

**Stochastische Differentialgleichungen:
Theorie, Numerik und Anwendungen**

Roland Pulch

Skript zur Vorlesung im Sommersemester 2023

Universität Greifswald

Institut für Mathematik und Informatik

Literaturempfehlungen:

- B. Øksendal:
Stochastic Differential Equations.
(6. Aufl.) Springer, Berlin 2010 (380 Seiten).
- M. Günther, A. Jüngel:
Finanzderivate mit MATLAB.
(2. Aufl.) Vieweg+Teubner, Wiesbaden 2010 (352 Seiten).
- D.J. Higham:
An Algorithmic Introduction to Numerical Simulation of Stochastic Differential Equations.
SIAM Review, Vol. 43, No. 3 (2001) S. 525-546 (22 Seiten).
- D.J. Higham, P.E. Kloeden:
An Introduction to the Numerical Simulation of Stochastic Differential Equations.
SIAM, 2021 (293 Seiten).
- P.E. Kloeden, E. Platen:
Numerical solution of stochastic differential equations.
(2. Aufl.) Springer, Berlin 1995 (632 Seiten).

Inhaltsverzeichnis

1	Einleitung	5
2	Zufallsprozesse	9
2.1	Grundlagen aus der Wahrscheinlichkeitsrechnung	9
2.2	Definition eines Zufallsprozesses	11
2.3	Wiener Prozess bzw. Brownsche Bewegung	12
3	Stochastische Integrale	22
3.1	Deterministische Integralbegriffe	22
3.2	Integrale mit Zufallsprozessen	24
3.3	Konstruktion des Itô-Integrals	26
3.4	Eigenschaften des Itô-Integrals	34
3.5	Verallgemeinerungen des Itô-Integrals	36
3.6	Stratonovich-Integral	39
4	Itô-Prozesse	42
4.1	Eindimensionale Itô-Prozesse und Itô-Formel	42
4.2	Mehrdimensionale Itô-Prozesse und Itô-Formel	47
4.3	Lineare Differentialgleichungen	49
4.4	Existenz und Eindeutigkeit	54
5	Numerische Verfahren	57
5.1	Euler-Maruyama-Verfahren	57
5.2	Starke und Schwache Konvergenz	60
5.3	Stochastische Taylor-Entwicklungen	63
5.4	Milstein-Verfahren	65

5.5	Verfahren höherer Ordnung	68
5.6	Simulationsbeispiel: SIR-Modell	70
6	Anwendungen in der Finanzmathematik	73
6.1	Modell für Aktienkurse	73
6.2	Optionspreise	75
6.3	Erweiterte Modelle für Aktienkurse	83
6.4	Optionen auf mehrere Aktien	89
7	Stochastische Resonanz	95
7.1	Klimamodell	95
7.2	Stochastische Periodizität	103

1 Einleitung

In diesem Abschnitt wird eine kurze Motivation zur Modellierung mit stochastischen Differentialgleichungen gegeben.

Stochastische Differentialgleichungen

Zu klassischen Differentialgleichungen existieren häufig stochastische Analogien. Beispielsweise werden Wachstumsprozesse beschrieben durch die gewöhnliche Differentialgleichung

$$\frac{dN(t)}{dt} = a(t)N(t) \quad (1.1)$$

für die Populationsgröße $N : [0, \infty) \rightarrow \mathbb{R}$ mit einer gegebenen Wachstumsrate $a : [0, \infty) \rightarrow \mathbb{R}$. Typischerweise werden Anfangswerte $N(0) = N_0 > 0$ vorgegeben. Ist die Wachstumsrate in (1.1) dagegen zufallsbehaftet, so liegt der Ansatz

$$a(t) = b(t) + R(t) \quad (1.2)$$

nahe, wobei $b : [0, \infty) \rightarrow \mathbb{R}$ fest gegeben ist, jedoch $R : [0, \infty) \rightarrow \mathbb{R}$ einen Zufallseffekt darstellt. Es kann R häufig als Rauschen interpretieren. Somit entsteht aus (1.2) die stochastische Variante

$$\frac{dN(t)}{dt} = b(t)N(t) + R(t)N(t), \quad (1.3)$$

wobei hier R eine relative Änderung bezüglich N bewirkt. Beeinflusst das Rauschen nicht die Wachstumsrate sondern das jeweilige Wachstum, so resultiert die Gleichung

$$\frac{dN(t)}{dt} = b(t)N(t) + R(t) \quad (1.4)$$

in der R eine absolute Änderung beschreibt. In den Gleichungen (1.3) und (1.4) liegt somit eine zeitabhängige Funktion R vor, die kontinuierliche Zufallseinflüsse enthält. Diese Funktion hängt damit nicht nur von der Zeit sondern auch von einem Wahrscheinlichkeitsraum ab. Zudem ist die Stetigkeit der Funktion R in der Zeit gewünscht.

Die Funktion R ist zu spezifizieren, um ein Modell für den stochastischen Einfluss zu erhalten. Ein naheliegender Ansatz erweist sich jedoch als ungeeignet, wie wir in einem späteren Abschnitt sehen werden. Daher wird das Rauschen in Zusammenhang mit der *Brownschen Bewegung* gebracht. Man nennt die entstehenden Modelle *stochastische Differentialgleichungen*.

Brownsche Bewegung

Im Jahr 1828 beobachtete der schottische Botaniker Robert Brown unter dem Mikroskop die Bewegung von Blütenpollen auf der Oberfläche einer Flüssigkeit. Er stellte eine irreguläre Bewegung der Pollen fest, die später über Zusammenstöße der Pollen mit den Molekülen in der Flüssigkeit erklärt werden konnte. Abbildung 1 skizziert drei Beispiele dieser Bewegungen in einer zweidimensionalen Ebene, wobei die Bewegungen jeweils im Nullpunkt starten.

Ein mathematisches Modell für diese Brownschen Bewegungen führt auf einen zeitabhängigen *Zufallsprozess* (auch: *stochastischen Prozess*). Dabei sind gewisse Eigenschaften im mathematischen Modell idealisiert, die zu qualitativen Abweichungen vom realen Verhalten der Pollenbahnen führen. Jedoch erweist sich die Brownsche Bewegung als Zufallsprozess geeignet zur Beschreibung von Rauschen. Abbildung 2 zeigt drei Beispiele der Brownschen Bewegung als zeitabhängige Funktionen im eindimensionalen Fall, wobei diese Rauscheffekte ersichtlich werden. Der unterschiedliche Verlauf der Bahnen in diesen Beispielen entsteht durch Einfluss von Zufall. Eine n -dimensionale Brownsche Bewegung resultiert, indem man n eindimensionale Bewegungen, die voneinander unabhängig sind, komponentenweise zu einem Vektor zusammenfasst.

Die ausgeprägte Bedeutung der Brownschen Bewegung im Gebiet der Stochastik erklärt sich aus ihrem engen Zusammenhang mit der Normalverteilung, welche für viele Problembeschreibungen grundlegend ist. Für weitere Details zur Brownschen Bewegung siehe z.B. [14].

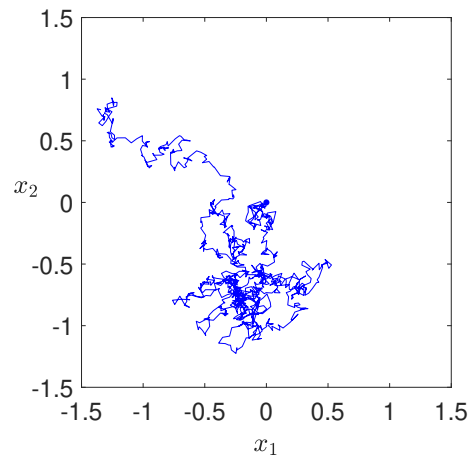
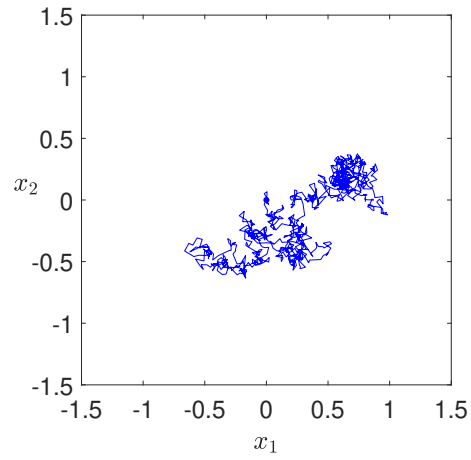
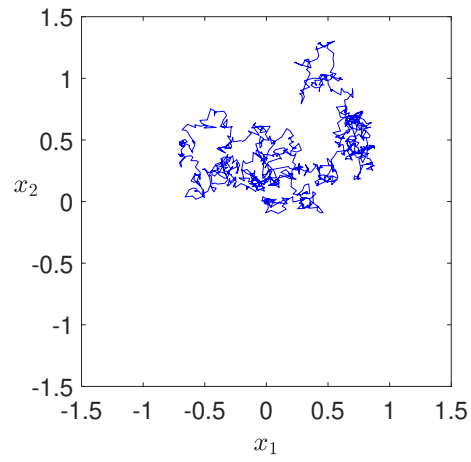


Abbildung 1: Approximationen von zweidimensionalen Brownschen Bewegungen im Phasendiagramm (Zeitintervall $t \in [0, 1]$, Zeitschrittweite $\Delta t = 10^{-3}$).

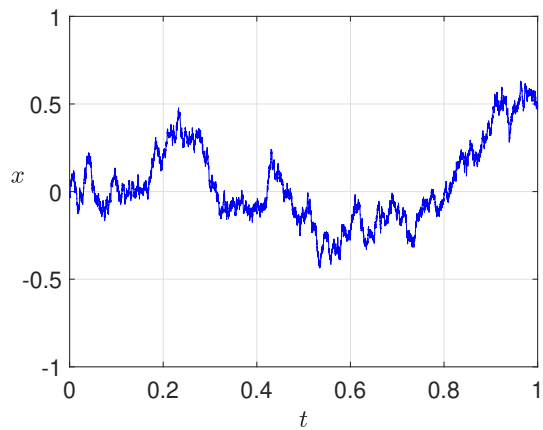
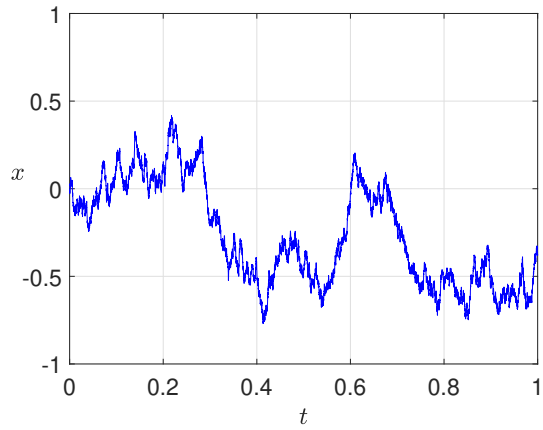
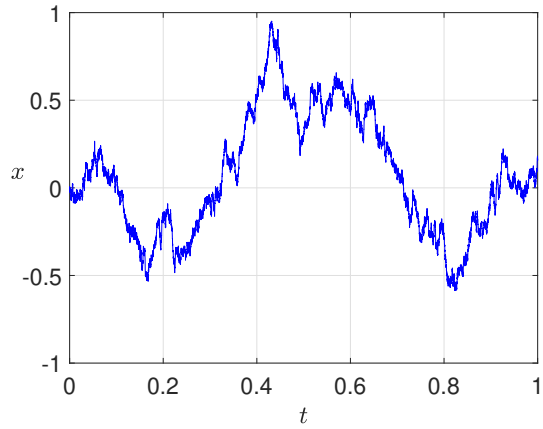


Abbildung 2: Approximationen von eindimensionalen Brownschen Bewegungen im Weg-Zeit-Diagramm (Zeitintervall $t \in [0, 1]$, Zeitschrittweite $\Delta t = 10^{-4}$).

2 Zufallsprozesse

In diesem Kapitel behandeln wir Zufallsprozesse (auch: stochastische Prozesse) und als einen Spezialfall davon den Wiener Prozess (auch: Brownsche Bewegung). Die Lösungen von stochastischen Differentialgleichungen werden selbst Zufallsprozesse sein.

2.1 Grundlagen aus der Wahrscheinlichkeitsrechnung

Eine σ -Algebra \mathcal{A} auf einer nichtleeren Menge Ω ist ein System aus Teilmengen von Ω , d.h. $\mathcal{A} \subseteq \mathcal{P}(\Omega)$ ($\mathcal{P}(\Omega)$: Potenzmenge von Ω), mit den drei Eigenschaften:

- (i) $\Omega \in \mathcal{A}$,
- (ii) $A \in \mathcal{A} \Rightarrow A^C = \Omega \setminus A \in \mathcal{A}$,
- (iii) $A_i \in \mathcal{A}, i \in \mathbb{N} \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$.

Man kann folgern, dass damit $\emptyset \in \mathcal{A}$ gilt und dass auch abzählbare Schnitte von Mengen wieder in \mathcal{A} liegen.

Ein *messbarer Raum* ist ein Paar (Ω, \mathcal{A}) bestehend aus einer nichtleeren Menge Ω und einer σ -Algebra \mathcal{A} auf Ω .

Ein *Maß* auf einem messbaren Raum (Ω, \mathcal{A}) ist eine Funktion $\mu : \mathcal{A} \rightarrow \mathbb{R}_0^+$ mit den Eigenschaften

- (i) $\mu(\emptyset) = 0$,
- (ii) $A_i \in \mathcal{A}, i \in \mathbb{N}, A_i \cap A_j = \emptyset$ für $i \neq j \Rightarrow \mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i)$.

Ein *Wahrscheinlichkeitsmaß* ist ein Maß $P : \mathcal{A} \rightarrow [0, 1]$ mit der zusätzlichen Eigenschaft $P(\Omega) = 1$. Eine Folgerung daraus ist das Maß des Komplements

$$P(A^C) = 1 - P(A) \quad \text{für alle } A \in \mathcal{A}.$$

Ein *Maßraum* ist ein Tripel $(\Omega, \mathcal{A}, \mu)$ bestehend aus einer nichtleeren Menge Ω , einer σ -Algebra \mathcal{A} auf Ω und einem Maß μ .

Ein *Wahrscheinlichkeitsraum* ist ein Tripel (Ω, \mathcal{A}, P) bestehend aus einer nicht-leeren Menge Ω , einer σ -Algebra \mathcal{A} auf Ω und einem Wahrscheinlichkeitsmaß P . Die Elemente $\omega \in \Omega$ nennen wir *Ergebnisse*. Die Teilmengen $A \in \mathcal{A}$ bezeichnen wir als *Ereignisse*. Das *Gegenereignis* zu $A \in \mathcal{A}$ ist das Komplement A^C . Der Wert $P(A)$ heißt die *Wahrscheinlichkeit* des Ereignisses $A \in \mathcal{A}$.

Sind $(\Omega_1, \mathcal{A}_1)$ und $(\Omega_2, \mathcal{A}_2)$ messbare Räume, dann nennen wir eine Abbildung $f : \Omega_1 \rightarrow \Omega_2$ *messbar*, wenn

$$f^{-1}(A) := \{\omega \in \Omega_1 : f(\omega) \in A\} \in \mathcal{A}_1 \quad \text{für alle } A \in \mathcal{A}_2$$

gilt. Man sagt auch kurz, dass f $(\mathcal{A}_1, \mathcal{A}_2)$ -messbar ist. Analog gilt diese Definition bei Funktionen zwischen Maßräumen. Eine messbare Abbildung $X : \Omega \rightarrow \mathbb{R}^n$ zwischen einem Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) und dem messbaren Raum $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$, wobei $\mathcal{B}(\mathbb{R}^n)$ die Borel-Algebra des \mathbb{R}^n bezeichnet, heißt *Zufallsvariable*. Der Fall $n = 1$ ist dabei zugelassen.

Zu einem Maßraum $(\Omega, \mathcal{A}, \mu)$ existiert die Theorie der Lebesgue-Integration. Eine messbare Funktion $f : \Omega \rightarrow \mathbb{R}$ heißt *integrierbar*, wenn ihr Positiv- und Negativteil jeweils ein endliches Integral besitzen. Dies sind genau die Funktionen aus der Menge $L^1(\Omega)$. Im folgenden betrachten wir einen Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) . Die messbaren Funktionen $f : \Omega \rightarrow \mathbb{R}$ sind damit Zufallsvariablen.

Für $f : \Omega \rightarrow \mathbb{R}$ mit $f \in L^1(\Omega)$ lautet der *Erwartungswert* (erstes Moment)

$$\mathbb{E}(f) := \int f \, dP = \int_{\Omega} f(\omega) \, dP(\omega).$$

Für eine messbare Funktion f gilt $f \in L^2(\Omega)$ falls $\mathbb{E}(f^2) < \infty$, d.h. f ist *quadratintegrierbar*. Es ist $L^2(\Omega)$ ein Hilbert-Raum mit dem Skalarprodukt

$$\langle f, g \rangle := \mathbb{E}(fg) = \int fg \, dP$$

für $f, g \in L^2(\Omega)$. Zudem gilt $L^2(\Omega) \subset L^1(\Omega)$ wegen $P(\Omega) < \infty$.

Zu einer Zufallsvariablen $f : \Omega \rightarrow \mathbb{R}$ haben wir die Kennzahlen

$$\begin{aligned} \mathbb{E}(f^n) & \quad n\text{-tes Moment (falls } f^n \in L^1(\Omega)), \\ \mathbb{E}(|f|^n) & \quad n\text{-tes absolutes Moment (falls } f^n \in L^1(\Omega)), \\ \mathbb{E}((f - \mathbb{E}(f))^n) & \quad n\text{-tes zentriertes Moment (falls } f, f^n \in L^1(\Omega)). \end{aligned}$$

Gilt $f \in L^2(\Omega)$, dann existiert die *Varianz* (zweites zentriertes Moment)

$$\text{Var}(f) := \mathbb{E}((f - \mathbb{E}(f))^2) = \mathbb{E}(f^2) - \mathbb{E}(f)^2.$$

Es ist $\sigma(f) := \sqrt{\text{Var}(f)}$ die *Standardabweichung*. Sind zwei Funktionen $f, g \in L^2(\Omega)$ gegeben, so ist deren *Covarianz* definiert durch

$$\text{Cov}(f, g) := \mathbb{E}((f - \mathbb{E}(f))(g - \mathbb{E}(g))) = \mathbb{E}(fg) - \mathbb{E}(f)\mathbb{E}(g).$$

Insbesondere gilt $\text{Cov}(f, f) = \text{Var}(f)$. Für $f : \Omega \rightarrow \mathbb{R}^n$, $f = (f_1, \dots, f_n)^\top$ werden die obigen Definitionen komponentenweise aufgefasst.

Ist eine Zufallsvariable $X : \Omega \rightarrow \mathbb{R}^n$ gegeben, so nennt man das induzierte Maß $\mu_X : \mathcal{B}(\mathbb{R}^n) \rightarrow [0, 1]$ mit

$$\mu_X(B) := P[X \in B] := P(X^{-1}(B)) \quad \text{für } B \in \mathcal{B}(\mathbb{R}^n)$$

die *Verteilung* von X . Die Messbarkeit von X stellt die Existenz dieses Maßes sicher. Im allgemeinen ist das Maß P nur in abstrakter Weise gegeben. Wir sind jedoch auch an der konkreten Berechnung dieser Wahrscheinlichkeiten interessiert. Diese Auswertung ist durchführbar, wenn die Verteilung durch eine Dichtefunktion $g : \mathbb{R}^n \rightarrow \mathbb{R}$ beschrieben werden kann. Für $B \in \mathcal{B}(\mathbb{R}^n)$ gilt dann

$$P[X \in B] = P(X^{-1}(B)) = \int_B g(x_1, \dots, x_n) dx_1 \cdots dx_n.$$

Für $(\mathcal{B}(\mathbb{R}^n), \mathcal{B}(\mathbb{R}^m))$ -messbare Funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ ist $f(X) = f \circ X$ dann $(\Omega, \mathcal{B}(\mathbb{R}^m))$ -messbar und es gilt im Fall $f \circ X \in L^1(\Omega)$

$$\begin{aligned} \mathbb{E}(f(X)) &= \int_{\Omega} f(X(\omega)) dP(\omega) \\ &= \int_{\mathbb{R}^n} f(x_1, \dots, x_n) \cdot g(x_1, \dots, x_n) dx_1 \cdots dx_n. \end{aligned}$$

Ist f die Identität ($m = n$), dann liegt komponentenweise der Erwartungswert von X vor. Das Problem der Auswertung von Wahrscheinlichkeiten im abstrakten Raum (Ω, \mathcal{A}, P) kann damit auf die übliche Lebesgue-Integration im \mathbb{R}^n zurückgeführt werden. Dies wird in Theorie und Numerik eingesetzt.

Für weitere Grundlagen zur Stochastik siehe [2, 3].

2.2 Definition eines Zufallsprozesses

Wir führen jetzt den grundlegenden Begriff eines Zufallsprozesses (auch: stochastischer Prozess) ein.

Definition 2.1 Sei (Ω, \mathcal{A}, P) ein Wahrscheinlichkeitsraum und $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ der messbare Raum mit der Borel-Algebra des \mathbb{R}^n . Ist $I \subseteq [0, \infty)$ eine Indexmenge, dann heißt eine Familie $(X_t)_{t \in I}$ aus messbaren Abbildungen $X_t : \Omega \rightarrow \mathbb{R}^n$ für $t \in I$ ein Zufallsprozess (mit Zustandsraum \mathbb{R}^n).

In allgemeinen wird t als Zeitparameter interpretiert und I ist ein Intervall. Somit ist für festes $t \in I$ die Abbildung

$$X_t : \Omega \rightarrow \mathbb{R}^n, \quad \omega \mapsto X_t(\omega)$$

eine n -dimensionale Zufallsvariable, wobei auch oft $n = 1$ betrachtet wird. Es liegt also eine Familie aus Zufallsvariablen parametrisiert durch t vor. Man hat jedoch auch eine alternative Interpretation für jeweils festes Ergebnis $\omega \in \Omega$. Die Funktion

$$(X_t)_{t \in I}(\omega) : I \rightarrow \mathbb{R}^n, \quad t \mapsto X_t(\omega) \tag{2.1}$$

nennt man einen *Pfad* von $(X_t)_{t \in I}$. Das Ergebnis ω kann man als Teilchen oder Experiment interpretieren. Der korrespondierende Pfad beschreibt die Bahn des Teilchens bzw. der Verlauf des Experiments mit der Zeit. Über eventuelle Stetigkeit oder Messbarkeit des Pfads ist hier noch keine Aussage enthalten.

Der Zufallsprozess kann auch betrachtet werden als eine Abbildung

$$(X_t)_{t \in I} : I \times \Omega \rightarrow \mathbb{R}^n, \quad (t, \omega) \mapsto X_t(\omega).$$

Wir könnten somit die Messbarkeit dieser Abbildung bezüglich des Produkt-raums $I \times \Omega$ mit entsprechender σ -Algebra untersuchen.

Die (endlichdimensionalen) *Verteilungen* eines Zufallsprozesses $(X_t)_{t \in I}$ sind die Maße

$$\mu_{t_1, \dots, t_k}(F_1 \times \dots \times F_k) := P[X_{t_1} \in F_1, \dots, X_{t_k} \in F_k] := P\left(\bigcap_{j=1}^k X_{t_j}^{-1}(F_j)\right)$$

definiert auf \mathbb{R}^{nk} , wobei $F_1, \dots, F_k \in \mathcal{B}(\mathbb{R}^n)$ und $t_1, \dots, t_k \in I$ gilt. Diese Familie aus Verteilungen bestimmt viele (aber nicht alle) wichtigen Eigenschaften des Prozesses.

2.3 Wiener Prozess bzw. Brownsche Bewegung

Die in der Einleitung beschriebene Brownsche Bewegung impliziert ein formales mathematisches Modell, das als *Wiener Prozess* bezeichnet wird.

Definition 2.2 Ein Zufallsprozess $(W_t)_{t \geq 0}$ mit $W_t : \Omega \rightarrow \mathbb{R}$ bezüglich eines Raums (Ω, \mathcal{A}, P) ist ein Wiener Prozess, wenn die folgenden Bedingungen gelten:

- (i) $W_0 = 0$ fast sicher,
- (ii) $(W_t)_{t \geq 0}$ besitzt unabhängige Inkremente, d.h. die beiden Zufallsvariablen $W_{t_2} - W_{t_1}$ und $W_{t_4} - W_{t_3}$ sind unabhängig für alle $t_1 < t_2 \leq t_3 < t_4$.
- (iii) Für $0 \leq s < t$ ist $W_t - W_s$ normalverteilt gemäß $N(0, t - s)$.

Aus diesen Eigenschaften folgern wir weiter für alle $t > 0$

$$\mathbb{E}(W_t) = \mathbb{E}(W_t - W_0) + \mathbb{E}(W_0) = 0 + 0 = 0$$

und für alle $0 \leq s < t$

$$\mathbb{E}((W_t - W_s)^2) = \text{Var}(W_t - W_s) + (\mathbb{E}(W_t - W_s))^2 = \text{Var}(W_t - W_s) = t - s$$

sowie (mit mehr Aufwand)

$$\text{Cov}(W_t, W_s) = \min\{s, t\} = s.$$

Für die entsprechende Korrelation gilt

$$\rho(W_t, W_s) = \frac{\text{Cov}(W_t, W_s)}{\sqrt{\text{Var}(W_t)\text{Var}(W_s)}} = \frac{s}{ts} = \frac{1}{t}.$$

Somit geht die Korrelation für festes s mit $t \rightarrow \infty$ gegen null, d.h. weit separierte Zeitpunkte hängen nur geringfügig voneinander ab.

Mit (i) und (iii) folgt wegen $W_t = W_t - W_0$ zudem für alle $a < b$

$$P[a \leq W_t \leq b] = \frac{1}{\sqrt{2\pi t}} \int_a^b \exp\left(-\frac{x^2}{2t}\right) dx.$$

Zudem ist ein Wiener Prozess ein *Gaußscher Prozess* im folgenden Sinn.

Definition 2.3 Ein Zufallsprozess $(X_t)_{t \geq 0}$ auf einem Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) ist ein Gaußscher Prozess, wenn für alle $t_1, \dots, t_n \geq 0$ die entsprechende n -dimensionale Zufallsvariable $(X_{t_1}, \dots, X_{t_n})^\top$ eine mehrdimensionale Normalverteilung besitzt.

Existenz des Wiener Prozesses

Wir haben noch zu zeigen, dass ein Zufallsprozess mit den Eigenschaften aus Definition 2.2 überhaupt existiert. Der *Erweiterungssatz von Kolmogorov* gibt eine allgemeinere Existenzaussage für Zufallsprozesse.

Satz 2.1 *Für alle $t_1, \dots, t_k \in [0, \infty)$ seien Wahrscheinlichkeitsmaße ν_{t_1, \dots, t_k} auf jeweils $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$ gegeben mit den Eigenschaften*

$$\nu_{t_1, \dots, t_k}(F_1 \times \dots \times F_k) = \nu_{t_{\sigma(1)}, \dots, t_{\sigma(k)}}(F_{\sigma(1)} \times \dots \times F_{\sigma(k)}) \quad (2.2)$$

für alle $F_i \in \mathcal{B}(\mathbb{R})$ und für alle Permutationen σ auf $\{1, 2, \dots, k\}$ sowie

$$\nu_{t_1, \dots, t_k}(F_1 \times \dots \times F_k) = \nu_{t_1, \dots, t_k, t_{k+1}, \dots, t_{k+m}}(F_1 \times \dots \times F_k \times \mathbb{R} \times \dots \times \mathbb{R}) \quad (2.3)$$

für alle weiteren $t_{k+1}, \dots, t_{k+m} \in [0, \infty)$. Dann existiert ein Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) und ein Zufallsprozess $(X_t)_{t \geq 0}$ auf diesem Raum mit

$$\nu_{t_1, \dots, t_k}(F_1 \times \dots \times F_k) = P[X_{t_1} \in F_1, \dots, X_{t_k} \in F_k] \quad (2.4)$$

für alle $t_1, \dots, t_k \in [0, \infty)$ und alle $F_i \in \mathcal{B}(\mathbb{R})$.

Da ein Wiener Prozess auch ein Gaußscher Prozess sein soll, werden nun Maße bezüglich mehrdimensionaler Normalverteilungen eingeführt. Die konkrete Gestalt ist motiviert durch das (komponentenweise) Verhalten der Brownschen Bewegung von Pollen, siehe Kapitel 1. Dabei sei $x_0 \in \mathbb{R}$ ein gegebener Anfangswert der Position. Wir setzen

$$p : [0, \infty) \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}, \quad p(t, x, y) := \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{(x-y)^2}{2t}\right).$$

Zu $0 \leq t_1 \leq t_2 \leq \dots \leq t_k$ definieren wir das Maß auf \mathbb{R}^k für $F := F_1 \times \dots \times F_k$ mit $F_i \in \mathcal{B}(\mathbb{R})$

$$\begin{aligned} \nu_{t_1, \dots, t_k}(F_1 \times \dots \times F_k) &:= \\ &\int_F p(t_1, x_0, x_1) p(t_2 - t_1, x_1, x_2) \dots p(t_k - t_{k-1}, x_{k-1}, x_k) dx_1 \dots dx_k, \end{aligned}$$

wobei die übliche Lebesgue-Integration im \mathbb{R}^k verwendet wird. Für den Spezialfall $t_i = t_{i+1}$ sei $p(0, x, y) = \delta_x(y)$ die Einheitsmasse an der Stelle x . Für beliebige $t_1, \dots, t_k \in [0, \infty)$ wird dann ν_{t_1, \dots, t_k} über (2.2) definiert. Wegen

$$\int_{\mathbb{R}} p(t, x, y) dy = \frac{1}{\sqrt{2\pi t}} \int_{\mathbb{R}} \exp\left(-\frac{(x-y)^2}{2t}\right) dy = 1$$

für alle $x \in \mathbb{R}$ und $t > 0$ ist die Bedingung (2.3) erfüllt. Nach dem Satz von Kolmogorov (Satz 2.1) existiert damit ein Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P_{x_0})$ und ein Zufallsprozess $(W_t)_{t \geq 0}$ mit der Eigenschaft, vergleiche (2.4),

$$P_{x_0}[W_{t_1} \in F_1, \dots, W_{t_k} \in F_k] := \int_F p(t_1, x_0, x_1) p(t_2 - t_1, x_1, x_2) \cdots p(t_k - t_{k-1}, x_{k-1}, x_k) dx_1 \cdots dx_k. \quad (2.5)$$

Ein solcher Prozess wird eine (eindimensionale) *Brownsche Bewegung* mit Startpunkt x_0 genannt. Es gilt nämlich $P_{x_0}[W_0 = x_0] = 1$. Jedoch gibt es mehrere Wahrscheinlichkeitsräume und korrespondierende Prozesse, für die (2.5) erfüllt ist. Wir setzen $x_0 = 0$ und $P = P_0$, damit die Bedingung (i) aus Definition 2.2 erfüllt ist. Aus der Darstellung (2.5) folgt direkt $W_t \sim N(0, t)$ für alle $t \geq 0$. Zudem impliziert (2.5), dass $(W_t)_{t \geq 0}$ ein Gaußscher-Prozess ist. Die Covarianzmatrix zu $(W_{t_1}, \dots, W_{t_k})$ für $0 \leq t_1 \leq \dots \leq t_k$ lautet

$$C := (\mathbb{E}(W_{t_i} W_{t_j}))_{i,j=1,\dots,k} = \begin{pmatrix} t_1 & t_1 & t_1 & \cdots & t_1 \\ t_1 & t_2 & t_2 & \cdots & t_2 \\ t_1 & t_2 & t_3 & \cdots & t_3 \\ \vdots & \vdots & \vdots & & \vdots \\ t_1 & t_2 & t_3 & \cdots & t_k \end{pmatrix}.$$

Eine Folgerung hieraus ist

$$\mathbb{E}(W_t W_s) = \min\{s, t\} \quad \text{für alle } s, t \geq 0 \quad (2.6)$$

und

$$\text{Cov}(W_t, W_s) = \mathbb{E}(W_t W_s) - \mathbb{E}(W_t)\mathbb{E}(W_s) = \mathbb{E}(W_t W_s) = \min\{s, t\}.$$

Für $u \leq v \leq s \leq t$ folgt

$$\begin{aligned} & \mathbb{E}((W_v - W_u)(W_t - W_s)) \\ &= \mathbb{E}(W_v W_t) - \mathbb{E}(W_u W_t) - \mathbb{E}(W_v W_s) + \mathbb{E}(W_u W_s) \\ &= v - u - v + u = 0. \end{aligned}$$

Zwei normalverteilte Zufallsvariablen sind unabhängig genau dann, wenn ihre Covarianz null ergibt. Somit sind die Inkremente unabhängig und Bedingung (ii) aus Definition 2.2 ist erfüllt. Das Inkrement $W_t - W_s$ für $0 \leq s < t$ ist als Differenz von normalverteilten Zufallsvariablen wieder normalverteilt mit

$$\begin{aligned} \mathbb{E}(W_t - W_s) &= \mathbb{E}(W_t) - \mathbb{E}(W_s) = 0 - 0 = 0 \\ \text{Var}(W_t - W_s) &= \text{Var}(W_t) + \text{Var}(W_s) - 2\text{Cov}(W_t, W_s) \\ &= t + s - 2s = t - s. \end{aligned}$$

Somit ist auch die Bedingung (iii) aus Definition 2.2 gegeben. Wir haben die Existenz eines Wiener Prozesses gezeigt.

Pfade eines Wiener Prozesses

Wir betrachten jetzt die Pfade eines Wiener Prozesses gemäß (2.1). Die Frage, ob die Pfade $W_t(\omega)$ eines Wiener Prozesses für fast alle $\omega \in \Omega$ stetig sind, kann nicht direkt behandelt werden. Jedoch ist die Existenz einer stetigen Version gesichert. Sind zwei Prozesse $(X_t)_{t \geq 0}$ und $(Y_t)_{t \geq 0}$ auf einem Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) gegeben, dann heißt $(Y_t)_{t \geq 0}$ eine Version oder Modifikation von $(X_t)_{t \geq 0}$, wenn gilt

$$P(\{\omega : X_t(\omega) = Y_t(\omega)\}) = 1 \quad \text{für alle } t \geq 0. \quad (2.7)$$

In diesem Fall besitzen $(X_t)_{t \geq 0}$ und $(Y_t)_{t \geq 0}$ die gleichen (endlichdimensionalen) Verteilungen. Ihre Pfadeigenschaften können jedoch unterschiedlich sein. Die Frage nach der Stetigkeit der Pfade zur Brownschen Bewegung kann mit folgendem *Stetigkeitssatz von Kolmogorov* beantwortet werden.

Satz 2.2 *Sei ein Zufallsprozess $(X_t)_{t \geq 0}$ auf einem Raum (Ω, \mathcal{A}, P) gegeben. Existieren zu jedem $T > 0$ Konstanten $\alpha, \beta, C > 0$, so dass*

$$\mathbb{E}(|X_t - X_s|^\alpha) \leq C \cdot |t - s|^{1+\beta} \quad \text{für } 0 \leq s, t \leq T$$

gilt, dann gibt es eine Version von $(X_t)_{t \geq 0}$ bei der fast alle Pfade stetig sind.

Der oben konstruierte Wiener Prozess erfüllt wegen $W_t - W_s \sim N(0, t - s)$ die Beziehung

$$\mathbb{E}(|W_t - W_s|^4) = 3|t - s|^2,$$

wodurch der Satz anwendbar ist. Somit haben wir das Ziel erreicht. Im folgenden gehen wir stets von einer Version eines Wiener Prozesses aus, bei der fast alle Pfade stetig sind. Über die Eindeutigkeit dieser stetigen Version liegt hier keine Aussage vor. Jedoch besitzen alle stetigen Versionen die gleichen Eigenschaften.

Es kann bewiesen werden, dass fast alle Pfade eines Wiener Prozesses fast nirgends differenzierbar sind, siehe [5]. Daraus folgt, dass fast alle Pfade eine unbeschränkte Variation in jedem Zeitintervall (s, t) mit $0 \leq s < t$ besitzen. Die Länge des Funktionsgraphen eines Pfads in (s, t) können wir somit als unendlich ansehen.

Weierstraß-Oszillator

Um die Eigenschaften eines Pfads zu einem Wiener Prozess besser zu verstehen, betrachten wir ein Beispiel für eine (deterministische) stetige Funktion mit unbeschränkter Variation: der *Weierstraß-Oszillator*. Die Sägezahn-Funktion lautet

$$s(x) := \text{dist}(x, \mathbb{Z}) := \min\{x - \lfloor x \rfloor, \lceil x \rceil - x\} \quad (2.8)$$

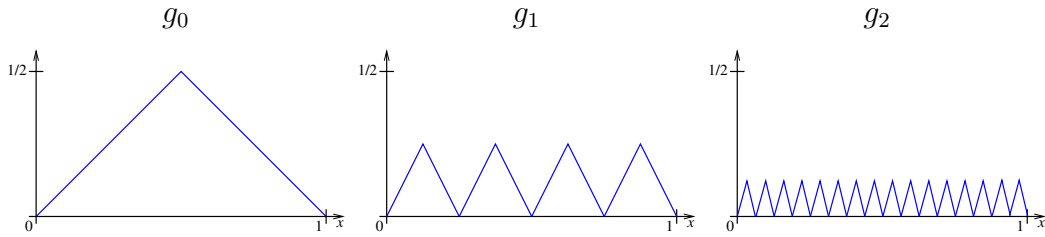


Abbildung 3: Funktionen $g_k(x)$ im Bereich $x \in [0, 1]$.

mit den Gauß-Klammern $[\cdot], [\cdot]$. Der Weierstraß-Oszillator ist der Grenzwert einer Funktionenreihe

$$\hat{f}(x) = \lim_{m \rightarrow \infty} f_m(x) \quad \text{für } x \in \mathbb{R} \quad (2.9)$$

mit den Partialsummen

$$f_m(x) = \sum_{k=0}^m 2^{-k} s(4^k x). \quad (2.10)$$

Offensichtlich gilt $|f_m(x) - f_{m+1}(x)| \leq 2^{-(m+1)}$ für alle x und somit ist die Reihe gleichmäßig konvergent. Die Grenzfunktion (2.9) ist dadurch stetig und besitzt zudem die Periode 1.

In den Partialsummen (2.10) treten die Funktionen

$$g_k(x) = 2^{-k} s(4^k x) \quad \text{für } k \in \mathbb{N}_0$$

auf, welche selbst eine Sägezahn-Form besitzen, siehe Abbildung 3. Offensichtlich gilt $g_k(x) \geq 0$ für alle x und alle k . Zudem ist

$$\max\{g_k(x) : x \in \mathbb{R}\} = 2^{-(k+1)}.$$

Wir können interpretieren, dass sich die Amplitude der Funktionen g_k von $k \rightarrow k + 1$ überall halbiert. Jetzt betrachten wir das Intervall $[0, 1]$. Die Sägezahn-Funktion $s = g_0$ aus (2.8) besitzt eine Dreiecksform. Wegen des Terms $s(4^k x)$ vervierfacht sich die Anzahl der Dreiecke von $k \rightarrow k + 1$. Wir können interpretieren, dass sich die Frequenz vervierfacht. Es folgt, dass die Länge des Graphen von g_k in $[0, 1]$ gegen unendlich geht für $k \rightarrow \infty$.

Man kann zeigen, dass \hat{f} nirgends differenzierbar ist und dass die Variation unbeschränkt ist auf allen Intervallen $[a, b]$ mit $a < b$. Abbildung 4 zeigt Partialsummen der Funktionenfolge. Wir erkennen die Zunahme der Variation in den Folgengliedern.

