Stochastische Differentialgleichungen

Rüdiger W. Braun

Wintersemester 2024/25

Inhaltsverzeichnis

1 Wahrscheinlichkeitstheorie

Die Vorlesung orientiert sich an dem Buch [Eva13] von C. L. Evans. Ein detaillierterer einführender Text ist das Buch [Øks03] von B. Øksendal.

Bezeichnung. 0 ist keine natürliche Zahl. Für ein Ereignis A in einem Ereignisraum Ω bezeichne ich das Komplement mit $A^c = \Omega \setminus A$.

In diesem Abschnitt wird ein Abriss der für die Vorlesung wichtigen Begriffe der Wahrscheinlichkeitstheorie gegeben. Die dafür benötigten Grundlagen der Maßtheorie übernehmen wir aus der Analysis III.

1.1 Definition. Ein *Wahrscheinlichkeitsraum* ist ein Maßraum $(\Omega, \mathcal{U}, P)$ mit $P(\Omega) = 1$. In diesem Fall bezeichnet man das Maß P als Wahrscheinlichkeitsmaß.

Die $\mathcal{U}$-messbaren Mengen bezeichnet man als *Ereignisse*. Anstelle der Formulierung "fast überall" benutzt man *"fast sicher"*.

1.2 Beispiel. Für einen topologischen Raum X bezeichnen wir mit $\mathcal{B}(X)$ die σ-Algebra der Borelmengen. Sei $\Omega \in \mathcal{B}(\mathbb{R}^n)$ und sei $f\colon \Omega \to [0, \infty[$ Lebesgue-messbar mit $\int f \, d\lambda_n = 1$. Dann wird durch

$$P(E) = \int_E f \, d\lambda_n$$

ein Wahrscheinlichkeitsmaß auf Ω gegeben. Wenn P von dieser Form ist, dann bezeichnet man f als *Dichte* von P.

Wenn f die Dichte des Maßes P ist, dann gilt für alle Lebesgue-messbaren Funktionen g

$$\int g \, dP = \int fg \, d\lambda_n.$$

1.3 Definition. Der $\mathbb{R}^n$ wird immer mit der σ-Algebra der Borelmengen versehen.

Eine *Zufallsvariable* ist eine messbare Abbildung $f\colon \Omega \to \mathbb{R}^n$.

1.4 Lemma. *Sei $X\colon \Omega \to \mathbb{R}^n$ eine Zufallsvariable. Dann ist*

$$\mathcal{U}(X) = \left\{ X^{-1}(B) \,\middle|\, B \in \mathcal{B}(\mathbb{R}^n) \right\}$$

eine σ-Algebra. Es handelt sich um die kleinste σ-Algebra, bzgl. derer X messbar ist. Man bezeichnet sie als die von X erzeugte *σ-Algebra.*

1.5 Lemma. *Seien* $X \colon \Omega \to \mathbb{R}^n$ *und* $Y \colon \Omega \to \mathbb{R}^m$ *zwei Zufallsvariablen. Wenn* Y *messbar bzgl. der* σ*-Algebra* $\mathcal{U}(X)$ *ist, dann existiert eine Abbildung* $\Phi \colon \mathbb{R}^n \to \mathbb{R}^m$*, so dass* $Y = \Phi \circ X$*.*

1.6 Definition. Ein *stochastischer Prozess* ist eine Familie $(X(t))_{t \geq 0}$ von Zufallsvariablen. Für festes $\omega \in \Omega$ ist $X(t, \omega)_{t \geq 0}$ ein *Pfad* des Prozesses.

Ein stochastischer Prozesses $(Y(t))_{t \geq 0}$ ist eine *Version* von $(X(t))_{t \geq 0}$, wenn für alle $t \geq 0$ die Gleichheit $X(t) = Y(t)$ fast sicher gilt.

1.7 Bezeichnung. (a) Für eine vektorwertige Zufallsvariable $X = (X^1, \dots, X^n) \colon \Omega \to \mathbb{R}^n$ mit integrierbaren Komponentenfunktionen definieren wir das Integral komponentenweise, also

$$\int X \, dP = \left(\int X^1 \, dP, \dots, \int X^n \, dP \right).$$

In [Eva13] wird der Index oben notiert, um die untere Position für partielle Ableitungen nutzen zu können.

(b) Eine Zufallsvariable X hat k*-te Momente,* wenn $|X|^j$ für $j = 0, \dots, k$ integrierbar ist. Mit $|\cdot|$ wird die euklidische Norm bezeichnet.

1.8 Definition. (a) Für eine integrierbare Zufallsvariable X bezeichnet

$$\mathbb{E}(X) = \int X \, dP$$

den *Erwartungswert.*

(b) Wenn die Zufallsvariable X zweite Momente hat, dann bezeichnet

$$\mathbb{V}(X) = \int |X - \mathbb{E}(X)|^2 \, dP$$

ihre *Varianz.*

Bemerkung. $\mathbb{V}(X) = \mathbb{E}(|X|^2) - |\mathbb{E}(X)|^2$.

1.9 Bezeichnung. Für $x, y \in \mathbb{R}^n$ schreiben wir $x \leq y$, wenn $x_j \leq y_j$ für $j = 1, \dots, n$.

1.10 Definition. Es sei $X \colon \Omega \to \mathbb{R}^n$ eine Zufallsvariable.

(a) Ihre *Verteilung* ist das durch

$$M \mapsto P(X \in M)$$

gegebene Wahrscheinlichkeitsmaß.

(b) Ihre *Verteilungsfunktion* ist gegeben durch

$$F_X\colon \mathbb{R}^n \to [0,1], \quad F_X(x) = P(X \le x).$$

Wenn n Zufallsvariablen $X_1, \dots, X_n$ gegeben sind, dann kombiniert man sie zu einer $\mathbb{R}^n$-wertigen Zufallsvariablen X und setzt $F_{X_1,\dots,X_n} = F_X$.

(c) Wenn es eine messbare Funktion f gibt, so dass für alle $A \in \mathcal{B}(\mathbb{R}^n)$

$$P(X \in A) = \int_A f \, d\lambda_n,$$

dann bezeichnet man f als *Verteilungsdichte* von X.

1.11 Beispiel. (a) Wenn $X\colon \Omega \to \mathbb{R}$ die Dichte

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}}$$

besitzt, dann sagt man, X sei $N(m, \sigma^2)$ verteilt.

(b) Wenn $C \in \mathbb{R}^{n\times n}$ invertierbar und symmetrisch ist und $X\colon \Omega \to \mathbb{R}^n$ die Dichte

$$f(x) = \frac{1}{\sqrt{(2\pi)^n \det C}} e^{-\frac{1}{2}(x-m)C^{-1}(x-m)}$$

besitzt, dann sagt man, X sei $N(m, C)$ verteilt.

Der Satz von Radon-Nikodym (s. [MV11], Theorem 13.12) gibt Auskunft über die Existenz von Dichten.

1.12 Satz. *Die Zufallsvariable* $X\colon \Omega \to \mathbb{R}^n$ *besitze die Dichte* f. *Ferner sei* $g\colon \mathbb{R}^n \to \mathbb{R}$ *eine messbare Funktion, für welche* $g \circ X$ *integrierbar ist. Dann gilt*

$$\mathbb{E}(g \circ X) = \int_{\mathbb{R}^n} gf \, d\lambda_n.$$

1.13 Korollar. *Wenn* X *die Dichte* f *besitzt und integrierbar ist bzw. zweite Momente hat, dann*

$$\mathbb{E}(X) = \int x f(x) \, d\lambda_n(x) \quad \textit{und} \quad \mathbb{V}(X) = \int |x - \mathbb{E}(X)|^2 f(x) \, d\lambda_n(x).$$

1.14 Beispiel. Eine $N(m, \sigma^2)$-verteilte Zufallsvariable hat Erwartungswert m und Varianz σ^2.

1.15 Definition. Sei $(\Omega, \mathcal{U}, P)$ ein Wahrscheinlichkeitsraum.

(a) Ereignisse $A_1, \dots, A_n$ heißen *unabhängig,* wenn für jede Wahl von $k_1 < k_2 < \cdots < k_m \le n$ gilt

$$P(A_{k_1} \cap \cdots \cap A_{k_m}) = P(A_{k_1}) \cdots P(A_{k_m}).$$

(b) Sei $(\mathcal{U}_k)_{k\in\mathbb{N}}$ eine Folge von σ-Algebren $\mathcal{U}_k \subseteq \mathcal{U}$. Sie heißt *unabhängig,* wenn für jede Wahl $k_1 < k_2 < \cdots < k_m$ und $A_i \in \mathcal{U}_{k_i}$ die Ereignisse $A_1, A_2, \dots, A_m$ unabhängig sind.

(c) Zufallsvariable $X_1, X_2, \dots$ heißen *unabhängig,* wenn die Folge $(\mathcal{U}(X_k))_{k\in\mathbb{N}}$ ihrer σ-Algebren unabhängig ist.

(d) Analog definiert man die Unabhängigkeit einer Zufallsvariablen von einer σ-Algebra und ähnliches.

1.16 Beispiel. Sei $\Omega = [0,1[$, versehen mit dem Lebesguemaß. Für $n \in \mathbb{N}$ definieren wir die n*-te Rademacher-Funktion* durch

$$X_n(\omega) = \begin{cases} 1, & \frac{k}{2^n} \leq \omega < \frac{k+1}{2^n} \text{ für ein gerades } k, \\ -1, & \frac{k}{2^n} \leq \omega < \frac{k+1}{2^n} \text{ für ein ungerades } k. \end{cases}$$

Die Rademacher-Funktionen sind unabhängig. Es gilt nämlich für jedes $x \in \{-1,1\}^k$, dass

$$P(X_1 = x_1, \dots, X_k = x_k) = 2^{-k},$$
$$P(X_1 = x_1) \cdots P(X_k = x_k) = 2^{-k}.$$

1.17 Satz. *Die* $\mathbb{R}^n$*-wertigen Zufallsvariablen* $X_1, \dots, X_k$ *sind genau dann unabhängig, wenn*

$$F_{X_1,\dots,X_k}(x_1, \dots, x_k) = F_{X_1}(x_1) \cdots F_{X_k}(x_k) \quad \textit{für alle } x.$$

Wenn alle X^i *Verteilungsdichten besitzen, dann gilt auch*

$$f_{X_1,\dots,X_k}(x_1, \dots, x_k) = f_{X_1}(x_1) \cdots f_{X_k}(x_k) \textit{ f. s.}$$

1.18 Bezeichnung. Es sei $(\Omega, \mathcal{U}, P)$ ein Wahrscheinlichkeitsraum und es sei $(A_n)_{n\in\mathbb{N}}$ eine Folge in $\mathcal{U}$. Das Ereignis

$$\limsup_{n\to\infty} A_n := \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m$$

bezeichnet man als "A_n unendlich oft".

Bemerkung. (a) Da die σ-Algebra keine Toplogie trägt, ist der Limes superior rein formal zu verstehen.

(b) ω liegt genau dann in dem Ereignis "A_n unendlich oft", wenn es unendlich viele n mit $\omega \in A_n$ gibt.

1.19 Theorem (Borel-Cantelli-Lemma). *Wenn* $\sum_{n=1}^{\infty} P(A_n) < \infty$*, dann*

$$P\left(\limsup_{n\to\infty} A_n\right) = 0.$$

1.20 Definition. Sei Q ein Wahrscheinlichkeitsmaß auf dem $\mathbb{R}^n$. Die *Fouriertransformierte* von Q ist gegeben durch

$$\hat{Q}(u) = \int_{\mathbb{R}^n} e^{i\langle u,x\rangle}\, dQ(x),$$

wobei $\langle u, x\rangle = \sum_{j=1}^n u_j x_j$.

1.21 Bemerkung. Das Wahrscheinlichkeitsmaß Q induziert via

$$\langle T_Q, \varphi\rangle = \int_{\mathbb{R}^n} \varphi(x)\, dQ(x), \quad \varphi \in \mathscr{S}(\mathbb{R}^n),$$

eine temperierte Distribution. Die Funktion $\hat{Q}$ entspricht bis auf der inversen Fouriertransformierten dieser temperierten Distribution. Da die Fouriertransformation ein Automorphismus des Raums der temperierten Distributionen ist, gilt der folgende Satz.

1.22 Satz. *Wenn zwei Wahrscheinlichkeitsmaße dieselbe Fouriertransformierte besitzen, dann stimmen sie überein.*

Beweis. Einen Beweis, der nicht auf Distributionstheorie zurückgreift — aber den Satz von Stone Weierstraß verwendet —, findet man in Klenke [Kle20], Satz 15.9. □

1.23 Definition. Sei $X\colon \Omega \to \mathbb{R}^n$ eine Zufallsvariable. Die Fouriertransformierte ihrer Verteilung bezeichnet man als *charakteristische Funktion* der Zufallsvariablen. Man bezeichnet die charakteristische Funktion mit φ_X.

1.24 Lemma. *Es sei* X *eine Zufallsvariable auf dem Wahrscheinlichkeitsraum* $(\Omega, \mathcal{F}, P)$. *Dann*

$$\varphi_X(u) = \mathbb{E}\big(e^{i\langle u,X\rangle}\big)\,.$$

1.25 Korollar. *Seien* $X_1, \dots, X_m$ *unabhängige Zufallsvariablen. Dann*

$$\varphi_{X_1+\dots+X_m}(u) = \varphi_{X_1}(u)\cdots\varphi_{X_m}(u), \quad u \in \mathbb{R}^n.$$

1.26 Satz. *Die Zufallsvariable* X *ist genau dann* $N(\mu, \sigma^2)$*-verteilt, wenn*

$$\varphi_X(u) = e^{i\mu u - \frac{u^2\sigma^2}{2}}, \quad u \in \mathbb{R}.$$

2 Bedingte Erwartung

2.1 Definition. Seien A, B Ereignisse, wobei $P(B) \neq 0$. Dann ist

$$P(A|B) = \frac{P(A \cap B}{P(B)}$$

die bedingte Wahrscheinlichkeit von A unter der Hypothese B.

2.2 Theorem. *Sei $(\Omega, \mathcal{U}, P)$ ein Wahrscheinlichkeitsraum und sei $\mathcal{V} \subseteq \mathcal{U}$ eine σ-Algebra auf Ω. Sei ferner $X\colon \Omega \to \mathbb{R}^n$ eine integrierbare Zufallsvariable. Dann existiert eine bis auf Nullmengen eindeutig bestimmte, $\mathcal{V}$-messbare Zufallsvariable Z, so dass*

$$\int_A X \, dP = \int_A Z \, dP \quad \textit{für alle } A \in \mathcal{V}. \tag{2.1}$$

2.3 Beispiel. $\mathcal{V} = \{\emptyset, \Omega\}$ ist eine σ-Algebra auf Ω. Wenn die Zufallsvariable $Z\colon \Omega \to \mathbb{R}$ $\mathcal{V}$-messbar sein soll, dann muss sie konstant sein. Damit (2.1) gilt, muss diese Konstante gleich $\mathbb{E}(X)$ sein. Also $Z(\omega) = \mathbb{E}(X)$ für alle ω.

2.4 Definition. Sei $(\Omega, \mathcal{U}, P)$ ein Wahrscheinlichkeitsraum, sei X eine Zufallsvariablen auf Ω und sei $\mathcal{V} \subseteq \mathcal{U}$ eine σ-Algebra. Die *bedingte Erwartung* $\mathbb{E}(X|\mathcal{V})$ von X gegeben $\mathcal{V}$ ist die durch folgende Eigenschaften bis auf Nullmengen eindeutig bestimmte Zufallsvariable

(a) $\mathbb{E}(X|\mathcal{V})$ ist $\mathcal{V}$-messbar.

(b) $\int_A X \, dP = \int_A \mathbb{E}(X|\mathcal{V}) \, dP$ für alle $A \in \mathcal{V}$.

Für eine Zufallsvariable Y setzen wir $\mathbb{E}(X|Y) = \mathbb{E}(X|\mathcal{U}(Y))$.

2.5 Bemerkung. (a) Theorem 2.2 zeigt, dass die bedingte Erwartung existiert.

(b) Für $A, B \in \mathcal{U}$ mit $P(B) \neq 0$ und $P(B^c) \neq 0$ gilt

$$\mathbb{E}(\chi_A|\chi_B) = P(A|B)\chi_B + P(A|B^c)\chi_{B^c}.$$

Beweis auf Blatt 2.

