

Skript zur Vorlesung:  
„Numerik stochastischer Differentialgleichungen“

Prof. Dr. P.E. Kloeden  
Fachbereich Mathematik

Johann Wolfgang Goethe Universität  
Zimmer 101, Robert-Mayer-Straße 10

Telefon: (069) 798 28622 — Sekretariat (069) 798 22422

email: [kloeden@math.uni-frankfurt.de](mailto:kloeden@math.uni-frankfurt.de)

31. März 2011



# Inhaltsverzeichnis

<b>1</b>	<b>Überblick</b>	<b>4</b>
1.1	Gewöhnliche Differentialgleichungen . . . . .	4
1.2	Stochastische Differentialgleichungen . . . . .	5
<b>2</b>	<b>Wahrscheinlichkeitstheorie</b>	<b>8</b>
2.1	Unabhängigkeit . . . . .	11
2.2	Konvergenz . . . . .	11
2.3	Pseudo-Zufallsvariablen-Erzeuger . . . . .	12
<b>3</b>	<b>Stochastische Prozesse</b>	<b>14</b>
3.1	Der Wiener Prozess . . . . .	15
<b>4</b>	<b>Stochastische Integrale</b>	<b>19</b>
4.1	Das Ito'sche stochastische Integral . . . . .	21
4.2	Das Stratonovich'sche stochastische Integral . . . . .	24
<b>5</b>	<b>Stochastische Differentialgleichungen</b>	<b>26</b>
5.1	Der Existenz- und Eindeutigkeitssatz . . . . .	28
<b>6</b>	<b>Das stochastische Euler-Verfahren</b>	<b>30</b>
<b>7</b>	<b>Starke und schwache Konvergenz</b>	<b>35</b>
7.1	Starke Approximationen . . . . .	36
7.2	Schwache Approximationen . . . . .	37
7.2.1	Das vereinfachte Euler-Verfahren . . . . .	38
7.2.2	Funktionalschätzung . . . . .	39
7.3	Konsistenz . . . . .	40
<b>8</b>	<b>Die stochastische Kettenregel</b>	<b>43</b>
8.1	Die Ito-Formel . . . . .	43
8.2	Die Stratonovich-Kettenregel . . . . .	45
8.3	Die Feynman–Kac-Formel . . . . .	47

<b>9 Konvergenzbeweise</b>	<b>48</b>
9.1 Starke Konvergenz . . . . .	48
9.1.1 Beweis-Skizze für starke Konvergenz . . . . .	50
9.2 Schwache Konvergenz . . . . .	52
9.2.1 Beweis-Skizze für schwache Konvergenz . . . . .	52
<b>10 Stochastische Taylor-Entwicklungen</b>	<b>56</b>
<b>11 Taylor-Bäume</b>	<b>61</b>
11.1 Die Notation von Wagner und Platen . . . . .	64
11.1.1 Multi-Indizes . . . . .	64
11.1.2 Iterierte Differential-Operatoren . . . . .	65
11.1.3 Iterierte Integrale . . . . .	66
11.1.4 Hierarchische Menge . . . . .	66
<b>12 Beispiele stochastischer Taylor-Verfahren</b>	<b>68</b>
12.1 Schwache Taylor-Verfahren . . . . .	71
12.2 Starke Taylor-Verfahren . . . . .	72
<b>13 Vektorwertige SDGLen</b>	<b>74</b>
13.1 Die Ito-Formel . . . . .	76
13.2 Die stochastische Taylor-Entwicklung . . . . .	78
<b>14 Approximation stochastischer Multi-Integrale</b>	<b>81</b>
<b>15 Stochastische Multi-Integrale mit bekannten Verteilungen</b>	<b>87</b>
<b>16 Pfadweisekonvergenz</b>	<b>92</b>
16.1 Restrictiveness of the standard assumptions . . . . .	93
16.1.1 Counterexamples for the Euler-Maruyama scheme . . . . .	94

# Kapitel 1

## Überblick

Hier geben wir einen kurzen Überblick zu der zweiwochenstündigen Vorlesung

„Numerik für stochastische Differentialgleichungen“

Meine andere Vorlesung in diesem Semester (vier Wochenstunden)

„Numerical Methods for Differential Equations“

betrachtet zwei Arten deterministischer Differentialgleichungen — gewöhnliche Differentialgleichungen und partielle Differentialgleichungen. In der Vorlesung „Numerik für stochastische Differentialgleichungen“ betrachten wir hauptsächlich Ito stochastische Differentialgleichungen, aber wir werden auch Stratonovich stochastische Differentialgleichungen verwenden.

Die beiden Vorlesungen sind im Prinzip unabhängig, obwohl sie ähnliche Themen untersuchen. Am besten sollte man beide gleichzeitig hören

### 1.1 Gewöhnliche Differentialgleichungen

Betrachte eine Anfangswertaufgabe für eine gewöhnliche Differentialgleichung:

$$\frac{dx}{dt} = f(t, x), \quad x(t_0) = x_0, \quad x \in \mathbb{R}^d, \quad t \in [t_0, T]. \quad (1.1)$$

Eine Lösung ist eine stetig differenzierbare Funktion  $x : [t_0, T] \rightarrow \mathbb{R}^d$  mit  $x(t_0) = x_0$ , die der Differentialgleichung für jedes  $t \in [t_0, T]$  genügt, d.h.

$$\frac{d}{dt}x(t) = f(t, x(t)), \quad \forall t \in [t_0, T].$$

Voraussetzung: *Es existiere eine eindeutige Lösung  $x(t) = x(t; t_0, x_0)$  für  $t \in [t_0, T]$  ( $T < \infty$  hier).*

Im Allgemeinen kann man die Lösung nicht explizit finden und muß eine numerische Approximation berechnen. Betrachte eine Zerlegung des Intervalls  $[t_0, T]$

$$t_0 < t_1 < \dots < t_n < t_{n+1} < \dots < t_N = T$$

mit Schrittweite  $h_n = t_{n+1} - t_n > 0$ .

Es gibt zwei Arten von rekursiven numerischen Verfahren, die für Berechnungen mit einem digitalen Rechner geeignet sind.

- 1-Schrittverfahren z.B., die Euler, Taylor und Runge-Kutta Verfahren

$$x_{n+1} = x_n + h_n F(h_n; t_n, x_n)$$

- Mehrschrittverfahren z.B., die Adams-Bashford und Adams-Moulton Verfahren ( $k$ -Schritt mit  $k \geq 2$ )

$$x_{n+1} = G(h_n, \dots, h_{n-k+1}, t_n, \dots, t_{n-k+1}, x_n, \dots, x_{n-k+1})$$

Das einfachste numerische Verfahren ist das Euler-Verfahren:

$$x_{n+1} = x_n + h_n f(t_n, x_n) \quad (1.2)$$

mit dem Anfangswert  $x(t_0) = x_0$  von der Anfangswertaufgabe. Hier soll  $x_n \approx x(t_n)$  sein. Der Fehler

$$\max_{n=0,1,\dots,N} |x_n - x(t_n)|$$

heißt globaler Diskretisierungsfehler. Für das Euler-Verfahren kann man zeigen, dass

$$\max_{n=0,1,\dots,N} |x_n - x(t_n)| \leq K_T h,$$

wobei  $K_T$  eine Konstante ist, die von  $T$  abhängt, aber nicht von der maximalen Schrittweite  $h := \max_{n=0,1,\dots,N} h_n$ , d.h., das Euler-Verfahren ist ein numerisches Verfahren erster Ordnung.

## 1.2 Stochastische Differentialgleichungen

Stochastische Differentialgleichungen sind Differentialgleichungen mit Rauschen, zum Beispiel

$$\frac{dx}{dt} = f(t, x) + \xi_t,$$

wobei  $\xi_t$  „weißes Rauschen“ ist, oder allgemeiner

$$\frac{dx}{dt} = f(t, x) + g(t, x) \xi_t.$$

Sei  $\{W_t, t \geq 0\}$  ein Wiener Prozeß, d.h. eine mathematische Formulierung einer physikalischen Brown'schen-Bewegung.

Dann gilt

$$\xi_t = \frac{dW_t}{dt} \quad !$$

und die obige geräuschte DGL lautet jetzt

$$\frac{dx}{dt} = f(t, x) + g(t, x) \frac{dW_t}{dt} \quad !$$

Wir schreiben „!“ hier, weil ein Wiener Prozeß nicht differenzierbar entweder pfadweise oder in Quadratmittel ist. Schlimmer noch, die Pfade eines Wiener Prozesses sind nicht von beschränkter Variation auf jedem beschränkten Intervall.

Deswegen schreiben wir die obige DGL

$$\boxed{dX_t = f(t, X_t) dt + g(t, X_t) dW_t}$$

Wir nennen eine solche DGL „stochastische Differentialgleichung“ (SDGL) und interpretieren diese SDGL als eine Integralgleichung:

$$X_t = X_{t_0} + \underbrace{\int_{t_0}^t f(s, X_s) ds}_{\text{pfadweise ein Riemansches Integral}} + \underbrace{\int_{t_0}^t g(s, X_s) dW_s}_{\text{stochastisches Integral}}$$

Ein stochastisches Integral kann pfadweise kein Riemann-Stieltjes'sches Integral sein weil die Pfade eines Wiener Prozesses ohne beschränkte Variation sind.

⇒ Wir brauchen eine neue Art Integral, das Ito'sche stochastische Integral (nach dem japanischen Mathematiker K. Ito).

Stochastisches Kalkül oder Ito'sches Kalkül hat komische Eigenschaften. z.B. die Kettenregel oder Ito-Formel enthält einen zusätzlichen Term. Sei  $X_t$  eine Lösung der SDGL und definiere

$$Y_t = U(t, X_t)$$

Die Ito-Formel lautet

$$\boxed{dY_t = L^0 U(t, X_t) dt + L^1 U(t, X_t) dW_t}$$

wobei

$$L^0 U(t, x) = \frac{\partial U}{\partial t}(t, x) + f(t, x) \frac{\partial U}{\partial x}(t, x) + \underbrace{\frac{1}{2} g^2(t, x) \frac{\partial^2 U}{\partial x^2}(t, x)}_{\text{zusätzlicher Term}}$$

$$L^1 U(t, x) = g(t, x) \frac{\partial U}{\partial x}(t, x)$$

⇒ viele angepasste deterministische numerische Verfahren sind ungeeignet für SDGLen, d.h. inkonsistent und nicht konvergent.

Wir können konsistente numerische Verfahren durch stochastische Taylor-Entwicklungen herleiten. Es gibt verschiedene Möglichkeiten: starke und schwache Verfahren, die von der Art der Approximation abhängen

- stark  $L^1$ -Approximation ⇒ eine Approximation der Pfade
- schwach Approximation der Wahrscheinlichkeitsmaße.

Wir werden beide Arten von Verfahren untersuchen und auch Implementierungsfragestellungen erörtern.

## Kapitel 2

# Wahrscheinlichkeitstheorie

Betrachte eine „verauschte“ DGL

$$\frac{dx}{dt} = f(t, x) + \text{„Rauschen“},$$

wobei Rauschen eine zeitabhängige zufällige Störung ist

z.B. weißes Gauß'sches Rauschen.

Der „deterministische“ Term  $f(t, x)$  heißt Drift. Im Allgemeinen hat das Rauschen  $\xi_t$  einen  $x$ -abhängigen Intensitätskoeffizienten  $g(t, x)$ . Die geräuschte DGL lautet jetzt

$$\frac{dx}{dt} = f(t, x) + g(t, x)\xi_t$$

Aber dies ist noch sehr vage. Um weiter zu gehen, brauchen wir ein paar Tatsachen aus der Stochastik, z.B. Begriffe wie Zufallsvariablen und stochastische Prozesse.

Im Hintergrund steht immer ein Wahrscheinlichkeitsraum

$$(\Omega, \mathcal{A}, P),$$

dabei sind:

- 1)  $\Omega$  der Stichprobenraum (= eine beliebige nichtleere Menge)
- 2)  $\mathcal{A}$  eine  $\sigma$ -Algebra von Ereignissen d.h. Teilmengen von  $\Omega$  mit Eigenschaften
  - (i)  $\emptyset \in \mathcal{A}$
  - (ii)  $A^c = \Omega \setminus A \in \mathcal{A}$ , falls  $A \in \mathcal{A}$
  - (iii)  $\cup_{i=1}^{\infty} A_i \in \mathcal{A}$ , falls  $A_1, A_2, \dots \in \mathcal{A}$
- 3)  $P$  ein Wahrscheinlichkeitsmaß (auf  $\mathcal{A}$ ), d.h.  $P : \mathcal{A} \rightarrow [0, 1]$  mit
  - (i)  $P(\emptyset) = 0, \quad P(\Omega) = 1$

$$(ii) P(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i), \text{ falls } A_i \cap A_j = \emptyset \quad \forall i \neq j$$

Der Wahrscheinlichkeitsraum ist meistens nicht explizit bekannt. Statt dessen haben wir nur Daten oder Messungen

$$X(\omega) \in \mathbb{R}, \quad \omega \in \Omega.$$

Eine Zufallsvariable oder Zufallsgröße  $X$  ist eine  $\mathcal{A}$ -messbare Funktion

$$X : \Omega \rightarrow \mathbb{R},$$

d.h mit

$$X^{-1}(a) := \{\omega \in \Omega : X(\omega) \leq a\} \in \mathcal{A} \quad \forall a \in \mathbb{R}$$

( $\mathcal{A}$ -messbar  $\cong$  konsistent mit der „Information“ der  $\sigma$ -Algebra  $\mathcal{A}$ .)

Sei  $X$  eine Zufallsvariable und definiere die Funktion

$$F_X : \mathbb{R} \rightarrow [0, 1]$$

durch

$$F_X(a) := P(\{\omega \in \Omega : X(\omega) \leq a\}), \quad \forall a \in \mathbb{R}.$$

Die Funktion  $F_X$  heißt Verteilung der ZV  $X$ . Sie besitzt die Eigenschaften

$$(i) 0 \leq F_X(a) \leq F_X(b) \leq 1, \quad a \leq b$$

$$(ii) \lim_{a \rightarrow -\infty} F_X(a) = 0$$

$$(iii) \lim_{a \rightarrow +\infty} F_X(a) = 1.$$

Die Funktion  $F_X$  ist manchmal absolut stetig mit einer Ableitung  $f_X(x) = F'_X(x)$  für fast alle  $x \in \mathbb{R}$ . Es gelten

$$F_X(b) - F_X(a) = \int_a^b f_X(x) dx, \quad a \leq b$$

und

$$F_X(a) = \int_{-\infty}^a f_X(x) dx$$

Die Funktion  $f_X$  heißt Dichte der Verteilung  $F_X$  (oder der ZV  $X$ ).

### Beispiele

1) Sei  $X$  auf  $[0, 1]$  gleichmäßig verteilt

$$f_X(x) = \begin{cases} 1, & \text{falls } x \in [0, 1] \\ 0, & \text{sonst.} \end{cases}$$

- 2) Sei  $X \sim N(0, 1)$ -verteilt, d.h. standard normal verteilt oder Standard-Gauß-verteilt mit der Dichte

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}.$$

Zufallsvariablen haben nützliche numerische Merkmale, z.B.

$$\text{Erwartungswert} \quad \mathbb{E}(X) := \int_{-\infty}^{\infty} x f_X(x) dx$$

$$p\text{-ter Moment} \quad \mathbb{E}(X^p) := \int_{-\infty}^{\infty} x^p f_X(x) dx$$

$$\text{Varianz} \quad \text{Var}(X) := \mathbb{E}((X - \mathbb{E}(X))^2)$$

Die Varianz charakterisiert die Streuung um den Erwartungswert. Es gilt auch

$$\text{Var}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2 \quad (2.1)$$

Beispiele

- 1)  $X$  ist auf  $[0, 1]$  gleichmäßig verteilt:

$$\mathbb{E}(X) = \frac{1}{2}, \quad \text{Var}(X) = \frac{1}{12}$$

- 2)  $X$  ist  $N(0, 1)$  verteilt:

$$\mathbb{E}(X) = 0, \quad \text{Var}(X) = 1$$

auch

$$\mathbb{E}(X^{2p+1}) = 0, \quad \mathbb{E}(X^{2p}) = (2p-1)!! = (2p-1)(2p-3)\cdots 3 \cdot 1.$$

Eine Gauß-verteilte Zufallsvariable  $X \sim N(\mu, \sigma^2)$  bedeutet:  $X$  ist eine Gauß-verteilte Zufallsvariable mit

$$\text{Erwartungswert} \quad \mathbb{E}(X) = \mu$$

$$\text{Varianz} \quad \text{Var}(X) = \sigma^2.$$

Die entsprechende Verteilungsdichte lautet

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/2\sigma^2}$$

$\Rightarrow$  Die zentralisierten Momente sind

$$\mathbb{E}((X - \mu)^{2p+1}) = 0, \quad \mathbb{E}((X - \mu)^{2p}) = (2p-1)!! \sigma^p$$

für  $p = 1, 2, \dots$ . Hier:  $(2p-1)!! = (2p-1)\cdots 3 \cdot 1$

d.h. alle (zentralisierten) Momente einer Gauß-verteilten Zufallsvariable sind durch die zwei Parameter  $\mu$  und  $\sigma$  bestimmt  $\Rightarrow$  sehr ungewöhnlich !

## 2.1 Unabhängigkeit

Zwei Zufallsvariable  $X$  und  $Y$  (auf demselben Wahrscheinlichkeitsraum) heißen unabhängig, falls

$$\begin{aligned} & P(\{\omega \in \Omega : X(\omega) \leq a\} \cap \{\omega \in \Omega : Y(\omega) \leq b\}) \\ &= P(\{\omega \in \Omega : X(\omega) \leq a\}) \cdot P(\{\omega \in \Omega : Y(\omega) \leq b\}) \end{aligned}$$

für alle  $a, b \in \mathbb{R}$

oder bezüglich Verteilungen und Dichten

$$F_{(X,Y)}(a, b) = F_X(a) \cdot F_Y(b)$$

$$f_{(X,Y)}(a, b) = f_X(a) \cdot f_Y(b)$$

für alle  $a, b \in \mathbb{R}$

Dann gelten

- (i)  $\mathbb{E}(XY) = \mathbb{E}(X) \cdot \mathbb{E}(Y)$
- (ii)  $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$

aber diese Bedingungen (i) und (ii) sind nur notwendige Bedingungen – sie sind nicht hinreichend für die Unabhängigkeit von  $X$  und  $Y$ .

## 2.2 Konvergenz

Betrachte jetzt eine ZV  $\bar{X}$  und eine Folge von ZVen  $\{X_n\}_{n \geq 0}$  auf demselben Wahrscheinlichkeitsraum.

Was bedeutet Konvergenz  $X_n \rightarrow \bar{X}$  für  $n \rightarrow \infty$ ?

Es gibt mehrere nützliche Definitionen.

- 1) Konvergenz mit Wahrscheinlichkeit 1 oder fast sicher)

$$P\left(\left\{\omega \in \Omega : \lim_{n \rightarrow \infty} X_n(\omega) = \bar{X}(\omega)\right\}\right) = 1$$

d.h.

$$\exists \hat{\Omega} \in \mathcal{A} \text{ mit } P(\hat{\Omega}) = 1 \quad \text{und} \quad X_n(\omega) \rightarrow \bar{X}(\omega) \text{ für } n \rightarrow \infty \text{ für jedes } \omega \in \hat{\Omega}.$$

Wir schreiben:  $X_n \rightarrow \bar{X}$  m.W. 1 oder f.s.

2) Quadratmittel-Konvergenz

$$\mathbb{E}(|X_n - \bar{X}|^2) \rightarrow 0 \quad \text{für } n \rightarrow \infty$$

Wir schreiben:  $X_n \rightarrow \bar{X}$  q.m.

Bemerkung

$$|\mathbb{E}(X_n) - \mathbb{E}(\bar{X})| \leq \mathbb{E}(|X_n - \bar{X}|) \leq \sqrt{\mathbb{E}(|X_n - \bar{X}|^2)}$$

Cauchy-Schwarz-Ungleichung!

$$\Rightarrow \quad \mathbb{E}(X_n) \rightarrow \mathbb{E}(\bar{X}), \quad \text{Var}(X_n - \bar{X}) \rightarrow 0 \quad \text{und} \quad \text{Var}(X_n) \rightarrow \text{Var}(\bar{X})$$

Daher wegen (2.1) gilt insgesamt

qm-Konvergenz  $\Leftrightarrow$  Konvergenz der Erwartungswerte und der Varianzen

3) Konvergenz in Wahrscheinlichkeit

$$\lim_{n \rightarrow \infty} P(\{\omega \in \Omega : |X_n(\omega) - \bar{X}(\omega)| \geq \epsilon\}) = 0 \quad \text{für alle } \epsilon > 0$$

auch genannt: stochastische Konvergenz

Bemerkungen1) qm-Konvergenz  $\Rightarrow$  Konvergenz in Wahrscheinlichkeit

wegen der Tschebyscheff-Ungleichung, d.h.

$$P(\{\omega \in \Omega : |X(\omega)| \geq \epsilon\}) \leq \frac{1}{\epsilon^2} \mathbb{E}(|X|^2)$$

2) Konvergenz m.W.1  $\Rightarrow$  qm-Konvergenz

nur unter gewissen Bedingungen, z.B.  $\bar{X}$  und die  $X_n$  sind gleichmäßig beschränkt (f.s.)

$$|\bar{X}(\omega)| \leq L, \quad |X_n(\omega)| \leq L \quad \forall \omega \in \hat{\Omega}, \quad \forall n \in \mathbb{N}$$

wo  $P(\hat{\Omega}) = 1$ .

**2.3 Pseudo-Zufallsvariablen-Erzeuger**

Pseudo-Zufallsvariablen-Erzeuger heißen pseudo, weil sie rein deterministisch sind, d.h. durch deterministische Differenzgleichungen definiert sind.

z.B. die Linearkongruent-PZVE sind der Form

$$X_{n+1} = aX_n + b \pmod{c}$$

für geeignete Parameter  $a, b, c \Rightarrow X_n \in [0, c]$

$$\Rightarrow U_n = X_n/c \in [0, 1]$$

Die Paare sukzessiver  $(U_n, U_{n+1})$  liegen auf Geraden mit Neigung  $a/c$  in dem Einheitsrechteck  $[0, 1] \times [0, 1]$

Ein Anfangswert  $X_0$  heißt „Samen“ ( Engl. seed)

Beispiel „RANDU“ von IBM in den 1960s

$$a = 2^{16} + 3, \quad b = 0, \quad c = 2^{31}$$

Potenzen von 2 sind bevorzugt wegen binärer Computer-Arithmetik.

# Kapitel 3

## Stochastische Prozesse

Sei  $(\Omega, \mathcal{A}, P)$  ein Wahrscheinlichkeitsraum. Sei  $\mathbb{T}$  eine nichtleere Teilmenge von  $\mathbb{R}$ , die wir als einen Zeitraum betrachten werden, z.B.

$$\mathbb{T} = \{0, 1, 2, \dots\} \quad \text{zeitdiskreter Fall}$$

$$\mathbb{T} = [0, T] \quad \text{zeitkontinuierlicher Fall}$$

Eine Funktion  $X : \mathbb{T} \times \Omega \rightarrow \mathbb{R}$ , wobei  $X(t, \cdot)$  eine Zufallsvariable für jedes  $t \in \mathbb{T}$  ist, heißt stochastischer Prozess (SP).

### Bemerkungen

- 1 Wir können  $X$  betrachten entweder als eine Familie von Zufallsvariablen  $\{X(t, \cdot)\}_{t \in \mathbb{T}}$  oder als eine Familie von Pfaden  $\{X(\cdot, \omega)\}_{\omega \in \Omega}$ . Jeder Pfad ist eine Realisierung des SPs.
- 2 In der Stochastik schreibt man  $X_t(\omega)$  für  $X(t, \omega)$

Beispiel Wir können eine Folge von Zufallsvariablen  $\{X_n\}_{n=0}^{\infty}$  als einen zeitdiskreten stochastischen Prozess interpretieren.

Es gibt viele Möglichkeiten für die Beziehung zwischen den verschiedenen Zufallsvariablen eines stochastischen Prozesses, d.h. zwischen  $X_{t_1}$  und  $X_{t_2}$  für alle  $t_1, t_2 \in \mathbb{T}$ . Zwei Extreme (in dem zeitdiskreten Fall) sind

- 1) Differenzgleichungen

$$X_{n+1}(\omega) = g(X_n(\omega)) \quad \forall n = 0, 1, 2, \dots \text{ und alle } \omega \in \Omega$$

wobei  $g : \mathbb{R} \rightarrow \mathbb{R}$  eine gegebene (deterministische) Funktion ist.

- 2) i.i.d. -Zufallsvariable: die  $X_n$  sind unabhängig (independent) und haben dieselbe Verteilung (identically distributed).

Im Allgemeinen betrachtet man eine ganze Familie von Verteilungen der Art

$$F(t_1, x_1; t_2, x_2; \dots; t_j, x_j) \\ = P(\{\omega : X_{t_1}(\omega) \leq x_1\} \cap \{\omega : X_{t_2}(\omega) \leq x_2\} \cap \dots \cap \{\omega : X_{t_j}(\omega) \leq x_j\}) \\ \text{für alle } t_1, t_2, \dots, t_j, \dots \in \mathbb{T} \text{ und } x_1, x_2, \dots, x_j, \dots \in \mathbb{R}$$

Beispiel Im i.i.d -Fall gilt

$$F(t_1, x_1; t_2, x_2; \dots; t_j, x_j) = F_X(x_1) \cdot F_X(x_2) \cdot \dots \cdot F_X(x_j),$$

wobei  $F_X$  die gemeinsame Verteilung aller Zufallsvariablen  $X_t, t \in \mathbb{T}$ , ist.

### 3.1 Der Wiener Prozess

Der Wiener-Prozess, oft Brown'sche Bewegung genannt, ist einer der Grundprozesse der Stochastik.

Ein Wiener-Prozess  $\{W_t, t \geq 0\}$  ist durch die folgenden Eigenschaften definiert

1)  $W_t$  ist Gauß-verteilt mit

- (i)  $W_0 = 0$  m.W. 1
- (ii)  $\mathbb{E}(W_t) = 0$  für jedes  $t \geq 0$
- (iii)  $\mathbb{E}(W_t^2) = t$  für jedes  $t \geq 0$ ,

$$\boxed{W_t \sim N(0, t) \text{ für jedes } t \geq 0}$$

2) die nichtüberlappenden Zuwächse von  $W_t$  sind unabhängig, d.h.

$$W_{t_2} - W_{t_1} \text{ und } W_{t_4} - W_{t_3}$$

sind unabhängige Zufallsvariablen für alle  $0 \leq t_1 < t_2 \leq t_3 < t_4$ .

$$\boxed{\text{Sätzchen} \quad \mathbb{E}((W_t - W_s)^2) = t - s \quad \text{für alle } 0 \leq s \leq t}$$

Beweis:

$$\begin{aligned} \mathbb{E}((W_t - W_s)^2) &= \mathbb{E}(W_t^2 - 2W_tW_s + W_s^2) \\ &= \mathbb{E}(W_t^2 - W_s^2 - 2W_tW_s + 2W_s^2) \\ &= \mathbb{E}(W_t^2) - \mathbb{E}(W_s^2) - 2\mathbb{E}((W_t - W_s)W_s) \\ &= t - s - 2\mathbb{E}((W_t - W_s)(W_s - W_0)) \end{aligned}$$

$$\begin{aligned}
& \text{nicht überlappende Zuwächse sind unabhängig} \\
& = t - s - 2\mathbb{E}(W_t - W_s) \cdot \mathbb{E}(W_s - W_0) \\
& = t - s - 2(\mathbb{E}(W_t) - \mathbb{E}(W_s)) \cdot (\mathbb{E}(W_s) - \mathbb{E}W_0) \\
& = t - s - 2(0 - 0) \cdot (0 - 0) = t - s
\end{aligned}$$

Offensichtlich gilt auch

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

und  $W_t - W_s$  ist auch Gauß-verteilt, weil lineare Kombinationen Gauß-verteilter Zufallsvariablen Gauß-verteilt sind

$$\Rightarrow W_t - W_s \sim \mathbb{N}(0, t - s)$$

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

Ein Kriterium von Kolmogoroff sagt dann,

die Pfade eines Wiener-Prozesses sind stetig

Aber von dem obigen Sätzchen folgt auch, dass

$$\mathbb{E} \left( \left( \frac{W_{t+h} - W_t}{h} \right)^2 \right) = \frac{E((W_{t+h} - W_t)^2)}{h^2} = \frac{h}{h^2} = \frac{1}{h}$$

für alle  $t \geq 0$  und  $h > 0$

$\Rightarrow$  ein Wiener-Prozess ist nirgendwo quadratmittel-differenzierbar

Schlimmer noch (und der Beweis ist viel komplizierter) ist

SATZ Die Pfade eines Wiener Prozesses sind nirgendwo differenzierbar, d.h.

$$\frac{d}{dt}W_t(\omega) \text{ existiert für kein } t \geq 0.$$

und noch schlimmer

**SATZ** Die Pfade eines Wiener-Prozesses haben unbeschränkte Variation auf jedem endlichen Intervall, d.h.

$$\sup \sum_{j=0}^{N_T-1} |W_{t_{j+1}}(\omega) - W_{t_j}(\omega)| = +\infty$$

wobei das Supremum über alle Unterteilungen  $t_0 \leq t_1 \leq \dots \leq t_{N_T} \leq T$  eines Intervalls  $[t_0, T]$  ist.

