Der Satz von Itô-Nisio

Jonas Köppl

Bachelorarbeit im Studiengang Mathematik

Betreuer: Prof. Dr. Thomas Müller-Gronbach

Universität Passau
Fakultät für Informatik und Mathematik
Lehrstuhl für Mathematische Stochastik und ihre Anwendungen

Abgegeben am: 5. Juli 2019

Inhaltsverzeichnis

Zusammenfassung

Sei $(X_n)_{n\in\mathbb{N}}$ eine Folge unabhängiger reellwertiger Zufallsvariablen und bezeichne

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

die Partialsummen. Nach einem auf P. Lévy [14] zurückgehenden Satz sind für die Folge $(S_n)_{n\in\mathbb{N}}$ die folgenden Aussagen 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.

Im Jahr 1968 gelang es K.Itô und M. Nisio in [10], dieses Resultat auf Folgen von Radon-Zufallsvariablen mit Werten in einem beliebigen separablen Banachraum zu verallgemeinern und im Fall unabhängiger symmetrischer Zufallsvariablen um weitere äquivalente Eigenschaften zu erweitern. Ziel dieser Arbeit ist es, zunächst eine (größtenteils) in sich geschlossene Einführung in die für den Beweis benötigten Hilfsmittel zu geben, und anschließend den Satz von Itô-Nisio zu beweisen. Im Anschluss an den zentralen Beweis werden zudem zwei leichte, aber interessante Anwendungen diskutiert. Neben den Inhalten der Grundvorlesungen zur Linearen Algebra und Analysis werden Grundlagen der maßtheoretischen Wahrscheinlichkeitstheorie im Umfang von [6, Kapitel I] als Vorkenntnisse angenommen.

1 Maßtheoretische Vorbereitungen

Bevor wir uns im späteren Verlauf der Arbeit mit zufälligen Reihen in Banachräumen beschäftigen können, benötigen wir ein paar maßtheoretische Vorbereitungen. Wir beginnen hierbei mit einigen grundlegenden Eigenschaften Borelscher σ-Algebren und darauf definierter Wahrscheinlichkeitsmaße. Danach gehen wir kurz auf messbare Vektorräume ein und führen dann den Begriff der Radon-Zufallsvariable mit Werten in einem Banachraum ein. Da die zusätzliche algebraische Struktur eines Banachraums für unsere Betrachtung zunächst nicht von Bedeutung ist, werden wir uns in den ersten Abschnitten dieses Kapitels mit dem allgemeineren Fall eines (vollständigen) metrischen Raumes beschäftigen. Die Darstellung der ersten drei Abschnitte orientiert sich an den beiden Standardwerken [16] und [4]. Eine gut strukturierte und lesbare Einführung zu Wahrscheinlichkeitsmaßen auf metrischen Räumen bietet zudem [5]. Der kleine Exkurs zur Topologie in Abschnitt 1.1 stammt aus [17]. Die Abschnitte 1.5 und 1.6 zu messbaren Vektorräumen und Zufallsvariablen mit Werten in einem Banachraum fußen auf [18] und [12].

1.1 Borelmengen in metrischen Räumen

Notation und Konventionen
Für einen metrischen Raum (X, d) bezeichne im Folgenden $\mathcal{B}(X)$ die Borelsche σ-Algebra in X. Zudem wird 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. Für ein Mengensystem $\mathcal{C} \subseteq \mathcal{P}(X)$ sei $\sigma(\mathcal{C})$ die *von $\mathcal{C}$ erzeugte σ-Algebra*, d.h. die kleinste σ-Algebra in $\mathcal{P}(X)$, die $\mathcal{C}$ enthält.

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

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

Man sieht leicht ein, dass $\mathcal{A}_2 = \mathcal{A}_1 \subseteq \mathcal{B}(X)$. Zu zeigen bleibt somit 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 $D \subseteq X$. Setze

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

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

Als nächstes wollen wir zeigen, dass für zwei separable metrische Räume (X_1, d_1) und (X_2, d_2), die Produkt-σ-Algebra $\mathcal{B}(X_1) \otimes \mathcal{B}(X_2)$, mit der von der Produkttopologie auf $X_1 \times X_2$ erzeugten Borelschen σ-Algebra $\mathcal{B}(X_1 \times X_2)$ übereinstimmt. Für den Beweis benötigen wir ein paar topologische Grundbegriffe und Hilfsmittel.

1.2 Definition.
Sei $(X, \mathcal{O})$ ein topologischer Raum. Eine Menge $\mathcal{U} \subseteq \mathcal{O}$ heißt *Basis* von $\mathcal{O}$, falls

$$\forall A \in \mathcal{O}\ \forall x \in A\ \exists U \in \mathcal{U}:\ x \in U \subseteq A.$$

1.3 Bemerkung.
Man zeigt leicht, dass ein Mengensystem $\mathcal{U} \subseteq \mathcal{O}$ genau dann eine Basis von $\mathcal{O}$ ist, wenn

$$\begin{aligned}\mathcal{O} &= \{A \subseteq X \mid \forall x \in A\ \exists B \in \mathcal{U}:\ x \in B \subseteq A\}\\ &= \{\cup_{i\in I} U_i \mid U_i \in \mathcal{U}, i \in I, I \text{ beliebig}\}.\end{aligned}$$

1.4 Lemma.
Sei X eine Menge, I eine nichtleere Indexmenge und $\mathcal{U} := \{B_i : B_i \in \mathcal{P}(X), i \in I\}$ mit $\bigcup_{i\in I} B_i = X$. Dann ist das Mengensystem

$$\mathcal{O}(\mathcal{U}) := \{\cup_{i\in\tilde{I}} B_i \mid \tilde{I} \subseteq I \text{ beliebig}\}$$

eine Topologie auf X und wird als die *von $\mathcal{U}$ erzeugte Topologie* bezeichnet. Nach Bemerkung 1.3 ist $\mathcal{U}$ dann insbesondere eine Basis von $\mathcal{O}(\mathcal{U})$.

Beweis.
Nach Voraussetzung gilt $X \in \mathcal{O}(\mathcal{U})$ und da auch $\tilde{I} = \emptyset$ möglich ist, gilt $\emptyset \in \mathcal{O}(\mathcal{U})$. Seien nun $n \in \mathbb{N}$ und $A_1, ..., A_n \in \mathcal{O}(\mathcal{U})$. Dann existieren Mengen $I_1, ..., I_n \subseteq I$ mit

$$A_j = \bigcup_{i\in I_j} B_i, \quad j = 1, ..., n.$$

Setze nun

$$I^{(1)} := \bigcap_{i=1}^{n} I_i.$$

Es gilt dann $I^{(1)} \subseteq I$ und

$$\bigcap_{i=1}^{n} A_i = \bigcup_{i\in I^{(1)}} B_i \in \mathcal{O}(\mathcal{U}).$$

Für eine Familie $(A_j)_{j\in J}$ von Mengen in $\mathcal{O}(\mathcal{U})$ gibt es für alle $j \in J$ eine Menge $I_j \subseteq I$ mit

$$A_j = \bigcup_{i\in I_j} B_i.$$

Setze

$$I^{(2)} := \bigcup_{j\in J} I_j.$$

Dann gilt $I^{(2)} \subseteq I$ und

$$\bigcup_{j\in J} A_j = \bigcup_{i\in I^{(2)}} B_i \in \mathcal{O}(\mathcal{U}).$$

Also ist $\mathcal{O}(\mathcal{U})$ eine Topologie. □

Man beachte ferner, dass die Topologie $\mathcal{O}(\mathcal{U})$ in der Situation von Lemma 1.4 die kleinste Topologie ist, die $\mathcal{U}$ enthält. D.h. für jede andere Topologie $\mathcal{O}$ auf X mit $\mathcal{U} \subseteq \mathcal{O}$ gilt $\mathcal{O}(\mathcal{U}) \subseteq \mathcal{O}$. Ist ferner $\mathcal{E}$ eine Basis von $\mathcal{O}$, so gilt nach Bemerkung 1.3 $\mathcal{O}(\mathcal{E}) = \mathcal{O}$.

Wie man leicht zeigt, ist für einen metrischen Raum (X, d) das Mengensystem

$$\mathcal{E} := \{B(x, \varepsilon) \mid \varepsilon > 0, x \in X\}$$

eine Basis der von d erzeugten Topologie, d.h. eine Basis von

$$\mathcal{O}_d := \{A \subseteq X \mid A \text{ ist offen bzgl. } d\}.$$

Für separable metrische Räume existiert sogar eine abzählbare Basis von $\mathcal{O}_d$.

1.5 Lemma.
Sei (X, d) ein separabler metrischer Raum und $\mathcal{O}_d$ die von d erzeugte Topologie. Weiter sei $D \subseteq X$ eine abzählbare und dichte Teilmenge von X. Dann ist

$$\tilde{\mathcal{E}} := \{B(x, \varepsilon) \mid \varepsilon \in \mathbb{Q} \cap (0, \infty), x \in D\}$$

eine abzählbare Basis von $\mathcal{O}_d$.

Beweis.
Die Abzählbarkeit von $\tilde{\mathcal{E}}$ ist klar und man sieht leicht ein, dass $\mathcal{O}(\tilde{\mathcal{E}}) \subseteq \mathcal{O}(\mathcal{E}) = \mathcal{O}_d$. Weiter gilt für eine Menge $B \in \mathcal{E}$

$$B = \bigcup_{U \in \tilde{\mathcal{E}}, U \subseteq B} U,$$

da D dicht in X liegt. Es gilt also $\mathcal{E} \subseteq \mathcal{O}(\tilde{\mathcal{E}})$ und folglich

$$\mathcal{O}_d = \mathcal{O}(\mathcal{E}) \subseteq \mathcal{O}(\tilde{\mathcal{E}}) \subseteq \mathcal{O}_d.$$

Insgesamt haben wir somit gezeigt, dass $\mathcal{O}_d = \mathcal{O}(\tilde{\mathcal{E}})$. □

1.6 Lemma.
Sei $(X, \mathcal{O})$ ein topologischer Raum mit einer abzählbaren Basis $\mathcal{C} \subseteq \mathcal{O}$. Dann gilt $\sigma(\mathcal{C}) = \mathcal{B}(X)$.

Beweis.
Aus $\mathcal{C} \subseteq \mathcal{O}$ folgt direkt $\sigma(\mathcal{C}) \subseteq \sigma(\mathcal{O}) = \mathcal{B}(X)$. Zu zeigen bleibt also nur $\mathcal{O} \subseteq \sigma(\mathcal{C})$. Sei dazu $A \in \mathcal{O}$. Da $\mathcal{C}$ eine abzählbare Basis von $\mathcal{O}$ ist, existieren Mengen $C_1, C_2, ... \in \mathcal{C}$ mit

$$A = \bigcup_{i=1}^{\infty} C_i.$$

Somit gilt $A \in \sigma(\mathcal{C})$. □

Nun können wir das zuvor angekündigte und für später wichtige Resultat über Borelsche σ-Algebren auf Produkten separabler metrischer Räume zeigen.

1.7 Proposition.
Seien (X_1, d_1) und (X_2, d_2) zwei separable metrische Räume. Dann gilt

$$\mathcal{B}(X_1 \times X_2) = \mathcal{B}(X_1) \otimes \mathcal{B}(X_2).$$

Beweis.
Seien $\mathcal{O}_1$ und $\mathcal{O}_2$ die von d_1 bzw. d_2 erzeugten Topologien und $\mathcal{O}$ die Produkttopologie auf $X_1 \times X_2$. Nach Lemma 1.5 existiert für $i = 1, 2$ eine abzählbare Basis $\mathcal{C}_i \subseteq \mathcal{O}_i$ von $\mathcal{O}_i$. Betrachte nun das Mengensystem

$$\mathcal{C} = \{A_1 \times A_2 \mid A_1 \in \mathcal{C}_1, A_2 \in \mathcal{C}_2\}.$$

Für $k = 1,2$ bezeichne π_k die Projektion auf die k-te Komponente. Nach Definition der Produkttopologie ist

$$\mathcal{Z} := \{\pi_1(O_1) \cap \pi_2(O_2) \mid O_1 \in \mathcal{O}_1, O_2 \in \mathcal{O}_2\} = \{O_1 \times O_2 \mid O_1 \in \mathcal{O}_1, O_2 \in \mathcal{O}_2\}$$

eine Basis von $\mathcal{O}$. Da $\mathcal{C}_1$ und $\mathcal{C}_2$ Basen von $\mathcal{O}_1$ bzw. $\mathcal{O}_2$ sind, zeigt man leicht, dass $\mathcal{O}(\mathcal{Z}) = \mathcal{O}(\mathcal{C})$. Folglich ist $\mathcal{C}$ eine abzählbare Basis von $\mathcal{O}$ und nach Lemma 1.6 gilt $\sigma(\mathcal{C}) = \mathcal{B}(X_1 \times X_2)$. Wir zeigen nun

$$\sigma(\mathcal{C}) = \mathcal{B}(X_1) \otimes \mathcal{B}(X_2).$$

Zu $\subseteq$: Diese Inklusion ist klar, wegen $\mathcal{C} \subseteq \mathcal{B}(X_1) \times \mathcal{B}(X_2) \subseteq \mathcal{B}(X_1) \otimes \mathcal{B}(X_2)$.
Zu $\supseteq$: Nach der Definition der Produkttopologie $\mathcal{O}$ sind die Projektionen π_1, π_2 stetig bzgl. $\mathcal{O}$, also insbesondere $\mathcal{B}(X_1 \times X_2)/\mathcal{B}(X_1)$- bzw. $\mathcal{B}(X_1 \times X_2)/\mathcal{B}(X_2)$-messbar. Somit gilt

$$\mathcal{B}(X_1) \otimes \mathcal{B}(X_2) \subseteq \mathcal{B}(X_1 \times X_2),$$

da $\mathcal{B}(X_1) \otimes \mathcal{B}(X_2)$ die Initial-σ-Algebra der Abbildungen π_1, π_2 ist. $\square$

1.8 Bemerkung.
Die Aussage von Proposition 1.7 lässt sich sogar auf abzählbare Produkte separabler metrischer Räume verallgemeinern. Das heißt, für eine Folge $((X_i, d_i))_{i \in \mathbb{N}}$ von separablen metrischen Räumen gilt

$$\mathcal{B}(X_1 \times X_2 \times ...) = \otimes_{i \in \mathbb{N}} \mathcal{B}(X_i).$$

Der Beweis funktioniert ähnlich wie der von Proposition 1.7 und ist beispielsweise in [11, Lemma 1.2] skizziert.

1.2 Borelmaße auf metrischen Räumen

Im folgenden Abschnitt beschäftigen wir uns mit Wahrscheinlichkeitsmaßen auf der Borelschen σ-Algebra $\mathcal{B}(X)$ eines metrischen Raums (X, d), welche teilweise auch als *Borelsche Wahrscheinlichkeitsmaße* bezeichnet werden. Zunächst interessieren wir uns hierbei für Regularitätseigenschaften solcher Maße, welche uns die spätere Arbeit erleichtern werden. Im zweiten Teil dieses Abschnitts untersuchen wir Folgen von Wahrscheinlichkeitsmaßen und den Begriff der schwachen Konvergenz.

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*, $\overline{A}$ der *Abschluss*, und

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

der *Rand der Menge*. Setze ferner

$$C_b(X) := \{f : X \to \mathbb{R} \mid f \text{ ist stetig und beschränkt}\}.$$

Für eine Folge $(A_n)_{n \in \mathbb{N}}$ von Mengen und eine Menge A schreiben wir $A_n \downarrow A$, falls $A_1 \supseteq A_2 \supseteq ...$ und $\cap_{n=1}^{\infty} A_n = A$.