2.6 Satz. *Sei $(\Omega, \mathcal{U}, P)$ ein Wahrscheinlichkeitsraum mit Zufallsvariablen X und Y und sei $\mathcal{V} \subseteq \mathcal{U}$ eine σ-Algebra.*

(a) *Für* $a, b \in \mathbb{R}$ *gilt* $\mathbb{E}(aX + bY|\mathcal{V}) = a\mathbb{E}(X|\mathcal{V}) + b\mathbb{E}(Y, \mathcal{V})$.

(b) *Wenn* X $\mathcal{V}$*-messbar ist, dann* $\mathbb{E}(X, \mathcal{V}) = X$ *f. s.*

(c) *Sei* X $\mathcal{V}$*-messbar und sei* XY *integrierbar, dann* $\mathbb{E}(XY|\mathcal{V}) = X\mathbb{E}(Y|\mathcal{V})$.

(d) *Wenn* X *unabhängig von* $\mathcal{V}$ *ist, dann* $\mathbb{E}(X|\mathcal{V}) = \mathbb{E}(X)$ *f. s.*

(e) *(Turmeigenschaft) Wenn* $\mathcal{W} \subseteq \mathcal{V}$ *eine weitere* σ*-Algebra ist, dann gilt*

$$\mathbb{E}(X|\mathcal{W}) = \mathbb{E}(\mathbb{E}(X|\mathcal{V})|\mathcal{W}) = \mathbb{E}(\mathbb{E}(X|\mathcal{W})|\mathcal{V}).$$

(f) *Wenn* $X \leq Y$*, dann* $\mathbb{E}(X|\mathcal{V}) \leq \mathbb{E}(Y|\mathcal{V})$ *f. s.*

3 Der Wienersche Prozess

Wir wiederholen den Maßfortsetzungssatz aus der Analysis III.

3.1 Definition. Seien X eine Menge und $\mathcal{R} \subseteq \mathcal{P}(X)$. $\mathcal{R}$ ist ein *Ring von Teilmengen* von X, wenn

(a) $\emptyset \in \mathcal{R}$.

(b) Sind $A, B \in \mathcal{R}$, so ist $A \cup B \in \mathcal{R}$.

(c) Sind $A, B \in \mathcal{R}$, so ist $A \setminus B \in \mathcal{R}$.

3.2 Definition. Seien X eine Menge und $\mathcal{R}$ ein Ring von Teilmengen von X. Eine Abbildung $\mu\colon \mathcal{R} \to \mathbb{R} \cup \{\infty\}$ ist ein *Prämaß,* wenn

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

(b) $\mu(A) \geq 0$ für alle $A \in \mathcal{R}$.

(iii′) (σ-Additivität) Sind $A_1, A_2, \dots \in \mathcal{R}$ paarweise disjunkt und ist $\bigcup_{m=1}^{\infty} A_m \in \mathcal{R}$, so gilt

$$\mu\Bigl(\bigcup_{m=1}^{\infty} A_m\Bigr) = \sum_{m=1}^{\infty} \mu(A_m).$$

3.3 Theorem (Maßfortsetzungssatz von Carathéodory). *Seien X eine Menge, $\mathcal{R}$ ein Ring von Teilmengen von X und μ ein Prämaß auf $\mathcal{R}$. Dann kann μ zu einem Maß fortgesetzt werden.*

3.4 Bezeichnung. Sei $T \subseteq [0, \infty[$ ein abgeschlossenes Intervall oder $T = \mathbb{N}_0$. Für jede Wahl von $k \in \mathbb{N}$ und $t_1, t_2, \dots, t_k \in T$ sei ein Wahrscheinlichkeitsmaß $\nu_{t_1,\dots,t_k}$ auf dem $\mathbb{R}^{nk}$ gegeben. Die Familie

$$\{\nu_{t_1,\dots,t_k} \mid k \in \mathbb{N},\ t_1, t_2, \dots, t_k \in T\}$$

heißt *konsistent,* wenn

(a) Für jedes $k \in \mathbb{N}$, jede Permutation σ von $\{1, \dots, k\}$, jede Wahl von $t_1, \dots, t_k \in T$ und Borelmengen $F_1, \dots, F_k \subseteq \mathbb{R}^n$ gilt

$$\nu_{t_{\sigma(1)},\dots,t_{\sigma(k)}}(F_1 \times \dots \times F_k) = \nu_{t_1,\dots,t_k}(F_{\sigma^{-1}(1)} \times \dots \times F_{\sigma^{-1}(k)}),$$

(b) Für $k, m \in \mathbb{N}$, jede Wahl von $t_1, \dots, t_{k+m} \in T$ und Borelmengen $F_1, \dots, F_k \subseteq \mathbb{R}^n$ gilt

$$\nu_{t_1,\dots,t_k}(F_1 \times \cdots \times F_k) = \nu_{t_1,\dots,t_{k+m}}(F_1 \times \cdots \times F_k \times \mathbb{R}^n \times \cdots \times \mathbb{R}^n).$$

3.5 Theorem (Fortsetzungssatz von Kolmogorow). *Es sei*

$$\{\nu_{t_1,\dots,t_k} \mid k \in \mathbb{N},\ t_1, t_2, \dots, t_k \in T\}$$

eine konsistente Familie von Wahrscheinlichkeitsmaßen. Dann existieren ein Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ *und ein stochastischer Prozess* $(X_t)_{t \in T}$ *auf* Ω*, so dass*

$$\nu_{t_1,\dots,t_k}(F_1 \times \cdots \times F_k) = P(X(t_1) \in F_1, \dots, X(t_k) \in F_k)\,,$$

für alle Wahlen von $k \in \mathbb{N}$*,* $t_1, \dots, t_k \in T$ *und Borelmengen* $F_1, \dots, F_k$*.*

Der Satz kann auf den Maßfortetzungssatz von Carathéodory zurückgeführt werden. Details findet man in Aliprantis und Border [AB99], Theorem 14.23.

3.6 Korollar. *Es gibt einen Wahrscheinlichkeitsraum* $(\Omega, \mathcal{F}, P)$*, der eine Folge* $(X^n)_{n \in \mathbb{N}}$ *von unabhängigen,* $N(0,1)$*-verteilten Zufallsvariablen besitzt.*

3.7 Definition. Ein $\mathbb{R}$-wertiger stochastischer Prozess $(W(t))_{t \geq 0}$ heißt (eindimensionale) *Brownsche Bewegung* oder *Wienerscher Prozess,* wenn

(a) $W(0) = 0$ f. s.

(b) Für $t \geq s \geq 0$ ist $W(t) - W(s)$ $N(0, t-s)$-verteilt.

(c) Für $0 < t_1 < \cdots < t_n$ sind die Zufallsvariablen $W(t_1)$, $W(t_2) - W(t_1)$, …, $W(t_n) - W(t_{n-1})$ unabhängig.

(d) Die Pfade sind fast sicher stetig.

3.8 Bezeichnung. Wir setzen

$$g(x, t|y) = \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{(x-y)^2}{2t}\right).$$

3.9 Theorem. *Sei* $(W(t))_{t \geq 0}$ *eine Brownsche Bewegung, sei* $n \in \mathbb{N}$*, seien* $0 = t_0 < t_1 < \cdots < t_n$ *und sei* $f\colon \mathbb{R}^n \to \mathbb{R}$ *Borel-messbar. Dann gilt*

$$\begin{aligned}&\mathbb{E}(f(W(t_1), \dots, W(t_n))\\ &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, \dots, x_n) g(x_1, t_1|0) g(x_2, t_2 - t_1|x_1) \dots g(x_n, t_n - t_{n-1}|x_{n-1})\, dx_n \dots dx_1.\end{aligned}$$

3.10 Bezeichnung. Die *Haar-Funktionen* sind definiert durch

$$
\begin{aligned}
h_0(t) &= 1 \quad 0 \le t \le 1,\\
h_1(t) &= \begin{cases} 1, & 0 \le t \le \frac{1}{2},\\ -1, & \frac{1}{2} < t \le 1, \end{cases}\\
h_k(t) &= \begin{cases} 2^{n/2}, & \frac{k-2^n}{2^n} \le t \le \frac{k-2^n+\frac{1}{2}}{2^n},\\ -2^{n/2}, & \frac{k-2^n+\frac{1}{2}}{2^n} < t \le \frac{k-2^n+1}{2^n},\\ 0, & \text{sonst}, \end{cases}
\end{aligned}
$$

wobei $2^n \le k < 2^{n+1}$.

3.11 Satz (Haar (1909)). *Die Haar-Funktionen bilden eine Orthonormalbasis des* $L^2[0,1]$.

Der Beweis orientiert sich an dem von Theorem 1.4 des Buchs von Wojtaszczyk [Woj97].

3.12 Bezeichnung. Für $k \in \mathbb{N}_0$ setzen wir

$$s_k(x) = \int_0^x h_k(t)\,dt$$

und $s_{-1} \equiv 1$. Die Folge $(s_k)_{k \in \mathbb{N}_0 \cup \{-1\}}$ ist das *Faber-Schauder-System.*

Faber konstruierte es 1910 und zeigte, dass es eine Schauderbasis für $C[0,1]$ ist. Schauder verallgemeinerte später diese Konstruktion.

3.13 Lemma. *(a) Für* $n \in \mathbb{N}_0$ *und* $2^n \le k < 2^{n+1}$ *gilt* $\|s_k\|_\infty = 2^{-\frac{n}{2}-1}$.

(b) Für $s, t \in [0,1]$ *gilt*

$$\sum_{k=0}^{\infty} s_k(s) s_k(t) = \min(s,t).$$

3.14 Definition. Sei $0 < \gamma \le 1$, sei $X \subset \mathbb{R}^n$ komkakt und sei $f\colon X \to \mathbb{K}$ eine stetige Funktion.

(a) f ist *Hölder-stetig* zum Exponenten γ in $x \in X$, falls es eine Konstante C gibt, so dass

$$|f(x) - f(y)| \le C|x-y|^\gamma \text{ für alle } y \in X.$$

(b) f erfüllt eine *Hölder-Bedingung* zum Exponenten γ, wenn es ein C gibt, so dass

$$|f(x) - f(y)| \le C|x-y|^\gamma$$

für alle $x, y \in X$.

(c) Der Raum aller Funktionen, welche eine Hölder-Bedingung zum Exponenten γ erfüllen, wird mit $C^{0,\gamma}(X)$ bezeichnet und mit der Norm

$$\|f\|_\gamma = \|f\|_\infty + \sup_{x,y\in X, x\neq y} \frac{|f(x)-f(y)|}{|x-y|^\gamma}$$

versehen.

3.15 Bemerkung. (a) Die Hölder-Bedingung zum Exponenten 1 wird üblicherweise als Lipschitz-Bedingung bezeichnet.

(b) Der Raum $C^{0,\gamma}(X)$ ist ein Banachraum. Das wird als Aufgabe 2.15 in Kaballo [Kab11] gezeigt.

3.16 Lemma. *Sei* $0 < \delta < \frac{1}{2}$. *Wenn* $|a_k| \leq Ck^{\frac{1}{2}-\delta-\epsilon}$ *für* $C > 0$ *und* $\epsilon > 0$ *geeignet, dann konvergiert die Reihe* $\sum_{k=0}^\infty a_k s_k$ *in* $C^{0,\delta}[0,1]$.

3.17 Lemma. *Sei* $(A_k)_{k\in\mathbb{N}}$ *eine Folge unabhängiger,* $N(0,1)$*-verteilter Zufallsvariablen. Dann existiert eine Nullmenge* N*, so dass es zu jedem* $\omega \notin N$ *ein* $C > 0$ *gibt, so dass für alle* k

$$|A_k(\omega)| \leq C\sqrt{\log(k+1)}$$

3.18 Satz. *Es sei* $(A_k)_{k\in\mathbb{N}}$ *eine Folge unabhängiger,* $N(0,1)$*-verteilter Zufallsvariablen. Für jedes* $\delta < \frac{1}{2}$ *konvergiert die Reihe*

$$W(t) = \sum_{k=0}^\infty A_k s_k(t), \quad 0 \leq t \leq 1,$$

fast sicher in $C^{0,\delta}[0,1]$. *Ferner gilt:*

(a) $(W(t))_{t\geq 0}$ *ist eine Brownsche Bewegung.*

(b) Für fast alle ω *ist der Pfad* $t \mapsto W(t,\omega)$ *Hölder-stetig zum Exponenten* δ*, insbesondere stetig.*

3.19 Satz. *Seien* $X_1,\dots,X_{n+m}$ *unabhängige,* $\mathbb{R}^k$*-wertige Zufallsvariablen und seien* $f\colon (\mathbb{R}^k)^n \to \mathbb{R}$ *und* $g\colon (\mathbb{R}^k)^m \to \mathbb{R}$ *messbar. Dann sind die Zufallsvariablen*

$$Y := f(X_1,\dots,X_n) \quad \text{und} \quad Z := g(X_{n+1},\dots,X_{n+m})$$

unabhängig.

Beweis. Steht in § 2.3 von [Eva13]. Er zitiert das Buch [Bre68] von Breiman. □

3.20 Theorem. *Es sei* $(\Omega,\mathcal{U},P)$ *ein Wahrscheinlichkeitsraum, auf dem es abzählbar viele unabhängige,* $N(0,1)$*-verteilte Zufallsvariablen gibt. Dann gibt es auf* Ω *einen eindimensionalen Wienerschen Prozess.*

3.21 Bezeichnung. Unter $\mathcal{U}(X(t)|t \in T)$ versteht man die kleinste σ-Algebra, die alle Mengen der Form $X(t)^{-1}(B)$, $t \in T$, $B \in \mathcal{B}(\mathbb{R}^n)$, enthält.

3.22 Definition. Ein $\mathbb{R}^n$-wertiger stochastischer Prozess ist eine n-dimensionale Brownsche Bewegung oder ein n-dimensionaler Wienerscher Prozess, falls

(a) W^k für $k = 1, \dots, n$ ein eindimensionaler Wienerscher Prozess ist, und

(b) die σ-Algebren $\mathcal{U}(W^k) = \mathcal{U}(W^k(t)|t \geq 0)$ unabhängig sind.

Es gibt einen n-dimensionalen Wienerschen Prozess. Dazu konstruiert man mit Theorem 3.20 n unabhängige Prozesse.

3.23 Lemma. *Sei W ein n-dimensionaler Wienerscher Prozess. Dann gelten für* $k, \ell = 1, \dots, n$

$$\mathbb{E}\big(W^k(t)W^\ell(s)\big) = \min(t, s)\delta_{k,\ell},$$
$$\mathbb{E}\big((W^k(t) - W^k(s))(W^\ell(t) - W^\ell(s))\big) = (t - s)\delta_{k,\ell}, \qquad 0 \leq t \leq s.$$

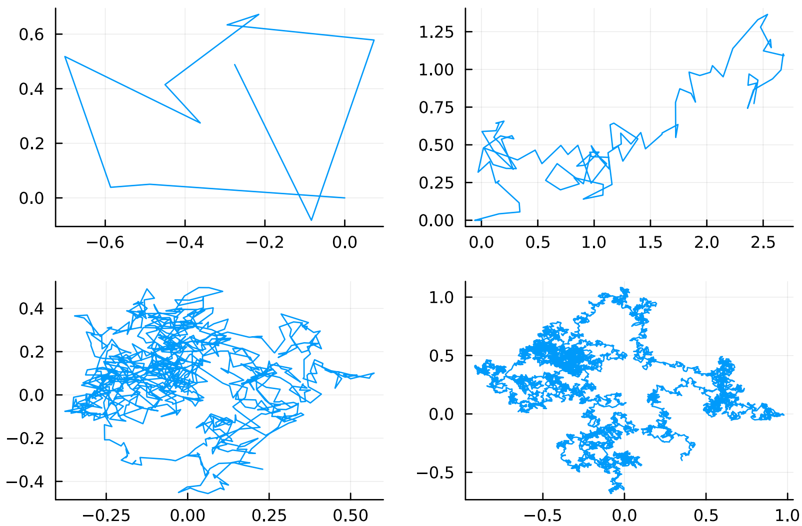

Abbildung 3.1: Simulationen einer Brownschen Bewegung mit Schrittweiten 10^{-1}, 10^{-2}, 10^{-3} und 10^{-4}

Abbildung 3.1 zeigt Simulationen der Brownschen Bewegung mit verschieden feinen Zeitschritten.

Bemerkung. Nach Konstruktion ist der hier definierte Wienersche Prozess Hölder-stetig zu allen Exponenten $\gamma < \frac{1}{2}$. Man kann Wienersche Prozesse aber auch anders erhalten. In Abschnitt 3.4.1 weist Evans [Eva13] nach, dass die Hölder-Stetigkeit bereits aus den Axiomen des Wienerschen Prozesses folgt.

