

Stochastische Prozesse

SoSe 2018

Übersicht 1. Vorlesungswoche (9.-13.4.2018)

1 Maßtheoretische Grundlagen

1.1 Maßtheorie

Definition (σ -Algebra) Sei Ω eine nichtleere Menge. Eine σ -Algebra \mathcal{F} (über Ω) ist eine Menge von Teilmengen von Ω (d. h. $\mathcal{F} \subset \mathcal{P}(\Omega)$) mit:

- $\Omega \in \mathcal{F}$
- $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$ (wobei $A^c = \{\omega \in \Omega \mid \omega \notin A\} = \Omega \setminus A$)
- $A_1, A_2, \dots \in \mathcal{F} \Rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$

Bemerkungen: Für jede σ -Algebra \mathcal{F} über Ω gilt:

- 1) $\emptyset = \Omega^c \in \mathcal{F}$
- 2) $A_1, A_2, \dots \in \mathcal{F} \Rightarrow \bigcap_{n=1}^{\infty} A_n \in \mathcal{F}$
- 3) $A_1, \dots, A_n \in \mathcal{F} \Rightarrow A_1 \cup \dots \cup A_n \in \mathcal{F}$ und $A_1 \cap \dots \cap A_n \in \mathcal{F}$
- 4) $A, B \in \mathcal{F} \Rightarrow A \setminus B \in \mathcal{F}$

Erste Beispiele: 1) $\mathcal{F} = \{\emptyset, \Omega\}$ (größte σ -Algebra über Ω)

2) $\mathcal{F} = \mathcal{P}(\Omega)$ (feinste σ -Algebra über Ω)

3) Für $A \subset \Omega$: $\mathcal{F} = \{\emptyset, A, A^c, \Omega\}$

Im Weiteren bezeichnet Ω immer eine nichtleere Menge.

Satz 1.1.1 Sei S eine Menge, die aus σ -Algebren über Ω besteht. Dann ist der Durchschnitt

$$\bigcap_{\mathcal{F} \in S} \mathcal{F}$$

wieder eine σ -Algebra über Ω .

Definition (erzeugte σ -Algebra) Sei $\mathcal{M} \subset \mathcal{P}(\Omega)$. Dann heißt

$$\sigma(\mathcal{M}) = \sigma_{\Omega}(\mathcal{M}) := \bigcap_{\substack{\mathcal{F} \text{ } \sigma\text{-Algebra über } \Omega \\ \mathcal{F} \supset \mathcal{M}}} \mathcal{F}$$

die von \mathcal{M} erzeugte σ -Algebra über Ω .

Definition (Borel- σ -Algebra) Sei (X, \mathcal{O}) ein topologischer Raum (\mathcal{O} bezeichnet das System der offenen Mengen). Dann heißt $\mathcal{B}(X) = \mathcal{B}(X, \mathcal{O}) = \sigma_X(\mathcal{O})$ die σ -Algebra der Borelmengen über X .

Wir betrachten die reellen Zahlen immer als topologischen Raum mit der üblichen lokalkompakten Topologie.

Satz 1.1.2

$$\begin{aligned} \mathcal{B}(\mathbb{R}) &= \sigma(\{]-\infty, \alpha[\mid \alpha \in \mathbb{R} \}) \\ &= \sigma(\{]-\infty, \alpha[\mid \alpha \in \mathbb{R} \}) \\ &= \sigma(\{ [\alpha, \infty[\mid \alpha \in \mathbb{R} \}) \\ &= \sigma(\{]\alpha, \infty[\mid \alpha \in \mathbb{R} \}) \\ &= \sigma(\{]\alpha, \beta[\mid \alpha, \beta \in \mathbb{R} \}) \\ &= \dots \end{aligned}$$

Definition (messbarer Raum) Ein *messbarer Raum* ist ein Paar (Ω, \mathcal{F}) bestehend aus

- einer nichtleeren Menge Ω und
- einer σ -Algebra \mathcal{F} über Ω .

Die Mengen aus \mathcal{F} heißen *messbar*.

Definition (messbare Abbildung) Seien (Ω, \mathcal{F}) und (E, \mathcal{E}) zwei messbare Räume. Eine Abbildung $f : \Omega \rightarrow E$ heißt *messbar*, genauer \mathcal{F} - \mathcal{E} -messbar, wenn gilt:

$$B \in \mathcal{E} \Rightarrow f^{-1}(B) \in \mathcal{F}, \tag{1.1.1}$$

wenn also Urbilder messbarer Mengen messbar sind.

Bemerkungen: 1) Die Hintereinanderschaltung messbarer Abbildungen ist messbar.

2) Konstante Abbildungen sind stets messbar. Es gilt ja für eine beliebige Teilmenge B von Ω und $x \in E$:

$$f(\omega) = x \quad \forall \omega \in \Omega \Rightarrow f^{-1}(B) = \begin{cases} \Omega & \text{falls } x \in B \\ \emptyset & \text{falls } x \notin B. \end{cases}$$

3) Für messbare Räume (Ω, \mathcal{F}) und (E, \mathcal{E}) gilt: Jede Abbildung $f : \Omega \rightarrow E$ ist $\mathcal{P}(\Omega)$ - \mathcal{E} -messbar und \mathcal{F} - $\{\emptyset, E\}$ -messbar.

Satz 1.1.3 Sei (Ω, \mathcal{F}) ein messbarer Raum, E eine nichtleere Menge und $\mathcal{N} \subset \mathcal{P}(E)$.

Dann ist $f : \Omega \rightarrow E$ genau dann \mathcal{F} - $\sigma_E(\mathcal{N})$ -messbar, wenn die folgende Bedingung erfüllt ist:

$$B \in \mathcal{N} \Rightarrow f^{-1}(B) \in \mathcal{F}.$$

Für die Messbarkeit muss man also Bedingung (1.1.1) nur für Mengen B aus einem Erzeugendensystem nachprüfen.

Korollar 1.1.4 (a) Seien (X, \mathcal{O}) und (Y, \mathcal{R}) topologische Räume.

Dann ist jede stetige Abbildung $f : X \rightarrow Y$ eine $\mathcal{B}(X)$ - $\mathcal{B}(Y)$ -messbare Abbildung.

(b) Eine Abbildung $f : \Omega \rightarrow \mathbb{R}$ ist genau dann \mathcal{F} - $\mathcal{B}(\mathbb{R})$ -messbar, wenn gilt

$$\{f \leq \alpha\} \in \mathcal{F} \quad \forall \alpha \in \mathbb{R}.$$

Definition (numerische Funktionen) Eine *numerische Funktion* auf Ω ist eine Funktion

$$f : \Omega \rightarrow \overline{\mathbb{R}} := \mathbb{R} \cup \{+\infty, -\infty\}.$$

Ist (Ω, \mathcal{F}) ein messbarer Raum, so heißt eine numerische Funktion f auf Ω (\mathcal{F} -)messbar, wenn sie \mathcal{F} - $\mathcal{B}(\overline{\mathbb{R}})$ -messbar ist, wobei $\overline{\mathbb{R}}$ mit der Topologie

$$\{U, U \cup \{+\infty\}, U \cup \{-\infty\}, U \cup \{\pm\infty\} \mid U \text{ offen in } \mathbb{R}\}$$

versehen wird.

Im Weiteren sei (Ω, \mathcal{F}) ein messbarer Raum.

Satz 1.1.5 Sei f_n eine Folge \mathcal{F} -messbarer numerischer Funktionen auf Ω .

Dann gilt:

(a) $\sup_{n \in \mathbb{N}} f_n$, $\inf_{n \in \mathbb{N}} f_n$, $\limsup_{n \rightarrow \infty} f_n$ und $\liminf_{n \rightarrow \infty} f_n$ sind \mathcal{F} -messbare numerische Funktionen auf Ω .

(b) Falls die Grenzfunktion $\lim_{n \rightarrow \infty} f_n$ (punktweise in $\overline{\mathbb{R}}$) existiert, so ist sie eine \mathcal{F} -messbare numerische Funktion.

Definition (Maß) Ein (positives) *Maß* μ auf dem messbaren Raum (Ω, \mathcal{F}) ist eine Abbildung

$$\mu : \mathcal{F} \rightarrow \overline{\mathbb{R}}_+ := \mathbb{R}_+ \cup \{\infty\},$$

mit:

· Für jede paarweise disjunkte Folge $A_1, A_2, \dots \in \mathcal{F}$ gilt

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n).$$

· Es existiert eine Menge $A \in \mathcal{F}$ mit $\mu(A) < \infty$.

Ein Tripel $(\Omega, \mathcal{F}, \mu)$ bestehend aus

- einem messbaren Raum (Ω, \mathcal{F})
- einem Maß μ auf (Ω, \mathcal{F})

heißt *Maßraum*.

Bemerkungen 1) Die zweite Bedingung in der Definition ist äquivalent zu:

$$\mu(\emptyset) = 0.$$

$$2) A_1, \dots, A_n \in \mathcal{F} \text{ paarweise disjunkt} \Rightarrow \mu(A_1 \cup \dots \cup A_n) = \mu(A_1) + \dots + \mu(A_n)$$

$$3) A, B \in \mathcal{F}, A \subset B \Rightarrow$$

$$\mu(A) \leq \mu(B).$$

Satz 1.1.6 (“Stetigkeit” von Maßen)

(a) Für jede Folge $A_1, A_2, \dots \in \mathcal{F}$ mit $A_1 \subset A_2 \subset \dots$ gilt

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n \rightarrow \infty} \mu(A_n).$$

(b) Für jede Folge $A_1, A_2, \dots \in \mathcal{F}$ mit $A_1 \supset A_2 \supset \dots$ und $\mu(A_1) < \infty$ gilt

$$\mu\left(\bigcap_{n=1}^{\infty} A_n\right) = \lim_{n \rightarrow \infty} \mu(A_n).$$

Definition (Wahrscheinlichkeitsraum) Ein Maßraum $(\Omega, \mathcal{F}, \mu)$ heißt *Wahrscheinlichkeitsraum*, wenn gilt

$$\mu(\Omega) = 1.$$

Wir schreiben im Falle eines Wahrscheinlichkeitsraumes häufig P statt μ .

Stochastische Prozesse
SoSe 2018
Übersicht 2. Vorlesungswoche (16.-20.4.2018)

Beispiel: 1) Sei $\Omega \neq \emptyset$ eine Menge. Wir setzen

$$\mu(A) = \begin{cases} \#A & \text{wenn } A \text{ endlich} \\ \infty & \text{sonst} \end{cases}$$

für $A \subset \Omega$. Dann ist μ ein Maß auf $(\Omega, \mathcal{P}(\Omega))$. Dieses Maß heißt *Zählmaß* auf Ω . Ist Ω endlich, $\#\Omega = n$, so können wir μ zu einem Wahrscheinlichkeitsmaß P normieren, indem wir $P(A) = \frac{\mu(A)}{n}$ definieren. In diesem Fall nennen wir $(\Omega, \mathcal{P}(\Omega), P)$ den *Laplaceschen Wahrscheinlichkeitsraum* über Ω . (Alle Elementarereignisse besitzen die gleiche Wahrscheinlichkeit $\frac{1}{n}$.)

2) Sei für $\omega_0 \in \Omega$

$$\mu(A) = \begin{cases} 1 & \text{falls } \omega_0 \in A \\ 0 & \text{falls } \omega_0 \notin A \end{cases}$$

für $A \subset \Omega$. Dann ist μ ein Wahrscheinlichkeitsmaß auf $(\Omega, \mathcal{P}(\Omega))$. Es heißt das im Punkte ω_0 konzentrierte *Dirac-Maß*. Wir bezeichnen dieses Maß mit δ_{ω_0} .

3) Es gibt genau ein Maß λ^d auf dem messbaren raum $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ mit der folgenden Eigenschaft:

$$\lambda^d([a, b]) = \prod_{i=1}^d (b_i - a_i),$$

wobei $a, b \in \mathbb{R}^d$ mit

$$a = (a_1, \dots, a_d), b = (b_1, \dots, b_d); a_i \leq b_i, i = 1, \dots, d.$$

Für die Existenz und Eindeutigkeit von λ^d siehe z. B. "H. Bauer: Maß- und Integrationstheorie" oder "Dudley: Real analysis and probability". λ^d heißt *Lebesgue-Maß* oder Lebesgue-Borel-Maß auf \mathbb{R}^d .

1.2 Integrationstheorie nach Lebesgue

Definition (einfache Funktion) Sei (Ω, \mathcal{F}) ein messbarer Raum. Eine Funktion

$$f : \Omega \rightarrow \mathbb{R}$$

heißt *einfach*, wenn gilt

- f ist $(\mathcal{F}\text{-}\mathcal{B}(\mathbb{R}))$ -messbar
- f nimmt nur endlich viele Werte an.

Bemerkung Für eine einfache Funktion f gilt

$$f(\omega) = \sum_{i=1}^n \alpha_i \mathbf{1}_{A_i}(\omega), \quad (1.2.1)$$

wobei $\alpha_1, \dots, \alpha_n$ die n paarweise verschiedenen Werte von f sind und

$$A_i = \{\omega \in \Omega \mid f(\omega) = \alpha_i\} = \{f = \alpha_i\}.$$

Hierbei bezeichnet für $A \subset \Omega$ das Symbol $\mathbf{1}_A$ die *Indikatorfunktion* von A , d. h.

$$\mathbf{1}_A(\omega) = \begin{cases} 1 & \text{falls } \omega \in A \\ 0 & \text{falls } \omega \notin A. \end{cases}$$

Wir bezeichnen die Darstellung (1.2.1) als *kanonische Darstellung* von f . Offenbar hat man

- $\alpha_1, \dots, \alpha_n \in \mathbb{R}$ paarweise verschieden
- A_1, \dots, A_n paarweise disjunkt
- $A_1, \dots, A_n \in \mathcal{F}$
- $A_1 \cup \dots \cup A_n = \Omega$

Eine Darstellung (1.2.1) der einfachen Funktion f mit den obigen 4 Eigenschaften ist *eindeutig*. Auf der anderen Seite ist jede Funktion f mit einer Darstellung (1.2.1) mit *beliebigen* $\alpha_1, \dots, \alpha_n, A_1, \dots, A_n \in \mathcal{F}$ einfach.

Definition (Funktionenräume für die Integrationstheorie)

Für einen messbaren Raum (Ω, \mathcal{F}) setzen wir:

- $F_{+,0}(\Omega, \mathcal{F}) = F_{+,0} := \{f : \Omega \rightarrow \mathbb{R} \mid f \text{ ist einfach und } f \geq 0\}$
- $\overline{F}_+(\Omega, \mathcal{F}) = \overline{F}_+ := \{f : \Omega \rightarrow \overline{\mathbb{R}} \mid f \text{ ist } \mathcal{F}\text{-}\mathcal{B}(\overline{\mathbb{R}})\text{-messbar und } f \geq 0\}$
- $\overline{F}(\Omega, \mathcal{F}) = \overline{F} := \{f : \Omega \rightarrow \overline{\mathbb{R}} \mid f \text{ ist } \mathcal{F}\text{-}\mathcal{B}(\overline{\mathbb{R}})\text{-messbar}\}$

Im weiteren sei immer ein Maßraum $(\Omega, \mathcal{F}, \mu)$ fest vorgegeben.

Satz 1.2.1 Für jede Funktion $f \in \overline{F}_+$ gibt es eine Folge $f_n, n \in \mathbb{N}$, von Funktionen aus $F_{+,0}$ mit

- $f_1 \leq f_2 \leq \dots \leq f$
- $\lim_{n \rightarrow \infty} f_n(\omega) = f(\omega) \quad \forall \omega \in \Omega.$

Definition (Lebesgue-Integral für einfache Funktionen ≥ 0)

Für eine Funktion $f \in F_{+,0}$ mit kanonischer Darstellung (1.2.1) definieren wir das Lebesgue-Integral von f (bzgl. des Maßes μ) durch

$$\int f \, d\mu := \sum_{i=1}^n \alpha_i \mu(A_i).$$

Dabei verwenden wir die *maßtheoretische Konvention*

$$0 \cdot \infty := 0.$$

Es ist also $\int f \, d\mu$ eine Zahl aus $\overline{\mathbb{R}}_+$.

Lemma 1.2.2 (Monotonie des Integrals)

Für $f, g \in F_{+,0}$ mit $f \leq g$ gilt

$$\int f \, d\mu \leq \int g \, d\mu.$$

Definition (Lebesgue-Integral für messbare numerische Funktionen ≥ 0)

Für eine Funktion $f \in \overline{F}_+$ definieren wir das Lebesgue-Integral von f (bzgl. des Maßes μ) durch

$$\int f \, d\mu := \sup\left\{ \int g \, d\mu \mid g \in F_{+,0}, g \leq f \right\}.$$

Bemerkungen: 1) Man überzeugt sich mit Hilfe von Lemma 1.2.2, dass die beiden Definitionen von $\int f \, d\mu$ für $f \in F_{+,0}$ übereinstimmen!

2) Es gilt allgemein (wieder nach Lemma 1.2.2) für $f, g \in \overline{F}_+$:

$$f \leq g \Rightarrow \int f \, d\mu \leq \int g \, d\mu$$

3) Für $A \in \mathcal{F}$ und $f \in \overline{F}_+$ ist die Funktion $\mathbf{1}_A f$ wieder aus \overline{F}_+ . Wir setzen $f \in \overline{F}_+$:

$$\int_A f \, d\mu := \int \mathbf{1}_A f \, d\mu$$

4) Es gilt für $f, g \in F_{+,0}$ und $c \in \mathbb{R}_+$:

$$\int (cf + g) \, d\mu = c \int f \, d\mu + \int g \, d\mu$$

(Linearität des Integrals gilt auch allgemein für Funktionen aus \overline{F}_+ . Folgt aus dem nächsten Satz; siehe Korollar.)

Satz 1.2.3 (Satz von der monotonen Konvergenz = Satz von Beppo Levi)
 Für $f_n \in \overline{F}_+$, $n \in \mathbb{N}$, mit

$$f_1 \leq f_2 \leq \dots$$

liegt

$$f := \lim_{n \rightarrow \infty} f_n = \sup_{n \in \mathbb{N}} f_n$$

wieder in \overline{F}_+ und es gilt

$$\int f \, d\mu = \lim_{n \rightarrow \infty} \int f_n \, d\mu = \sup_{n \in \mathbb{N}} \int f_n \, d\mu.$$

Bemerkung: Ein Vergleich mit Satz 1.2.1 zeigt nun: Wir können $\int f \, d\mu$ für $f \in \overline{F}_+$ auch als

$$\lim_{n \rightarrow \infty} \int f_n \, d\mu$$

definieren, wobei f_n eine (beliebige, nach Satz 1.2.1 existierende) Folge von Funktionen aus $F_{+,0}$ mit $f_1 \leq f_2 \leq \dots$ und $f = \lim_{n \rightarrow \infty} f_n$ ist.

Korollar 1.2.4 (Linearität des Lebesgue-Integrals)

Es gilt für $f, g \in F_{+,0}$ und $c \in \mathbb{R}_+$:

$$\int (cf + g) \, d\mu = c \int f \, d\mu + \int g \, d\mu$$

Satz 1.2.5 Für $f \in \overline{F}_+$ gilt:

$$\int f \, d\mu = 0 \Leftrightarrow \mu(f > 0) = 0.$$

Definition (Lebesgue-Integral für integrierbare Funktionen)

Für $f \in \overline{F}$ setzen wir:

$$f_+(\omega) = \begin{cases} f(\omega) & \text{falls } f(\omega) \geq 0 \\ 0 & \text{falls } f(\omega) < 0, \end{cases}$$

d. h. $f_+ = \mathbf{1}_{\{f \geq 0\}} f$. Außerdem: $f_- := (-f)_+$. Eine Funktion $f \in \overline{F}$ heißt (μ) -integrierbar, wenn gilt:

$$\int f_+ \, d\mu < \infty \text{ \underline{und} } \int f_- \, d\mu < \infty.$$

Wenn f integrierbar ist, dann heißt die reelle Zahl

$$\int f \, d\mu := \left(\int f_+ \, d\mu \right) - \left(\int f_- \, d\mu \right)$$

das Integral von f bzgl. μ . Wir setzen

$$L^1(\mu) = L^1(\Omega, \mathcal{F}, \mu) := \{f \in \overline{F} \mid f \text{ } \mu\text{-integrierbar}\}.$$

Stochastische Prozesse

SoSe 2018

Übersicht 3. Vorlesungswoche (23.-27.4.2018)

Satz 1.2.7 Für $f \in \bar{F}$ gilt:

$$f \in L^1(\mu) \Leftrightarrow |f| \in L^1(\mu) \Leftrightarrow \int |f| d\mu < \infty$$

Bemerkungen: 1) Für $f \in L^1(\mu)$ gilt die Ungleichung $|\int f d\mu| \leq \int |f| d\mu$.
 2) (Linearität des Lebesgue-Integrals) $L^1(\mu)$ ist ein Vektorraum, d. h.

$$f, g \in L^1(\mu), c \in \mathbb{R} \Rightarrow cf + g \in L^1(\mu)$$

3) $f \in L^1(\mu) \Rightarrow \mu(|f| = \infty) = 0$

4) (Monotonie des Lebesgue-Integrals) $f, g \in L^1(\mu), f \leq g \Rightarrow \int f d\mu \leq \int g d\mu$

Falls für $f \in \bar{F}$ eine Funktion $g \in L^1(\mu)$ mit $|f| \leq g$ existiert, so folgt: $f \in L^1(\mu)$.

Lemma 1.2.8 (Lemma von Fatou)

Für eine Folge f_n von Funktionen aus \bar{F}_+ gilt stets die Abschätzung:

$$\int (\liminf_{n \rightarrow \infty} f_n) d\mu \leq \liminf_{n \rightarrow \infty} \int f_n d\mu$$

Satz 1.2.9 (Satz von der majorisierten Konvergenz = Lebesguescher Konvergenzsatz)

Seien $f_n, f \in \bar{F}$ mit $f(\omega) = \lim_{n \rightarrow \infty} f_n(\omega) \forall \omega \in \Omega$ (d. h. f_n konvergiert punktweise gegen f).

Außerdem existiere eine Funktion $g \in L^1(\mu)$ mit

$$|f_n| \leq g \quad \forall n \in \mathbb{N}.$$

Dann folgt $f_n, f \in L^1(\mu)$, und es gilt:

$$\lim_{n \rightarrow \infty} \int f_n d\mu = \int f d\mu \tag{1.2.1}$$

sowie

$$\lim_{n \rightarrow \infty} \int |f_n - f| d\mu = 0. \tag{1.2.2}$$

Bemerkungen: 1) Aus (1.2.2) folgt (1.2.1).

2) Für Anwendungen in der Wahrscheinlichkeitstheorie ist der folgende Spezialfall des Lebesgueschen Konvergenzsatzes wichtig: Nehmen wir an, dass $(\Omega, \mathcal{F}, \mu) = (\Omega, \mathcal{F}, P)$ ein Wahrscheinlichkeitsraum ist. Sind dann die Funktionen $f_n \in \bar{F}$ gleichmäßig in n beschränkt, d. h. es existiert ein Konstante $C \geq 0$, so dass $|f_n(\omega)| \leq C \forall \omega \in \Omega \forall n \in \mathbb{N}$, dann folgt aus der punktweisen Konvergenz von f_n gegen eine Funktion $f \in \bar{F}$ die Integrierbarkeit der f_n , f und die Gültigkeit von (1.2.2) und damit von (1.2.1).

2 Die Konstruktion von Daniell-Kolmogorov

2.1 Grundlegende Begriffe der Theorie der stochastischen Prozesse

Gegenüberstellung von Begriffen der Maß- und der Wahrscheinlichkeitstheorie:

Maßtheorie	Wahrscheinlichkeitstheorie
$(\Omega, \mathcal{F}, \mu)$ Maßraum	(Ω, \mathcal{F}, P) Wahrscheinlichkeitsraum
$f : \Omega \rightarrow \mathbb{R}$ messbare Funktion	$X : \Omega \rightarrow \mathbb{R}$ Zufallsvariable
$\int f d\mu$ Integral von f	$\int X dP = E(X)$ Erwartungswert von X
$f(\mu)$ Bildmaß von μ unter f	$X(P) = P_X$ Verteilung von X

Dabei:

Definition (Bildmaß) Für einen Maßraum $(\Omega, \mathcal{F}, \mu)$, einen messbaren Raum (E, \mathcal{E}) und eine messbare Abbildung $f : \Omega \rightarrow E$ wir durch

$$f(\mu)(B) := \mu(f^{-1}(B)) = \mu(f \in B), \quad B \in \mathcal{E}$$

ein Maß auf (E, \mathcal{E}) definiert (Beweis?). Dieses Maß bezeichnen wir als das *Bildmaß* von μ (unter der Abbildung f).

Bemerkung: In Falle eines Wahrscheinlichkeitsraumes (Ω, \mathcal{F}, P) ist $f(P)$ ein Wahrscheinlichkeitsmaß.

Definition (Zufallsvariable) Eine \mathcal{F} - $\mathcal{B}(\mathbb{R})$ -messbare Abbildung $X : \Omega \rightarrow \mathbb{R}$ heißt (reelle) *Zufallsvariable* (über dem Wahrscheinlichkeitsraum (Ω, \mathcal{F}, P)). Allgemeiner bezeichnen wir eine messbare Abbildung $X : \Omega \rightarrow E$ (mit "Werten" in dem messbaren Raum (E, \mathcal{E})) als E -wertige (oder (E, \mathcal{E}) -wertige) Zufallsvariable. Im Falle $(E, \mathcal{E}) = (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ sprechen wir auch von einer d -dimensionalen Zufallsvariablen. Zufallsvariablen sind also (wenn sonst

nichts gesagt wird) $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ -Zufallsvariablen und auch 1-dimensionale Zufallsvariablen.

Bemerkung: Für eine Zufallsvariable X ist $X(P)$ also ein Wahrscheinlichkeitsmaß auf \mathbb{R} . Wir schreiben für dieses Wahrscheinlichkeitsmaß wie üblich P_X . Das Bildmaß $X(P) = P_X$ heißt *Verteilung* von X .

Definition (stochastischer Prozess)

Sei T eine beliebige Menge $\neq \emptyset$. Ein *stochastischer Prozess* mit Indexmenge T , über einem Wahrscheinlichkeitsraum (Ω, \mathcal{F}, P) , ist eine Familie

$$\mathbf{X} = (X_t)_{t \in T}$$

von Zufallsvariablen über (Ω, \mathcal{F}, P) .

Definition (gemeinsame Verteilung)

Für $n \in \mathbb{N}$ Zufallsvariablen X_1, \dots, X_n über dem (gleichen) Wahrscheinlichkeitsraum (Ω, \mathcal{F}, P) heißt das Bildmaß von P unter der n -dimensionalen Zufallsvariablen

$$(X_1, \dots, X_n) : \Omega \rightarrow \mathbb{R}^d; (X_1, \dots, X_n)(\omega) := (X_1(\omega), \dots, X_n(\omega))$$

die *gemeinsame Verteilung* der X_1, \dots, X_n .

Die gemeinsame Verteilung ist ein Wahrscheinlichkeitsmaß auf \mathbb{R}^d . Sie wird mit $P_{(X_1, \dots, X_n)}$ bezeichnet.

Definition (endlich-dimensionale Verteilungen)

Sei $(X_t)_{t \in T}$ ein stochastischer Prozess. Setze

$$\mathcal{P}_0(T) = \{I \subset T \mid \#I < \infty, I \neq \emptyset\} \subset \mathcal{P}(T).$$

Für $I \in \mathcal{P}_0(T)$, $I = \{t_1, \dots, t_n\}$, $n = \#I$, bezeichnen wir mit P_I die gemeinsame Verteilung der Zufallsvariablen X_{t_1}, \dots, X_{t_n} .

Die Familie $(P_I)_{I \in \mathcal{P}_0(T)}$ heißt *Familie der endlich-dimensionalen Verteilungen* des stochastischen Prozesses $(X_t)_{t \in T}$.

Bemerkung: Es ist also P_I ein Wahrscheinlichkeitsmaß auf \mathbb{R}^n mit $n = \#I$. Wir verwenden auch folgende Notation: Für $I \subset T$, $I \neq \emptyset$, setzen wir

$$\mathbb{R}^I = \text{Abb}(I, \mathbb{R}) := \{\alpha : I \rightarrow \mathbb{R}\} = \{(\alpha_t)_{t \in I} \mid \alpha_s \in \mathbb{R}, s \in I\}.$$

Dann können wir \mathbb{R}^I für $I \in \mathcal{P}_0(T)$ mit \mathbb{R}^n identifizieren (sobald wir eine Reihenfolge der Elemente von I festgelegt haben). In diesem Sinne ist auch der messbare Raum $(\mathbb{R}^I, \mathcal{B}(\mathbb{R}^I))$ zu interpretieren.

Des Weiteren setzen wir ganz allgemein für eine Familie $(\Omega_t)_{t \in T}$ nichtleerer Mengen Ω_t und $B_t \subset \Omega_t$, $t \in I$, $I \subset T$, $I \neq \emptyset$,

$$\prod_{t \in I} B_t := \{(\alpha_t)_{t \in I} \mid \alpha_s \in B_s, s \in I\},$$

d. h. $\prod_{t \in I} B_t$ ist das kartesische Produkt der B_t , $t \in I$, und eine Teilmenge von $\prod_{t \in I} \Omega_t$.

Definition (stochastische Äquivalenz)

Zwei stochastische Prozesse $\mathbf{X} = (X_t)_{t \in T}$ (über $(\omega^{(1)}, \mathcal{F}^{(1)}, P^{(1)})$) und $\mathbf{Y} = (Y_t)_{t \in T}$ (über $(\omega^{(2)}, \mathcal{F}^{(2)}, P^{(2)})$) (insbesondere müssen \mathbf{X} und \mathbf{Y} dieselbe Indexmenge haben!) heißen *stochastisch äquivalent*, wenn sie dieselbe Familie endlich-dimensionaler Verteilungen besitzen, wenn also gilt

$$P^{(1)}(X_{t_1} \in B_1, \dots, X_{t_n} \in B_n) = P^{(2)}(Y_{t_1} \in B_1, \dots, Y_{t_n} \in B_n)$$

für alle $n \in \mathbb{N}$, $B_i \in \mathcal{B}(\mathbb{R})$, $t_i \in T$, $i = 1, \dots, n$.