Trotzdem definieren Physiker und Ingenieure weißes Gauß'sches Rauschen  $\xi_t$  durch

$$\xi_t(\omega) = \frac{d}{dt} W_t(\omega)$$

Sie meinen einen „stochastischen“ Prozess, wobei die  $\xi_t$  unabhängig und alle  $N(0; 1)$ -verteilt sind — ein i.i.d.-Prozess!

( $\xi_t$  existiert nur im Sinne verallgemeinerter Funktionen oder Distributionen).

Dann schreiben unsere Kollegen die geräuschte DGL

$$\frac{dx}{dt} = f(t, x) + g(t, x)\xi_t$$

um:

$$\frac{dx}{dt} = f(t, x) + g(t, x) \frac{dW_t}{dt}$$

oder als

$$dX_t = f(t, X_t)dt + g(t, X_t)dW_t$$

Eine solche DGL heißt stochastische Differentialgleichung (SDGL). Sie existiert nur symbolisch als eine Differentialgleichung und ist tatsächlich nur eine Integralgleichung

$$X_t = X_{t_0} + \int_{t_0}^t f(s, X_s)ds + \int_{t_0}^t g(s, X_s)dW_s$$

Aber wir müssen sagen, wie die Integrale hier sind.

Das erste Integral ist pfadweise ein deterministisches Riemann-Integral

$$\int_{t_0}^t f(s, X_s(\omega))ds \quad (\omega \text{ fest})$$

Das zweite Integral sieht aus wie ein pfadweises deterministisches Riemann-Stieltjes-Integral.

Riemann-Integral

$$\int_a^b f(t)dt = \lim_{N \rightarrow \infty} \sum_{l=0}^{N-1} f(\tau_l)(t_{l+1} - t_l)$$

Riemann-Stieltjes-Integral

$$\int_a^b f(t)dg(t) = \lim_{N \rightarrow \infty} \sum_{l=0}^{N-1} f(\tau_l) \{g(t_{l+1}) - g(t_l)\}$$

(in beiden Fällen) für beliebige Auswertungsstellen  $\tau_l \in [t_l, t_{l+1}]$

— aber das ist nicht möglich, weil die Pfade von  $W_t$  unbeschränkte Variationen haben.

Wir brauchen ein stochastisches Integral für das zweite Integral.

# Kapitel 4

## Stochastische Integrale

Wir wollen ein stochastisches Integral

$$\int_0^T f(s, \omega) dW_s(\omega)$$

für geeignete (auch zufällige!) Funktionen  $f$  definieren.

Wie? Das pfadweise definierte Riemann-Stieltjes-Integral existiert nicht!, d.h. die Summen

$$S_N(\omega) = \sum_{j=0}^{N-1} f(\tau_j, \omega) \cdot \{W_{t_{j+1}}(\omega) - W_{t_j}(\omega)\} \quad \text{beliebiges } \tau_j \in [t_j, t_{j+1}]$$

konvergieren nicht (fast sicher).

*Könnten wir Quadratmittelkonvergenz benutzen?*

NEIN! Betrachte  $f(t, \omega) \equiv W_t(\omega)$

$\begin{aligned} \text{qm - } \lim_{N \rightarrow \infty} \sum_{j=0}^{N-1} W_{\tau_j}(\omega) \cdot \{W_{t_{j+1}}(\omega) - W_{t_j}(\omega)\} \\ \text{SATZ} \quad = \begin{cases} \frac{1}{2} W_T^2(\omega) - \frac{1}{2} T, & \text{falls } \tau_j \equiv t_j \\ \frac{1}{2} W_T^2(\omega), & \text{falls } \tau_j \equiv t_j + \frac{1}{2}(t_{j+1} - t_j) \\ \frac{1}{2} W_T^2(\omega) + \frac{1}{2} T, & \text{falls } \tau_j \equiv t_{j+1} \end{cases} \end{aligned}$
---

Beweis: Definiere

$$S_N(\omega) = \sum_{j=0}^{N-1} W_{t_j}(\omega) \cdot \{W_{t_{j+1}}(\omega) - W_{t_j}(\omega)\}$$

(Jetzt werden wir das „ $\omega$ “ nicht schreiben).

Wir müssen beweisen, dass

$$\mathbb{E} \left( \left| S_N - \frac{1}{2} W_T^2 + \frac{1}{2} T \right|^2 \right) \rightarrow 0 \text{ für } N \rightarrow \infty.$$

Definiere

$$\Delta_j = t_{j+1} - t_j = \frac{T}{N}, \quad \Delta W_j = W_{t_{j+1}} - W_{t_j} \sim N(0, \Delta_j)$$

Dann gilt

$$S_N = \frac{1}{2} \sum_{j=0}^{N-1} (W_{t_{j+1}}^2 - W_{t_j}^2) - \frac{1}{2} \sum_{j=0}^{N-1} (\Delta W_j)^2,$$

weil

$$W_{t_{j+1}}^2 = (W_{t_j} + \Delta W_j)^2 = W_{t_j}^2 + 2W_{t_j} \cdot \Delta W_j + (\Delta W_j)^2$$

d.h.

$$S_N = \frac{1}{2} W_T^2 - \frac{1}{2} W_0^2 - \frac{1}{2} \sum_{j=0}^{N-1} (\Delta W_j)^2$$

weil  $t_N = T$  und  $t_0 = 0$ .

$$\begin{aligned} \Rightarrow \left| S_N - \frac{1}{2} W_T^2 + \frac{1}{2} T \right|^2 &= \frac{1}{4} \left| T - \sum_{j=0}^{N-1} (\Delta W_j)^2 \right|^2 \\ &= \frac{1}{4} \left| \sum_{j=0}^{N-1} \{ \Delta_j - (\Delta W_j)^2 \} \right|^2 \end{aligned}$$

Definiere  $Z_j = \Delta_j - (\Delta W_j)^2$ . Dann gelten

$$\Rightarrow \left\{ \begin{array}{l} \mathbb{E}(Z_j) = \Delta_j - \mathbb{E}((\Delta W_j)^2) = \Delta_j - \Delta_j = 0 \\ \text{Var}(Z_j) = \mathbb{E}(Z_j^2) \quad \text{weil } \mathbb{E}(Z_j) = 0 \text{ hier} \\ = \mathbb{E}(\Delta_j^2 - 2\Delta_j \cdot (\Delta W_j)^2 + (\Delta W_j)^4) \\ = \Delta_j^2 - 2\Delta_j \mathbb{E}((\Delta W_j)^2) + \mathbb{E}((\Delta W_j)^4) \\ = \Delta_j^2 - 2 \cdot \Delta_j \Delta_j + 3\Delta_j^2 \\ = 2\Delta_j^2 \quad \text{weil } \Delta W_j \sim N(0, \Delta_j) \text{ ist} \end{array} \right.$$

Wichtig: die  $\Delta W_i$  und  $\Delta W_j$  sind unabhängig für  $i \neq j$  (nicht überlappende Zuwächse!)

$\Rightarrow Z_i$  und  $Z_j$  sind unabhängig für  $i \neq j$

$$\Rightarrow \text{Var} \left( \sum_{j=0}^{N-1} Z_j \right) = \sum_{j=0}^{N-1} \text{Var} (Z_j) = \sum_{j=0}^{N-1} 2(\Delta_j)^2 = 2 \sum_{j=0}^{N-1} \frac{T^2}{N^2} = \frac{2T^2}{N}$$

Aber  $\mathbb{E} \left( \sum_{j=0}^{N-1} Z_j \right) = 0$  hier  $\Rightarrow$

$$\begin{aligned} \mathbb{E} \left( \left| S_N - \frac{1}{2} W_T^2 - \frac{1}{2} T \right|^2 \right) &= \frac{1}{4} \text{Var} \left( \sum_{j=0}^{N-1} Z_j \right) \\ &= \frac{T^2}{2N} \rightarrow 0 \text{ für } N \rightarrow \infty \end{aligned}$$

Die dritte Behauptung ( $t_j \equiv t_{j+1}$ ) folgt durch

$$\sum W_{t_{j+1}} (W_{t_{j+1}} - W_{t_j}) = \underbrace{\sum (W_{t_{j+1}} - W_{t_j})^2}_{\text{qm} \rightarrow T} + \underbrace{\sum W_{t_j} (W_{t_{j+1}} - W_{t_j})}_{\text{qm} \rightarrow \frac{1}{2} W_T^2 - \frac{1}{2} T}$$

Ähnlicherweise beweist man die zweite Behauptung für  $\tau_j = t_j + \frac{1}{2}(t_{j+1} - t_j)$ . (Die Details sind etwas komplizierter und werden hier nicht ausgeschrieben).

Diese Ergebnisse zeigen: ein durch qm-Konvergenz definiertes Riemann-Stieltjes stochastisches Integral existiert auch nicht, weil die Grenzwerte von der Wahl der Auswertungsstellen  $\tau_j$  abhängen.

Was ist zu tun?      Stochastische Konvergenz?    nein !

qm-Konvergenz immer mit derselben Auswertungsstelle

- $\tau_j \equiv t_i$  für das Ito-Integral
- $\tau_j \equiv t_j + \frac{1}{2}(t_{j+1} - t_j)$  für das Stratonovich-Integral

## 4.1 Das Ito'sche stochastische Integral

Sei  $\{W_t, t \geq 0\}$  ein Wiener-Prozess. Um das stochastische Integral

$$\int_0^T f(s, \omega) dW_s(\omega)$$

zu definieren, brauchen wir eine Klasse geeigneter Integrand-Funktionen

$$f : [0, T] \times \Omega \rightarrow \mathbb{R}.$$

Wir setzen voraus, dass  $f$  die folgenden Eigenschaften besitzt:

1  $f_t(\cdot) \equiv f(t, \dots)$  ist eine Zufallsvariable (d.h.  $\mathcal{A}$ -messbar) mit

$$\mathbb{E}(f_t^2) < \infty$$

für jedes  $t \in [0, T]$

2 die Pfade  $t \rightarrow f_t(\omega)$  sind stetig auf  $[0, T]$  fast sicher

3 die  $\{f_t, t \in [0, T]\}$  sind nichtantizipativ bezüglich des Wiener-Prozesses  $\{W_t, t \geq 0\}$ , d.h.  $f_t$  und  $W_{t+h} - W_t$  sind unabhängig für jedes  $h > 0, t \in [0, T]$ .

Bemerkung Im Allgemeinen sind die Pfade  $t \rightarrow f_t(\omega)$  nur Lebesgue-messbar. Aber dieser Begriff braucht eine Kenntnis von Maß-Theorie. Auch hier brauchen wir nur stetige Pfade.

Wir definieren das Ito-Integral einer Funktion  $f$  von obiger Klasse durch

$$\int_0^T f(t, \omega) dW_t(\omega) = \text{qm} - \lim_{N \rightarrow \infty} \sum_{j=0}^{N-1} f(t_j, \omega) \cdot \{W_{t_{j+1}}(\omega) - W_{t_j}(\omega)\}$$

Der Grenzwert hier hängt nicht von der Folge der Unterteilungen des Intervalls  $[0, T]$  ab - wir brauchen nur, dass  $\Delta^{(N)} = \max_{j=0, \dots, N-1} (t_{j+1} - t_j) \rightarrow 0$  für  $N \rightarrow \infty$ .

Das Ito-Integral  $\int_0^T f_t dW_t$  besitzt die folgenden wichtigen Eigenschaften

SATZ

- |                          |  |
|--------------------------|--|
| 1. <u>Linearität</u>     | $\int_0^T (\alpha f_t + \beta g_t) dW_t = \alpha \int_0^T f_t dW_t + \beta \int_0^T g_t dW_t$<br>für alle $f, g$ von der Klasse und $\alpha, \beta \in \mathbb{R}$ |
| 2. <u>Erwartungswert</u> | $\mathbb{E} \left( \int_0^T f_t dW_t \right) = 0$  |
| 3. <u>Varianz</u>        | $\mathbb{E} \left( \left( \int_0^T f_t dW_t \right)^2 \right) = \int_0^T \mathbb{E}(f_t^2) dt$   |

Die dritte Eigenschaft (4.1) heißt die Ito-Isometrie.

Beweis (Skizze): Wir zeigen hier nur, dass jede Summe

$$\sum_{j=0}^{N-1} f_{t_j} \{W_{t_{j+1}} - W_{t_j}\}$$

der behaupteten Eigenschaft genügt – dann benutzt man qm-Konvergenz, um die Aussage zu erhalten.

Schreibe

$$f_j = f_{t_j}, \quad g_j = g_{t_j}, \quad \Delta W_j = W_{t_{j+1}} - W_{t_j}.$$

1)  $\alpha f + \beta g$  gehört zu der obigen Klasse. Dann

$$\begin{aligned} \sum_{j=0}^{N-1} (\alpha f + \beta g)_j \Delta W_j &= \sum_{j=0}^{N-1} (\alpha f_j + \beta g_j) \cdot \Delta W_j \\ &= \alpha \sum_{j=0}^{N-1} f_j \Delta W_j + \beta \sum_{j=0}^{N-1} g_j \Delta W_j \end{aligned}$$

2)  $f_j$  und  $\Delta W_j$  sind unabhängig wegen der Nichtantizipativität von  $f$ . Dann

$$\begin{aligned} \mathbb{E} \left( \sum_{j=0}^{N-1} f_j \Delta W_j \right) &= \sum_{j=0}^{N-1} \mathbb{E}(f_j \cdot \Delta W_j) \\ &= \sum_{j=0}^{N-1} \mathbb{E}(f_j) \cdot \mathbb{E}(\Delta W_j) \quad [\text{unabhängig !}] \\ &= \sum_{j=0}^{N-1} \mathbb{E}(f_j) \cdot 0 = 0 \end{aligned}$$

3)  $f_j \Delta W_j$  und  $f_i \Delta W_i$  sind unabhängig für  $i \neq j$  wegen der Nichtantizipativität von  $f$  und der Unabhängigkeit der nichtüberlappenden Zuwächse von  $W_t$ . Dann

$$\begin{aligned} \mathbb{E} \left( \left( \sum_{j=0}^{N-1} f_j \Delta W_j \right)^2 \right) &= \text{Var} \left( \sum_{j=0}^{N-1} f_j \Delta W_j \right), \quad \left[ \text{weil } \mathbb{E} \left( \sum_{j=0}^{N-1} f_j \Delta W_j \right) = 0 \text{ ist} \right] \\ &= \sum_{j=0}^{N-1} \text{Var}(f_j \Delta W_j) \quad [\text{unabhängige Zufallsvariablen}] \\ &= \sum_{j=0}^{N-1} \mathbb{E}((f_j \Delta W_j)^2) \\ &= \sum_{j=0}^{N-1} \mathbb{E}(f_j^2) \cdot \mathbb{E}((\Delta W_j)^2) \quad [\text{Nichtantizipativität}] \\ &= \sum_{j=0}^{N-1} \mathbb{E}(f_j^2) \cdot \Delta_j \end{aligned}$$

wobei  $\Delta_j = t_{j+1} - t_j$ .

## 4.2 Das Stratonovich'sche stochastische Integral

Das Ito-Integral genügt auch der komischen Eigenschaft

$$\int_0^T W_t dW_t = \frac{1}{2}W_T^2 - \frac{1}{2}T \quad (W_0 = 0 \text{ hier})$$

verglichen mit dem deterministischen Kalkül, wo

$$\begin{aligned} \int_{u(0)}^{u(T)} u du &= \int_0^T u(t)u'(t) dt, & u' \text{ Ableitung} \\ &= \frac{1}{2}u(T)^2 - \frac{1}{2}u(0)^2 \\ &= \frac{1}{2}u(T)^2 \quad \text{falls } u(0) = 0. \end{aligned}$$

Dasselbe Ergebnis gilt auch in dem stochastischen Fall bezüglich der Auswertungsstellen  $\bar{t}_j = t_j + \frac{1}{2}(t_{j+1} - t_j) \equiv \frac{1}{2}(t_j + t_{j+1})$ .

**Definition:** Sei  $f$  von der obigen Klasse. Das Stratonovich-Integral

von  $f$  ist definiert durch

$$\int_0^T f_t \circ dW_t = \text{qm} - \lim_{N \rightarrow \infty} \sum_{j=0}^{N-1} f\left(t_j + \frac{1}{2}(t_{j+1} - t_j)\right) \{W_{t_{j+1}} - W_{t_j}\}$$

Die Limeszufallsvariable  $\int_0^T f_t \circ dW_t$  hängt nicht von der Folge der Unterteilungen des Intervalls  $[0, T]$  ab.

Wir haben

$$\int_0^T W_t \circ dW_t = \frac{1}{2}W_T^2.$$

Das Stratonovich-Integral genügt auch der Linearitätseigenschaft

$$\int_0^T (\alpha f_t + \beta g_t) \circ dW_t = \alpha \int_0^T f_t \circ dW_t + \beta \int_0^T g_t \circ dW_t$$

Aber die anderen günstigen Eigenschaften des Ito-Integrals gelten im Allgemeinen nicht, d.h. im Allgemeinen

$$\begin{aligned} \mathbb{E} \left( \int_0^T f_t \circ dW_t \right) &\neq 0 \\ \mathbb{E} \left( \left( \int_0^T f_t \circ dW_t \right)^2 \right) &\neq \int_0^T \mathbb{E} (f_t^2) dt, \end{aligned}$$

z.B. (speziell für  $f_t = W_t$ )

$$\begin{aligned}\mathbb{E}\left(\int_0^T W_t \circ dW_t\right) &= \mathbb{E}\left(\frac{1}{2}W_T^2\right) = \frac{1}{2}T \neq 0 \\ \mathbb{E}\left(\left(\int_0^T W_t \circ dW_t\right)^2\right) &= \mathbb{E}\left(\frac{1}{4}W_T^4\right) = \frac{3}{4}T^2 \neq \frac{1}{2}T^2 = \int_0^T t dt = \int_0^T \mathbb{E}(W_t^2) dt.\end{aligned}$$

Im Gegensatz dazu, genügt das Ito-Integral

$$\begin{aligned}\mathbb{E}\left(\int_0^T W_t dW_t\right) &= \mathbb{E}\left(\frac{1}{2}W_T^2 - \frac{1}{2}T\right) = \mathbb{E}\left(\frac{1}{2}W_T^2\right) - \frac{1}{2} = \frac{1}{2}T - \frac{1}{2}T = 0 \\ \mathbb{E}\left(\left(\int_0^T W_t dW_t\right)^2\right) &= \mathbb{E}\left(\left(\frac{1}{2}W_T^2 - \frac{1}{2}T\right)^2\right) \\ &= \frac{1}{4}\left(\mathbb{E}(W_T^4) - 2T\mathbb{E}(W_T^2) + T^2\right) \\ &= \frac{1}{4}(3T^2 - 2T \cdot T + T^2) \\ &= \frac{1}{2}T^2 = \int_0^T t dt = \int_0^T \mathbb{E}(W_t^2) dt.\end{aligned}$$

## Kapitel 5

# Stochastische Differentialgleichungen

Stochastische Differentialgleichungen (SDGLen) sind eigentlich Integralgleichungen – die Differentialdarstellung ist nur symbolisch.

Es gibt zwei Arten von SDGLen

- 1) Ito'sche SDGLen mit dem Ito-Integral

$$\begin{cases} dX_t = f(t, X_t)dt + g(t, X_t)dW_t \\ X_t = X_{t_0} + \int_{t_0}^t f(s, X_s)ds + \int_{t_0}^t g(s, X_s)dW_s \end{cases}$$

- 2) Stratonovich'sche SDGLen mit dem Stratonovich-Integral ( $\circ$  = Stratonovich)

$$\begin{cases} dX_t = f(t, X_t)dt + g(t, X_t) \circ dW_t \\ X_t = X_{t_0} + \int_{t_0}^t f(s, X_s)ds + \int_{t_0}^t g(s, X_s) \circ dW_s \end{cases}$$

Eine Lösung  $X_t$  ist ein stochastischer Prozess, welcher der stochastischen Integralgleichung der SDGL genügt, mit folgenden Eigenschaften

- (i)  $X_t$  ist eine Zufallsvariable mit  $\mathbb{E}(X_t^2) < \infty$  für jedes  $t \in [t_0, T]$
- (ii)  $X_t$  ist nichtantizipativ bzgl. des Wiener Prozesses  $W_t$
- (iii) Die Pfade  $t \rightarrow X_t(\omega)$  sind stetig (f.s.)

### Bemerkungen

Die Lösungen einer Ito'schen SDGL und der Stratonovich'schen SDGL mit denselben Koeffizienten sind nicht immer identisch, z.B.

(I) die Ito-SDGL  $dX_t = aX_t dt + bX_t dW_t$  besitzt die Lösung

$$X_t = X_0 e^{(a - \frac{1}{2}b^2)t + bW_t}$$

(II) Die Stratonovich-SDGL  $dX_t = aX_t dt + bX_t \circ dW_t$  besitzt die Lösung

$$X_t = X_0 e^{at + bW_t}$$

$\Rightarrow$  das Verhalten der Lösungen kann ganz anders sein!

Frage Sollen wir die Ito'sche oder Stratonovich'sche Version einer SDGL benutzen?

Die Antwort hat mehr mit Modellierung als mit Mathematik zu tun – im Allgemeinen gibt's nur Daumenregeln.

Aber die Lösungen der Ito-SDGL

$$dX_t = f(t, X_t) dt + g(t, X_t) dW_t$$

und der modifizierten Stratonovich-SDGL

$$dX_t = \underline{f}(t, X_t) dt + g(t, X_t) \circ dW_t$$

wobei

$$\underline{f}(t, x) \equiv f(t, x) - \frac{1}{2}g(t, x) \frac{\partial g}{\partial x}(t, x)$$

sind identisch!

Beispiel

$$\begin{cases} dX_t = aX_t dt + bX_t dW_t & \text{Ito} \\ dX_t = (a - \frac{1}{2}b^2) X_t dt + bX_t \circ dW_t & \text{Stratonovich} \end{cases}$$

haben dieselbe Lösung

$$X_t = X_0 e^{(a - \frac{1}{2}b^2)t + bW_t}.$$

Beispiel Wir sagen, dass das Rauschen additiv ist, falls

$$\frac{\partial g}{\partial x}(t, x) \equiv 0.$$

In diesem Fall gilt  $\underline{f}(t, x) \equiv f(t, x)$ . Dann haben die Ito-SDGL und die Stratonovich-SDGL mit denselben Koeffizienten dieselben Lösungen.

Das Rauschen heißt multiplikativ, wenn  $\frac{\partial g}{\partial x}(t, x) \neq 0$  ist.

Bemerkungen Die beiden Versionen der stochastischen Integralrechnung haben ihre Vorteile und Nachteile – aber wir können immer von einer in die andere wechseln, um die Vorteile auszubenten.

z.B.

- die Ito-Version ist günstiger für Abschätzungen und Beweise wegen der Sondereigenschaften des Ito-Integrals.
- die Stratonovich-Version ist günstiger, wenn wir explizite Lösungen finden wollen – die Stratonovich-Kettenregel (aber nicht die Ito'sche) ist genau wie in der deterministischen Integral-/Differentialrechnung.

## 5.1 Der Existenz- und Eindeutigkeitssatz

1. Die Koeffizienten  $f, g : [t_0, T] \times \mathbb{R} \rightarrow \mathbb{R}$  genügen den folgenden Eigenschaften

(i) stetig in  $(t, x)$

(ii) Linearwachstum

$$|f(t, x)| \leq K(1 + |x|)$$

$$|g(t, x)| \leq K(1 + |x|)$$

(iii) Lipschitz-Bedingung

$$|f(t, x) - f(t, y)| \leq L|x - y|$$

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

gleichmäßig in  $t \in [t_0, T]$

2. Die Anfangszufallsvariable  $X_{t_0}$  sei nichtantizipativ bzgl. des Wiener Prozesses  $W_t$  mit

$$\mathbb{E}(X_{t_0}^2) < \infty$$

Unter diesen beiden Voraussetzungen besitzt die Ito-SDGL

$$dX_t = f(t, X_t) dt + g(t, X_t) dW_t$$

eine eindeutige Lösung auf  $[t_0, T]$  mit Anfangswert  $X_{t_0}$ .

Idee des Beweises Man zeigt, dass die sukzessiven Approximationen

$$X_t^{(n+1)} = X_{t_0} + \int_{t_0}^t f(s, X_s^{(n)}) ds + \int_{t_0}^t g(s, X_s^{(n)}) dW_s, \quad n = 0, 1, 2, \dots,$$

mit  $X_t^{(0)} \equiv X_{t_0}$  einen eindeutigen qm-Grenzwert  $\bar{X}_t$  besitzen und dann (schwieriger!), dass  $\bar{X}_t$  die behaupteten Eigenschaften besitzt. Fixpunktsatz!

Ein großer Unterschied zum deterministischen Fall ist die Benutzung der Lipschitz-Bedingung in dem Ito-Integral. Statt

$$\begin{aligned} \left| \int_{t_0}^t f(s, X_s) ds - \int_{t_0}^t f(s, Y_s) ds \right| &\leq \int_{t_0}^t |f(s, X_s) - f(s, Y_s)| ds && \text{(für } t_0 \leq t) \\ &\leq \int_{t_0}^t L |X_s - Y_s| ds && \text{Lipschitz-Bedingung} \end{aligned}$$

für den Drift-Term, müssen wir die folgende Abschätzung benutzen:

$$\begin{aligned} &\mathbb{E} \left( \left| \int_{t_0}^t g(s, X_s) dW_s - \int_{t_0}^t g(s, Y_s) dW_s \right|^2 \right) \\ &= \mathbb{E} \left( \left| \int_{t_0}^t (g(s, X_s) - g(s, Y_s)) dW_s \right|^2 \right) && \text{Linearität} \\ &= \int_{t_0}^t \mathbb{E} \left( |g(s, X_s) - g(s, Y_s)|^2 \right) ds && \text{Varianz/Isometrie des Ito-Integrals} \\ &\leq \int_{t_0}^t L^2 \cdot \mathbb{E} \left( |X_s - Y_s|^2 \right) ds && \text{Lipschitz-Bedingung} \end{aligned}$$

Bemerkungen Die Voraussetzungen des Satzes sind ziemlich streng, z.B. die Linearwachstumsbedingungen. Schwächere Voraussetzungen sind auch möglich!