1.9 Definition.
Ein Wahrscheinlichkeitsmaß μ auf $\mathcal{B}(X)$ heißt *regulär*, falls

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

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

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

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

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

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

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

$$\mu(O_n) - 2^{-n}\varepsilon < \mu(A_n) < \mu(C_n) + 2^{n+1}\varepsilon. \tag{1.1}$$

Insbesondere gilt somit

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

Zudem ist $\bigcup_{n=1}^{\infty} O_n$ offen und es gilt nach (1.1)

$$\begin{aligned} \mu\left(\bigcup_{n=1}^{\infty} O_n\right) - \mu\left(\bigcup_{n=1}^{\infty} A_n\right) &\le \mu\left(\bigcup_{n=1}^{\infty} O_n \setminus \bigcup_{n=1}^{\infty} A_n\right) \\ &\le \mu\left(\bigcup_{n=1}^{\infty} (O_n \setminus A_n)\right) \le \sum_{n=1}^{\infty} \mu(O_n \setminus A_n) < \sum_{n=1}^{\infty} 2^{-n}\varepsilon = \varepsilon. \end{aligned} \tag{1.2}$$

Wegen

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

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

$$\mu(\cup_{n=1}^{\infty} C_n) - \mu(\cup_{n=1}^{k} C_n) < \frac{\varepsilon}{2}. \tag{1.3}$$

Die Menge $C := \cup_{n=1}^{k} C_n$ ist als endliche Vereinigung abgeschlossener Mengen abgeschlossen und nach Konstruktion in $\cup_{n=1}^{\infty} A_n$ enthalten. Ferner gilt wegen (1.1) und (1.3)

$$\begin{aligned}
\mu(\bigcup_{n=1}^{\infty} A_n) - \mu(C) &< \mu(\bigcup_{n=1}^{\infty} A_n) - \mu(\bigcup_{n=1}^{\infty} C_n) + \frac{\varepsilon}{2} \\
&\leq \mu(\bigcup_{n=1}^{\infty} A_n \setminus \bigcup_{n=1}^{\infty} C_n) + \frac{\varepsilon}{2} \\
&\leq \mu(\bigcup_{n=1}^{\infty} (A_n \setminus C_n)) + \frac{\varepsilon}{2} \\
&\leq \sum_{n=1}^{\infty} \mu(A_n \setminus C_n) + \frac{\varepsilon}{2} = \varepsilon. \qquad (1.4)
\end{aligned}$$

Da ε beliebig gewählt war, folgt aus (1.2) und (1.4), dass $\cup_{n=1}^{\infty} A_n \in \mathcal{R}$. Folglich ist $\mathcal{R}$ eine σ-Algebra. Es bleibt noch zu zeigen, dass $\mathcal{R}$ alle abgeschlossenen Mengen enthält. Da die leere Menge abgeschlossen und offen ist, folgt aus der Monotonie von μ direkt $\emptyset \in \mathcal{R}$. Sei also $\emptyset \neq A \subseteq X$ abgeschlossen. Die Bedingung

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

ist erfüllt, weil μ monoton ist. 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 $O_n \downarrow A$, da A abgeschlossen ist. Mit der σ-Stetigkeit von μ folgt letztendlich

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

Infolgedessen gilt $A \in \mathcal{R}$. Da die abgeschlossenen Mengen $\mathcal{B}(X)$ erzeugen, gilt $\mathcal{R} = \mathcal{B}(X)$ und folglich ist μ regulär. □

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

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

1.12 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 weil μ nach Proposition 1.10 regulär ist, gibt es eine abgeschlossene Menge $C \subseteq A$ mit $\mu(C) > \mu(A) - \frac{\varepsilon}{2}$. Die Menge $K_\varepsilon \cap C$ ist als abgeschlossene Teilmenge einer

kompakten Menge wiederum kompakt und es gilt

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

Da $\varepsilon > 0$ beliebig gewählt war, folgt daraus

$$\mu(A) = \sup\{\mu(K) \ : \ K \subseteq A, \ K \text{ kompakt}\}.$$

□

1.13 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* bezeichnet.

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

Wir verwenden zum Beweis von Proposition 1.14 die folgende Charakterisierung kompakter Teilmengen metrischer Räume. Der Beweis dieses Lemmas wird mittels der in metrischen Räumen geltenden Äquivalenz von Überdeckungskompaktheit und Folgenkompaktheit geführt und ist beispielsweise in [2, Theorem III.3.10] zu finden.

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

(a) K ist vollständig und

(b) 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. Insbesondere gilt dann für jedes $q \in \mathbb{N}$

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

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

$$\mu\Big(\bigcup_{i=1}^{N_q} \overline{B}(x_i, 2^{-q})\Big) \geq 1 - \varepsilon 2^{-q}. \tag{1.5}$$

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 $K \subseteq \cup_{i=1}^{N_q} B(x_i, 2^{-q}) \subseteq \cup_{i=1}^{N_q} B(x_i, \varepsilon)$. Nach Lemma 1.15 ist K somit kompakt. Zudem gilt wegen (1.5)

$$\begin{aligned}\mu(K) = 1 - \mu\big(\bigcup_{q=1}^{\infty} \bigcap_{i=1}^{N_q} \overline{B}(x_i, 2^{-q})^c\big) &\geq 1 - \sum_{q=1}^{\infty} \mu\big(\bigcap_{i=1}^{N_q} \overline{B}(x_i, 2^{-q})^c\big) \\ &\geq 1 - \sum_{q=1}^{\infty} \varepsilon 2^{-q} = 1 - \varepsilon.\end{aligned}$$

Also ist μ straff. □

1.16 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 $E \in \mathcal{B}(X)$ mit $\mu(E) = 1$.

Beweis.
Zu (i) $\Rightarrow$ (ii): Für alle $n \in \mathbb{N}$ existiert eine kompakte Menge $K_n \subseteq X$ mit $\mu(K_n) \geq 1 - \frac{1}{n}$. Da endliche Vereinigungen kompakter Mengen wiederum kompakt sind, können wir ohne Einschränkung annehmen, dass $K_n \subseteq K_{n+1}$. Wegen der σ-Stetigkeit von μ gilt

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

Da kompakte Teilmengen metrischer Räume insbesondere separabel sind, ist $\overline{E} := \overline{\cup_{n=1}^{\infty} K_n}$ als topologischer Abschluss einer abzählbaren Vereinigung separabler Mengen ebenso separabel und als abgeschlossene Menge messbar.
Zu (ii) $\Rightarrow$ (i): Analog zum Beweis von Proposition 1.14. □

Insgesamt erhalten wir aus unseren bisherigen Überlegungen den folgenden Satz.

1.17 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 $E \in \mathcal{B}(X)$ mit $\mu(E) = 1$.

(iii) μ ist ein Radon-Maß, d.h.

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

1.18 Bemerkung.
Sei $(E, ||\cdot||)$ ein Banachraum und $A \subseteq E$ eine separable Teilmenge. Dann ist nach Korollar $A.5$ auch der abgeschlossenen Untervektorraum $\overline{lin(A)}$ separabel und es gilt $A \subseteq \overline{lin(A)}$. Für ein Wahrscheinlichkeitsmaß μ auf $\mathcal{B}(E)$ sind also äquivalent:

(i) μ ist straff.

(ii) Es gibt einen separablen und abgeschlossenen Untervektorraum E_0 mit $\mu(E_0) = 1$.

(iii) μ ist ein Radon-Maß.

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

1.19 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: $\mu_n \rightharpoonup \mu$.

Um die schwache Konvergenz von Wahrscheinlichkeitsmaßen auf metrischen Räumen zu beschreiben, gibt es zahlreiche äquivalente Formulierungen. Ein paar davon werden im folgenden Satz, der meist als *Portmanteau-Theorem* bezeichnet wird, zusammengefasst.

1.20 Satz. (Portmanteau-Theorem)
Sei (X, d) ein metrischer Raum und seien $\mu, \mu_1, \mu_2, ...$ 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

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

(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

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

Dann ist die Menge U_m für alle $m \in \mathbb{N}$ offen und es gilt $U_m \downarrow A$, weil A abgeschlossen ist. Aufgrund der σ-Stetigkeit von μ existiert also ein $k \in \mathbb{N}$ mit

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

Betrachte nun die Abbildung

$$f : X \to \mathbb{R}, \quad x \mapsto \max\{1 - k \inf_{y\in A} d(x, y), 0\}.$$

Offensichtlich ist f beschränkt. Ferner ist die Abbildung

$$d(\cdot, A) : X \to \mathbb{R}, \quad x \mapsto d(x, A) := \inf_{y \in A} d(x, y)$$

nach der umgekehrten Dreiecksungleichung stetig. Da die Abbildung

$$x \mapsto \max\{x, 0\}, \quad x \in \mathbb{R},$$

stetig ist, ist somit f als Komposition stetiger Funktionen 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 $\mu(\overline{C}) = \mu(C) = \mu(\mathring{C})$. Da (iii) auch (ii) impliziert, erhalten wir somit

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

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 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 Prinzip von Cavalieri, vgl. [6, Satz 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

Nachdem wir im letzten Abschnitt damit begonnen haben, Folgen von Wahrscheinlichkeitsmaßen auf schwache Konvergenz zu untersuchen, wollen wir im Folgenden ein weiteres wichtiges Hilfsmittel für dieses Unterfangen besprechen. Zunächst einmal sei dafür an den Begriff der relativen Kompaktheit erinnert.

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

kompakt ist.

Mittels der Teilfolgencharakterisierung kompakter Teilmengen metrischer Räume lassen sich die relativ kompakten Teilmengen eines metrischen Raumes auch wie folgt beschreiben.

1.22 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 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:

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

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

Ziel dieses Abschnitts ist es daher, zunächst eine adäquate Metrik zu konstruieren, die die schwache Konvergenz von Wahrscheinlichkeitsmaßen beschreibt. Anschließend besprechen wir eine nützliche Charakterisierung der relativen Kompaktheit einer Menge von Wahrscheinlichkeitsmaßen.

Notation und Konventionen
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 $\varepsilon > 0$

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

Man zeigt leicht, dass B^ε für alle $B \in \mathcal{B}(X)$ offen ist. Daher ist B^ε insbesondere messbar.

Betrachte nun die Abbildung

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

definiert durch

$$(\mu,\nu) \mapsto \inf\big\{\alpha > 0 \mid \forall B \in \mathcal{B}(X) : (\mu(B) \leq \nu(B^\alpha) + \alpha) \;\wedge\; (\nu(B) \leq \mu(B^\alpha) + \alpha)\big\}. \tag{1.6}$$

1.23 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 auf der rechten Seite von (1.6) ist nicht leer, weil 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 monoton fallende Folge $(\alpha_n)_{n\in\mathbb{N}}$ mit $\lim_{n\to\infty} \alpha_n = 0$ und

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

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

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

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

$$\begin{aligned} \forall A \in \mathcal{B}(X): \quad & \mu(A) \leq \eta(A^\alpha) + \alpha \quad \wedge \quad \eta(A) \leq \mu(A^\alpha) + \alpha, && (1.7)\\ & \nu(A) \leq \eta(A^\beta) + \beta \quad \wedge \quad \eta(A) \leq \nu(A^\beta) + \beta. && (1.8) \end{aligned}$$

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

$$\rho(\mu, \nu) \leq \alpha + \beta.$$

Da diese Ungleichung für alle $\alpha, \beta > 0$ mit (1.7) und (1.8) erfüllt ist, folgt durch Bilden der Infima

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

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 $B \subseteq (A^\alpha)^c$. Für $B \in \mathcal{B}(X)$ setze also $A := (B^\alpha)^c$ 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 monoton fallende Folge $(\alpha_n)_{n\in\mathbb{N}}$ mit $\lim_{n\to\infty} \alpha_n = 0$ 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 mit der σ-Stetigkeit von μ folgt

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

Nach Satz 1.20 gilt somit $\mu_n \rightharpoonup \mu$. $\square$

Falls der zugrunde liegende metrische Raum separabel ist, so gilt auch die Umkehrung von Proposition 1.23(iii). Zum Beweis benötigen wir das folgende technische Lemma, welches uns die Existenz spezieller Überdeckungen eines separablen metrischen Raumes garantiert.

1.24 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

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

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

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

Da die Abbildung

$$d(x_n, \cdot) : X \to \mathbb{R}, \quad y \mapsto d(x_n, y)$$

stetig ist, folgt aus

$$S(x_n, r) = d(x_n, \cdot)^{-1}(\{r\})$$

direkt die Messbarkeit von $S(x_n, r)$. Zudem 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 $\mu(S(x_n, r_n)) = 0$. Es gilt ferner

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

weil D dicht in X liegt. Folglich erfüllen $(x_n)_{n\in\mathbb{N}}$ und $(r_n)_{n\in\mathbb{N}}$ die geforderten Bedingungen. □

1.25 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 1.23(iii).
Zu $\Leftarrow$: Wir zeigen

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

Daraus folgt zusammen mit Proposition 1.23(ii) die Behauptung. Sei $\delta := \frac{\varepsilon}{4} > 0$. Dann existiert nach Lemma 1.24 eine Folge $(B_n)_{n\in\mathbb{N}} := (B(x_n, r_n))_{n\in\mathbb{N}}$ offener Kugeln mit Radius $0 < r_n < \frac{\delta}{2}$ und $\mu(\partial B_n) = 0$ für alle $n \in \mathbb{N}$, sowie $X = \cup_{n\in\mathbb{N}} B_n$. Wegen Letzterem existiert ein $k \in \mathbb{N}$ mit

$$\mu\left(\bigcup_{i=1}^{k} B_i\right) \geq 1 - \delta. \tag{1.9}$$

Betrachte nun das endliche Mengensystem

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

Dann gilt für alle $A \in \mathcal{C}$

$$\partial A \subseteq \bigcup_{i=1}^{k} \partial B_i,$$

also $\mu(\partial A) = 0$. Nach Satz 1.20 gilt somit $\mu_n(A) \to \mu(A)$ für alle $A \in \mathcal{C}$. Da $\mathcal{C}$ endlich ist, existiert infolgedessen $N \in \mathbb{N}$ mit

$$\forall n \geq N \ \forall A \in \mathcal{C} : \quad |\mu(A) - \mu_n(A)| < \delta. \tag{1.10}$$

Insbesondere gilt daher

$$\forall n \geq N : \quad \mu_n \left(\bigcup_{i=1}^{k} B_i \right) \geq \mu \left(\bigcup_{i=1}^{k} B_i \right) - \delta \geq 1 - 2\delta. \tag{1.11}$$

Für $B \in \mathcal{B}(X)$ betrachte die Menge

$$A := \bigcup_{j \in J} B_j,$$

wobei $J := \{j \in \{1, ..., k\} : B_j \cap B \neq \emptyset\}$. Für jedes $x \in A$ existiert dann ein $j \in J$ mit $x \in B_j = B(x_j, r_j)$. Nach Definition von J gibt es ferner ein $y \in B$ mit $y \in B_j$. Mit der Dreiecksungleichung erhalten wir somit

$$d(x, y) \leq d(x, x_j) + d(x_j, y) < 2r_j < \delta.$$

Also gilt $A \subseteq B^\delta$. Unter Verwendung von $A \subseteq B^\delta \subseteq B^\varepsilon$, (1.10) und (1.11) für alle $n \geq N$ ergibt sich nun

$$\mu(B) \leq \mu(A) + \mu\left((\cup_{i=1}^{k} B_i)^c\right) \leq \mu(A) + 2\delta \leq \mu_n(A) + 3\delta \leq \mu_n(B^\varepsilon) + \varepsilon.$$

□

Nachdem wir nun eine passende Metrik gefunden haben, um für separable metrische Räume (X, d) die schwache Konvergenz von Wahrscheinlichkeitsmaßen in $\mathcal{M}(X)$ zu beschreiben, wollen wir noch kurz auf ein Resultat von Prokhorov eingehen, welches uns eine äußerst nützliche Charakterisierung der relativ kompakten Teilmengen von $(\mathcal{M}(X), \rho)$ liefert. Dazu führen wir den Begriff der *gleichmäßigen Straffheit* ein.

1.26 Definition.
Eine Menge $M \subseteq \mathcal{M}(X)$ 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. Den Beweis findet man beispielsweise in [16, Theorem II.6.7] oder [4, Theorem 5.1, Theorem 5.2].

1.27 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) alle (schwach) konvergenten Teilfolgen von $(\mu_n)_{n\in\mathbb{N}}$ konvergieren gegen den selben Grenzwert.

1.28 Bemerkung.
In [13] findet sich ein alternativer Zugang zur Prokhorov-Metrik. Hier wird die Metrik auf $\mathcal{M}(X)$ nicht wie in (1.6) über die Mengen aus $\mathcal{B}(X)$ definiert, sondern über Funktionen $f \in C_b(\mathbb{R})$. Es wird dort gezeigt, dass es für einen vollständigen und separablen metrischen Raum (X, d) eine Folge $(f_n)_{n\in\mathbb{N}}$ in der Einheitskugel von $(C_b(\mathbb{R}), ||\cdot||_\infty)$ gibt, sodass durch

$$(\mu, \nu) \mapsto \sum_{n=1}^{\infty} \frac{1}{2^n} \left| \int_X f_n d\mu - \int_X f_n d\nu \right|, \quad \mu, \nu \in \mathcal{M}(X),$$

eine Metrik definiert ist, die die schwache Konvergenz charakterisiert.

