Inhaltsverzeichnis

1 Zufallsvariablen in Banachräumen

In diesem ersten Kapitel befassen wir uns zunächst mit einigen grundlegenden Eigenschaften Borelscher σ-Algebren und darauf definierten Wahrscheinlichkeitsmaßen. Da die algebraische Struktur eines Banachraums hierfür zunächst nicht von Bedeutung ist lassen sich viele Ergebnisse im allgemeineren Kontext von vollständigen metrischen Räumen zeigen. **TODO**

1.1 Borelmengen in metrischen Räumen

Für einen metrischen Raum (X,d) bezeichne im Folgenden $\mathcal{B}(X)$ die Borel-σ-algebra in X. Zudem werden für $x \in X$ und $r > 0$ mit $B(x,r)$ bzw. $\overline{B}(x,r)$ die offene bzw. abgeschlossene Kugel um x mit Radius r bezeichnet.

1.1 Proposition
Sei (X,d) ein separabler metrischer Raum. Dann gilt

$$\mathcal{B}(X) = \sigma(\{B(x,r) : x \in X, r > 0\}) = \sigma(\{\overline{B}(x,r) : x \in X, r > 0\}).$$

Beweis.
Setze

$$\mathcal{A}_1 := \sigma(\{B(x,r) : x \in X, r > 0\}),$$
$$\mathcal{A}_2 := \sigma(\{\overline{B}(x,r) : x \in X, r > 0\}).$$

Nach Definition gilt $\mathcal{A}_2 = \mathcal{A}_1 \subseteq \mathcal{B}(X)$. Zu zeigen ist also nur die Inklusion $\mathcal{B}(X) \subseteq \mathcal{A}_1$. Sei dazu $U \subseteq X$ offen und $x \in U$. Nach Voraussetzung existiert eine abzählbare dichte Teilmenge $D \subseteq X$. Definiere

$$R := \{(y,r) : y \in U \cap D, r > 0, r \in \mathbb{Q}, B(y,r) \subseteq U\}.$$

Dann ist R abzählbar und da D dicht in X liegt gilt $U = \bigcup_{(y,r)\in R} B(y,r)$. Also gilt $U \in \mathcal{A}_1$ und da $\mathcal{B}(X)$ von den offenen Teilmengen von X erzeugt wird folgt die Behauptung. □

1.2 Proposition
Für $i \in \mathbb{N}$ sei (X_i, d_i) ein separabler metrischer Raum. Dann gilt

$$\mathcal{B}(X_1 \times X_2 \times ...) = \otimes_{i=1}^{\infty} \mathcal{B}(X_i)$$

Beweis.
Setze $X := \times_{k\in\mathbb{N}} X_k$ und bezeichne $p_k : X \to X_k$ die Projektion auf die k-te Komponente. Betrachte das Mengensystem

$$\mathcal{E} : = \{ \bigcap_{k \in K} p_k^{-1}(O) | \forall k \in K : O_k \subseteq X_k \text{ offen}, K \subseteq \mathbb{N} \text{ endlich}\}.$$

Offensichtlich gilt $\otimes_{k\in\mathbb{N}} \mathcal{B}(X_k) = \sigma(\mathcal{E})$. Ferner ist X ein separabler metrischer Raum und $\mathcal{E}$ eine Basis der Produkttopologie auf X, vgl. [7][3.7]. Also lässt sich jede offene Menge $O \subset X$ als abzählbare Vereinigung von Elementen aus $\mathcal{E}$ darstellen. Dies impliziert $\mathcal{B}(X) = \sigma(\mathcal{E}) = \otimes_{k\in N} \mathcal{B}(X_k)$. □

1.2 Borelmaße auf metrischen Räumen

Notation und Konventionen
Bis auf weiteres sei (X,d) ein metrischer Raum mit Borel-σ-algebra $\mathcal{B}(X)$. Für eine Teilmenge $A \subseteq X$ sei $\mathring{A}$ das *Innere der Menge* und $\overline{A}$ der *Abschluss*. Ferner bezeichne

$$\partial A := \overline{A} \setminus \mathring{A}$$

den *Rand der Menge*.
Im Folgenden Abschnitt beschäftigen wir uns mit Wahrscheinlichkeitsmaßen auf $\mathcal{B}(X)$, welche teilweise auch als *Borel-Maße* bezeichnet werden. Die Bezeichnung wird in der Literatur allerdings nicht einheitlich verwendet.

1.3 Definition
Ein Maß μ auf $\mathcal{B}(X)$ heißt *regulär* , falls

$$\begin{aligned} \forall B \in \mathcal{B}(X): \quad \mu(B) &= \sup\{\mu(C) : C \subseteq B,\ C \text{ abgeschlossen}\} \\ &= \inf\{\mu(O) : B \subseteq O,\ O \text{ offen}\}. \end{aligned}$$

1.4 Proposition
Sei μ ein Wahrscheinlichkeitsmaß auf $\mathcal{B}(X)$. Dann ist μ regulär.

Beweis.
Wir verwenden zum Beweis das Good-Set-Principle. Setze dazu

$$\mathcal{R} := \{A \in \mathcal{B}(X) : \mu(A) = \sup\{\mu(C) : C \subseteq A,\ C \text{ abgeschlossen}\} = \inf\{\mu(O) : A \subseteq O,\ O \text{ offen}\}\}.$$

Wir zeigen zunächst, dass $\mathcal{R}$ eine σ-Algebra ist. Offensichtlich gilt $\emptyset \in \mathcal{R}$. Sei nun $A \in \mathcal{R}$ und $\varepsilon > 0$. Dann existieren eine offene Menge O und eine abgeschlossene Menge C mit $C \subseteq A \subseteq O$ und

$$\mu(0) - \varepsilon < \mu(A) < \mu(C) + \varepsilon.$$

Es gilt also $O^c \subseteq A^c \subseteq C^c$ und

$$\mu(C^c) - \varepsilon = 1 - \mu(C) - \varepsilon < 1 - \mu(A) = \mu(A^c) = 1 - \mu(A) < 1 - \mu(O) + \varepsilon = \mu(O^c) + \varepsilon.$$

Da O^c abgeschlossen ist und C^c offen ist, folgt $A^c \in \mathcal{R}$.
Seien nun $A_1, A_2, ... \in \mathcal{R}$ und $\varepsilon > 0$ Für jedes $n \in \mathbb{N}$ existieren dann eine offene Menge O_n und eine abgeschlossene Menge C_n mit $C_n \subseteq A_n \subseteq O_n$ und

$$\mu(O_n) - 2^{-n}\varepsilon < \mu(A_n) < \mu(C_n) + 2^{n+1}\varepsilon.$$

Es gilt also

$$\bigcup_{n=1}^{\infty} C_n \subseteq \bigcup_{n=1}^{\infty} A_n \subseteq \bigcup_{n=1}^{\infty} C_n$$

und

$$\begin{aligned} \mu\left(\bigcup_{n=1}^{\infty} U_n\right) - \mu\left(\bigcup_{n=1}^{\infty} A_n\right) &\leq \mu\left(\bigcup_{n=1}^{\infty} U_n \setminus \bigcup_{n=1}^{\infty} A_n\right) \\ &\leq \mu\left(\bigcup_{n=1}^{\infty} (U_n \setminus A_n)\right) \leq \sum_{i=1}^{\infty} \mu(U_n \setminus A_n) < \sum_{i=1}^{\infty} 2^n\varepsilon = \varepsilon. \end{aligned} \tag{1.1}$$

Zudem gilt

$$\mu(\cup_{n=1}^{\infty} C_n) = \lim_{k\to\infty} \mu(\cup_{n=1}^{k} C_n),$$

also existiert ein $k \in \mathbb{N}$ mit

$$\mu(\cup_{n=1}^{\infty} C_n) - \mu(\cup_{n=1}^{k} C_n) < \frac{\varepsilon}{2}.$$

Die Menge $C := \cup_{n=1}^{k} C_n$ ist als endliche Vereinigung abgeschlossener Mengen selbst abgeschlossen und offensichtlich in $\cup_{n=1}^{\infty} A_n$ enthalten. Ferner gilt

$$\begin{aligned}
\mu(\bigcup_{n=1}^{\infty} A_n) - \mu(C) &< \mu(\bigcup_{n=1}^{\infty} A_n) - \mu(\bigcup_{n=1}^{\infty} C_n) + \frac{\varepsilon}{2} \\
&\leq \mu(\bigcup_{n=1}^{\infty} A_n \setminus \bigcup_{n=1}^{\infty} C_n) + \frac{\varepsilon}{2} \\
&\leq \mu(\bigcup_{n=1}^{\infty} (A_n \setminus C_n)) + \frac{\varepsilon}{2} \\
&\leq \sum_{n=1}^{\infty} \mu(A_n \setminus C_n) + \frac{\varepsilon}{2} = \varepsilon. \qquad (1.2)
\end{aligned}$$

Da ε beliebig gewählt war folgt nun aus (1.1) und (1.2), dass $\cup_{n=1}^{\infty} A_n \in \mathcal{R}$. Also ist $\mathcal{R}$ eine σ-Algebra. Wir zeigen nun, dass $\mathcal{R}$ alle abgeschlossenen Mengen enthält. Sei also $\emptyset \neq A \subseteq X$ abgeschlossen. Die Bedingung

$$\mu(A) = \sup\{\mu(C) : C \subseteq A, C \text{ abgeschlossen } \}$$

folgt direkt aus der Monotonie von μ. Um die zweite Bedingung zu zeigen setze für $n \in \mathbb{N}$

$$O_n := \{x \in X : \inf_{y \in A} d(x,y) < \frac{1}{n}\}.$$

Dann ist O_n für jedes $n \in \mathbb{N}$ offen und es gilt $O_1 \supseteq O_2 \supseteq ...$, sowie $\cup_{n=1}^{\infty} O_n = A$, da A abgeschlossen ist. Mit der σ-Stetigkeit von μ folgt letztendlich

$$\mu(A) \leq \inf\{\mu(O) : A \subseteq O, O \text{ offen}\} \leq \inf_{n \in \mathbb{N}} \mu(O_n) = \lim_{n \to \infty} \mu(O_n) = \mu(A).$$

Da die abgeschlossenen Mengen $\mathcal{B}(X)$ erzeugen folgt also $\mathcal{R} = \mathcal{B}(X)$ und folglich ist μ regulär. □

1.5 Definition
Ein Maß μ auf $\mathcal{B}(X)$ heißt *straff*, falls es für alle $\varepsilon > 0$ eine kompakte Menge $K \subseteq X$ gibt mit

$$\mu(K) \geq 1 - \varepsilon.$$

1.6 Korollar
Sei μ ein straffes Wahrscheinlichkeitsmaß auf $\mathcal{B}(X)$. Dann gilt

$$\forall A \in \mathcal{B}(X) : \quad \mu(A) = \sup\{\mu(K) : K \subseteq A,\ K \text{ kompakt}\}.$$

Beweis.
Sei $A \in \mathcal{B}(X)$ und $\varepsilon > 0$. Wegen der Straffheit von μ existiert eine kompakte Menge $K_\varepsilon \subseteq X$ mit $\mu(K_\varepsilon) \geq 1 - \frac{\varepsilon}{2}$, und da μ nach Proposition 1.4 regulär ist gibt es eine abgeschlossene Menge $C \subseteq A$ mit $\mu(C) > \mu(A) - \frac{\varepsilon}{2}$. Dann ist die Menge $K_\varepsilon \cap C$ wiederum kompakt und es gilt

$$\mu(A) \geq \mu(K_\varepsilon \cap C) > \mu(C) - \frac{\varepsilon}{2} > \mu(A) - \varepsilon.$$

□

1.7 Bemerkung
Ein Wahrscheinlichkeitsmaß μ auf $\mathcal{B}(X)$ mit der Eigenschaft

$$\forall A \in \mathcal{B}(X) : \quad \mu(A) = \sup\{\mu(K) : K \subseteq A,\ K \text{ kompakt}\}.$$

wird auch als *Radon-Wahrscheinlichkeitsmaß* oder *Radon-Maß* bezeichnet.