Definition

Sei $(\Omega_t)_{t \in T}$ eine Familie von nichtleeren Mengen. Wir führen für $I, J \subset T$, $I \subset J$, $I \neq \emptyset$, die *Projektionen*

$$\begin{aligned} p_I &: \prod_{t \in T} \Omega_t \rightarrow \prod_{t \in I} \Omega_t \\ p_{J,I} &: \prod_{t \in J} \Omega_t \rightarrow \prod_{t \in I} \Omega_t \end{aligned}$$

durch

$$\begin{aligned} p_I((\omega_t)_{t \in T}) &= (\omega_t)_{t \in I} \\ p_{J,I}((\omega_t)_{t \in J}) &= (\omega_t)_{t \in I} \end{aligned}$$

ein. (Also ist $p_I = p_{T,I}$.)

Satz 2.1.1 Für einen stochastischen Prozess $(X_t)_{t \in T}$ gilt

$$P_I = p_{J,I}(P_J) \tag{2.1.1}$$

für alle $I, J \in \mathcal{P}_0(T)$, $I \subset J$.

2.2 Der Satz von Daniell-Kolmogorov

Definition (projektive Familie von Wahrscheinlichkeitsmaßen)

Eine Familie $(P_I)_{I \in \mathcal{P}_0(T)}$, P_I Wahrscheinlichkeitsmaß auf $(\mathbb{R}^I, \mathcal{B}(\mathbb{R}^I))$, heißt *projektiv*, wenn (2.1.1) gilt.

Definition (Produkt messbarer Räume)

Gegeben sei eine Familie $((\Omega_t, \mathcal{F}_t))_{t \in T}$ messbarer Räume. Wir setzen $\Omega := \prod_{t \in T} \Omega_t$ und

$$\bigotimes_{t \in T} \mathcal{F}_t := \sigma_{\Omega}(\{\prod_{t \in T} A_t \mid A_s \in \mathcal{F}_s, s \in T, \text{ und:} \quad (2.2.1)$$

$$\exists I \in \mathcal{P}_0(T) \text{ mit } A_s = \Omega_s \text{ für } s \in I^c\}). \quad (2.2.2)$$

Dann heißt

$$\bigotimes_{t \in T} ((\Omega_t, \mathcal{F}_t)) := (\Omega, \bigotimes_{t \in T} \mathcal{F}_t)$$

Produkt der messbaren Räume $(\Omega_t, \mathcal{F}_t)$, $t \in T$.

Die Mengen $\prod_{t \in T} A_t$ der echten Seite von (2.2.1) heißen *Rechteckzylindermengen* von $\bigotimes_{t \in T} ((\Omega_t, \mathcal{F}_t))$.

Bemerkung: Falls $T = \{1, \dots, n\}$, $n \in \mathbb{N}$, so ist

$$\bigotimes_{t \in T} (\Omega_t, \mathcal{F}_t) = (\Omega_1 \times \dots \times \Omega_n, \mathcal{F}_1 \otimes \dots \otimes \mathcal{F}_n)$$

und

$$\mathcal{F}_1 \otimes \dots \otimes \mathcal{F}_n = \sigma_{\Omega_1 \times \dots \times \Omega_n}(\{\mathcal{B}_1 \times \dots \times \mathcal{B}_n \mid B_i \in \mathcal{B}(\mathbb{R}), i = 1, \dots, n\}).$$

Behauptung: Es gilt für $I \subset T$, $I \neq \emptyset$,

$$p_I^{-1}(B) \in \bigotimes_{t \in T} \mathcal{F}_t$$

für alle $B \in \bigotimes_{t \in I} \mathcal{F}_t$, d. h. die Abbildungen p_I sind $\bigotimes_{t \in T} \mathcal{F}_t$ - $\bigotimes_{t \in I} \mathcal{F}_t$ -messbar (und somit die $p_{J,I}$ messbar bzgl. $\bigotimes_{t \in J} \mathcal{F}_t$ und $\bigotimes_{t \in I} \mathcal{F}_t$).

Definiton Die Mengen der Gestalt $p_I^{-1}(B)$ mit $B \in \bigotimes_{t \in I} \mathcal{F}_t$ und $I \in \mathcal{P}_0(T)$ heißen *Zylindermengen* von $\bigotimes_{t \in T} ((\Omega_t, \mathcal{F}_t))$. Rechteckzylindermengen sind also spezielle Zylindermengen.

Lemma 2.2.1 (Reduktion auf Abzählbarkeit)

$$\bigotimes_{t \in T} \mathcal{F}_t = \bigcup_{\substack{I \subset T, I \neq \emptyset \\ I \text{ abzählbar}}} p_I^{-1}(\bigotimes_{t \in T} \mathcal{F}_t)$$

Stochastische Prozesse

SoSe 2018

Übersicht 4. Vorlesungswoche (30.4.-4.5.2018)

Wir verwenden (siehe z. B. “J. Elstroth: Maß- und Integrationstheorie” oder “H. Bauer: Maßtheorie”) den

Satz (von Carathéodory)

Sei $\Omega (\neq \emptyset)$ eine Menge und \mathcal{A} eine Algebra über Ω .

Dann kann jedes σ -endliche Prämaß auf \mathcal{A} auf genau eine Weise zu einem Maß auf $\sigma_\Omega(\mathcal{A})$ fortgesetzt werden.

Bemerkungen: 1) $\mathcal{A} \subset \mathcal{P}(\Omega)$ heißt *Algebra*, wenn gilt:

- $\Omega \in \mathcal{A}$
- $A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$
- $A, B \in \mathcal{A} \Rightarrow A \cup B \in \mathcal{A}$

(Die letzte Bedingung ist offenbar äquivalent zu: $A_1, \dots, A_n \in \mathcal{A} \Rightarrow A_1 \cup \dots \cup A_n \in \mathcal{A}$.)

2) $\mu : \mathcal{A} \rightarrow \overline{\mathbb{R}}_+$, \mathcal{A} Algebra, heißt *Inhalt* auf \mathcal{A} , wenn

- $\mu(\emptyset) = 0$
- $\mu(A \cup B) = \mu(A) + \mu(B)$ für $A, B \in \mathcal{A}$ mit $A \cap B = \emptyset$.

(Die letzte Bedingung ist offenbar äquivalent zu: $\mu(A_1 \cup \dots \cup A_n) = \mu(A_1) + \dots + \mu(A_n)$ für $A_1, \dots, A_n \in \mathcal{A}$ paarweise disjunkt, $n \in \mathbb{N}$.)

3) Ein Inhalt heißt *Prämaß*, wenn die stärkere Bedingung

$$A_1, A_2, \dots \in \mathcal{A} \text{ paarweise disjunkt mit } \bigcup_{n=1}^{\infty} A_n \in \mathcal{A} \\ \Rightarrow \mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n)$$

erfüllt ist.

4) Ein Maßraum heißt *σ -endlich*, wenn es $A_1, A_2, \dots \in \mathcal{A}$ gibt mit

- $\mu(A_n) < \infty$
- $\bigcup_{n=1}^{\infty} A_n = \Omega$.

DanKol **Satz 2.2.2** (von Daniell-Kolmogorov)

Sei $(P_I)_{I \in \mathcal{P}_0(T)}$ eine projektive Familie von Wahrscheinlichkeitsmaßen.

Dann gibt es ein eindeutig bestimmtes Wahrscheinlichkeitsmaß P auf dem messbaren Raum $(\mathbb{R}^T, \mathcal{B}(\mathbb{R}^T))$ mit

$$p_I(P) = P_I \quad \forall I \in \mathcal{P}_0(T).$$

Beweisskizze: Falls T endlich ist, tut es offenbar $P = P_T$. Auch die Eindeutigkeit ist klar.-

Nun nehmen wir an, dass T *abzählbar unendlich* ist, $T = \mathbb{N}$. Betrachte das Mengensystem

$$\mathcal{Z} := \bigcup_{I \in \mathcal{P}_0(\mathbb{N})} p_I^{-1}(\mathcal{B}(\mathbb{R}^I))$$

der Zylindermengen in $\mathcal{B}(\mathbb{R}^{\mathbb{N}})$. Da jede endliche Teilmenge von \mathbb{N} in einer Menge der Form $\{1, \dots, n\}$ liegt, ist

$$\mathcal{Z} = \{A \times \mathbb{R} \times \mathbb{R} \times \dots \mid A \in \mathcal{B}(\mathbb{R}^n), n \in \mathbb{N}\}.$$

Behauptung: \mathcal{Z} ist eine Algebra über \mathbb{R}^T .

Beweis: $\Omega \in \mathcal{Z}$ folgt z. B. mit $A = \mathbb{R}$ und $n = 1$. Außerdem ist \mathcal{Z} stabil gegenüber Komplementbildung, weil gilt: $(A \times \mathbb{R} \times \mathbb{R} \times \dots)^c = A^c \times \mathbb{R} \times \mathbb{R} \times \dots$. Seien $A \times \mathbb{R} \times \mathbb{R} \times \dots$ und $B \times \mathbb{R} \times \mathbb{R} \times \dots$ in \mathcal{Z} mit $A \in \mathcal{B}(\mathbb{R}^m)$, $B \in \mathcal{B}(\mathbb{R}^n)$, $m \leq n$. Dann gilt:

$$\begin{aligned} & (A \times \mathbb{R} \times \mathbb{R} \times \dots) \cup (B \times \mathbb{R} \times \mathbb{R} \times \dots) \\ &= (A \times \mathbb{R}^{n-m} \cup B) \times \mathbb{R} \times \mathbb{R} \times \dots \in \mathcal{Z}, \end{aligned}$$

da $A \times \mathbb{R}^{n-m} \cup B \in \mathcal{B}(\mathbb{R}^n)$. ✓

Wir definieren nun P auf \mathcal{Z} durch

$$P(A \times \mathbb{R} \times \mathbb{R} \times \dots) := P_I(A)$$

für $I = \{1, \dots, n\}$, $\#I = n \in \mathbb{N}$, $A \in \mathcal{B}(\mathbb{R}^I) = \mathcal{B}(\mathbb{R}^n)$. Die Eigenschaft (2.1.1) einer projektiven Familie sorgt dafür, dass diese Definition von P auf der Algebra \mathcal{Z} sinnvoll ist (Beweis?!)

Der schwierige Teil des Beweises besteht nun darin zu zeigen, dass \mathcal{P} ein Prämaß auf \mathcal{Z} ist (wir verweisen auf “A. Klenke: Wahrscheinlichkeitstheorie”, Satz 14.32, oder “H. Bauer: Wahrscheinlichkeitstheorie”); vgl. die Konstruktion des Lebesgue-Maßes auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ oder die Konstruktion des Produktmaßes.

Zum Fall einer allgemeinen unendlichen Indexmenge T : Wir führen diesen Fall mit Hilfe von Lemma 2.2.1 auf den abzählbar unendlichen zurück. Für $J \subset T$, $J \neq \emptyset$, J *abzählbar*, gibt es nach dem vorhergehenden Teil des Beweises ein eindeutig bestimmtes Wahrscheinlichkeitsmaß P_J auf $(\mathbb{R}^J, \mathcal{B}(\mathbb{R}^J))$ mit $p_{J,I}(P_J) = P_I$ für jedes $I \in \mathcal{P}_0(J)$. Nach Lemma 2.2.1 gibt es für $A \in \mathcal{B}(\mathbb{R}^T)$ eine *abzählbare* Teilmenge J , $J \neq \emptyset$, von T mit $A \in p_J^{-1}(\mathcal{B}(\mathbb{R}^J))$, d. h. $A = p_J^{-1}(B)$ für ein geeignetes $B \in \mathcal{B}(\mathbb{R}^J)$. Setze

$$P(A) := P_J(B).$$

Die Eigenschaft (2.1.1) einer projektiven Familie sorgt wieder dafür, dass diese Definition sinnvoll ist (Beweis?)!

Wir müssen noch zeigen, dass P ein Wahrscheinlichkeitsmaß ist. Klar, dass gilt $P(\Omega) = P_J(\Omega) = 1$. Fehlt der Nachweis der σ -Additivität von P . Seien dazu $A_1, A_2, \dots \in \mathcal{B}(\mathbb{R}^T)$ paarweise disjunkt. Jetzt wird ähnlich wie im Beweis von Lemma 2.2.1 argumentiert. Es ist $A_n = p_{J_n}^{-1}(B_n)$, $n = 1, 2, \dots$, für geeignete abzählbare nichtleere $J_n \subset T$ und $B_n \in \mathcal{B}(\mathbb{R}^{J_n})$. Setze $J := J_1 \cup J_2 \cup \dots$. Dann ist J abzählbar. Es gilt: $A_n \in p_J^{-1}(\mathcal{B}(\mathbb{R}^J))$ und

$$\begin{aligned} P\left(\bigcup_{n=1}^{\infty} A_n\right) &= P_J\left(\bigcup_{n=1}^{\infty} A_n\right) \\ &= \sum_{n=1}^{\infty} P_J(A_n) \\ &= \sum_{n=1}^{\infty} P(A_n), \end{aligned}$$

wobei die vorletzte Gleichheit aus der σ -Additivität von P_J folgt. \square

Bemerkung: Der Satz gilt viel allgemeiner (siehe z. B. "Bauer"). Sei

$$(\Omega_t, \mathcal{F}_t)_{t \in T}$$

eine Familie messbarer Räume. Dabei müssen die $(\Omega_t, \mathcal{F}_t)$ von einem speziellen Typ sein. Es wird vorausgesetzt, dass jeder Raum ein Borelraum $(E, \mathcal{B}(E))$ ist mit einem *polnischen* topologischen Raum E . Dabei heißt ein topologischer Raum polnisch, wenn er ein vollständiger metrischer Raum mit abzählbarer Basis der Topologie ist. Dann heißt eine Familie $(P_I)_{I \in \mathcal{P}_0(T)}$ von Wahrscheinlichkeitsmaßen P_I auf $(\prod_{t \in I} \Omega_t, \otimes_{t \in I} \mathcal{F}_t)$ projektiv, wenn wieder $p_{J,I}(P_J) = P_I$ gilt. Die allgemeinere Version des Satzes 2.2.2 besagt, dass es ein eindeutig bestimmtes Wahrscheinlichkeitsmaß P auf $(\prod_{t \in T} \Omega_t, \otimes_{t \in T} \mathcal{F}_t)$ gibt mit $p_I(P) = P_I$ für alle $I \in \mathcal{P}_0(T)$.

Definition P heißt der *projektive Limes* der projektiven Familie $(P_I)_{I \in \mathcal{P}_0(T)}$.

2.3 Folgerungen

Satz 2.3.1 Sei $(P_I)_{I \in \mathcal{P}_0(T)}$ eine projektive Familie von Wahrscheinlichkeitsmaßen.

Dann gibt es einen Wahrscheinlichkeitsraum (Ω, \mathcal{F}, P) und einen stochastischen Prozess $\mathbf{X} = (X_t)_{t \in T}$ über (Ω, \mathcal{F}, P) , so dass $(P_I)_{I \in \mathcal{P}_0(T)}$ die Familie der endlich-dimensionalen Verteilungen von \mathbf{X} ist.

Beweis: Setze:

- $\Omega = \mathbb{R}^T$
- $\mathcal{F} = \mathcal{B}(\mathbb{R}^T)$
- P : das Wahrscheinlichkeitsmaß aus Satz ^{DanKol} 2.2.2
- $X_t = p_{\{t\}} : \mathbb{R}^T \rightarrow \mathbb{R}^{\{t\}} = \mathbb{R}$

Dann gilt für $n \in \mathbb{N}$, $t_1, \dots, t_n \in T$, $B_1, \dots, B_n \in \mathcal{B}(\mathbb{R})$:

$$\begin{aligned} P_{(X_{t_1}, \dots, X_{t_n})}(B_1 \times \dots \times B_n) &= P(X_{t_1} \in B_1, \dots, X_{t_n} \in B_n) \\ &= P(\{(\alpha_t)_{t \in T} \in \mathbb{R}^T \mid \alpha_{t_1} \in B_1, \dots, \alpha_{t_n} \in B_n\}) \\ &= P(p_{\{t_1, \dots, t_n\}}^{-1}(B_1 \times \dots \times B_n)) \\ &= P_{\{t_1, \dots, t_n\}}(B_1 \times \dots \times B_n). \square \end{aligned}$$

Bemerkung: Der Beweis gibt eine Konstruktion des gesuchten stochastischen Prozesses an. Allerdings sind die Zufallsvariablen X_t des Prozesses auf einem “riesigen” Raum definiert, nämlich auf

$$\mathbb{R}^T = \{(\alpha_t)_{t \in T} \mid \alpha_s \in \mathbb{R}\} = \text{Abb}(T, \mathbb{R}),$$

dem Raum *aller* Abbildungen von T nach \mathbb{R} . Wenn z. B. $T = \mathbb{R}_+$, so ist der Raum der Pfade von X_t der Raum *aller* reellwertigen Funktionen auf \mathbb{R}_+ . In jedem Einzelfall wird man versuchen, diesen Raum einzugrenzen.

Stochastische Prozesse

SoSe 2018

Übersicht 5. Vorlesungswoche (7.5.-11.5.2018)

Bemerkungen: 1) (Produktmaße) Sei $(\Omega_t, \mathcal{F}_t, \mu_t)_{t \in T}$ eine Familie von σ -endlichen Maßräumen. Dann gibt es genau ein Maß

$$\mu := \bigotimes_{t \in T} \mu_t \text{ auf } \bigotimes_{t \in T} (\Omega_t, \mathcal{F}_t, \mu_t)$$

mit

$$\mu(p_I^{-1}(\prod_{t \in I} B_t)) = \prod_{t \in I} \mu_t(B_t)$$

für alle $I \in \mathcal{P}_0(T)$, $B_t \in \mathcal{F}_t$; siehe "Elstrodt" oder "Bauer". Das Maß $\bigotimes_{t \in T} \mu_t$ heißt das *Produktmaß* der Familie $(\Omega_t, \mathcal{F}_t, \mu_t)_{t \in T}$. Das Produktmaß für Wahrscheinlichkeitsräume $(\Omega_t, \mathcal{F}_t, P_t)_{t \in T}$ ist wieder ein Wahrscheinlichkeitsmaß. Im Fall von zwei Maßräumen (d. h. $T = \{1, 2\}$) ist das Produktmaß $\mu_1 \otimes \mu_2$ ein Maß auf $(\Omega_1 \times \Omega_2, \mathcal{F}_1 \otimes \mathcal{F}_2)$. Insbesondere ist das Produkt zweier Wahrscheinlichkeitsmaße auf \mathbb{R} ein Wahrscheinlichkeitsmaß auf \mathbb{R}^2 .

2) (Satz von Fubini) Wir verwenden die Sätze von Tonelli und Fubini:

Satz (von Tonelli)

Für zwei σ -endliche Maßräume $(\Omega, \mathcal{F}, \mu)$ und (E, \mathcal{E}, ν) und $f \in \bar{F}_+(\Omega \times E, \mathcal{F} \otimes \mathcal{E})$ gilt:

$$y \mapsto f(x, y)$$

$$x \mapsto f(x, y)$$

sind für $x \in \Omega$ Funktionen aus $\bar{F}_+(E, \mathcal{E})$ bzw. für $y \in E$ Funktionen aus $\bar{F}_+(\Omega, \mathcal{F})$.

Außerdem sind

$$x \mapsto \int f(x, y) \nu(dy)$$

$$y \mapsto \int f(x, y) \mu(dx)$$

Funktionen aus $\bar{F}_+(\Omega, \mathcal{F})$ bzw. $\bar{F}_+(E, \mathcal{E})$ und es bestehen die Gleichheiten

$$\begin{aligned} \int f(x, y) d(\mu \otimes \nu)(x, y) &= \int \left(\int f(x, y) d\mu(x) \right) d\nu(y) \\ &= \int \left(\int f(x, y) d\nu(y) \right) d\mu(x). \end{aligned}$$

Satz (von Fubini)

Für zwei σ -endliche Maßräume $(\Omega, \mathcal{F}, \mu)$ und (E, \mathcal{E}, ν) und $f \in L^1(\Omega \times E, \mathcal{F} \otimes \mathcal{E}, \mu \otimes \nu)$ gilt:

$$y \mapsto f(x, y)$$

$$x \mapsto f(x, y)$$

sind für $x \in \Omega$ Funktionen aus $L^1(E, \mathcal{E}, \nu)$ bzw. für $y \in E$ Funktionen aus $L^1(\Omega, \mathcal{F}, \mu)$. Außerdem sind

$$x \mapsto \int f(x, y) \nu(dy)$$

$$y \mapsto \int f(x, y) \mu(dx)$$

Funktionen aus $L^1(\Omega, \mathcal{F}, \mu)$ bzw. $L^1(E, \mathcal{E}, \nu)$ und es bestehen die Gleichheiten

$$\begin{aligned} \int f(x, y) d(\mu \otimes \nu)(x, y) &= \int \left(\int f(x, y) d\mu(x) \right) d\nu(y) \\ &= \int \left(\int f(x, y) d\nu(y) \right) d\mu(x). \end{aligned}$$

Definition (Faltung von Wahrscheinlichkeitsmaßen auf \mathbb{R})

Seien μ und ν Wahrscheinlichkeitsmaße auf \mathbb{R} . Betrachte die Abbildungen

$$\alpha_n : \mathbb{R}^n \rightarrow \mathbb{R}$$

$$\alpha_n(x_1, \dots, x_n) = x_1 + \dots + x_n.$$

Das Wahrscheinlichkeitsmaß

$$\alpha_2(\mu \otimes \nu) =: \mu \star \nu$$

auf \mathbb{R} heißt *Faltung* von μ mit ν .

Bemerkung: Man zeigt leicht:

$$\mu \star \nu = \nu \star \mu$$

$$(\mu_1 \star \mu_2) \star \mu_3 = \alpha_3(\mu_1 \otimes \mu_2 \otimes \mu_3) = \mu_1 \star (\mu_2 \star \mu_3)$$

Definition (Faltungshalbgruppe)

Eine Familie $(\mu_t)_{t \in \mathbb{R}_+}$ von Wahrscheinlichkeitsmaßen auf \mathbb{R} heißt *Faltungshalbgruppe*, wenn gilt:

$$\mu_s \star \mu_t = \mu_{s+t} \quad \forall t \in \mathbb{R}_+ \quad (2.3.1)$$

$$\mu_0 = \delta_0 \quad (2.3.2)$$

Eine Faltungshalbgruppe heißt *stetig*, wenn gilt:

μ_t konvergiert gegen δ_0 schwach

Wir schreiben oft μ_t für $(\mu_t)_{t \in \mathbb{R}_+}$.

Bemerkung: Die Bedingung (2.3.2) folgt aus der Bedingung (2.3.1); siehe z. B. "Bauer".

Definiton (Lévy-Prozess)

Ein stochastischer Prozess $(X_t)_{t \in \mathbb{R}_+}$ heißt *Lévy-Prozess*, wenn gilt:

- $X_0 = 0$ f.s.
 - $X_{t_1, t_2}, \dots, X_{t_n, t_{n+1}}$ unabhängig für alle $0 \leq t_1 < t_2 < \dots < t_{n+1}$, $n \in \mathbb{N}$.
 - $P_{X_{s,t}} = P_{X_{t-s}}$ für alle $s, t \in \mathbb{R}_+$, $s \leq t$
- (Dabei steht $X_{s,t}$ für den "Zuwachs" $X_t - X_s$.)

Satz 2.3.2 Für einen Lévy-Prozess $(X_t)_{t \in \mathbb{R}_+}$ bilden die Verteilungen

$$\mu_t := P_{X_t}, \quad t \in \mathbb{R}_+,$$

eine Faltungshalbgruppe.

Wenn X_t für $t \rightarrow 0+$ stochastisch gegen 0 konvergiert, so ist $(\mu_t)_{t \in \mathbb{R}_+}$ eine stetige Faltungshalbgruppe.

Wir definieren nun:

$$\begin{aligned} \tilde{\alpha}_n &: \mathbb{R}^n \rightarrow \mathbb{R}^n \\ \tilde{\alpha}_n(x_1, \dots, x_n) &:= (x_1, x_1 + x_2, \dots, x_1 + \dots + x_n) \end{aligned}$$

Satz 2.3.3 (a) Sei μ_t eine Faltungshalbgruppe von Wahrscheinlichkeitsmaßen auf \mathbb{R} . Dann wird durch

$$P_I := \tilde{\alpha}(\mu_{t_1} \otimes \mu_{t_2-t_1} \otimes \dots \otimes \mu_{t_n-t_{n-1}}),$$

$I = \{0 \leq t_1 < \dots < t_n\} \in \mathcal{P}_0(\mathbb{R}_+)$, eine projektive Familie von Wahrscheinlichkeitsmaßen definiert.

(b) Der durch Satz 2.3.1 gegebene Prozess $(X_t)_{t \in \mathbb{R}_+}$, der zu der projektiven Familie aus (a) gehört, ist ein Lévy-Prozess mit

$$P_{X_t} = \mu_t, \quad t \in \mathbb{R}_+,$$

(c) Ist μ_t eine stetige Faltungshalbgruppe, so konvergiert X_t für den Lévy-Prozess aus (b) für $t \rightarrow 0+$ stochastisch gegen 0.

Stochastische Prozesse

SoSe 2018

Übersicht 6. Vorlesungswoche (14.5.-18.5.2018)

Beispiele: 1) (Dirac-Maße; deterministischer Prozess) Das Dirac-Maß δ_x im Punkte $x \in \mathbb{R}^d$ ist gegeben durch ($B \subset \mathbb{R}^d$)

$$\delta_x(B) = \mathbf{1}_B(x) = \begin{cases} 1 & \text{falls } x \in B \\ 0 & \text{falls } x \notin B. \end{cases}$$

δ_x ist ein Wahrscheinlichkeitsmaß auf $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ (sogar auf $(\mathbb{R}^d, \mathcal{P}(\mathbb{R}^d))$). Außerdem

$$\delta_x \star \delta_y = \delta_{x+y} \quad \forall x, y \in \mathbb{R}^d,$$

so dass z. B. $(\delta_t)_{t \in \mathbb{R}_+}$ eine Faltungshalbgruppe von Wahrscheinlichkeitsmaßen auf \mathbb{R} ist. Es gilt

$$\int f(y) d\delta_x(y) = f(x) \quad \forall f \in \mathcal{C}_b(\mathbb{R}^d),$$

so dass δ_t eine stetige Faltungshalbgruppe ist. Wenn die Zufallsvariable X die Verteilung δ_t besitzt, so bedeutet dies gerade, dass gilt: $X = t$ f.s. Also können wir den Lévy-Prozess zur Faltungshalbgruppe δ_t einfach durch den deterministischen Prozess $X_t = t$ realisieren. Der nach Satz 2.3.3 mit Hilfe der Daniell-Kolmogorovschen Konstruktion gegebene Lévy-Prozess “lebt” allerdings auf dem Raum *aller* Pfade. Dieses zugegeben etwas “pathologische” Beispiel zeigt besonders deutlich, dass der “kanonische” Prozess, den man durch den Satz 2.3.1 erhält, auf einem viel zu großen Raum definiert sein kann.

2) (Poisson-Prozess) Wir erinnern an dieser Stelle an die wichtige *Fourier-Transformation von Wahrscheinlichkeitsmaßen*. Jedem Wahrscheinlichkeitsmaß μ auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ (geht auch allgemeiner auf $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$) wird dabei seine Fourier-Transformierte

$$\hat{\mu} : \mathbb{R} \rightarrow \mathbb{C}$$

$$\hat{\mu}(x) = \int e^{ixt} d\mu(t), \quad x \in \mathbb{R}$$

zugeordnet; siehe z. B. “Bauer”. Die Funktion $\hat{\mu}$ bestimmt das Wahrscheinlichkeitsmaß μ eindeutig, d. h. die Abbildung $\mu \mapsto \hat{\mu}$ ist injektiv. Man kann zeigen:

- $\widehat{\mu \star \nu}(x) = \hat{\mu}(x)\hat{\nu}(x)$, $x \in \mathbb{R}$, für Wahrscheinlichkeitsmaße μ, ν auf \mathbb{R}
 - X, Y unabhängige Zufallsvariablen $\Rightarrow P_{X+Y} = P_X \star P_Y \Rightarrow \widehat{P_{X+Y}} = \widehat{P_X} \widehat{P_Y}$
- Für die Poisson-Verteilung π_λ mit Intensität $\lambda \in \mathbb{R}_+^*$ gilt

$$\widehat{\pi}_\lambda(x) = e^{\lambda(e^{ix} - 1)}$$

und somit

$$\widehat{\pi}_{t\lambda}(x) \widehat{\pi}_{s\lambda}(x) = e^{(t+s)\lambda(e^{ix} - 1)} = \widehat{\pi}_{(t+s)\lambda}(x).$$

Also bilden $\pi_{t\lambda}$, $t \in \mathbb{R}_+$, eine stetige (!) Faltungshalbgruppe, die so genannte *Poissonsche Faltungshalbgruppe* (zur Intensität λ). Dazu gehört wieder ein kanonischer Lévy-Prozess nach Satz 2.3.1, den man – mit zusätzlichen Forderungen über Eigenschaften der Pfade (siehe Abschnitt 3.2 unten) – als *Poisson-Prozess* bezeichnet.

3) (Brownsche Bewegung; Wiener-Prozess) Die Normalverteilung $N(\mu, \sigma^2)$ zu den Parametern $\mu \in \mathbb{R}$ (Erwartungswert; wir verwenden wieder den Buchstabe μ , da die Verwechslung mit einem Maß μ kaum möglich) und $\sigma^2 \in \mathbb{R}_+^*$ (Varianz) ist bekanntlich durch die Dichte

$$\phi(\mu, \sigma^2)(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad x \in \mathbb{R},$$

gegeben. Es gilt

$$\widehat{N}(\mu, \sigma^2)(x) = e^{i\mu x} e^{-\frac{\sigma^2 x^2}{2}}$$

und darum

$$\widehat{N}(0, t)(x) = e^{-t\frac{x^2}{2}}$$

sowie

$$\widehat{N}(0, s)(x) \widehat{N}(0, t)(x) = e^{-(s+t)\frac{x^2}{2}} = \widehat{N}(0, s+t)(x).$$

Also bilden $N(0, t)$, $t \in \mathbb{R}_+$, eine stetige (!) Faltungshalbgruppe, die *Gaußsche Faltungshalbgruppe*. Dazu gehört nach Satz 2.3.1 ebenfalls ein kanonischer Lévy-Prozess. Mit Zusatzforderungen an die Pfade (siehe Abschnitt 4.1 unten) – erhält man die Definition der *Brownschen Bewegung*, die auch als *Wiener-Prozess* bezeichnet wird.

