Mathematical Finance and Stochastic Integration

Mathias Vetter

July 5, 2022

Contents

Chapter 1

Brownian motion

In this chapter we discuss Brownian motion, the prime example for a stochastic process in continuous time. One of its defining properties is that its future development is independent of its current state which makes it extremely useful for the modelling of a lot of time-dependent phenomena both in economics and in the life sciences.

Definition 1.1. Let T be an arbitrary index set. A family of (real-valued) random variables

$$X = (X_t)_{t \in T} = \{X(t) \mid t \in T\}$$

on $(\Omega, \mathcal{F}, \mathbb{P})$ is called a *stochastic process.*

Remark 1.2.

(i) Typical index sets for stochastic processes are $T = \mathbb{N}_0$ (discrete time) or $T = [0, \infty)$ (continuous time).

(ii) In analogy to standard random variables

$$\omega \mapsto X(\omega),$$

where each ω maps to a (real-valued) number, we have that

$$\omega \mapsto (X_t(\omega))_{t \in T}$$

maps ω to a function from T to $\mathbb{R}$. This explains the notion of a stochastic process.

Definition 1.3. A real-valued process $B = (B_t)_{t \geq 0}$ is called a *Brownian motion* with start in x if:

(i) $B_0 = x$ almost surely;

(ii) the process has independent increments, i.e. for all $0 = t_0 \leq t_1 \leq \ldots \leq t_n$ we have that the random variables $B_{t_1} - B_{t_0}, \ldots, B_{t_n} - B_{t_{n-1}}$ are independent;

(iii) for any $t \geq 0$ and any $h > 0$ we have that the increments $B_{t+h} - B_t$ are normal distributed with expectation 0 and variance h;

(iv) the process $t \mapsto B_t(\omega)$ is almost surely continuous, i.e.

$$\mathbb{P}(\{\omega \mid t \mapsto B_t(\omega) \text{ is continuous}\}) = 1.$$

Theorem 1.4. (Wiener's theorem) *Brownian motion exists.*

Proof: cf. for example Theorem 1.3 in Mörters and Peres (2010). $\square$

Remark 1.5.

(i) Unless stated otherwise we will always assume that Brownian motion starts in $x = 0$.

(ii) The definition of a d-dimensional Brownian motion B works analogously, with the components $B^{(i)}$, $i = 1, \dots, d$, being independent one-dimensional Brownian motions.

(iii) Brownian motion is a subclass of the bigger class of *Lévy processes* which are defined similarly to Definition 1.3, but without normality of increments and with a little more generality than just continuous paths.

Definition 1.6. Let T be an ordered index set and X a stochastic process. Then the distributions of

$$(X_{t_1}, \dots, X_{t_n})^T \text{ for all } t_1 \le t_2 \le \dots \le t_n \text{ with } t_1, t_2, \dots, t_n \in T, \text{ any } n \ge 1,$$

are called the *finite-dimensional distributions* of X.

Example 1.7. Let B be a Brownian motion and let $U \sim \mathcal{U}[0, 1]$ be an independent uniformly distributed random variable. Then $(B^*_t)_{t \ge 0}$ with

$$B^*_t = \begin{cases} B_t, & t \ne U, \\ 0, & t = U, \end{cases}$$

has the same finite-dimensional distributions as a Brownian motion, but is, because of $B_t \sim \mathcal{N}(0, t)$ for all t, almost surely not continuous.

Remark 1.8. While Definition 1.3(i)–(iii) determines the finite-dimensional distributions of a Brownian motion, Definition 1.3(iv) is an additional condition on the structure of almost all *paths* $t \mapsto B_t(\omega)$, which according to Example 1.7 is not implied by the finite-dimensional distributions alone. This phenomenon is typical for stochastic processes in continuous time.

Definition 1.9. A stochastic process X is called a *Gaussian process* if its finite-dimensional distributions are all normal distributions.

Theorem 1.10. *A Brownian motion with start in x is a Gaussian process.*

Proof: For any $0 \le t_1 \le t_2 \le \dots \le t_n$ we have

$$\begin{pmatrix} B_{t_1} \\ \vdots \\ B_{t_n} \end{pmatrix} = \begin{pmatrix} B_0 \\ \vdots \\ B_0 \end{pmatrix} + \begin{pmatrix} 1 & 0 & \dots & 0 \\ 1 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 1 & \dots & 1 \end{pmatrix} \begin{pmatrix} B_{t_1} - B_0 \\ \vdots \\ B_{t_n} - B_{t_{n-1}} \end{pmatrix},$$

and affine linear combinations of normal distributed vectors are again normal distributed. $\square$

Definition 1.11. Two stochastic processes $X = (X_t)_{t\in I}$ and $Y = (Y_s)_{s\in J}$ are called *independent* if for all $t_1, t_2, \ldots, t_n \in I$ and all $s_1, s_2 \ldots, s_m \in J$ and all n and m the vectors $(X_{t_1}, \ldots, X_{t_n})^T$ and $(Y_{s_1}, \ldots, Y_{s_m})^T$ are independent.

Theorem 1.12. (Markov property) *Let B be a Brownian motion with start in x and $t \geq 0$. Then the process $(Z_s)_{s\geq 0}$ with $Z_s = B_{t+s} - B_t$ is a Brownian motion with start in 0 and independent of $(B_s)_{s\leq t}$.*

Proof: Obviously, $Z_{s_2} - Z_{s_1} = B_{t+s_2} - B_{t+s_1}$, so that increments of $(Z_s)_{s\geq 0}$ equal the increments of a classical Brownian motion. In particular, we have (ii) and (iii) of Definition 1.3 directly. Start in 0 and almost sure continuity of the paths is immediate.

That both processes are independent follows with a similar argument as in the proof of Theorem 1.10: Any two vectors $(B_{s_1}, \ldots, B_{s_m})^T$, $s_1, s_2 \ldots, s_m \leq t$, and $(B_{t_1} - B_t, \ldots, B_{t_n} - B_t)^T$, $t_1, t_2, \ldots, t_n \geq 0$, can be written as linear combinations of increments until t and after t, respectively, which according to Definition 1.3 are independent from each other. □

Definition 1.13. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space.

(i) A family of σ-fields $(\mathcal{F}_t)_{t\geq 0}$ with

$$\mathcal{F}_s \subset \mathcal{F}_t \subset \mathcal{F} \quad \text{for all } s \leq t$$

is called a *filtration*. $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ is then called a *filtered probability space*.

(ii) A stochastic process $X = (X_t)_{t\geq 0}$ is called *adapted* to a filtration $(\mathcal{F}_t)_{t\geq 0}$ if X_t is $\mathcal{F}_t$-measurable for every $t \geq 0$.

Remark 1.14.

(i) For every stochastic process $X = (X_t)_{t\geq 0}$ on $(\Omega, \mathcal{F}, \mathbb{P})$ a filtration can be defined via $\mathcal{F}_t = \sigma(\{X_s \mid s \leq t\})$ to which X is naturally adapted. It is called the *filtration generated by* X or *the natural filtration*. Intuitively it contains all information about X until time t.

(ii) The definition of filtrations and generated filtrations works analogously in discrete time.