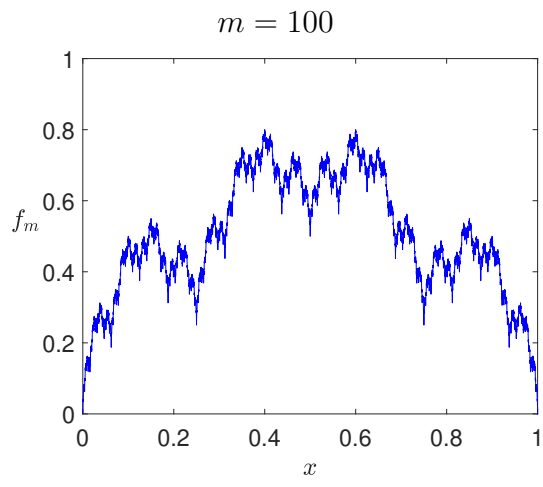
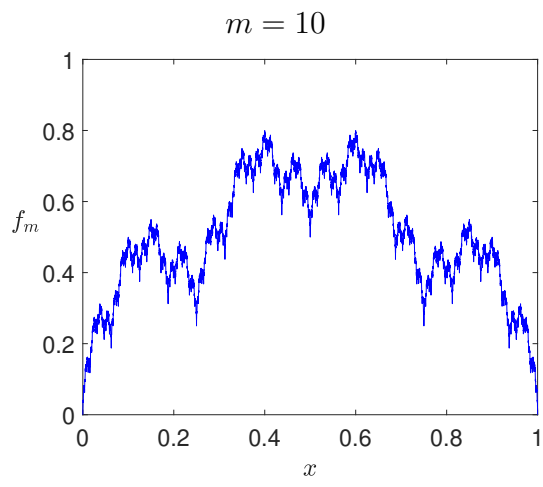
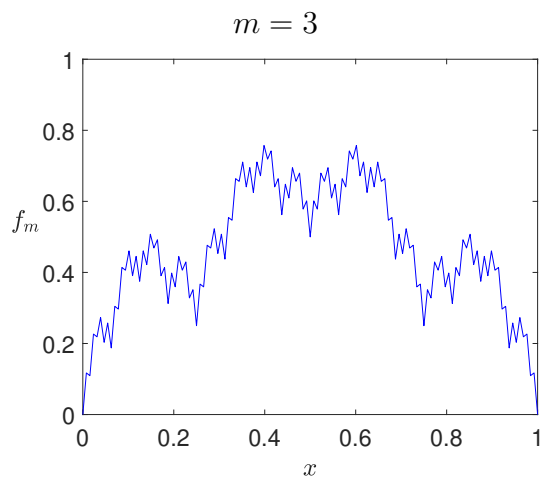


Abbildung 4: Partialsummen des Weierstraß-Oszillators.

Mehrdimensionale Wiener Prozesse

Ein mehrdimensionaler Wiener Prozess entsteht einfach durch komponentenweise eindimensionale Prozesse.

Definition 2.4 Ein Prozess $(W_t)_{t \geq 0}$ auf einem Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) mit $W_t : \Omega \rightarrow \mathbb{R}^n$ und $W_t = (W_1(t), \dots, W_n(t))^\top$ heißt *n*-dimensionaler Wiener Prozess, wenn $(W_i(t))_{t \geq 0}$ für jedes i ein Wiener Prozess ist und die Prozesse $(W_i(t))_{t \geq 0}$ für $i = 1, \dots, n$ unabhängig sind.

Folgerungen hieraus sind:

- (i) $(W_1(0), \dots, W_n(0)) = (0, \dots, 0)$ fast sicher,
- (ii) $(W_1(t), \dots, W_n(t))_{t \geq 0}^\top$ besitzt unabhängige Inkremente im \mathbb{R}^n ,
- (iii) $(W_1(t), \dots, W_n(t))^\top - (W_1(s), \dots, W_n(s))^\top$ ist für $0 \leq s < t$ nun n -dimensional normalverteilt nach $N(0, (t-s)I)$ mit dem Nullvektor als Erwartungswert und der Covarianzmatrix $(t-s)I$ (I : Einheitsmatrix).

Die Existenz eines n -dimensionalen Wiener Prozesses ergibt sich aus der bereits gezeigten Existenz eines eindimensionalen Wiener Prozesses. Analog verwenden wir stets eine Version eines mehrdimensionalen Wiener Prozesses, bei dem alle Pfade (komponentenweise) stetig sind.

Numerische Simulation eines Wiener Prozesses

Wir können jeweils Näherungen von Pfaden eines Wiener Prozesses erzeugen in einem endlichen Intervall $[0, T]$. Dazu werden Gitterpunkte im Intervall eingeführt

$$0 = t_0 < t_1 < t_2 < \dots < t_{m-1} < t_m = T.$$

Eine einfache und übliche Wahl sind äquidistante Gitterpunkte $t_j = j\Delta t$ mit Zeitschrittweite $\Delta t = \frac{T}{m}$. Berechnet werden Werte $s_j = W_{t_j}(\omega)$ für $j = 1, \dots, m$ und ein festes Ergebnis $\omega \in \Omega$, d.h. ein Pfad.

Hierzu werden Zufallszahlen verwendet. Seien Z_1, Z_2, \dots, Z_m unabhängige identisch verteilte Zufallsvariablen mit Standard-Normalverteilung ($Z_j \sim N(0, 1)$). Ideale Zufallszahlen sind reelle Zahlen z_1, z_2, \dots, z_m , wobei $z_j \in \mathbb{R}$ eine Realisierung der Zufallsvariable Z_j ist. In der Praxis sind wir dabei jedoch auf Pseudo-Zufallszahlen angewiesen. Bezüglich der Erzeugung von Pseudo-Zufallszahlen zur Normalverteilung, siehe [6].

Laut Definition 2.2 ist jedes Inkrement $W_{t_j} - W_{t_{j-1}}$ eines Wiener Prozesses eine normalverteilte Zufallsvariable mit Erwartungswert null und Varianz $t_j - t_{j-1}$. Wir definieren Zufallsvariablen rekursiv durch

$$S_j = S_{j-1} + \sqrt{t_j - t_{j-1}} Z_j \quad \text{für } j = 1, \dots, m.$$

Der Anfangswert ist $S_0 = 0$. Es gilt

$$\begin{aligned} \mathbb{E}(\sqrt{t_j - t_{j-1}} Z_j) &= \sqrt{t_j - t_{j-1}} \underbrace{\mathbb{E}(Z_j)}_{=0} = 0 \\ \text{Var}(\sqrt{t_j - t_{j-1}} Z_j) &= \sqrt{t_j - t_{j-1}}^2 \underbrace{\text{Var}(Z_j)}_{=1} = t_j - t_{j-1}. \end{aligned}$$

Dadurch liegt eine Normalverteilung vor: $\sqrt{t_j - t_{j-1}} Z_j \sim N(0, t_j - t_{j-1})$.

Einen Pfad erhalten wir durch die Rekursion mit den Zufallszahlen

$$s_j = s_{j-1} + \sqrt{t_j - t_{j-1}} z_j \quad \text{für } j = 1, \dots, m$$

und dem Anfangswert $s_0 = 0$.

Wären z_1, z_2, \dots, z_m ideale Zufallszahlen, dann wären die Werte s_j exakte Realisierungen aus einem Pfad eines Wiener Prozesses in den Gitterpunkten. Die Verwendung von Pseudo-Zufallszahlen macht daraus streng genommen nur eine Näherung. Es ist zur Darstellung der Pfade sinnvoll, die zweidimensionalen Punkte (t_j, s_j) für $j = 0, 1, \dots, m$ mit einem Streckenzug (Polygonzug) zu verbinden, d.h. eine stückweise lineare Interpolation. Jedoch ist dies nur eine Näherung des Pfades in einem offenen Teilintervall (t_{j-1}, t_j) , da der exakte Pfad z.B. keine endliche Länge in einem Intervall besitzt.

Dieses Verfahren kann beliebig oft wiederholt werden. Somit können wir eine Vielzahl an Pfaden erzeugen. In der Praxis werden häufig nur die Werte an einem Endzeitpunkt T aus diesen Pfaden benötigt. Abbildung 5 zeigt Beispiele für Pfade eines Wiener Prozesses aus diesem Verfahren.

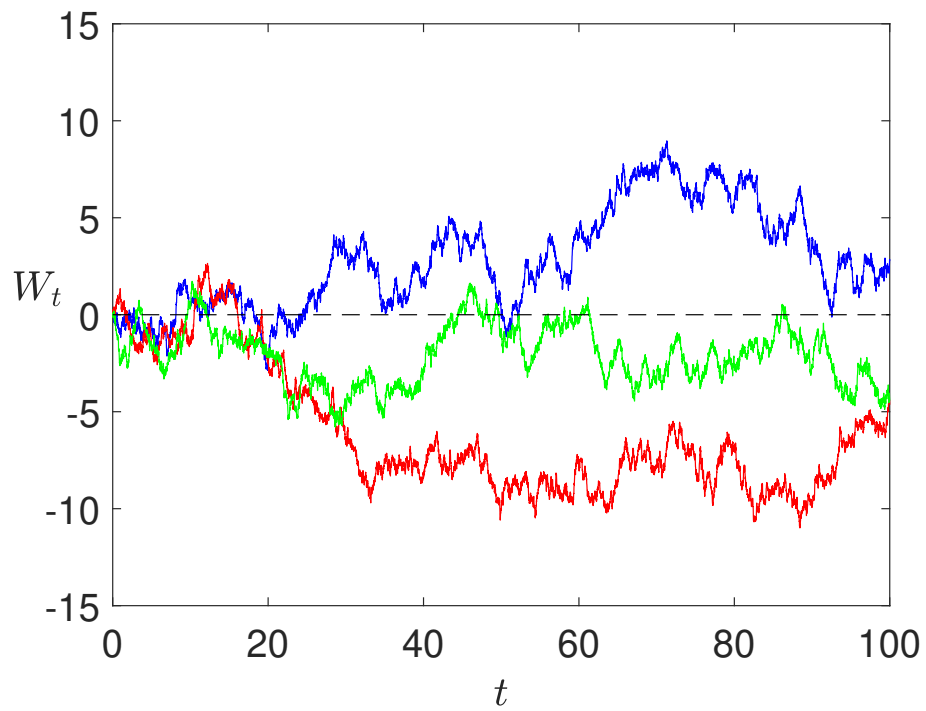
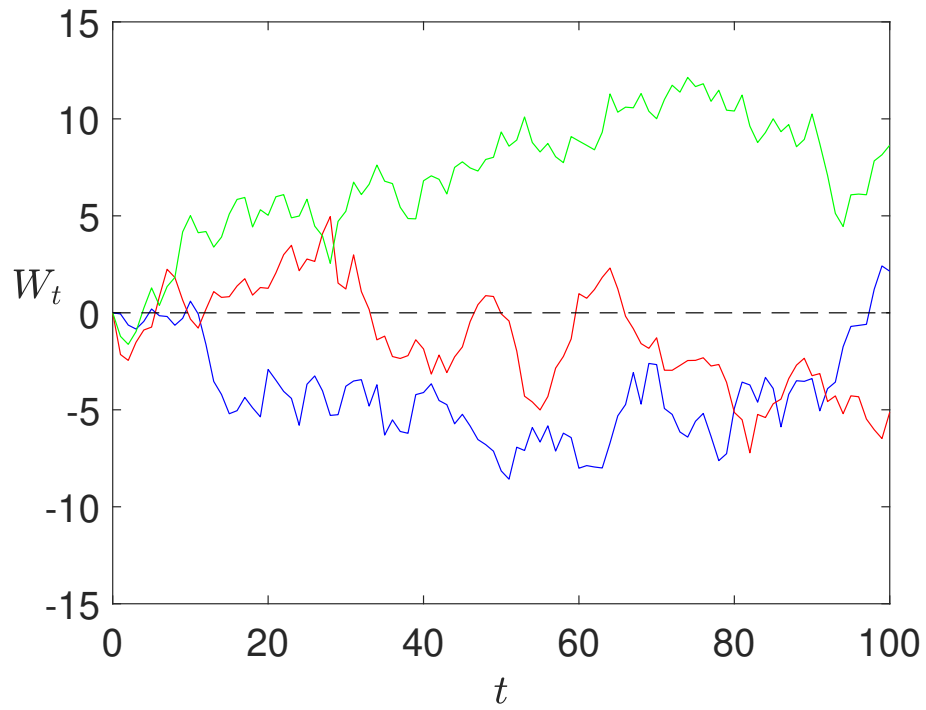


Abbildung 5: Näherungen von Pfaden eines Wiener Prozesses im Intervall $[0, 100]$ mit Zeitschrittweite $\Delta t = 1$ (oben) und $\Delta t = 0.01$ (unten).

3 Stochastische Integrale

In diesem Kapitel führen wir einen Integralbegriff für Zufallsprozesse ein. Dieses Integral dient dann zur Definition von Integralgleichungen, die als stochastische Differentialgleichungen bezeichnet werden.

3.1 Deterministische Integralbegriffe

Für Anfangswertprobleme gewöhnlicher Differentialgleichungen

$$y'(t) = f(t, y(t)), \quad y(t_0) = y_0 \tag{3.1}$$

existiert mit glatten Funktionen $y : \mathbb{R} \rightarrow \mathbb{R}^n$ die äquivalente *Integralgleichung*

$$y(t) = y_0 + \int_{t_0}^t f(s, y(s)) \, ds. \tag{3.2}$$

Für eine stetige Funktion, welche nicht überall differenzierbar ist, kann die Differentialgleichung (3.1) nicht ausgewertet werden. Dagegen könnte die Funktion eine Lösung der Integralgleichung (3.2) sein. Die Integralgleichung erlaubt somit eine größere Lösungsmenge. Da beim Wiener Prozess bzw. der Brownschen Bewegung nur stetige aber nicht differenzierbare Pfade auftreten, diskutieren wir im folgenden Integralbegriffe.

Eine Folge aus endlichen Zerlegungen von $[a, b] \subset \mathbb{R}$ (o.B.d.A. $a < b$) in jeweils n Teilintervalle sei gegeben

$$a = t_0^{(n)} < t_1^{(n)} < t_2^{(n)} < \dots < t_{n-1}^{(n)} < t_n^{(n)} = b, \tag{3.3}$$

welche die Voraussetzung

$$\lim_{n \rightarrow \infty} \left(\max_{i=1, \dots, n} t_i^{(n)} - t_{i-1}^{(n)} \right) = 0 \tag{3.4}$$

erfüllt. Wir wählen Zwischenstellen $\xi_i^{(n)} \in [t_{i-1}^{(n)}, t_i^{(n)}]$. Es sei $f : [a, b] \rightarrow \mathbb{R}$ eine stetige Funktion. Dann gilt

$$\int_a^b f(t) \, dt := \lim_{n \rightarrow \infty} \sum_{i=1}^n f(\xi_i^{(n)}) \cdot (t_i^{(n)} - t_{i-1}^{(n)}), \tag{3.5}$$

d.h. alle Grenzwerte existieren und sind identisch. Auf diese Weise wird das *Riemann-Integral* (R-Integral) definiert. Man beachte, dass dabei alle Grenzwerte unabhängig von der Wahl der Zerlegungen und der Zwischenstellen existieren und identisch sein müssen. Bei einer stetigen Funktion ist dies immer gegeben.

Spezielle Wahlen für die Zwischenstellen sind beispielsweise die Intervallränder

$$\int_a^b f(t) dt = \lim_{n \rightarrow \infty} \sum_{i=1}^n f(t_{i-1}^{(n)}) \cdot (t_i^{(n)} - t_{i-1}^{(n)})$$

oder die Intervallmittelpunkte

$$\int_a^b f(t) dt = \lim_{n \rightarrow \infty} \sum_{i=1}^n f\left(\frac{1}{2}(t_{i-1}^{(n)} + t_i^{(n)})\right) \cdot (t_i^{(n)} - t_{i-1}^{(n)}).$$

Eine Verallgemeinerung des Riemann-Integrals (3.5) besteht darin, dass das Integrationsgebiet kein Intervall sondern eine Menge von Werten einer anderen Funktion ist. Es sei wieder eine Folge von Zerlegungen (3.3) mit der Eigenschaft (3.4) gegeben. Zudem sei $g : \mathbb{R} \rightarrow \mathbb{R}$ eine vordefinierte Funktion, genannt der *Integrator*. Das *Riemann-Stieltjes-Integral* (RS-Integral) vom Integrand f bezüglich Integrator g wird definiert durch

$$\int_a^b f(t) dg(t) := \lim_{n \rightarrow \infty} \sum_{i=1}^n f(\xi_i^{(n)}) \cdot (g(t_i^{(n)}) - g(t_{i-1}^{(n)})). \quad (3.6)$$

Wieder wird gefordert, dass die Grenzwerte für alle Zerlegungen und Zwischenstellen existieren und identisch sind. Für $g(t) \equiv t$ entsteht gerade das gewöhnliche R-Integral.

Die Stetigkeit von g ist nicht hinreichend für die Existenz des RS-Integral. Dagegen ist g stückweise monoton hinreichend. Allgemeiner existiert das RS-Integral für alle Funktionen g , die eine beschränkte *Variation* besitzen. Die Variation auf einem Intervall $[a, b]$ wird definiert mittels

$$V_a^b(g) := \sup \left\{ \sum_{i=1}^m |g(t_i) - g(t_{i-1})| : \begin{array}{l} a = t_0 < t_1 < \dots \\ < t_{m-1} < t_m = b \end{array} \right\}. \quad (3.7)$$

Ist die Variation von g auf $[a, b]$ beschränkt, dann ist zudem g fast überall differenzierbar in $[a, b]$. Ist dagegen die Variation von g nicht beschränkt, so ist die Existenz des RS-Integrals für allgemeines stetiges f nicht gesichert. Funktionen mit endlicher Variation können als Differenz zweier monoton wachsender Funktionen dargestellt werden, d.h.

$$V_a^b(g) < \infty \quad \Rightarrow \quad g = g_1 - g_2 \quad \text{mit } g_1, g_2 \text{ monoton wachsend.}$$

Somit folgt als Integral

$$\int_a^b f(t) dg(t) = \int_a^b f(t) dg_1(t) - \int_a^b f(t) dg_2(t).$$

Es gilt insbesondere die Abschätzung

$$\left| \int_a^b f(t) \, dg(t) \right| \leq \left(\max_{t \in [a,b]} |f(t)| \right) \cdot V_a^b(g).$$

Ist die Funktion g stetig differenzierbar, so gilt für deren Variation

$$V_a^b(g) = \int_a^b |g'(t)| \, dt \leq (b - a) \cdot \max_{t \in [a,b]} |g'(t)| < \infty.$$

Man hat dann mit dem Mittelwertsatz der Differentialrechnung die Darstellung

$$\begin{aligned} \int_a^b f(t) \, dg(t) &= \lim_{n \rightarrow \infty} \sum_{i=1}^n f(\xi_i^{(n)}) \cdot \frac{g(t_i^{(n)}) - g(t_{i-1}^{(n)})}{t_i^{(n)} - t_{i-1}^{(n)}} \cdot (t_i^{(n)} - t_{i-1}^{(n)}) \\ &= \lim_{n \rightarrow \infty} \sum_{i=1}^n f(\xi_i^{(n)}) \cdot g'(\tilde{\xi}_i^{(n)}) \cdot (t_i^{(n)} - t_{i-1}^{(n)}) \\ &= \int_a^b f(t)g'(t) \, dt. \end{aligned}$$

Das RS-Integral kann in diesem Fall als die R-Integration bezüglich einer Dichtefunktion interpretiert werden.

Wir betrachten später als Funktion g Realisierungen bzw. Pfade eines Wiener Prozesses, für die V_a^b unbeschränkt für alle $a < b$ ist. Daher ist ein alternativer Integralbegriff notwendig. So wie der Begriff des Riemann-Integrals zum *Lebesgue-Integral* erweitert werden kann, so existiert auch zum Riemann-Stieltjes-Integral die Verallgemeinerung zum *Lebesgue-Stieltjes-Integral*. Dieses liefert jedoch ebenfalls keinen Integralbegriff bei Auftreten von Funktionen mit unbeschränkter Variation.

3.2 Integrale mit Zufallsprozessen

Als Motivation betrachten wir eine Differentialgleichung mit einem zufallsabhängigen Rauschterm. Wir sind an der Herleitung einer Differentialgleichung der Form

$$\frac{dX_t}{dt} = b(t, X_t) + \sigma(t, X_t)R_t \tag{3.8}$$

mit $t \geq 0$ und Anfangswert X_0 interessiert, wobei $b, \sigma : [0, \infty) \times \mathbb{R} \rightarrow \mathbb{R}$ gegebene stetige Funktionen sind und die Funktion $R : [0, \infty) \rightarrow \mathbb{R}$, $t \mapsto R_t$ die Bedeutung eines Rauschterms hat. Dadurch ist es sinnvoll einen geeigneten Zufallsprozess $(R_t)_{t \geq 0}$ bezüglich eines Wahrscheinlichkeitsraums (Ω, \mathcal{A}, P) anzusetzen. Jeder Pfad $R_t(\omega)$ soll (fast sicher) stetig sein. Die Lösung $X_t : [0, \infty) \rightarrow \mathbb{R}$ wird somit auch als Zufallsprozess $(X_t)_{t \geq 0}$ zu interpretieren sein.

Aus den Anwendungen folgen drei natürliche Forderungen an den Prozess $(R_t)_{t \geq 0}$:

- (i) R_{t_1} und R_{t_2} sind unabhängig für $t_1 \neq t_2$.
- (ii) Die Familie $(R_t)_{t \geq 0}$ ist stationär, d.h. die (gemeinsame) Verteilung von $\{R_{t_1+\tau}, \dots, R_{t_k+\tau}\}$ ist identisch für alle $\tau \geq 0$.
- (iii) $\mathbb{E}(R_t) = 0$ für alle $t \geq 0$.

Die Bedingung (i) bedeutet, dass das Rauschen in den Zeitpunkten separat vorliegt. Die Forderung (ii) kann man so interpretieren, dass das Rauschen in den Zeitpunkten sich qualitativ gleich verhält. Die Bedingung (iii) lässt sich stets erfüllen, indem man den Prozess standardisiert. Dies bedeutet, wir setzen $\tilde{R}_t = R_t - \mathbb{E}(R_t)$ falls der Erwartungswert ungleich null ist. Jedoch kann ein Zufallsprozess, der die Bedingungen (i) und (ii) erfüllt, keine stetigen Pfade besitzen. Damit sind auch für die Lösung X_t keine geeigneten Pfade möglich.

Würde ein Zufallsprozess $(R_t)_{t \geq 0}$ mit stetigen Pfaden und den Bedingungen (i)-(iii) existieren, dann könnten wir eine Realisierung $R_t(\omega)$ a priori (näherungsweise) bestimmen und diese in die Gleichung (3.8) einsetzen. Dann läge eine gewöhnliche Differentialgleichung mit stetiger rechter Seite vor und wir könnten übliche numerische Verfahren zur Bestimmung von $X_t(\omega)$ verwenden. Aussagen über den Zufallsprozess X_t würden aus einer Monte-Carlo-Simulation folgen.

Es ist also ein alternativer Ansatz notwendig, um eine zu (3.8) analoge Gleichung für Lösungen mit stetigen Pfaden herzuleiten. Dazu schreiben wir die Differentialgleichung (3.8) in diskretisierter Form auf dem Zeitgitter

$$0 = t_0 < t_1 < \dots < t_k = t, \quad X_j := X_{t_j}, \quad R_j := R_{t_j}, \quad \Delta t_j := t_{j+1} - t_j. \quad (3.9)$$

Es folgt (analog zum expliziten Euler-Verfahren)

$$\frac{X_{j+1} - X_j}{t_{j+1} - t_j} = b(t_j, X_j) + \sigma(t_j, X_j)R_j$$

und damit

$$X_{j+1} = X_j + b(t_j, X_j)\Delta t_j + \sigma(t_j, X_j)R_j\Delta t_j. \quad (3.10)$$

Die Idee ist nun, die Inkremente $R_j\Delta t_j$ durch Zuwächse $\Delta W_j := W_{t_{j+1}} - W_{t_j}$ mit einem geeigneten Zufallsprozess $(W_t)_{t \geq 0}$ zu ersetzen. Die ursprünglichen Forderungen an den Zufallsprozess $(R_t)_{t \geq 0}$ übertragen sich nun auf die Inkremente von $(W_t)_{t \geq 0}$ wie folgt:

- (i) Zu einer endlichen Zerlegung $0 \leq t_1 < t_2 < \dots < t_k$ sind die korrespondierenden Inkremente $W_{t_2} - W_{t_1}, \dots, W_{t_k} - W_{t_{k-1}}$ unabhängig.
- (ii) Die Inkremente sind stationär, d.h. die Verteilung von $W_t - W_s$ für $0 \leq s < t$ ist identisch mit der Verteilung von $W_{t+\tau} - W_{s+\tau}$ für $\tau \geq 0$. Die Verteilung von $W_t - W_s$ hängt somit nur von $t - s$ ab.

(iii) $\mathbb{E}(W_t - W_s) = 0$ für alle $0 \leq s < t$.

Anstatt auf der Ebene der Differentialgleichung (3.8) Forderungen direkt an einen Zufallsprozess zu stellen, resultieren in der Formulierung (3.10) Bedingungen an die Inkremente eines Zufallsprozesses.

Der einzige Zufallsprozess, der die neuen drei Forderungen erfüllt und (fast sicher) stetige Pfade besitzt, ist der in Abschnitt 2.3 eingeführte Wiener Prozess $(W_t)_{t \geq 0}$. Dessen Pfade sind jedoch (fast sicher) nirgends differenzierbar und besitzen in jedem offenen Zeitintervall eine unbeschränkte Variation.

Aus dem Ansatz (3.10) folgt die Gleichung

$$X_{j+1} = X_j + b(t_j, X_j)\Delta t_j + \sigma(t_j, X_j)\Delta W_j$$

und durch rekursives Einsetzen

$$X_k = X_0 + \sum_{j=0}^{k-1} b(t_j, X_j)\Delta t_j + \sum_{j=0}^{k-1} \sigma(t_j, X_j)\Delta W_j.$$

Im Grenzübergang $\Delta t_j \rightarrow 0$, siehe (3.9), möchten wir eine Darstellung der Form

$$X_t = X_0 + \int_0^t b(s, X_s) ds + \int_0^t \sigma(s, X_s) dW_s \quad (3.11)$$

erhalten. Man beachte, dass diese Gleichung für jedes Ergebnis $\omega \in \Omega$ separat besteht, d.h. ausführlicher lautet die Notation

$$X_t(\omega) = X_0(\omega) + \int_0^t b(s, X_s(\omega)) ds + \int_0^t \sigma(s, X_s(\omega)) dW_s(\omega).$$

Für einen stetigen Pfad $X_t(\omega)$ und stetige Funktion b existiert das erste Integral in der Formulierung (3.11) als Riemann-Integral. Trotz $X_t(\omega)$ und σ stetig kann das zweite Integral in (3.11) nicht als Riemann-Stieltjes-Integral definiert werden, da die Pfade des Wiener Prozesses in jedem Zeitintervall eine unbeschränkte Variation besitzen. Das Riemann-Stieltjes-Integral bezeichnet einen Grenzwert von reellen Zahlen. In \mathbb{R} existiert nur ein Konvergenzbegriff, während in Funktionenräumen im allgemeinen mehrere Grenzwertbildungen diskutiert werden können. Daher ist eine alternative Definition des Integrals in (3.11) durchzuführen, bei der eine Konvergenz auf der gesamten Menge Ω heranzuziehen ist.

3.3 Konstruktion des Itô-Integrals

Die hier gegebene Herleitung des stochastischen Integrals folgt der Vorgehensweise in [14]. Nachfolgende Definitionen werden benötigt.

Definition 3.1 Sei $[a, b] \subset [0, \infty)$ gegeben. Ist \mathcal{A} eine σ -Algebra über Ω , dann heißt ein System $\{\mathcal{F}_q : q \in [a, b]\}$ eine Filtration in \mathcal{A} , wenn gilt:

- (i) \mathcal{F}_q ist eine σ -Algebra über Ω für jedes $q \in [a, b]$,
- (ii) $\mathcal{F}_q \subseteq \mathcal{A}$ für alle $q \in [a, b]$,
- (iii) $\mathcal{F}_q \subseteq \mathcal{F}_r$ für alle $q, r \in [a, b]$ mit $q \leq r$.

Eine triviale Filtration ist $\mathcal{F}_q = \mathcal{A}$ für alle q , welche jedoch nicht verwendet wird.

Definition 3.2 Ein Zufallsprozess $(X_t)_{t \in [a, b]}$ auf einem Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) mit $X_t : \Omega \rightarrow \mathbb{R}$ heißt einer gegebenen Filtration $\{\mathcal{F}_q : q \in [a, b]\}$ adaptiert, wenn X_t jeweils $(\mathcal{F}_t, \mathcal{B}(\mathbb{R}))$ -messbar für alle $t \in [a, b]$ ist.

Umgekehrt hat man zu einem beliebigen Zufallsprozess $(X_t)_{t \in [a, b]}$ eines Wahrscheinlichkeitsraums (Ω, \mathcal{A}, P) die kanonische Filtration, bei der \mathcal{F}_t die kleinste σ -Algebra ist, bezüglich der alle Zufallsvariablen X_s für $a \leq s \leq t$ messbar sind. Diese σ -Algebra ist somit der Durchschnitt aller σ -Algebren, die die Mengen

$$X_s^{-1}(B) \quad \text{für alle } a \leq s \leq t \quad \text{und alle } B \in \mathcal{B}(\mathbb{R})$$

enthalten. Wir schreiben als Abkürzung $\mathcal{F}_t = \sigma(X_s; a \leq s \leq t)$ für die kanonische Filtration. Die drei Bedingungen aus Definition 3.1 sind für die kanonische Filtration erfüllt. Im Fall des Wiener Prozesses $(W_t)_{t \geq 0}$ gilt für die kanonische Filtration zusätzlich $\mathcal{F}_q \subset \mathcal{F}_r$ für $q < r$, d.h. eine aufsteigende Folge liegt vor.

Um einen geeigneten Integralbegriff zu konstruieren, müssen wir uns auf folgende Menge von Funktionen einschränken.

Definition 3.3 Sei $(W_t)_{t \geq 0}$ ein Wiener Prozess auf einem Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) und $[a, b] \subset [0, \infty)$ ein Intervall. Die Menge $\mathcal{V} = \mathcal{V}(a, b)$ bestehe genau aus den Funktionen $f : [a, b] \times \Omega \rightarrow \mathbb{R}$ mit den Eigenschaften:

- (i) f ist $(\mathcal{B}([a, b]) \otimes \mathcal{A}, \mathcal{B}(\mathbb{R}))$ -messbar,
- (ii) $f \in L^2([a, b] \times \Omega, \mathcal{B}([a, b]) \otimes \mathcal{A}, \lambda \otimes P)$ mit dem Lebesgue-Maß λ auf $[a, b]$,
- (iii) $(f(t, \cdot))_{t \in [a, b]}$ ist der Filtration $\mathcal{F}_t := \sigma(W_s : 0 \leq s \leq t)$ adaptiert.

Ist $f \in \mathcal{V}$, dann liefert mit der Eigenschaft (ii) der Satz von Fubini-Tonelli wegen $|f|^2 \in L^1([a, b] \times \Omega, \mathcal{B}([a, b]) \otimes \mathcal{A}, \lambda \otimes P)$ die Existenz der Integrale

$$\int_{[a, b] \times \Omega} f(t, \omega)^2 d(\lambda \otimes P) = \mathbb{E} \left(\int_a^b f(t, \cdot)^2 dt \right) = \int_a^b \mathbb{E}(f(t, \cdot)^2) dt. \quad (3.12)$$

Insbesondere muss nur eines dieser Integrale existieren und endlich sein, damit alle existieren und identisch sind. Die Gleichheit (3.12) bedeutet die Vertauschbarkeit der Integrationsreihenfolge

$$\int_{\Omega} \left(\int_a^b f(t, \omega)^2 dt \right) dP(\omega) = \int_a^b \left(\int_{\Omega} f(t, \omega)^2 dP(\omega) \right) dt.$$

Bei der Konstruktion von stochastischen Integralen sind Folgen von Zufallsvariablen auf dem Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) zu betrachten.

Definition 3.4 *Eine Folge $(X_n)_{n \in \mathbb{N}}$ aus reellen Zufallsvariablen auf einem Raum (Ω, \mathcal{A}, P) konvergiert im quadratischen Mittel, d.h. in $L^2(\Omega)$, gegen eine Zufallsvariable \hat{X} auf dem Raum, wenn gilt*

$$\lim_{n \rightarrow \infty} \mathbb{E} \left((X_n(\cdot) - \hat{X}(\cdot))^2 \right) = \lim_{n \rightarrow \infty} \int_{\Omega} (X_n(\omega) - \hat{X}(\omega))^2 dP(\omega) = 0.$$

Analog zur Konstruktion des Lebesgue-Integrals definieren wir stochastische Integral nur für eine Menge aus *Elementarfunktionen*. Für allgemeines $f \in \mathcal{V}$ wird dann eine Folge aus Elementarfunktionen konstruiert, die gegen f konvergiert. Der Grenzwert der korrespondierenden Integrale ist dann das Integral der Funktion f .

Die Elementarfunktionen sollen die Gestalt

$$\phi : [0, \infty) \times \Omega \rightarrow \mathbb{R}, \quad \phi(t, \omega) = \sum_{j \geq 0} e_j(\omega) \cdot \chi_{[j2^{-n}, (j+1)2^{-n})}(t) \quad (3.13)$$

besitzen mit der charakteristischen Funktion

$$\chi_M(t) = \begin{cases} 1 & \text{für } t \in M \\ 0 & \text{sonst} \end{cases} \quad (3.14)$$

für Teilmengen $M \subseteq \mathbb{R}$ und den Zufallsvariablen $e_j : \Omega \rightarrow \mathbb{R}$ zu jeweils einem festen $n \in \mathbb{N}$. In der Darstellung (3.13) sind die Zeitvariable t und die Ergebnisse ω separiert. Man beachte, dass für festes $t \in [0, \infty)$ höchstens ein Summand ungleich null ist, d.h. die Reihe dient nur als Schreibweise und benötigt keine Konvergenzanalyse. Offensichtlich bilden die Elementarfunktionen einen Vektorraum

über \mathbb{R} . Eine sinnvolle Definition des Integrals bezüglich des Wiener Prozesses (bzw. Brownsche Bewegung) ist

$$\int_a^b \phi(t, \omega) dW_t(\omega) = \sum_{j \geq 0} e_j(\omega) \left(W_{t_{j+1}^{(n)}}(\omega) - W_{t_j^{(n)}}(\omega) \right) \quad (3.15)$$

mit

$$t_k^{(n)} := \begin{cases} a & \text{für } k2^{-n} < a \\ k2^{-n} & \text{für } a \leq k2^{-n} \leq b \\ b & \text{für } k2^{-n} > b. \end{cases} \quad (3.16)$$

Wir lassen den oberen Index n im folgenden weg. Mit dieser Definition sind in der Reihe nur endlich viele Glieder ungleich null. Das Integral (3.15) ist somit eine Zufallsvariable auf Ω . Bei der Approximation eines Zufallsprozesses durch Elementarfunktionen ergeben sich jedoch Probleme. Betrachten wir als zu approximierenden Prozess den Wiener Prozess selbst, d.h. $\phi(t, \omega) \approx W_t(\omega)$. Wir machen den Ansatz

$$e_j(\omega) := (1 - \tau)W_{t_j}(\omega) + \tau W_{t_{j+1}}(\omega) \quad \text{für } j \geq 0 \quad (3.17)$$

mit einem festen $\tau \in [0, 1]$. Für das Integral (3.15) erhalten wir damit

$$\begin{aligned} I_\tau &= \sum_{j \geq 0} ((1 - \tau)W_{t_j} + \tau W_{t_{j+1}}) (W_{t_{j+1}} - W_{t_j}) \\ &= \sum_{j \geq 0} \frac{1}{2}(W_{t_{j+1}}^2 - W_{t_j}^2) + (\tau - \frac{1}{2})(W_{t_{j+1}} - W_{t_j})^2 \\ &= \frac{1}{2}(W_b^2 - W_a^2) + (\tau - \frac{1}{2}) \sum_{j \geq 0} (W_{t_{j+1}} - W_{t_j})^2. \end{aligned} \quad (3.18)$$

Der Erwartungswert hierzu liefert

$$\begin{aligned} \mathbb{E}(I_\tau) &= \frac{1}{2}(\mathbb{E}(W_b^2) - \mathbb{E}(W_a^2)) + (\tau - \frac{1}{2}) \sum_{j \geq 0} \mathbb{E}((W_{t_{j+1}} - W_{t_j})^2) \\ &= \frac{1}{2}(b - a) + (\tau - \frac{1}{2}) \sum_{j \geq 0} t_{j+1} - t_j \\ &= \frac{1}{2}(b - a) + (\tau - \frac{1}{2})(b - a) = \tau(b - a). \end{aligned} \quad (3.19)$$

Somit sind die Erwartungswerte für verschiedene τ unterschiedlich und die Grenzwerte der Folgen können nicht identisch sein. Genauer kann man zeigen, dass die Folge der Zufallsvariablen (3.18) im quadratischen Mittel (siehe Definition 3.4) gegen

$$\hat{I}_\tau = \frac{1}{2}(W_b^2 - W_a^2) + (\tau - \frac{1}{2})(b - a) \quad (3.20)$$

konvergiert. Obwohl (3.17) für alle $\tau \in [0, 1]$ eine sinnvolle Approximation des Zufallsprozesses liefert, fallen die korrespondierenden Integralwerte (3.20) auseinander. Analog machen wir als Approximation den alternativen Ansatz

$$e_j(\omega) := W_{(1-\tau)t_j + \tau t_{j+1}}(\omega) \quad \text{für } j \geq 0 \quad (3.21)$$

und $\tau \in [0, 1]$ anwenden und erhält als Integral

$$I_\tau = \sum_{j \geq 0} W_{(1-\tau)t_j + \tau t_{j+1}} (W_{t_{j+1}} - W_{t_j}). \quad (3.22)$$

Für den Ansatz (3.21) folgt als Erwartungswert von (3.22) hier mit (2.6)

$$\begin{aligned} \mathbb{E}(I_\tau) &= \sum_{j \geq 0} \mathbb{E}(W_{(1-\tau)t_j + \tau t_{j+1}} W_{t_{j+1}}) - \mathbb{E}(W_{(1-\tau)t_j + \tau t_{j+1}} W_{t_j}) \\ &= \sum_{j \geq 0} \min\{(1-\tau)t_j + \tau t_{j+1}, t_{j+1}\} - \min\{(1-\tau)t_j + \tau t_{j+1}, t_j\} \\ &= \sum_{j \geq 0} (1-\tau)t_j + \tau t_{j+1} - t_j = \tau \sum_{j \geq 0} t_{j+1} - t_j = \tau(b-a). \end{aligned} \quad (3.23)$$

Das Ergebnis ist mit (3.19) identisch. Somit hängt das Integral von der Stelle τ ab, an der der Zufallsprozess für eine konstante Approximation ausgewertet wird.

Für eine allgemeine messbare Funktion $f : [a, b] \times \Omega \rightarrow \mathbb{R}$, deren Pfade $f(\cdot, \omega)$ fast alle stetig sind, kann das Integral definiert werden durch

$$I_\tau^{(n)}(\omega) = \sum_{j \geq 0} f((1-\tau)t_j + \tau t_{j+1}, \omega) (W_{t_{j+1}}(\omega) - W_{t_j}(\omega)). \quad (3.24)$$

Die Abhängigkeit von n besteht durch die Gitterpunkte (3.16) in der Zeit. Für $\tau = 0$ und $f \in \mathcal{V}$ ist die Konvergenz in $L^2(\Omega)$ gesichert. Bei $\tau > 0$ gilt dies nicht immer bzw. hängt von f ab. Zudem muss die zu (3.24) korrespondierende Folge aus Elementarfunktionen für $\tau \in [0, 1]$ nicht gegen f in $L^2([a, b] \times \Omega)$ konvergieren. Der Grenzwert für das Integral bei n gegen unendlich (wenn er existiert) hängt im allgemeinen von der Wahl der Zwischenstellen $\tau \in [0, 1]$ ab. Man kann zwei qualitativ unterschiedliche Fälle klassifizieren:

- (i) $\tau = 0$: Der jeweils linke Intervallrand wird verwendet und der korrespondierende Begriff wird *Itô-Integral* genannt. Diese Situation entspricht einem nicht in der Zeit vorgreifen. Das Integral ist in diesem Fall ein *Martingal*, siehe [14].
- (ii) $0 < \tau \leq 1$: Hier geht die günstige Martingal-Eigenschaft verloren. In diesem Fall ist die Wahl $\tau = \frac{1}{2}$ am vorteilhaftesten, da zum einen sinnvolle Interpretationen dieses Modells möglich sind und zum anderen einfachere Rechenregeln als im Fall (i) vorliegen. Es entsteht das *Stratonovich-Integral*.

