STOCHASTISCHE PROZESSE

II: Martingale und Brownsche Bewegung

Wolfgang König

Vorlesungsskript

Universität Leipzig

Wintersemester 2005/6

Inhaltsverzeichnis

Vorwort

Dies ist das Skript zum zweiten Teil einer Vorlesung über Stochastische Prozesse, gehalten an der Universität Leipzig im Wintersemester 2005/06. Die beiden Themen sind Martingale in diskreter Zeit und die Brown'sche Bewegung. Wir bringen zunächst die allgemeine Theorie der Martingale und etliche Beispiele und wenden dann diese Theorie an fünf Beispielen aus verschiedenen Teilgebieten der Wahrscheinlichkeitstheorie und verwandter Gebiete an. Zum Schluss bringen wir eine Darstellung der Brown'schen Bewegung und ihrer wichtigsten Eigenschaften.

Der Martingalteil dieses Skriptes basiert auf einem Skript von Peter Eichelsbacher (Bochum), das wiederum auf einen Text von Erwin Bolthausen (Zürich) zurück geht, welches seinerseits viele Anleihen in dem Buch von D. Williams macht. Ich danke allen hier Genannten für die Überlassung des Textes. Der Skriptteil über die Brown'sche Bewegung stützt sich auf das (leider vergriffene) Büchlein [Pa84] sowie auf die entsprechenden Kapitel des im Werden befindlichen Lehrbuches [Kl06]. Ich bedanke mich bei Achim Klenke (Mainz) für die freundliche Überlassung einer vorläufigen Fassung von [Kl06].

Leipzig, im Dezember 2005

Kapitel 1

Theorie der Martingale

Eine der ursprünglichen Anwendungen und Motivationen für eine Beschäftigung mit Martingalen ist die formale Beschreibung eines fairen Glücksspiels, genauer gesagt: des Gewinns in einer Serie von fairen Glücksspielen. Es stellte sich heraus, dass die stochastische Struktur einer Vielzahl von stochastischen Prozessen der eines Martingals entspricht, so dass zusammenfassend eine Vielzahl von Situationen betrachtet werden kann. Eine sehr angenehme Eigenschaft von Martingalen ist, dass sie unter sehr schwachen Voraussetzungen bei divergierender Zeit fast sicher konvergieren. Das Interessante daran ist, dass diese Konvergenz nicht durch Borel-Cantelli-Argumente (die üblicher Weise für fast sichere Konvergenzaussagen verwendet werden) bewiesen wird, sondern durch völlig andersartige. Unter nur wenig stärkeren Voraussetzungen liegt die Konvergenz auch im $\mathcal{L}^p$-Sinn vor. Daher ist die Theorie der Martingale zentral für die Untersuchung der Konvergenz von Folgen von Zufallsvariablen.

Dem ganzen Kapitel liege ein Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbb{P})$ zu Grunde.

1.1 Definition und einfache Eigenschaften

Wir betrachten zunächst eine Standardsituation:

Es sei $(Y_n)_{n\in\mathbb{N}}$ eine Folge unabhängiger integrierbarer reeller Zufallsvariablen, und sei $X_n := Y_1+\cdots+Y_n$. Offenbar gilt $\sigma(X_1,\ldots,X_n) = \sigma(Y_1,\ldots,Y_n)$. Da Y_{n+1} unabhängig von $\sigma(Y_1,\ldots,Y_n)$ ist, also auch von $\sigma(X_1,\ldots,X_n)$, folgt

$$\mathbb{E}\big(Y_{n+1}|X_1,\ldots,X_n\big) = \mathbb{E}(Y_{n+1}) \qquad \mathbb{P}\text{-fast sicher.}$$

Weiter gilt $\mathbb{E}(Y_i|X_1,\ldots,X_n) = Y_i$ $\mathbb{P}$-fast sicher für alle $i = 1,\ldots,n$. Also folgt durch Summation

$$\mathbb{E}(X_{n+1}|X_1,\ldots,X_n) = X_n + \mathbb{E}(Y_{n+1}) \qquad \mathbb{P}\text{-fast sicher.}$$

Sind alle Y_n zentriert, so folgt

$$\mathbb{E}(X_{n+1}|X_1,\ldots,X_n) = X_n \qquad \mathbb{P}\text{-fast sicher.}$$

Gilt $\mathbb{E}(Y_n) \leq 0$ (bzw. ≥ 0) für alle n, so folgt entsprechend

$$\mathbb{E}(X_{n+1}|X_1,\ldots,X_n) \leq X_n \quad (\geq X_n) \qquad \mathbb{P}\text{-fast sicher,}$$

jeweils für alle $n \in \mathbb{N}$.

Nun folgt eine Standardanwendung der Idee der Martingale, die uns als motivierendes Beispiel dienen wird:

Beispiel 1.1.1 (Glückspielserie). Die unabhängigen Zufallsgrößen $Y_1, Y_2, \dots$ seien nun ± 1-wertig mit $\mathbb{P}(Y_n = 1) = p$ und somit $\mathbb{E}(Y_n) = 2p - 1$. Wir interpretieren Y_n als den Ausgang des n-ten Spiels: Erfolg oder kein Erfolg. Der Spieler verfolge die folgende Strategie. Vor Eintritt in das Spiel wählt er eine Folge $(e_n)_{n\in\mathbb{N}}$ von Abbildungen $e_n : \{-1, +1\}^n \to [0, \infty)$. Für die $(n + 1)$-te Spielrunde leistet er den Einsatz $e_n(Y_1, \dots, Y_n)$, der also von den Ausgängen aller vorangegangenen Spiele abhängt. Die erste Runde wird ohne vorherigen Einsatz gespielt. Der Gegenspieler leiste keine Einsätze. Mit X_n bezeichnen wir den Gesamtgewinn des Spielers nach der n-ten Runde. Also gelten

$$X_1 = Y_1 \qquad \text{und} \qquad X_{n+1} = X_n + e_n(Y_1, \dots, Y_n)Y_{n+1}.$$

Dann errechnet man leicht:

$$\begin{aligned}\mathbb{E}(X_{n+1}|Y_1, \dots, Y_n) &= X_n + e_n(Y_1, \dots, Y_n)\mathbb{E}(Y_{n+1}|Y_1, \dots, Y_n)\\ &= X_n + e_n(Y_1, \dots, Y_n)\mathbb{E}(Y_{n+1})\\ &= X_n + (2p - 1)e_n(Y_1, \dots, Y_n).\end{aligned}$$

Im Fall $p = 1/2$ folgt

$$\mathbb{E}(X_{n+1}|Y_1, \dots, Y_n) = X_n \qquad \mathbb{P}\text{-fast sicher.}$$

Im Falle von Chancengleichheit zwischen dem Spieler und der Bank ($p = 1/2$) spielt der Spieler also ein faires Spiel, denn der erwartete Gewinn im nächsten Spiel gegeben den gesamten bisherigen Spielverlauf ist so groß wie der aktuelle gewonnene Betrag. Analog gilt für $p < 1/2$ bzw. $p > 1/2$:

$$\mathbb{E}(X_{n+1}|Y_1, \dots, Y_n) \le X_n \quad (\ge X_n) \qquad \mathbb{P}\text{-f.s.}$$

$\diamond$

Wir nehmen diese Beobachtungen zum Anlass für die folgenden Definitionen.

Definition 1.1.2 (Filtrierung, Martingal). *(i) Eine* Filtrierung *von $\mathcal{F}$ ist eine aufsteigende Folge $(\mathcal{F}_n)_n$ von Teil-σ-Algebren von $\mathcal{F}$ (d. h. $\mathcal{F}_m \subset \mathcal{F}_n$ für alle $m \le n$). Eine Folge von Zufallsvariablen $(X_n)_n$ auf $(\Omega, \mathcal{F}, \mathbb{P})$ heißt* $(\mathcal{F}_n)_n$-adaptiert *(oder* angepasst*), wenn X_n für jedes $n \in \mathbb{N}$ $\mathcal{F}_n$-messbar ist.*

(ii) Sei $(X_n)_n$ eine $(\mathcal{F}_n)_n$-adaptierte Folge von integrierbaren Zufallsgrößen. $(X_n)_n$ heißt

- $(\mathcal{F}_n)_n$-Martingal, *falls $\mathbb{E}(X_n|\mathcal{F}_m) = X_m$ $\mathbb{P}$-fast sicher für alle $m, n \in \mathbb{N}$ mit $m \le n$ gilt.*
- $(\mathcal{F}_n)_n$-Submartingal, *falls $\mathbb{E}(X_n|\mathcal{F}_m) \ge X_m$ $\mathbb{P}$-fast sicher für alle $m, n \in \mathbb{N}$ mit $m \le n$ gilt.*
- $(\mathcal{F}_n)_n$-Supermartingal, *falls $\mathbb{E}(X_n|\mathcal{F}_m) \le X_m$ $\mathbb{P}$-fast sicher für alle $m, n \in \mathbb{N}$ mit $m \le n$ gilt.*

Bemerkung 1.1.3. (i) Die *kanonische Filtrierung* für eine gegebene Folge $\mathbb{X} = (X_n)_{n\in\mathbb{N}}$ von Zufallsgrößen ist definiert als $(\mathcal{F}_n^{\mathbb{X}})_{n\in\mathbb{N}}$ mit $\mathcal{F}_n^{\mathbb{X}} := \sigma(X_1, \dots, X_n)$. Offensichtlich ist $(X_n)_{n\in\mathbb{N}}$ adaptiert an $(\mathcal{F}_n^{\mathbb{X}})_{n\in\mathbb{N}}$.

(ii) Um nachzuprüfen, dass $(X_n)_n$ ein Martingal ist, reicht es natürlich zu zeigen, dass $\mathbb{P}$-fast sicher gilt:

$$\mathbb{E}(X_{n+1}|\mathcal{F}_n) = X_n, \qquad n \in \mathbb{N}.$$

Dies folgt aus der $\mathbb{P}$-fast sicheren Gleichheit

$$\mathbb{E}(X_{n+2}|\mathcal{F}_n) = \mathbb{E}\big(\mathbb{E}(X_{n+2}|\mathcal{F}_{n+1})|\mathcal{F}_n\big) = \mathbb{E}(X_{n+1}|\mathcal{F}_n) = X_n.$$

Entsprechendes gilt für Sub- bzw. Supermartingale.

(iii) Das Wort „Martingal“ hat mehrere Bedeutungen. Es steht für „Hilfszügel beim Zaumzeug“ eines Reitpferdes, welches zu starke Kopfbewegungen des Pferdes verhindert, aber auch für ein die Takelage bei Segelschiffen absicherndes Seil. Vor allem bedeutet es aber wohl eine Spielstrategie beim Roulettespiel, im Provenzalischen genannt „jouga a la martegalo“. Diese Strategie besteht in der jeweiligen Verdoppelung des beim vorausgegangenen Spiel verlorenen Einsatzes, ein Spezialfall des in den Vorbemerkungen betrachteten Beispiels, welches wir in Beispiel 1.1.4(7) genauer betrachten.

◇

Bei der Wahl der kanonischen Filtrierung $(\mathcal{F}_n^{\mathbb{X}})_n$ nennen wir die Filtrierung nicht extra.

Wir sammeln Beispiele:

Beispiel 1.1.4. (1) Die eindimensionale symmetrische Irrfahrt auf $\mathbb{Z}$ mit Zeitachse $\mathbb{N}_0$ ist ein Martingal.

(2) Sei $(X_n)_n$ die symmetrische Irrfahrt auf $\mathbb{N}$ mit Start in 1 und Absorption in 0. Also $p_{i,i+1} = p_{i,i-1} = 1/2$ für $i \in \mathbb{N}$ und $p_{0,0} = 1$. Also gilt für $n \in \mathbb{N}_0$ fast sicher

$$\mathbb{E}(X_{n+1}|\mathcal{F}_n^{\mathbb{X}}) = \mathbb{E}(X_{n+1}|X_1, \dots, X_n) = \sum_{j\in\mathbb{N}_0} j p_{X_n,j} = X_n\,,$$

also ist diese Irrfahrt ein Martingal.

(3) In den Vorbemerkungen haben wir schon die Verallgemeinerung von (1) gesehen: Seien $\xi_1, \xi_2, \dots$ unabhängige und identisch verteilte, integrierbare Zufallsgrößen mit $\mathbb{E}(\xi_j) = 0$. Wir setzen $X_0 := 0$ und $X_n := \sum_{j=1}^n \xi_j$ für $n \in \mathbb{N}$, und $\mathcal{F}_0 := \{\emptyset, \Omega\}$ und $\mathcal{F}_n := \sigma(\xi_1, \dots, \xi_n)$. Dann ist $(X_n)_n$ ein $(\mathcal{F}_n)_n$-Martingal. Mit anderen Worten: Jede Irrfahrt mit zentrierten Schritten ist ein Martingal bezüglich seiner natürlichen Filtrierung.

(4) (*Levys Martingal*) Es sei $(\mathcal{F}_n)_n$ eine beliebige Filtrierung und $X \in \mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$. Dann ist die Folge der bedingten Erwartungswerte

$$X_n := \mathbb{E}(X|\mathcal{F}_n)\,, \qquad n \in \mathbb{N}\,,$$

ein Martingal, denn es gilt $\mathbb{P}$-fast sicher

$$\mathbb{E}(X_{n+1}|\mathcal{F}_n) = \mathbb{E}\big(\mathbb{E}(X|\mathcal{F}_{n+1})|\mathcal{F}_n\big) = \mathbb{E}(X|\mathcal{F}_n)\,.$$

Wegen der Einfachheit und leichten Handhabbarkeit dieses Typs von Martingalen fragt man sich vielleicht, ob jedes Martingal in dieser Weise dargestellt werden kann. Dies ist leider falsch, wie wir noch sehen werden (aber in einigen Fällen richtig).

(5) (*exponentielles Martingal*) Die $\xi_1, \xi_2, \dots$ und die Filtrierung seien wie in (3) gewählt. Es sei $\mathbb{E}\xi_j \neq 0$ zugelassen. Es existiere ein $\lambda_0 > 0$, so dass die Momenten erzeugende Funktion $M(\lambda) := \mathbb{E}(\mathrm{e}^{\lambda\xi_j})$ endlich sei für $|\lambda| \leq \lambda_0$. Sei $X_0 = 1$ und

$$X_n := \exp\Big(\lambda \sum_{j=1}^n \xi_j\Big) M(\lambda)^{-n}, \qquad n \in \mathbb{N}.$$

Dann ist $(X_n)_n$ ein $(\mathcal{F}_n)_n$-Martingal, denn es gilt $\mathbb{P}$-fast sicher

$$\begin{aligned}\mathbb{E}(X_{n+1}|\mathcal{F}_n) &= M(\lambda)^{-n-1} \exp\Big(\lambda \sum_{j=1}^n \xi_j\Big) \mathbb{E}(e^{\lambda\xi_{n+1}}|\mathcal{F}_n) \\ &= M(\lambda)^{-n-1} \exp\Big(\lambda \sum_{j=1}^n \xi_j\Big) \mathbb{E}(e^{\lambda\xi_{n+1}}) \\ &= M(\lambda)^{-n} \exp\Big(\lambda \sum_{j=1}^n \xi_j\Big) = X_n\,.\end{aligned}$$

(6) Durch Differentiation nach λ in $\lambda = 0$ lassen sich aus Beispiel (5) neue Martingale gewinnen. Wir erinnern uns, dass $M'(0) = \mathbb{E}[\xi_j]$ und $M''(0) = \mathbb{E}[\xi_j^2]$ gelten. Einmalige Differentiation liefert Beispiel (3), und zweimaliges Differenzieren liefert für $S_n := \sum_{j=1}^n \xi_j$:

$$\frac{\mathrm{d}^2}{\mathrm{d}\lambda^2} X_n\Big|_{\lambda=0} = S_n^2 - 2nS_n M'(0) + n(n+1)M'(0)^2 - nM''(0).$$

Die rechte Seite ist ein $(\mathcal{F}_n)_n$-Martingal (Übungsaufgabe).

(7) (*Martingal-Wettstrategie*) Sei $(X_n)_n$ die Folge ± 1-wertiger unabhängiger Zufallsgrößen mit $\mathbb{P}(X_i = 1) = 1/2$. Der Spieler setzt eine Einheit Einsatz zu Beginn und verdoppelt seinen Einsatz so lange, bis zum ersten Mal 1 erscheint, um dann aufzuhören. M_n bezeichne den Gewinn nach n Würfen mit $M_0 = 0$. Im Fall eines Gewinns hört der Spieler auf, also $\mathbb{P}(M_{n+1} = 1 | M_n = 1) = 1$. Erscheint n Mal stets -1, beträgt der Verlust $1 + 2 + 4 + \cdots + 2^{n-1} = 2^n - 1$ Einheiten. Dann wird der Einsatz auf 2^n verdoppelt. Es gilt:

$$\mathbb{P}\big(M_{n+1} = 1 | M_n = -(2^n - 1)\big) = \mathbb{P}(X_{n+1} = 1) = \frac{1}{2}$$

und

$$\mathbb{P}\big(M_{n+1} = -(2^{n+1} - 1) | M_n = -(2^n - 1)\big) = \mathbb{P}(X_{n+1} = -1) = \frac{1}{2},$$

also

$$\mathbb{E}\big(M_{n+1} | M_n = -(2^n - 1)\big) = 1 \cdot \frac{1}{2} + (-2^{n+1} + 1) \cdot \frac{1}{2} = -2^n + 1,$$

also ist $(M_n)_n$ ein Martingal bezüglich $(\mathcal{F}_n)_n$ mit $\mathcal{F}_n := \sigma(M_0, X_1, \dots, X_n)$.

$\diamond$

Wir sammeln ein paar einfache, aber nützliche Eigenschaften von Martingalen:

Lemma 1.1.5. *Sei $(X_n)_n$ ein $(\mathcal{F}_n)_n$-Martingal (bzw. Submartingal).*

(i) Ist $(\mathcal{B}_n)_n$ eine weitere Filtrierung von $\mathcal{F}$ mit $\mathcal{B}_n \subset \mathcal{F}_n$ für alle $n \in \mathbb{N}$ und ist $(X_n)_n$ an $(\mathcal{B}_n)_n$ adaptiert, so ist $(X_n)_n$ auch ein $(\mathcal{B}_n)_n$-Martingal (bzw. Submartingal)

(ii) Es gilt $\mathbb{E}(X_n) = \mathbb{E}(X_m)$ (bzw. $\mathbb{E}(X_n) \geq \mathbb{E}(X_m)$) für alle $m, n \in \mathbb{N}$ mit $m \leq n$.

(iii) Ist φ eine konvexe Funktion mit $\mathbb{E}\,|\varphi(X_n)| < \infty$, $n \in \mathbb{N}$, und $(X_n)_n$ ein $(\mathcal{F}_n)_n$-Martingal, so ist $\big(\varphi(X_n)\big)_n$ ein Submartingal bezüglich $(\mathcal{F}_n)_n$.

Beweis. (i) und (ii) folgen leicht aus Definition 1.1.2. Für den Beweis von (iii) verwende man die Jensen'sche Ungleichung für bedingte Erwartungen. □

1.2 Stoppzeiten und Optional Sampling

Martingale beschreiben im Sinne der einführenden Beispiele faire Spiele. Beispiel 1.1.4(7) beschreibt eine Strategie, mit der in einem fairen Spiel Gewinn gemacht werden kann: Sei $\tau = \inf\{n \geq 0 \colon M_n = 1\}$ die Wartezeit auf den ersten Gewinn, dann gilt stets $M_\tau = 1$, während $M_0 = 0$ ist. Also hat man zu der zufälligen Zeit τ einen Gewinn erzielt. Unter gewissen Einschränkungen an eine zufällige Zeit τ werden wir aber sehen, dass man in einem fairen Spiel keinen Gewinn zum Zeitpunkt τ machen kann. Dies wird der Inhalt des sogenannten *Optional-Stopping-Theorems* werden. Diejenigen Zufallszeiten, die wir betrachten werden, werden fast alle die folgende Abhängigkeitsstruktur aufweisen:

Definition 1.2.1 (Stoppzeit). *Sei $(\mathcal{F}_n)_n$ eine Filtrierung von $\mathcal{F}$. Eine Abbildung $\tau \colon \Omega \to \mathbb{N} \cup \{+\infty\}$ heißt* Stoppzeit *bezüglich $(\mathcal{F}_n)_n$, falls $\{\tau \leq n\} \in \mathcal{F}_n$ für alle $n \in \mathbb{N}$ gilt.*

Wir schreiben im Folgenden $\overline{\mathbb{N}} = \mathbb{N} \cup \{+\infty\}$.

Bemerkung 1.2.2. (i) $\tau \colon \Omega \to \overline{\mathbb{N}}$ ist genau dann eine Stoppzeit, wenn $\{\tau = n\} \in \mathcal{F}_n$ für alle $n \in \mathbb{N}$ gilt. Denn

$$\{\tau = n\} = \{\tau \leq n\} \backslash \{\tau \leq n-1\} \in \mathcal{F}_n \,,$$

falls τ eine Stoppzeit ist. Außerdem ist $\{\tau \leq n\} = \bigcup_{m=1}^n \{\tau = m\} \in \mathcal{F}_n$ für jedes $n \in \mathbb{N}$, falls alle $\{\tau = n\}$ in $\mathcal{F}_n$ liegen.

(ii) Ist τ eine Stoppzeit, so gilt

$$\{\tau = \infty\} = \Big(\bigcup_{n \in \mathbb{N}} \{\tau \leq n\}\Big)^c \in \sigma\Big(\bigcup_{n \in \mathbb{N}} \mathcal{F}_n\Big) \,.$$

◇

Beispiel 1.2.3. (i) Konstante Abbildungen $\tau \colon \Omega \to \overline{\mathbb{N}}$ sind Stoppzeiten.

(ii) $(X_n)_{n \in \mathbb{N}}$ sei eine Folge von Zufallsvariablen mit Werten in einem messbaren Raum $(E, \mathcal{E})$, und $\mathcal{F}_n = \sigma(X_1, \ldots, X_n)$ sei die kanonische Filtrierung. Sei $A \in \mathcal{E}$, dann ist die Ersteintrittszeit in A,

$$\tau_A = \inf\big\{n \in \mathbb{N} \colon X_n \in A\big\} \,.$$

eine Stoppzeit, wobei wir wie üblich $\inf \emptyset = \infty$ setzen, denn

$$\{\tau_A \le n\} = \bigcup_{m=1}^{n} \{X_m \in A\} \in \mathcal{F}_n\,, \qquad n \in \mathbb{N}\,.$$

◇

Für jede Stoppzeit τ setzen wir

$$\mathcal{F}_\tau := \Big\{ A \in \sigma\Big(\bigcup_{n\in\mathbb{N}} \mathcal{F}_n\Big) \,\Big|\, A \cap \{\tau \le n\} \in \mathcal{F}_n \text{ für alle } n \in \mathbb{N} \Big\}\,.$$

Anschaulich enthält dieses Mengensystem diejenigen Ereignisse, deren Eintreffen oder Nichteintreffen mit der Information, die die σ-Algebra $\mathcal{F}_n$ bis zum Zeitpunkt $n = \tau$ enthält, entschieden werden kann. Man zeigt leicht, dass $\mathcal{F}_\tau$ eine σ-Algebra ist und dass $\mathcal{F}_\tau = \mathcal{F}_n$ ist, falls $\tau = n$ die konstante Stoppzeit ist. Weitere einfache Eigenschaften von Stoppzeiten werden nun gesammelt:

Lemma 1.2.4. *Es seien τ und σ zwei Stoppzeiten.*

(i) Das Minimum $\tau \wedge \sigma := \min\{\tau, \sigma\}$ und das Maximum $\tau \vee \sigma := \max\{\tau, \sigma\}$ sind Stoppzeiten.

(ii) Gilt $\sigma(\omega) \le \tau(\omega)$ für alle $\omega \in \Omega$, so gilt $\mathcal{F}_\sigma \subset \mathcal{F}_\tau$.

(iii) τ ist $\mathcal{F}_\tau$-messbar.

(iv) Für jedes $A \in \mathcal{F}_\sigma$ liegt das Ereignis $A \cap \{\sigma \le \tau\}$ in $\mathcal{F}_\sigma \cap \mathcal{F}_\tau$. Insbesondere liegt $\{\sigma \le \tau\}$ in $\mathcal{F}_\sigma \cap \mathcal{F}_\tau$.

(v) Ist $\tau' \colon \Omega \to \overline{\mathbb{N}}$ eine $\mathcal{F}_\tau$-messbare Abbildung und gilt $\tau' \ge \tau$, so ist τ' eine Stoppzeit.

(vi) $\tau + \sigma$ ist eine Stoppzeit.

Beweis.

(i) Für jedes $n \in \mathbb{N}$ gelten $\{\tau \vee \sigma \le n\} = \{\tau \le n\} \cap \{\sigma \le n\}$ und $\{\tau \wedge \sigma \le n\} = \{\tau \le n\} \cup \{\sigma \le n\}$.

(ii) Sei $A \in \mathcal{F}_\sigma$. Mit $\{\tau \le n\} \subset \{\sigma \le n\}$ und $A \cap \{\sigma \le n\} \in \mathcal{F}_n$ folgt $A \cap \{\tau \le n\} = \big(A \cap \{\sigma \le n\}\big) \cap \{\tau \le n\} \in \mathcal{F}_n$ für jedes $n \in \mathbb{N}$.

(iii) Für $n, m \in \mathbb{N}$ gilt

$$\{\tau \le n\} \cap \{\tau \le m\} = \{\tau \le n \wedge m\} \in \mathcal{F}_{n \wedge m} \subset \mathcal{F}_m\,.$$

Dies gilt für alle m, also $\{\tau \le n\} \in \mathcal{F}_\tau$ für alle n. Damit ist τ $\mathcal{F}_\tau$-messbar.

(iv) Wir zeigen: $A \cap \{\sigma \le \tau\} \in \mathcal{F}_{\sigma\wedge\tau}$ für alle $A \in \mathcal{F}_\sigma$. Mit (i) und (ii) ist $\mathcal{F}_{\sigma\wedge\tau} \subset \mathcal{F}_\sigma \cap \mathcal{F}_\tau$. Für $n \in \mathbb{N}$ ist

$$\begin{aligned} A \cap \{\sigma \le \tau\} \cap \{\sigma \wedge \tau \le n\} &= A \cap \{\sigma \le \tau\} \cap \{\sigma \le n\} \\ &= A \cap \{\sigma \le n\} \cap \bigcup_{m=1}^{n} \{\sigma \le m\} \cap \{\tau \ge m\} \in \mathcal{F}_n\,, \end{aligned}$$

denn $A \cap \{\sigma \le n\} \in \mathcal{F}_n$ und $\{\sigma \le m\} \cap \{\tau \ge m\} \in \mathcal{F}_m \subset \mathcal{F}_n$. Also $A \cap \{\sigma \le \tau\} \in \mathcal{F}_{\sigma\wedge\tau}$.

(v) Für $n \in \mathbb{N}$ gilt $\{\tau' \le n\} = \{\tau' \le n\} \cap \{\tau \le n\} \in \mathcal{F}_n$, da $\{\tau' \le n\}$ nach Voraussetzung in $\mathcal{F}_\tau$ ist.

(vi)
$$\{\tau + \sigma \le n\} = \bigcup_{\substack{k,l \in \mathbb{N}:\\ k+l \le n}} \{\tau \le k\} \cap \{\sigma \le l\} \in \mathcal{F}_n.$$

□

Wir möchten gerne die Martingaleigenschaft $X_n = \mathbb{E}[X_m \mid \mathcal{F}_n]$ (für $n \le m$) auch an Stoppzeiten an Stelle von n und m haben. Als einen ersten Schritt ersetzen wir n durch eine Stoppzeit:

Lemma 1.2.5. *Sei $(X_n)_{n \in \mathbb{N}_0}$ ein Martingal, $T \in \mathbb{N}$ und τ eine Stoppzeit mit $\tau \le T$ fast sicher. Dann gilt $X_\tau = \mathbb{E}[X_T \mid \mathcal{F}_\tau]$ und insbesondere $\mathbb{E}[X_\tau] = \mathbb{E}[X_0]$.*

Beweis. Es ist leicht zu sehen, dass X_τ messbar bezüglich $\mathcal{F}_\tau$ ist. Wir müssen also nur noch zeigen, dass für jedes $A \in \mathcal{F}_\tau$ gilt: $\mathbb{E}[X_T \mathbb{1}_A] = \mathbb{E}[X_\tau \mathbb{1}_A]$. Wir vergessen nicht, dass $\{\tau = n\} \cap A \in \mathcal{F}_n$ für jedes $n \in \mathbb{N}$ gilt, und errechnen

$$\mathbb{E}[X_\tau \mathbb{1}_A] = \sum_{n=0}^{T} \mathbb{E}[X_\tau \mathbb{1}_{\{\tau=n\}} \mathbb{1}_A] = \sum_{n=0}^{T} \mathbb{E}\big[\mathbb{E}[X_T \mid \mathcal{F}_n] \mathbb{1}_{\{\tau=n\}\cap A}\big] = \sum_{n=0}^{T} \mathbb{E}[X_T \mathbb{1}_A \mathbb{1}_{\{\tau=n\}}] = \mathbb{E}[X_T \mathbb{1}_A].$$

□

Nun können wir die Martingaleigenschaft auch für Stoppzeiten erhalten:

Satz 1.2.6 (Optional-Sampling-Theorem). *Es seien $X = (X_n)_{n \in \mathbb{N}_0}$ ein Martingal und $\sigma \le \tau$ zwei Stoppzeiten.*

(i) Falls es ein $T \in \mathbb{N}$ gibt mit $\tau \le T$ fast sicher, dann ist $X_\sigma = \mathbb{E}[X_\tau \mid \mathcal{F}_\sigma]$, also insbesondere $\mathbb{E}[X_\sigma] = \mathbb{E}[X_\tau]$.

