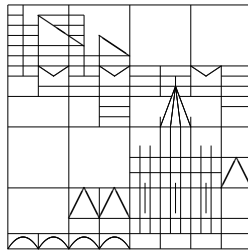


Skript zur Vorlesung

Numerik stochastischer
Differentialgleichungen

Wintersemester 2025/26

Johannes Schropp



Universität Konstanz

Fachbereich Mathematik und Statistik

© Johannes Schropp, 19. Januar 2026

Inhaltsverzeichnis

1	Grundlagen stochastischer Prozesse	1
1.1	Einführung	1
1.1.1	Beispiel Teilchenbewegung	1
1.1.2	Stochastische Modellierung des zeitlich gemittelten Prozesses	2
1.1.3	Stochastische Prozesse	4
1.1.4	Verallgemeinerung unseres Beispiels	5
1.2	Stochastisches Integral (Itô-Integral) und Itô-Prozesse	7
1.2.1	Beispiel	7
1.2.2	Technische Voraussetzungen dazu	7
1.2.3	Eigenschaften des Itô-Integrals	8
1.2.4	Prinzipielles weiteres Vorgehen	10
1.2.5	Rechenregeln für stochastische Integrale	10
1.2.6	Itô-Prozesse	11
1.2.7	Das Lemma von Itô	13
1.3	Modellierung von Prozessen mit Rauschen	14
1.3.1	Additives Rauschen	14
1.3.2	Multiplikatives Rauschen	15
1.3.3	Modellierung in der Finanzökonomie	17
1.3.4	Ausgangspunkt für die Numerik	17
1.3.5	Optionspreisberechnung	18
1.3.6	Monte-Carlo-Simulation	20
1.4	Pseudozufallszahlen	21
1.4.1	Einführung	21
1.4.2	Beispiel: Monte-Carlo-Integration	21
1.4.3	Konstruktion von Pseudozufallszahlen	23
1.4.4	Konstruktion von Pseudozufallsvektoren (PZV)	26
1.4.5	Die Box-Muller-Methode	27
1.4.6	Korrelierte Normalverteilung (korreliert normalverteilte PZV)	28
1.4.7	Die Cholesky-Zerlegung	28
1.4.8	Konstruktion von PZZ mit Gleichverteilung auf $[0, 1]$	29
2	Numerische Methoden für stochastische Differentialgleichungen	31
2.1	Das Euler-Maruyama-Verfahren	31
2.1.1	Numerische Methoden	31
2.1.2	Güte der Approximation	31
2.1.3	Untersuchung der Doppelintegrale	32
2.2	Konvergenztheorie für das Euler-Maruyama-Verfahren	34
2.2.1	Exakte Lösung vs. diskretes Verfahren	34
2.2.2	Analyse des L^2 -Fehlers	34
2.2.3	Separate Analyse der Terme (I)-(IV)	35
2.2.4	Das diskrete Gronwall-Lemma	38
2.2.5	Starke Konvergenz	39

2.2.6	Numerische Bestimmung der Konvergenzordnung	40
2.2.7	Fehler bei der Erwartungswertberechnung	41
2.3	Weitere numerische Verfahren und schwache Konvergenz	43
2.3.1	Das Milstein-Verfahren	43
2.3.2	Schwache Konvergenz	45
3	Parabolische partielle Differentialgleichungen und Optionspreisbewer-	
	tung	50
3.1	Die Fokker-Planck-Gleichung	50
3.1.1	Einführung	50
3.1.2	Die 1-dimensionale Fokker-Planck-Gleichung	51
3.1.3	Weiteres zu stochastischen Prozessen und PDGL	52
3.1.4	Verallgemeinerung auf vektorwertige Prozesse	53
3.1.5	Beispiel: Optionsbewertung eines Capped-Symmetric-Power-Calls .	53
3.2	Finite-Differenzen-Verfahren für die Wärmeleitungsgleichung	55
3.2.1	Einführung	55
3.2.2	Das Prinzip der Linienmethode	55
3.2.3	Das Euler-Cauchy-Verfahren	57
3.2.4	Das ϑ -Verfahren	58
3.2.5	Auflösung der Gleichungssysteme auf jedem Zeitlevel τ_j	58
3.2.6	Praktische Lösung des Gleichungssystems	60
3.3	Konsistenz, Stabilität und Konvergenz	61
3.3.1	Einführung	61
3.3.2	Der Konsistenzfehler des ϑ -Verfahrens	63
3.3.3	Stabilitätsanalyse für das ϑ -Verfahren	64

1 Grundlagen stochastischer Prozesse

1.1 Einführung

1.1.1 Beispiel Teilchenbewegung

Im Folgenden soll die Bewegung eines mikroskopischen Teilchens in einer ruhenden Flüssigkeit beschrieben werden. Die Bewegung des Teilchens in der Flüssigkeit ist auf die Interaktion mit Flüssigkeitsmolekülen zurückzuführen.

Sei $p(t) = m \cdot v(t)$ der Impuls des Teilchens und $v(t)$ seine Geschwindigkeit zur Zeit t . Die Position des Teilchens im $\mathbb{R}^3$ zur Zeit t sei durch $x(t)$ gegeben.

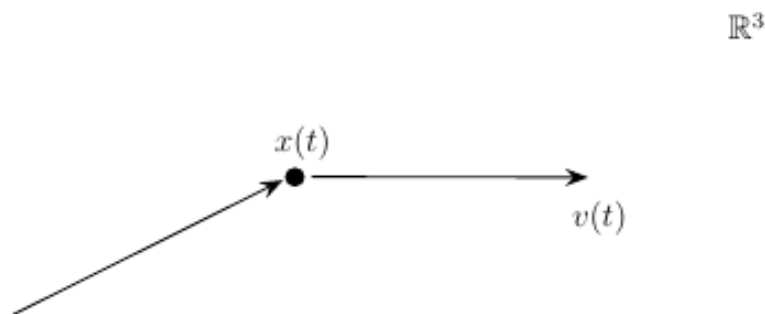


Abbildung 1.1: Impuls des Teilchens.

Newton'sches Prinzip:

$$p'(t) = m \cdot v'(t) = m \cdot x''(t) = F(t),$$

wobei m die Teilchenmasse ist und $F(t)$ die auf das Teilchen wirkende Kraft.

Kraftquellen:

- (i) Interaktion mit Flüssigkeitsmolekülen
- (ii) Schwerkraft
- (iii) Druck, Auftrieb

Die Punkte (ii) und (iii) sind vernachlässigbar.

Annahmen für ein Molekülmodell:

- Flüssigkeitsmoleküle bewegen sich ungeordnet

1.1 Einführung

- die Geschwindigkeitsverteilung hat den Mittelwert 0
- ein Teilchen erfährt ca. 10^8 Molekülstöße pro Sekunde
- ein Molekülstoß entspricht einer extrem kurzen Krafteinwirkung

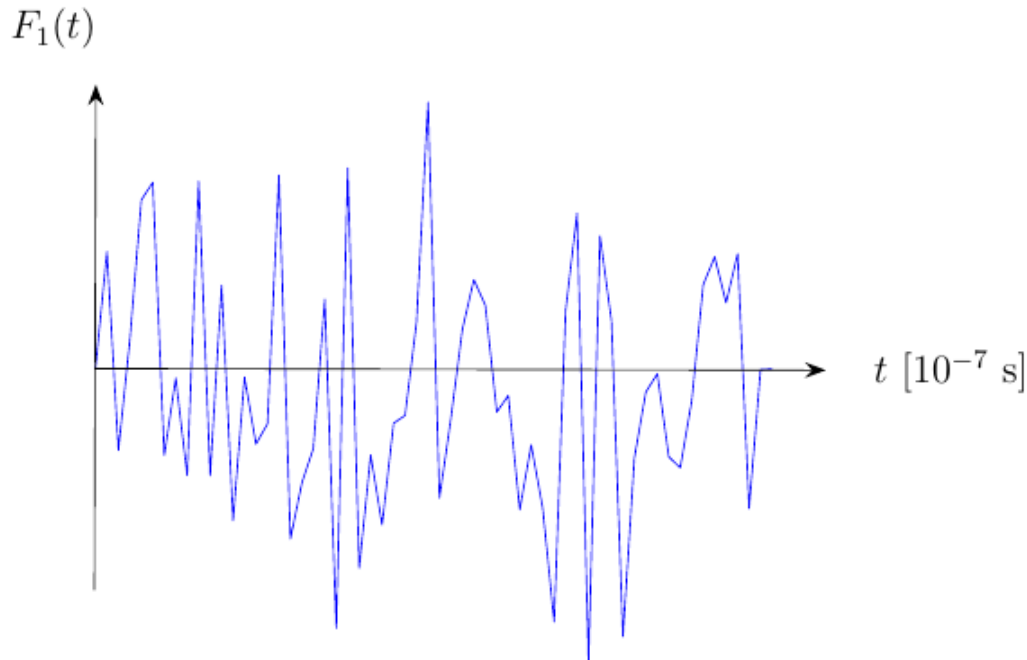


Abbildung 1.2: Krafteinwirkung für die erste Komponente.

Der Prognosezeitraum ist z. B. 100 Sekunden, das entspricht ca. 10^{10} Stöße im Prognosezeitraum. Die Details der Kraftübertragung sind unklar.

1.1.2 Stochastische Modellierung des zeitlich gemittelten Prozesses

Aus $p'(t) = F(t)$ folgt

$$p(t + \Delta t) = p(t) + \int_t^{t+\Delta t} p'(s) ds = p(t) + \int_t^{t+\Delta t} F(s) ds.$$

Ansatz mit N Einzelstößen in $[t, t + \Delta t]$ mit $t_1 := t < t_2 < \dots < t_N < t_{N+1} := t + \Delta t$:

$$\int_t^{t+\Delta t} F(s) ds = \sum_{i=1}^N \int_{t_i}^{t_{i+1}} F(s) ds = \sum_{i=1}^N p_i = N\mu + \sum_{i=1}^N (p_i - \mu),$$

wobei $p_i := \int_{t_i}^{t_{i+1}} F(s) ds$ und μ der mittlere Impulsbeitrag ist.

Annahmen und Konsequenzen:

- $N \sim \Delta t$, solange Δt hinreichend groß ist (z. B. $\Delta t = 10^{-2}$), d. h. $N = \gamma \cdot \Delta t$.

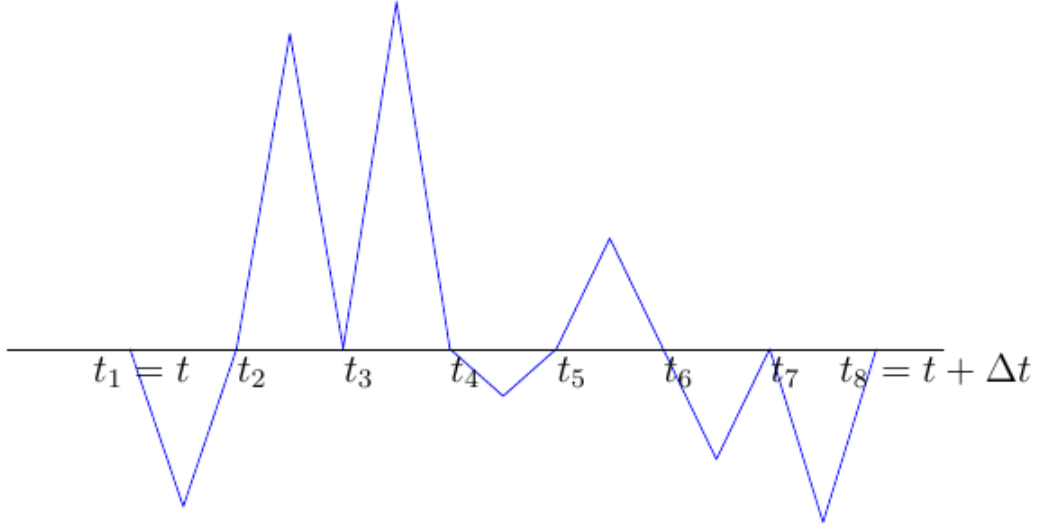


Abbildung 1.3: Beispiel mit $N = 7$ Einzelstößen in $[t, t + \Delta t]$.

- $p_i - \mu$ ist eine unabhängige, identisch verteilte Zufallsvariable (ZV) mit

$$E(p_i - \mu) = 0 \quad \text{und} \quad V(p_i - \mu) = \alpha I < \infty$$

für $i = 1, 2, \dots$

- Für die ZV $S_N := \frac{1}{\sqrt{\alpha N}} \sum_{i=1}^N (p_i - \mu)$ gilt:

$$E(S_N) = \frac{1}{\sqrt{\alpha N}} \sum_{i=1}^N E(p_i - \mu) = 0,$$

$$V(S_N) = \left(\frac{1}{\sqrt{\alpha N}} \right)^2 \sum_{i=1}^N V(p_i - \mu) = \frac{1}{\alpha N} \sum_{i=1}^N \alpha I_3 = I_3.$$

- Für $N \rightarrow \infty$ gilt nach dem zentralen Grenzwertsatz

$$S_N \xrightarrow{d} \xi \quad \text{in Verteilung und} \quad \xi \sim \mathcal{N}(0, I_3)$$

$$\Rightarrow \sum_{i=1}^N (p_i - \mu) = \sqrt{\alpha N} \frac{1}{\sqrt{\alpha N}} \sum_{i=1}^N (p_i - \mu) = \sqrt{\alpha N} \cdot S_N \approx \sqrt{\alpha N} \cdot \xi = \sqrt{\alpha \gamma \Delta t} \cdot \xi,$$

mit $N = \gamma \Delta t$ und der standardnormalverteilten ZV ξ .

- Der mittlere Impulsbeitrag μ entsteht bei Bewegung des Teilchens. Der Impuls ist entgegengesetzt zu $v(t)$ gerichtet, d. h. es gilt für ein $\beta > 0$

$$\mu = -\beta v(t).$$

Also folgt: $N\mu = -\gamma \Delta t \beta v(t)$.

Resultierende Impulsevolution:

$$mv_{n+1} = mv_n + (-\beta \gamma v_n \Delta t) + \sqrt{\alpha \gamma} \sqrt{\Delta t} \xi_n, \quad (1.1)$$

1.1 Einführung

wobei v_n die Geschwindigkeit zum Zeitpunkt $t_n := n\Delta t$ ist und $\xi_1, \xi_2, \dots, \xi_n$ unabhängige, identisch $\mathcal{N}(0, I_3)$ -verteilte ZV sind.

Beachte dabei, dass $(v_n)_{n \in \mathbb{N}}$ deshalb ein zeitdiskreter stochastischer Prozess ist.

1.1.3 Stochastische Prozesse

1.1 Definition (Stochastischer Prozess)

a) Ein stochastischer Prozess ist eine Familie $(X_t)_{t \in I}$ von $\mathbb{R}^d$ -wertigen Zufallsvariablen $X_t : \Omega \rightarrow \mathbb{R}^d$ auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$. Für $I = \mathbb{N}$ spricht man von einem zeitdiskreten, für $I = [a, b]$ von einem zeitkontinuierlichen Prozess.

b) Die Abbildung $t \mapsto X_t(\omega)$ für $\omega \in \Omega$ fest heißt Pfad des Prozesses $(X_t)_{t \in I}$.

Spezialfall: Teilchenmasse m klein im Vergleich zu $\beta\gamma\Delta t$ und $\sqrt{\alpha\gamma\Delta t}$. Setze $m = 0$ in (1.1) und erhalte

$$0 = -\beta\gamma v_n \Delta t + \sqrt{\alpha\gamma} \sqrt{\Delta t} \xi_n,$$

d. h. es gilt

$$\Delta t v_n = \frac{\sqrt{\alpha\gamma}}{\beta\gamma} \sqrt{\Delta t} \xi_n = c \sqrt{\Delta t} \xi_n, \quad (1.2)$$

mit $c := \frac{\sqrt{\alpha\gamma}}{\beta\gamma}$.

Auf dem Ortslevel x finden wir mit (1.2)

$$\begin{aligned} x_{n+1} &= x_n + \Delta t v_n = x_n + c \sqrt{\Delta t} \xi_n, \quad n = 0, 1, 2, \dots \\ x_0 &= 0. \end{aligned} \quad (1.3)$$

Induktion liefert dann

$$x_n = c \sqrt{\Delta t} \sum_{k=0}^{n-1} \xi_k = c b_n, \quad (1.4)$$

mit $b_n := \sqrt{\Delta t} \sum_{k=0}^{n-1} \xi_k$ für $n \geq 1$. Ferner setzen wir $b_0 := 0$.

Da die ξ_k $\mathcal{N}(0, I_3)$ -verteilt sind, gilt $\sqrt{\Delta t} \xi_k \sim \mathcal{N}(0, \Delta t I_3)$. Damit erhalten wir für den Prozess $(b_n)_{n \in \mathbb{N}}$

- $b_0 = 0$,
- für $m > n$ gilt: $b_m - b_n = \sqrt{\Delta t} \sum_{k=n}^{m-1} \xi_k \sim \mathcal{N}(0, \Delta t(m-n)I_3)$.

Die Folge $(b_n)_{n \in \mathbb{N}}$ entspricht einer zeitdiskreten, dreidimensionalen Brownschen Bewegung.

1.2 Definition (Brownsche Bewegung)

a) Eine eindimensionale Brownsche Bewegung ist ein reellwertiger stochastischer Prozess $(W_t)_{t \geq 0}$ auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ mit folgenden Eigenschaften:

- (i) $W_0 = 0$ P -fast sicher,

1.1 Einführung

- (ii) $(W_t)_{t \geq 0}$ hat stetige Pfade $t \mapsto W_t(\omega)$,
- (iii) für $t > s \geq 0$ gilt $(W_t - W_s) \sim \mathcal{N}(0, t - s)$,
- (iv) $(W_t)_{t \geq 0}$ hat unabhängige Zuwächse, d. h. für beliebige $0 \leq t_0 < t_1 < \dots < t_n$ sind

$$W_{t_1} - W_{t_0}, W_{t_2} - W_{t_1}, \dots, W_{t_n} - W_{t_{n-1}}$$

unabhängige ZV.

- b) Als d -dimensionale Brownsche Bewegung bezeichnen wir einen $\mathbb{R}^d$ -wertigen Prozess $(W_t)_{t \geq 0} = (W_t^1, \dots, W_t^d)_{t \geq 0}$, dessen Komponenten $(W_t^i)_{t \geq 0}$ für $i = 1, \dots, d$ unabhängige eindimensionale Brownsche Bewegungen sind.

Für unser Beispiel folgt mit (1.4)

$$x_n = cW_{t_n}, \quad n \in \mathbb{N}.$$

Dabei sollte Δt hinreichend groß sein und $(W_t)_{t \geq 0}$ eine dreidimensionale Brownsche Bewegung sein.

Simulation mit Matlab:

```
% Zweidimensionale Brownsche Bewegung
N = 1000;
dt = 1/N;
figure(1);
W = [0;0];
dW = randn(2,N)*sqrt(dt);
for i = 1:N
    W = W + dW(:,i);
    plot(W(1),W(2), 'b. ');
    axis([-3 3 -3 3]);
    drawnow;
end
```

1.1.4 Verallgemeinerung unseres Beispiels

Gemäß (1.2) und (1.3) gilt

$$x_{n+1} = x_n + \frac{\sqrt{\alpha\gamma}}{\beta\gamma} \sqrt{\Delta t} \xi_n.$$

Sei nun die Impulsvarianz α der Stöße von der Temperatur und damit vom Ort abhängig, d. h. $\alpha = \alpha(x)$. Damit ergibt sich für das Orts- und Geschwindigkeitslevel

$$x_{n+1} = x_n + \Delta t v_n, \tag{1.5}$$

$$v_n \Delta t = \frac{\sqrt{\alpha(x_n)}}{\beta\sqrt{\gamma}} \sqrt{\Delta t} \xi_n. \tag{1.6}$$

Gleichung (1.5) wird für $\Delta t \rightarrow 0$ immer genauer, während Gleichung (1.6) verlangt, dass Δt nicht zu klein ist, damit die Anzahl der Stöße in jedem Δt -Intervall in etwa gleich und

1.1 Einführung

hinreichend groß ist.

Trick: Betrachte nun den Grenzübergang $\Delta t \rightarrow 0$ auf hinreichend langen Zeitintervallen (zentraler Grenzwertsatz muss gültig sein). Mit

$$c(x_n) := \frac{\sqrt{\alpha(x_n)}}{\beta\sqrt{\gamma}} \quad \text{und} \\ W_{t_{n+1}} - W_{t_n} := \sqrt{\Delta t}\xi_n$$

folgt

$$x_{n+1} = x_n + c(x_n) (W_{t_{n+1}} - W_{t_n}), \quad n = 0, 1, 2, \dots \quad (1.7) \\ x_0 = 0.$$

1.3 Bemerkung

Die Differenzengleichung (1.7) für x_n kann **nicht** in eine Differentialgleichung für einen Grenzprozess X_t umgewandelt werden! Betrachte hierzu

$$\frac{x_{n+1} - x_n}{t_{n+1} - t_n} = c(x_n) \frac{W_{t_{n+1}} - W_{t_n}}{t_{n+1} - t_n}.$$

Dann folgt

$$V\left(\frac{W_{t_{n+1}} - W_{t_n}}{t_{n+1} - t_n}\right) = V\left(\frac{W_{t_{n+1}} - W_{t_n}}{\Delta t}\right) = \frac{1}{\Delta t^2} V(W_{t_{n+1}} - W_{t_n}) \\ = \frac{1}{\Delta t^2} (t_{n+1} - t_n) I_3 = \frac{1}{\Delta t} I_3 \rightarrow \infty \quad \text{für } t \rightarrow 0.$$

Da eine Umwandlung in eine Differentialgleichung nicht möglich ist, wird die Differenzengleichung (1.7) in eine Integralgleichung umgewandelt:

$$x_N - x_0 = \sum_{n=0}^{N-1} (x_{n+1} - x_n) = \sum_{n=0}^{N-1} c(x_n) \Delta W_{t_n},$$

mit $\Delta W_{t_n} := W_{t_{n+1}} - W_{t_n}$. Setze nun $t := t_N$, bilde formal den Grenzwert $\Delta t \rightarrow 0$, $N \cdot \Delta t = t$ und „finde“

$$\sum_{n=0}^{N-1} c(x_n) \Delta W_{t_n} \xrightarrow{\Delta t \rightarrow 0} \int_0^t c(X_s) dW_s.$$

Den Ausdruck auf der rechten Seite nennen wir „stochastisches Integral“. Wir erhalten also formal die Gleichheit

$$X_t = X_0 + \int_0^t c(X_s) dW_s, \quad (1.8)$$

falls X_t eindeutig bestimmt ist. (1.8) ist eine Integralgleichung für den Grenzprozess X_t . Da der Integrator selbst ein stochastischer Prozess ist, müssen wir diesem sogenannten „stochastischen Integral“ einen Sinn geben.

1.2 Stochastisches Integral (Itô-Integral) und Itô-Prozesse

1.2.1 Beispiel

Wir betrachten wieder unser Beispiel aus dem letzten Kapitel:

$$\begin{aligned} x_{n+1} &= x_n + c(x_n)\Delta W_{t_n} \quad (n \in \mathbb{N}), \\ x_0 &= 0. \end{aligned}$$

Dies ist äquivalent zu

$$x_N = \sum_{n=0}^{N-1} c(x_n)\Delta W_{t_n} \quad (N \in \mathbb{N}). \quad (1.9)$$

Beobachtung:

- x_n hängt von ΔW_{t_i} ab, $i = 0, \dots, n-1$,
- der Zuwachs ΔW_{t_n} ist unabhängig von x_n .

Betrachte nun die Summe in (1.9) als Integral einer zeitlichen Treppenfunktion. Sei $0 = t_0 < t_1 < \dots < t_N$, und sei

$$f_t(\omega) := \sum_{n=0}^{N-1} c(x_n(\omega))\mathbb{I}_{[t_n, t_{n+1}[}(t) + c(x_N(\omega))\mathbb{I}_{\{t_N\}},$$

mit der Indikatorfunktion

$$\mathbb{I}_M(t) := \begin{cases} 1, & \text{für } t \in M, \\ 0, & \text{sonst.} \end{cases}$$

Damit gilt bezüglich Integration nach der Zeit

$$\int_0^{t_N} f_s ds = \sum_{n=0}^{N-1} c(x_n)(t_{n+1} - t_n) = \sum_{n=0}^{N-1} c(x_n)\Delta t_n \quad \text{mit } \Delta t_n = t_{n+1} - t_n.$$

Im Folgenden soll jetzt dem Ausdruck $\int_a^b f_t dW_t$ ein Sinn gegeben werden.

1.2.2 Technische Voraussetzungen dazu

- Sei $(\Omega, \mathcal{F}, \mathcal{P})$ ein Wahrscheinlichkeitsraum.
- $(t_0, t_1, \dots, t_N)$ sei eine Zerlegung von $[a, b]$ mit $a = t_0 < t_1 < \dots < t_N = b$.
- $(W_t)_{t \geq 0}$ sei eine Brownsche Bewegung, wobei W_t $\mathcal{A}_t$ -messbar ist mit einer Familie $(\mathcal{A})_{t \geq 0}$ von Unter- σ -Algebren von $\mathcal{F}$ mit $\mathcal{A}_s \subset \mathcal{A}_t$, falls $s < t$ ($(\mathcal{A})_{t \geq 0}$ ist eine sogenannte Filtration von $\mathcal{F}$). Ferner enthalte $\mathcal{A}_0$ alle $\mathcal{F}$ -Nullmengen und $W_t - W_s$ sei von $\mathcal{A}_s$ unabhängig, falls $s < t$.
- $e_0, \dots, e_N$ seien ZV, wobei e_j $\mathcal{A}_{t_j}$ -messbar sei, $j = 0, \dots, N$.

1.2 Stochastisches Integral (Itô-Integral) und Itô-Prozesse

Sei f eine Treppenfunktion mit zufälligen Sprunghöhen, d. h.

$$f_t(\omega) := \sum_{j=0}^{N-1} e_j(\omega) \mathbb{I}_{[t_j, t_{j+1}[}(t) + e_N(\omega) \mathbb{I}_{\{t_N\}}(t), \quad (1.10)$$

In Abbildung ?? ist eine Treppenfunktion für $N = 5$ dargestellt mit vorgegebenen Sprungstellen t_i ($i = 0, \dots, 5$). Hierbei ist z. B. die Sprunghöhe $e_2(\omega) - e_1(\omega)$ zufällig.

