

Stetige Modelle in der stochastischen Finanzmathematik

Frank Oertel

Departement T - Mathematik und Physik

Zürcher Hochschule Winterthur ZHW

CH-8401 Winterthur

frank.oertel@zhwin.ch

WS 03/04

Vorlesungsskript – Grobstruktur in Bearbeitung

1 Modelle und Konzepte der erweiterten Wahrscheinlichkeitsrechnung

- σ -Algebren, Maße, Wahrscheinlichkeitsmaße (vgl. z. B. [Bau], [Bmd], [Bin-Kie], [Hac], [Kre], [Mik] und [Wil]): Hierbei handelt es sich um diejenigen Grundlagen, ohne die ein Verständnis der stochastischen Analysis und der damit verbundenen stetigen Modellbildung in der Finanzmathematik nicht möglich ist. Weshalb also ist dieses Grundgerüst so bedeutsam?

- ↪ Formalisierung der elementargeometrischen "Maßzahlen" Länge, Fläche und Volumen (Addition von Volumina; Berechnung der Fläche eines Kreises durch Approximation mittels n -Eck-Flächen, etc.);
- ↪ Formalisierung der intuitiv erfaßbaren Begriffe "Wahrscheinlichkeit", "Erwartungswert";
- ↪ Banach-Tarski-Paradoxon: Betrachten wir die Oberfläche $S^2 := \{x \in \mathbb{R}^3 : \|x\| = 1\}$ der 3-dimensionalen Einheitskugel, dann läßt sich – unter Voraussetzung des Auswahlaxioms – eine Teilmenge F von S^2 konstruieren, so daß sich für *jede* (!) natürliche Zahl $k \geq 3$ die Oberfläche S^2 als eine disjunkte Vereinigung von k euklidisch-kongruenten Kopien von F darstellen läßt:

$$S^2 \stackrel{!}{=} \bigcup_{i=1}^k \tau_i^{(k)}(F).$$

Dabei beschreibt jede Abbildung $\tau_i^{(k)}$ eine Drehung der Einheitskugel. Schreiben wir nun dem Flächenstück einen bestimmten – anschaulich erfaßbaren – Flächeninhalt $\mu(F) > 0$ zu, der dann natürlich durch eine Drehung nicht verändert wird, dann erhielten wir also wegen der disjunkten Vereinigung:

$$4\pi = k \cdot \mu(F) \quad \forall k = 3, 4, \dots$$

Also würde $\mu(F)$ *simultan* die Werte $4\pi/3, 4\pi/4, \dots$ annehmen, was natürlich nicht möglich sein kann. Demzufolge läßt sich diesem sehr kompliziert aussehendem Flächenstück kein anschaulich erfaßbarer "Flächeninhalt" zuordnen! Welche sinnvollen Struktureigenschaften sollte also ein anschaulich erfaßbarer Flächeninhalt aufweisen?

- ↪ Erinnern wir uns an den bisher verwendeten (Riemannsches) Integral-Begriff, dann ist damit eine direkte Berechnung von elementargeometrischen Flächen und Volumina verbunden. Welche Bedeutung spielt also das Integral im Zusammenhang mit der Berechnung von "Flächen"? Wie läßt sich sinnvoll das Riemann-Integral erweitern?
- ↪ Formalisierung struktureller Eigenschaften von stochastischen Prozessen (bedingter Erwartungswert, Martingalbegriff, etc.) und damit verknüpfter Modellaufbau in der modernen Finanzmathematik.

Um diese Fragen anzugehen, benötigen wir folgende grundlegende

Definition 1.1 (σ -Algebra) Ein System \mathcal{F} von Teilmengen einer Menge Ω heißt eine σ -Algebra (in Ω), wenn es folgende Eigenschaften besitzt:

- (i) $\Omega \in \mathcal{F}$;
- (ii) $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$;
- (iii) Für jede Folge $(A_n)_{n \in \mathbb{N}}$ von Mengen aus \mathcal{F} liegt $\bigcup_{n=1}^{\infty} A_n$ in \mathcal{F} .

Ist \mathcal{F} eine σ -Algebra (in Ω), dann werden die Elemente von \mathcal{F} als \mathcal{F} -meßbare Mengen in Ω bezeichnet.

Beispiel 1.1 Stets ist das System $\mathcal{F} := \{\emptyset, \Omega\}$ eine σ -Algebra in Ω , und zwar die kleinste.

Beispiel 1.2 Stets ist das System $\mathcal{F} := \mathcal{P}(\Omega)(=: 2^\Omega)$ aller Teilmengen von Ω eine σ -Algebra, und zwar die größte.

Beispiel 1.3 Seien Ω und Ω' zwei Mengen, \mathcal{A}' eine σ -Algebra in Ω' und $f : \Omega \rightarrow \Omega'$ eine Abbildung von Ω in Ω' . Dann ist das Mengensystem $\mathcal{F} := \{f^{-1}(B) : B \in \mathcal{A}'\}$ eine σ -Algebra in Ω .

Es ist oft schwierig, auf direktem Wege festzustellen, ob ein vorgegebenes Mengensystem eine σ -Algebra ist. Oftmals jedoch können wir explizit ein "überschaubares" Mengensystem \mathcal{M} angeben, mit dessen Hilfe man daraus eine "große" σ -Algebra konstruieren kann, die \mathcal{M} umfaßt. Sehr wichtig für die gesamte Vorlesung ist daher das folgende Erzeugungsprinzip für σ -Algebren, das darauf beruht, daß der Durchschnitt von σ -Algebren stets wieder eine σ -Algebra ist:

Satz und Definition 1.1 Sei \mathcal{M} ein beliebig vorgegebenes System von Teilmengen von Ω . Dann gibt es eine kleinste σ -Algebra in Ω , die das System \mathcal{M} enthält. Sie heißt die durch \mathcal{M} erzeugte σ -Algebra und wird mit dem Symbol $\sigma(\mathcal{M})$ bezeichnet.

BEWEIS: Setzen wir

$$\mathbb{M} := \{\mathcal{F} : \mathcal{F} \text{ ist eine } \sigma\text{-Algebra mit } \mathcal{M} \subseteq \mathcal{F}\},$$

so ist $\mathbb{M} \neq \emptyset$ (da $\mathcal{P}(\Omega) \in \mathbb{M}$) und also

$$\sigma(\mathcal{M}) := \bigcap_{\mathcal{F} \in \mathbb{M}} \mathcal{F}$$

die gesuchte σ -Algebra. □

Bemerkung 1.1 *Im allgemeinen ist die Vereinigung zweier σ -Algebren in Ω keine σ -Algebra in Ω mehr.*

BEWEIS: Betrachte z. B. für $\Omega := \mathbb{R}$ $\mathcal{F}_1 := \sigma([0, 1])$ und $\mathcal{F}_2 := \sigma((1, 2])$. Dann ist $[0, 2] = [0, 1] \cup (1, 2] \notin \mathcal{F}_1 \cup \mathcal{F}_2$. □

Beispiel 1.4 (Produkt von σ -Algebren) Sei \mathcal{F}_1 eine σ -Algebra in Ω_1 und \mathcal{F}_2 eine σ -Algebra in Ω_2 . Auf dem kartesischen Produkt $\Omega_1 \times \Omega_2$ betrachten wir das folgende Mengensystem:

$$\mathcal{Z} := \{A_1 \times A_2 : A_1 \in \mathcal{F}_1 \text{ und } A_2 \in \mathcal{F}_2\}.$$

Dann wird durch

$$\mathcal{F}_1 \otimes \mathcal{F}_2 := \sigma(\mathcal{Z})$$

eine σ -Algebra in $\Omega_1 \times \Omega_2$ definiert. Sie heißt *Produkt- σ -Algebra* (in $\Omega_1 \times \Omega_2$).

Eines der wichtigsten Beispiele ist die Borelsche σ -Algebra (vgl. [Wil], Beispiel 1.2 auf S. 17 und Beispiel 8.5 auf S. 80), die fundamental für die Erweiterung des Riemannsches Integralbegriffes ist.

Beispiel 1.5 (Borelsche σ -Algebra in \mathbb{R}) Sei $\mathcal{H} := \{(-\infty, x] : x \in \mathbb{R}\}$. Dann heißt

$$\mathcal{B}(\mathbb{R}) := \sigma(\mathcal{H})$$

Borelsche σ -Algebra und deren Elemente die *Borelschen Mengen*¹.

Beispiel 1.6 (Borelsche σ -Algebra in \mathbb{R}^n) Sei $n \in \mathbb{N}$ und \mathcal{M} ein Mengensystem von Teilmengen in \mathbb{R}^n , dessen Elemente sich jeweils als "abgeschlossene Hyperwürfel" $W \in \mathbb{R}^n$ darstellen lassen:

$$\mathcal{M} := \{W : W = [a, b] \times [a, b] \times \cdots \times [a, b]; a < b\}.$$

¹Borel, Émile (1871 – 1956): französischer Mathematiker

Dann heißt

$$\mathcal{B}(\mathbb{R}^n) := \sigma(\mathcal{M})$$

Borelsche σ -Algebra in \mathbb{R}^n und deren Elemente die *Borelschen Mengen* in \mathbb{R}^n . Es läßt sich dann zeigen, daß $\mathcal{B}(\mathbb{R}^n) \stackrel{!}{=} \mathcal{B}(\mathbb{R}) \otimes \dots \otimes \mathcal{B}(\mathbb{R})$.

Eigentlich jede Teilmenge in \mathbb{R}^n , die man üblicherweise verwendet, ist eine Borelsche Menge – auch sehr kompliziert aussehende. Trotzdem läßt sich explizit, jedoch nur auf schwierigem Wege, eine Teilmenge in \mathbb{R}^n konstruieren, die nicht Borelsch ist. Mit anderen Worten: $\mathcal{B}(\mathbb{R}^n) \neq \mathcal{P}(\mathbb{R}^n)$! Um insbesondere also eine sinnvolle Verallgemeinerung des Riemannsches Integralbegriffs bereitstellen zu können, dürfen wir nicht mehr über alle Teilmengen von \mathbb{R} integrieren (vgl mit Theorem 1.1)!

Nun formalisieren wir den anschaulich erfaßbaren Begriff "Volumen" durch Einführung des Maßbegriffes:

Definition 1.2 Sei \mathcal{F} eine σ -Algebra (in Ω). Dann heißt eine Funktion $\mu : \mathcal{F} \longrightarrow [0, \infty]$ ein Maß auf Ω , falls folgende Eigenschaften erfüllt sind:

- (i) $\mu(\emptyset) = 0$;
- (ii) Für jede Folge (A_n) paarweise verschiedener Mengen aus \mathcal{F} gilt die σ -Additivität:

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n).$$

Das Tripel $(\Omega, \mathcal{F}, \mu)$ heißt Maßraum. Ist $\mu(\Omega) < \infty$, dann heißt μ endliches Maß.

Beispiel 1.7 Sei Ω eine beliebige nicht-leere Menge und $\mathcal{F} := \mathcal{P}(\Omega)$. Durch $\mu(A) := \#(A)$ (Anzahl der Elemente von A , falls A aus endlich vielen Elementen besteht, und sonst $+\infty$), $A \in \mathcal{F}$, wird ein Maß auf Ω definiert – das *Zählmaß*.