Die Integraldefinition (3.15) ist für Elementarfunktionen (3.13) mit beliebigen Zufallsvariablen e_j möglich. Wichtige Eigenschaften des Integrals entstehen jedoch durch die zusätzliche Forderung $\phi \in \mathcal{V}$ für die Elementarfunktionen. Wesentlich

dafür ist die Bedingung (iii) aus Definition 3.3. Eine reelle Zufallsvariable h ist \mathcal{F}_t -messbar genau dann, wenn man sie darstellen kann als punktweisen Grenzwert von Funktionen der Form

$$g_1(W_{t_1})g_2(W_{t_2})\cdots g_k(W_{t_k}) \quad \text{für } k = 1, 2, \dots$$

mit stetigen beschränkten Funktionen g_1, \dots, g_k für $t_j \leq t$ und $j \leq k$. Anschaulich bedeutet die \mathcal{F}_t -Messbarkeit, dass sich h aus den Werten des Wiener Prozesses $W_s(\omega)$ für $s \leq t$ rekonstruieren lässt. Es ist also $h_1 := W_{t/2}$ hier \mathcal{F}_t -messbar, während dies für $h_2 := W_{2t}$ nicht gilt. Mit $\phi \in \mathcal{V}$ für Elementarfunktionen (3.13) kann man beweisen (siehe [5]), dass e_j jeweils von $W_{t_{i+1}} - W_{t_i}$ in gewissem Sinne unabhängig ist für alle $i \geq j$. In der Definition (3.17) ist die korrespondierende Elementarfunktion für $\tau = 0$ in \mathcal{V} , für $\tau > 0$ jedoch nicht.

Eine wichtige Folgerung ist die *Itô-Isometrie*, die zwischen den normierten Räumen $L^2([a, b] \times \Omega, \mathcal{B}([a, b]) \otimes \mathcal{A}, \lambda \otimes P)$ und $L^2(\Omega, \mathcal{A}, P)$ besteht.

Satz 3.1 (Itô-Isometrie) *Ist $\phi \in \mathcal{V}$ eine beschränkte Elementarfunktion, dann gilt*

$$\mathbb{E} \left[\left(\int_a^b \phi(t, \cdot) dW_t(\cdot) \right)^2 \right] = \mathbb{E} \left[\int_a^b \phi(t, \cdot)^2 dt \right]. \quad (3.25)$$

Beweis:

Zur Abkürzung sei $\Delta W_j := W_{t_{j+1}} - W_{t_j}$. Die Elementarfunktion ϕ besitzt die Gestalt (3.13). Da ϕ beschränkt ist, existieren alle Momente der Zufallsvariablen e_j . Es sind in gewissem Sinn $e_i e_j \Delta W_i$ und ΔW_j unabhängig für $i < j$ sowie e_j^2 und $(\Delta W_j)^2$ unabhängig für alle j wegen Eigenschaft (iii) aus Definition 3.3. Es folgt

$$\begin{aligned} & \mathbb{E} \left[\left(\int_a^b \phi dW_t \right)^2 \right] = \mathbb{E} \left[\left(\sum_{j \geq 0} e_j \Delta W_j \right)^2 \right] = \mathbb{E} \left[\sum_{i, j \geq 0} (e_i e_j \Delta W_i \Delta W_j) \right] \\ &= \sum_{i, j \geq 0} \mathbb{E} (e_i e_j \Delta W_i \Delta W_j) = \sum_{j \geq 0} \mathbb{E} (e_j^2 (\Delta W_j)^2) + 2 \sum_{i < j} \mathbb{E} (e_i e_j \Delta W_i \Delta W_j) \\ &= \sum_{j \geq 0} \mathbb{E} (e_j^2) \underbrace{\mathbb{E}((\Delta W_j)^2)}_{=t_{j+1}-t_j} + \sum_{i \neq j} \mathbb{E}(e_i e_j \Delta W_i) \underbrace{\mathbb{E}(\Delta W_j)}_{=0} = \sum_{j \geq 0} \mathbb{E} (e_j^2) (t_{j+1} - t_j) \\ &= \mathbb{E} \left[\sum_{j \geq 0} e_j^2 (t_{j+1} - t_j) \right] = \mathbb{E} \left[\int_a^b \phi^2 dt \right], \end{aligned}$$

da in den Reihen nur endlich viele Summanden ungleich null sind. \square

Allgemeiner kann man noch zeigen (siehe [5]), dass für beschränkte Elementarfunktionen $\phi, \psi \in \mathcal{V}$ gilt

$$\mathbb{E} \left[\left(\int_a^b \phi(t, \cdot) dW_t(\cdot) \right) \left(\int_a^b \psi(t, \cdot) dW_t(\cdot) \right) \right] = \mathbb{E} \left[\int_a^b \phi(t, \cdot) \psi(t, \cdot) dt \right],$$

wobei diese Ausdrücke jeweils die Skalarprodukte in den Räumen darstellen.

Die Definition des Itô-Integrals für beliebige Funktionen $f \in \mathcal{V}$ wird nun in drei Schritten durchgeführt.

1. Schritt:

Sei $g \in \mathcal{V}$ eine beschränkte Funktion und jeder Pfad $g(\cdot, \omega)$ sei stetig. Wir definieren eine Folge $(\phi_n)_{n \in \mathbb{N}}$ aus Elementarfunktionen durch

$$\phi_n(t, \omega) = \sum_{j \geq 0} g(t_j, \omega) \chi_{[t_j, t_{j+1})}(t). \quad (3.26)$$

Da $g \in \mathcal{V}$ gefordert ist, liegt auch die Bedingung (iii) aus Definition 3.3 für diese Folge vor. Die Forderung (i) folgt direkt. Da die Pfade von g stetig sind, folgt $g(\cdot, \omega) \in L^2([a, b])$ und

$$\lim_{n \rightarrow \infty} \int_a^b (g(t, \omega) - \phi_n(t, \omega))^2 dt = 0 \quad \text{für jedes } \omega \in \Omega.$$

Es sei $|g(t, \omega)| \leq \hat{g}$ für alle $t \in [a, b]$ und $\omega \in \Omega$. Dann ist auch $|\phi_n(t, \omega)| \leq \hat{g}$. Eine grobe Abschätzung liefert

$$\left| \int_a^b (g - \phi_n)^2 dt \right| \leq \int_a^b (|g| + |\phi_n|)^2 dt \leq 4(b-a)\hat{g}.$$

Der Satz von Fubini-Tonelli impliziert die Bedingung (ii) aus Definition 3.3. Insgesamt folgt mit dem Satz über dominierte Konvergenz, dass

$$\lim_{n \rightarrow \infty} \mathbb{E} \left(\int_a^b (g(t, \omega) - \phi_n(t, \omega))^2 dt \right) = 0.$$

2. Schritt:

Sei $h \in \mathcal{V}$ beschränkt, d.h. $|h(t, \omega)| \leq \hat{h}$ für alle $t \in [a, b]$ und $\omega \in \Omega$. Für $n \in \mathbb{N}$ existiert eine nicht-negative stetige Funktion $\psi_n : \mathbb{R} \rightarrow \mathbb{R}$ mit Träger $\mathcal{T}(\psi_n) \subset (-\frac{1}{n}, 0)$ und

$$\int_{-\infty}^{+\infty} \psi_n(x) dx = 1.$$

Wir definieren über eine Faltung die Folge

$$g_n(t, \omega) := \int_a^t \psi_n(s-t)h(s, \omega) ds. \quad (3.27)$$

Dann ist jeder Pfad $g_n(\cdot, \omega)$ stetig wegen der glättenden Wirkung der Faltung. Zudem gilt überall $|g_n(t, \omega)| \leq \hat{h}$. Da $h \in \mathcal{V}$ gefordert wird, kann (mit viel Aufwand) gezeigt werden, dass g_n aus (3.27) die Bedingung (iii) aus Definition 3.3 erfüllt. Zudem gilt punktweise

$$\lim_{n \rightarrow \infty} \int_a^b (g_n(t, \omega) - h(t, \omega))^2 dt = 0 \quad \text{für jedes } \omega \in \Omega,$$

da die Funktionenfolge $(\psi_n)_{n \in \mathbb{N}}$ die Identität immer besser approximiert. Weil h beschränkt ist, folgt wie im 1. Schritt die Aussage

$$\lim_{n \rightarrow \infty} \mathbb{E} \left(\int_a^b (g_n(t, \omega) - h(t, \omega))^2 dt \right) = 0$$

nach dem Satz über dominierte Konvergenz.

3. Schritt:

Nun sei nur $f \in \mathcal{V}$ gegeben. Wir definieren die Funktionenfolge

$$h_n(t, \omega) := \begin{cases} -n & \text{für } f(t, \omega) < -n \\ f(t, \omega) & \text{für } -n \leq f(t, \omega) \leq n \\ n & \text{für } f(t, \omega) > n. \end{cases} \quad (3.28)$$

Die Eigenschaften (i)-(iii) aus Definition 3.3 übertragen sich direkt von der Funktion f auf die Abbildungen h_n . Es gilt $h_n(t, \omega)^2 \leq h_m(t, \omega)^2 \leq f(t, \omega)^2$ für $t \in [a, b]$ und $\omega \in \Omega$ sowie $n \leq m$. Weiter ist punktweise

$$\lim_{m \rightarrow \infty} h_m(t, \omega)^2 = \sup_{m \in \mathbb{N}} h_m(t, \omega)^2 = f(t, \omega)^2.$$

Wir haben auf dem Raum $[a, b] \times \Omega$ mit der Bedingung (ii) aus Definition 3.3 und dem Satz über monotone Konvergenz (man beachte $h_m h_n \geq 0$ für alle m, n)

$$\begin{aligned} & \int (f - h_n)^2 d(\lambda \otimes P) \\ &= \int f^2 d(\lambda \otimes P) - 2 \int f h_n d(\lambda \otimes P) + \int h_n^2 d(\lambda \otimes P) \\ &= \int f^2 d(\lambda \otimes P) - 2 \int \sup_{m \in \mathbb{N}} h_m h_n d(\lambda \otimes P) + \int h_n^2 d(\lambda \otimes P) \\ &= \int f^2 d(\lambda \otimes P) - 2 \sup_{m \in \mathbb{N}} \int h_m h_n d(\lambda \otimes P) + \int h_n^2 d(\lambda \otimes P) \\ &\leq \int f^2 d(\lambda \otimes P) - 2 \int h_n^2 d(\lambda \otimes P) + \int h_n^2 d(\lambda \otimes P) \\ &= \int f^2 d(\lambda \otimes P) - \int h_n^2 d(\lambda \otimes P). \end{aligned}$$

Erneute Anwendung des Satzes über monotone Konvergenz liefert somit für (3.28)

$$\lim_{n \rightarrow \infty} \int (f - h_n)^2 d(\lambda \otimes P) = 0.$$

Zu jedem $f \in \mathcal{V}$ können wir also durch Abarbeiten der Schritte 1-3 eine Folge von beschränkten Elementarfunktionen $(\phi_n)_{n \in \mathbb{N}} \subset \mathcal{V}$ angeben, die in $L^2([a, b] \times \Omega)$ gegen f konvergiert, d.h.

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[\int_a^b (f - \phi_n)^2 dt \right] = 0.$$

Insbesondere liegt damit eine Cauchy-Folge in $L^2([a, b] \times \Omega)$ vor, d.h.

$$\mathbb{E} \left[\int_a^b (\phi_m - \phi_n)^2 dt \right] < \varepsilon$$

für alle $n, m \geq N(\varepsilon)$.

Definition 3.5 Besteht $(\phi_n)_{n \in \mathbb{N}} \subset \mathcal{V}$ aus beschränkten Elementarfunktionen, die gegen $f \in \mathcal{V}$ in $L^2([a, b] \times \Omega)$ konvergieren, dann lautet das Itô-Integral von f

$$\mathcal{I}[f](\omega) := \int_a^b f(t, \omega) dW_t(\omega) := \lim_{n \rightarrow \infty} \int_a^b \phi_n(t, \omega) dW_t(\omega). \quad (3.29)$$

Mit der Itô-Isometrie (3.25) aus Satz 3.1 gilt für $n, m \in \mathbb{N}$

$$\mathbb{E} \left[\left(\int_a^b \phi_n dW_t - \int_a^b \phi_m dW_t \right)^2 \right] = \mathbb{E} \left[\int_a^b |\phi_n - \phi_m|^2 dt \right].$$

Somit stellt (3.29) eine Cauchyfolge in $L^2(\Omega)$ dar, die gegen eine Zufallsvariable konvergiert, nämlich das Itô-Integral. Die Itô-Isometrie liefert weiter, dass der Grenzwert (3.29) unabhängig ist von der Wahl der Folge aus Elementarfunktionen, die gegen f konvergieren.

3.4 Eigenschaften des Itô-Integrals

Das Itô-Integral ist als Grenzwert von Zufallsvariablen wieder eine Zufallsvariable. Wichtige Eigenschaften des Itô-Integrals sind:

- (i) Linearität: Für $\alpha, \beta \in \mathbb{R}$ und $f, g \in \mathcal{V}$ gilt

$$\mathcal{I}[\alpha f + \beta g] = \alpha \mathcal{I}[f] + \beta \mathcal{I}[g].$$

(ii) Zerlegung: Mit $a < c < b$ folgt mit $f \in \mathcal{V}$ für fast alle $\omega \in \Omega$

$$\int_a^b f \, dW_t = \int_a^c f \, dW_t + \int_c^b f \, dW_t.$$

(iii) Messbarkeit: Für alle $f \in \mathcal{V}$ ist $\mathcal{I}[f]$ dann $(\mathcal{F}_b, \mathcal{B}(\mathbb{R}))$ -messbar.

(iv) Erwartungswert: Für alle $f \in \mathcal{V}$ gilt $\mathbb{E}(\mathcal{I}[f]) = 0$.

(v) Varianz: Für alle $f \in \mathcal{V}$ gilt $\text{Var}(\mathcal{I}[f]) = \|f\|_{L^2([a,b] \times \Omega)}^2$.

Die Aussagen (i) und (ii) sind analog zur üblichen Integration. Die Eigenschaften (iii) und (iv) stehen in Zusammenhang mit der Bedingung (iii) aus Definition 3.3. Bewiesen werden alle vier Aussagen, indem man sie für Elementarfunktionen nachweist und dann den Grenzwert gegen ein allgemeines $f \in \mathcal{V}$ bildet. Die Eigenschaft (v) folgt aus der Itô-Isometrie in Satz 3.2.

Satz 3.2 (Itô-Isometrie) *Für beliebiges $f \in \mathcal{V}$ gilt die Itô-Isometrie*

$$\mathbb{E} \left[\left(\int_a^b f \, dW_t \right)^2 \right] = \mathbb{E} \left[\int_a^b f^2 \, dt \right].$$

Beweis:

Für beliebiges $f \in \mathcal{V}$ folgt durch Grenzübergang

$$\begin{aligned} \mathbb{E} \left[\left(\int_a^b f \, dW_t \right)^2 \right] &= \mathbb{E} \left[\left(\lim_{n \rightarrow \infty} \int_a^b \phi_n \, dW_t \right)^2 \right] = \lim_{n \rightarrow \infty} \mathbb{E} \left[\left(\int_a^b \phi_n \, dW_t \right)^2 \right] \\ &= \lim_{n \rightarrow \infty} \mathbb{E} \left[\int_a^b \phi_n^2 \, dt \right] = \mathbb{E} \left[\int_a^b f^2 \, dt \right]. \end{aligned}$$

Die Vertauschung von Integration und Grenzwertbildung ist wegen der Konvergenz der Itô-Integrale zu den Elementarfunktionen in $L^2(\Omega)$ erlaubt. Analog folgt die letzte Gleichheit aus der Konvergenz von $(\phi_n)_{n \in \mathbb{N}}$ gegen f im Hilbert-Raum $L^2([a, b] \times \Omega)$. In der dritten Gleichheit geht die Itô-Isometrie für Elementarfunktionen aus Satz 3.1 ein. \square

Wir benötigen noch den folgenden Begriff. Zur Definition eines bedingten Erwartungswerts zu einem Zufallsprozess siehe Anhang B in [14].

Definition 3.6 Sei (Ω, \mathcal{A}, P) ein Wahrscheinlichkeitsraum und $(\mathcal{F}_t)_{t \in I}$ mit $I = [a, b]$ oder $I = [0, \infty)$ eine Filtration. Ein Zufallsprozess $(X_t)_{t \in I}$, der zu dieser Filtration adaptiert ist und $\mathbb{E}(|X_t|) < \infty$ für alle $t \in I$ erfüllt, heißt ein Martingal, wenn für den bedingten Erwartungswert fast sicher gilt

$$\mathbb{E}(X_t | \mathcal{F}_s) = X_s$$

für alle $s < t$.

Ein Martingal ist ein Zufallsprozess, der in gewisser Weise nicht in die Zukunft vorgreift. Diese Eigenschaft entspricht in vielen Aufgabenstellungen den physikalischen Rahmenbedingungen. Eine Konsequenz der Martingal-Eigenschaft ist

$$\mathbb{E}(X_t) = \mathbb{E}(\mathbb{E}(X_t | \mathcal{F}_s)) = \mathbb{E}(X_s)$$

für alle $s \leq t$. Ein Beispiel für ein Martingal ist ein Wiener Prozess. Für weitere Erläuterungen zum Begriff des Martingals siehe auch [3].

Definition 3.3 impliziert, dass mit $f \in \mathcal{V}(a, b)$ auch $f \in \mathcal{V}(a, c)$ und $f \in \mathcal{V}(c, b)$ für alle $a < c < b$ gilt. Dadurch wird die Abbildung

$$M : [a, b] \rightarrow L^2(\Omega), \quad t \mapsto \int_a^t f(s, \omega) dW_s(\omega) \quad (3.30)$$

wohldefiniert für festes $f \in \mathcal{V}(a, b)$. Es ist $(M_t)_{t \in [a, b]}$ ein Zufallsprozess. Man kann zeigen (siehe [14]), dass eine Version (siehe (2.7)) von $(M_t)_{t \in [a, b]}$ existiert, bei der fast alle Pfade stetig sind. O.E.d.A. sei mit $(M_t)_{t \in [a, b]}$ im folgenden diese Version bezeichnet. Zudem ist $(M_t)_{t \in [a, b]}$ ein Martingal gemäß Definition 3.6. Folglich wird in stochastischen Differentialgleichungen bzw. Integralgleichungen meist der Itô-Integralbegriff für den letzten Term in (3.11) verwendet, da die anderen stochastischen Integralbegriffe keine Martingale liefern.

3.5 Verallgemeinerungen des Itô-Integrals

Wir verallgemeinern das Itô-Integral für eine größere Menge von Funktionen und insbesondere mehrdimensionale Funktionen.

Erweiterung des Funktionenraums

Wir erweitern den Funktionenraum \mathcal{V} aus Definition 3.3, wobei die Bedingungen (ii) und (iii) leicht abgeschwächt werden.

Definition 3.7 Sei $(W_t)_{t \geq 0}$ ein Wiener Prozess auf einem Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) und $[a, b] \subset [0, \infty)$ ein Intervall. Die Menge $\mathcal{W} = \mathcal{W}(a, b)$ bestehe genau aus den Funktionen $f : [a, b] \times \Omega \rightarrow \mathbb{R}$ mit den Eigenschaften:

- (i) f ist $(\mathcal{B}([a, b]) \otimes \mathcal{A}, \mathcal{B}(\mathbb{R}))$ -messbar,
- (ii)' $P \left[\int_a^b f(t, \omega)^2 dt < \infty \right] := P \left(\left\{ \omega \in \Omega : \int_a^b f(t, \omega)^2 dt < \infty \right\} \right) = 1$,
- (iii)' $(f(t, \cdot))_{t \in [a, b]}$ ist einer Filtration \mathcal{G}_t adaptiert und der Wiener Prozess $(W_t)_{t \geq 0}$ ist ein Martingal bezüglich dieser Filtration.

Zu (ii)': Aus der Bedingung (ii) in Definition 3.3 folgt die Bedingung (ii)' in Definition 3.7, was leicht durch Kontraposition zu sehen ist. Die Umkehrung gilt jedoch nicht. Die Verallgemeinerung (ii)' bewirkt, dass Grenzwerte von Zufallsvariablen nicht mehr im quadratischen Mittel zu betrachten sind sondern wie folgt.

Definition 3.8 Eine Folge $(X_n)_{n \in \mathbb{N}}$ bestehend aus reellen Zufallsvariablen auf einem Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) konvergiert P -stochastisch gegen eine Zufallsvariable \hat{X} auf dem Wahrscheinlichkeitsraum, wenn für jedes $\delta > 0$ gilt

$$\lim_{n \rightarrow \infty} P \left[|X_n - \hat{X}| > \delta \right] = \lim_{n \rightarrow \infty} P \left(\left\{ \omega \in \Omega : |X_n(\omega) - \hat{X}(\omega)| > \delta \right\} \right) = 0.$$

Man bezeichnet dann \hat{X} als den stochastischen Grenzwert.

Zu $f \in \mathcal{W}$ existiert nämlich eine Folge aus Elementarfunktionen $(\phi_n)_{n \in \mathbb{N}} \subset \mathcal{W}$ mit

$$\lim_{n \rightarrow \infty} \int_a^b (\phi_n(t, \omega) - f(t, \omega))^2 dt = 0$$

als stochastischer Grenzwert. Die Folge der korrespondierenden Itô-Integrale besitzt dann ebenfalls einen stochastischen Grenzwert, der nur von f und nicht der Folge $(\phi_n)_{n \in \mathbb{N}}$ abhängt. Dadurch kann man für $f \in \mathcal{W}$ setzen

$$\int_a^b f(t, \omega) dW_t(\omega) = \lim_{n \rightarrow \infty} \int_a^b \phi_n(t, \omega) dW_t(\omega)$$

mit Konvergenz im Sinn von Definition 3.8. Zum korrespondierenden Zufallsprozess (3.30) existiert wieder eine Version, bei der fast alle Pfade stetig sind. Jedoch ist (3.30) kein Martingal, sondern nur ein lokales Martingal.

Zu (iii)': Dass der Wiener Prozess zur Filtration ein Martingal ist impliziert $\mathcal{F}_t \subseteq \mathcal{G}_t$ für alle $t \geq 0$ mit der kanonischen Filtration $\mathcal{F}_t = \sigma(W_s; 0 \leq s \leq t)$. Der Zufallsprozess $(f(t, \cdot))_{t \in [a, b]}$ darf somit von mehr abhängen als nur vom Wiener Prozess, solange der Wiener Prozess ein Martingal bleibt. In dieser Situation kann ein Itô-Integral wie zuvor konstruiert werden und die wesentlichen Eigenschaften bleiben bestehen.

Mehrdimensionale Itô-Integration

Mehrdimensionale Itô-Integrale werden zur Definition von Systemen aus stochastischen Differentialgleichungen benötigt, welche mehr als eine Rauschquelle einbeziehen.

Wir verwenden einen n -dimensionalen Wiener Prozess $(W_1(t), \dots, W_n(t))$ auf einem Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) , siehe Definition 2.4. Dann sei $(\mathcal{F}_t^{(n)})_{t \geq 0}$ die zugehörige (kanonische) Filtration, d.h. \mathcal{F}_t ist die kleinste σ -Algebra, welche alle Mengen

$$W_i(s)^{-1}(B) \in \mathcal{A} \quad \text{für } 0 \leq s \leq t \quad \text{und } B \in \mathcal{B}(\mathbb{R}) \quad \text{und } i = 1, \dots, n$$

enthält. Dann ist jeder der eindimensionalen Wiener Prozesse $W_i(t)$ ein Martingal bezüglich dieser Filtration. Die Erweiterung (iii)' aus Definition 3.7 erlaubt es somit, ein Itô-Integral für $\mathcal{F}_t^{(n)}$ -adaptierte Funktionen zu erhalten. Dies ermöglicht wiederum die Auswertung von Integralen wie

$$\int_a^b W_2 \, dW_1, \quad \int_a^b e^{W_1^2 + W_2^2} \, dW_2, \quad \text{etc.}$$

Wir führen ein entsprechendes *mehrdimensionales Itô-Integral* ein.

Definition 3.9 Sei (W_1, \dots, W_n) ein n -dimensionaler Wiener Prozess. Die Menge $\mathcal{V}^{m \times n} = \mathcal{V}^{m \times n}(a, b)$ umfasst die $(m \times n)$ -Matrizen $v = (v_{ij})$, deren Einträge Funktionen $v_{ij} : [a, b] \times \Omega \rightarrow \mathbb{R}$ sind, welche Bedingung (i) und (ii) aus Definition 3.3 sowie Bedingung (iii)' aus Definition 3.7 mit der Filtration $\mathcal{G}_t = \mathcal{F}_t^{(n)}$ erfüllen. Das korrespondierende mehrdimensionale Itô-Integral

$$\int_a^b v \, dW = \int_a^b \begin{pmatrix} v_{11} & \cdots & v_{1n} \\ \vdots & & \vdots \\ v_{m1} & \cdots & v_{mn} \end{pmatrix} \begin{pmatrix} dW_1 \\ \vdots \\ dW_n \end{pmatrix} \quad (3.31)$$

besteht aus den Komponenten

$$\sum_{j=1}^n \int_a^b v_{ij}(t, \omega) \, dW_j(t, \omega) \quad \text{für } i = 1, \dots, m. \quad (3.32)$$

Man beachte, dass sich jede Komponente des mehrdimensionalen Integrals qualitativ gleich verhält. Funktionen, für die das Integral über die gesamte Zeit $t \geq 0$ existiert, werden bezeichnet mit

$$\mathcal{V}^{m \times n}(0, \infty) := \bigcap_{T > 0} \mathcal{V}^{m \times n}(0, T).$$

Desweiteren kann auch eine mehrdimensionale Form der Funktionenmenge \mathcal{W} aus Definition 3.7 konstruiert werden.

Definition 3.10 *Der Funktionenraum $\mathcal{W}^{m \times n} = \mathcal{W}^{m \times n}(a, b)$ bezeichnet hier alle $(m \times n)$ -Matrizen, deren Komponenten Funktionen aus $\mathcal{W}(a, b)$ gemäß Definition 3.7 sind.*

Für Funktionen aus $\mathcal{W}^{m \times n}$ ergibt sich das zugehörige mehrdimensionale Itô-Integral wieder aus (3.31) und (3.32).

3.6 Stratonovich-Integral

Ein alternativer Integralbegriff für den letzten Term in der stochastischen Differentialgleichung bzw. Integralgleichung (3.11) entsteht, wenn man statt dem linken Intervallrand $\tau = 0$ den Intervallmittelpunkt $\tau = \frac{1}{2}$ in (3.24) wählt. Es folgt das *Stratonovich-Integral*. Ein Vorteil dieser Variante besteht darin, dass zugehörige Differentiationsregeln die übliche Struktur wie im deterministischen Fall besitzen. Beispielsweise liefert (3.23) für $\tau = \frac{1}{2}$ eine Formel für das Integral bezüglich des Wiener Prozesses wie im klassischen Kalkül. Der korrespondierende Zufallsprozess (3.30) ist jedoch kein Martingal.

Das Stratonovich-Integral kann für alle Funktionen $f \in \mathcal{V}$ laut Definition 3.3 gebildet werden, siehe [13]. Sind zudem die Pfade von f fast sicher stetig, dann ist das Stratonovich-Integral darstellbar als der Grenzwert (3.24) in $L^2(\Omega)$, d.h. im quadratischen Mittel.

Aus dem 1. Schritt bei der Konstruktion des Itô-Integrals folgt, dass für einen beschränkten Zufallsprozess $f \in \mathcal{V}$, bei dem die Pfade fast sicher stetig sind, dann die Integralbegriffe für alle $\tau \in [0, 1]$ identisch sind. Die Approximation (3.26) liefert dann nämlich für alle Zwischenstellen die gleichen Eigenschaften. Daher muss notwendigerweise ein Prozess unbeschränkt sein oder unstetige Pfade besitzen, damit die Integralbegriffe verschieden sind.

Stratonovich-Integrale werden durch eine spezielle Notation gekennzeichnet. Die Schreibweise in der Integralgleichung (3.11) bezeichnet stets das Itô-Integral für

den letzten Term, d.h. eine Itô-Differentialgleichung. Dagegen schreibt sich die korrespondierende Stratonovich-Differentialgleichung

$$X_t = X_0 + \int_0^t b(s, X_s) ds + \int_0^t \sigma(s, X_s) \circ dW_s \quad (3.33)$$

mit dem Stratonovich-Integral als letzten Term.

Eine sinnvolle physikalische Interpretation ist für das Stratonovich-Integral wie folgt möglich. Fast alle Pfade $W_t(\omega)$ eines Wiener Prozesses sind stetig. Auf einem kompakten Intervall $[a, b]$ kann man daher für fast alle Pfade eine Folge von glatten Funktionen $(W_t^{(n)}(\omega))_{n \in \mathbb{N}}$ (z.B. nach dem Weierstraßschen Approximationssatz durch Polynome) angeben mit

$$\lim_{n \rightarrow \infty} \max_{t \in [a, b]} |W_t^{(n)}(\omega) - W_t(\omega)| = 0 \quad (3.34)$$

für jeweils fast alle $\omega \in \Omega$. Die stochastische Differentialgleichung (3.11) wird dann ersetzt durch die Folge von (deterministischen) Differentialgleichungen

$$\frac{dX_t^{(n)}}{dt}(\omega) = b(t, X_t^{(n)}(\omega)) + \sigma(t, X_t^{(n)}(\omega)) \frac{dW_t^{(n)}}{dt}(\omega)$$

für jeweils fast alle $\omega \in \Omega$. Anfangswerte sind $X_0^{(n)}(\omega) = X_0(\omega)$. Da die rechten Seiten für festes ω stetig sind, existiert punktweise eine Lösung im deterministischen Sinn. Es kann gezeigt werden, dass die Folge der Lösungen $(X_t^{(n)}(\omega))_{n \in \mathbb{N}}$ gleichmäßig in t auf $[a, b]$ (wie in (3.34)) gegen eine Zufallsvariable $X_t(\omega)$ konvergiert. Weiter folgt, dass diese punktweise für ω definierte Zufallsvariable X_t dann der Lösung zur Stratonovich-Differentialgleichung (3.33) entspricht.

Diese Interpretation legt nahe, dass das Stratonovich-Integral der geeignete Begriff zur Modellierung von Rauschen ist. Eine analoge Interpretation für das Itô-Integral existiert nämlich nicht. Jedoch besitzt das Itô-Integral die Martingaleigenschaft, die unabhängig von der Modellierung des Rauschens physikalisch sinnvoll ist. Man beachte, dass zudem der Fall $\tau = \frac{1}{2}$ keine höhere Approximationsgüte erwarten lässt als andere Werte für τ . In deterministischen Integralen entspricht $\tau = \frac{1}{2}$ der Mittelpunkregel, welche Ordnung 2 im Gegensatz zu Ordnung 1 bei Rechteckregeln besitzt. Dies gilt jedoch nur, wenn der Integrand zweimal stetig differenzierbar ist. Der Wiener Prozess und damit die Lösungen der stochastischen Differentialgleichungen haben dagegen im allgemeinen nirgends differenzierbare Pfade.

Desweiteren kann die Stratonovich-Differentialgleichung (3.33) transformiert werden in eine Itô-Differentialgleichung

$$X_t = X_0 + \int_0^t b(s, X_s) ds + \frac{1}{2} \int_0^t \frac{\partial \sigma}{\partial x}(s, X_s) \sigma(s, X_s) ds + \int_0^t \sigma(s, X_s) dW_s, \quad (3.35)$$

d.h. beide Integralgleichungen besitzen die gleichen Lösungen. Voraussetzung dafür ist b stetig und σ stetig differenzierbar. Die Modifikation findet dabei nur im deterministischen Teil statt. Für $\frac{\partial \sigma}{\partial x} \equiv 0$, d.h. σ ist unabhängig von X_t , fallen die beiden Integralbegriffe sogar zusammen. Wegen dieser Transformationseigenschaft brauchen wir im folgenden nur Itô-Integrale zu untersuchen.

Ist $h : \mathbb{R} \rightarrow \mathbb{R}$ eine glatte Funktion, dann hat man die Transformationsformel

$$\int_a^b h(W_t) \circ dW_t = \frac{1}{2} \int_a^b h'(W_t) dt + \int_a^b h(W_t) dW_t.$$

Für $h(X_t)$ mit allgemeinem Zufallsprozess $(X_t)_{t \in [a,b]}$ gilt dies jedoch nicht immer.

4 Itô-Prozesse

In diesem Abschnitt führen wir die Itô-Formel ein, welche das zentrale Hilfsmittel zur analytischen Lösung von Itô-Differentialgleichungen darstellt.

4.1 Eindimensionale Itô-Prozesse und Itô-Formel

Mit der Definition des Itô-Integrals haben wir beispielsweise die Aussage

$$\int_0^t W_s \, dW_s = \frac{1}{2}W_t^2 - \frac{1}{2}t \quad \Leftrightarrow \quad \frac{1}{2}W_t^2 = \int_0^t \frac{1}{2} \, ds + \int_0^t W_s \, dW_s. \quad (4.1)$$

Das Bild des Itô-Integrals $W_t = \int_0^t dW_s$ unter der Abbildung $g(x) = \frac{1}{2}x^2$ ist somit kein Itô-Integral allein, sondern die Summe aus einem deterministischen Integral und einem Itô-Integral. Dies motiviert die folgende Definition.

Definition 4.1 Sei $(W_t)_{t \geq 0}$ ein eindimensionaler Wiener Prozess auf dem Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) . Ein eindimensionaler Itô-Prozess bezeichnet einen Zufallsprozess $(X_t)_{t \geq 0}$ auf (Ω, \mathcal{A}, P) , welcher eine Integralgleichung

$$X_t = X_0 + \int_0^t u(s, \omega) \, ds + \int_0^t v(s, \omega) \, dW_s \quad (4.2)$$

erfüllt, wobei $v \in \mathcal{W}(0, t)$ für alle $t \geq 0$ gemäß Definition 3.7 gilt mit

$$P \left[\int_0^t v(s, \omega)^2 \, ds < \infty \text{ für alle } t \geq 0 \right] = 1$$

und die messbare Funktion u auch Bedingung (iii)' aus Definition 3.7 leistet mit

$$P \left[\int_0^t |u(s, \omega)| \, ds < \infty \text{ für alle } t \geq 0 \right] = 1.$$

Mit dieser Definition können wir die eindimensionale *Itô-Formel* angeben.

Satz 4.1 Sei $(X_t)_{t \geq 0}$ ein Itô-Prozess, der die Integralgleichung (4.2) erfüllt. Sei $g : [0, \infty) \times \mathbb{R} \rightarrow \mathbb{R}$, $(t, x) \mapsto g(t, x)$ eine Funktion, die C^1 bzgl. t und C^2 bzgl. x ist, dann ist $Y_t := g(t, X_t)$ wieder ein Itô-Prozess, der die Integralgleichung

$$\begin{aligned} g(t, X_t) &= g(0, X_0) \\ &+ \int_0^t \left(\frac{\partial g}{\partial t}(s, X_s) + u_s \frac{\partial g}{\partial x}(s, X_s) + \frac{1}{2}v_s^2 \frac{\partial^2 g}{\partial x^2}(s, X_s) \right) \, ds \\ &+ \int_0^t v_s \frac{\partial g}{\partial x}(s, X_s) \, dW_s \end{aligned} \quad (4.3)$$

mit $u_s := u(s, \omega)$, $v_s := v(s, \omega)$ löst.

Schreiben wir die Integralgleichung (4.2) in der Kurzform

$$dX_t = u dt + v dW_t,$$

dann kann man die Itô-Formel (4.3) abkürzen durch

$$dY_t = \left(\frac{\partial g}{\partial t} + u \frac{\partial g}{\partial x} + \frac{1}{2} v^2 \frac{\partial^2 g}{\partial x^2} \right) dt + v \frac{\partial g}{\partial x} dW_t. \quad (4.4)$$

Als Vorstufe hierzu haben wir auch die symbolische Form

$$dY_t = \frac{\partial g}{\partial t}(t, X_t) dt + \frac{\partial g}{\partial x}(t, X_t) dX_t + \frac{1}{2} \frac{\partial^2 g}{\partial x^2}(t, X_t) (dX_t)^2 \quad (4.5)$$

darstellen, wobei für die Differentiale die Rechenregeln

$$dt \cdot dt = dt \cdot dW_t = dW_t \cdot dt = 0, \quad dW_t \cdot dW_t = dt$$

gelten.

Wir vergleichen die Itô-Formel mit dem deterministischen Fall. Ist eine Funktion $x : [0, \infty) \rightarrow \mathbb{R}$ Lösung der gewöhnlichen Differentialgleichung

$$x'(t) = u(t, x(t)) + v(t, x(t))R(t) \quad (4.6)$$

mit stetigen Funktionen u, v und stetiger Funktion R , dann kann man für die neue Funktion $y(t) := g(t, x(t))$ mit einer stetig differenzierbaren Funktion g herleiten

$$\begin{aligned} y'(t) &= \frac{d}{dt} g(t, x(t)) \\ &= \frac{\partial g}{\partial t}(t, x(t)) + \frac{\partial g}{\partial x}(t, x(t)) x'(t) \\ &= \frac{\partial g}{\partial t}(t, x(t)) + \frac{\partial g}{\partial x}(t, x(t)) u(t, x(t)) + \frac{\partial g}{\partial x}(t, x(t)) v(t, x(t)) R(t). \end{aligned} \quad (4.7)$$

Führt $y(t) = g(t, x(t))$ auf eine Formel $x(t) = h(t, y(t))$ mit stetig differenzierbarer Funktion h , dann resultiert eine gewöhnliche Differentialgleichung für y .

Sei nun X_t ein Itô-Prozess, der die stochastische Differentialgleichung

$$dX_t = u(t, X_t) dt + v(t, X_t) dW_t$$

erfüllt. Definieren wir den Zufallsprozess $Y_t := g(t, X_t)$ mit $g \in C^2$, dann ist auch Y_t ein Itô-Prozess und nach der Itô-Formel (4.3) gilt (4.4). Für $v \equiv 0$ entsteht eine deterministische Differentialgleichung und die Transformation entspricht (4.7). Bemerkenswert ist, dass für $v \neq 0$ im Vergleich zu (4.7) ein Korrekturterm im deterministischen Teil der stochastischen Differentialgleichung auftritt.

Beweisskizze zu Satz 4.1:

Zur Vereinfachung sei $g \in C^3([0, \infty) \times \mathbb{R})$. Zudem nehmen wir an, dass die Funktion g und ihre betrachteten Ableitungen beschränkt sind, da man anderenfalls

mit Folgen aus beschränkten Funktionen approximieren kann, die auf kompakten Teilmengen von $[0, \infty) \times \mathbb{R}$ gleichmäßig konvergieren.

Die Voraussetzungen an die Funktionen u und v sind derart, dass man sie als Grenzwerte von Elementarfunktionen (3.13) darstellen kann. Im folgenden seien zu einer Zerlegung des Zeitintervalls $[0, t]$ mit u und v die korrespondierenden Elementarfunktionen bezeichnet.