1.8 Proposition
Sei (X, d) ein vollständiger separabler metrischer Raum. Dann ist jedes Wahrscheinlichkeitsmaß μ auf $\mathcal{B}(X)$ straff.

Wir verwenden zum Beweis der Proposition die folgende Charakterisierung kompakter Teilmengen metrischer Räume. Der Beweis wird mittels der in metrischen Räumen geltenden Äquivalenz von Überdeckungskompaktheit und Folgenkompaktheit geführt und findet sich etwa in [2][Theorem 3.10].

1.9 Lemma
Sei (X, d) ein metrischer Raum. Eine Menge $K \subseteq X$ ist genau dann kompakt, wenn sie die folgenden beiden Eigenschaften erfüllt:

(i) K ist vollständig,

(ii) K ist total-beschränkt, d.h.

$$\forall \varepsilon > 0 \; \exists x_1, ..., x_n \in K : \; K \subseteq \cup_{i=1}^n B(x_i, \varepsilon).$$

Beweis.
Sei $\varepsilon > 0$. Nach Voraussetzung existiert eine abzählbare dichte Teilmenge $D = \{x_1, x_2, ...\}$ von X. Also gilt insbesondere für $q \in \mathbb{N}$

$$\bigcup_{i \in \mathbb{N}} \overline{B}(x_i, 2^{-q}) = X.$$

Wegen der σ-Stetigkeit von μ existiert also ein $N_q \in \mathbb{N}$ mit

$$\mu(\cup_{i=1}^{N_q} \overline{B}(x_i, 2^{-q}) \geq 1 - \varepsilon 2^{-q}.$$

Setze nun

$$K := \bigcap_{q=1}^{\infty} \bigcup_{i=1}^{N_q} \overline{B}(x_i, 2^{-q}).$$

Dann ist K als Schnitt abgeschlossener Teilmengen abgeschlossen, und da X vollständig ist, folgt daraus bereits die Vollständigkeit von K. Ferner ist K total-beschränkt, denn zu $\varepsilon > 0$ existiert ein $q \in \mathbb{N}$ mit $2^{-q} < \varepsilon$ und $K \subseteq \cup_{i=1}^{N_q} B(x_i, 2^{-q}) \subseteq \cup_{i=1}^{N_q} B(x_i, \varepsilon)$. Zudem gilt

$$\begin{aligned} \mu(K) = 1 - \mu(\cup_{q \in \mathbb{N}} \cap_{i=1}^{N_q} \overline{B}(x_i, 2^{-q})^c) &\geq 1 - \sum_{q=1}^{\infty} \mu(\cap_{i=1}^{N_q} \overline{B}(x_i, 2^{-q})^c) \\ &\geq 1 - \sum_{q=1}^{\infty} \varepsilon 2^{-q} = 1 - \varepsilon. \end{aligned}$$

Also ist μ straff. □

1.10 Proposition
Sei (X, d) ein vollständiger metrischer Raum und μ ein Wahrscheinlichkeitsmaß auf $\mathcal{B}(X)$. Dann sind äquivalent

(i) μ ist straff.

(ii) Es gibt eine separable Teilmenge $E \subseteq X$ mit $\mu(E) = 1$.

Beweis.
zu (i) $\Rightarrow$ (ii): Für alle $n \in \mathbb{N}$ existiert $K_n \subseteq X$ kompakt mit $\mu(K_n) \geq 1 - \frac{1}{n}$, o.E. gelte $K_n \subseteq K_{n+1}$. Es folgt

$$\mu\big(\cup_{n=1}^{\infty} K_n\big) = \lim_{n \to \infty} \mu\big(K_{n+1}\big) = 1.$$

Da kompakte Teilmengen metrischer Räume insbesondere separabel sind, ist $E := \cup_{n=1}^{\infty} K_n$ als abzählbare Vereinigung separabler Mengen ebenso separabel.
zu (ii) $\Rightarrow$ (i): Analog zum Beweis von Proposition 1.8. □

Insgesamt haben wir also gezeigt

1.11 Satz
Für ein Wahrscheinlichkeitsmaß μ auf der Borelschen σ-Algebra eines vollständigen metrischen Raumes (X, d) sind äquivalent

(i) μ ist straff,

(ii) Es gibt eine separable Menge $E \subseteq X$ mit $\mu(E) = 1$,

(iii) μ ist ein Radon-Maß, d.h.

$$\forall A \in \mathcal{B}(X): \quad \mu(A) = \sup\{\mu(K) : K \subseteq A, \ K \text{ kompakt}\}.$$

Nachdem wir uns nun ausgiebig mit den Eigenschaften einzelner Wahrscheinlichkeitsmaße beschäftigt haben, möchten wir uns nun mit Folgen von Wahrscheinlichkeitsmaßen und deren Konvergenz beschäftigen.

1.12 Definition
Eine Folge $(\mu_n)_{n\in\mathbb{N}}$ von Wahrscheinlichkeitsmaßen auf $\mathcal{B}(X)$ heißt *schwach konvergent* gegen ein Wahrscheinlichkeitsmaß μ auf $\mathcal{B}(X)$, falls

$$\forall f \in C_b(X): \quad \lim_{n\to\infty} \int_X f d\mu_n = \int_X f d\mu.$$

Bezeichnung: $\mu_n \rightharpoonup \mu$.

Als nützliches Hilfsmittel für viele Beweise dient der folgende Satz, der meist als *Portmanteau-Theorem* bezeichnet wird.

1.13 Satz (Portmanteau-Theorem)
Sei (X, d) ein metrischer Raum und $\mu, \mu_1, \mu_2, ...$ Wahrscheinlichkeitsmaße auf $\mathcal{B}(X)$. Dann sind äquivalent

(i) $(\mu_n)_{n\in\mathbb{N}}$ konvergiert schwach gegen μ,

(ii) Für alle abgeschlossenen Teilmengen $A \subseteq X$ gilt

$$\limsup_{n\to\infty} \mu_n(A) \leq \mu(A)$$

(iii) Für alle offenen Teilmengen $B \subseteq X$ gilt

$$\liminf_{n\to\infty} \mu_n(A) \geq \mu(A)$$

(iv) Für alle Borelmengen $C \in \mathcal{B}(X)$ mit $\mu(\partial C) = 0$ gilt

$$\lim_{n\to\infty} \mu_n(C) = \mu(C)$$

Beweis.
zu $(i) \Rightarrow (ii)$: Sei $A \subseteq X$ abgeschlossen und $\varepsilon > 0$. Da die Aussage für $A = \emptyset$ trivialerweise erfüllt ist können wir ohne Beschränkung der Allgemeinheit annehmen, dass $A \neq \emptyset$. Für $m \in \mathbb{N}$ setze

$$U_m := \{x \in X : d(x, A) < \frac{1}{m}\}.$$

Dann sind die Mengen U_m offen und es für alle $m \in \mathbb{N}$ gilt $A \subseteq U_m$. Da A abgeschlossen ist erhalten wir zudem $A = \cap_{m\in\mathbb{N}} U_m$. Aufgrund der σ-Stetigkeit von μ existiert ein $k \in \mathbb{N}$ mit

$$\mu(U_k) < \mu(A) + \varepsilon.$$

Betrachte nun die Abbildung

$$f : X \to \mathbb{R}, \quad x \mapsto \max\{1 - kd(x, A), 0\}.$$

Offensichtlich ist f beschränkt und nach der umgekehrten Dreiecksungleichung auch stetig. Wegen $1_A \leq f \leq 1_{U_k}$ erhalten wir zusammen mit der Voraussetzung

$$\limsup_{n\to\infty} \mu_n(A) \leq \lim_{n\to\infty} \int_X f d\mu_n = \int_X f d\mu \leq \mu(U_k) \leq \mu(A) + \varepsilon.$$

Da diese Ungleichung für alle $\varepsilon > 0$ erfüllt ist folgt die Behauptung.
zu $(ii) \iff (iii)$: Folgt unmittelbar durch Komplementbildung.
zu $(iii) \Rightarrow (iv)$: Sei $C \in \mathcal{B}(X)$ mit $\mu(\partial C) = 0$. Dann gilt insbesondere $\mu(\overline{C}) = \mu(\mathring{C}))$. Da (iii) auch (ii) impliziert erhalten wir somit

$$\mu(C) = \mu(int(C)) \leq \liminf_{n\to\infty} \mu_n(C) \leq \limsup_{n\to\infty} \mu_n(C) \leq \mu(\overline{C}) = \mu(C).$$

zu $(iv) \Rightarrow (i)$: Sei $f \in C_b(X)$ beschränkt durch $M > 0$. Wegen der Linearität des Integrals können wir ohne Einschränkung annehmen, dass $f \geq 0$. Wegen der Stetigkeit von f erhalten wir zunächst für alle $t > 0$

$$\partial\{f > t\} \subseteq \{f = t\}.$$

Da μ ein Wahrscheinlichkeitsmaß ist, gibt es eine abzählbare Menge $C \subseteq \mathbb{R}$ mit

$$\forall t \in \mathbb{R} \setminus C: \quad \mu(\{f = t\}) = 0.$$

Also gilt für alle $t \in \mathbb{R} \setminus C$ nach Voraussetzung

$$\lim_{n\to\infty} \mu_n(\{f > t\}) = \mu(\{f > t\})$$

Mit dem Satz von Cavalieri, vgl. [4][1.8.20], erhalten wir schließlich per dominierter Konvergenz

$$\lim_{n\to\infty} \int_X f d\mu_n = \lim_{n\to\infty} \int_0^M \mu_n(\{f > t\}) d\lambda(t) = \int_0^M \mu(\{f > t\}) d\lambda(t) = \int_X f d\mu.$$

□

1.3 Die Prokhorov Metrik

1.14 Definition
Eine Teilmenge A eines metrischen Raums (X, d) heißt *relativ kompakt*, falls ihr Abschluss $\overline{A}$ kompakt ist.

Mittels der Teilfolgencharakterisierung von kompakten Teilmengen metrischer Räume erhält man unmittelbar

1.15 Proposition
Eine Teilmenge A eines metrischen Raums (X, d) ist genau dann relativ kompakt, wenn jede Folge $(x_n)_{n\in\mathbb{N}}$ in A eine in $\overline{A}$ konvergente Teilfolge enthält.