3 Der Poisson-Prozess

3.1 Der Poisson-Prozess als Lévy-Prozess

Zur Einleitung diskutieren wir ein mögliches Modell für einen stochastischen Prozess, der sich mit der Zeit $t \geq 0$ entwickelt und der die *Anzahl der Sternschnuppen* in einer Nacht angibt:

- X_t : Anzahl der beobachteten Sternschnuppen bis zum Zeitpunkt t
- $X_0 = 0$
- $X_{st} := X_t - X_s, s \leq t$: Anzahl der beobachteten Sternschnuppen im Zeitintervall $]s, t]$

Wir nehmen an, dass

- sich die Sternschnuppen nicht gegenseitig beeinflussen
- die Anzahl der Sternschnuppen nur von der Länge des Beobachtungszeitraums abhängt.

Auftreten der Poisson-Verteilung: Wir zerlegen das Einheitsintervall in immer kleinere gleichlange Teilintervalle. Die Wahrscheinlichkeit, in einem Intervall $] \frac{k-1}{n}, \frac{k}{n}]$, $k = 1, \dots, n$, der n -ten Zerlegung eine Sternschnuppe zu beobachten, sei p_n . Wenn die Zerlegung fein genug ist, kann man annehmen, dass entweder *eine* oder *keine* Sternschnuppe beobachtet wird. Wir erhalten die Bernoulli-Zufallsvariablen $X_k^{(n)}$, $k = 1, \dots, n$, mit

$$X_k^{(n)} = \begin{cases} 1 & \text{falls Sternschnuppe im } k\text{-ten Intervall auftritt} \\ 0 & \text{falls nicht} \end{cases}$$

$$P(X_k^{(n)} = 1) = p_n, \quad P(X_k^{(n)} = 0) = 1 - p_n.$$

Nun nehmen wir noch an, dass p_n proportional zur Länge $\frac{1}{n}$ der Intervalle der n -ten Zerlegung ist (Proportionalitätsfaktor: λ):

$$p_n = \frac{\lambda}{n}, \quad np_n = \lambda.$$

Die Zufallsvariable $X^{(n)} = X_1^{(n)} + \dots + X_n^{(n)}$ gibt die Anzahl der Sternschnuppen im Einheitsintervall an. $X^{(n)}$ ist binomialverteilt mit Parametern n und p_n .

Behauptung: *Es gilt für $k \in \mathbb{N}_0$:*

$$\lim_{n \rightarrow \infty} P(X^{(n)} = k) = e^{-\lambda} \frac{\lambda^k}{k!} = \pi_\lambda(k)$$

Beweis: siehe z. B. "Klenke", Satz 3.7.

Dies führt zu der Annahme

$$X_t \sim \pi_{\lambda t}.$$

Wir fassen zusammen:

- Der Prozess X_t besitzt rechtsseitig stetige monoton wachsende Pfade mit Sprüngen der Größe 1.
- $X_0 = 0$

- $X_{st} \sim \pi_{\lambda(t-s)}$
- $X_{t_1}, X_{t_1, t_2}, \dots, X_{t_{n+1}, t_n}$ unabhängig für $0 \leq t_1 < \dots < t_{n+1}$

Wir werden sehen, dass es (genau) einen Prozess mit diesen Eigenschaften gibt, nämlich den Poisson-Prozess (mit Intensität λ). Wir wissen schon, wie wir einen Prozess mit den letzten drei Eigenschaften der Liste konstruieren können: Mit der Poissonschen Faltungshalbgruppe $\pi_{\lambda t}$ ergibt sich ein kanonischer Lévy-Prozess aus der projektiven Familie, die zur dieser Faltungshalbgruppe gehört. Die letzten drei Bedingungen sind für einen Lévy-Prozess immer erfüllt. Im nächsten Abschnitt geben wir eine Konstruktion des Poisson-Prozesses an.

3.2 Der Poisson-Prozess als Sprungprozess

Bemerkung: Allgemein gilt: Bei vorgegebenen Verteilungen μ_t , $t \in T$, kann man stets einen Wahrscheinlichkeitsraum (Ω, \mathcal{F}, P) und Zufallsvariablen X_t , $t \in T$, konstruieren, so dass:

- $(X_t)_{t \in T}$ unabhängig
- $X_t \sim \mu_t$

Geht mit:

·

$$(\Omega, \mathcal{F}) = (\mathbb{R}^T, \mathcal{B}(\mathbb{R}^T)) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))^{\otimes T}$$

- P: Produktmaß $\bigotimes_{t \in T} \mu_t$ der Wahrscheinlichkeitsmaße μ_t auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$
- $X_t((\alpha_s)_{s \in T}) = \alpha_t$ (Projektion auf die t -Koordinate)

Definition (Exponentialverteilung)

Die *Exponentialverteilung zum Parameter* $\lambda \in \mathbb{R}_+$ ist das durch die Wahrscheinlichkeitsdichte

$$x \mapsto \begin{cases} \lambda e^{-\lambda x} & \text{für } x \geq 0 \\ 0 & \text{für } x < 0 \end{cases} \quad (3.2.1)$$

gegebene Wahrscheinlichkeitsmaß E_λ auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

Bemerkung: Man rechnet sofort nach, dass durch (3.2.1) wirklich eine Wahrscheinlichkeitsdichte definiert wird. E_λ ist auf \mathbb{R}_+ konzentriert. Es gilt $P(X \leq t) = e^{-\lambda t}$ für eine nach E_λ verteilte Zufallsvariable X .

Nach der obigen Eingangsbemerkung ist die Existenz einer Folge

$$S_1, S_2, \dots$$

identisch verteilter unabhängiger Zufallsvariablen mit

$$S_1 \sim E_\lambda$$

gesichert.

Definition (Poissonscher Sprungprozess)

Für $\lambda \in \mathbb{R}_+^*$ und eine gegebene Folge S_1, S_2, \dots wie oben setzen wir

$$\begin{aligned} J_0 &:= 0 \\ J_k &:= S_1 + \dots + S_k, \quad k \in \mathbb{N}, \end{aligned}$$

und

$$N_t := \sum_{k=1}^{\infty} \mathbf{1}_{]0,t]}(J_k).$$

Der (numerische) Prozess N_t heißt *Poissonscher Sprungprozess* oder einfach *Poisson-Prozess* zur Intensität λ .

Bemerkungen: 1) Offenbar gilt

$$N_t(\omega) = n \in \mathbb{N}_0 \Leftrightarrow J_n(\omega) \leq t < J_{n+1}(\omega),$$

so dass die N_t jedenfalls numerische Zufallsvariablen, d. h. messbar bzgl. $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, sind.

2) Folgende Überlegung zeigt, dass für alle $t \in \mathbb{R}_+$ f.s. $N_t < \infty$ gelten muss.

Behauptung:

$$P(N_t = \infty) = 0 \quad \forall t \in \mathbb{R}_+$$

Beweis: Es gilt:

$$\begin{aligned} \{N_t = \infty\} &= \bigcap_{n \in \mathbb{N}} \{J_n \leq t\} \\ &= \bigcap_{n \in \mathbb{N}} \{S_1 + \dots + S_n \leq t\} \\ &= \lim_{n \rightarrow \infty} P(A_n) \end{aligned}$$

mit

$$A_n := \{S_1 + \dots + S_n \leq t\}.$$

Die n -fache Faltung der Exponentialverteilung E_λ ergibt aber die Gamma-Verteilung $\Gamma(\lambda, n)$, die wiederum durch die Wahrscheinlichkeitsdichte

$$x \mapsto \frac{\lambda^n}{(n-1)!} x^{n-1} e^{-\lambda x}$$

gegeben ist; siehe z. B. "Klenke". Also folgt

$$P(A_n) = (E_\lambda)^{*n}[0, t] = \int_0^t \frac{\lambda^n}{(n-1)!} x^{n-1} e^{-\lambda x} dx.$$

Der letzte Ausdruck strebt für $n \rightarrow \infty$ gegen 0. \square

Wir fassen zusammen:

- $J_k(\omega)$: Zeitpunkt des k -ten Sprunges; Sprünge besitzen die Größe 1
- $S_k(\omega)$: "Wartezeit" zwischen dem $(k-1)$ -ten und dem k -ten Sprung
- $N_t(\omega)$: Anzahl der Sprünge bis zum Zeitpunkt t (einschließlich)
- $\{N_t = k\} = \{J_k \leq t < J_{k+1}\}$

Offenbar gilt:

- Die Pfade von N_t sind monoton wachsend.
- Ein Pfad beginnt bei 0.
- Es gibt nur Sprünge der Größe 1 nach oben.
- Zwischen zwei Sprüngen ist der Pfad konstant.
- Die Pfade sind càdlàg.

Satz 3.2.1 $(N_t)_{t \geq 0}$ ist ein Lévy-Prozess. Seine Faltungshalbgruppe ist durch die Poissonsche Faltungshalbgruppe zur Intensität λ gegeben.

Beweis: Man überzeugt sich, dass es genügt zu zeigen:

$$\begin{aligned} & P(N_{t_2} - N_{t_1} = k_1, N_{t_3} - N_{t_2} = k_2, \dots, N_{t_{n+1}} - N_{t_n} = k_n) \\ & = \pi_{\lambda(t_2-t_1)}(k_1) \dots \pi_{\lambda(t_{n+1}-t_n)}(k_n) \end{aligned}$$

für alle $n \in \mathbb{N}$, $0 \leq t_1 < t_2 < \dots < t_{n+1}$, $k_1, \dots, k_n \in \mathbb{N}_0$. Das ist dann eine längere Rechnung ... \square

Stochastische Prozesse
SoSe 2018
Übersicht 7. Vorlesungswoche (28.5.-1.6.2018)

4 Die Brownsche Bewegung

4.1 Die Brownsche Bewegung als Lévy-Prozess

Definition (stetiger stochastischer Prozess)

Wir nennen einen stochastischen Prozess $(X_t)_{t \in \mathbb{R}_+}$ *stetig*, wenn alle seine Pfade

$$t \mapsto X_t(\omega)$$

$(\omega \in \Omega)$ stetige Funktionen von \mathbb{R}_+ nach \mathbb{R} sind.

Bemerkung: Ist $(X_t)_{t \in \mathbb{R}_+}$ ein stochastischer Prozess, für den es eine Menge $\tilde{\Omega} \subset \Omega$ gibt, $\tilde{\Omega} \in \mathcal{F}$, mit

- $\omega \in \tilde{\Omega} \Rightarrow t \mapsto X_t(\omega)$ stetig
- $P(\tilde{\Omega}) = 1$,

dann gilt für den stochastischen Prozess $(\tilde{X}_t)_{t \in \mathbb{R}_+}$ mit

$$\begin{aligned}\tilde{X}_t &: \tilde{\Omega} \rightarrow \mathbb{R} \\ \tilde{X}_t &= X_t \upharpoonright \tilde{\Omega}\end{aligned}$$

über dem Wahrscheinlichkeitsraum $(\tilde{\Omega}, \mathcal{F} \upharpoonright \tilde{\Omega}, P \upharpoonright \tilde{\Omega})$:

$$\begin{aligned}P(\tilde{X}_{t_1} \in B_1, \dots, \tilde{X}_{t_n} \in B_n) &= P(\{X_{t_1} \in B_1, \dots, X_{t_n} \in B_n\} \cap \tilde{\Omega}) \\ &= P(X_{t_1} \in B_1, \dots, X_{t_n} \in B_n)\end{aligned}$$

für alle $n \in \mathbb{N}$, $t_1, \dots, t_n \in \mathbb{R}_+$, $B_1, \dots, B_n \in \mathcal{B}(\mathbb{R})$, d. h. \tilde{X}_t ist eine *stetige Version* von X_t .

Es gilt das folgende Kriterium für die Existenz einer stetigen Version eines gegebenen stochastischen Prozesses mit Indexmenge \mathbb{R}_+ :

Satz 4.1.1 (Stetigkeitssatz von Kolmogorov; z. B. “Klenke” Satz 21.6)
Sei $(X_t)_{t \in \mathbb{R}_+}$ ein stochastischer Prozess mit der folgenden Eigenschaft:

Für alle $T \in \mathbb{R}_+^*$ existieren Konstanten $\alpha, \beta, C \in \mathbb{R}_+^*$, so dass

$$\mathbb{E}(|X_t - X_s|^\alpha) \leq C |t - s|^{1+\beta} \quad \forall s, t \in [0, T]. \quad (4.1.1)$$

Dann besitzt X_t eine stetige Version.

Korollar 4.1.2 Der kanonische Lévy-Prozess zur Gaußschen Faltungshalbgruppe $(N(0, t))_{t \in \mathbb{R}_+}$ (siehe Beispiel 3 in Abschnitt 2.3) besitzt eine stetige Version.

Beweis: Die Bedingung (4.1.1) aus Satz 4.1.1 ist mit $\alpha = 4$, $C = 3$ und $\beta = 1$ erfüllt, da

$$\mathbb{E}(|X_t - X_s|^4) = 3(t - s)^2. \square$$

Definition (Brownsche Bewegung)

Ein stochastischer Prozess $(B_t)_{t \in \mathbb{R}_+}$ heißt *Brownsche Bewegung* oder *Wiener-Prozess*, wenn gilt:

- B_t ist ein Lévy-Prozess mit Faltungshalbgruppe $N(0, t)$.
- B_t ist stetig.

Nach Korollar 4.1.2 existiert eine Brownsche Bewegung!

Bemerkung: Später betrachten wir auch Fälle, in denen die Indexmenge nur eine Teilmenge M von \mathbb{R}_+ ist. Dann nennen wir $(X_t)_{t \in M}$ eine Brownsche Bewegung *auf* M , wenn gilt:

- $X_0 = 0$
- $X_{t_2} - X_{t_1}, \dots, X_{t_{n+1}} - X_{t_n}$ unabhängig für alle $t_1 < \dots < t_{n+1}, t_1, \dots, t_n \in M$
- $X_t - X_s \sim N(0, t - s)$ für alle $s < t, s, t \in M$
- Die Funktionen $t \rightarrow X_t(\omega)$ von M nach \mathbb{R} sind stetig für alle $\omega \in \Omega$.

(Wenn M diskret ist, dann ist die Stetigkeitsbedingung natürlich immer erfüllt.)

4.2 P. Lévy's pfadweise Konstruktion der Brownschen Bewegung

Plan: Setze

$$D_n := \left\{ \frac{k}{2^n} \mid k \in \mathbb{N}_0 \right\}$$

$$D := \bigcup_{n \in \mathbb{N}_0} D_n.$$

D_{n+1} entsteht dann aus D_n durch Hinzunahme der Mittelpunkte zweier benachbarter Punkte. Es ist z. B.

$$\begin{aligned} D_0 &= \mathbb{N}_0 \\ D_1 &= \{0, \frac{1}{2}, 1, \frac{3}{2}, 2, \dots\} \\ D_2 &= \{0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1, \frac{5}{4}, \frac{3}{2}, \dots\}. \end{aligned}$$

Ausgangspunkt der Konstruktion nach P. Lévy ist eine durch die Menge D indizierte Familie $(Y_t)_{t \in D}$ unabhängiger Zufallsvariablen mit Standard-Normalverteilung. Mit Hilfe dieser Familie werden induktiv Brownsche Bewegungen auf D_n definiert. Setzt man nun die Brownschen Bewegungen auf D_n durch *lineare Interpolation* zu stochastischen Prozessen mit Indexmenge \mathbb{R}_+ fort, so kann man zeigen, dass mit Wahrscheinlichkeit 1 diese Prozesse auf jedem beschränkten Intervall gleichmäßig gegen eine Brownsche Bewegung konvergieren.

Wir benötigen noch eine spezielle Eigenschaft von ‘‘Gauß-Systemen’’.

Definition (Gaußsystem)

Eine Familie $(X_i)_{i \in I}$ von Zufallsvariablen heißt *Gauß-System*, wenn jede reelle Linearkombination der X_i eine Gauß-Verteilung $N(\mu, \sigma^2)$ besitzt. (Dabei lassen wir auch $N(\mu, 0) = \delta_\mu$ zu.)

Bemerkungen: 1) Für ein Gauß-System gilt also:

$$\begin{aligned} n \in \mathbb{N}, \alpha_1, \dots, \alpha_n \in \mathbb{R}, i_1, \dots, i_n \in I \\ \Rightarrow \exists (\mu, \sigma^2) \in \mathbb{R} \times \mathbb{R}_+ \text{ mit } \sum_{l=1}^n \alpha_l X_{i_l} \sim N(\mu, \sigma^2) \end{aligned}$$

2) Teilsysteme von Gauß-Systemen sind offenbar wieder Gauß-Systeme.

Beispiel: Sei $(Y_t)_{t \in M}$ (wobei M irgendeine Indexmenge bezeichne) eine Familie von Zufallsvariablen mit

- $(Y_t)_{t \in M}$ unabhängig
 - $Y_t, t \in M$, Gauß-verteilt, d. h. $Y_t \sim N(\mu_t, \sigma_t^2)$ für $(\mu_t, \sigma_t^2) \in \mathbb{R} \times \mathbb{R}_+$ geeignet.
- Dann ist die Familie

$$\left(\sum_{t \in I} \alpha_t Y_t \right)_{(I, \alpha)},$$

wobei $I \in \mathcal{P}_0(M)$ und $\alpha = (\alpha_t)_{t \in I}, \alpha_t \in \mathbb{R}$, ein Gauß-System von Zufallsvariablen.

Satz 4.2.1 Ein Gauß-System $(X_i)_{i \in I}$ ist genau dann unabhängig, wenn gilt

$$E((X_i - EX_i)(X_j - EX_j)) = 0 \quad \forall i, j \in I \text{ mit } i \neq j,$$

d. h. wenn die X_i paarweise unkorreliert sind.

Nun zurück zur pfadweisen Konstruktion der Brownschen Bewegung.

Definition Gegeben sei eine Familie $(Y_t)_{t \in D}$ unabhängiger Zufallsvariablen mit $Y_t \sim N(0, 1)$ für alle $t \in D$. (Eine solche Familie existiert nach der Bemerkung zu Beginn des Abschnitts 3.2.)

Wir definieren die Prozesse $(B_t^{(n)})_{t \in D_n}$, $n \in \mathbb{N}_0$, induktiv:

$$\begin{aligned} B_0^{(0)} &:= 0 \\ B_t^{(0)} &:= Y_1 + Y_2 + \dots + Y_t \text{ für } t \in D_0 = \mathbb{N}_0 \\ &\vdots \\ B_t^{(n+1)} &:= B_t^{(n)} \text{ für } t \in D_n \\ B_t^{(n+1)} &:= \frac{1}{2} \left(B_{t-\frac{1}{2^{n+1}}}^{(n)} + B_{t+\frac{1}{2^{n+1}}}^{(n)} \right) + Z_t \text{ für } t \in D_{n+1} \setminus D_n \end{aligned}$$

Dabei ist

$$Z_t := \frac{Y_t}{\sqrt{2^{n+2}}} = 2^{-\frac{n+2}{2}} Y_t.$$

Satz 4.2.2 $(B_t^{(n)})_{t \in D_n}$ ist eine Brownsche Bewegung auf D_n .

Beweis: Wir führen den Beweis mittels Induktion über $n \in \mathbb{N}_0$.

$n = 0$: Es ist $B_0^{(0)} = 0$. Für den Zuwachs von s nach t gilt:

$$B_t^{(0)} - B_s^{(0)} = Y_{s+1} + \dots + Y_t,$$

weswegen die Unabhängigkeit der Zuwächse aus der Unabhängigkeit der Y_1, Y_2, \dots folgt. Da außerdem $Y_{s+1} + \dots + Y_t \sim N(0, t - s)$ gilt, haben wir nachgewiesen, dass $B_t^{(0)}$ eine Brownsche Bewegung auf \mathbb{N}_0 ist.

$n \rightarrow n + 1$: Wir nehmen an, dass $(B_t^{(n)})_{t \in D_n}$ eine Brownsche Bewegung auf D_n ist.

Behauptung: $(B_t^{(n+1)})_{t \in D_{n+1}}$ hat dann unabhängige Zuwächse.

Beweis davon: Offenbar genügt es wegen

$$\begin{aligned} B_t^{(n+1)} - B_s^{(n+1)} &= (B_t^{(n+1)} - B_{t-\frac{1}{2^{n+1}}}^{(n+1)}) + (B_{t-\frac{1}{2^{n+1}}}^{(n+1)} - B_{t-\frac{2}{2^{n+1}}}^{(n+1)} + \dots + (B_{s+\frac{1}{2^{n+1}}}^{(n+1)} - B_s^{(n+1)})) \end{aligned}$$

die Unabhängigkeit der Zuwächse

$$B_{\frac{1}{2^{n+1}}}^{(n+1)}, B_{\frac{2}{2^{n+1}}}^{(n+1)} - B_{\frac{1}{2^{n+1}}}^{(n+1)}, B_{\frac{3}{2^{n+1}}}^{(n+1)} - B_{\frac{2}{2^{n+1}}}^{(n+1)}, \dots \quad (4.2.1)$$

der Spanne 1 nachzuweisen. Diese Zuwächse bilden ein Teilsystem des Systems

$$\left(\sum_{t \in I} \alpha_t Y_t \right)_{(I, \alpha)},$$

welches nach obigem Beispiel wegen der Unabhängigkeit der $(Y_t)_{t \in D}$ ein Gauß-System ist. Nach Satz 4.2.1 genügt es zum Nachweis der Unabhängigkeit von (4.2.1) die Gültigkeit von

$$\mathbb{E}\left(\left(B_{\frac{k+1}{2^{n+1}}}^{(n+1)} - B_{\frac{k}{2^{n+1}}}^{(n+1)} \right) \left(B_{\frac{l+1}{2^{n+1}}}^{(n+1)} - B_{\frac{l}{2^{n+1}}}^{(n+1)} \right) \right) = 0 \quad (4.2.2)$$

für alle $k, l \in \mathbb{N}_0$, $k < l$, nachzuweisen. Für $k+1 < l$ folgt (4.2.1) mit Induktion daraus, dass die Zuwächse aus den entsprechenden Zuwächsen der Stufe n und *zwei verschiedenen* Zufallsvariablen aus den Y_t , $t \in D_{n+1} \setminus D_n$ aufgebaut sind. Einzig der Fall der Zuwächse der Form

$$B_r^{(n+1)} - B_s^{(n+1)}, B_t^{(n+1)} - B_r^{(n+1)}$$

mit $r \in D_{n+1}$ und $s = r - \frac{1}{2^{n+1}}$, $t = r + \frac{1}{2^{n+1}}$ bleibt zu untersuchen ...

Stochastische Prozesse

SoSe 2018

Übersicht 8. Vorlesungswoche (4.6.-8.6.2018)

Beweis von Satz 4.2.2, Fortsetzung: Jetzt verwenden wir zum ersten Mal die Wahl des Faktors $2^{-\frac{n+2}{2}}$ bei der Definition von Z_t .

Mit

$$s = r - \frac{1}{\sqrt{2^{n+1}}}, \quad t = r + \frac{1}{\sqrt{2^{n+1}}}$$

erhalten wir:

$$\begin{aligned} & \mathbb{E}((B_r^{(n+1)} - B_s^{(n+1)})(B_t^{(n+1)} - B_r^{(n+1)})) \\ &= \mathbb{E}\left(\left(\frac{1}{2}(B_t^{(n)} - B_s^{(n)}) + Z_r\right)\left(\frac{1}{2}(B_t^{(n)} - B_s^{(n)}) - Z_r\right)\right) \\ &= \frac{1}{4}\mathbb{E}((B_t^{(n)} - B_s^{(n)})^2) - \mathbb{E}(Z_r^2) \\ &= \frac{1}{4}(t - s) - \frac{1}{2^{n+2}} = 0 \end{aligned}$$

Zur Stationarität der Zuwächse von $B_t^{(n+1)}$: Es gilt:

$$\begin{aligned} B_r^{(n+1)} - B_s^{(n+1)} &= \frac{1}{2}(B_t^{(n)} - B_s^{(n)}) + Z_r \\ &\sim N\left(0, \frac{1}{4} \frac{1}{2^n} + \frac{1}{2^{n+2}}\right) \\ &= N\left(0, \frac{1}{2^{n+1}}\right) = N(0, r - s) \end{aligned}$$

und

$$\begin{aligned} B_t^{(n+1)} - B_r^{(n+1)} &= \frac{1}{2}(B_t^{(n)} - B_s^{(n)}) - Z_r \\ &\sim N\left(0, \frac{1}{2^{n+1}}\right) = N(0, t - r) \square \end{aligned}$$

Wir dehnen nun jeden der Prozesse $(B_t^{(n)})_{t \in D_n}$ zu einem Prozess $(B_t^{(n)})_{t \in \mathbb{R}_+}$ aus, indem wir linear interpolieren:

Definition Sei $t \in \mathbb{R}_+ \setminus D_n$. Dann gibt es ein eindeutiges $k \in \mathbb{N}_0$, so dass

$$\frac{k}{2^n} < t < \frac{k+1}{2^n}.$$

Wir setzen

$$B_t^{(n)}(\omega) := B_{\frac{k}{2^n}}^{(n)}(\omega) + \left(t - \frac{k}{2^n}\right) 2^n \left(B_{\frac{k+1}{2^n}}^{(n)}(\omega) - B_{\frac{k}{2^n}}^{(n)}(\omega)\right)$$

und

$$Z_t^{(n+1)} := B_t^{(n+1)} - B_t^{(n)}.$$

Behauptung: Es gilt $Z_t^{(n+1)} = Z_t$ für alle $t \in D_{n+1} - D_n$. Außerdem:

$$\sup\{|Z_t^{(n+1)}| \mid t \in [0, 1]\} = \sup\{|Z_t^{(n+1)}| \mid t \in (D_{n+1} \setminus D_n) \cap [0, 1]\} \quad (4.2.1)$$

(Gleichheit (4.2.1) macht man sich anhand einer Skizze klar!)

Wir setzen ($n \in \mathbb{N}_0$)

$$M_{n+1} := \sup\{|Z_t^{(n+1)}| \mid t \in [0, 1]\}.$$

Lemma 4.2.3

$$\mathbb{E}\left(\sum_{k=1}^{\infty} M_k\right) < \infty$$

Beweis: Geht mit Hilfe der Gleichung

$$\mathbb{E}(X^p) = \int_0^{\infty} p\lambda^{p-1} \mathbb{P}(X > \lambda) \, d\lambda, \quad (4.2.2)$$

die allgemein für Zufallsvariablen $X \geq 0$ und für $p \in \mathbb{R}, p > 1$ gilt. Außerdem:

$$\begin{aligned} \mathbb{P}(M_{n+1} > 2^{-\frac{n+2}{2}} \lambda) &= \mathbb{P}\left(\bigcup_{x \in (D_{n+1} \setminus D_n) \cap [0, 1]} \{|Y_x| > \lambda\}\right) \\ &\leq \sum_{x \in (D_{n+1} \setminus D_n) \cap [0, 1]} \mathbb{P}(|Y_x| > \lambda) \\ &= 2^n \mathbb{P}(|Y_1| > \lambda) \end{aligned} \quad (4.2.3)$$

Mit $X = 2^{\frac{n+2}{2}} M_{n+1}$ ergibt sich aus (4.2.2) und (4.2.3):

$$2^{\frac{p(n+2)}{2}} \mathbb{E}(M_{n+1}^p) \leq 2^n \mathbb{E}(|Y_1|^p) \quad (4.2.4)$$

Jetzt folgt die Abschätzung ($p > 1$)

$$\begin{aligned} \mathbb{E}\left(\sum_{k=1}^{\infty} M_k\right) &= \sum_{k=1}^{\infty} \mathbb{E}(M_k) \\ &\leq \sum_{k=1}^{\infty} \mathbb{E}(M_k^p)^{\frac{1}{p}} \\ &\leq (\mathbb{E}(|Y_1|^p))^{\frac{1}{p}} \sum_{k=1}^{\infty} \left(\frac{1}{2}\right)^{\frac{k+1}{2} - \frac{k-1}{p}}, \end{aligned}$$

wobei die erste Ungleichung sich aus der Hölderschen Ungleichung $\mathbb{E}(M_k) = \mathbb{E}(M_k \mathbf{1}) \leq \mathbb{E}(M_k^p)^{\frac{1}{p}} \mathbb{E}(\mathbf{1}^q)^{\frac{1}{q}}$ ergibt. Für $p = 3$ erhalten wir

$$\mathbb{E}\left(\sum_{k=1}^{\infty} M_k\right) \leq (\mathbb{E}(|Y_1|^p))^{\frac{1}{p}} \left(\frac{1}{2}\right)^{\frac{5}{6}} \sum_{k=1}^{\infty} \left(\frac{1}{2^{\frac{1}{6}}}\right)^k < \infty. \square$$

Satz 4.2.4 (Konvergenz der Interpolationsprozesse)

$B_t^{(n)}$ konvergiert gleichmäßig in $t \in [0, 1]$ mit Wahrscheinlichkeit 1.

Genauer: Es gibt eine Menge $\tilde{\Omega}$ mit $\mathbb{P}(\tilde{\Omega}) = 1$, so dass für alle $\omega \in \tilde{\Omega}$ die Folge $B_t^{(n)}(\omega)$ eine Cauchy-Folge ist.

Beweis: Wir zeigen, dass das Ereignis

$$A := \bigcap_{l \in \mathbb{N}} \bigcup_{N \in \mathbb{N}} \bigcap_{n, m \geq N} \left\{ \sup_{t \in [0, 1]} |B_t^{(n)} - B_t^{(m)}| \leq \frac{1}{l} \right\}$$

die Wahrscheinlichkeit 1 hat.