Taylor-Entwicklung der Funktion $g \in C^3$ liefert uns zu den Zeitpunkten (3.16) mit den Abkürzungen $\Delta t_j = t_{j+1} - t_j$, $\Delta X_j = X_{t_{j+1}} - X_{t_j}$, $\Delta W_j = W_{t_{j+1}} - W_{t_j}$

$$\begin{aligned} g(t, X_t) &= g(0, X_0) + \sum_j g(t_{j+1}, X_{t_{j+1}}) - g(t_j, X_{t_j}) \\ &= g(0, X_0) + \sum_j \frac{\partial g}{\partial t} \Delta t_j + \sum_j \frac{\partial g}{\partial x} \Delta X_j + \frac{1}{2} \sum_j \frac{\partial^2 g}{\partial t^2} (\Delta t_j)^2 \\ &\quad + \sum_j \frac{\partial^2 g}{\partial t \partial x} (\Delta t_j) (\Delta X_j) + \frac{1}{2} \sum_j \frac{\partial^2 g}{\partial x^2} (\Delta X_j)^2 + \sum_j R_j, \end{aligned}$$

wobei die partiellen Ableitungen jeweils in (t_j, X_{t_j}) ausgewertet werden. Man beachte, dass die Anzahl der Summanden ungleich null wieder endlich ist. Für die Restglieder gilt

$$R_j = o(|\Delta t_j|^2 + |\Delta X_j|^2) \quad \text{für alle } j.$$

Da X_t ein Itô-Prozess ist, können fast alle Pfade als stetig angenommen werden. Es folgt damit punktweise

$$\lim_{\Delta t_j \rightarrow 0} \left| \sum_j \frac{\partial g}{\partial t}(t_j, X_{t_j}(\omega)) \Delta t_j - \int_0^t \frac{\partial g}{\partial t}(s, X_s(\omega)) \, ds \right| = 0.$$

Da die Funktionen als beschränkt vorausgesetzt sind, folgt nach dem Satz über dominierte Konvergenz

$$\lim_{\Delta t_j \rightarrow 0} \sum_j \frac{\partial g}{\partial t}(t_j, X_{t_j}(\omega)) \Delta t_j = \int_0^t \frac{\partial g}{\partial t}(s, X_s(\omega)) \, ds$$

in jeder Norm $L^p(\Omega)$ mit $p \geq 1$. Da u, v Elementarfunktionen sind, hat man mit einer geeigneten Zerlegung des Zeitintervalls

$$\Delta X_j(\omega) = u(t_j, \omega) \Delta t_j + v(t_j, \omega) \Delta W_j(\omega).$$

Es sei im folgenden $u_j := u(t_j, \omega)$ und $v_j := v(t_j, \omega)$. Mit der Definition des Itô-Integrals folgt desweiteren in $L^2(\Omega)$

$$\lim_{\Delta t_j \rightarrow 0} \sum_j \frac{\partial g}{\partial x}(t_j, X_{t_j}) \Delta X_j = \int_0^t \frac{\partial g}{\partial x}(s, X_s) u_s \, ds + \int_0^t \frac{\partial g}{\partial x}(s, X_s) v_s \, dW_s.$$

Mit der Beschränktheit der Funktionen gilt punktweise

$$\left| \sum_j \frac{\partial^2 g}{\partial t^2} (\Delta t_j)^2 \right| \leq K \sum_j (\Delta t_j)^2 = Kt \left(\max_i \Delta t_i \right), \quad (4.8)$$

wodurch dieser Term in $L^p(\Omega)$ für $p \geq 1$ gegen null konvergiert. Der Ausdruck

$$\sum_j \frac{\partial^2 g}{\partial t \partial x} (\Delta t_j) (\Delta X_j) = \sum_j \frac{\partial^2 g}{\partial t \partial x} u_j (\Delta t_j)^2 + \sum_j \frac{\partial^2 g}{\partial t \partial x} v_j (\Delta t_j) (\Delta W_j)$$

konvergiert ebenfalls gegen null in $L^2(\Omega)$, was für den ersten Term analog zu (4.8) folgt und beim zweiten Term mit der Abkürzung $a_j := \frac{\partial^2 g}{\partial t \partial x} v_j$ sowie der Eigenschaft $\mathbb{E}((\Delta W_j)^2) = \Delta t_j$ resultiert aus

$$\begin{aligned} \mathbb{E} \left(\left(\sum_j a_j (\Delta t_j) (\Delta W_j) \right)^2 \right) &= \sum_{i,j} \mathbb{E} (a_i a_j (\Delta t_i) (\Delta t_j) (\Delta W_i) (\Delta W_j)) \\ &= \sum_j \mathbb{E} (a_j^2) (\Delta t_j)^3. \end{aligned}$$

Dabei wurde eine gewisse Unabhängigkeit der beteiligten Funktionen bzw. Inkremente verwendet. Schließlich betrachten wir noch

$$\sum_j \frac{\partial^2 g}{\partial x^2} (\Delta X_j)^2 = \sum_j \frac{\partial^2 g}{\partial x^2} (u_j^2 (\Delta t_j)^2 + 2u_j v_j (\Delta t_j) (\Delta W_j) + v_j^2 (\Delta W_j)^2).$$

Die ersten beiden Terme konvergieren analog zu oben gegen null. Für den dritten Term behaupten wir

$$\lim_{\Delta t_j \rightarrow 0} \sum_j \frac{\partial^2 g}{\partial x^2} v_j^2 (\Delta W_j)^2 = \int_0^t \frac{\partial^2 g}{\partial x^2} v_s^2 ds \quad (4.9)$$

in $L^2(\Omega)$. Es sei $b_j := \frac{\partial^2 g}{\partial x^2} v_j^2$. Dann haben wir wieder mit gewissen Unabhängigkeitseigenschaften

$$\begin{aligned} &\mathbb{E} \left(\left(\sum_j b_j (\Delta W_j)^2 - \sum_j b_j \Delta t_j \right)^2 \right) \\ &= \sum_{i,j} \mathbb{E} (b_i b_j ((\Delta W_i)^2 - \Delta t_i) ((\Delta W_j)^2 - \Delta t_j)) \\ &= \sum_j \mathbb{E} (b_j^2) \mathbb{E} (((\Delta W_j)^2 - \Delta t_j)^2) \\ &= \sum_j \mathbb{E} (b_j^2) \mathbb{E} ((\Delta W_j)^4 - 2(\Delta W_j)^2 (\Delta t_j) + (\Delta t_j)^2) \\ &= \sum_j \mathbb{E} (b_j^2) (3(\Delta t_j)^2 - 2(\Delta t_j)^2 + (\Delta t_j)^2) = 2 \sum_j \mathbb{E} (b_j^2) (\Delta t_j)^2. \end{aligned}$$

Damit folgt die Konvergenz (4.9). Analog begründet man, dass die Summe der Restglieder gegen null in $L^2(\Omega)$ konvergiert. Somit ist die Itô-Formel gezeigt. \square

Bemerkung: Für die stochastische Differentialgleichung nach Stratonovich

$$dX_t = u(t, X_t) dt + v(t, X_t) \circ dW_t$$

folgt mit $Y_t := g(t, X_t)$ die Transformationsformel

$$dY_t = \left(\frac{\partial g}{\partial t} + u \frac{\partial g}{\partial x} \right) dt + v \frac{\partial g}{\partial x} \circ dW_t,$$

d.h. das Ergebnis ist analog zu (4.7). Es gilt die Kettenregel im klassischen Sinn.

Beispiel: Als einen Anwendungsfall greifen wir das motivierende Beispiel (4.1) auf. Mit $X_t := W_t$ gilt in (4.2) dann $u \equiv 0$ und $v \equiv 1$. Mit $g(t, x) := \frac{1}{2}x^2$ hat man $\frac{\partial g}{\partial t} = 0$, $\frac{\partial g}{\partial x} = x$, $\frac{\partial^2 g}{\partial x^2} = 1$. Mit der Itô-Formel (4.3) gilt dann

$$\frac{1}{2}W_t^2 = \frac{1}{2}W_0^2 + \int_0^t \frac{1}{2} ds + \int_0^t W_s dW_s = \frac{1}{2}t + \int_0^t W_s dW_s.$$

Analog verwenden wir $g(t, x) := \frac{1}{3}x^3$ und haben $\frac{\partial g}{\partial t} = 0$, $\frac{\partial g}{\partial x} = x^2$, $\frac{\partial^2 g}{\partial x^2} = 2x$. Es folgt

$$\frac{1}{3}W_t^3 = \frac{1}{3}W_0^3 + \int_0^t \frac{1}{2} \cdot 2W_s ds + \int_0^t W_s^2 dW_s = \int_0^t W_s ds + \int_0^t W_s^2 dW_s.$$

Auf diese Weise können wir also die Itô-Integrale mit Potenzen des Wiener Prozesses als Integranden bestimmen. Mit $g(t, x) = \frac{1}{n}x^n$ folgt für $n \geq 2$

$$\frac{1}{n}W_t^n = \int_0^t \frac{n-1}{2}W_s^{n-2} ds + \int_0^t W_s^{n-1} dW_s.$$

Die Itô-Formel kann angewendet werden, um eine nützliche Formel zur partiellen Integration bei stochastischen Integralen herzuleiten. Dabei ist wesentlich, dass der Integrand im Itô-Integral nicht zufallsabhängig ist.

Satz 4.2 *Ist $f : [0, t] \rightarrow \mathbb{R}$ eine stetig differenzierbare (deterministische) Funktion, dann gilt*

$$\int_0^t f(s) dW_s = f(t)W_t - \int_0^t W_s df_s = f(t)W_t - \int_0^t W_s f'(s) ds.$$

Beweis:

Wir wenden die Itô-Formel mit $X_t = W_t$ und der Funktion $g(t, x) = f(t)x$ an. Dann gilt $u \equiv 0$, $v \equiv 1$ und $\frac{\partial g}{\partial t} = f'(t)x$, $\frac{\partial g}{\partial x} = f(t)$, $\frac{\partial^2 g}{\partial x^2} = 0$. Somit erfüllt $Y_t = g(t, W_t) = f(t)W_t$ die Integralgleichung

$$f(t)W_t = f(0)W_0 + \int_0^t f'(s)W_s \, ds + \int_0^t f(s) \, dW_s.$$

Mit $W_0 = 0$ ist die Behauptung gezeigt. \square

Der Satz gilt analog, wenn f keine stetig differenzierbare Funktion ist, aber eine endliche Variation besitzt, so dass das korrespondierende Riemann-Stieltjes Integral existiert (siehe Abschnitt 3.1).

Als Anwendungsbeispiel bestimmen wir mit $f(t) \equiv t$

$$\int_0^t s \, dW_s = tW_t - \int_0^t W_s \, ds.$$

4.2 Mehrdimensionale Itô-Prozesse und Itô-Formel

Analog zum eindimensionalen Itô-Prozess definieren wir den *mehrdimensionalen Itô-Prozess*.

Definition 4.2 Sei $(W_1(t), \dots, W_n(t))_{t \geq 0}$ ein n -dimensionaler Wiener Prozess, auf dem Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) . Ein m -dimensionaler Itô-Prozess ist ein Zufallsprozess $(X_1(t), \dots, X_m(t))_{t \geq 0}$ auf (Ω, \mathcal{A}, P) , welcher ein Integralgleichungssystem (in Kurzform geschrieben)

$$\begin{aligned} dX_1 &= u_1 \, dt + v_{11} \, dW_1 + \dots + v_{1n} \, dW_n \\ &\vdots \\ dX_m &= u_m \, dt + v_{m1} \, dW_1 + \dots + v_{mn} \, dW_n \end{aligned} \tag{4.10}$$

erfüllt, wobei die enthaltenen Funktionen $u_i(t, \omega)$ und $v_{ij}(t, \omega)$ für $i = 1, \dots, m$ und $j = 1, \dots, n$ die Eigenschaften wie in Definition 4.1 besitzen.

Mit $X(t) = (X_i(t))$, $W(t) = (W_j(t))$, $u = (u_i)$ und $v = (v_{ij})$ lautet das System (4.10) als Kurzform in Matrix-Vektor-Notation

$$dX(t) = u \, dt + v \, dW(t).$$

Nun können wir die *mehrdimensionale Itô-Formel* aufstellen.

Satz 4.3 *Es sei $(X(t))_{t \geq 0}$ ein m -dimensionaler Itô-Prozess, der ein Integralgleichungssystem (4.10) erfüllt. Ist $g : [0, \infty) \times \mathbb{R}^m \rightarrow \mathbb{R}^p$, $(t, x) \mapsto g(t, x)$ eine Funktion, die C^1 bzgl. t und C^2 bzgl. $x \in \mathbb{R}^m$ ist, dann ist $Y(t) := g(t, X(t))$ ein p -dimensionaler Itô-Prozess, dessen Komponenten Lösungen des Integralgleichungssystems (in Kurzform dargestellt)*

$$\begin{aligned} dY_k(t) &= \left(\frac{\partial g_k}{\partial t} + \sum_{i=1}^m u_i \frac{\partial g_k}{\partial x_i} + \frac{1}{2} \sum_{i, \ell=1}^m \frac{\partial^2 g_k}{\partial x_i \partial x_\ell} \sum_{j=1}^n v_{ij} v_{\ell j} \right) dt \\ &\quad + \sum_{i=1}^m \sum_{j=1}^n v_{ij} \frac{\partial g_k}{\partial x_i} dW_j(t) \end{aligned} \tag{4.11}$$

für $k = 1, \dots, p$ repräsentieren, wobei die partiellen Ableitungen in (s, X_s) ausgewertet werden.

Der Beweis dieses Satzes verläuft analog zur Aussage im eindimensionalen Fall aus Satz 4.1. Entsprechend zu Formel (4.5) haben wir bei (4.11) die Darstellung

$$dY_k(t) = \frac{\partial g_k}{\partial t} dt + \sum_{i=1}^m \frac{\partial g_k}{\partial x_i} (dX_i(t)) + \frac{1}{2} \sum_{i, \ell=1}^m \frac{\partial^2 g_k}{\partial x_i \partial x_\ell} (dX_i(t))(dX_\ell(t))$$

mit den Rechenregeln für Differentiale

$$dt \cdot dt = dt \cdot dW_j = dW_j \cdot dt = 0, \quad dW_i \cdot dW_j = \delta_{ij} dt, \quad i, j = 1, \dots, n.$$

Beispiel: Für eine Anwendung der mehrdimensionalen Itô-Formel betrachten wir $X(t) = W(t)$ mit einem n -dimensionalen Wiener Prozess, d.h. es ist u der Nullvektor und v die Einheitsmatrix. Wir verwenden die Betragsfunktion

$$g : [0, \infty) \times \mathbb{R}^n \rightarrow \mathbb{R}, \quad g(t, x) = |x| = \sqrt{x_1^2 + \dots + x_n^2}.$$

Die Betragsfunktion ist stetig differenzierbar in allen Punkten $x \neq 0$. Damit gilt

$$\frac{\partial g}{\partial t} = 0, \quad \frac{\partial g}{\partial x_i} = \frac{x_i}{|x|}, \quad \frac{\partial^2 g}{\partial x_i \partial x_\ell} = \frac{\delta_{i\ell} |x|^2 - x_i x_\ell}{|x|^3}.$$

Es folgt, dass der Prozess

$$|W(t)| = \sqrt{W_1(t)^2 + \dots + W_n(t)^2}$$

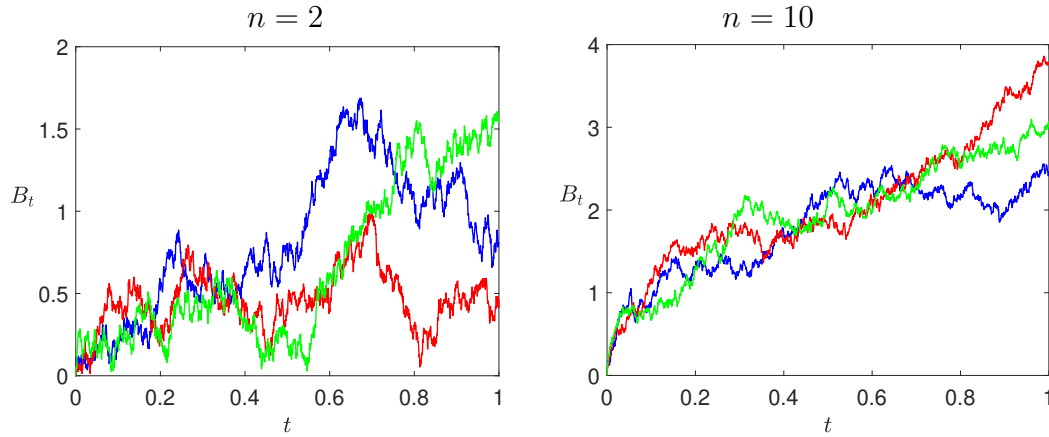


Abbildung 6: Pfade zu einem n -dimensionalen Bessel Prozess.

die Differentialgleichung

$$\begin{aligned}
 d|W| &= \left(\frac{1}{2} \sum_{i,\ell=1}^n \frac{\delta_{i\ell} |W|^2 - W_i W_\ell}{|W|^3} \sum_{j=1}^n \delta_{ij} \delta_{\ell j} \right) dt + \sum_{i=1}^n \sum_{j=1}^n \delta_{ij} \frac{W_i}{|W|} dW_j \\
 &= \left(\frac{1}{2} \sum_{i=1}^n \frac{|W|^2 - W_i^2}{|W|^3} \right) dt + \sum_{i=1}^n \frac{W_i}{|W|} dW_i \\
 &= \frac{n-1}{2|W|} dt + \sum_{i=1}^n \frac{W_i}{|W|} dW_i
 \end{aligned}$$

erfüllt. Die korrespondierende Integralgleichung ist jedoch nur in Zeitintervallen (t_0, ∞) mit $t_0 > 0$ sinnvoll, da $|W(0)| = 0$ gilt. Es ist fast sicher $|W(t)| \neq 0$ für $t > 0$ bei $n > 1$. Der Zufallsprozess $|W(t)|$ wird n -dimensionaler Bessel Prozess genannt. Abbildung 6 zeigt Pfade zu Bessel Prozessen.

4.3 Lineare Differentialgleichungen

In diesem Abschnitt diskutieren wir skalare lineare stochastische Differentialgleichungen.

Arithmetische Brownsche Bewegung

Eine *arithmetische Brownsche Bewegung* ist eine Lösung der Itô-Integralgleichung

$$M_t = M_0 + \int_0^t \alpha ds + \int_0^t \beta dW_s. \quad (4.12)$$

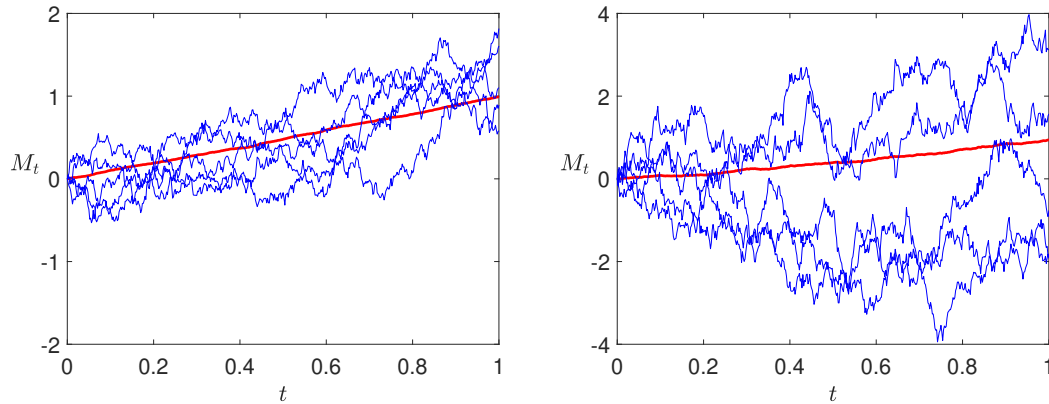


Abbildung 7: Pfade einer arithmetischen Brownschen Bewegung (blau) sowie Approximation des Erwartungswerts (rot) zu Parametern $\alpha = 1, \beta = 0.1$ (links) und $\alpha = 1, \beta = 0.5$ (rechts).

mit Konstanten $\alpha, \beta \in \mathbb{R}$. Als Lösung von (4.12) folgt sofort

$$M_t = M_0 + \alpha t + \beta W_t. \quad (4.13)$$

Der Spezialfall für $\alpha = 0$ und $\beta = 1$ liefert den Wiener Prozess bzw. die Brownsche Bewegung selbst. Als Erwartungswert ergibt sich

$$\mathbb{E}(M_t) = \mathbb{E}(M_0) + \alpha t, \quad (4.14)$$

da $\mathbb{E}(W_t) = 0$ für alle $t \geq 0$ gilt. Als Varianz erhalten wir mit der Annahme, dass M_0 und $(W_t)_{t \geq 0}$ unabhängig sind,

$$\text{Var}(M_t) = \text{Var}(M_0) + \beta^2 t,$$

weil $\text{Var}(W_t) = t$ vorliegt.

Die zur Itô-Integralgleichung (4.12) analoge Stratonovich-Integralgleichung besitzt ebenfalls die Lösung (4.13), da der Integrand im stochastischen Integral konstant ist.

Zur Interpretation setzen wir die Zeitpunkte $t_j = j\Delta t$ für $j = 0, 1, 2, \dots$ mit einer Schrittweite $\Delta t > 0$ fest. Es folgt

$$M_{t_{j+1}} - M_{t_j} = \alpha \Delta t + \beta (W_{t_{j+1}} - W_{t_j}).$$

Darin ist das Inkrement $W_{t_{j+1}} - W_{t_j}$ normalverteilt mit Erwartungswert null und Varianz Δt . Die Differenz $M_{t_{j+1}} - M_{t_j}$ besitzt daher stets die gleiche Verteilung.

Abbildung 7 zeigt fünf Pfade einer arithmetischen Brownschen Bewegung zu verschiedenen Parametern und Anfangswert $M_0 = 0$. Zusätzlich ist der Mittelwert von 1000 Pfaden dargestellt, welcher den Erwartungswert (4.14) approximiert.

Geometrische Brownsche Bewegung

Eine *geometrische Brownsche Bewegung* ist eine Lösung der Itô-Integralgleichung

$$M_t = M_0 + \int_0^t \alpha M_s \, ds + \int_0^t \beta M_s \, dW_s. \quad (4.15)$$

Darin bezeichnet die Konstanten $\alpha \in \mathbb{R}$ die Driftrate und die Konstante $\beta \in \mathbb{R}$ die Diffusionsrate. Zur Lösung dieser Gleichung soll die Itô-Formel (4.3) eingesetzt werden. Wir prüfen, ob $Y_t := \ln M_t$ eine noch einfachere Integralgleichung löst. Mit $g(t, x) = \ln x$ gilt $\frac{\partial g}{\partial t} = 0$, $\frac{\partial g}{\partial x} = \frac{1}{x}$, $\frac{\partial^2 g}{\partial x^2} = -\frac{1}{x^2}$. Die Itô-Formel zeigt mit $u = \alpha M_t$ und $v = \beta M_t$

$$\begin{aligned} \ln M_t &= \ln M_0 + \int_0^t \alpha M_s \frac{1}{M_s} - \frac{1}{2} \beta^2 M_s^2 \frac{1}{M_s^2} \, ds + \int_0^t \beta M_s \frac{1}{M_s} \, dW_s \\ &= \ln M_0 + \int_0^t \alpha - \frac{1}{2} \beta^2 \, ds + \int_0^t \beta \, dW_s \\ &= \ln M_0 + (\alpha - \frac{1}{2} \beta^2) t + \beta W_t. \end{aligned}$$

Wir erhalten als Lösung der linearen Itô-Integralgleichung (4.15)

$$M_t = M_0 e^{(\alpha - \frac{1}{2} \beta^2) t + \beta W_t}. \quad (4.16)$$

Insbesondere ist die dieser Zufallsprozess positiv falls $M_0 > 0$ gilt.

Um die Lösung der Stratonovich-Integralgleichung

$$N_t = N_0 + \int_0^t \alpha N_t \, ds + \int_0^t \beta N_t \circ dW_s \quad (4.17)$$

zu bestimmen bedienen wir uns der Transformationsformel (3.35). Dabei gilt $\sigma(t, x) = \beta x$ und $\frac{\partial \sigma}{\partial x} = \beta$. Die Lösung von (4.17) erfüllt demnach die Itô-Integralgleichung

$$N_t = N_0 + \int_0^t (\alpha + \frac{1}{2} \beta^2) N_s \, ds + \int_0^t \beta N_s \, dW_s.$$

Aus der Formel (4.16) folgt daher

$$N_t = N_0 e^{\alpha t + \beta W_t}.$$

Zur Interpretation setzen wir wieder die Zeitpunkte $t_j = j \Delta t$ für $j = 0, 1, 2, \dots$ mit einer Schrittweite $\Delta t > 0$ fest. Für die Lösung der linearen Itô-Integralgleichung (4.15) gilt

$$\frac{M_{t_{j+1}}}{M_{t_j}} = \frac{M_0 e^{(\alpha - \frac{1}{2} \beta^2) t_{j+1} + \beta W_{t_{j+1}}}}{M_0 e^{(\alpha - \frac{1}{2} \beta^2) t_j + \beta W_{t_j}}} = e^{(\alpha - \frac{1}{2} \beta^2) \Delta t} e^{\beta (W_{t_{j+1}} - W_{t_j})}.$$

Das Inkrement $W_{t_{j+1}} - W_{t_j}$ ist normalverteilt mit Erwartungswert null und Varianz Δt . Der Quotient $M_{t_{j+1}}/M_{t_j}$ besitzt daher stets die gleiche Verteilung.

Eigenschaften der geometrischen Brownschen Bewegung

Allgemeiner betrachten wir Zufallsprozesse der Form

$$X_t = X_0 e^{\mu t + \sigma W_t} \quad \text{mit } \mu, \sigma \in \mathbb{R}, \quad (4.18)$$

welche qualitativ einer geometrischen Brownschen Bewegung entsprechen. Das *Gesetz vom iterierten Logarithmus* besagt für einen Wiener Prozess

$$\limsup_{t \rightarrow \infty} \frac{W_t(\omega)}{\sqrt{2t \ln \ln t}} = 1, \quad \liminf_{t \rightarrow \infty} \frac{W_t(\omega)}{\sqrt{2t \ln \ln t}} = -1$$

für fast alle $\omega \in \Omega$. Eine Konsequenz hieraus ist, dass für jedes $\varepsilon > 0$ und fast jedes $\omega \in \Omega$ ein Zeitpunkt $t_0 = t_0(\varepsilon, \omega) > 0$ existiert mit

$$-(1 + \varepsilon)\sqrt{2t \ln \ln t} \leq W_t(\omega) \leq (1 + \varepsilon)\sqrt{2t \ln \ln t} \quad \text{für alle } t > t_0.$$

Daraus folgt, dass fast alle Pfade des Wiener Prozesses für $t \rightarrow \infty$ nur sublinear steigen bzw. fallen. Folglich ist für das Verhalten nach langer Zeit im Prozess (4.18) nur der Parameters μ bestimmend. Für gewisse Parameterwahlen α, β im Itô-Modell (4.15) bzw. Stratonovich-Modell (4.17) besitzen die jeweiligen Lösungen somit ein qualitativ unterschiedliches Verhalten bei $t \rightarrow \infty$.

Wir interessieren uns für den Erwartungswert einer geometrischen Brownschen Bewegung (4.18). Unter der Voraussetzung, dass X_0 und $(W_t)_{t \geq 0}$ unabhängig sind, folgt

$$\mathbb{E}(X_t) = \mathbb{E}(X_0) e^{\mu t} \mathbb{E}(e^{\sigma W_t}). \quad (4.19)$$

Es ist also der Erwartungswert von $Y_t = e^{\sigma W_t}$ zu bestimmen. Anwendung der Itô-Formel (4.3) liefert mit $g(t, x) = e^{\sigma x}$ sowie $\frac{\partial g}{\partial t} = 0$, $\frac{\partial g}{\partial x} = \sigma e^{\sigma x}$, $\frac{\partial^2 g}{\partial x^2} = \sigma^2 e^{\sigma x}$.

$$\begin{aligned} Y_t &= Y_0 + \int_0^t \frac{1}{2} \sigma^2 e^{\sigma W_s} ds + \int_0^t \sigma e^{\sigma W_s} dW_s \\ &= Y_0 + \int_0^t \frac{1}{2} \sigma^2 Y_s ds + \int_0^t \sigma Y_s dW_s. \end{aligned}$$

Mit den Eigenschaften des Itô-Integrals folgt

$$\mathbb{E}(Y_t) = \mathbb{E}(Y_0) + \frac{1}{2} \sigma^2 \int_0^t \mathbb{E}(Y_s) ds.$$

Wegen $Y_0 = e^{\sigma W_0} = 1$ ist die Lösung dieser deterministischen Integralgleichung

$$\mathbb{E}(e^{\sigma W_t}) = e^{\frac{1}{2} \sigma^2 t}. \quad (4.20)$$

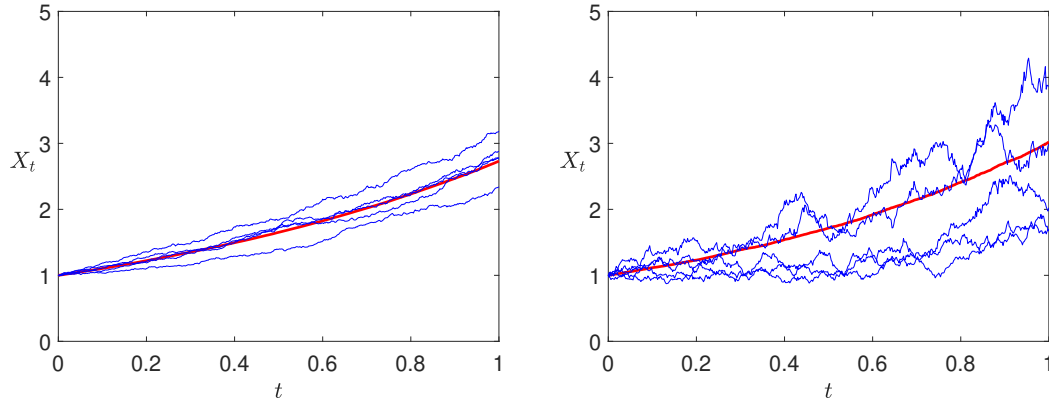


Abbildung 8: Pfade einer geometrischen Brownschen Bewegung (blau) sowie Approximation des Erwartungswerts (rot) zu Parametern $\mu = 1, \sigma = 0.1$ (links) und $\mu = 1, \sigma = 0.5$ (rechts).

Somit haben wir aus (4.19)

$$\mathbb{E}(X_t) = \mathbb{E}(X_0) e^{(\mu + \frac{1}{2}\sigma^2)t}. \quad (4.21)$$

Analog bestimmen wir

$$\mathbb{E}(X_t^2) = \mathbb{E}(X_0^2 e^{2\mu t + 2\sigma W_t}) = \mathbb{E}(X_0^2) \mathbb{E}(e^{2\mu t + 2\sigma W_t}) = \mathbb{E}(X_0^2) e^{2\mu t + 2\sigma^2 W_t}.$$

Als Varianz erhalten wir daraus im Fall von X_0 konstant

$$\text{Var}(X_t) = X_0^2 e^{(2\mu + \sigma^2)t} \left[e^{\sigma^2 t} - 1 \right].$$

Wir bemerken, dass der Parameter σ in diesen Formeln stets quadratisch auftritt, d.h. sein Vorzeichen spielt keine Rolle. Dies spiegelt wider, dass beim Wiener Prozess $\mathbb{E}(W_t) = 0$ gilt, also keine Tendenz bevorzugt wird.

Damit folgt für die Lösung der Itô-Integralgleichung (4.15)

$$\mathbb{E}(M_t) = \mathbb{E}(M_0) e^{\alpha t},$$

d.h. der Parameter β beeinflusst den Erwartungswert nicht. Für die Lösung der Stratonovich-Integralgleichung (4.17) haben wir dagegen

$$\mathbb{E}(N_t) = \mathbb{E}(N_0) e^{(\alpha + \frac{1}{2}\beta^2)t}.$$

In der Varianz der Lösungen macht sich in beiden Fällen der Parameter β bemerkbar.

Abbildung 8 zeigt fünf Pfade einer geometrischen Brownschen Bewegung (4.18) zu verschiedenen Parametern und Anfangswert $X_0 = 1$. Zusätzlich ist der Mittelwert von 1000 Pfaden dargestellt, welcher den Erwartungswert (4.21) approximiert.

Die Resultate in diesem Abschnitt zeigen insbesondere, dass zu linearen stochastischen Differentialgleichungen stets eine Lösung auf dem gesamten Zeitbereich $t \geq 0$ existiert.

4.4 Existenz und Eindeutigkeit

Wir betrachten ein allgemeines System aus Itô-Differentialgleichungen, d.h. m Integralgleichungen

$$X_i(t) = X_i(0) + \int_0^t u_i(s, X(s)) ds + \sum_{j=1}^n \int_0^t v_{ij}(s, X(s)) dW_j(s) \quad (4.22)$$

für $i = 1, \dots, m$ mit einem n -dimensionalen Wiener Prozess. Wir verwenden die Euklidische Vektornorm $\|\cdot\|$ und die Schur-(Matrix-)Norm $\|\cdot\|_S$. Es gilt das folgende Resultat über Existenz und Eindeutigkeit von Lösungen.

Satz 4.4 *Sei $T > 0$. Die folgenden Voraussetzungen seien erfüllt.*

- i) Die Koeffizienten $u : [0, T] \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ und $v : [0, T] \times \mathbb{R}^m \rightarrow \mathbb{R}^{m \times n}$ seien (komponentenweise) messbare Funktionen.*
- ii) Die Koeffizientenfunktionen besitzen die Schranken*

$$\begin{aligned} \|u(t, x)\| &\leq C_u(1 + \|x\|) \\ \|v(t, x)\|_S &\leq C_v(1 + \|x\|) \end{aligned}$$

für alle $t \in [0, T]$ und alle $x \in \mathbb{R}^m$ mit Konstanten $C_u, C_v > 0$.

- iii) Die Koeffizientenfunktionen sind global Lipschitz-stetig bzgl. x , d.h.*

$$\begin{aligned} \|u(t, x) - u(t, y)\| &\leq L_u \|x - y\| \\ \|v(t, x) - v(t, y)\|_S &\leq L_v \|x - y\| \end{aligned}$$

für alle $t \in [0, T]$ und alle $x, y \in \mathbb{R}^m$ mit Konstanten $L_u, L_v > 0$.

- iv) Der Anfangswert $X(0) = Z$ ist unabhängig bezüglich der vom Wiener Prozess erzeugten kanonischen Filtration $(\mathcal{F}_t^{(n)})_{t \in [0, T]}$. Zudem gilt $\mathbb{E}(\|Z\|^2) < \infty$.*

Dann besitzt das Itô-Differentialgleichungssystem (4.22) eine eindeutige Lösung $(X(t))_{t \in [0, T]}$ mit den Eigenschaften:

1. Die Pfade von $(X(t))_{t \in [0, T]}$ sind fast sicher stetig.
2. Der Zufallsprozess $(X(t))_{t \in [0, T]}$ ist adaptiert bezüglich der von Z und $(W(t))_{t \in [0, T]}$ erzeugten kanonischen Filtration.
3. Im Produktraum $[0, T] \times \Omega$ ist die Norm von $(X(t))_{t \in [0, T]}$ endlich, d.h.

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

Der Beweis kann in Abschnitt 5.2 aus [14] gefunden werden.

Die Voraussetzung (ii) ist insbesondere für global beschränkte Funktionen erfüllt. Die Voraussetzungen (iv) sind sofort erfüllt, falls die Anfangswerte $X(0)$ konstant (in ω) gewählt werden. Desweiteren gilt für $Z = (Z_1, \dots, Z_m)^\top$ die Bedingung $\mathbb{E}(\|Z\|^2) < \infty$ genau dann, wenn $\mathbb{E}(Z_i^2) < \infty$ für alle $i = 1, \dots, m$.

Im Fall $v \equiv 0$ reduziert sich das System (4.22) zu einer Schar aus deterministischen Integralgleichungen. Ist zudem u stetig, dann ist das System (4.22) äquivalent zu einer Schar aus gewöhnlichen Differentialgleichungssystemen (Scharparameter ist $\omega \in \Omega$). Daher müssen auch die Bedingungen für Existenz und Eindeutigkeit einer Lösung zu gewöhnlichen Differentialgleichungen für u erfüllt sein. Die Beschränkung (ii) entspricht einer Nicht-Explosions-Bedingung, welche die Existenz auf dem gesamten Intervall $[0, T]$ garantiert. Die Lipschitz-Stetigkeit (iii) impliziert die Eindeutigkeit der Lösung.

Wir betrachten ein allgemeines lineares Itô-Differentialgleichungssystem

$$X_i(t) = X_i(0) + \int_0^t \sum_{k=1}^m \alpha_{ik} X_k(s) ds + \sum_{j=1}^n \int_0^t \sum_{k=1}^m \beta_{ijk} X_k(s) dW_j(s) \quad (4.23)$$

mit Konstanten $\alpha_{ik}, \beta_{ijk} \in \mathbb{R}$ für $i, k = 1, \dots, m$ und $j = 1, \dots, n$. Die Koeffizientenfunktionen lauten daher komponentenweise

$$u_i(t, x) = \sum_{k=1}^m \alpha_{ik} x_k, \quad v_{ij}(t, x) = \sum_{k=1}^m \beta_{ijk} x_k.$$

In Satz 4.4 sind dann die Voraussetzungen (i)-(iii) erfüllt. Die Messbarkeit (i) ist offensichtlich, da die Funktionen stetig sind. Wir zeigen die Bedingung (ii) für u . Dazu sei

$$a = \max_{i, k=1, \dots, m} |\alpha_{ik}|.$$

Es gilt

$$|u_i| = \left| \sum_{k=1}^m \alpha_{ik} x_k \right| \leq \sum_{k=1}^m |\alpha_{ik}| \cdot |x_k| \leq a \sum_{k=1}^m |x_k| \leq am \|x\|.$$

Dadurch erhalten wir

$$\|u\|^2 = \sum_{i=1}^m u_i^2 = \sum_{i=1}^m |u_i|^2 \leq \sum_{i=1}^m a^2 m^2 \|x\|^2 = a^2 m^3 \|x\|^2.$$

Also ist

$$\|u\| \leq am^{\frac{3}{2}} \|x\| \leq am^{\frac{3}{2}} (1 + \|x\|),$$

d.h. Bedingung (ii) ist erfüllt mit $C_u = am^{\frac{3}{2}}$. Bezüglich Bedingung (iii) für u schätzen wir ab

$$\begin{aligned} |u_i(x) - u_i(y)| &= \left| \sum_{k=1}^m \alpha_{ik} x_k - \sum_{k=1}^m \alpha_{ik} y_k \right| = \left| \sum_{k=1}^m \alpha_{ik} (x_k - y_k) \right| \\ &\leq \sum_{k=1}^m |\alpha_{ik}| \cdot |x_k - y_k| \leq a \sum_{k=1}^m |x_k - y_k| \leq am \|x - y\|. \end{aligned}$$

Wie oben folgt daraus

$$\|u(x) - u(y)\| \leq am^{\frac{3}{2}} \|x - y\|,$$

d.h. die Lipschitz-Stetigkeit von u mit Lipschitz-Konstante $L_u = am^{\frac{3}{2}}$. Für v folgen die Bedingungen (ii) und (iii) analog, da die Schur-Norm (auch: Frobenius-Norm) einer Matrix $A \in \mathbb{R}^{m \times n}$ gegeben ist durch

$$\|A\|_S = \sqrt{\sum_{i=1}^m \sum_{j=1}^n a_{ij}^2}.$$

Die Bedingung (iv) hängt vom Anfangswert $X(0)$ ab.