Eine Folge $(x_n)_{n\in\mathbb{N}}$ in einem metrischen Raum konvergiert also genau dann, wenn die beiden folgenden Bedingungen erfüllt sind

(i) Die Menge $\{x_n : n \in \mathbb{N}\}$ ist relativ kompakt und

(ii) alle konvergenten Teilfolgen von $(x_n)_{n\in\mathbb{N}}$ haben den selben Grenzwert.

Ziel dieses Abschnitts ist es daher zunächst eine adäquate Metrik für die schwache Konvergenz von Wahrscheinlichkeitsmaßen zu konstruieren und anschließend den Begriff der relativen Kompaktheit bzgl. dieser Metrik zu charakterisieren.

Für einen metrischen Raum (X, d) bezeichne $\mathcal{M}(X)$ die Menge aller Wahrscheinlichkeitsmaße auf der Borelschen σ-Algebra $\mathcal{B}(X)$ und für eine Menge $B \in \mathcal{B}(X)$ definiere für $\varepsilon > 0$

$$B^\varepsilon = \begin{cases} \{x \in X : \inf_{y \in B} d(x,y) < \varepsilon\}, \text{ falls } B \neq \emptyset, \\ \emptyset, \text{ sonst.} \end{cases}$$

Betrachte nun die Abbildung

$$\rho : \mathcal{M}(X) \times \mathcal{M}(X) \to [0, \infty),$$

definiert durch

$$(\mu, \nu) \mapsto \inf\{\alpha > 0 | \forall B \in \mathcal{B}(X) : \mu(B) \leq \nu(B^\alpha) + \alpha \ \wedge \ \nu(B) \leq \mu(B^\alpha) + \alpha\}.$$

1.16 Proposition

(i) Die Abbildung ρ ist wohldefiniert und eine Metrik auf $\mathcal{M}(X)$.

(ii) Für $\alpha > 0$ und $\mu, \nu \in \mathcal{M}(X)$ gilt

$$(\forall B \in \mathcal{B}(X) : \ \mu(B) \leq \nu(B^\alpha) + \alpha) \Rightarrow (\forall B \in \mathcal{B}(X) : \ \nu(B) \leq \mu(B^\alpha) + \alpha).$$

(iii) Für $\mu, \mu_1, \mu_2, ... \in \mathcal{M}(X)$ gilt

$$\lim_{n \to \infty} \rho(\mu_n, \mu) = 0 \ \Rightarrow \ \mu_n \rightharpoonup \mu$$

Beweis.
zu (i): Die Menge ist nicht leer, da die Bedingung für alle $\alpha \geq 1$ erfüllt ist. Offensichtlich gilt ferner $\rho(\mu, \nu) \geq 0$ und $\rho(\mu, \nu) = \rho(\nu, \mu)$. Wegen $B \subseteq B^\alpha$ für alle $B \in \mathcal{B}(X)$ und $\alpha > 0$ gilt

$$\forall B \in \mathcal{B} \ \forall \alpha > 0 : \quad \mu(A) \leq \mu(A^\alpha) + \alpha.$$

Also gilt $\rho(\mu, \mu) = 0$. Gilt andererseits $\rho(\mu, \nu) = 0$, so existiert eine Folge $(\alpha_n)_{n \in \mathbb{N}}$ mit

$$\forall B \in \mathcal{B} \ \forall n \in \mathbb{N} : \quad \mu(A) \leq \nu(A^{\alpha_n}) + \alpha_n \quad \wedge \nu(A) \leq \mu(A^{\alpha_n}) + \alpha_n.$$

Insbesondere gilt für alle abgeschlossenen Mengen $A \subseteq X$, dass $A^{\alpha_n} \downarrow A$ und daher

$$\mu(A) = \lim_{n \to \infty} \mu(A^{\alpha_n}) + \alpha_n \geq \nu(\alpha) = \lim_{n \to \infty} \nu(A^{\alpha_n}) + \alpha_n \geq \mu(\alpha).$$

Da die abgeschlossenen Mengen ein schnittstabiler Erzeuger von $\mathcal{B}(X)$ sind folgt somit $\mu = \nu$.
zur Dreiecksungleichung: Seien $\mu, \nu, \eta \in \mathcal{M}(X)$ und $\alpha, \beta > 0$ mit

$$\begin{aligned} \forall B \in \mathcal{B}(X) : \quad & \mu(A) \leq \eta(A^\alpha) + \alpha \quad \wedge \quad \eta(A) \leq \mu(A^\alpha) + \alpha, \\ & \nu(A) \leq \eta(A^\beta) + \beta \quad \wedge \quad \eta(A) \leq \nu(A^\beta) + \beta. \end{aligned}$$

Dann gilt für $A \in \mathcal{B}(X)$

$$\begin{aligned} \mu(A) &\leq \eta(A^\alpha) + \alpha \leq \nu((A^\alpha)^\beta) + \alpha + \beta, \\ \nu(A) &\leq \eta(A^\beta) + \beta \leq \mu((A^\alpha)^\beta) + \alpha + \beta. \end{aligned}$$

Nach der Dreiecksungleichung für d gilt $(A^\alpha)^\beta \subseteq A^{\alpha+\beta}$ und $(A^\beta)^\alpha \subseteq A^{\alpha+\beta}$. Also ergibt sich mit obigem $\rho(\mu, \nu) \leq \alpha + \beta$. Bilden der Infima auf der rechten Seite liefert

$$\rho(\mu, \nu) \leq \rho(\mu, \eta) + \rho(\eta, \nu).$$

zu (ii): Sei $\alpha > 0$ mit

$$\forall B \in \mathcal{B}(X) : \ \mu(B) \leq \nu(B^\alpha) + \alpha.$$

Für zwei Mengen $A, B \in \mathcal{B}(X)$ gilt nach Definition $A \subseteq (B^\alpha)^c$ genau dann, wenn $B \subseteq (A^\alpha)^c$. Zu $B \in \mathcal{B}(X)$ setze also $A := (A^\alpha)^c$ und erhalte

$$\mu(B^\alpha) = 1 - \mu(A) \geq 1 - \nu(A^\alpha) - \alpha = \nu((A^\alpha)^c) - \alpha \geq \nu(B) - \alpha.$$

zu (iii): Wegen $\rho(\mu_n, \mu) \to 0$ existiert eine Folge $(\alpha_n)_{n\in\mathbb{N}}$ mit $\alpha_n \downarrow 0$ und

$$\forall n \in \mathbb{N}\ \forall B \in \mathcal{B}(X): \quad \mu_n(B) \leq \mu(B^{\alpha_n}) + \alpha_n.$$

Sei nun $A \subseteq X$ abgeschlossen. Dann gilt $A^{\alpha_n} \downarrow A$ und folglich

$$\limsup_{n\to\infty} \mu_n(A) \leq \limsup_{n\to\infty} \mu(A^{\alpha_n}) + \alpha_n = \lim_{n\to\infty} \mu(A^{\alpha_n}) = \mu(A).$$

Nach dem Portmanteau-Theorem gilt also $\mu_n \rightharpoonup \mu$. □

Falls der zugrunde liegende metrische Raum separabel ist, so gilt auch die Umkehrung von (ii). Zum Beweis dieser Tatsache benötigen wir zunächst das folgende technische Lemma.

1.17 Lemma
Sei X ein separabler metrischer Raum und μ ein Wahrscheinlichkeitsmaß auf $\mathcal{B}(X)$. Dann existiert für alle $\delta > 0$ eine Folge $(x_n)_{n\in\mathbb{N}}$ in X und eine Folge $(r_n)_{n\in\mathbb{N}}$ in $(0, \delta)$ mit

$$X = \bigcap_{n=1}^{\infty} B(x_n, r_n) \text{ und } \mu(\partial B(x_n, r_n)) = 0 \text{ für alle } n \in \mathbb{N}.$$

Beweis.
Sei $D = \{x_1, x_2, ...\} \subseteq X$ eine abzählbare dichte Teilmenge von X und $x_n \in D$. Setze

$$S(x_n, r) = \{y \in X : d(x_n, y) = r\}.$$

Dann gilt $\partial B(x_n, r) \subseteq S(x_n, r)$ und für gegebenes $\delta > 0$ ist die Mengenfamilie

$$\mathcal{S}_n := \{S(x_n, r) : \frac{\delta}{2} < r < \delta\}$$

disjunkt und überabzählbar. Da μ ein Wahrscheinlichkeitsmaß ist enthält $\mathcal{S}_n$ aber höchstens abzählbar viele Mengen S mit $\mu(S) > 0$. Also existiert insbesondere ein $r_n \in (\frac{\delta}{2}, \delta)$ mit $\mu(S(x_n, r_n)) = 0$. Da D dicht in X ist gilt

$$X \subseteq \bigcup_{n\in\mathbb{N}} B(x_n, r_n)$$

und folglich erfüllen $(x_n)_{n\in\mathbb{N}}$ und $(r_n)_{n\in\mathbb{N}}$ die geforderten Bedingungen. □

1.18 Satz
Sei (X, d) ein separabler metrischer Raum. Dann gilt für $\mu, \mu_1, \mu_2, ... \in \mathcal{M}(X)$

$$\lim_{n\to\infty} \rho(\mu_n, \mu) = 0 \iff \mu_n \rightharpoonup \mu$$

Beweis.
zu $\Rightarrow$: Siehe Proposition.
zu $\Leftarrow$: Wir zeigen

$$\forall \varepsilon > 0\ \exists N \in \mathbb{N}\ \forall n \geq N: \quad \forall B \in \mathcal{B}: \ \mu(B) \leq \mu_n(B^{\varepsilon}) + \varepsilon.$$

Daraus folgt nach Proposition $XXX.(ii)$ die Behauptung. Sei $\delta := \frac{\varepsilon}{4} > 0$. Dann existiert nach vorherigem Lemma eine Folge $(B_n)_{n\in\mathbb{N}}) := (B(x_n, r_n))_{n\in\mathbb{N}}$ offener Kugeln mit Radius $r_n < \frac{\delta}{2}$ und $\mu(\partial B_n) = 0$ für alle $n \in \mathbb{N}$, sowie $X = \cup_{n\in\mathbb{N}} B_n$. Ferner existiert $k \in N$ mit

$$\mu\left(\bigcup_{i=1}^{k} B_i\right) \geq 1 - \delta.$$

Betrachte nun das endliche Mengensystem

$$\mathcal{C} := \{\cup_{j\in J} B_j : J \subseteq \{1, .., k\}\}.$$

Dann gilt für alle $A \in \mathcal{C}$
$$\partial A \subseteq \bigcup_{i=1}^{k} \partial B_i,$$
also $\mu(\partial A) = 0$. Nach dem Portmanteau-Theorem gilt also $\mu_n(A) \to \mu(A)$ für alle $A \in \mathcal{C}$. Da $\mathcal{C}$ endlich ist existiert daher $N \in \mathbb{N}$ mit
$$\forall n \geq N\ \forall A \in \mathcal{A}: \quad |\mu(A) - \mu_n(A)| < \delta. \tag{1.3}$$
Insbesondere gilt also
$$\forall n \geq N: \quad \mu_n\left(\bigcup_{i=1}^{k} B_i\right) \geq \mu\left(\bigcup_{i=1}^{k} B_i\right) - \delta \geq 1 - 2\delta.$$
Zu $B \in \mathcal{B}$ betrachte die Menge
$$A := \bigcup_{j \in J} B_j,$$
wobei $J := \{j \in \{1, ..k\} : B_j \cap B \neq 0\}$. Dann ergibt sich unter Verwendung von $A \subseteq B^\delta \subseteq B^\varepsilon$ und (1.1) für alle $n \geq N$
$$\mu(B) \leq \mu(A) + \mu\left((\cup_{i=1}^{k} B_i)^c\right) \leq \mu(A) + 2\delta \leq \mu_n + 3\delta \leq \mu_n(B^\varepsilon) + \varepsilon.$$
□

Nachdem wir nun eine passende Metrik gefunden haben, um die schwache Konvergenz von Wahrscheinlichkeitsmaßen in $\mathcal{M}(X)$ zu beschreiben wollen wir nun noch kurz auf ein Resultat von Prokhorov eingehen. Dazu benötigen wir die folgende

1.19 Definition
Eine Familie $M \subseteq \mathcal{M}(X)$ von Wahrscheinlichkeitsmaßen heißt *gleichmäßig straff*, falls es für alle $\varepsilon > 0$ eine kompakte Menge $K \subseteq X$ gibt mit
$$\forall \mu \in M : \mu(K) \geq 1 - \varepsilon.$$
Mittels des Begriffs der gleichmäßigen Straffheit lässt sich die relative Kompaktheit von Teilmengen von $\mathcal{M}(X)$ für einen vollständigen und separablen metrischen Raum X wie folgt charakterisieren. Ein Beweis findet sich etwa in [6][Theorem 6.7].