1.4 Flache Konzentrierung von Wahrscheinlichkeitsmaßen

Für unsere spätere Betrachtung von Zufallsvariablen mit Werten in einem Banachraum lohnt es sich hier, noch eine weitere Charakterisierung der relativen Kompaktheit von Familien von Wahrscheinlichkeitsmaßen zu diskutieren. Der Begriff der *flachen Konzentrierung* von Wahrscheinlichkeitsmaßen wurde zuerst von A.D. de Acosta betrachtet, vgl. [1]. Die Darstellung hier orientiert sich an [18].

Notation und Konventionen
Im Folgenden sei $(E, ||\cdot||)$ ein separabler reeller Banachraum und $(E', ||\cdot||_{op})$ der zugehörige Dualraum. Der metrische Raum der Wahrscheinlichkeitsmaße auf $\mathcal{B}(E)$ werde mit $(\mathcal{M}(E), \rho)$ bezeichnet, wobei ρ die durch (1.6) definierte Prokhorov-Metrik ist. Für ein Maß μ auf $\mathcal{B}(E)$, einen messbaren Raum $(X, \mathcal{X})$ und eine messbare Abbildung $f : E \to X$ bezeichne μ^f das *Bildmaß* von μ unter f. Dieses ist definiert als

$$\mu^f(A) := \mu(f^{-1}(A)), \quad A \in \mathcal{X}.$$

Es sei ferner daran erinnert, dass jeder endlichdimensionale Untervektorraum $S \subseteq E$ abgeschlossen ist und eine Menge $A \subseteq S$ nach dem Satz von Heine-Borel genau dann kompakt ist, wenn sie abgeschlossen und beschränkt ist. Insbesondere besitzt jede beschränkte Folge in S eine konvergente Teilfolge.

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

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

1.30 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 $\{x_1, x_2, ...\} \subseteq \overline{A}$. Für $\varepsilon > 0$ ist $(B(x_n, \varepsilon))_{n\in\mathbb{N}}$ somit eine offene Überdeckung von $\overline{A}$. Wegen der Kompaktheit von $\overline{A}$ existiert $I \subseteq \mathbb{N}$ endlich mit

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

Sei daher S der von $\{x_i : i \in I\}$ erzeugte endlichdimensionale 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 existiert insbesondere eine Folge $(y_n)_{n\in\mathbb{N}}$ in S mit

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

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 existiert demnach $N(\varepsilon) \in \mathbb{N}$, sodass für $k, m \geq N(\varepsilon) \in \mathbb{N}$

$$||x_{n_k}^{(0)} - x_{n_m}^{(0)}|| \leq ||x_{n_k}^{(0)} - y_{n_k}|| + ||y_{n_k} - y_{n_m}|| + ||x_{n_m}^{(0)} - y_{n_m}|| \leq 3\varepsilon$$

gilt. Durch Entfernen endlich vieler Folgenglieder können wir ohne Einschränkung annehmen, dass

$$\forall k, m \in \mathbb{N} : \quad ||x_{n_k}^{(0)} - x_{n_m}^{(0)}|| \leq 3\varepsilon.$$

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

$$\forall m, n \in \mathbb{N} : \quad ||x_n^{(N)} - x_m^{(N)}|| \leq \frac{3}{N}.$$

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. $\square$

1.31 Lemma.
Sei $S \subseteq E$ ein endlichdimensionaler Untervektorraum und $f_1, ..., f_n \in E'$ Funktionale mit

$$\forall x, y \in S: \; x \neq y \;\Rightarrow\; \exists k \in \{1, ..., n\}: \quad f_k(x) \neq f_k(y). \tag{1.12}$$

Dann ist die Menge

$$B := \overline{S^\varepsilon} \cap \{x \in E : |f_1(x)| \leq r_1, ..., |f_n(x)| \leq r_n\}$$

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

Beweis.
Wegen (1.12) definiert

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

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 $(x_m)_{m \in \mathbb{N}}$ in B mit

$$\lim_{m \to \infty} ||x_m|| = \infty.$$

Wegen $B \subseteq \overline{S^\varepsilon}$ gibt es somit eine Folge $(y_m)_{m \in \mathbb{N}}$ in S mit

$$\forall m \in \mathbb{N}: \quad ||x_m - y_m|| \leq \varepsilon.$$

Mittels der Dreiecksungleichung erhält man daraus

$$\lim_{m \to \infty} ||y_m|| = \infty.$$

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

$$\lim_{m \to \infty} |f_k(y_m)| = \infty. \tag{1.13}$$

Da $f_1, ..., f_n$ stetig und linear sind, existiert ferner ein $K > 0$ mit

$$\forall j \in \{1, ..., n\} \; \forall x, y \in E: \quad |f_j(x) - f_j(y)| \leq K||x - y||.$$

Nach Definition von B gilt

$$\forall m \in \mathbb{N}: \quad |f_k(x_m)| \leq r_k.$$

Also folgt für alle $m \in \mathbb{N}$ aus der Dreiecksungleichung

$$|f_k(y_m)| \leq |f_k(y_m) - f_k(x_m)| + |f_k(x_m)| \leq K\varepsilon + r_k.$$

Im Widerspruch zu (1.13). $\square$

1.32 Bemerkung.
Sei $S \subseteq E$ ein endlichdimensionaler Untervektorraum von E mit $dim(S) = r$. Man sagt eine Familie von Abbildungen mit der Eigenschaft (1.12) *trenne die Punkte von S*. Ist $\{x_1, ..., x_r\}$ eine Basis von S, so lässt sich jedes Element $x \in S$ eindeutig darstellen als $x = \sum_{i=1}^r \lambda_i(x) x_i$, wobei $\lambda_1(x), ..., \lambda_r(x) \in \mathbb{R}$. Definiere nun für $k \in \{1, ..., r\}$

$$\tilde{f}_k : S \to \mathbb{R}, \quad x \mapsto \lambda_k(x).$$

Dann sind $\tilde{f}_1, ..., \tilde{f}_r$ offensichtlich linear und da S endlichdimensional ist, folgt daraus bereits die Stetigkeit. Ferner trennen $\tilde{f}_1, ..., \tilde{f}_r$ die Punkte von S, da $\{x_1, ..., x_r\}$ eine Basis von S ist. Mit Satz A.9 erhalten wir nun Funktionale $f_1, ..., f_r \in E'$ mit $f_j|S = \tilde{f}_j$ für alle $j = 1, ..., r$. Da $\tilde{f}_1, ..., \tilde{f}_r$ die Punkte von S trennen, gilt dies insbesondere auch für $f_1, ..., f_r \in E'$. Wir haben also gezeigt, dass es zu jedem endlichdimensionalen Untervektorraum S von E tatsächlich endlich viele stetige lineare Funktionale gibt, die die Punkte von S trennen.

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

(a) Für alle $f \in E'$ ist $\{\mu^f : \mu \in M\} \subseteq \mathcal{M}(\mathbb{R})$ relativ kompakt und

(b) M ist flach konzentriert.

Beweis.
Zu $\Rightarrow$: Sei $M \subseteq \mathcal{M}(E)$ relativ kompakt. Nach Satz 1.27 ist M dann insbesondere gleichmäßig straff. Daher existiert zu $\varepsilon > 0$ eine kompakte Menge $K \subseteq E$ mit

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

Da alle $f \in E'$ stetig sind, ist jeweils auch $f(K)$ kompakt und es gilt

$$\forall \mu \in M : \quad \mu^f(f(K)) = \mu(f^{-1}(f(K))) \geq \mu(K) \geq 1 - \varepsilon.$$

Also ist $\{\mu^f : \mu \in M\}$ für alle $f \in E'$ gleichmäßig straff. Erneutes Anwenden von Satz 1.27 liefert (a). Da K insbesondere relativ kompakt ist, liefert Lemma 1.30 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 $\Leftarrow$: Sei $\varepsilon > 0$. Dann existiert für alle $n \in \mathbb{N}$ ein endlichdimensionaler Untervektorraum $S_{n,\varepsilon} \subseteq E$ mit

$$\forall \mu \in M : \quad \mu(S_{n,\varepsilon}^{\frac{\varepsilon}{2^{n+1}}}) \geq 1 - \frac{\varepsilon}{2^{n+1}}. \tag{1.14}$$

Nach Bemerkung 1.32 existieren $f_1^{(n)}, ..., f_{k_n}^{(n)} \in E'$ mit

$$\forall x, y \in S_{n,\varepsilon} : \quad x \neq y \Rightarrow \ \exists j \in \{1, ..., k_n\} : \quad f_j(x) \neq f_j(y).$$

Nach Voraussetzung (a) können wir zudem $r_1^{(n)}, ..., r_{k_n}^{(n)} \in (0, \infty)$ so wählen, dass

$$\inf_{\mu \in M} \mu^{f_i}\big([-r_i^{(n)}, r_i^{(n)}]\big) \geq 1 - \frac{\varepsilon}{k_n 2^{n+1}}, \quad i = 1, ..., k_n. \tag{1.15}$$

Setze ferner

$$F_{n,\varepsilon} := \overline{\big(S_{n,\varepsilon}^{\frac{\varepsilon}{2^{n+1}}}\big)}.$$

Dann ist die Menge

$$K := \bigcap_{n=1}^{\infty} \big(F_{n,\varepsilon} \cap \{x \in E : \ \left|f_1^{(n)}\right| \leq r_1^{(n)}, ..., \left|f_{k_n}^{(n)}\right| \leq r_{k_n}^{(n)}\}\big)$$

nach Lemma 1.31 beschränkt und als Schnitt abgeschlossener Mengen abgeschlossen. Ferner gilt für alle $n \in \mathbb{N}$
$$K \subseteq F_{n,\varepsilon} \subseteq S_{n,\varepsilon}^{\frac{\varepsilon}{2^n}}.$$
Also ist K nach Lemma 1.30 kompakt. Zudem gilt für $\mu \in M$ nach (1.14) und (1.15)

$$\begin{aligned}
\mu(K^c) &\leq \sum_{n=1}^{\infty} \mu(F_{n,\varepsilon}^c) + \sum_{i=1}^{k_n} \mu^{f_i^{(n)}}\left([-r_i^{(n)}, r_i^{(n)}]^c\right) \\
&\leq \sum_{n=1}^{\infty} \mu(F_{n,\varepsilon}^c) + \frac{\varepsilon}{2^{n+1}} \\
&\leq \sum_{n=1}^{\infty} (\frac{\varepsilon}{2^{n+1}} + \frac{\varepsilon}{2^{n+1}}) = \varepsilon.
\end{aligned}$$

Folglich ist M gleichmäßig straff und somit nach Satz 1.27 relativ kompakt. $\square$

1.34 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.
Zu $\Rightarrow$: Nach Satz 1.27 und Satz 1.33 ist M flach konzentriert und gleichmäßig straff. Somit existiert für jedes $\varepsilon > 0$ eine kompakte Menge $K \subseteq E$ mit

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

Als kompakte Menge ist K insbesondere beschränkt und die Behauptung ist gezeigt.
Zu $\Leftarrow$: Nach Voraussetzung ist M flach konzentriert, es genügt daher nach Satz 1.33 zu zeigen, dass $\{\mu^f : \mu \in M\}$ für alle $f \in E'$ relativ kompakt ist. Sei dazu $\varepsilon > 0$ und $f \in E'$. Dann existiert eine beschränkte Teilmenge $L \subseteq E$ mit

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

Da f ein stetiger linearer Operator ist, existiert ein $R > 0$ mit

$$\forall x \in E: \quad |f(x)| \leq R||x||.$$

Somit folgt aus der Beschränktheit von L, dass $f(L)$ eine beschränkte Teilmenge von $\mathbb{R}$ ist und nach dem Satz von Heine-Borel ist $\overline{f(L)}$ kompakt. Es gilt ferner

$$\forall \mu \in M: \quad \mu^f(\overline{f(L)}) = \mu(\{x \in E: \ f(x) \in \overline{f(L)}\}) \geq \mu(L) \geq 1 - \varepsilon.$$

Also ist $\{\mu^f : \mu \in M\}$ gleichmäßig straff und nach Satz 1.27 relativ kompakt. $\square$

1.5 Messbare Vektorräume

Bislang haben wir uns fast ausschließlich mit dem Zusammenspiel von Maßen und den topologischen Eigenschaften der zugrunde liegenden Räume beschäftigt. In Banachräumen steht uns aber auch die algebraische Struktur eines Vektorraums zur Verfügung. Allerdings ist per se nicht klar, ob die algebraischen Operationen mit der messbaren Struktur kompatibel, also messbar, sind. Diese Überlegung führt direkt zur Definition eines *messbaren Vektorraums*. Wir beschäftigen uns im Folgenden der Einfachheit halber ausschließlich mit Vektorräumen über $\mathbb{R}$. Die Sätze aus den folgenden Abschnitten lassen sich aber analog für $\mathbb{C}$-Vektorräume formulieren und beweisen.

1.35 Definition.
Sei X ein Vektorraum und $\mathcal{C}$ eine σ-Algebra auf X. Das Tupel $(X, \mathcal{C})$ heißt *messbarer Vektorraum*, falls die folgenden beiden Bedingungen erfüllt sind:

(a) Die Abbildung

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

ist $\mathcal{C} \otimes \mathcal{C}/\mathcal{C}$-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.

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

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

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

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

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

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

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

Beweis.
Aus der Messbarkeit der Skalarmultiplikation und Vektoraddition folgt zunächst

$$\begin{aligned} \forall A \in \mathcal{C} : \quad A^{(1)} &:= \{(\alpha, x) \in \mathbb{R} \times X : \ \alpha x \in A\} \in \mathcal{B}(\mathbb{R}) \otimes \mathcal{C}, \\ A^{(2)} &:= \{(x, y) \in X \times X : \ x + y \in A\} \in \mathcal{C} \otimes \mathcal{C}. \end{aligned}$$

Man beachte nun, 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 $\omega_1 \in \Omega_1$

$$A(\omega_1) = \{\omega_2 \in \Omega_2 : (\omega_1, \omega_2) \in A\} \in \mathcal{A}_2$$

gilt. In unserem Fall erhalten wir für festes $\alpha \in R$ und $y \in X$

$$\begin{aligned} f_\alpha^{-1}(A) &= \{x \in X : \ \alpha x \in A\} = A^{(1)}(\alpha) \in \mathcal{C}, \\ g_y^{-1}(A) &= \{x \in X : \ x + y \in A\} = A^{(2)}(y) \in \mathcal{C}. \end{aligned}$$

□

Da die Komposition messbarer Abbildungen wiederum messbar ist, erhält man unmittelbar

1.37 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.

Bislang wissen wir noch nicht einmal, ob es überhaupt nicht-triviale Beispiele messbarer Vektorräume gibt. Das wollen wir nun ändern.

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

Beweis.
Nach Proposition 1.7 gilt $\mathcal{B}(E \times E) = \mathcal{B}(E) \otimes \mathcal{B}(E)$ und $\mathcal{B}(\mathbb{R} \times E) = \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(E)$. Ferner sind die Abbildungen

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

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

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

Im Folgenden bezeichne $(X', ||\cdot||_{op})$ den Dualraum eines normierten Vektorraums $(X, ||\cdot||)$.