Definition 1.15. For a given filtration $(\mathcal{F}_t)_{t\geq 0}$ one can define another filtration $(\mathcal{F}_t^+)_{t\geq 0}$ via

$$\mathcal{F}_t^+ = \bigcap_{u>t} \mathcal{F}_u,$$

which compared to $(\mathcal{F}_t)_{t\geq 0}$ allows for an infinitesimal look into the future. If $\mathcal{F}_t = \mathcal{F}_t^+$ holds for all $t \geq 0$, then $(\mathcal{F}_t)_{t\geq 0}$ is called *right continuous*.

Lemma 1.16. *Let $(\mathcal{F}_t)_{t\geq 0}$ be a filtration and $(\mathcal{F}_t^+)_{t\geq 0}$ its* infinitesimal extension*. Then $(\mathcal{F}_t^+)_{t\geq 0}$ is right continuous.*

Proof: Set $\mathcal{G}_t = \mathcal{F}_t^+$. From

$$\mathcal{G}_t^+ = \bigcap_{r>t} \mathcal{G}_r = \bigcap_{r>t} \bigcap_{u>r} \mathcal{F}_u = \bigcap_{u>t} \mathcal{F}_u = \mathcal{F}_t^+ = \mathcal{G}_t$$

the claim follows. $\square$

Theorem 1.17. *Let $B = (B_t)_{t\geq 0}$ be a Brownian motion with start in x, generated filtration $(\mathcal{F}_t)_{t\geq 0}$ and accompanying infinitesimal extension $(\mathcal{F}_t^+)_{t\geq 0}$. Then, for every $t \geq 0$, we have that the process $(Z_s)_{s\geq 0}$ given by $Z_s = B_{t+s} - B_t$ is independent of $\mathcal{F}_t^+$.*

Proof: Let us first assume that we are given a sequence $(X_n)_n$ of random variables and a σ-field $\mathcal{A}$ such that X_n and $\mathcal{A}$ are independent for every n. We first prove that if $(X_n)_n$ converges to X almost surely, then X and $\mathcal{A}$ are independent as well. To this end let Ω_0 be a set with $\mathbb{P}(\Omega_0) = 1$ such that $X_n(\omega) \to X(\omega)$ for all $\omega \in \Omega_0$. If one now sets

$$\widetilde{X}_n(\omega) = \begin{cases} X_n(\omega), & \omega \in \Omega_0, \\ 0, & \omega \notin \Omega_0, \end{cases} \quad \text{and} \quad \widetilde{X}(\omega) = \begin{cases} X(\omega), & \omega \in \Omega_0, \\ 0, & \omega \notin \Omega_0, \end{cases}$$

then, because of $\mathbb{P}(\Omega_0) = 1$, we have that any $\widetilde{X}_n$ is independent from $\mathcal{A}$ as well. But now $\widetilde{X}_n$ converges pointwise to $\widetilde{X}$, so $\widetilde{X} \subset \sigma(\widetilde{X}_n \mid n \in \mathbb{N})$, and therefore $\widetilde{X}$ is independent from $\mathcal{A}$. Because of $X = \widetilde{X}$ $\mathbb{P}$-almost surely we finally obtain

$$\mathbb{P}(\{X \in B\} \cap A) = \mathbb{P}(\{\widetilde{X} \in B\} \cap A) = \mathbb{P}(\{\widetilde{X} \in B\})\mathbb{P}(A) = \mathbb{P}(\{X \in B\})\mathbb{P}(A)$$

for all $B \in \mathcal{B}$ and all A in $\mathcal{A}$.

In our situation we apply the previous result to the almost surely continuous Brownian motion, for which

$$B_{t+s} - B_t = \lim_{m\to\infty} (B_{t_m+s} - B_{t_m}) \text{ a.s.}$$

with a strictly decreasing sequence $t_m \to t$. Theorem 1.12 proves that $B_{t_m+s} - B_{t_m}$ is independent from $\mathcal{F}_{t_m}$ for every m, and because of $\mathcal{F}_t^+ \subset \mathcal{F}_{t_m}$ from $\mathcal{F}_t^+$ as well. This property then transfers to $B_{t+s} - B_t$, and the general proof for all finite-dimensional distributions works similarly. $\square$

Theorem 1.18. (Blumenthal's 0-1 law) *Let $B = (B_t)_{t\geq 0}$ be a Brownian motion with start in x, generated filtration $(\mathcal{F}_t)_{t\geq 0}$ and accompanying infinitesimal extension $(\mathcal{F}_t^+)_{t\geq 0}$. Then*

$$\mathbb{P}(A) \in \{0, 1\}$$

for every $A \in \mathcal{F}_0^+$.

Proof: Theorem 1.17 and $B_0 = x$ almost surely imply that any

$$A \in \sigma(\{B_t \mid t \geq 0\})$$

is independent from $\mathcal{F}_0^+$. This property applies in particular to $A \in \mathcal{F}_0^+ \subset \sigma(\{B_t \mid t \geq 0)$. Thus, A is independent of itself, and $\mathbb{P}(A) = \mathbb{P}(A \cap A) = \mathbb{P}(A)^2$ gives the claim. $\square$

Theorem 1.19. *Let $B = (B_t)_{t\geq 0}$ be a Brownian motion and let $\sigma = \inf\{t > 0 \mid B_t > 0\}$ and $\tau = \inf\{t > 0 \mid B_t = 0\}$. Then*

$$\mathbb{P}(\sigma = 0) = \mathbb{P}(\tau = 0) = 1.$$

Proof: The event

$$\{\sigma = 0\} = \bigcap_{n \geq 1} \{\text{there exists some } 0 < \varepsilon < 1/n \text{ with } B_\varepsilon > 0\}$$

obviously is in $\mathcal{F}_0^+$. We also have

$$\mathbb{P}(\sigma \leq t) \geq \mathbb{P}(B_t > 0) = \frac{1}{2}$$

for all $t > 0$, so that continuity from above implies

$$\mathbb{P}(\sigma = 0) = \lim_{t \to 0} \mathbb{P}(\sigma \leq t) \geq \frac{1}{2}.$$

Theorem 1.18 then gives $\mathbb{P}(\sigma = 0) = 1$. The same argument works for $B_t < 0$, and almost sure continuity of the paths of B finally gives the result for τ. □

Theorem 1.20. *Let $(B_t)_{t \in [0,1]}$ be a Brownian motion with start in x. The following claims hold almost surely:*

(i) Every local maximum of B is a strict local maximum.

(ii) The set of all local maxima of B is countable and dense.

(iii) The global maximum of B is unique.