1.20 Satz (Satz von Prokhorov)
Sei (X, d) ein vollständiger separabler metrischer Raum und $M \subseteq \mathcal{M}(X)$. Dann sind äquivalent

(i) M ist relativ kompakt,

(ii) M ist gleichmäßig straff.

Um die schwache Konvergenz einer Folge $(\mu_n)_{n\in\mathbb{N}}$ von Wahrscheinlichkeitsmaßen zu zeigen ist es in einem vollständigen und separablen metrischen Raum also ausreichend die folgenden beiden Bedingungen zu prüfen

(a) $\{\mu_n : n \in \mathbb{N}\}$ ist gleichmäßig straff und

(b) jede konvergente Teilfolge von $(\mu_n)_{n\in\mathbb{N}}$ konvergiert gegen den selben Grenzwert.

1.4 Flache Konzentrierung von Wahrscheinlichkeitsmaßen

Da wir uns im Laufe der Arbeit hauptsächlich mit Zufallsvariablen auf Banachräumen beschäftigen werden, wollen wir noch eine weitere für diesen Kontext nützliche Charakterisierung der relativen Kompaktheit betrachten. Quelle: [1], Originalpaper: de Acosta **(TODO: zitieren)** Für eine nichtleere Teilmenge $A \subseteq E$ setze
$$A^\varepsilon := \{x \in E : \inf_{y \in A} ||x - y|| < \varepsilon\}.$$

Ferner sei daran erinnert, dass jeder endlich dimensionale Untervektorraum $S \subseteq E$ abgeschlossen ist und eine Menge $A \subseteq S$ genau dann kompakt ist, wenn sie abgeschlossen und beschränkt ist. Insbesondere besitzt jede beschränkte Folge in S eine konvergente Teilfolge.

1.21 Definition
Eine Familie $M \subseteq \mathcal{M}(E)$ von Wahrscheinlichkeitsmaßen auf $\mathcal{B}(E)$ heißt *flach konzentriert*, falls es für alle $\varepsilon > 0$ einen endlichdimensionalen Untervektorraum $S \subseteq E$ gibt mit

$$\forall \mu \in M : \quad \mu(S^\varepsilon) \geq 1 - \varepsilon.$$

1.22 Lemma
Eine Teilmenge A von E ist genau dann relativ kompakt, wenn A beschränkt ist und es für alle $\varepsilon > 0$ einen endlichdimensionalen Untervektorraum $S \subseteq E$ gibt mit

$$A \subseteq S^\varepsilon$$

Beweis.
zu $\Rightarrow$: Sei $A \subseteq E$ relativ kompakt. Dann ist $\overline{A}$ kompakt und folglich beschränkt, woraus wir direkt die Beschränktheit von A erhalten. Ferner ist $\overline{A}$ separabel, also existiert eine abzählbare dichte Teilmenge $\{x_1, x_2, ...\} \subseteq A$. Für $\varepsilon > 0$ ist daher $(B(x_n, \varepsilon))_{n\in\mathbb{N}}$ eine offene Überdeckung von $\overline{A}$. Wegen der Kompaktheit existiert $I \subseteq \mathbb{N}$ endlich mit

$$\overline{A} \subseteq \bigcup_{i\in I} B(x_i, \varepsilon).$$

Sei also S der von $\{x_i : i \in I\}$ erzeugte endlich dimensionale Untervektorraum von E. Dann gilt

$$A \subseteq \overline{A} \subseteq \bigcup_{i\in I} B(x_i, \varepsilon) \subseteq S^\varepsilon.$$

zu $\Leftarrow$: Wir zeigen, dass jede Folge in A eine konvergente Teilfolge besitzt, der Grenzwert der Teilfolge muss hierbei nicht in A liegen. Sei dazu $(x_n^{(0)})_{n\in\mathbb{N}}$ eine Folge in A, $\varepsilon > 0$ und $S \subseteq E$ ein endlichdimensionaler Untervektorraum mit $A \subseteq S^\varepsilon$. Dann existieren insbesondere eine Folge $(y_n)_{n\in\mathbb{N}}$ in S mit

$$\forall n \in \mathbb{N} : \quad d(x_n^{(0)}, y_n) \leq 2\varepsilon.$$

Aus der Beschränktheit von $(x_n^{(0)})_{n\in\mathbb{N}}$ erhalten wir direkt die Beschränktheit von $(y_n)_{n\in\mathbb{N}}$ und da S endlichdimensional ist existiert eine konvergente Teilfolge $(y_{n_k})_{k\in\mathbb{N}}$. Es gilt also für $k, m \geq N(\varepsilon) \in \mathbb{N}$

$$||x_{n_k}^{(0)} - x_{n_m}^{(0)}|| \leq ||x_{n_k}^{(0)} - y_{n_k}|| + ||y_{n_k} - y_{n_m}|| + ||x_{n_m}^{(0)} - y_{n_m}|| \leq 5\varepsilon.$$

Ohne Einschränkung können wir durch entfernen endlich vieler Folgenglieder annehmen, dass

$$\forall k, m \in \mathbb{N} : \quad ||x_{n_k}^{(0)} - x_{n_m}^{(0)}|| \leq 5\varepsilon.$$

Durch obiges Verfahren können wir für $N \in \mathbb{N}$ und $\varepsilon_N = \frac{1}{N}$ induktiv eine Teilfolge $(x_n^{(N)})_{n\in\mathbb{N}}$ von $(x_n^{(N-1)})_{n\in\mathbb{N}}$ gewinnen mit

$$\forall m, n \in \mathbb{N} : \quad ||x_n^{(N)} - x_m^{(N)}|| \leq \frac{5}{N}.$$

Durch bilden der Diagonalfolge $(x_N^{(N)})_{N\in\mathbb{N}}$ erhalten wir somit eine Teilfolge der Ausgangsfolge $(x_n^{(0)})_{n\in\mathbb{N}}$ die eine Cauchy-Folge ist und daher in E konvergiert. □

1.23 Lemma
Sei $S \subseteq E$ ein endlichdimensionaler Untervektorraum und $l_1, ..., l_n \in E'$ Funktionale mit

$$\forall x, y \in S \ \exists k \in \{1, ..., n\} : \quad l_k(x) \neq l_k(y). \tag{1.4}$$

Dann ist die Menge

$$B := S^\varepsilon \cap \{x \in E : |l_1(x)| \leq r_1, ..., |l_n(x)| \leq r_n\}$$

für alle $\varepsilon, r_1, ..., r_n \in (0, \infty)$ beschränkt.

Beweis.
Wegen (1.4) definiert

$$p(x) := \max_{1 \leq k \leq n} |l_k(x)|, \quad x \in S,$$

eine Norm auf S. Da S endlichdimensional ist, ist diese insbesondere äquivalent zur Einschränkung von $||\cdot||$ auf S. Angenommen die Menge B ist nicht beschränkt. Dann existiert eine Folge $(x_n)_{n\in\mathbb{N}}$ in B mit

$$\lim_{n\to\infty} ||x_n|| = \infty.$$

Wegen der Normäquivalenz existiert dann ein $k \in \{1, ..., n\}$ mit

$$\lim_{n\to\infty} |l_k(x_n)| = \infty.$$

Im Widerspruch zur Definition von B. □

1.24 Satz
Sei $\Gamma \subseteq E'$, sodass

$$\forall x, y \in E \; \exists l \in \Gamma : \quad l(x) \neq l(y). \tag{1.5}$$

Eine Menge $M \subseteq \mathcal{M}(E)$ von Wahrscheinlichkeitsmaßen ist genau dann relativ kompakt in $(\mathcal{M}, \rho)$ wenn die folgenden beiden Bedingungen erfüllt sind

(a) Für alle $l \in \Gamma$ ist $\{\mu^l : \mu \in M\} \subseteq \mathcal{M}(\mathbb{R})$ relativ kompakt,

(b) M ist flach konzentriert.

Beweis.
zu $\Rightarrow$: Sei $M \subseteq \mathcal{M}(E)$ relativ kompakt. Nach dem Satz von Prokhorov ist M dann insbesondere gleichmäßig straff. Also existiert zu $\varepsilon > 0$ eine kompakte Menge $K \subseteq E$ mit

$$\forall \mu \in M : \quad \mu(K) \geq 1 - \varepsilon.$$

Aus der Stetigkeit von $l \in \Gamma$ erhalten wir somit direkt die gleichmäßige Straffheit von $\{\mu^l : \mu \in \Gamma\}$. Erneutes anwenden des Satzes von Prokhorov liefert (a). Da K insbesondere relativ kompakt ist liefert **Lemma zwei davor, TODO** einen endlichdimensionalen Untervektorraum $S \subseteq E$ mit $K \subseteq S^\varepsilon$. Somit gilt

$$\forall \mu \in M : \quad \mu(S^\varepsilon) \geq \mu(K) \geq 1 - \varepsilon.$$

Folglich ist M flach konzentriert.
zu $\Leftarrow$: **TODO**

1.25 Korollar
Eine Menge $M \subseteq \mathcal{M}(E)$ ist genau dann relativ kompakt, wenn M flach konzentriert ist und für alle $\varepsilon > 0$ eine beschränkte Menge $L \subseteq E$ existiert mit

$$\forall \mu \in M : \quad \mu(L) \geq 1 - \varepsilon.$$

Beweis.
Die Richtung $\Rightarrow$ ist klar und für $\Leftarrow$ reicht es anzumerken, dass beschränkte Mengen durch lineare Abbildungen auf beschränkte Mengen abgebildet werden. Nach dem Satz von Heine-Borel sind beschränkte Teilmengen endlich dimensionaler normierter Vektorräume aber insbesondere relativ kompakt. □

1.26 Bemerkung
Man sagt eine Familie von Abbildungen mit der Eigenschaft (1.2) (bzw. (1.1)) *trenne die Punkte von* E (bzw S). Die Existenz einer solchen Menge $\Gamma \subseteq E'$ ist durch den Satz von Hahn-Banach sichergestellt. Insbesondere erfüllt E' selbst(1.1).