Es gilt für $n, m \geq N$

$$\begin{aligned} |B_t^{(n)} - B_t^{(m)}| &= \left| \sum_{k=(n \wedge m)+1}^{n \vee m} Z_t^{(k)} \right| \\ &\leq \sum_{k=N}^{\infty} |Z_t^{(k)}| \end{aligned}$$

und damit

$$\begin{aligned} \sup_{t \in [0, 1]} |B_t^{(n)} - B_t^{(m)}| &\leq \sum_{k=N}^{\infty} \sup_{t \in [0, 1]} |Z_t^{(k)}| \\ &= \sum_{k=N}^{\infty} M_k. \end{aligned}$$

Nun setzen wir Lemma 4.2.3 ein:

$$\mathbb{E}\left(\sum_{k=1}^{\infty} M_k\right) < \infty \Rightarrow \mathbb{P}\left(\sum_{k=1}^{\infty} M_k < \infty\right) = 1.$$

Wegen

$$\bigcap_{l \in \mathbb{N}} \bigcup_{N \in \mathbb{N}} \left\{ \sum_{k=N}^{\infty} M_k \leq \frac{1}{l} \right\} = \left\{ \sum_{k=1}^{\infty} M_k < \infty \right\}$$

und weil für $n, m \geq N$

$$\left\{ \sum_{k=N}^{\infty} M_k \leq \frac{1}{l} \right\} \subset \left\{ \sup_{t \in [0,1]} |B_t^{(n)} - B_t^{(m)}| \leq \frac{1}{l} \right\},$$

folgt

$$\mathbb{P}(A) \geq \mathbb{P}\left(\bigcap_{l \in \mathbb{N}} \bigcup_{N \in \mathbb{N}} \left\{ \sum_{k=N}^{\infty} M_k \leq \frac{1}{l} \right\}\right) = 1. \square$$

Ausdehnung auf \mathbb{R}_+ : Zunächst gilt genauso für $[0, T]$, $T \in \mathbb{R}_+$, dass $B_t^{(n)}$ auf $[0, T]$ gleichmäßig mit Wahrscheinlichkeit 1 konvergiert. Also: Für $T \in \mathbb{R}_+$ existiert $\tilde{\Omega}_T \in \mathcal{F}$ mit $\mathbb{P}(\tilde{\Omega}_T) = 1$, so dass

$$\omega \in \tilde{\Omega}_T \Rightarrow B_t^{(n)}(\omega) \text{ konvergiert gleichmäßig auf } [0, T].$$

Mit $\tilde{\Omega} := \bigcap_{N \in \mathbb{N}} \tilde{\Omega}_N$ gilt dann:

$$\omega \in \tilde{\Omega} \Rightarrow B_t^{(n)}(\omega) \text{ konvergiert gleichmäßig auf } [0, T] \quad \forall T \in \mathbb{R}_+.$$

Dann ist ($\omega \in \tilde{\Omega}$)

$$B_t(\omega) := \lim_{n \rightarrow \infty} B_t^{(n)}(\omega)$$

ein stetiger Prozess über $(\tilde{\Omega}, \mathcal{F} \upharpoonright \tilde{\Omega}, \mathbb{P} \upharpoonright \tilde{\Omega})$.

Satz 4.2.5 *Der stochastische Prozess $(B_t)_{t \in \mathbb{R}_+}$ (über $(\tilde{\Omega}, \mathcal{F} \upharpoonright \tilde{\Omega}, \mathbb{P} \upharpoonright \tilde{\Omega})$) ist eine Brownsche Bewegung.*

Beweis: Die Stetigkeit von B_t ist bereits gezeigt. Wir müssen nur noch zeigen, dass die Zuwächse die gewünschte gemeinsame Verteilung besitzen, d. h. dass gilt

$$\mathbb{P}_{(B_{t_2-t_1}, \dots, B_{t_{m+1}-B_{t_m}})} = N(0, t_2 - t_1) \otimes \dots \otimes N(0, t_{m+1} - t_m) \quad (4.2.5)$$

für alle $m \in \mathbb{N}$, $0 \leq t_1 < \dots < t_{m+1}$.

Wähle für $k = 1, \dots, m + 1$ Folgen $(t_k^{(n)})_{n \in \mathbb{N}}$ in \mathbb{R} mit

· $t_k^{(n)} \rightarrow t_k$ für $n \rightarrow \infty$

· $t_k^{(n)} \in D = \bigcup_{l \in \mathbb{N}} D_l$.

Zunächst gilt

$$B_{t_k^{(n)}} \rightarrow B_{t_k} \text{ punktweise}$$

nach der Definition von B_t . Daraus folgt

$$(B_{t_2^{(n)}} - B_{t_1^{(n)}}, \dots, B_{t_{m+1}^{(n)}} - B_{t_m^{(n)}}) \rightarrow (B_{t_2} - B_{t_1}, \dots, B_{t_{m+1}} - B_{t_m}) \text{ punktweise.}$$

Da die f. s. Konvergenz die Konvergenz in Verteilung zur Folge hat, erhalten wir:

$$P_{(B_{t_2^{(n)}} - B_{t_1^{(n)}}), \dots, B_{t_{m+1}^{(n)}} - B_{t_m^{(n)}})} \rightarrow P_{(B_{t_2} - B_{t_1}), \dots, B_{t_{m+1}} - B_{t_m}} \text{ schwach.}$$

Wir können den Limes der Verteilungen

$$P_{(B_{t_2^{(n)}} - B_{t_1^{(n)}}), \dots, B_{t_{m+1}^{(n)}} - B_{t_m^{(n)}}} \tag{4.2.6}$$

noch auf eine andere Weise berechnen. Da die $t_k^{(n)} \in D_N$, $n \in \mathbb{N}$, $k = 1, \dots, m + 1$, für N groß genug und weil $B_t^{(N)}$, $t \in D_N$, eine Brownsche Bewegung auf D_N ist, muss gelten, dass (4.2.6) gleich

$$N(0, t_2^{(n)} - t_1^{(n)}) \otimes \dots \otimes N(0, t_{m+1}^{(n)} - t_m^{(n)}) \tag{4.2.7}$$

ist. Nun argumentiert man mit den Verteilungsfunktionen (oder den Fourier-Transformierten) von (4.2.7) um zu sehen, dass (4.2.7) für $n \rightarrow \infty$ schwach gegen

$$N(0, t_2 - t_1) \otimes \dots \otimes N(0, t_{m+1} - t_m)$$

konvergiert.

Zusammen folgt jetzt (4.2.5). \square

Stochastische Prozesse

SoSe 2018

Übersicht 9. Vorlesungswoche (11.6.-15.6.2018)

Als nächstes wollen wir zeigen, dass die Brownsche Bewegung unter den stochastischen Prozessen eine ähnlich *zentrale Rolle* spielt wie die Normalverteilung unter den Verteilungen (Zentraler Grenzwertsatz). Zuerst einige wiederholende Anmerkungen zum Begriff "schwache Konvergenz".

Bemerkungen: 1) (schwache Konvergenz, mehrdimensional)

Definition Seien μ_n, μ Wahrscheinlichkeitsmaße (= Verteilungen) auf dem messbaren Raum $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$, $d \in \mathbb{N}$. Wir sagen, dass μ_n (für $n \rightarrow \infty$) *schwach gegen μ konvergiert*, wenn gilt:

$$\int f d\mu_n \xrightarrow{n \rightarrow \infty} \int f d\mu \quad \forall f \in \mathcal{C}_b(\mathbb{R}^d),$$

wobei

$$\mathcal{C}_b(\mathbb{R}^d) = \{f : \mathbb{R}^d \rightarrow \mathbb{R} \mid f \text{ stetig und beschränkt}\}.$$

Für Zufallsvariablen $X_n, X : \Omega \rightarrow \mathbb{R}^d$ sagen wir, dass X_n *in Verteilung gegen X konvergiert*, wenn die gemeinsamen Verteilungen $RP_{(X_n^{(1)}, \dots, X_n^{(d)})}$ schwach gegen die gemeinsame Verteilung $RP_{(X^{(1)}, \dots, X^{(d)})}$ konvergieren.

2) (Konvergenz im p -ten Mittel)

Definition Für Zufallsvariablen $X_n, X : \Omega \rightarrow \mathbb{R}^d$ und $p \in \mathbb{R}_+^*$ sagen wir, dass X_n *im p -ten Mittel gegen X konvergiert*, wenn gilt

$$\|X_n - X\|_p \xrightarrow{n \rightarrow \infty} 0,$$

wobei die p -Norm einer Zufallsvariable $Y : \Omega \rightarrow \mathbb{R}^d$ durch

$$\|Y\|_p = \left(\int \|Y\|^p dP \right)^{\frac{1}{p}}$$

definiert wird. Hier bezeichnet $\|Y\|$ die euklidische Norm, d. h.

$$\|Y\|^2(\omega) = (Y^{(1)}(\omega))^2 + \dots + (Y^{(d)}(\omega))^2.$$

Im Falle $p = 2$ spricht man von *Konvergenz im quadratischen Mittel*.

3) (Stetigkeitssatz von Paul Lévy, mehrdimensional)

Definition Die *Fourier-Transformierte* $\widehat{\mu} : \mathbb{R}^d \rightarrow \mathbb{C}$ eines Wahrscheinlichkeitsmaßes μ auf $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$, $d \in \mathbb{N}$ wird durch

$$\widehat{\mu}(t_1, \dots, t_d) = \int e^{i\langle t, s \rangle} d\mu(s),$$

wobei für $t = (t_1, \dots, t_d)$ und $s = (s_1, \dots, s_d)$ mit $\langle t, s \rangle$ das kanonische Skalarprodukt von t mit s bezeichnet wird. Es gilt (siehe z. B. "Bauer"):

Satz (Stetigkeitssatz von Paul Lévy)

$$\mu_n \xrightarrow[n \rightarrow \infty]{} \mu \text{ schwach} \Leftrightarrow \widehat{\mu}_n(t) \xrightarrow[n \rightarrow \infty]{} \widehat{\mu}(t) \quad \forall t \in \mathbb{R}^d.$$

Die Fourier-Transformierte von P_X wird mit φ_X bezeichnet und heißt auch *charakteristische Funktion* der Zufallsvariablen $X : \Omega \rightarrow \mathbb{R}^d$. Nach dem Satz gilt

Satz (Stetigkeitssatz von Paul Lévy; Version für Zufallsvariablen)

$$X_n \xrightarrow[n \rightarrow \infty]{} X \text{ in Verteilung} \Leftrightarrow \varphi_{X_n}(t) \xrightarrow[n \rightarrow \infty]{} \varphi_X(t) \quad \forall t \in \mathbb{R}^d.$$

4) Als eine Folgerung aus dem Stetigkeitssatz erhalten wir:

Satz Seien für jedes $n \in \mathbb{N}$ unabhängige Zufallsvariablen

$$X_n^{(1)}, \dots, X_n^{(d)} \tag{4.2.1}$$

gegeben.

Des Weiteren seien

$$X^{(1)}, \dots, X^{(d)} \tag{4.2.2}$$

unabhängige Zufallsvariablen.

Dann gilt mit

$$\begin{aligned} X_n &:= (X_n^{(1)}, \dots, X_n^{(d)}) \\ X &:= (X^{(1)}, \dots, X^{(d)}) \end{aligned}$$

die Äquivalenz von

$$X_n^{(k)} \xrightarrow[n \rightarrow \infty]{} X^{(k)} \text{ in Verteilung für } k = 1, \dots, d$$

und

$$X_n \xrightarrow[n \rightarrow \infty]{} X \text{ in Verteilung.}$$

5) (Zentraler Grenzwertsatz)

Seien X_1, X_2, \dots unabhängige, identisch verteilte Zufallsvariablen mit

$$E(X_1^2) < \infty$$

und $V(X_1) > 0$. Setze $\mu := E(X_1)$, $\sigma^2 = V(X_1)$.

Dann gilt

$$P_{S_n} \xrightarrow[n \rightarrow \infty]{} N(0, \sigma^2) \text{ schwach,} \quad (4.2.3)$$

wobei

$$S_n := \sum_{l=1}^n \frac{(X_l - \mu)}{\sqrt{n}}.$$

Nach dem Stetigkeitssatz von Lévy gilt (4.2.3) genau dann, wenn

$$\varphi_{S_n}(t) \xrightarrow[n \rightarrow \infty]{} e^{-\frac{\sigma^2 t^2}{2}} \quad \forall t \in \mathbb{R}.$$

Definition (Irrfahrt)

Gegeben sei eine Folge Y_1, Y_2, \dots unabhängiger, identisch verteilter Zufallsvariablen. Setze

$$\begin{aligned} X_0 &:= 0 \\ X_n &:= Y_1 + \dots + Y_n \text{ für } n \in \mathbb{N}. \end{aligned}$$

Wir nennen $(X_n)_{n \in \mathbb{N}_0}$ eine *Irrfahrt auf \mathbb{R}* und Y_1, Y_2, \dots die *Schritte* der Irrfahrt.

Bemerkung: Wegen

$$Y_n = X_n - X_{n-1}, \quad n \in \mathbb{N},$$

erhält man die Schritte der Irrfahrt zurück.

Definition (Brownsche Bewegung mit Varianz σ^2)

Sei $\sigma^2 \in \mathbb{R}_+^*$. Ein stochastischer Prozess $(B_t)_{t \in \mathbb{R}_+}$ heißt *Brownsche Bewegung mit Varianz σ^2* , wenn gilt:

- $(B_t)_{t \in \mathbb{R}_+}$ ist ein Lévy-Prozess mit Faltungshalbgruppe $N(0, \sigma^2 t)$
- $(B_t)_{t \in \mathbb{R}_+}$ ist stetig.

Bemerkung: Man kann genau wie für $\sigma^2 = 1$ zeigen, dass es eine Brownsche Bewegung mit Varianz σ^2 gibt.

Satz 4.2.6 (Invarianzprinzip)

Sei $(X_n)_{n \in \mathbb{N}_0}$ eine Irrfahrt auf \mathbb{R} , deren Schritte quadrat-integrierbar sind mit Erwartungswert 0 und Varianz $\sigma^2 > 0$.

Der reskalierte Prozess $X_t^{(c)}$ für $c \in \mathbb{R}_+^*$ wird definiert durch

$$X_t^{(c)} := \frac{1}{\sqrt{c}} X_{ct}, \quad t \in \mathbb{R}_+,$$

wobei wir $(X_n)_{n \in \mathbb{N}_0}$ durch lineare Interpolation zu einem Prozess $(X_t)_{t \in \mathbb{R}_+}$ fortsetzen.

Dann gilt für alle $d \in \mathbb{N}$, $t_1, \dots, t_d \in \mathbb{R}_+$

$$(X_{t_1}^{(c)}, \dots, X_{t_d}^{(c)}) \xrightarrow{c \rightarrow \infty} (B_{t_1}, \dots, B_{t_d}) \text{ in Verteilung,} \quad (4.2.4)$$

wobei $(B_t)_{t \in \mathbb{R}_+}$ eine Brownsche Bewegung mit Varianz σ^2 bezeichnet.

Beweis: Zuerst definieren wir

$$\tilde{X}_t^{(c)} := \frac{1}{\sqrt{c}} X_{[ct]}, \quad t \in \mathbb{R}_+.$$

Dabei bezeichnet $[\alpha]$ für $\alpha \in \mathbb{R}$ das durch

$$[\alpha] := \max\{m \in \mathbb{Z} \mid m \leq \alpha\}$$

definierte Gauß-Symbol von α . Dann überlegt man sich, dass die Abschätzung

$$|X_{t_1}^{(c)} - \tilde{X}_{t_1}^{(c)}| \leq \frac{1}{\sqrt{c}} |Y_{[ct_1]+1}|$$

gilt (Skizze). Daraus schließt man

$$\|(X_{t_1}^{(c)}, \dots, X_{t_d}^{(c)}) - (X_{t_1}^{(c)}, \dots, X_{t_d}^{(c)})\|_2^2 \xrightarrow{c \rightarrow \infty} 0,$$

d. h.

$$(X_{t_1}^{(c)}, \dots, X_{t_d}^{(c)}) - (\tilde{X}_{t_1}^{(c)}, \dots, \tilde{X}_{t_d}^{(c)}) \xrightarrow{c \rightarrow \infty} 0 \quad (4.2.5)$$

im quadratischen Mittel. Da die Konvergenz im quadratischen Mittel die Konvergenz in Verteilung impliziert, folgt die schwache Konvergenz (4.2.5). Jetzt schließt man, dass es genügt,

$$(\tilde{X}_{t_1}^{(c)}, \dots, \tilde{X}_{t_d}^{(c)}) \xrightarrow{c \rightarrow \infty} (B_{t_1}, \dots, B_{t_d}) \text{ in Verteilung}$$

zu zeigen.

Setze dazu ($t_0 := 0$)

$$\begin{aligned} U_k^{(c)} &:= \tilde{X}_{t_k}^{(c)} - \tilde{X}_{t_{k-1}}^{(c)} \\ Z_k &:= B_{t_k} - B_{t_{k-1}} \end{aligned}$$

für $k = 1, \dots, d$. Da $\tilde{X}_0^{(c)}(0) = X_0 = 0 = B_0$, genügt es zu zeigen, dass

$$(U_1^{(c)}, \dots, U_d^{(c)}) \xrightarrow{c \rightarrow \infty} (Z_1, \dots, Z_d) \text{ in Verteilung.}$$

Da nun $U_1^{(c)}, \dots, U_d^{(c)}$ unabhängig (X_n ist Irrfahrt) und Z_1, \dots, Z_n unabhängig, folgt (siehe Bem. 4) oben), dass es ausreicht,

$$U_k^{(c)} \xrightarrow{c \rightarrow \infty} Z_k \text{ in Verteilung}$$

nachzuweisen für $k = 1, \dots, d$.

Es besitzt

$$U_k^{(c)} = \frac{1}{\sqrt{c}} \sum_{l=[ct_{k+1}]+1}^{[ct_k]} Y_l$$

dieselbe Verteilung wie

$$\frac{1}{\sqrt{c}} \sum_l^{[ct_k]-[ct_{k-1}]} Y_l,$$

da Y_1, Y_2, \dots , identisch verteilt sind nach Voraussetzung. Es gilt:

$$\begin{aligned} & \frac{1}{\sqrt{c}} \sum_l^{[ct_k]-[ct_{k-1}]} Y_l \\ &= \frac{\sqrt{[ct_k]-[ct_{k-1}]}}{\sqrt{c}} \frac{1}{\sqrt{[ct_k]-[ct_{k-1}]}} (Y_1 + \dots + Y_{[ct_k]-[ct_{k-1}]}) \end{aligned} \quad (4.2.6)$$

Behauptung: Für $\alpha, \beta \in \mathbb{R}$ gilt stets

$$\frac{[n\alpha] - [n\beta]}{n} \xrightarrow{n \rightarrow \infty} \alpha - \beta.$$

Beweis: Folgt aus der Einschachtelung

$$\frac{n\alpha - 1 - n\beta}{n} \leq \frac{[n\alpha] - [n\beta]}{n} \leq \frac{n\alpha - (n\beta - 1)}{n}. \checkmark$$

Also:

$$\frac{[ct_k] - [ct_{k-1}]}{c} \xrightarrow{c \rightarrow \infty} t_k - t_{k-1} \quad (4.2.7)$$

und insbesondere

$$[ct_k] - [ct_{k-1}] \xrightarrow{c \rightarrow \infty} \infty.$$

Nach dem Zentralen Grenzwertsatz (siehe Bem. 5) oben) konvergiert

$$\frac{1}{\sqrt{[ct_k] - [ct_{k-1}]}} (Y_1 + \dots + Y_{[ct_k] - [ct_{k-1}]}) \quad (4.2.8)$$

für $c \rightarrow \infty$ in Verteilung gegen eine Zufallsvariable, die normalverteilt mit Erwartungswert 0 und Varianz σ^2 ist. Wir haben aber

$$\frac{1}{\sqrt{t_k - t_{k-1}}} Z_k = \frac{1}{\sqrt{t_k - t_{k-1}}} (B_{t_k} - B_{t_{k-}}) \sim N(0, \sigma^2).$$

Also konvergiert (4.2.8) gegen

$$\frac{1}{\sqrt{t_k - t_{k-1}}} Z_k.$$

Nun gilt allgemein:

Lemma: Sei X_n eine Folge von Zufallsvariablen mit

$$P_{X_n} \xrightarrow[n \rightarrow \infty]{} N(\mu, \sigma^2) \text{ schwach.} \quad (4.2.9)$$

Außerdem sei α_n eine Folge in \mathbb{R} mit

$$\alpha_n \xrightarrow[n \rightarrow \infty]{} \alpha \in \mathbb{R}_+^*.$$

Dann gilt

$$P_{\alpha_n X_n} \xrightarrow[n \rightarrow \infty]{} N(\alpha\mu, \alpha^2\sigma^2) \text{ schwach.}$$

Beweis: Bezeichne mit F_n die Verteilungsfunktion von X_n und mit $\Phi(\mu, \sigma^2)$ die Verteilungsfunktion von $N(\mu, \sigma^2)$. Es ist (4.2.9) gleichbedeutend mit

$$\lim_{n \rightarrow \infty} F_n(t) = \Phi(\mu, \sigma^2)(t) \quad \forall t \in \mathbb{R}.$$

Sei $t \geq 0$. Für $\varepsilon > 0$ mit $\alpha - \varepsilon > 0$ sei n_0 so, dass

$$\alpha - \varepsilon \leq \alpha_n \leq \alpha + \varepsilon \quad \forall n \geq n_0.$$

Da

$$(\alpha - \varepsilon)X_n(\omega) \leq \alpha_n X_n(\omega) \leq (\alpha + \varepsilon)X_n(\omega),$$

wenn $X_n(\omega) \geq 0$, was für $\omega \in \{(\alpha - \varepsilon)X_n > t\}$, $\omega \in \{\alpha_n X_n > t\}$ oder $\omega \in \{(\alpha + \varepsilon)X_n > t\}$ der Fall ist, folgt

$$\{(\alpha - \varepsilon)X_n > t\} \subset \{\alpha_n X_n > t\} \subset \{(\alpha + \varepsilon)X_n > t\}.$$

Also:

$$P((\alpha - \varepsilon)X_n > t) \leq P(\alpha_n X_n > t) \leq P((\alpha + \varepsilon)X_n > t) \quad (4.2.10)$$

Indem wir in (4.2.10) mit n gegen ∞ gehen, erhalten wir

$$\begin{aligned} 1 - \Phi(\mu, \sigma^2)\left(\frac{t}{\alpha - \varepsilon}\right) &\leq \liminf_{n \rightarrow \infty} P(\alpha_n X_n > t) \\ &\leq \limsup_{n \rightarrow \infty} P(\alpha_n X_n > t) \leq 1 - \Phi(\mu, \sigma^2)\left(\frac{t}{\alpha + \varepsilon}\right). \end{aligned}$$

Mit $\varepsilon \downarrow 0$ folgt hieraus aus der Stetigkeit von $\Phi(\mu, \sigma^2)$, dass $P(\alpha_n X_n > t)$ konvergiert mit

$$\lim_{n \rightarrow \infty} P(\alpha_n X_n > t) = 1 - \Phi(\mu, \sigma^2)\left(\frac{t}{\alpha}\right),$$

d. h.

$$\lim_{n \rightarrow \infty} P(\alpha_n X_n \leq t) = \Phi(\mu, \sigma^2)\left(\frac{t}{\alpha}\right) = \Phi(\alpha\mu, \alpha^2\sigma^2)(t). \quad (4.2.11)$$

Mit einem analogen Argument zeigt man (4.2.11) für $t < 0$. Aus der Konvergenz der Verteilungsfunktion von $\alpha_n X_n$ gegen die Verteilungsfunktion von $\Phi(\alpha\mu, \alpha^2\sigma^2)$ in allen Punkten, folgt die schwache Konvergenz der Verteilung von $\alpha_n X_n$ gegen $N(\alpha\mu, \alpha^2\sigma^2)$. ✓

Aus dem Lemma folgt wegen

$$\frac{\sqrt{[ct_k] - [ct_{k-1}]}}{\sqrt{c}} \xrightarrow{c \rightarrow \infty} \sqrt{t_k - t_{k-1}}$$

und (4.2.6) die Gültigkeit von

$$\frac{1}{\sqrt{c}} \sum_l^{[ct_k] - [ct_{k-1}]} Y_l \xrightarrow{c \rightarrow \infty} Z_k \text{ in Verteilung.}$$

Also auch $U_k^{(c)} \xrightarrow{c \rightarrow \infty} Z_k$ in Verteilung. □

Bemerkung: Die “Invarianz” in Satz 4.2.6 besteht darin, dass der Limesprozess B_t nur von σ^2 , nicht aber von der genauen Verteilung der ursprünglichen Schritte Y_1, Y_2, \dots der Irrfahrt abhängt.

Wir haben gezeigt, dass die endlich dimensionalen Verteilungen von $\frac{1}{\sqrt{c}}X_{ct}$ für $c \rightarrow \infty$ schwach gegen die endlich dimensionalen Verteilungen von B_t

konvergieren. Das ist noch nicht das *Invarianzprinzip von Donsker*, aber geht in diese Richtung. Wir haben gesehen, dass der Lévy-Prozess über

$$(\mathbb{R}^{\mathbb{R}_+}, \mathcal{B}(\mathbb{R}^{\mathbb{R}_+}), P)$$

mit Faltungshalbgruppe $N(0, \sigma^2 t)$ auf $\mathcal{C}(\mathbb{R}_+) \subset \mathbb{R}^{\mathbb{R}_+}$ “konzentriert” ist. Man kann zeigen, dass man die Brownsche Bewegung auch als Prozess

$$X_t(f) = f(t), \quad f \in \mathcal{C}(\mathbb{R}_+),$$

über

$$(\mathcal{C}(\mathbb{R}_+), \mathcal{B}(\mathcal{C}(\mathbb{R}_+)), P_W)$$

mit einem geeigneten Wahrscheinlichkeitsmaß P_W , dem so genannten *Wiener-Maß*, realisieren kann. (Dabei ist $\mathcal{C}(\mathbb{R}_+)$ mit der Topologie der gleichmäßigen Konvergenz auf Kompakta zu versehen.) Entsprechend kann man $\frac{1}{\sqrt{c}} X_{ct}$ auf

$$(\mathcal{C}(\mathbb{R}_+), \mathcal{B}(\mathcal{C}(\mathbb{R}_+)), P_c)$$

mit einem (von c abhängigen) Wahrscheinlichkeitsmaß P_c realisieren. Das Invarianzprinzip von Donsker besagt nun, dass

$$P_c \xrightarrow{c \rightarrow \infty} P_W \text{ schwach.}$$

Das impliziert die Konvergenz der endlich-dimensionalen Verteilungen und ist sogar äquivalent dazu, weil die Folge $(P_c)_{c \in \mathbb{N}}$ “straff” ist. Die Straffheit ist nicht so leicht zu zeigen; siehe z. B. “Klenke”.

4.3 Weitere Eigenschaften der Pfade einer Brownschen Bewegung

Wir beginnen mit eine bemerkenswerten Tatsache:

Behauptung: Sei B_t eine Brownsche Bewegung (mit Varianz 1). Außerdem sei $c \in \mathbb{R}_+^*$. Dann ist durch

$$X_t := \frac{1}{c} B_{c^2 t}, \quad t \in \mathbb{R}_+,$$

ebenfalls eine Brownsche Bewegung gegeben.

Beweis: Da

$$B_{c^2 t_2} - B_{c^2 t_1}, \dots, B_{c^2 t_{n+1}} - B_{c^2 t_n}$$

unabhängig für alle $n \in \mathbb{N}$, $0 \leq t_1 < \dots < t_{n+1}$, folgt jedenfalls, dass X_t unabhängige Zuwächse hat. Weiterhin für $0 \leq s < t$:

$$P_{X_t - X_s} = P_{\frac{1}{c}(B_{c^2 t} - B_{c^2 s})} = N\left(0, \frac{1}{c^2}(c^2 t - c^2 s)\right) = N(0, t - s),$$

womit gezeigt ist, dass X_t auch stationäre Zuwächse besitzt mit $P_{X_t} = N(0, t)$. Also ist X_t ein Lévy-Prozess mit Faltungshalbgruppe $N(0, t)$. Die Stetigkeit der Pfade $t \mapsto \frac{1}{c}B_{c^2 t}(\omega)$ folgt aus der Stetigkeit der Pfade $t \mapsto B_t(\omega)$. ✓

Satz 4.3.1 *Fast alle Pfade einer Brownschen Bewegung sind nirgends differenzierbar. Genauer: Es gibt eine messbare Menge $\tilde{\Omega} \subset \Omega$ mit:*

- $P(\tilde{\Omega}) = 1$
- $\omega \in \tilde{\Omega} \Rightarrow$ Die Funktion $t \mapsto B_t(\omega)$ von \mathbb{R}_+ nach \mathbb{R} ist an keiner Stelle differenzierbar.

Beweis: Behauptung: Es genügt zu zeigen, dass f. a. Pfade $t \mapsto B_t(\omega)$ von $[0, 1]$ nach \mathbb{R} nirgends differenzierbar sind.

Beweis davon: Da $(\frac{1}{N}B_{N^2 t})_{t \in \mathbb{R}_+}$ für $N \in \mathbb{N}$ eine Brownsche Bewegung ist, gäbe es dann eine Menge $\tilde{\Omega}_N \in \mathcal{F}$ mit

- $P(\tilde{\Omega}_N) = 1$
- $\omega \in \tilde{\Omega}_N \Rightarrow$ Die Funktion $t \mapsto B_t(N^2 \omega)$ von $[0, 1]$ nach \mathbb{R} ist nirgends differenzierbar.