Proof: We first show that the maxima m_1 and m_2 over two "disjoint" intervals $[a_1, b_1]$ and $[a_2, b_2]$ with $a_1 < b_1 \leq a_2 < b_2$ are almost surely different: Using Theorem 1.12 and Theorem 1.19 we first have $B_{a_2} < m_2$ almost surely, so that for almost all ω we can find some $n \geq 1$ such that m_2 equals the maximum over $[a_2 + 1/n, b_2]$. Thus, we can assume $b_1 < a_2$ without loss of generality. Applying Theorem 1.12 twice now proves that $B_{b_1} - m_1$, $B_{a_2} - B_{b_1}$ and $m_2 - B_{a_2}$ are all independent. If now $m_1 = m_2$, then

$$B_{a_2} - B_{b_1} = (m_1 - B_{b_1}) - (m_2 - B_{a_2}).$$

Conditional on $m_1 - B_{b_1}$ and $m_2 - B_{a_2}$ the left hand side has a continuous distribution while the right hand side is constant. This contradiction gives $\mathbb{P}(m_1 = m_2) = 0$.

(i) From the claim above, if follows (almost surely) that any two disjoint non-empty compact intervals with endpoints in $\mathbb{Q}$ have different maxima. But if a Brownian motion on $[0, 1]$ has a non-strict local maximum, then there exist two disjoint non-empty compact intervals with endpoints in $\mathbb{Q}$ having the same maxima. Thus, this only happens with probability zero.

(ii) We have also seen above that almost surely no maximum over $[a, b]$ with rational $a < b$ lies at one of the endpoints. Thus, any of these intervals contains a true local maximum, and its set then lies dense. Also, as every local maximum is strict, it has to be unique in a neighbourhood. Thus the number of intervals $[a, b]$ as above is an upper bound for the number of local maxima.

(iii) The maxima over $[0, q]$ and $[q, 1]$ are different for every rational q, almost surely. If there exist t_1 and t_2 with $B_{t_1} = B_{t_2}$ being the same global maximum, then there exists a $q \in \mathbb{Q}$ with $t_1 < q < t_2$ such that the maxima over $[0, q]$ and $[q, 1]$ coincide. □

Remark 1.21. Both random variables σ and τ from Theorem 1.19 are examples for stopping times which are used to model random times whose exact value depends only on the filtration $(\mathcal{F}_t)_{t\geq 0}$ (in this case on the filtration generated by B) and where it is known by time t whether $\{\tau \leq t\}$ has occured yet or not.

Definition 1.22. Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ be a filtered probability space. A random variable $\tau : \Omega \to [0, \infty]$ is called a *stopping time* (with respect to the filtration $(\mathcal{F}_t)_{t\geq 0}$) if

$$\{\tau \leq t\} \in \mathcal{F}_t$$

for all $t \geq 0$.

Remark 1.23.

(i) Deterministic times are stopping times. Also, if σ and τ are stopping times, so is $\tau \wedge \sigma := \min(\tau, \sigma)$ and $\tau \vee \sigma := \max(\tau, \sigma)$.

(ii) If $(\tau_n)_{n\geq 1}$ is an increasing sequence of stopping times with respect to $(\mathcal{F}_t)_{t\geq 0}$, satisfying $\tau_n \to \tau$ almost surely, then τ is, because of

$$\{\tau \leq t\} = \bigcap_{n\geq 1} \{\tau_n \leq t\} \in \mathcal{F}_t,$$

a stopping time as well.

Example 1.24. The classical examples for stopping times are entry times. For example, one can show for a Brownian motion B and a closed set $H \subset \mathbb{R}$ that $\tau = \inf\{t \geq 0 \mid B_t \in H\}$ is a stopping time with respect to the filtration $(\mathcal{F}_t)_{t\geq 0}$ generated by B.

We will only sketch the proof: By closedness of H we have that $\tau_n \leq t$ is equivalent to the existence of some $r \leq t$ with $B_r \in H$. Continuity of B proves that the latter is equivalent to finding a sequence $s_n \leq t$ of rational numbers such that $B_{s_n} \to y \in H$. One can then conclude

$$\{\tau \leq t\} = \bigcap_{n\geq 1} \bigcup_{s\in\mathbb{Q}\cap(0,t)} \bigcup_{x\in\mathbb{Q}\cap H} \{B_s \in \mathcal{B}(x, 1/n)\} \in \mathcal{F}_t,$$

where $\mathcal{B}(y, \varepsilon)$ denotes the open ball with radius ε around y.

Remark 1.25. Let $\Lambda \subset \mathbb{R}$ be a Borel set and X be an arbitrary stochastic process. In general, one cannot prove that $\tau = \inf\{t \geq 0 \mid X_t \in \Lambda\}$ is a stopping time with respect to the filtration $(\mathcal{F}_t)_{t\geq 0}$ generated by X, so additional structural conditions need to be imposed:

(i) The paths of X are almost surely *càdlàg* (right continuous with left limits);

(ii) The stopping time is defined with respect to the right continuous filtration $(\mathcal{F}_t^+)_{t\geq 0}$.

Unless stated otherwise, we will assume from now on that filtrations are right continuous. For technical reasons we also often assume *completeness of the filtration*, i.e. that $\mathcal{F}_0$ contains all $\mathbb{P}$-zero sets.

Theorem 1.26. *Let $(\mathcal{F}_t)_{t\geq 0}$ be a right continuous filtration. A random variable $\tau : \Omega \to [0,\infty]$ is a stopping time with respect to $(\mathcal{F}_t)_{t\geq 0}$ if and only if*

$$\{\tau < t\} \in \mathcal{F}_t$$

for all $t \geq 0$.

Proof:

$\Longrightarrow$ It holds

$$\{\tau < t\} = \bigcup_{0<\varepsilon<t, \varepsilon\in\mathbb{Q}} \{\tau \leq t-\varepsilon\},$$

and for every stopping time $\{\tau \leq t-\varepsilon\} \in \mathcal{F}_{t-\varepsilon} \subset \mathcal{F}_t$. Thus $\{\tau < t\} \in \mathcal{F}_t$.

$\Longleftarrow$ On the other hand,

$$\{\tau \leq t\} = \bigcap_{t<u<t+\varepsilon, u\in\mathbb{Q}} \{\tau < u\}$$

for all $\varepsilon > 0$. Therefore

$$\{\tau \leq t\} \in \bigcap_{t<u} \mathcal{F}_u = \mathcal{F}_t$$

where we have used right continuity of the filtration. □

Definition 1.27. Let τ be a stopping time with respect to a right continuous filtration $(\mathcal{F}_t)_{t\geq 0}$. Then

$$\mathcal{F}_\tau = \{A \in \mathcal{F} \mid A \cap \{\tau \leq t\} \in \mathcal{F}_t \text{ for all } t \geq 0\}$$

is called the σ-field generated by τ.

Lemma 1.28. *Let $(\mathcal{F}_t)_{t\geq 0}$ be a (right continuous) filtration and let σ and τ be stopping times. Then:*

(i) If $\tau \leq \sigma$ almost surely then $\mathcal{F}_\tau \subset \mathcal{F}_\sigma$.

(ii) τ is $\mathcal{F}_\tau$-measurable.

(iii) If $X = (X_t)_{t\geq 0}$ is càdlàg and adapted, then $X_\tau 1_{\{\tau<\infty\}}$ is $\mathcal{F}_\tau$-measurable.

Proof: See e.g. *Probability and Statistics II* for the proof in discrete time. Part (iii) will be shown in full detail later in Theorem 2.29. □

Remark 1.29.

(i) The σ-field $\mathcal{F}_\tau$ contains all events which are known to have occured or not by time τ.