## Kapitel 6

# Das stochastische Euler-Verfahren

Die Lösung  $x(t) = x(t, t_0, x_0)$  der deterministischen Anfangswertaufgabe

$$\frac{dx}{dt} = f(t, x), \quad x(t_0) = x_0$$

genügt der Integralgleichung

$$x(t) = x_0 + \int_{t_0}^t f(s, x(s)) ds$$

auf dem Intervall  $[t_0, T]$ , sowie der Integralgleichung

$$x(t_{n+1}) = x(t_n) + \int_{t_n}^{t_{n+1}} f(s, x(s)) ds$$

auf jedem Teilintervall  $[t_n, t_{n+1}]$  von  $[t_0, T]$ . Daraus erhalten wir die Approximation

$$x(t_{n+1}) \approx x(t_n) + \int_{t_n}^{t_{n+1}} \underbrace{f(t_n, x(t_n))}_{\substack{\text{Auswertungsstelle} \\ s = t_n}} ds$$

$$\begin{aligned} \text{d.h.} \quad x(t_{n+1}) &\approx x(t_n) + f(t_n, x(t_n)) \int_{t_n}^{t_{n+1}} 1 ds \\ &= x(t_n) + f(t_n, x(t_n)) \cdot \Delta_n, \quad n = 0, 1, \dots, N-1 \end{aligned}$$

wobei  $\Delta_n = t_{n+1} - t_n = \int_{t_n}^{t_{n+1}} 1 ds$ .

Dies ist die Motivierung für das Euler-Verfahren

$$\boxed{x_{n+1} = x_n + f(t_n, x_n) \Delta_n} \quad n = 0, 1, \dots, N-1$$

bezüglich einer Unterteilung  $t_0 < t_1 < \dots < t_n < \dots < t_N = T$  des Intervalls  $[t_0, T]$ . Hier soll  $x_n \approx x(t_n)$  sein.

Der globale Diskretisierungsfehler genügt der Abschätzung

$$|x(t_n) - x_n| \leq K_T \Delta,$$

wobei  $\Delta = \max_n \Delta_n > 0$  ist, für die Klasse von stetig differenzierbaren Vektorfeldfunktionen  $f$ . Deshalb besitzt das deterministische Euler-Verfahren die Ordnung  $p = 1$ .

Betrachte jetzt eine Ito-SDGL

$$dX_t = f(t, X_t) dt + g(t, X_t) dW_t$$

auf dem Intervall  $[t_0, T]$  mit Anfangswert  $X_{t_0}$ . Die Lösung  $X_t$  genügt der stochastischen Ito-Integralgleichung

$$X_t = X_{t_0} + \int_{t_0}^t f(s, X_s) ds + \int_{t_0}^t g(s, X_s) dW_s$$

auf dem Intervall  $[t_0, T]$ , sowie der Integralgleichung

$$X_{t_{n+1}} = X_{t_n} + \int_{t_n}^{t_{n+1}} f(s, X_s) ds + \int_{t_n}^{t_{n+1}} g(s, X_s) dW_s$$

auf jedem Teilintervall  $[t_n, t_{n+1}]$  von  $[t_0, T]$ .

Wie im deterministischen Fall „einfrieren“ wir die Integrandenfunktionen  $f$  und  $g$  zu der Auswertungsstelle  $s = t_n$ . Wir erhalten die Approximation

$$\begin{aligned} X_{t_{n+1}} &\approx X_{t_n} + \int_{t_n}^{t_{n+1}} f(t_n, X_{t_n}) ds + \int_{t_n}^{t_{n+1}} g(t_n, X_{t_n}) dW_s \\ &= X_{t_n} + f(t_n, X_{t_n}) \int_{t_n}^{t_{n+1}} 1 ds + g(t_n, X_{t_n}) \int_{t_n}^{t_{n+1}} 1 dW_s \\ &= X_{t_n} + f(t_n, X_{t_n}) \Delta_n + g(t_n, X_{t_n}) \Delta W_n, \end{aligned}$$

wobei

$$\begin{aligned} \Delta_n &= t_{n+1} - t_n = \int_{t_n}^{t_{n+1}} 1 \cdot ds \\ \Delta W_n &= W_{t_{n+1}} - W_{t_n} = \int_{t_n}^{t_{n+1}} 1 \cdot dW_s \end{aligned}$$

Diese Approximation scheint konsistent mit der Definition des Ito-Integrals zu sein!

Daraus erhalten wir das stochastische Euler-Verfahren

$$\boxed{X_{n+1} = X_n + f(t_n, X_n)\Delta_n + g(t_n, X_n) \cdot \Delta W_n} \quad n = 0, 1, \dots, N - 1$$

Dieses Verfahren heißt auch Euler-Maruyama-Verfahren. Hier soll  $X_n \approx X_{t_n}$  in einem geeigneten Sinne sein – aber welcher Sinn?

Die Vorschrift des Euler-Maruyama-Verfahrens ist eine stochastische Differenzengleichung, die eine Zufallsvariable  $X_n$  in eine neue Zufallsvariable  $X_{n+1}$  abbildet. Aber in einem Computer können wir nur einzelne Realisierungen einer Zufallsvariablen bzw. einzelne Pfade eines (zeitdiskreten) stochastischen Prozesses berechnen, d.h.  $X_n(\omega) \rightarrow X_{n+1}(\omega)$  für festes  $\omega$ . Dafür brauchen wir die entsprechenden Zuwächse  $\Delta W_n(\omega)$  des Pfades  $W_t(\omega)$  des Wiener Prozesses  $W_t$ . Die Berechnungen laufen rekursiv.

$$\begin{array}{ccccccc} \Delta W_0(\omega) & & \Delta W_1(\omega) & & \Delta W_2(\omega) & & \dots \\ & & \searrow & & \searrow & & \searrow \\ & & & & & & \dots \\ X_0(\omega) & \longrightarrow & X_1(\omega) & \longrightarrow & X_2(\omega) & \longrightarrow & X_3(\omega) \dots \end{array}$$

Bemerkung

$$\boxed{W_{t_{n+1}} = W_{t_0} + \sum_{j=0}^n \Delta W_j}$$

Wir simulieren die Zuwachsrealisierungen  $\Delta W_n(\omega)$  durch einen Pseudo-Zufallsvariablen-Erzeuger (PZVE) und die Box-Muller-Methode.

- 1) PZVE Jeder Computer besitzt einen eingebauten PZVE, z.B. namens RAN, RAND, usw (random= zufällig auf Englisch)  
Betrachte 2 sukzessive Verwendungen von RAN

$$RAN \rightarrow U_1, \quad RAN \rightarrow U_2$$

$U_1$  und  $U_2$  sind (sollen sein!) Realisierungen von 2 unabhängigen gleichmäßig auf  $[0, 1]$  verteilten Zufallsvariablen.

- 2) Box-Muller-Methode Berechne

$$G_1 = \sqrt{-2 \ln(U_1)} \cdot \cos(2\pi U_2)$$

$$G_2 = \sqrt{-2 \ln(U_1)} \cdot \sin(2\pi U_2)$$

$G_1$  und  $G_2$  sind (sollen sein!) Realisierungen von 2 unabhängigen  $N(0, 1)$ -verteilten Zufallsvariablen.

3) Zuwächse Berechne

$$\Delta W_{2n}(\omega) = G_1 \sqrt{\Delta_{2n}}, \quad \Delta W_{2n+1}(\omega) = G_2 \sqrt{\Delta_{2n+1}}$$

d.h. Realisierungen der unabhängigen Zuwächse  $\Delta W_{2n} \sim N(0, \Delta_{2n})$  und  $\Delta W_{2n+1} \sim N(0, \Delta_{2n+1})$  auf den Teilintervallen  $[t_{2n}, t_{2n+1}]$  und  $[t_{2n+1}, t_{2n+2}]$  für dasselbe  $\omega$  (für denselben Pfad  $W_t(\omega)$  des Wiener Prozesses).

N.B. Box-Muller  $\Rightarrow$  aufwendig wegen  $\sqrt{\cdot}$ ,  $\ln$ ,  $\cos$ ,  $\sin$  (Polar-Marsaglia besser?)

Wir können den numerischen Pfad  $X_j(\omega)$ ,  $j = 0, 1, \dots, N$ , besser durch einen Polygonzugprozess -Pfad veranschaulichen. Definiere  $X_t^{(\Delta)}$ ,  $t \in [t_0, T]$ ,  $\Delta = \max_n \Delta_n$  pfadweise durch

$$X_t^{(\Delta)}(\omega) = X_n(\omega) + \frac{t - t_n}{\Delta_n} \cdot [X_{n+1}(\omega) - X_n(\omega)] \text{ für } t \in [t_n, t_{n+1}].$$

Für das Euler-Maruyama-Verfahren gilt

$$\begin{aligned} X_t^{(\Delta)}(\omega) &= X_n(\omega) + \frac{t - t_n}{\Delta_n} [f(t_n, X_n(\omega)) \cdot \Delta_n + g(t_n, X_n(\omega)) \Delta W_n(\omega)] \\ &= X_n(\omega) + f(t_n, X_n(\omega)) \cdot (t - t_n) + g(t_n, X_n(\omega)) \cdot \frac{t - t_n}{\Delta_n} \cdot \Delta W_n(\omega) \end{aligned}$$

Bemerkung Dieser stochastische Prozess ist nicht nichtantizipativ (d.h. er ist antizipativ bzgl.  $W_t$ ) wegen  $\Delta W_n = W_{t_{n+1}} - W_{t_n}$ .

Frage: Wie können wir denselben Pfad (d.h. mit demselben  $\omega$ ) für eine feinere Unterteilung von  $[t_0, T]$  berechnen?

z.B mit halbierter Schrittweite:

- alte Unterteilung:  $t_0 < t_1 < \dots < t_n < t_{n+1} < \dots < t_N = T$
- neue Unterteilung:  $t'_0 = t_0 < t'_1 < t'_2 < \dots < t'_{2N} = T$

$$\text{mit } \begin{cases} t'_{2n} = t_n \\ t'_{2n+1} = t_n + \frac{1}{2}(t_{n+1} - t_n) \end{cases}$$

Dann können wir die Lévy-Konstruktion verwenden, um einen geeigneten Wert von  $W_{t'_{2n+1}}(\omega)$  zu berechnen, wenn  $W_{t_n}(\omega)$  und  $W_{t_{n+1}}(\omega)$  bekannt sind.

Definiere

$$W_{t'_{2n+1}}(\omega) = \frac{1}{2} \{W_{t_n}(\omega) + W_{t_{n+1}}(\omega)\} + \frac{\sqrt{\Delta_n}}{2} \epsilon_n(\omega)$$

wobei  $\epsilon_n \sim N(0, 1)$  ist – und natürlich sind die verschiedenen  $\epsilon_n$  unabhängig.

Beweis-Hinweis Die neuen Zuwächse

$$\Delta'W_{2n} = W_{t'_{2n+1}} - W_{t'_{2n}} = W_{t'_{2n+1}} - W_{t_n} = \frac{\sqrt{\Delta_n}}{2} \epsilon_n + \frac{1}{2} \Delta W_n$$

$$\Delta'W_{2n+1} = W_{t'_{2n+2}} - W_{t'_{2n+1}} = W_{t_{n+1}} - W_{t'_{2n+1}} = -\frac{\sqrt{\Delta_n}}{2} \epsilon_n + \frac{1}{2} \Delta W_n$$

sind unabhängig und  $N(0, \Delta_n/2)$  verteilt mit

$$\Delta'W_{2n} + \Delta'W_{2n+1} = \Delta W_n$$

NB

$$\begin{aligned} \text{Var}(\Delta'W_{2n} + \Delta'W_{2n+1}) &= \text{Var}(\Delta'W_{2n}) + \text{Var}(\Delta'W_{2n+1}) \quad \text{unabhängig !} \\ &= \Delta_n/2 + \Delta_n/2 = \Delta_n = \text{Var}(\Delta W_n). \end{aligned}$$

# Kapitel 7

## Starke und schwache Konvergenz

Betrachte eine Ito-SDG

$$dX_t = f(t, X_t)dt + g(t, X_t) dW_t$$

auf einem Intervall  $[t_0, T]$  mit Lösung  $X_t$ .

Betrachte auch eine Unterteilung  $t_0 < t_1 < \dots < t_n < \dots < t_N = T$  des Intervalls  $[t_0, T]$  mit Schrittweiten  $\Delta_n = t_{n+1} - t_n > 0, n = 0, 1, \dots, N - 1$  und maximaler Schrittweite

$$\Delta = \max_n \Delta_n$$

Das Euler-Maruyama-Verfahren hier lautet

$$X_{n+1} = X_n + f(t_n, X_n) \Delta_n + g(t_n, X_n) \Delta W_n,$$

wobei die Zuwächse  $\Delta W_n \sim N(0, \Delta_n)$  und unabhängig sind.

Die Vorschrift hier ist eine Differenzgleichung für Zufallsvariablen. In der Praxis können wir nur einzelne Realisierungen von Zufallsvariablen berechnen, d.h.

$$X_{n+1}(\omega) = X_n(\omega) + f(t_n, X_n(\omega)) \Delta_n + g(t_n, X_n(\omega)) \Delta W_n(\omega),$$

$n = 0, 1, \dots$ , für ein festes  $\omega$ . (Dann können/müssen wir die Berechnungen für ein neues  $\omega$  wiederholen). Dafür simulieren wir die  $\Delta W_n(\omega)$  durch PZVE und die Box-Muller-Methode.

Wir hoffen, dass die  $X_n(\omega)$  irgendwie die  $X_{t_n}(\omega)$  approximieren. Aber, in welchem Sinn?

Beispiel: Sei  $X$  eine 2-Punkt-Zufallsvariable mit

$$P(\{\omega \in \Omega : X(\omega) = -1\}) = 1/2, \quad P(\{\omega \in \Omega : X(\omega) = +1\}) = 1/2$$

und definiere  $Y(\omega) \equiv -X(\omega)$  für alle  $\omega \in \Omega$ . Dann sind die Zufallsvariablen  $X$  und  $Y$  identisch verteilt, aber

$$|X(\omega) - Y(\omega)| = 2, \quad \forall \omega \in \Omega.$$

Was wollen wir approximieren?

Für einige Anwendungen sollen die Realisierungen nah sein. Für andere reicht es, dass nur die Wahrscheinlichkeiten verschiedener Ereignisse (oder die Verteilungen) nah sind.

Wir werden 2 Arten stochastischer Approximationen unterscheiden:

starke Approximationen und schwache Approximationen.

## 7.1 Starke Approximationen

Im Prinzip können wir den Fehler  $|X_n(\omega) - X_{t_n}(\omega)|$  für jedes  $\omega \in \Omega$  betrachten.

Aber dann sind die Fehlerabschätzungen technisch sehr kompliziert (vielleicht unmöglich!) und viele Fälle passieren fast nie oder nur mit geringer Wahrscheinlichkeit.

In der Praxis ist der qm-Fehler (qm = quadrat-mittel)

$$\mathbb{E}(|X_n - X_{t_n}|^2)$$

günstiger, z.B Abschätzungen sind technisch leichter durchzuführen.

Aber: wir werden der erste Moment

$$\mathbb{E}(|X_n - X_{t_n}|)$$

benutzen, weil es die natürliche Verallgemeinerung des deterministischen Diskretisierungsfehlers ist. Wir können ihn auch durch den qm-Fehler abschätzen:

$$\mathbb{E}(|X_n - X_{t_n}|) \leq \sqrt{\mathbb{E}(|X_n - X_{t_n}|^2)}$$

Schreibe  $X_n^{(\Delta)}$  für eine numerische Approximation mit maximaler Schrittweite  $\Delta$ .

Wir definieren die starke Konvergenz solcher numerischen Approximationen durch

$$\max_n \mathbb{E}(|X_n^{(\Delta)} - X_{t_n}|) \rightarrow 0 \text{ für } \Delta \rightarrow 0.$$

Wir sagen, dass ein numerisches Verfahren starke Ordnung  $\gamma > 0$  besitzt, falls

$$\max_n \mathbb{E} \left( \left| X_n^{(\Delta)} - X_{t_n} \right| \right) \leq K_T \Delta^\gamma,$$

für eine Konstante, die von  $T$  abhängt.

Das Euler-Maruyama-Verfahren besitzt starke Ordnung

$$\boxed{\gamma = 1/2}$$

bezüglich der Klasse von Ito-SDGLen mit Koeffizienten  $f$  und  $g$  wie in dem Existenz/Eindeutigkeitssatz.

Verglichen mit dem deterministischen Fall, ist der Beweis schwieriger, weil die Lösungen nicht pfadweise gleichmäßig beschränkt sind, d.h. wir können keine gemeinsame kompakte Menge für z.B. lokale Lipschitz-Bedingungen benutzen.

Hinweis:

$$\mathbb{E} \left( (\Delta W_n)^2 \right) = \Delta_n \quad \Rightarrow \quad \mathbb{E} (|\Delta W_n|) \sim \sqrt{\Delta_n}$$

## 7.2 Schwache Approximationen

Jetzt wollen wir die Verteilungen von  $X_n^{(\Delta)}$  und  $X_{t_n}$  vergleichen, oder die entsprechenden Wahrscheinlichkeiten für verschiedene Ereignisse. Diese sind meistens explizit nicht bekannt – aber die Momente sind oft bekannt.

Wir definieren die schwache Konvergenz einer numerischen Approximation  $X_n^{(\Delta)}$  durch

$$\max_n \left| \mathbb{E} \left( P \left( X_n^{(\Delta)} \right) \right) - \mathbb{E} \left( P \left( X_{t_n} \right) \right) \right| \rightarrow 0$$

für  $\Delta \rightarrow 0$ , für alle Polynome  $P$  (oder von einer anderen geeigneten Klasse von Testfunktionen).

Wir sagen, dass ein numerisches Verfahren die schwache Ordnung  $\beta > 0$  besitzt, falls

$$\max_n \left| \mathbb{E} \left( P \left( X_n^{(\Delta)} \right) \right) - \mathbb{E} \left( P \left( X_{t_n} \right) \right) \right| \leq K_{P,T} \cdot \Delta^\beta$$

wobei hier die Konstante auch von  $P$  abhängt.

Das Euler-Maruyama-Verfahren besitzt schwache Ordnung

$$\boxed{\beta = 1}$$

bzgl. der obigen Klasse von Ito-SDGLen.

Bemerkung: Schwache Konvergenz ist tatsächlich schwächer als starke Konvergenz.

Dies ist leicht zu sehen, wenn  $P$  einer globalen Lipschitz-Bedingungen genügt, z.B.  $P(x) = x$ . Dann gilt

$$\begin{aligned} |\mathbb{E}(P(X_n)) - \mathbb{E}(P(X_{t_n}))| &= |\mathbb{E}(P(X_n) - P(X_{t_n}))| \\ &\leq \mathbb{E}(|P(X_n) - P(X_{t_n})|) \\ &\leq \mathbb{E}(L_P |X_n - X_{t_n}|) = L_P \mathbb{E}(|X_n - X_{t_n}|) \end{aligned}$$

$\Rightarrow$  ein Verfahren konvergiert schwach, wenn es stark konvergiert.

Umgekehrt – meistens falsch!

z.B. Sei  $X_n \equiv -X_{t_n}$  mit  $\mathbb{E}(X_n) = 0$  and  $E(|X_n|) \neq 0$ . Dann haben wir

$$|\mathbb{E}(X_n) - \mathbb{E}(X_{t_n})| = |0 - 0| = 0$$

aber

$$\mathbb{E}(|X_n - X_{t_n}|) = 2\mathbb{E}(|X_n|) \neq 0.$$

### 7.2.1 Das vereinfachte Euler-Verfahren

Für schwache Konvergenz können wir die 2-Punkt-Zufallsvariablen  $\widetilde{\Delta W}_n$  mit

$$P\left(\left\{\omega \in \Omega : \widetilde{\Delta W}_n(\omega) = -\sqrt{\Delta_n}\right\}\right) = 1/2$$

$$P\left(\left\{\omega \in \Omega : \widetilde{\Delta W}_n(\omega) = +\sqrt{\Delta_n}\right\}\right) = 1/2$$

statt Gauß-verteilter  $\Delta W_n \sim N(0, \Delta_n)$  in dem Euler-Verfahren benutzen.

Dann heißt das Euler-Verfahren das vereinfachte Euler-Verfahren und lautet

$$X_{n+1} = X_n + f(t_n, X_n) \Delta_n + g(t_n, X_n) \widetilde{\Delta W}_n$$

Dieses Verfahren besitzt auch die schwache Ordnung  $\beta = 1$ , aber konvergiert nicht in dem starken Sinne.

Bemerkung Die ersten 3 Momente von  $\Delta W_n$  und  $\widetilde{\Delta W}_n$  sind identisch – diese Tatsache reicht für eine schwache Approximation erster Ordnung.

$$\mathbb{E}\left(\widetilde{\Delta W}_n\right) = \mathbb{E}(\Delta W_n) = 0$$

$$\mathbb{E} \left( \left( \widetilde{\Delta W}_n \right)^2 \right) = \mathbb{E} \left( (\Delta W_n)^2 \right) = \Delta_n$$

$$\mathbb{E} \left( \left( \widetilde{\Delta W}_n \right)^3 \right) = \mathbb{E} \left( (\Delta W_n)^3 \right) = 0$$

aber  $\mathbb{E} \left( \left( \widetilde{\Delta W}_n \right)^4 \right) = \Delta_n^2 \neq \mathbb{E} \left( (\Delta W_n)^4 \right) = 3\Delta_n^2.$

Bemerkung Die 2-Punkt-Zufallsvariablen  $\widetilde{\Delta W}_n$  sind sehr leicht (und billig!) zu simulieren.

(1) PZVE  $\rightarrow U$  auf  $[0, 1]$  gleichmäßig verteilt

(2)

$$\begin{cases} U \in (0, 1/2) & \Rightarrow \widetilde{\Delta W}_n(\omega) = -\sqrt{\Delta_n} \\ U \in [1/2, 1] & \Rightarrow \widetilde{\Delta W}_n(\omega) = +\sqrt{\Delta_n} \end{cases}$$

d.h. wir müssen keine aufwendigen Funktionen  $\sqrt{\cdot}, \ln, \cos, \sin$  wie in der Box-Muller-Methode auswerten.

## 7.2.2 Funktionalschätzung

Oft wollen wir ein Funktional einer Lösung  $X_t$  einer Ito-SDG schätzen, z.B.

$$\mathbb{E}(q(X_T)),$$

wobei  $q: \mathbb{R} \rightarrow \mathbb{R}$  eine gegebene Funktion ist und  $X_t$  die Lösung einer Ito-SDGL auf einem Intervall  $[t_0, T]$ .

Wir können ein schwach konvergierendes numerisches Verfahren benutzen, um  $X_N^{(\Delta)}(\omega)$  für  $\omega = \omega_1, \dots, \omega_K$  zu berechnen. Dann ist das arithmetische Mittel

$$\hat{q} = \frac{1}{K} \sum_{j=1}^K q \left( X_N^{(\Delta)}(\omega_j) \right)$$

eine Schätzung für  $\mathbb{E} \left( q \left( X_N^{(\Delta)} \right) \right)$ .

Das arithmetische Mittel  $\hat{q}$  ist selbst zufällig, aber wir können ein Konfidenz-Intervall der Form

$$[\hat{q} - r_{\Delta, K, \alpha}, \hat{q} + r_{\Delta, K, \alpha}]$$

finden. Der genaue Wert  $\mathbb{E} \left( q \left( X_N^{(\Delta)} \right) \right)$  liegt in diesem Intervall mit Wahrscheinlichkeit  $1 - \alpha$  und der erwünschte Wert

$$\mathbb{E}(q(X_T)) = \mathbb{E} \left( q \left( X_N^{(\Delta)} \right) \right) + O(\Delta^\beta),$$

wobei  $\beta$  die schwache Ordnung des numerischen Verfahrens ist.

Aber  $K$  muss sehr groß sein, wenn  $\alpha$  und  $r_{\Delta, K, \alpha}$  klein sein sollen.

### 7.3 Konsistenz

Die starke Ordnung  $\gamma = 1/2$  und schwache Ordnung  $\beta = 1$  des stochastischen Euler-Verfahrens sind nicht sehr günstig, weil wir viele verschiedene Pfade berechnen müssen. Daher brauchen wir stochastische numerische Verfahren höherer Ordnung. Aber wie können wir solche Verfahren herleiten? Angepasste Versionen deterministischer Verfahren (z.B. des Heun-Verfahrens) konvergieren entweder nicht oder nur mit niedriger Ordnung.

Beispiel Das Heun-Verfahren für die Ito-SDGL

$$dX_t = f(X_t) dt + g(X_t) dW_t$$

lautet

$$\begin{aligned} X_{n+1} &= X_n + \frac{1}{2} [f(X_n) + f(X_n + f(X_n)\Delta_n + g(X_n)\Delta W_n)] \Delta_n \\ &\quad + \frac{1}{2} [g(X_n) + g(X_n + f(X_n)\Delta_n + g(X_n)\Delta W_n)] \Delta W_n \end{aligned}$$

Für multiplikatives Rauschen, d.h.  $g(x) \not\equiv 0$ , konvergiert das Heun-Verfahren weder stark noch schwach. Dies ist keine Überraschung, weil der stochastische Term des Heun-Verfahrens offensichtlich nicht konsistent mit den Definitionen des Ito-Integrals ist, z.B.

$$\begin{cases} dX_t = 2X_t dW_t & f(x) \equiv 0, \quad g(x) \equiv 2x \\ X_0 = 1 \end{cases}$$

mit der "Lösung"  $X_t = 1 + \int_0^t 2X_s dW_s$ , für welche gilt

$$\mathbb{E}(X_t) = 1 + \mathbb{E}\left(\int_0^t 2X_s dW_s\right) = 1 + 0 = 1 \quad \forall t \geq 0, \quad \text{Ito-Integral!}$$

Das Heun-Verfahren hier lautet

$$\begin{aligned} X_{n+1} &= X_n + \frac{1}{2} [2X_n + 2(X_n + 2X_n \Delta W_n)] \Delta W_n \\ &= X_n [1 + 2\Delta W_n + 2(\Delta W_n)^2] \end{aligned}$$

$$\Rightarrow X_n = \prod_{j=0}^{n-1} (1 + 2\Delta W_j + 2(\Delta W_j)^2) \quad \text{unabhängige ZVen} \quad (X_0 \equiv 1)$$

$$\begin{aligned}
\Rightarrow \mathbb{E}(X_n) &= \prod_{j=0}^{n-1} \mathbb{E}(1 + 2\Delta W_j + 2(\Delta W_j)^2) \\
&= \prod_{j=0}^{n-1} (1 + 2\mathbb{E}(\Delta W_j) + 2(\mathbb{E}(\Delta W_j)^2)) \\
&= \prod_{j=0}^{n-1} (1 + 2\Delta) \\
&= (1 + 2\Delta)^n \\
&= e^{2n\Delta} + O(\Delta) = e^{2t_n} + O(\Delta), \quad t_n = n\Delta,
\end{aligned}$$

(Betrachte den globalen Diskretisierungsfehler des Euler-Verfahrens für die deterministische DGL  $x' = 2x$ .)

aber  $\mathbb{E}(X_{t_n}) = 1 \neq e^{2t_n} \Rightarrow$  Das Heun-Verfahren konvergiert nicht schwach (und deshalb auch nicht stark) in diesem Beispiel.

Für additives Rauschen, d.h.  $g'(x) \equiv 0$ , konvergiert das Heun-Verfahren mit Ordnungen  $\gamma = \beta = 1$ .

Es geht nicht besser mit anderen Runge-Kutta-Verfahren: entweder keine Konvergenz oder Konvergenz nur mit Ordnungen  $\gamma, \beta < 2$ .

Ein stochastisches numerisches Verfahren höherer Ordnung soll mindestens die entsprechenden statistischen Eigenschaften des Euler-Verfahrens besitzen. Diese Beobachtung motiviert die folgenden Definitionen starker und schwacher Konsistenz.

Betrachte eine Ito-SDGL

$$dX_t = f(t, X_t) dt + g(t, X_t) dW_t$$

Schreibe  $X_1^{(\Delta, t, X)}$  für den ersten Schritt eines numerischen Verfahrens mit Anfangsbedingungen  $(t, X)$  und Schrittweite  $\Delta$ .