Beispiel 1.8 (Produkt-Maße) Seien $(\Omega_1, \mathcal{F}_1, \mu_1)$ und $(\Omega_2, \mathcal{F}_2, \mu_2)$ zwei endliche Maßräume. Dann existiert genau ein Maß $\mu_1 \otimes \mu_2$ auf der Produkt- σ -Algebra $\mathcal{F}_1 \otimes \mathcal{F}_2$ in $\Omega_1 \times \Omega_2$, das die folgende Eigenschaft erfüllt:

$$\mu_1 \otimes \mu_2(A_1 \times A_2) = \mu_1(A_1) \cdot \mu_2(A_2) \quad (A_1 \in \mathcal{F}_1, A_2 \in \mathcal{F}_2).$$

Dieses Maß heißt das *Produktmaß* von μ_1 und μ_2 .

Sowohl der Begriff der Produkt- σ -Algebra als auch der Begriff des Produktmaßes können direkt auf $n > 2$ "Faktoren" übertragen werden.

Ohne Beweis notieren wir die folgende nicht-triviale Aussage:

Theorem 1.1 Auf der Borelschen σ -Algebra $\mathcal{B}(\mathbb{R}^n)$ gibt es genau ein Maß

$$\lambda_n : \mathcal{B}(\mathbb{R}^n) \longrightarrow [0, \infty],$$

das jedem "halboffenen Hyperwürfel" in \mathbb{R}^n dessen n -dimensionalen Elementarinhalt zuordnet:

$$\lambda_n\left([a_1, b_1) \times [a_2, b_2) \times \cdots \times [a_n, b_n)\right) = \prod_{i=1}^n (b_i - a_i).$$

Dieses eindeutig bestimmte Maß λ_n heißt das Lebesgue-Maß² auf \mathbb{R}^n .

Anschaulich ist es klar, daß man einer einpunktigen Menge $\{a\}$ in \mathbb{R} keinen positiven Inhalt zuordnen kann. Für das Lebesgue-Maß λ_1 läßt sich diese Eigenschaft beweisen. Es gilt nämlich für alle $a, b \in \mathbb{R}$ mit $a \leq b$:

$$\lambda_1\left([a, b)\right) = \lambda_1\left([a, b]\right) = \lambda_1\left((a, b)\right) = \lambda_1\left((a, b]\right) = b - a.$$

Mit anderen Worten: $\lambda_1(\{a\}) = 0 \quad \forall a \in \mathbb{R}$.

Damit gelangen wir zu einem weiteren zentralen Begriff:

Definition 1.3 Sei $(\Omega, \mathcal{F}, \mu)$ ein Maßraum. Eine Teilmenge N von Ω heißt $(\mu-)$ Nullmenge, falls $N \in \mathcal{F}$ und $\mu(N) = 0$ ist.

Offensichtlich sind abzählbare Vereinigungen von Nullmengen und Teilmengen von Nullmengen wieder Nullmengen (man verwende dazu die geeigneten Maßeigenschaften).

Ist $A = A(\omega)$ eine Aussage, die von $\omega \in \Omega$ abhängt, so sagen wir, daß A μ -fast überall (μ -f.ü.) zutrifft, falls die Menge

$$N := \{\omega \in \Omega : A(\omega) \text{ trifft nicht zu}\}$$

eine μ -Nullmenge ist.

Kehren wir zurück zur allgemeinen Definition eines Maßes und setzen nun noch *zusätzlich* voraus, daß $\mu(\Omega) = 1$ ist, so erhalten wir damit die fundamentale Definition des Wahrscheinlichkeitsmaßes:

Definition 1.4 Ein endliches Maß $\mathbb{P} : \mathcal{F} \longrightarrow [0, 1]$ auf einem Maßraum $(\Omega, \mathcal{F}, \mathbb{P})$ heißt Wahrscheinlichkeitsmaß, falls $\mathbb{P}(\Omega) = 1$. In diesem Falle heißt das Tripel $(\Omega, \mathcal{F}, \mathbb{P})$ Wahrscheinlichkeitsraum.

Wie ist ein Wahrscheinlichkeitsraum zu interpretieren? Durch einen Wahrscheinlichkeitsraum werden mathematisch alle wichtigen Bestandteile eines *Zufallsexperimentes* (wie beispielsweise der Wurf mit einem Würfel) erfaßt. Dabei beschreibt die *Stichprobenmenge* Ω die Gesamtheit aller denkbar möglichen Ergebnisse, die durch dieses Zufallsexperiment erzielt werden kann. Bei obigem Würfelwurf wäre $\Omega := \{1, 2, 3, 4, 5, 6\}$. Man bezeichnet häufig $\omega \in \Omega$ auch als *Ergebnis* bzw. *Resultat*. \mathcal{F} ist ein System aus

²Lebesgue, Henri (1875 – 1941): französischer Mathematiker

bestimmten Mengen. Jede solche Menge $A \in \mathcal{F}$, setzt sich dabei aus einer bestimmten Gruppe von Ergebnissen des Experimentes zusammen und wird daher als *Ereignismenge* bzw. kürzer als *Ereignis* bezeichnet. \mathcal{F} umfaßt also alle Ereignisse, die (sinnvollerweise) die Eigenschaften (i), (ii) und (iii) der Definition 1.1 erfüllen. Beim Würfelexperiment käme jede Teilmenge $A \subseteq \Omega$ als Ereignis in Frage, so daß wir also hier $\mathcal{F} := \mathcal{P}(\Omega)$ setzen. Das Wahrscheinlichkeitsmaß \mathbb{P} liefert nun diejenige Maßzahl $0 \leq \mathbb{P}(A) \leq 1$, die angibt, mit welcher "relativen Häufigkeit" das Ereignis A auftreten kann und das deshalb (sinnvollerweise) die Eigenschaften (i) und (ii) der Definition 1.2 erfüllen sollte. Beim Würfelexperiment hätten wir

$$\mathbb{P}(A) := \frac{\#(A)}{\#(\Omega)} = \frac{\#(A)}{6} \quad (A \in \mathcal{F}).$$

Proposition 1.1 *Sei $(\Omega, \mathcal{F}, \mathbb{P})$ ein Wahrscheinlichkeitsraum. Dann besitzt das Wahrscheinlichkeitsmaß \mathbb{P} die folgenden Eigenschaften:*

- Für disjunkte Mengen $A, B \in \mathcal{F}$ gilt: $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$;
- Seien $A, B \in \mathcal{F}$ mit $A \subseteq B$. Dann gilt $\mathbb{P}(B \setminus A) = \mathbb{P}(B) - \mathbb{P}(A)$;
- Seien $A, B \in \mathcal{F}$ mit $A \subseteq B$. Dann gilt $\mathbb{P}(A) \leq \mathbb{P}(B)$;
- $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$ ($A, B \in \mathcal{F}$).

BEWEIS: Übung (vgl. Übungsblatt 1). □

...TBC

- *Zufallsvariablen und ihre Verteilungen* (S. 9 – 12, 85 – 88, 96 – 98 von [Bmd], S. 138 – 139 von [Kre] und S. 14 – 17 von [Mik]): Zufallsvariable, Funktionen von Zufallsvariablen (Summen, Produkte, Maximum, Minimum, etc.), Zufallsvektor(!), kumulative Verteilungsfunktion, W’dichte, marginale Dichte, normalverteilte Zufallsvariable
- *Erwartungswert, Varianz und Kovarianz* (S. 50, 88 – 90, 98 – 101, 110 – 112 von [Bmd] und S. 12 – 13, 17 – 18 von [Mik]) : \mathbb{E} für diskrete ZVen, \mathbb{E} für ZVen mit Dichte, \mathbb{E} für allgemeine ZVen (vgl. dazu mit S. 110 – 112 von [Bmd]) und S. 38 – 41 von [Bin-Kie]), Momente, Varianz, Kovarianz, Kovarianzmatrix
- *Abhängigkeit, bedingte Wahrscheinlichkeiten und Unabhängigkeit* (S. 12 – 17, 105 – 108 von [Bmd], S. 21 – 27, 140 – 142 von [Kre] und S. 19 – 23 von [Mik]): unabhängige Ereignisse, unabhängige ZVen (man beachte hier den sehr wichtigen Satz 1.14 in [Sch] auf S. 86 bzw. die Betrachtung am Ende von Kapitel 3.2 in [Bmd] auf S. 109), Korrelation und Abhängigkeit

2 Gauß’sche Zufallsvektoren

Um den Begriff des Gauß’schen Zufallsvektors richtig zu erfassen, rufen wir uns den folgenden Begriff aus der linearen Algebra in Erinnerung:

Definition 2.1 Sei $A \in M(n \times n; \mathbb{R})$ eine symmetrische $n \times n$ -Matrix. A heißt positiv definit (symbolisch: $A > 0$), falls

$$\langle x, Ax \rangle > 0 \quad \forall x \in \mathbb{R}^n \setminus \{0\}.$$

Ohne Beweis zitieren wir den fundamentalen Spektraldarstellungssatz für symmetrische Matrizen:

Theorem 2.1 Sei $A \in M(n \times n; \mathbb{R})$ eine symmetrische $n \times n$ -Matrix mit den (reellen) Eigenwerten $\lambda_1, \dots, \lambda_n$. Sei $\Delta := \text{diag}(\lambda_1, \dots, \lambda_n)$. Dann existiert eine orthogonale Matrix $S \in O(n \times n; \mathbb{R})$, so daß

$$A = S\Delta S^\top.$$

Ist A zusätzlich positiv definit, dann sind sämtliche Eigenwerte von A strikt positiv: $\lambda_i > 0 \quad \forall i = 1, \dots, n$.

Hieraus erkennen wir unmittelbar das folgende

Korollar 2.1 Sei $A > 0$ eine beliebige positiv definite Matrix. Dann gilt:

- (1) A ist invertierbar – mit $\det(A) > 0$;
- (2) Es gibt eine symmetrische Matrix $B \in M(n \times n; \mathbb{R})$, so dass $A = B^2$.

BEWEIS: Aussage (1) folgt direkt durch Anwendung des Determinantenmultiplikationssatzes auf die obige Spektraldarstellung von A . Aussage (2) erhalten wir folgendermaßen: Setzen wir $B := S\Delta^{1/2}S^\top$, wobei $\Delta^{1/2} := \text{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n})$, dann ergibt sich aus dem obigen Spektraldarstellungssatz:

$$A = S\Delta S^\top = S\Delta^{1/2}\Delta^{1/2}S^\top = S\Delta^{1/2}S^\top S\Delta^{1/2}S^\top = B^2,$$

da $S^\top S = S^{-1}S = E_n$ (aufgrund der Orthogonalität von S). □

Wir führen nun Gauß'sche Zufallsvektoren (GZVRen) $\underline{X} \sim N_n(\mu, C)$ mittels der von Mikosch benutzten Definition ein (s. [Mik], Beispiel 1.1.8 auf S. 17):

Definition 2.2 Sei $\underline{X} = (X_1, \dots, X_n)^\top : \Omega \longrightarrow \mathbb{R}^n$ ein n -dimensionaler Zufallsvektor. \underline{X} heißt n -dimensionaler Gauß'scher Zufallsvektor (bzw. n -dimensionaler multivariat-normalverteilter Zufallsvektor), falls \underline{X} eine multivariate Dichte $f_{\underline{X}}$ mit der folgenden Struktur besitzt:

$$f_{\underline{X}}(\xi) := (2\pi)^{-n/2} (\det(C))^{-1/2} \exp\left(-\frac{1}{2} \langle \xi - \mu, C^{-1}(\xi - \mu) \rangle\right) \quad (\xi \in \mathbb{R}^n),$$

mit $\mu \in \mathbb{R}^n$ und positiv definiten $n \times n$ -Matrix $C > 0$. Wir verwenden folgende Kurzschreibweise: $\underline{X} \sim N_n(\mu, C)$.

