchsbeweis, indem wir annehmen, dass $(\mathcal{X}, \rho)$ von erster Kategorie ist, d.h. es gibt eine Darstellung

$$\mathcal{X} = \bigcup_{i=1}^{\infty} A_i \text{ mit nirgends dichten Mengen } A_i \subseteq \mathcal{X}.$$

Wir konstruieren rekursiv eine Folge $\{f_n\}_{n\in\mathbb{N}}$, deren Grenzwert f nicht in $\mathcal{X}$ liegt. Dazu wählen wir zunächst $0 < r_0 < 2$ und $f_0 \in \mathcal{X}$ beliebig. Weil A_1 nirgends dicht ist, gibt es ein $0 < r_1 < 1$ und einen Punkt $f_1 \in \mathcal{X}$ mit

$$\overline{K}_{r_1}(f_1) \subseteq \overline{K}_{r_0}(f_0) \text{ und } \overline{K}_{r_1}(f_1) \cap A_1 = \varnothing.$$

Sind nun $f_1, \cdots, f_n$ für ein $n \in \mathbb{N}$ bereits konstruiert, so gibt es ein $0 < r_{n+1} < \dfrac{1}{n+1}$ und einen Punkt $f_{n+1} \in \mathcal{X}$ mit

$$\overline{K}_{r_{n+1}}(f_{n+1}) \subseteq \overline{K}_{r_n}(f_n) \text{ und } \overline{K}_{r_{n+1}}(f_{n+1}) \cap A_{n+1} = \varnothing.$$

Die Mengenfolge $\left\{\overline{K}_{r_n}(f_n)\right\}_{n\in\mathbb{N}}$ ist nun offensichtlich monoton fallend mit

$$0 < \delta\left(\overline{K}_{r_n}(f_n)\right) < 2r_n < \frac{2}{n} \text{ für alle } n \in \mathbb{N},$$

so dass sich aus dem Cantor'schen Durchschnittssatz 5 die Existenz eines Punktes $f \in \bigcap_{n=1}^{\infty} \overline{K}_{r_n}(f_n)$ ergibt mit $A_k \cap \bigcap_{n=1}^{\infty} \overline{K}_{r_n}(f_n) \subseteq A_k \cap \overline{K}_{r_k}(f_k) = \varnothing$ für alle $k \in \mathbb{N}$. Damit ist aber $f \notin A_k$ für alle $k \in \mathbb{N}$ und somit auch $f \notin \bigcup_{k=1}^{\infty} A_k = \mathcal{X}$: Widerspruch! Also ist $(\mathcal{X}, \rho)$ von zweiter Kategorie. ■

Definition 13 (Kompaktheit)**.** Es sei $(\mathcal{X}, \rho)$ ein metrischer Raum und $A \subseteq \mathcal{X}$.

a) $(\mathcal{X}, \rho)$ heißt *abzählbar kompakt*, falls jede unendliche Teilmenge von $\mathcal{X}$ mindestens einen Häufungspunkt in $\mathcal{X}$ besitzt. A heißt abzählbar kompakt, falls jede unendliche Teilmenge von A mindestens einen Häufungspunkt in A besitzt.

b) A heißt *folgenkompakt*, wenn jede Folge in A eine in A konvergente Teilfolge besitzt.

c) A heißt *(Heine-Borel-)kompakt*, falls jede offene Überdeckung von A eine endliche Teilüberdeckung besitzt.

d) A heißt *totalbeschränkt* (oder auch *präkompakt*), falls zu jedem $\varepsilon > 0$ eine endliche Teilmenge $F_\varepsilon \subseteq A$ existiert mit $A \subseteq \bigcup \{K_\varepsilon(f) \mid f \in F_\varepsilon\}$.

Die folgenden Sätze, die wir hier nur ohne Beweis vorstellen (vgl. u.a. BACHMAN / NARICI (2000), Chapter 5, DUGUNDJI (1966), Chapters IX und XI oder ROYDEN (1988), Chapter 9), gibt eine Übersicht über die verschiedenen Beziehungen zwischen diesen Kompaktheitsbegriffen in metrischen Räumen.

Satz 7 (Zusammenhang zwischen Kompaktheitsbegriffen I)**.** Es sei $(\mathcal{X}, \rho)$ ein metrischer Raum und $A \subseteq \mathcal{X}.$ Dann gelten folgende Äquivalenzen:

a) $(\mathcal{X}, \rho)$ ist abzählbar kompakt

b) $(\mathcal{X}, \rho)$ ist folgenkompakt

c) $(\mathcal{X}, \rho)$ ist kompakt

d) $(\mathcal{X}, \rho)$ ist vollständig und totalbeschränkt.

Die Äquivalenz der Aussagen a) bis c) gilt sinngemäß auch für die Teilmenge A. Ist $(\mathcal{X}, \rho)$ zusätzlich vollständig, so sind in diesem Fall a) bis c) auch äquivalent zu

e) A ist abgeschlossen und totalbeschränkt.

Bemerkung: Eine totalbeschränke Menge ist stets auch beschränkt; im metrischen Raum $(\mathcal{X}, \rho)$ aus Beispiel 1, Teil b) ist jede beschränkte Menge umgekehrt auch totalbeschränkt.

Satz 8 (Zusammenhang zwischen Kompaktheitsbegriffen II)**.** Es seien $(\mathcal{X}, \rho_\mathcal{X})$ und $(\mathcal{Y}, \rho_\mathcal{Y})$ metrische Räume und $A \subseteq \mathcal{X}.$ Dann gilt:

a) Ist A kompakt, so auch abgeschlossen und beschränkt, im Allgemeinen jedoch nicht umgekehrt.

b) Ist A kompakt, so auch jede abgeschlossene Teilmenge von A.

c) Ist $T: \mathcal{X} \to \mathcal{Y}$ stetig und A kompakt (in $\mathcal{X}$), so ist $T(A)$ kompakt (in $\mathcal{Y}$).

d) Ist in der Situation c) speziell $(\mathcal{Y}, \rho_\mathcal{Y})$ der metrische Raum aus Beispiel 1, Teil a), so nimmt T auf A Maximum und Minimum an, wenn A kompakt ist.

Definition 14 (Separabilität)**.** Ein metrischer Raum $(\mathcal{X},\rho)$ heißt *separabel*, wenn $\mathcal{X}$ eine abzählbare Teilmenge A enthält, die dicht in $\mathcal{X}$ ist, d.h. für die gilt: $\overline{A}=\mathcal{X}$.

Lemma 2. Jeder kompakte metrische Raum $(\mathcal{X},\rho)$ ist separabel.

Beweis: Als kompakter metrischer Raum ist nach Satz 7 $(\mathcal{X},\rho)$ insbesondere totalbeschränkt. Zu jedem $n\in\mathbb{N}$ existiert also eine endliche Teilmenge $F_n\subseteq\mathcal{X}$ mit $\mathcal{X}=\bigcup\{K_{1/n}(f)\mid f\in F_n\}$. Dann ist $A:=\bigcup_{n=1}^{\infty}F_n$ eine abzählbare Menge, die in $\mathcal{X}$ dicht ist. Sei dazu $f\in A^c$ beliebig. f ist dann ein Häufungspunkt von *A*, weil zu jedem $n\in\mathbb{N}$ nach obigem ein $f_n\in F_n$ existiert mit $\rho(f_n,f)<\dfrac{1}{n}$. Also gilt $\overline{A}=\mathcal{X}$. ■

Beispiel 5. Die metrischen Räume aus Beispiel 1 sind sämtlich separabel. Für die Teile a) bis c) ist nämlich $\mathbb{Q}$ bzw. $\mathbb{Q}^d$ eine abzählbare dichte Teilmenge. Für die Teile d) und e) kann man die Menge $A=\left\{f\in\mathbb{Q}^{\mathbb{N}}\mid f_i=0 \text{ für alle bis auf endlich viele } i\right\}$ wählen. Für die Teile f) und g) kann man als Teilmenge A die Menge aller Polynome auf $[a,b]$ mit rationalen Koeffizienten wählen. Dagegen ist i. Allg. der metrische Raum aus Beispiel 2, Teil a) nicht separabel. Dies trifft beispielsweise schon auf den Fall $(\Omega,\mathcal{A})=(\mathbb{R}^1,\mathcal{B}^1)$ zu. Zum Nachweis der Nicht-Separabilität reicht es nämlich, eine überabzählbare Teilmenge A von $\mathcal{X}$ anzugeben mit der Eigenschaft

$$\rho(f,g)\geq\varepsilon>0 \text{ für alle } f,g\in A \text{ mit } f\neq g \text{ und ein geeignetes } \varepsilon>0.$$

Hier leistet beispielsweise die Menge $A:=\{\varepsilon_x\mid x\in\mathbb{R}\}$ aller Einpunkt-Verteilungen das Gewünschte: es ist nämlich

$$\rho(\varepsilon_x,\varepsilon_y)=\sup\left\{\left|\varepsilon_x(B)-\varepsilon_y(B)\right|\mid B\in\mathcal{B}^1\right\}\geq\left|\varepsilon_x(\{x\})-\varepsilon_y(\{x\})\right|=1 \text{ für } x\neq y.$$

Für abzählbare Grundmengen Ω ist – mit $\mathcal{A}=\mathfrak{P}(\Omega)$ – dagegen $(\mathcal{X},\rho)$ stets separabel, weil in diesem Fall $(\mathcal{X},\rho)$ isometrisch ist zu einem Teilraum des metrischen Raums aus Beispiel 1, Teil d) aufgrund der Beziehung (vgl. auch die Bemerkung zu dem späteren Lemma 6)

$$\rho(P,Q)=\frac{1}{2}\sum_{\omega\in\Omega}\left|P(\{\omega\})-Q(\{\omega\})\right| \text{ für alle } P,Q\in\mathcal{X}.$$

I. 1.3. Lineare metrische Räume

In diesem Abschnitt betrachten wir speziellere metrische Räume, nämlich solche, deren Punktmenge $\mathcal{X}$ einen Vektorraum über einem der Körper $K=\mathbb{R}$ oder $K=\mathbb{C}$ bildet. Dazu sind an die Metrik ρ weitere Anforderungen zu stellen, damit eine sinnvolle Beziehung zur linearen Struktur von $\mathcal{X}$ erreicht wird. Mit **0** werde im Folgenden der Nullvektor des jeweiligen Vektorraums bezeichnet, mit 0 wie üblich die Null im jeweiligen Körper .

Definition 15 (Linearer metrischer Raum)**.** Ein metrischer Raum $(\mathcal{X},\rho)$ heißt *linearer metrischer Raum*, wenn folgende Bedingungen erfüllt sind:

a) ist ein Vektorraum über dem Körper K.
b) Es gilt $\rho(f,g)=\rho(f-g,\mathbf{0})$ für alle $f,g\in\mathcal{X}$.
c) Ist $\{f_n\}_{n\in\mathbb{N}}$ eine Folge in $\mathcal{X}$ mit Grenzwert $f\in\mathcal{X}$ und $\{\alpha_n\}_{n\in\mathbb{N}}$ eine Folge in K mit Grenzwert $\alpha\in K,$ so ist $\alpha f\in\mathcal{X}$ Grenzwert der Folge $\{\alpha_n f_n\}_{n\in\mathbb{N}}$ in $\mathcal{X}$.

Lemma 3. Es sei $(\mathcal{X},\rho)$ ein linearer metrischer Raum. Die Bedingung c) aus Definition 15 ist dann äquivalent zu folgenden drei Aussagen:

Für jede Folge $\{f_n\}_{n\in\mathbb{N}}$ in $\mathcal{X}$ und jede Folge $\{\alpha_n\}_{n\in\mathbb{N}}$ in K gilt:

a) $\lim\limits_{n\to\infty} f_n=\mathbf{0}\ \Rightarrow\ \lim\limits_{n\to\infty}\alpha f_n=\mathbf{0}$ für alle $\alpha\in K$
b) $\lim\limits_{n\to\infty}\alpha_n=0\ \Rightarrow\ \lim\limits_{n\to\infty}\alpha_n f=\mathbf{0}$ für alle $f\in\mathcal{X}$
c) $\lim\limits_{n\to\infty} f_n=\mathbf{0}$ und $\lim\limits_{n\to\infty}\alpha_n=0\Rightarrow\ \lim\limits_{n\to\infty}\alpha_n f_n=\mathbf{0}.$

Beweis: "$\Rightarrow$": Wähle für Teil a) die Folge mit $\alpha_n=\alpha$ für alle $n\in\mathbb{N}$ mit $f=\mathbf{0}$ und für Teil b) die Folge mit $f_n=f$ für alle $n\in\mathbb{N}$ mit $\alpha=0.$ Teil c) ist der Spezialfall mit $f=\mathbf{0}$ und $\alpha=0.$

"$\Leftarrow$": Ist $\{f_n\}_{n\in\mathbb{N}}$ eine Folge in $\mathcal{X}$ mit Grenzwert $f\in\mathcal{X}$ und $\{\alpha_n\}_{n\in\mathbb{N}}$ eine Folge in K mit Grenzwert $\alpha\in K,$ so gilt:

$$\alpha_n f_n-\alpha f=\underbrace{(\alpha_n-\alpha)(f_n-f)}_{\to 0\text{ nach Teil c)}}+\underbrace{(\alpha_n-\alpha)f}_{\to 0\text{ nach Teil b)}}+\underbrace{\alpha(f_n-f)}_{\to 0\text{ nach Teil a)}}\to 0\ \text{ mit } n\to\infty. \qquad\blacksquare$$

Lemma 4. In einem linearen metrischen Raum gelten folgende Eigenschaften:

a) $\rho(f+h,g+h)=\rho(f,g)=\rho(-f,-g)$ für alle $f,g,h\in\mathcal{X}$
b) Ist $\{f_n\}_{n\in\mathbb{N}}$ eine Folge in $\mathcal{X}$ mit Grenzwert $f\in\mathcal{X}$ und $\{g_n\}_{n\in\mathbb{N}}$ eine Folge in $\mathcal{X}$ mit Grenzwert $g\in\mathcal{X},$ so ist $f+g\in\mathcal{X}$ Grenzwert der Folge $\{f_n+g_n\}_{n\in\mathbb{N}}$ in $\mathcal{X}$.

Beweis:

a) Es ist $\rho(f+h,g+h)=\rho((f+h)-(g+h),\mathbf{0})=\rho(f-g,\mathbf{0})=\rho(f,g)$ für alle $f,g,h\in\mathcal{X}.$ Wählt man speziell $h=-f-g,$ so ergibt sich aufgrund der Symmetrie hieraus noch $\rho(f+h,g+h)=\rho(-g,-f)=\rho(-f,-g)$ für alle $f,g\in\mathcal{X}.$

b) Mit der Dreiecksungleichung und Teil a) oben gilt:

$$\rho(f_n+g_n,f+g)\le\rho(f_n+g_n,f_n+g)+\rho(f_n+g,f+g)=\rho(g_n,g)+\rho(f_n,f),$$

woraus die Behauptung folgt. $\blacksquare$

Bemerkung: In einem linearen metrischen Raum $(\mathcal{X},\rho)$ ist die Operation (Abbildung) $T: K\times\mathcal{X}\to\mathcal{X}:(\alpha, f)\mapsto \alpha f$ der skalaren Multiplikation stetig auf dem Produktraum $K\times\mathcal{X}$ bezüglich der Metrik

$$\rho_{K\times\mathcal{X}}\big((\alpha, f),(\beta,g)\big):=\rho_K(\alpha,\beta)+\rho(f,g) \text{ für alle } \alpha,\beta\in K,\ \ f,g\in\mathcal{X},$$

wobei ρ_K die euklidische Metrik aus Beispiel 1, Teil a) bezeichne, d.h. es gilt

$$\rho_K(\alpha,\beta)=|\alpha-\beta| \text{ für alle } \alpha,\beta\in K.$$

Es gilt dann nämlich: ist $\{(\alpha_n, f_n)\}_{n\in\mathbb{N}}$ eine Folge in $K\times\mathcal{X}$ mit $\rho_{K\times\mathcal{X}}$-Grenzwert (α, f) in $K\times\mathcal{X}$, so gilt $\alpha_n\to\alpha$ in K und $f_n\to f$ in $\mathcal{X}$, d.h. auch $T(\alpha_n, f_n)=\alpha_n f_n\to\alpha f=T(\alpha, f)$ in $K\times\mathcal{X}$, nach Definition 15 c).

Entsprechend zeigt man, dass die Operation (Abbildung) $S:\mathcal{X}\times\mathcal{X}\to\mathcal{X}:(f,g)\mapsto f+g$ der (vektoriellen) Addition stetig auf dem Produktraum $\mathcal{X}\times\mathcal{X}$ ist bezüglich der Metrik

$$\rho_{\mathcal{X}\times\mathcal{X}}\big((f_1,g_1),(f_2,g_2)\big):=\rho(f_1,f_2)+\rho(g_1,g_2) \text{ für alle } f_1,f_2,g_1,g_2\in\mathcal{X}.$$

In der Funktionalanalysis spielen im Gegensatz zur "klassischen" Linearen Algebra vor allem unendlich-dimensionale Räume eine wichtige Rolle. Deshalb sind Begriffe wie "lineare Unabhängigkeit" und "Basis" hier etwas allgemeiner zu fassen.

Definition 16 (lineare Unabhängigkeit; Basis). In einem linearen metrischen Raum $(\mathcal{X},\rho)$ heißen endlich viele Elemente $f_1,\cdots,f_n\in\mathcal{X}$ *linear unabhängig*, wenn gilt:

$$\sum_{i=1}^{n}\alpha_i f_i=\mathbf{0}\ \Rightarrow\ \alpha_i=0 \text{ für alle } i\in\{1,\cdots,n\}.$$

Anderenfalls heißen $f_1,\cdots,f_n$ linear abhängig. Eine Teilmenge $A\subseteq\mathcal{X}$ heißt *linear unabhängig*, wenn jede endliche Teilmenge von A linear unabhängig ist. $(\mathcal{X},\rho)$ heißt *n-dimensional* (mit $n\in\mathbb{N}$), wenn $\mathcal{X}$ eine linear unabhängige Teilmenge $A\subseteq\mathcal{X}$ aus n Elementen enthält und jede $n+1$-elementige Teilmenge von $\mathcal{X}$ linear abhängig ist. $(\mathcal{X},\rho)$ heißt *unendlich-dimensional*, wenn $\mathcal{X}$ nicht endlich-dimensional ist. Eine Teilmenge $A\subseteq\mathcal{X}$ heißt *Hamel-Basis* von $\mathcal{X}$, wenn A linear unabhängig ist und $\mathcal{X}$ aufspannt, d.h. wenn gilt

$$\operatorname{span}(A):=\left\{\sum_{i=1}^{n}\alpha_i f_i\ \middle|\ \alpha_1,\cdots,\alpha_n\in K,\ f_1,\cdots,f_n\in A;\ \ n\in\mathbb{N}\right\}=\mathcal{X}.$$

Man sagt ferner, $\mathcal{X}$ besitze eine *abzählbare Basis* $A=\{f_n\,|\,n\in\mathbb{N}\}$, wenn es zu jedem $f\in\mathcal{X}$ eine Familie $\{\alpha_n\,|\,n\in\mathbb{N}\}\subseteq K$ gibt derart, dass

$$f=\sum_{n=1}^{\infty}\alpha_n f_n=\rho\text{-}\lim_{n\to\infty}\sum_{k=1}^{n}\alpha_k f_k$$

gilt. Eine solche Basis heißt *Schauder-Basis*, wenn die Koeffizienten $\{\alpha_n \mid n \in \mathbb{N}\} \subseteq K$ für die Darstellung von f stets eindeutig bestimmt sind. Eine Teilmenge $A \subseteq \mathcal{X}$ heißt ferner *fundamental in* $\mathcal{X}$, wenn gilt:

$$\overline{\operatorname{span}(A)} = \mathcal{X}.$$

Bemerkung: Die "Konstruktion" einer Hamel-Basis ist oft nur mit dem *Auswahlaxiom* (vgl. das Skript STOCHASTIK, Abschnitt I.1.) oder äquivalenten Axiomen, z.B. dem *Zorn'schen Lemma*, möglich. Hamel-Basen entsprechen nämlich auch *maximal linear-unabhängigen* Teilmengen A von $\mathcal{X}$, also solchen linear-unabhängigen Teilmengen A, für die gilt: ist $B \subseteq \mathcal{X}$ mit $B \supseteq A$ linear unabhängig, so ist $B = A$. Es gibt stets linear unabhängige Teilmengen $A \subseteq \mathcal{X}$; z.B. sind alle einelementigen Teilmengen $A = \{f\}$ mit $\mathbf{0} \neq f \in \mathcal{X}$ linear unabhängig. Da die Potenzmenge $\mathfrak{P}(\mathcal{X})$ durch die Enthaltenseins-Relation "$\subseteq$" geordnet ist und jede Kette in $(\mathfrak{P}(\mathcal{X}), \subseteq)$ die obere Schranke $\mathcal{X} \in \mathfrak{P}(\mathcal{X})$ besitzt, lässt sich also nach dem Zorn'schen Lemma jede linear unabhängige Teilmenge zu einer Hamel-Basis als maximales Element in $(\mathfrak{P}(\mathcal{X}), \subseteq)$ vervollständigen (siehe auch TAYLOR (1963), Abschnitt 1.72).

Beipielsweise kann $\mathbb{R}$ als Vektorraum über $\mathbb{Q}$ aufgefasst werden. Dann ist etwa die Menge $A = \{1, \sqrt{2}, \ln(5), \pi\}$ linear unabhängig und kann mit dem Zorn'schen Lemma zu einer Hamel-Basis vervollständigt werden.

Beispiel 6. Die Räume in Beispiel 1, Teile a) bis c) sind ein- bzw. *d*-dimensional. Die Räume in Beispiel 1, Teile d) bis g) sind unendlich-dimensional. Die Räume in Beispiel 1, Teile d) und e) besitzen Schauder-Basen; sie sind gegeben durch

$$A = \{f_n \mid n \in \mathbb{N}\} \text{ mit } f_{nk} = \begin{cases} 1, & k = n \\ 0, & k \neq n \end{cases} \text{ für alle } n, k \in \mathbb{N}.$$

Für $g = \{g_n\}_{n \in \mathbb{N}} \in \mathcal{X}$ sind die entsprechenden Koeffizienten nämlich genau durch $\alpha_n = g_n$ für alle $n \in \mathbb{N}$ gegeben. Die Räume in Beispiel 1, Teile f) und g) besitzen die *Monome* als Fundamentalsystem, d.h. die Abbildungen der Form $x \mapsto x^n$, $x \in [a, b]$ mit $n \in \mathbb{Z}^+ = \mathbb{N} \cup \{0\}$.

Bemerkung: Die Räume in Beispiel 2 sind keine *linearen* metrischen Räume, weil die Menge der Wahrscheinlichkeitsmaße auf einer σ-Algebra keinen Vektorraum bilden. Man kann aber z.B. den dort in Teil a) angegebenen metrischen Raum zu einem linearen metrischen Raum ergänzen. Dazu benötigt man allerdings den Begriff des signierten Maßes, das in der folgenden Definition behandelt wird.

Definition 17 (signiertes Maß)**.** Es sei $(\Omega, \mathcal{A})$ ein Messraum. Unter einem *signierten Maß* versteht man eine Abbildung $\nu : \mathcal{A} \to \overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, \infty\}$ mit den folgenden Eigenschaften:

a) ν nimmt höchstens einen der Werte $-\infty, \infty$ an.
b) $\nu(\varnothing) = 0$ (Normiertheit).
c) $\nu\left(\bigoplus_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \nu(A_n)$ für jede Folge $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathcal{A}$ paarweise disjunkter messbarer Mengen, wobei die Reihe auf der rechten Seite entweder absolut (gegen einen endlichen Wert) konvergiert oder bestimmt (gegen $-\infty$ oder ∞) divergiert (σ-Additivität).

Das Tripel $(\Omega, \mathcal{A}, \nu)$ heißt auch *signierter Maßraum*.

Bemerkung: Der Begriff des signierten Maßes umfasst den des üblichen Maßbegriffs, d.h. jedes signierte Maß mit der Eigenschaft

$$\nu(A) \geq 0 \text{ für alle } A \in \mathcal{A}$$

ist ein Maß. Sind ferner μ_1 und μ_2 Maße im üblichen Sinne, die nicht zugleich den Wert ∞ annehmen, so ist $\nu := \mu_1 - \mu_2$ ein signiertes Maß. Mit ν ist zugleich auch $\alpha\nu$ für jedes $\alpha \in \mathbb{R}$ ein signiertes Maß, insbesondere also auch $-\nu$.

Definition 18 (positive Menge)**.** In einem signierten Maßraum $(\Omega, \mathcal{A}, \nu)$ heißt eine Menge $A \in \mathcal{A}$ *positiv* (bezüglich ν), wenn gilt:

$$\nu(B) \geq 0 \text{ für alle } B \in \mathcal{A},\ B \subseteq A.$$

Entsprechend heißt eine Menge $A \in \mathcal{A}$ *negativ* (bezüglich ν), wenn gilt:

$$\nu(B) \leq 0 \text{ für alle } B \in \mathcal{A},\ B \subseteq A.$$

Eine sowohl positive wie negative Menge $A \in \mathcal{A}$ (für die dann notwendig $\nu(A) = 0$ gilt) heißt (ν-)*Nullmenge*.

Lemma 5. In einem signierten Maßraum $(\Omega, \mathcal{A}, \nu)$ gilt:

a) Jede Teilmenge einer positiven (negativen, (ν-)Null-)Menge ist selbst auch positiv (negativ, (ν-)Nullmenge).
b) Ist $A \in \mathcal{A}$ positiv, so ist die Abbildung $\mu : \mathcal{A} \to \mathbb{R}^+ \cup \{\infty\} : B \mapsto \nu(A \cap B)$ ein Maß.
c) Die abzählbare Vereinigung von positiven (negativen, (ν-)Null-)Mengen ist selbst auch positiv (negativ, (ν-)Nullmenge).

Beweis:

a) Folgt unmittelbar aus der Definition positiver (negativer, (ν-) Null-)Mengen.

b) Offensichtlich ist mit ν auch μ σ-additiv. Die Nicht-Negativität von μ folgt aus Teil a) des Lemmas.

c) Sei zunächst $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathcal{A}$ eine Folge positiver messbarer Mengen und $A := \bigcup_{n=1}^{\infty} A_n$. Für jede messbare Teilmenge $B \subseteq A$ definiere $B_n := B \cap A_n \setminus \bigcup_{i=1}^{n-1} A_i$, $n \in \mathbb{N}$. Dann ist jedes B_n messbare Teilmenge von A_n und somit positiv. Wegen $B = \bigoplus_{n=1}^{\infty} B_n$ ist also $\nu(B) = \sum_{n=1}^{\infty} \nu(B_n) \geq 0$, womit die Behauptung für positive Mengen gezeigt ist. Für negative Mengen argumentiert man völlig analog. Ist $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathcal{A}$ eine Folge von Nullmengen, so ist jedes A_n sowohl positiv wie negativ, d.h. $A := \bigcup_{n=1}^{\infty} A_n$ ist nach vorigem auch positiv wie negativ und damit Nullmenge, was zu zeigen war. ■

Bemerkung: Der Begriff der "Nullmenge" darf nicht mit dem Begriff einer "Menge vom Maß Null" verwechselt werden. Wählen wir beispielsweise $\Omega = \{1,2,3\}$ und $\mathcal{A} := \mathfrak{P}(\Omega)$, so wird durch $\nu(\{\omega\}) := 2 - \omega$, $\omega \in \Omega$ ein signiertes Maß definiert vermöge $\nu(A) := \sum_{\omega \in A} \nu(\{\omega\})$, $A \in \mathfrak{P}(\Omega)$. Für die Menge $B = \{1,3\}$ gilt dann zwar $\nu(B) = 0$, aber es ist $\nu(\{1\}) = -1$, $\nu(\{3\}) = 1$ und damit B eine Menge vom Maß Null, aber keine Nullmenge im Sinne der Definition.

Satz 9 (positive Teilmengen)**.** Es sei $(\Omega, \mathcal{A}, \nu)$ ein signierter Maßraum und $A \in \mathcal{A}$ eine Menge mit $0 < \nu(A) < \infty$. Dann existiert eine positive Menge $B \in \mathcal{A}$ mit $B \subseteq A$ und $\nu(B) \geq \nu(A) > 0$.

Beweis: Wir führen den Beweis konstruktiv. Dazu definieren wir

$$\mathcal{S}(B,\varepsilon) := \{C \in \mathcal{A} \mid C \subseteq B \wedge \nu(C) < -\varepsilon\} \text{ für alle } B \in \mathcal{A} \text{ und } \varepsilon > 0.$$ [8]

Wir konstruieren nun rekursiv natürliche Zahlen $k_1, k_2, \cdots$ sowie paarweise disjunkte Mengen $A_1, A_2, \cdots \subseteq \mathcal{A}$ mit jeweils negativem Maß wie folgt:

Wenn die Menge A schon positiv ist, sind wir mit dem Beweis fertig (wähle $B = A$). Anderenfalls gibt es ein $\varepsilon > 0$ mit $\mathcal{S}(A,\varepsilon) \neq \varnothing$. Wähle dann

$$k_1 := \min\left\{k \in \mathbb{N} \,\middle|\, \mathcal{S}\left(A, \frac{1}{k}\right) \neq \varnothing\right\} \text{ und eine Menge } A_1 \subseteq A \text{ mit } \nu(A_1) < -\frac{1}{k_1}.$$

Wenn nun die Menge $A \setminus A_1$ positiv ist, so sind wir mit dem Beweis fertig, denn wir können dann $B := A \setminus A_1$ wählen. B hat nämlich auch positives Maß, weil $A = A_1 \oplus B$ ist und somit

[8] Falls keine Mengen $C \subseteq B$ mit der Eigenschaft $\nu(C) < -\varepsilon$ existieren, ist natürlich $\mathcal{S}(B,\varepsilon) = \varnothing$.

$$0 < \nu(A) = \nu(B) + \underbrace{\nu\left(A_1\right)}_{<0} < \nu(B)$$

ist. Anderenfalls wird der Prozess folgendermaßen fortgesetzt: sind für ein $n \in \mathbb{N}$ die Zahlen $k_1, \cdots, k_n$ und die – paarweise disjunkten – Mengen $A_1, \cdots, A_n \subseteq A$ mit $v\left(A_i\right) < -\frac{1}{k_i}$ für $i = 1, \cdots, n$ bereits konstruiert, so ist entweder schon $A \setminus \bigoplus_{i=1}^{n} A$ positiv, und wir sind mit dem Beweis fertig, da $B := A \setminus \bigoplus_{i=1}^{n} A$ gewählt werden kann, weil *B* aufgrund von $A = B \oplus \bigoplus_{i=1}^{n} A_i$ und daraus folgend

$$0 < \nu(A) = \underbrace{\nu(B) + \sum_{i=1}^{n} \underbrace{\nu\left(A_i\right)}_{<0}}_{<\infty} < \nu(B) - \sum_{i=1}^{n} \frac{1}{k_i} < \nu(B)$$

positives Maß hat. Anderenfalls gibt es wieder ein $\varepsilon > 0$ mit $\mathcal{S}\left(A \setminus \bigoplus_{i=1}^{n} A_i, \varepsilon\right) \neq \varnothing$. Wähle dann

$$k_{n+1} := \min\left\{k \in \mathbb{N} \,\middle|\, \mathcal{S}\left(A \setminus \bigoplus_{i=1}^{n} A_i, \frac{1}{k}\right) \neq \varnothing\right\} \text{ und eine Menge } A_{n+1} \subseteq A \setminus \bigoplus_{i=1}^{n} A_i \text{ mit } v\left(A_{n+1}\right) < -\frac{1}{k_{n+1}}.$$

Wenn das Verfahren nach endlich vielen Schritten endet, sind wir mit dem Beweis fertig. Anderenfalls setzen wir $B := A \setminus \bigoplus_{i=1}^{\infty} A_i$, woraus $A = B \oplus \bigoplus_{i=1}^{\infty} A_i$ und somit

$$0 < \nu(A) = \underbrace{\nu(B) + \sum_{i=1}^{\infty} \underbrace{\nu\left(A_i\right)}_{<0}}_{<\infty} < \nu(B) - \sum_{i=1}^{\infty} \frac{1}{k_i} < \nu(B)$$

folgt, also $\sum_{i=1}^{\infty} \frac{1}{k_i} < \infty$ ist und demnach $k_n \to \infty$ strebt für $n \to \infty$. Jedenfalls ist $\nu(B) > \nu(A) > 0$. Sei nun $\varepsilon > 0$ gegeben. Wegen der Divergenz von $\{k_n\}_{n \in \mathbb{N}}$ können wir *n* so groß wählen, dass $\frac{1}{k_n - 1} < \varepsilon$ wird. Sei ferner $C \subseteq B$ eine beliebige messbare Teilmenge von *B*. Nach Wahl von k_n ist das Mengensystem $\mathcal{S}\left(A \setminus \bigoplus_{i=1}^{n} A_i, \frac{1}{k_n - 1}\right)$ leer, d.h. wegen $C \subseteq B = A \setminus \bigoplus_{i=1}^{\infty} A_i \subseteq A \setminus \bigoplus_{i=1}^{n} A_i$ muss $\nu(C) \geq -\frac{1}{k_n - 1} > -\varepsilon$ sein. Da $\varepsilon > 0$ beliebig war, folgt also $\nu(C) \geq 0$, d.h. *B* enthält nur messbare Teilmengen mit nicht-negativem Maß. Damit ist aber *B* eine positive Teilmenge positiven Maßes von *A* (mit $\nu(B) > \nu(A)$), und der Satz somit bewiesen. ■

Satz 10 (Hahn'scher Zerlegungssatz)**.** In einem signierten Maßraum $(\Omega, \mathcal{A}, \nu)$ existieren stets eine positive Menge *A* und eine negative Menge *B* mit $A \cap B = \varnothing$ und $A \oplus B = \Omega$. Für diese Mengen gilt ferner:

$$\nu(A) = \sup_{C \in \mathcal{A}} \nu(C) \quad \text{sowie} \quad \nu(B) = \inf_{D \in \mathcal{A}} \nu(D).$$

Beweis: O.B.d.A. können wir annehmen, dass $\nu < \infty$ ist (sonst Übergang zum signierten Maß $-\nu$). Sei $\mathcal{P}_\nu := \{C \in \mathcal{A} \mid C \text{ ist positiv}\}$ und $\lambda := \sup\{\nu(C) \mid C \in \mathcal{P}_\nu\} \geq 0$ (wegen $\varnothing \in \mathcal{P}_\nu$). Es existiert also eine Folge positiver Mengen $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathcal{P}_\nu$ mit $\lambda = \lim_{n\to\infty} \nu(A_n)$. Dann ist $A := \bigcup_{n\in\mathbb{N}} A_n$ eine mögliche Wahl für die gesuchte positive Menge, denn nach Lemma 5, Teil c) ist erstens A positiv und somit $\lambda \geq \nu(A)$. Die Mengen $C_n := A \setminus A_n$, $n \in \mathbb{N}$ erfüllen damit als Teilmengen der positiven Menge A die Beziehung $\nu(C_n) \geq 0$ mit $\nu(A) = \nu(A_n) + \nu(C_n) \geq \nu(A_n)$ für alle $n \in \mathbb{N}$, so dass auch $\nu(A) \geq \lim_{n\to\infty} \nu(A_n) = \lambda$ und somit schließlich $\lambda = \nu(A) < \infty$ folgt. Es bleibt zweitens zu zeigen, dass die Menge $B := A^c$ negativ ist. Sei dazu $D \subseteq B$ eine beliebige messbare *positive* Teilmenge von B. Dann ist $D \cap A = \varnothing$ und folglich

$$\lambda \geq \nu(A \cup D) = \nu(A) + \nu(D) = \lambda + \nu(D),$$

also $\nu(D) = 0.$ Also enthält B keine *positiven* Teilmengen positiven Maßes und damit nach Satz 9 auch keine (sonstigen) Teilmengen positiven Maßes, so dass B in der Tat eine negative Menge ist. womit der Satz bewiesen ist.
Zum Nachweis der beiden weiteren angegebenen Beziehungen reicht es, die erste nachzuweisen (die andere folgt dann z.B. durch den Übergang von ν zu $-\nu$). Falls $\nu(A) = 0$ ist, haben alle messbaren Teilmengen von Ω nicht-positives Maß (d.h. Ω ist dann negativ), so dass die Aussage wegen $\nu(\varnothing) = 0$ trivial ist. Anderenfalls existiert bei $\nu < \infty$ nach Satz 9 zu jeder messbaren Menge D mit $\nu(D) > 0$ eine positive Menge E mit $0 < \nu(D) \leq \nu(E) \leq \nu(A) = \sup\{\nu(C) \mid C \in \mathcal{P}_\nu\}$ nach obigem. Wenn es dagegen eine messbare Menge C mit $\nu(C) = \infty$ gibt, ist die Beziehung ebenfalls trivialerweise erfüllt, womit alles gezeigt ist. ∎

Bemerkung: Die Zerlegung im Hahn'schen Satz ist im Allgemeinen nicht eindeutig, weil die ν-Nullmengen als Teilmengen sowohl der Menge A als auch der Menge B zugeordnet werden können. Allerdings besteht Eindeutigkeit, wenn man die Zerlegung in der Sprache der Maße formuliert.

Definition 19 (Singularität von Maßen)**.** Zwei Maße μ_1 und μ_2 in einem Messraum $(\Omega, \mathcal{A})$ heißen (zueinander) *orthogonal* oder *singulär* (in Zeichen: $\mu_1 \perp \mu_2$), wenn eine Zerlegung $\Omega = A \oplus B$ in messbare Mengen existiert mit $\mu_1(A) = \mu_2(B) = 0.$

Satz 11 (Jordan-Zerlegung)**.** In einem signierten Maßraum $(\Omega, \mathcal{A}, \nu)$ existieren stets zwei eindeutig bestimmte singuläre Maße ν^+ und ν^- auf $\mathcal{A}$ mit $\nu = \nu^+ - \nu^-$.

Beweis: Es sei $\Omega = A \oplus B$ eine Hahn-Zerlegung im Sinne von Satz 9. Dann leisten die Maße

$$\nu^+ := \nu(\bullet \cap A) \quad \text{und} \quad \nu^- := -\nu(\bullet \cap B)$$

das Gewünschte, wobei die Eindeutigkeit daraus resultiert, dass die Hahn-Zerlegung bis auf die Zuordnung von ν-Nullmengen eindeutig ist. ∎

Definition 20 (Totalvariation)**.** In der nach Satz 11 in einem signierten Maßraum $(\Omega, \mathcal{A}, \nu)$ bestehende Jordan-Zerlegung heißt ν^+ der *Positivteil* und ν^- der *Negativteil* von ν. Das durch $\|\nu\| := \nu^+ + \nu^-$ auf $\mathcal{A}$ definierte Maß heißt die *Totalvariation* von ν.

Ähnlich wie für Maße lässt sich auch für signierte Maße ν eine Integrationstheorie entwickeln. Es ist nahe liegend, ein solches Integral für messbare Abbildungen f durch

$$\int f\,d\nu = \int f\,d\nu^{+} - \int f\,d\nu^{-}$$

zu definieren, sofern die rechte Seite eindeutig bestimmt ist.

Definition 21 (ν-Integral)**.** Es sei $(\Omega, \mathcal{A}, \nu)$ ein signierter Maßraum und $f:(\Omega,\mathcal{A}) \to (\mathbb{R}^1, \mathcal{B}^1)$ eine messbare Abbildung, mit der Borel'schen σ-Algebra $\mathcal{B}^1$. f heißt *ν-integrierbar,* wenn f integrierbar ist bezüglich der Maße ν^{+} und ν^{-} und nicht zugleich $\int f\,d\nu^{+} = \int f\,d\nu^{-} \in \{-\infty, \infty\}$ gilt. In diesem Fall ist das ν-Integral von f definiert durch

$$\int f\,d\nu := \int f\,d\nu^{+} - \int f\,d\nu^{-}.$$

f heißt *eigentlich ν-integrierbar*, wenn sowohl $\int f\,d\nu^{+}$ als auch $\int f\,d\nu^{-}$ endlich sind. Die Menge der eigentlich ν-integrierbaren Abbildungen wird mit $L^1(\Omega, \mathcal{A}, \nu)$ oder kürzer mit $L^1(\nu)$ bezeichnet.

Lemma 6. Das ν-Integral für ein signiertes Maß ν besitzt folgende Eigenschaften:

a) Eine Abbildung f ist genau dann ν-integrierbar, wenn $|f|$ ν-integrierbar ist.

b) Es gilt $\left|\int f\,d\nu\right| \leq \int |f|\,d\|\nu\|$ für alle ν-integrierbaren Abbildungen f.

c) Es ist $\int \mathbb{1}_A\,d\nu = \nu(A)$ für alle $A \in \mathcal{A}$.

Beweis:

a) Folgt aus der Darstellung

$$\int f\,d\nu = \int f\,d\nu^{+} - \int f\,d\nu^{-} = \int f^{+}\,d\nu^{+} - \int f^{-}\,d\nu^{+} - \int f^{+}\,d\nu^{-} + \int f^{-}\,d\nu^{-}$$

mit $\int f\,d\nu^{+}$ oder $\int f\,d\nu^{-}$ endlich, sowie

$$\int |f|\,d\nu = \int |f|\,d\nu^{+} - \int |f|\,d\nu^{-} = \int f^{+}\,d\nu^{+} + \int f^{-}\,d\nu^{+} - \int f^{+}\,d\nu^{-} - \int f^{-}\,d\nu^{-}.$$

b) Nach der vorigen Beziehung ist

$$\left|\int f\,d\nu\right| \leq \int f^{+}\,d\nu^{+} + \int f^{-}\,d\nu^{+} + \int f^{+}\,d\nu^{-} + \int f^{-}\,d\nu^{-} = \int |f|\,d\|\nu\|.$$

c) Es ist $\int \mathbb{1}_A\,d\nu = \int \mathbb{1}_A\,d\nu^{+} - \int \mathbb{1}_A\,d\nu^{-} = \nu^{+}(A) - \nu^{-}(A) = \nu(A)$ für alle $A \in \mathcal{A}$. ■

Wir kommen nun zu der angekündigten Erweiterung des in Beispiel 2, Teil a) angegebenen metrischen Raums zu einem linearen metrischen Raum.

Lemma 7. Es sei $(\Omega, \mathcal{A})$ ein Messraum und $\mathfrak{N}$ die Menge der *endlichen* signierten Maße auf $\mathcal{A}$. Dann ist $(\mathfrak{N}, \rho_T)$ ein linearer metrischer Raum über dem Körper $\mathbb{R}$ mit der *Totalvariation* als Metrik, gegeben durch

$$\rho_T(\nu, \upsilon) := \sup_{E \in \mathcal{A}} \left|\nu(E) - \upsilon(E)\right| \text{ für alle } \nu, \upsilon \in \mathfrak{N}.$$

Beweis: Aufgrund der Endlichkeitsannahme ist mit $\nu, \upsilon \in \mathfrak{N}$ auch $\alpha\nu + \beta\upsilon \in \mathfrak{N}$ für alle $\alpha, \beta \in \mathbb{R}$, also $\mathfrak{N}$ ein Vektorraum über $\mathbb{R}$. Die Eigenschaften einer Metrik für ρ_T ergeben sich aus unmittelbar aus den Eigenschaften der euklidischen Metrik. ■

Bemerkung: Das Supremum in der Definition der Totalvariation wird auch angenommen. Für das signierte Maß $\tau := \nu - \upsilon$ existieren ja nach Satz 10 eine positive Menge A und eine dazu disjunkte negative Menge B mit $\tau(A) = \sup_{C \in \mathcal{A}} \tau(C) \geq 0$ und $\tau(B) = \inf_{D \in \mathcal{A}} \tau(D) \leq 0,$ also ist

$$\rho_T(\nu, \upsilon) = \sup_{E \in \mathcal{A}} \left|\nu(E) - \upsilon(E)\right| = \sup_{E \in \mathcal{A}} \left|\tau(E)\right| = \max\left\{\tau(A), -\tau(B)\right\}.$$

Im Fall $\nu(\Omega) = \upsilon(\Omega),$ also $\tau(\Omega) = 0$ wie in Beispiel 2, Teil a) reduziert sich dies wegen $B = A^c$ noch auf

$$\rho_T(\nu, \upsilon) = \sup_{E \in \mathcal{A}} \left|\nu(E) - \upsilon(E)\right| = \sup_{E \in \mathcal{A}} \left|\tau(E)\right| = \tau(A).$$

Betrachtet man spezieller den metrischen Raum $\mathcal{X}$ der Wahrscheinlichkeitsmaße P über $\mathcal{A}$, die eine Dichte f_P bezüglich eines (festen) Maßes μ auf $\mathcal{A}$ besitzen, so ergibt sich noch

$$\rho_T(P, Q) = \frac{1}{2} \int \left| f_P - f_Q \right| d\mu \text{ für alle } P, Q \in \mathcal{X}.$$

Eine geeignete positive Menge für das signierte Maß $\tau = P - Q$ ist hier nämlich gegeben durch $A := \left\{ f_P \geq f_Q \right\} \in \mathcal{A}$ mit

$$\rho_T(P, Q) = \tau(A) = \int_A \left(f_P - f_Q \right) d\mu = \int_A \left| f_P - f_Q \right| d\mu = -\tau\left(A^c\right) = -\int_{A^c} \left(f_P - f_Q \right) d\mu = \int_{A^c} \left| f_P - f_Q \right| d\mu,$$

also

$$2\rho_T(P, Q) = \int_A \left| f_P - f_Q \right| d\mu + \int_{A^c} \left| f_P - f_Q \right| d\mu = \int_\Omega \left| f_P - f_Q \right| d\mu,$$

woraus die angegebene Beziehung folgt.

Hieraus ergibt sich u.A. auch die (im diskreten stets gültige) Darstellung

$$\rho(P, Q) = \frac{1}{2} \sum_{\omega \in \Omega} \left| P\left(\{\omega\}\right) - Q\left(\{\omega\}\right) \right|$$

in Beispiel 5, wenn man als μ das abzählende Maß auf $\mathfrak{P}(\Omega)$ wählt.

In der Literatur wird neben der Totalvariation $\rho_T(\nu,\upsilon) = \max\{\tau(A), -\tau(B)\} = \max\{|\tau(A)|, |\tau(B)|\}$ häufig noch die durch

$$\rho_S(\nu,\upsilon) = \tau(A) - \tau(B) = |\tau(A)| + |\tau(B)| = \|\nu - \upsilon\|(\Omega), \quad \nu,\upsilon \in \mathfrak{N}$$

definierte Metrik betrachtet. Wegen

$$\rho_T(\nu,\upsilon) \le \rho_S(\nu,\upsilon) \le 2\rho_T(\nu,\upsilon), \;\; \nu,\upsilon \in \mathfrak{N}$$

sind diese Metriken aber *äquivalent*, d.h. Konvergenz in der einen impliziert Konvergenz in der anderen Metrik und umgekehrt. [Beide Metriken erzeugen dann auch dieselbe *Topologie*, d.h. dasselbe System offener Mengen; vgl. auch die spätere Bemerkung am Ende von Abschnitt I.2.1.]

I. 1.4. Banach- und Hilbert-Räume

In diesem Abschnitt befassen wir uns mit dem fundamentalen Begriff der *Norm*, der für alles Weitere wesentlich sein wird.

Definition 22 (normierter Vektorraum)**.** Ein Vektorraum $\mathcal{X}$ über dem Körper K heißt *normiert*, wenn eine Abbildung $\|\bullet\|: \mathcal{X} \to \mathbb{R}$ (genannt "Norm") existiert mit den Eigenschaften:

a) $\|f\| \ge 0$ für alle $f \in \mathcal{X}$
b) $\|f\| = 0 \;\Leftrightarrow\; f = \mathbf{0}$ für alle $f \in \mathcal{X}$
c) $\|f + g\| \le \|f\| + \|g\|$ für alle $f, g \in \mathcal{X}$
d) $\|\alpha f\| = |\alpha| \cdot \|f\|$ für alle $\alpha \in K$ und $f \in \mathcal{X}$.

Lemma 8. Sei $\mathcal{X}$ ein normierter Vektorraum über dem Körper K. Dann wird durch

$$\rho(f,g) := \|f - g\|, \;\; f, g \in \mathcal{X}$$

$(\mathcal{X}, \rho)$ zu einem linearen metrischen Raum.

Beweis: Nach a) ist trivialerweise $\rho(f,g) \ge 0$ für alle $f, g \in \mathcal{X}$, mit $\rho(f,g) = 0 \;\Leftrightarrow\; f = g$ für alle $f, g \in \mathcal{X}$ nach b). Ferner ist $\rho(f,g) = \|f - g\| = \|(-1)\cdot(g - f)\| = |-1| \cdot \|g - f\| = \rho(g, f)$ für alle $f, g \in \mathcal{X}$ nach d). Die Dreiecksungleichung ergibt sich unmittelbar aus c), d.h. ρ ist eine Metrik. Ferner gilt $\rho(f,g) = \|f - g\| = \|(f - g) - \mathbf{0}\| = \rho(f - g, \mathbf{0})$ für alle $f, g \in \mathcal{X}$. Ist schließlich $\{f_n\}_{n\in\mathbb{N}}$ eine Folge in $\mathcal{X}$ mit Grenzwert $f \in \mathcal{X}$ und $\{\alpha_n\}_{n\in\mathbb{N}}$ eine Folge in K mit Grenzwert $\alpha \in K$, so ist $\alpha f \in \mathcal{X}$ Grenzwert der Folge $\{\alpha_n f_n\}_{n\in\mathbb{N}}$ in $\mathcal{X}$ wegen

$$\rho(\alpha_n f_n, \alpha f) = \|\alpha_n f_n - \alpha f\| = \|\alpha_n (f_n - f) - (\alpha - \alpha_n) f\| \le |\alpha_n| \cdot \underbrace{\|f_n - f\|}_{\to 0} + \underbrace{|\alpha - \alpha_n|}_{\to 0} \cdot \|f\| \to 0$$

mit $n \to \infty$, da die konvergente Folge $\{\alpha_n\}_{n\in\mathbb{N}}$ beschränkt ist. ■

Definition 23 (linearer normierter Raum, Banach-Raum)**.** Ein normierter Vektorraum $\mathcal{X}$ über dem Körper K mit der in Lemma 8 definierten Metrik heißt *linearer normierter Raum* und wird mit $(\mathcal{X}, \|\bullet\|)$ bezeichnet. Ein vollständiger linearer normierter Raum $(\mathcal{X}, \|\bullet\|)$ heißt *Banach-Raum*.

Bemerkung: Definitionsgemäß ist jeder lineare normierte Raum auch ein linearer metrischer Raum, aber nicht notwendig umgekehrt. Dazu ist nämlich eine weitere Bedingung (Homogenitätsbedingung) notwendig.

Lemma 9. Ein linearer metrischer Raum ist genau dann ein linearer normierter Raum, wenn für die Metrik ρ die folgende Homogenitätsbedingung gilt:

$$\rho(\alpha f, \mathbf{0}) = |\alpha| \rho(f, \mathbf{0}) \text{ für alle } f \in \mathcal{X} \text{ und } \alpha \in K.$$

Beweis: Es ist nur zu zeigen, dass unter der Homogenitätsbedingung der lineare metrische Raum $(\mathcal{X}, \rho)$ zu einem linearen normierten Raum $(\mathcal{X}, \|\bullet\|)$ wird. Dies wird aber durch die Definition

$$\|f\| := \rho(f, \mathbf{0}) \text{ für alle } f \in \mathcal{X}$$

geleistet:

a) Trivialerweise ist $\| f \| \geq 0$ für alle $f \in \mathcal{X}$.
b) Es gilt $\| f \| = 0 \Leftrightarrow \rho(f, \mathbf{0}) \Leftrightarrow f = \mathbf{0}$ für alle $f \in \mathcal{X}$.
c) Es gilt $\| f + g \| = \rho(f + g, \mathbf{0}) = \rho(f, -g) \leq \rho(f, \mathbf{0}) + \rho(-g, \mathbf{0}) = \rho(f, \mathbf{0}) + \rho(g, \mathbf{0}) = \| f \| + \| g \|$ für alle $f, g \in \mathcal{X}$.
d) Es ist $\| \alpha f \| = \rho(\alpha f, \mathbf{0}) = |\alpha| \rho(f, \mathbf{0}) = |\alpha| \cdot \| f \|$ für alle $\alpha \in K$ und $f \in \mathcal{X}$. ■

Beispiel 7. Die (vollständigen) linearen metrischen Räume aus Beispiel 1, Teile a) bis f) sind sämtlich Banach-Räume, weil die zugehörigen Metriken alle die Homogenitätsbedingung aus Lemma 9 erfüllen. Der lineare metrische Raum aus Beispiel 1, Teil g) ist dagegen nur normiert (da nicht vollständig nach Beispiel 3). Der lineare metrischer Raum $(\mathfrak{N}, \rho_T)$ aus Lemma 7 sowie der Raum $(\mathfrak{N}, \rho_S)$ (vgl. die daran anschließende Bemerkung) Bemerkung sind ebenfalls normiert, da die Totalvariation ρ_T wie auch die Metrik ρ_S offensichtlich die Homogenitätsbedingung erfüllt.

Lemma 10 (Vervollständigungssatz II)**.** Jeder (unvollständige) lineare normierte Raum $(\mathcal{X}, \|\bullet\|)$ besitzt (mindestens) eine Vervollständigung. Die Vervollständigung ist eindeutig bis auf Kongruenz, d.h. bis auf eine lineare Isometrie.

Beweis: Nach Satz 4 besitzt $(\mathcal{X}, \|\bullet\|)$ jedenfalls als linearer metrischer Raum eine Vervollständigung. Die dort im Beweis benutzte Metrik auf den Äquivalenzklassen von Cauchy-Folgen erfüllt aber offensichtlich die Homogenitätsbedingung aus Lemma 8, so dass die dort konstruierte Vervollständigung ein linearer normierter Raum ist. Da ferner die Grenzwertbildung linear ist, sind die im Beweis verwendeten Isometrien selbst auch linear, womit das Lemma gezeigt ist. Für Einzelheiten verweisen wir auf BACHMAN / NARICI (2000), Abschnitt 8.3 oder TAYLOR (1963), Abschnitt 3.13. ■

Definition 24 (Prä-Hilbert-Raum)**.** Ein Vektorraum $\mathcal{X}$ über K heißt *Prä-Hilbert-Raum*, wenn auf $\mathcal{X}\times\mathcal{X}$ eine sesquilineare Abbildung $(\bullet,\bullet)$ mit Werten in K definiert ist (*Skalarprodukt, inneres Produkt*)[9] mit der Eigenschaft

$$(f,f)\geq 0 \text{ für alle } f\in\mathcal{X} \quad \text{mit} \quad (f,f)=0 \Leftrightarrow f=\mathbf{0}.$$

Bemerkung: In einem Prä-Hilbert-Raum gilt trivialerweise noch:

$$(f,\alpha g)=\overline{(\alpha g,f)}=\overline{\alpha}\cdot\overline{(g,f)}=\overline{\alpha}(f,g) \text{ für alle } f,g\in\mathcal{X},\ \alpha\in K$$

sowie

$$(f,\mathbf{0})=(\mathbf{0},f)=0 \text{ für alle } f\in\mathcal{X}.$$

Lemma 11 (Ungleichungen von Schwarz und Minkowski)**.** In einem Prä-Hilbert-Raum $\mathcal{X}$ gilt:

$$|(f,g)|\leq\sqrt{(f,f)}\cdot\sqrt{(g,g)} \quad \text{sowie}$$
$$\sqrt{(f+g,f+g)}\leq\sqrt{(f,f)}+\sqrt{(g,g)}$$

für alle $f,g\in\mathcal{X}$ mit Gleichheit in der oberen Ungleichung genau dann, wenn f und g linear abhängig sind.

Beweis: Die Aussage ist für $g=\mathbf{0}$ trivial; für $g\neq\mathbf{0}$ setze $\alpha:=-\frac{(f,g)}{(g,g)}$. Dann gilt:

$$\begin{aligned}0&\leq(f+\alpha g,f+\alpha g)=(f,f)+\alpha(g,f)+\overline{\alpha}(f,g)+\alpha\overline{\alpha}(g,g)\\&=(f,f)-\frac{(f,g)}{(g,g)}(g,f)-\frac{(g,f)}{(g,g)}(f,g)+\frac{(f,g)}{(g,g)}\frac{(g,f)}{\cancel{(g,g)}}\cancel{(g,g)}\\&=(f,f)-\frac{(f,g)}{(g,g)}(g,f),\end{aligned}$$

also

$$(f,f)(g,g)\geq(f,g)(g,f)=(f,g)\overline{(f,g)}=|(f,g)|^2$$

---

[9] D.h. es gilt:

$(f,g)=\overline{(g,f)}$ für alle $f,g\in\mathcal{X}$ [konjugiert-komplexe Zahl]

$(\alpha f,g)=\alpha(f,g)$ für alle $f,g\in\mathcal{X},\ \alpha\in K$

$(f+g,h)=(f,h)+(g,h)$ für alle $f,g,h\in\mathcal{X}$.

mit Gleichheit genau dann, wenn $(f+\alpha g, f+\alpha g)=0$, also $f+\alpha g=\mathbf{0}$ ist und damit f und g linear abhängig sind. Die beweist die erste Ungleichung. Die zweite Ungleichung ergibt sich aus[10]

$$(f,g)+(g,f)=(f,g)+\overline{(f,g)}\le 2\left|(f,g)\right|\le 2\sqrt{(f,f)}\cdot\sqrt{(g,g)}$$

und damit

$$\begin{aligned}(f+g,f+g)&=(f,f)+(f,g)+(g,f)+(g,g)\\&\le(f,f)+2\sqrt{(f,f)}\cdot\sqrt{(g,g)}+(g,g)=\left(\sqrt{(f,f)}+\sqrt{(g,g)}\right)^2\end{aligned}$$

für alle $f,g\in\mathcal{X}$. ■

Lemma 12. Es sei $\mathcal{X}$ ein Prä-Hilbert-Raum. Dann wird durch

$$\|f\|:=\sqrt{(f,f)}\quad\text{für alle } f\in\mathcal{X}$$

eine Norm auf $\mathcal{X}$ definiert, unter der $(\mathcal{X},\|\bullet\|)$ ein linearer normierter Raum wird.

Beweis: Es sind nur die beiden letzten Eigenschaften aus Definition 22 nachzuweisen:

a) $\|f+g\|=\sqrt{(f+g,f+g)}\le\sqrt{(f,f)}+\sqrt{(g,g)}=\|f\|+\|g\|$ ist nach Lemma 10 für alle $f,g\in\mathcal{X}$ erfüllt;
b) $\|\alpha f\|^2=(\alpha f,\alpha f)=\alpha\bar{\alpha}(f,f)=|\alpha|^2\cdot\|f\|^2$ für alle $\alpha\in K$ und $f\in\mathcal{X}$ ist nach Voraussetzung (Sesquilinearität) erfüllt. ■

Definition 25 (Hilbert-Raum)**.** Ein vollständiger Prä-Hilbert-Raum heißt *Hilbert-Raum*.

Bemerkung: Jeder Hilbert-Raum ist also ein Banach-Raum, aber nicht notwendig umgekehrt. Die Abbildung $(\bullet,\bullet)$ ist darüber hinaus stetig auf $\mathcal{X}\times\mathcal{X}$.

Satz 12 (P. Jordan / J. von Neumann)**.** Ein Banach-Raum $(\mathcal{X},\|\bullet\|)$ lässt sich genau dann mit einem Skalarprodukt $(\bullet,\bullet)$ mit der Eigenschaft $(f,f)=\|f\|^2$ für alle $f\in\mathcal{X}$ versehen, we*nn die Parallelogramm-Identität* gilt:

$$\|f+g\|^2+\|f-g\|^2=2\left\{\|f\|^2+\|g\|^2\right\}\quad\text{für alle } f,g\in\mathcal{X}.$$

Das Skalarprodukt ist dabei eindeutig bestimmt.

[10] Für $z=x+iy$ ist $|x|=|\Re(z)|\le\sqrt{x^2+y^2}=|z|$.

Beweis:

"$\Rightarrow$": Für $f, g \in \mathcal{X}$ erhält man durch Ausrechnen:

$$\begin{aligned}\|f+g\|^2+\|f-g\|^2 &= (f+g, f+g)+(f-g, f-g) \\ &= (f,f)+(g,f)+(f,g)+(g,g)+(f,f)-(g,f)-(f,g)+(g,g) \\ &= 2\{(f,f)+(g,g)\} = 2\left\{\|f\|^2+\|g\|^2\right\}.\end{aligned}$$

"$\Leftarrow$": Definiere $(\bullet,\bullet)$ durch

$$(f,g) := \begin{cases} \frac{1}{4}\left(\|f+g\|^2-\|f-g\|^2\right) & \text{für } K=\mathbb{R} \\ \frac{1}{4}\left(\|f+g\|^2-\|f-g\|^2\right)-\frac{i}{4}\left(\|i\cdot f+g\|^2-\|i\cdot f-g\|^2\right) & \text{für } K=\mathbb{C} \end{cases} \quad \text{für } f,g\in\mathcal{X}.$$

Wir zeigen jetzt nur exemplarisch für den Fall $K=\mathbb{R}$, dass hierdurch tatsächlich ein Skalarprodukt definiert wird. Für $f,g,h\in\mathcal{X}$ gilt zunächst:

$$\begin{aligned}\|f+g+h\|^2 &= 2\|f+h\|^2+2\|g\|^2-\|f-g+h\|^2 =: a \quad \text{sowie} \\ \|f+g+h\|^2 &= 2\|g+h\|^2+2\|f\|^2-\|-f+g+h\|^2 =: b.\end{aligned}$$

Damit ergibt sich auch sofort

$$\|f+g+h\|^2 = \frac{a+b}{2} = \|f+h\|^2+\|f\|^2+\|g+h\|^2+\|g\|^2-\frac{1}{2}\left(\|f-g+h\|^2+\|-f+g+h\|^2\right).$$

Ersetzt man hier h durch $-h$, erhält man analog

$$\begin{aligned}\|f+g-h\|^2 &= \|f-h\|^2+\|f\|^2+\|g-h\|^2+\|g\|^2-\frac{1}{2}\left(\|f-g-h\|^2+\|-f+g-h\|^2\right) \\ &= \|f-h\|^2+\|f\|^2+\|g-h\|^2+\|g\|^2-\frac{1}{2}\left(\|-f+g+h\|^2+\|f-g+h\|^2\right).\end{aligned}$$

Es folgt

$$\begin{aligned}(f+g,h) &= \frac{1}{4}\left(\|f+g+h\|^2-\|f+g-h\|^2\right) = \frac{1}{4}\left(\|f+h\|^2+\|g+h\|^2-\|f-h\|^2-\|g-h\|^2\right) \\ &= (f,h)+(g,h).\end{aligned}$$

Induktiv folgt hieraus auch

$$(nf,g) = ((n-1)f,g)+(f,g) = \cdots = n(f,g) \text{ für alle } n\in\mathbb{N}$$

und wegen $(f,g)=(-f,g)$ sowie $(\mathbf{0},g)=0$ nach Definition sogar

$$(nf,g) = n(f,g) \text{ für alle } n\in\mathbb{Z}^+.$$

Für $\alpha=\frac{m}{n}\in\mathbb{Q}$ mit $n\in\mathbb{N}$, $m\in\mathbb{Z}^+$ ergibt sich somit

$$(\alpha f,g) = \frac{1}{n}(n\alpha f,g) = \frac{1}{n}(mf,g) = \frac{m}{n}(f,g) = \alpha(f,g).$$

Für $\alpha \in \mathbb{R}$ existiert allgemeiner eine Folge $\{\alpha_n\}_{n\in\mathbb{N}} \subseteq \mathbb{Q}$ mit $\lim_{n\to\infty} \alpha_n = \alpha$, so dass

$$\begin{aligned}(\alpha f, g) &= \frac{1}{4}\left(\|\alpha f + g\|^2 - \|\alpha f - g\|^2\right) = \lim_{n\to\infty} \frac{1}{4}\left(\|\alpha_n f + g\|^2 - \|\alpha_n f - g\|^2\right) \\ &= \lim_{n\to\infty} (\alpha_n f, g) = \lim_{n\to\infty} \{\alpha_n (f, g)\} = \alpha (f, g)\end{aligned}$$

folgt. Damit ist (aus Symmetriegründen) die Bilinearität von $(\bullet,\bullet)$ gezeigt.

Wegen $(f, f) = \|f\|^2 \geq 0$ für $f \in \mathcal{X}$ mit Gleichheit genau dann, wenn $\|f\|^2 = 0 \Leftrightarrow f = \mathbf{0}$, ist der Satz also (für den reellen Fall) insgesamt gezeigt. ■

Beispiel 8. Der in Beispiel 1, Teil e) angegebene metrische Raum ist ein Hilbert-Raum mit dem Skalarprodukt[11]

$$(f, g) := \sum_{n=1}^{\infty} f_n \overline{g_n} \quad \text{für} \quad f, g \in \ell^2,$$

was wegen

$$\left|\sum_{n=1}^{\infty} f_n \overline{g_n}\right| \leq \sqrt{\sum_{n=1}^{\infty} |f_n|^2} \sqrt{\sum_{n=1}^{\infty} |g_n{}^2|} < \infty \quad \text{für} \quad f, g \in \ell^2$$

wohldefiniert ist. Der Raum $C[a,b]$ mit dem Skalarprodukt[12]

$$(f, g) := \int_a^b f(x)\overline{g(x)}\,dx, \quad f, g \in C[a,b]$$

ist dagegen nur ein Prä-Hilbert-Raum, da der Raum nicht vollständig ist. Die Vervollständigung zu einem Hilbert-Raum ist der bezüglich des Lebesgue-Maßes $\mathfrak{m}^1$ über $[a,b]$ absolut quadratisch integrierbaren Funktionen[13], mit derselben formalen Definition der Norm, gegeben durch

$$\|f\| := \int_a^b |f|^2 \, d\mathfrak{m}^1$$

(vgl. auch die entsprechende Bemerkung in Beispiel 4).

[11] Hier sind auch $\mathbb{C}$-wertige Folgen zugelassen.
[12] Hier sind auch $\mathbb{C}$-wertige stetige Funktionen zugelassen.
[13] Genauer gesagt, der Raums der Äquivalenzklassen von fast überall gleichen quadratisch integrierbaren Funktionen.

Definition 26 (Orthogonalität)**.** In einem Prä-Hilbert-Raum $\mathcal{X}$ heißen zwei Vektoren *f* und *g* *orthogonal*, wenn gilt:

$$(f,g)=0,$$

in Zeichen: $f \perp g$. Für eine Teilmenge $A \subseteq \mathcal{X}$ heißt

$$A^{\perp} := \left\{ g \in \mathcal{X} \mid \forall\, f \in A : (f,g)=0 \right\} = \left\{ g \in \mathcal{X} \mid \forall\, f \in A : f \perp g \right\}$$

das *orthogonale Komplement* von *A* in $\mathcal{X}$.