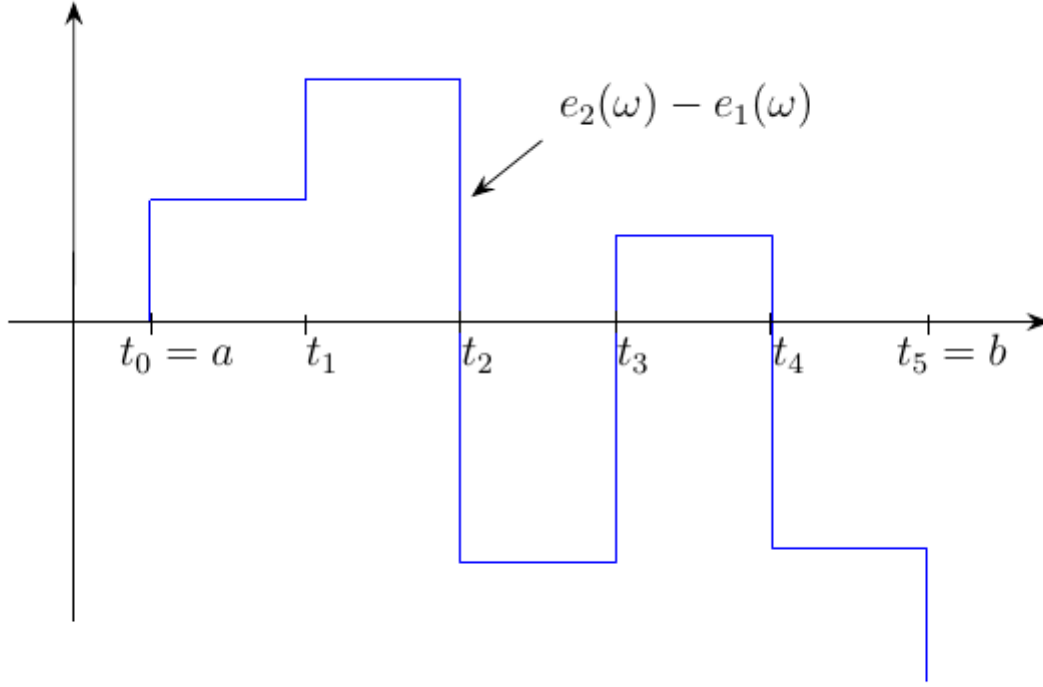


Abbildung 1.4: Beispiel einer Treppenfunktion für $N = 5$.

Für eine Treppenfunktion wie in (1.10) setzen wir

$$\int_a^b f_t dW_t := \sum_{j=0}^{N-1} e_j (W_{t_{j+1}} - W_{t_j}) = \sum_{j=0}^{N-1} e_j \Delta W_{t_j}, \quad (1.11)$$

mit $\Delta W_{t_j} := (W_{t_{j+1}} - W_{t_j})$.

1.2.3 Eigenschaften des Itô-Integrals

Das folgende Lemma zeigt zwei zentrale Eigenschaften des Itô-Integrals von Treppenfunktionen für eine Fortsetzung auf allgemeine Prozesse.

1.4 Lemma

Sei f eine Treppenfunktion der Form (1.10) mit $e_j \in L^2(\Omega)$, $j = 0, \dots, N$. Dann gelten

1.2 Stochastisches Integral (Itô-Integral) und Itô-Prozesse

folgende Gleichheiten:

$$\begin{aligned} i) \quad & E \left(\int_a^b f_t dW_t \right) = 0 \quad (\text{„Erwartungswert“}), \\ ii) \quad & E \left(\left(\int_a^b f_t dW_t \right)^2 \right) = \int_a^b E(f_t^2) dt \quad (\text{„Varianz (Itô-Isometrie)“}). \end{aligned}$$

Beweis. Sei $f_t(\omega) := \sum_{j=0}^{N-1} e_j(\omega) \mathbb{I}_{[t_j, t_{j+1}[}(t) + e_N(\omega) \mathbb{I}_{\{t_N\}}(t)$ mit $e_j \in L^2(\Omega)$ für $j = 0, \dots, N$, wobei $L^2(\Omega) = \{f : \Omega \rightarrow \mathbb{R}, f \text{ messbar}, f^2 \text{ integrierbar}\}$.

Zu i):

$$E \left(\int_a^b f_t dW_t \right) = E \left(\sum_{j=0}^{N-1} e_j \Delta W_{t_j} \right) = \sum_{j=0}^{N-1} E(e_j) E(\Delta W_{t_j}) = 0,$$

da e_j und ΔW_{t_j} unabhängig sind und $E(\Delta W_{t_j}) = 0$.

Zu ii): Es gilt für Treppenfunktionen

$$\left(\int_a^b f_t dW_t \right)^2 = \sum_{i,j=0}^{N-1} e_i e_j \Delta W_{t_i} \Delta W_{t_j} = \sum_{i=0}^{N-1} e_i^2 (\Delta W_{t_i})^2 + 2 \sum_{i < j} e_i e_j \Delta W_{t_i} \Delta W_{t_j}.$$

Nun sind e_i und ΔW_{t_i} unabhängig, und damit auch e_i^2 und $\Delta W_{t_i}^2$. Also folgt

$$E(e_i^2 \Delta W_{t_i}^2) = E(e_i^2) E(\Delta W_{t_i}^2).$$

Da $e_i \in L^2(\Omega)$, gilt $E(e_i^2) < \infty$. Mit $E(\Delta W_{t_i}^2) = E(\Delta W_{t_i})^2 + V(\Delta W_{t_i}) = 0 + t_{i+1} - t_i = \Delta t_i$ folgt

$$E(e_i^2 \Delta W_{t_i}^2) = \Delta t_i E(e_i^2) < \infty. \quad (1.12)$$

Nach der Hölderschen Ungleichung folgt mit $e_i \in L^2(\Omega)$ außerdem $e_i e_j \Delta W_{t_j} \in L^1(\Omega)$ und somit gilt

$$E(e_i e_j \Delta W_{t_i} \Delta W_{t_j}) = E(e_i e_j \Delta W_{t_i}) \underbrace{E(\Delta W_{t_j})}_{=0} = 0.$$

Damit erhalten wir

$$\begin{aligned} E \left(\left(\int_a^b f_t dW_t \right)^2 \right) &= E \left(\sum_{i,j=0}^{N-1} e_i e_j \Delta W_{t_i} \Delta W_{t_j} \right) = E \left(\sum_{i=0}^{N-1} e_i^2 (\Delta W_{t_i})^2 \right) \\ &= \sum_{i=0}^{N-1} E(e_i^2 (\Delta W_{t_i})^2) \stackrel{(1.12)}{=} \sum_{i=0}^{N-1} \Delta t_i E(e_i^2) = \sum_{i=0}^{N-1} \Delta t_i E(f_t^2|_{[t_i, t_{i+1}[}) \\ &= \int_a^b E(f_t^2) dt, \end{aligned}$$

da f_t eine Treppenfunktion ist. □

1.2.4 Prinzipielles weiteres Vorgehen

Ist $(f^{(n)})_{n \in \mathbb{N}}$ eine Folge von Treppenfunktionen mit

$$f^{(n)} \rightarrow f \quad \text{in } L^2([a, b] \times \Omega),$$

so ist $\left(\int_a^b f_t^{(n)} dW_t\right)_{n \in \mathbb{N}}$ eine Cauchy-Folge in $L^2(\Omega)$, d. h. es gibt einen Grenzwert dieser Folge in $L^2(\Omega)$, der mit $\int_a^b f_t dW_t$ bezeichnet wird (Itô-Integral).

Des Weiteren soll das Itô-Integral auf pfadweise L^2 -Prozesse übertragen werden.

1.5 Definition

Sei $(\Omega, \mathcal{F}, P)$ ein Wahrscheinlichkeitsraum, $p \in [1, \infty[$ und $[a, b] \subset \mathbb{R}$ ein Intervall. Weiter sei $L_\omega^p([a, b])$ die Menge aller messbaren Funktionen $f : [a, b] \times \Omega \rightarrow \mathbb{R}$ mit

- $\int_a^b |f|^p(t, \omega) dt < \infty$ P -fast sicher,
- $f_t := f(t, \cdot)$ ist $\mathcal{A}_t$ -messbar für $a \leq t \leq b$.

Für $f \in L_\omega^p([a, b])$ setzen wir

$$\int_a^b f_t dW_t := P\text{-}\lim_{n \rightarrow \infty} \int_a^b f_t^{(n)} dW_t$$

wobei $(f^{(n)})_{n \in \mathbb{N}}$ eine Folge von Treppenfunktionen ist mit

$$P\text{-}\lim_{n \rightarrow \infty} \int_a^b |f_t^{(n)} - f_t|^2 dt = 0.$$

In der letzten Definition taucht der Begriff der stochastischen Konvergenz auf, dazu folgende

1.6 Definition (Stochastische Konvergenz)

Für eine Folge von ZV $(X_n)_{n \in \mathbb{N}}$ und eine ZV X definiert man

$$P\text{-}\lim_{n \rightarrow \infty} X_n = X \quad :\Longleftrightarrow \quad \forall \varepsilon > 0 : \lim_{n \rightarrow \infty} P(|X_n - X| > \varepsilon) = 0.$$

Ausgeschrieben ist dies äquivalent zu

$$\forall \varepsilon > 0 : \lim_{n \rightarrow \infty} P(\{\omega \in \Omega : |X_n(\omega) - X(\omega)| > \varepsilon\}) = 0.$$

1.2.5 Rechenregeln für stochastische Integrale

Wir notieren einige Rechenregeln für stochastische Integrale:

- Linearität: Für $f, g \in L_\omega^2([a, b])$, $\alpha, \beta \in \mathbb{R}$ gilt

$$\int_a^b \alpha f_t + \beta g_t dW_t = \alpha \int_a^b f_t dW_t + \beta \int_a^b g_t dW_t.$$

- Für $I \subset [a, b]$ messbar setzen wir

$$\int_I f_t dW_t := \int_a^b \mathbb{I}_I(t) f_t dW_t.$$

1.2 Stochastisches Integral (Itô-Integral) und Itô-Prozesse

- Teleskopsumme:

$$\int_a^b dW_t = W_b - W_a.$$

- Erwartungswert:

$$E \left(\int_a^b f_t dW_t \right) = 0.$$

- Varianz (Itô-Isometrie):

$$E \left(\left(\int_a^b f_t dW_t \right)^2 \right) = \int_a^b E(f_t^2) dt.$$

1.7 Bemerkung

Diese Eigenschaften folgen aus den entsprechenden Eigenschaften der Treppenfunktionen und der Konvergenz von

$$\int_a^b f_t^{(n)} dW_t \rightarrow \int_a^b f_t dW_t \quad \text{in } L^2(\Omega).$$

1.2.6 Itô-Prozesse

1.8 Definition

Ein Itô-Prozess $(X_t)_{a \leq t \leq b}$ ist ein stetiger $\mathcal{A}_t$ -adaptierter Prozess, der für $a \leq t \leq b$ eine Darstellung der Form

$$X_t = X_a + \int_a^t f_\tau d\tau + \int_a^t g_\tau dW_\tau \quad P\text{-fast sicher}$$

besitzt mit $f \in L_\omega^1([a, b])$ und $g \in L_\omega^2([a, b])$.

1.9 Bemerkung

a) Wegen $W_t = W_a + \int_a^t dW_\tau$ ist die Brownsche Bewegung ein Itô-Prozess mit $f = 0$ und $g = 1$.

b) Mit $\mathbb{I}_{[a,t]} = \mathbb{I}_{[a,s]} + \mathbb{I}_{[s,t]}$ folgt für $a \leq s \leq t$

$$X_t = X_a + \underbrace{\int_a^s f_\tau d\tau + \int_a^s g_\tau dW_\tau}_{=X_s} + \int_s^t f_\tau d\tau + \int_s^t g_\tau dW_\tau,$$

d. h. es gilt

$$X_t - X_s = \int_s^t f_\tau d\tau + \int_s^t g_\tau dW_\tau. \quad (1.13)$$

1.2 Stochastisches Integral (Itô-Integral) und Itô-Prozesse

Gleichung (1.13) besagt, dass der Zuwachs $X_t - X_s$ selbst ein Itô-Prozess ist. Wir schreiben in diesem Fall

$$X_t - X_s =: \int_s^t dX_\tau.$$

Konsequenz:

$$\int_s^t dX_\tau = \int_s^t f_\tau d\tau + \int_s^t g_\tau dW_\tau.$$

Abkürzende Schreibweise:

$$dX_t = f_t dt + g_t dW_t. \quad (1.14)$$

Die Gleichung (1.14) wird zur Differentialgleichung (DGL), wenn f oder g von X abhängen, d. h. $f_t = A(t, X_t)$ oder $g_t = B(t, X_t)$. Man erhält damit die „stochastische DGL“ (SDGL)

$$dX_t = A(t, X_t)dt + B(t, X_t)dW_t. \quad (1.15)$$

Bezeichnung:

$A(t, X_t)dt$ heißt „Drift-Term“ und $B(t, X_t)dW_t$ „Diffusionsterm“.

1.10 Definition

Seien $A, B \in C([0, T] \times \mathbb{R})$. Ein stetiger $\mathcal{A}_t$ -adaptierter Prozess $(X_t)_{0 \leq t \leq T}$ erfüllt die SDGL (1.15) mit Anfangswert X_0 , falls für $0 \leq t \leq T$

$$X_t = X_0 + \int_0^t A(s, X_s)ds + \int_0^t B(s, X_s)dW_s \quad P\text{-fast sicher.}$$

1.11 Bemerkung

$A, B \in C([0, T] \times \mathbb{R})$ und $(X_t)_{0 \leq t \leq T}$ Itô-Prozess impliziert

$$A(\cdot, X_\cdot) \in L_\omega^1([0, T]), \quad B(\cdot, X_\cdot) \in L_\omega^2([0, T]).$$

1.12 Satz (Existenz- und Eindeutigkeitssatz (ohne Beweis))

Es gelte

- (i) $A, B \in C([0, T] \times \mathbb{R})$,
- (ii) $|A(t, x)| \leq K(1 + |x|)$ und $|B(t, x)| \leq K(1 + |x|)$ für $x \in \mathbb{R}$ („lineares Wachstum“),
- (iii) $|A(t, x) - A(t, y)| \leq L|x - y|$ und $|B(t, x) - B(t, y)| \leq L|x - y|$ gleichmäßig in $[0, T]$ für $x, y \in \mathbb{R}$ („Lipschitz-Bedingung“).

Ist dann $E(X_0^2) < \infty$, so besitzt die SDGL

$$dX_t = A(t, X_t)dt + B(t, X_t)dW_t$$

eine eindeutige Lösung $(X_t)_{0 \leq t \leq T}$ auf $[0, T]$ mit Anfangswert X_0 .

1.13 Bemerkung

Die Bedingung (ii) aus obigem Satz ist einschränkend, es gibt aber auch allgemeinere Versionen von diesem Satz.

1.2.7 Das Lemma von Itô

1.14 Satz (Itô's Lemma (ohne Beweis))

Seien $f \in L^1_\omega([a, b])$, $g \in L^2_\omega([a, b])$ und X_t ein Itô-Prozess mit $dX_t = f_t dt + g_t dW_t$. Ist $F \in C^{1,2}([a, b] \times \mathbb{R})$ und $F_t := F(t, X_t)$, so gilt die sogenannte Itô-Formel:

$$\begin{aligned} dF(t, X_t) &= \left(\frac{\partial F}{\partial t}(t, X_t) + \frac{\partial F}{\partial X}(t, X_t)f_t + \frac{1}{2} \frac{\partial^2 F}{\partial X^2}(t, X_t)g_t^2 \right) dt + \frac{\partial F}{\partial X}(t, X_t)g_t dW_t \\ &= \frac{\partial F}{\partial t}(t, X_t)dt + \frac{\partial F}{\partial X}(t, X_t)dX_t + \frac{1}{2} \frac{\partial^2 F}{\partial X^2}(t, X_t)g_t^2 dt. \end{aligned}$$

1.15 Bemerkung

Die Itô-Formel entspricht der Kettenregel aus der Differentialrechnung. Falls f , g und w glatt sind, dann gilt

$$\int_s^t g dw = \int_s^t g_r w'(r) dr.$$

Damit gilt für den Prozess X_t

$$\begin{aligned} X_t &= X_0 + \int_0^t f_r dr + \int_0^t g_r w'(r) dr \quad \text{bzw. in Differentialform} \\ dX_t &= f_t dt + g_t w'_t dt, \end{aligned}$$

und wir erhalten

$$\begin{aligned} \frac{d}{dt} F(t, X_t) &= \frac{\partial F}{\partial t}(t, X_t) + \frac{\partial F}{\partial X}(t, X_t)(f_t + g_t w'_t) \quad \text{bzw.} \\ dF_t &= \frac{\partial F}{\partial t} dt + \frac{\partial F}{\partial X} dX_t. \end{aligned}$$

Es gibt also einen Zusatzterm $\frac{1}{2} \frac{\partial^2 F}{\partial X^2} g_t^2 dt$ in der Itô-Formel. Dieser entsteht, weil W_t nicht glatt ist. Die Brownsche Bewegung ist so „zappelig“, dass der Zusatzterm nicht vernachlässigbar ist.

1.16 Korollar (Produktregel)

Seien X_t und Y_t zwei Itô-Prozesse mit

$$\begin{aligned} dX_t &= f_t^X dt + g_t^X dW_t, \\ dY_t &= f_t^Y dt + g_t^Y dW_t. \end{aligned}$$

Dann gilt

$$d(X_t \cdot Y_t) = X_t dY_t + Y_t dX_t + g_t^X g_t^Y dt.$$

Beweis.

$$\begin{aligned} d(X_t \cdot Y_t) &= d \left(\frac{1}{4} (4X_t Y_t + X_t^2 - X_t^2 + Y_t^2 - Y_t^2) \right) \\ &= \frac{1}{4} [d((X_t + Y_t)^2) - d((X_t - Y_t)^2)]. \end{aligned}$$

1.3 Modellierung von Prozessen mit Rauschen

Wende nun Itô's Lemma an mit $Z_t := X_t + Y_t$, $F_1(t, Z) := Z^2$ und $W_t := X_t - Y_t$, $F_2(t, W) := W^2$. Damit folgt dann

$$\begin{aligned} d(X_t \cdot Y_t) &= \frac{1}{4} \left[2(X_t + Y_t)d(X_t + Y_t) + 2 \cdot \frac{1}{2}(g_t^X + g_t^Y)^2 dt - 2(X_t - Y_t)d(X_t - Y_t) \right. \\ &\quad \left. - 2 \cdot \frac{1}{2}(g_t^X - g_t^Y)^2 dt \right] \\ &= \frac{1}{4} [2X_t dY_t + 2Y_t dX_t + 2g_t^X g_t^Y dt + 2X_t dY_t + 2Y_t dX_t + 2g_t^X g_t^Y dt] \\ &= X_t dY_t + Y_t dX_t + g_t^X g_t^Y dt. \end{aligned}$$

□

1.3 Modellierung von Prozessen mit Rauschen

1.3.1 Additives Rauschen

Ein Beispiel für additives Rauschen ist unser Startbeispiel, d. h. die Bewegung eines (masselosen) Mikroteilchens in einer Flüssigkeit. Die dazugehörige SDGL lautet

$$\begin{aligned} dX_t &= aX_t dt + b(t)dW_t, \\ X_0 &= \bar{X}, \end{aligned} \tag{1.16}$$

bzw. in Integralform

$$X_t - X_0 = \int_0^t aX_s ds + \int_0^t b(r)dW_r.$$

Der Term $b(t)dW_t$ wird als **additives Rauschen** bezeichnet.

1.17 Lemma

Die Lösung der Anfangswertaufgabe (1.16) lautet

$$X_t = \exp(at)\bar{X} + \exp(at) \int_0^t \exp(-as)b(s)dW_s.$$

Beweis. Definiere die Itô-Prozesse

$$\begin{aligned} Y_t &:= \exp(at) = 1 + \int_0^t a \exp(as) ds + \int_0^t \underbrace{0}_{=g_s^Y} dW_s, \\ Z_t &:= \int_0^t \exp(-as)b(s)dW_s = 0 + \int_0^t 0 ds + \int_0^t \underbrace{\exp(-as)b(s)}_{=g_s^Z} dW_s \end{aligned}$$

und finde

$$X_t = \bar{X}Y_t + Y_t Z_t.$$

1.3 Modellierung von Prozessen mit Rauschen

Bildet man nun das totale Differential unter Benutzung von Korollar 1.16 und $d\bar{X} = 0$, so folgt

$$\begin{aligned} dX_t &= d(\bar{X}Y_t) + d(Y_tZ_t) = \bar{X}dY_t + Y_t dZ_t + dY_t Z_t + g_t^Y g_t^Z dt \\ &= \bar{X}(aY_t dt) + Y_t(\exp(-at)b(t)dW_t) + aY_t dt Z_t + 0 \\ &= (\bar{X}aY_t + aY_t Z_t)dt + Y_t \exp(-at)b(t)dW_t \\ &= aX_t dt + \exp(at) \exp(-at)b(t)dW_t = aX_t dt + b(t)dW_t. \end{aligned}$$

Ferner gilt $X_0 = \bar{X} \underbrace{Y_0}_{=1} + Y_0 \underbrace{Z_0}_{=0} = \bar{X}$. □

1.3.2 Multiplikatives Rauschen

Im Gegensatz zum additiven Rauschen nimmt die Volatilität des stochastischen Terms mit der Größe der Variable X_t zu bzw. ab. In der Finanzökonomie verwendet man multiplikatives Rauschen oft um Aktienkurse zu modellieren. Die SDGL lautet in diesem Fall

$$\begin{aligned} dX_t &= (\alpha_t + \mu_t X_t)dt + (\beta_t + \sigma_t X_t)dW_t, \\ X_0 &= \bar{X}. \end{aligned} \tag{1.17}$$

Unterteile (1.17) in einen homogenen und einen inhomogenen Teil:

$$\begin{aligned} dX_t &= \underbrace{\mu_t X_t dt + \sigma_t X_t dW_t}_{\text{homogener Teil}} + \underbrace{\alpha_t dt + \beta_t dW_t}_{\text{inhomogener Teil}}, \\ X_0 &= \bar{X}. \end{aligned}$$

Wie bei gewöhnlichen Differentialgleichungen finden wir auch hier die Lösung mittels Variation der Konstanten.

1.18 Lemma

Die Lösung der Anfangswertaufgabe (1.17) lautet

$$X_t = c_t H_t,$$

wobei

$$\begin{aligned} c_t &= \bar{X} + \int_0^t \frac{1}{H_s} (\alpha_s - \beta_s \sigma_s) ds + \int_0^t \frac{1}{H_s} \beta_s dW_s, \\ H_t &= \exp \left(\int_0^t \left(\mu_s - \frac{1}{2} \sigma_s^2 \right) ds + \int_0^t \sigma_s dW_s \right). \end{aligned}$$

Beweis. (i) Löse die homogene Gleichung

$$\begin{aligned} dH_t &= \mu_t H_t dt + \sigma_t H_t dW_t, \\ H_0 &= 1. \end{aligned} \tag{1.18}$$

Dabei ist (1.18) ein Itô-Prozess mit $f_t^H := \mu_t H_t$ und $g_t^H := \sigma_t H_t$. Wende nun die Itô-Formel an auf $F(t, H_t) := \ln H_t$ mit $\frac{\partial}{\partial H} F = \frac{1}{H}$ sowie $\frac{\partial^2}{\partial H^2} F = -\frac{1}{H^2}$ und finde

$$\begin{aligned} d(\ln H_t) &= \frac{1}{H_t} dH_t + \frac{1}{2} \frac{(-1)}{H_t^2} \underbrace{(g_t^H)^2}_{=(\sigma_t H_t)^2} dt \stackrel{(1.18)}{=} \frac{1}{H_t} (\mu_t H_t dt + \sigma_t H_t dW_t) - \frac{\sigma_t^2 H_t^2}{2H_t^2} dt \\ &= \left(\mu_t - \frac{\sigma_t^2}{2} \right) dt + \sigma_t dW_t. \end{aligned}$$

1.3 Modellierung von Prozessen mit Rauschen

Dies bedeutet

$$\begin{aligned}\ln(H_t) &= \ln(H_0) + \int_0^t \left(\mu_s - \frac{1}{2} \sigma_s^2 \right) ds + \int_0^t \sigma_s dW_s, \text{ d. h.} \\ H_t &= \underbrace{H_0}_{=1} \cdot \exp \left(\int_0^t \left(\mu_s - \frac{1}{2} \sigma_s^2 \right) ds + \int_0^t \sigma_s dW_s \right) \\ &= \exp \left(\int_0^t \left(\mu_s - \frac{1}{2} \sigma_s^2 \right) ds + \int_0^t \sigma_s dW_s \right).\end{aligned}\tag{1.19}$$

(Gleichung (1.19) impliziert $H_t > 0$, falls $H_0 > 0$. Man beachte, dass dies eine notwendige Bedingung ist für einen Aktienpreis.)