3.24 Theorem (Dvoretzky, Erdős und Kakutani). *Für $\gamma > \frac{1}{2}$ ist der Pfad $t \mapsto W(t, \omega)$ f. s. nirgends Hölder-stetig zum Exponent γ.*

Bemerkung. Simon zitiert in den Notizen zu § 4.16 von [Sim15] die Aussage, dass die Pfade des Wienerschen Prozesses fast sicher nicht Hölder-stetig zum Exponenten $\frac{1}{2}$ sind.

3.25 Korollar. *Die Pfade der Brownschen Bewegung sind fast sicher nirgends differenzierbar.*

3.26 Bemerkung. Eine *Zerlegung* des Intervalls $[a, b]$ ist ein Tupel $p = (p_0, \ldots, p_n)$ mir $a = p_0 < p_1 < \cdots < p_n = b$, wobei $n \in \mathbb{N}$ beliebig ist. Für $f \colon [a, b] \to \mathbb{K}$ definieren wir

$$V_p(f) = \sum_{j=1}^{n} |f(p_j) - f(p_{j-1})|.$$

Die *Variation* von f ist definiert als $\sup_p V_p(f)$, wobei p über alle Zerlegungen von $[a, b]$ variiert. Wenn die Variation von f endlich ist, dann sagt man, f sei eine Funktion mit *beschränkter Variation.*

Für Funktionen beschränkter Variation hat Stieltjes einen Maßbegriff eingeführt, der zum Riemann-Stieltjes-Integral führt. Details findet man z. B. in §4.1 von Simon [Sim15].

3.27 Satz. *Die Pfade des Wienerschen Prozesses sind fast sicher nicht von beschränkter Variation.*

Dieser Satz impliziert, dass das Integral bzgl. W nicht als Riemann-Stieltjes-Integral konstruiert werden kann.

3.28 Definition. Eine *Filtration* ist eine monoton wachsende Familie $(\mathcal{V}(t))_{t \geq 0}$ von σ-Algebren.

3.29 Beispiel. Es sei X in stochastischer Prozess und es sei $s \geq 0$. Die kleinste σ-Algebra, die alle $\mathcal{U}(X(t))$ für $0 \leq t \leq s$ enthält, bezeichnen wir als *Geschichte* des Prozesses bis einschließlich der Zeit s. Wir schreiben $\mathcal{U}(s)$ dafür.

$(\mathcal{U}(s))_{s \geq 0}$ ist eine Filtration.

3.30 Definition. Sei $(\Omega, \mathcal{U}, P)$ ein Wahrscheinlichkeitsraum und sei $\mathcal{V} \subseteq U$ eine σ-Algebra. Die Zufallsvariable

$$P(A|\mathcal{V}) = \mathbb{E}(\chi_A|\mathcal{V}), \quad A \in \mathcal{V},$$

ist die *bedingte Wahrscheinlichkeit* von A, gegeben $\mathcal{V}$.

3.31 Definition. Ein $\mathbb{R}^n$-wertiger stochastischer Prozess ist ein *Markow-Prozess,* wenn für jede Wahl von $0 \leq s \leq t$ und jede Borelmenge B gilt

$$P(X(t) \in B|\mathcal{U}(s)) = P(X(t) \in B|X(s)).$$

3.32 Theorem. *Der Wienersche Prozess ist Markowsch.*

Beweis. Das ist ein Spezialfall von Theorem 9.6, welches wir später zeigen werden, ohne auf diesen Satz zurückzugreifen. □

Die Markow-Eigenschaft des Prozesses bedeutet, dass er sich nicht daran "erinnert", wie er an einen Punkt gelangt ist. Speziell weiß er nicht, aus welcher Richtung er gekommen ist und kann daher nicht in derselben Richtung weiterfahren. Das erklärt heuristisch, dass Pfade f. s. nicht differenzierbar sind.

4 Das Itô-Integral

Ab jetzt ist W immer der Wienersche Prozess auf $(\Omega, \mathcal{U}, P)$.

Ziel: Eine stochastische Differentialgleichung soll die folgende Form haben

$$\begin{aligned} dX &= b(X,t)\,dt + B(X,t)\,dW, \\ X(0) &= X_0 \end{aligned}$$

wobei X, X_0 und B Zufallsvariablen und W die Brownsche Bewegung sind. Für $B = 0$ hat man eine gewöhnliche Differentialgleichung in laxer Schreibweise, die man wie in der Analysis II in eine Integralgleichung verwenden kann. Für $B \neq 0$ spielt der letzte Term die Rolle eines Rauschens. Die zugehörige Integralgleichung ist dann

$$X(t) = X_0 + \int_0^t b(X,s)\,ds + \int_0^t B(X,s))\,dW.$$

Das erste Integral kann man pfadweise verstehen. Die Interpretation des zweiten Integrals ist Thema dieses Kapitels. Da die Pfade von W unbeschränkte Variation besitzen, ist die Interpretation als Stieltjes-Integral nicht möglich.

4.1 Definition (Paley, Wiener und Zygmund (1933)). Sei $g\colon [0,T] \to \mathbb{R}$ stetig differenzierbar mit $g(0) = g(T) = 0$. Dann definieren wir

$$\int_0^t g\,dW = -\int_0^t g' W\,dt.$$

Die Funktion g ist also deterministisch, aber das Integral ist eine Zufallsvariable.

4.2 Lemma. *Unter den Voraussetzungen der Definition gelten*

$$\mathbb{E}\left(\int_0^T g\,dW\right) = 0.$$

$$\mathbb{E}\left(\left(\int_0^T g\,dW\right)^2\right) = \int_0^T g(t)^2\,dt.$$

4.3 Bemerkung. Das Lemma zeigt die Existenz einer Abbildung

$$\Psi\colon C^1[0,T] \to L^2(\Omega, \mathcal{U}, P), \quad g \mapsto \int_0^T g\,dW.$$

Wenn wir den $C^1[0,T]$ mit der L^2-Norm versehen, dann ist diese Abbildung sogar stetig, genauer

$$\|\Psi(g)\|_2 \leq \|g\|_2.$$

Sie setzt sich also fort zu einer stetigen Abbildung $L^2[0,T] \to L^2(\Omega,\mathcal{U},P)$.

Die Definition von Paley, Wiener und Zygmund ist somit auf den $L^2[0,T]$ fortgesetzt worden.

4.4 Bezeichnung. Es sei Z die durch $0 = t_0 < t_1 < \cdots < t_m = T$ gegebene Zerlegung von $[0,T]$.

(a) Die *Feinheit* dieser Zerlegung ist

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

(b) Für $0 \leq \lambda \leq 1$ ist die *Riemannsumme* zu $\int_0^T W\,dW$ definiert als

$$R(Z,\lambda) = \sum_{k=0}^{m-1} W(\tau_k)\big(W(t_{k+1}) - W(t_k)\big),$$

wobei $\tau_k = (1-\lambda)t_k + \lambda t_{k+1}$.

4.5 Lemma (Quadratische Variation). *Seien $0 \leq a < b$. Für $n \in \mathbb{N}$ sei durch $a = t_0^n < t_1^n < \cdots < t_{m_n}^n = b$ eine Zerlegung Z^n von $[a,b]$ gegeben. (Das n ist also ein Index, kein Exponent.) Für festes $\lambda \in [0,1]$ sei $\tau_k^n = (1-\lambda)t_k^n + \lambda t_{k+1}^n$. Wenn $\lim_{n\to\infty}|Z^n| = 0$, dann konvergiert*

$$\sum_{k=0}^{m_n-1} \big(W(\tau_k^n) - W(t_k^n)\big)^2$$

im $L^2(\Omega,\mathcal{U},P)$ gegen $\lambda(b-a)$.

4.6 Lemma. *Sei $0 \leq \lambda \leq 1$ und sei $(Z^n)_{n\in\mathbb{N}}$ eine Folge von Zerlegungen von $[0,T]$ mit $\lim_{n\to\infty}|Z^n| = 0$. Dann*

$$\lim_{n\to\infty} R(Z^n,\lambda) = \frac{1}{2}W(T)^2 + \left(\lambda - \frac{1}{2}\right)T$$

in $L^2(\Omega,\mathcal{U},P)$.

Bemerkung. Das Itôsche Integral entspricht dem Fall $\lambda = 0$. Die Alternative $\lambda = \frac{1}{2}$ führt zum Stratonovich-Integral.

4.7 Bezeichnung. (a) Wir bezeichnen die Geschichte des Wienerschen Prozesses bis zur Zeit t mit $\mathcal{W}(t)$.

(b) Die *Zukunft* des Wienerschen Prozesses ab der Zeit t ist die kleinste σ-Algebra, für die alle Zufallsvariablen $W(s) - W(t)$, $s \geq t$, messbar sind. Wir bezeichnen sie mit $\mathcal{W}^+(t)$.

4.8 Satz (Klenke [Kle20], Satz 2.26). *Sei* K *eine Menge, seien* I_k, $k \in K$, *paarweise disjunkte Indexmengen und sei* $I = \bigcup_{k \in K} I_k$. *Durch* $(X_i)_{i \in I}$ *sei eine Familie unabhängiger Zufallsvariablen gegeben. Für* $k \in K$ *sei* $\mathcal{U}_k$ *die kleinste* σ*-Algebra, bzgl. derer alle* X_i *mit* $i \in K$ *messbar sind. Dann sind die* σ*-Algebren* $\mathcal{U}_k$, $k \in K$, *unabhängig.*

4.9 Bezeichnung. Seien $\mathcal{A}_1$ und $\mathcal{A}_2$ σ-Algebren. Die *Produkt-σ-Algebra* $\mathcal{A}_1 \otimes \mathcal{A}_2$ ist die kleinste σ-Algebra, die alle Produkte $M_1 \times M_2$ mit $M_j \in \mathcal{A}_j$ enthält.

4.10 Lemma. *Seien* $0 \leq t < s$.

(a) Dann sind $W(s) - W(t)$ *und* $\mathcal{W}(t)$ *unabhängig.*

(b) $W(s)$ *ist nicht* $\mathcal{W}(t)$*-messbar.*

4.11 Definition. Sei $(\mathcal{F}(t))_{t \geq 0}$ eine Filtration. Ein $\mathbb{R}^n$-wertiger Prozess $(X(t))_{t \geq 0}$ heißt $\mathcal{F}(t)$*-adaptiert,* wenn für jedes $t \geq 0$ die Zufallsvariable $X(t)$ messbar bzgl. $\mathcal{F}(t)$ ist.

4.12 Beispiel. Der durch $X(t) = W(\frac{t}{2})$ gegebene Prozess ist $\mathcal{W}(t)$-adaptiert, der durch $Y(t) = W(2t)$ gegebene nicht.

4.13 Bezeichnung. Sei $p = 1$ oder $p = 2$ und sei $T > 0$. Ein $\mathbb{R}^n$-wertiger Prozess $X = (X(t))_{t \geq 0}$ liegt in $\mathbb{L}^p[0, T]$, wenn

(a) $X\colon [0, \infty[\times \Omega \to \mathbb{R}^n$ messbar bzgl. $\mathcal{B}([0, \infty[) \otimes \mathcal{U}$ ist,

(b) X $\mathcal{W}(t)$-adaptiert ist,

(c) $\mathbb{E}\left(\int_0^T |X(t)|^p \, dt\right) < \infty$.

Bemerkung. Einen solchen Prozess bezeichnet man als *progressiv messbar.* Das wird im Buch von Klenke [Kle20] ausgeführt.

4.14 Lemma. $\mathbb{L}^p[0, T]$, $p = 1, 2$ *ist ein* $\mathbb{R}$*-Vektorraum.*

4.15 Bezeichnung. Ein Prozess G ist *elementar,* wenn es eine Zerlegung $Z = (t_0, t_1, \ldots, t_m)$ von $[0, T]$ sowie Zufallsvariable G_k, $k = 0, \ldots, m-1$, gibt. so dass

$$G(t) = g_k \text{ für } t_k \leq t < t_{k+1}.$$

Für einen elementaren Prozess $G \in \mathbb{L}^2[0, T]$ definieren wir das Itô-Integral durch

$$\int_0^T G \, dW = \sum_{k=0}^{m-1} G_k(W(t_{k+1}) - W(t_k))). \tag{4.1}$$

4.16 Bemerkung. (a) Weil der Prozess adaptiert ist, ist G_k messbar bzgl. $\mathcal{W}(t_k)$.

(b)

$$\sum_{k=1}^{m-1} W(t_k)\chi_{[t_k,t_{k+1}]} \in \mathbb{L}^2[0,T]$$

aber

$$\sum_{k=1}^{m-1} W\left(\frac{1}{2}(t_k+t_{k+1})\right)\chi_{[t_k,t_{k+1}]} \notin \mathbb{L}^2[0,T].$$

4.17 Lemma. *Seien* $a, b \in \mathbb{R}$ *und seien* $G, H \in \mathbb{L}^2[0,T]$ *elementare Prozesse. Dann gelten*

$$\int_0^T (aG+bH)\,dW = a\int_0^T G\,dW + b\int_0^T b\,dW,$$

$$\mathbb{E}\left(\int_0^T G\,dW\right) = 0,$$

$$\mathbb{E}\left(\left(\int_0^T G\,dW\right)^2\right) = \mathbb{E}\left(\int_0^T G^2\,dt\right) \qquad \textit{(Itô-Isomorphie für elementare Prozesse).}$$

4.18 Lemma ([Øks03], § 3.1)**.** *Sei* $X \in \mathbb{L}^2[0,T]$*. Dann existiert eine Folge* $(X_n)_{n\in\mathbb{N}}$ *beschränkter, elementarer stochastischer Prozesse in* $\mathbb{L}^2[0,T]$*, die in dem Sinn gegen* X *konvergiert, dass*

$$\lim_{n\to\infty} \mathbb{E}\left(\int_0^T |X-X_n|^2\,dt\right) = 0.$$

"Beschränkt" bedeutet in diesem Zusammenhang, dass es zu jedem n *ein* C *gibt, so dass* $|X_n| \le C$ *f. s.*

Sei $S^2[0,T]$ der Unterraum der elementaren Prozesse in $\mathbb{L}^2[0,T]$. Wegen Lemma 4.17 kann die Abbildung

$$S^2[0,T] \to L^2(\Omega,\mathcal{U},P), \quad X \mapsto \int_0^T X\,dW,$$

zu einer Abbildung $\mathbb{L}^2[0,T] \to L^2(\Omega,\mathcal{U},P)$ fortgesetzt werden.

4.19 Definition. Diese Fortsetzung ist das Itô-Integral.

Analog definiert man $\int_R^S X\,dW$ für $0 \le R \le S$.

Wenn die Integrationsvariable benannt werden muss, schreibt man auch

$$\int_0^T X(t)\,dW(t).$$

4.20 Theorem (Itô-Isomorphie). *Für* $X \in \mathbb{L}^2[0,T]$ *gilt*

$$\mathbb{E}\left(\left(\int_0^T X\,dW\right)^2\right) = \mathbb{E}\left(\int_0^T X^2\,dt\right).$$

4.21 Korollar. *Wenn* $X, X_1, X_2, \dots \in \mathbb{L}^2[0,T]$ *und*

$$\mathbb{E}\left(\int_0^T (X_n - X)^2\,dt\right) \to 0,$$

dann

$$\int_0^T X_n\,dW \to \int_0^T X\,dW.$$

4.22 Beispiel.

$$\int_0^T W\,dW = \frac{1}{2}W(T)^2 - \frac{T}{2}.$$

4.23 Satz. *Seien* $X, Y \in \mathbb{L}^2[0,T]$ *und* $0 \le S < U < T$.

(a) $\displaystyle\int_S^T X\,dW = \int_S^U X\,dW + \int_U^T X\,dW.$

(b) $\displaystyle\int_S^T (cX + Y)\,dW = c\int_S^T X\,dW + \int_S^T Y\,dW.$

(c) $\mathbb{E}\left(\int_S^T X\,dW\right) = 0.$

(d) $\int_S^T X\,dW$ *ist messbar bzgl.* $\mathcal{W}(T)$.

4.24 Definition. Sei $(\Omega, \mathcal{U}, P)$ ein beliebiger Wahrscheinlichkeitsraum und sei $(\mathcal{V}(t))_{t\ge0}$ eine Filtration.