Sei  $c : \mathbb{R} \rightarrow \mathbb{R}^+$  eine stetige Funktion mit

$$\lim_{\Delta \downarrow 0} c(\Delta) = 0$$

Ein numerisches Verfahren heißt stark konsistent, falls

$$1) \left| \frac{1}{\Delta} \mathbb{E} \left( X_1^{(\Delta, t, X)} - X \right) - f(t, X) \right|^2 \leq c(\Delta)$$

$$2) \mathbb{E} \left( \frac{1}{\Delta} \left| X_1^{(\Delta, t, X)} - X - \mathbb{E} \left( X_1^{(\Delta, t, X)} - X \right) - g(t, X) \Delta W \right|^2 \right) \leq c(\Delta),$$

und schwach konsistent, falls (1) und

$$2') \left| \frac{1}{\Delta} \mathbb{E} \left( \left| X_1^{(\Delta, t, X)} - X \right|^2 \right) - g(t, X)^2 \right| \leq c(\Delta)$$

SATZ (1) starke Konsistenz  $\Rightarrow$  starke Konvergenz  
(2) schwache Konsistenz  $\Rightarrow$  schwache Konvergenz

Beispiele

- 1) Das Euler-Maruyama-Verfahren mit  $\Delta W \sim N(0, \Delta)$  ist stark und schwach konsistent.
- 2) Das vereinfachte Euler-Verfahren ist schwach konsistent.
- 3) Das Heun-Verfahren ist stark/schwach konsistent genau dann, wenn das Rauschen additiv ist.

# Kapitel 8

## Die stochastische Kettenregel

Wir brauchen stochastische Taylor-Entwicklungen, um konsistente stochastische Approximationen höherer Ordnung systematisch herzuleiten. Diese sind durch die stochastische Kettenregel oder Ito-Formel konstruiert. Die Ito-Formel hat viele andere wichtige Anwendungen, z.B. die Herleitung der Feynman-Kac-Formel, die wir auch in diesem Kapitel betrachten werden.

### 8.1 Die Ito-Formel

Betrachte eine Ito-SDGL

$$dX_t = f(t, X_t) dt + g(t, X_t) dW_t$$

mit Lösung  $X_t$  und definiere

$$Y_t = U(t, X_t),$$

wobei  $U : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$  eine 2-mal stetig differenzierbare Funktion ist. Dann gilt:

$$\boxed{dY_t = L^0 U(t, X_t) dt + L^1 U(t, X_t) dW_t}$$

oder in Integralform

$$\boxed{Y_t = U(t_0, X_{t_0}) + \int_{t_0}^t L^0 U(s, X_s) ds + \int_{t_0}^t L^1 U(s, X_s) dW_s}$$

mit den Differential-Operatoren

$$L^0U = \frac{\partial U}{\partial t} + f \frac{\partial U}{\partial x} + \frac{1}{2} g^2 \frac{\partial^2 U}{\partial x^2}, \quad L^1U = g \frac{\partial U}{\partial x}$$

Dieser Ausdruck ist die stochastische Kettenregel und heißt Ito-Formel.

Beispiel  $Y_t = W_t^2$  Nimm  $X_t \equiv W_t$ , d.h.

$$dX_t = 0 dt + 1 dW_t \quad \Rightarrow \quad f(t, X) \equiv 0, \quad g(t, X) \equiv 1$$

Wir haben  $Y_t = W_t^2 \equiv X_t^2 = U(t, X_t)$  mit  $U(t, x) = x^2$

$$\Rightarrow \quad \frac{\partial U}{\partial t} = 0, \quad \frac{\partial U}{\partial x} = 2x, \quad \frac{\partial^2 U}{\partial x^2} = 2$$

$$\Rightarrow \quad L^0U = 0 + 0 \cdot 2x + \frac{1}{2} \cdot 1^2 \cdot 2 \equiv 1, \quad L^1U = 1 \cdot 2x = 2x$$

Die Ito-Formel lautet  $dY_t = 1 dt + 2X_t dW_t$

hier  $\Rightarrow \quad d(W_t^2) = 1 dt + 2W_t dW_t$

$$\begin{aligned} \Rightarrow \quad W_t^2 &= W_0^2 + \int_0^t 1 ds + 2 \int_0^t W_s dW_s = 0 + t + 2 \int_0^t W_s dW_s \\ &\Rightarrow \quad \int_0^t W_s dW_s = \frac{1}{2} W_t^2 - \frac{1}{2} t \end{aligned}$$

Bemerkung Bei einem Vergleich mit der totalen Ableitung

$$\frac{d}{dt} U(t, x) = \frac{\partial U}{\partial t}(t, x) + f \frac{\partial U}{\partial x}(t, x)$$

der deterministischen Kettenregel, enthält der Operator  $L^0$  einen zusätzlichen Term  $\frac{1}{2} g^2 \frac{\partial^2}{\partial x^2}$ . Der Ursprung dieses Terms liegt in der die Tatsache, dass

$$\mathbb{E}((\Delta W)^2) = \Delta t$$

Beweis (Skizze) Schreibe  $\Delta X = X_{t+\Delta t} - X_t$ . Dann gilt

$$\begin{aligned} \Delta Y = Y_{t+\Delta t} - Y_t &= U(t + \Delta t, X_{t+\Delta t}) - U(t, X_t) \\ &= U(t + \Delta t, X_t + \Delta X) - U(t, X_t) \\ &= \frac{\partial U}{\partial t} \Delta t + \frac{\partial U}{\partial x} \Delta X + \frac{1}{2} \frac{\partial^2 U}{\partial x^2} (\Delta t)^2 \\ &\quad + \frac{\partial^2 U}{\partial t \partial x} \Delta t \Delta X + \frac{1}{2} \frac{\partial^2 U}{\partial x^2} (\Delta X)^2 + \dots \end{aligned}$$

Aber  $\Delta X = f(t, X_t) \Delta t + g(t, X_t) \Delta W_t + \dots$ . Schreibe  $\Delta X = f \Delta t + g \Delta W_t + \dots$

$$\begin{aligned} \Rightarrow \Delta Y &= \frac{\partial U}{\partial t} \Delta t + \frac{\partial U}{\partial x} (f \Delta t + g \Delta W_t) \\ &+ \frac{1}{2} \frac{\partial^2 U}{\partial t^2} (\Delta t)^2 + \frac{\partial^2 U}{\partial t \partial x} \Delta t (f \Delta t + g \Delta W_t) \\ &+ \frac{1}{2} \frac{\partial^2 U}{\partial x^2} (f^2 \Delta t^2 + 2f \Delta t \cdot g \Delta W_t + g^2 (\Delta W_t)^2) + \dots \\ &= \left( \frac{\partial U}{\partial t} + f \frac{\partial U}{\partial x} + \frac{1}{2} g^2 \frac{\partial^2 U}{\partial x^2} \right) \Delta t + g \frac{\partial U}{\partial x} \Delta W_t \\ &+ \frac{1}{2} g^2 \frac{\partial^2 U}{\partial x^2} ((\Delta W_t)^2 - \Delta t) + \dots \\ \Rightarrow \Delta Y &= L^0 U \cdot \Delta t + L^1 U \Delta W_t + \frac{1}{2} g^2 \frac{\partial^2 U}{\partial x^2} \{(\Delta W_t)^2 - \Delta t\} \\ &+ \text{Terme h\"oherer Ordnungen} \end{aligned}$$

Nur die ersten 2 Terme bleiben nach qm-Konvergenz \u00fcbriigs.

## 8.2 Die Stratonovich-Kettenregel

Die Kettenregel f\u00fcr Stratonovich SDGLen ist eine direkte Verallgemeinerung der deterministischen Kettenregel. Betrachte eine Stratonovich SDGL

$$dX_t = f(t, X_t) dt + g(t, X_t) \circ dW_t$$

und eine Abbildung

$$Y_t = U(t, X_t)$$

Dann lautet die Stratonovich-Kettenregel

$$\boxed{dY_t = \underline{L}^0 U(t, X_t) dt + \underline{L}^1 U(t, X_t) \circ dW_t}$$

oder in Integralform

$$\boxed{Y_t = U(t_0, X_{t_0}) + \int_{t_0}^t \underline{L}^0 U(s, X_s) ds + \int_{t_0}^t \underline{L}^1 U(s, X_s) dW_s}$$

mit den Differential-Operatoren

$$\underline{L}^0 U = \frac{\partial U}{\partial t} + f \frac{\partial U}{\partial x}, \quad \underline{L}^1 U = g \frac{\partial U}{\partial x}$$

Dies bedeutet, dass wir Stratonovich-SDGLen mit Integrationstricks des deterministischen Kalküls lösen können.

Beispiel Stratonovich-SDGL:  $dX_t = aX_t dt + bX_t \circ dW_t$

$$\begin{aligned} \Rightarrow \frac{dX_t}{X_t} &= a dt + b \circ dW_t \\ \text{integriere} \Rightarrow \ln(X_t/X_0) &= a t + b W_t \\ X_t &= X_0 e^{at+bW_t} \end{aligned}$$

Wir können eine Ito-SDGL über die entsprechende Stratonovich-SDGL integrieren.

Beispiel Ito-SDGL:  $dX_t = aX_t dt + bX_t dW_t$  mit modifiziertem Drift

$$\begin{aligned} \underline{f}(t, x) &= f(t, x) - \frac{1}{2} g(t, x) \frac{\partial g}{\partial x}(t, x) \\ &= ax - \frac{1}{2} bx \frac{\partial}{\partial x}(bx) = \left(a - \frac{1}{2} b^2\right) x \end{aligned}$$

Entsprechende Stratonovich-SDGL:  $dX_t = \left(a - \frac{1}{2} b^2\right) X_t dt + bX_t \circ dW_t$

$$\begin{aligned} \Rightarrow \frac{dX_t}{X_t} &= \left(a - \frac{1}{2} b^2\right) dt + b \circ dW_t \\ \text{integriere} \Rightarrow \ln(X_t/X_0) &= \left(a - \frac{1}{2} b^2\right) (t - 0) + b(W_t - W_0) \\ &= \left(a - \frac{1}{2} b^2\right) t + b \cdot W_t \\ \Rightarrow X_t &= X_0 \exp\left(\left(a - \frac{1}{2} b^2\right) t + b W_t\right) \end{aligned}$$

Umgekehrt: Beweise für Stratonovich-SDGLen sind im Allgemeinen über die entsprechenden Ito-SDGLen durchgeführt.

### 8.3 Die Feynman–Kac-Formel

Die Feynman–Kac-Formel ist eine einfache, aber sehr nützliche Formel, welche einen Zusammenhang zwischen der Lösung einer deterministischen partiellen DGL und einem Erwartungswert  $\mathbb{E}[\Phi(X_T)]$  eines Funktionals der Lösung einer stochastischen DGL herstellt. Wir werden diese Formel in dem Beweis der schwachen Konvergenz des Euler–Maruyama-Verfahrens verwenden.

Hier wollen wir zeigen, wie wir diesen Erwartungswert durch die Lösung einer zugehörigen partiellen DGL (deterministisch!) darstellen können.

Wir betrachten eine Anfangswertaufgabe für eine skalare SDGL

$$dX_t = f(X_t) dt + g(X_t) dW_t, \quad X_\tau = x. \quad (8.1)$$

auf einem Zeitintervall  $[\tau, T]$ , wobei  $\tau$ ,  $T$  und  $x$  fest sind.

Sei  $u(t, x)$  eine Lösung der deterministischen PDGL

$$u_t(t, x) + f(x) u_x(t, x) + \frac{1}{2} g(x)^2 u_{xx}(t, x) = 0, \quad t \in [0, T], \quad x \in \mathbb{R}, \quad (8.2)$$

mit der Endzeitbedingung, d.h. zu der Endzeit  $T$ ,

$$u(T, x) = \Phi(x). \quad (8.3)$$

Die Ito-Formel für  $u(t, X_t)$  ergibt

$$\begin{aligned} du(t, X_t) &= \left( u_t(t, X_t) + f(X_t) u_x(t, X_t) + \frac{1}{2} g(X_t)^2 u_{xx}(t, X_t) \right) dt \\ &\quad + g(t, X_t) u_x(t, X_t) dW_t \end{aligned}$$

Wegen (8.2) ist der erste Term gleich Null und wir haben

$$du(t, X_t) = g(t, X_t) u_x(t, X_t) dW_t,$$

oder das stochastische Integral

$$u(T, X_T) - u(\tau, x) = \int_\tau^T g(s, X_s) u_x(s, X_s) dW_s.$$

Aber der Erwartungswert eines Ito-Integrals ist gleich Null, d.h.

$$\mathbb{E}[u(T, X_T)] = u(\tau, x).$$

Wegen (8.3) gilt somit

$$\boxed{\mathbb{E}[\Phi(X_T)] = u(\tau, x)} \quad (8.4)$$

welche Feynman–Kac-Formel heißt.

Die PDGL (8.2) mit Endzeitbedingung (8.3) heißt Feynman–Kac-PDGL. Wenn wir diese PDGL rückwärts in der Zeit von der Endzeit  $T$  lösen, erhalten wir durch die Feynman–Kac-Formel (8.4) den Erwartungswert  $\mathbb{E}[\Phi(X_T)]$ , wobei  $X_t$  die Lösung der SDGL (8.1) mit dem Anfangswert  $X_\tau = x$  ist.

## Kapitel 9

# Konvergenzbeweise für das Euler-Maruyama-Verfahren

Der Einfachheit wegen betrachten wir eine skalare SDGL

$$dX_t = f(X_t) dt + g(X_t) dW_t \quad (9.1)$$

und das entsprechende Euler-Maruyama-Verfahren

$$X_{n+1} = X_n + f(X_n) \Delta + g(X_n) \Delta W_n. \quad (9.2)$$

mit konstanter Schrittweite. Wir werden oft  $X_n^{(\Delta)}$  statt  $X_n$  schreiben

Wir setzen voraus, dass die Koeffizientenfunktionen  $f$  und  $g$  globalen Lipschitz-Bedingungen genügen, d.h.

$$|f(x) - f(y)| \leq L|x - y|, \quad |g(x) - g(y)| \leq L|x - y|, \quad \forall x, y \in \mathbb{R}.$$

Daher genügen  $f$  und  $g$  auch den linearen Wachstumsbedingungen

$$|f(x)|^2 \leq \widehat{L}(1 + |x|^2), \quad |g(x)|^2 \leq \widehat{L}(1 + |x|^2), \quad \forall x \in \mathbb{R}, \quad (9.3)$$

für eine geeignete Konstante  $\widehat{L}$ .

### 9.1 Starke Konvergenz

Wir wollen zeigen, dass der starke Diskretisierungsfehler des Euler-Maruyama-Verfahrens der folgenden Abschätzung genügt

$$\sup_{0 \leq n \leq N} \mathbb{E} \left( |X_n^{(\Delta)} - X_{t_n}| \right) \leq K_T \Delta^{\frac{1}{2}}, \quad (9.4)$$

falls  $\Delta$  klein genug ist, wobei die Konstante  $K_T$  nicht von  $\Delta$  abhängt.

Dafür ist es günstig eine zeitkontinuierliche Erweiterung der “Lösung” des Euler-Maruyama-Verfahrens einzuführen, nämlich den stückweise konstanten Prozess  $\bar{X}_t^{(\Delta)}$  definiert durch

$$\bar{X}_t^{(\Delta)} = X_n^{(\Delta)}, \quad \text{für } t_n \leq t < t_{n+1}. \quad (9.5)$$

Von den obigen Voraussetzungen über die Koeffizientenfunktionen kann man zeigen, dass die zweiten Momente  $\mathbb{E}(X_t^2)$  und  $\mathbb{E}((\bar{X}_t^{(\Delta)})^2)$  für  $t \in [0, T]$  beschränkt sind.

Insbesondere werden wir beweisen, dass

$$Z(t) \leq K_T^2 \Delta, \quad t \in [0, T], \quad (9.6)$$

falls  $\Delta$  genügend klein ist, wobei

$$Z(t) := \sup_{0 \leq s \leq t} \mathbb{E} \left( \left| \bar{X}_s^{(\Delta)} - X_s \right|^2 \right).$$

Mit einer Anwendung der Ljapunov-Ungleichung, d.h.  $\mathbb{E}(X) \leq \sqrt{\mathbb{E}(X^2)}$ , erhalten wir dann die Fehlerabschätzung

$$\sup_{0 \leq s \leq T} \mathbb{E} \left( \left| \bar{X}_s^{(\Delta)} - X_s \right| \right) \leq K_T \Delta^{\frac{1}{2}}.$$

Es folgt dann die gewünschte Fehlerabschätzung (9.4), weil  $\bar{X}_t^{(\Delta)} = X_n^{(\Delta)}$  für  $0 \leq t_n \leq t < t_{n+1} \leq T$ .

Ein wichtiges Werkzeug in dem Beweis ist die Ungleichung von Gronwall.

**Lemma 1** Sei  $\alpha : [0, T] \rightarrow \mathbb{R}$  eine integrierbare Funktion mit

$$0 \leq \alpha(t) \leq A + B \int_0^t \alpha(r) dr, \quad 0 \leq t \leq T,$$

wobei  $A, B > 0$  Konstanten sind. Dann ist

$$\alpha(t) \leq Ae^{Bt}, \quad 0 \leq t \leq T.$$

Wir werden eine Ungleichung der Form

$$Z(t) \leq A\Delta + B \int_0^t Z(r) dr, \quad 0 \leq t \leq T, \quad (9.7)$$

herleiten und dann die Gronwall-Ungleichung verwenden, um die Ungleichung (9.6) mit  $K_T^2 = Ae^{BT}$  zu erhalten.

### 9.1.1 Beweis-Skizze für starke Konvergenz

Sei  $s \in [0, T]$  und definiere  $n_s$  als die ganze Zahl, so dass  $s \in [t_{n_s}, t_{n_s+1})$ , wobei  $t_n := n\Delta$ . Dann gilt  $\bar{X}_s = X_{n_s}$  und wir haben

$$\begin{aligned}
\bar{X}_s - X_s &= X_{n_s} - X_s \\
&= X_{n_s} - \left( X_0 + \int_0^s f(X_r) dr + \int_0^s g(X_r) dW_r \right) \\
&= \sum_{i=0}^{n_s-1} (X_{i+1} - X_i) - \int_0^s f(X_r) dr - \int_0^s g(X_r) dW_r \\
&= \sum_{i=0}^{n_s-1} f(X_i)\Delta + \sum_{i=0}^{n_s-1} g(X_i)\Delta W_i - \int_0^s f(X_r) dr - \int_0^s g(X_r) dW_r. \quad (9.8)
\end{aligned}$$

Die Definition von  $\bar{X}_t$  ergibt

$$\int_{t_i}^{t_{i+1}} f(\bar{X}_t) dt = \int_{t_i}^{t_{i+1}} f(X_i) dt = f(X_i) \int_{t_i}^{t_{i+1}} dt = f(X_i)\Delta, \quad (9.9)$$

und ähnlicherweise

$$\int_{t_i}^{t_{i+1}} g(\bar{X}_t) dW_t = g(X_i)\Delta W_i, \quad (9.10)$$

die wir in (9.8) einfügen und erhalten

$$\begin{aligned}
\bar{X}_s - X_s &= \int_0^{t_{n_s}} f(\bar{X}_r) dr + \int_0^{t_{n_s}} g(\bar{X}_r) dW_r - \int_0^s f(X_r) dr - \int_0^s g(X_r) dW_r \\
&= \int_0^{t_{n_s}} (f(\bar{X}_r) - f(X_r)) dr + \int_0^{t_{n_s}} (g(\bar{X}_r) - g(X_r)) dW_r \quad (9.11)
\end{aligned}$$

$$- \int_{t_{n_s}}^s f(X_r) dr - \int_{t_{n_s}}^s g(X_r) dW_r. \quad (9.12)$$

Die Integrale in der Zeile (9.11) enthalten den Fehler des Euler-Maruyama-Verfahrens in dem Zeitintervall  $[0, t_{n_s}]$  und die Integrale in der Zeile (9.12) enthalten den Fehler der Approximation von  $X_s$  durch  $X_{n_s} = \bar{X}_s$ .

Wie brauchen die folgende Ungleichung:

$$(a + b + c + d)^2 \leq 4(a^2 + b^2 + c^2 + d^2), \quad a, b, c, d \in \mathbb{R} \quad (9.13)$$

Wir quadrieren die beiden Seiten von (9.11)-(9.12) und nehmen die Erwartungswerte. Dann erhalten wir

$$\begin{aligned}
\mathbb{E} \left( (\bar{X}_s - X_s)^2 \right) &\leq 4 \left\{ \mathbb{E} \left( \left( \int_0^{t_{n_s}} (f(\bar{X}_r) - f(X_r)) dr \right)^2 \right) \right. \\
&\quad \left. + \mathbb{E} \left( \left( \int_0^{t_{n_s}} (g(\bar{X}_r) - g(X_r)) dW_r \right)^2 \right) \right\}
\end{aligned}$$

$$\begin{aligned}
& +\mathbb{E} \left( \left( \int_{t_{n_s}}^s f(X_r) dr \right)^2 \right) \\
& +\mathbb{E} \left( \left( \int_{t_{n_s}}^s g(X_r) dW_r \right)^2 \right) \}. \quad (9.14)
\end{aligned}$$

Wir werden jetzt die vier Integrale an der rechten Seite abschätzen. Für das erste Integral benutzen wir die Cauchy-Schwarz-Ungleichung und die Lipschitz-Bedingung, um zu erhalten

$$\begin{aligned}
\mathbb{E} \left( \left( \int_0^{t_{n_s}} (f(\bar{X}_r) - f(X_r)) , dr \right)^2 \right) & \leq t_{n_s} \int_0^{t_{n_s}} \mathbb{E} \left( (f(\bar{X}_r) - f(X_r))^2 \right) dr \\
& \leq TL^2 \int_0^{t_{n_s}} \mathbb{E} \left( (\bar{X}_r - X_r)^2 \right) dr \\
& \leq TL^2 \int_0^s Z(r) dr. \quad (9.15)
\end{aligned}$$

Analog erhalten wir durch die Cauchy-Schwarz-Ungleichung und die lineare Wachstumsbedingung (9.3)

$$\begin{aligned}
\mathbb{E} \left( \left( \int_{t_{n_s}}^s f(X_r) dr \right)^2 \right) & \leq (s - t_{n_s}) \int_{t_{n_s}}^s \mathbb{E} (f(X_r)^2) dr \\
& \leq \Delta \widehat{L} \int_{t_{n_s}}^s (1 + \mathbb{E} (X_r^2)) dr \\
& \leq C_1 \Delta^2, \quad (9.16)
\end{aligned}$$

für eine Konstante  $C_1$ . Hier brauchen wir die Beschränktheit des Erwartungswertes  $\mathbb{E}(X_r^2)$ .

Für die stochastischen Integrale benutzen wir die Ito-Isometrie (4.1). Dann erhalten wir mit der Lipschitz-Bedingung

$$\begin{aligned}
\mathbb{E} \left( \left( \int_0^{t_{n_s}} (g(\bar{X}_r) - g(X_r)) dW_r \right)^2 \right) & = \int_0^{t_{n_s}} \mathbb{E} \left( (g(\bar{X}_r) - g(X_r))^2 \right) dr \\
& \leq L^2 \int_0^{t_{n_s}} \mathbb{E} \left( (\bar{X}_r - X_r)^2 \right) dr \\
& \leq L^2 \int_0^s Z(r) dr. \quad (9.17)
\end{aligned}$$

Mit der linearen Wachstumsbedingung (9.3) folgt nun

$$\begin{aligned}
\mathbb{E} \left( \left( \int_{t_{n_s}}^s g(X_r) dW_r \right)^2 \right) & = \int_{t_{n_s}}^s \mathbb{E} (g(X_r)^2) dr \\
& \leq \widehat{L} \int_{t_{n_s}}^s (1 + \mathbb{E} (X_r^2)) dr \\
& \leq C_2 \Delta, \quad (9.18)
\end{aligned}$$

für eine Konstante  $C_2$ .

Die Abschätzungen (9.15)–(9.18) in der Ungleichung (9.14) ergeben die erwünschte Fehlerabschätzung (9.7),  $\square$

**Bemerkung zum Beweis** Die Abschätzungen (9.16) und (9.18) der deterministischen und stochastischen Integrale zeigen den wesentlichen Unterschied zwischen der deterministischen und stochastischen Fehleranalyse. Ohne das stochastische Integral in (9.18) würden wir eine Gesamtfehlerabschätzung der Form  $O(\Delta)$  erhalten, genau wie für das deterministische Euler-Verfahren für eine gewöhnliche Differentialgleichung.

## 9.2 Schwache Konvergenz

Hier betrachten wir nochmal die skalare SDGL (9.1) und das entsprechende Euler–Maruyama-Verfahren (9.2) unter den obigen Voraussetzungen. Unsere Hauptwerkzeuge sind die Feynman–Kac-PDGL (8.2) und die Feynman–Kac-Formel (8.4).

### 9.2.1 Beweis-Skizze für schwache Konvergenz

Der Einfachheit wegen betrachten wir den Fall mit einem deterministischen Anfangswert  $X_0 = x_0$ . Wir werden auch voraussetzen, dass alle Erwartungswerte im folgenden Beweis beschränkt sind — in einem vollständigen Beweis kann man diese Eigenschaften mit Hilfe der globalen Lipschitz-Bedingungen, unter anderem, beweisen.

Wir werden auch nur den Fehler

$$|\mathbb{E}(\Phi(X_N)) - \mathbb{E}(\Phi(X_T))| \quad (9.19)$$

abschätzen, d.h. zu der Endzeit  $T$ , statt

$$\max_{n=0,1,\dots,N} |\mathbb{E}(\Phi(X_n)) - \mathbb{E}(\Phi(X_{t_n}))|.$$

Wie im Beweis der starken Konvergenz ist es auch günstig Interpolationen der Iterationen  $X_n$  des Euler–Maruyama-Verfahrens auf dem ganzen Intervall  $[0, T]$  einzuführen. Wir haben schon den stückweise konstanten Prozess  $\bar{X}_t$  definiert durch

$$\bar{X}_t = X_n, \quad \text{für } t_n \leq t < t_{n+1},$$

benutzt. Wir brauchen jetzt auch den stückweise linear-interpolierten Prozess

$$\hat{X}_t = X_n + f(X_n)(t - t_n) + g(X_n)(W_t - W_{t_n}), \quad t_n \leq t < t_{n+1}, \quad (9.20)$$

oder in der Integraldarstellung

$$\hat{X}_t = X_0 + \int_0^t f(\bar{X}_s) ds + \int_0^t g(\bar{X}_s) dW_s, \quad t \in [0, T]. \quad (9.21)$$

Der Prozess  $\widehat{X}_t$  ist offensichtlich pfadweise stetig.

Merke, dass  $\widehat{X}_{t_n+\gamma}$  für  $\gamma \in [0, \Delta)$  der Wert des Euler–Maruyama-Verfahrens mit Schrittweite  $\gamma$  und Anfangswert  $(t_n, X_n)$  ist. Es gilt  $\widehat{X}_{t_n} = X_n$  für  $n = 0, 1, 2, \dots, N$ . Insbesondere  $\widehat{X}_T = X_N$ . Daher lautet der obige schwache Diskretisierungsfehler (9.19)

$$\left| \mathbb{E}(\Phi(\widehat{X}_T)) - \mathbb{E}(\Phi(X_T)) \right|.$$

Ein Vorteil des Prozesses  $\widehat{X}_t$  ist: wir können ihn in der Ito-Formel benutzen.

Sei  $u(t, x)$  die Lösung der Feynman-Kac-PDGL (8.2),

$$u_t(t, x) + f(x) u_x(t, x) + \frac{1}{2} g(x)^2 u_{xx}(t, x) = 0, \quad t \in [0, T], \quad x \in \mathbb{R},$$

mit dem Endzeitwert  $u(T, x) = \Phi(x)$ .

Wir verwenden die Ito-Formel für  $u(t, \widehat{X}_t)$  und erhalten

$$\begin{aligned} du(t, \widehat{X}_t) &= \left( u_t(t, \widehat{X}_t) + u_x(t, \widehat{X}_t) f(\bar{X}_t) + \frac{1}{2} u_{xx}(t, \widehat{X}_t) g(\bar{X}_t)^2 \right) dt \\ &\quad + u_x(t, \widehat{X}_t) g(\bar{X}_t) dW_t. \end{aligned} \quad (9.22)$$

(Vorsicht mit  $\widehat{X}_t$  und  $\bar{X}_t$  hier! Warum ist es so?)