(ii) Lösung der inhomogenen Gleichung: Mache dazu den Ansatz

$$X_t = c_t H_t, \text{ d. h. } c_t = \frac{X_t}{H_t}.$$

Wähle nun c_t geeignet, d. h. so, dass X_t Gleichung (1.17) löst. Für $c_t = \frac{X_t}{H_t}$ folgt mit der Produktregel

$$dc_t = \frac{1}{H_t} dX_t + X_t d\left(\frac{1}{H_t}\right) + g_t^X g_t^{\frac{1}{H}} dt.\tag{1.20}$$

Mit $G(t, Z) := \frac{1}{Z}$, $\frac{\partial}{\partial Z} G = -\frac{1}{Z^2}$, $\frac{\partial^2}{\partial Z^2} G = \frac{2}{Z^3}$ folgt mit der Itô-Formel

$$\begin{aligned}d\left(\frac{1}{H_t}\right) &= -\frac{1}{H_t^2} dH_t + \frac{1}{2} \frac{2}{H_t^3} (g_t^H)^2 dt \stackrel{(1.18)}{=} -\frac{1}{H_t^2} (\mu_t H_t dt + \sigma_t H_t dW_t) + \frac{1}{H_t^3} \sigma_t^2 H_t^2 dt \\ &= -\frac{\mu_t}{H_t} dt - \underbrace{\frac{\sigma_t}{H_t} dW_t}_{=: g_t^{\frac{1}{H}}} + \frac{\sigma_t^2}{H_t} dt.\end{aligned}$$

Setze jetzt die letzte Gleichung in (1.20) ein und erhalte

$$\begin{aligned}dc_t &= \frac{1}{H_t} (\alpha_t + \mu_t X_t) dt + \frac{1}{H_t} (\beta_t + \sigma_t X_t) dW_t + X_t \left(-\frac{\mu_t}{H_t} dt + \frac{\sigma_t^2}{H_t} dt - \frac{\sigma_t}{H_t} dW_t \right) \\ &\quad + \underbrace{(\beta_t + \sigma_t X_t)}_{=: g_t^X} \underbrace{\left(-\frac{\sigma_t}{H_t} \right)}_{=: g_t^{\frac{1}{H}}} dt \\ &= \frac{1}{H_t} (\alpha_t - \beta_t \sigma_t) dt + \frac{1}{H_t} \beta_t dW_t.\end{aligned}$$

Somit finden wir mit $c_0 = \bar{X}$ sofort

$$c_t = \bar{X} + \int_0^t \frac{1}{H_s} (\alpha_s - \beta_s \sigma_s) ds + \int_0^t \frac{1}{H_s} \beta_s dW_s.$$

□

1.3.3 Modellierung in der Finanzökonomie

Als Beispiel betrachten wir nun die Modellierung risikobehafteter Papiere in der Finanzökonomie. Für den Wert risikiofreier Papiere setzt man an

$$\frac{\Delta B(t)}{B(t)} = r\Delta t,$$

wobei $\frac{\Delta B(t)}{B(t)}$ der relative Wertzuwachs im Intervall $[t, t + \Delta t]$ ist, d. h. $\Delta B(t) = B(t + \Delta t) - B(t)$. r ist die Zinsrate bei kontinuierlicher Verzinsung. Bildet man den Grenzwert $\Delta t \rightarrow 0$, so erhält man

$$dB(t) = rB(t)dt.$$

Modelliere nun ein risikobehaftetes Papier als Überlagerung eines risikolosen Papiers mit zufälligem relativen Wertzuwachs

$$\frac{\Delta S(t)}{S(t)} = \mu\Delta t + \sigma\Delta W_t.$$

Der Grenzwert $\Delta t \rightarrow 0$ liefert

$$\begin{aligned} dS(t) &= \mu S(t)dt + \sigma S(t)dW_t, \\ S_0 &= \bar{S}. \end{aligned}$$

Dies entspricht der SDGL (1.17) mit $\alpha_t = 0, \mu_t = \mu, \beta_t = 0, \sigma_t = \sigma$. Die Lösung lautet hier also nach Lemma 1.18

$$\begin{aligned} S(t) &= \bar{S} \cdot \exp \left(\int_0^t \left(\mu - \frac{1}{2}\sigma^2 \right) ds + \int_0^t \sigma dW_s \right) \\ &= \bar{S} \cdot \exp \left(\left(\mu - \frac{1}{2}\sigma^2 \right) t + \sigma W_t \right), \end{aligned}$$

denn $\int_0^t (\mu - \frac{1}{2}\sigma^2) ds = (\mu - \frac{1}{2}\sigma^2) t$ und $\int_0^t \sigma dW_s = \sigma(W_t - W_0) = \sigma W_t$.

1.19 Bemerkung

Bei komplizierteren Modellen können μ und σ zeit- und auch preisabhängig sein.

1.3.4 Ausgangspunkt für die Numerik

Typische Aufgabenstellungen: Berechnung von Erwartungswerten von Prozessen, welche durch eine SDGL gesteuert werden.

Sei $(X_t)_{0 \leq t \leq T}$ ein Itô-Prozess, so gilt für $0 \leq t \leq T$

$$E(g(X_t)) = \int_{\Omega} g(X(t, \omega)) dP(\omega).$$

Häufig ist $(X_t)_{0 \leq t \leq T}$ durch eine skalare oder vektorwertige SDGL gegeben, z. B.

$$\begin{aligned} dX_t &= a(t, X_t)dt + b(t, X_t)dW_t, \\ X_0 &= \bar{X}. \end{aligned} \tag{1.21}$$

1.3 Modellierung von Prozessen mit Rauschen

Im Allgemeinen ist $a(t, X_t) \in \mathbb{R}^d$ und $b(t, X_t) \in \mathbb{R}^{d \times m}$ und W_t ist eine m -dimensionale Brownsche Bewegung. Gleichung (1.21) ist eine Abkürzung für

$$(X_t)_i = \bar{X}_i + \int_0^t a_i(s, X_s) ds + \sum_{j=1}^m \int_0^t b_{ij}(s, X_s) (dW_s)_j,$$
$$(X_0)_i = \bar{X}_i,$$

für $i = 1, \dots, d$.

Beispiele:

- Innere Spannung eines Polymermoleküls in einer Flüssigkeit
- Optionspreisberechnung

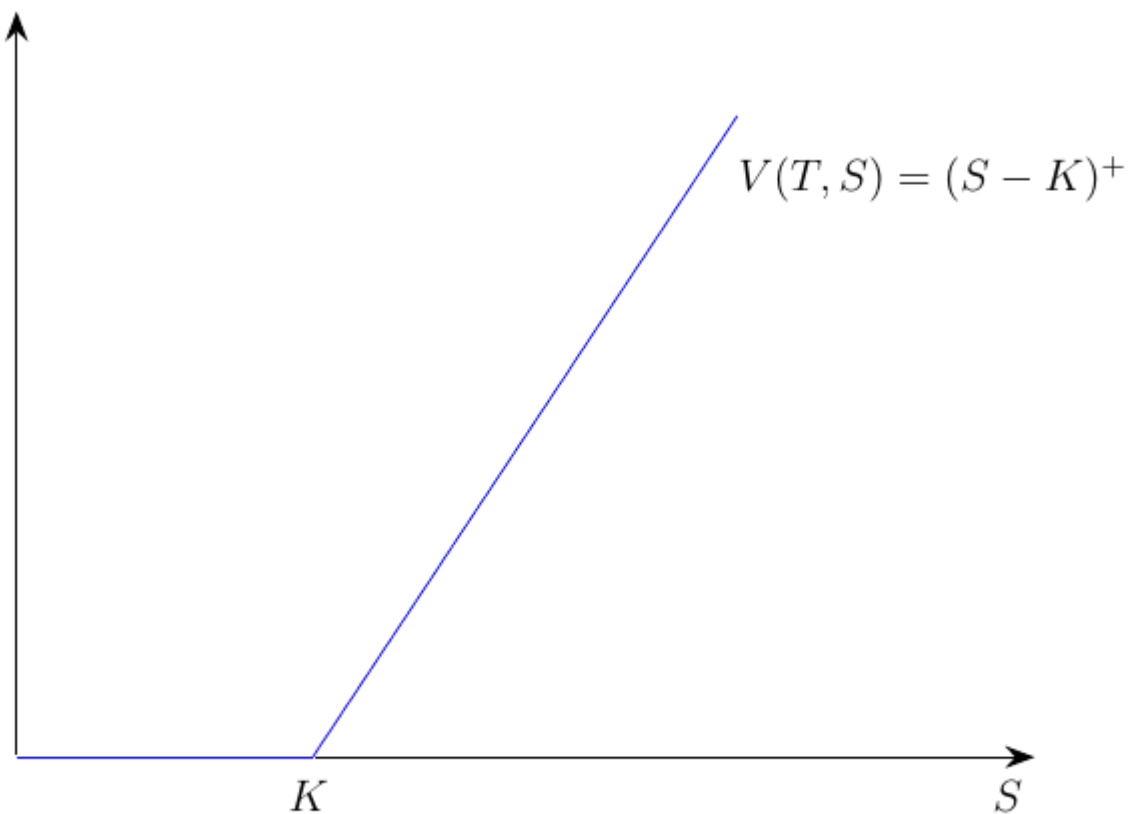


Abbildung 1.5: Europäischer Call ($t = T$).

1.3.5 Optionspreisberechnung

Eine Option ist das Recht auf Kauf (Call-Option) bzw. Verkauf (Put-Option) eines Basiswertes (Waren, Aktien etc.) zur Verfallszeit T (europäisch) bzw. bis zur Verfallszeit T (amerikanisch) zu einem bestimmten Preis (Ausübungspreis).

Der Optionspreis ist der Wert, dieses Recht wahrnehmen zu dürfen.

Notationen:

1.3 Modellierung von Prozessen mit Rauschen

- Kurs des Basiswertes zur Zeit t : $S(t)$
- Wert der Option zur Zeit t : $V(t, S(t))$
- Ausübungspreis: K
- Wert einer Call-Option in $t = T$: $V(T, S) = \max \{0, S - K\} =: (S - K)^+$
- Wert einer Put-Option in $t = T$: $V(T, S) = \max \{0, K - S\} = (K - S)^+$

Idee der Call-Option:

Man erhält den Basiswert zum Preis K und verkauft sofort zum Preis S , wenn $S > K$. Der Gewinn (Wert der Option) ist in diesem Fall $(S - K)^+$.

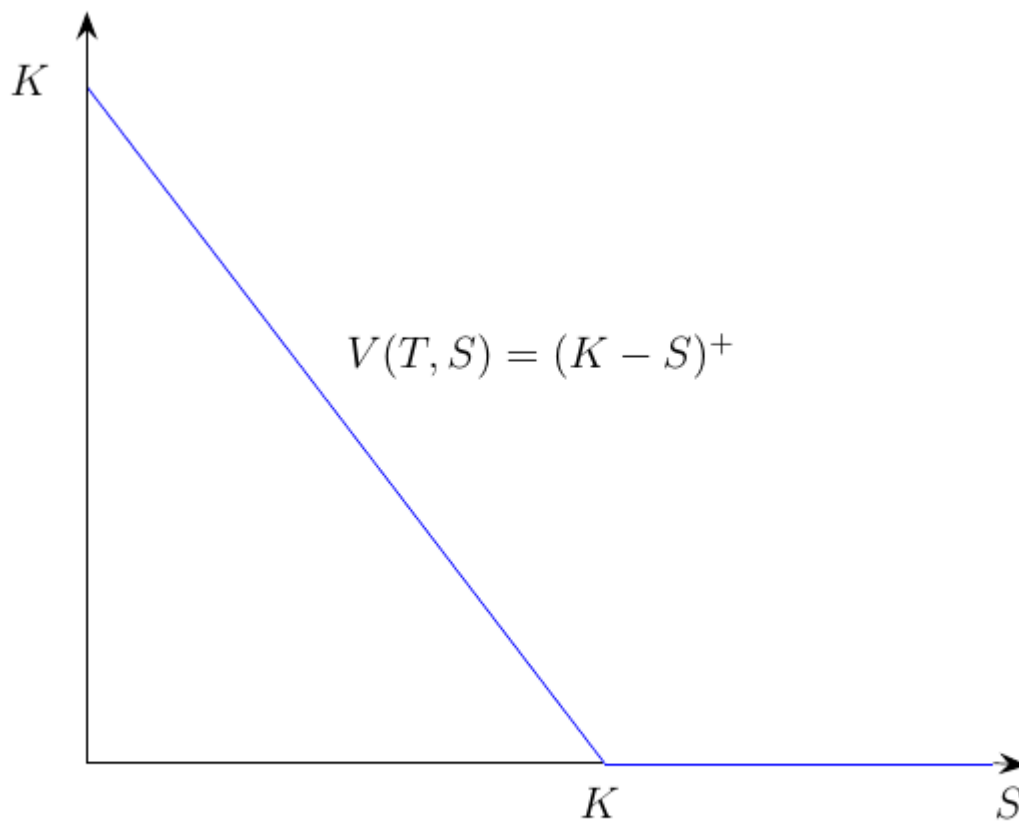


Abbildung 1.6: Europäischer Put ($t = T$).

Barrier-Option (Down and Out Call):

Dies ist eine Option, die wertlos (oder wertvoll) wird, wenn der Kurs des Basiswertes eine vorher festgelegte Schranke B vor dem Verfallstag über- bzw. unterschreitet.

Die zugehörige Auszahlungsfunktion (Call) lautet

$$V(T, S|_{[0, T]}) = \begin{cases} (S(T) - K)^+, & \text{falls } \min \{S(t) : 0 \leq t \leq T\} \geq B, \\ 0, & \text{sonst.} \end{cases}$$

1.3 Modellierung von Prozessen mit Rauschen

Gesucht ist nun der Wert der Option zum Verkaufszeitpunkt $V(0, S(0))$, wobei die Größe $S(0)$ bekannt ist.

Zur Berechnung von Optionspreisen führen wir ein einfaches Marktmodell ein mit folgenden Annahmen:

- risikoloser Zinssatz $r(t) \geq 0$ für Geldanlagen und Kredite
- der Basiswert entwickle sich gemäß

$$\begin{aligned} d\bar{S}(t) &= \mu(t)\bar{S}(t)dt + \sigma(t)\bar{S}(t)dW_t, \\ \bar{S}(0) &= S_0 \end{aligned} \tag{1.22}$$

- kontinuierlicher Handel ohne Transaktionskosten

Damit lässt sich für eine europäische Option zeigen:

$$V(0, S_0) = E \left(\exp \left(\int_0^T -r(s)ds \right) \cdot V(T, S(T)) \right),$$

wobei S die SDGL (1.22) löst. Für konstantes r erhalten wir

$$V(0, S_0) = \exp(-rT) E(V(T, S(T))).$$

Man sagt: „Der Optionspreis zur Zeit $t = 0$ ist gegeben durch den abdiskontierten Erwartungswert des Optionspreises zur Zeit $t = T$.“

1.3.6 Monte-Carlo-Simulation

Vorgelegt sei die SDGL

$$dS_t = \mu S_t dt + \sigma S_t dW_t, \quad S_0 > 0.$$

Gesucht ist der Prozess $(S_t)_{0 \leq t \leq T}$.

Naheliegende Diskretisierung:

- Ersetze $[0, T]$ durch das Gitter $\{t_0, \dots, t_N\}$, $t_j = j\Delta t$, $\Delta t = \frac{T}{N}$, $j = 0, \dots, N$.

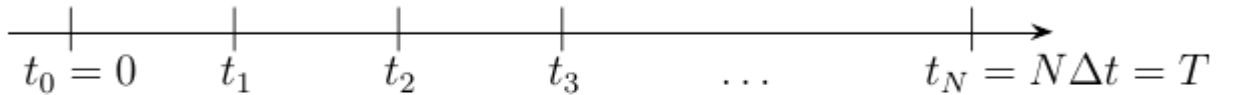


Abbildung 1.7: Numerisches Gitter $\{t_0, \dots, t_N\}$.

- Ersetze den kontinuierlichen Prozess $S : [0, T] \times \Omega \rightarrow \mathbb{R}$ durch den diskreten Prozess $\hat{S} : \{t_0, \dots, t_N\} \times \Omega \rightarrow \mathbb{R}$.

1.4 Pseudozufallszahlen

- Ersetze die SDGL für S durch die stochastische Differenzengleichung

$$\hat{S}(t_{n+1}) - \hat{S}(t_n) = \mu \hat{S}(t_n) \Delta t + \sigma \hat{S}(t_n) \Delta W_{t_n},$$

wobei $\Delta W_{t_n} := W(t_{n+1}) - W(t_n)$ und ΔW_{t_n} unabhängige, $\mathcal{N}(0, \Delta t)$ -verteilte ZV sind.

Diese Ersetzungen führen auf den „**Euler-Maruyama-Algorithmus**“

$$\begin{aligned} \hat{S}(t_{n+1}) &= \hat{S}(t_n) + \mu \Delta t \hat{S}(t_n) + \sigma \hat{S}(t_n) \hat{\xi}(t_n), \quad n = 0, \dots, N-1, \\ \hat{S}(0) &= S_0. \end{aligned}$$

Dabei sind die $\hat{\xi}(t_n)$ i. i. d. und $\mathcal{N}(0, \Delta t)$ -verteilt. Dieser Algorithmus ist eine Iteration von ZV!

Zur Berechnung eines approximativen Pfades $\hat{S}(t_j, \omega)$, $j = 0, \dots, N$, $\omega \in \Omega$ durch

$$\begin{aligned} \hat{S}(t_{n+1}, \omega) &= \hat{S}(t_n, \omega) + \mu \Delta t \hat{S}(t_n, \omega) + \sigma \hat{S}(t_n, \omega) \hat{\xi}(t_n, \omega), \quad n = 0, \dots, N-1, \\ \hat{S}(0, \omega) &= S_0(\omega) \end{aligned}$$

benötigt man also die Zahlen $\hat{\xi}(t_i, \omega)$, $i = 0, \dots, N-1$.

Fragen:

- Wie erhält man diese Zahlen?
- Wie gut ist der Euler-Maruyama-Algorithmus?
- Gibt es andere Verfahren zur approximativen Lösung einer SDGL?

1.4 Pseudozufallszahlen

1.4.1 Einführung

Sei $X : \Omega \rightarrow \mathbb{R}^d$ eine vektorwertige ZV auf dem Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$ und sei $(X_i)_{i \in \mathbb{N}}$ eine Folge unabhängiger, identisch wie X verteilter ZV. Für jedes $\omega \in \Omega$ heißt dann $(X_i(\omega))_{i \in \mathbb{N}}$ eine Folge von Zufallszahlen ($d = 1$) bzw. eine Folge von Zufallsvektoren ($d > 1$).

Deterministisch, d. h. durch einen Algorithmus berechenbare Folgen $(Y_i)_{i \in \mathbb{N}}$, die sich näherungsweise wie Zufallszahlen bzw. Zufallsvektoren verhalten, heißen Pseudozufallszahlen bzw. Pseudozufallsvektoren.

1.4.2 Beispiel: Monte-Carlo-Integration

Sei $X : \Omega \rightarrow \mathbb{R}^d$ eine vektorwertige ZV und sei $g : \mathbb{R}^d \rightarrow \mathbb{R}$ messbar. Dann gilt

$$E(g(X)) = \int_{\Omega} g(X) dP = \int_{\mathbb{R}^d} g(v) f(v) dv.$$

Dabei gilt das zweite Gleichheitszeichen, falls X die Dichte f bezüglich des Lebesgue-Maßes besitzt.

1.4 Pseudozufallszahlen

Wir nehmen nun an, zur Verfügung stehen vektorwertige ZV $X_1, X_2, \dots, X_N$, welche unabhängig und identisch wie X verteilt sind. Betrachtet man den Mittelwert

$$S_N = \frac{1}{N} \sum_{i=1}^N g(X_i),$$

so gilt:

$$E(S_N) = \frac{1}{N} \sum_{i=1}^N E(g(X_i)) = \frac{1}{N} \sum_{i=1}^N E(g(X)) = E(g(X)).$$

Das zweite Gleichheitszeichen gilt, da g messbar ist und die X_i die gleiche Verteilung haben wie X . Unter den Annahmen $V(g(X)) < \infty$, X_i , $i = 1, \dots, N$ unabhängig gilt überdies

$$V(S_N) = V\left(\frac{1}{N} \sum_{i=1}^N g(X_i)\right) \stackrel{(*)}{=} \frac{1}{N^2} \sum_{i=1}^N V(g(X_i)) = \frac{1}{N^2} \sum_{i=1}^N V(g(X)) = \frac{1}{N} V(g(X)),$$

wobei $(*)$ gilt, da die $g(X_i)$ unabhängig sind (da X_i unabhängig und g messbar). Das darauf folgende Gleichheitszeichen gilt wiederum wegen der identischen Verteilung von X_i wie X und der Messbarkeit von g . Damit folgt

$$V(S_N) = \frac{1}{N} V(g(X)) \rightarrow 0 \text{ für } N \rightarrow \infty.$$

S_N variiert also immer weniger um $E(S_N) = E(g(X))$ für $N \rightarrow \infty$.

Das starke Gesetz der großen Zahlen liefert

$$\lim_{N \rightarrow \infty} S_N = E(g(X)) \quad P\text{-fast sicher,}$$

d. h. für fast jedes $\omega \in \Omega$ gilt

$$\lim_{N \rightarrow \infty} S_N(\omega) = E(g(X)).$$

Für „gute“ Pseudozufallsvektoren $Y_1, Y_2, \dots, Y_N$ gilt also für große N

$$\frac{1}{N} \sum_{i=1}^N g(Y_i) \approx E(g(X)) = \int_{\mathbb{R}^d} g(v) f(v) dv \quad (1.23)$$

Die $Y_1, Y_2, \dots$ übernehmen also die Rolle von $X_1(\omega), X_2(\omega), \dots$. Gleichung (1.23) liefert die Idee der Monte-Carlo-Integration; diese ist besonders interessant für d groß.

Wichtiger Spezialfall:

Sei $d = 1$ und seien die Pseudozufallszahlen $Y_1, Y_2, \dots \in \mathbb{R}$ verteilt wie die ZV X . Für $g(v) := \mathbb{I}_{]-\infty, z]}(v)$ ($z \in \mathbb{R}$ fest) folgt

$$\frac{1}{N} \sum_{i=1}^N g(Y_i) = \frac{1}{N} \sum_{i=1}^N \mathbb{I}_{]-\infty, z]}(Y_i) \stackrel{N \text{ groß}}{\approx} \int_{\Omega} \mathbb{I}_{]-\infty, z]}(X) dP = \int_{\{X \leq z\}} dP = P(X \leq z) = F(z),$$

wobei F die Verteilungsfunktion von X ist.

1.20 Bemerkung

Es gilt $\mathbb{I}_{]-\infty, z]}(Y_i) = 1$ genau dann, wenn $Y_i \leq z$. Also gibt der Ausdruck $\frac{1}{N} \sum_{i=1}^N \mathbb{I}_{]-\infty, z]}(Y_i)$ die relative Häufigkeit der $Y_i \leq z$ an und

$$F_N(z) := \frac{1}{N} \sum_{i=1}^N \mathbb{I}_{]-\infty, z]}(Y_i)$$

ist für N hinreichend groß eine gute Approximation der Verteilungsfunktion von X an der Stelle z .

1.4.3 Konstruktion von Pseudozufallszahlen

Nun soll es um die Konstruktion von Pseudozufallszahlen (PZZ) gehen bei vorgegebener Verteilung. Wichtig dafür ist folgende zentrale Beobachtung.

1.21 Satz

Sei $F : \mathbb{R} \rightarrow [0, 1]$ streng monoton wachsend mit $\lim_{x \rightarrow -\infty} F(x) = 0$, $\lim_{x \rightarrow \infty} F(x) = 1$ und seien $X_1, X_2, \dots$ unabhängig und identisch, gleichverteilte ZV auf $[0, 1]$. Dann sind

$$F^{-1}(X_1), F^{-1}(X_2), F^{-1}(X_3), \dots$$

unabhängig, identisch verteilt mit Verteilungsfunktion F .

Beweis. Da X gleichverteilt auf $[0, 1]$, hat X die Dichtefunktion $f(x) = \frac{1}{1-0} = 1$ für $x \in [0, 1]$ und $f(x) = 0$, sonst. Damit folgt

$$P(F^{-1}(X_i) \leq z) = P(X_i \leq F(z)) = \int_0^{F(z)} 1 dx = F(z).$$

□

1.22 Bemerkung

- Sind also $X_1, X_2, \dots$ gute gleichverteilte PZZ auf $]0, 1[$, so sind $F^{-1}(X_1), F^{-1}(X_2), F^{-1}(X_3), \dots$ gute PZZ mit Verteilungsfunktion F . Numerisch ist dies allerdings nur brauchbar, wenn F^{-1} direkt verfügbar ist und nicht erst durch Lösen einer nichtlinearen Gleichung $F(x) = y$ berechnet werden muss.
- Gute gleichverteilte PZZ auf $]0, 1[$ bekommt man in Matlab mit dem Befehl **rand** (genauer: **rand(n,m)** erzeugt eine $n \times m$ -Matrix mit gleichverteilten PZZ auf $]0, 1[$).

1.23 Beispiel (Standardnormalverteilung)

$$\begin{aligned} F(z) &= \int_{-\infty}^z \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx = \frac{1}{2} + \int_0^z \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx \\ &= \frac{1}{2} \left(1 + \sqrt{\frac{2}{\pi}} \int_0^z \exp\left(-\frac{x^2}{2}\right) dx \right) \stackrel{s:=\frac{x}{\sqrt{2}}}{=} \frac{1}{2} \left(1 + \frac{2}{\sqrt{\pi}} \int_0^{z/\sqrt{2}} \exp(-s^2) ds \right) \\ &= \frac{1}{2} \left(1 + \operatorname{erf}\left(\frac{z}{\sqrt{2}}\right) \right), \end{aligned}$$

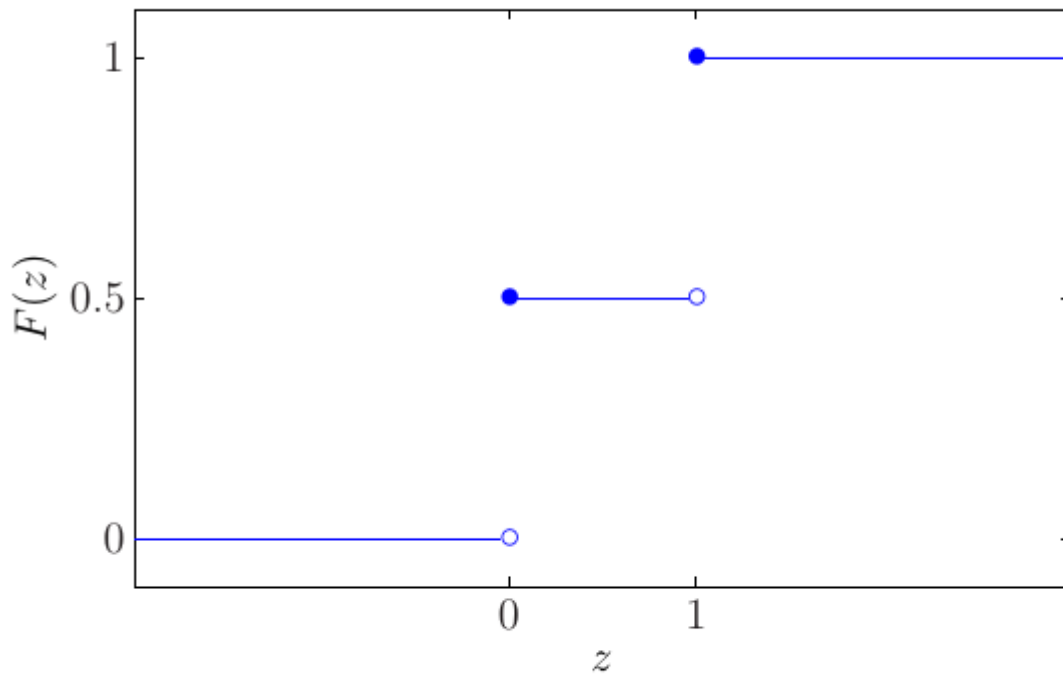


Abbildung 1.8: Diskrete Verteilung.

wobei $\operatorname{erf}\left(\frac{z}{\sqrt{2}}\right) := \frac{2}{\sqrt{\pi}} \int_0^{z/\sqrt{2}} \exp(-s^2) ds$ und erf das Gauß'sche Fehlerintegral ist. Damit folgt mit $z := F^{-1}(y)$

$$y = F(F^{-1}(y)) = \frac{1}{2} \left(1 + \operatorname{erf}\left(\frac{F^{-1}(y)}{\sqrt{2}}\right) \right),$$

das bedeutet

$$2y - 1 = \operatorname{erf}\left(\frac{F^{-1}(y)}{\sqrt{2}}\right) \quad \text{bzw.} \\ F^{-1}(y) = \sqrt{2} \operatorname{erf}^{-1}(2y - 1).$$

1.24 Bemerkung

Noch nicht erfasst sind monotone, aber nicht streng monotone Verteilungen, vgl. Abbildungen 1.8 und 1.9. Diese Verteilungsfunktionen

$$F(z) = P(X \leq z) = \int_{-\infty}^z f(s) ds$$

sind rechtsstetig, d.h. $\lim_{h \rightarrow 0, h > 0} F(x + h) = F(x)$.