1.5 Meßbare Vektorräume

1.27 Definition
Sei X ein Vektorraum und $\mathcal{C}$ eine σ-Algebra auf X. Das Tupel $(X, \mathcal{C})$ heißt *messbarer Vektorraum*, falls

(a) Die Abbildung

$$+ : X \times X \to X, \quad (x, y) \mapsto x + y$$

ist $\mathcal{A} \otimes \mathcal{C}/\mathcal{C}$-messbar, und

(b) die Abbildung

$$\cdot : \mathbb{R} \times X \to X, \quad (\alpha, x) \mapsto \alpha x$$

ist $\mathcal{B}(\mathbb{R}) \otimes \mathcal{C}/\mathcal{C}$-messbar.

1.28 Bemerkung
Sei $(X, \mathcal{C})$ ein messbarer Vektorraum. Dann gilt

(i) Für alle $\alpha \in \mathbb{R}$ ist die Abbildung

$$f_\alpha : X \to X, x \mapsto \alpha x$$

$\mathcal{C}/\mathcal{C}$-messbar.

(ii) Für alle $y \in X$ ist die Abbildung

$$g_y : X \to X, x \mapsto x + y$$

$\mathcal{C}/\mathcal{C}$-messbar.

Da die Komposition messbarer Abbildungen wiederum messbar ist erhält man unmittelbar

1.29 Proposition
Sei $(X, \mathcal{C})$ ein messbarer Vektorraum und $(\Omega, \mathcal{A})$ ein messbarer Raum. Sind $X, Y : \Omega \to X$ zwei $\mathcal{A}/\mathcal{C}$-messbare Abbildungen und $\alpha, \beta \in \mathbb{R}$, so ist auch $\alpha X + \beta Y$ $\mathcal{A}/\mathcal{C}$-messbar.

Beweis.
Man beachte, dass für beliebige messbare Räume $(\Omega_1, \mathcal{A}_1), (\Omega_2, \mathcal{A}_2)$, Mengen $A \in \mathcal{A}_1 \otimes \mathcal{A}_2$ und $\omega_1 \in \Omega_1$

$$A(\omega_1) = \{\omega_2 : (\omega_1, \omega_2) \in A\} \in \mathcal{A}_2$$

gilt. □

1.30 Proposition
Sei X ein separabler Banachraum. Dann ist $(X, \mathcal{B}(X))$ ein messbarer Vektorraum.

Beweis.
Nach Proposition 1.2 gilt $\mathcal{B}(X \times X) = \mathcal{B}(X) \otimes \mathcal{B}(X)$ und $\mathcal{B}(\mathbb{R} \times X) = \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(X)$. Ferner sind die Abbildungen

$$\begin{aligned} + : X \times X \to X, &\quad (x, y) \mapsto x + y, \\ \cdot : \mathbb{R} \times X \to X, &\quad (\alpha, x) \mapsto \alpha x \end{aligned}$$

stetig bzgl. der jeweiligen Produkttopologien und somit insbesondere $\mathcal{B}(X \times X)/\mathcal{B}(X)$- bzw. $\mathcal{B}(\mathbb{R} \times X)/\mathcal{B}(X)$-messbar. □

1.31 Beispiel
Für $d \in \mathbb{N}$ ist $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ ein messbarer Vektorraum.

Im Folgenden sei $(X, ||\cdot||)$ ein Banachraum und $(X', ||\cdot||_{op})$ der zugehörige Dualraum.

1.32 Proposition
Sei $\emptyset \neq \Gamma \subseteq X'$. Dann ist $(X, \sigma(\Gamma))$ ein messbarer Vektorraum.

Beweis.
TODO

1.33 Proposition
Sei X ein separabler Banachraum. Dann gilt $\sigma(X') = \mathcal{B}(X)$.

1.6 Zufallsvariablen mit Werten in Banachräumen

Sei $(\Omega, \mathcal{A}, P)$ ein vollständiger Wahrscheinlichkeitsraum und E ein Banachraum mit Norm $||\cdot||$.

1.34 Definition
Eine Abbildung $X : \Omega \to E$ heißt *(Radon-)Zufallsvariable* falls

(a) X ist $\mathcal{A}/\mathcal{B}(E)$-messbar und

(b) es existiert ein separabler Untervektorraum $E_0 \subseteq E$ mit $P(\{X \in E_0\}) = 1$.

1.35 Bemerkung
Nach Abschnitt 1.2 sind für eine $\mathcal{A}/\mathcal{B}(X)$-messbare Abbildung $X : \Omega \to E$ äquivalent

(i) es existiert ein separabler Untervektorraum $E_0 \subseteq E$ mit $P(\{X \in E_0\}) = 1$,

(ii) P^X ist ein Radon-Maß auf $\mathcal{B}(X)$,

(iii) P^X ist straff.

Hiermit erklärt sich auch die Bezeichnung Radon-Zufallsvariable.

TODO: Einführung einfache Funktionen im Banach-Kontext

1.36 Proposition
TODO: Summen von Radon-Variablen sind messbar, Approximation durch einfache Funktionen, Grenzwerte sind messbar

Bezeichne $\mathcal{L}_0(\Omega, \mathcal{A}, P; E)$ den Raum der E-wertigen Radon-Zufallsvariablen auf $(\Omega, \mathcal{A}, P)$ und sei $L_0(\Omega, \mathcal{A}, P; E)$ der Raum der Äquivalenzklassen bezüglich fast sicherer Gleichheit. Falls klar ist welcher Wahrscheinlichkeitsraum gemeint ist, so schreiben wir auch $\mathcal{L}_0(E)$ oder $L_0(E)$.

2 Konvergenzarten

2.1 Fast sichere Konvergenz

TODO: überarbeiten, Benennungen einheitlich. Ggf. Konvention $(X_n)_n$ ZVen nach-vorne ziehen

2.1 Definition
Seien $X, X_1, X_2, ...$ E-wertige Zufallsvariablen. Die Folge $(X_n)_{n\in\mathbb{N}}$ *konvergiert fast sicher* gegen X, falls

$$P\left(\left\{\lim_{n\to\infty} ||X_n - X|| = 0\right\}\right)$$

Notation: $X_n \xrightarrow{f.s.} X$.

Analog zum Spezialfall $E = \mathbb{R}^d$ zeigt man für Folgen von Zufallsvariablen mit Werten in einem allgemeinen Banachraum E die fast sichere Eindeutigkeit des fast sicheren Grenzwerts.

2.2 Proposition
Für $X, Y, X_1, X_2, ... \in \mathcal{L}_0(E)$ gilt

(i) $(X_n \xrightarrow{f.s.} X \wedge X_n \xrightarrow{f.s.} Y) \Rightarrow X = Y$ fast sicher.

(ii) $(X_n \xrightarrow{f.s.} X \wedge X = Y$ fast sicher$) \Rightarrow X_n \xrightarrow{f.s.} Y$.

(iii) $(X_n \xrightarrow{f.s.} X \wedge f : E \to \mathbb{R}$ stetig$) \Rightarrow f(X_n) \xrightarrow{f.s.} f(X)$.

Ferner lassen sich auch die folgenden beiden Konvergenzkriterien vollkommen problemlos in das allgemeinere Setting übertragen.

2.3 Satz
Seien $X, X_1, X_2, ... \in \mathcal{L}_0(E)$. Dann gilt

$$X_n \xrightarrow{f.s.} X \iff \forall\varepsilon > 0 \lim_{N\to\infty} P\left(\left\{\sup_{n\geq N} ||X_n - X|| > \varepsilon\right\}\right) = 0.$$

2.4 Satz (Cauchy-Kriterium für fast sichere Konvergenz)
Für eine Folge $(X_n)_{n\in\mathbb{N}}$ in $\mathcal{L}_0(E)$ existiert genau dann eine Zufallsvariable $X \in \mathcal{L}_0(E)$ mit $X_n \xrightarrow{f.s.} X$, wenn

$$\forall\varepsilon > 0 : \quad \lim_{N\to\infty} P\left(\left\{\sup_{n\geq N} ||X_n - X_N|| > \varepsilon\right\}\right) = 0.$$

Wie im skalaren Fall betrachtet man neben dem Begriff der fast sicheren Konvergenz auch den der *stochastischen Konvergenz*

2.5 Definition
Eine Folge $(X_n)_{n\in\mathbb{N}}$ von E-wertigen Zufallsvariablen *konvergiert stochastisch* gegen eine Zufallsvariable X falls

$$\forall\varepsilon > 0 : \quad \lim_{n\to\infty} P\left(\{||X_n - X|| > \varepsilon\}\right) = 0.$$

Man erhält unmittelbar die folgende Charakterisierung.

2.6 Satz (Teilfolgenkriterium)
Seien $X, X_1, X_2, ... \in \mathcal{L}_0(E)$. Die Folge $(X_n)_{n\in\mathbb{N}}$ konvergiert genau dann stochastisch gegen X, wenn es jede Teilfolge $(n_k)_{k\in\mathbb{N}}$ von $(n)_{n\in\mathbb{N}}$ eine Teilfolge $(n_{k_l})_{l\in\mathbb{N}}$ besitzt, sodass $X_{n_{k_l}} \xrightarrow{f.s.} X$.

2.7 Korollar
Für Zufallsvariablen $X, Y, X_1, X_2, ... \in \mathcal{L}_0(E)$ und eine stetige Abbildung $f : E \to \mathbb{R}$ gilt

(i) $X_n \xrightarrow{f.s.} X \;\Rightarrow\; X_n \xrightarrow{st} X$,

(ii) $(X_n \xrightarrow{st} X \wedge X_n \xrightarrow{st} Y) \Rightarrow X = Y$ fast sicher.

(iii) $(X_n \xrightarrow{st} X \wedge X = Y$ fast sicher$) \Rightarrow X_n \xrightarrow{st} Y$.

(iv) $(X_n \xrightarrow{st} X \wedge f : E \to \mathbb{R}$ stetig$) \Rightarrow f(X_n) \xrightarrow{st} f(X)$.

Wie auch bei der fast sicheren Konvergenz lassen sich die beiden folgenden Charakterisierungen komplett analog zum skalaren Fall beweisen.