(ii) Falls $X_n \ge 0$ für jedes $n \in \mathbb{N}_0$ und $\tau < \infty$ fast sicher, so gelten $\mathbb{E}[X_\tau] \le \mathbb{E}[X_0]$, $\mathbb{E}[X_\sigma] \le \mathbb{E}[X_0]$ und $X_\sigma \ge \mathbb{E}[X_\tau \mid \mathcal{F}_\sigma]$ fast sicher.

Beweis. (i) Wir benutzen Lemma 1.2.5, erinnern uns daran, dass $\mathcal{F}_\sigma \subset \mathcal{F}_\tau$ gilt (siehe Lemma 1.2.4(ii)), und sehen mit Hilfe der Projektionseigenschaft der bedingten Erwartung, dass fast sicher gilt:
$$\mathbb{E}[X_\tau \mid \mathcal{F}_\sigma] = \mathbb{E}\big[\mathbb{E}[X_T \mid \mathcal{F}_\tau] \mid \mathcal{F}_\sigma\big] = \mathbb{E}[X_T \mid \mathcal{F}_\sigma] = X_\sigma.$$

(ii) Für jedes $n \in \mathbb{N}_0$ ist $\tau \wedge n$ eine Stoppzeit, also liefert (i) (mit $\sigma = 0$), dass $\mathbb{E}[X_{\tau\wedge n}] = \mathbb{E}[X_0]$ gilt. Da $\lim_{n\to\infty} X_{\tau\wedge n} = X_\tau$ fast sicher gilt, liefert das Lemma von Fatou die Aussage $\mathbb{E}[X_\tau] = \mathbb{E}[\lim_{n\to\infty} X_{\tau\wedge n}] \le \liminf_{n\to\infty} \mathbb{E}[X_{\tau\wedge n}] = \mathbb{E}[X_0]$. Analog gilt $\mathbb{E}[X_\sigma] \le \mathbb{E}[X_0]$. Für den Beweis der letzten Aussage wählen wir $n, m \in \mathbb{N}_0$ mit $n \le m$, wenden (i) auf die Stoppzeiten $\sigma \wedge n \le \tau \wedge m$ an und erhalten $X_{\sigma\wedge n} = \mathbb{E}[X_{\tau\wedge m} \mid \mathcal{F}_{\sigma\wedge n}]$. Für $A \in \mathcal{F}_\sigma$ ist $A \cap \{\sigma < n\} \in \mathcal{F}_{\sigma\wedge n}$, also ist
$$\mathbb{E}[X_\sigma \mathbb{1}_{A\cap\{\sigma<n\}}] = \mathbb{E}[X_{\sigma\wedge n} \mathbb{1}_{A\cap\{\sigma<n\}}] = \mathbb{E}[X_{\tau\wedge m} \mathbb{1}_{A\cap\{\sigma<n\}}].$$

Das Lemma von Fatou liefert

$$\mathbb{E}[X_\tau \mathbb{1}_{A\cap\{\sigma<n\}}] \leq \liminf_{m\to\infty} \mathbb{E}[X_{\tau\wedge m} \mathbb{1}_{A\cap\{\sigma<n\}}] = \mathbb{E}[X_\sigma \mathbb{1}_{A\cap\{\sigma<n\}}].$$

Nun folgt die Aussage durch den Grenzübergang mit Hilfe des Satzes von der monotonen Konvergenz. □

1.3 Das Optional-Stopping-Theorem

Im Folgenden wollen wir das Beispiel 1.1.1 formalisieren. Wir erinnern uns an die Interpretation von X_n als der Gewinn nach dem n-ten Glücksspiel, Y_n der Ausgang des n-ten Spiels und $e_n(Y_1, \dots, Y_n)$ als der Einsatz im $(n+1)$-ten Spiel. Im Folgenden wird Y_n nicht mehr explizit auftauchen, sondern statt dieser Folge werden wir eine Filtrierung $(\mathcal{F}_n)_{n\in\mathbb{N}_0}$ von $\mathcal{F}$ betrachten, die der 'Vorrat' von Ereignissen ist, die bis zu einem Zeitpunkt $n \in \mathbb{N}$ eintreten können.

Definition 1.3.1 (Vorhersehbarkeit, Martingaltransformation). *(i) Eine Folge $\mathbb{V} := (V_n)_{n\in\mathbb{N}}$ von Zufallsgrößen heißt* vorhersehbar *bezüglich $(\mathcal{F}_n)_{n\in\mathbb{N}_0}$, wenn V_n für jedes $n \in \mathbb{N}$ bezüglich $\mathcal{F}_{n-1}$ messbar ist.*

(ii) Sind $\mathbb{V}$ und $\mathbb{X} = (X_n)_{n\in\mathbb{N}_0}$ zwei Folgen von Zufallsgrößen, so definieren wir eine Folge $\mathbb{V} \bullet \mathbb{X} = (Y_n)_{n\in\mathbb{N}_0}$ durch $Y_0 = 0$ und $Y_n := \sum_{k=1}^n V_k(X_k - X_{k-1})$ für alle $n \in \mathbb{N}$.

Wenn $\mathbb{X}$ ein Martingal ist, so heißt $\mathbb{V} \bullet \mathbb{X}$ die Martingaltransformation *von $\mathbb{X}$ unter $\mathbb{V}$, siehe Lemma 1.3.2(i).*

Die Zuwächse (Martingaldifferenzen) $X_k - X_{k-1}$ spielen in Beispiel 1.1.1 die Rolle der Spielausgänge Y_k, die Gewichtungen V_k die der Einsätze e_{k-1}, die ja von der bisherigen 'Geschichte' bis zum Zeitpunkt $k-1$ abhängen.

Lemma 1.3.2. *(i) Sei $\mathbb{X} = (X_n)_{n\in\mathbb{N}_0}$ ein Supermartingal und $\mathbb{V} = (V_n)_{n\in\mathbb{N}}$ ein vorhersehbarer Prozess mit $V_n \geq 0$ und $\|V_n\|_\infty < \infty$ für alle $n \in \mathbb{N}$. Dann ist $\mathbb{V}\bullet\mathbb{X}$ ein Supermartingal. Ist $\mathbb{X}$ ein Martingal, so auch $\mathbb{V} \bullet \mathbb{X}$.*

(ii) In (i) kann die Voraussetzung $\|V_n\|_\infty < \infty$ ersetzt werden durch $\|V_n\|_2 < \infty$, falls $\|X_n\|_2 < \infty$ für alle $n \in \mathbb{N}_0$ gilt.

Beweis. Sei $\mathbb{V} \bullet \mathbb{X} = (Y_n)_{n\in\mathbb{N}_0}$. Für $n \in \mathbb{N}$ folgt unter den Voraussetzungen in (i) oder (ii) die Integrierbarkeit von Y_n. Weiter gilt $\mathbb{P}$-fast sicher

$$\mathbb{E}(Y_n|\mathcal{F}_{n-1}) = Y_{n-1} + V_n\mathbb{E}(X_n - X_{n-1}|\mathcal{F}_{n-1}) \leq Y_{n-1}\,.$$

Die letzte Ungleichung ist eine Gleichung, falls $\mathbb{X}$ ein Martingal ist. □

Wir untersuchen nun den wichtigen Spezialfall von Martingaltransformationen mittels einer Stoppzeit τ. Für $n \in \mathbb{N}$ sei $V_n^\tau = \mathbb{1}_{\{n\leq\tau\}}$. Da $\{n \leq \tau\} = \{\tau \leq n-1\}^{\mathrm{c}} \in \mathcal{F}_{n-1}$, ist $\mathbb{V}^\tau = (V_n^\tau)_{n\in\mathbb{N}}$ vorhersehbar. Für jede Folge $(X_n)_{n\in\mathbb{N}_0}$ gilt

$$(\mathbb{V}^\tau \bullet \mathbb{X})_n = \sum_{k=1}^n \mathbb{1}_{\{\tau>k-1\}}(X_k - X_{k-1}) = \sum_{k=1}^{\tau\wedge n}(X_k - X_{k-1}) = X_{\tau\wedge n} - X_0\,. \tag{1.3.1}$$

Wir definieren den zur Zeit τ *gestoppten Prozess* $\mathbb{X}^\tau$ durch $\mathbb{X}^\tau = (X_{\tau\wedge n})_{n\in\mathbb{N}}$. Also ergibt eine Anwendung von Lemma 1.3.2 den folgenden berühmten Satz.

Satz 1.3.3 (Optional-Stopping-Theorem). *Es seien $\mathbb{X}$ ein Supermartingal und τ eine Stoppzeit. Dann ist $\mathbb{X}^\tau$ ein Supermartingal. Ist $\mathbb{X}$ ein Martingal, so auch $\mathbb{X}^\tau$.*

Am Anfang von Abschnitt 1.2 warfen wir die Frage auf, ob man in einem fairen Spiel zu einer Stoppzeit einen Gewinn machen kann, d. h. unter welchen Umständen $\mathbb{E}(X_\tau) \neq \mathbb{E}(X_0)$ gelten kann. Aus dem Optional-Stopping-Theorem folgt immerhin, dass $\mathbb{E}(X_{\tau\wedge n}) = \mathbb{E}(X_0)$ für alle $n \in \mathbb{N}_0$ gilt, wenn $\mathbb{X}$ ein Martingal und τ eine Stoppzeit ist. Aber daraus folgt im Allgemeinen nicht $\mathbb{E}(X_\tau) = \mathbb{E}(X_0)$, siehe etwa Beispiel 1.1.4(7) oder das folgende.

Beispiel 1.3.4. $\mathbb{X} = (X_n)_{n\in\mathbb{N}_0}$ sei die symmetrische Irrfahrt auf $\mathbb{N}_0$ mit Start in 1 und Absorbtion in 0. Nach Beispiel 1.1.4(2) ist $\mathbb{X}$ ein $(\mathcal{F}_n^{\mathbb{X}})_{n\in\mathbb{N}_0}$-Martingal. Sei $\tau = \inf\{n \in \mathbb{N}_0 \colon X_n = 0\}$ die Rückkehrzeit zum Ursprung, dann ist τ eine Stoppzeit bezüglich $(\mathcal{F}_n^{\mathbb{X}})_{n\in\mathbb{N}_0}$. Wegen der Rekurrenz der eindimensionalen Irrfahrt gilt $\mathbb{P}(\tau < \infty) = 1$. Es gilt $\mathbb{E}(X_0) = 1$, also $\mathbb{E}(X_{\tau\wedge n}) = 1$ für alle $n \in \mathbb{N}_0$ mittels Satz 1.3.3, aber offensichtlich ist $\mathbb{E}(X_\tau) = 0$. $\diamond$

Nun folgen hinreichende Kriterien für $\mathbb{E}(X_\tau) = \mathbb{E}(X_0)$.

Satz 1.3.5. *Es seien τ eine Stoppzeit und $(X_n)_{n\in\mathbb{N}_0}$ ein Supermartingal (bzw. ein Martingal), die eine der folgenden Bedingungen erfüllen:*

(i) τ *ist beschränkt,*

(ii) τ *ist $\mathbb{P}$-fast sicher endlich und $(X_n)_{n\in\mathbb{N}_0}$ ist beschränkt.*

(iii) $\mathbb{E}(\tau) < \infty$ *und* $\sup_{n\in\mathbb{N}} \|X_n - X_{n-1}\|_\infty < \infty$.

Dann ist X_τ integrierbar und es gilt

$$\mathbb{E}(X_\tau) \leq \mathbb{E}(X_0) \qquad (\textit{bzw. } \mathbb{E}(X_\tau) = \mathbb{E}(X_0))\,.$$

Beweis. Nach Satz 1.3.3 gilt $\mathbb{E}(X_{\tau\wedge n}) \leq \mathbb{E}(X_0)$ (bzw. $\mathbb{E}(X_{\tau\wedge n}) = \mathbb{E}(X_0)$) für jedes $n \in \mathbb{N}$. Ist (i) erfüllt, so ist $\tau \wedge n = \tau$ für ein genügend großes $n \in \mathbb{N}$. Unter (ii) gilt $\lim_{n\to\infty} \mathbb{E}(X_{\tau\wedge n}) = \mathbb{E}(X_\tau)$ nach dem Satz von der majorisierten Konvergenz. Ist (iii) erfüllt, so existiert ein $K > 0$ mit $\mathbb{P}(|X_n - X_{n-1}| \leq K) = 1$ für alle $n \in \mathbb{N}$. Also $|X_{\tau\wedge n} - X_0| \leq K\tau$ fast sicher. Wegen $\mathbb{E}(\tau) < \infty$ folgt $\lim_{n\to\infty} \mathbb{E}(X_{\tau\wedge n}) = \mathbb{E}(X_\tau)$ ebenfalls nach dem Satz von der majorisierten Konvergenz. $\square$

Korollar 1.3.6. *Seien $(X_n)_{n\in\mathbb{N}_0}$ ein positives Supermartingal und τ eine endliche Stoppzeit. Dann gilt $\mathbb{E}(X_\tau) \leq \mathbb{E}(X_0)$.*

Beweis. Nach Satz 1.3.3 gilt $\mathbb{E}(X_{\tau\wedge n}) \leq \mathbb{E}(X_0)$ und mittels des Lemmas von Fatou folgt

$$\mathbb{E}(X_\tau) = \mathbb{E}\Big(\lim_{n\to\infty} X_{\tau\wedge n}\Big) \leq \lim_{n\to\infty} \mathbb{E}(X_{\tau\wedge n}) \leq \mathbb{E}(X_0)\,. \qquad \square$$

1.4 Konvergenz von Martingalen

Es ist einer der großen Vorzüge von Martingalen, dass ihre Konvergenz unter sehr schwachen Voraussetzungen bewiesen werden kann. Den fundamentalen Satz zu diesem Thema, den Doob'schen Konvergenzsatz, stellen wir nun vor. Als Vorbereitung zum Beweis diskutieren wir zunächst, wie oft ein Martingal ein gegebenes Intervall aufsteigend übersteigt.

Seien $\alpha = (\alpha_n)_{n\in\mathbb{N}_0}$ eine Folge reeller Zahlen und $a, b \in \mathbb{R}$ mit $a < b$. Für $n \in \mathbb{N}_0$ ist die Anzahl der *aufsteigenden Überschreitungen* des Intervalls $[a, b]$ durch die Folge α bis zum Zeitpunkt n definiert durch

$$U_n[a,b](\alpha) := \sup\big\{k \in \mathbb{N} \,|\, \text{es gibt } 0 \le s_1 < t_1 < s_2 < t_2 < \cdots < s_k < t_k \le n \\ \text{mit } \alpha_{s_i} < a < b < \alpha_{t_i} \text{ für alle } i \in \{1,\dots,k\}\big\}\,,$$

wobei $\sup\emptyset := 0$ gesetzt wird.

Natürlich ist $U_n[a,b](\alpha) \le n$, und $U_n[a,b](\alpha)$ ist monoton steigend in n. Wir setzen

$$U_\infty[a,b](\alpha) := \sup_{n\in\mathbb{N}_0} U_n[a,b](\alpha) \in \overline{\mathbb{N}}\,.$$

Die Konvergenz einer Folge α kann man mit Hilfe von $U_\infty[a,b](\alpha)$ entscheiden:

Lemma 1.4.1. *Eine Folge $\alpha = (\alpha_n)_{n\in\mathbb{N}_0}$ konvergiert genau dann in $\overline{\mathbb{R}}$, wenn $U_\infty[a,b](\alpha) < \infty$ für alle $a, b \in \mathbb{Q}$ mit $a < b$ gilt.*

Beweis. Wir argumentieren indirekt. Die Folge α konvergiert genau dann nicht, wenn gilt: $\liminf_{n\to\infty} \alpha_n < \limsup_{n\to\infty} \alpha_n$. Dies ist genau dann der Fall, wenn es rationale Zahlen $a < b$ gibt mit $\liminf_{n\to\infty} \alpha_n < a < b < \limsup_{n\to\infty} \alpha_n$. Die letzteren Ungleichungen sind aber äquivalent zu $U_\infty[a,b](\alpha) = \infty$. □

Also benötigen wir effektive obere Abschätzungen für $U_n[a,b](\mathbb{X})$, wenn $\mathbb{X}$ ein Martingal ist. Eine solche lautet wie folgt. Wir setzen wie üblich $x^- := \max\{0, -x\}$ für alle $x \in \mathbb{R}$.

Satz 1.4.2 (Aufkreuzungsungleichung). *Es sei $\mathbb{X} = (X_n)_{n\in\mathbb{N}_0}$ ein Supermartingal und $a < b$. Dann gilt für jedes $n \in \mathbb{N}_0$:*

$$\mathbb{E}\big(U_n[a,b](\mathbb{X})\big) \le \frac{1}{b-a}\mathbb{E}\big((X_n - a)^-\big)\,.$$

Bevor wir den Beweis geben, ziehen wir Konsequenzen.

Satz 1.4.3 (Doob'scher Konvergenzsatz). *Sei $\mathbb{X} = (X_n)_{n\in\mathbb{N}_0}$ ein $\mathcal{L}^1$-beschränktes Supermartingal (d. h. $\sup_{n\in\mathbb{N}_0} \mathbb{E}|X_n| < \infty$). Dann existiert $X_\infty := \lim_{n\to\infty} X_n$ $\mathbb{P}$-fast sicher und ist integrierbar.*

Beweis. Seien $a < b$ zwei reelle Zahlen. Nach dem Satz von der monotonen Konvergenz gilt

$$(b-a)\mathbb{E}\big(U_\infty[a,b](\mathbb{X})\big) = (b-a)\lim_{n\to\infty} \mathbb{E}\big(U_n[a,b](\mathbb{X})\big)\,.$$

Die rechte Seite kann mit Hilfe von Satz 1.4.2 nach oben durch

$$\sup_{n\in\mathbb{N}_0} \mathbb{E}\big((X_n - a)^-\big) \le |a| + \sup_{n\in\mathbb{N}_0} \mathbb{E}|X_n| < \infty$$

abgeschätzt werden. Also ist $\mathbb{P}\big(U_\infty[a,b](\mathbb{X}) < \infty\big) = 1$, also

$$\mathbb{P}\Big(\bigcap_{\substack{a,b\in\mathbb{Q}\\ a<b}} \big\{U_\infty[a,b](\mathbb{X}) < \infty\big\}\Big) = 1\,.$$

Somit existiert nach Lemma 1.4.1 $X_\infty = \lim_{n\to\infty} X_n$ in $\overline{\mathbb{R}}$ fast sicher. Nach dem Lemma von Fatou ist

$$\mathbb{E}|X_\infty| \le \liminf_{n\to\infty} \mathbb{E}|X_n| \le \sup_{n\in\mathbb{N}_0} \mathbb{E}|X_n| < \infty\,. \qquad \square$$

Korollar 1.4.4. *Jedes nicht-negative Supermartingal* $(X_n)_{n\in\mathbb{N}_0}$ *konvergiert fast sicher gegen eine integrierbare Zufallsgröße.*

Beweis. Aus $X_n \ge 0$ folgt $\mathbb{E}|X_n| = \mathbb{E}(X_n) \le \mathbb{E}(X_0)$, also $\sup_{n\in\mathbb{N}_0} \mathbb{E}|X_n| < \infty$. $\square$

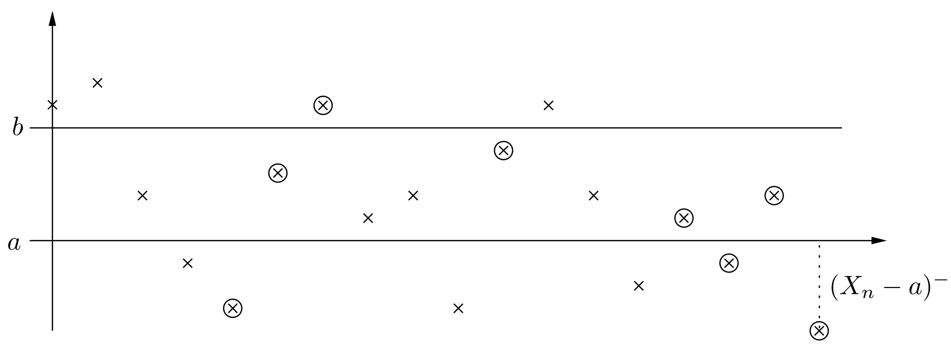
Bemerkung 1.4.5. Die Bedingung $\sup_{n\in\mathbb{N}_0} \mathbb{E}|X_n| < \infty$ reicht im Allgemeinen nicht für $\mathcal{L}^1$-Konvergenz aus. Siehe dazu Beispiel 1.3.4: Dort gilt $\mathbb{E}[X_n] = 1$ für alle $n \in \mathbb{N}$, $X_n \to 0$ fast sicher (siehe Korollar 1.4.4), aber die Erwartungswerte konvergieren nicht gegen Null, also liegt keine $\mathcal{L}^1$-Konvergenz vor. $\diamond$

Beweis von Satz 1.4.2. Wir konstruieren eine Martingaltransformation $\mathbb{Y} = (Y_n)_{n\in\mathbb{N}_0}$ von $\mathbb{X}$. Mit Hilfe der nachfolgenden Regeln nutzt dabei $\mathbb{Y}$ alle aufsteigenden Überschreitungen des Supermartingals $\mathbb{X}$, um möglichst weit nach oben zu gelangen:

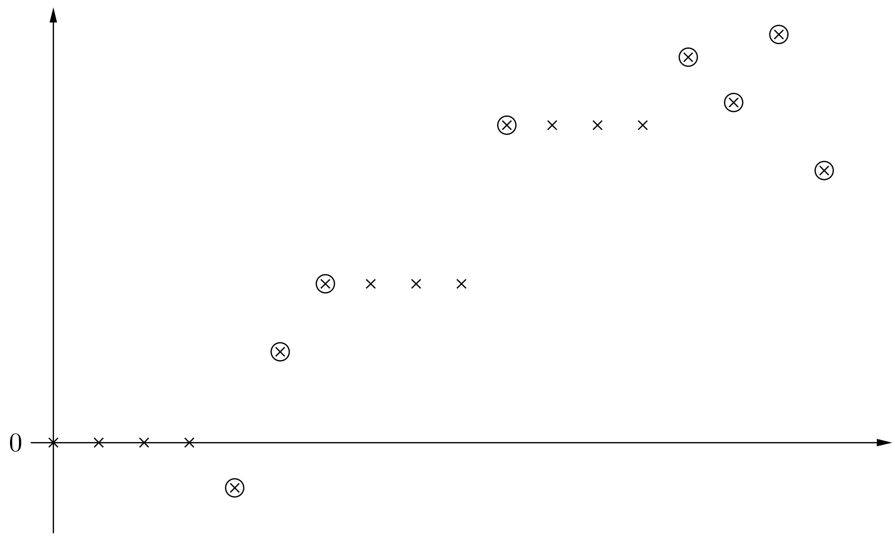
(i) Starte mit $Y_0(\omega) = 0$. Ist $X_0(\omega) \ge a$, benutze Regel (ii), andernfalls Regel (iii).

(ii) Warte so lange, das heißt, setze $Y_n(\omega) = Y_{n-1}(\omega)$, bis $X_n(\omega) < a$ ist. Benutze dann für den nächsten Schritt Regel (iii).

(iii) Nutze die Zuwächse, das heißt, setze $Y_n(\omega) = Y_{n-1}(\omega) + X_n(\omega) - X_{n-1}(\omega)$, bis $X_n(\omega) > b$. Benutze dann für den nächsten Schritt Regel (ii).



Pfad $\mathbb{N}_0 \ni n \mapsto X_n(\omega)$ eines Supermartingals.



Der zugehörige Pfad $\mathbb{N}_0 \ni n \mapsto Y_n(\omega)$ der Martingaltransformation.

Da der Prozess $\mathbb{Y}$ jedes Mal mindestens die Höhe $(b-a)$ gewinnt, wenn $\mathbb{X}$ das Intervall $[a,b]$ aufsteigend überschreitet, und $\mathbb{Y}$ seit der letzten Überschreitung höchstens die Höhe $(X_n - a)^-$ verloren haben kann, gilt für jedes $n \in \mathbb{N}_0$

$$Y_n \geq (b-a)U_n[a,b](\mathbb{X}) - (X_n - a)^- \,. \tag{1.4.1}$$

$\mathbb{Y}$ ist eine Martingaltransformation von $\mathbb{X}$: Der Prozess $\mathbb{V} = (V_n)_{n\in\mathbb{N}}$ mit

$$V_n = \begin{cases} 1\!\!1_{\{X_0<a\}} & \text{für } n = 1\,, \\ 1\!\!1_{\{V_{n-1}=1, X_{n-1}\leq b\}} + 1\!\!1_{\{V_{n-1}=0, X_{n-1}<a\}} & \text{für } n \geq 2\,, \end{cases}$$

ist vorhersehbar, und es gilt

$$Y_n = \sum_{k=1}^{n} V_k(X_k - X_{k-1}), \qquad n \in \mathbb{N}_0.$$

Insbesondere ist $\mathbb{Y}$ ein Supermartingal nach Lemma 1.3.2(i). Mithin gilt $\mathbb{E}(Y_n) \leq \mathbb{E}(Y_0) = 0$ für jedes $n \in \mathbb{N}_0$, und dies liefert mit (1.4.1) die behauptete Ungleichung. $\square$

Beispiel 1.4.6 (Polyas Urnenschema). In einer Urne liegen R_n rote und S_n schwarze Kugeln zum Zeitpunkt $n \in \mathbb{N}_0$. Im Zeitintervall $(n, n+1)$ wird die Urne gut gemischt, eine Kugel zufällig gezogen und zusammen mit einer zusätzlichen Kugel der gleichen Farbe zurückgelegt. Zum Zeitpunkt 0 sei $R_0 = S_0 = 1$. Dann ist $((R_n, S_n))_{n\in\mathbb{N}_0}$ eine Markovkette mit Zustandsraum $\mathbb{N}^2$ und Übergangswahrscheinlichkeiten $p\big((r,s),(r+1,s)\big) = r/(r+s)$ und $p\big((r,s),(r,s+1)\big) = s/(r+s)$ für $(r,s) \in \mathbb{N}^2$. Man beachte, dass für jedes $n \in \mathbb{N}_0$ gilt: $R_n + S_n = n+2$. Wir setzen $\mathcal{F}_n = \sigma((R_1,S_1),\dots,(R_n,S_n))$ und

$$X_n = \frac{R_n}{n+2}\,, \qquad n \in \mathbb{N}_0\,.$$

Dann ist $(X_n)_{n\in\mathbb{N}_0}$ ein $(\mathcal{F}_n)_{n\in\mathbb{N}_0}$-Martingal, denn für jedes $n \in \mathbb{N}_0$ gilt:

$$\mathbb{E}(X_{n+1}|\mathcal{F}_n) = \mathbb{E}\Big(\frac{R_{n+1}}{n+3}\,\Big|\,(R_n, S_n)\Big) = \frac{R_n}{n+2}\frac{R_n+1}{n+3} + \frac{(n+2-R_n)}{n+2}\frac{R_n}{n+3} = \frac{R_n}{n+2} = X_n\,.$$

Es gilt $X_n \geq 0$, also existiert $X_\infty := \lim_{n\to\infty} X_n$ fast sicher nach Korollar 1.4.4.

Etwas Kombinatorik führt zu $\mathbb{P}(R_n = j) = 1/(n+1)$ für jedes $j \in \{1, \dots, n+1\}$. Ist $f\colon [0,1] \to \mathbb{R}$ stetig, so gilt

$$\mathbb{E}f(X_n) = \frac{1}{n+1}\sum_{j=1}^{n+1} f\Big(\frac{j}{n+2}\Big)\,,$$

was für $n \to \infty$ gegen $\int_0^1 f(x)\,\mathrm{d}x$ konvergiert.