Wir werden jetzt die folgenden Abkürzungen benutzen:  $\widehat{u}$  statt  $u(t, \widehat{X}_t)$  und  $\bar{f}$  statt  $f(\bar{X}_t)$ , usw. Dann lautet (9.22)

$$d\widehat{u} = \left( \widehat{u}_t + \widehat{u}_x \bar{f} + \frac{1}{2} \widehat{u}_{xx} \bar{g}^2 \right) dt + \widehat{u}_x \bar{g} dW_t.$$

Wegen der Feynman-Kac-PDGL können wir  $\widehat{u}_t$  durch  $-\widehat{u}_x \widehat{f} - \frac{1}{2} \widehat{u}_{xx} \widehat{g}^2$  ersetzen, um zu erhalten

$$d\widehat{u} = \left( (\bar{f} - \widehat{f}) \widehat{u}_x + \frac{1}{2} (\bar{g}^2 - \widehat{g}^2) \widehat{u}_{xx} \right) dt + \widehat{u}_x \bar{g} dW_t$$

deren Integraldarstellung lautet

$$\begin{aligned} u(T, \widehat{X}_T) - u(0, \widehat{X}_0) &= \int_0^T (\bar{f} - \widehat{f}) \widehat{u}_x dt + \int_0^T \frac{1}{2} (\bar{g}^2 - \widehat{g}^2) \widehat{u}_{xx} dt \\ &\quad + \int_0^T \bar{g} \widehat{u}_x dW_t. \end{aligned} \quad (9.23)$$

Von den obigen Definitionen haben wir  $\widehat{X}_T = X_N$  und  $\widehat{X}(0) = X_0 = x_0$ , sowie  $u(T, X_N) = \Phi(X_N)$  (Endzeitbedingung der Feynman-Kac-PDGL). Überdies ergibt die Feynman-Kac-Formel (8.4)

$$u(0, x_0) = u(0, X(0)) = \mathbb{E}(\Phi(X_T)).$$

Dann nehmen wir den Erwartungswert der beiden Seiten der Gleichung (9.23) und wir erhalten

$$\mathbb{E}(\Phi(X_N)) - \mathbb{E}(\Phi(X_T)) = \mathbb{E} \left( \int_0^T (\bar{f} - \hat{f}) \hat{u}_x dt + \int_0^T \frac{1}{2} (\bar{g}^2 - \hat{g}^2) \hat{u}_{xx} dt \right), \quad (9.24)$$

weil der Erwartungswert eines Ito-Integrals gleich Null ist. Daher haben wir

$$|\mathbb{E}(\Phi(X_N)) - \mathbb{E}(\Phi(X_T))| \leq \left| \mathbb{E} \left[ \int_0^T (\bar{f} - \hat{f}) \hat{u}_x dt \right] \right| + \left| \mathbb{E} \left[ \int_0^T \frac{1}{2} (\bar{g}^2 - \hat{g}^2) \hat{u}_{xx} dt \right] \right|. \quad (9.25)$$

Die beiden Integrale auf der rechten enthalten die Differenzen  $\bar{f} - \hat{f}$  und  $\bar{g}^2 - \hat{g}^2$ , d.h. die Differenzen der Funktionen ausgewertet für die stückweise konstante Interpolation  $\bar{X}_t$  und die lineare Interpolation  $\hat{X}_t$  der numerischen Lösung. Die beiden Interpolationen sind identisch an den die Diskretisierungszeiten  $t_n$ . Wir können hoffen, dass die Erwartungswerte der Integrale (9.25) klein sind, wenn die Schrittweite  $\Delta$  klein ist. Der Beweis folgt.

Um den ersten Term auf der rechten Seite von (9.25) zu behandeln, definieren wir

$$e(t, x) := (f(X_n) - f(x)) u_x(t, x), \quad t \in [t_n, t_{n+1}). \quad (9.26)$$

Wir verwenden die Ito-Formel für  $e(t, \hat{X}_t)$  und erhalten

$$de(t, \hat{X}_t) = \left( \hat{e}_t + \hat{e}_x \bar{f} + \frac{1}{2} \hat{e}_{xx} \bar{g}^2 \right) dt + \hat{e}_x \bar{g} dW_t.$$

Integration von  $t_n$  bis  $s \in [t_n, t_{n+1})$  ergibt

$$e(s, \hat{X}_s) = \int_{t_n}^s \left( \hat{e}_t + \hat{e}_x \bar{f} + \frac{1}{2} \hat{e}_{xx} \bar{g}^2 \right) dt + \int_{t_n}^s \hat{e}_x \bar{g} dW_t \quad (9.27)$$

weil

$$e(t_n, \hat{X}_{t_n}) = (f(X_n) - f(X_n)) u_x(t_n, X_n) = 0$$

ist. Dann nehmen wir den Erwartungswert von beiden Seiten der Gleichung (9.27) und vertauschen die Reihenfolge von Erwartungswert und Integral. Wir erhalten

$$\begin{aligned} \mathbb{E} \left( e(s, \hat{X}_s) \right) &= \mathbb{E} \left( \int_{t_n}^s \left( \hat{e}_t + \hat{e}_x \bar{f} + \frac{1}{2} \hat{e}_{xx} \bar{g}^2 \right) dt \right) \\ &= \int_{t_n}^s \left( \mathbb{E}(\hat{e}_t) + \mathbb{E}(\hat{e}_x \bar{f}) + \frac{1}{2} \mathbb{E}(\hat{e}_{xx} \bar{g}^2) \right) dt. \end{aligned}$$

Wir haben am Anfang vorausgesetzt, dass alle Erwartungswerte hier beschränkt sind. Daher gibt es eine Konstante  $K_1$ , so dass

$$\left| \mathbb{E} \left( e(s, \hat{X}_s) \right) \right| \leq K_1 \Delta,$$

weil die Integrale sind auf einem Intervall von Länge  $s - t_n \leq \Delta$ .

Überdies gilt diese Abschätzung für jedes  $s \in [0, T]$ . Daher haben wir

$$\begin{aligned} \left| \mathbb{E} \left( \int_0^T (\bar{f} - \hat{f}) \hat{u}_x dt \right) \right| &= \left| \int_0^T \mathbb{E} \left( (\bar{f} - \hat{f}) \hat{u}_x \right) dt \right| \\ &= \left| \int_0^T \mathbb{E} \left( e(s, \hat{X}_s) \right) dt \right| \\ &\leq \left| \int_0^T K_1 \Delta dt \right| \leq K_1 T \Delta. \end{aligned}$$

Analog können wir zeigen, dass

$$\left| \mathbb{E} \left( \int_0^T \frac{1}{2} (\bar{g}^2 - \hat{g}^2) \hat{u}_{xx} dt \right) \right| \leq K_2 T \Delta$$

für eine Konstante  $K_2$ .

Jetzt sind wir fertig. Wir fügen die beiden Abschätzungen in (9.25) ein und erhalten das erwünschte Ergebnis:

$$|\mathbb{E}[\Phi(X_N)] - \mathbb{E}[\Phi(X_T)]| \leq (K_1 + K_2) T \Delta.$$

## Kapitel 10

# Stochastische Taylor-Entwicklungen

Betrachte den deterministischen Fall nochmal: d.h., eine gewöhnliche DGL

$$\frac{dx}{dt} = f(t, x)$$

Eine Lösung  $x(t)$  auf  $[0, T]$  genügt der Integralgleichung

$$\boxed{x(t) = x(0) + \int_0^t f(s, x(s)) ds} \quad (*)$$

Definiere  $y(t) = u(t, x(t))$ , wobei  $u : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$  stetig differenzierbar ist. Dann lautet die deterministische Kettenregel

$$\frac{d}{dt} u(t, x(t)) = Lu(t, x(t)),$$

wobei

$$Lu = \frac{\partial u}{\partial t} + f \frac{\partial u}{\partial x}$$

die totale Ableitung bzgl. der obigen DGL ist.

Die äquivalente Integralgleichung lautet:

$$\boxed{u(t, x(t)) = u(0, x(0)) + \int_0^t Lu(s, x(s)) ds} \quad (**)$$

Wir verwenden  $(**)$  auf dem Intervall  $[0, s]$  statt  $[0, t]$  für  $u \equiv f$ , d.h.,

$$f(s, x(s)) = f(0, x(0)) + \int_0^s Lf(\tau, x(\tau)) d\tau.$$

Mit diesem Ausdruck für  $f(s, x(s))$  können wir (\*) umschreiben:

$$\begin{aligned} x(t) &= x(0) + \int_0^t \left[ f(0, x(0)) + \int_0^s Lf(\tau, x(\tau)) d\tau \right] ds \\ &= \underbrace{x(0) + f(0, x(0)) \int_0^t ds}_{\text{Taylor-Approximation}} + \underbrace{\int_0^t \int_0^s Lf(\tau, x(\tau)) d\tau ds}_{\text{Rest}} \end{aligned}$$

Dies ist die einfachste nicht-triviale Taylor-Entwicklung der Lösung der obigen DGL.

Wir können weitergehen: verwende (\*\*) auf dem Intervall  $[0, \tau]$  für  $u = Lf$ , d.h.

$$Lf(\tau, x(\tau)) = Lf(0, x(0)) + \int_0^\tau LLf(p, x(p)) dp,$$

dann folgt

$$\begin{aligned} x(t) &= \underbrace{x(0) + f(0, x(0)) \int_0^t ds + Lf(0, x(0)) \int_0^t \int_0^s d\tau ds}_{\text{nächstfolgende Taylor-Approximation}} \\ &\quad + \underbrace{\int_0^t \int_0^s \int_0^\tau LLf(p, x(p)) dp d\tau ds}_{\text{neuer Rest}}. \end{aligned}$$

Der stochastische Fall ist formell ähnlich, aber jedes Mal haben wir mehrere Integrale, die wir entwickeln können. Betrachte die Ito-SDGL

$$dX_t = f(t, X_t) dt + g(t, X_t) dW_t.$$

Eine Lösung  $X_t$  auf dem Intervall  $[0, T]$  genügt der stochastischen Integralgleichung

$$\boxed{X_t = X_0 + \int_0^t f(s, X_s) ds + \int_0^t g(s, X_s) dW_s}. \quad (***)$$

Definiere  $Y_t = U(t, X_t)$ , wobei  $U : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$  zwei-mal stetig differenzierbar ist. Die stochastische Kettenregel oder Ito-Formel lautet

$$\boxed{dY_t = dU(t, X_t) = L^0 U(t, X_t) dt + L^1 U(t, X_t) dW_t}$$

oder in Integralform

$$\boxed{U(t, X_t) = U(0, X_0) + \int_0^t L^0 U(s, X_s) ds + \int_0^t L^1 U(s, X_s) dW_s} \quad (***)$$

wobei

$$L^0 U = \frac{\partial U}{\partial t} + f \frac{\partial U}{\partial x} + \frac{1}{2} g^2 \frac{\partial^2 U}{\partial x^2}, \quad L^1 U = g \frac{\partial U}{\partial x}$$

Jetzt haben wir 2 Integrale in (\*\*\*) , deren Integranden wir entwickeln können.

Wir verwenden (\*\*\*) auf dem Intervall  $[0, s]$  mit

(1)  $U \equiv f$  in dem Riemann-Integral

$$\int_0^t f(s, X_s) ds = \int_0^t \left( f(0, X_0) + \int_0^s L^0 f(\tau, X_\tau) d\tau + \int_0^s L^1 f(\tau, X_\tau) dW_\tau \right) ds$$

(2)  $U \equiv g$  in dem Ito-Integral

$$\int_0^t g(s, X_s) dW_s = \int_0^t \left( g(0, X_0) + \int_0^s L^0 g(\tau, X_\tau) d\tau + \int_0^s L^1 g(\tau, X_\tau) dW_\tau \right) dW_s$$

Nach Umschreibung lautet (\*\*\*) dann

$$\begin{aligned} X_t &= \underbrace{X_0 + f(0, X_0) \int_0^t dt + g(0, X_0) \int_0^t dW_s}_{\text{Taylor-Approximation}} \\ &+ \left. \begin{aligned} &+ \int_0^t \int_0^s L^0 f(\tau, X_\tau) d\tau ds + \int_0^t \int_0^s L^1 f(\tau, X_\tau) dW_\tau ds \\ &+ \int_0^t \int_0^s L^0 g(\tau, X_\tau) d\tau dW_s + \int_0^t \int_0^s L^1 g(\tau, X_\tau) dW_\tau dW_s \end{aligned} \right\} \text{Restterme} \end{aligned}$$

Dies ist die einfachste nicht-triviale Taylor-Entwicklung.

Von der Taylor-Approximation

$$\boxed{X_{t_{n+1}} \approx X_{t_n} + f(t_n, X_{t_n}) \int_{t_n}^{t_{n+1}} ds + g(t_n, X_{t_n}) \int_{t_n}^{t_{n+1}} dW_s}$$

auf einem Intervall  $[t_n, t_{n+1}]$  (statt  $[0, t]$ ) erhalten wir das stochastische Euler-Verfahren

$$\boxed{Y_{n+1} = Y_n + f(t_n, Y_n) \Delta_n + g(t_n, Y_n) \Delta W_n}$$

mit

$$\Delta_n = \int_{t_n}^{t_{n+1}} ds = t_{n+1} - t_n, \quad \Delta W_n = \int_{t_n}^{t_{n+1}} dW_s = W_{t_{n+1}} - W_{t_n}.$$

Der Rest in der obigen Taylor-Entwicklung enthält jetzt 4 Doppelintegrale (deterministisch, stochastisch und gemischt) im Vergleich mit nur einem Integral im deterministischen Fall.

Frage Welche Integrale des Rests sollen wir weiter entwickeln?

Im Prinzip alle, aber in der Praxis brauchen wir nur gewisse Integrale zu entwickeln - die Wahl bestimmt eine Taylor-Entwicklung. Daher können wir viele verschiedene stochastische Taylor-Entwicklungen herleiten.

Zum Beispiel betrachte das doppelte stochastische Integral

$$\int_0^t \int_0^s L^1 g(\tau, X_\tau) dW_\tau dW_s$$

Wir verwenden die Ito-Formel auf  $[0, \tau]$  für  $U \equiv L^1 g$  in diesem Integral und erhalten

$$\begin{aligned} & \int_0^t \int_0^s L^1 g(\tau, X_\tau) dW_\tau dW_s \\ = & \int_0^t \int_0^s \left[ L^1 g(0, X_0) + \int_0^\tau L^0 L^1 g(p, X_p) dp + \int_0^\tau L^1 L^1 g(p, X_p) dW_p \right] dW_\tau dW_s \\ = & \underbrace{L^1 g(0, X_0) \int_0^t \int_0^s dW_\tau dW_s}_{\text{neuer Term in der Taylor-Approximation}} \\ & + \int_0^t \int_0^s \int_0^p L^0 L^1 g(p, X_p) dp dW_\tau dW_s \\ & + \int_0^t \int_0^s \int_0^p L^1 L^1 g(p, X_p) dW_p dW_\tau dW_s \end{aligned}$$

neue Restterme

Wir haben dann die folgende Taylor-Approximation

$$X_t \approx X_0 + f(0, X_0) \int_0^t ds + g(0, X_0) \int_0^t dW_s + L^1 g(0, X_0) \int_0^t \int_0^s dW_\tau dW_s$$

Das entsprechende numerische Verfahren

$$\begin{aligned} Y_{n+1} = & Y_n + f(t_n, Y_n) \Delta_n + g(t_n, Y_n) \Delta W_n \\ & + L^1 g(t_n, Y_n) \int_{t_n}^{t_{n+1}} \int_{t_n}^s dW_\tau dW_s \end{aligned}$$

auf dem Intervall  $[t_n, t_{n+1}]$  heißt Milstein-Verfahren. Es besitzt die starke Ordnung  $\gamma = 1$  und die schwache Ordnung  $\beta = 1$ .

$\Rightarrow$  eine Verbesserung in dem starken Sinne. Der Grund dafür ist: das doppelte stochastische Integral enthält mehr Information über das Verhalten des Wiener-Prozesses innerhalb des Teilintervalls  $[t_n, t_{n+1}]$  als der einfache Zuwachs

$$\Delta W_n = W_{t_{n+1}} - W_{t_n} = \int_{t_n}^{t_{n+1}} dW_s$$

$$\text{Hilfssatz } \int_{t_n}^{t_{n+1}} \int_{t_n}^t dW_s dW_t = \frac{1}{2}(\Delta W_n)^2 - \frac{1}{2}\Delta_n$$

Beweis:

$$\begin{aligned} \int_{t_n}^{t_{n+1}} \left[ \int_{t_n}^t dW_s \right] dW_t &= \int_{t_n}^{t_{n+1}} (W_t - W_{t_n}) dW_t \\ &= \int_{t_n}^{t_{n+1}} W_t dW_t - W_{t_n} \int_{t_n}^{t_{n+1}} dW_t \\ &= \left( \frac{1}{2}W_t^2 - \frac{1}{2}t \right) \Big|_{t_n}^{t_{n+1}} - W_{t_n} (W_{t_{n+1}} + W_{t_n}) \\ &= \frac{1}{2}W_{t_{n+1}}^2 - \frac{1}{2}t_{n+1} - \frac{1}{2}W_{t_n}^2 + \frac{1}{2}t_n - W_{t_n} (W_{t_{n+1}} + W_{t_n}) \\ &= \frac{1}{2} (W_{t_{n+1}} - W_{t_n})^2 - \frac{1}{2}(t_{n+1} - t_n) \\ &= \frac{1}{2}(\Delta W_n)^2 - \frac{1}{2}\Delta_n \end{aligned}$$

Bemerkung Es gilt

$$L^1 g(t, x) = g(t, x) \frac{\partial g}{\partial x}(t, x)$$

Daher lautet das Milstein-Verfahren

$$\begin{aligned} Y_{n+1} &= Y_n + f(t_n, Y_n) \Delta_n + g(t_n, Y_n) \Delta W_n \\ &\quad + \frac{1}{2}g(t_n, Y_n) \frac{\partial g}{\partial x}(t_n, Y_n) \{(\Delta W_n)^2 - \Delta_n\} \end{aligned}$$

d.h. wir brauchen nur die Rauschenzuwächse  $\Delta W_n$  hier.

Bemerkung Der letzte Term des Milstein-Verfahrens verschwindet bei additivem Rauschen, d.h.

$$\frac{\partial g}{\partial x}(t, x) \equiv 0.$$

Für solche SDGLen sind das Euler-Verfahren und das Milstein-Verfahren identisch. In diesem Fall hat das Euler-Verfahren starke Ordnung  $\gamma = 1$ .

# Kapitel 11

## Taylor-Bäume

Wir werden eine Ito-SDGL

$$dX_t = f(t, X_t) dt + g(t, X_t) dW_t$$

mit günstiger Notation umschreiben.

Statt  $f, g, dt, dW_t$  schreiben wir  $f^0, f^1, dW_t^0, dW_t^1$ . Dann lautet die SDGL

$$dX_t = \sum_{i=0}^1 f^i(t, X_t) dW_t^i$$

und die entsprechende Integralgleichung lautet

$$X_t = X_0 + \sum_{i=0}^1 \int_0^t f^i(s, X_s) dW_s^i$$

⇒ Index  $i = 0$  Riemann Integration,  $i = 1$  Ito-Integration

Wir können die Ito-Formel ähnlich umschreiben

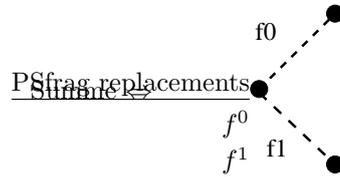
$$U(t, X_t) = U(0, X_0) + \sum_{i=0}^1 \int_0^t L^i U(s, X_s) dW_s^i$$

mit den Differential-Operatoren

$$L^0 U = \frac{\partial U}{\partial t} + f^0 \frac{\partial U}{\partial x} + \frac{1}{2} (f^1)^2 \frac{\partial^2 U}{\partial x^2}, \quad L^1 U = f^1 \frac{\partial U}{\partial x}$$

Wir können die obigen Ito-Integralgleichungen mit einem Graphen darstellen.

$$X_t - X_0 = \sum_{i=0}^1 \int_0^t f^i(s, X_s) dW_s^i$$



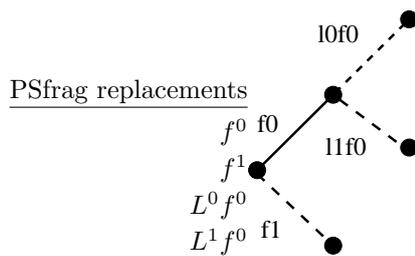
Interpretation

- (1) zunehmende Linie ↗ ⇔ Riemann-Integral ( $i = 0$ )
- (2) abfallende Linie ↘ ⇔ Ito-Integral ( $i = 1$ )
- (3) gestrichelte Linie  $\overset{h}{-}$  ⇔ Integrand  $h(s, X_s)$
- (4) durchgezogene Linie  $\overset{h}{-}$  ⇔ Integrand  $h(0, X_0)$

Verwende die Ito-Formel mit  $U = f^i$  innerhalb eines der obigen Integrale , d.h

$$\begin{aligned} \int_0^t f^i(s, X_s) dW_s^i &= \int_0^t \left[ f^i(0, X_0) + \sum_{j=0}^1 \int_0^s L^j f^i(\tau, X_\tau) dW_\tau^j \right] dW_s^i \\ &= f^i(0, X_0) \int_0^t dW_s^i + \sum_{j=0}^1 \int_0^t \int_0^s L^j f^i(\tau, X_\tau) dW_\tau^j dW_s^i \end{aligned}$$

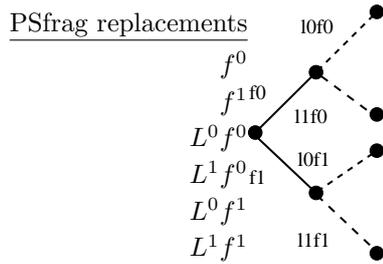
Wir können den obigen Graphen ergänzen, z.B für  $i = 0$



$$\begin{aligned} X_t - X_0 &= f^0(0, X_0) \int_0^t dW_s^0 \\ &+ \int_0^t f^1(s, W_s) dW_s^1 \\ &+ \int_0^t \int_0^s L^0 f^0(\tau, X_\tau) dW_\tau^0 dW_s^0 \quad \left[ \Leftrightarrow \int_0^t \int_0^s L^0 f^0(\tau, X_\tau) d\tau ds \right] \end{aligned}$$

$$+ \int_0^t \int_0^s L^1 f^0(\tau, X_\tau) dW_\tau^1 dW_s^0 \quad \left[ \Leftrightarrow \int_0^t \int_0^s L^1 f^0(\tau, X_\tau) dW_\tau^1 ds \right]$$

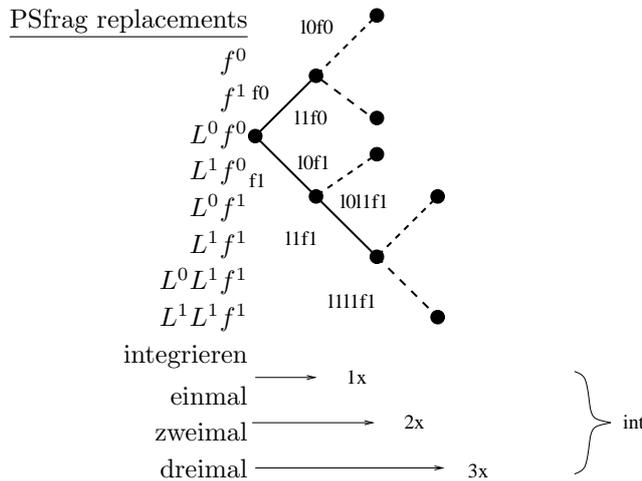
Mit den beiden ( $i_1 = 0$  und  $1$ ) erhalten wir den Taylor-Baum für die Euler-Approximation



$$X_t - X_0 = \sum_i^1 f^i(0, X_0) \int_0^t dW_s^i$$

$$\sum_{i_1=0}^1 \sum_{i_2=0}^1 \int_0^t \int_0^s L^{i_1} f^{i_2}(\tau, X_\tau) dW_\tau^{i_1} dW_s^{i_2}$$

Verwenden wir die Ito-Formel mit  $U = L^1 f^1$  hier, dann erhalten wir den Taylor-Baum für die Milstein-Approximation.



\* Eine neue Kante  $\Rightarrow$  eine weitere Integration.

\* Die durchgezogenen Kanten bestimmen die Terme der Taylor-Approximation und die gestrichelten Kanten die Terme des entsprechenden Rests.

## 11.1 Die Notation von Wagner und Platen

Die obigen Taylor-Entwicklungen und Taylor-Bäume enthalten sukzessive Mehrfach-iterierte Integrale der Form

$$\begin{aligned} & \int_0^t f^{i_1}(s, X_s) dW_s^{i_1} \\ & \int_0^t \int_0^s L^{i_1} f^{i_2}(\tau, X_\tau) dW_\tau^{i_1} dW_s^{i_2} \\ & \int_0^t \int_0^s \int_0^\tau L^{i_1} L^{i_2} f^{i_3}(p, X_p) dW_p^{i_1} dW_\tau^{i_2} dW_s^{i_3} \end{aligned}$$

oder mit konstanter Integrandfunktion  $\equiv 1$ .

Im Allgemeinen

$$\underbrace{\int_0^t \int_0^{s_1} \dots \int_0^{s_{l-1}}}_{l\text{-fach}} L^{i_1} \dots L^{i_{l-1}} f^{i_l}(s_1, X_{s_1}) dW_{s_1}^{i_1} \dots dW_{s_{l-1}}^{i_{l-1}} dW_{s_l}^{i_l}$$

Bemerkungen: Die Integrandfunktionen sind auch gleich

$$L^{i_1} \dots L^{i_{l-1}} L^{i_l} id_X(\cdot, X),$$

weil  $L^{i_l} id_X = f^{i_l}$ , wobei  $id_X$  die Identitätsfunktion  $f(x) \equiv x$  ist.

d.h. wir betrachten Integrale der Form

$$\int_0^t \int_0^{s_1} \dots \int_0^{s_{l-1}} L^{i_1} \dots L^{i_{l-1}} L^{i_l} id_X(s_1, X_{s_1}) dW_{s_1}^{i_1} \dots dW_{s_{l-1}}^{i_{l-1}} dW_{s_l}^{i_l}$$

Der  $l$ -dimensionale Vektor  $(i_1, \dots, i_l)$  enthält wichtige Informationen über Art und Folge der Integrale.

### 11.1.1 Multi-Indizes

Ein  $l$ -dimensionaler Vektor  $\alpha = (i_1, \dots, i_l)$  mit  $i_1, \dots, i_l \in \{0, 1\}$  und  $l \geq 1$  heißt Multi-Index der Länge  $l(\alpha) = l$ .

Definiere

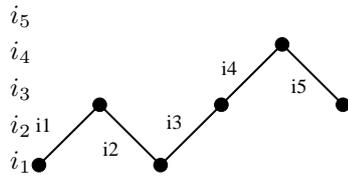
- (1) die Menge  $\mathcal{M}_1^{(l)}$  aller Multi-Indizes der Länge  $l \in \{1, 2, \dots\} = \mathbb{N}$ .