Wiederum ohne Beweis³ notieren wir die folgende sehr wichtige und nicht-triviale Aussage, deren vollständiger Beweis die Verwendung charakteristischer Funktionen von Zufallsvektoren (also Fourier-Transformierte von Bildmaßen) benötigt⁴:

³Die Untersuchung Gauß'scher Zufallsvektoren beansprucht eine eigene Theorie, die hier natürlich nicht im Detail erläutert werden kann (vgl. auch [Mar-Ken-Bib], S. 36 – 38) und [Mey], S. 106 – 107!

⁴Vereinbarung: Alle in dieser Vorlesung verwendeten Vektoren werden als Spaltenvektoren aufgefaßt und als solche gekennzeichnet

Theorem 2.2 Seien $\mu \in \mathbb{R}^n$ und $C \in M(n \times n; \mathbb{R})$ eine positiv definite Matrix. Sei $\underline{X} = (X_1, \dots, X_n)^\top : \Omega \longrightarrow \mathbb{R}^n$ ein n -dimensionaler Zufallsvektor. Dann sind folgende Aussagen äquivalent:

- (i) $\underline{X} \sim N_n(\mu, C)$
- (ii) $\langle a, \underline{X} \rangle := \sum_{i=1}^n a_i X_i \sim N(\langle a, \mu \rangle, \langle a, Ca \rangle)$ für alle $a \in \mathbb{R}^n$.

Korollar 2.2 Seien $\mu = (\mu_1, \dots, \mu_n)^\top \in \mathbb{R}^n$ und $C = (c_{ij}) \in M(n \times n; \mathbb{R})$ eine positiv definite Matrix. Ist $\underline{X} = (X_1, \dots, X_n)^\top \sim N_n(\mu, C)$ ein n -dimensionaler GZVR, so ist jede Komponente X_i von \underline{X} normalverteilt mit $X_i \sim N(\mu_i, c_{ii})$, $\mu_i = \mathbb{E}(X_i)$ und $\text{Cov}(X_i, X_j) = c_{ij}$ für jedes $i, j \in \{1, 2, \dots, n\}$.

BEWEIS: Anwendung des Theorems auf alle Einheitsvektoren $a = e_i$ liefert: $X_i = \langle e_i, \underline{X} \rangle \sim N(\langle e_i, \mu \rangle, \langle e_i, Ce_i \rangle)$. Also folgt $X_i \sim N(\mu_i, c_{ii})$ und insbesondere $\mu_i = \mathbb{E}(X_i)$. Sei $\Sigma_{ij} := \text{Cov}(X_i, X_j)$. Zur Berechnung von C beachten wir zunächst, daß $\langle a, Ca \rangle = \text{Var}(\sum_{i=1}^n a_i X_i)$ für jedes $a \in \mathbb{R}^n$ gilt (wegen Theorem 2.2) und somit für jedes $a \in \mathbb{R}^n$ folgende Gleichheit erfüllt ist: $\langle a, \Sigma a \rangle = \text{Var}(\sum_{i=1}^n a_i X_i) = \langle a, Ca \rangle$. Insbesondere gilt $\langle a, \Theta a \rangle = 0$ für alle $a \in \mathbb{R}^n$, wobei $\Theta := \Sigma - C$. Mit $\varphi : \mathbb{R}^n \longrightarrow \mathbb{R}$, $a \mapsto \langle a, \Theta a \rangle$ ergibt sich also $\varphi \equiv 0$ (auf ganz \mathbb{R}^n). Differentiation liefert $\nabla \varphi(a) = 2\Theta a = 0 \forall a \in \mathbb{R}^n$ (φ ist ja stetig partiell differenzierbar), und wir erhalten somit $\Theta = 0$. \square

Satz und Definition 2.1 Sei $\underline{X} = (X_1, \dots, X_n)^\top : \Omega \longrightarrow \mathbb{R}^n$ ein n -dimensionaler Zufallsvektor mit $\mathbb{E}^2(X_i) < \infty$ für jedes $i \in \{1, 2, \dots, n\}$. Dann existieren der Erwartungswertvektor $\mathbb{E}(\underline{X}) := (\mathbb{E}(X_1), \dots, \mathbb{E}(X_n))^\top$ von \underline{X} und die Kovarianzmatrix $\Sigma(\underline{X}) := (\text{Cov}(X_i, X_j))_{ij}$ von \underline{X} . Ist $\underline{X} \sim N_n(\mu, C)$ sogar ein n -dimensionaler GZVR mit positiv definiten Matrix $C \in M(n \times n; \mathbb{R})$ und $\mu \in \mathbb{R}^n$, dann gilt: $\mathbb{E}(\underline{X}) = \mu$ und $\Sigma(\underline{X}) = C$.

Sehr wichtig (insbesondere zur Untersuchung von Eigenschaften der Brownschen Bewegung) ist der folgende Transformationssatz für GZVRen:

Theorem 2.3 Seien $\mu \in \mathbb{R}^n$ und $C = (c_{ij}) \in M(n \times n; \mathbb{R})$ eine positiv definite Matrix und $\underline{X} \sim N_n(\mu, C)$ ein n -dimensionaler GZVR. Sei $k \in \mathbb{N}$ beliebig gewählt. Dann ist für jedes $b \in \mathbb{R}^k$ und für jede Matrix $A \in M(k \times n; \mathbb{R})$ der k -dimensionale Zufallsvektor $b + A\underline{X}$ ebenfalls ein GZVR mit:

$$\mathbb{E}(b + A\underline{X}) = b + A\mathbb{E}(\underline{X}) = b + A\mu$$

und

$$\Sigma(b + A\underline{X}) = A\Sigma(\underline{X})A^\top = ACA^\top.$$

BEWEIS: Sei $a \in \mathbb{R}^k$ ein beliebig gewählter Vektor. Wir setzen $\underline{Y} := b + A\underline{X}$. Dann gilt:

$$\langle a, \underline{Y} \rangle = \langle a, b \rangle + \langle a, A\underline{X} \rangle = \langle a, b \rangle + \langle A^\top a, \underline{X} \rangle.$$

Setzen wir $\tilde{a} := A^\top a$, $\tilde{\mu} := \langle a, b \rangle$ und $\tilde{Z} := \langle A^\top a, \underline{X} \rangle$, so erhalten wir zunächst $\tilde{Z} \sim N(\langle \tilde{a}, \mu \rangle, \langle \tilde{a}, C\tilde{a} \rangle)$ (wegen Theorem 2.2) und also $\langle a, \underline{Y} \rangle = \tilde{\mu} + \tilde{Z} \sim N(\tilde{\mu} + \langle \tilde{a}, \mu \rangle, \langle \tilde{a}, C\tilde{a} \rangle)$. Mittels linearer Algebra erhalten wir nun

$$\tilde{\mu} + \langle \tilde{a}, \mu \rangle = \langle a, b \rangle + \langle A^\top a, \mu \rangle = \langle a, b + A\mu \rangle$$

und

$$\langle \tilde{a}, C\tilde{a} \rangle = \langle A^\top a, CA^\top a \rangle = \langle a, (ACA^\top)a \rangle.$$

Theorem 2.2 beendet damit den Beweis. \square

Theorem 2.4 Sei $\underline{X} = (X_1, \dots, X_n)^\top : \Omega \rightarrow \mathbb{R}^n$ ein n -dimensionaler Zufallsvektor, und seien $\sigma_i > 0$ und $\mu_i \in \mathbb{R}$ gegeben ($i \in \{1, \dots, n\}$). Dann sind folgende Aussagen äquivalent:

- (i) Die Zufallsvariablen X_1, \dots, X_n sind unabhängig, und für jedes $i \in \{1, \dots, n\}$ ist X_i eine normalverteilte ZV mit $X_i \sim N(\mu_i, \sigma_i^2)$;
- (ii) \underline{X} ist ein n -dimensionaler GZVR mit $\mathbb{E}(\underline{X}) = (\mu_1, \dots, \mu_n)^\top$ und $\Sigma(\underline{X}) = \text{diag}(\sigma_1^2, \dots, \sigma_n^2)$;
- (iii) $\langle a, \underline{X} \rangle = \sum_{i=1}^n a_i X_i \sim N(\langle a, \mu \rangle, \sum_{i=1}^n \sigma_i^2 a_i^2)$ für alle $a \in \mathbb{R}^n$.

BEWEIS: "(i) \Rightarrow (ii)": Es gelte also (i). Wir setzen $\mu = (\mu_1, \dots, \mu_n)^\top$ und $C := \text{diag}(\sigma_1^2, \dots, \sigma_n^2)$. Unter Beachtung der vorausgesetzten Unabhängigkeit der normalverteilten Zufallsvariablen X_1, \dots, X_n , erhalten wir gemäß Theorem 1.4 nach einer kleinen, jedoch einfachen, Rechnung die folgende Dichte:

$$f_{\underline{X}}(\xi) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left(-\frac{(\xi_i - \mu_i)^2}{2\sigma_i^2}\right) = (2\pi)^{-n/2} (\det(C))^{-1/2} \exp\left(-\frac{1}{2} \langle \xi - \mu, C^{-1}(\xi - \mu) \rangle\right). \quad (1)$$

für alle $\xi = (\xi_1, \dots, \xi_n)^\top \in \mathbb{R}^n$.

Also folgt (nach Definition eines GZVRs) $\underline{X} \sim N_n(\mu, C)$ und damit Aussage (ii).

"(ii) \Rightarrow (i)": Zunächst folgt sofort aus Korollar 2.2, daß jede der Komponenten des GZVRs X_i eine normalverteilte ZV ist mit $X_i \sim N(\mu_i, \sigma_i^2)$ und somit die folgende Dichte besitzt:

$$f_{X_i}(u) := \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left(-\frac{(u - \mu_i)^2}{2\sigma_i^2}\right) \quad (u \in \mathbb{R}).$$

Nach Voraussetzung besitzt nun die Dichte des GZVRs \underline{X} die Darstellung (1), wobei $C := \text{diag}(\sigma_1^2, \dots, \sigma_n^2)$.

Also erhalten wir für alle $(\xi_1, \dots, \xi_n)^\top \in \mathbb{R}^n$:

$$f_{\underline{X}}((\xi_1, \dots, \xi_n)^\top) = \prod_{i=1}^n f_{X_i}(\xi_i)$$

und damit (nach Theorem 1.4) die Unabhängigkeit der normalverteilten Zufallsvariablen X_1, \dots, X_n .