Ein Prozess $(X(t))_{t\ge0}$ ist ein *Martingal* bzgl. $(\mathcal{V}(t))_{t\ge0}$, wenn die folgenden Bedingungen gelten:

(a) Für jedes t ist $X(t)$ messbar bzgl. $\mathcal{V}(t)$.

(b) Für jedes t gilt $\mathbb{E}(|X(t)|) < \infty$.

(c) $\mathbb{E}(X(s)|\mathcal{V}(t)) = X(t)$ falls $s \ge t$.

Wenn die Filtration nicht genannt wird, dann ist die Geschichte des Prozesses gemeint.

4.25 Satz. *Der Wienersche Prozess ist ein Martingal.*

4.26 Theorem (Doobsche Ungleichung für Martingale). *Sei* $(X(t))_{t\geq 0}$ *ein Martingal bzgl. einer Filtration* $(\mathcal{V}(t))_{t\geq 0}$*, dessen Pfade fast sicher stetig sind, dann gilt für* $p \geq 1$*,* $T \geq 0$ *und* $\lambda > 0$

$$P\left(\sup_{0\leq t\leq T} |X(t)| \geq \lambda\right) \leq \frac{1}{\lambda^p}\mathbb{E}(|X(T)|^p)\,.$$

4.27 Lemma. *Sei* $X \in \mathbb{L}^2[0,T]$ *elementar. Dann ist* $\left(\int_0^t X\,dW\right)_{0\leq t\leq T}$ *ein Martingal bzgl.* $\mathcal{W}(t)$*.*

4.28 Theorem. *Sei* $X \in \mathbb{L}^2[0,T]$*. Dann besitzt der Prozess* $\left(\int_0^t X\,dW\right)_{0\leq t\leq T}$ *eine Version mit stetigen Pfaden.*

4.29 Korollar. *Sei* X *für jedes* T *in* $\mathbb{L}^2[0,T]$*. Dann ist der durch* $M(t) = \int_0^t X\,dW$ *gegebene Prozess ein Martingal bzgl.* $(\mathcal{W}(t))_{t\geq 0}$ *und es gilt für jedes* $\lambda > 0$

$$P\left(\sup_{0\leq t\leq T} |M(t)| \geq \lambda\right) \leq \frac{1}{\lambda^2}\mathbb{E}\left(\int_0^T X(t)^2\,dt\right).$$

In Øksendal [Øks03] und Friedman [Fri06] werden die Klassen $\mathbb{L}^p[0,T]$ verallgemeinert. Die Forderungen an X werden abgeschwächt. Die Eigenschaft (b) der Bezeichnung 4.13 wird verallgemeinert zu

(b) Es gibt eine Filtration $(\mathcal{H}(t))_{t\geq 0}$ derart, dass
   (i) $W(t)$ ist ein Martingal bzgl. $\mathcal{H}(t)$.
   (ii) $X(t)$ ist $\mathcal{H}(t)$-adaptiert.

Evans [Eva13] und Friedman [Fri06] regeln diesen Punkt über die Unabhängigkeit von der Zukunft.

Die Bedingung (c) aus 4.13 wird abgeschwächt zu

(c) $\int_0^T f(s)^2\,ds < \infty$ fast sicher.

4.30 Definition. Wir bezeichnen mit $\mathbb{W}_{\mathcal{H}}[S,T]$ die Klasse der eindimensionalen Prozesse $X = X(t)_{S\leq t\leq T}$ mit den Eigenschaften (a) aus 4.13 und (b) und (c) von oben. Ferner schreiben wir $\mathbb{W}_{\mathcal{H}}$ für $\bigcap_{t\geq 0} \mathbb{W}_{\mathcal{H}}[0,T]$. Wenn $\mathcal{H}(t) = \mathcal{W}(t)$, dann lassen wir den Index weg.

4.31 Bemerkung. Sei $\mathcal{W}$ ein n-dimensionaler Wienerscher Prozess und sei $\mathcal{W}(t)$ seine Geschichte. Für $1 \leq k \leq n$ bezeichnen wir mit $\mathcal{W}_k$ die k-te Komponente von $\mathcal{W}$. Dann zeigt der Beweis von Satz 4.25, dass $W_k \in \mathbb{W}_{\mathcal{W}}$.

4.32 Definition. Eine Folge $(X_n)_{n\in\mathbb{N}}$ von Zufallsvariablen konvergiert *im Maße* gegen eine Zufallsvariable X, wenn für jedes $\epsilon > 0$

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

4.33 Theorem ([Øks03], § 3.3, [Fri06], Abschnitt 4). *Für* $X \in \mathbb{W}_{\mathcal{H}}$ *existiert eine Folge* $(X_n)_{n\in\mathbb{N}}$ *von elementaren Prozessen in* $\mathbb{W}_{\mathcal{H}}$*, so dass*

$$\int_0^t |X_n(s) - X(s)|^2 \, ds \to 0 \text{ im Maße.}$$

Für X_n *kann dann* $\int_0^t X_n \, dW$ *wie zuvor konstruiert werden. Die Folge* $\int_0^t X_n \, dW$ *konvergiert im Maße gegen eine Zufallsvariable, die man als Itô-Integral* $\int_0^t X \, dW$ *schreibt. Dieser Grenzwert ist unabhängig von der Wahl der Folge.*

Auch das erweiterte Itô-Integral besitzt eine Version mit stetigen Pfaden.

Bemerkung. Für $X \in \mathbb{W}_{\mathcal{H}}$ braucht das Itô-Integral kein Martingal zu sein.

5 Die Itô-Formel

5.1 Definition. Sei W ein eindimensionaler Wienerscher Prozess. Ein *eindimensionaler Itô-Prozess* ist ein stochastischer Prozess X der Form

$$X(t) = X(0) + \int_0^t F(s)\,ds + \int_0^t G(s)\,dW$$

mit $G \in \mathbb{W}_{\mathcal{H}}$, so dass

$$P\left(\left\{\omega \,\middle|\, \forall t : \int_0^t G(s,\omega)^2\,ds < \infty\right\}\right) = 1, \tag{5.1}$$

und $\mathcal{H}(t)$-adaptiertem F, so dass

$$P\left(\left\{\omega \,\middle|\, \forall t : \int_0^t |F(s,\omega)|\,ds < \infty\right\}\right) = 1. \tag{5.2}$$

In diesem Fall schreiben wir

$$dX = F\,dt + G\,dW.$$

5.2 Bemerkung. Die Bedingungen (5.1) und (5.2) sind erfüllt, falls F und G fast sicher stetige Pfade besitzen.

5.3 Lemma.

$$\begin{aligned} d(W^2) &= 2W\,dW + dt, \\ d(tW) &= W\,dt + t\,dW. \end{aligned}$$

5.4 Theorem (Itôsche Produktregel). *Seien X_1 und X_2 Itô-Prozesse mit*

$$\begin{aligned} dX_1 &= F_1\,dt + G_1\,dW, \\ dX_2 &= F_2\,dt + G_2\,dW, \end{aligned}$$

wobei F_j und G_j wie in Definition 5.1. Dann gilt

$$d(X_1X_2) = X_2\,dX_1 + X_1\,dX_2 + G_1G_2\,dt.$$

5.5 Bemerkung. (a) Was tatsächlich gezeigt werden muss, ist

$$\int_s^r X_2\,dX_1 = X_1(r)X_2(r) - X_1(s)X_2(s) - \int_s^r X_1\,dX_2 - \int_s^r G_1G_2\,dt. \tag{5.3}$$

Das ist die partielle Integration für das Itô-Integral.

(b) Dazu muss man zuerst die Existenz dieser drei Integrale zeigen. Dazu seien ohne Einschränkung $s = 0$ und $X_1(0) = X_2(0) = 0$. Wenn $I_2(t) = \int_0^t G_2(s)\,dW(s)$, dann

$$\int_0^r X_2\,dX_1 = \int_0^r \int_0^t F_2(s)\,ds F_1(t)\,dt + \int_0^r I_2(t)F_1(t)\,dt \\ + \int_0^r \int_0^t F_2(s)\,ds G_1(t)\,dW + \int_0^r G_1(t)I_2(t)\,dW(t).$$

Aus den Voraussetzungen folgt, dass X_1 und X_2 fast sicher stetige Pfade haben. Also existieren die Integrale bzgl. dt und die Integranden erfüllen (5.2). Setzen wir nun $G(t) = G_1(t)\int_0^t F_2(s)\,ds$, so müssen wir $G \in \mathbb{W}_{\mathcal{H}}$ zeigen. Die Messbarkeitseigenschaften bekommt man aus der Approximation. Noch zu zeigen ist

$$\int_0^r G(t)^2\,ds < \infty \text{ fast sicher.}$$

Weil $\int_0^t |F_2(s)|\,ds < \infty$ fast sicher, folgt das aus der entsprechenden Eigenschaft von G_1.

Jetzt setzen wir noch $\widetilde{G}(t) = G_1(t)I_2(t)$. Dann $\int_0^r \widetilde{G}(t)^2 < \infty$ fast sicher, weil I_2 stetige Pfade hat. Nach [Fri06], § 4.3, ist I_2 adaptiert an $\mathcal{W}(t)$. Daher ist G adaptiert an $\mathcal{H}(t)$.

Beweis. Schritt 1: Wir zeigen (5.3), aber nur für $s = 0$ und $X_1(0) = X_2(0) = 0$, um weniger schreiben zu müssen. Im ersten Schritt sind F_j und G_j außerdem zeitunabhängig. Das bedeutet

$$X_i(t) = F_i t + G_i W(t).$$

Es gilt

$$\begin{aligned}
&\int_0^r (X_2\,dX_1 + X_1\,dX_2 + G_1G_2)\;dt \\
&\quad = \int_0^r (X_1F_2 + X_2F_1)\;dt + \int_0^r (X_1G_2 + X_2G_1)\;dW + \int_0^r G_1G_2\,dt \\
&\quad = \int_0^r ((F_1t + G_1W)\,F_2 + (F_2t + G_2W)\,F_1)\;dt \\
&\qquad + \int_0^r ((F_1t + G_1W)\,G_2 + (F_2t + G_2W)\,G_1)\;dW + G_1G_2r \\
&\quad = F_1F_2r^2 + (G_1F_2 + G_2F_1)\left(\int_0^r W\,dt + \int_0^r t\,dW\right) + 2G_1G_2\int_0^r W\,dW + G_1G_2r.
\end{aligned}$$

Wegen des Lemmas können wie die Klammer ersetzen durch $rW(r)$ und das andere Integral durch $\frac{1}{2}\left(W^2(r) - r\right)$. Damit erhalten wir

$$\begin{aligned}
\int_0^r (X_2\,dX_1 + X_1\,dX_2 + G_1G_2)\;dt &= F_1F_2r^2 + (G_1F_2 + G_2F_1)rW(r) + G_1G_2W^2(r) \\
&= X_1(r)X_2(r).
\end{aligned}$$

Damit ist (5.3) für diesen Spezialfall gezeigt.

Schritt 2: Wenn F_j und G_j elementare Prozesse sind, dann führen wir den ersten Schritt auf denjenigen Intervallen aus, auf denen alle Prozesse zeitunabhängig sind. Die erhaltene Summe enthält eine Teleskopsumme.

Schritt 3: Wir approximieren F_j und G_j durch elementare Prozesse und gehen zum Grenzwert über. □

5.6 Bemerkung. Sei $g \in C_0^1[0,T]$. Wenn wir setzen $X_1 = g$ und $X_2 = W$, dann $dX_1 = g'\,dt$ und $dX_2 = dW$. Dann zeigt Theorem 5.4, dass für X_1 das Itô-Integral mit dem Paley-Wiener-Zygmund Integral übereinstimmt.

5.7 Lemma. *Sei* $dX = F\,dt + G\,dW$ *mit* F *und* G *wie in Definition 5.1. Dann gilt für* $m \in \mathbb{N} \setminus \{1\}$

$$d(X^m) = mX^{m-1} + \frac{m(m-1)}{2} X^{m-2} G^2\,dt.$$

Beweis durch vollständige Induktion als Übung.

5.8 Bezeichnung. Wir bezeichnen partielle Ableitungen duch Indices, also

$$u_t = \frac{\partial u}{\partial t}, \quad u_{xx} = \frac{\partial^2 u}{\partial u^2} \text{ usw.}$$

5.9 Theorem (Itôsche Kettenregel, Formel von Itô). *Sei* $dX = F\,dt + G\,dW$ *mit* F *und* G *wie in Definition 5.1. Ferner sei* $u \in C^1(\mathbb{R} \times [0,T])$ *und die Ableitung* u_{xx} *existiere ebenfalls und sei stetig. Dann besitzt*

$$Y = u(X(t), t)$$

das stochastische Differential

$$dY = u_t\,dt + u_x\,dX + \frac{1}{2} u_{xx} G^2\,dt.$$

5.10 Bemerkung. Das bedeutet für $0 \le s \le r \le T$

$$Y(r) - Y(s) = \int_s^r \left(u_t(X(t),t) + u_x(X(t),t)F(t) + \frac{1}{2} u_{xx}(X(t),t)G^2(t) \right) dt + \int_s^r u_x(X(t),t)G(t)\,dW.$$

5.11 Beispiel. Wir setzen $X = W$. Dann $dX = dW$ und daher $F = 0$ und $G = 1$. Wir wenden die Itô-Formel an auf $u(x,t) = x^2$ und erhalten

$$W^2(T) = \int_0^T dt + \int_0^T 2W\,dW,$$

also die Formel, die wir bereits durch die Riemannsumme erhalten haben.

Für ein beliebiges stochastisches Differential dX und $u(x,t) = x^m$, $m \ge 2$, erhalten wir das Ergebnis, welches wir in Lemma 5.7 aus der Produktformel hergeleitet hatten. Dieses Ergebnis wird allerdings beim Beweis der Itô-Formel verwendet.

5.12 Beispiel. Sei $X = W$, also $F \equiv 0$ und $G \equiv 1$, und sei $u(x,t) = \exp\Big(\lambda x - \frac{\lambda^2 t}{2}\Big)$. Wir setzen

$$Y(t) = \exp\left(\lambda W(t) - \frac{\lambda^2 t}{2}\right).$$

Dann

$$dY = \left(-\frac{\lambda^2}{2}Y(t) + \frac{\lambda^2}{2}Y(t)\right) dt + \lambda Y(t)\, dW = \lambda Y(t)\, dW.$$

Wir können das auch anders berechnen, nämlich

$$X(t) = -\frac{\lambda^2 t}{2} + \lambda W, \quad u(x,t) = e^x,$$

also

$$F(t) = -\frac{\lambda^2}{2}, \quad G(t) = \lambda.$$

Dann

$$\begin{aligned} dY &= e^X\, dX + \frac{1}{2} e^X \lambda^2\, dt \\ &= \exp\left(-\frac{\lambda^2 t}{2} + \lambda W(t)\right)(\, dt + \lambda\, dW) + \frac{\lambda^2}{2}\exp\left(-\frac{\lambda^2 t}{2} + \lambda W(t)\right) dt \\ &= \lambda Y(t)\, dW. \end{aligned}$$

5.13 Lemma. *Für u wie in Theorem 5.9 und $T > 0$ gibt es eine Folge $(u^n)_{n\in\mathbb{N}}$ von Polynomen in x und t, so dass*

$$u^n \to u, u^n_t \to u_t, u^n_x \to u_x \textit{ und } u^n_{xx} \to u_{xx}$$

gleichmäßig auf den kompakten Teilmengen von $\mathbb{R} \times [0,T]$.

5.14 Definition. Die *Hermite-Polynome* sind definiert durch

$$h_n(x,t) = \frac{(-t)^n}{n!} e^{\frac{x^2}{2t}} \frac{d^n}{dx^n} e^{-\frac{x^2}{2t}}, \quad n \in \mathbb{N}_0.$$

5.15 Bemerkung. In der Einführung in die Funktionalanalysis werden ebenfalls Hermite-Polynome erklärt, dann aber als Spezialfall $t = \frac{1}{2}$ und mit anderen Faktoren.

Für festes t bilden die Hermite-Polynome $h_n(\cdot, t)$ eine Orthonormalbasis in einem geeignet gewichteten $L^2(\mathbb{R})$.