Letzteres bedeutet aber, dass $t \mapsto B_t(\omega)$ als Funktion von $[0, N]$ nach \mathbb{R} nirgends differenzierbar ist. Mit $\tilde{\Omega} := \bigcap_{N \in \mathbb{N}} \tilde{\Omega}_N$ würde dann der Satz folgen. ✓

Lemma: Wenn für $\omega \in \Omega$ gilt, dass $t \mapsto B_t(\omega)$ in wenigstens einem Punkt $s \in [0, 1]$ differenzierbar ist, dann folgt:

$\exists C, m \in \mathbb{N}$, so dass für alle $n \geq m$ ein i , $1 \leq i \leq n + 1$, existiert mit:

$$j \in \{i, i + 1, i + 2\} \Rightarrow |B(\frac{j+1}{n}, \omega) - B(\frac{j}{n}, \omega)| \leq \frac{8C}{n}.$$

Beweis des Lemmas: nächste Woche

Stochastische Prozesse

SoSe 2018

Übersicht 10. Vorlesungswoche (18.6.-22.6.2018)

Fortsetzung des Beweises von Satz 4.3.1:

Lemma: Wenn für $\omega \in \Omega$ gilt, dass $t \mapsto B_t(\omega)$ in wenigstens einem Punkt $s \in [0, 1]$ differenzierbar ist, dann folgt:

$\exists C, m \in \mathbb{N}$, so dass für alle $n \geq m$ ein i , $1 \leq i \leq n + 1$, existiert mit:

$$j \in \{i, i + 1, i + 2\} \Rightarrow |B(\frac{j+1}{n}, \omega) - B(\frac{j}{n}, \omega)| \leq \frac{8C}{n}.$$

Beweis des Lemmas: Sei $t \mapsto B(t, \omega)$ in $s \in [0, 1]$ differenzierbar. (ω ist jetzt fest!) Dann existiert

$$\lim_{t \rightarrow s^+} \frac{|B(t, \omega) - B(s, \omega)|}{t - s} =: c.$$

Setze $C := [c] + 1$. Dann gibt es ein $\varepsilon > 0$ mit

$$\frac{|B(t, \omega) - B(s, \omega)|}{t - s} \leq C$$

für alle $t > s$ mit $t - s \leq \varepsilon$ oder

$$|B(t, \omega) - B(s, \omega)| \leq C(t - s) \quad \forall t \geq s \text{ mit } t - s \leq \varepsilon. \quad (4.3.1)$$

Wähle nun $m \in \mathbb{N}$ so groß, dass

$$\frac{1}{n} < \frac{\varepsilon}{4} \quad (\Leftrightarrow \frac{4}{n} < \varepsilon) \quad \forall n \geq m.$$

Setze als nächstes $i := [ns] + 1$. Dann gilt wegen $s \in [0, 1]$ jedenfalls $1 \leq i \leq n + 1$ und

$$\frac{i}{n} = \frac{[ns] + 1}{n} \geq \frac{ns}{n} = s$$

sowie

$$\frac{i}{n} = \frac{[ns] + 1}{n} \leq \frac{ns + 1}{n} = s + \frac{1}{n},$$

d. h.

$$s \leq \frac{i}{n} \leq s + \frac{1}{n} \leq s + \varepsilon.$$

Nach dreimaliger Addition von $\frac{1}{n}$ erhalten wir daraus die Ungleichungen

$$\begin{aligned} s + \frac{1}{n} &\leq \frac{i+1}{n} \leq s + \frac{2}{n} \leq s + \varepsilon \\ s + \frac{2}{n} &\leq \frac{i+2}{n} \leq s + \frac{3}{n} \leq s + \varepsilon \\ s + \frac{3}{n} &\leq \frac{i+3}{n} \leq s + \frac{4}{n} \leq s + \varepsilon. \end{aligned}$$

Also:

$$s \leq \frac{j}{n} \leq s + \varepsilon \text{ für } j = i, i+1, i+2, i+3. \quad (4.3.2)$$

Wegen (4.3.2) folgt nun mit (4.3.1):

$$\begin{aligned} |B(\frac{j}{n}, \omega) - B(\frac{j+1}{n}, \omega)| &\leq |B(\frac{j}{n}, \omega) - B(s, \omega)| \\ &\quad + |B(s, \omega) - B(\frac{j+1}{n}, \omega)| \\ &\leq C(\frac{j}{n} - s) + C(\frac{j+1}{n} - s) \\ &\leq C\frac{4}{n} + C\frac{4}{n} = \frac{8C}{n} \end{aligned}$$

für $j = i, i+1, i+2$. ✓

Zurück zum Beweis von Satz 4.3.1:

Setze für $C, n, j \in \mathbb{N}$

$$A_{C,n}^j := \{\omega \in \Omega \mid |B(\frac{j}{n}, \omega) - B(\frac{j+1}{n}, \omega)| \leq \frac{8C}{n}\}$$

und

$$A := \bigcup_{C \in \mathbb{N}} \bigcup_{\substack{m \in \mathbb{N} \\ n \geq m}} \bigcap_{\substack{n \in \mathbb{N} \\ n \geq m}} \bigcup_{1 \leq i \leq n+1} \bigcap_{i \leq j \leq i+2} A_{C,n}^j. \quad (4.3.3)$$

Nach dem obigen Lemma enthält $A \in \mathcal{F}$ alle $\omega \in \Omega$, für die $t \mapsto B(t, \omega)$ in wenigstens einem Punkt $s \in [0, 1]$ differenzierbar ist. Wenn wir also $P(A) = 0$ nachweisen, gilt für $\tilde{\Omega} = A^c$:

· $P(\tilde{\Omega}) = 1$

· $\omega \in \tilde{\Omega} \Rightarrow t \mapsto B(t, \omega)$ ist in keinem Punkt von $[0, 1]$ differenzierbar,

Damit wäre nach unserer Vorüberlegung der Beweis von Satz 4.3.1 abgeschlossen.

Beweis von $P(A) = 0$:

Wir haben für $C, m \in \mathbb{N}$:

$$\mathbb{P}\left(\bigcap_{\substack{n \in \mathbb{N} \\ n \geq m}} \bigcup_{1 \leq i \leq n+1} \bigcap_{i \leq j \leq i+2} A_{C,n}^j\right) \leq \mathbb{P}\left(\bigcup_{1 \leq i \leq n_0+1} \bigcap_{i \leq j \leq i+2} A_{C,n_0}^j\right) \quad (4.3.4)$$

für alle $n_0 \geq m$. Da $\bigcup_{1 \leq i \leq n+1} \bigcap_{i \leq j \leq i+2} A_{C,n}^j$ eine in $n \in \mathbb{N}$ aufsteigende Folge von messbaren Mengen ist, folgt aus (4.3.4)

$$\mathbb{P}\left(\bigcap_{\substack{n \in \mathbb{N} \\ n \geq m}} \bigcup_{1 \leq i \leq n+1} \bigcap_{i \leq j \leq i+2} A_{C,n}^j\right) \leq \lim_{n \rightarrow \infty} \mathbb{P}\left(\bigcup_{1 \leq i \leq n+1} \bigcap_{i \leq j \leq i+2} A_{C,n}^j\right) \quad (4.3.5)$$

Nun schätzen wir weiter ab:

$$\begin{aligned} & \mathbb{P}\left(\bigcup_{1 \leq i \leq n+1} \bigcap_{i \leq j \leq i+2} A_{C,n}^j\right) \quad (4.3.6) \\ & \leq \sum_{i=1}^{n+1} \mathbb{P}\left(\bigcap_{i \leq j \leq i+2} A_{C,n}^j\right) \\ & = \sum_{i=1}^{n+1} \mathbb{P}\left(|B\left(\frac{i}{n}\right) - B\left(\frac{i+1}{n}\right)| \leq \frac{8C}{n}, |B\left(\frac{i+1}{n}\right) - B\left(\frac{i+2}{n}\right)| \leq \frac{8C}{n}, \right. \\ & \quad \left. |B\left(\frac{i+2}{n}\right) - B\left(\frac{i+3}{n}\right)| \leq \frac{8C}{n}\right) \\ & = (n+1) \left(\mathbb{P}\left(|B\left(\frac{1}{n}\right)| \leq \frac{8C}{n}\right)\right)^3 \end{aligned}$$

und, da $\sqrt{n}B\left(\frac{1}{n}\right) \sim N(0, 1)$:

$$\begin{aligned} \mathbb{P}\left(|B\left(\frac{1}{n}\right)| \leq \frac{8C}{n}\right) &= \mathbb{P}\left(\sqrt{n} |B\left(\frac{1}{n}\right)| \leq \frac{8C}{\sqrt{n}}\right) \\ &= \frac{2}{\sqrt{2\pi}} \int_0^{\frac{8C}{\sqrt{n}}} e^{-\frac{x^2}{2}} dx \\ &\leq \frac{2}{\sqrt{2\pi}} \frac{8C}{\sqrt{n}} \end{aligned}$$

Jetzt folgt mit (4.3.5) und (4.3.6)

$$\begin{aligned} & \mathbb{P}\left(\bigcap_{\substack{n \in \mathbb{N} \\ n \geq m}} \bigcup_{1 \leq i \leq n+1} \bigcap_{i \leq j \leq i+2} A_{C,n}^j\right) \\ & \leq \lim_{n \rightarrow \infty} (n+1) \left(\frac{2}{\sqrt{2\pi}} \frac{8C}{\sqrt{n}}\right)^3 \\ & = 0. \end{aligned}$$

Es ist also A als abzählbare Vereinigung (4.3.3) von Nullmengen eine Nullmenge. \square

Definition (p -te Variation)

Sei $p \in \mathbb{R}_+^*$. Eine Funktion ($a, b \in \mathbb{R}, a \leq b$)

$$f : [a, b] \rightarrow \mathbb{R}$$

heißt von *beschränkter p -ter Variation*, falls es eine Konstante $C \in \mathbb{R}_+$ gibt, so dass für jede Zerlegung \mathfrak{Z} ,

$$\mathfrak{Z} = \{a = t_0 < t_1 < \dots < t_m < t_{m+1} = b\},$$

von $[a, b]$ gilt:

$$V_p^{\mathfrak{Z}}(f) := \sum_{k=0}^m |f(t_{k+1}) - f(t_k)|^p \leq C.$$

Setze für eine Funktion $f : [a, b] \rightarrow \mathbb{R}$:

$$V_p(f) := \sup\{V_p^{\mathfrak{Z}} \mid \mathfrak{Z} \text{ Zerlegung von } [a, b]\}$$

$V_p(f)$ heißt p -te Variation von f . Es ist also f genau dann von beschränkter p -ter Variation, wenn $V_p(f) < \infty$ gilt, Im Fall $p = 1$ heißt $V_p(f)$ *totale Variation*, im Fall $p = 2$ *quadratische Variation* von f .

Bemerkungen: 1) Die Zerlegungen \mathfrak{Z} von $[a, b]$ bilden mit

$$\mathfrak{Z}_1 \prec \mathfrak{Z}_2 \Leftrightarrow \mathfrak{Z}_1 \subset \mathfrak{Z}_2$$

eine gerichtete Menge, und es gilt für $p \in [0, 1]$: Das Netz $(V_p^{\mathfrak{Z}}(f))_{\mathfrak{Z}}$ konvergiert genau dann, wenn f von beschränkter p -ter Variation ist, und in diesem Fall gilt

$$V_p^{\mathfrak{Z}}(f) \rightarrow V_p(f);$$

für $p = 1$ siehe "Heuser: Analysis I" (sollte für $p < 1$ genauso gehen).

Definition Bezeichne mit $\mathbf{Par}_{a,b}$ die Menge aller Zerlegungen des Intervalls $[a, b]$. Setze für $\mathfrak{Z} \in \mathbf{Par}_{a,b}$

$$\|\mathfrak{Z}\| := \max_{k=0, \dots, m} (t_{k+1} - t_k).$$

Dann heißt $\|\mathfrak{Z}\|$ *Maschenweite* der Zerlegung \mathfrak{Z} . Für einen normierten Vektorraum \mathcal{V} und ein Netz $(v_{\mathfrak{Z}})_{\mathfrak{Z} \in \mathbf{Par}_{a,b}}$ von Vektoren $v_{\mathfrak{Z}}$ in \mathcal{V} schreiben wir

$$\lim_{\|\mathfrak{Z}\| \rightarrow 0} v_{\mathfrak{Z}} = v \tag{4.3.7}$$

mit $v \in \mathcal{V}$, wenn gilt: Für alle $\varepsilon > 0$ existiert ein $\delta > 0$, so dass

$$\mathfrak{Z} \in \mathbf{Par}_{a,b}, \|\mathfrak{Z}\| \leq \delta \Rightarrow \|v_{\mathfrak{Z}} - v\| \leq \varepsilon. \quad (4.3.8)$$

Bemerkung:

Behauptung Aus (4.3.7) folgt, dass das Netz $(v_{\mathfrak{Z}})_{\mathfrak{Z} \in \mathbf{Par}_{a,b}}$ gegen v konvergiert.

Beweis: Gelte (4.3.7). Dann ist zu zeigen: Für jedes $\varepsilon > 0$ existiert ein $\mathfrak{Z}_0 \in \mathbf{Par}_{a,b}$ mit: $\mathfrak{Z}_0 \prec \mathfrak{Z} \Rightarrow \|v_{\mathfrak{Z}} - v\| \leq \varepsilon$. Jedes \mathfrak{Z}_0 mit $\|\mathfrak{Z}_0\| \leq \delta$, δ wie in (4.3.8), erfüllt diese Bedingung. ✓

Satz 4.3.2 Sei B_t eine Brownsche Bewegung und sei $t \in \mathbb{R}_+^*$. Betrachte die Zufallsvariablen

$$V_{2,t}^{\mathfrak{Z}}(B), \mathfrak{Z} \in \mathbf{Par}_{0,t},$$

wobei wir für $\mathfrak{Z} = \{0 = t_0 < \dots < t_{m+1} = t\}$

$$V_{2,t}^{\mathfrak{Z}}(B)(\omega) := \sum_{k=0}^m (B(t_{k+1}, \omega) - B(t_k, \omega))^2, \omega \in \Omega,$$

setzen.

Dann gilt

$$\lim_{\|\mathfrak{Z}\| \rightarrow 0} V_{2,t}^{\mathfrak{Z}}(B) = t \quad (4.3.9)$$

in der Norm von $L^2(\Omega, \mathcal{F}, P)$, d. h.:

Für alle $\varepsilon > 0$ existiert ein $\delta > 0$, so dass

$$\|V_{2,t}^{\mathfrak{Z}}(B) - t\|_2 = (E((V_{2,t}^{\mathfrak{Z}}(B) - t)^2))^{\frac{1}{2}} \leq \varepsilon$$

falls nur $\|\mathfrak{Z}\| \leq \delta$.

Beweis: Es gilt für $\mathfrak{Z} \in \mathbf{Par}_{0,t}$:

$$\begin{aligned} \|V_{2,t}^{\mathfrak{Z}}(B) - t\|_2^2 &= E\left(\left(\sum_{k=0}^m (B(t_{k+1}) - B(t_k))^2 - t\right)^2\right) & (4.3.10) \\ &= E\left(\left(\sum_{k=0}^m (B(t_{k+1}) - B(t_k))^2\right)^2\right) \\ &\quad - 2t E\left(\sum_{k=0}^m (B(t_{k+1}) - B(t_k))^2\right) + t^2 \\ &= E\left(\left(\sum_{k=0}^m (B(t_{k+1}) - B(t_k))^2\right)^2\right) - t^2 \end{aligned}$$

Als nächstes haben wir

$$\begin{aligned} \mathbb{E}\left(\left(\sum_{k=0}^m (B(t_{k+1}) - B(t_k))^2\right)^2\right) &= \mathbb{E}\left(\sum_{k=0}^m (B(t_{k+1}) - B(t_k))^4\right) \\ &+ \sum_{\substack{i,j \\ i \neq j}} (t_{i+1} - t_i)(t_{j+1} - t_j), \end{aligned} \quad (4.3.11)$$

weil disjunkte Zuwächse der Brownschen Bewegung unabhängig sind und da $B(t_{k+1}) - B(t_k) \sim N(0, t_{k+1} - t_k)$.

Außerdem:

$$\begin{aligned} \sum_{\substack{i,j \\ i \neq j}} (t_{i+1} - t_i)(t_{j+1} - t_j) &= \sum_{i,j} (t_{i+1} - t_i)(t_{j+1} - t_j) - \sum_i (t_{i+1} - t_i)^2 \\ &= t^2 - \sum_i (t_{i+1} - t_i)^2 \end{aligned} \quad (4.3.12)$$

Des Weiteren gilt für $0 \leq r \leq s$

$$\begin{aligned} \mathbb{E}((B_r - B_s)^4) &= \mathbb{E}(B_{r-s}^4) \\ &= \int x^4 dN(0, r-s)(x) \\ &= 3(r-s)^2 \end{aligned}$$

und deshalb

$$\mathbb{E}\left(\sum_{k=0}^m (B(t_{k+1}) - B(t_k))^4\right) = 3 \sum_{k=0}^m (t_{k+1} - t_k)^2 \quad (4.3.13)$$

Nun erhalten wir schließlich:

$$\begin{aligned}
\mathbb{E}\left(\left(\sum_{k=0}^m (B(t_{k+1}) - B(t_k))^2 - t\right)^2\right) &\stackrel{(4.3.10)}{=} \mathbb{E}\left(\left(\sum_{k=0}^m (B(t_{k+1}) - B(t_k))^2\right)^2\right) - t^2 \\
&\stackrel{(4.3.11)}{=} \mathbb{E}\left(\sum_{k=0}^m (B(t_{k+1}) - B(t_k))^4\right) \\
&\stackrel{(4.3.12)}{=} \mathbb{E}\left(\sum_{k=0}^m (B(t_{k+1}) - B(t_k))^4\right) \\
&\quad + t^2 - \sum_{k=0}^m (t_{k+1} - t_k)^2 - t^2 \\
&\stackrel{(4.3.13)}{=} 3 \sum_{k=0}^m (t_{k+1} - t_k)^2 - \sum_{k=0}^m (t_{k+1} - t_k)^2 \\
&= 2 \sum_{k=0}^m (t_{k+1} - t_k)^2 \\
&\leq 2\|\mathfrak{Z}\| \sum_{k=0}^m (t_{k+1} - t_k) \\
&= 2\|\mathfrak{Z}\|t,
\end{aligned}$$

d . h.

$$\|V_{2,t}^{\mathfrak{Z}}(B) - t\|_2 \leq \sqrt{2t} \sqrt{\|\mathfrak{Z}\|},$$

womit (4.3.9) folgt. \square

Bemerkung: Da aus der Konvergenz im quadratischen Mittel die stochastische Konvergenz folgt, haben wir mit Satz 4.3.2 die *stochastische Konvergenz* von $V_{2,t}^{\mathfrak{Z}}(B)$ gegen t für $\|\mathfrak{Z}\| \rightarrow 0$ bewiesen. (Was das genau bedeutet, müsste man noch präzisieren. Geht aber sehr gut, da die stochastische Konvergenz von einer Metrik kommt.) Man sagt: *Für die Brownsche Bewegung existiert der Prozess der quadratischen Variation.* (Er ist durch den deterministischen Prozess $X_t = t$ gegeben!) Allerdings bedeutet dies nicht, dass etwa f. a. Pfade $s \mapsto B(s, \omega)$, $s \in [0, t]$, einer Brownschen Bewegung von beschränkter quadratischer Variation sind! Es gilt sogar, dass f. a. Pfade *für kein* t von beschränkter quadratischer Variation sind.

Korollar 4.3.3 Sei $t \in \mathbb{R}_+^*$. Mit

$$I^{\mathfrak{Z}}(B) := \sum_{k=0}^m B_{t_k} (B_{t_{k+1}} - B_{t_k}), \quad \mathfrak{Z} \in \mathbf{Par}_{0,t}, \quad (4.3.14)$$

gilt

$$I^{\mathfrak{Z}}(B) \xrightarrow{\|\mathfrak{Z}\| \rightarrow 0} \frac{1}{2} B_t^2 - \frac{1}{2} t$$

im quadratischen Mittel.

Beweis: Wegen $a^2 - b^2 = (a - b)^2 + 2b(a - b)$ gilt (mit Teleskop)

$$\begin{aligned} B_t^2 &= \sum_{k=0}^m (B_{t_{k+1}}^2 - B_{t_k}^2) \\ &= \sum_{k=0}^m (B_{t_{k+1}} - B_{t_k})^2 + 2 \sum_{k=0}^m B_{t_k} (B_{t_{k+1}} - B_{t_k}) \end{aligned}$$

und nach Auflösen:

$$\sum_{k=0}^m B_{t_k} (B_{t_{k+1}} - B_{t_k}) = \frac{1}{2} B_t^2 - \frac{1}{2} \sum_{k=0}^m (B_{t_{k+1}} - B_{t_k})^2$$

Nach Satz 4.3.2 konvergiert die Summe auf der rechten Seite dieser Gleichung im quadratischen Mittel gegen t . \square

Bemerkungen: 1) Für Funktionen $\alpha : [a, b] \rightarrow \mathbb{R}$ von beschränkter Variation konvergieren die *Riemann-Stieltjes-Summen*

$$\Sigma^{\tilde{\mathfrak{Z}}} = \sum_{k=0}^m f(\tau_k) (\alpha(t_{k+1}) - \alpha(t_k))$$

mit

$$\tilde{\mathfrak{Z}} = (\mathfrak{Z}, Z), Z = \{\tau_0, \tau_1, \dots, \tau_m\}, t_k \leq \tau_k \leq t_{k+1},$$

jedenfalls z. B. für stetige Funktionen f , für $\|\tilde{\mathfrak{Z}}\| = \|\mathfrak{Z}\| \rightarrow 0$ gegen einen Wert, der mit

$$\int_a^b f d\alpha$$

bezeichnet wird (Riemann-Stieltjes-Integral von f bzgl. α). Wir werden im nächsten Satz sehen, dass die Pfade einer Brownschen Bewegung aber fast alle nicht von beschränkter Variation sind, so dass die pfadweise Definition eines Integrals $\int_0^t F_s dB_s$ als Riemann-Stieltjes-Integral nicht in Frage kommt.

2) Man kann $I^{\tilde{\mathfrak{Z}}}(B)$ in (4.3.14) als Riemann-Stieltjes-Summen *mit der Itô-Wahl der Zwischenpunkte* (linker Intervall-Eckpunkt) deuten. Dann wird

$$\frac{1}{2} B(t)^2 - \frac{t}{2} =: \int_0^t B(s) dB(s) \quad (\text{f. s.}) \quad (4.3.15)$$

das *Itô-Integral* der Brownschen Bewegung B_t bzgl. der Brownschen Bewegung. Ein Vergleich mit der "klassischen" Formel

$$\int_0^t f(s) df(s) = \int_0^t f(s) f'(s) ds = \frac{1}{2} f(t)^2$$

(für eine stetig differenzierbare Funktionen f mit $f(0) = 0$) zeigt, dass beim Itô-Integral der Term $-\frac{t}{2}$ hinzutritt.

Als Vorbereitung für Satz 4.3.5 benötigen wir das folgende Lemma.

Lemma 4.3.4 *Gegeben seien Zufallsvariablen X_n, X mit*

$$\sum_{n=1}^{\infty} \mathbb{E}(|X_n - X|^p) < \infty \quad (4.3.16)$$

für ein $p \in \mathbb{R}_+^*$.
Dann folgt

$$X_n \xrightarrow[n \rightarrow \infty]{} X \text{ f. s.}$$

Beweis: Die Bedingung (4.3.16) bedeutet, dass $\sum_{n=1}^{\infty} (|X_n - X|^p)$ integrierbar ist. Dies impliziert

$$\mathbb{P}\left(\sum_{n=1}^{\infty} (|X_n - X|^p) < \infty\right) = 1$$

und

$$\mathbb{P}(|X_n - X| \xrightarrow[n \rightarrow \infty]{} 0) = 1. \square$$

Stochastische Prozesse

SoSe 2018

Übersicht 11. Vorlesungswoche (25.6.-29.6.2018)

Satz 4.3.5 *Fast alle Pfade einer Brownschen Bewegung sind auf jedem Intervall $[0, t]$, $t \in \mathbb{R}_+$, nicht von beschränkter (totaler) Variation.*

Beweis: Wir zeigen, dass es eine Folge \mathfrak{Z}_n in $\mathbf{Par}_{0,t}$ gibt mit

- $\|\mathfrak{Z}_n\| \xrightarrow{n \rightarrow \infty} 0$
- $V_1^{\mathfrak{Z}_n}(B(\omega)) \xrightarrow{n \rightarrow \infty} \infty$ für fast alle $\omega \in \Omega$.

Wähle dazu $\mathfrak{Z}_n \in \mathbf{Par}_{0,t}$ mit

$$\sum_{n=1}^{\infty} \|\mathfrak{Z}_n\| < \infty,$$

z. B.

$$\mathfrak{Z}_n = \left\{0, \frac{1}{2^n}t, \frac{2}{2^n}t, \dots, \frac{2^n - 1}{2^n}t, t\right\}.$$

Im Beweis von Satz 4.3.2 wurde die Ungleichung

$$\mathbb{E}\left(\left(\sum_{k=0}^{m_n} (B_{t_{k+1}^{(n)}} - B_{t_k^{(n)}})^2 - t\right)^2\right) \leq 2\|\mathfrak{Z}_n\|t$$

bewiesen. Also:

$$\sum_{n=1}^{\infty} \mathbb{E}\left(\left(\sum_{k=0}^{m_n} (B_{t_{k+1}^{(n)}} - B_{t_k^{(n)}})^2 - t\right)^2\right) \leq 2t \sum_{n=1}^{\infty} \|\mathfrak{Z}_n\| < \infty.$$

Nach Lemma 4.3.4 mit

$$X_n = \sum_{k=0}^{m_n} (B_{t_{k+1}^{(n)}} - B_{t_k^{(n)}})^2$$

und $X = t$ folgt

$$\sum_{k=0}^{m_n} (B_{t_{k+1}^{(n)}} - B_{t_k^{(n)}})^2 \xrightarrow[n \rightarrow \infty]{\text{f.s.}} t.$$

Nun können wir wie folgt argumentieren: Es gilt

$$\begin{aligned} & \sum_{k=0}^{m_n} (B_{t_{k+1}^{(n)}} - B_{t_k^{(n)}})^2 \\ & \leq \max_{0 \leq k \leq m_n} |B_{t_{k+1}^{(n)}} - B_{t_k^{(n)}}| \sum_{k=0}^{m_n} |B_{t_{k+1}^{(n)}} - B_{t_k^{(n)}}|. \end{aligned} \quad (4.3.1)$$

Aus der f. s. Stetigkeit der Pfade einer Brownschen Bewegung folgt

$$\max_{0 \leq k \leq m_n} |B_{t_{k+1}^{(n)}} - B_{t_k^{(n)}}| \xrightarrow[n \rightarrow \infty]{\text{f.s.}} 0.$$

Da die linke Seite von (4.3.1) f. s. konvergiert, und zwar gegen $t > 0$, muss f. s.

$$V_1^{\mathfrak{Z}_n}(B) = \sum_{k=0}^{m_n} |B_{t_{k+1}^{(n)}} - B_{t_k^{(n)}}| \xrightarrow[n \rightarrow \infty]{} \infty$$

gelten. \square

Bemerkung: In dem obigen Beweis haben wir gesehen, dass mit der dyadischen Zerlegung \mathfrak{Z}_n von $[0, t]$ die Konvergenz

$$V_2^{\mathfrak{Z}_n}(B) = \sum_{k=0}^{m_n} (B_{t_{k+1}^{(n)}} - B_{t_k^{(n)}})^2 \xrightarrow[n \rightarrow \infty]{\text{f.s.}} t$$

gilt. Man kann zeigen, dass die Pfade einer Brownschen Bewegung dennoch nicht immer von beschränkter quadratischer Variation sind. Es gilt sogar, dass f. a. Pfade für $p \in]0, 2]$ nicht von beschränkter p -ter Variation sind. (Dagegen sind sie es für $p \geq 2$.)

5 Die Markov-Eigenschaft

In diesem Abschnitt sei $T = \mathbb{N}_0$ oder $T = \mathbb{R}_+$.

Definition (Filtration, Adaptiertheit)

Sei $X = (X_t)_{t \in T}$ ein stochastischer Prozess. Eine Familie $(\mathcal{F}_t)_{t \in T}$ von σ -Unteralgebren (über Ω) von \mathcal{F} heißt *Filtration*, wenn gilt:

$$s, t \in T, s \leq t \Rightarrow \mathcal{F}_s \subset \mathcal{F}_t$$

(Im Falle $T = \mathbb{N}_0$ ist dies äquivalent zu: $\mathcal{F}_n \subset \mathcal{F}_{n+1}$ für alle $n \in \mathbb{N}_0$)

Der Prozess X heißt *adaptiert* an die Filtration \mathcal{F}_t , wenn gilt: X_t ist \mathcal{F}_t -messbar für alle $t \in T$.

Bemerkung: Es ist stets

$$\mathcal{F}_t := \sigma(X_s \mid s \leq t)$$

eine Filtration, und X ist immer an diese Filtration, die als *die von X erzeugte* bezeichnet wird, adaptiert.

Eine Filtration \mathcal{F}_t wird interpretiert als *Geschichte bis zum Zeitpunkt t* . Dass X adaptiert an \mathcal{F}_t ist, bedeutet dann, dass \mathcal{F}_t alle Information des Prozesses bis zum Zeitpunkt t enthält.

Ein *Markov-Prozess* ist ein stochastischer Prozess $X = (X_t)_{t \in T}$, für den

$$P(X_t \in B \mid X_{t_1}, \dots, X_{t_n}) = P(X_t \in B \mid X_{t_n}) \quad (5.0.2)$$

gilt für alle $n \in \mathbb{N}$, $0 \leq t_1 < t_2 < \dots < t_n < t$, $B \in \mathcal{B}(\mathbb{R})$. Dabei bedeuten $P(X_t \in B \mid X_{t_1}, \dots, X_{t_n})$ die *bedingte Wahrscheinlichkeit von $\{X_t \in B\}$ gegeben alle Information über den Prozess zu den Zeitpunkten t_1, \dots, t_n* und

$P(X_t \in B \mid X_{t_n})$ die *bedingte Wahrscheinlichkeit von $\{X_t \in B\}$ gegeben alle Information über den Prozess zum Zeitpunkt t_n* .

Die bedingten Wahrscheinlichkeiten sind Zufallsvariablen; für die genau Definition siehe Abschnitt 5.1.

Interpretation der Eigenschaft (5.0.2): Wenn alles über den Prozess zum Zeitpunkt t_n bekannt ist, spielt es für das Verhalten des Prozesses zum Zeitpunkt $t > t_n$ keine Rolle, was noch weiter in der Vergangenheit (t_1, \dots, t_{n-1}) geschehen ist.

Um diese *Markov-Eigenschaft* (5.0.2) exakt formulieren zu können, benötigen wir den Begriff der *bedingten Erwartung*.