Andererseits folgt aus $X_n \to X_\infty$ und dem Satz von Lebesgue von der majorisierten Konvergenz

$$\mathbb{E}f(X_\infty) = \lim_{n\to\infty}\mathbb{E}f(X_n)\,.$$

Also ist $\mathbb{E}f(X_\infty) = \int_0^1 f(x)\,\mathrm{d}x$. Dies gilt für jede stetige Funktion, und somit ist die Verteilung von X_∞ das Lebesgue-Maß auf $[0,1]$. $\diamond$

1.5 $\mathcal{L}^2$-Konvergenz

Wir diskutieren nun die $\mathcal{L}^2$-Konvergenz von Martingalen. Wir werden sehen, dass $\mathcal{L}^2$-Beschränktheit ausreicht, um $\mathcal{L}^2$-Konvergenz zu erhalten. Dies ist also eine stärkere Eigenschaft als bei der $\mathcal{L}^1$-Beschränktheit, die ja nicht für die $\mathcal{L}^1$-Konvergenz hinreicht, siehe Beispiel 1.3.4.

Definition 1.5.1. *Ein Martingal $(X_n)_{n\in\mathbb{N}_0}$ heißt $\mathcal{L}^2$-*Martingal*, wenn für jedes $n \in \mathbb{N}_0$ die Zufallsgröße X_n quadratisch integrierbar ist.*

Satz 1.5.2. *Sei $\mathbb{X} = (X_n)_{n\in\mathbb{N}_0}$ ein $\mathcal{L}^2$-Martingal. Folgende Aussagen sind äquivalent:*

(i) $\sup_{n\in\mathbb{N}_0}\mathbb{E}(X_n^2) < \infty$,

(ii) $\sum_{k=1}^\infty \mathbb{E}\big((X_k - X_{k-1})^2\big) < \infty$,

(iii) $(X_n)_{n\in\mathbb{N}_0}$ *konvergiert $\mathbb{P}$-fast sicher und in $\mathcal{L}^2$.*

Beweis. Quadratisch integrierbare Martingale haben unkorrelierte Zuwächse: Für alle $m, n \in \mathbb{N}_0$ mit $m < n$ gilt

$$\mathbb{E}\big((X_n - X_m)^2\big) = \sum_{k=m+1}^{n}\mathbb{E}\big((X_k - X_{k-1})^2\big)\,. \tag{1.5.1}$$

Wir zeigen (1.5.1) via Induktion nach n: Für $n = m+1$ ist (1.5.1) klar und für $n \geq m+2$ gilt

$$\begin{aligned}\mathbb{E}\big((X_n - X_m)^2\big) &= \mathbb{E}\big((X_n - X_{n-1})^2\big) + \mathbb{E}\big((X_{n-1} - X_m)^2\big)\\ &\quad + 2\mathbb{E}\big((X_n - X_{n-1})(X_{n-1} - X_m)\big)\,.\end{aligned}$$

Für den dritten Summanden liefert das Einschieben eines bedingten Erwartungswertes:

$$\begin{aligned}\mathbb{E}\big((X_n - X_{n-1})(X_{n-1} - X_m)\big) &= \mathbb{E}\Big(\mathbb{E}\big((X_n - X_{n-1})(X_{n-1} - X_m)|\mathcal{F}_{n-1}\big)\Big)\\ &= \mathbb{E}\big((X_{n-1} - X_m)\mathbb{E}(X_n - X_{n-1}|\mathcal{F}_{n-1})\big) = 0.\end{aligned}$$

Mit der Induktionsvoraussetzung für $n-1$ folgt (1.5.1) für n.

Für jedes $n \in \mathbb{N}_0$ folgt aus $\mathbb{E}(X_n\, X_0) = \mathbb{E}\big(\mathbb{E}(X_n|\mathcal{F}_0)X_0\big) = \mathbb{E}(X_0^2)$ und (1.5.1):

$$\mathbb{E}(X_n^2) - \mathbb{E}(X_0^2) = \mathbb{E}\big((X_n - X_0)^2\big) = \sum_{k=1}^{n} \mathbb{E}\big((X_k - X_{k-1})^2\big)\,,$$

also die Äquivalenz von (i) und (ii). (i) folgt aus der $\mathcal{L}^2$-Konvergenz in (iii).

Wir zeigen noch, dass (iii) aus (i) und (ii) folgt: Aus (i) folgt $\sup_{n\in\mathbb{N}_0} \mathbb{E}|X_n| < \infty$ und somit mit Satz 1.4.3, dass $X_\infty := \lim_{n\to\infty} X_n$ $\mathbb{P}$-fast sicher existiert. Aus dem Lemma von Fatou und (1.5.1) folgt

$$\mathbb{E}\big((X_\infty - X_m)^2\big) \le \liminf_{n\to\infty} \mathbb{E}\big((X_n - X_m)^2\big) = \sum_{k=m+1}^{\infty} \mathbb{E}\big((X_k - X_{k-1})^2\big)\,,$$

was nach (ii) für $m \to \infty$ gegen Null konvergiert. □

1.6 $\mathcal{L}^1$-Konvergenz und gleichgradige Integrierbarkeit

Wir diskutieren nun, unter welchen Zusatzbedingungen neben der $\mathcal{L}^1$-Beschränktheit ein Martingal in $\mathcal{L}^1$ konvergiert. Diese Untersuchung werden wir mit dem folgenden wichtigen Begriff kombinieren.

Definition 1.6.1. *Eine Teilmenge Γ von $\mathcal{L}^1 = \mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P})$ heißt* gleichgradig integrierbar*, falls*

$$\lim_{n\to\infty} \sup_{X\in\Gamma} \mathbb{E}\big(|X| 1\!\!1_{\{|X|\ge n\}}\big) = 0.$$

In der Literatur wird oft die Notation $\mathbb{E}(X; A) = \mathbb{E}(X 1\!\!1_A)$ für eine integrierbare Zufallsgröße X und ein Ereignis A benutzt, aber wir benutzen sie im Folgenden nicht.

Lemma 1.6.2. *Es sei $X \in \mathcal{L}^1$. Dann gilt*

$$\lim_{\varepsilon\downarrow 0} \sup\big\{\mathbb{E}(|X| 1\!\!1_A) | A \in \mathcal{F}\,,\, \mathbb{P}(A) \le \varepsilon\big\} = 0\,.$$

Beweis. Angenommen, es existiert eine Folge $(A_n)_{n\in\mathbb{N}}$ in $\mathcal{F}$ mit $\mathbb{P}(A_n) \le 2^{-n}$ für alle $n \in \mathbb{N}$ und $\limsup_{n\to\infty} \mathbb{E}(|X| 1\!\!1_{A_n}) > 0$. Für jedes $n \in \mathbb{N}$ sei $B_n := \bigcup_{k=n}^{\infty} A_k$. Dann gilt $B_n \downarrow A$ für $n \to \infty$ mit einem $A \in \mathcal{F}$, das das Maß $\mathbb{P}(A) = \lim_{n\to\infty} \mathbb{P}(B_n) \le \lim_{n\to\infty} \sum_{k=n}^{\infty} 2^{-k} = 0$ hat. Mit dem Satz von der majorisierten Konvergenz folgt

$$\limsup_{n\to\infty} \mathbb{E}(|X| 1\!\!1_{A_n}) \le \lim_{n\to\infty} \mathbb{E}(|X| 1\!\!1_{B_n}) = 0\,.$$

Dies ist ein Widerspruch. □

Lemma 1.6.3. *Jede der folgenden Bedingungen ist hinreichend für die gleichgradige Integrierbarkeit einer Familie $\Gamma \subset \mathcal{L}^1$:*

(i) Es existiert $p \in (1, \infty)$ mit $\sup_{X\in\Gamma} \|X\|_p < \infty$.

(ii) Es existiert eine Zufallsgröße $Y \in \mathcal{L}^1$ mit $|X| \leq Y$ fast sicher für alle $X \in \Gamma$.

(iii) Es existiert $Y \in \mathcal{L}^1$ und eine Familie Φ von Teil-σ-Algebren von $\mathcal{F}$ mit

$$\Gamma = \{\mathbb{E}(Y|\mathcal{G})\colon \mathcal{G} \in \Phi\}\,.$$

Beweis. *(i):* Für alle $n \in \mathbb{N}$ und $X \in \Gamma$ gilt

$$\mathbb{E}(|X|\mathbb{1}_{\{|X|\geq n\}}) \leq \mathbb{E}\Big(|X|\frac{|X|^{p-1}}{n^{p-1}}\mathbb{1}_{\{|X|\geq n\}}\Big) \leq \frac{\|X\|_p^p}{n^{p-1}} \leq \frac{1}{n^{p-1}}\sup_{X\in\Gamma}\|X\|_p^p,$$

und dies konvergiert gegen Null für $n \to \infty$.

(ii): Für alle $X \in \Gamma$ und $n \in \mathbb{N}$ gilt

$$\mathbb{E}(|X|\mathbb{1}_{\{|X|\geq n\}}) \leq \mathbb{E}(Y\mathbb{1}_{\{Y\geq n\}})\,,$$

und dies konvergiert nach dem Satz von der majorisierten Konvergenz gegen Null für $n \to \infty$.

(iii): Für $n \in \mathbb{N}$ und $\mathcal{G} \in \Phi$ sei

$$A_n(\mathcal{G}) := \big\{|\mathbb{E}(Y|\mathcal{G})| \geq n\big\}\,.$$

Es gilt $|\mathbb{E}(Y|\mathcal{G})| \leq \mathbb{E}(|Y|\,|\,\mathcal{G})$. Somit liefert die Markov'sche Ungleichung

$$\mathbb{P}\big(A_n(\mathcal{G})\big) \leq \frac{1}{n}\mathbb{E}\big(|\mathbb{E}(Y|\mathcal{G})|\big) \leq \frac{1}{n}\mathbb{E}\big(\mathbb{E}(|Y|\,|\,\mathcal{G})\big) = \frac{1}{n}\mathbb{E}(|Y|)\,.$$

Also gilt $\lim_{n\to\infty}\sup_{\mathcal{G}\in\Phi}\mathbb{P}\big(A_n(\mathcal{G})\big) = 0$. Nun ist $A_n(\mathcal{G}) \in \mathcal{G}$, also gilt

$$\mathbb{E}\big(|\mathbb{E}(Y|\mathcal{G})|\mathbb{1}_{A_n(\mathcal{G})}\big) \leq \mathbb{E}\big(\mathbb{E}(|Y|\,|\,\mathcal{G})\mathbb{1}_{A_n(\mathcal{G})}\big) = \mathbb{E}\big(|Y|\mathbb{1}_{A_n(\mathcal{G})}\big)\,.$$

Somit liefert Lemma 1.6.2 die gleichgradige Integrierbarkeit von Γ. □

Der wichtigste Zusammenhang zwischen gleichgradiger Integrierbarkeit und $\mathcal{L}^1$-Konvergenz ist der folgende.

Satz 1.6.4. *Es seien $(X_n)_{n\in\mathbb{N}}$ eine Folge in $\mathcal{L}^1$ und $X \in \mathcal{L}^1$. Die Folge $(X_n)_{n\in\mathbb{N}}$ konvergiert genau dann in $\mathcal{L}^1$ gegen X, wenn $(X_n)_{n\in\mathbb{N}}$ in Wahrscheinlichkeit gegen X konvergiert und $(X_n)_n$ gleichgradig integrierbar ist.*

Beweis. '$\Longrightarrow$': $(X_n)_n$ konvergiere in $\mathcal{L}^1$ gegen X. Dann konvergiert $(X_n)_n$ insbesondere in Wahrscheinlichkeit gegen X. Für $k, n \in \mathbb{N}$ ist

$$\mathbb{P}(|X_n| \geq k) \leq \frac{\|X_n\|_1}{k}\,.$$

Konvergiert $(X_n)_n$ im ersten Mittel, so ist $\sup_{n\in\mathbb{N}}\|X_n\|_1 < \infty$. Damit folgt

$$\lim_{k\to\infty}\sup_{n\in\mathbb{N}}\mathbb{P}(|X_n| \geq k) = 0\,. \tag{1.6.1}$$

Ist $N \in \mathbb{N}$, so gilt

$$\sup_{n\in\mathbb{N}}\mathbb{E}(|X_n|\mathbb{1}_{\{|X_n|\geq k\}}) \leq \sup_{n\leq N}\mathbb{E}(|X_n|\mathbb{1}_{\{|X_n|\geq k\}}) + \sup_{n\geq N}\|X_n - X\|_1 + \sup_{n\geq N}\mathbb{E}(|X|\mathbb{1}_{\{|X_n|\geq k\}})\,.$$

Mit (1.6.1) und Lemma 1.6.2 konvergiert der dritte Summand für $k \to \infty$ gegen Null. Jede endliche Familie von integrierbaren Zufallsgrößen ist natürlich gleichgradig integrierbar, also konvergiert der erste Summand für $k \to \infty$ gegen Null. Da N beliebig war, folgt die gleichgradige Integrierbarkeit von $(X_n)_n$.

'$\Longleftarrow$': Sei nun $(X_n)_n$ gleichgradig integrierbar und konvergiere in Wahrscheinlichkeit gegen X. Für $k \in \mathbb{N}$ und $x \in \mathbb{R}$ sei

$$\varphi_k(x) := (-k) \vee (x \wedge k)\,.$$

Für $\varepsilon > 0$ und $k, n \in \mathbb{N}$ gilt dann

$$\|\varphi_k(X_n) - \varphi_k(X)\|_1 \le \varepsilon + 2k\mathbb{P}(|X_n - X| > \varepsilon)\,.$$

Die Konvergenz in Wahrscheinlichkeit liefert

$$\lim_{n\to\infty} \|\varphi_k(X_n) - \varphi_k(X)\|_1 = 0, \qquad k \in \mathbb{N}\,,$$

da $\varepsilon > 0$ beliebig war. Nun ist

$$\begin{aligned}\|X_n - X\|_1 &\le \|X_n - \varphi_k(X_n)\|_1 + \|\varphi_k(X_n) - \varphi_k(X)\|_1 + \|\varphi_k(X) - X\|_1\\ &\le 2\,\mathbb{E}(|X_n|\mathbb{1}_{\{|X_n|\ge k\}}) + \|\varphi_k(X_n) - \varphi_k(X)\|_1 + 2\,\mathbb{E}(|X|\mathbb{1}_{\{|X|\ge k\}})\,,\end{aligned}$$

also

$$\limsup_{n\to\infty} \|X_n - X\|_1 \le 2 \sup_{n\in\mathbb{N}} E(|X_n|\mathbb{1}_{\{|X_n|\ge k\}}) + 2\,\mathbb{E}(|X|\mathbb{1}_{\{|X|\ge k\}})\,.$$

Da k beliebig ist, folgt aus der gleichgradigen Integrierbarkeit, dass $\|X_n - X\|_1$ für $n \to \infty$ gegen Null konvergiert. □

Nun zur $\mathcal{L}^1$-Konvergenz von Martingalen:

Satz 1.6.5. *Es sei $\mathbb{X} = (X_n)_{n\in\mathbb{N}}$ ein $(\mathcal{F}_n)_n$-Martingal. Dann sind die folgenden Aussagen äquivalent:*

(i) Das Martingal $(X_n)_n$ ist gleichgradig integrierbar.

(ii) Das Martingal $(X_n)_n$ konvergiert $\mathbb{P}$-fast sicher und in $\mathcal{L}^1$.

(iii) Es existiert ein $Y \in \mathcal{L}^1$ mit $X_n = \mathbb{E}(Y|\mathcal{F}_n)$ $\mathbb{P}$-fast sicher für alle $n \in \mathbb{N}$.

Erfüllt $(X_n)_n$ eine dieser Bedingungen, so kann für Y in (iii) insbesondere der $\mathbb{P}$-fast sichere und $\mathcal{L}^1$-Limes des Martingals $(X_n)_n$ gewählt werden.

Beweis. (iii) $\Rightarrow$ (i) folgt aus Lemma 1.6.3(iii).

(i) $\Rightarrow$ (ii) folgt aus Satz 1.6.4, wenn wir zeigen, dass $(X_n)_n$ $\mathbb{P}$-fast sicher und somit auch in Wahrscheinlichkeit konvergiert: Es existiert ein $k \in \mathbb{N}$ mit $\mathbb{E}(|X_n|\mathbb{1}_{\{|X_n|>k\}}) \le 1$ für alle $n \in \mathbb{N}$. Also

$$\sup_{n\in\mathbb{N}} \|X_n\|_1 \le \sup_{n\in\mathbb{N}} \big(\mathbb{E}(|X_n|\mathbb{1}_{\{|X_n|\le k\}}) + \mathbb{E}(|X_n|\mathbb{1}_{\{|X_n|>k\}})\big) \le k + 1 < \infty\,.$$

Nun liefert Satz 1.4.3 die $\mathbb{P}$-fast sichere Konvergenz von $(X_n)_n$.

(ii) $\Rightarrow$ (iii): Seien X_∞ der $\mathcal{L}^1$-Limes der Folge $(X_n)_n$ und $n \in \mathbb{N}$. Dann gilt für jedes $m \ge n$

$$\mathbb{E}\big(|X_n - \mathbb{E}(X_\infty|\mathcal{F}_n)|\big) = \mathbb{E}\big(|\mathbb{E}(X_m - X_\infty)|\mathcal{F}_n)|\big) \quad \le \mathbb{E}\big(\mathbb{E}(|X_m - X_\infty|\,|\,\mathcal{F}_n)\big) = \|X_m - X_\infty\|_1\,.$$

Für $m \to \infty$ folgt $X_n = \mathbb{E}(X_\infty|\mathcal{F}_n)$ $\mathbb{P}$-fast sicher. Hiermit ist auch die weitere Aussage bewiesen. □

Korollar 1.6.6. *Für $Y \in \mathcal{L}^1$ ist $\big(\mathbb{E}(Y|\mathcal{F}_n)\big)_n$ ein $(\mathcal{F}_n)_n$-Martingal, das fast sicher und in $\mathcal{L}^1$ gegen $\mathbb{E}(Y|\mathcal{F}_\infty)$ konvergiert.*

Beweis. Für $n \in \mathbb{N}$ sei $X_n := \mathbb{E}(Y|\mathcal{F}_n)$. Dann ist $(X_n)_n$ ein Martingal. Nach Satz 1.6.5 konvergiert dieses Martingal $\mathbb{P}$-fast sicher und in $\mathcal{L}^1$ gegen eine Zufallsgröße X_∞, die $\mathcal{F}_\infty$-messbar ist. Zu zeigen ist also nun noch: $X_\infty = \mathbb{E}(Y|\mathcal{F}_\infty)$. Da X_∞ messbar ist bezüglich $\mathcal{F}_\infty$, müssen wir nur noch zeigen, dass $\mathbb{E}(X_\infty 1\!\!1_A) = \mathbb{E}(Y 1\!\!1_A)$ für alle $A \in \mathcal{F}_\infty$ gilt. Mit Satz 1.6.5 folgt dies für alle $A \in \mathcal{F}_n$. Da $\bigcup_{n\in\mathbb{N}} \mathcal{F}_n$ ein durchschnittstabiles Erzeugendensystem von $\mathcal{F}_\infty$ ist, folgt $\mathbb{E}(X_\infty 1\!\!1_A) = \mathbb{E}(Y 1\!\!1_A)$ für alle $A \in \mathcal{F}_\infty$. □

1.7 Doob'sche Ungleichung

Eine weitere angenehme Eigenschaft von Martingalen ist, dass eine gute Kontrolle über das Maximum, das sie über ein gegebenes Zeitintervall annehmen, möglich ist. Sei also $(X_n)_{n\in\mathbb{N}}$ ein Submartingal, und wir betrachten die Maxima

$$X_n^* = \max\{X_1, \dots, X_n\} \qquad \text{sowie} \qquad |X|_n^* = \max\{|X_1|, \dots, |X_n|\}, \qquad n \in \mathbb{N}.$$

Die folgende Aussage ist schon der Kern der Kontrolle über X_n^*.

Lemma 1.7.1. *Sei $(X_n)_{n\in\mathbb{N}}$ ein Submartingal, dann gilt für jedes $\lambda \in (0,\infty)$*

$$\lambda \mathbb{P}(X_n^* \geq \lambda) \leq \mathbb{E}[X_n 1\!\!1_{\{X_n^* \geq \lambda\}}] \leq \mathbb{E}[|X_n| 1\!\!1_{\{X_n^* \geq \lambda\}}].$$

Beweis. Die zweite Ungleichung ist trivial. Für den Beweis der ersten betrachten wir

$$\tau = \inf\{k \in \mathbb{N} \colon X_k \geq \lambda\} \wedge n.$$

Dann ist τ eine Stoppzeit mit $\tau \leq n$. Wir machen uns zunächst klar, dass wir die Ungleichung $\mathbb{E}[X_n] \geq \mathbb{E}[X_\tau]$ haben: Für den Fall, dass $(X_n)_n$ ein Martingal ist, zeigen wir nun sogar, dass $X_\tau = \mathbb{E}[X_n \mid \mathcal{F}_\tau]$ gilt. Da X_τ messbar bezüglich $\mathcal{F}_\tau$ ist, muss man nur noch zeigen, dass für jedes $A \in \mathcal{F}_\tau$ gilt: $\mathbb{E}[X_\tau 1\!\!1_A] = \mathbb{E}[X_n 1\!\!1_A]$. Man beachtet, dass $A \cap \{\tau = t\} \in \mathcal{F}_t$ für jedes $t \in \mathbb{N}$ gilt, und erhält

$$\begin{aligned}\mathbb{E}[X_\tau 1\!\!1_A] &= \sum_{t\leq n} \mathbb{E}[X_t 1\!\!1_{A\cap\{\tau=t\}}] = \sum_{t\leq n} \mathbb{E}\big[\mathbb{E}[X_n \mid \mathcal{F}_t] 1\!\!1_{A\cap\{\tau=t\}}\big] \\ &= \sum_{t\leq n} \mathbb{E}\big[\mathbb{E}[X_n 1\!\!1_{A\cap\{\tau=t\}} \mid \mathcal{F}_t]\big] = \sum_{t\leq n} \mathbb{E}[X_n 1\!\!1_{A\cap\{\tau=t\}}] = \mathbb{E}[X_n 1\!\!1_A].\end{aligned}$$

Dies zeigt, dass $\mathbb{E}[X_\tau 1\!\!1_A] = \mathbb{E}[X_n 1\!\!1_A]$, falls $(X_n)_n$ ein Martingal ist. Der Beweis für $\mathbb{E}[X_\tau 1\!\!1_A] \leq \mathbb{E}[X_n 1\!\!1_A]$ für ein Submartingal $(X_n)_n$ wird hier weggelassen.

Auf dem Ereignis $\{X_n^* < \lambda\}$ gilt $\tau = n$, und auf dem Gegenereignis gilt $\tau = \inf\{k \in \mathbb{N} \colon X_k \geq \lambda\}$, also insbesondere $X_\tau \geq \lambda$. Also haben wir

$$\mathbb{E}[X_n] \geq \mathbb{E}[X_\tau] = \mathbb{E}[X_\tau 1\!\!1_{\{X_n^* \geq \lambda\}}] + \mathbb{E}[X_\tau 1\!\!1_{\{X_n^* < \lambda\}}] \geq \lambda \mathbb{P}(X_n^* \geq \lambda) + \mathbb{E}[X_n 1\!\!1_{\{X_n^* < \lambda\}}].$$

Nun folgt die erste Ungleichung durch Subtraktion von $\mathbb{E}[X_n 1\!\!1_{\{X_n^* < \lambda\}}]$. □

Nun folgt eine Variante bei höherer Integrierbarkeit.

Satz 1.7.2 (Doob'sche L^p-Ungleichung). *Sei $(X_n)_{n\in\mathbb{N}}$ ein Martingal oder ein positives Submartingal, dann gelten:*

(i) Für jedes $p \ge 1$ und jedes $\lambda \in (0,\infty)$ gilt $\lambda^p \mathbb{P}(|X_n|^* \ge \lambda) \le \mathbb{E}[|X_n|^p]$.

(ii) Für jedes $p > 1$ gilt $\mathbb{E}[|X_n|^p] \le \mathbb{E}[(|X|_n^*)^p] \le (\frac{p}{p-1})^p \mathbb{E}[|X_n|^p]$.

Beweis. *(i).* Nach Lemma 1.1.5(iii) ist $(|X_n|^p)_{n\in\mathbb{N}}$ ein Submartingal, also folgt (i) aus Lemma 1.7.1.

(ii). Wir brauchen nur die zweite Ungleichung zu beweisen. Wir halten zunächst fest, dass wir aus Lemma 1.1.5(iii), zusamen mit Lemma 1.7.1, auch haben, dass gilt:

$$\lambda \mathbb{P}(|X|_n^* \ge \lambda) \le \mathbb{E}[|X_n| \mathbb{1}_{\{|X|_n^* > \lambda\}}].$$

Um sicher zu sein, dass die folgend auftretenden Terme endlich sind, schneiden wir bei einem großen $K \in (0,\infty)$ ab und errechnen:

$$\begin{aligned}
\mathbb{E}[(|X|_n^* \wedge K)^p] &= \mathbb{E}\Big[\int_0^{|X|_n^* \wedge K} p\lambda^{p-1}\, \mathrm{d}\lambda\Big] = \mathbb{E}\Big[\int_0^K p\lambda^{p-1} \mathbb{1}_{\{|X|_n^* \ge \lambda\}}\, \mathrm{d}\lambda\Big] \\
&= \int_0^K p\lambda^{p-1} \mathbb{P}(|X|_n^* \ge \lambda)\, \mathrm{d}\lambda \le \int_0^K p\lambda^{p-2} \mathbb{E}[|X_n| \mathbb{1}_{\{|X|_n^* > \lambda\}}]\, \mathrm{d}\lambda \\
&= p\mathbb{E}\Big[|X_n| \int_0^{|X|_n^* \wedge K} \lambda^{p-2}\, \mathrm{d}\lambda\Big] = \frac{p}{p-1} \mathbb{E}\big[|X_n| \big(|X|_n^* \wedge K\big)^{p-1}\big].
\end{aligned}$$

Die Hölder'sche Ungleichung liefert nun

$$\mathbb{E}[(|X|_n^* \wedge K)^p] \le \frac{p}{p-1} \mathbb{E}\big[\big(|X|_n^* \wedge K\big)^p\big]^{(p-1)/p} \mathbb{E}[|X_n|^p]^{1/p}.$$

Nun erheben wir beide Seiten zur Potenz p und teilen dann durch $\mathbb{E}[(|X|_n^* \wedge K)^p]^{p-1}$. Dies ergibt

$$\mathbb{E}[(|X|_n^* \wedge K)^p] \le \Big(\frac{p}{p-1}\Big)^p \mathbb{E}[|X_n|^p].$$

Nun folgt die Aussage durch den Grenzübergang $K \to \infty$. □

Sei nun $(X_n)_{n\in\mathbb{N}}$ ein $\mathcal{L}^p$-beschränktes Martingal, d. h. es gelte $\sup_{n\in\mathbb{N}} \mathbb{E}[|X_n|^p] < \infty$. Im Allgemeinen folgt aus der $\mathcal{L}^p$-Beschränktheit noch nicht die gleichgradige Integrierbarkeit von $(|X_n|^p)_{n\in\mathbb{N}}$. Falls aber $(X_n)_{n\in\mathbb{N}}$ ein Martingal ist und $p > 1$, so folgt dies aus der Doob'schen Ungleichung. Insbesondere folgt dann aus fast sicherer Konvergenz auch schon die L^p-Konvergenz:

Korollar 1.7.3 (L^p-Konvergenzsatz für Martingale). *Sei $p > 1$ und $(X_n)_{n\in\mathbb{N}}$ ein $\mathcal{L}^p$-beschränktes Martingal. Dann existiert eine Zufallsgröße $X_\infty \in \mathcal{L}^p$, so dass $\lim_{n\to\infty} X_n = X_\infty$ fast sicher und in L^p. Insbesondere ist $(|X_n|^p)_{n\in\mathbb{N}}$ gleichgradig integrierbar.*

Beweis. Nach Lemma 1.6.3(i) ist $(X_n)_{n\in\mathbb{N}}$ gleichgradig integrierbar. Nach Satz 1.6.5 konvergiert X_n fast sicher und in $\mathcal{L}^1$. Aus der $\mathcal{L}^p$-Beschränktheit und der Doob'schen Ungleichung folgt, dass $\sup_{k\in\mathbb{N}} |X_k|^p$ einen endlichen Erwartungswert besitzt. Nach Lemma 1.6.3(ii) ist daher auch $(|X_n|^p)_{n\in\mathbb{N}}$ gleichgradig integrierbar. Mit Hilfe des majorisierten Konvergenzsatzes erhält man, dass $X_\infty \in \mathcal{L}^p$ und dass die Konvergenz von X_n gegen X_∞ auch in L^p vorliegt. □

Kapitel 2

Anwendungen von Martingalen

Wir wollen in diesem Kapitel Anwendungen der Martingaltheorie behandeln. Wir betrachten *Produktmartingale*, den *Satz von Kakutani* und Anwendungen in der *asymptotischen Statistik*; wir beweisen den *Satz von Radon-Nikodym* mittels Martingaltheorie; wir diskutieren *Kolmogorovs Kriterium* zu starken Gesetzen der großen Zahlen, untersuchen *verrauschte Beobachtungen* (ein einfaches Modell der sogenannten *Filtertheorie*) und betrachten einfache *Verzweigungsprozesse.*

2.1 Der Satz von Kakutani

Der Satz von Kakutani macht Aussagen über die Existenz und Positivität von Dichten zwischen zwei unendlichen Produktmaßen, wenn ihre Projektionen jeweils Dichten besitzen.