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

Beweis.
Zur Messbarkeit der Addition: Es genügt zu zeigen, dass

$$\forall f \in \Gamma : \big(g : X \times X \to \mathbb{R},\ (x, y) \mapsto f(x + y)\big) \text{ ist messbar.}$$

Sei dazu $f \in \Gamma$. Wegen der Linearität von f gilt für $(x, y) \in X \times X$

$$f(x + y) = (f(x) + f(y)).$$

Weiter ist die Abbildung

$$h : X \times X \to \mathbb{R}, (x, y) \mapsto f(x) + f(y)$$

als Komposition der messbaren Funktionen

$$\begin{aligned} &h_1 : X \times X \to \mathbb{R} \times \mathbb{R}, (x, y) \mapsto (f(x), f(y)), \\ &h_2 : \mathbb{R} \times \mathbb{R} \to \mathbb{R}, (x, y) \mapsto x + y \end{aligned}$$

messbar. Die Abbildung h_2 ist messbar, weil $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ nach Beispiel 1.39 ein messbarer Vektorraum ist. Die Messbarkeit der Skalarmultiplikation zeigt man ähnlich. □

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

Beweis.
Da alle $f \in E'$ stetig sind, gilt offensichtlich $\sigma(E') \subseteq \mathcal{B}(E)$. Wegen der Separabilität von E wird $\mathcal{B}(E)$ nach Proposition 1.1 von den abgeschlossenen Kugeln erzeugt und weil $(E, \sigma(E'))$ nach Proposition 1.40 ein messbarer Vektorraum ist, genügt es nach Bemerkung 1.36 zu zeigen, dass $\overline{B}(0,1)$ in $\sigma(E')$ enthalten ist. Da E separabel ist, existiert nach Korollar A.12 eine Folge $(f_n)_{n\in\mathbb{N}}$ in E' mit

$$\forall x \in E : \quad ||x|| = \sup_{n\in\mathbb{N}} |f_n(x)|\,.$$

Es gilt also

$$\overline{B}(0,1) = \{x \in E : ||x|| \leq 1\} = \{x \in E : \sup_{n\in\mathbb{N}} |f_n(x)| \leq 1\} = \bigcap_{n=1}^{\infty} \{x \in E : |f_n(x)| \leq 1\} \in \sigma(E').$$

□

1.6 Zufallsvariablen mit Werten in Banachräumen

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

1.42 Definition.
Eine Abbildung $X : \Omega \to E$ heißt *(E-wertige) Radon-Zufallsvariable*, falls sie die folgenden beiden Bedingungen erfüllt:

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

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

1.43 Bemerkung.
Nach Bemerkung 1.18 sind für eine $\mathcal{A}/\mathcal{B}(X)$-messbare Abbildung $X : \Omega \to E$ äquivalent:

(i) Es existiert ein abgeschlossener 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 aus [12] stammende Bezeichnung Radon-Zufallsvariable.
Da $(\mathbb{R}^d, ||\cdot||)$ für alle $d \in \mathbb{N}$ und jede Norm $||\cdot||$ auf $\mathbb{R}^d$ ein separabler Banachraum ist, ist jede $\mathbb{R}^d$-wertige Zufallsvariable auch eine Radon-Zufallsvariable. Der Begriff der Radon-Zufallsvariable ist also eine Verallgemeinerung des aus dem endlichdimensionalen Fall bekannten Begriffs.
Ist ferner $(X_n)_{n\mathbb{N}}$ eine Folge von Radon-Zufallsvariablen mit Werten in E, so dass die zugehörige Folge der Verteilungen $(P^{X_n})_{n\in\mathbb{N}}$ gleichmäßig straff ist, dann nennen wir $(X_n)_{n\in\mathbb{N}}$ gleichmäßig straff.

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 $\mathcal{L}_0(E)$ bzw. $L_0(E)$.

1.44 Proposition.
Für alle $X, Y \in \mathcal{L}_0(E)$ und $\alpha, \beta \in \mathbb{R}$ gilt $\alpha X + \beta Y \in \mathcal{L}_0(E)$.

Beweis.
Da X, Y Radon-Zufallsvariablen sind, existieren zwei abgeschlossene separable Untervektorräume E_X und E_Y mit

$$P(\{X \in E_X\}) = P(\{Y \in E_Y\}) = 1.$$

Nach Korollar $A.4$ ist dann auch $E_0 := \overline{lin(E_X \cup E_Y)}$ separabel und es gilt

$$P(\{X \in E_0\}) = P(\{Y \in E_0\}) = 1.$$

Da $\mathcal{A}$ vollständig ist, können wir also ohne Beschränkung der Allgemeinheit annehmen, dass E selbst separabel ist. Nach Proposition 1.38 ist dann $(E, \mathcal{B}(E))$ ein messbarer Vektorraum und die Behauptung folgt nun aus Proposition 1.37. $\square$

Wie im skalaren Fall lassen sich Radon-Zufallsvariablen durch sogenannte *einfache Zufallsvariablen* approximieren.

1.45 Definition.
Eine Abbildung $X : \Omega \to E$ von der Form

$$X = \sum_{i=1}^{n} 1_{A_i} x_i$$

mit $x_1, ..., x_n \in E$ und $A_1, ..., A_n \in \mathcal{A}$ heißt *einfache Zufallsvariable*. Offensichtlich gilt $X \in \mathcal{L}_0(E)$.

1.46 Proposition.
Sei $X \in \mathcal{L}_0(E)$ und $Y : \Omega \to E$ eine Abbildung.

(i) Dann existiert eine Folge einfacher Zufallsvariablen $(X_n)_{n\in\mathbb{N}}$ mit $X_n \xrightarrow{f.s.} X$ und

$$||X_n(\omega)|| \leq 2||X(\omega)||$$

für alle $n \in \mathbb{N}$ und fast alle $\omega \in \Omega$.

(ii) Sei $(Y_n)_{n\in\mathbb{N}}$ eine Folge in $\mathcal{L}_0(E)$, sodass eine Menge $\Omega^* \in \mathcal{A}$ existiert mit $P(\Omega^*) = 1$ und

$$\forall \omega \in \Omega^* : \quad \lim_{n\to\infty} Y_n(\omega) = Y(\omega).$$

Dann gilt $Y \in \mathcal{L}_0(E)$.

Beweis.
Zu (i): Da X eine Radon-Zufallsvariable ist, existiert ein abgeschlossener separabler Untervektorraum $E_0 \subseteq E$ mit $P(\{X \in E_0\}) = 1$. Da $\mathcal{A}$ vollständig ist können wir daher ohne Einschränkung annehmen, dass E selbst bereits separabel ist. Sei $\{x_1, x_2, ...\} \subseteq E$ eine dichte Teilmenge. Für $n \in \mathbb{N}$ betrachte die Abbildung

$$T_n : \Omega \to \mathbb{N}, \quad \omega \mapsto \inf\{k \leq n : ||X(\omega) - x_k|| = \min_{1\leq l\leq n} ||X(\omega) - x_l||\}.$$

Aus der Messbarkeit von X und der Stetigkeit von $||\cdot - x||$ für $x \in E$ folgt direkt die Messbarkeit von T_n, da für alle $i, j \in \{1, ..., n\}$ die Menge $\{||X - x_i|| < ||X - x_j||\}$ messbar ist. Setze nun

$$X_n := \sum_{k=1}^{n} 1_{\{T_n=k\}} x_k, \quad n \in \mathbb{N}.$$

Dann ist $(X_n)_{n\in\mathbb{N}}$ eine Folge einfacher Zufallsvariablen und es gilt für $n \in \mathbb{N}$ und $\omega \in \Omega$

$$||X(\omega) - X_n(\omega)|| = \min_{1\leq k\leq n} ||X(\omega) - x_k||.$$

Da $\{x_1, x_2, ...\}$ dicht in E liegt, folgt für alle $\omega \in \Omega$

$$\lim_{n\to\infty} ||X(\omega) - X_n(\omega)|| = \lim_{n\to\infty} \min_{1\leq k\leq n} ||X(\omega) - x_k|| = \inf_{n\in\mathbb{N}} ||X(\omega) - x_n|| = 0.$$

Zu (ii): Da Y_n für jedes $n \in \mathbb{N}$ eine Radon-Zufallsvariable ist, existiert eine Folge $(E_n)_{n\in\mathbb{N}}$ von abgeschlossenen separablen Untervektorräumen von E mit

$$\forall n \in \mathbb{N} : \quad P(\{Y_n \in E_n\}) = 1.$$

Nach Korollar $A.4$ ist $E_0 := \overline{lin(\cup_{n\in\mathbb{N}} E_n)}$ ein abgeschlossener separabler Untervektorraum von E und es gilt

$$\forall n \in \mathbb{N} : \quad P(\{Y_n \in E_0\}) = 1.$$

Da $\mathcal{A}$ vollständig ist, können wir daher ohne Einschränkung annehmen, dass E separabel ist. Ferner können wir wegen der Vollständigkeit von $\mathcal{A}$ annehmen, dass $\lim_{n\to\infty} Y_n(\omega) = Y(\omega)$ für alle $\omega \in \Omega$. Wir zeigen, dass für jede abgeschlossene Menge $A \subseteq E$ das Urbild $Y^{-1}(A)$ in $\mathcal{A}$ liegt. Da $\mathcal{B}(E)$ von den abgeschlossenen Mengen erzeugt wird folgt daraus die Behauptung. Für eine abgeschlossene Menge $\emptyset \neq A \subseteq E$ betrachte für $k \in \mathbb{N}$ die offene, also auch messbare, Menge

$$A_k := \{x \in E : \inf_{y\in A} ||x - y|| < \frac{1}{k}\}.$$

Da A abgeschlossen ist gilt

$$\{Y \in A\} = \bigcap_{k=1}^{\infty} \liminf_{n\to\infty} \{Y_n \in A_k\} = \bigcap_{k=1}^{\infty} \bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} \{Y_n \in A_k\} \in \mathcal{A}.$$

$\square$

1.47 Bemerkung.
Neben dem Begriff der Radon-Zufallsvariable findet man in der Literatur auch den der *starken Messbarkeit*. Eine Abbildung $f : \Omega \to E$ heißt *stark messbar*, falls es eine Folge $(f_n)_{n\in\mathbb{N}}$ einfacher Zufallsvariablen gibt, die punktweise gegen f konvergiert. Proposition 1.46 zeigt, dass für eine Abbildung $X : \Omega \to E$ die folgenden beiden Bedingungen äquivalent sind:

(i) X ist eine Radon-Zufallsvariable.

(ii) Es existiert eine stark messbare Abbildung $f : \Omega \to E$ mit $P(\{X = f\}) = 1$.

Mit Hilfe von Proposition 1.46 lässt sich zudem auch ohne Rückgriff auf Abschnitt 1.5 leicht zeigen, dass die Summe zweier Radon-Zufallsvariablen wieder eine Radon-Zufallsvariable ist. Die starke Messbarkeit und die Approximation durch einfache Zufallsvariablen sind ferner für die Definition des Bochner-Integrals nützlich. Eine ausführliche Konstruktion des Bochner-Integrals findet sich etwa in [8].

2 Konvergenzarten

Notation und Konventionen
Sei $(E, ||\cdot||)$ ein Banachraum über $\mathbb{R}$ und $(\Omega, \mathcal{A}, P)$ ein vollständiger Wahrscheinlichkeitsraum.

2.1 Fast sichere und stochastische Konvergenz

In diesem Abschnitt wollen wir kurz auf die beiden Konzepte der fast sicheren und der stochastischen Konvergenz von Radon-Zufallsvariablen eingehen, die uns schon für $\mathcal{R}^d$-wertige Zufallsvariablen bekannt sind. Aufgrund der Vollständigkeit von E lassen sich die Beweise in diesem Abschnitt wie im skalaren Fall durchführen, daher werden wir hier nur kurz die für den weiteren Verlauf wichtigen Ergebnisse zusammenfassen. Die Beweise für den skalaren Fall findet man beispielsweise in [6].

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

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

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

Analog zum Spezialfall $E = \mathbb{R}^d$ zeigt man für Folgen von Radon-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 ein $X \in \mathcal{L}_0(E)$ 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 Radon-Zufallsvariablen *konvergiert stochastisch* gegen eine Radon-Zufallsvariable X falls

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

Bezeichnung: $X_n \xrightarrow{st} X$.

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 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 E-wertige Radon-Zufallsvariablen $X, Y, X_1, X_2, ...$ 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 ein $X \in \mathcal{L}_0(E)$ mit $X_n \xrightarrow{st} X$, wenn

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

2.2 Konvergenz in Verteilung und charakteristische Funktionale

Neben der fast sicheren und der stochastischen Konvergenz beschäftigt man sich im skalaren Fall auch mit der Konvergenz in Verteilung einer Folge von Zufallsvariablen. Auch dieser Konvergenzbegriff lässt sich leicht aus dem skalaren Fall in das allgemeinere Setting von Radon-Zufallsvariblen mit Werten in einem Banachraum übertragen.

2.10 Definition.
Eine Folge $(X_n)_{n\in\mathbb{N}}$ von E-wertigen Radon-Zufallsvariablen *konvergiert in Verteilung* gegen ein $X \in \mathcal{L}_0(E)$, falls die Folge $(P^{X_n})_{n\in\mathbb{N}}$ von Verteilungen schwach gegen P^X konvergiert.
Bezeichnung: $X_n \xrightarrow{w} X$.

Da die Verteilungskonvergenz lediglich von den Verteilungen der beteiligten Radon-Zufallsvariablen abhängt, erhalten wir aus Satz 1.20 unmittelbar die folgenden Charakterisierungen.

2.11 Proposition.
Für $X, X_1, X_2, ... \in \mathcal{L}_0(E)$ sind äquivalent:

(i) $(X_n)_{n\in\mathbb{N}}$ konvergiert in Verteilung gegen X.

(ii) Für alle abgeschlossenen Mengen $A \subseteq E$ gilt

$$\limsup_{n\to\infty} P\left(\{X_n \in A\}\right) \leq P\left(\{X \in A\}\right).$$

(iii) Für alle offenen Mengen $O \subseteq E$ gilt

$$\liminf_{n\to\infty} P\left(\{X_n \in O\}\right) \geq P\left(\{X \in O\}\right).$$

(iv) Für alle Mengen $C \in \mathcal{B}(X)$ mit $P\left(\{X \in \partial C\}\right) = 0$ gilt

$$\lim_{n\to\infty} P\left(\{X_n \in C\}\right) = P\left(\{X \in C\}\right).$$

Falls E separabel ist können wir ferner auf die Prokhorov-Metrik und die Sätze 1.27 und 1.33 von Prokhorov und de Acosta zurückgreifen. Wie im Fall $\mathbb{R}^d$-wertiger Zufallsvariablen zeigt man ferner

2.12 Proposition.
Seien $X, X_1, X_2, ... \in \mathcal{L}_0(E)$ mit $X_n \xrightarrow{st} X$. Dann gilt insbesondere $X_n \xrightarrow{w} X$.

Ein wichtiges Hilfsmittel zur Untersuchung von Verteilungskonvergenz sind die sogenannten *charakteristischen Funktionale*. Sie sind eine Verallgemeinerung der aus dem skalaren Fall bekannten charakteristischen Funktionen. Die folgende Darstellung orientiert sich an [15], [13] und [10].

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

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

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 Radon-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

$$f \mapsto \mathbb{E}(e^{if(X)}), \quad f \in E'.$$

2.14 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 Fréchet-Riesz, vgl. [19, Theorem V.3.6], ist dies auch für allgemeine Hilberträume E möglich.

2.15 Proposition.
Sei μ ein Wahrscheinlichkeitsmaß auf $\mathcal{B}(E)$. Dann ist $\widehat{\mu} : E' \to \mathbb{C}$ stetig bzgl. der Norm $||\cdot||_{op}$.

Beweis.
Sei $f \in E'$ und $(f_n)_{n\in\mathbb{N}}$ eine Folge in E' mit $\lim_{n\to\infty}||f - f_n||_{op} = 0$. Dann konvergiert $(f_n)_{n\in\mathbb{N}}$ insbesondere punktweise gegen f und mit dem Satz von der dominierten Konvergenz erhalten wir

$$\lim_{n\to\infty} \widehat{\mu}(f_n) = \int_E \lim_{n\to\infty} e^{if_n(x)}\mu(dx) = \int_E e^{if(x)}\mu(dx) = \widehat{\mu}(f).$$

$\square$

Wie schon im skalaren Fall dienen die charakteristischen Funktionale als nützliches Hilfsmittel zur Untersuchung von Wahrscheinlichkeitsmaßen und schwacher Konvergenz. Es sei daran erinnert, dass nach Proposition 1.41 für einen separablen Banachraum E die beiden σ-Algebren $\mathcal{B}(E)$ und $\sigma(E')$ übereinstimmen. Wie man leicht einsieht ist zudem

$$\mathcal{C}(E) := \{\bigcap_{i=1}^{n} f_i^{-1}(A) : \ n \in \mathbb{N}, \ f_1, ..., f_n \in E', \ A_1, ..., A_n \in \mathcal{B}(\mathbb{R})\}$$

ein schnitt-stabiler Erzeuger von $\sigma(E')$.

2.16 Satz. (Eindeutigkeitssatz)
Sei E ein separabler Banachraum und 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 $f_1, ..., f_d \in E'$ gilt

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

wobei

$$T : E \to \mathbb{R}^d, \quad x \mapsto (f_1(x), ..., f_d(x)).$$

Aus der Voraussetzung folgt also $\widehat{\mu^T} = \widehat{\nu^T}$ und mit dem Eindeutigkeitssatz für charakteristische Funktionen im $\mathbb{R}^d$, vgl. [6, Satz 8.7.1], erhalten wir $\mu^T = \nu^T$. Es gilt somit

$$\forall d \in \mathbb{N}\ \forall f_1, ..., f_d \in E'\ \forall A \in \mathcal{B}(\mathbb{R}^d) : \ \mu((f_1, ..., f_d)^{-1}(A)) = \nu((f_1, ..., f_d)^{-1}(A)).$$

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

2.17 Satz.
Sei E ein separabler Banachraum, $(X_n)_{n\in\mathbb{N}}$ eine Folge in $\mathcal{L}_0(E)$ und $X \in \mathcal{L}_0(E)$. Sei ferner $\mu_n := P^{X_n}$ und $\mu := P^X$. 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) $\{\mu_n : \ n \in \mathbb{N}\}$ ist gleichmäßig straff.