5.16 Theorem. *Für $n \in \mathbb{N}_0$ gilt*

$$dh_{n+1}(W(t), t) = h_n(W(t), t)\, dW.$$

5.17 Bezeichnung. Wenn wir setzen

$$\begin{aligned} dt\,dt &= 0,\\ dt\,dW &= 0,\\ dW\,dW &= dt, \end{aligned}$$

dann haben Produkt- und Kettenregel im eindimensionalen Fall jeweils die folgende Form

$$\begin{aligned} d(u(X,t)) &= u_t\,dt + u_x\,dX + \frac{1}{2}u_{xx}\,dX\,dX,\\ d(X_1X_2) &= X_1\,dX_2 + X_2\,dX_1 + dX_1\,dX_2. \end{aligned}$$

5.18 Bezeichnung. Im mehrdimensionalen Fall sei $W(t) = (W^1(t),\dots,W^m(t))$ ein m-dimensionaler Wienerscher Prozess. Wenn F^i, $i = 1,\dots,n$, und G^{ij}, $i = 1,\dots,n$, $j = 1,\dots,m$, die Voraussetzungen aus 5.1 erfüllen, dann setzen wir

$$\begin{aligned} dX^1 &= F^1\,dt + G^{11}\,dW^1 + \dots + G^{1m}\,dW^m,\\ &\vdots\\ dX^n &= F^n\,dt + G^{n1}\,dW^1 + \dots + G^{nm}\,dW^m, \end{aligned}$$

und kürzen das ab als

$$dX = F\,dt + G\,dW.$$

In diesem Fall ist X ein *n-dimensionaler Itô-Prozess.*

5.19 Theorem (Itô-Formel). *Sei X ein n-dimensionaler Itô-Prozess und sei $u\colon [0,\infty[\times\mathbb{R}^n \to \mathbb{R}^p$ von der Klasse C^2. Dann ist der Prozess $Y(t) = u(X(t),t)$ wieder ein Itô-Prozess und es gilt für $k = 1,\dots,p$*

$$dY^k = u^k_t\,dt + \sum_{i=1}^n u^k_{x_i}\,dX^i + \frac{1}{2}\sum_{i,j=1}^n u^k_{x_ix_j}\,dX^i\,dX^j,$$

wobei das Produkt formal definiert ist als

$$\begin{aligned} dt\,dt &= 0,\\ dt\,dW^i &= 0,\\ dW^i\,dW^j &= \delta_{ij}\,dt. \end{aligned}\tag{5.4}$$

Wir beweisen nur den folgenden Spezialfall.

5.20 Beispiel. Seien W^1 und W^2 unabhängige Wienersche Prozesse. Dann

$$d(W^1W^2) = W^1\,dW^2 + W^2\,dW^1.$$

5.21 Beispiel. Sei $n \geq 2$ und sei W ein n-dimensionaler Wienerscher Prozess und sei

$$R(t) = |W(t)|.$$

Wir wenden die Itô-Formel an, obwohl die Funktion $u(x) = |x|$ im Ursprung nicht differenzierbar ist. In Exercise 9.7 von [Øks03] wird gezeigt, dass W f. s. den Ursprung kein zweites Mal trifft. Man kann sich überlegen, dass daher die Anwendung der Itô-Formel gerechtfertigt ist.

Dann ist F gleich 0 und G ist die Einheitsmatrix. Ferner ist $u\colon \mathbb{R}^n \times [0, \infty[\to \mathbb{R}$ gegeben durch $u(x,t) = \sqrt{(x^1)^2 + \cdots + (x^n)^2}$. Dann

$$u_t = 0, \quad u_{x_j} = \frac{x_j}{|x|}, \quad u_{x_i x_j} = \frac{\delta_{ij}|x|^2 - x_i x_j}{|x|^3}.$$

Die Itô-Formel ergibt nun

$$dR = \sum_{i=1}^{n} \frac{W^i}{R} \, dW^i + \frac{n-1}{2R} \, dt.$$

Diesen Prozess bezeichnet man als n-dimensionalen *Bessel-Prozess.*

6 Stochastische Differentialgleichungen

6.1 Beispiel. Für $\alpha, r \in \mathbb{R}$ betrachten wir die Differentialgleichung eines stochastischen Wachstums- bzw. Abklingprozesses

$$dN(t) = N(t)(r\,dt + \alpha\,dW). \tag{6.1}$$

Wir schreiben das formal um zu

$$\frac{dN(t)}{N(t)} = r\,dt + \alpha\,dW.$$

Wenn $N(t)$ deterministisch wäre, dann wäre die linke Seite gleich $d(\ln N(t))$. Wir wendend daher die Itô-Formel auf diesen Ausdruck an, in der Hoffnung, den Korrekturterm später verrechnen zu können. Wir setzen also $u(x,t) = \ln x$ und haben dann $u_t = 0$. $u_x = \frac{1}{x}$ und $u_{xx} = \frac{-1}{x^2}$. Daher

$$d(\ln N(t)) = \frac{1}{N(t)}\,dN - \frac{1}{2N^2(t)}\alpha^2 N^2(t)\,dt.$$

Also

$$\frac{dN(t)}{N} = d(\ln N(t)) + \frac{\alpha^2}{2}\,dt.$$

Das setzen wir in die Differentialgleichung ein und erhalten

$$d(\ln N(t)) = \left(r - \frac{\alpha^2}{2}\right)dt + \alpha\,dW.$$

Das integrieren wir jetzt und erhalten

$$\ln\frac{N(t)}{N(0)} = \left(r - \frac{\alpha^2}{2}\right)t + \alpha W(t)$$

und damit schließlich

$$N(t) = N(0)e^{\left(r - \frac{\alpha^2}{2}\right)t + \alpha W(t)}$$

6.2 Definition. Ein Prozess der Form

$$X(t) = X(0)e^{\mu t + \alpha W(t)} \tag{6.2}$$

mit Konstanten μ und α ist eine *geometrische Brownsche Bewegung.*

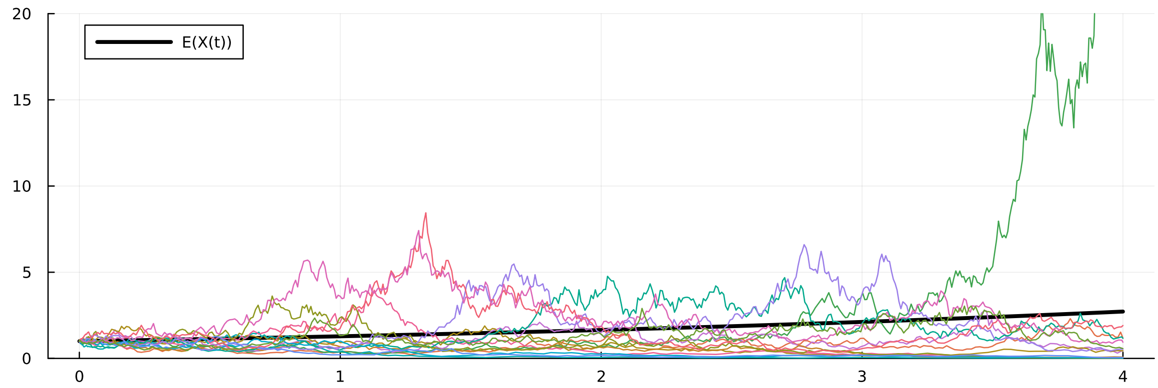

Abbildung 6.1: Pfade der geometrischen Brownschen Bewegung aus Beispiel 6.7

6.3 Satz. *Sei* $N(t)$ *wie in Beispiel 6.1. Ferner sei* $N(0)$ *unabhängig von* $(W(t))_{t\geq 0}$ *und* $\mathbb{E}(|N(0)|)$ *sei endlich. Dann*

$$\mathbb{E}(N(t)) = \mathbb{E}(N(0))\, e^{rt}.$$

Bemerkung. Wenn man das Stratonovich-Integral benutzt, dann erhält man für den Erwartungswert den Wert $e^{(r+\frac{1}{2}\alpha^2)t}$.

6.4 Korollar. *Der Erwartungswert der Lösung der stochastischen Differentialgleichung* (6.1) *wächst unbeschränkt, falls* $r > 0$*, und fällt gegen* 0*, falls* $r < 0$.

Man kann aber auch Aussagen über die Pfade machen, wenn man das folgende Resultat verwendet:

6.5 Theorem (Gesetz des iterierten Logarithmus, [Kle20], Satz 22.1).

$$\limsup_{t\to\infty} \frac{W(t)}{\sqrt{2t\ln\ln t}} = 1 \textit{ f. s.}$$

6.6 Satz. *Sei* $N(t)$ *eine Lösung von* (6.1) *mit* $N(0) > 0$ *f. s.*

(a) Falls $2r > \alpha^2$*, dann* $\lim_{t\to\infty} N(t) = \infty$ *f. s.*

(b) Falls $2r < \alpha^2$*, dann* $\lim_{t\to\infty} N(t) = 0$ *f. s.*

(c) Falls $2r = \alpha^2$*, dann* $\limsup_{t\to\infty} N(t) = \infty$ *f. s. und* $\liminf_{t\to\infty} N(t) = 0$ *f. s.*

6.7 Beispiel. Im Fall $0 < r < \frac{1}{2}\alpha^2$ haben wir also einen wachsenden, unbeschränkten Erwartungswert, während die Pfade fast sicher gegen 0 konvergieren.

Abbildung 6.1 zeigt einen solchen Prozesse mit $r = \frac{1}{4}$ und $\alpha = 1$.

6.8 Beispiel (Harmonischer Oszillator). Für $\mu \geq 0$ und $\omega_0 > 0$ ist

$$y'' + 2\mu y' + \omega_0^2 y = F$$

die Differentialgleichung des gedämpften harmonischen Oszillators (bzw. des ungedämpften, wenn $\mu = 0$). Wir gehen von einer verrauschten Zwangskraft der Form

$$F(t) = G(t) + \alpha W(t)$$

aus, wobei α eine Konstante und G eine deterministische Funktion ist. Wir führen einen zweidimensionalen Prozess $X = \binom{X_0}{X_1}$ ein und schreiben die Differentialgleichung als System

$$\begin{aligned} X_0' &= X_1 \\ X_1' &= -\omega_0^2 X_0 - 2\mu X_1 + G + \alpha W. \end{aligned}$$

In Matrixschreibweise

$$dX = AX\,dt + H\,dt + K\,dW,$$

wobei

$$A = \begin{pmatrix} 0 & 1 \\ -\omega_0^2 & -2\mu \end{pmatrix}, \quad H = \begin{pmatrix} 0 \\ G \end{pmatrix} \quad K = \begin{pmatrix} 0 \\ \alpha \end{pmatrix},$$

und W weiterhin ein eindimensionaler Prozess ist. Diese Gleichung multiplizieren wir mit dem Matrixexponential e^{-At} und erhalten

$$\exp(-At)\,dX - \exp(-At)AX\,dt = \exp(-At)H\,dt + \exp(-At)K\,dW.$$

Wir wollen nun $\exp(-At)X$ als Itô-Prozess schreiben. Dazu verwenden wir $u\colon \mathbb{R}^2 \times [0, \infty[\to \mathbb{R}^2$, $u(x,t) = \exp(-At)x$. Diese Abbildung ist linear in x, führt also nicht zu einem Korrekturterm in der Itô-Formel. Daher

$$d(\exp(-At)X) = -A\exp(-At)X\,dt + \exp(-At)\,dX$$

und schließlich

$$\exp(-At)X(t) - X(0) = \int_0^t \exp(-As)H(s)\,ds + \int_0^t \exp(-As)K(s)\,dW(s). \tag{6.3}$$

Das zweite Integral wollen wir partiell integrieren. Dazu setzen wir $\exp(-As) = \begin{pmatrix} a(s) & b(s) \\ c(s) & f(s) \end{pmatrix}$. Dann

$$\begin{aligned} \int_0^t \exp(-As)K\,dW &= \int_0^t \begin{pmatrix} a & b \\ c & f \end{pmatrix} \begin{pmatrix} 0 \\ \alpha \end{pmatrix} dW \\ &= \int_0^t \alpha \begin{pmatrix} b \\ f \end{pmatrix} dW \\ &= \alpha \begin{pmatrix} b(t) \\ f(t) \end{pmatrix} W(t) - \int_0^t \alpha \begin{pmatrix} b'(s) \\ f'(s) \end{pmatrix} W\,ds. \end{aligned}$$

Aus der Differentialgleichung des Matrixexponentials bekommen wir

$$\begin{pmatrix} b'(s) \\ f'(s) \end{pmatrix} = \frac{d}{dt} \exp(-As)K = -\exp(-As)AK.$$

Das setzen wir oben ein und erhalten

$$\int_0^t \exp(-As)K\,dW = \exp(-At)KW(t) + \int_0^t \exp(-As)AKW\,ds.$$

Schließlich setzen wir diese Formel in (6.3) ein und multiplizieren mit $\exp(At)$

$$X(t) = \exp(At)\left(X(0) + \exp(-At)KW(t) + \int_0^t \exp(-As)\left(H(s) + AKW(s)\right)\,ds\right).$$

Man beachte, dass man dieses Beispiel auch ohne das Itô-Integral direkt auf den Pfaden hätte rechnen können.

In Beispiel 6.1 musste man noch folgendes wissen:

6.9 Lemma. *Für* $\mu, \alpha \in \mathbb{R}$ *sei* $X(t) = \mu t + \alpha W(t)$. *Dann* $e^{X(t)} \in \mathbb{L}^2[0, T]$ *für jedes* T.

6.10 Beispiel. Für einen eindimensionalen Wienerschen Prozess $W(t)$ setzen wir

$$Y(t) = (\cos(W(t)), \sin(W(t))).$$

Wir zeigen zunächst, dass W ein Itô-Prozess im Sinne von 5.18 ist. Es gilt nämlich

$$\begin{aligned} d(\cos(W(t))) &= -\sin(W(t))\,dW - \frac{1}{2}\cos(W(t))\,dt, \\ d(\sin(W(t))) &= \cos(W(t))\,dW - \frac{1}{2}\sin(W(t))\,dt. \end{aligned}$$

Das bedeutet

$$\begin{aligned} dY_1 &= -\frac{1}{2}Y_1\,dt - Y_2\,dW, \\ dY_2 &= -\frac{1}{2}Y_2\,dt + Y_1\,dW, \end{aligned}$$

oder in Matrixnotation

$$dY = -\frac{1}{2}Y\,dt + KY\,dW \quad \text{mit } K = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right).$$

Y ist die *Brownsche Bewegung auf dem Kreisrand.*

6.11 Beispiel. Für eine stetige Funktion g sei die folgende Differentialgleichung gegeben

$$\begin{aligned} dX &= gX\,dW \\ X(0) &= 1. \end{aligned}$$

Wir gehen vor wie in Beispiel 6.1 und bestimmen zuerst

$$d(\ln X) = \frac{1}{X}\,dX - \frac{g^2}{2}\,dt = g\,dW - \frac{g^2}{2}\,dt.$$

Das führt wegen $X(0) = 1$ dann zu

$$\ln X(t) = \int_0^1 g(s)\,dW - \frac{1}{2}g^2(s)\,ds$$

und

$$X(t) = \exp\left(\int_0^t g(s)\,dW - \frac{1}{2}\int_0^t g^2(s)\,ds\right).$$

6.12 Bezeichnung. Wenn man eine stochastischen Differentialgleichung $dX = f(X)\,dt + g(X)\,dW$ formal durch dt teilt, dann erhält man eine Formel der Form

$$\frac{dX}{dt} = f(X) + g(X)\frac{dW}{dt},$$

bzw., wenn man die Zeitableitung mit dem Punkt schreibt

$$\dot{X} = f(X) + g(X)\dot{W}.$$

In [Øks03] wird am Anfang von § 3.1 auf einige Arbeiten verwiesen, in denen die Ableitung $\dot{W}$ im Distributionssinn präzisiert wird.

Die — bei uns nur formale — Größe $\dot{W}$ bezeichnet man als *weißes Rauschen.*

6.13 Beispiel (Langevinsche Gleichung). Seien $b > 0$ und $\sigma \in \mathbb{R}$ und sei ξ das weiße Rauschen. Die eindimensionale Gleichung

$$\dot{X} = -bX + \sigma\xi$$

ist die Gleichung von Langevin. Als stochastische Differentialgleichung geschrieben, lautet sie

$$dX = -bX\,dt + \sigma\,dW.$$

Die Interpretation ist, dass X die Geschwindigkeit eines Teilchens ist und $\sigma\xi$ die "fluktuierende Kraft" modelliert.