5.1 Bedingte Erwartung, Stopp-Zeiten

Definition (Elementare bedingte Wahrscheinlichkeit)

Für Ereignisse $A, B \in \mathcal{F}$ heißt

$$E(A \mid B) := \begin{cases} \frac{P(A \cap B)}{P(B)} & \text{falls } P(B) > 0 \\ 0 & \text{falls } P(B) = 0 \end{cases}$$

bedingte Wahrscheinlichkeit von A gegeben B .

Bemerkungen: 1) Falls $P(B) > 0$, so ist $A \mapsto E(A \mid B)$ wieder ein Wahrscheinlichkeitsmaß auf (Ω, \mathcal{F}) .

2) Wenn X eine integrierbare Zufallsvariable ist, dann ist X auch bzgl. $P(\cdot \mid B)$ integrierbar (alles im Fall $P(B) > 0$). Der Erwartungswert von

X bzgl. $P(\cdot | B)$ heißt *bedingte Erwartung von X gegeben B* und wird mit $E(X | B)$ bezeichnet.

3) Sei nun A_1, A_2, \dots eine *messbare Zerlegung von (Ω, \mathcal{F})* , d. h.:

- $A_1, A_2, \dots \in \mathcal{F}$
- $A_i \cap A_j = \emptyset$ für $i, j \in \mathbb{N}, i \neq j$
- $\bigcup_{i=1}^{\infty} A_i = \Omega$

Setze

$$\mathcal{A} := \sigma(A_i | i \in \mathbb{N}).$$

Gesucht ist ein ‘‘Ersatz’’ $E(X | \mathcal{A})$ für X , der \mathcal{A} -messbar ist. Dazu ersetzen wir in naheliegender Weise X auf der Menge A_i durch den Mittelwert $E(X | A_i)$:

$$E(X | \mathcal{A}) := \sum_{i=1}^{\infty} E(X | A_i) \mathbf{1}_{A_i}$$

Behauptung *Es gilt:*

- (a) $E(X | \mathcal{A})$ ist \mathcal{A} -messbar.
- (b) $E(X | \mathcal{A})$ ist integrierbar.
- (c)

$$\int_A E(X | \mathcal{A}) dP = \int_A X dP \quad \forall A \in \mathcal{A}$$

Beweis: Übungsaufgabe

Motiviert durch die obige Behauptung definieren wir:

Definition (Bedingte Erwartung)

Sei X eine integrierbare Zufallsvariable und sei $\mathcal{A} \subset \mathcal{F}$ eine σ -Algebra über der Grundmenge Ω .

Eine Zufallsvariable Y heißt *bedingte Erwartung von X gegeben \mathcal{A}* , wenn gilt:

- Y ist \mathcal{A} -messbar.
- Y ist integrierbar.
-

$$\int_A Y dP = \int_A X dP \quad \forall A \in \mathcal{A} \tag{5.1.1}$$

Bemerkung: Diese dritte Eigenschaft (5.1.1) beinhaltet (mit $A = \Omega$) die Gültigkeit von $E(E(X | \mathcal{A})) = E(X)$.

Satz 5.1.1 (Existenz und Eindeutigkeit der bedingten Erwartung)

Für jede integrierbare Zufallsvariable X und jede σ -Algebra $\mathcal{A} \subset \mathcal{F}$ über Ω existiert eine bedingte Erwartung.

Sind Y und Y' bedingte Erwartungen von X , dann folgt

$$Y = Y' \text{ f. s.}$$

Beweis: z. B. “Bauer” oder “Klenke”

Wir stellen die wichtigsten *elementaren Eigenschaften* der bedingten Erwartung zusammen (ohne Beweis); siehe z. B. “Klenke”.

Satz 5.1.2 Seien X, Y integrierbare Zufallsvariablen und \mathcal{A}, \mathcal{C} σ -Algebren über Ω mit $\mathcal{C} \subset \mathcal{A} \subset \mathcal{F}$.

Dann gilt:

(a) (Linearität)

$$E(\lambda X + Y | \mathcal{A}) \stackrel{f.s.}{=} \lambda E(X | \mathcal{A}) + E(Y | \mathcal{A}) \quad \forall \lambda \in \mathbb{R}$$

(b) (Isotonie)

$$X \geq Y \Rightarrow E(X | \mathcal{A}) \stackrel{f.s.}{\geq} E(Y | \mathcal{A})$$

(c)

$$|E(X | \mathcal{A})| \stackrel{f.s.}{\leq} E(|X| | \mathcal{A})$$

(d) Falls XY integrierbar:

$$Y \text{ } \mathcal{A}\text{-messbar} \Rightarrow E(XY | \mathcal{A}) \stackrel{f.s.}{=} Y E(X | \mathcal{A})$$

(e) (Turmeigenschaft)

$$E(X | \mathcal{C}) \stackrel{f.s.}{=} E(E(X | \mathcal{A}) | \mathcal{C}) \stackrel{f.s.}{=} E(E(X | \mathcal{C}) | \mathcal{A})$$

(f) Falls X und \mathcal{A} unabhängig:

$$E(X | \mathcal{A}) = E(X)$$

Bemerkung: Da eine konstante Zufallsvariable unabhängig von allen σ -Algebren ist, folgt aus Satz 5.1.2 (f):

$$E(c | \mathcal{A}) = c \quad \forall c \in \mathbb{R}$$

Definition (Stopp-Zeit)

Gegeben seien eine Indexmenge $T \subset \mathbb{R}$, hier $T = \mathbb{N}_0$ oder $T = \mathbb{R}_+$, und eine Filtration $(\mathcal{F}_t)_{t \in T}$.

Eine (numerische) Zufallsvariable

$$\tau : \Omega \rightarrow \bar{T} := T \cup \{\infty\} \subset \bar{\mathbb{R}}_+$$

heißt *Stopp-Zeit* (bzgl. der Filtration $(\mathcal{F}_t)_{t \in T}$), falls gilt:

$$\{\tau \leq t\} \in \mathcal{F}_t \quad \forall t \in T. \tag{5.1.2}$$

Bemerkungen: 1) Interpretation von (5.1.2): Für die Entscheidung, ob $\tau \leq t$ oder $\tau > t$ genügt die Kenntnis der Verhältnisse bis zum Zeitpunkt t .
 2) Im Fall $T = \mathbb{N}_0$ ist (5.1.2) äquivalent zu (Übungsaufgabe)

$$\{\tau = n\} \in \mathcal{F}_n \quad \forall n \in \mathbb{N}_0.$$

Beispiel: Sei $(\mathcal{F}_n)_{n \in \mathbb{N}_0}$ eine Filtration und sei $(X_n)_{n \in \mathbb{N}_0}$ ein an \mathcal{F}_n adaptierter stochastischer Prozess. Für eine Menge $B \in \mathcal{B}(\mathbb{R})$ setzen wir ($\inf \emptyset := \infty$)

$$\tau_B := \inf\{n \in \mathbb{N}_0 \mid X_n \in B\}.$$

τ ist also der Zeitpunkt, zu dem X_n zum ersten Mal in B ist. Die Stopp-Zeit-Bedingung (5.1.2) besagt, dass man entscheiden kann, ob $\tau \leq n$ ist oder nicht, wenn man den Prozess bis zum Zeitpunkt n kennt (diese Information steht in \mathcal{F}_n). Natürlich kann man entscheiden, ob der Prozess bis zum Zeitpunkt n in B war ($\{\tau_B \leq n\}$), wenn man den Prozess bis zum Zeitpunkt n kennt! Also sollte unsere *erste Eintrittszeit* τ_B eine Stopp-Zeit sein. In der Tat:

$$\{\tau_B = n\} = \{X_n \in B\} \cap \bigcap_{k=0}^{n-1} \{X_k \notin B\}$$

liegt in \mathcal{F}_n . (Im Fall $T = \mathbb{R}_+$ benötigt man zusätzliche Bedingungen über die Pfade des Prozesses (z. B. càdlàg) und/oder die Filtration.)

Man mache sich klar, dass z. B.

$$\sup\{n \in \mathbb{N}_0 \mid X_n \in B\}$$

i. A. keine Stopp-Zeit ist.

Satz 5.1.3 Seien σ und τ Stopp-Zeiten bzgl. einer Filtration $(\mathcal{F}_t)_{t \in T}$, $T = \mathbb{N}_0$ oder $T = \mathbb{R}_+$.

Dann gilt:

(a) $\sigma \wedge \tau$ und $\sigma \vee \tau$ sind Stopp-Zeiten.

(b) $\sigma + \tau$ ist eine Stopp-Zeit.

(c)

$$\mathcal{F}_\tau := \{A \in \mathcal{F} \mid A \cap \{\tau \leq t\} \in \mathcal{F}_t \quad \forall t \in T\}$$

ist eine in \mathcal{F} enthaltene σ -Algebra über Ω .

(d) $\sigma \leq \tau \Rightarrow \mathcal{F}_\sigma \subset \mathcal{F}_\tau$

Beweis: Nur (b), ansonsten siehe "Klenke".

(b): Für $t \in T$ sind $\sigma \wedge t$ und $\tau \wedge t$ nach (a) Stopp-Zeiten (da konstantes $\tau = t$ immer eine Stopp-Zeit ist!). Also:

$$\{\sigma \wedge t \leq s\} \in \mathcal{F}_s \subset \mathcal{F}_t \quad \forall s, t \in T, s \leq t.$$

Hieraus folgt, dass $\sigma \wedge t : \Omega \rightarrow T$ messbar bzgl. \mathcal{F}_t ist. Damit ist auch

$$X := (\sigma \wedge t) + \mathbf{1}_{\{\sigma > t\}}$$

\mathcal{F}_t -messbar. Genauso ist

$$Y := (\tau \wedge t) + \mathbf{1}_{\{\tau > t\}}$$

\mathcal{F}_t -messbar. Somit ist $X + Y$ messbar bzgl. \mathcal{F}_t . Nun überzeugt man sich mittels Fallunterscheidung, dass

$$X + Y \leq t \Leftrightarrow \sigma + \tau \leq t$$

gilt. \square

Definition Für eine *endliche* Stopp-Zeit $\tau : \Omega \rightarrow T$ setzen wir:

$$X_\tau := X_{\tau(\omega)}(\omega), \quad \omega \in \Omega$$

Satz 5.1.4 Für einen an eine gegebene Filtration $(\mathcal{F}_n)_{n \in \mathbb{N}_0}$ adaptierten Prozess $(X_n)_{n \in \mathbb{N}_0}$ und eine endliche Stopp-Zeit τ gilt:

X_τ ist \mathcal{F}_τ -messbar.

Beweis: Für $B \in \mathcal{B}(\mathbb{R})$ und $n \in \mathbb{N}_0$ haben wir:

$$\begin{aligned} \{X_\tau \in B\} \cap \{\tau \leq n\} &= \bigcup_{k=0}^n (\{X_\tau \in B\} \cap \{\tau = k\}) \\ &= \bigcup_{k=0}^n (\{X_k \in B\} \cap \{\tau = k\}) \\ &\in \mathcal{F}_n \square \end{aligned}$$

5.2 Markov-Halbgruppen

Definition (Markov-Kern)

Für einen messbaren Raum (E, \mathcal{E}) heißt die Abbildung

$$K : E \times \mathcal{E} \rightarrow [0, 1]$$

Markov-Kern auf (E, \mathcal{E}) , wenn gilt:

$\cdot x \mapsto K(x, A)$ ist \mathcal{E} -messbar für alle $A \in \mathcal{E}$.

· $A \mapsto K(x, A)$ ist ein Wahrscheinlichkeitsmaß auf (E, \mathcal{E}) für alle $x \in E$.

Definition (Markov-Halbgruppe)

Eine *Markov-Halbgruppe* auf dem messbaren Raum (E, \mathcal{E}) ist eine Familie $(K_t)_{t \in \mathbb{R}_+}$ mit:

· $K_t, t \in \mathbb{R}_+$ ist ein Markov-Kern auf (E, \mathcal{E}) .

$$K_{s+t}(x, A) = \int K_t(y, A) K_s(x, dy) \quad \forall s, t \in \mathbb{R}_+ \quad (5.2.1)$$

Bemerkungen: 1) Die linke Seite von (5.2.1) ergibt Sinn, weil

$$y \mapsto K_t(y, A) \in [0, 1]$$

messbar und beschränkt, diese Funktion somit $K_s(x, \cdot)$ -integrierbar ist.

2) Wegen $s + t = t + s$ gilt für eine Markov-Halbgruppe auch:

$$K_{s+t}(x, A) = \int K_s(y, A) K_t(x, dy) \quad \forall s, t \in \mathbb{R}_+$$

3) Die Gleichungen (5.2.1) heißen *Chapman-Kolmogorov-Gleichungen*. Interpretation: $K_t(x, A)$ ist die Wahrscheinlichkeit, nach t Zeiteinheiten in A zu sein, wenn man sich gerade in x befindet. Die Gleichung (5.2.1) besagt, dass man die Wahrscheinlichkeiten für die Zeitspanne $s + t$ durch Integration aller möglichen *Zwischenaufenthalte* nach t Zeiteinheiten bekommt.

Beispiele: 1) (Endliche Markov-Ketten in kontinuierlicher Zeit)

Sei $\mathbf{X} = (X_t)_{t \in \mathbb{R}_+}$ ein stochastischer Prozess mit *endlichem Zustandsraum* S . (Man kann offenbar immer $S = \{1, \dots, n\}$ falls $n = \#S$ annehmen.) Ist \mathbf{X} eine (zeitlich homogene) Markov-Kette, so gilt

$$P(X_t = i \mid X_{t_1} = i_1, \dots, X_{t_m} = i_m) = P(X_t = i \mid X_{t_m} = i_m)$$

für alle $m \in \mathbb{N}$,

$$i, i_1, \dots, i_m \in S, 0 \leq i_1 \leq i_2 \leq \dots \leq i_m \leq i.$$

(Hier stehen elementare bedingte Wahrscheinlichkeiten!) Die $n \times n$ -Matrizen $(t \in \mathbb{R}_+)$

$$(P_t)_{ij} := P(X_{s+t} = j \mid X_s = i)$$

hängen nicht von $s \in \mathbb{R}_+$ ab. Sie heißen *Matrizen der Übergangswahrscheinlichkeiten* von \mathbf{X} . Setze noch $\mu_i := P(X_0 = i), i \in S$. Dann heißt $\mu = (\mu_1, \dots, \mu_n) \in \mathbb{R}^n$ *Vektor der Anfangswahrscheinlichkeiten* von \mathbf{X} .

Die Markov-Kette ist durch $(P_t)_{t \in \mathbb{R}_+}$ und μ bis auf stochastische Äquivalenz eindeutig bestimmt und es gilt

$$P_{s+t} = P_s P_t, \quad (5.2.2)$$

d. h. die P_t bilden eine Matrix-Halbgruppe, die *Halbgruppe der Übergangswahrscheinlichkeiten* von \mathbf{X} .

Die Gleichungen (5.2.2) sind genau die Chapman-Kolmogorov-Gleichungen:

· $K_s(i, \cdot)$ ist das durch den i -ten Zeilenvektor

$$((P_s)_{i1}, \dots, (P_s)_{in})$$

von P_s gegebene Wahrscheinlichkeitsmaß auf $E = S = \{1, \dots, n\}$.

· Aus (5.2.1) wird

$$(P_{s+t})_{ij} = \sum_{l=1}^n (P_s)_{il} (P_t)_{lj},$$

d. h. (5.2.2).

Stochastische Prozesse

SoSe 2018

Übersicht 12. Vorlesungswoche (2.7.-6.7.2018)

Beispiele: 2) (Faltungshalbgruppen von Wahrscheinlichkeitsmaßen)

Sei $(\mu_t)_{t \in \mathbb{R}_+}$ eine Faltungshalbgruppe von Wahrscheinlichkeitsmaßen auf dem messbaren Raum $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$; siehe Abschnitt 2.3.

Setze für $x \in \mathbb{R}$, $B \in \mathcal{B}(\mathbb{R})$:

$$K_t(x, B) := \mu_t(B - x).$$

Behauptung: $(K_t)_{t \in \mathbb{R}_+}$ ist eine Markov-Halbgruppe auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

Beweis: Zunächst überlegen wir uns, dass die K_t Markov-Kerne sind: Die Messbarkeit von $x \mapsto \mu(B - x)$ für jedes feste B folgt daraus, dass

$$\mu(B - x) = \int \mathbf{1}_{B-x}(y) \mu(dy) = \int \mathbf{1}_B(x + y) \mu(dy)$$

gilt: Die Funktion $(x, y) \mapsto \mathbf{1}_B(x + y)$ und damit auch die Funktion

$$x \mapsto \int \mathbf{1}_B(x + y) \mu(dy)$$

sind messbar (siehe Satz von Fubini). Dass $B \mapsto K_t(x, B)$ für jedes feste x ein Wahrscheinlichkeitsmaß ist, folgt daraus, dass $K_t(x, \cdot)$ das Bildmaß von μ unter der Abbildung $y \mapsto x + y$ ist. ✓

Bleiben noch die Chapman-Kolmogorov-Gleichungen (5.2.1) zu zeigen: Für $f \in L^\infty(\Omega, \mathcal{F}, P)$ gilt

$$\int f(y) K_t(x, dy) = \int f(x + y) \mu_t(dy), \quad (5.2.2)$$

insbesondere

$$K_t(y, A) = \int \mathbf{1}_A(z) K_t(y, dz) = \int \mathbf{1}_A(y + z) \mu_t(dz),$$

und somit

$$\begin{aligned}
\int K_t(y, A) K_s(x, dy) &\stackrel{(5.2.1)}{=} \int \left(\int \mathbf{1}_A(y+z) \mu_t(dz) \right) K_s(x, dy) \\
&\stackrel{(5.2.1)}{=} \int \int \mathbf{1}_A(x+y+z) d\mu_t(z) d\mu_s(y) \\
&= \int \mathbf{1}_A(x+z) d(\mu_s \star \mu_t)(z) \\
&= \int \mathbf{1}_A(x+z) d\mu_{s+t}(z) \\
&= \int \mathbf{1}_A(z) K_{s+t}(x, dz) \\
&= K_{s+t}(x, A). \checkmark \checkmark
\end{aligned}$$

Bemerkung: Wir starten umgekehrt mit einer Markov-Halbgruppen $(K_t)_{t \in \mathbb{R}_+}$ auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, die *räumlich homogen* ist, d. h. für die gilt:

$$K_t(x, B) = K_t(x+z, B+z) \quad \forall x, z \in \mathbb{R} \quad \forall B \in \mathcal{B}(\mathbb{R}). \quad (5.2.3)$$

Setze $\mu_t(B) := K_t(0, B)$. Dann folgt aus den Chapman-Kolmogorov-Gleichungen für die K_t , dass die μ_t eine Faltungshalbgruppe von Wahrscheinlichkeitsmaßen bilden (Übungsaufgabe). Natürlich gilt $K_t(x, B) = \mu_t(B-x)$. Faltungshalbgruppen von Wahrscheinlichkeitsmaßen und räumlich homogene Markov-Halbgruppen stehen also in 1-1-Korrespondenz.

Es ergeben sich zwei naheliegende Fragen: 1. Kann man einer Markov-Halbgruppe auf $(K_t)_{t \in \mathbb{R}_+}$ ähnlich wie im Fall einer Faltungshalbgruppe von Wahrscheinlichkeitsmaßen in kanonischer Weise eine projektive Familie von Wahrscheinlichkeitsmaßen (siehe Satz 2.3.3) – und damit einen stochastischen Prozess – zuordnen? 2. Wenn dem so ist, welche Eigenschaften zeichnen diese so aus den Markov-Halbgruppen konstruierten Prozesse aus (im Fall einer Faltungshalbgruppe ergaben sich gerade die Lévy-Prozesse)?

Zur ersten Frage :

Satz 5.2.1 *Gegeben seien eine Markov-Halbgruppe $(K_t)_{t \in \mathbb{R}_+}$ und ein Wahrscheinlichkeitsmaß μ auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.*

Dann wird durch $(I = \{t_1 < \dots < t_n\}, f \in L^\infty(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n)))$

$$\begin{aligned}
&\int_{\mathbb{R}^n} f dP_I \\
&= \int_{\mathbb{R}} \int_{\mathbb{R}^n} f(x_1, \dots, x_n) K_{t_n-t_{n-1}}(x_{n-1}, dx_n) K_{t_{n-1}-t_{n-2}}(x_{n-2}, dx_{n-1}) \dots \\
&\quad \dots K_{t_2-t_1}(x_1, dx_2) K_{t_1}(x, dx_1) \mu(dx)
\end{aligned}$$

eine projektive Familie $(P_I)_{I \in \mathcal{P}_0(\mathbb{R}_+)}$ von Wahrscheinlichkeitsmaßen definiert.

Beweis: Analog zum Beweis von Satz 2.3.3.

Definition In der Situation von Satz 5.2.1 bezeichnen wir das nach Satz 2.2.2 (Satz von Daniell-Kolmogorov) existierende Wahrscheinlichkeitsmaß auf $(\mathbb{R}^{\mathbb{R}_+}, \mathcal{B}(\mathbb{R}^{\mathbb{R}_+}))$ mit P^μ . Es hängt bei gegebener Markov-Halbgruppe nur von der ‘‘Startwahrscheinlichkeit’’ μ ab. Insbesondere setzen wir

$$P^x := P^{\delta_x}$$

für $x \in \mathbb{R}$, wobei δ_x wieder das Dirac-Maß im Punkte x bezeichnet.

Bemerkung: Wir erhalten also eine Familie $(P^x)_{x \in \mathbb{R}}$ von Wahrscheinlichkeitsmaßen auf $(\Omega, \mathcal{F}) = (\mathbb{R}^{\mathbb{R}_+}, \mathcal{B}(\mathbb{R}^{\mathbb{R}_+}))$. Der stochastische Prozess

$$X_t = p_{\{t\}} : \Omega \rightarrow \mathbb{R},$$

über $(\Omega, \mathcal{F}, P^x)$, d. h. die Projektion auf die t -te Koordinate, besitzt dann als Familie der endlich-dimensionalen Verteilungen die nach Satz 5.2.1 zu K_t und $\mu = \delta_x$ gehörige projektive Familie.

5.3 Elementare, schwache und starke Markov-Eigenschaft

Definition (Elementare Markov-Eigenschaft)

Ein stochastischer Prozess $(X_t)_{t \in \mathbb{R}_+}$ besitzt die *elementare Markov-Eigenschaft*, wenn gilt:

$$P(X_t \in B \mid X_{t_1}, \dots, X_{t_n}) \stackrel{f.s.}{=} P(X_t \in B \mid X_{t_n}) \quad (5.3.1)$$

für alle $n \in \mathbb{N}$, $0 \leq t_1 \leq t_2 \leq \dots \leq t_n \leq t$ und alle $B \in \mathcal{B}(\mathbb{R})$.

Satz 5.3.1 (Voraussetzungen wie in Satz 5.2.1)

Der Prozess $X_t = p_{\{t\}}$ besitzt die *elementare Markov-Eigenschaft* und es gilt:

$$P^\mu(X_t \in B \mid X_{t_1}, \dots, X_{t_n}) = K_{t-t_n}(X_{t_n}, B) \quad (5.3.2)$$

P^μ -f.s. für alle $n \in \mathbb{N}$, $0 \leq t_1 \leq t_2 \leq \dots \leq t_n \leq t$ und alle $B \in \mathcal{B}(\mathbb{R})$.

Beweis: Zum Beweis von (5.3.2) müssen wir zeigen, dass die Zufallsvariable $K_{t-t_n}(X_{t_n}, B)$ eine Version der bedingten Erwartung

$$E^\mu(\mathbf{1}_{\{X_t \in B\}} \mid X_{t_1}, \dots, X_{t_n})$$

ist. Da $K_{t-t_n}(\cdot, B)$ nach Voraussetzung $\mathcal{B}(\mathbb{R})$ -messbar und X_{t_n} natürlich $\sigma(X_{t_1}, \dots, X_{t_n})$ -messbar ist, erhalten wir jedenfalls die $\sigma(X_{t_1}, \dots, X_{t_n})$ -Messbarkeit von $K_{t-t_n}(X_{t_n}(\cdot), B)$. Noch zu zeigen ist

$$\begin{aligned} \int_A K_{t-t_n}(X_{t_n}, B) dP^\mu &\stackrel{!}{=} \int_A \mathbf{1}_{\{X_t \in B\}} dP^\mu \\ &= P^\mu(\{X_t \in B\} \cap A) \end{aligned} \quad (5.3.3)$$

für alle $A \in \mathcal{A} := \sigma(X_{t_1}, \dots, X_{t_n})$. Nach einem Standardargument genügt es dabei den Fall

$$A = \{X_{t_1} \in B_1, \dots, X_{t_n} \in B_n\}, \quad B_1, \dots, B_n \in \mathcal{B}(\mathbb{R}), \quad (5.3.4)$$

zu betrachten (die Mengen der Gestalt (5.3.4) bilden einen \cap -stabilen Erzeuger von \mathcal{A}). Es gilt ($I = \{t_1 < t_2 < \dots < t_n\}$):

$$\begin{aligned} &\int_A K_{t-t_n}(X_{t_n}, B) dP^\mu && (5.3.5) \\ &= \int K_{t-t_n}(X_{t_n}, B) \mathbf{1}_A dP^\mu \\ &= \int K_{t-t_n}(X_{t_n}, B) \mathbf{1}_{B_1 \times \dots \times B_n}(X_{t_1}, \dots, X_{t_n}) dP^\mu \\ &= \int_{\mathbb{R}^n} K_{t-t_n}(X_{t_n}, B) \mathbf{1}_{B_1}(x_1) \dots \mathbf{1}_{B_n}(x_n) dP_{(X_{t_1}, \dots, X_{t_n})}^\mu(x_1, \dots, x_n) \\ &= \int_{\mathbb{R}^n} K_{t-t_n}(X_{t_n}, B) \mathbf{1}_{B_1}(x_1) \dots \mathbf{1}_{B_n}(x_n) dP_I^\mu(x_1, \dots, x_n) \\ &= \int_{\mathbb{R}^{n+1}} K_{t-t_n}(x_n, B) \mathbf{1}_{B_1}(x_1) \dots \mathbf{1}_{B_n}(x_n) \\ &\quad K_{t_n-t_{n-1}}(x_{n-1}, dx_n) \dots K_{t_1}(x, dx_1) \mu(dx) \\ &= \int_{\mathbb{R}} \int_{B_1} \dots \int_{B_n} \int_B K_{t-t_n}(x_n, dx_{n+1}) \\ &\quad K_{t_n-t_{n-1}}(x_{n-1}, dx_n) \dots K_{t_1}(x, dx_1) \mu(dx) \\ &= P_{I \cup \{t\}}^\mu(B_1 \times \dots \times B_n \times B) \\ &= P^\mu(X_{t_1} \in B_1, \dots, X_{t_n} \in B_n, X_t \in B) \\ &= P^\mu(\{X_t \in B\} \cap A). \end{aligned}$$

Die elementare Markov-Eigenschaft ergibt sich nun aus

$$P^\mu(X_t \in B | X_{t_1}, \dots, X_{t_n}) \stackrel{f.s.}{=} K_{t-t_n}(X_{t_n}, B) \stackrel{f.s.}{=} P^\mu(X_t \in B | X_{t_n}). \square$$

Definition (Markov-Prozess)

Für eine Familie

$$(\Omega, \mathcal{F}, \mathbb{P}^x)_{x \in \mathbb{R}} = (\Omega, \mathcal{F}, (\mathbb{P}^x)_{x \in \mathbb{R}})$$

von Wahrscheinlichkeitsräumen (nur das Wahrscheinlichkeitsmaß variiert!) heißt eine Familie

$$(X_t)_{t \in \mathbb{R}_+}$$

von Zufallsvariablen (d. h. die $X_t : \Omega \rightarrow \mathbb{R}$ sind \mathcal{F} - $\mathcal{B}(\mathbb{R})$ -messbar für alle $t \in \mathbb{R}_+$) *Markov-Prozess*, wenn gilt:

- Die Abbildungen $x \mapsto \mathbb{P}^x(A)$ von \mathbb{R} nach $[0, 1]$ sind messbar für alle $A \in \mathcal{F}$
- Es gilt die *schwache Markov-Eigenschaft*

$$\mathbb{P}^x(X_t \in B \mid X_{t_1}, \dots, X_{t_n}) \stackrel{f.s.}{=} \mathbb{P}^{X_{t_n}}(X_{t-t_n} \in B) \quad (5.3.6)$$

für alle $n \in \mathbb{N}$, $0 \leq t_1 \leq t_2 \leq \dots \leq t_n \leq t$, $B \in \mathcal{B}(\mathbb{R})$.

Bemerkung: 1) Die schwache Markov-Eigenschaft (5.3.6) ist äquivalent zu

$$\mathbb{P}^x(X_t \in B \mid \mathcal{F}_s) \stackrel{f.s.}{=} \mathbb{P}^{X_s}(X_{t-s} \in B) \quad (5.3.7)$$

für alle $0 \leq s \leq t$, $B \in \mathcal{B}(\mathbb{R})$ oder

$$\mathbb{P}^x(X_{s+t} \in B \mid \mathcal{F}_s) \stackrel{f.s.}{=} \mathbb{P}^{X_s}(X_t \in B) \quad (5.3.8)$$

für alle $s, t \in \mathbb{R}_+$, $B \in \mathcal{B}(\mathbb{R})$. Dabei ist $\mathcal{F}_s = \sigma(X_r \mid r \leq s)$.

2) Zwei Markov-Prozesse heißen stochastisch äquivalent, wenn die Prozesse für jedes $x \in \mathbb{R}$ stochastisch äquivalent sind.