"(ii) \iff (iii)": Diese Äquivalenz folgt sofort aus Theorem 2.2. □

Hieraus erhalten wir unmittelbar zwei wichtige Folgerungen:

Korollar 2.3 Sei $\underline{X} = (X_1, \dots, X_n)^\top : \Omega \longrightarrow \mathbb{R}^n$ ein n -dimensionaler Gauß'scher Zufallsvektor. Dann sind folgende Aussagen äquivalent:

- (i) Die ZVen X_1, \dots, X_n sind paarweise unkorreliert: $\text{Cov}(X_i, X_j) = 0 \forall i \neq j$.
- (ii) Die ZVen X_1, \dots, X_n sind unabhängig.

Diese Aussage bedeutet jedoch *nicht*, daß paarweise unkorrelierte und normalverteilte Zufallsvariablen X_1, \dots, X_n notwendigerweise schon unabhängig sind! Diese Aussage ist *falsch*, wie das folgende Gegenbeispiel beweist (vgl. Übung E4 auf S. 134 von [Bmd]):

Beispiel 2.1 Seien Y und U unabhängige Zufallsvariablen, so daß $Y \sim N(0, 1)$ und $\mathbb{P}(U = 1) = 1/2 = \mathbb{P}(U = -1)$. Setzen wir $Z := UY$, dann sind die ZVen $Y \sim N(0, 1)$ und $Z \sim N(0, 1)$ abhängig und jeweils standard-normalverteilt. Jedoch sind die ZVen Y, U, Z paarweise unkorreliert.

Korollar 2.4 Seien X_1, \dots, X_n unabhängige ZVen. Es gelte

$$X_i \sim N(\mu_i, \sigma_i^2) \quad \forall i \in \{1, \dots, n\},$$

wobei $\mu_i \in \mathbb{R}$ und $\sigma_i > 0$ ($i \in \{1, \dots, n\}$). Dann ist auch die Summe $\sum_{i=1}^n X_i$ normalverteilt mit $\sum_{i=1}^n X_i \sim N(\mu, \sigma^2)$, wobei $\mu := \sum_{i=1}^n \mu_i$ und $\sigma := \sqrt{\sum_{i=1}^n \sigma_i^2}$.

Mit diesem (erweiterten) Abschnitt über Gauß'sche Zufallsvektoren haben wir nun eines der wichtigsten Hilfsmittel bereitgestellt, um später Mikoschs elegante Vorgehensweise zur Untersuchung der Brownschen Bewegung vollständig übernehmen zu können. Doch zunächst benötigen wir den allgemeinen Rahmen.

3 Stochastische Prozesse

- *Allgemeine stochastische Prozesse* (S. 23 – 32 von [Mik]) : Definition eines stochastischen Prozesses, Verteilung eines stochastischen Prozesses (FDF), Gauß'scher Prozess, Erwartungswert-Funktion und Kovarianz-Funktion, strikte Stationarität, stationäre Zuwächse, unabhängige Zuwächse, Homogener Poisson Prozess

- *Brownsche Bewegung* (S. 33 – 44 von [Mik]) : Definition und charakteristische Eigenschaften der standardisierten Brownschen Bewegung (SBM)⁵, SBM B besitzt beschränkte quadratische Variation $[B]_t = t$ \mathbb{P} -f. s. und unbeschränkte Variation (s. S. 97 – 98 und S. 189 – 190 von [Mik] und S. 18 – 20 von [Ptt]), Nicht-Differenzierbarkeit von SBM, \rightsquigarrow Anwendung des klassischen (pfadweise definierten) Riemann–Stieltjes-Integrals mit SBM als Integrator ist nicht möglich! \rightsquigarrow stochastische Analysis; abgeleitete Prozesse aus der SBM, Brownsche Bewegung mit Drift, geometrische Brownsche Bewegung
- *Simulation der Pfade einer Brownschen Bewegung* (Als Folie projizieren: S. 44 – 52 von [Mik])
- *Der bedingte Erwartungswert für diskrete Zufallsvariablen* (S. 56 – 62 von [Mik]) : Hier können wir uns wieder (fast) vollständig an das Buch [Mik] halten, wobei wir die Abschnitte über σ -Algebren und Borelmengen auf den Seiten 62 – 65 nicht mehr behandeln müssen, da sie ja bereits zu Beginn dieser Vorlesung eingeführt wurden. Neu jedoch, ist der Begriff der durch eine ZVe generierten σ -Algebra. Da wir folgende Aussage noch mehrmals benötigen, formulieren wir sie als

Lemma 3.1 *Sei $X : \Omega \longrightarrow \mathbb{R}$ eine diskrete Zufallsvariable, die die reellen Werte x_1, x_2, \dots annimmt, derart daß $\mathbb{E}(|X|) < \infty$. Sei $B \subseteq \Omega$. Dann gilt:*

$$\mathbb{E}(X \mathbf{1}_B) = \sum_{k=1}^{\infty} x_k \mathbb{P}(\{X = x_k\} \cap B).$$

Insbesondere gilt:

$$\mathbb{E}(X|B) = \sum_{k=1}^{\infty} x_k \mathbb{P}(X = x_k|B).$$

BEWEIS: Setzen wir $Z := X \mathbf{1}_B$, dann gilt zunächst:

$$\mathbb{E}(Z) = \sum_{k=1}^{\infty} x_k \mathbb{P}(Z = x_k). \quad (2)$$

Somit müssen wir alle W'keiten $\mathbb{P}(Z = x_k)$ berechnen. Da für jedes $k \in \mathbb{N}$ $\{Z = x_k\} = (\{X = x_k\} \cap B) \cup (\{0 = x_k\} \cap B^c)$, folgt also, daß $\mathbb{P}(Z = x_k) = \mathbb{P}(\{X = x_k\} \cap B) + \mathbb{P}(\{0 = x_k\} \cap B^c)$. Beachten wir nun noch, daß $x_k \cdot \mathbb{P}(\{0 = x_k\} \cap B^c) = 0$ für jedes $k \in \mathbb{N}$ gilt, so folgt aus (2)

$$\mathbb{E}(Z) = \sum_{k=1}^{\infty} x_k \mathbb{P}(\{X = x_k\} \cap B) + 0$$

⁵Die Definition der SBM beinhaltet nicht den Bezug auf eine beliebig vorgegebene Informationsstruktur (Filtration). Jedoch ist die SBM B ein Wiener-Prozess bez. der durch die SBM erzeugten kanonischen Filtration \mathbb{F}^B (vgl. Kapitel 1.5 von [Nls], S. 17 – 19).

und damit die Behauptung. \square

Obwohl die Darstellung von $\mathbb{E}(X|B)$ für ZVen X mit Dichte f_X auf S. 58 im allgemeinen so nicht geschrieben werden kann, können wir trotzdem das Beispiel 1.4.2 auf der Seite 58 verwenden, da ja dort $X(\omega) := \omega$ mit $\omega \in (0, 1]$ gesetzt wurde und somit $X(\omega)\mathbf{1}_{A_i}(\omega)$ als $X(\omega)\mathbf{1}_{A_i}(\omega) = (\psi \circ X)(\omega)$ geschrieben werden kann, wobei $\psi(x) := x \cdot \mathbf{1}_{A_i}(x)$ ($x \in \mathbb{R}$). Somit ergibt die Transformationsformel:

$$\mathbb{E}(X\mathbf{1}_{A_i}) = \int_{\Omega} (\psi \circ X)(\omega) d\mathbb{P}(\omega) = \int_{-\infty}^{\infty} \psi(x) f_X(x) dx = \int_{A_i} x f_X(x) dx.$$

Da wir bereits den Meßbarkeitsbegriff zu Beginn dieser Vorlesung behandelt haben, können wir nach Einführung der erzeugten σ -Algebra vom Typ $\sigma(X_1, X_2, \dots, X_n)$ (als der kleinsten σ -Algebra, bezüglich der jede Funktion X_i meßbar ist⁶) an dieser Stelle die folgende Aussage formulieren:

Bemerkung 3.1 Seien n Funktionen $X_i : \Omega \rightarrow \mathbb{R}$ von einer Menge Ω nach \mathbb{R} gegeben ($i = 1, 2, \dots, n$). Sei \mathcal{F} eine σ -Algebra in Ω . Dann sind folgende Aussagen äquivalent:

- (i) Jede der Funktionen X_i ist \mathcal{F} -meßbar.
- (ii) Die Abbildung $(X_1, \dots, X_n)^\top : \Omega \rightarrow \mathbb{R}^n$ ist $\mathcal{F} - \mathcal{B}(\mathbb{R}^n)$ -meßbar.
- (iii) $\sigma(X_1, X_2, \dots, X_n) \subseteq \mathcal{F}$. \square

BEWEISSKIZZE Für $n = 1$, folgt die Aussage direkt aus der Darstellung von $\sigma(X) = \{X^{-1}(B) : B \in \mathcal{B}(\mathbb{R})\}$. Durch die Darstellung aller Borel-Mengen in \mathbb{R}^n als n -faches kartesisches Produkt $\mathcal{B}(\mathbb{R}^n) = \mathcal{B}(\mathbb{R}) \times \dots \times \mathcal{B}(\mathbb{R})$ und durch die Darstellung von $\sigma(X_1, X_2, \dots, X_n)$ als $\sigma(X_1, X_2, \dots, X_n) = \sigma(\bigcup_{i=1}^n \sigma(X_i))$, können wir die Aussage für den mehrdimensionalen Fall $n > 1$ auf den eindimensionalen Fall zurückführen (s. [Wil], A3.2., Seite 207). $\dots \square$

Einen weiteren wichtigen und nicht-trivialen Satz, der sich als sehr nützlich zur konkreten Berechnung bedingter Erwartungswerte herausstellen wird, ist der folgende (vgl. dazu die Ausführungen auf den Seiten 206 – 207 von [Wil]):

Theorem 3.1 Seien n Funktionen $Y_i : \Omega \rightarrow \mathbb{R}$ von einer Menge Ω nach \mathbb{R} gegeben ($i = 1, 2, \dots, n$). Eine Funktion $X : \Omega \rightarrow \mathbb{R}$ ist genau dann $\sigma(Y_1, Y_2, \dots, Y_n)$ -meßbar, wenn es eine meßbare Funktion $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}$ gibt, so daß $X = \varphi \circ (Y_1, \dots, Y_n)$.

BEWEISSKIZZE FÜR $n = 1$: Wir geben hier nur eine Beweisskizze für die nicht-triviale Implikation " \implies " an. Zunächst beachten wir, daß für jede Teilmenge $B \subseteq \Omega$ gilt: (vgl. Übungsaufgabe):

$$\mathbf{1}_B : (\Omega, \mathcal{F}) \rightarrow \mathbb{R} \text{ ist } \mathcal{F}\text{-meßbar} \iff B \in \mathcal{F}.$$

⁶Formal erkennen wir, daß $\sigma(X_1, X_2, \dots, X_n) = \sigma(\bigcup_{i=1}^n \sigma(X_i))$.