Wir beginnen mit Betrachtungen von Martingalen, die als Partialproduktfolge von unabhängig und identisch verteilten Zufallsgrößen definiert sind, und diskutieren die Existenz und Positivität ihres Grenzwertes. Es sei $(\xi_n)_n$ eine Folge von unabhängigen nicht-negativen Zufallsgrößen mit $\mathbb{E}(\xi_n) = 1$. Für $n \in \mathbb{N}$ seien ferner $\mathcal{F}_n = \sigma(\xi_1, \dots, \xi_n)$ und $M_n = \prod_{i=1}^n \xi_i$. Die Folge $(M_n)_n$ ist ein Martingal, denn $\mathbb{P}$-fast sicher gilt

$$\mathbb{E}(M_{n+1}|\mathcal{F}_n) = \mathbb{E}(M_n\xi_{n+1}|\mathcal{F}_n) = M_n\mathbb{E}(\xi_{n+1}|\mathcal{F}_n) = M_n\mathbb{E}(\xi_{n+1}) = M_n \ . \tag{2.1.1}$$

Korollar 1.4.4 besagt, dass $(M_n)_n$ $\mathbb{P}$-fast sicher gegen eine nicht-negative Zufallsgröße M_∞ konvergiert. Das Lemma von Fatou liefert $\mathbb{E}(M_\infty) \leq \liminf_{n\to\infty} \mathbb{E}(M_n) = 1$. Nach Satz 1.6.5 konvergiert $(M_n)_n$ genau dann in $\mathcal{L}^1$ gegen M_∞, wenn $(M_n)_n$ gleichgradig integrierbar ist. (In diesem Fall ist dann $\mathbb{E}(M_\infty) = 1$.) Von großer Wichtigkeit für das Folgende ist zu wissen, unter welchen Umständen dies vorliegt und unter welchen Umständen der Grenzwert fast sicher positiv ist:

Lemma 2.1.1. *(i) Das Martingal $(M_n)_n$ ist genau dann gleichgradig integrierbar, wenn*

$$\prod_{n\in\mathbb{N}} \mathbb{E}\big(\sqrt{\xi_n}\big) > 0 \ .$$

(ii) Ist $(M_n)_n$ nicht gleichgradig integrierbar, so gilt $M_\infty = 0$ $\mathbb{P}$-fast sicher.

(iii) Ist $(M_n)_n$ gleichgradig integrierbar und $\xi_n > 0$ $\mathbb{P}$-fast sicher für alle $n \in \mathbb{N}$, so gilt auch $M_\infty > 0$ $\mathbb{P}$-fast sicher.

Beweis. Für $n \in \mathbb{N}$ seien $a_n = \mathbb{E}(\sqrt{\xi_n})$ und $b_n = \prod_{i=1}^n a_i$. Mittels der Cauchy-Schwarz-Ungleichung folgt $a_n \leq 1$, also ist $(b_n)_{n\in\mathbb{N}}$ fallend und konvergiert daher gegen ein $b_\infty \geq 0$. Wegen $\mathbb{P}(\xi_n = 0) \neq 1$ ist $a_n > 0$, also auch $b_n > 0$.

Für $n \in \mathbb{N}$ sei $N_n = \prod_{i=1}^n \frac{\sqrt{\xi_i}}{a_i}$. Dann ist $(N_n)_n$ ein $(\mathcal{F}_n)_n$-Martingal (dies sieht man wie in (2.1.1)). Es konvergiert nach Korollar 1.4.4 also $\mathbb{P}$-fast sicher gegen eine nicht-negative Zufallsgröße N_∞. Ist $b_\infty = \prod_{i=1}^\infty a_i > 0$, so folgt

$$K = \sup_{n\in\mathbb{N}} \mathbb{E}(N_n^2) = \sup_{n\in\mathbb{N}} \prod_{i=1}^n \frac{\mathbb{E}(\xi_i)}{a_i^2} = \prod_{i=1}^\infty \frac{1}{a_i^2} = \frac{1}{b_\infty^2} < \infty,$$

und nach Satz 1.5.2 konvergiert $(N_n)_n$ gegen N_∞ in $\mathcal{L}^2$. Es gilt $M_n = b_n^2 N_n^2$ für jedes $n \in \mathbb{N}$, also $M_\infty = b_\infty^2 N_\infty^2$ $\mathbb{P}$-fast sicher. Weiter liefert die Cauchy-Schwarz-Ungleichung

$$\begin{aligned}
||M_n - M_\infty||_1 &= ||(b_n N_n + b_\infty N_\infty)(b_n N_n - b_\infty N_\infty)||_1 \\
&\leq ||b_n N_n + b_\infty N_\infty||_2 \; ||b_n N_n - b_\infty N_\infty||_2 \\
&\leq \Big(\|b_n N_n\|_2 + \|b_\infty N_\infty\|_2\Big)\Big(\|b_n N_n - b_n N_\infty\|_2 + \|b_n N_\infty - b_\infty N_\infty\|_2\Big) \\
&\leq 2\sqrt{K}\Big(b_n ||N_n - N_\infty||_2 + |b_n - b_\infty| \; ||N_\infty||_2\Big) .
\end{aligned}$$

Also konvergiert $(M_n)_n$ gegen M_∞ im ersten Mittel. Nach Satz 1.6.5 ist $(M_n)_n$ gleichgradig integrierbar.

Ist $b_\infty = \prod_{i=1}^\infty a_i = 0$, so folgt aus der $\mathbb{P}$-fast sicheren Konvergenz von $(N_n)_n$ gegen N_∞, dass $(M_n)_n$ $\mathbb{P}$-fast sicher gegen Null konvergiert. Also gilt $M_\infty = 0$ $\mathbb{P}$-fast sicher und $(M_n)_n$ konvergiert nicht im ersten Mittel, kann also nach Satz 1.6.5 nicht gleichgradig integrierbar sein. Somit sind (i) und (ii) bewiesen.

Nun kommen wir zum Beweis von (iii): Wir wissen schon, dass $(M_n)_n$ $\mathbb{P}$-fast sicher und in $\mathcal{L}^1$ gegen M_∞ konvergiert. Also ist $\mathbb{E}(M_\infty) = 1$ und $\mathbb{P}(M_\infty = 0) \neq 1$. Sei $B_n := \{\prod_{i=n}^\infty \xi_i = 0\}$. Wegen $\prod_{i=1}^{n-1} \xi_i > 0$ $\mathbb{P}$-fast sicher gilt $\mathbb{P}(\{M_\infty = 0\} \triangle B_n) = 0$. Ferner gilt

$$\mathbb{P}\Big(\{M_\infty = 0\} \triangle \Big(\bigcap_{n\in\mathbb{N}} B_n\Big)\Big) \leq \sum_{n\in\mathbb{N}} \mathbb{P}\Big(\{M_\infty = 0\} \triangle B_n\Big) = 0 .$$

Da $(B_n)_n$ eine absteigende Folge ist, gilt

$$\bigcap_{n\in\mathbb{N}} B_n = \limsup_{n\to\infty} B_n \in \bigcap_{k\in\mathbb{N}} \sigma(\xi_j, j \geq k).$$

Nach Kolmogorovs 0-1-Gesetz ist $\mathbb{P}(\bigcap_{n\in\mathbb{N}} B_n) \in \{0, 1\}$. Wegen $\mathbb{P}(M_\infty = 0) \neq 1$ folgt somit $\mathbb{P}(M_\infty = 0) = 0$. $\square$

Nun kommen wir zum Satz von Kakutani. Wir betrachten nun die folgende Situation: Es seien $(E, \mathcal{E})$ ein messbarer Raum und $(\mu_n)_n$ und $(\nu_n)_n$ zwei Folgen von Wahrscheinlichkeitsmaßen auf $(E, \mathcal{E})$. Für jedes $n \in \mathbb{N}$ besitze ν_n eine Dichte f_n bezüglich μ_n und umgekehrt μ_n eine Dichte bezüglich ν_n. Also können wir ohne Einschränkung annehmen, dass $f_n(x) > 0$ für alle $x \in E$ gilt. Auf $(\Omega, \mathcal{F}) := (E^{\mathbb{N}}, \mathcal{E}^{\mathbb{N}})$ seien $\mathbb{P} := \otimes_{n\in\mathbb{N}} \mu_n$ und $Q := \otimes_{n\in\mathbb{N}} \nu_n$ die zugehörigen Produktmaße.

Hat $\mathbb{P}$ eine Dichte bezüglich Q und umgekehrt Q eine Dichte bezüglich $\mathbb{P}$? Kakutanis Satz gibt ein notwendiges und hinreichendes Kriterium.

Für $n \in \mathbb{N}$ seien $\pi_n : \Omega \to E$ die Projektion auf den n-ten Faktor und $\mathcal{F}_n = \sigma(\pi_1, \ldots, \pi_n)$. Dann ist $(\mathcal{F}_n)_n$ eine Filtrierung von $\mathcal{F}$. Sei $\xi_n = f_n \circ \pi_n$. Unter $\mathbb{P}$ ist $(\xi_n)_n$ eine Folge unabhängiger Zufallsgrößen mit $\mathbb{E}(\xi_n) = 1$ für alle $n \in \mathbb{N}$. Sei $a_n := \mathbb{E}\big(\sqrt{\xi_n}\big) = \int_E \sqrt{f_n}\, \mathrm{d}\mu_n$.

Satz 2.1.2 (Kakutani). *(i) Ist $\prod_{n=1}^\infty a_n > 0$, so sind $\mathbb{P}$ und Q äquivalent, d.h. es existiert eine Dichte von $\mathbb{P}$ bzgl. Q und eine Dichte von Q bzgl. $\mathbb{P}$. Es gilt*

$$\frac{\mathrm{d}Q}{\mathrm{d}\mathbb{P}} = M_\infty := \lim_{n\to\infty} \prod_{i=1}^n \xi_i \qquad \textit{und} \qquad \frac{\mathrm{d}\mathbb{P}}{\mathrm{d}Q} = M_\infty^{-1} \qquad \mathbb{P}\textit{-fast sicher.}$$

(ii) Ist $\prod_{n=1}^\infty a_n = 0$, so ist Q nicht absolutstetig bezüglich $\mathbb{P}$. Für die Menge $A := \{\lim_{n\to\infty} M_n = 0\} \in \mathcal{F}$ gilt sogar $\mathbb{P}(A) = 1$ und $Q(A) = 0$.

Beweis. *(i).* Für jedes $n \in \mathbb{N}$ seien $M_n := \prod_{i=1}^n \xi_i$ und $\mathbb{P}_n = \mathbb{P}|_{\mathcal{F}_n}$ und $Q_n = Q|_{\mathcal{F}_n}$. Dann sind $\mathbb{P}_n$ und Q_n zueinander äquivalent mit $\frac{\mathrm{d}Q_n}{\mathrm{d}\mathbb{P}_n} = M_n$ $\mathbb{P}$-fast sicher. Nach Lemma 2.1.1 und Satz 1.6.5 konvergiert $(M_n)_n$ $\mathbb{P}$-fast sicher und in $\mathcal{L}^1(\mathbb{P})$ gegen M_∞. Weiter gilt $M_n = \mathbb{E}(M_\infty | \mathcal{F}_n)$ $\mathbb{P}$-fast sicher für jedes $n \in \mathbb{N}$. Also gilt für $A \in \mathcal{F}_n$

$$Q(A) = Q_n(A) = \int_A M_n \,\mathrm{d}\mathbb{P}_n = \int_A M_n \mathrm{d}\mathbb{P} = \int_A M_\infty \,\mathrm{d}\mathbb{P} \;.$$

Dies gilt für alle $A \in \bigcup_{n\in\mathbb{N}} \mathcal{F}_n$, und $\bigcup_{n\in\mathbb{N}} \mathcal{F}_n$ ist ein durchschnittstabiles Erzeugendensystem von $\mathcal{F}$. Also gilt die Gleichheit für alle $A \in \mathcal{F}$. Damit ist $\frac{\mathrm{d}Q}{\mathrm{d}\mathbb{P}} = M_\infty$ $\mathbb{P}$-fast sicher. Wegen $M_\infty > 0$ $\mathbb{P}$-fast sicher (siehe Lemma 2.1.1(iii)) ist $\frac{\mathrm{d}\mathbb{P}}{\mathrm{d}Q} = M_\infty^{-1}$.

(ii). Aus Lemma 2.1.1 wissen wir, dass $(M_n)_n$ $\mathbb{P}$-fast sicher gegen Null konvergiert. Wir zeigen nun, dass $(M_n)_n$ Q-fast sicher gegen unendlich konvergiert. Da $(N_n)_n$ mit $N_n := \frac{\sqrt{M_n}}{\prod_{i=1}^n a_i}$ ein $\mathbb{P}$-Martingal ist und $a_i \le 1$ für alle $i \in \mathbb{N}$ gilt, ist $(\sqrt{M_n})_n$ ein $\mathbb{P}$-Supermartingal, denn es gilt $\mathbb{P}$-fast sicher

$$\mathbb{E}\big(\sqrt{M_n} | \mathcal{F}_{n-1}\big) = \Big(\prod_{i=1}^n a_i\Big)\, \mathbb{E}(N_n | \mathcal{F}_{n-1}) = a_n \sqrt{M_{n-1}} \le \sqrt{M_{n-1}}.$$

Für $B \in \mathcal{F}_n$ gilt

$$\int_B \frac{1}{\sqrt{M_n}} \,\mathrm{d}Q = \int_B \frac{1}{\sqrt{M_n}} M_n \,\mathrm{d}\mathbb{P} = \int_B \sqrt{M_n} \,\mathrm{d}\mathbb{P} \ge \int_B \sqrt{M_{n+1}} \,\mathrm{d}\mathbb{P} = \int_B \frac{1}{\sqrt{M_{n+1}}} \,\mathrm{d}Q.$$

(M_n und M_{n+1} sind überall strikt positiv gewählt). Also ist $(M_n^{-1/2})_n$ ein positives Q-Supermartingal und konvergiert gemäß Korollar 1.4.4 Q-fast sicher in $\overline{\mathbb{R}}$. Wegen $\int_\Omega M_n^{-1/2} \,\mathrm{d}Q = \int_\Omega \sqrt{M_n}\, \mathrm{d}\mathbb{P} = \prod_{i=1}^n a_i \downarrow 0$ für $n \to \infty$ folgt $\lim_{n\to\infty} 1/\sqrt{M_n} = 0$ Q-fast sicher, also $\lim_{n\to\infty} M_n = \infty$ Q-fast sicher. Somit ist $A = \{\omega \in \Omega \colon \lim_{n\to\infty} M_n(\omega) = 0\}$ eine Q-Nullmenge mit $\mathbb{P}(A) = 1$. □

Beispiel 2.1.3 (Normalverteilungen). Sei $(E, \mathcal{E}) = (\mathbb{R}, \mathcal{B})$. Für $n \in \mathbb{N}$ seien μ_n die Standardnormalverteilung und ν_n die Normalverteilung mit Erwartungswert $\alpha_n \in \mathbb{R}$ und Varianz 1. Dann gilt für jedes $n \in \mathbb{N}$

$$\frac{\mathrm{d}\nu_n}{\mathrm{d}\mu_n}(x) = \frac{\exp\big(-(x-\alpha_n)^2/2\big)}{\exp\big(-x^2/2\big)} = \exp\big(\alpha_n x - \alpha_n^2/2\big) \;, \qquad x \in \mathbb{R} \;.$$

Weiter ist

$$\begin{aligned} a_n &:= \int_{\mathbb{R}} \sqrt{\frac{\mathrm{d}\nu_n}{\mathrm{d}\mu_n}}\,\mathrm{d}\mu_n = \int_{\mathbb{R}} \exp\Big(\frac{\alpha_n x}{2} - \frac{\alpha_n^2}{4}\Big)\frac{1}{\sqrt{2\pi}}\exp\big(-x^2/2\big)\,\mathrm{d}x \\ &= \exp\big(-\alpha_n^2/8\big)\int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}}\exp\Big(-\frac{1}{2}\big(x-\frac{\alpha_n}{2}\big)^2\Big)\,\mathrm{d}x = \exp\big(-\alpha_n^2/8\big)\ . \end{aligned}$$

Somit ist $\prod_{n=1}^\infty a_n > 0$ genau dann, wenn $\sum_{n=1}^\infty \alpha_n^2 < \infty$. Nach Satz 2.1.2 sind $\mathbb{P}$ und Q genau dann äquivalent, wenn $(\alpha_n)_n \in \ell_2$. In diesem Fall ist

$$\frac{\mathrm{d}Q}{\mathrm{d}\mathbb{P}}(x) = \exp\Big(\sum_{n=1}^\infty \alpha_n x_n - \frac{1}{2}\sum_{n=1}^\infty \alpha_n^2\Big)$$

für $\mathbb{P}$-fast alle $x = (x_n)_n \in \mathbb{R}^{\mathbb{N}}$. Die Reihe $\sum_{n=1}^\infty \alpha_n x_n$ konvergiert nur $\mathbb{P}$-fast sicher. Außerhalb dieser Menge von $\mathbb{P}$- und Q-Maß 1 können wir die Dichtefunktion nach Belieben setzen, zum Beispiel gleich 1. $\diamond$

Beispiel 2.1.4 (Asymptotische Statistik). Wir untersuchen das Verhalten von Folgen von Dichtequotienten. Sind $\mathbb{P}$ und Q die obigen Produktmaße auf $(\Omega, \mathcal{F})$, so sind die Projektionen $\pi_1, \pi_2, \dots$ unabhängige Zufallsgrößen. In der Statistik möchte man auf der Basis von Beobachtungen entscheiden, ob $\mathbb{P}$ oder Q vorliegt. Es sei nun $(\rho_n)_n$ eine weitere Folge von Wahrscheinlichkeitsmaßen auf $(E, \mathcal{E})$, so dass für jedes $n \in \mathbb{N}$ μ_n eine strikt positive Dichte f_n bezüglich ρ_n besitzt und ν_n eine strikt positive Dichte g_n bezüglich ρ_n. Wir setzen $\xi_i = \frac{g_i}{f_i} \circ \pi_i$. Dann ist (siehe Notation wie im Beweis von Satz 2.1.2)

$$\frac{\mathrm{d}Q_n}{\mathrm{d}\mathbb{P}_n} = \prod_{i=1}^n \frac{g_i}{f_i} \circ \pi_i = \prod_{i=1}^n \xi_i \quad \text{fast sicher}$$

und $\mathbb{E}(\xi_i) = \int_E \frac{g_i(x)}{f_i(x)} f_i(x)\,\rho_i(\mathrm{d}x) = \int_E g_i(x)\,\rho_i(\mathrm{d}x) = \nu_i(E) = 1$, wie bereits gesehen. Weiter ist

$$\mathbb{E}(\sqrt{\xi_i}) = \int \frac{\sqrt{g_i(x)}}{\sqrt{f_i(x)}} f_i(x)\,\rho_i(\mathrm{d}x) = \int \sqrt{g_i(x) f_i(x)}\,\rho_i(\mathrm{d}x)\ .$$

Die Bedingung $\prod_{i=1}^\infty a_n > 0$ im Satz von Kakutani liest sich hier also zu

$$\prod_{i=1}^\infty \int_E \sqrt{g_i(x) f_i(x)}\,\rho_i(\mathrm{d}x) > 0\ , \tag{2.1.2}$$

was übrigens zu $\sum_{i\in\mathbb{N}} \int_E \big(f_i(x)^{1/2} - g_i(x)^{1/2}\big)^2 \rho_i(\mathrm{d}x) < \infty$ äquivalent ist. $\diamond$

Dies können wir einsetzen, um die Konsistenz einer natürlichen Folge von Schätzern zu zeigen:

Beispiel 2.1.5 (Konsistenz). Wir bleiben in der Notation von Beispiel 2.1.4 und betrachten den Fall $\mu_1 = \mu_2 = \cdots$ und $\nu_1 = \nu_2 = \cdots$ und $\rho_1 = \rho_2 = \cdots$; somit sind alle Dichten f_n, sowie g_n, gleich, in Notation f und g. Also ist

$$M_n = \prod_{k=1}^n \frac{g(\pi_k)}{f(\pi_k)}, \qquad n \in \mathbb{N}_0,$$

eine Dichte von $Q_n = \nu_1^{\otimes n}$ bezüglich $\mathbb{P}_n = \mu_1^{\otimes n}$. Um Trivialitäten zu vermeiden, setzen wir voraus, dass $\mu_1 \neq \nu_1$. Somit ist $\mathbb{P}$ singulär zu Q, siehe Satz 2.1.2(ii).

Wir betrachten eine Standardaufgabe aus der Testtheorie. Aufgrund von n unabhängigen Beobachtungen $X_1, \dots, X_n$ möchte man entscheiden, ob diesen Zufallsexperimenten die Verteilung μ_1 oder ν_1 zugrundeliegt. Wenn der Wert der Dichte $\mathrm{d}Q_n/\mathrm{d}\mathbb{P}_n$ im Vektor $(X_1, \dots, X_n)$ groß ist, sollte man meinen, dass ν_1 vorliegt, ansonsten, dass μ_1 vorliegt. Also betrachtet man die Entscheidungsfunktion (den *Schätzer*)

$$\varphi_n(x_1, \dots, x_n) := 1\!\!1_{(k_n,\infty)}\Big(\prod_{k=1}^{n} \frac{g(x_k)}{f(x_k)}\Big) = 1\!\!1_{\{M_n > k_n\}}\big((x_i)_{i\in\mathbb{N}}\big). \tag{2.1.3}$$

und verfolgt die Strategie: Wenn $\varphi(X_1, \dots, X_n) = 1$, so tippe ich auf ν_1, sonst auf μ_1. Wir möchten dabei zwei Dinge minimieren: die Wahrscheinlichkeit für einen *Fehler Erster Art* (d. h. die Wahrscheinlichkeit unter $\mathbb{P}$ für die Wahl von ν_1, obwohl μ_1 richtig ist) sowie die für einen *Fehler Zweiter Art* (die Wahrscheinlichkeit unter Q für die Wahl von μ_1, obwohl ν_1 richtig ist). Da man nicht gleichzeitig über zwei Dinge optimieren kann, fordern wir, dass die Wahrscheinlichkeit für einen Fehler Erster Art unter einer gegebenen Schranke $\alpha \in (0, 1)$ liegen soll, und optimieren dann die Wahrscheinlichkeit für einen Fehler Zweiter Art. In anderen Worten, wir wählen k_n möglichst groß, so dass gilt:

$$\mu_1^{\otimes n}\Big(\Big\{(x_1, \dots, x_n) \in \mathbb{R}^n \colon \prod_{k=1}^{n} \frac{g(x_k)}{f(x_k)} > k_n\Big\}\Big) = \mathbb{P}(M_n > k_n) \leq \alpha.$$

Damit haben wir die Symmetrie zwischen $\mathbb{P}$ und Q zerstört: Wir erwarten die *Hypothese* $\mathbb{P}$ und testen sie gegen die *Alternative* Q. Die Schranke nennt man dann auch ein *Irrtumsniveau.* Eines der wichtigsten Ergebnisse in der Testtheorie ist das *Neyman-Pearson-Lemma*, das besagt, dass der Schätzer φ in (2.1.3) die obige Optimierungsaufgabe löst.

Da nach Satz 2.1.2 $\mathbb{P}(M_n \to 0) = 1$, folgt $k_n \to 0$. Damit konvergiert die Wahrscheinlichkeit für einen Fehler Zweiter Art, $Q(M_n \leq k_n)$, für $n \to \infty$ gegen 0, denn

$$Q(M_n \leq k_n) = \int_{\{M_n \leq k_n\}} M_n \,\mathrm{d}\mathbb{P} \leq k_n \to 0 \ .$$

Dies nennt man *Konsistenz* der Testfolge $(\varphi_n)_n$. Analog folgt durch Vertauschen der Rollen von μ_1 und ν_1 die Konsistenz des analogen Testes zum Niveau α für ν_1 gegen μ_1. Die Folge der Dichtequotienten heißt auch *Likelihood-Prozess.* ◇

2.2 Der Satz von Radon-Nikodym

Wir zeigen hier eine (leicht eingeschränkte) Version des *Satzes von Radon-Nikodym.* Wir erinnern, dass ein Maß Q absolutstetig bezüglich eines Maßes $\mathbb{P}$ heißt, wenn jede $\mathbb{P}$-Nullmenge auch eine Q-Nullmenge ist, und wir schreiben dann $Q \ll \mathbb{P}$. Eine σ-Algebra heißt *separabel*, wenn sie durch abzählbar viele Mengen erzeugt wird.

Satz 2.2.1 (Satz von Radon-Nikodym für separable σ-Algebren). *Es sei $(\Omega, \mathcal{F}, \mathbb{P})$ ein Wahrscheinlichkeitsraum, und $\mathcal{F}$ sei separabel. Weiter sei Q ein endliches Maß auf $(\Omega, \mathcal{F})$, und Q sei absolutstetig bezüglich $\mathbb{P}$. Dann existiert eine Zufallsvariable X auf $(\Omega, \mathcal{F})$ mit*

$$Q(A) = \int_A X(\omega)\, \mathbb{P}(\mathrm{d}\omega) \qquad \textit{für jedes } A \in \mathcal{F} \ .$$

Viele σ-Algebren sind separabel, zum Beispiel die Borel-σ-Algebra eines jeden separablen metrischen Raums. Der Satz von Radon-Nikodym gilt allerdings auch ohne diese Voraussetzung, wie man in jedem guten Buch über Maßtheorie nachlesen kann.

Im Beweis von Satz 2.2.1 werden wir eine allgemeine (technische) Charakterisierung der Absolutstetigkeit benötigen:

Lemma 2.2.2. *Es seien $\mathbb{P}$ und Q zwei endliche Maße auf $(\Omega, \mathcal{F})$. Dann gilt $Q \ll \mathbb{P}$ genau dann, wenn zu jedem $\varepsilon > 0$ ein $\delta > 0$ existiert mit $Q(A) \leq \varepsilon$ für jedes $A \in \mathcal{F}$, das $\mathbb{P}(A) \leq \delta$ erfüllt.*

Beweis. „$\Longleftarrow$“: Aus der Bedingung folgt $Q(A) \leq \varepsilon$ für jede $\mathbb{P}$-Nullmenge $A \in \mathcal{F}$ und jedes $\varepsilon > 0$. Daher ist $Q(A) = 0$.

„$\Longrightarrow$“: Angenommen, die Bedingung gilt nicht. Dann existiert ein $\varepsilon > 0$ und eine Folge $(A_n)_n$ in $\mathcal{F}$ mit $\mathbb{P}(A_n) \leq 2^{-n}$ und $Q(A_n) > \varepsilon$ für jedes $n \in \mathbb{N}$. Wähle

$$A = \limsup_{n\to\infty} A_n = \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m \ ,$$

so ist $A \in \mathcal{F}$ und

$$\mathbb{P}(A) \leq \mathbb{P}\Big(\bigcup_{m=n}^{\infty} A_m\Big) \leq \sum_{m=n}^{\infty} \mathbb{P}(A_m) \leq \sum_{m=n}^{\infty} 2^{-m} = 2^{-n+1} \ , \qquad n \in \mathbb{N}.$$

Also gilt $\mathbb{P}(A) = 0$. Andererseits gilt wegen der Endlichkeit von Q

$$Q(A) \geq \limsup_{n\to\infty} Q(A_n) \geq \varepsilon > 0 \ ,$$

im Widerspruch zur Annahme, dass jede $\mathbb{P}$-Nullmenge eine Q-Nullmenge ist. □

Beweis von Satz 2.2.1. Auf Grund der Voraussetzung der Separabilität können wir eine Folge $A_1, A_2, \dots$ von messbaren Mengen wählen, so dass $\mathcal{F} = \sigma(A_n : n \in \mathbb{N})$. Damit konstruieren wir leicht eine Folge $\mathcal{A}_1 \subset \mathcal{A}_2 \subset \mathcal{A}_3 \subset \cdots$ von Zerlegungen $\mathcal{A}_n = \{A_{n,1}, \dots, A_{n,k_n}\}$ von Ω in disjunkte messbare Mengen $A_{n,i}$, die jeweils eine (disjunkte) Vereinigung von Elementen der Zerlegung $\mathcal{A}_{n+1}$ ist. Mit einer geeigneten Indexmenge $I(n,i)$ haben wir also

$$A_{n,i} = \bigcup_{j\in I(n,i)} A_{n+1,j}, \qquad n \in \mathbb{N}, i \in \{1, \dots, k_n\}.$$

Seien nun $\mathbb{P}$ und Q wie in Satz 2.2.1 gegeben. Dann definieren wir $\mathcal{F}_n = \sigma(\mathcal{A}_n)$ und

$$M_n = \sum_{i=1}^{k_n} \frac{Q(A_{n,i})}{\mathbb{P}(A_{n,i})} 1\!\!1_{A_{n,i}} \ .$$

Dann ist $(M_n)_n$ unter $\mathbb{P}$ ein Martingal bezüglich $(\mathcal{F}_n)_n$, wie wir uns nun überlegen wollen. Zunächst ist M_n offensichtlich messbar bezüglich $\sigma(\mathcal{A}_n) = \mathcal{F}_n$. Weiterhin ist jedes $G \in \mathcal{F}_n$ eine disjunkte Vereinigung von geeigneten $A_{n,i}$'s. Ohne Einschränkung ist $G = A_{n,i}$ für ein i. Dann gilt

$$\int_G M_{n+1} \,\mathrm{d}\mathbb{P} = \sum_{j\in I(n,i)} \frac{Q(A_{n+1,j})}{\mathbb{P}(A_{n+1,j})} \mathbb{P}(A_{n+1,j}) = Q(A_{n,i}) = \int_G M_n \,\mathrm{d}\mathbb{P} \ .$$

Also folgt $\mathbb{E}[M_{n+1} \mid \mathcal{A}_n] = M_n$ fast sicher, d. h., $(M_n)_n$ ist ein $(\mathcal{F}_n)_n$-Martingal unter $\mathbb{P}$.