Satz 5.3.2 (a) *Zu jeder Markov-Halbgruppe $(K_t)_{t \in \mathbb{R}_+}$ existiert ein Markov-Prozess*

$$(\Omega, \mathcal{F}, (\mathbb{P}^x)_{x \in \mathbb{R}}, (X_t)_{t \in \mathbb{R}_+}), \quad (5.3.9)$$

so dass gilt

$$K_t(x, B) = \mathbb{P}^x(X_t \in B) \quad \forall t \in \mathbb{R}_+, \forall x \in \mathbb{R}, \forall B \in \mathcal{B}(\mathbb{R}). \quad (5.3.10)$$

(b) *Für jeden Markov-Prozess (5.3.9) wird durch (5.3.10) eine Markov-Halbgruppe definiert.*

(c) *Zwei Markov-Prozesse sind genau dann stochastisch äquivalent, wenn die gemäß (b) gegebenen Markov-Halbgruppen übereinstimmen.*

Beweis: (a): Wir nehmen natürlich

$$(\Omega, \mathcal{F}, (\mathbb{P}^x)_{x \in \mathbb{R}}, (X_t)_{t \in \mathbb{R}_+}) = (\mathbb{R}^{\mathbb{R}_+}, \mathcal{B}(\mathbb{R}^{\mathbb{R}_+}), (\mathbb{P}^x)_{x \in \mathbb{R}}, (p_{\{t\}})_{t \in \mathbb{R}_+})$$

gemäß Satz 5.2.1 mit $\mu = \delta_x$. Dann gilt

$$\begin{aligned} & \mathbb{P}^x(X_{t_1} \in B_1, \dots, X_{t_n} \in B_n) \\ &= \int_{B_1} \dots \int_{B_n} K_{t_n - t_{n-1}}(x_{n-1}, dx_n) \dots K_{t_1}(x, dx_1), \end{aligned}$$

was die Messbarkeit der Abbildung

$$x \mapsto \mathbb{P}^x(X_{t_1} \in B_1, \dots, X_{t_n} \in B_n)$$

zeigt. Mit einem Standardargument folgt die Messbarkeit der Abbildung $x \mapsto \mathbb{P}^x(A)$ für jedes feste $A \in \mathcal{B}(\mathbb{R}^{\mathbb{R}_+})$, da $\mathcal{B}(\mathbb{R}^{\mathbb{R}_+})$ von den Mengen

$$\{X_{t_1} \in B_1, \dots, X_{t_n} \in B_n\}$$

erzeugt wird. ✓

Gleichung (5.3.10): Es gilt für $t \geq 0$ und $B \in \mathcal{B}(\mathbb{R})$:

$$\begin{aligned} \mathbb{P}_{\{t\}}^x(B) &= \int K_t(y, B) \delta_x(dy) & (5.3.11) \\ &= K_t(x, B) \\ &= \mathbb{P}^x(X_t \in B). \quad \checkmark \end{aligned}$$

X_t ist ein Markov-Prozess: X_t besitzt nach Satz 5.2.1 als Prozess über

$$(\mathbb{R}^{\mathbb{R}_+}, \mathcal{B}(\mathbb{R}^{\mathbb{R}_+}), \mathbb{P}^x)$$

die elementare Markov-Eigenschaft und also:

$$\begin{aligned} \mathbb{P}^x(X_t \in B \mid X_{t_1}, \dots, X_{t_n}) &= K_{t-t_n}(X_{t_n}, B) \\ &\stackrel{(5.3.11)}{=} \mathbb{P}^{X_{t_n}}(X_{t-t_n} \in B). \end{aligned}$$

(b): Da $x \mapsto \mathbb{P}^x(A)$ messbar ist für alle $A \in \mathcal{F}$, ist insbesondere

$$x \mapsto \mathbb{P}^x(X_t \in B) = K_t(x, B)$$

messbar für alle $B \in \mathcal{B}(\mathbb{R})$. Außerdem ist

$$B \mapsto K_t(x, B) = \mathbb{P}^x(X_t \in B) \quad (5.3.12)$$

gerade die Verteilung der Zufallsvariablen X_t bzgl. P^x , d. h. (5.3.12) ist ein Wahrscheinlichkeitsmaß. Es folgt, dass die K_t jedenfalls Markov-Kerne sind. ✓

Chapman-Komlogorov-Gleichungen:

$$\begin{aligned}
K_{s+t}(x, B) &= P^x(X_{s+t} \in B) \\
&= E^x(\mathbf{1}_{\{X_{s+t} \in B\}}) \\
&= E^x(E^x(\mathbf{1}_{\{X_{s+t} \in B\}} | \mathcal{F}_s)) \\
&= E^x(P^x(X_{s+t} \in B) | \mathcal{F}_s) \\
&= E^x(P^{X_s}(X_t \in B)) \\
&= E^x(K_t(X_s, B)) \\
&= \int K_t(X_s(\omega), B) P^x(d\omega) \\
&= \int K_t(y, B) P_{X_s}^x(dy) \\
&= \int K_t(y, B) K_s(x, dy),
\end{aligned}$$

da die Verteilung von X_s bzgl. P^x durch $K_s(x, \cdot)$ gegeben ist. ✓

(c): Die endlich-dimensionalen Verteilungen des Prozesses X_t über $(\Omega, \mathcal{F}, P^x)$ werden für einen Markov-Prozess durch seine Markov-Halbgruppe (5.3.10) festgelegt, da

$$P^x(X_{t_1} \in B_1) = K_{t_1}(x, B_1)$$

gilt und für $n \geq 2$

$$\begin{aligned}
&P^x(X_{t_1} \in B_1, \dots, X_{t_n} \in B_n) \\
&= \int_{\{X_{t_1} \in B_1, \dots, X_{t_{n-1}} \in B_{n-1}\}} \mathbf{1}_{\{X_{t_n} \in B_n\}} dP^x \\
&= \int_{\{X_{t_1} \in B_1, \dots, X_{t_{n-1}} \in B_{n-1}\}} P^{X_{t_{n-1}}}(X_{t_n - t_{n-1}} \in B) dP^x \\
&= \int_{\{X_{t_1} \in B_1, \dots, X_{t_{n-1}} \in B_{n-1}\}} K_{t_n - t_{n-1}}(X_{t_{n-1}}, B) dP^x. \square
\end{aligned}$$

Definition (Starke Markov-Eigenschaft)

Unter zusätzlichen technischen Voraussetzungen (siehe den Abschnitt über Stopp-Zeiten) heißt

$$(\Omega, \mathcal{F}, (P^x)_{x \in \mathbb{R}}, (X_t)_{t \in \mathbb{R}_+})$$

starker Markov-Prozess, wenn die Bedingung (5.3.8) durch die folgende stärkere Bedingung ersetzt wird:

$$P^x(X_{\sigma+t} \in B \mid \mathcal{F}_\sigma) \stackrel{f.s.}{=} P^{X_\sigma}(X_t \in B) \quad (5.3.13)$$

für alle $t \geq 0$, $B \in \mathcal{B}(\mathbb{R})$ und *alle endlichen Stopp-Zeiten* σ .

Beispiele für Prozesse, die die starke Markov-Eigenschaft besitzen, erhält man durch die Lévy-Prozesse, insbesondere also die Brownsche Bewegung und den Poisson-Prozess.

Bemerkung: Zusammenfassend haben wir:

- Markov-Prozesse und Markov-Halbgruppen stehen in 1-1-Korrespondenz (bis auf stochastische Äquivalenz bei den Markov-Prozessen).
- Faltungshalbgruppen liefern (spezielle!) Beispiele von Markov-Halbgruppen und damit von Markov-Prozessen, nämlich die räumlich homogenen. Lévy-Prozesse besitzen die elementare Markov-Eigenschaft und können als die räumlich homogenen Markov-Prozesse angesehen werden.
- Matrix-Halbgruppen von Übergangswahrscheinlichkeiten liefern ebenfalls Beispiele von Markov-Halbgruppen und damit von Markov-Prozessen, nämlich die endlichen Markov-Ketten in stetiger Zeit.
- Die Brownsche Bewegung (sogar alle Lévy-Prozesse) besitzen die starke Markov-Eigenschaft.

Stochastische Prozesse

SoSe 2018

Übersicht 13. Vorlesungswoche 9.7.-613.7.2018)

6 Itô-Isometrie und stochastische IntegrationZiel: Wir suchen eine “gute” Definition des Integrals ($T \in \mathbb{R}_+$)

$$\int_0^T F_t dB_t \tag{6.0.2}$$

mit einer Brownschen Bewegung B_t (über einem Wahrscheinlichkeitsraum (Ω, \mathcal{F}, P)) als *Integrator-Prozess* und für eine möglichst große Klasse stochastischer Prozesse F_t (über (Ω, \mathcal{F}, P)) als *Integranden-Prozesse*. Es geht also um die Definition der Zufallsvariablen

$$\omega \mapsto \left(\int_0^T F_t dB_t \right)(\omega), \quad \omega \in \Omega.$$

Wenn wir schreiben

$$\omega \mapsto \left(\int_0^T F_t dB_t \right)(\omega) = \int_0^T F_t(\omega) dB_t,$$

so mag dies eine pfadweise Definition des Integrals (6.0.2) suggerieren. Wie wir aber in Satz 4.3.5 gesehen haben, ist eine solche pfadweise Definition – jedenfalls als Riemann-Stieltjes-Integral – zumindest problematisch, da die Pfade der Brownschen Bewegung nicht von beschränkter Variation sind.

Wir werden bald sehen, dass dagegen eine Integration *adaptierter* Prozesse in einem bestimmten Sinne möglich ist. Dabei bedeutet “adaptiert” adaptiert an die Filtration \mathcal{F}_t der Brownschen Bewegung, d. h. ein Prozess heißt adaptiert, wenn F_t messbar bzgl. \mathcal{F}_t ist (wir schreiben etwas missbräuchlich: $F_t \in \mathcal{F}_t$) für alle $t \in \mathbb{R}_+$.

Die Integrale werden dabei in einem “ L^2 -Sinne” definiert, und zwar zunächst für *einfache adaptierte Prozesse*. Das sind Prozesse der Gestalt

$$F_t(\omega) = \sum_{k=0}^m X_k(\omega) \mathbf{1}_{[t_k, t_{k+1}]}(t) \tag{6.0.3}$$

mit $\{0 = t_0 < t_1 < \dots < t_m\} \in \mathbf{Par}_{0,T}$ und $X_k \in \mathcal{F}_{t_k}$. Dabei ist das *Itô-Integral* von F_t bzgl. der Brownschen Bewegung durch

$$\left(\int_0^T F_t dB_t\right)(\omega) := \sum_{k=0}^m X_k(\omega) (B_{t_{k+1}}(\omega) - B_{t_k}(\omega)) \quad (6.0.4)$$

gegeben, d. h. als eine Riemann-Stieltjes-Summe. Beachte aber die Zusatzbedingung " $X_k \in \mathcal{F}_{t_k}$ ", die sich als wesentlich erweist und die gerade bedeutet, dass der einfache Prozess (6.0.3) adaptiert ist.

Als wesentliches Mittel zur Definition von (6.0.2) verwenden wir die *Itô-Isometrie*

$$\int_{\Omega} \left(\int_0^T F_t dB_t\right)^2(\omega) dP(\omega) = \int_0^T \int_{\Omega} (F_t(\omega))^2 dP(\omega) dt, \quad (6.0.5)$$

die eine Ausdehnung der Definition des Itô-Integrals von einfachen adaptierten Prozessen (wie in (6.0.4)) auf alle adaptierten Prozesse

$$F_t \in L^2(\Omega \times [0, T], \mathcal{F} \otimes \mathcal{B}([0, T]), P \otimes \lambda \upharpoonright [0, T])$$

ermöglicht.

6.1 Einige Grundbegriffe aus der Theorie der Hilberträume

Definition (Hilbertraum)

Ein Vektorraum H (über \mathbb{C}) mit Skalarprodukt

$$\begin{aligned} (v, w) &\mapsto \langle v, w \rangle \\ H \times H &\rightarrow \mathbb{C}, \end{aligned} \quad (6.1.1)$$

der bzgl. der durch das Skalarprodukt gegebenen Norm

$$v \mapsto \|v\| := \sqrt{\langle v, v \rangle}$$

ein vollständiger normierter Vektorraum (d. h. ein *Banachraum*) ist, heißt *Hilbertraum*.

Skalarprodukte sind hier anti-linear im ersten und linear im zweiten Eingang. Oft schließen wir stillschweigend aus, dass H trivial ist, d. h. wir nehmen $H \neq \{0\}$.

Beispiel Sei $(\Omega, \mathcal{F}, \mu)$ ein Maßraum. Setze

$$L^2(\Omega, \mathcal{F}, \mu) = L^2(\mu) := \{[f] \mid f : \Omega \rightarrow \mathbb{C} \text{ messbar und } \int |f|^2 d\mu < \infty\},$$

wobei

$$[f] := \{g : \Omega \rightarrow \mathbb{C} \mid g \text{ messbar und } \mu(f \neq g) = 0\}.$$

(Wir schreiben ab jetzt wieder f für $[f]$.) Dann ist für $f, g \in L^2(\mu)$ das Produkt fg μ -integrierbar, so dass wir

$$\langle f, g \rangle := \int \bar{f}g \, d\mu \tag{6.1.2}$$

setzen können; siehe z. B. “Elstrodt”. $L^2(\mu)$ wird mit dem Skalarprodukt (6.1.2) ein Hilbertraum. Wir betrachten einige Spezialfälle:

1) Für $d \in \mathbb{N}$ betrachte den Maßraum

$$(\{1, \dots, d\}, \mathcal{P}(\{1, \dots, d\}), \alpha),$$

wobei α das Zählmaß auf $\mathcal{P}(\{1, \dots, d\})$ bezeichnet. Die Elemente von $L^2(\alpha)$ kann man mittels der Zuordnung

$$f \leftrightarrow (f(1), \dots, f(d))^T$$

als (Spalten-)Vektoren von \mathbb{C}^d auffassen. Der Hilbertraum $L^2(\alpha)$ ist der \mathbb{C}^d mit dem üblichen Skalarprodukt

$$\langle (x_1, \dots, x_d)^T, (y_1, \dots, y_d)^T \rangle = \sum_{n=1}^d \bar{x}_n y_n.$$

2) Nun zum Maßraum

$$(\mathbb{N}, \mathcal{P}(\mathbb{N}), \alpha)$$

(α wieder das Zählmaß). In diesem Fall schreibt man auch $l^2(\mathbb{N})$ für den Hilbertraum $L^2(\alpha)$, und es ist

$$l^2(\mathbb{N}) = \{(c_n)_{n \in \mathbb{N}} \mid c_n \in \mathbb{C}, n \in \mathbb{N}, \text{ mit } \sum_{n=1}^{\infty} |c_n|^2 < \infty\}$$

der Raum der quadrat-summierbaren Folgen komplexer Zahlen. Aus (6.1.2) wird

$$\langle (c_n)_{n \in \mathbb{N}}, (d_n)_{n \in \mathbb{N}} \rangle = \sum_{n=1}^{\infty} \bar{c}_n d_n.$$

3) Betrachte als nächstes den Maßraum

$$(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+), \lambda \upharpoonright \mathbb{R}_+)$$

mit dem Lebesgue-Maß λ . Jetzt schreiben wir $L^2(\mathbb{R}_+)$ für $L^2(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+), \lambda \upharpoonright \mathbb{R}_+)$. Wir erhalten den Hilbertraum der quadrat-integrierbaren komplexwertigen Funktionen auf \mathbb{R}_+ . Stetige Funktionen mit kompaktem Träger liegen z. B. in $L^2(\mathbb{R}_+)$. Ganz entsprechend kann man $L^2(\mathbb{R})$ oder $L^2([0, T])$, $T \in \mathbb{R}_+$, definieren.

4) Natürlich kann man auch $L^2(\Omega, \mathcal{F}, P)$ für einen Wahrscheinlichkeitsraum (Ω, \mathcal{F}, P) betrachten. Dann wird aus (6.1.2)

$$\langle X, Y \rangle = E(\overline{XY})$$

für (komplexwertige) quadrat-integrierbare Zufallsvariablen X und Y . Wir schreiben in diesem Fall oft $L^2(\Omega)$ für $L^2(\Omega, \mathcal{F}, P)$. Der Raum $L^2(\Omega)$ enthält alle beschränkten Zufallsvariablen.

Grundfakten: 0) Wenn für $v, u \in H$ gilt $\langle v, w \rangle = \langle u, w \rangle$ für alle $w \in H$, dann folgt $\langle v - u, v - u \rangle = 0$ und damit $v = u$.

1) Es gilt die *Cauchy-Schwarzsche Ungleichung*

$$|\langle v, w \rangle| \leq \|v\| \|w\|$$

für alle $v, w \in H$.

2) Eine lineare Abbildung $T : H \rightarrow K$, H, K Hilberträume, ist *genau dann stetig*, wenn sie *beschränkt* ist, d. h. wenn

$$\|T\| := \sup\{\|Tv\| \mid v \in H, \|v\| = 1\} < \infty$$

gilt. Durch $\|T\|$ wird der Vektorraum

$$B(H, K) := \{T : H \rightarrow K \mid T \text{ linear und stetig}\}$$

zu einem Banachraum, d. h. zu einem vollständigen normierten Vektorraum. Die Elemente von $B(H) := B(H, H)$ heißen *lineare Operatoren auf H* .

3) Nach dem *Satz von Riesz* sind die stetigen linearen Abbildungen von H nach \mathbb{C} genau die Abbildungen der Form

$$v \mapsto \langle u, v \rangle,$$

wobei u ein geeignetes Element von H ist. Dabei ist u durch die stetige lineare Abbildung von H nach \mathbb{C} eindeutig bestimmt und die Norm der Abbildung ist gleich der Norm $\|u\|$ von u . Genauer: Für $u \in H$ wird durch

$$\varphi_u(v) = \langle u, v \rangle, v \in H,$$

eine stetige lineare Abbildung φ_u mit $\|\varphi_u\| = \|u\|$ gegeben. Umgekehrt gibt es zu jeder stetigen linearen Abbildung $\varphi : H \rightarrow \mathbb{C}$ ein eindeutig durch φ bestimmtes Element u von H mit $\varphi = \varphi_u$.

Die linearen Abbildungen $\varphi : H \rightarrow \mathbb{C}$ heißen *stetige lineare Funktionale auf H* und

$$H^* := \{\varphi \mid \varphi \text{ stetiges lineares Funktional auf } H\}$$

heißt *Dualraum von H* . Der Dualraum eines Hilbertraumes kann also mit dem Hilbertraum selbst identifiziert werden.

4) Mit Hilfe des Satzes von Riesz macht man sich klar, dass es zu jedem $T \in B(H, K)$ ein (eindeutig bestimmtes) $T^* \in B(K, H)$ gibt, so dass

$$\langle v, Tw \rangle = \langle T^*v, w \rangle \quad \forall v \in K, w \in H.$$

T^* heißt *die zu T adjungierte Abbildung*.

Definition (Isometrie)

Seien \mathcal{V} und \mathcal{W} normierte Vektorräume. Eine Abbildung $U : H \rightarrow K$ heißt *Isometrie*, wenn gilt

$$\|Uv\| = \|v\| \quad \forall v \in \mathcal{V}.$$

Bemerkung: Für jede Sesquilinearform $\beta : \mathcal{V} \times \mathcal{V} \rightarrow \mathbb{C}$ auf einem komplexen Vektorraum \mathcal{V} gilt die *Polarisationsgleichung*

$$\begin{aligned} \beta(v, w) = \frac{1}{4} & (\beta(v+w, v+w) - \beta(v-w, v-w) \\ & - i\beta(v+iw, v+iw) + i\beta(v-iw, v-iw)) \end{aligned} \quad (6.1.3)$$

Angewandt auf $\beta(v, w) = \langle Uv, Uw \rangle$ erhält man, dass jede Isometrie U zwischen zwei Hilberträumen H und K *das Skalarprodukt invariant lässt*:

$$\langle Uv, Uw \rangle = \langle v, w \rangle \quad \forall v, w \in H. \quad (6.1.4)$$

Eine lineare Abbildung $U : H \rightarrow K$ ist also genau dann eine Isometrie, wenn (6.1.4) gilt. Offenbar ist jede Isometrie beschränkt mit $\|U\| = 1$.

Definition (Unitäre Abbildung)

Eine Abbildung $U : H \rightarrow K$, H, K Hilberträume, heißt *unitär*, wenn gilt:

- U ist eine Isometrie.
- U ist bijektiv.

Bemerkung: Eine Abbildung $U : H \rightarrow K$ ist also genau dann unitär, wenn sie eine Bijektion ist und die Gleichungen (6.1.4) erfüllt.

5)

Definition (Vollständiges Orthonormalsystem)

Sei H ein Hilbertraum, I eine Indexmenge. Eine Familie $(v_i)_{i \in I}$ von Vektoren $v_i \in H$ heißt *vollständiges Orthonormalsystem von H* , wenn gilt:

- $\langle v_i, v_j \rangle = \delta_{ij}$ für alle $i, j \in I$.
- Die v_i sind *total* in H , d. h. der Abschluss

$$\overline{\text{Lin}\{v_i \mid i \in I\}}$$

der linearen Hülle der v_i ist gleich H .

Bemerkung: Jeder Hilbertraum besitzt ein vollständiges Orthonormalsystem, und die Kardinalität von I ist eindeutig durch H bestimmt.

Definition Ein Hilbertraum heißt *separabel*, wenn es in H eine *abzählbare* dichte Teilmenge gibt. Dies ist genau dann der Fall, wenn für H ein abzählbares Orthonormalsystem existiert.

Bemerkung: Im Weiteren setzen wir in der Regel voraus, dass H separabel ist.

6) Ist $\{v_n \mid n \geq 1\}$ ein vollständiges Orthonormalsystem des Hilbertraumes H , so besitzt jeder Vektor v in H die *Fourier-Entwicklung*

$$v = \sum_{n \geq 1} \langle v_n, v \rangle v_n, \quad (6.1.5)$$

wobei im Falle eines abzählbar *unendlichen* vollständigen Orthonormalsystems die unendliche Summe in (6.1.5) normkonvergent ist. Es gilt für alle $v, w \in H$:

$$\|v\|^2 = \sum_{n \geq 1} |\langle v_n, v \rangle|^2$$

und

$$\langle v, w \rangle = \sum_{n \geq 1} \langle v, v_n \rangle \langle v_n, w \rangle.$$

(Die letzte Summe ist nach Cauchy-Schwarz absolut konvergent.)

Lemma 6.1.1 (Isometrie-Lemma)

Seien H und K Hilberträume, I eine Indexmenge. Gegeben seien zwei Familien

$$\begin{aligned} (v_i)_{i \in I}, v_i \in H \\ (w_i)_{i \in I}, w_i \in K \end{aligned}$$

mit

$$\langle v_i, v_j \rangle = \langle w_i, w_j \rangle \quad \forall i, j \in I. \quad (6.1.6)$$

Bezeichne mit \mathcal{V} und mit \mathcal{W} die linearen Hüllen der v_i und der w_i , d. h.

$$\begin{aligned}\mathcal{V} &= \text{Lin} \{v_i \mid i \in I\} \\ \mathcal{W} &= \text{Lin} \{w_i \mid i \in I\},\end{aligned}$$

und mit $\overline{\mathcal{V}}$ und $\overline{\mathcal{W}}$ den Abschluss von \mathcal{V} und \mathcal{W} in H bzw. K . (\mathcal{V} und \mathcal{W} sind Vektorräume mit Skalarprodukt, $\overline{\mathcal{V}}$ und $\overline{\mathcal{W}}$ sind Hilberträume!)

Dann gibt es eine eindeutig bestimmte lineare Abbildung

$$U : \mathcal{V} \rightarrow \mathcal{W}$$

mit

$$Uv_i = w_i \quad \forall i \in I.$$

Es ist U eine bijektive Isometrie.

Es gibt eine eindeutig bestimmte unitäre Abbildung

$$\overline{U} : \overline{\mathcal{V}} \rightarrow \overline{\mathcal{W}}$$

mit $\overline{U}v_i = w_i \quad \forall i \in I$.

Beweis:

Behauptung: Für $J \in \mathcal{P}_0(I)$ und $\lambda_i \in \mathbb{C}$, $i \in I$, gilt

$$\sum_{i \in J} \lambda_i v_i = 0 \Leftrightarrow \sum_{i \in J} \lambda_i w_i = 0.$$

Beweis:

$$\begin{aligned}\sum_{i \in J} \lambda_i v_i &= 0 \Leftrightarrow \\ \left\| \sum_{i \in J} \lambda_i v_i \right\|^2 &= \sum_{i,j \in J} \overline{\lambda_i} \lambda_j \langle v_i, v_j \rangle = 0 \\ &\stackrel{(6.1.6)}{=} \sum_{i,j \in J} \overline{\lambda_i} \lambda_j \langle w_i, w_j \rangle \\ &= \left\| \sum_{i \in J} \lambda_i w_i \right\|^2 \\ \Leftrightarrow \sum_{i \in J} \lambda_i w_i &= 0 \checkmark\end{aligned}$$

Nun folgt, dass durch

$$U\left(\sum_{i \in J} \lambda_i v_i\right) = \sum_{i \in J} \lambda_i w_i$$

eine (surjektive) lineare Abbildung von \mathcal{V} nach \mathcal{W} definiert wird. Natürlich gilt $Uv_i = w_i$ und U ist die einzige lineare Abbildung von \mathcal{V} nach \mathcal{W} , die v_i auf w_i abbildet. Offenbar ist U eine Isometrie und damit injektiv. Der Rest ist klar. \square

Als erste Anwendung von Lemma 6.1.1 betrachten wir Verallgemeinerungen von $L^2(\mathbb{R}_+)$ und $L^2([0, T])$.

Definition (Tensorprodukt von Hilberträumen)
Für zwei Hilberträume H_1 und H_2 wird durch

$$\langle v_1 \otimes v_2, w_1 \otimes w_2 \rangle := \langle v_1, w_1 \rangle \langle v_2, w_2 \rangle, \quad v_1, w_1 \in H_1, \quad v_2, w_2 \in H_2 \quad (6.1.7)$$

ein Skalarprodukt auf dem Tensorprodukt $H_1 \otimes H_2$ der Vektorräume H_1 und H_2 definiert (folgt mit dem Satz, dass das Schur-Produkt zweier positiv definiter Matrizen wieder positiv definit ist). Die Vervollständigung

$$H_1 \overline{\otimes} H_2 := \overline{H_1 \otimes H_2}$$

von $H_1 \otimes H_2$ bzgl. der von diesem Skalarprodukt auf $H_1 \otimes H_2$ gegebenen Norm heißt *Tensorprodukt der Hilberträume H_1 und H_2* .

Beispiel: Sei nun \mathcal{H} ein separabler Hilbertraum. Das ist z. B. für $L^2(\mathbb{R}_+)$ und für $L^2([0, T])$ der Fall, aber auch für $L^2(\Omega, \mathcal{F}, P)$, wenn $((\Omega, \mathcal{F}, P))$ den Raum der kanonischen Realisierung der Brownschen Bewegung aus Abschnitt 4.2 bezeichnet.

Betrachte den Raum

$$L^2(\mathbb{R}_+, \mathcal{H}) := \{f : \mathbb{R}_+ \rightarrow \mathcal{H} \mid t \mapsto \langle f(t), v \rangle \text{ messbar} \quad (6.1.8)$$

$$\text{für alle } v \in \mathcal{H} \text{ und } \int_0^\infty \|f(t)\|^2 dt < \infty\}. \quad (6.1.9)$$

Die Definition ist sinnvoll, weil aus der Messbarkeit von $\langle f(\cdot), v \rangle$ für alle $v \in \mathcal{H}$ mit Hilfe der Fourier-Entwicklung folgt, dass die Abbildung $t \mapsto \|f(t)\|^2$ messbar ist (hier wird die Separabilität von \mathcal{H} gebraucht). Aus ähnlichen Gründen ist die Definition

$$\langle f, g \rangle := \int_0^\infty \langle f(t), g(t) \rangle dt \quad (6.1.10)$$

sinnvoll (Cauchy-Schwarz zweimal!). Mit einfachen Argumenten zeigt man, dass $L^2(\mathbb{R}_+, \mathcal{H})$ mit dem Skalarprodukt (6.1.10) vollständig und damit ein Hilbertraum ist.

Stochastische Prozesse

SoSe 2018

Übersicht 14. Vorlesungswoche 16.7.-20.7.2018)

Auf der anderen Seite betrachte den Hilbertraum

$$L^2(\mathbb{R}_+) \overline{\otimes} \mathcal{H}.$$

Für $\alpha, \beta \in L^2(\mathbb{R}_+)$, $u, v \in \mathcal{H}$ haben wir

$$\begin{aligned} \langle \alpha \otimes u, \beta \otimes v \rangle &= \langle \alpha, \beta \rangle \langle u, v \rangle \\ &= \left(\int_0^t \overline{\alpha(t)} \beta(t) dt \right) \langle u, v \rangle \\ &= \int_0^t \langle \alpha(t)u, \beta(t)v \rangle dt \\ &= \langle \alpha(t)u, \beta(t)v \rangle, \end{aligned}$$

wobei $\alpha u, \beta v \in L^2(\mathbb{R}_+, \mathcal{H})$ mit

$$\begin{aligned} (\alpha u)(t) &= \alpha(t)u \\ (\beta v)(t) &= \beta(t)v. \end{aligned}$$

Jetzt wenden wir das Isometrie-Lemma 6.1.1 auf die Indexmenge

$$I = \{(\alpha, u) \mid \alpha \in L^2(\mathbb{R}_+), u \in \mathcal{H}\}$$

und die Familien

$$\begin{aligned} (\alpha \otimes u)_{(\alpha, u) \in I}, \alpha \otimes u &\in L^2(\mathbb{R}_+) \otimes \mathcal{H} \\ (\alpha u)_{(\alpha, u) \in I}, \alpha u &\in L^2(\mathbb{R}_+, \mathcal{H}) \end{aligned}$$

an, um eine Isometrie

$$U : L^2(\mathbb{R}_+) \overline{\otimes} \mathcal{H} \rightarrow L^2(\mathbb{R}_+, \mathcal{H})$$

zu erhalten. Für jedes $f \in L^2(\mathbb{R}_+, \mathcal{H})$ gilt die Entwicklung

$$f(t) = \sum_{n \geq 1} \langle e_n, f(t) \rangle e_n.$$

(Dabei bezeichnet $(e_n)_{n \geq 1}$ ein vollständiges Orthonormalsystem von \mathcal{H} . Es wurde ja vorausgesetzt, dass \mathcal{H} separabel ist!) Also ist $\{\alpha u \mid (\alpha, u) \in I\}$ total in $L^2(\mathbb{R}_+, \mathcal{H})$ und wir erhalten, dass U unitär ist, d. h.

$$L^2(\mathbb{R}_+) \overline{\otimes} \mathcal{H} \cong L^2(\mathbb{R}_+, \mathcal{H}).$$

6.2 Itô-Integrationstheorie für L^2 -Prozesse

Hier, zum Teil wiederholend, einige Bezeichnungen, die im Weiteren verwendet werden:

1) Wir bezeichnen mit \mathcal{F}_t , die σ -Algebra

$$\sigma(B_s \mid 0 \leq s \leq t),$$

die von einer Brownsche Bewegung B_t über (Ω, \mathcal{F}, P) erzeugt wird. Dabei nehmen wir an, dass B_t und (Ω, \mathcal{F}, P) so gewählt sind, dass $L^2(\Omega, \mathcal{F}, P)$ separabel ist. Dies ist z. B. der Fall für die Realisierung von B_t aus Abschnitt 4.2. (Beweisen wir hier nicht; siehe etwa "T. Hida: Brownian Motion".)