Beweis.
Zu $\Rightarrow$: 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 somit nach Satz 1.27 gleichmäßig straff. Ferner ist für $f \in E'$ die Abbildung

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

beschränkt und stetig. Insbesondere sind also der Realteil $Re(g_f)$ und der Imaginärteil $Im(g_f)$ von g_f stetig und beschränkt. 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 $\{\mu_n : n \in \mathbb{N}\}$ gleichmäßig straff, also nach Satz 1.27 relativ kompakt. Es genügt daher 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

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

Satz 2.16 liefert nun $\mu = \nu$ und somit die Behauptung. □

2.18 Bemerkung.
Im endlichdimensionalen Fall muss die gleichmäßige Straffheit von $(\mu_n)_{n\in\mathbb{N}}$ nicht zusätzlich gefordert werden, vgl. [6, Satz 8.7.5].

Bevor wir uns im folgenden Kapitel mit der Konvergenz zufälliger Reihen in Banachräumen beschäftigen, halten wir noch zwei für die Beweise der Sätze 3.6 und 3.9 nützliche Ergebnisse fest, die sich mit unserem bisher gesammelten Wissen über charakteristische Funktionale leicht beweisen lassen.

2.19 Proposition.
Sei E ein separabler Banachraum und 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.
Nach Satz 2.6 können wir ohne Einschränkung annehmen, dass $X_n \xrightarrow{f.s.} X$ und $Y_n \xrightarrow{f.s.} Y$. Betrachte die $(E \times E)$-wertigen Radon-Zufallsvariablen $Z_n := (X_n, Y_n)$, $n \in \mathbb{N}$, und $Z := (X, Y)$. Für $f \in (E \times E)'$ setze $f_1 := f(\cdot, 0)$ und $f_2 := f(0, \cdot)$. Für $(x, y) \in (E \times E)$ gilt dann $f(x, y) = f_1(x) + f_2(y)$. Mittels dominierter Konvergenz und dem Satz von Fubini erhalten wir

$$\begin{aligned}
\widehat{P^Z}(f) = \mathbb{E}(e^{i(f_1(X)+f_2(Y))}) &= \lim_{n\to\infty} \mathbb{E}(e^{i(f_1(X_n)+f_2(Y_n))}) \\
&= \lim_{n\to\infty} \mathbb{E}(e^{if_1(X_n)})\mathbb{E}(e^{if_2(Y_n)}) \\
&= \mathbb{E}(e^{if_1(X)})\mathbb{E}(e^{if_2(Y)}) \\
&= \widehat{P^X}(f_1)\widehat{P^Y}(P)(f_2) = \widehat{P^X \times P^Y}(f).
\end{aligned}$$

Nach Satz 2.16 gilt somit $P^X \times P^Y = P^{(X,Y)}$. Also sind X, Y unabhängig. □

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

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

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

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

so gilt $\mu = \delta_0$.

Beweis.
Sei $f \in E' \setminus \{0\}$ beliebig, aber fest. Betrachte die Abbildung

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

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

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

Per Hölder-Ungleichung und der Definition des Absolutbetrags erhalten wir

$$\begin{aligned}
\int_E \left| e^{isf(x)}(e^{i(t-s)f(x)} - 1) \right| \mu(dx) &\leq \left(\int_E \left| e^{isf(x)}(e^{i(t-s)f(x)} - 1) \right|^2 \mu(dx) \right)^{\frac{1}{2}} \\
&= \left(\int_E \left| e^{i(t-s)f(x)} - 1 \right|^2 \mu(dx) \right)^{\frac{1}{2}} \\
&= \left(\int_E \left| \cos((t-s)f(x)) + i \sin((t-s)f(x)) - 1 \right|^2 \mu(dx) \right)^{\frac{1}{2}} \\
&= \left(\int_E \big(2 - 2\cos((t-s)f(x))\big) \mu(dx) \right)^{\frac{1}{2}} \\
&= \sqrt{2(1 - Re(\phi(t-s)))} \\
&\leq \sqrt{2|1 - \phi(t-s)|}.
\end{aligned}$$

Es gilt also

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

Daraus folgt

$$\forall t, s \in \mathbb{R} : \quad |t - s| \leq \frac{r}{||f||_{op}} \ \Rightarrow \ \phi(t) = \phi(s).$$

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

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

Nach Satz 2.16 gilt somit $\mu = \delta_0$. □

2.21 Bemerkung.
Auch wenn im Rahmen dieser Arbeit nur reelle Vektorräume betrachtet werden, sei darauf hingewiesen, dass sich viele der vorhergehenden Resultate analog für $\mathbb{C}$-Vektorräume formulieren und beweisen lassen. Ist E ein $\mathbb{C}$-Vektorraum und μ ein Wahrscheinlichkeitsmaß auf $\mathcal{B}(E)$, so

definiert man das charakteristische Funktional $\widehat{\mu}$ von μ wie folgt:

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

In [9, Lemma E.1.15] ist ferner beschrieben, wie sich die in diesem Abschnitt gesammelten Resultate aus dem reellen Fall in den komplexen Fall übertragen lassen. Durch die Einschränkung auf reelle Banachräume büßen wir also nichts an Allgemeinheit ein.

3 Maximal-Ungleichungen und Konvergenz zufälliger Reihen

Notation und Konventionen
Im Folgenden sei E ein separabler Banachraum über $\mathbb{R}$ und $(\Omega, \mathcal{A}, P)$ ein vollständiger Wahrscheinlichkeitsraum. Wegen der Separabilität von E ist jede messbare Abbildung $X : (\Omega, \mathcal{A}) \to (E, \mathcal{B}(E))$ eine Radon-Zufallsvariable. Es sei ferner daran erinnert, dass nach Proposition 1.44 die Summe endlich vieler E-wertiger Radon-Zufallsvariablen wieder eine Radon-Zufallsvariable ist.

3.1 Maximalungleichungen

Um den Satz von Itô-Nisio zu beweisen, benötigen wir noch eine auf P.Lévy zurückgehende Ungleichung für Summen unabhängiger symmetrischer Radon-Zufallsvariablen. Die Beweise in diesem Abschnitt orientieren sich an [13] und [3].

3.1 Definition.
Eine E-wertige Radon-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 2.16 ist $X \in \mathcal{L}_0(E)$ genau dann symmetrisch, wenn

$$\forall f \in E' : \quad \widehat{P^X}(f) = \widehat{P^{-X}}(f) = \widehat{P^X}(-f).$$

3.3 Satz. (Lévys Maximal-Ungleichung)
Sei $N \in \mathbb{N}$ und seien $X_1, ..., X_N \in \mathcal{L}_0(E)$ unabhängige und symmetrische Radon-Zufallsvariablen. Setze

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

Dann gilt für alle $t > 0$

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

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

Beweis.
Zu (3.1): Setze

$$T_1(\omega) := \inf\{k \leq N : ||S_k|| > t\} \in [0, \infty], \quad \omega \in \Omega.$$

Dann ist T_1 messbar, da für jedes $i = 1, ..., N$ die Menge $\{||S_i|| \leq t\}$ messbar ist. Wir zeigen zunächst

$$\{T_1 = n\} \subseteq \{||S_N|| > t,\ T_1 = n\} \cup \{||2S_n - S_N|| > t,\ T_1 = n\}, \quad n \in \{1, ..., N\}, \tag{3.3}$$
$$P\left(\{||S_N|| > t,\ T_1 = n\}\right) = P\left(\{||2S_n - S_N|| > t,\ T_1 = n\}\right), \quad n \in \{1, ..., N\}. \tag{3.4}$$

Zu (3.3): Für $\omega \in \Omega$ mit $T_1(\omega) = n$ und $||S_N(\omega)|| \leq t$ liefert die umgekehrte Dreiecksungleichung

$$||2S_n(\omega) - S_N(\omega)|| \geq 2||S_n(\omega)|| - ||S_N(\omega)|| > 2t - t = t.$$

Zu (3.4): Setze $\varepsilon_1 = \varepsilon_2 = ... = \varepsilon_n = 1$ und $\varepsilon_{n+1} = ... = \varepsilon_N = -1$, sowie

$$S'_j := \sum_{i=1}^{j} \varepsilon_i X_i, \quad j \in \{1, ..., N\}.$$

Dann gilt $S_j = S'_j$ für alle $j \leq n$ und

$$2S_n - S_N = 2\sum_{i=1}^{n} X_i - \sum_{i=1}^{N} X_i = \sum_{i=1}^{n} X_i - \sum_{i=n+1}^{N} X_i = S'_N.$$

Wegen der Symmetrie und Unabhängigkeit von $X_1, ..., X_N$ sind $(S_1, ..., S_N)$ und $(S'_1, ..., S'_N)$ identisch verteilt. Also ergibt sich

$$\begin{aligned} P\left(\{||S_N|| > t, T_1 = n\}\right) &= P\left(\{||S_1|| \leq t, ..., ||S_{n-1}|| \leq t, ||S_n|| > t, ||S_N|| > t\}\right) \\ &= P\left(\left\{||S'_1|| \leq t, ..., ||S'_{n-1}|| \leq t, ||S'_n|| > t, ||S'_N|| > t\right\}\right) \\ &= P\left(\{||2S_n - S_N|| > t, T_1 = n\}\right). \end{aligned}$$

Wir erhalten nun mit (3.3) und (3.4)

$$P\left(\{T_1 = n\}\right) \leq 2P\left(\{||S_N|| > t,\ T_1 = n\}\right).$$

Daraus folgt schließlich

$$\begin{aligned} P\left(\left\{\max_{1 \leq n \leq N} ||S_n|| > t\right\}\right) \leq \sum_{n=1}^{N} P\left(\{T_1 = n\}\right) &\leq \sum_{n=1}^{N} 2P\left(\{||S_N|| > t,\ T_1 = n\}\right) \\ &= 2P\left(\{||S_N|| > t,\ T_1 \leq N\}\right) \leq 2P\left(\{||S_N|| > t\}\right). \end{aligned}$$

Zu (3.2): Setze
$$T_2(\omega) := \inf\{k \leq N : ||X_k(\omega)|| > t\} \in [0, \infty], \quad \omega \in \Omega.$$
Analog zum Beweis von (3.3) und (3.4) zeigt man

$$\{T_2 = n\} \subseteq \{||S_N|| > t, T_2 = n\} \cup \{||2X_n - S_N|| > t, T_2 = n\}, \quad n \in \{1, ..., N\},$$
$$P\left(\{||S_N|| > t, T_2 = n\}\right) = P\left(\{||2X_n - S_N||, T_2 = n\}\right), \quad n \in \{1, ..., N\},$$

unter der Verwendung der Symmetrie und Unabhängigkeit von $X_1, ..., X_N$ und folgert daraus letztendlich

$$P\left(\left\{\max_{1\leq n\leq N}||X_n||>t\right\}\right)=\sum_{n=1}^{N}P\left(\{T_2=n\}\right)\leq\sum_{n=1}^{N}2P\left(\{||S_N||>t,T=n\}\right)$$
$$=2P\left(\{||S_N||>t,T_2\leq N\}\right)\leq 2P\left(\{||S_N||>t\}\right).$$

□

3.4 Korollar.
Sei $(X_n)_{n\in\mathbb{N}}$ eine Folge unabhängiger und symmetrischer E-wertiger Radon-Zufallsvariablen. Definiere

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

Falls $S_n\xrightarrow{w}S$ für $S\in\mathcal{L}_0(E)$, dann gilt für λ-fast-alle $t\in\mathbb{R}$

$$P\left(\left\{\sup_{n\in\mathbb{N}}||S_n||>t\right\}\right)\leq 2P\left(\{||S||>t\}\right).$$

Wobei λ das Lebesgue-Maß auf $\mathcal{B}(\mathbb{R})$ bezeichnet.

Beweis.
Für alle $n\in\mathbb{N}$ gilt nach Ungleichung (3.1)

$$P\left(\left\{\max_{1\leq k\leq n}||S_k||>t\right\}\right)\leq 2P\left(\{||S_n||>t\}\right).$$

Und wegen $||S_n||\xrightarrow{w}||S||$ folgt für alle $t\in\mathbb{R}$, in denen die Verteilungsfunktion $F_{||S||}$ von $||S||$ stetig ist

$$P\left(\left\{\sup_{n\in\mathbb{N}}||S_n||>t\right\}\right)=\lim_{n\to\infty}P\left(\left\{\max_{1\leq k\leq n}||S_k||>t\right\}\right)\leq 2\limsup_{n\to\infty}P\left(\{||S_n||>t\}\right)=2P\left(\{||S||>t\}\right).$$

□

Mit einer ähnlichen Beweismethode wie im Beweis von Satz 3.3 erhalten wir für nicht-symmetrische Radon-Zufallsvariablen die folgende, auf Giuseppe Ottaviani und Anatoli Skorohod zurückgehende, Maximal-Ungleichung.

3.5 Satz. (Maximal-Ungleichung von Ottaviani-Skorohod)
Sei $N\in\mathbb{N}$ und seien $X_1,...,X_N\in\mathcal{L}_0(E)$ unabhängig. Setze

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

Dann gilt für alle $s,t>0$

$$P(\{\max_{1\leq k\leq N}||S_k||>s+t\})\leq\frac{P(\{||S_N||>t\})}{1-\max_{1\leq k\leq N}P(\{||S_N-S_k||>s\})}\,. \tag{3.5}$$

Beweis.
Setze
$$T(\omega) := \inf\{k \leq N : ||S_k(\omega)|| > s+t\} \in [0,\infty], \quad \omega \in \Omega.$$
Dann ist T messbar und es gilt $\{T = k\} \in \sigma(X_1, ..., X_k)$. Weiter gilt
$$\sum_{k=1}^{N} P\left(\{T = k\}\right) = P\left(\left\{\max_{1\leq k\leq N} ||S_k|| > s+t\right\}\right).$$
Für $\omega \in \Omega$ mit $T(\omega) = k$ und $||S_N(\omega) - S_k(\omega)|| \leq s$ gilt zudem $||S_N(\omega)|| > t$, denn mit der umgekehrten Dreiecksungleichung erhält man in diesem Fall
$$s \geq ||S_N(\omega) - S_k(\omega)|| \geq \Big|||S_N(\omega)|| - ||S_k(\omega)||\Big| > (s+t) - ||S_N(\omega)||.$$
Die Unabhängigkeit von $\sigma(X_1, ..., X_k)$ und $\sigma(X_{k+1}, ..., X_N)$ für alle $k \in \{1, ..., N\}$ liefert schließlich
$$\begin{aligned} P\left(\{||S_N|| > t\}\right) &\geq \sum_{k=1}^{N} P\left(\{T = k, ||S_N|| > t\}\right) \\ &\geq \sum_{k=1}^{N} P\left(\{T = k, ||S_N - S_k|| \leq s\}\right) \\ &\geq \min_{1\leq k\leq N} P\left(\{||S_N - S_k|| \leq s\}\right) \sum_{k=1}^{N} P\left(\{T = k\}\right). \end{aligned}$$
Umstellen und Beachten von
$$\min_{1\leq k\leq N} P\left(\{||S_N - S_k|| \leq s\}\right) = 1 - \max_{1\leq k\leq N} P\left(\{||S_N - S_k|| > s\}\right)$$
liefert nun die Behauptung. □

3.2 Der Satz von Itô-Nisio

Wir wollen nun die bisher erarbeitete Technik zur Untersuchung der Konvergenz zufälliger Reihen in Banachräumen anwenden. Die Beweise in diesem Abschnitt orientieren sich an [10], [12], [13], [9] und [7].