(ii) For discrete stopping times τ with values in $\{0,1,\ldots\} \cup \{\infty\}$ (or any other discrete ordered set) it is clear that

$$\{\tau = t\} = \{\tau \leq t\} \backslash \{\tau \leq t-1\}$$

and

$$\{\tau \leq t\} = \bigcup_{s\leq t} \{\tau = s\}$$

give

$$\mathcal{F}_\tau = \{A \in \mathcal{F} \mid A \cap \{\tau = t\} \in \mathcal{F}_t \text{ for all } t \in T\}.$$

Theorem 1.30. (Strong Markov property) *Let $B = (B_t)_{t\geq 0}$ be a Brownian motion with start in x, generated filtration $(\mathcal{F}_t)_{t\geq 0}$ and accompanying infinitesimal extension $(\mathcal{F}_t^+)_{t\geq 0}$ and let τ be an almost surely finite stopping time with respect to $(\mathcal{F}_t)_{t\geq 0}$. Then the process $(Z_t)_{t\geq 0}$ with*

$$Z_t = B_{\tau+t} - B_\tau$$

is a Brownian motion with start in 0 and independent of $\mathcal{F}_\tau^+$.

Proof: We will first discuss a discrete approximation of τ, namely

$$\tau_n = (m+1)2^{-n}, \quad \text{if } m2^{-n} \leq \tau < (m+1)2^{-n},\ m \in \mathbb{N}_0.$$

Note that $\{\tau_n \leq (\ell+1)2^{-n}\} = \{\tau < (\ell+1)2^{-n}\}$, so with a similar argument as in the proof of Theorem 1.26 each τ_n is a stopping time with respect to $(\mathcal{F}_t)_{t\geq 0}$ as well. Then we define $\{B_k(t) \mid t \geq 0\}$ via

$$B_k(t) = B(t + k2^{-n}) - B(k2^{-n})$$

and $\{B_*(t) \mid t \geq 0\}$ via

$$B_*(t) = B(t+\tau_n) - B(\tau_n).$$

Obviously, B_* has continuous paths. So if we want to check the Strong Markov property for τ_n we need to show that its finite dimensional distributions equal the ones of a Brownian motion B^0 starting in zero and that B_* is independent of $\mathcal{F}_{\tau_n}^+$. So let $E \in \mathcal{F}_{\tau_n}^+$ and A be measurable of arbitrary finite dimension. Then

$$\begin{aligned}\mathbb{P}(\{B_* \in A\} \cap E) &= \sum_{k=1}^{\infty} \mathbb{P}(\{B_k \in A\} \cap E \cap \{\tau_n = k2^{-n}\}) \\ &= \sum_{k=1}^{\infty} \mathbb{P}(\{B_k \in A\})\mathbb{P}(E \cap \{\tau_n = k2^{-n}\}),\end{aligned}$$

because B_k is independent from $E \cap \{\tau_n = k2^{-n}\} \in \mathcal{F}_{k2^{-n}}^+$ by Theorem 1.17 and Remark 1.29(ii). Also

$$\mathbb{P}(\{B_k \in A\}) = \mathbb{P}(\{B^0 \in A\})$$

by Theorem 1.12. Thus,

$$\begin{aligned}\sum_{k=1}^{\infty} \mathbb{P}(\{B_k \in A\})\mathbb{P}(E \cap \{\tau_n = k2^{-n}\}) &= \mathbb{P}(\{B^0 \in A\}) \sum_{k=1}^{\infty} \mathbb{P}(E \cap \{\tau_n = k2^{-n}\}) \\ &= \mathbb{P}(\{B^0 \in A\})\mathbb{P}(E).\end{aligned}$$

The general case uses $\tau_n \searrow \tau$ almost surely as well as that

$$B(t+\tau_n) - B(\tau_n)$$

is a Brownian motion independent of $\mathcal{F}_{\tau_n}^+ \supset \mathcal{F}_\tau^+$. In particular, the increments

$$B(t+s+\tau) - B(t+\tau) = \lim_{n\to\infty} (B(t+s+\tau_n) - B(t+\tau_n))$$

of $\{B(r+\tau) - B(\tau) \mid r \geq 0\}$ are independent over disjoint intervals and normal distributed with expectation 0 and variance s. The process is also almost surely continuous, thus a Brownian motion. Finally, all increments of $\{B(r+\tau) - B(\tau) \mid r \geq 0\}$, and thus the process itself, is independent of $\mathcal{F}_\tau^+$, passing to the limit again (compare the beginning of the proof of Theorem 1.17). □

Example 1.31. Let $\tau = \inf\{t \geq 0 \mid B_t = \max_{0 \leq s \leq 1} B_s\}$. It is intuitively clear that τ is not a stopping time, because at no time $t \in (0,1)$ we can be sure that we do not observe greater values of B later on.

Formally, note that because of the proof of Theorem 1.20 we have $\tau < 1$ almost surely. In this case it follows that $Z_t = B_{t+\tau} - B_\tau$ is non-positive in a neighbourhood of τ. But according to Theorem 1.19 Z then cannot be a Brownian motion with start in 0, so Theorem 1.30 proves that τ is not a stopping time.

Definition 1.32. Let $(B_t)_{t\geq 0}$ be a Brownian motion and τ a stopping time. Then the process $(W_t)_{t\geq 0}$ with

$$W_t = B_t 1_{\{t \leq \tau\}} + (2B_\tau - B_t) 1_{\{t > \tau\}}$$

is called a *Brownian motion reflected in* τ.

Theorem 1.33. (Reflection principle) $(W_t)_{t\geq 0}$ *is a Brownian motion.*

Proof: Clearly, W_t is a process with continuous paths and starting in zero, so we only need to show that distribution of any vector $(W_{t_1}, \ldots, W_{t_n})^T$ equals the one of $(B_{t_1}, \ldots, B_{t_n})^T$ to deduce the remaining defining properties of Brownian motion. For simplicity, we do this only in the one-dimensional case.

Let $a \in \mathbb{R}$ be arbitrary. Note by definition of W that

$$\mathbb{P}(W_t \leq a) = \mathbb{P}(W_t \leq a, \tau \geq t) + \mathbb{P}(W_t \leq a, \tau < t) = \mathbb{P}(B_t \leq a, \tau \geq t) + \mathbb{P}(W_t \leq a, \tau < t),$$

and for $\tau < t$ we write $W_t = B_\tau - Z_{t-\tau}$ with $Z_s = B_{\tau+s} - B_\tau$ as in Theorem 1.12. Iterated expectation plus the fact that both τ and B_τ are $\mathcal{F}_\tau$-measurable gives

$$\mathbb{P}(W_t \leq a, \tau < t) = \mathbb{E}[\mathbb{P}(-Z_{t-\tau} \leq a - B_\tau | \mathcal{F}_\tau) 1_{\{\tau < t\}}].$$

Now, Theorem 1.30 gives

$$\mathbb{P}(-Z_{t-\tau} \leq a - B_\tau | \mathcal{F}_\tau) = \mathbb{P}(Z_{t-\tau} \leq a - B_\tau | \mathcal{F}_\tau).$$

Repeating the same steps as above we obtain

$$\mathbb{P}(W_t \leq a, \tau < t) = \mathbb{E}[\mathbb{P}(Z_{t-\tau} \leq a - B_\tau | \mathcal{F}_\tau) 1_{\{\tau < t\}}] = \mathbb{P}(B_t \leq a, \tau < t)$$

and $\mathbb{P}(W_t \leq a) = \mathbb{P}(B_t \leq a)$ follows. $\square$

Theorem 1.34. *Let B be a Brownian motion and let $M_t = \max_{0 \leq s \leq t} B_s$. Then*

$$\mathbb{P}(M_t > \lambda) = 2\mathbb{P}(B_t > \lambda) = \mathbb{P}(|B_t| > \lambda)$$

for all $\lambda > 0$.