Per Definition ist $\mathcal{F}_n$ gleich der Menge aller 2^{k_n} möglichen Vereinigungen der Mengen $A_{n,1}, \dots, A_{n,k_n}$. Für $A \in \mathcal{F}_n$ gibt es also ein $I \subset \{1, \dots, k_n\}$ mit $A = \bigcup_{i \in I} A_{n,i}$. Dann ist

$$Q(A) = \sum_{i \in I} Q(A_{n,i}) = \sum_{i \in I} \frac{Q(A_{n,i})}{\mathbb{P}(A_{n,i})} \mathbb{P}(A_{n,i}) = \sum_{i \in I} \int_{A_{n,i}} M_n \, \mathrm{d}\mathbb{P} = \int_A M_n \, \mathrm{d}\mathbb{P} ,$$

wobei wir beachteten, dass im Fall $\mathbb{P}(A_{n,i}) = 0$ auch $Q(A_{n,i}) = 0$ folgt, $\frac{Q(A_{n,i})}{\mathbb{P}(A_{n,i})}$ also beliebig gesetzt werden kann. Also ist M_n eine Dichte von $Q|_{\mathcal{F}_n}$ bezüglich $\mathbb{P}|_{\mathcal{F}_n}$. $(M_n)_n$ ist ein nichtnegatives Martingal, konvergiert also fast sicher gegen eine Zufallsgröße M_∞. Zu $\varepsilon > 0$ wähle $\delta > 0$ wie in Lemma 2.2.2, und sei $K \in (0, \infty)$ so, dass $K > \frac{Q(\Omega)}{\delta}$. Dann ist $\mathbb{P}(M_n > K) \leq \frac{1}{K}\mathbb{E}(M_n) = \frac{1}{K} Q(\Omega) < \delta$ und somit

$$\int_{\{M_n > K\}} M_n \, \mathrm{d}\mathbb{P} = Q(M_n > K) \leq \varepsilon .$$

Also ist $(M_n)_n$ gleichgradig integrierbar, und somit konvergiert $(M_n)_n$ in $\mathcal{L}^1$ gegen M_∞. Also gilt für jedes $m \in \mathbb{N}$ und für jedes $A \in \mathcal{F}_m$

$$\int_A M_\infty \, \mathrm{d}\mathbb{P} = \lim_{n \to \infty} \int_A M_n \, \mathrm{d}\mathbb{P} = Q(A) .$$

Da $\bigcup_n \mathcal{F}_n$ ein durchschnittsstabiler Erzeuger von $\mathcal{F}$ ist, folgt die Aussage für alle $A \in \mathcal{F}$, also ist M_∞ die gesuchte Dichte. $\square$

2.3 Kolmogorovs Kriterium

Mit Hilfe der Borel-Cantelli-Lemmata kann man die folgende Variante des Starken Gesetzes der Großen Zahlen beweisen:

Satz 2.3.1 (Kolmogorov). *Es sei $(X_n)_n$ eine Folge von unabhängigen zentrierten Zufallsgrößen und $(a_n)_n$ eine steigende Zahlenfolge in $(0, \infty)$ mit $\lim_{n \to \infty} a_n = \infty$. Dann gilt die folgende Implikation:*

$$\sum_{n \in \mathbb{N}} \frac{1}{a_n^2} \operatorname{Var}(X_n) < \infty \qquad \Longrightarrow \qquad \lim_{n \to \infty} \frac{1}{a_n} \sum_{i=1}^n X_i = 0 \quad \mathbb{P}\text{-}\textit{fast sicher.}$$

Wir diskutieren dieses Kriterium nun mittels der Martingaltheorie. Wir werden einen Beweis erhalten, der sogar ohne die Unabhängigkeit der $X_1, X_2, \dots$ auskommt, andererseits aber die Voraussetzung $\sum_{n \in \mathbb{N}} \frac{1}{a_n^2} \operatorname{Var}(X_n) < \infty$ modifiziert.

Sei $(X_n)_n$ eine Folge integrierbarer reeller Zufallsvariablen mit

$$\mathbb{E}(X_{n+1} | X_1, \dots, X_n) = 0, \qquad n \in \mathbb{N}, \ \mathbb{P}\text{-fast sicher.} \tag{2.3.1}$$

(Wir ersetzen also die Voraussetzung der Unabhängigkeit durch die schwächere Bedingung (2.3.1).) Insbesondere sind die X_i zentriert. Wir betrachten die Zufallsgrößen $\widetilde{X}_i = \frac{1}{a_i} X_i$ und die Partialsummen $S_n = \widetilde{X}_1 + \dots + \widetilde{X}_n$. Es gilt dann auch

$$\mathbb{E}(\widetilde{X}_{n+1} | \widetilde{X}_1, \dots, \widetilde{X}_n) = 0, \qquad n \in \mathbb{N}, \ \mathbb{P}\text{-fast sicher.} \tag{2.3.2}$$

Dann ist $(S_n)_{n\in\mathbb{N}}$ ein Martingal bezüglich der natürlichen Filtrierung

$$\sigma(S_1,\dots,S_n) = \sigma(\widetilde{X}_1,\dots,\widetilde{X}_n) = \sigma(X_1,\dots,X_n),$$

wie wir uns nun klar machen wollen. Zu (2.3.2) addieren wir die Gleichungen $\mathbb{E}(\widetilde{X}_i|\widetilde{X}_1,\dots,\widetilde{X}_n) = \widetilde{X}_i$ für $i = 1,\dots,n$ und erhalten

$$\begin{aligned}\mathbb{E}\big[S_{n+1} \mid S_0,\dots,S_n\big] &= \mathbb{E}(\widetilde{X}_{n+1} \mid \widetilde{X}_1,\dots,\widetilde{X}_n) + \sum_{i=1}^n \mathbb{E}\big[\widetilde{X}_i \mid \widetilde{X}_1,\dots,\widetilde{X}_n\big]\\ &= 0 + \sum_{i=1}^n \widetilde{X}_i = S_n.\end{aligned}$$

(In dem Spezialfall, wo die $X_1, X_2, \dots$ unabhängig sind, hatten wir die Martingaleigenschaft zu Beginn von Abschnitt 1.1 gesehen.)

Nun setzen wir an Stelle der Bedingung $\sum_{n\in\mathbb{N}} \frac{1}{a_n^2}\operatorname{Var}(X_n) < \infty$ voraus, dass gilt:

$$\sup_{n\in\mathbb{N}} \mathbb{E}\Big(\Big|\sum_{i=1}^n \frac{1}{a_i} X_i\Big|\Big) < \infty. \tag{2.3.3}$$

Diese Voraussetzung ist äquivalent zur Bedingung $\sup_{n\in\mathbb{N}} \mathbb{E}|S_n| < \infty$, und diese ist nach Satz 1.4.3 hinreichend für fast sichere Konvergenz von S_n für $n \to \infty$. Somit liefert das Kronecker'sche Lemma[1]

$$\lim_{n\to\infty} \frac{1}{a_n} \sum_{i=1}^n X_i = 0 \quad \text{fast sicher},$$

also haben wir eine Variante von Satz 2.3.1 bewiesen, in dem die Unabhängigkeit durch die Bedingung (2.3.1) ersetzt wurde und die Voraussetzung $\sum_{n=1}^\infty \operatorname{Var}(\widetilde{X}_n) < \infty$ durch (2.3.3).

Im Fall der Unabhängigkeit ist die Bedingung $\sum_{n=1}^\infty \operatorname{Var}(\widetilde{X}_n) < \infty$ aus Satz 2.3.1 ebenfalls hinreichend für die fast sichere Konvergenz von S_n, denn

$$\big(\mathbb{E}|S_n|\big)^2 \le \operatorname{Var}(S_n) = \sum_{i=1}^n \operatorname{Var}(\widetilde{X}_i) \le \sum_{i=1}^\infty \operatorname{Var}(\widetilde{X}_i)\ .$$

2.4 Verrauschte Beobachtungen (Filtertheorie)

Es seien X und $\eta_1, \eta_2, \dots$ unabhängige zentrierte normalverteilte Zufallsgrößen, wobei σ^2 die Varianz von X sei, und diejenige der η_n sei 1. Es sei X die Zufallsgröße, an der wir interessiert sind. Angenommen, sie kann nicht direkt beobachtet werden, aber in *verrauschter* Form. Wir nehmen dazu an, dass man die Zufallsgrößen

$$Y_n = X + c_n\eta_n, \qquad n \in \mathbb{N}\ ,$$

beobachten kann, wobei $c_1, c_2, \dots \in (0,\infty)$ sind. Ein natürlicher Schätzer, der auf den Beobachtungen $Y_1,\dots,Y_n$ basieren soll, ist vermutlich $M_n = \mathbb{E}(X|\mathcal{F}_n)$, wobei $\mathcal{F}_n = \sigma(Y_1,\dots,Y_n)$. Die Folge $(M_n)_n$ ist ein quadratisch integrierbares Martingal mit

$$\mathbb{E}(M_n^2) \le \mathbb{E}\big(\mathbb{E}(X^2|\mathcal{F}_n)\big) = \mathbb{E}(X^2) = \sigma^2, \qquad n \in \mathbb{N}.$$

[1] Das Kronecker'sche Lemma besagt, dass $\lim_{n\to\infty} \frac{y_n}{a_n} = 0$, falls die Folge $(\sum_{i=1}^n (y_i - y_{i-1})/a_i)_{n\in\mathbb{N}}$ in $\mathbb{R}$ konvergiert; hierbei ist $(a_n)_{n\in\mathbb{N}}$ eine Folge wie in Satz 2.3.1 und $(y_n)_{n\in\mathbb{N}_0}$ eine beliebige Folge reeller Zahlen.

Mit Satz 1.5.2 und Korollar 1.6.6 folgt $\lim_{n\to\infty} M_n = M_\infty = \mathbb{E}(X|\mathcal{F}_\infty)$ in $\mathcal{L}^2$ und $\mathbb{P}$-fast sicher, wobei $\mathcal{F}_\infty = \sigma(Y_n : n \in \mathbb{N})$.

Die Frage, die wir hier diskutieren wollen, ist: Unter welchen Umständen ist $M_\infty = X$ $\mathbb{P}$-fast sicher? Dazu werden wir zunächst M_n und die Varianzen von $X - M_n$ berechnen. Wir bestimmen dazu $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ so, dass

$$X - \sum_{i=1}^{n} \lambda_i Y_i =: Z$$

und $\mathcal{F}_n$ unabhängig sind. Da alle Zufallsgrößen gemeinsam normalverteilt sind, ist die gewünschte Unabhängigkeit gleichbedeutend mit

$$\operatorname{Cov}(Z, Y_i) = 0 \quad \text{für jedes } i \in \{1, \ldots, n\} \ .$$

Eine elementare Rechnung ergibt

$$\operatorname{Cov}(Z, Y_i) = \sigma^2 \Big(1 - \sum_{j=1}^{n} \lambda_j\Big) - \lambda_i c_i^2 \ .$$

Wir wollen also das Gleichungssystem

$$\sigma^2 \Big(1 - \sum_{j=1}^{n} \lambda_j\Big) - \lambda_i c_i^2 = 0 \ , \qquad i = 1, \ldots, n$$

lösen. Es ergibt sich die Lösung $\lambda_i = \frac{\sigma^2 c_i^{-2}}{1+\sigma^2 \sum_{j=1}^n c_j^{-2}}$, $i \in \{1, \ldots, n\}$. Für diese $\lambda_1, \ldots, \lambda_n$ ist also Z unabhängig von $\mathcal{F}_n$. Dies impliziert $\mathbb{E}(Z|\mathcal{F}_n) = \mathbb{E}(Z) = 0$, also

$$M_n = \mathbb{E}(X|\mathcal{F}_n) = \sum_{i=1}^{n} \lambda_i Y_i$$

sowie

$$\mathbb{E}\big((X - M_n)^2\big) = \mathbb{E}Z^2 = \sigma^2 \Big(1 - \sum_{j=1}^{n} \lambda_j\Big)^2 + \sum_{i=1}^{n} \lambda_i^2 c_i^2 = \sigma^2 \Big(1 + \sigma^2 \sum_{i=1}^{n} c_i^{-2}\Big)^{-1} \ .$$

Dies konvergiert genau dann gegen Null für $n \to \infty$, wenn $\lim_{n\to\infty} \sum_{i=1}^n c_i^{-2} = \infty$ gilt. Somit erhalten wir

Satz 2.4.1. *$(M_n)_n$ konvergiert gegen X $\mathbb{P}$-fast sicher und in $\mathcal{L}^2$ genau dann, wenn*

$$\sum_{i=1}^{\infty} \frac{1}{c_i^2} = \infty \ .$$

Allgemeine Filterprobleme lesen sich etwa so:

Es seien $\alpha, \beta, \gamma, \rho, \nu$ reelle Konstanten und X_0, $(\varepsilon_n)_n$, $(\eta_n)_n$ stochastisch unabhängige Zufallsgrößen mit $X_0 \sim \mathcal{N}(\mu_0, \sigma_0^2)$ sowie $\varepsilon_n \sim \eta_n \sim \mathcal{N}(0, 1)$ für jedes $n \in \mathbb{N}$. Der Zustand X_n eines linearen stochastischen Systems zum Zeitpunkt $n \in \mathbb{N}$ sei durch

$$X_n - X_{n-1} = \alpha X_{n-1} + \beta \varepsilon_n + \gamma$$

beschrieben. Statt X_n selbst kann aber nur eine gestörte (verrauschte) Version Y_n, beschrieben durch $Y_n - Y_{n-1} = \rho X_n + \nu\eta_n$, $Y_0 = 0$, beobachtet werden. Der sogenannte *Kalman-Bucy-Filter* ist ein Verfahren, mit dessen Hilfe der tatsächliche Systemzustand erstaunlich gut geschätzt werden kann. Dies findet Anwendungen bei der Analyse von Zeitreihen oder der Signalverarbeitung (Ortung beweglicher Objekte).

2.5 Verzweigungsprozesse

Sei $(X_{n,i})_{n\in\mathbb{N}_0, i\in\mathbb{N}}$ eine Familie unabhängiger und identisch verteilter Zufallsvariablen mit Werten in $\mathbb{N}_0$. Für $k \in \mathbb{N}$ setze $p_k = \mathbb{P}(X_{1,1} = k)$. Seien $m = \mathbb{E}(X_{1,1}) = \sum_{k\in\mathbb{N}} kp_k < \infty$ und $\sigma^2 := \mathrm{Var}(X_{1,1}) \in (0,\infty)$. Wir definieren den stochastischen Prozess $Z = (Z_n)_{n\in\mathbb{N}_0}$ durch

$$Z_0 = 1 \qquad \text{und} \qquad Z_{n+1} = \sum_{i=1}^{Z_n} X_{n,i}, \quad n \in \mathbb{N}\ .$$

Eine mögliche Interpretation ist: Z_n ist die Größe einer Population zur Zeit n. Das i-te Individuum hat $X_{n,i}$ Nachkommen in der $(n+1)$-ten Generation. Z heißt *Galton-Watson-Prozess* oder *Verzweigungsprozess* mit Nachkommenverteilung $p = (p_k)_{k\in\mathbb{N}_0}$. Um Trivialitäten auszuschließen, setzen wir voraus, dass $p_1 \neq 1$ gilt.

Mit Hilfe erzeugender Funktionen kann man zeigen, dass die *Aussterbewahrscheinlichkeit* $\lim_{n\to\infty} \mathbb{P}(Z_n = 0)$ genau dann strikt kleiner als 1 ist, wenn die erwartete Nachkommenschaft m pro Individuum strikt größer als 1 ist. Im Fall $m \leq 1$ ist fast sicher also sogar $Z_n = 0$ für alle genügend großen n, denn Z_n ist ganzzahlig.

Wir betrachten dieses Modell jetzt im Rahmen der Martingaltheorie: Sei $\mathcal{F}_n := \sigma(X_{k,i} \colon k < n, i \in \mathbb{N})$. Dann ist $Z = (Z_n)_{n\in\mathbb{N}_0}$ an $(\mathcal{F}_n)_{n\in\mathbb{N}_0}$ adaptiert. Setze

$$W_n = \frac{Z_n}{m^n}\ .$$

Lemma 2.5.1. $\mathbb{W} = (W_n)_{n\in\mathbb{N}_0}$ *ist ein Martingal.*

Beweis. Wir errechnen, dass fast sicher gilt:

$$\begin{aligned}
\mathbb{E}(W_{n+1}|\mathcal{F}_n) &= m^{-(n+1)}\mathbb{E}(Z_{n+1}|\mathcal{F}_n) = m^{-(n+1)}\mathbb{E}\Big(\sum_{i=1}^{Z_n} X_{n,i}|\mathcal{F}_n\Big) \\
&= m^{-(n+1)}\sum_{k=1}^{\infty}\mathbb{E}\big(1_{\{Z_n=k\}}\,k\,X_{n,1}|\mathcal{F}_n\big) \qquad (2.5.1)\\
&= m^{-n}\sum_{k=1}^{\infty}\mathbb{E}\big(k\cdot 1_{\{Z_n=k\}}|\mathcal{F}_n\big) = m^{-n}Z_n = W_n.
\end{aligned}$$

□

Also wittern wir, dass das Martingal $\mathbb{W}$ vielleicht konvergieren könnte. Dies ist tatsächlich ohne weitere Vorausssetzungen richtig, und wir können sogar eine Dichotomie beweisen und charakterisieren: Entweder ist der Grenzwert fast sicher konstant gleich Null oder fast sicher positiv, und Letzteres ist genau im Fall $m > 1$ richtig.

Satz 2.5.2. *Es existiert der fast sichere Grenzwert* $W_\infty = \lim_{n\to\infty} W_n$. *Es sind äquivalent:*

$$(a)\ m > 1, \qquad (b)\ \mathbb{E}(W_\infty) = 1, \qquad (c)\ \mathbb{E}(W_\infty) > 0.$$

Beweis. Der fast sichere Grenzwert W_∞ existiert, da $(W_n)_{n\in\mathbb{N}_0}$ ein nicht-negatives Martingal ist, siehe Korollar 1.4.4. Ist $m \leq 1$, so ist gilt $Z_n \to 0$ fast sicher, also sogar $Z_n = 0$ für alle genügend großen n, denn Z_n ist ja ganzzahlig. Also ist auch $W_n = 0$ für alle genügend großen n, d. h. $W_\infty = 0$.

Es sei nun $m > 1$. Dann wollen wir die Varianz von W_n bestimmen, also die von Z_n. Dazu benutzen wir eine Formel, die man auch den *Satz von Blackwell-Girshick* nennt, den wir im Anschluss formulieren und beweisen:

$$\operatorname{Var}(W_n) = m^{-2n}\big(\sigma^2\mathbb{E}(Z_{n-1}) + m^2 \operatorname{Var}(Z_{n-1})\big) = \sigma^2 m^{-(n+1)} + \operatorname{Var}(W_{n-1}) \ .$$

Induktiv folgt

$$\operatorname{Var}(W_n) = \sigma^2 \sum_{k=2}^{n+1} m^{-k} \leq \frac{\sigma^2 m}{m-1} < \infty \ .$$

Also ist $(W_n)_n$ in $\mathcal{L}^2$ beschränkt, also folgt mit Satz 1.5.2 $W_n \to W_\infty$ in $\mathcal{L}^2$, somit auch in $\mathcal{L}^1$ und speziell $\mathbb{E}(W_\infty) = \mathbb{E}(W_0) = 1$. □

Satz 2.5.3 (Wald'sche Gleichheit, Satz von Blackwell-Girshick). *Es sei* $(X_n)_n$ *eine an eine Filtrierung* $(\mathcal{F}_n)_n$ *in* $\mathcal{F}$ *adaptierte Folge identisch verteilter, integrierbarer, reeller Zufallsvariablen über* $(\Omega, \mathcal{F}, \mathbb{P})$, *wobei* X_{n+1} *für jedes* $n \in \mathbb{N}$ *von* $\mathcal{F}_n$ *unabhängig sei.* $(S_n)_n$ *sei die Folge der Partialsummen* $S_n := X_1 + \cdots + X_n$. T *sei eine integrierbare Stoppzeit bzgl.* $(\mathcal{F}_n)_n$ *mit Werten in* $\mathbb{N}$. *Dann gilt:*

(i) S_T *ist integrierbar und*

$$\mathbb{E}(S_T) = \mathbb{E}(T)\,\mathbb{E}(X_1) \ . \qquad \textit{(Wald'sche Identität).}$$

(ii) Sind die X_n *sogar quadratisch integrierbar und zentriert, so trifft dies auch auf* S_T *zu und es gilt*

$$\mathbb{E}(S_T^2) = \mathbb{E}(T)\,\mathbb{E}(X_1^2) \ . \qquad \textit{(Blackwell-Girshick).}$$

Wäre die Stoppzeit T beschränkt, so würde man das Martingal $S_n^* = S_n - n\mathbb{E}(X_1)$ betrachten und erhielte mit Satz 1.3.5 $\mathbb{E}(S_T^*) = \mathbb{E}(S_1^*) = 0$, also (i). Die Bedingung $\mathbb{E}(T) < \infty$ erzwingt ein anderes Vorgehen:

Beweis. *(i):* Wir errechnen

$$\begin{aligned}\mathbb{E}(|S_T|) &= \sum_{n=1}^\infty \mathbb{E}(|S_n|\mathbb{1}_{\{T=n\}}) \leq \sum_{n=1}^\infty \sum_{i=1}^n \mathbb{E}(|X_i|\mathbb{1}_{\{T=n\}}) = \sum_{i=1}^\infty \sum_{n=i}^\infty \mathbb{E}(|X_i|\mathbb{1}_{\{T=n\}}) \\ &= \sum_{i=1}^\infty \mathbb{E}(|X_i|\mathbb{1}_{\{T\geq i\}}) \ .\end{aligned} \tag{2.5.2}$$

Für $i \geq 2$ ist X_i ist von $\mathcal{F}_{i-1}$ unabhängig, also von $\mathbb{1}_{\{T\geq i\}}$ (man beachte, dass $\{T < i\} = \{T \leq i-1\} \in \mathcal{F}_{i-1}$ gilt). Es gilt $\{T \geq 1\} = \Omega$. Somit

$$\mathbb{E}(|X_i|\mathbb{1}_{\{T\geq i\}}) = \mathbb{P}(T \geq i)\,\mathbb{E}(|X_i|) = \mathbb{P}(T \geq i)\,\mathbb{E}(|X_1|) \ ,$$

also

$$\mathbb{E}(|S_T|) \le \sum_{i=1}^{\infty} \mathbb{P}(T \ge i)\, \mathbb{E}(|X_1|) = \mathbb{E}(T)\, \mathbb{E}(|X_1|) < \infty \ ,$$

da $\mathbb{P}(T < \infty) = 1$.

Also ist S_T integrierbar. Es folgt die Wald'sche Identität, wenn man S_T anstelle von $|S_T|$ betrachtet. Wegen der bewiesenen Konvergenz der Reihe $\sum_{n=1}^{\infty} \sum_{i=1}^{n} \left|\mathbb{E}(X_i \mathbb{1}_{\{T=n\}})\right|$ ist die Vertauschung der Summen in (2.5.2) gerechtfertigt.

(ii): Sei $Y_n := X_n \mathbb{1}_{\{n \le T\}} \in \mathcal{L}^2$. Wir machen uns klar, dass $\mathbb{E}(Y_m Y_n) = 0$ für $m \neq n$ gilt: Für $m < n$ ist $Y_m Y_n = X_n X_m \mathbb{1}_{\{n \le T\}}$ wegen $\{n \le T\} \subset \{m \le T\}$. Nun folgt aus der Unabhängigkeit von X_n und $X_m \mathbb{1}_{\{n \le T\}}$:

$$\mathbb{E}(Y_m Y_n) = \mathbb{E}(X_n)\, \mathbb{E}(X_m \mathbb{1}_{\{n \le T\}}) = 0, \qquad m \not n \ . \tag{2.5.3}$$

Auch X_n^2 und $\mathbb{1}_{\{n \le T\}}$ sind unabhängig, also folgt

$$\mathbb{E}(Y_n^2) = \mathbb{E}(X_n^2)\, \mathbb{P}(n \le T) = \mathbb{E}(X_1^2)\, \mathbb{P}(n \le T) \ , \quad n \in \mathbb{N} \ .$$

Somit gilt

$$\sum_{n \in \mathbb{N}} \mathbb{E}(Y_n^2) = \mathbb{E}(X_1^2)\, \mathbb{E}(T) < \infty \ . \tag{2.5.4}$$

Aus (2.5.3) und (2.5.4) folgt, dass die Reihe $\sum_{i=1}^{\infty} Y_i$ in $\mathcal{L}^2$ konvergiert, denn man sieht leicht, dass $\lim_{n,m \to \infty} \mathbb{E}((Y_m + \cdots + Y_n)^2) = 0$.

Es gilt $S_T = \sum_{i=1}^{\infty} X_i 1_{\{i \le T\}} = \sum_{i=1}^{\infty} Y_i$ fast sicher. Also ist $S_T \in \mathcal{L}^2$ und

$$\lim_{n \to \infty} \sum_{i=1}^{n} Y_i = S_T \qquad \text{in } \mathcal{L}^2 \ .$$

Schließlich gilt nach dem Satz von der majorisierten Konvergenz

$$\mathbb{E}(S_T^2) = \lim_{n \to \infty} \mathbb{E}\Big(\Big(\sum_{i=1}^{n} Y_i\Big)^2\Big) = \lim_{n \to \infty} \sum_{i=1}^{n} \mathbb{E}(Y_i^2) = \sum_{i=1}^{\infty} \mathbb{E}(Y_i^2) \ ,$$

also folgt die Identität in (ii) aus (2.5.4), und die Wald'sche Identität liefert die Zentriertheit von S_T. □

Ist X_1 in Teil (ii) von Satz 2.5.3 nicht zentriert und bezeichnet μ den endlichen Erwartungswert und σ^2 die endliche Varianz, kann man mittels eines analogen Beweises zeigen:

$$\mathrm{Var}(S_T) = \sigma^2\, \mathbb{E}(T) + \mu^2\, \mathrm{Var}(T).$$

Tatsächlich haben wir diese Formel zur Bestimmung von $\mathrm{Var}(W_n)$ verwendet.

Kapitel 3

Die Brown'sche Bewegung

Die Brown'sche Bewegung ist der bei weitem wichtigste stochastische Prozess in stetiger Zeit, und ihre Verteilung ist das wichtigste Wahrscheinlichkeitsmaß in der Wahrscheinlichkeitstheorie. Sie ist unabdingbar für die Beschreibung vieler zufälliger Phänomene in vielen Bereichen (Physik, Statistik, Finanzmathematik, Chemie) und bietet unerschöpfliche offene Forschungsgebiete in vielerlei Richtungen. In diesem Kapitel stellen wir die Brown'sche Bewegung vor, geben mehrere ihrer Konstruktionen an, formulieren und beweisen einige ihrer interessantesten Eigenschaften und Beziehungen zu anderen Prozessen.

3.1 Definition und allgemeine Bemerkungen

Definition 3.1.1 (Brown'sche Bewegung). *Eine Brown'sche Bewegung ist ein reellwertiger stochastischer Prozess $B = (B_t)_{t\in[0,\infty)}$ in stetiger Zeit mit den folgenden sechs charakteristischen Eigenschaften:*

(i) $B_0 = 0$ *fast sicher,*

(ii) Gauß'scher Prozess:*Prozess!Gauß'scher Für jedes $m \in \mathbb{N}$ und jede $0 \leq t_1 < \cdots < t_m$ ist der Vektor $(B_{t_1}, \dots, B_{t_m})$ Gauß'sch verteilt,*

(iii) Unabhängigkeit der Zuwächse: *Für jedes $m \in \mathbb{N}$ und jede $0 \leq t_0 < t_1 < \cdots < t_m$ sind die Größen $B(t_m) - B(t_{m-1}), B(t_{m-1}) - B(t_{m-2}), \dots, B(t_1) - B(t_0)$ unabhängig,*

(iv) Stationarität der Zuwächse: *Für jedes $0 \leq s \leq t$ stimmen die Verteilungen von $B_t - B_s$ und B_{t-s} überein,*

(v) Für jedes $t > 0$ ist B_t $\mathcal{N}(0, t)$-verteilt,

(vi) Stetigkeit: *Die Pfade $t \mapsto B_t$ sind fast sicher stetig.*

Also ist eine Brown'sche Bewegung ein stetiger, in Null startender Gauß'scher Prozess mit unabhängigen und stationären Zuwächsen, der zum Zeitpunkt Eins die Standardnormalverteilung besitzt. Wir können $B = (B_t)_{t\in[0,\infty)}$ als eine Zufallsgröße mit Werten in der Menge $\mathcal{C}([0, \infty))$

der stetigen Funktionen $[0, \infty) \to \mathbb{R}$ mit Start in 0 auffassen.