Notation und Konventionen
Für eine Folge von E-wertigen Radon-Zufallsvariablen $(X_n)_{n\in\mathbb{N}}$ setze für $n \in \mathbb{N}$
$$S_n := \sum_{i=1}^{n} X_i, \quad \mu_n := P^{S_n}.$$

3.6 Satz. (Satz von Itô-Nisio)
Für eine Folge $(X_n)_{n\in\mathbb{N}}$ unabhängiger E-wertiger Radon-Zufallsvariablen 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 $(i) \Rightarrow (ii) \Rightarrow (iii)$ gelten nach Korollar 2.7 und Proposition 2.12. 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 verteilt.
Für $n, N \in \mathbb{N}$ mit $n < N$ setze

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

Seien $\varepsilon, t > 0$. Da $(S_n)_{n\in\mathbb{N}}$ nach Voraussetzung stochastisch konvergiert, existiert nach Satz 2.9 ein $n_0 := n_0(\varepsilon, t) \in \mathbb{N}$, sodass für alle $N > n \geq n_0$

$$P(\{||S_N - S_n|| > t\}) \leq \frac{\varepsilon}{2}.$$

Zusammen mit Lévys Maximal-Ungleichung (3.1) erhalten wir somit für alle $N > n \geq n_0$

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

Daraus folgt

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

für $n \geq n_0$. Demnach gilt $Y_n \xrightarrow{st} 0$. Nach Satz 2.4 konvergiert $(S_n)_{n\in\mathbb{N}}$ daher fast sicher.
Fall B: Allgemeiner Fall.
Wir verwenden für diesen Fall 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 nach Konstruktion

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

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

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

Mit dem Satz von Fubini erhalten wir

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

wobei $\Omega^*(\omega_2) = \{\omega_1 \in \Omega : (\omega_1, \omega_2) \in \Omega^*\}$. Somit existiert 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 gilt daher

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

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

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

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

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

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 $(\mathcal{M}(E), \rho)$ und somit nach Satz 1.27 gleichmäßig straff. Folglich 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

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

Somit ist auch $\{\mu_{m,n} : m, n \in \mathbb{N}, m < n\}$ gleichmäßig straff und nach Satz 1.27 relativ kompakt in $(\mathcal{M}(E), \rho)$. Wir zeigen nun

$$\forall \varepsilon > 0\ \exists N \in \mathbb{N}:\ \forall n > m \geq N: \quad P\left(\{||S_n - S_m|| < \varepsilon\}\right) = \mu_{m,n}(B(0, \varepsilon)) > 1 - \varepsilon. \tag{3.6}$$

Nach Satz 2.9 folgt daraus die stochastische Konvergenz der Folge $(S_n)_{n\in\mathbb{N}}$. Angenommen (3.6) 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 schwach 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 Satz 1.20

$$\nu(B(0, \varepsilon)) \leq \liminf_{N\to\infty} \mu_{m(N),n(N)}(B(0, \varepsilon)) \leq 1 - \varepsilon. \tag{3.7}$$

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

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

Mit Grenzübergang $N \to \infty$ folgt aus Satz 2.17 wegen $\mu_N \rightharpoonup \mu$ und $\mu_{m(N),n(N)} \rightharpoonup \nu$

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

Aufgrund von $\widehat{\mu}(0_{E'}) = 1$ und der Stetigkeit von $\widehat{\mu}$ existiert zudem ein $r > 0$ mit

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

Also gilt

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

Aus Proposition 2.20 folgt nun $\nu = \delta_0$. Im Widerspruch zu (3.7). Es gilt folglich (3.6) und $(S_n)_{n\in\mathbb{N}}$ konvergiert demnach stochastisch. □

3.7 Bemerkung.
Mittels der Maximal-Ungleichung (3.5) von Ottaviani-Skorohod ist auch ein Beweis der Implikation $(ii) \Rightarrow (i)$ ohne Fallunterscheidung möglich. Man betrachtet hierzu etwa die Ereignisse

$$A_N := \bigcap_{m\in\mathbb{N}} \{ \sup_{n>m} ||S_n - S_m|| > \frac{1}{N}\}, \quad N \in \mathbb{N}.$$

Dann ist $A := (\cup_{N=1}^{\infty} A_N)^c$ das Ereignis, dass $(S_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge ist und man zeigt

$$\forall N \in \mathbb{N}: \quad P(A_N) = 0.$$

Ein Beweis mit dieser Vorgehensweise für den skalaren Fall findet sich etwa [3, Theorem 14.2]. Der allgemeine Fall funktioniert vollkommen analog.

Unter der zusätzlichen Voraussetzung, dass die unabhängigen Zufallsvariablen $(X_n)_{n\in\mathbb{N}}$ symmetrisch verteilt sind, lässt sich der Satz von Itô-Nisio auf drei noch schwächere Annahmen als die Verteilungskonvergenz von $(S_n)_{n\in\mathbb{N}}$ erweitern. Für den Beweis dient uns unter anderem der folgende Satz.

3.8 Satz.
Sei $(X_n)_{n\in\mathbb{N}}$ eine Folge unabhängiger E-wertiger Radon-Zufallsvariablen und 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 für $n \in \mathbb{N}$

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

sowie

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

Nach Konstruktion sind die Zufallsvariablen $\widetilde{S}_n$, $\widetilde{T}_n$ und S_n für festes $n \in \mathbb{N}$ identisch verteilt. 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 $\tilde{K} \subseteq E$ mit

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

Wegen der Stetigkeit der Abbildung

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

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

$$\begin{aligned}
\mu_{U_n}(K) = (P \times P)\{\widetilde{S}_n - \widetilde{T}_n \in K\} &\geq (P \times P)\{\widetilde{S}_n \in \tilde{K}, \widetilde{T}_n \in \tilde{K}\} \\
&\geq 1 - (P \times P)\{\widetilde{S}_n \notin \tilde{K}\} - (P \times P)\{\widetilde{T}_n \notin \tilde{K}\} \\
&= 1 - 2\mu_n(\tilde{K}^c) \\
&\geq 1 - 2\varepsilon.
\end{aligned}$$

Demnach ist $\{\mu_{U_n} :\ n \in \mathbb{N}\}$ straff. Als nächstes zeigen wir, dass $(\widehat{\mu_{U_n}}(f))_{n\in\mathbb{N}}$ für alle $f \in E'$ konvergiert. Sei dazu $f \in E'$ beliebig, aber fest. Die Unabhängigkeit von $(\widetilde{Y}_n)_{n\in\mathbb{N}}$ und $(\widetilde{X}_n)_{n\in\mathbb{N}}$ liefert direkt die Unabhängigkeit von $(\widetilde{X}_n - \widetilde{Y}_n)_{n\in\mathbb{N}}$ und nach Konstruktion sind $\tilde{Y}_n$ und $\tilde{X}_n$ für $n \in \mathbb{N}$ identisch verteilt. Es gilt demzufolge

$$\begin{aligned}
\widehat{\mu_{U_n}}(f) = E(e^{if(U_n)}) &= \mathbb{E}\Big(\prod_{j=1}^{n}(e^{if(\widetilde{X}_j - \widetilde{Y}_j)})\Big) \\
&= \prod_{j=1}^{n}\mathbb{E}(e^{if(\widetilde{X}_j - \widetilde{Y}_j)}) = \prod_{j=1}^{n}\mathbb{E}(e^{if(\widetilde{X}_j)})\mathbb{E}(e^{-if(\widetilde{Y}_j)}) = \prod_{j=1}^{n}\left|\mathbb{E}(e^{if(\widetilde{X}_j)})\right|^2.
\end{aligned}$$

Wegen $0 \leq \left|\mathbb{E}(e^{if(\widetilde{X}_j)})\right| \leq 1$ für alle $j \in \mathbb{N}$ folgt daraus die Konvergenz von $(\widehat{\mu_{U_n}}(f))_{n\in\mathbb{N}}$. Wir zeigen nun, dass die Folge $(\mu_{U_n})_{n\in\mathbb{N}}$ schwach konvergiert. Wie zuvor gezeigt wurde, ist $\{\mu_{U_n} : n \in \mathbb{N}\}$ gleichmäßig straff, also nach Satz 1.27 relativ kompakt in $(\mathcal{M}(E), \rho)$. Es genügt also zu zeigen, dass alle konvergenten Teilfolgen von $(\mu_{U_n})_{n\in\mathbb{N}}$ gegen den gleichen Grenzwert konvergieren. Seien dazu $(\nu_n)_{n\in\mathbb{N}}$ und $(\eta_n)_{n\in\mathbb{N}}$ zwei konvergente Teilfolgen von $(\mu_{U_n})_{n\in\mathbb{N}}$ mit schwachem Grenzwert ν bzw. η. Insbesondere gilt also nach Satz 2.17 für alle $f \in E'$

$$\lim_{n\to\infty}\widehat{\nu_n}(f) = \widehat{\nu}(f) \text{ und } \lim_{n\to\infty}\widehat{\eta_n}(f) = \widehat{\eta}(f).$$

Da für alle $f \in E'$ die Folge $(\widehat{\mu_{U_n}}(f))_{n\in\mathbb{N}}$ konvergiert, gilt

$$\forall f \in E' :\ \nu(f) = \lim_{n\to\infty}\widehat{\nu_n}(f) = \lim_{n\to\infty}\widehat{\mu_{U_n}}(f) = \lim_{n\to\infty}\widehat{\eta_n}(f) = \widehat{\eta}(f).$$

Nach Satz 2.16 folgt daraus $\eta = \nu$. Also ist die Folge $(\mu_{U_n})_{n\in\mathbb{N}}$ schwach konvergent, d.h. $(U_n)_{n\in\mathbb{N}}$ konvergiert in Verteilung. Aus Satz 3.6 folgt nun, dass $(U_n)_{n\in\mathbb{N}}$ fast sicher konvergiert. Daher existiert eine Menge $\Omega^* \in \mathcal{A} \otimes \mathcal{A}$ mit $(P \times P)(\Omega^*) = 1$ und

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

Mit dem Satz von Fubini erhalten wir wie im Beweis von Satz 3.6 $(ii) \Rightarrow (i)$ ein $\omega' \in \Omega$, sodass $(S_n - S_n(\omega'))_{n\in\mathbb{N}}$ fast sicher konvergiert. Die Folge $(c_n)_{n\in\mathbb{N}}$, definiert durch $c_n := S_n(\omega')$, $n \in \mathbb{N}$, erfüllt nun die gewünschte Eigenschaft. $\square$

3.9 Satz. (Satz von Itô-Nisio für Folgen symmetrischer Radon-Zufallsvariablen)
Sei $(X_n)_{n\in\mathbb{N}}$ eine Folge unabhängiger und symmetrischer E-wertiger Radon-Zufallsvariablen. 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 Radon-Zufallsvariable $S \in \mathcal{L}_0(E)$, sodass

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

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

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

Beweis.
Die Äquivalenz $(i) \iff (ii) \iff (iii)$ folgt aus Satz 3.6. Die Implikation $(iii) \Rightarrow (iv)$ gilt nach Satz 1.27. Ferner erhalten wir $(i) \Rightarrow (v)$ aus Proposition 2.12. Wir zeigen noch $(v) \Rightarrow (vi)$, $(vi) \Rightarrow (iv)$ und $(v) \Rightarrow (iv) \Rightarrow (i)$.
Zu $(v) \Rightarrow (vi)$: Setze $\mu := P^S$ und sei $f \in E'$. Wegen Satz 2.6 können wir ohne Einschränkung annehmen, dass $(f(S_n))_{n\in\mathbb{N}}$ fast sicher gegen $f(S)$ konvergiert. Der Satz von der dominierten Konvergenz liefert dann

$$\lim_{n\to\infty} \widehat{\mu_n}(f) = \lim_{n\to\infty} \mathbb{E}(e^{if(S_n)}) = \mathbb{E}(e^{f(S)}) = \widehat{\mu}(f).$$

Zu $(iv) \Rightarrow (i)$: Nach Satz 3.8 existiert eine Folge $(c_n)_{n\in\mathbb{N}}$ in E, sodass $(S_n - c_n)_{n\in\mathbb{N}}$ fast sicher konvergiert. Setze nun $P^X := P^{(X_1,X_2,\ldots)}$ und $P^{-X} := P^{(-X_1,-X_2,\ldots)}$. Dann sind P^X und P^{-X} jeweils Wahrscheinlichkeitsmaße auf dem Produktraum $(E^{\mathbb{N}}, \otimes_{n\in\mathbb{N}} B(E))$ und wegen der Unabhängigkeit und Symmetrie von $(X_n)_{n\in\mathbb{N}}$ erhalten wir direkt $P^X = P^{-X}$. Daher gilt für $N \in \mathbb{N}$ und $\varepsilon > 0$

$$\begin{aligned}
&P\left(\left\{\sup_{n\geq N} ||(S_n - c_n) - (S_N - c_N)|| > \varepsilon\right\}\right) \\
&= P^X\big(\{(x_n)_{n\in\mathbb{N}} \in E^{\mathbb{N}} : \sup_{n\geq N} ||x_n + x_{n-1} + \ldots + x_{N+1} + (c_N - c_n)|| > \varepsilon\}\big) \\
&= P^{-X}\big(\{(x_n)_{n\in\mathbb{N}} \in E^{\mathbb{N}} : \sup_{n\geq N} ||x_n + x_{n-1} + \ldots + x_{N+1} + (c_N - c_n)|| > \varepsilon\}\big) \\
&= P\left(\left\{\sup_{n\geq N} ||(-S_n - c_n) - (-S_N - c_N)|| > \varepsilon\right\}\right). \qquad (3.8)
\end{aligned}$$

Die Messbarkeit der Menge $\{(x_n)_{n\in\mathbb{N}} \in E^{\mathbb{N}} : \sup_{n\geq N} ||x_n + x_{n-1} + \ldots + x_{N+1} + (c_N - c_n)|| > \varepsilon\}$ folgt aus der Messbarkeit der Projektionen

$$\pi_k : E^{\mathbb{N}} \to E, \quad (x_n)_{n\in\mathbb{N}} \mapsto x_k, \quad k \in \mathbb{N},$$

der Messbarkeit der Addition in dem messbaren Vektorraum $(E, \mathcal{B}(E))$ und der Stetigkeit der Norm. Nach Satz 2.4 folgt aus (3.8) die fast sichere Konvergenz von $(-S_n - c_n)_{n\in\mathbb{N}}$. Daraus folgt nun die fast sichere Konvergenz von $(S_n)_{n\in\mathbb{N}}$, denn für $n \in \mathbb{N}$ gilt

$$S_n = \frac{1}{2}\big((S_n - c_n) - (-S_n - c_n)\big).$$

Zu $(v) \Rightarrow (iv)$: Wegen der Unabhängigkeit von $(X_n)_{n\in\mathbb{N}}$ sind für alle $f \in E'$ und $m \geq n$ die Zufallsvariablen $f(S_m - S_n)$ und $f(S_n)$ unabhängig. Nach Proposition 2.19 sind somit $f(S - S_n)$ und $f(S_n)$ für alle $n \in \mathbb{N}$ unabhängig. Zusammen mit der Symmetrie von S_n ergibt sich deshalb für $f \in E'$

$$\begin{aligned}\widehat{P^S}(f) = \mathbb{E}(e^{if(S)}) &= \mathbb{E}(e^{if(S-S_n)})\mathbb{E}(e^{if(S_n)}) \\ &= \mathbb{E}(e^{if(S-S_n)})\mathbb{E}(e^{if(-S_n)}) = \mathbb{E}(e^{if(S-2S_n)}) = \widehat{P^{S-2S_n}}(f).\end{aligned}$$

Folglich sind S und $S - 2S_n$ nach Satz 2.16 für alle $n \in \mathbb{N}$ identisch verteilt. Da P^S straff ist, existiert zu $\varepsilon > 0$ eine kompakte Menge $K \subseteq E$ mit $P(\{S \in K\}) \geq 1 - \varepsilon$. Aus Stetigkeitsgründen ist auch die Menge $L := \{\frac{1}{2}(x-y) : x, y \in K\}$ kompakt und es gilt für alle $n \in \mathbb{N}$

$$P(\{S_n \in L\}) \geq P(\{S \in K,\ S - 2S_n \in K\}) \geq 1 - P(\{S \notin K\}) - P(\{S - 2S_n \notin K\}) \geq 1 - 2\varepsilon.$$

Somit ist $\{\mu_n :\ n \in \mathbb{N}\}$ gleichmäßig straff.
Zu $(vi) \Rightarrow (iv)$: Sei $f \in E'$ beliebig, aber fest. Betrachte die Abbildungen

$$\begin{aligned}\varphi_n : \mathbb{R} \to \mathbb{C}, \quad t &\mapsto \int_E e^{itf(x)}\mu_n(dx) = \widehat{\mu_n}(tf), \quad n \in \mathbb{N}, \\ \varphi : \mathbb{R} \to \mathbb{C}, \quad t &\mapsto \int_E e^{itf(x)}\mu(dx) = \widehat{\mu}(tf).\end{aligned}$$