Sei $C \in \sigma(Y)$ beliebig gewählt. Dann gibt es eine Borel-Menge B , so daß $C = Y^{-1}(B)$. Damit ist $X := \mathbf{1}_C = \mathbf{1}_B \circ Y$ $\sigma(Y)$ -meßbar, und X läßt sich schreiben als $X = \varphi \circ Y$ mit der meßbaren Funktion $\varphi := \mathbf{1}_B$. Durch Bilden aller möglichen Linearkombinationen $\sum \alpha_i \mathbf{1}_{C_i}$ solcher charakteristischer Funktionen und anschließendem Grenzübergang erhalten wir dann die gewünschte Behauptung (vgl. [Bau], Faktorisierungslemma 11.7 auf S. 71). ... \square

- *Der bedingte Erwartungswert für allgemeine Zufallsvariablen* (S. 67 – 76 von [Mik] und S. 84 – 92 von [Wil]) Hier übernehmen wir **nicht** Mikoschs Definition auf S. 68, da diese zu Mißverständnissen führen kann: Was bedeutet hier die genaue Bedeutung des Begriffes "Zufallsvariable"? Meßbarkeit? Was genau ist hier \mathcal{F} ? Wieso übernimmt Mikosch nicht die vollständige Definition? Deshalb verwenden wir nun zur Einführung des allgemeinen bedingten Erwartungswertes die übliche Definition und (voraussichtlich ohne Beweis) das Existenztheorem von Kolmogorov ⁷ (s. [Wil], S. 84 – 87). Die Beispiele 1.4.11 auf Seite 68 – 69 von [Mik] sind sehr suggestiv, die wir wieder in unsere Vorlesung integrieren. Zusätzlich dazu berechnen wir den bedingten Erwartungswert $\mathbb{E}(X|\mathcal{B})$ für $\mathcal{B} := \{\emptyset, \Omega\}$:

Übungsaufgabe 3.1 Sei $X : \Omega \longrightarrow \mathbb{R}$ eine Zufallsvariable und $\mathcal{B} := \{\emptyset, \Omega\}$. Dann gilt:

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

LÖSUNG: Setze $Z := \mathbb{E}(X|\mathcal{B})$. Nach Definition des bedingten Erwartungswertes, müssen folgende Eigenschaften erfüllt sein

$$(a) \sigma(Z) \subseteq \mathcal{B}$$

und

$$(b) \mathbb{E}(Z\mathbf{1}_B) = \mathbb{E}(X\mathbf{1}_B) \quad \forall B \in \mathcal{B}.$$

Wegen (a) gilt insbesondere: $\{Z = a\} \in \mathcal{B} = \{\emptyset, \Omega\} \quad \forall a \in \mathbb{R}$. Wir wählen nun ein festes $\omega_0 \in \Omega$ und setzen $a_0 := Z(\omega_0)$. Dann gilt $\{Z = a_0\} \neq \emptyset$ (da ja $\omega_0 \in \{Z = a_0\}$) und somit $\{Z = a_0\} = \Omega$. Also ist $Z(\omega) = a_0 = \text{const.} \quad \forall \omega \in \Omega$. Wenden wir nun (b) auf die Menge $B := \Omega$ an, erhalten wir somit die Behauptung. \square

Bedeutung des bedingten Erwartungswertes (Interpretation): "Wächst die Information \mathcal{B} , dann erhalten wir dabei auch eine bessere Information über die Zufallsvariable X in $\mathbb{E}(X|\mathcal{B})$."

Anschließend daran formulieren wir die bekannten und wichtigen Rechenregeln für bedingte Erwartungswerte, wobei wir uns durchaus wieder an den Ausführungen in [Mik] orientieren können, obwohl diese etwas knapp formuliert sind. Einen präzisen und trotzdem verständlichen Zugang erhalten wir mittels [Wil], S. 88 – 92. Für die Bewertung von Optionen sehr wichtig ist dabei das folgende Theorem (vgl. [Lam-Lap], proposition A.2.5 auf S. 177, [Sch], Satz 4.5.1 auf S. 169 und [Wil], Abschnitt 9.10. auf S. 91 – 92), das wir ohne Beweis angeben:

⁷Interessant an Williams' Existenz-Beweis ist, daß er dazu nicht den Satz von Radon-Nikodym benötigt, sondern "nur" mittels der Fortsetzung von L^2 auf L^1 arbeitet und dabei die geometrische Struktur von L^2 benutzt - mit dem Skalarprodukt $\langle X, Y \rangle := \mathbb{E}(XY)$.

Theorem 3.2 Seien $(\Omega, \mathcal{F}, \mathbb{P})$ ein beliebiger Wahrscheinlichkeitsraum, \mathcal{B} eine Teil- σ -Algebra von \mathcal{F} und $X, Y : \Omega \rightarrow \mathbb{R}$ zwei Zufallsvariablen. Ist $\sigma(X) \subseteq \mathcal{B}$ und Y unabhängig von \mathcal{B} , dann gilt für jede meßbare und beschränkte (bzw. positive) Funktion $h : \mathbb{R}^2 \rightarrow \mathbb{R}$ fast überall⁸:

$$\mathbb{E}(h \circ (X, Y) | \mathcal{B})(\omega) = \mathbb{E}(h(X(\omega), Y)) \forall \omega \in \Omega.$$

Mit anderen Worten⁹:

$$\mathbb{E}(h \circ (X, Y) | \mathcal{B}) = f \circ X \text{ fast überall,}$$

wobei $f(x) := \mathbb{E}(h(x, Y))$ ($x \in \mathbb{R}$).

Zweifellos folgen nun die unverzichtbaren Beispiele 1.4.14 und 1.4.15 von [Mik] auf den Seiten 73 – 74. Sehr schön ist der elementare Zugang von [Mik] zur Projektionseigenschaft des bedingten Erwartungswertes in L^2 , deren einfachen Beweis wir vollständig übernehmen können (s. [Mik], S. 74 – 76)!

- *Martingale* (S. 77 – 85 von [Mik]): Filtration, adaptierte stochastische Prozesse, Martingale (zeitstetige \sim und diskrete \sim), Beispiele, diskrete vorhersehbare Prozesse, Martingale-Transformation, Interpretation eines Martingals als faires Spiel

4 Das stochastische Integral

- *Erinnerungen an das Riemannsches Integral und das RS-Integral* (S. 88 – 96 von [Mik]): Wir können wiederum problemlos das Buch von Mikosch verwenden. Zu Beispiel 2.1.4 sei nur folgendes hinzuzufügen: Ist $q > 0, p > 2$ und $1/q + 1/p > 1$, so folgt notwendigerweise $q < 2$ (denn in diesem Falle gilt: $1/q > 1 - 1/p > 1 - 1/2 = 1/2$). Jedoch erhalten wir $1/q_0 + 1/p_0 \leq 1$, beispielsweise für $q_0 := 199/100 < 2$ und $p_0 := 102 > 2$, so daß also die Bedingung " $q < 2$ " nicht hinreichend für die Bedingung " $1/q + 1/p > 1$ " ist.
- *Das Itô-Integral: Motivation* (S. 96 – 101 von [Mik]): Zunächst erwähnen wir zu einem besseren Verständnis der folgenden Überlegungen die unterschiedlichen Konvergenzbegriffe und deren Bedeutung (fast sichere Konvergenz, L^p -Konvergenz, stochastische Konvergenz, Konvergenz in Verteilung). Dies läßt sich beispielsweise sehr schön und verständlich nachlesen in Kapitel 5 von [Bmd] auf den Seiten 163 – 180. Insbesondere wird es damit klarer, weshalb man über die "Konvergenz im Mittel" das stochastische Integral definieren kann. An dieser Stelle führen wir auch die Schreibweise $\|X\|_2 := (\mathbb{E}(X^2))^{1/2}$ ein.

Unter den Voraussetzungen (2.4) und (2.5) von [Mik] (s. S. 97) gilt:

⁸Hierbei ist zu beachten, daß $h \circ (X, Y)(\omega) := h(X(\omega), Y(\omega)) \forall \omega \in \Omega$ (vgl. [Sch], Bemerkung auf S. 49 oben).

⁹Mikoschs "Rule 7" auf S. 72 ist eine andere Formulierung von Theorem 3.2.

Theorem 4.1 Seien $n \in \mathbb{N}$ und $t \in \mathbb{R}_+$ gegeben. Seien $S_n(t) := \sum_{i=1}^n B(t_{i-1}) \Delta_i B$ und

$$Q_n(t) := \sum_{i=1}^n (\Delta_i B)^2. \text{ Dann gilt:}$$

- (i) $\mathbb{E}(Q_n(t)) = t;$
- (ii) $\lim_{n \rightarrow \infty} \text{Var}(Q_n(t)) = 0;$
- (iii) $\lim_{n \rightarrow \infty} \|t - Q_n(t)\|_2 = 0;$
- (iv) $\lim_{n \rightarrow \infty} \left\| \frac{B(t)^2 - t}{2} - S_n(t) \right\|_2 = 0.$

BEWEIS: Den Beweis von (i) können wir direkt von [Mik] übernehmen. Zum Beweis von (ii) ist nur folgendes zu ergänzen: Aus der Unabhängigkeit der ZVen $\Delta_i B = B(t_i) - B(t_{i-1})$ und der Borel-Meßbarkeit der Funktion $f(y) := y^2$ folgt die Unabhängigkeit der ZVen $(\Delta_i B)^2$ ($i \in \{0, \dots, n\}$). Also gilt: $\text{Var}(Q_n(t)) = \sum_{i=1}^n \text{Var}((\Delta_i B)^2)$, so daß wir nun auch hier den Beweis von [Mik] übernehmen können. (iii) folgt nun sofort aus (i) und (ii). Zum Beweis von (iv) beachten wir zunächst, daß:

$$\begin{aligned} Q_n(t) &= \sum_{i=1}^n \left(B(t_i) - B(t_{i-1}) \right)^2 = \sum_{i=1}^n \left(B(t_i)^2 - 2B(t_{i-1})B(t_i) + B(t_{i-1})^2 \right) \\ &= \sum_{i=1}^n \left(B(t_i)^2 - 2B(t_{i-1}) \cdot (B(t_i) - B(t_{i-1})) - B(t_{i-1})^2 \right) \\ &= \sum_{i=1}^n B(t_i)^2 - 2 \sum_{i=1}^n B(t_{i-1}) \Delta_i B - \sum_{i=1}^n B(t_{i-1})^2 \\ &= \sum_{i=1}^n B(t_i)^2 - 2 \sum_{i=1}^n B(t_{i-1}) \Delta_i B - \sum_{i=0}^{n-1} B(t_i)^2 \\ &= B(t_n)^2 - 2 \sum_{i=1}^n B(t_{i-1}) \Delta_i B. \end{aligned}$$

Also erhalten wir $Q_n(t) = B(t)^2 - 2S_n(t)$ (auf ganz Ω) und insbesondere $S_n(t) = \frac{1}{2}(B(t)^2 - Q_n(t))$.