Bemerkung: Wegen $(f,g)=\overline{(g,f)}$ für $f,g \in \mathcal{X}$ ist die Relation "$\perp$" symmetrisch. Ist ferner $\mathcal{X}$ ein Hilbert-Raum und *A* abgeschlossen, so auch $A^{\perp}$.

Ein wichtiges Resultat bezüglich Orthogonalität im Hilbert-Raum gibt der nachfolgende Satz, den wir hier nur ohne Beweis zitieren (vgl. RUDIN (1987), Theorem 4.11).

Satz 13 (Projektionssatz)**.** Es sei $\mathcal{Y}$ ein abgeschlossener Teilraum eines Hilbert-Raumes $\mathcal{X}$. Dann gilt:

a) Jedes $f \in \mathcal{X}$ besitzt eine eindeutige Zerlegung der Form

$$f = Tf + Sf \text{ mit } Tf \in \mathcal{Y} \text{ und } Sf \in \mathcal{Y}^{\perp}.$$

Ferner gilt:

$$\|Tf - f\| = \min\left\{ \|g - f\| \mid g \in \mathcal{Y} \right\} \text{ und } \|Sf - f\| = \min\left\{ \|g - f\| \mid g \in \mathcal{Y}^{\perp} \right\},$$

d.h. Tf und Sf sind diejenigen Punkte aus $\mathcal{Y}$ bzw. $\mathcal{Y}^{\perp}$, die dem Punkt f am nächsten liegen.

b) Die Abbildungen *T* und *S* sind linear.

c) Es gilt: $\|f\|^2 = \|Tf\|^2 + \|Sf\|^2$ für alle $f \in \mathcal{X}$.

Die Abbildungen *T* und *S* heißen *(die) orthogonale(n) Projektionen* von $\mathcal{X}$ auf $\mathcal{Y}$ bzw. $\mathcal{Y}^{\perp}$.

Satz 14 (Trennungssatz)**:** Es sei *C* eine nicht-leere, kompakte, konvexe[14] Teilmenge in einem reellen[15] Hilbert-Raum $\mathcal{X}$ und $\mathcal{V} \subseteq \mathcal{X}$ ein abgeschlossener linearer Teilraum mit $C \cap \mathcal{V} = \varnothing$. Dann existiert ein $f \in \mathcal{V}^{\perp}$ mit

$$\|f\| = \inf\left\{ \|g-h\| \mid g \in C,\ h \in \mathcal{V} \right\} =: \operatorname{dist}(C,\mathcal{V}) > 0$$

und

$$(f,g) \geq \|f\|^2 > 0 \text{ für alle } g \in C.$$

[14] D.h. mit zwei Punkten $f,g \in C$ ist auch $tf + (1-t)g \in C$ für alle $t \in [0,1]$.

[15] D.h. über dem Körper $K = \mathbb{R}$.

Beweis: Sei zunächst $\mathcal{W} = \{\mathbf{0}\}$, wobei $\mathbf{0}$ wieder den Nullvektor bezeichne [diese Menge ist ein abgeschlossener Teilraum von $\mathcal{X}$], und *F* eine beliebige, nicht-leere, konvexe, abgeschlossene Teilmenge von $\mathcal{X}$ mit $F \cap \mathcal{W} = \varnothing$, d.h. $\mathbf{0} \notin F$. Dann existiert ein $f \in F$ mit

$$\|f\| = \operatorname{dist}(F, \mathcal{W}) = \inf\{\|g - h\| \mid g \in F,\, h \in \mathcal{W}\} = \inf\{\|g\| \mid g \in F\} > 0.$$

Es gibt nämlich eine Folge $\{g_n\}_{n\in\mathbb{N}} \subseteq F$ mit $\lim_{n\to\infty}\|g_n\| = \operatorname{dist}(F, \mathcal{W}) =: \Delta$. Nach der Paralelogramm-Identität gilt nun

$$\|g_n + g_m\|^2 + \|g_n - g_m\|^2 = 2\left\{\|g_n\|^2 + \|g_m\|^2\right\} \text{ bzw.}$$

$$\|g_n - g_m\|^2 = 2\left\{\|g_n\|^2 + \|g_m\|^2\right\} - 4\left\|\frac{1}{2}(g_n + g_m)\right\|^2 \le 2\left\{\|g_n\|^2 + \|g_m\|^2\right\} - 4\Delta^2 \text{ für alle } n, m \in \mathbb{N},$$

weil wegen der Konvexität von *F* jeder Punkt $\frac{1}{2}(g_n + g_m)$ zu *F* gehört, also $\left\|\frac{1}{2}(g_n + g_m)\right\| \ge \Delta$ erfüllt. Nach Voraussetzung gilt also

$$0 \le \lim_{n,m\to\infty}\|g_n - g_m\|^2 = \le 2\left\{\lim_{n\to\infty}\|g_n\|^2 + \lim_{m\to\infty}\|g_m\|^2\right\} - 4\Delta^2 = 0,$$

womit $\{g_n\}_{n\in\mathbb{N}} \subseteq F$ eine Cauchy-Folge ist und deshalb in $\mathcal{X}$ mit Grenzwert f konvergiert. Da *F* abgeschlossen ist, gehört aber f zu *F*. Wegen $f \ne \mathbf{0}$ nach Voraussetzung ist dann

$$\|f\| = \operatorname{dist}(F, \mathcal{W}) = \inf\{\|g\| \mid g \in F\} > 0,$$

wie behauptet. Da *F* konvex ist, gehören für $0 < t \le 1$ dann mit *f* auch alle Punkte $f + t(u - f)$ mit $u \in F$ ebenfalls zu *F*, mit

$$\begin{aligned}
&\|f\|^2 + t^2\|u - f\|^2 + 2t(f, u - f) \;=\; \|f + t(u - f)\|^2 \ge \|f\|^2 \quad \forall t \in (0,1] \\
\Leftrightarrow\;& t\|u - f\|^2 + 2(f, u - f) \;\ge\; 0 \quad \forall t \in (0,1] \\
\Leftrightarrow\;& (f, u) \;\ge\; (f, f) \;=\; \|f\|^2 > 0;
\end{aligned}$$

nach Voraussetzung folgt also

$$(f, u) \;\ge\; \|f\|^2 > 0 \text{ für alle } u \in F.$$

Damit ist der Satz in der speziellen Situation $\mathcal{V} = \mathcal{W} = \{\mathbf{0}\}$ und $C = F$ (für sogar nur abgeschlossene) konvexe Mengen *F* bewiesen, da hier trivialerweise

$$(f, g) \;=\; 0 \text{ für alle } g \in \mathcal{V}, \quad \text{also } f \in \mathcal{V}^\perp$$

gilt.

Sei nun $\mathcal{V}$ ein beliebiger, von $\{\mathbf{0}\}$ verschiedener abgeschlossener Teilraum und *C* kompakt. Wir betrachten dann die Menge $L := C - \mathcal{V} = \{g - h \mid g \in C,\, h \in \mathcal{V}\}$. Diese ist abgeschlossen und konvex, mit

$$\operatorname{dist}\left(L,\{\mathbf{0}\}\right)=\operatorname{dist}\left(C,\mathcal{V}\right)>0,$$

also insbesondere $L\cap\{\mathbf{0}\}=\varnothing.$ [Dies zeigen wir zum Ende des Beweises separat.] Ersetzt man nun in der obigen Argumentation *F* durch *L*, so folgt die Existenz eines Vektors $f\in L$ mit

$$\left(f,g-\beta h\right)\;=\;\left(f,g\right)-\beta\left(f,h\right)\;\geq\;\|f\|^2>\;0$$

wegen $f-\beta g\in L$ für alle $g\in C,\; h\in\mathcal{V},\; \beta\in\mathbb{R}.$ Da β beliebig gewählt werden kann, folgt nun direkt $\left(f,h\right)=\;0$ für alle $h\in\mathcal{V}$ sowie $\left(f,g\right)\geq\;\|f\|^2\;>\;0$ für alle $g\in C,$ wie behauptet.
Es bleibt noch zu zeigen, dass $L=C-\mathcal{V}=\left\{g-h\;\middle|\;g\in C,\,h\in\mathcal{V}\right\}$ abgeschlossen und konvex ist. Sei dazu $f\in\mathcal{X}$ Häufungspunkt von *L*. Dann gibt es eine Folge $\left\{f_n\right\}_{n\in\mathbb{N}}\subseteq L$ mit $f=\lim\limits_{n\to\infty}f_n.$ Jedes f_n besitzt dabei eine Darstellung der Form $f_n=g_n-h_n$ mit $g_n\in C,\,h_n\in\mathcal{V}.$ Da *C* als kompakte Menge insbesondere folgenkompakt ist, gibt es ein $g\in C$ und eine gegen *g* konvergente Teilfolge $\left\{g_{n_k}\right\}_{k\in\mathbb{N}}.$ Es folgt $f=\lim\limits_{n\to\infty}f_n=\lim\limits_{k\to\infty}\left(g_{n_k}-h_{n_k}\right)=g-\lim\limits_{k\to\infty}h_{n_k},$ d.h. die Folge $\left\{h_n\right\}_{n\in\mathbb{N}}\subseteq\mathcal{V}$ besitzt eine in $\mathcal{X}$ konvergente Teilfolge $\left\{h_{n_k}\right\}_{k\in\mathbb{N}}$ mit $\lim\limits_{k\to\infty}h_{n_k}=g-f.$ Da $\mathcal{V}$ abgeschlossen ist, ist also $\lim\limits_{k\to\infty}h_{n_k}=g-f\in\mathcal{V}$ und somit $f\in C-\mathcal{V}=L,$ d.h. *L* ist abgeschlossen. *L* ist aber auch konvex, denn für $g_1,g_2\in C,\,h_1,h_2\in\mathcal{V}$ ist auch

$$t\left(g_1-h_1\right)+(1-t)\left(g_2-h_2\right)=\underbrace{\left\{tg_1+(1-t)g_2\right\}}_{\in C,\text{ da }C\text{ konvex}}-\underbrace{\left\{th_1+(1-t)h_2\right\}}_{\in\mathcal{V},\text{ da }\mathcal{V}\text{ Teilraum}}\in C-\mathcal{V}\;\text{ für }\;0<t<1.$$

Der Satz ist damit vollständig bewiesen. ■

Bemerkung: Der obige Trennungssatz ist *das* zentrale Beweismittel für den so genannten *Fundamentalsatz der Optionspreistheorie* in der Stochastischen Finanzmathematik; vgl. das Skript STOCHASTISCHE FINANZMATHEMATIK I, Satz 8.

Auf die Voraussetzung der Konvexität kann im ersten Beweisteil für Satz 13 übrigens nicht verzichtet werden, wie das Beispiel 1, Teil e) für die Menge $F=\left\{f_n\right\}_{n\in\mathbb{N}}$ mit $f_{nk}=\begin{cases}1+\dfrac{1}{n}, & k=n\\ 0, & k\neq n\end{cases}$ für $n,k\in\mathbb{N}$ zeigt. Hier ist *F* abgeschlossen mit

$$\operatorname{dist}\left(F,\mathcal{W}\right)=\inf\left\{\|g\|\;\middle|\;g\in F\right\}=1,$$

aber es gibt kein $f\in F$ mit $\|f\|=1.$

I. 1.5. ℓ^p- und L^p-Räume

In diesem Abschnitt sollen die fundamentalen *Lebesgue-Räume*, die schon in Beispiel 1, Teile d) und e) sowie in Beispiel 4, Teil c) behandelt wurden, genauer diskutiert werden. Insbesondere der für $p = 2$ resultierende Hilbert-Raum ist für die stochastische Finanzmathematik von besonderem Interesse, weil sich das später zu erklärende *Stochastische Integral* als Grenzwert in einem solchem Raum herausstellt.

Für das Folgende setzen wir einen Maßraum $(\Omega, \mathcal{A}, \mu)$ als gegeben voraus und bezeichnen für $1 \leq p < \infty$ mit

$$\|f\|_p := \left\{ \int |f|^p \, d\mu \right\}^{1/p} \quad \text{für alle } \mathcal{A}\text{-messbaren Abildungen } f \text{ auf } \Omega.$$

Für $p = \infty$ definieren wir speziell

$$\|f\|_\infty := \operatorname{ess\,sup} |f| := \inf \left\{ M \geq 0 \mid \mu \left\{ \omega \in \Omega \,\big|\, |f(\omega)| > M \right\} = 0 \right\},$$

das so genannte *essentielle Supremum* der Abbildung $|f|$.

Definition 27 (Lebesgue-Räume)**.** Für $1 \leq p \leq \infty$ sind die Lebesgue-Räume definiert durch

$$L^p(\Omega, \mathcal{A}, \mu) := \left\{ f : \Omega \to \mathbb{R} \,\big|\, f \text{ ist } \mathcal{A}\text{-messbar und } \|f\|_p < \infty \right\}$$

sowie

$$\mathfrak{L}^p(\Omega, \mathcal{A}, \mu) := \left\{ \tilde{f} \mid f \in L^p(\Omega, \mathcal{A}, \mu) \right\},$$

wobei $\tilde{f}$ die zu f gehörende Äquivalenzklasse bezüglich der Äquivalenzrelation ~, definiert durch

$$f \sim g \;:\Leftrightarrow\; f = g \;\; \mu\text{-fast überall} \quad \text{für alle } \mathcal{A}\text{-messbaren Abbildungen } f, g \text{ auf } \Omega,$$

bezeichne. Ferner bezeichne

$$\left\| \tilde{f} \right\|_p := \|f\|_p \;\text{ für } 1 \leq p \leq \infty \text{ und alle } \tilde{f} \in \mathfrak{L}^p(\Omega, \mathcal{A}, \mu).$$

Bemerkung: die Äquivalenzrelation ~ ist verträglich mit der Vektorraumstruktur von $L^p(\Omega, \mathcal{A}, \mu)$, d.h. es gilt

$$\widetilde{\alpha f + g} = \alpha \tilde{f} + \tilde{g} \;\text{ für alle } f, g \in L^p(\Omega, \mathcal{A}, \mu) \text{ und } \alpha \in \mathbb{R}.$$

Ferner ist $\left\| \tilde{f} \right\|_p$ unabhängig von der Wahl des Repräsentanten von $\tilde{f}$, d.h .es gilt

$$\left\| \tilde{f} \right\|_p = \|\tilde{g}\|_p \;\text{ für alle } g \in \tilde{f}.$$

Unser Ziel ist es, zu zeigen, dass $\left(\mathfrak{L}^p(\Omega,\mathcal{A},\mu),\|\!\|\bullet\|\!\|_p\right)$ ein Banach-Raum ist für alle $1 \le p \le \infty$. Um dabei Trivialfälle auszuschließen, wollen wir für das Folgende stets annehmen, dass das Maß μ nicht das Nullmaß ist, weil sonst der Raum $\left(\mathfrak{L}^p(\Omega,\mathcal{A},\mu),\|\!\|\bullet\|\!\|_p\right)$ trivial wäre, da in diesem Fall alle messbaren Abbildungen fast überall gleich der Nullabbildung wären.

Wir benötigen dazu einige vorbereitende Hilfsresultate.

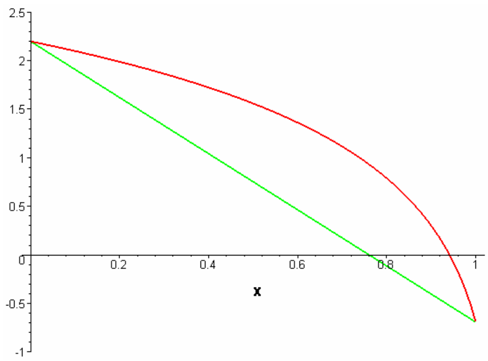
Lemma 13. Es seien $\alpha,\beta \ge 0.$ Dann gilt

$$\alpha^x \beta^{1-x} \le \alpha x + \beta(1-x) \text{ für } 0 < x < 1,$$

mit Gleichheit genau dann, wenn $\alpha = \beta.$

Beweis: O.B.d.A. können wir $\alpha,\beta > 0$ annehmen (sonst ist die linke Seite Null und die Ungleichung trivialerweise erfüllt). Durch Logarithmieren ergibt sich

$$x\ln\alpha + (1-x)\ln\beta \le \ln\left(\alpha x + \beta(1-x)\right),\ 0 < x < 1,$$



da der natürliche Logarithmus (streng) konkav ist (in der Zeichnung rot dargestellt) und damit vollständig oberhalb der Sekante verläuft (in der Zeichnung grün dargestellt; hier: $\alpha = 1/2,\ \beta = 9$). Durch Anwendung der Exponentialfunktion auf diese Ungleichung ergibt sich das gewünschte Ergebnis.

Die Gleichheit wird hier nur erreicht, wenn die rote und grüne Kurve zusammenfallen, d.h. wenn die linke Seite von x unabhängig wird, also (nur) für $\alpha = \beta.$ ■

Lemma 14 (Hölder-Ungleichung I)**.** Es seien p und $q \in (1,\infty)$ so gewählt, dass $\frac{1}{p} + \frac{1}{q} = 1$ gilt (so genannte *konjugierte Indizes*). Dann gilt:

$$f \cdot g \in L^1(\Omega,\mathcal{A},\mu) \text{ mit } \|f \cdot g\|_1 \le \|f\|_p \cdot \|g\|_q \text{ für alle } f \in L^p(\Omega,\mathcal{A},\mu) \text{ und } g \in L^q(\Omega,\mathcal{A},\mu).$$

Beweis: Seien zunächst f und g so gewählt, dass $\|f\|_p = \|g\|_q = 1$ ist. Mit Lemma 13 folgt dann mit der Wahl $\alpha := |f(\omega)|^p,\ \beta := |g(\omega)|^q,\ x := \frac{1}{p},\ 1 - x = \frac{1}{q}$:

$$|f(\omega)\cdot g(\omega)| \le x|f(\omega)|^p + (1-x)|g(\omega)|^q \text{ für alle } \omega \in \Omega,$$

so dass nach Integration resultiert:

$$\int |f\cdot g|\,d\mu \le x\int |f|^p\,d\mu + (1-x)\int |g|^q\,d\mu = x\|f\|_p + (1-x)\|g\|_q = 1.$$

Die Hölder-Ungleichung ist trivialerweise erfüllt, wenn $\|f\|_p = 0$ oder $\|g\|_q = 0$ ist. Für die folgende Rechnung können wir also o.B.d.A. $\|f\|_p > 0$ und $\|g\|_q > 0$ annehmen. Dann haben wir aber $\left\|\frac{f}{\|f\|_p}\right\|_p = \left\|\frac{g}{\|g\|_q}\right\|_q = 1,$ und durch Einsetzen das das vorige Ergebnis erhalten wir

$$\int \frac{|f\cdot g|}{\|f\|_p\|g\|_q}\,d\mu \le 1 \text{ und somit } \|f\cdot g\|_1 = \int |f\cdot g|\,d\mu \le \|f\|_p\|g\|_q,$$

wie gewünscht. ■

Lemma 15 (Hölder-Ungleichung II)**.** Es sei $p = 1$ und $q = \infty$. Dann gilt ebenfalls

$$f\cdot g \in L^1(\Omega, \mathcal{A}, \mu) \text{ mit } \|f\cdot g\|_1 \le \|f\|_p \cdot \|g\|_q \text{ für alle } f \in L^p(\Omega, \mathcal{A}, \mu) \text{ und } g \in L^q(\Omega, \mathcal{A}, \mu).$$

Beweis: Es ist $|f(\omega)\cdot g(\omega)| \le |f(\omega)| \cdot \operatorname{ess\,sup}|g|$ für μ-fast alle $\omega \in \Omega$, woraus die Ungleichung sofort folgt. ■

Lemma 16 (Minkowki-Ungleichung)**.** Für $1 \le p \le \infty$ gilt:

$$f+g \in L^p(\Omega, \mathcal{A}, \mu) \text{ mit } \|f+g\|_p \le \|f\|_p + \|g\|_p \text{ für alle } f, g \in L^p(\Omega, \mathcal{A}, \mu).$$

Beweis: Die Aussage ist klar für $p = \infty$ (Dreiecksungleichung) bzw. für $\|f+g\|_p = 0$. Für $p < \infty$ sei nun $q \in (1, \infty)$ so gewählt, dass $\frac{1}{p} + \frac{1}{q} = 1$ ist. Für den Fall $\|f+g\|_p > 0$ argumentiert man so: wegen $|f+g|^p \le 2^p\left(|f|^p + |g|^p\right)$ ist mit $f, g \in L^p(\Omega, \mathcal{A}, \mu)$ auch $f+g \in L^p(\Omega, \mathcal{A}, \mu)$. Ferner gilt

$$\int |f+g|^p\,d\mu \le \int |f+g|^{p-1}|f|\,d\mu + \int |f+g|^{p-1}|g|\,d\mu \text{ für } f, g \in L^p(\Omega, \mathcal{A}, \mu).$$

Eine zweimalige Anwendung der Hölder-Ungleichung liefert, für $f, g \in L^p(\Omega, \mathcal{A}, \mu)$, unter Beachtung der Beziehung $p = q(p-1)$,

$$\int |f+g|^{p-1}|f|\,d\mu \le \|f\|_p \left\||f+g|^{p-1}\right\|_q, \int |f+g|^{p-1}|g|\,d\mu \le \|g\|_p \left\||f+g|^{p-1}\right\|_q, \text{ mit}$$

$$\left\||f+g|^{p-1}\right\|_q = \left\{\int |f+g|^{(p-1)q}\,d\mu\right\}^{1/q} = \left\{\int |f+g|^p\,d\mu\right\}^{1/q} = \|f+g\|_p^{p/q}, \text{ also}$$

$$\|f+g\|_p^p \le \left\{\|f\|_p + \|g\|_p\right\}\|f+g\|_p^{p/q} \text{ und damit nach Division}$$

$$\|f+g\|_p = \frac{\|f+g\|_p^p}{\|f+g\|_p^{p/q}} \le \|f\|_p + \|g\|_p, \text{ was zu zeigen war.} \qquad \blacksquare$$

Definition 28 (Pseudometrik, Halbnorm)**.** Erfüllt die Abbildung $\rho: \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ in Definition 1 nur die schwächere Bedingung

$$f = g \;\Rightarrow\; \rho(f, g) = 0 \text{ für alle } f, g \in \mathcal{X},$$

So heißt ρ *Pseudometrik*. Erfüllt die Abbildung $\|\bullet\|: \mathcal{X} \to K$ aus Definition 22 nicht die Bedingung b), sondern nur die – aus der aus Bedingung d) ableitbare – schwächere Bedingung

$$f = \mathbf{0} \;\Rightarrow\; \|f\| = 0,$$

so heißt $\|\bullet\|$ *Halbnorm* (vgl. hierzu auch die spätere Definition 33).

Bemerkung: Die Abbildung $\|\bullet\|_p$ auf $L^p(\Omega, \mathcal{A}, \mu)$ ist für $1 \le p \le \infty$ lediglich eine Halbnorm, weil aus $\|f\|_p = 0$ nur $f = 0$ μ-fast überall folgt. Entsprechend ist die Abbildung $\rho_p := \|\bullet - \bullet\|_p$ nur eine Pseudometrik. Dies folgt sofort aus der Definition dieser Größen sowie aus der Minkowski-Ungleichung ($\to$ "Pseudo"-Dreiecksungleichung). Durch Übergang zu den Äquivalenzklassen bezüglich der obigen Relation $\sim$, also durch Übergang von $L^p(\Omega, \mathcal{A}, \mu)$ zu $\mathfrak{L}^p(\Omega, \mathcal{A}, \mu)$, ergibt sich aber ein "echter" linearer normierter bzw. metrischer Raum. Weil die (exakte) Gleichheit in diesen Fällen aber lediglich durch den sehr ähnlichen Begriff der "μ-fast überall-Gleichheit" ersetzt wird, kann man in allen praktischen Anwendungsfällen mit den pseudometrischen bzw. halbnormierten Räumen $L^p(\Omega, \mathcal{A}, \mu)$ rechnen. Entsprechend sind die vorher eingeführten Begriffe wie "Häufungspunkt", "Cauchy-Folge", "Vollständigkeit", "Hilbert-Raum" usw. in der Pseudometrik bzw. Halbnorm zu interpretieren. Wir werden deshalb – dem allgmeinen Gebrauch folgend – im Folgenden mit den "Räumen" $L^p(\Omega, \mathcal{A}, \mu)$ statt $\mathfrak{L}^p(\Omega, \mathcal{A}, \mu)$ weiterarbeiten. Wir fassen das Gesagte wie folgt zusammen:

Lemma 17. Die Räume $\mathfrak{L}^p(\Omega, \mathcal{A}, \mu)$ [bzw. $L^p(\Omega, \mathcal{A}, \mu)$ mit der obigen Interpretation der Pseudometrik bzw. Halbnorm] sind für alle $1 \le p \le \infty$ lineare normierte bzw. lineare metrische Räume.

Satz 15 (Satz von Riesz-Fischer; Vollständigkeit der Lebesgue-Räume)**.** Die Räume $\mathfrak{L}^p(\Omega, \mathcal{A}, \mu)$ [bzw. $L^p(\Omega, \mathcal{A}, \mu)$] sind für alle $1 \le p \le \infty$ vollständig, also Banach-Räume.

Beweis: Wir nehmen zunächst $1 \le p < \infty$ an. Sei nun $\{f_n\}_{n\in\mathbb{N}}$ eine beliebige Cauchy-Folge in $L^p(\Omega, \mathcal{A}, \mu)$. Dann existiert eine Teilfolge $\{f_{n_k}\}_{k\in\mathbb{N}}$ mit der Eigenschaft

$$\left\| f_{n_{k+1}} - f_{n_k} \right\| < \frac{1}{2^k} \text{ für alle } k \in \mathbb{N}.$$

Setze

$$g_k := \sum_{i=1}^{k} \left| f_{n_{i+1}} - f_{n_i} \right| \text{ für alle } k \in \mathbb{N} \text{ und } g := \sum_{i=1}^{\infty} \left| f_{n_{i+1}} - f_{n_i} \right|.$$

Auf Grund der Minkowski-Ungleichung erhalten wir

$$\|g_k\|_p \leq \sum_{i=1}^{k} \left\| f_{n_{i+1}} - f_{n_i} \right\| \leq \sum_{i=1}^{k} \frac{1}{2^i} \leq 1 \text{ für alle } k \in \mathbb{N}.$$

Eine Anwendung des Lemmas von Fatou (siehe Skript STOCHASTIK, Lemma 22) auf die Folge $\left\{g_k^p\right\}_{k\in\mathbb{N}}$ ergibt dann auch

$$\|g\|_p \leq 1.$$

Insbesondere ist also g μ-fast überall endlich, so dass die Reihe

$$f_{n_k} = f_{n_1} + \sum_{i=1}^{k-1} \left(f_{n_{i+1}} - f_{n_i} \right)$$

für $k \to \infty$ μ-fast überall gegen eine messbare Abbildung f konvergiert. Wir wollen zeigen, dass f der gewünschte Grenzwert der ursprünglichen Cauchy-Folge ist. Sei dazu $\varepsilon > 0.$ Dann existiert ein $N \in \mathbb{N},$ so dass $\|f_n - f_m\|_p < \varepsilon$ ist für alle $n,m \geq N.$ Für $m \geq N$ zeigt eine erneute Anwendung vom Lemma von Fatou

$$\int |f - f_m|^p \, d\mu \leq \liminf_{k\to\infty} \int \left| f_{n_k} - f_m \right|^p d\mu \leq \varepsilon^p,$$

also $f - f_m \in L^p(\Omega, \mathcal{A}, \mu)$ für alle $m \geq N$ und damit auch $f = (f - f_m) + f_m \in L^p(\Omega, \mathcal{A}, \mu).$ Damit haben wir aber zugleich auch $\|f_n - f\|_p \to 0$ für $n \to \infty$ gezeigt, so dass in der Tat $f = \lim_{n\to\infty} f_n$ ist.
Für $p = \infty$ argumentieren wir so: Sei $\left\{f_n\right\}_{n\in\mathbb{N}}$ eine beliebige Cauchy-Folge in $L^\infty(\Omega, \mathcal{A}, \mu).$ Nach Definition des essentiellen Supremums gibt es dann μ-Nullmengen A_k und $B_{m,n}$ mit

$$|f_k(x)| \leq \|f_k\|_\infty \text{ für alle } x \in A_k^c \quad \text{und} \quad |f_n(x) - f_m(x)| \leq \|f_n - f_m\|_\infty \text{ für alle } x \in B_{n,m}^c,$$

für alle $k,n,m \in \mathbb{N}.$ Sei $E := \bigcup_{k=1}^{\infty} A_k \cup \bigcup_{n=1}^{\infty} \bigcup_{m=1}^{\infty} B_{n,m}.$ Dann ist auch E eine μ-Nullmenge, und auf E^c konvergiert die Folge $\left\{f_n\right\}_{n\in\mathbb{N}}$ punktweise und gleichmäßig gegen eine beschränkte Funktion f mit o.B.d.A. $f(x) = 0$ für $x \in E.$ Dieses f ist damit der gesuchte Grenzwert. ■

Als Nebenresultat des vorangegangenen Beweises erhalten wir noch folgende, auch für sich allein interessante Aussage.

Satz 16. Jede Cauchy-Folge in $\mathfrak{L}^p(\Omega, \mathcal{A}, \mu)$ [bzw. $L^p(\Omega, \mathcal{A}, \mu)$] mit $1 \leq p \leq \infty$ besitzt eine μ-fast überall punktweise konvergente Teilfolge.

Bemerkung: Wie bereits in Beispiel 4, Teil c) angedeutet, sind die Räume $\mathfrak{L}^p(\Omega, \mathcal{A}, \mu)$ [bzw. $L^p(\Omega, \mathcal{A}, \mu)$] mit $1 \leq p < \infty$ separabel für den Fall $\Omega = [a,b]$ mit der Spur-σ-Algebra

$\Omega \cap \mathcal{B}^1 = \{A \in \mathcal{B}^1 \mid A \subseteq [a,b]\}$ und dem auf $[a,b]$ eingeschränkten Lebesgue-Maß, denn $C[a,b]$ ist hier eine dichte Teilmenge (vgl. WERNER (2004), Korollar I.2.1.). Dagegen ist $\mathfrak{L}^\infty(\Omega,\mathcal{A},\mu)$ [bzw. $L^\infty(\Omega,\mathcal{A},\mu)$] über diesem Maßraum nicht separabel, denn das Funktionensystem $\mathcal{F} := \{\mathbb{1}_{[a,t]} \mid a < t \leq b\} \subseteq L^\infty(\Omega,\mathcal{A},\mu)$ ist beispielsweise überabzählbar mit $\|f-g\|_\infty = 1$ für alle $f,g \in \mathcal{F},\ f \neq g$ (vgl. Beispiel 5).

Satz 17. Der Raum $\mathfrak{L}^2(\Omega,\mathcal{A},\mu)$ [bzw. $L^2(\Omega,\mathcal{A},\mu)$] ist ein Hilbert-Raum mit dem Skalarprodukt

$$(f,g) := \int fg\,d\mu \text{ für alle } f,g \in L^2(\Omega,\mathcal{A},\mu)$$

mit der Einschränkung, dass hier nur gilt:

$$(f,f) \geq 0 \text{ für alle } f \in L^2(\Omega,\mathcal{A},\mu) \quad \text{mit} \quad (f,f) = 0 \Leftrightarrow f = \mathbf{0}\ \mu\text{-fast überall.}$$

Beweis: Folgt sofort aus den Additivitäts-Eigenschaften des μ-Integrals und der Tatsache, dass für $f \in L^2(\Omega,\mathcal{A},\mu)$ gilt:

$$(f,f) = \int f^2\,d\mu = 0 \Leftrightarrow f^2 = \mathbf{0}\ \mu\text{-fast überall} \Leftrightarrow f = \mathbf{0}\ \mu\text{-fast überall.} \qquad \blacksquare$$

Bemerkung: Für den speziellen Messraum $(\Omega,\mathcal{A},\mu) = ([a,b],[a,b]\cap\mathcal{B}^1,\mathfrak{m}^1)$ mit reellen Zahlen $a < b$ und dem (auf $[a,b]\cap\mathcal{B}^1$ eingeschränkten) Lebesgue-Maß $\mathfrak{m}^1$ schreiben wir für $1 \leq p \leq \infty$ statt $\mathfrak{L}^p(\Omega,\mathcal{A},\mu)$ bzw. $L^p(\Omega,\mathcal{A},\mu)$ kürzer auch $\mathfrak{L}^p[a,b]$ bzw. $L^p[a,b]$, analog für den Fall $(\Omega,\mathcal{A},\mu) = (\mathbb{R}^1,\mathcal{B}^1,\mathfrak{m}^1)$ auch $\mathfrak{L}^p(\mathbb{R})$ bzw. $L^p(\mathbb{R})$.
Für den speziellen Messraum $(\Omega,\mathcal{A},\mu) = (\mathbb{N},\mathfrak{P}(\mathbb{N}),\#)$ mit dem abzählenden Maß # schreiben wir statt $L^p(\Omega,\mathcal{A},\mu)$ kürzer auch ℓ^p. Dieser Sonderfall ist strukturell etwas einfacher als der allgemeine Fall, weil hier die leere Menge die einzige Nullmenge ist, wie das folgende Resultat zeigt.

Satz 18. Für $1 \leq p \leq \infty$ ist der Raum $(\ell^p, \|\bullet\|_p)$ mit

$$\|f\|_p = \begin{cases} \sqrt[p]{\sum_{n=1}^\infty |f_n|^p} & \text{für } 1 \leq p < \infty \\ \sup_{n\in\mathbb{N}}\{|f_n|\} & \text{für } p = \infty, \end{cases} \qquad f \in \ell^p$$

ein Banach-Raum und für $1 \leq p < \infty$ auch separabel. Für $p = \infty$ ist $(\ell^\infty, \|\bullet\|_\infty)$ nicht separabel. Für $p = 2$ ist $(\ell^2, \|\bullet\|_2)$ ein Hilbert-Raum mit dem Skalarprodukt

$$(f,g) = \sum_{n=1}^\infty f_n g_n \text{ für } f,g \in \ell^2.$$

Beweis: Es gilt offensichtlich $f = g$ #-fast überall genau dann, wenn $f = g$ (überall) für $f, g \in \ell^p$ gilt, weil $\#(A) = 0 \Leftrightarrow A = \varnothing$ ist für alle $A \subseteq \mathbb{N}$. Damit ist der Übergang zu Äquivalenzklassen fast überall gleicher Elemente unnötig. Ferner ist hier

$$\|f\|_\infty = \operatorname{ess\,sup}|f| = \inf\left\{M \geq 0 \,|\, \#\left\{n \in \mathbb{N} \,|\, |f_n| > M\right\} = 0\right\} = \sup_{n\in\mathbb{N}}\{|f_n|\}$$

für sogar beliebige Folgen $f = \{f_n\}_{n\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$. Nach Satz 15 ist also $\left(\ell^p, \|\bullet\|_p\right)$ für $1 \leq p \leq \infty$ ein Banach-Raum. Für $1 \leq p < \infty$ ist der Raum auch separabel, denn die Mengen

$$A_n := \left\{f \in \mathbb{Q}^{\mathbb{N}} \,|\, \forall\, k \geq n: \, f_k = 0\right\} \subseteq \ell^p, \; n \in \mathbb{N}$$

sind sämtlich abzählbar und damit auch $A := \bigcup_{n=1}^{\infty} A_n \subseteq \ell^p$. A ist aber dicht in ℓ^p, wie man folgendermaßen sieht: Sei $\varepsilon > 0$ gegeben und $f \in \ell^p$ beliebig. Dann existiert wegen der Dichtheit von $\mathbb{Q}$ in $\mathbb{R}$ eine Folge rationaler Zahlen $\{g_n\}_{n\in\mathbb{N}}$ mit $|f_n - g_n| < \frac{\varepsilon}{\sqrt[p]{2^n}}$ für alle $n \in \mathbb{N}$. Wegen $\|f\|_p^p = \sum_{n=1}^{\infty}|f_n|^p < \infty$ existiert ferner ein $N \in \mathbb{N}$ mit $\sum_{n=N}^{\infty}|f_n|^p < \varepsilon^p$. Damit ergibt sich insgesamt für die Folge $h \in A_N \subseteq A$ mit $h_n := \begin{cases} g_n & \text{für } n < N \\ 0 & \text{für } n \geq N \end{cases}$ die Abschätzung

$$\|f - h\|_p^p = \sum_{n=1}^{\infty}|f_n - h_n|^p = \sum_{n=1}^{N-1}|f_n - g_n|^p + \sum_{n=N}^{\infty}|f_n|^p < 2\varepsilon^p,$$

d.h. zu $f \in \ell^p$ existiert ein Element $h \in A$ mit $\|f - h\|_p < 2\varepsilon$, was zu zeigen war. Der Raum $\left(\ell^\infty, \|\bullet\|_\infty\right)$ ist dagegen nicht separabel, weil etwa für die überabzählbare Menge $B := \{0,1\}^{\mathbb{N}} \subseteq \ell^\infty$ gilt: $\|f - g\|_\infty = \sup_{n\in\mathbb{N}}\{|f_n - g_n|\} = 1$ für alle $f, g \in B$ mit $f \neq g$ (vgl. Beispiel 5). Die Hilbert-Raum-Eigenschaft folgt aus Satz 17 unter Beachtung der Tatsache, dass hier für alle $f \in \ell^2$ gilt:

$$(f, f) = \sum_{n=1}^{\infty} f_n^2 = 0 \;\Leftrightarrow\; f = \mathbf{0}.$$

Damit ist der Satz vollständig bewiesen. ■

I. 2. Lineare Operatoren

In der gesamten Mathematik spielen vor allem "lineare" Abbildungen zwischen geeigneten Strukturen eine zentrale Rolle. Dies trifft in besonderem Maße auf lineare metrische bzw. normierte Räume zu, bei denen zur "Linearität" in natürlicher Weise auch die Eigenschaft der "Stetigkeit" hinzukommt. Lineare Operatoren spielen daher auch bei der Charakterisierung gewisser linearer metrischer bzw. normierter Räume eine wesentliche Rolle.

Definition 29 (linearer Operator)**.** Es seien $(\mathcal{X}, \rho_{\mathcal{X}})$ und $(\mathcal{Y}, \rho_{\mathcal{Y}})$ lineare metrische Räume über dem Körper K und $T: \mathcal{X} \to \mathcal{Y}$ eine Abbildung, genannt *Operator*. T heißt

- *additiv*, wenn gilt: $T(f+g)=Tf+Tg$ für alle $f,g\in\mathcal{X}$;
- *homogen*, wenn gilt: $T(\alpha f)=\alpha Tf$ für alle $f\in\mathcal{X}$ und $\alpha\in K$;
- *linear*, wenn T additiv und homogen ist.

Die Abbildung T mit $Tf=\mathbf{0}\in\mathcal{Y}$ für alle $f\in\mathcal{X}$ heißt *Nulloperator* und wird ebenfalls mit $\mathbf{0}$ bezeichnet. Ist speziell $\mathcal{X}\subseteq\mathcal{Y}$, so heißt die Abbildung T mit $Tf=f$ für alle $f\in\mathcal{X}$ *Identitätsoperator* und wird mit $\mathbf{I}$ bezeichnet. Die Menge

$$\mathcal{N}=\{f\in\mathcal{X}\,|\,Tf=\mathbf{0}\}$$

heißt *Kern von T* (Nullraum), die Menge

$$\mathcal{B}=\{Tf\mid f\in\mathcal{X}\}$$

heißt *Bild von T*.

Lemma 18 (Eigenschaften eines linearen Operators)**.** Es seien $(\mathcal{X},\rho_{\mathcal{X}})$ und $(\mathcal{Y},\rho_{\mathcal{Y}})$ lineare metrische Räume über dem Körper K und $T:\mathcal{X}\to\mathcal{Y}$ ein linearer Operator. Dann gilt:

a) Der Kern von T ist ein linearer (metrischer) Teilraum von $(\mathcal{X},\rho_{\mathcal{X}})$, das Bild von T ist ein linearer (metrischer) Teilraum von $(\mathcal{Y},\rho_{\mathcal{Y}})$.
b) Es gilt für alle $f,g\in\mathcal{X}$: $Tf=Tg \Leftrightarrow f-g\in\mathcal{N}$.
c) T ist injektiv genau dann, wenn $\mathcal{N}=\{\mathbf{0}\}$ gilt und surjektiv genau dann, wenn $\mathcal{B}=\mathcal{Y}$ gilt.

Beweis: elementar, durch Nachprüfen der Eigenschaften eines Vektorraums unter Ausnutzung der Linearität von T (wie in der Linearen Algebra üblich). ■

Beispiel 9. Es sei $(\mathcal{X},\rho_{\mathcal{X}})=(\mathcal{Y},\rho_{\mathcal{Y}})=\ell^p$ mit $1\le p\le\infty$. Dann ist die Abbildung T mit

$$T\left(\{f_n\}_{n\in\mathbb{N}}\right):=\{f_{n+1}\}_{n\in\mathbb{N}} \text{ für alle } f=\{f_n\}_{n\in\mathbb{N}}\in\ell^p$$

ein linearer Operator (so genannter *Shift-Operator*). Der Kern ist hier gegeben durch die Menge

$$\mathcal{N}=\{f\in\mathcal{X}\,|\,Tf=\mathbf{0}\}=\left\{\{f_n\}_{n\in\mathbb{N}}\in\ell^p\,\middle|\,\forall k\ge 2:\ f_k=0\right\}.$$

T ist damit nicht injektiv, jedoch – wie man leicht sieht – surjektiv. Dagegen ist der (lineare) Shift-Operator S mit

$$S\left(\{f_n\}_{n\in\mathbb{N}}\right):=\{0,f_n\}_{n\in\mathbb{N}} \text{ für alle } f=\{f_n\}_{n\in\mathbb{N}}\in\ell^p$$

injektiv, aber nicht surjektiv.

Als Besonderheit ist hier festzuhalten, dass gilt: $\|Tf\|_p\le\|f\|_p$ für alle $f\in\ell^p$ mit Gleichheit genau dann, wenn $f_1=0$ ist, und $\|Sf\|_p=\|f\|_p$ für alle $f\in\ell^p$ (d.h. S ist insbesondere eine Isometrie).

Derartige Ungleichungen führen insbesondere bei linearen normierten Räumen auf den sehr wichtigen Begriff der Norm eines Operators, der im folgenden Abschnitt genauer behandelt wird.

I. 2.1. Die Norm eines Operators

Definition 30 (Norm eines Operators)**.** Es seien $(\mathcal{X}, \|\bullet\|_{\mathcal{X}})$ und $(\mathcal{Y}, \|\bullet\|_{\mathcal{Y}})$ lineare normierte Räume über dem Körper K und $T : \mathcal{X} \to \mathcal{Y}$ ein linearer Operator. T heißt *beschränkt*, wenn es eine universelle reelle Zahl $M > 0$ gibt mit der Eigenschaft:

$$\|Tf\|_{\mathcal{Y}} \leq M \|f\|_{\mathcal{X}} \text{ für alle } f \in \mathcal{X}.$$

Das Infimum aller solcher Zahlen $M > 0$ heißt *Norm von T* und wird mit $\|T\|$ bezeichnet, d.h. es gilt

$$\|T\| := \inf \left\{ M \geq 0 \mid \forall\, f \in \mathcal{X} : \ \|Tf\|_{\mathcal{Y}} \leq M \|f\|_{\mathcal{X}} \right\}.$$

Bemerkung: Nach der obigen Definition ergibt sich also stets folgende Ungleichung:

$$\|Tf\|_{\mathcal{Y}} \leq \|T\| \, \|f\|_{\mathcal{X}} \text{ für alle } f \in \mathcal{X}.$$

Lemma 19. Es seien $(\mathcal{X}, \|\bullet\|_{\mathcal{X}})$ und $(\mathcal{Y}, \|\bullet\|_{\mathcal{Y}})$ nicht-triviale[16] lineare normierte Räume über dem Körper K und $T : \mathcal{X} \to \mathcal{Y}$ ein beschränkter linearer Operator. Dann gilt:

$$\|T\| = \sup_{f \neq \mathbf{0}} \frac{\|Tf\|_{\mathcal{Y}}}{\|f\|_{\mathcal{X}}} = \sup_{\|f\|_{\mathcal{X}} = 1} \|Tf\|_{\mathcal{Y}} = \sup_{0 \leq \|f\|_{\mathcal{X}} \leq 1} \|Tf\|_{\mathcal{Y}} = \sup_{0 < \|f\|_{\mathcal{X}} < 1} \|Tf\|_{\mathcal{Y}}.$$

Beweis: Die erste Gleichung folgt sofort aus der Definition der Operatornorm, weil trivialerweise die Beziehung $0 = \|\mathbf{0}\|_{\mathcal{Y}} = \|T(\mathbf{0})\|_{\mathcal{Y}} \leq M \|\mathbf{0}\|_{\mathcal{X}} = M \cdot 0$ für alle reellen $M > 0$ besteht. Ferner gilt aufgrund der Homogenität von T und der Homogenität der Norm für jedes $f \in \mathcal{X}$ mit $f \neq \mathbf{0}$:

$$\frac{\|Tf\|_{\mathcal{Y}}}{\|f\|_{\mathcal{X}}} = \left\| T\left(\frac{1}{\|f\|_{\mathcal{X}}} f \right) \right\|_{\mathcal{Y}} \text{ mit } \left\| \frac{1}{\|f\|_{\mathcal{X}}} f \right\|_{\mathcal{X}} = 1,$$

woraus die zweite Gleichheit folgt. Die dritte Gleichheit folgt aus den Ungleichungen

$$\sup_{0 \leq \|f\|_{\mathcal{X}} \leq 1} \|Tf\|_{\mathcal{Y}} \geq \sup_{\|f\|_{\mathcal{X}} = 1} \|Tf\|_{\mathcal{Y}} = \|T\| \text{ und } \|Tf\|_{\mathcal{Y}} \leq \|T\| \, \|f\|_{\mathcal{X}} \leq \|T\| \text{ für alle } f \in \mathcal{X} \text{ mit } 0 \leq \|f\|_{\mathcal{X}} \leq 1.$$

Die vierte Gleichheit ergibt sich aus folgender Überlegung: ist $\{f_n\}_{n \in \mathbb{N}}$ eine Folge in $\mathcal{X}$ mit $\|f_n\|_{\mathcal{X}} = 1$ und $\lim_{n \to \infty} \|Tf_n\|_{\mathcal{Y}} = \sup_{\|f\|_{\mathcal{X}} = 1} \|Tf\|_{\mathcal{Y}} = \|T\|$ und wählen wir eine Folge $\{\alpha_n\}_{n \in \mathbb{N}} \subseteq (0,1)$ mit $\lim_{n \to \infty} \alpha_n = 1,$ so ist $0 < \|\alpha_n f_n\|_{\mathcal{X}} < 1$ und

[16] D.h. außer dem Nullvektor enhalten $\mathcal{X}$ und $\mathcal{Y}$ mindestes ein weiteres (und damit je unendlich viele) Elemente.

$$\|T\| = \lim_{n\to\infty} \|Tf_n\|_{\mathcal{Y}} = \lim_{n\to\infty} \frac{1}{\alpha_n} \|T(\alpha_n f_n)\|_{\mathcal{Y}} = \underbrace{\left(\lim_{n\to\infty} \frac{1}{\alpha_n}\right)}_{=1} \cdot \left(\lim_{n\to\infty} \|T(\alpha_n f_n)\|_{\mathcal{Y}}\right)$$
$$= \lim_{n\to\infty} \|T(\alpha_n f_n)\|_{\mathcal{Y}} \le \sup_{0<\|f\|_{\mathcal{X}}<1} \|Tf\|_{\mathcal{Y}} \le \sup_{0\le\|f\|_{\mathcal{X}}\le 1} \|Tf\|_{\mathcal{Y}} = \|T\|,$$

womit alles gezeigt ist. ■

Beispiel 10 (Fortsetzung von Beispiel 9)**.** Für die beiden linearen Operatoren T und S gilt jeweils: $\|T\| = \|S\| = 1$. Dies ist nur für T nicht-trivial und kann wie folgt eingesehen werden: aufgrund der Ungleichung $\|Tf\|_p \le \|f\|_p$ für alle $f \in \ell^p$ ist jedenfalls $\|T\| \le 1$. Sei nun $g := (0,1,0,0,\cdots) \in \ell^p$, dann ist $\|g\|_p = 1$ und $Tg = (1,0,0,\cdots) \in \ell^p$ mit $\|Tg\|_p = 1$, also folgt

$$1 \ge \|T\| = \sup_{\|f\|_p = 1} \|Tf\|_p \ge \|Tg\|_p = 1 \text{ und damit } \|T\| = 1.$$

Bemerkung: das vorangehende Beispiel zeigt, wie man auch allgemein eine Gleichheit der Form $\|T\| = M$ nachweisen kann: man zeige zuerst die Gültigkeit von $\|Tf\|_{\mathcal{Y}} \le M \|f\|_{\mathcal{X}}$ für alle $f \in \mathcal{X}$ und dann $\|Tg\|_{\mathcal{Y}} = M \|g\|_{\mathcal{X}}$ [oder auch $\|Tg\|_{\mathcal{Y}} \ge M \|g\|_{\mathcal{X}}$] für ein spezielles $g \in \mathcal{X}$.

Satz 19 (Stetigkeit)**.** Es seien $(\mathcal{X}, \|\bullet\|_{\mathcal{X}})$ und $(\mathcal{Y}, \|\bullet\|_{\mathcal{Y}})$ nicht-triviale lineare normierte Räume über dem Körper K und $T : \mathcal{X} \to \mathcal{Y}$ ein linearer Operator.

a) Folgende Aussagen sind äquivalent:
   i) T ist stetig in einem (beliebigen) Punkt $f \in \mathcal{X}$.
   ii) T ist *gleichmäßig stetig auf* $\mathcal{X}$, d.h. es gilt: zu jedem $\varepsilon > 0$ gibt es ein $\delta > 0$ mit der Eigenschaft:
   $$\|f - g\|_{\mathcal{X}} < \delta \Rightarrow \|Tf - Tg\|_{\mathcal{Y}} < \varepsilon \text{ für alle } f, g \in \mathcal{X}.$$
   iii) T ist beschränkt.

b) T ist beschränkt genau dann, wenn T beschränkte Mengen in $\mathcal{X}$ auf beschränkte Mengen in $\mathcal{Y}$ abbildet. In diesem Fall ist auch das Bild kompakter Mengen in $\mathcal{X}$ unter T kompakt in $\mathcal{Y}$.

Beweis: Für den Teil a) führen wir folgenden Ringschluss:

i) $\Rightarrow$ iii): Sei T stetig ein einem Punkt $f \in \mathcal{X}$. Dann gibt es ein $\delta > 0$ mit der Eigenschaft:

$$\|g\|_{\mathcal{X}} = \|f-(f-g)\|_{\mathcal{X}} < \delta \;\Rightarrow \left\|T\left(f-(f-g)\right)\right\|_{\mathcal{Y}} = \|Tg\|_{\mathcal{Y}} < 1 \text{ für alle } g \in \mathcal{X}.$$

Sei nun $h \in \mathcal{X}$ beliebig, mit $h \neq \mathbf{0}$. Setze $\alpha := \dfrac{\delta}{2\|h\|_{\mathcal{X}}} > 0$ und $g := \alpha h$, so ist $\|g\|_{\mathcal{X}} = \dfrac{\delta}{2} < \delta$ und somit $\|Th\|_{\mathcal{Y}} = \dfrac{1}{\alpha}\|Tg\|_{\mathcal{Y}} < \dfrac{1}{\alpha} = \dfrac{2}{\delta}\|h\|_{\mathcal{X}}$, woraus folgt, dass T beschränkt ist mit $\|T\| \leq \dfrac{2}{\delta}$.

iii) $\Rightarrow$ ii): Sei $\varepsilon > 0$ gegeben; wähle $\delta := \dfrac{\varepsilon}{\|T\|}$. Dann gilt für alle $f, g \in \mathcal{X}$:

$$\|f-g\|_{\mathcal{X}} < \delta \;\Rightarrow \|Tf-Tg\|_{\mathcal{Y}} = \|T(f-g)\|_{\mathcal{Y}} \leq \|T\|\|f-g\|_{\mathcal{X}} < \varepsilon,$$

d.h. T ist gleichmäßig stetig.

ii) $\Rightarrow$ i): trivial.

Zum Beweis von Teil b) sei zunächst T beschränkt und $A \subseteq \mathcal{X}$ eine beschränkte Menge. Dann gilt:

$$\begin{aligned}\delta\left(T(A)\right) &= \sup\left\{\|Tf-Tg\|_{\mathcal{Y}} \mid f,g \in A\right\} = \sup\left\{\|T(f-g)\|_{\mathcal{Y}} \mid f,g \in A\right\}\\ &\leq \|T\| \sup\left\{\|f-g\|_{\mathcal{X}} \mid f,g \in A\right\} = \|T\|\,\delta(A) < \infty,\end{aligned}$$

d.h. $T(A)$ ist beschränkt.

Für die Umkehrung betrachten wir die beschränkte Menge $A := \left\{f \in \mathcal{X} \mid \|f\|_{\mathcal{X}} = 1\right\} \subseteq \bar{K}_1(\mathbf{0})$. Nach Voraussetzung ist dann $T(A) = \left\{Tf \in \mathcal{X} \mid f \in A\right\}$ ebenfalls beschränkt, d.h. es ist

$$\delta\left(T(A)\right) = \sup\left\{\|Tf-Tg\|_{\mathcal{Y}} \mid f,g \in A\right\} < \infty.$$

Für die spezielle Wahl $g = -f$ erhält man somit

$$2\sup\left\{\|Tf\|_{\mathcal{Y}} \mid f \in A\right\} \leq \delta(A) \quad \text{und somit} \quad \|Tf\|_{\mathcal{Y}} \leq \frac{\delta(A)}{2} =: M \quad \text{für alle} \quad f \in A.$$

Sei nun $g \in \mathcal{X}$ beliebig mit $g \neq \mathbf{0}$ und $f := \dfrac{1}{\|g\|_{\mathcal{X}}} g$. Dann ist $f \in A$, und es folgt

$$\|Tg\|_{\mathcal{Y}} = \|g\|_{\mathcal{X}} \|Tf\|_{\mathcal{Y}} \leq M \|g\|_{\mathcal{X}}, \text{ also } \|T\| \leq M < \infty.$$

Die letzte Aussage folgt aus der gleichmäßigen Stetigkeit beschränkter linearer Operatoren (man denke etwa an den Begriff der Folgen-Kompaktheit). ∎

Satz 20 (Beschränktheit der Umkehrabbildung)**.** Es seien $\left(\mathcal{X},\|\bullet\|_{\mathcal{X}}\right)$ und $\left(\mathcal{Y},\|\bullet\|_{\mathcal{Y}}\right)$ nicht-triviale lineare normierte Räume über dem Körper K und $T:\mathcal{X}\to\mathcal{Y}$ ein linearer Operator. Die Umkehrabbildung $T^{-1}:T(\mathcal{X})\to\mathcal{X}$ existiert und ist stetig auf $T(\mathcal{X})$ genau dann, wenn eine reelle Zahl $m>0$ existiert mit der Eigenschaft:

$$\|Tf\|_{\mathcal{Y}}\ge m\|f\|_{\mathcal{X}} \text{ für alle } f\in\mathcal{X}.$$

Beweis: "$\Leftarrow$": T ist injektiv, weil gilt:

$$Tf=\mathbf{0} \;\Rightarrow\; \|Tf\|_{\mathcal{Y}}=0 \;\Rightarrow\; \|f\|_{\mathcal{X}}=0 \;\Rightarrow\; f=\mathbf{0},$$

also ist der Nullraum $\mathcal{N}=\{\mathbf{0}\}$ trivial. Damit existiert T^{-1} auf $T(\mathcal{X})$ als lineare Abbildung und ist dort beschränkt (und damit nach Satz 18 stetig) wegen

$$\left\|T^{-1}(Tf)\right\|_{\mathcal{X}}=\|f\|_{\mathcal{X}}=\le\frac{1}{m}\|Tf\|_{\mathcal{Y}},\text{ d.h. es ist } \left\|T^{-1}\right\|\le\frac{1}{m}<\infty.$$

"$\Rightarrow$": Aufgrund der Stetigkeit und damit Beschränktheit von T^{-1} auf $T(\mathcal{X})$ existiert ein $M>0$ mit

$$\left\|T^{-1}g\right\|_{\mathcal{X}}\le M\|g\|_{\mathcal{Y}} \text{ für alle } g\in T(\mathcal{X}).$$

Damit folgt

$$\|f\|_{\mathcal{X}}=\left\|T^{-1}(Tf)\right\|_{\mathcal{X}}=\le M\|Tf\|_{\mathcal{Y}},\text{ also } \|Tf\|_{\mathcal{Y}}\ge\frac{1}{M}\|f\|_{\mathcal{X}} \text{ für alle } f\in\mathcal{X}.$$

Mit der Wahl $m:=\frac{1}{M}>0$ ist der Satz bewiesen. ■

Beispiel 11. Wir betrachten die Abbildung $T:\ell^1\to\ell^\infty: f\mapsto f$. Diese Abbildung ist beschränkt bzw. stetig wegen

$$\|f\|_\infty=\sup\{|f_n|\,|\,n\in\mathbb{N}\}\le\sum_{n=1}^{\infty}|f_n|=\|f\|_1$$

mit $\|T\|\le 1$. Die spezielle Wahl $f=(1,0,0,0,\cdots)\in\ell^1$ zeigt ferner $\|Tf\|_\infty=1=\|f\|_1$, also gilt sogar $\|T\|=1$. Die Umkehrabbildung $T^{-1}:\ell^\infty\to\ell^1: f\mapsto f$ ist dagegen *nicht stetig* bzw. nicht beschränkt, weil die dafür notwendige Bedingung in Satz 20 nicht erfüllt ist. Dies sieht man an den Folgen $f_n:=\frac{1}{n}\Big(\underbrace{1,1,1,\cdots,1}_{n\text{ mal}},0,0,0,\cdots\Big)\in\ell^\infty$ mit $\|Tf_n\|_\infty=\frac{1}{n}=\frac{1}{n}\|f_n\|_1$ für $n\in\mathbb{N}$.

Lemma 20. Es seien $\left(\mathcal{X}, \|\bullet\|_{\mathcal{X}}\right)$ und $\left(\mathcal{Y}, \|\bullet\|_{\mathcal{Y}}\right)$ nicht-triviale unvollständige lineare normierte Räume über dem Körper K und $T:\mathcal{X}\to\mathcal{Y}$ ein beschränkter [bzw. stetiger] linearer Operator. Dann kann T zu einem beschränkten [bzw. stetigen] linearen Operator $\tilde{T}:\tilde{\mathcal{X}}\to\tilde{\mathcal{Y}}$ fortgesetzt werden mit $\tilde{T}\big|_{\mathcal{X}} = T$ und $\|\tilde{T}\| = \|T\|$, wobei $\tilde{\mathcal{X}}$ und $\tilde{\mathcal{Y}}$ jeweilige Vervollständigungen der gegebenen Räume seien.

Beweis: Zur Vereinfachung nehmen wir an, dass die Mengen $\mathcal{X}$ und $\mathcal{Y}$ selbst schon dichte Teilmengen in $\tilde{\mathcal{X}}$ und $\tilde{\mathcal{Y}}$ sind. Dann ist eine Fortsetzung von T nur für Häufungspunkte $\tilde{f}\in\tilde{\mathcal{X}}$, die nicht selbst schon Element von $\mathcal{X}$ sind, notwendig. Wir definieren $\tilde{T}\tilde{f} := \lim\limits_{n\to\infty} Tf_n$ für Folgen $\{f_n\}_{n\in\mathbb{N}} \subseteq \mathcal{X}$ mit $\tilde{f} := \lim\limits_{n\to\infty} f_n$. Diese Festlegung ist unabhängig von der speziellen Wahl der Folge und ergibt jeweils ein Element $\tilde{T}\tilde{f}\in\tilde{\mathcal{Y}}$. Es bleibt nur zu zeigen, dass $\|\tilde{T}\| = \|T\|$ gilt. Dafür reicht es, für Häufungspunkte $\tilde{f}\in\tilde{\mathcal{X}}$, die nicht selbst schon Element von $\mathcal{X}$ sind, die Ungleichung

$$\|\tilde{T}\tilde{f}\|_{\tilde{\mathcal{Y}}} \le \|T\| \|\tilde{f}\|_{\tilde{\mathcal{X}}}$$

nachzuweisen. Wegen der Stetigkeit der Norm gilt aber

$$\|\tilde{T}\tilde{f}\|_{\tilde{\mathcal{Y}}} = \lim_{n\to\infty} \|Tf_n\|_{\mathcal{Y}} \le \|T\| \lim_{n\to\infty} \|f_n\|_{\mathcal{X}} = \|T\| \|\tilde{f}\|_{\tilde{\mathcal{X}}},$$

womit alles gezeigt ist. (Für einen Alternativbeweis siehe Taylor (1963), Theorem 3.13-A.) ■

Beispiel 12.

a) Wir betrachten die Räume $\mathcal{X} = \mathcal{Y} = C[a,b]$ mit der Supremumsnorm $\|\bullet\|_C$. Es sei $\chi:[a,b]\times[a,b]\to\mathbb{R}$ stetig und der Operator $T:\mathcal{X}\to\mathcal{Y}$ definiert durch

$$Tf(x) := \int_a^b f(u)\chi(x,u)\,du \text{ für } f\in\mathcal{X}.$$

Dann ist T linear und beschränkt wegen

$$\|Tf\|_C = \sup_{a\le x\le b} \left| \int_a^b f(u)\chi(x,u)\,du \right| \le \|f\|_C \sup_{a\le x\le b} \int_a^b |\chi(x,u)|\,du,$$

d.h. es ist $\|T\| \le \sup\limits_{a\le x\le b} \int_a^b |\chi(x,u)|\,du$. Es gilt hier aber sogar Gleichheit, denn die Abbildung $x\mapsto \int_a^b |\chi(x,u)|\,du$ ist stetig auf $[a,b]$ und nimmt daher in einem Punkt $x_0\in[a,b]$ ihr Maximum an. Wenn wir nun zunächst die (ggf. unstetige) Funktion g auf $[a,b]$ definieren durch

$$g(u) := \begin{cases} 1, & \chi(x_0,u) \ge 0 \\ -1, & \chi(x_0,u) < 0, \end{cases}$$

so ist $\sup\limits_{a\le u\le b}|g(u)|\le 1$ und $\int_a^b g(u)\chi(x_0,u)\,du=\int_a^b|\chi(x_0,u)|\,du=\sup\limits_{a\le x\le b}\int_a^b|\chi(x,u)|\,du.$ Die Abbildung g ist aber Element von $L^1[a,b]$ und lässt sich daher in der L^1-Norm durch stetige Funktionen g_n mit $\|g_n\|_C\le 1$ approximieren, mit

$$\left|\int_a^b g(u)\chi(x_0,u)\,du-\int_a^b g_n(u)\chi(x_0,u)\,du\right|\le\int_a^b|g(u)-g_n(u)|\cdot|\chi(x_0,u)|\,du$$
$$\le\|g-g_n\|_{L^1}\sup_{a\le u\le b}|\chi(x_0,u)|\to 0 \text{ für } n\to\infty,$$

so dass

$$\lim_{n\to\infty}\int_a^b g_n(u)\chi(x_0,u)\,du=\int_a^b g(u)\chi(x_0,u)\,du=\sup_{a\le x\le b}\int_a^b|\chi(x,u)|\,du$$

resultiert, was $\|T\|=\sup\limits_{a\le x\le b}\int_a^b|\chi(x,u)|\,du$ impliziert.

Ist speziell $\chi(x,u)=h(x-u)$ mit einer Funktion $h\in C[a-b,b-a]$, so spricht man bei der Operation

$$Tf(x):=\int_a^b f(u)\chi(x,u)\,du=\int_a^b f(u)h(x-u)\,du \text{ für } f\in\mathcal{X}$$

auch von der *Faltung* von *f* und *h*, in Zeichen: $Tf=f*h.$ Hier ist dann

$$\|T\|=\sup_{a\le x\le b}\int_a^b|\chi(x,u)|\,du=\sup_{a\le x\le b}\int_a^b|h(x-u)|\,du=\sup_{a\le x\le b}\int_{x-b}^{x-a}|h(v)|\,dv.$$

Wählen wir beispielsweise $a=0,\ b=1$ und $h(v)=1-v^2,\ -1\le v\le 1,$ so ist

$$\int_{x-1}^{x}|h(v)|\,dv=\frac{2+3x-3x^2}{3} \text{ mit Maximum bei } x_0=\frac{1}{2} \text{ und Wert } \|T\|=\int_{-1/2}^{1/2}|h(v)|\,dv=\frac{11}{12}.$$

b) Es sei $\mathcal{X}=C^1[a,b]=\{f\in C[a,b]\mid f'\in C[a,b]\}$ der Raum der auf $[a,b]$ stetig differenzierbaren Funktionen mit Norm $\|\bullet\|_C$ und $\mathcal{Y}$ wie unter a). Dann ist der Ableitungs-Operator $T:\mathcal{X}\to\mathcal{Y}: f\mapsto f'$ linear, aber nicht beschränkt (also nicht stetig). Dies zeigen die Funktionen $f_n(x):=x^n$, mit $Tf_n(x):=nx^{n-1},\ x\in[0,1],\ n\in\mathbb{N}.$ Hier ist

$$\|f_n\|_C=1, \text{ aber } \|Tf_n\|_C=n=n\|f_n\|_C \text{ für alle } n\in\mathbb{N}.$$

Wenn wir jedoch den Raum $\mathcal{X}=C^1[a,b]$ mit der Norm $\|\bullet\|_{C^1}$ ausstatten, definiert durch

$$\|f\|_{C^1}:=\|f\|_C+\|f'\|_C \text{ für alle } f\in C^1[a,b],$$

so ist

$$\|Tf\|_C = \|f'\|_C \le \|f\|_C + \|f'\|_C = \|f\|_{C^1} \text{ für alle } f \in C^1[a,b], \text{ also } \|T\| \le 1.$$

Bemerkung. Mit Hilfe von Satz 20 kann auch gezeigt werden, dass zwei lineare normierte Räume $\left(\mathcal{X}, \|\bullet\|_{\mathcal{X}}\right)$ und $\left(\mathcal{Y}, \|\bullet\|_{\mathcal{Y}}\right)$ derselben *endlichen* Dimension $d \in \mathbb{N}$ äquivalent sind. Dazu brauchen wir nur zu zeigen, dass dies auf den Spezialfall $\left(\mathcal{X}, \|\bullet\|_{\mathcal{X}}\right) = \left(\mathbb{R}^d, \|\bullet\|_1\right)$ mit $K = \mathbb{R}$ und der Norm

$$\|f\|_1 := \sum_{i=1}^{d} |f_i| \text{ für alle } f \in \mathbb{R}^d$$

zutrifft. Wir betrachten nun eine Basis $\{g_1, \cdots, g_d\} \in \mathcal{Y}$ sowie den linearen Operator $T : \mathbb{R}^d \to \mathcal{Y}$ mit

$$Tf = \sum_{i=1}^{d} f_i g_i \text{ für } f = (f_1, \cdots, f_d) \in \mathbb{R}^d.$$

Insbesondere gilt also $Te_k = g_k$ für den k-ten Einheitsvektor e_k in $\mathbb{R}^d$, für $k = 1, \cdots, d$. (*T* beschreibt also den Übergang von der Standard-Basis in $\mathbb{R}^d$ zu einer Basis des Zielraums $\mathcal{Y}$, ist damit also bijektiv.) Mit $M := \max\left\{\|g_1\|_{\mathcal{Y}}, \cdots, \|g_d\|_{\mathcal{Y}}\right\} \in (0, \infty)$ folgt nun

$$\|Tf\|_{\mathcal{Y}} \le \sum_{i=1}^{d} |\alpha_i| \|g_i\|_{\mathcal{Y}} \le M \sum_{i=1}^{d} |\alpha_i| = M \|f\|_1 \text{ für alle } f \in \mathbb{R}^d \text{ und somit } \|T\| \le M.$$

Es bleibt also nur noch zu zeigen, dass T^{-1} stetig bzw. beschränkt ist. Dazu nutzen wir aus, dass der Rand der abgeschlossenen Einheitskugel $\partial\overline{K}_1(\mathbf{0}) = \left\{ f \in \mathbb{R}^d \mid \|f\|_1 = 1 \right\}$ in $\left(\mathbb{R}^d, \|\bullet\|_1\right)$ (als abgeschlossene und beschränkte Menge) kompakt ist. Die Abbildung $S : \mathbb{R}^d \to \mathbb{R}$ mit

$$Sf := \|Tf\|_{\mathcal{Y}} \text{ für } f \in \mathbb{R}^d$$

ist aber stetig auf $\mathbb{R}^d$ wegen

$$|Sf - Sh| = \left| \|Tf\|_{\mathcal{Y}} - \|Th\|_{\mathcal{Y}} \right| \le \|Tf - Th\|_{\mathcal{Y}} \le M \|f - h\|_1 \text{ für alle } f, h \in \mathbb{R}^d.$$

Damit nimmt $\|Tf\|_{\mathcal{Y}}$ auf der kompakten Menge $\partial\overline{K}_1(\mathbf{0})$ ein (endliches) Minimum an, welches wegen $\mathbf{0} \notin \partial\overline{K}_1(\mathbf{0})$ einen positiven Wert $m > 0$ hat. Wie im Beweis zu Lemma 19 folgt nun

$$m = \inf_{\|f\|_1 = 1} \|Tf\|_{\mathcal{Y}} = \inf_{0 \le \|f\|_1 \le 1} \|Tf\|_{\mathcal{Y}}, \text{ und somit auch } \|Tf\|_{\mathcal{Y}} \ge m \|f\|_1 \text{ für alle } f \in \mathcal{X}.$$

Also ist T^{-1} beschränkt, und es folgt

$$m \|f\|_1 \le \|Tf\|_{\mathcal{Y}} \le M \|f\|_1 \text{ für alle } f \in \mathcal{X},$$

womit alles gezeigt ist.