Das berühmte *Wiener-Maß* ist die Verteilung einer Brown'schen Bewegung:

Definition 3.1.2 (Wiener-Maß). *Sei B eine Brown'sche Bewegung, also eine Zufallsgröße mit Werten in $\mathcal{C}([0, \infty))$, definiert auf einem geeigneten Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbb{P})$. Dann heißt das Bildmaß $\mathbb{P} \circ B^{-1}$ auf $\mathcal{C}([0, \infty))$ das* Wiener-Maß.

Die Brown'sche Bewegung erhielt ihren Namen nach dem Botaniker R. Brown, der 1827 in Flüssigkeiten aufgeschwemmte kleine Teilchen beobachtete, die unablässig wirre unregelmäßige Bewegungen vollführen. In der zweiten Hälfte des 19. Jahrhunderts versuchte sich eine Vielzahl von Wissenschaftlern unterschiedlicher Ausrichtung an der Erklärung dieses Phänomens. Erst Anfang des 20. Jahrhunderts machten Einstein, Langevin und Smoluchovski wesentliche Fortschritte in der quantitativen Beschreibung dieses Phänomens. Man ging vereinfachend davon aus, dass die wirren Bewegungen durch Zusammenstöße mit zufälligen unregelmäßigen und sich unabhängig voneinander bewegenden Flüssigkeitsmolekülen verursacht wird. Dabei werden Wechselwirkungen in den Zeitintervallen zwischen den Zusammenstößen vernachlässigt. Dadurch erhält man näherungsweise einen unregelmäßigen Polygonzug, der sich mit Wahrscheinlichkeit $p_t(x)$ zum Zeitpunkt t im Punkte x aufhalten sollte. Unter ein paar plausiblen vereinfachenden Annahmen zeigte Einstein, dass diese Aufenthaltsdichte die Gleichung

$$\partial_t p_t(x) = D \partial_x^2 p_t(x) \tag{3.1.1}$$

erfüllen sollte, wobei $D \in (0, \infty)$ eine Konstante, die *Diffusionskonstante*, ist. Wegen der Anfangsbedingung $p_0(x) = \delta_0(x)$ (d. h. die Delta-Distribution) erhält man die Lösung

$$p_t(x) = \frac{1}{\sqrt{4\pi t D}} \mathrm{e}^{-\frac{x^2}{4Dt}}, \qquad t > 0, x \in \mathbb{R}, \tag{3.1.2}$$

also die Gauß'sche Dichte. Die Diffusionskonstante D hängt ab von physikalischen Eigenschaften der betrachteten Flüssigkeit und der Teilchen. Man stellte heuristische Anforderungen an das mathematische Modell auf, die in Definition 3.1.1 zusammengestellt werden, wobei man allerdings aus physikalischen Gründen postulierte, dass die Aussagen in (ii) und (iii) nur für t_i's mit $t_i - t_{i-1} \geq t^*$ gelten solle, wobei $t^* > 0$ von D abhängt. Definition 3.1.1 stellt also eine Idealisierung dar. Dort haben wir auch $D = \frac{1}{2}$ gesetzt, was die 'Geschwindigkeit' der Brown'schen Bewegung so steuert, dass sie zum Zeitpunkt Eins *Standard*normalverteilt ist.

Als ein mathematisches Objekt wurde die Brown'sche Bewegung 1923 von Norbert Wiener in die Wahrscheinlichkeitstheorie eingeführt, der die Existenz eines solchen stochastischen Prozesses als eine $\mathcal{C}([0, \infty))$-wertige Zufallsgröße bewies. Daher heißt dieser Prozess auch der *Wiener'sche Prozess*. In den 1930er bis 1950er Jahren wurden von Paul Lévy viele tiefliegende Resultate über die Brown'sche Bewegung erzielt, und es entstand eine stochastische Theorie der Diffusionsprozesse.

Die überragende Bedeutung der Brown'schen Bewegung in der Wahrscheinlichkeitstheorie basiert auf einer Vielzahl von Gründen:

(i) Historisch diente die Brown'sche Bewegung als wesentliche Orientierung für die systematische Erforschung spezieller Eigenschaften stochastischer Prozesse, für die die Brown'sche Bewegung jeweils das Standardbeispiel ist.

(ii) Viele stochastische Prozesse lassen sich durch Transformationen aus der Brown'schen Bewegung erzeugen.

(iii) Als Funktionalvariante des Zentralen Grenzwertsatzes entsteht die Brown'sche Bewegung als natürlicher universeller Reskalierungsgrenzwert einer großen Zahl von zeitdiskreten Prozessen.

3.2 Konstruktionen der Brown'schen Bewegung

Es gibt (mindestens) drei alternative Konstruktionen der Brown'schen Bewegung:

(i) Die gemeinsamen Verteilungen der Zufallsgrößen $B(t_0), B(t_1), \dots, B(t_m)$ mit $m \in \mathbb{N}$ und $0 \leq t_0 < t_1 < \cdots < t_m$ bilden eine konsistente Familie von Verteilungen. Nach dem Kolmogorov'schen Erweiterungssatz gibt es eine Verteilung $\mathbb{P}$ auf $\mathbb{R}^{[0,\infty)}$, deren Projektionen auf endliche Teilkoordinaten gerade diese Verteilungen sind. Mit Hilfe des Kolmogorov-Chentsov'schen Satzes über die Existenz einer stetigen Modifikation sieht man, dass $\mathbb{P}$ auf $\mathcal{C}([0,\infty))$ konzentriert ist.

(ii) Wir betrachten eine Orthonormalbasis $(b_n)_{n\in\mathbb{N}}$ im Raum $L^2([0,1])$ und eine Folge $(X_n)_{n\in\mathbb{N}}$ unabhängiger $\mathcal{N}$-verteilter Zufallsgrößen. Dann zeigt man, dass die Folge

$$B_t^{(n)} = \sum_{m=1}^{n} X_m \langle \mathbb{1}_{[0,t]}, b_m \rangle$$

für jedes $t \in [0,1]$ im L^2-Sinn eine Cauchy-Folge ist, und zeigt, dass der L^2-Grenzwert $B_t = \lim_{n\to\infty} B_t^{(n)}$ die Eigenschaften (i), (ii) und (iii) in Definition 3.1.1 erfüllt. Die Stetigkeit der Abbildung $t \mapsto B_t$ erreicht man durch geschickte Wahl der Orthonormalbasis $(b_n)_{n\in\mathbb{N}}$ oder wie in (i) mit Hilfe des Satzes von Kolmogorov-Chentsov.

(iii) Wir reskalieren eine gewöhnliche Irrfahrt $(S_n)_{n\in\mathbb{N}_0}$ auf $\mathbb{Z}$, indem wir $B_t^{(n)} = n^{-\frac{1}{2}} S_{\lfloor nt \rfloor}$ betrachten oder aber auch $\widetilde{B}_{i/n}^{(n)} = n^{-\frac{1}{2}} S_i$ mit linearer Interpolation. Der Zentrale Grenzwertsatz zeigt, dass die endlich-dimensionalen Verteilungen von $B^{(n)} = (B_t^{(n)})_{t\in[0,1]}$ und auch die von $\widetilde{B}^{(n)} = (\widetilde{B}_t^{(n)})_{t\in[0,1]}$ gegen die einer Brown'schen Bewegung konvergieren. Mit Hilfe des Kompaktheitskriteriums von Prohorov für Folgen von Verteilungen und dem Satz von Arzelà-Ascoli erhält man eine Grenzwertverteilung dieser beiden Prozesse für $n \to \infty$. Der Grenzprozess entpuppt sich als eine Brown'sche Bewegung.

Wir werden im Folgenden diese Konstruktionen ausführlich skizzieren.

3.2.1 Konstruktion mit dem Kolmogorov'schen Erweiterungssatz

Unser Ziel ist, eine Familie von reellwertigen Zufallsgrößen B_t zu konstruieren, die (unter Anderem) durch gemeinsame Verteilungen von je endlich vielen von ihnen gegeben sind.

Definition 3.2.1 (Endlich-dimensionale Verteilungen). *Die* endlich-dimensionalen Verteilungen *eines stochastischen Prozesses* $X = (X_t)_{t\in[0,\infty)}$ *sind die Verteilungen der Vektoren*

$(X(t_0), X(t_1), \dots, X(t_m))$ *mit* $m \in \mathbb{N}$ *und* $0 \le t_0 < t_1 < \cdots < t_m$*, also die Verteilungen der Projektionen*

$$\pi_T : \mathbb{R}^{[0,\infty)} \to \mathbb{R}^T, \qquad f \mapsto (f(t))_{t \in T}, \qquad \textit{mit } T \sqsubset [0,\infty),$$

wobei $\sqsubset$ *heißt, dass* T *eine endliche Teilmenge ist.*

Die endlich-dimensionalen Verteilungen legen die Verteilung des stochastischen Prozesses $X = (X_t)_{t \in [0,\infty)}$ als eine $\mathbb{R}^{[0,\infty)}$-wertige Zufallsgröße eindeutig fest, denn die Produkt-σ-Algebra auf $\mathbb{R}^{[0,\infty)}$ wird per Definition durch die Urbild-σ-Algebren der Projektionen π_T mit $T \sqsubset [0,\infty)$ erzeugt.

In Definition 3.1.1 sind zwar nicht die endlich-dimensionalen Verteilungen gegeben, aber man kann sie leicht aus der Bedingung (ii) errechnen, was wir auch gleich machen werden.

Definition 3.2.2 (Gauß'scher Prozess, zentrierter Prozess, Kovarianzfunktion). *Es sei* $X = (X_t)_{t \in [0,\infty)}$ *ein stochastischer Prozess.*

(i) Wir nennen X *einen* Gauß'schen Prozess*, Prozess!Gauß'scherwenn alle endlich-dimensionalen Verteilungen von* X *Gauß'sch sind.*

(ii) Der Prozess X *heißt* zentriert*, wenn alle* X_t *den Erwartungswert Null haben.*

(iii) Falls X *quadratisch integrierbar ist (d. h.* $\mathbb{E}[X_t^2] < \infty$ *für jedes* $t \ge 0$*), so nennen wir die Abbildung* $K : [0,\infty)^2 \to \mathbb{R}$*, definiert durch* $K(s,t) = \mathrm{Cov}(X_s, X_t)$*, die* Kovarianzfunktion *des Prozesses.*

Also ist die Verteilung eines zentrierten Gauß'schen Prozesses eindeutig festgelegt durch seine Kovarianzfunktion. Zunächst machen wir uns klar, dass wir eine Brown'sche Bewegung auch wie folgt charakterisieren können:

Lemma 3.2.3. *Ein Prozess* $B = (B_t)_{t \in [0,\infty)}$ *ist genau dann eine Brown'sche Bewegung, wenn* B *ein in Null startender zentrierter Gauß'scher Prozess mit Kovarianzfunktion* $K(s,t) = \min\{s,t\}$ *und stetigen Pfaden ist. Insbesondere ist für jede* $0 \le t_1 < \cdots < t_m$ *der Vektor* $(B_{t_1}, \dots, B_{t_m})$ *ein* m*-dimensionaler zentrierter normalverteilter Vektor mit Kovarianzmatrix* $(\min\{t_i, t_j\})_{i,j=1,\dots,m}$.

Beweisskizze. Falls B eine Brown'sche Bewegung ist, so sind für $0 \le s \le t$ insbesondere B_s und $B_t - B_s$ unabhängig, zentriert und normalverteilt mit Varianzen s bzw. $t - s$. Also ist auch (B_s, B_t) zentriert und normalverteilt, und es gilt

$$\mathrm{Cov}(B_s, B_t) = \mathrm{Cov}(B_s, B_t - B_s) + \mathrm{Cov}(B_s, B_s) = \mathbb{V}(B_s) = s = \min\{s,t\}.$$

Wenn andersherum B ein in Null startender zentrierter Gauß'scher Prozess mit Kovarianzfunktion $K(s,t) = \min\{s,t\}$ und stetigen Pfaden ist, so ist zunächst klar, dass Eigenschaften (i), (ii), (v) und (vi) in Definition 3.1.1 erfüllt sind. Dass die Zuwächse dann auch unabhängig sind, sieht man z. B., indem man die Verteilung des Vektors $(B(t_m) - B(t_{m-1}), B(t_{m-1}) - B(t_{m-2}), \dots, B(t_1) - B(t_0))$ mit Hilfe einer linearen Abbildung aus $(B_{t_0}, \dots, B_{t_m})$ erzeugt und das Bildmaß der Gauß'schen Verteilung als $\mathcal{N}(0, \mathrm{Diag}(t_m - t_{m-1}, \dots, t_1 - t_0))$ identifiziert. Ähnlich sieht man, dass die Zuwächse auch stationär sind. $\square$

Nun wollen wir beweisen, dass ein solcher Prozess existiert. Die Konstruktion basiert auf dem folgenden Existenz- und Eindeutigkeitssatz. Wir nennen eine Familie $(\mu_T)_{T \sqsubset [0,\infty)}$ von Verteilungen μ_T auf $\mathbb{R}^T$ *konsistent*, wenn für je zwei Mengen $S \subset T \sqsubset [0,\infty)$ das Maß μ_S gleich der Projektion von μ_T auf $\mathbb{R}^S$ ist, d. h. wenn $\mu_S = \mu_T \circ \pi_{S,T}^{-1}$ das Bildmaß von μ_T unter der Abbildung $\pi_{S,T} : \mathbb{R}^T \to \mathbb{R}^S$, $\pi_{S,T}((x_t)_{t\in T}) = (x_t)_{t\in S}$, ist. In Termen von Zufallsvektoren $(X_t^{(T)})_{t\in T}$ mit Verteilung μ_T heißt diese Bedingung, dass die Verteilungen von $(X_t^{(T)})_{t\in S}$ und $(X_t^{(S)})_{t\in S}$ identisch sein müssen.

Satz 3.2.4 (Kolmogorov'scher Erweiterungssatz). *Es sei $(\mu_T)_{T\sqsubset[0,\infty)}$ eine konsistente Familie von Verteilungen μ_T auf $\mathbb{R}^T$. Dann existiert ein Wahrscheinlichkeitsmaß μ auf $\mathbb{R}^{[0,\infty)}$, sodass für jedes $T \sqsubset [0,\infty)$ das Maß μ_T gleich dem Bildmaß von μ unter der Projektion $\mathbb{R}^{[0,\infty)} \to \mathbb{R}^T$ ist.*

Beweise dieses Satzes finden sich in vielen Büchern über Maßtheorie oder Wahrscheinlichkeitstheorie. Es handelt sich um die zeitlich stetige Variante eines Spezialfalls des *Satzes von Ionescu Tulcea*, der die Konstruktion allgemeinerer stochastische Prozesse mit Pfaden in beliebigen polnischen Räumen (an Stelle von $\mathbb{R}$) klärt.

Um Satz 3.2.4 auf die Brown'sche Bewegung anzuwenden, wählt man μ_T für $T = \{t_1, \dots, t_m\}$ mit $m \in \mathbb{N}$ und $0 \le t_1 < \dots < t_m$ als die zentrierte Normalverteilung auf $\mathbb{R}^m$ mit Kovarianzmatrix $(\min\{t_i, t_j\})_{i,j=1,\dots,m}$ und prüft die Konsistenz dieser Familie. Eine elementare, aber lästige formale Rechnung zeigt, dass die Familie dieser Verteilungen konsistent ist: Wenn der Vektor $(X_{t_1}, \dots, X_{t_m})$ die Verteilung μ_T hat, so hat für jedes $i \in \{1, \dots, m\}$ der Vektor $X' = (X_{t_1}, \dots, X_{t_{i-1}}, X_{t_{i+1}}, \dots, X_{t_m})$ die Verteilung $\mu_{T\setminus\{t_i\}}$ auf dem $\mathbb{R}^{m-1}$. Iteration dieser Aussage ergibt die behauptete Konsistenz. Also existiert nach dem Kolmogorov'schen Erweiterungssatz ein zentrierter Gauß'scher Prozess B mit Start in Null und der Kovarianzfunktion $K(s,t) = \min\{s,t\}$. Als zugrundeliegenden messbaren Raum kann man z. B. die Menge $\mathbb{R}^{[0,\infty)}$ mit der Produkt-σ-Algebra wählen.

Allerdings ist noch nicht klar, dass auch die Eigenschaft (vi) erfüllt ist, also die Stetigkeit der Pfade. Um diese zu erhalten, muss man zu einer geeigneten Modifikation übergehen.

Definition 3.2.5 (Modifikation, ununterscheidbar). *Seien X und Y zwei reellwertige stochastische Prozesse, die auf dem selben Wahrscheinlichkeitsraum definiert sind.*

(i) X und Y heißen Modifikationen voneinander*, wenn für jedes $t \in [0,\infty)$ fast sicher gilt: $X_t = Y_t$.*

(ii) X und Y heißen ununterscheidbar*, wenn eine Nullmenge N existiert mit $\{X_t \ne Y_t\} \subset N$ für jedes $t \in [0,\infty)$.*

Der Unterschied zwischen diesen Begriffen ist also nur, dass die Nullmenge $\{X_t \ne Y_t\}$ bei Modifikationen von t abhängen darf, aber ihre (überabzählbare!) Vereinigung über t beim Begriff der Ununterscheidbarkeit immer noch eine Nullmenge sein muss. Man kann zeigen, dass die beiden Begriffe zusammenfallen, wenn beide Prozesse fast sicher rechtsseitig stetig sind.

Unser Ziel ist also, aus dem Prozess, dessen Existenz durch Satz 3.2.4 gesichert ist, eine Modifikation mit stetigen Pfaden zu konstruieren. Dies wird durchgeführt im Beweis des folgenden Satzes. Wir nennen eine Funktion $f : [0,\infty) \to \mathbb{R}$ *lokal zum Parameter γ Hölder-stetig*, wenn für jedes $t \in [0,\infty)$ ein $\varepsilon > 0$ existiert, so dass f in $[t-\varepsilon, t+\varepsilon] \cap [0,\infty)$ Hölder-stetig zum Parameter γ ist, also so dass ein $C > 0$ existiert mit

$$|f(s) - f(r)| \le C|s-r|^\gamma, \qquad s, r \in [t-\varepsilon, t+\varepsilon] \cap [0,\infty).$$

Der folgende Satz sagt, dass man einen Prozess auf einer Nullmenge so abändern kann, dass er diese Eigenschaft mit Wahrscheinlichkeit Eins hat, falls seine Erwartungswerte eine gewisse Stetigkeit aufweisen.

Satz 3.2.6 (Kolmogorov-Chentsov). *Sei $X = (X_t)_{t\in[0,\infty)}$ ein reellwertiger stochastischer Prozess. Für jedes $T > 0$ gebe es Zahlen $\alpha, \beta, C > 0$ mit*

$$\mathbb{E}\big[|X_t - X_s|^\alpha\big] \leq C|t-s|^{1+\beta} \qquad s, t \in [0, T]. \tag{3.2.1}$$

Dann gilt für jedes $\gamma \in (0, \frac{\beta}{\alpha})$:

(i) Es gibt eine Modifikation $\widetilde{X}$ von X, die lokal zum Parameter γ Hölder-stetige Pfade hat.

(ii) Zu jedem $\varepsilon > 0$ und $T \in (0, \infty)$ gibt es ein $K > 0$, so dass

$$\mathbb{P}\Big(|X_t - X_s| \leq K|t-s|^\gamma \text{ für alle } s, t \in [0, T]\Big) \geq 1 - \varepsilon. \tag{3.2.2}$$

Der Beweis dieses Satzes ist lang und technisch, man findet ihn in vielen Büchern über stochastische Prozesse. Um ihn auf die Brown'sche Bewegung anzuwenden, muss man prüfen, ob man für jedes $T > 0$ geeignete $\alpha, \beta, C > 0$ findet, so dass (3.2.1) erfüllt ist. Dies ist relativ einfach. Wenn B ein Prozess ist, der die Eigenschaften (i) – (v) aus Definition 3.1.1 hat, so ist für alle $n \in \mathbb{N}$ und alle $s, t \in [0, \infty)$ für ein geeignetes C_n (das man mit Hilfe von Fakultäten explizit ausdrücken kann)

$$\mathbb{E}\big[|B_t - B_s|^{2n}\big] = \mathbb{E}\big[\big(\sqrt{|t-s|}|B_1|\big)^{2n}\big] = C_n|t-s|^n, \tag{3.2.3}$$

wie eine beliebte Übungsaufgabe über Normalverteilungen zeigt. Sei nun $n \geq 2$, dann können wir $\alpha = 2n$ und $\beta = n - 1$ setzen und erhalten aus Satz 3.2.6 die Existenz einer lokal zu jedem Parameter $\gamma \in (0, \frac{n-1}{2n})$ Hölder-stetigen Modifikation von B. Da $n \geq 2$ beliebig ist, erhalten wir zusammenfassend die folgende Aussage.

Korollar 3.2.7. *Es existiert eine Brown'sche Bewegung B auf einem geeigneten Wahrscheinlichkeitsraum. Für jedes $\gamma \in (0, \frac{1}{2})$ hat diese Brown'sche Bewegung lokal zum Parameter γ Hölder-stetige Pfade, und die Aussage in Satz 3.2.6(ii) gilt für B an Stelle von X.*

Damit ist die (Skizze der) ersten Konstruktion der Brown'schen Bewegung abgeschlossen.

3.2.2 Funktionalanalytische Konstruktion

In diesem Abschnitt konstruieren wir die Brown'sche Bewegung zunächst nur auf dem Zeitintervall $[0, 1]$ statt $[0, \infty)$. Wir betrachten den Hilbertraum $H = L^2[0, 1]$, also die Menge aller bezüglich des Lebesgue-Maßes λ quadratintegrierbaren Funktionen, zusammen mit dem Skalarprodukt $\langle f, g\rangle = \int_{[0,1]} f(x)g(x)\,\lambda(\mathrm{d}x)$ und der davon erzeugten Norm $\|f\| = \sqrt{\langle f, f\rangle}$. Sei $(b_n)_{n\in\mathbb{N}}$

eine Orthonormalbasis von H, also $\langle b_n, b_m\rangle = \delta_{n,m}$ (Kroneckersymbol) für alle $n, m \in \mathbb{N}$ und

$$\lim_{n\to\infty} \Big\| f - \sum_{m=1}^{n} \langle f, b_m\rangle\, b_m \Big\| = 0, \qquad f \in H.$$

Insbesondere gelten

$$\langle f, g\rangle = \sum_{m\in\mathbb{N}} \langle f, b_m\rangle\, \langle g, b_m\rangle, \qquad f, g \in H,$$

und der Spezialfall

$$\|f\|^2 = \sum_{m\in\mathbb{N}} \langle f, b_m\rangle^2, \qquad f \in H,$$

die *Parseval'sche Gleichung*. Sei nun $(X_n)_{n\in\mathbb{N}}$ eine Folge unabhängiger, $\mathcal{N}$-verteilter Zufallsgrößen, die auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbb{P})$ definiert seien. Wir setzen

$$B_t^{(n)} = \int_{[0,1]} \mathbb{1}_{[0,t]}(s) \sum_{m=1}^{n} X_m b_m(s)\,\mathrm{d}s = \sum_{m=1}^{n} X_m \langle \mathbb{1}_{[0,t]}, b_m\rangle, \qquad n \in \mathbb{N}, t \in [0,1].$$

Dann sind die $B_t^{(n)}$ quadratintegrierbare Zufallsgrößen, denn man sieht leicht, dass für jedes $n, m \in \mathbb{N}$ mit $m \le n$ gilt

$$\begin{aligned}\mathbb{E}\big[(B_t^{(n)} - B_t^{(m)})^2\big] &= \mathbb{E}\Big[\sum_{k=m+1}^{n} X_k \langle \mathbb{1}_{[0,t]}, b_k\rangle \sum_{l=m+1}^{n} X_l \langle \mathbb{1}_{[0,t]}, b_l\rangle\Big] = \sum_{k=m+1}^{n} \langle \mathbb{1}_{[0,t]}, b_k\rangle^2 \\ &\le \sum_{k=m+1}^{\infty} \langle \mathbb{1}_{[0,t]}, b_k\rangle^2.\end{aligned} \tag{3.2.4}$$

Da die Reihe auf der rechten Seite für $m = 0$ wegen der Parseval'schen Gleichung gleich $\|\mathbb{1}_{[0,t]}\|^2 = t$ ist, gilt $B_t^{(n)} \in \mathcal{L}^2$ für jedes $t \in [0,1]$ und jedes $n \in \mathbb{N}$. Ferner sieht man auch direkt von (3.2.4), dass $(B_t^{(n)})_{n\in\mathbb{N}}$ eine Cauchy-Folge in $\mathcal{L}^2$ ist. Wegen der Vollständigkeit des Raumes $L^2(\mathbb{P})$ existiert ein L^2-Grenzwert B_t. Natürlich konvergiert für jedes $m \in \mathbb{N}$ und jede $0 \le t_1 < \cdots < t_m$ auch der Vektor $(B_{t_1}^{(n)}, \dots, B_{t_m}^{(n)})$ gegen den Vektor $(B_{t_1}, \dots, B_{t_m})$. Daher ist letzterer Gauß'sch verteilt und zentriert. Die Kovarianzen errechnen sich zu

$$\begin{aligned}\mathrm{Cov}(B_t^{(n)}, B_s^{(n)}) &= \mathbb{E}\Big[\Big(\sum_{m=1}^{n} X_m \langle \mathbb{1}_{[0,t]}, b_m\rangle\Big)\Big(\sum_{m=1}^{n} X_m \langle \mathbb{1}_{[0,s]}, b_m\rangle\Big)\Big] \\ &= \sum_{k,l=1}^{n} \mathbb{E}[X_k X_l] \langle \mathbb{1}_{[0,t]}, b_k\rangle \langle \mathbb{1}_{[0,s]}, b_l\rangle = \sum_{k=1}^{n} \langle \mathbb{1}_{[0,t]}, b_k\rangle \langle \mathbb{1}_{[0,s]}, b_k\rangle.\end{aligned}$$

Für $n \to \infty$ konvergiert dies gegen $\langle \mathbb{1}_{[0,s]}, \mathbb{1}_{[0,t]}\rangle = \min\{s, t\}$. Also ist $B = (B_t)_{t\in[0,1]}$ ein zentrierter Gauß-Prozess mit Kovarianzfunktion $K(s,t) = \min\{s,t\}$. Eine stetige Version erhält man durch eine Anwendung des Satzes 3.2.6 von Kolmogorov und Chentsov. Nun muss man noch die Konstruktion auf das Zeitintervall $[0, \infty)$ erweitern, was wir hier nicht ausführen wollen. Damit ist auch die Skizze dieser Konstruktion abgeschlossen.

3.2.3 Konstruktion mit reskalierten Irrfahrten

Die Skizze der dritten Konstruktion halten wir kurz, da sie Gegenstand der Vorlesung *Wahrscheinlichkeitstheorie II* ist. Sie ist sehr wichtig, da sie durch eine approximative Konstruktion zustande kommt und daher praktischen Wert für Simulationen hat. Außerdem gibt sie universelle Antworten auf viele Fragen über das asymptotische Verhalten stochastischer Prozesse in diskreter Zeit.

Wir betrachten eine Irrfahrt $(S_n)_{n\in\mathbb{N}_0}$ auf $\mathbb{R}$ mit Start in 0. Also $S_n = X_1 + \cdots + X_n$, wobei $(X_i)_{i\in\mathbb{N}}$ eine Folge unabhängiger Zufallsgrößen ist, die wir als standardisiert voraussetzen, d. h. es gelte $\mathbb{E}[X_i] = 0$ und $\mathbb{E}[X_i^2] = 1$. Also ist auch $S_n/\sqrt{n}$ standardisiert. Das Standardbeispiel ist die sogenannte einfache Irrfahrt, wo X_i die Werte 1 und -1 mit gleicher Wahrscheinlichkeit annimmt, und in diesem Fall kann man viele explizite Rechnungen anstellen.