- (2) der Leer-Index  $\emptyset$  der Länge  $l(\emptyset) = 0$ .
- (3) die Menge  $\mathcal{M}_1$  aller Multi-Indizes aller möglichen Längen  $l = 0, 1, 2, \dots$

$$\mathcal{M}_1 = \{\emptyset\} \cup \bigcup_{l=1}^{\infty} \mathcal{M}_1^{(l)}$$

Ein Multi-Index  $\alpha = (i_1, \dots, i_l)$  der Länge  $l \geq 1$  beschreibt

PSfrag replacements (1) einen Ast eines Taylor-Baums



hier  
 $(i_1, i_2, i_3, i_4, i_5) = (1, 0, 0, 1, 0)$

- (2) ein  $l$ -faches Integral der Taylor-Entwicklung,

- (i) falls die letzte Kante des Astes des Taylor-Baums gestrichelt ist

$$\int_0^t \int_0^{s_1} \dots \int_0^{s_{l-1}} L^{i_1} \dots L^{i_{l-1}} L^{i_l} id_X(s_l, X_{s_l}) dW_{s_1}^{i_1} \dots dW_{s_{l-1}}^{i_{l-1}} dW_{s_l}^{i_l}$$

- (ii) falls die letzte Kante des Astes des Taylor-Baums durchgezogen ist

$$L^{i_1} \dots L^{i_{l-1}} L^{i_l} id_X(0, X_0) \cdot \int_0^t \int_0^{s_1} \dots \int_0^{s_{l-1}} dW_{s_1}^{i_1} \dots dW_{s_{l-1}}^{i_{l-1}} dW_{s_l}^{i_l}$$

Wir definieren die folgenden Operatoren auf  $\mathcal{M}_1 \setminus \{\emptyset\}$ , d.h., für  $\alpha = (i_1, \dots, i_l) \in \mathcal{M}_1 \setminus \{\emptyset\}$  mit Länge  $l(\alpha) \geq 1$

$$(1) -\alpha = -(i_1, \dots, i_l) = \begin{cases} \emptyset & \text{falls } l(\alpha) = 1 \\ (i_2, \dots, i_l) & \text{falls } l(\alpha) \geq 2 \end{cases}$$

$$(2) \alpha- = (i_1, \dots, i_l)- = \begin{cases} \emptyset & \text{falls } l(\alpha) = 1 \\ (i_1, \dots, i_{l-1}) & \text{falls } l(\alpha) \geq 2 \end{cases}.$$

Merke  $l(-\alpha) = l(\alpha-) = l(\alpha) - 1$

### 11.1.2 Iterierte Differential-Operatoren

Sei  $\alpha \in \mathcal{M}_1$  und definiere

$$L^\alpha U = \begin{cases} U & \text{falls } \alpha = \emptyset \\ L^{i_1} L^{-\alpha} U & \text{falls } l(\alpha) \geq 1 \end{cases}$$

wobei  $\alpha = (i_1, \dots, i_l)$  in dem zweiten Fall

$$\Rightarrow \boxed{L^{(i_1, \dots, i_l)} = L^{i_1} L^{i_2} \dots L^{i_l}}$$

$\Rightarrow L^\alpha id_X(\cdot, X)$  ist die Integrandfunktion des durch  $\alpha$  beschriebenen Mehrfach-Integrals.

### 11.1.3 Iterierte Integrale

Sei  $\alpha \in \mathcal{M}_1$  mit  $\alpha = (i_1, \dots, i_l)$ , falls  $l(\alpha) = l \geq 1$ .

Wir definieren das durch  $\alpha$  beschriebene iterierte Mehrfach-Integral einer Funktion  $h$  durch

$$I_{\alpha,0,t}[h(\cdot)] = \begin{cases} h(t) & \text{falls } \alpha = \emptyset \\ \int_0^t I_{\alpha-,0,s_l}[h(\cdot)] dW_{s_l}^{i_l} & \text{falls } l(\alpha) \geq 1 \end{cases}$$

Dann erhalten wir, falls  $l(\alpha) \geq 1$ ,

$$I_{\alpha,0,t}[L^\alpha id_X(\cdot, X)] = \int_0^t \int_0^{s_l} \dots \int_0^{s_2} L^{i_1} \dots L^{i_{l-1}} L^{i_l} id_X(s_1, X_{s_1}) dW_{s_1}^{i_1} \dots dW_{s_{l-1}}^{i_{l-1}} dW_{s_l}^{i_l}$$

$$I_{\alpha,0,t}[L^\alpha id_X(0, X_0)] = L^{i_1} \dots L^{i_{l-1}} L^{i_l} id_X(0, X_0) \cdot \int_0^t \int_0^{s_l} \dots \int_0^{s_2} dW_{s_1}^{i_1} \dots dW_{s_{l-1}}^{i_{l-1}} dW_{s_l}^{i_l}$$

und

$$I_{\emptyset,0,t}[id_X(0, X_0)] = id_X(0, X_0) = X_0,$$

falls  $\alpha = \emptyset$ .

### 11.1.4 Hierarchische Menge

Frage Welche  $\alpha$  beschreiben einen Taylor-Baum? Welche den Rest des Taylor-Baums?

Eine Teilmenge  $\mathcal{A}$  von  $\mathcal{M}_1$  heißt hierarchische Menge, falls sie die folgenden Eigenschaften besitzt

- (1)  $\mathcal{A}$  ist nichtleer – mindestens  $\emptyset \in \mathcal{A}$
- (2)  $\mathcal{A}$  ist gleichmäßig beschränkt  $\sup_{\alpha \in \mathcal{A}} l(\alpha) < \infty$
- (3)  $\mathcal{A}$  ist hierarchisch, d.h.  $-\alpha \in \mathcal{A}$  für alle  $\alpha \in \mathcal{A} \setminus \{\emptyset\}$

Die Menge

$$\mathcal{B}(\mathcal{A}) := \{\alpha \in \mathcal{M}_1 \setminus \mathcal{A} : -\alpha \in \mathcal{A}\}$$

heißt Restmenge der hierarchischen Menge  $\mathcal{A}$

Beispiele

(1) Ito-SDGL  $\mathcal{A} = \{\emptyset\}$  mit  $\mathcal{B}(\mathcal{A}) = \{(0), (1)\}$

(2) Euler-Approximation  $\mathcal{A} = \{\emptyset, (0), (1)\}$  mit

$$\mathcal{B}(\mathcal{A}) = \{(0, 0), (1, 0), (0, 1), (1, 1)\}$$

(3) Milstein-Approximation  $\mathcal{A} = \{\emptyset, (0), (1), (1, 1)\}$  mit

$$\mathcal{B}(\mathcal{A}) = \{(0, 0), (1, 0), (0, 1), (0, 1, 1), (1, 1, 1)\}$$

Definition Die  $\mathcal{A}$ -Taylor-Entwicklung für eine hierarchische Menge  $\mathcal{A}$  mit Restmenge  $\mathcal{B}(\mathcal{A})$  lautet

$$X_t = \sum_{\alpha \in \mathcal{A}} L^\alpha id_X(0, X_0) \cdot I_{\alpha, 0, t}[1] + \sum_{\alpha \in \mathcal{B}(\mathcal{A})} I_{\alpha, 0, t}[L^\alpha id_X(\cdot, X.)]$$

Der erste Summe hier ist die  $\mathcal{A}$ -Taylor-Approximation und die zweite Summe der Rest der  $\mathcal{A}$ -Taylor-Entwicklung.

## Kapitel 12

# Beispiele stochastischer Taylor-Verfahren

Wir haben eine allgemeine Darstellung einer stochastischen Taylor-Entwicklung, nämlich

$$X_t = \sum_{\alpha \in \mathcal{A}} L^\alpha id_X(0, X_0) \cdot I_{\alpha,0,t}[1] \quad \underline{\text{Taylor-Approximation}}$$
$$+ \sum_{\alpha \in \mathcal{B}(\mathcal{A})} I_{\alpha,0,t}[L^\alpha id_X(\cdot, X_\cdot)] \quad \underline{\text{Rest}}$$

für jede hierarchische Menge  $\mathcal{A}$  und deren entsprechende Restmenge  $\mathcal{B}(\mathcal{A})$ .

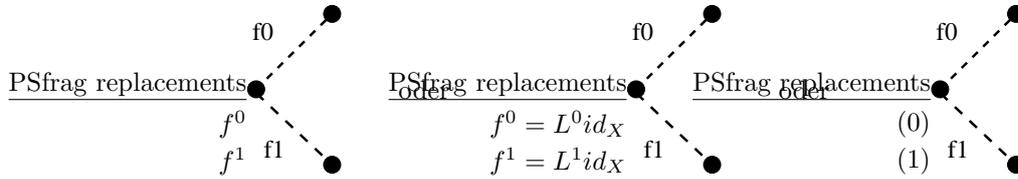
Wir haben auch eine äquivalente graphentheoretische Darstellung jeder stochastischen Taylor-Entwicklung durch einen Taylor-Baum mit durchgezogenen und gestrichelten Kanten.

Beispiel 1 Die ursprüngliche Ito-SDGL lautet

$$X_t = X_0 + \sum_{i=0}^1 \int_0^t f^i(s, X_s) dW_s^i$$

mit  $dt = dW_t^0$ ,  $f = f^0$ ,  $dW_t = dW_t^1$  und  $g = f^1$ .

Die Graphendarstellung lautet



wobei eine gestrichelte Kante einen Restterm darstellt. (Hier:  $L^i id_X = f^i$ )

Wir können die entsprechende hierarchische Menge  $\mathcal{A}$  und deren Restmenge  $\mathcal{B}(\mathcal{A})$  von dem letzten Graphen ablesen.

$$\begin{aligned} \mathcal{A} &= \{\emptyset\} \quad \text{nur } \emptyset, \text{ weil der Graph keine durchgezogenen Kanten besitzt} \\ \mathcal{B}(\mathcal{A}) &= \{(0), (1)\} \quad \text{d.h. die gestrichelten Kanten} \end{aligned}$$

Die Taylor-Entwicklung lautet

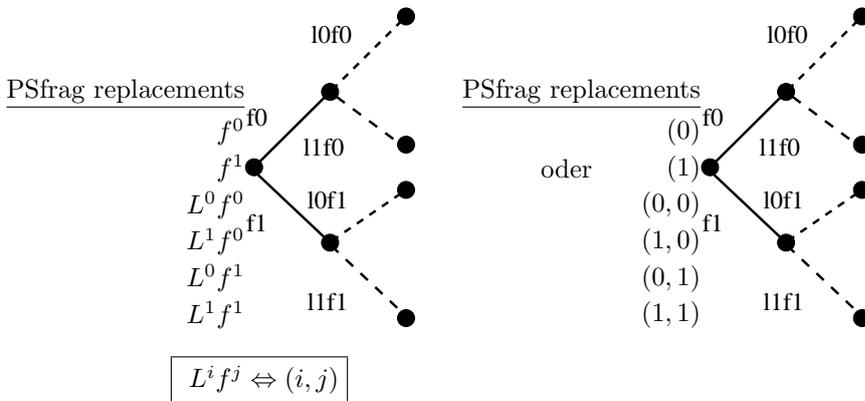
$$\begin{aligned} X_t &= L^0 id_X(0, X_0) I_{\emptyset, 0, t}[1] + \sum_{i=0}^1 I_{(i), 0, t}[L^{(i)} id_X(\cdot, X_0)] \\ &= X_0 + \sum_{i=0}^1 \int_0^t f^i(s, X_s) dW_s^i \end{aligned}$$

d.h. die ursprüngliche Ito-SDGL !

Beispiel 2 Wenn wir die beiden Integrandfunktionen  $f^0 = L^0 id_X$  und  $f^1 = L^1 id_X$  durch die Ito-Formel entwickeln, erhalten wir die Taylor-Entwicklung

$$\begin{aligned} X_t &= X_0 + \sum_{i=0}^1 f^i(0, X_0) \int_0^t dW_s^i \quad \text{Taylor-Approximation} \\ &\quad + \sum_{i=0}^1 \sum_{j=0}^1 \int_0^t \int_0^s L^i f^j(\tau, X_\tau) dW_\tau^i dW_s^j \quad \text{Rest} \end{aligned}$$

Die Graphendarstellungen lauten



Die hierarchische Menge  $\mathcal{A}$  hier entspricht den durchgezogenen Kanten, d.h.

$$\mathcal{A} = \{\emptyset, (0), (1)\},$$

und die Restmenge  $\mathcal{B}(\mathcal{A})$  den gestrichelten Kanten, d.h.

$$\mathcal{B}(\mathcal{A}) = \{(0, 0), (0, 1), (1, 0), (1, 1)\}$$

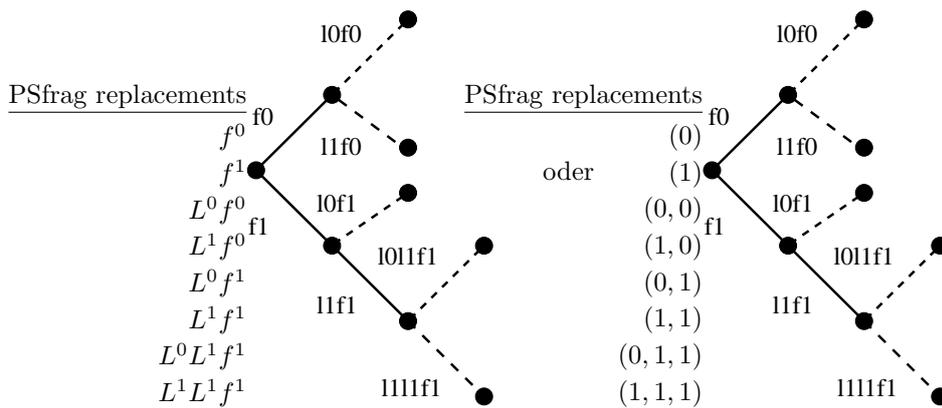
Hier gilt

$$\begin{aligned} L^{(i,j)}id_X(\tau, X_\tau) &= L^i[L^{(j)}id_X(\tau, X_\tau)] && (j) = -(i, j) \\ &= L^iL^j[L^\emptyset id_X(\tau, X_\tau)] && \emptyset = -(j) \\ &= L^iL^j id_X(\tau, X_\tau) && L^\emptyset h \equiv h \\ &= L^i f^j(\tau, X_\tau) && L^j id_X = f^j \end{aligned}$$

Daher lautet

$$\begin{aligned} I_{(i,j),0,t}[L^{(i,j)}id_X(\cdot, X_\cdot)] &= \int_0^t I_{(i),0,s}[L^{(i,j)}id_X(\cdot, X_\cdot)] dW_s^j && (i, j) - = (i) \\ &= \int_0^t \left[ \int_0^s I_{\emptyset,0,\tau}[L^{(i,j)}id_X(\cdot, X_\cdot)] dW_\tau^i \right] dW_s^j \\ &= \int_0^t \int_0^s L^{(i,j)}id_X(\tau, X_\tau) dW_\tau^i dW_s^j \\ &= \int_0^t \int_0^s L^i f^j(\tau, X_\tau) dW_\tau^i dW_s^j && \text{von oben} \end{aligned}$$

Beispiel 3 Wir erhalten die Milstein-Approximation aus der Euler-Approximation (Beispiel 2) durch Verwendung der Ito-Formel auf  $U = L^1 f^1$ . Der neue Taylor-Baum lautet



und daher lauten die hierarchische Menge und die Restmenge

$$\mathcal{A} = \{\emptyset, (0), (1), (1, 1)\}$$

$$\mathcal{B}(\mathcal{A}) = \{(0, 0), (0, 1), (1, 0), (0, 1, 1), (1, 1, 1)\}$$

Frage: Welche hierarchische Menge entspricht einer Approximation starker Ordnung  $\gamma$  oder schwacher Ordnung  $\beta$ ?

Wir werden eine  $\mathcal{A}$ -Taylor-Approximation auf einem Intervall  $[t_n, t_{n+1}]$  benutzen, d.h.

$$X_{t_{n+1}} \approx \sum_{\alpha \in \mathcal{A}} L^\alpha id_X(t_n, X_{t_n}) I_{\alpha, t_n, t_{n+1}}[1]$$

um das  $\mathcal{A}$ -Taylor-Verfahren, d.h.

$$Y_{n+1} = \sum_{\alpha \in \mathcal{A}} L^\alpha id_X(t_n, Y_n) I_{\alpha, t_n, t_{n+1}}[1]$$

zu definieren.

## 12.1 Schwache Taylor-Verfahren

Das Taylor-Verfahren der hierarchischen Menge

$$\Gamma_\beta := \{\alpha \in \mathcal{M}_1 : 0 \leq l(\alpha) \leq \beta\},$$

mit  $\beta = 1, 2, 3, \dots$ , besitzt die schwache Konvergenzordnung  $\beta$ . Es heißt schwaches Taylor-Verfahren von Ordnung  $\beta$  oder Taylor-Verfahren schwacher Ordnung  $\beta$ .

Beispiel:  $\beta = 1$

$$\Gamma_1 = \{\alpha \in \mathcal{M}_1 : 0 \leq l(\alpha) \leq 1\} = \{\emptyset, (0), (1)\}$$

$$\begin{aligned} Y_{n+1} &= \sum_{\alpha \in \Gamma_1} L^\alpha id_X(t_n, Y_n) I_{\alpha, t_n, t_{n+1}}[1] \\ &= \underbrace{Y_n}_{\alpha=\emptyset} + \underbrace{f^0(t_n, Y_n) \int_{t_n}^{t_{n+1}} dW_t^0}_{\alpha=(0)} + \underbrace{f^1(t_n, Y_n) \int_{t_n}^{t_{n+1}} dW_t^1}_{\alpha=(1)} \end{aligned}$$

d.h.  $\boxed{Y_n + f^0(t_n, Y_n) \Delta_n + f^1(t_n, Y_n) \cdot \Delta W_n}$  das Euler-Verfahren !

mit

$$\Delta_n = \int_{t_n}^{t_{n+1}} dW_t^0 = t_{n+1} - t_n, \quad \Delta W_n = \int_{t_n}^{t_{n+1}} dW_t^1 = W_{t_{n+1}} - W_{t_n}.$$

## 12.2 Starke Taylor-Verfahren

Die hierarchischen Mengen der stark konvergierenden Taylor-Verfahren sind etwas komplizierter. Sei  $\alpha = (i_1, \dots, i_l) \in \mathcal{M}_1 \setminus \{\emptyset\}$  und definiere

$$\boxed{n(\alpha) = \#\{i_j = 0\}} = \text{die \underline{Anzahl} der Komponenten von } \alpha, \text{ die gleich } 0 \text{ sind}$$

Definiere auch  $n(\emptyset) = 0$ . Offensichtlich gilt:  $0 \leq n(\alpha) \leq l(\alpha)$ .

Die Menge

$$\Lambda_\gamma := \left\{ \alpha \in \mathcal{M}_1 : l(\alpha) + n(\alpha) \leq 2\gamma \quad \text{oder} \quad l(\alpha) = n(\alpha) = \gamma + \frac{1}{2} \right\}$$

ist eine hierarchische Menge für  $\gamma = \frac{1}{2}, 1, \frac{3}{2}, 2, \frac{5}{2}, \dots$ .

Das  $\Lambda_\gamma$ -Taylor-Verfahren besitzt starke Konvergenzordnung  $\gamma$ . Es heißt starkes Taylor-Verfahren der Ordnung  $\gamma$  oder Taylor-Verfahren starker Ordnung  $\gamma$ .

Beispiel 1     $\boxed{\gamma = \frac{1}{2}}$

$$\begin{aligned} \Lambda_{\frac{1}{2}} &= \{ \alpha \in \mathcal{M}_1 : n(\alpha) + l(\alpha) \leq 1 \quad \underline{\text{oder}} \quad n(\alpha) = l(\alpha) = 1 \} \\ &= \{ \emptyset, (0), (1) \} \end{aligned}$$

weil

- (a)  $\alpha = \emptyset$     :     $n(\emptyset) + l(\emptyset) = 0 + 0 = 0 \leq 1$
- (b)  $\alpha = (0)$     :     $n((0)) + l((0)) = 1 + 1 = 2 > 1$     aber     $n((0)) = l((0)) = 1$
- (c)  $\alpha = (1)$     :     $n((1)) + l((1)) = 0 + 1 = 1$ .

$\Rightarrow$  das starke Taylor-Verfahren von Ordnung  $\gamma = \frac{1}{2}$  ist das Euler-Verfahren!

Beispiel 2     $\boxed{\gamma = 1}$      $\Rightarrow$      $\Lambda_1 = \{ \alpha \in \mathcal{M}_1 : n(\alpha) + l(\alpha) \leq 2 \}$ .

Die Bedingung  $n(\alpha) = l(\alpha) = \frac{3}{2}$  ist hier nicht möglich und fällt weg, d.h.,

$$\Lambda_1 = \{ \emptyset, (0), (1), (1, 1) \}$$

weil

$$\begin{aligned} n(\emptyset) + l(\emptyset) &= 0 + 0 \leq 2 \\ n((0)) + l((0)) &= 1 + 1 = 2 \\ n((1)) + l((1)) &= 0 + 1 = 1 \leq 2 \\ n((1, 1)) + l((1, 1)) &= 0 + 2 = 2 \end{aligned}$$

und kein anderer Multi-Index genügt der obigen Bedingung.

Das starke Taylor-Verfahren von Ordnung  $\gamma = 1$  ist das Milstein-Verfahren

$$\begin{aligned} Y_{n+1} &= \sum_{\alpha \in \Lambda_1} L^\alpha \text{id}_X(t_n, Y_n) I_{\alpha, t_n, t_{n+1}}[1] \\ &= Y_n + f^0(t_n, Y_n) \int_{t_n}^{t_{n+1}} dW_t^0 \\ &\quad + f^1(t_n, Y_n) \int_{t_n}^{t_{n+1}} dW_t^1 + L^1 f^1(t_n, Y_n) \int_{t_n}^{t_{n+1}} \int_{t_n}^t dW_s^1 dW_t^1 \end{aligned}$$

d.h. das Milstein-Verfahren lautet

$$Y_{n+1} = Y_n + f^0(t_n, Y_n) \Delta_n + f^1(t_n, Y_n) \Delta W_n + \frac{1}{2} L^1 f^1(t_n, Y_n) \cdot \{(\Delta W_n)^2 - \Delta_n\}$$

weil

$$\Delta_n = \int_{t_n}^{t_{n+1}} dW_t^0 = \int_{t_n}^{t_{n+1}} dt = t_{n+1} - t_n$$

$$\Delta W_n = \int_{t_n}^{t_{n+1}} dW_t^1 = \int_{t_n}^{t_{n+1}} dW_t = W_{t_{n+1}} - W_{t_n}$$

und

$$\int_{t_n}^{t_{n+1}} \int_{t_n}^t dW_s^1 dW_t^1 = \int_{t_n}^{t_{n+1}} \int_{t_n}^t dW_s dW_t = \frac{1}{2} (\Delta W_n)^2 - \frac{1}{2} \Delta_n.$$

# Kapitel 13

## Vektorwertige SDGLen

Jetzt betrachten wir den allgemeinen Fall einer vektorwertigen SDGL mit

- (1)  $d$ -dimensionalem Zustandsvektor

$$X = (X^1, \dots, X^d)^\top$$

- (2)  $m$ -dimensionalem Wiener-Prozess

$$W_t = (W_t^1, \dots, W_t^m)^\top$$

d.h., die  $W_t^i$  sind skalare Wiener-Prozesse, die paarweise unabhängig sind – insbesondere gilt

$$\mathbb{E} \left( W_t^i W_t^j \right) = t \delta_{i,j} \quad i, j = 1, \dots, m$$

mit dem Kronecker-delta Symbol

$$\delta_{i,j} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

Die vektorwertige Ito-SDGL lautet

$$\boxed{dX_t = f(t, X_t) dt + \sum_{j=1}^m g_j(t, X_t) dW_t^j} \quad (*)$$

wobei

$$f(t, X) = \begin{pmatrix} f^1(t, X^1, \dots, X^d) \\ \vdots \\ f^d(t, X^1, \dots, X^d) \end{pmatrix}$$
$$g_j(t, X) = \begin{pmatrix} g_j^1(t, X^1, \dots, X^d) \\ \vdots \\ g_j^d(t, X^1, \dots, X^d) \end{pmatrix}, \quad j = 1, \dots, m$$

d.h. mit den Abbildungen  $f : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ ,  $g_j : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ ,  $j = 1, \dots, m$

ODER komponentenweise als ein System von SDGLen

$$\boxed{dX_t^i = f^i(t, X_t^1, \dots, X_t^d) dt + \sum_{j=1}^m g_j^i(t, X_t^1, \dots, X_t^d) dW_t^j} \quad i = 1, \dots, d$$

das wir als folgendes System stochastischer Integralgleichungen interpretieren:

$$X_t^i = X_{t_0}^i + \underbrace{\int_{t_0}^t f^i(s, X_s^1, \dots, X_s^d) ds}_{\text{Riemann-Integrale}} + \underbrace{\sum_{j=1}^m \int_{t_0}^t g_j^i(s, X_s^1, \dots, X_s^d) dW_s^j}_{\text{Ito-Integrale}}$$

for  $i = 1, \dots, d$

Beispiel Wir können die gewöhnliche DGL zweiter Ordnung mit weißem Rauschen

$$\frac{d^2x}{dt^2} + a \frac{dx}{dt} + bx = \zeta_t \quad (a, b \text{ Konstanten})$$

umschreiben als das 2-dimensionale System

$$\begin{aligned} \frac{dX_1}{dt} &= X_2 \\ \frac{dX_2}{dt} &= -bX_1 - aX_2 + \zeta_t \end{aligned}$$

wobei  $X_1 = x$  und  $X_2 = \frac{dx}{dt}$

Das entsprechende System von Ito-SDGLen lautet

$$\begin{aligned} dX_t^1 &= X_t^2 dt \\ dX_t^2 &= (-bX_t^1 - aX_t^2) dt + dW_t^1 \end{aligned}$$

d.h. die 2-dimensionale Ito-SDGL mit 1-dimensionalem Wiener-Prozess  $W_t^1$  (hier  $d = 2, m = 1$ ):

$$d \begin{pmatrix} X_t^1 \\ X_t^2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -b & -a \end{pmatrix} \begin{pmatrix} X_t^1 \\ X_t^2 \end{pmatrix} dt + \begin{pmatrix} 0 \\ 1 \end{pmatrix} dW_t^1$$

Die entsprechende vektorwertige Stratonovich-SDGL einer vektorwertigen Ito-SDGL (\*) lautet

$$\boxed{dX_t = \underline{f}(t, X_t) dt + \sum_{j=1}^m g_j(t, X_t) \circ dW_t^j}$$

mit dem modifizierten Driftterm  $\underline{f}(t, X) = \left( \underline{f}^1(t, X), \dots, \underline{f}^d(t, X) \right)^\top$ , wobei

$$\underline{f}^i(t, x_1, \dots, x_d) = f^i(t, x_1, \dots, x_d) - c^i(t, x_1, \dots, x_d)$$

mit

$$c^i(t, x_1, \dots, x_d) = \frac{1}{2} \sum_{k=1}^d \sum_{l=1}^m g_l^k(t, x_1, \dots, x_d) \cdot \frac{\partial g_l^i}{\partial x_k}(t, x_1, \dots, x_d)$$

for  $i = 1, \dots, d$

Sonderfall: Additives Rauschen  $\frac{\partial}{\partial x_k} g_l^i(t, x_1, \dots, x_d) \equiv 0 \quad \forall i, k, l$

$$\Rightarrow g_l^i(t, x_1, \dots, x_d) \equiv g_l^i(t) \quad \forall i, l,$$

$$\Rightarrow \underline{f}(t, x_1, \dots, x_d) \equiv f(t, x_1, \dots, x_d)$$

In diesem Fall haben die entsprechenden Ito- und Stratonovich-SDGLEN die selben Koeffizienten.

In dem vektorwertigen Fall können die entsprechenden Ito- und Stratonovich-SDGLEN identisch sein, auch wenn das Rauschen nicht additiv ist. z.B Seien  $d = 2, m = 1, f$  beliebig, und

$$g_1(t, x_1, x_2) = \begin{pmatrix} g_1^1(t, x_1, x_2) \\ g_1^2(t, x_1, x_2) \end{pmatrix} = \begin{pmatrix} x_1 + x_2 \\ -x_1 - x_2 \end{pmatrix}$$

$$\Rightarrow c^1(t, x_1, x_2) \equiv c^2(t, x_1, x_2) \equiv 0 \quad \Rightarrow \underline{f} \equiv f$$

### 13.1 Die Ito-Formel

Sei  $X_t = (X_t^1, \dots, X_t^d)^\top$  eine Lösung der  $d$ -dimensionalen Ito-SDGL

$$dX_t = f(t, X_t) dt + \sum_{j=1}^m g_j(t, X_t) dW_t^j,$$

wobei  $W_t = (W_t^1, \dots, W_t^m)^\top$  ein  $m$ -dimensionaler Wiener-Prozess ist.