I. 2.2. Der Raum der beschränkten linearen Operatoren

In diesem Abschnitt betrachten wir etwas genauer die Struktur beschränkter linearer Operatoren zwischen nicht-trivialen linearen normierten Räumen $\left(\mathcal{X},\|\bullet\|_{\mathcal{X}}\right)$ und $\left(\mathcal{Y},\|\bullet\|_{\mathcal{Y}}\right)$ über dem Körper *K*. Die Menge dieser Operatoren bildet offensichtlich auch einen Vektorraum über *K*, wobei die Addition und skalare Multiplikation wie bei Funktionern üblich erklärt ist. Im Folgenden wird die Menge dieser Operatoren mit dem Symbol $\mathfrak{L}[\mathcal{X},\mathcal{Y}]$ bezeichnet.

Satz 21. Sind $\left(\mathcal{X},\|\bullet\|_{\mathcal{X}}\right)$ und $\left(\mathcal{Y},\|\bullet\|_{\mathcal{Y}}\right)$ nicht-triviale lineare normierte Räume über dem Körper *K*, so ist die Menge $\mathfrak{L}[\mathcal{X},\mathcal{Y}]$ der beschränkten linearen Operatoren ebenfalls ein nicht-trivialer linearer normierter Raum unter der (Operator-)Norm[17]

$$\|T\|=\sup_{0\leq\|f\|_{\mathcal{X}}\leq 1}\|Tf\|_{\mathcal{Y}} \text{ für alle } T\in\mathfrak{L}[\mathcal{X},\mathcal{Y}].$$

Ist darüber hinaus $\left(\mathcal{Y},\|\bullet\|_{\mathcal{Y}}\right)$ vollständig, so auch $\mathfrak{L}[\mathcal{X},\mathcal{Y}]$.

Beweis: Wir weisen zuerst die Eigenschaften einer Norm nach. Offensichtlich ist $\|T\|\geq 0$ für alle $T\in\mathfrak{L}[\mathcal{X},\mathcal{Y}]$ mit Gleichheit genau dann, wenn $\|Tf\|_{\mathcal{Y}}=0$ ist für alle $f\in\mathcal{X}$ (nach Definition 30), d.h. genau dann, wenn $T=\mathbf{0}$ ist. Die Dreiecksungleichung ergibt sich so: für $S,T\in\mathfrak{L}[\mathcal{X},\mathcal{Y}]$ ist

$$\|S+T\|=\sup_{0\leq\|f\|_{\mathcal{X}}\leq 1}\|Sf+Tf\|_{\mathcal{Y}}\leq\sup_{0\leq\|f\|_{\mathcal{X}}\leq 1}\left(\|Sf\|_{\mathcal{Y}}+\|Tf\|_{\mathcal{Y}}\right)\leq\sup_{0\leq\|f\|_{\mathcal{X}}\leq 1}\|Sf\|_{\mathcal{Y}}+\sup_{0\leq\|f\|_{\mathcal{X}}\leq 1}\|Tf\|_{\mathcal{Y}}=\|S\|+\|T\|.$$

Die letzte Eigenschaft folgt aus

$$\|\alpha T\|=\sup_{0\leq\|f\|_{\mathcal{X}}\leq 1}\|\alpha Tf\|_{\mathcal{Y}}=\sup_{0\leq\|f\|_{\mathcal{X}}\leq 1}|\alpha|\|Tf\|_{\mathcal{Y}}=|\alpha|\sup_{0\leq\|f\|_{\mathcal{X}}\leq 1}\|Tf\|_{\mathcal{Y}}=|\alpha|\|T\| \text{ für alle } T\in\mathfrak{L}[\mathcal{X},\mathcal{Y}],\ \alpha\in K.$$

Sei nun $\left(\mathcal{Y},\|\bullet\|_{\mathcal{Y}}\right)$ vollständig und $\{T_m\}_{m\in\mathbb{N}}$ Cauchy-Folge in $\mathfrak{L}[\mathcal{X},\mathcal{Y}]$. Dann ist auch $\{T_m f\}_{m\in\mathbb{N}}$ Cauchy-Folge in $\mathcal{X}$ für jedes $f\in\mathcal{X}$ wegen

$$\|T_m f-T_k f\|_{\mathcal{Y}}\leq\|f\|_{\mathcal{X}}\|T_m-T_k\| \text{ für alle } m,k\in\mathbb{N},\ f\in\mathcal{X}.$$

Mit der Festlegung

$$Tf:=\lim_{m\to\infty}T_m f \text{ für alle } f\in\mathcal{X}$$

wird nun ein Element $T\in\mathfrak{L}[\mathcal{X},\mathcal{Y}]$ definiert mit $T=\lim_{m\to\infty}T_m$ in $\mathfrak{L}[\mathcal{X},\mathcal{Y}]$. Dies zeigen wir in folgenden Einzelschritten:

[17] Hier kann jede der nach Lemma 17 äquivalenten Definitionen der Operatornorm verwendet werden.

$\underline{T \text{ ist linear:}}$ Seien $\alpha, \beta \in K,\ f, g \in \mathcal{X},\ m \in \mathbb{N}$. Dann ist

$$\begin{aligned}&\|T(\alpha f + \beta g) - \alpha Tf - \beta Tg\|_{\mathcal{Y}} = \|T(\alpha f + \beta g) - T_m(\alpha f + \beta g) + \alpha T_m f - \beta T_m g + \alpha Tf - \beta Tg\|_{\mathcal{Y}} \\ &\le \underbrace{\|T(\alpha f + \beta g) - T_m(\alpha f + \beta g)\|_{\mathcal{Y}}}_{\to 0} + |\alpha| \underbrace{\|T_m f - Tf\|_{\mathcal{Y}}}_{\to 0} + |\beta| \underbrace{\|T_m g - Tg\|_{\mathcal{Y}}}_{\to 0} \to 0 \text{ für } n \to \infty,\end{aligned}$$

also gilt $T(\alpha f + \beta g) = \alpha Tf + \beta Tg$.

$\underline{T \text{ ist beschränkt:}}$ Seien $m, k \in \mathbb{N}$. Dann gilt $\big|\|T_m\| - \|T_k\|\big| \le \|T_m - T_k\|$, d.h. $\{\|T_m\|\}_{m \in \mathbb{N}}$ ist Cauchy-Folge in $\mathbb{R}$ und damit konvergent, also insbesondere beschränkt, etwa durch $0 < M < \infty$. Es folgt, für alle $f \in \mathcal{X}$ mit $f \ne \mathbf{0}$ und $m \in \mathbb{N}$:

$$\|Tf\|_{\mathcal{Y}} \le \|(T - T_m)f\|_{\mathcal{Y}} + \|T_m f\|_{\mathcal{Y}} \le \underbrace{\|(T - T_m)f\|_{\mathcal{Y}}}_{\to 0} + \|T_m\| \|f\|_{\mathcal{X}} \le \left(\frac{\|(T - T_m)f\|_{\mathcal{Y}}}{\|f\|_{\mathcal{X}}} + M \right) \|f\|_{\mathcal{X}},$$

also $\|Tf\|_{\mathcal{Y}} \le M \|f\|_{\mathcal{X}}$ und somit $\|T\| \le M < \infty$.

$\underline{T \text{ ist Grenzwert:}}$ Zu $\varepsilon > 0$ gibt es nach Voraussetzung ein $N \in \mathbb{N}$ mit

$$\|T_m - T_k\| < \varepsilon \text{ für alle } n, k \ge N.$$

Es folgt

$$\|Tf - T_k f\|_{\mathcal{Y}} \le \|T_m f - T_k f\|_{\mathcal{Y}} + \|Tf - T_m f\|_{\mathcal{Y}} \le \|T_m - T_k\| \|f\|_{\mathcal{X}} + \|Tf - T_m f\|_{\mathcal{Y}} \le \varepsilon \|f\|_{\mathcal{X}} + \underbrace{\|Tf - T_m f\|_{\mathcal{Y}}}_{\to 0 \text{ für } m \to \infty}$$

für $n, k \ge N,\ f \in \mathcal{X}$ und somit

$$\|Tf - T_k f\|_{\mathcal{Y}} \le \varepsilon \|f\|_{\mathcal{X}} \text{ für alle } k \ge N,\ f \in \mathcal{X} \quad \text{bzw.} \quad \|T - T_k\| \le \varepsilon \text{ für alle } k \ge N,$$

woraus $\lim\limits_{k \to \infty} \|T - T_k\| = 0$ folgt. Der Satz ist damit bewiesen. ■

Definition 31 (Konvergenzarten)**.** Es seien $(\mathcal{X}, \|\bullet\|_{\mathcal{X}})$ und $(\mathcal{Y}, \|\bullet\|_{\mathcal{Y}})$ nicht-triviale lineare normierte Räume über dem Körper *K*. Eine Folge $\{T_m\}_{m \in \mathbb{N}} \subseteq \mathfrak{L}[\mathcal{X}, \mathcal{Y}]$ beschränkter linearer Operatoren heißt *gleichmäßige Cauchy-Folge*, wenn gilt:

$$\lim_{m,k \to \infty} \|T_m - T_k\| = 0.$$

Sie heißt *gleichmäßig konvergent gegen* $T \in \mathfrak{L}[\mathcal{X}, \mathcal{Y}]$, wenn gilt:

$$\lim_{m \to \infty} \|T_m - T\| = 0.$$

Die Folge $\{T_m\}_{m \in \mathbb{N}} \subseteq \mathfrak{L}[\mathcal{X}, \mathcal{Y}]$ heißt *starke Cauchy-Folge*, wenn gilt:

$$\lim_{m,k \to \infty} \|T_m f - T_k f\|_{\mathcal{Y}} = 0 \text{ für alle } f \in \mathcal{X}.$$

Sie heißt *stark konvergent gegen* $T \in \mathfrak{L}[\mathcal{X}, \mathcal{Y}]$, wenn gilt:

$$\lim_{m \to \infty} \|T_m f - Tf\|_{\mathcal{Y}} = 0 \text{ für alle } f \in \mathcal{X}.$$

Lemma 21. Unter den Voraussetzungen von Definition 31 gilt:

a) Eine Folge $\{T_m\}_{m\in\mathbb{N}} \subseteq \mathfrak{L}[\mathcal{X},\mathcal{Y}]$ ist genau dann gleichmäßige Cauchy-Folge bzw. gleichmäßig konvergent gegen $T \in \mathfrak{L}[\mathcal{X},\mathcal{Y}]$, wenn gilt: $\{T_m f\}_{m\in\mathbb{N}}$ ist Cauchy-Folge in $\mathcal{Y}$ bzw. konvergent in $\mathcal{Y}$ gleichmäßig in $f \in \bar{K}_r(\mathbf{0})$ für ein $r > 0$.

b) "Gleichmäßigkeit" impliziert "Starkheit", aber nicht notwendig umgekehrt.

Beweis:

a) "$\Rightarrow$": Sei $f \in \bar{K}_r(\mathbf{0})$ für ein $r > 0$. Dann folgt

$$\|T_m f - T_k f\|_{\mathcal{Y}} \le \|T_m - T_k\| \|f\|_{\mathcal{X}} \le r\|T_m - T_k\| \text{ für alle } m,k \in \mathbb{N}, \text{ unabhängig von } f.$$

Dies zeigt die Behauptung für eine gleichmäßige Cauchy-Folge; analog bei gleichmäßiger Konvergenz (ersetze T_k durch T).

"$\Leftarrow$": Sei für $f \in \mathcal{X}$, $f \neq \mathbf{0}$ und $r > 0$ die Zahl $\alpha := \dfrac{r}{\|f\|_{\mathcal{X}}} > 0$ gewählt. Dann ist $\alpha f \in \bar{K}_r(\mathbf{0})$, und es folgt

$$\|T_m f - T_k f\|_{\mathcal{Y}} = \frac{1}{\alpha}\|T_m(\alpha f) - T_k(\alpha f)\|_{\mathcal{Y}} = \frac{1}{r}\|T_m(\alpha f) - T_k(\alpha f)\|_{\mathcal{Y}} \|f\|_{\mathcal{X}},$$

also

$$\|T_m - T_k\| \le \frac{1}{r}\|T_m(\alpha f) - T_k(\alpha f)\|_{\mathcal{Y}} \to 0 \text{ für } m,k \to \infty.$$

Dies zeigt die Behauptung für eine in $f \in \bar{K}_r(\mathbf{0})$ gleichmäßige Cauchy-Folge; analog bei in $f \in \bar{K}_r(\mathbf{0})$ gleichmäßiger Konvergenz (ersetze T_k durch T).

b) Folgt aus der Ungleichung

$$\|T_m f - T_k f\|_{\mathcal{Y}} \le \|T_m - T_k\| \|f\|_{\mathcal{X}} \text{ bzw. } \|T_m f - Tf\|_{\mathcal{Y}} \le \|T_m - T\| \|f\|_{\mathcal{X}} \text{ für } m,k \in \mathbb{N},\ f \in \mathcal{X}. \qquad \blacksquare$$

Beispiel 13. Die Umkehrung in Lemma 21 Teil b) gilt nicht, wie man an folgendem Beispiel sieht: Wir betrachten den Raum ℓ^1 und darauf die linearen beschränkten Operatoren $T_n : \ell^1 \to \ell^1$ mit

$$T_n f := (0,0,\cdots,0, f_{n+1}, f_{n+2},\cdots) \text{ für } f \in \ell^1 \text{ und } n \in \mathbb{N}.$$

Offenbar ist $\|T_n f\|_1 = \sum_{k>n} |f_k| \le \|f\|_1$ für $f \in \ell^1$ und $n \in \mathbb{N}$ und somit $\|T_n\| \le 1$. Die spezielle Folge $\tilde{f} = \Big(\underbrace{0,0,\cdots,0}_{n \text{ mal}},1,0,0,\cdots\Big)$ zeigt aber $\|T_n \tilde{f}\|_1 = \|\tilde{f}\|_1$ und somit $\|T_n\| = 1$ für alle $n \in \mathbb{N}$. Es gilt aber wegen der absoluten Konvergenz der Reihe: $\lim_{n\to\infty}\|T_n f\|_1 = \lim_{n\to\infty}\sum_{k>n}|f_k| = 0$ für alle $f \in \ell^1$ und somit $T_n \to \mathbf{0}$ im starken Sinne, aber nicht im gleichmäßigen Sinne.

Definition 32 (Algebra)**.** Ein Vektorraum $\mathcal{X}$ über dem Körper K heißt *Algebra über K*, wenn eine Abbildung $\circ : \mathcal{X} \times \mathcal{X} \to \mathcal{X}$ (Produkt) gegeben ist mit den folgenden distributiven Eigenschaften:

a) $f \circ (g+h) = f \circ g + f \circ h$ für alle $f, g, h \in \mathcal{X}$
b) $(f+g) \circ h = f \circ h + g \circ h$ für alle $f, g, h \in \mathcal{X}$
c) $\alpha(f \circ g) = \alpha f \circ g = f \circ \alpha g$ für alle $f, g \in \mathcal{X}$, $\alpha \in K$
d) $(f \circ g) \circ h = f \circ (g \circ h)$ für alle $f, g, h \in \mathcal{X}$.

Existiert ferner ein Element $\mathbf{e} \in \mathcal{X}$ mit $\mathbf{e} \circ f = f \circ \mathbf{e} = f$ für alle $f \in \mathcal{X}$, so heißt $\mathbf{e}$ das (eindeutig bestimmte) *Einselement* von $\mathcal{X}$. Eine Algebra heißt *kommutativ*, wenn gilt:

e) $f \circ g = g \circ f$ für alle $f, g \in \mathcal{X}$.

$\mathcal{X}$ heißt *normierte Algebra* über K, wenn zusätzlich gilt:

f) $\|f \circ g\|_{\mathcal{X}} \le \|f\|_{\mathcal{X}} \|g\|_{\mathcal{X}}$ für alle $f, g \in \mathcal{X}$.

Ist $\mathcal{X}$ vollständig (also ein Banach-Raum), so heißt $\mathcal{X}$ auch *Banach-Algebra*.

Lemma 22 (Operator-Algebra)**.** Es sei $\left(\mathcal{X}, \|\bullet\|_{\mathcal{X}}\right)$ ein linearer normierter Raum über K. Dann wird durch

$$(T \circ S) f := T\left(S(f)\right) \text{ für alle } f \in \mathcal{X},\ S, T \in \mathfrak{L}[\mathcal{X}, \mathcal{X}]$$

eine Abbildung $\circ : \mathfrak{L}[\mathcal{X}, \mathcal{X}] \times \mathfrak{L}[\mathcal{X}, \mathcal{X}] \to \mathfrak{L}[\mathcal{X}, \mathcal{X}]$ definiert, unter der $\mathfrak{L}[\mathcal{X}, \mathcal{X}]$ zu einer normierten Algebra wird. Ist $\mathcal{X}$ ein Banach-Raum, so ist $\mathfrak{L}[\mathcal{X}, \mathcal{X}]$ eine Banach-Algebra.

Beweis: Die Eigenschaften a) bis d) aus Definition 32 ergeben sich unmittelbar aus der Linearität der Operatoren. Das Einselement ist hier der Identitätsoperator (vgl. Definition 29). Die Eigenschaft f) ergibt sich aus

$$\|T \circ S\| = \sup_{0 \le \|f\|_{\mathcal{X}} \le 1} \left\|T\left(S(f)\right)\right\|_{\mathcal{X}} \le \sup_{0 \le \|f\|_{\mathcal{X}} \le 1} \left\{\|T\| \|Sf\|_{\mathcal{X}}\right\} \le \sup_{0 \le \|f\|_{\mathcal{X}} \le 1} \left\{\|T\| \|S\| \|f\|_{\mathcal{X}}\right\} = \|T\| \|S\|$$

für alle $S, T \in \mathfrak{L}[\mathcal{X}, \mathcal{X}]$. Nach Satz 21 ist $\mathfrak{L}[\mathcal{X}, \mathcal{X}]$ vollständig, wenn $\mathcal{X}$ vollständig ist. Damit ist das Lemma bewiesen. ■

Beispiel 14 (Fortsetzung von Beispiel 13)**.** Die Teilmenge $\mathfrak{T} := \operatorname{span}\left\{T_n \mid n \in \mathbb{Z}^+\right\} \subseteq \mathfrak{L}\left[\ell^1, \ell^1\right]$ mit $T_0 := \mathbf{I}$ ist eine Banach-Algebra mit dem in Lemma 22 eingeführten Produkt. Hier gilt insbesondere

$$T_n \circ T_m = T_{\max\{n,m\}} = T_m \circ T_n \text{ für alle } n, m \in \mathbb{Z}^+,$$

und damit ist $\mathfrak{T}$ kommutativ wegen

$$\left(\sum_{i=1}^{n} \alpha_i T_i\right) \circ \left(\sum_{j=1}^{m} \beta_j T_j\right) = \sum_{i=1}^{n} \sum_{j=1}^{m} \alpha_i \beta_j\, T_i \circ T_j = \sum_{j=1}^{m} \sum_{i=1}^{n} \beta_j \alpha_i\, T_j \circ T_i = \left(\sum_{j=1}^{m} \beta_j T_j\right) \circ \left(\sum_{i=1}^{n} \alpha_i T_i\right)$$

für alle $\alpha_i, \beta_j \in K,\ i=1,\cdots,n,\ j=1,\cdots,m;\ n,m \in \mathbb{N}.$ Allerdings ist die größere Banach-Algebra $\mathfrak{L}\left[\ell^1, \ell^1\right]$ nicht kommutativ, denn bezeichnet wieder *T* den (ersten) Shift-Operator aus Beispiel 9 und $S = T_1 \in \mathfrak{T},$ so gilt für alle $f \in \ell^1$:

$$f = (f_1, f_2, f_3, \cdots) \xrightarrow{T} (f_2, f_3, f_4, \cdots) \xrightarrow{S} (0, f_3, f_4, \cdots), \text{ aber}$$
$$f = (f_1, f_2, f_3, \cdots) \xrightarrow{S} (0, f_2, f_3, \cdots) \xrightarrow{T} (f_2, f_3, f_4, \cdots),$$

so dass hier $T \circ S \neq S \circ T$ gilt.

I. 2.3. Folgen von beschränkten linearen Operatoren

In diesem Abschnitt soll die zuvor begonnene Diskussion von Konvergenzbeziehungen bei Operatoren vertieft fortgesetzt werden.

Satz 22 ("Uniform-Boundedness-Theorem")**.** Es sei $\left(\mathcal{X}, \|\bullet\|_{\mathcal{X}}\right)$ ein nicht-trivialer Banach-Raum und $\left(\mathcal{Y}, \|\bullet\|_{\mathcal{Y}}\right)$ ein linearer normierter Raum über dem Körper *K*, $\{T_m\}_{m\in\mathbb{N}} \subseteq \mathfrak{L}[\mathcal{X}, \mathcal{Y}]$ eine Folge beschränkter linearer Operatoren. Ist dann für jedes $f \in \mathcal{X}$ die Folge $\left\{\|T_m f\|_{\mathcal{Y}}\right\}_{m\in\mathbb{N}}$ punktweise in $\mathcal{Y}$ beschränkt, dann ist auch die Folge $\left\{\|T_m\|\right\}_{m\in\mathbb{N}}$ der Operatornormen beschränkt.

Beweis: Für $m, k \in \mathbb{N}$ sei

$$\mathcal{X}_{mk} := \left\{ f \in \mathcal{X} \mid \|T_k f\|_{\mathcal{Y}} \le m \right\} = T_k^{-1}\left(\bar{K}_m(\mathbf{0})\right).$$

Dann sind alle Mengen $\mathcal{X}_{mk}$ abgeschlossen in $\mathcal{X}$, da jeder Operator T_k als beschränkter Operator auch stetig ist. Setze

$$\mathcal{X}_m := \bigcap_{k=1}^{\infty} \mathcal{X}_{mk} = \left\{ f \in \mathcal{X} \mid \forall\, k \in \mathbb{N} : \|T_k f\|_{\mathcal{Y}} \le m \right\}.$$

Nach Satz 1, Teil b) sind die Mengen $\mathcal{X}_m$ ebenfalls abgeschlossen mit

$$\mathcal{X} = \bigcup_{m=1}^{\infty} \mathcal{X}_m.$$

Als Banach-Raum ist $\mathcal{X}$ von zweiter Kategorie nach Satz 6, also *nicht* abzählbare Vereinigung nirgends dichter Mengen. Also gibt es ein $m_0 \in \mathbb{N},$ so dass $\mathcal{X}_{m_0}$ *nicht* nirgends dicht ist, d.h. es ist $\mathcal{X}_{m_0}{}^{o} \neq \varnothing.$ Damit enthält $\mathcal{X}_{m_0}$ eine offene und damit als weitere Teilmenge auch abgeschlossene Kugel $\bar{K}_r(f) \subseteq \mathcal{X}_{m_0}$ für ein $f \in \mathcal{X}_{m_0}$ und ein $r > 0.$ Für beliebige $g \in \mathcal{X}$ mit $g \neq \mathbf{0}$ und $k \in \mathbb{N}$ folgt dann mit $\alpha := \dfrac{r}{\|g\|_{\mathcal{X}}} > 0$:

$$\|T_k g\|_{\mathcal{Y}} = \frac{1}{\alpha}\|T_k(\alpha g)\|_{\mathcal{Y}} \le \frac{\|g\|_{\mathcal{X}}}{r}\left\{\Bigg\|T_k\underbrace{(\alpha g - f)}_{\in \bar{K}_r(f)}\Bigg\|_{\mathcal{Y}} + \|T_k f\|_{\mathcal{Y}}\right\} \le \frac{2m_0}{r}\|g\|_{\mathcal{X}} \text{ und somit } \|T_k\| \le \frac{2m_0}{r}.$$

Also ist die Folge $\{\|T_m\|\}_{m\in\mathbb{N}}$ der Operatornormen durch $\dfrac{2m_0}{r}$ beschränkt, was zu zeigen war. ■

Lemma 23. Seien $(\mathcal{X}, \|\bullet\|_{\mathcal{X}})$ und $(\mathcal{Y}, \|\bullet\|_{\mathcal{Y}})$ nicht-triviale Banach-Räume und $\{T_m\}_{m\in\mathbb{N}} \subseteq \mathfrak{L}[\mathcal{X},\mathcal{Y}]$ eine starke Cauchy-Folge beschränkter linearer Operatoren. Dann existiert ein $M>0$, so dass $\|T_k\| \le M$ ist für alle $k\in\mathbb{N}$, und $\{T_m\}_{m\in\mathbb{N}}$ ist stark konvergent gegen ein $T\in\mathfrak{L}[\mathcal{X},\mathcal{Y}]$ mit $\|T\|\le M$.

Beweis: Für beliebige $f\in\mathcal{X}$ gilt $\left|\|T_m f\|_{\mathcal{Y}} - \|T_k f\|_{\mathcal{Y}}\right| \le \|T_m f - T_k f\|_{\mathcal{Y}} \to 0$ mit $n,m\to\infty$, also ist $\{\|T_k f\|_{\mathcal{Y}}\}_{k\in\mathbb{N}}$ reelle Cauchy-Folge und somit beschränkt. Nach Satz 22 existiert also ein $M>0$ mit $\|T_k\|\le M$ für alle $k\in\mathbb{N}$. Da $\mathcal{Y}$ vollständig ist, hat die Folge $\{\|T_k f\|_{\mathcal{Y}}\}_{k\in\mathbb{N}}$ einen Grenzwert in $\mathcal{Y}$; durch

$$Tf := \lim_{m\to\infty} T_m f \text{ für alle } f\in\mathcal{X}$$

wird also wie in Satz 21 ein beschränkter linearer Operator $T\in\mathfrak{L}[\mathcal{X},\mathcal{Y}]$ mit $\|T\|\le M$ definiert. ■

Satz 23 (Banach-Steinhaus)**.** Es seien $(\mathcal{X}, \|\bullet\|_{\mathcal{X}})$ und $(\mathcal{Y}, \|\bullet\|_{\mathcal{Y}})$ nicht-triviale Banach-Räume und $\{T_m\}_{m\in\mathbb{N}} \subseteq \mathfrak{L}[\mathcal{X},\mathcal{Y}]$ eine Folge beschränkter linearer Operatoren mit den Eigenschaften

a) es existiert ein $M>0$ mit $\|T_k\|\le M$ für alle $k\in\mathbb{N}$

b) $\{T_m\}_{m\in\mathbb{N}}$ ist starke Cauchy-Folge auf einer Fundamentalmenge $A\subseteq\mathcal{X}$.

Dann ist $\{T_m\}_{m\in\mathbb{N}}$ stark konvergent auf ganz $\mathcal{X}$ gegen einen linearen beschränkten Operator $T\in\mathfrak{L}[\mathcal{X},\mathcal{Y}]$ mit $\|T\|\le M$.

Beweis: Wir zeigen zunächst, dass $\{T_m\}_{m\in\mathbb{N}}$ stark konvergent auf $\mathrm{span}(A)$ ist. Seien dazu $\alpha_1,\cdots,\alpha_n\in K$ und $f_1,\cdots,f_n\in\mathcal{X}$ für $n\in\mathbb{N}$. Dann gilt:

$$\left\|T_m\left(\sum_{i=1}^n \alpha_i f_i\right) - T_k\left(\sum_{i=1}^n \alpha_i f_i\right)\right\|_{\mathcal{Y}} = \left\|\sum_{i=1}^n \alpha_i (T_m - T_k)(f_i)\right\|_{\mathcal{Y}} \le \sum_{i=1}^n |\alpha_i| \|(T_m - T_k)(f_i)\|_{\mathcal{Y}} \to 0 \text{ mit } n,m\to\infty,$$

womit diese erste Behauptung bewiesen ist. Die Aussage des Satzes folgt jetzt unter Beachtung von $\overline{\mathrm{span}(A)} = \mathcal{X}$ so: Sei $\varepsilon>0$ und $f\in\mathcal{X}$ beliebig. Dann existiert ein $h\in\mathrm{span}(A)$ mit $\|f-h\|_{\mathcal{X}}<\varepsilon$ und ein $N\in\mathbb{N}$ mit $\|T_m h - T_k h\|_{\mathcal{Y}}<\varepsilon$ für $m,n\ge N$. Damit erhält man aber

$$\begin{aligned}\|T_m f - T_k f\|_{\mathcal{Y}} &\le \|T_m f - T_m h\|_{\mathcal{Y}} + \|T_m h - T_k h\|_{\mathcal{Y}} + \|T_k h - T_k f\|_{\mathcal{Y}} \\ &\le \{\|T_m\| + \|T_k\|\}\|f-h\|_{\mathcal{X}} + \|T_m h - T_k h\|_{\mathcal{Y}} \le \{2M+1\}\varepsilon\end{aligned}$$

für $m,k \ge N,$ also ist $\{T_m\}_{m\in\mathbb{N}}$ starke Cauchy-Folge auf ganz $\mathcal{X}$. Die Konvergenzaussage ergibt sich jetzt aus Lemma 23. ■

Wir können jetzt die wesentlichen Ergebnisse dieses Abschnitts folgendermaßen zusammenfassen:

Satz 24. Es seien $(\mathcal{X},\|\bullet\|_{\mathcal{X}})$ und $(\mathcal{Y},\|\bullet\|_{\mathcal{Y}})$ nicht-triviale Banach-Räume und $\{T_m\}_{m\in\mathbb{N}} \subseteq \mathfrak{L}[\mathcal{X},\mathcal{Y}]$ eine Folge beschränkter linearer Operatoren. Notwendig und hinreichend dafür, dass $\{T_m\}_{m\in\mathbb{N}}$ eine starke Cauchy-Folge [bzw. stark konvergent] auf ganz $\mathcal{X}$ ist, sind die folgenden beiden Bedingungen:

a) es existiert ein $M>0$ mit $\|T_k\| \le M$ für alle $k\in\mathbb{N}$

b) $\{T_m\}_{m\in\mathbb{N}}$ ist starke Cauchy-Folge [bzw. stark konvergent] auf einer Fundamentalmenge $A\subseteq\mathcal{X}$.

Beispiel 15 (Fortsetzung von Beispiel 14)**.** Wir wollen mit Satz 24 zeigen, dass die Operatoren $\{T_m\}_{m\in\mathbb{N}}$ aus Beispiel 14 stark gegen **0** konvergieren. Dazu überlegen wir uns zunächst, dass die Menge der Einheitsvektoren $A := \{e_n \mid n\in\mathbb{N}\}$ mit $e_{nk} := \begin{cases}1, & k=n\\ 0, & k\ne n\end{cases}$ für $n,k\in\mathbb{N}$ eine Fundamentalmenge in ℓ^1 bildet. Es ist nämlich $\operatorname{span}(A) = \{f\in\ell^1 \mid \exists n\in\mathbb{N}: f_k = 0 \text{ für alle } k\ge n\}$ mit $\overline{\operatorname{span}(A)} = \ell^1,$ denn zu jedem $\varepsilon>0$ und $f\in\ell^1$ existiert ein $N\in\mathbb{N}$ mit $\sum_{k=N+1}^{\infty}|f_k| < \varepsilon.$ Wähle $h\in\ell^1$ mit $h_k := \begin{cases}f_k, & k\le N\\ 0, & k>N\end{cases}$ für $k\in\mathbb{N}.$ Dann ist $h\in\operatorname{span}(A)$ mit $\|f-h\|_1 < \varepsilon.$ Also ist A fundamental in ℓ^1. Wegen $\|T_k\| = 1$ für alle $k\in\mathbb{N}$ ist Voraussetzung a) von Satz 24 erfüllt. Ferner ist $T_m(e_k) = \mathbf{0}$ für $m>k;$ also ist auch Voraussetzung b) in Satz 24 erfüllt. Damit konvergiert $\{T_m\}_{m\in\mathbb{N}}$ stark gegen **0** auf A und damit also auch auf ganz ℓ^1.

I. 3. Lineare Funktionale

In diesem Abschnitt befassen wir uns genauer mit linearen Operatoren, die Werte in $\mathbb{R}$ annehmen. Solche Operatoren heißen *lineare Funktionale* und spielen eine wichtige Rolle u.a. für die Charakterisierung von Banach-Räumen.

Definition 33 (sublineares Funktional)**.** Es sei $\mathcal{X}$ ein reeller Vektorraum und $p:\mathcal{X}\to\mathbb{R}$ eine Abbildung. p heißt

a) *sublineares Funktional*, wenn gilt:

1.) $p(f+g)\le p(f)+p(g)$ für alle $f,g\in\mathcal{X}$ [Sub-Additivität]

2.) $p(\alpha f)=\alpha\, p(f)$ für alle $f\in\mathcal{X},\ \alpha\ge 0$ [positive Homogenität]

b) *Halbnorm*, wenn statt 2.) in Teil a) gilt:

3.) $p(\alpha f)=|\alpha|\, p(f)$ für alle $f\in\mathcal{X},\ \alpha\in\mathbb{R}$.

Bemerkung: Offensichtlich sind alle Normen und Halbnormen auch sublineare Funktionale. Ferner folgt aus der Definition eines sublinearen Funktionals sofort auch

$$p(\mathbf{0})=0 \quad [\text{für die Wahl } \alpha=0 \text{ in 2.)}] \quad \text{sowie}$$
$$-p(-f)\le p(f) \ \text{ für alle } f\in\mathcal{X} \quad [\text{für die Wahl } f=-g \text{ in 1.)}].$$

Definition 34 (lineares Funktional)**.** Es sei $(\mathcal{X},\|\bullet\|_{\mathcal{X}})$ ein linearer normierter Raum über K und $(\mathcal{Y},\|\bullet\|_{\mathcal{Y}})=(K,|\bullet|)$. Dann heißt ein linearer Operator $T:\mathcal{X}\to\mathcal{Y}$ *lineares Funktional auf* $\mathcal{X}$.

Bemerkung: Üblicherweise – und erst aus weiter unten ersichtlichen Gründen – werden lineare Funktionale mit einem Stern gekennzeichnet, also in der Form $f^*,g^*,\cdots$ geschrieben.
Da $(K,|\bullet|)$ ein Banach-Raum[18] ist, können alle Definitionen und Sätze über lineare Operatoren auf lineare Funktionale angewendet werden, insbesondere Definition 30 und Satz 20. So gilt etwa für ein lineares Funktional f^* auf $\mathcal{X}$:

$$\|f^*\|=\sup_{\|f\|_{\mathcal{X}}=1}|f^*(f)|.$$

Lemma 24. In einem Hilbert-Raum $(\mathcal{X},\|\bullet\|_{\mathcal{X}})$ mit Skalarprodukt $(\bullet,\bullet)$ wird durch jedes Element $g\in\mathcal{X}$ ein beschränktes lineares Funktional g^* mit $\|g^*\|=\|g\|_{\mathcal{X}}$ definiert vermöge

$$g^*(f):=(f,g) \ \text{ für alle } \ f\in\mathcal{X}.$$

Beweis: Die Linearität des Funktionals folgt aus der Sesquilinearität des Skalarprodukts. Mit der Schwarz'schen Ungleichung (Lemma 11) folgt ferner

[18] Zur Erinnerung: für $K=\mathbb{C}$ ist $|a+bi|^2=(a+bi)\cdot\overline{(a+bi)}=a^2+b^2$.

$$\left\|g^*\right\| = \sup_{\|f\|_{\mathcal{X}}=1} \left|g^*(f)\right| = \sup_{\|f\|_{\mathcal{X}}=1} \left|(f,g)\right| \le \sup_{\|f\|_{\mathcal{X}}=1} \left|\|f\|_{\mathcal{X}} \cdot \|g\|_{\mathcal{X}}\right| = \|g\|_{\mathcal{X}}.$$

Für $g \neq \mathbf{0}$ erhält man mit der speziellen Wahl $f = \dfrac{1}{\|g\|_{\mathcal{X}}} g$ auch noch

$$\left|(f,g)\right| = \frac{\|g\|_{\mathcal{X}}^2}{\|g\|_{\mathcal{X}}} = \|g\|_{\mathcal{X}}, \text{ also insgesamt } \left\|g^*\right\| = \|g\|_{\mathcal{X}}. \qquad \blacksquare$$

Die obige Darstellung linearer Funktionale auf Hilbert-Räumen ist damit gerade die Motivation für die allgemeine *-Schreibweise.

Definition 35 (Dualraum)**.** Der Raum $\mathfrak{L}[\mathcal{X},\mathbb{R}]$ aller *beschränkten* linearen Funktionale auf einem linearen normierten Raum $\left(\mathcal{X},\|\bullet\|_{\mathcal{X}}\right)$ mit Werten in $\left(\mathbb{R},|\bullet|\right)$ heißt der *zu $\mathcal{X}$ duale* (konjugierte) Raum und wird mit $\mathcal{X}^*$ bezeichnet.

Lemma 25. Der zu einem linearen normierten Raum $\left(\mathcal{X},\|\bullet\|_{\mathcal{X}}\right)$ duale Raum $\left(\mathcal{X}^*,\|\bullet\|_{\mathcal{X}^*}\right)$ mit der Operatornorm $\|\bullet\|_{\mathcal{X}^*}$ (vgl. die Bemerkung im Anschluss an Definition 34) ist vollständig, also ein Banach-Raum.

Beweis: Unmittelbar klar nach Satz 21. ■

I. 3.1. Fortsetzung linearer Funktionale

In diesem Abschnitt behandeln wir einige der wichtigsten Sätze in Bezug auf lineare Funktionale, die bekannt sind als die *Sätze von Hahn* und *Banach*.

Satz 25 (Hahn-Banach-Fortsetzungssatz I; Banach 1929)**.** Es sei *p* ein sublineares Funktional auf einem reellen Vektorraum $\mathcal{X}$; ferner sei $A \subseteq \mathcal{X}$ ein linearer Teilraum von $\mathcal{X}$ und f^* ein (reelles) lineares Funktional auf *A* mit der Eigenschaft

$$f^*(f) \le p(f) \text{ für alle } f \in A.$$

Dann existiert ein lineares Funktional F^* auf ganz $\mathcal{X}$ (Fortsetzung von f^*) mit den Eigenschaften

$$F^*(f) = f^*(f) \text{ für alle } f \in A \text{ und } F^*(f) \le p(f) \text{ für alle } f \in \mathcal{X}.$$

Beweis: O.B.d.A. können wir $A \neq \mathcal{X}$ annehmen. Wir konstruieren zunächst eine Fortsetzung mit den gewünschten Eigenschaften auf den linearen Teilraum $B := \operatorname{span}\left(A \cup \{g\}\right)$ für ein festes Element $g \in A^c$. Jedes Element $h \in B$ ist dann eindeutig darstellbar als $h = f + \alpha g$ mit geeigneten

$f \in A,\ \alpha \in \mathbb{R},$ denn aus $f + \alpha g = \hat{f} + \hat{\alpha} g$ mit $f, \hat{f} \in A,\ \alpha, \hat{\alpha} \in \mathbb{R}$ folgt $f - \hat{f} = (\hat{\alpha} - \alpha) g$ und damit $\alpha = \hat{\alpha}$ und $f = \hat{f}$ [sonst wäre nämlich $g = \dfrac{1}{\hat{\alpha} - \alpha}\left(f - \hat{f}\right) \in A$: Widerspruch!].

Für $f_1, f_2 \in A$ erhalten wir zunächst

$$f^*(f_1) - f^*(f_2) = f^*(f_1 - f_2) \le p(f_1 - f_2) = p(f_1 + g - f_2 - g) \le p(f_1 + g) + p(-f_2 - g),$$

also

$$-f^*(f_2) - p(-f_2 - g) \le -f^*(f_1) + p(f_1 + g)$$

und somit

$$m := \sup_{f_2 \in A}\left\{-f^*(f_2) - p(-f_2 - g)\right\} \le M := \inf_{f_1 \in A}\left\{-f^*(f_1) + p(f_1 + g)\right\}.$$

Eine der gewünschten Fortsetzungen F_1^* von f^* auf B kann jetzt mit jedem $\xi \in [m, M]$ wie folgt erreicht werden:

$$F_1^*(f + \alpha g) := f^*(f) + \alpha \cdot \xi \text{ für alle } f \in A,\ \alpha \in \mathbb{R}.$$

Nach Wahl von ξ gilt nämlich:

$$\begin{aligned} -f^*(f) - p(-f - g) \le \xi \le -f^*(f) + p(f + g) \text{ bzw.} \\ -p(-f - g) \le f^*(f) + \xi \le p(f + g) \text{ für alle } f \in A, \end{aligned} \qquad (*)$$

also ist F_1^* Fortsetzung von f^* auf B (setze $\alpha = 0$) und linear wegen der Linearität von f^* auf A. Es bleibt zu zeigen, dass $F_1^*(h) \le p(h)$ gilt für alle $h \in B$. Sei dazu $h = f + \alpha g$ mit $\alpha \ne 0$ (sonst ist die Aussage trivial). Wir unterscheiden zwei Fälle:

$\underline{\alpha > 0}$: Hier ist

$$F_1^*(h) = F_1^*(f + \alpha g) = \alpha F_1^*\left(\frac{1}{\alpha} f + g\right) = \alpha\left(f^*\left(\frac{1}{\alpha} f\right) + \xi\right) \overset{(*)}{\le} \alpha\, p\left(\frac{1}{\alpha} f + g\right) = p(f + \alpha g) = p(h).$$

$\underline{\alpha < 0}$: Hier ist

$$F_1^*(h) = F_1^*(f + \alpha g) = \alpha F_1^*\left(\frac{1}{\alpha} f + g\right) = \alpha\left(f^*\left(\frac{1}{\alpha} f\right) + \xi\right) \overset{(*)}{\le} -\alpha\, p\left(-\frac{1}{\alpha} f - g\right) = p(f + \alpha g) = p(h).$$

Damit ist alles bewiesen, sofern schon $B = \mathcal{X}$ ist.

Für den Fall, dass $\mathcal{X}$ endlich-dimensional ist, bricht dieses Verfahren nach endlich vielen Schritten ab, denn man kann bei $B \ne \mathcal{X}$ eine geeignete Fortsetzung F_2^* auf $C := \operatorname{span}\left(B \cup \{k\}\right)$ finden für $k \in B^c$ usw. Für unendlich-dimensionale Vektorräume $\mathcal{X}$ lässt sich dagegen i. Allg. eine vollständige Fotsetzung auf ganz $\mathcal{X}$ nur mit dem Zorn'schen Lemma erreichen, und zwar auf folgende Weise: Bezeichnet $\mathfrak{S}$ die Menge aller linearen Funktionale F^*, die f^* mit den genannten Bedingungen auf eine Menge $D(F^*)$ fortsetzen, so ist wegen $f^* \in \mathfrak{S}$ mit $D(f^*) = A$ die Menge $\mathfrak{S}$ nichtleer. Wir führen auf $\mathfrak{S}$ eine partielle Ordnung $\prec$ ein vermöge

$$F_1^* \prec F_2^* \;:\Leftrightarrow\; D\left(F_2^*\right) \supseteq D\left(F_1^*\right) \text{ und } F_2^*\Big|_{D\left(F_1^*\right)} = F_1^*.$$

Mit Hilfe dieser Festlegung lässt sich zeigen, dass $\mathfrak{S}$ sogar induktiv geordnet ist, d.h. jede total geordnete Teilmenge $\left\{F_i^* \mid i \in I\right\}$ von $\mathfrak{S}$ besitzt eine obere Schranke. Diese ist von der Form G^* derart, dass $D\left(G^*\right) = \bigcup_{i \in I} D\left(F_i^*\right)$ und $G^*(f) = F_j^*(f)$ für $f \in D\left(F_j^*\right) \subseteq D\left(G^*\right)$. Nach dem Zorn'schen Lemma besitzt $\left(\mathfrak{S}, \prec\right)$ damit ein maximales Element H^*, dessen Definitionsbereich $D\left(H^*\right)$ nachweisbar ganz $\mathcal{X}$ ist, und das der gewünschten Fortsetzung entspricht. Für Details dieses Beweisschritts vgl. etwa BACHMAN / NARICI (2000), Theorem 11.1. ■

Satz 26 (Hahn-Banach-Fortsetzungssatz II; Hahn 1927)**.** Es sei A linearer Teilraum eines linearen normierten Raums $\left(\mathcal{X}, \|\bullet\|_{\mathcal{X}}\right)$ und $f^* \in A^*$ ein beschränktes lineares Funktional auf A mit der Norm

$$\left\|f^*\right\|_{A^*} = \sup\left\{\left|f^*(f)\right| \,\middle|\, f \in A, \|f\|_{\mathcal{X}} = 1\right\}.$$

Dann existiert ein beschränktes lineares Funktional $F^* \in \mathcal{X}^*$ mit

$$F^*(f) = f^*(f) \text{ für alle } f \in A \text{ und } \left\|F^*\right\|_{\mathcal{X}^*} = \left\|f^*\right\|_{A^*}.$$

Beweis: Die durch $p(f) := \left\|f^*\right\|_{A^*} \|f\|_{\mathcal{X}}$ für $f \in \mathcal{X}$ definierte Abbildung ist sublinear mit

$$f^*(f) \le \left|f^*(f)\right| \le \left\|f^*\right\|_{A^*} \|f\|_{\mathcal{X}} = p(f) \text{ für alle } f \in A.$$

Nach Satz 25 existiert nun eine Fortsetzung F^* von f^* auf $\mathcal{X}$ mit

$$F^*(f) \le p(f) = \left\|f^*\right\|_{A^*} \|f\|_{\mathcal{X}} \text{ für alle } f \in \mathcal{X}$$

und damit wegen $-F^*(f) = F^*(-f) \le p(-f) = p(f)$ auch

$$\left|F^*(f)\right| \le p(f) = \left\|f^*\right\|_{A^*} \|f\|_{\mathcal{X}} \text{ für alle } f \in \mathcal{X} \text{ und somit}$$

$$\left\|F^*\right\|_{\mathcal{X}^*} \le \left\|f^*\right\|_{A^*} = \sup\left\{\left|f^*(f)\right| \,\middle|\, f \in A, \|f\|_{\mathcal{X}} = 1\right\} \le \sup\left\{\left|F^*(f)\right| \,\middle|\, f \in \mathcal{X}, \|f\|_{\mathcal{X}} = 1\right\} = \left\|F^*\right\|_{\mathcal{X}^*},$$

also $\left\|F^*\right\|_{\mathcal{X}^*} = \left\|f^*\right\|_{A^*}$, was zu zeigen war. ■

Die folgenden Sätze behandeln einige unmittelbare Konsequenzen aus den Hahn-Banach-Fortsetzungssätzen.

Satz 27. Es sei $\left(\mathcal{X}, \|\bullet\|_{\mathcal{X}}\right)$ ein linearer normierter Raum und $\mathbf{0} \neq f_0 \in \mathcal{X}$. Dann existiert ein beschränktes lineares Funktional $F^* \in \mathcal{X}^*$ mit $\left\|F^*\right\|_{\mathcal{X}^*} = 1$ und $F^*\left(f_0\right) = \|f_0\|_{\mathcal{X}}$.

Beweis: Die Menge $A := \operatorname{span}\left\{f_0\right\} = \left\{\alpha f_0 \mid \alpha \in \mathbb{R}\right\}$ ist ein linearer Teilraum von $\mathcal{X}$. Definiere ein beschränktes lineares Funktional $f^* \in \mathcal{X}^*$ durch

$$f^*\left(\alpha f_0\right) := \alpha \|f_0\|_{\mathcal{X}} \text{ für alle } \alpha \in \mathbb{R}.$$

Dann ist

$$\|f^*\|_{A^*} = \sup\left\{\left|f^*(f)\right| \,\middle|\, f \in A, \|f\|_{\mathcal{X}} = 1\right\} = \sup\left\{|\alpha| \|f_0\|_{\mathcal{X}} \,\middle|\, \alpha \in \mathbb{R}, \|\alpha f_0\|_{\mathcal{X}} = 1\right\} = 1,$$

so dass sich die gewünschte Fortsetzung nach Satz 26 sofort ergibt. ■

Lemma 26. Es sei $\left(\mathcal{X}, \|\bullet\|_{\mathcal{X}}\right)$ ein linearer normierter Raum.

a) Ist $\mathcal{X}$ nicht-trivial, so auch $\mathcal{X}^*$.
b) Gilt $f^*\left(f_0\right) = 0$ für alle beschränkten linearen Funktionale $f^* \in \mathcal{X}^*$ für ein $f_0 \in \mathcal{X}$, so ist $f_0 = \mathbf{0}$.

Beweis: Teil a) folgt sofort aus Satz 27, Teil b) ebenso, weil für $f_0 \neq \mathbf{0}$ nach Satz 27 ein beschränktes lineares Funktional $F^* \in \mathcal{X}^*$ mit $\left\|F^*\right\|_{\mathcal{X}^*} = 1$ und $F^*\left(f_0\right) = \|f_0\|_{\mathcal{X}} > 0$ existieren würde: Widerspruch! ■

Satz 28. Es sei $\left(\mathcal{X}, \|\bullet\|_{\mathcal{X}}\right)$ ein linearer normierter Raum und $A \subseteq \mathcal{X}$ ein linearer Teilraum. Hat ein Element $f_0 \in A^c$ einen positiven Abstand zu A, d.h. gilt

$$r := \operatorname{dist}\left(\left\{f_0\right\}, A\right) = \inf_{f \in A} \|f - f_0\|_{\mathcal{X}} > 0,$$

so existiert ein beschränktes lineares Funktional $F^* \in \mathcal{X}^*$ mit den Eigenschaften:

a) $F^*(f) = 0$ für alle $f \in A$;
b) $F^*\left(f_0\right) = 1$;
c) $\left\|F^*\right\|_{\mathcal{X}^*} = \dfrac{1}{r}$.

Beweis: Sei $B := \operatorname{span}\left\{A \cup \left\{f_0\right\}\right\} = \left\{f + \alpha f_0 \mid f \in A, \alpha \in \mathbb{R}\right\}$. Definiere f^* auf B vermöge

$$f^*\left(f + \alpha f_0\right) := \alpha \text{ für alle } f \in A, \alpha \in \mathbb{R}.$$

Dann ist f^* ein lineares Funktional auf B mit $f^*(f) = 0$ für alle $f \in A$ (für $\alpha = 0$) und $f^*\left(f_0\right) = 1$. Ferner gilt:

$$\left\| f^* \right\|_{B^*} = \sup\left\{ |\alpha| \,\middle|\, f \in A, \alpha \in \mathbb{R}, \left\| f + \alpha f_0 \right\|_{\mathcal{X}} = 1 \right\} = \sup\left\{ |\alpha| \,\middle|\, f \in A, \alpha \neq 0, |\alpha| \left\| \frac{1}{\alpha} f + f_0 \right\|_{\mathcal{X}} = 1 \right\}$$

$$= \sup\left\{ |\alpha| \,\middle|\, g \in A, \alpha \neq 0, |\alpha| = \frac{1}{\left\| -g + f_0 \right\|_{\mathcal{X}}} \right\} = \sup_{g \in A} \frac{1}{\left\| g - f_0 \right\|_{\mathcal{X}}} = \frac{1}{\inf\limits_{g \in A} \left\| g - f_0 \right\|_{\mathcal{X}}} = \frac{1}{r}.$$

Die gewünschte Fortsetzung ergibt sich jetzt wieder nach Satz 26. ■

Bemerkung. Der Dualraum eines endlich-dimensionalen linearen normierten Raums ist ebenfalls endlich-dimensional, und beide Dimensionen sind gleich. Ist der Dualraum eines linearen normierten Raums endlich-dimensional, so auch der ursprüngliche linearen normierten Raum, und beide Dimensionen sind gleich [vgl. TAYLOR (1963), Section 1.61]. Ferner sind sämtliche linearen Funktionale auf einem endlich-dimensionalen linearen normierten Raum $\left(\mathcal{X}, \|\bullet\|_{\mathcal{X}}\right)$ beschränkt [vgl. BACHMAN / NARICI (2000), Theorem 11.5]. Dies folgt aus der Beobachtung, dass sich bei gegebener Basis $B = \{g_1, \cdots, g_n\}$ von $\mathcal{X}$ jedes Element $f \in \mathcal{X}$ schreiben lässt als

$$f = \sum_{i=1}^{n} \alpha_i g_i \text{ mit geeigneten } \alpha_1, \cdots, \alpha_n \in \mathbb{R}.$$

Für ein beliebiges lineares Funktional f^* auf $\mathcal{X}$ gilt dann

$$f^*(f) = \sum_{i=1}^{n} \alpha_i f^*(g_i) \text{ mit } \left| f^*(f) \right| \leq \sum_{i=1}^{n} |\alpha_i| \left| f^*(g_i) \right| \leq \| f \|_1 \cdot M \text{ für } f \in \mathcal{X}$$

mit $M := \max\left\{ \left| f^*(g_1) \right|, \cdots, \left| f^*(g_n) \right| \right\}$ und der Norm $\| f \|_1$ wie in der Bemerkung im Anschluss an Beispiel 12. Wegen der Äquivalenz aller Normen in endlich-dimensionalen linearen normierten Räumen folgt damit die Behauptung.

Satz 29. Es sei $\left(\mathcal{X}, \|\bullet\|_{\mathcal{X}}\right)$ ein linearer normierter Raum und der Dualraum $\mathcal{X}^*$ separabel. Dann ist auch $\mathcal{X}$ separabel.

Beweis: Nach Voraussetzung gibt es eine Menge $\left\{ g_k^* \mid k \in \mathbb{N} \right\}$ in $\mathcal{X}^*$, die dicht in $\mathcal{X}^*$ ist. O.B.d.A. kann $\left\| g_k^* \right\|_{\mathcal{X}^*} > 0$ für alle $k \in \mathbb{N}$ angenommen werden (der Nullvektor kann entfallen). Setze $f_k^* := \frac{1}{\left\| g_k^* \right\|_{\mathcal{X}^*}} g_k^*$ für alle $k \in \mathbb{N}$. Dann ist $\left\{ f_k^* \mid k \in \mathbb{N} \right\}$ dicht in der Menge

$$\mathcal{S} := \left\{ f_k^* \in \mathcal{X}^* \,\middle|\, \left\| f_k^* \right\|_{\mathcal{X}^*} = 1 \right\}.$$

Wegen $1 = \left\| f_k^* \right\|_{\mathcal{X}^*} = \sup\limits_{\|f\|_{\mathcal{X}} = 1} \left| f_k^*(f) \right|$ gibt es zu jedem $k \in \mathbb{N}$ ein $f_k \in \mathcal{X}$ mit $\left| f_k^*(f_k) \right| \geq \frac{1}{2}$. Die Menge $\operatorname{span}\{ f_k \mid k \in \mathbb{N} \}$ ist aber dicht in $\mathcal{X}$, d.h. es gilt $A := \overline{\operatorname{span}\{ f_k \mid k \in \mathbb{N} \}} = \mathcal{X}$. Anderenfalls gäbe es wegen der Abgeschlossenheit von A ein $f_0 \in A^c$ mit $r := \operatorname{dist}(\{ f_0 \}, A) > 0$. Nach Satz 28 gäbe es dann ein lineares Funktional $F^* \in \mathcal{X}^*$ mit $F^*(f) = 0$ für alle $f \in A$; $F^*(f_0) = 1$; $\|F^*\|_{\mathcal{X}^*} = \frac{1}{r}$. Das lineare Funktional $g^* := r \cdot F^*$ hat dann die Eigenschaft $\|g^*\|_{\mathcal{X}^*} = 1$, d.h. es ist $g^* \in \mathcal{S}$, mit $g^*(f_0) = r > 0$ und $g^*(f) = 0$ für alle $f \in A$. Damit ergäbe sich aber

$$\frac{1}{2} \leq \left| f_k^*(f_k) \right| = \left| f_k^*(f_k) - \underbrace{g^*(f_k)}_{=0} \right| = \left| \left(f_k^* - g^*\right)(f_k) \right| \leq \left\| f_k^* - g^* \right\|_{\mathcal{X}^*} \underbrace{\| f_k \|_{\mathcal{X}}}_{=1} = \left\| f_k^* - g^* \right\|_{\mathcal{X}^*} \text{ für alle } k \in \mathbb{N}$$

im Widerspruch zur Dichtheit von $\left\{ f_k^* \mid k \in \mathbb{N} \right\}$ in $\mathcal{S}$! Also ist $A = \mathcal{X}$ und damit $\mathcal{X}$ separabel, da mit $\operatorname{span}\left\{ f_k \mid k \in \mathbb{N} \right\}$ auch die abzählbare Menge

$$\tilde{A} := \left\{ \sum_{i \in I} \alpha_i f_i \,\middle|\, I \subset \mathbb{N} \text{ endlich, } \forall i \in I : \alpha_i \in \mathbb{Q} \right\}$$

dicht in $\mathcal{X}$ ist. ■

Die Sätze von Hahn-Banach haben neben den oben angedeuteten Anwendungen in der Funktionalanalysis auch eine zentrale Bedeutung in der Maßtheorie. Der folgende Satz 30, den wir hier ohne Beweis nur zitieren [vgl. ROYDEN (1988), Chapter 10, Section 3, Proposition 5 oder BACHMAN / NARICI (2000), Chapter 11, Appendix, insbesondere Theorem A.2], zeigt dies im Besonderen auf.

Definition 36 (Halbgruppe von Operatoren)**.** Es sei $\mathcal{X}$ ein Vektorraum und $\mathfrak{G}$ eine Familie von linearen Operatoren von $\mathcal{X}$ nach $\mathcal{X}$. $\mathfrak{G}$ heißt *Halbgruppe*, wenn folgende Eigenschaften erfüllt sind:

a) $\mathbf{I} \in \mathfrak{G}$
b) $S \circ T \in \mathfrak{G}$ für alle $S, T \in \mathfrak{G}$.

Die Halbgruppe $\mathfrak{G}$ heißt *abelsch*, wenn zusätzlich noch gilt:

c) $S \circ T = T \circ S$ für alle $S, T \in \mathfrak{G}$.

Beispiel 16 (Fortsetzung von Beispiel 9; vgl. auch Beispiele 13 und 14)**.** Es sei wieder $(\mathcal{X}, \rho_{\mathcal{X}}) = \ell^p$ mit $1 \leq p \leq \infty$. Dann bilden die Shift-Operatoren T_k mit

$$T_k\left(\{f_n\}_{n \in \mathbb{N}}\right) := \{f_{n+k}\}_{n \in \mathbb{N}} \text{ für alle } f = \{f_n\}_{n \in \mathbb{N}} \in \ell^p \text{ und } k \in \mathbb{Z}^+$$

offensichtlich eine abelsche Halbgruppe $\mathfrak{G} = \left\{ T_k \mid k \in \mathbb{Z}^+ \right\}$ mit $T_0 = \mathbf{I}$ und $T_k \circ T_m = T_{k+m}$ für $k, m \in \mathbb{Z}^+$, ebenso die Shift-Operatoren $\mathfrak{H} = \left\{ S_k \mid k \in \mathbb{Z}^+ \right\}$ mit

$$S_k\left(\{f_n\}_{n \in \mathbb{N}}\right) := \left\{ \underbrace{0, \cdots, 0}_{k\text{-mal}}, f_n \right\}_{n \in \mathbb{N}} \text{ für alle } f = \{f_n\}_{n \in \mathbb{N}} \in \ell^p \text{ und } k \in \mathbb{Z}^+$$

mit $S_0 = \mathbf{I}$ und $S_k \circ S_m = S_{k+m}$ für $k, m \in \mathbb{Z}^+$.

Satz 30 (Hahn-Banach-Fortsetzungssatz III)**.** Es sei p ein sublineares Funktional auf einem reellen Vektorraum $\mathcal{X}$; ferner sei $\mathfrak{G}$ eine abelsche Halbgruppe von Operatoren auf $\mathcal{X}$, $A \subseteq \mathcal{X}$ ein linearer Teilraum von $\mathcal{X}$ und f^* ein (reelles) lineares Funktional auf A. Es gelten die Eigenschaften

a) $p(Tf) \leq p(f)$ für alle $f \in \mathcal{X}$ und $T \in \mathfrak{G}$

b) $T(f) \in A$ für alle $f \in A$

c) $f^*(Tf) = f^*(f)$ für alle $f \in A$ und $T \in \mathfrak{G}$

d) $f^*(f) \leq p(f)$ für alle $f \in A$.

Dann existiert ein lineares Funktional F^* auf ganz $\mathcal{X}$ (Fortsetzung von f^*) mit den Eigenschaften

e) $F^*(f) = f^*(f)$ für alle $f \in A$

f) $F^*(f) \leq p(f)$ für alle $f \in \mathcal{X}$

g) $F^*(Tf) = F^*(f)$ für alle $f \in \mathcal{X}$ und $T \in \mathfrak{G}$.

Die beiden folgenden Ergebnisse, die wir auch nur ohne Beweis zitieren, beruhen auf Satz 30 [vgl. ROYDEN (1988), Chapter 10, Section 3, Problems 20 und 21 oder BACHMAN / NARICI (2000), Chapter 11, Appendix].

Lemma 27 (Banach-Limes)**.** Auf dem Banach-Raum ℓ^∞ der beschränkten Zahlenfolgen kann ein beschränktes lineares Funktional $F^* =: \mathrm{Lim}$ definiert werden mit folgenden Eigenschaften:

a) $\liminf\limits_{n\to\infty} f_n \leq F^*\left(\{f_n\}_{n\in\mathbb{N}}\right) \leq \limsup\limits_{n\to\infty} f_n$ für alle Folgen $\{f_n\}_{n\in\mathbb{N}} \in \ell^\infty$;

b) $F^*\left(\{f_n\}_{n\in\mathbb{N}}\right) \leq F^*\left(\{g_n\}_{n\in\mathbb{N}}\right)$, falls $f_n \leq g_n$ für alle $n \in \mathbb{N}$, für alle $\{f_n\}_{n\in\mathbb{N}}, \{g_n\}_{n\in\mathbb{N}} \in \ell^\infty$;

c) $F^*\left(\{f_n\}_{n\in\mathbb{N}}\right) = F^*\left(\{f_{n+k}\}_{n\in\mathbb{N}}\right)$ für alle Folgen $\{f_n\}_{n\in\mathbb{N}} \in \ell^\infty$ und alle $k \in \mathbb{N}$.

Jedes solche (i. Allg. nicht eindeutige) Funktional F^* heißt *Banach-Limes* der Folge $\{f_n\}_{n\in\mathbb{N}} \in \ell^\infty$.

Bemerkung: Die Idees des Beweises beruht auf einer Fortsetzung des üblichen "Limes" als beschränktem linearem Funktional f^* auf dem Teilraum $c \subset \ell^\infty$ der *konvergenten* Zahlenfolgen auf ganz ℓ^∞, wobei die in Satz 30 betrachtete Halbgruppe hier konkret die (erste) Shift-Halbgruppe aus Beispiel 16 ist. Beziehung a) besagt ja gerade, dass $F^*\left(\{f_n\}_{n\in\mathbb{N}}\right) = f^*\left(\{f_n\}_{n\in\mathbb{N}}\right) = \lim\limits_{n\to\infty} f_n$ für $\{f_n\}_{n\in\mathbb{N}} \in c$ ist. Die Beziehung c) von Lemma 27 spiegelt dabei die Vorstellung wider, dass ein verallgemeinerter Grenzwert-Begriff nicht von den ersten endlich-vielen Gliedern der Folge abhängen sollte; Beziehung b) beschreibt die (wünschenswerte) Monotonie-Eigenschaft des Banach-Limes.

Lemma 28 (Translationsinvariante additive Fortsetzung des Lebesgue-Maßes)**.** Auf der Menge $\mathfrak{B}^d := \{ A \subset \mathbb{R}^d \mid A \text{ beschränkt} \}$ (mit $d \in \mathbb{N}$) kann ein endlich-additiver Inhalt μ^d definiert werden[19] mit folgenden Eigenschaften:

a) $\mu^d(T_t(A)) = \mu^d(A)$ für alle $A \in \mathfrak{B}^d$ mit $T_t(f) := f + t$ für alle $f, t \in \mathbb{R}^d$;

b) $\mu^d(B) = \mathfrak{m}^d(B)$ für alle beschränkten Borel-Mengen $B \in \mathcal{B}^d$, wobei $\mathfrak{m}^d$ das Lebesgue-Maß auf der Borel'schen σ-Algebra $\mathcal{B}^d$ bezeichne.