Man geht ähnlich vor wie beim harmonischen Oszillator und teilt durch die Lösung der zugehörigen deterministischen Gleichung, betrachtet also

$$d\left(e^{bt}X\right) = be^{bt}X\,dt + e^{bt}\,dX = be^{bt}X\,dt - be^{bt}X\,dt + \sigma e^{bt}\,dW = \sigma e^{bt}\,dW.$$

Die Lösung der Langevinschen Differentialgleichung ist also

$$X(t) = X(0)e^{-bt} + \sigma\int_0^t e^{(s-t)b}\,dW.$$

6.14 Satz. *Sei* $X(0)$ *unabhängig von* $\mathcal{W}(t)$ *und* $X(0)$ *besitze zweite Momente. Dann gelten für die Lösung* $X(t)$ *der Langevinschen Differentialgleichung*

$$\begin{aligned}\mathbb{E}(X(t)) &= e^{-bt}\mathbb{E}(X(0)),\\ \mathbb{E}(X^2(t)) &= e^{-2bt}\mathbb{E}(X^2(0)) + \frac{\sigma^2}{2b}\left(1-e^{-2bt}\right).\end{aligned}$$

6.15 Bemerkung. Für die Varianz bedeutet dies

$$\operatorname{Var}(X(t)) = \mathbb{E}(X^2(t)) - \mathbb{E}(X(t))^2 = e^{-2bt}\operatorname{Var}(X(0)) + \frac{\sigma^2}{2b}\left(1-e^{-2bt}\right).$$

Der Prozess kommt also nicht zur Ruhe.

6.16 Beispiel (Ornstein-Uhlenbeck Prozess). Der Ornstein-Uhlenbeck modelliert die Bewegung eines Teilchens, dessen Geschwindigkeit durch die Langevinsche Differentialgleichung beschrieben wird. Die Gleichung lautet

$$\begin{aligned}\ddot{Y} &= -b\dot{Y} + \sigma\xi,\\ Y(0) &= Y_0,\\ \dot{Y}(0) &= Y_1.\end{aligned}$$

Y_0 und Y_1 sind dabei unabhängig von $\mathcal{W}(t)$ und normalverteilt. Für $X = \dot{Y}$ gilt dann nach dem vorstehenden Beispiel

$$X(t) = e^{-bt}Y_1 + \sigma\int_0^t e^{b(s-t)}\,dW.$$

Daraus

$$Y(t) = Y_0 + \int_0^t X(s)\,ds.$$

und

$$\mathbb{E}(Y(t)) = \mathbb{E}(Y_0) + \int_0^t \mathbb{E}(X(s))\,ds = \mathbb{E}(y_0) + \int_0^t e^{-bs}\mathbb{E}(Y_1)\,ds = \mathbb{E}(Y_0) + \frac{1-e^{-bt}}{b}\mathbb{E}(Y_1).$$

In Example 6 von § 5.1 gibt Evans [Eva13] auch die Varianz an.

7 Existenz- und Eindeutigkeitssätze

7.1 Beispiel. Sei $b \in C^1(\mathbb{R})$ eine Funktion mit beschränkter Ableitung und sei $x_0 \in \mathbb{R}$. Dann besitzt die stochastische Differentialgleichung

$$\begin{aligned} dX &= b(X)\,dt + dW, \\ X(0) &= x_0 \end{aligned}$$

eine Lösung.

7.2 Lemma. *Sei X_0 eine von der Geschichte $(\mathcal{W}(t))_{t\geq 0}$ des Wienerschen Prozesses unabhängige Zufallsvariable. Für jedes t sei $\mathcal{H}(t)$ die kleinste σ-Algebra, die $\mathcal{W}(t)$ und $\mathcal{U}(X_0)$ umfasst. Dann ist $(\mathcal{H}(t))_{t\geq 0}$ eine Filtration, bzgl. derer der Wienersche Prozess ein Martingal ist.*

Wir benötigen diese Beobachtung, um im folgenden die Itô-Integrale überhaupt bilden zu können. In Friedman [Fri06] wird gezeigt, dass unter der Voraussetzung der Quadratintegrierbarkeit die Itô-Isomorphie weiterhin besteht.

7.3 Theorem (Existenz- und Eindeutigkeitssatz). *Sei W ein m-dimensionaler Wienerscher Prozess und seien $b\colon \mathbb{R}^n \times [0,T] \to \mathbb{R}^n$ und $B\colon \mathbb{R}^n \times [0,T] \to \mathbb{R}^{m\times n}$ stetig mit:*

(a) Für $0 \leq t \leq T$ und $x, y \in \mathbb{R}^n$

$$\begin{aligned} |b(x,t) - b(y,t)| &\leq L|x-y| \\ |B(x,t) - B(y,t)| &\leq L|x-y|. \end{aligned}$$

(b) Für $0 \leq t \leq T$ und $x \in \mathbb{R}^n$

$$\begin{aligned} |b(x,t)| &\leq L(1+|x|), \\ |B(x,t)| &\leq L(1+|x|). \end{aligned}$$

Ferner sei X_0 eine $\mathbb{R}^n$-wertige Zufallsvariable mit zweiten Momenten, welche unabhängig von der Geschichte $\mathcal{W}(t)$ des Wienerschen Prozesses ist.

Dann existiert eine eindeutige Lösung X der stochastischen Differentialgleichung

$$\begin{aligned} dX &= b(X,t)\,dt + B(X,t)\,dW \\ X(0) &= X_0. \end{aligned} \tag{7.1}$$

Diese Lösung ist adaptiert an die Filtration, die durch X_0 *und* $(\mathcal{W}(t))_{t\geq 0}$ *erzeugt wird, und erfüllt*

$$\mathbb{E}\left(\int_0^T |X(t)|^2 \, dt\right) < \infty.$$

Eindeutig bedeutet dabei folgendes: Wenn X und Y die stochastische Differentialgleichung lösen, dann

$$P(\{\omega \mid X(t,\omega) = Y(t,\omega) \text{ für } 0 \leq t \leq T\}) = 1.$$

Wir beweisen zuerst die Existenz ausgehend von derselben Idee wie zuvor:

Das Lemma von Gronwall wird in der Analysis II gezeigt. Es gibt verschiedene Versionen, hier ist die aus [Eva13].

7.4 Theorem (Lemma von Gronwall). *Sei* φ *und* f *stetige, nicht-negative Funktionen auf* $[0, T]$ *und sei* $C_0 \geq 0$ *konstant. Falls*

$$\varphi(t) \leq C_0 + \int_0^t f(s)\varphi(s) \, ds, \quad 0 \leq t \leq T,$$

dann

$$\varphi(t) \leq C_0 \exp\left(\int_0^t f(s) \, ds\right).$$

7.5 Satz. *Seien* b *und* B *wie in Theorem 7.3 und seien* X *und* Y *Lösungen von* (7.1), *allerdings mit* $Y(0) = Y_0$. *Dann existiert* $C \geq 0$, *so dass*

$$\mathbb{E}(|X(t) - Y(t)|^2) \leq 9\mathbb{E}(|X_0 - Y_0|^2)e^{Ct}. \tag{7.2}$$

7.6 Theorem ([Eva13], § 5.3, [Fri06], Theorem 2.3 in § 5). *Seien* b, B *und* X_0 *wie im Existenz- und Eindeutigkeitssatz 7.3 und sei* $p \in \mathbb{N} \setminus \{1\}$. *Falls* $\mathbb{E}(|X_0|^{2p}) < \infty$, *dann gibt es Konstanten* C_1 *und* C_2, *die nur von* T, L, m, n *und* p *abhängen, so dass für die Lösung* $X(t)$ *von* (7.1) *gilt*

$$\mathbb{E}(|X(t)|^{2p}) \leq C_2(1 + \mathbb{E}(|X_0|^{2p}))e^{C_1 t},$$
$$\mathbb{E}(|X(t) - X_0|^{2p}) \leq C_2(1 + \mathbb{E}(|X_0|^{2p}))t^p e^{C_1 t}.$$

7.7 Theorem (Stetigkeitssatz von Kolmogoroff [Eva13] 3.4.1, Beweis in Appendix D). *Sei* $(X(t))_{t\geq 0}$ *ein stochastischer Prozess mit fast sicher stetigen Pfaden, so dass es Konstanten* $\alpha, \beta, C > 0$ *gibt, so dass für alle* $s, t \geq 0$

$$\mathbb{E}(|X(t) - X(s)|^\beta) \leq C|t - s|^{1+\alpha}.$$

Sei ferner $0 < \gamma < \frac{\alpha}{\beta}$. *Dann gibt es für jedes* $T > 0$ *und fast jeden Pfad* ω *ein* K, *so dass*

$$|X(t,\omega) - X(s,\omega| \leq K|t - s|^\gamma, \quad 0 \leq s, t \leq T.$$

7.8 Korollar. *Unter den Voraussetzungen von Theorem 7.6 sind die Pfade von* X *fast sicher Hölder-stetig zu allen Exponenten* $\gamma < \frac{p-1}{2p}$.

Analog kann man die Hölder-Stetigkeit der Pfade des Wienerschen Prozesses zeigen, wenn man ihn auf eine andere Weise als in Abschnitt 3 konstruiert hat.

7.9 Theorem (Parameterabhängigkeit, [Eva13], § 5.3, [Fri06], Theorem 5.2 in §5). *Für* $k \in \mathbb{N}$ *seien* b^k, B^k *und* X_0^k *gegeben, die jeweils die Voraussetzungen von Theorem 7.3 erfüllen, und zwar alle mit demselben* L. *Ferner sollen gelten*

$$\lim_{k\to\infty} \mathbb{E}\big(|X_0^k - X_0|^2\big) = 0$$

und für jedes $M > 0$

$$\lim_{k\to\infty} \sup_{\substack{0\le t\le T\\ |x|\le M}} \big(|b^k(x,t) - b(x,t)| + |B^k(x,t) - B(x,t)|\big) = 0.$$

Schließlich gelte

$$\begin{aligned} dX^k &= b^k(X^k,t)\,dt + B^k(X^k,t)\,dW, \\ X^k(0) &= X_0^k. \end{aligned}$$

Dann gilt

$$\lim_{k\to\infty} \mathbb{E}\left(\sup_{0\le t\le T} |X^k(t) - X(t)|^2\right) = 0;$$

dabei ist X *die Lösung von*

$$\begin{aligned} dX &= b(X,t)\,dt + B(X,t)\,dW, \\ X(0) &= X_0. \end{aligned}$$

7.10 Beispiel. Für $x_0 \in \mathbb{R}$, b wie in Theorem 7.3 und $k \in \mathbb{N}$ sei X^k die Lösung von

$$\begin{aligned} dX^k &= b(X^k,t)\,dt + \frac{1}{k}\,dW, \\ X^k(0) &= x_0. \end{aligned}$$

Dann gibt es eine Teilfolge $(X^{k_j})_{j\in\mathbb{N}}$, so dass für jedes $T > 0$ die Pfade $X^{k_j}(\cdot,\omega)$ fast sicher gleichmäßig gegen die Lösung der gewöhnlichen Differentialgleichung

$$\begin{aligned} \dot{x} &= b(x), \\ x(0) &= x_0 \end{aligned}$$

konvergieren.

Wenn man auf die globale Lipschitz-Stetigkeit verzichtet, sieht vieles ganz anders aus, wie die beiden folgenden Ergebnisse von Scheutzow zeigen:

7.11 Beispiel (Scheutzow [Sch93], Example 1). Es gibt eine C^∞-Funktion $b\colon \mathbb{R}^2 \to \mathbb{R}^2$, so dass für jede Anfangsbedingung $x_0 \in \mathbb{R}$ die Lösung der gewöhnlichen Differentialgleichung

$$\dot{x} = b(x), \quad x(0) = x_0,$$

in endlicher Zeit explodiert, während für jedes $\epsilon > 0$ die Lösung der stochastischen Differentialgleichung

$$dX = b(X(t)) + \epsilon\, dW, \quad X(0) = x_0,$$

fast sicher für alle $t \geq 0$ erklärt ist.

7.12 Beispiel (Scheutzow [Sch93], Example 2). Es gibt eine C^∞-Funktion $b\colon \mathbb{R}^2 \to \mathbb{R}^2$, so dass für jedes $\epsilon > 0$ und jede Anfangsbedingung $x_0 \in \mathbb{R}$ die Lösung der stochastischen Differentialgleichung

$$dX = b(X(t)) + \epsilon\, dW, \quad X(0) = x_0,$$

fast sicher in endlicher Zeit explodiert, während die Lösung der gewöhnlichen Differentialgleichung

$$\dot{x} = b(x), \quad x(0) = 0,$$

für alle $t \geq 0$ erklärt ist.

7.13 Theorem. *Seien $c\colon [0,T] \to \mathbb{R}^n$ und $D, E\colon [0,T] \to \mathbb{R}^{n\times m}$ stetig und sei X_0 eine von $\mathcal{W}$ unabhängige Zufallsvariable. Dann besitzt die stochastische Differentialgleichung*

$$\begin{aligned} dX &= (c(t) + D(t)X)\, dt + E(t)\, dW, \\ X(0) &= X_0 \end{aligned}$$

die eindeutig bestimmt Lösung

$$X(t) = \Phi(t)\left(X_0 + \int_0^t \Phi(s)^{-1} c(s)\, ds + \int_0^t \Phi(s)^{-1} E(s)\, dW\right).$$

Dabei ist Φ die Lösung der gewöhnlichen Differentialgleichung

$$\begin{aligned} \dot{\Phi} &= D(t)\Phi \\ \Phi(0) &= \mathrm{id}\,. \end{aligned}$$

Bemerkung. Wenn D nicht von t abhängt, dann kann Φ mit den aus der Analysis II bekannten Mitteln bestimmt werden.

8 Schwache Lösungen

Wir betrachten weiterhin die stochastische Differentialgleichung (7.1)

$$\begin{aligned} dX &= b(X,t)\,dt + B(X,t)\,dW \\ X(0) &= X_0. \end{aligned}$$

Hierbei ist X_0 eine Zufallsvariable auf dem Wahrscheinlichkeitsraum $(\Omega, \mathcal{U}, P)$.

8.1 Definition. Sei $(\mathcal{H}^{X_0}(t))_{t\geq 0}$ die von $(\mathcal{W}(t))_{t\geq 0}$ und $\mathcal{U}(X_0)$ erzeugte Filtration. Eine $\mathcal{H}(t)$-adaptierte Zufallsvariable, welche die stochastische Differentialgleichung (7.1) löst, bezeichnet man als *starke Lösung*.

8.2 Definition. Eine *schwache Lösung* von (7.1) besteht aus

(a) einem Wahrscheinlichkeitsraum $(\widetilde{\Omega}, \widetilde{\mathcal{U}}, \widetilde{P})$,

(b) einer Filtration $(H(t))_{t\geq 0}$ auf $\widetilde{\Omega}$,

(c) einem Wienerschen Prozess $(\widetilde{W}(t))_{t\geq 0}$ auf $\widetilde{\Omega}$, die ein Martingal bzgl. $(\mathcal{H}(t))_{t\geq 0}$ ist,

(d) einem $\mathcal{H}(t)$-adaptierten Prozess $(\widetilde{X}(t))_{t\geq 0}$,

so dass

$$\begin{aligned} d\widetilde{X} &= b(\widetilde{X}, t)\,dt + B(\widetilde{X}, t)\,d\widetilde{W}, \\ \widetilde{P}(\widetilde{X}(0) \in B) &= P(X_0 \in B) \text{ für alle } B \in \mathcal{B}(\mathbb{R}^n). \end{aligned} \tag{8.1}$$

8.3 Bemerkung. (a) Unter den dort angegebene Bedingungen zeigt Theorem 7.3 die Existenz starker Lösungen.

(b) Da $(\widetilde{W}(t)_{t\geq 0})$ ein Martingal bzgl. $\mathcal{H}(t)$ ist, ist das Itô-Integral $\int_0^t B(\widetilde{X}, t)\,d\widetilde{W}$ erklärt.

(c) Es braucht kein Zusammenhang zu bestehen zwischen $(\Omega, \mathcal{U}, P)$ und $(\widetilde{\Omega}, \widetilde{\mathcal{U}}, \widetilde{P})$. Insbesondere ist X_0 keine Zufallsvariable auf $\widetilde{\Omega}$ und die Anfangsbedingung kann nur über die Verteilung formuliert werden.

Für das nächste Beispiel benötigen wir den folgenden Satz:

8.4 Satz ([Øks03], Theorem 8.4.2). *Sei W ein m-dimensionaler Wienerscher Prozess und sei B eine $(n \times m)$-Matrix von Funktionen, die in jedem $\mathbb{L}^2[0,T]$ liegen. Der durch*

$$dY = B(t)\,dW$$

gegebene Prozess hat genau dann die Verteilung eines n-dimensionalen Wienerschen Prozesses, wenn

$$G\,G^{\mathsf{T}} = \mathrm{id}_n \ \textit{f. s.}$$

8.5 Beispiel (Tanaka). Die stochastische Differentialgleichung

$$\begin{aligned} dX &= \mathrm{sign}(X(t))\,dW, \\ X(0) &= 0. \end{aligned}$$

besitzt keine starke, aber eine schwache Lösung.