Sei  $U : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^1$  und definiere  $Y_t = U(t, X_t)$ .

Dann lautet die Ito-Formel:

$$dY_t = L^0 U(t, X_t) dt + \sum_{j=1}^m L^j U(t, X_t) dW_t^j$$

oder die entsprechenden Integralversion

$$U(t, X_t) = U(t_0, X_{t_0}) + \int_{t_0}^t L^0 U(s, X_s) ds + \sum_{j=1}^m \int_{t_0}^t L^j U(s, X_s) dW_s^j,$$

wobei

$$L^0 U = \frac{\partial U}{\partial t} + \sum_{k=1}^d f^k \frac{\partial U}{\partial x_k} + \frac{1}{2} \sum_{k,l=1}^d \sum_{j=1}^m g_j^k g_j^l \frac{\partial^2 U}{\partial x_k \partial x_l}$$

$$L^j U = \sum_{k=1}^d g_j^k \frac{\partial U}{\partial x_k}, \quad j = 1, \dots, m.$$

Beispiel 1 Wir haben schon gesehen, dass

$$d(W_t)^2 = 1 dt + 2W_t dW_t$$

Aber sind  $W_t^1$  und  $W_t^2$  unabhängige Wiener Prozesse, dann gilt

$$d(W_t^1 W_t^2) = W_t^1 dW_t^2 + W_t^2 dW_t^1$$

Beweis Benutze  $Y_t = X_t^1 X_t^2 = U(t, X_t^1, X_t^2)$  mit  $U(t, X_1, X_2) = X_1 X_2$  wobei  $X_t^1 = W_t^1, X_t^2 = W_t^2$ , d.h.

$$d \begin{pmatrix} X_t^1 \\ X_t^2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} dt + \begin{pmatrix} 1 \\ 0 \end{pmatrix} dW_t^1 + \begin{pmatrix} 0 \\ 1 \end{pmatrix} dW_t^2$$

d.h. mit

$$f = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad g_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad g_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

und deswegen sind

$$L^0 U = 0 + \sum_{i=1}^2 0 \cdot \frac{\partial U}{\partial x_i} + \frac{1}{2} \cdot 0 \equiv 0, \quad L^1 U = X_t^2, \quad L^2 U = X_t^1$$

$$\Rightarrow d(X_t^1 X_t^2) = X_t^2 dW_t^1 + X_t^1 dW_t^2$$

Beispiel 2 Wir können auch vektorwertige  $Y_t = U(t, X_t)$  mit  $U : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^n$  betrachten. Alles geht komponentenweise durch. Wir verwenden die Ito-Formel komponentenweise, d.h. für  $Y_t^p = U^p(t, X_t), p = 1, \dots, n$ .

z.B. Für  $Y_t \equiv X_t$  haben wir  $Y_t^p = U^p(t, X_t) \equiv X_t^p, p = 1, \dots, n (= d \text{ hier})$ .

$$\Rightarrow dY_t^p = dX_t^p = L^0 U^p dt + \sum_{j=1}^m L^j U^p dW_t^j$$

$$\Rightarrow dX_t^p = f^p(t, X_t) dt + \sum_{j=1}^m g_j^p(t, X_t) dW_t^j \quad \text{die ursprüngliche SDGL !}$$

## 13.2 Die stochastische Taylor-Entwicklung

Wir betrachten eine  $d$ -dimensionale Ito-SDGL mit einem  $m$ -dimensionalen Wiener-Prozess und schreiben diese SDGL jetzt in der Form

$$dX_t = \sum_{j=0}^m g_j(t, X_t) dW_t^j$$

d.h. mit  $g_0(t, X) \equiv f(t, X)$  und  $dW_t^0 \equiv dt$ .

Die Ito-Formel für  $Y_t = U(t, X_t)$ , wobei  $U : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^1$ , lautet jetzt

$$dY_t = \sum_{j=0}^m L^j U(t, X_t) dW_t^j$$

oder in Integralform

$$U(t, X_t) = U(t_0, X_{t_0}) + \sum_{j=0}^m \int_{t_0}^t L^j U(s, X_s) dW_s^j$$

$\Rightarrow$  der Taylor-Baum besitzt nun  $m + 1$  Äste

$\Rightarrow$  die Multi-Indize haben jetzt Komponenten  $i_j \in \{0, 1, \dots, m\}$ .

Sei  $\mathcal{M}_m$  die Menge der Multi-Indize, d.h.,

$$\mathcal{M}_m = \{\emptyset\} \cup \{\alpha = (i_1, \dots, i_l) : i_j \in \{0, 1, \dots, m\}, j = 1, \dots, l, l = l(\alpha) = 1, 2, \dots\}$$

$\Rightarrow$  jedes Mal, wenn wir die Ito-Formel verwenden, erhalten wir jetzt  $m + 1$  neue Restterme – statt 2 wie vorher.

Sonst ist fast alles identisch wie in dem skalaren Fall, z.B., die Definitionen und Eigenschaften der hierarchischen Mengen, Restmengen, der iterierten Operatoren und Multi-Integrale.

Sei  $\mathcal{A}$  eine hierarchische Menge in  $\mathcal{M}_m$  mit Restmenge  $\mathcal{B}(\mathcal{A})$  und sei  $U : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^1$  oft genug stetig differenzierbar. Dann lautet die  $\mathcal{A}$ -Taylor Entwicklung des Prozesses  $Y_t = U(t, X_t)$ :

$$U(t, X_t) = \sum_{\alpha \in \mathcal{A}} L^\alpha U(t_0, X_{t_0}) I_{\alpha, t_0, t}[1] + \sum_{\alpha \in \mathcal{B}(\mathcal{A})} I_{\alpha, t_0, t}[L^\alpha U(\cdot, X)]$$

Mit  $U(t, X) \equiv X^i \equiv id_X^i$  erhalten wir die  $i$ -te Komponenten der  $\mathcal{A}$ -Taylor-Entwicklung für den Prozess  $X_t$ ,

$$X_t^i = \sum_{\alpha \in \mathcal{A}} L^\alpha id_X^i(t_0, X_{t_0}) I_{\alpha, t_0, t}[1] + \sum_{\alpha \in \mathcal{B}(\mathcal{A})} I_{\alpha, t_0, t}[L^\alpha id_X^i(\cdot, X)]$$

für  $i = 1, \dots, d$ .

Verwenden wir diese Entwicklung auf einem Teilintervall  $[t_n, t_{n+1}]$  und schneiden wir den Rest ab, dann erhalten wir das  $\mathcal{A}$ -Taylor-Verfahren

$$Y_{n+1}^i = \sum_{\alpha \in \mathcal{A}} L^\alpha id_X^i(t_n, Y_n) I_{\alpha, t_n, t_{n+1}}[1], \quad i = 1, \dots, d,$$

wobei  $Y_n = (Y_n^1, \dots, Y_n^d)^\top$ .

Die hierarchischen Menge für starke Konvergenz mit Ordnung  $\gamma = \frac{1}{2}, 1, \frac{3}{2}, 2, \frac{5}{2}, \dots$  und schwache Konvergenz mit Ordnung  $\beta = 1, 2, 3, \dots$  sind in dem Fall  $m \geq 1$ :

$$\Lambda_\gamma = \left\{ \alpha \in \mathcal{M}_m : l(\alpha) + n(\alpha) \leq 2\gamma \text{ oder } l(\alpha) = n(\alpha) = \gamma + \frac{1}{2} \right\}$$

und

$$\Gamma_\beta = \{ \alpha \in \mathcal{M}_m : 0 \leq l(\alpha) \leq \beta \}$$

Beispiel  $d = 1$  (d.h.,  $x$  skalar!),  $m = 2$  and  $\gamma = 1$

$$\begin{aligned} \Lambda_1 &= \{ \alpha \in \mathcal{M}_2 : l(\alpha) + n(\alpha) \leq 2 \} \\ &= \{ \emptyset, (0), (1), (2), (1, 1), (1, 2), (2, 1), (2, 2) \} \end{aligned}$$

Die Bedingung  $l(\alpha) = n(\alpha) = \frac{3}{2}$  ist bedeutungslos hier.

Das  $\Lambda_1$ -Taylor-Verfahren ist das Milstein-Verfahren

$$\begin{aligned} Y_{n+1} &= \sum_{\alpha \in \Lambda_1} L^\alpha id_X(t_n, Y_n) \cdot I_{\alpha, t_n, t_{n+1}}[1] \\ &= Y_n \quad [\alpha = \emptyset] \\ &\quad + \sum_{i=0}^2 L^i id_X(t_n, Y_n) I_{(i), t_n, t_{n+1}}[1] \quad [\alpha = (0), (1), (2)] \\ &\quad + \sum_{i,j=1}^2 L^{(i,j)} id_X(t_n, Y_n) I_{(i,j), t_n, t_{n+1}}[1] \quad [\alpha = (1, 1), (1, 2), (2, 1), (2, 2)] \end{aligned}$$

oder

$$Y_{n+1} = Y_n + \sum_{i=0}^2 g_i(t_n, Y_n) \int_{t_n}^{t_{n+1}} dW_t^i + \sum_{i,j=1}^2 g_i(t_n, Y_n) \frac{\partial g_j}{\partial x}(t_n, Y_n) \int_{t_n}^{t_{n+1}} \int_{t_n}^t dW_s^i dW_t^j$$

weil

$$L^0 id_X \equiv id_X, \quad L^i id_X = g_i, \quad L^{(i,j)} id_X = L^i(L^j id_X) = L^i g_j = g_i \frac{\partial g_j}{\partial x}$$

Wir haben

$$\Delta_n = \int_{t_n}^{t_{n+1}} dt, \quad \Delta W_n^i = \int_{t_n}^{t_{n+1}} dW_t^i \quad (i = 1, 2)$$

und

$$\int_{t_n}^{t_{n+1}} \int_{t_n}^t dW_s^i dW_t^i = \frac{1}{2} (\Delta W_n^i)^2 - \frac{1}{2} \Delta_n, \quad i = 1, 2$$

ABER es gibt keine einfache Formel für

$$\int_{t_n}^{t_{n+1}} \int_{t_n}^t dW_s^i dW_t^j \quad \text{falls } i \neq j.$$

## Kapitel 14

# Approximation stochastischer Multi-Integrale

Stochastische Taylor-Verfahren höherer Ordnung enthalten stochastische Multi-Integrale, die wir auswerten müssen, um das Verfahren zu verwenden. Diese Aufgabe ist nicht immer einfach, z.B., das Milstein-Verfahren für eine SDGL mit  $m \geq 2$  unabhängigen Wiener-Prozessen  $W_t^1, \dots, W_t^m$  enthält die Multi-Integrale

$$I_{(i_1, i_2), t_n, t_{n+1}} = \int_{t_n}^{t_{n+1}} \int_{t_n}^t dW_s^{i_1} dW_t^{i_2}, \quad i_1, i_2 \in \{1, \dots, m\}$$

Für  $i_1 = i_2 = j \in \{1, \dots, m\}$  haben wir die einfache Formel

$$\begin{aligned} I_{(j, j), t_n, t_{n+1}} &= \int_{t_n}^{t_{n+1}} \int_{t_n}^t dW_s^j dW_t^j \\ &= \frac{1}{2} \left( \int_{t_n}^{t_{n+1}} dW_t^j \right)^2 - \frac{1}{2} \int_{t_n}^{t_{n+1}} dt = \frac{1}{2} (\Delta W_n^j)^2 - \frac{1}{2} \Delta_n \end{aligned}$$

$\Rightarrow$  einfach im Sinne, dass sie nur die einfachen Zuwächse  $\Delta W_n^j$  und  $\Delta_n$  enthält.

Für den Fall  $i_1 \neq i_2$  gibt's keine einfache Formel dieser Art.

Versuchen wir das Multi-Integral  $I_{(i_1, i_2), t_n, t_{n+1}}$  durch  $\frac{1}{2} \Delta W_n^{i_1} \Delta W_n^{i_2}$  (oder etwas ähnliches) zu ersetzen, dann verlieren wir die behauptete starke Ordnung  $\gamma = 1$  des Milstein-Verfahrens.

Beispiel Betrachte die SDGL ( $d = m = 2$ )

$$dX_t^1 = dW_t^1, \quad dX_t^2 = X_t^1 dW_t^2$$

d.h

$$d \begin{pmatrix} X_t^1 \\ X_t^2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} dt + \begin{pmatrix} 1 \\ 0 \end{pmatrix} dW_t^1 + \begin{pmatrix} 0 \\ X_t^1 \end{pmatrix} dW_t^2$$

Die Lösung mit dem Anfangswert  $X_0^1 = X_0^2 = 0$  ist tatsächlich

$$X_t^1 = W_t^1, \quad X_t^2 = I_{(1,2),0,t} = \int_0^t W_s^1 dW_s^2$$

Das Milstein-Verfahren hier lautet

$$\begin{aligned} Y_{n+1}^1 &= Y_n^1 + \Delta W_n^1 \\ Y_{n+1}^2 &= Y_n^2 + Y_n^1 \Delta W_n^2 + I_{(1,2),t_n,t_{n+1}} \end{aligned}$$

Dieses Verfahren besitzt die starke Ordnung  $\gamma = 1$ , aber wir müssen die Lösung auswerten, um das Verfahren auszuführen!

Betrachte jetzt das modifizierte Verfahren

$$\begin{aligned} Z_{n+1}^1 &= Z_n^1 + \Delta W_n^1 \\ Z_{n+1}^2 &= Z_n^2 + Z_n^1 \Delta W_n^2 + \frac{1}{2} \Delta W_n^1 \Delta W_n^2 \end{aligned}$$

mit konstanter Schrittweite  $\Delta_n \equiv \Delta = T/N$  in dem Intervall  $[0, T]$ .

SATZ (Cameron + Clark, 1980) :  $\mathbb{E} (|X_T^2 - Z_N^2|^2) = \frac{1}{4} T \Delta$

$\Rightarrow$  nur die starke Ordnung  $\gamma = \frac{1}{2}$  statt  $\gamma = 1$ !

Beweis

$$Z_0^1 = 0 \quad \Rightarrow \quad Z_n^1 = \sum_{j=0}^{n-1} \Delta W_j^1 = W_{t_n}^1, \quad n \geq 1$$

$$\Rightarrow Z_{n+1}^2 = Z_n^2 + W_{t_n}^1 \Delta W_n^2 + \frac{1}{2} \Delta W_n^1 \Delta W_n^2 = Z_n^2 + \frac{1}{2} (W_{t_n}^1 + W_{t_{n+1}}^1) \Delta W_n^2$$

Mit  $Z_0^2 = 0$ , erhalten wir

$$Z_N^2 = \sum_{n=0}^{N-1} \frac{1}{2} (W_{t_n}^1 + W_{t_{n+1}}^1) \Delta W_n^2 = \sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} \frac{1}{2} (W_{t_n}^1 + W_{t_{n+1}}^1) dW_t^2$$

$W_t^1$  und  $W_t^2$  sind unabhängig  $\Rightarrow$  die Konstante  $\frac{1}{2} (W_{t_n}^1 + W_{t_{n+1}}^1)$  ist nicht antizipativ bzgl.  $W_t^2$ .  $\Rightarrow$  Integral OK!

Aber

$$X_T^2 = \int_0^T W_t^1 dW_t^2 = \sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} W_t^1 dW_t^2$$

$$\Rightarrow \mathbb{E} \left( |X_T^2 - Z_N^2|^2 \right) = \mathbb{E} \left( \left| \sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} \left\{ W_t^1 - \frac{1}{2} (W_{t_n}^1 + W_{t_{n+1}}^1) \right\} dW_t^2 \right|^2 \right)$$

Integrand hier nicht antizipativ bzgl.  $W_t^2$

$$= \sum_{n=0}^{N-1} \mathbb{E} \left( \left| \int_{t_n}^{t_{n+1}} \left\{ W_t^1 - \frac{1}{2} (W_{t_n}^1 + W_{t_{n+1}}^1) \right\} dW_t^2 \right|^2 \right)$$

weil die ZVen  $\int_{t_n}^{t_{n+1}} (*) dW_t^2$  hier unabhängig sind und Erwartungswert  $\mathbb{E}(\int_{t_n}^{t_{n+1}} (*) dW_t^2) = 0$  haben. Von einer Eigenschaft des Ito-Integrals erhalten wir dann

$$\mathbb{E} \left( |X_T^2 - Z_N^2|^2 \right) = \sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} \mathbb{E} \left( \left| W_t^1 - \frac{1}{2} (W_{t_n}^1 + W_{t_{n+1}}^1) \right|^2 \right) dt$$

Aber

$$\begin{aligned} \mathbb{E} \left( \left| W_t^1 - \frac{1}{2} (W_{t_n}^1 + W_{t_{n+1}}^1) \right|^2 \right) &= \mathbb{E} \left( \underbrace{\left| \frac{1}{2} (W_t^1 - W_{t_n}^1) + \frac{1}{2} (W_t^1 - W_{t_{n+1}}^1) \right|^2}_{\text{unabhängige ZVen mit } \mathbb{E}(\cdot)=0} \right) \\ &= \frac{1}{4} \mathbb{E} \left( |W_t^1 - W_{t_n}^1|^2 \right) + \frac{1}{4} \mathbb{E} \left( |W_t^1 - W_{t_{n+1}}^1|^2 \right) \\ &= \frac{1}{4} (t - t_n) + \frac{1}{4} (t_{n+1} - t) \\ &= \frac{1}{4} (t_{n+1} - t_n) = \frac{1}{4} \Delta \end{aligned}$$

Deshalb gilt

$$\mathbb{E} \left( |X_T^2 - Z_N^2|^2 \right) = \sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} \frac{1}{4} \Delta dt = N \cdot \frac{1}{4} \Delta^2 = \frac{1}{4} T \Delta$$

weil  $N \Delta = T$  ist.

### Was ist zu tun?

Oft können wir die Struktur einer SDGL benutzen, um das Verfahren zu vereinfachen und die Auswertung der Multi-Integrale zu vermeiden, z.B. eine SDGL

$$dX_t = f(t, X_t) dt + \sum_{j=1}^m g_j(t, X_t) dW_t^j$$

besitzt kommutatives Rauschen (erster Art), falls

$$\boxed{L^i g_j(t, x) \equiv L^j g_i(t, x)}, \quad \forall t, x, \quad \forall i, j \in \{1, \dots, m\}$$

In diesem Fall benutzen wir die Identität (Beweis später)

$$\boxed{I_{(i,j),t_n,t_{n+1}} + I_{(j,i),t_n,t_{n+1}} = I_{(i),t_n,t_{n+1}} I_{(j),t_n,t_{n+1}}}$$

d.h.

$$\begin{aligned} \int_{t_n}^{t_{n+1}} \int_{t_n}^t dW_s^i dW_t^j + \int_{t_n}^{t_{n+1}} \int_{t_n}^t dW_s^j dW_t^i &= \int_{t_n}^{t_{n+1}} dW_t^i \int_{t_n}^{t_{n+1}} dW_t^j \\ &= \Delta W_n^i \Delta W_n^j \end{aligned}$$

um die Summe

$$L^i g_j(t_n, Y_n) I_{(i,j),t_n,t_{n+1}} + L^j g_i(t_n, Y_n) I_{(j,i),t_n,t_{n+1}}$$

in dem Milstein-Verfahren zu vereinfachen. Wir erhalten nun

$$\begin{aligned} \text{Summe} &= L^i g_j(t_n, Y_n) \{ I_{(i,j),t_n,t_{n+1}} + I_{(j,i),t_n,t_{n+1}} \} \\ &= L^i g_j(t_n, Y_n) I_{(i),t_n,t_{n+1}} I_{(j),t_n,t_{n+1}} \\ &= L^i g_j(t_n, Y_n) \Delta W_n^i \Delta W_n^j \end{aligned}$$

Beweis der Identität  $W_t^1, W_t^2$  unabhängig  $\Rightarrow$

$$d(W_t^1 W_t^2) = W_t^1 dW_t^2 + W_t^2 dW_t^1$$

Integrieren und umformen !

Beispiel Lineare SDGL mit  $(d = 1, m = 2)$

$$dX_t = aX_t dt + bX_t dW_t^1 + cX_t dW_t^2$$

$$g_1 = bx \Rightarrow L^2 g_1 = cx \frac{d}{dx}(bx) = bcx$$

$$g_2 = cx \Rightarrow L^1 g_2 = bx \frac{d}{dx}(cx) = bcx$$

$\Rightarrow$  kommutatives Rauschen.

In diesem Fall lautet das Milstein-Verfahren

$$\begin{aligned} Y_{n+1} &= Y_n + aY_n \Delta_n + bY_n \Delta W_n^1 + cY_n \Delta W_n^2 \\ &\quad + \frac{1}{2} b^2 Y_n \{ (\Delta W_n^1)^2 - \Delta_n \} + \frac{1}{2} c^2 Y_n \{ (\Delta W_n^2)^2 - \Delta_n \} \\ &\quad + bcY_n \Delta W_n^1 \Delta W_n^2. \end{aligned}$$

ABER nicht jede SDGL besitzt kommutatives Rauschen, z.B.

$$d \begin{pmatrix} X_t^1 \\ X_t^2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} dt + \begin{pmatrix} 1 \\ 0 \end{pmatrix} dW_t^1 + \begin{pmatrix} 0 \\ X_t^1 \end{pmatrix} dW_t^2$$

Dann müssen wir irgendwie die Multi-Integrale  $I_{(i,j),t_n,t_{n+1}}$  ( $i \neq j$ ) simulieren oder approximieren.

Beispiel Für das Integral  $I_{(1,2),0,\Delta}$  benutze das Euler-Verfahren für die SDGL

$$dX_t^1 = dW_t^1, \quad dX_t^2 = X_t^1 dW_t^2$$

mit Anfangswert  $X_0^1 = X_0^2 = 0$  auf den Intervallen  $[k\delta, (k+1)\delta]$  mit Schrittweite  $\delta = \Delta/K$ , d.h.

$$Y_{k+1}^1 = Y_k^1 + \delta W_k^1, \quad Y_{k+1}^2 = Y_k^2 + Y_k^1 \delta W_k^2$$

mit  $Y_0^1 = Y_0^2 = 0$  und  $\delta W_k^i \sim N(0, \delta)$

Dann soll  $Y_K^2$  eine Approximation für  $X_\Delta^2 = I_{(1,2),0,\Delta}$  sein. Für den globalen Diskretisierungsfehler für das Euler-Verfahren erhalten wir

$$\mathbb{E}(|X_\Delta^2 - Y_K^2|) \leq K_T^{(E)} \delta^{1/2}$$

mit globaler Konstante  $K_T^{(E)}$  auf dem Intervall  $[0, T]$ .

Wir können  $Y_K^2$  in dem Milstein-Verfahren für eine beliebige SDGL statt  $I_{(1,2),0,\Delta}$  benutzen und die starke Ordnung  $\gamma = 1$  des Milstein-Verfahrens erhalten, wenn wir  $\delta$  wählen mit

$$K_T^{(E)} \delta^{1/2} \leq K_T^{(M)} \Delta \quad \Rightarrow \quad \delta \sim \Delta^2$$

Hier ist  $K_T^{(M)}$  die Konstante des globalen Diskretisierungsfehlers des Milstein-Verfahrens und  $\delta \ll \Delta$ , weil  $\Delta \ll 1$

Aber diese Methode ist aufwendig im Rechner.

Alternative mit zufälligen Fourier-Reihen

$$W_t^j - \frac{t}{\Delta} W_\Delta^j = \frac{1}{2} a_0 + \sum_{k=1}^{\infty} \left\{ a_k \cos\left(\frac{2\pi k}{\Delta} t\right) + b_k \sin\left(\frac{2\pi k}{\Delta} t\right) \right\}$$

wobei die  $a_0, a_k, b_k \sim N(0, S_k)$  sind (für geeignete  $S_k$ )

Differenziere komponentenweise  $\Rightarrow$

$$dW_t^j = \frac{1}{\Delta} W_\Delta^j + \frac{d}{dt} \sum_{k=1}^{\infty} \{***\}$$

Multipliziere die Reihen  $\Rightarrow$

$$\int_0^\Delta W_t^i dW_t^j = \int_0^\Delta (\text{Produkt der Reihen}) dt = \sum_{k=0}^{\infty} \{***\}$$

abschneide  $\sum_{k=0}^{\infty} \{***\}$  für eine Approximation.

Siehe Kloeden/Platen: Seite 198 ff.

## Kapitel 15

# Stochastische Multi-Integrale mit bekannten Verteilungen

Nicht alle stochastischen Multi-Integrale sind schwierig zu simulieren, z.B. Multi-Integrale wie  $I_{(0,1),t_n,t_{n+1}}$  und  $I_{(1,0),t_n,t_{n+1}}$ , die in dem  $\gamma = \frac{3}{2}$ -Taylor-Verfahren entstehen.

Betrachte die hierarchische Menge  $\Lambda_{\frac{3}{2}}$  mit  $m = 1$ , d.h.

$$\begin{aligned}\Lambda_{\frac{3}{2}} &= \{\alpha \in \mathcal{M}_1 : l(\alpha) + n(\alpha) \leq 3 \text{ oder } l(\alpha) = n(\alpha) = 2\} \\ &= \underbrace{\{\emptyset, (0), (1), (1, 1)\}}_{\text{Milstein } \Lambda_1}, \underbrace{\{(0, 0), (0, 1), (1, 0), (1, 1, 1)\}}_{\text{neu}}\end{aligned}$$

Wir haben einfache Formeln für  $I_{(0,0)}$  und  $I_{(1,1,1)}$ .

$$\begin{aligned}I_{(0,0),t_n,t_{n+1}} &= \int_{t_n}^{t_{n+1}} \int_{t_n}^t dW_s^0 dW_t^0 \\ &= \int_{t_n}^{t_{n+1}} \int_{t_n}^t ds dt \quad \text{weil } dW_t^0 = dt \\ &= \frac{1}{2} (t_{n+1} - t_n)^2 = \frac{1}{2} \Delta_n^2\end{aligned}$$

und

$$\begin{aligned}I_{(1,1,1),t_n,t_{n+1}} &= \int_{t_n}^{t_{n+1}} \int_{t_n}^t \int_{t_n}^s dW_\tau^1 dW_s^1 dW_t^1 \\ &= \frac{1}{6} \left\{ (\Delta W_n^1)^3 - 3\Delta_n \Delta W_n^1 \right\}\end{aligned}$$

Es gibt keine einfachen Formeln für  $I_{(0,1)}$  oder  $I_{(1,0)}$ . Aber die folgenden Eigenschaften sind bekannt.

(1) Identität

$$\boxed{I_{(0,1),t_n,t_{n+1}} + I_{(1,0),t_n,t_{n+1}} = I_{(0),t_n,t_{n+1}} I_{(1),t_n,t_{n+1}}}$$

d.h.

$$\int_{t_n}^{t_{n+1}} \int_{t_n}^t ds dW_t^1 + \int_{t_n}^{t_{n+1}} \int_{t_n}^t dW_s^1 dt = \int_{t_n}^{t_{n+1}} dt \int_{t_n}^{t_{n+1}} dW_t^1 = \Delta_n \Delta W_n^1$$

(2) Verteilungen

$$I_{(1,0),t_n,t_{n+1}} \sim N\left(0, \frac{1}{3}\Delta_n^3\right), \quad I_{(0,1),t_n,t_{n+1}} \sim N\left(0, \frac{1}{3}\Delta_n^3\right)$$

mit den Korrelationen

$$\mathbb{E}\left(I_{(1,0),t_n,t_{n+1}} I_{(1),t_n,t_{n+1}}\right) = \frac{1}{2}\Delta_n^2$$

$$\mathbb{E}\left(I_{(0,1),t_n,t_{n+1}} I_{(1),t_n,t_{n+1}}\right) = \frac{1}{2}\Delta_n^2$$

$$\mathbb{E}\left(I_{(0,1),t_n,t_{n+1}} I_{(1,0),t_n,t_{n+1}}\right) = \frac{1}{6}\Delta_n^3.$$

Beweise Die Identität folgt von der Ito-Formel:

$$d(t W_t^1) = t dW_t^1 + W_t^1 dt$$

dann integriere auf  $[t_n, t_{n+1}]$  und forme um !