Bemerkung: Die Idee des Beweises beruht auf einer Fortsetzung $F^* =: (\mathrm{B})\int \bullet \, d\mu^d$ des Lebesgue-Integrals $\int \bullet \, d\mathfrak{m}^d$ als beschränktem linearem Funktional f^* auf dem Teilraum

$$\mathfrak{F}^0(\mathbb{R}^d, \mathcal{B}^d) := \left\{ f : \mathbb{R}^d \to \mathbb{R} \mid f \text{ Borel-messbar, beschränkt und } f\big|_{A^c} = 0 \text{ für ein } A \in \mathfrak{B}^d \right\}$$

von $\mathfrak{L}^1(\mathbb{R}^d, \mathcal{B}^d, \mathfrak{m}^d)$ mit der Eigenschaft $(\mathrm{B})\int \mathbb{1}_A \, d\mu^d = \mu^d(A)$ für $A \in \mathfrak{B}^d$. Die in Satz 30 benötigte Halbgruppe ist hier die obige Halbgruppe $\{T_t \mid t \in \mathbb{R}\}$ mit $T_0 = \mathbf{I}$.

Es ist jedoch zu bemerken, dass sowohl in Lemma 27 als auch in Lemma 28 das Zorn'sche Lemma eine wesentliche Anwendung findet, die dort beschriebenen Fortsetzungen also nicht-konstruktiv sind.

I. 3.2. Darstellungssätze für beschränkte lineare Funktionale

In diesem Abschnitt wollen wir insbesondere die Dualräume der allgemeinen Lebesgue-Räume aus Abschnitt I.1.5 und die Dualräume von Räumen stetiger Funktionen auf geeigneten Mengen genauer untersuchen. Dazu benötigen wir vorab einige Hilfsresultate.

Lemma 29. Es sei $1 \leq p \leq \infty$ und q der zu p konjugierte Index, also entweder $q = \frac{p}{p-1}$ für $1 < p < \infty$ oder $q = \infty$ für $p = 1$ oder $q = 1$ für $p = \infty$. Ferner sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum mit nicht-trivialem Maß μ. Dann gilt: jedes $g \in L^q(\Omega, \mathcal{A}, \mu)$ induziert ein beschränktes lineares Funktional g^* auf $L^p(\Omega, \mathcal{A}, \mu)$ vermöge

$$g^*(f) := \int fg \, d\mu \text{ für alle } f \in L^p(\Omega, \mathcal{A}, \mu) \text{ mit } \|g^*\| \leq \|g\|_q$$

und Gleichheit für $p > 1$ (ohne weitere Einschränkung) oder $p = 1$, wenn zugleich μ σ-endlich ist.

---

[19] D.h. μ^d erfüllt folgende Eigenschaften: $\mu^d \geq 0$ mit $\mu^d(\varnothing) = 0$, und $\mu^d(A \oplus B) = \mu^d(A) + \mu^d(B)$ für alle paarweise disjunkten Mengen $A, B \in \mathfrak{B}$ (endliche Additivität).

Beweis: Die definierte Abbildung ist offensichtlich linear, mit $f \cdot g \in L^1(\Omega, \mathcal{A}, \mu)$ und

$$\left|g^*(f)\right| = \left|\int fg\,d\mu\right| \le \int |fg|\,d\mu = \|fg\|_1 \le \|f\|_p \|g\|_q \text{ für alle } f \in L^p(\Omega, \mathcal{A}, \mu)$$

gemäß den Hölder-Ungleichungen (Lemmata 14 und 15), woraus zunächst $\|g^*\| \le \|g\|_q$ folgt. Zum Nachweis der Gleichheit betrachten wir zunächst den Fall $1 < p < \infty$ sowie die Abbildung

$$f := \operatorname{sgn}(g) \cdot |g|^{q/p} \in L^p(\Omega, \mathcal{A}, \mu)$$

wegen $|f|^p = |g|^q$. Damit ist $f \cdot g = |g| \cdot |g|^{q/p} = |g|^{1+q/p} = |g|^q$, also

$$g^*(f) := \int fg\,d\mu = \int |g|^q\,d\mu = \|g\|_q^q = \|g\|_q^{q-1}\|g\|_q = \|g\|_q^{q/p}\|g\|_q = \|f\|_p\|g\|_q,$$

womit $\|g^*\| = \|g\|_q$ gezeigt ist. Für $p = \infty$ argumentiert man völlig analog mit der Wahl

$$f := \operatorname{sgn}(g) \in L^\infty(\Omega, \mathcal{A}, \mu) \text{ mit } \|f\|_\infty = 1 \text{ für } \mu(\{g \ne \mathbf{0}\}) > 0$$

und $g^*(f) := \int fg\,d\mu = \int |g|\,d\mu = \|g\|_1 = \|f\|_\infty \|g\|_1$. Sei schließlich $p = 1$ und μ zunächst endlich. Für $\varepsilon > 0$ sei $A_\varepsilon := \{g > \|g\|_\infty - \varepsilon\}$ und $f := \mathbb{1}_{A_\varepsilon}$. Dann ist

$$g^*(f) := \int fg\,d\mu = \int_{A_\varepsilon} g\,d\mu \ge \left(\|g\|_\infty - \varepsilon\right)\mu(A_\varepsilon) = \left(\|g\|_\infty - \varepsilon\right)\|f\|_1,$$

also $\|g^*\| \ge \|g\|_\infty - \varepsilon$ für jedes $\varepsilon > 0$ und damit $\|g^*\| = \|g\|_\infty$. Ist μ σ-endlich, so existiert eine disjunkte Zerlegung $\Omega = \bigoplus_{i \in I} E_i$ mit abzählbarer Indexmenge I und $0 < \mu(E_i) < \infty$ für alle $i \in I$. Wegen $0 < \mu(A_\varepsilon) = \sum_{i \in I} \mu(A_\varepsilon \cap E_i)$ existiert ein $k \in I$ mit $\mu(A_\varepsilon \cap E_k) > 0$. Insbesondere ist also $A_\varepsilon \cap E_k \ne \varnothing$ und Teilmenge von A_ε. Mit $f := \mathbb{1}_{A_\varepsilon \cap E_k}$ folgt nun wie zuvor:

$$g^*(f) := \int fg\,d\mu = \int_{A_\varepsilon \cap E_k} g\,d\mu \ge \left(\|g\|_\infty - \varepsilon\right)\mu(A_\varepsilon \cap E_k) = \left(\|g\|_\infty - \varepsilon\right)\|f\|_1,$$

also auch hier schließlich $\|g^*\| = \|g\|_\infty$. Damit ist das Lemma bewiesen. ■

Unser Ziel wird es sein, zu zeigen, dass unter der Voraussetzung $1 < p < \infty$ der Dualraum $\mathfrak{L}^p(\Omega, \mathcal{A}, \mu)^*$ in geeigneter Weise mit ganz $\mathfrak{L}^q(\Omega, \mathcal{A}, \mu)$ identifiziert werden kann. Diese Aussage ist als einer der *Riesz'schen Darstellungssätze* bekannt. [Für $p = 1$ und $p = \infty$ gelten andere bzw. nicht so allgemeine Aussagen.]

Lemma 30 ("Littlewood's second principle")**.** Es sei $1 \le p < \infty$ und $f \in L^p(\Omega, \mathcal{A}, \mu)$. Dann gibt es zu jedem $\varepsilon > 0$ eine Elementarfunktion[20] $\varphi \in L^p(\Omega, \mathcal{A}, \mu)$, die außerhalb einer Menge von endlichem Maß verschwindet und für die $\|f - \varphi\|_p < \varepsilon$ ist.

[20] D.h. eine messbare Funktion, die nur endlich viele Werte annimmt.

Beweis: [siehe auch TAYLOR (1963), Lemmata 7.3-B, 7.3-C und 7.3-D] Für $n \in \mathbb{N}$ setzen wir $A_n := \left\{ |f^p| > \frac{1}{n} \right\} \in \mathcal{A}$. Wegen

$$\frac{1}{n}\mu(A_n) \leq \int_{A_n} |f^p| \, d\mu \leq \int |f^p| \, d\mu = \|f\|_p^p < \infty \text{ für alle } n \in \mathbb{N}$$

haben alle Mengen A_n endliches Maß, mit $\mu(A_n) \leq n \|f\|_p^p$ für alle $n \in \mathbb{N}$. Setze $f_n := f \cdot \mathbb{1}_{A_n}$ für alle $n \in \mathbb{N}$. Dann gilt

$$\|f - f_n\|_p^p = \int_{A_n^c} |f| \, d\mu \quad \text{mit } A := \lim_{n\to\infty} A_n^c = \bigcap_{n=1}^{\infty} A_n^c = \left(\bigcup_{n=1}^{\infty} A_n \right)^c = \{|f| = 0\},$$

also

$$\lim_{n\to\infty} \|f - f_n\|_p^p = \int_A |f| \, d\mu = 0,$$

so dass $f = \lim_{n\to\infty} f_n$ in $L^p(\Omega, \mathcal{A}, \mu)$. Für jedes $\varepsilon > 0$ existiert also (zunächst) ein $g \in L^p(\Omega, \mathcal{A}, \mu)$, das außerhalb einer Menge *B* von endlichem Maß verschwindet und für das $\|f - g\|_p < \varepsilon$ ist. Für solche *g* definieren wir weiter $B_n := \{|g| > n\}$ sowie $g_n := g \cdot \mathbb{1}_{B_n^c} + n \cdot \operatorname{sgn}(g) \cdot \mathbb{1}_{B_n}$ für $n \in \mathbb{N}$. Es folgt

$$n^p \mu(B_n) \leq \int_{B_n} |g|^p \, d\mu \leq \int |g|^p \, d\mu < \infty,$$

d.h. es gilt $\lim_{n\to\infty} \mu(B_n) = 0$. Damit erhält man mit der Minkowski-Ungleichung (Lemma 16)

$$\int |g - g_n|^p \, d\mu = \int_{B_n} |g - g_n|^p \, d\mu \leq \int_{B_n} (|g| + |g_n|)^p \, d\mu \leq 2^p \int_{B_n} |g|^p \, d\mu,$$

so dass $g = \lim_{n\to\infty} g_n$ in $L^p(\Omega, \mathcal{A}, \mu)$. Für jedes $\varepsilon > 0$ existiert also sogar ein (überall) beschränktes $h \in L^p(\Omega, \mathcal{A}, \mu)$, das außerhalb einer Menge *B* von endlichem Maß verschwindet, und für das $\|f - h\|_p < 2\varepsilon$ ist. Für solche *h* mit $|h| \leq M < \infty$ definieren wir weiter $a_0 := -M$ sowie $a_k := a_0 + \frac{2kM}{n}$ für $k = 1, \cdots, n$ mit den Intervallen $C_1 := [a_0, a_1]$ und $C_k := (a_{k-1}, a_k]$ sowie die messbare Abbildung $h_n := \sum_{k=1}^{n} a_k \cdot \mathbb{1}_{C_k}(h) \cdot \mathbb{1}_B$. Dann ist h_n eine Elementarfunktion mit der Eigenschaft $\sup_{\omega \in \Omega} |h(\omega) - h_n(\omega)| \leq \frac{2M}{n}$. Es folgt

$$\int |h - h_n|^p \, d\mu = \int_B |h - h_n|^p \, d\mu \leq \left(\frac{2M}{n} \right)^p \mu(B),$$

so dass $h = \lim_{n\to\infty} h_n$ in $L^p(\Omega, \mathcal{A}, \mu)$. Für jedes $\varepsilon > 0$ existiert also insgesamt eine Elementarfunktion $\varphi \in L^p(\Omega, \mathcal{A}, \mu)$, die außerhalb einer Menge von endlichem Maß verschwindet und für die $\|f - \varphi\|_p < 3\varepsilon$ ist. Damit ist die Aussage bewiesen. ■

Lemma 31. Es sei $1 \le p < \infty$ und q der zugehörige konjugierte Index. Ferner sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum mit endlichem Maß μ. Dann gilt: ist $g \in L^1(\Omega, \mathcal{A}, \mu)$ und $M < \infty$ mit der Eigenschaft

$$\left|\int g\varphi \, d\mu\right| \le M \|\varphi\|_p \text{ für alle Elementarfunktionen } \varphi,$$

so ist $g \in L^q(\Omega, \mathcal{A}, \mu)$.

Beweis: Auf Grund der Endlichkeit von μ ist jedenfalls $\|\varphi\|_p = \int |\varphi|^p \, d\mu < \infty$ für jede Elementarfunktion φ. Sei nun zunächst $p > 1$ und $\{\zeta_n\}_{n \in \mathbb{N}}$ eine monoton wachsende Folge nicht-negativer Elementarfunktionen mit $\sup\limits_{n \in \mathbb{N}} \zeta_n = |g|^q$. Setze ferner $\varphi_n := \zeta_n^{1/p} \operatorname{sgn} g$ für alle $n \in \mathbb{N}$. Dann sind auch alle φ_n Elementarfunktionen mit

$$\|\varphi_n\|_p = \left\{\int \zeta_n \, d\mu\right\}^{1/p} \text{ für alle } n \in \mathbb{N}.$$

Wegen

$$\varphi_n g = \zeta_n^{1/p} |g| \ge \zeta_n^{1/p} \zeta_n^{1/q} = \zeta_n^{1/p + 1/q} = \zeta_n$$

folgt

$$\int \zeta_n \, d\mu \le \int \varphi_n g \, d\mu \le M \|\varphi_n\|_p = M \cdot \left\{\int \zeta_n \, d\mu\right\}^{1/p}, \text{ also } \left\{\int \zeta_n \, d\mu\right\}^{1/q} \le M \text{ bzw. } \int \zeta_n \, d\mu \le M^q$$

für alle $n \in \mathbb{N}$. Wegen

$$\int |g|^q \, d\mu = \sup_{n \in \mathbb{N}} \int \zeta_n \, d\mu \le M^q$$

nach dem *Satz von der monotonen Konvergenz* (vgl. das Skript STOCHASTIK, Satz 18) folgt also $g \in L^q(\Omega, \mathcal{A}, \mu)$, wie behauptet. Sei schließlich $p = 1$. Für $n \in \mathbb{N}$ sei $A_n := \{|g| > n\}$ und $\varphi_n := \mathbb{1}_{A_n} \cdot \operatorname{sgn} g$. Dann sind wieder alle φ_n Elementarfunktionen mit

$$\|\varphi_n\|_1 = \int |\varphi_n| \, d\mu = \int \mathbb{1}_{A_n} \, d\mu = \mu(A_n) \text{ für alle } n \in \mathbb{N}.$$

Es folgt

$$n \cdot \mu(A_n) \le \int_{A_n} |g| \, d\mu = \int \varphi_n g \, d\mu \le M \|\varphi_n\|_1 = M \cdot \mu(A_n) \text{ für alle } n \in \mathbb{N}$$

und somit $\mu(A_n) = 0$ für $n > M$, also $\|g\|_\infty \le M + 1$, also $g \in L^\infty(\Omega, \mathcal{A}, \mu)$, wie behauptet. ■

Lemma 32. Es sei $1 \le p < \infty$ und $(\Omega, \mathcal{A}, \mu)$ ein Maßraum sowie $\{E_n\}_{n \in \mathbb{N}} \subseteq \mathcal{A}$ eine Folge paarweise disjunkter messbarer Mengen. Sind dann $f_n \in L^p(\Omega, \mathcal{A}, \mu)$ mit $f_n\big|_{E_n^c} = 0$ und ist $f := \sum\limits_{n=1}^{\infty} f_n$, so gilt $f \in L^p(\Omega, \mathcal{A}, \mu)$ genau dann, wenn $\sum\limits_{n=1}^{\infty} \|f_n\|_p^p < \infty$ ist. In diesem Fall ist die Reihe $\sum\limits_{n=1}^{\infty} f_n$ in $L^p(\Omega, \mathcal{A}, \mu)$ konvergent, d.h. es gilt $\lim\limits_{m \to \infty} \left\| f - \sum\limits_{n=1}^{m} f_n \right\|_p = 0$ mit $\|f\|_p^p = \sum\limits_{n=1}^{\infty} \|f_n\|_p^p < \infty$.

Beweis: Dies folgt sofort aus der Beziehung

$$\int |f|^p \, d\mu = \sum_{n=1}^{\infty} \int_{E_n} |f|^p \, d\mu = \sum_{n=1}^{\infty} \int_{E_n} |f_n|^p \, d\mu = \sum_{n=1}^{\infty} \|f_n\|_p^p$$

unter Verwendung des *Satzes von der majorisierten Konvergenz* (vgl. das Skript STOCHASTIK, Satz 19). ■

Satz 31 (Riesz'scher Darstellungssatz I)**.** Es sei $1 \le p < \infty$ und q der zu p konjugierte Index. Ferner sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum mit nicht-trivialem σ-endlichem Maß μ. Dann gilt: zu jedem beschränkten linearen Funktional $\tilde{g}^* \in \mathfrak{L}^p(\Omega, \mathcal{A}, \mu)^*$ existiert genau ein $\tilde{g} \in \mathfrak{L}^q(\Omega, \mathcal{A}, \mu)$ [21] mit

$$\tilde{g}^*\left(\tilde{f}\right) = \int fg \, d\mu \text{ für alle } \tilde{f} \in \mathfrak{L}^p(\Omega, \mathcal{A}, \mu) \text{ mit } \|g^*\| = \|g\|_q.$$

Damit existiert also ein isometrischer Isomorphismus zwischen $\mathfrak{L}^p(\Omega, \mathcal{A}, \mu)^*$ und $\mathfrak{L}^q(\Omega, \mathcal{A}, \mu)$. In diesem Sinn kann der Dualraum $\mathfrak{L}^p(\Omega, \mathcal{A}, \mu)^*$ mit $\mathfrak{L}^q(\Omega, \mathcal{A}, \mu)$ identifiziert werden.

Beweis: In Lemma 29 wurde bereits gezeigt, dass jedes $g \in L^q(\Omega, \mathcal{A}, \mu)$ ein $\tilde{g} \in \mathfrak{L}^q(\Omega, \mathcal{A}, \mu)$ induziert mit der obigen Eigenschaft. Es bleibt zu zeigen, dass sich auf diese Weise *alle* beschränkten linearen Funktionale ergeben. Dazu nehmen wir zunächst wieder an, dass das Maß μ endlich ist. Dann sind alle beschränkten messbaren Abbildungen Elemente von $L^p(\Omega, \mathcal{A}, \mu)$. Sei nun $\tilde{g}^* \in \mathfrak{L}^p(\Omega, \mathcal{A}, \mu)^*$ bzw. $g^* \in L^p(\Omega, \mathcal{A}, \mu)^*$. Wir setzen

$$\nu(A) := g^*\left(1\!\!1_A\right) \text{ für alle } A \in \mathcal{A}.$$

Ist $\{E_n\}_{n \in \mathbb{N}} \subseteq \mathcal{A}$ eine Folge paarweise disjunkter messbarer Mengen, setze

$$f := \sum_{n=1}^{\infty} \operatorname{sgn}\left(g^*\left(1\!\!1_{E_n}\right)\right) \cdot 1\!\!1_{E_n} = \sum_{n=1}^{\infty} \operatorname{sgn}\left(\nu\left(E_n\right)\right) \cdot 1\!\!1_{E_n}.$$

Nach Lemma 32 ist dann $f \in L^p(\Omega, \mathcal{A}, \mu)$ mit

$$\sum_{n=1}^{\infty} \left|\nu\left(E_n\right)\right| = g^*(f) \le \|g^*\| \cdot \|f\|_p < \infty \text{ und}$$

$$\sum_{n=1}^{\infty} \nu\left(E_n\right) = g^*\left(1\!\!1_E\right) = \nu(E) \text{ mit } E = \bigoplus_{n=1}^{\infty} E_n.$$

Also ist ν ein endliches signiertes Maß mit der aus seiner Definition folgenden Eigenschaft, dass sowohl der Positivteil ν^+ als auch der Negativteil ν^- von μ dominiert werden. Dies folgt aus der Beobachtung, dass für jede μ-Nullmenge $N \in \mathcal{A}$ $1\!\!1_N = \mathbf{0}$ μ-fast überall ist und somit auch $\nu(N) = g^*\left(1\!\!1_N\right) = g^*(\mathbf{0}) = 0$ gilt. Ist insbesondere A^+ eine der möglichen positiven Trägermengen für ν^+ nach dem Hahn'schen Zerlegungssatz (Satz 10), so zeigt dies, dass speziell jede μ-

[21] Die Klassen $\tilde{f}$ und $\tilde{g}$ usw. beziehen sich auf die früher behandelte Äquivalenzrelation der μ-fast-überall-Gleichheit, siehe Abschnitt I.1.5.

Nullmenge $N = \left(N \cap A^+\right) \oplus \left(N \cap \left(A^+\right)^c\right)$ auch ν^+-Nullmenge ist, also μ ν^+ dominiert; analog für ν^-. Nach dem Satz von Radon-Nikodym (vgl. das Skript STOCHASTIK, Satz 21) existieren also nicht-negative messbare Abbildungen g^+ und g^- derart, dass

$$\nu^+(A) = \int_A g^+ \, d\mu \text{ und } \nu^-(A) = \int_A g^- \, d\mu \text{ und damit } \nu(A) = \nu^+(A) - \nu^-(A) = \int_A \left(g^+ - g^-\right) d\mu$$

für alle $A \in \mathcal{A}$. Setze $g := g^+ - g^-$. Dann gilt also

$$\nu(A) = \int_A g \, d\mu \text{ für alle } A \in \mathcal{A}.$$

Wegen der Endlichkeit von ν ist also insbesondere g μ-integrierbar. Für jede Elementarfunktion φ folgt nun aus der Linearität von g^*

$$g^*(\varphi) = \int g\varphi \, d\mu \text{ mit } \left|g^*(\varphi)\right| = \left|\int g\varphi \, d\mu\right| \le \left\|g^*\right\| \left\|\varphi\right\|_p.$$

Nach Lemma ist also $g \in L^q\left(\Omega, \mathcal{A}, \mu\right)$. Definiere hiermit das beschränkte lineare Funktional

$$G^*(f) := \int fg \, d\mu \text{ für alle } f \in L^p\left(\Omega, \mathcal{A}, \mu\right).$$

Dann ist auch $G^* - g^*$ ein beschränktes lineares Funktional, das auf dem Teilraum der Elementarfunktionen verschwindet. Dieser ist aber nach Lemma 30 dicht in $L^p\left(\Omega, \mathcal{A}, \mu\right)$, d.h. es folgt $G^* - g^* = 0$ auf ganz $L^p\left(\Omega, \mathcal{A}, \mu\right)$. Also ist das ursprüngliche beschränkte lineare Funktional g^* von der gewünschten Form. Ist h^* ein weiteres solches Funktional, d.h. gilt

$$g^*(f) = \int fg \, d\mu = \int fh \, d\mu = h^*(f) \text{ für alle } f \in L^p\left(\Omega, \mathcal{A}, \mu\right)$$

mit $h \in L^q\left(\Omega, \mathcal{A}, \mu\right)$, so ist offensichtlich $g^* - h^*$ das Nullfunktional auf $L^p\left(\Omega, \mathcal{A}, \mu\right)$, d.h. es gilt $\left\|g^* - h^*\right\| = 0$ und damit $g^* = h^*$ μ-fast überall. Also ist $\tilde{g}^* = \tilde{h}^*$, was die Aussage für endliche Maße μ zeigt.

Sei nun μ σ-endlich. Dann existiert eine monoton wachsende Mengenfolge $\left\{\Omega_n\right\}_{n \in \mathbb{N}}$ mit $\Omega = \bigcup_{n=1}^{\infty} \Omega_n$ und $\mu\left(\Omega_n\right) < \infty$ für alle $n \in \mathbb{N}$. Für ein $f \in L^p\left(\Omega, \mathcal{A}, \mu\right)$ mit $f\big|_{\Omega_n^c} = 0$ liefert das obige Resultat nun für jedes $n \in \mathbb{N}$ ein $g_n \in L^q\left(\Omega, \mathcal{A}, \mu\right)$ mit $g_n\big|_{\Omega_n^c} = 0$ und

$$g^*(f) = \int f g_n \, d\mu \text{ mit } \left\|g_n\right\|_q \le \left\|g^*\right\|.$$

Wegen der Monotonie der Mengenfolge $\left\{\Omega_n\right\}_{n \in \mathbb{N}}$ und der fast überall eindeutigen Bestimmtheit der g_n können wir O.B.d.A. annehmen, dass $g_{n+1}\big|_{\Omega_n} = g_n$ ist. Definiere nun die Abbildung g durch

$$g(\omega) := \sum_{n=1}^{\infty} g_n(\omega) \cdot \mathbb{1}_{\Omega_n}(\omega) \text{ für } \omega \in \Omega.$$

Nach Konstruktion ist diese Abbildung wohldefiniert mit $|g| = \sup_{n\in\mathbb{N}} |g_n|$. Nach dem schon benutzen *Satz von der monotonen Konvergenz* folgt

$$\int |g|^q \, d\mu = \lim_{n\to\infty} \int |g_n|^q \, d\mu \le \|g^*\| \text{ und somit } g \in L^q(\Omega, \mathcal{A}, \mu).$$

Für beliebiges $f \in L^p(\Omega, \mathcal{A}, \mu)$ setzen wir $f_n := f \cdot \mathbb{1}_{\Omega_n}$ für alle $n \in \mathbb{N}$. Dann konvergiert $\{f_n\}_{n\in\mathbb{N}}$ punktweise gegen *f*. Wegen $g \in L^q(\Omega, \mathcal{A}, \mu)$ ist aber $|fg| \in L^1(\Omega, \mathcal{A}, \mu)$ mit $|f_n g| \le |fg|$ für alle $n \in \mathbb{N}$. Nach dem ebenfalls schon benutzen *Satz von der majorisierten Konvergenz* ergibt sich demnach schließlich

$$\int fg \, d\mu = \lim_{n\to\infty} \int f_n g \, d\mu = \lim_{n\to\infty} \int f_n g_n \, d\mu = \lim_{n\to\infty} g^*(f_n) = g^*(f),$$

womit der Satz für den σ-endlichen Fall bewiesen ist.

Für den allgemeinen Fall (und $p > 1$) kann man noch so argumentieren: wir nennen eine messbare Menge $A \in \mathcal{A}$ *von σ-endlichem Maß*, wenn eine disjunkte Zerlegung von $A = \bigoplus_{n=1}^{\infty} A_n$ in messbare Mengen A_n mit $\mu(A_n) < \infty$ existiert. Für jede solche Menge *A* existiert nach obigem ein μ-fast überall eindeutiges $g_A \in L^q(\Omega, \mathcal{A}, \mu)$ mit $g_A|_{A^c} = 0$ und

$$g^*(f) = \int fg_A \, d\mu \text{ für alle } f \in L^p(\Omega, \mathcal{A}, \mu) \text{ mit } f|_{A^c} = 0.$$

Fern gilt $g_B|_A = g_A$ μ-fast überall für messbare Mengen *A* und *B* mit $A \subseteq B$. Für Mengen $A \in \mathcal{A}$ von σ-endlichem Maß setze $\lambda(A) := \int |g_A|^q \, d\mu$. Dann gilt auch $\lambda(A) \le \lambda(B) \le \|g^*\|$ für $A \subseteq B$. Wähle nun eine Folge $\{A_n\}_{n\in\mathbb{N}}$ von Mengen von σ-endlichem Maß mit $s = \lim_{n\to\infty} \lambda(A_n)$, wobei *s* das Supremum aller Werte von $\lambda(A)$ für Mengen *A* von σ-endlichem Maß bezeichne. Dann ist auch $H := \bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$ eine Menge von σ-endlichem Maß, mit $\lambda(H) = s$. Wir setzen jetzt $g := g_H \cdot \mathbb{1}_H$. Dann ist $g \in L^q(\Omega, \mathcal{A}, \mu)$. Ist nun *E* eine beliebige messbare Menge von σ-endlichem Maß mit $E \supseteq H$, so ist $g_E|_H = g_H$ μ-fast überall mit

$$\lambda(E) = \int |g_E|^q \, d\mu \le \lambda(H) = \int |g_H|^q \, d\mu,$$

also $g_E = 0$ μ-fast überall auf $E \setminus H$. Damit ist aber $g_E = g$ μ-fast überall auf *E*. Für beliebiges $f \in L^p(\Omega, \mathcal{A}, \mu)$ ist nun aber $C := \{f \ne 0\} = \bigcup_{k=1}^{\infty} \left\{ |f| > \frac{1}{k} \right\}$ eine Menge von σ-endlichem Maß, also auch $E = C \cup H$. Also gilt

$$g^*(f) = \int fg_E \, d\mu = \int fg \, d\mu, \text{ mit } \|g^*\| = \|g\|_q \text{ wie zuvor.}$$

Damit ist der Satz vollständig bewiesen. ■

Bemerkung: Satz 31 besagt *nicht*, dass beschränkte lineare Funktionale auf $L^p(\Omega,\mathcal{A},\mu)$ *ausschließlich* durch Elemente aus $g\in L^q(\Omega,\mathcal{A},\mu)$ beschrieben werden können. Zumindest in der Situation eines nicht σ-endlichen Maßes μ können auch Elemente anderer Banach-Räume dazu verwendet werden; sie liefern allerdings keine anderen *Funktionale*. Wir verdeutlichen dies an folgendem

Beispiel 17. Es sei $\Omega=\mathbb{Z}^+$, $\mathcal{A}=\mathfrak{P}(\Omega)$ und $\mu(\{\omega\})=\begin{cases}1, & \omega=0\\ \infty, & \omega>0\end{cases}$ für $\omega\in\Omega$. Dann ist μ nicht σ-endlich, und es gilt $f\in L^p(\Omega,\mathcal{A},\mu)$ genau dann, wenn $f(\omega)=0$ für alle $\omega\in\mathbb{N}$, mit $f(0)\in\mathbb{R}$. Die beschränkten linearen Funktionale für $1\le p<\infty$ sind hier trivialerweise genau von der Form

$$g^*(f)=\alpha\cdot f(0) \text{ mit } \alpha\in\mathbb{R}, \text{ für } f\in L^p(\Omega,\mathcal{A},\mu), \text{ mit } \|g^*\|=|\alpha|.$$

Andererseits gilt aber auch

$$g^*(f)=\int fg\,d\mu \text{ für } f\in L^p(\Omega,\mathcal{A},\mu)$$

mit beliebigen, lediglich messbaren, endlichen (aber nicht notwendig wesentlich beschränkten) Abbildungen $g:\Omega\to\mathbb{R}$ mit $g(0)=\alpha$, etwa mit $g(\omega)=\omega$ für $\omega\in\mathbb{N}$. Solche g liegen ganz offensichtlich in keinem $L^q(\Omega,\mathcal{A},\mu)$ für eine beliebige Wahl von $1\le q\le\infty$. Ein möglicher Banach-Raum, dem solche g angehören können, ist beispielsweise der gewichtete ℓ^q_α-Raum mit

$$\ell^q_\alpha=\left\{\{f_n\}_{n\in\mathbb{N}}\subseteq\mathbb{R}^{\mathbb{N}} \mid \sum_{n=1}^{\infty}|\alpha_n f_n|^q<\infty\right\} \text{ mit } \|f\|_{\alpha,q}:=\sum_{n=1}^{\infty}|\alpha_n f_n|^q$$

mit einer geeigneten Folge $\alpha\in\ell^\infty$, für die $\alpha_n\neq 0$ für alle $n\in\mathbb{N}$ ist; etwa $\alpha_n=n^{-\beta}$ mit $\beta\ge 3$.

Dieses Beispiel zeigt zugleich auch, warum die σ-Endlichkeit im Fall $p=1$ eine wesentliche Rolle spielt, denn in der obigen Situation sind *alle* $g\in L^\infty(\Omega,\mathcal{A},\mu)$ geeignet, vermöge

$$g^*(f):=\int fg\,d\mu=f(0)g(0) \text{ für } f\in L^p(\Omega,\mathcal{A},\mu)$$

ein beschränktes lineares Funktional auf $L^1(\Omega,\mathcal{A},\mu)$ zu definieren, mit $\|g^*\|=|g(0)|\le\|g\|_\infty$. Offensichtlich gibt es hier sogar unendlich viele solcher Wahlen mit

$$\|g^*\|=|g(0)|=|\alpha|<\|g\|_\infty=\sup_{\omega\in\Omega}|g(\omega)|.$$

Eine ausführlichere Diskussion über den Dualraum $\mathfrak{L}^1(\Omega,\mathcal{A},\mu)^*$ bzw. $L^1(\Omega,\mathcal{A},\mu)^*$ findet man in ROYDEN (1988), Chapter 11, Section 7.

Der Dualraum $\mathfrak{L}^\infty(\Omega,\mathcal{A},\mu)^*$ ist dagegen wesentlich von $\mathfrak{L}^1(\Omega,\mathcal{A},\mu)$ verschieden; vgl. ROYDEN (1988), Chapter 11, Section 7, Problem 46 oder HEWITT / STROMBERG (1965), S. 354ff. Dies ist auch aus Lemma 27 ersichtlich, denn der dort definierte Banach-Limes ist i. Allg. nicht in Integralform ausdrückbar.

Der nachfolgende Satz enthält als Umkehrung von Lemma 24 ein zu Satz 31 analoges Ergebnis für Hilbert-Räume.

Satz 32 (Dualräume von Hilbert-Räumen; Satz von Riesz-Fréchet)**.** Es sei $\left(\mathcal{X}, \|\bullet\|_{\mathcal{X}}\right)$ ein Hilbert-Raum mit Skalarprodukt $(\bullet,\bullet)$. Dann gilt: zu jedem beschränkten linearen Funktional $g^* \in \mathcal{X}^*$ existiert genau ein Element $g \in \mathcal{X}$ mit $\|g^*\| = \|g\|_{\mathcal{X}}$ und

$$g^*(f) := (f, g) \text{ für alle } f \in \mathcal{X},$$

d.h. jeder Hilbert-Raum ist vermöge einer konjugiert-linearen Abbildung isometrisch isomorph zu seinem Dualraum.

Beweis: Für $g^* = \mathbf{0}$ ergibt sich trivialerweise $g = \mathbf{0}$. Für $g^* \neq \mathbf{0}$ betrachten wir den Kern $\mathcal{N} = \left(g^*\right)^{-1}\left(\{0\}\right) \neq \mathcal{X}$. Dieser ist ein abgeschlossener Teilraum von $\mathcal{X}$, also existiert nach Satz 14 ein nicht-triviales Element $h \in \mathcal{N}^{\perp}$ (wähle als kompakte und konvexe Menge etwa $C = \{c\}$ mit $c \notin \mathcal{N}$.) Es folgt $g^*(f)\cdot h - g^*(h)\cdot f \in \mathcal{N}$ für jedes $f \in \mathcal{X}$ (wegen $g^*\left(g^*(f)\cdot h - g^*(h)\cdot f\right) = g^*(f)\cdot g^*(h) - g^*(h)\cdot g^*(f) = 0$) und somit

$$g^*(f)\cdot(h,h) - g^*(h)\cdot(f,h) = 0 \text{ für jedes } f \in \mathcal{X},$$

so dass

$$g^*(f) = \frac{g^*(h)}{(h,h)}(f,h) = \left(f, \frac{\overline{g^*(h)}}{\|h\|_{\mathcal{X}}^2}\cdot h\right) = (f, g) \text{ für alle } f \in \mathcal{X}$$

folgt mit $g = \dfrac{\overline{g^*(h)}}{\|h\|_{\mathcal{X}}^2}\cdot h \in \mathcal{X}$. Die Eindeutigkeit ergibt sich so: ist $z \in \mathcal{X}$ ein weiteres Element mit

$$g^*(f) := (f, z) \text{ für alle } f \in \mathcal{X},$$

so ist $(f, g - z) = \mathbf{0}$ für alle $f \in \mathcal{X}$ und somit nach Lemma 26, Teil b) $g - z = \mathbf{0}$, also $g = z$. ■

Der Rest dieses Abschnitts ist dem Studium des Dualraums von $C[a,b]$ bzw. allgemeinerer Räume stetiger Funktionen auf geeigneten metrischen Räumen gewidmet. Zur Vorbereitung wollen wir den Themenkreis der Riemann-Stieltjes- und Lebesgue-Stieltjes-Integrale etwas genauer beleuchten. Diese Vertiefung ist insbesondere auch für die Einführung des stochastischen Integrals von Bedeutung. Wir folgen zunächst im Wesentlichen HEWITT / STROMBERG (1965), CHAPTER III.

Definition 37 (Riemann-Stieltjes-Integral I)**.** Es sei $[a,b] \subset \mathbb{R}$ ein kompaktes Intervall und $f:[a,b] \to \mathbb{R}$ eine beliebige sowie $g:[a,b] \to \mathbb{R}$ eine schwach monoton wachsende Funktion. Unter einer *Zerlegung von* $[a,b]$ verstehen wir eine angeordnete Menge $\Delta_n = \{t_0, \cdots, t_n\}$ mit $a = t_0 < t_1 < \cdots < t_n = b.$ Die Menge aller solchen Zerlegungen werde mit $\mathfrak{D}[a,b]$ bezeichnet. Für jede Zerlegung $\Delta_n \in \mathfrak{D}[a,b]$ sei ferner

$$U(f,g,\Delta_n) := \sum_{i=1}^{n} \inf\{f(x) | t_{i-1} \le x \le t_i\} \cdot \left(g(t_i) - g(t_{i-1})\right) \text{ und}$$
$$O(f,g,\Delta_n) := \sum_{i=1}^{n} \sup\{f(x) | t_{i-1} \le x \le t_i\} \cdot \left(g(t_i) - g(t_{i-1})\right).$$

[Diese Größen heißen auch *untere* und *obere Darboux-Summen*.] Falls für jedes $\varepsilon > 0$ ein $\Delta_n \in \mathfrak{D}[a,b]$ existiert mit

$$O(f,g,\Delta_n) - U(f,g,\Delta_n) < \varepsilon,$$

so heißt *f Riemann-Stieltjes-integrierbar bezüglich g*.

Satz 33 (Riemann-Stieltjes-Integral I)**.** Unter den Voraussetzungen von Definition 37 sei f Riemann-Stieltjes-integrierbar bezüglich g. Dann gilt:

$$\inf_{\Delta_n \in \mathfrak{D}[a,b]} \{O(f,g,\Delta_n)\} = \sup_{\Delta_n \in \mathfrak{D}[a,b]} \{U(f,g,\Delta_n)\}.$$

Diese Zahl heißt dann das *Riemann-Stieltjes-Integral* von f bezüglich g, in Zeichen:

$$\int_a^b f(x)\, dg(x).$$

Beweis: Durch Verfeinerung der Zerlegungen sind die unteren Darboux-Summen monoton wachsend und die oberen monoton fallend. Es gilt also jedenfalls

$$o := \inf_{\Delta_n \in \mathfrak{D}[a,b]} \{O(f,g,\Delta_n)\} \ge \sup_{\Delta_n \in \mathfrak{D}[a,b]} \{U(f,g,\Delta_n)\} =: u.$$

Wäre nun $o > u,$ so gäbe es zu $\varepsilon < o - u$ eine weitere Zerlegung $\Delta_m \in \mathfrak{D}[a,b]$ mit

$$O(f,g,\Delta_m) - U(f,g,\Delta_m) < \varepsilon : \text{ Widerspruch!}$$

Damit ist der Satz bewiesen. ■

Bemerkung: Man nennt in diesem Zusammenhang die Funktion f gelegentlich auch den *Integranden* und die Funktion g den *Integrator*.

Satz 34 (Riemann-Stieltjes-Integral II)**.** Unter den Voraussetzungen von Definition 37 gelte $f \in C[a,b]$. Dann ist f Riemann-Stieltjes-integrierbar bezüglich g.

Beweis: Da $[a,b]$ kompakt ist, ist f dort sogar gleichmäßig stetig. Zu $\varepsilon > 0$ existiert also ein $\delta > 0$ mit der Eigenschaft:

$$|x-y| < \delta \;\Rightarrow\; |f(x)-f(y)| < \eta := \frac{\varepsilon}{\alpha(b)-\alpha(a)+1} \text{ für alle } x,y \in [a,b].$$

Wähle nun eine Zerlegung $\Delta_n \in \mathfrak{D}[a,b]$ mit $t_i - t_{i-1} < \delta$ sowie Punkte $x_i, y_i \in [t_{i-1}, t_i]$ mit

$$f(x_i) = \max\{f(x) | t_{i-1} \le x \le t_i\},\; f(y_i) = \min\{f(x) | t_{i-1} \le x \le t_i\} \text{ für } i = 1,\cdots,n$$

(dies geht, da f auf Kompakta seine Extremwerte annimmt). Also ist

$$0 \le f(x_i) - f(y_i) \le \eta \quad \text{für } i = 1,\cdots,n,$$

und es folgt

$$\begin{aligned} O(f,g,\Delta_n) - U(f,g,\Delta_n) &= \sum_{i=1}^{n} (f(x_i)-f(y_i)) \cdot (g(t_i) - g(t_{i-1})) \\ &< \eta \cdot \sum_{i=1}^{n} (g(t_i) - g(t_{i-1})) = \eta \cdot (g(b)-g(a)) < \varepsilon, \end{aligned}$$

Womit der Satz bewiesen ist. ■

Definition 38 (Totalvariation einer Funktion)**.** Es sei $[a,b] \subset \mathbb{R}$ ein kompaktes Intervall und $g:[a,b] \to \mathbb{R}$ eine Funktion. Dann heißt mit der Notation von Definition 37

$$V_a^b(g) := \sup_{\Delta_n \in \mathfrak{D}[a,b]} \left\{ \sum_{i=1}^{n} |g(t_i) - g(t_{i-1})| \right\} =: \int_a^b |dg(x)|$$

Totalvariation von g über dem Intervall $[a,b]$. Gilt hier $V_a^b(g) < \infty$, so nennt man g *von beschränkter Variation* auf dem Intervall $[a,b]$. Entsprechend heißen

$$V_a^b(g)^+ := \sup_{\Delta_n \in \mathfrak{D}[a,b]} \left\{ \sum_{i=1}^{n} (g(t_i) - g(t_{i-1}))^+ \right\} \text{ bzw. } V_a^b(g)^- := \sup_{\Delta_n \in \mathfrak{D}[a,b]} \left\{ \sum_{i=1}^{n} (g(t_i) - g(t_{i-1}))^- \right\}$$

positive bzw. *negative Variation* von g über dem Intervall $[a,b]$. Eine Funktion g beschränkter Variation heißt *normalisiert*, wenn $g(a) = 0$ gilt und g auf dem offenen (!) Intervall (a,b) rechtsseitig stetig ist.

Wir listen im Folgenden einige wichtige Eigenschaften der Total- bzw. positiven bzw. negativen Variation von Funktionen (ohne Beweis) auf, die größtenteils unmittelbar einsichtig sind. Für Details verweisen wir auf ELSTRODT (1996), Kapitel VII, §1, Aufgabe 1.10 sowie TAYLOR (1963), S. 195ff und BACHMAN / NARICI (2000), Chapter 13.

Lemma 33 (Eigenschaften der Variation einer Funktion)**.** Unter den Voraussetzungen von Definition 38 gilt:

a) Ist g monoton auf $[a,b]$, so folgt $V_a^b(g)=|g(b)-g(a)|$. Ist g monoton wachsend, so gilt $V_a^b(g)^+ = g(b)-g(a)$ und $V_a^b(g)^- = 0$; ist g monoton fallend, so gilt $V_a^b(g)^+ = 0$ und $V_a^b(g)^- = g(a)-g(b)$.

b) Die Abbildungen $V_a^x(g), V_a^x(g)^+$ und $V_a^x(g)^-$ sind als Funktion von x jeweils monoton wachsend in $x\in[a,b]$ mit

$$-V_a^x(g) \le -V_a^x(g)^- \le g(x) \le V_a^x(g)^+ \le V_a^x(g) \text{ für alle } x\in[a,b].$$

c) Jede Funktion g von beschränkter Variation auf dem Intervall $[a,b]$ lässt sich darstellen als Differenz zweier monoton wachsender Funktionen. Genauer gilt:

$$g(x) = g(a) + V_a^x(g)^+ - V_a^x(g)^- \text{ für alle } x\in[a,b].$$

d) Jede Funktion g von beschränkter Variation auf dem Intervall $[a,b]$ kann durch die Festlegung

$$\underline{g}(x) := \begin{cases} 0, & x=a \\ \lim\limits_{u\downarrow x} g(u)-g(a), & a<x<b \\ g(b)-g(a), & x=b \end{cases}$$

normalisiert werden, mit $V_a^b\left(\underline{g}\right) \le V_a^b(g)$. Für stetige Funktionen $f\in C[a,b]$ gilt dabei:

$$\int_a^b f(x)\,dg(x) = \int_a^b f(x)\,d\underline{g}(x).$$

e) Jede Funktion g von beschränkter Variation auf dem Intervall $[a,b]$ besitzt höchstens abzählbar viele Unstetigkeitsstellen.

f) Ist g auf $[a,b]$ differenzierbar, so gilt $V_a^b(g) = \int_a^b |g'(x)|\,dx$.

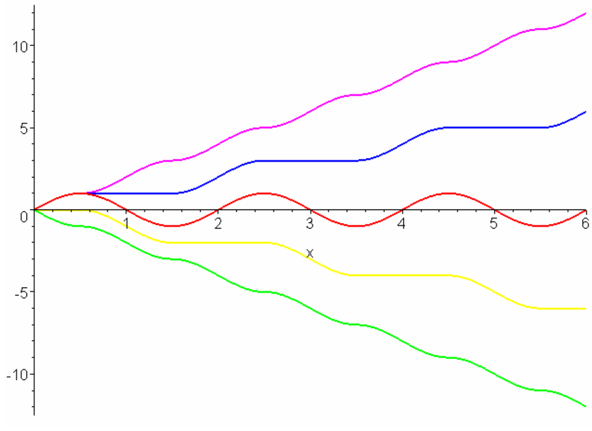
g) Die Menge $BV[a,b]$ aller Funktionen von beschränkter Variation auf dem Intervall $[a,b]$ bildet einen Banach-Raum unter der Norm

$$\|g\|_{BV} := |g(a)| + V_a^b(g).$$

h) Die Menge $NBV[a,b]$ aller normalisierter Funktionen von beschränkter Variation auf dem Intervall $[a,b]$ bildet einen Banach-Raum unter der Norm

$$\|g\|_{NBV} := V_a^b(g).$$

Beispiel 18: die folgende Graphik zeigt den Verlauf von $-V_a^x(g)$, $-V_a^x(g)^-$, $V_a^x(g)^+$ und $V_a^x(g)$ für die Funktion $g(x)=\sin(\pi x)$ im Intervall $[a,b]=[0,6]$.



Legende:

$-V_a^x(g)$: grün

$-V_a^x(g)^-$: gelb

$g(x)$: rot

$V_a^x(g)^+$: blau

$V_a^x(g)$: magenta

Definition 39 (Riemann-Stieltjes-Integral II)**.** Es sei $[a,b]\subset\mathbb{R}$ ein kompaktes Intervall und $f\,[a,b]\to\mathbb{R}$ eine beliebige sowie $g:[a,b]\to\mathbb{R}$ eine Funktion von beschränkter Variation. Dann ist das Riemann-Stieltjes-Integral von f bezüglich g definiert durch

$$\int_a^b f(x)\,dg(x):=\int_a^b f(x)\,d\,V_a^x(g)^+-\int_a^b f(x)\,d\,V_a^x(g)^-,$$

sofern beide Integrale auf der rechten Seite für sich allein als Riemann-Stieltjes-Integrale existieren.

Lemma 34 (Eigenschaften des Riemann-Stieltjes-Integrals)**.** Unter den Voraussetzungen von Definition 39 gilt:

a) Existiert das Riemann-Stieltjes-Integral von f bezüglich g, so gilt auch

$$\int_a^b f(x)\,dg(x)=\int_a^b f(x)\,d\,g_1(x)-\int_a^b f(x)\,d\,g_2(x)$$

für jede Wahl monoton wachsender reeller Funktionen g_1,g_2 auf $[a,b]$ mit $g=g_1-g_2$.

b) Ist $f\in C[a,b]$, so existiert das Riemann-Stieltjes-Integral von f bezüglich g.

c) Ist g auf $[a,b]$ differenzierbar und existiert das Riemann-Stieltjes-Integral von f bezüglich g, so gilt

$$\int_a^b f(x)\,dg(x)=\int_a^b f(x)\cdot g'(x)\,dx.$$

d) Existiert das Riemann-Stieltjes-Integral von f bezüglich g, so gilt

$$\left|\int_a^b f(x)\,dg(x)\right|\le\int_a^b |f(x)|\cdot|dg(x)|:=\int_a^b |f(x)|\;dV_a^x(g).$$

e) Sind $f, g \in C[a,b]$, so gilt die *Regel von der partiellen Integration*:

$$\int_a^b f(x)\,dg(x) = f(x)g(x)\Big|_a^b - \int_a^b g(x)\,df(x).$$

f) Sind g_1, g_2 von beschränkter Variation auf $[a,b]$ und $\beta_1, \beta_2 \in \mathbb{R}$, so ist auch $g := \beta_1 g_1 + \beta_2 g_2$ von beschränkter Variation auf $[a,b]$. Sind ferner f_1, f_2 Riemann-Stieltjes-integrierbar bzgl. g_1 und g_2, so auch bzgl. g, und es gilt: $f := \alpha_1 f_1 + \alpha_2 f_2$ ist Riemann-Stieltjes-integrierbar bzgl. g_1, g_2 und g für alle $\alpha_1, \alpha_2 \in \mathbb{R}$, mit

$$\int_a^b f(x)\,dg(x) = \sum_{i=1}^{2}\sum_{j=1}^{2} \alpha_i \beta_j \int_a^b f_i(x)\,dg_j(x),$$

d.h. das Riemann-Stieltjes-Integral ist linear im Integranden wie im Integrator.

Wir wollen jetzt einen fundamentalen Zusammenhang zwischen normalisierten Funktionen beschränkter Variation auf $[a,b]$ und signierten Maßen mit Träger $[a,b]$ untersuchen. Dazu betrachten wir zunächst Riemann-Stieltjes-Integrale mit speziellen Integranden, nämlich Indikatorfunktionen über Teilintervallen.

Lemma 35. Unter den Voraussetzungen von Definition 39 gilt für jede normalisierte Funktion g beschränkter Variation und jede Indikatorfunktion $f = \mathbb{1}_{(c,d]}$ mit $a < c < d \leq b$:

$$\int_a^b f(x)\,dg(x) = g(d) - g(c).$$

Beweis: Sei zunächst $f = \mathbb{1}_{(c,d]}$ mit $a < c < d \leq b$, g monoton wachsend und auf dem Intervall (a,b) rechtsseitig stetig sowie $\Delta_n \in \mathfrak{D}[a,b]$ eine Zerlegung, für die die Punkte c, $c+h$ mit $0 < h < d-c$ und d Teilpunkte der Zerlegung sind (dies kann ggf. durch Verfeinerung der Zerlegung immer erreicht werden). Dann gilt:

$$U(f,g,\Delta_n) = \sum_{i=1}^{n} \inf\{f(x) | t_{i-1} \leq x \leq t_i\} \cdot \big(g(t_i) - g(t_{i-1})\big) \geq g(d) - g(c+h) \text{ und}$$
$$O(f,g,\Delta_n) = \sum_{i=1}^{n} \sup\{f(x) | t_{i-1} \leq x \leq t_i\} \cdot \big(g(t_i) - g(t_{i-1})\big) \leq g(d) - g(c).$$

Wegen der rechtsseitigen Stetigkeit von g folgt somit

$$\begin{aligned} g(d) - g(c) &\geq \inf_{\Delta_n \in \mathfrak{D}[a,b]} \{O(f,g,\Delta_n)\} \geq \sup_{\Delta_n \in \mathfrak{D}[a,b]} \{U(f,g,\Delta_n)\} \\ &\geq \sup_{0<h<d-c} \big(g(d) - g(c+h)\big) = g(d) - g(c) \end{aligned}$$

und daher $\int_a^b f(x)\,dg(x) = g(d) - g(c)$, wie behauptet. Ist allgemeiner g nicht monoton, so existieren aber jeweils monoton wachsende (und rechtsseitig stetige) Funktionen g^+ und g^- mit

$g = g^+ - g^-$ (z.B. die positive und negative Variation von *g*), so dass nach Lemma 34, Teil a) auch hier folgt:

$$\int_a^b f(x)\,dg(x) = \int f(x)\,dg^+(x) - \int f(x)\,dg^-(x) = \left(g^+(d) - g^+(c)\right) - \left(g^-(d) - g^-(c)\right)$$
$$= g(d) - g(c).$$

Damit ist das Lemma bewiesen. ■

Bemerkung. Auf Grund der Definition der Normalisiertheit von *g* ergeben sich Besonderheiten für Indikatorfunktionen des Typs $f = \mathbb{1}_{(a,d]}$ mit $a < d \leq b.$ Wegen Lemma 34, Teil a) existiert aber der rechtsseitige Grenzwert $g(a+) := \lim\limits_{x \downarrow a} g(x).$ Mit einer Argumentation wie im Beweis von Lemma 35 erhält man damit

$$\int_a^b \mathbb{1}_{(a,d]}(x)\,dg(x) = g(d) - g(a+).$$

Analog lässt sich zeigen, dass

$$\int_a^b \mathbb{1}_{\{a\}}(x)\,dg(x) = g(a+) - g(a)$$

gilt, man also allgemeiner auch noch

$$\int_a^b \mathbb{1}_{[a,d]}(x)\,dg(x) = g(d) - g(a)$$

erhält. Ist ferner $(c,d] \subset \mathbb{R}$ ein *beliebiges* Intervall, so kann für $f = \mathbb{1}_{(c,d]}$ das Riemann-Stieltjes-Integral offensichtlich entsprechend allgemein definiert werden durch

$$\int_a^b \mathbb{1}_{(c,d]}(x)\,dg(x) := \int_a^b \mathbb{1}_{(c,d]}(x) \cdot \mathbb{1}_{[a,b]}(x)\,dg(x) = \int_a^b \mathbb{1}_{(c,d] \cap [a,b]}(x)\,dg(x).$$

Dadurch wird aber auf dem Semi-Ring $\mathcal{E}^1$ der Standard-Intervalle (vgl. Skript STOCHASTIK, Definition 14) durch

$$\nu(E) := \int_a^b \mathbb{1}_E(x)\,dg(x) \text{ für alle } E \in \mathcal{E}^1$$

eine (zunächst) *endlich-additive* Mengenfunktion im Sinne der Definition 17 eines signierten Maßes erklärt. Wir wollen im nächsten Schritt zeigen, dass diese Mengenfunktion sogar σ-additiv auf $\mathcal{E}^1$ ist, so dass sie zu einem signierten Maß auf die Borel'sche σ -Algebra $\mathcal{B}^1$ fortgesetzt werden kann.

Satz 35. Es sei $g \in NBV[a,b]$ eine normalisierte Funktion beschränkter Variation. Dann existiert genau ein (endliches) signiertes Maß ν auf der Borel'schen σ-Algebra $\mathcal{B}^1$ mit

$$\nu(E) = \int_a^b \mathbb{1}_E(x)\, dg(x) \text{ für alle } E \in \mathcal{E}^1.$$

Insbesondere sind damit alle messbaren Teilmengen von $[a,b]^c$ ν-Nullmengen. Bezeichnet speziell $g^+ := V_a^\bullet(g)^+$ die positive und $g^- := V_a^\bullet(g)^-$ die negative Variation von g, so sind auch $g^+, g^- \in NBV[a,b]$, und es gilt

$$\nu^+(E) := \int_a^b \mathbb{1}_E(x)\, dg^+(x),\ \nu^-(E) := \int_a^b \mathbb{1}_E(x)\, dg^-(x) \text{ für alle } E \in \mathcal{E}^1,$$

wobei ν^+ und ν^- den Positiv- und Negativteil von ν gemäß dem Zerlegungs-Satz 11 von Jordan bezeichnen. Für jede bzgl. g Riemann-Stieltjes-integrierbare Funktion f gilt darüber hinaus:

$$\int_a^b f(x)\, dg(x) = \int \overline{f}\, d\nu$$

für jede messbare Fortsetzung $\overline{f}$ von f auf ganz $\mathbb{R}$, d.h. jede messbare Funktion mit der Eigenschaft $\overline{f}\big|_{[a,b]} = f$.

Beweis: Die Funktionen g^+ und g^- sind auf dem offenen Intervall (a,b) nach Konstruktion schwach monoton wachsend und rechtsseitig stetig mit $g^+(a) = g^-(a) = 0$, liegen also in $NBV[a,b]$. Wir nehmen zunächst $g^\pm(b) > 0$ an und definieren

$$F^\pm(x) := \begin{cases} 0, & x < a \\ \dfrac{g^\pm(x)}{g^\pm(b)}, & a \le x < b \\ 1, & x \ge b. \end{cases}$$

Dann ist $F^\pm$ jeweils Verteilungsfunktion eines (eindeutig bestimmten) Wahrscheinlichkeitsmaßes $P^\pm$ auf der Borel'schen σ-Algebra $\mathcal{B}^1$. Die Maße $\nu^\pm$ ergeben sich hieraus durch die Festlegung

$$\nu^\pm := g^\pm(b) \cdot P^\pm.$$

Es ist dann nämlich $\nu := \nu^+ - \nu^-$ ein signiertes Maß mit der Eigenschaft

$$\begin{aligned}\nu(E) &= g^+(b)\cdot P^+(E) - g^-(b)\cdot P^-(E) = \big(g^+(d) - g^+(c)\big) - \big(g^-(d) - g^-(c)\big) \\ &= g(d) - g(c) = \int_a^b \mathbb{1}_E(x)\, dg(x)\end{aligned}$$

für alle $E = (c,d] \subseteq (a,b]$. Für den allgemeinen Fall betrachtet man bei entsprechender Argumentation wie in der vorangehenden Bemerkung die Schnitte $(c,d] \cap [a,b]$. Die Maße ν^+ und ν^- sind auf Grund der Eigenschaften der positiven und negativen Variation von g außerdem orthogonal, bilden also genau die behauptete Jordan-Zerlegung. Für den Fall, dass $g^+(b)=0$ und / oder $g^-(b)=0$ gilt, erhält man natürlich die Trivialfälle $\nu^+=0$ und / oder $\nu^-=0$.

Aus der Definition des Riemann-Stieltjes-Integrals folgt ferner, dass jede bzgl. g Riemann-Stieltjes-integrierbare Funktion f auch messbar ist. Insbesondere ist für jede Zerlegung $\Delta_n \in \mathfrak{D}[a,b]$

$$f_n = \sum_{i=1}^{n} \inf\{f(x) | t_{i-1} \le x \le t_i\} \cdot \mathbb{1}_{(t_{i-1},t_i]} + f(a) \cdot \mathbb{1}_{\{a\}}$$

eine Elementarfunktion, die mit Verfeinerung der Zerlegung ν-fast überall monoton gegen f konvergiert, woraus der Rest der Aussage des Satzes folgt. ■

Wir kommen jetzt zu dem angekündigten Darstellungssatz für die beschränkten linearen Funktionale auf $C[a,b]$.

Satz 36 (Riesz'scher Darstellungssatz II)**.** Es bezeichne $\mathfrak{N}[a,b]$ die Menge der endlichen signierten Maße ν auf der Borel'schen σ-Algebra $\mathcal{B}^1$ mit der Eigenschaft $\nu^{\pm}\left([a,b]^c\right)=0.$ Dann gilt:

a) Jedes endliche signierte Maß $\nu \in \mathfrak{N}[a,b]$ definiert ein beschränktes lineares Funktional f^* auf $C[a,b]$ vermöge

$$f^*(f) := \int f\, d\nu \text{ für alle } f \in C[a,b] \text{ mit } \|f^*\| = \|\nu\|([a,b]),$$

wobei $\|\nu\|$ die Totalvariation von ν im Sinne von Definition 20 bezeichne.

b) Ist umgekehrt f^* ein beschränktes lineares Funktional auf $C[a,b]$, so existiert ein eindeutig bestimmtes endliches signiertes Maß $\nu \in \mathfrak{N}[a,b]$ mit

$$f^*(f) := \int f\, d\nu \text{ für alle } f \in C[a,b].$$

Damit existiert also ein isometrischer Isomorphismus zwischen $C[a,b]^*$ und $\mathfrak{N}[a,b]$. In diesem Sinn kann der Dualraum $C[a,b]^*$ mit $\mathfrak{N}[a,b]$ identifiziert werden.

Beweis:

a) Die Linearität des Integrals impliziert die des Funktionals, mit

$$\left|f^*(f)\right| = \left|\int f\, d\nu\right| \le \int |f|\, d\|\nu\| \le \|f\|_C \int d\|\nu\| = \|f\|_C \|\nu\|([a,b]) \text{ für alle } f \in C[a,b],$$

woraus zunächst $\|f^*\| \le \|\nu\|([a,b])$ folgt. Sind nun $\nu^+ := \nu(\bullet \cap A)$ und $\nu^- := -\nu(\bullet \cap B)$ die Positiv- und Negativteile von ν nach den Zerlegungssätzen von Hahn und Jordan (vgl. den Beweis zu Satz 11) mit geeigneten Auswahlen für die zugehörigen disjunkten positiven bzw. negativen Mengen A, B (die unter den getroffenen Annahmen so gewählt werden können, dass $A, B \subseteq [a,b]$ ist), so folgt mit $f := \mathbb{1}_A - \mathbb{1}_B$ sofort

$$\left|\int f\,d\nu\right| = \left|\int (\mathbb{1}_A - \mathbb{1}_B)\,d\nu\right| = \left|\nu^+(A) + \nu^-(B)\right| = \|\nu\|([a,b]).$$

Mit Hilfe des schon in Beispiel 4, Teil c) verwendeten *Satzes von Lusin* lässt sich aber (sogar allgemeiner als in Beispiel 4) zu jedem $\varepsilon > 0$ eine messbare Menge $A_\varepsilon \subseteq [a,b]$ mit $\|\nu\|(A_\varepsilon) < \varepsilon$ sowie eine stetige Funktion $f_\varepsilon \in C[a,b]$ mit $\mathbb{1}_A(x) - \mathbb{1}_B(x) = f(x) = f_\varepsilon(x)$ für alle $x \in A_\varepsilon^c$ finden; es folgt damit

$$\left|f^*(f_\varepsilon) - \int f\,d\nu\right| = \left|\int (f_\varepsilon - f)\,d\nu\right| \le \int |f_\varepsilon - f|\,d\|\nu\| \le \left\{\|f_\varepsilon\|_C + 1\right\}\int \mathbb{1}_{A_\varepsilon}\,d\|\nu\| \le \varepsilon\left\{\|f_\varepsilon\|_C + 1\right\}.$$

Dabei kann o.B.d.A. angenommen werden, dass $\|f_\varepsilon\|_C = 1$ ist, da mit f_ε auch die Funktion $\tilde{f}_\varepsilon := \max\left\{\min\{f_\varepsilon, 1\}, -1\right\}$ stetig ist und auf A_ε^c mit f übereinstimmt (dort ist ja $|f|$ durch 1 beschränkt). Damit folgt

$$\left|f^*(f_\varepsilon)\right| \ge \left|\int f\,d\nu\right| - 2\varepsilon = \|f_\varepsilon\|_C\,\|\nu\|([a,b]) - 2\varepsilon,$$

also letztlich $\|f^*\| = \|\nu\|([a,b])$, wie behauptet.

b) Wir betrachten den Banach-Raum $B[a,b]$ der beschränkten Funktionen auf $[a,b]$, ausgestattet mit der üblichen Supremumsnorm

$$\|f\|_B := \sup_{a \le x \le b}\left\{|f(x)|\right\} \text{ für alle } f \in B[a,b].$$

Dann ist $C[a,b]$ ein Unterraum von $B[a,b]$, so dass sich nach dem Hahn-Banach-Fortsetzungssatz II (Satz 26) das beschränkte lineare Funktional f^* auf $C[a,b]$ zu einem beschränkten linearen Funktional F^* auf $B[a,b]$ mit gleicher Norm fortsetzen lässt. Wir betrachten jetzt die Funktion

$$g(x) := \begin{cases} F^*\left(\mathbb{1}_{[a,x]}\right), & x > a \\ 0 = F^*(\mathbf{0}), & x = a \end{cases} \quad \text{für } x \in [a,b].$$

Diese ist von beschränkter Variation mit

$$V_a^b(g) = \|g\|_{BV} \le \|f^*\|,$$

wie man auf folgende Weise sehen kann:

Ist $\Delta_n \in \mathfrak{D}[a,b]$ eine Zerlegung und setzt man

$$h := s_1 \mathbb{1}_{[a,t_1]} + \sum_{i=2}^n s_i \mathbb{1}_{(t_{i-1},t_i]} \text{ mit } s_j := \operatorname{sgn}\left\{g(t_j) - g(t_{j-1})\right\},\ j = 1,\cdots,n,$$

so ist $h \in B[a,b]$ mit $\|h\|_B \le 1$ und

$$F^*(h) = s_1 F^*\left(\mathbb{1}_{[a,t_1]}\right) + \sum_{i=2}^n s_i F^*\left(\mathbb{1}_{(t_{i-1},t_i]}\right) = \sum_{i=1}^n s_i\left(g(t_i) - g(t_{i-1})\right) = \sum_{i=1}^n \left|g(t_i) - g(t_{i-1})\right| \le \|f^*\|_{C[a,b]^*}$$

wegen

$$\left|F^*(h)\right| \le \left\|F^*\right\|_{B[a,b]^*} \left\|h\right\|_B \le \left\|F^*\right\|_{B[a,b]^*} = \left\|f^*\right\|_{C[a,b]^*}.$$

Damit folgt aber nach Supremumsbildung gerade $V_a^b(g) = \|g\|_{BV} \le \left\|f^*\right\|_{C[a,b]^*}$. Mit $\Delta_n \in \mathfrak{D}[a,b]$ definiere nun für $f \in C[a,b]$

$$f_n := \sum_{i=1}^{n} f\left(t_{i-1}\right) \mathbb{1}_{\left(t_{i-1}, t_i\right]}.$$

Da f auf dem Kompaktum $[a,b]$ sogar gleichmäßig stetig ist, konvergiert nun die Folge $\left\{f_n\right\}_{n\in\mathbb{N}}$ in $B[a,b]$ gegen *f*, wenn die Zerlegungen feiner werden, so dass dann auch

$$\lim_{n\to\infty} \left|F^*(f) - F^*\left(f_n\right)\right| \le \lim_{n\to\infty} \left(\left\|F^*\right\|_{B[a,b]^*} \cdot \left\|f - f_n\right\|_B\right) = 0$$

resultiert. Es folgt daher

$$f^*(f) = F^*(f) = \lim_{n\to\infty} F^*\left(f_n\right) = \lim_{n\to\infty} \sum_{i=1}^{n} f\left(t_{i-1}\right)\left(g\left(t_i\right) - g\left(t_{i-1}\right)\right) = \int_a^b f(x)\, dg(x).$$

Nach Lemma 33, Teil d) und Satz 35 existiert also ein eindeutig bestimmtes endliches signiertes Maß $\nu \in \mathfrak{N}[a,b]$ mit

$$f^*(f) = \int_a^b f(x)\, dg(x) = \int_a^b f(x)\, d\underline{g}(x) = \int f\, d\nu \text{ für alle } f \in C[a,b],$$

wie behauptet. ■

Bemerkung. Der vorangehende Beweis zeigt sogar genauer, dass $V_a^b(g) = \|g\|_{BV} = \left\|f^*\right\|_{C[a,b]^*}$ ist, und dass wegen Lemma 33, Teil d) auch $V_a^b(g) = V_a^b\left(\underline{g}\right) = \left\|\underline{g}\right\|_{NBV} = \left\|f^*\right\|_{C[a,b]^*}$ gilt. Damit ergibt sich eine Alternative zum Argument mit dem Satz von Lusin. Es existiert also nicht nur ein isometrischer Isomorphismus zwischen $C[a,b]^*$ und $\mathfrak{N}[a,b]$, sondern auch zwischen $C[a,b]^*$ und $NBV[a,b]$. In diesem Sinn kann der Dualraum $C[a,b]^*$ auch mit $NBV[a,b]$ bzw. $\mathfrak{N}[a,b]$ mit $NBV[a,b]$ identifiziert werden.