Etwas allgemeiner ist der folgende

1.25 Satz

Sei $F : \mathbb{R} \rightarrow [0, 1]$ eine Verteilungsfunktion, d. h.

1. F ist monoton wachsend,

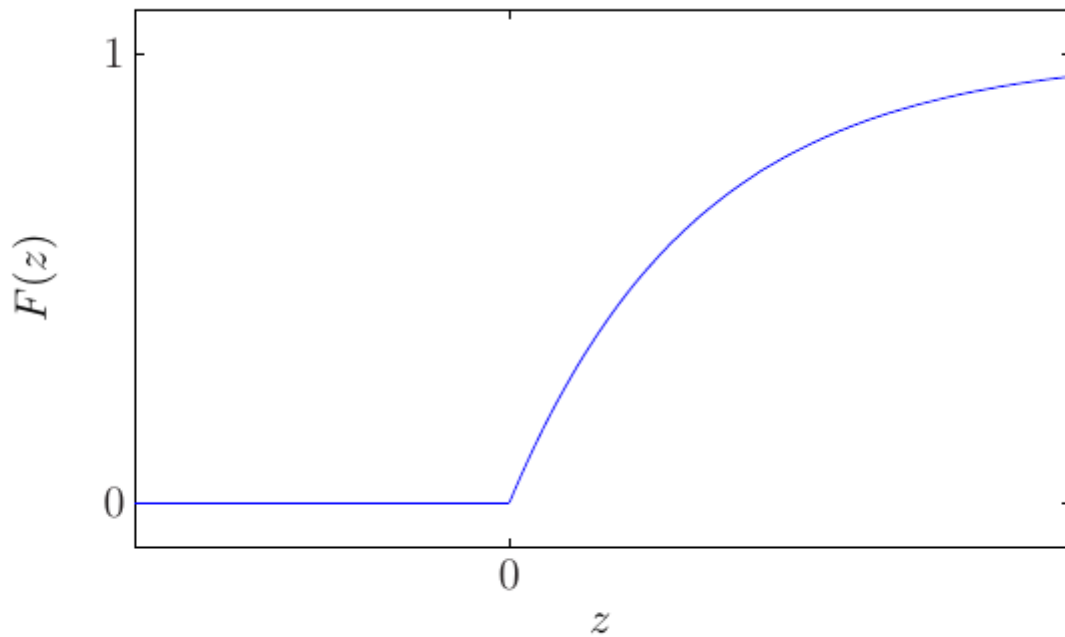


Abbildung 1.9: Exponentielle Verteilung.

2. F ist rechtsstetig,

3. $\lim_{x \rightarrow -\infty} F(x) = 0$ und $\lim_{x \rightarrow \infty} F(x) = 1$.

Weiter sei $G(y) := \inf \{x \in \mathbb{R} : F(x) \geq y\}$, $0 < y < 1$. Dann gilt

$$G(y) \leq z \iff y \leq F(z).$$

Beweis. „ $\Leftarrow$ “: Sei $y \leq F(z)$, dann folgt $z \in \{x \in \mathbb{R} : F(x) \geq y\}$ und somit

$$z \geq \inf \{x \in \mathbb{R} : F(x) \geq y\} = G(y).$$

„ $\Rightarrow$ “: Sei nun $G(y) = \inf \{x \in \mathbb{R} : F(x) \geq y\} \leq z$. Dann existiert zu $\varepsilon > 0$ hinreichend klein ein x mit $F(x) \geq y$ und $x \leq G(y) + \varepsilon$. Insbesondere haben wir nach Voraussetzung $x \leq G(y) + \varepsilon \leq z + \varepsilon$. Die Monotonie von F liefert dann

$$y \leq F(x) \leq F(z + \varepsilon)$$

und mit der Rechtsstetigkeit von F und $\varepsilon \rightarrow 0$ folgt

$$y \leq F(z).$$

□

Konsequenz:

Seien $X_1, X_2, \dots$ i.i.d. und gleichverteilt auf $]0, 1[$. Dann gilt mit Satz 1.25

$$P(G(X_i) \leq z) = P(X_i \leq F(z)) = F(z).$$

Das letzte Gleichheitszeichen gilt aufgrund der Gleichverteilung von X_i auf $]0, 1[$. Sind also $X_1, X_2, \dots$ i.i.d. und gleichverteilt auf $]0, 1[$, so sind $G(X_1), G(X_2), \dots$ i.i.d. mit Verteilungsfunktion F .

1.26 Beispiel (Pseudo-Münze)

Sei $X(\omega) \in \{0, 1\}$, $P(X = 0) = P(X = 1) = \frac{1}{2}$. Für die Verteilungsfunktion F gilt (vergleiche Abbildung 1.8)

$$F(z) = \begin{cases} 0, & \text{für } z < 0, \\ \frac{1}{2}, & \text{für } 0 \leq z < 1, \\ 1, & \text{für } 1 \leq z. \end{cases}$$

Dann erhalten wir für G

$$G(y) = \inf \{x \in \mathbb{R} : F(x) \geq y\} = \begin{cases} 0, & \text{für } y \leq \frac{1}{2}, \\ 1, & \text{für } y > \frac{1}{2}, \end{cases} \quad 0 < y < 1.$$

1.27 Beispiel (Exponentialverteilung)

Die Verteilungsfunktion der Exponentialverteilung ist für $\mu > 0$ gegeben durch

$$F(z) = \int_{-\infty}^z f(s)ds \quad \text{mit} \quad f(s) = \begin{cases} \frac{1}{\mu} \exp(-s/\mu), & \text{für } s \geq 0, \\ 0, & \text{für } s < 0. \end{cases}$$

Also folgt

$$F(z) = \begin{cases} 1 - \exp(-z/\mu), & \text{für } z \geq 0, \\ 0, & \text{für } z < 0 \end{cases}$$

und $F^{-1}(z) = -\mu \ln(1 - z)$ für $z \geq 0$. Damit erhalten wir für G

$$G(y) = -\mu \ln(1 - y), \quad 0 < y < 1.$$

Gleichverteilte PZZ sind also der Ausgangspunkt für anders verteilte PZZ.

1.4.4 Konstruktion von Pseudozufallsvektoren (PZV)

Seien $X_1, X_2, X_3, \dots$ unabhängig und identisch verteilte ZV wie $X : \Omega \rightarrow \mathbb{R}$. Setze dann

$$Y_1 := \begin{pmatrix} X_1 \\ \vdots \\ X_m \end{pmatrix}, Y_2 := \begin{pmatrix} X_{m+1} \\ \vdots \\ X_{2m} \end{pmatrix}, \dots$$

$Y_1, Y_2, \dots$ sind i. i. d. mit Verteilung $P_{Y_i} = P_X \otimes \dots \otimes P_X =: P_Y$, d. h. $Y_1(\omega), Y_2(\omega), \dots$ sind Zufallsvektoren mit Verteilung P_Y . Entsprechend sind $Y_1, Y_2, \dots$ gute PZV, falls $X_1, X_2, \dots$ gute PZZ sind.

Allgemeinere PZV findet man mit

1.28 Satz

Sei $X : \Omega \rightarrow \mathbb{R}^m$ eine vektorwertige ZV mit Dichte $f : \mathbb{R}^m \rightarrow \mathbb{R}$ und seien $A := \{x \in \mathbb{R}^m : f(x) > 0\}$ und $h : A \rightarrow h(A) \subset \mathbb{R}^m$ umkehrbar und h^{-1} differenzierbar. Dann hat die transformierte Zufallsabbildung $Y := h(X)$ die Dichte

$$\hat{f}(y) := f(h^{-1}(y)) \cdot \left| \det \left(\frac{\partial h^{-1}}{\partial y}(y) \right) \right|, \quad y \in h(A).$$

1.4 Pseudozufallszahlen

Beweisidee. Für $M \subset \mathbb{R}^m$ gilt

$$\begin{aligned} P(Y \in M) &= P(h(X) \in M) = P(X \in h^{-1}(M)) = \int_{h^{-1}(M)} f(x) dx \\ &= \int_M f(h^{-1}(y)) \left| \det \left(\frac{\partial h^{-1}}{\partial y}(y) \right) \right| dy. \end{aligned}$$

Beim letzten Gleichheitszeichen wurde der Transformationssatz im $\mathbb{R}^m$ angewendet. $\square$

1.29 Beispiel

Typische Anwendung: $f(x) = \mathbb{I}_{[0,1]^m}(x)$, d. h. X ist gleichverteilt auf $]0, 1[^m$ (die Komponenten sind unabhängig und gleichverteilt). Es gilt also: $f = 1$ auf $A := [0, 1[^m$. Die transformierte Abbildung $Y = h(X)$ ist dann verteilt mit Dichte

$$\hat{f}(y) = \left| \det \left(\frac{\partial h^{-1}}{\partial y}(y) \right) \right|, \quad y \in h(A) = h([0, 1[^m).$$

1.4.5 Die Box-Muller-Methode

Betrachte die Abbildung $h :]0, 1[^2 \rightarrow \mathbb{R}^2$, $h(x_1, x_2) = \sqrt{-2 \ln(x_1)} \begin{pmatrix} \cos(2\pi x_2) \\ \sin(2\pi x_2) \end{pmatrix}$. Die Umkehrabbildung von h hat die Gestalt

$$h^{-1}(y) = \begin{pmatrix} \exp(-(y_1^2 + y_2^2)/2) \\ \arctan\left(\frac{y_2}{y_1}\right)/2\pi \end{pmatrix}.$$

Somit folgt

$$\left| \det \left(\frac{\partial h^{-1}}{\partial y}(y) \right) \right| = \frac{1}{2\pi} \exp(-(y_1^2 + y_2^2)/2).$$

Ist also X gleichverteilt auf $]0, 1[^2$, so ist $h(X)$ nach Satz 1.28 standardnormalverteilt auf $\mathbb{R}^2$. Beachte dabei, dass für die Dichte f von X gilt: $f = \mathbb{I}_{[0,1]^2}$.

Der Box-Muller-Algorithmus:

- Generiere $]0, 1[$ -gleichverteilte PZZ $X_1, X_2, X_3, \dots$
- Dann sind $h_1(X_1, X_2), h_2(X_1, X_2), h_1(X_3, X_4), h_2(X_3, X_4), \dots$ $\mathcal{N}(0, 1)$ -verteilte PZZ.

1.30 Bemerkung

Der Befehl **randn** erzeugt $\mathcal{N}(0, 1)$ -verteilte PZZ in Matlab; genauer: **randn(n, m)** erzeugt eine $n \times m$ -Matrix von $\mathcal{N}(0, 1)$ -verteilten PZZ. Dieser Befehl benutzt den sogenannten Ziggurat-Algorithmus, welcher ohne die Funktionen $\sqrt{\cdot}$, $\ln(\cdot)$, $\cos(\cdot)$ und $\sin(\cdot)$ auskommt.

1.4.6 Korrelierte Normalverteilung (korreliert normalverteilte PZV)

Sei X m -dimensional und $\mathcal{N}(0, I_m)$ -verteilt, d. h. X hat die Dichtefunktion

$$f(x) = \frac{1}{(2\pi)^{m/2}} \exp\left(-\frac{1}{2} \langle x, x \rangle\right).$$

Derartige PZV erhält man z. B. mit der Box-Muller-Methode.

Es sei $h(X) := \mu + AX$ mit $A \in \mathbb{R}^{m \times m}$ invertierbar. Wir finden $h^{-1}(Y) = A^{-1}(Y - \mu)$ und damit $\frac{\partial h^{-1}}{\partial Y} = A^{-1}$. Also hat $Y = h(X)$ nach Satz 1.28 die Dichte

$$\begin{aligned} \hat{f}(y) &= \frac{1}{(2\pi)^{m/2}} \exp\left(-\frac{1}{2} \langle A^{-1}(y - \mu), A^{-1}(y - \mu) \rangle\right) \cdot |\det(A^{-1})| \\ &= \frac{1}{(2\pi)^{m/2} |\det(A)|} \exp\left(-\frac{1}{2} \langle (AA^T)^{-1}(y - \mu), y - \mu \rangle\right). \end{aligned}$$

Mit $\Sigma := AA^T$ folgt $\det(\Sigma) = \det(A)^2$ und somit

$$\hat{f}(y) = \frac{1}{(2\pi)^{m/2} \sqrt{\det(\Sigma)}} \exp\left(-\frac{1}{2} \langle \Sigma^{-1}(y - \mu), y - \mu \rangle\right),$$

d. h. Y ist $\mathcal{N}(\mu, \Sigma)$ -verteilt.

Konkretes Vorgehen:

Sei $\mu \in \mathbb{R}^m$ und $\Sigma \in \mathbb{R}^{m \times m}$ positiv definit gegeben.

1. Bestimme $A \in \mathbb{R}^{m \times m}$ mit $\Sigma = AA^T$.
2. Konstruiere $\mathcal{N}(0, I_m)$ -verteilte PZV $X_1, X_2, \dots$ (z. B. mit der Box-Muller-Methode).
3. Bilde $Y_i = \mu + AX_i$, $i = 1, 2, \dots$

Dann sind $Y_1, Y_2, \dots$ $\mathcal{N}(\mu, \Sigma)$ -verteilte PZV.

1.4.7 Die Cholesky-Zerlegung

Um aus unkorrelierten ZV $\mathcal{N}(\mu, \Sigma)$ -verteilte ZV zu erhalten, benötigt man die Zerlegung $\Sigma = AA^T$.

1.31 Satz

Sei $C \in \mathbb{R}^{N \times N}$ symmetrisch und positiv definit. Dann existiert eine eindeutig bestimmte linke untere Dreiecksmatrix L mit strikt positiven Diagonalelementen, so dass $C = L \cdot L^T$.

Beweis. Eindeutigkeit: Seien L_1 und L_2 linke untere Dreiecksmatrizen mit positiver Hauptdiagonale, und es gelte

$$L_1 L_1^T = C = L_2 L_2^T.$$

Da linke untere (rechte obere) Dreiecksmatrizen eine Algebra mit Einheit I bilden, sind sie strukturerhaltend unter Addition, Multiplikation und Invertierung. Also folgt

$$L_2^{-1} L_1 = L_2^T (L_1^T)^{-1} = \left[(L_2^{-1} L_1)^T \right]^{-1},$$

1.4 Pseudozufallszahlen

wobei $L_2^{-1} L_1$ eine linke untere und $\left[(L_2^{-1} L_1)^T \right]^{-1}$ eine rechte obere Dreiecksmatrix ist. D.h. $H := L_2^{-1} L_1$ erfüllt die Gleichung $H = (H^T)^{-1}$ und ist gleichzeitig linke untere und rechte obere Dreiecksmatrix. Also ist H diagonal. Damit folgt

$$H^2 = HH^T = HH^{-1} = I \Rightarrow H = I,$$

da L_1 und L_2 strikt positive Diagonalelemente haben. Somit gilt $L_1 = L_2$.

Die Existenz wird durch das Cholesky-Verfahren gesichert (vgl. Numerik I). $\square$

In Matlab erfolgt die Berechnung von $C = L \cdot L^T$ für symmetrisches und positiv definites C mit folgenden Programmzeilen:

```
% Cholesky-Zerlegung  
R=chol(C);  
L=R';
```

1.4.8 Konstruktion von PZZ mit Gleichverteilung auf $[0, 1]$

Sind gleichverteilte PZZ erzeugt, so können wir mit den bereits beschriebenen Methoden PZZ mit anderer Verteilung generieren.

Lineare Kongruenzmethode:

Erzeuge Zahlen $z_1, \dots, z_N$ gemäß

$$z_{i+1} = (az_i + c) \mod m, \quad i = 0, \dots, N-1,$$

mit $z_0 \in \{0, 1, \dots, m-1\}$ („seed“). Setze dann

$$x_i := \frac{z_i}{m} \in [0, 1[$$

und erhalte so PZZ.

1.32 Bemerkung

a) Der Ausdruck $x \mod y$ entspricht dem Rest bei Division von x mit y .

b) Illustration für $a = c = 3$ und $m = 5$:

$$\begin{aligned} z_0 &= 3, \\ z_1 &= (3 \cdot 3 + 3) \mod 5 = 2, \\ z_2 &= (3 \cdot 2 + 3) \mod 5 = 4, \\ z_3 &= (3 \cdot 4 + 3) \mod 5 = 0, \\ z_4 &= (3 \cdot 0 + 3) \mod 5 = 3 = z_0. \end{aligned}$$

Dieser Generator generiert eine 4-periodische Zahlenfolge. Die Frage ist nun, wann eine solche Rekursion gute PZZ liefert. Offensichtlich müssen a , c und m sorgfältig gewählt werden. Insbesondere muss z. B. m groß sein!

1.33 Bemerkung

Jeder solche Algorithmus ist periodisch. Wähle also a , c und m derart, dass die Periode möglichst groß wird. Aus der Zahlentheorie ist folgender Sachverhalt bekannt:

Keine Schleifenlänge $< m$, falls

1. c und m keine gemeinsamen Teiler haben (z. B. $m = 2^{31}$ und c ungerade, bei 32-bit-Rechnern).
2. $a \bmod g = 1$ für jeden Primfaktor g von m (z. B. $m = 2^{31}$ und $g = 2$, d. h. a hat die Form $a = 2^r + 1$).
3. $a \bmod 4 = 1$, wenn m ein Vielfaches von 4 ist (z. B. $m = 2^{31}$ und $a = 2^r + 1$, $r \geq 2 \Rightarrow a = 2^2 2^{r-2} + 1$).

Den Beweis dieser Aussage findet man in Knuth, *The Art of Computer Programming* (1969). Anzufügen bleibt, dass eine möglichst große Periode nur ein notwendiges, aber kein hinreichendes Kriterium für gute auf $[0, 1[$ gleichverteilte PZZ ist.

1.34 Bemerkung

Der Zufallsgenerator in Matlab basiert auf dem „lagged“ Fibonacci-Generator (Tausworth, 1965)

$$z_i = (z_{i-\mu} + z_{i-\gamma}) \bmod m,$$

$$x_i = \frac{z_i}{m}, \quad i = 0, 1, 2, \dots$$

Es müssen also $\max\{\mu, \gamma\}$ viele Zahlen aus $\{0, \dots, m-1\}$ vorgegeben werden. In Matlab ist $\max\{\mu, \gamma\} = 32$, d. h. der Zustand des Generators besteht aus 32 Zahlen und ist erhältlich mit dem Befehl **rand('state')** und zu setzen mit **rand('state', myseed)**.

2 Numerische Methoden für stochastische Differentialgleichungen

2.1 Das Euler-Maruyama-Verfahren

2.1.1 Numerische Methoden

Betrachte die SDGL

$$\begin{aligned} dX_t &= a(t, X_t)dt + b(t, X_t)dW_t, \quad 0 \leq t \leq T, \\ X_0 &= \xi. \end{aligned} \tag{2.1}$$

Numerische Methoden zur Lösung von (2.1):

Sei $N \in \mathbb{N}$ gegeben und für $h := \frac{T}{N}$ und $t_n := nh$, $n = 0, 1, \dots, N$ sei $\hat{X}_{t_n}$ die Approximation für $X_{t_n} = X_{nh}$. Wir betrachten nun die Diskretisierung von (2.1):

$$\begin{aligned} \hat{X}_{t_{n+1}} &= \hat{X}_{t_n} + a(t_n, \hat{X}_{t_n})h + b(t_n, \hat{X}_{t_n})\Delta W_{t_n}, \\ \hat{X}_0 &= \xi, \end{aligned}$$

mit $\Delta W_{t_n} := W_{t_{n+1}} - W_{t_n} \sim \mathcal{N}(0, h)$.

2.1.2 Güte der Approximation

Ziel ist es nun, die Güte der approximativen Lösung zu beurteilen. Dazu untersuchen wir Teilstücke der exakten Lösung und wählen zur technischen Vereinfachung das autonome Problem

$$\begin{aligned} dX_t &= a(X_t)dt + b(X_t)dW_t, \\ X_0 &= \xi. \end{aligned} \tag{2.2}$$

Gleichung (2.2) gilt für jedes Teilintervall von $[0, T]$, also auch für $[t_n, t_{n+1}]$, d. h. wir haben

$$X_{t_{n+1}} = X_{t_n} + \int_{t_n}^{t_{n+1}} a(X_t)dt + \int_{t_n}^{t_{n+1}} b(X_t)dW_t.$$

Mit der Itô-Formel gilt für $a, b \in C^2$

$$\begin{aligned} da(X_t) &= a'(X_t)dX_t + \frac{1}{2}a''(X_t)b^2(X_t)dt, \\ db(X_t) &= b'(X_t)dX_t + \frac{1}{2}b''(X_t)b^2(X_t)dt. \end{aligned}$$

2.1 Das Euler-Maruyama-Verfahren

Mit $dX_t = a(X_t)dt + b(X_t)dW_t$ folgt somit durch Integration über $[t_n, t_{n+1}]$

$$a(X_t) = \underbrace{a(X_{t_n})}_{= \text{konstant}} + \int_{t_n}^t a'(X_s)a(X_s) + \frac{1}{2}a''(X_s)b^2(X_s)ds + \int_{t_n}^t a'(X_s)b(X_s)dW_s, \quad (2.3)$$

$$b(X_t) = \underbrace{b(X_{t_n})}_{= \text{konstant}} + \int_{t_n}^t b'(X_s)a(X_s) + \frac{1}{2}b''(X_s)b^2(X_s)ds + \int_{t_n}^t b'(X_s)b(X_s)dW_s. \quad (2.4)$$

Einsetzen von (2.3) und (2.4) in (2.2) liefert die Struktur

$$\begin{aligned} X_{t_{n+1}} &= X_{t_n} + \int_{t_n}^{t_{n+1}} a(X_{t_n})dt + \int_{t_n}^{t_{n+1}} b(X_{t_n})dW_t + \text{Rest}_{t_n} \\ &= X_{t_n} + a(X_{t_n})h + b(X_{t_n})\Delta W_{t_n} + \text{Rest}_{t_n}, \end{aligned} \quad (2.5)$$

mit

$$\begin{aligned} \text{Rest}_{t_n} &:= \int_{t_n}^{t_{n+1}} \left[\int_{t_n}^t a'(X_s)a(X_s) + \frac{1}{2}a''(X_s)b^2(X_s)ds + \int_{t_n}^t a'(X_s)b(X_s)dW_s \right] dt \\ &\quad + \int_{t_n}^{t_{n+1}} \left[\int_{t_n}^t b'(X_s)a(X_s) + \frac{1}{2}b''(X_s)b^2(X_s)ds + \int_{t_n}^t b'(X_s)b(X_s)dW_s \right] dW_t. \end{aligned} \quad (2.6)$$

2.1 Bemerkung

- a) Rest_{t_n} besteht aus sechs Doppelintegralen!
- b) Vergleiche die Struktur von (2.5) für die SDGL (2.2) mit dem Euler-Maruyama-Verfahren.

2.1.3 Untersuchung der Doppelintegrale

Wir untersuchen im Folgenden die Doppelintegrale genauer. Betrachte nun den Term mit zweifacher stochastischer Integration, d. h.

$$\int_{t_n}^{t_{n+1}} \int_{t_n}^t \underbrace{b'(X_s)b(X_s)}_{=: f(X_s)} dW_s dW_t. \quad (2.7)$$

Für f gilt mit $dX_s = a(X_s)ds + b(X_s)dW_s$ nach der Itô-Formel

$$df(X_s) = f'(X_s)dX_s + \frac{1}{2}f''(X_s)b^2(X_s)ds$$

und somit

$$f(X_s) = f(X_{t_n}) + \int_{t_n}^s f'(X_\eta)a(X_\eta) + \frac{1}{2}f''(X_\eta)b^2(X_\eta)d\eta + \int_{t_n}^s f'(X_\eta)b(X_\eta)dW_\eta. \quad (2.8)$$

Einsetzen von (2.8) in (2.7) liefert

$$\int_{t_n}^{t_{n+1}} \int_{t_n}^t \underbrace{b'(X_s)b(X_s)}_{=: f(X_s)} dW_s dW_t = \underbrace{b'(X_{t_n})b(X_{t_n})}_{=: f(X_{t_n})} \int_{t_n}^{t_{n+1}} \int_{t_n}^t dW_s dW_t + \text{Rest}_{3,t_n}, \quad (2.9)$$

2.1 Das Euler-Maruyama-Verfahren

wobei Rest_{3,t_n} aus zwei Dreifachintegralen besteht.