Die Verteilungen folgen von der Identität und den Eigenschaften des Ito-Integrals.

Insbesondere sind die beiden Integrale  $I_{(1,0)}$  und  $I_{(0,1)}$  Gauß-verteilt, weil sie Grenzwerte linearer Kombinationen Gauß-verteilter Zufallsvariablen sind.

Erwartungswerte

$$\mathbb{E}\left(\int_{t_n}^{t_{n+1}} \int_{t_n}^t ds dW_t^1\right) = \mathbb{E}\left(\int_{t_n}^{t_{n+1}} (t - t_n) dW_t^1\right) = 0 \quad \text{Ito-Integral}$$

$$\begin{aligned} \mathbb{E}\left(\int_{t_n}^{t_{n+1}} \int_{t_n}^t dW_s^1 dt\right) &= \mathbb{E}\left(\int_{t_n}^{t_{n+1}} (W_t^1 - W_{t_n}^1) dt\right) \\ &= \int_{t_n}^{t_{n+1}} \mathbb{E}(W_t^1 - W_{t_n}^1) dt = \int_{t_n}^{t_{n+1}} 0 \cdot dt = 0 \end{aligned}$$

Korrelationen (1. Teil)

$$\begin{aligned} \mathbb{E} (I_{(0,1),t_n,t_{n+1}} I_{(1),t_n,t_{n+1}}) &= \mathbb{E} \left( \int_{t_n}^{t_{n+1}} (t - t_n) dW_t^1 \int_{t_n}^{t_{n+1}} 1 dW_t^1 \right) \\ &= \int_{t_n}^{t_{n+1}} \mathbb{E} ((t - t_n) 1) dt \\ &= \frac{1}{2} (t_{n+1} - t_n)^2 = \frac{1}{2} \Delta_n^2 \end{aligned}$$

wegen der folgenden Eigenschaft des Ito-Integrals

$$\mathbb{E} \left( \int f dW_t \int g dW_t \right) = \int \mathbb{E} (fg) dt$$

Beweis: benutze die Identitäten

$$\mathbb{E} \left( \int f dW_t \right)^2 = \int \mathbb{E} (f^2) dt, \quad 4AB = (A + B)^2 - (A - B)^2.$$

Varianzen (hier = 2-te Momente)

$$\begin{aligned} \mathbb{E} \left( |I_{(0,1),t_n,t_{n+1}}|^2 \right) &= \mathbb{E} \left( \left| \int_{t_n}^{t_{n+1}} (t - t_n) dW_t^1 \right|^2 \right) \\ &= \int_{t_n}^{t_{n+1}} \mathbb{E} (|t - t_n|^2) dt \quad \text{Ito-Integral} \\ &= \int_{t_n}^{t_{n+1}} (t - t_n)^2 dt \\ &= \frac{1}{3} (t_{n+1} - t_n)^3 = \frac{1}{3} \Delta_n^3 \end{aligned}$$

und dann

$$\begin{aligned} \mathbb{E} \left( |I_{(1,0),t_n,t_{n+1}}|^2 \right) &= \mathbb{E} \left( |I_{(0)} I_{(1)} - I_{(0,1)}|^2 \right) \\ &= \mathbb{E} \left( |I_{(0)} I_{(1)}|^2 \right) - 2\mathbb{E} (I_{(0)} I_{(1)} I_{(0,1)}) + \mathbb{E} \left( |I_{(0,1)}|^2 \right) \\ &= \mathbb{E} \left( \Delta_n^2 |\Delta W_n^1|^2 \right) - 2\mathbb{E} (\Delta_n I_{(1)} I_{(0,1)}) + \frac{1}{3} \Delta_n^3 \\ &= \Delta_n^2 \mathbb{E} \left( |\Delta W_n^1|^2 \right) - 2\Delta_n \underbrace{\mathbb{E} (I_{(1)} I_{(0,1)})}_{=\frac{1}{2} \Delta_n^2 \text{ Korrelation!}} + \frac{1}{3} \Delta_n^3 \\ &= \Delta_n^3 - \Delta_n^3 + \frac{1}{3} \Delta_n^3 = \frac{1}{3} \Delta_n^3 ! \end{aligned}$$

Korrelation (2. Teil) Wir benutzen die Identität

$$I_{(0,1)} = I_{(0)}I_{(1)} - I_{(1,0)}$$

$$\begin{aligned} \Rightarrow \mathbb{E} \left( |I_{(0,1)}|^2 \right) &= \mathbb{E} \left( |I_{(0)}I_{(1)} - I_{(1,0)}|^2 \right) \\ &= \mathbb{E} \left( |I_{(0)}I_{(1)}|^2 \right) - 2\mathbb{E} \left( I_{(0)}I_{(1)}I_{(1,0)} \right) + \mathbb{E} \left( |I_{(1,0)}|^2 \right) \\ \Rightarrow \frac{1}{3}\Delta_n^3 &= \Delta_n^3 - 2\Delta_n \mathbb{E} \left( I_{(1)}I_{(1,0)} \right) + \frac{1}{3}\Delta_n^3 \end{aligned}$$

deshalb gilt  $\mathbb{E} \left( I_{(1)}I_{(1,0)} \right) = \frac{1}{2}\Delta_n^2$ .

$$\begin{aligned} \mathbb{E} \left( I_{(1,0)}I_{(0,1)} \right) &= \mathbb{E} \left( I_{(1,0)}(I_0I_1 - I_{(1,0)}) \right) \\ &= \mathbb{E} \left( I_{(1,0)}I_0I_1 \right) - \mathbb{E} \left( (I_{(1,0)})^2 \right) \\ &= \frac{1}{2}\Delta_n^3 - \frac{1}{3}\Delta_n^3 = \frac{1}{6}\Delta_n^3 \end{aligned}$$

Schreibe  $\Delta Z_n = I_{(1,0),t_n,t_{n+1}}$  und  $\Delta W_n = I_{(1),t_n,t_{n+1}}$

$$\Rightarrow \Delta Z_n \sim N \left( 0, \Delta_n^3/3 \right), \quad \Delta W_n \sim N \left( 0, \Delta_n \right)$$

mit Korrelation

$$\mathbb{E} \left( \Delta Z_n \Delta W_n \right) = \frac{1}{2}\Delta_n^2$$

SATZ Sei  $G_1, G_2 \sim N(0, 1)$  und unabhängig. Dann gilt

$$\begin{cases} \Delta W_n &= \Delta_n^{1/2}G_1 \\ \Delta Z_n &= \frac{1}{2}\Delta_n^{3/2}G_1 + \frac{1}{\sqrt{12}}\Delta_n^{3/2}G_2 \end{cases}$$

Beweis: direkte Bestätigung !

Wir können  $G_1$  und  $G_2$  durch die Box-Müller-Methode erzeugen. Dann haben wir  $\Delta W_n$  und  $\Delta Z_n$  von den obigen Formeln, und von der Identität

$$I_{(0,1)} = \Delta_n \cdot \Delta W_n - \Delta Z_n.$$

Sonderstruktur der SDGL

Wenn die SDGL  $dX_t = f(t, X_t)dt + g(t, X_t)dW_t^1$  der Bedingung

$$\boxed{L^0 g \equiv L^1 f}$$

genügt, dann können wir die Identität verwenden und das  $\gamma = \frac{3}{2}$ -Taylor-Verfahren vereinfachen. Insbesondere

$$\begin{aligned} & L^0 g(t_n, Y_n) \cdot I_{(0,1),t_n,t_{n+1}} + L^1 f(t_n, Y_n) I_{(1,0),t_n,t_{n+1}} \\ &= L^0 g(t_n, Y_n) \{ I_{(0,1),t_n,t_{n+1}} + I_{(1,0),t_n,t_{n+1}} \} \\ &= L^0 g(t_n, Y_n) I_{(0),t_n,t_{n+1}} I_{(1),t_n,t_{n+1}} \\ &= L^0 g(t_n, Y_n) \Delta_n \Delta W_n^1 \end{aligned}$$

d.h. wir brauchen  $\Delta Z_n$  nicht zu erzeugen!.

Die obige Bedingung ist eine Art kommutatives Rauschen !

# Kapitel 16

## Pfadweiskonvergenz

Pathwise convergence is interesting for SODEs because numerical calculations of the random variables  $Y_n$  in the numerical scheme above are carried out path by path. Itô calculus is, however, an  $L^2$  or a mean-square calculus and not a pathwise calculus. Nevertheless some results for pathwise approximation of SODE are known. For example, in 1983 Talay showed that the Milstein scheme SODE with a scalar Brownian motion has the pathwise error estimate

$$\sup_{n=0, \dots, N_T} \left| X_{t_n}(\omega) - Y_n^{(Mil)}(\omega) \right| \leq K_{\epsilon, T}^{(Mil)}(\omega) \Delta^{\frac{1}{2} - \epsilon},$$

for all  $\epsilon > 0$  and almost all  $\omega \in \Omega$ , i.e., is pathwise of order  $\frac{1}{2} - \epsilon$ . Later Gyöngy and Fleury showed that the Euler-Maruyama scheme has the same pathwise convergence order. Note that the error constants here depend on  $\omega$ , so they are in fact random variables. The nature of their statistical properties is an interesting question, on which less is known theoretically so far and which requires further investigations.

Given that the sample paths of a Wiener process are Hölder continuous with exponent  $\frac{1}{2} - \epsilon$  one may ask: *Is the convergence order  $\frac{1}{2} - \epsilon$  “sharp for pathwise approximation?”* The answer is NO! Kloeden & Neuenkirch [13] showed that an arbitrary pathwise convergence order is possible.

**Theorem 1** *Under standard assumptions the Itô-Taylor scheme of strong order  $\gamma > 0$  converges pathwise with order  $\gamma - \epsilon$  for all  $\epsilon > 0$ , i.e.,*

$$\sup_{i=0, \dots, N_T} \left| X_{t_n}(\omega) - Y_n^{(\gamma)}(\omega) \right| \leq K_{\epsilon, T}^{(\gamma)}(\omega) \cdot \Delta^{\gamma - \epsilon}$$

for almost all  $\omega \in \Omega$ .

Thus, for example, the Milstein scheme has pathwise order  $1 - \epsilon$  rather than the lower order  $\frac{1}{2} - \epsilon$  obtained in Talay (which was a consequence of the proof used there).

The proof of Theorem 1, which will not be given here, is based on the Burkholder-Davis-Gundy inequality

$$\mathbb{E} \sup_{s \in [0, t]} \left| \int_0^s X_\tau dW_\tau \right|^p \leq C_p \cdot \mathbb{E} \left| \int_0^t X_\tau^2 d\tau \right|^{p/2}$$

and a Borel-Cantelli argument in the following lemma.

**Lemma 2** *Let  $\gamma > 0$  and  $c_p \geq 0$  for  $p \geq 1$ . If  $\{Z_n\}_{n \in \mathbb{N}}$  is a sequence of random variables with*

$$(\mathbb{E}|Z_n|^p)^{1/p} \leq c_p \cdot n^{-\gamma}$$

*for all  $p \geq 1$  and  $n \in \mathbb{N}$ , then for each  $\epsilon > 0$  there exists a finite non-negative random  $K_\epsilon$  such that*

$$|Z_n(\omega)| \leq K_\epsilon(\omega) \cdot n^{-\gamma+\epsilon} \quad a.s.$$

*for all  $n \in \mathbb{N}$ .*

## 16.1 Restrictiveness of the standard assumptions

Proofs in the literature of the above convergence orders, e.g., in the monographs Kloeden & Platen [14] and Milstein [16], assume that the coefficient functions  $f_\alpha$  in the Itô-Taylor schemes are *uniformly bounded* on  $\mathbb{R}^d$ , i.e., the partial derivatives of appropriately high order of the SODE coefficient functions  $a, b_1, \dots, b_m$  are *uniformly bounded* on  $\mathbb{R}^d$ .

This assumption is not satisfied for many SODEs in important applications such as the stochastic Ginzburg-Landau equation (the constants here are all positive)

$$dX_t = \left( \left( \nu + \frac{1}{2}\sigma^2 \right) X_t - \lambda X_t^3 \right) dt + \sigma X_t dW_t, \quad (16.1)$$

or even the simpler SODE (16.5) with additive noise with a cubic nonlinearity. Matters are even worse for the Fisher-Wright equation

$$dX_t = [\kappa_1(1 - X_t) - \kappa_2 X_t] dt + \sqrt{X_t(1 - X_t)} dW_t, \quad (16.2)$$

the Feller diffusion with logistic growth SODE

$$dX_t = \lambda X_t (K - X_t) dt + \sigma \sqrt{X_t} dW_t \quad (16.3)$$

and the Cox-Ingersoll-Ross equation

$$dV_t = \kappa(\lambda - V_t) dt + \theta \sqrt{V_t} dW_t, \quad (16.4)$$

since square root function is not differentiable at zero and requires the expression under it to remain non-negative for the SODE to make sense.

### 16.1.1 Counterexamples for the Euler-Maruyama scheme

The scalar SODE

$$dX_t = -X_t^3 dt + dW_t, \quad (16.5)$$

with the cubic drift and additive noise has a globally pathwise asymptotically stable stochastic stationary solution, since the drift satisfies a one sided dissipative Lipschitz condition (outside any arbitrarily small neighbourhood of the origin). Its solution on the time interval  $[0, 1]$  for initial value  $X_0 = 0$  satisfies the stochastic integral equation

$$X_t = - \int_0^t X_s^3 ds + W_t \quad (16.6)$$

for every  $t \in [0, 1]$  and has finite first moment  $\mathbb{E}|X_1| < \infty$ .

The corresponding Euler-Maruyama scheme with constant step size  $\Delta = \frac{1}{N}$  is given by

$$Y_{k+1}^{(N)} = Y_k^{(N)} - \left(Y_k^{(N)}\right)^3 \Delta + \Delta W_k(\omega), \quad (16.7)$$

with

$$\Delta W_k = W_{(k+1)\Delta} - W_{k\Delta}.$$

The next theorem, due to Hutzenthaler, Jentzen & Kloeden[8] shows that the Euler-Maruyama scheme does not converge strongly for the SODE (16.6). In addition, the remark following the proof shows that it also does not converge weakly.

**Theorem 2** *The solution  $X_t$  of (16.6) and its Euler-Maruyama approximation  $Y_k^K$  satisfy*

$$\lim_{N \rightarrow \infty} \mathbb{E} \left| X_1 - Y_N^{(N)} \right| = \infty. \quad (16.8)$$

**Proof:** Let  $N \in \mathbb{N}$  be arbitrary, define  $r_N := \max\{3N, 2\}$  and consider the event

$$\Omega_N := \left\{ \omega \in \Omega : \sup_{k=1, \dots, N-1} |\Delta W_k(\omega)| \leq 1, \quad |\Delta W_0(\omega)| \geq r_N \right\}.$$

Then, it follows by induction that

$$\left| Y_k^{(N)}(\omega) \right| \geq r_N^{2^{k-1}}, \quad \forall \omega \in \Omega_N \quad (16.9)$$

for every  $k = 1, 2, \dots, N$ . Clearly, for  $K = 1$ ,

$$\left| Y_1^{(N)}(\omega) \right| = |\Delta W_0(\omega)| \geq r_N \quad (16.10)$$

for each  $\omega \in \Omega_N$  from the definition of the Euler-Maruyama scheme (16.7) and the definition of the set  $\Omega_N$ . Therefore assume that equation (16.9) holds for some  $k \in \{1, 2, \dots, N-1\}$ . Then

$$\left| Y_k^{(N)}(\omega) \right| \geq r_N^{2^{k-1}} \geq r_N \geq 1 \quad (16.11)$$

for every  $\omega \in \Omega_N$ . Now, by the definition of the set  $\Omega_N$ ,

$$|\Delta W_k(\omega)| \leq 1 \quad (16.12)$$

for every  $\omega \in \Omega_N$ , so by (16.7) it follows that

$$\begin{aligned} |Y_{k+1}^{(N)}(\omega)| &= \left| Y_k^{(N)}(\omega) - \left( Y_k^{(N)}(\omega) \right)^3 \Delta + \Delta W_k(\omega) \right| \\ &\geq \left| \left( Y_k^{(N)}(\omega) \right)^3 \Delta - \left| Y_k^{(N)}(\omega) \right| - |\Delta W_k(\omega)| \right| \\ &\geq \left| Y_k^{(N)}(\omega) \right|^3 \Delta - \left| Y_k^{(N)}(\omega) \right| - 1 \end{aligned}$$

for every  $\omega \in \Omega_N$ . Hence equation (16.11) implies that

$$\begin{aligned} |Y_{k+1}^{(N)}(\omega)| &\geq \left| Y_k^{(N)}(\omega) \right|^3 \Delta - 2 \left| Y_k^{(N)}(\omega) \right|^2 \\ &\geq \left| Y_k^{(N)}(\omega) \right|^2 \left( \left| Y_k^{(N)}(\omega) \right| \Delta - 2 \right) \\ &\geq \left| Y_k^{(N)}(\omega) \right|^2 (r_N \Delta - 2) \geq \left| Y_k^{(N)}(\omega) \right|^2 \end{aligned}$$

for every  $\omega \in \Omega_N$ , since  $r_N \Delta - 2 \geq 1$  by definition. Thus, the induction hypothesis, i.e., that (16.9) holds for  $k$ , implies that

$$\left| Y_{k+1}^{(N)}(\omega) \right| \geq \left( r_N^{2^{k-1}} \right)^2 = r_N^{2^k}$$

for all  $\omega \in \Omega_N$ , which shows that the inequality (16.9) holds for  $k+1$ .

Hence, inequality (16.9) holds for every  $k = 1, 2, \dots, N$ . In particular,

$$\left| Y_N^{(N)}(\omega) \right| \geq r_N^{2^{N-1}}$$

for every  $\omega \in \Omega_N$  and  $N \in \mathbb{N}$  and it follows that

$$\mathbb{E} \left| Y_N^{(N)} \right| = \int_{\Omega} \left| Y_N^{(N)}(\omega) \right| \mathbb{P}(d\omega) \geq \int_{\Omega_N} \left| Y_N^{(N)}(\omega) \right| \mathbb{P}(d\omega) \geq \mathbb{P}[\Omega_N] \cdot 2^{(2^{N-1})}$$

for every  $N \in \mathbb{N}$ . Now

$$\begin{aligned} \mathbb{P}[\Omega_N] &= \mathbb{P} \left[ \sup_{k=1, \dots, N-1} |\Delta W_k| \leq 1 \right] \cdot \mathbb{P}[|W_1| \geq r_N] \\ &\geq \mathbb{P} \left[ \sup_{0 \leq t \leq 1} |W_t| \leq \frac{1}{2} \right] \cdot \mathbb{P}[|W_1| \geq r_N] \\ &= \mathbb{P} \left[ \sup_{0 \leq t \leq 1} |W_t| \leq \frac{1}{2} \right] \cdot \mathbb{P} \left[ \sqrt{N} |W_1| \geq \sqrt{N} r_N \right] \\ &\geq \mathbb{P} \left[ \sup_{0 \leq t \leq 1} |W_t| \leq \frac{1}{2} \right] \cdot \frac{1}{4} \sqrt{N} r_N e^{-(\sqrt{N} r_N)^2} \\ &\geq \frac{1}{4} \cdot \mathbb{P} \left[ \sup_{0 \leq t \leq 1} |W_t| \leq \frac{1}{2} \right] \cdot e^{-N r_N^2} \end{aligned}$$

for every  $N \in \mathbb{N}$ . The next step needs the following lemma, which was proved in [8].

**Lemma 3** *Let  $(\Omega, \mathcal{F}, \mathbb{P})$  be a probability space and let  $Z: \Omega \rightarrow \mathbb{R}$  be an  $\mathcal{F}/\mathcal{B}(\mathbb{R})$ -measurable mapping which is standard normally distributed. Then*

$$\mathbb{P}\left[|Z| \geq x\right] \geq \frac{x e^{-x^2}}{4}, \quad \mathbb{P}\left[|Z| \in [x, 2x]\right] \geq \frac{x e^{-2x^2}}{2} \quad (16.13)$$

for every  $x \in [0, \infty)$ .

It follows from Lemma 3 that

$$\mathbb{E}\left|Y_N^{(N)}\right| \geq \frac{1}{4} \cdot \mathbb{P}\left[\sup_{0 \leq t \leq 1} |W_t| \leq \frac{1}{2}\right] \cdot e^{-Nr_N^2} \cdot 2^{2^{N-1}}$$

for every  $N \in \mathbb{N}$ . Hence

$$\begin{aligned} \lim_{N \rightarrow \infty} \mathbb{E}\left|Y_N^{(N)}\right| &\geq \frac{1}{4} \cdot \mathbb{P}\left[\sup_{0 \leq t \leq 1} |W_t| \leq \frac{1}{2}\right] \cdot \lim_{N \rightarrow \infty} e^{-Nr_N^2} \cdot 2^{2^{N-1}} \\ &= \frac{1}{4} \cdot \mathbb{P}\left[\sup_{0 \leq t \leq 1} |W_t| \leq \frac{1}{2}\right] \cdot \lim_{N \rightarrow \infty} e^{-9N^3} \cdot 2^{2^{N-1}} = \infty. \end{aligned}$$

Finally, since  $\mathbb{E}|X_1|$  is finite, it follows that

$$\lim_{N \rightarrow \infty} \mathbb{E}\left|X_1 - Y_N^{(N)}\right| \geq \lim_{N \rightarrow \infty} \mathbb{E}\left|Y_N^{(N)}\right| - \mathbb{E}|X_1| = \infty,$$

which is the assertion.  $\square$

**Remark 1** *Theorem 2 implies, by Jensen's inequality, that*

$$\lim_{N \rightarrow \infty} \mathbb{E}\left[\left|X_1 - Y_N^{(N)}\right|^p\right] = \infty \quad (16.14)$$

for every  $p \in [1, \infty)$ . Moreover, since  $\mathbb{E}|X_1|^p < \infty$  for every  $p \in [1, \infty)$ , it follows that

$$\begin{aligned} \left(\mathbb{E}\left|Y_N^{(N)}\right|^p\right)^{\frac{1}{p}} &= \left(\mathbb{E}\left|Y_N^{(N)} - X_1 + X_1\right|^p\right)^{\frac{1}{p}} \\ &\geq \left(\mathbb{E}\left|Y_N^{(N)} - X_1\right|^p\right)^{\frac{1}{p}} - \left(\mathbb{E}|X_1|^p\right)^{\frac{1}{p}} \rightarrow \infty \end{aligned}$$

as  $N \rightarrow \infty$  by (16.14). Hence

$$\lim_{N \rightarrow \infty} \left|\mathbb{E}|X_1|^p - \mathbb{E}\left|Y_N^{(N)}\right|^p\right| = \infty \quad (16.15)$$

for every  $p \in [1, \infty)$ .

The above result is a special case of a more general result for scalar SODEs

$$dX_t = a(X_t) dt + b(X_t) dW_t, \quad t \in [0, 1], \quad (16.16)$$

for which at least one coefficient functions grows super-linearly, assuming that a solution exists, i.e., a predictable stochastic process  $X_t$  satisfying

$$\mathbb{P}\left[\int_0^1 |a(X_s)| + |b(X_s)|^2 ds < \infty\right] = 1 \quad (16.17)$$

and

$$\mathbb{P} \left[ X_t = \xi + \int_0^t a(X_s) ds + \int_0^t g(X_s) dW_s \right] = 1 \quad (16.18)$$

for every  $t \in [0, 1]$ , where the initial value is  $X_0 = \xi$ .

**Theorem 3** *Suppose that there exist constants  $C \geq 1$  and  $\beta > \alpha > 1$  such that*

$$\max(|a(x)|, |b(x)|) \geq \frac{|x|^\beta}{C}, \quad \min(|a(x)|, |b(x)|) \leq C|x|^\alpha \quad (16.19)$$

for all  $|x| \geq C$ . If the solution  $X_t$  of the scalar SODE

$$dX_t = a(X_t) dt + b(X_t) dW_t, \quad X_0 = \xi, \quad (16.20)$$

for  $t \in [0, 1]$  satisfies  $\mathbb{E}|X_1|^p < \infty$  for some  $p \in [1, \infty)$  and if the initial condition satisfies  $\mathbb{P}[g(\xi) \neq 0] > 0$ , then

$$\lim_{N \rightarrow \infty} \mathbb{E} \left| X_1 - Y_N^{(N)} \right|^p = \infty \quad \text{and} \quad \lim_{N \rightarrow \infty} \left| \mathbb{E}|X_1|^p - \mathbb{E}|Y_N^{(N)}|^p \right| = \infty, \quad (16.21)$$

where  $Y_N^{(N)}$  denotes the  $N$ th iterate of the corresponding Euler-Maruyama scheme with constant time step  $\Delta = \frac{1}{N}$ .

The assumption that the diffusion function does not vanish on the starting point ensures the presence of noise in the first time step.

The proof of Theorem 3 can be found in Hutzenthaler, Jentzen & Kloeden [8]. The estimates in the proof require  $\beta > 1$ , so do not apply to a super-linear drift function such as  $a(x) = x \log x$ . However, the theorem does apply to the stochastic Ginzburg-Landau equation (16.1) and the Feller diffusion with logistic growth SODE (16.3), amongst others.

# Literaturverzeichnis

- [1] L. Arnold, *Stochastische Differentialgleichungen*, R. Oldenbourg-Verlag, 1973.
- [2] S. Cyganowski, P.E. Kloeden and J. Ombach, *From Elementary Probability to Stochastic Differential Equations with MAPLE*, Heidelberg: Springer 2002.
- [3] T.C. Gard, *Introduction to Stochastic Differential Equations*, Marcel-Dekker, 1989.
- [4] L. Grüne und P.E. Kloeden, Higher order numerical schemes for affinely controlled nonlinear systems. *Numer. Math.* **89** (2001), 669–690
- [5] D.J. Higham, An algorithmic introduction to numerical simulation of stochastic differential equations, *SIAM Review*, **43** (2001), 525–546.
- [6] D.J. Higham and P.E. Kloeden, MAPLE and MATLAB for Stochastic Differential Equations in Finance, in *Programming Languages and Systems in Computational Economics and Finance* Editor: S.S. Nielsen. Kluwer Academic Publishers, Amsterdam (2002), pp.233–269.
- [7] J. Honerkamp, *Stochastische Dynamische Systeme – Konzepte, numerische Methoden, Datenanalysen*, VCH, Weinheim, 1990.
- [8] M. HUTZENTHALER, A. JENTZEN AND P.E. KLOEDEN, Strong and weak divergence in finite time of Euler’s method for SDEs with non-globally Lipschitz coefficients, *Proc. Roy. Soc. London A* (to appear)
- [9] M. HUTZENTHALER AND A. JENTZEN, Convergence of the stochastic Euler scheme for locally Lipschitz coefficients, *Foundations of Computational Math.* (to appear)
- [10] A. JENTZEN, P.E. KLOEDEN AND A. NEUENKIRCH, *Convergence of numerical approximations of stochastic differential equations on domains: higher order convergence rates without global Lipschitz coefficients*, *Numerische Mathematik* 112 (2009), no. 1, 41–64.

- [11] P.E. Kloeden, *Skript zur Vorlesung: Numerical Methods for Differential Equations*, Universität Frankfurt SS2011. (Im Internet)
- [12] P.E. Kloeden, The systematic deviation of higher order numerical methods for stochastic differential equations, *Milan J. Math.*, **70** (2002), 187–207.
- [13] P.E. KLOEDEN AND A. NEUENKIRCH, *The pathwise convergence of approximation schemes for stochastic differential equations*, LMS J. Comp. Math. 10 (2007), 235-253.
- [14] P.E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations*, Heidelberg: Springer 1992 (und später).
- [15] P.E. Kloeden, E. Platen und H. Schurz, *Numerical Solution of SDEs through Computer Experiments*, Heidelberg: Springer 1994 (und später).
- [16] G. Milstein, *Numerical Integration of Stochastic Differential Equations*, Dordrecht: Kluwer 1994.
- [17] B. Øksendal, *Stochastic Differential Equations*, Springer-Verlag, 1985 (und neuere Aufgaben).