Der Riesz'sche Darstellungssatz gilt in allgemeinerer Form auch für Banch-Räume stetiger Funktionen auf gewissen *lokalkompakten Hausdorff-Räumen*, vgl. ELSTRODT (1996), Kapitel VII, § 2, Satz 2.26 oder ROYDEN (1988), Chapter 13, Section 5, Theorem 25; siehe auch TAYLOR (1963), § 7.7. Die Beweistechnik erfordert allerdings einen Aufwand, den wir hier nicht leisten können. Eine einfachere Verallgemeinerung ist dabei die Erweiterung ins Mehrdimensionale. Ist etwa $\mathcal{X} \subset \mathbb{R}^d$ eine kompakte Menge mit $d \in \mathbb{N}$, so ist entsprechend der Dualraum $C(\mathcal{X})^*$ isometrisch isomorph zur Menge $\mathfrak{N}(\mathcal{X})$ der signierten Maße ν auf der Borel'schen σ-Algebra $\mathcal{B}^d$ mit der Eigenschaft $\nu^{\pm}\left(\mathcal{X}^c\right) = 0.$

I. 3.3. Zweite konjugierte und reflexive Räume

In den bisherigen Überlegungen haben wir bei dem Dualraum $\mathcal{X}^*$ eines linearen normierten Raumes $\left(\mathcal{X}, \|\bullet\|_{\mathcal{X}}\right)$ das lineare Funktional $f^* \in \mathcal{X}^*$ festgehalten und die Elemente $f \in \mathcal{X}$ variiert. Man kann aber auch umgekehrt ein Element $f \in \mathcal{X}$ festhalten und die linearen Funktionale $f^* \in \mathcal{X}^*$ variieren. So erhält man in natürlicher Weise durch die Elemente des ursprünglichen Raumes lineare Funktionale auf dem Dualraum $\left(\mathcal{X}^*, \|\bullet\|_{\mathcal{X}^*}\right)$. Diese Betrachtungsweise führt zum Begriff des zweiten konjugierten Raumes.

Definition 40 (zweiter konjugierter Raum)**.** Es sei $\left(\mathcal{X}, \|\bullet\|_{\mathcal{X}}\right)$ ein linearer normierter Raum und $\left(\mathcal{X}^*, \|\bullet\|_{\mathcal{X}^*}\right)$ der zugehörige Dualraum. Dann heißt $\mathcal{X}^{**} := \left(\mathcal{X}^*\right)^*$ der (zu $\mathcal{X}$) *zweite konjugierte Raum* oder *Bidualraum*.

Lemma 36. Durch jedes $f \in \mathcal{X}$ wird ein beschränktes lineares Funktional f^{**} auf $\left(\mathcal{X}^*, \|\bullet\|_{\mathcal{X}^*}\right)$ definiert vermöge

$$f^{**}\left(f^*\right) := f^*(f) \quad \text{mit } \left\|f^{**}\right\| \leq \|f\|_{\mathcal{X}} \quad \text{für alle } f^* \in \mathcal{X}^*,$$

d.h. es gilt $f^{**} \in \mathcal{X}^{**}$ für alle $f^* \in \mathcal{X}^*$ und somit $\mathcal{X} \subseteq \mathcal{X}^{**}$.

Beweis: Die Linearität von f^{**} ist klar. Die Beschränktheit folgt aus der Ungleichung

$$\left|f^{**}\left(f^*\right)\right| = \left|f^*(f)\right| \leq \left\|f^*\right\|_{\mathcal{X}^*} \|f\|_{\mathcal{X}} \quad \text{für alle } f \in \mathcal{X} \text{ und } f^* \in \mathcal{X}^*. \qquad \blacksquare$$

Definition 41 (kanonische Abbildung)**.** Die durch Lemma 36 induzierte Abbildung $J : \mathcal{X} \to \mathcal{X}^{**} : f \mapsto f^{**}$ heißt *die kanonische Abbildung* zwischen $\mathcal{X}$ und $\mathcal{X}^{**}$.

Lemma 37. Die kanonische Abbildung *J* zwischen einem nicht-trivialen linearen normierten Raum und seinem Bidualraum ist linear, isometrisch und injektiv.

Beweis: Für alle $\alpha, \beta \in K$, $f, g \in \mathcal{X}$ und $f^* \in \mathcal{X}^*$ gilt:

$$\begin{aligned} J(\alpha f + \beta g)\left(f^*\right) &= (\alpha f + \beta g)^{**}\left(f^*\right) = f^*(\alpha f + \beta g) = \alpha f^*(f) + \beta f^*(g) \\ &= \alpha f^{**}\left(f^*\right) + \beta g^{**}\left(f^*\right) = \left(\alpha f^{**} + \beta g^{**}\right)\left(f^*\right) = \left(\alpha J(f) + \beta J(g)\right)\left(f^*\right), \end{aligned}$$

d.h. *J* ist linear. Nach Lemma 36 gilt ferner $\|Jf\| = \left\|f^{**}\right\| \leq \|f\|_{\mathcal{X}}$ für alle $f \in \mathcal{X}$, d.h. *J* ist beschränkt mit $\|J\| \leq 1$. Sei nun $f_0 \in \mathcal{X}$, $f_0 \neq \mathbf{0}$. Dann existiert nach Satz 27 ein $F^* \in \mathcal{X}^*$ mit $\left\|F^*\right\|_{\mathcal{X}^*} = 1$ und $F^*\left(f_0\right) = \|f_0\|_{\mathcal{X}}$, so dass

$$\|f_0\|_{\mathcal{X}} = F^*\left(f_0\right) = f_0^{**}\left(F^*\right) \leq \left\|f_0^{**}\right\| \cdot \underbrace{\left\|F^*\right\|_{\mathcal{X}^*}}_{=1} = \left\|f_0^{**}\right\| \leq \|f_0\|_{\mathcal{X}}, \text{ also } \|f_0\|_{\mathcal{X}} = \left\|f_0^{**}\right\| = \|Jf_0\|$$

ist (und damit $\|J\| = 1$ folgt). Also ist *J* isometrisch. Die Injektivität folgt aus Lemma 26. $\blacksquare$

Definition 42 (Reflexivität)**.** Es sei $\left(\mathcal{X}, \|\bullet\|_{\mathcal{X}}\right)$ ein linearer normierter Raum und $\left(\mathcal{X}^*, \|\bullet\|_{\mathcal{X}^*}\right)$ der zugehörige Dualraum sowie J die kanonische Abbildung zwischen $\mathcal{X}$ und $\mathcal{X}^{**}$. $\left(\mathcal{X}, \|\bullet\|_{\mathcal{X}}\right)$ heißt *reflexiv*, wenn J surjektiv ist, also $\mathcal{X}^{**} = J(\mathcal{X})$ gilt.

Bemerkung: Bei einem reflexiven Raum kann also der Bidualraum mit dem ursprünglichen Raum identifiziert werden. Nicht vollständige lineare normierte Räume können dabei nicht reflexiv sein, weil nach Lemma 25 ein Bidualraum (als Dualraum des Dualraums) stets vollständig ist.

Beipiele reflexiver Räume sind:

- die Lebesgue-Räume $\mathfrak{L}^p(\Omega, \mathcal{A}, \mu)$ mit $1 < p < \infty$
- jeder endlich-dimensionale lineare normierte Raum
- jeder Hilbert-Raum (vgl. Satz 32).

Den nachfolgenden Satz zitieren wir der Vollständigkeit halber nur ohne Beweis; vgl. WERNER (2004), Satz III.3.4 und Korollar III.3.5).

Satz 37. Es sei $\left(\mathcal{X}, \|\bullet\|_{\mathcal{X}}\right)$ ein linearer normierter Raum. Dann gilt:

a) Ist $\mathcal{X}$ reflexiv und A ein abgeschlossener Teilraum von $\mathcal{X}$, so ist auch A reflexiv.
b) Ist $\mathcal{X}$ vollständig (also ein Banach-Raum), so ist $\mathcal{X}$ genau dann reflexiv, wenn $\mathcal{X}^*$ reflexiv ist.
c) Ist $\mathcal{X}$ reflexiv, so ist $\mathcal{X}$ genau dann separabel, wenn $\mathcal{X}^*$ separabel ist.

Bemerkung: Ein Konsequenz des vorigen Satzes, Teil c) ist, dass die Räume ℓ^p nicht reflexiv sind und i. Allg. auch nicht die Lebesgue-Räume $\mathfrak{L}^p(\Omega, \mathcal{A}, \mu)$ für $p \in \{1, \infty\}$. [Diese sind aber reflexiv, falls sie endlich-dimensional sind, was jedoch von der genauen Gestalt der Grundmenge Ω, der σ-Algebra $\mathcal{A}$ bzw. dem Maß μ abhängt.] Ebenso ist $C[a,b]$ nicht reflexiv.

I. 3.4. Schwache und schwach-$*$-Konvergenz

In diesem Abschnitt wollen wir kurz einige topologische Aspekte von Dual- und Bidualräumen ansprechen, die das Konzept der Norm-Topologie abschwächen und auch in der Stochastik (z.B. bei Grenzwertsätzen von Verteilungen oder für Stochastische Prozesse, vgl. POLLARD (1984)) von Bedeutung sind.

Definition 43 (schwache Konvergenz)**.** Es sei $\left(\mathcal{X}, \|\bullet\|_{\mathcal{X}}\right)$ ein linearer normierter Raum. Eine Folge $\{f_n\}_{n\in\mathbb{N}} \subseteq \mathcal{X}$ heißt *schwach konvergent* gegen $f \in \mathcal{X}$, wenn $\left\{f^*(f_n)\right\}_{n\in\mathbb{N}}$ (in $\mathbb{R}$) gegen $f^*(f)$ konvergiert für alle $f^* \in \mathcal{X}^*$; Schreibweise: $f_n \xrightarrow{w} f$ oder $f = w\text{-}\lim\limits_{n\to\infty} f_n$.

Bemerkung. Zur besseren begrifflichen Abgrenzung wird die übliche Konvergenz in einem linearer normierter Raum auch als *starke Konvergenz* bezeichnet.

Lemma 38. Es sei $\left(\mathcal{X}, \|\bullet\|_{\mathcal{X}}\right)$ ein linearer normierter Raum. Dann gilt:

a) Der schwache Grenzwert einer Folge $\{f_n\}_{n\in\mathbb{N}} \subseteq \mathcal{X}$ ist eindeutig bestimmt und linear auf der Menge der Folgen aus $\mathcal{X}$.

b) Jede stark konvergente Folge $\{f_n\}_{n\in\mathbb{N}} \subseteq \mathcal{X}$ ist auch schwach konvergent.

Beweis: Der erste Teil folgt aus Lemma 26 und der Linearität von $f^* \in \mathcal{X}^*$, der zweite unmittelbar aus der Ungleichung

$$\left|f^*(f_n) - f^*(f)\right| \le \left\|f^*\right\|_{\mathcal{X}^*} \left\|f_n - f\right\|_{\mathcal{X}}$$

für alle $f^* \in \mathcal{X}^*$. ■

Dass die Umkehrung in Lemma 38, Teil b) i. Allg. nicht gilt, zeigt das folgende

Beispiel 19. Wir betrachten den Banach-Raum $\mathcal{X} = \ell^p$ mit $1 < p < \infty$. Ferner sei $q = \dfrac{p}{p-1}$ der zu p konjugierte Index. Nach dem Riesz'schen Darstellungssatz I (Satz 31) kann jedes beschränkte lineare Funktional $g^* \in \mathcal{X}^*$ dargestellt werden als

$$g^*(f) = \sum_{n=1}^{\infty} f_n g_n$$

mit einem geeigneten $g \in \ell^q$. Bezeichnen nun e_k mit ähnlich wie in Beispiel 15 wieder die kanonischen Einheitsvektoren in ℓ^p, so gilt

$$g^*(e_k) = f_k \to 0 \text{ für } k \to \infty, k \in \mathbb{N}$$

d.h. die Folge $\{e_k\}_{k\in\mathbb{N}}$ ist schwach gegen **0** konvergent. Sie ist aber wegen $\|e_k\|_p = 1$ für alle $k \in \mathbb{N}$ nicht stark gegen **0** konvergent.

Definition 44 (schwach-$*$-Konvergenz)**.** Es sei $\left(\mathcal{X}, \|\bullet\|_{\mathcal{X}}\right)$ ein linearer normierter Raum. Eine Folge $\left\{f_n^*\right\}_{n\in\mathbb{N}} \subseteq \mathcal{X}^*$ heißt *schwach-$*$-konvergent* gegen $f^* \in \mathcal{X}^*$, wenn $\left\{f_n^*(f)\right\}_{n\in\mathbb{N}}$ (in $\mathbb{R}$) gegen $f^*(f)$ konvergiert für alle $f \in \mathcal{X}$; Schreibweise: $f_n^* \xrightarrow{w^*} f^*$ oder $f^* = w^*\text{-}\lim\limits_{n\to\infty} f_n^*$.

Bemerkung. In dem Dualraum $\left(\mathcal{X}^*, \|\bullet\|_{\mathcal{X}^*}\right)$ existieren also (mindestens) drei verschiedene Arten von Konvergenz: stark, schwach und schwach-$*$. Diese sind in der Regel auch voneinander verschiedenen, wie weiter unten gezeigt werden wird.

Lemma 39. Es sei $\left(\mathcal{X}, \|\bullet\|_{\mathcal{X}}\right)$ ein linearer normierter Raum. Dann gilt:

a) Der schwach-$*$-Grenzwert einer Folge $\left\{f_n^*\right\}_{n\in\mathbb{N}} \subseteq \mathcal{X}^*$ ist eindeutig bestimmt und linear auf der Menge der Folgen aus $\mathcal{X}^*$.

b) Jede (in $\mathcal{X}^*$) schwach konvergente Folge $\left\{f_n^*\right\}_{n\in\mathbb{N}} \subseteq \mathcal{X}^*$ ist auch schwach-$*$ konvergent.

Beweis: Teil a) ergibt sich so: sind $f^*, g^* \in \mathcal{X}^*$ mit $f_n^* \xrightarrow{w^*} f^*$ und $f_n^* \xrightarrow{w^*} g^*$, so folgt

$$\left|f^*(f) - g^*(f)\right| \leq \left|f_n^*(f) - f^*(f)\right| + \left|f_n^*(f) - g^*(f)\right| \to 0 \text{ mit } n \to \infty \text{ für alle } f \in \mathcal{X},$$

d.h. es ist $\left|f^*(f) - g^*(f)\right| = 0$ für alle $f \in \mathcal{X}$ und damit $f^* = g^*$. Die Linearität des schwach-$*$ Grenzwerts folgt aus der Linearität der Funktionale. Teil b) ergibt sich aus Lemma 36. ■

Dass die Umkehrung in Lemma 39, Teil b) i. Allg. nicht gilt, zeigt das folgende

Beispiel 20. Wir betrachten den Banach-Raum c_0 der reellen Nullfolgen, ausgestattet mit der ℓ^∞-Norm. Es lässt sich zeigen, dass der Dualraum $c_0{}^*$ isometrisch isomorph zu ℓ^1 ist, dessen Dualraum nach dem Riesz'schen Darstellungssatz I aber isometrisch isomorph zu $\ell^\infty \neq c_0$ ist. (Damit ist insbesondere c_0 nicht reflexiv.) Zu jedem linearen Funktional $g^* \in c_0{}^*$ existiert also ein $g \in \ell^1$ mit

$$g^*(f) = \sum_{n=1}^{\infty} f_n g_n \text{ für alle } f \in c_0.$$

Wir betrachten nun in ℓ^1 wieder die kanonischen Einheitsvektoren e_k für $k \in \mathbb{N}$, aufgefasst im obigen Sinne als beschränkte lineare Funktionale e_k^* auf c_0. Dann konvergiert die Folge $\left\{e_k^*\right\}_{k\in\mathbb{N}}$ schwach-$*$ gegen **0** wegen

$$e_k^*(f) = \sum_{n=1}^{\infty} e_{kn} f_n = f_k \to 0 \text{ mit } n \to \infty \text{ für alle } f \in c_0.$$

Die Folge ist aber nicht schwach konvergent, weil für $h = (1,1,1,\cdots) \in \ell^\infty$ gilt:

$$h^{**}\left(e_k^*\right) = \sum_{n=1}^{\infty} e_{kn} h_n = h_k = 1 \text{ mit Grenzwert } 1 \neq 0.$$

Bemerkung. In einem *reflexiven* linearen normierten Raum sind natürlich auf Grund der Definition von Reflexivität die schwache und schwach-$*$-Konvergenz äquivalent.

Definition 45 (schwach-$*$-Folgenkompaktheit)**.** Eine Teilmenge $A \subseteq \mathcal{X}^*$ im Dualraum eines linearen normierten Raumes $\left(\mathcal{X}, \|\bullet\|_{\mathcal{X}}\right)$ heißt *schwach-$*$-folgenkompakt*, wenn jede Folge in A eine in A schwach-$*$-konvergente Teilfolge besitzt.

Satz 38 (Alaoglu)**.** Der lineare normierte Raum $\left(\mathcal{X}, \|\bullet\|_{\mathcal{X}}\right)$ sei separabel. Dann ist die abgeschlossene Einheitskugel $\bar{K}_1^*(\mathbf{0}) = \left\{ f^* \in \mathcal{X}^* \mid \|f^*\|_{\mathcal{X}^*} \leq 1 \right\}$ im Dualraum $\mathcal{X}^*$ von $\mathcal{X}$ schwach-$*$-folgenkompakt.

Beweis: Es sei $\left\{ f_n^* \right\}_{n \in \mathbb{N}} \subseteq \bar{K}_1^*(\mathbf{0})$ eine Folge durch 1 beschränkter linearer Funktionale und $A = \left\{ f_n \mid n \in \mathbb{N} \right\} \subseteq \mathcal{X}$ eine abzählbare, in $\mathcal{X}$ dichte Teilmenge. Die reelle Folge $\left\{ f_n^*\left(f_1\right) \right\}_{n \in \mathbb{N}}$ ist wegen $\left| f_n^*\left(f_1\right) \right| \leq \|f_n^*\| \cdot \|f_1\|_{\mathcal{X}} = \|f_1\|_{\mathcal{X}}$ beschränkt, enthält also eine in $\mathbb{R}$ konvergent Teilfolge, etwa $\left\{ f_{n1}^*\left(f_1\right) \right\}_{n \in \mathbb{N}}$. Entsprechend enthält die reelle Folge $\left\{ f_{n1}^*\left(f_2\right) \right\}_{n \in \mathbb{N}}$ eine konvergente Teilfolge, etwa $\left\{ f_{n2}^*\left(f_2\right) \right\}_{n \in \mathbb{N}}$. Auf diese Weise erhält man rekursiv ein System $\left\{ f_{nk}^*\left(f_k\right) \right\}_{n \in \mathbb{N}}$ von konvergenten Teilfolgen. Wir betrachten jetzt die Diagonalfolge $\left\{ f_{nn}^* \right\}_{n \in \mathbb{N}} \subseteq \bar{K}_1^*(\mathbf{0})$, für die nach Konstruktion alle Folgen $\left\{ f_{nn}^*\left(f\right) \right\}_{n \in \mathbb{N}}$ für jedes $f \in A$ konvergieren. Wir zeigen nur noch, dass $\left\{ f_{nn}^* \right\}_{n \in \mathbb{N}}$ in $\mathcal{X}^*$ schwach-$*$-konvergent ist mit einem Grenzwert $f^* \in \bar{K}_1^*(\mathbf{0})$. Sei dazu $\varepsilon > 0$ vorgegeben sowie $f \in \mathcal{X}$ beliebig. Dann gibt es wegen der Dichtheit von A in $\mathcal{X}$ ein $k \in \mathbb{N}$ mit $\|f - f_k\|_{\mathcal{X}} < \varepsilon$. Ferner gilt für alle $n, m \in \mathbb{N}$:

$$\begin{aligned}
\left| f_{nn}^*(f) - f_{mm}^*(f) \right| &\leq \left| f_{nn}^*(f) - f_{nn}^*\left(f_k\right) \right| + \left| f_{nn}^*\left(f_k\right) - f_{mm}^*\left(f_k\right) \right| + \left| f_{mm}^*\left(f_k\right) - f_{mm}^*(f) \right| \\
&\leq \left\{ \|f_{nn}^*\| + \|f_{mm}^*\| \right\} \|f - f_k\|_{\mathcal{X}} + \left| f_{nn}^*\left(f_k\right) - f_{mm}^*\left(f_k\right) \right| \leq 2\varepsilon + \left| f_{nn}^*\left(f_k\right) - f_{mm}^*\left(f_k\right) \right|.
\end{aligned}$$

Da $\left\{ f_{nn}^*\left(f_k\right) \right\}_{n \in \mathbb{N}}$ für alle $k \in \mathbb{N}$ konvergiert, ist diese Folge auch Cauchy-Folge, so dass $\left| f_{nn}^*\left(f_k\right) - f_{mm}^*\left(f_k\right) \right| < \varepsilon$ ist für genügend große $n, m,$ etwa $n, m \geq N.$ Also ist auch $\left\{ f_{nn}^*(f) \right\}_{n \in \mathbb{N}}$ Cauchy-Folge und somit konvergent gegen eine reelle Zahl $f^*(f)$. f^* ist damit klarerweise ein lineares Funktional mit $\|f^*\| \leq 1$ nach dem Satz von Banach-Steinhaus (Satz 23). ■

Bemerkung: In der Literatur wird in der Regel der Begriff der *schwach-$*$-Kompaktheit* verwendet, der bei separablen Räumen zu dem der schwach-$*$-Folgenkompaktheit äquivalent ist. I. Allg. sind diese Begriffe jedoch verschieden. Die schwach-$*$-Kompaktheit erfordert ein geeignetes System offener Mengen; dazu muss zunächst die schwach-$*$-Topologie als *Topologie der punktweisen Konvergenz* eingeführt werden [vgl. BACHMAN / NARICI (2000), Section 19.7 oder WERNER (2004), Kapitel VIII]. Der Satz von Alaoglu gilt dann auch allgemeiner für nicht-separable Räume in dem Sinne, dass die schwach-$*$-Einheitskugel stets schwach-$*$- kompakt ist (vgl. etwa ROYDEN (1988), Chapter 10, Section 7, Theorem 17). Allerdings gilt der Satz nicht mehr allgemein mit der schwach-$*$-Folgenkompaktheit, wie das folgende Beispiel zeigt.

Beispiel 21. Wir betrachten den (nicht separablen) Raum $\mathcal{X} = \ell^\infty$ und seinen Dualraum $\mathcal{X}^*$ (der "größer" ist als ℓ^1). Nach Lemma 29 wird jedenfalls durch jedes $g \in \ell^1$ ein beschränktes lineares Funktional g^* auf $\mathcal{X} = \ell^\infty$ definiert vermöge

$$g^*(f) = \sum_{n=1}^{\infty} g_n f_n \text{ für alle } f \in \mathcal{X} = \ell^\infty \text{ mit } \left\|g^*\right\|_{\mathcal{X}^*} = \left\|g\right\|_1 .$$

Die Folge $\left\{g_n^*\right\}_{n\in\mathbb{N}} \subset \bar{K}_1^*(\mathbf{0})$ der durch die kanonischen Einheitsvektoren aus ℓ^1 induzierten beschränkten linearen Funktionale mit

$$g_n^*(f) = \sum_{k=1}^{\infty} e_k f_k = f_n \text{ für alle } f \in \mathcal{X} = \ell^\infty \text{ mit } \left\|g_n^*\right\|_{\mathcal{X}^*} = \left\|e_n\right\|_1 = 1$$

(vgl. die Beispiele 15 und 20) besitzt aber keine schwach-$*$-konvergente Teilfolge. Denn sei $\left\{n_k\right\}_{\in\mathbb{N}} \subseteq \mathbb{N}$ beliebig vorgegeben, so betrachte die Indexmenge $M := \left\{n_{2k} \mid k \in \mathbb{N}\right\}$ sowie das spezielle Element $f \in \mathcal{X} = \ell^\infty$ mit $f_m = \begin{cases} 1, & m \in M \\ 0, & m \notin M \end{cases}$ für $m \in \mathbb{N}$. Dann gilt:

$$g_{n_{2k}}^*(f) = f_{n_{2k}} = 1, \; g_{n_{2k-1}}^*(f) = f_{n_{2k-1}} = 0 \text{ für alle } k \in \mathbb{N},$$

d.h. die Folge $\left\{g_{n_k}^*(f)\right\}_{k\in\mathbb{N}}$ ist oszillierend und damit nicht konvergent.

Der Satz von Alaoglu hat eine Reihe von Anwendungen in der Optimierungstheorie, weil die abgeschlossene Einheitskugel eines linearen normierten Raumes nur im endlich-dimensionalen Fall kompakt ist, man aber gern mit dem Annehmen von Extremwerten stetiger Abbildungen auf Kompakta argumentiert. Man kann die Stetigkeit hier zu einer schwach-$*$-Stetigkeit abschwächen und erhält dann das analoge Ergebnis, dass schwach-$*$-stetige Abbildungen auf schwach-$*$-kompakten Mengen ihre Extremwerte annehmen. Für weitere anwendungsrelevante Details siehe etwa LUENBERGER (1968).

I. 4. Abgeschlossene Operatoren

In diesem Abschnitt geht es um wichtige Zusammenhänge zwischen der Beschränktheit (Stetigkeit) von linearen Operatoren auf Teilräumen linearer normierter Räume und deren Abgeschlossenheit bzw. der Abgeschlossenheit dieser Teilräume. Diese Zusammenhänge sind insbesondere in der nachfolgenden Halbgruppentheorie von Interesse.

Definition 46 (abgeschlossener Operator)**.** Es seien $\left(\mathcal{X}, \left\|\bullet\right\|_{\mathcal{X}}\right)$ und $\left(\mathcal{Y}, \left\|\bullet\right\|_{\mathcal{Y}}\right)$ lineare normierte Räume und $D \subseteq \mathcal{X}$ ein linearer Teilraum. Ein linearer Operator $T : D \to \mathcal{Y}$ heißt *abgeschlossen*, wenn für alle $f \in \mathcal{X}$, $g \in \mathcal{Y}$ und jede Folge $\left\{f_n\right\}_{n\in\mathbb{N}}$ in D mit $f_n \xrightarrow{s} f$ (d.h. $\left\|f_n - f\right\|_{\mathcal{X}} \to 0$) und $Tf_n \xrightarrow{s} g$ (d.h. $\left\|Tf_n - g\right\|_{\mathcal{Y}} \to 0$) folgt: $f \in D$ und $Tf = g$.

Lemma 40. Es seien $\left(\mathcal{X}, \|\bullet\|_{\mathcal{X}}\right)$ und $\left(\mathcal{Y}, \|\bullet\|_{\mathcal{Y}}\right)$ lineare normierte Räume. Dann gilt:

a) Ist $T \in \mathfrak{L}[\mathcal{X}, \mathcal{Y}]$ ein beschränkter linearer Operator, so ist *T* abgeschlossen.

b) Ist $D \subseteq \mathcal{X}$ ein abgeschlossener linearer Teilraum und $T \in \mathfrak{L}[D, \mathcal{Y}]$, so ist *T* abgeschlossen.

Beweis: Teil a) folgt sofort aus der zur Beschränktheit äquivalenten Stetigkeit von *T*, da der starke Grenzwert einer Folge eindeutig bestimmt ist. Teil b) ergibt sich analog, weil aus $f_n \xrightarrow{s} f$ für jede Folge $\{f_n\}_{n\in\mathbb{N}}$ in *D* sofort $f \in D$ auf Grund der Abgeschlossenheit von *D* folgt. ■

Bemerkung: es gibt abgeschlossene lineare Operatoren, die nicht stetig sind, wie Beispiel 12, Teil b) für den Teilraum $D = C^1[a,b] = \{ f \in C[a,b] \mid f' \in C[a,b]\} \subset C[a,b] = \mathcal{X}$ der auf $[a,b]$ stetig differenzierbaren Funktionen mit Norm $\|\bullet\|_C$ zeigt. Denn sei $\{f_n\}_{n\in\mathbb{N}}$ in D mit $f_n \xrightarrow{s} f$ und $Tf_n \xrightarrow{s} g$, so bedeutet dies, dass $\{f_n\}_{n\in\mathbb{N}}$ gleichmäßig gegen *f* und $\{f_n'\}_{n\in\mathbb{N}}$ gleichmäßig gegen *g* strebt. In diesem Fall dürfen Limes und Ableitung vertauscht werden, so dass *f* stetig differenzierbar (und damit in *D*) ist mit $\lim\limits_{n\to\infty} f_n' = g$, also der Ableitungsoperator tatsächlich abgeschlossen ist.

Definition 47 (Graph eines Operators)**.** Es seien $\left(\mathcal{X}, \|\bullet\|_{\mathcal{X}}\right)$ und $\left(\mathcal{Y}, \|\bullet\|_{\mathcal{Y}}\right)$ lineare normierte Räume und $D \subseteq \mathcal{X}$ ein linearer Teilraum sowie $T: D \to \mathcal{Y}$ ein (nicht notwendig linearer) Operator. Dann heißt die Menge

$$G_T := \{(f, Tf) \mid f \in D\} \subseteq \mathcal{X} \times \mathcal{Y}$$

Graph von T.

Bemerkung: Ist *T* linear, so ist offensichtlich G_T ein linearer Teilraum von $\mathcal{X} \times \mathcal{Y}$.

Lemma 41. Es seien $\left(\mathcal{X}, \|\bullet\|_{\mathcal{X}}\right)$ und $\left(\mathcal{Y}, \|\bullet\|_{\mathcal{Y}}\right)$ lineare normierte Räume und $D \subseteq \mathcal{X}$ ein linearer Teilraum sowie $T: D \to \mathcal{Y}$ ein linearer Operator. Dann gilt: *T* ist genau dann abgeschlossen, wenn der Graph von *T* abgeschlossen ist.

Beweis: Wir können $\mathcal{X} \times \mathcal{Y}$ mit der folgenden (oder einer anderen äquivalenten) Norm versehen:

$$\|(f,g)\|_{\mathcal{X}\times\mathcal{Y}} := \|f\|_{\mathcal{X}} + \|g\|_{\mathcal{Y}} \quad \text{für alle } (f,g) \in \mathcal{X} \times \mathcal{Y}.$$

"$\Rightarrow$": Sei $\{(f_n, Tf_n)\}_{n\in\mathbb{N}}$ eine Folge in G_T und $(f,g) \in \mathcal{X} \times \mathcal{Y}$ mit $(f_n, Tf_n) \xrightarrow{s} (f,g)$. Dann folgt nach Definition der obigen Norm $D \ni f_n \xrightarrow{s} f$ und $Tf_n \xrightarrow{s} g$. Da *T* abgeschlossen ist, folgt $f \in D$ und $Tf = g$, also $(f,g) \in G_T$. Damit ist G_T abgeschlossen.

"$\Leftarrow$": Sei $\{f_n\}_{n\in\mathbb{N}}$ Folge in D mit $f_n \xrightarrow{s} f$ und $Tf_n \xrightarrow{s} g$. Dann folgt nach Definition der obigen Norm $\|(f_n, Tf_n) - (f,g)\|_{\mathcal{X}\times\mathcal{Y}} = \|f_n - f\|_{\mathcal{X}} + \|Tf_n - g\|_{\mathcal{Y}} \to 0$, so dass $(f,g) \in G_T$ wegen der Abgeschlossenheit von G_T. Also ist $f \in D$ und $Tf = g$ und somit T abgeschlossen.

Lemma 42. Es seien $\left(\mathcal{X}, \|\bullet\|_{\mathcal{X}}\right)$ und $\left(\mathcal{Y}, \|\bullet\|_{\mathcal{Y}}\right)$ lineare normierte Räume und $D \subseteq \mathcal{X}$ ein linearer Teilraum. $T: D \to \mathcal{Y}$ sei ein abgeschlossener linearer Operator. Dann gilt: existiert der inverse Operator T^{-1}, so ist auch T^{-1} abgeschlossen.

Beweis: Die Linearität von T impliziert, dass $T(D)$ ein linearer Teilraum von $\mathcal{Y}$ ist. Ferner gilt:

$$G_T = \{(f, Tf) \mid f \in D\} = \left\{\left(T^{-1}g, g\right) \mid g \in T(D)\right\},$$

und dies ist abgeschlossen nach Lemma 41. Also ist auch

$$G_{T^{-1}} = \left\{\left(g, T^{-1}g\right) \mid g \in T(D)\right\}$$

abgeschlossen (Vertauschung der Komponenten) und damit nach Lemma 41 auch der Operator T^{-1}. ■

Lemma 43. Es sei $\left(\mathcal{X}, \|\bullet\|_{\mathcal{X}}\right)$ ein linearer normierter Raum und $\left(\mathcal{Y}, \|\bullet\|_{\mathcal{Y}}\right)$ ein Banach-Raum. Ferner sei $D \subseteq \mathcal{X}$ ein linearer Teilraum und $T: D \to \mathcal{Y}$ ein abgeschlossener und beschränkter linearer Operator. Dann ist D abgeschlossen.

Beweis: Es sei $\{f_n\}_{n\in\mathbb{N}}$ eine Folge in D und $f \in \mathcal{X}$ mit $f_n \xrightarrow{s} f$. Dann ist aber $\{Tf_n\}_{n\in\mathbb{N}}$ eine Cauchy-Folge in $\mathcal{Y}$ wegen

$$\|Tf_n - Tf_m\| \le \|T\| \|f_n - f_m\|_{\mathcal{X}} \le \|T\| \left(\|f_n - f\|_{\mathcal{X}} + \|f_m - f\|_{\mathcal{X}}\right) \text{ für alle } n, m \in \mathbb{N}.$$

Da $\mathcal{Y}$ voraussetzungsgemäß vollständig ist, existiert also ein $g \in \mathcal{Y}$ mit $Tf_n \xrightarrow{s} g$. Die Abgeschlossenheit von T impliziert nun $f \in D$ (mit $Tf = g$) und damit auch die Abgeschlossenheit von D. ■

Satz 39 (von der Beschränktheit der Umkehrtransformation)**.** Es sei $\left(\mathcal{X}, \|\bullet\|_{\mathcal{X}}\right)$ ein Banach-Raum und $\left(\mathcal{Y}, \|\bullet\|_{\mathcal{Y}}\right)$ ein linearer normierter Raum. Ferner sei $D \subseteq \mathcal{X}$ ein linearer Teilraum und $T: D \to \mathcal{Y}$ ein abgeschlossener und beschränkter linearer Operator mit der Eigenschaft, dass $T(D)$ von zweiter Kategorie in $\mathcal{Y}$ ist. Dann gilt:

a) $T(D) = \mathcal{Y}$, d.h. T ist surjektiv.
b) Es existiert eine reelle Zahl $m > 0$ derart, dass gilt: zu jedem $g \in \mathcal{Y}$ existiert ein $f \in \mathcal{X}$ mit $Tf = g$ und $\|f\|_{\mathcal{X}} \le m\|g\|_{\mathcal{Y}}$.
c) Existiert der inverse Operator T^{-1}, so ist T^{-1} beschränkt.

Beweis: Wir folgen der Beweisidee für den Teil a) des Satzes in BACHMAN / NARICI (2000), Theorem 16.5 oder TAYLOR (1963), Theorem 4.2-G. Ein wesentliches Zwischenergebnis ist hier, dass es ein universelles $\delta > 0$ gibt derart, dass zu jedem von $\mathbf{0}$ verschiedenen $g \in \mathcal{Y}$ ein $h \in \mathcal{X}$ mit

$\|h\|_{\mathcal{X}} \le 2$ und $g = T\left(\frac{4\|g\|_{\mathcal{Y}}}{\delta} h\right)$ existiert. Damit kann Teil b) gezeigt werden, denn mit der Wahl von $f = \frac{4\|g\|_{\mathcal{Y}}}{\delta} h$ ergibt sich

$$\|f\|_{\mathcal{X}} = \frac{4}{\delta}\|h\|_{\mathcal{X}}\|g\|_{\mathcal{Y}} \le \frac{8}{\delta}\|g\|_{\mathcal{Y}}.$$

Damit ist Teil b) gezeigt, mit $m = \frac{8}{\delta}$. Für $g = \mathbf{0}$ kann direkt $f = \mathbf{0}$ gewählt werden, so dass die Ungleichung $\|f\|_{\mathcal{X}} \le m\|g\|_{\mathcal{Y}}$ hier trivialerweise richtig ist. Teil c) folgt hieraus ebenfalls unmittelbar, da dieselbe Ungleichung nun

$$\|Tf\|_{\mathcal{Y}} \ge \frac{1}{m}\|f\| \text{ für alle } f \in \mathcal{X}$$

impliziert, also T^{-1} stetig ist nach Satz 20.

Es bleibt also nur noch Teil a) des Satzes bzw. das oben erwähnte Zwischenergebnis zu zeigen. Wir zerlegen den Beweis in 6 Schritte:

<u>1. Schritt:</u>

Für $r > 0$ seien

$$S_r := \bar{K}_r(\mathbf{0}) = \left\{ f \in \mathcal{X} \mid \|f\|_{\mathcal{X}} \le r \right\},\ D_r := D \cap S_r.$$

Dann ist $D = \bigcup_{n=1}^{\infty} D_n$, also $T(D) = \bigcup_{n=1}^{\infty} T(D_n)$. Da $T(D)$ von zweiter Kategorie ist, gibt es also ein $n_0 \in \mathbb{N}$, so dass $T\left(D_{n_0}\right)$ nicht nirgends dicht ist, also das Innere des Abschlusses $\overline{T\left(D_{n_0}\right)}^{o} \ne \varnothing$ ist. Demnach existiert ein $h_0 \in T\left(D_{n_0}\right)$ und ein $\delta_0 > 0$ mit

$$K_{\delta_0}(h_0) := \left\{ g \in \mathcal{Y} \mid \|g - h_0\|_{\mathcal{Y}} < \delta_0 \right\} \subseteq \overline{T\left(D_{n_0}\right)}.$$

<u>2. Schritt:</u>

Mit $K_{\delta_0}(h_0) \subseteq \overline{T\left(D_{n_0}\right)}$ gilt auch $K_{\delta_0/n_0}\left(\frac{1}{n_0} h_0\right) \subseteq \overline{T(D_1)}$ (vgl. das "scaling-down"-Argument in BACHMAN / NARICI (2000), S. 273). Mit $\delta_1 := \frac{\delta_0}{2n_0}$ folgt also: es gibt ein $h_1 \in T(D_1)$ mit $K_{\delta_1}(h_1) \subseteq \overline{T(D_1)}$ und somit auch ein $f_1 \in D_1$ mit $h_1 = Tf_1$. (dieses $\delta_1 = \frac{\delta_0}{2n_0}$ ist zugleich das obige "universelle" δ.)

<u>3. Schritt:</u>

Es gilt $\overline{T(D_1)} - h_1 \subseteq \overline{T(D_2)}$: Zunächst ist $D_1 - f_1 = \{ f - f_1 \mid f \in D_1 \} \subseteq D_2$, da für alle $f \in D_1$ gilt:

$$\|f - f_1\|_{\mathcal{X}} \leq \|f\|_{\mathcal{X}} + \|f_1\|_{\mathcal{X}} \leq 1+1=2,$$

also $f - f_1 \in S_2$. Andererseits gilt auch $D_1 - f_1 \subseteq D - f_1 = D$, also $D_1 - f_1 \subseteq D \cap S_2 = D_2$. Damit ergibt sich:

$$\overline{T(D_1)} - h_1 = \overline{T(D_1) - h_1} = \overline{T(D_1) - Tf_1} = \overline{T(D_1 - f_1)} \subseteq \overline{T(D_2)}.$$

<u>4. Schritt:</u>

Es gilt $K_{\delta_1}(\mathbf{0}) \subseteq \overline{T(D_2)}$: Für $g \in K_{\delta_1}(\mathbf{0})$ ist nämlich wegen $\|g\|_{\mathcal{Y}} = \|(g+h_1)-h_1\|_{\mathcal{Y}} < \delta_1$ jedenfalls $g + h_1 \in K_{\delta_1}(h_1) \subseteq \overline{T(D_1)}$ nach Schritt 2 und somit auch $g \in \overline{T(D_1)} - h_1 \subseteq \overline{T(D_2)}$ nach Schritt 3.

<u>5. Schritt:</u>

Es gilt $K_{\delta_1/2}(\mathbf{0}) \subseteq T(D_2) \subseteq T(D)$: Aus dem 4. Schritt erhält man zunächst

$$K_{\delta_1/2^{k+1}}(\mathbf{0}) \subseteq \overline{T\left(D_{1/2^k}\right)} \text{ für alle } k \in \mathbb{N}$$

(wieder mit dem "scaling-down"-Argument in BACHMAN / NARICI (2000), S. 273). Für gegebenes $g \in K_{\delta_1/2}(\mathbf{0})$ konstruieren wir nun induktiv eine Folge $\{v_k\}_{k\in\mathbb{N}}$ mit

$$v_k \in D_{1/2^{k-1}} \quad \text{und} \quad \left\| g - \sum_{j=1}^{k} T\left(v_j\right) \right\|_{\mathcal{Y}} < \frac{\delta_1}{2^{k+1}} \quad \text{für alle } k \in \mathbb{N}.$$

Das erste Folgenglied erhalten wir so: Wegen $g \in K_{\delta_1/2}(\mathbf{0}) \subseteq \overline{T(D_1)}$ existiert ein $w \in T(D_1)$ mit $\|g - w\|_{\mathcal{Y}} < \frac{\delta_1}{4}$, weil $T(D_1)$ in $\overline{T(D_1)}$ dicht liegt. Demnach gibt es ein $z \in D_1$ mit $w = Tz$. Wähle $v_1 := z$.

Seien nun $v_1, \cdots, v_k$ konstruiert mit $v_m \in D_{1/2^{m-1}}$ und $\left\| g - \sum_{j=1}^{m} T\left(v_j\right) \right\|_{\mathcal{Y}} < \frac{\delta_1}{2^{m+1}}$ für $1 \leq m \leq k \in \mathbb{N}$. Dann ist insbesondere

$$g - \sum_{j=1}^{k} T\left(v_j\right) \in K_{\delta_1/2^{k+1}}(\mathbf{0}) \subseteq \overline{T\left(D_{1/2^k}\right)},$$

also gibt es analog wie zuvor ein $w \in T\left(D_{1/2^k}\right)$ mit $\left\| g - \sum_{j=1}^{k} T\left(v_j\right) - w \right\|_{\mathcal{Y}} < \frac{\delta_1}{2^{k+2}}$. Demnach gibt es ein $z \in D_{1/2^k}$ mit $w = Tz$. Wähle $v_{k+1} := z$.

Mit Hilfe dieser Folge argumentieren wir so weiter: Setze $s_n := \sum_{k=1}^{n} v_k$ für alle $n \in \mathbb{N}$. Dann liegen alle $s_n \in D$ mit $\|s_n\|_{\mathcal{X}} \leq \sum_{k=1}^{n} \|v_k\|_{\mathcal{X}} \leq \sum_{k=1}^{n} \frac{1}{2^{k-1}} < 2$ für alle $n \in \mathbb{N}$, d.h. es ist $s_n \in D \cap S_2 = D_2$ für alle $n \in \mathbb{N}$. Ferner ist $\{s_n\}_{\in\mathbb{N}}$ Cauchy-Folge in $\mathcal{X}$, weil gilt:

$$\left\|s_n - s_m\right\|_{\mathcal{X}} \le \sum_{k=n+1}^{m} \left\|v_k\right\|_{\mathcal{X}} \le \sum_{k=n+1}^{m} \frac{1}{2^{k-1}} < \sum_{k=n+1}^{\infty} \frac{1}{2^{k-1}} = \frac{1}{2^{n-1}} \quad \text{für alle } n,m \in \mathbb{N},\ m > n.$$

Da $\mathcal{X}$ als Banach-Raum vollständig ist, existiert also ein $f \in \mathcal{X}$ mit $\lim\limits_{n\to\infty} \left\|s_n - f\right\|_{\mathcal{X}} = 0$ und

$$\|f\|_{\mathcal{X}} \le \lim_{n\to\infty} \left(\|f - s_n\|_{\mathcal{X}} + \|s_n\|_{\mathcal{X}}\right) \le 2,$$

d.h. es ist $f \in D_2$ mit

$$\lim_{n\to\infty} \left\|g - Ts_n\right\|_{\mathcal{Y}} = \lim_{n\to\infty} \left\|g - \sum_{j=1}^{n} T\left(v_j\right)\right\|_{\mathcal{Y}} \le \lim_{n\to\infty} \frac{\delta_1}{2^{n+1}} = 0.$$

Also haben wir die Beziehungen $s_n \xrightarrow{s} f$ und $Ts_n \xrightarrow{s} g$, woraus nach Voraussetzung der Abgeschlossenheit von T folgt: $f \in D$ (trivial) und $Tf = g$. Wegen $f \in D_2$ und $g \in K_{\delta_1/2}(\mathbf{0})$ folgt damit die obige Aussage $K_{\delta_1/2}(\mathbf{0}) \subseteq T\left(D_2\right) \subseteq T(D)$.

<u>6. Schritt:</u>

Es gilt $T(D) = \mathcal{Y}$: Nach dem 5. Schritt gilt für $g \in \mathcal{Y},\ g \ne \mathbf{0}$

$$\frac{\delta_1}{4\|g\|_{\mathcal{Y}}} g \in K_{\delta_1/2}(\mathbf{0}) \subseteq T\left(D_2\right).$$

Also existiert ein $f \in D_2$ mit $g = T\left(\dfrac{4}{\|g\|_{\mathcal{Y}}} f\right)$, womit $g \in T(D)$ folgt. Wegen $T(\mathbf{0}) = \mathbf{0}$ ergibt sich somit $T(D) = \mathcal{Y}$, wie gewünscht.

Der Beweis des Satzes ist damit vollständig erbracht. ■

Lemma 44. Es seien $\left(\mathcal{X}, \|\bullet\|_{\mathcal{X}}\right)$ und $\left(\mathcal{Y}, \|\bullet\|_{\mathcal{Y}}\right)$ Banach-Räume und $T:\ \mathcal{X} \to \mathcal{Y}$ ein linearer Operator. Dann gilt:

a) Ist T beschränkt und existiert der inverse Operator T^{-1} (auf ganz $\mathcal{Y}$), so ist T^{-1} beschränkt.
b) Ist T abgeschlossen und injektiv, so ist T beschränkt.

Beweis:

a) Nach Voraussetzung ist T surjektiv, also $T\left(\mathcal{X}\right) = \mathcal{Y}$ von zweiter Kategorie. Die Aussage folgt damit aus Lemma 40, Teil a) und Satz 39, Teil c).
b) Nach Voraussetzung existiert T^{-1} zumindest auf dem linearen Teilraum $T\left(\mathcal{X}\right) \subseteq \mathcal{Y}$. Nach Lemma 42 ist T^{-1} also (dort) abgeschlossen. Es ist jedenfalls voraussetzungsgemäß $T^{-1}\left(T\left(\mathcal{X}\right)\right) = \mathcal{X}$ von zweiter Kategorie, also kann Satz 39, Teil c) auf die Situation $T^{-1} : T\left(\mathcal{X}\right) \to \mathcal{X}$ angewendet werden, woraus wegen $T = \left(T^{-1}\right)^{-1}$ die Behauptung folgt. ■

Satz 40 (Satz vom abgeschlossenen Graphen; Closed Graph Theorem)**.** Es seien $\left(\mathcal{X}, \|\bullet\|_{\mathcal{X}}\right)$ und $\left(\mathcal{Y}, \|\bullet\|_{\mathcal{Y}}\right)$ Banach-Räume und $T: \mathcal{X} \to \mathcal{Y}$ ein abgeschlossener linearer Operator. Dann ist T beschränkt.

Beweis: Unter der im Beweis zu Lemma 41 definierten Norm wird $\mathcal{X} \times \mathcal{Y}$ zu einem Banach-Raum. Nach Lemma 41 ist der Graph G_T abgeschlossen, also ebenfalls ein Banach-(Teil-)Raum. Definiere $S : G_T \to \mathcal{X} : (f, Tf) \mapsto f$. Dann ist S linear und beschränkt wegen

$$\|S\| = \sup_{\|(f,Tf)\|_{G_T}=1} \|S(f,Tf)\| = \sup_{\|f\|_{\mathcal{X}}+\|Tf\|_{\mathcal{Y}}=1} \|f\|_{\mathcal{X}} \leq 1.$$

Wegen $S(G_T) = \mathcal{X}$ ist S surjektiv. S ist aber auch injektiv (und damit bijektiv), weil aus $f = S(f,Tf) = \mathbf{0}$ folgt: $f = \mathbf{0}$ und damit auch $Tf = \mathbf{0}$, also $(f,Tf) = \mathbf{0}$ (in G_T). Damit existiert $S^{-1} : \mathcal{X} \to G_T : f \mapsto (f,Tf)$. Nach Lemma 44, Teil a) ist also S^{-1} beschränkt, d.h. es gilt:

$$\|Tf\|_{\mathcal{Y}} \leq \|Tf\|_{\mathcal{Y}} + \|f\|_{\mathcal{X}} = \|(f,Tf)\|_{G_T} = \left\|S^{-1}(f)\right\|_{G_T} \leq \left\|S^{-1}\right\| \|f\|_{\mathcal{X}} \quad \text{für alle } f \in \mathcal{X},$$

d.h. T ist beschränkt mit $\|T\| \leq \left\|S^{-1}\right\|$. ■

Die Ergebnisse dieses Abschnitts können nun folgendermaßen zusammengefasst werden:

Satz 41. Es seien $\left(\mathcal{X}, \|\bullet\|_{\mathcal{X}}\right)$ und $\left(\mathcal{Y}, \|\bullet\|_{\mathcal{Y}}\right)$ Banach-Räume und $D \subseteq \mathcal{X}$ ein linearer Teilraum. Ferner sei $T: D \to \mathcal{Y}$ ein linearer Operator. Es bezeichne

$\boxed{1}$ T ist abgeschlossen

$\boxed{2}$ D ist abgeschlossen

$\boxed{3}$ T ist beschränkt.

Dann gilt: aus je zwei dieser Eigenschaften folgt die dritte.

Beweis:

Aus $\boxed{1}$ und $\boxed{2}$ folgt $\boxed{3}$: Satz 40 (da bei Abgeschlossenheit auch D ein Banach-Raum ist mit gleicher Norm wie $\mathcal{X}$)

Aus $\boxed{1}$ und $\boxed{3}$ folgt $\boxed{2}$: Lemma 43

Aus $\boxed{2}$ und $\boxed{3}$ folgt $\boxed{1}$: Lemma 40, Teil b). ■

II Halbgruppen von Operatoren

Die Theorie der Halbgruppen linearer Operatoren ist eng verknüpft mit Fragestellungen der Physik, die insbesondere auch stochastische Aspekte berühren, wie das Problem der Wärmeleitung in festen Körpern, das allgemeiner der physikalischen Diffusionstheorie zuzurechnen ist. Dieses Problem hat schon Albert Einstein 1905 durch partielle Differentialgleichungen zweiter Ordnung beschrieben; es hat einen unmittelbaren Zusammenhang zur Brown'schen Bewegung bzw. dem Wiener-Prozess und damit auch zur modernen Stochastischen Finanzmathematik, was u.a. erklärt, warum in Banken in diesem Bereich heutzutage verstärkt Physiker tätig sind.

Eine systematische Mathematisierung des Gebiets erfolgte etwa ab 1930, vor allem durch Einar Hille, Ralph Phillips und Kôsaku Yosida; es nimmt heute innerhalb der Funktionalanalysis durchaus einen eigenständigen Raum ein (vgl. etwa die Monographien von BUTZER / BERENS (1967) oder GOLDSTEIN (1985)).

Für das Folgende wollen wir stets annehmen, dass $\left(\mathcal{X}, \|\bullet\|_{\mathcal{X}}\right)$ ein reeller oder komplexer Banach-Raum ist. Die folgende Definition ist im Zusammenhang mit Definition 36 zu sehen.

Definition 47 (Halbgruppe von Operatoren)**.** Eine Familie $\mathfrak{T} = \left\{T(t) \mid t \geq 0\right\} \subseteq \mathfrak{L}[\mathcal{X}, \mathcal{X}]$ beschränkter linearer Operatoren heißt *(Operator-)Halbgruppe*, wenn gilt:

a) $T(s) \circ T(t) = T(s+t)$ für alle $s, t \geq 0$
b) $T(0) = \mathbf{I}$.

Die Operator-Halbgruppe $\mathfrak{T}$ heißt *von der Klasse* $\left(C_0\right)$, wenn zusätzlich gilt:

c) $T(t) f \xrightarrow{s} f$ bei $t \downarrow 0$, für alle $f \in \mathcal{X}$

(d.h. die Abbildung $T(t) f$ ist für jedes $f \in \mathcal{X}$ stark stetig in $t = 0$).

Lemma 45. Für eine Operator-Halbgruppe der Klasse $\left(C_0\right)$ bestehen folgende Eigenschaften:

a) Die Operator-Normen $\|T(t)\|$ sind in jedem endlichen Teilintervall von $\mathbb{R}^+$ beschränkt.
b) Die Abbildung $T(t) f$ ist für jedes $f \in \mathcal{X}$ stark stetig in jedem Punkt $t = t_0 \in (0, \infty)$.
c) Es gilt

$$\omega_0 := \inf_{t>0}\left\{\frac{1}{t}\ln\|T(t)\|\right\} = \lim_{t\to\infty}\left\{\frac{1}{t}\ln\|T(t)\|\right\} \in [-\infty, \infty)$$

d) Zu jedem $\omega > \omega_0$ existiert ein $M(\omega) \geq 1$ mit

$$\|T(t)\| \leq M(\omega) e^{\omega t} \text{ für alle } t \geq 0.$$

Beweis:

a) Wir zeigen zunächst:

$$\exists\, t_0 > 0,\ M \geq 1:\ \|T(t)\| \leq M \quad \text{für alle } t \in [0, t_0].$$

Anderenfalls gäbe es eine Folge $\{t_n\}_{n\in\mathbb{N}} \subset \mathbb{R}^+$ mit $t_n \downarrow 0$ und $\|T(t_n)\| \geq n$ für alle $n \in \mathbb{N}$. Nach dem Uniform-Boundedness-Theorem (Satz 22) wäre dann mit $T_n := T(t_n)$, $n \in \mathbb{N}$ die Folge $\{\|T_n f\|_{\mathcal{X}}\}_{n\in\mathbb{N}}$ für wenigstens ein $f \in \mathcal{X}$ unbeschränkt im Widerspruch zu Definition 47, Teil c).
Für ein beliebiges $t > 0$ existiert nun eine Darstellung $t = m t_0 + \tau$ mit $m \in \mathbb{Z}^+$ und $\tau \in [0, t_0)$, so dass folgt:

$$\|T(t)\| = \|T(m t_0 + \tau)\| \leq \|T(m t_0)\| \cdot \|T(\tau)\| \leq \|T(t_0)\|^m \cdot \|T(\tau)\| \leq M^m M \leq M \cdot M^{t/t_0} = Me^{\eta t}$$

für $\eta := \dfrac{\ln M}{t_0}$. Hieraus folgt die Aussage.

b) Es sei $t > 0$ beliebig. Dann gilt für alle $f \in \mathcal{X}$:

$$\|T(t+s)f - T(t)f\| = \|T(t) \circ (T(s) - \mathbf{I}) f\| \leq \|T(t)\| \cdot \|(T(s) - \mathbf{I}) f\| \to 0 \text{ für } s \downarrow 0,$$

d.h. die Halbgruppe ist in *t* (stark) rechtsseitig stetig. Ferner gilt für $0 < s < t$:

$$\|T(t-s)f - T(t)f\| = \|T(t-s) \circ (T(s) - \mathbf{I}) f\| \leq \|T(t-s)\| \cdot \|(T(s) - \mathbf{I}) f\| \to 0 \text{ für } s \downarrow 0,$$

d.h. die Halbgruppe ist in *t* auch (stark) linksseitig stetig und damit insgesamt (stark) stetig.

c) Für $t > 0$ sei $p(t) := \ln\|T(t)\|$. Nach dem Beweis zu Teil a) oben folgt dann

$$-\infty \leq p(t) \leq \ln\left(Me^{\eta t}\right) = \eta t + \ln M < \infty \quad \text{für alle } t > 0 \text{ mit}$$

$$-\infty \leq \frac{p(t)}{t} \leq \eta + \frac{\ln M}{t} < \infty \text{ für alle } t > 0.$$

Damit erhält man zunächst

$$\omega_0 = \inf_{t>0} \left\{ \frac{1}{t} \ln\|T(t)\| \right\} \in [-\infty, \infty).$$

Aus der Halbgruppeneigenschaft Teil a), Definition 47 ergibt sich ferner

$$p(s+t) = \ln\|T(s+t)\| \leq \ln\left(\|T(s)\| \cdot \|T(t)\|\right) = p(s) + p(t) \text{ für alle } s, t > 0.$$

Sei nun zunächst $\omega_0 > -\infty$. Zu jedem $\varepsilon > 0$ existiert dann definitionsgemäß ein $t_\varepsilon > 0$ mit $\dfrac{p(t_\varepsilon)}{t_\varepsilon} \leq \omega_0 + \varepsilon$. Zu beliebigem $t > 0$ existiert dann wie im Beweis zu Teil a) wieder eine Darstellung $t = m t_\varepsilon + \tau$ mit $m \in \mathbb{Z}^+$ und $\tau \in [0, t_\varepsilon)$, so dass

$$p(t) = p(mt_\varepsilon + \tau) \leq mp(t_\varepsilon) + p(\tau)$$

folgt. Mit der obigen ersten Abschätzung ergibt sich ferner

$$p(\tau) = \ln\|T(\tau)\| \leq \eta\tau + \ln M.$$

Wegen $\dfrac{mt_\varepsilon}{t} \leq 1$ folgt nun:

$$\omega_0 \le \frac{p(t)}{t} \le \frac{mp(t_\varepsilon)+p(\tau)}{t} \le \frac{m\,t_\varepsilon}{t}\frac{p(t_\varepsilon)}{t_\varepsilon}+\frac{p(\tau)}{t} \le \frac{p(t_\varepsilon)}{t_\varepsilon}+\frac{\eta\tau+\ln M}{t}$$
$$\le \omega_0+\varepsilon+\frac{\eta t_\varepsilon+\ln M}{t} \to \omega_0+\varepsilon \quad \text{für } t\to\infty,$$

so dass

$$\omega_0 \le \limsup_{t\to\infty}\frac{p(t)}{t} \le \omega_0+\varepsilon \quad \text{für alle } \varepsilon>0$$

und damit die Behauptung resultiert. Im Fall $\omega_0=-\infty$ argumentiert man analog so: Zu jedem $N>0$ existiert ein $t_N>0$ mit $\frac{p(t_N)}{t_N}\le -N.$ Mit einer analogen Darstellung $t=m\,t_N+\tau$ mit $m\in\mathbb{Z}^+$ und $\tau\in[0,t_N)$ ergibt sich dann

$$\frac{p(t)}{t} \le \frac{mp(t_N)+p(\tau)}{t} \le \frac{m\,t_N}{t}\frac{p(t_N)}{t_N}+\frac{p(\tau)}{t} \le \frac{p(t_N)}{t_N}+\frac{\eta\tau+\ln M}{t}$$
$$\le -N+\frac{\eta t_\varepsilon+\ln M}{t} \to -N \quad \text{für } t\to\infty,$$

also auch hier $\lim_{t\to\infty}\left\{\frac{1}{t}\ln\|T(t)\|\right\}=-\infty=\omega_0.$

d) Zu $\omega>\omega_0$ gibt es nach Teil c) ein $t_\omega>0$ mit $\frac{p(t)}{t}<\omega$ für alle $t>t_\omega,$ so dass

$$\|T(t)\|=e^{\ln\|T(t)\|}=e^{p(t)}<e^{\omega t} \quad \text{für alle } t>t_\omega$$

resultiert. Für $t\in[0,t_\omega]$ existiert nach Teil a) ein $M\ge 1$ mit $\|T(t)\|\le M.$ Insgesamt ergibt sich:

$$\|T(t)\|\le M(\omega)e^{\omega t} \quad \text{für alle } t\ge 0$$

mit $M(\omega):=\begin{cases} M, & \omega\ge 0 \\ Me^{-\omega t_\omega}, & \omega<0. \end{cases}$ Damit ist das Lemma bewiesen. ■

Beispiel 22 (Translations-Halbgruppe)**.** Wir bezeichnen mit $UCB(\mathbb{R})$[23] den Banach-Raum der auf $\mathbb{R}$ gleichmäßig stetigen beschränkten Funktionen mit der üblichen Supremums-Norm

$$\|f\|_{UCB}:=\sup\left\{|f(x)|\,\middle|\,x\in\mathbb{R}\right\} \quad \text{für } f\in UCB(\mathbb{R}).$$

Die durch

$$T(t)f(x):=f(x+t),\ x\in\mathbb{R},\ t\ge 0 \quad \text{für } f\in UCB(\mathbb{R})$$

definierte Halbgruppe der Klasse (C_0) heißt *Halbgruppe der Linkstranslationen*. Hier gilt für alle $t\ge 0$ die Gleichheit $\|T(t)f\|_{UCB}=\|f\|_{UCB}$ für alle $f\in UCB(\mathbb{R})$ und somit $\|T(t)\|=1$ für alle $t\ge 0.$ Dies entspricht der Situation $\omega_0=0$ in Lemma 45.

[23] Die Abkürzung steht für "**u**niformly **c**ontinuous and **b**ounded".

Beispiel 23. Es sei $\left(\mathcal{X},\|\bullet\|_{\mathcal{X}}\right)$ ein beliebiger Banach-Raum und $\lambda\in\mathbb{R}$ fest gewählt. Für die durch

$$T(t)f := e^{\lambda t}\cdot f,\ t\geq 0 \ \text{ für } f\in\mathcal{X}$$

definierte Halbgruppe der Klasse $\left(C_0\right)$ gilt offensichtlich: $\|T(t)f\|_{\mathcal{X}} = e^{\lambda t}\|f\|_{\mathcal{X}}$ und somit $\|T(t)\| = e^{\lambda t}$ für alle $t\geq 0$. Hier ist also $\omega_0=\lambda$. Zum Nachweis der $\left(C_0\right)$-Eigenschaft beachte: $\lim_{t\downarrow 0}\|T(t)f-f\|_{\mathcal{X}} = \left\{\lim_{t\downarrow 0}\left|e^{\lambda t}-1\right|\right\}\|f\|_{\mathcal{X}} = 0$ für $f\in\mathcal{X}$.

Beispiel 24. Wir betrachten den Banach-Raum $\mathcal{X}=\mathbb{R}^2$ mit der $\ell^1(2)$-Norm. Dann wird für feste $\mu,\lambda\in\mathbb{R}$ durch

$$T(t)f := \frac{1}{2}\begin{bmatrix} e^{\mu t}+e^{\lambda t} & e^{\mu t}-e^{\lambda t} \\ e^{\mu t}-e^{\lambda t} & e^{\mu t}+e^{\lambda t}\end{bmatrix}\cdot\begin{bmatrix} f_1 \\ f_2\end{bmatrix},\ t\geq 0 \ \text{ für alle } f=\begin{bmatrix} f_1 \\ f_2\end{bmatrix}\in\mathcal{X}$$

eine Halbgruppe der Klasse $\left(C_0\right)$ definiert. Dies folgt aus der leicht nachprüfbaren Beziehung

$$\begin{bmatrix} e^{\mu s}+e^{\lambda s} & e^{\mu s}-e^{\lambda s} \\ e^{\mu s}-e^{\lambda s} & e^{\mu s}+e^{\lambda s}\end{bmatrix}\cdot\begin{bmatrix} e^{\mu t}+e^{\lambda t} & e^{\mu t}-e^{\lambda t} \\ e^{\mu t}-e^{\lambda t} & e^{\mu t}+e^{\lambda t}\end{bmatrix} = 2\begin{bmatrix} e^{\mu(s+t)}+e^{\lambda(s+t)} & e^{\mu(s+t)}-e^{\lambda(s+t)} \\ e^{\mu(s+t)}-e^{\lambda(s+t)} & e^{\mu(s+t)}+e^{\lambda(s+t)}\end{bmatrix}$$

für $s,t\geq 0$ sowie

$$\lim_{t\downarrow 0}\frac{1}{2}\begin{bmatrix} e^{\mu t}+e^{\lambda t} & e^{\mu t}-e^{\lambda t} \\ e^{\mu t}-e^{\lambda t} & e^{\mu t}+e^{\lambda t}\end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1\end{bmatrix} = \mathbf{I}.$$

Hier gilt

$$\|T(t)f\|_1 \leq \frac{1}{2}\left\{\left|e^{\mu t}+e^{\lambda t}\right|+\left|e^{\mu t}-e^{\lambda t}\right|\right\}\|f\|_1 \ \text{ für } t\geq 0 \ \text{ und } f\in\mathcal{X},$$

wobei Gleichheit bei geeigneter Wahl der Vorzeichen von f_1 und f_2 erreicht werden kann. Die $\left(C_0\right)$-Eigenschaft folgt aus $\lim_{t\downarrow 0}\|T(t)f-f\|_1 \leq \frac{1}{2}\left\{\lim_{t\downarrow 0}\left|e^{\lambda t}+e^{\mu t}-2\right|+\lim_{t\downarrow 0}\left|e^{\lambda t}-e^{\mu t}\right|\right\}\|f\|_1 = 0,\ f\in\mathcal{X}.$

Hieraus ergibt sich

$$\|T(t)\| = \begin{cases} e^{\mu t}, & \lambda\leq\mu \\ e^{\lambda t}, & \lambda\geq\mu\end{cases} \ \text{ für } t\geq 0 \quad \text{und somit } \omega_0 = \begin{cases} \mu, & \lambda\leq\mu \\ \lambda, & \lambda\geq\mu\end{cases} = \max\left\{\mu,\lambda\right\}.$$

Beispiel 25. Es sei $UCB\left(\mathbb{R}^+\right)$ der Banach-Raum der auf $\mathbb{R}^+$ gleichmäßig stetigen beschränkten Funktionen (mit der gleichen Norm wie $UCB(\mathbb{R})$). Wir betrachten den Teilraum

$$\mathcal{X} := \left\{f\in UCB\left(\mathbb{R}^+\right)\,\middle|\, f(x)=0 \text{ für alle } x\geq 1\right\}$$

und hier die Translations-Halbgruppe wie in Beispiel 22:

$$T(t)f(x) := f(x+t),\ x\geq 0,\ t\geq 0 \ \text{ für } f\in\mathcal{X}.$$

Im Gegensatz zu vorher ist jetzt aber $T(t)f=\mathbf{0}$ für $f\in\mathcal{X}$ und $t\geq 1$, also gilt $\omega_0=-\infty$.