2.8 Satz
Seien $X, X_1, X_2, ... \in \mathcal{L}_0(E)$. Dann gilt

$$X_n \xrightarrow{st} X \iff \forall \varepsilon > 0 \lim_{N\to\infty} \sup_{n\geq N} P(\{||X_n - X|| > \varepsilon\}) = 0.$$

2.9 Satz (Cauchy-Kriterium für stochastische Konvergenz)
Für eine Folge $(X_n)_{n\in\mathbb{N}}$ in $\mathcal{L}_0(E)$ existiert genau dann eine Zufallsvariable $X \in \mathcal{L}_0(E)$ mit $X_n \xrightarrow{st} X$, wenn

$$\forall \varepsilon > 0 : \quad \lim_{N\to\infty} \sup_{n\geq N} P(\{||X_n - X_N|| > \varepsilon\}) = 0.$$

Wie bereits aus dem skalaren Fall bekannt, lässt sich die fast sichere Konvergenz nicht durch eine Metrik beschreiben. Die stochastische Konvergenz allerdings schon. Definiere dazu eine Abbildung

$$d_P : L_0(E) \times L_0(E) \to [0, \infty)$$

durch

$$d_P([X], [Y]) := E\left(\frac{||X - Y||}{1 + ||X + Y||}\right), \quad [X], [Y] \in L_0(E). \tag{2.1}$$

Dann ist d_P wegen der Stetigkeit von $||\cdot||$ und $\left|\frac{x}{1+x}\right| \leq 1$ für alle $x \in [0, \infty)$ wohldefiniert und liefert eine weitere Charakterisierung der stochastischen Konvergenz.

2.10 Proposition
(i) d_P ist eine Metrik auf $L_0(E)$,

(ii) $X_n \xrightarrow{st} X \iff \lim_{n\to\infty} d_P([X_n], [X]) = 0$,

(iii) $(L_0(E), d_P)$ ist vollständig.

2.2 Verteilungskonvergenz und charakteristische Funktionale

2.11 Definition
Eine Folge $(X_n)_{n\in\mathbb{N}}$ von E-wertigen Zufallsvariablen *konvergiert in Verteilung* gegen eine Zufallsvariable X, falls die Folge $(P^{X_n})_{n\in\mathbb{N}}$ von Verteilungen schwach gegen P^X konvergiert.
Bezeichnung: $X_n \xrightarrow{w} X$.

2.12 Definition
Sei μ ein Wahrscheinlichkeitsmaß auf $\mathcal{B}(E)$. Das *charakteristische Funktional* $\widehat{\mu}$ von μ ist definiert durch

$$\widehat{\mu} : E' \to \mathbb{C}, \quad l \mapsto \int_E e^{il(x)} \mu(dx).$$

Wegen der Stetigkeit von $e^{i\cdot}$ und $\left|e^{ix}\right| \leq 1$ für alle $x \in \mathbb{R}$ ist $\widehat{\mu}$ wohldefiniert. Für eine Zufallsvariable $X \in \mathcal{L}_0(E)$ bezeichne $\widehat{\mu_X}$ das charakteristische Funktional der Verteilung von X. In diesem Fall lässt sich die Abbildung auch schreiben als

$$l \mapsto E(e^{il(X)}), \quad l \in E'.$$

2.13 Bemerkung
Im Fall $E = \mathbb{R}^d$ lässt sich E' auf kanonische Weise mit $\mathbb{R}^d$ identifizieren, daher lassen sich die charakteristischen Funktionale hier in der Form

$$y \mapsto \int_{\mathbb{R}^d} e^{i\langle x,y\rangle}\mu(dx)$$

schreiben. Nach dem Darstellungssatz von Riesz ist dies auch für allgemeine Hilberträume E möglich.

Wie schon im skalaren Fall dienen die charakteristischen Funktionale als nützliches Hilfsmittel zur Untersuchung von Wahrscheinlichkeitsmaßen und schwacher Konvergenz.

2.14 Satz (Eindeutigkeitssatz)
Seien μ und ν zwei Wahrscheinlichkeitsmaße auf $\mathcal{B}(E)$ mit $\widehat{\mu} = \widehat{\nu}$. Dann gilt $\mu = \nu$.

Beweis.
Bemerke zunächst. dass für ein Wahrscheinlichkeitsmaß μ auf $\mathcal{B}(X)$, $d \in \mathbb{N}$, $t = (t_1, ..., t_d) \in \mathbb{R}^d$ und $l_1, ..., l_d \in E'$ gilt

$$\widehat{\mu}(\sum_{j=1}^d t_j l_j) = \int_E e^{i\sum_{j=1}^d t_j l_j(x)}\mu(dx) = \int_E e^{i\langle t,\xi\rangle}\mu^T(d\xi) = \widehat{\mu^T}(t),$$

wobei

$$T : E \to \mathbb{R}^d, \quad x \mapsto (l_1(x), ..., l_d(x))^T.$$

Nach Voraussetzung folgt also $\widehat{\mu^T} = \widehat{\nu^T}$ und aus **APPENDIX** folgt $\mu^T = \nu^T$. Es gilt somit

$$\forall d \in \mathbb{N}\ \forall l_1, ..., l_d \in E' \forall A \in \mathcal{B}(\mathbb{R}^d) :\ \mu((l_1, ..l_d)^{-1}(A)) = \nu((l_1, ..l_d)^{-1}(A)).$$

Also stimmen μ und ν auf dem schnittstabilen Erzeuger $\mathcal{C}(E)$ von $\mathcal{B}(E)$ und somit auf ganz $\mathcal{B}(E)$ überein. □

2.15 Satz
Sei $(X_n)_{n\in\mathbb{N}}$ eine Folge in $\mathcal{L}_0(E)$ und $X \in \mathcal{L}_0(E)$. Dann konvergiert $(X_n)_{n\in\mathbb{N}}$ genau dann in Verteilung gegen X, wenn die folgenden beiden Bedingungen erfüllt sind

(a) $(\widehat{\mu_{X_n}})_{n\in\mathbb{N}}$ konvergiert punktweise gegen $\widehat{\mu_X}$ und

(b) $(X_n)_{n\in\mathbb{N}}$ ist gleichmäßig straff.

Beweis.
zu $\Rightarrow$: Für $n \in \mathbb{N}$ sei μ_n die Verteilung von X_n und μ die Verteilung von X. Da die Folge $(\mu_n)_{n\in\mathbb{N}}$ nach Voraussetzung schwach gegen μ konvergiert, ist $(\mu_n)_{n\in\mathbb{N}}$ insbesondere relativ kompakt in $(\mathcal{M}(E), \rho)$, und nach dem Satz von Prokhorov gleichmäßig straff. Ferner ist für $l \in E'$ die Abbildung

$$g_l : E \to \mathbb{C},\ x \mapsto e^{il(x)}$$

beschränkt und stetig. Durch Zerlegung in Real und Imaginärteil liefern dadurch die Linearität des Integrals und die Definition der schwachen Konvergenz die puntkweise Konvergenz der charakteristischen Funktionale.
zu $\Leftarrow$: Nach Voraussetzung ist $\{\mu_n : n \in \mathbb{N}\}$ relativ kompakt. Es genügt also zu zeigen, dass jede konvergente Teilfolge von $(\mu_n)_{n\in\mathbb{N}}$ gegen μ konvergiert. Sei dazu $(\mu_{n_k})_{k\in\mathbb{N}}$ eine konvergente Teilfolge und $\nu \in \mathcal{M}(E)$ mit $\mu_{n_k} \rightharpoonup \nu$. Nach Voraussetzung und der Hinrichtung gilt also

$$\forall l \in E' :\ \widehat{\mu}(l) = \lim_{k\to\infty} \widehat{\mu_{n_k}}(l) = \widehat{\nu}(l).$$

Der Eindeutigkeitssatz liefert nun $\mu = \nu$ und somit die Behauptung. □

2.16 Bemerkung
Erinnert man sich an den endlich dimensionalen Fall, vgl. [3][Theorem 8.8.1], so sieht man, dass die gleichmäßige Straffheit der Folge $(\mu_n)_{n\in\mathbb{N}}$ dort nicht explizit gefordert wird. Hier ist die

gleichmäßige Straffheit eine Konsequenz aus der punktweisen Konvergenz der charakteristischen Funktionen $(\widehat{\mu_n})_{n\in\mathbb{N}}$ gegen $\widehat{\mu}$. **TODO: Gegenbeispiel? Weitere Argumentation wieso das im unendlich dimensionalen Fall nicht so ist?**

Bevor wir uns im Folgenden Kapitel mit der Konvergenz zufälliger Reihen in Banachräumen beschäftigen, halten wir noch zwei für den späteren Beweis des Satzes von Itô-Nisio nützliche Ergebnisse fest, die sich aus unserem bisher gesammelten Wissen über charakteristische Funktionen leicht beweisen lassen.

2.17 Proposition
Seien $(X_n)_{n\in\mathbb{N}}$ und $(Y_n)_{n\in\mathbb{N}}$ Folgen in $\mathcal{L}_0(E)$ und $X, Y \in \mathcal{L}_0(E)$ mit $X_n \xrightarrow{st} X$, sowie $Y_n \xrightarrow{st} Y$. Falls für alle $n \in \mathbb{N}$ X_n und Y_n unabhängig sind, so sind auch X, Y unabhängig.

Beweis.
Wegen des Teilfolgenkriteriums für stochastische Konvergenz können wir ohne Einschränkung annehmen, dass $X_n \xrightarrow{f.s.} X$ und $Y_n \xrightarrow{f.s.} Y$. Betrachte die $E \times E$-wertige Zufallsvariablen $Z_n := (X_n, Y_n)$, $n \in \mathbb{N}$, und $Z := (X, Y)$. Identifiziere $(E \times E)^* = E^* \times E^*$ und erhalte mittels dominierter Konvergenz und Fubini für $(l_1, l_2) \in E^* \times E^*$.

$$\begin{aligned}
\widehat{P^Z}(x,y) = E(e^{i(l_1(X)+l_2(Y))}) &= \lim_{n\to\infty} E(e^{i(l_1(X_n)+l_2(Y_n))}) \\
&= \lim_{n\to\infty} E(e^{il_1(X_n)}E(e^{il_2(Y_n)}) \\
&= E(e^{il_1(X)})E(e^{il_2(Y)}) \\
&= \widehat{P^X}(l_1)\widehat{P^Y}(P)(l_2) = \widehat{P^X \times P^Y}(l_1, l_2).
\end{aligned}$$

Der Eindeutigkeitssatz für charakteristische Funktionen liefert unmittelbar $P^X \times P^Y = P^{(X,Y)}$. Also sind X, Y unabhängig. $\square$

Sei 0 das neutrale Element der Addition in E und δ_0 die Einpunktverteilung auf 0, d.h.

$$\delta_0(A) = 1_A(0), \quad A \in \mathcal{B}(E).$$

2.18 Proposition
Sei μ ein Wahrscheinlichkeitsmaß auf $\mathcal{B}(E)$. Falls ein $r > 0$ existiert mit

$$\forall l \in E' : \ ||l||_{op} \leq r \ \Rightarrow \ \widehat{\mu}(l) = 1,$$

so gilt $\mu = \delta_0$.