Proof: Let $\tau = \inf\{t \geq 0 \mid B_t = \lambda\}$ and let W be the Brownian motion reflected at τ. Then

$$\{M_t > \lambda\} = \{M_t > \lambda, B_t > \lambda\} \cup \{M_t > \lambda, B_t \leq \lambda\}.$$

Both events are disjoint, and we also know that $\{M_t > \lambda, B_t \leq \lambda\} = \{M_t > \lambda, W_t \geq \lambda\}$ by construction of W. Since both $B_t > \lambda$ and $W_t \geq \lambda$ imply $M_t > \lambda$, we get

$$\mathbb{P}(M_t > \lambda) = \mathbb{P}(B_t > \lambda) + \mathbb{P}(W_t \geq \lambda).$$

From Theorem 1.33 and $\mathbb{P}(W_t = \lambda) = 0$ we obtain

$$\mathbb{P}(M_t > \lambda) = 2\mathbb{P}(B_t > \lambda).$$

The second claim follows by symmetry of the normal distribution. $\square$

Remark 1.35. It is not difficult to see that $X \sim \mathcal{N}(0,1)$ satisfies the inequality

$$\mathbb{P}(X > x) \leq \frac{1}{x}\frac{1}{\sqrt{2\pi}}\exp(-x^2/2).$$

From 1.34 we thus conclude

$$\mathbb{P}(M_t > \lambda) \leq \frac{\sqrt{2t}}{\lambda\sqrt{\pi}}\exp(-\lambda^2/(2t)).$$

Chapter 2

Martingale theory

This chapter discusses martingales, a rich class of stochastic processes which are used to model fair games. The first part of the chapter reviews the discrete time case where we will refer to the lecture *Probability and Statistics II* for proofs. The second part extends this situation to continuous time.

Definition 2.1. Let $T \subset \mathbb{R}$ be an ordered index set and let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in T}, \mathbb{P})$ be a filtered probability space. A stochastic process $X = (X_t)_{t\in T}$ is called a *martingale* with respect to $(\mathcal{F}_t)_{t\in T}$ if

(i) X is adapted to $(\mathcal{F}_t)_{t\in T}$;

(ii) $\mathbb{E}[|X_t|] < \infty$ for all $t \in T$;

(iii) $\mathbb{E}[X_t|\mathcal{F}_s] = X_s$ for all $s, t \in T$ with $s \le t$.

Remark 2.2.

(i) The defining relation $\mathbb{E}[X_t|\mathcal{F}_s] = X_s$ for all $s \le t$ means essentially that by knowledge of the process X at time s one does not expect further gains or losses until t.

(ii) If $\mathbb{E}[X_t|\mathcal{F}_s] \le X_s$ or $\mathbb{E}[X_t|\mathcal{F}_s] \ge X_s$ for all $s \le t$, then the processes are called *super-* or *submartingales*, respectively.

Theorem 2.3. *Let X be a stochastic process, $(\mathcal{F}_t)_{t\in T}$ be a filtration and $\varphi : \mathbb{R} \to \mathbb{R}$ be a convex mapping with $\mathbb{E}[|\varphi(X_t)|] < \infty$ for all $t \in T$. If*

(i) X is a martingale or

(ii) X is a submartingale and $\varphi : \mathbb{R} \to \mathbb{R}$ is increasing,

then $(\varphi(X_t))_{t\in T}$ is a submartingale.

Example 2.4.

(i) Let $B = (B_t)_{t\ge 0}$ be a Brownian motion with start in x and $(\mathcal{F}_t)_{t\ge 0}$ its generated filtration. Then we have for all $0 \le s \le t$

$$\mathbb{E}[B_t|\mathcal{F}_s] = \mathbb{E}[B_s|\mathcal{F}_s] + \mathbb{E}[B_t - B_s|\mathcal{F}_s] = B_s + \mathbb{E}[B_t - B_s] = B_s,$$

because we know from Theorem 1.12 that $B_t - B_s$ is a Brownian motion with start in 0 and independent of $\mathcal{F}_s$.

(ii) Let $X_1, X_2, \dots$ be i.i.d. with $\mathbb{E}[|X_i|] < \infty$. Then the sequence $(S_n)_{n\in\mathbb{N}_0}$ of partial sums $S_n = \sum_{i=1}^n X_i$ with its generated filtration $(\mathcal{F}_n)_{n\in\mathbb{N}_0}$ satisfies

$$\mathbb{E}[S_n|\mathcal{F}_m] = S_m + \mathbb{E}[S_n - S_m] = S_m + (n-m)\mathbb{E}[X_i]$$

for all $0 \le m \le n$. Thus $(S_n)_{n\in\mathbb{N}}$ is a martingale if $\mathbb{E}[X_i] = 0$ holds. The other cases $\mathbb{E}[X_i] \le 0$ and $\mathbb{E}[X_i] \ge 0$ lead to super- and submartingales, respectively.

(iii) Let X be a random variable on $(\Omega, \mathcal{F}, \mathbb{P})$ with $\mathbb{E}[|X|] < \infty$ and let $(\mathcal{F}_t)_{t\ge 0}$ be any filtration. Then $(Y_t)_{t\ge 0}$ with

$$Y_t = \mathbb{E}[X|\mathcal{F}_t]$$

is a martingale, as one sees from

$$\mathbb{E}[|Y_t|] \le \mathbb{E}[\mathbb{E}[|X||\mathcal{F}_t]] = \mathbb{E}[|X|] < \infty$$

and

$$\mathbb{E}[Y_t|\mathcal{F}_s] = \mathbb{E}[\mathbb{E}[X|\mathcal{F}_t]|\mathcal{F}_s] = \mathbb{E}[X|\mathcal{F}_s] = Y_s, \quad s \le t,$$

using the conditional Jensen inequality, iterated expectation and the tower property. A similar result holds in discrete time.

Definition 2.5. Let $(\mathcal{F}_t)_{t\in\mathbb{N}_0}$ be a filtration.

(i) A process $V = (V_t)_{t\in\mathbb{N}}$ is called *predictable* with respect to $(\mathcal{F}_t)_{t\in\mathbb{N}_0}$ if any V_t is measurable with respect to $\mathcal{F}_{t-1}$.