5 Numerische Verfahren

Nichtlineare stochastische Differentialgleichungen lassen sich analog zu gewöhnlichen Differentialgleichungen häufig nicht analytisch lösen. Dadurch sind wir auf numerische Methoden angewiesen, die uns zu einer beliebigen stochastischen Differentialgleichung aus einer gewissen Klasse eine geeignete Näherungslösung liefern.

5.1 Euler-Maruyama-Verfahren

Wir betrachten ein (nichtlineares) System aus Itô-Integralgleichungen

$$X_t = X_0 + \int_0^t a(s, X_s) ds + \int_0^t b(s, X_s) dW_s \quad \text{für } t \in [0, T] \quad (5.1)$$

mit vorgegebenem Anfangswert X_0 und Funktionen $a, b : [0, T] \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ sowie einem eindimensionalen Wiener Prozess $(W_t)_{t \geq 0}$. Das Zeitintervall $[0, T]$ sei äquidistant zerlegt mit der Schrittweite $h = \Delta t = \frac{T}{N}$. Das explizite Euler-Verfahren angewandt auf das Anfangswertproblem eines Systems gewöhnlicher Differentialgleichungen

$$\frac{dx}{dt}(t) = a(t, x(t)), \quad x(0) = x_0$$

für $t \in [0, T]$ mit Lösung $x : [0, T] \rightarrow \mathbb{R}^k$ liefert für die Näherungen $\tilde{x}_i \approx x(t_i)$ mit $t_i = ih$ die Rekursion

$$\tilde{x}_{i+1} = \tilde{x}_i + a(t_i, \tilde{x}_i)h \quad \text{für } i = 0, 1, \dots, N-1. \quad (5.2)$$

Wir konstruieren ein entsprechendes Verfahren für die Integralgleichung (5.1). Es seien $\Delta W_i = W_{t_{i+1}} - W_{t_i}$ die Inkremente des Wiener Prozesses zur Zerlegung des Zeitintervalls. Der Anfangswert $X_0 : \Omega \rightarrow \mathbb{R}^m$ ist vorgegeben. Wir approximieren wir die Funktionen a, b im Teilintervall $[t_0, t_1]$ durch konstante Vektoren

$$\bar{a} = a(t_0, X_0) \in \mathbb{R}^m \quad \text{und} \quad \bar{b} = b(t_0, X_0) \in \mathbb{R}^m.$$

Es folgt aus (5.1)

$$\begin{aligned} X_{t_1} &= X_{t_0} + \int_{t_0}^{t_1} a(s, X_s) ds + \int_{t_0}^{t_1} b(s, X_s) dW_s \\ &\approx X_0 + \int_{t_0}^{t_1} \bar{a} ds + \int_{t_0}^{t_1} \bar{b} dW_s \\ &= X_0 + \bar{a}h + \bar{b}(W_1 - W_0). \end{aligned}$$

Wir erhalten also als Näherung für X_{t_1}

$$\tilde{X}_1 = X_0 + a(t_0, X_0)h + b(t_0, X_0)\Delta W_0.$$

Dieses Konstruktionsprinzip kann rekursiv fortgesetzt werden, um Näherungen $\tilde{X}_i \approx X_{t_i}$ zu erhalten. Das *Euler-Maruyama-Verfahren* für die stochastische Differentialgleichung (5.1) lautet

$$\tilde{X}_{i+1} = \tilde{X}_i + a(t_i, \tilde{X}_i)h + b(t_i, \tilde{X}_i)\Delta W_i \quad \text{für } i = 0, 1, \dots, N-1 \quad (5.3)$$

mit Anfangswert $\tilde{X}_0 = X_0$ und $W_0 = 0$. In dieser Form liegt eine Rekursion für eine Folge von Zufallsvariablen vor.

Algorithmus für einen Pfad

Um ein numerisches Verfahren zu erhalten, müssen wir pfadweise Näherungslösungen berechnen, d.h. punktweise für $\omega \in \Omega$. Für ein benötigtes Inkrement gilt $\Delta W_i \sim N(0, h)$. Daher bestimmt sich aus Zufallszahlen gemäß dieser Normalverteilung eine Realisierung der Näherungen \tilde{X}_i . Die Erzeugung der Zufallszahlen erfolgt auf dem Rechner näherungsweise über Pseudo-Zufallszahlen, siehe [6].

Eine Näherung für einen Pfad $X_t(\omega)$ der Lösung von (5.1) (d.h. ω ist fest) ergibt sich durch folgenden Algorithmus basierend auf dem Schema (5.3).

1. Wähle ein $N \in \mathbb{N}$ und setze $h = \frac{T}{N}$ sowie $\tilde{x}_0 = X_0(\omega)$.
2. Erzeuge Pseudo-Zufallszahlen z_0, z_1, \dots, z_{N-1} gemäß (unabhängiger) Standard-Normalverteilungen $N(0, 1)$.
3. Für $i = 0, 1, \dots, N-1$ berechne

$$\tilde{x}_{i+1} = \tilde{x}_i + a(t_i, \tilde{x}_i)h + b(t_i, \tilde{x}_i)(\sqrt{h}z_i).$$

4. Ausgabe ist $\tilde{x}_0, \tilde{x}_1, \dots, \tilde{x}_{N-1}, \tilde{x}_N$.

Für die Resultate gilt dann $\tilde{x}_i \approx X_{t_i}(\omega)$ falls h hinreichend klein. Eine Funktion $\tilde{x}(t) \approx X_t(\omega)$ für $t \in [0, T]$ kann wieder durch lineare Interpolation gebildet werden.

Der Algorithmus kann beliebig oft mit unterschiedlichen Pseudo-Zufallszahlen wiederholt werden, um Näherungen für eine Vielzahl $X_t(\omega_1), \dots, X_t(\omega_K)$ an Pfaden im Zeitintervall $[0, T]$ zu erzeugen.

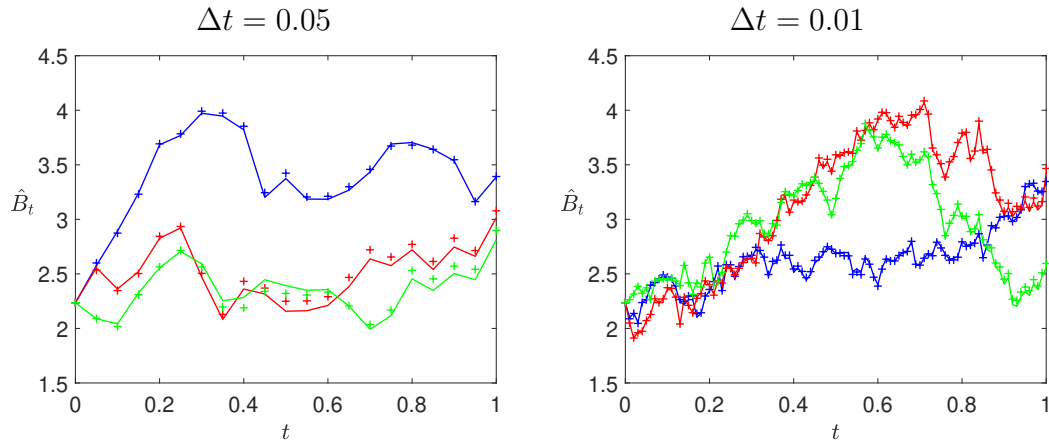


Abbildung 9: Pfade eines Bessel Prozesses zu einem 5-dimensionalen Wiener Prozess: direkte Simulation (Linien) und Näherungen aus dem Euler-Maruyama-Verfahren (Symbol +).

Verallgemeinerungen des Euler-Maruyama-Verfahrens

Das Euler-Maruyama-Verfahren ist wegen seiner Einfachheit die einzige Methode, die direkt auf Systeme von Differentialgleichungen mit mehrdimensionalem Wiener Prozess übertragen werden kann. Gilt in (5.1) nun $b : [0, T] \times \mathbb{R}^m \rightarrow \mathbb{R}^{m \times n}$ mit einem n -dimensionalen Wiener Prozess, dann brauchen wir ΔW_i in (5.3) nur durch einen Vektor mit Pseudo-Zufallszahlen z_1, \dots, z_n jeweils nach $N(0, h)$ verteilt zu ersetzen.

In manchen Problemen (5.1) tritt ein steifer deterministischer Anteil a auf. In diesem Fall können wir ein implizites Euler-Maruyama-Verfahren

$$\tilde{X}_{i+1} = \tilde{X}_i + a(t_{i+1}, \tilde{X}_{i+1})h + b(t_i, \tilde{X}_i)\Delta W_i \quad \text{für } i = 0, 1, \dots, N-1$$

verwenden. Hier ist ein nichtlineares Gleichungssystem für die jeweilige Approximation zu lösen. Der stochastische Anteil b kann jedoch nicht implizit behandelt werden. Die Definition des Itô-Integrals erzwingt die Auswertung der Funktion b am jeweils linken Intervallrand.

Beispiel: Bessel Prozess

Wir betrachten den Bessel Prozess aus Abschnitt 4.2. Sei dazu $(W(t))_{t \geq 0}$ ein n -dimensionaler Wiener Prozess. Um Probleme bei Division durch null zu vermeiden setzen wir den Anfangswert $W(0) = (1, \dots, 1)^\top$ fest. Der zugehörige Bessel Prozess

$$\hat{B}_t = \sqrt{W_1(t)^2 + \dots + W_n(t)^2} \quad (5.4)$$

erfüllt die nichtlineare Itô-Integralgleichung

$$\hat{B}_t = \hat{B}_0 + \int_0^t \frac{n-1}{2\hat{B}_s} ds + \sum_{i=1}^n \int_0^t \frac{W_i(s)}{\hat{B}_s} dW_i(s). \quad (5.5)$$

Der Anfangswert ergibt sich zu $\hat{B}_0 = \sqrt{n}$.

Wir setzen als Dimension $n = 5$ und betrachten das Zeitintervall $[0, 1]$. Zum einen werden Pfade des Bessel Prozesses berechnet über simulierte Werte eines 5-dimensionalen Wiener Prozesses gemäß Abschnitt 2.3 mit einer konstanten Schrittweite Δt und Auswertung von (5.4). Zum anderen wenden wir das Euler-Maruyama-Verfahren auf die Integralgleichung (5.5) an, wobei als Schrittweite wieder $h = \Delta t$ gesetzt wird. Es werden die gleichen Pseudo-Zufallszahlen für die Inkremente verwendet wie bei der Simulation des Wiener Prozesses selbst. Abbildung 9 zeigt die Näherungen für jeweils drei Pfade, die zu verschiedenen Schrittweiten $\Delta t = 0.05$ und $\Delta t = 0.01$ berechnet wurden. Wir erkennen, dass mit kleinerer Schrittweite die Differenz in den Ergebnissen abnimmt.

5.2 Starke und Schwache Konvergenz

Um die Güte eines numerischen Verfahrens für Differentialgleichungen zu beurteilen ist die Konvergenzordnung von Bedeutung. Im Gegensatz zu gewöhnlichen Differentialgleichungen ergeben sich bei stochastischen Differentialgleichungen zwei verschiedene Konzepte.

Die Lösung einer stochastischen Differentialgleichung ist eine Funktion in der Zeit $t \in [0, \infty)$ und in den Ergebnissen $\omega \in \Omega$. Die Differentialgleichung bzw. Integralgleichung beschreibt die zeitliche Änderung. An einem Zeitpunkt $T > 0$ können wir daher den gemittelten absoluten Fehler über alle Ergebnisse betrachten. Dies entspricht gerade der Konvergenz in der L^1 -Norm zum Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) .

Definition 5.1 Sei X_t die Lösung einer stochastischen Differentialgleichung und X_t^h die Näherung aus einem numerischen Verfahren mit der Schrittweite h . Die Approximation X_T^h konvergiert stark mit Ordnung (mindestens) $p > 0$ gegen X_T , falls für alle hinreichend kleinen Schrittweiten h gilt

$$\mathbb{E}(|X_T^h - X_T|) \leq Ch^p$$

mit einer von h unabhängigen Konstanten $C > 0$.

Ein weiteres Konzept ergibt sich daraus, dass wir häufig nicht an der Lösung einer stochastischen Differentialgleichung selbst interessiert sind, sondern nur an

Kennzahlen wie den Momenten der Zufallsvariable. Auf den Erwartungswert der Lösung bezogen resultiert das folgende Konzept.

Definition 5.2 Sei X_t die Lösung einer stochastischen Differentialgleichung und X_t^h die Näherung aus einem numerischen Verfahren mit der Schrittweite h . Sei eine Funktion $g : \mathbb{R} \rightarrow \mathbb{R}$ gegeben. Die Approximation $g(X_T^h)$ konvergiert schwach mit Ordnung (mindestens) $q > 0$ gegen $g(X_T)$, falls für alle hinreichend kleinen Schrittweiten h gilt

$$|\mathbb{E}(g(X_T^h)) - \mathbb{E}(g(X_T))| \leq Dh^q$$

mit einer von h unabhängigen Konstanten $D > 0$. Die Approximation X_T^h konvergiert schwach gegen X_T , wenn diese Eigenschaft für $g(x) \equiv x$ gilt.

Die Definitionen 5.1 und 5.2 gelten bei Differentialgleichungssystemen analog mit entsprechenden Vektornormen. Analog heißt ein numerisches Verfahren für eine bestimmte Klasse von stochastischen Differentialgleichungen stark bzw. schwach konvergent, wenn alle von ihm gelieferten Näherungen stark bzw. schwach konvergent sind.

Die schwache Konvergenz von X_T^h gegen X_T bedeutet auch

$$|\mathbb{E}(X_T^h - X_T)| = |\mathbb{E}(X_T^h) - \mathbb{E}(X_T)| \leq Dh^q.$$

Die schwache Konvergenz bezüglich $g(x) = x^k$ führt auf die Bedingung

$$|\mathbb{E}((X_T^h)^k) - \mathbb{E}((X_T)^k)| \leq Dh^q.$$

Dies ist gerade die Konvergenz des k -ten Moments.

Satz 5.1 Sei $g : \mathbb{R} \rightarrow \mathbb{R}$ global Lipschitz-stetig. Dann folgt aus der starken Konvergenz mit Ordnung p die schwache Konvergenz mit Ordnung mindestens p .

Beweis:

Die globale Lipschitz-Stetigkeit von g bedeutet

$$|g(x) - g(y)| \leq L|x - y|$$

für alle $x, y \in \mathbb{R}$ mit einer Konstante $L > 0$.

Unter der Annahme der starken Konvergenz schätzen wir ab

$$\begin{aligned}
|\mathbb{E}(g(X_T^h)) - \mathbb{E}(g(X_T))| &= |\mathbb{E}(g(X_T^h) - g(X_T))| \\
&= \left| \int_{\Omega} g(X_T^h) - g(X_T) \, dP \right| \\
&\leq \int_{\Omega} |g(X_T^h) - g(X_T)| \, dP \\
&\leq \int_{\Omega} L |X_T^h - X_T| \, dP \\
&= L \mathbb{E}(|X_T^h - X_T|) \leq LCh^p.
\end{aligned}$$

Diese Abschätzung zeigt die schwache Konvergenz mit Konstante $D = LC$. \square

Satz 5.1 gilt insbesondere für die Funktion $g(x) \equiv x$ mit Lipschitz-Konstante $L = 1$.

Das Euler-Maruyama-Verfahren aus dem vorhergehenden Abschnitt besitzt die schwache Konvergenzordnung 1 aber nur eine starke Konvergenzordnung $\frac{1}{2}$. Eine Erörterung hierzu erfolgt im nächsten Abschnitt. Die schwache Konvergenz des Euler-Maruyama-Verfahrens weisen wir im Fall der stochastischen Differentialgleichung zur geometrischen Brownschen Bewegung direkt nach. Die lineare Integralgleichung (4.15) besitzt die Lösung (4.16) mit dem Erwartungswert $\mathbb{E}(M_t) = \mathbb{E}(M_0)e^{\alpha t}$. Das Euler-Maruyama-Verfahren liefert für die Näherungen \tilde{M}_i von M_{ih} die rekursive Vorschrift

$$\tilde{M}_{i+1} = \tilde{M}_i (1 + \alpha h + \beta \Delta W_i) \quad \text{für } i = 0, \dots, N-1$$

mit $\Delta W_i \sim N(0, h)$. Anfangswert ist $\tilde{M}_0 = M_0$. Die Rekursion ineinander eingesetzt ergibt für den Endwert die Formel

$$\tilde{M}_N = M_0 \prod_{i=0}^{N-1} (1 + \alpha h + \beta \Delta W_i).$$

Nach Voraussetzung sind die Realisierungen der Inkremente ΔW_i alle unabhängig voneinander, wodurch wir die Bildung des Erwartungswerts und das Produkt vertauschen können.

Es folgt

$$\begin{aligned}
\mathbb{E}(\tilde{M}_N) &= \mathbb{E}(M_0) \mathbb{E}\left(\prod_{i=0}^{N-1} (1 + \alpha h + \beta \Delta W_i)\right) \\
&= \mathbb{E}(M_0) \left(\prod_{i=0}^{N-1} \mathbb{E}(1 + \alpha h + \beta \Delta W_i)\right) \\
&= \mathbb{E}(M_0) \left(\prod_{i=0}^{N-1} \left(1 + \alpha h + \beta \underbrace{\mathbb{E}(\Delta W_i)}_{=0}\right)\right) \\
&= \mathbb{E}(M_0) (1 + \alpha h)^N.
\end{aligned}$$

Mit $h = \frac{T}{N}$ erhalten wir

$$\lim_{N \rightarrow \infty} \mathbb{E}(\tilde{M}_N) = \mathbb{E}(M_0) \lim_{N \rightarrow \infty} \left(1 + \frac{\alpha T}{N}\right)^N = \mathbb{E}(M_0) e^{\alpha T} = \mathbb{E}(M_T).$$

Um die Konvergenzordnung einzusehen beachte man, dass die Werte $\mathbb{E}(M_t)$ Lösungen der gewöhnlichen Dgl. $y'(t) = \alpha y(t)$ mit $y(0) = \mathbb{E}(M_0)$ sind. Die Werte $\mathbb{E}(M_i)$ sind dann die Näherungen, die das klassische Euler-Verfahren (5.2) mit Schrittweite h für diese Dgl. liefert. Das klassische Euler-Schema besitzt die Konvergenzordnung 1. Damit folgt die schwache Konvergenzordnung $q = 1$ für das Euler-Maruyama-Verfahren im linearen Fall.

5.3 Stochastische Taylor-Entwicklungen

Um numerische Verfahren höherer Ordnung zu konstruieren, betrachten wir Methoden basierend auf Entwicklungen analog zu Taylor-Reihen. Verfahren von hoher Ordnung dieser Art erweisen sich bei gewöhnlichen Differentialgleichungen als ineffizient. Jedoch wollen wir später nur Verfahren bis zur starken Konvergenzordnung 1 konstruieren.

Eindimensionaler Fall

Wir betrachten zunächst autonome Gleichungen der Gestalt

$$X_t = X_0 + \int_0^t a(X_s) ds + \int_0^t b(X_s) dW_s, \quad t \in [0, T] \quad (5.6)$$

mit $a, b : \mathbb{R} \rightarrow \mathbb{R}$ hinreichend glatt. Für eine stochastische Taylor-Entwicklung von niedriger Ordnung liefert die Itô-Formel eine Darstellung. Mit zweimal stetig differenzierbarer Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ ist $f(X_t)$ wieder ein Itô-Prozess und erfüllt

die Integralgleichung

$$\begin{aligned}
f(X_t) &= f(X_{t_0}) + \int_{t_0}^t (f'(X_s)a(X_s) + \frac{1}{2}f''(X_s)b(X_s)^2) ds \\
&\quad + \int_{t_0}^t f'(X_s)b(X_s) dW_s.
\end{aligned} \tag{5.7}$$

Wir setzen nun (5.7) für $f = a$ und $f = b$ in (5.6) ein, wodurch folgt

$$\begin{aligned}
X_t &= X_{t_0} + \int_{t_0}^t \left(a(X_{t_0}) + \int_{t_0}^s (a'a + \frac{1}{2}a''b^2) du + \int_{t_0}^s a'b dW_u \right) ds \\
&\quad + \int_{t_0}^t \left(b(X_{t_0}) + \int_{t_0}^s (b'a + \frac{1}{2}b''b^2) du + \int_{t_0}^s b'b dW_u \right) dW_s.
\end{aligned}$$

Diese Formel können wir umschreiben in

$$X_t = X_{t_0} + a(X_{t_0})(t - t_0) + b(X_{t_0}) \int_{t_0}^t dW_s + R,$$

wobei der Restterm R alle Doppelintegrale enthält. Bei Entfallen des Restterms erhalten wir das Euler-Maruyama-Verfahren (5.3). Den Restterm spalten wir auf in

$$R = \int_{t_0}^t \int_{t_0}^s b'b dW_u dW_s + \tilde{R}. \tag{5.8}$$

Man kann zeigen, dass für den neuen Restterm $\tilde{R} = \mathcal{O}(h^{\frac{3}{2}})$ gilt. Das Doppelintegral in (5.8) ist dagegen $\mathcal{O}(h)$, denn man rechnet nach

$$\begin{aligned}
\int_{t_0}^t \int_{t_0}^s dW_u dW_s &= \int_{t_0}^t (W_s - W_{t_0}) dW_s \\
&= \int_{t_0}^t W_s dW_s - W_{t_0} \int_{t_0}^t dW_s \\
&= \frac{1}{2}(W_t^2 - W_{t_0}^2) - \frac{1}{2}(t - t_0) - W_{t_0}(W_t - W_{t_0}) \\
&= \frac{1}{2}((W_t - W_{t_0})^2 - (t - t_0))
\end{aligned} \tag{5.9}$$

und es gilt für $t \geq t_0$

$$\begin{aligned}
\mathbb{E}(|(W_t - W_{t_0})^2 - (t - t_0)|) &\leq \mathbb{E}(|(W_t - W_{t_0})^2| + |t - t_0|) \\
&= \mathbb{E}((W_t - W_{t_0})^2) + \mathbb{E}(t - t_0) \\
&= 2(t - t_0).
\end{aligned}$$

Somit repräsentiert im Restglied (5.8) dieses Doppelintegral den dominanten Term.

Mehrdimensionaler Fall

Sei $(W(t))_{t \geq 0}$ ein n -dimensionaler Wiener Prozess. Wir betrachten ein autonomes System aus Itô-Integralgleichungen

$$X_k(t) = X_k(t_0) + \int_{t_0}^t a_k(X(s)) \, ds + \sum_{j=1}^n \int_{t_0}^t b_{kj}(X(s)) \, dW_j(s) \quad (5.10)$$

mit Komponenten $k = 1, \dots, m$. Die (mehrdimensionale) Itô-Formel (4.11) in Bezug auf das System (5.10) liefert für $Y(t) = f(X(t))$ mit $f : \mathbb{R}^m \rightarrow \mathbb{R}$ die Aussage

$$\begin{aligned} f(X(t)) &= f(X(t_0)) + \int_{t_0}^t \sum_{i=1}^m a_i \frac{\partial f}{\partial x_i} + \frac{1}{2} \sum_{i,\ell=1}^m \frac{\partial^2 f}{\partial x_i \partial x_\ell} \sum_{j=1}^n b_{ij} b_{\ell j} \, ds \\ &\quad + \int_{t_0}^t \sum_{i=1}^m \sum_{j=1}^n b_{ij} \frac{\partial f}{\partial x_i} \, dW_j(s). \end{aligned} \quad (5.11)$$

Für diese Funktion können wir nun $f = a_k$ oder $f = b_{kp}$ anwenden. Einsetzen der jeweiligen Formeln (5.11) in (5.10) führt auf die Integralgleichungen

$$\begin{aligned} X_k(t) &= X_k(t_0) + \int_{t_0}^t a_k(X(s)) \, ds + \int_{t_0}^s \sum_{i=1}^m a_k \frac{\partial a_k}{\partial x_i} + \frac{1}{2} \sum_{i,\ell=1}^m \frac{\partial^2 a_k}{\partial x_i \partial x_\ell} \sum_{j=1}^n b_{ij} b_{\ell j} \, du \\ &\quad + \int_{t_0}^s \sum_{i=1}^m \sum_{j=1}^n b_{ij} \frac{\partial a_k}{\partial x_i} \, dW_j(u) \, ds \\ &\quad + \sum_{p=1}^n \int_{t_0}^t b_{kp}(X(s)) \, ds + \int_{t_0}^s \sum_{i=1}^m a_i \frac{\partial b_{kp}}{\partial x_i} + \frac{1}{2} \sum_{i,\ell=1}^m \frac{\partial^2 b_{kp}}{\partial x_i \partial x_\ell} \sum_{j=1}^n b_{ij} b_{\ell j} \, du \\ &\quad + \int_{t_0}^s \sum_{i=1}^m \sum_{j=1}^n b_{ij} \frac{\partial b_{kp}}{\partial x_i} \, dW_j(u) \, dW_p(s) \end{aligned} \quad (5.12)$$

für $k = 1, \dots, m$.

5.4 Milstein-Verfahren

Wir erhalten eine Methode höherer Ordnung als das Euler-Maruyama-Verfahren, wenn wir den dominanten Term von (5.8) in die Approximation einbeziehen. Ersetzen wir die Funktionen b' im Doppelintegral durch ihre Auswertungen am linken Intervallrand, dann entsteht nur ein Fehler höherer Ordnung. Es folgt das *Milstein-Verfahren* für die autonome Itô-Integralgleichung (5.6)

$$\tilde{X}_{i+1} = \tilde{X}_i + a(\tilde{X}_i)h + b(\tilde{X}_i)\Delta W_i + \frac{1}{2}b'(\tilde{X}_i)b(\tilde{X}_i)((\Delta W_i)^2 - h). \quad (5.13)$$

Das Milstein-Verfahren für das (nichtautonome) System aus Itô-Integralgleichungen (5.1) mit einem eindimensionalen Wiener Prozess lautet analog

$$\tilde{X}_{i+1} = \tilde{X}_i + a(t_i, \tilde{X}_i)h + b(t_i, \tilde{X}_i)\Delta W_i + \frac{1}{2}Db(t_i, \tilde{X}_i)b(t_i, \tilde{X}_i)((\Delta W_i)^2 - h)$$

für $i = 0, 1, \dots, N - 1$. Im Fall eines eindimensionalen Wiener Prozesses erhöht sich der Aufwand gegenüber dem Euler-Maruyama-Verfahren nur durch die Bildung der Funktionalmatrix $Db \in \mathbb{R}^{m \times m}$. Die Berechnung dieser Funktionalmatrix kann zudem durch geeignete Approximationen mit Differenzenquotienten vermieden werden, wobei dennoch die Konvergenzordnung 1 erhalten bleibt, siehe [6].

Wir leiten nun das Milstein-Verfahren für ein autonomes System aus Itô-Integralgleichungen (5.10) her. Die (mehrdimensionale) Itô-Formel führte auf die Gleichungen (5.12). Setzen wir dort $t_0 = t_q$ und $t = t_{q+1}$, dann erhalten wir mit der Schrittweite $h = t_{q+1} - t_q$ die Aussage

$$\begin{aligned} X_k(t_{q+1}) &= X_k(t_q) + \int_{t_q}^{t_{q+1}} a_k(X(t_q)) ds + \mathcal{O}(h^2) + \mathcal{O}(h^{\frac{3}{2}}) \\ &\quad + \sum_{p=1}^n \int_{t_q}^{t_{q+1}} b_{kp}(X(t_q)) dW_p(s) + \mathcal{O}(h^{\frac{3}{2}}) \\ &\quad + \sum_{p=1}^n \sum_{j=1}^n \int_{t_q}^{t_{q+1}} \int_{t_q}^s \left(\sum_{i=1}^m b_{ij} \frac{\partial b_{kp}}{\partial x_i} \right) dW_j(u) dW_p(s). \end{aligned}$$

Wenn die Restterme höherer Ordnung entfallen und die Funktionen im letzten Doppelintegral durch die Auswertungen am linken Intervallrand ersetzt werden, dann folgt das Milstein-Verfahren

$$\begin{aligned} \tilde{X}_k(t_{q+1}) &= \tilde{X}_k(t_q) + a_k(\tilde{X}(t_q))h + \sum_{p=1}^n b_{kp}(\tilde{X}(t_q))\Delta W_p \\ &\quad + \sum_{p,j=1}^n \left(\sum_{i=1}^m b_{ij}(\tilde{X}(t_q)) \frac{\partial b_{kp}}{\partial x_i}(\tilde{X}(t_q)) \right) \int_{t_q}^{t_{q+1}} \int_{t_q}^s dW_j(u) dW_p(s) \end{aligned} \tag{5.14}$$

für $k = 1, \dots, m$, wobei \tilde{X}_k die numerische Approximation bezeichnet. Die Formel gilt analog im nichtautonomen Fall. Um das Verfahren durchzuführen, sind noch die Doppelintegrale bezüglich der Wiener Prozesse auszuwerten. Dies kann leider im Gegensatz zum eindimensionalen Fall nicht mehr explizit durchgeführt werden. Stattdessen müssen wir eine numerische Approximation dieser Doppelintegrale herleiten, deren Fehlerordnung mindestens $\mathcal{O}(h)$ beträgt, damit wir die starke Konvergenzordnung $p = 1$ im Milstein-Verfahren erreichen.

Es sei zur Abkürzung

$$I_{ij} := \int_{t_q}^{t_{q+1}} \int_{t_q}^s dW_i(u) dW_j(s) \quad \text{für } i, j = 1, \dots, n. \tag{5.15}$$

Analog zu (5.9) folgt für die Doppelintegrale bezüglich nur eines Wiener Prozesses

$$I_{ii} := \int_{t_q}^{t_{q+1}} \int_{t_q}^s dW_i(u) dW_i(s) = \frac{1}{2} [(\Delta W_i)^2 - h] \quad \text{für } i = 1, \dots, n. \quad (5.16)$$

Kompliziert sind nur die Auswertungen für $i \neq j$. Man beachte auch, dass im allgemeinen $I_{ij} \neq I_{ji}$ gilt.

Wir können die Doppelintegrale I_{ij} für $i \neq j$ als Lösungen von einfachen stochastischen Differentialgleichungen darstellen. O.E.d.A. sei I_{21} betrachtet. Dann stellen wir das System

$$Y_1(t) = \int_{t_q}^t Y_2(s) dW_1(s), \quad Y_2(t) = \int_{t_q}^t dW_2(s) \quad (5.17)$$

für $t_q \leq t \leq t_{q+1}$ auf. Für die zweite Komponente gilt

$$Y_2(t) = W_2(t) - W_2(t_q) \quad \Rightarrow \quad Y_2(t_{q+1}) = \Delta W_2.$$

Für die erste Komponente folgt

$$Y_1(t_{q+1}) = \int_{t_q}^{t_{q+1}} Y_2(s) dW_1(s) = \int_{t_q}^{t_{q+1}} \int_{t_q}^s dW_2(u) dW_1(s).$$

Die erste Komponente der Lösung von (5.17) an der Stelle t_{q+1} entspricht somit I_{21} . Um diese Lösung näherungsweise zu erhalten, führen wir M Schritte des Euler-Maruyama-Verfahrens im Intervall $[t_q, t_{q+1}]$ durch, d.h. die Schrittweite beträgt $\tilde{h} = \frac{h}{M}$. Auf diese Weise erhalten wir eine Näherung $Z^{\tilde{h}}$ der Lösung von (5.17) an der Stelle t_{q+1} . Wir verwenden die kleinste Schrittzahl M mit $\frac{1}{h} \leq M$. Da das Euler-Maruyama-Verfahren die starke Konvergenzordnung $\frac{1}{2}$ besitzt, folgt

$$\mathbb{E} \left(\left| Z_1^{\tilde{h}} - I_{21} \right| \right) \leq C \sqrt{\tilde{h}} = C \sqrt{\frac{h}{M}} \leq C \sqrt{h^2} = Ch.$$

Insgesamt haben wir im Milstein-Verfahren eine starke Konvergenzordnung $p = 1$ erreicht. Als Preis dafür steigt der Aufwand zur Bestimmung der Doppelintegrale (5.15) hier quadratisch mit der Anzahl der Integrationsschritte auf einem Zeitintervall $[0, T]$ an.

In manchen Spezialfällen von Koeffizientenfunktionen b_{ij} (kommutatives Rauschen, lineares Rauschen) lassen sich die Doppelintegrale auf die Inkremente ΔW_i zurückführen, was den Aufwand erheblich vereinfacht. Im Fall der einer Diagonalmatrix b_{ij} (diagonales Rauschen) treten nur die Doppelintegrale (5.16) auf. Hängen alle b_{ij} nicht von x ab (additives Rauschen), dann reduziert sich das

Milstein-Verfahren zum Euler-Maruyama-Verfahren. Das Euler-Maruyama-Verfahren besitzt somit bei additivem Rauschen die starke Konvergenzordnung 1.

Wieder können im Milstein-Verfahren (5.14) für Systeme die Terme mit Ableitungen von b durch Differenzenquotienten approximiert werden, siehe [6]. Der resultierende Aufwand im Verfahren ist dann deutlich geringer als wenn die Ableitungen direkt durch numerische Differentiation bestimmt werden.

5.5 Verfahren höherer Ordnung

Über stochastische Taylor-Entwicklungen höherer Ordnung können auch Verfahren höherer Konvergenzordnung hergeleitet werden. Wir wollen zwei Beispiele betrachten ohne die Konstruktion nachzuvollziehen. Dabei betrachten wir nur einzelne Itô-Integralgleichungen mit einem eindimensionalen Wiener Prozess. Zu Zeitpunkten $t_i = ih$ für $i = 0, 1, \dots, N$ mit $h = \frac{T}{N}$ werden Näherungen \tilde{X}_i berechnet. Es bezeichnet wieder $\Delta W_i = W_{t_{i+1}} - W_{t_i}$, d.h. $\Delta W_i \sim N(0, h)$.

Methode mit starker Konvergenzordnung 1.5

Gegeben sei die Itô-Differentialgleichung $dX_t = a(t, X_t)dt + b(t, X_t)dW_t$ mit Anfangswert X_0 . Ein numerisches Verfahren wird festgesetzt durch die Rekursion

$$\begin{aligned}\tilde{X}_{i+1} &= \tilde{X}_i + ah + b\Delta W_i + \frac{1}{2}bb'(\Delta W_i - h) + a'b\Delta V_i \\ &\quad + \frac{1}{2}(aa' + \frac{1}{2}b^2a'')h^2 + (ab' + \frac{1}{2}b^2b'')(h\Delta W_i - \Delta V_i) \\ &\quad + \frac{1}{2}b(bb'' + b'^2)(\frac{1}{3}(\Delta W_i^2) - h)\Delta W_i\end{aligned}$$

für $i = 0, 1, \dots, N - 1$. Darin werden alle Koeffizientenfunktionen und deren Ableitungen bei $t = t_i$ und $x = \tilde{X}_i$ ausgewertet. Die Ableitungen a', b', a'', b'' bezeichnen die partiellen Ableitungen von a bzw. b nach x . Die neue Zufallsvariable ΔV_i ist normalverteilt und korreliert mit ΔW_i . Es muss gelten

$$\begin{aligned}\mathbb{E}(\Delta V_i) &= 0 \\ \mathbb{E}(\Delta V_i^2) &= \frac{1}{3}h^3 \\ \mathbb{E}(\Delta V_i\Delta W_i) &= \frac{1}{2}h^2.\end{aligned}$$

Dies kann erreicht werden durch eine unabhängige Zufallsvariable $\Delta U_i \sim N(0, h)$ oder äquivalent $\Delta U_i = \sqrt{h}Z_i$ mit $Z_i \sim N(0, 1)$. Setzen wir

$$\Delta V_i = \frac{1}{2}h \left(\Delta W_i + \frac{1}{\sqrt{3}} \Delta U_i \right),$$

dann sind die drei Bedingungen erfüllt.

Im Vergleich zum Euler-Maruyama-Verfahren ergibt sich pro Schritt ein Mehraufwand durch: zwei Pseudo-Zufallszahlen statt einer und Auswertung von vier Ableitungen a', a'', b', b'' .

Methoden mit schwacher Konvergenzordnung 2

Gegeben sei die autonome Itô-Differentialgleichung $dX_t = a(X_t)dt + b(X_t)dW_t$ mit Anfangswert X_0 . Ein numerisches Verfahren wird rekursiv festgelegt. Dazu sei

$$\begin{aligned} c(x) &= \frac{1}{2}a'(x)a(x) + \frac{1}{4}a''(x)b(x)^2 \\ d(x) &= \frac{1}{2}a'(x)b(x) + \frac{1}{2}b'(x)a(x) + \frac{1}{4}b(x)^2b''(x) \end{aligned}$$

und damit $\hat{a}_h = a + hc$ sowie $\hat{b}_h = b + hd$. Die Rekursion lautet

$$\tilde{X}_{i+1} = \tilde{X}_i + \hat{a}_h(\tilde{X}_i)h + \hat{b}_h(\tilde{X}_i)\Delta W_i + \frac{1}{2}b'(\tilde{X}_i)b(\tilde{X}_i)((\Delta W_i)^2 - h)$$

für $i = 0, 1, \dots, N - 1$. Diese Methode besitzt die Form des Milstein-Verfahrens (5.13), wobei die Koeffizientenfunktionen a und b abgeändert werden in \hat{a}_h und \hat{b}_h . Diese Erweiterung bewirkt, dass das Verfahren jetzt die schwache Konvergenzordnung 2 besitzt. Im Vergleich zum Euler-Maruyama-Verfahren besteht pro Schritt ein Mehraufwand durch die Auswertung der vier Ableitungen von a', a'', b', b'' .

Wir weisen die schwache Konvergenzordnung 2 für den Fall der geometrischen Brownschen Bewegung direkt nach. Die geometrische Brownsche Bewegung (4.16) erfüllt die lineare Integralgleichung (4.15). Es gilt hier $a(x) = \alpha x$ und $b(x) = \beta x$. Damit ist $\hat{a}_h(x) = (\alpha + \frac{1}{2}h\alpha^2)x$ und $\hat{b}_h(x) = (\beta + h\alpha\beta)x$. Es folgt die Rekursion

$$\begin{aligned} \tilde{M}_{i+1} &= \tilde{M}_i + (h\alpha + \frac{1}{2}h^2\alpha^2)\tilde{M}_i + (\beta + h\alpha\beta)\Delta W_i\tilde{M}_i + \frac{1}{2}\beta^2((\Delta W_i)^2 - h)\tilde{M}_i \\ &= \tilde{M}_i \left((1 + h\alpha + \frac{1}{2}h^2\alpha^2) + (\beta + h\alpha\beta)\Delta W_i + \frac{1}{2}\beta^2((\Delta W_i)^2 - h) \right). \end{aligned}$$

Sukzessive erhalten wir

$$\tilde{M}_N = M_0 \prod_{i=0}^{N-1} \left(1 + h\alpha + \frac{1}{2}h^2\alpha^2 + (\beta + h\alpha\beta)\Delta W_i + \frac{1}{2}\beta^2((\Delta W_i)^2 - h) \right).$$

Bildung des Erwartungswerts führt wegen $\mathbb{E}(\Delta W_i) = 0$ und $\mathbb{E}(\Delta W_i^2) = h$ auf

$$\mathbb{E}(\tilde{M}_N) = \mathbb{E}(M_0) \prod_{i=0}^{N-1} \left(1 + h\alpha + \frac{1}{2}h^2\alpha^2 \right) = \mathbb{E}(M_0) (1 + h\alpha + \frac{1}{2}(h\alpha)^2)^N.$$

Wir erkennen hier die ersten Terme der Taylor-Entwicklung von $e^{\alpha t}$ um $t = 0$. Es folgt, dass diese Approximation mit Konvergenzordnung 2 gegen $\mathbb{E}(M_T)$ läuft.