II. 1. Der infinitesimale Erzeuger

In diesem Abschnitt soll näher untersucht werden, ob und in welchem Sinn für gewisse $f \in \mathcal{X}$ "der" Grenzwert $Af := \lim\limits_{t \downarrow 0} \frac{1}{t}\left(T(t) - \mathbf{I}\right) f$ existiert. Man gelangt damit zum Begriff des *infinitesimalen Erzeugers A* einer Halbgruppe, der i. Allg. nur auf einem Teilraum $D \subseteq \mathcal{X}$ definiert und dort auch i. Allg. nur abgeschlossen und nicht beschränkt ist. Allerdings spielt er für die Charakterisierung von Halbgruppen eine fundamentale Rolle, gerade auch im Zusammenhang mit Markoff-Prozessen (wie dem Wiener-Prozess), und hat damit z.B. Auswirkungen auf das Berechnen von Optionspreisen in zeitstetigen Finanzmarktmodellen.

Ein wesentliches Hilfsmittel zur Behandlung des infinitesimalen Erzeugers sind gewisse verallgemeinerte Differentiations- und Integrationstechniken, die deshalb zuerst im Einzelnen ausführlicher behandelt werden.

II. 1.1. Abstrakte Differentiation und Integration

In diesem Abschnitt werden wir kurz die Begriffe der starken und schwachen Stetigkeit, Differenzierbarkeit und Integrierbarkeit von Abbildungen mit Werten in Banach-Räumen behandeln. Die Darstellung orientiert sich an dem Appendix, Teil II in BUTZER / BEHRENS (1967).

Definition 48 (starke / schwache Stetigkeit)**.** Es sei $E \subseteq \mathbb{R}$ ein (ggf. unendliches) Intervall, $\left(\mathcal{X}, \|\bullet\|_{\mathcal{X}}\right)$ ein Banach-Raum und $f : E \to \mathcal{X}$ eine Abbildung. f heißt

a) *stark stetig* in $t_0 \in E,$ wenn gilt:

$$\lim_{t \to t_0} \left\| f(t) - f\left(t_0\right) \right\|_{\mathcal{X}} = 0;$$

b) *schwach stetig* in $t_0 \in E,$ wenn gilt:

$$\lim_{t \to t_0} \left| f^*\left(f(t)\right) - f^*\left(f\left(t_0\right)\right) \right| = 0 \text{ für alle } f^* \in \mathcal{X}^*.$$

Ist $\left(\mathcal{Y}, \|\bullet\|_{\mathcal{Y}}\right)$ ein weiterer Banach-Raum und $T : E \to \mathfrak{L}[\mathcal{X}, \mathcal{Y}],$ so heißt T

c) *gleichmäßig stetig* in $t_0 \in E,$ wenn gilt:

$$\lim_{t \to t_0} \left\| T(t) - T\left(t_0\right) \right\| = 0;$$

d) *stark stetig* in $t_0 \in E,$ wenn gilt:

$$\lim_{t \to t_0} \left\| T(t) f - T\left(t_0\right) f \right\|_{\mathcal{Y}} = 0 \text{ für alle } f \in \mathcal{X};$$

e) *schwach stetig* in $t_0 \in E,$ wenn gilt:

$$\lim_{t \to t_0} \left| f^*\left(T(t) f\right) - f^*\left(T\left(t_0\right) f\right) \right| = 0 \text{ für alle } f \in \mathcal{X} \text{ und } f^* \in \mathcal{Y}^*.$$

Definition 49 (starke / schwache Differenzierbarkeit)**.** Es sei $E \subseteq \mathbb{R}$ ein (ggf. unendliches) Intervall, $\left(\mathcal{X}, \|\bullet\|_{\mathcal{X}}\right)$ ein Banach-Raum und $f : E \to \mathcal{X}$ eine Abbildung. f heißt

a) *stark differenzierbar* in $t_0 \in E,$ wenn ein $g_0 \in \mathcal{X}$ existiert mit

$$\lim_{h\to 0}\left\|\frac{1}{h}\left\{f\left(t_0+h\right)-f\left(t_0\right)\right\}-g_0\right\|_{\mathcal{X}}=0;$$

g_0 heißt dann auch *starke Ableitung von f im Punkt* t_0, in Zeichen: $g_0 = \frac{\mathbf{s}d}{dt} f\left(t_0\right);$

b) *schwach differenzierbar* in $t_0 \in E,$ wenn ein $g_0 \in \mathcal{X}$ existiert mit

$$\lim_{h\to 0}\left|\frac{1}{h}\left\{f^*\left(f\left(t_0+h\right)\right)-f^*\left(f\left(t_0\right)\right)\right\}-f^*\left(g_0\right)\right|=0 \text{ für alle } f^* \in \mathcal{X}^*;$$

g_0 heißt dann auch *schwache Ableitung von f im Punkt* t_0, in Zeichen: $g_0 = \frac{\mathbf{w}d}{dt} f\left(t_0\right).$

Ist $\left(\mathcal{Y}, \|\bullet\|_{\mathcal{Y}}\right)$ ein weiterer Banach-Raum und $T : E \to \mathfrak{L}[\mathcal{X},\mathcal{Y}],$ so heißt T

c) *gleichmäßig differenzierbar* in $t_0 \in E,$ wenn ein $G_0 \in \mathfrak{L}[\mathcal{X},\mathcal{Y}]$ existiert mit

$$\lim_{h\to 0}\left\|\frac{1}{h}\left\{T\left(t_0+h\right)-T\left(t_0\right)\right\}-G_0\right\|=0;$$

G_0 heißt dann auch *gleichmäßige Ableitung von T im Punkt* t_0, in Zeichen: $G_0 = \frac{\mathbf{u}d}{dt} T\left(t_0\right);$[24]

d) *stark differenzierbar* in $t_0 \in E,$ wenn gilt:

$$\lim_{t\to t_0}\left\|\frac{1}{h}\left\{T\left(t_0+h\right)-T\left(t_0\right)\right\}f-G_0 f\right\|_{\mathcal{Y}}=0 \text{ für alle } f \in \mathcal{X};$$

G_0 heißt dann auch *starke Ableitung von T im Punkt* t_0, in Zeichen: $G_0 = \frac{\mathbf{s}d}{dt} T\left(t_0\right);$

e) *schwach differenzierbar* in $t_0 \in E,$ wenn gilt:

$$\lim_{t\to t_0}\left|\frac{1}{h}f^*\left[\left\{T\left(t_0+h\right)-T\left(t_0\right)\right\}f\right]-f^*\left(G_0 f\right)\right|=0 \text{ für alle } f \in \mathcal{X} \text{ und } f^* \in \mathcal{Y}^*;$$

G_0 heißt dann auch *schwache Ableitung von T im Punkt* t_0, in Zeichen: $G_0 = \frac{\mathbf{w}d}{dt} T\left(t_0\right).$

Bemerkung: Sofern keine Missverständnisse entstehen (können), wird die starke (schwache) Ableitung auch nur mit $\frac{d}{dt} f\left(t_0\right)$ bezeichnet.

[24] Die Symbole **u**, **s** und **w** stehen hier für *uniform(ly)*, *strong(ly)* und *weak(ly)*.

Bemerkung: Beim Begriff der Stetigkeit bzw. der Differenzierbarkeit gilt jeweils, dass die gleichmäßige die starke und diese die schwache Eigenschaft impliziert; jedoch gilt i. Allg. nicht die Umkehrung. Beispielsweise impliziert (mit der Notation von Definition 48) die starke Stetigkeit die schwache wegen

$$\lim_{t\to t_0}\left|f^*\left(f(t)\right)-f^*\left(f\left(t_0\right)\right)\right| \le \left\|f^*\right\|\cdot\lim_{t\to t_0}\left\|f(t)-f\left(t_0\right)\right\|_{\mathcal{X}}=0 \text{ für alle } f^*\in\mathcal{X}^*.$$

Beispiel 26. Wir betrachten für $1\le p\le\infty$ den Raum $\mathcal{X}:=L^p\left[-1,1\right]$[25] mit

$$g(x;t):=\mathbb{1}_{(t,1]}(x),\ x\in\left[-1,1\right],\, t\in\left(0,1\right).$$

Dann definiert mit $E:=\left(0,1\right)$ die Abbildung

$$f:t\mapsto g(\,\bullet\,;t)=\mathbb{1}_{(t,1]}$$

eine Abbildung von E nach $\mathcal{X}$. Diese Abbildung ist für $1\le p<\infty$ stark und damit auch schwach stetig in $t_0:=0,$ denn es gilt:

$$\lim_{t\to0}\left\|f(t)-f\left(0\right)\right\|_{\mathcal{X}}=\lim_{t\to0}\left\|-\mathbb{1}_{(0,t]}\right\|_p=\lim_{t\to0}t^{1/p}=0.$$

Für $p=\infty$ ist diese Abbildung dagegen nicht stark stetig wegen

$$\left\|f(t)-f\left(0\right)\right\|_\infty=\left\|-\mathbb{1}_{(0,t]}\right\|_\infty=1 \text{ für alle } t\in\left(0,1\right).$$

Sie ist aber auch nicht schwach stetig; der entsprechende Nachweis ist jedoch nicht-trivial, denn es gilt z.B. für alle $h\in L^1\left[-1,1\right]"\subset"\mathcal{X}^*$ (vgl. Abschnitt I.3.2)

$$\lim_{t\to0}\left|\int_{-1}^{1}\mathbb{1}_{(0,t]}(x)\,h(x)\,dx\right|\le\lim_{t\to0}\int_0^t\left|h(x)\right|dx=0.$$

Für einen exakten Beweis ist der Dualraum $\mathcal{X}^*$ genau zu bestimmen. Dieser ist aber isometrisch isomorph zu einem gewissen Banach-Raum signierter Inhalte (engl.: *charges*) über $\left[-1,1\right]$ (siehe etwa TAYLOR (1963), Abschnitt 7.8). Beispielsweise existiert im üblichen Sinne ein Inhalt ν über $\left[-1,1\right]$ mit der Eigenschaft

$$\nu\left(\left[-1,0\right]\right)=0,\ \nu\left(\left(0,b\right]\right)=1 \text{ und } \nu\left(\left(b,1\right]\right)=0 \text{ für alle } b\in\left(0,1\right].$$

Dieser ist offensichtlich nicht σ-additiv wegen $\varnothing=\bigcap_{n=1}^{\infty}\left(0,\frac{1}{n}\right]$ mit $\nu\left(\varnothing\right)=0,$ aber $\lim_{n\to\infty}\nu\left(\left(0,\frac{1}{n}\right]\right)=1.$

Hier gilt nun

$$\int_{-1}^{1}\mathbb{1}_{(0,t]}\,d\nu=\nu\left(\left(0,t\right]\right)=1 \text{ für alle } t\in\left(0,1\right),$$[26]

also

---

[25] Genauer müsste hier der Banach-Raum $\mathfrak{L}^p\left[-1,1\right]$ betrachtet werden; vgl. dazu die Diskussion in Abschnitt I.1.5.

[26] Zur Definition eines (Lebesgue-)Integrals bzgl. eines signierten Inhalts vgl. TAYLOR (1963), Abschnitte 7.8 und 7.9.

$$\lim_{t\to 0}\left|f^*\left(f(t)\right)-f^*\left(f(0)\right)\right|=1$$

für das beschränkte lineare Funktional

$$f^*(h)=\int_{-1}^{1} h\,d\nu,\;\; h\in L^\infty[-1,1].$$

Damit ist aber die Abbildung $f(t)$ nicht schwach stetig für $p=\infty$.

Definiert man als weiteres Beispiel oben

$$g(x;t):=|x-t|,\;\; x\in[-1,1],\, t\in(0,1)$$

und wieder mit $E:=(0,1)$ die Abbildung

$$f:t\mapsto g(\,\bullet\,;t),$$

so ist diese Abbildung für $1\le p<\infty$ stark und damit auch schwach differenzierbar in $t_0:=0$ mit

$$g_0=\frac{\mathbf{s}d}{dt}f(0)=2\cdot\mathbb{1}_{[0,1]}-1\quad\text{(fast überall)},$$

denn es gilt:

$$\lim_{h\to 0}\left\|\frac{1}{h}\{f(h)-f(0)\}-g_0\right\|_p^p=\lim_{h\to 0}\left|\int_{-h}^{0}\left\{\frac{2(x+h)}{h}\right\}^p dx\right|=\lim_{h\to 0}\frac{2^p h}{p+1}=0.$$

Für $p=\infty$ ist diese Abbildung dagegen nicht stark differenzierbar wegen

$$\lim_{h\to 0}\left\|\frac{1}{h}\{f(h)-f(0)\}-g_0\right\|_\infty=\lim_{h\to 0}\left\{\operatorname*{ess\,sup}_{-h\le x\le 0}\left|\frac{2(x+h)}{h}\right|\right\}=2.$$

Mit ähnlichen Argumenten wie oben kann man auch hier zeigen, dass $f(t)$ nicht schwach im Nullpunkt differenzierbar ist, wobei allerdings wieder eine Integration nach signierten Inhalten notwendig ist.

Der aus der klassischen Analysis bekannte Sachverhalt, dass Differenzierbarkeit die Stetigkeit impliziert, gilt mit den allgemeineren Begriffen auch hier.

Lemma 46. Unter den Voraussetzungen von Definition 49 gilt: Ist die Abbildung $f(t)$ in $t_0\in E$ stark (schwach) differenzierbar, so ist $f(t)$ dort auch stark (schwach) stetig.

Beweis: Die starke Stetigkeit folgt aus der Ungleichung

$$\lim_{t\to t_0}\left\|f(t)-f(t_0)\right\|_{\mathcal{X}}=\lim_{h\to 0}\left\|h\cdot\left\{\frac{1}{h}\left[f(t_0+h)-f(t_0)\right]-g_0\right\}+h\cdot g_0\right\|_{\mathcal{X}}=0;$$

analog argumentiert man bei schwacher Stetigkeit. ■

Lemma 47. Unter den Voraussetzungen von Definition 49 gilt: Ist die Abbildung $f(t)$ in jedem Punkt $t \in E$ stark bzw. schwach differenzierbar mit $\frac{\mathbf{w}d}{dt} f = \mathbf{0}$, so existiert ein $g \in \mathcal{X}$ mit $f(t) = g$ für alle $t \in E$.

Beweis: Nach Definition der schwachen Ableitung ist für jedes beschränkte lineare Funktional $f^* \in \mathcal{X}^*$ die Funktion $f^*\left(f(\bullet)\right)$ im üblichen Sinne differenzierbar mit

$$\frac{d}{dt} f^*\left(f(t)\right) = f^*\left(\frac{\mathbf{w}d}{dt} f(t)\right) = f^*\left(\mathbf{0}\right) = 0 \text{ für alle } t \in E.$$

Also ist die Abbildung $f^*\left(f(\bullet)\right)$ auf E konstant. Für festes $t_0 \in E$ ist somit $f^*\left(f(t)\right) = f^*\left(f\left(t_0\right)\right)$ für alle $t \in E$ bzw.

$$f^*\left(f(t)\right) - f^*\left(f\left(t_0\right)\right) = f^*\left(f(t) - f\left(t_0\right)\right) = 0 \text{ für alle } t \in E$$

und somit $f(t) - f\left(t_0\right) = \mathbf{0}$ bzw. $f(t) = f\left(t_0\right) =: g$ für alle $t \in E$ nach Lemma 26. ■

Definition 50 (abstraktes Riemann- Integral). Gegeben seien die Voraussetzungen von Definition 49, wobei zusätzlich $E = [a,b]$ mit reellen $a < b$ ein kompaktes Intervall und $f(t)$ stark (schwach) stetig auf E sei. Für jede Zerlegung $\Delta_n \in \mathfrak{D}[a,b]$ und Zwischenpunkte $\xi_i \in \left[t_{i-1}, t_i\right]$, $i = 1, \cdots, n$ sei ferner

$$S\left(f, \left\{\xi_i\right\}_{i=1}^n, \Delta_n\right) := \sum_{i=1}^n f\left(\xi_i\right) \cdot \left(t_i - t_{i-1}\right).$$

Die Größe $\delta_n := \max\left\{t_i - t_{i-1} \mid 1 \le i \le n\right\}$ bezeichne die *Feinheit der Zerlegung* $\Delta_n \in \mathfrak{D}[a,b]$. Existiert dann

$$\lim_{\delta_n \to 0} S\left(f, \left\{\xi_i\right\}_{i=1}^n, \Delta_n\right) = \lim_{\delta_n \to 0} \sum_{i=1}^n f\left(\xi_i\right) \cdot \left(t_i - t_{i-1}\right)$$

unabhängig von der Wahl der Zwischenpunkte im starken (schwachen) Sinn in $\mathcal{X}$, so heißt *f stark (schwach) Riemann-Stieltjes-integrierbar*, und der obige Grenzwert wird mit

$$\mathbf{s}\int_a^b f(t)\,dt \quad \text{bzw.} \quad \mathbf{w}\int_a^b f(t)\,dt$$

bezeichnet.

Bemerkung: Sofern keine Missverständnisse entstehen (können), wird das starke (schwache) Riemann-Integral auch nur mit $\int_a^b f(t)\,dt$ bezeichnet.

Lemma 48 (Eigenschaften des abstrakten Integrals)**.** Es sei $E = [a,b] \subset \mathbb{R}$ ein kompaktes Intervall, $(\mathcal{X}, \|\bullet\|_{\mathcal{X}})$ ein Banach-Raum und $f, f_1, f_2 : E \to \mathcal{X}$ seien Abbildungen. Dann gilt:

a) Ist *f* stark (schwach) stetig, so ist *f* auch stark (schwach) Riemann-integrierbar.

b) Das starke (schwache) Riemann-Integral ist linear, d.h. sind f_1, f_2 stark (schwach) Riemann-integrierbar, so auch $\alpha_1 f_1 + \alpha_2 f_2$ für alle $\alpha_1, \alpha_2 \in \mathbb{R}$ mit

$$\int_a^b \alpha_1 f_1(t) + \alpha_2 f_2(t)\, dt = \int_a^b \alpha_1 f_1(t)\, dt + \int_a^b \alpha_2 f_2(t)\, dt.$$

c) Ist *f* stark (schwach) Riemann-integrierbar über $[a,b]$, so auch über jedem kompakten Teilintervall von $[a,b]$. Insbesondere gilt:

$$\int_a^b f(t)\, dt = \int_a^c f(t)\, dt + \int_c^b f(t)\, dt \text{ für alle } c \in (a,b).$$

d) Ist *f* stark Riemann-integrierbar über $[a,b]$, so auch $\|f(\bullet)\|_{\mathcal{X}}$ (im üblichen Sinne), und es gilt

$$\left\| \int_a^b f(t)\, dt \right\|_{\mathcal{X}} \leq \int_a^b \|f(t)\|_{\mathcal{X}}\, dt \leq (b-a) \max_{a \leq x \leq b} \|f(t)\|_{\mathcal{X}}.$$

e) Ist *f* stark (schwach) Riemann-integrierbar über $[a,b]$, so auch $f^*(f(\bullet))$ für alle $f^* \in \mathcal{X}^*$ (im üblichen Sinne), und es gilt

$$f^*\left(\int_a^b f(t)\, dt \right) = \int_a^b f^*(f(t))\, dt \text{ für alle } f^* \in \mathcal{X}^*.$$

f) Ist *f* stark (schwach) stetig, so ist die durch

$$F(t) := \int_a^t f(s)\, ds, \ a \leq t \leq b$$

definierte Abbildung $F : E \to \mathcal{X}$ stark (schwach) differenzierbar mit

$$\frac{d}{dt} F(t) = f(t) \text{ für alle } a \leq t \leq b.$$

g) Besitzt *f* eine stark stetige Ableitung auf $[a,b]$, so gilt

$$\int_a^t f'(s)\, ds = f(t) - f(a), \ a \leq t \leq b.$$

h) Ist $\{f_n(\bullet)\}_{n\in\mathbb{N}}$ eine Folge Riemann-integrierbarer Funktionen auf $[a,b]$ mit

$$\lim_{n\to\infty}\sup_{a\le t\le b}\|f_n(t)-f(t)\|=0,$$

so gilt

$$\int_a^b f(t)\,dt=\mathbf{s}\text{-}\lim_{n\to\infty}\int_a^b f_n(t)\,dt.$$

i) Ist $D\subseteq\mathcal{X}$ ein linearer Teilraum, $(\mathcal{Y},\|\bullet\|_{\mathcal{Y}})$ ein weiterer Banach-Raum, $T:D\to\mathcal{Y}$ ein *beschränkter* linearer Operator und f stark stetig mit Werten in D, so ist die Abbildung $(Tf)(t):=T\left(f(t)\right)$ für $t\in E$ stark stetig mit

$$\int_a^b Tf(t)\,dt=T\left(\int_a^b f(t)\,dt\right).$$

j) Ist $D\subseteq\mathcal{X}$ ein linearer Teilraum, $(\mathcal{Y},\|\bullet\|_{\mathcal{Y}})$ ein weiterer Banach-Raum, $T:D\to\mathcal{Y}$ ein *abgeschlossener* linearer Operator und f stark stetig mit Werten in D, so ist die Abbildung $(Tf)(t):=T\left(f(t)\right)$ für $t\in E$ Riemann-integrierbar mit

$$\int_a^b Tf(t)\,dt=T\left(\int_a^b f(t)\,dt\right).$$

Beweisskizze:

a) Für jedes beschränkte lineare Funktional $f^*\in\mathcal{X}^*$ ist die Abbildung $f^*\left(f(t)\right)$ definitionsgemäß stetig in $t\in E,$ also dort auch Riemann-integrierbar. Daraus folgt die Existenz des schwachen Riemann-Integrals für schwach stetiges f wegen

$$f^*\left(S\left(f,\{\xi_i\}_{i=1}^n,\Delta_n\right)\right):=f^*\left(\sum_{i=1}^n f(\xi_i)\cdot(t_i-t_{i-1})\right)=\sum_{i=1}^n f^*\left(f(\xi_i)\right)\cdot(t_i-t_{i-1})$$

für alle Zerlegungen $\Delta_n\in\mathfrak{D}[a,b]$ und Zwischenpunkte $\xi_i\in[t_{i-1},t_i],\ i=1,\cdots,n,$ mit dem Grenzwert $\int_a^b f^*\left(f(t)\right)dt$ auf der rechten Seite für $\delta_n\to0$ und der (hier nicht im Detail gezeigten) Tatsache, dass es ein eindeutig bestimmtes $I\in\mathcal{X}$ gibt mit

$$f^*(I)=\int_a^b f^*\left(f(t)\right)dt,$$

so dass diesem Fall $I=\mathbf{w}\int_a^b f(t)\,dt$ ist. Für die starke Riemann-Integrierbarkeit muss man noch zusätzlich zeigen, dass die Summen $S\left(f,\{\xi_i\}_{i=1}^n,\Delta_n\right)$ stark gegen I konvergieren.

b) Folgt unmittelbar aus der Definition des Riemann-Integrals, ebenso Teil c).

d) Folgt aus der Ungleichung

$$\left\|S\left(f,\{\xi_i\}_{i=1}^n,\Delta_n\right)\right\|_{\mathcal{X}} \leq \sum_{i=1}^n \left\|f(\xi_i)\right\|_{\mathcal{X}} \cdot (t_i - t_{i-1}) \leq \max_{a\leq x\leq b}\|f(t)\|_{\mathcal{X}} \sum_{i=1}^n (t_i - t_{i-1}) = (b-a)\max_{a\leq x\leq b}\|f(t)\|_{\mathcal{X}}$$

mit $\lim\limits_{\delta_n\to 0}\sum\limits_{i=1}^n \left\|f(\xi_i)\right\|_{\mathcal{X}} \cdot (t_i - t_{i-1}) = \int\limits_a^b \|f(t)\|_{\mathcal{X}}\, dt$ (man beachte, dass wegen der Kompaktheit von $[a,b]$ die stetige Funktion $\|f(\bullet)\|_{\mathcal{X}}$ dort ihre Extremwerte annimmt).

e) Folgt aus dem unter Teil a) angedeuteten Beweis.

f) Ist *f* schwach stetig, so gilt für $a < t < b,\ a-t < h < b-t$:

$$f^*\left(\frac{1}{h}\int\limits_t^{t+h} f(s)\,ds\right) = \frac{1}{h}\int\limits_t^{t+h} f^*\left(f(s)\right)ds \to f^*\left(f(t)\right) \text{ für } h\to 0,$$

d.h. es ist F auf (a,b) schwach differenzierbar (und damit schwach stetig) mit

$$\frac{\mathbf{w}d}{dt}F(t) = f(t) \text{ für alle } a<t<b.$$

Für die Randpunkte *a* und *b* argumentiert man analog.

Ist *f* stark stetig, so gilt für $a < t < b,\ 0 < h < b-t$:

$$\begin{aligned}\left\|\frac{1}{h}\int\limits_t^{t+h} f(s)\,ds - f(t)\right\|_{\mathcal{X}} &= \frac{1}{h}\left\|\int\limits_t^{t+h}\left(f(s)-f(t)\right)ds\right\|_{\mathcal{X}} \leq \frac{1}{h}\int\limits_t^{t+h}\|f(s)-f(t)\|_{\mathcal{X}}\,ds\\ &\leq \max_{t\leq s\leq t+h}\|f(s)-f(t)\|_{\mathcal{X}} \to 0 \text{ für } h\to 0,\end{aligned}$$

analog für negative *h*; d.h. es ist *F* auf (a,b) stark differenzierbar (und damit stark stetig) mit

$$\frac{\mathbf{s}d}{dt}F(t) = f(t) \text{ für alle } a<t<b.$$

Für die Randpunkte *a* und *b* argumentiert man analog.

g) Nach Teil e) gilt für alle beschränkten linearen Funktionale $f^* \in \mathcal{X}^*$

$$\begin{aligned} f^*\left(\int\limits_a^t f'(s)\,ds\right) &= \int\limits_a^t f^*\left(f'(s)\right)ds = \int\limits_a^t \frac{d}{ds}f^*\left(f(s)\right)ds = f^*\left(f(t)\right) - f^*\left(f(a)\right)\\ &= f^*\left(f(t)-f(a)\right),\ a\leq t\leq b,\end{aligned}$$

so dass die Behauptung wieder aus Lemma 26 folgt.

h) Folgt aus der Ungleichung

$$\left\|\int\limits_a^b f(t)\,dt - \int\limits_a^b f_n(t)\,dt\right\|_{\mathcal{X}} \leq \int\limits_a^b \|f(t)-f_n(t)\|_{\mathcal{X}}\,dt \leq (b-a)\sup_{a\leq t\leq b}\|f_n(t)-f(t)\| \to 0 \text{ für } n\to\infty.$$

i) Folgt aus der Beziehung

$$T\left(\int_a^b f(t)\,dt\right) = \lim_{\delta_n \to 0} T\left(S\left(f, \{\xi_i\}_{i=1}^n, \Delta_n\right)\right) = \lim_{\delta_n \to 0} S\left(Tf, \{\xi_i\}_{i=1}^n, \Delta_n\right) = \int_a^b Tf(t)\,dt,$$

da mit f auch Tf stark stetig ist wegen

$$\|Tf(t) - Tf(s)\|_{\mathcal{Y}} \le \|T\| \cdot \|f(t) - f(s)\|_{\mathcal{X}} \text{ für } s, t \in [a, b].$$

j) Es gilt

$$\lim_{\delta_n \to 0} S\left(f, \{\xi_i\}_{i=1}^n, \Delta_n\right) = \mathbf{s}\int_a^b f(t)\,dt \quad \text{und}$$

$$\lim_{\delta_n \to 0} T\left(S\left(f, \{\xi_i\}_{i=1}^n, \Delta_n\right)\right) = \lim_{\delta_n \to 0} S\left(Tf, \{\xi_i\}_{i=1}^n, \Delta_n\right) = \int_a^b Tf(t)\,dt,$$

also, da T abgeschlossen ist, $\mathbf{s}\int_a^b f(t)\,dt \in D$ und $\int_a^b Tf(t)\,dt = T\left(\int_a^b f(t)\,dt\right)$. ■

Das folgende Lemma ist eine Ergänzung der obigen Aussagen für Operator-wertige Abbildungen und kann analog bewiesen werden, weswegen wir hier auf einen detaillierten Beweis verzichten.

Lemma 49. Sind $\left(\mathcal{X}, \|\bullet\|_{\mathcal{X}}\right)$ und $\left(\mathcal{Y}, \|\bullet\|_{\mathcal{Y}}\right)$ Banach-Räume und ist $T : E \to \mathfrak{L}[\mathcal{X}, \mathcal{Y}]$ stark stetig, dann existiert das Riemann-Integral $\int_a^b T(t)\,dt$ im Sinne der gleichmäßigen Konvergenz, und es gilt

$$\left(\int_a^b T(t)\,dt\right) f = \int_a^b \left(T(t)\right) f\,dt.$$

Ferner existiert die Abbildung

$$F(t) := \int_a^t T(s)\,ds, \ \ a \le t \le b$$

punktweise im Sinne der gleichmäßigen Konvergenz, und ist gleichmäßig differenzierbar mit

$$\frac{d}{dt} F(t) = \frac{d}{dt}\int_a^t T(s)\,ds = T(t), \ \ a \le t \le b.$$

II. 1.2. Potenzen des infinitesimalen Erzeugers

Für das Folgende wollen wir wieder grundsätzlich annehmen, dass $\mathfrak{T} = \{T(t) \mid t \geq 0\} \subseteq \mathfrak{L}[\mathcal{X}, \mathcal{X}]$ eine Operator-Halbgruppe der Klasse (C_0) ist.

Definition 51 (infinitesimaler Erzeuger)**.** Der *inifinitesimale Erzeuger A* einer Operator-Halbgruppe der Klasse (C_0) ist definiert durch

$$Af := \mathbf{s}\text{-}\lim_{h \downarrow 0} \frac{1}{h}[T(h) - \mathbf{I}]f$$

für alle $f \in \mathcal{X}$, für die dieser Grenzwert existiert. Die Menge dieser $f \in \mathcal{X}$ wird mit $D(A)$ bezeichnet (*Definitionsbereich von A*).

Aus Vereinfachungsgründen wollen wir im Folgenden die Abkürzung

$$A_h := \frac{1}{h}[T(h) - \mathbf{I}] \text{ für alle } h > 0$$

verwenden.

Lemma 50 (Eigenschaften des infinitesimalen Erzeugers)**.** Der infinitesimale Erzeuger A und sein Definitionsbereich $D(A)$ besitzen folgende Eigenschaften:

a) $D(A)$ ist ein linearer Teilraum von $\mathcal{X}$, und A ist ein linearer Operator auf $D(A)$.
b) Es ist $T(t)f \in D(A)$ für alle $f \in D(A)$ und $t \geq 0$, mit

$$\frac{d}{dt} T(t)f = A\big(T(t)f\big) = T(t)(Af) \text{ für alle } f \in D(A) \text{ und } t \geq 0.$$

c) Es gilt

$$T(t)f - f = \int_0^t T(s)Af\, ds \quad \text{für alle } t \geq 0.$$

d) $D(A)$ ist dicht in $\mathcal{X}$ und A ist abgeschlossen.

Beweis:

a) Dies folgt unmittelbar aus der Definition von A.
b) Aus der Halbgruppeneigenschaft folgt sofort

$$T(t)\big(A_h f\big) = T(t)\left(\frac{1}{h}\{T(h)f - f\}\right) = \frac{1}{h}\big(T(t+h)f - T(t)f\big) = A_h\big(T(t)f\big)$$

für alle $f \in \mathcal{X}$ und $t \geq 0$, $h > 0$. Damit ergibt sich

$$T(t)\left(Af\right)=T(t)\left(\mathbf{s}\text{-}\lim_{h\downarrow 0} A_h f\right)=\mathbf{s}\text{-}\lim_{h\downarrow 0} T(t)\left(A_h f\right)=\mathbf{s}\text{-}\lim_{h\downarrow 0} A_h\left(T(t)f\right) \text{ für } f\in D(A) \text{ und } t\geq 0.$$

Für den linksseitigen Grenzwert argumentiert man so: Für $f\in\mathcal{X}$, $0<h\leq t$ ist

$$\frac{1}{h}\big(T(t-h)-T(t)\big)f=-T(t-h)\left\{\frac{1}{h}\big(T(h)-\mathbf{I}\big)f\right\},$$

also für $f\in D(A)$, $0<h\leq t$ nach mehrmaliger Anwendung der Dreiecks-Ungleichung

$$\begin{aligned}\left\|T(t-h)\left(A_h f\right)-T(t)(Af)\right\|_{\mathcal{X}}&\leq\left\|T(t)\right\|\cdot\left\|A_h f-Af\right\|_{\mathcal{X}}+\left\|T(t-h)-T(t)\right\|\cdot\left\|A_h f-Af\right\|_{\mathcal{X}}+\cdots\\&+\left\|\big(T(t-h)-T(t)\big)(Af)\right\|_{\mathcal{X}}\end{aligned}$$

mit $\lim_{h\downarrow 0}\left\|A_h f-Af\right\|_{\mathcal{X}}=0$ nach Definition von A, und $\lim_{h\downarrow 0}\left\|\big(T(t-h)-T(t)\big)(Af)\right\|_{\mathcal{X}}$ nach Lemma 45, Teil b), woraus schließlich auch

$$\lim_{h\downarrow 0}\left\|T(t-h)\left(A_h f\right)-T(t)(Af)\right\|_{\mathcal{X}}=0$$

folgt. Somit ist $T(t)f\in D(A)$ mit $\dfrac{d}{dt}T(t)f=A\big(T(t)f\big)=T(t)(Af)$ für alle $f\in D(A)$ und $t\geq 0$, nach dem obigen Beweisgang.

c) Für $f\in D(A)$, $t>0$ und $f^*\in\mathcal{X}^*$ ist nach Teil b) die Abbildung $t\mapsto f^*\left(T(t)f\right)$ differenzierbar mit

$$\begin{aligned}f^*\left(T(t)f-f\right)&=f^*\left(T(t)f-T(0)f\right)=f^*\left(T(t)f\right)-f^*\left(T(0)f\right)=\int_0^t\frac{d}{ds}f^*\left(T(s)f\right)ds\\&=\int_0^t f^*\left(\frac{d}{ds}T(s)f\right)ds=f^*\left(\int_0^t\frac{d}{ds}T(s)f\,ds\right)=f^*\left(\int_0^t T(s)(Af)\,ds\right),\end{aligned}$$

so dass mit Lemma 26 die Behauptung folgt.

d) Für $f\in\mathcal{X}$, $t,h>0$ gilt:

$$\begin{aligned}A_h\left(\int_0^t T(s)f\,ds\right)&=\int_0^t A_h\left(T(s)f\right)ds=\frac{1}{h}\int_0^t T(s+h)f\,ds-\frac{1}{h}\int_0^t T(s)f\,ds\\&=\frac{1}{h}\int_h^{t+h}T(s)f\,ds-\frac{1}{h}\int_0^t T(s)f\,ds=\frac{1}{h}\int_t^{t+h}T(s)f\,ds-\frac{1}{h}\int_0^h T(s)f\,ds\\&\to T(t)f-T(0)f\quad\text{für } h\downarrow 0,\end{aligned}$$

so dass also

$$A\left(\int_0^t T(s)f\,ds\right)=T(t)f-T(0)f$$

folgt mit $\int_0^t T(s)f\,ds\in D(A)$. Also ist auch $\frac{1}{t}\int_0^t T(s)f\,ds\in D(A)$ für alle $t>0$, mit

$$\mathbf{s}\text{-}\lim_{t\downarrow 0}\frac{1}{t}\int_0^t T(s)f\,ds = T(0)f = f.$$

Folglich ist für beliebiges $f\in\mathcal{X}$ auch $f\in\overline{D(A)}$, also $\mathcal{X}=\overline{D(A)}$ und somit $D(A)$ dicht in $\mathcal{X}$. Die Abgeschlossenheit von A ergibt sich nun noch so: Ist $\{f_n\}_{n\in\mathbb{N}}$ eine Folge in $D(A)$ mit starkem Grenzwert $f\in\mathcal{X}$ und starkem Grenzwert $g\in\mathcal{X}$ für die Folge $\{Af_n\}_{n\in\mathbb{N}}$, so gilt für jedes $t>0$:

$$T(t)f-f=\mathbf{s}\text{-}\lim_{n\to\infty}\{T(t)f_n-f_n\}=\mathbf{s}\text{-}\lim_{n\to\infty}\int_0^t T(s)\left(Af_n\right)ds=\int_0^t T(s)g\,ds$$

wegen

$$\left\|\int_0^t T(s)\left(Af_n-g\right)ds\right\|_{\mathcal{X}}\le t\cdot\max_{0\le s\le t}\|T(s)\|\cdot\|Af_n-g\|_{\mathcal{X}}\to 0\ \text{ für } n\to\infty.$$

Damit ist aber

$$Af=\mathbf{s}\text{-}\lim_{t\downarrow 0}\frac{1}{t}\{T(t)f-f\}=\mathbf{s}\text{-}\lim_{t\downarrow 0}\frac{1}{t}\int_0^t T(s)g\,ds=T(0)g=g,$$

also $f\in D(A)$ und $Af=g$, d.h. A ist abgeschlossen. ■

Die Nacheinander-Ausführung des infinitesimalen Erzeugers A führt damit in natürlicher Weise auf höhere Potenzen von A, wie die nachfolgende Definition zeigt.

Definition 52 (höhere Potenzen von A)**.** Für $r\in\mathbb{Z}^+=\{0,1,2,\cdots\}$ werden die (höheren) Potenzen des infinitesimalen Erzeugers A einer Operator-Halbgruppe der Klasse (C_0) folgendermaßen rekursiv definiert:

$$A^0:=\mathbf{I}\quad\text{und}\quad A^r f:=A\left(A^{r-1}f\right)\quad\text{für } r\in\mathbb{N},\ f\in D\left(A^r\right)\ \text{ mit}$$

$$D\left(A^r\right):=\left\{f\in\mathcal{X}\mid f\in D\left(A^{r-1}\right)\wedge A^{r-1}f\in D(A)\right\}.$$

Satz 42 (Eigenschaften der Potenzen des Erzeugers)**.** Die Potenzen A^r des infinitesimalen Erzeugers A und die zugehörigen Definitionsbereiche $D(A^r)$ besitzen folgende Eigenschaften:

a) $D(A^r)$ ist ein linearer Teilraum von $\mathcal{X}$, und A^r ist ein linearer Operator auf $D(A^r)$ für alle $r \in \mathbb{N}$.

b) Für jedes $r \in \mathbb{N}$ ist $T(t)f \in D(A^r)$ für alle $f \in D(A^r)$ und $t \geq 0$, mit

$$T(t)f - \sum_{k=0}^{r-1} \frac{t^k}{k!} A^k f = \frac{1}{(r-1)!} \int_0^t (t-s)^{r-1} T(s) A^r f \, ds \quad \text{für alle } f \in D(A^r) \text{ und } t \geq 0.$$

Ferner gilt für jedes $r \in \mathbb{N}$ und alle $h > 0$:

$$A_h^r f = \frac{1}{h^r} \int_0^h \cdots \int_0^h T(s_1 + \cdots + s_r) A^r f \, ds_1 \cdots ds_r \quad \text{für alle } f \in D(A^r).$$

c) $D(A^r)$ ist dicht in $\mathcal{X}$ für jedes $r \in \mathbb{N}$, ebenso $\bigcap_{r \in \mathbb{N}} D(A^r)$, und A^r ist abgeschlossen.

d) $D(A^r)$ ist ein Banach-Raum unter der Norm

$$\|f\|_{D(A^r)} := \sum_{k=0}^{r} \|A^k f\|_{\mathcal{X}} \quad \text{für alle } f \in D(A^r).$$

Beweis:

a) Dies folgt unmittelbar aus der Definition von A^r.

b) Die erste Aussage ergibt sich mittels vollständiger Induktion aus Lemma 50, Teil b). Für den weiteren Teil benutzen wir folgenden, aus der "klassischen" Analysis wohlbekannten Satz von Taylor: Ist g eine auf $[0, t]$ definierte und dort r-mal geeignet differenzierbare Funktion, so gilt

$$g(t) = \sum_{k=0}^{r-1} \frac{t^k}{k!} g^{(k)}(0) + \frac{1}{(r-1)!} \int_0^t (t-s)^{r-1} g^{(r)}(s) \, ds.$$

Man kann dies ebenfalls mit vollständiger Induktion zeigen, wobei man im Induktionsschritt mit partieller Integration argumentiert: es ist nämlich

$$\begin{aligned}\frac{1}{(r-1)!} \int_0^t (t-s)^{r-1} g^{(r)}(s) \, ds &= -\frac{(t-s)^r}{r!} g^{(r)}(s) \Big|_0^t + \frac{1}{r!} \int_0^t (t-s)^r g^{(r+1)}(s) \, ds \\ &= \frac{1}{r!} \int_0^t (t-s)^r g^{(r+1)}(s) \, ds + \frac{t^r}{r!} g^{(r)}(t).\end{aligned}$$

Eine Anwendung beschränkter linearer Funktionale zeigt dann, dass hier wie im Beweisteil c) von Lemma 50 gilt:

$$f^*\left(T(t)f\right)=\sum_{k=0}^{r-1}\frac{t^k}{k!}f^*\left(A^kf\right)+\frac{1}{(r-1)!}\int_0^t (t-s)^{r-1}f^*\left(T(s)\left(A^rf\right)\right)ds$$
$$=f^*\left(\sum_{k=0}^{r-1}\frac{t^k}{k!}A^kf+\frac{1}{(r-1)!}\int_0^t (t-s)^{r-1}T(s)\left(A^rf\right)ds\right)$$

und sich somit wieder mit Lemma 26 die Aussage ergibt. Für die zweite Beziehung ist nur zu zeigen:

$$\left(T(h)-\mathbf{I}\right)^r f=\int_0^h\cdots\int_0^h T\left(s_1+\cdots+s_r\right)A^rf\,ds_1\cdots ds_r \text{ für alle } f\in D\left(A^r\right).$$

Auch hier folgt das Ergebnis mit vollständiger Induktion: Für $r=1$ ist dies die Aussage aus Lemma 50, Teil c). Im Induktionsschritt rechnet man so:

$$\left(T(h)-\mathbf{I}\right)^{r+1}f=\left(T(h)-\mathbf{I}\right)\left(T(h)-\mathbf{I}\right)^r f=\left(T(h)-\mathbf{I}\right)\circ\left(\int_0^h\cdots\int_0^h T\left(s_1+\cdots+s_r\right)A^rf\,ds_1\cdots ds_r\right)$$
$$=\int_0^h T\left(s_{r+1}\right)A\left(\int_0^h\cdots\int_0^h T\left(s_1+\cdots+s_r\right)A^rf\,ds_1\cdots ds_r\right)ds_{r+1}$$
$$=\int_0^h\left(\int_0^h\cdots\int_0^h T\left(s_{r+1}\right)\circ T\left(s_1+\cdots+s_r\right)A^{r+1}f\,ds_1\cdots ds_r\right)ds_{r+1}$$
$$=\int_0^h\cdots\int_0^h T\left(s_1+\cdots+s_r+s_{r+1}\right)A^{r+1}f\,ds_1\cdots ds_r\,ds_{r+1}.$$

c) Der einzig nicht-triviale Teil ist hier die Aussage bzgl. des Durchschnitts $\bigcap_{r\in\mathbb{N}}D\left(A^r\right)$. Dazu betrachten wir die Menge $C_{00}^{\infty}\left(\mathbb{R}^+\right)$ derjenigen reellwertigen Funktionen auf $\mathbb{R}^+$, die stetige Ableitungen aller Ordnungen besitzen und außerhalb eines kompakten Intervalls $[a,b]\subset\mathbb{R}$ mit $a>0$ verschwinden. Mit φ ist auch $\varphi^{(r)}$ Element von $C_{00}^{\infty}\left(\mathbb{R}^+\right)$ für alle $r\in\mathbb{N}$, und die durch $s\mapsto\varphi(s)T(s)f$ für jedes $f\in\mathcal{X}$ definierte Abbildung ist stark stetig. Wir definieren

$$\mathcal{X}_{00}:=\left\{\int_0^\infty\varphi(s)T(s)f\,ds\;\middle|\;\varphi\in C_{00}^{\infty}\left(\mathbb{R}^+\right),\ f\in\mathcal{X}\right\}.$$

Man beachte, dass hierbei $\int_0^\infty\varphi(s)T(s)f\,ds$ wohldefiniert ist, weil das Integral in Abhängigkeit von der Wahl für φ auf einen kompakten Integrationsbereich beschränkt werden kann. Offensichtlich ist $\mathcal{X}_{00}$ ein linearer Teilraum von $\mathcal{X}$. Wir werden jetzt zeigen, dass $\mathcal{X}_{00}\subseteq D\left(A^r\right)$ ist für jedes $r\in\mathbb{N}$, und dass $\mathcal{X}_{00}$ dicht in $\mathcal{X}$ ist, womit die Aussage gezeigt ist. Dazu weisen wir zunächst die Gültigkeit von

$$A\left(\int_0^\infty\varphi(s)T(s)f\,ds\right)=-\int_0^\infty\varphi'(s)T(s)f\,ds \quad\text{für }\varphi\in C_{00}^{\infty}\left(\mathbb{R}^+\right)\text{ und } f\in\mathcal{X} \qquad (*)$$

nach. Für $h>0$ ist nämlich, mit der Abkürzung $g=\int_0^\infty \varphi(s)T(s)f\,ds$ und der kanonischen Fortsetzung von φ auf ganz $\mathbb{R}$ mit $\varphi(s)=0,\ s<0$:

$$A_h g=\frac{1}{h}\int_0^\infty \varphi(s)\big(T(s+h)-T(s)\big)f\,ds=\frac{1}{h}\int_0^\infty \{\varphi(s-h)-\varphi(s)\}T(s)f\,ds$$

mit dem im Träger von φ gleichmäßigen Lines $-\varphi'(s)=\lim_{h\downarrow 0}\frac{\varphi(s-h)-\varphi(s)}{h}$, woraus $(*)$ folgt. Eine mehrfache Anwendung von $(*)$ zeigt schließlich

$$A^r\left(\int_0^\infty \varphi(s)T(s)f\,ds\right)=(-1)^r\int_0^\infty \varphi^{(r)}(s)T(s)f\,ds \quad \text{für } \varphi\in C_{00}^\infty\left(\mathbb{R}^+\right) \text{ und } f\in\mathcal{X}, \text{ für alle } r\in\mathbb{N},$$

woraus $\mathcal{X}_{00}\subseteq D\left(A^r\right)$ für alle $r\in\mathbb{N}$ und somit auch $\mathcal{X}_{00}\subseteq\bigcap_{r\in\mathbb{N}}D\left(A^r\right)$ folgt. Angenommen, $\mathcal{X}_{00}$ sei nicht dicht in $\mathcal{X}$. Dann existiert ein $f_0\in\mathcal{X}$ mit $\operatorname{dist}\left(\{f_0\},\mathcal{X}_{00}\right)=\inf_{f\in\mathcal{X}_{00}}\left\{\|f_0-f\|_{\mathcal{X}}\right\}>0$. Nach Satz 28 existiert dann aber ein beschränktes lineares Funktional $F^*\in\mathcal{X}^*$ mit $F^*=\mathbf{0}$ auf $\mathcal{X}_{00}$ und $F^*\left(f_0\right)=1$. Dies impliziert

$$F^*\left(\int_0^\infty \varphi(s)T(s)f\,ds\right)=\int_0^\infty \varphi(s)F^*\left(T(s)f\right)ds=0 \text{ für alle } \varphi\in C_{00}^\infty\left(\mathbb{R}^+\right) \text{ und } f\in\mathcal{X}.$$

Nun ist aber die Abbildung $s\mapsto F^*\left(T(s)f_0\right)$ stetig auf $\mathbb{R}^+$ mit $\lim_{s\downarrow 0}F^*\left(T(s)f_0\right)=F^*\left(f_0\right)=1$. Also existiert ein $\delta>0$ mit der Eigenschaft $F^*\left(T(s)f_0\right)\geq\frac{1}{2}$ für $0\leq s\leq\delta$. Ferner existiert ein nicht-negatives $\varphi_\delta\in C_{00}^\infty\left(\mathbb{R}^+\right)$ mit der Eigenschaft $\varphi_\delta(s)\geq 1$ für $\frac{\delta}{4}\leq s\leq\frac{3\delta}{4}$ und $\varphi_\delta(s)=0$ für $s\leq\frac{\delta}{8}$ oder $s\geq\frac{7\delta}{8}$. Damit ist aber

$$F^*\left(\int_0^\infty \varphi_\delta(s)T(s)f_0\,ds\right)=\int_{\delta/8}^{7\delta/8}\varphi_\delta(s)F^*\left(T(s)f_0\right)ds\geq\frac{1}{2}\int_{\delta/4}^{3\delta/4}\varphi_\delta(s)\,ds=\frac{\delta}{4}>0: \text{ Widerspruch!}$$

Also ist $\mathcal{X}_{00}$ und damit auch $\bigcap_{r\in\mathbb{N}}D\left(A^r\right)$ tatsächlich dicht in $\mathcal{X}$.

d) Wie im Beweis von Lemma 50, Teil d) ergibt sich auch hier, dass A^r abgeschlossen ist für jedes $r\in\mathbb{N}$. Sei nun zunächst $r=1$. Nach Lemma 41 ist dann mit A auch der Graph $G_A=\{(f,Af)\mid f\in D(A)\}$ abgeschlossen in $\mathcal{X}\times\mathcal{X}$, mit der Norm

$$\|(f,Af)\|_{G_A}=\|f\|_{\mathcal{X}}+\|Af\|_{\mathcal{X}} \text{ für } f\in D(A).$$

Damit folgt, dass $D(A)$ mit dieser Norm ein Banach-Raum ist. Die allgemeinere Aussage ergibt sich induktiv durch Betrachtung des Teil-Graphen $G_A=\left\{\left(A^{r-1}f,A^rf\right)\mid f\in D\left(A^{r-1}\right)\right\}$ mit der Norm

$$\left\|\left(A^{r-1}f, A^r f\right)\right\|_{G_A} = \left\|A^{r-1}f\right\|_{\mathcal{X}} + \left\|A^r f\right\|_{\mathcal{X}} = \left\|A^{r-2}f\right\|_{\mathcal{X}} + \left\|A^{r-1}f\right\|_{\mathcal{X}} + \left\|A^r f\right\|_{\mathcal{X}} = \cdots$$

$$= \left\|f\right\|_{\mathcal{X}} + \cdots + \left\|A^r f\right\|_{\mathcal{X}} = \sum_{k=0}^{r} \left\|A^k f\right\|_{\mathcal{X}} \text{ für } f \in D\left(A^r\right).$$

Damit ist der Satz vollständig bewiesen. ■

Zum Abschluss dieses Abschnitts wollen wir noch kurz die infinitesimalen Erzeuger und ihre Potenzen der in den Beispielen 22 bis 24 vorgestellten Halbgruppen untersuchen.

Beispiel 27 (Fortsetzung von Beispiel 22)**.** Hier gilt offensichtlich:

$$\lim_{h\downarrow 0} A_h f(x) = \lim_{h\downarrow 0} \frac{f(x+h)-f(x)}{h} = f'(x) = Af(x) \text{ für } x \in \mathbb{R},$$

mit

$$D(A) = \left\{ f \in UCB(\mathbb{R}) \,|\, f' \in UCB(\mathbb{R}) \right\}$$

(man beachte, dass die Konvergenz hier nicht nur punktweise, sondern gleichmäßig sein muss). Für die höheren Potenzen ergibt sich entsprechend

$$A^r f = f^{(r)} \text{ mit } D\left(A^r\right) = \left\{ f \in UCB(\mathbb{R}) \,|\, f^{(r)} \in UCB(\mathbb{R}) \right\}, \; r \in \mathbb{N}.$$

Beispiel 28 (Fortsetzung von Beispiel 23)**.** Hier gilt offensichtlich:

$$\mathbf{s}\text{-}\lim_{h\downarrow 0} A_h f = \lim_{h\downarrow 0} \frac{e^{\lambda h}-1}{h} f = \lambda f = Af,$$

mit $D(A) = \mathcal{X}$. Für die höheren Potenzen ergibt sich analog

$$A^r f = \lambda^r f \quad \text{mit } D\left(A^r\right) = \mathcal{X}, \; r \in \mathbb{N}.$$

Beispiel 29 (Fortsetzung von Beispiel 24)**.** Hier gilt, mit $B_t := \frac{1}{2}\begin{bmatrix} e^{\mu t}+e^{\lambda t} & e^{\mu t}-e^{\lambda t} \\ e^{\mu t}-e^{\lambda t} & e^{\mu t}+e^{\lambda t} \end{bmatrix}, \; t \geq 0:$

$$\lim_{h\downarrow 0} \frac{1}{t}\left(B_t - \mathbf{I}\right) = \frac{1}{2} \lim_{h\downarrow 0} \begin{bmatrix} \frac{e^{\mu t}-1}{t} + \frac{e^{\lambda t}-1}{t} & \frac{e^{\mu t}-1}{t} - \frac{e^{\lambda t}-1}{t} \\ \frac{e^{\mu t}-1}{t} - \frac{e^{\lambda t}-1}{t} & \frac{e^{\mu t}-1}{t} + \frac{e^{\lambda t}-1}{t} \end{bmatrix} = \frac{1}{2}\begin{bmatrix} \mu+\lambda & \mu-\lambda \\ \mu-\lambda & \mu+\lambda \end{bmatrix} =: B.$$

Damit erhält man für den infinitesimalen Erzeuger und seine Potenzen:

$$A^r f = B^r \cdot f, \text{ mit } D\left(A^r\right) = \mathcal{X}, \; r \in \mathbb{N}.$$

Die Potenzen der Matrix B können in diesem Fall explizit angegeben werden; es gilt nämlich:

$$B^r = \frac{1}{2}\begin{bmatrix} \mu^r + \lambda^r & \mu^r - \lambda^r \\ \mu^r - \lambda^r & \mu^r + \lambda^r \end{bmatrix} \text{ für } r \in \mathbb{N}.$$

Dies kann z.B. durch vollständige Induktion leicht verifiziert werden.

II.2. Darstellungssätze für Operatorhalbgruppen

In diesem Abschnitt werden wir unterschiedliche Repräsentationen von Operator-Halbgruppen studieren, die sich wesentlich danach unterscheiden, ob der infinitesimale Erzeuger A beschränkt (d.h. stetig) ist oder lediglich abgeschlossen. Erstaunlicherweise ergibt sich im letzteren Fall ein interessanter Zusammenhang zum Gesetz der Großen Zahlen aus der Stochastik.

II. 2.1. Analytische Darstellungssätze

Wir beginnen mit dem Fall, dass der infinitesimale Erzeuger A beschränkt ist. Dann besitzt die Operatorhalbgruppe eine eindeutige Darstellung als Exponentialfunktion.

Satz 43 (Darstellungssatz für beschränkte Erzeuger)**.** Es sei $A \in \mathfrak{L}[\mathcal{X}, \mathcal{X}]$ ein beschränkter linearer Operator. Dann wird durch

$$T(t) := e^{tA} := \sum_{k=0}^{\infty} \frac{t^k}{k!} A^k, \ t \geq 0$$

eine Halbgruppe $\mathfrak{T}$ der Klasse (C_0) definiert, deren infinitesimaler Erzeuger gerade A ist, mit $D(A) = \mathcal{X}$. Ferner gilt dann

$$\|T(t)\| \leq e^{t\|A\|}, \ t \geq 0.$$

Ist umgekehrt $\mathfrak{T} = \{T(t) \,|\, t \geq 0\} \subseteq \mathfrak{L}[\mathcal{X}, \mathcal{X}]$ eine Halbgruppe der Klasse (C_0) mit beschränktem infinitesimalem Erzeuger A, so ist notwendig $D(A) = \mathcal{X}$, und die Halbgruppe besitzt die oben gegebene eindeutige Darstellung.

Beweis: Wir zeigen für den ersten Teil zunächst, dass die Reihe $\sum_{k=0}^{\infty} \frac{t^k}{k!} A^k$ für jedes $t \geq 0$ stark konvergiert. Für festes $t \geq 0$, $f \in \mathcal{X}$ und beliebige $n, m \in \mathbb{N}$ mit $n < m$ gilt nämlich:

$$\left\| \sum_{k=0}^{m} \frac{t^k}{k!} A^k f - \sum_{k=0}^{n} \frac{t^k}{k!} A^k f \right\|_{\mathcal{X}} = \left\| \sum_{k=n+1}^{m} \frac{t^k}{k!} A^k f \right\|_{\mathcal{X}} \leq \|f\|_{\mathcal{X}} \sum_{k=n+1}^{m} \frac{t^k}{k!} \|A\|^k \to 0 \ \text{ für } n, m \to \infty,$$

d.h. die Partialsummenfolge $\left\{\sum_{k=0}^{n}\frac{t^k}{k!}A^k f\right\}_{n\in\mathbb{N}}$ ist starke Cauchy-Folge in $\mathcal{X}$ und damit die Reihe $\sum_{k=0}^{\infty}\frac{t^k}{k!}A^k$ für jedes $t\ge 0$ nach Lemma 23 stark konvergent. Ähnlich folgt

$$\|T(t)\|\le\sum_{k=0}^{\infty}\frac{t^k}{k!}\|A\|^k=e^{t\|A\|},\ t\ge 0.$$

Wir weisen jetzt zunächst die eigentliche Halbgruppen-Eigenschaft nach. Dazu seien die beschränkten linearen Operatoren $S_n(t)$ definiert durch

$$S_n(t):=\sum_{k=0}^{n}\frac{t^k}{k!}A^k,\ t\ge 0,\ n\in\mathbb{N}.$$

Ähnlich wie in der "klassischen" Analysis beim Produkt von Polynomen ergibt sich jetzt für die Hintereinanderausführung zweier solcher Operatoren für $n,m\in\mathbb{N},\ n\le m$:

$$S_n(t)\circ S_m(s)=\sum_{k=0}^{n+m}c_k A^k,\ t,s\ge 0,\ \text{mit}$$
$$c_k=\sum_{i=0\vee(k-m)}^{n\wedge k}\frac{t^i}{i!}\frac{s^{k-i}}{(k-i)!},\ k=0,\cdots,n+m$$

(mit der üblichen Konvention $a\vee b=\max\{a,b\},\ a\wedge b=\min\{a,b\}$ für $a,b\in\mathbb{R}$). Speziell ergibt sich noch

$$c_k=\sum_{i=0}^{k}\frac{t^i}{i!}\frac{s^{k-i}}{(k-i)!}=\frac{1}{k!}\sum_{i=0}^{k}k!\frac{t^i}{i!}\frac{s^{k-i}}{(k-i)!}=\frac{1}{k!}\sum_{i=0}^{k}\binom{k}{i}t^i s^{k-i}=\frac{(s+t)^k}{k!},\ k=0,\cdots,n.$$

Hieraus erhält man wie oben:

$$T(t)\circ T(s)=\mathbf{s}\text{-}\lim_{n,m\to\infty}S_n(t)\circ S_m(s)=\sum_{k=0}^{\infty}c_k A^k=\sum_{k=0}^{\infty}\frac{(s+t)^k}{k!}A^k=T(s+t),\ t,s\ge 0,$$

mit $T(0)=A^0=\mathbf{I}$. Die Halbgruppe ist ferner von der Klasse (C_0) wegen

$$\|T(t)f-f\|_{\mathcal{X}}=\left\|\sum_{k=1}^{\infty}\frac{t^k}{k!}A^k f\right\|_{\mathcal{X}}\le\|f\|_{\mathcal{X}}\sum_{k=1}^{\infty}\frac{t^k}{k!}\|A\|^k=\|f\|_{\mathcal{X}}\left(e^{t\|A\|}-1\right)\to 0\ \text{ mit } t\downarrow 0 \text{ für alle } f\in\mathcal{X}.$$

Ferner gilt für alle $f\in\mathcal{X}$:

$$\mathbf{s}\text{-}\lim_{t\downarrow 0}\left\{\frac{1}{t}\left(e^{tA}-\mathbf{I}\right)-A\right\}f=\mathbf{s}\text{-}\lim_{t\downarrow 0}t\cdot\left(\sum_{k=2}^{\infty}\frac{t^{k-2}}{k!}A^k f\right)=0$$

wegen

$$\left\|\sum_{k=2}^{\infty}\frac{t^{k-2}}{k!}A^k f\right\|_{\mathcal{X}}\le\|f\|_{\mathcal{X}}\sum_{k=1}^{\infty}\frac{t^{k-2}}{k!}\|A\|^k\le\|f\|_{\mathcal{X}}\|A\|^2 e^{t\|A\|}<\infty.$$

Damit ist zugleich A der infinitesimale Erzeuger der Halbgruppe, mit $D(A)=\mathcal{X}$.

Für den zweiten Teil sei A der beschränkte infinitesimale Erzeuger. Dann wird durch

$$S(t) := T(t) \circ e^{-tA},\ t \geq 0$$

eine weitere Halbgruppe der Klasse (C_0) erzeugt (man beachte hierbei, dass die Operatoren $T(t)$ und $e^{-tA} = \sum_{k=0}^{\infty} (-1)^k \frac{t^k}{k!} A^k$ auf Grund von Lemma 50, Teil b) kommutieren!). Hier gilt nun:

$$(T(t)-\mathbf{I})(e^{-tA}-\mathbf{I}) = (T(t)e^{-tA}-\mathbf{I}) - (T(t)-\mathbf{I}) - (e^{-tA}-\mathbf{I}) \text{ für alle } t \geq 0, \text{ folglich}$$

$$\frac{1}{t}(T(t)e^{-tA}-\mathbf{I}) = \left\{(T(t)-\mathbf{I})\frac{1}{t}(e^{-tA}-\mathbf{I})\right\} + \frac{1}{t}(T(t)-\mathbf{I}) + \frac{1}{t}(e^{-tA}-\mathbf{I}) \text{ für alle } t > 0, \text{ also}$$

$$\mathbf{s}\text{-}\lim_{t\downarrow 0}\frac{1}{t}(T(t)e^{-tA}-\mathbf{I})f = \mathbf{s}\text{-}\lim_{t\downarrow 0}\frac{1}{t}(T(t)-\mathbf{I})f + \mathbf{s}\text{-}\lim_{t\downarrow 0}\frac{1}{t}(e^{-tA}-\mathbf{I})f = Af - Af = \mathbf{0} \text{ für alle } f \in D(A).$$

Damit ist der Erzeuger der Halbgruppe $\mathfrak{S} := \{S(t) \mid t \geq 0\}$ aber der (beschränkte) Null-Operator auf dem Banach-Raum $D(A)$, so dass nach dem ersten Teil folgt:

$$S(t)f = T(t)e^{-tA}f = \mathbf{I}f = f \quad \text{und somit } T(t)f = e^{tA}f \text{ für alle } f \in D(A).$$

Nach Lemma 50, Teil d) ist aber $D(A)$ dicht in $\mathcal{X}$. Zu jedem $f \in \mathcal{X}$ existiert also eine Folge $\{f_n\}_{n\in\mathbb{N}}$ in $D(A)$ mit starkem Grenzwert f. Somit folgt

$$T(t)f = \mathbf{s}\text{-}\lim_{n\to\infty} T(t)f_n = \mathbf{s}\text{-}\lim_{n\to\infty} e^{tA}f_n = e^{tA}f \text{ für alle } t \geq 0, \text{ also } T(t) = e^{tA} \text{ für alle } t \geq 0,$$

wie behauptet. Damit ist der Satz vollständig bewiesen. ■

Bemerkung: Der vorangegangene Beweis zeigt, dass bei beschränktem Erzeuger die die Halbgruppe $\mathfrak{T} = \{e^{tA} \mid t \geq 0\}$ darstellende Potenzreihe sogar gleichmäßig konvergiert, und die Halbgruppe auch gleichmäßig differenzierbar ist, d.h. es gilt

$$\frac{1}{t}(T(t)-\mathbf{I}) = \frac{1}{t}(e^{A}-\mathbf{I}) \to A \text{ gleichmäßig für } t \downarrow 0 \text{ (in der Operator-Norm),}$$

vgl. Definition 31.

"Der" infinitesimale Erzeuger einer beliebigen Halbgruppe der Klasse (C_0) ist stets eindeutig bestimmt, nicht nur in dem speziellen Fall des Satzes 43 (vgl. Remark 2.14 in GOLDSTEIN (1985) oder WERNER (2004), Korollar VII.4.8).

Ferner gilt allgemeiner: sind $\mathfrak{T}_i = \{T_i(t) \mid t \geq 0\}$ für $i = 1, \cdots, n$ kommutierende Operator-Halbgruppen der Klasse (C_0) auf demselben Banach-Raum $\mathcal{X}$ mit den infinitesimalen Erzeugern $A_i,\ i = 1, \cdots n,$ so ist auch $\mathfrak{T} = \left\{\prod_{i=1}^{n} T_i(t) \mid t \geq 0\right\}$ eine Operator-Halbgruppe der Klasse (C_0), wobei hier $\prod$ die Hintereinanderausführung der Operatoren bezeichne. Diese Halbgruppe besitzt den infinitesimalen Erzeuger $A=\sum_{i=1}^{n}A_i$ mit $D(A)=\bigcap_{i=1}^{n}D\left(A_i\right)$. Für den Fall der Beschränktheit der Erzeuger entspricht dies der klassischen Rechenregel für die Exponentialfunktion, nämlich

$$\prod_{i=1}^{n}e^{tA_i}=\exp\left\{\sum_{i=1}^{n}tA_i\right\},\ t\geq 0.$$

Beispiel 30 (Fortsetzung von Beispiel 29)**.** Der hier gegebene infinitesimale Erzeuger ist offensichtlich darstellbar als Summe zweier beschränkter linearer Operatoren, in Matrix-Form gegeben durch

$$B_1:=\frac{\mu}{2}\begin{bmatrix}1&1\\1&1\end{bmatrix}\text{ und }B_2:=\frac{\lambda}{2}\begin{bmatrix}1&-1\\-1&1\end{bmatrix},\text{ mit }B_1{}^k:=\frac{\mu^k}{2}\begin{bmatrix}1&1\\1&1\end{bmatrix}\text{ und }B_2{}^k:=\frac{\lambda^k}{2}\begin{bmatrix}1&-1\\-1&1\end{bmatrix}\text{ für }k\in\mathbb{N}.$$

Die hierdurch einzeln erzeugten Halbgruppen sind (in Matrix-Form):

$$e^{tB_1}=\sum_{k=0}^{\infty}\frac{t^k}{k!}B_1{}^k=\mathbf{I}+\frac{1}{2}\left(\sum_{k=1}^{\infty}\frac{(\mu t)^k}{k!}\right)\begin{bmatrix}1&1\\1&1\end{bmatrix}=\mathbf{I}+\frac{e^{\mu t}-1}{2}\begin{bmatrix}1&1\\1&1\end{bmatrix}=\frac{1}{2}\begin{bmatrix}e^{\mu t}+1&e^{\mu t}-1\\e^{\mu t}-1&e^{\mu t}+1\end{bmatrix}$$

$$e^{tB_2}=\sum_{k=0}^{\infty}\frac{t^k}{k!}B_2{}^k=\mathbf{I}+\frac{1}{2}\left(\sum_{k=1}^{\infty}\frac{(\lambda t)^k}{k!}\right)\begin{bmatrix}1&-1\\-1&1\end{bmatrix}=\mathbf{I}+\frac{e^{\lambda t}-1}{2}\begin{bmatrix}1&-1\\-1&1\end{bmatrix}=\frac{1}{2}\begin{bmatrix}e^{\lambda t}+1&-e^{\lambda t}+1\\-e^{\lambda t}+1&e^{\lambda t}+1\end{bmatrix}$$

für $t\geq 0,$ mit der Produkt-Darstellung

$$e^{t(B_1+B_2)}=e^{tB_1}e^{tB_2}=\frac{1}{4}\begin{bmatrix}e^{\mu t}+1&e^{\mu t}-1\\e^{\mu t}-1&e^{\mu t}+1\end{bmatrix}\cdot\begin{bmatrix}e^{\lambda t}+1&-e^{\lambda t}+1\\-e^{\lambda t}+1&e^{\lambda t}+1\end{bmatrix}=\frac{1}{2}\begin{bmatrix}e^{\mu t}+e^{\lambda t}&e^{\mu t}-e^{\lambda t}\\e^{\mu t}-e^{\lambda t}&e^{\mu t}+e^{\lambda t}\end{bmatrix}\text{ für }t\geq 0,$$

wie erwartet.