(ii) If $X = (X_t)_{t\in\mathbb{N}_0}$ is another adapted process, then we define the *stochastic integral* $Y = V \bullet X$ via $Y_0 = 0$ and

$$Y_t = \sum_{s=1}^t V_s(X_s - X_{s-1}), \quad t \in \mathbb{N}.$$

(iii) A process Z is called *increasing* or *decreasing* if $t \mapsto Z_t(\omega)$ is almost surely increasing or decreasing, respectively.

Remark 2.6. The stochastic integral can be interpreted as follows: If X is a martingale, then in each round the *martingale differences* $X_s - X_{s-1}$ are weighted with the wager V_s whose value is known one period in advance. Y_t then is the total gain (or loss) by time t. Intuitively it is clear that most strategies V lead (on average) neither to gains nor to losses.

Lemma 2.7. *Let $X = (X_t)_{t\in\mathbb{N}_0}$ be a (sub-)martingale and $V = (V_t)_{t\in\mathbb{N}}$ be predictable with $V_t \ge 0$ and $||V_t||_\infty < \infty$ for all $t \in \mathbb{N}$. Then $V \bullet X$ is a (sub-)martingale as well.*

Theorem 2.8. (Doob decomposition) *Let $X = (X_t)_{t\in\mathbb{N}_0}$ be a stochastic process adapted to $(\mathcal{F}_t)_{t\in\mathbb{N}_0}$ with $\mathbb{E}[|X_t|] < \infty$ for all t. Then there exists a martingale $M = (M_t)_{t\in\mathbb{N}_0}$ with $M_0 = 0$ and a predictable process $V = (V_t)_{t\in\mathbb{N}_0}$ with $V_0 = 0$ such that*

$$X = X_0 + M + V$$

holds. The decomposition is almost surely unique.

Lemma 2.7 proves that a fair game remains a fair game if one plays it with a predictable strategy. A similar result holds in terms of stopping times if one imposes a minor additional structural assumption. Here we used boundedness of τ, but other conditions involving properties of X and τ work as well.

Theorem 2.9. (Optional Sampling Theorem) *Let $X = (X_t)_{t\in\mathbb{N}_0}$ be adapted to $(\mathcal{F}_t)_{t\in\mathbb{N}_0}$. Then the following are equivalent:*

(i) X is a martingale.

(ii) For all bounded stopping times τ we have $\mathbb{E}[X_\tau] = \mathbb{E}[X_0]$.

Example 2.10. The classical counterexample to Theorem 2.9 is the *martingale strategy* which uses an unbounded stopping time. Let $(X_t)_{t\in\mathbb{N}}$ be a sequence of i.i.d. random variables with $\mathbb{P}(X_t = 1) = \mathbb{P}(X_t = -1) = \frac{1}{2}$. Using Example 2.4(ii) we know that $(S_t)_{t\in\mathbb{N}}$ with

$$S_t = \sum_{s=1}^{t} X_s$$

is a martingale. In each round the wager is doubled, i.e. $V_s = 2^{s-1}$, so the resulting stochastic integral $Y = V \bullet S$ becomes

$$Y_t = \sum_{s=1}^{t} V_s(S_s - S_{s-1}) = \sum_{s=1}^{t} 2^{s-1} X_s$$

and is according to Lemma 2.7 again a martingale. The stopping time is chosen as $\tau = \inf\{t \in \mathbb{N} \mid X_t = 1\}$. Obviously $\tau < \infty$ almost surely, and we have

$$Y_\tau 1_{\{\tau=t\}} = \Big(2^{t-1} - \sum_{s=0}^{t-2} 2^s\Big) 1_{\{\tau=t\}} = 1_{\{\tau=t\}}$$

for all $t \in \mathbb{N}$. Thus $Y_\tau = 1$ almost surely, but τ is unbounded.

Theorem 2.11. (Optional Sampling Theorem) *Let $X = (X_t)_{t\in\mathbb{N}_0}$ be a (sub-) martingale with respect to $(\mathcal{F}_t)_{t\in\mathbb{N}_0}$, and let σ, τ be bounded stopping times with $\sigma \le \tau$. Then*

$$\mathbb{E}[X_\tau|\mathcal{F}_\sigma] \ge X_\sigma,$$

and we have equality for martingales.

We are now interested in the convergence of martingales and related processes. For supermartingales one can use an intuition from basic calculus: Intuitively, they are processes with a negative trend. Such sequences usually converge if they are bounded from below.

Theorem 2.12. (Doob's Convergence Theorem) *Let $X = (X_t)_{t\in\mathbb{N}_0}$ be a supermartingale with $\sup_{t\in\mathbb{N}_0} \mathbb{E}[|X_t|] < \infty$. Then $X_\infty = \lim_{n\to\infty} X_n$ exists almost surely and is integrable.*

Corollary 2.13. *Any non-negative supermartingale $(X_t)_{t\in\mathbb{N}_0}$ converges almost surely to an integrable limit.*

Remark 2.14. The standard proof of Theorem 2.12 is based on the *upcrossing inequality* which we will state below for later purposes. *Upcrossings* themselves are defined as follows: For a real-valued sequence $\alpha = (\alpha_n)_{n\geq 0}$, $n \geq 0$, $a, b \in \mathbb{R}$ with $a < b$ and a finite subset $F \subset \mathbb{N}_0$ one sets

$$\begin{aligned} U_F[a,b](\alpha) = \sup\{k \in \mathbb{N} \mid & \text{ there exist } 0 \leq s_1 < t_1 < \ldots < s_k < t_k \\ & \text{with } s_i, t_i \in F \text{ and with } \alpha_{s_i} < a, \alpha_{t_i} > b \text{ for all } i \in \{1, \ldots, k\}\} \end{aligned}$$

and

$$U_{\mathbb{N}_0}[a,b](\alpha) = \sup\{U_F[a,b](\alpha) \mid F \subset \mathbb{N}_0 \text{ finite}\}.$$

A simple result from basic calculus then gives that $\alpha = (\alpha_n)_{n\geq 0}$ converges in $\overline{\mathbb{R}}$ if and only if $U_{\mathbb{N}_0}[a,b](\alpha) < \infty$ for all $a, b \in \mathbb{Q}$ with $a < b$.

Lemma 2.15. **(Upcrossing inequality)** *Let $X = (X_t)_{t\in\mathbb{N}_0}$ be a supermartingale with $a < b$. Then*

$$\mathbb{E}[U_F[a,b](X)] \leq \frac{\sup_{t\in\mathbb{N}_0} \mathbb{E}[(X_t - a)^-]}{b - a}$$

for any finite $F \subset \mathbb{N}_0$.

Remark 2.16. While Theorem 2.12 deals with almost sure convergence of supermartingales, the situation is a little more complicated if one is interested in L^p convergence of martingales. So let $X = (X_t)_{t\in\mathbb{N}_0}$ be a martingale with respect to $(\mathcal{F}_t)_{t\in\mathbb{N}_0}$.

(i) For L^1 the following are equivalent:

  (a) X is uniformly integrable, i.e. $\lim_{k\to\infty} \sup_{t\in\mathbb{N}_0} \mathbb{E}[|X_t| 1_{\{|X_t|\geq k\}}] = 0$;

  (b) X converges almost surely and in L^1;

  (c) There exists a random variable $Y \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ such that $X_t = \mathbb{E}[Y|\mathcal{F}_t]$.