Wir berechnen nun in einer Nebenrechnung folgendes Integral:

$$\begin{aligned} \int_{t_n}^{t_{n+1}} \underbrace{\left(\int_{t_n}^t dW_s \right)}_{=W_t - W_{t_n}} dW_t &= \int_{t_n}^{t_{n+1}} (W_t - W_{t_n}) dW_t = \int_{t_n}^{t_{n+1}} W_t dW_t - W_{t_n} \underbrace{\int_{t_n}^{t_{n+1}} dW_t}_{=\Delta W_{t_n}} \\ &= \int_{t_n}^{t_{n+1}} W_t dW_t - W_{t_n} \Delta W_{t_n}. \end{aligned}$$

Nach dem Lemma von Itô gilt weiter für $F(X) := X^2$ und $X_t := W_t$

$$dW_t^2 = 2W_t dW_t + \frac{1}{2} \cdot 2 \cdot 1^2 dt \iff \frac{1}{2} dW_t^2 = W_t dW_t + \frac{1}{2} dt.$$

Also folgt mit Integration über $[t_n, t_{n+1}]$

$$\frac{1}{2} W_{t_{n+1}}^2 = \frac{1}{2} W_{t_n}^2 + \int_{t_n}^{t_{n+1}} W_t dW_t + \underbrace{\int_{t_n}^{t_{n+1}} \frac{1}{2} dt}_{=h/2},$$

bzw. äquivalent dazu

$$\int_{t_n}^{t_{n+1}} W_t dW_t = \frac{1}{2} (W_{t_{n+1}}^2 - W_{t_n}^2) - \frac{h}{2}.$$

Wir finden somit

$$\begin{aligned} \int_{t_n}^{t_{n+1}} \int_{t_n}^t dW_s dW_t &= \frac{1}{2} (W_{t_{n+1}}^2 - W_{t_n}^2 - h) - W_{t_n} \Delta W_{t_n} \\ &= \frac{1}{2} ((W_{t_{n+1}} + W_{t_n}) \Delta W_{t_n} - h) - W_{t_n} \Delta W_{t_n} \\ &= \frac{1}{2} ((W_{t_{n+1}} - W_{t_n}) \Delta W_{t_n} - h) \\ &= \frac{1}{2} (\Delta W_{t_n}^2 - h). \end{aligned} \tag{2.10}$$

Einsetzen von (2.10) in (2.9) liefert

$$\int_{t_n}^{t_{n+1}} \int_{t_n}^t b'(X_s) b(X_s) dW_s dW_t = b'(X_{t_n}) b(X_{t_n}) \cdot \frac{1}{2} (\Delta W_{t_n}^2 - h) + \text{Rest}_{3,t_n}.$$

Daraus folgt mit (2.5) und (2.6)

$$X_{t_{n+1}} = X_{t_n} + a(X_{t_n})h + b(X_{t_n})\Delta W_{t_n} + b'(X_n)b(X_n)\frac{1}{2}(\Delta W_{t_n}^2 - h) + \text{Rest}'_{t_n},$$

wobei

$$\begin{aligned} \text{Rest}'_{t_n} &:= \int_{t_n}^{t_{n+1}} \left[\int_{t_n}^t a'(X_s) a(X_s) + \frac{1}{2} a''(X_s) b^2(X_s) ds + \int_{t_n}^t a'(X_s) b(X_s) dW_s \right] dt \\ &\quad + \int_{t_n}^{t_{n+1}} \left[\int_{t_n}^t b'(X_s) a(X_s) + \frac{1}{2} b''(X_s) b^2(X_s) ds \right] dW_t + \text{Rest}_{3,t_n} \\ &=: \text{Rest}_{t_n} - b'(X_n) b(X_n) \frac{1}{2} (\Delta W_{t_n}^2 - h). \end{aligned}$$

Zu bemerken ist noch, dass Rest'_{t_n} aus fünf Doppelintegralen besteht mit mindestens einer Zeitintegration und zwei Dreifachintegrale enthält (vgl. Definition Rest_{3,t_n}).

2.2 Konvergenztheorie für das Euler-Maruyama-Verfahren

2.2.1 Exakte Lösung vs. diskretes Verfahren

Für die exakte Lösung X gilt nach (2.5)

$$\begin{aligned} X_{t_{n+1}} - X_{t_n} &= a(X_{t_n})h + b(X_{t_n})\Delta W_{t_n} + \text{Rest}_{t_n}, \quad n = 0, 1, 2, \dots, \\ X_0 &= \xi. \end{aligned}$$

Für das Euler-Maruyama-Verfahren gilt

$$\begin{aligned} \hat{X}_{t_{n+1}} - \hat{X}_{t_n} &= a(\hat{X}_{t_n})h + b(\hat{X}_{t_n})\Delta W_{t_n}, \quad n = 0, 1, 2, \dots, \\ \hat{X}_0 &= \xi. \end{aligned}$$

2.2.2 Analyse des L^2 -Fehlers

Wir untersuchen im Folgenden den L^2 -Fehler $e_{n+1} := E \left(\left| X_{t_{n+1}} - \hat{X}_{t_{n+1}} \right|^2 \right)$ des Euler-Maruyama-Verfahrens:

$$\begin{aligned} e_{n+1} &= E \left(\left| X_{t_{n+1}} - \hat{X}_{t_{n+1}} \right|^2 \right) \\ &= E \left(\left| \left(\sum_{k=0}^n (X_{t_{k+1}} - X_{t_k}) + \underbrace{X_0}_{=\xi} \right) - \left(\sum_{k=0}^n (\hat{X}_{t_{k+1}} - \hat{X}_{t_k}) + \underbrace{\hat{X}_0}_{=\xi} \right) \right|^2 \right) \\ &= E \left(\left| \sum_{k=0}^n \left((X_{t_{k+1}} - X_{t_k}) - (\hat{X}_{t_{k+1}} - \hat{X}_{t_k}) \right) \right|^2 \right) \\ &= E \left(\left| \sum_{k=0}^n \left(a(X_{t_k}) - a(\hat{X}_{t_k}) \right) h + \sum_{k=0}^n \left(b(X_{t_k}) - b(\hat{X}_{t_k}) \right) \Delta W_{t_k} \right. \right. \\ &\quad \left. \left. + \sum_{k=0}^n b'(X_{t_k})b(X_{t_k}) \frac{1}{2} (\Delta W_{t_k}^2 - h) + \sum_{k=0}^n \text{Rest}'_{t_k} \right|^2 \right). \end{aligned}$$

Für eine endliche Folge a_i mit festem m gilt

$$(a_1 + a_2 + \dots + a_m)^2 \leq m (a_1^2 + a_2^2 + \dots + a_m^2). \quad (2.11)$$

Somit folgt für den Fehler e_{n+1} mit $m = 4$

$$\begin{aligned} e_{n+1} &\leq 4E \left(\left| \sum_{k=0}^n \left(a(X_{t_k}) - a(\hat{X}_{t_k}) \right) h \right|^2 \right) + 4E \left(\left| \sum_{k=0}^n \left(b(X_{t_k}) - b(\hat{X}_{t_k}) \right) \Delta W_{t_k} \right|^2 \right) \\ &\quad + 4E \left(\left| \sum_{k=0}^n b'(X_{t_k})b(X_{t_k}) \frac{1}{2} (\Delta W_{t_k}^2 - h) \right|^2 \right) + 4E \left(\left| \sum_{k=0}^n \text{Rest}'_{t_k} \right|^2 \right) \\ &=: 4 \{ (I) + (II) + (III) + (IV) \}. \end{aligned} \quad (2.12)$$

2.2.3 Separate Analyse der Terme (I)-(IV)

Summand (I):

$$\begin{aligned}
 \left| \sum_{k=0}^n \left(a(X_{t_k}) - a(\hat{X}_{t_k}) \right) h \right|^2 &\stackrel{(2.11)}{\leq} (n+1)h^2 \sum_{k=0}^n \left(a(X_{t_k}) - a(\hat{X}_{t_k}) \right)^2 \\
 &\leq \underbrace{(n+1)h}_{=t_{n+1} \leq T} \sum_{k=0}^n h \left(a(X_{t_k}) - a(\hat{X}_{t_k}) \right)^2 \\
 &\leq Th \sum_{k=0}^n L_a^2 |X_{t_k} - \hat{X}_{t_k}|^2,
 \end{aligned}$$

falls a Lipschitz-stetig ist mit Konstante L_a . Somit erhalten wir für den Erwartungswert

$$\begin{aligned}
 E \left(\left| \sum_{k=0}^n \left(a(X_{t_k}) - a(\hat{X}_{t_k}) \right) h \right|^2 \right) &\leq TL_a^2 h \sum_{k=0}^n \underbrace{E \left(|X_{t_k} - \hat{X}_{t_k}|^2 \right)}_{=e_k} \\
 &= TL_a^2 h \sum_{k=0}^n e_k.
 \end{aligned} \tag{2.13}$$

Summand (II):

Sei $S_k := \left(b(X_{t_k}) - b(\hat{X}_{t_k}) \right) \Delta W_{t_k}$ für $k = 0, \dots, N$. Dann gilt

$$\begin{aligned}
 E(S_k) &= E \left(b(X_{t_k}) - b(\hat{X}_{t_k}) \right) \cdot \underbrace{E(\Delta W_{t_k})}_{=0, \text{ da } \Delta W_{t_k} \sim \mathcal{N}(0, h)} = 0, \\
 E(S_k^2) &= E \left(\left(b(X_{t_k}) - b(\hat{X}_{t_k}) \right)^2 \right) \cdot \underbrace{E(\Delta W_{t_k}^2)}_{=h \text{ nach Aufgabe 3, Blatt 1}} \\
 &= E \left(\left(b(X_{t_k}) - b(\hat{X}_{t_k}) \right)^2 \right) \cdot h.
 \end{aligned}$$

Wir finden somit

$$\begin{aligned}
 E \left(\left| \sum_{k=0}^n \left(b(X_{t_k}) - b(\hat{X}_{t_k}) \right) \Delta W_{t_k} \right|^2 \right) &= E \left(\left| \sum_{k=0}^n S_k \right|^2 \right) \\
 &= E \left(\sum_{k=0}^n S_k^2 \right) + 2 \sum_{k < l} E(S_k \cdot S_l) \\
 &= \sum_{k=0}^n E(S_k^2) + 2 \sum_{k < l} E(S_k \cdot S_l).
 \end{aligned}$$

2.2 Konvergenztheorie für das Euler-Maruyama-Verfahren

Da S_k und S_l stochastisch unabhängig sind, gilt $E(S_k \cdot S_l) = E(S_k)E(S_l) = 0$ und somit folgt

$$\begin{aligned}
 E \left(\left| \sum_{k=0}^n \left(b(X_{t_k}) - b(\hat{X}_{t_k}) \right) \Delta W_{t_k} \right|^2 \right) &= \sum_{k=0}^n E \left(\left(b(X_{t_k}) - b(\hat{X}_{t_k}) \right)^2 \right) \cdot h \\
 &\leq L_b^2 h \sum_{k=0}^n \underbrace{E \left(|X_{t_k} - \hat{X}_{t_k}|^2 \right)}_{=e_k} \\
 &= L_b^2 h \sum_{k=0}^n e_k,
 \end{aligned} \tag{2.14}$$

falls b Lipschitz-stetig ist mit Konstante L_b .

Summand (III):

Sei $S_k := \frac{1}{2}b'(X_{t_k})b(X_{t_k})(\Delta W_{t_k}^2 - h)$ für $k = 0, \dots, N$. Wie oben folgt mit Aufgabe 3, Blatt 1

$$\begin{aligned}
 E(\Delta W_{t_k}^2) &= h \quad \text{und} \quad E(\Delta W_{t_k}^4) = 3h^2, \quad \text{d. h.} \\
 E \left(\frac{1}{2}(\Delta W_{t_k}^2 - h) \right) &= \frac{1}{2}(E(\Delta W_{t_k}^2) - h) = 0 \quad \text{und} \\
 E \left(\frac{1}{2}(\Delta W_{t_k}^2 - h)^2 \right) &= \frac{3}{2}h^2 - h^2 + \frac{1}{2}h^2 = h^2.
 \end{aligned}$$

Somit gilt für $k = 0, \dots, N$

$$E(S_k) = 0 \quad \text{und} \quad E(S_k^2) \leq Ch^2.$$

Ferner sind S_k und S_l für $k < l$ stochastisch unabhängig und damit erhalten wir also

$$\begin{aligned}
 E \left(\left| \sum_{k=0}^n \frac{1}{2}b'(X_{t_k})b(X_{t_k})(\Delta W_{t_k}^2 - h) \right|^2 \right) &= E \left(\left| \sum_{k=0}^n S_k \right|^2 \right) \\
 &= \sum_{k=0}^n \underbrace{E(S_k^2)}_{\leq Ch^2} + 2 \sum_{k < l} \underbrace{E(S_k \cdot S_l)}_{=E(S_k)E(S_l)=0} \\
 &\leq C \sum_{k=0}^n h^2 \\
 &= C(n+1)h^2 \\
 &\leq CTh.
 \end{aligned} \tag{2.15}$$

Summand (IV):

$$\begin{aligned}
 E \left(\left| \sum_{k=0}^n \text{Rest}'_{t_k} \right|^2 \right) &= \sum_{k=0}^n E \left((\text{Rest}'_{t_k})^2 \right) + 2 \sum_{j < k} \underbrace{E \left(\text{Rest}'_{t_j} \cdot \text{Rest}'_{t_k} \right)}_{=E(\text{Rest}'_{t_j})E(\text{Rest}'_{t_k})} \\
 &= \sum_{k=0}^n E \left((\text{Rest}'_{t_k})^2 \right) + 2 \sum_{j < k} E \left(\text{Rest}'_{t_j} \right) E \left(\text{Rest}'_{t_k} \right).
 \end{aligned} \tag{2.16}$$

2.2 Konvergenztheorie für das Euler-Maruyama-Verfahren

Dabei sind die Einzeltermen in Rest'_{t_j} und Rest'_{t_k} entweder mindestens doppelte reine Zeitintegrale oder es ist mindestens ein stochastisches Integral dabei. Somit folgt für jeden Einzelterm ET

$$E(\text{ET}) \leq \begin{cases} Ch^2, & \text{falls doppeltes Zeitintegral,} \\ 0, & \text{für } E\left(\int_a^b f_t dW_t\right) = 0. \end{cases}$$

Insgesamt erhalten wir für alle l

$$E(\text{Rest}'_{t_l}) \leq C_1 h^2.$$

Somit folgt

$$\begin{aligned} \sum_{\substack{j,k=0 \\ j < k}}^n E(\text{Rest}'_{t_j}) E(\text{Rest}'_{t_k}) &\leq \sum_{\substack{j,k=0 \\ j < k}}^n C_1 h^2 \cdot C_1 h^2 = C_1^2 h^4 \sum_{\substack{j,k=0 \\ j < k}}^n 1 = C_1^2 h^4 \frac{n(n+1)}{2} \\ &\leq \frac{1}{2} C_1^2 h^2 \underbrace{(n^2 + n)}_{\leq 2n^2} h^2 \leq C_1^2 h^2 \underbrace{n^2 h^2}_{\leq T^2} \leq C_1^2 h^2 T^2. \end{aligned} \quad (2.17)$$

Ferner lässt sich nun das Folgende zeigen:

$$E\left((\text{Rest}'_{t_k})^2\right) \leq C_0 h^3, \quad (2.18)$$

falls a, a', a'' und b, b', b'' beschränkt sind. Dabei sind Terme der folgenden Formen zu berechnen.

a) Doppel-Zeit-Integral:

$$E\left(\left|\int_{t_n}^{t_{n+1}} \int_{t_n}^t f_s ds dt\right|^2\right) = ?$$

Mit $|f_s| \leq C$ folgt $\left|\int_{t_n}^{t_{n+1}} \int_{t_n}^t f_s ds dt\right| \leq C \left|\int_{t_n}^{t_{n+1}} \int_{t_n}^t ds dt\right| = \frac{C}{2} h^2$ und somit

$$E\left(\left|\int_{t_n}^{t_{n+1}} \int_{t_n}^t f_s ds dt\right|^2\right) \leq C_1 h^4.$$

b) Zeit-Itô-Integral:

$$E\left(\left|\int_{t_n}^{t_{n+1}} \int_{t_n}^t f_s ds dW_t\right|^2\right) = ?$$

Mit Hilfe der Itô-Isometrie, d. h.

$$E\left(\left(\int_c^d f_s dW_s\right)^2\right) = \int_c^d E(f_s^2) ds$$

und $|f_s| \leq C$ folgt

$$\begin{aligned} E\left(\left|\int_{t_n}^{t_{n+1}} \left(\int_{t_n}^t f_s ds\right) dW_t\right|^2\right) &= \int_{t_n}^{t_{n+1}} E\left(\left|\int_{t_n}^t f_s ds\right|^2\right) dt \\ &\leq \int_{t_n}^{t_{n+1}} E(C_2 h^2) dt \leq C_2 h^3. \end{aligned}$$

c) Dreifach-Itô-Integral:

$$E \left(\left| \int_{t_n}^{t_{n+1}} \int_{t_n}^t \int_{t_n}^s f_\eta dW_\eta dW_s dW_t \right|^2 \right) = ?$$

Wie in b) folgt mit dreimaliger Anwendung der Itô-Isometrie und $|f_\eta|^2 \leq C$

$$\begin{aligned} E \left(\left| \int_{t_n}^{t_{n+1}} \int_{t_n}^t \int_{t_n}^s f_\eta dW_\eta dW_s dW_t \right|^2 \right) &= \int_{t_n}^{t_{n+1}} \int_{t_n}^t \int_{t_n}^s E(f_\eta^2) d\eta ds dt \\ &\leq \int_{t_n}^{t_{n+1}} \int_{t_n}^t \int_{t_n}^s C d\eta ds dt \leq C_3 h^3. \end{aligned}$$

Somit folgt für C_0 geeignet die Abschätzung (2.18). Wir erhalten also

$$\sum_{k=0}^n E \left((\text{Rest}'_{t_k})^2 \right) \leq \sum_{k=0}^n C_0 h^3 = \underbrace{(n+1)h}_{\leq T} C_0 h^2 \leq C_0 T h^2. \quad (2.19)$$

Einsetzen von (2.19) und (2.17) in (2.16) liefert

$$\begin{aligned} E \left(\left| \sum_{k=0}^n \text{Rest}'_{t_k} \right|^2 \right) &= \sum_{k=0}^n E \left((\text{Rest}'_{t_k})^2 \right) + 2 \sum_{j < k} E \left(\text{Rest}'_{t_j} \right) E \left(\text{Rest}'_{t_k} \right) \\ &\leq C_0 T h^2 + 2 C_1^2 T^2 h^2 \leq C h^2. \end{aligned} \quad (2.20)$$

2.2.4 Das diskrete Gronwall-Lemma

Mit einer generischen Konstante $C > 0$ erhalten wir aus (2.12), (2.13), (2.14), (2.15) und (2.20) die Abschätzung

$$e_{n+1} \leq C h^p + C h \sum_{k=0}^n e_k \quad (2.21)$$

mit $p = 1$.

2.2 Bemerkung

Man beachte, dass ohne den Term (III) die Abschätzung (2.21) mit $p = 2$ gelten würde.

Für die weitere Abschätzung von e_{n+1} benötigen wir noch das folgende Lemma.

2.3 Lemma (Diskretes Gronwall-Lemma)

Seien $b, d \geq 0$ und $(u_n)_{n \in \mathbb{N}}$ eine reelle Folge mit $u_n \geq 0$. Ferner gelten folgende Abschätzungen:

$$u_0 \leq d \quad \text{und} \quad u_n \leq d + b \sum_{k=0}^{n-1} u_k, \quad n = 1, 2, \dots$$

Dann gilt

$$u_n \leq d \exp(nb), \quad n = 1, 2, \dots$$

Beweis. Siehe Übung. □

2.4 Korollar

Für den Erwartungswert der Differenz zwischen exakter und diskreter Lösung gilt

$$E \left(\left| X_{t_n} - \hat{X}_{t_n} \right| \right) \leq \sqrt{C \exp(CT)} \cdot \sqrt{h},$$

wobei $n \in \mathbb{N}$ mit $nh \leq T$.

Beweis. Für die Abschätzung (2.21) gilt nach dem diskreten Gronwall-Lemma mit $d := Ch$ und $b := Ch$

$$e_n \leq Ch \exp(Cnh) \leq C \exp(CT) \cdot h.$$

Mit der Jensenschen Ungleichung für den Erwartungswert, d. h.

$$E(|A|)^2 \leq E(|A|^2) \quad (A \text{ ist integrierbare ZV, } \varphi(x) = x^2)$$

folgt somit

$$E \left(\left| X_{t_n} - \hat{X}_{t_n} \right| \right) \leq \sqrt{E \left(\left| X_{t_n} - \hat{X}_{t_n} \right|^2 \right)} \leq \sqrt{e_n} \leq \sqrt{C \exp(CT)} \cdot \sqrt{h}, \quad n = 0, \dots, N.$$

□

2.2.5 Starke Konvergenz

Wir kommen nun zum Begriff der starken Konvergenz.

2.5 Definition (Starke Konvergenz)

Sei $(X_t)_{0 \leq t \leq T}$ die Lösung der SDGL

$$\begin{aligned} dX_t &= a(t, X_t)dt + b(t, X_t)dW_t, \\ X_0 &= \bar{X} \end{aligned}$$

und sei $\hat{X}_{t_n}$, $t_n = nh$, $0 \leq nh \leq T$ eine Approximation von $(X_t)_{0 \leq t \leq T}$. Wir sagen $\hat{X}$ konvergiert stark gegen X , wenn gilt:

$$\varepsilon(h) := \sup \left\{ E \left(\left| X_{t_n} - \hat{X}_{t_n} \right| \right) : 0 \leq nh \leq T \right\} \rightarrow 0 \quad \text{für } h \rightarrow 0.$$

Die Konvergenz ist von der Ordnung $\gamma > 0$, falls eine Konstante $C > 0$ existiert mit

$$\varepsilon(h) \leq Ch^\gamma, \quad 0 < h \leq h_0$$

für ein $h_0 > 0$.

Nach dem bisher Gezeigten erhalten wir nun folgendes Konvergenzresultat.

2.6 Satz

Für Lipschitz-stetige C^2 -Funktionen a und b , deren partielle Ableitungen bis zur Ordnung 2 ebenfalls beschränkt sind, konvergiert das Euler-Maruyama-Verfahren stark mit der Ordnung $\frac{1}{2}$.

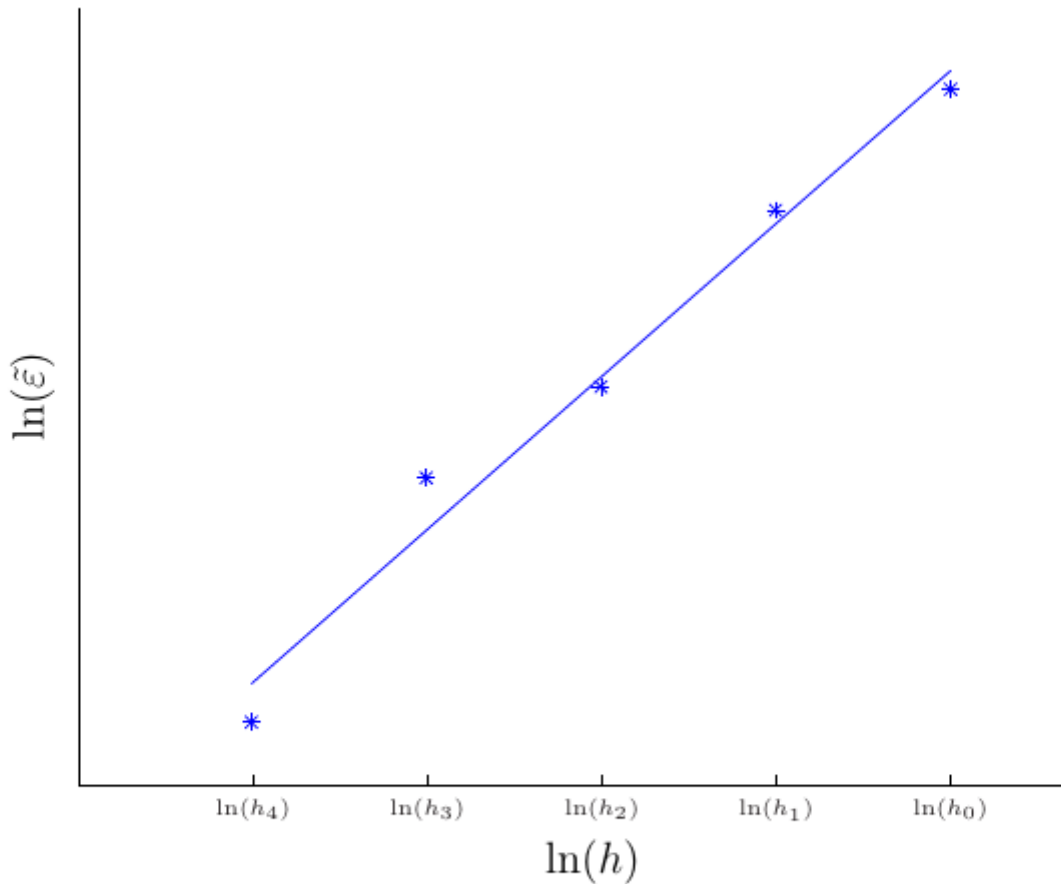
2.2.6 Numerische Bestimmung der Konvergenzordnung

Seien $h_0 > h_1 > h_2 > \dots > h_m$ vorgegeben. Wir berechnen hierzu eine möglichst gute Approximation $\tilde{\varepsilon}(h_i)$ des maximalen zeitlichen Fehlers $\varepsilon(h_i)$ ($i = 0, \dots, m$) und zeichnen das Paar $(\ln(h_i), \ln(\tilde{\varepsilon}(h_i)))$ in ein $(\ln(h), \ln(\tilde{\varepsilon}))$ -Diagramm, vgl. Abbildung ??.

Die Steigung der Ausgleichsgeraden entspricht annähernd der Konvergenzordnung, denn für $\varepsilon(h) = Ch^\gamma$ gilt

$$\ln(\varepsilon(h)) = \ln(C) + \ln(h^\gamma) = \gamma \cdot \ln(h) + \ln(C),$$

wobei γ die Steigung der Geraden ist.



Abbildungung 2.1: Ausgleichsgerade mit 4 verschiedenen Schrittweiten.

2.7 Bemerkung (Berechnung von $\tilde{\varepsilon}(h)$)

- a) Da wir zur Berechnung vom Erwartungswert $E\left(\left|X_{t_n} - \hat{X}_{t_n}\right|\right)$ die exakte Lösung einer SDGL benötigen, wählen wir für das Experiment eine einfache SDGL, von welcher die Lösung bekannt ist, z. B.

$$\begin{aligned} dX_t &= aX_t dt + bX_t dW_t, \\ X_0 &= \bar{X} \end{aligned}$$

2.2 Konvergenztheorie für das Euler-Maruyama-Verfahren

mit der Lösung $X_t = \bar{X} \exp\left(\left(a - \frac{1}{2}b^2\right)t + bW_t\right)$.

- b) Wir brauchen $E\left(\left|X_{t_n} - \hat{X}_{t_n}\right|\right)$, $t_n = nh$, $0 \leq t_n \leq T$. Den Erwartungswert können wir jedoch nur approximieren. Man benötigt hierzu PZZ η_i ($i = 1, \dots, M$), welche wie $\left|X_{t_n} - \hat{X}_{t_n}\right|$ verteilt sind. Dann gilt

$$E\left(\left|X_{t_n} - \hat{X}_{t_n}\right|\right) \approx \frac{1}{M} \sum_{i=1}^M \eta_i,$$

für $M \in \mathbb{N}$ hinreichend groß.

Die PZZ η_i ($i = 1, \dots, M$) konstruieren wir durch M unabhängige Berechnungen von $\left|X_{t_n}(\omega) - \hat{X}_{t_n}(\omega)\right|$. Hierfür benötigen wir M unabhängige Brownsche Bewegungen

$$W_t^{(1)}(\omega), \dots, W_t^{(M)}(\omega).$$

Es gilt

$$W_{t_m}^{(k)}(\omega) = \sum_{j=0}^{m-1} \Delta W_{t_j}^{(k)}(\omega) =: \sum_{j=0}^{m-1} \xi_{t_j}^{(k)}, \quad k = 1, \dots, M, \quad (2.22)$$

wobei $\Delta W_{t_j}^{(k)} \sim \mathcal{N}(0, h)$.

Generiere also die $N \cdot M$ PZZ

$$\xi_{t_0}^{(1)}, \dots, \xi_{t_{N-1}}^{(1)}, \xi_{t_0}^{(2)}, \dots, \xi_{t_{N-1}}^{(2)}, \dots, \xi_{t_0}^{(M)}, \dots, \xi_{t_{N-1}}^{(M)},$$

wobei alle $\mathcal{N}(0, h)$ -verteilt sind, und bilde damit $W_{t_m}^{(k)}(\omega)$, $k = 1, \dots, M$, $m = 1, \dots, N$ gemäß Gleichung (2.22). Dabei muss M hinreichend groß gewählt werden, damit $\tilde{\varepsilon}(h) \approx \varepsilon(h)$.

2.2.7 Fehler bei der Erwartungswertberechnung

Seien $X_1, X_2, X_3, \dots$ i. i. d. wie X mit $E(X) = \mu$ und $V(X) = \sigma^2 < \infty$. Setze

$$S_M := \frac{1}{M} \sum_{i=1}^M X_i$$

und finde

$$\begin{aligned} E(S_M) &= \frac{1}{M} \sum_{i=1}^M E(X_i) = \frac{1}{M} \sum_{i=1}^M E(X) = \mu, \\ V(S_M) &= \frac{1}{M^2} \sum_{i=1}^M V(X_i) = \frac{1}{M^2} \sum_{i=1}^M V(X) = \frac{\sigma^2}{M}. \end{aligned}$$

Ferner ist nach dem zentralen Grenzwertsatz

$$\hat{S}_M := \frac{S_M - \mu}{\sqrt{V(S_M)}}$$

2.2 Konvergenztheorie für das Euler-Maruyama-Verfahren

für M hinreichend groß approximativ $\mathcal{N}(0, 1)$ -verteilt.

Herleitung eines Konfidenzintervalls:

Sei α vorgegeben. Bestimme dazu $r > 0$ mit

$$P\left(\left|\hat{S}_M\right| \leq r\right) = 1 - \alpha.$$

Da $\hat{S}_M$ approximativ $\mathcal{N}(0, 1)$ -verteilt ist, folgt

$$\begin{aligned} 1 - \alpha &= \int_{-r}^r \frac{1}{\sqrt{2\pi}} \exp(-x^2/2) dx \\ &= \frac{2}{\sqrt{2\pi}} \int_0^r \exp(-x^2/2) dx \\ &\stackrel{(*)}{=} \frac{2}{\sqrt{2\pi}} \int_0^{r/\sqrt{2}} \exp(-w^2) \sqrt{2} dw \\ &= \frac{2}{\sqrt{\pi}} \int_0^{r/\sqrt{2}} \exp(-w^2) dw \\ &= \operatorname{erf}\left(\frac{r}{\sqrt{2}}\right). \end{aligned}$$

Dabei wurde in $(*)$ die Substitution $w := \frac{x}{\sqrt{2}}$, $dw = \frac{dx}{\sqrt{2}}$ vorgenommen. Damit folgt für r

$$\sqrt{2} \cdot \operatorname{erf}^{-1}(1 - \alpha) = r.$$

Wir erhalten also mit einer Wahrscheinlichkeit von $1 - \alpha$

$$\hat{S}_M \in [-r, r] = r[-1, 1].$$

Mit

$$\hat{S}_M = \frac{S_M - \mu}{\sqrt{V(S_M)}} = \frac{S_M - \mu}{\frac{\sigma}{\sqrt{M}}} = \frac{\sqrt{M}(S_M - \mu)}{\sigma}$$

folgt

$$S_M = \frac{\sigma \hat{S}_M}{\sqrt{M}} + \mu$$

und damit

$$S_M \in \mu + \frac{\sigma}{\sqrt{M}} r[-1, 1] = \mu + \frac{\sigma}{\sqrt{M}} \sqrt{2} \cdot \operatorname{erf}^{-1}(1 - \alpha)[-1, 1]$$

mit einer Wahrscheinlichkeit von $1 - \alpha$.

Für $\alpha = 0.01$ erhalten wir $\sqrt{2} \operatorname{erf}^{-1}(1 - \alpha) \approx 2.6$ und somit

$$S_M \in \mu + \frac{2.6\sigma}{\sqrt{M}}[-1, 1]$$

mit einer Wahrscheinlichkeit von 0.99.

2.3 Weitere numerische Verfahren und schwache Konvergenz

Ferner lesen wir ab: Die Intervallbreite verhält sich in M wie $\frac{1}{\sqrt{M}}$ und in der Varianz wie $\sigma = \sqrt{V(X)}$. In unserem Fall gilt

$$\tilde{\varepsilon}(t_n, h) = \frac{1}{M} \sum_{i=1}^M \eta_i(t_n, h),$$

wobei $\tilde{\varepsilon}(t_n, h)$ die Approximation des zeitlichen Fehlers $\varepsilon(t_n, h) = E \left(\left| X_{t_n} - \hat{X}_{t_n} \right| \right)$ zur Zeit t_n mit Schrittweite h ist. Dabei sind die $\eta_i(t_n, h)$ verteilt wie $\left| X_{t_n} - \hat{X}_{t_n} \right|$.

Aus der Konvergenzanalyse folgt (vgl. Beweis zu Korollar 2.4)

$$E \left(\left| X_{t_n} - \hat{X}_{t_n} \right|^2 \right) \leq Ch$$

und somit

$$V \left(\left| X_{t_n} - \hat{X}_{t_n} \right| \right) = \underbrace{E \left(\left| X_{t_n} - \hat{X}_{t_n} \right|^2 \right)}_{\leq Ch} - \underbrace{E \left(\left| X_{t_n} - \hat{X}_{t_n} \right| \right)^2}_{\geq 0} \leq Ch,$$

d. h.

$$\sigma_{t_n} = \sqrt{V \left(\left| X_{t_n} - \hat{X}_{t_n} \right| \right)} \leq \hat{C} \sqrt{h}.$$

Mit einer Wahrscheinlichkeit von 0.99 erhalten wir also

$$\tilde{\varepsilon}(t_n, h) - \varepsilon(t_n, h) \in \frac{2.6\sigma_{t_n}}{\sqrt{M}} [-1, 1] \subset C_1 \frac{\sqrt{h}}{\sqrt{M}} [-1, 1].$$

Insgesamt folgt dann

$$\begin{aligned} |\tilde{\varepsilon}(t_n, h)| &\leq |\tilde{\varepsilon}(t_n, h) - \varepsilon(t_n, h)| + |\varepsilon(t_n, h)| \\ &\leq C_1 \frac{\sqrt{h}}{\sqrt{M}} + C_2 \sqrt{h} = \sqrt{h} \left(\frac{C_1}{\sqrt{M}} + C_2 \right) \\ &\leq C_3 \sqrt{h} \text{ unabhängig von } M. \end{aligned}$$

Man beachte dabei aber, dass der zentrale Grenzwertsatz greifen sollte, d. h. M hinreichend groß gewählt ist.

2.3 Weitere numerische Verfahren und schwache Konvergenz

2.3.1 Das Milstein-Verfahren

Betrachte die SDGL

$$\begin{aligned} dX_t &= a(X_t)dt + b(X_t)dW_t, \quad 0 \leq t \leq T, \\ X_0 &= \xi. \end{aligned}$$

2.3 Weitere numerische Verfahren und schwache Konvergenz

Als Verbesserung des Euler-Maruyama-Verfahrens setzen wir für $N \in \mathbb{N}$, $h = \frac{T}{N} > 0$, $t_n = nh$, $n = 0, \dots, N$ folgendermaßen an:

$$\begin{aligned}\hat{X}_{t_{n+1}} &= \hat{X}_{t_n} + a(\hat{X}_{t_n})h + b(\hat{X}_{t_n})\Delta W_{t_n} + \frac{1}{2}b'(\hat{X}_{t_n})b(\hat{X}_{t_n})(\Delta W_{t_n}^2 - h), \\ \hat{X}_0 &= \xi.\end{aligned}$$

Dies ist das sogenannte „Milstein-Verfahren“.

Für den L^2 -Fehler $e_{n+1} = E\left(\left|X_{t_{n+1}} - \hat{X}_{t_{n+1}}\right|^2\right)$ erhält man nun

$$\begin{aligned}e_{n+1} &\leq 4E\left(\left|\sum_{k=0}^n (a(X_{t_k}) - a(\hat{X}_{t_k}))h\right|^2\right) \\ &\quad + 4E\left(\left|\sum_{k=0}^n (b(X_{t_k}) - b(\hat{X}_{t_k}))\Delta W_{t_k}\right|^2\right) \\ &\quad + 4E\left(\left|\sum_{k=0}^n (b'(X_{t_k})b(X_{t_k}) - b'(\hat{X}_{t_k})b(\hat{X}_{t_k}))\frac{1}{2}(\Delta W_{t_k}^2 - h)\right|^2\right) \\ &\quad + 4E\left(\left|\sum_{k=0}^n \text{Rest}'_{t_k}\right|^2\right) \\ &=: 4\{(I) + (II) + (III) + (IV)\},\end{aligned}$$

wobei sich im Vergleich zum Euler-Maruyama-Verfahren nur der Term (III) geändert hat. Unter der Voraussetzung, dass die Funktion $b'(x)b(x)$ global Lipschitz-stetig ist, erhält man nun für diesen Term eine Abschätzung wie für die Terme (I) und (II), d. h.

$$(III) \leq \hat{C}h \sum_{k=0}^n e_k.$$

Insgesamt erhält man damit aus (I)–(IV) die Abschätzung

$$e_{n+1} \leq Ch^p + Ch \sum_{k=0}^n e_k$$

mit $p=2$ und einer universellen Konstanten $C > 0$ (vgl. (2.21) für das Euler-Maruyama-Verfahren).

Analog zum Euler-Maruyama-Verfahren folgt dann

$$E\left(\left|X_{t_n} - \hat{X}_{t_n}\right|\right) \leq \sqrt{C \exp(CT)} \cdot h,$$

für $n \in \mathbb{N}$ mit $nh \leq T$, d. h. wir erhalten starke Konvergenz der Ordnung 1 unter geeigneten Voraussetzungen.

2.8 Bemerkung

- a) Ersetzt man den Term $b'(X_t)b(X_t)$ durch eine Differenzenapproximation, so ergeben sich Verfahren vom Runge-Kutta-Typ.

2.3 Weitere numerische Verfahren und schwache Konvergenz

- b) Die Übertragung des Milstein-Verfahrens auf Systeme von SDGL ist nicht trivial, denn die Größe

$$\int_{t_n}^{t_{n+1}} \int_{t_n}^t dW_s^{(i)} dW_t^{(j)}$$

lässt sich nicht analytisch berechnen, wenn $W^{(i)}$ und $W^{(j)}$ unabhängige Brownsche Bewegungen sind.

2.3.2 Schwache Konvergenz

Starke Konvergenz bezieht sich auf die maximal erwartete Trajektorienabweichung, wobei das Fehlermaß durch

$$\varepsilon(h) = \sup \left\{ E \left(\left| X_{t_n} - \hat{X}_{t_n} \right| \right) : 0 \leq nh \leq T \right\}$$

gegeben ist. Oft ist aber lediglich die Größe $E(g(X_t))$ für gewisse Funktionen g von Interesse. Der relevante Fehler lautet dann

$$\omega_g(h) := \sup \left\{ \left| E(g(X_{t_n})) - E(g(\hat{X}_{t_n})) \right| : 0 \leq nh \leq T \right\}.$$

Dies führt zu folgender

2.9 Definition (Schwache Konvergenz)

Sei $X = (X_t)_{0 \leq t \leq T}$ die Lösung einer SDGL auf $[0, T]$ und sei $\hat{X} = (\hat{X}_{t_n})_{0 \leq nh \leq T}$, $h = \frac{T}{N} > 0$ eine Approximation. Wir sagen $\hat{X}$ konvergiert schwach gegen X bezüglich einer Klasse $\mathcal{C}$ von Testfunktionen, falls für den Fehler

$$\omega_g(h) = \sup \left\{ \left| E(g(X_{t_n})) - E(g(\hat{X}_{t_n})) \right| : 0 \leq nh \leq T \right\}$$

gilt

$$\forall g \in \mathcal{C} : \lim_{h \rightarrow 0} \omega_g(h) = 0.$$

Die Konvergenz ist von der Ordnung $\gamma > 0$, falls zu jedem $g \in \mathcal{C}$ eine Konstante $C > 0$ existiert mit

$$\omega_g(h) \leq Ch^\gamma, \quad 0 < h \leq h_0$$

für ein $h_0 > 0$.

2.10 Bemerkung

Für Lipschitz-stetige Funktionen g mit Konstante L_g gilt

$$\begin{aligned} \omega_g(h) &= \sup \left\{ \underbrace{\left| E(g(X_{t_n})) - E(g(\hat{X}_{t_n})) \right|}_{\leq E(|g(X_{t_n}) - g(\hat{X}_{t_n})|)} : 0 \leq nh \leq T \right\} \\ &\leq L_g \sup \left\{ E(|X_{t_n} - \hat{X}_{t_n}|) : 0 \leq nh \leq T \right\} = L_g \varepsilon(h). \end{aligned}$$

Dies bedeutet, dass die schwache Konvergenzordnung mindestens so groß ist wie die starke Konvergenzordnung.

2.11 Beispiel

a) Gesucht sei $E(X_t)$, d. h. $E(g(X_t))$ mit $g(x) = x$, und X_t genüge der SDGL

$$\begin{aligned} dX_t &= aX_t dt + bX_t dW_t, \\ X_0 &= \xi \end{aligned}$$

bzw.

$$X_t = \underbrace{X_0}_{=\xi} + \int_0^t aX_s ds + \int_0^t bX_s dW_s.$$

Für den Erwartungswert gilt somit

$$\begin{aligned} E(X_t) &= E(\xi) + E\left(\int_0^t aX_s ds\right) + \underbrace{E\left(\int_0^t bX_s dW_s\right)}_{=0} = E(\xi) + \int_0^t aE(X_s) ds \\ \Rightarrow \frac{d}{dt}E(X_t) &= aE(X_t), \quad E(X_0) = E(\xi), \end{aligned}$$

d. h.

$$E(X_t) = E(\xi) \exp(at).$$

Wenden wir das Euler-Maruyama-Verfahren an, so gilt

$$\begin{aligned} E(\hat{X}_{t_{n+1}}) &= E(\hat{X}_{t_n}) + aE(\hat{X}_{t_n})h, \quad n = 0, 1, 2, \dots, \\ E(\hat{X}_0) &= E(\xi). \end{aligned}$$

Dies entspricht dem Euler-Verfahren für gewöhnliche DGL der Form

$$\begin{aligned} \dot{u} &= au, \\ u(0) &= E(\xi). \end{aligned}$$

Das Euler-Verfahren konvergiert mit Ordnung 1 gegen die Lösung $E(X_t) = E(\xi) \cdot \exp(at)$, d. h.

$$\omega_{\text{id}}(h) = \sup \left\{ \left| E(X_{t_n}) - E(\hat{X}_{t_n}) \right| : 0 \leq nh \leq T \right\} \leq Ch.$$

In diesem Beispiel ist die schwache Konvergenzordnung 1 also besser als die starke Konvergenzordnung $\frac{1}{2}$.

b) Bestimmung der schwachen Konvergenzordnung bei rein numerischer Auswertung von $E(\hat{X}_{t_n})$:

Bestimme durch M -fache Wiederholung des Euler-Maruyama-Algorithmus mit unabhängigen Zuwächsen die Größen $\hat{X}_{t_n}^{(1)}, \hat{X}_{t_n}^{(2)}, \dots, \hat{X}_{t_n}^{(M)}$. Dann gilt

$$\left| E(\hat{X}_{t_n}) - \frac{1}{M} \sum_{i=1}^M \hat{X}_{t_n}^{(i)} \right| \leq C \cdot \frac{\sigma(\hat{X}_{t_n})}{\sqrt{M}} = C \cdot \frac{\sqrt{V(\hat{X}_{t_n})}}{\sqrt{M}}$$

2.3 Weitere numerische Verfahren und schwache Konvergenz

mit hoher Wahrscheinlichkeit (vgl. Konfidenzintervallbestimmung). Somit folgt für $g = \text{id}$

$$\begin{aligned}\tilde{\omega}_{\text{id}}(t_n, h) &= \left| E(X_{t_n}) - \frac{1}{M} \sum_{i=1}^M \hat{X}_{t_n}^{(i)} \right| \\ &\leq \underbrace{\left| E(X_{t_n}) - E(\hat{X}_{t_n}) \right|}_{=\omega_{\text{id}}(t_n, h)} + \left| E(\hat{X}_{t_n}) - \frac{1}{M} \sum_{i=1}^M \hat{X}_{t_n}^{(i)} \right| \\ &\leq \omega_{\text{id}}(t_n, h) + C \cdot \frac{\sqrt{V(\hat{X}_{t_n})}}{\sqrt{M}} = \omega_{\text{id}}(t_n, h) + \frac{\hat{C}}{\sqrt{M}}.\end{aligned}$$

Beachte dabei: $V(\hat{X}_{t_n})$ verschwindet nicht für $h \rightarrow 0$, denn

$$V(\hat{X}_{t_n}) \rightarrow V(X_{t_n}) \quad \text{für } h \rightarrow 0.$$

Mit $\omega_{\text{id}}(t_n, h) \leq \hat{C}h$ folgt damit

$$\tilde{\omega}_{\text{id}}(t_n, h) \leq \hat{C} \left(h + \frac{1}{\sqrt{M}} \right).$$

Konsequenz: $M \sim \frac{1}{h^2}$, damit $\tilde{\omega}_{\text{id}}(t_n, h) \leq Ch$. Ansonsten sehen wir bei zu klein gewähltem M eine entsprechend niedrigere Ordnung für $\tilde{\omega}_{\text{id}}(t_n, h)$.

Wie sieht es für Funktionen g mit $g(x) \neq x$ aus? Unter geeigneten Voraussetzungen lässt sich für die Funktionenklasse

$$C_p^4(\mathbb{R}) := \{f \in C^4(\mathbb{R}) : f \text{ besitzt polynomiales Wachstum}\}$$

zeigen, dass das Euler-Maruyama-Verfahren schwach konvergent der Ordnung 1 bezüglich aller $g \in C_p^4(\mathbb{R})$ ist.

Weitere Konsequenzen:

Strebt man schwache Konvergenz an, so müssen Pfade von $\hat{X}$ keine gute Approximation der Pfade von X sein. Nur die Verteilung von $\hat{X}$ muss passen; dies bedeutet Folgendes: Sei

$$\begin{aligned}\hat{Y}_{t_{n+1}} &= \hat{Y}_{t_n} + a(\hat{Y}_{t_n})h + b(\hat{Y}_{t_n})\sqrt{h}\xi_n, \quad n = 0, 1, 2, \dots, N-1, \\ \hat{Y}_0 &= \eta\end{aligned}\tag{2.23}$$

mit $\xi_1, \xi_2, \dots, \xi_N$ i. i. d. wie ξ mit $E(\xi) = 0$ und $V(\xi) = 1$ gegeben. Dies entspricht dem Euler-Maruyama-Algorithmus für

$$\begin{aligned}dX_t &= a(X_t)dt + b(X_t)dW_t, \\ X_0 &= \eta,\end{aligned}$$

falls $\xi_1, \xi_2, \dots, \xi_N$ i. i. d. wie ξ mit $\xi \sim \mathcal{N}(0, 1)$. Dann gilt: Die „Approximation (2.23)“ besitzt die schwache Konvergenzordnung 1.

2.3 Weitere numerische Verfahren und schwache Konvergenz

Mögliche Realisierung der ZV ξ_i in (2.23): $\xi \in \{-1, 1\}$, $P(\xi = 1) = P(\xi = -1) = \frac{1}{2}$ und somit $E(\xi) = 0$ und $V(\xi) = 1$.

Anwendung bei Optionsbewertung:

Betrachte eine europäische Option in einem einfachen Marktmodell. Der Preisprozess des Basiswertes S sei gegeben durch die SDGL

$$\begin{aligned} dS_t &= rS_t dt + \sigma S_t dW_t, \quad 0 \leq t \leq T, \\ S_0 &= \xi \end{aligned}$$

und die Auszahlungsfunktionen zur Verfallszeit T seien

- Call: $V(S) = V(T, S) = (S - K)^+$,
- Put: $V(S) = V(T, S) = (K - S)^+$,

wobei K der Ausübungspreis ist. Der Optionswert zur Zeit $t = 0$ ist dann gegeben durch den Erwartungswert

$$V_0 = E(\exp(-rT)V(S_T)).$$

Er ist von der Form $E(g(S_T))$ mit $g(s) = \exp(-rT)V(s)$. Beachte hier, dass $g \notin C_p^4(\mathbb{R})$, d. h. das Konvergenzresultat für Funktionen der Klasse $\mathcal{C} = C_p^4(\mathbb{R})$ ist nicht anwendbar. Ob ein schwaches Konvergenzresultat trotzdem gültig ist, kann numerisch getestet werden. Die Optionsbewertung ist ein gutes Beispiel hierfür, da wegen der Black-Scholes-Formeln die exakten Erwartungswerte bekannt sind.

2.12 Bemerkung (Black-Scholes-Optionsbewertung)

Call:

$$V_0 = \xi \Phi(d_1) - K \exp(-rT) \Phi(d_2)$$

mit

$$\begin{aligned} d_1 &= \frac{1}{\sigma\sqrt{T}} \left(\ln\left(\frac{\xi}{K}\right) + \left(r + \frac{1}{2}\sigma^2\right)T \right), \\ d_2 &= d_1 - \sigma\sqrt{T} \end{aligned}$$

und Φ ist die Verteilungsfunktion der Standardnormalverteilung.

Put:

$$V_0 = K \exp(-rT) \Phi(-d_2) - \xi \Phi(-d_1).$$

Hierbei ist T die Verfallszeit und ξ der Anfangswert der SDGL.

Allgemeinere Situation:

Sei der Preisprozess modelliert durch die SDGL

$$\begin{aligned} dS_t &= r_t S_t dt + \sigma_t S_t dW_t, \\ S_0 &= \xi. \end{aligned} \tag{2.24}$$

Der Optionswert zur Zeit $t = 0$ ist dann gegeben durch

$$V_0 = E \left(\exp \left(- \int_0^T r_s ds \right) V(S_T) \right)$$

mit der Abzinsung $\exp \left(- \int_0^T r_s ds \right)$ und der Auszahlung $V(S_T)$.

Berechnungsalgorithmus:

- (i) Simuliere M unabhängige Realisierungen $V^{(i)}$ der Auszahlung $V : \omega \mapsto V(S_T(\omega))$ und $e^{(i)}$ der Abzinsung $\omega \mapsto \exp\left(-\int_0^T r_s(\omega) ds\right)$.
- (ii) Wähle die Größe $\frac{1}{M} \sum_{i=1}^M e^{(i)} V^{(i)}$ als Approximation von V_0 .

Zu (i):

Simuliere zunächst einen approximativen Pfad $\hat{S}^{(i)}$ durch Lösen der SDGL (2.24) mit dem Euler-Maruyama-Verfahren. Berechne dann $V^{(i)}$ aus $\hat{S}^{(i)}$.

Für die Auszahlungsfunktion

$$V(S_T) = \left(S_T - \frac{1}{T} \int_0^T S_t dt \right)^+$$

einer Asiatischen Option muss

$$\frac{1}{T} \int_0^T S_t dt$$

entsprechend approximiert werden, z. B.

$$\frac{1}{T} \int_0^T S_t dt \approx \frac{1}{T} \sum_{k=0}^{N-1} \hat{S}_{t_k}^{(i)} \cdot h, \quad h = \frac{T}{N}.$$

2.13 Bemerkung

Der Fehler bei der Erwartungswertberechnung ist proportional zu $\frac{1}{\sqrt{M}}$ und zur Standardabweichung (d. h. $\sqrt{\text{Varianz}}$) der relevanten ZV (vgl. Konfidenzintervallbestimmung). Er kann also verkleinert werden, in dem M entsprechend groß gewählt wird oder man versucht mit geeigneten Verfahren die Varianz der relevanten ZV zu reduzieren.

3 Parabolische partielle Differentialgleichungen und Optionspreisbewertung

3.1 Die Fokker-Planck-Gleichung

3.1.1 Einführung

Betrachte die skalare SDGL

$$\begin{aligned} dX_t &= a(t, X_t)dt + b(t, X_t)dW_t, \quad 0 \leq t \leq T, \\ X_0 &= \xi, \quad \xi \in \mathbb{R}. \end{aligned}$$

Angenommen, $\psi(t, x)$ sei die Dichte der Verteilung von X_t . Dann gilt

$$E(g(X_t)) = \int_{\mathbb{R}} g(x)\psi(t, x)dx.$$

Um ψ zu bestimmen, verwenden wir beliebige, glatte Testfunktionen

$$f : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}.$$

Die Itô-Formel liefert dann

$$\begin{aligned} df(t, X_t) &= \left(\frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial x} dX_t + \frac{1}{2} \frac{\partial^2 f}{\partial x^2} b^2 dt \right) (t, X_t) \\ &= \left(\frac{\partial f}{\partial t} + a \frac{\partial f}{\partial x} + \frac{1}{2} b^2 \frac{\partial^2 f}{\partial x^2} \right) (t, X_t) dt + \left(b \frac{\partial f}{\partial x} \right) (t, X_t) dW_t. \end{aligned}$$

Weiter gilt

$$\begin{aligned} \int_{\mathbb{R}} f(t, x)\psi(t, x)dx &= E(f(t, X_t)) = E\left(f(0, X_0) + \int_0^t df(\tau, X_\tau)\right) \\ &= \underbrace{E(f(0, X_0))}_{=\int_{\mathbb{R}} f(0, x)\psi(0, x)dx} + E\left(\int_0^t \left(\frac{\partial f}{\partial t} + a \frac{\partial f}{\partial x} + \frac{1}{2} b^2 \frac{\partial^2 f}{\partial x^2}\right) (\tau, X_\tau) d\tau\right) \\ &\quad + \underbrace{E\left(\int_0^t \left(b \frac{\partial f}{\partial x}\right) (\tau, X_\tau) dW_\tau\right)}_{\stackrel{(*)}{=} 0}. \end{aligned}$$

Dabei gilt $(*)$ aufgrund der Erwartungswerteigenschaft des stochastischen Integrals. Es folgt also

$$\begin{aligned} \int_{\mathbb{R}} f(t, x)\psi(t, x)dx &= \int_{\mathbb{R}} f(0, x)\psi(0, x)dx \\ &\quad + \int_{\mathbb{R}} \int_0^t \left(\frac{\partial f}{\partial t} + a \frac{\partial f}{\partial x} + \frac{1}{2} b^2 \frac{\partial^2 f}{\partial x^2} \right) (\tau, x) \cdot \psi(\tau, x) d\tau dx \quad (3.1) \end{aligned}$$

3.1 Die Fokker-Planck-Gleichung

für beliebige, glatte Funktionen f .

Die Dichte ψ ist somit schwache Lösung einer partiellen DGL (PDGL)!

3.1.2 Die 1-dimensionale Fokker-Planck-Gleichung

Sei ψ hinreichend glatt und die Testfunktionen f besitzen einen kompakten Träger. Dann lässt sich die schwache Formulierung der PDGL in eine starke umschreiben. Es gilt

$$\int_{\mathbb{R}} \int_0^t \left(\frac{\partial f}{\partial t} \cdot \psi \right) (\tau, x) d\tau dx = \int_{\mathbb{R}} [(f \cdot \psi)(\tau, x)]_{\tau=0}^{\tau=t} dx - \int_{\mathbb{R}} \int_0^t \left(f \cdot \frac{\partial \psi}{\partial t} \right) (\tau, x) d\tau dx.$$

Analog gilt

$$\begin{aligned} \int_0^t \int_{\mathbb{R}} \left(a \cdot \frac{\partial f}{\partial x} \cdot \psi \right) (\tau, x) dx d\tau &= - \int_0^t \int_{\mathbb{R}} \left(f \cdot \frac{\partial}{\partial x} (a \cdot \psi) \right) (\tau, x) dx d\tau, \\ \int_0^t \int_{\mathbb{R}} \left(\psi \cdot \frac{1}{2} b^2 \frac{\partial^2 f}{\partial x^2} \right) (\tau, x) dx d\tau &= \int_0^t \int_{\mathbb{R}} \left(f \cdot \frac{1}{2} \frac{\partial^2}{\partial x^2} (b^2 \cdot \psi) \right) (\tau, x) dx d\tau \end{aligned}$$

für alle glatten Testfunktionen f mit kompaktem Träger. Insgesamt erhalten wir aus (3.1)

$$0 = \int_0^t \int_{\mathbb{R}} \left(-\frac{\partial \psi}{\partial t} - \frac{\partial}{\partial x} (a\psi) + \frac{1}{2} \frac{\partial^2}{\partial x^2} (b^2 \psi) \right) (\tau, x) f(\tau, x) dx d\tau$$

für alle Testfunktionen f . Das Variationslemma impliziert dann

$$\frac{\partial \psi}{\partial t}(t, x) + \frac{\partial}{\partial x} (a\psi)(t, x) = \frac{1}{2} \frac{\partial^2}{\partial x^2} (b^2 \psi)(t, x), \quad 0 \leq t \leq T, \quad x \in \mathbb{R}.$$

Dies ist die „1-dimensionale Fokker-Planck-Gleichung“.

3.1 Bemerkung

Die Fokker-Planck-Gleichung beschreibt also die zeitliche Entwicklung der Dichtefunktion eines Itô-Prozesses, welcher der SDGL $dX_t = a(t, X_t)dt + b(t, X_t)dW_t$ genügt.

Alternative zur Berechnung von $E(g(X_t))$:

- Löse die Fokker-Planck-Gleichung

$$\begin{aligned} \frac{\partial \psi}{\partial t} + \frac{\partial}{\partial x} (a\psi) &= \frac{1}{2} \frac{\partial^2}{\partial x^2} (b^2 \psi), \\ \psi|_{t=0} &= \text{Dichte des Startwertes der SDGL} = \xi, \quad \xi \in \mathbb{R}. \end{aligned}$$

- Berechne das Integral

$$\int_{\mathbb{R}} g(x) \psi(t, x) dx.$$

3.1.3 Weiteres zu stochastischen Prozessen und PDGL

Im Folgenden soll es um weitere Zusammenhänge zwischen stochastischen Prozessen und PDGL gehen.

Sei $\tilde{X}_s^{t,\xi}$ die Lösung der SDGL

$$dX_s = a(s, X_s)ds + b(s, X_s)dW_s, \quad 0 \leq s \leq T$$

mit der Anfangsbedingung

$$X_t = \xi, \quad 0 \leq t \leq s \leq T,$$

wobei $\xi \in \mathbb{R}$ eine konstante ZV ist. Setze

$$X_s := \tilde{X}_s^{0,\xi}, \quad 0 \leq s \leq T.$$

Dann liefert die Itô-Formel

$$\begin{aligned} f\left(s, \tilde{X}_s^{t,\xi}\right) &= f\left(t, \underbrace{\tilde{X}_t^{t,\xi}}_{=\xi}\right) + \int_t^s \left(\frac{\partial f}{\partial t} + a \frac{\partial f}{\partial x} + \frac{1}{2} b^2 \frac{\partial^2 f}{\partial x^2} \right) (\tau, \tilde{X}_\tau^{t,\xi}) d\tau \\ &\quad + \int_t^s \left(b \frac{\partial f}{\partial x} \right) (\tau, \tilde{X}_\tau^{t,\xi}) dW_\tau, \end{aligned}$$

d. h. wie oben folgt

$$E\left(f\left(s, \tilde{X}_s^{t,\xi}\right)\right) = \underbrace{E\left(f(t, \xi)\right)}_{=f(t,\xi)} + E\left(\int_t^s \left(\frac{\partial f}{\partial t} + a \frac{\partial f}{\partial x} + \frac{1}{2} b^2 \frac{\partial^2 f}{\partial x^2} \right) (\tau, \tilde{X}_\tau^{t,\xi}) d\tau\right),$$

und man kann Folgendes einsehen:

Ist f eine Lösung des Endwertproblems

$$\begin{aligned} \frac{\partial f}{\partial t} + a \frac{\partial f}{\partial x} + \frac{1}{2} b^2 \frac{\partial^2 f}{\partial x^2} &= 0, \\ f(T, x) &= g(x), \end{aligned} \tag{3.2}$$

so folgt mit $s = T$

$$E\left(\underbrace{f\left(T, \tilde{X}_T^{t,\xi}\right)}_{=g(\tilde{X}_T^{t,\xi})}\right) = f(t, \xi) + 0 = f(t, \xi),$$

d. h.

$$E\left(g\left(\tilde{X}_T^{t,\xi}\right)\right) = f(t, \xi).$$

3.2 Bemerkung

Dies ist ein Spezialfall der sogenannten Feynman-Kac-Formel. Das Lösen einer bestimmten Klasse von PDGL kann also auf das Lösen von SDGL zurückgeführt werden und umgekehrt.

Dies liefert eine weitere Lösungsmethode zu Bestimmung von $E(g(X_T))$:

- Löse das Endwertproblem (3.2) rückwärts in der Zeit.
- Bestimme $f(0, \xi) = E\left(g\left(\tilde{X}_T^{0,\xi}\right)\right) = E(g(X_T))$.

3.1.4 Verallgemeinerung auf vektorwertige Prozesse

Betrachte das System von SDGL

$$dX_t = \underbrace{a(t, X_t)}_{\in \mathbb{R}^d} dt + \underbrace{b(t, X_t)}_{\in \mathbb{R}^{d \times m}} dW_t, \quad 0 \leq t \leq T,$$

$$X_0 = \xi.$$

Dabei ist $W_t = (W_t^{(1)}, \dots, W_t^{(m)}) \in \mathbb{R}^m$ eine m -dimensionale Brownsche Bewegung.

Die Dichte von $X_t = (X_t^1, \dots, X_t^d)$, $0 \leq t \leq T$ sei durch

$$\psi(t, x) \in \mathbb{R}, \quad x \in \mathbb{R}^d$$

gegeben, d. h. für den Erwartungswert gilt

$$E(g(X_t)) = \int_{\mathbb{R}^d} g(x) \psi(t, x) dx.$$

Mit den im 1-dimensionalen Fall eingeführten Techniken erhält man, dass ψ die d -dimensionale Fokker-Planck-Gleichung

$$\frac{\partial \psi}{\partial t}(t, x) + \sum_{i=1}^d \frac{\partial}{\partial x_i} (a_i \psi)(t, x) = \frac{1}{2} \sum_{i,k=1}^d \frac{\partial^2}{\partial x_i \partial x_k} (d_{ik} \psi)(t, x)$$

erfüllt. Dabei ist $d_{ik} := \sum_{j=1}^m b_{ij} b_{kj} = (bb^T)_{ik}$.

3.1.5 Beispiel: Optionsbewertung eines Capped-Symmetric-Power-Calls

Die Auszahlungsfunktion eines Calls ist in diesem Fall gegeben durch

$$C(S) = \min \{ ((S - K)^+)^p, L \}, \quad L, p > 0,$$

wobei S der Kurs des Basiswertes, K der Ausübungspreis und T die Verfallszeit sind.

Für $\sigma, r > 0$ genüge nun $S_\eta^{t,s}$ dem Preisprozess

$$dS_\eta = rS_\eta d\eta + \sigma S_\eta dW_\eta, \quad 0 \leq \eta \leq T,$$

$$S_t = s \in \mathbb{R} \quad (\text{konstante ZV}), \quad t \in [0, T].$$

Der Wert der Option zur Zeit t zu einem Wert des Basiswertes S ist gegeben durch

$$V(S, t) = E(\exp(-r(T-t)) \cdot C(S_T^{t,s})).$$

Ist also $f(t, s)$ eine Lösung des Endwertproblems

$$\frac{\partial f}{\partial t} + rs \frac{\partial f}{\partial s} + \frac{1}{2} \sigma^2 s^2 \frac{\partial^2 f}{\partial s^2} = 0,$$

$$f(T, s) = C(s), \tag{3.3}$$

3.1 Die Fokker-Planck-Gleichung

so gilt

$$f(t, s) = E \left(C \left(S_T^{t,s} \right) \right),$$

d. h.

$$V(s, t) = \exp(-r(T-t)) \cdot f(t, s). \quad \text{„Optionspreisbewertung mit PDGL“}$$

Umformulierung von (3.3) als Gleichung für den Optionswert V: Mit

$$\begin{aligned} \frac{\partial V}{\partial t} &= r \exp(-r(T-t)) f + \exp(-r(T-t)) \frac{\partial f}{\partial t} \\ &= \left(r f + \frac{\partial f}{\partial t} \right) \exp(-r(T-t)) \iff \\ \frac{\partial V}{\partial t} - rV &= \frac{\partial f}{\partial t} \exp(-r(T-t)) \end{aligned}$$

und

$$\frac{\partial^j V}{\partial S^j} = \exp(-r(T-t)) \frac{\partial^j f}{\partial S^j}, \quad j = 1, 2$$

folgt die „**Black-Scholes-Gleichung**“

$$\frac{\partial V}{\partial t} + rS \frac{\partial V}{\partial S} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} - rV = 0, \quad (3.4)$$

mit der Endbedingung

$$V(S, T) = \exp(-r(T-T)) f(T, S) = f(T, S) = C(S).$$

Als geeignete Randwerte findet man

$$V(0, t) = 0 \quad \text{für } 0 < t < T \quad (\text{Call wertlos, wenn } S = 0), \quad (3.5)$$

$$V(S, t) \rightarrow L \quad \text{für } S \rightarrow \infty, \quad 0 < t < T \quad (\text{„capped“}). \quad (3.6)$$

Statt L kann man in (3.6) auch den diskontierten Wert $L \exp(-r(t-T))$ vorschreiben.

Folgendes lässt sich zeigen:

V erfüllt die Black-Scholes-Gleichung (3.4), falls $u = u(x, \tau)$ die Diffusionsgleichung (Wärmeleitungsgleichung)

$$\frac{\partial u}{\partial \tau} = \frac{\partial^2 u}{\partial x^2}$$

löst und

$$V(S, t) = K \exp \left(\frac{1-q}{2} x - \frac{(1+q)^2}{4} \tau \right) u(x, \tau) \quad (3.7)$$

mit $x = \ln(S/K)$, $\tau = \sigma^2 \left(\frac{T-t}{2} \right)$, $q = \frac{2r}{\sigma^2}$.

Der Endwert für V übersetzt sich in einen Anfangswert für u , d. h. mit $t = T$ folgt

$$C(S) = V(S, T) = K \exp \left(\frac{1-q}{2} x \right) \underbrace{u(x, 0)}_{=u_0(x)}$$

und somit

$$u_0(x) = u(x, 0) = \frac{1}{K} C (K \exp(x)) \exp\left(\frac{q-1}{2}x\right).$$

Für die Randbedingungen ergeben sich (vgl. (3.5), (3.6) und (3.7))

$$\begin{aligned} \lim_{x \rightarrow -\infty} u(x, \tau) &= 0, \\ \lim_{x \rightarrow \infty} u(x, \tau) &= \lim_{x \rightarrow \infty} \frac{V(K \exp(x), T - \frac{2\tau}{\sigma^2})}{K \exp\left(\frac{1-q}{2}x - \frac{(1+q)^2}{4}\tau\right)} \\ &= \lim_{x \rightarrow \infty} \underbrace{\frac{V(K \exp(x), T - \frac{2\tau}{\sigma^2})}{L}}_{=1} \cdot \lim_{x \rightarrow \infty} \frac{L}{K} \exp\left(\frac{q-1}{2}x + \frac{(1+q)^2}{4}\tau\right) \\ &= \lim_{x \rightarrow \infty} g(x, \tau), \end{aligned}$$

$$\text{mit } g(x, \tau) := \frac{L}{K} \exp\left(\frac{q-1}{2}x + \frac{(1+q)^2}{4}\tau\right).$$

3.2 Finite-Differenzen-Verfahren für die Wärmeleitungsgleichung

3.2.1 Einführung

Vorgelegt sei die PDGL

$$\frac{\partial u}{\partial \tau} = \frac{\partial^2 u}{\partial x^2}, \quad (x, \tau) \in \Omega =]-\infty, \infty[\times [0, \tau_{\max}], \quad (3.8)$$

mit $\tau_{\max} = \frac{\sigma^2}{2}T$. Numerisch ist das Gebiet abzuschneiden, d.h. betrachte die Gleichung (3.8) für $(x, \tau) \in \Omega =]-a, a[\times [0, \tau_{\max}]$, wobei $a > 0$ hinreichend groß. Ergänze diese Aufgabe durch

$$u(x, 0) = u_0(x) = \frac{1}{K} C (K \exp(x)) \exp\left(\frac{q-1}{2}x\right), \quad -a \leq x \leq a \quad (3.9)$$

und die abgeschnittenen Randbedingungen

$$\begin{aligned} u(-a, \tau) &= 0, \\ u(a, \tau) &= g(a, \tau) \end{aligned} \quad (3.10)$$

für $0 \leq \tau \leq \frac{\sigma^2}{2}T$.

3.2.2 Das Prinzip der Linienmethode

Diskretisiere zuerst die Raumvariable und danach die Zeitvariable.

Sei $M \in \mathbb{N}$ und sei $\Delta x := \frac{a}{M}$ sowie $x_i := i\Delta x = \frac{ia}{M}$ für $i = -M, -M+1, \dots, M$. Für

3.2 Finite-Differenzen-Verfahren für die Wärmeleitungsgleichung

$\tau \in \left[0, \frac{\sigma^2}{2}T\right]$ und $x_i = i\Delta x, i \in \{-M+1, \dots, M-1\}$ fest, ersetzen (in Zeichen „ $\sim$ “) wir die partielle Ableitung nach x mit der zentralen Differenz, d. h.

$$\begin{aligned}\frac{\partial^2}{\partial x^2}u(x_i, \tau) &\sim \frac{\frac{\partial}{\partial x}u\left(x_i + \frac{\Delta x}{2}, \tau\right) - \frac{\partial}{\partial x}u\left(x_i - \frac{\Delta x}{2}, \tau\right)}{\Delta x} \\ &\sim \frac{\frac{u(x_i + \Delta x, \tau) - u(x_i, \tau)}{\Delta x} - \frac{u(x_i, \tau) - u(x_i - \Delta x, \tau)}{\Delta x}}{\Delta x} \\ &= \frac{1}{\Delta x^2} (u(x_{i+1}, \tau) - 2u(x_i, \tau) + u(x_{i-1}, \tau)).\end{aligned}$$

Dabei ersetzen wir im Fall $i = -M+1$ die Größe $u(-x_M, \tau)$ gemäß

$$u(-x_M, \tau) = u(-a, \tau) = 0$$

und im Fall $i = M-1$ die Größe $u(x_M, \tau)$ durch

$$u(x_M, \tau) = u(a, \tau) = g(a, \tau)$$

gemäß den Randbedingungen (3.10).

Sei nun

$$\begin{aligned}v(\tau) &:= (u(x_{-M+1}, \tau), u(x_{-M+2}, \tau), \dots, u(x_{M-1}, \tau)) \\ &= (v_1(\tau), v_2(\tau), \dots, v_{2M-1}(\tau)) \in \mathbb{R}^{2M-1}.\end{aligned}$$

Ersetze also den Ausdruck

$$\frac{\partial^2}{\partial x^2}u(x, \tau), \quad u(-a, \tau) = 0, \quad u(a, \tau) = g(a, \tau)$$

durch

$$-\frac{1}{\Delta x^2} \begin{pmatrix} 2 & -1 & 0 & \dots & 0 & 0 & 0 \\ -1 & 2 & -1 & \dots & 0 & 0 & 0 \\ 0 & -1 & 2 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 2 & -1 & 0 \\ 0 & 0 & 0 & \dots & -1 & 2 & -1 \\ 0 & 0 & 0 & \dots & 0 & -1 & 2 \end{pmatrix} \begin{pmatrix} v_1(\tau) \\ v_2(\tau) \\ v_3(\tau) \\ \vdots \\ v_{2M-3}(\tau) \\ v_{2M-2}(\tau) \\ v_{2M-1}(\tau) \end{pmatrix} + \frac{1}{\Delta x^2} \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ g(a, \tau) \end{pmatrix}.$$

Mit

$$v'(\tau) = \left(\frac{\partial u}{\partial \tau}(x_{-M+1}, \tau), \frac{\partial u}{\partial \tau}(x_{-M+2}, \tau), \dots, \frac{\partial u}{\partial \tau}(x_{M-1}, \tau) \right)$$

und

$$\frac{\partial u}{\partial \tau} = \frac{\partial^2 u}{\partial x^2}$$

folgt damit das System gewöhnlicher DGL

$$v'(\tau) = -A_{\Delta x}v(\tau) + r^{\Delta x}(\tau) =: F_{\Delta x}(v(\tau), \tau), \quad 0 \leq \tau \leq \frac{\sigma^2}{2}T, \quad (3.11)$$

mit der symmetrischen Tridiagonalmatrix

$$A_{\Delta x} := \frac{1}{\Delta x^2} \begin{pmatrix} 2 & -1 & 0 & \dots & 0 & 0 & 0 \\ -1 & 2 & -1 & \dots & 0 & 0 & 0 \\ 0 & -1 & 2 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 2 & -1 & 0 \\ 0 & 0 & 0 & \dots & -1 & 2 & -1 \\ 0 & 0 & 0 & \dots & 0 & -1 & 2 \end{pmatrix} \in \mathbb{R}^{(2M-1) \times (2M-1)} \quad (3.12)$$

und

$$r^{\Delta x}(\tau) := \frac{1}{\Delta x^2} (0, \dots, 0, g(a, \tau)). \quad (3.13)$$

Als Anfangsbedingung erhalten wir

$$\begin{aligned} v(0) &= (u(x_{-M+1}, 0), u(x_{-M+2}, 0), \dots, u(x_{M-1}, 0)) \\ &= (u_0(x_{-M+1}), u_0(x_{-M+2}), \dots, u_0(x_{M-1})) = v^0. \end{aligned} \quad (3.14)$$

3.2.3 Das Euler-Cauchy-Verfahren

Das Anfangswertproblem (3.11)-(3.14) heißt Liniensystem zur PDGL (3.8)-(3.10). Der Übergang von (3.8)-(3.10) zu (3.11)-(3.14) wird als Semidiskretisierung bezeichnet. Zu einer vollständigen Diskretisierung von (3.8)-(3.10) kommen wir, wenn wir auf die Anfangswertaufgabe (3.11)-(3.14) eines der bekannten numerischen Verfahren zur Lösung von Anfangswertaufgaben anwenden. Bei der Durchführung dieser Verfahren sollte man sich jedoch die spezielle Gestalt von $F_{\Delta x}$ zunutze machen, denn das System (3.11)-(3.14) wird für $\Delta x \rightarrow 0$ immer größer. Das einfachste Verfahren für (3.11)-(3.14) ist das Euler-Cauchy-Verfahren:

Seien $\Delta \tau := \frac{\sigma^2 T}{2N} > 0$, v^0 vorgegeben und v^j approximiere $v(j\Delta \tau)$ für $j = 0, 1, \dots, N$. Dann finden wir

$$\begin{aligned} \frac{v^{j+1} - v^j}{\Delta \tau} &= F_{\Delta x}(v^j, j\Delta \tau) = -A_{\Delta x} v^j + r^{\Delta x}(j\Delta \tau), \quad j = 0, 1, \dots, N-1, \\ v^0 &= (u_0(x_{-M+1}), u_0(x_{-M+2}), \dots, u_0(x_{M-1})) \in \mathbb{R}^{2M-1}. \end{aligned}$$

Explizit folgt mit

$$v^j = (u_{-M+1}^j, u_{-M+2}^j, \dots, u_{M-1}^j), \quad u_i^j \sim u(i\Delta x, j\Delta \tau)$$

die Iteration

$$\begin{aligned} \begin{pmatrix} u_{-M+1}^{j+1} \\ \vdots \\ u_i^{j+1} \\ \vdots \\ u_{M-1}^{j+1} \end{pmatrix} &= \begin{pmatrix} u_{-M+1}^j \\ \vdots \\ u_i^j \\ \vdots \\ u_{M-1}^j \end{pmatrix} + \Delta \tau \left\{ \frac{1}{\Delta x^2} \begin{pmatrix} -2u_{-M+1}^j + u_{-M+2}^j \\ \vdots \\ u_{i-1}^j - 2u_i^j + u_{i+1}^j \\ \vdots \\ u_{M-2}^j - 2u_{M-1}^j \end{pmatrix} \right. \\ &\quad \left. + \frac{1}{\Delta x^2} \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ g(a, j\Delta \tau) \end{pmatrix} \right\}, \quad j = 0, 1, \dots, N-1, \end{aligned} \quad (3.15)$$

3.2 Finite-Differenzen-Verfahren für die Wärmeleitungsgleichung

mit der Anfangsbedingung

$$\begin{pmatrix} u_{-M+1}^0 \\ \vdots \\ u_i^0 \\ \vdots \\ u_{M-1}^0 \end{pmatrix} = \begin{pmatrix} u_0(x_{-M+1}) \\ \vdots \\ u_0(x_i) \\ \vdots \\ u_0(x_{M-1}) \end{pmatrix}. \quad (3.16)$$

Das System (3.15) und (3.16) heißt auch explizites Differenzenverfahren zur Aufgabe (3.8)-(3.10). Man erhält u_i^{j+1} aus u_{i-1}^j , u_i^j und u_{i+1}^j , was man durch das Schema des „Differenzensterns“ andeutet, vgl. Abbildung 3.1.

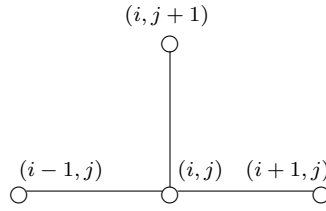


Abbildung 3.1: Differenzenstern für das Euler-Cauchy-Verfahren.

3.2.4 Das ϑ -Verfahren

Wir betrachten als nächstes für das Liniensystem (3.11)-(3.14) das von einem Parameter $\vartheta \in [0, 1]$ abhängige Verfahren, das sogenannte ϑ -Verfahren:

$$\frac{v^{j+1} - v^j}{\Delta\tau} = \vartheta F_{\Delta x}(v^{j+1}, (j+1)\Delta\tau) + (1 - \vartheta) F_{\Delta x}(v^j, j\Delta\tau), \quad j = 0, 1, \dots, N-1, \quad (3.17)$$

$$v^0 = (u_0(x_{-M+1}), u_0(x_{-M+2}), \dots, u_0(x_{M-1})).$$

Spezialfälle des ϑ -Verfahrens für das Liniensystem:

- $\vartheta = 0$: Euler-Cauchy-Verfahren (ECV)
- $\vartheta = \frac{1}{2}$: Trapez-Methode oder das sogenannte „Crank-Nicolson-Verfahren“ (CNV)
- $\vartheta = 1$: Implizites Euler-Cauchy-Verfahren (IECV)

Für $\vartheta > 0$ erfordert (3.17) wegen $F_{\Delta x}(v, \tau) = -A_{\Delta x}v + r^{\Delta x}(\tau)$ die Auflösung eines linearen Gleichungssystems und das Verfahren heißt daher implizit. Die Abbildung 3.2 stellt die zugehörigen Differenzensterne dar.

3.2.5 Auflösung der Gleichungssysteme auf jedem Zeitlevel τ_j

Mit $F_{\Delta x}(v, \tau) = -A_{\Delta x}v + r^{\Delta x}(\tau)$ schreibt sich (3.17) um zu

$$\frac{v^{j+1} - v^j}{\Delta\tau} = \vartheta (-A_{\Delta x}v^{j+1} + r^{\Delta x}((j+1)\Delta\tau)) + (1 - \vartheta) (-A_{\Delta x}v^j + r^{\Delta x}(j\Delta\tau)),$$

$$j = 0, 1, \dots, N-1,$$

3.2 Finite-Differenzen-Verfahren für die Wärmeleitungsgleichung

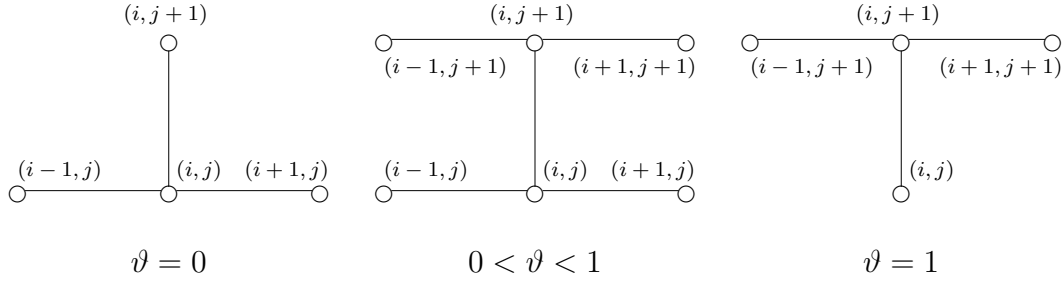


Abbildung 3.2: Differenzensterne für $\vartheta \in [0, 1]$.

bzw. äquivalent dazu

$$\left(\frac{1}{\Delta\tau} I + \vartheta A_{\Delta x} \right) v^{j+1} = \left(\frac{1}{\Delta\tau} I - (1 - \vartheta) A_{\Delta x} \right) v^j + \vartheta r^{\Delta x} ((j+1)\Delta\tau) + (1 - \vartheta) r^{\Delta x} (j\Delta\tau),$$

$$j = 0, 1, \dots, N-1. \quad (3.18)$$

In jedem Schritt ist also für $\vartheta > 0$ ein tridiagonales Gleichungssystem mit der Matrix

$$C_{\Delta x, \Delta\tau} := \left(\frac{1}{\Delta\tau} I + \vartheta A_{\Delta x} \right)$$

$$= \begin{pmatrix} \frac{1}{\Delta\tau} + \frac{2\vartheta}{\Delta x^2} & \frac{-\vartheta}{\Delta x^2} & 0 & \dots & 0 & 0 & 0 \\ \frac{-\vartheta}{\Delta x^2} & \frac{1}{\Delta\tau} + \frac{2\vartheta}{\Delta x^2} & \frac{-\vartheta}{\Delta x^2} & \dots & 0 & 0 & 0 \\ 0 & \frac{-\vartheta}{\Delta x^2} & \frac{1}{\Delta\tau} + \frac{2\vartheta}{\Delta x^2} & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \frac{1}{\Delta\tau} + \frac{2\vartheta}{\Delta x^2} & \frac{-\vartheta}{\Delta x^2} & 0 \\ 0 & 0 & 0 & \dots & \frac{-\vartheta}{\Delta x^2} & \frac{1}{\Delta\tau} + \frac{2\vartheta}{\Delta x^2} & \frac{-\vartheta}{\Delta x^2} \\ 0 & 0 & 0 & \dots & 0 & \frac{-\vartheta}{\Delta x^2} & \frac{1}{\Delta\tau} + \frac{2\vartheta}{\Delta x^2} \end{pmatrix} \quad (3.19)$$

zu lösen, wobei $C_{\Delta x, \Delta\tau} \in \mathbb{R}^{(2M-1) \times (2M-1)}$.

3.3 Definition

Eine Matrix $D \in \mathbb{R}^{m \times m}$ heißt strikt diagonal-dominant, falls

$$|d_{ii}| > \sum_{\substack{j=1 \\ j \neq i}}^m |d_{ij}|, \quad i = 1, \dots, m.$$

3.4 Lemma

Sei $D \in \mathbb{R}^{m \times m}$ strikt diagonal-dominant. Dann ist D invertierbar.

Beweis. Annahme: D sei nicht invertierbar. Dann existiert ein $x \in \mathbb{R}^m \setminus \{0\}$ mit $Dx = 0$. Ohne Einschränkung sei

$$x_r = \max \{|x_i| : i = 1, \dots, m\} = 1$$

3.2 Finite-Differenzen-Verfahren für die Wärmeleitungsgleichung

für ein $r \in \{1, \dots, m\}$. Aus

$$0 = (Dx)_r \iff d_{rr}x_r = - \sum_{\substack{j=1 \\ j \neq r}}^m d_{rj}x_j$$

folgt mit $|x_r| = 1$ sofort

$$|d_{rr}| = |d_{rr}| |x_r| \leq \sum_{\substack{j=1 \\ j \neq r}}^m |d_{rj}| \underbrace{|x_j|}_{\leq 1} \leq \sum_{\substack{j=1 \\ j \neq r}}^m |d_{rj}|,$$

Widerspruch. □

Die Matrix $C_{\Delta x, \Delta \tau}$ aus (3.19) ist diagonal-dominant, d.h. nach Lemma 3.4 ist $C_{\Delta x, \Delta \tau}$ invertierbar und das ϑ -Verfahren (3.18) ist somit durchführbar.

3.2.6 Praktische Lösung des Gleichungssystems

Auf jedem Zeitlevel $j\Delta\tau, j = 0, \dots, N-1$ kann das Lösen des Gleichungssystems z.B. mit dem LU-Verfahren für Tridiagonalmatrizen erfolgen. Dazu machen wir folgende Beobachtung:

$$\begin{pmatrix} 1 & 0 & 0 & \dots & 0 & 0 & 0 \\ l_1 & 1 & 0 & \dots & 0 & 0 & 0 \\ 0 & l_2 & 1 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 & 0 \\ 0 & 0 & 0 & \dots & l_{m-2} & 1 & 0 \\ 0 & 0 & 0 & \dots & 0 & l_{m-1} & 1 \end{pmatrix} \cdot \begin{pmatrix} d_1 & r_1 & 0 & \dots & 0 & 0 & 0 \\ 0 & d_2 & r_2 & \dots & 0 & 0 & 0 \\ 0 & 0 & d_3 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & d_{m-2} & r_{m-2} & 0 \\ 0 & 0 & 0 & \dots & 0 & d_{m-1} & r_{m-1} \\ 0 & 0 & 0 & \dots & 0 & 0 & d_m \end{pmatrix} =$$

$$\begin{pmatrix} d_1 & r_1 & 0 & \dots & 0 & 0 & 0 \\ d_1 l_1 & r_1 l_1 + d_2 & r_2 & \dots & 0 & 0 & 0 \\ 0 & d_2 l_2 & r_2 l_2 + d_3 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & r_{m-3} l_{m-3} + d_{m-2} & r_{m-2} & 0 \\ 0 & 0 & 0 & \dots & d_{m-2} l_{m-2} & r_{m-2} l_{m-2} + d_{m-1} & r_{m-1} \\ 0 & 0 & 0 & \dots & 0 & d_{m-1} l_{m-1} & r_{m-1} l_{m-1} + d_m \end{pmatrix}$$

$$=: (c_{ij})_{1 \leq i, j \leq m} = C.$$

Koeffizientenvergleich liefert

$$d_1 = c_{11}, \quad r_k = c_{k, k+1}, \quad k = 1, \dots, m-1$$

und

$$l_{k-1} = \frac{c_{k, k-1}}{d_{k-1}}, \quad d_k = c_{kk} - r_{k-1} l_{k-1}, \quad k = 2, \dots, m.$$

3.5 Bemerkung

- a) Es lässt sich zeigen, dass der LU-Tridiagonalalgorithmus für strikt diagonal-dominante Matrizen immer durchführbar ist.
- b) In Matlab erhält man die LU-Zerlegung einer Matrix C mit $[L, U] = \text{lu}(C)$.
- c) Das Gleichungssystem $Cx = b$ kann dann als $LUx = b$ geschrieben werden und in zwei Schritten aufgelöst werden:
 - (i) „Vorwärts auflösen“:
 $Ly = b$, d. h. $y_1 = b_1$ und $y_i = b_i - l_{i-1}y_{i-1}$, $i = 2, \dots, m$.
 - (ii) „Rückwärts auflösen“:
 $Ux = y$, d. h. $x_m = \frac{y_m}{d_m}$ und $x_i = \frac{y_i - r_i x_{i+1}}{d_i}$, $i = m-1, \dots, 1$.
- d) In Matlab lautet dies $y = L \backslash b$ und $x = U \backslash y$. Man beachte dabei, dass man Gleichungssysteme lediglich löst, aber niemals eine Matrix invertiert.

3.3 Konsistenz, Stabilität und Konvergenz

3.3.1 Einführung

Wir möchten in diesem Paragraphen beschreiben mit welcher Güte die Approximationen w_i^j des ϑ -Verfahrens die Lösung $\bar{u}(i\Delta x, j\Delta\tau)$ der Wärmeleitungsgleichung approximieren. Wir schreiben dazu das ϑ -Verfahren nun als Nullstellenproblem.

Sei $h = (\Delta x, \Delta\tau)$, $\Delta x = \frac{a}{M} > 0$ und $\Delta\tau = \frac{\sigma^2 T}{2N} > 0$. Ersetze das Gebiet

$$\Omega =]-a, a[\times \left[0, \frac{\sigma^2 T}{2} \right]$$

durch das Gitter

$$\Omega_h := \{(i\Delta x, j\Delta\tau) : i = -M+1, \dots, M-1, j = 0, \dots, N\}.$$

Ein Element

$$w = (w_{-M+1}^0, \dots, w_{M-1}^0, w_{-M+1}^1, \dots, w_{M-1}^1, \dots, w_{-M+1}^N, \dots, w_{M-1}^N) \in \mathbb{R}^{\Omega_h}$$

heißt Gitterfunktion. Wir versehen den Raum $\mathbb{R}^{\Omega_h}$ mit der Maximumsnorm, d. h.

$$\|w\|_\infty = \max \{|w_i^j| : i = -M+1, \dots, M-1, j = 0, \dots, N\}.$$

Wir schreiben nun das ϑ -Verfahren als Nullstellenproblem

$$T^h(u) = 0$$

mit dem Operator

$$T^h : \mathbb{R}^{\Omega_h} \rightarrow \mathbb{R}^{\Omega_h}$$

3.3 Konsistenz, Stabilität und Konvergenz

wie folgt:

$$(T^h(u))_i^j := \begin{cases} u_i^0 - u_0(i\Delta x), & i = -M+1, \dots, M-1, j=0, \\ \frac{1}{\Delta\tau} (u_i^j - u_i^{j-1}) - \vartheta \frac{1}{\Delta x^2} (u_{i-1}^j - 2u_i^j + u_{i+1}^j) \\ \quad - (1-\vartheta) \frac{1}{\Delta x^2} (u_{i-1}^{j-1} - 2u_i^{j-1} + u_{i+1}^{j-1}), & i = -M+1, \dots, M-1, j=1, \dots, N, \end{cases} \quad (3.20)$$

wobei $T^h(u) \in \mathbb{R}^{\Omega_h}$. In der Definition (3.20) ist dabei

$$u_{-M}^j = 0, \quad u_{-M}^{j-1} = 0, \quad u_M^j = g(a, j\Delta\tau), \quad u_M^{j-1} = g(a, (j-1)\Delta\tau)$$

zu setzen.

Sei $\bar{u} \in C^2(\bar{\Omega})$ klassische Lösung von

$$\begin{aligned} \frac{\partial u}{\partial \tau} &= \frac{\partial^2 u}{\partial x^2} \quad \text{in } \Omega =]-a, a[\times \left[0, \frac{\sigma^2}{2}T\right], \\ u(x, 0) &= u_0(x), \quad -a \leq x \leq a, \\ u(-a, \tau) &= 0, \quad u(a, \tau) = g(a, \tau), \quad \tau \in \left[0, \frac{\sigma^2}{2}T\right] \end{aligned} \quad (3.21)$$

und sei

$$\bar{u}_h := \bar{u}|_{\Omega_h}$$

die Einschränkung von $\bar{u}$ auf das Gitter Ω_h . Ferner sei u^h die Lösung von $T^h(u^h) = 0$, d. h. die Approximation von $\bar{u}$ mit dem ϑ -Verfahren mit der Schrittweite $h = (\Delta x, \Delta\tau)$.

3.6 Definition

a) Der Fehler des numerischen Verfahrens ist definiert durch

$$\begin{aligned} e(h) &:= \|\bar{u}_h - u^h\|_\infty \\ &= \max \left\{ \left| \bar{u}_h(i\Delta x, j\Delta\tau) - (u^h)_i^j \right| : i = -M+1, \dots, M-1, j = 0, \dots, N \right\}. \end{aligned}$$

Das numerische Verfahren heißt konvergent (der Ordnung p), falls $e(h) \rightarrow 0$ für $h \rightarrow 0$ ($e(h) = O(h^p)$).

b) Der Term $\|T^h(\bar{u}_h)\|_\infty$ heißt Konsistenzfehler. Ein Verfahren heißt konsistent (der Ordnung p), falls $\|T^h(\bar{u}_h)\|_\infty \rightarrow 0$ für $h \rightarrow 0$ ($\|T^h(\bar{u}_h)\|_\infty = O(h^p)$).

c) Das numerische Verfahren $T^h(u) = 0$ heißt stabil, falls mit einem $h_0 > 0$ die Gleichung für $0 < h \leq h_0$ eine Lösung u^h besitzt und falls ein $C > 0$ existiert mit

$$\|w - v\|_\infty \leq C \|T^h(w) - T^h(v)\|_\infty,$$

für alle $v, w \in \mathbb{R}^{\Omega_h}$, $0 < h \leq h_0$.

3.7 Korollar

Ist ein numerisches Verfahren stabil und konsistent (der Ordnung p), so ist es auch konvergent (der Ordnung p).

3.3 Konsistenz, Stabilität und Konvergenz

Beweis. Wende die Stabilitätsungleichung an mit $w = \bar{u}_h$ und $v = u^h$. Dies liefert

$$\begin{aligned} e(h) = \|\bar{u}_h - u^h\|_\infty &\leq C \|T(\bar{u}_h) - \underbrace{T^h(u^h)}_{=0}\|_\infty \\ &= C \|T^h(\bar{u}_h)\|_\infty \rightarrow 0 \quad \text{für } h \rightarrow 0 \end{aligned}$$

bzw.

$$e(h) = \|\bar{u}_h - u^h\|_\infty \leq C \|T^h(\bar{u}_h)\|_\infty = O(h^p).$$

□

Slogan:

Konsistenz + Stabilität $\implies$ Konvergenz

3.3.2 Der Konsistenzfehler des ϑ -Verfahrens

Sei $\bar{u} \in C^2(\bar{\Omega})$ klassische Lösung und $\bar{u}_h = \bar{u}|_{\Omega_h}$.

Für $j = 0$ und $i = -M + 1, \dots, M - 1$ gilt

$$(T^h(\bar{u}_h))_i^0 = \bar{u}_i^0 - u_0(i\Delta x) = 0.$$

Für $j = 1, \dots, N$ und $i = -M + 1, \dots, M - 1$ erhalten wir

$$\begin{aligned} (T^h(\bar{u}_h))_i^j &= \frac{1}{\Delta\tau} (\bar{u}_i^j - \bar{u}_i^{j-1}) - \vartheta \frac{1}{\Delta x^2} (\bar{u}_{i-1}^j - 2\bar{u}_i^j + \bar{u}_{i+1}^j) \\ &\quad - (1 - \vartheta) \frac{1}{\Delta x^2} (\bar{u}_{i-1}^{j-1} - 2\bar{u}_i^{j-1} + \bar{u}_{i+1}^{j-1}) \\ &= \frac{1}{\Delta\tau} (\bar{u}_i^j - \bar{u}_i^{j-1}) - \vartheta (\bar{u}_\tau)_i^j - \vartheta \left(\frac{1}{\Delta x^2} (\bar{u}_{i-1}^j - 2\bar{u}_i^j + \bar{u}_{i+1}^j) - (\bar{u}_{xx})_i^j \right) \\ &\quad - (1 - \vartheta) (\bar{u}_\tau)_i^{j-1} - (1 - \vartheta) \left(\frac{1}{\Delta x^2} (\bar{u}_{i-1}^{j-1} - 2\bar{u}_i^{j-1} + \bar{u}_{i+1}^{j-1}) - (\bar{u}_{xx})_i^{j-1} \right) \\ &= \frac{1}{\Delta\tau} (\bar{u}_i^j - \bar{u}_i^{j-1}) - \vartheta (\bar{u}_\tau)_i^j - (1 - \vartheta) (\bar{u}_\tau)_i^{j-1} \\ &\quad + \vartheta \left[(\bar{u}_{xx})_i^j - \frac{1}{\Delta x^2} (\bar{u}_{i-1}^j - 2\bar{u}_i^j + \bar{u}_{i+1}^j) \right] \\ &\quad + (1 - \vartheta) \left[(\bar{u}_{xx})_i^{j-1} - \frac{1}{\Delta x^2} (\bar{u}_{i-1}^{j-1} - 2\bar{u}_i^{j-1} + \bar{u}_{i+1}^{j-1}) \right] \\ &=: (I) + \vartheta(II) + (1 - \vartheta)(III). \end{aligned}$$

Für die Terme (I)-(III) gilt dabei

$$\begin{aligned} (I) &= \Delta\tau \left(\frac{1}{2} - \vartheta \right) (\bar{u}_{\tau\tau})_i^{j-1} + O(\Delta\tau^2), \\ (II) &= O(\Delta x^2), \\ (III) &= O(\Delta x^2), \end{aligned}$$

3.3 Konsistenz, Stabilität und Konvergenz

falls $\bar{u} \in C^4(\bar{\Omega})$. Somit erhalten wir insgesamt

$$(T^h(\bar{u}_h))_i^j = \Delta\tau \left(\frac{1}{2} - \vartheta \right) (\bar{u}_{\tau\tau})_i^{j-1} + O(\Delta\tau^2) + O(\Delta x^2),$$

falls $\bar{u} \in C^4(\bar{\Omega})$. Also gilt für den Konsistenzfehler des ϑ -Verfahrens

$$\|T^h(\bar{u}_h)\|_\infty = \begin{cases} O(\Delta\tau + \Delta x^2) & \text{für } \vartheta \neq \frac{1}{2}, \\ O(\Delta\tau^2 + \Delta x^2) & \text{für } \vartheta = \frac{1}{2}, \end{cases}$$

falls $\bar{u} \in C^4(\bar{\Omega})$.

3.3.3 Stabilitätsanalyse für das ϑ -Verfahren

Sei u_i^j die Approximation für $\bar{u}(i\Delta x, j\Delta\tau)$. Setze wieder $v^j = (u_{-M+1}^j, \dots, u_{M-1}^j) \in \mathbb{R}^{2M-1}$. Dann lautet das ϑ -Verfahren (3.18) für $j = 0, \dots, N-1$

$$C_{\Delta x, \Delta\tau} v^{j+1} = D_{\Delta x, \Delta\tau} v^j + \hat{r}^{j+1}, \quad (3.22)$$

wobei $C_{\Delta x, \Delta\tau}$ wie in (3.19) definiert ist und

$$\begin{aligned} D_{\Delta x, \Delta\tau} &:= \frac{1}{\Delta\tau} I - (1 - \vartheta) A_{\Delta x}, \\ \hat{r}^{j+1} &:= \vartheta r^{\Delta x}((j+1)\Delta\tau) + (1 - \vartheta) r^{\Delta x}(j\Delta\tau). \end{aligned}$$

Die Anfangsbedingung ist gegeben durch

$$\hat{r}^0 := v^0 = (u_0((-M+1)\Delta x), \dots, u_0((M-1)\Delta x)).$$

In Matrixform lautet (3.22) damit

$$\begin{pmatrix} I & 0 & 0 & \dots & 0 & 0 & 0 \\ -D_{\Delta x, \Delta\tau} & C_{\Delta x, \Delta\tau} & 0 & \dots & 0 & 0 & 0 \\ 0 & -D_{\Delta x, \Delta\tau} & C_{\Delta x, \Delta\tau} & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & C_{\Delta x, \Delta\tau} & 0 & 0 \\ 0 & 0 & 0 & \dots & -D_{\Delta x, \Delta\tau} & C_{\Delta x, \Delta\tau} & 0 \\ 0 & 0 & 0 & \dots & 0 & -D_{\Delta x, \Delta\tau} & C_{\Delta x, \Delta\tau} \end{pmatrix} \begin{pmatrix} v^0 \\ v^1 \\ v^2 \\ \vdots \\ v^{N-2} \\ v^{N-1} \\ v^N \end{pmatrix} - \begin{pmatrix} \hat{r}^0 \\ \hat{r}^1 \\ \hat{r}^2 \\ \vdots \\ \hat{r}^{N-2} \\ \hat{r}^{N-1} \\ \hat{r}^N \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 0 \end{pmatrix}. \quad (3.23)$$

Wir müssen also die Stabilität von Dreiecksblockmatrizen der Form wie in (3.23) untersuchen.

3.3 Konsistenz, Stabilität und Konvergenz

Es lässt sich zeigen, dass Folgendes hinreichend für Stabilität ist:

$$\begin{aligned} \|C_{\Delta x, \Delta \tau}^{-1}\|_{\infty} &= \Delta \tau, \\ \| (C_{\Delta x, \Delta \tau}^{-1}) (D_{\Delta x, \Delta \tau}) \|_{\infty} &\leq 1. \end{aligned} \quad (3.24)$$

Dabei ist für eine Matrix $E \in \mathbb{R}^{m \times m}$

$$\|E\|_{\infty} = \sup \{ \|Ex\|_{\infty} : x \in \mathbb{R}^m, \|x\|_{\infty} = 1 \}$$

die der $\|\cdot\|_{\infty}$ -Norm zugeordnete Matrixnorm. Die Abschätzungen in (3.24) lassen sich zeigen, falls

$$(D_{\Delta x, \Delta \tau})_{kl} \geq 0, \quad -M + 1 \leq k, l \leq M - 1. \quad (3.25)$$

Wegen $(D_{\Delta x, \Delta \tau})_{kl} \geq 0$ für $k \neq l$ ist (3.25) gleichbedeutend zu

$$\begin{aligned} (D_{\Delta x, \Delta \tau})_{kk} &= \frac{1}{\Delta \tau} - (1 - \vartheta) \frac{2}{\Delta x^2} \geq 0 && \Longleftrightarrow \\ 1 &\geq 2(1 - \vartheta) \frac{\Delta \tau}{\Delta x^2} && \Longleftrightarrow \\ \frac{\Delta \tau}{\Delta x^2} &\leq \frac{1}{2(1 - \vartheta)}. \end{aligned}$$

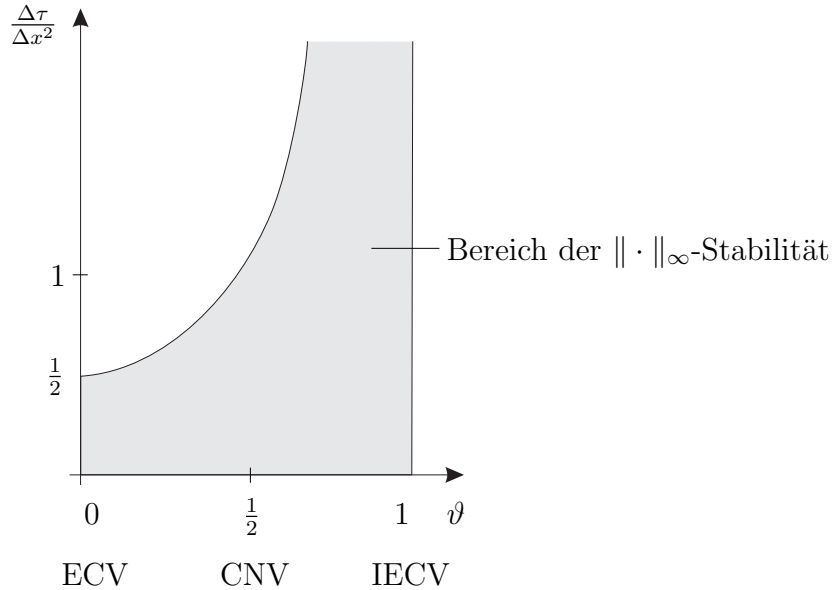


Abbildung 3.3: Grafische Interpretation der $\|\cdot\|_{\infty}$ -Stabilität.

Damit erhalten wir folgenden Stabilitätssatz.

3.8 Satz

Unter der Bedingung

$$\frac{\Delta \tau}{\Delta x^2} \leq \frac{1}{2(1 - \vartheta)} \quad (3.26)$$

ist das ϑ -Verfahren für die Wärmeleitungsgleichung bezüglich der $\|\cdot\|_{\infty}$ -Norm auf $\mathbb{R}^{\Omega_h}$ stabil.

3.3 Konsistenz, Stabilität und Konvergenz

Die Stabilitätsbedingung (3.26) für das ϑ -Verfahren lässt sich grafisch interpretieren, in dem man die Größen ϑ und $\frac{\Delta\tau}{\Delta x^2}$ in ein Schaubild zeichnet, vgl. Abbildung 3.3.

3.9 Korollar

Unter der Voraussetzung (3.26) ist das ϑ -Verfahren, $\vartheta \in [0, 1]$, bezüglich der Maximumsnorm konvergent der Ordnung $O(\Delta\tau + \Delta x^2)$ für $\vartheta \neq \frac{1}{2}$ bzw. $O(\Delta\tau^2 + \Delta x^2)$ für $\vartheta = \frac{1}{2}$ an jeder Lösung $\bar{u} \in C^4(\bar{\Omega})$ der Wärmeleitungsgleichung.