Dann ist φ_n für jedes $n \in \mathbb{N}$ die charakteristische Funktion von μ_n^f und φ die charakteristische Funktion von μ^f. Nach Voraussetzung konvergiert $(\varphi_n)_{n\in\mathbb{N}}$ punktweise gegen φ und nach dem Stetigkeitssatz von Lévy, vgl. [6, Satz 8.7.5], gilt daher $\mu_n^f \rightharpoonup \mu^f$. Für festes $f \in E'$ ist $\{\mu_n^f : n \in \mathbb{N}\}$ somit insbesondere relativ kompakt. Nach Satz 1.33 genügt es folglich zu zeigen, dass $\{\mu_n :\ n \in \mathbb{N}\}$ flach konzentriert ist. Da $\{\mu\}$ flach konzentriert ist, genügt es dafür zu zeigen, dass für jeden endlichdimensionalen Untervektorraum $F \subseteq E$, jedes $\varepsilon > 0$ und alle $n \in \mathbb{N}$ gilt

$$\mu_n(\{x \in E :\ \inf_{y\in F}||x - y|| > \varepsilon\}) \leq 2\mu(\{x \in E :\ \inf_{y\in F}||x - y|| > \varepsilon\}).$$

Man beachte hierbei, dass die Menge $\{x \in E :\ \inf_{y\in F}||x-y|| > \varepsilon\}$ als offene Menge insbesondere messbar ist. Sei also $F \subseteq E$ ein endlichdimensionaler Untervektorraum und $\varepsilon > 0$. Nach Korollar A.13 existiert eine Folge $(f_n)_{n\in\mathbb{N}}$ in E' mit

$$\forall x \in E : \quad d(x, F) := \inf_{y\in F}||x - y|| = \sup_{n\in\mathbb{N}} |f_n(x)|\,.$$

Sei zunächst $m \in \mathbb{N}$ festgewählt. Mittels charakteristischer Funktionale und dem Satz von der dominierten Konvergenz prüft man leicht, dass $\mu_n^{(f_1,\dots,f_m)} \rightharpoonup \mu^{(f_1,\dots,f_m)}$. Wegen der Linearität von $f_1, ..., f_m$ können wir Satz 3.6 auf die Folge $(T_n^{(m)})_{n\in\mathbb{N}} := ((f_1, ..., f_m) \circ S_n)_{n\in\mathbb{N}}$ anwenden und erhalten eine $\mathbb{R}^m$-wertige Zufallsvariable $T^{(m)}$ auf $(\Omega, \mathcal{A}, P)$ mit $T_n^{(m)} \xrightarrow{st} T^{(m)}$. Insbesondere gilt daher $T_n^{(m)} \xrightarrow{w} T^{(m)}$ und somit $P^{T^{(m)}} = \mu^{(f_1,\dots,f_m)}$. Ferner ist T_n wegen der Linearität von

$f_1, ..., f_m$ und der Symmetrie von S_n für alle $n \in \mathbb{N}$ symmetrisch . Mit Korollar 3.4, angewendet auf $(T_n^{(m)})_{n\in\mathbb{N}}$, erhalten wir im Banachraum $(\mathbb{R}^m, ||\cdot||_\infty)$ für fast alle $\varepsilon > 0$

$$\begin{aligned} P\left(\left\{\max_{1\leq i\leq m} |f_i(S_n)| > \varepsilon\right\}\right) &= P\left(\left\{||T_n^{(m)}||_\infty > \varepsilon\right\}\right) \\ &\leq 2P\left(\left\{||T^{(m)}||_\infty > \varepsilon\right\}\right) = 2\mu\big(\{x \in E : \max_{1\leq i\leq m} |f_i(x)| > \varepsilon\}\big). \end{aligned}$$

Mit der σ-Stetigkeit von Wahrscheinlichkeitsmaßen folgt schließlich

$$\begin{aligned} \mu_n(\{x \in E : \inf_{y\in F} ||x-y|| > \varepsilon\}) = P\left(\{d(S_n, F) > \varepsilon\}\right) &= \lim_{m\to\infty} P\left(\left\{\max_{1\leq i\leq m} |f_i(S_n)| > \varepsilon\right\}\right) \\ &\leq \lim_{m\to\infty} 2\mu(\{x \in E : \max_{1\leq i\leq m} |f_i(x)| > \varepsilon\}) \\ &= 2\mu(\{x \in E : \inf_{y\in F} ||x-y|| > \varepsilon\}). \end{aligned}$$

□

3.10 Bemerkung.
Wie das folgende Gegenbeispiel aus [10] zeigt, kann auf die Annahme der Symmetrie in Satz 3.9 im Allgemeinen nicht verzichtet werden. Sei etwa $(E, \langle\cdot,\cdot\rangle)$ ein Hilbertraum mit Orthonormalbasis $(e_n)_{n\in\mathbb{N}}$. Man bemerke zunächst, dass nach dem Darstellungsatz von Fréchet-Riesz, vgl. [19, Theorem V.3.6], für jedes $f \in E'$ ein $z \in E$ mit $f = \langle\cdot, z\rangle$ existiert. Setze nun

$$X_1(\omega) = e_1, \quad X_n(\omega) = e_n - e_{n-1},\ n \geq 2, \quad \omega \in \Omega.$$

Dann gilt offensichtlich $S_n(\omega) = e_n$ für alle $\omega \in \Omega$ und da $(e_n)_{n\in\mathbb{N}}$ eine Orthonormalbasis von E ist, gilt für alle $z \in E$

$$\lim_{n\to\infty} \langle z, S_n\rangle = 0 = \langle z, S\rangle,$$

wobei $S(\omega) = 0$ für alle $\omega \in \Omega$. Nach dem Satz von Fréchet-Riesz gilt also

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

Wegen $||S_n|| = ||e_n|| = 1$ für alle $n \in \mathbb{N}$ konvergiert $(S_n)_{n\in\mathbb{N}}$ aber weder fast-sicher noch stochastisch gegen S. ◊

Für Folgen $(a_n)_{n\in\mathbb{N}}$ in $[0, \infty)$ kennt man aus der reellen Analysis die Äquivalenz

$$\left(\sum_{i=1}^n a_i\right)_{n\in\mathbb{N}} \text{ konvergiert.} \iff \left(\sum_{i=1}^n a_i\right)_{n\in\mathbb{N}} \text{ ist beschränkt.}$$

Mit Hilfe des Satzes von Itô-Nisio können wir nun ein ähnliches Resultat für symmetrische $\mathbb{R}^d$-wertige Zufallsvariablen formulieren.

3.11 Definition.
Eine Folge $(X_n)_{n\in\mathbb{N}}$ in $\mathcal{L}_0(E)$ heißt *stochastisch beschränkt*, falls

$$\forall\varepsilon > 0\ \exists R > 0 : \quad \sup_{n\in\mathbb{N}} P\left(\{||X_n|| > R\}\right) < \varepsilon.$$

3.12 Lemma.
Sei $(X_n)_{n\in\mathbb{N}}$ eine Folge E-wertiger Radon-Zufallsvariablen und $X \in \mathcal{L}_0(E)$ mit $X_n \xrightarrow{st} X$. Dann ist $(X_n)_{n\in\mathbb{N}}$ stochastisch beschränkt.

Beweis.
Sei $\varepsilon > 0$. Dann existiert nach Satz 2.8 ein $N \in \mathbb{N}$ mit

$$\sup_{n\geq N} P\left(\{||X_n - X|| > \varepsilon\}\right) < \frac{\varepsilon}{2}.$$

Wie man leicht einsieht, existiert ferner ein $R_0 > 0$ mit

$$P\left(\{||X|| > R_0\}\right) < \frac{\varepsilon}{2} \quad \text{und} \quad \max_{1\leq n\leq N-1} P\left(\{||X_n|| > R_0\}\right) < \varepsilon.$$

Für $R := \max\{R_0, \varepsilon\}$ gilt also

$$\begin{aligned}
\sup_{n\geq N} P\left(\{||X_n|| > R\}\right) &\leq \sup_{n\geq N} P\left(\{||X_n - X|| + ||X|| > R\}\right) \\
&\leq \sup_{n\geq N} P\left(\left\{||X_n - X|| > \frac{R}{2}\right\}\right) + P\left(\left\{||X|| > \frac{R}{2}\right\}\right) \\
&< \sup_{n\geq N} P\left(\left\{||X_n - X|| > \frac{R}{2}\right\}\right) + \frac{\varepsilon}{2} < \varepsilon.
\end{aligned}$$

Zudem gilt nach Konstruktion

$$\max_{1\leq n\leq N-1} P\left(\{||X_n|| > R\}\right) < \varepsilon.$$

Insgesamt erhalten wir somit

$$\sup_{n\in\mathbb{N}} P\left(\{||X_n|| > R\}\right) = \max\left(\max_{1\leq n\leq N-1} P\left(\{||X_n|| > R\}\right), \sup_{n\geq N} P\left(\{||X_n|| > R\}\right)\right) < \varepsilon.$$

□

3.13 Korollar.
Sei $d \in \mathbb{N}$ und $(X_n)_{n\in\mathbb{N}}$ eine unabhängige und symmetrische Folge $\mathbb{R}^d$-wertiger Zufallsvariablen. Dann sind äquivalent:

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

(ii) $(S_n)_{n\in\mathbb{N}}$ ist stochastisch beschränkt.

Beweis.
Zu $(i) \Rightarrow (ii)$ Nach Korollar 2.7 konvergiert $(S_n)_{n\in\mathbb{N}}$ insbesondere stochastisch und nach Lemma 3.12 ist $(S_n)_{n\in\mathbb{N}}$ folglich stochastisch beschränkt.
Zu $(ii) \Rightarrow (i)$: Aus (ii) erhalten wir unmittelbar

$$\forall \varepsilon > 0 \; \exists R > 0 : \quad \inf_{n\in\mathbb{N}} P\left(\left\{S_n \in \overline{B}(0, R)\right\}\right) \geq 1 - \varepsilon.$$

Da abgeschlossene und beschränkte Teilmengen des $\mathbb{R}^d$ nach dem Satz von Heine-Borel kompakt sind, folgt daraus die gleichmäßige Straffheit von $(\mu_n)_{n\in\mathbb{N}}$. Nach Satz 3.9 konvergiert $(S_n)_{n\in\mathbb{N}}$ also fast sicher. □

Wir wollen nun noch eine zweite Anwendung von Satz 3.9 geben, welche in der Literatur teilweise als *qualitative Version des Kontraktions-Prinzips* bezeichnet wird. Die Darstellung orientiert sich hierbei an [13]. Im Folgenden sei $(X_n)_{n\in\mathbb{N}}$ eine unabhängige Folge symmetrischer Radon-Zufallsvariablen aus $\mathcal{L}_0(E)$ und $(\lambda_n)_{n\in\mathbb{N}}$ eine beschränkte Folge in $\mathbb{R}$. Für $n \in \mathbb{N}$ setze

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

Aus technischen Gründen seien ferner $S_0 = T_0 = 0$. Unser Ziel ist es, zu zeigen, dass die fast sichere Konvergenz von $(S_n)_{n\in\mathbb{N}}$ auch die fast sichere Konvergenz von $(T_n)_{n\in\mathbb{N}}$ impliziert. Der Beweis beruht hauptsächlich auf der folgenden Abschätzung.

3.14 Lemma.
Für alle $t > 0$ und $N \in \mathbb{N}$ gilt

$$P\left(\{||T_N|| > t\}\right) \leq 2P\left(\{||S_N|| > t\}\right). \tag{3.9}$$

Beweis.
Wegen der Symmetrie von X_n sind $\lambda_n X_n$ und $|\lambda_n|\, X_n$ für $n \in \mathbb{N}$ identisch verteilt. Ferner ist $(\lambda_n)_{n\in\mathbb{N}}$ nach Voraussetzung beschränkt. Wir können also ohne Einschränkung annehmen, dass $0 \leq \lambda_n \leq 1$ für alle $n \in \mathbb{N}$.
<u>Fall A</u>: $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_N$.
Wie man leicht nachrechnet, gilt

$$T_N = \sum_{n=1}^{N} \lambda_n (S_n - S_{n-1}) = \sum_{n=1}^{N-1} (\lambda_n - \lambda_{n+1}) S_n + \lambda_N S_N.$$

Mit der Dreiecksungleichung erhalten wir somit

$$\begin{aligned} ||T_N|| &\leq \sum_{n=1}^{N-1} (\lambda_n - \lambda_{n+1}) ||S_n|| + \lambda_N ||S_N|| \\ &\leq \max_{1\leq n\leq N} ||S_n|| \left(\sum_{n=1}^{N-1} (\lambda_n - \lambda_{n+1}) + \lambda_N \right) \\ &= \max_{1\leq n\leq N} ||S_n|| \lambda_1 \leq \max_{1\leq n\leq N} ||S_n||. \end{aligned}$$

Mit Lévys Maximalungleichung (3.1) folgt daraus

$$P\left(\{||T_N|| > t\}\right) \leq P\left(\left\{ \max_{1\leq n\leq N} ||S_n|| > t \right\}\right) \leq 2P\left(\{||S_N|| > t\}\right).$$

<u>Fall B</u>: Allgemeiner Fall.
Mittels einer Permutation σ erhält man $\lambda_{\sigma(1)} \geq ... \geq \lambda_{\sigma(N)}$. Man beachte schließlich, dass

$$\sum_{n=1}^{N} \lambda_{\sigma(n)} X_{\sigma(n)} = T_N \quad \text{und} \quad \sum_{n=1}^{N} X_{\sigma(n)} = S_N.$$

Aus Fall A folgt nun die Behauptung. □

3.15 Satz. (Kontraktions-Prinzip, Qualitative Version)
Falls $(S_n)_{n\in\mathbb{N}}$ fast sicher konvergiert, dann konvergiert auch $(T_n)_{n\in\mathbb{N}}$ fast sicher.

Beweis.
Sei $\varepsilon > 0$. Nach Lemma 3.13 gilt für $m < n$

$$P\left(\{||T_n - T_m|| > \varepsilon\}\right) \leq 2P\left(\{||S_n - S_m|| > \varepsilon\}\right). \tag{3.10}$$

Da $(S_n)_{n\in\mathbb{N}}$ fast sicher konvergiert, gilt nach Korollar 2.7 und Satz 2.9

$$\lim_{N\to\infty} \sup_{n\geq N} P\left(\{||S_n - S_N|| > \varepsilon\}\right) = 0.$$

Wegen (3.10) folgt also

$$\lim_{N\to\infty} \sup_{n\geq N} P\left(\{||T_n - T_N|| > \varepsilon\}\right) = 0.$$

Da ε beliebig gewählt war, konvergiert $(T_n)_{n\in\mathbb{N}}$ nach Satz 2.9 stochastisch und nach Satz 3.9 insbesondere fast sicher. □

3.16 Bemerkung.
In Satz 3.15 kann nicht auf die Symmetrie der Radon-Zufallsvariablen $X_1, X_2, ...$ verzichtet werden. Betrachte dazu etwa die Folge $(X_n)_{n\in\mathbb{N}}$, definiert durch

$$X_n(\omega) := (-1)^n \frac{1}{n}, \quad \omega \in \Omega,\ n \in \mathbb{N},$$

und die beschränkte Folge $((-1)^n)_{n\in\mathbb{N}}$. Dann ist die Folge $(\sum_{k=1}^n X_n)_{n\in\mathbb{N}}$ fast sicher konvergent, aber die Folge $(\sum_{k=1}^n \lambda_k X_k)_{n\in\mathbb{N}}$ divergiert fast sicher. ◊

A Funktionalanalytische und topologische Hilfsmittel

A.1 Separabilität

In diesem Abschnitt wollen wir einige nützliche Fakten über separable Vektorräume sammeln, die vor allem in Abschnitt 1.6 Anwendung finden. Wir beginnen zunächst mit ein paar grundlegenden Definitionen. Sei dazu $(X, ||\cdot||)$ ein normierter Vektorraum.

A.1 Definition.
Eine Teilmenge $U \subseteq X$ heißt *separabel*, falls eine abzählbare Menge $A \subseteq U$ mit $\overline{A} = U$ existiert.

A.2 Definition.
Sei $A \subseteq X$. Dann heißt

$$lin(A) := \left\{ \sum_{i=1}^{n} \alpha_i x_i \mid n \in \mathbb{N}, \alpha_1, ..., \alpha_n \in \mathbb{R}, x_1, ..., x_n \in A \right\}$$

die *lineare Hülle von* A.

Eines der für diese Arbeit wichtigsten Resultate ist die folgende Charakterisierung separabler normierter Räume. Den Beweis findet man etwa in [19, Lemma I.2.9].

A.3 Satz.
Es sind äquivalent:

(i) X ist separabel.

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

A.4 Korollar.
Sei $(E_n)_{n \in \mathbb{N}}$ eine Folge separabler Untervektorräume von X. Dann ist auch

$$\overline{lin(\bigcup_{n \in \mathbb{N}} E_n)}$$

ein separabler Untervektorraum von X.