8.6 Bemerkung. (a) Das Beispiel von Tanaka zeigt auch, dass die schwache Löung auch für den Wienerschen Prozess $(\widetilde{W}(t))_{t\geq 0}$ keine starke Lösung ist.

(b) Die Fage nach der pfadweisen Eindeutigkeit macht gar keinen Sinn, weil Differentialgleichung und Lösung i. a. auf verschiedenen Wahrscheinlichkeitsräumen definiert sind.

(c) Selbst wenn Differentialgleichung und Lösung auf demselben Wahrscheinlichkeitsraum erklärt sind, ist die pfadweise Eindeutigkeit nicht gegeben. Das zeigt die Differentialgleichung

$$dX = dW, \quad X(0) = 0,$$

welche im schwachen Sinn die beiden Lösungen $\pm W(t)$ besitzt.

8.7 Satz (Schwacher Eindeutigkeitssatz). *Es seien b, B und X_0 wie in Theorem 7.3. Ferner seien $(\hat{X}(t))_{t\geq 0}$ und $(\widetilde{Y}(t))_{t\geq 0}$ zwei Prozesse, definiert auf den Wahrscheinlichkeitsräumen $(\hat{\Omega}, \hat{\mathcal{U}}, \hat{P})$ und $(\widetilde{\Omega}, \widetilde{\mathcal{U}}, \widetilde{P})$, welche beide die stochastische Differentialgleichung (8.1) lösen. Dann besitzen $\hat{X}(t)$ und $\widetilde{Y}(t)$ für jedes t dieselbe Verteilung.*

9 Die Markow-Eigenschaft

9.1 Bezeichnung. Es seien b und B wie im Existenz- und Eindeutigkeitssatz 7.3, sei $x \in \mathbb{R}$ und sei $s \geq 0$. Die eindeutige Lösung der stochastischen Differentialgleichung

$$\begin{aligned} dX &= b(X(t), t)\, dX + B(X(t), t)\, dW, \\ X(s) &= x \end{aligned}$$

bezeichnen wir mit $X^{s,x}$.

Im Fall $s = 0$ schreiben wir X^x.

9.2 Definition. Die stochastische Differentialgleichung heißt *autonom,* wenn b und B nicht von t abhängen.

9.3 Bemerkung. Wenn die Differentialgleichung autonom ist, dann

$$\begin{aligned} X^{s,x}(s+h) &= x + \int_s^{s+h} b(X^{s,x}(u))\, du + \int_s^{s+h} B(X^{s,x}(u))\, dW(u) \\ &= x + \int_0^h b(X^{s,x}(v))\, dv + \int_0^h B(X^{s,x}(v))\, d\widetilde{W}(v), \end{aligned}$$

wobei $\widetilde{W}(v) = W(s+v) - W(s)$. Die zweite Substitution sieht man durch Betrachtung der elementaren Funktionen, denn $h(u_k)(W(u_{k+1}) - W(u_k)) = h(s+v_k)((W(s+v_{k+1}) - W(s)) - (W(s+v_k) - W(s))$ für $v_k = u_k - s$. Man überlegt sich leicht, dass $\widetilde{W}$ ebenfalls ein Wienerscher Prozess ist. Andererseits gilt

$$X^{0,x}(h) = x + \int_0^h b(X^{0,x}(v))\, dv + \int_0^t b(X^{0,x}(v))\, dW(v).$$

Weil W und $\widetilde{W}$ zwar beides Wienersche Prozesse sind, aber nicht dieselben, kann die Eindeutigkeitsaussage von Theorem 7.3 nicht angewandt werden. Wegen Satz 8.7 besitzen $X^{s,x}(s+h)$ und $X^x(h)$ aber dieselben Verteilungen.

9.4 Bezeichnung. Wir bezeichnen die Verteilung von $X^{0,x}$ mit Q^x und den zugehörigen Erwartungswert mit $\mathbb{E}^x$. In etwas formaler Schreibweise schreiben wir also

$$\mathbb{E}^x(f_1(X(t_1), \ldots, f_k(X(t_k)))) = \mathbb{E}(f_1(X^x(t_1), \ldots, f_k(X^x(t_k)))).$$

Die folgenden Aussagen und Beweise orientieren sich an § 5.3 von [Fri06].

9.5 Bezeichnung. Für $0 \le s \le t$, $x \in \mathbb{R}^n$ und $A \in \mathcal{B}(\mathbb{R}^n)$ setzen wir

$$p(s,x,t,A) = P(X^{s,x}(t) \in A).$$

9.6 Theorem (Markow-Eigenschaft). *Seien* b*,* B *und* X_0 *wie beim Existenz- und Eindeutigkeitssatz 7.3 und sei* X *die Lösung von*

$$\begin{aligned} dX &= b(X(t),t)\,dt + B(X(t),t)\,dW, \\ X(0) &= X_0. \end{aligned}$$

Für $t \ge s$ *bezeichnen wir mit* $\mathcal{H}(t)$ *die von* X_0 *und den* $W(s)$*,* $0 \le s \le t$*, erzeugte* σ*-Algebra. Dann gilt für jede Borelmenge* A *und* $s \le t$

$$P(X(t) \in A \mid \mathcal{H}(s)) = P(X(t) \in A \mid X(s)) = p(s, X(s), t, A).$$

9.7 Bemerkung. (a) Im Beweis wurde für stetige, beschränkte Funktionen f gezeigt

$$\mathbb{E}(f(X(t)) \mid \mathcal{H}(s)) = \mathbb{E}(f(X^{s,x}(t)))\big|_{x=X(s)}. \tag{9.1}$$

Durch Approximation gilt diese Aussage auch für beschränktes, messbares f.

(b) Wenn das System *autonom* ist, dann kann (9.1) auch wie folgt geschrieben werden

$$\mathbb{E}(f(X(s+h)) \mid \mathcal{H}(s)) = \mathbb{E}(f(X^{x}(h)))\big|_{x=X(s)}. \tag{9.2}$$

Das ist Theorem 7.1.2 aus [Øks03].

(c) Mit Theorem 9.6 ist auch Theorem 3.32 gezeigt, denn der Wienersche Prozess löst die stochastische Differentialgleichung $dX = dW$, $X(0) = 0$, und wir hatten die Markow-Eigenschaft des Wienerschen Prozesses beim Beweis von Theorem 9.6 nicht verwendet.

9.8 Definition ([Fri06], § 2.1). Eine *Markowsche Übergangswahrscheinlichkeit* ist eine Abbildung $p(s,x,t,A)$, wobei $0 \le s \le t$, $x \in \mathbb{R}^n$ und $A \in \mathcal{B}(\mathbb{R}^n)$, welche folgende Eigenschaften besitzt:

(a) $p(s,\cdot,t,A)$ ist borelmessbar,

(b) $p(s,x,t,\cdot)$ ist ein Wahrscheinlichkeitsmaß, manchmal auch geschrieben als $P_{s,x,t}$,

(c) p erfüllt die Chapman-Kolmogoroff-Gleichung

$$p(s,x,t,A) = \int_{\mathbb{R}^n} p(\lambda,y,t,A)\,dP_{s,x,\lambda}(y), \qquad s < \lambda < t.$$

9.9 Bemerkung. (a) Häufig findet wird auch gefordert, dass $P_{s,x,t}$ eine Verteilungsdichte besitzt. Dann meint man mit $p(s,x,t,y)$ diese Verteilungsdichte. Die Verteilungsdichte wird nicht von allen Autoren in derselben Form hingeschrieben.

(b) Friedman [Fri06] schreibt $p(s, x, t, dy)$ für $dP_{s,x,t}(y)$.

(c) In Friedman [Fri06], Theorem 2.1.1, wird gezeigt, dass es zu jeder Markowschen Übergangswahrscheinlichkeit einen Markowschen Prozess gibt, der die zu der gegebenen Übergangswahrscheinlichkeit gehörende Chapman-Kolmogoroff-Gleichung erfüllt.

9.10 Satz. *Das* p *aus Bezeichnung 9.5 ist eine Markowsche Übergangswahrscheinlichkeit.*

9.11 Definition. Ein 5-Tupel $\big(\Omega, \mathcal{U}, (\mathcal{U}_t^s)_{s\le t}, x, (P_{x,s})_{x,s}\big)$ ist ein *Markowscher Prozess* mit Übergangswahrscheinlichkeit p, wenn

(a) $(\Omega, \mathcal{U})$ ist ein Messraum, für $s \le t$ ist $\mathcal{U}_t^s$ eine σ-Algebra, wobei für $s' \le s$ und $t \le t'$ gilt $\mathcal{U}_t^s \subseteq \mathcal{U}_{t'}^{s'}$, und $\mathcal{U}$ umfasst alle $\mathcal{U}_t^s$.

(b) x ist eine Funktion $[0, \infty[\times \Omega \to \mathbb{R}^n$, die bzgl. jedes $\mathcal{U}_t^s$ messbar ist.

(c) Für $x \in \mathbb{R}^n$ und $s \ge 0$ ist $P_{x,s}$ ein Wahrscheinlichkeitsmaß auf dem Messraum $(\Omega, \mathcal{U}_\infty^s)$, so dass

$$P_{x,s}\,(x(s) = x) = 1,$$
$$P_{x,s}\,(x(t+h,\omega) \in A \mid \mathcal{U}_t^s) = p(t, x(t,\omega), t+h, A) \text{ f. s.}$$

9.12 Bemerkung. Friedman [Fri06] schreibt den Lösungsprozess einer stochastischen Differentialgleichung in § 5.3 mehr oder weniger konkret hin. Wie bereits früher vereinbart, werden Versionen der Lösungen mit stetigen Pfaden gewählt.

Dann setzt er $\mathcal{C} = C([0, \infty[, \mathbb{R}^n)$ und definiert $\mathcal{U}_t^s$ für $s \le t \le \infty$ als die kleinste σ-Algebra, welche für jedes $A \in \mathcal{B}(\mathbb{R}^n)$ und jedes $u \in [s, t]$ die Menge

$$\{f \in \Omega \mid \forall u \in [s,t] : f(u) \in A\}$$

enthält.

Ferner definiert er einen Prozess X auf Ω durch

$$X(t, f) = f(t).$$

Es muss gezeigt werden, dass $X(t)$ messbar bzgl. $\mathcal{U}_t^s$ ist. Dazu sei $A \in \mathcal{B}(\mathbb{R}^n)$. Dann $X(t)^{-1}(A) = \{f \in \Omega \mid f(t) \in A\} \in \mathcal{U}_t^t \subseteq \mathcal{U}_t^s$.

Für $x \in \mathbb{R}^n$, $0 \le s \le t$ und $B \in \mathcal{U}_t^s$ erklärt er

$$P_{x,s}(B) = P(X^{s,x} \in B);$$

wenn z. B. B von der Form $\{f \in \Omega \mid \forall u \in [s,t] : f(u) \in A\}$ ist, dann bedeutet $X^{s,x} \in B$, dass $X^{s,x}(u) \in A$ für $s \le u \le t$. Also

$$P_{x,s}(X(s) = x) = P_{x,s}(\{f \mid f(s) \in \{x\}\}) = P(X^{s,x}(s) \in \{x\}) = 1.$$

Dass dieses $P_{x,s}$ ein Wahrscheinlichkeitsmaß ist, ist klar.

Nun ist zu zeigen

$$P_{x,s}\,(X(t+h) \in A \mid \mathcal{U}_t^s) = p(t, X(t), t+h, A). \tag{9.3}$$

Wenn wir mit $\mathcal{H}_{s,x}$ die Geschichte von $X^{s,x}$ bezeichnen, dann haben wir wegen Theorem 9.6

$$P(X^{s,x}(t+h) \in A \mid \mathcal{H}_{s,x}(t)) = p(t, X^{s,x}(t), t+h, A).$$

Seien nun $s < t_1 < \cdots < t_m \leq t$ und seien $A_1, \ldots, A_m \in \mathcal{B}(\mathbb{R}^n)$. Wegen $D := \bigcap_{j=1}^m \{X^{s,x}(t_j) \in A_j\} \in \mathcal{H}_{s,x}(t)$ und der Definition der bedingten Erwartung gilt

$$\begin{aligned}
&P(X^{s,x}(t+h) \in A, X^{s,x}(t_1) \in A_1, \ldots, X^{s,x}(t_m) \in A_m) \\
&\quad = \int_D \chi_{\{X^{s,x}(t+h)\in A\}}\, dP \\
&\quad = \int_D \mathbb{E}\big(\chi_{\{X^{s,x}(t+h)\in A\}} \mid \mathcal{H}_{s,x}(t)\big)\; dP \\
&\quad = \int_D p(t, X^{s,x}(t), t+h, A)\, dP \\
&\quad = \int \chi_{A_1}(X^{s,x}(t_1)) \cdots \chi_{A_m}(X^{s,x}(t_m)) p(t, X^{s,x}(t), t+h, A)\, dP.
\end{aligned}$$

Um (9.3) zu zeigen, muss man für jedes $B \in \mathcal{U}_t^s$ folgendes zeigen

$$P_{s,x}(\{X(t+h) \in A\} \cap B) = \int_B p(t, X(t), t+h, A)\, dP_{s,x}.$$

Wir machen das nur für ein Erzeugendensystem, also nur für B von der Form $\{X(t_1) \in A_1, \ldots, X(t_m) \in A_m\}$. Aus dem, was wir gerade ausgerechnet haben folgt

$$\begin{aligned}
P_{s,x}(\{X(t+h) \in A\} \cap B) &= \int \chi_{A_1}(X^{s,x}(t_1)) \cdots \chi_{A_m}(X^{s,x}(t_m)) p(t, X^{s,x}(t), t+h, A)\, dP \\
&= \int \chi_{A_1}(X(t_1)) \cdots \chi_{A_m}(X(t_m)) p(t, X(t), t+h, A)\, dP_{s,x} \\
&= \int_B p(t, X(t), t+h, A)\, dP_{s,x}.
\end{aligned}$$

Damit haben wir gezeigt, dass $(\mathcal{C}, \mathcal{U}, (\mathcal{U}_t^s)_{s \leq t}, X, (P_{s,x})_{s,x})$ ein Markowscher Prozess mit Übergangswahrscheinlichkeit $p(s, x, t, A)$ ist.

10 Stoppzeiten

10.1 Definition. Eine Zufallsvariable $\tau \colon \Omega \to [0, \infty]$ ist eine *Stoppzeit* bzgl. einer Filtration $\mathcal{F}$, wenn

$$\{\tau \leq t\} \in \mathcal{F}(t) \text{ für jedes } t \geq 0.$$

Wenn die Filtration nicht genannt wird, handelt es sich um die Geschichte des Wienerschen Prozesses.

Beispielsweise ist jede konstante Zufallsvariable mit Werten in $[0, \infty]$ eine Stoppzeit.

10.2 Satz. *Seien τ_1 und τ_2 Stoppzeiten bzgl. $\mathcal{F}$.*

(a) Dann gelten $\{\tau < t\} \in \mathcal{F}(t)$ und $\{\tau = t\} \in \mathcal{F}(t)$ für jedes t.

(b) $\min(\tau_1, \tau_2)$ und $\max(\tau_1, \tau_2)$ sind Stoppzeiten.

10.3 Satz. *Es sei $(X(t))_{t \geq 0}$ ein an die Filtration $(\mathcal{F}(t))_{t \geq 0}$ adaptierter Prozess mit stetigen Pfaden. Wenn E eine abgeschlossene Menge ist, dann ist*

$$\tau = \inf\{t \geq 0 \mid X(t) \in E\}$$

eine Stoppzeit bzgl. der Filtration $(\mathcal{F}(t))_{t \geq 0}$.

10.4 Bemerkung. Die Stoppzeit aus Satz 10.3 bezeichnet man als *Eintrittszeit* (engl. *hitting time*). Wenn E offen ist, dann ist offenbar auch

$$\tau = \inf\{t \geq 0 \mid X(t) \notin E\}$$

eine Stoppzeit. Man bezeichnet sie als *Austrittszeit* (engl. *exit time*). Mit Austrittszeit ist immer die Zeit des erstmaligen Verlassens gemeint. Die Zeit des letztmaligen Verlassens ist i. a. keine Stoppzeit.