Beispiel 31 (Ergänzung zu Beispiel 28)**.** Ist $\mathfrak{T}=\{T(t)\,|\,t\geq 0\}$ eine Halbgruppe der Klasse $\left(C_0\right)$ mit infinitesimalem Erzeuger A, so ist offensichtlich auch die Halbgruppe $\mathfrak{T}_\lambda=\left\{e^{\lambda t}T(t)\,|\,t\geq 0\right\}$ von der Klasse $\left(C_0\right)$ für beliebiges $\lambda\in\mathbb{R}$, mit dem infinitesimalem Erzeuger $A_\lambda=\lambda\mathbf{I}+A$ wegen

$$\frac{1}{t}\left(e^{\lambda t}T(t)-\mathbf{I}\right)f=\left\{\left(T(t)-\mathbf{I}\right)\frac{1}{t}\left(e^{\lambda t}\mathbf{I}-\mathbf{I}\right)\right\}f+\frac{1}{t}\left(T(t)-\mathbf{I}\right)f+\frac{1}{t}\left(e^{\lambda t}\mathbf{I}-\mathbf{I}\right)f\to Af+\lambda f$$

für $t\downarrow 0$ und $f\in D(A)$, mit $D\left(A_\lambda\right)=D(A)$. Wählt man speziell $\lambda\leq-\omega$ mit der Konstanten $M(\omega)$ aus Lemma 45, Teil d), so erhält man noch die Abschätzung

$$\left\|e^{\lambda t}T(t)\right\|\leq M(\omega)\ \text{ für alle }t\geq 0.$$

Halbgruppen mit dieser Eigenschaft heißen *gleichmäßig beschränkt* (engl.: equi-bounded); im Fall $M(\omega)=1$ heißen sie auch *Kontraktions-Halbgruppe*.

Wir kommen nun zu einem der ersten wichtigen nicht-trivialen Darstellungssätze für Operator-Halbgruppen, der auch als "Hille’s First Exponential Formula" bekannt ist. Dazu benötigen wir den Begriff des *Stetigkeitsmoduls*.

Definition 53 (Stetigkeitsmodul)**.** Ist $\mathfrak{T} = \{T(t) \mid t \geq 0\}$ eine Halbgruppe der Klasse (C_0) und $L > 0,$ so heißt für $\delta > 0$ und $f \in \mathcal{X}$ die Größe

$$\omega_L(f;\delta) := \sup\left\{\left\|\big(T(t) - T(s)\big) f\right\| \,\middle|\, |t-s| < \delta,\ 0 \leq s,t \leq L\right\}$$

der rektifizierte Stetigkeitsmodul (der Halbgruppe) für $f \in \mathcal{X}$ im Intervall $[0, L]$.

Satz 44 (Hille's erste Exponentialformel, E. Hille (1942))**.** Es sei $\mathfrak{T} = \{T(t) \mid t \geq 0\}$ eine Halbgruppe der Klasse (C_0). Dann gilt:

$$T(t)f = \mathbf{s}\text{-}\lim_{h \downarrow 0} \exp(tA_h) f \quad \text{für alle } f \in \mathcal{X},$$

wobei der Grenzwert gleichmäßig bzgl. t in jedem kompakten Teilintervall von $\mathbb{R}^+$ existiert. Ist ferner $L > 1$ vorgegeben, so gibt es eine – nur von L und dem charakteristischen Exponenten ω_0 der Halbgruppe[27] abhängige – universelle Konstante K derart, dass genauer gilt:

$$\left\|T(t)f - \exp(tA_h) f\right\|_{\mathcal{X}} \leq \omega_L\left(f; h^{1/3}\right) + h^{1/3} K \left\|f\right\|_{\mathcal{X}} \text{ für alle } t \in [0, L-1] \text{ und } f \in \mathcal{X},$$

für genügend kleine $h > 0.$

Beweis: Wir folgen im Wesentlichen dem Beweisgang des fast gleich lautenden Theorems 1.1.2 in BUTZER / BEHRENS (1967), werden aber schon hier das später noch ausführlicher diskutierte probabilistische Beweisargument einsetzen.

Dazu betrachten wir eine mit dem Parameter $\lambda \geq 0$ Poisson-verteilte Zufallsvariable X_λ (wobei für den Entartungsfall $\lambda = 0$ definitionsgemäß $X_0 \equiv 0$ gelte). Dann gilt für Erwartungswert und Varianz von X_λ:

$$E(X_\lambda) = \lambda,\ Var(X_\lambda) = \lambda \quad \text{für } \lambda \geq 0.$$

Ferner gilt jedes $t > 0, h > 0$ und $f \in \mathcal{X}$ mit $\lambda := \dfrac{t}{h}$:

$$T(t)f - \exp(tA_h) f = T(t)f - e^{-\lambda} \exp(\lambda T(h)) f = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!}\left[T(t) - T(h)^k\right] f$$

$$= e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!}\left[T(t) - T(kh)\right] f.$$

Es werde nun ein $\omega > \max\{\omega_0, 0\}$ beliebig gewählt. Nach Lemma 45, Teil d) existiert dann ein $M = M(\omega) \geq 1$ mit

$$\left\|T(s)\right\| \leq M\, e^{\omega s} \quad \text{für alle } s \geq 0,$$

so dass gilt:

$$\left\|T(kh)\right\| \leq M\, e^{\omega kh} = M \tau^k \quad \text{für alle } k \in \mathbb{Z}^+, \text{ mit } \tau = e^{\omega h}.$$

[27] Vgl. Lemma 45.

Sei nun $L > 1$ zunächst fest gewählt sowie $h > 0$ derart, dass $h \le \min\left\{\frac{1}{\omega}\ln\left(1+\sqrt{h}\right),1\right\}$ ist. Dies ist für genügend kleine Werte von h stets möglich, da asymptotisch $\frac{1}{\omega}\ln\left(1+\sqrt{h}\right) \approx \frac{\sqrt{h}}{\omega}$ für $h \downarrow 0$ gilt. Dann ist insbesondere $\tau - 1 = e^{\omega h} - 1 \le \min\left\{1,\sqrt{h}\right\}$. Setzt man $\delta := h^{1/3}$, erhält man noch die Abschätzung $\lambda(\tau-1) = t\frac{e^{\omega h}-1}{h} \le t \cdot e^{\omega}$. Ferner sei die Indexmenge J gegeben durch $J := \left\{k \in \mathbb{Z}^+ \,\middle|\, |k-\lambda| \le h^{-2/3}\right\}$. Dann gilt, unter Anwendung der Tschebyscheff-Ungleichung bzw. des Verschiebungssatzes von Steiner, für jedes $f \in \mathcal{X}$ und $t \in [0, L-1]$:

$$
\begin{aligned}
\left\|T(t)f - \exp\left(tA_h\right)f\right\|_{\mathcal{X}} &= \left\|e^{-\lambda}\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}\left[T(t)-T(kh)\right]f\right\|_{\mathcal{X}} \le e^{-\lambda}\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}\left\|\left[T(t)-T(kh)\right]f\right\|_{\mathcal{X}} \\
&= e^{-\lambda}\sum_{k\in J}\frac{\lambda^k}{k!}\left\|\left[T(t)-T(kh)\right]f\right\|_{\mathcal{X}} + \|f\|_{\mathcal{X}}\, e^{-\lambda}\sum_{k\notin J}\frac{\lambda^k}{k!}\left\|T(t)-T(kh)\right\| \\
&\le \omega_L(f;\delta)\, e^{-\lambda}\sum_{k\in J}\frac{\lambda^k}{k!} + \|T(t)\|\cdot\|f\|_{\mathcal{X}}\, e^{-\lambda}\sum_{k\notin J}\frac{\lambda^k}{k!} + \|f\|_{\mathcal{X}} e^{-\lambda}\sum_{k\notin J}\frac{\lambda^k}{k!}\|T(kh)\| \\
&\le \omega_L(f;\delta) + \|T(t)\|\cdot\|f\|_{\mathcal{X}}\cdot P\left(\left|X_\lambda - \lambda\right| > h^{-2/3}\right) + M\|f\|_{\mathcal{X}}\, e^{-\lambda}\sum_{k\notin J}\frac{(\lambda\tau)^k}{k!} \\
&\le \omega_L(f;\delta) + \|T(t)\|\cdot\|f\|_{\mathcal{X}}\cdot h^{4/3}\, Var\left(X_\lambda\right) + M\|f\|_{\mathcal{X}}\cdot e^{\lambda(\tau-1)}P\left(\left|X_{\lambda\tau} - \lambda\right| > h^{-2/3}\right) \\
&\le \omega_L(f;\delta) + \|T(t)\|\cdot\|f\|_{\mathcal{X}}\cdot t\cdot h^{1/3} + M\|f\|_{\mathcal{X}}\cdot e^{\lambda(\tau-1)}\cdot h^{4/3} E\left(X_{\lambda\tau} - \lambda\right)^2 \\
&\le \omega_L(f;\delta) + \|T(t)\|\cdot\|f\|_{\mathcal{X}}\cdot t\cdot h^{1/3} + M\|f\|_{\mathcal{X}}\cdot e^{\lambda(\tau-1)}\cdot h^{4/3}\left\{Var\left(X_{\lambda\tau}\right) + \lambda^2(\tau-1)^2\right\} \\
&\le \omega_L(f;\delta) + \|T(t)\|\cdot\|f\|_{\mathcal{X}}\cdot t\cdot h^{1/3} + M\|f\|_{\mathcal{X}}\cdot e^{\lambda(\tau-1)}\cdot h^{4/3}\left\{\lambda\tau + \lambda^2(\tau-1)^2\right\} \\
&\le \omega_L(f;\delta) + h^{1/3}\|f\|_{\mathcal{X}}\left\{M(L-1)e^{\omega(L-1)} + ML^2\exp\left((L-1)e^{\omega}\right)\right\} \\
&= \omega_L\left(f;h^{1/3}\right) + h^{1/3}K\|f\|_{\mathcal{X}}
\end{aligned}
$$

mit $K := M(L-1)e^{\omega(L-1)} + ML^2\exp\left((L-1)e^{\omega}\right)$. Dabei ist zu beachten, dass $k \in J$ genau dann gilt, wenn $|kh - t| \le h^{1/3} \le 1$ ist, was wegen $t \le L-1$ seinerseits $kh \le t+1 \le L$ für $k \in J$ impliziert. Im vorletzten Schritt wurde ferner von der Ungleichung

$$h^{4/3}\left(\lambda\tau + \lambda^2(\tau-1)^2\right) \le h^{4/3}\left(2\lambda + \lambda^2 h\right) = h^{4/3}\frac{2t+t^2}{h} \le h^{1/3}(t+1)^2 \le h^{1/3}L^2$$

Gebrauch gemacht. Damit ist die zweite Aussage des Satzes bewiesen, aus der die erste unmittelbar folgt. ■

Bemerkung: Der Bereich für h, in dem die im Satz angegebene Ungleichung gültig ist, kann nach dem obigen Beweisgang explizit angegeben werden, indem man die Gleichung $\omega h = \ln\left(1+\sqrt{h}\right)$ nach h mit der (unter den gegebenen Voraussetzungen) eindeutigen Lösung, etwa $h = h_\omega$, auflöst. Die Ungleichung gilt dann im Intervall $h \in \left[0, h_\omega\right]$. Für (etwa durch $M \ge 1$) gleichmäßig beschränkte bzw. Kontraktions-Halbgruppen zeigt der obige Beweis, dass die Abschätzung in Satz 44 noch vereinfacht werden kann zu

$$\left\|T(t)f-\exp\left(tA_h\right)f\right\|_{\mathcal{X}} \le \omega_L\left(f;h^{1/3}\right)+2h^{1/3}M\left\|f\right\|_{\mathcal{X}} \text{ für alle } t\in[0,L-1],\ h>0 \text{ und } f\in\mathcal{X}.$$

Eine besonders wichtige Anwendung dieses Satzes ergibt sich für den Fall der Translations-Halbgruppe aus Beispiel 22.

Satz 45. Es sei $\mathcal{X}=UCB(\mathbb{R})$ der Banach-Raum der auf $\mathbb{R}$ definierten, gleichmäßig stetigen und beschränkten Funktionen. Dann gilt die verallgemeinerte Taylor-Formel:

$$f(x+t)=\lim_{h\downarrow 0}\sum_{k=0}^{\infty}\frac{t^k}{k!}\delta_h^k f(x) \text{ für jedes } f\in\mathcal{X} \text{ und } x\in\mathbb{R},$$

gleichmäßig bzgl. t in jedem kompakten Teilintervall von $\mathbb{R}^+$. Dabei ist

$$\delta_h f(x) := \frac{f(x+h)-f(x)}{h} \text{ für } x\in\mathbb{R} \text{ und } h>0.$$

Beweis: Mit der Halbgruppe der Linkstranslationen (Beispiel 22) ergibt sich mit Satz 44 für jedes $f\in\mathcal{X}$:

$$f(x+t)=T(t)f(x)=\mathbf{s}\text{-}\lim_{h\downarrow 0}\exp\left(tA_h\right)f(x)=\lim_{h\downarrow 0}\sum_{k=0}^{\infty}\frac{t^k}{k!}\delta_h^k f(x) \text{ für jedes } x\in\mathbb{R} \text{ und } t\ge 0,$$

da

$$A_h f(x)=\frac{f(x+h)-f(x)}{h}=\delta_h f(x) \text{ ist für jedes } x\in\mathbb{R} \text{ und } h>0,$$

und $\exp\left(tA_h\right)f=\sum_{k=0}^{\infty}\frac{t^k}{k!}A_h^k f$ gilt. Die Gleichmäßigkeitsaussage folgt aus dem zweiten Teil von Satz 44. ■

Bemerkung: Satz 45 ist deswegen bemerkenswert, weil hier eine approximative Reihenentwicklung ohne jegliche Differenzierbarkeitsvoraussetzungen möglich ist. Für die iterierten Differenzen gilt dabei noch folgende, aus dem binomischen Lehrsatz folgende Darstellung:

$$\delta_h^k f(x)=\frac{1}{h^k}\sum_{i=0}^{k}(-1)^{k-i}\binom{k}{i}f(x+ih) \text{ für alle } x\in\mathbb{R},\ k\in\mathbb{N} \text{ und } h>0.$$

II. 2.2. Stochastische Darstellungssätze

In diesem Abschnitt wollen wir zeigen, dass Satz 44 auf sehr elegante Art und Weise erweitert werden kann, indem das im dortigen Beweis verwendete probabilistische Argument entsprechend verallgemeinert wird. Die Darstellung folgt hier im Wesentlichen der Original-Arbeit von PFEIFER (1984).

Ein wichtiges Werkzeug sind dabei "Erwartungswerte" Banach-Raum- bzw. Operator-wertiger Abbildungen. Allerdings treten hier Messbarkeitsprobleme auf, weil diese Abbildungen in der Regel nicht separabel-wertig sind.

Für das Folgende betrachten wir verschiedene Borel'sche σ-Algebren: zum einen die von den offenen Mengen in $\mathcal{X}$ erzeugte σ-Algebra $\mathfrak{B}(\mathcal{X})$, zum anderen die von den offenen Mengen der Banach-Algebra $\mathfrak{L}[\mathcal{X},\mathcal{X}]$ erzeugte σ-Algebra $\mathfrak{B}[\mathcal{X},\mathcal{X}]$.

Lemma 51. Es sei $\mathfrak{T}=\{T(t)\,|\,t\geq 0\}$ eine Halbgruppe der Klasse (C_0). Ist die Halbgruppe in einem genügend kleinen Intervall $[0,\delta]$ mit $\delta>0$ injektiv und existiert ein $t_0>0$ mit der Eigenschaft $\liminf\limits_{t\downarrow t_0}\|T(t)-T(t_0)\|>0,$ so ist die Operator-wertige Abbildung $t\mapsto T(t)$ nicht $\mathfrak{B}[\mathcal{X},\mathcal{X}]$-messbar.

Beweis: Nach Voraussetzung existiert ein $c>0$ mit $\inf\limits_{t_0\leq t\leq t_0+\delta}\|T(t)-T(t_0)\|\geq c,$ so dass mit der Abschätzung aus Lemma 45, Teil d) für ein $\omega>\omega_0$ folgt:

$$\inf_{s\leq t\leq s+\delta}\|T(t)-T(t_0)\|\geq\frac{c}{M(\omega)}e^{-\omega t_0}=:c^*>0.$$

O.B.d.A. können wir $\delta\leq t_0$ annehmen, so dass die letzte Ungleichung auch

$$\|T(t)-T(s)\|\geq c^*\quad\text{für } s,t\in[0,\delta]$$

impliziert. Sei nun $N\subset[0,\delta]$ eine nicht-Borel'sche Menge (eine solche kann mit Hilfe des Auswahlaxioms konstruiert werden; vgl. das Skript STOCHASTIK, Satz 15). Dann besteht die Menge $T(N)=\{T(t)\,|\,t\in N\}$ aus überabzählbar vielen, paarweise separierten Punkten, d.h. $T(N)$ ist abgeschlossen in $\mathfrak{L}[\mathcal{X},\mathcal{X}]$ (aber nicht separabel) und gehört daher zu $\mathfrak{B}[\mathcal{X},\mathcal{X}]$. Das Urbild dieser Menge, $N=T^{-1}(T(N)),$ ist aber nach Wahl nicht-messbar, womit das Lemma bewiesen ist. ■

Beispiel 32. Eine Halbgruppe, auf die die Voraussetzungen von Lemma 51 zutrifft, ist die Halbgruppe der Linkstranslationen aus Beispiel 22. hier gilt nämlich:

$$\|T(t)-T(s)\| = 2 \text{ für alle } s,t \geq 0,\ s \neq t.$$

Für jedes $\varepsilon > 0$ gibt es nämlich eine schwach monoton wachsende Funktion $f_\varepsilon \in UCB(\mathbb{R})$ mit der Eigenschaft: $f_\varepsilon(x) = \begin{cases} -1, & x \leq -\varepsilon \\ 1, & x \geq \varepsilon. \end{cases}$ Jede solche Funktion besitzt offensichtlich die Eigenschaft $\|f_\varepsilon\|_{UCB(\mathbb{R})} = 1$. Für $s \neq t$ mit O.B.d.A. $s < t$ und $0 < \varepsilon < \frac{t-s}{2}$ ist nun

$$\|T(t)f_\varepsilon - T(s)f_\varepsilon\|_{UCB(\mathbb{R})} = \sup_{x\in\mathbb{R}} |f(x+t) - f(x+s)| \geq |f(x_0+t) - f(x_0+s)| = 2$$

für die Wahl $x_0 := -\frac{t+s}{2}$, weil dann

$$f(x_0+t) = f\left(\frac{t-s}{2}\right) = f(\varepsilon) = 1 \text{ und } f(x_0+s) = f\left(-\frac{t-s}{2}\right) = f(-\varepsilon) = -1$$

gilt. Andererseits ist $\|T(t)-T(s)\| \leq \|T(t)\| + \|T(s)\| = 2,$ womit alles gezeigt ist. Ferner ist die Abbildung $t \mapsto T(t)$ trivialerweise injektiv.

Für die Bildung von "Erwartungswerten" für Abbildungen der Art $T(X)$ mit Zufallsvariablen X, die Werte in $\mathbb{R}^+$ annehmen, benötigen wir allerdings noch weitere Integralbegriffe, die über das in Abschnitt II.1.1 definierte Riemann-Integral hinausgehen. Wir beginnen mit dem so genannten *Bochner-Integral* (eingeführt 1932 von Salomon Bochner, polnischer Mathematiker, 1899 – 1982).

Definition 54. Es sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum und $(\mathcal{X}, \|\bullet\|_{\mathcal{X}})$ ein Banach-Raum. Eine Abbildung $\varphi : \Omega \to \mathcal{X}$ heißt *Elementarfunktion*, wenn sie in der Form

$$\varphi(\omega) = \sum_{i=1}^{n} \alpha_i \mathbb{1}_{A_i}(\omega),\ \omega \in \Omega$$

mit $\alpha_1, \cdots, \alpha_n \in \mathcal{X}$ und $n \in \mathbb{N}$ darstellbar ist. Sie heißt *stark messbar*, wenn es eine Folge $\{\varphi_n\}_{n\in\mathbb{N}}$ von Elementarfunktionen gibt mit $\varphi = \mathbf{s}\text{-}\lim_{n\to\infty} \varphi_n$. Das *Bochner-Integral* einer Elementarfunktion φ ist definiert als

$$B\int \varphi\, d\mu := \sum_{i=1}^{n} \mu(A_i)\alpha_i.$$

Eine stark messbare Abbildung φ heißt *Bochner-integrierbar*, wenn es eine Folge $\{\varphi_n\}_{n\in\mathbb{N}}$ von Elementarfunktionen gibt mit

$$\varphi = \mathbf{s}\text{-}\lim_{n\to\infty} \varphi_n \quad \mu\text{-fast überall} \quad \text{und} \quad \lim_{n\to\infty} \int \|\varphi - \varphi_n\|_{\mathcal{X}}\, d\mu = 0.$$ [28]

[28] Genauer sind hier die punktweisen Normen gemeint, d.h. Integrale der Form $\int \|\varphi(\omega) - \varphi_n(\omega)\|_{\mathcal{X}}\, \mu(d\omega)$.

In diesem Fall existiert der Grenzwert $\mathbf{s}\text{-}\lim_{n\to\infty} B\int \varphi_n \, d\mu$ in $\mathcal{X}$ und wird mit $B\int \varphi \, d\mu$ bezeichnet (*Bochner-Integral* von φ bzgl. μ).

Bemerkung: Das Bochner-Integral ähnelt in seinem Aufbau dem Lebesgue-Integral; es verfügt auch über analoge Eigenschaften. Für weitergehende Details, die wir hier nicht behandeln, sei auf MIKUSINSKI (1978) verwiesen.

Lemma 52. Unter den Voraussetzungen von Definition 54 gilt:

a) Ist φ stark messbar, so ist $\|\varphi\|_{\mathcal{X}}$ messbar im üblichen Sinne.

b) Für Elementarfunktionen φ gilt: $\left\| B\int \varphi \, d\mu \right\|_{\mathcal{X}} \le \int \|\varphi\|_{\mathcal{X}} \, d\mu$.

c) Sind φ_1, φ_2 Elementarfunktionen und $\beta_1, \beta_2 \in \mathbb{R}$, so ist $\varphi = \beta_1\varphi_1 + \beta_2\varphi_2$ ebenfalls eine Elementarfunktion, und es gilt

$$B\int \varphi \, d\mu = \beta_1 \cdot B\int \varphi_1 \, d\mu + \beta_2 \cdot B\int \varphi_2 \, d\mu.$$

d) Das Bochner-Integral ist eindeutig bestimmt.

Beweis:

a) Folgt aus der Darstellung

$$\varphi^{-1}\big(K_r(\mathbf{0})\big) = \{\omega \in \Omega \mid \|\varphi(\omega)\| < r\} = \bigcup_{k\in\mathbb{N}} \bigcup_{m\in\mathbb{N}} \bigcap_{n\ge m} \left\{\omega \in \Omega \mid \|\varphi_n(\omega)\| < r - \frac{1}{k}\right\}$$

für $r > 0$ und $\varphi = \mathbf{s}\text{-}\lim_{n\to\infty} \varphi_n$ mit Elementarfunktionen $\{\varphi_n\}_{n\in\mathbb{N}}$.

b) Trivial wegen $\left\| B\int \varphi \, d\mu \right\|_{\mathcal{X}} = \left\| \sum_{i=1}^{n} \mu(A_i)\alpha_i \right\|_{\mathcal{X}} \le \sum_{i=1}^{n} \mu(A_i)\|\alpha_i\|_{\mathcal{X}} = \int \|\varphi\|_{\mathcal{X}} \, d\mu.$

c) Folgt unmittelbar aus der Definition des Bochner-Integrals.

Die Wohldefiniertheit des Bochner-Integrals für Elementarfunktionen folgt wie beim Lebesgue-Integral, vgl. das Skript STOCHASTIK, Lemma 18. Ist nun $\varphi = \mathbf{s}\text{-}\lim_{n\to\infty} \varphi_n$ mit Elementarfunktionen $\{\varphi_n\}_{n\in\mathbb{N}}$ Bochner-integrierbar, so gilt

$$\left\| B\int \varphi_n \, d\mu - B\int \varphi_m \, d\mu \right\|_{\mathcal{X}} \le \int \|\varphi_n - \varphi_m\|_{\mathcal{X}} \, d\mu \le \int \|\varphi - \varphi_n\|_{\mathcal{X}} \, d\mu + \int \|\varphi - \varphi_m\|_{\mathcal{X}} \, d\mu \to 0$$

für $n, m \to \infty$, d.h. $\left\{B\int \varphi_n \, d\mu\right\}_{n\in\mathbb{N}}$ ist Cauchy-Folge in $\mathcal{X}$ und damit stark konvergent. Ist $\{\psi_n\}_{n\in\mathbb{N}}$ eine weitere Folge von Elementarfunktionen mit $\varphi = \mathbf{s}\text{-}\lim_{n\to\infty} \psi_n$, so folgt ähnlich

$$\left\| B\int \varphi_n \, d\mu - B\int \psi_n \, d\mu \right\|_{\mathcal{X}} \le \int \|\varphi - \varphi_n\|_{\mathcal{X}} \, d\mu + \int \|\varphi - \psi_n\|_{\mathcal{X}} \, d\mu \to 0 \text{ für } n \to \infty,$$

d.h. die Grenzwerte $\mathbf{s}\text{-}\lim_{n\to\infty} B\int \varphi_n \, d\mu$ und $\mathbf{s}\text{-}\lim_{n\to\infty} B\int \psi_n \, d\mu$ stimmen überein. ■

Bemerkung: Ähnlich wie beim Lebesgue-Integral wird als Konvention definiert:

$$B\int_A \varphi\, d\mu := B\int \mathbb{1}_A \cdot \varphi\, d\mu.$$

Satz 46 (Bochner 1932)**.** Es sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum und $\left(\mathcal{X}, \|\bullet\|_{\mathcal{X}}\right)$ ein Banach-Raum. Eine stark messbare Abbildung φ ist genau dann Bochner-integrierbar, wenn $\|\varphi(\bullet)\|_{\mathcal{X}}$ im gewöhnlichen Sinne integrierbar ist. Ferner gilt: Ist $\{\varphi_n\}_{n\in\mathbb{N}}$ eine Folge von Elementarfunktionen mit

$$\varphi = \mathbf{s}\text{-}\lim_{n\to\infty} \varphi_n \quad \mu\text{-fast überall} \quad \text{und} \quad \lim_{n\to\infty} \int \|\varphi - \varphi_n\|_{\mathcal{X}}\, d\mu = 0,$$

so gilt auch

$$\int \|\varphi\|_{\mathcal{X}}\, d\mu = \lim_{n\to\infty} \int \|\varphi_n\|_{\mathcal{X}}\, d\mu.$$

Beweis:

"$\Rightarrow$": Folgt aus $\|\varphi\|_{\mathcal{X}} \leq \|\varphi - \varphi_n\|_{\mathcal{X}} + \|\varphi_n\|_{\mathcal{X}}$ für alle $n \in \mathbb{N}$, wenn $\varphi = \mathbf{s}\text{-}\lim_{n\to\infty} \varphi_n$ ist.

"$\Leftarrow$": Für stark messbares φ existiert eine Folge von Elementarfunktionen mit $\varphi = \mathbf{s}\text{-}\lim_{n\to\infty} \varphi_n$ μ-f.ü. , also gilt auch $\|\varphi\|_{\mathcal{X}} = \lim_{n\to\infty} \|\varphi_n\|_{\mathcal{X}}$ μ-f.ü. Definiere

$$\psi_n(\omega) := \begin{cases} \varphi_n(\omega), & \text{falls } \|\varphi_n(\omega)\|_{\mathcal{X}} < 2\|\varphi(\omega)\|_{\mathcal{X}} \\ \mathbf{0}, & \text{sonst.} \end{cases}$$

Dann sind auch alle ψ_n Elementarfunktionen mit $\varphi = \mathbf{s}\text{-}\lim_{n\to\infty} \psi_n$ μ-f.ü. und

$$\|\varphi - \psi_n\|_{\mathcal{X}} \leq 3\|\varphi\|_{\mathcal{X}} \quad \mu\text{-f.ü. für alle } n \in \mathbb{N}.$$

Mit dem Satz von Lebesgue von der majorisierten Konvergenz folgt nun:

$$\lim_{n\to\infty} \int \|\varphi - \psi_n\|_{\mathcal{X}}\, d\mu = \int \lim_{n\to\infty} \|\varphi - \psi_n\|_{\mathcal{X}}\, d\mu = 0.$$

Nach Definition 54 ist also φ Bochner-integrierbar. ■

Lemma 53. Für eine Bochner-integrierbare Abbildung φ gilt: $\left\| B\int \varphi\, d\mu \right\|_{\mathcal{X}} \leq \int \|\varphi\|_{\mathcal{X}}\, d\mu.$

Beweis: Nach Definition ist $\varphi = \mathbf{s}\text{-}\lim_{n\to\infty} \varphi_n$ starker Grenzwert von Elementarfunktionen, woraus folgt:

$$\left\| B\int \varphi\, d\mu \right\|_{\mathcal{X}} = \lim_{n\to\infty} \left\| B\int \varphi_n\, d\mu \right\|_{\mathcal{X}} \leq \lim_{n\to\infty} \int \|\varphi_n\|_{\mathcal{X}}\, d\mu = \int \|\varphi\|_{\mathcal{X}}\, d\mu$$

nach Satz 46. ■

Lemma 54. Ist g eine nicht-negative Lebesgue-integrierbare Abbildung und φ stark messbar mit $\|\varphi(\bullet)\|_{\mathcal{X}} \le g$ μ-f.ü., so ist φ Bochner-integrierbar. Sind ferner $\{\psi_n\}_{n\in\mathbb{N}}$ und ψ stark messbar mit $\|\psi_n(\bullet)\|_{\mathcal{X}} \le g$ und $\varphi = \mathbf{s}\text{-}\lim_{n\to\infty} \psi_n$ μ-f.ü., so sind alle ψ_n und ψ Bochner-integrierbar. Ferner gilt:

$$\lim_{n\to\infty} \int \|\psi - \psi_n\|_{\mathcal{X}}\, d\mu = 0 \quad \text{impliziert} \quad B\int \psi\, d\mu = \mathbf{s}\text{-}\lim_{n\to\infty} B\int \psi_n\, d\mu.$$

Beweis: Im ersten Teil ist $\|\varphi(\bullet)\|_{\mathcal{X}}$ messbar und daher im gewöhnlichen Sinne integrierbar; die Aussage folgt dann unmittelbar aus dem Satz 46 von Bochner. Im zweiten Teil setze $g_n := g - \frac{1}{2}\|\psi - \psi_n\|_{\mathcal{X}} \ge 0;$ dann ist $\int g_n\, d\mu \le \int g\, d\mu < \infty$. Nach dem Lemma von Fatou (vgl. das Skript STOCHASTIK, Lemma 22) gilt aber auch

$$\int g\, d\mu \le \liminf_{n\to\infty} \int g_n\, d\mu = \int g\, d\mu - \limsup_{n\to\infty} \int \|\psi - \psi_n\|_{\mathcal{X}}\, d\mu.$$

Wegen $\int g\, d\mu < \infty$ folgt hieraus $\limsup_{n\to\infty} \int \|\psi - \psi_n\|_{\mathcal{X}}\, d\mu = 0$ und damit die Behauptung. ■

Abschließend wollen wir noch einen Zusammenhang zwischen der starken Messbarkeit und der Messbarkeit bezüglich der Borel'schen σ-Algebra $\mathfrak{B}(\mathcal{X})$ aufzeigen. Wir formulieren die Ergebnisse hier ohne Beweis, für Details sei auf MIKUSINSKI (1978) und KELLEY / SRINIVASAN (1988) verwiesen.

Lemma 55. Ist $(\mathcal{Y}, \|\bullet\|_{\mathcal{Y}})$ ein weiterer Banach-Raum und $T \in \mathfrak{L}[\mathcal{X},\mathcal{Y}]$ ein beschränkter linearer Operator, dann ist für eine $\mathcal{X}$-wertige Bochner-integrierbare Abbildung φ die $\mathcal{Y}$-wertige Abbildung $T\varphi$ ebenfalls Bochner-integrierbar mit

$$B\int T\varphi\, d\mu = T\left(B\int \varphi\, d\mu\right).$$

Beweis: Zu φ existiert eine Folge $\{\varphi_n\}_{n\in\mathbb{N}}$ von Elementarfunktionen mit $\varphi = \mathbf{s}\text{-}\lim_{n\to\infty} \varphi_n$. Damit folgt aber $T\varphi = \mathbf{s}\text{-}\lim_{n\to\infty} T\varphi_n$ wegen

$$\|T\varphi - T\varphi_n\|_{\mathcal{Y}} \le \|T\| \cdot \|\varphi - \varphi_n\|_{\mathcal{X}} \to 0 \quad \mu\text{-f.ü.},$$

d.h. $T\varphi$ ist stark messbar. Aus der Linearität von T ergibt sich ferner $B\int T\varphi_n\, d\mu = T\left(B\int \varphi_n\, d\mu\right)$ und damit nach Grenzübergang die Behauptung. ■

Lemma 56. Ist der Maßraum $(\Omega, \mathcal{A}, \mu)$ vollständig[29] und σ-endlich, dann sind folgende Aussagen für eine Abbildung $\varphi : \Omega \to \mathcal{X}$ äquivalent:

a) Es gibt eine μ-Nullmenge $\mathcal{N} \subseteq \mathcal{X}$ derart, dass $\varphi : \Omega \setminus \mathcal{N} \to \mathcal{X}$ $\mathfrak{B}(\mathcal{X})$-messbar ist und $\varphi(\Omega \setminus \mathcal{N})$ eine abzählbare und in $\mathcal{X}$ dichte Teilmenge enthält (d.h. in diesem Sinn ist φ μ-f.ü. separabel-wertig).

b) Es existiert eine μ-Nullmenge $\mathcal{N} \subseteq \mathcal{X}$ und eine Folge $\{\varphi_n\}_{n\in\mathbb{N}}$ von Elementarfunktionen mit $\varphi(\omega) = \mathbf{s}\text{-}\lim_{n\to\infty} \varphi_n(\omega)$ für alle $\omega \in \Omega \setminus \mathcal{N}$.

Insbesondere ist die Abbildung φ genau dann stark messbar im Sinne von Definition 54, wenn sie $\mathfrak{B}(\mathcal{X})$-messbar ist und μ-f.ü. separabel-wertig.

Wir wollen jetzt die Grundlagen für die angekündigten stochastischen Darstellungssätze mit Hilfe des Bochner-Integrals präzisieren.

Lemma 57. Es sei X eine nicht-negative Zufallsvariable auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$ und $\mathfrak{T} = \{T(t) \mid t \geq 0\}$ eine Halbgruppe der Klasse (C_0). Ferner existiere die momenterzeugende Funktion $\psi_X(t) = E\left(e^{tX}\right)$ für $t \in [0, \eta]$[30] mit einem geeigneten $\eta > \omega_0$, dem charakteristischen Exponenten der Halbgruppe nach Lemma 45. Dann existiert für jedes $f \in \mathcal{X}$ das Bochner-Integral (Erwartungswert)

$$E[T(X)f] := B\int T(X)f\, dP,$$

mit

$$\left\| E[T(X)f] \right\|_{\mathcal{X}} \leq \int \left\| T(X)f \right\|_{\mathcal{X}} dP \leq M(\eta)\psi_X(\eta) \left\| f \right\|_{\mathcal{X}}.$$

Beweis: Die Abbildung $\omega \mapsto T(X(\omega))f$ ist stark messbar für jedes $f \in \mathcal{X}$ nach Lemma 56, da $T(t)f$ in $t \geq 0$ stark stetig (und damit $\mathfrak{B}(\mathcal{X})$-messbar) ist und X im üblichen Sinne messbar. Wegen $\left\| T(X(\omega))f \right\|_{\mathcal{X}} \leq M(\eta) e^{\eta X(\omega)} \left\| f \right\|_{\mathcal{X}}$ und der vorausgesetzten Existenz der momenterzeugenden Funktion von X ist also $M(\eta) e^{\eta X} \left\| f \right\|_{\mathcal{X}}$ eine integrierbare Majorante, so dass die Bochner-Integrierbarkeit von $T(X)f$ zusammen mit der angegebenen Abschätzung aus Lemma 53 und 54 folgt. ■

---

[29] D.h. alle Teilmengen von μ-Nullmengen gehören zu $\mathcal{A}$.

[30] Auf Grund der Nichtnegativität von X ist dies äquivalent damit, dass die momenterzeugende Funktion dann auch für alle reellen $t \leq \eta$ existiert.

Bemerkung: Ohne die Anwendung auf ein Element $f \in \mathcal{X}$ ist die Abbildung $\omega \mapsto T\big(X(\omega)\big)$ i. Allg. nicht stark messbar, wie Lemma 51 zeigt, da die Abbildung $t \mapsto T(t)$ i. Allg. nicht separabel-wertig ist. Daher kann auch auf diesem Weg i. Allg. kein "Erwartungswert" $E\big[T(X)\big]$ definiert werden. Hierzu ist ein weiterer Integral-Begriff nötig, nämlich der des *Pettis-Integrals*.

Definition 55 (Pettis-Integral, B.J. Pettis 1938). Es sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum und $\left(\mathcal{X}, \|\bullet\|_{\mathcal{X}}\right)$ ein Banach-Raum. Eine Abbildung $\varphi : \Omega \to \mathcal{X}$ heißt *schwach messbar*, wenn $f^*(\varphi)$ messbar ist für alle beschränkten linearen Funktionale $f^* \in \mathcal{X}^*$. φ heißt *Pettis-integrierbar*, wenn φ schwach messbar ist und ein Element $J \in \mathcal{X}$ existiert mit

$$f^*(J) = \int f^*(\varphi)\, d\mu \text{ für alle } f^* \in \mathcal{X}^*.$$

In diesem Fall heißt J das *Pettis-Integral* von φ bezüglich μ, in Zeichen: $J = P\int \varphi\, d\mu$.

Lemma 58. Das Pettis-Integral besitzt u.a. folgende Eigenschaften:

a) Sind φ_1, φ_2 Pettis-integrierbar und $\beta_1, \beta_2 \in \mathbb{R}$, so ist $\varphi = \beta_1\varphi_1 + \beta_2\varphi_2$ ebenfalls Pettis-integrierbar, und es gilt

$$P\int \varphi\, d\mu = \beta_1 \cdot P\int \varphi_1\, d\mu + \beta_2 \cdot P\int \varphi_2\, d\mu.$$

b) Das Pettis-Integral ist im Falle der Existenz eindeutig bestimmt.

c) Ist φ Bochner-integrierbar, so auch Pettis-integrierbar, und es gilt

$$B\int \varphi\, d\mu = P\int \varphi\, d\mu.$$

Beweis:

a) Folgt aus der Linearität von f^* für alle $f^* \in \mathcal{X}^*$.

b) Folgt aus Lemma 26.

c) Sei zunächst $\varphi = \sum_{i=1}^{n} \alpha_i \mathbb{1}_{A_i}$ eine Elementarfunktion. Dann ist

$$P\int \varphi\, d\mu = \sum_{i=1}^{n} \mu(A_i)\alpha_i = B\int \varphi\, d\mu,$$

da $f^*(\varphi) = \sum_{i=1}^{n} f^*(\alpha_i)\mathbb{1}_{A_i}$ für alle $f^* \in \mathcal{X}^*$ eine Elementarfunktion im üblichen Sinne und damit messbar, also φ schwach messbar ist mit

$$f^*\left(B\int \varphi\, d\mu\right) = f^*\left(\sum_{i=1}^{n} \mu(A_i)\alpha_i\right) = \sum_{i=1}^{n} \mu(A_i) f^*(\alpha_i) = \int f^*(\varphi)\, d\mu.$$

Ist nun φ Bochner-integrierbar, so gibt es eine Folge $\{\varphi_n\}_{n\in\mathbb{N}}$ von Elementarfunktionen mit $\varphi = \mathbf{s}\text{-}\lim\limits_{n\to\infty} \varphi_n$ und $B\int \varphi\, d\mu = \mathbf{s}\text{-}\lim\limits_{n\to\infty} B\int \varphi_n\, d\mu$. Damit existiert aber für jedes $f^* \in \mathcal{X}^*$ auch der Grenzwert

$$\lim_{n\to\infty} f^*\left(B\int \varphi_n\, d\mu\right) = \lim_{n\to\infty} \int f^*(\varphi_n)\, d\mu = f^*\left(B\int \varphi\, d\mu\right),$$

d.h. es gilt $B\int \varphi\, d\mu = P\int \varphi\, d\mu$. ∎

Das Pettis-Integral ist im Unterschied zum Bochner-Integral nicht "konstruktiv"; seine Bestimmung hängt wesentlich von der Struktur des Dualraums $\mathcal{X}^*$ ab. Dies kann z.B. bei Banach-Algebren wie $\mathfrak{L}[\mathcal{X},\mathcal{Y}]$ und deren Dualräumen zu Problemen führen. Man kann diesen Integralbegriff aber abschwächen, wenn man sich auf einen geeigneten Teilraum des Dualraums zurückzieht, der noch die Eindeutigkeit im Fall der Existenz (Lemma 58, Teil b)) garantiert. Dies führt zum Begriff des modifizierten Pettis-Integrals.

Definition 56 (modifiziertes Pettis-Integral). Es sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum, $(\mathcal{X}, \|\bullet\|_{\mathcal{X}})$ und $(\mathcal{Y}, \|\bullet\|_{\mathcal{Y}})$ seien Banach-Räume. Eine Abbildung $S:\Omega \to \mathfrak{L}[\mathcal{X},\mathcal{Y}]$ heißt *modifizert schwach messbar*, wenn $f^*(S(\bullet)(f))$ messbar ist für alle beschränkten linearen Funktionale $f^* \in \mathcal{Y}^*$ und alle $f \in \mathcal{X}$. S heißt *modifiziert Pettis-integrierbar*, wenn S modifiziert schwach messbar ist und ein Element $J \in \mathfrak{L}[\mathcal{X},\mathcal{Y}]$ existiert mit

$$f^*(J(f)) = \int f^*(S(\bullet)(f))\, d\mu = \int f^*(S(\omega)(f))\, \mu(d\omega) \text{ für alle } f^* \in \mathcal{Y}^* \text{ und } f \in \mathcal{X}.$$

In diesem Fall heißt J das *modifizierte Pettis-Integral* von S bezüglich μ, in Zeichen: $J = P^*\int S\, d\mu$.

Lemma 59. Das modifizierte Pettis-Integral besitzt u.a. folgende Eigenschaften:

a) Ist S modifiziert Pettis-integrierbar, so ist $S(\bullet)(f):\Omega \to \mathcal{Y}$ Pettis-integrierbar für alle $f \in \mathcal{X}$. In diesem Falle ist

$$\left(P^*\int S\, d\mu\right)(f) = P\int S(\bullet)(f)\, d\mu.$$

b) Ist S Pettis-integrierbar im üblichen Sinn oder Bochner-integreirbar, dann auch modifiziert Pettis-integrierbar, und beide Integrale sind gleich.

c) Sind S_1, S_2 Pettis-integrierbar und $\beta_1, \beta_2 \in \mathbb{R}$, so ist $S = \beta_1 S_1 + \beta_2 S_2$ ebenfalls Pettis-integrierbar, und es gilt

$$P^*\int S\, d\mu = \beta_1 \cdot P^*\int S_1\, d\mu + \beta_2 \cdot P^*\int S_2\, d\mu.$$

d) Das modifizierte Pettis-Integral ist im Falle der Existenz eindeutig bestimmt.

Beweis:

a) Es sei *S* modifiziert Pettis-integrierbar. Dann existiert ein $J \in \mathfrak{L}[\mathcal{X},\mathcal{Y}]$ mit

$$f^*\left(J(f)\right) = \int f^*\left(S(\bullet)(f)\right) d\mu = \int f^*\left(S(\omega)(f)\right) \mu(d\omega) \text{ für alle } f^* \in \mathcal{Y}^* \text{ und } f \in \mathcal{X}.$$

Damit ist aber nach Definition des Pettis-Integrals für jedes (feste) $f \in \mathcal{X}$

$$J(f) = P\int S(\bullet)(f)\, d\mu.$$

b) Die erste Aussage folgt aus der Beobachtung, dass für jedes $T \in \mathfrak{L}[\mathcal{X},\mathcal{Y}]$ durch $F^*(T) := f^*\left(T(f)\right)$ für alle $f^* \in \mathcal{Y}^*$ und $f \in \mathcal{X}$ ein beschränktes lineares Funktional $F^* \in \mathfrak{L}[\mathcal{X},\mathcal{Y}]^*$ definiert wird mit $\left\|F^*\right\| \le \left\|f^*\right\| \|f\|_{\mathcal{X}}$. Die zweite Aussage folgt aus der ersten wegen Lemma 58, Teil c).

c) Folgt aus der Linearität von f^* für alle $f^* \in \mathcal{Y}^*$ und der Linearität von S_1, S_2 und *S*.

d) Nach Lemma 26 gilt für jedes $T \in \mathfrak{L}[\mathcal{X},\mathcal{Y}]$: ist $f^*\left(T(f)\right) = 0$ für alle $f^* \in \mathcal{Y}^*$ und $f \in \mathcal{X}$, so ist $T = \mathbf{0}$. Hieraus folgt die Behauptung. ■

Lemma 60. Es sei $(\Omega, \mathcal{A}, \mu)$ ein Maßraum, $\left(\mathcal{X}, \|\bullet\|_{\mathcal{X}}\right)$ und $\left(\mathcal{Y}, \|\bullet\|_{\mathcal{Y}}\right)$ seien Banach-Räume. Ferner sei eine Abbildung $S : \Omega \to \mathfrak{L}[\mathcal{X},\mathcal{Y}]$ derart gegeben, dass $S(\bullet)(f)$ stark messbar sei für jedes $f \in \mathcal{X}$, und dass $\|S(\bullet)\|$ durch eine μ-integrierbare Funktion $g \ge 0$ dominiert sei. Dann ist *S* modifiziert Pettis-integrierbar, und es gilt

$$\left\|P^*\int S\, d\mu\right\| \le \int g\, d\mu.$$

Beweis: Wegen $\|S(\bullet)(f)\|_{\mathcal{Y}} \le \|S(\bullet)\| \|f\|_{\mathcal{X}} \le g \|f\|_{\mathcal{X}}$ für jedes $f \in \mathcal{X}$ ist nach Satz 46 $S(\bullet)(f)$ Bochner-integrierbar und damit nach Lemma 58, Teil c) auch Pettis-integrierbar für alle $f \in \mathcal{X}$, mit

$$\left\|P\int S(\bullet)(f)\, d\mu\right\| = \left\|B\int S(\bullet)(f)\, d\mu\right\| \le \|f\|_{\mathcal{X}} \int g\, d\mu.$$

Die Abbildung $J : \mathcal{X} \to \mathcal{Y} : f \mapsto P\int S(\bullet)(f)\, d\mu$ definiert damit einen beschränken linearen Operator $J \in \mathfrak{L}[\mathcal{X},\mathcal{Y}]$ mit

$$\|J\| \le \int g\, d\mu.$$

Ferner gilt für alle $f^* \in \mathcal{Y}^*$ und $f \in \mathcal{X}$:

$$f^*\left(J(f)\right) = f^*\left(\int S(\bullet)(f)\, d\mu\right) = \int f^*\left(S(\bullet)(f)\right) d\mu,$$

d.h. $S(\bullet)$ ist modifiziert Pettis-integrierbar mit

$$J = P^*\int S\, d\mu \quad \text{und} \quad \|J\| = \left\|P^*\int S\, d\mu\right\| \le \int g\, d\mu,$$

wie behauptet. ■

Satz 47. Es sei X eine nicht-negative Zufallsvariable auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$ und $\mathfrak{T} = \{T(t) \mid t \geq 0\}$ eine Halbgruppe der Klasse (C_0). Ferner existiere die momenterzeugende Funktion $\psi_X(t) = E\left(e^{tX}\right)$ für $t = \eta$ mit einem geeigneten $\eta > \omega_0$, dem charakteristischen Exponenten der Halbgruppe nach Lemma 45. Dann existiert für $T(X)$ das modifizierte Pettis-Integral (Erwartungswert)

$$E[T(X)] := P^* \int T(X)\, dP,$$

mit

$$\|E[T(X)]\| \leq M(\eta)\psi_X(\eta) \quad \text{und} \quad E[T(X)]f = E[T(X)f] \quad \text{für alle } f \in \mathcal{X}.$$

Beweis: Nach Voraussetzung ist $T(X(\bullet))f$ messbar und P-f.s. separabelwertig für jedes $f \in \mathcal{X}$, also nach Lemma 56 stark messbar für jedes $f \in \mathcal{X}$ (vgl. den Beweis zu Lemma 57). Die Aussage ergibt sich jetzt mit Lemma 60, mit der dominierenden Funktion $g = M(\eta)e^{\eta X}$. ■

Das zentrale Ergebnis dieses Abschnittes ist

Satz 48. Es seien X und Y nicht-negative, stochastisch unabhängige Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$ und $\mathfrak{T} = \{T(t) \mid t \geq 0\}$ eine Halbgruppe der Klasse (C_0). Ferner existiere die momenterzeugende Funktion $\psi_X(t) = E\left(e^{tX}\right)$ für $t = \eta$ mit einem geeigneten $\eta > \omega_0$, dem charakteristischen Exponenten der Halbgruppe nach Lemma 45. Dann sind $T(X)$, $T(Y)$ und $T(X) \circ T(Y)$ modifizierte Pettis-integrierbar, und es gilt die "Produkt-Formel"

$$E[T(X) \circ T(Y)] = E[T(X)] \circ E[T(Y)].$$

Beweis: Wegen $\psi_{X+Y}(t) = E\left(e^{t(X+Y)}\right) = \psi_X(t)\psi_Y(t)$ existiert auch die momenterzeugende Funktion von $X+Y$ für $t = \eta$, womit nach Satz 47 $E[T(X) \circ T(Y)] = E[T(X+Y)]$ als Pettis-Integral existiert. Ferner wird für festes $f^* \in \mathcal{X}^*$ durch

$$h^*(h) := f^*\left(E[T(X)]h\right) = f^*\left(E[T(X)h]\right) = E\left[f^*\left(T(X)h\right)\right], \; h \in \mathcal{X}$$

Ein beschränktes lineares Funktional $h^* \in \mathcal{X}^*$ definiert mit

$$\|h^*\| \leq \|f^*\| \|E[T(X)]\| \leq M(\eta)\psi_X(\eta)\|f^*\|.$$

Es folgt nun, für $f \in \mathcal{X}$ und $f^* \in \mathcal{X}^*$:

$$\begin{aligned} f^*\left(E[T(X)] \circ E[T(Y)]f\right) &= h^*\left(E[T(Y)]f\right) = E\left[h^*\left(T(Y)f\right)\right] = E_X\left[E_Y\left[f^*\left(T(X) \circ T(Y)f\right)\right]\right] \\ &= E_X\left[E_Y\left[f^*\left(T(X+Y)f\right)\right]\right] = \int\int f^*\left(T(x+y)f\right)P^{(X,Y)}(dx,dy) \\ &= E\left[f^*\left(T(X+Y)f\right)\right] = E\left[f^*\left(T(X) \circ T(Y)f\right)\right] = f^*\left(E[T(X) \circ T(Y)]f\right) \end{aligned}$$

unter Anwendung des Satzes von Fubini, woraus mit Lemma 26 die behauptete Gleichheit $E[T(X)\circ T(Y)] = E[T(X)]\circ E[T(Y)]$ folgt. Mit den Symbolen E_X und E_Y wurde dabei die Integration bzgl. X bzw. Y bezeichnet. ■

Satz 49. Es sei N eine nicht-negative ganzzahlige und Y eine nicht-negative Zufallsvariable auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$ und $\mathfrak{T} = \{T(t) \mid t \geq 0\}$ eine Halbgruppe der Klasse (C_0). Ferner existiere ein $\eta > \omega_0$, so dass die momenterzeugende Funktion $\psi_Y(t) = E\left(e^{tY}\right)$ sowie die Komposition $\varphi_N\left(\psi_Y(t)\right)$ für $t = \eta$ existiere, wobei $\varphi_N(s) = E\left(s^Y\right)$ die wahrscheinlichkeitserzeugende Funktion von N bezeichne. Ist dann $\{Y_k\}_{k\in\mathbb{N}}$ eine stochastisch unabhängige Folge von Zufallsvariablen, die wie Y verteilt und auch von N unabhängig sind, und bezeichnet $X := \sum_{k=1}^{N} Y_k$, so ist $T(X)$ modifiziert Pettis-integrierbar, und es gilt

$$E[T(X)] = \varphi_N\left(E[T(Y)]\right) = \sum_{k=0}^{\infty} P(N=k)\{E[T(Y)]\}^k,$$

mit der üblichen Konvention $\{E[T(Y)]\}^0 = \mathbf{I}$.

Beweis: Zunächst ist $\varphi_N\left(\psi_Y(t)\right)$ die momenterzeugende Funktion von X mit $\varphi_N\left(\psi_Y(\eta)\right) < \infty$; damit existiert nach Satz 47 der Operator $E[T(X)]$ in $\mathfrak{L}[\mathcal{X}, \mathcal{X}]$. Ferner gilt nach Satz 48 für alle $k \in \mathbb{N}$:

$$\left\|\{E[T(Y)]\}^k\right\| = \left\|E\left[T\left(\sum_{i=1}^{k} Y_i\right)\right]\right\| \leq M(\eta)\,\psi_{\sum_{i=1}^{m} Y_i}(\eta) = M(\eta)\{\psi_Y(\eta)\}^k,$$

was wegen $M(\eta) \geq 1$ auch noch für $k = 0$ gilt. Damit erhalten wir

$$\sum_{k=0}^{\infty} P(N=k)\left\|\{E[T(Y)]\}^k\right\| \leq M(\eta)\sum_{k=0}^{\infty} P(N=k)\{\psi_Y(\eta)\}^k = M(\eta)\varphi_N\left(\psi_Y(\eta)\right) < \infty,$$

was impliziert, dass $\left\{\sum_{k=0}^{n} P(N=k)\{E[T(Y)]\}^k\right\}_{n\in\mathbb{N}}$ eine Cauchy-Folge ist wegen

$$\left\|\sum_{k=n}^{n+m} P(N=k)\{E[T(Y)]\}^k\right\| \leq \sum_{k=n}^{n+m} P(N=k)\left\|\{E[T(Y)]\}^k\right\| \leq M(\eta)\sum_{k=n}^{n+m} P(N=k)\{\psi_Y(\eta)\}^k \to 0$$

für $n, m \to \infty$ und damit in $\mathfrak{L}[\mathcal{X}, \mathcal{X}]$ gegen

$$\sum_{k=0}^{\infty} P(N=k)\{E[T(Y)]\}^k = \varphi_N\left(E[T(Y)]\right)$$

konvergiert. Für $f \in \mathcal{X}$ und $f^* \in \mathcal{X}^*$ folgt also

$$f^*\left(E[T(X)]f\right) = E\left[f^*\left(T(X)f\right)\right] = \sum_{k=0}^{\infty} P(N=k)\, E\left[f^*\left(T\left(\sum_{i=1}^{k} Y_i\right)f\right)\right]$$
$$= \sum_{k=0}^{\infty} P(N=k)\, f^*\left(E\left[T\left(\sum_{i=1}^{k} Y_i\right)\right]f\right) = \sum_{k=0}^{\infty} P(N=k) f^*\left(\left\{E[T(Y)]\right\}^k f\right)$$
$$= f^*\left(\sum_{k=0}^{\infty} P(N=k)\left\{E[T(Y)]\right\}^k f\right) = f^*\left(\varphi_N\left(E[T(Y)]\right)f\right),$$

woraus noch die Beziehung

$$E[T(X)] = \varphi_N\left(E[T(Y)]\right) = \sum_{k=0}^{\infty} P(N=k)\left\{E[T(Y)]\right\}^k$$

folgt. Damit ist der Satz bewiesen. ■

Um nun den zentralen Darstellungssatz für Operatorhalbgruppen formulieren zu können, benötigen wir noch eine etwas allgemeinere Version des Gesetzes der großen Zahlen.

Lemma 61. Es sei $\{N(\tau) \mid \tau \geq 0\}$ eine Familie nicht-negativer ganzzahliger Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$ derart, dass $\frac{1}{\tau}N(\tau)$ für $\tau \to \infty$ stochastisch gegen eine reelle Zahl $\zeta \geq 0$ konvergiere. Ist dann $\{Y_k\}_{k\in\mathbb{N}}$ eine stochastisch unabhängige Folge von Zufallsvariablen, die wie eine nicht-negative Zufallsvariable Y mit $E(Y) = \gamma \geq 0$ verteilt und auch von der Familie $\{N(\tau) \mid \tau \geq 0\}$ unabhängig sind, und bezeichnet $X(\tau) := \sum_{k=1}^{N(\tau)} Y_k$ für jedes $\tau \geq 0,$ so gilt:

$$\frac{1}{\tau}X(\tau) = \frac{1}{\tau}\sum_{k=1}^{N(\tau)} Y_k \to \varsigma\gamma \text{ stochastisch für } \tau \to \infty.$$

Beweis: Es bezeichne $A_\tau := \{N(\tau) > \zeta\tau\},\ B_\tau := \{N(\tau) < \zeta\tau\}$ für $\tau \geq 0.$ Dann gilt

$$X(\tau) = \sum_{1\leq k\leq \zeta\tau} Y_k + 1\!\!1_{A_\tau} \sum_{\zeta\tau < k \leq N(\tau)} Y_k - 1\!\!1_{B_\tau} \sum_{N(\tau) < k \leq \zeta\tau} Y_k \text{ für alle } \tau \geq 0.$$

Nach dem (starken) Gesetz der großen Zahlen gilt für den ersten Summanden, falls $\zeta > 0$:

$$\frac{1}{\tau}\sum_{1\leq k\leq \zeta\tau} Y_k = \zeta\frac{1}{\zeta\tau}\sum_{1\leq k\leq \zeta\tau} Y_k \to \zeta\gamma \quad P\text{-f.s. für } \tau \to \infty$$

(und damit auch stochastisch). Diese Konvergenzaussage ist aber trivialerweise auch für $\zeta = 0$ gültig. Für den zweiten Term wähle $\varepsilon > 0$ und $0 < \vartheta \leq \frac{\varepsilon}{1+\gamma}.$ Dann gilt, wieder mit dem starken Gesetz der großen Zahlen:

$$\frac{1}{\tau}\sum_{\zeta\tau < k \leq (\zeta+\vartheta)\tau} Y_k = (\zeta+\vartheta)\frac{1}{(\zeta+\vartheta)\tau}\sum_{k=1}^{(\zeta+\vartheta)\tau} Y_k - \zeta\frac{1}{\zeta\tau}\sum_{k=1}^{\zeta\tau} Y_k \to (\zeta+\vartheta)\gamma - \zeta\gamma = \vartheta\gamma < \varepsilon \quad P\text{-f.s. für } \tau \to \infty$$

(und damit auch stochastisch). Damit ergibt sich:

$$P\left(\mathbb{1}_{A_\tau}\frac{1}{\tau}\sum_{\zeta\tau<k\le N(\tau)}Y_k>\varepsilon\right)=P\left(A_\tau\cap\left\{\frac{1}{\tau}\sum_{\zeta\tau<k\le N(\tau)}Y_k>\varepsilon\right\}\right)$$
$$=P\left(\{\zeta\tau<N(\tau)\le(\zeta+\vartheta)\tau\}\cap\left\{\frac{1}{\tau}\sum_{\zeta\tau<k\le N(\tau)}Y_k>\varepsilon\right\}\right)+P\left(\{(\zeta+\vartheta)\tau<N(\tau)\}\cap\left\{\frac{1}{\tau}\sum_{\zeta\tau<k\le N(\tau)}Y_k>\varepsilon\right\}\right)$$
$$\le P\left(\frac{1}{\tau}\sum_{\zeta\tau<k\le(\zeta+\vartheta)\tau}Y_k>\varepsilon\right)+P\left(\frac{1}{\tau}N(\tau)>\zeta+\vartheta\right)\to 0+0=0\quad\text{für } \tau\to\infty$$

Nach dem Vorigen und der ersten Voraussetzung des Lemmas. Für den dritten Summanden argumentiert man analog. Damit gilt insgesamt:

$$\frac{1}{\tau}X(\tau)=\frac{1}{\tau}\sum_{1\le k\le\zeta\tau}Y_k+\mathbb{1}_{A_\tau}\frac{1}{\tau}\sum_{\zeta\tau<k\le N(\tau)}Y_k-\mathbb{1}_{B_\tau}\frac{1}{\tau}\sum_{N(\tau)<k\le\zeta\tau}Y_k\to\zeta\gamma\quad\text{stochastisch für } \tau\to\infty,$$

womit das Lemma bewiesen ist. ■

Lemma 62. Es sei $\{Z(\tau)\,|\,\tau\ge 0\}$ eine Familie nicht-negativer ganzzahliger Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\Omega,\mathcal{A},P)$, die für $\tau\to\infty$ stochastisch gegen eine reelle Zahl $\xi\ge 0$ konvergiere, und $\mathfrak{T}=\{T(t)\,|\,t\ge 0\}$ eine Halbgruppe der Klasse (C_0). Ferner existiere ein $\eta>\omega_0$, so dass für die momenterzeugenden Funktionen der Ausdruck $s(\eta,\tau_0):=\sup_{\tau\ge\tau_0}\{\psi_{Z(\tau)}(\eta)\}<\infty$ für ein $\tau_0\ge 0$ ist. Dann gilt:

$$E[T(Z(\tau))]f\xrightarrow{s}T(\xi)f\quad\text{für } \tau\to\infty,\text{ für alle } f\in\mathcal{X}.$$

Beweis: Wähle ein $\eta_0\in(\omega_0,\eta)$ und setze $\delta:=\dfrac{\eta-\eta_0}{\eta_0}$. Es gilt unter Verwendung der Hölder-Ungleichung (Lemma 14) und der Eigenschaften des Bochner-Integrals für $\varepsilon\in(0,1)$ und $f\in\mathcal{X}$:

$$\begin{aligned}
\left\|E[T(Z(\tau))]f-T(\xi)f\right\|_{\mathcal{X}}&=\left\|E[T(Z(\tau))-T(\xi)]f\right\|_{\mathcal{X}}\le E\left\|[T(Z(\tau))-T(\xi)]f\right\|_{\mathcal{X}}\\
&=\int\left\|[T(Z(\tau))-T(\xi)]f\right\|_{\mathcal{X}}dP\\
&=\int_{\{|Z(\tau)-\xi|<\varepsilon\}}\left\|[T(Z(\tau))-T(\xi)]f\right\|_{\mathcal{X}}dP+\int_{\{|Z(\tau)-\xi|\ge\varepsilon\}}\left\|[T(Z(\tau))-T(\xi)]f\right\|_{\mathcal{X}}dP\\
&\le\int_{\{|Z(\tau)-\xi|<\varepsilon\}}\sup_{|\zeta-\xi|<\varepsilon}\left\|[T(\zeta)-T(\xi)]f\right\|_{\mathcal{X}}dP+\|f\|_{\mathcal{X}}\int_{\{|Z(\tau)-\xi|\ge\varepsilon\}}\left\{\|T(Z(\tau))\|+\|T(\xi)\|\right\}dP\\
&\le\sup_{|\zeta-\xi|<\varepsilon}\left\|[T(\zeta)-T(\xi)]f\right\|_{\mathcal{X}}+\|f\|_{\mathcal{X}}\int_{\{|Z(\tau)-\xi|\ge\varepsilon\}}\|T(Z(\tau))\|dP+P(|Z(\tau)-\xi|\ge\varepsilon)\|T(\xi)\|\|f\|_{\mathcal{X}}\\
&\le\omega_{\xi+1}(f;\varepsilon)+\|f\|_{\mathcal{X}}P(|Z(\tau)-\xi|\ge\varepsilon)\left[\left\{\int\|T(Z(\tau))\|^{1+\delta}dP\right\}^{1/(1+\delta)}+\|T(\xi)\|\right]\\
&\le\omega_{\xi+1}(f;\varepsilon)+\|f\|_{\mathcal{X}}P(|Z(\tau)-\xi|\ge\varepsilon)M(\eta_0)\left[\{\psi_{Z(\tau)}(\eta)\}^{\eta_0/\eta}+e^{\eta_0\xi}\right]\\
&\le\omega_{\xi+1}(f;\varepsilon)+\|f\|_{\mathcal{X}}P(|Z(\tau)-\xi|\ge\varepsilon)M(\eta_0)\left[\{s(\eta)\}^{\eta_0/\eta}+e^{\eta_0\xi}\right],
\end{aligned}$$

also

$$\lim_{\tau\to\infty}\left\|E\left[T\left(Z(\tau)\right)\right]f-T(\xi)f\right\|_{\mathcal{X}}\leq\omega_{\xi+1}\left(f;\varepsilon\right)\text{ für alle }\varepsilon\in(0,1)\text{ und }f\in\mathcal{X}$$

und somit

$$\lim_{\tau\to\infty}\left\|E\left[T\left(Z(\tau)\right)\right]f-T(\xi)f\right\|_{\mathcal{X}}=0\text{ für alle }f\in\mathcal{X},$$

womit das Lemma bewiesen ist. ■

Satz 50 (Stochastischer Darstellungssatz für Operatorhalbgruppen). Es sei $\mathfrak{T}=\{T(t)\,|\,t\geq 0\}$ eine Halbgruppe der Klasse (C_0) und $\{N(\tau)\,|\,\tau\geq 0\}$ eine Familie nicht-negativer ganzzahliger Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\Omega,\mathcal{A},P)$ derart, dass $\frac{1}{\tau}N(\tau)$ für $\tau\to\infty$ stochastisch gegen eine reelle Zahl $\zeta\geq 0$ konvergiere. Ferner sei $\{Y_k\}_{k\in\mathbb{N}}$ eine stochastisch unabhängige Folge von Zufallsvariablen, die wie eine nicht-negative Zufallsvariable Y mit $E(Y)=\gamma\geq 0$ verteilt und auch von der Familie $\{N(\tau)\,|\,\tau\geq 0\}$ unabhängig sind. Existiert dann ein $\eta>\omega_0$, so dass für die jeweiligen erzeugenden Funktionen der Ausdruck

$$s(\eta,\tau_0):=\sup_{\tau\geq\tau_0}\left\{\varphi_{N(\tau)}\left(\psi_{Y/\tau}(\eta)\right)\right\}<\infty\text{ für ein }\tau_0\geq 0\text{ ist, so gilt:}$$

$$\varphi_{N(\tau)}\left(E\left[T\left(\frac{Y}{\tau}\right)\right]\right)f\xrightarrow{s}T(\xi)f\quad\text{für }\tau\to\infty,\text{ für alle }f\in\mathcal{X},\text{ mit }\xi=\varsigma\gamma.$$

Beweis: Folgt unmittelbar aus Satz 49, Lemma 61 und Lemma 62. ■

Satz 50 erlaubt eine Vielzahl von Darstellungssätzen, die praktisch alle in der Halbgruppentheorie bekannten Ergebnisse umfassen; vgl. PFEIFER (1984), BUTZER / BERENS (1967) oder GOLDSTEIN (1985)). Wir zeigen hier nur exemplarisch, wie *Hille's erste Exponentialformel* (Satz 44) leicht aus Satz 50 gewonnen werden kann. Dazu betrachten wir einen homogenen Poisson-Prozess $\{N(\tau)\,|\,\tau\geq 0\}$ mit Parameter $\xi>0$ (für Details siehe z.B. das Skript STOCHASTIK, Kapitel IV), für den die Voraussetzung von Satz 50 mit $\zeta=\xi$ erfüllt ist, sowie die konstante Zufallsvariable $Y=1$. Da für jedes $\tau>0$ die Zufallsvariable $N(\tau)$ eine Poisson-Verteilung mit Parameter $\tau\xi$ besitzt, folgt

$$\begin{aligned}s(\eta,\tau_0)&:=\sup_{\tau\geq\tau_0}\left\{\varphi_{N(\tau)}\left(\psi_{Y/\tau}(\eta)\right)\right\}=\sup_{\tau\geq\tau_0}\left\{\exp\left(\tau\xi\left(\psi_{Y/\tau}(\eta)-1\right)\right)\right\}=\sup_{\tau\geq\tau_0}\left\{\exp\left(\tau\xi\left(e^{\eta/\tau}-1\right)\right)\right\}\\&\leq\sup_{\tau\geq\tau_0}\left\{\exp\left(\xi\eta e^{1/\tau}\right)\right\}=\exp\left(\xi\eta e^{1/\tau_0}\right)<\infty\end{aligned}$$

mit beliebigem $\tau_0>0$. Es folgt nun

$$T(\xi)f=s\text{-}\lim_{\tau\to\infty}\varphi_{N(\tau)}\left(E\left[T\left(\frac{Y}{\tau}\right)\right]\right)f=s\text{-}\lim_{\tau\to\infty}\exp\left(\xi\tau\left[T\left(\frac{1}{\tau}\right)-I\right]\right)f=s\text{-}\lim_{\tau\to\infty}\exp\left(\xi A_{1/\tau}\right)f,$$

für $\tau\to\infty$ und für alle $f\in\mathcal{X}$, was gerade der Aussage des Satzes 44 entspricht.