5.6 Simulationsbeispiel: SIR-Modell

Ein einfaches epidemiologisches Modell für die Dynamik einer Infektionskrankheit ist das SIR-Modell, siehe [4]. Dabei treten drei Populationsgruppen auf: Suszeptible (S), Infizierte (I) und Resistente (R). Es entsteht ein System aus gewöhnlichen Dgln.

$$\begin{aligned}\frac{dS}{dt}(t) &= -\beta S(t)I(t) \\ \frac{dI}{dt}(t) &= \beta S(t)I(t) - \gamma I(t) \\ \frac{dR}{dt}(t) &= \gamma I(t)\end{aligned}\tag{5.18}$$

mit zwei positiven Konstanten, nämlich die Ansteckungsrate β und die Genesungsrate γ . Nichtnegative Anfangswerte geben wir o.E.d.A. bei $t = 0$ vor. Dann sind alle drei Populationen stets nichtnegativ. Zudem ist die Summe $S + I + R$ in der Zeit konstant. Daher können wir auch o.E.d.A. $S(0) + I(0) + R(0) = 1$ setzen. Typischerweise entstehen für $t \rightarrow \infty$ konstante Lösungen in der Zeit.

Wir fügen nun zufallsabhängige Anteile hinzu, die als ein Rauschen interpretiert werden können. Es folgt ein autonomes System aus Itô-Dgln.

$$\begin{aligned}dS_t &= -\beta S_t I_t dt - \sigma_1 (\beta S_t I_t)^q dW_1(t) \\ dI_t &= (\beta S_t I_t - \gamma I_t) dt + \sigma_1 (\beta S_t I_t)^q dW_1(t) - \sigma_2 (\gamma I_t)^q dW_2(t) \\ dR_t &= \gamma I_t dt + \sigma_2 (\gamma I_t)^q dW_2(t)\end{aligned}\tag{5.19}$$

mit Diffusionskonstanten $\sigma_1, \sigma_2 > 0$, einem Exponenten $q > 0$ sowie einem zweidimensionalen Wiener Prozess (W_1, W_2) . Dieses System besitzt die Gestalt (5.10) mit $m = 3$ Komponenten und $n = 2$ als Dimension des Wiener Prozesses.

In der Literatur wird oft der Exponent $q = \frac{1}{2}$, d.h. Wurzelfunktionen, verwendet, siehe z.B. [1]. Jedoch ist dann die Nichtnegativität der Lösungen nicht gegeben. Deshalb verwenden wir im folgenden den Exponenten $q = 1$. Als Anfangswerte geben wir vor

$$S_0 = 0.9, \quad I_0 = 0.1, \quad R_0 = 0.$$

Als weitere Parameter wird gesetzt

$$\beta = 0.2, \quad \gamma = 0.1, \quad \sigma_1 = \sigma_2 = 0.8.$$

Bei dieser Wahl endet das Infektionsgeschehen bevor sich alle Suszeptiblen infizieren.

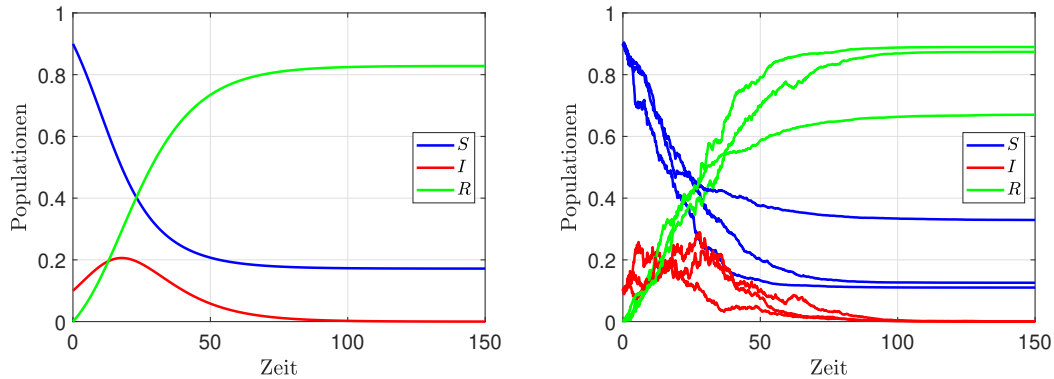


Abbildung 10: Lösung des deterministischen SIR-Modells (5.18) (links) und drei Pfade des stochastischen SIR-Modells (5.19) (rechts) bei Parametern $\beta = 0.2$, $\gamma = 0.1$.

Abbildung 10 (links) zeigt die Populationen aus dem zugehörigen deterministischen Modell (5.18), welches mit einem numerischen Verfahren höherer Ordnung für gewöhnliche Dgln. berechnet wurde. Wir lösen das Modell (5.19) numerisch mit dem Euler-Maruyama-Verfahren aus Abschnitt 5.1. Als Schrittweite wird $h = \Delta t = 0.1$ verwendet. Abbildung 10 (rechts) stellt drei berechnete Pfade zu den Populationen dar.

Desweiteren führen wir eine Monte-Carlo-Simulation durch, wobei wir Näherungen zu 10^4 Pfaden des stochastischen Modells (5.19) bestimmen. Als Approximationen für den Erwartungswert und die Varianz zu den Populationen wird der Mittelwert und die Stichprobenvarianz gebildet. Die Ergebnisse sind in Abbildung 11 enthalten.

Zum Vergleich ändern wir die Genesungsrate zu $\gamma = 0.01$, während die anderen Einstellungen identisch bleiben. Als Konsequenz werden alle Suszeptiblen mit der Zeit infiziert. Wieder erstellen wir eine Monte-Carlo-Simulation mit 10^4 Pfaden zum stochastischen Modell (5.19). Die Lösungen werden über ein längeres Zeitintervall betrachtet. Abbildung 12 zeigt die Ergebnisse für den Erwartungswert und die Varianz. Dabei unterscheidet sich der Erwartungswert nun kaum von der Lösung des deterministischen Modells (5.18).

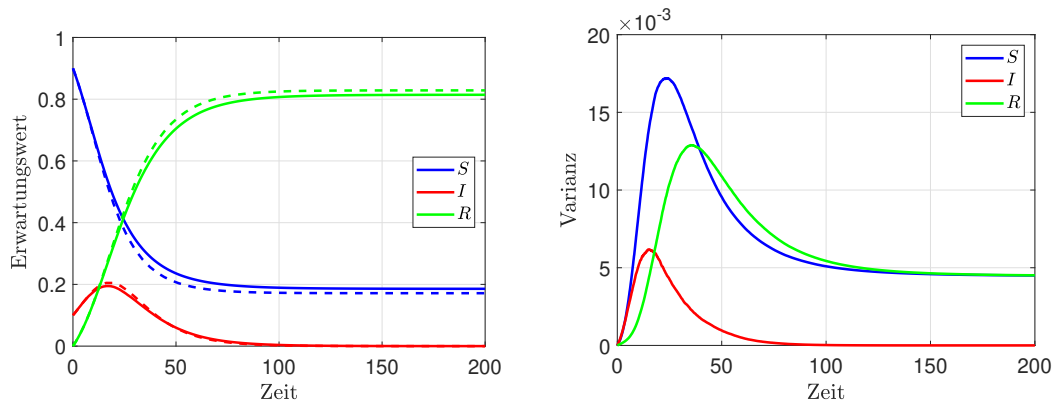


Abbildung 11: Erwartungswerte (durchgezogenen Linien) und deterministische Lösung (gestrichelte Linien) (links) sowie Varianz (rechts) im stochastischen SIR-Modell bei Parametern $\beta = 0.2$, $\gamma = 0.1$.

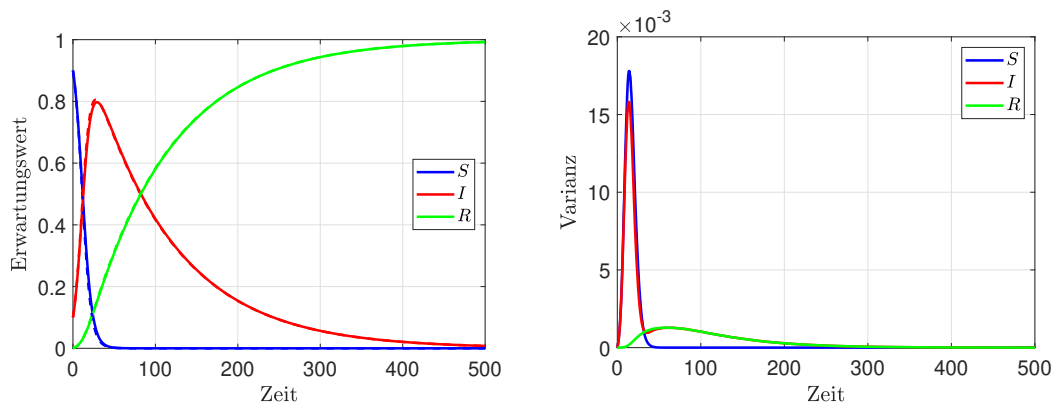


Abbildung 12: Erwartungswerte (durchgezogenen Linien) und deterministische Lösung (gestrichelte Linien) (links) sowie Varianz (rechts) im stochastischen SIR-Modell bei Parametern $\beta = 0.2$, $\gamma = 0.01$.

6 Anwendungen in der Finanzmathematik

In diesem Kapitel behandeln wir Anwendungen von Modellen mit stochastischen Differentialgleichungen in der Finanzmathematik. Dabei werden Aktienkurse und die Bepreisung von zugehörigen Optionen dargestellt.

6.1 Modell für Aktienkurse

Abbildung 13 zeigt die Aktienkurse der Firmen Siemens und Volkswagen an den Börsentagen in einem Zeitraum von 10 Jahren, wobei ein Wert pro Tag im Datensatz enthalten ist. Wir beobachten zwei Eigenschaften in diesen Aktienkursen:

1. Die Kurse besitzen schnelle Schwankungen auf einer kleinen Zeitskala. Dieses Bild ist ähnlich zum lokalen Verhalten der Pfade eines Wiener Prozesses oder eines Itô-Prozesses.
2. Bis auf kurzfristige Ausnahmen besteht weitgehend ein Aufwärtstrend bei den Kursen. Ein Wiener Prozess besitzt eine solche langfristige Tendenz nicht.

Obgleich für die einzelnen Kursschwankungen oft konkrete Gründe vorliegen (z.B. Entwicklung der Gesamtwirtschaft, Maßnahmen in einer einzelnen Firma, etc.), ist eine zufallsabhängige Modellierung der Aktienkurse sinnvoll.

Die obigen Beobachtungen legen als Vermutung nahe, den Aktienkurs durch eine geometrische Brownsche Bewegung geeignet beschreiben zu können. Die geome-

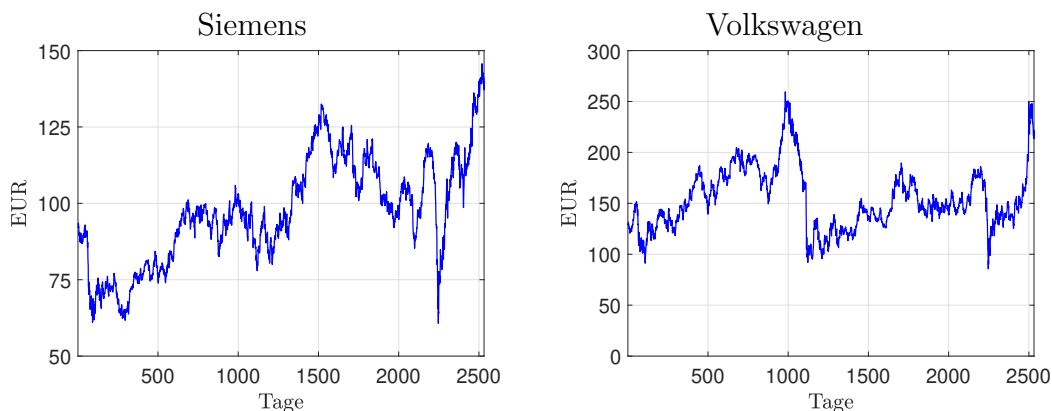


Abbildung 13: Aktienkurse der Firmen Siemens und Volkswagen an den Börsentagen im Zeitraum 6.5.2011 bis 6.5.2021.

trische Brownsche Bewegung ist die Lösung (4.16) der linearen Itô-Differentialgleichung (4.15). Wir schreiben diese Gleichung mit den in der Finanzmathematik üblichen Symbolen

$$S_t = S_0 + \int_0^t \mu S_s \, ds + \int_0^t \sigma S_s \, dW_s \quad (6.1)$$

für $t \geq 0$. Die Konstante $\mu > 0$ stellt einen Driftparameter dar. Die Konstante $\sigma > 0$ wird als Diffusionsparameter oder Volatilität bezeichnet. Vom Anfangswert wird $S_0 > 0$ gefordert. Die Lösung der Gleichung (6.1) lautet

$$S_t = S_0 e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma W_t} \quad (6.2)$$

mit eindimensionalen Wiener Prozess $(W_t)_{t \geq 0}$.

Wir führen eine Diskretisierung in der Zeit durch über $t_i = i\Delta t$ für $i = 0, 1, 2, \dots$ mit Schrittweite $\Delta t > 0$. Es sei $\Delta W_i = W_{t_{i+1}} - W_{t_i}$, wodurch gilt $\Delta W_i \sim N(0, \Delta t)$ für alle i . Mit (6.2) folgt

$$\frac{S_{t_{i+1}}}{S_{t_i}} = \frac{S_0 e^{(\mu - \frac{1}{2}\sigma^2)t_{i+1} + \sigma W_{t_{i+1}}}}{S_0 e^{(\mu - \frac{1}{2}\sigma^2)t_i + \sigma W_{t_i}}} = e^{(\mu - \frac{1}{2}\sigma^2)\Delta t + \sigma \Delta W_i}.$$

Weiter erhalten wir

$$L_i := \ln \left(\frac{S_{t_{i+1}}}{S_{t_i}} \right) = (\mu - \frac{1}{2}\sigma^2)\Delta t + \sigma \Delta W_i. \quad (6.3)$$

Dadurch liegt die Normalverteilung $L_i \sim N((\mu - \frac{1}{2}\sigma^2)\Delta t, \sigma^2 \Delta t)$ vor. Diese Verteilung hängt nicht mehr von i ab. Mit der Annahme $\mu > \frac{1}{2}\sigma^2$ gilt $\mathbb{E}(L_i) > 0$ und es besteht ein Aufwärtstrend.

Nun betrachten wir reale Kurse zu Aktien. Die Daten seien bezeichnet mit (i, \hat{S}_i) für $i = 1, 2, \dots$, d.h. $t_i = i$ und $\Delta t = 1$. Die Werte

$$\ell_i := \ln \left(\frac{\hat{S}_{i+1}}{\hat{S}_i} \right) \quad (6.4)$$

entsprechen der Konstruktion in (6.3). Abbildung 14 zeigt diese Werte für die beiden Aktien von Siemens und Volkswagen aus Abbildung 13. Wir erkennen, dass bis auf kurzfristige Ausnahmen die Daten ähnlich einer Normalverteilung sind. Statistische Untersuchungen eines Aktienkurses in diesem Kontext finden sich auch in Abschnitt 5.3 aus [9]. Dies motiviert als Modell für den zufallsabhängigen Aktienkurs die geometrische Brownsche Bewegung zu verwenden.

Als zusätzliches Beispiel betrachten wir den Deutschen Aktien-Index (DAX) mit einem Wert pro Börsentag in einem Zeitraum von 5 Jahren. Abbildung 15 stellt den DAX-Kurs sowie die zugehörigen Daten (6.4) dar. Das Verhalten entspricht den vorherigen Beispielen.

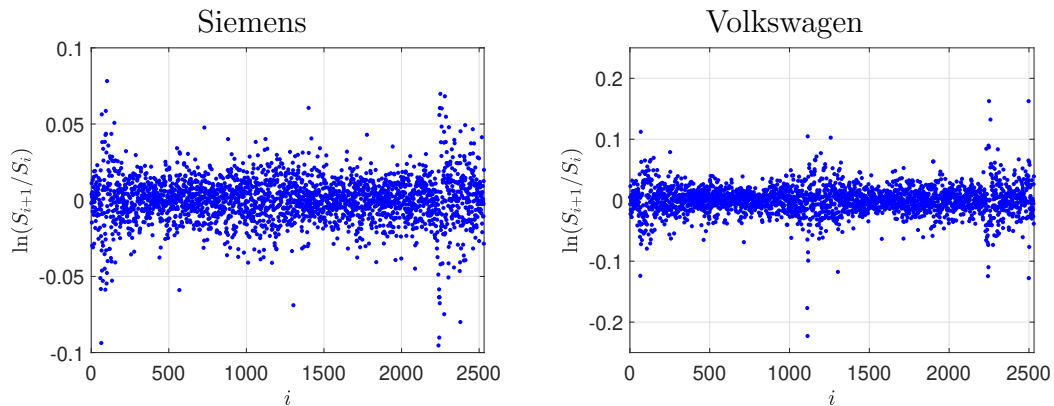


Abbildung 14: Daten (6.4) zu den Aktienkursen der Firmen Siemens und Volkswagen an den Börsentagen im Zeitraum 6.5.2011 bis 6.5.2021.

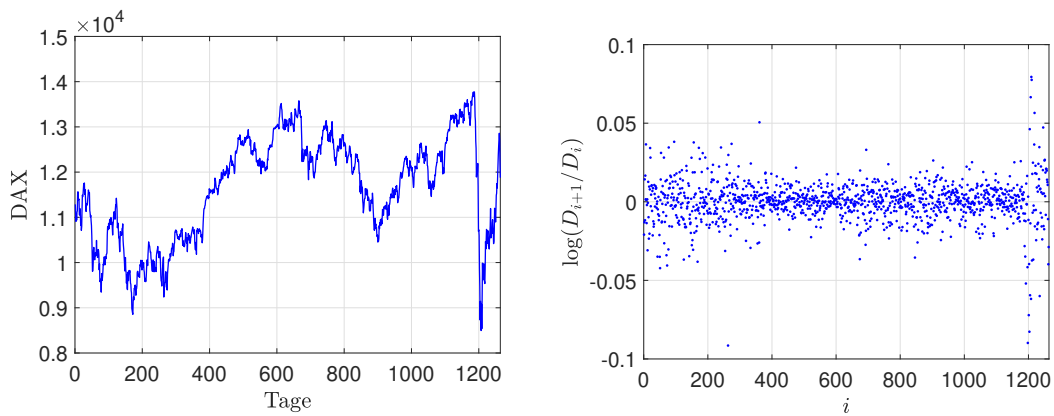


Abbildung 15: Werte des Deutschen Aktien-Index (DAX) an den Börsentagen im Zeitraum 12.6.2015 bis 12.6.2020 sowie zugehörige Daten (6.4).

6.2 Optionspreise

Folgende Modellannahmen werden für den Markt gemacht:

1. Geld kann risikofrei auf Bankkonten mit fester Zinsrate $r > 0$ angelegt werden.
2. Kredite können von Banken mit gleicher fester Zinsrate $r > 0$ in beliebiger Höhe erhalten werden.
3. Verzinsung erfolgt kontinuierlich über die Zeit.

4. Arbitrage-Freiheit: Es kann kein risikofreier Gewinn gemacht werden, der über die Zinsrate r hinausgeht.
5. Markt ist liquide und Handel jederzeit möglich.
6. Markt ist friktionslos (d.h. keine Transaktionskosten, keine Steuern, etc.).
7. Leerverkäufe von Aktien sind möglich.

Diese Annahmen sind teilweise erhebliche Idealisierungen.

Optionen

Jetzt charakterisieren wir Optionen auf jeweils eine feste Aktie. Zudem vernachlässigen wir Dividendenzahlungen bei den Untersuchungen.

Definition 6.1 *Eine Europäische Call-Option gibt dem Besitzer die Möglichkeit (nicht die Verpflichtung) eine bestimmte Aktie zu einem bestimmten Preis an einem bestimmten Zeitpunkt in der Zukunft zu kaufen.*

Eine Europäische Put-Option gibt dem Besitzer die Möglichkeit (nicht die Verpflichtung) eine bestimmte Aktie zu einem bestimmten Preis an einem bestimmten Zeitpunkt in der Zukunft zu verkaufen.

Es liegen (grob) folgende Typen von Optionen vor:

- Europäische Optionen,
- Amerikanische Optionen,
- Exotische Optionen
 - Asiatische Optionen,
 - Barriere-Optionen,
 - und andere.

Diese Papiere gibt es jeweils als Call-Option (Möglichkeit zu kaufen) und als Put-Option (Möglichkeit zu verkaufen).

Wir behandeln den Fall von Europäischen Call-Optionen aus Definition 6.1 näher. Der heutige Zeitpunkt sei $t = 0$, während der Ausübungszeitpunkt der Option bei $t = T > 0$ liegt. Der Ausübungspreis (Kaufpreis der Aktie über Option) sei $E > 0$. Zudem sei S_T der Aktienpreis am Endzeitpunkt. Es ergibt sich die Fallunterscheidung:

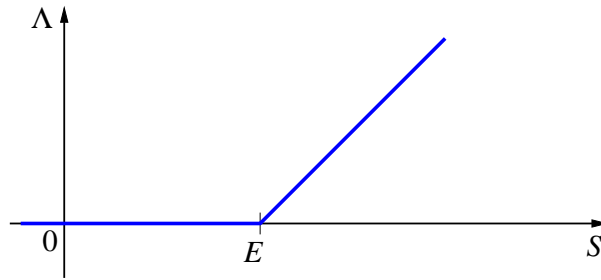


Abbildung 16: Auszahlungsfunktion bei einer Europäischen Call-Option.

1. Fall: $S_T > E$
Option wird ausgeübt. (Aktie wird zum Preis E gekauft.)
2. Fall: $S_T < E$
Option wird nicht ausgeübt. (Aktie wird nicht zum Preis E gekauft.)
3. Fall: $S_T = E$
Ausübung der Option hat weder Vorteil noch Nachteil.

Der eventuelle Gewinn am Endzeitpunkt T ist $\Lambda(S_T)$ mit der Auszahlungsfunktion

$$\Lambda(S) = \max \{S - E, 0\}, \quad (6.5)$$

welche in Abbildung 16 dargestellt ist. Die Funktion Λ ist nichtlinear (nur stückweise linear), jedoch global Lipschitz-stetig. In diesem Gewinn sind jedoch noch nicht die Kosten zum Kauf der Option berücksichtigt.

Um einen fairen Optionspreis zu motivieren definieren wir: Bei einem *fairen Spiel* ist der Erwartungswert des Gewinns null.

Beispiel: Münzwurf

Der Spieler setzt N EUR ein. Bei Ergebnis „Zahl“ erhält der Spieler $2N$ EUR zurück. Bei Ergebnis „Kopf“ erhält der Spieler nichts zurück, d.h. der Einsatz wird einbehalten. Als Erwartungswert folgt

$$\mathbb{E}(\text{Gewinn}) = \mathbb{E}(\text{Auszahlung}) - \text{Einsatz} = \left(\frac{1}{2} \cdot 2N + \frac{1}{2} \cdot 0\right) - N = 0.$$

Das festgelegte Spiel ist somit fair.

Im Kontext der Optionen auf Aktien ist der Kaufpreis einer Option wie ein Einsatz zu betrachten. Wir kennen den heutigen Aktienkurs S_0 . Der Aktienkurs S_T

sei eine Zufallsvariable mit gegebener Verteilung. Dadurch wird auch die Auszahlungsfunktion (6.5) zu einer Zufallsvariable. Es bezeichnet V_0 den heutigen Kaufpreis für die Option. Wir fordern in Analogie eines fairen Spiels nun

$$\mathbb{E}(„Gewinn“) = \mathbb{E}(\Lambda(S_T)) - V_0 \stackrel{!}{=} 0$$

mit der Auszahlungsfunktion (6.5) für Europäische Call-Optionen oder anderen Auszahlungsfunktionen für andere Optionstypen. Als Optionspreis würden wir daraus $V_0 = \mathbb{E}(\Lambda(S_T))$ erhalten. Jedoch ist die Zeit T bis zur Ausübungsmöglichkeit mit zu betrachten, weil in dieser Zeitspanne auch ein Geldbetrag V_0 auf einem Konto zur Zinsrate r risikofrei angelegt werden könnte. Als fairer Optionspreis ergibt sich daher der diskontierte Gewinn am Endzeitpunkt, d.h.

$$V_0 = e^{-rT} \mathbb{E}(\Lambda(S_T)). \quad (6.6)$$

Dieser Preis ist sowohl für Käufer als auch Verkäufer der Option in gewissem Sinne fair.

Black-Scholes-Formel

Die Beschreibung Europäischer Optionen ist noch einfach genug, so dass in diesem Fall eine Formel für den fairen Preis hergeleitet werden kann. Dabei wird die Annahme $\mu = r$ gemacht, d.h. die Driftrate zum Aktienkurs entspricht der risikofreien Zinsrate. Diese Annahme beschreibt eine Risikoneutralität, siehe Abschnitt 12.3 in [9].

Als Abkürzung setzen wir

$$d_1 = \frac{\ln\left(\frac{S_0}{E}\right) + \left(r + \frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}} \quad \text{und} \quad d_2 = \frac{\ln\left(\frac{S_0}{E}\right) + \left(r - \frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}.$$

Desweiteren bezeichnet

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{1}{2}s^2} ds$$

die Verteilungsfunktion der Standard-Normalverteilung. Die Black-Scholes-Formel für den fairen Preis einer Europäischen Call-Option lautet dann

$$V_0 = S_0 \Phi(d_1) - E e^{-rT} \Phi(d_2). \quad (6.7)$$

Eine direkte Auswertung dieser Formel ist daher möglich, wenn die fünf Parameter r, σ, T, S_0, E eingegeben werden. Eine Formel dieser Art existiert nicht für Amerikanische Optionen oder Exotische Optionen, d.h. dort sind wir auf Näherungsverfahren angewiesen.

Monte-Carlo-Simulation

Wir erzeugen Stichproben des Aktienkurses zum Endzeitpunkt T . Gemäß der geometrischen Brownschen Bewegung (6.2) folgt

$$s_j = S_0 e^{(\mu - \frac{1}{2}\sigma^2)T + \sigma\sqrt{T}z_j}$$

für $j = 1, \dots, M$, wobei die Pseudo-Zufallszahlen z_j zu unabhängigen Standard-Normalverteilungen gehören.

Als Näherung des Erwartungswerts zur Auszahlungsfunktion folgt der Mittelwert

$$\mathbb{E}(\Lambda(S_T)) \approx \frac{1}{M} \sum_{j=1}^M \Lambda(s_j).$$

Die Approximation zum fairen Preis (6.6) einer Option lautet somit

$$\tilde{V}_0 = \frac{e^{-rT}}{M} \sum_{j=1}^M \Lambda(s_j).$$

Diese Näherung gilt für alle Optionstypen mit jeweiliger Auszahlungsfunktion Λ am Endzeitpunkt. Diese numerische Methode für eine Näherung in einem zufallsabhängigen Modell wird als *Monte-Carlo-Simulation* bezeichnet.

Als Beispiel setzen wir die Parameter auf $r = \mu = 0.02$, $\sigma = 0.1$, $T = 1$, $S_0 = 10$, $E = 9$. Als Stichprobenumfänge wird $M = 2^k$ mit $k = 7, 8, \dots, 23$ verwendet. Zum Vergleich wird für die obigen Parameter auch die Black-Scholes-Formel (6.7) ausgewertet. Abbildung 17 stellt die Ergebnisse dar. Wir beobachten, dass für steigende Stichprobenanzahl die Näherungen aus der Monte-Carlo-Simulation gegen die Auswertung der Black-Scholes-Formel konvergieren. Abbildung 18 zeigt die Differenzen zwischen den Näherungen aus der Monte-Carlo Simulation und dem Wert aus der Black-Scholes-Formel. Wir bemerken, dass dieser Fehler etwa proportional zu $\frac{1}{\sqrt{M}}$ gegen null konvergiert.

Asiatische Optionen

Bei Asiatischen Optionen handelt es sich um *pfadabhängige Optionen*, d.h. die Auszahlungsfunktion hängt nicht nur vom Aktienkurs am Endzeitpunkt T ab sondern von allen Aktienkursen $S_{[0,T]} = \{(t, S_t) : t \in [0, T]\}$. Wir beschränken uns im folgenden auf Call-Optionen, da die Untersuchung für Put-Optionen analog verläuft.

Es bezeichne \bar{S} einen Mittelwert des Aktienkurses. Der arithmetische Mittelwert und der geometrische Mittelwert lauten

$$\bar{S}_{\text{arith}} = \frac{1}{T} \int_0^T S_t dt \quad \text{bzw.} \quad \bar{S}_{\text{geom}} = \exp\left(\frac{1}{T} \int_0^T \ln(S_t) dt\right).$$

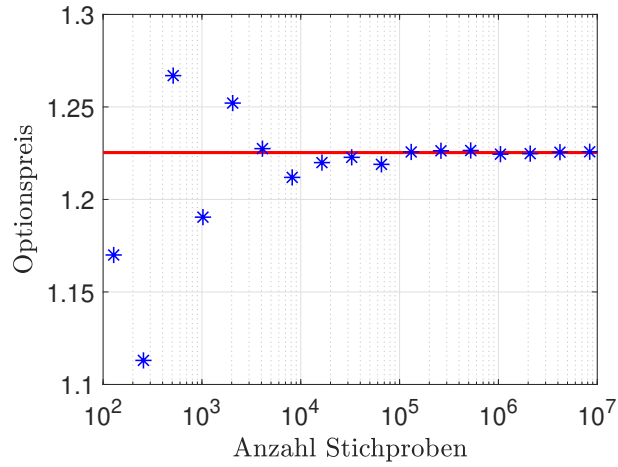


Abbildung 17: Näherungen (blau) aus Monte-Carlo-Simulation zu verschiedenen Anzahlen der Stichproben und Wert (rot) aus Black-Scholes-Formel bezüglich des Preises einer Europäischen Option.

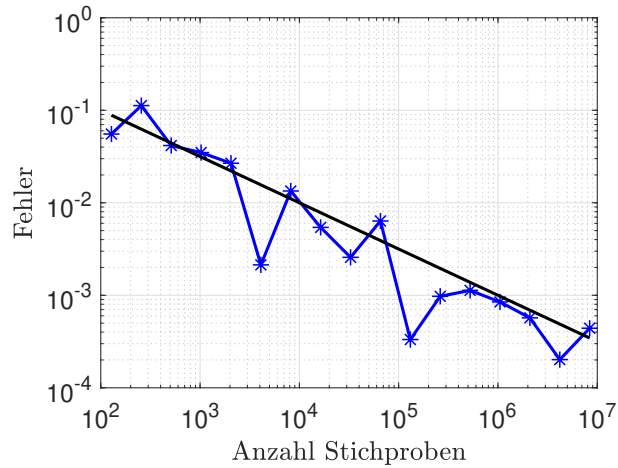


Abbildung 18: Differenzen (blau) zwischen Näherungen aus der Monte-Carlo-Simulation und Wert aus Black-Scholes-Formel bezüglich des Preises einer Europäischen Option sowie Funktion $\frac{1}{\sqrt{M}}$ (schwarz) in doppelt-logarithmischer Skala.

Mit den Zeitpunkten $t_i = i\Delta t$ für $i = 0, 1, \dots, N$ und Zeitschrittweite $\Delta t = \frac{T}{N}$ folgen als diskrete Approximationen

$$\widehat{S}_{\text{arith}} = \frac{1}{N} \sum_{i=1}^N S_{t_i} \quad \text{bzw.} \quad \widehat{S}_{\text{geom}} = \left(\prod_{i=1}^N S_{t_i} \right)^{\frac{1}{N}}. \quad (6.8)$$

Bei den *Rate-Optionen* (auch: *fixed strike option*) wird anstelle der Auszahlungsfunktion (6.5) hier

$$\Lambda(S_{[0,T]}) = \max \{ \bar{S} - E, 0 \} \quad (6.9)$$

mit einem Mittelwert \bar{S} gesetzt. Dieser Optionstyp ist im Vergleich zu Europäischen Optionen weniger anfällig gegenüber eventuellen Manipulationen am Endzeitpunkt.

Ein anderer Typ von Asiatischen Optionen sind die *Strike-Optionen* (auch: *floating strike option*). Die Auszahlungsfunktion lautet dabei

$$\Lambda(S_{[0,T]}) = \max \{ S_T - \bar{S}, 0 \} \quad (6.10)$$

mit einem Mittelwert \bar{S} . Der Gewinn ist somit umso höher, je mehr der Aktienkurs am Endzeitpunkt den mittleren Aktienkurs übertrifft.

Der Mittelwert \bar{S} in (6.9) oder (6.10) muss durch eine Approximation wie (6.8) ersetzt werden. Daher sind die Werte S_{t_i} für $i = 1, \dots, N$ zu bestimmen. Wir verwenden als Modell für den Aktienkurs wieder die geometrische Brownsche Bewegung. Hier müssen ganze Pfade erzeugt werden. Dies geschieht über die numerische Simulation von Pfaden $W_t(\omega_j)$ eines Wiener Prozesses wie bereits in Abschnitt 2.3 beschrieben. Wir setzen rekursiv

$$w_{i+1,j} = w_{i,j} + \sqrt{\Delta t} z_{i,j} \quad (6.11)$$

für $i = 0, 1, \dots, N-1$ und $j = 1, \dots, M$, wobei z_{ij} Pseudo-Zufallszahlen zu unabhängigen Standard-Normalverteilungen bezeichnen. Die Anfangswerte sind $w_{0,j} = 0$ für alle j . Mit der Formel (6.2) für die geometrische Brownsche Bewegung ergibt sich mit den Realisierungen (6.11) für $s_{i,j} = S_{t_i}(\omega_j)$ nun

$$s_{i,j} = S_0 e^{(\mu - \frac{1}{2}\sigma^2)t_i + \sigma w_{i,j}} \quad (6.12)$$

für $i = 1, \dots, N$ und $j = 1, \dots, M$.

Die Monte-Carlo-Simulation führt im Fall der Strike-Option mit Auszahlungsfunktion (6.10) für den fairen Preis auf die Näherung

$$\tilde{V}_0 = \frac{e^{-rT}}{M} \sum_{j=1}^M \max \left\{ s_{N,j} - \frac{1}{N} \sum_{i=1}^N s_{i,j}, 0 \right\}.$$

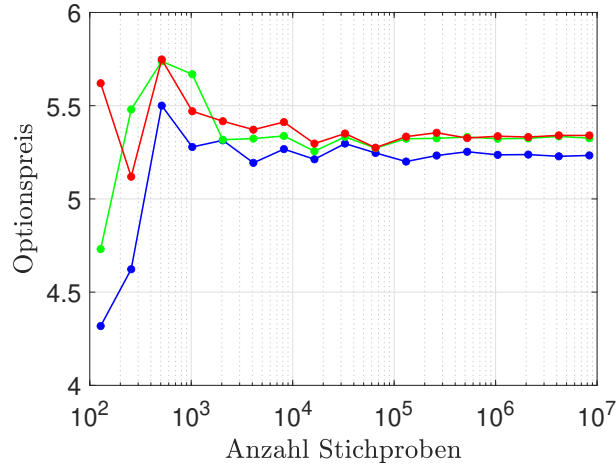


Abbildung 19: Näherungen für den Preis einer Asiatischen Option aus Monte-Carlo-Simulation zu verschiedenen Anzahlen der Stichproben und verschiedenen Zeitschrittweiten $\Delta t = 0.1$ (blau), $\Delta t = 0.01$ (grün) und $\Delta t = 0.001$ (rot).

Um eine hinreichend gute Näherung zu erhalten muss nicht nur die Stichprobenanzahl M hinreichend hoch sein sondern auch die Zeitschrittweite Δt hinreichend klein sein für die Mittelwertbildung.

Die Auswertungen der Exponentialfunktion in (6.12) können vermieden werden durch Verwendung des Euler-Maruyama Verfahrens aus Abschnitt 5.1. Wir erhalten die Rekursion

$$s_{i+1,j} = s_{i,j} \left(1 + \mu\Delta t + \sigma\sqrt{\Delta t} z_{i,j} \right) \quad (6.13)$$

für $i = 0, 1, \dots, N - 1$ und $j = 1, \dots, M$. Startwert ist $s_{0,j} = S_0$ für alle j . Bei dieser Näherung entsteht ein zusätzlicher kleiner Fehler gegenüber (6.12).

Als Beispiel betrachten wir eine Strike-Option mit dem diskreten arithmetischen Mittel aus (6.8). Die verwendeten Parameter sind $r = \mu = 0.03$, $\sigma = 0.2$, $T = 1$, $S_0 = 100$. Wir führen Monte-Carlo-Simulationen mit den Stichprobenumfängen $M = 2^k$ für $k = 7, 8, \dots, 23$ durch. Das Euler-Maruyama-Verfahren (6.13) wird eingesetzt. Zudem werden verschiedene Zeitschrittweiten $\Delta t = 10^{-\ell}$ für $\ell = 1, 2, 3$ verwendet. Wir bemerken, dass in allen drei Fällen verschiedener Zeitschrittweiten die Näherungen für steigende Stichprobenanzahlen jeweils konvergieren. Für $\Delta t = 0.01$ und $\Delta t = 0.001$ liegen für hohe Stichprobenumfänge die Näherungen nahe beieinander. Dieses Verhalten deutet darauf hin, dass die Zeitschrittweite $\Delta t = 0.01$ bereits hinreichend genaue Näherungen liefert.

6.3 Erweiterte Modelle für Aktienkurse

Wir betrachten verfeinerte Modelle für Aktienpreise im Vergleich zur geometrischen Brownschen Bewegung.

Stochastische Zinsrate

Die zeitliche Entwicklung eines Aktienkurses S_t unter Zufallseinfluss kann durch die lineare Itô-Differentialgleichung (6.1) modelliert werden. Darin treten die Driftrate μ und die Volatilität σ als positive Konstanten auf. Wir nehmen $\mu = r$ im Sinne der Risikoneutralität an, wobei r die konstante Zinsrate bzw. den Zinssatz bezeichnet.

Ein verfeinertes Modell entsteht, wenn die Zinsrate zeitabhängig und unter Zufallseinfluss betrachtet wird. Somit wird die Zinsrate zu einem eigenen Zufallsprozess R_t . Das *Modell von Vasicek* lautet

$$\begin{aligned}dS_t &= R_t S_t dt + \sigma S_t dW_1(t) \\dR_t &= \kappa(\theta - R_t) dt + \xi dW_2(t),\end{aligned}\tag{6.14}$$

mit Konstanten $\kappa, \theta, \xi > 0$. Dabei ist $\theta > 0$ ein Mittelwert zu dem R_t tendiert. Der Parameter $\kappa > 0$ bestimmt die Geschwindigkeit in dieser Tendenz. Der Parameter $\xi > 0$ beschreibt die Volatilität der Zinsrate. Somit entsteht ein System aus zwei stochastischen Differentialgleichungen mit einem zweidimensionalen Wiener Prozess. Im Modell (6.14) kann die Zinsrate leider negativ werden.

Alternativ wird im *Modell von Cox-Ingersoll-Ross* angesetzt

$$\begin{aligned}dS_t &= R_t S_t dt + \sigma S_t dW_1(t) \\dR_t &= \kappa(\theta - R_t) dt + \xi \sqrt{R_t} dW_2(t).\end{aligned}\tag{6.15}$$

Im Vergleich zu (6.14) liegt nur eine Änderung im stochastischen Teil zur Zinsrate vor. Die Bedingung $\kappa\theta > \frac{1}{2}\xi^2$ stellt sicher, dass Lösungen des Modells (6.15) positiv bleiben. Wir bemerken, dass die Struktur des Systems (6.15) ein diagonales Rauschen darstellt.