10.5 Definition. Sei $G \in \mathbb{L}^2[0, T]$ und sei τ eine Stoppzeit mit $0 \leq \tau \leq T$. Dann setzen wir

$$\int_0^\tau G \, dW = \int_0^T \chi_{t \leq \tau} G \, dW.$$

Bemerkung. Weil $\min(\tau, T)$ eine Stoppzeit ist, ist die Forderung $0 \leq \tau \leq T$ keine Einschränkung.

10.6 Satz. *Sei* $G \in \mathbb{L}^2[0,T]$ *und sei* $0 \le \tau \le T$ *eine Stoppzeit. Dann*

$$\mathbb{E}\left(\int_0^\tau G\,dW\right) = 0,$$

$$\mathbb{E}\left(\left(\int_0^\tau G\,dW\right)^2\right) = \mathbb{E}\left(\int_0^\tau G^2\,dt\right).$$

10.7 Definition. (a) Mit $\mathcal{M}(\mathbb{R}^n)$ bezeichnen wir den Raum der beschränkten, borelmessbaren Funktion $f\colon \mathbb{R}^n \to \mathbb{R}$. Er wird versehen mit der Supremumsnorm.

(b) Es sei p die Übergangswahrscheinlichkeit eines Markowschen Prozesses. Für $0 \le s < t < \infty$ setzen wir

$$T_{s,t}\colon \mathcal{M}(\mathbb{R}^n) \to M(\mathbb{R}^n), \quad T_{s,t}(f)(x) = \int_{\mathbb{R}_n} f\,dP_{s,x,t}.$$

Ferner setzen wir $T_{t,t} = \mathrm{id}$.

10.8 Bemerkung. (a) $\|T_{s,t}\| = 1$.

(b) Für $0 \le s < t < u$ gilt

$$T_{s,t}T_{t,u} = T_{s,u}.$$

(c) Man bezeichnet $(T_{s,t})_{s<t}$ als die zu dem Markowschen Prozess *assoziierte Halbgruppe*.

10.9 Definition. (a) Ein Markowscher Prozess hat die starke Markow-Eigenschaft, wenn für jede Stoppzeit τ gilt

$$P_{x,s}\left(x(t+\tau) \in A \mid \mathcal{U}_\tau^s\right) = p(\tau, x(\tau), t+\tau, A) \text{ f. s.}$$

(b) Wenn für jede stetige, beschränkte Funktion $f\colon \mathbb{R}^n \to \mathbb{R}$ und jedes $\lambda > 0$ die Abbildung

$$(t,z) \mapsto \int_{\mathbb{R}^n} fP_{t,z,t+\lambda}$$

stetig ist, dann besitzt der Markowsche Prozess die *Feller-Eigenschaft.*

10.10 Theorem ([Fri06], Cor. 2.2.6). *Sei* $\left(\Omega, \mathcal{U}, (\mathcal{U}_t^s)_{s\le t}, x, (P_{x,s})_{x,s}\right)$ *ein Markowscher Prozess mit f. s. stetigen Pfaden, welcher die Feller Eigenschaft besitzt. Dann besitzt der Prozess die starke Markow-Eigenschaft.*

10.11 Lemma. *Es gelten die Voraussetzungen des Existenz- und Eindeutigkeitssatzes 7.3. Zu jeder Wahl von* $R, T > 0$ *gibt es* $C > 0$*, so dass*

$$\mathbb{E}\left(\sup_{\tau\le t\le T} |X^{s,x}(t) - X^{\tau,y}(t)|^2\right) \le C\left(|x-y|^2 + |s-\tau|\right),$$

falls $|x|, |y| \le R$ *und* $0 \le s \le \tau \le T$.

10.12 Theorem ([Fri06], Thm. 5.3.4). *Der in Bemerkung 9.12 konstruierte Markowsche Prozess besitzt die Feller-Eigenschaft und damit auch die starke Markow-Eigenschaft.*

Der folgende Satz mit der zugehörigen Definition gehört thematisch zur Martingalungleichung.

10.13 Definition. Eine Folge $(X_n)_{n\in\mathbb{N}}$ von Zufallsvariablen *konvergiert dem Maße nach* gegen eine Zufallsvariable X, wenn für jedes $\epsilon > 0$

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

Wir schreiben $X_n \xrightarrow{P} X$.

10.14 Bemerkung. (a) L^1-Konvergenz, also Konvergenz im Mittel, impliziert Konvergenz dem Maße nach. Die Umkehrung gilt i. a. nicht.

(b) Der Grenzwert dem Maße nach ist eindeutig bis auf eine Nullfunktion.

10.15 Theorem. *Sei* $G \in \mathbb{L}^2[0,T]$ *und seien* $\epsilon, N > 0$. *Dann*

$$P\left(\sup_{0\le t\le T}\left|\int_0^t G(s)\,dW\right| > \epsilon\right) \le P\left(\int_0^T G^2(s)\,ds > N\right) + \frac{N}{\epsilon^2}.$$

10.16 Theorem. *Seien* G_n *und* G *in* $\mathbb{L}^2[0,T]$ *und sei* $\int_0^T |G_n(t) - G(t)|^2\,dt \xrightarrow{P} 0$. *Dann*

$$\sup_{0\le t\le T}\left|\int_0^t G_n\,dW - \int_0^t G\,dW\right| \xrightarrow{P} 0.$$

10.17 Definition. Eine Folge $(X_n)_{n\in\mathbb{N}}$ von Zufallsvariablen heißt *gleichmäßig integrierbar,* wenn $\mathbb{E}(|X_n|) < \infty$ für jedes n und

$$\lim_{\lambda\to\infty}\limsup_{n\to\infty}\int_{|X_n|>\lambda} |X_n|\,dP = 0.$$

10.18 Theorem ([Fri06], Lemma 1.3.6). *Wenn* $(X_n)_{n\in\mathbb{N}}$ *gleichmäßig integrierbar ist, dann* $\limsup_{n\to\infty}\mathbb{E}(|X_n|) < \infty$. *Wenn außerdem noch* $X_n \to X$ *f. s. oder im Maße, dann*

$$\mathbb{E}(|X_n - X|) \to 0 \quad \textit{und} \quad \mathbb{E}(X_n) \to \mathbb{E}(X).$$

10.19 Satz. *Sei* $G \in \mathbb{L}^2[0,T]$, *sei* $0 \le \tau \le T$ *eine Stoppzeit und sei* $X(t) = \int_0^t G\,dW$. *Dann gilt*

$$X(\tau) = \int_0^\tau G(s)\,dW(s). \tag{10.1}$$

.

11 Zeithomogene Markowsche Prozesse

Es sei $(\Omega, \mathcal{U}, (\mathcal{U}_t^s)_{s\leq t}, X, (P_{s,x})_{x,s})$ ein Markowscher Prozess im Sinne von Definition 9.11.

11.1 Definition. Eine Übergangswahrscheinlichkeit p heißt *stationär,* wenn

$$p(s, x, t, A) = p(0, x, t - s, A).$$

Einen Markowscher Prozess mit stationärer Übergangswahrscheinlichkeit bezeichnet man als *zeithomogen.*

11.2 Bezeichnung. In diesem Fall lässt man das Argument s in der Regel weg und schreibt

$$P_s = P_{x,0}, \quad \mathcal{U}_t = \mathcal{U}_t^0, \quad p(t, x, A) = p(0, x, t, A).$$

Wir schreiben auch, anders als Friedman, $\widetilde{P}_{t,x}(A) = p(t, x, A)$, falls wir nach $p(t, x, \cdot)$ integrieren müssen.

11.3 Bemerkung. Gegeben sei eine autonome stochastische Differentialgleichungen. In Bemerkung 9.3 hatten wir gesehen, dass der durch die Lösung gegebene Markowsche Prozess zeithomogen ist. In (9.2) hatten wir gesehen, dass

$$\mathbb{E}(f(X(s + h)) \mid \mathcal{H}(s)) = \mathbb{E}(f(X^x(h)))\big|_{x=X(s)}.$$

11.4 Bezeichnung. Die zugehörige Halbgruppe $T_t \colon \mathcal{M}(\mathbb{R}^n) \to \mathcal{M}(\mathbb{R}^n)$ wird jetzt wie folgt geschrieben:

$$(T_t f)(x) = \int_{\mathbb{R}^n} f(y)\, d\widetilde{P}_{t,x}(y).$$

11.5 Beispiel. Was bedeutet das, wenn der Markowsche Prozess aus den Lösungen einer autonomen stochastischen Differentialgleichung besteht?

Im Beweis von Satz 1.12 hatten wir gesehen, dass für $\psi \in \mathcal{M}(\mathbb{R}^n)$ gilt

$$\int \psi \, d\widetilde{P}_{t,x} = \int \psi(X^x(t))\, dP.$$

Also

$$(T_t f)(x) = \int f(X^x(t))\, dP.$$

11.6 Bezeichnung. Zu einer gegebenen Halbgruppe wid mit $\mathcal{M}_0(\mathbb{R}^n)$ wird der Raum derjenigen Elemente $f \in \mathcal{M}(\mathbb{R}^n)$ bezeichnet, für den

$$\|T_t f - f\|_\infty \to 0 \text{ für } t \searrow 0,$$

wobei

$$\|g\|_\infty = \inf\{C > 0 \mid \lambda_n(|g| \geq C) = 0\}$$

das *wesentliche Supremum* von g bezeichnet.

11.7 Definition. Der *Erzeuger* eines zeithomogenen Markowschen Prozesses ist gegeben durch

$$\mathscr{P} f = \lim_{t \searrow 0} \frac{T_t f - f}{t}, \quad f \in D(\mathscr{P}),$$

wobei $D(\mathscr{P})$ aus denjenigen $f \in \mathcal{M}(\mathbb{R}^n)$ besteht, für die der Grenzwert existiert.

12 Itô-Diffusion

12.1 Definition. Ein n-dimensionler Markow-Prozess mit stetigen Pfaden und stationärer Übergangswahrscheinlichkeit $p(h, x, A)$ ist ein *Diffusionsprozess,* wenn

(a) für alle Wahlen von $\epsilon > 0$ und $x \in \mathbb{R}^n$

$$\lim_{h \searrow 0} \frac{1}{h} p(h, x, \mathbb{R}^n \setminus \overline{B}_\epsilon(x)) = 0,$$

(b) es gibt Abbildungen $b \colon \mathbb{R}^n \to \mathbb{R}^n$ und $a \colon \mathbb{R}^n \to \mathbb{R}^{n \times n}$, so dass für alle $\epsilon > 0$ und $x \in \mathbb{R}^n$

$$\lim_{h \searrow 0} \frac{1}{h} \int_{|y-x|<\epsilon} (y_i - x_i) \, d\widetilde{P}_{h,x}(y) = b_i(x), \qquad i = 1, \dots, n,$$
$$\lim_{h \searrow 0} \frac{1}{h} \int_{|y-x|<\epsilon} (y_i - x_i)(y_j - x_j) \, d\widetilde{P}_{h,x}(y) = a_{i,j}(x), \qquad i, j = 1, \dots, n.$$

Dann bezeichnet man b als *Drift* und a als *Diffusionsmatrix* des Prozesses.

12.2 Lemma. *Die folgenden beiden Bedingungen implizieren die Bedingungen (a) und (b) aus der Definition:*

(a) Es gibt* $\delta > 0$ *und* $x \in \mathbb{R}^n$*, so dass*

$$\lim_{h \searrow 0} \frac{1}{h} \int_{\mathbb{R}^n} |x - y|^{2+\delta} \, d\widetilde{P}_{h,x}(y) = 0,$$

(b) Für alle* $x \in \mathbb{R}^n$

$$\lim_{h \searrow 0} \frac{1}{h} \int_{\mathbb{R}^n} (y_i - x_i) \, d\widetilde{P}_{x,t}(y) = b_i(x), \qquad i = 1, \dots, n,$$
$$\lim_{h \searrow 0} \frac{1}{h} \int_{\mathbb{R}^n} (y_i - x_i)(y_j - x_j) \, d\widetilde{P}_{x,t}(y) = a_{i,j}(x), \qquad i, j = 1, \dots, n.$$

12.3 Theorem. *Es gelten die Voraussetzungen des Existenz- und Eindeutigkeitssatzes 7.3 mit zeitunbhängigen* b *und* B*. Dann ist der gemäß Definition 9.11 konstruierte zeithomogene Markowsche Prozess eine Itô-Diffusion mit Drift* b *und Diffusionsmatrix* $a = BB^T$*.*

Bemerkung. In [Fri06] wird das alles auch im zeitinhomogenen Fall gemacht.

Wie in Definition 11.7 bezeichnen wir mit $\mathscr{P}$ den Erzeuger der Halbgruppe und mit $D(\mathscr{P})$ seinen Definitionsbereich.

12.4 Theorem. *Jede beschränkte Funktion* $f\colon \mathbb{R}^n \to \mathbb{R}$ *von der Klasse* C^2 *liegt in* $D(\mathscr{P})$ *und es gilt*

$$\mathscr{P}(f)(x) = \frac{1}{2}\sum_{i,j=1}^{n} a_{i,j}(x)\frac{\partial^2 f}{\partial x_i \partial x_j} + \sum_{i,j=1}^{n} b_i(x)\frac{\partial f}{\partial x_i}.$$

Literatur

[AB99] Charalambos D. Aliprantis und Kim C. Border. *Infinite-dimensional analysis*. Second. A hitchhiker's guide. Springer-Verlag, Berlin, 1999, S. xx+672. ISBN: 3-540-65854-8. DOI: 10.1007/978-3-662-03961-8. URL: https://doi.org/10.1007/978-3-662-03961-8.

[Bre68] Leo Breiman. *Probability*. Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont., 1968, S. ix+421.

[Eva13] Lawrence C. Evans. *An introduction to stochastic differential equations*. American Mathematical Society, Providence, RI, 2013, S. viii+151. ISBN: 978-1-4704-1054-4. DOI: 10.1090/mbk/082. URL: https://doi.org/10.1090/mbk/082.

[Fri06] Avner Friedman. *Stochastic differential equations and applications*. Two volumes bound as one, Reprint of the 1975 and 1976 original published in two volumes. Dover Publications, Inc., Mineola, NY, 2006, S. xvi+531. ISBN: 0-486-45359-6.

[Kab11] Winfried Kaballo. *Grundkurs Funktionalanalysis*. German. Heidelberg: Spektrum Akademischer Verlag, 2011, S. xii + 345. ISBN: 978-3-8274-2149-4/pbk; 978-3-8274-2721-2/ebook. DOI: 10.1007/978-3-8274-2721-2.

[Kle20] Achim Klenke. *Wahrscheinlichkeitstheorie*. German. 4th revised and supplemented edition. Masterclass. Berlin: Springer Spektrum, 2020. ISBN: 978-3-662-62088-5; 978-3-662-62089-2. DOI: 10.1007/978-3-662-62089-2.

[MV11] Reinhold Meise und Dietmar Vogt. *Einführung in die Funktionalanalysis*. German. 2nd revised ed. Wiesbaden: Vieweg+Teubner, 2011, S. x + 273. ISBN: 978-3-8348-1872-0/pbk.

[Øks03] Bernt Øksendal. *Stochastic differential equations*. Sixth. Universitext. An introduction with applications. Springer-Verlag, Berlin, 2003, S. xxiv+360. ISBN: 3-540-04758-1. DOI: 10.1007/978-3-642-14394-6. URL: https://doi.org/10.1007/978-3-642-14394-6.

[Sch93] M. Scheutzow. „Stabilization and destabilization by noise in the plane". In: *Stochastic Anal. Appl.* 11.1 (1993), S. 97–113. ISSN: 0736-2994,1532-9356. DOI: 10.1080/07362999308809304. URL: https://doi.org/10.1080/07362999308809304.

[Sim15] Barry Simon. *Real analysis*. Bd. Part 1. A Comprehensive Course in Analysis. With a 68 page companion booklet. American Mathematical Society, Providence, RI, 2015, S. xx+789. ISBN: 978-1-4704-1099-5. DOI: 10.1090/simon/001. URL: https://doi.org/10.1090/simon/001.

[Woj97] P. Wojtaszczyk. *A mathematical introduction to wavelets*. Bd. 37. London Mathematical Society Student Texts. Cambridge University Press, Cambridge, 1997, S. xii+261. ISBN: 0-521-57020-4; 0-521-57894-9. DOI: 10.1017/CBO9780511623790. URL: https://doi.org/10.1017/CBO9780511623790.