Der Zentrale Grenzwertsatz sagt, dass die Verteilung von $n^{-\frac{1}{2}} S_n$ gegen die Standardnormalverteilung $\mathcal{N}$ konvergiert. Mit Hilfe dieses Satzes kann man auch leicht zeigen, dass die endlich-dimensionalen Verteilungen des Prozesses $B^{(n)} = (B_t^{(n)})_{t\in[0,1]}$ mit

$$B_t^{(n)} = n^{-\frac{1}{2}}\Big(S_{\lfloor tn\rfloor} + \big(tn - \lfloor tn\rfloor\big) X_{\lfloor tn\rfloor+1}\Big), \qquad n \in \mathbb{N}, t \in [0,1],$$

gerade gegen die endlich-dimensionalen Verteilungen einer Brown'schen Bewegung konvergieren. (Man beachte, dass $B^{(n)} \in \mathcal{C}([0,\infty))$ die stückweise lineare Interpolation der Punkte $(\frac{i}{n}, n^{-\frac{1}{2}} S_i)$ für $i = 0, \dots, n$ ist, sein Graph also ein Polygonzug.) Dies legt die Vermutung nahe, dass $B^{(n)}$ als eine $\mathcal{C}([0,1])$-wertige Zufallsgröße gegen eine Brown'sche Bewegung konvergieren sollte. Die Frage, ob Verteilungskonvergenz vorliegt, wird üblicherweise mit Hilfe des folgenden Satzes behandelt.

Satz 3.2.8 (Satz von Prohorov, hinreichender Teil). *Sei E ein metrischer Raum, und sei $(Y_n)_{n\in\mathbb{N}}$ eine Folge E-wertiger Zufallsgrößen. Wenn $(Y_n)_{n\in\mathbb{N}}$ straff ist (also wenn für jedes $\varepsilon > 0$ eine kompakte Menge $K \subset E$ existiert mit $\sup_{n\in\mathbb{N}} \mathbb{P}(Y_n \in K^c) < \varepsilon$), so gibt es eine Teilfolge $(n_k)_{k\in\mathbb{N}}$ von $(n)_{n\in\mathbb{N}}$ und eine E-wertige Zufallsgröße Y, gegen die Y_{n_k} in Verteilung konvergiert.*

Wir wollen Satz 3.2.8 anwenden auf $E = \mathcal{C}([0,1])$ und müssen also nur noch die Straffheit der Folge $(B^{(n)})_{n\in\mathbb{N}}$ zeigen, denn dann folgt die Existenz eines $\mathcal{C}([0,1])$-wertigen Grenzprozesses $B = (B_t)_{t\in[0,1]}$, und dieser muss die endlich-dimensionalen Verteilungen einer Brown'schen Bewegung haben, also selber eine sein.

Um die Straffheit zu zeigen, müssen wir uns über die kompakten Mengen in $\mathcal{C}([0,1])$ klar sein. Dies liefert der folgende berühmte Satz aus der Funktionalanalysis. Wir führen die folgende Maßgröße für die Stetigkeit einer Funktion ein:

$$w_\delta(f) = \sup\big\{|f(s) - f(t)| \colon s,t \in [0,1], |s-t| \le \delta\big\}, \qquad f \colon [0,1] \to \mathbb{R}, \delta > 0. \tag{3.2.5}$$

Eine Funktion $f \colon [0,1] \to \mathbb{R}$ ist also genau dann gleichmäßig stetig, wenn $\lim_{\delta\downarrow 0} w_\delta(f) = 0$ ist.

Satz 3.2.9 (Arzelà-Ascoli). *Eine Teilmenge K von $\mathcal{C}([0,1])$ hat genau dann einen kompakten Abschluss, wenn gelten:*

$$(i) \quad \sup_{f\in K} |f(0)| < \infty \qquad und \qquad (ii) \quad \lim_{\delta\downarrow 0} \sup_{f\in K} w_\delta(f) = 0.$$

Die Eigenschaft (ii) wird die gleichmäßige gleichgradige Stetigkeit von K genannt. Wenn wir nun die Straffheit der Folge $(B^{(n)})_{n\in\mathbb{N}}$ zeigen wollen, so ist nur diese Eigenschaft zu beachten, denn die $B^{(n)}$ starten alle in Null. Die Aufgabe lautet also, zu jedem $\varepsilon > 0$ eine kompakte Menge $K \subset \mathcal{C}([0,1])$ zu finden, so dass $\mathbb{P}(B^{(n)} \in K^{\mathrm{c}}) < \varepsilon$ für jedes $n \in \mathbb{N}$ gilt. Diesen Punkt leistet der Beweis des folgenden berühmten Satzes.

Satz 3.2.10 (Satz von Donsker, Invarianzprinzip). *Die Folge $(B^{(n)})_{n\in\mathbb{N}}$ ist straff. Insbesondere konvergiert $B^{(n)}$ für $n \to \infty$ in Verteilung, und der Grenzwert ist eine Brown'sche Bewegung.*

Nachdem die Straffheit gezeigt worden ist, weiß man, dass $(B^{(n)})_{n\in\mathbb{N}}$ Häufungspunkte besitzt. Aber wir wussten auch schon, dass die endlich-dimensionalen Verteilungen jedes dieser Häufungspunkte die selben sein müssen, nämlich die der Brown'schen Bewegung. Also stimmen alle Häufungspunkte miteinander überein, d. h. die ganze Folge konvergiert.

Man nennt den Satz 3.2.10 auch ein *Invarianzprinzip* oder einen *funktionalen Grenzwertsatz*. Der erste Begriff betont, dass der Grenzprozess invariant ist unter der zugrunde liegenden Schrittverteilung der Irrfahrt (solange sie standardisiert ist), und der zweite weist darauf hin, dass die Aussage eine Ausweitung des Zentralen Grenzwertsatzes auf zufällige Funktionen ist. Der Satz von Donsker kann für die Identifikation der Verteilungen einiger Funktionale des Brown'schen Pfades benutzt werden (etwa die Zeit, die sie in $(0,\infty)$ verbringt oder das Maximum $\max\{B_t : t \in [0,1]\}$), wenn man dieses Funktional asymptotisch für die einfache Irrfahrt explizit berechnen kann.

3.3 Eigenschaften der Brown'schen Bewegung

Die Brown'sche Bewegung B hat eine Reihe interessanter, teils angenehmer, teils komplizierter Eigenschaften, die zum Teil durchaus auf dem ersten Blick verwirrend sein können. Hier ist eine kleine Liste der wichtigsten Eigenschaften:

- Der reskalierte Prozess $(\frac{1}{c}B_{tc^2})_{t\in[0,\infty)}$ hat die selbe Verteilung wie B, ebenso der Prozess $(tB_{1/t})_{t\in[0,\infty)}$.
- B ist ein Markovprozess und besitzt sogar die starke Markoveigenschaft.
- Die Pfade von B sind nirgends differenzierbar.
- Ereignisse, die in der infinitesimalen Zukunft liegen (also durch die Ereignisse jedes Zeitintervalls $[0,t]$ mit $t > 0$ bestimmt werden), haben Wahrscheinlichkeit Null oder Eins.

- Für $t \to \infty$ approximiert $B_t/\sqrt{2t \log\log t}$ fast sicher jeden Wert aus $[-1,1]$, aber keine anderen.

In diesem Abschnitt werden wir diese und andere Eigenschaften beweisen.

3.3.1 Skalierung und Zeitstürzung

Eine der einfachsten Eigenschaften einer Brown'schen Bewegung B ist, dass auch $-B$ eine ist. Fahren wir fort mit der Skalierungs- und der Zeitstürzungseigenschaft.

Lemma 3.3.1 (Skalierungs- und Zeitstürzungseigenschaft). *Sei $B = (B_t)_{t\in[0,\infty)}$ eine Brown'sche Bewegung. Dann sind auch die folgenden Prozesse Brown'sche Bewegungen:*

(i) $(\frac{1}{c}B_{tc^2})_{t\in[0,\infty)}$ *mit* $c \in (0,\infty)$,

(ii) $\widetilde{B} = (\widetilde{B}_t)_{t\in[0,\infty)}$ *mit* $\widetilde{B}_0 = 0$ *und* $\widetilde{B}_t = tB_{1/t}$ *für* $t \in (0,\infty)$.

Beweisskizze. Wir benutzen Lemma 3.2.3 und müssen nur zeigen, dass die jeweiligen Prozesse zentrierte Gauß'sche Prozesse sind mit stetigen Pfaden, Start in Null und der Kovarianzfunktion $(s,t) \mapsto \min\{s,t\}$.

(i) Mit Hilfe der offensichtlichen Skalierungseigenschaften der mehrdimensionalen Normalverteilung sieht man leicht, dass $(\frac{1}{c}B_{tc^2})_{t\in[0,\infty)}$ ein zentrierter Gauß'scher Prozess mit Start in Null und Kovarianzfunktion $(s,t) \mapsto \min\{s,t\}$ ist, und die Stetigkeit der Pfade ist ebenfalls offensichtlich.

(ii) Dass $\widetilde{B}$ ein zentrierter Gauß'scher Prozess ist, ist leicht zu sehen, und die Stetigkeit der Pfade in $(0,\infty)$ ebenfalls. Für $s,t \in (0,\infty)$ ist

$$\mathrm{Cov}(\widetilde{B}_s, \widetilde{B}_t) = ts\,\mathrm{Cov}(B_{1/s}, B_{1/t}) = ts\,\min\{1/s, 1/t\} = \min\{s,t\},$$

also hat $\widetilde{B}$ die angegebene Kovarianzfunktion. Die Stetigkeit von $\widetilde{B}$ in Null beweist man wie folgt. Wir schätzen ab:

$$\limsup_{t\downarrow 0} \widetilde{B}_t = \limsup_{t\to\infty} \frac{1}{t}B_t \le \limsup_{n\to\infty} \frac{1}{n}B_n + \limsup_{n\to\infty} \frac{1}{n} \sup\big\{B_t - B_n \colon t \in [n,n+1]\big\}.$$

Nach dem Starken Gesetz der Großen Zahlen ist $\lim_{n\to\infty} \frac{1}{n}B_n = 0$ fast sicher, denn die Zufallsgrößen $B_1 - B_0, B_2 - B_1, \dots, B_n - B_{n-1}$ sind unabhängige identisch verteilte zentrierte Zufallsgrößen. Die Zufallsgrößen $\frac{1}{n}(\max_{[n,n+1]} B - B_n)$ sind unabhängig und haben jeweils die Verteilung von $\frac{1}{n}\max_{[0,1]} B$. Wir wollen das 1. Borel-Cantelli-Lemma anwenden, um zu zeigen, dass ihr limes superior gleich Null ist. Aus dem Spiegelungsprinzip (siehe Satz 3.3.8) folgt für jedes $\varepsilon \in (0,1)$:

$$\begin{aligned}\mathbb{P}\Big(\frac{1}{n}\Big(\max_{[n,n+1]} B - B_n\Big) > n^{-\varepsilon}\Big) &= \mathbb{P}\big(\max_{[0,1]} B > n^{1-\varepsilon}\big) = 2\mathbb{P}(B_1 > n^{1-\varepsilon}) = \frac{2}{\sqrt{2\pi}} \int_{n^{1-\varepsilon}}^{\infty} \mathrm{e}^{-\frac{1}{2}u^2}\,\mathrm{d}u \\ &\le n^{-1+\varepsilon}\mathrm{e}^{-\frac{1}{2}n^{2(1-\varepsilon)}}.\end{aligned}$$

Da die rechte Seite summierbar über $n \in \mathbb{N}$ ist, folgt aus dem 1. Borel-Cantelli-Lemma, dass $\limsup_{n\to\infty} \frac{1}{n}(\max_{[n,n+1]} B - B_n) \le 0$ fast sicher. Also folgt $\limsup_{t\to\infty} \widetilde{B}_t \le 0$. Da auch $-B$

eine Brown'sche Bewegung ist, folgt auch die umgekehrte Ungleichung, $\liminf_{t\to\infty} \widetilde{B}_t \geq 0$. Dies beendet den Beweis der Stetigkeit von $\widetilde{B}$ in Null. □

3.3.2 0-1-Gesetz

Nun betrachten wir die sogenannte *infinitesimale Zukunft*. Mit $(\mathcal{F}_t)_{t\in[0,\infty)}$ bezeichnen wir die durch die Brownsche Bewegung B erzeugte Filtrierung, also $\mathcal{F}_t = \sigma(B_s\colon s \in [0,t])$. Dann heißt

$$\mathcal{F}_0^+ = \bigcap_{t\in(0,\infty)} \mathcal{F}_t \tag{3.3.1}$$

die σ-Algebra der infinitesimalen Zukunft: Sie enthält alle Ereignisse, die sich vollständig durch alle beliebig kleinen Zeitintervalle $(0,t)$ mit $t > 0$ charakterisieren lassen. Zum Beispiel liegt das Ereignis $\{\limsup_{t\downarrow 0} \frac{1}{t} B_t = 0\}$ in $\mathcal{F}_0^+$.

Satz 3.3.2 (Blumenthals 0-1-Gesetz). *Sei B eine Brown'sche Bewegung mit zugehöriger Filtrierung $(\mathcal{F}_t)_{t\in[0,\infty)}$. Dann besitzen alle Ereignisse in $\mathcal{F}_0^+$ nur die Wahrscheinlichkeiten Null oder Eins.*

Beweis. Wir betrachten $Y^{(n)} = (B_{2^{-n}+t} - B_{2^{-n}})_{t\in[0,2^{-n}]}$ für $n \in \mathbb{N}$. Dann ist $(Y^{(n)})_{n\in\mathbb{N}}$ eine unabhängige Familie von Zufallsgrößen mit Werten in der Menge $\mathcal{C}([0,2^{-n}])$ der stetigen Abbildungen $[0,2^{-n}] \to \mathbb{R}$. Die terminale σ-Algebra $\mathcal{T} = \bigcap_{n\in\mathbb{N}} \sigma(Y^{(m)}\colon m \geq n)$ ist nach dem Kolmogorov'schen 0-1-Gesetz $\mathbb{P}$-trivial, d. h. alle ihre Ereignisse haben nur die Wahrscheinlichkeiten Null oder Eins. Wegen $\mathcal{F}_0^+ = \bigcap_{t\in(0,\infty)} \mathcal{F}_t = \bigcap_{n\in\mathbb{N}} \mathcal{F}_{2^{-n+1}} = \mathcal{T}$ folgt die Aussage. □

3.3.3 Hölder-Unstetigkeit der Pfade

Als Anwendung des Blumenthal'schen 0-1-Gesetzes zeigen wir, dass die reskalierte Brown'sche Bewegung $B_t/\sqrt{t}$ jeden noch so großen Wert in beliebig kurzer Zeit erreicht:

Beispiel 3.3.3. Wir betrachten die Ersterreichenszeit eines Punktes $K > 0$ durch den Prozess $(B_t/\sqrt{t})_{t\in[0,\infty)}$, also $\tau = \inf\{t > 0\colon B_t > K\sqrt{t}\}$. Wir wollen zeigen, dass fast sicher $\tau = 0$ gilt. Um dies einzusehen, bemerken wir, dass das Ereignis $\{\tau < t\}$ in $\mathcal{F}_t$ liegt, und haben dann, dass

$$\{\tau = 0\} = \bigcap_{t>0} \{\tau < t\} \in \bigcap_{t>0} \mathcal{F}_t = \mathcal{F}_0^+.$$

Nach dem Blumenthal'schen 0-1-Gesetz ist $\mathbb{P}(\tau = 0) \in \{0,1\}$. Wegen der Skalierungseigenschaft der Brown'schen Bewegung gilt $\mathbb{P}(\tau = 0) = \inf_{t>0} \mathbb{P}(\tau < t) \geq \inf_{t>0} \mathbb{P}(B_t > 2K\sqrt{t}) = \mathbb{P}(B_1 > 2K) > 0$, also scheidet die Möglichkeit $\mathbb{P}(\tau = 0) = 0$ aus. ◇

Dieses Beispiel zeigt insbesondere, dass fast sicher $\limsup_{t\downarrow 0} B_t/\sqrt{t} \geq K$ ist für jedes $K > 0$, also dass $\limsup_{t\downarrow 0} B_t/\sqrt{t} = \infty$. (Als eine Übungsaufgabe münze man diese Aussage mit Hilfe der Zeitstürzungseigenschaft in eine Aussage für *große* Zeiten um.) Mit Hilfe der Tatsache, dass $-B$ ebenfalls eine Brown'sche Bewegung ist, sieht man, dass auch $\liminf_{t\downarrow 0} B_t/\sqrt{t} = -\infty$ gilt. Insbesondere sind die Pfade fast sicher in Null nicht differenzierbar, denn $\limsup_{t\downarrow 0} \frac{1}{t}(B_t - B_0) = \infty$. Dann ist auch klar, dass die Pfade für jedes $s \in (0,\infty)$ fast sicher nicht in s differenzierbar sind. Diese Aussage über die (mangelnde) Glattheit der Pfade wird in dem folgenden Satz sehr verstärkt.

Satz 3.3.4 (Hölder-Unstetigkeit der Pfade). *Sei $\gamma \in (\frac{1}{2}, \infty)$, dann gilt für eine Brown'sche Bewegung $B = (B_t)_{t\in[0,\infty)}$ fast sicher Folgendes: Der Pfad $t \mapsto B_t$ ist in keinem Punkt Hölder-stetig von der Ordnung γ. Insbesondere ist der Pfad auch in keinem Punkt differenzierbar.*

Man beachte also, dass es eine *von t unabhängige* Ausnahmemenge in dem Wahrscheinlichkeitsraum gibt, so dass der Pfad außerhalb dieser Menge in jedem t nicht differenzierbar ist.

Beweis. Es reicht, $(B_t)_{t\in[0,1]}$ zu betrachten. Es sei $H_{\gamma,t}$ die Menge aller in t zum Parameter γ Hölder-stetigen Funktionen $[0,1] \to \mathbb{R}$, und es sei $H_\gamma = \bigcup_{t\in[0,1]} H_{\gamma,t}$. Dann ist also zu zeigen, dass H_γ das Maß Null unter der Verteilung der Brown'schen Bewegung hat, d. h. $\mathbb{P}((B_t)_{t\in[0,1]} \in H_\gamma) = 0$.

Für jedes $t \in [0,1)$ und $w \in H_{\gamma,t}$ gibt es ein $\delta > 0$ und ein $C > 0$ mit $|w_s - w_t| \leq C|s-t|^\gamma$ für jedes $s \in [0,1]$ mit $|s-t| \leq \delta$. Wir wählen ein $k \in (\frac{2}{2\gamma-1}, \infty) \cap \mathbb{N}$ und setzen $n_0 = \lfloor (k+1)/\delta \rfloor$. Speziell ist dann für jedes $n \in \mathbb{N}$ mit $n \geq n_0$, $i = \lfloor tn \rfloor + 1$ und $l \in \{0, \dots, k-1\}$:

$$|w_{(i+l+1)/n} - w_{(i+l)/n}| \leq |w_{(i+l+1)/n} - w_t| + |w_t - w_{(i+l)/n}| \leq (2k+1)Cn^{-\gamma}.$$

Für $N \geq (2k+1)C$ und $n \geq n_0$ liegt dann also w in der Menge

$$A_{n,N,i} = \bigcap_{l=0}^{k-1} \Big\{ w \colon |w_{(i+l+1)/n} - w_{(i+l)/n}| \leq Nn^{-\gamma} \Big\}.$$

Wir setzen $A_{n,N} = \bigcup_{i=1}^n A_{n,N,i}$ sowie $A_N^{(n_0)} = \bigcap_{n=n_0}^\infty A_{n,N}$ und $A = \bigcup_{N,n_0\in\mathbb{N}} A_N^{(n_0)}$, dann ist leicht zu sehen, dass H_γ eine Teilmenge von A ist. Wegen der Unabhängigkeit der Zuwächse und weil die Dichte der Standardnormalverteilung durch 1 beschränkt ist, gilt

$$\mathbb{P}(B \in A_{n,N,i}) = \mathbb{P}(|B_{1/n}| \leq Nn^{-\gamma})^k = \mathbb{P}(|B_1| \leq Nn^{-\gamma+\frac{1}{2}})^k \leq (2N)^k n^{k(\frac{1}{2}-\gamma)}.$$

Da $A_N^{(n_0)}$ in jedem $A_{n,N}$ für jedes genügend große n enthalten ist, gilt

$$\mathbb{P}(B \in A_N^{(n_0)}) = \limsup_{n\to\infty} \mathbb{P}(B \in A_{n,N}) \leq \limsup_{n\to\infty} n(2N)^k n^{k(\frac{1}{2}-\gamma)} = 0,$$

wobei wir die Wahl von k beachtet haben, d. h. dass $1 + k(\frac{1}{2} - \gamma) < 0$. Da A die abzählbare Vereinigung der Mengen $A_N^{(n_0)}$ mit $N, n_0 \in \mathbb{N}$ ist, haben wir $\mathbb{P}((B_t)_{t\in[0,1]} \in A) = 0$. Wegen $H_\gamma \subset A$ ist der Beweis beendet. □

Bemerkung 3.3.5 (Hölder-Stetigkeit). In Abschnitt 3.3.6 wird sich heraus stellen, dass die Brown'schen Pfade Hölder-stetig zu jedem Parameter $\gamma \in (0, \frac{1}{2})$ sind, aber nicht zum Parameter $\gamma = \frac{1}{2}$. ◇

3.3.4 Die starke Markoveigenschaft

Um die Markoveigenschaft zu formulieren, benötigen wir den Begriff einer Brown'schen Bewegung, die in einem festen gegebenen Punkt $x \in \mathbb{R}$ startet. Nichts leichter als das: Wir nennen

$\mathbb{P}_x$ das Maß, unter dem $B = (B_t)_{t\geq 0}$ eine in x gestartete Brown'sche Bewegung ist. Also ist $\mathbb{P}_x$-fast sicher $(B_t - x)_{t\geq 0}$ eine Brown'sche Bewegung im Sinne der Definition 3.1.1. Man spricht oft von der Brown'schen Bewegung als der Familie $(B, (\mathbb{P}_x)_{x\in\mathbb{R}})$ oder sogar $(\Omega, \mathcal{F}, (\mathbb{P}_x)_{x\in\mathbb{R}}, B)$, wenn man den Wahrscheinlichkeitsraum betonen möchte. Dann ist die (gewöhnliche) Markoveigenschaft leicht zu formulieren und auch leicht zu beweisen (als Übungsaufgabe):

Lemma 3.3.6 (Markoveigenschaft). *Sei B eine Brown'sche Bewegung, und sei $t > 0$. Dann ist der Prozess $(B_{s+t} - B_t)_{s\in[0,\infty)}$ ebenfalls eine Brown'sche Bewegung, und sie ist unabhängig von $(B_s)_{s\in[0,t]}$. Insbesondere gilt für alle $n \in \mathbb{N}$, alle $0 < t_1 < t_2 < \cdots < t_n < t_{n+1}$ und jede messbare Menge $A \subset \mathbb{R}$ fast sicher:*

$$\mathbb{P}\big(B_{t_{n+1}} \in A \mid B_{t_0}, \dots, B_{t_n}\big) = \mathbb{P}\big(B_{t_{n+1}} \in A \mid B_{t_n}\big) = \mathbb{P}_{B_{t_n}}(B_{t_{n+1}-t_n} \in A).$$

Mit anderen Worten, zu jedem Zeitpunkt $t > 0$ kann man die Bewegung stoppen und neu los laufen lassen. Sie muss nur wissen, wo sie gerade zum Zeitpunkt t ist, dann verhält sie sich wie eine Brown'sche Bewegung mit diesem Startpunkt und hat alles, was vor dem Zeitpunkt t war, vergessen.

Nun wollen wir diese Stopp-Eigenschaft auch für *zufällige* Zeiten τ an Stelle von t haben. Dies kann aber nur klappen, wenn τ nicht in die Zukunft sehen kann, also wenn sie eine Stoppzeit ist. Die Definition dieses Begriffes in stetiger Zeit ist wie bei diskretzeitigen Prozessen: Eine $[0, \infty]$-wertige Zufallsgröße τ heißt eine *Stoppzeit*, falls für jedes $t > 0$ das Ereignis $\{\tau \leq t\}$ in der σ-Algebra $\mathcal{F}_t = \sigma(B_s : s \in [0, t])$ liegt. Die σ-Algebra $\mathcal{F}_\tau = \{A \in \mathcal{F} : A \cap \{\tau \leq t\} \in \mathcal{F}_t\}$ enthält alle Ereignisse, die vor dem Eintreffen der Stoppzeit τ liegen.

Dann können wir die starke Markoveigenschaft der Brown'schen Bewegung beweisen:

Satz 3.3.7 (Starke Markoveigenschaft). *Sei B eine Brown'sche Bewegung, und sei τ eine fast sicher endliche Stoppzeit. Dann ist der Prozess $(B_{t+\tau} - B_\tau)_{t\in[0,\infty)}$ ebenfalls eine Brown'sche Bewegung, und sie ist unabhängig von $(B_s)_{s\in[0,\tau]}$. Genauer gesagt, es gilt für jede beschränkte messbare Funktion $F : \mathbb{R}^{[0,\infty)} \to \mathbb{R}$ und für jedes $x \in \mathbb{R}$:*

$$\mathbb{E}_x\big[F\big((B_{t+\tau})_{t\in[0,\infty)}\big) \mid \mathcal{F}_\tau\big] = \mathbb{E}_{B_\tau}[F(B)] \qquad \mathbb{P}_x\text{-fast sicher.} \tag{3.3.2}$$

Beweisskizze. Die Gleichung (3.3.2) muss man nur für beschränkte stetige Funktionen F zeigen, die nur von endlich vielen Koordinaten $t_1, \dots, t_n$ abhängen, denn die Verteilung des Prozesses wird durch die Gesamtheit aller dieser Koordinatentupel bestimmt. Wir betrachten also nur eine Funktion F von der Form $F(B) = f(B_{t_1}, \dots, B_{t_N})$ mit $0 \leq t_1 < \cdots < t_N$ und einem beschränkten stetigen $f : \mathbb{R}^N \to \mathbb{R}$.

Die Abbildung $x \mapsto \mathbb{E}_x[F(B)] = \mathbb{E}_0[f(B_{t_1} + x, \dots, B_{t_N} + x)]$ ist offensichtlich stetig und beschränkt. Wir approximieren τ mit $\tau_n = 2^{-n}\lfloor 2^n\tau + 1\rfloor$. Dann gilt $\tau_n \downarrow \tau$ fast sicher für $n \to \infty$, also auch $B_{\tau_n} \to B_\tau$. Außerdem ist τ_n eine Stoppzeit. Für festes $n \in \mathbb{N}$ schränken wir nun B ein auf die Zeitmenge aller positiven ganzzahligen Linearkombinationen von $2^{-n}, t_1, \dots, t_N$ und benutzen, dass dieser Prozess in diskreter Zeit eine Markovkette auf $\mathbb{R}$ ist. Also haben wir

$$\mathbb{E}_x\big[F\big((B_{\tau_n+t})_{t\in[0,\infty)}\big) \mid \mathcal{F}_{\tau_n}\big] = \mathbb{E}_x\big[f(B_{\tau_n+t_1}, \dots, B_{\tau_n+t_N}) \mid \mathcal{F}_{\tau_n}\big] = \mathbb{E}_{B_{\tau_n}}\big[f(B_{t_1}, \dots, B_{t_N})\big]. \tag{3.3.3}$$

Für $n \to \infty$ konvergiert die rechte Seite gegen $\mathbb{E}_{B_\tau}[f(B_{t_1}, \dots, B_{t_N})]$. Aufgrund der Stetigkeit konvergiert $F((B_{\tau_n+t})_{t\in[0,\infty)})$ fast sicher und in $\mathcal{L}^1$ gegen $F((B_{\tau+t})_{t\in[0,\infty)})$, also folgt

$$\lim_{m\to\infty} \mathbb{E}_x[F((B_{\tau_m+t})_{t\in[0,\infty)}) \mid \mathcal{F}_{\tau_n}] = \mathbb{E}_x[F((B_{\tau+t})_{t\in[0,\infty)}) \mid \mathcal{F}_{\tau_n}] \qquad \text{in } \mathcal{L}^1,$$

gleichmäßig in $n \in \mathbb{N}$. Ferner fällt die Folge der σ-Algebren $\mathcal{F}_{\tau_n}$ für $n \to \infty$ gegen die σ-Algebra

$$\mathcal{F}_{\tau+} = \bigcap_{\sigma>\tau \text{ Stoppzeit}} \mathcal{F}_\sigma,$$

die die σ-Algebra $\mathcal{F}_\tau$ übrigens enthält. Nun muss der Konvergenzsatz für sogenannte Rückwärtsmartingale angewendet werden, den wir in Kapitel 1 nicht behandelten. (Es handelt sich hier um eine Version des Doob'schen Konvergenzsatzes 1.4.3.) Dieser liefert

$$\begin{aligned}\mathbb{E}_{B_\tau}[F(B)] &= \lim_{n\to\infty} \mathbb{E}_x\big[F\big((B_{\tau_n+t})_{t\in[0,\infty)}\big) \mid \mathcal{F}_{\tau_n}\big] = \lim_{n\to\infty} \mathbb{E}_x\big[F\big((B_{\tau+t})_{t\in[0,\infty)}\big) \mid \mathcal{F}_{\tau_n}\big]\\ &= \mathbb{E}_x\big[F\big((B_{\tau+t})_{t\in[0,\infty)}\big) \mid \mathcal{F}_{\tau+}\big].\end{aligned}$$

Da die linke Seite dieser Gleichung $\mathcal{F}_\tau$-messbar ist, kann auf der rechten Seite $\mathcal{F}_{\tau+}$ durch $\mathcal{F}_\tau$ ersetzt werden. Dann landen wir bei (3.3.2), und der Beweis ist beendet. □

Eine schöne und nützliche Anwendung gibt uns die Möglichkeit, die Verteilung des Maximums des Pfades bis zu einem festen Zeitpunkt zu identifizieren:

Satz 3.3.8 (Spiegelungsprinzip). *Für jedes $T > 0$ und $a > 0$ gilt*

$$\mathbb{P}\Big(\max_{t\in[0,T]} B_t > a\Big) = 2\mathbb{P}(B_T > a). \tag{3.3.4}$$

Beweis. Wegen der Skalierungseigenschaft (siehe Satz 3.3.1(i)) können wir $T = 1$ annehmen. Sei $\tau = \inf\{t \geq 0 \colon B_t \geq a\} \wedge 1$ die Ersterreichenszeit von a, nach oben abgeschnitten bei 1. Dann ist die linke Seite von (3.3.4) gleich $\mathbb{P}(\tau < 1)$.