Stochastische Volatilität

Analog kann in der Itô-Differentialgleichung (6.1) statt des Parameters μ die Volatilität σ als ein Zufallsprozess σ_t dargestellt werden. Es sei $Y_t = \sigma_t^2$. Eine Möglichkeit hierzu ist das *Modell von Heston* [8]. Dieses beinhaltet ein System aus zwei Itô-Differentialgleichungen

$$\begin{aligned}dS_t &= \mu S_t dt + \sqrt{Y_t} S_t dW_t \\dY_t &= \kappa(\theta - Y_t) dt + \nu \sqrt{Y_t} dV_t\end{aligned}\tag{6.16}$$

mit zwei Wiener Prozessen $(W_t)_{t \geq 0}$ und $(V_t)_{t \geq 0}$. Jedoch ist $(V_t)_{t \geq 0}$ zu $(W_t)_{t \geq 0}$ korreliert mit der konstanten Korrelation $\rho = \rho(W_t, V_t)$. Dabei wird typischerweise eine negative Korrelation $\rho < 0$ gesetzt wegen des sogenannten Leverage-Effekts. Es ist $\theta > 0$ ein Mittelwert zu dem Y_t tendiert. Der Parameter $\kappa > 0$ beschreibt die Geschwindigkeit in dieser Tendenz. Die Konstante $\nu > 0$ bestimmt die Schwankungen in der Volatilität bzw. in deren Quadrat. Diese Parameter müssen geeignet an die Marktdaten angepasst werden. Die Korrelation der Prozesse kann erreicht werden durch

$$V_t = \rho W_1(t) + \sqrt{1 - \rho^2} W_2(t)$$

mit $W_1(t) = W_t$ und einem neuen Wiener Prozess $W_2(t)$ unabhängig von W_t . Wir erhalten das zu (6.16) äquivalente System

$$\begin{aligned} dS_t &= \mu S_t dt + \sqrt{Y_t} S_t dW_1(t) \\ dY_t &= \kappa(\theta - Y_t) dt + \nu \sqrt{Y_t} \rho dW_1(t) + \nu \sqrt{Y_t(1 - \rho^2)} dW_2(t). \end{aligned} \quad (6.17)$$

Es ist $Y_t \geq 0$ in diesem System sichergestellt. Durch die Tendenz zum Mittelwert θ erwarten wir, dass sich Lösungen des Heston-Modells (6.17) grob ähnlich verhalten wie Lösungen zu (6.1) mit $\sigma = \sqrt{\theta}$.

Wir führen eine numerische Simulation des Heston-Modells (6.17) durch mit den Parametern $\mu = 0.05$, $\theta = 0.09$, $\kappa = 2$, $\nu = 0.2$, $\rho = -0.2$ sowie Anfangswerten $S_0 = 100$ und $Y_0 = 0.5$ durch. Zum Vergleich wird die geometrischen Brownschen Bewegung (6.1) mit μ und S_0 wie oben sowie $\sigma = \sqrt{\theta} = 0.3$ betrachtet. Es wird das Euler-Maruyama-Verfahren mit Schrittweite $\Delta t = 0.005$ eingesetzt. Abbildung 20 zeigt zu den beiden Modellen jeweils drei Pfade und den Mittelwert aus 10 000 Pfaden im Intervall $[0, T] \in [0, 5]$. Wir erkennen, wie sich die stochastische Volatilität im Heston-Modell dem Mittelwert θ annähert.

Wir berechnen mit einer Monte-Carlo Simulation zu 10^6 Pfaden den fairen Preis einer Europäischen Call-Option mit den obigen Parametern und dem Ausübungspreis $E = 90$. Dabei wird $r = \mu$ gesetzt. Im Modell mit konstanter Volatilität folgt aus der Black-Scholes Formel $V_0 = 40.21$, während das Heston-Modell $\check{V}_0 = 39.77$ liefert. Der Unterschied dieser Werte beträgt etwa 1.1%.

Sprung-Modelle

Wir beziehen zufällige Sprünge in das Aktienkursmodell ein. Dazu werden folgende Wahrscheinlichkeitsverteilungen bzw. Zufallsprozesse benötigt. Eine Exponentialverteilung mit Parameter $\lambda > 0$ ist gekennzeichnet durch die Wahrscheinlichkeitsdichte

$$\rho_{\text{exp}}(x) = \begin{cases} \lambda e^{-\lambda x} & \text{für } x \geq 0 \\ 0 & \text{für } x < 0. \end{cases}$$

Der Erwartungswert zu dieser Verteilung ist $\mu = \frac{1}{\lambda}$.

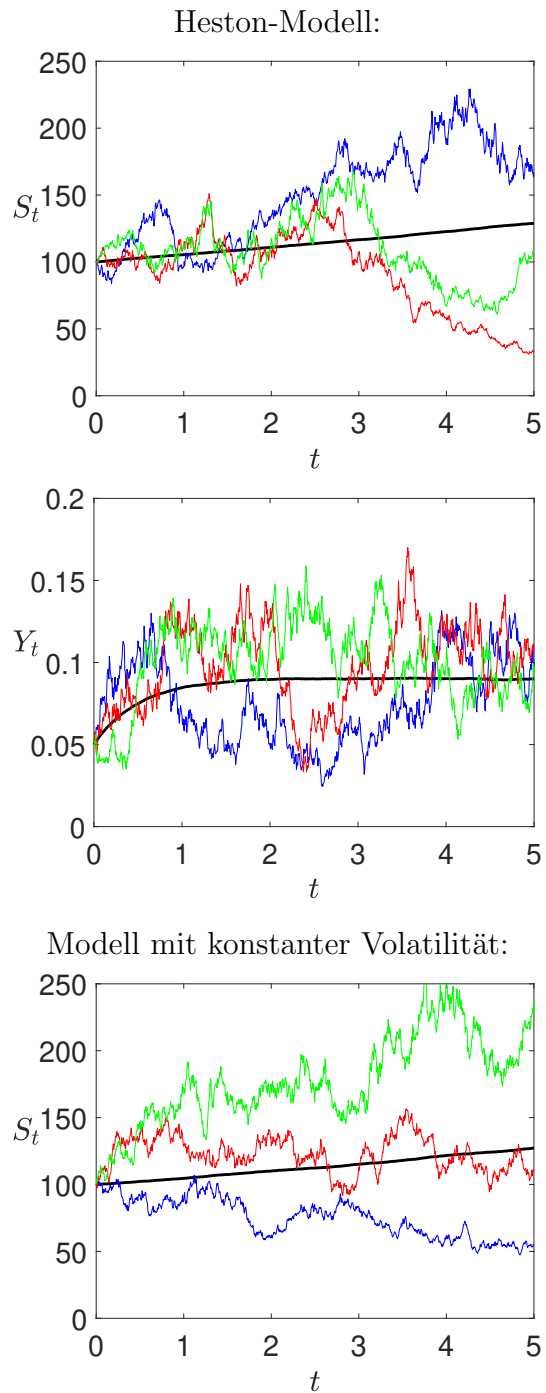


Abbildung 20: Simulationen der Modelle zum Aktienkurs mit Beispielpfaden und Mittelwert aus 10 000 Pfaden (schwarze Linie).

Definition 6.2 Sei $(\tau_j)_{j \in \mathbb{N}}$ eine Folge aus unabhängigen und identisch verteilten Zufallsvariablen mit Exponentialverteilung zu Parameter λ . Die zugehörigen Sprungzeiten sind die Zufallsvariablen

$$T_k = \sum_{i=1}^k \tau_i \quad \text{für } k = 1, 2, \dots \quad (6.18)$$

Ein Poisson Prozess mit Intensität λ ist ein Zufallsprozess $(N_t)_{t \geq 0}$ mit

$$N_t = \begin{cases} 0 & \text{falls } t < T_1 \\ k & \text{falls } t \in [T_k, T_{k+1}). \end{cases}$$

Ein Poisson Prozess $(N_t)_{t \geq 0}$ ist ein Spezialfall eines *Zähl-Prozesses*.

Eigenschaften eines Poisson Prozesses sind:

- i) Für $t \geq 0$ ist fast sicher N_t endlich, insbesondere $N_t \in \mathbb{N}_0$.
- ii) Für fast alle $\omega \in \Omega$ ist der Pfad $t \mapsto N_t(\omega)$ stückweise konstant mit Sprüngen der Höhe 1.
- iii) Für $t \geq 0$ ist N_t Poisson-verteilt mit Parameter λt , d.h.

$$P[N_t = k] = e^{-\lambda t} \frac{(\lambda t)^k}{k!} \quad \text{für } k = 0, 1, 2, \dots$$

- iv) Je endlich viele Inkremente $N_{t_j} - N_{s_j}$ sind unabhängig voneinander für $s_j < t_j$ und $t_j \leq s_{j+1}$.
- v) Die Inkremente sind stationär, d.h. die Verteilung von $N_t - N_s$ entspricht der Verteilung von N_{t-s} für $t > s$.

Die Eigenschaften (iv) und (v) gelten auch für einen Wiener Prozess. Aus Eigenschaft (iii) folgt $N_0 = 0$ fast sicher (setze $t = 0$ und $k = 0$).

Wir erweitern das Aktienkursmodell aus der geometrischen Brownsche Bewegung (6.2) zu

$$S_t = S_0 \exp \left(\left(\mu - \frac{1}{2} \sigma^2 \right) + \sigma W_t + \sum_{i=1}^{N_t} U_i \right) \quad (6.19)$$

für $t \geq 0$ mit einem Poisson Prozess $(N_t)_{t \geq 0}$. Darin bezeichnet $(U_i)_{i \in \mathbb{N}}$ eine Folge aus unabhängigen und identisch verteilten Zufallsvariablen, welche die zufälligen Sprunghöhen ergeben. Üblicherweise werden die Zufallsvariablen $(U_i)_{i \in \mathbb{N}}$ sowie

die Zufallsprozesse $(W_t)_{t \geq 0}$ und $(N_t)_{t \geq 0}$ als voneinander unabhängig angenommen. Im Sprung-Modell von Merton wird eine Normalverteilung $U_i \sim N(\mu_U, \sigma_U^2)$ verwendet. Im Sprung-Modell von Kou dagegen eine Laplace-Verteilung (doppelte Exponentialverteilung).

Identisch zu (6.19) ist die Darstellung

$$S_t = S_0 e^{(\mu - \frac{1}{2}\sigma^2) + \sigma W_t} \prod_{i=1}^{N_t} e^{U_i}.$$

Sind die Zufallsvariablen U_i normalverteilt, dann besitzen die Zufallsvariablen e^{U_i} eine Log-Normalverteilung. Desweiteren definieren wir den Zufallsprozess $(U_t)_{t \geq 0}$ durch

$$U_t = \sum_{i=1}^{N_t} U_i \chi_{[T_i, \infty)}(t) \quad (6.20)$$

mit der charakteristischen Funktion $\chi_{[T_i, \infty)}$, siehe (3.14), unter Verwendung der Sprungzeiten (6.18). Der Fall $N_t = 0$ ist als leere Summe zu interpretieren, d.h. der Wert der Summe ist null. Die Pfade des Zufallsprozesses (6.20) sind fast sicher stückweise konstant sowie rechtsseitig stetig. Der Zufallsprozess (6.19) kann geschrieben werden als

$$S_t = S_0 e^{(\mu - \frac{1}{2}\sigma^2) + \sigma W_t + U_t}. \quad (6.21)$$

Für eine numerische Simulation eines Pfads zum Zufallsprozess (6.21), d.h. für ein festes $\omega \in \Omega$, in einem Zeitintervall $[0, T]$ benötigen wir Realisierungen des Zufallsprozesses (6.20). Wir generieren Pseudo-Zufallszahlen ξ_1, \dots, ξ_K bezüglich der Exponentialverteilung mit Parameter λ und bilden mit ihnen die Sprungzeiten $\hat{T}_1, \hat{T}_2, \dots, \hat{T}_K$ aus (6.18). Dabei werden so viele Zufallszahlen erzeugt, dass $\hat{T}_{K-1} \leq T$ und $\hat{T}_K > T$ gilt. Für $t \in [0, T]$ folgt $0 \leq N_t(\omega) \leq K - 1$. Wir generieren jetzt Pseudo-Zufallszahlen u_1, u_2, \dots, u_{K-1} gemäß der Verteilung der Zufallsvariablen U_i . Dadurch können wir (6.20) auswerten.

Als Beispiel betrachten wir die Parameterwahl $\mu = 0.05$, $\sigma = 0.25$, $\lambda = 1$. Für die unabhängigen Zufallsvariablen U_i werden Normalverteilungen mit Erwartungswert $\mu_U = -0.05$ und Standardabweichung $\sigma_U = 0.1$ verwendet. Der Endzeitpunkt ist $T = 4$ und der Anfangswert $S_0 = 100$. Wir diskretisieren mit Schrittweite $\Delta t = \frac{1}{250}$. Abbildung 21 zeigt Näherungen von Pfaden zum Sprungprozess (6.19) bzw. (6.21).

Als Verallgemeinerung betrachten wir den Aktienkurs als die Lösung einer stochastischen Integralgleichung

$$S_t = S_0 + \int_0^t u_s ds + \int_0^t v_s dW_s + \int_0^t q_s dN_s. \quad (6.22)$$

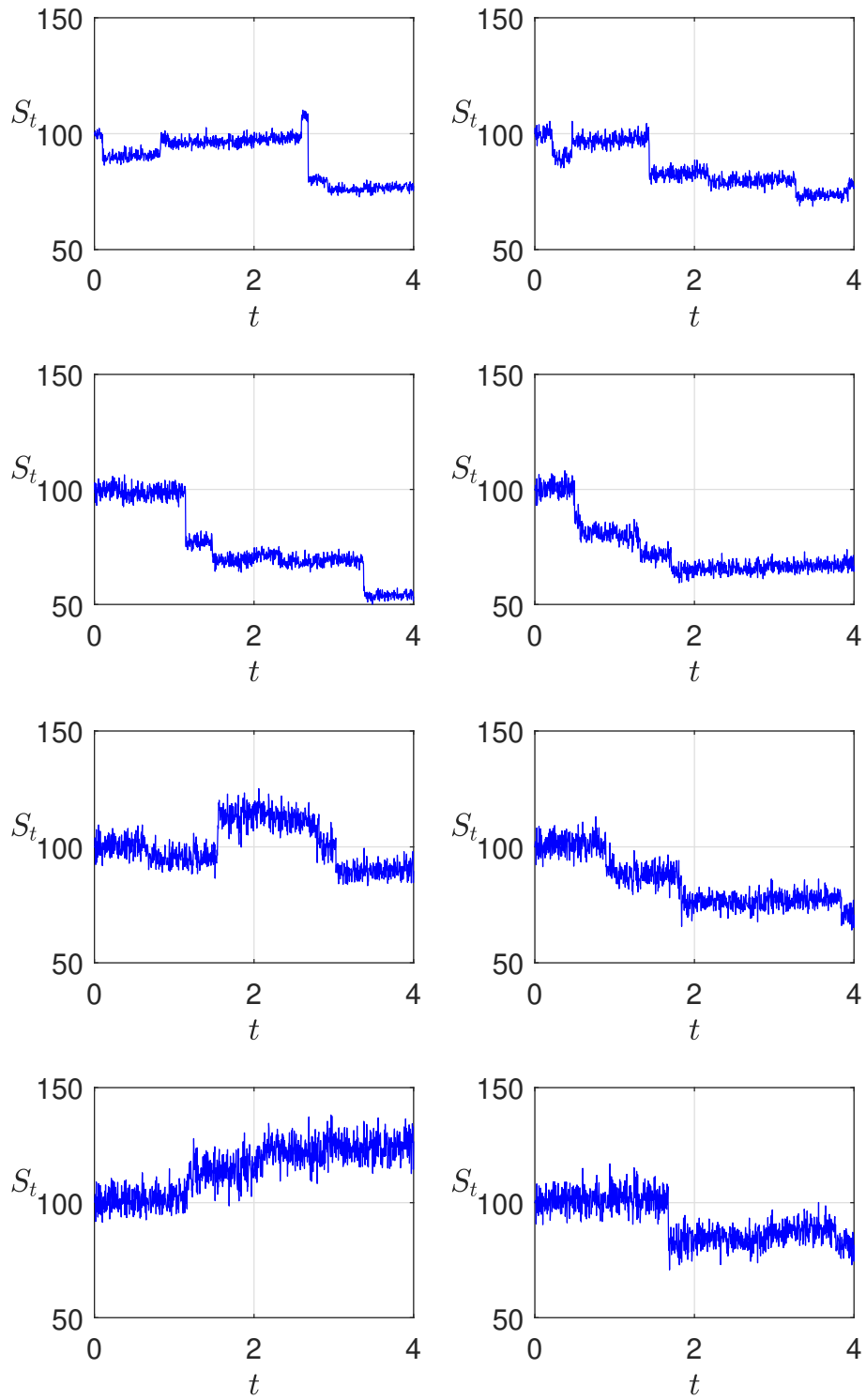


Abbildung 21: Pfade eines Sprung-Prozesses konstruiert aus einer geometrischen Brownschen Bewegung und einem Poisson Prozess.

Es liegen ein deterministischer Anteil und zwei stochastische Anteile vor. Das stochastische Integral bezüglich des Poisson Prozesses ist noch zu definieren, was wir hier nicht ausführen. Im Heston-Modell (6.16) würde in (6.22) $u_t = \mu S_t$ und $v_t = \sqrt{Y_t} S_t$ vorliegen, wobei Y_t eine weitere Itô-Integralgleichung erfüllt.

Wir gehen darauf ein, wie in einer numerischen Methode mit Zeitpunkten $t_i = i\Delta t$ für $i = 0, 1, \dots, N$ und Zeitschrittweite $\Delta t = \frac{T}{N}$ der Sprung-Anteil einbezogen werden kann. Es ist $N_0 = 0$ fast sicher. Bei $t = \Delta t$ folgt mit Eigenschaft (iii) eines Poisson Prozesses

$$\begin{aligned} P[N_{\Delta t} = 0] &= e^{-\lambda\Delta t} = 1 - \lambda\Delta t + \mathcal{O}((\Delta t)^2) \\ P[N_{\Delta t} = 1] &= e^{-\lambda\Delta t} \lambda\Delta t = \lambda\Delta t + \mathcal{O}((\Delta t)^2) \\ P[N_{\Delta t} > 1] &= 1 - (1 - \lambda\Delta t + \mathcal{O}((\Delta t)^2)) - (\lambda\Delta t + \mathcal{O}((\Delta t)^2)) = \mathcal{O}((\Delta t)^2). \end{aligned}$$

Wir setzen eine kleine Schrittweite Δt voraus, insbesondere $\Delta t < \frac{1}{\lambda}$. Daher können wir den Fall von mehr als einem Sprung vernachlässigen. Desweiteren sind die Wahrscheinlichkeiten für Sprünge im Intervall $[t_i, t_{i+1}]$ identisch wegen der Eigenschaften (iv) und (v) eines Poisson Prozesses. Daher setzen wir als Näherung an

$$\begin{aligned} \text{kein Sprung:} \quad & P[N_{t_{i+1}} = N_{t_i}] = 1 - \lambda\Delta t \\ \text{ein Sprung:} \quad & P[N_{t_{i+1}} = N_{t_i} + 1] = \lambda\Delta t. \end{aligned}$$

Diese Verteilung entspricht einem Bernoulli-Experiment mit Trefferwahrscheinlichkeit $p = \lambda\Delta t$. Wir bilden jeweils eine Pseudo-Zufallszahl u_i gemäß einer Gleichverteilung auf dem Intervall $[0, 1]$. Ist $u_i \leq p$, so findet der Sprung statt und der Wert $S_{t_{i+1}}$ wird entsprechend modifiziert. Ist $u_i > p$, dann liegt kein Sprung vor und die Näherung $S_{t_{i+1}}$ wird wie üblich festgesetzt.

6.4 Optionen auf mehrere Aktien

Jetzt werden mehrere Basiswerte S_1, S_2, \dots, S_n einbezogen, welche jeweils den Preis für eine Aktie, für eine Fremdwährung oder für eine bestimmte Menge eines Guts (z.B. Gold, Rohöl) darstellen. Zu einer Option hängt der Preis hier von diesen Basiswerten ab.

Mehrdimensionale Normalverteilung

Gegeben seien k standard-normalverteilte Zufallsvariablen $X = (X_1, \dots, X_k)^\top$. Sind diese Zufallsvariablen unabhängig, dann lautet ihre Dichtefunktion

$$f : \mathbb{R}^k \rightarrow \mathbb{R}, \quad f(x) = \frac{1}{\sqrt{2\pi}^k} \exp\left(-\frac{x^\top x}{2}\right). \quad (6.23)$$

Wir führen mit $Y = (Y_1, \dots, Y_k)^\top$ die lineare Transformation

$$Y = AX + b, \quad A \in \mathbb{R}^{k \times k}, \quad \det(A) \neq 0, \quad b \in \mathbb{R}^k \quad (6.24)$$

durch. Die Zufallsvariable Y ist dann k -dimensional normalverteilt. Offensichtlich gilt $\mathbb{E}(Y) = b$. Mit $x = A^{-1}(y - b)$ transformiert sich ein Integral über obige Dichtefunktion wie folgt

$$\begin{aligned} & \frac{1}{\sqrt{2\pi}^k} \int \exp\left(-\frac{x^\top x}{2}\right) dx \\ &= \frac{1}{\sqrt{2\pi}^k} \int \exp\left(-\frac{1}{2}(A^{-1}(y-b))^\top (A^{-1}(y-b))\right) \det(A^{-1}) dy \quad (6.25) \\ &= \frac{1}{\sqrt{2\pi}^k \det(A)} \int \exp\left(-\frac{1}{2}(y-b)^\top (AA^\top)^{-1}(y-b)\right) dy. \end{aligned}$$

Die Matrix AA^\top ist symmetrisch und (wegen $\det(A) \neq 0$) positiv definit. Desweiteren gilt $\det(AA^\top) = \det(A)^2$. Die Matrix $AA^\top = (\sigma_{ij})$ enthält die Kovarianzen der Zufallsvariablen, denn mit $\mathbb{E}(X_n X_m) = \delta_{nm}$ folgt

$$\begin{aligned} \text{Cov}(Y_i, Y_j) &= \mathbb{E}((Y_i - \mathbb{E}(Y_i))(Y_j - \mathbb{E}(Y_j))) \\ &= \mathbb{E}\left(\left(\sum_{n=1}^k a_{in} X_n\right) \left(\sum_{m=1}^k a_{jm} X_m\right)\right) \quad (6.26) \\ &= \sum_{n=1}^k \sum_{m=1}^n a_{in} a_{jm} \mathbb{E}(X_n X_m) = \sum_{m=1}^k a_{im} a_{jm}. \end{aligned}$$

Umgekehrt kann eine mehrdimensionale normalverteilte Zufallsvariable durch ihre Dichtefunktion definiert werden.

Definition 6.3 Eine Zufallsvariable $Y : \Omega \rightarrow \mathbb{R}^k$ heißt normalverteilt, wenn eine korrespondierende Dichtefunktion $g : \mathbb{R}^k \rightarrow \mathbb{R}$ der Form

$$g(y) = \frac{1}{\sqrt{2\pi}^k \sqrt{\det(\Sigma)}} \exp\left(-\frac{1}{2}(y-b)^\top \Sigma^{-1}(y-b)\right) dy \quad (6.27)$$

mit einer symmetrischen positiv definiten Matrix Σ existiert.

Es folgt dann, dass $\mathbb{E}(Y) = b$ gilt und die Matrix $\Sigma = (\sigma_{ij})$ die Kovarianzen von Y wie oben enthält. Wir schreiben dann $Y \sim N(b, \Sigma)$. Aus der Kovarianzmatrix Σ folgt auch die Korrelationsmatrix $R = (\rho_{ij}) \in \mathbb{R}^{k \times k}$ mit den Einträgen

$$\rho_{ij} = \frac{\sigma_{ij}}{\sqrt{\sigma_{ii}\sigma_{jj}}}$$

für $i, j = 1, \dots, k$. Diese Matrix ist wieder symmetrisch und ihre Diagonale besteht aus Einsen.

Aus der Matrix Σ können wir auch eine Matrix A zur linearen Transformation (6.24) aus standard-normalverteilten Zufallsvariablen konstruieren. Die Matrix A ist dabei nicht eindeutig, da jede symmetrische Zerlegung $\Sigma = AA^\top$ verwendet werden kann. Gebräuchliche symmetrische Zerlegungen sind:

i) *Cholesky-Zerlegung*:

$\Sigma = LL^\top$ mit einer unteren Dreiecksmatrix L , siehe [16].

ii) *Eigenwert-Zerlegung*:

$\Sigma = TDT^\top$ mit Diagonalmatrix $D = \text{diag}(\lambda_1, \dots, \lambda_k)$ bestehend aus den positiven Eigenwerten und orthogonaler Matrix T . Es folgt mit $D = D^{\frac{1}{2}}D^{\frac{1}{2}}$ hier $\Sigma = (TD^{\frac{1}{2}})(TD^{\frac{1}{2}})^\top$, d.h. $A = TD^{\frac{1}{2}}$.

Die numerische Berechnung der Cholesky-Zerlegung ist weniger aufwändig als die Eigenwert-Zerlegung.

Modell für mehrere Aktienkurse

Zu Basiswerten S_1, S_2, \dots, S_n betrachten wir das lineare System aus Itô-Integralgleichungen

$$S_i(t) = S_i(0) + \int_0^t \mu_i S_i(s) ds + \int_0^t \sigma_i S_i(s) d\widehat{W}_i(s) \quad (6.28)$$

für $i = 1, \dots, n$. Es liegen Driftparameter $\mu_i > 0$ und Volatilitäten $\sigma_i > 0$ vor. In einem mehrdimensionalen Wiener Prozess $W = (W_1, \dots, W_n)^\top$ sind die Komponenten nach Definition 2.4 unabhängig. Bei (6.28) tritt ein korrelierter Wiener Prozess $\widehat{W} = (\widehat{W}_1, \dots, \widehat{W}_n)^\top$ auf. Dieser Wiener Prozess modelliert Abhängigkeiten zwischen den Basiswerten. Die Kursänderung eines Basiswerts kann den Kurs eines anderen Basiswerts beeinflussen. Die Korrelationen zwischen den (eindimensionalen) Wiener Prozessen übertragen sich direkt auf Korrelationen zwischen den zugehörigen Inkrementen.

Sei $t \geq 0$. Für die Zufallsvariablen $\widehat{W}_i(t)$ gelte wie üblich $\mathbb{E}(\widehat{W}_i(t)) = 0$ sowie $\text{Var}(\widehat{W}_i(t)) = t$. Wir definieren Korrelationen zwischen den Wiener Prozessen durch

$$\text{Cor}(\widehat{W}_i(t), \widehat{W}_j(t)) = \rho_{ij}.$$

Wegen der Symmetrie müssen nur die Parameter ρ_{ij} für $j < i$ und $i = 1, \dots, n$ festgelegt werden. Es gilt stets $\rho_{ii} = 1$ für $i = 1, \dots, n$. Wir fassen die Korrelationen in der Matrix $R = (\rho_{ij}) \in \mathbb{R}^{n \times n}$ zusammen.

Aus der allgemein gültigen Formel

$$\text{Cor} \left(\widehat{W}_i(t), \widehat{W}_j(t) \right) = \frac{\text{Cov} \left(\widehat{W}_i(t), \widehat{W}_j(t) \right)}{\sqrt{\text{Var} \left(\widehat{W}_i(t) \right) \text{Var} \left(\widehat{W}_j(t) \right)}}$$

folgt hier

$$\text{Cov} \left(\widehat{W}_i(t), \widehat{W}_j(t) \right) = \rho_{ij}t.$$

Beispiel: Im zweidimensionalen Fall ($n = 2$) lautet die Korrelationsmatrix

$$R = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$$

mit nur einem Freiheitsgrad $\rho \in (-1, 1)$. Die Cholesky-Zerlegung $R = LL^\top$ ist gegeben durch

$$L = \begin{pmatrix} 1 & 0 \\ \rho & \sqrt{1 - \rho^2} \end{pmatrix},$$

wie sofort durch Matrizenmultiplikation nachgeprüft werden kann.

Das System (6.28) können wir in kompakter Form schreiben als

$$S(t) = S(0) + \int_0^t u(S(t)) dt + \int_0^t v(S(t)) d\widehat{W}(s)$$

mit $S = (S_1, \dots, S_n)^\top$. Darin sind die Funktionen $u : \mathbb{R}^n \rightarrow \mathbb{R}^n$ und $v : \mathbb{R}^n \rightarrow \mathbb{R}^{n \times n}$ enthalten. Die matrix-wertige Funktion v ist hier diagonal, d.h. $v_{ij} = 0$ für $i \neq j$. Sei $R = AA^\top$ mit einer konstanten Matrix $A \in \mathbb{R}^{n \times n}$. Die Transformation $\widehat{W}(t) = \sqrt{t}AW(t)$ impliziert für die Differentiale dann $d\widehat{W} = A dW$. Es folgt das äquivalente System

$$S(t) = S(0) + \int_0^t u(S(t)) dt + \int_0^t v(S(t))A dW(s).$$

Jetzt liegt die übliche Gestalt mit einem unkorrelierten Wiener Prozess vor. Wir können somit die numerischen Verfahren aus Kapitel 5 einsetzen. Im stochastischen Anteil ist die matrix-wertige Funktion $\tilde{v} = vA$ jetzt im allgemeinen voll besetzt. Bei Verwendung der Cholesky-Zerlegung ist \tilde{v} hier eine untere Dreiecksmatrix.

Optionstypen

Wir betrachten übliche Optionstypen für eine Menge S_1, S_2, \dots, S_n an Basiswerten. Für die Auszahlungsfunktion gilt jeweils eine Abhängigkeit $\Lambda(S_1, S_2, \dots, S_n)$.

Tabelle 1: Korrelationen zwischen den Umtauschkursen der Währungen Euro (E), US Dollar (D), Britisches Pfund (P) und Japanischer Yen (Y) aus dem Jahr 2001.

	P/D	D/Y	P/Y	E/D	E/P	E/Y
P/D	1	-0.47	0.42	0.71	-0.19	0.27
D/Y	-0.47	1	0.6	-0.53	-0.18	0.45
P/Y	0.42	0.6	1	0.1	-0.36	0.71
E/D	0.71	-0.53	0.1	1	0.55	0.52
E/P	-0.19	-0.18	-0.36	0.55	1	0.4
E/Y	0.27	0.45	0.71	0.52	0.4	1

Eine Linearkombination der Basiswerte lautet

$$\bar{S}_t = \sum_{j=1}^n \alpha_j S_j(t) \quad (6.29)$$

mit reellen Koeffizienten $\alpha_j > 0$ für $j = 1, \dots, n$. Bei Aktien ist die Wahl $\alpha_j \in \mathbb{N}$ sinnvoll. Eine solche Kombination wird auch als „Korb“ mit Basiswerten bezeichnet.

Eine *Basket-Option* vom Call-Typ gibt die Möglichkeit, diese Kombination von Basiswerten an einem Endzeitpunkt T zum Ausübungspreis E zu kaufen. Die Option vom Put-Typ ist analog definiert. Es ergeben sich die Auszahlungsfunktionen

$$\Lambda(S_1, S_2, \dots, S_n) = \max\{\bar{S} - E, 0\} \quad (\text{Call})$$

$$\Lambda(S_1, S_2, \dots, S_n) = \max\{E - \bar{S}, 0\} \quad (\text{Put})$$

mit (6.29) ausgewertet bei $t = T$, d.h. $S_j(T)$ für $j = 1, \dots, n$ und \bar{S}_T .

Als Beispiel betrachten wir die Umtauschkurse zwischen den vier Währungen: Euro, US Dollar, Britisches Pfund und Japanischer Yen. Tabelle 1 zeigt die Korrelationen dieser Umtauschkurse, welche aus Marktdaten im Jahr 2001 bestimmt wurden, siehe [7]. Beziehen wir uns als Referenz auf die Euro-Währung, dann lautet die Korrelationsmatrix zwischen den drei anderen Währungen

$$R = \begin{pmatrix} 1 & 0.55 & 0.52 \\ 0.55 & 1 & 0.4 \\ 0.52 & 0.4 & 1 \end{pmatrix}.$$

Eine entsprechende Basket-Option bezieht sich auf eine Kombination aus den drei Währungen Dollar, Pfund und Yen und der Ausübungspreis E ist in Euro festgelegt.

Eine weitere Klasse von Optionen auf mehrere Basiswerte sind die *Rainbow-Optionen*. Die Basiswerte aus einer Menge $\{S_1, S_2, \dots, S_n\}$ werden auch als Farben des „Regenbogens“ bezeichnet. Diese Optionen sind durch ihre Auszahlungsfunktion ausgewertet am Endzeitpunkt T und den Ausübungspreis E definiert.

Rainbow-Optionen vom Call-Typ:

i) Call auf Maximum

$$\Lambda(S_1, S_2, \dots, S_n) = \max\{\max\{S_1, S_2, \dots, S_n\} - E, 0\}$$

ii) Call auf Minimum

$$\Lambda(S_1, S_2, \dots, S_n) = \max\{\min\{S_1, S_2, \dots, S_n\} - E, 0\}$$

iii) Bester Basiswert oder Bargeld

$$\Lambda(S_1, S_2, \dots, S_n) = \max\{S_1, S_2, \dots, S_n, E\}$$

Rainbow-Optionen vom Put-Typ:

i) Put auf Maximum

$$\Lambda(S_1, S_2, \dots, S_n) = \max\{E - \max\{S_1, S_2, \dots, S_n\}, 0\}$$

ii) Put auf Minimum

$$\Lambda(S_1, S_2, \dots, S_n) = \max\{E - \min\{S_1, S_2, \dots, S_n\}, 0\}$$

Es existieren noch weitere Optionen vom Rainbow-Typ.

7 Stochastische Resonanz

Wir befassen uns mit stochastischer Periodizität und stochastischer Resonanz. Als Anwendungsbeispiel wird ein mathematisches Klimamodell diskutiert.

7.1 Klimamodell

In diesem Abschnitt betrachten wir ein Modell zur qualitativen Erklärung des periodischen Wechsels zwischen Eiszeit und Warmzeit in den letzten 700.000 Jahren. Dabei zeigt sich, dass eine ausschließlich deterministische Modellbildung nicht ausreicht, um die realen Effekte zu reproduzieren. Eine stochastische Differentialgleichung ist zur Beschreibung des Problems aufzustellen. Für eine allgemeine Einführung in *Klimamodelle* siehe [15] und die Literaturverweise dort. Tabelle 2 zeigt die benötigten Variablen im resultierenden mathematischen Modell.

Die Konzentration des Sauerstoff-Isotops O-18 in Ablagerungen von Gestein oder Eis ergibt eine grobe Schätzung der globalen Temperatur auf der Erde in der Vergangenheit. Es zeigt sich, dass in den letzten 700.000 Jahren ein periodischer Wechsel zwischen Eiszeit und Warmzeit stattgefunden hat. Die Eiszeit besitzt eine globale Temperatur von ca. 7°C , während die Warmzeit durch ca. 15°C gekennzeichnet ist. Die Periode dieses Zyklus beträgt ca. 100.000 Jahre. Es ergibt sich die Frage, woher der Wert dieser Periode resultiert.

Untersuchungen aus der Astrophysik deuten darauf hin, dass die jährliche Umlaufbahn der Erde als Ellipse nicht konstant ist. Durch Wechselwirkung mit anderen Planeten (vor allem Jupiter) ergeben sich geringe Schwankungen im mittleren jährlichen Abstand von der Sonne. Diese Konstellation zwischen den Planeten ändert sich in einem periodischen Zyklus von ca. 100.000 Jahren, was die beobachtete Periode erklärt. Ein geringerer Abstand der Erde zur Sonne bewirkt eine erhöhte Wärmeeinstrahlung im Gegensatz zu einem weiteren Abstand. Jedoch sind die Unterschiede im Abstand während des Zyklus so gering, dass eine Differenz von ca. 8 Grad in der globalen Temperatur nicht reproduziert werden kann, sondern nur weniger als 1 Grad. Es ist also eine alternative Erklärung notwendig.

Die von der Sonne einstrahlende Energie wird in der Erdatmosphäre nicht vollständig absorbiert, da ein Teil zurück reflektiert wird. Die Albedo bezeichnet das Rückstrahlvermögen. Die planetare Albedo auf der Erde ist jedoch selbst temperaturabhängig. Befindet sich die Erde in der Eiszeit, so sind weite Teile mit Schnee oder Eis bedeckt, wodurch eine maximale Reflexion des Sonnenlichts eintritt. In der Warmzeit dagegen wird einem Treibhauseffekt entsprechend ein maximaler Anteil an Sonnenenergie absorbiert, d.h. die Reflexion nimmt einen minimalen Wert an. Wir modellieren diese Temperaturabhängigkeit in der Albedo

Tabelle 2: Variablen in mathematischen Klimamodell.

Variable	Einheit	Bedeutung
T	[K]	globale Temperatur
t	[s]	Zeit
c	[Wsm ⁻² K ⁻¹]	spezifische Wärmekapazität der Atmosphäre
a	—	planetare Albedo
σ	[Wm ⁻² K ⁻⁴]	Boltzmann-Konstante
Q	[Wm ⁻²]	Energieeinstrahlung durch die Sonne

durch die einfache stückweise lineare Funktion

$$a(T) = \begin{cases} a_{\max} & \text{für } T \leq T_1 \\ \left(1 - \frac{T-T_1}{T_2-T_1}\right)a_{\max} + \frac{T-T_1}{T_2-T_1}a_{\min} & \text{für } T_1 < T < T_2 \\ a_{\min} & \text{für } T \geq T_2 \end{cases} \quad (7.1)$$

mit Konstanten $T_1 < T_2$ und $0 < a_{\min} < a_{\max} < 1$.

Die Änderung der globalen Temperatur mit der Zeit ergibt sich aus der Differenz zwischen Erwärmung und Abkühlung. Die Erwärmung entsteht durch die Absorption von Sonnenenergie in der Erdatmosphäre. Die Abkühlung resultiert durch Abgabe von Wärme aus der Erdatmosphäre an das Weltall. Insgesamt erhalten wir als mathematisches Modell die gewöhnliche Differentialgleichung

$$c \frac{dT}{dt} = Q(t)(1 - a(T)) - \sigma T^4 \quad (7.2)$$

für die zeitabhängige Temperatur T . Im folgenden beschreibt $f(T) = Q(1 - a(T))$ die Erwärmung und $g(T) = \sigma T^4$ die Abkühlung. Die Funktion $Q(t)$ kann entweder konstant vorausgesetzt werden oder eine entsprechende zeitabhängige Funktion ist zu spezifizieren.

Wir interessieren uns für stationäre Lösungen der Differentialgleichung (7.2). Dazu nehmen wir $Q(t) \equiv Q_0$ konstant an. Aufwärmung und Abkühlung heben sich gegenseitig auf, wenn gilt

$$\frac{dT}{dt} = 0 \quad \Leftrightarrow \quad Q_0(1 - a(T)) = \sigma T^4.$$

Abbildung 22 (links) verdeutlicht die Funktionen f und g für Aufwärmung bzw. Abkühlung bei entsprechenden realistischen Parametern. Wir erkennen drei stationäre Temperaturen T_E, T_I, T_W mit den Relationen

$$T_E < T_1 < T_I < T_2 < T_W.$$

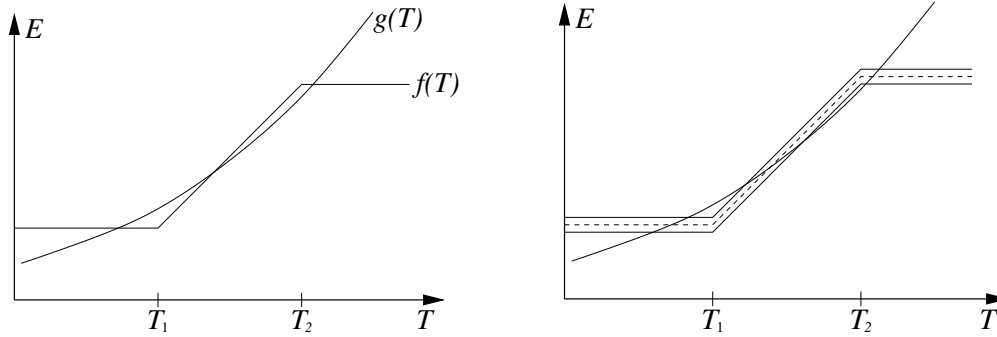


Abbildung 22: Temperaturabhängiger Energieaustausch.