  If any of these conditions holds, Y from (c) equals the limit almost surely.

(ii) For L^2 the following are equivalent:

  (a) $\sup_{t\in\mathbb{N}_0} \mathbb{E}[X_t^2] < \infty$;

  (b) $(X_t)_{t\in\mathbb{N}_0}$ converges almost surely and in L^2.

Definition 2.17. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $(\mathcal{F}_n)_{n\leq 0}$ be a filtration. A process $X = (X_n)_{n\leq 0}$ is called a *backwards martingale*, if

(i) X is adapted to $(\mathcal{F}_n)_{n\leq 0}$;

(ii) $\mathbb{E}[|X_n|] < \infty$ for all $n \leq 0$;

(iii) $\mathbb{E}[X_n|\mathcal{F}_m] = X_m$ for all $m \leq n \leq 0$.

Theorem 2.18. *A backwards martingale $X = (X_n)_{n\leq 0}$ converges almost surely and in L^1 to $X_{-\infty}$. In particular,*

$$X_{-\infty} = \mathbb{E}[X_0|\mathcal{F}_{-\infty}] \quad \textit{with} \quad \mathcal{F}_{-\infty} = \bigcap_{n\leq 0} \mathcal{F}_n.$$

The remaining part of the chapter deals with martingales in continuous time. We will first use the convergence results in discrete time to prove helpful path properties of continuous time martingales. Later on we discuss Optional Sampling Theorems. This discussion will not involve sub- or supermartingales, which one can treat as well, using a continuous time version of Theorem 2.8.

Lemma 2.19. *Let $X = (X_t)_{t\geq 0}$ be a martingale. Then there exists $\Omega_0 \subset \Omega$ with $\mathbb{P}(\Omega_0) = 1$ such that*

$$\lim_{r\nearrow t, r\in\mathbb{Q}} X_r(\omega) \quad \textit{and} \quad \lim_{r\searrow t, r\in\mathbb{Q}} X_r(\omega)$$

exist for all $t \geq 0$ and all $\omega \in \Omega_0$.

Proof: Since probability measures are continuous from below it is sufficient to prove the result for all $t \in [0, k]$, with $k \in \mathbb{N}$ being arbitrary.

First, because $x \mapsto x^- = -x1_{\{x<0\}}$ is a convex function, we obtain from Theorem 2.3 that the process X_t^- is a submartingale. Thus

$$\mathbb{E}[(X_t - a)^-] \leq \mathbb{E}[X_t^-] + |a| \leq \mathbb{E}[X_k^-] + |a|.$$

We now apply Lemma 2.15 in a version for martingales indexed by rational numbers. Setting $T = [0, k] \cap \mathbb{Q}$, the same proof gives

$$(b-a)\mathbb{E}[U_F[a,b](X)] \leq \sup_{t\in T} \mathbb{E}[(X_t - a)^-]$$

for every finite $F \subset T$. We obtain

$$(b-a)\mathbb{E}[U_{[0,k]\cap\mathbb{Q}}[a,b](X)] = (b-a)\sup\{U_F[a,b](\alpha) \mid F \subset T \text{ finite}\} \leq \mathbb{E}[X_k^-] + |a|,$$

and we conclude that there exists Ω_0 with $\mathbb{P}(\Omega_0) = 1$ so that $U_{[0,k]\cap\mathbb{Q}}[a,b](X) < \infty$ for all $\omega \in \Omega_0$. Remark 2.14 then proves that the limits exist as claimed. □

Lemma 2.20. *Let $X = (X_t)_{t\geq 0}$ be a martingale. Then $X_{t+} = \lim_{r\searrow t, r\in\mathbb{Q}} X_r$ satisfies*

$$\mathbb{E}[X_{t+}|\mathcal{F}_t] = X_t$$

almost surely.

Proof: Let $t \in \mathbb{R}$. Then there exists a sequence $(t_n)_{n\geq 0}$ of rational numbers with $t_n \searrow t$. By construction $(X_{t_n})_{n\geq 0}$ is a backwards martingale, and Theorem 2.18 gives

$$X_{t+} = \lim_{n\to\infty} X_{t_n} \text{ in } L^1.$$

Thus, for every $F \in \mathcal{F}_t$

$$\mathbb{E}[X_t 1_F] = \lim_{n\to\infty} \mathbb{E}[X_{t_n} 1_F] = \mathbb{E}[X_{t+} 1_F],$$

where we have used $\mathbb{E}[X_{t_n}|\mathcal{F}_t] = X_t$ in the first identity. □

Remark 2.21. X_{t+} is measurable with respect to to all $\mathcal{F}_{t_n}$ by construction, thus also with respect to $\bigcap_{u>t} \mathcal{F}_u = \bigcap_{n\geq 1} \mathcal{F}_{t_n}$. Since we assume right continuity of the filtration, measurability of X_{t+} with respect to $\mathcal{F}_t$ follows. Lemma 2.20 then gives $X_{t+} = X_t$ almost surely.

Definition 2.22. Two stochastic processes X and Y on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ are called *modifications* if for every $t \geq 0$

$$X_t(\omega) = Y_t(\omega)$$

holds almost surely.

Theorem 2.23. *Every martingale $X = (X_t)_{t\geq 0}$ has a modification which is a martingale itself and has càdlàg paths.*

Proof: Let Ω_0 be the set from Lemma 2.19 on which the left and right limits exist. We then set

$$\widetilde{X_t}(\omega) = \begin{cases} X_{t+}(\omega), & \text{if } \omega \in \Omega_0, \\ 0, & \text{else.} \end{cases}$$

This means that, on Ω_0, we decide for the limit from the right if X_{t-} and X_{t+} do not coincide, so $t \mapsto X_{t+}(\omega)$ is càdlàg. This property obviously also holds for $t \mapsto \widetilde{X_t}(\omega)$. In particular,

$$\widetilde{X_t}(\omega) = X_{t+}(\omega) \text{ almost surely}$$

using Lemma 2.19, and Remark 2.21 gives

$$X_{t+}(\omega) = X_t(\omega) \text{ almost surely.}$$

Therefore $\widetilde{X_t}$ is a modification of X_t.

We finally prove that $(\widetilde{X_t})_{t\geq 0}$ is a martingale with respect to $(\mathcal{F}_t)_{t\geq 0}$, and it is sufficient to prove the claim for $(X_{t+})_{t\geq 0}$. Let $s < t$. We choose a sequence $(s_n)_n$ of rational numbers with $t > s_n \searrow s$. The martingale property of $(X_t)_{t\geq 0}$ and Lemma 2.20 give

$$X_{s_n} = \mathbb{E}[X_t|\mathcal{F}_{s_n}] = \mathbb{E}[\mathbb{E}[X_{t_+}|\mathcal{F}_t]|\mathcal{F}_{s_n}] = \mathbb{E}[X_{t_+}|\mathcal{F}_{s_n}]$$

for every n. It follows that $\dots, X_{s_n}, \dots, X_{s_1}, X_{t+}$ is a backwards martingale, and Theorem 2.18 gives

$$X_{s+} = \lim_{n\to\infty} X_{s_n} = \mathbb{E}[X_{t_+}|\bigcap_n \mathcal{F}_{s_n}]$$

almost surely. The result then follows from $\bigcap_n \mathcal{F}_{s_n} = \bigcap_{u>s} \mathcal{F}_u = \mathcal{F}_s$. □

Remark 2.24. Unless explicitly state otherwise, we always work with the càdlàg modification of a martingale.