Falls $\tau < 1$, also falls B den Punkt a im Zeitintervall $[0, 1)$ erreicht, so kann man den Pfadteil $(B_t)_{t\in[\tau,1]}$ um seinen Startpunkt $B_\tau = a$ herum spiegeln, so dass man einen Pfadteil $(B'_t)_{t\in[\tau,1]}$ erhält, der also zum Zeitpunkt 1 in $B'_1 = a + (a - B_1) = 2a - B_1$ landet. Aufgrund der starken Markoveigenschaft und der Stetigkeit der Pfade ist der gespiegelte Pfad eine in a startende Brown'sche Bewegung, die unabhängig ist von $(B_t)_{t\in[0,\tau]}$. Er endet also mit Wahrscheinlichkeit $\frac{1}{2}$ in (a, ∞), sonst in $(-\infty, a]$. Die Spiegelung ist eine bijektive Abbildung von $\{\tau < 1, B_1 < a\}$ auf $\{B_1 > a\}$. (Die Umkehrabbildung erhält man leicht, indem man am Zeitpunkt τ den späteren Pfadteil zurückspiegelt.) Leicht ist auch zu sehen, dass die Spiegelung die Verteilung erhält, also die Wahrscheinlichkeiten nicht verändert. Daher ist $\mathbb{P}(\tau < 1, B_1 < a) = \mathbb{P}(B_1 > a)$. Wegen des obigen Arguments ist $\mathbb{P}(\tau < 1, B_1 < a) = \mathbb{P}(\tau < 1, B_1 \geq a) = \frac{1}{2}\mathbb{P}(\tau < 1)$. Dies zeigt (3.3.4). □

Bemerkung 3.3.9 (Alternativer Beweis). Das Spiegelungsprinzip kann auch bewiesen werden, indem man die diskrete Variante für die einfache Irrfahrt auf $\mathbb{Z}$ beweist (die Aussage und der Beweis sind analog zum Obigen) und dann einen reskalierten Grenzübergang im Geiste des Donsker'schen Invarianzprinzips (Satz 3.2.10) durchführt. ◇

Bemerkung 3.3.10 (Ersterreichenszeiten). Mit Hilfe der Maxima und Minima kann man auch die Verteilung der *Ersterreichenszeiten*

$$\tau_a = \inf\{t \geq 0 \colon B_t = a\}$$

leicht errechnen, denn für $a > 0$ haben wir $\{\tau_a > t\} = \{\max_{s\in[0,t]} B_s < a\}$. $\diamond$

Dass das Spiegelungsprinzip nützlich ist, liegt auch daran, dass man die Wahrscheinlichkeit auf der rechten Seite von (3.3.4) elementar und klar in beide Richtungen abschätzen kann:

Lemma 3.3.11. *Sei X eine standardnormalverteilte Zufallsgröße. Dann gilt für jedes $a > 0$:*

$$\frac{1}{\sqrt{2\pi}}\,\frac{a}{1+a^2}\mathrm{e}^{-\frac{1}{2}a^2} \leq \mathbb{P}(X \geq a) \leq \frac{1}{\sqrt{2\pi}}\,\frac{1}{a}\mathrm{e}^{-\frac{1}{2}a^2}.$$

Beweis. Der Beweis der ersten Ungleichung ist eine Übungsaufgabe. Die zweite sieht man zum Beispiel so ein:

$$\mathbb{P}(X \geq a) = \frac{1}{\sqrt{2\pi}}\int_a^\infty \mathrm{e}^{-\frac{1}{2}x^2}\,\mathrm{d}x \leq \frac{1}{\sqrt{2\pi}}\,\frac{1}{a}\int_a^\infty x\mathrm{e}^{-\frac{1}{2}x^2}\,\mathrm{d}x = \frac{1}{\sqrt{2\pi}}\,\frac{1}{a}\,\mathrm{e}^{-a^2/2}.$$

$\square$

Als eine Anwendung des Spiegelungsprinzips identifizieren wir die Verteilung des Zeitpunkts des letzten Besuchs in Null vor einem gegebenen Zeitpunkt:

Satz 3.3.12 (Lévy'sches Arkussinus-Gesetz). *Sei $T > 0$ und $\zeta_T = \sup\{t \in [0,T] \colon B_t = 0\}$. Dann gilt*

$$\mathbb{P}(\zeta_T \leq t) = \frac{2}{\pi}\arcsin\sqrt{t/T}, \qquad t \in [0,T].$$

Beweis. Wegen der Skalierungseigenschaft reicht es, den Beweis nur für $T = 1$ zu führen. Wir schreiben ζ statt ζ_1. Mit Hilfe der Markoveigenschaft, der Tatsache, dass B und $-B$ die selbe Verteilung haben, der Stetigkeit der Pfade und des Spiegelungsprinzips erhalten wir für jedes $t \in [0,1]$

$$\begin{aligned}
\mathbb{P}(\zeta \leq t) &= \mathbb{P}\big(B_s \neq 0 \text{ für jedes } s \in [t,1]\big)\\
&= \int_{\mathbb{R}} \mathbb{P}\big(B_s \neq 0 \text{ für jedes } s \in [t,1] \,\big|\, B_t = a\big)\,\mathbb{P}(B_t \in \mathrm{d}a)\\
&= \int_{\mathbb{R}} \mathbb{P}_{|a|}\big(B_s > 0 \text{ für jedes } s \in [0,1-t]\big)\,\mathbb{P}(B_t \in \mathrm{d}a)\\
&= \int_{\mathbb{R}} \mathbb{P}\Big(\sup_{s\in[0,1-t]} B_s < |a|\Big)\,\mathbb{P}(B_t \in \mathrm{d}a)\\
&= \int_{\mathbb{R}} \mathbb{P}(|B_{1-t}| < |a|)\,\mathbb{P}(B_t \in \mathrm{d}a)\\
&= \mathbb{P}\big(|B_{1-t}| < |\widetilde{B}_t|\big),
\end{aligned}$$

wobei $\widetilde{B}$ eine zweite, von B unabhängige Brown'sche Bewegung ist. Wir schreiben die beiden unabhängigen Zufallsgrößen B_{1-t} und $\widetilde{B}_t$ als $\sqrt{1-t}X$ und $\sqrt{t}Y$, wobei X und Y zwei unabhängige standardnormalverteilte Zufallgrößen sind. Dann können wir weitermachen mit

$$\begin{aligned}\mathbb{P}(\zeta \le t) &= \mathbb{P}\big(|\sqrt{1-t}X| < |\sqrt{t}Y|\big) = \mathbb{P}\big(X^2 < t(X^2+Y^2)\big)\\ &= \frac{1}{2\pi}\int_{\mathbb{R}} \mathrm{d}x \int_{\mathbb{R}} \mathrm{d}y\, \mathrm{e}^{-\frac{1}{2}(x^2+y^2)} \mathbb{1}_{\{x^2 \le t(x^2+y^2)\}}.\end{aligned}$$

Durch Polarkoordinatentransformation errechnet sich das letzte Integral als

$$\mathbb{P}(\zeta \le t) = \frac{1}{2\pi}\int_0^\infty \mathrm{d}r\, r\mathrm{e}^{-\frac{1}{2}r^2} \int_0^{2\pi} \mathrm{d}\varphi\, \mathbb{1}_{\{(\sin\varphi)^2 \le t\}} = \frac{2}{\pi}\arcsin\sqrt{t}.$$

□

Bemerkung 3.3.13 (Ein anderes Arkussinus-Gesetz). Die Verteilung eines weiteren Brown'schen Funktionals lässt sich als die Arkussinus-Verteilung identifizieren, und zwar die von

$$\xi = \int_0^1 \mathbb{1}_{\{B_t \in (0,\infty)\}}\, \mathrm{d}t,$$

der Zeit, die die Brown'sche Bewegung im Zeitintervall $[0,1]$ im positiven Bereich verbringt. Man kann zeigen, dass

$$\mathbb{P}(\xi \le t) = \frac{2}{\pi}\arcsin\sqrt{t}, \qquad t \in [0,1].$$

Ein möglicher Beweis geht über das Donsker'sche Invarianzprinzip (Satz 3.2.10) und die Approximation mit der einfachen Irrfahrt $(S_n)_{n\in\mathbb{N}_0}$: Mit Hilfe elementarer kombinatorischer Ideen kann man zeigen, dass für die Zufallsgröße $X_{2n} = \#\{i \in \{0,\dots,n\} \colon S_{2i-1} > 0\}$ gilt:

$$\mathbb{P}\big(X_{2n} = k\big) = u_{2k}u_{2(n-k)}, \qquad k \in \{0,\dots,n\}, n \in \mathbb{N},$$

wobei $u_{2k} = 2^{-2k}\binom{2k}{k}$. Mit Hilfe der Stirling'schen Formel erhält man die Asymptotik $u_{2k} \sim 1/\sqrt{\pi k}$ für $k \to \infty$, also kann man mit ein wenig Arbeit zeigen, dass die Zufallsgröße $\frac{1}{2n}X_{2n}$ in Verteilung gegen die Arkussinus-Verteilung konvergiert.

Die Zufallsgröße $\frac{1}{2n}X_{2n}$ kann als $\Phi(B^{(n)})$ geschrieben werden, wobei die stückweise lineare Funktion $B^{(n)}$ durch $B^{(n)}_{i/n} = S_i/\sqrt{n}$ definiert ist und wir $\Phi(f) = \int_0^1 \mathbb{1}_{\{f(t)\in(0,\infty)\}}\,\mathrm{d}t$ gesetzt haben. Da nach dem Donsker'schen Invarianzprinzip $B^{(n)}$ gegen B in Verteilung konvergiert und da die Menge der Unstetigkeitsstellen von Φ unter dem Wiener-Maß die Masse Null hat (hier muss ein wenig gearbeitet werden), konvergiert auch die Verteilung von $\frac{1}{2n}X_{2n} = \Phi(B^{(n)})$ gegen $\Phi(B) = \xi$. Also hat ξ die Arkussinus-Verteilung. ◇

3.3.5 Das Gesetz vom Iterierten Logarithmus

In Beispiel 3.3.3 haben wir gesehen, dass fast sicher $\limsup_{t\downarrow 0} B_t/\sqrt{t} = \infty$ gilt. Da nach Satz 3.3.1(ii) $(tB_{1/t})_{t\ge 0}$ auch eine Brown'sche Bewegung ist, gilt also auch

$$\limsup_{t\to\infty} \frac{B_t}{\sqrt{t}} = \infty \qquad \text{fast sicher.}$$

(Da B und $-B$ die selbe Verteilung haben, gilt natürlich auch $\liminf_{t\to\infty} B_t/\sqrt{t} = -\infty$ fast sicher.) In diesem Abschnitt wollen wir die Skalierungsfunktion $\sqrt{t}$ durch eine andere ersetzen, so dass der limes superior endlich wird und positiv bleibt.

Satz 3.3.14 (Gesetz vom Iterierten Logarithmus). *Es gilt*

$$\limsup_{t\to\infty} \frac{B_t}{\sqrt{2t\log\log t}} = 1 \qquad \textit{fast sicher.} \tag{3.3.5}$$

Beweis. '$\leq$': Es reicht, für jedes $\alpha > 1$ zu zeigen, dass $\limsup_{t\to\infty} B_t/\sqrt{2t\log\log t} \leq \alpha$ fast sicher gilt. Sei also $\alpha > 1$, und wir setzen $f(t) = 2\alpha^2 \log\log t$. Wir betrachten die Folge $t_n = \alpha^n$. Nach dem Spiegelungsprinzip (Satz 3.3.8), zusammen mit Lemma 3.3.11, gilt

$$\begin{aligned}\mathbb{P}\Big(\sup_{s\in[t_n,t_{n+1}]} B_s > \sqrt{t_n f(t_n)}\Big) &\leq \mathbb{P}\Big(t_{n+1}^{-1/2}\sup_{s\in[0,t_{n+1}]} B_s > \sqrt{f(t_n)/\alpha}\Big) \leq \sqrt{\frac{\alpha}{f(t_{n+1})}}\,\mathrm{e}^{-\frac12 f(t_n)/\alpha}\\ &= (\log\alpha)^{-\alpha}\sqrt{\frac{\alpha}{f(t_{n+1})}}\,n^{-\alpha}.\end{aligned} \tag{3.3.6}$$

Man beachte, dass

$$\frac{f(t_n)}{2\alpha} = \alpha\log(n\log\alpha) = \alpha\log n + \alpha\log\log\alpha \to \infty \qquad \text{für } n\to\infty.$$

Also ist die rechte Seite von (3.3.6) summierbar über $n\in\mathbb{N}$. Daher liefert das Erste Borel-Cantelli-Lemma, dass das Ereignis $\{\sup_{s\in[t_n,t_{n+1}]} B_s > \sqrt{t_n f(t_n)}\}$ fast sicher nur für endlich viele n eintritt. Also gilt die umgekehrte Ungleichung fast sicher für fast alle $n\in\mathbb{N}$. Da die Abbildung $t\mapsto\sqrt{tf(t)}$ steigend ist, folgt fast sicher für alle genügend großen $t\in(0,\infty)$, dass

$$B_t \leq \sup_{s\in[t_n,t_{n+1}]} B_s \leq \sqrt{t_n f(t_n)} \leq \sqrt{tf(t)}$$

gilt, wobei $n\in\mathbb{N}$ so gewählt wurde, dass $t\in[t_n,t_{n+1}]$. Dies zeigt, dass

$$1 \geq \limsup_{t\to\infty} B_t/\sqrt{tf(t)} = \limsup_{t\to\infty} B_t/\sqrt{2\alpha^2 t\log\log t},$$

und der Beweis von '$\leq$' ist beendet.

'$\geq$': Wieder fixieren wir ein $\alpha > 1$ und betrachten die Folge der $t_n = \alpha^n$. Diesmal wird α sehr groß gewählt werden. Setze $\beta = \alpha/(\alpha-1) \in (1,\infty)$ und $g(t) = \frac{2}{\beta^2}\log\log t$. Wähle n_0 so groß, dass $\beta g(t_n) > 1$ für alle $n \geq n_0$ gilt. Man beachte, dass $t_n - t_{n-1} = t_n/\beta$. Deshalb erhalten wir aus der Brown'schen Skalierungseigenschaft und Lemma 3.3.11, dass

$$\begin{aligned}\mathbb{P}\Big(B_{t_n} - B_{t_{n-1}} > \sqrt{t_n g(t_n)}\Big) &= \mathbb{P}\big(B_1 > \sqrt{\beta g(t_n)}\big) \geq \frac{1}{\sqrt{2\pi}}\,\frac12\,\frac{1}{\sqrt{\beta g(t_n)}}\mathrm{e}^{-\frac12\beta g(t_n)}\\ &= \frac{1}{\sqrt{2\pi}}\,\frac12\,(\log\alpha)^{-1/\beta}\,\frac{1}{\sqrt{\beta g(t_n)}}\,n^{-1/\beta}.\end{aligned}$$

Also ist die linke Seite nicht summierbar über $n\in\mathbb{N}$. Die betrachteten Ereignisse sind unabhängig, also liefert das Zweite Borel-Cantelli-Lemma, dass die Ungleichung fast sicher für unendlich viele $n\in\mathbb{N}$ eintritt, d. h. es gilt fast sicher

$$B_{t_n} - B_{t_{n-1}} > \sqrt{t_n g(t_n)} \qquad \text{für unendlich viele } n\in\mathbb{N}. \tag{3.3.7}$$

Man beachte, dass

$$\lim_{n\to\infty} \frac{t_n g(t_n)}{t_{n-1} g(t_{n-1})} = \lim_{n\to\infty} \frac{t_n \log\log t_n}{t_{n-1}\log\log t_{n-1}} = \alpha.$$

Außdem wissen wir, da ja in (3.3.5) die Ungleichung '$\leq$' schon bewiesen worden ist, dass für jedes $\varepsilon > 0$ fast sicher gilt (man gehe von B zu $-B$ über):

$$B_{t_{n-1}} > -(1+\varepsilon)\alpha^{-1/2}\sqrt{2t_n \log\log t_n}, \qquad \text{für fast jedes } n \in \mathbb{N}. \tag{3.3.8}$$

Wenn wir (3.3.7) und (3.3.8) kombinieren, erhalten wir

$$\limsup_{n\to\infty} \frac{B_{t_n}}{\sqrt{2t_n \log\log t_n}} \geq \frac{1}{\beta} - (1+\varepsilon)\alpha^{-1/2} = \frac{\alpha - 1}{\alpha} - (1+\varepsilon)\alpha^{-1/2}.$$

Nun lassen wir $\alpha \to \infty$ gehen und sehen, dass die rechte Seite gegen Eins konvergiert. Das beendet den Beweis. □

Mit Hilfe der Stürzungseigenschaft (Satz 3.3.1(ii)) erhalten wir leicht ein Gesetz des Iterierten Logarithmus für kleine Zeiten:

Korollar 3.3.15. *Es gilt fast sicher*

$$\limsup_{t\downarrow 0} \frac{B_t}{\sqrt{2t\log\log(1/t)}} = 1.$$

3.3.6 Der Lévy'sche Stetigkeitsmodul

Aus dem Gesetz des Iterierten Logarithmus folgt, dass der Pfad mit Wahrscheinlichkeit Eins in Null nicht Hölder-stetig mit Parameter $\gamma = \frac{1}{2}$ ist. Nun wollen wir die schwierigere Frage stellen, was die Güte der gleichmäßigen Stetigkeit der Pfade mit Wahrscheinlichkeit Eins auf einem ganzen Zeitintervall ist, etwa auf $[0,1]$. Mit anderen Worten, wir wollen die fast sichere Asymptotik von $w_\delta(B)$ für $\delta \downarrow 0$ bestimmen, wobei $w_\delta(f)$ in (3.2.5) eingeführt wurde. Es stellt sich heraus, dass $w_\delta(B)$ wiederum von der Größenordnung $\sqrt{2\delta}$ ist, diesmal aber mit einer anderen logarithmischen Korrektur als im Gesetz des Iterierten Logarithmus.

Satz 3.3.16 (Lévy'scher Stetigkeitsmodul). *Mit Wahrscheinlichkeit Eins gilt*

$$\limsup_{\delta\downarrow 0} \frac{\max_{t\in[0,1-\delta]} |B_{t+\delta} - B_t|}{\sqrt{2\delta\log(1/\delta)}} = 1. \tag{3.3.9}$$

Insbesondere ist der Pfad $(B_t)_{t\in[0,1]}$ fast sicher nirgends Hölder-stetig mit Parameter $\gamma = \frac{1}{2}$, aber er ist fast sicher Hölder-stetig mit jedem Parameter $\gamma \in (0, \frac{1}{2})$.

Beweisskizze. Zunächst zeigen wir, dass die linke Seite von (3.3.9) nicht kleiner als Eins ist. (Dies ist die leichtere Aufgabe, da wir nur eine geeignete Teilfolge $\delta_n \downarrow 0$ konstruieren müssen.) Wir kürzen ab: $g(\delta) = \sqrt{2\delta\log(1/\delta)}$ und setzen $\delta_n = 2^{-n}$. Sei $\varepsilon > 0$ klein, dann reicht es zu zeigen, dass der Limes Superior der linken Seite von (3.3.9) mit $\delta = \delta_n = 2^{-n}$ nicht kleiner als $1 - \varepsilon$ ist. Zunächst sieht man, dass

$$\begin{aligned} A_n &:= \Big\{ \max_{t\in[0,1-2^{-n}]} |B_{t+2^{-n}} - B_t| < (1-\varepsilon) g(2^{-n}) \Big\} \\ &\subset \Big\{ \max_{k\in\mathbb{N}_0\colon k<2^n} \big(B_{(k+1)2^{-n}} - B_{k2^{-n}}\big) \leq (1-\varepsilon) g(2^{-n}) \Big\}. \end{aligned}$$

Mit Hilfe der Unabhängigkeit der Zuwächse, der Tatsache, dass $2^{n/2}B_{2^{-n}} \sim \mathcal{N}$, der Abschätzung $1-x \leq \mathrm{e}^{-x}$ und von Lemma 3.3.11 können wir die Wahrscheinlichkeit von A_n abschätzen:

$$\begin{aligned}
\mathbb{P}(A_n) &\leq \Big(1-\mathbb{P}\big(B_{2^{-n}} \geq (1-\varepsilon)g(2^{-n})\big)\Big)^{2^n}\\
&\leq \exp\Big(-2^n\mathcal{N}\big([(1-\varepsilon)\sqrt{2n\log 2},\infty)\big)\Big)\\
&\leq \exp\Big(-\frac{2^n}{\sqrt{2\pi}}\frac{(1-\varepsilon)\sqrt{2n\log 2}}{1+(1-\varepsilon)^2 2n\log 2}\,\mathrm{e}^{-(1-\varepsilon)^2 n\log 2}\Big)\\
&\leq \exp\Big(-\frac{C}{\sqrt{n}}2^{(2\varepsilon-\varepsilon^2)n}\Big),
\end{aligned}$$

wobei $C>0$ nur von ε abhängt. Da die rechte Seite für genügend kleines $\varepsilon>0$ summierbar über n ist, folgt aus dem Ersten Borel-Cantelli-Lemma, dass A_n fast sicher nur für endlich viele n eintritt, d. h. $\mathbb{P}(\limsup_{n\to\infty} A_n)=0$. Dies zeigt, dass der Limes Superior der linken Seite von (3.3.9) mit $\delta=\delta_n=2^{-n}$ nicht kleiner als $1-\varepsilon$ ist, und beendet den Beweis von '$\geq$' in (3.3.9).

Wir bringen hier nicht den vollen Beweis von '$\leq$' in (3.3.9), sondern beweisen nur, dass für jedes $\varepsilon>0$ gilt:

$$\max_{k\in\mathbb{N}\colon k<2^{n\delta}}\max_{i=0}^{2^n-k}\frac{\big|B_{(i+k)2^{-n}}-B_{i2^{-n}}\big|}{g(k2^{-n})}\leq 1+\varepsilon, \qquad \text{für fast jedes } n\in\mathbb{N}, \tag{3.3.10}$$

wobei $\delta>0$ mit $\varepsilon>\frac{1+\delta}{1-\delta}-1$ gewählt wurde. Mit anderen Worten, wir beweisen die Stetigkeitsaussage in (3.3.9) nur für binär-rationale t und δ. Mit einer geeigneten Approximationstechnik (die keine wahrscheinlichkeitstheoretischen Überlegungen benötigt) verstärkt man (3.3.10) zu (3.3.9), worum wir uns hier aber nicht kümmern wollen.

Nach dem Ersten Borel-Cantelli-Lemma ist nur die Summierbarkeit der Wahrscheinlichkeit für das Komplement des Ereignisses in (3.3.10) über $n\in\mathbb{N}$ zu zeigen. Dies machen wir mit ähnlichen Hilfsmitteln wie im ersten Teil des Beweises wie folgt. Für jedes $n\in\mathbb{N}$ gilt

$$\begin{aligned}
&\mathbb{P}\Big(\max_{k\in\mathbb{N}\colon k<2^{n\delta}}\max_{i=0}^{2^n-k}\frac{\big|B_{(i+k)2^{-n}}-B_{i2^{-n}}\big|}{g(k2^{-n})}\geq 1+\varepsilon\Big)\\
&\quad\leq \sum_{k=1}^{2^{\lfloor n\delta\rfloor}}\sum_{i=0}^{2^n-k}\frac{2}{\sqrt{2\pi}}\int_{(1+\varepsilon)\sqrt{2\log(2^n/k)}}^{\infty}\mathrm{e}^{-\frac12 x^2}\,\mathrm{d}x\\
&\quad\leq \sum_{k=1}^{2^{\lfloor n\delta\rfloor}}2^n\frac{2}{\sqrt{2\pi}}\,\frac{1}{(1+\varepsilon)\sqrt{2\log(2^n/k)}}\mathrm{e}^{-(1+\varepsilon)^2\log(2^n/k)}\,\mathrm{d}x\\
&\quad\leq \frac{C}{\sqrt{n}}2^n\,2^{n\delta}\,2^{-n(1-\delta)(1+\varepsilon)^2}
\end{aligned}\tag{3.3.11}$$

für ein geeignetes, nur von ε und δ abhängiges $C>0$. (Die letzte Abschätzung erhält man, indem man k durch $2^{n\delta}$ abschätzt und zusammenfasst.) Nun beachte man, dass wegen $\varepsilon>\frac{1+\delta}{1-\delta}-1$ gilt:

$$(1-\delta)(1+\varepsilon)^2>(1-\delta)\frac{(1+\delta)^2}{(1-\delta)^2}>(1+\delta)^2>1+\delta.$$

Also ist die rechte Seite von (3.3.11) summierbar, und der Beweis von (3.3.10) ist beendet. □

Bibliographie

[Ba91] H. Bauer, *Maß- und Integrationstheorie,* 2. Auflage, Walter de Gruyter, Berlin–New York, 1991.

[Co80] D.L. Cohn, *Measure Theory,* Birkhäuser, Basel 1980.

[Du84] R. Durrett, *Brownian Motion and Martingales in Analysis*, Wadsworth, Belmont (1984).

[Fe68] W. Feller, *An Introduction to Probability Theory and its Applications*, Vol. I und II, 3rd ed., Wiley, New York, 1968.

[Ge02] H.-O. Georgii, *Stochastik. Einführung in die Wahrscheinlichkeitstheorie und Statistik*, Walter de Gruyter, 2002.

[GS75] I. I. Gikhman und A. V. Skorohod, *The Theory of Stochastic Processes II*, Springer, Berlin (1975).

[KS91] I. Karatzas und S.E. Shreve, *Brownian Motion and Stochastic Calculus*, 2nd edition. Springer, New York (1991).

[KT75] S. Karlin und H. M. Taylor, *A First Course in Stochastic Processes*, Academic Press, New York, 1975.

[KT81] S. Karlin und H. M. Taylor, *A Second Course in Stochastic Processes*, Academic Press, New York, 1975.

[Kl06] A. Klenke, *Einführung in die Wahrscheinlichkeitstheorie*, Springer, Berlin (2006).

[Pa84] L. Partzsch, *Vorlesungen zum eindimensionalen Wienerschen Prozeß*, Teubner-Texte zur Mathematik, Leipzig (1984).

[RY99] D. Revuz und M. Yor, *Continuous Martingales and Brownian Motion*, (3rd edition), Springer, Berlin (1999).

[Sch98] K. Schürger, *Wahrscheinlichkeitstheorie*, Oldenbourg, München, 1998.

[Wi91] D. Williams, *Probability with Martingales*, Cambridge (1991).

Das Lehrbuch [Ge02] ist hauptsächlich der Elementaren Wahrscheinlichkeitstheorie gewidmet und dient dem Nacharbeiten von wichtigen Beispielen und Anwendungen der bekannten diskreten und absolutstetigen Verteilungen. Die Lehrbücher [Fe68], [Sch98] und [Kl06] dienen dem Nachschlagen benötigter Hilfsmittel der Wahrscheinlichkeitstheorie, und [Co80] und [Ba91] informieren über die benötigte Maßtheorie.

Die jeweiligen Kapitel in [Fe68] und [Kl06] über Martingale sind gute Einführungen in das Thema. Eine ausführliche Behandlung ist [Wi91]. Gut lesbare Einführungen in die Brown'sche Bewegung sind das (nicht mehr erhältliche) Büchlein [Pa84] und das entsprechende Kapitel aus [Kl06]. Eine enorme Fülle von Material über die Brown'sche Bewegung sowie über Martingale in stetiger Zeit (mit knappen Beweisen) findet man in [KS91], [Du84], [Wi91] und [RY99]. Die allgemeine Theorie der Stochastischen Prozesse wird in [GS75] wiedergegeben.

Index