Beweis.
Wähle $l \in E'$ beliebig aber fest und betrachte die Abbildung

$$\phi : \mathbb{R} \to \mathbb{C}, \ t \mapsto \widehat{\mu}(tl).$$

Dann ist ϕ die charakteristische Funktion von μ^l und nach Voraussetzung gilt $\phi(t) = 1$ für alle $t \in \mathbb{R}$ mit $|t| \leq \frac{r}{||l||_{op}}$. Weiter gilt für $s, t \in \mathbb{R}$

$$|\phi(t) - \phi(s)| = \left| \int_E e^{isl(x)}(e^{i(t-s)l(x)} - 1)\mu(dx) \right| \leq \int_E \left| e^{isl(x)}(e^{i(t-s)l(x)} - 1) \right| \mu(dx).$$

Per Hölder-Ungleichung und der Definition des Absolutbetrags erhalten wir

$$\begin{aligned}
\int_E \left| e^{isl(x)}(e^{i(t-s)l(x)} - 1) \right| \mu(dx) &\leq \left(\int_E \left| e^{isl(x)}(e^{i(t-s)l(x)} - 1) \right|^2 \mu(dx) \right)^{\frac{1}{2}} \\
&= \sqrt{2 - \phi(t-s) - \overline{\phi(t-s)}} \\
&\leq \sqrt{2\,|1 - \phi(t-s)|},
\end{aligned}$$

und somit insgesamt

$$|\phi(t) - \phi(s)| \leq \sqrt{2\,|1 - \phi(t-s)|}.$$

Also muss ϕ konstant sein mit $\phi(t) = 1$ für alle $t \in \mathbb{R}$. Da l beliebig gewählt war erhalten wir

$$\forall l \in E' : \quad \widehat{\mu}(l) = 1 = \int_E e^{il(x)} \delta_0(dx) = \widehat{\delta_0}(l).$$

Der Eindeutigkeitssatz für charakteristische Funktionale liefert nun $\mu = \delta$. $\square$

3 Maximal-Ungleichungen und Konvergenz zufälliger Reihen

3.1 Maximalungleichungen

Bezeichne $L_0(E)$ den Vektorraum der E-wertigen-Zufallsvariablen.

3.1 Definition
Eine E-wertige Zufallsvariable X heißt *symmetrisch*, falls $-X$ die selbe Verteilung besitzt wie X, d.h.

$$\forall A \in \mathcal{B}(E) : P(\{X \in A\}) = P(\{-X \in A\}).$$

3.2 Bemerkung
Nach dem Eindeutigkeitssatz für charakteristische Funktionale ist eine Zufallsvariable $X \in L_0(E)$ genau dann symmetrisch, wenn

$$\forall z \in E' : \quad \widehat{\mu_X}(z) = \widehat{\mu_{-X}}(z).$$

3.3 Definition
Eine Folge $(X_n)_{n\in\mathbb{N}}$ von E-wertigen Zufallsvariablen heißt *symmetrisch*, falls $(\varepsilon_1 X_1, \varepsilon_2 X_2, ...)$ für jede Wahl von $\varepsilon_i = \pm 1$ die gleiche Verteilung hat wie $(X_1, X_2, ...)$.

3.4 Bemerkung
Sind $X_1, X_2, ...$ unabhängige E-wertige Zufallsvariablen, sodass X_n für alle $n \in \mathbb{N}$ symmetrisch ist, dann ist $(X_1, X_2, ...)$ symmetrisch.

3.5 Satz (Lévys Maximal-Ungleichung)
Seien $X_1, ..., X_N \in L_0(E)$ unabhängige und symmetrische Zufallsvariablen und setze

$$S_n := \sum_{i=1}^{n} X_i, \quad 1 \le n \le N.$$

Dann gilt für alle $t > 0$

$$P\big(\{\max_{1\le n\le N} ||S_n|| > t\}\big) \le 2P\big(\{||S_N|| > t\}\big), \tag{3.1}$$

$$P\big(\{\max_{1\le n\le N} ||X_n|| > t\}\big) \le 2P\big(\{||S_N|| > t\}\big). \tag{3.2}$$

Beweis.
TODO

Für nicht-symmetrische Zufallsvariablen erhalten wir mit einer ähnlichen Beweismethode die folgende auf Giuseppe Ottaviani und Anatoli Skorohod zurückgehende Maximal-Ungleichung, vgl. [5][Lemma 6.2].

3.6 Satz (Maximal-Ungleichung von Ottaviani-Skorohod)
Seien $X_1, ..., X_N$ unabhängige E-wertige Zufallsvariablen, $N \in \mathbb{N}$. Setze

$$S_k := \sum_{i=1}^{k} X_i, \quad k = 1, ..., N.$$

Dann gilt für alle $s, t > 0$

$$P(\{\max_{1\le k\le N} ||S_k|| > s + t\}) \le \frac{P(\{||S_N|| > t\})}{1 - \max_{1\le k\le N} P(\{||S_N - S_k|| > s\})}. \tag{3.3}$$

Beweis.
TODO

3.2 Der Satz von Itô-Nisio

Sei $(X_n)_{n\in\mathbb{N}}$ eine Folge unabhängiger Zufallsvariablen in $\mathcal{L}_0(E)$. Für $n \in \mathbb{N}$ setze

$$S_n := \sum_{i=1}^{n} X_i, \quad \mu_n := P^{S_n}.$$

3.7 Satz (Itô-Nisio)
Es sind äquivalent

(i) $(S_n)_{n\in\mathbb{N}}$ konvergiert fast sicher,

(ii) $(S_n)_{n\in\mathbb{N}}$ konvergiert stochastisch,

(iii) $(S_n)_{n\in\mathbb{N}}$ konvergiert in Verteilung.

Beweis.
Die Implikationen $(i) \Rightarrow (ii) \Rightarrow (iii)$ wurden bereits in Kapitel 2 gezeigt, es genügt also $(ii) \Rightarrow (i)$ und $(iii) \Rightarrow (ii)$ zu zeigen.
zu $(ii) \Rightarrow (i)$: Fall A: Für alle $n \in \mathbb{N}$ ist X_n symmetrisch.
Für $n, N \in \mathbb{N}$ mit $n < N$ setze

$$Y_{n,N} := \max_{n<k\leq N} ||S_k - S_n||,$$
$$Y_n := \lim_{N\to\infty} Y_{n,N} = \sup_{k>n} ||S_k - S_n||.$$

Seien $\varepsilon, t > 0$. Mit dem Cauchy-Kriterium für stochastische Konvergenz und Lévys Maximal-Ungleichung erhalten wir für $N > n \geq n_0 := n_0(\varepsilon, t) \in \mathbb{N}$

$$P(\{Y_{n,N} > t\}) \leq 2P(\{||S_N - S_n|| > t\}) \leq \varepsilon.$$

Es folgt somit

$$P(\{Y_n > t\}) = \lim_{n\to\infty} P(\{Y_{n,N} > t\}) \leq \varepsilon$$

für $n \geq n_0$. Also gilt $Y_n \xrightarrow{st} 0$, nach dem Cauchy-Kriterium für fast sichere Konvergenz folgt daraus die fast sichere Konvergenz von $(S_n)_{n\in\mathbb{N}}$.
Fall B: Allgemeiner Fall.
Wir verwenden hierzu die Beweistechnik der Symmetrisierung. Betrachte hierzu den Produktraum $(\Omega \times \Omega, \mathcal{A} \otimes \mathcal{A}, P \times P)$. Für eine Zufallsvariable X auf $(\Omega, \mathcal{A}, P)$ bezeichne

$$\overline{X}(\omega_1, \omega_2) := X(\omega_1) - X(\omega_2), \quad (\omega_1, \omega_2) \in \Omega \times \Omega$$

die Symmetrisierung von X. Wie man mittels der charakteristischen Funktionale von $\overline{X}$ und $-\overline{X}$ leicht einsieht ist $\overline{X}$ tatsächlich symmetrisch. Sei nun S eine Zufallsvariable auf $(Omega, \mathcal{A}, P)$ mit $\S_n \xrightarrow{st} S$. Dann folgt direkt $\overline{S_n} \xrightarrow{st} \overline{S}$, denn für $\varepsilon > 0$ gilt

$$(P \times P)(\{||\overline{S_n} - \overline{S}||\}) \leq 2P(\{||S_n - S|| > \frac{\varepsilon}{2}\}).$$

Nach Fall A gilt also insbesondere $\overline{S_n} \xrightarrow{f.s.} \overline{S}$. Es existiert also eine Menge $\Omega^* \in \mathcal{A} \otimes \mathcal{A}$ mit $P \times P(\Omega^*) = 1$ und

$$\forall(\omega_1, \omega_2) \in \Omega^* : \quad \overline{S_n}(\omega_1, \omega_2) \text{ konvergiert.}$$

Mit dem Satz von Fubini erhalten wir

$$0 = \int_{\Omega\times\Omega} 1 - 1_{\Omega^*} d(P \times P) = \int_\Omega \int_\Omega 1 - 1_{\Omega^*(\omega_2)}(\omega_1) dP(\omega_1) dP(\omega_2).$$

wobei $\Omega^*(\omega_2) = \{\omega_1 \in \Omega : (\omega_1, \omega_2) \in \Omega^*\}$. Also existiert (mindestens) ein $\omega_2 \in \Omega$ mit $P(\Omega^*(\omega_2)) = 1$ und es gilt

$$\forall \omega_1 \in \Omega^*(\omega_2) : \quad S_n(\omega_1) - S_n(\omega_2) \text{ konvergiert.}$$

Setze nun $x_n := S_n(\omega_2)$, $n \in \mathbb{N}$. Dann existiert eine Zufallsvariable L auf $(\Omega, \mathcal{A}, P)$ mit $S_n - x_n \xrightarrow{f.s.} L$. Nach Voraussetzung erhalten wir also

$$x_n \xrightarrow{st} S - L$$

wobei wir x_n für $n \in \mathbb{N}$ als konstante Zufallsvariable auf $(\Omega, \mathcal{A}, P)$ auffassen. Also existiert ein $x \in E$ mit

$$\lim_{n \to \infty} x_n = x$$

und folglich erhalten wir $S_n \xrightarrow{f.s.} L + x$.
zu $(iii) \Rightarrow (ii)$: Für $1 \leq m < n$ bezeichne

$$\mu_{m,n} := P^{S_n - S_m}$$

die Verteilung von $S_n - S_m$. Da $(\mu_n)_{n \in \mathbb{N}}$ nach Voraussetzung schwach gegen ein Wahrscheinlichkeitsmaß μ konvergiert ist die Menge $\{\mu_n : n \in \mathbb{N}\}$ relativ kompakt in $(\mathcal{M}, \rho)$ und nach dem Satz von Prokhorov somit gleichmäßig straff. Also existiert zu $\varepsilon > 0$ eine kompakte Teilmenge $K \subseteq E$ mit

$$\forall n \in \mathbb{N} : \quad \mu_n(K) \geq 1 - \varepsilon.$$

Wegen der Stetigkeit von

$$(x, y) \mapsto x - y, \quad (x, y) \in E \times E$$

ist die Menge $\tilde{K} := \{x - y : x, y \in K\}$ wiederum kompakt, also insbesondere messbar und es gilt

$$\mu_{m,n}(\tilde{K}) \geq P(\{S_n \in K, S_m \in K\}) \geq 1 - P(\{S_n \in K^c\}) - P(\{S_m \in K^c\}) \geq 1 - 2\varepsilon.$$

Also ist auch $\{\mu_{m,n} : m, n \in \mathbb{N}, m < n\}$ gleichmäßig straff und folglich relativ kompakt in $(\mathcal{M}, \rho)$. Wir zeigen nun

$$\forall \varepsilon > 0 \; \exists N \in \mathbb{N} : \; \forall n > m \geq N : \quad \mu_{m,n}(B(0, \varepsilon)) > 1 - \varepsilon. \tag{3.4}$$