2)

$$\mathcal{H} = L^2(\Omega, \mathcal{F}, P) = L^2(\Omega)$$

3)

$$\begin{aligned} L^2(\mathbb{R}_+) &= L^2(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+), \lambda \upharpoonright \mathbb{R}_+) \\ L^2([0, T]) &= L^2([0, T], \mathcal{B}([0, T]), \lambda \upharpoonright [0, T]), \quad T \in \mathbb{R}_+ \end{aligned}$$

Bemerkungen:

1) Es gilt (wie oben für \mathbb{R}_+), dass

$$\begin{aligned} L^2([0, T], \mathcal{H}) &:= \{f : [0, T] \rightarrow \mathcal{H} \mid t \mapsto \langle f(t), v \rangle \text{ messbar} \\ &\text{für alle } v \in \mathcal{H} \text{ und } \int_0^T \|f(t)\|^2 dt < \infty\}. \end{aligned}$$

ein Hilbertraum ist, isomorph zu

$$L^2([0, T]) \otimes \overline{\mathcal{H}}.$$

Wir haben natürlich auch

$$L^2([0, T], \mathcal{H}) \cong \mathcal{H} \otimes L^2([0, T]).$$

Diese Identifikation ziehen wir jetzt vor, da wir Prozesse der Gestalt $X \mathbf{1}_B$ mit Zufallsvariable X aus $\mathcal{H} = L^2(\Omega)$ und $B \subset [0, T]$ betrachten wollen.

2) Man kann $L^2([0, T], \mathcal{H})$ auch mit dem Produktraum

$$L^2(\Omega \times [0, t], \mathcal{F} \otimes \mathcal{B}([0, T]), P \otimes \lambda \upharpoonright [0, T])$$

identifizieren: Es gilt für $X, Y \in \mathcal{H}, f, g \in L^2([0, T])$

$$\langle X \otimes f, Y \otimes g \rangle = \int_0^T \int_{\Omega} \overline{X(\omega) f(t)} Y(\omega) g(t) dP(\omega) dt,$$

so dass durch

$$U(X \otimes f)(\omega, t) = X(\omega)f(t)$$

eine Isometrie (weitere Anwendung des Isometrie-Lemmas 6.1.1)

$$U : L^2(\Omega) \overline{\otimes} L^2([0, T]) \rightarrow L^2(\Omega \times [0, T])$$

gegeben wird. Wieder gilt, dass

$$\{(\omega, t) \mapsto X(\omega)f(t) \mid X \in L^2(\Omega), f \in L^2([0, T])\}$$

total in $L^2(\Omega \times [0, T])$ liegen, so dass U unitär ist. Insbesondere ist ein “Prozess” $F \in \mathcal{H} \otimes L^2([0, T])$ genau dann = 0, wenn

$$F_t(\omega) = 0 \text{ für } (P \otimes \lambda)\text{-f.a. } (\omega, t).$$

Im Weiteren sei $T \geq 0$ fest gewählt.

Definition (Einfache adaptierte Prozesse)

Ein Element F aus dem (Vektorraum-)Tensorprodukt

$$\mathcal{H} \otimes L^2([0, T]) \subset \mathcal{H} \overline{\otimes} L^2([0, T])$$

heißt *einfacher adaptierter Prozess*, wenn F von der Form

$$F = \sum_{k=0}^m X_k \otimes \mathbf{1}_{[t_k, t_{k+1}[}$$

ist mit:

- $\{0 = t_0 < t_1 < \dots < t_{m+1} = T\} \in \mathbf{Par}_{0,T}$
- $X_k \in \mathcal{F}_{t_k}$

(Achtung: Letzteres ist eine etwas missbräuchliche Schreibweise für “ X_k ist \mathcal{F}_{t_k} -messbar”!)

Bezeichnung: Wir bezeichnen den Raum aller einfachen adaptierten Prozesse mit \mathcal{E}_T .

Wir kommen nun zu einem zentralen Ergebnis:

Satz 6.2.1 (Itô-Isometrie)

Gegeben seien $u, s, t, v \in \mathbb{R}_+$ mit

$$u \leq s \leq t \leq v$$

und quadrat-integrierbare Zufallsvariablen X und Y , so dass

$$X \in \mathcal{F}_u \text{ und } Y \in \mathcal{F}_s.$$

Dann gilt:

$X (B(t) - B(u))$ und $Y (B(v) - B(s))$ liegen in $L^2(\Omega)$

und

$$\langle X, Y \rangle (t - s) = \langle X (B(t) - B(u)), Y (B(v) - B(s)) \rangle$$

Beweis: Zunächst gilt wegen $u \leq s$

$$\bar{X} \in \mathcal{F}_u, Y \in \mathcal{F}_s \Rightarrow \bar{X} Y \in \mathcal{F}_s,$$

woraus zusammen mit der Unabhängigkeit der Zuwächse einer Brownschem Bewegung und weil $u \leq s \leq t \leq v$ folgt:

$$\bar{X} Y, B(t) - B(s), B(v) - B(t) \text{ unabhängig} \quad (6.2.1)$$

$$\bar{X} Y, (B(t) - B(s))^2 \text{ unabhängig}$$

$$\bar{X} Y (B(s) - B(u)) \in \mathcal{F}_s$$

$$\Rightarrow \bar{X} Y (B(s) - B(u)), B(v) - B(t) \text{ unabhängig}$$

$$\bar{X} Y (B(s) - B(u)) \in \mathcal{F}_s$$

$$\Rightarrow \bar{X} Y (B(s) - B(u)), B(t) - B(s) \text{ unabhängig}$$

Wir erhalten:

$$\begin{aligned} & \langle X (B(t) - B(u)), Y (B(v) - B(s)) \rangle \\ &= E(\bar{X} Y (B(t) - B(u)) (B(v) - B(s))) \\ &= E(\bar{X} Y (B(t) - B(s) + B(s) - B(u)) \\ & \quad (B(v) - B(t) + B(t) - B(s))) \\ &= E(\bar{X} Y (B(t) - B(s)) (B(v) - B(t))) \\ & \quad + E(\bar{X} Y (B(t) - B(s))^2) \\ & \quad + E(\bar{X} Y (B(s) - B(u)) (B(v) - B(t))) \\ & \quad + E(\bar{X} Y (B(s) - B(u)) (B(t) - B(s))) \\ &\stackrel{(6.2.1)}{=} E(\bar{X} Y) E(B(t) - B(s)) E(B(v) - B(t)) \\ & \quad + E(\bar{X} Y) E((B(t) - B(s))^2) \\ & \quad + E(\bar{X} Y (B(s) - B(u))) E(B(v) - B(t)) \\ & \quad + E(\bar{X} Y (B(s) - B(u))) E(B(t) - B(s)) \\ &= E(\bar{X} Y) E((B(t) - B(s))^2) \\ &= \langle X, Y \rangle E((B(t) - B(s))^2) \\ &= \langle X, Y \rangle E((B_{t-s})^2) \\ &= \langle X, Y \rangle (t - s). \end{aligned}$$

Die letzten beiden Gleichheiten folgen dabei aus der Stationarität der Zuwächse einer Brownschen Bewegung bzw. aus $B_{t-s} \sim N(0, t-s)$. \square

Wichtige Folgerung: Für $u, s, t, v \in \mathbb{R}_+$, $u \leq s \leq t \leq v$, und $X \in \mathcal{F}_u$, $Y \in \mathcal{F}_s$ betrachte die einfachen Prozesse

$$X \otimes \mathbf{1}_{[u,t[}, Y \otimes \mathbf{1}_{[s,v[}.$$

Dann haben wir:

$$\begin{aligned} \langle X \otimes \mathbf{1}_{[u,t[}, Y \otimes \mathbf{1}_{[s,v[} \rangle &= \langle X, Y \rangle \langle \mathbf{1}_{[u,t[}, \mathbf{1}_{[s,v[} \rangle \\ &= \langle X, Y \rangle \int_0^T \mathbf{1}_{[u,t[}(\tau) \mathbf{1}_{[s,v[}(\tau) \, d\tau \\ &= \langle X, Y \rangle \int_0^T \mathbf{1}_{[s,t[}(\tau) \, d\tau \\ &= \langle X, Y \rangle (t-s) \\ &\stackrel{(6.2.1)}{=} \langle X (B(t) - B(u)), Y (B(v) - B(s)) \rangle \end{aligned}$$

Behauptung: Die Gleichung

$$\langle X \otimes \mathbf{1}_{[u,t[}, Y \otimes \mathbf{1}_{[s,v[} \rangle = \langle X (B(t) - B(u)), Y (B(v) - B(s)) \rangle \quad (6.2.2)$$

gilt für alle $u, s, t, v \in \mathbb{R}_+$ mit $u \leq t$ und $s \leq v$.

Beweis: Sind $[u, t[$ und $[s, v[$ disjunkt, so steht auf beiden Seiten von (6.2.2) der Wert 0. Im Falle $s \leq u \leq v \leq t$ müssen wir nur (s, v) durch (u, t) ersetzen, um wieder im Fall von Satz 6.2.1 zu sein. Falls $u \leq s \leq v \leq t$ führt die Zerlegung

$$B(t) - B(u) = (B(t) - B(v)) + (B(v) - B(s)) + (B(s) - B(u))$$

zum Ziel. \checkmark

Nun wenden wir das Isometrie-lemma 6.1.1 auf die folgende Situation an:

$$\begin{aligned} H &= L^2(\Omega) \overline{\otimes} L^2([0, T]) \cong L^2([0, T], L^2(\Omega)) \\ K &= L^2(\Omega) \\ I &= \{(X, s, t) \mid 0 \leq s \leq t, X \in \mathcal{F}_s\} \\ v_{(X,s,t)} &= X \otimes \mathbf{1}_{[s,t[} \\ w_{(X,s,t)} &= X (B(t) - B(s)). \end{aligned}$$

Dann gilt, wie wir soeben gesehen haben,

$$\langle v_{(X,s,t)}, v_{(X',s',t')} \rangle = \langle w_{(X,s,t)}, w_{(X',s',t')} \rangle.$$

Wir können also durch

$$U(X \otimes \mathbf{1}_{[s,t]}) := X (B(t) - B(s))$$

eine Isometrie

$$U : \overline{\text{Lin } M} \rightarrow L^2(\Omega)$$

definieren, wobei

$$M := \{v_{(X,s,t)} \mid (X, s, t) \in I\}.$$

Bemerkung: Eine einfache Überlegung zeigt:

Behauptung: $\text{Lin } M = \mathcal{E}_T$

Beweis: Sicherlich sind einfache adaptierte Prozesse Linearkombinationen von Elementen aus M .–

Umgekehrt kann jede Linearkombination $\sum_{i=0}^n X_i \mathbf{1}_{[s_i, t_i[}$ von Elementen aus M in die Form $\sum_{j=0}^m Y_j \mathbf{1}_{[r_j, r_{j+1}[}$ gebracht werden mit

$$\{0 = r_0 < r_1 < \dots < r_m = T\} \in \mathbf{Par}_{0,T},$$

indem man zur disjunkten Zerlegung von $[0, T[$ übergeht, die durch Verfeinerung aus den Intervallen $[s_i, t_i[$, $i = 1, \dots, n$ entsteht. ✓

Wir erhalten also eine Isometrie

$$U : \mathcal{E}_T \rightarrow L^2(\Omega).$$

Definition (Stochastisches Itô-Integral für einfache adaptierte Prozesse)

Für $F \in \mathcal{E}_T$ heißt die Zufallsvariable $U(F) \in L^2(\Omega) = \mathcal{H}$ *stochastisches Integral von F* .

Schreibweise:

$$U(F) =: \int_0^T F_t dB_t = \int_0^T F dB$$

Bemerkung: Für festes $T \in \mathbb{R}_+$ und $t \in [0, T]$ können wir

$$F \in L^2([0, t], \mathcal{H})$$

das Element

$$F \upharpoonright [[0, t] : [0, t] \rightarrow \mathcal{H}$$

aus \mathcal{E}_t zuordnen: Der Wert an der Stelle t kann gleich 0 gesetzt werden, ohne dass sich die Äquivalenzklasse von $F \upharpoonright [[0, t] \in L^2([0, t], \mathcal{H})$ ändert.

Definition (Stochastisches Itô-Integral als Prozess)

Wir setzen für $F \in \mathcal{E}_T$ und für $t \in [0, T]$

$$\int_0^t F_s dB_s := \int_0^t (F \upharpoonright [0, t])_s dB_s.$$

Der Prozess

$$\left(\int_0^t F_s dB_s \right)_{t \in [0, T]} = \left(\int_0^t F dB \right)_{t \in [0, T]}$$

heißt *stochastisches Integral des Prozesses F* .

Definition (Quadrat-integrierbare adaptierte Prozesse)

Ein stochastischer Prozess $(X_t)_{t \in \mathbb{R}_+}$ heißt *quadrat-integrierbar adaptiert*, wenn die Bedingungen

· (Quadrat-Integrierbarkeit “im Ganzen”)

$$X \in L^2(\Omega \times [0, T])$$

· (Adaptiertheit)

$$X_t \in \mathcal{F}_t \quad \forall t \in [0, T]$$

erfüllt sind.

Bezeichnung: Wir bezeichnen den Raum aller quadrat-integrierbaren adaptierten Prozesse mit Θ_T .

Bemerkungen: 1) Es gilt für $F \in \mathcal{E}_T$:

$$\int_0^T F dB = \sum_{k=0}^m F(t_k) (B(t_{k+1}) - B(t_k)) \quad (6.2.3)$$

Damit ist das Integral $\int_0^T F dB$ durch eine *Riemann-Stieltjes-Summe* definiert; vgl. Bemerkungen nach Korollar 4.3.3. Dabei ist die \mathcal{F}_{t_k} -Messbarkeit von $F(t_k)$ wichtig. Sie bedeutet gerade, dass F_t adaptiert ist. Etwas anders interpretiert ist (6.2.3) eine besondere Riemann-Stieltjes-Summe, nämlich eine solche mit der speziellen *Itôschen Wahl der Zwischenpunkte* $\tau_k, t_k \leq \tau_k \leq t_{k+1}$: linke Intervallgrenze, d. h. $\tau_k = t_k$.

2) Wir dehnen unser stochastisches Integral aus auf Elemente aus

$$\overline{\mathcal{E}}_T \subset L^2(\Omega) \otimes L^2([0, T]), \quad (6.2.4)$$

indem wir die Isometrie $U : \mathcal{E}_T \rightarrow L^2(\Omega)$ in eindeutiger Weise fortsetzen zu einer Isometrie $U : \overline{\mathcal{E}}_T \rightarrow L^2(\Omega)$. Hierbei bezeichnet $\overline{\mathcal{E}}_T$ den Abschluss von \mathcal{E}_T in der Norm von $L^2(\Omega) \otimes L^2([0, T])$. Wir schreiben wieder $\int_0^T F dB$ für $U(F)$, $F \in \overline{\mathcal{E}}_T$.

Es stellt sich die Frage, welche Prozesse in $\overline{\mathcal{E}}_T$ liegen.

Behauptung: $\overline{\mathcal{E}}_T \subset \Theta_T$

Beweis: Wegen (6.2.4) muss es sich bei $F \in \overline{\mathcal{E}}_T$ jedenfalls um einen quadratintegrierbaren Prozess handeln. Des Weiteren impliziert $X_n \xrightarrow[n \rightarrow \infty]{\|\cdot\|_2} X$ für $X_n \in \mathcal{E}_T$, $X \in \overline{\mathcal{E}}_T$, die Existenz einer Teilfolge $(X_{n_k})_{k \in \mathbb{N}}$, die $(\mathbb{P} \otimes \lambda)$ -f.ü. gegen X konvergiert (Konvergenz in $L^2 \Rightarrow$ Existenz einer f.ü-konvergenten Teilfolge).
Aber:

$$X_{n_k} \xrightarrow[k \rightarrow \infty]{} X \text{ } (\mathbb{P} \otimes \lambda)\text{-f.ü.} \Rightarrow X_{n_k}(t) \xrightarrow[k \rightarrow \infty]{} X(t) \text{ } \mathbb{P}\text{-f.s. für f.a. } t$$

Da aber $X_{n_k}(t) \in \mathcal{F}_t$ für $X_n \in \mathcal{E}_T$, muss auch $X(t) \in \mathcal{F}_t$ gelten. ✓

3) Offenbar ist $(B_t)_{t \in [0, T]}$ in Θ_T :

$$\int_0^T \left(\int_{\Omega} B_t^2 d\mathbb{P} \right) dt = \int_0^T t dt = \frac{T^2}{2} < \infty$$

und $B_t \in \mathcal{F}_t$ nach Definition von \mathcal{F}_t .

Behauptung: Für

$$B^{(n)} = \sum_{k=0}^{m_n} B(t_k^{(n)}) \mathbf{1}_{[t_k^{(n)}, t_{k+1}^{(n)}[}$$

mit

$$\mathfrak{Z}_n = \{0 = t_0^{(n)} < t_1^{(n)} < \dots < t_{m_n}^{(n)} = T\}, \quad |\mathfrak{Z}_n| \xrightarrow[n \rightarrow \infty]{} 0,$$

gilt

$$B^{(n)} \xrightarrow[n \rightarrow \infty]{\|\cdot\|_2} B,$$

d. h. insbesondere $(B_t)_{t \in [0, T]} \in \overline{\mathcal{E}}_T$.

Beweis: Übung!

Nach Satz 4.3.3 folgt jetzt

$$\begin{aligned} \int_0^T B_t dB_t &= \lim_{n \rightarrow \infty} \int_0^T B_t^{(n)} dB_t & (6.2.5) \\ &= \lim_{n \rightarrow \infty} \sum_{k=0}^{m_n} B(t_k^{(n)}) (B(t_{k+1}^{(n)}) - B(t_k^{(n)})) \\ &= \frac{1}{2} B_T^2 - \frac{1}{2} T. \end{aligned}$$

Satz 6.2.2

$$\overline{\mathcal{E}}_T = \Theta_T$$

Beweis: (Skizze) Zu zeigen ist: Für $F \in \Theta_T$ existiert eine Folge $F_n, F_n \in \mathcal{E}_T$, mit $F_n \xrightarrow[n \rightarrow \infty]{\|\cdot\|_2} F$, d. h. mit

$$\int_{\Omega} \int_0^T |F_n(\omega, t) - F(\omega, t)|^2 dt dP \xrightarrow[n \rightarrow \infty]{} 0.$$

1. Schritt: Der Satz wird für alle Prozesse $F \in \Theta_T$ mit
 · F beschränkt, d. h. $|F(\omega, t)| \leq C$ für alle $\omega \in \Omega$ und für alle $t \in [0, T]$, C geeignete Konstante
 · $t \mapsto F(\omega, t)$ ist eine stetige Abbildung von $[0, T]$ nach \mathbb{C} für alle $\omega \in \Omega$
 bewiesen. Mit Hilfe des Konvergenzsatzes von Lebesgue zeigt man dazu, dass F_n mit

$$F_n(\omega, t) := \sum_{k=0}^{m_n} F(\omega, t_k^{(n)}) \mathbf{1}_{[t_k^{(n)}, t_{k+1}^{(n)}[}$$

die gewünschte Approximation liefert.

2. Schritt: Mit etwas größerem Aufwand zeigt man den Satz für beschränktes $F \in \Theta_T$ (ohne Forderungen an die Stetigkeit von F): Hierbei wird F durch *Faltung* mit geeigneten Funktionen g_n "stetig gemacht": Die Folge

$$F_n(t, \omega) := \int_0^t g_n(s-t) F(s, \omega) ds$$

approximiert dann F in $L^2(\Omega \times [0, T])$.

3. Schritt: Für beliebiges $F \in \Theta_T$ schneidet man den Prozess F mittels

$$F_n(t, \omega) := \begin{cases} -n & \text{falls } F(t, \omega) < -n \\ F(t, \omega) & \text{falls } -n \leq F(t, \omega) \leq +n \\ +n & \text{falls } F(t, \omega) > +n \end{cases}$$

ab, um eine Folge beschränkter Prozesse $F_n \in \Theta_T$ zu erhalten, die in $L^2(\Omega \times [0, T])$ gegen F konvergiert. \square

Definition (Itô-Integral für Prozesse aus Θ_T)

Nach Bemerkung 2 oben ist nun das stochastische Integral für alle quadratintegrierbaren adaptierten Prozesse, d. h. für alle Prozesse aus Θ_T definiert. Für jedes $F \in \Theta_T$ gibt es eine Folge $F_n, F_n \in \mathcal{E}_T$, so dass

$$\int_0^T F dB = \lim_{n \rightarrow \infty} \int_0^T F_n dB.$$

Wir stellen einige Eigenschaften des stochastischen Integrals zusammen.

Satz 6.2.3 Für $F, G \in \Theta_T$, $T \in \mathbb{R}_+$, gilt:

(a)

$$\int_0^T (\alpha F + G) dB = \alpha \int_0^T F dB + \int_0^T G dB \quad \forall \alpha \in \mathbb{C}$$

(b)

$$\mathbb{E}\left(\int_0^T F dB\right) = 0$$

(c)

$$\int_0^T F dB \in \mathcal{F}_T$$

(d)

$$\mathbb{E}\left(\left|\int_0^T F dB\right|^2\right) = \mathbb{E}\left(\int_0^T |F|^2 dt\right)$$

(e) Für

$$X_t := \int_0^t F dB$$

gilt $X = (X_t)_{t \in [0, T]} \in \Theta_T$.

(f) Das stochastische Integral

$$\int : \Theta_T \rightarrow \Theta_T$$

ist stetig bzgl. der L^2 -Norm auf $L^2(\Omega \times [0, T])$, d. h.

$$F_n \xrightarrow[n \rightarrow \infty]{\|\cdot\|_2} F \Rightarrow X_n \xrightarrow[n \rightarrow \infty]{\|\cdot\|_2} X,$$

wobei

$$X_n(t) = \int_0^t F_n dB; \quad X(t) = \int_0^t F dB.$$

Beweis: (a)-(d) folgen daraus, dass sie für einfache adaptierte Prozesse gelten.

(e) und (f) folgen mit Fubini und der Itô-Isometrie. \square

6.3 Die Itô-Formel

Zunächst halten wir fest:

Satz 6.3.1

$$F, G \in \Theta_T, \quad F \text{ beschränkt} \Rightarrow FG \in \Theta_T$$

Beweis: Falls F durch C beschränkt:

$$\int_{\Omega} \int_0^T |FG|^2 dt dP \leq C^2 \int_{\Omega} \int_0^T |G|^2 dt dP < \infty$$

und somit $FG \in L^2(\Omega \times [0, T])$. Außerdem:

$$F_t, G_t \in \mathcal{F}_t \Rightarrow (FG)_t \in \mathcal{F}_t \square$$

Satz 6.3.2 (Regel von der partiellen Integration für das Itô-Integral)

Seien $F, G \in \Theta_T$ beschränkt. Setze:

$$X_t = \int_0^t F dB$$

$$Y_t = \int_0^t G dB$$

Dann gilt:

$$X_t Y_t = \int_0^t X_s G_s dB_s + \int_0^t F_s Y_s dB_s \quad (6.3.1)$$

$$+ \int_0^t F_s G_s ds \quad (6.3.2)$$

Bemerkungen 1) Wegen Satz 6.2.3 (e) und wegen Satz 6.3.1 sind XG und FY in Θ_T , so dass die in (6.3.1) auftretenden Integrale jedenfalls Sinn ergeben.

2) Das Integral in (6.3.2) ist pfadweise zu verstehen.

Beweis: (Skizze) Wir nehmen zunächst an, dass

$$F = C \mathbf{1}_{[s,t]}$$

$$G = D \mathbf{1}_{[s,t]}$$

mit $0 \leq s \leq t \leq T$ und beschränkten Zufallsvariablen $C, D \in \mathcal{F}_s$. Dann gilt:

$$X_r = \begin{cases} 0 & \text{falls } r \leq s \\ C(B_r - B_s) & \text{falls } s \leq r \leq t \\ C(B_t - B_s) & \text{falls } t \leq r \leq T \end{cases}$$

$$Y_r = \begin{cases} 0 & \text{falls } r \leq s \\ D(B_r - B_s) & \text{falls } s \leq r \leq t \\ D(B_t - B_s) & \text{falls } t \leq r \leq T \end{cases}$$

$$X_r Y_r = \begin{cases} 0 & \text{falls } r \leq s \\ CD(B_r - B_s)^2 & \text{falls } s \leq r \leq t \\ CD(B_t - B_s)^2 & \text{falls } t \leq r \leq T \end{cases}$$

Weiterhin:

$$\begin{aligned} X_r G_r &= CD(B_r - B_s) \mathbf{1}_{[s,t]}(r) \\ F_r Y_r &= DC(B_r - B_s) \mathbf{1}_{[s,t]}(r) \\ F_r G_r &= CD \mathbf{1}_{[s,t]}(r) \end{aligned}$$

Wir wissen schon (für $S \geq t$, was wir der Einfachheit halber annehmen):

$$\begin{aligned} \int_0^S (B_r - B_s) \mathbf{1}_{[s,t]}(r) dB_r &= \int_0^t (B_r - B_s) dB_r - \int_0^s (B_r - B_s) dB_r \\ &\stackrel{6.2.5}{=} \frac{1}{2} B_t^2 - \frac{1}{2} t - B_s B_t - \frac{1}{2} B_s^2 + \frac{1}{2} s + B_s^2 \\ &= \frac{1}{2} B_t^2 + \frac{1}{2} B_s^2 - \frac{1}{2} (t - s) - B_s B_t \end{aligned}$$

Also:

$$\begin{aligned} &\int_0^S X_r G_r dB_r + \int_0^S F_r Y_r dB_r + \int_0^S F_r G_r dr \\ &= 2CD \left(\frac{1}{2} B_t^2 + \frac{1}{2} B_s^2 - \frac{1}{2} (t - s) - B_s B_t \right) + CD(t - s) \\ &= CD (B_t^2 + B_s^2 - 2 B_s B_t) \\ &= CD (B_t - B_s)^2 \\ &= X_S Y_S \end{aligned}$$

Ähnlich beweist man die Formel für

$$\begin{aligned} F &= C \mathbf{1}_{[s_1, t_1]} \\ G &= D \mathbf{1}_{[s_2, t_2]} \end{aligned}$$

und damit für $F, G \in \mathcal{E}_T$.

Nun folgen Approximationsargumente, die wir aus zeitlichen Gründen hier nicht aufführen ... \square

Bemerkungen: 1) Wir starten etwas allgemeiner mit den Integralen

$$\begin{aligned} X_t &= X_0 + \int_0^t F_s^{(1)} dB_s + \int_0^t F_s^{(2)} ds \\ Y_t &= Y_0 + \int_0^t G_s^{(1)} dB_s + \int_0^t G_s^{(2)} ds \end{aligned}$$

wobei z. B. $F^{(2)}, G^{(2)} \in \Theta_T$ (und $F^{(1)}, G^{(1)} \in \Theta_T$ wieder beschränkt). Dann sind $\int_0^t F_s^{(2)} ds$ und $\int_0^t G_s^{(2)} ds$ wieder pfadweise zu verstehen, was geht, da z. B. $t \mapsto F_t(\omega)$ eine $L^2([0, T])$ - und damit $L^1([0, T])$ -Funktion ist für f.a. ω . Dann gilt:

$$\begin{aligned} X_t Y_t &= X_0 Y_0 + \int_0^t X_s G_s^{(1)} dB_s + \int_0^t X_s G_s^{(2)} ds \\ &\quad + \int_0^t F_s^{(1)} Y_s dB_s + \int_0^t F_s^{(2)} Y_s ds \\ &\quad + \int_0^t F_s^{(1)} G_s^{(1)} ds \end{aligned}$$

In (formaler) *differentieller Schreibweise* erhalten wir etwas “griffiger”:

$$\begin{aligned} dX_t &= F_t^{(1)} dB_t + F_t^{(2)} dt \\ dY_t &= G_t^{(1)} dB_t + G_t^{(2)} dt \\ d(X_t Y_t) &= X_t G_t^{(1)} dB_t + X_t G_t^{(2)} dt \\ &\quad + F_t^{(1)} Y_t dB_t + F_t^{(2)} Y_t dt \\ &\quad + F_t^{(1)} G_t^{(1)} dt \end{aligned} \tag{6.3.3}$$

Wir können (6.3.3) zu

$$d(X_t Y_t) = X_t dY_t + Y_t dX_t + (dX_t)(dY_t), \tag{Itô}$$

zusammenfassen, wobei der Itô-Term $(dX_t)(dY_t)$ gemäß der Tabelle

		dB _t		dt
dB _t		dt		0
dt		0		0

zu berechnen ist.

2) Unter der *Itô-Formel*, die für Spezialfälle – z. B. wenn f ein Polynom ist – aus der Formel (Itô) abgeleitet werden kann, versteht man die Formel

$$f(X_t) - f(X_0) = \int_0^t f'(X_s) dX_s + \frac{1}{2} \int_0^t f''(X_s) (dX_s)^2$$

für $f : \mathbb{R} \rightarrow \mathbb{R}$ zweimal stetig differenzierbar und

$$X_t = X_0 + \int_0^t F_s^{(1)} dB_s + \int_0^t F_s^{(2)} ds.$$

Für eine wesentlich allgemeinere und weiterführende Theorie der stochastischen Integration (“stochastische Analysis”) siehe z. B. die Bücher “Thomas Deck: Der Itô-Kalkül” oder “Philip Protter: Stochastic Integration and Differential Equations”.