Die lokale Stabilität dieser stationären Lösungen entscheidet sich am Vorzeichen der Ableitung der rechten Seite von (7.2). Es gilt

$$h(T) = f'(T) - g'(T) = -Q_0 a'(T) - 4\sigma T^3.$$

Wir erhalten $a'(T_E) = a'(T_W) = 0$ und $a'(T_I) < 0$. Es folgt $h(T_E), h(T_W) < 0$ und bei entsprechenden Parametern $h(T_I) > 0$. Somit sind T_E und T_W stabile Lösungen, während die stationäre Lösung T_I instabil ist. Es sind offensichtlich T_E und T_W die Zustände bezüglich der Eiszeit bzw. Warmzeit.

Es ist noch zu erklären, warum es zu einem periodischen Wechsel zwischen beiden Zuständen kommt. Liegt eine zeitabhängige Funktion $Q(t)$ vor, dann betrachten wir die stationären Lösungen jeweils für festes t . Wir verwenden den Ansatz

$$Q(t) = Q_0 \left(1 + \alpha \cos\left(\frac{2\pi}{P}t\right)\right) \quad (7.3)$$

mit Konstanten $Q_0 > 0$ und $0 < \alpha < 1$. Durch die physikalischen Ergebnisse motiviert setzen wir als Periode stets $P = 10^5$ Jahre. Bei starken Schwankungen, d.h. relativ hohen Werten α sind die extremen Situationen in Abbildung 23 veranschaulicht. Es kommt jeweils zum Verlust von einer der stabilen stationären Lösungen. Folglich wird sich die globale Temperatur zum noch bestehenden stabilen Gleichgewicht hin entwickeln und bei Erreichen dort zunächst verbleiben.

Jedoch sind die real auftretenden Schwankungen sehr klein, nämlich im Bereich $\alpha \in [0.001, 0.01]$. Daher existieren tatsächlich zu allen Zeiten jeweils die drei stationären Lösungen T_E, T_I, T_W , wie in Abbildung 22 (rechts) dargestellt ist. Abbildung 24 zeigt uns zwei Lösungen von Anfangswertproblemen zu (7.2) mit (7.3) und $\alpha = 0.005$ über einen langen Zeitbereich. Wir erkennen leichte Schwankungen um den Temperaturwert zur Eiszeit bzw. Warmzeit mit der Periode P . Ein Übergang zwischen beiden Zuständen wird jedoch nicht beobachtet. Hierfür ist eine weitere Modifikation im Modell notwendig.

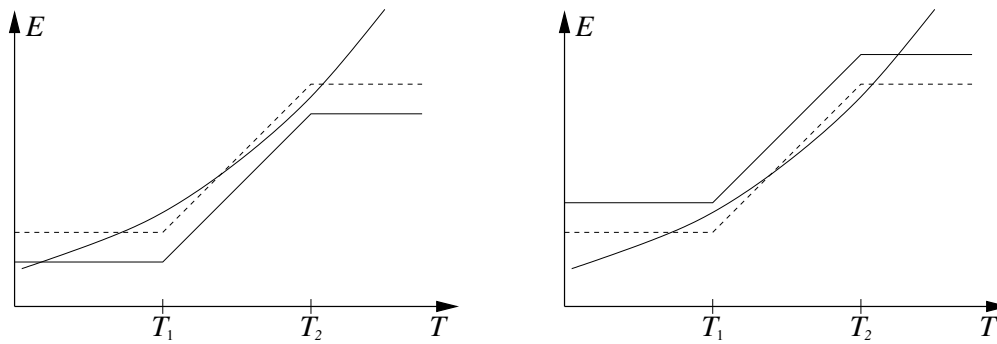


Abbildung 23: Temperaturabhängiger Energieaustausch bei unrealistisch hohen Schwankungen in der Albedo.

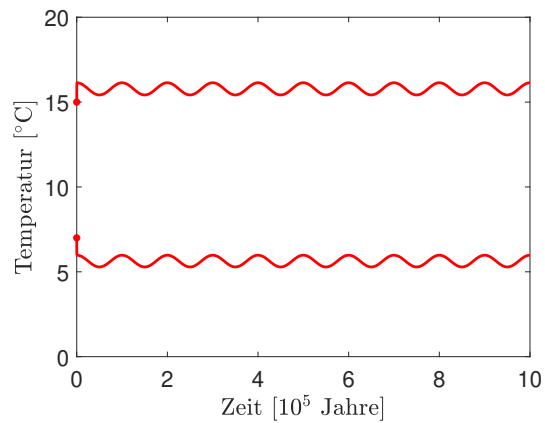


Abbildung 24: Zeitabhängige Lösungen zu verschiedenen Anfangswerten.

Eine Möglichkeit, die Stabilität der stationären Lösungen quantitativ zu untersuchen, besteht in der Einführung eines Potentials. Wir definieren bezüglich der rechten Seite von (7.2) das Potential

$$U(t, T) = - \int_{T_0}^T Q(t)(1 - a(\tau)) - \sigma\tau^4 d\tau \quad (7.4)$$

für jeweils feste Zeit $t \geq 0$. Abbildung 25 verdeutlicht die Potentiale (7.4) zu bestimmten Zeitpunkten. Lokale Minima bzw. Maxima des Potentials stellen stabile bzw. instabile stationäre Lösungen dar. Die Potentialdifferenz zwischen zwei Temperaturen T_a und T_b kann als zu verrichtende Arbeit beim Übergang zwischen den Zuständen interpretiert werden. Je nach Vorzeichen ist die Arbeit dem System zuzuführen bzw. wird vom System geleistet. Im Zyklus ist somit jeweils abwechselnd einer der beiden Gleichgewichte quantitativ stabiler. Im Modell (7.2) existiert jedoch kein Quellterm, der eine Arbeit für einen Wechsel zwischen den Zuständen hinzufügen könnte.

Nun erfolgt die entscheidende Modellierung mittels eines stochastischen Anteils. Die globale Temperatur hängt verständlicherweise nicht nur von der absorbierten Sonnenenergie und der Abkühlung ans Weltall ab. Eine Reihe weiterer Faktoren bewirken kleinere Schwankungen in der globalen Temperatur. Diese Effekte nehmen wir als zufallsabhängig an, nämlich über einen (eindimensionalen) Wiener Prozess als korrespondierendes Rauschen. Dementsprechend modellieren wir die globale Temperatur mittels einer stochastischen Differentialgleichung. Aus (7.2) entsteht das neue Modell

$$dT_t = \frac{1}{c} [Q(t)(1 - a(T_t)) - \sigma T_t^4] dt + \sqrt{\varepsilon} dW_t \quad (7.5)$$

mit einer geeigneten Konstanten $\varepsilon > 0$. Da die Funktion im stochastischen Anteil konstant ist, stimmen hier Itô-Integral und Stratonovich-Integral überein.

Die zufälligen Schwankungen können nun bewirken, dass eine stabile globale Temperatur verlassen wird und möglicherweise der jeweils andere stabile Zustand erreicht wird. Anhand Abbildung 25 erkennen wir, dass ein alternierender Wechsel mit der Periode P wahrscheinlich wird.

Der deterministische Teil der Differentialgleichung (7.5) wird mit weiteren Konstanten genauer festgelegt als, siehe [15],

$$dT_t = \frac{1}{4h\rho c} [Q(t)(1 - a(T_t)) - 4e\sigma T_t^4] dt + \sqrt{\varepsilon} dW_t, \quad (7.6)$$

worin $a(T)$ in (7.1) und $Q(t)$ in (7.3) gegeben ist. Die Bedeutung und Werte der Konstanten sind in Tabelle 3 aufgelistet.

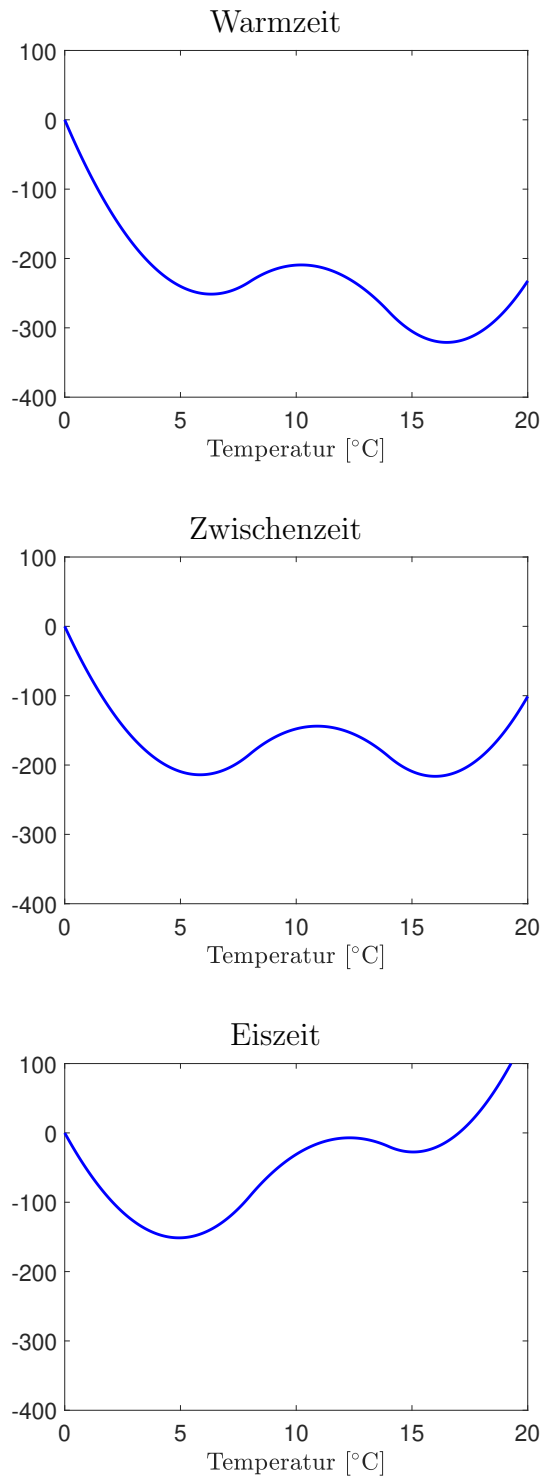


Abbildung 25: Potentiale $U(t, \cdot)$ zu Zeiten $t = 0$ (oben), $t = 0.2P$ (mitte) und $t = 0.5P$ (unten).

Tabelle 3: Konstanten in mathematischen Klimamodell im SI-System (Zeit: Sekunde s, Länge: Meter m, Temperatur: Kelvin K, etc.).

Konstante	Wert	Bedeutung
c	1000	spezifische Wärmekapazität der Luft
ρ	1200	Dichte von Luft
h	8300	vertikale Ausdehnung der Troposphäre
a_{\min}	0.25	minimale Reflexion
a_{\max}	0.35	maximale Reflexion
T_1	281	Temperaturschranke
T_2	287	Temperaturschranke
σ	$5.67 \cdot 10^{-8}$	Boltzmann-Konstante
e	0.65	Emissivität
Q_0	1367	Energieeinstrahlung der Sonne / Solarkonstante
P	$3 \cdot 10^{12}$	Periode

Wir testen das stochastische Modell (7.6) in einer numerischen Simulation. Das Euler-Maruyama-Verfahren liefert die korrespondierende Näherungslösung. Für die spezielle Gleichung (7.5) bzw. (7.6) besitzt das Euler-Maruyama-Verfahren die starke Konvergenzordnung 1. Wir setzen $\sqrt{\varepsilon} = 3.2 \cdot 10^{-5}$ sowie $\alpha = 0.005$ in (7.3). Als Anfangswert dient $T_0 = 288$. Abbildung 26 (links) zeigt das Simulationsergebnis für einen Pfad der globalen Temperatur. Zusätzlich ist die relative Aufwärmung dargestellt durch eine Cosinusfunktion mit Periode P . Wir erkennen einen periodischen Wechsel zwischen Eiszeit und Warmzeit entsprechend den Schwankungen in der Aufwärmung. Damit ist eine Erklärung des Klimawandels durch das mathematische Modell basierend auf einer stochastischen Differentialgleichung gefunden. Desweiteren enthält Abbildung 26 (rechts) den Mittelwert aus 10.000 Pfaden. Wir beobachten, dass dieser Mittelwert eine nahezu periodische Funktion darstellt.

In obiger Simulation wurde der Parameter ε so gewählt, dass Lösungen der stochastischen Differentialgleichung im wesentlichen periodisch sind. Bei kleinen Werten ε finden Übergänge zwischen den zwei stabilen Zuständen weniger häufig statt. Die Periode P ist dadurch nicht mehr feststellbar, sondern nur Wechsel nach (jeweils unterschiedlichen) ganzzahligen Vielfachen von P . Dagegen führen große Parameter ε zu einer Vielzahl von Übergängen, welche die Periodizität ebenfalls verschwinden lassen. Abbildung 27 demonstriert dieses Verhalten.

Wir können daher nach einem optimalen Wert ε fragen, zu dem sich Lösungen der stochastischen Differentialgleichung möglichst periodisch verhalten. Diese Aufgabenstellung führt auf die *stochastische Resonanz*. Bei der Resonanz von mecha-

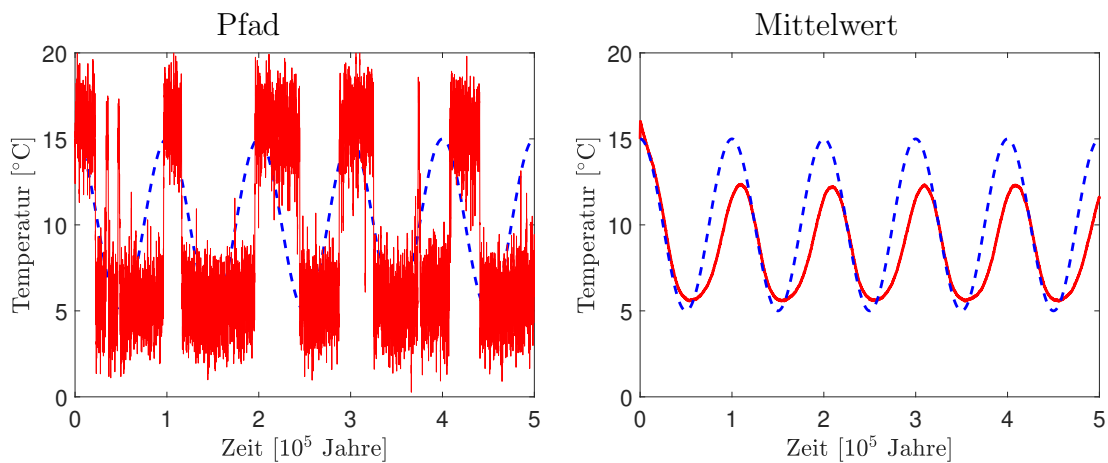


Abbildung 26: Ein Pfad (rot) sowie Mittelwert (rot) der globalen Temperatur und schematischer Verlauf der Aufwärmung (blau).

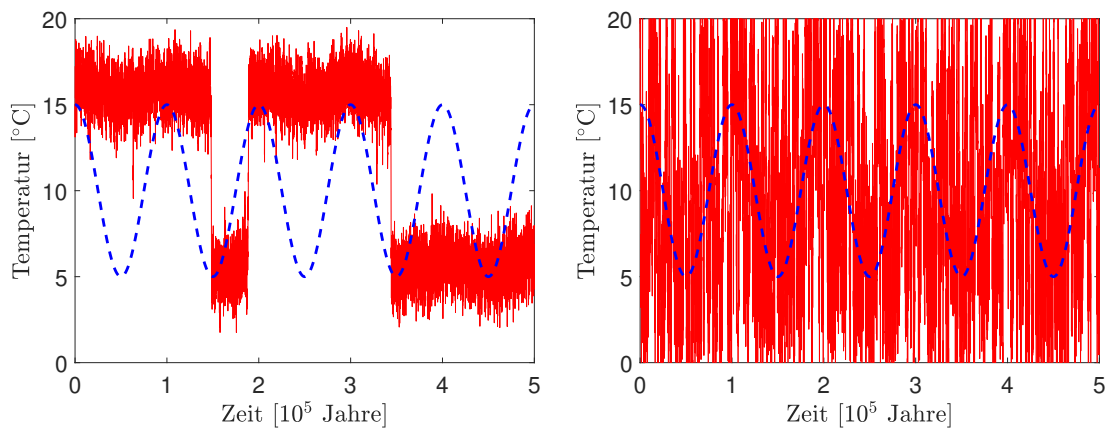


Abbildung 27: Pfade der globalen Temperatur sowohl bei kleinem Parameter $\sqrt{\varepsilon} = 2.5 \cdot 10^{-5}$ (links) als auch bei großem Parameter $\sqrt{\varepsilon} = 7 \cdot 10^{-5}$ (rechts).

nischen Systemen erhalten wir eine maximale Amplitude, wenn die von außen erzwungene Frequenz mit der Eigenfrequenz des Systems übereinstimmt. Analog möchten wir hier den Parameter ε so vorgeben, dass der im System enthaltene periodische Anteil so maximal wie möglich verstärkt wird. Auf diese Weise kann ein sehr geringer periodischer Effekt, der dem System inhärent ist, überhaupt erst signifikant werden. Es ist bemerkenswert, dass bei den Klimawechseln von Eis- bzw. Warmzeit die Größenordnung der zufälligen Schwankungen sich im optimalen Bereich befindet, so dass der durch interplanetare Wechselwirkung entstehende geringfügige Effekt erhebliche Änderungen verursacht. Stochastische Resonanz tritt auch in anderen Anwendungsgebieten auf, z.B. bei chemischen Reaktionssystemen, bei biologischen Interaktionen, bei Supraleitern oder bei optischen Systemen. Weiterführende Literatur findet sich in [11, 12] bzw. den dortigen Verweisen.

7.2 Stochastische Periodizität

Im folgenden gehen wir noch auf das Phänomen der stochastischen Resonanz in einem vereinfachten Modell ein. Eine entscheidende Frage ist, wie die Periodizität eines stochastischen Prozesses charakterisiert wird. Bei Rauscheffekten können wir nämlich nie erwarten, dass sich ein stochastischer Prozess strikt periodisch im üblichen Sinn verhält.

Gegeben sei die stochastische Differentialgleichung

$$dX_t = -U' \left(\frac{t}{2T}, X_t \right) dt + \sqrt{\varepsilon} dW_t \quad (7.7)$$

mit Parameter $\varepsilon > 0$ und vorgegebener glatter Funktion U . Der Wiener Prozess $(W_t)_{t \geq 0}$ sei auf dem Wahrscheinlichkeitsraum (Ω, \mathcal{A}, P) gegeben. Hier stimmen Itô-Integral und Stratonovich-Integral überein. Ziel ist es, diese Gleichung für kleine ε zu diskutieren, d.h. im Grenzfall $\varepsilon \rightarrow 0$.

Die zugehörige Funktion $U : [0, \infty) \times \mathbb{R} \rightarrow \mathbb{R}$ konstruieren wir wie folgt. Es sei $\hat{U} : \mathbb{R} \rightarrow \mathbb{R}$ eine glatte Funktion mit lokalen Minima an den Stellen $x = -1$ und $x = 1$ und einem lokalen Maximum an der Stelle $x = 0$. Sonst treten keine lokalen Optima auf. Zudem sei $x = 0$ eine Nullstelle der Funktion. Es gelte $\hat{U} \rightarrow \infty$ für $x \rightarrow \pm\infty$. Desweiteren sei $\hat{U}(1) = -\frac{a}{2}$ und $\hat{U}(-1) = -\frac{A}{2}$ mit $A > a > 0$. Damit definieren wir die Funktion

$$U(t, x) = \begin{cases} \hat{U}(x) & \text{für } t \in [0, \frac{1}{2}) \\ \hat{U}(-x) & \text{für } t \in [\frac{1}{2}, 1) \end{cases} \quad (7.8)$$

und setzen diese periodisch auf $t \in \mathbb{R}$ mit Periode 1 fort. Abbildung 28 illustriert die Konstruktion. In (7.7) stellt $T > 0$ somit eine halbe Periode in der Zeit dar.

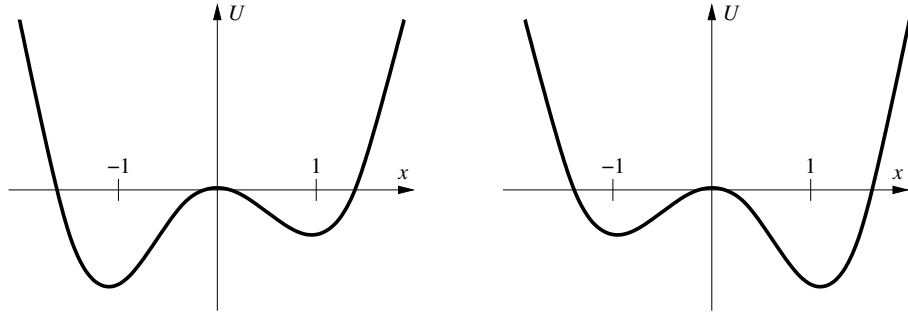


Abbildung 28: Potential im Zeitbereich $t \in [0, \frac{1}{2}]$ (links) und $t \in [\frac{1}{2}, 1]$ (rechts).

Zeitlich konstantes Potential

Wir betrachten zunächst ein autonomes Problem der Form

$$dX_t^{\varepsilon, x} = -U'(X_t) dt + \sqrt{\varepsilon} dW_t, \quad X_0^{\varepsilon, x} = x \quad (7.9)$$

mit Parameter $\varepsilon > 0$, konstantem Anfangswert $x \in \mathbb{R}$ und einer beliebigen glatten Funktion $U : \mathbb{R} \rightarrow \mathbb{R}$.

Wir betrachten den deterministischen Anteil in (7.9), d.h. die autonome gew. Dgl. $\dot{x} = -U'(x)$. Für eine Funktion $g \in C^1([0, T])$ lautet das zugehörige Residuum

$$r(t) = \dot{g}(t) + U'(g(t)). \quad (7.10)$$

Damit löst g die nichtautonome gew. Dgl. $\dot{x} = -U'(x) + r(t)$.

Für $T > 0$ definieren wir ein korrespondierendes Funktional $S_T : C([0, T]) \rightarrow \mathbb{R}_{\geq 0} \cup \{+\infty\}$ durch

$$S_T(g) := \begin{cases} \frac{1}{2} \int_0^T (\dot{g}(s) + U'(g(s)))^2 ds & \text{falls } g \text{ absolut-stetig} \\ +\infty & \text{sonst.} \end{cases} \quad (7.11)$$

Hier liegt das (halbe) Norm-Quadrat des Residuums (7.10) zum Raum $L^2([0, T])$ vor. Offensichtlich gilt $S_T \geq 0$ und $S_T(g) = 0$ genau dann, wenn g eine Lösung der gew. Dgl. $\dot{x} = -U'(x)$ auf dem Zeitintervall $[0, T]$ ist. Damit können wir ein Quasi-Potential $V : \mathbb{R}^2 \rightarrow \mathbb{R}_{\geq 0} \cup \{+\infty\}$ definieren über

$$V(x, y) = \inf\{S_T(g) : g \in C([0, T]), g(0) = x, g(T) = y, T > 0\}.$$

Das Quasi-Potential beschreibt die Arbeit, die ein Partikel leisten muss, um sich in der Landschaft des Potentials U vom Punkt x zum Punkt y zu bewegen. Wir setzen

$$\tau_{x,y}^{\varepsilon} := \inf\{t \geq 0 : X_t^{\varepsilon, x} = y\},$$

welches eine Zufallsvariable ist. Damit gilt die Aussage

$$V(x, y) = \lim_{T \rightarrow \infty} \lim_{\varepsilon \rightarrow 0} -\ln P [\tau_{x,y}^\varepsilon \leq T].$$

Über den Grenzfall für kleine ε gibt der folgende Satz Auskunft.

Satz 7.1 Sei $[x_1, x_2] \subset \mathbb{R}$ ein endliches Intervall und $0 \in [x_1, x_2]$ die einzige Nullstelle von U' in diesem Intervall. Zudem sei $x = 0$ eine asymptotisch stabile stationäre Lösung der gew. Dgl. $\dot{x} = -U'(x)$. Mit

$$\tau_x^\varepsilon := \inf \{t > 0 : X_t^{\varepsilon, x} \notin (x_1, x_2)\}$$

gilt für jedes $x \in (x_1, x_2)$

$$\lim_{\varepsilon \rightarrow 0} \varepsilon \ln \mathbb{E}(\tau_x^\varepsilon) = \min\{V(0, x_1), V(0, x_2)\} =: V_0$$

und für beliebiges $\delta > 0$

$$\lim_{\varepsilon \rightarrow 0} P [e^{(V_0 - \delta)/\varepsilon} < \tau_x^\varepsilon < e^{(V_0 + \delta)/\varepsilon}] = 1. \quad (7.12)$$

Nun setzen wir für U die obige Funktion \hat{U} . Es existieren zwei asymptotisch stabile stationäre Lösungen, auch Gleichgewichte genannt, in jeweils einer Senke. Befinden sich x und y in der gleichen Senke ($x, y > 0$ oder $x, y < 0$), dann gilt

$$V(x, y) = 2 \max\{U(y) - U(x), 0\}.$$

Der Faktor 2 erklärt die Normierung $U(-1) = -\frac{A}{2}$ und $U(1) = -\frac{a}{2}$. Im Fall $U(y) < U(x)$ folgt $V(x, y) = 0$ und ein Partikel bewegt sich in der Landschaft des Potentials von x nach y auf einer deterministischen Trajektorie, d.h. das Funktional (7.11) liefert den Wert null. Gilt dagegen $U(x) < U(y)$, so ist eine Arbeit zu leisten, um die Potentialdifferenz zu überwinden.

Liegen x und y in verschiedenen Senken, beispielsweise $-1 \leq x < 0 < y \leq 1$, dann ist eine Potentialdifferenz $U(0) - U(x)$ beim Übergang von x nach 0 zu überwinden, während die weitere Bewegung von 0 nach y keine Arbeit benötigt. Dadurch haben wir

$$V(x, y) = 2(U(0) - U(x)) \quad \text{und} \quad V(y, x) = 2(U(0) - U(y))$$

und in den Spezialfällen der Gleichgewichte wegen $U(0) = 0$

$$V(x = -1, y) = A \quad \text{und} \quad V(y = 1, x) = a.$$

Eine Interpretation der Aussage (7.12) ist, dass die mittlere Zeit, die ein Partikel benötigt um ein Intervall mit einem stabilen Gleichgewicht zu verlassen, exponentiell mit ε^{-1} ansteigt. In unserem Beispiel erwarten wir daher die Größenordnung $e^{A/\varepsilon}$ bzw. $e^{a/\varepsilon}$ beim Sprung von $x = -1$ zu $y = 1$ bzw. umgekehrt. Die zu betrachtenden Zeitskalen sollten daher exponentiell mit ε^{-1} ansteigen. Der nächste Satz berücksichtigt dies.

Satz 7.2 Für $\varepsilon > 0$ sei $T = T(\varepsilon)$ derart, dass

$$\lim_{\varepsilon \rightarrow 0} \varepsilon \ln T(\varepsilon) = \vartheta > 0. \quad (7.13)$$

Seien $\delta, K > 0$ und $x \neq 0$ beliebig. Es bezeichnet λ das Lebesgue-Maß auf \mathbb{R} . Für $\vartheta > a$ gilt

$$\lim_{\varepsilon \rightarrow 0} \lambda \left(\left\{ t \in [0, K] : |X_{tT(\varepsilon)}^{\varepsilon, x} - (-1)| > \delta \right\} \right) = 0 \quad (7.14)$$

und für $\vartheta < a$ folgt

$$\lim_{\varepsilon \rightarrow 0} \lambda \left(\left\{ t \in [0, K] : |X_{tT(\varepsilon)}^{\varepsilon, x} - \text{sign}(x)| > \delta \right\} \right) = 0$$

jeweils als stochastischer Grenzwert.

Die Konvergenzaussage als stochastischer Grenzwert bedeutet, dass mit beliebigen $\gamma > 0$ und hier $z = -1$ oder $z = 1$ gilt

$$\lim_{\varepsilon \rightarrow 0} P \left(\left\{ \omega \in \Omega : \lambda \left(\left\{ t \in [0, K] : |X_{tT(\varepsilon)}^{\varepsilon, x}(\omega) - z| > \delta \right\} \right) > \gamma \right\} \right) = 0.$$

Satz 7.2 besagt, dass bei einem Zeitintervall der Länge proportional zu $e^{\vartheta/\varepsilon}$ im Fall $\vartheta > a$ die Dynamik genug Zeit hat, um das Gleichgewicht mit tieferem Potential zu erreichen. Der Prozess verlässt zwar immer noch zeitweise eine δ -Umgebung dieses Gleichgewichts, die korrespondierenden Zeiten sind aber im Grenzübergang unbedeutend. Betrachten wir dagegen $\vartheta < a$, so hat die Dynamik nicht genug Zeit die Potentialdifferenz zwischen beiden Gleichgewichten zu überwinden. Der Prozess wird je nach $x > 0$ oder $x < 0$ beim näherliegenden Gleichgewicht mit hoher Wahrscheinlichkeit verbleiben. Auffällig bei Satz 7.2 ist noch, dass die Aussagen nur vom kleineren der Werte $a < A$ abhängt.

Zeitabhängiges Potential

Nun untersuchen wir die Dynamik in einem zeitabhängigen Potential. Wir betrachten daher statt (7.9) die Familie aus stochastischen Differentialgleichungen

$$dX_t^{\varepsilon, x} = -U' \left(\frac{t}{2T}, X_t^{\varepsilon, x} \right) dt + \sqrt{\varepsilon} dW_t, \quad X_0^{\varepsilon, x} = x \quad (7.15)$$

mit $\varepsilon > 0$ und $x \in \mathbb{R}$. Die Halbperiode hängt im folgenden von ε ab, d.h. $T = T(\varepsilon)$. Für die Funktion U wählen wir das Potential (7.8). Folglich ist der Prozess (7.15) auf den Teilintervallen $[kT, (k+1)T)$ für $k \in \mathbb{N}$ jeweils autonom. Damit kann Satz 7.2 auf jedem der Teilintervalle angewendet werden. Wir erhalten das folgende Resultat.

Satz 7.3 Für $\varepsilon > 0$ sei $T = T(\varepsilon)$ derart, dass (7.13) erfüllt ist. Seien $\delta, K > 0$ und $x \neq 0$ beliebig und λ das Lebesgue-Maß auf \mathbb{R} . Wir definieren

$$\phi(t) = \begin{cases} -1 & \text{für } t \in [k, k + \frac{1}{2}) \\ 1 & \text{für } t \in [k + \frac{1}{2}, k + 1) \end{cases} \quad \text{mit } k \in \mathbb{N}. \quad (7.16)$$

Es gilt für $\vartheta > a$

$$\lim_{\varepsilon \rightarrow 0} \lambda \left(\left\{ t \in [0, K] : |X_{t2T(\varepsilon)}^{\varepsilon, x} - \phi(t)| > \delta \right\} \right) = 0 \quad (7.17)$$

und für $\vartheta < a$

$$\lim_{\varepsilon \rightarrow 0} \lambda \left(\left\{ t \in [0, K] : |X_{t2T(\varepsilon)}^{\varepsilon, x} - \text{sign}(x)| > \delta \right\} \right) = 0$$

jeweils als stochastischer Grenzwert.

Die Funktion ϕ gibt die Position des jeweiligen globalen Minimums von U zur Zeit $2T(\varepsilon)t$ an. Gilt $\vartheta > a$, dann hat die Dynamik genug Zeit, in jeder Halbperiode dieses globale Minimum zu erreichen. Bei $\vartheta < a$ reicht die Zeit jedoch nicht zum Verlassen der ursprünglichen Senke aus.

Satz 7.3 gibt uns erste Informationen zur quantitativen Wahl des Parameters ε , um periodische Vorgänge zu erzeugen. Gilt $\vartheta < a$, dann wird sich die zeitliche Periodizität in (7.16) kaum bemerkbar machen. Wir erhalten eine untere Schranke, nämlich dass $\varepsilon \ln T(\varepsilon) > a$ sein muss.

Ist uns die Zeit T vorgegeben, dann müssen wir ein entsprechendes ε wählen. Beim Klimamodell im vorhergehenden Abschnitt war es erstaunlich, dass sich ein geeignetes Verhältnis zwischen ε und T in der Natur eingestellt hat, obwohl beide Größen durch jeweils unterschiedliche Effekte entstehen.

Der Satz 7.3 liefert jedoch keine Angaben bezüglich einer oberen Schranke für den Wert $T(\varepsilon)$. Offensichtlich wird der Prozess in (7.15) für hohes ε viele Sprünge zwischen den beiden stabilen Gleichgewichten durchführen. Die Zeit, in der er bei dem flacheren Potential verbleibt, wird dann zwar gemäß (7.17) geringfügig sein, das periodische Verhalten wird jedoch vollkommen zerstört, siehe Abbildung 27 (rechts).

Spektrale Leistungsverstärkung

Ein in der Physik beliebtes Hilfsmittel, um stochastische Resonanz zu untersuchen, ist die *spektrale Leistungsverstärkung*. Der entsprechende Koeffizient beschreibt zu den gemittelten Trajektorien eines Prozesses die Energie, die sich

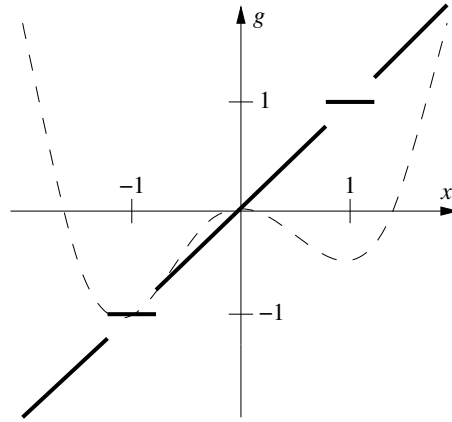


Abbildung 29: Modifizierte Identitätsfunktion.

im spektralen Anteil mit der Periode $2T$ befindet. Zur Lösung der stochastischen Differentialgleichung (7.15) ergibt sich der Koeffizient somit zu

$$\eta^X(\varepsilon, T) = \left| \int_0^1 \mathbb{E}(X_{2Ts}^{\varepsilon, x}) e^{i2\pi s} ds \right|^2. \quad (7.18)$$

(Genauer liegt hier ein Erwartungswert bezüglich eines bestimmten invarianten Maßes vor. Wir vermeiden im folgenden die Einzelheiten.) Für vorgegebene Halbperiode T tritt stochastische Resonanz auf, wenn die Abbildung $\varepsilon \mapsto \eta^X(\varepsilon, T)$ ein lokales Maximum $\varepsilon_0 > 0$ besitzt.

Erstaunlicherweise liefert dieses weitverbreitete Hilfsmittel für unser Problem mit dem Potential (7.8) keine Information, da im relevanten Parameterbereich der Koeffizient (7.18) kein lokales Maximum besitzt. Grund hierfür ist, dass bei kleinem ε die Fluktuationen beim Prozess in (7.15) hauptsächlich um die beiden stabilen Gleichgewichte mit geringen Amplituden stattfinden. Diese umfangreichen Oszillationen verschmieren die relevanten Informationen.

Ein Ausweg hieraus ist, die Kenngröße (7.18) so zu modifizieren, dass die kleinen Oszillationen um die Gleichgewichte nicht mehr berücksichtigt werden. Wir verwenden daher den modifizierten Koeffizienten

$$\tilde{\eta}^X(\varepsilon, T) = \left| \int_0^1 \mathbb{E}(g(X_{2Ts}^{\varepsilon, x})) e^{i2\pi s} ds \right|^2 \quad (7.19)$$

mit einer Funktion $g : \mathbb{R} \rightarrow \mathbb{R}$. Für $g(x) \equiv x$ entsteht der ursprüngliche Ausdruck (7.18). Nun setzen wir

$$g(x) = \begin{cases} x & \text{für } x \in (-\infty, x_1] \cup [x_2, y_1] \cup [y_2, +\infty) \\ -1 & \text{für } x \in (x_1, x_2) \\ 1 & \text{für } x \in (y_1, y_2) \end{cases}$$

mit $x_1 < -1 < x_2 < 0 < y_1 < 1 < y_2$. Abbildung 29 verdeutlicht grob das Konstruktionsprinzip. Zudem soll gelten $U(x_1) = U(x_2) = -\frac{A}{4}$ und $U(y_1) = U(y_2) = -\frac{a}{4}$.

Es folgt (unter gewissen Voraussetzungen), dass der modifizierte Koeffizient (7.19) für festes T ein lokales Maximum ε_0 besitzt. Genauer erhalten wir für jedes $\gamma > 1$ und hinreichend hohe $T > T(\gamma)$

$$\varepsilon_0(T) \in \left[\frac{1}{\gamma} \frac{A+a}{2 \ln T}, \gamma \frac{A+a}{2 \ln T} \right]. \quad (7.20)$$

Es gilt also im Grenzfall $T \rightarrow \infty$ ungefähr

$$\varepsilon_0(T) \approx \frac{A+a}{2 \ln T}.$$

Bei diesem Parameter liegt also stochastische Resonanz vor. Das Resultat (7.20) steht im Einklang mit dem aus Satz 7.3 ersichtlichen Parameterbereich, denn es gilt entsprechend zu (7.13)

$$\lim_{T \rightarrow \infty} \varepsilon_0(T) \ln T = \lim_{T \rightarrow \infty} \frac{A+a}{2 \ln T} \ln T = \frac{A+a}{2} > a$$

wegen $A > a > 0$.

Für nähere Erläuterungen zu diesem Abschnitt siehe [11].

Literatur

- [1] L.J.S. Allen: *A primer on stochastic epidemic models: Formulation, numerical simulation, and analysis*. Infect. Dis. Model., Vol. 2 (2017), S. 128-142.
- [2] H. Bauer: *Maß- und Integrationstheorie*. (2. Aufl.), de Gruyter, Berlin 1992.
- [3] H. Bauer: *Wahrscheinlichkeitstheorie*. (5. Aufl.), de Gruyter, Berlin 2002.
- [4] F. Brauer, C. Castillo-Chavez: *Mathematical models in population biology and epidemiology*. Springer, New York 2001.
- [5] V. Capasso, D. Bakstein: *An Introduction to Continuous-Time Stochastic Processes*. Birkhäuser, Boston 2005.
- [6] M. Günther, A. Jüngel: *Finanzderivate mit MATLAB*. (2. Aufl.) Vieweg + Teubner, Wiesbaden 2010.
- [7] J. Hakala, U. Wystруп: *FX Basket Options*. Frankfurt School of Finance & Management, 2008.
- [8] S. Heston: *A closed-form solution for options with stochastic volatility with application to bond and currency options*. Rev. Financ. Studies, Vol. 6 (1993), S. 327-343.
- [9] D.J. Higham: *An Introduction to Financial Option Valuation*. Cambridge University Press, 2004.
- [10] D.J. Higham: *An algorithmic introduction to numerical simulation of stochastic differential equations*. SIAM Review, Vol. 43, No. 3 (2001), S. 525-546.
- [11] P. Imkeller, I. Pavlyukevich: *Model reduction and stochastic resonance*. Stochastics and Dynamics, Vol. 2 No. 4 (2002) S. 463-506.
- [12] P. Imkeller, I. Pavlyukevich: *Stochastic resonance: a comparative study of two-state models*. In: R.C. Dalang, M. Dozzi, F. Russo, Seminar on Stochastic Analysis, Random Fields and Applications IV, Birkhäuser, Basel, 2004, S. 141-154.
- [13] P.E. Kloeden, E. Platen: *Numerical Solution of Stochastic Differential Equations*. (2. Aufl.) Springer, Berlin, 1995.
- [14] B. Øksendal: *Stochastic Differential Equations*. (6. Aufl.) Springer, Berlin, 2010.
- [15] T. Stocker: *Einführung in die Klimamodellierung*. Skript, Physikalisches Institut, Universität Bern, 2008.
- [16] J. Stoer, R. Bulirsch: *Introduction to Numerical Analysis*. (3. Aufl.), Springer, New York, 2002.