Mit dem Cauchy-Kriterium folgt daraus die stochastische Konvergenz der Folge $(S_n)_{n \in \mathbb{N}}$. Angenommen (3.4) ist nicht erfüllt, dann gilt

$$\exists \varepsilon > 0 \; \forall N \in \mathbb{N} : \; \exists n(N) > m(N) \geq N : \mu_{m(N),n(N)}(B(0, \varepsilon)) \leq 1 - \varepsilon.$$

Da $\{\mu m, n : m, n \in \mathbb{N}, m < n\}$ relativ kompakt ist existiert insbesondere eine Teilfolge von $(\mu_{m(N),n(N)})_{N \in \mathbb{N}}$ die gegen ein Wahrscheinlichkeitsmaß ν auf $\mathcal{B}(E)$ konvergiert. Ohne Einschränkung können wir annehmen, dass bereits $\mu_{m(N),n(N)} \rightharpoonup \nu$ gilt. Da $B(0, \varepsilon)$ offen ist liefert das Portmanteau-Theorem

$$\nu(B(0, \varepsilon)) \leq \lim_{N \to \infty} \mu_{m(N),n(N)}(B(0, \varepsilon)) \leq 1 - \varepsilon. \tag{3.5}$$

Andererseits gilt für $z \in E'$ wegen der Unabhängigkeit von $(X_n)_{n \in \mathbb{N}}$

$$\begin{aligned}
\widehat{\mu_{n(N)}} = E(e^{iz(S_{n(N)})}) &= E(e^{iz(S_{m(N)})} e^{iz(S_{n(N)} - S_{m(N)})}) \\
&= E(e^{iz(S_{m(N)})}) E(e^{iz(S_{n(N)} - S_{m(N)})}) \\
&= \widehat{\mu_{m(N}}(z) \widehat{\mu_{m(N),n(N)}}(z).
\end{aligned}$$

Mit Grenzübergang $N \to \infty$ folgt daraus wegen $\mu_N \rightharpoonup \mu$ und $\mu_{m(N),n(N)} \rightharpoonup \nu$

$$\forall z \in E' : \quad \widehat{\mu}(z) = \widehat{\mu}(z) \widehat{\nu}(z)$$

Wegen der Stetigkeit von $\widehat{\mu}$ und $\widehat{\mu}(0) = 1$ existiert ein $r > 0$ mit

$$\widehat{\mu}(z) \neq 0, \qquad \text{für alle } z \in E' \text{ mit } ||z||_{op} \leq r.$$

Also muss

$$\widehat{\nu}(z) = 1, \qquad \text{für alle } z \in E' \text{ mit } ||z||_{op} \leq r.$$

gelten. Nach **Proposition(TODO)** gilt also $\nu = \delta_0$. Im Widerspruch zu (3.5). Also gilt (3.4) und folglich konvergiert $(S_n)_{n\in\mathbb{N}}$ stochastisch. □

3.8 Bemerkung
TODO: Auch direkter Beweis von stochastisch ⇒ fast sicher per Ottaviani-Skorohod möglich.

3.9 Satz
Sei $(\mu_n)_{n\in\mathbb{N}}$ gleichmäßig straff. Dann existiert eine Folge $(c_n)_{n\in\mathbb{N}}$ in E sodass $(S_n - c_n)_{n\in\mathbb{N}}$ fast sicher konvergiert.

Beweis.
Betrachte den Produktraum $(\Omega \times \Omega, \mathcal{A} \otimes \mathcal{A}, P \times P)$ und definiere

$$\tilde{X}_n : \Omega \times \Omega \to E,\ (\omega_1, \omega_2) \mapsto X_n(\omega_1),$$
$$\tilde{Y}_n : \Omega \times \Omega \to E,\ (\omega_1, \omega_2) \mapsto X_n(\omega_2),$$

sowie

$$\tilde{S}_n := \sum_{i=1}^{n} \tilde{X}_i, \quad \tilde{T}_n := \sum_{i=1}^{n} \tilde{Y}_n, \quad U_n := \tilde{S}_n - \tilde{T}_n, \quad \mu_{U_n} := (P \times P)^{U_n}.$$

Nach Konstruktion gilt $\tilde{S}_n \sim \tilde{T}_n \sim S_n$. Wir zeigen zunächst, dass $(\mu_{U_n})_{n\in\mathbb{N}}$ gleichmäßig straff ist. Sei dazu $\varepsilon > 0$ Dann existiert nach Voraussetzung eine kompakte Menge $K_\varepsilon \subseteq E$ mit

$$\forall n \in \mathbb{N} : \mu_n(K_\varepsilon) \geq 1 - \varepsilon.$$

Wegen der Stetigkeit der Abbildung

$$E \times E \to E,\ (x, y) \mapsto x - y$$

ist auch die Menge $K := \{x - y : x, y \in K_\varepsilon\}$ kompakt und somit insbesondere messbar. Ferner gilt

$$\begin{aligned}
\mu_{U_n}(K) = P\left(\left\{\tilde{S}_n - \tilde{T}_n \in K\right\}\right) &\geq P\left(\left\{\tilde{S}_n \in K_\varepsilon, \tilde{T}_n \in \varepsilon\right\}\right) \\
&\geq 1 - P\left(\left\{\tilde{S}_n \in K_\varepsilon\right\}\right) - P\left(\left\{\tilde{T}_n \in K_\varepsilon\right\}\right) \\
&\geq 1 - 2\varepsilon.
\end{aligned}$$

Also ist $(\mu_{U_n})_{n\in\mathbb{N}}$ straff. Als nächstes zeigen wir, dass $(\widehat{\mu_{U_n}}(l))_{n\in\mathbb{N}}$ für alle $l \in E'$ konvergiert. Sei dazu $l \in E'$ beliebig aber fest. Die Unabhängigkeit von $(\tilde{Y}_n)_{n\in\mathbb{N}}$ und $(\tilde{X}_n)_{n\in\mathbb{N}}$ liefert direkt die Unabhängigkeit von $(\tilde{X}_n - \tilde{Y}_n)_{n\in\mathbb{N}}$ und nach Konstruktion gilt zudem $Y_n \sim X_n$. Also folgt

$$\begin{aligned}
\widehat{\mu_{U_n}}(l) = E(e^{il(U_n)}) = E\Big(\prod_{j=1}^{n}(e^{il(\tilde{X}_j - \tilde{Y}_j)})\Big) &= \prod_{j=1}^{n} E(e^{il(\tilde{X}_j - \tilde{Y}_j)}) \\
&= \prod_{j=1}^{n} E(e^{il(\tilde{X}_j)})E(e^{-il(\tilde{Y}_j)}) \\
&= \prod_{j=1}^{n} \left|E(e^{il(\tilde{X}_j)})\right|^2
\end{aligned}$$

Wegen $0 \leq \left|E(e^{il(\tilde{X}_j)})\right| \leq 1$ für alle $j \in \mathbb{N}$ folgt somit die Konvergenz von $(\widehat{\mu_{U_n}}(l))_{n\in\mathbb{N}}$. Seien nun

$(\mu_{U_{n_k}})_{k\in\mathbb{N}}$ und $(\mu_{U_{m_k}})_{k\in\mathbb{N}}$ zwei konvergente Teilfolgen von $(\mu_{U_n})_{n\in\mathbb{N}}$ mit Grenzwert μ' bzw. μ''. Für $l \in E'$ gilt also

$$\widehat{\mu'}(l) = \lim_{k\to\infty} \widehat{\mu_{U_{n_k}}}(l) = \lim_{k\to\infty} \widehat{\mu_{U_{m_k}}}(l) = \widehat{\mu''}(l).$$

Mit dem Eindeutigkeitssatz für charakteristische Funktionale erhalten wir somit $\mu' = \mu''$. Nach Kapitel 2 konvergiert $(U_n)_{n\in\mathbb{N}}$ also in Verteilung und somit nach dem Satz von Itô-Nisio insbesondere fast sicher. Daher existiert eine Menge $\Omega^* \in \mathcal{A} \otimes \mathcal{A}$ mit $(P \times P)(\Omega^*) = 1$ und

$$\forall(\omega_1, \omega_2) \in \Omega^* : \quad U_n(\omega_1, \omega_2) = S_n(\omega_1) - S_n(\omega_2) \text{ konvergiert.}$$

Mit dem Satz von Fubini erhalten wir wie im Beweis des Satzes von Itô-Nisio ein $\omega' \in \Omega$, sodass $S_n - S_n(\omega')$ fast sicher konvergiert. Also erfüllt die Folge $(c_n)_{n\in\mathbb{N}}$ definiert durch $c_n := S_n(\omega')$, $n \in \mathbb{N}$, die gewünschte Eigenschaft. $\square$

3.10 Satz (Satz von Itô-Nisio für symmetrische Folgen)
Sei $(X_n)_{n\in\mathbb{N}}$ eine Folge unabhängiger und symmetrischer Zufallsvariablen in $\mathcal{L}_0(E)$. Dann sind äquivalent

(i) $(S_n)_{n\in\mathbb{N}}$ konvergiert fast sicher,

(ii) $(S_n)_{n\in\mathbb{N}}$ konvergiert stochastisch,

(iii) $(S_n)_{n\in\mathbb{N}}$ konvergiert in Verteilung,

(iv) $(\mu_n)_{n\in\mathbb{N}}$ ist gleichmäßig straff,

(v) Es gibt eine Zufallsvariable $S \in \mathcal{L}_0(E)$, sodass

$$\forall l \in E' : \quad l(S_n) \xrightarrow{st} l(S),$$

(vi) Es gibt ein Wahrscheinlichkeitsmaß μ auf $\mathcal{B}(E)$, sodass

$$\forall l \in E' : \quad \lim_{n\to\infty} \widehat{\mu_n}(l) = \widehat{\mu}(l).$$

Beweis.
Die Äquivalenz $(i) \iff (ii) \iff (iii)$ wurde bereits im allgemeinen Fall nicht-symmetrischer Zufallsvariablen gezeigt und die Implikationen $(iii) \Rightarrow (iv)$, $(i) \Rightarrow (v) \Rightarrow (vi)$ sind klar. Wir zeigen noch $(vi) \Rightarrow (v) \Rightarrow (iv) \Rightarrow (i)$.
zu $(iv) \Rightarrow (i)$: **TODO**
zu $(v) \Rightarrow (iv)$: **TODO**
zu $(vi) \Rightarrow (v)$: **TODO** $\square$

3.3 Das Kontraktions-Prinzip

Literaturverzeichnis

[1] *Probability Distributions on Banach Spaces.* D. Reidel Publishing Company, Dordrecht, 1987

[2] Amann, H. ; Escher, J.: *Analysis I.* Birkhäuser, Basel, 2006

[3] Bogachev, V.I.: *Measure Theory, Volume II.* Springer Verlag, Berlin - Heidelberg, 2007

[4] Gänssler, P. ; Stute, W.: *Wahrscheinlichkeitstheorie.* Springer Verlag, Berlin, 1977

[5] Ledoux, M. ; Talagrand, M.: *Probability in Banach Spaces.* Springer-Verlag, Berlin, 1991

[6] Parthasarathy, K.R: *Probability Measures on Metric Spaces.* Academic Press, New York – London, 1967

[7] Querenburg, B.v.: *Mengentheoretische Topologie.* Springer-Verlag, Berlin, 2001