Beweis.
Nach Definition existiert für alle $n \in \mathbb{N}$ eine abzählbare Teilmenge $A_n \subset E_n$ mit $\overline{A_n} = E_n$. Wir zeigen:

$$\overline{lin(\bigcup_{n \in \mathbb{N}} E_n)} = \overline{lin(\bigcup_{n \in \mathbb{N}} A_n)}.$$

Zu $\supseteq$: Klar, weil $\bigcup_{n \in \mathbb{N}} A_n \subseteq \bigcup_{n \in \mathbb{N}} E_n$.
Zu $\subseteq$: Nach der Definition des topologischen Abschlusses genügt es zu zeigen, dass

$$lin(\bigcup_{n \in \mathbb{N}} E_n) \subseteq \overline{lin(\bigcup_{n \in \mathbb{N}} A_n)}.$$

Sei also $y \in lin(\bigcup_{n\in\mathbb{N}} E_n)$. Dann existiert ein $N \in \mathbb{N}$, $\lambda_1, ..., \lambda_N \in \mathbb{R}$ und $y_1, ..., y_N \in \bigcup_{n\in\mathbb{N}} E_n$ mit

$$y = \sum_{k=1}^{N} \lambda_k y_k.$$

Durch eventuelles Umnummerieren können wir ohne Einschränkung annehmen, dass $y_k \in E_k$ für $k = 1, ..., N$. Da A_n für $n \in \mathbb{N}$ dicht in E_n liegt, existiert für $k \in \{1, ..., N\}$ eine Folge $(x_m^{(k)})_{m\in\mathbb{N}}$ in A_k mit $\lim_{m\to\infty} x_m^{(k)} = y_k$. Also ist die Folge $(\sum_{k=1}^{N} \lambda_k x_m^{(k)})_{m\in\mathbb{N}}$ in $lin(\bigcup_{n\in\mathbb{N}} A_n)$ und nach den Rechenregeln für Grenzwerte gilt

$$\lim_{m\to\infty} \left(\sum_{k=1}^{N} \lambda_k x_m^{(k)} \right) = \sum_{k=1}^{N} \lambda_k y_k = y.$$

Also gilt $y \in \overline{lin(\bigcup_{n\in\mathbb{N}} A_n)}$. Insgesamt folgt also $\overline{lin(\bigcup_{n\in\mathbb{N}} E_n)} = \overline{lin(\bigcup_{n\in\mathbb{N}} A_n)}$ und da $\bigcup_{n\in\mathbb{N}} A_n$ abzählbar ist, liefert Satz $A.3$ die Behauptung. □

A.5 Korollar.
Sei $E \subseteq X$ eine separable Teilmenge von X. Dann ist auch $\overline{lin(E)}$ separabel.

Beweis.
Nach Voraussetzung existiert eine dichte abzählbare Menge $A \subseteq E$. Wir zeigen

$$lin(E) \subseteq \overline{lin(A)}.$$

Zusammen mit $A \subseteq E$ und der Definition des topologischen Abschlusses einer Menge folgt daraus

$$\overline{lin(A)} \subseteq \overline{lin(E)} \subseteq \overline{lin(A)}$$

und somit die Behauptung. Sei also $x \in lin(E)$. Dann existieren $n \in \mathbb{N}, x_1, ..., x_n \in E$ und $\lambda_1, ..., \lambda_n \in \mathbb{R}$ mit

$$x = \sum_{i=1}^{n} \lambda_i x_i.$$

Da A dicht in E liegt, existiert für jedes $i \in \{1, ..., n\}$ eine Folge $(x_m^{(i)})_{m\in\mathbb{N}}$ in A mit $\lim_{m\to\infty} x_m^{(i)} = x_i$. Für alle $m \in \mathbb{N}$ gilt zudem $\sum_{i=1}^{n} \lambda_i x_m^{(i)} \in lin(A)$ und wegen

$$\lim_{m\to\infty} \sum_{i=1}^{n} \lambda_i x_m^{(i)} = \sum_{i=1}^{n} \lambda_i \lim_{m\to\infty} x_m^{(i)} = \sum_{i=1}^{n} \lambda_i x_i$$

folgt schließlich $x \in \overline{lin(A)}$. □

A.2 Der Satz von Hahn-Banach

Der Fortsetzungssatz von Satz von Hahn-Banach ist wohl einer der wichtigsten Sätze aus der Funktionalanalysis. Da in den Abschnitten 1.4, 1.5 und 3.2 einige Konsequenzen des Satzes benutzt werden, soll dieser Abschnitt als kurze Zusammenfassung der für uns nützlichsten Resultate dienen. Die ausgelassenen Beweise findet man beispielsweise in [19, Kapitel III.1].

Notation und Konventionen
Für einen normierten Vektorraum $(X, ||\cdot||)$ sei $(X', ||\cdot||_{op})$ der zugehörige (topologische) Dualraum. Ferner bezeichne

$$B_{X'} := \{f \in X' \;:\; ||f||_{op} \leq 1\}$$

die abgeschlossene Einheitskugel in X'.

A.6 Definition.
Sei X ein Vektorraum. Eine Abbildung $p : X \to \mathbb{R}$ heißt *sublinear*, falls

(a) $p(\lambda x) = \lambda p(x)$ für alle $\lambda \geq 0, x \in X$,

(b) $p(x + y) \leq p(x) + p(y)$ für alle $x, y \in X$.

A.7 Proposition.
Sei X ein normierter Vektorraum und $F \subseteq X$ ein abgeschlossener Untervektorraum. Dann ist die Abbildung

$$d(\cdot, F) : X \to \mathbb{R}, \quad y \mapsto d(y, F) := \inf_{x \in F} ||x - y||$$

sublinear.

Beweis.
Da F ein Untervektorraum von X ist, gilt für $\lambda \geq 0$ und $y \in X$

$$d(\lambda y, F) = \inf_{x \in F} ||\lambda y - x|| = \inf_{x \in F} ||\lambda y - \lambda x|| = \lambda \inf_{x \in F} ||y - x|| = \lambda d(y, F).$$

Ferner liefert die Dreiecksungleichung für $y, z \in X$

$$d(y+z, F) = \inf_{x \in F} ||y+z-x|| \leq \inf_{x \in F} (||y-\frac{x}{2}||+||z-\frac{x}{2}||) \leq \inf_{x \in F} ||y-x||+\inf_{x \in F} ||z-x|| = d(y, F)+d(z, F).$$

Also ist $d(\cdot, F)$ sublinear. $\square$

A.8 Satz. (Hahn-Banach, Version der Linearen Algebra)
Sei X ein Vektorraum und U ein Untervektorraum von X. Ferner seien $p : X \to \mathbb{R}$ sublinear und $f : U \to \mathbb{R}$ linear mit

$$\forall x \in U : \quad f(x) \leq p(x).$$

Dann existiert eine lineare Fortsetzung $F : X \to \mathbb{R}$, mit $F|_U = f$ und

$$\forall x \in X : \quad F(x) \leq p(x).$$

A.9 Satz. (Hahn-Banach, Fortsetzungsversion)
Sei X ein normierter Raum und U ein Untervektorraum. Zu jedem stetigen linearen Funktional $f : U \to \mathbb{R}$ existiert ein stetiges lineares Funktional $F : X \to \mathbb{R}$ mit

$$F|_U = f \text{ und } ||F||_{op} = ||f||_{op}.$$

Jedes stetige Funktional kann also normgleich fortgesetzt werden.

A.10 Korollar.
In jedem normierten Raum X existiert zu jedem $x \in X$, $x \neq 0$, ein Funktional $f_x \in X'$ mit $||f_x||_{op} = 1$ und $f_x(x) = ||x||$. Speziell trennt X' die Punkte von X, d.h.

$$\forall x_1, x_2 \in X \text{ mit } x_1 \neq x_2 \; \exists f \in X' : \quad f(x_1) \neq f(x_2).$$

A.11 Korollar.
In jedem normierten Raum X gilt

$$\forall x \in X: \quad ||x|| = \sup_{f \in B_{X'}} |f(x)|.$$

A.12 Korollar.
Falls X separabel ist, so existiert eine abzählbare Menge $D \subseteq B_{X'}$ mit

$$\forall x \in X: \quad ||x|| = \sup_{f \in D} |f(x)|.$$

Beweis.
Sei $E \subseteq X$ eine abzählbare dichte Teilmenge von X. Nach Korollar $A.10$ existiert für jedes $x \in E$ ein $f_x \in X'$ mit $||f_x||_{op} = 1$ und $f_x(x) = ||x||$. Setze also $D := \{f_x : x \in E\}$. Sei nun $x \in X$ beliebig und $\varepsilon > 0$. Dann gilt zunächst

$$\forall n \in \mathbb{N}: \quad |f_n(x)| \leq ||f_n||_{op}||x|| = ||x||.$$

Also

$$\sup_{f \in D} |f(x)| \leq ||x||.$$

Ferner existiert eine Folge $(y_n)_{n \in \mathbb{N}}$ in E mit $\lim_{n \to \infty} ||x - y_n|| = 0$. Insbesondere existiert $N \in \mathbb{N}$, sodass für alle $n \geq N$

$$||y_n - x|| < \frac{\varepsilon}{2}. \tag{A.1}$$

Zudem gilt für alle $n \geq N$

$$|f_{y_n}(y_n)| \leq |f_{y_n}(y_n) - f_{y_n}(x)| + |f_{y_n}(x)| \leq \frac{\varepsilon}{2} + |f_{y_n}(x)|. \tag{A.2}$$

Insgesamt ergibt sich also wegen $(A.1)$ und $(A.2)$ für $n \geq N$

$$||x|| \leq ||x - y_n|| + ||y_n|| \leq \frac{\varepsilon}{2} + |f_{y_n}(y_n)| \leq |f_{y_n}(x)| + \varepsilon.$$

Da $\varepsilon > 0$ und $x \in X$ beliebig gewählt waren, gilt also

$$\forall x \in X: \ ||x|| = \sup_{f \in D} |f(x)|.$$

A.13 Korollar.
Sei X separabel und $F \subseteq X$ ein abgeschlossener Untervektorraum. Dann existiert eine Folge $(f_n)_{n \in \mathbb{N}}$ in X' mit $||f_n||_{op} \leq 1$ für alle $n \in \mathbb{N}$ und

$$\forall x \in X: \quad d(x, F) = \inf_{y \in F} ||x - y|| = \sup_{n \in \mathbb{N}} |f_n(x)|.$$

Beweis.
Da X separabel ist, existiert eine abzählbare dichte Teilmenge $D = \{x_1, x_2, ...\} \subseteq X$. Sei nun $n \in \mathbb{N}$. Dann gilt $lin(\{x_n\}) = \{\lambda x_n : \lambda \in \mathbb{R}\}$ und die Abbildung

$$g_n : lin(\{x_n\}) \to \mathbb{R}, \quad \lambda x_n \mapsto \lambda d(x_n, F)$$

ist, wie man leicht nachrechnet, linear. Nach Proposition $A.7$ ist ferner die Abbildung

$$d(\cdot, F) : X \to \mathbb{R}, \quad y \mapsto d(y, F) := \inf_{x \in F} ||x - y||$$

sublinear. Für $z = \lambda x_n \in lin(\{x_n\})$ gilt zudem

$$g_n(z) = \lambda d(x_n, F) \leq |\lambda| \inf_{y \in F} ||x - y|| = \inf_{y \in F} ||\lambda x - \lambda y|| = d(z, F).$$

Also existiert nach Satz $A.8$ für alle $n \in \mathbb{N}$ ein Funktional $f_n \in X'$ mit $f_n|_{lin(\{x_n\})} = g_n$ und

$$\forall y \in X : \quad f_n(y) \leq d(y, F) = \inf_{x \in F} ||x - y|| \leq ||y||.$$

Somit gilt insbesondere $||f_n||_{op} \leq 1$. Betrachte nun die so gewonnene Folge $(f_n)_{n \in \mathbb{N}}$ in X'. Für $x \in X$ gilt dann nach Konstruktion

$$\forall n \in \mathbb{N} : \quad f_n(x) \leq d(x, F).$$

Daraus folgt direkt

$$\sup_{n \in \mathbb{N}} |f_n(x)| \leq d(x, F).$$

Sei nun $\varepsilon > 0$. Da D dicht in X liegt, gibt es ein $n \in \mathbb{N}$, sodass $||x_n - x|| < \frac{\varepsilon}{2}$. Daher gilt wegen $||f_n||_{op} \leq 1$

$$|f_n(x_n)| \leq |f_n(x_n) - f_n(y)| + |f_n(y)| \leq |f_n(y)| + \frac{\varepsilon}{2}.$$

Die Sublinearität von $d(\cdot, F)$ liefert nun

$$d(x, F) \leq d(x - x_n, F) + d(x_n, F) \leq ||x - x_n|| + |f_n(x_n)| \leq |f_n(y)| + \varepsilon.$$

Da $\varepsilon > 0$ beliebig gewählt war, folgt insgesamt

$$\sup_{n \in \mathbb{N}} |f_n(x)| = d(x, F).$$

$\square$

Literaturverzeichnis

[1] Acosta, A. D. de: *Existence and Convergence of Probability Measures in Banach Spaces.* Transactions of the American Mathematical Society, 152(1):273–298, November 1970.

[2] Amann, H. und Escher, J.: *Analysis I.* Birkhäuser, Basel, 3. Auflage, 2006.

[3] Bauer, H.: *Wahrscheinlichkeitstheorie.* Walter de Gruyter, Berlin – New York, 5. Auflage, 2002.

[4] Billingsley, P.: *Convergence of Probability Measures.* John Wiley and Sons, New York, 2. Auflage, 1999.

[5] Gaans, O. van: *Probability measures on metric spaces.* Vorlesungsskript, 2003. `https://www.math.leidenuniv.nl/~vangaans/jancol1.pdf`.

[6] Gänssler, P. und Stute, W.: *Wahrscheinlichkeitstheorie.* Springer-Verlag, Berlin, 1. Auflage, 1977.

[7] Heyer, H.: *Structural Aspects In The Theory of Probability.* World Scientific, 2. Auflage, 2009.

[8] Hytönen, T., Neerven, J. van, Veraar, M. und Weis, L.: *Analysis in Banach Spaces Volume I: Martingales and Littlewood-Paley Theory.* Springer-Verlag, Cham, 1. Auflage, 2016.

[9] Hytönen, T., Neerven, J. van, Veraar, M. und Weis, L.: *Analysis in Banach Spaces Volume II: Probabilistic Methods and Operator Theory.* Springer-Verlag, Cham, 1. Auflage, 2017.

[10] Itô, K. und Nisio, M.: *On the Convergence of Sums of Independent Banach Space Valued Random Variables.* Osaka Journal of Mathematics, 5(1):35–48, 1968.

[11] Kallenberg, O.: *Foundations of Modern Probability, Second Edition.* Springer-Verlag, New York – Berlin – Heidelberg, 2. Auflage, 2002.

[12] Ledoux, M. und Talagrand, M.: *Probability in Banach Spaces.* Springer-Verlag, Berlin, 1. Auflage, 1991.

[13] Li, D. und Queffélec, H.: *Introduction to Banach Spaces: Analysis and Probability Volume 1.* Cambridge University Press, Cambridge, 1. Auflage, 2018.

[14] Lévy, P.: *Théorie de l'addition des variables aléatoires.* Gauthier-Villars, Paris, 1. Auflage, 1937.

[15] Neerven, Jan van: *Stochastic Evolution Equations.* Vorlesungsskript, 2008. `https://fa.its.tudelft.nl/~neerven/publications/notes/ISEM.pdf`.

[16] Parthasarathy, K.R.: *Probability Measures on Metric Spaces.* Academic Press, New York – London, 1. Auflage, 1967.

[17] Preuß, G.: *Allgemeine Topologie.* Springer-Verlag, Berlin – Heidelberg, 2. Auflage, 1975.

[18] Vakhania, N., Tarieladze, V. und Chobanyan, S.: *Probability Distributions on Banach Spaces.* D. Reidel Publishing Company, Dordrecht, 1. Auflage, 1987.

[19] Werner, D.: *Funktionalanalysis.* Springer-Verlag, Berlin, 7. Auflage, 2007.

Eigenständigkeitserklärung

Hiermit bestätige ich, Jonas Köppl, dass ich die vorliegende Arbeit selbstständig und ohne unzulässige Hilfe verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt sowie die wörtlich und sinngemäß übernommenen Passagen aus anderen Werken kenntlich gemacht habe. Die Arbeit ist weder von mir noch von einer anderen Person an der Universität Passau oder an einer anderen Hochschule zur Erlangung eines akademischen Grades bereits eingereicht worden.

Ort, Datum:

Unterschrift: