

# Stochastische Analysis

Martin Keller-Ressel

23. März 2015

## Inhaltsverzeichnis

|          |  |           |
|----------|--|-----------|
| <b>1</b> | <b>Stochastische Prozesse in stetiger Zeit</b>                             | <b>1</b>  |
| 1.1      | Grundlagen . . . . .   | 1         |
| 1.2      | Gleichheit von stochastischen Prozessen . . . . .                          | 2         |
| 1.3      | Filtrationen . . . . .   | 4         |
| 1.4      | Martingale . . . . .   | 5         |
| 1.5      | Stoppzeiten . . . . .  | 7         |
| 1.6      | Gleichgradige Integrierbarkeit . . . . .                                   | 10        |
| <b>2</b> | <b>Martingale in stetiger Zeit</b>   | <b>13</b> |
| 2.1      | Martingale in diskreter Zeit . . . . .                                     | 13        |
| 2.2      | Optionales Sampling und Optionales Stoppen in stetiger Zeit . . . . .      | 14        |
| 2.3      | Doob'sche Ungleichungen und Martingalkonvergenz in stetiger Zeit . . . . . | 16        |
| 2.4      | Martingale und Brownsche Bewegung . . . . .                                | 19        |
| <b>3</b> | <b>Stochastische Integration</b>   | <b>23</b> |
| 3.1      | Probleme der Pfadweisen Integration . . . . .                              | 24        |
| 3.2      | Ito-Isometrie & Ito-Integral . . . . .                                     | 27        |
| 3.3      | Das Ito-Integral als stetiges Martingal . . . . .                          | 29        |
| 3.4      | Stoppen unter dem Integralzeichen . . . . .                                | 32        |
| <b>4</b> | <b>Lokalisierung und lokale Martingale</b>                                 | <b>35</b> |
| 4.1      | Das Ito-Integral auf $\mathcal{L}_{loc}^2$ . . . . .                       | 35        |
| 4.2      | Riemann-Darstellung und Gaussche Prozesse . . . . .                        | 37        |
| 4.3      | Lokale Martingale . . . . .  | 39        |

|  |            |
|--|------------|
| <b>5 Die Ito-Formel</b>  | <b>42</b>  |
| 5.1 Herleitung der Ito-Formel . . . . .                        | 42         |
| 5.2 Erste Folgerungen und Anwendungen . . . . .                | 45         |
| 5.3 Die multivariate Ito-Formel und Anwendungen . . . . .      | 48         |
| 5.4 Ito-Prozesse . . . . .                                     | 52         |
| 5.5 Die quadratische Variation von Ito-Prozessen . . . . .     | 55         |
| <b>6 Stochastische Differentialgleichungen</b>                 | <b>60</b>  |
| 6.1 Beispiele Stochastischer Differentialgleichungen . . . . . | 60         |
| 6.2 Existenz und Eindeutigkeit von Lösungen . . . . .          | 65         |
| 6.3 Numerische Lösungsmethoden . . . . .                       | 72         |
| <b>7 Der Satz von Girsanov</b>                                 | <b>75</b>  |
| 7.1 Importance Sampling und Exponentiells Kippen . . . . .     | 75         |
| 7.2 Maßwechsel auf dem Pfadraum $C[0, T]$ . . . . .            | 77         |
| 7.3 Der Satz von Girsanov . . . . .                            | 80         |
| 7.4 Die Novikov-Bedingung . . . . .                            | 84         |
| <b>8 Martingaldarstellung und Doob-Meyer Zerlegung</b>         | <b>86</b>  |
| 8.1 Der Martingaldarstellungssatz . . . . .                    | 86         |
| 8.2 Doob-Meyer-Zerlegung . . . . .                             | 90         |
| <b>9 Die Diffusionsgleichung</b>                               | <b>93</b>  |
| 9.1 Herleitung der Diffusionsgleichung . . . . .               | 93         |
| 9.2 Lösen der Diffusionsgleichung . . . . .                    | 95         |
| 9.3 Eindeutigkeit der Lösung . . . . .                         | 100        |
| <b>10 Die Feynman-Kac-Formel und das Dirichlet-Problem</b>     | <b>106</b> |
| 10.1 Feynman-Kac-Formel für Brownsche Bewegung . . . . .       | 107        |
| 10.2 Feynman-Kac-Formel für Diffusionsprozesse . . . . .       | 111        |
| 10.3 Harmonische Funktionen und das Dirichletproblem . . . . . | 114        |

# 1 Stochastische Prozesse in stetiger Zeit

## 1.1 Grundlagen

Wir betrachten zufällige Prozesse, definiert auf einem Wahrscheinlichkeitsraum  $(\Omega, \mathcal{F}, \mathbb{P})$ , welche Werte in einem fest gewählten Zustandsraum annehmen. Der Zustandsraum ist im Allgemeinen zumindest ein messbarer Raum  $(S, \mathcal{S})$ ; wir werden jedoch in diesem Skript nur den Fall  $S = \mathbb{R}^d$  mit der Borel-Algebra  $\mathcal{S} = \mathcal{B}(\mathbb{R}^d)$  behandeln. Die meisten Resultate aus dem ersten Kapitel lassen sich ohne wesentliche Änderungen auf den Fall übertragen, dass  $S$  ein sogenannter polnischer Raum ist.<sup>1</sup>

**Definition.** Ein **stochastischer Prozess** ist eine messbare Abbildung  $X : \mathbb{R}_{\geq 0} \times \Omega \rightarrow S$ ,  $(t, \omega) \mapsto X(t, \omega)$ .

Als messbare Struktur auf  $\mathbb{R}_{\geq 0} \times \Omega$  wird dabei die Produkt- $\sigma$ -Algebra  $\mathcal{B}(\mathbb{R}_{\geq 0}) \otimes \mathcal{F}$  festgesetzt. Je nach Bedarf wird die Notation  $X(t, \omega)$  auch mit  $X_t(\omega)$  oder  $X_t$  abgekürzt. Der Parameter  $t$  wird stets als *Zeit* interpretiert, was unserer Vorstellung eines stochastischen *Prozesses* entspricht.

**Definition.** Für festes  $\omega \in \Omega$  heisst die Abbildung  $t \mapsto X_t(\omega)$  **Pfad** von  $X$ .

Diese Definition stellt die Abhängigkeit vom Zeitparameter  $t$  in den Vordergrund. Umgekehrt können wir auch für fest gewählte Zeitpunkte  $t_1 < \dots < t_n$  den Zufallsvektor  $\omega \mapsto (X_{t_1}(\omega), \dots, X_{t_n}(\omega))$  betrachten. Die Verteilung von  $(X_{t_1}, \dots, X_{t_n})$  heisst endlich-dimensionale Randverteilung von  $X$  (zu den Zeitpunkten  $(t_1, \dots, t_n)$ ). Viele interessante Aspekte von stochastischen Prozessen ergeben sich aus dem Zusammenspiel zwischen ‘Pfadeigenschaften’ (d.h. Eigenschaften der Abbildungen  $t \mapsto X_t(\omega)$ ) und Eigenschaften der Randverteilungen. Im allgemeinen werden wir an die Pfade der betrachteten stochastischen Prozesse gewisse Regularitätsanforderungen stellen:

**Definition.** Existiert eine  $\mathbb{P}$ -Nullmenge  $A$  sodass für alle  $\omega \in \Omega \setminus A$  der Pfad  $t \mapsto X_t(\omega)$  stetig ist, so heisst  $X$  **stetiger stochastischer Prozess**.

Stetigkeit ist eine Pfadeigenschaft von  $X$  die sich nicht aus den endlich-dimensionalen Randverteilungen von  $X$  alleine ablesen lässt. Analog zur obigen Definition lassen sich auch **rechtsstetige stochastische Prozesse** definieren. Rechtssteigkeit des Pfads  $t \mapsto X_t(\omega)$  bedeutet dass an jeder Stelle  $t \geq 0$  der rechtsseitige Limes  $\lim_{s \downarrow t} X_s(\omega)$  existiert und mit  $X_t(\omega)$  übereinstimmt.

Wichtigstes Beispiel für einen stetigen Prozess ist die (bereits bekannte) Brownsche Bewegung.

---

<sup>1</sup>Ein topologischer Raum heisst *polnisch*, falls er separabel ist und sich mit einer vollständigen Metrik ausstatten lässt. Separabel bedeutet hier dass der Raum eine abzählbare dichte Teilmenge enthält.

**Definition.** Ein  $\mathbb{R}$ -wertiger stochastischer Prozess  $(B_t)_{t \geq 0}$  heisst **Brownsche Bewegung**, wenn er folgende Eigenschaften besitzt:

- (a)  $B_0 = 0$ ;
- (b)  $B$  hat unabhängige Zuwächse, d.h. für alle  $0 \leq t_1 < \dots < t_n, n \in \mathbb{N}$  sind die Zufallsvariablen

$$B_{t_2} - B_{t_1}, B_{t_3} - B_{t_2}, \dots, B_{t_n} - B_{t_{n-1}}$$

unabhängig;

- (c) Für jedes  $t \geq s \geq 0$  ist  $B_t - B_s$  normalverteilt mit Mittelwert 0 und Varianz  $t - s$ ;
- (d) Es existiert  $A \in \mathcal{F}$  mit  $\mathbb{P}(A) = 1$ , sodas  $t \mapsto B_t(\omega)$  stetig ist für alle  $\omega \in A$ .

Für die Existenz der Brownschen Bewegung sei auf Steele [2003] oder Bolthausen [2009] verwiesen. Die Forderung nach stetigen Pfaden schliesst einige wichtige stochastische Prozesse aus die lediglich rechtsstetige Pfade besitzen. Ein wichtiges Beispiel ist der Poisson-Prozess, welcher sich folgendermaßen definieren lässt: Es sei  $(E_k)_{k \in \mathbb{N}}$  eine Familie von unabhängigen exponentialverteilten Zufallsvariablen mit Parameter  $\lambda > 0$ . Definiere

$$N_t(\omega) := \max \left\{ n \in \mathbb{N} : \sum_{k=1}^n E_k(\omega) < t \right\}.$$

Dann heisst  $(N_t)_{t \geq 0}$  Poisson-Prozess mit Rate  $\lambda$ . Es lässt sich leicht überprüfen, dass  $N$  ein rechtsstetiger aber kein stetiger stochastischer Prozess ist.

## 1.2 Gleichheit von stochastischen Prozessen

Schon bei gewöhnlichen Zufallsvariablen müssen wir zwischen Gleichheit *fast sicher* und *in Verteilung* unterscheiden. Für stochastische Prozesse in stetiger Zeit ist eine noch feinere Unterscheidung notwendig.

**Definition.** (a)  $X$  und  $Y$  sind **Versionen**<sup>2</sup> voneinander, falls sie die selben endlich-dimensionalen Randverteilungen besitzen, d.h.

$$\mathbb{P}(X_{t_1} \in A_1, \dots, X_{t_n} \in A_n) = \mathbb{P}(Y_{t_1} \in A_1, \dots, Y_{t_n} \in A_n)$$

für alle  $n \in \mathbb{N}$ ,  $(t_1, \dots, t_n) \in \mathbb{R}_{\geq 0}^n$  und  $A_1, \dots, A_n \in \mathcal{B}(\mathbb{R}^d)$ .

<sup>2</sup>In einem Teil der Literatur wird ‘Version’ synonym mit ‘Modifikation’ im Sinne von Punkt (b) verwendet und Prozesse die (a) erfüllen einfach als ‘Prozesse mit den selben endlich-dimensionalen Randverteilungen’ bezeichnet.

(b)  $X$  heisst **Modifikation** von  $Y$  falls

$$\mathbb{P}(X_t = Y_t) = 1, \quad \forall t \geq 0.$$

(c)  $X$  heisst **ununterscheidbar** von  $Y$  falls

$$\mathbb{P}(X_t = Y_t, \forall t \geq 0) = 1.$$

Es gelten die Implikationen (c)  $\Rightarrow$  (b)  $\Rightarrow$  (a). Für Gleichheit im Sinne von (a) müssen die Prozesse  $X$  und  $Y$  nicht auf dem selben Wahrscheinlichkeitsraum definiert sein; für (b) und (c) hingegen schon. Das folgende Beispiel zeigt dass (b) und (c) nicht äquivalent sind.

*Beispiel.* Sei  $T$  exponentialverteilt mit Rate 1. Definiere  $X_t(\omega) \equiv 0$  für alle  $(t, \omega) \in \mathbb{R}_{\geq 0} \times \Omega$  und

$$Y_t = \begin{cases} 0 & t \neq T \\ 1 & t = T. \end{cases}$$

Dann gilt:  $X$  ist eine Modifikation von  $Y$ , aber  $X$  ist nicht ununterscheidbar von  $Y$ . Außerdem ist  $X$  ein stetiger Prozess,  $Y$  jedoch nicht.

Die Notwendigkeit zwischen Gleichheit im Sinne von (b) und (c) zu unterscheiden ist eine Folge der Überabzählbarkeit der verwendeten Zeitskala  $\mathbb{R}_{\geq 0}$ . Falls wir lediglich stochastische Prozesse in diskreter Zeit – d.h. mit Zeitskala  $\mathbb{N}$  – betrachten würden, so würde aus der  $\sigma$ -Additivität des Wahrscheinlichkeitsmaßes  $\mathbb{P}$  folgen dass

$$\mathbb{P}(X_n = Y_n) = 1 \quad \forall n \in \mathbb{N} \quad \iff \quad \mathbb{P}(X_n = Y_n, \quad \forall n \in \mathbb{N}) = 1.$$

In manchen Fällen lässt sich dieses ‘Überabzählbarkeitsproblem’ durch Pfadeigenschaften der Prozesse umgehen, wie das folgende Lemma zeigt.

**Lemma 1.1.** *Es seien  $X$  und  $Y$  rechtsstetige stochastische Prozesse und Modifikationen voneinander. Dann sind  $X$  und  $Y$  ununterscheidbar.*

*Beweis.* Wir betrachten eine abzählbare dichte Teilmenge von  $\mathbb{R}_{\geq 0}$ , z.B.  $\mathbb{Q}_+ := \mathbb{Q} \cap \mathbb{R}_{\geq 0}$ . Sei  $A = \{\omega \in \Omega : X_s = Y_s \quad \forall s \in \mathbb{Q}_+\}$ . Dann gilt wegen der  $\sigma$ -Additivität von  $\mathbb{P}$

$$\mathbb{P}(A^c) = \mathbb{P}\left(\bigcup_{s \in \mathbb{Q}_+} \{\omega \in \Omega : X_s(\omega) \neq Y_s(\omega)\}\right) \leq \sum_{s \in \mathbb{Q}_+} \mathbb{P}(X_s \neq Y_s) = 0,$$

weil  $Y$  eine Modifikation von  $X$  ist, und wir folgern  $\mathbb{P}(A) = 1$ . Da  $X$  und  $Y$  rechtsstetige Prozesse sind, existiert eine Menge  $B \in \mathcal{F}$  mit  $\mathbb{P}(B) = 1$  sodass die Pfade  $t \mapsto X_t(\omega)$

und  $t \mapsto Y_t(\omega)$  für alle  $\omega \in B$  rechtsstetige Funktionen sind. Jedes  $t \in \mathbb{R}_{\geq 0}$  lässt sich durch eine fallende Folge  $t_n \downarrow t$  in  $\mathbb{Q}_+$  approximieren, und es gilt

$$X_t(\omega) = \lim_{t_n \in \mathbb{Q}_+, t_n \downarrow t} X_{t_n}(\omega) = \lim_{t_n \in \mathbb{Q}_+, t_n \downarrow t} Y_{t_n}(\omega) = Y_t(\omega)$$

für alle  $\omega \in A \cap B$ . Die Behauptung folgt nun aus

$$\mathbb{P}(X_t = Y_t \forall t \in \mathbb{R}_{\geq 0}) \leq \mathbb{P}(A \cap B) = 1. \quad \square$$

### 1.3 Filtrationen

Den Parameter  $t$  haben wir im letzten Abschnitt als Zeitparameter eingeführt. Der Begriff der Filtration wird der Zeitachse nun eine Richtung verleihen und der Tatsache Rechnung tragen, dass die Zeit nur vorwärts fließt und sich nicht ohne weiteres umkehren lässt.

**Definition.** Eine Familie von  $\sigma$ -Algebren  $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$  mit  $\mathcal{F}_s \subseteq \mathcal{F}_t$  und  $\mathcal{F}_t \subseteq \mathcal{F}$  für alle  $s \leq t$  heisst **Filtration** (auf  $(\Omega, \mathcal{F})$ ).

Eine Filtration lässt sich als  $\sigma$ -Algebra der zum Zeitpunkt  $t$  bekannten Ereignisse auffassen, wobei einmal bekannt gewordene Ereignisse auch niemals vergessen werden. Letzteres folgt aus der ansteigenden Eigenschaft  $\mathcal{F}_s \subseteq \mathcal{F}_t$  für  $s \leq t$ , welche dem unumkehrbaren Vorwärtsfluss der Zeit entspricht.

**Definition.** Ein stochastischer Prozess  $X$  heisst **adaptiert** (an  $\mathbb{F}$ ) falls  $X_t$  messbar bezüglich  $\mathcal{F}_t$  für alle  $t \geq 0$  ist.

Für ‘ $X_t$  messbar bezüglich  $\mathcal{F}_t$ ’ schreiben wir auch kurz  $X_t \in \mathcal{F}_t$ . Die Eigenschaft der Adaptiertheit lässt sich so interpretieren, dass für einen Beobachter dem alle Ereignisse in  $\mathcal{F}_t$  bekannt sind auch der Zustand  $X_t$  des Prozesses  $X$  zum Zeitpunkt  $t$  bekannt ist. Zu jedem stochastischen Prozess  $X$  existiert die *natürliche Filtration*  $\mathcal{F}_t^X = \sigma(X_s, s \leq t)$ . Dies ist die kleinste  $\sigma$ -Algebra an die  $X$  adaptiert ist.

Um mit bestimmten Schwierigkeiten die sich aus der Überabzählbarkeit der Zeitachse ergeben eleganter umgehen zu können werden wir am gegebenen Wahrscheinlichkeitsraum  $(\Omega, \mathcal{F}, \mathbb{P})$  einige Änderungen vornehmen. Sei  $A \in \mathcal{F}$  mit  $\mathbb{P}(A) = 0$ . Dann heisst jede Untermenge  $B \subseteq A$  *Nullmenge*. Falls jede Nullmenge ebenfalls in  $\mathcal{F}$  ist (und damit automatisch Maß 0 hat) so heisst  $(\Omega, \mathcal{F}, \mathbb{P})$  *vollständig*. Jeder Wahrscheinlichkeitsraum lässt sich vervollständigen, indem man  $\mathcal{F}$  durch die  $\sigma$ -Algebra die von  $\mathcal{F}$  gemeinsam mit allen Nullmengen erzeugt wird ersetzt und das Wahrscheinlichkeitsmaß  $\mathbb{P}$  fortsetzt indem jeder Nullmenge das Maß 0 zugeordnet wird. Auch die Filtration  $\mathbb{F}$  ändern wir entsprechend.

**Definition.** Auf einem vollständigen Wahrscheinlichkeitsraum  $(\Omega, \mathcal{F}, \mathbb{P})$  mit Brownscher Bewegung definieren wir die **Brownsche Standardfiltration** (BSF) als Filtration die von der Brownschen Bewegung zusammen mit allen  $\mathbb{P}$ -Nullmengen erzeugt wird.

Die Brownsche Standardfiltration besitzt folgende zwei Eigenschaften, für die sich der wenig aussagekräftige Name ‘übliche Bedingungen’ (*usual conditions*) durchgesetzt hat.

**Definition.** Eine Filtration  $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$  genügt den **üblichen Bedingungen** falls folgendes gilt:

- (a)  $\mathcal{F}_0$  enthält alle  $\mathbb{P}$ -Nullmengen;
- (b)  $\mathbb{F}$  ist rechtsstetig, d.h.

$$\mathcal{F}_t = \bigcap_{s:t < s} \mathcal{F}_s.$$

Ein Wahrscheinlichkeitsraum mit Filtration die den üblichen Bedingungen genügt ist automatisch vollständig, da  $\mathcal{F}_0 \subset \mathcal{F}$  und somit laut (a) alle  $\mathbb{P}$ -Nullmengen in  $\mathcal{F}$  enthalten sind. Aus (a) ergibt sich auch folgende nützliche Folgerung:

**Lemma 1.2.** *Sei  $X$  ein stochastischer Prozess der bezüglich einer Filtration  $\mathbb{F}$  adaptiert ist, welche den üblichen Bedingungen genügt. Wenn  $Y$  von  $X$  ununterscheidbar ist, so ist auch  $Y$   $\mathbb{F}$ -adaptiert.*

Der Beweis ist dem Leser überlassen und benutzt nur Punkt (a) der üblichen Bedingungen. Punkt (b) wird unter anderem im Beweis des Stoppsatzes für Martingale verwendet werden.

## 1.4 Martingale

Der bereits aus diskreter Zeit bekannte Begriff des Martingals lässt sich ohne wesentliche Änderungen in stetige Zeit übertragen. Auch dort wird er sich als zentral für die Theorie der stochastischen Prozesse herausstellen.

**Definition.** Ein adaptierter stochastischer Prozess  $X$  auf einem gefilterten Wahrscheinlichkeitsraum heisst **Martingal** falls gilt

- (a)  $\mathbb{E}[|X_t|] < \infty, \quad \forall t \geq 0$
- (b)  $X_s = \mathbb{E}[X_t | \mathcal{F}_s], \quad \forall 0 \leq s \leq t.$

Wenn  $X$  stetig ist, sprechen wir von einem **stetigen Martingal**.

Viele Resultate lassen sich etwas allgemeiner auch für **rechtsstetige Martingale** formulieren. Für Prozesse mit Zustandsraum  $\mathbb{R}$  können wir die Gleichung in Punkt (b) durch eine Ungleichung ersetzen und erhalten zwei verwandte Klassen von stochastischen Prozessen.

**Definition.** Ein adaptierter stochastischer Prozess  $X$  mit Zustandsraum  $\mathbb{R}$  heisst **Submartingal** wenn in der vorangegangenen Definition statt (b) die Eigenschaft

$$(b\text{-sub}) \quad X_s \leq \mathbb{E}[X_t | \mathcal{F}_s], \quad \forall 0 \leq s \leq t$$

gilt und heisst **Supermartingal** wenn statt (b) die Eigenschaft

$$(b\text{-super}) \quad X_s \geq \mathbb{E}[X_t | \mathcal{F}_s], \quad \forall 0 \leq s \leq t$$

gilt.

Nach Anwendung der (unbedingten) Erwartung und der Turmregel auf beiden Seiten von (b) sehen wir dass für ein Martingal  $\mathbb{E}[X_s] = \mathbb{E}[X_t]$  für alle  $0 \leq s \leq t$  gilt. Analog folgt  $\mathbb{E}[X_s] \leq \mathbb{E}[X_t]$  für Submartingale und  $\mathbb{E}[X_s] \geq \mathbb{E}[X_t]$  für Supermartingale. Es gilt also die Merkregel

‘Das Leben ist ein Supermartingal; die Erwartungen sinken mit der Zeit’.

Wenn wir wieder die Filtration  $\mathcal{F}_s$  als ‘zum Zeitpunkt  $s$  bekannte Ereignisse’ interpretieren und  $t \geq s$  als ‘noch unbekanntes Zukunft’ so ergibt sich folgende Interpretation der Martingaleigenschaft: Die Beste mit der Information zum Zeitpunkt  $s$  mögliche Schätzung des zukünftigen Werts  $X_t$  ist der momentane Wert  $X_s$ . Für Submartingale hingegen ist der momentane Wert  $X_s$  ein *nach unten* verzerrter Schätzer für  $X_t$ ; für Supermartingale ein *nach oben* verzerrter.

Submartingale treten häufig als konvexe Transformationen von Martingalen auf. Auch wenn man eine steigende konvexe Transformation auf ein Submartingal anwendet erhält man wieder ein Submartingal; in diesem Sinne sind Submartingale also etwas ‘robuster’ als Martingale.

**Proposition 1.3.** 1. Sei  $X$  ein Martingal und  $\Phi : \mathbb{R}^d \rightarrow \mathbb{R}$  eine konvexe Funktion, sodass  $\mathbb{E}|\Phi(X_t)| < \infty$  für alle  $t \geq 0$ . Dann ist  $Y_t = \Phi(X_t)$  ein Submartingal.

2. Sei  $X$  ein Submartingal und  $\Phi : \mathbb{R} \rightarrow \mathbb{R}$  eine wachsende konvexe Funktion, sodass  $\mathbb{E}|\Phi(X_t)| < \infty$  für alle  $t \geq 0$ . Dann ist  $Y_t = \Phi(X_t)$  ebenfalls ein Submartingal.

*Beweis.* Die Behauptung ist in beiden Fällen eine einfache Folgerung aus der Jensen’schen Ungleichung für bedingte Erwartungswerte. Im Fall (a) gilt

$$\mathbb{E}[Y_t | \mathcal{F}_s] = \mathbb{E}[\Phi(X_t) | \mathcal{F}_s] \geq \Phi(\mathbb{E}[X_t | \mathcal{F}_s]) = \Phi(X_s) = Y_s;$$



in Fall (b) gilt

$$\mathbb{E}[Y_t | \mathcal{F}_s] = \mathbb{E}[\Phi(X_t) | \mathcal{F}_s] \geq \Phi(\mathbb{E}[X_t | \mathcal{F}_s]) \geq \Phi(X_s) = Y_s,$$

wobei für das letzte Ungleichungszeichen die Monotonie von  $\Phi$  benötigt wird. In beiden Fällen folgt die Submartingaleigenschaft von  $Y$ .  $\square$

Folgende Beispiele treten häufig auf:

- Sei  $(X_t)$  ein Martingal; dann ist seine euklidische Norm  $(|X_t|)$  stets ein Submartingal.
- Sei  $(X_t)$  ein Martingal und  $p > 1$ . Falls  $\mathbb{E}|X_t|^p < \infty$  so ist auch  $(|X_t|^p)$  ein Submartingal.

## 1.5 Stoppzeiten

Gemeinsam mit Martingalen gehören Stoppzeiten zu den wichtigsten Werkzeugen um mit stochastischen Prozessen zu arbeiten. Kaum einer der folgenden Beweise wird ohne einen dieser beiden Begriffe auskommen. Das Lemma zum ‘optionalen Stoppen’ von Martingalen im nächsten Kapitel und der Begriff des ‘lokalen Martingals’ werden Martingale und Stoppzeiten eng miteinander verbinden.

**Definition.** Eine Zufallsvariable  $\tau : \Omega \rightarrow \mathbb{R}_{\geq 0} \cup \{+\infty\}; \omega \mapsto \tau(\omega)$  heisst **Stoppzeit**, falls

$$\{\tau \leq t\} \in \mathcal{F}_t, \quad \forall t \geq 0.$$

Mit unserer Interpretation von  $\mathcal{F}_t$  als zum Zeitpunkt  $t$  bekannten Ereignissen, lässt sich diese Definition folgendermaßen lesen: Ein zufälliger Zeitpunkt  $\tau$  heisst Stoppzeit, falls zu jedem Zeitpunkt  $t$  mit Sicherheit bekannt ist ob  $\tau$  bereits in der Vergangenheit liegt oder nicht. Man beachte dass auch der Wert  $\tau(\omega) = +\infty$  zugelassen ist. Stoppzeiten sind stabil unter einer Reihe von Operationen:

**Lemma 1.4.** (a) *Es seien  $\tau, \sigma$  Stoppzeiten bezüglich einer Filtration  $\mathbb{F}$ . Dann sind auch  $\sigma \wedge \tau$ ,  $\sigma \vee \tau$  und  $\sigma + \tau$   $\mathbb{F}$ -Stoppzeiten.*

(b) *Es seien  $(\tau_n)_{n \in \mathbb{N}}$  Stoppzeiten bezüglich einer Filtration  $\mathbb{F}$ . Dann ist auch  $\sup_{n \in \mathbb{N}} \tau_n$  eine  $\mathbb{F}$ -Stoppzeit. Falls  $\mathbb{F}$  rechtsstetig ist, so ist auch  $\inf_{n \in \mathbb{N}} \tau_n$  eine  $\mathbb{F}$ -Stoppzeit.*

*Beweis.* Der Beweis von (a) sei dem Leser überlassen. Für (b) schreibe

$$\left\{ \sup_{n \in \mathbb{N}} \tau_n \leq t \right\} = \bigcap_{n \in \mathbb{N}} \{ \tau_n \leq t \} \quad \text{und} \quad \left\{ \inf_{n \in \mathbb{N}} \tau_n < s \right\} = \bigcup_{n \in \mathbb{N}} \{ \tau_n < s \};$$

wegen  $\{\tau_n \leq t\} \in \mathcal{F}_t$  ist damit ist der Beweis für  $\sup_n \tau_n$  beendet. Für das Infimum beachte dass wegen  $\{\tau_n < s\} = \bigcup_{k \in \mathbb{N}} \{\tau_n \leq s - \frac{1}{k}\}$  auch  $\{\tau_n < s\} \in \mathcal{F}_s$  gilt. Schliesslich verwenden wir die Rechtsstetigkeit von  $\mathbb{F}$  und schreiben

$$\left\{ \inf_n \tau_n \leq t \right\} = \bigcap_{s: s > t} \left\{ \inf_n \tau_n < s \right\} \in \mathcal{F}_t. \quad \square$$

Gegeben eine Stoppzeit  $\tau$  nennen wir ein Ereignis  $A \in \mathcal{F}$  **bekannt vor**  $\tau$  wenn

$$A \cap \{\tau \leq t\} \in \mathcal{F}_t, \quad \forall t \geq 0,$$

und definieren die  $\sigma$ -Algebra  $\mathcal{F}_\tau$  als

$$\mathcal{F}_\tau = \sigma(A \in \mathcal{F} : A \text{ bekannt vor } \tau).^3$$

Die Rechtsstetigkeit von  $\mathbb{F}$  bleibt auch entlang von Stoppzeiten erhalten:

**Lemma 1.5.** *Sei  $\mathbb{F}$  eine rechtsstetige Filtration mit Stoppzeiten  $\tau_n \downarrow \tau$ . Dann gilt*

$$\mathcal{F}_\tau = \bigcap_{n \in \mathbb{N}} \mathcal{F}_{\tau_n}.$$

Der Beweis lässt sich sehr ähnlich zum Beweis des vorangegangenen Lemmas führen und bleibt dem Leser überlassen. Später werden wir auch noch folgendes Resultat benötigen:

**Lemma 1.6.** *Sei  $X$  ein rechtsstetiger stochastischer Prozess, adaptiert an eine rechtsstetige Filtration  $\mathbb{F}$ , und  $\tau$  eine  $\mathbb{F}$ -Stoppzeit sodass  $\tau < \infty$  fast sicher gilt. Dann ist  $X_\tau$  eine wohl-definierte Zufallsvariable und es gilt  $X_\tau \in \mathcal{F}_\tau$ .*

*Beweis.* Die Funktionen  $(t, \omega) \mapsto X(t, \omega)$  und  $\omega \mapsto \tau(\omega)$  sind messbar. Falls  $\tau < \infty$  so ist auch ihre Verknüpfung  $X_\tau(\omega) = X(\tau(\omega), \omega)$  wohldefiniert, somit messbar und folglich eine Zufallsvariable auf  $\Omega$ .

Für den zweiten Teil der Behauptung müssen wir zeigen, dass

$$\{X_\tau \in B\} \cap \{\tau \leq t\} \in \mathcal{F}_t \tag{1.1}$$

für alle  $B \in \mathcal{B}(\mathbb{R}^d)$  und  $t \geq 0$ . Da  $\mathcal{B}(\mathbb{R}^d)$  von allen abgeschlossenen Mengen in  $\mathbb{R}^d$  erzeugt wird, genügt es die Gleichung für abgeschlossenes  $B$  zu zeigen. Definiere die ‘dyadische Partition’

$$S(n) := \{k/2^n : k \in \mathbb{N}\} \subset \mathbb{R}_{\geq 0}.$$

<sup>3</sup>Es lässt sich zeigen, dass die vor  $A$  bekannten Ereignisse bereits eine  $\sigma$ -Algebra bilden und somit der  $\sigma(\cdot)$ -Operator auch weggelassen werden kann.

Für jedes  $\omega \in \Omega$  bezeichne  $\tau_n(\omega)$  das kleinste Element in  $S(n)$  welches grösser als  $\tau(\omega)$  ist, und  $t_n$  sei das kleinste Element in  $S(n)$  größer  $t$ . Es gilt  $\{\tau_n \leq t_n\} = \{\tau \leq t_n\} \in \mathcal{F}_{t_n}$ . Des weiteren gilt für jedes  $n \in \mathbb{N}$

$$\{X_{\tau_n} \in B\} \cap \{\tau_n \leq t_n\} = \bigcup_{k \in \mathbb{N}} \{X_{k/2^n} \in B; k2^{-n} \leq t_n; \tau_n = k/2^n\} \in \mathcal{F}_{t_n}.$$

Wir bilden auf beiden Seiten den Durchschnitt über alle  $n \in \mathbb{N}$ . Auf der linken Seite erhalten wir aufgrund der Rechtsstetigkeit von  $X$  den Ausdruck  $\{X_\tau \in B\} \cap \{\tau \leq t\}$ ; auf der rechten Seite erhalten wir wegen der Rechtsstetigkeit von  $\mathbb{F}$  die  $\sigma$ -Algebra  $\mathcal{F}_t = \bigcap_{n \in \mathbb{N}} \mathcal{F}_{t_n}$  und (1.1) folgt.  $\square$

Die wichtigsten Beispiele für Stoppzeiten die wir kennenlernen werden, sind *Eintrittszeiten* von stochastischen Prozessen

**Definition.** Sei  $X$  ein adaptierter stochastischer Prozess und  $B$  eine Borel-Menge in  $\mathbb{R}^d$ . Dann heisst

$$\tau_B := \inf \{t > 0 : X_t \in B\}$$

**Eintrittszeit** von  $X$  in  $B$ .

Wir verwenden hier die Konvention dass das Infimum der leeren Menge gleich  $+\infty$  ist. Eine wichtige und tiefgreifende Frage ist, ob Eintrittszeiten tatsächlich immer Stoppzeiten sind. Wir begnügen uns mit folgendem Resultat für den Fall dass  $B$  eine offene bzw. abgeschlossene Menge ist.<sup>4</sup>

**Theorem 1.7.** (a) Sei  $X$  ein stetiger Prozess, adaptiert an eine Filtration  $\mathbb{F}$  und  $B$  eine abgeschlossene Menge. Dann ist die Eintrittszeit  $\tau_B$  eine Stoppzeit bezüglich  $\mathbb{F}$ .

(b) Sei  $X$  ein rechtsstetiger Prozess, adaptiert an eine rechtsstetige Filtration  $\mathbb{F}$  und  $B$  eine offene Menge. Dann ist die Eintrittszeit  $\tau_B$  eine Stoppzeit bezüglich  $\mathbb{F}$ .

*Beweis.* Sei  $B$  zunächst eine offene Menge. Dann gilt wegen der Rechtssteigkeit von  $X$  die Äquivalenz:  $X_s \in B$  für ein  $s \in [0, t) \Leftrightarrow X_s \in B$  für ein  $s \in \mathbb{Q} \cap [0, t)$ . Es folgt

$$\{\tau_B < t\} = \bigcup_{s: s \in \mathbb{Q} \cap [0, t)} \{X_s \in B\} \in \mathcal{F}_t.$$

Für (b) dürfen wir nun zusätzlich annehmen dass  $\mathbb{F}$  rechtsstetig ist und erhalten die Behauptung indem wir  $\{\tau_B \leq t\} = \bigcap_{s: s > t} \{\tau_B < s\}$  schreiben. Für (a) sei  $B$  eine abgeschlossene Menge und  $d(x, B) := \inf_{y \in B} |x - y|$  der euklidische Abstand von  $x$

<sup>4</sup>Eine allgemeinere Antwort gibt das sogenannte *debut theorem* (siehe Kallenberg [1997, Thm. 7.7]), welches jedoch einen weitaus komplizierteren Beweis erfordert.

und  $B$ . Dann sind  $B_n = \{x \in \mathbb{R}^d : d(x, B) < 1/n\}$  offene Mengen mit  $B_n \downarrow B$ . Unter der Annahme der (links-)Stetigkeit von  $X$  gilt

$$\{\tau_B \leq t\} = \bigcap_{n \in \mathbb{N}} \{\tau_{B_n} < t\} \in \mathcal{F}_t,$$

wobei wir auf  $\{\tau_{B_n} < t\}$  den ersten Teil des Beweises anwenden können.  $\square$

Oft sind wir auch an Eintrittszeiten in ‘zeitabhängige’ Mengen  $B \subset \mathbb{R}_{\geq 0} \times \mathbb{R}^d$  interessiert, d.h. an Eintrittszeiten vom Typ

$$\tau_B := \inf \{t \geq 0 : (t, X_t) \in B\}.$$

Hier bietet es sich an dem Prozess eine (deterministische) Komponente hinzuzufügen und den Prozess  $\tilde{X}_t := (t, X_t)$  auf dem entsprechend angepassten Zustandsraum  $\mathbb{R}^{d+1}$  zu betrachten. Falls  $X$  zur Filtration  $\mathbb{F}$  adaptiert ist, so ist auch  $\tilde{X}$  adaptiert; genauso bleibt Stetigkeit und Rechtsstetigkeit erhalten und Theorem 1.7 lässt sich anwenden.

## 1.6 Gleichgradige Integrierbarkeit

**Definition.** Sei  $\mathcal{C}$  eine Familie von Zufallsvariablen und definiere

$$\rho(r) = \sup_{Z \in \mathcal{C}} \mathbb{E}[|Z| \mathbf{1}\{|Z| > r\}], \quad (r \geq 0). \quad (1.2)$$

Die Familie  $\mathcal{C}$  heisst **gleichgradig Integrierbar** wenn

$$\lim_{r \rightarrow \infty} \rho(r) = 0.$$

Für großes  $r$  lässt sich der Ausdruck  $\mathbb{E}[|Z| \mathbf{1}\{|Z| > r\}]$  als Wahrscheinlichkeitsmasse im äussersten Verteilungsende von  $Z$  interpretieren. Gleichgradige Integrierbarkeit besagt dass diese Masse *gleichmässig* über alle  $Z \in \mathcal{C}$  hinweg verschwindet, wenn  $r \rightarrow \infty$ . Eine gleichgradig integrierbare Familie von Zufallsvariablen ist auch automatisch gleichmässig in  $L^1$  beschränkt, denn  $\mathbb{E}[|Z|] \leq \rho(r) + r$  für alle  $Z \in \mathcal{C}$  und  $r \geq 0$ . Wichtigste Anwendung der gleichgradigen Integrierbarkeit ist der Schluss von fast sicherer Konvergenz auf  $L^1$ -Konvergenz:

**Lemma 1.8.** *Sei  $(Z_n)_{n \in \mathbb{N}}$  gleichgradig integrierbar und es gelte  $Z_n \rightarrow Z$  fast sicher. Dann gilt auch  $\mathbb{E}[|Z_n - Z|] \rightarrow 0$  für  $n \rightarrow \infty$ .*

Das folgende Beispiel zeigt, dass sogar gleichmäßige Beschränktheit der  $L_1$ -Norm nicht für den Schluss von fast sicherer Konvergenz auf  $L_1$ -Konvergenz ausreicht.

*Beispiel.* Die Zufallsvariablen  $(Z_n)_{n \in \mathbb{N}}$  seien unabhängig mit der Verteilung

$$Z_n = \begin{cases} 0 & \text{mit Wahrscheinlichkeit } \frac{n-1}{n} \\ n & \text{mit Wahrscheinlichkeit } \frac{1}{n}. \end{cases}$$

Dann gilt  $\mathbb{E}[|Z_n|] = 1$ , die Zufallsvariablen sind also gleichmäßig in  $L_1$  beschränkt. Außerdem gilt  $Z_n \rightarrow 0$  fast sicher. Offensichtlich gilt aber  $\lim_{n \rightarrow \infty} \mathbb{E}[|Z_n - 0|] = 1$ , sodass keine  $L_1$ -Konvergenz stattfindet. Wir folgern aus Lemma 1.8, dass die  $Z_n$  nicht gleichgradig integrierbar sein können; und tatsächlich gilt  $\rho(r) = 1$  für alle  $r \geq 0$ .

*Beweis von Lemma 1.8.* Aus der gleichgradigen Integrierbarkeit von  $(Z_n)$  und (1.2) folgt mit dem Lemma von Fatou

$$\mathbb{E}[|Z| \mathbf{1}\{|Z| > r\}] \leq \rho(r).$$

Insbesondere gilt  $\mathbb{E}[|Z|] < \infty$  und die Schranke  $\mathbb{E}[|Z|] < \rho(r) + r$  für alle  $r \geq 0$ . Wir schätzen

$$|Z_n - Z| \leq |Z_n - Z| \mathbf{1}\{|Z_n| \leq r\} + |Z| \mathbf{1}\{|Z_n| > r\} + |Z_n| \mathbf{1}\{|Z_n| > r\}$$

ab. Der erste Summand ist durch  $r + Z$  beschränkt und dominierte Konvergenz liefert die  $L^1$ -Konvergenz gegen 0. Auch auf den zweiten Summanden lässt sich dominierte Konvergenz anwenden und wir erhalten dass  $\lim_{n \rightarrow \infty} \mathbb{E}[|Z| \mathbf{1}\{|Z_n| > r\}] = \lim_{n \rightarrow \infty} \mathbb{E}[|Z_n| \mathbf{1}\{|Z_n| > r\}] \leq \rho(r)$ . Insgesamt ergibt sich dass

$$\lim_{n \rightarrow \infty} \mathbb{E}[|Z_n - Z|] \leq 2\rho(r), \quad \forall r \geq 0$$

und die Behauptung folgt aus  $\lim_{r \rightarrow \infty} \rho(r) = 0$ . □

Lemma 1.8 lässt sich sehr leicht auf bedingte Erwartungswerte ausweiten. Mit

$$\mathbb{E} \left[ |\mathbb{E}[Z_n | \mathcal{G}] - \mathbb{E}[Z | \mathcal{G}]| \right] \leq \mathbb{E}[\mathbb{E}[|Z_n - Z| | \mathcal{G}]] = \mathbb{E}[|Z_n - Z|]$$

erhält man:

**Korollar 1.9.** *Sei  $\mathcal{G}$  eine Unter- $\sigma$ -Algebra von  $\mathcal{F}$ . Unter den Voraussetzungen von Lemma 1.8 konvergiert auch  $\mathbb{E}[Z_n | \mathcal{G}]$  in  $L^1$  gegen  $\mathbb{E}[Z | \mathcal{G}]$ .*

Eine nützliche hinreichende Bedingung für gleichgradige Integrierbarkeit ist die folgende.

**Lemma 1.10.** Sei  $\Phi : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$  eine Funktion mit  $\lim_{r \rightarrow \infty} \frac{\Phi(r)}{r} = \infty$  und  $\mathcal{C}$  eine Familie von Zufallsvariablen mit

$$\mathbb{E}[\Phi(|Z|)] \leq B, \quad \forall Z \in \mathcal{C}.$$

Dann ist  $\mathcal{C}$  gleichgradig integrierbar.

Der Beweis bleibt dem Leser überlassen. Hinweis: Setze  $\Psi(r) = \Phi(r)/r$  und schätze  $|Z|$  mit  $\frac{\Phi(|Z|)}{\Psi(|Z|)}$  ab.

Schließlich sei noch ein Resultat zum Auffinden einer geeigneten Funktion  $\Phi$  erwähnt.

**Lemma 1.11.** Sei  $Z$  eine Zufallsvariable mit  $\mathbb{E}[|Z|] < \infty$ . Dann existiert eine konvexe Funktion  $\Phi : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$  mit  $\lim_{r \rightarrow \infty} \frac{\Phi(r)}{r} = \infty$ , sodass  $\mathbb{E}[\Phi(|Z|)] < \infty$ .

Auch dieser Beweis bleibt dem Leser überlassen. Hinweis: Als Funktion  $\Phi(r)$  lässt sich  $\Phi(r) := \int_0^r a(x)dx$  verwenden, mit  $a(x) = \left\{ \int_x^\infty \mathbb{P}[|Z| > u] du \right\}^{-1/2}$ . Aus dem obigen Lemma folgt als wichtige Konsequenz die gleichgradige Integrierbarkeit der Familie aller bedingten Erwartungen einer Zufallsvariable.

**Lemma 1.12.** Sei  $Z$  eine Zufallsvariable mit  $\mathbb{E}[|Z|] < \infty$ . Dann ist die Familie

$$\mathcal{C} := \{\mathbb{E}[Z|\mathcal{G}] : \mathcal{G} \text{ ist } \sigma\text{-Algebra, } \mathcal{G} \subseteq \mathcal{F}\}$$

gleichgradig integrierbar.

*Beweis.* Mit  $\Phi$  aus Lemma 1.11 gilt nach Anwendung der Jensenschen Ungleichung und der Turmeigenschaft

$$\mathbb{E}[\Phi(\mathbb{E}[Z|\mathcal{G}])] \leq \mathbb{E}[\Phi(Z)] \leq B$$

für jede beliebige  $\sigma$ -Algebra  $\mathcal{G} \subset \mathcal{F}$ . Die Behauptung folgt nun aus Lemma 1.10.  $\square$

## 2 Martingale in stetiger Zeit

Ziel dieses Abschnitts ist es die wichtigsten Resultate für Martingale aus diskreter Zeit in stetige Zeit zu übertragen. Wie zu erwarten ist treten in stetiger Zeit einige beweistechnische Hindernisse auf die in diskreter Zeit nicht zu finden sind. Die meisten davon lassen sich elegant überwinden indem wir Martingale mit *stetigen* oder etwas allgemeiner auch nur *rechtsstetigen* Pfaden betrachten. Zunächst fassen wir nochmals die wichtigsten Resultate für Martingale in diskreter Zeit zusammen.

### 2.1 Martingale in diskreter Zeit

Für Martingale in diskreter Zeit lassen sich folgende Resultate auf recht elementare Weise zeigen. Für die Beweise sei auf die Wahrscheinlichkeitstheorie 2 und auf Klenke [2008] verwiesen.

**Theorem 2.1** (Optionales Sampling – Diskrete Zeit). *Sei  $(M_n)_{n \in \mathbb{N}}$  ein Martingal bezüglich einer Filtration  $\mathbb{F}$  und  $\sigma, \zeta$  zwei  $\mathbb{F}$ -Stoppzeiten, wobei  $\zeta$  fast sicher beschränkt sei. Dann gilt*

$$M_{\sigma \wedge \zeta} = \mathbb{E}[M_\zeta | \mathcal{F}_\sigma].$$

*Die Aussage gilt analog für Submartingale mit ' $\leq$ ' und Supermartingale mit ' $\geq$ ' anstatt von ' $=$ '.*

**Theorem 2.2** (Optionales Stoppen – Diskrete Zeit). *Sei  $(M_n)_{n \in \mathbb{N}}$  ein Martingal bezüglich einer Filtration  $\mathbb{F}$  und  $\tau$  eine  $\mathbb{F}$ -Stoppzeit. Dann ist der gestoppte Prozess  $(M_{n \wedge \tau})_{n \in \mathbb{N}}$  wieder ein Martingal bezüglich  $\mathbb{F}$ . Die Aussage gilt analog für Sub- und Supermartingale.*

Für die nächsten Resultate führen wir folgende Notation ein:  $Z^+(\omega) := \max(Z(\omega), 0)$  bezeichnet den positiven Anteil einer Zufallsvariable und  $Z^-(\omega) := \max(-Z(\omega), 0)$  den negativen Anteil. Es gilt  $Z = Z^+ - Z^-$  und  $|Z| = Z^+ + Z^-$ . Des weiteren sei

$$\|Z\|_p := (E|Z|^p)^{1/p}, \quad p \geq 1$$

die  $L^p$ -norm von  $Z$ . Wir formulieren die folgenden Sätze für Submartingale, was natürlich den Martingalfall einschliesst.

**Theorem 2.3** (Doob'sche Ungleichungen – Diskrete Zeit). *Sei  $(M_n)_{n \in \mathbb{N}}$  ein Submartingal und  $M_n^* = \sup_{0 \leq m \leq n} M_m$  der zugehörige Maximumprozess.*

(a) *Für alle  $\lambda > 0$  und  $n \in \mathbb{N}$  gilt*

$$\mathbb{P}[M_n^* \geq \lambda] \leq \frac{1}{\lambda} \mathbb{E}[M_n^+].$$

(b) Ist  $M_n$  positiv, dann gilt weiters

$$\|(M_n^*)\|_p \leq \frac{p}{p-1} \|M_n\|_p$$

für alle  $p > 1$ .

Zum beweisen der Martingalkonvergenzsätze werden wir schliesslich noch die Aufkreuzungsungleichung benötigen.

**Definition.** Sei  $a < b$ ,  $N \in \mathbb{N} \cup \{+\infty\}$  und  $(\alpha_n)_{n \in \mathbb{N}}$  eine Folge reeller Zahlen. Dann bezeichnet  $U_N[a, b](\alpha)$  die Anzahl der **ansteigenden Überschreitungen** von  $[a, b]$  durch die Folge  $\alpha$  bis zum Zeitpunkt  $N$ . Das heisst  $U_N[a, b](\alpha)$  ist die größte Zahl  $K \in \mathbb{N}$ , sodass sich Indices

$$0 \leq q_1 < p_1 < q_2 < p_2 < \dots < q_K < p_K \leq N$$

finden lassen, mit

$$\alpha_{q_i} < a < b < \alpha_{p_i}, \quad \forall i \in \{1, \dots, K\}.$$

Wenn  $\alpha$  keine Folge, sondern eine Funktion von  $\mathbb{Q}_+$  nach  $\mathbb{R}$  ist, so lässt sich obige Definition leicht verallgemeinern, indem wir  $N \in [0, +\infty]$  und die  $q_i, p_i$  in  $\mathbb{Q}_+$  wählen. Diese Verallgemeinerung wird für den Beweis der Martingalkonvergenz benötigt werden.

Für Submartingale gilt folgendes Resultat

**Lemma 2.4** (Aufkreuzungsungleichung). Sei  $(M_n)_{n \in \mathbb{N}}$  ein Submartingal,  $a < b$  und  $N \in \mathbb{N}$ . Dann gilt

$$\mathbb{E}[U_N[a, b](M)] \leq \frac{\mathbb{E}[(M_N - a)^+] - \mathbb{E}[(M_0 - a)^+]}{b - a}$$

## 2.2 Optionales Sampling und Optionales Stoppen in stetiger Zeit

Beim den Sätzen vom Optionalen Sampling und Optionalen Stoppen handelt es sich um die im letzten Kaptiel angekündigte Verbindung von Martingalen und Stoppzeiten. Der Name kommt daher, dass Stoppzeiten manchmal auch als ‘optionale Zeiten’ bezeichnet werden.<sup>5</sup>

**Theorem 2.5** (Optionales Sampling). Sei  $M$  ein rechtsstetiges Martingal bezüglich einer Filtration  $\mathbb{F}$  welche den üblichen Bedingungen genügt. Des weiteren seien  $\sigma, \zeta$

<sup>5</sup>Genau genommen stimmen die Begriffe Stoppzeit und optionale Zeit nur auf rechtsstetigen Filtrationen überein. Es gibt eine allgemeinere Version des Satzes vom optionalen Stoppen die sich auch auf nicht rechtsstetige Filtrationen anwenden lässt.



zwei  $\mathbb{F}$ -Stoppzeiten, wobei  $\zeta$  fast sicher beschränkt sei. Dann gilt

$$M_{\sigma \wedge \tau} = \mathbb{E}[M_\zeta | \mathcal{F}_\sigma]. \quad (2.1)$$

Die Aussage gilt analog für Submartingale mit ' $\leq$ ' und Supermartingale mit ' $\geq$ ' anstatt von '='.

Insbesondere lässt sich der Satz auf zwei geordnete Stoppzeiten anwenden und besagt dann dass wir in der Martingaleigenschaft die deterministischen Zeitpunkte  $s \leq t$  durch (beschränkte) Stoppzeiten  $\sigma \leq \zeta$  ersetzen dürfen. Ein wichtiges Korollar ergibt sich wenn wir eine beliebige (möglicherweise unbeschränkte) Stoppzeit  $\tau$  betrachten und im obigen Satz  $\zeta = t \wedge \tau$  und  $\sigma = s$  setzen.

**Korollar 2.6** (Optionales Stoppen). *Sei  $M$  ein rechtsstetiges Martingal bezüglich einer Filtration  $\mathbb{F}$  welche den üblichen Bedingungen genügt und  $\tau$  eine  $\mathbb{F}$ -Stoppzeit. Dann ist auch der gestoppte Prozess*

$$M_t^\tau := M_{\tau \wedge t}$$

ein rechtsstetiges  $\mathbb{F}$ -Martingal. Die Aussage gilt analog für Sub- und Supermartingale.

*Beweis.* Wir schreiben  $T > 0$  für die Konstante die  $\zeta$  fast sicher von oben beschränkt und definieren für  $n \in \mathbb{N}$  die dyadische Partition

$$S(n) := \{k/2^n : k \in \mathbb{N}\} \subset \mathbb{R}_{\geq 0}.$$

Für jedes  $\omega \in \Omega$  bezeichne  $\zeta_n(\omega)$  bzw.  $\sigma_n(\omega)$  das kleinste Element von  $S(n)$  welches grösser als  $\zeta(\omega)$  bzw.  $\sigma(\omega)$  ist. Es ist klar, dass sich auch  $\zeta_n$  durch deterministische  $T_n \in S(n)$  beschränken lassen, welche so gewählt werden können, dass  $T_n \downarrow T$  gilt. Schreibe  $(M^n, \mathbb{F}^n)$  für das Martingal  $M$  und die Filtration  $\mathbb{F}$  eingeschränkt auf Zeitpunkte in  $S(n)$ . Jedes  $M^n$  ist ein  $\mathbb{F}^n$ -Martingal mit diskretem Zeitparameter; folglich ist auch  $|M^n|$  ein  $\mathbb{F}^n$ -Submartingal in diskreter Zeit. Weiters gilt für jedes  $t_n \in S(n)$  dass

$$\{\zeta_n \leq t_n\} = \{\zeta \leq t_n\} \in \mathcal{F}_{t_n} = \mathcal{F}_{t_n}^n,$$

sodass  $\zeta_n$  eine  $\mathbb{F}^n$ -Stoppzeit ist; genauso wie  $\sigma_n$ . Mit dem Stoppsatz aus diskreter Zeit angewandt auf das Submartingal  $|M^n|$  folgt nun

$$\mathbb{E}[|M_{\zeta_n}^n|] = \mathbb{E}[|M_{\zeta_n}^n|] \leq \mathbb{E}[\mathbb{E}[|M_{T_n}^n| | \zeta_n]] = \mathbb{E}[|M_{T_n}^n|] = \mathbb{E}[|M_{T_n}|] < \infty. \quad (2.2)$$

Der Grenzübergang  $n \rightarrow \infty$  liefert aufgrund der Rechtsstetigkeit von  $M$  und dem Lemma von Fatou, dass  $E|M_\zeta| < \infty$ . Analog zu (2.2) können wir den Stoppsatz aus

diskreter Zeit auch auf das Martingal  $M^n$  direkt anwenden und erhalten

$$\mathbb{E}[M_{\zeta_n} | \mathcal{F}_{\sigma_n}] = \mathbb{E}[M_{\zeta_n}^n | \mathcal{F}_{\sigma_n}] = M_{\sigma_n \wedge \tau_n}^n = M_{\sigma_n \wedge \tau_n}. \quad (2.3)$$

Nach Definition der bedingten Erwartung ist dies gleichbedeutend mit  $\mathbb{E}[M_{\zeta_n} A] = \mathbb{E}[M_{\sigma_n \wedge \zeta_n} A]$  für alle  $A \in \mathcal{F}_{\sigma_n}$ ,  $n \in \mathbb{N}$ . Aufgrund von Lemma 1.5 gilt  $\mathcal{F}_\sigma = \bigcap_{n \in \mathbb{N}} \mathcal{F}_{\sigma_n}$ , wir können  $A \in \mathcal{F}_{\sigma_n}$  durch  $A \in \mathcal{F}_\sigma$  ersetzen und erhalten

$$\mathbb{E}[M_{\zeta_n} | \mathcal{F}_\sigma] = \mathbb{E}[M_{\sigma_n \wedge \zeta_n} | \mathcal{F}_\sigma]. \quad (2.4)$$

Aufgrund der Rechtsstetigkeit von  $M$  gilt  $\lim_{n \rightarrow \infty} M_{\zeta_n} = M_\zeta$  fast sicher; analog für  $\sigma_n$ . Es bleibt zu zeigen, dass auch die bedingten Erwartungswerte in (2.4) konvergieren. Nach Lemma 1.11 können wir eine konvexe Funktion  $\Phi$  mit  $\lim_{x \rightarrow \infty} \frac{\Phi(x)}{x} = \infty$  finden, sodass  $\mathbb{E}[\Phi(M_{T_0})] < \infty$ . Wir betrachten wieder das diskrete Submartingal  $\Phi(M^n)$  bezüglich der diskreten Filtration  $\mathbb{F}^n$  und erhalten analog zu (2.5), dass

$$\mathbb{E}[\Phi(M_{\zeta_n})] \leq \mathbb{E}[\Phi(M_{T_0})] < \infty \quad (2.5)$$

für alle  $n \in \mathbb{N}$ . Es folgt mit Lemma 1.10 die gleichgradige Integrierbarkeit sowohl von  $(M_{\zeta_n})_{n \in \mathbb{N}}$  als auch von  $(M_{\sigma_n \wedge \zeta_n})_{n \in \mathbb{N}}$  und damit nach Korollar 1.9 die Konvergenz der bedingten Erwartungswerte  $\mathbb{E}[M_{\zeta_n} | \mathcal{F}_\sigma] \rightarrow \mathbb{E}[M_\zeta | \mathcal{F}_\sigma]$  und  $\mathbb{E}[M_{\sigma_n \wedge \zeta_n} | \mathcal{F}_\sigma] \rightarrow \mathbb{E}[M_{\sigma \wedge \zeta} | \mathcal{F}_\sigma] = M_{\sigma \wedge \zeta}$ . Die Behauptung folgt nun aus (2.4) mit  $n \rightarrow \infty$ .  $\square$

### 2.3 Doob'sche Ungleichungen und Martingalkonvergenz in stetiger Zeit

Gegeben einen rechtsstetigen stochastischen Prozess  $M$  definieren wir den zugehörigen Supremumssprozess

$$M_t^* = \sup_{0 \leq s \leq t} M_s.$$

Wie klar ersichtlich ist hängt der Supremumssprozess vom gesamten Pfad von  $M$  bis zum Zeitpunkt  $t$  ab. Die erstaunliche Aussage der Doob'schen Ungleichung ist nun dass im Fall dass  $M$  ein Submartingal ist, sich diese pfadabhängige Größe alleine durch die Randverteilung von  $M$  zum Zeitpunkt  $t$  kontrollieren lässt.

**Theorem 2.7** (Doob'sche Ungleichungen). *Sei  $M$  ein rechtsstetiges Submartingal und  $M_t^* = \sup_{0 \leq s \leq t} M_s$  der zugehörige Maximumprozess.*

(a) Für alle  $\lambda > 0$  und  $T \geq 0$  gilt

$$\mathbb{P}[M_T^* \geq \lambda] \leq \frac{1}{\lambda} \mathbb{E}[M_T^+].$$

(b) Ist  $M$  zusätzlich positiv oder Martingal, dann gilt auch

$$\|(M_T^*)\|_p \leq \frac{p}{p-1} \|M_T\|_p$$

für alle  $p > 1$ .

*Beweis.* Definiere für  $n \in \mathbb{N}$  die ‘dyadische Partition’

$$S(n, T) := \{Tk/2^n : k \in \{0, \dots, 2^n\}\} \subseteq [0, T].$$

Aufgrund der Rechtssteigkeit von  $M$  gilt  $\lim_{n \rightarrow \infty} \sup_{t \in S(n, T)} M_t = M_T^*$  fast sicher. Die Doob’sche Ungleichung für diskrete Submartingale liefert

$$\lambda \mathbb{P} \left[ \sup_{t \in S(n, T)} M_t \geq \lambda \right] \leq \mathbb{E} [M_T^+]$$

und Behauptung (a) folgt nach Anwendung von Fatou’s Lemma. Analog lässt sich (b) aus der Doob’schen  $L^p$ -Ungleichung in diskreter Zeit herleiten.  $\square$

Als nächstes Resultat zeigen wir den Konvergenzsatz für Martingale. Wie in diskreter Zeit gibt es eine Version für  $L^1$ -beschränkte Martingale und eine für  $L^p$ -beschränkte ( $p > 1$ ) mit einer etwas stärkeren Folgerung.

**Theorem 2.8** (Martingalkonvergenz). *Sei  $M$  ein rechtsstetiges Submartingal, sodass  $\mathbb{E} [M_t^+] \leq B < \infty$  für alle  $t \in \mathbb{R}_{\geq 0}$ . Dann existiert eine Zufallsvariable  $M_\infty$  mit  $\mathbb{E} [|M_\infty|] < \infty$  und es gilt*

$$\lim_{t \rightarrow \infty} M_t = M_\infty \quad f.s.$$

*Beweis.* Wegen der Rechtsstetigkeit von  $M$  gilt

$$\liminf_{t \rightarrow \infty} M_t = \liminf_{q \in \mathbb{Q}_+, q \rightarrow \infty} M_q \quad \text{und} \quad \limsup_{t \rightarrow \infty} M_t = \limsup_{q \in \mathbb{Q}_+, q \rightarrow \infty} M_q.$$

Falls der Grenzwert  $\lim_{t \rightarrow \infty} M_t$  also mit Wahrscheinlichkeit  $> 0$  nicht als erweiterte reelle Zahl in  $[-\infty, \infty]$  existiert, so lassen sich  $a < b$  in  $\mathbb{Q}_+$  finden, sodass

$$\liminf_{q \in \mathbb{Q}_+, q \rightarrow \infty} M_t < a < b < \limsup_{q \in \mathbb{Q}_+, q \rightarrow \infty} M_t. \quad (2.6)$$

mit strikt positiver Wahrscheinlichkeit. Wir bezeichnen mit  $U_N[a, b](M(\omega)|\mathbb{Q}_+)$  die aufsteigenden Überschreitungen des Intervalls  $[a, b]$  durch die Funktion  $M(\omega) : \mathbb{Q}_+ \rightarrow \mathbb{R}, q \mapsto M_q(\omega)$  bis zum Zeitpunkt  $N$ . Sei weiters  $(D_k)_{k \in \mathbb{N}}$  eine aufsteigende Folge von endlichen Teilmengen von  $\mathbb{Q}_+$ , sodass  $\mathbb{Q}_+ = \bigcup_{k \in \mathbb{N}} D_k$ . Dann gilt aufgrund von

monotoner Konvergenz und Lemma 2.4

$$\begin{aligned} \mathbb{E}[U_\infty[a, b](M(\omega)|\mathbb{Q}_+)] &= \lim_{N \rightarrow \infty} \lim_{k \rightarrow \infty} \mathbb{E}[U_N[a, b](M(\omega)|D(k))] \leq \\ &\leq \lim_{N \rightarrow \infty} \frac{\mathbb{E}[(M_N - a)^+] + \mathbb{E}[(M_0 - a)^+]}{b - a} \leq 2 \frac{\max(B, a)}{b - a} < \infty. \end{aligned}$$

Daraus folgt dass die Wahrscheinlichkeiten der Ereignisse  $\{\omega : U_\infty[a, b](M(\omega)|\mathbb{Q}_+) = \infty\}$  für alle  $a < b$  in  $\mathbb{Q}_+$  gleich Null sind. Nach Bilden einer abzählbaren Vereinigung erhalten wir

$$\mathbb{P} \left[ \bigcup_{a < b; a, b \in \mathbb{Q}_+} \{\omega : U_\infty[a, b](M(\omega)|\mathbb{Q}_+) = \infty\} \right] = 0,$$

im Widerspruch zu (2.6). Folglich existiert  $\lim_{t \rightarrow \infty} M_t =: M_\infty$  fast sicher in  $[-\infty, \infty]$ . Es bleibt zu zeigen, dass  $\mathbb{E}[|M_\infty|] < \infty$ . Dazu wenden wir zweimal das Lemma von Fatou an: Aus

$$\mathbb{E}[M_\infty^+] \leq \liminf_{t \rightarrow \infty} \mathbb{E}[M_t^+] \leq B$$

und

$$\mathbb{E}[M_\infty^-] \leq \liminf_{t \rightarrow \infty} \mathbb{E}[M_t^-] = \liminf_{t \rightarrow \infty} (\mathbb{E}[M_t^+] - \mathbb{E}[M_t]) \leq B - \mathbb{E}[M_0]$$

folgt die Behauptung.  $\square$

Bei genauerer Betrachtung des Konvergenzsatzes fällt auf, dass wir zwar eine Zufallsvariable  $M_\infty$  in  $L^1$  als Grenzwert erhalten, allerdings konvergiert  $M_t$  nur fast sicher, aber nicht unbedingt in  $L^1$ . Mit welcher zusätzlichen Bedingung lässt sich der Satz also verstärken um auch  $L^1$ -Konvergenz zu erhalten? Es stellt sich heraus dass gleichgradige Integrierbarkeit hinreichend und notwendig für  $L^1$ -Konvergenz ist.

**Theorem 2.9.** *Sei  $M$  ein rechtsstetiges  $\mathbb{F}$ -Martingal. Dann sind folgende Aussagen äquivalent*

- (a) *Es existiert  $M_\infty$  mit  $\mathbb{E}[|M_\infty|] < \infty$  sodass  $\lim_{t \rightarrow \infty} \|M_t - M_\infty\|_1 = 0$ .*
- (b) *Es existiert  $M_\infty$  mit  $\mathbb{E}[|M_\infty|] < \infty$  sodass  $M_t = \mathbb{E}[M_\infty | \mathcal{F}_t]$  für alle  $t \geq 0$ .*
- (c) *Die Familie  $(M_t)_{t \geq 0}$  ist gleichgradig integrierbar.*

*Der Satz gilt analog für positive Submartingale, wenn in (b) '=' durch ' $\leq$ ' ersetzt wird.*

*Beweis.* Der Beweis sei dem Leser überlassen und besteht im Wesentlichen aus der Kombination der oben bewiesenen Resultate zur Martingalkonvergenz und der Resultate zur gleichgradigen Integrierbarkeit aus Abschnitt 1.6.  $\square$

Als letztes zeigen wir Konvergenzsatz für  $L^p$ -beschränkte Martingale ( $p > 1$ ). Anders als in  $L^1$  fallen hier fast sichere und  $L^p$ -Konvergenz stets zusammen.

**Theorem 2.10** ( $L^p$ -Martingalkonvergenz). *Sei  $M$  ein rechtsstetiges Martingal, sodass  $\|M_t\|_p \leq B < \infty$  für ein  $p > 1$  und alle  $t \in \mathbb{R}_{\geq 0}$ . Dann existiert eine Zufallsvariable  $M_\infty$  mit  $\|M_\infty\|_p \leq B < \infty$ , sodass*

$$\lim_{t \rightarrow \infty} M_t = M_\infty \quad \text{f.s.} \quad \text{und} \quad \lim_{t \rightarrow \infty} \|M_t - M_\infty\|_p = 0.$$

Der Satz gilt auch unter der Voraussetzung, dass  $M$  ein rechtsstetiges, positives Submartingal ist.

*Beweis.* Aus der Beschränktheit der  $L^p$ -Norm folgt mit der Jensenschen Ungleichung, dass auch  $\mathbb{E}[|M_t|] \leq B$  gilt. Weiterhin gilt

$$\mathbb{E}[M_t^+] = \frac{\mathbb{E}[|M_t| + M_t]}{2} = \frac{B + \mathbb{E}[M_0]}{2},$$

sodass  $M$  den Voraussetzungen von Theorem 2.8 genügt. Es existiert also  $M_\infty$  mit  $\mathbb{E}[|M_\infty|] < \infty$  sodass  $M_t \rightarrow M_\infty$  fast sicher. Um auch die Konvergenz in  $L^p$  zu zeigen, schätzen wir zunächst mit

$$|M_t - M_\infty|^p \leq 2^{p-1} (|M_t|^p + |M_\infty|^p) \leq 2^p \sup_{t \geq 0} |M_t|^p$$

ab. Mit monotoner Konvergenz und der Doobschen  $L^p$ -Ungleichung gilt

$$\mathbb{E} \left[ \sup_{s \geq 0} |M_s|^p \right]^{1/p} = \lim_{t \rightarrow \infty} \|M_t^*\|_p \leq \frac{p}{p-1} \lim_{t \rightarrow \infty} \|M_t\|_p \leq \frac{p}{p-1} B < \infty.$$

Es folgt also mit dominierter Konvergenz, dass

$$\lim_{t \rightarrow \infty} \mathbb{E}[|M_t - M_\infty|^p] = 0$$

und insbesondere dass  $\|M_\infty\|_p \leq B$ . Der Beweis für Submartingale ist analog zu führen.  $\square$

## 2.4 Martingale und Brownsche Bewegung

Einige wichtige Beispiele für Martingale lassen sich aus der Brownschen Bewegung ableiten.

**Theorem 2.11.** *Die folgenden Prozesse sind stetige Martingale bezüglich der Brownschen Standardfiltration:*

(a) Die Brownsche Bewegung  $B_t$  selbst;

(b)  $B_t^2 - t$ ;

(c)  $\exp\left(\vartheta B_t - \frac{\vartheta^2}{2}t\right)$  für jedes  $\vartheta \in \mathbb{R}$ .

*Beweis.* Der Beweis von (a) und (b) sei dem Leser überlassen. Aussage (c) folgt mit  $M_t = \exp\left(\vartheta B_t - \frac{\vartheta^2}{2}t\right)$  aus

$$\mathbb{E}[M_t | \mathcal{F}_s] = M_s \mathbb{E}\left[\exp\left(\vartheta(B_t - B_s) - \frac{\vartheta^2}{2}(t - s)\right) \middle| \mathcal{F}_s\right] = M_s,$$

wobei verwendet wurde dass  $B_t - B_s$  unabhängig von  $\mathcal{F}_s$  ist.  $\square$

Wir wollen nun die in Theorem 2.11 definierten Martingale und die in den vorangehenden Abschnitten bewiesenen Resultate verwenden um folgende Fragen zum Verhalten der Brownschen Bewegung zu beantworten (wir nehmen dazu an, dass  $a, b > 0$ ):

(a) Wie groß ist die Wahrscheinlichkeit, dass die Brownsche Bewegung einen Wert  $a$  erreicht, bevor sie  $-b$  erreicht?

(b) Wie lange ist die durchschnittliche Wartezeit bis entweder  $a$  oder  $-b$  erreicht wird?

(c) Was ist die Verteilung der Eintrittszeit der Brownschen Bewegung in  $[a, \infty)$ ?

Sei also  $a, b > 0$  und sei

$$\tau := \inf\{t > 0 : B_t = a \text{ oder } B_t = -b\}$$

der erste Zeitpunkt zu dem die Brownsche Bewegung entweder  $a$  oder  $-b$  erreicht. Wir können  $\tau$  als Eintrittszeit in die abgeschlossene Menge  $(-\infty, -b] \cup [a, \infty)$  auffassen und  $\tau$  ist nach Lemma 1.4 eine Stoppzeit. Wir zeigen zunächst dass  $\tau$  fast sicher endlich ist und sogar endliche Momente beliebiger Ordnung besitzt. Für  $n \in \mathbb{N}$  definiere die Ereignisse  $E_n = \{|B_{n+1} - B_n| > a + b\}$ . Wegen der Unabhängigkeit und Stationarität der Inkremente von  $B$  gilt

$$\mathbb{P}[\tau > n + 1] \leq \mathbb{P}\left[\bigcap_{k=0}^{n-1} E_k^c\right] = (1 - \epsilon)^n,$$

mit  $\epsilon = \mathbb{P}[E_n] > 0$ . Für beliebiges  $K \in \mathbb{N}$  schätzen wir mit

$$\mathbb{E}[\tau^K] \leq K \int_0^\infty x^{K-1} \mathbb{P}[\tau > x] dx \leq K \sum_{i=1}^\infty n^{K-1} (1 - \epsilon)^{n-1}$$

ab. Das Quotientenkriterium zeigt dass die Reihe unabhängig von  $K$  konvergiert und somit dass  $\mathbb{E}[\tau^K] < \infty$ . Insbesondere ist  $\tau$  fast sicher endlich.

Nach Definition von  $\tau$  gilt nun

$$\mathbb{E}[B_\tau] = a \mathbb{P}[B_\tau = a] - b(1 - \mathbb{P}[B_\tau = a]). \quad (2.7)$$

Da die Brownsche Bewegung ein Martingal ist, ist auch der gestoppte Prozess  $B_{t \wedge \tau}$  ein Martingal. Dominierte Konvergenz liefert

$$\mathbb{E}[B_\tau] = \lim_{t \rightarrow \infty} \mathbb{E}[B_{t \wedge \tau}] = 0.$$

Die Gleichung (2.7) lässt sich nun nach  $\mathbb{P}[B_\tau = a]$  auflösen und wir erhalten  $\mathbb{P}[B_\tau = a] = \frac{b}{a+b}$ . Die Antwort auf die erste Frage lautet also

- Die Wahrscheinlichkeit der Brownschen Bewegung  $a$  vor  $-b$  zu erreichen ist  $\frac{b}{a+b}$ .

Wir wenden die selben Überlegungen auf das Martingal  $M_t = B_t^2 - t$  an. Es gilt

$$\mathbb{E}[M_\tau] = a^2 \mathbb{P}[B_\tau = a] + b^2 \mathbb{P}[B_\tau = b] - \mathbb{E}[\tau] = ab - \mathbb{E}[\tau]. \quad (2.8)$$

Gleichzeitig gilt aufgrund der Martingaleigenschaft des gestoppten Prozesses  $M_{t \wedge \tau}$

$$\mathbb{E}[M_\tau] = \lim_{t \rightarrow \infty} \mathbb{E}[M_{t \wedge \tau}] = M_0 = 0$$

und wir erhalten als Antwort auf unsere zweite Frage

- Die durchschnittliche Wartezeit bis zum Erreichen von  $a$  oder  $-b$  ist  $\mathbb{E}[\tau] = ab$ .

Wir wechseln nun von der zweiseitigen Eintrittszeit  $\tau$  zur einseitigen Eintrittszeit

$$\tau_a = \inf \{t > 0 : B_t = a\}. \quad (2.9)$$

Wir berechnen die Laplace-Transformierte  $\mathbb{E}[e^{-u\tau_a}]$ , ( $u \geq 0$ ) welche bekanntlich die Verteilung von  $\tau_a$  eindeutig bestimmt. Zuerst zeigen wir, dass  $\tau_a$  fast sicher endlich ist. Aus

$$\mathbb{P}[\tau_a < \infty] \geq \mathbb{P}[B_\tau = a] = \frac{b}{b+a}$$

folgt mit  $b \rightarrow \infty$ , dass  $\mathbb{P}[\tau_a < \infty] = 1$ . Als nächstes betrachten wir das dritte Martingal aus Theorem 2.11,  $M_t^\vartheta = \exp(\vartheta B_t - \frac{\vartheta^2}{2}t)$  und erhalten

$$\mathbb{E}[M_{\tau_a}^\vartheta] = \mathbb{E}\left[\exp\left(\vartheta a - \frac{\vartheta^2}{2}\tau_a\right)\right] = e^{\vartheta a} \mathbb{E}\left[e^{-\frac{\vartheta^2}{2}\tau_a}\right] \quad (2.10)$$

für alle  $\vartheta \in \mathbb{R}$ . Gleichzeitig gilt

$$\mathbb{E}[M_{\tau_a}^\vartheta] = \lim_{t \rightarrow \infty} \mathbb{E}[M_{t \wedge \tau_a}^\vartheta] = M_0 = 1.$$

Setzen wir nun  $\theta = \sqrt{2u}$  in (2.10) und formen um, erhalten wir die Laplace-Transformierte  $\mathbb{E}[e^{-u\tau_a}] = e^{-a\sqrt{2u}}$ . Durch nachrechnen (oder nachschlagen in der passenden Tabelle von Laplace-Integralen) lässt sich folgende Antwort leicht überprüfen

- Die Verteilung der einseitigen Eintrittszeit  $\tau_a$  hat die Dichte

$$f_{\tau_a}(t) = \frac{a}{\sqrt{2\pi t^3}} \exp\left(-\frac{a^2}{2t}\right), \quad (t \geq 0).$$

Indem wir die Laplace-Transformierte ableiten lässt sich ausserdem leicht der Erwartungswert von  $\tau_a$  berechnen:

$$\mathbb{E}[\tau_a] = -\lim_{u \rightarrow 0} \frac{\partial}{\partial u} \mathbb{E}[e^{-u\tau_a}] = \lim_{u \rightarrow 0} \frac{a}{\sqrt{2u}} e^{-a\sqrt{2u}} = \infty.$$



### 3 Stochastische Integration

Auch das stochastische Integral lässt sich als stetige Verallgemeinerung der *Martingaltransformation* aus diskreter Zeit motivieren. Prototyp des Martingals in diskreter Zeit ist die symmetrische Irrfahrt  $S_n = \sum_{k=1}^n X_k$  wobei  $(X_k)_{k \in \mathbb{N}}$  unabhängige Zufallsvariablen sind, die mit Wahrscheinlichkeit von jeweils  $1/2$  die Werte  $\pm 1$  annehmen. Das Martingal  $S_n$  lässt sich als Gewinn/Verlust aus einem fairen Spiel nach  $n$  Runden interpretieren. So könnte  $X_n$  beispielsweise einem Münzwurf entsprechen wobei bei Kopf Spieler  $A$  an Spieler  $B$  einen Euro zahlt, und bei Zahl Spieler  $B$  an Spieler  $A$ . Erlauben wir nun Spieler  $A$  den Einsatz in jeder Runde frei zu wählen (bis zu einer absoluten Obergrenze  $K$ ) und auch nach Belieben von Kopf nach Zahl und zurück zu wechseln, so lässt sich dies mathematisch durch festlegen einer (Spiel-)Strategie  $(A_k)_{k \in \mathbb{N}}$  festlegen, wobei jedes  $A_k$  eine  $\mathcal{F}_{k-1}$ -messbare Zufallsgrösse ist. Das bedeutet, dass die Strategie noch nicht vor Beginn des Spiels festgelegt werden muss, aber dass der Einsatz für Runde  $k$  nur die Information aus Runde 1 bis  $k-1$  nutzen kann. Der Gewinn aus dem Spiel mit Strategie  $A$  lässt sich dann als

$$(A \bullet S)_n = \sum_{k=1}^n A_k(S_k - S_{k-1}), \quad (A_k \in \mathcal{F}_{k-1}) \quad (3.1)$$

schreiben. Es ist sehr leicht zu zeigen dass  $(A \bullet S)$  ebenfalls ein Martingal ist; ein faires Spiel lässt sich also auch durch eine noch so ausgeklügelte Strategie nicht in ein vorteilhaftes Spiel verwandeln. In stetiger Zeit tritt die Brownsche Bewegung  $B$  an Stelle der Irrfahrt  $S$  und ein stochastischer Prozess  $f(\omega, t)$  an Stelle der Strategie  $A$ . Ziel ist es als stetige Verallgemeinerung von (3.1) ein ‘stochastisches Integral’

$$I(f)(\omega) = \int_0^T f(\omega, t) dB_t(\omega) \quad (3.2)$$

zu definieren. Als Minimalanforderungen an einen sinnvollen Integralbegriff legen wir fest:

- Linearität:  $I(\alpha f + \beta g) = \alpha I(f) + \beta I(g)$  für alle  $\alpha, \beta \in \mathbb{R}$ .
- Für die Indikatorfunktion  $f(\omega, t) = \mathbf{1}\{a < t \leq b\}$  soll gelten

$$I(f) = \int_0^T \mathbf{1}\{a < t \leq b\} dB_t = \int_a^b dB_t = B_b - B_a.$$

- Zumindest jeder beschränkte, adaptierte und stückweise stetige stochastische Prozess  $f$  soll sich gegen die Brownsche Bewegung integrieren lassen.

Aus (a) und (b) ergibt sich, dass für jeden Integranden der Form

$$f(\omega, t) := \sum_{i=0}^{n-1} a_i(\omega) \mathbf{1}_{\{t_i < t \leq t_{i+1}\}}$$

gilt, dass

$$I(f)(\omega) = \sum_{i=0}^{n-1} a_i(\omega) (B_{t_{i+1}} - B_{t_i}). \quad (3.3)$$

Wenn wir nur adaptierte Integranden  $f$  zulassen, dann ist notwendigerweise  $a_i \in \mathcal{F}_{t_i}$  und (3.3) entspricht genau der diskreten Martingaltransformation (3.1). Wie lässt sich das Integral  $I$  nun auf allgemeine Integranden  $f$  ausdehnen? Es zeigt sich dass gerade der naheliegendste ‘pfadweise’ Ansatz problematisch ist.

### 3.1 Probleme der Pfadweisen Integration

Der naheliegendste (aber letztlich nicht zielführende) Ansatz zur Definition des stochastischen Integrals (3.2) für allgemeine (stetige, beschränkte und adaptierte) Integranden ist es pfadweise zu betrachten, also für jedes  $\omega \in \Omega$  die Funktion  $t \mapsto f(\omega, t)$  gegen das ‘Signal’  $t \mapsto B(t, \omega)$  zu integrieren. Der zugehörige aus der deterministischen Welt bekannte Integralbegriff ist das Riemann-Stieltjes-Integral

**Definition.** Seien  $f$  und  $g$  Funktionen von  $[0, T]$  nach  $\mathbb{R}$ . Zusätzlich sei  $\mathcal{P}^n$  eine Folge von Zerlegungen  $0 = t_0^n < t_1^n < \dots < t_N^n = T$  mit Netzweite  $|\mathcal{P}^n| = \sup_i |t_{i+1}^n - t_i^n|$  und  $|\mathcal{P}^n| \rightarrow 0$  und  $\mathcal{B}^n : (b_i^n)_{i=1 \dots N}$  eine zugehörige Folge von Belegungen  $b_i^n \in [t_i^n, t_{i+1}^n]$ . Wenn der Limes

$$\lim_{n \rightarrow \infty} \sum_{i=0}^{N-1} f(b_i^n) (g(t_{i+1}^n) - g(t_i^n)),$$

unabhängig von der Wahl der Zerlegungen  $\mathcal{P}^n$  und Belegungen  $\mathcal{B}^n$  existiert und denselben Wert annimmt, so heisst dieser Wert das **Riemann-Stieltjes-Integral**  $\int_0^T f dg$  von  $f$  gegen  $g$ .

Es ist bekannt, dass zumindest für stetige Integranden  $f$  und Signale  $g$  von beschränkter Totalvariation das Riemann-Stieltjes-Integral  $\int_0^T f dg$  stets existiert. Für die Brownsche Bewegung gilt jedoch folgendes:

**Lemma 3.1.** *Die Pfade der Brownschen Bewegung haben fast sicher unendliche Totalvariation: Sei  $(\mathcal{P}^n)_{n \in \mathbb{N}}$  eine Folge von Zerlegungen von  $[0, T]$  mit  $|\mathcal{P}^n| \rightarrow 0$ , dann existiert  $A \in \mathcal{F}$  mit  $\mathbb{P}[A] = 1$ , sodass*

$$\lim_{n \rightarrow \infty} \sum_{t_i \in \mathcal{P}_n} |B_{t_{i+1}}(\omega) - B_{t_i}(\omega)| = \infty, \quad \forall \omega \in A.$$

*Beweis.* Wir zeigen das Resultat für die Folge  $(\mathcal{P}_n)$  der dyadischen Partitionen  $\mathcal{P}_n = \{Tk2^{-n} : k = 0, \dots, 2^n\}$ , der allgemeine Beweis lässt sich ähnlich führen. Definiere

$$V_n := \sum_{i=0}^{2^n-1} |B_{T(i+1)2^{-n}} - B_{Ti2^{-n}}|.$$

Dann gilt unter Verwendung der Skalierungseigenschaft der Brownschen Bewegung

$$\mathbb{E}[V_n] = 2^n \mathbb{E}[|B_{T2^{-n}}|] = 2^n \sqrt{T} 2^{-n/2} \mathbb{E}[|B_1|] = 2^{n/2} \sqrt{\frac{2T}{\pi}}.$$

Des weiteren gilt

$$\text{Var}(V_n^2) = T,$$

und mit der Ungleichung von Tschebyscheff folgt

$$\mathbb{P}\left[\left|V_n - 2^{n/2} \sqrt{\frac{2T}{\pi}}\right| \geq \lambda\right] \leq \frac{T}{\lambda^2}, \quad \forall \lambda > 0.$$

Setzen wir nun  $\lambda = \frac{\sqrt{T}}{2} 2^{n/2}$  erhalten wir

$$\mathbb{P}\left[\left|V_n - 2^{n/2} \sqrt{\frac{2T}{\pi}}\right| \geq \frac{\sqrt{T}}{2} 2^{n/2}\right] \leq \frac{4}{2^n}.$$

Diese Wahrscheinlichkeiten sind über  $n \in \mathbb{N}$  summierbar und mit dem Lemma von Borel-Cantelli folgt, dass eine Menge  $A \in \mathcal{F}$  mit  $\mathbb{P}[A] = 1$  existiert, sowie zu jedem  $\omega \in A$  ein  $N(\omega) \in \mathbb{N}$ , sodass

$$2^{n/2} \sqrt{T} \left(\sqrt{\frac{2}{\pi}} - \frac{1}{2}\right) \leq V_n \leq 2^{n/2} \sqrt{T} \left(\sqrt{\frac{2}{\pi}} + \frac{1}{2}\right), \quad \forall n \geq N(\omega), \omega \in A.$$

Es folgt  $V_n \rightarrow \infty$  fast sicher, wie behauptet.  $\square$

Ein Ausweg würde sich ergeben, wenn endliche Totalvariation zwar eine hinreichende, aber nicht unbedingt notwendige Bedingung zur Existenz des Riemann-Stieltjes-Integrals wäre. Diese Hoffnung ist aber ebenfalls trügerisch, wie folgendes Resultat zeigt.

**Theorem 3.2.** *Sei  $g : [0, T] \rightarrow \mathbb{R}$  eine Funktion von unbeschränkter Totalvariation. Dann gibt es eine stetige beschränkte Funktion  $f : [0, T] \rightarrow \mathbb{R}$ , sodass das Riemann-Stieltjes-Integral  $\int_0^T f dg$  nicht existiert.*

Für den Beweis verwenden wir den aus der Funktionalanalysis bekannten Satz von

Banach-Steinhaus über die gleichmässige Beschränktheit.

**Theorem 3.3** (Banach-Steinhaus). *Sei  $X$  ein Banachraum,  $Y$  ein normierter Vektorraum und  $\{T_\alpha\}_{\alpha \in I}$  eine Familie von beschränkten linearen Operatoren von  $X$  nach  $Y$ . Wenn die Menge  $\{T_\alpha x\}_{\alpha \in I}$  für jedes  $x \in X$  beschränkt ist, dann ist auch  $\{\|T_\alpha\|\}_{\alpha \in I}$  beschränkt.*

Hierbei ist  $\|T_\alpha\|$  die übliche Operatornorm  $\|T_\alpha\| = \sup\{\|Tx\| / \|x\| : x \in X\}$ .

*Beweis von Theorem 3.2.* Wir wählen  $X$  als den Banachraum der beschränkten stetigen Funktionen auf  $[0, T]$ , ausgestattet mit der Supremumsnorm.  $Y$  sei der Vektorraum der reellen Zahlen mit der euklidischen Norm, und  $(\mathcal{P}^n)_{n \in \mathbb{N}}$  eine Folge von Zerlegungen von  $[0, T]$  mit  $|\mathcal{P}^n| \rightarrow 0$ . Dann ist für jede Zerlegung  $\mathcal{P}$  mit  $N$  Teilungspunkten ein beschränktes lineares Funktional  $T_\mathcal{P} : X \rightarrow Y$  durch

$$T_\mathcal{P} f := \sum_{i=0}^{N-1} f(t_i) (g(t_{i+1}) - g(t_i)), \quad f \in X$$

definiert. Des weiteren können wir für jede Zerlegung  $\mathcal{P}$  ein  $f^* \in X$  finden, sodass  $f^*(t_i) = \text{sign}(g(t_{i+1}) - g(t_i))$ , wobei sich  $f^*$  so wählen lässt, dass  $\|f^*\| = 1$ . Die Funktion  $f^*$  erfüllt

$$T_\mathcal{P} f^* = \sum_{i=1}^n |g(t_{i+1}) - g(t_i)|,$$

das heisst  $T_\mathcal{P} f^*$  entspricht genau der absoluten Variation von  $g$  entlang der Partition  $\mathcal{P}$ . Da  $g$  unendliche Totalvariation hat, ist also  $\|T_{\mathcal{P}^n}\|_{n \in \mathbb{N}}$  unbeschränkt. Mit dem Satz von Banach-Steinhaus folgt, dass ein  $f \in X$  existiert, sodass auch  $|T_{\mathcal{P}^n} f|_{n \in \mathbb{N}}$  unbeschränkt ist. Es kann also das Riemann-Stieltjes-Integral  $\int_0^T f dg$  nicht existieren und die Behauptung ist bewiesen.  $\square$

Was bleibt also für ein Ausweg um doch noch ein stochastisches Integral zu definieren, das die Wunschliste am Anfang des Kapitels erfüllt? Aus einer funktionalanalytischen Perspektive suggeriert der eben bewiesene Satz, dass wir auf  $X$  (dem Raum der Integranden) beziehungsweise  $Y$  (dem Wertebereich des Integrals) die falschen Topologien gewählt haben. Ein möglicher Ausweg ist also die Topologie auf dem Wertebereich  $Y$  zu ändern. Hier können wir insbesondere die pfadweise Perspektive aufgeben und nutzen dass Werte der ‘elementaren Integrale’ in (3.3) Zufallsvariablen sind, und wir für Zufallsvariablen noch andere Konvergenzarten als die fast-sichere Konvergenz kennen. Die Topologie der  $L^2$ -Konvergenz führt zum Begriff des *Ito-Integrals*, den wir hier weiterverfolgen wollen.

### 3.2 Ito-Isometrie & Ito-Integral

Wir setzen ab nun stets einen Wahrscheinlichkeitsraum  $(\Omega, \mathcal{F}, \mathbb{P})$  mit Brownscher Bewegung  $B$  als gegeben voraus und schreiben  $\mathbb{F}$  für die Brownsche Standardfiltration. Des weiteren setzen wir einen fixen Zeithorizont  $T > 0$  fest.

**Definition.** Es sei  $\mathcal{H}^2[0, T]$  der Vektorraum aller  $\mathbb{F}$ -adaptierten stochastischen Prozesse  $f$ , welche die Integrierbarkeitsbedingung

$$\mathbb{E} \left[ \int_0^T f(\omega, t)^2 dt \right] < \infty \quad (3.4)$$

erfüllen.

Der linke Seite von (3.4) lässt sich als  $L^2$ -norm des Hilbertraums  $L^2(d\mathbb{P} \otimes dt)$  über  $(\Omega, \mathcal{F}_T, \mathbb{P})$  lesen. Demnach ist  $\mathcal{H}^2[0, T]$  also ein linearer Unterraum von  $L^2(d\mathbb{P} \otimes dt)$ . Es wird sich herausstellen dass alle Elemente von  $\mathcal{H}^2[0, T]$  geeignete Integranden für das zu definierenden Ito-Integral sind. Wir betrachten einen weiteren Teilraum.

**Definition.** Es sei  $\mathcal{H}_0^2[0, T]$  der Vektorraum aller stückweise konstanten Funktionen in  $\mathcal{H}^2[0, T]$ , d.h.  $f \in \mathcal{H}_0^2[0, T]$  wenn eine Zerlegung  $0 = t_0 < t_1 < \dots < t_n = T$  existiert, sodass

$$f(\omega, t) = \sum_{i=0}^{n-1} a_i(\omega) \mathbf{1}_{\{t \in (t_i, t_{i+1}]\}}. \quad (3.5)$$

Wir nennen  $f \in \mathcal{H}_0^2[0, T]$  auch **einfache** oder **elementare Integranden**.

Da  $f$  adaptiert sein und die Integrierbarkeitsbedingung (3.4) erfüllen muss, folgt automatisch  $a_i \in \mathcal{F}_{t_i}$  und  $\mathbb{E}[a_i^2] < \infty$  für alle  $i \in \{0, \dots, n-1\}$ . Für  $f \in \mathcal{H}_0^2[0, T]$  lässt sich das Ito-Integral  $I(f)$  – entsprechend unseren Minimalforderungen am Anfang des Kapitels – explizit definieren und wir legen fest

$$I(f) = \sum_{i=0}^{n-1} a_i (B_{t_{i+1}} - B_{t_i}), \quad \forall f \in \mathcal{H}_0^2[0, T].$$

Wir machen folgende wichtige Beobachtung:

**Lemma 3.4** (Ito-Isometrie auf  $\mathcal{H}_0^2$ ). *Das elementare Ito-Integral  $I$  ist eine Isometrie von  $\mathcal{H}_0^2 \subset L^2(d\mathbb{P} \otimes dt)$  in  $L^2(d\mathbb{P})$ , das heisst*

$$\|I(f)\|_{L^2(d\mathbb{P})} = \|f\|_{L^2(d\mathbb{P} \otimes dt)} \quad \forall f \in \mathcal{H}_0^2.$$

*Beweis.* Für die  $L^2$ -Norm von  $I(f)$  gilt

$$\|I(f)\|_{L^2(dP)}^2 = \mathbb{E} \left[ \left( \sum_{i=0}^{n-1} a_i (B_{t_{i+1}} - B_{t_i}) \right)^2 \right].$$

Beim Quadrieren der Summe lässt sich ausnutzen dass aufgrund der Martingaleigenschaft der Brownschen Bewegung alle gemischten Produktterme unter dem Erwartungswert verschwinden: Es gilt für  $i < j$  dass

$$\mathbb{E} [a_i a_j (B_{t_{i+1}} - B_{t_i}) (B_{t_{j+1}} - B_{t_j})] = \mathbb{E} [a_i a_j (B_{t_{i+1}} - B_{t_i}) \mathbb{E} [(B_{t_{j+1}} - B_{t_j}) | \mathcal{F}_{t_j}]]] = 0.$$

Außerdem ist

$$\mathbb{E} [a_i^2 (B_{t_{i+1}} - B_{t_i})^2] = \mathbb{E} [a_i^2] (t_{i+1} - t_i),$$

und somit

$$\|I(f)\|_{L^2(dP)}^2 = \sum_{i=0}^{n-1} \mathbb{E} [a_i^2] (t_{i+1} - t_i) = \|f\|_{L^2(dP \times dt)}^2. \quad \square$$

Man beachte dass beim obigen Beweis die  $\mathbb{F}$ -Adaptiertheit von  $f$  entscheidend eingeht. Nächstes Ziel ist es das Ito-Integral  $I$  von  $\mathcal{H}_0^2$  auf  $\mathcal{H}^2$  auszudehnen. Dazu ist das nächste Resultat hilfreich.

**Lemma 3.5.**  $\mathcal{H}_0^2$  ist dicht in  $\mathcal{H}^2$ , das heisst für jedes  $f \in \mathcal{H}^2$  existiert eine Folge  $f_n \in \mathcal{H}_0^2$  sodass

$$\|f - f_n\|_{L^2(dP \times dt)} \rightarrow 0.$$

Den Beweis führen wir hier nicht, es sei auf Steele [2003, Kap. 6.6] verwiesen. Für  $f \in \mathcal{H}^2$  definieren wir das Ito-Integral  $I(f)$  folgendermassen:

- Wähle eine  $L^2$ -Approximationsfolge  $f_n \in \mathcal{H}_0^2$  mit  $f_n \rightarrow f$  gemäss Lemma 3.5
- Setze  $I(f)$  gleich dem Limes von  $I(f_n)$  in  $L^2(d\mathbb{P})$ .

Es bleibt natürlich zu zeigen, dass  $I(f)$  auf diese Weise wohldefiniert ist, also stets existiert und eindeutig ist. Der Nachweis ist einfach:

- **Existenz:** Nachdem  $f_n$  eine  $L^2$ -konvergente Folge ist, ist es auch eine Cauchy-folge, d.h. für jedes  $\epsilon > 0$  existiert  $N \in \mathbb{N}$  sodass  $\|f_n - f_m\|_{L^2(dP \times dt)} \leq \epsilon$  für alle  $n, m \geq N$ . Mit der Ito-Isometrie gilt

$$\|I(f_n) - I(f_m)\|_{L^2(dP)} = \|f_n - f_m\|_{L^2(dP \times dt)} \leq \epsilon \quad \forall m, n \geq N.$$

Es ist also auch  $I(f_n)$  eine Cauchy-Folge in  $L^2(dP)$  und es existiert der  $L^2$ -Limes  $I(f)$ .

- **Eindeutigkeit:** Sei  $f'_n \in \mathcal{H}_0^2$  eine weitere  $L^2$ -Approximationsfolge für  $f$ . Dann gilt nach der Ito-Isometrie und der Dreiecksungleichung

$$\begin{aligned} \|I(f_n) - I(f'_n)\|_{L^2(dP)} &= \|f_n - f'_n\|_{L^2(dP \times dt)} \leq \\ &\leq \|f_n - f\|_{L^2(dP \times dt)} + \|f'_n - f\|_{L^2(dP \times dt)} \rightarrow 0. \end{aligned}$$

Das Ito-Integral  $I(f)$  ist also eindeutig (als Element von  $L^2(d\mathbb{P})$ ) definiert.

Wir fassen die soeben Beschriebene Konstruktion zu folgender Definition zusammen.

**Definition.** Das **Ito-Integral**  $I$  ist auf  $\mathcal{H}^2$  als eindeutige stetige Fortsetzung des elementaren Ito-Integrals auf  $\mathcal{H}_0^2$  definiert.

Die folgenden Eigenschaften folgen nun direkt aus der Konstruktion des Ito-Integrals:

**Proposition 3.6.** Für  $f, g \in \mathcal{H}^2[0, T]$  gilt

- (a)  $I(f) \in \mathcal{F}_T$ ;
- (b)  $I(cf + g) = cI(f) + I(g)$  für alle  $c \in \mathbb{R}$ .

Auch die Ito-Isometrie lässt sich von  $\mathcal{H}_0^2$  auf  $\mathcal{H}^2$  fortsetzen

**Theorem 3.7** (Ito-Isometrie auf  $\mathcal{H}^2$ ). Das Ito-Integral  $I$  ist eine Isometrie von  $\mathcal{H}^2 \subset L^2(d\mathbb{P} \otimes dt)$  in  $L^2(dP)$ , das heisst

$$\|I(f)\|_{L^2(dP)} = \|f\|_{L^2(dP \times dt)} \quad \forall f \in \mathcal{H}^2.$$

*Beweis.* Wenig überraschend erfolgt der Beweis durch  $L^2$ -Approximation. Sei also  $f \in \mathcal{H}^2$  und  $f_n \in \mathcal{H}_0^2$  eine  $L^2$ -Approximationsfolge, sodass  $\|f - f_n\|_{L^2(dP \times dt)} \rightarrow 0$ . Mit der Dreiecksungleichung folgt dass auch  $\|f_n\|_{L^2(dP \times dt)} \rightarrow \|f\|_{L^2(dP \times dt)}$ . Genauso folgt aus  $\|I(f) - I(f_n)\|_{L^2(dP)} \rightarrow 0$  dass  $\|I(f_n)\|_{L^2(dP)} \rightarrow \|I(f)\|_{L^2(dP)}$ . Laut Lemma 3.4 gilt  $\|f_n\|_{L^2(dP \times dt)} = \|I(f_n)\|_{L^2(dP)}$  für alle  $n \in \mathbb{N}$  und die Behauptung folgt mit  $n \rightarrow \infty$ .  $\square$

### 3.3 Das Ito-Integral als stetiges Martingal

Sei  $f \in \mathcal{H}^2[0, T]$  und definiere die Abschneidefunktion

$$m_t(s) := \begin{cases} 1 & s \leq t \\ 0 & s > t. \end{cases}$$

Klarerweise ist für  $t \in [0, T]$  auch  $m_t f \in \mathcal{H}^2[0, T]$  und es macht Sinn

$$X'_t = I(m_t f), \quad t \in [0, T]$$

zu definieren. Es liegt Nahe  $X'_t$  als stochastischen Prozess zu interpretieren. Dabei tritt jedoch wieder ein grundlegendes ‘Überabzählbarkeitsproblem’ zu Tage: Für jedes  $t \in [0, T]$  ist das Ito-Integral  $I(m_t f)$  als Element von  $L^2(d\mathbb{P})$  definiert. Das heisst  $I(m_t f)$  lässt sich auf einer  $\mathbb{P}$ -Nullmenge  $A_t$  beliebig abändern ohne die Äquivalenzklasse in  $L^2(d\mathbb{P})$  zu verlassen. Insgesamt ist also  $X'$  nur auf einer Vereinigung überabzählbar vieler  $\mathbb{P}$ -Nullmengen  $(A_t)_{t \in [0, T]}$  eindeutig definiert. Die Vereinigung überabzählbar vieler  $\mathbb{P}$ -Nullmengen ist im Allgemeinen keine  $\mathbb{P}$ -Nullmenge und somit haben wir mit der Auffassung von  $X'$  als stochastischen Prozess ein Problem. Wie das nächste Resultat zeigt gibt es für dieses Problem eine überaus elegante Lösung.

**Theorem 3.8.** *Für jedes  $f \in \mathcal{H}^2[0, T]$  existiert ein stetiges Martingal  $X$ , adaptiert bezüglich der Brownschen Standardfiltration sodass  $X$  eine Modifikation von  $t \mapsto I(m_t f)$  ist, d.h.*

$$\mathbb{P}[X_t = I(m_t f)] = 1, \quad \forall t \in [0, T].$$

*Der Prozess  $X$  ist eindeutig bis auf Ununterscheidbarkeit.*

*Beweis.* Zunächst wählen wir gemäß Lemma 3.5 eine Approximationsfolge  $f_n \in \mathcal{H}_0^2$  mit  $\|f_n - f\|_{L^2(dP \times dt)} \rightarrow 0$ . Klarerweise ist für alle  $t \in [0, T], n \in \mathbb{N}$  auch  $m_t f_n \in \mathcal{H}_0^2$ . Für die elementaren Integranden  $m_t f_n$  ist das Ito-Integral  $X_t^n = I(m_t f_n)$  explizit definiert und lässt sich als

$$X_t^n = a_k(B_t - B_{t_k}) + \sum_{i=0}^{k-1} a_i(B_{t_{i+1}} - B_{t_i}), \quad t \in (t_k, t_{k+1}]$$

schreiben. Es lässt sich direkt ablesen dass jeder Prozess  $t \mapsto X_t^n$  ein stetiges Martingal ist. Für  $n \geq m$  können wir die Doobsche Ungleichung auf das stetige Submartingal  $|X^n - X^m|$  anwenden und erhalten

$$\mathbb{P} \left[ \sup_{t \in [0, T]} |X_t^n - X_t^m| > \epsilon \right] \leq \frac{1}{\epsilon^2} \mathbb{E} [(X_T^n - X_T^m)^2] \leq \frac{1}{\epsilon^2} \|f_n - f_m\|_{L^2(dP \times dt)}^2,$$

wobei im letzten Schritt die Ito-Isometrie verwendet wurde. Wir wählen nun eine steigende Teilfolge  $(n_k)_{k \in \mathbb{N}}$  sodass  $\|f_{n_{k+1}} - f_{n_k}\|_{L^2(dP \times dt)}^2 \leq 2^{-3k}$  für alle  $k \in \mathbb{N}$ . Dann folgt mit  $\epsilon = 2^{-k}$  aus obiger Abschätzung, dass

$$\mathbb{P} \left[ \sup_{t \in [0, T]} |X_t^{n_{k+1}} - X_t^{n_k}| > 2^{-k} \right] \leq 2^{-k}.$$

Diese Folge von Wahrscheinlichkeiten ist summierbar, und es folgt mit dem Lemma von Borel-Cantelli, dass mit Wahrscheinlichkeit 1 nur endlich viele der Ereignisse  $\sup_{t \in [0, T]} |X_t^{n_{k+1}} - X_t^{n_k}| > 2^{-k}$  eintreten. In anderen Worten, es existiert  $\Omega_0 \in \mathcal{F}$  mit



$\mathbb{P}[\Omega_0] = 1$  und eine Zufallsvariable  $N(\omega)$ , sodass

$$\sup_{t \in [0, T]} |X_t^{n_{k+1}}(\omega) - X_t^{n_k}(\omega)| \leq 2^{-k}, \quad \forall k \geq N(\omega), \omega \in \Omega_0.$$

Demnach bildet für jedes  $\omega \in \Omega_0$  die Folge  $(X_t^{n_k}(\omega))_{k \in \mathbb{N}}$  eine Cauchyfolge bezüglich der Supremumsnorm im Raum  $C[0, T]$  der stetigen Funktionen auf  $[0, T]$ . Wir folgern dass  $X(\omega) \in C[0, T]$  existiert, sodass  $X_t^{n_k}(\omega) \rightarrow X_t(\omega)$  gleichmäßig für  $t \in [0, T]$  und für alle  $\omega \in \Omega_0$ . Da nach Definition des Ito-Integrals  $X_t^{n_k} \rightarrow I(m_t f)$  in  $L^2$ , folgern wir aus der Eindeutigkeit des  $L^2$ -Limes dass

$$\mathbb{P}[X_t = I(m_t f)] = 1, \quad \forall t \in [0, T].$$

Aus der Martingaleigenschaft  $X_s^{n_k} = \mathbb{E}[X_t^{n_k} | \mathcal{F}_s]$  folgt mit der fast sicheren und  $L^2$ -Konvergenz von  $X_t^{n_k}$  gegen  $X_t$  auch die Martingaleigenschaft von  $X$ , sodass  $X$  tatsächlich ein stetiges Martingal ist. Die Eindeutigkeit von  $X$  ist eine Folgerung aus Lemma 1.1.  $\square$

Mit Theorem 3.8 haben wir nun die endgültige Rechtfertigung von der Notation  $I(m_t f)$  zur **Integralnotation** überzugehen. Wir schreiben also ab nun

$$\int_0^t f(\omega, s) dB_s := X_t, \quad f \in \mathcal{H}^2[0, T], t \in [0, T] \quad (3.6)$$

für das in Theorem 3.8 definierte stetige Martingal  $X$  und genauso auch

$$\int_s^t f(\omega, s) dB_s := X_t - X_s, \quad f \in \mathcal{H}^2[0, T], s, t \in [0, T].$$

Mit dieser Notation bekommt die Ito-Isometrie aus Theorem 3.7 folgende Gestalt:

$$\mathbb{E} \left[ \left( \int_0^t f(\omega, s) dB_s \right)^2 \right] = \mathbb{E} \left[ \int_0^t f(\omega, s)^2 ds \right], \quad \forall f \in \mathcal{H}^2[0, T].$$

Relativ leicht lässt sich nun auch folgende bedingte Version der Ito-Isometrie zeigen.

**Korollar 3.9.** *Für jedes  $f \in \mathcal{H}^2$  und  $0 \leq s \leq t$  gilt*

$$\mathbb{E} \left[ \left( \int_s^t f(\omega, u) dB_u \right)^2 \middle| \mathcal{F}_s \right] = \mathbb{E} \left[ \int_s^t f(\omega, u)^2 du \middle| \mathcal{F}_s \right].$$

*Beweis.* Die Aussage der Korollars ist äquivalent zu

$$\mathbb{E} \left[ \left( \mathbf{1}_{\{\omega \in A\}} \int_s^t f(\omega, u) dB_u \right)^2 \right] = \mathbb{E} \left[ \mathbf{1}_{\{\omega \in A\}} \int_s^t f(\omega, u)^2 du \right], \quad \forall A \in \mathcal{F}_s.$$

Definiere

$$f'(\omega, u) = \begin{cases} 0 & u \in [0, s] \\ \mathbf{1}_{\{\omega \in A\}} f(\omega, u) & u \in (s, t] \end{cases}.$$

Wegen  $A \in \mathcal{F}_s$  ist auch  $f' \in \mathcal{H}^2$  und die obige Gleichung ist äquivalent zur Ito-Isometrie angewandt auf  $f'$ .  $\square$

Aus der bedingten Ito-Isometrie folgt dass für jedes  $f \in \mathcal{H}^2$

$$M_t = \left( \int_0^t f(s) dB_s \right)^2 - \int_0^t f(s)^2 ds$$

ein Martingal ist. Für  $f \equiv 1$  erhalten wir das bekannte Martingal  $B_t^2 - t$  aus Theorem 2.11.

### 3.4 Stoppen unter dem Integralzeichen

In diesem Abschnitt interessieren wir uns für das Verhalten des Ito-Integrals unter Stoppen. In Analogie zur Definition würden wir erwarten, dass wir auch Stoppzeiten ‘unter das Integralzeichen ziehen’ dürfen, das heisst dass für jede Stoppzeit  $\tau$

$$\int_0^{T \wedge \tau} f(\omega, s) dB_s = \int_0^T f(\omega, s) \mathbf{1}_{\{s \leq \tau\}} dB_s \quad (3.7)$$

gilt. Auf den ersten Blick scheint diese Gleichung geradezu trivial, und man ist versucht zu argumentieren dass man ja ‘ $\omega$  festhalten könne’, und dann  $\tau(\omega)$  einfach wie eine deterministische Größe behandeln, wonach (3.7) aus der Definition des Ito-Integrals folgen würde. Diese Argumentation fußt jedoch auf einer pfadweisen Interpretation des Ito-Integrals, welche bekanntlich falsch ist. Das Ito-Integral ist eben nicht  $\omega$ -weise definiert sondern lediglich als  $L^2$ -Limes. Ein Hinweis auf die Probleme bei der obigen Argumentation ist, dass nie verwendet wird dass  $\tau$  eine Stoppzeit ist –  $\tau$  könnte demnach eine beliebige Zufallsvariable  $\tau : \omega \rightarrow \mathbb{R}_{\geq 0}$  sein. Dann wäre aber  $f(\omega, s) \mathbf{1}_{\{s \leq \tau\}}$  im allgemeinen kein adaptierter Prozess und  $\int_0^T f(\omega, s) \mathbf{1}_{\{s \leq \tau\}} dB_s$  somit gar nicht definiert. Wir müssen (3.7) also sauber beweisen. Hier soll sogar ein etwas allgemeineres Resultat gezeigt werden.

**Theorem 3.10.** Seien  $g, h \in \mathcal{H}^2[0, T]$  und  $\tau \leq T$  eine Stoppzeit, sodass  $g(\omega, t) = h(\omega, t)$  für fast alle  $\{\omega : t \leq \tau(\omega)\}$  gilt. Dann gilt auch

$$\int_0^t g(\omega, s) dB_s = \int_0^t h(\omega, s) dB_s$$

für fast alle  $\{\omega : t \leq \tau(\omega)\}$ .

Mit  $g = f$  und  $h(t) = f(t)\mathbf{1}\{t \leq \tau\}$  folgt aus dem Theorem natürlich (3.7).

*Beweis.* Wir setzen  $f = h - g$ ; es reicht offensichtlich zu zeigen, dass für alle  $t \in [0, T]$  aus  $f(\omega, t) = 0$  auf  $A_t = \{\omega : t \leq \tau(\omega)\}$  folgt dass  $\int_0^t f(\omega, s) dB_s = 0$  auf  $A_t$ . Für einfache Integranden  $f \in \mathcal{H}_0^2$  lässt sich die explizite Form des Ito-Integrals verwenden und die Behauptung durch direktes Nachrechnen zeigen. Für allgemeine  $f \in \mathcal{H}^2$  reduzieren wir die Aussage zunächst einmal auf beschränkte  $f$ : Falls nämlich die Behauptung für beschränkte Funktionen gilt, setzen wir  $f_n = f\mathbf{1}\{|f| \leq n\}$ . Es gilt  $f_n \rightarrow f$  in  $L^2(dP \times dt)$  und nach der Ito-Isometrie auch  $I(f_n) \rightarrow I(f)$  sodass sich das Resultat sofort auf allgemeine  $f \in \mathcal{H}^2$  übertragen lässt. Wir dürfen also annehmen dass  $|f|$  durch  $C > 0$  beschränkt ist und wählen wie üblich eine Approximationsfolge  $f_n \in \mathcal{H}_0^2$  mit  $\|f - f_n\|_2 \rightarrow 0$ . Jedes  $f_n$  ist von der Form

$$f_n(\omega, t) = \sum_{i=0}^{n-1} a_i^n(\omega) \mathbf{1}\{t \in (t_i^n, t_{i+1}^n]\}.$$

Leider folgt aus  $f(\omega, t) = 0$  auf  $A_t = \{\omega : t \leq \tau(\omega)\}$  nicht dass auch  $f_n(\omega, t) = 0$  auf  $A_t$  gilt (In diesem Fall wäre der Beweis hier zu Ende). Wir müssen die Approximationsfolge also noch auf geeignete Art und Weise modifizieren. Sei  $\Delta_n$  die Feinheit der zu  $f_n$  gehörigen Partition  $0 = t_0^n < t_1^n < \dots < t_N^n = T$ , das heisst  $\Delta_n := \max_i |t_{i+1}^n - t_i^n|$ . Ohne Beschränkung der Allgemeinheit dürfen wir annehmen dass  $\Delta_n \rightarrow 0$ . Definiere

$$\hat{f}_n(\omega, t) = \sum_{i=0}^{n-1} a_i^n(\omega) \mathbf{1}\{t \in (t_i^n, t_{i+1}^n]\} \mathbf{1}\{\tau \leq t_i^n\}.$$

Aus  $\tau(\omega) \geq t$  und  $t \in (t_i^n, t_{i+1}^n]$  folgt  $\tau(\omega) > t_i^n$ ; es gilt also tatsächlich dass  $\hat{f}_n(\omega, t) = 0$  auf  $A_t$ . Damit lässt sich das Resultat für einfache Integranden anwenden und  $\int_0^t \hat{f}_n(s) dB_s = 0$ . Falls wir noch zeigen können dass  $\|\hat{f}_n - f\|_{L^2(dP \times dt)} \rightarrow 0$  dann folgt die Behauptung. Die Funktionen  $\hat{f}_n$  und  $f_n \mathbf{1}\{\tau \leq t\}$  unterscheiden sich nur auf jenem Intervall  $(t_k^n, t_{k+1}^n]$  in welchem  $\tau(\omega)$  liegt. Wir können also folgendermassen abschätzen:

$$\|\hat{f}_n - f_n\|_{L^2(dP \times dt)}^2 = \int_0^T \left| \hat{f}_n(s) - f_n(s) \mathbf{1}\{\tau \leq s\} \right|^2 ds \leq B^2 \Delta_n^2 \rightarrow 0.$$

Mit der Dreiecksungleichung folgt die  $L^2$ -Konvergenz von  $\hat{f}_n$  gegen  $f$  und die Behauptung ist bewiesen.  $\square$

## 4 Lokalisierung und lokale Martingale

Das im letzten Kapitel definierte Ito-Integral bietet noch einigen Raum für Verbesserungen. Ein Schwachpunkt ist, dass auch für einige sehr ‘freundliche’ Integranden das Ito-Integral nicht unbedingt definiert ist. So würde man etwa hoffen, dass sich auch

$$\int_0^T f(B_t)dB_t, \quad \text{mit } f : \mathbb{R} \rightarrow \mathbb{R} \text{ stetig} \quad (4.1)$$

sinnvoll definieren lässt. Mit der Definition aus dem letzten Kapitel ist dies jedoch nur der Fall wenn  $f(B_t(\omega)) \in \mathcal{H}^2[0, T]$ , was z.B. für  $f(x) = \exp(x^2)$  nicht gilt. Gerade Integrale der Form (4.1) treten aber in Anwendungen besonders häufig auf und wir hätten sie gerne ohne Einschränkungen definiert. Wie sich zeigen wird können wir den Definitionsbereich des Ito-Integrals nochmals ein wenig ausdehnen, so dass auch alle Integrale der Gestalt (4.1) wohldefiniert sind. Diese Ausdehnung erfolgt durch die Methode der **Lokalisierung**, welche sich auch in anderen Zusammenhängen als nützlich erweist. So führt sie zum Begriff des **lokalen Martingals**, welcher eine wichtige Erweiterung des Martingalbegriffs darstellt.

### 4.1 Das Ito-Integral auf $\mathcal{L}_{\text{loc}}^2$

**Definition.** <sup>6</sup> Eine Folge  $(\tau_n)_{n \in \mathbb{N}}$  von Stoppzeiten heisst **lokalisierend**, wenn sie aufsteigend ist und  $\tau_n \rightarrow \infty$  fast sicher gilt.

Eine typische Anwendung von lokalisierenden Folgen ist die Approximation eines stochastischen Prozesses durch *beschränkte* stochastische Prozesse: Sei  $X$  ein stetiger stochastischer Prozess. Dann definiert

$$\tau_n := \inf \{t > 0 : |X_t| \geq n\}$$

eine lokalisierende Folge von Stoppzeiten. Die gestoppten Prozesse  $X_t^n := X_{t \wedge \tau_n}$  sind nun beschränkt, da  $|X_t^n| \leq n$ , und sie approximieren  $X$  in dem Sinne dass  $\lim_{n \rightarrow \infty} X_t^n = X_t$  fast sicher für alle  $t \geq 0$  gilt. Wir wollen die Methode der Lokalisierung nun zur Verallgemeinerung des Ito-Integrals verwenden.

**Definition.** Es sei  $\mathcal{L}_{\text{loc}}^2[0, T]$  der Vektorraum aller  $\mathbb{F}$ -adaptierten stochastischen Prozesse  $f$ , welche die Integrierbarkeitsbedingung

$$\mathbb{P} \left[ \int_0^T f(\omega, t)^2 dt < \infty \right] = 1 \quad (4.2)$$

---

<sup>6</sup>Je nach Anwendungsgebiet werden oft auch noch weitere Eigenschaften von der Folge  $(\tau_n)_{n \in \mathbb{N}}$  verlangt; die hier verwendete Definition ist sozusagen die ‘Minimaldefinition’.

erfüllen.

Offensichtlich gilt  $\mathcal{H}^2[0, T] \subset \mathcal{L}_{\text{loc}}^2[0, T]$  und  $\mathcal{L}_{\text{loc}}^2[0, T]$  erweitert den Raum der zulässigen Integranden für das Ito-Integral. Wegen der pfadweisen Stetigkeit der Brownschen Bewegung gilt ausserdem, dass  $\int_0^T f(B_t(\omega))^2 dt < \infty$  fast sicher für stetiges  $f$ . Daher sind wie gewünscht alle Integranden der Form  $f(B_t(\omega))$  in  $\mathcal{L}_{\text{loc}}^2[0, T]$ . Andererseits ist  $\mathcal{L}_{\text{loc}}^2[0, T]$  kein Unterraum von  $L^2(d\mathbb{P} \times dt)$  mehr und die Methoden der  $L^2$ -Approximation aus dem letzten Kapitel lassen sich nicht mehr unmittelbar anwenden.

Wir stellen nun den Zusammenhang zu lokalisierenden Folgen her.

**Lemma 4.1.** *Sei  $f \in \mathcal{L}_{\text{loc}}^2[0, T]$ . Dann ist*

$$\tau_n := \inf \left\{ t \in [0, T] : \int_0^t f(\omega, s)^2 ds \geq n \right\}$$

eine lokalisierende Folge von Stoppzeiten und es gilt

$$f(\omega, t) \mathbf{1}_{\{t \leq \tau_n\}} \in \mathcal{H}^2[0, T], \quad \forall n \in \mathbb{N}.$$

Das Resultat folgt direkt aus der Definition der Stoppzeiten  $\tau_n$ ; dabei ist zu beachten dass wir  $\inf \emptyset = \infty$  setzen.

Der Weg zur **Definition des Ito-Integrals auf  $\mathcal{L}_{\text{loc}}^2[0, T]$**  ist nun klar:

- Für  $f \in \mathcal{L}_{\text{loc}}^2[0, T]$  wähle eine lokalisierende Folge  $\tau_n$  wie in Lemma 4.1. Es gilt  $f(\omega, t) \mathbf{1}_{\{t \leq \tau_n\}} \in \mathcal{H}^2[0, T]$  und

$$X_t^n := \int_0^t f(\omega, s) \mathbf{1}_{\{s \leq \tau_n\}} dB_s \tag{4.3}$$

ist wohldefiniert für alle  $t \in [0, T]$ .

- Für  $t \in [0, T]$  definiere das Ito-Integral von  $f$  als

$$\int_0^t f(\omega, s) dB_s := \lim_{n \rightarrow \infty} X_t^n \tag{4.4}$$

wobei die Konvergenz fast sicher gilt.

Zu zeigen ist bei dieser Definition lediglich dass der Limes (4.4) tatsächlich existiert.

**Lemma 4.2.** *Sei  $f \in \mathcal{L}_{\text{loc}}^2[0, T]$  und  $X^n$  wie in (4.3) definiert. Dann existiert ein stetiger adaptierter stochastischer Prozess  $X$ , sodass*

$$\mathbb{P} \left[ X_t = \lim_{n \rightarrow \infty} X_t^n \quad \forall t \in [0, T] \right] = 1.$$

*Beweis.* Wir setzen  $N(\omega) = \inf \{n \in \mathbb{N} : \tau_n = \infty\}$ . Wegen

$$\mathbb{P}[N(\omega) < \infty] = \mathbb{P}\left[\int_0^T f(\omega, t)^2 dt < \infty\right] = 1$$

ist  $N$  fast sicher endlich. Wir finden also eine Menge  $A \in \mathcal{F}$  mit  $\mathbb{P}[A] = 1$  sodass für  $\omega \in A$  alle Pfade  $t \mapsto X_t^n(\omega)$  stetig sind und  $N(\omega) < \infty$ . Für  $\omega \in A$  definieren wir

$$X_t(\omega) := X_t^{N(\omega)}(\omega), \quad t \in [0, T].$$

Es ist klar dass  $X$  ein stetiger stochastischer Prozess ist. Außerdem gilt für alle  $m \geq N(\omega)$ , dass  $\tau_m(\omega) = \infty$  und daher  $X_t^m(\omega) = X_t(\omega)$ . Es folgt dass

$$X_t(\omega) = X_t^{N(\omega)}(\omega) = \lim_{m \rightarrow \infty} X_t^m(\omega), \quad \forall \omega \in A. \quad \square$$

Wir sehen also dass Adaptiertheit, Stetigkeit und Linearität des Ito-Integrals bei der Erweiterung auf  $\mathcal{L}_{\text{loc}}^2[0, T]$  erhalten bleiben. Zudem lässt sich zeigen dass auch Theorem 3.10 zum Stoppen unter dem Integralzeichen weiterhin gilt. Die Frage ob auch die Martingaleigenschaft erhalten bleibt wird uns in Abschnitt 4.3 beschäftigen.

## 4.2 Riemann-Darstellung und Gaussche Prozesse

Mit der Erweiterung des Ito-Integrals auf  $\mathcal{L}_{\text{loc}}^2[0, T]$  ist nun wie gewünscht jedes Integral der Form  $\int_0^T f(B_t)dB_t$  für stetiges  $f$  definiert. Wir zeigen dass sich Integrale dieser Form auch als Limes von Riemann-Summen interpretieren lassen, allerdings als Limes *in Wahrscheinlichkeit*.

**Theorem 4.3.** *Sei  $f : \mathbb{R} \rightarrow \mathbb{R}$  stetig und sei  $(\mathcal{P}^n)_{n \in \mathbb{N}}$  eine Folge von Partitionen von  $[0, T]$  mit  $|\mathcal{P}^n| \rightarrow 0$ . Dann gilt*

$$\int_0^T f(B_t)dB_t = \lim_{n \rightarrow \infty} \sum_{t_i \in \mathcal{P}^n} f(B_{t_i})(B_{t_{i+1}} - B_{t_i}) \quad (4.5)$$

*mit Konvergenz in Wahrscheinlichkeit.*

*Beweis.* Wir definieren eine lokalisierende Folge von Stoppzeiten durch

$$\tau_m := \inf \{t \in [0, T] : |B_t| \geq m\}.$$

Wir setzen  $f_m(x) := f(x)\mathbf{1}\{|x| \leq m\}$ . Da  $f$  stetig ist, ist  $f_m$  beschränkt und es gilt

$f_m(B_t(\omega)) \in \mathcal{H}^2[0, T]$ . Es lässt sich zeigen (siehe [Steele, 2003, S. 100] für Details) dass

$$\phi_n^m(\omega, t) = \sum_{i=0}^n f_m(B_{t_i}) \mathbf{1}_{\{t_i < t \leq t_{i+1}\}}$$

eine  $L_2$ -Approximationsfolge von einfachen Integranden für  $f_m$  ist, d.h.

$$\lim_{n \rightarrow \infty} \|\phi_n^m - f_m\|_{L_2(dP \times dt)} = 0, \quad \forall m \in \mathbb{N}.$$

Mit der Ito-Isometrie folgt daraus

$$\lim_{n \rightarrow \infty} \|I(\phi_n^m) - I(f_m)\|_{L_2(dP)} = 0, \quad \forall m \in \mathbb{N}. \quad (4.6)$$

Andererseits gilt

$$I(\phi_n^m) = \sum_{i=0}^n f_m(B_{t_i})(B_{t_{i+1}} - B_{t_i})$$

und wir erhalten die Riemann-Darstellung (4.5) für  $f_m$  und mit Konvergenz in  $L_2(d\mathbb{P})$  statt in Wahrscheinlichkeit.

Als nächstes stellen wir fest, dass  $f(B_t(\omega)) = f_m(B_t(\omega))$  auf der Menge  $\{\omega : t \leq \tau_m(\omega)\}$  gilt. Mit dem Satz zum Stoppen unter dem Integralzeichen folgt dass auch

$$\int_0^t f(B_s(\omega)) dB_s = \int_0^t f_m(B_s(\omega)) dB_s, \quad \text{auf } \{\omega : t \leq \tau_m\}.$$

Damit nun (4.5) gilt müssen die Wahrscheinlichkeiten der Ereignisse

$$A_n(\epsilon) = \left\{ \omega : \left| \sum_{i=0}^n f(B_{t_i})(B_{t_{i+1}} - B_{t_i}) - \int_0^T f(B_s) dB_s \right| \geq \epsilon \right\}$$

für jedes  $\epsilon > 0$  mit  $n \rightarrow \infty$  gegen Null gehen. Wir schreiben  $A_n^m(\epsilon)$  für den selben Ausdruck mit  $f_m$  statt  $f$ . Folgende Abschätzung folgt aus der Gleichheit  $f(B_t(\omega)) = f_m(B_t(\omega))$  auf  $\{\omega : t \leq \tau_m(\omega)\}$ :

$$\mathbb{P}[A_n(\epsilon)] \leq \mathbb{P}[T > \tau_m] + \mathbb{P}[A_n^m(\epsilon)].$$

Da  $\tau_m \rightarrow \infty$  fast sicher, können wir den ersten Term beliebig klein machen und es genügt den zweiten Term abzuschätzen. Mit der Ungleichung von Tschebyscheff gilt

$$\mathbb{P}[A_n^m(\epsilon)] \leq \frac{1}{\epsilon^2} \left\| \sum_{i=0}^n f_m(B_{t_i})(B_{t_{i+1}} - B_{t_i}) - \int_0^T f_m(B_s) dB_s \right\|_{L_2(dP)}^2$$



und die Aussage folgt mit  $n \rightarrow \infty$  aus der gezeigten  $L_2$ -Konvergenz in (4.6).  $\square$

Ein weiterer wichtiger Spezialfall des Ito-Integrals ergibt sich wenn wir *deterministische* Integranden betrachten. Das Ito-Integral ist in diesem Fall ein **Gausscher Prozess**, d.h. ein Prozess dessen endlichdimensionale Randverteilungen (multivariate) Normalverteilungen sind. Ein solcher Prozess ist (in Verteilung) durch seinen Mittelwert und seine Autokovarianz bereits eindeutig charakterisiert.

**Theorem 4.4.** *Sei  $f \in C[0, T]$ . Dann ist*

$$X_t = \int_0^t f(s)dB_s, \quad t \in [0, T]$$

ein Gausscher Prozess mit unabhängigen Inkrementen, Mittelwert Null und Autokovarianz

$$\text{Cov}(X_s, X_t) = \int_0^{s \wedge t} f(u)^2 du.$$

Der Beweis bleibt dem Leser überlassen und ist leicht zu führen wenn die gewünschten Eigenschaften zuerst für die Approximationsfolge  $f_n(\omega, t) = \sum_{i=0}^{2^n-1} f(t_i) \mathbf{1}_{\{t_i < t \leq t_{i+1}\}}$  gezeigt werden.

### 4.3 Lokale Martingale

Im letzten Kapitel haben wir gezeigt, dass das Ito-Integral für Integranden in  $\mathcal{H}^2[0, T]$  immer ein stetiges Martingal ist. Gilt diese Eigenschaft auch noch für Integranden in  $\mathcal{L}_{\text{loc}}^2[0, T]$ ? Die Antwort ist nein, jedoch ist das Ito-Integral auch auf  $\mathcal{L}_{\text{loc}}^2[0, T]$  ‘nahe daran’ ein Martingal zu sein. Wir machen diese Aussage durch folgende Definition präzise.

**Definition.** Ein adaptierter stochastischer Prozess  $X$  heisst **lokales Martingal**, wenn eine lokalisierende Folge von Stoppzeiten  $(\tau_n)_{n \in \mathbb{N}}$  existiert, sodass jeder der gestoppten Prozesse

$$X_t^n := X_{t \wedge \tau_n}, \quad n \in \mathbb{N}$$

ein Martingal ist.

Analog lassen sich lokale Sub- und Supermartingale definieren. Es gilt nun folgendes

**Theorem 4.5.** *Sei  $f \in \mathcal{L}_{\text{loc}}^2[0, T]$ . Dann ist das Ito-Integral  $\int_0^t f(\omega, s)dB_s$  ein stetiges lokales Martingal.*

*Beweis.* Nach Definition des Ito-Integrals existiert eine lokalisierende Folge  $\tau_n$  sodass  $f(\omega, s)\mathbf{1}\{s \leq \tau_n\} \in \mathcal{H}^2[0, T]$ . Folglich sind die Prozesse

$$X_{t \wedge \tau_n} = \int_0^{\tau_n \wedge t} f(\omega, s) dB_s = \int_0^t f(\omega, s)\mathbf{1}\{s \leq \tau_n\} dB_s$$

Martingale und die Behauptung folgt.  $\square$

Viele Resultate zu lokalen Martingalen beruhen darauf die lokalisierende Folge  $\tau_n$  zum richtigen Zeitpunkt durch Limesbildung verschwinden zu lassen. Als Beispiel zeigen wir folgendes Resultat, welches das Beispiel zur Brownschen Bewegung aus Abschnitt 2.4 verallgemeinert.

**Proposition 4.6.** *Sei  $X$  ein stetiges lokales Martingal mit  $X_0 = 0$ . Des weiteren sei  $a, b > 0$  und*

$$\tau = \inf \{t > 0 : X_t = a, \text{ oder } X_t = -b\}.$$

*Wenn  $\tau$  fast sicher endlich ist, dann gilt*

$$\mathbb{P}[X_\tau = a] = \frac{b}{a + b}.$$

*Beweis.* Sei  $\tau_n$  eine lokalisierende Folge, sodass die gestoppten Prozesse  $X_t^n := X_{t \wedge \tau_n}$  Martingale sind. Nach dem Satz vom optionalen Stoppen ist auch jedes  $X_{t \wedge \tau}^n$  ein Martingal. Da  $\tau$  fast sicher endlich ist und  $X^n$  durch  $\max(a, b)$  beschränkt ist, gilt mit dominanter Konvergenz

$$\mathbb{E}[X_\tau^n] = \lim_{t \rightarrow \infty} \mathbb{E}[X_{t \wedge \tau}^n] = \mathbb{E}[X_0] = 0.$$

Außerdem gilt wegen  $\tau_n \uparrow \infty$ , dass  $\lim_{n \rightarrow \infty} X_\tau^n = X_\tau$  und wir erhalten nach erneuter Anwendung der dominierten Konvergenz  $\mathbb{E}[X_\tau] = 0$ . Die Behauptung folgt nun aus

$$0 = \mathbb{E}[X_\tau] = a\mathbb{P}[X_\tau = a] - b(1 - \mathbb{P}[X_\tau = a]). \quad \square$$

Oft ist es von Interesse zu wissen wann ein lokales Martingal ein ‘echtes’ Martingal ist. Wir beginnen mit zwei einfachen hinreichenden Bedingungen

**Proposition 4.7.** *Sei  $M$  ein lokales Martingal.*

- (a) *Wenn  $M$  beschränkt ist, d.h.  $|M_t| \leq B < \infty$  für alle  $t \in \mathbb{R}_{\geq 0}$ , dann ist  $M$  ein Martingal.*
- (b) *Wenn  $M$  von unten beschränkt ist, d.h.  $M_t^- \leq B < \infty$  für alle  $t \in \mathbb{R}_{\geq 0}$ , dann ist  $M$  ein Supermartingal. Wenn zusätzlich  $t \mapsto \mathbb{E}[M_t]$  konstant ist, so ist  $M$  ein Martingal.*

*Beweis.* Der Beweis von (a) sei dem Leser überlassen. Für (b) sei  $\tau_n$  eine lokalisierende Folge von Stoppzeiten, sodass jedes  $t \mapsto M_{t \wedge \tau_n}$  ein Martingal ist. Aus  $M_{s \wedge \tau_n} = \mathbb{E}[M_{t \wedge \tau_n} | \mathcal{F}_s]$  folgt mit dem Lemma von Fatou die Supermartingaleigenschaft

$$M_s \geq \mathbb{E}[M_t | \mathcal{F}_s], \quad s \leq t. \quad (4.7)$$

Insbesondere gilt  $\mathbb{E}[M_0] \geq \mathbb{E}[M_s] \geq \mathbb{E}[M_t]$  und aus  $\mathbb{E}[M_t^+] \leq \mathbb{E}[M_0] + \mathbb{E}[M_t^-]$  folgt die Integrierbarkeit von  $|M_t|$ . Nehme nun an dass  $t \mapsto \mathbb{E}[M_t]$  konstant ist. Falls in (4.7) mit strikt positiver Wahrscheinlichkeit ' $>$ ' gilt, dann folgt  $\mathbb{E}[M_s] > \mathbb{E}[M_t]$  im Widerspruch zur Annahme. Also gilt Gleichheit fast sicher und  $M$  ist ein Martingal.  $\square$

Eine wichtige Folgerung aus dieser Proposition ist, dass jedes positive lokale Martingal ein Supermartingal ist.

Die folgende Bedingung für die Martingaleigenschaft ist notwendig und hinreichend aber in dem meisten Fällen nur schwer zu überprüfen.

**Proposition 4.8.** *Sei  $M$  ein lokales Martingal bezüglich einer Filtration  $\mathbb{F}$  und sei  $\Pi_a$  die Familie aller durch  $a$  beschränkten  $\mathbb{F}$ -Stoppzeiten. Falls jede der Familien*

$$\{M_\tau : \tau \in \Pi_a\}, \quad a \geq 0$$

*gleichgradig integrierbar ist<sup>7</sup>, so ist  $M$  ein Martingal. Sind  $M$  und die Filtration  $\mathbb{F}$  rechtsstetig, so gilt auch die Umkehrung.*

*Beweis.* Sei  $\tau_n$  eine lokalisierende Folge von Stoppzeiten. Für  $s \leq t$  gilt die Martingaleigenschaft

$$M_{s \wedge \tau_n} = \mathbb{E}[M_{t \wedge \tau_n} | \mathcal{F}_s]. \quad (4.8)$$

Des weiteren ist  $\lim_{n \rightarrow \infty} M_{t \wedge \tau_n} = M_t$  fast sicher, und genauso für  $s$ . Die Stoppzeiten  $(t \wedge \tau_n)$  sind in  $\Pi_t$  und nach Voraussetzung ist also  $(M_{t \wedge \tau_n})_{n \in \mathbb{N}}$  eine gleichgradig integrierbare Folge von Zufallsvariablen. Mit Korollar 1.9 folgt somit aus (4.8) die Martingaleigenschaft für  $M$ . Für die Umkehrung wenden wir den Satz vom Optionalen Sampling an (hier wird die Rechtsstetigkeit benötigt) und erhalten

$$M_\tau = \mathbb{E}[M_a | \mathcal{F}_\tau], \quad \forall \tau \in \Pi_a.$$

Familien von bedingten Erwartungen sind nach Lemma 1.12 gleichgradig integrierbar und die Behauptung folgt.  $\square$

---

<sup>7</sup>Diese Bedingung tritt auch in einigen anderen Resultaten der stochastischen Analysis auf. Die Klasse der Prozesse die diese Bedingung erfüllt nennt man auch 'Klasse (DL)'.

## 5 Die Ito-Formel

Bei der Anwendung der gewöhnlichen Integralrechnung müssen wir nur selten auf die explizite Definition des Integrals als Limes von Riemann-Summen zurückgreifen. Rechenregeln wie der Hauptsatz der Integralrechnung oder die Substitutionsregel vereinfachen den Umgang mit gewöhnlichen Integralen entscheidend. Auch das Ito-Integral wäre für Anwendungen nur von beschränktem Interesse wenn nicht ein ähnliches Kalkül zur Verfügung stehen würde. Herzstück dieses Kalküls ist die **Ito-Formel** – eigentlich eine Reihe von verwandten Formeln – die als Entsprechung des Hauptsatzes der Integralrechnung angesehen werden kann. In ihrer einfachsten Form besagt sie folgendes:

**Theorem 5.1** (Ito-Formel – Einfachste Form). *Sei  $f : \mathbb{R} \rightarrow \mathbb{R}$  eine zweifach stetig differenzierbare Funktion. Dann gilt*

$$f(B_t) = f(0) + \int_0^t f'(B_s)dB_s + \frac{1}{2} \int_0^t f''(B_s)ds, \quad t \in \mathbb{R}_{\geq 0}, \quad (5.1)$$

*fast sicher.*

Bei dieser Formel ist vor allem der zweite Integralterm bemerkenswert, durch den sich die Ito-Formel in ihrer Form wesentlich vom Hauptsatz der gewöhnlichen Integralrechnung unterscheidet.

Die Ito-Formel hat eine interessante Interpretation als Zerlegung des Prozesses  $f(B_t)$  in eine Trendkomponente und in ‘richtungsloses’ Rauschen. Die erste Integralkomponente ist – als Ito-Integral eines  $\mathcal{L}_{\text{loc}}^2$ -Integranden – ein lokales Martingal, und besitzt daher (bis auf Lokalisierung) Erwartungswert Null. Die zweite Integralkomponente muss also jeglichen Trend (d.h Ansteigen oder Absteigen der Erwartung) von  $f(B_t)$  abbilden. Bei den lokalen Fluktuationen von  $f(B_t)$  verhält es sich genau umgekehrt: Der zweite Integralterm hat beschränkte Totalvariation, sodass lokale Variationen höherer Ordnung – wie sie ja bei Funktionalen der Brownschen Bewegung zu erwarten sind – gänzlich vom ersten Integralterm abgebildet werden müssen.

Eine weitere wichtige Beobachtung ist, dass die Integrale auf der rechten Seite der Formel alleine aufgrund der Stetigkeit von  $f'$  und  $f''$  wohldefiniert sind; die Erweiterung des Ito-Integrals von  $\mathcal{H}^2$  auf  $\mathcal{L}_{\text{loc}}^2$  hat also Früchte getragen.

### 5.1 Herleitung der Ito-Formel

Wir beweisen die Ito-Formel in mehreren Schritten.

*Beweis. Schritt 1:* Wir nehmen zunächst an, dass die Funktion  $f$  kompakten Träger

hat. Wir setzen  $t_i = it/n$  und stellen  $f(B_t) - f(0)$  als Teleskopsumme dar

$$f(B_t) - f(0) = \sum_{i=0}^{n-1} (f(B_{t_{i+1}}) - f(B_{t_i})). \quad (5.2)$$

Nachdem  $f$  zweifach stetig differenzierbar ist, gilt die Taylorentwicklung:

$$f(y) - f(x) = f'(x)(y - x) + \frac{1}{2}f''(x)(y - x)^2 + h(x, y)(y - x)^2, \quad x, y \in \mathbb{R}, \quad (5.3)$$

mit  $h$  gleichmässig stetig und  $h(x, x) = 0$  für alle  $x \in \mathbb{R}$ . Wir setzen (5.3) in (5.2) ein, und erhalten eine Darstellung von  $f(B_t) - f(0)$  als Summe dreier Terme  $I_n$ ,  $J_n$  und  $K_n$ , gegeben durch

$$I_n = \sum_{i=0}^{n-1} f'(B_{t_i})(B_{t_{i+1}} - B_{t_i}), \quad J_n = \frac{1}{2} \sum_{i=0}^{n-1} f''(B_{t_i})(B_{t_{i+1}} - B_{t_i})^2$$

und

$$K_n = \sum_{i=0}^{n-1} h(B_{t_i}, B_{t_{i+1}})(B_{t_{i+1}} - B_{t_i})^2.$$

Für den ersten Term folgt aus Theorem 4.3 zur Riemann-Darstellung von stochastischen Integralen wegen der Stetigkeit von  $f'$ , dass

$$\lim_{n \rightarrow \infty} I_n = \int_0^t f'(B_s) dB_s, \quad \text{in Wahrscheinlichkeit.} \quad (5.4)$$

Wir schreiben den zweiten Term  $J_n$  um, indem wir die Zufallsvariablen  $(B_{t_{i+1}} - B_{t_i})^2$  zentrieren:

$$J_n = \frac{1}{2} \sum_{i=0}^{n-1} f''(B_{t_i})(t_{i+1} - t_i) + \frac{1}{2} \sum_{i=0}^{n-1} f''(B_{t_i}) \{ (B_{t_{i+1}} - B_{t_i})^2 - (t_{i+1} - t_i) \}. \quad (5.5)$$

Der erste Summand konvergiert mit Wahrscheinlichkeit 1 gegen ein gewöhnliches Integral; es gilt

$$\lim_{n \rightarrow \infty} \frac{1}{2} \sum_{i=0}^{n-1} f''(B_{t_i})(t_{i+1} - t_i) = \frac{1}{2} \int_0^t f''(B_s) ds.$$

Wir schreiben  $\widehat{J}_n$  für die zweite Summe in (5.5). Mit einem nun schon aus mehreren vorangehenden Beweisen bekannten Argument, beruhend auf der Orthogonalität der

einzelnen Summanden, erhalten wir

$$\mathbb{E} \left[ \widehat{J}_n^2 \right] = \frac{1}{4} \sum_{i=0}^{n-1} \mathbb{E} \left[ f''(B_{t_i})^2 \{ (B_{t_{i+1}} - B_{t_i})^2 - (t_{i+1} - t_i) \}^2 \right] \leq \|f''\|_\infty^2 \frac{t^2}{2n},$$

wobei in der letzten Abschätzung  $\text{Var}((B_{t_{i+1}} - B_{t_i})^2) = \frac{t^2}{n^2} \text{Var}(B_1^2) = 2 \frac{t^2}{n^2}$  verwendet wurde. Mit der Markovschen Ungleichung folgt aus dieser Abschätzung  $\widehat{J}_n \rightarrow 0$  in Wahrscheinlichkeit. Insgesamt gilt also

$$\lim_{n \rightarrow \infty} J_n = \frac{1}{2} \int_0^t f''(B_s) ds, \quad \text{in Wahrscheinlichkeit.} \quad (5.6)$$

Schritt 2: Im nächsten Schritt schätzen wir das Restglied  $K_n$  ab. Mit der Cauchy-schen Ungleichung gilt

$$\mathbb{E} [|K_n|] \leq \sum_{i=0}^{n-1} \mathbb{E} [(B_{t_{i+1}} - B_{t_i})^4]^{1/2} \mathbb{E} [h^2(B_{t_i}, B_{t_{i+1}})]^{1/2}.$$

Für den ersten Faktor erhalten wir

$$\mathbb{E} [(B_{t_{i+1}} - B_{t_i})^4] = \frac{t^2}{n^2} \mathbb{E} [B_1^4] = \frac{3t^2}{n^2}.$$

Wegen der gleichmäßigen Stetigkeit von  $h$  existiert für jedes  $\epsilon > 0$  ein  $\delta > 0$ , sodass für alle  $x, y \in \mathbb{R}$  mit  $|x - y| < \delta$  auch  $h(x, y) < \epsilon$  gilt. Wir können also unter Verwendung der Markovschen Ungleichung folgendermaßen abschätzen:

$$\begin{aligned} \mathbb{E} [h^2(B_{t_i}, B_{t_{i+1}})] &\leq \epsilon^2 + \|h\|_\infty^2 \mathbb{P} [|B_{t_{i+1}} - B_{t_i}| \geq \delta] \\ &\leq \epsilon^2 + \|h\|_\infty^2 \delta^{-2} \mathbb{E} [(B_{t_{i+1}} - B_{t_i})^2] \\ &\leq \epsilon^2 + \|h\|_\infty^2 \frac{t}{n\delta^2}, \end{aligned}$$

und insgesamt ergibt sich

$$\mathbb{E} [|K_n|] \leq n \sqrt{3t^2/n^2} \sqrt{\epsilon^2 + \|h\|_\infty^2 t/(n\delta^2)}.$$

Es folgt  $\limsup_{n \rightarrow \infty} \mathbb{E} [|K_n|] \leq \sqrt{3}t\epsilon$  und wegen der beliebigen Wahl von  $\epsilon > 0$  gilt daher  $\lim_{n \rightarrow \infty} \mathbb{E} [|K_n|] \rightarrow 0$ . Mit der Markovschen Ungleichung folgt schließlich

$$\lim_{n \rightarrow \infty} K_n = 0, \quad \text{in Wahrscheinlichkeit.} \quad (5.7)$$

Schritt 3: Jede in Wahrscheinlichkeit konvergierende Folge von Zufallsvariablen enthält

eine fast sicher konvergierende Teilfolge. Nach der Wahl der passenden Teilfolge lassen sich die oben gezeigten Limiten (5.4), (5.6) und (5.7) als fast sichere Grenzwerte auffassen. Wir haben somit gezeigt, dass für Funktionen  $f$  mit kompaktem Träger die Ito-Formel für jedes  $t$  mit Wahrscheinlichkeit 1 gilt. Anders gesagt haben wir gezeigt, dass die linke Seite der Ito-Formel (5.1) eine Modifikation der rechten ist. Da aber sowohl auf der linken Seite als auch auf der rechten Seite ein stetiger stochastischer Prozess aufscheint, gilt mit Lemma 1.1 eine stärkere Aussage, nämlich Gleichheit bis auf Ununterscheidbarkeit. Es gilt nun nur noch, die Annahme des kompakten Trägers von  $f$  loszuwerden. Dazu verwenden wir Lokalisierung: Für jedes  $f \in C^2(\mathbb{R})$  können wir eine Funktion  $f_m \in C^2(\mathbb{R})$  mit kompaktem Träger finden, welche für  $|x| \leq m$  mit  $f$  übereinstimmt. Nach dem bisher Gezeigten gilt also

$$f_m(B_t) = f_m(0) + \int_0^t f'_m(B_s)dB_s + \frac{1}{2} \int_0^t f''_m(B_s)ds, \quad t \in \mathbb{R}_{\geq 0}.$$

Wir definieren eine lokalisierende Folge von Stoppzeiten  $\tau_m = \inf \{t > 0 : |B_t| \geq m\}$ . Auf der Menge  $\{\omega : t \leq \tau_m(\omega)\}$  gilt  $f'(B_s) = f'_m(B_s)$  und nach dem Theorem zum Stoppen unter dem Integralzeichen auch

$$\int_0^t f'(B_s)dB_s = \int_0^t f'_m(B_s)dB_s.$$

Genauso gilt auf  $\{\omega : t \leq \tau_m(\omega)\}$

$$f(B_t) = f_m(B_t) \quad \text{und} \quad \int_0^t f''(B_s)ds = \int_0^t f''_m(B_s)ds,$$

und daher schließlich auch die Ito-Formel (5.1) für  $f$ . Wir bilden den Grenzwert  $m \rightarrow \infty$  und erhalten wegen  $\tau_m \rightarrow \infty$  fast sicher die Gültigkeit der Ito-Formel mit Wahrscheinlichkeit 1.  $\square$

## 5.2 Erste Folgerungen und Anwendungen

Eine erste Verallgemeinerung der Ito-Formel erhalten wir wenn wir Funktionen  $f : \mathbb{R}_{\geq 0} \times \mathbb{R} \rightarrow \mathbb{R}$  betrachten welche von einer Zeit- und einer Ortskomponente abhängen.

**Theorem 5.2** (Ito-Formel mit Zeitabhängigkeit). *Sei  $f \in C^{1,2}(\mathbb{R}_{\geq 0} \times \mathbb{R})$ . Dann gilt*

$$f(t, B_t) = f(0, 0) + \int_0^t \frac{\partial f}{\partial t}(s, B_s)ds + \int_0^t \frac{\partial f}{\partial x}(s, B_s)dB_s + \frac{1}{2} \int_0^t \frac{\partial^2 f}{\partial x^2}(s, B_s)ds \quad (5.8)$$

*fast sicher.*

Die Notation  $C^{1,2}(\mathbb{R}_{\geq 0} \times \mathbb{R})$  bezeichnet dabei Funktionen die in der ersten Variable einmal stetig differenzierbar, in der zweiten Variable zweimal stetig differenzierbar sind. Der Beweis wird wie für Theorem 5.1 geführt, nur wird die gemischte Taylorentwicklung

$$f(s, y) - f(t, x) = (s - t) \frac{\partial f}{\partial t}(t, x) + (y - x) \frac{\partial f}{\partial x}(t, x) + \frac{(y - x)^2}{2} \frac{\partial^2 f}{\partial x^2}(t, x) + (s - t)(y - x)^2 h(s, t, x, y)$$

verwendet.

Aus der Ito-Formel mit Zeitabhängigkeit ergibt sich eine interessante Folgerung.

**Korollar 5.3.** Sei  $f \in C^{1,2}(\mathbb{R}_{\geq 0} \times \mathbb{R})$  eine Funktion die

$$\frac{\partial f}{\partial t} = -\frac{1}{2} \frac{\partial^2 f}{\partial x^2} \quad (5.9)$$

erfüllt. Dann ist  $X_t = f(t, B_t)$  ein lokales Martingal. Gilt außerdem

$$\mathbb{E} \left[ \int_0^T \left( \frac{\partial f}{\partial x}(t, B_t) \right)^2 dt \right] < \infty, \quad (5.10)$$

so ist  $X$  ein Martingal auf  $[0, T]$ .

Dieses Resultat stellt einen Zusammenhang zwischen der partiellen Differentialgleichung (5.9) und Eigenschaften des stochastischen Prozesses  $X$  her. Diese Verbindung zwischen partiellen Differentialgleichungen (PDEs) und stochastischen Prozessen erweist sich als äußerst fruchtbar und wir werden sie in den Kapiteln 9 und 10 weiter verfolgen.

*Beweis.* Unter Voraussetzung (5.9) folgt aus der Ito-Formel mit Zeitabhängigkeit

$$X_t = f(t, B_t) = f(0, 0) + \int_0^t \frac{\partial f}{\partial x}(s, B_s) dB_s.$$

Da Ito-Integrale lokale Martingale sind, ist auch  $X$  eines. Unter Voraussetzung (5.10) ist der Integrand sogar in  $\mathcal{H}^2[0, T]$  und es folgt die Martingaleigenschaft von  $X$  auf  $[0, T]$ .  $\square$

Als erste Anwendung von Korollar 5.3 betrachten wir den aus Abschnitt 2.4 wohl-bekanntesten Prozess

$$M_t = \exp \left( \vartheta B_t - \frac{\vartheta^2}{2} t \right), \quad \vartheta \in \mathbb{R}.$$



Er lässt sich mit  $f(t, x) = \exp\left(\vartheta x - \frac{\vartheta^2}{2}t\right)$  als  $M_t = f(t, B_t)$  schreiben. Es gilt

$$\frac{\partial f}{\partial t} = -\frac{1}{2}\vartheta^2 f \quad \text{und} \quad \frac{\partial^2 f}{\partial x^2} = \vartheta^2 f,$$

die Funktion  $f$  erfüllt also die PDE (5.9) und  $M$  ist ein lokales Martingal. Da auch Bedingung (5.10) erfüllt ist, haben wir erneut nachgewiesen, dass  $M$  ein Martingal ist.

Als nächstes kehren wir zu dem ebenfalls in Abschnitt 2.4 behandelten Problem der Austrittszeit der Brownschen Bewegung aus einem Intervall  $(-b, a)$  zurück. Diesmal jedoch untersuchen wir die (skalierte) Brownsche Bewegung mit Drift, das heisst den Prozess

$$X_t = \sigma B_t + \mu t, \quad \sigma > 0, \mu \in \mathbb{R}.$$

Wie zuvor definieren wir die Stoppzeit

$$\tau := \inf \{t > 0 : X_t = a \quad \text{oder} \quad X_t = -b\}, \quad a, b > 0$$

und sind an der Wahrscheinlichkeit interessiert dass  $X_\tau = a$  gilt. Die Lösung aus Abschnitt 2.4 legt es nahe ein Martingal  $M$  der Form  $M_t = h(X_t)$  zu suchen. Wenn  $M$  zudem beschränkt ist, und  $h$  die Randbedingungen  $h(-b) = 0$  und  $h(a) = 1$  erfüllt, so erhalten wir mit optionalem Stoppen

$$\mathbb{E}[M_\tau] = \mathbb{E}[M_0] \quad \text{und somit} \quad \mathbb{P}[X_\tau = a] = h(0).$$

Wir müssen also überprüfen, wann  $M_t = h(X_t) = h(\sigma B_t + \mu t)$  ein Martingal ist.  $M$  ist von der Form  $f(t, B_t)$  mit  $f(t, x) = h(\sigma x + \mu t)$ . Wir wenden Korollar 5.3 an, und erhalten aus der PDE-Bedingung

$$\mu h'(\sigma x + \mu t) = -\frac{\sigma^2}{2} h''(\sigma x + \mu t).$$

Die partielle Differentialgleichung reduziert sich also auf eine gewöhnliche Differentialgleichung für  $h$ , für die wir auch noch die Randbedingungen  $h(-b) = 0$  und  $h(a) = 1$  vorgegeben haben. Nach einmaliger Integration ergibt sich eine lineare Differentialgleichung, welche sich einfach lösen lässt und wir erhalten

$$h(x) = \frac{\exp(-2\mu x/\sigma^2) - \exp(2\mu b/\sigma^2)}{\exp(-2\mu a/\sigma^2) - \exp(2\mu b/\sigma^2)}. \quad (5.11)$$

Die Wahrscheinlichkeit, dass die Brownsche Bewegung mit Drift  $X_t = \sigma B_t + \mu t$  also  $a$

vor  $-b$  erreicht ist

$$\mathbb{P}[X_\tau = a] = \frac{\exp(-2\mu b/\sigma^2) - 1}{\exp(-2\mu(a+b)/\sigma^2) - 1}.$$

Beachte, dass sich dieses Ergebnis als

$$\mathbb{P}[X_\tau = a] = \frac{W(b)}{W(a+b)}$$

schreiben lässt, mit  $W(x) = \frac{1}{\mu}(1 - \exp(-2\mu x/\sigma^2))$ . Für  $\mu \rightarrow 0$  ergibt sich  $W(x) = 2x$  und somit die bekannte Lösung für den Fall ohne Drift. Die Funktion  $W$  heisst **Skalenfunktion** von  $X_t = \sigma B_t + \mu t$ . Eine interessante Folgerung ergibt sich nun mit  $b \rightarrow \infty$ . Im Grenzfall wird aus der Wahrscheinlichkeit  $a$  vor  $-b$  zu erreichen die Wahrscheinlichkeit  $a$  überhaupt in endlicher Zeit jemals zu erreichen. Diese Wahrscheinlichkeit lässt sich mit dem Supremumsprozess von  $X$  ausdrücken und wir erhalten

$$\mathbb{P}\left[\sup_{t \geq 0} X_t \leq a\right] = 1 - \lim_{b \rightarrow \infty} \frac{\exp(-2\mu b/\sigma^2) - 1}{\exp(-2\mu(a+b)/\sigma^2) - 1}.$$

Im Fall  $\mu < 0$  folgt daraus

$$\mathbb{P}\left[\sup_{t \geq 0} X_t \leq a\right] = 1 - \exp(2\mu a/\sigma^2), \quad a > 0, \mu < 0.$$

Die Verteilung des ultimativen Supremums  $\sup_{t \geq 0} X_t$  von  $X$  ist also im Fall  $\mu < 0$  eine Exponentialverteilung mit Mittelwert  $\frac{\sigma^2}{2|\mu|}$ .

Ein alternativer Zugang zum Austrittsproblem für Brownsche Bewegung mit Drift wird sich in einem späteren Kapitel mit dem Satz von Girsanov ergeben.

### 5.3 Die multivariate Ito-Formel und Anwendungen

Bevor wir zur multivariaten Ito-Formel kommen, bietet es sich an die **Differentialschreibweise** der Ito-Formel einzuführen. Nach der Ito-Formel mit Zeitabhängigkeit können wir den Prozess  $X_t = f(t, B_t)$  mit  $f \in C^{1,2}(\mathbb{R}_{\geq 0}, \mathbb{R})$  als

$$X_t = X_0 + \int_0^t \frac{\partial f}{\partial t}(s, B_s) ds + \int_0^t \frac{\partial f}{\partial x}(s, B_s) dB_s + \frac{1}{2} \int_0^t \frac{\partial^2 f}{\partial x^2}(s, B_s) ds \quad (5.12)$$

schreiben. Diese Formel schreiben wir ab nun in Differentialschreibweise als

$$dX_t = \frac{\partial f}{\partial t}(t, B_t) dt + \frac{\partial f}{\partial x}(t, B_t) dB_t + \frac{1}{2} \frac{\partial^2 f}{\partial x^2}(t, B_t) dt. \quad (5.13)$$

Die Motivation diese Schreibweise einzuführen ist ähnlich wie die der Leibnizschen Notation für Differentialquotienten in der gewöhnlichen Analysis: Sie ist kurz, prägnant und erlaubt eine für die Intuition hilfreiche Interpretation von  $dX_t$  als ‘infinitesimal kleine’ Differenz  $X_{t+\Delta t} - X_t$ . Dennoch sollte man sich immer bewusst sein, dass die mathematisch korrekte Leseart der Gleichung eben die Integralgleichung und damit die Rückführung auf das Ito-Integral ist.

Um die multivariate Ito-Formel einzuführen machen wir folgende Definition

**Definition.** Die  $d$ -dimensionale Brownsche Bewegung ist der  $\mathbb{R}^d$ -wertige stochastische Prozess

$$B_t = (B_t^1, \dots, B_t^d),$$

wobei  $B^1, \dots, B^d$  unabhängige (eindimensionale) Brownsche Bewegungen sind.

Es sollte in den meisten Fällen aus dem Zusammenhang klar sein, ob  $B$  nun eine eindimensionale oder eine  $d$ -dimensionale Brownsche Bewegung bezeichnet. Die Brownsche Standardfiltration der  $d$ -dimensionalen Brownschen Bewegung ist die Filtration die von  $B^1, \dots, B^d$  gemeinsam mit allen  $\mathbb{P}$ -Nullmengen erzeugt wird.

Wir betrachten nun Funktionen  $f : \mathbb{R}_{\geq 0} \times \mathbb{R}^d \rightarrow \mathbb{R}; (t, x) \mapsto f(t, x)$  mit einer Zeit- und einer  $d$ -dimensionalen Ortskomponente. Es sei an die Notation

$$\nabla f = \left( \frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_d} \right)$$

für den Gradienten von  $f$ , und

$$\Delta f = \frac{\partial^2 f}{\partial x_1^2} + \dots + \frac{\partial^2 f}{\partial x_d^2}$$

für den Laplace-Operator erinnert. Es gilt folgendes Resultat:

**Theorem 5.4** (Multivariate Ito-Formel). *Sei  $f \in C^{1,2}(\mathbb{R}_{\geq 0} \times \mathbb{R}^d)$  und  $B$  eine  $d$ -dimensionale Brownsche Bewegung. Dann gilt*

$$df(t, B_t) = \frac{\partial f}{\partial t}(t, B_t)dt + \nabla f(t, B_t)dB_t + \frac{1}{2}\Delta f(t, B_t)dt.$$

Das Zusammentreffen zweier Vektoren in  $\nabla f(t, B_t)dB_t$  ist als Vektorprodukt zu lesen. Der Beweis der Multivariaten Ito-Formel erfolgt analog zum Beweis der einfachen Ito-Formel; diesmal unter Verwendung der multivariaten Taylorentwicklung und der Unabhängigkeit der einzelnen Brownschen Bewegungen  $B^1, \dots, B^d$ . Ähnlich wie im

letzten Abschnitt ergibt sich aus der multivariaten Ito-Formel eine Verbindung zu partiellen Differentialgleichungen.

**Korollar 5.5.** Sei  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  eine zweifach stetig differenzierbare Funktion, welche die Laplace-Gleichung

$$\Delta f = 0 \tag{5.14}$$

erfüllt und  $B$  eine  $d$ -dimensionale Brownsche Bewegung. Dann ist  $X_t = f(B_t)$  ein lokales Martingal.

Funktionen die auf einem Gebiet  $D$  die Laplace-Gleichung  $\Delta f = 0$  erfüllen, heissen auch **harmonische Funktionen**.

Wir wollen nun die multivariate Ito-Formel verwenden um ein Phänomen der Brownschen Bewegung zu zeigen, das der Mathematiker George Pòlya folgendermaßen beschrieben hat:

“Ein betrunkenener Mensch findet immer nach Hause; ein betrunkenener Vogel möglicherweise nie.”

Diese Aussage spielt auf das unterschiedliche Verhalten der Brownschen Bewegung (der Pfad des Betrunkenen) in Dimension 2 und Dimension 3 an. Das mathematische Äquivalent der Aussage ist:

Die Brownsche Bewegung ist in Dimension 2 rekurrent; in Dimension 3 jedoch transient.

**Rekurrenz** und **Transienz** haben folgende Bedeutung: Wir starten einen stochastischen Prozess  $X$  im Ursprung und betrachten eine offene Kugel  $B_r$  von beliebigem Radius  $r > 0$  um den Startpunkt. Wenn der Prozess nach dem erstmaligen Verlassen von  $B_r$  mit Wahrscheinlichkeit 1 in endlicher Zeit wieder nach  $B_r$  zurückkehrt, so heisst er *rekurrent*. Wenn der Prozess nach dem erstmaligen Verlassen von  $B_r$  mit strikt positiver Wahrscheinlichkeit nie mehr nach  $B_r$  zurückkehrt, so heisst er *transient*.

Wir zeigen zuerst die Rekurrenz der 2-dimensionalen Brownschen Bewegung. Betrachte die Funktion  $f(x) = \log |x| = \log \sqrt{x_1^2 + x_2^2}$ , definiert auf  $\mathbb{R}^2 \setminus \{0\}$ . Wie sich einfach nachrechnen lässt, ist  $f$  eine harmonische Funktion. Wir wählen  $0 < r < R$  und betrachten den Kreisring  $A = \{x \in \mathbb{R}^2 : r < |x| < R\}$ . Durch skalieren von  $f$  erhalten wir die Funktion

$$h(x) = \frac{\log R - \log |x|}{\log R - \log r},$$

welche harmonisch auf  $A$  ist und folgende Randbedingungen erfüllt:

$$h(x) = \begin{cases} 1, & \text{wenn } |x| = r \\ 0, & \text{wenn } |x| = R. \end{cases}$$

Wir wählen einen beliebigen Punkt  $x_0 \in A$  und definieren

$$\tau_r = \inf \{t > 0 : |B_t + x_0| = r\}, \quad \tau_R = \inf \{t > 0 : |B_t + x_0| = R\}$$

und  $\tau = \tau_r \wedge \tau_R$ . Diese Stoppzeiten entsprechen den Austrittszeiten der in  $x_0 \in A$  gestarteten Brownschen Bewegung aus dem Kreisring  $A$  am inneren, äußeren bzw. an einem beliebigen Rand. Aus der multivariaten Ito-Formel und der harmonischen Eigenschaft von  $h$  erhalten wir, dass  $h(B_{t \wedge \tau} + x_0)$  ein lokales Martingal ist. Als beschränktes lokales Martingal ist es sogar ein echtes Martingal und wir erhalten mit  $t \rightarrow \infty$  und dominierter Konvergenz

$$\mathbb{E}[h(B_\tau + x_0)] = h(x_0) = \frac{\log R - \log |x_0|}{\log R - \log r}.$$

Aufgrund der Randbedingungen an  $h$  gilt aber auch

$$\mathbb{E}[h(B_\tau + x_0)] = \mathbb{P}[\tau_r < \tau_R]$$

und die obige Gleichung gibt uns die Wahrscheinlichkeit an, dass die bei  $x_0 \in A$  gestartete 2-dimensionale Brownsche Bewegung den inneren Rand des Kreisrings  $A$  erreicht (und somit in die offene Kugel  $B_r$  eintritt!) bevor sie den äußeren Rand erreicht. Um nun die Wahrscheinlichkeit zu berechnen, dass die Brownsche Bewegung überhaupt jemals die offene Kugel  $B_r$  erreicht, lassen wir  $R$  gegen unendlich gehen und erhalten

$$\mathbb{P}[\tau_r < \infty] = \lim_{R \rightarrow \infty} \mathbb{P}[\tau_r < \tau_R] = \lim_{R \rightarrow \infty} \frac{\log R - \log |x_0|}{\log R - \log r} = 1,$$

was die Rekurrenz der 2-dimensionalen Brownschen Bewegung zeigt.

Als nächstes wenden wir uns der Transienz der 3-dimensionalen Brownschen Bewegung zu. Der Allgemeinheit halber betrachten wir gleich die  $d$ -dimensionale Brownsche Bewegung mit  $d \geq 3$ . Wir stellen fest, dass nun die Funktion  $f(x) = |x|^{2-d} = \left(\sum_{i=1}^d x_i^2\right)^{(2-d)/2}$ , definiert auf  $\mathbb{R}^d \setminus \{0\}$  eine harmonische Funktion ist. Durch skalieren erhalten wir mit

$$h(x) = \frac{R^{2-d} - |x|^{2-d}}{R^{2-d} - r^{2-d}},$$

eine harmonische Funktion auf der Hohlkugel  $A = \{x \in \mathbb{R}^d : r < |x| < R\}$ , die eben-

falls den Randbedingungen

$$h(x) = \begin{cases} 1, & \text{wenn } |x| = r \\ 0, & \text{wenn } |x| = R \end{cases}$$

genügt. Dasselbe Martingalargument wie zuvor liefert

$$\mathbb{P}[\tau_r < \tau_R] = \mathbb{E}[h(B_\tau + x_0)] = h(x_0) = \frac{|R|^{2-d} - |x_0|^{2-d}}{R^{2-d} - r^{2-d}},$$

für  $x_0 \in A$ . Mit  $R \rightarrow \infty$  ergibt sich

$$\mathbb{P}[\tau_r < \infty] = \lim_{R \rightarrow \infty} \mathbb{P}[\tau_r < \tau_R] = \left(\frac{r}{|x_0|}\right)^{d-2} < 1$$

und die Transienz der  $d$ -dimensionalen Brownschen Bewegung für  $d \geq 3$  ist gezeigt.

## 5.4 Ito-Prozesse

Bis jetzt haben wir uns bei der Ito-Formel auf Funktionen  $f(t, B_t)$  der Brownschen Bewegung beschränkt. Nun wollen wir den Fall  $f(t, X_t)$  behandeln, wobei der Prozess  $X$  selbst durch ein stochastisches Integral gegeben ist. Die resultierende Formel ist die stochastische Entsprechung der Substitutionsregel, bzw. in Differentialnotation betrachtet, der Kettenregel. Zuerst definieren wir die geeignete Klasse von Prozessen  $X$ .

**Definition.** Ein **eindimensionaler Ito-Prozess** ist ein stochastischer Prozess der Gestalt

$$X_t = x_0 + \int_0^t a(\omega, s) ds + \int_0^t b(\omega, s) dB_s \quad (5.15)$$

wobei  $a$  und  $b$  reellwertige, adaptierte stochastischer Prozesse mit

$$\mathbb{P}\left[\int_0^t |a(\omega, s)| ds < \infty\right] = 1 \quad \text{und} \quad \mathbb{P}\left[\int_0^t b(\omega, s)^2 ds < \infty\right] = 1, \quad \forall t \geq 0 \quad (5.16)$$

sind.

Die Bedingungen (5.16) garantieren, dass sowohl das gewöhnliche als auch das Ito-Integral in (5.15) mit Wahrscheinlichkeit 1 wohldefiniert und endlich sind. Wenn nun  $f(\omega, t)$  ein weiterer adaptierter stochastischer Prozess ist, mit den Eigenschaften

$$\mathbb{P}\left[\int_0^t |a(\omega, s)f(\omega, s)| ds < \infty\right] = 1 \quad \text{und} \quad \mathbb{P}\left[\int_0^t (b(\omega, s)f(\omega, s))^2 ds < \infty\right] = 1, \quad (5.17)$$

dann können wir ein stochastisches Integral von  $f(\omega, t)$  bezüglich  $X$  folgendermaßen definieren

$$\int_0^t f(s, \omega) dX_s := \int_0^t f(\omega, s) a(\omega, s) ds + \int_0^t f(\omega, s) b(\omega, s) dB_s.$$

Diese Konvention liest sich erheblich leichter in Differentialnotation. Da besagt sie nämlich dass wir

$$dX_t = a(\omega, t)dt + b(\omega, t)dB_t$$

unter der Integrierbarkeitsbedingung (5.17) in die Gleichung

$$f(\omega, t)dX_t = f(\omega, t)a(\omega, t)dt + f(\omega, t)b(\omega, t)dB_t$$

überführen dürfen.

Wir können nun eine verallgemeinerte Form der Ito-Formel für Funktionen  $f(t, X_t)$  eines eindimensionalen Ito-Prozesses  $X$  angeben.

**Theorem 5.6** (Ito-Formel für Ito-Prozesse). *Sei  $X$  ein eindimensionaler Ito-Prozess und  $f \in C^{1,2}(\mathbb{R}_{\geq 0} \times \mathbb{R})$ . Dann ist auch  $f(t, X_t)$  ein Ito-Prozess und es gilt*

$$df(t, X_t) = \frac{\partial f}{\partial t}(t, X_t)dt + \frac{\partial f}{\partial x}(t, X_t)dX_t + \frac{1}{2} \frac{\partial^2 f}{\partial x^2}(t, X_t)b(\omega, t)^2 dt.$$

Der Beweis der Ito-Formel für Ito-Prozesse läuft nach dem selben Schema ab wie der Beweis der Ito-Formel für Brownsche Bewegung. Die entscheidende Beobachtung ist, dass wir statt der Martingaleigenschaft von  $B_t^2 - t$  – welche im Beweis der Ito-Formel für Brownsche Bewegung benötigt wird – nun verwenden können dass der Prozess

$$\left( \int_0^t b(\omega, s) dB_s \right)^2 - \int_0^t b(\omega, s)^2 ds$$

ein lokales Martingal ist, was wir durch die bedingte Ito-Isometrie bereits nachgewiesen haben.

Die Ito-Formel für Ito-Prozesse lässt sich in noch prägnantere Form bringen wenn wir auf den quadratischen Variationsprozess von  $X$  zurückgreifen.

**Definition.** Sei  $X$  ein stochastischer Prozess, und  $\mathcal{P}$  eine Partition von  $[0, t]$ . Die quadratische Variation von  $X$  entlang der Partition  $\mathcal{P}$  ist gegeben durch

$$Q_{\mathcal{P}}(X)_t = \sum_{t_i \in \mathcal{P}} (X_{t_{i+1}} - X_{t_i})^2.$$

Wenn für jede Partitionsfolge  $(\mathcal{P}^n)_{n \in \mathbb{N}}$  mit  $|\mathcal{P}^n| \rightarrow 0$  der Limes

$$[X, X]_t := \lim_{|\mathcal{P}^n| \rightarrow 0} Q_{\mathcal{P}^n}(X)_t, \quad \text{in Wahrscheinlichkeit}$$

existiert und für alle Partitionsfolgen fast sicher den gleichen Wert annimmt, so heisst  $[X, X]_t$  **quadratische Variation** von  $X$  und  $t \mapsto [X, X]_t$  **quadratischer Variationsprozess** von  $X$ .

Wie wir nächsten Abschnitt zeigen werden existiert die quadratische Variation eines Ito-Prozesses und es gilt

$$[X, X]_t = \int_0^t b(\omega, s)^2 ds.$$

Damit können wir die Ito-Formel für Ito-Prozesse folgendermaßen umschreiben:

**Korollar 5.7** (Ito-Formel für Ito-Prozesse – alternative Form). *Sei  $X$  ein Ito-Prozess und  $f \in C^{1,2}(\mathbb{R}_{\geq 0} \times \mathbb{R})$ . Dann ist auch  $f(t, X_t)$  ein Ito-Prozess und es gilt*

$$df(t, X_t) = \frac{\partial f}{\partial t}(t, X_t)dt + \frac{\partial f}{\partial x}(t, X_t)dX_t + \frac{1}{2} \frac{\partial^2 f}{\partial x^2}(t, X_t)d[X, X]_t.$$

Schließlich betrachten wir noch die multivariate Verallgemeinerung der Ito-Formel für Ito-Prozesse.

**Definition.** Sei  $B$  eine  $d$ -dimensionale Brownsche Bewegung, und seien  $a(\omega, t)$  und  $b(\omega, t)$  adaptierte stochastische Prozesse, welche Werte in  $\mathbb{R}^k$  bzw. den  $k \times d$ -Matrizen  $\mathbb{R}^{(k \times d)}$  annehmen. Zusätzlich gelte mit  $\rho(\omega, t) = b(\omega, t)b(\omega, t)^\top$

$$\mathbb{P} \left[ \int_0^t |a_i(\omega, s)| ds < \infty \right] = 1 \quad \text{und} \quad \mathbb{P} \left[ \int_0^t \rho_{ii}(\omega, s) ds < \infty \right] = 1 \quad (5.18)$$

für alle  $i \in \{1, \dots, k\}$ . Dann heisst der  $\mathbb{R}^k$ -wertige Prozess

$$X_t = x_0 + \int_0^t a(\omega, s) ds + \int_0^t b(\omega, s) dB_s$$

(multivariater) **Ito-Prozess**.

Der Leser ist an dieser Stelle aufgefordert sich zu überzeugen, dass die Dimensionen sämtlicher Vektoren und Matrizen in obiger Definition tatsächlich zusammenpassen. Die Brownsche Bewegung, und die Prozesse  $X$  und  $a$  sind dabei als Spaltenvektoren zu betrachten. Um die Ito-Formel für den multivariaten Ito-Prozess  $X$  zu formulieren benötigen wir noch die **quadratische Kovariation**  $[X^i, X^j]$  seiner Komponenten  $(X^1, \dots, X^k)$ . Diese ist analog zur quadratischen Variation über den Limes in Wahr-



scheinlichkeit der Summen

$$C_{\mathcal{P}}(X^i, X^j)_t = \sum_{t_l \in \mathcal{P}} (X_{t_{l+1}}^i - X_{t_l}^i)(X_{t_{l+1}}^j - X_{t_l}^j)$$

definiert. Für Ito-Prozesse existiert die quadratische Kovariation, und es gilt

$$[X^i, X^j]_t = \int_0^t \rho_{ij}(s, \omega) ds \quad \text{mit} \quad \rho(\omega, s) = b(\omega, s)b(\omega, s)^\top.$$

**Theorem 5.8** (Ito-Formel für multivariate Ito-Prozesse). *Sei  $X$  ein  $k$ -dimensionaler Ito-Prozess und  $f \in C^{1,2}(\mathbb{R}_{\geq 0} \times \mathbb{R}^k)$ . Dann ist auch  $f(t, X_t)$  ein Ito-Prozess und es gilt*

$$df(t, X_t) = \frac{\partial f}{\partial t}(t, X_t)dt + \nabla f(t, X_t)dX_t + \frac{1}{2} \sum_{i,j=0}^k \frac{\partial^2 f}{\partial x_i \partial x_j}(t, X_t)d[X^i, X^j]_t.$$

Im Spezialfall  $k = 2$  und  $f(y, z) = yz$  ergibt sich eine Formel die als partielle Integration für das Ito-Integral gelesen werden kann.

**Korollar 5.9.** *Es sei  $(Y, Z)$  ein  $\mathbb{R}^2$ -wertiger Ito-Prozess. Dann gilt*

$$Y_t Z_t = Y_0 Z_0 + \int_0^t Y_s dZ_s + \int_0^t Z_s dY_s + [Y, Z]_t.$$

## 5.5 Die quadratische Variation von Ito-Prozessen

In diesem Abschnitt beweisen wir die zuvor aufgestellten Behauptungen über die quadratische Variation von Ito-Prozessen.

**Theorem 5.10.** *Sei  $X$  ein eindimensionaler Ito-Prozess mit der Darstellung*

$$X_t = x_0 + \int_0^t a(\omega, s)ds + \int_0^t b(\omega, s)dB_s.$$

*Dann existiert die quadratische Variation von  $X$  und ist gegeben durch*

$$[X, X]_t = \int_0^t b(\omega, s)^2 ds.$$

Für den Beweis verwenden wir folgendes Lemma:

**Lemma 5.11.** *Sei  $a(\omega, s)$  ein stochastischer Prozess mit  $\int_0^t |a(\omega, s)|ds < \infty$  fast sicher. Dann ist der quadratische Variationsprozess von  $A = \int_0^t a(\omega, s)ds$  gleich 0.*

*Beweis.* Wir betrachten den Stetigkeitsmodul

$$R_\delta(\omega) := \sup \left\{ \int_r^s |a(\omega, u)| du : r, s \in [0, t] \text{ und } |r - s| < \delta \right\}.$$

Aufgrund der Stetigkeit von  $s \mapsto \int_0^s |a(\omega, u)| du$  gilt  $\lim_{\delta \rightarrow 0} R_\delta = 0$  fast sicher. Aus der Abschätzung

$$\mathcal{Q}_{\mathcal{P}}(A)_t = \sum_{t_i \in \mathcal{P}} \left( \int_{t_i}^{t_{i+1}} a(\omega, s) ds \right)^2 \leq R_{|\mathcal{P}|} \int_0^t |a(\omega, s)| ds$$

folgt mit  $|\mathcal{P}| \rightarrow 0$  die Behauptung.  $\square$

*Beweis von Theorem 5.10. Schritt 1:* Wir zerlegen den Prozess  $X$  in  $X_t = x_0 + A_t + Z_t$  mit  $A_t = \int_0^t a(\omega, s) ds$  und  $Z_t = \int_0^t b(\omega, s) dB_s$ . Zunächst nehmen wir an, dass eine Konstante  $C$  existiert, sodass

$$|Z_s| \leq C \quad \text{und} \quad \int_0^t b(\omega, s)^2 ds \leq C, \quad \forall 0 \leq s \leq t \quad (5.19)$$

fast sicher gilt und zeigen die Abschätzungen

$$\sup_{\mathcal{P}} \mathbb{E} [Q_{\mathcal{P}}^2(Z)_t] \leq 12C^2 \mathbb{E} [(Z_t - Z_0)^2] \quad (5.20)$$

und

$$\lim_{|\mathcal{P}| \rightarrow 0} \sum_{i=1}^n \mathbb{E} [(Z_{t_{i+1}} - Z_{t_i})^4] = 0. \quad (5.21)$$

Da das lokale Martingal  $Z$  beschränkt ist, ist es ein echtes Martingal und es gilt

$$\begin{aligned} \mathbb{E} [(Z_{t_{j+1}} - Z_{t_j})^2 | \mathcal{F}_{t_i}] &= \mathbb{E} \left[ \mathbb{E} \left[ Z_{t_{j+1}}^2 - 2Z_{t_{j+1}}Z_{t_j} + Z_{t_j}^2 \middle| \mathcal{F}_{t_j} \right] \middle| \mathcal{F}_{t_i} \right] = \\ &= \mathbb{E} \left[ Z_{t_{j+1}}^2 - Z_{t_j}^2 \middle| \mathcal{F}_{t_i} \right] \end{aligned}$$

für alle  $i \leq j$  in  $\{1, \dots, n\}$ . Daraus folgt, dass

$$\mathbb{E} \left[ \sum_{j=i}^{n-1} (Z_{t_{j+1}} - Z_{t_j})^2 \middle| \mathcal{F}_{t_i} \right] = \mathbb{E} \left[ \sum_{j=i}^{n-1} (Z_{t_{j+1}}^2 - Z_{t_j}^2) \middle| \mathcal{F}_{t_i} \right] = \mathbb{E} [Z_t^2 - Z_{t_i}^2 | \mathcal{F}_{t_i}] \leq 4C^2. \quad (5.22)$$

Das Quadrat von  $Q_{\mathcal{P}}(Z)_t$  lässt sich umschreiben als

$$Q_{\mathcal{P}}(Z)_t^2 = \sum_{i=0}^{n-1} (Z_{t_{i+1}} - Z_{t_i})^4 + 2 \sum_{i=0}^{n-1} \left( (Z_{t_{i+1}} - Z_{t_i})^2 \sum_{j=i+1}^{n-1} (Z_{t_{j+1}} - Z_{t_j})^2 \right)$$

und nach mehrfacher Anwendung von (5.22) erhalten wir

$$\begin{aligned} \mathbb{E} [Q_{\mathcal{P}}(X)_t^2] &\leq 4C^2 \mathbb{E} \left[ \sum_{i=0}^{n-1} (Z_{t_{i+1}} - Z_{t_i})^2 \right] + 8C^2 \mathbb{E} \left[ \sum_{i=0}^{n-1} (Z_{t_{i+1}} - Z_{t_i})^2 \right] = \\ &= 12C^2 \mathbb{E} [(Z_t - Z_0)^2] \end{aligned}$$

wie in (5.20) behauptet. Als nächstes betrachten wir den Stetigkeitsmodul

$$R_{\delta}(\omega) := \sup \{ |Z_r - Z_s| : r, s \in [0, t] \text{ und } |r - s| < \delta \}.$$

Es gilt  $|R_{\delta}| \leq 2C$  und aufgrund der Stetigkeit der Pfade  $t \mapsto Z_t(\omega)$  auch  $\lim_{\delta \rightarrow 0} R_{\delta} = 0$  fast sicher. Wir schätzen mit

$$\mathbb{E} \left[ \sum_{i=1}^n (Z_{t_{i+1}} - Z_{t_i})^4 \right] \leq \mathbb{E} [R_{|\mathcal{P}|}^2] \mathbb{E} [Q_{\mathcal{P}}^2(Z)_t]$$

ab. Der Term  $\mathbb{E} [Q_{\mathcal{P}}^2(Z)_t]$  ist durch  $12C^2$  beschränkt; mit dominierter Konvergenz gilt  $\lim_{|\mathcal{P}| \rightarrow 0} \mathbb{E} [R_{|\mathcal{P}|}^2] = 0$  und wir folgern (5.21).

Schritt 2: Definiere

$$\Delta_{\mathcal{P}} = Q_{\mathcal{P}}(Z)_t - \int_0^t b(\omega, s) ds = \sum_{i=0}^{n-1} \left( (Z_{t_{i+1}} - Z_{t_i})^2 - \int_{t_i}^{t_{i+1}} b^2(\omega, s) ds \right).$$

Wir wollen zeigen dass  $\lim_{|\mathcal{P}| \rightarrow 0} \Delta_{\mathcal{P}} = 0$  in Wahrscheinlichkeit gilt. Aus der Martingaleigenschaft von  $Z_t^2 - \int_0^t b(\omega, s) ds$  folgern wir die Orthogonalität der Summanden und erhalten

$$\mathbb{E} [\Delta_{\mathcal{P}}^2] = \sum_{i=0}^{n-1} \mathbb{E} \left[ \left( (Z_{t_{i+1}} - Z_{t_i})^2 - \int_{t_i}^{t_{i+1}} b^2(\omega, s) ds \right)^2 \right].$$

Mit der Ungleichung  $(x + y)^2 \leq 2x^2 + 2y^2$  folgt

$$\mathbb{E} [\Delta_{\mathcal{P}}^2] \leq 2 \sum_{i=0}^{n-1} \mathbb{E} [(Z_{t_{i+1}} - Z_{t_i})^4] + 2 \sum_{i=0}^{n-1} \mathbb{E} \left[ \left( \int_{t_i}^{t_{i+1}} b^2(\omega, s) ds \right)^2 \right].$$

Wir haben bereits in Schritt 1 gezeigt, dass der erste Summand gegen 0 konvergiert.

Für den zweiten Summanden betrachten wir den Stetigkeitsmodul

$$R'_\delta(\omega) := \sup \left\{ \int_r^s b^2(\omega, u) du : r, s \in [0, t] \text{ und } |r - s| < \delta \right\}.$$

Wieder gilt  $R'_\delta \leq C$  und  $\lim_{\delta \rightarrow 0} R'_\delta = 0$  fast sicher. Wir folgern

$$\sum_{i=0}^{n-1} \mathbb{E} \left[ \left( \int_{t_i}^{t_{i+1}} b^2(\omega, s) ds \right)^2 \right] \leq \mathbb{E} [R'_{|\mathcal{P}|}] \mathbb{E} \left[ \int_0^t b^2(\omega, s) ds \right] \leq C \mathbb{E} [R'_{|\mathcal{P}|}]$$

und damit  $\mathbb{E} [\Delta_{\mathcal{P}}^2] \rightarrow 0$ . Es folgt die Konvergenz

$$\lim_{|\mathcal{P}| \rightarrow 0} Q_{\mathcal{P}}(Z)_t \rightarrow \int_0^t b^2(\omega, s) ds, \quad \text{in Wahrscheinlichkeit.}$$

Schritt 3: Wir kehren nun zum Prozess  $X_t = x_0 + A_t + Z_t$  zurück. Für diesen können wir

$$Q_{\mathcal{P}}(X)_t = Q_{\mathcal{P}}(A)_t + Q_{\mathcal{P}}(Z)_t + C_{\mathcal{P}}(A, Z)_t,$$

schreiben, mit

$$C_{\mathcal{P}}(A, Z)_t = \sum_{i=0}^{n-1} \left( \int_{t_i}^{t_{i+1}} a(\omega, s) ds \right) \left( \int_{t_i}^{t_{i+1}} b(\omega, s) dB_s \right).$$

Nach Lemma 5.11 gilt  $Q_{\mathcal{P}}(A)_t \rightarrow 0$ ; mit der Cauchy-Schwarzschen Ungleichung erhalten wir

$$C_{\mathcal{P}}(A, Z)_t^2 \leq Q_{\mathcal{P}}(A)_t \cdot Q_{\mathcal{P}}(Z)_t$$

und damit  $C_{\mathcal{P}}(A, Z) \rightarrow 0$ . Insgesamt haben wir nun

$$\lim_{|\mathcal{P}| \rightarrow 0} Q_{\mathcal{P}}(X)_t = \int_0^t b(\omega, s)^2 ds$$

gezeigt, unter den Beschränktheitsvoraussetzungen in (5.19).

Schritt 4: Der letzte Schritt ist ein Lokalisierungsargument um die Annahmen (5.19) aufzuheben. Dieses Argument soll hier nicht näher ausgeführt werden, siehe z.B. Steele [2003, Seiten 132f].  $\square$

Mit einem ähnlichen Beweis lässt sich auch folgende multivariate Verallgemeinerung zeigen.

**Theorem 5.12.** *Sei  $X$  ein  $k$ -dimensionaler Ito-Prozess mit der Darstellung*

$$X_t = x_0 + \int_0^t a(\omega, s) ds + \int_0^t b(\omega, s) dB_s.$$

*Dann existieren die quadratischen Kovariationen  $[X^i, X^j]_t$  und sind gegeben durch*

$$[X^i, X^j]_t = \int_0^t \rho_{ij}(\omega, s) ds, \quad i, j \in \{1, \dots, k\}$$

*mit  $\rho(\omega, t) = b(\omega, t)b(\omega, t)^\top$ .*

## 6 Stochastische Differentialgleichungen

Viele deterministische Modelle der Natur- und der Wirtschaftswissenschaften lassen sich mit Hilfe von Differentialgleichungen ausdrücken. Mit dem Ito-Integral und der Ito-Formel haben wir die Grundlagen für **stochastische Differentialgleichungen** gelegt, mit denen sich nun auch viele stochastische Modelle mathematisch formulieren lassen. Gegeben einen Wahrscheinlichkeitsraum mit Brownscher Bewegung und Brownscher Standardfiltration sowie zwei stetige Funktionen  $\mu : \mathbb{R}_{\geq 0} \times \mathbb{R} \rightarrow \mathbb{R}$  und  $\sigma : \mathbb{R}_{\geq 0} \times \mathbb{R} \rightarrow \mathbb{R}_{\geq 0}$  nennen wir die Gleichung

$$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dB_t, \quad X_0 = x_0$$

**stochastische Differentialgleichung.** Für diesen Begriff verwenden durchwegs die englische Abkürzung SDE (*stochastic differential equation*). In Integralform geschrieben lautet die Gleichung

$$X_t = x_0 + \int_0^t \mu(s, X_s)ds + \int_0^t \sigma(s, X_s)dB_s$$

und ein stetiger adaptierter stochastischer Prozess  $X$  der diese Gleichung erfüllt heißt Lösung der SDE. Es ist einfach zu sehen, dass jede Lösung automatisch ein Ito-Prozess im Sinne des letzten Kapitels ist. Wir nennen  $\mu(t, X_t)$  Driftkoeffizient und  $\sigma(t, X_t)$  Diffusionskoeffizient der Gleichung. Es sei nochmal an die unterschiedliche Rolle der zwei Terme bei der Beschreibung des ‘lokalen Trends’ und der ‘lokalen Variabilität’ von  $X$  erinnert, wie sie am Beginn von Kapitel 5 diskutiert wurde.

### 6.1 Beispiele Stochastischer Differentialgleichungen

#### Geometrische Brownsche Bewegung

Wir betrachten die stochastische Differentialgleichung

$$dX_t = \mu X_t dt + \sigma X_t dB_t, \quad X_0 = x_0 \tag{6.1}$$

mit  $\mu \in \mathbb{R}, \sigma > 0$  und  $x_0 \geq 0$ . Diese SDE beschreibt einen stochastischen Prozess dessen lokales Wachstum sich genauso wie seine lokale Variabilität *proportional* zum augenblicklichen Wert verhält. Anders gesagt sind die *relativen* Änderungen des Prozesses von konstanter Größenordnung. Diese Hypothese ist etwa in finanzmathematischen Modellen, z.B. bei der Beschreibung eines Aktienkurses zutreffend, wo Kursgewinne bzw. -verluste immer relativ zum Basiswert zu interpretieren sind.

Als einfachster Ansatz zum Auffinden einer Lösung bietet sich  $X_t = f(t, B_t)$  mit  $f$

ausreichend differenzierbar an. Nach der Ito-Formel gilt<sup>8</sup>

$$dX_t = \left( f_t(t, X_t) + \frac{1}{2} f_{xx}(t, X_t) \right) dt + f_x(t, X_t) dB_t,$$

und ein Koeffizientenvergleich mit (6.1) liefert

$$\sigma f(t, x) = f_x(t, x), \quad \mu f(t, x) = f_t(t, x) + \frac{1}{2} f_{xx}(t, x).$$

Die Lösung der ersten Gleichung ist von der Gestalt  $f(t, x) = \exp(\sigma x + g(t))$ ; einsetzen in die zweite Gleichung liefert  $g'(t) = \mu - \frac{\sigma^2}{2}$ . Wir erhalten also als Lösung der stochastischen Differentialgleichung (6.1) den Prozess

$$X_t = x_0 \exp \left( \sigma B_t + \left( \mu - \frac{\sigma^2}{2} \right) t \right),$$

auch **geometrische Brownsche Bewegung** genannt und Basis für das Finanzmarktmodell von Black & Scholes. Es sei bemerkt, dass wir an dieser Stelle noch nicht ausschließen können, dass auch noch andere Lösungen der SDE (6.1) existieren; wir werden aber im nächsten Abschnitt zeigen, dass dies nicht der Fall ist.

### Der Ornstein-Uhlenbeck-Prozess

Als nächstes wenden wir uns der stochastischen Differentialgleichung

$$dX_t = -\lambda X_t dt + \sigma dB_t, \quad X_0 = x_0 \tag{6.2}$$

mit  $\lambda, \sigma > 0$  und  $x_0 \in \mathbb{R}$  zu. Wir stellen fest, dass bei dieser SDE der Diffusionskoeffizient im Gegensatz zu (6.1) *nicht* proportional zum augenblicklichen Wert, sondern immer von konstanter Größenordnung  $\sigma$  ist. Der Driftterm  $-\lambda X_t$  ist negativ, wenn  $X_t$  positiv ist und positiv wenn  $X_t$  negativ ist, mit absoluter Größe proportional zu  $|X_t|$ . Wir können also davon ausgehen, dass der Prozess  $X$  langfristig gesehen trotz stochastischer Fluktuationen dazu tendiert in die Nähe der Null zurückzukehren. Diese Phänomen wird auch mit *mean-reversion* (Rückkehr zum Mittelwert) bezeichnet.

Das stochastische Modell (6.2) wurde von den Physikern Ornstein und Uhlenbeck zur Modellierung der Geschwindigkeit einzelner Gasmoleküle verwendet. Zugrundeliegende Annahme ist dabei dass die Geschwindigkeit eines einzelnen Moleküls zwar zufällig fluktuiert, aber letztendlich nicht zu stark von der Durchschnittsgeschwindigkeit der Gesamtheit aller Moleküle abweichen darf, und daher eine Tendenz der Rückkehr zum

<sup>8</sup>Zur Vereinfachung der Notation schreiben wir ab nun die partiellen Ableitungen mittels tiefgestellten Indices, d.h.  $f_t := \frac{\partial f}{\partial t}$ ,  $f_{xx} := \frac{\partial^2 f}{\partial x^2}$ , etc.

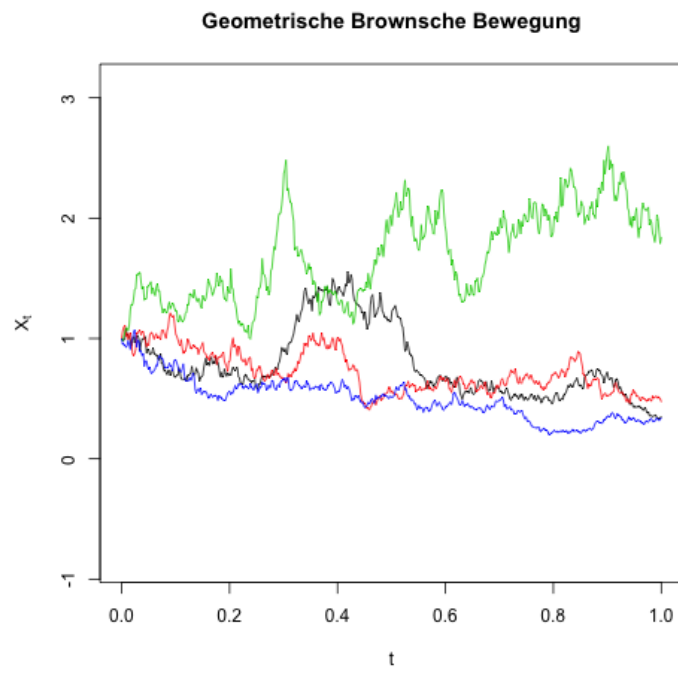


Abbildung 1: Vier Pfade einer geometrischen Brownschen Bewegung mit Parametern  $\mu = -0.1, \sigma = 1$ .



Mittelwert (mean-reversion) aufweisen muss. Eine weitere wichtige Anwendung findet die SDE (6.5) in der Finanzmathematik, wo sie als Vasicek-Modell für Zinsraten bekannt ist.

Der einfache Ansatz  $X_t = f(t, B_t)$  führt bei dieser SDE nicht mehr zum Ziel. Man könnte jedoch vermuten, dass es sich aufgrund der einfachen Struktur des Diffusions-terms bei  $X$  um einen Gaußschen Prozess handelt. Wir wissen aus Kapitel 4, dass Ito-Integrale von deterministischen Integranden, d.h. von der Gestalt  $\int_0^t b(s)dB_s$  Beispiele für Gaußsche Prozesse sind. Um eine etwas allgemeinere Klasse von Gaußschen Prozessen abzudecken wählen wir den Ansatz

$$X_t = a(t) \left\{ x_0 + \int_0^t b(s)dB_s \right\} \quad (6.3)$$

mit  $a$  stetig differenzierbar und Startwert  $a(0) = 1$ . Beachte dass  $X$  das Produkt der zwei Ito-Prozesse  $Y_t = a(t)$  und  $Z_t = x_0 + \int_0^t b(s)dB_s$  ist. Des weiteren gilt  $[Y, Z]_t = 0$  und wir erhalten nach Anwendung der Produktformel aus Korrolar 5.9

$$dX_t = a'(t) \left\{ x_0 + \int_0^t b(s)dB_s \right\} dt + a(t)b(t)dB_t.$$

Mit der Annahme  $a(t) > 0$  lässt sich dies umschreiben als

$$dX_t = \frac{a'(t)}{a(t)} X_t dt + a(t)b(t)dB_t.$$

Koeffizientenvergleich mit (6.2) liefert

$$\frac{a'(t)}{a(t)} = -\lambda, \quad a(t)b(t) = \sigma.$$

Die Lösungen dieser Gleichungen sind  $a(t) = e^{-\lambda t}$  und  $b(t) = \sigma e^{\lambda t}$ . Aus dem Ansatz (6.3) erhalten wir also

$$X_t = e^{-\lambda t} \left\{ x_0 + \sigma \int_0^t e^{\lambda s} dB_s \right\} = e^{-\lambda t} x_0 + \sigma \int_0^t e^{\lambda(s-t)} dB_s$$

als Lösung der Ornstein-Uhlenbeck SDE (6.2). Aus dieser Darstellung folgt mit den Eigenschaften des Ito-Integrals

$$\mathbb{E}[X_t] = e^{-\lambda t} x_0 \quad \text{und} \quad \text{Var}(X_t) = \sigma^2 \int_0^t e^{2\lambda(s-t)} ds = \frac{\sigma^2}{2\lambda} (1 - e^{-2\lambda t}).$$

Für  $t \rightarrow \infty$  gilt also  $\mathbb{E}[X_t] \rightarrow 0$  und  $\text{Var}(X_t) \rightarrow \frac{\sigma^2}{2\lambda}$ . Da  $X$  ein Gaußscher Prozess ist,

schließen wir daraus dass die Verteilung von  $X_t$  für  $t \rightarrow \infty$  gegen eine Normalverteilung mit Mittelwert 0 und Varianz  $\frac{\sigma^2}{2\lambda}$  konvergiert. Die Existenz solch einer Grenzverteilung für  $t \rightarrow \infty$  unterscheidet den Ornstein-Uhlenbeck wesentlich von anderen stochastischen Prozessen wie der (gewöhnlichen) Brownschen Bewegung oder der geometrischen Brownschen Bewegung.

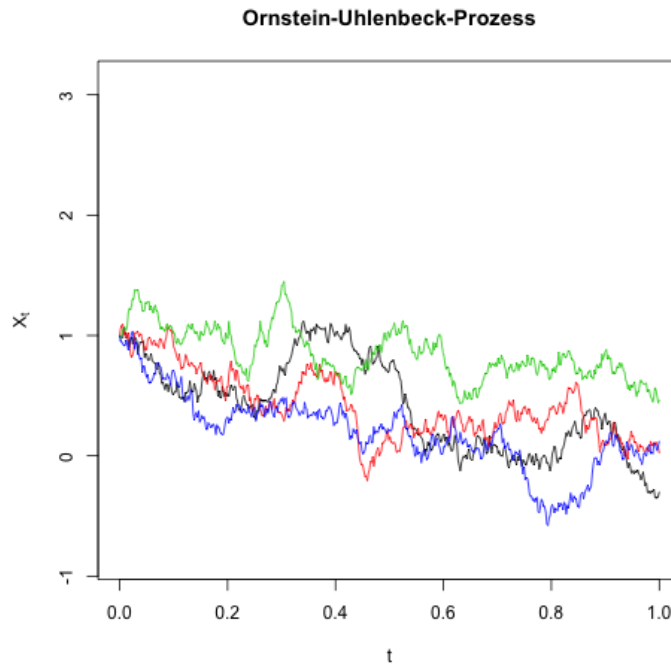


Abbildung 2: Vier Pfade eines Ornstein-Uhlenbeck-Prozesses mit Parametern  $\lambda = 2$ ,  $\sigma = 1$ .

### Die Brownsche Brücke

Wir betrachten die Differentialgleichung

$$dX_t = -\frac{1}{1-t}X_t dt + \sigma dB_t, \quad X_0 = x_0, t \in [0, 1]. \quad (6.4)$$

Der Einfachheit halber lassen wir  $X$  bei 0 starten, d.h. wir setzen  $x_0 = 0$ . Wie schon beim Ornstein-Uhlenbeck-Prozess ist das Vorzeichen des Driftterms dem Vorzeichen von  $X_t$  genau entgegengesetzt, der Prozess zeigt also eine Tendenz zur Null

zurückzukehren. Hier wächst aber der Driftterm umso stärker je näher sich  $t$  der Eins annähert, und ist für  $t = 1$  sogar singulär. Aus diesem Grund betrachten wir die SDE auch nur auf dem Intervall  $[0, 1)$ . Wegen der ähnlichen Gestalt können wir auch hier probieren den Ansatz (6.3) anzuwenden und erhalten die Gleichungen

$$\frac{a'(t)}{a(t)} = -\frac{1}{1-t}, \quad a(t)b(t) = \sigma.$$

Die Lösungen dieser Gleichungen sind

$$a(t) = 1-t \quad \text{und} \quad b(t) = \frac{\sigma}{1-t}$$

und somit gilt

$$X_t = \sigma(1-t) \int_0^t \frac{1}{1-s} dB_s.$$

Wieder lassen sich Erwartungswert und Autokovarianz dieses Gaußschen Prozesses leicht berechnen und wir erhalten mit  $0 \leq s \leq t < 1$

$$\mathbb{E}[X_t] = 0 \quad \text{und} \quad \text{Cov}(X_t, X_s) = \sigma^2(1-t)(1-s) \int_0^{s \wedge t} \frac{1}{(1-u)^2} du = \sigma^2 s(1-t).$$

Insbesondere gilt  $\mathbb{E}[X_t^2] = \sigma^2 t(1-t)$  und somit  $\lim_{t \rightarrow 1} \mathbb{E}[X_t^2] = 0$ . Wir können sogar zeigen, dass der Prozess  $X_t$  für  $t \rightarrow 1$  fast sicher gegen 0 konvergiert. Angenommen es gäbe eine Folge  $t_n \rightarrow 1$ , sodass mit positiver Wahrscheinlichkeit  $\lim_{n \rightarrow \infty} X_{t_n}^2 > 0$  gilt. Dann erhalten wir mit dem Lemma von Fatou

$$0 < \mathbb{E} \left[ \liminf_{n \rightarrow \infty} X_{t_n}^2 \right] \leq \liminf_{n \rightarrow \infty} \mathbb{E}[X_{t_n}^2] = 0$$

und somit einen Widerspruch. Wir können also  $X$  zu einem stetigen stochastischen Prozess auf dem abgeschlossenen Intervall  $[0, 1]$  fortsetzen, indem wir den Wert  $X_1 = 0$  hinzufügen. Die so definierte **Brownsche Brücke** ist ein Gaußscher Prozess der bei  $X_0 = 0$  startet und auch fast sicher wieder bei  $X_1 = 0$  endet.

## 6.2 Existenz und Eindeutigkeit von Lösungen

Häufig lassen sich die Lösungen von stochastischen Differentialgleichungen nicht in geschlossener Form angeben. Dennoch können wir hinreichende Bedingungen für die Existenz und Eindeutigkeit einer Lösung formulieren.

**Theorem 6.1.** *Wenn die Koeffizienten der stochastischen Differentialgleichung*

$$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dB_t, \quad X_0 = x_0, t \in [0, T] \quad (6.5)$$

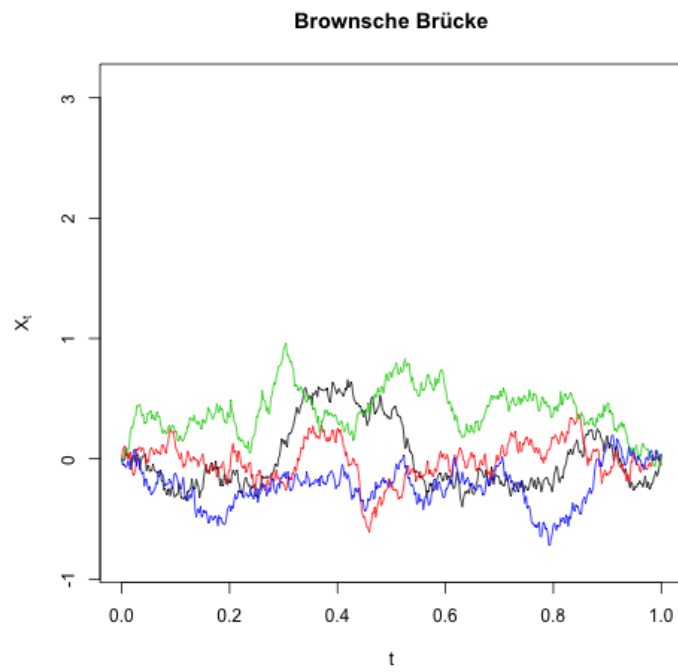


Abbildung 3: Vier Pfade einer Brownschen Brücke mit Parameter  $\sigma = 1$ .

die Lipschitz-Bedingung

$$|\mu(t, x) - \mu(t, y)|^2 + |\sigma(t, x) - \sigma(t, y)|^2 \leq K |x - y|^2 \quad (6.6)$$

und die Wachstumsbedingung

$$|\mu(t, x)|^2 + |\sigma(t, x)|^2 \leq K(1 + |x|^2) \quad (6.7)$$

für alle  $(t, x) \in [0, T] \times \mathbb{R}$  erfüllen, dann existiert ein stetiger adaptierter Prozess der (6.5) erfüllt und welcher gleichmäßig in  $L^2(dP)$  beschränkt ist:

$$\sup_{t \in [0, T]} \mathbb{E} [X_t^2] < \infty.$$

Sei  $Y$  eine weitere stetige und gleichmäßig in  $L^2(dP)$  beschränkte Lösung von (6.5), dann sind  $X$  und  $Y$  ununterscheidbar.

*Beweis von Theorem 6.1: Eindeutigkeit.* Der einfachere Teil des Beweises ist die Eindeutigkeit. Wir nehmen an dass zwei stetige gleichmäßig in  $L^2(dP)$  beschränkte Lösungen  $X$  und  $Y$  der SDE (6.5) auf  $[0, T]$  existieren und betrachten deren Differenz

$$X_t - Y_t = \int_0^t (\mu(s, X_s) - \mu(s, Y_s)) ds + \int_0^t (\sigma(s, X_s) - \sigma(s, Y_s)) dB_s.$$

Mithilfe der Ungleichung  $(a + b)^2 \leq 2a^2 + 2b^2$  erhalten wir

$$\mathbb{E} [(X_t - Y_t)^2] \leq 2\mathbb{E} \left[ \left( \int_0^t \mu(s, X_s) - \mu(s, Y_s) ds \right)^2 \right] + 2\mathbb{E} \left[ \left( \int_0^t \sigma(s, X_s) - \sigma(s, Y_s) dB_s \right)^2 \right]. \quad (6.8)$$

Aufgrund der Lipschitz-Bedingung an  $\sigma(t, x)$  gilt

$$(\sigma(s, X_s) - \sigma(s, Y_s))^2 \leq K(X_s - Y_s)^2,$$

woraus gemeinsam mit der gleichmäßigen  $L^2(dP)$ -Beschränktheit von  $X$  und  $Y$  folgt dass  $\sigma(s, X_s) - \sigma(s, Y_s) \in \mathcal{H}^2[0, T]$ . Wir können also auf das zweite Integral in (6.8) die Ito-Isometrie anwenden. Das erste Integral schätzen wir mit der Cauchy-Schwarzschen Ungleichung ab und erhalten unter erneuter Verwendung der Lipschitz-Bedingung

$$\begin{aligned} \mathbb{E} [(X_t - Y_t)^2] &\leq 2t\mathbb{E} \left[ \int_0^t (\mu(s, X_s) - \mu(s, Y_s))^2 ds \right] + 2\mathbb{E} \left[ \int_0^t (\sigma(s, X_s) - \sigma(s, Y_s))^2 ds \right] \leq \\ &\leq C \int_0^t \mathbb{E} [(X_s - Y_s)^2] ds < \infty, \quad \forall t \in [0, T] \end{aligned}$$

mit  $C = 2K \max(1, T)$ . Wir sehen also dass  $g(t) = \mathbb{E} [(X_t - Y_t)^2]$  die Integralungleichung<sup>9</sup>

$$0 \leq g(t) \leq C \int_0^t g(s) ds$$

erfüllt. Setze nun  $M = \sup_{t \in [0, T]} g(t)$ . Aus der Ungleichung folgt  $g(t) \leq MCt$ . Erneutes Anwenden liefert  $g(t) \leq MC^2 \frac{t^2}{2}$  und nach  $n$  Iterationen erhalten wir  $g(t) \leq MC^n t^n / n!$ . Da  $n!$  stärker wächst als  $t^n$  folgt mit  $n \rightarrow \infty$ , dass  $g(t) = 0$  für alle  $t \in [0, T]$ . Damit gilt  $X_t = Y_t$  fast sicher für alle  $t \in [0, T]$ . Da  $X$  und  $Y$  stetige Prozesse sind, folgt mit Lemma 1.1 die Ununterscheidbarkeit von  $X$  und  $Y$ .  $\square$

Für den Existenzbeweis können wir eine Idee aus der Theorie der gewöhnlichen Differentialgleichungen übernehmen – die **Picard-Iteration**. Wir setzen  $X_t^{(0)} \equiv x_0$  und definieren eine Folge  $(X^{(n)})_{n \in \mathbb{N}}$  von stochastischen Prozessen mittels

$$X_t^{(n+1)} = x_0 + \int_0^t \mu(s, X_s^{(n)}) ds + \int_0^t \sigma(s, X_s^{(n)}) dB_s. \quad (6.9)$$

Zuerst ist zu zeigen, dass die Prozesse  $X^{(n)}$  tatsächlich wohldefiniert sind.

**Lemma 6.2.** *Wenn  $X_t^{(n)}$  auf  $[0, T]$  in  $L^2(dP)$  beschränkt ist, dann gilt*

$$\sigma(t, X_t^{(n)}) \in \mathcal{H}^2[0, T] \quad \text{und} \quad \mu(t, X_t^{(n)}) \in L^2(dt \times dP).$$

*Zudem ist auch der durch (6.9) definierte Prozess  $X_t^{(n+1)}$  auf  $[0, T]$  in  $L_2(dP)$  beschränkt.*

*Beweis.* Aufgrund der  $L^2$ -Beschränktheit von  $X^{(n)}$  existiert ein  $C > 0$ , sodass  $\sup_{t \in [0, T]} \mathbb{E} \left[ \left( X_t^{(n)} \right)^2 \right] = C < \infty$ . Aus der Wachstumsbedingung (6.7) folgt

$$\mathbb{E} \left[ \int_0^T \left( \sigma(t, X_t^{(n)}) \right)^2 dt \right] \leq TK(1 + C)$$

und damit  $\sigma(t, X_t^{(n)}) \in \mathcal{H}^2[0, T]$ . Aus der Ito-Isometrie erhalten wir also

$$\mathbb{E} \left[ \left( \int_0^T \sigma(t, X_t^{(n)}) dB_t \right)^2 \right] = \mathbb{E} \left[ \int_0^T \left( \sigma(t, X_t^{(n)}) \right)^2 dt \right] \leq TK(1 + C).$$

Ebenso folgt aus der Wachstumsbedingung (6.7)  $\mu(t, x)^2 \leq K(1 + x^2)$  und damit nach

<sup>9</sup>Diese Integralungleichung ist eine besonders einfache Form der Gronwall-Ungleichung die auch in der Theorie der gewöhnlichen DGLen eine Rolle spielt.

Anwendung der Cauchy-Schwarzschen Ungleichung

$$\mathbb{E} \left[ \left( \int_0^T \mu(t, X_t^{(n)}) dt \right)^2 \right] \leq TK(1+C).$$

Das Zusammenfügen der Abschätzungen für Drift- und Diffusionsterm zeigt nun die  $L^2$ -Beschränktheit von  $X_t^{(n+1)}$ .  $\square$

Der nächste Schritt besteht darin zu zeigen, dass die Folge der Picard-Iterationen tatsächlich gegen einen stetigen Prozess  $X$  konvergiert. Ähnlich wie in der Konstruktion des Ito-Integrals als stetigen Prozess in Theorem 3.8 benötigen wir dafür pfadweise Konvergenz in der Supremumsnorm auf dem Raum  $C[0, T]$  der stetigen Funktionen. Schlüssel dazu ist folgende Abschätzung:

**Lemma 6.3.** *Wenn die Koeffizienten  $\mu$  und  $\sigma$  in (6.9) die Lipschitz-Bedingung (6.6) erfüllen, dann existiert ein  $C > 0$  sodass für die Prozesse  $X^{(n)}$  der Picard-Iteration (6.9) die Ungleichung*

$$\mathbb{E} \left[ \sup_{0 \leq s \leq t} \left( X_s^{(n+1)} - X_s^{(n)} \right)^2 \right] \leq C \int_0^t \mathbb{E} \left[ \left( X_s^{(n)} - X_s^{(n-1)} \right)^2 \right] ds \quad (6.10)$$

gilt.

*Beweis.* Wir zerlegen  $X_s^{(n+1)} - X_s^{(n)}$  in den Beitrag der Driftterme

$$D_t = \int_0^t \left( \mu(s, X_s^{(n)}) - \mu(s, X_s^{(n-1)}) \right) ds$$

und den Beitrag der Diffusionsterme

$$M_t = \int_0^t \left( \sigma(s, X_s^{(n)}) - \sigma(s, X_s^{(n-1)}) \right) dB_s.$$

Wir haben bereits gezeigt, dass  $\sigma(s, X_s^{(n)})$  in  $\mathcal{H}^2[0, T]$  liegt und folgern daraus, dass  $M_t$  ein Martingal ist. Es gilt

$$\sup_{0 \leq s \leq t} \left( X_s^{(n+1)} - X_s^{(n)} \right)^2 \leq 2 \sup_{0 \leq s \leq t} D_s^2 + 2 \sup_{0 \leq s \leq t} M_s^2.$$

Wir schätzen zuerst den Driftterm mit der Cauchy-Schwarzschen Ungleichung ab, und erhalten

$$\sup_{0 \leq s \leq t} D_s^2 \leq t \int_0^t \left( \mu(s, X_s^{(n)}) - \mu(s, X_s^{(n-1)}) \right)^2 ds. \quad (6.11)$$

Auf den Martingalterm wenden wir die Doobsche Ungleichung und die Ito-Isometrie an und erhalten

$$\mathbb{E} \left[ \sup_{0 \leq s \leq t} M_s^2 \right] \leq 4\mathbb{E} [M_t^2] = 4\mathbb{E} \left[ \int_0^t \left( \sigma(s, X_s^{(n)}) - \sigma(s, X_s^{(n-1)}) \right)^2 ds \right]. \quad (6.12)$$

Die Schranken (6.11) und (6.12) gemeinsam mit der Lipschitz-Bedingung (6.6) und der Wahl der Konstanten  $C := 8K \max(1, T)$  liefern die Behauptung.  $\square$

Es bleibt nur noch die Konvergenz der Picard-Iteration gegen eine Lösung der stochastischen Differentialgleichung zu zeigen.

*Beweis von Theorem 6.1: Existenz.* Wir definieren

$$g_n(t) = \mathbb{E} \left[ \sup_{0 \leq s \leq t} \left| X_s^{(n+1)} - X_s^{(n)} \right|^2 \right]. \quad (6.13)$$

Mit der Abschätzung aus Lemma 6.3 gilt

$$g_n(t) \leq C \int_0^t g_{n-1}(s) ds \quad \forall n \geq 1. \quad (6.14)$$

Aus den Bedingungen an  $\mu$  und  $\sigma$  folgt zunächst, dass eine Konstante  $M$  existiert, sodass  $g_0(t) \leq M$  für  $t \in [0, T]$ . Mit der Ungleichung (6.14) folgt  $g_1(t) \leq MCt$  und nach  $n$ -maliger Anwendung schließlich

$$0 \leq g_n(t) \leq MC^n t^n / n!.$$

Die Markowsche Ungleichung liefert daraus

$$\mathbb{P} \left[ \sup_{0 \leq s \leq t} \left| X_s^{(n+1)} - X_s^{(n)} \right|^2 \geq 2^{-n} \right] \leq g_n(t) 2^n \leq M \frac{(2CT)^n}{n!}.$$

Die rechte Seite ist über  $n \in \mathbb{N}$  summierbar und aus dem Lemma von Borel-Cantelli folgt die Existenz einer Menge  $A \in \mathcal{F}$  mit  $\mathbb{P}[A] = 1$ , sodass die Folge der Funktionen  $t \mapsto X_t^{(n)}(\omega)$  für jedes  $\omega \in A$  eine Cauchy-Folge in  $C[0, T]$  bezüglich der Supremumsnorm bildet. Wir schließen dass eine stetige Funktion  $t \mapsto X_t(\omega)$  existiert mit  $\sup_{t \in [0, T]} \left| X_t - X_t^{(n)} \right| \rightarrow 0$  fast sicher.

Es bleibt zu zeigen, dass der konstruierte Prozess  $X$  in  $L^2(dP)$  beschränkt und tatsächlich eine Lösung der SDE 6.5 ist. Aus (6.13) erhalten wir

$$\left\| X_t^{(n+1)} - X_t^{(n)} \right\|_{L^2(dP)}^2 \leq T g_n(T), \quad \forall t \in [0, T].$$



Aus (6.14) folgt sogar, dass  $\sum_{n=0}^{\infty} g_n(T) < \infty$  und damit, dass die  $X_t^{(n)}$  auch in  $L^2(dP)$  eine Cauchy-Folge bilden. Wir haben also gezeigt dass  $\lim_{n \rightarrow \infty} X_t^{(n)} = X_t$  auch in  $L^2(dP)$  für alle  $t \in [0, T]$  gilt. Aus der Lipschitzbedingung (6.6) erhalten wir

$$\begin{aligned} \mathbb{E} \left[ \int_0^T \left( \sigma(s, X_s^{(n)}) - \sigma(s, X_s^{(n-1)}) \right)^2 ds \right] &\leq K \int_0^T \mathbb{E} \left[ (X_s^{(n)} - X_s^{(n-1)})^2 \right] ds \leq \\ &\leq K \int_0^T g_{n-1}(s) ds. \end{aligned}$$

Da  $g_n(s)$  eine steigende Funktion ist, folgt daraus

$$\left\| \sigma(s, X_s^{(n)}) - \sigma(s, X_s^{(n-1)}) \right\|_{L^2(dP \times dt)}^2 \leq KT g_{n-1}(T).$$

Analog folgt für den Driftterm die Abschätzung

$$\left\| \mu(s, X_s^{(n)}) - \mu(s, X_s^{(n-1)}) \right\|_{L^2(dP \times dt)}^2 \leq KT g_{n-1}(T).$$

Aus der ersten Abschätzung schließen wir unter Verwendung der Ito-Isometrie die Konvergenz

$$\int_0^t \sigma(s, X_s^{(n)}) dB_s \rightarrow \int_0^t \sigma(s, X_s) dB_s, \quad \text{in } L^2(dP) \quad (6.15)$$

und aus der zweiten Abschätzung die Konvergenz

$$\int_0^t \mu(s, X_s^{(n)}) ds \rightarrow \int_0^t \mu(s, X_s) ds, \quad \text{in } L^2(dP). \quad (6.16)$$

Wir betrachten nun die Picard-Iteration (6.9) für ein fixes  $t \in [0, T]$ . Auf der linken Seite gilt  $X_t^{(n)} \rightarrow X_t$  fast sicher. Die gezeigte  $L^2$ -Konvergenz in (6.15) und (6.16) lässt sich durch Auswahl einer Teilfolge  $(n_k)_{k \in \mathbb{N}}$  zu fast sicherer Konvergenz verstärken. Mit  $n_k \rightarrow \infty$  erhalten wir also

$$X_t = x_0 + \int_0^t \mu(s, X_s) ds + \int_0^t \sigma(s, X_s) dB_s, \quad \text{fast sicher.}$$

Da beide Seiten der Gleichung stetig in  $t$  sind erhalten wir mit Lemma 1.1 sogar Gleichheit bis auf Ununterscheidbarkeit und haben gezeigt, dass der konstruierte Prozess  $X$  tatsächlich die SDE (6.5) löst.  $\square$

## 6.3 Numerische Lösungsmethoden

### Euler-Schema

Bei der numerischen Lösung von stochastischen Differentialgleichungen sucht man auf einem diskreten Gitter  $0 = t_0 < t_1 < \dots < t_N = T$  nach einer möglichst guten Näherung  $\hat{X}(t_i)$  für die tatsächliche Lösung  $X(t_i)$  der SDE. Auch hier können wir Anleihen bei den Methoden für gewöhnliche Differentialgleichungen nehmen. Für die einfachste Methode, das **Euler-Schema** ersetzen wir in der SDE einfach die infinitesimalen Terme  $dt$  und  $dB_t$  durch die finiten Differenzen  $t_{i+1} - t_i$  bzw.  $B_{t_{i+1}} - B_{t_i}$ . Zusätzlich können wir die Skalierungseigenschaft und die Unabhängigkeit der Inkremente der Brownschen Bewegung ausnutzen und erhalten folgende Rekursion

$$\hat{X}(t_{i+1}) = \hat{X}(t_i) + a(t_i, \hat{X}(t_i)) [t_{i+1} - t_i] + b(t_i, \hat{X}(t_i)) \sqrt{t_{i+1} - t_i} Z_i, \quad (6.17)$$

mit Startwert  $\hat{X}(0) = x_0$ , wobei  $(Z_i)_{i \in \mathbb{N}}$  unabhängige standardnormalverteilte Zufallsgrößen sind. Da praktisch jede Numeriksoftware die Erzeugung von normalverteilten Zufallszahlen ermöglicht, ist das Euler-Schema (6.17) sehr leicht zu implementieren. Wählt man die Gitterweite  $h = t_{i+1} - t_i$  konstant, so lässt sich das Euler-Schema noch kompakter als

$$\hat{X}((i+1)h) = \hat{X}(ih) + a(ih, \hat{X}(ih)) h + b(ih, \hat{X}(ih)) \sqrt{h} Z_i,$$

schreiben.

### Milstein-Schema

Wie bei den Methoden für gewöhnliche Differentialgleichungen kann man versuchen das numerische Näherungsverfahren durch Hinzufügen von Termen höherer Ordnung zu verbessern. Bei genauerer Betrachtung des Euler-Schemas fällt auf dass die Approximation des Driftterms von Größenordnung  $O(h)$  ist, die Approximation des Diffusionsterms hingegen von Größenordnung  $O(\sqrt{h})$  (in Wahrscheinlichkeit). Es liegt also Nahe, sich bei der Verbesserung des Verfahrens zunächst einmal auf die Approximation des Diffusionsterms zu konzentrieren. In Integralschreibweise geschrieben verwendet das Euler-Schema für diesen Term die Näherung

$$\int_t^{t+h} b(s, X_s) dB_s \approx b(t, X_t) \cdot [B_{t+h} - B_t]. \quad (6.18)$$

Zur Vereinfachung betrachten wir nun den Fall, dass  $a(s, X_s)$  und  $b(s, X_s)$  nicht explizit vom Zeitparameter  $t$  abhängen und von der Form  $a(X_s)$  und  $b(X_s)$  sind. Wenn wir

auf  $b(X_s)$  die Ito-Formel anwenden erhalten wir

$$b(X_s) = b(X_t) + \int_t^s \left( b'(X_u) + \frac{1}{2} b''(X_u) b^2(X_u) \right) du + \int_t^s b'(X_u) b(X_u) dB_u.$$

Wegen seiner höheren Größenordnung vernachlässigen wir den  $dt$ -term und erhalten die Approximation

$$b(X_s) \approx b(X_t) + b'(X_t) b(X_t) (B_s - B_t).$$

Eingesetzt in die linke Seite von (6.18) ergibt sich

$$\int_t^{t+h} b(X_s) dB_s \approx b(X_t) (B_{t+h} - B_t) + b'(X_t) b(X_t) \int_t^{t+h} (B_s - B_t) dB_s.$$

Wir berechnen

$$\begin{aligned} \int_t^{t+h} (B_s - B_t) dB_s &= \int_t^{t+h} B_s dB_s - B_t (B_{t+h} - B_t) = \\ &= \frac{1}{2} (B_{t+h}^2 - (t+h)) - \frac{1}{2} (B_t^2 - t) - B_t (B_{t+h} - B_t) = \\ &= \frac{1}{2} (B_{t+h} - B_t)^2 - \frac{h}{2}, \end{aligned}$$

und wir erhalten schließlich das **Milstein-Schema**

$$\begin{aligned} \hat{X}((i+1)h) &= \hat{X}(ih) + a(\hat{X}(ih)) h + b(\hat{X}(ih)) \sqrt{h} Z_i + \\ &\quad + \frac{1}{2} b'(\hat{X}(ih)) b(\hat{X}(ih)) h (Z_i^2 - 1) \end{aligned}$$

wobei  $(Z_i)_{i \in \mathbb{N}}$  wie zuvor unabhängige standardnormalverteilte Zufallsgrößen sind. Man beachte, dass die beiden stochastischen Terme im Milstein-Schema unkorreliert sind, denn  $\mathbb{E}[Z_i(Z_i^2 - 1)] = 0$ .

### Fehlerordnung

Natürlich ist bei den vorgestellten Verfahren die Konvergenz zur tatsächlichen Lösung  $X$  für  $h \rightarrow 0$  zu zeigen und auch die Größe des Diskretisierungsfehlers ist von Interesse. Man nennt ein Diskretisierungsverfahren von **starker Fehlerordnung**  $\delta > 0$ , wenn

eine Konstante  $C$  existiert, sodass<sup>10</sup>

$$\mathbb{E} \left[ \left| X(T) - \hat{X}(Nh) \right| \right] \leq Ch^\delta$$

für alle ausreichend kleinen  $h$ . Für eine festgelegte Funktionenklasse  $\mathcal{G}$  – meistens die Klasse der stetig differenzierbaren Funktionen mit polynomiell beschränkten Ableitungen – heißt ein Diskretisierungsverfahren von **schwacher Fehlerordnung**  $\delta > 0$ , wenn für jedes  $f \in \mathcal{G}$  eine Konstante  $C_f$  existiert, sodass

$$\left| \mathbb{E} [f(X(T))] - \mathbb{E} [f(\hat{X}(Nh))] \right| \leq C_f h^\delta,$$

für alle ausreichend kleinen  $h$  gilt. Wenn die Klasse  $\mathcal{G}$  nur Lipschitzstetige Funktionen enthält gilt

$$\left| \mathbb{E} [f(X(T))] - \mathbb{E} [f(\hat{X}(Nh))] \right| \leq K_f \mathbb{E} \left[ \left| X(T) - \hat{X}(Nh) \right| \right]$$

für passendes  $K_f$  und die schwache Fehlerordnung ist stets größer als die starke Fehlerordnung. Die Motivation den Begriff der schwachen Fehlerordnung einzuführen ist, dass in vielen Anwendungen, etwa in der Finanzmathematik, die Auswertung von Erwartungswerten der Form  $\mathbb{E} [f(X(T))]$  das letztendliche Ziel der numerischen Berechnung ist. Bei manchen Verfahren ist unter bestimmten Voraussetzungen die schwache Fehlerordnung höher als die starke und sie sind daher numerisch gutartiger als die starke Fehlerordnung suggeriert. Es gilt folgendes Resultat:

**Theorem 6.4.** *Gegeben sei eine stochastische Differentialgleichung*

$$dX_t = a(X_t)dt + b(X_t)dB_t, \quad X_0 = x_0$$

*deren Koeffizienten die Voraussetzungen von Theorem 6.1 erfüllen. Dann ist das Euler-Schema von starker Konvergenzordnung  $\frac{1}{2}$ . Wenn zusätzlich  $a \in C^1(\mathbb{R})$  und  $b \in C^2(\mathbb{R})$  gilt, so ist das Milstein-Schema von starker Konvergenzordnung 1.*

Ein Beweis findet sich in Kloeden and Platen [1995, Theorem 10.2.2, Theorem 10.6.3]. Für zeitabhängige Koeffizienten  $a(t, X_t)$  und  $b(t, X_t)$  gilt das Resultat unter milden zusätzlichen Voraussetzungen ebenfalls. Wie oben bemerkt, ist die schwache Konvergenzordnung der beschriebenen Verfahren unter Umständen höher. Falls beispielsweise  $a$  und  $b$  viermal stetig differenzierbar mit polynomiell beschränkten Ableitungen sind, so ist die schwache Konvergenzordnung des Euler-Verfahrens 1 (siehe Glasserman [2004, Kap. 6.1.2]).

<sup>10</sup>In der Fachliteratur werden auch leicht abweichende Formen der Fehlerordnung verwendet. Wir orientieren uns an der Terminologie von Kloeden and Platen [1995].

## 7 Der Satz von Girsanov

Der Satz von Girsanov wird uns eine neue Perspektive auf die Rolle des Drifts liefern. Die Prozesse

- Brownsche Bewegung  $B_t$ ,
- Brownsche Bewegung mit Drift  $X_t = B_t + \mu t$

haben wir bis jetzt einfach als zwei unterschiedliche stochastische Prozesse auf dem Wahrscheinlichkeitsraum  $(\Omega, \mathcal{F}, \mathbb{P})$  betrachtet. Der Satz von Girsanov wird uns auch folgenden Blickwinkel ermöglichen:  $B$  und  $X$  können als *derselbe* Prozess auf dem messbaren Raum  $(\Omega, \mathcal{F})$  aufgefasst werden, aber unter unterschiedlichen Maßen  $\mathbb{P}$  und  $\mathbb{Q}$ . Dieser Perspektivenwechsel ist vor allem in der Finanzmathematik von großer Wichtigkeit – dort müssen wir nämlich stets zwischen dem statistischen Maß  $\mathbb{P}$  und dem risikoneutralen Maß  $\mathbb{Q}$  unterscheiden. Der Satz von Girsanov sagt uns dann wie sich die auf dem Wahrscheinlichkeitsraum  $(\Omega, \mathcal{F}, \mathbb{P})$  definierten stochastischen Prozesse unter dem Maßwechsel zu  $\mathbb{Q}$  verhalten.

### 7.1 Importance Sampling und Exponentiells Kippen

Der Satz von Girsanov soll durch die Technik des *Importance Sampling* motiviert werden, welche in der numerischen Berechnung von Erwartungswerten verwendet wird. Es sei  $X$  eine standardnormalverteilte Zufallsgröße und  $(Y_i)_{i \in \mathbb{N}}$  unabhängige Kopien von  $X$ . Wir sind daran interessiert den Erwartungswert  $\mathbb{E}[f(X)]$  mit  $f$  messbar und beschränkt zu berechnen. Bei dem Verfahren der Monte-Carlo-Simulation erzeugen wir  $N$  unabhängige normalverteilte Zufallsvariablen  $(Y_i)_{i=1, \dots, N}$  und Approximieren

$$\mathbb{E}[f(X)] \approx \frac{1}{N} \sum_{i=1}^N f(Y_i). \quad (7.1)$$

Das starke Gesetz der großen Zahlen garantiert dass die Approximation der rechten Seite für  $N \rightarrow \infty$  tatsächlich gegen  $\mathbb{E}[f(X)]$  konvergiert. Oft liefert das Verfahren auch schon für moderate Werte von  $N$  eine brauchbare Approximation. Was passiert aber, wenn wir uns nun für die Funktion  $f(x) = \mathbf{1}\{x \geq 30\}$  interessieren? Der Erwartungswert  $\mathbb{E}[f(X)]$  wird sehr nahe an 0 liegen und wir wollen zumindest die ersten 3 signifikanten Stellen<sup>11</sup> berechnen. Die Anzahl der Simulationen  $N$  bis auch nur ein einziger Summand in (7.1) ungleich Null ist, ist geometrisch verteilt mit Parameter  $p = \mathbb{P}[X \geq 30]$ . Das bedeutet, dass die erwartete Anzahl von Simulationen bis zum

<sup>11</sup>Signifikante Stellen sind jene Nachkommastellen die ungleich Null sind.

ersten Summanden ungleich Null gegeben ist durch

$$\mathbb{E}[N] = 1/\mathbb{P}[X \geq 30] \approx 1.5 \times 10^{196}.$$

In der Praxis ist es also unmöglich, auch nur eine signifikante Stelle von  $\mathbb{E}[f(X)]$  mit der naiven Monte-Carlo-Simulation (7.1) zu berechnen. Eine Lösung für dieses Problem bietet das *Importance Sampling*. Wir schreiben  $\mathbb{E}_\mu$  für den Erwartungswert unter der Modellannahme  $X \sim \mathcal{N}(\mu, 1)$ . Durch Ergänzen auf das vollständige Quadrat im Exponenten erhalten wir

$$\begin{aligned} \mathbb{E}_0[f(X)] &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x)e^{-x^2/2} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x)e^{-(x-\mu)^2/2} e^{-\mu x + \mu^2/2} dx = \\ &= \mathbb{E}_\mu \left[ f(X)e^{-\mu X + \mu^2/2} \right]. \end{aligned}$$

Diese Identität wird auch als **exponentielles Kippen** von  $X$  bezeichnet. Durch Multiplikation mit dem Exponential  $\exp(-\mu X + \mu^2/2)$  verschieben wir den Mittelwert der Zufallsvariable  $X$ . Übersetzt man die Identität zurück in ein numerisches Verfahren erhält man das **Importance Sampling**

$$\mathbb{E}_0[f(X)] \approx \frac{1}{N} \sum_{i=1}^N g(\tilde{Y}_i) \quad \text{mit} \quad g(x) = f(x)e^{-\mu x + \mu^2/2}, \mu \in \mathbb{R} \quad (7.2)$$

und  $\tilde{Y}_i$  unabhängig normalverteilt mit Mittelwert  $\mu$  und Varianz 1. Der Parameter  $\mu$  kann nun der Funktion  $f$  angepasst werden; für  $f(x) = \mathbf{1}\{x \geq 30\}$  können wir z.B.  $\mu = 30$  wählen und erhalten deutlich bessere Konvergenz als mit  $\mu = 0$ .

Der nächste Schritt ist es das exponentielle Kippen auf stochastische Prozesse auszuweiten. Sei also  $X_t = B_t + \mu t$  Brownsche Bewegung mit Drift und  $\mathbb{E}[f(X_{t_1}, \dots, X_{t_n})]$  ein Erwartungswert der von der endlichdimensionalen Randverteilung von  $X$  abhängt. Es gilt folgendes Resultat.

**Proposition 7.1.** *Sei  $X_t = B_t + \mu t$ ,  $0 \leq t_1 \leq \dots \leq t_n \leq T$  und  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  beschränkt und messbar. Dann gilt*

$$\mathbb{E}[f(X_{t_1}, \dots, X_{t_n})] = \mathbb{E}[f(B_{t_1}, \dots, B_{t_n})M_T]$$

mit

$$M_T = \exp\left(\mu B_T - \frac{\mu^2}{2}T\right).$$

*Beweis.* Die Dichtefunktion  $\phi$  des Vektors  $(X_{t_1}, \dots, X_{t_n})$  können wir unter Verwen-

dung der Unabhängigkeit der Inkremente von  $X$  folgendermaßen schreiben:

$$\phi(x_1, \dots, x_n) = C \prod_{i=1}^n \exp \left( -\frac{[(x_i - x_{i-1}) - \mu(t_i - t_{i-1})]^2}{2(t_i - t_{i-1})} \right),$$

wobei wir  $x_0 = 0$  setzten und  $C = (2\pi)^{-n/2} \prod_{i=1}^n (t_i - t_{i-1})^{-1/2}$ . Durch Umformen des Exponenten ergibt sich

$$\begin{aligned} \phi(x_1, \dots, x_n) &= C \prod_{i=1}^n \exp \left( -\frac{(x_i - x_{i-1})^2}{2(t_i - t_{i-1})} \right) \prod_{i=1}^n \exp \left( \mu(x_i - x_{i-1}) - \frac{1}{2}\mu^2(t_i - t_{i-1}) \right) = \\ &= C \prod_{i=1}^n \exp \left( -\frac{(x_i - x_{i-1})^2}{2(t_i - t_{i-1})} \right) \exp \left( \mu x_n - \frac{1}{2}\mu^2 t_n \right). \end{aligned}$$

Der erste Faktor ist nun genau die Dichte des Vektors  $(B_{t_1}, \dots, B_{t_n})$ , also erhalten wir

$$\mathbb{E}[f(X_{t_1}, \dots, X_{t_n})] = \mathbb{E}[f(B_{t_1}, \dots, B_{t_n})M_{t_n}].$$

Aufgrund der Martingaleigenschaft von  $M$  gilt  $M_{t_n} = \mathbb{E}[M_T | \mathcal{F}_{t_n}]$  und die Behauptung folgt nach Anwendung der Turmregel.  $\square$

## 7.2 Maßwechsel auf dem Pfadraum $C[0, T]$

Die obige Proposition lässt sich auf Funktionale  $\mathbb{E}[f(X_{t_1}, \dots, X_{t_n})]$  anwenden, die von der Verteilung von  $X$  zu *endlich vielen* Zeitpunkten  $t_1, \dots, t_n$  abhängen. Was ist aber mit pfadabhängigen Funktionalen wie z.B.  $\mathbb{E}\left[f(\sup_{t \in [0, T]} X_t)\right]$ , die sich nicht auf Beobachtungen zu endlich vielen Zeitpunkten reduzieren lassen? Hier gibt es eine elegante Verallgemeinerung, die Proposition 7.1 auf den **Pfadraum** von  $X$  überträgt. Der hier betrachtete Prozess  $X$  ist ein stetiger stochastischer Prozess. Auf einem festen Intervall  $[0, T]$  sind seine Pfade  $t \mapsto X_t(\omega)$  mit Wahrscheinlichkeit eins Elemente von  $C[0, T]$ , dem Raum der stetigen Funktionen auf  $[0, T]$ . Wir können also  $C[0, T]$  als den Pfadraum von  $X$  bezeichnen. Mit der Supremumsnorm  $\|\cdot\|_\infty$  versehen ist  $(C[0, T], \|\cdot\|_\infty)$  ein vollständiger, normierter Raum, also ein Banachraum.<sup>12</sup> Die offenen Mengen in diesem normierten Raum erzeugen die Borel- $\sigma$ -Algebra  $\mathcal{B}$  auf  $(C[0, T], \|\cdot\|_\infty)$ , sodass wir es mit einem messbaren Raum zu tun haben. Zu einem Wahrscheinlichkeitsraum fehlt also nur mehr ein Wahrscheinlichkeitsmaß welches wir folgendermaßen definieren:

**Definition.** Sei  $(X_t)_{t \in [0, T]}$  ein stetiger stochastischer Prozess auf einem Wahrscheinlichkeitsraum  $(\Omega, \mathcal{F}, \mathbb{P})$ . Dann lässt sich ein Maß  $\mathbb{P}_X$  auf dem Pfadraum  $(C[0, T], \mathcal{B})$

<sup>12</sup> $(C[0, T], \|\cdot\|_\infty)$  enthält zudem eine abzählbare dichte Teilmenge und ist somit ein polnischer Raum, was für einige maßtheoretische Überlegungen von Bedeutung ist.

mittels

$$\mathbb{P}_X(A) := \mathbb{P}(X^{-1}(A)), \quad A \in \mathcal{B}$$

definieren. Das Maß  $\mathbb{P}_X$  heisst von  $X$  induziertes Maß und  $(C[0, T], \mathcal{B}, \mathbb{P}_X)$  bildet einen Wahrscheinlichkeitsraum.

In dieser Definition steht  $X^{-1}(A)$  für das Urbild der Menge  $A$  unter der messbaren Abbildung  $X : \Omega \rightarrow C[0, T], \omega \mapsto (t \mapsto X_t(\omega))$ . Von besonderem Interesse ist der Fall, dass  $X$  die Brownsche Bewegung ist, d.h.  $X = B$ . Dann heisst  $\mathbb{P}_B$  **Wiener-Maß** und  $(C[0, T], \mathcal{B}, \mathbb{P}_B)$  **Wiener-Raum**. Wir können nun das ‘unendlichdimensionale’ Analogon von Proposition 7.1 formulieren.

**Theorem 7.2.** *Sei  $\mathbb{P}_\mu$  das von  $X_t = B_t + \mu t$  induzierte Maß auf dem Pfadraum  $(C[0, T], \mathcal{B})$ . Dann gilt für jede beschränkte, Borel-messbare Funktion  $W : C[0, T] \rightarrow \mathbb{R}$*

$$\mathbb{E}_\mu [W] = \mathbb{E}_0 [W M_T] \tag{7.3}$$

mit

$$M_T : C[0, T] \rightarrow \mathbb{R}, \quad \eta \mapsto \exp\left(\mu\eta_T - \frac{\mu^2}{2}T\right).$$

Für den Beweis kombinieren wir Proposition 7.1 mit einem Resultat der Maßtheorie, dem Hauptsatz über Dynkin-Systeme.

**Definition.** Ein System  $\mathcal{C}$  von Teilmengen von  $\Omega$  heisst **Dynkin-System** wenn folgendes gilt:

- (a)  $\Omega \in \mathcal{C}$ ;
- (b) Aus  $A \in \mathcal{C}$  folgt  $A^c \in \mathcal{C}$ ;
- (c) Wenn  $(A_n)_{n \in \mathbb{N}}$  eine Folge von disjunkten Mengen in  $\mathcal{C}$  ist, so ist auch  $\bigcup_{i=1}^{\infty} A_n$  in  $\mathcal{C}$ .

**Theorem 7.3** (Hauptsatz über Dynkin-Systeme). *Sei  $\mathcal{C}$  ein Dynkin-System und  $\mathcal{I}$  eine durchschnittsstabile Teilmenge  $\mathcal{I} \subset \mathcal{C}$ . Dann ist auch die von  $\mathcal{I}$  erzeugte  $\sigma$ -Algebra in  $\mathcal{C}$  enthalten, d.h.  $\sigma(\mathcal{I}) \subset \mathcal{C}$ .*

*Beweis von Theorem 7.2.* Jede Borel-messbare Funktion  $W$  lässt sich als Limes von elementaren Funktionen der Form  $\sum_{i=1}^n \theta_i \mathbf{1}_{A_i}$  mit  $A_i \in \mathcal{B}$  darstellen. Es reicht also, die Gleichung (7.3) für Funktionen der Gestalt  $W = \mathbf{1}_A$  mit  $A \in \mathcal{B}$  zu zeigen. Sei  $\mathcal{C}$  das System aller Borel-Mengen  $A$ , deren Indikatorfunktion  $W = \mathbf{1}_A$  die Gleichung (7.3) erfüllt. Wir wollen zeigen, dass  $\mathcal{C}$  ein Dynkin-System ist. Für  $A = \Omega$  ist  $W = 1$ . Die Gleichung (7.3) ist erfüllt, denn  $\mathbb{E}_0 [M_T] = 1$ , also ist  $\Omega \in \mathcal{C}$ . Sei  $A \in \mathcal{C}$ ; dann ist  $\mathbf{1}_{A^c} = 1 - \mathbf{1}_A$  und (7.3) folgt aus der Linearität des Erwartungswerts. Sei schliesslich



$(A_n)_{n \in \mathbb{N}}$  eine Folge von disjunkten Mengen in  $\mathcal{C}$  und  $A_\infty = \bigcup_{i=1}^{\infty} A_n$  ihre Vereinigung. Dann gilt  $\mathbf{1}_{A_\infty} = \sum_{n=1}^{\infty} \mathbf{1}_{A_n}$  und (7.3) folgt mit monotoner Konvergenz. Wir haben also gezeigt, dass  $\mathcal{C}$  ein Dynkin-System ist. Betrachte nun das System  $\mathcal{I}$  aller Mengen  $A \in \mathcal{B}$  von folgender Gestalt:

$$A = \{\eta \in C[0, T] : \eta(t_i) \in [a_i, b_i] \quad \text{mit} \quad 0 \leq t_1 \cdots t_n \leq T; a_i \leq b_i \in \mathbb{R}\}.$$

Es ist leicht einzusehen, dass  $\mathcal{I}$  durchschnittstabil ist und die Borel- $\sigma$ -Algebra  $\mathcal{B}$  erzeugt. Zudem lässt sich für jede Funktion  $\eta \in C[0, T]$  alleine an endlich vielen Stellen  $t_1, \dots, t_n$  ablesen ob sie in einer Menge  $A \in \mathcal{I}$  liegt oder nicht. Wir folgern also aus Proposition 7.1, dass  $W = \mathbf{1}_A$  für  $A \in \mathcal{I}$  die Gleichung (7.3) erfüllt. Damit gilt  $\mathcal{I} \subset \mathcal{C}$  und es folgt aus dem Hauptsatz über Dynkin-Systeme, dass  $\sigma(\mathcal{I}) \subset \mathcal{C}$ . Es ist aber  $\sigma(\mathcal{I}) = \mathcal{B}$  und wir haben daher wie gewünscht gezeigt, dass für jede Borel-Menge  $A \in \mathcal{B}$  die Funktion  $W = \mathbf{1}_A$  Gleichung (7.3) erfüllt.  $\square$

Es sei daran erinnert, dass zwei Maße  $\mathbb{P}, \mathbb{Q}$  äquivalent heissen (Notation:  $\mathbb{P} \sim \mathbb{Q}$ ) wenn sie dieselben Nullmengen besitzen, d.h  $\mathbb{P}(A) = 0 \Leftrightarrow \mathbb{Q}(A) = 0$  für alle  $A \in \mathcal{B}$ . Hingegen heissen  $\mathbb{P}$  und  $\mathbb{Q}$  singulär (Notation:  $\mathbb{P} \perp \mathbb{Q}$ ), wenn es eine Menge  $A \in \mathcal{B}$  gibt mit  $\mathbb{P}(A) = 0$  und  $\mathbb{Q}(A^c) = 0$ . Es ist klar, dass sich Äquivalenz und Singularität gegenseitig ausschliessen. Einsetzen von  $\mathbf{1}_A$  bzw.  $\mathbf{1}_A \frac{1}{M_T}$  in Gleichung (7.3), liefert dass die induzierten Maße  $\mathbb{P}_0$  und  $\mathbb{P}_\mu$  äquivalent sind. Wie verhält es sich nun, wenn wir neben dem Drift auch noch den Diffusionskoeffizienten variieren, also die von  $X_t = \sigma B_t + \mu t$  induzierten Maße betrachten? Das folgende Resultat beantwortet die Frage.

**Korollar 7.4.** *Sei  $\mathbb{P}_{\mu, \sigma}$  das von  $X_t = \sigma B_t + \mu t$  induzierte Maß auf  $(C[0, T], \mathcal{B})$ . Dann gilt*

- (a)  $\mathbb{P}_{\mu, \sigma} \sim \mathbb{P}_{\tilde{\mu}, \tilde{\sigma}}$  genau dann wenn  $\sigma = \tilde{\sigma}$ ;
- (b)  $\mathbb{P}_{\mu, \sigma} \perp \mathbb{P}_{\tilde{\mu}, \tilde{\sigma}}$  genau dann wenn  $\sigma \neq \tilde{\sigma}$ .

*Beweis.* Die Richtung ‘(a)  $\Leftarrow$ ’ folgt aus Theorem 7.2. (Genau genommen deckt Theorem 7.2 nur den Fall  $\sigma = 1$  ab; der Fall  $\sigma \neq 1$  ist aber komplett analog zu behandeln.) Durch Umkehrung folgt daraus auch die Richtung ‘(b)  $\Rightarrow$ ’. Es reicht also ‘(b)  $\Leftarrow$ ’ zu zeigen, dann folgt durch Umkehrung auch ‘(a)  $\Rightarrow$ ’ und der Beweis ist komplett. Wir verwenden für den Beweis die aus Kapitel 5 bekannten Resultate zur quadratischen Variation  $[X, X]_T$  von  $X$ . Sei  $\mathcal{P}^n$  die Folge der dyadischen Partitionen von  $[0, T]$  und definiere

$$\mathcal{Q}_{\mathcal{P}^n} : C[0, T] \rightarrow \mathbb{R}_{\geq 0}, \quad \eta \mapsto \sum_{i=0}^{2^n-1} \{\eta((i+1)2^{-n}) - \eta(i2^{-n})\}^2.$$

Die Abbildung  $\mathcal{Q}_{\mathcal{P}^n}$  ist stetig und daher Borel-messbar. Wir folgern dass auch  $\limsup_{n \rightarrow \infty} \mathcal{Q}_{\mathcal{P}^n}$  eine Borel-messbare Abbildung ist. Wir setzen

$$A_{\tilde{\sigma}} := \left\{ \eta \in C[0, T] : \limsup_{n \rightarrow \infty} \mathcal{Q}_{\mathcal{P}^n}(\eta) = \tilde{\sigma}T \right\}.$$

Es gilt

$$\mathbb{P}_{\mu, \sigma} [A_{\tilde{\sigma}}] = \mathbb{P} \left[ \limsup_{n \rightarrow \infty} \mathcal{Q}_{\mathcal{P}^n}(X)_T = \tilde{\sigma}T \right] = \mathbb{P} [[X, X]_T = \tilde{\sigma}T] = \begin{cases} 1 & \text{wenn } \tilde{\sigma} = \sigma \\ 0 & \text{wenn } \tilde{\sigma} \neq \sigma. \end{cases}$$

und '(b)  $\Leftarrow$ ' folgt. □

### 7.3 Der Satz von Girsanov

Der Satz von Girsanov verallgemeinert die Resultate aus den vorangehenden Abschnitten auf Ito-Prozesse. Wir formulieren ihn zunächst für einen Ito-Prozess mit konstantem Diffusionskoeffizient  $\sigma(\omega, t) \equiv 1$ . Der anschliessende Korollar liefert dann die Verallgemeinerung.

**Theorem 7.5** (Satz von Girsanov). *Sei  $\mu \in \mathcal{L}_{loc}^2$  und  $X$  ein Ito-Prozess gegeben durch*

$$X_t = B_t + \int_0^t \mu(\omega, s) ds, \quad t \in [0, T]$$

mit  $B$  Brownsche Bewegung auf  $(\Omega, \mathcal{F}, \mathbb{P})$ . Wenn

$$M_t = \exp \left( - \int_0^t \mu(\omega, s) dB_s - \frac{1}{2} \int_0^t \mu^2(\omega, s) ds \right), \quad t \in [0, T] \quad (7.4)$$

ein  $\mathbb{P}$ -Martingal ist, dann definiert  $\mathbb{Q}(A) := \mathbb{E}_{\mathbb{P}} [\mathbf{1}_A M_T]$  für  $A \in \mathcal{F}_T$  ein Wahrscheinlichkeitsmaß  $\mathbb{Q} \sim \mathbb{P}$  auf  $(\Omega, \mathcal{F}_T)$  und  $X$  ist Brownsche Bewegung auf  $(\Omega, \mathcal{F}_T, \mathbb{Q})$ .

Zuerst überzeugen wir uns, dass durch  $\mathbb{Q}(A) := \mathbb{E}_{\mathbb{P}} [\mathbf{1}_A M_T]$  tatsächlich ein Wahrscheinlichkeitsmaß auf  $(\Omega, \mathcal{F}_T)$  definiert wird. Wegen der Martingaleigenschaft von  $M$  gilt  $\mathbb{Q}(\Omega) = \mathbb{E}_{\mathbb{P}} [M_T] = 1$ , außerdem ist  $\mathbb{Q}(A) \geq 0$ , denn  $M$  ist positiv. Es bleibt die  $\sigma$ -Additivität von  $\mathbb{Q}$  nachzuweisen; diese folgt jedoch ohne Schwierigkeiten mit dem Lemma von Fatou aus der Definition von  $\mathbb{Q}$ . Die Zufallsvariable  $M_T$  ist also die Radon-Nikodym-Ableitung auf  $(\Omega, \mathcal{F}_T)$  von  $\mathbb{Q}$  bezüglich  $\mathbb{P}$ , kurz geschrieben

$$\left. \frac{d\mathbb{Q}}{d\mathbb{P}} \right|_{\mathcal{F}_T} = M_T.$$

Durch Approximation von messbaren Funktionen durch elementare Funktionen folgt auch das zwischen den Erwartungswerten unter  $\mathbb{Q}$  und  $\mathbb{P}$  die Relationen

$$\mathbb{E}_{\mathbb{Q}}[X] = \mathbb{E}_{\mathbb{P}}[X M_T] \quad \text{und} \quad \mathbb{E}_{\mathbb{Q}}\left[X \frac{1}{M_T}\right] = \mathbb{E}_{\mathbb{P}}[X]$$

für jedes beschränkte  $X \in \mathcal{F}_T$  gelten.

*Beweis.* Es ist zu beweisen, dass  $(X_t)_{t \in [0, T]}$  eine Brownsche Bewegung auf  $(\Omega, \mathcal{F}_T, \mathbb{Q})$  ist. Klarerweise ist  $X$  ein stetiger stochastischer Prozess ist, also reicht es zu zeigen, dass für beliebige Partitionen  $0 \leq t_1 \leq \dots \leq t_n \leq T$  die Inkremente  $(X_{t_1} - X_{t_2}, \dots, X_{t_{n-1}} - X_{t_n})$  unter  $\mathbb{Q}$  unabhängig sind und normalverteilt mit Mittelwert 0 und Varianz  $t_{i+1} - t_i$ . Wir werden dies mithilfe der charakteristischen Funktion nachweisen. Als vorbereitendes Resultat nehmen wir an, dass  $f : \mathbb{R} \rightarrow \mathbb{R}$  eine deterministische, messbare und beschränkte Funktion ist und zeigen<sup>13</sup>

$$\mathbb{E}_{\mathbb{Q}} \left[ \exp \left( i \int_0^T f(s) dX_s \right) \right] = \exp \left( -\frac{1}{2} \int_0^T f^2(s) ds \right). \quad (7.5)$$

Wir wenden die Definition von  $\mathbb{Q}$  auf die linke Seite an und erhalten nach Ergänzen auf vollständige Quadrate

$$\begin{aligned} & \mathbb{E}_{\mathbb{Q}} \left[ \exp \left( i \int_0^T f(s) dX_s \right) \right] = \\ & = \mathbb{E}_{\mathbb{P}} \left[ \exp \left( i \int_0^T f(s) dX_s \right) \exp \left( -\int_0^T \mu(\omega, s) dB_s - \frac{1}{2} \int_0^T \mu^2(\omega, s) ds \right) \right] = \\ & = \exp \left( -\frac{1}{2} \int_0^T f^2(s) ds \right) \mathbb{E}_{\mathbb{P}} \left[ \exp \left( \int_0^T (i f(s) - \mu(\omega, s)) dB_s - \frac{1}{2} \int_0^T (i f(s) - \mu(\omega, s))^2 ds \right) \right]. \end{aligned}$$

Wir bezeichnen den Term im Inneren des Erwartungswertes mit  $N_T$ . Wenn  $\mathbb{E}_{\mathbb{P}}[N_T] = 1$  gilt, dann folgt die gewünschte Formel (7.5). Wir schreiben  $N_t = \exp(L_t)$  und wenden die Ito-Formel auf den Ito-Prozess  $L_t$  an:

$$\begin{aligned} dN_t &= N_t dL_t + \frac{1}{2} N_t (i f(t) - \mu(\omega, t))^2 dt = \\ &= N_t (i f(t) - \mu(\omega, t)) dB_t - \frac{1}{2} N_t (i f(t) - \mu(\omega, t))^2 dt + \frac{1}{2} N_t (i f(t) - \mu(\omega, t))^2 dt = \\ &= N_t (i f(t) - \mu(\omega, t)) dB_t. \end{aligned}$$

<sup>13</sup>Hier bezeichnet  $i$  die imaginäre Einheit  $i = \sqrt{-1}$ .

Es folgt, dass der Prozess  $N_t$  ein lokales Martingal ist. Wegen  $|\exp(x + iy)| = \exp(x)$  gilt

$$|N_t| = \exp\left(\frac{1}{2} \int_0^T f^2(s) ds\right) M_t$$

mit  $M$  aus (7.4). Laut Voraussetzung ist  $M$  ein Martingal und somit gilt auch  $\mathbb{E}[M_t] = 1$ . Sei  $\tau_n$  eine lokalisierende Folge für das lokale Martingal  $N$ , dann folgt mit dominanter Konvergenz

$$\mathbb{E}[N_T] = \mathbb{E}\left[\lim_{n \rightarrow \infty} N_{T \wedge \tau_n}\right] = \lim_{n \rightarrow \infty} \mathbb{E}[N_{T \wedge \tau_n}] = 1$$

und wir haben die Gleichung (7.5) gezeigt.

Setzen wir nun die Funktion  $f(t) = \sum_{i=0}^{n-1} \theta_i \mathbf{1}_{\{t_i < t \leq t_{i+1}\}}$  in (7.5) ein, so erhalten wir

$$\begin{aligned} \mathbb{E}_{\mathbb{Q}} \left[ \exp\left(i \sum_{k=0}^n \theta_k (X_{t_{k+1}} - X_{t_k})\right) \right] &= \exp\left(-\frac{1}{2} \sum_{k=0}^n \theta_k (t_{k+1} - t_k)\right) = \\ &= \prod_{k=0}^n \exp\left(-\frac{1}{2} \theta_k (t_{k+1} - t_k)\right). \end{aligned}$$

Die linke Seite ist die charakteristische Funktion des Vektors  $(X_{t_1} - X_{t_0}, \dots, X_{t_{n-1}} - X_{t_{n-2}})$  unter  $\mathbb{Q}$ . Die rechte Seite ist das Produkt der charakteristischen Funktionen von normalverteilten Zufallsvariablen mit Erwartungswert 0 und Varianz  $t_{k+1} - t_k$ . Wir folgern dass  $X$  eine Brownsche Bewegung auf  $(\Omega, \mathcal{F}_T, \mathbb{Q})$  ist und die Behauptung ist bewiesen.  $\square$

Mit dem Satz von Girsanov lassen sich viele Resultate zur Brownschen Bewegung in Resultate zur Brownschen Bewegung mit Drift umwandeln. Beispielsweise haben wir am Ende von Kapitel 2 gezeigt, dass die Eintrittszeit der Brownschen Bewegung in das Intervall  $[a, \infty)$  die Dichte

$$f_a(t) = \frac{a}{\sqrt{2\pi t^3}} \exp\left(-\frac{a^2}{2t}\right), \quad (t \geq 0) \quad (7.6)$$

besitzt. Wir setzen nun  $X_t = B_t + \mu t$  sowie

$$\tau_a = \inf\{t > 0 : X_t = a\}$$

und versuchen ein analoges Resultat für die Brownsche Bewegung mit Drift zu erhalten.

Mit dem Satz von Girsanov gilt

$$\begin{aligned}
\mathbb{P}[\tau_a \leq t] &= \mathbb{E}_{\mathbb{P}}[\mathbf{1}\{\tau_a \leq t\}] = \mathbb{E}_{\mathbb{Q}}\left[\mathbf{1}\{\tau_a \leq t\} \frac{1}{M_{\tau_a \wedge t}}\right] = \\
&= \mathbb{E}_{\mathbb{Q}}\left[\mathbf{1}\{\tau_a \leq t\} \exp\left(\mu B_{\tau_a} + \frac{\mu^2}{2} \tau_a\right)\right] = \\
&= \mathbb{E}_{\mathbb{Q}}\left[\mathbf{1}\{\tau_a \leq t\} \exp\left(\mu X_{\tau_a} - \frac{\mu^2}{2} \tau_a\right)\right] = \\
&= \mathbb{E}_{\mathbb{Q}}\left[\mathbf{1}\{\tau_a \leq t\} \exp\left(\mu a - \frac{\mu^2}{2} \tau_a\right)\right]
\end{aligned}$$

Unter dem Maß  $\mathbb{Q}$  ist  $X$  eine Brownsche Bewegung, also hat auch  $\tau_a$  unter  $\mathbb{Q}$  die Dichte  $f_a(t)$  aus (7.6). Damit gilt weiter

$$\begin{aligned}
\mathbb{E}_{\mathbb{Q}}\left[\mathbf{1}\{\tau_a \leq t\} \exp\left(\mu a - \frac{\mu^2}{2} \tau_a\right)\right] &= \int_0^t \frac{a}{\sqrt{2\pi s^3}} \exp\left(-\frac{a^2}{2s} + \mu a - \frac{\mu^2}{2} 2s\right) ds = \\
&= \int_0^t \frac{a}{\sqrt{2\pi s^3}} \exp\left(-\frac{(a - \mu s)^2}{2s}\right) ds,
\end{aligned}$$

und wir haben gezeigt: Die Eintrittszeit der Brownschen Bewegung mit Drift  $B_t + \mu t$  in das Intervall  $[a, \infty)$  besitzt die Dichte

$$f_{a,\mu}(t) = \frac{a}{\sqrt{2\pi t^3}} \exp\left(-\frac{(a - \mu t)^2}{2t}\right), \quad (t \geq 0).$$

Eine einfache Folgerung aus dem Satz von Girsanov ist der folgende Korollar, mit dem sich der Drift  $\mu(\omega, t)$  eines allgemeinen Ito-Prozesses durch Maßwechsel gegen einen neuen Drift  $\nu(\omega, t)$  austauschen lässt.

**Korollar 7.6.** *Sei  $X$  ein Ito-Prozess von der Form*

$$X_t = x_0 + \int_0^t \mu(\omega, s) ds + \int_0^t \sigma(\omega, s) dB_s, \quad t \in [0, T]$$

mit  $B$   $\mathbb{P}$ -Brownsche Bewegung. Außerdem sei  $\theta(\omega, t) = \frac{\mu(\omega, t) - \nu(\omega, t)}{\sigma(\omega, t)}$  in  $\mathcal{L}_{loc}^2[0, T]$  und

$$M_t = \exp\left(-\int_0^t \theta(\omega, s) dB_s - \frac{1}{2} \int_0^t \theta^2(\omega, s) ds\right), \quad t \in [0, T] \quad (7.7)$$

ein  $\mathbb{P}$ -Martingal. Dann definiert  $\mathbb{Q}(A) := \mathbb{E}_{\mathbb{P}}[\mathbf{1}_A M_T]$  ein Wahrscheinlichkeitsmaß auf

$(\Omega, \mathcal{F}_T)$ , der Prozess

$$\tilde{B}_t = B_t + \int_0^t \theta(\omega, s) ds, \quad t \in [0, T]$$

ist eine  $\mathbb{Q}$ -Brownsche Bewegung und  $X$  ist ein Ito-Prozess unter  $\mathbb{Q}$  von der Form

$$X_t = x_0 + \int_0^t \nu(\omega, s) ds + \int_0^t \sigma(\omega, s) d\tilde{B}_s, \quad t \in [0, T].$$

*Beweis.* Aus Theorem 7.5 folgt, dass  $\tilde{B}$  eine  $\mathbb{Q}$ -Brownsche Bewegung ist. Es gilt

$$\begin{aligned} dX_t &= \mu(\omega, t) dt + \sigma(\omega, t) dB_t = \\ &= \mu(\omega, t) dt + \sigma(\omega, t) (d\tilde{B}_t - \theta(\omega, t) dt) = \\ &= \mu(\omega, t) dt + \sigma(\omega, t) d\tilde{B}_t + (\nu(\omega, t) - \mu(\omega, t)) dt = \\ &= \nu(\omega, t) dt + \sigma(\omega, t) d\tilde{B}_t \end{aligned}$$

und das Korollar ist bewiesen.  $\square$

## 7.4 Die Novikov-Bedingung

Die entscheidende Voraussetzung im Satz von Girsanov ist, dass der Prozess  $M$  in (7.4) ein Martingal ist. Wie aber lässt sich diese Voraussetzung überprüfen? Die Novikov-Bedingung liefert ein hinreichendes Kriterium.

**Theorem 7.7** (Novikov-Bedingung). Sei  $\mu \in \mathcal{L}_{loc}^2$  und

$$M(\mu)_t = \exp \left( - \int_0^t \mu(\omega, s) dB_s - \frac{1}{2} \int_0^t \mu^2(\omega, s) ds \right), \quad t \in [0, T]. \quad (7.8)$$

Wenn die Bedingung

$$\mathbb{E} \left[ \exp \left( \frac{1}{2} \int_0^T \mu^2(\omega, s) ds \right) \right] < \infty \quad (7.9)$$

erfüllt ist, dann ist  $(M(\mu)_t)_{0 \leq t \leq T}$  ein Martingal.

Wir zeigen die Aussage des Theorems zunächst unter einer etwas stärkeren Voraussetzung:

**Lemma 7.8** ( $\epsilon$ -Novikov). Sei  $\epsilon > 0$ . Wenn die Bedingung

$$\mathbb{E} \left[ \exp \left( \frac{1}{2} (1 + \epsilon) \int_0^T \mu^2(\omega, s) ds \right) \right] < \infty \quad (7.10)$$

erfüllt ist, dann ist  $(M(\mu)_t)_{0 \leq t \leq T}$  aus Gleichung (7.8) ein Martingal.

*Beweis von Theorem 7.7.* Definiere

$$X_t = - \int_0^t \mu(\omega, s) dB_s, \quad Y_t = \int_0^t \mu(\omega, s)^2 ds. \quad (7.11)$$

Der Prozeß  $M(\mu)$  ist ein lokales Martingal und ein Supermartingal, es gilt also  $\mathbb{E}[M(\mu)_t] \leq 1$ . Aus Proposition 4.7 folgt, dass die Aussage  $\mathbb{E}[M(\mu)_t] \geq 1$  ausreicht um die Martingaleigenschaft von  $M(\mu)$  zu zeigen. Wir setzen

$$N(\mu) = \mathbb{E} \left[ \exp \left( \frac{1}{2} \int_0^T \mu^2(\omega, s) ds \right) \right].$$

Aus der Jensenschen Ungleichung folgt  $N(\alpha\mu) \leq N(\mu)$  für  $\alpha \in [0, 1]$  und laut Voraussetzung gilt  $N(\mu) < \infty$ . Mit  $\beta = (1 - \delta^2)^2$  erhalten wir

$$N(\beta\mu) = \mathbb{E} \left[ \exp \left( \frac{1}{2} (1 - \delta^2)^2 \int_0^T \mu^2(\omega, s) ds \right) \right] = \mathbb{E} \left[ \exp \left( \frac{1}{2} (1 + \delta)^2 \int_0^T (1 - \delta)^2 \mu^2(\omega, s) ds \right) \right].$$

Wenn wir nun  $\epsilon$  so wählen, dass  $(1 + \delta)^2 = 1 + \epsilon^2$  gilt, dann lässt sich die rechte Seite genau als  $\epsilon$ -Novikov-Bedingung für  $M((1 - \delta)\mu)$  interpretieren. Aus der Endlichkeit von  $N(\beta\mu)$  folgt daher mit Lemma 7.8, dass  $M((1 - \delta)\mu)$  ein Martingal ist. Mit Hilfe der Hölderschen Ungleichung erhält man

$$\begin{aligned} 1 = \mathbb{E}[M((1 - \delta)\mu)_t] &= \mathbb{E} \left[ \exp \left( (1 - \delta) \left( X_t - \frac{1}{2} Y_t \right) \right) \exp \left( \frac{\delta(1 - \delta)}{2} Y_t \right) \right] \leq \\ &\leq (M(\mu)_t)^{1 - \delta} N((1 - \delta)\mu)^\delta \leq \mathbb{E}[M(\mu)_t]^{1 - \delta} N(\mu)^\delta \end{aligned}$$

für alle  $t \in [0, T]$ . Lassen wir nun  $\delta$  gegen Null gehen folgt  $1 \leq \mathbb{E}[M(\mu)_t]$  und damit die Martingaleigenschaft von  $(M_t(\mu))_{0 \leq t \leq T}$ .  $\square$

## 8 Martingaldarstellung und Doob-Meyer Zerlegung

### 8.1 Der Martingaldarstellungssatz

In Kapitel 3 haben wir gezeigt, dass das Ito-Integral eines  $\mathcal{H}^2$ -Integranden ein stetiges Martingal ist. Der Martingaldarstellungssatz beschäftigt sich mit der Umkehrung dieser Aussage: Lässt sich ein stetiges Martingal als Ito-Integral darstellen? Für Martingale in der Brownschen Standardfiltration die einer quadratischen Integrierbarkeitsbedingung genügen, lautet die Antwort Ja.

**Theorem 8.1.** *Sei  $X$  ein Martingal bezüglich der Brownschen Standardfiltration mit  $\mathbb{E}[X_T^2] < \infty$ . Dann existiert  $\phi(\omega, t) \in \mathcal{H}^2[0, T]$ , sodass*

$$X_t = x_0 + \int_0^t \phi(\omega, s) dB_s, \quad \forall t \in [0, T] \quad (8.1)$$

*gilt. Der Integrand  $\phi(\omega, t)$  ist eindeutig bis auf  $d\mathbb{P} \times dt$ -Nullmengen.*

*Beweis der Eindeutigkeit in Thm. 8.1.* Dies ist der einfachere Teil des Beweises. Es gelte die Darstellung (8.1) einerseits mit  $\phi(\omega, t) \in \mathcal{H}^2[0, T]$  und andererseits mit  $\psi(\omega, t) \in \mathcal{H}^2[0, T]$ . Dann ist auch  $\phi(\omega, t) - \psi(\omega, t)$  in  $\mathcal{H}^2[0, T]$  und es gilt

$$0 = \int_0^T (\phi(\omega, s) - \psi(\omega, s)) dB_s.$$

Wir berechnen die Varianz beider Seiten und erhalten mit der Ito-Isometrie

$$0 = \mathbb{E} \left[ \int_0^T (\phi(\omega, s) - \psi(\omega, s))^2 ds \right].$$

Der Integrand der rechten Seite ist positiv, daher muss er  $d\mathbb{P} \times dt$ -fast überall Null sein und die Eindeutigkeit von  $\phi$  ist bewiesen.  $\square$

Für den Beweis der Darstellung (8.1) stellen wir zunächst mal fest, dass es ausreicht sie für  $t = T$  zu zeigen. Aus

$$X_T = x_0 + \int_0^T \phi(\omega, s) dB_s \quad (8.2)$$

folgt nämlich nach Anwendung der bedingten Erwartung  $\mathbb{E}[\cdot | \mathcal{F}_t]$  auf beiden Seiten die Gleichung (8.1). Der restliche Beweis erfolgt in zwei Schritten: Zeige erstens die Darstellung (8.2) für eine große Klasse von Zufallsvariablen und zeige dann zweitens, dass diese Klasse von Zufallsvariablen bereits dicht in  $L_2(\Omega, \mathcal{F}_T, \mathbb{P})$  ist, womit dann



(8.2) mittels Approximation folgt. Wir formulieren diese Schritte in den folgenden zwei Lemmas genauer aus.

**Lemma 8.2.** *Sei  $0 \leq t_0 \leq \dots \leq t_n = T$ . Dann besitzt die Zufallsvariable*

$$Z = \prod_{j=0}^{n-1} \exp(\mathbf{i}\theta_j(B_{t_{j+1}} - B_{t_j}))$$

eine Darstellung

$$Z = \mathbb{E}[Z] + \int_0^T \phi(\omega, t) dB_t \quad \text{mit } \phi \in \mathcal{H}^2[0, T]. \quad (8.3)$$

*Beweis.* Wir setzen  $Z_j = \exp(\mathbf{i}\theta_j(B_{t_{j+1}} - B_{t_j}))$  und zeigen die gewünschte Darstellung zuerst für die einzelnen Faktoren  $Z_j$  und dann für das Produkt  $Z = \prod_{j=0}^{n-1} Z_j$ . Wir modifizieren das altbekannte Martingal  $\exp(-\frac{1}{2}\sigma^2 t + \sigma B_t)$  ein wenig und ersetzen den reellen Parameter  $\sigma$  durch einen imaginären Parameter  $\mathbf{i}\theta$ . So erhält man das komplexwertige Martingal

$$X_t = \exp\left(\frac{1}{2}\theta^2 t + \mathbf{i}\theta B_t\right)$$

welches die SDE  $dX_t = \mathbf{i}\theta X_t dB_t$  erfüllt. Durch Integration der SDE bekommen wir

$$X_{t+s} = X_s + \int_s^{t+s} \mathbf{i}\theta X_u dB_u$$

und nach Einsetzen der expliziten Gestalt von  $X_t$  ergibt sich

$$\exp\left(\frac{1}{2}\theta^2(t+s) + \mathbf{i}\theta B_{t+s}\right) = \exp\left(\frac{1}{2}\theta^2 s + \mathbf{i}\theta B_s\right) + \int_s^{t+s} \mathbf{i}\theta \exp\left(\frac{\theta^2}{2}u + \mathbf{i}\theta B_u\right) dB_u.$$

Diese Gleichung lässt sich umformen auf

$$\exp(\mathbf{i}\theta(B_{s+t} - B_s)) = \exp\left(-\frac{\theta^2}{2}t\right) + \int_s^{s+t} \mathbf{i}\theta \exp\left(\frac{\theta^2}{2}(s+t-u) + \mathbf{i}\theta B_u\right) dB_u$$

und wir erhalten somit für jeden der Faktoren  $Z_j := \exp(\mathbf{i}\theta(B_{t_{j+1}} - B_{t_j}))$  die Darstellung

$$Z_j = a_j + \int_0^T \psi_j(\omega, s) dB_s$$

mit  $\psi_j(\omega, s) = 0$  außerhalb von  $[t_j, t_{j+1})$ . Wir können nun auch

$$Z_j(t) := \mathbb{E}[Z_j | \mathcal{F}_t] = a_j + \int_0^t \psi_j(\omega, s) dB_s$$

definieren und erhalten mit der Produktformel für Ito-Prozesse

$$\begin{aligned} d(Z_j(t)Z_k(t)) &= Z_j(t) dZ_k(t) + Z_k(t) dZ_j(t) + \psi_j(\omega, t)\psi_k(\omega, t)dt = \\ &= (Z_j(t)\psi_k(\omega, t) + Z_k(t)\psi_j(\omega, t)) dB_t, \quad (j \neq k) \end{aligned}$$

und somit wieder eine Darstellung der Form (8.3). Die Eigenschaft  $(Z_j\psi_k + Z_k\psi_j) \in \mathcal{H}^2[0, T]$  folgt dabei aus der Beschränktheit der  $Z_j(t)$ . Das  $n$ -fache Wiederholen dieser Überlegung liefert dann die Behauptung.  $\square$

**Lemma 8.3.** *Sei  $\mathcal{S}$  die Menge aller Zufallsvariablen welche sich als Linearkombinationen von Zufallsvariablen der Form*

$$Z = \prod_{j=0}^{n-1} \exp(i\theta_j(B_{t_{j+1}} - B_{t_j}))$$

darstellen lassen. Dann ist  $\mathcal{S}$  eine dichte Teilmenge des Hilbertraums

$$L_2(\Omega, \mathcal{F}_T, \mathbb{P}) = \{X \in \mathcal{F}_T : \mathbb{E}[X_T^2] < \infty\}.$$

*Beweis.* Wir skizzieren den Beweis nur. Sei  $\mathcal{D}$  der Raum aller Zufallsvariablen in  $L_2(\Omega, \mathcal{F}_T, \mathbb{P})$ , welche sich in der Form

$$X = f(B_{t_1}, B_{t_2} - B_{t_1}, \dots, B_{t_n} - B_{t_{n-1}}), \quad f : \mathbb{R}^n \rightarrow \mathbb{R} \text{ messbar}$$

darstellen lassen. Dann folgt zunächst die Dichtheit von  $\mathcal{S}$  in  $\mathcal{D}$  aus klassischen Argumenten der Fourieranalysis bzw. Hilbertraumtheorie.<sup>14</sup> Mit dem Satz über monotone Klassen aus der Maßtheorie folgt, dass  $\mathcal{D}$  dicht in  $L_2(\Omega, \mathcal{F}_T, \mathbb{P})$  liegt. An dieser Stelle geht entscheidend ein, dass  $\mathbb{F}$  die Brownsche Standardfiltration ist, d.h. alleine von der Brownschen Bewegung und allen  $\mathbb{P}$ -Nullmengen erzeugt wird.  $\square$

*Beweis von Theorem 8.1.* Nach Lemma 8.2 und 8.3 existiert eine Folge von Zufallsvariablen  $X^n$  in  $L_2(\Omega, \mathcal{F}_T, \mathbb{P})$  mit  $X^n \rightarrow X_T$  und den Darstellungen

$$X^n = \mathbb{E}[X^n] + \int_0^T \phi_n(\omega, s) dB_s, \quad \phi_n \in \mathcal{H}^2[0, T]. \quad (8.4)$$

<sup>14</sup> $L_2(\Omega, \mathcal{F}_T, \mathbb{P})$  ist ein Hilbertraum und  $\mathbb{E}[\exp(i\theta_i(B_t - B_s))]$  kann man als Fouriertransformation der Dichte von  $B_t - B_s$  interpretieren.

Aus der  $L_2$ -Konvergenz folgt  $\mathbb{E}[X^n] \rightarrow \mathbb{E}[X_T]$  und somit ist  $X^n - \mathbb{E}[X^n]$  eine Cauchyfolge in  $L_2(d\mathbb{P})$ . Mit der Ito-Isometrie folgt dass  $\phi_n(\omega, s)$  ebenfalls eine Cauchyfolge in  $L_2(d\mathbb{P} \times dt)$  ist. Da  $\mathcal{H}^2[0, T]$  ein vollständiger Teilraum von  $L_2(\Omega, \mathcal{F}_T, \mathbb{P})$  ist, existiert  $\phi(\omega, t) \in \mathcal{H}^2[0, T]$  mit  $\phi_n \rightarrow \phi$  in  $L_2$ . Wir können nun auf beiden Seiten von (8.4) den  $L_2$ -Limes bilden und erhalten die Behauptung.  $\square$

Als Anwendung des Martingaldarstellungssatzes beweisen wir **Lévy's Charakterisierung der Brownschen Bewegung**. Diese besagt im Wesentlichen, dass ein stetiges Martingal dessen quadratische Variation der einer Brownschen Bewegung entspricht, auch eine Brownsche Bewegung sein muss.

**Theorem 8.4** (Lévy's Charakterisierung der Brownschen Bewegung). *Sei  $M$  ein stetiges Martingal in der Brownschen Standardfiltration mit  $M_0 = 0$ ,  $\mathbb{E}[M_t^2] < \infty$  und quadratischer Variation  $[M, M]_t = t$  für alle  $t \in \mathbb{R}_{\geq 0}$ . Dann ist  $M$  eine Brownsche Bewegung.*

*Beweis.* Mit dem Martingaldarstellungssatz existiert  $\phi(\omega, t) \in \mathcal{H}^2[0, T]$  mit  $dM_t = \phi(\omega, t)dB_t$ . Wir wenden die Ito-Formel für Ito-Prozesse an und erhalten

$$\begin{aligned} d\left(e^{i\theta(M_t - M_s)}\right) &= i\theta e^{i\theta(M_t - M_s)}dM_t - \frac{1}{2}\theta^2 e^{i\theta(M_t - M_s)}d[M, M]_t = \\ &= i\theta e^{i\theta(M_t - M_s)}\phi(\omega, t)dB_t - \frac{1}{2}\theta^2 e^{i\theta(M_t - M_s)}dt. \end{aligned}$$

In Integralform geschrieben, gilt also

$$e^{i\theta(M_t - M_s)} - 1 = \int_s^t i\theta e^{i\theta(M_u - M_s)}\phi(\omega, u)dB_u - \frac{1}{2}\theta^2 \int_s^t e^{i\theta(M_u - M_s)}du$$

und unter dem bedingten Erwartungswert folgt daraus

$$\mathbb{E}\left[e^{i\theta(M_t - M_s)} \middle| \mathcal{F}_s\right] - 1 = -\frac{1}{2}\theta^2 \int_s^t \mathbb{E}\left[e^{i\theta(M_u - M_s)} \middle| \mathcal{F}_s\right] du.$$

Dies bedeutet, die Funktion  $\alpha(t) = \mathbb{E}\left[e^{i\theta(M_t - M_s)} \middle| \mathcal{F}_s\right]$  erfüllt die DGL

$$\alpha'(t) = -\frac{\theta^2}{2}\alpha(t), \quad \alpha(s) = 1$$

mit der einzigen Lösung  $\alpha(t) = e^{-\frac{\theta^2}{2}(t-s)}$  und wir haben gezeigt, dass

$$\mathbb{E}\left[e^{i\theta(M_t - M_s)} \middle| \mathcal{F}_s\right] = \exp\left(-\frac{\theta^2}{2}(t-s)\right)$$

gilt. Es folgt, dass die Inkremente von  $M$  unabhängig und normalverteilt sind, mit Erwartungswert 0 und Varianz  $t - s$ . Somit ist  $M$  eine Brownsche Bewegung, wie behauptet.  $\square$

## 8.2 Doob-Meyer-Zerlegung

Wir zeigen nun ein weiteres wichtiges Darstellungsergebnis, welches für positive Submartingale gilt.

**Theorem 8.5** (Doob-Meyer-Zerlegung). *Sei  $X$  ein positives stetiges Submartingal. Dann existiert ein stetiges Martingal  $M$  und ein fast sicher steigender stetiger und adaptierter Prozess  $A$  mit  $A_0 = 0$ , sodass*

$$X_t = M_t + A_t, \quad t \geq 0.$$

*Die Zerlegung ist eindeutig bis auf Ununterscheidbarkeit.*

Es sei betont, dass wir im Unterschied zum Martingaldarstellungssatz und den meisten anderen Resultaten der letzten Kapitel bei diesem Theorem keine Bedingung an die Filtration stellen. Insbesondere muss  $\mathbb{F}$  nicht die Brownsche Standardfiltration sein. Die hier dargestellte Version der Doob-Meyer-Zerlegung ist die Version für stetige Submartingale. Es existiert eine allgemeinere Version der Doob-Meyer-Zerlegung für *rechtsstetige* Submartingale, welche auch noch die wichtige Information liefert, dass  $A$  ein sogenannter *vorhersehbarer* Prozess ist. Vorhersehbarkeit ist eine Verschärfung der Adaptiertheit von stochastischen Prozessen, die für stetige Prozesse jedoch mit dieser zusammenfällt. Der Beweis der Doob-Meyer-Zerlegung ist anspruchsvoll, sodass wir ihn hier nur skizzieren.<sup>15</sup>

*Beweisskizze Doob-Meyer-Zerlegung.* Für diskrete Submartingale ist die Doob-Meyer-Zerlegung leicht herzuleiten und wird dort auch einfach Doob-Zerlegung genannt. Sei also  $(X_n)_{n \in \mathbb{N}}$  ein diskretes Submartingal bezüglich einer diskreten Filtration  $(\mathcal{F}_n)_{n \in \mathbb{N}}$ . Wir definieren  $A_0 = 0$  und

$$A_{n+1} := A_n + \mathbb{E}[X_{n+1} - X_n | \mathcal{F}_n] = \sum_{j=0}^n \mathbb{E}[X_{j+1} - X_j | \mathcal{F}_j].$$

Aus der Submartingaleigenschaft von  $X$  folgt, dass  $A_n$  fast sicher steigend ist. Wir

<sup>15</sup>Ein gut lesbarer und relativ kurzer vollständiger Beweis findet sich in dem Artikel ‘A short proof of the Doob-Meyer theorem’ von Beiglböck, Schachermayer und Veliyev (<http://arxiv.org/abs/1012.5292>).

setzen  $M_n = X_n - A_n$  und erhalten mit

$$\mathbb{E}[M_n | \mathcal{F}_k] = \sum_{j=0}^{n-1} (\mathbb{E}[X_{j+1} - X_j | \mathcal{F}_k] - \mathbb{E}[\mathbb{E}[X_{j+1} - X_j | \mathcal{F}_j] | \mathcal{F}_k]) = M_k$$

die Martingaleigenschaft von  $M$ . Wir haben also die Doob-Zerlegung für diskrete Submartingale gezeigt. Sei nun  $X$  ein stetiges Submartingal auf  $[0, T]$ . Mit Lemma 1.11 existiert eine konvexe Funktion  $\Phi : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$  mit  $\lim_{r \rightarrow \infty} \frac{\Phi(r)}{r} = \infty$ , sodass  $\mathbb{E}[\Phi(X_T)] < \infty$ . Mit dem Satz vom optionalen Sampling gilt für jede Stoppzeit  $\tau \leq T$  die Ungleichung

$$\mathbb{E}[\Phi(X_\tau)] \leq \mathbb{E}[\Phi(X_T)] < \infty.$$

Mit dem Kriterium aus Lemma 1.10 für die gleichgradige Integrierbarkeit folgt daraus, dass  $X$  ‘von Klasse DL’<sup>16</sup> ist, das heisst die Menge

$$\{X_\tau : \tau \text{ Stoppzeit}, \tau \leq T\} \quad (8.5)$$

ist gleichmässig Integrierbar. Sei nun  $\mathcal{P}^n$  die Folge der verfeinernden dyadischen Partitionen von  $[0, T]$ . Mit der diskreten Doob-Zerlegung erhalten wir eine Folge von Prozessen  $(M^n, A^n)_{n \in \mathbb{N}}$  auf den dyadischen Partitionen  $\mathcal{P}^n$ . Jeder dieser Prozesse lässt sich zu einem Prozess in stetiger Zeit erweitern, indem wir  $M_t^n = \mathbb{E}[M_T^n | \mathcal{F}_t]$  setzen und die Prozesse  $A^n$  zwischen den Gitterpunkten rechtstetig konstant interpolieren. Wir erhalten also für jedes  $n \in \mathbb{N}$  ein Martingal  $M_t^n$  und einen steigenden Prozess  $A_t^n$  welche auf den Gitterpunkten der dyadischen Partition  $\mathcal{P}^n$  mit der diskreten Doob-Zerlegung übereinstimmen. Als nächsten Schritt wählen wir aus dieser Folge von Prozessen eine konvergente Teilfolge aus. Es lässt sich zeigen, dass aus der gleichgradigen Integrierbarkeit (8.5) auch die gleichgradige Integrierbarkeit von  $\{A_T^n\}_{n \in \mathbb{N}}$  folgt. (Dies ist der langwierigste Teil des Beweises, den wir hier nicht näher ausführen.) Mit dem Satz von Dunford & Pettis folgt aus der gleichgradigen Integrierbarkeit die Existenz einer in  $L_1(d\mathbb{P})$  konvergenten Teilfolge  $A_T^{n_k} \rightarrow A_T$ . Wegen  $M_T^n = X_T - A_T^n$  erhalten wir auch die  $L_1(d\mathbb{P})$ -Konvergenz  $M_T^{n_k} \rightarrow M_T = X_T - A_T$ . Nun definieren wir  $M_t = \mathbb{E}[M_T | \mathcal{F}_t]$  und  $A_t = X_t - M_t$  erhalten die Konvergenz

$$A_t^n = (S_t - M_t^n) \rightarrow (S_t - M_t) = A_t, \quad \text{in } L_1(d\mathbb{P})$$

auf den Gitterpunkten von  $\mathcal{P}^n$ . Durch Auswahl einer weiteren Teilfolge erhalten wir eine fast sicher konvergente Teilfolge und schliessen, dass  $t \mapsto A_t$  steigend ist. Es bleibt nun noch zu zeigen, dass die Zerlegung eindeutig ist, dass  $A$  und  $M$  stetige Prozesse sind und das Resultat von  $[0, T]$  auf das unbeschränkte Intervall  $\mathbb{R}_{\geq 0}$  auszuweiten.

---

<sup>16</sup>siehe auch Kapitel 4

Auch diese letzten Schritte überspringen wir in dieser Beweisskizze.  $\square$

Wir besprechen nun zwei wichtige Anwendungen der Doob-Meyer-Zerlegung.

- Sei  $M$  ein stetiges Martingal. Dann ist  $(M_t^2)_{t \geq 0}$  als konvexe Transformation ein positives Submartingal. Der steigende Prozess  $A_t$  in der Doob-Meyer-Zerlegung von  $M_t^2$  wird mit  $\langle M, M \rangle_t$  bezeichnet und heisst **vorhersehbare quadratische Variation**. Es existiert also stets ein stetiger, steigender Prozess  $\langle M, M \rangle_t$ , sodass  $M_t^2 - \langle M, M \rangle_t$  ein Martingal ist. Nehmen wir nun an  $M$  sei ein Ito-Prozess von der Form  $M_t = \int_0^t b(\omega, s) dB_s$ . Dann wissen wir aus Kapitel 2, dass

$$M_t^2 - \int_0^t b(\omega, s)^2 ds$$

ein Martingal ist. Aus der Eindeutigkeit der Doob-Meyer-Zerlegung folgern wir, dass

$$\langle M, M \rangle_t = \int_0^t b(\omega, s)^2 ds$$

gilt. Insbesondere stimmt die vorhersehbare quadratische Variation mit der quadratischen Variation überein, es gilt also<sup>17</sup>

$$\langle M, M \rangle_t = [M, M]_t.$$

Was haben wir durch diese Überlegungen gewonnen? Die Erkenntnis ist, dass auch im Fall, dass das stetige Martingal  $M$  keine Darstellung  $M_t = \int_0^t b(\omega, s) ds$  besitzt ein Prozess  $\langle M, M \rangle_t$  existiert, sodass  $M_t^2 - \langle M, M \rangle_t$  ein Martingal ist. Die Bedeutung dieser Erkenntnis ist vor allem dann einzusehen, wenn wir uns in Erinnerung rufen wie wichtig Martingale von der Gestalt  $M_t^2 - \langle M, M \rangle_t$  im Beweis der Ito-Formel waren. Mit der Existenz des Martingals  $M_t^2 - \langle M, M \rangle_t$  haben wir einen Ansatz gewonnen um die Ito-Formel jenseits von Ito-Prozessen auf beliebige stetige Martingale zu erweitern.

- Sei  $B$  Brownsche Bewegung. Dann ist  $|B_t|_{t \geq 0}$  als konvexe Transformation von  $B$  ein positives Submartingal. Eine interessante Eigenschaft von  $|B_t|$  ist, dass sich dieser Prozess ausserhalb einer Umgebung der Null wie Brownsche Bewegung verhält. Mit der Doob-Meyer-Zerlegung schliessen wir, dass ein stetiger steigender Prozess  $L_t^0$  existiert, genannt **Lokale Zeit bei 0**, sodass  $|B_t| - L_t^0$  ein Martingal ist.

---

<sup>17</sup>Falls  $M$  ein rechtsstetiges Martingal ist, so gilt im allgemeinen aber  $\langle M, M \rangle_t \neq [M, M]_t$ , was auch die unterschiedliche Notation erklärt.

## 9 Die Diffusionsgleichung

Die Diffusionsgleichung ist ein Beispiel für die erstaunliche Effizienz der Mathematik bei der Beschreibung unterschiedlichster naturwissenschaftlicher Phänomene.<sup>18</sup> Die Diffusionsgleichung

$$\frac{\partial u}{\partial t} = \lambda \frac{\partial^2 u}{\partial x^2}, \quad \lambda > 0$$

ist eine partielle Differentialgleichung und tritt in der Theorie der Wärmeleitung, der Ausbreitung von Gasen und in der Beschreibung der Bewegung von Mikroorganismen auf. Eine ähnliche Differentialgleichung (allerdings mit negativem Vorzeichen vor  $\lambda$ ) ist bereits in Kapitel 5 im Zusammenhang mit der Ito-Formel aufgetaucht. Bevor wir uns im nächsten Kapitel mit den Verbindungen zur Stochastik beschäftigen, werden wir zuerst die Diffusionsgleichung aus einer analytischen Perspektive betrachten. Diese Reihenfolge entspricht auch der historischen Entwicklung der mathematischen Theorie.

### 9.1 Herleitung der Diffusionsgleichung

Wir wollen das Phänomen der Diffusion anhand der Ausbreitung von einzelligen Organismen, in diesem Fall Amöben beschreiben. Der Einfachheit halber nehmen wir an, dass die Amöben auf einem langgezogenen streifenförmigen Gebiet leben, sodass eine einzige Koordinate  $x$  zu ihrer Positionsbestimmung ausreicht. Des weiteren nehmen wir an, dass keine nennenswerten Nahrungsquellen in dem Gebiet zu finden sind (Amöben ernähren sich beispielsweise von Bakterien indem sie diese einschliessen). Die Abwesenheit von Nahrungsquellen hat zwei Auswirkungen: Die Amöben stellen die Zellteilung ein (wir dürfen also annehmen dass die Population konstant ist) und anstatt an den Nahrungsquellen zu verharren beginnen Sie eine Suchbewegung bei der sie dünn besiedelte Gebiete gegenüber dicht besiedelten bevorzugen. Wir finden also zwei zur Herleitung der Diffusionsgleichung wesentliche Eigenschaften:

- Eine **Erhaltungsgröße** – in diesem Fall die Gesamtpopulation
- Ein **Fluss** von grösserer zu kleiner Konzentration

Zu einem Zeitpunkt  $t$  und an einem Punkt  $x$  schreiben wir  $u(t, x)$  für die lokale Konzentration der Amöbenpopulation und  $q(t, x)$  für die Rate mit der die Amöben den Punkt  $x$  von links nach rechts überqueren. Aufgrund der Erhaltung der Gesamtpopulation lässt sich die Anzahl der Amöben die ein Intervall  $[x_0, x_0 + \Delta]$  verlassen als

$$\frac{\partial}{\partial t} \int_{x_0}^{x_0 + \Delta x} u(t, x) dx = q(t, x_0) - q(t, x_0 + \Delta x)$$

---

<sup>18</sup>‘The unreasonable efficiency of mathematics in the natural sciences’ ist der Titel eines bekannten Aufsatzes des Nobelpreisträgers Eugene Wigner.

schreiben. Division durch  $\Delta x$  und der Grenzübergang  $\Delta x \rightarrow 0$  liefert die **Erhaltungsgleichung**

$$\frac{\partial}{\partial t} u(t, x) = -\frac{\partial}{\partial x} q(t, x). \quad (9.1)$$

Die Rate  $q(t, x)$  der Bewegung wird durch die lokale Änderung der Konzentration bestimmt, lässt sich also als Funktion

$$q(t, x) = F\left(\frac{\partial}{\partial x} u(t, x)\right)$$

schreiben. Da bei konstanter Konzentration keine Bewegung stattfindet muss  $F(0) = 0$  gelten. Weil die Amöben sich von grösserer zu kleinerer Konzentration bewegen ergibt sich in erster Näherung der lineare Ansatz

$$q(t, x) = -\lambda \frac{\partial u}{\partial x}(t, x) \quad (9.2)$$

wobei  $\lambda$  die Motilität (Bewegungsfreude) der Amöben beschreibt. Einsetzen dieser Gleichung in die Erhaltungsgleichung (9.1) liefert die Diffusionsgleichung

$$\frac{\partial u}{\partial t} = \lambda \frac{\partial^2 u}{\partial x^2}. \quad (9.3)$$

Die zwei wesentlichen Modelleigenschaften – Erhaltungsgrösse und Fluss von grösserer zu kleiner Konzentration – finden sich in vielen anderen physikalischen und naturwissenschaftlichen Phänomenen. Betrachtet man etwa in einer Trägerflüssigkeit gelöste Farbpartikel und deren lokale Konzentration  $u(t, x)$  so ist klarerweise die Anzahl der Partikel eine Erhaltungsgrösse und (9.2) entspricht dem sogenannten Fickschen Gesetz der Konzentrationsveränderung, womit auch hier die Diffusionsgleichung (9.3) gilt. Beschreibt  $u(t, x)$  die Temperaturverteilung in einem stabförmigen Medium, so gilt (9.1) aufgrund der Energieerhaltung und (9.2) entspricht dem Newtonschen Abkühlungsgesetz. Aus diesem Grund wird die Diffusionsgleichung auch oft **Wärmeleitungsgleichung** genannt.

Wir werden uns vorerst mit der Lösung der einfachsten Form (9.3) der Diffusionsgleichung beschäftigen. Jedoch sind sofort verschiedene Erweiterungen des besprochenen Modells denkbar. So könnte die Motilität  $\lambda$  der Amöben von der lokalen Beschaffenheit der Umgebung am Punkt  $x$  sowie vom Zeitpunkt  $t$  abhängen und somit durch eine Funktion  $\lambda(t, x)$  gegeben sein. Ausserdem könnte neben der Bewegung von größerer zu kleinerer Populationsdichte auch eine systematische Wanderungsbewegung – etwa aufgrund von chemischen Botenstoffen – mit der (zeit- und ortsabhängigen) Rate  $a(t, x)$



stattfinden, womit dann anstatt von (9.2) die allgemeinere Gleichung

$$q(t, x) = -\lambda(t, x) \frac{\partial u}{\partial x}(t, x) - a(t, x)u(t, x) \quad (9.4)$$

gelten würde. Wir könnten auch die Annahme der konstanten Population hinterfragen und postulieren, dass der Saldo aus Vermehrungs- und Sterberate am Ort  $x$  und zum Zeitpunkt  $t$  durch eine Funktion  $c(t, x)$  gegeben ist. In diesem Fall erhalten wir anstatt der Erhaltungsgleichung (9.1) die Gleichung

$$\frac{\partial}{\partial t} u(t, x) = -\frac{\partial}{\partial x} q(t, x) + c(t, x)u(t, x). \quad (9.5)$$

und gemeinsam mit (9.4) die verallgemeinerte Diffusionsgleichung

$$\frac{\partial u}{\partial t}(t, x) = \lambda(t, x) \frac{\partial^2 u}{\partial x^2}(t, x) + \tilde{a}(t, x) \frac{\partial u}{\partial x}(t, x) + \tilde{c}(t, x)u(t, x). \quad (9.6)$$

wobei  $\tilde{a}(t, x) = a(t, x) + \frac{\partial}{\partial x} \lambda(t, x)$  und  $\tilde{c}(t, x) = c(t, x) + \frac{\partial}{\partial x} a(t, x)$ .

## 9.2 Lösen der Diffusionsgleichung

Wir betrachten die Diffusionsgleichung

$$\frac{\partial u}{\partial t} = \lambda \frac{\partial^2 u}{\partial x^2} \quad (9.7)$$

für  $u \in C^{1,2}(\mathbb{R}_{\geq 0} \times \mathbb{R})$  und mit vorgegebener Anfangsbedingung

$$u(0, x) = f(x), \quad x \in \mathbb{R}.$$

### Methode 1: Fouriertransformation

Eine Methode diese Diffusionsgleichung zu lösen, basiert auf der Fouriertransformation. Gegeben eine Funktion  $\phi \in L^1(\mathbb{R})$  schreiben wir

$$\hat{\phi}(\theta) = \int_{-\infty}^{\infty} e^{i\theta x} \phi(x) dx, \quad \theta \in \mathbb{R}$$

für ihre Fouriertransformation. Falls wir eine Funktion von  $t$  und  $x$  betrachten, so wird nur in der Variable  $x$  transformiert, das heisst wir schreiben

$$\hat{g}(t, \theta) = \int_{-\infty}^{\infty} e^{i\theta x} g(t, x) dx, \quad \theta \in \mathbb{R}.$$

Die entscheidende Beobachtung ist, dass die Anwendung der Fouriertransformation die partielle Differentialgleichung (9.7) in eine *gewöhnliche* Differentialgleichung mit bekannter Lösung überführt. Nach Vertauschen von Integration und Differentiation erhalten wir auf der Linken Seite von (9.7)

$$\int_{-\infty}^{\infty} e^{i\theta x} u_t(t, x) dx = \hat{u}_t(\theta, x)$$

und auf der rechten Seite folgt mit partieller Integration unter der Annahme das  $u_x$  und  $u$  ausreichend schnell gegen 0 abfallen

$$\int_{-\infty}^{\infty} e^{i\theta x} u_{xx}(t, x) dx = -\theta^2 \int_{-\infty}^{\infty} e^{i\theta x} u(t, x) dx = -\theta^2 \hat{u}(t, \theta).$$

Weiters erhalten wir aus der Anfangsbedingung

$$\hat{u}(0, \theta) = \hat{f}(\theta).$$

Es gilt also die gewöhnliche Differentialgleichung

$$\hat{u}_t(t, \theta) = -\lambda\theta^2 \hat{u}(t, \theta), \quad \hat{u}(0, \theta) = \hat{f}(\theta)$$

zu lösen. Die Lösung dieser DGL ist natürlich

$$\hat{u}(t, \theta) = \hat{f}(\theta) e^{-\lambda\theta^2 t}$$

und wir müssen nur noch die Fouriertransformation invertieren um  $u(t, x)$  zu erhalten. Für diese Inversion machen wir uns zwei Eigenschaften der Fouriertransformation zunutze:

- Sei  $\gamma(x)$  die Faltung der Funktionen  $\psi(x)$  und  $\phi(x)$ , d.h.

$$\gamma(x) = \psi \star \phi(x) = \int_{-\infty}^{\infty} \psi(x-y)\phi(y) dy.$$

Dann ist die Fouriertransformation von  $\gamma$  gleich dem Produkt der Fouriertransformierten von  $\psi$  und  $\phi$ , d.h.

$$\hat{\gamma}(\theta) = \widehat{\phi \star \psi}(\theta) = \hat{\phi}(\theta) \cdot \hat{\psi}(\theta), \quad \theta \in \mathbb{R}.$$

- Sei  $k(x, \lambda, t)$  die Dichte der Normalverteilung mit Mittelwert 0 und Varianz  $2\lambda t$ , d.h.

$$k(x, \lambda, t) = \frac{1}{2\sqrt{\pi\lambda t}} e^{-\frac{x^2}{4\lambda t}}. \quad (9.8)$$

Dann gilt  $\hat{k}(\theta, \lambda, t) = e^{-\lambda\theta^2 t}$ .

Wegen seiner Bedeutung in der Theorie der Diffusionsgleichung wird  $k(x, \lambda, t)$  auch Wärmeleitungskern ('heat kernel') genannt. Durch Kombination der beiden diskutierten Eigenschaften erhalten wir

$$\widehat{k \star f}(\theta) = \hat{f}(\theta) \cdot e^{-\lambda\theta^2 t} = \hat{u}(t, \theta).$$

Da die Fouriertransformation die zugrundeliegende Funktion eindeutig bestimmt, muss also gelten

$$u(t, x) = k \star f(x) = \int_{-\infty}^{\infty} f(y) \frac{1}{2\sqrt{\pi\lambda t}} e^{-\frac{(x-y)^2}{4\lambda t}} dy.$$

Wir schliessen, dass sich die Lösung der Diffusionsgleichung durch Faltung der Anfangsbedingung mit dem Wärmeleitungskern bestimmen lässt. Folgendes Theorem fasst die gewonnene Erkenntnis zusammen.

**Theorem 9.1.** *Sei  $f \in C(\mathbb{R})$  eine Funktion die der Wachstumsbedingung*

$$|f(x)| \leq C e^{Bx^\rho}, \quad \forall x \in \mathbb{R}$$

mit  $B, C \in \mathbb{R}$  und  $\rho < 2$  genügt. Definiere

$$u(t, x) = f \star k(x, \lambda, t) = \int_{-\infty}^{\infty} f(y) k(x - y, \lambda, t) dx, \quad (t, x) \in (0, \infty) \times \mathbb{R}$$

Dann ist  $u \in C^{1,2}((0, \infty) \times \mathbb{R})$  und erfüllt die Diffusionsgleichung

$$u_t(t, x) = \lambda u_{xx}(t, x), \quad \forall (t, x) \in (0, \infty) \times \mathbb{R}.$$

Weiters kann  $u(t, x)$  stetig auf  $\mathbb{R}_{\geq 0} \times \mathbb{R}$  fortgesetzt werden, sodass

$$u(0, x) = f(x), \quad \forall x \in \mathbb{R}$$

gilt.

*Beweis.* Durch direktes Nachrechnen erhält man für die Ableitungen des Wärmeleitungskerns

$$k_t(x, \lambda, t) = \left(-\frac{1}{2t} + \frac{x^2}{4\lambda t^2}\right) k(x, \lambda, t), \quad \text{und} \quad k_{xx}(x, \lambda, t) = \left(-\frac{1}{2\lambda t} + \frac{x^2}{4\lambda^2 t^2}\right) k(x, \lambda, t)$$

und es ist ersichtlich, dass  $k(x, \lambda, t)$  die Diffusionsgleichung erfüllt, d.h.

$$k_t(x, \lambda, t) = \lambda k_{xx}(x, \lambda, t), \quad \forall (t, x) \in (0, \infty) \times \mathbb{R}.$$

Wegen der Voraussetzung an  $f$  sind auch  $f(y)k_t(x-y, \lambda, t)$  und  $f(y)k_{xx}(x-y, \lambda, t)$  integrierbar; wir dürfen also Integration und Ableitung vertauschen und damit erfüllt auch  $u(t, x)$  die Diffusionsgleichung. Indem wir eine standardnormalverteilte Zufallsgröße  $Z$  einführen, können wir  $u(t, x)$  in kompakter Form als

$$u(t, x) = \mathbb{E} \left[ f(x + \sqrt{2\lambda t}Z) \right]$$

schreiben, da die Dichte von  $\sqrt{2\lambda t}Z$  genau dem Wärmeleitungskern  $k(x, \lambda, t)$  entspricht. Für  $t = 0$  gilt  $u(0, x) = f(x)$ , für  $t > 0$  erfüllt  $u(t, x)$  die Diffusionsgleichung. Es bleibt also nur noch zu zeigen, dass  $u(t, x)$  für  $t \rightarrow 0$  stetig ist. Für  $t < 1/(2\lambda)$  gilt die Abschätzung

$$\left| f(x + \sqrt{2\lambda t}Z) \right| \leq C \exp(B(|x| + |Z|)^p)$$

und die Zufallsgröße auf der rechten Seite besitzt einen endlichen Erwartungswert. Wir können also dominierte Konvergenz anwenden und es folgt mit

$$\lim_{t \rightarrow 0} \mathbb{E} \left[ f(x + \sqrt{2\lambda t}Z) \right] = \mathbb{E} \left[ \lim_{t \rightarrow 0} f(x + \sqrt{2\lambda t}Z) \right] = f(x)$$

die Stetigkeit von  $u(t, x)$  auf  $\mathbb{R}_{\geq 0} \times \mathbb{R}$ . □

## Methode 2: Potenzreihenentwicklung

Ein anderer möglicher Ansatz ergibt sich aus der Annahme, dass die Lösung  $u(t, x)$  eine Taylorentwicklung um  $t = 0$  besitzt:

$$u(t, x) = \sum_{n=0}^{\infty} \frac{t^n}{n!} \frac{\partial^n u}{\partial t^n}(0, x). \quad (9.9)$$

Mit Iteration der Diffusionsgleichung (9.7) erhalten wir

$$\frac{\partial^n u}{\partial t^n}(t, x) = \lambda^n \frac{\partial^{2n} u}{\partial x^{2n}}(t, x), \quad (t, x) \in \mathbb{R}_{\geq 0} \times \mathbb{R}.$$

Einsetzen von  $t = 0$  und die Anfangsbedingung  $u(0, x) = f(x)$  liefern

$$\frac{\partial^n u}{\partial t^n}(t, x) = \lambda^n f^{(2n)}(x), \quad x \in \mathbb{R}$$

und wir erhalten aus dem Ansatz (9.9) die Lösung

$$u(t, x) = \sum_{n=0}^{\infty} \frac{t^n \lambda^n}{n!} f^{(2n)}(x). \quad (9.10)$$

Für diese Lösungsmethode muss natürlich ausreichende Differenzierbarkeit von  $f$  und die Konvergenz der auftretenden Taylorentwicklungen vorausgesetzt werden. Im konkreten Fall liefert die Methode jedoch oft geschlossenen Ausdrücke für  $u(t, x)$  von denen sich direkt überprüfen lässt ob sie Lösungen von (9.7) sind. Falls zum Beispiel  $f$  ein Polynom ist, so bricht die Entwicklung (9.10) schon nach endlich vielen Gliedern ab und liefert einen geschlossenen Ausdruck. Ein weiteres Beispiel ist die Anfangsbedingung  $f(x) = e^{\alpha x}$ . In diesem Fall erhalten wir mit

$$u(t, x) = \sum_{n=0}^{\infty} \frac{t^n \lambda^n}{n!} \alpha^{2n} e^{\alpha x} = e^{\lambda t \alpha^2 + \alpha x}$$

eine Lösung der Diffusionsgleichung. Für  $f(x) = \cos x$  gilt  $f^{(2n)}(x) = (-1)^n \cos x$  und der Potenzreihenansatz liefert

$$u(t, x) = \sum_{n=0}^{\infty} \frac{t^n \lambda^n}{n!} (-1)^n \cos x = e^{-\lambda t} \cos x.$$

### Method 3: Symmetrie

Wir suchen nach Symmetrien, das heisst Transformationen des Variablenraums  $\mathbb{R}_{\geq 0} \times \mathbb{R}$  welche die Differentialgleichung (9.7) unverändert lassen. Solche Transformationen müssen demnach jede Lösung in eine weitere Lösung überführen, wodurch sich ein weiterer Ansatz ergeben wird. Wir betrachten eine besonders einfache Art der Transformation, nämlich die linearen Streckungen

$$\tau = \alpha t, \quad y = \beta x.$$

Weiters schreiben wir  $v(\tau, y) = u(t, x)$  für die Lösung in Abhängigkeit von den transformierten Variablen. Die partiellen Ableitungen von  $v$  erfüllen

$$u_t = \alpha v_\tau, \quad u_x = \beta v_y, \quad u_{xx} = \beta^2 v_{yy}.$$

Wenn wir  $\alpha = \beta^2$  wählen, sehen wir, dass  $v(\tau, y)$  wiederum die Diffusionsgleichung

$$v_\tau = \lambda v_{yy}$$

erfüllt. Jede Lösung der Diffusionsgleichung wird also unter den Transformationen

$$t \mapsto \alpha t, \quad x \mapsto \alpha^{1/2} x$$

wieder in eine Lösung übergeführt; man spricht von einer Symmetriegruppe<sup>19</sup> der Diffusionsgleichung (9.7). Wir parametrisieren die Lösung mit dieser Symmetriegruppe und wählen den Ansatz

$$u_\alpha(t, x) = u(\alpha t, \sqrt{\alpha}x).$$

Jede Lösung dieser Art lässt sich als Funktion  $u(t, x) = \psi(x/\sqrt{t})$  schreiben. Einsetzen in (9.7) liefert

$$-\frac{1}{2}xt^{-3/2}\psi'(x/\sqrt{t}) = \lambda\frac{1}{t}\psi''(x/\sqrt{t}).$$

Nach Umformen und der Parametrisierung  $y = x/\sqrt{t}$  ergibt sich die gewöhnliche Differentialgleichung  $y\psi'(y) = -2\lambda\psi''(y)$  mit der allgemeinen Lösung  $\psi'(y) = c \exp\left(-\frac{y^2}{4\lambda}\right)$ . Nach nochmaliger Integration ergibt sich als eine mögliche Lösung

$$u(t, x) = \Phi\left(\frac{x}{\sqrt{2\lambda t}}\right),$$

wobei  $\Phi$  die Verteilungsfunktion der Standardnormalverteilung bezeichnet. Mit  $t \rightarrow 0$  sehen wir, dass diese Lösung der Anfangsbedingung

$$u(0, x) = \begin{cases} 1 & x > 0 \\ 1/2 & x = 0 \\ 0 & x < 0 \end{cases}$$

entspricht und mit der Lösung übereinstimmt die wir aus dieser Anfangsbedingung mit Theorem (9.1) erhalten.

### 9.3 Eindeutigkeit der Lösung

Wir beginnen wieder mit einem Potenzreihenansatz und zwar mit

$$u(t, x) = \sum_{n=0}^{\infty} \frac{x^{2n}}{(2n)!} \frac{\partial^n \phi}{\partial t^n}(t), \tag{9.11}$$

---

<sup>19</sup>Es ist leicht zu sehen, dass die Transformationen eine Gruppe bilden; mit  $\alpha = 1$  erhält man das neutrale Element und mit  $\alpha \mapsto \alpha^{-1}$  das Inverse.

wobei  $\phi \in C^\infty(\mathbb{R}_{\geq 0})$ . Wir erhalten

$$\begin{aligned} u_{xx}(t, x) &= \sum_{n=0}^{\infty} \frac{1}{(2n)!} 2n(2n-1)x^{2n-2} \frac{\partial^n \phi}{\partial t^n}(t) = \\ &= \sum_{n=0}^{\infty} \frac{x^{2n}}{(2n)!} \frac{\partial^{n+1} \phi}{\partial t^{n+1}}(t) = u_t(t, x) \end{aligned}$$

und sehen, dass  $u(t, x)$  eine Lösung der Diffusionsgleichung ist. Mit  $\phi \equiv 0$  erhalten wir natürlich zur Anfangsbedingung  $u(0, x) = 0$  die Lösung  $u(t, x) \equiv 0$ . Interessanter wird es wenn wir die Funktion

$$\phi(t) = \begin{cases} e^{-1/t^2}, & t > 0 \\ 0, & t \leq 0 \end{cases}$$

betrachten. Diese Funktion liegt einigen Gegenbeispielen der Analysis zugrunde und eine ihrer interessanten Eigenschaften ist, dass sie in  $C^\infty(\mathbb{R}_{\geq 0})$  liegt, mit Ableitungen  $\phi^{(n)}(0) = 0$  für alle  $n \in \mathbb{N}$ . Dies hat zur Konsequenz, dass die Taylorentwicklung von  $\phi$  am Punkt 0 die konstante Nullfunktion ergibt und somit in keiner noch so kleinen Umgebung von  $t = 0$  gegen  $\phi$  konvergiert. In unserem Fall setzten wir  $\phi$  in den Ansatz (9.11) ein und erhalten eine weitere Lösung  $u(t, x)$  der Diffusionsgleichung mit Anfangsbedingung  $u(0, x) = 0$ , jedoch mit  $u(t, x) \neq 0$  für  $t > 0$ . Durch Linearkombination mit dieser Lösung folgt, dass die Diffusionsgleichung für jede Anfangsbedingung  $u(0, x) = f(x)$  mehrere Lösungen besitzt. Wie wir sehen werden ist es jedoch möglich unter zusätzlichen Anforderungen an die Lösung Eindeutigkeit zu erzwingen. Schlüssel zu diesen Überlegungen ist das sogenannte **parabolische Maximumsprinzip**.<sup>20</sup>

Für das parabolische Maximumsprinzip betrachten wir einen offenen Raum-Zeit-Zylinder

$$D = \{(t, x) : 0 < t < T, A < x < B\}$$

und definieren dessen **parabolischen Rand** als

$$\partial_0 D = \partial D \setminus \{(T, x) : A < x < B\}.$$

Der parabolische Rand von  $D$  entspricht also dem gewöhnlichen Rand von  $D$  abzüglich der Verbindungsstrecke der Punkte  $(T, A)$  und  $(T, B)$ . Das parabolische Maximumsprinzip besagt nun, dass die Lösung  $u(t, x)$  der Diffusionsgleichung ihr Maximum über den Raum-Zeit-Zylinder  $D$  stets auf dessen parabolischen Rand  $\partial_0 D$  annimmt. Wieso sollte solch ein Maximumsprinzip gelten? In unserer Herleitung über die Amöbenpopulation würde ein Maximum an einem Punkt  $(t_0, x_0) \in D$  ausserhalb des parabolischen

<sup>20</sup>Die Diffusionsgleichung gehört zur Kategorie der ‘parabolischen’ partiellen Differentialgleichungen.

schen Rands bedeuten, dass sich im Zeitverlauf die Populationsdichte lokal am Punkt  $x_0$  gegenüber den umliegenden Punkten konzentriert hat. Dies stellt einen Widerspruch zu unserer Modellannahme dar, dass die Amöben sich stets von höherer Dichte zu niedriger Dichte bewegen und sich nicht vermehren. Ähnlich kann bei der Modellierung der Wärmeleitung argumentiert werden, dass bei Abwesenheit von Wärmequellen die Temperatur sich lokal nicht gegenüber allen umliegenden Punkten spontan erhöhen kann. Wir formulieren das Maximumsprinzip etwas allgemeiner für Funktionen die statt der Diffusionsgleichung eine ‘Diffusionsungleichung’ erfüllen.