Unter Beachtung von (i) folgt somit:

$$\begin{aligned} \left\| \frac{B(t)^2 - t}{2} - S_n(t) \right\|_2^2 &= \left\| \frac{B(t)^2 - t}{2} - \frac{B(t)^2 - Q_n(t)}{2} \right\|_2^2 \\ &= \mathbb{E} \left(\frac{Q_n(t) - \mathbb{E}(Q_n(t))}{2} \right)^2 \\ &= \frac{1}{4} \text{Var}(Q_n(t)). \end{aligned}$$

(ii) ergibt nun das gewünschte Resultat. □

Schauen wir uns den vorhergehenden Beweis genauer an, können wir uns sogar im folgenden Sinne von den strukturellen Eigenschaften der Brownschen Bewegung lösen. Angenommen, wir ersetzen die Brownsche Bewegung B durch einen (stetigen) stochastischen Prozess X mit $X_0 = 0$ und der folgenden Eigenschaft: Es existiere der *quadratische Variationsprozess* $\langle X, X \rangle$, der bei gegebener Zerlegungsnullfolge¹⁰ $0 = t_0^{(n)} < t_1^{(n)} < \dots < t_{k_n}^{(n)} = t$ (im Mittel) definiert sei durch

$$\langle X, X \rangle(t) := L_2\text{-}\lim_{n \rightarrow \infty} \sum_{i=1}^{k_n} (\Delta_i^{(n)} X)^2 < \infty.$$

Setzen wir dann wie oben $S_{k_n}(t) := \sum_{i=1}^{k_n} X(t_{i-1}^{(n)}) \cdot \Delta_i^{(n)} X$, so erhalten wir analog:

$$\lim_{n \rightarrow \infty} \left\| \frac{X(t)^2 - \langle X, X \rangle(t)}{2} - S_{k_n}(t) \right\|_2 = 0.$$

Korollar 4.1 *Der quadratische Variationsprozess $\langle B, B \rangle$ der Brownschen Bewegung B existiert, und es gilt: $\langle B, B \rangle(t) = t$ \mathbb{P} -fast überall für jedes $t \in [0, \infty)$.*

- *Das Itô-Integral für einfache Prozesse* (S. 101 – 107 von [Mik] und S. 21 – 25 von [Øks]¹¹):

Unter Beibehaltung der Notation von [Mik] können wir bei gegebener Zerlegung $\tau_n : 0 = t_0 < t_1 < \dots < t_n = T$ einfache Prozesse C geschlossen schreiben als:

$$C_t(\omega) = \sum_{i=1}^n Z_i(\omega) \cdot \mathbf{1}_{[t_{i-1}, t_i)}(t) + Z_n(\omega) \cdot \mathbf{1}_{\{T\}}(t) \quad (3)$$

$$= \sum_{i=1}^{n-1} Z_i(\omega) \cdot \mathbf{1}_{[t_{i-1}, t_i)}(t) + Z_n(\omega) \cdot \mathbf{1}_{[t_n, T]}(t). \quad (4)$$

Dabei ist insbesondere für jedes $i \in \{1, 2, \dots, n\}$ und für jede $\mathcal{F}_{t_{i-1}}$ -meßbare L_2 -Zufallsvariable Z der Baustein

$$D_t(\omega) := Z(\omega) \cdot \mathbf{1}_{[t_{i-1}, t_i)}(t) \quad (5)$$

ein einfacher Prozess, dessen Interpretation im Rahmen der stochastischen Finanzmathematik die folgende ist: In einem "festgehaltenen" Zustand ω kauft eine Händlerin zum Zeitpunkt t_{i-1} $Z(\omega)$ Anteile eines Finanzproduktes (z. B. eine Aktie oder eine Option), dessen Preisprozess hier zunächst (!) durch die Brown'sche Bewegung B beschrieben wird. Die Händlerin behält nun bis kurz vor Ablauf des darauffolgenden Zeitpunktes t_i ihre Handelsposition, die sie erst kurz vor Eintreten von t_i durch (eventuellen) Verkauf von $Z(\omega)$ (eventuell) ändert.¹² Setzen wir nun bei fest

¹⁰Beispielsweise fortlaufende Intervall-Halbierung von $[0, t]$ mit $k_n := 2^n$ und $t_i^{(n)} := t \cdot \frac{i}{2^n}$ ($i \in \{0, 1, \dots, k_n\}$)

¹¹Die Seiten 21 – 25 von [Øks] dienen sehr gut als Einstieg und Motivation.

¹²Eine noch realitätsnähere Modellierung wäre die Verwendung von *Stoppzeiten*, so daß also jede Zeitvariable t_i selbst wieder eine Zufallsvariable darstellt, deren inhärente Information von der Vergangenheit abhängt und somit diejenige Situation in den Finanzmärkten widerspiegelt, daß vergangene Ereignisse (und die damit verknüpfte Information) die Wahl von Zeitpunkten zur Ausübung einer Aktion beeinflussen. Zunächst nehmen wir jedoch an, daß die Zeitpunkte t_i keine Stoppzeiten sind.

gewähltem Zustand ω die einzelnen Bausteine (5) über die Zeit hinweg zusammen, so erhalten wir (4), basierend auf der Handelsstrategie $Z_1(\omega), Z_2(\omega), \dots, Z_n(\omega)$, wobei der Handel nur in den Zeitpunkten $0 = t_0, t_1, \dots, t_n = T$ stattfindet. Damit ergibt sich zum Zeitpunkt T der Nettogewinn bzw. Nettoverlust der Gesamtposition, die durch den Handel mit dem Preisprozess B entstanden ist, gerade als das Itô-Integral $I_T(C)(\omega) := (\int_0^T C_s dB_s)(\omega) = \sum_{i=1}^n Z_i(\omega) \cdot (B(\omega, t_i) - B(\omega, t_{i-1}))!$

Sei nun $0 \leq t < T$ beliebig und τ_n obige Zerlegung von $[0, T]$. Dann existiert genau ein $k \in \{1, \dots, n\}$, so daß $t_{k-1} \leq t < t_k$. Da $[t_{i-1}, t_i] \cap [0, t] = [t_{i-1}, t_i]$ für $i \in \{1, \dots, k-1\}$, $[t_{k-1}, t_k] \cap [0, t] = [t_{k-1}, t]$ und $[t_{i-1}, t_i] \cap [0, t] = \emptyset$ für $i \in \{k+1, \dots, n\}$, erhalten wir somit für alle $s \in [0, T]$:

$$C_s(\omega) \cdot \mathbf{1}_{[0,t]}(s) = \sum_{i=1}^{k-1} Z_i(\omega) \cdot \mathbf{1}_{[t_{i-1}, t_i]}(s) + Z_k(\omega) \cdot \mathbf{1}_{[t_{k-1}, t]}(s).$$

Insbesondere ist nach Konstruktion $(\omega, s) \mapsto C_s(\omega) \cdot \mathbf{1}_{[0,t]}(s)$ wieder ein einfacher Prozess bezüglich der Zerlegung $0 = \tilde{t}_0 < \tilde{t}_1 < \dots < \tilde{t}_{k-1} < \tilde{t}_k = t$, wobei $\tilde{t}_i := t_i$ für $i = 1, 2, \dots, k-1$ und $\tilde{t}_k = t$ (vgl. Gleichung (4)). Somit wird durch $I_t(C) := \int_0^t C_s dB_s := \int_0^T C_s \mathbf{1}_{[0,t]}(s) dB_s$ wieder ein Itô-Integral $I_t(C)$ definiert, dessen Wert unter Beachtung der obigen Überlegungen folgende Gestalt annimmt:

$$\begin{aligned} I_t(C)(\omega) &= \left(\int_0^T C_s \mathbf{1}_{[0,t]}(s) dB_s \right)(\omega) \\ &= \sum_{i=1}^{k-1} Z_i(\omega) \cdot (B(\omega, t_i) - B(\omega, t_{i-1})) + Z_k(\omega) \cdot (B(\omega, t) - B(\omega, t_{k-1})) \\ &= I_{t_{k-1}}(C)(\omega) + Z_k(\omega) \cdot (B(\omega, t) - B(\omega, t_{k-1})). \end{aligned}$$

Kürzer läßt sich dies auch schreiben als:

$$I_t(C) = \int_0^T C_s \mathbf{1}_{[0,t]}(s) dB_s = I_{t_{k-1}}(C) + Z_k \cdot (B(t) - B(t_{k-1})).$$

Damit können wir nun problemlos die Martingaleigenschaft des stochastischen Prozesses $\{I_t(C) : t \in [0, T]\}$ zeigen. Seien also $s, t \in [0, T]$ mit $s < t$. Dann gibt es also genau einen Wert $l \in \{1, \dots, n\}$, so daß $t_{l-1} \leq s < t_l$. Wir unterscheiden nun zwei Fälle:

- (i) $s \in [t_{k-1}, t_k)$,
- (ii) $s \notin [t_{k-1}, t_k)$.

Im Fall (i) muß $k = l$ sein und somit

$$\begin{aligned} I_t(C) - I_s(C) &= I_{t_{k-1}}(C) + Z_k \cdot (B(t) - B(t_{k-1})) - (I_{t_{k-1}}(C) + Z_k \cdot (B(s) - B(t_{k-1}))) \\ &= Z_k \cdot (B(t) - B(s)). \end{aligned}$$

Da $\sigma(Z_k) \subseteq \mathcal{F}_{t_{k-1}} \subseteq \mathcal{F}_s$ folgt somit die \mathcal{F}_s -Meßbarkeit von Z_k . Also ergibt sich aus der Martingaleigenschaft von B : $\mathbb{E}_s(I_t(C) - I_s(C)) = Z_k \cdot \mathbb{E}_s(B(t) - B(s)) = Z_k \cdot 0$. Damit erkennen wir in Fall (i) die Martingaleigenschaft von $\{I_t(C) : t \in [0, T]\}$.

Es gelte nun der Fall (ii). Dann ist offensichtlich $l \in \{1, \dots, k-1\}$. Wie zuvor erkennen wir, daß

$$\begin{aligned} I_t(C) - I_s(C) &= I_{t_{k-1}}(C) + Z_k \cdot (B(t) - B(t_{k-1})) - (I_{t_{l-1}}(C) + Z_l \cdot (B(s) - B(t_{l-1}))) \\ &= \sum_{i=l}^{k-1} Z_i \cdot (B(t_i) - B(t_{i-1})) + Z_k \cdot (B(t) - B(t_{k-1})) - Z_l \cdot (B(s) - B(t_{l-1})) \\ &= \sum_{i=l+1}^{k-1} Z_i \cdot \Delta_i B + Z_k \cdot (B(t) - B(t_{k-1})) - Z_l \cdot (B(t_l) - B(s)). \end{aligned}$$

Nun gilt nach Voraussetzung (ii):

$$\mathcal{F}_{t_{l-1}} \subseteq \mathcal{F}_s \subseteq \mathcal{F}_{t_l} \subseteq \mathcal{F}_{t_{k-1}}$$

und demzufolge

$$\mathcal{F}_s \subseteq \mathcal{F}_{t_{i-1}} \quad \forall i \in \{l+1, \dots, k-1\}.$$

Berechnen wir somit nun die bedingten Erwartungswerte \mathbb{E}_s und berücksichtigen dabei jeweils die "Turmeigenschaft", dann erhalten wir der Reihe nach die folgenden Werte:

$$\mathbb{E}_s(Z_i \cdot \Delta_i B) = \mathbb{E}_s(Z_i \cdot \mathbb{E}_{t_{i-1}}(\Delta_i B)) = 0 \quad \forall i \in \{l+1, \dots, k-1\},$$

$$\mathbb{E}_s(Z_k \cdot (B(t) - B(t_{k-1}))) = \mathbb{E}_s(Z_k \cdot \mathbb{E}_{t_{k-1}}(B(t) - B(t_{k-1}))) = 0,$$

und analog ergibt sich

$$\mathbb{E}_s(Z_l \cdot (B(t_l) - B(s))) = 0.$$

Wegen der Linearität des Operators \mathbb{E}_s haben wir somit auch im zweiten Fall (ii) die Martingaleigenschaft des Prozesses $\{I_t(C) : t \in [0, T]\}$ erhalten! \square

Den Beweis und die Formulierung der sehr wichtigen *Isometrie-Eigenschaft* des Itô-Integrals und dessen erste Rechengesetze können wir wieder bedenkenlos von [Mik] übernehmen.

- *Das allgemeine Itô-Integral* (S. 107 – 112 von [Mik] und S. 25 – 33 von [Øks]): An dieser Stelle erscheint es sehr sinnvoll, sowohl [Mik] als auch [Øks] abwechselnd einzusetzen. Wir benutzen hierzu Øksendals Bezeichnung $\mathcal{V}(0, T)$ für die Klasse derjenigen Funktionen, die als Integranden für die das Itô-Integral wohldefiniert sind (s. Definition 3.1.4 auf S. 25 von [Øks]). Offensichtlich ist $\mathcal{V}(0, T)$ gerade diejenige Klasse von Funktionen, die von Mikosch als Integranden-Prozesse verwendet werden (s. [Mik], S. 108). Sehr interessant und nützlich ist dabei die erste Aussage von Aufgabe 3.9 in [Øks], die sich nach Lokalisationstechniken¹³ direkt aus dem ersten Schritt der Erweiterung und der Definition des allgemeinen Itô-Integrals ergibt (s. [Øks], Aufgabe 3.9, S. 39, "Step 1" auf S. 27 und Definition 3.1.6 auf S. 29). Diese Überlegungen werden für die Skizzierung des Beweises der einfachsten Itô-Formel noch sehr nützlich sein. Dabei sollte sehr deutlich auf die *L²-Konstruktion* des Itô-Integrals hingewiesen werden!

¹³Mittels Lokalisation dürfen wir nämlich die Brownsche Bewegung "stoppen" und somit o.b.d.A. die Beschränktheit von f voraussetzen (vgl. [Øks], Aufgabe 4.9 auf S. 57)

- *Das Itô-Kalkül* (S. 112 – 122 von [Mik] und S. 43 – 49 von [Øks]): Wir gelangen nun an das zentrale Resultat der stochastischen Analysis, das die gesamte stochastische Finanzmathematik durchdringt: die sogenannte Itô-Formel. Wir werden zumindestens für die einfachste Version eine Beweisskizze angeben, die sehr nützlich zum Verständnis der Itô-Formel ist, ohne dabei zu sehr auf technische Details einzugehen.

Theorem 4.2 (Itô-Formel: einfachste Version) Sei $f : \mathbb{R} \rightarrow \mathbb{R}$ eine zweimal stetig-partiell differenzierbare Funktion. Dann gilt \mathbb{P} -fast sicher für alle $s, t \in [0, \infty)$ mit $s \leq t$:

$$f(B_t(\omega)) - f(B_s(\omega)) = \int_s^t f'(B_u(\omega)) dB_u(\omega) + \frac{1}{2} \int_s^t f''(B_u(\omega)) du.$$

BEWEISSKIZZE: O.b.d.A. sei $s = 0$. Der Beweisaufbau unterteilt sich in ... Schritte:

- (1) Lokalisation
- (2) Taylorreihenentwicklung bis zur Ordnung 2
- (3) Abschätzung der einzelnen Teilsummen der Taylorreihenentwicklung in der L^2 -Norm.

Ad (1): Wir werden hier nur die zugrundeliegende Idee angeben. Jeder bezüglich der Brownschen Bewegung adaptierte und stetige stochastische Prozess X mit $X_0 = 0$ ist in dem folgenden Sinne *lokal beschränkt*: Es gibt eine Folge von Stoppzeiten (τ_n) mit $\tau_n \nearrow \infty$, so daß der gestoppte Prozess X^{τ_n} beschränkt ist. Dazu brauchen wir nur für beliebiges $n \in \mathbb{N}$ und für beliebiges $\omega \in \Omega$

$$\tau_n(\omega) := \inf\{t \in [0, \infty) : |X_t(\omega)| \geq n\}$$

zu setzen. Es gilt dann nämlich zunächst nach Definition von τ_n für jedes $\omega \in \Omega$ und für jedes $\varepsilon > 0$: $|X_{\tau_n(\omega)-\varepsilon}(\omega)| < n$ und somit (wegen der linksseitigen Stetigkeit von X) $|X_{\tau_n(\omega)}(\omega)| = \lim_{\varepsilon \rightarrow 0} |X_{\tau_n(\omega)-\varepsilon}(\omega)| \leq n$. Also erhalten wir insgesamt $|X_t^{\tau_n(\omega)}(\omega)| = |X_{t \wedge \tau_n(\omega)}(\omega)| \leq n \quad \forall (t, \omega) \in [0, \infty) \times \Omega$. Insbesondere läßt sich damit dann zeigen, daß wir o.B.d.A. die Beschränktheit von f , f' und von f'' auf \mathbb{R} annehmen können – mit jeweils gleicher oberer Schranke (vgl. mit Satz 3.18 auf S. 96 von [H-T] und mit dem Beweis von Theorem 3.3 in [K-S], S. 149 – 154). Wegen der vorausgesetzten Stetigkeit und der Beschränktheit von f' können wir daher bei gegebener Zerlegung $0 = t_0 < t_1 < \dots < t_n = t$ gemäß Schritt 1 in [Øks] nun durch

$$C_t(\omega) = \sum_{i=1}^n f'(B_{t_{i-1}}(\omega)) \cdot \mathbf{1}_{[t_{i-1}, t_i)}(s) + f'(B_{t_{n-1}}(\omega)) \cdot \mathbf{1}_{\{t\}}(s)$$

einen einfachen Prozess $C \in \mathcal{V}(0, t)$ konstruieren, aus dem wir dann durch immer feinere Unterteilung des Intervalles $[0, t]$ mittels einer *Zerlegungsnullfolge* $Z_n : 0 = t_0^{(n)} < t_1^{(n)} < \dots < t_{k_n}^{(n)} = t$ eine Folge $C^{(n)}$ solcher elementaren Prozesse erhalten, die folgenden Grenzübergang erfüllt:

$$\int_0^t f' \circ B_u dB_u = \lim_{n \rightarrow \infty} \int_0^t f' \circ C_u^{(n)} dB_u = \lim_{n \rightarrow \infty} \sum_{i=1}^{k_n} (f' \circ B_{t_{i-1}^{(n)}}) \cdot \Delta_i^{(n)} B \quad (\text{Limes in } L^2),$$

wobei natürlich $\Delta_i^{(n)} B := B_{t_i^{(n)}} - B_{t_{i-1}^{(n)}}$.

Ad (2): Da nach Voraussetzung f'' stetig auf \mathbb{R} ist, also auf jedem kompakten Teilintervall von \mathbb{R} sogar gleichmäßig stetig ist, können wir die folgende Version der Taylorreihenentwicklung anwenden:

$$f(x + \xi) - f(x) = f'(x) \cdot \xi + \frac{1}{2} \cdot f''(x) \cdot \xi^2 + r(x, \xi) \cdot \xi^2,$$

wobei für jedes $M > 0$ $\sup_{|x| \leq M} |r(x, \xi)| \rightarrow 0$ für $\xi \rightarrow 0, \xi \neq 0$. Ausgehend von einer

beliebigen Zerlegungsnullfolge $Z_n : 0 = t_0^{(n)} < t_1^{(n)} < \dots < t_{k_n}^{(n)} = t$ erhalten wir damit folgende Teleskopsumme ¹⁴:

$$f \circ B_t - f \circ B_0 = \sum_{i=1}^{k_n} (f \circ B_{t_i} - f \circ B_{t_{i-1}}) = I_1^{(n)} + \frac{1}{2} I_2^{(n)} + I_3^{(n)}, \quad (6)$$

wobei $I_1^{(n)} := \sum_{i=1}^{k_n} (f' \circ B_{t_{i-1}}) \cdot \Delta_i B$, $I_2^{(n)} := \sum_{i=1}^{k_n} (f'' \circ B_{t_{i-1}}) \cdot (\Delta_i B)^2$ und

$I_3^{(n)} := \sum_{i=1}^{k_n} (r \circ (B_{t_{i-1}}, \Delta_i B)) \cdot (\Delta_i B)^2$. Aufgrund der obigen Überlegungen in (1), erkennen wir sofort, daß

$$I_1^{(n)} \rightarrow \int_0^t f' \circ B_u dB_u \text{ für } n \rightarrow \infty \text{ (Limes in } L^2).$$

Insbesondere konvergiert $I_1^{(n)}$ also stochastisch:

$$\int_0^t f' \circ B_u dB_u = \mathbb{P} - \lim_{n \rightarrow \infty} I_1^{(n)}. \quad (7)$$

Zur Abschätzung von $I_2^{(n)}$ setzen wir für jedes i $f'' \circ B_{t_{i-1}} =: a_{i-1}$. Damit folgt zunächst (nach Definition der Brownschen Bewegung), daß für jedes i die ZVe a_{i-1} unabhängig ist von der ZVe $X_i := (\Delta_i B)^2 - (t_i - t_{i-1})$ und für jedes $i < j$ die ZVe $a_{i-1} a_{j-1} X_i$ unabhängig ist von der ZVe X_j . Da $\mathbb{E}(X_i) = 0$ für jedes i , gilt somit:

$$\begin{aligned} \mathbb{E} \left(\left(I_2^{(n)} - \sum_{i=1}^{k_n} a_{i-1} (t_i - t_{i-1}) \right)^2 \right) &= \mathbb{E} \left(\left(\sum_{i=1}^{k_n} a_{i-1} \cdot X_i \right)^2 \right) \\ &= \sum_{i=1}^{k_n} \sum_{j=1}^{k_n} \mathbb{E}(a_{i-1} a_{j-1} X_i X_j) = \sum_{i=1}^{k_n} \mathbb{E}(a_{i-1}^2) \cdot \mathbb{E}(X_i^2). \end{aligned}$$

¹⁴Um die Darstellung möglichst übersichtlich zu halten, schreiben wir für $i = 0, 1, \dots, k_n$ anstatt $t_i^{(n)}$ kürzer t_i .

Die direkte Berechnung von $\mathbb{E}(X_i^2)$ liefert $\mathbb{E}(X_i^2) = 2(t_i - t_{i-1})^2$ (vgl. [Øks], S. 47 unten), und somit konvergiert $\mathbb{E}\left(\left(I_2^{(n)} - \sum_{i=1}^{k_n} a_{i-1}(t_i - t_{i-1})\right)^2\right) = 2 \sum_{i=1}^{k_n} \mathbb{E}(a_{i-1}^2)(t_i - t_{i-1})^2$ wegen der angenommenen Beschränktheit von f'' gegen Null für $n \rightarrow \infty$. Mit anderen Worten:

$$I_2^{(n)} \rightarrow \int_0^t f'' \circ B_u du \text{ für } n \rightarrow \infty \text{ (Limes in } L^2)$$

und damit insbesondere

$$\int_0^t f'' \circ B_u du = \mathbb{P} - \lim_{n \rightarrow \infty} I_2^{(n)}. \quad (8)$$

Bleibt also nur noch $I_3^{(n)}$ abzuschätzen. Dazu beachten wir folgende Ungleichung, die für jedes $\omega \in \Omega$ erfüllt ist:

$$|I_3^{(n)}(\omega)| \leq \max_{1 \leq i \leq k_n} |r(B_{t_{i-1}}(\omega), \Delta_i B(\omega))| \cdot \sum_{i=1}^{k_n} (\Delta_i B(\omega))^2$$

Wählt man $M := \max_{0 \leq s \leq t} |B_s(\omega)|$, so folgt nach Darstellung des Restgliedes, daß

$\max_{1 \leq i \leq k_n} |r(B_{t_{i-1}}(\omega), \Delta_i B(\omega))| \rightarrow 0$ für $n \rightarrow \infty$. Setzen wir (wie früher) $Q_n(t) := \sum_{i=1}^{k_n} (\Delta_i B)^2$, dann konvergiert $Q_n(t)$ in L_2 (und damit insbesondere stochastisch) für $n \rightarrow \infty$ gegen den (endlichen) Wert t , woraus wir erkennen können, daß

$$0 = \mathbb{P} - \lim_{n \rightarrow \infty} I_3^{(n)}. \quad (9)$$

Addition der stochastischen Limes (6), (7) und (8) liefert damit – unter Beachtung der Teleskopsumme (6) und der Stetigkeit der Prozesse $(t, \omega) \mapsto f(B_t(\omega)) - f(B_0(\omega))$ und $(t, \omega) \mapsto \int_0^t f'(B_u(\omega)) dB_u(\omega) + \frac{1}{2} \int_0^t f''(B_u(\omega)) du$ – deren Ununterscheidbarkeit, also die Behauptung. \square

Die Erweiterungen der Itô-Formel können problemlos (ohne Beweis) von [Mik] übernommen werden. Sehr gut merken sollte man sich dabei die Struktur der Erweiterung I, da diese oft in der Praxis benötigt wird! Die Erweiterung II der Itô-Formel auf S. 120 von [Mik] läßt sich dann sofort als Merkregel herleiten, indem man in der Erweiterung I die Brownsche Bewegung dB_t und deren quadratische Variation $d\langle B, B \rangle_t = dt$ jeweils durch den Itô-Prozess $dX_t = A_t^1 dt + A_t^2 dB_t$ und dessen quadratische Variation $d\langle X, X \rangle_t \stackrel{!}{=} (A_t^2)^2 dt$ "ersetzt":

$$f(t, X_t) = f(s, X_s) + \int_s^t f_1(u, X_u) du + \int_s^t f_2(u, X_u) dX_u + \int_s^t \frac{1}{2} f_{22}(u, X_u) d\langle X, X \rangle_u.$$

Um Itô-Prozesse genauer zu charakterisieren, benötigen wir an dieser Stelle einen möglichst grossen Definitionsbereich von stochastischen Prozessen, die als Integrand für das Itô-Integral gewählt werden dürfen und die $\mathcal{V}(0, T)$ als Teilmenge enthalten. Dazu betrachten wir die folgenden Mengen: (vgl. [Nls], S. 30 und S. 35):

- \mathcal{L}^1 sei die Menge aller meßbaren und bezüglich der Brownschen Bewegung adaptierten Prozesse A , so daß für alle $t \in [0, T]$ gilt: $\int_0^t |A_s| ds < \infty$ \mathbb{P} -fast überall.
- \mathcal{L}^2 sei die Menge aller meßbaren und bezüglich der Brownschen Bewegung adaptierten Prozesse D , so daß für alle $t \in [0, T]$ gilt: $\int_0^t |D_s|^2 ds < \infty$ \mathbb{P} -fast überall.

Wir benötigen nun folgende nicht-triviale Beschreibung von Itô-Prozessen, die wir ohne Beweis hier angeben (vgl. [Nls], S. 52):

Satz und Definition 4.1 *Seien $A^{(1)} \in \mathcal{L}^1$ und $A^{(2)} \in \mathcal{L}^2$. Dann sind die folgenden Integrale wohldefiniert: das RS-Integral $(t, \omega) \mapsto \int_0^t A^{(1)}(s, \omega) ds$ und das Itô-Integral $(t, \omega) \mapsto \int_0^t A^{(2)}(s, \omega) dB(s, \omega)$. Insbesondere existiert damit für jede \mathcal{F}_0 -meßbare ZVe X_0 der stochastische Prozess*

$$X = X_0 + \int_0^\bullet A_s^{(1)} ds + \int_0^\bullet A_s^{(2)} dB_s.$$

Ein stochastischer Prozess X , der eine solche Darstellung besitzt, heißt Itô-Prozess.

Wir können nun unter (trickreicher) Verwendung der Erweiterung II der Itô-Formel die Eindeutigkeit der Koeffizienten des Itô-Prozesses zeigen (vgl. [Nls], Proposition 2.6 auf S. 60):

Proposition 4.1 *Seien $A^{(1)}, D^{(1)} \in \mathcal{L}^1$ und $A^{(2)}, D^{(2)} \in \mathcal{L}^2$. Seien X_0 und Y_0 zwei \mathcal{F}_0 -meßbare ZVen. Besitzt ein Itô-Prozess X zwei ununterscheidbare Darstellungen $X_t = X_0 + \int_0^t A_s^{(1)} ds + \int_0^t A_s^{(2)} dB_s$ und $X_t = Y_0 + \int_0^t D_s^{(1)} ds + \int_0^t D_s^{(2)} dB_s$, dann gilt: $X_0 = Y_0$ fast-überall, $A^{(1)} = D^{(1)}$ (fast-sicher identisch) und $A^{(2)} = D^{(2)}$ (fast-sicher identisch).*

BEWEIS: Seien also $X_0 + \int_0^t A_s^{(1)} ds + \int_0^t A_s^{(2)} dB_s$ und $Y_0 + \int_0^t D_s^{(1)} ds + \int_0^t D_s^{(2)} dB_s$ zwei ununterscheidbare Itô-Prozesse. Dann gilt insbesondere $X_0 = Y_0$ fast-überall. Also wird durch die Differenz $Z_t := \int_0^t (A_s^{(1)} - D_s^{(1)}) ds - \int_0^t (A_s^{(2)} - D_s^{(2)}) dB_s$ ein Itô-Prozess Z definiert, der ununterscheidbar vom Null-Prozess 0 ist. Setzen wir $G^{(1)} := A^{(1)} - D^{(1)}$ und $G^{(2)} := A^{(2)} - D^{(2)}$, so müssen wir nur noch zeigen, daß $G^{(1)} \stackrel{!}{=} 0$ fast-sicher und $G^{(2)} \stackrel{!}{=} 0$ fast-sicher. Dazu wenden wir nun die Erweiterung II der Itô-Formel auf den transformierten Prozess $f(t, Z_t)$ an, mit der Funktion $f(t, x) := e^x$, und erhalten bis auf Ununterscheidbarkeit:

$$e^Z = e^{Z_0} + \int_0^\bullet \left(G_s^{(1)} e^{Z_s} + \frac{1}{2} (G_s^{(2)})^2 e^{Z_s} \right) ds - \int_0^\bullet G_s^{(2)} e^{Z_s} dB_s.$$

Da jedoch $Z = 0$ (ununterscheidbar), folgt somit durch Einsetzen in die obere Gleichung (bis auf Ununterscheidbarkeit):

$$1 = 1 + \int_0^\bullet \left(G_s^{(1)} + \frac{1}{2} (G_s^{(2)})^2 \right) ds - \int_0^\bullet G_s^{(2)} dB_s \stackrel{(!)}{=} 1 + Z + \int_0^\bullet \frac{1}{2} (G_s^{(2)})^2 ds.$$

Also erhalten wir $\int_0^\bullet \frac{1}{2} (G_s^{(2)})^2 ds = 0$ (bis auf Ununterscheidbarkeit) und damit $G^{(2)} = 0$ (fast-sicher identisch). Ergo erhalten wir (vgl. [Nls], Proposition 1.34 auf S. 38) $\int_0^\bullet G_s^{(2)} dB_s = 0$ (bis auf Ununterscheidbarkeit) und damit $Z = 0 = \int_0^\bullet G_s^{(1)} ds$ (bis auf Ununterscheidbarkeit). Also ist insbesondere $G^{(1)} = 0$ fast-sicher identisch (vgl. [Nls], Proposition 1.26 auf S. 31). \square

... TBC

Literatur

- [Bau] Bauer, H., *Maß – und Integrationstheorie*, 2. überarbeitete Auflage, de Gruyter, 1992.
- [Bin-Kie] Bingham, N. H. und R. Kiesel, *Risk-neutral valuation*, Springer Finance, 1998.
- [Bmd] Brémaud, P., *An introduction to probalistic modeling*, Undergraduate Texts in Mathematics, Springer 1988.
- [Hac] Hackenbroch, W., *Integrationstheorie*, Teubner 1987.
- [H-T] Hackenbroch, W. und A. Thalmaier, *Stochastische Analysis*, Teubner 1994.
- [K-S] Karatzas I. und Shreve, S. E., *Brownian motion and stochastic calculus*, 2nd ed., Springer, 1991.
- [Kre] Krenzel, U., *Einführung in die Wahrscheinlichkeitstheorie und Statistik*, Aufbaukurs Mathematik, Vieweg 1988.
- [Lam-Lap] Lamberton, D. und B. Lapeyre, *Introduction to stochastic calculus applied to finance*, Chapman & Hall, 1996.
- [Mar-Ken-Bib] Mardia, K. V., Kent, J. T. und J. M. Bibby, *Multivariate analysis*, Academic Press, 1979.
- [Mey] Meyer, M., *Continuous stochastic calculus with applications to finance*, Chapman & Hall/CRC, 2001.
- [Mik] Mikosch, T., *Elementary stochastic calculus with finance in view*, World Scientific, 1998.
- [Nls] Nielsen, L. T., *Pricing and hedging of derivative securities*, Oxford University Press, 1999.

- [Øks] Øksendal, B., *Stochastic Differential Equations: an introduction with applications*, fifth edition, Springer, 1998.
- [Ptt] Protter, P., *Stochastic integration and differential equations*, Springer, 1990.
- [Sch] Schürger, K., *Wahrscheinlichkeitstheorie*, Lehr- und Handbücher der Statistik, Oldenbourg, 1998.
- [Wil] Williams, D., *Probability with martingales*, Cambridge University Press, 1991.