II. 3. Spezielle Halbgruppen

In der Stochastik spielen Halbgruppen von Operatoren eine wichtige Rolle, insbesondere bei den so genannten (homogenen) *Markoff-Prozessen*. Darauf gehen wir in Abschnitt II. 3.3. noch einmal gesondert ein. Vor allem in der *Stochastischen Finanzmathematik*, wo der Wienerprozess eine fundamentale Rolle für die Beschreibung von Aktienkursen und die Bestimmung von Preisen für Derivate in stetiger Zeit spielt, werden Methoden der Halbgruppentheorie – neben der Stochastischen Integration – verstärkt eingesetzt. Für weitere Darstellungen in Bezug auf Markoff-Ketten und -Prozesse siehe auch ÇINLAR (1975), KARLIN / TAYLOR (1981), ROSS (1996) oder APPLEBAUM (2004).

II. 3.1. Die Poisson'sche Faltungshalbgruppe

Als Vorstufe zu den Halbgruppen, die bei Markoff-Prozessen wichtig sind, wollen wir zunächst zwei Typen so genannter Faltungs-Halbgruppen studieren; in diesem Abschnitt betrifft dies insbesondere Zufallsvariablen mit diskretem Zustandsraum. Dazu bezeichnen wir hier leicht abweichend mit ℓ^1 den Banach-Raum aller Folgen $f \in \mathbb{R}^{\mathbb{Z}^+}$ mit $\|f\|_1 := \sum_{n=0}^{\infty} |f_n| < \infty.$[31]

Definition 57 (Faltung). Für $f, g \in \ell^1$ ist die *Faltung* $f * g$ als Folge definiert durch

$$(f * g)(k) := \sum_{j=0}^{k} f(j)\, g(k-j) = \sum_{j=0}^{k} f(k-j)\, g(j),\ k \in \mathbb{Z}^+.$$

Lemma 63. Der Banach-Raum ℓ^1 ist abgeschlossen gegenüber der Faltungsbildung, d.h. es gilt

$$f * g \in \ell^1 \text{ für alle } f, g \in \ell^1$$

mit

$$\|f * g\|_1 \le \|f\|_1 \cdot \|g\|_1,\ f, g \in \ell^1.$$

Ferner ist die Faltungsoperation kommutativ und assoziativ.

Beweis: Es ist für $k \in \mathbb{Z}^+$

$$|(f * g)(k)| := \left| \sum_{j=0}^{k} f(j)\, g(k-j) \right| \le \sum_{j=0}^{k} |f(j)| \cdot |g(k-j)|$$

und somit

$$\begin{aligned}\sum_{k=0}^{\infty} |(f * g)(k)| &\le \sum_{k=0}^{\infty} \sum_{j=0}^{k} |f(j)| \cdot |g(k-j)| = \sum_{0 \le j \le k} \sum |f(j)| \cdot |g(k-j)| = \sum_{j=0}^{\infty} \sum_{k=j}^{\infty} |f(j)| \cdot |g(k-j)| \\ &= \sum_{j=0}^{\infty} |f(j)| \sum_{k=j}^{\infty} |g(k-j)| = \sum_{j=0}^{\infty} |f(j)| \cdot \|g\|_1 = \|g\|_1 \sum_{j=0}^{\infty} |f(j)| = \|f\|_1 \cdot \|g\|_1\end{aligned}$$

für alle $f, g \in \ell^1$. Die Kommutativität ist unmittelbar einsichtig, die Assoziativität folgt so:

[31] D.h. wir beginnen hier die Indizierung der Folgen mit 0 statt mit 1.

$$
\begin{aligned}
(f*g)*h(k) &= \sum_{j=0}^{k}(f*g)(j)\,h(k-j) = \sum_{j=0}^{k}\sum_{i=0}^{j} f(i)\,g(j-i)\,h(k-j) = \sum_{0\le i\le j\le k}\sum f(i)\,g(j-i)\,h(k-j) \\
&= \sum_{i=0}^{k}\sum_{j=i}^{k} f(i)\,g(j-i)\,h(k-j) = \sum_{i=0}^{k}\sum_{j=0}^{k-i} f(i)\,g(j)\,h(k-j-i) = \sum_{i=0}^{k} f(i)\big(g*h\big)(k-i) \\
&= f*\big(g*h\big)(k)
\end{aligned}
$$

für alle $f,g,h\in\ell^1,\ k\in\mathbb{Z}^+$. ■

Durch die Faltung mit einem (festen) Element $f\in\ell^1$ wird also ein linearer beschränkter Operator $T_f\in\mathfrak{L}\left[\ell^1,\ell^1\right]$ definiert vermöge

$$T_f g := f*g,\ \ g\in\ell^1 \ \text{ mit } \left\|T_f\right\| = \left\|f\right\|_1,$$

denn nach Lemma 63 gilt jedenfalls $\left\|T_f\right\|\le\left\|f\right\|_1$, und für die spezielle Wahl $g=(1,0,0,0,\ldots)\in\ell^1$ ergibt sich noch wegen $T_f g = f*g = f$: $\left\|T_f g\right\|_1 = \left\|f\right\|_1$.

In der Stochastik spielt die Faltung eine wesentliche Rolle bei Summen unabhängiger Zufallsvariablen, wie das folgende Ergebnis zeigt.

Lemma 64 (Summen unabhängiger Zufallsvariablen)**.** Es seien X und Y stochastisch unabhängige Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\Omega,\mathcal{A},P)$ mit Werten in $\mathbb{Z}^+$. Dann sind die Folgen $f_X := \left\{P(X=k)\right\}_{k\in\mathbb{Z}^+}$ und $f_Y := \left\{P(Y=k)\right\}_{k\in\mathbb{Z}^+}$ Elemente von ℓ^1 mit $\left\|f_X\right\|_1 = \left\|f_Y\right\|_1 = 1$, und es gilt

$$f_{X+Y} = f_X * f_Y.$$

Beweis: Es gilt

$$
\begin{aligned}
f_{X+Y}(k) &= P(X+Y=k) = \sum_{i=0}^{k} P(X=i,\,Y=k-i) = \sum_{i=0}^{k} P(X=i)\ P(Y=k-i) \\
&= \sum_{i=0}^{k} f_X(i)\ f_Y(k-i) = f_X * f_Y(k)
\end{aligned}
$$

für alle $k\in\mathbb{Z}^+$. ■

Lemma 65 (Poisson-Verteilungen)**.** Sind unter den Voraussetzungen von Lemma 64 die Zufallsvariablen X und Y Poisson-verteilt mit $P^X = \mathcal{P}(\lambda),\, P^Y = \mathcal{P}(\mu)$ mit $\lambda, \mu \geq 0,$[32] so ist auch $X+Y$ Poisson-verteilt mit $P^{X+Y} = \mathcal{P}(\lambda+\mu).$

Beweis: Wir verwenden wahrscheinlichkeitserzeugende Funktionen, d.h.

$$\varphi_X(t) = E\left(t^X\right) = \sum_{k=0}^{\infty} t^k\, P(X=k) = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda t)^k}{k!} = e^{\lambda(t-1)},\ t \in \mathbb{R}.$$

Dann gilt

$$\varphi_{X+Y}(t) = E\left(t^{X+Y}\right) = E\left(t^X\right) E\left(t^Y\right) = e^{\lambda(t-1)} e^{\mu(t-1)} = e^{(\lambda+\mu)(t-1)},\ t \in \mathbb{R},$$

woraus durch Reihenentwicklung die Behauptung folgt. ■

Für das Folgende verwenden wir der Einfachheit halber die abkürzende Schreibweise

$$\mathfrak{p}_\lambda = f_X \text{ für eine Poisson-verteilte Zufallsvariable mit } P^X = \mathcal{P}(\lambda),\ \lambda \geq 0.$$

Satz 51 (Poisson'sche Faltungshalbgruppe)**.** Für jedes $\lambda > 0$ wird durch

$$T_\lambda(t) f := \mathfrak{p}_{\lambda t} * f,\ t \geq 0$$

auf dem Banach-Raum ℓ^1 eine Operator-Halbgruppe der Klasse (C_0) definiert. Diese besitzt den beschränkten infinitesimalen Erzeuger

$$A = \lambda(B - \mathbf{I})$$

mit dem Shift-Operator B, definiert durch[33]

$$Bf(k) = \begin{cases} f(k-1), & k \in \mathbb{N} \\ 0, & k = 0 \end{cases} \quad \text{für } f \in \ell^1.$$

Diese Halbgruppe wird auch als *Poisson'sche Faltungshalbgruppe* bezeichnet.

Beweis: Nach Lemmata 63 und 65 gilt zunächst

$$T_\lambda(s+t)f = \mathfrak{p}_{\lambda(s+t)} * f = \left(\mathfrak{p}_{\lambda s} * \mathfrak{p}_{\lambda t}\right) * f = \mathfrak{p}_{\lambda s} * \mathfrak{p}_{\lambda t} * f = \mathfrak{p}_{\lambda s} * \left(\mathfrak{p}_{\lambda t} * f\right) = T_\lambda(s)\left[T_\lambda(t) f\right],\ t \geq 0,\ f \in \ell^1.$$

Dies ist die Halbgruppen-Eigenschaft. Ferner ist

$$T_\lambda(0) f = \mathfrak{p}_0 * f = \varepsilon_0 * f = f \quad \text{für alle } f \in \ell^1.$$

Die Stetigkeitseigenschaft ergibt sich z.B. so: es ist

[32] Für $\lambda = 0$ entspricht die Poisson-Verteilung definitionsgemäß der Einpunktverteilung ε_0 im Nullpunkt. Wir werden zur Vereinfachung hier und im Folgenden die Einpunktverteilungen im Punkt $k \in \mathbb{Z}^+$ mit dem k-ten Einheitsvektor $(0, \cdots, 0, 1, 0, \cdots) \in \ell^1$ identifizieren.

[33] Dies ist gerade der (zweite) Shift-Operator S aus Beispiel 9.

$$
\begin{aligned}
\left\|T_\lambda(t)f-f\right\|_1 &= \sum_{k=0}^{\infty}\left|\sum_{j=0}^{k} e^{-\lambda t}\frac{(\lambda t)^{k-j}}{(k-j)!}f(j)-f(k)\right| = \left(1-e^{-\lambda t}\right)|f(0)| + \sum_{k=1}^{\infty}\left|\sum_{j=0}^{k} e^{-\lambda t}\frac{(\lambda t)^{k-j}}{(k-j)!}f(j)-f(k)\right| \\
&\le \left(1-e^{-\lambda t}\right)|f(0)| + \left(1-e^{-\lambda t}\right)\sum_{k=1}^{\infty}|f(k)| + \sum_{k=1}^{\infty}\left|\sum_{j=0}^{k-1} e^{-\lambda t}\frac{(\lambda t)^{k-j}}{(k-j)!}f(j)\right| \\
&\le \left(1-e^{-\lambda t}\right)\sum_{k=0}^{\infty}|f(k)| + \sum_{k=0}^{\infty}\left|\sum_{j=0}^{k} e^{-\lambda t}\frac{(\lambda t)^{k+1-j}}{(k+1-j)!}f(j)\right| \\
&\le \left(1-e^{-\lambda t}\right)\|f\|_1 + \lambda t\sum_{k=0}^{\infty}\left|\sum_{j=0}^{k} e^{-\lambda t}\frac{(\lambda t)^{k-j}}{(k-j)!}f(j)\right| \\
&\le \left(1-e^{-\lambda t}\right)\|f\|_1 + \lambda t\left\|\mathfrak{p}_{\lambda t}*f\right\|_1 \le \left(1-e^{-\lambda t}+\lambda t\right)\|f\|_1
\end{aligned}
$$

für alle $t \ge 0$ und $f \in \ell^1$, woraus $s\text{-}\lim_{t\downarrow 0}\left\|T_\lambda(t)f-f\right\|_1 = 0$ folgt.

Eine einfachere Argumentation, die zugleich den Rest der Aussage beweist, ist folgende: die Operatoren $-\lambda\mathbf{I}$ und λB sind beschränkt, also erzeugen sie jeweils eine Halbgruppe der Klasse $\left(C_0\right)$, nämlich $U_\lambda(t) = e^{-\lambda t\mathbf{I}} = e^{-\lambda t}\mathbf{I}$ und $V_\lambda(t) = e^{\lambda t B}$ für $t \ge 0$. Dann ist aber auch das Produkt $W_\lambda(t) := U_\lambda(t)\circ V_\lambda(t) = e^{-\lambda t}e^{\lambda t B}$ für $t \ge 0$ eine solche Halbgruppe, da $U_\lambda(t)$ und $V_\lambda(t)$ für alle $t \ge 0$ kommutieren. Es ist leicht zu sehen, dass gilt:

$$
\begin{aligned}
W_\lambda(t)f = e^{-\lambda t}e^{\lambda t B}f &= \sum_{k=0}^{\infty} e^{-\lambda t}\frac{(\lambda t)^k}{k!}B^k f = \sum_{k=0}^{\infty} e^{-\lambda t}\frac{(\lambda t)^k}{k!}\left(\varepsilon_k * f\right) \\
&= \left(\sum_{k=0}^{\infty} e^{-\lambda t}\frac{(\lambda t)^k}{k!}\varepsilon_k\right) * f = \mathfrak{p}_{\lambda t} * f
\end{aligned}
$$

für $t \ge 0$ und $f \in \ell^1$, so dass in der Tat $W_\lambda(t) = T_\lambda(t)$ für alle $t \ge 0$ gilt, diese beiden Halbgruppen wegen der Eindeutigkeit des Erzeugers also identisch sind. Damit ist der Satz bewiesen. ■

Die Faltungsoperation auf ℓ^1 kann auch durch eine Matrixoperation beschrieben werden. Wählen wir nämlich für $f \in \ell^1$ die (unendlich-dimensionale) quadratische Matrix

$$
\Pi_f := \begin{bmatrix}
f(0) & f(1) & f(2) & f(3) & \cdots \\
0 & f(0) & f(1) & f(2) & \cdots \\
0 & 0 & f(0) & f(1) & \cdots \\
0 & 0 & 0 & f(0) & \cdots \\
\vdots & \vdots & \vdots & \vdots & \ddots
\end{bmatrix}
$$

und fassen wir die Folgen $g \in \ell^1$ als (unendlich-dimensionale) *Zeilenvektoren* auf, so gilt in diesem Sinne auch:

$$
f * g = g * f = g \cdot \Pi_f = f \cdot \Pi_g \quad \text{für alle } f, g \in \ell^1.
$$

Die Poisson'sche Faltungshalbgruppe kann daher auch dargestellt werden als

$$T(t)f = f \cdot \Pi_{\lambda t} = f \cdot \Pi_\lambda^t,\ t \geq 0,\ f \in \ell^1$$

$$\text{mit } \Pi_\lambda = \begin{bmatrix} e^{-\lambda} & \lambda e^{-\lambda} & \frac{\lambda^2}{2} e^{-\lambda} & \cdots \\ 0 & e^{-\lambda} & \lambda e^{-\lambda} & \cdots \\ 0 & 0 & e^{-\lambda} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

Die (auch nicht-ganzzahligen) Potenzen Π_λ^t besitzen dabei die Darstellung

$$\Pi_\lambda^t = e^{tQ_\lambda} = \sum_{k=0}^{\infty} \frac{t^k}{k!} Q_\lambda^k \quad \text{für } t \geq 0$$

mit der so genannten *Fundamentalmatrix*[34]

$$Q_\lambda := \begin{bmatrix} -\lambda & \lambda & 0 & 0 & \cdots \\ 0 & -\lambda & \lambda & 0 & \cdots \\ 0 & 0 & -\lambda & \lambda & \cdots \\ 0 & 0 & 0 & -\lambda & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

Die Poisson'sche Faltungshalbgruppe spielt eine wichtige Rolle u.a. bei der Approximation von Summen unabhängiger, aber ggf. unterschiedlich binomialverteilter Zufallsvariablen durch Poisson-verteilte Zufallsvariablen. So kann im Sinne der Faltung die Binomialverteilung $\mathcal{B}(1,p)$ mit dem Faltungs-Operator $\mathbf{I} + pA$ mit dem infinitesimalen Erzeuger A aus Satz 51 (für $\lambda = 1$) identifiziert werden. Sind nun $X_1, \cdots, X_n$ stochastisch unabhängige Zufallsvariablen mit $P^{X_i} = B(1, p_i)$ mit $p_1, \cdots, p_n \in (0,1)$, so kann die Verteilung von $\sum_{i=1}^n X_i$ mit dem Faltungs-Operator $\prod_{i=1}^n (\mathbf{I} + p_i A)$ identifiziert werden, der seinerseits durch den Faltungs-Operator $\exp\left\{\left(\sum_{i=1}^n p_i\right) A\right\}$ approximiert werden kann. Eine Abschätzung des entsprechenden Abstands in der ℓ^1-Norm führt dann im Sinne von Beispiel 2 Teil a) zu einer Abschätzung des Totalvariationsabstands zwischen den involvierten Verteilungen, die z.B. in der Personenversicherungsmathematik (einjährige Wahrscheinlichkeitsverteilung für die Anzahl der Versicherungsfälle bei altersgemischten Kollektiven) eine wichtige Anwendung findet.

---

[34] Die wird in Abschnitt II. 3.3 im Rahmen der Markoff-Prozesse deutlicher, wo Π_λ^t die Rolle der *t*-Schritt-Übergangsmatrix für den homogenen Poisson-Prozess spielt.

II. 3.2. Die Brown'sche Faltungshalbgruppe

In diesem Abschnitt behandeln wir eine Faltungs-Halbgruppe, die im Zusammenhang mit stetig verteilten Zufallsvariablen steht und daher auch einen anderen Banach-Raum als ℓ^1 benötigt. Wir arbeiten hier mit dem schon bekannten Raum $UCB(\mathbb{R})$ der gleichmäßig stetigen und beschränkten Funktionen auf $\mathbb{R}$, es sind aber auch andere Räume denkbar (vgl. BUTZER / BERENS (1967) oder GOLDSTEIN (1985)).

Definition 58 (Faltung II)**.** Für $f \in L^1(\mathbb{R},\mathcal{B},\mathfrak{m})$[35] und $g \in UCB(\mathbb{R})$ oder $g \in L^1(\mathbb{R},\mathcal{B},\mathfrak{m})$ ist die *Faltung* $f * g$ als Funktion definiert durch

$$(f*g)(x) := \int_{-\infty}^{\infty} f(y)\, g(x-y)\, \mathfrak{m}(dy) \quad \text{für } x \in \mathbb{R}.$$

Lemma 66. Die Faltung im Sinne von Definition 58 besitzt folgende Eigenschaften:

$$f*g \in \begin{cases} UCB(\mathbb{R}) \\ L^1(\mathbb{R},\mathcal{B},\mathfrak{m}) \end{cases} \text{ für alle } f \in L^1(\mathbb{R},\mathcal{B},\mathfrak{m}) \text{ und } g \in \begin{cases} UCB(\mathbb{R}) \\ L^1(\mathbb{R},\mathcal{B},\mathfrak{m}) \end{cases}$$

mit

$$\|f*g\|_{UCB} \le \|f\|_1 \cdot \|g\|_{UCB},\ f \in L^1(\mathbb{R},\mathcal{B},\mathfrak{m}),\ g \in UCB(\mathbb{R}) \text{ bzw.}$$

$$\|f*g\|_1 \le \|f\|_1 \cdot \|g\|_1,\ f,g \in L^1(\mathbb{R},\mathcal{B},\mathfrak{m}).$$

Ferner ist die Faltungsoperation assoziativ in folgendem Sinn:

$$(f*g)*h = f*(g*h) \text{ für alle } f,g \in L^1(\mathbb{R},\mathcal{B},\mathfrak{m}),\ h \in \begin{cases} UCB(\mathbb{R}) \\ L^1(\mathbb{R},\mathcal{B},\mathfrak{m}) \end{cases}.$$

Beweis: Es ist im ersten Fall

$$|(f*g)(x)| \le \int_{-\infty}^{\infty} |f(y)||g(x-y)|\,\mathfrak{m}(dy) \le \|g\|_{UCB} \int_{-\infty}^{\infty} |f(y)|\,\mathfrak{m}(dy) = \|f\|_1 \|g\|_{UCB},$$

woraus $\|f*g\|_{UCB} \le \|f\|_1 \cdot \|g\|_{UCB}$, $f \in L^1(\mathbb{R},\mathcal{B},\mathfrak{m})$, $g \in UCB(\mathbb{R})$ folgt. $f*g$ ist auch gleichmäßig stetig, denn sei $\varepsilon > 0$, dann existiert ein $\delta > 0$ mit der Eigenschaft:

$$|x-y| < \delta \ \Rightarrow\ |g(x)-g(y)| < \varepsilon \quad \text{für alle } x,y \in \mathbb{R}.$$

Dies impliziert ferner:

$$|x-y| < \delta \ \Rightarrow\ |f*g(x) - f*g(y)| \le \int_{-\infty}^{\infty} |f(z)| \underbrace{|g(x-z)-g(y-z)|}_{<\varepsilon} \mathfrak{m}(dz) \le \varepsilon \int_{-\infty}^{\infty} |f(z)|\,\mathfrak{m}(dz) \le \varepsilon \|f\|_1$$

für alle $x,y \in \mathbb{R}$, d.h. die Faltung ist gleichmäßig stetig. Im zweiten Fall argumentiert man so:

[35] Hier bezeichnet wie üblich $\mathcal{B}$ die Borel'sche σ-Algebra über $\mathbb{R}$ und $\mathfrak{m}$ das zugehörige Lebesgue-Maß.

$$\int_{-\infty}^{\infty} |(f*g)(x)|\,\mathfrak{m}(dx) \le \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |f(y)||g(x-y)|\,\mathfrak{m}(dy)\,\mathfrak{m}(dx) = \int_{-\infty}^{\infty} |f(y)| \int_{-\infty}^{\infty} |g(x-y)|\,\mathfrak{m}(dx)\,\mathfrak{m}(dy)$$
$$= \int_{-\infty}^{\infty} |f(y)| \int_{-\infty}^{\infty} |g(x)|\,\mathfrak{m}(dx)\,\mathfrak{m}(dy) = \|f\|_1 \|g\|_1,$$

woraus $\|f*g\|_1 \le \|f\|_1 \cdot \|g\|_1,\ f,g \in L^1(\mathbb{R},\mathcal{B},\mathfrak{m})$ folgt.

Die Assoziativität ergibt sich ähnlich wie in Lemma 63 aus

$$\begin{aligned}
(f*g)*h(x) &= \int_{-\infty}^{\infty} (f*g)(y)\,h(x-y)\,\mathfrak{m}(dy) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(z)\,g(y-z)\,h(x-y)\,\mathfrak{m}(dy)\,\mathfrak{m}(dz)\\
&= \int_{-\infty}^{\infty} f(z) \int_{-\infty}^{\infty} g(y-z)\,h(x-y)\,\mathfrak{m}(dy)\,\mathfrak{m}(dz) = \int_{-\infty}^{\infty} f(z) \int_{-\infty}^{\infty} g(y)\,h(x-y-z)\,\mathfrak{m}(dy)\,\mathfrak{m}(dz)\\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(z)\,g(y)\,h(x-y-z)\,\mathfrak{m}(dy)\,\mathfrak{m}(dz) = \int_{-\infty}^{\infty} f(z)(g*h)(x-z)\,\mathfrak{m}(dz)\,\mathfrak{m}(dy)\\
&= f*(g*h)(x)
\end{aligned}$$

für alle $f,g \in L^1(\mathbb{R},\mathcal{B},\mathfrak{m}),\ h \in UCB(\mathbb{R}),$ weil hier einerseits $f*g \in L^1(\mathbb{R},\mathcal{B},\mathfrak{m})$ und $h \in UCB(\mathbb{R})$, andererseits $f \in L^1(\mathbb{R},\mathcal{B},\mathfrak{m})$ und $g*h \in UCB(\mathbb{R})$ ist; im Fall $f,g,h \in L^1(\mathbb{R},\mathcal{B},\mathfrak{m})$ ist die Aussage analog. Damit ist das Lemma bewiesen. ■

Auch hier wird durch die Faltung mit einem (festen) Element $f \in L^1(\mathbb{R},\mathcal{B},\mathfrak{m})$ für $\mathcal{X} = L^1(\mathbb{R},\mathcal{B},\mathfrak{m}),\ \mathcal{Y} \in \{L^1(\mathbb{R},\mathcal{B},\mathfrak{m}),\ UCB(\mathbb{R})\}$ ein linearer beschränkter Operator $T_f \in \mathfrak{L}[\mathcal{Y},\mathcal{Y}]$ definiert vermöge

$$T_f g := f*g,\ g \in \mathcal{Y} \text{ mit } \|T_f\| \le \|f\|_1.$$

Ähnlich wie in Lemma 64 kann auch hier gezeigt werden, dass $f*g$ eine Lebesgue-Dichte der Summe unabhängiger reeller Zufallsvariablen X,Y ist, wenn X die Dichte f und Y die Dichte g besitzt.

Für das Folgende verwenden wir die in der Stochastik übliche Bezeichnung

$$\varphi_{\mu,\sigma^2}(x) := \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right),\ x \in \mathbb{R}$$

für die Dichte der $\mathcal{N}(\mu,\sigma^2)$-Normalverteilung, mit $\mu \in \mathbb{R}$ und $\sigma^2 > 0.$ Klarerweise gilt:

$$\varphi_{\mu,\sigma^2} \in L^1(\mathbb{R},\mathcal{B},\mathfrak{m}) \cap UCB(\mathbb{R}).$$

Lemma 67 (Normalverteilungen)**.** Für die Dichten der Normalverteilung gilt:

$$\varphi_{\mu,\sigma^2} * \varphi_{\nu,\tau^2} = \varphi_{\mu+\nu,\sigma^2+\tau^2} \text{ für alle } \mu,\nu \in \mathbb{R} \text{ und } \sigma^2,\tau^2 > 0,$$

d.h. sind X und Y stochastisch unabhängige normalverteilte Zufallsvariablen mit $P^X = N\left(\mu,\sigma^2\right)$, $P^Y = N\left(\nu,\tau^2\right)$, so ist $X+Y$ ebenfalls normalverteilt mit $P^{X+Y} = N\left(\mu+\nu,\sigma^2+\tau^2\right)$.

Beweis: Wir verwenden momenterzeugende bzw. charakteristische Funktionen, d.h.

$$\begin{aligned}\psi_X(t) = E\left(e^{tX}\right) &= \int_{-\infty}^{\infty} e^{tx}\varphi_{\mu,\sigma^2}(x)\,dx = \frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty} e^{tx}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)dx \\ &= \frac{e^{\mu t+t^2\sigma^2/2}}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty}\exp\left(-\frac{(x-\mu-\sigma^2 t/2)^2}{2\sigma^2}\right)dx = \exp\left(\mu t+\frac{t^2\sigma^2}{2}\right),\ t\in\mathbb{R}.\end{aligned}$$

Die charakteristische Funktion (Fourier-Transformierte) ergibt sich hieraus durch Übergang von $t \in \mathbb{R}$ zum komplexen Argument $it \in \mathbb{C}$.

Damit gilt

$$\begin{aligned}\psi_{X+Y}(t) = E\left(e^{t(X+Y)}\right) = E\left(e^{tX}\right)E\left(e^{tY}\right) &= \exp\left(\mu t+\frac{t^2\sigma^2}{2}\right)\exp\left(\nu t+\frac{t^2\tau^2}{2}\right) \\ &= \exp\left((\mu+\nu)t+\frac{t^2\left(\sigma^2+\tau^2\right)}{2}\right),\ t\in\mathbb{R},\end{aligned}$$

woraus die Behauptung folgt, weil sich die Varianzen unabhängiger Zufallsvariablen addieren. ■

Satz 52 (Brown'sche Faltungshalbgruppe)**.** Für jedes Paar $\mu \in \mathbb{R}$, $\sigma^2 > 0$ wird durch

$$T_{\mu,\sigma^2}(t)f := \varphi_{-\mu t,\sigma^2 t} * f,\ t>0 \text{ und } T_{\mu,\sigma^2}(0)f := f$$

auf dem Banach-Raum $UCB(\mathbb{R})$ eine Operator-Halbgruppe der Klasse $\left(C_0\right)$ definiert. Diese besitzt den unbeschränkten infinitesimalen Erzeuger

$$Af = \mu f' + \frac{\sigma^2}{2}f'',\quad f \in \mathcal{D}(A) := \left\{f \in UCB(\mathbb{R}) \mid f'' \in UCB(\mathbb{R})\right\}.$$

Diese Halbgruppe wird auch als *Brown'sche Faltungshalbgruppe* bezeichnet.

Beweis: Die Halbgruppeneigenschaft folgt unmittelbar aus Lemma 67. Zum Nachweis der Stetigkeit zeigen wir zunächst: Ist X_t eine normalverteilte Zufallsvariable mit $P^{X_t} = \mathcal{N}\left(\mu t,\sigma^2 t\right)$, so gilt

$$T_{\mu,\sigma^2}(t)f(x) = \varphi_{-\mu t,\sigma^2 t} * f(x) = E\left(f\left(x+X_t\right)\right),\ t\geq 0,\ f\in UCB(\mathbb{R}),\ x\in\mathbb{R}.$$ [36]

[36] Hier gilt wie bei der Poisson-Verteilung per Konvention $\mathcal{N}(0,0) = \varepsilon_0$.

Es ist nämlich wegen der Symmetrie der Dichte der Normalverteilung $\mathcal{N}\left(0,\sigma^2 t\right)$ um Null:

$$E\left(f\left(x+X_t\right)\right)=\int_{-\infty}^{\infty} f(x+z)\varphi_{\mu t,\sigma^2 t}(z)\,dz=\int_{-\infty}^{\infty} f(x+\mu t+z)\varphi_{0,\sigma^2 t}(z)\,dz=\int_{-\infty}^{\infty} f(x+\mu t+z)\varphi_{0,\sigma^2 t}(-z)\,dz$$

$$=\int_{-\infty}^{\infty} f(x+\mu t-z)\varphi_{0,\sigma^2 t}(z)\,dz=\int_{-\infty}^{\infty} f(x-z)\varphi_{-\mu t,\sigma^2 t}(z)\,dz=\varphi_{-\mu t,\sigma^2 t}*f(x)$$

$$=T_{\mu,\sigma^2}(t)f(x)$$

für alle $t>0,\ f\in UCB(\mathbb{R}),\ x\in\mathbb{R}.$ Auf Grund der Definition der Operatoren gilt die Gleichung

$$T_{\mu,\sigma^2}(t)f(x)=E\left(f\left(x+X_t\right)\right)$$

auch für $t=0,\ f\in UCB(\mathbb{R}),\ x\in\mathbb{R},$ weil $X_0=0$ fast sicher ist. Sei nun $f\in UCB(\mathbb{R})$ und $\varepsilon>0.$ Dann gibt es ein $\delta>0$ mit der Eigenschaft:

$$|x-y|<\delta\ \Rightarrow\ |f(x)-f(y)|<\varepsilon\quad\text{für alle } x,y\in\mathbb{R}.$$

Für $t>0$ sei nun $C_t:=\left\{|X_t|<\delta\right\}$. Es folgt mit der Markoff-Ungleichung für beliebiges $x\in\mathbb{R}$:

$$\left|T_{\mu,\sigma^2}(t)f(x)-f(x)\right|=\left|E\left(f\left(x+X_t\right)\right)-f(x)\right|=\left|E\left(f\left(x+X_t\right)-f(x)\right)\right|\le E\left|f\left(x+X_t\right)-f(x)\right|$$

$$\le E\left(\left|f\left(x+X_t\right)-f(x)\right|\cdot 1\!\!1_{C_t}\right)+E\left(\left|f\left(x+X_t\right)-f(x)\right|\cdot 1\!\!1_{C_t^c}\right)$$

$$\le\varepsilon+2\|f\|_{UCB}\,P\left(|X_t|\ge\delta\right)\le\varepsilon+2\|f\|_{UCB}\frac{E\left(X_t^2\right)}{\delta^2}\le\varepsilon+2\|f\|_{UCB}\frac{\sigma^2 t+\mu^2 t^2}{\delta^2}<2\varepsilon$$

für genügend kleines $t>0,$ gleichmäßig in $x\in\mathbb{R}.$ Daraus folgt aber

$$\left\|T_{\mu,\sigma^2}(t)f-f\right\|_{UCB}\le 2\varepsilon\ \text{ für genügend kleines } t>0,$$

womit die Stetigkeitseigenschaft gezeigt ist.

Für die Herleitung des infinitesimalen Erzeugers sei nun $f\in\mathcal{D}(A)$ mit der obigen Definition von $\mathcal{D}(A)$. Nach dem Satz von Taylor gilt dann für $x,h\in\mathbb{R}$:

$$f(x+h)=f(x)+hf'(x)+\frac{h^2}{2}f''\left(x+\xi(x,h)\right)$$

mit einem geeigneten (messbaren) Zwischenpunkt $\xi(x,h)\in\left[-|h|,|h|\right]$. Es folgt

$$\frac{1}{t}\left(T_{\mu,\sigma^2}(t)f(x)-f(x)\right)=\frac{1}{t}E\left(f\left(x+X_t\right)-f(x)\right)=\frac{1}{t}E\left(X_t\right)f'(x)+\frac{1}{t}E\left(\frac{X_t^2}{2}f''\left(x+\xi(x,X_t)\right)\right)$$

$$=\mu f'(x)+\frac{\sigma^2+\mu^2 t}{2}f''(x)+\frac{1}{2t}E\left(X_t^2\left[f''\left(x+\xi(x,X_t)\right)-f''(x)\right]\right)$$

und somit

$$\left\| \frac{1}{t}\left(T_{\mu,\sigma^2}(t)f - f\right) - \mu f' - \frac{\sigma^2}{2} f'' \right\|_{UCB} \leq \frac{\mu^2 t}{2} \| f'' \|_{UCB} + \frac{1}{2t} \sup_{x\in\mathbb{R}} E\left(X_t^2 \left| f''\left(x + \xi(x, X_t)\right) - f''(x) \right| \right).$$

Mit einem Argument ähnlich dem zuvor und den dort verwendeten Bezeichnungen lässt sich aber zeigen, dass auf Grund der gleichmäßigen Stetigkeit von f und wegen $|\xi(x, X_t)| \leq |X_t|$ sowie der Hölder-Ungleichung (Lemma 15)

$$\begin{aligned}
E\left(X_t^2 \left| f''\left(x + \xi(x, X_t)\right) - f''(x) \right| \right) &\leq E\left(X_t^2 \left| f''\left(x + \xi(x, X_t)\right) - f''(x) \right| \mathbb{1}_{C_t} \right) + 2\| f'' \|_{UCB} E\left(X_t^2 \mathbb{1}_{C_t^c} \right) \\
&\leq t\left(\sigma^2 + \mu^2 t\right)\varepsilon + 2\| f'' \|_{UCB} \sqrt{E\left(X_t^4\right)} \sqrt{E\left(\mathbb{1}_{C_t^c}\right)} \\
&\leq t\left(\sigma^2 + \mu^2 t\right)\varepsilon + 2t\| f'' \|_{UCB} \sqrt{t^2\mu^4 + 6t\mu^2\sigma^2 + 3\sigma^4} \sqrt{\frac{E\left(X_t^2\right)}{\delta^2}} \\
&= t\left\{ \left(\sigma^2 + \mu^2 t\right)\varepsilon + 2\| f'' \|_{UCB} \sqrt{t^2\mu^4 + 6t\mu^2\sigma^2 + 3\sigma^4} \, \frac{\sqrt{\sigma^2 t + \mu^2 t^2}}{\delta} \right\} \leq 2\varepsilon t
\end{aligned}$$

für genügend kleines $t > 0$, gleichmäßig in $x \in \mathbb{R}$ ist, woraus

$$Af = s\text{-}\lim_{t\downarrow 0} \frac{1}{t}\left(T_{\mu,\sigma^2}(t)f - f\right) = \mu f' + \frac{\sigma^2}{2} f'' \text{ für } f \in \mathcal{D}(A)$$

folgt. Der Satz ist damit bewiesen. ■

Die Brown'sche Halbgruppe spielt eine zentrale Rolle in der mathematischen Physik, insbesondere der Physik der *Diffusionsprozesse* und der *Wärmeleitung*, sowie der inhaltlich verwandten *Stochastischen Finanzmathematik* (vgl. insbesondere BUTZER / BERENS (1967), GOLDSTEIN (1985) oder ETHERIDGE (2002)). Sie ist damit insbesondere geeignet, Lösungen für bestimmte partielle Differenzialgleichungen zweiter Ordnung zu bestimmen. Betrachten wir etwa die Differenzialgleichung für eine Funktion *u* von zwei Variablen (Ort *x* und Zeit *t*)

$$\frac{\partial}{\partial t} u(x,t) = \mu \frac{\partial}{\partial x} u(x,t) + \frac{\sigma^2}{2} \frac{\partial^2}{\partial x^2} u(x,t) \quad \text{mit der Anfangsbedingung } u(x,0) = f(x)$$

für eine Funktion $f \in UCB(\mathbb{R})$, mit reellen Konstanten $\mu \in \mathbb{R}$ und $\sigma^2 > 0$. Solche Differenzialgleichungen hat schon Albert Einstein 1905 in einer seiner drei berühmten Arbeiten aus jenem Jahr studiert und ihre Lösung in einer Form angegeben, die dem heutigen moderneren Standard durchaus entspricht.[37] Mit dem Ansatz

$$u(x,t) = T_{\mu,\sigma^2}(t) f(x), \; t \geq 0, \; x \in \mathbb{R}$$

kommen wir hier nämlich zum Ziel: nach Lemma 50 und Satz 52 gilt ja gerade

$$\frac{\partial}{\partial t} u(x,t) = \frac{d}{dt} T_{\mu,\sigma^2}(t) f(x) = A\left(T_{\mu,\sigma^2}(t) f\right)(x) = \mu \frac{\partial}{\partial x} u(x,t) + \frac{\sigma^2}{2} \frac{\partial^2}{\partial x^2} u(x,t),$$

mit $u(x,0) = T_{\mu,\sigma^2}(0) f(x) = f(x)$, wie gewünscht.

[37] Tatsächlich hat schon L. Bachelier im Jahr 1900 an diesem Problem gearbeitet; vgl. den entsprechenden Hinweis im Vorwort der Monographie von ITÔ / MCKEAN (1974). Seine Arbeit "Théorie de la spéculation" befasste sich übrigens als erste mit der Modellierung von Aktienkursen.

II. 3.3. Markoff'sche Halbgruppen

Die Ausführungen der vorigen beiden Abschnitte haben einen natürlichen Zusammenhang mit den so genannten *homogenen Markoff-Prozessen* der Stochastik, insbesondere dem homogenen Poisson-Prozess, der ja schon bei den stochastischen Darstellungssätzen für Operatorhalbgruppen eine wesentliche Rolle spielte, und dem Wiener-Prozess (auch Brown'sche Bewegung genannt). Wir wollen deshalb in diesem letzten Abschnitt noch kurz darlegen, wie der allgemeine Zusammenhang zwischen (homogenen) Markoff-Prozessen und den zugehörigen Halbgruppen aussieht. Für den benötigten Hintergrund in Bezug auf bedingte Wahrscheinlichkeiten bzw. ihrer abstrakten Verallgemeinerung als bedingte Erwartungen sei auf das Skript STOCHASTIK, Kapitel II.8 oder die Monographie von BREIMAN (1968) verwiesen.

Unter einem Markoff-Prozess versteht man allgemein einen Stochastischen Prozess $\{X_t | t \geq 0\}$ auf einem geeigneten Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$ mit Werten in $\mathbb{R}$, für den gilt:

$$P^{X_t}\left(\bullet \mid \sigma\{X_u, 0 \leq u \leq s\}\right) = P^{X_t}\left(\bullet \mid \sigma\{X_s\}\right) \text{ für } 0 \leq s < t \text{ fast sicher,}$$

d.h. das zukünftige stochastische Verhalten des Prozesses hängt bezüglich der Vergangenheit immer nur vom letzten bekannten Zustand und nicht von der Entwicklung davor ab. Der Prozess entscheidet sich also bezüglich seiner zukünftigen Entwicklung zu jedem Zeitpunkt „spontan", wobei die Entwicklung nur von aktuellen Zustand abhängt, unabhängig davon, wie der Prozess aus der Vergangenheit heraus in diesen Zustand gelangt ist. Der Prozess heißt darüber hinaus (zeitlich) homogen, wenn für die bedingten Verteilungen gilt:

$$P^{X_{t+h}}\left(\bullet \mid \sigma\{X_{s+h}\}\right) = P^{X_t}\left(\bullet \mid \sigma\{X_s\}\right) \text{ für alle } 0 \leq s < t \text{ und } h \geq 0.$$

Insbesondere gilt damit

$$P^{X_t}\left(\bullet \mid \sigma\{X_s\}\right) = P^{X_{t-s}}\left(\bullet \mid \sigma\{X_0\}\right).$$

Unter gewissen Regularitätsbedingungen an die Übergangswahrscheinlichkeiten des Prozesses und der Existenz rechtsstetiger Pfade des Prozesses erhält man dann eine typische Halbgruppe der Klasse (C_0) durch die Setzung[38]

$$T(t)f(x) := E\left[f(X_t) | X_0 = x\right] = \int f(y) P^{X_t}(dy | X_0 = x) \quad \text{für } t \geq 0,\, x \in \mathbb{R} \text{ und } f \in UCB(\mathbb{R}).$$

Die Halbgruppeneigenschaft ergibt sich dabei so: es ist

$$\begin{aligned} T(s)[T(t)f](x) &= \int T(t)f(y) P^{X_s}(dy | X_0 = x) = \int\int f(z) P^{X_t}(dz | X_0 = y) P^{X_s}(dy | X_0 = x) \\ &= \int\int f(z) P^{X_{t+s}}(dz | X_s = y) P^{X_s}(dy | X_0 = x) \\ &= \int\int f(z) P^{X_{t+s}}(dz | X_s = y, X_0 = x) P^{X_s}(dy | X_0 = x) \\ &= \int\int f(z) P^{(X_{t+s}, X_s)}(dz, dy | X_0 = x) = \int f(z) P^{X_{t+s}}(dz | X_0 = x) = T(s+t)f(x) \end{aligned}$$

für $t \geq 0,\, x \in \mathbb{R}$ und $f \in UCB(\mathbb{R})$ (zur Umrechnung der Integrale vgl. etwa das Skript zur STOCHASTIK, Satz 48). Die gleichmäßige Stetigkeit von f und die

[38] Nach dem Faktorisierungssatz für bedingte Erwartungen (die bedingte Wahrscheinlichkeiten als Spezialfall enthalten) unter den von Zufallsvariablen X erzeugten σ-Algebren $\sigma\{X\}$ ist stets $E\left(\bullet | \sigma\{X\}\right) = g(X)$ für eine geeignete messbare Abbildung g, der so genannten *Faktorisierungsfunktion*. Unter $E\left(\bullet | X = x\right)$ wird dann gerade $g(x)$ selbst verstanden.

Rechtsstetigkeit der Pfade liefern dazu die (C_0)-Eigenschaft, denn es gilt trivialerweise

$$T(0)f(x) = E\left[f\left(X_0\right)|X_0 = x\right] = f(x),\ x \in \mathbb{R}.$$

Beispiel 33 (homogener Poisson-Prozess). In der Standard-Form ist der homogene Poisson-Prozess ein Markoff-Prozess mit stetiger Zeit und dem abzählbaren Zustandsraum $\mathcal{S} = \mathbb{Z}^+$. Die Übergangswahrscheinlichkeiten sind gegeben durch die stochastischen Matrizen (vgl. Seite 146 oben)

$$\Pi_\lambda^t = \begin{bmatrix} e^{-\lambda t} & \lambda t\, e^{-\lambda t} & \dfrac{(\lambda t)^2}{2} e^{-\lambda t} & \cdots \\ 0 & e^{-\lambda t} & \lambda t e^{-\lambda t} & \cdots \\ 0 & 0 & e^{-\lambda t} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} = \left[\pi_{ij}(t;\lambda)\right]_{i,j\in\mathbb{Z}^+} \text{ mit } \pi_{ij}(t;\lambda) = P\left(X_t = j | X_0 = i\right),\ i,j \in \mathbb{Z}^+.$$

Eine Realisierung des Prozesses erhält man dadurch, dass man in einem Anfangszustand $i \in \mathbb{Z}^+$ beginnt und jeweils nach einer mit Parameter $\lambda > 0$ verteilten „Wartezeit" vom aktuellen Zustand $k \in \mathbb{Z}^+$ in den Zustand $k+1 \in \mathbb{Z}^+$ wechselt, und so fort. Die Wartezeitenfolge ist dabei stochastisch unabhängig.

Fasst man hier eine Folge $f \in \ell^\infty$ als (unendlich-dimensionalen) *Spaltenvektor* auf, so ergibt sich die diskrete Version der oben eingeführten Prozess-Halbgruppe gerade zu

$$T_\lambda(t)f(x) := E\left[f\left(X_t\right)|X_0 = x\right] = \left(\Pi_\lambda^t \cdot f\right)(x) \text{ für } x \in \mathbb{Z}^+.$$

Der infinitesimale Erzeuger ist hier wieder gegeben durch die Fundamentalmatrix

$$Q_\lambda := \begin{bmatrix} -\lambda & \lambda & 0 & 0 & \cdots \\ 0 & -\lambda & \lambda & 0 & \cdots \\ 0 & 0 & -\lambda & \lambda & \cdots \\ 0 & 0 & 0 & -\lambda & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

vermöge $Af = Q_\lambda f,\ f \in \ell^\infty$, oder in expliziter Form: $Af(x) = \lambda\left[f(x+1) - f(x)\right],\ x \in \mathbb{Z}^+.$ Man beachte, dass die stochastischen Matrizen Π_λ^t also zwei unterschiedliche Halbgruppen induzieren, nämlich zum einen die im vorigen Abschnitt vorgestellte *Faltungshalbgruppe* (bei *Links*multiplikation mit einem *Zeilen*vektor aus ℓ^1), zum anderen die hier besprochene *Markoff'sche Prozesshalbgruppe* (bei *Rechts*multiplikation mit einem *Spalten*vektor aus ℓ^∞)![39]

Eine Verallgemeinerung des homogenen Poisson-Prozesses auf den Zustandsraum $\mathbb{R}$ erhält man nun leicht, wenn die Sprunghöhen der Größe 1 beibehalten werden, man aber erlaubt, dass der Prozess in einem beliebigen reellen Zustand $x \in \mathbb{R}$ startet. Dann gilt analog

$$T_\lambda(t)f(x) := E\left[f\left(X_t\right)|X_0 = x\right] = E\left[f\left(x + Y_t\right)\right] \text{ für } t \geq 0,\ x \in \mathbb{R},\ f \in UCB(\mathbb{R}),$$

wobei hier $\left\{Y_t | t \geq 0\right\}$ einen Standard-Poisson-Prozess mit Zustandsraum $\mathcal{S} = \mathbb{Z}^+$ und Anfangszustand Null bezeichne. Man erhält hier durch direkte Rechnung

[39] Es handelt sich hierbei um so genannte *duale Halbgruppen*; vgl. BUTZER / BERENS (1967), Abschnitt 1.4.

$$T_\lambda(t)f(x) = E\left[f\left(x+Y_t\right)\right] = e^{-\lambda t}\sum_{k=0}^{\infty} f(x+k)\frac{(\lambda t)^k}{k!},\quad t\geq 0,\ x\in\mathbb{R},\ f\in UCB(\mathbb{R}),$$

woraus sogleich auch

$$\begin{aligned}\frac{1}{t}\left[T_\lambda(t)f(x)-f(x)\right] &= -\frac{1-e^{-\lambda t}}{t}f(x)+\lambda e^{-\lambda t}f(x+1)+e^{-\lambda t}\sum_{k=2}^{\infty}f(x+k)\frac{\lambda^k t^{k-1}}{k!}\\ &= \lambda e^{-\lambda t}f(x+1)-\frac{1-e^{-\lambda t}}{t}f(x)+O\left(t\right),\end{aligned}$$

gleichmäßig in $x\in\mathbb{R}$, folgt, was wieder zu dem (beschränkten) infinitesimalen Erzeuger (Differenzenoperator)

$$Af(x) = \lambda\left[f(x+1)-f(x)\right],\ x\in\mathbb{R}\quad \text{für alle } f\in UCB(\mathbb{R})$$

führt.

Beispiel 34 (homogener Wiener-Prozess). In der Grundform ist der homogene Wiener-Prozess $\left\{X_t \mid t\geq 0\right\}$ ein Stochastischer Prozess mit unabhängigen, stationären, normalverteilten Zuwächsen mit Varianz σ^2 für die Zeitperiode 1, der im Zustand Null startet. Seine Kovarianzen sind damit durch $Kov\left(X_s,X_t\right)=\sigma^2\min(s,t)$ für $s,t\geq 0$ gegeben. Es ist also insbesondere $P^{X_t}=\mathcal{N}\left(0,\sigma^2 t\right)$. Man kann zeigen, dass es Versionen des Wiener-Prozesses gibt, die stetige Pfade besitzten, was aber keine triviale Forderung ist, da dies z.B. schon nicht mit der Messbarkeit bezüglich der (sogar überabzählbaren) Produkt-σ-Algebra vereinbar ist. Eine besonders schöne explizite Konstruktion des Wiener-Prozesses geht auf Reihenentwicklungen bezüglich bestimmter Orthogonal-Systeme zurück, siehe etwa BREIMAN (1968), Proposition 12.14 und ITÔ / MCKEAN (1974), Section 1.5, Problem 3. In leicht abgewandelter Form sieht diese Konstruktion über dem Zeitintervall $\left[0,T\right]$ mit $T>0$ so aus:

Es sei $\left\{Y_n\right\}_{n\in\mathbb{Z}^+}$ eine unabhängige Folge $\mathcal{N}(0,1)$-verteilter Zufallsvariablen. Dann wird durch

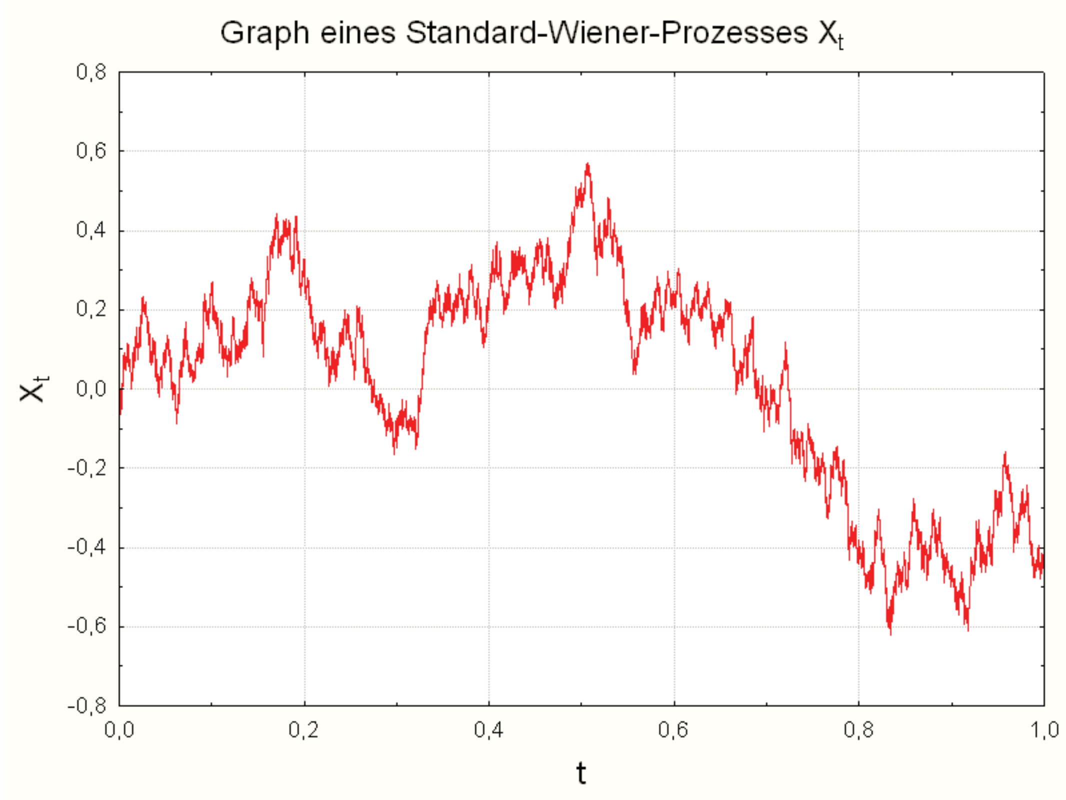
$$W_t := \frac{t}{\sqrt{T}}Y_0+\frac{\sqrt{2T}}{\pi}\sum_{n=1}^{\infty}\frac{1}{n}\sin\left(\frac{n\pi t}{T}\right)Y_n,\ 0\leq t\leq T$$

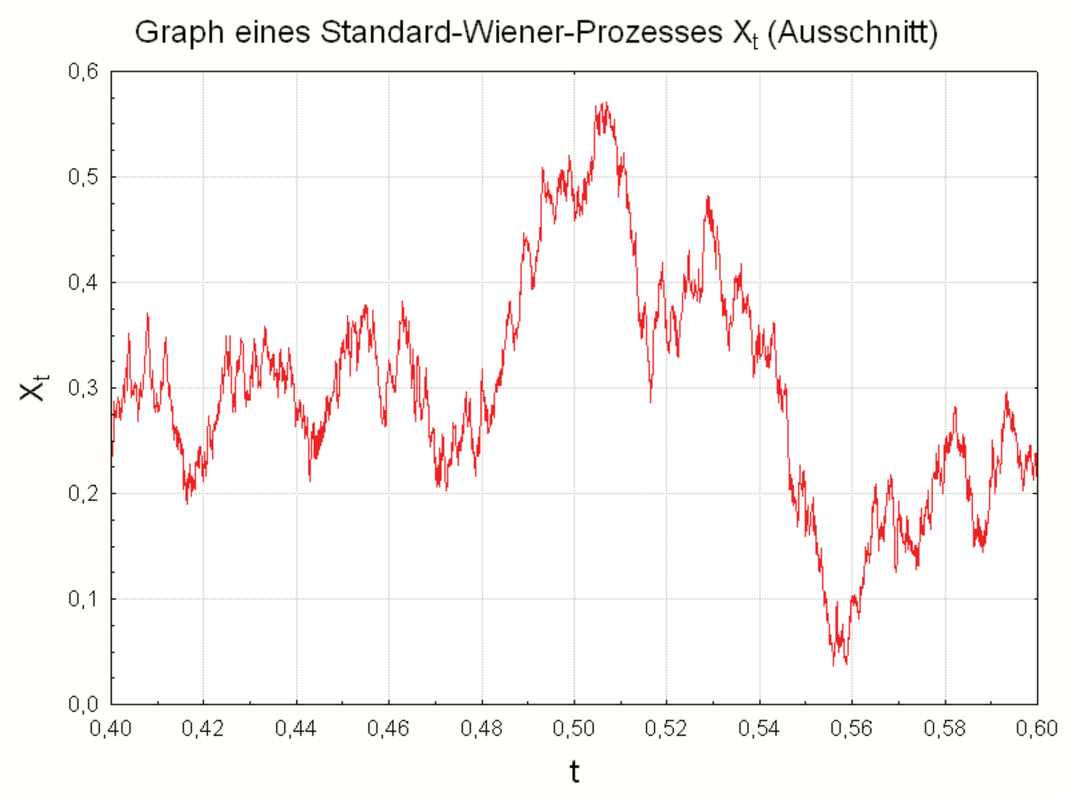
ein Standard-Wiener-Prozess (d.h. mit $\sigma^2=1$) definiert. Einen Wiener-Prozess mit beliebiger Varianz $\sigma^2>0$ und beliebigem Anfangszustand $x\in\mathbb{R}$ erhält man hieraus z.B. durch die Transformation $X_t=x+\sigma W_t,\ t\geq 0$. Durch weitere Addition des Terms μt erhält man den so genannten Wiener-Prozess mit *Drift* $\mu\in\mathbb{R}$.

Die zugehörige kanonische Halbgruppe hat – inklusive Driftterm - die Form

$$T_{\mu,\sigma^2}(t)f(x) := E\left[f\left(X_t\right)\mid X_0=x\right] = E\left[f\left(x+Z_t\right)\right]\ \text{für } t\geq 0,\ x\in\mathbb{R},\ f\in UCB(\mathbb{R}),$$

wobei $\left\{Z_t \mid t\geq 0\right\}$ einen Wiener-Prozess mit Drift bezeichne, der im Nullpunkt startet. Damit ergibt sich exakt die Halbgruppe aus Satz 52, weil hier $P^{Z_t}=\mathcal{N}\left(\mu t,\sigma^2 t\right)$ für $t>0$ gilt.


Graph eines Standard-Wiener-Prozesses X_t
0,8
0,6
0,4
0,2
0,0
-0,2
-0,4
-0,6
-0,8
X_t
0,0
0,2
0,4
0,6
0,8
1,0
t


Graph eines Standard-Wiener-Prozesses X_t (Ausschnitt)
0,6
0,5
0,4
0,3
0,2
0,1
0,0
X_t
0,40
0,42
0,44
0,46
0,48
0,50
0,52
0,54
0,56
0,58
0,60
t

Verzeichnis der Definitionen, Beispiele, Sätze und Lemmata

Definition	Beispiel	Satz	Lemma	Seite
	15			61
33				62
34				62
			24	62
35				63
			25	63
		25		63
		26		65
		27		66
			26	66
		28		66
		29		67
36				68
	16			68
		30		69
			27	69
			28	69
			29	70
			30	71
			31	73
			32	73
		31		74
	17			77
		32		78
37				79
		33		79
		34		80
38				80
			33	81
	18			82
39				82
			34	82
			35	83
		35		85
		36		86
40				89
			36	89
41				89
			37	89
42				90
		37		90
43				91
			38	91
	19			91
44				92
			39	92

Definition	Beispiel	Satz	Lemma	Seite
	20			92
45				93
		38		93
	21			94
46				94
			40	95
			41	95
			42	96
			43	96
		39		96
			44	99
		40		100
		41		100
47				101
			45	101
	22			103
	23			104
	24			104
	25			104
48				105
49				106
	26			107
			46	108
			47	109
50				109
			48	110
			49	113
51				114
			50	114
52				116
		42		117
	27			120
	28			120
	29			120
		43		121
	30			124
	31			124
53				125
		44		125
		45		127
			51	128
	32			129
54				129
			52	130
		46		131
			53	131

Definition	Beispiel	Satz	Lemma	Seite
			54	132
			55	132
			56	133
			57	133
55				134
			58	134
56				135
			59	135
			60	136
		47		137
		48		137
		49		138
			61	139
			62	140
		50		141
57				142
			63	142
			64	143
			65	144
		51		144
58				147
			66	147
			67	149
		52		149
	33			150
	34			154

Literatur

[1] D. APPLEBAUM (2004): *Lévy Processes and Stochastic Calculus.* Cambridge University Press, Cambridge.

[2] G. BACHMAN AND L. NARICI (2000): *Functional Analysis*. Dover Edition, Dover Publications, Mineola, N.Y.

[3] H. BAUER (2002): *Wahrscheinlichkeitstheorie*. 5. Aufl., Walter de Gruyter, Berlin.

[4] H. BAUER (1992): *Maß- und Integrationstheorie*. 2. Aufl., Walter de Gruyter, Berlin.

[5] P. BILLINGSLEY (1986): *Probability and Measure.* 2nd ed., Wiley, N.Y.

[6] P.L. BUTZER AND H. BERENS (1967): *Semi-Groups of Operators and Approximation*. Springer, N.Y.

[7] L. BREIMAN (1968): *Probability*. Addison-Wesley, Reading, Massachusetts.

[8] E. ÇINLAR (1975): *Introduction to Stochastic Processes*. Prentice-Hall, Englewood Cliffs.

[9] J. DUGUNDJI (1966): *Topology*. Allyn and Bacon, Boston.

[10] J.A. GOLDSTEIN (1985): *Semigroups of Linear Operators and Applications*. Oxford University Press, N.Y.

[11] J. ELSTRODT (1996): *Maß- und Integrationstheorie.* Springer, Berlin.

[12] A. ETHERIDGE (2002): *A Course in Financial Calculus*. Cambridge University Press, Cambridge.

[13] E. HEWITT AND K. STROMBERG (1965): *Real and Abstract Analysis*. Springer, N.Y.

[14] K. ITÔ AND H.P. MCKEAN (1974): *Diffusion Processes and their Sample Paths*. Springer, N.Y.

[15] I. KARATZAS AND S.E. SHREVE (1991): *Brownian Motion and Stochastic Calculus*. Springer, N.Y.

[16] S. KARLIN AND H.M. TAYLOR (1981): *A Second Course in Stochastic Processes*. Academic Press, N.Y.

[17] J.L. KELLEY AND T.P SRINIVASAN (1988): *Measure and Integral*. Vol 1. Springer, N.Y.

[18] D. LAMBERTON AND B. LAPEYRE (1997): *Stochastic Calculus Applied to Finance*. Chapman & Hall, London.

[19] D.G. LUENBERGER (1968): *Optimization by Vector Space Methods*. Wiley, N.Y.

[20] P. MALLIAVIN (1997): *Stochastic Analysis*. Springer, N.Y.

[21] M. MEYER (2001): *Continuous Stochastic Calculus with Application to Finance*. Chapman & Hall, Boca Raton.

[22] J. MIKUSINSKI (1978): *The Bochner Integral*. Birkhäuser, Basel.

[23] D. PFEIFER (1984): *Probabilistic representations of operator semigroups – a unifying approach*. Semigroup Forum 30, 17 – 34.

[24] D. POLLARD (1984): *Convergence of Stochastic Processes*. Springer, N.Y.

[25] S.M. ROSS (1996): *Stochastic Processes*. Second Edition. Wiley, N.Y.

[26] H.L. ROYDEN (1988): *Real Analysis*. 3rd ed., Prentice-Hall, New Jersey.

[27] W. RUDIN (1991): *Functional Analysis*. McGraw-Hill, N.Y.

[28] W. RUDIN (1987): *Real & Complex Analysis*. McGraw-Hill, N.Y.

[29] J.M. STEELE (2001): *Stochastic Calculus and Financial Applications*. Springer, N.Y.

[30] A. E. TAYLOR (1963): *Introduction to Functional Analysis*. 3rd ed., Wiley, N.Y.

[31] D. WERNER (2004): *Funktionalanalysis*. 5. Auflage, Springer, Berlin.

[32] J. WLOKA (1971): *Funktionalanalysis und Anwendungen*. De Gruyter, Berlin.

Danksagung

Herr Dr. Peter Harmand hat das vorliegende Skript freundlicherweise gründlich durchgesehen und an vielen Stellen konstruktive Verbesserungsvorschläge gemacht. Herr M.Sc. Feras Karakit hat mich auf einen Fehler in einer älteren Version des Skripts aufmerksam gemacht. Dafür sei beiden an dieser Stelle herzlich gedankt.