**Theorem 9.2** (Parabolisches Maximumsprinzip). *Sei  $D$  ein Raum-Zeit-Zylinder und  $u \in C^{1,2}(D) \cap C(\overline{D})$  eine Funktion die*

$$\frac{\partial u}{\partial t}(t, x) \leq \frac{\partial^2 u}{\partial x^2}(t, x) \quad \forall (t, x) \in D \quad (9.12)$$

*erfüllt. Dann nimmt  $u$  das Maximum über  $\overline{D}$  stets auf dem parabolischen Rand  $\partial_0 D$  von  $D$  an, d.h.*

$$\max_{(t,x) \in \overline{D}} u(t, x) = \max_{(t,x) \in \partial_0 D} u(t, x).$$

*Beweis.* Wir wählen  $\delta > 0$   $\epsilon \in [0, T]$  und zeigen das Maximumsprinzip zunächst für die Funktion  $v(t, x) = u(t, x) - \delta t$  auf dem verkleinerten Raum-Zeit-Zylinder

$$D(\epsilon) = \{(t, x) : 0 < t < T - \epsilon, A < x < B\}.$$

Aus (9.12) folgt

$$v_t(t, x) < v_{xx}(t, x), \quad \forall (t, x) \in \overline{D(\epsilon)}. \quad (9.13)$$

Wir nehmen nun an, dass das Maximum von  $v$  über  $\overline{D(\epsilon)}$  an einem Punkt  $(t_0, x_0)$  im offenen Zylinder  $D(\epsilon)$  angenommen wird. Dann besitzt die Funktion  $x \mapsto v(t_0, x)$  ein lokales Maximum bei  $x = x_0$  und es gilt  $v_{xx}(t_0, x_0) \leq 0$ ; genauso besitzt  $t \mapsto v(t, x_0)$  ein lokales Maximum bei  $t = t_0$  und es gilt  $v_t(t_0, x_0) = 0$ , im Widerspruch zu (9.13). Als nächstes nehmen wir an, dass das Maximum von  $v$  über  $\overline{D(\epsilon)}$  an einem Punkt  $(T - \epsilon, x_0)$  angenommen wird. In diesem Fall schliessen wir wie vorhin  $v_{xx}(T - \epsilon, x_0) \leq 0$  und erhalten außerdem  $v_t(T - \epsilon, x_0) \geq 0$ , wieder im Widerspruch zu (9.13). Wir haben also gezeigt, dass  $v$  das Maximum über  $\overline{D(\epsilon)}$  auf  $\partial_0 D(\epsilon)$  annimmt, d.h. es gilt

$$\max_{(t,x) \in \overline{D(\epsilon)}} v(t, x) = \max_{(t,x) \in \partial_0 D(\epsilon)} v(t, x).$$

Da  $v$  stetig auf der kompakten Menge  $\overline{D}$  ist, gilt mit  $\epsilon \rightarrow 0$

$$\max_{(t,x) \in \overline{D}} v(t, x) = \max_{(t,x) \in \partial_0 D} v(t, x).$$



Des weiteren gilt wegen  $v(t, x) = u(t, x) - \delta t$

$$\begin{aligned} -\delta T + \max_{(t,x) \in \bar{D}} u(t, x) &\leq \max_{(t,x) \in \bar{D}} (u(t, x) - \delta t) = \\ &= \max_{(t,x) \in \partial_0 D} (u(t, x) - \delta t) \leq \max_{(t,x) \in \partial_0 D} u(t, x) \end{aligned}$$

und mit  $\delta \rightarrow 0$  folgt die Behauptung.  $\square$

Aus dem eben gezeigten Maximumsprinzip lässt sich eine Eindeutigkeitsaussage für die Diffusionsgleichung herleiten. Wieder zeigen wir zuerst ein etwas allgemeineres Resultat, dass mit Ungleichungen statt Gleichungen arbeitet.

**Theorem 9.3.** *Sei  $u \in C^{1,2}((0, \infty) \times \mathbb{R}) \cap C(\mathbb{R}_{\geq 0} \times \mathbb{R})$  und seien  $B, C \in \mathbb{R}$ , sodass  $u$  die Wachstumsbedingung*

$$|u(t, x)| \leq Ce^{Bx^2}, \quad \forall (t, x) \in \mathbb{R}_{\geq 0} \times \mathbb{R}$$

erfüllt. Wenn  $u$  die Differentialungleichung

$$u_t(t, x) \leq u_{xx}(t, x), \quad \forall (t, x) \in (0, \infty) \times \mathbb{R}$$

und die Anfangsbedingung  $u(0, x) \leq 0$  für alle  $x \in \mathbb{R}$  erfüllt, dann gilt

$$u(t, x) \leq 0, \quad \forall (t, x) \in \mathbb{R}_{\geq 0} \times \mathbb{R}.$$

Aus diesem Theorem lässt sich leichte folgendes Korollar ableiten.

**Korollar 9.4.** *Sei  $u$  die in Theorem 9.1 definierte Lösung der Diffusionsgleichung*

$$u_t(t, x) = \lambda u_{xx}(t, x), \quad \forall (t, x) \in (0, \infty) \times \mathbb{R}$$

mit Anfangsbedingung  $u(0, x) = f(x)$ ,  $f \in C(\mathbb{R})$  und sei  $v(t, x)$  eine weitere Lösung die derselben Anfangsbedingung genügt und  $|v(t, x)| \leq Ce^{Bx^2}$  erfüllt. Dann gilt

$$u(t, x) = v(t, x), \quad \forall (t, x) \in \mathbb{R}_{\geq 0} \times \mathbb{R}.$$

*Beweis von Theorem 9.3.* Es reicht die Behauptung für fixes  $T$  auf dem Gebiet  $[0, T] \times \mathbb{R}$  zu zeigen, da mit der Transformation  $t' = t - T$  die Aussage auf  $[0, 2T] \times \mathbb{R}$  und iterativ auf ganz  $\mathbb{R}_{\geq 0} \times \mathbb{R}$  ausgedehnt werden kann. Den Zeithorizont  $T$  wählen wir dabei ausreichend klein, sodass  $T < 1/(8B)$  gilt. Für  $\delta > 0$  und  $y \in \mathbb{R}$  definieren wir

$$v(t, x) = u(t, x) - \delta(2T - t)^{-1/2} \exp\left(\frac{(x - y)^2}{4(2T - t)}\right).$$

Da die Differenz  $v(t, x) - u(t, x)$  die Diffusionsgleichung mit  $\lambda = 1$  erfüllt gilt

$$v_t(t, x) \leq v_{xx}(t, x), \quad \forall (t, x) \in [0, T] \times \mathbb{R}.$$

Für beliebiges  $h > 0$  betrachten wir den Raum-Zeit-Zylinder

$$D(y, h) = \{(t, x) : 0 < t < T, y - h < x < y + h\}.$$

Mit dem parabolischen Maximumsprinzip gilt

$$\max_{(t,x) \in \bar{D}(y,h)} v(t, x) = \max_{(t,x) \in \partial_0 D(y,h)} v(t, x). \quad (9.14)$$

Wir untersuchen das Verhalten von  $v(t, x)$  auf dem parabolischen Rand  $\partial_0 D$ . An der oberen Seite gilt  $x = y + h$ ,  $t \in [0, T]$  und somit

$$\begin{aligned} v(t, x) &\leq C \exp(B(y+h)^2) - \delta(2T-t)^{-1/2} \exp\left(\frac{(x-y)^2}{4(2T-t)}\right) \leq \\ &\leq C \exp(B(y+h)^2) - \delta(2T)^{-1/2} \exp(h^2/(8T)). \end{aligned}$$

Wegen  $B < 1/(8T)$  wird für  $h$  ausreichend groß die rechte Seite negativ. Dasselbe Argument liefert, dass auch auf der unteren Seite des parabolischen Randes für  $h$  ausreichend groß  $v(t, x)$  negativ wird. Auf dem verbleibenden Teil des parabolischen Randes gilt  $t = 0$  und somit laut Anfangsbedingung

$$v(0, x) \leq u(0, x) \leq 0.$$

Wir haben gezeigt, dass  $v(t, x)$  auf dem gesamten parabolischen Rand  $\partial_0 D(y, h)$  für ausreichend großes  $h > 0$  negativ ist und mit dem Maximumsprinzip (9.14) folgt dass  $v(t, x)$  auf dem gesamten Zylinder  $\bar{D}(y, h)$  negativ ist. Für  $h$  ausreichend groß folgt also

$$u(t, x) - \delta(2T-t)^{-1/2} \exp\left(\frac{(x-y)^2}{4(2T-t)}\right) = v(t, x) \leq 0 \quad \forall y \in \mathbb{R}, (t, x) \in D(y, h)$$

und mit  $\delta \rightarrow 0$  erhalten wir die gewünschte Aussage.  $\square$

*Beweis von Korollar 9.4.* Wie im Beweis von Theorem 9.1 schreiben wir  $u(t/\lambda, x) = \mathbb{E}[f(x + \sqrt{2}Z)]$  mit  $Z$  standardnormalverteilt und schätzen mit

$$\begin{aligned} |u(t/\lambda, x)| &\leq \mathbb{E}[|f(x + \sqrt{2}Z)|] \leq C \mathbb{E}\left[\exp\left(B(|x| + \sqrt{2}|Z|)^\rho\right)\right] \leq \\ &\leq A \exp\left(2^{\rho-1}B|x|^\rho\right) \mathbb{E}\left[\exp\left(2^{3\rho/2-1}B|Z|^\rho\right)\right] \end{aligned}$$

ab, wobei wir ohne Beschränkung der Allgemeinheit annehmen können, dass  $\rho \in [1, 2)$ . Damit erfüllen  $u(t/\lambda, x) - v(t/\lambda, x)$  und  $v(t/\lambda, x) - u(t/\lambda, x)$  die Voraussetzungen von Theorem 9.3 und die Behauptung folgt.  $\square$

## 10 Die Feynman-Kac-Formel und das Dirichlet-Problem

Die Feynman-Kac-Formel (nach dem Physiker und Nobelpreisträger Richard Feynman und dem Mathematiker Mark Kac) stellt den angekündigten Zusammenhang zwischen der Diffusionsgleichung und der Brownschen Bewegung oder allgemeiner zwischen der Diffusionsgleichung und stochastischen Differentialgleichungen her. Dieser Zusammenhang soll hier motiviert werden indem wir nochmals die Amöbenpopulation aus dem letzten Kapitel betrachten. Wie zuvor setzten wir die Populationsdichte  $f(x)$  zum Zeitpunkt 0 als bekannt voraus und wollen die Populationsdichte  $u(t, x)$  zum Zeitpunkt  $t$  modellieren. Anstatt die Dichte  $u(t, x)$  direkt zu betrachten konzentrieren wir uns nun auf das Verhalten einer einzelnen Amöbe und nehmen einen stochastischen Blickwinkel ein. Wir machen folgende Modellannahmen.

- Die Suchbewegung einer einzelnen Amöbe folgt einer Brownschen Bewegung mit Diffusionskoeffizient  $\sigma$ .
- Die Amöben bewegen sich unabhängig voneinander und unabhängig von ihrer Startposition zum Zeitpunkt 0.

Wir wählen aus der Gesamtpopulation zum Zeitpunkt  $t$  eine Amöbe zufällig aus und schreiben  $X_t$  für ihre Position zum Zeitpunkt  $t$  und  $X_0$  für ihre Startposition zum Zeitpunkt 0. Sowohl  $X_t$  als auch  $X_0$  sind Zufallsvariablen und nach unserer Modellannahme gilt

$$X_t = X_0 + \sigma B_t.$$

Die Populationsdichte  $u(t, x)$  lässt sich schreiben als

$$u(t, x) = \lim_{\Delta x \rightarrow 0} \frac{1}{\Delta x} \mathbb{P}[x \leq X_t < x + \Delta x].$$

Gleichzeitig gilt

$$\begin{aligned} \mathbb{P}[x \leq X_t < x + \Delta x] &= \mathbb{P}[x \leq X_0 + \sigma B_t < x + \Delta x] = \\ &= \mathbb{P}\left[x + \sigma \tilde{B}_t \leq X_0 < x + \sigma \tilde{B}_t + \Delta x\right] \end{aligned}$$

wobei  $\tilde{B}_t = -B_t$  wieder eine Brownsche Bewegung ist. Wegen der Anfangsbedingung an die Populationsdichte gilt außerdem

$$\mathbb{P}[a \leq X_0 < a + \Delta x] = f(a)\Delta x + o(\Delta x).$$

Nach Bedingen auf die Brownsche Bewegung  $\tilde{B}$  erhalten wir aus diesen Gleichungen

$$\begin{aligned} u(t, x) &= \lim_{\Delta x \rightarrow 0} \frac{1}{\Delta x} \mathbb{E} \left[ \mathbb{P} \left[ x + \sigma \tilde{B}_t \leq X_0 < x + \sigma \tilde{B}_t + \Delta x \mid \sigma((\tilde{B}_t)_{t \geq 0}) \right] \right] = \\ &= \lim_{\Delta x \rightarrow 0} \frac{1}{\Delta x} \mathbb{E} \left[ f(x + \sigma \tilde{B}_t) \Delta x + o(\Delta x) \right] = \\ &= \mathbb{E} [f(x + \sigma B_t)], \end{aligned}$$

wobei wir  $\tilde{B}_t$  unter dem Erwartungswert wieder durch  $B_t$  ersetzen können, da die Verteilungen übereinstimmen. Wir haben also für die Populationsdichte die Darstellung

$$u(t, x) = \mathbb{E} [f(x + \sigma B_t)] \quad (10.1)$$

hergeleitet. Mithilfe der Dichte der Normalverteilung lässt sich dies schreiben als

$$u(t, x) = \int_{-\infty}^{\infty} f(x - y) \frac{1}{\sqrt{2\pi t \sigma^2}} \exp\left(-\frac{y^2}{2t\sigma^2}\right) dy$$

und wir sehen, dass  $u(t, x)$  mit der in Theorem 9.1 hergeleiteten Lösung der Diffusionsgleichung

$$u_t(t, x) = \frac{\sigma^2}{2} u_{xx}(t, x), \quad u(0, x) = f(x)$$

übereinstimmt! Der stochastische Ansatz zur Beschreibung der Amöbenpopulation führt also zum genau gleichen Ergebnis wie der analytische Ansatz wenn wir  $\lambda = \sigma^2/2$  setzen. Zudem haben wir mit (10.1) gezeigt, dass sich die Lösung der Diffusionsgleichung als Erwartungswert eines Funktionals der Brownschen Bewegung schreiben lässt. Diese Art der Darstellung nennt man Feynman-Kac-Formel.

## 10.1 Feynman-Kac-Formel für Brownsche Bewegung

**Theorem 10.1** (Feynman-Kac für Brownsche Bewegung). *Sei  $q : \mathbb{R} \rightarrow \mathbb{R}$  von oben beschränkt und sei  $u(t, x)$  eine Lösung der Diffusionsgleichung*

$$u_t(t, x) = \frac{\sigma^2}{2} u_{xx}(t, x) + q(x)u(t, x), \quad u(0, x) = f(x)$$

welche für jedes  $T \geq 0$  auf  $[0, T] \times \mathbb{R}$  beschränkt ist. Dann gilt die Darstellung

$$u(t, x) = \mathbb{E} \left[ f(x + \sigma B_t) \exp \left( \int_0^t q(x + \sigma B_s) ds \right) \right].$$

Im Unterschied zur obigen Herleitung tritt hier nun auch der Potentialterm  $q(x)u(t, x)$  in der Diffusionsgleichung auf. Nach der Diskussion im letzten Kapitel wissen wir dass

ein solcher Term sich als Sterbe- bzw. Vermehrungsrate im Populationsmodell interpretieren lässt. In der stochastischen Darstellung von  $u(t, x)$  schlägt sich der Potentialterm als exponentieller Faktor nieder. In der Finanzmathematik treten solche exponentiellen Faktoren typischerweise als Abzinsungsfaktoren auf; in diesen Fällen lässt sich der Potentialterm als Zinsrate interpretieren.

*Beweis.* Wir definieren

$$M_s = \underbrace{u(t-s, x + \sigma B_s)}_{:=U_s} \exp \left( \underbrace{\int_0^s q(x + \sigma B_r) dr}_{:=I_s} \right), \quad s \leq t$$

mit der Ito-Formel und unter Verwendung der partiellen Differentialgleichung für  $u(t, x)$  gilt

$$\begin{aligned} dU_s &= -u_t(t-s, x + \sigma B_s) ds + u_x(t-s, x + \sigma B_s) \sigma dB_s + \\ &+ \frac{1}{2} u_{xx}(t-s, x + \sigma B_s) \sigma^2 ds = \\ &= u_x(t-s, x + \sigma B_s) dB_s - q(x + \sigma B_s) u(t-s, x + \sigma B_s) ds. \end{aligned}$$

Mit der Produktregel für Ito-Prozesse gilt weiter

$$dM_s = U_s dI_s + I_s dU_s = \exp \left( \int_0^s q(x + \sigma B_r) dr \right) u_x(t-s, x + \sigma B_s) dB_s$$

und wir folgern, dass  $M$  ein lokales Martingal ist. Aufgrund der Annahmen können wir mit

$$\sup_{0 \leq s \leq t} |M_s| \leq \sup_{0 \leq s \leq t} |u(t, x)| \exp(t \|q^+\|_\infty) < \infty$$

abschätzen. Der Prozess  $M$  ist daher ein beschränktes lokales Martingal und somit ein echtes Martingal. Es gilt also

$$u(t, x) = M_0 = \mathbb{E}[M_t] = \mathbb{E} \left[ f(x + \sigma B_s) \exp \left( \int_0^t q(x + \sigma B_s) ds \right) \right]$$

und die Behauptung ist bewiesen.  $\square$

Als erste Anwendung der gezeigten Feynman-Kac-Formel leiten wir Paul Lévy's **Arcus-Sinus-Gesetz** für die Brownsche Bewegung her. Das Arcus-Sinus-Gesetz beschreibt die Verteilung der Zeit die die Brownsche Bewegung im Intervall  $[0, \infty)$  ver-

bringt, das heisst die Verteilung der Zufallsvariable

$$T_t = \int_0^t \mathbf{1}\{B_s \geq 0\} ds.$$

Wenn wir die Brownsche Bewegung als Kontostand aus einem fairen Spiel betrachten, so ist  $T_t/t$  die Proportion der Zeit die wir ‘im Plus’ verbringen. Es ist klar, dass  $T_t/t$  in  $[0, 1]$  liegt und aus Symmetriegründen gilt  $\mathbb{E}[T_t] = t/2$ . Das Arcus-Sinus-Gesetz von Lévy beschreibt die Verteilung von  $T_t$  folgendermaßen.

**Proposition 10.2** (Arcus-Sinus-Gesetz). *Die Zufallsvariable  $T_t = \int_0^t \mathbf{1}\{B_s \geq 0\} ds$  folgt der Verteilung*

$$\mathbb{P}\left[\frac{T_t}{t} \leq p\right] = \frac{2}{\pi} \arcsin \sqrt{p} = \frac{1}{\pi} \int_0^p \frac{1}{\sqrt{u(1-u)}} du, \quad p \in [0, 1]. \quad (10.2)$$

Interessant ist dass die Dichte  $\frac{1}{\pi} \frac{1}{\sqrt{u(1-u)}}$  ihr Minimum bei  $u = 1/2$  annimmt und der Erwartungswert somit gleichzeitig der unwahrscheinlichste Wert ist. Die Dichte konzentriert sich an den Punkten  $u = 0$  und  $u = 1$  was bedeutet, dass in den wahrscheinlichsten Fällen die Brownsche Bewegung sich entweder durchgehend in  $[0, \infty)$  oder durchgehend in  $(-\infty, 0]$  aufhält.

*Beweis.* Wir berechnen zunächst die Laplace-Transformierte  $\mathbb{E}[e^{-\lambda T_t}]$  von  $T_t$ . Wenn wir in Theorem 10.1  $f = 1$  und  $q(x) = -\lambda \mathbf{1}\{x \geq 0\}$  setzen, so erhalten wir

$$u(t, x) = \mathbb{E}\left[\exp\left(-\lambda \int_0^t \mathbf{1}\{B_s + x \geq 0\} ds\right)\right]$$

und somit  $\mathbb{E}[e^{-\lambda T_t}] = u(t, 0)$ . Als Voraussetzung muss  $u(t, x)$  die partielle Differentialgleichung

$$u_t(t, x) = \frac{1}{2} u_{xx}(t, x) - \lambda \mathbf{1}\{x \geq 0\} u(t, x), \quad u(0, x) = 1 \quad (10.3)$$

erfüllen. Um diese Gleichung zu lösen wenden wir auf beiden Seiten nochmals die Laplace-Transformation an.<sup>21</sup> Wir definieren

$$\hat{u}(\alpha, x) = \int_0^\infty e^{-t\alpha} u(t, x) dt$$

<sup>21</sup>Diese Technik ähnelt der Fourier-Methode zur Lösung der Diffusionsgleichung in Kapitel 9.

Mit partieller Integration gilt

$$\int_0^{\infty} e^{-t\alpha} u_t(t, x) dt = \alpha \hat{u}(\alpha, x) - 1$$

und Laplace-Transformation auf beiden Seiten von (10.3) ergibt nach Umordnen

$$\hat{u}_{xx}(\alpha, x) = \begin{cases} 2(\alpha + \lambda)\hat{u}(\alpha, x) - 2, & x \geq 0 \\ 2\alpha\hat{u}(\alpha, x) - 2 & x < 0. \end{cases} \quad (10.4)$$

Dies ist eine gewöhnliche Differentialgleichung vom Typ  $f''(x) = Af(x) + B$  deren allgemeine Lösung durch

$$f(x) = -\frac{B}{A} + C_1 \exp(-x\sqrt{A}) + C_2 \exp(x\sqrt{A})$$

gegeben ist. Um eine in  $x$  beschränkte Lösung  $u(t, x)$  zu erhalten muss auch  $\hat{u}(\alpha, x)$  in  $x$  beschränkt sein und wir erhalten als allgemeine beschränkte Lösung von (10.4)

$$\hat{u}(\alpha, x) = \begin{cases} \frac{1}{\alpha + \lambda} + C_1 \exp(-x\sqrt{2(\alpha + \lambda)}) & x \geq 0 \\ \frac{1}{\alpha} + C_2 \exp(x\sqrt{2\alpha}) & x < 0. \end{cases}$$

Es bleiben noch die Konstanten  $C_1$  und  $C_2$  zu bestimmen. Dazu verwenden wir das Prinzip des glatten Übergangs: Jede Lösung  $u(t, x)$  der Diffusionsgleichung muss in  $C^{1,2}(\mathbb{R}_{\geq 0} \times \mathbb{R})$  liegen und damit muss auch  $x \mapsto \hat{u}(\alpha, x)$  zweifach stetig differenzierbar sein. Es gilt

$$\lim_{x \downarrow 0} \hat{u}(\alpha, x) = \frac{1}{\alpha + \lambda} + C_1, \quad \lim_{x \uparrow 0} \hat{u}(\alpha, x) = \frac{1}{\alpha} + C_2$$

und außerdem

$$\lim_{x \downarrow 0} \hat{u}_x(\alpha, x) = -C_1 \sqrt{2(\alpha + \lambda)}, \quad \lim_{x \uparrow 0} \hat{u}_x(\alpha, x) = C_2 \sqrt{2\alpha}.$$

Die linken und rechten Limiten müssen jeweils übereinstimmen und wir erhalten insgesamt ein lineares Gleichungssystem für  $C_1, C_2$  mit den Lösungen

$$C_1 = \frac{\sqrt{\alpha + \lambda} - \sqrt{a}}{\sqrt{a}(\alpha + \lambda)}, \quad C_2 = \frac{\sqrt{a} - \sqrt{\alpha + \lambda}}{a\sqrt{(\alpha + \lambda)}}.$$

Somit haben wir gezeigt, dass

$$\int_0^{\infty} \mathbb{E}[e^{-\lambda T_t}] e^{-\alpha t} dt = \hat{u}(\alpha, 0) = \frac{1}{\sqrt{\alpha(\alpha + \lambda)}}.$$



Wir setzen nun die behauptete Dichte von  $T_t$  aus (10.2) ein. Wenn wir zeigen können, dass

$$\int_0^\infty \int_0^t \frac{1}{\pi} \frac{1}{x(t-x)} e^{-\lambda x} e^{-\alpha t} dx dt = \frac{1}{\sqrt{\alpha(\alpha + \lambda)}} \quad (10.5)$$

gilt, dann folgt aus der Eindeutigkeit der Laplace-Transformierten, dass (10.2) tatsächlich die Verteilung von  $T_t$  beschreibt. Durch Vertauschen der Integrationsreihenfolge und Substitution ergibt sich

$$\int_0^\infty \int_0^t \frac{1}{\pi} \frac{1}{x(t-x)} e^{-\lambda x} e^{-\alpha t} dx dt = \frac{1}{\pi} \int_0^\infty \frac{e^{-(\lambda+\alpha)x}}{\sqrt{x}} dx \int_0^\infty \frac{e^{-\alpha s}}{\sqrt{s}} ds.$$

Diese Integrale lassen sich mit der Gamma-Funktion  $\Gamma(z) = \int_0^\infty e^{-t} t^{z-1} dt$  auswerten, welche  $\Gamma(1/2) = \sqrt{\pi}$  erfüllt und wir erhalten

$$\frac{1}{\pi} \int_0^\infty \frac{e^{-(\lambda+\alpha)x}}{\sqrt{x}} dx \int_0^\infty \frac{e^{-\alpha s}}{\sqrt{s}} ds = \frac{1}{\pi} \frac{\Gamma(1/2) \cdot \Gamma(1/2)}{\sqrt{\alpha(\alpha + \lambda)}} = \frac{1}{\sqrt{\alpha(\alpha + \lambda)}},$$

womit Gleichung (10.5) bewiesen ist.  $\square$

## 10.2 Feynman-Kac-Formel für Diffusionsprozesse

Wir leiten nun zwei allgemeinere Formen der Feynman-Kac-Formel her, bei denen statt der Brownschen Bewegung die Lösung einer stochastischen Differentialgleichung auftritt.

**Theorem 10.3** (Feynman-Kac für Diffusionsprozesse). *Sei  $q : \mathbb{R} \rightarrow \mathbb{R}$  von oben beschränkt und sei  $u(t, x)$  eine Lösung der verallgemeinerten Diffusionsgleichung*

$$u_t(t, x) = \frac{1}{2} \sigma(x)^2 u_{xx}(t, x) + \mu(x) u_x(t, x) + q(x) u(t, x), \quad u(0, x) = f(x) \quad (10.6)$$

welche für jedes  $T \geq 0$  auf  $[0, T] \times \mathbb{R}$  beschränkt ist. Falls  $\sigma$  und  $\mu$  die Bedingungen aus Theorem 6.1 erfüllen, so hat die stochastische Differentialgleichung

$$dX_t = \mu(X_t) dt + \sigma(X_t) dB_t, \quad X_0 = 0 \quad (10.7)$$

eine eindeutige Lösung und es gilt

$$u(t, x) = \mathbb{E} \left[ f(x + X_t) \exp \left( \int_0^t q(x + X_s) ds \right) \right]$$

für alle  $(t, x) \in \mathbb{R}_{\geq 0} \times \mathbb{R}$ .

Das Resultat wird hier mit relativ starken Voraussetzungen formuliert, da dann der

Beweis vergleichsweise einfach ist. Es existieren Versionen der Feynman-Kac-Formel welche mit schwächeren Voraussetzungen auskommen oder auch stärkere Schlussfolgerungen liefern. Hier geht es vor allem um den prinzipiellen Zusammenhang den das Resultat aufzeigt: Die Lösung einer partiellen Differentialgleichung kann über den Erwartungswert eines Prozesses  $X$  dargestellt werden, welcher seinerseits eine stochastische Differentialgleichung löst. Die partielle Differentialgleichung (10.6) wird manchmal auch Kolmogorov'sche Rückwärtsgleichung zu (10.7) genannt.

*Beweis.* Wie auch schon im Beweis von Theorem 10.1 definieren wir den Prozess

$$M_s = \underbrace{u(t-s, x+X_s)}_{:=U_s} \exp \left( \underbrace{\int_0^s q(x+X_r) dr}_{:=I_s} \right), \quad s \leq t.$$

Mit der Ito-Formel für Ito-Prozesse und unter Verwendung der partiellen Differentialgleichung für  $u(t, x)$  gilt

$$\begin{aligned} dU_s &= -u_t(t-s, x+X_s)ds + u_x(t-s, x+X_s)\mu(x+X_s)ds + \\ &+ u_x(t-s, x+X_s)\sigma(x+X_s)dB_s + \frac{1}{2}u_{xx}(t-s, x+X_s)\sigma(x+X_s)^2ds = \\ &= u_x(t-s, x+X_s)\sigma(x+X_s)dB_s - q(x+X_s)u(t-s, x+X_s)ds. \end{aligned}$$

Mit der Produktregel für Ito-Prozesse gilt weiter

$$dM_s = U_s dI_s + I_s dU_s = I_s u_x(t-s, x+X_s)\sigma(x+X_s)dB_s$$

und wir folgern, dass  $M$  ein lokales Martingal ist. Aufgrund der Annahmen können wir mit

$$\sup_{0 \leq s \leq t} |M_s| \leq \sup_{0 \leq s \leq t} |u(t, x)| \exp(t \|q^+\|_\infty) < \infty$$

abschätzen. Der Prozess  $M$  ist daher ein beschränktes lokales Martingal und somit ein echtes Martingal. Es gilt also

$$u(t, x) = M_0 = \mathbb{E}[M_t] = \mathbb{E} \left[ f(x+X_s) \exp \left( \int_0^t q(x+X_s) ds \right) \right]$$

und die Behauptung ist bewiesen.  $\square$

In Problemstellungen der Finanzmathematik tritt die Diffusionsgleichung häufig mit einer Endbedingung statt mit einer Anfangsbedingung auf. Außerdem ist es von Interesse die Koeffizienten  $\mu$  und  $\sigma$  zeit- und ortsabhängig wählen zu können. Wir be-

trachten also das Endwertproblem

$$\begin{aligned} u_t(t, x) &= -\frac{1}{2}\sigma^2(t, x)u_{xx}(t, x) - \mu(t, x)u_x(t, x) - q(t, x)u(t, x), \\ u(T, x) &= h(x). \end{aligned} \quad (10.8)$$

Auch für diese partielle Differentialgleichung gilt eine Feynman-Kac-Formel.

**Theorem 10.4** (Feynman-Kac für Endwertprobleme). *Sei  $q : \mathbb{R} \rightarrow \mathbb{R}$  von oben beschränkt und sei  $u(t, x)$  eine Lösung des Endwertproblems (10.8) welche auf  $[0, T] \times \mathbb{R}$  beschränkt ist. Falls  $\sigma$  und  $\mu$  die Bedingungen aus Theorem 6.1 erfüllen, so hat die stochastische Differentialgleichung*

$$dX_s^{t,x} = \mu(s, X_s^{t,x})ds + \sigma(s, X_s^{t,x})dB_s, \quad X_t^{t,x} = x, \quad t \leq s \leq T. \quad (10.9)$$

für jedes  $(t, x) \in [0, T] \times \mathbb{R}$  eine eindeutige Lösung und es gilt

$$u(t, x) = \mathbb{E} \left[ h(X_T^{t,x}) \exp \left( \int_t^T q(X_s^{t,x}) ds \right) \right]$$

für alle  $(t, x) \in [0, T] \times \mathbb{R}$ .

*Beweis.* Wir definieren den Prozess

$$M_s = \underbrace{u(s, X_s^{t,x})}_{:=U_s} \underbrace{\exp \left( \int_s^T q(X_r^{t,x}) dr \right)}_{:=I_s}, \quad t \leq s \leq T.$$

Um die Notation zu vereinfachen halten wir in der folgenden Rechnung  $(t, x)$  fest und schreiben  $X_s \doteq X_s^{t,x}$ . Mit der Ito-Formel für Ito-Prozesse und unter Verwendung der partiellen Differentialgleichung für  $u(t, x)$  gilt

$$\begin{aligned} dU_s &= u_t(s, X_s)ds + u_x(s, X_s)\mu(s, X_s)ds + \\ &+ u_x(s, X_s)\sigma(s, X_s)dB_s + \frac{1}{2}u_{xx}(s, X_s)\sigma(s, X_s)^2ds = \\ &= u_x(s, X_s)\sigma(s, X_s)dB_s - q(s, X_s)u(s, X_s)ds. \end{aligned}$$

Mit der Produktregel für Ito-Prozesse gilt weiter

$$dM_s = U_s dI_s + I_s dU_s = I_s u_x(s, X_s) \sigma(s, X_s) dB_s$$

und wir folgern, dass  $(M_s)_{t \leq s \leq T}$  ein lokales Martingal ist. Aufgrund der Annahmen

können wir mit

$$\sup_{t \leq s \leq T} |M_s| \leq \sup_{0 \leq s \leq T} |u(t, x)| \exp(T \|q^+\|_\infty) < \infty$$

abschätzen. Der Prozess  $M$  ist daher ein beschränktes lokales Martingal und somit ein echtes Martingal. Es gilt also

$$u(t, x) = M_t = \mathbb{E}[M_T] = \mathbb{E} \left[ h(X_T^{t,x}) \exp \left( \int_t^T q(s, X_s^{t,x}) ds \right) \right]$$

und die Behauptung ist bewiesen.  $\square$

### 10.3 Harmonische Funktionen und das Dirichletproblem

In diesem Abschnitt beschäftigen wir uns mit einer weiteren Verbindung von partiellen Differentialgleichungen und stochastischen Prozessen. Sei  $D \subset \mathbb{R}^d$  eine offene Menge. Eine Funktion  $u : D \rightarrow \mathbb{R}$  heisst **harmonisch** wenn  $u \in C^2(D)$  und

$$\Delta u = 0, \quad \text{auf } D$$

gilt. Hier ist  $\Delta = \frac{\partial^2}{\partial x_1^2} + \dots + \frac{\partial^2}{\partial x_d^2}$  der aus Kapitel 5 bekannte Laplace-Operator. Wir bezeichnen mit  $K_r$  die offene Kugel  $\{x \in \mathbb{R}^d : |x| < r\}$  um 0 mit Radius  $r$  und mit  $V_r$  ihr Volumen. Wir betrachten einen messbaren Raum  $(\Omega, \mathcal{F})$  mit einer Familie von Wahrscheinlichkeitsmaßen  $(\mathbb{P}_x)_{x \in D}$  sodass  $B$  unter dem Maß  $\mathbb{P}_x$  eine Brownsche Bewegung mit Startwert  $B(0) = x$  ist.<sup>22</sup> Für jede Borel-messbare Menge  $A \subset \mathbb{R}^d$  definieren wir die Austrittszeit

$$\tau_A = \inf \{t \geq 0 : B_t \in A^c\}.$$

Auf dem Rand  $\partial K_r$  der Kugel  $K_r$  definieren wir ein Wahrscheinlichkeitsmaß  $\mu_r$  mittels

$$\mu_r(A) = \mathbb{P}_0 [B(\tau_{K_r}) \in A], \quad A \in \mathcal{B}(\partial K_r). \quad (10.10)$$

Das Maß  $\mu_r$  ordnet also jeder Borelmenge  $A$  auf der Oberfläche der Kugel  $K_r$  die Wahrscheinlichkeit zu, dass die Brownsche Bewegung beim ersten Verlassen der Kugel  $K_r$  durch einen Punkt in  $A$  austritt. Aufgrund der Rotationssymmetrie der Brownschen Bewegung lässt sich leicht zeigen, dass  $\mu_r$  genau die Gleichverteilung auf  $\partial K_r$  ist. Folgendes Lemma stellt den Zusammenhang zwischen den eingeführten Begriffen her.

<sup>22</sup>Die Maße  $\mathbb{P}_x$  lassen sich z.B. mit ähnlichen Überlegungen wie in Kapitel 7 als induzierte Maße auf dem Raum der stetigen Pfade realisieren.

**Lemma 10.5.** Sei  $u$  harmonisch auf  $D$ . Dann gilt die Mittelwerteigenschaft

$$u(a) = \int_{\partial K_r} u(a+x) \mu_r(dx) \quad (10.11)$$

für alle  $a \in D$  mit  $a + \overline{K}_r \subset D$ .

Die Mittelwerteigenschaft besagt, dass für jede abgeschlossene Kugel  $a + \overline{K}_r \subset D$  der Wert von  $u$  im Zentrum der Kugel dem Mittelwert über alle Funktionswerte von  $u$  auf der Oberfläche der Kugel entspricht. Durch Umformung des Integrals kann man zeigen, dass aus (10.11) auch

$$u(a) = \frac{1}{V_r} \int_{K_r} u(a+x) dx$$

folgt. Das bedeutet, dass die Mittelwerteigenschaft auch gilt, wenn man anstatt über die Oberfläche der Kugel über das gesamte Kugelvolumen mittelt. Der Beweis von Lemma 10.5 lässt sich wieder mit stochastischen Methoden führen.

*Beweis.* Mit der multivariaten Ito-Formel gilt

$$\begin{aligned} u(B(t \wedge \tau_{K_r})) &= u(B(0)) + \int_0^{t \wedge \tau_{K_r}} \nabla u(B_s) \cdot dB_s + \frac{1}{2} \int_0^{t \wedge \tau_{K_r}} \Delta u(B_s) ds = \\ &= u(B(0)) + \int_0^{t \wedge \tau_{K_r}} \nabla u(B_s) \cdot dB_s, \end{aligned}$$

wobei wir in der zweiten Zeile die harmonische Eigenschaft  $\Delta u = 0$  verwendet haben. Wir stellen zudem fest, dass wegen  $u \in C^2(D)$  der Gradient  $\nabla u$  beschränkt auf  $K_r$  ist. Unter dem Erwartungswert folgt aus der obigen Gleichung also

$$\mathbb{E}_a [u(B(t \wedge \tau_{K_r}))] = u(a)$$

Des weiteren gilt  $\mathbb{P}_a [\tau_{K_r} < \infty] = 1$ , da die Brownsche Bewegung jede beschränkte Menge fast sicher in endlicher Zeit verlässt. Mit  $t \rightarrow \infty$  und dominierter Konvergenz folgt also

$$\mathbb{E}_a [u(B(\tau_{K_r}))] = u(a).$$

Der linke Erwartungswert lässt sich als Integral bezüglich des in (10.10) definierten Maßes  $\mu_r$  schreiben und wir erhalten

$$\mathbb{E}_a [u(B(\tau_{K_r}))] = \int_{\partial K_r} u(a+x) \mu_r(dx) = u(a). \quad \square$$

Aus Lemma 10.5 ergibt sich als interessante Folgerung das **harmonische Maximumsprinzip**.

**Theorem 10.6** (Harmonisches Maximumsprinzip). *Sei  $D$  zusammenhängend und beschränkt und sei  $u \in C^2(D) \cap C(\bar{D})$  harmonisch auf  $D$ . Dann nimmt  $u$  das Maximum über  $\bar{D}$  auf  $\partial D$  an.*

*Beweis.* Definiere  $M = \sup_{x \in D} u(x)$  und  $D_M = \{x \in D : u(x) = M\}$ . Wegen der Stetigkeit von  $u$  auf  $\bar{D}$  ist  $D_M$  abgeschlossen in  $D$ . Falls  $D_M$  leer ist, so folgt dass das Maximum von  $u$  auf  $\partial D$  angenommen wird und der Beweis ist fertig. Es sei also  $D_M$  nichtleer und wir wählen  $a \in D_M$  und  $r > 0$  mit  $a + \bar{K}_r \subset D$ . Aus der Mittelwertseigenschaft von  $u$  folgt

$$M = u(a) = \frac{1}{V_r} \int_{K_r} u(a+x) dx.$$

Da  $M$  das Supremum der Funktion  $u$  auf  $D$  ist folgern wir daraus, dass  $u(x) = M$  für alle  $a + K_r$  gilt. Insbesondere muss  $D_M$  also offen in  $D$  sein. Da  $D_M$  gleichzeitig offen und abgeschlossen in  $D$  und ausserdem nichtleer ist, gilt  $D = D_M$ . Wir folgern, dass  $u$  konstant auf  $D$  ist. Wegen der Stetigkeit von  $u$  auf  $\bar{D}$  ist  $u$  also auch auf  $\bar{D}$  konstant und das Maximum wird ebenfalls auf  $\partial D$  angenommen.  $\square$

Nach diesen Überlegungen zu harmonischen Funktionen wenden wir uns dem Dirichlet-Problem zu.

**Definition.** Sei  $D$  eine offene Teilmenge von  $\mathbb{R}^d$  und  $f : \partial D \rightarrow \mathbb{R}$  eine stetige Funktion. Das **Dirichlet-Problem** besteht darin eine Funktion  $u \in C^2(D) \cap C(\bar{D})$  zu finden, welche

$$\begin{aligned} \Delta u(x) &= 0, & x \in D, \\ u(x) &= f(x), & x \in \partial D \end{aligned}$$

erfüllt.

Das Dirichlet-Problem zu lösen bedeutet also eine auf  $D$  harmonische Funktion zu finden, welche sich stetig auf den Rand  $\partial D$  fortsetzen lässt und dort die vorgegebene Randbedingung  $u(x) = f(x)$  erfüllt. Im Unterschied zur Diffusionsgleichung welche wir mit Anfangs- und Endbedingung betrachtet haben, haben wir es nun mit einer partiellen Differentialgleichung mit *Randbedingung* zu tun. In Kapitel 5 haben wir bereits die Lösung eines Dirichlet-Problems konstruiert. Im Beweis der Rekurrenz (in Dimension 2) und Transienz (in Dimension 3 und höher) der Brownschen Bewegung war der entscheidende Schritt eine harmonische Funktion auf einem gegebenen Kreisring (bzw. einer Hohlkugel) zu konstruieren, welche am inneren Rand den Wert 1 und am äusseren

Rand den Wert 0 annimmt. Folgendes Resultat liefert in Analogie zur Feynman-Kac-Formel eine stochastische Darstellung der Lösung des Dirichlet-Problems.

**Theorem 10.7.** *Sei  $f : \partial D \rightarrow \mathbb{R}$  beschränkt, es gelte  $\mathbb{P}_x[\tau_D < \infty] = 1$  für alle  $x \in D$  und es sei  $u$  eine beschränkte Lösung des Dirichlet-Problems. Dann gilt*

$$u(x) = \mathbb{E}_x[f(B_{\tau_D})], \quad x \in D.$$

Falls das Gebiet  $D$  beschränkt ist, so ist die Bedingung  $\mathbb{P}_a[\tau_D < \infty] = 1$  automatisch erfüllt.

*Beweis.* Sei  $K_n$  die Kugel um 0 mit Radius  $n$ . Wir definieren

$$D_n = \left\{ x \in D : \inf_{y \in \partial D} |x - y| > 1/n \right\}.$$

Die Menge  $D_n$  besteht aus allen Punkten in  $D$  die vom Rand  $\partial D$  mindestens den Abstand  $1/n$  haben. Es gilt  $D = \bigcup_{n \in \mathbb{N}} D_n$  und damit  $\tau_{D_n} \rightarrow \tau_D$  fast sicher unter jedem  $\mathbb{P}^x, x \in D$ . Mit der Ito-Formel und der harmonischen Eigenschaft von  $u$  erhalten wir

$$u(B(t \wedge \tau_{K_n} \wedge \tau_{D_n})) = u(B(0)) + \int_0^{t \wedge \tau_{K_n} \wedge \tau_{D_n}} \nabla u(B_s) \cdot dB_s.$$

Wegen  $u \in C^2(D)$  ist  $\nabla u$  beschränkt auf kompakten Teilmengen von  $D$ . Unter dem Erwartungswert folgt also

$$\mathbb{E}_x[u(B(t \wedge \tau_{K_n} \wedge \tau_{D_n}))] = u(x)$$

für alle  $x \in D$ . Im Limes  $t \rightarrow \infty$  erhalten wir mit dominierter Konvergenz

$$\mathbb{E}_x[u(B(\tau_{K_n} \wedge \tau_{D_n}))] = u(x).$$

Nun lassen wir auch noch  $n$  gegen unendlich gehen. Es gilt  $\tau_{K_n} \rightarrow \infty, \tau_{D_n} \rightarrow \tau_D$  und  $\mathbb{P}_x[\tau_D < \infty] = 1$ . Des weiteren wissen wir, dass  $u$  sich stetig auf  $\partial D$  fortsetzen lässt und dort mit  $f$  übereinstimmt. Es folgt also wie behauptet

$$\mathbb{E}_x[f(B(\tau_D))] = u(x). \quad \square$$

## Literatur

Erwin Bolthausen. Skript für Wahrscheinlichkeitstheorie, TU Berlin. 2009.

Paul Glasserman. *Monte Carlo Methods in Financial Engineering*. Springer, 2004.

Olav Kallenberg. *Foundations of Modern Probability*. Springer, 1997.

Achim Klenke. *Wahrscheinlichkeitstheorie*. Springer, 2008.

Peter E. Kloeden and Eckhard Platen. *Numerical Solution of Stochastic Differential Equations*. Springer, 2nd ed. edition, 1995.

J. Michael Steele. *Stochastic Calculus and Financial Applications*. Springer, 2003.