Definition 2.25. An adapted process $(X_t)_{t\geq 0}$ is called *progressively measurable* if

$$\begin{aligned} \Omega \times [0,t] &\to \mathbb{R} \\ (\omega, s) &\mapsto X_s(\omega) \end{aligned}$$

is $(\mathcal{F}_t \otimes \mathcal{B}_{[0,t]})$-measurable for all $t \geq 0$.

Remark 2.26. Progressive measurability is a much stronger requirement than just adaptedness. We not only assume that a fixed X_s is measurable with respect to $\mathcal{F}_s$ but also that the resulting functions $s \mapsto X_s(\omega)$ are measurable functions on $[0,t]$ for any fixed ω and t. We will see, however, that the two notions are equivalent if the entire process is already determined by its values at countably many times, as the next result shows.

Lemma 2.27. *Every adapted right continuous process $(X_t)_{t\geq 0}$ is progressively measurable.*

Proof: We set $(X_s^n)_{s\in[0,t]}$ for every $n\in\mathbb{N}$ via

$$X_s^n(\omega) = X_0(\omega)1_{\{0\}}(s) + \sum_{k=1}^{2^n} X_{\frac{kt}{2^n}}(\omega)1_{\left(\frac{(k-1)t}{2^n},\frac{kt}{2^n}\right]}(s).$$

Since every $(\omega,s)\mapsto X_{\frac{kt}{2^n}}(\omega)$ is $(\mathcal{F}_t\otimes\mathcal{B}_{[0,t]})$-measurable, so is $(\omega,s)\mapsto X_s^n(\omega)$. Finally, by right continuity $\lim_{n\to\infty}X_s^n(\omega)=X_s(\omega)$, so the limit $(\omega,s)\mapsto X_s(\omega)$ is $(\mathcal{F}_t\otimes\mathcal{B}_{[0,t]})$-measurable as well. □

Theorem 2.28. *Let $(X_t)_{t\geq 0}$ be an adapted right continuous process and τ a stopping time. Then $X_\tau 1_{\{\tau<\infty\}}$ is $\mathcal{F}_\tau$-measurable.*

Proof: By Lemma 2.27 we have to prove that $X_\tau 1_{\{\tau<\infty\}}$ is $\mathcal{F}_\tau$-measurable whenever X is progressively measurable. Let $B\in\mathcal{B}$ and $t\geq 0$. The mapping

$$\begin{aligned}\sigma:\Omega&\to\Omega\times[0,t]\\ \omega&\mapsto(\omega,\tau(\omega)\wedge t)\end{aligned}$$

satisfies, for all $F\in\mathcal{F}_t$ and all $s\leq t$

$$\sigma^{-1}(F\times[0,s]) = F\cap\{\tau\leq s\}\in\mathcal{F}_t,$$

so it is $\mathcal{F}_t-(\mathcal{F}_t\otimes\mathcal{B}_{[0,t]})$-measurable. In particular,

$$\{X\circ\sigma\in B\}\in\mathcal{F}_t$$

because X is progressively measurable. Then

$$\{X_\tau 1_{\{\tau<\infty\}}\in B\}\cap\{\tau\leq t\} = \{X_{\tau\wedge t}\in B\}\cap\{\tau\leq t\} = \{X\circ\sigma\in B\}\cap\{\tau\leq t\}\in\mathcal{F}_t$$

gives the claim. □

Theorem 2.29. (Optional Sampling Theorem) *Let $X=(X_t)_{t\geq 0}$ be a right continuous (sub-)martingale with respect to $(\mathcal{F}_t)_{t\geq 0}$ and let σ, τ be bounded stopping times with $\sigma\leq\tau$. Then*

$$\mathbb{E}[X_\tau|\mathcal{F}_\sigma]\geq X_\sigma$$

with equality for martingales.

Proof: As in the proof of Theorem 1.30 we approximate τ via

$$\tau_n = (m+1)2^{-n},\quad \text{if } m2^{-n}\leq\tau<(m+1)2^{-n},\ m\in\mathbb{N}_0,$$

and every τ_n is a stopping time itself with $\tau_n\searrow\tau$. The same procedure is done for σ. By assumption, the discrete process $(X_{m2^{-n}})_{m\geq 0}$ is a (sub-)martingale with respect to $(\mathcal{F}_{m2^{-n}})_{m\geq 0}$, and Theorem 2.11 gives

$$\mathbb{E}[X_{\tau_n}|\mathcal{F}_{\sigma_n}]\geq X_{\sigma_n}$$

for all n. In particular,

$$\mathbb{E}[X_{\tau_n}1_A]\geq\mathbb{E}[X_{\sigma_n}1_A]$$

for all $A \in \mathcal{F}_\sigma \subset \mathcal{F}_{\sigma_n}$. (In both cases we have equality for martingales.) The claim then follows from the definition of conditional expectation as long as

$$\mathbb{E}[X_{\tau_n} 1_A] \to \mathbb{E}[X_\tau 1_A] \text{ for all } A \in \mathcal{F}_\sigma \tag{2.1}$$

can be shown. (The analogous claim with σ instead of τ can be shown similarly.)

For the proof of (2.1) we first use right continuity of X and $\tau_n \searrow \tau$, from which $X_{\tau_n} \searrow X_\tau$ almost surely follows. In the martingale case we also have, because of $\tau \leq N$ by boundedness,

$$\mathbb{E}[X_N | \mathcal{F}_{\tau_n}] = X_{\tau_n}.$$

Such sequences $(X_{\tau_n})_{n\geq 0}$ can be shown to be uniformly integrable, and then $X_{\tau_n} \searrow X_\tau$ in L^1 from Remark 2.16(i) as well.

In the true submartingale case we first assume additionally that $X_t \geq c$ for all $t \geq 0$. Then

$$\mathbb{E}[X_N | \mathcal{F}_{\tau_n}] \geq X_{\tau_n} \geq c.$$

Obviously $(|X_{\tau_n}|)_n$ is then bounded by a sequence of uniformly integrable random variables, hence uniformly integrable as well. (2.1) then follows, using that almost sure convergence and uniform integrability implies convergence in L^1.

Finally, we discuss the sequence $X_t^k = X_t \vee (-k)$. As $x \mapsto x \vee (-k)$ is convex and increasing, we obtain from Theorem 2.3 that $(X_t^k)_{t\geq 0}$ is a submartingale as well. Thus

$$\mathbb{E}[X_\tau^k | \mathcal{F}_\sigma] \geq X_\sigma^k,$$

using the claim for submartingales bounded from below. Monotone convergence, applied in its conditional version to $(-X_\tau^k)_{k\in\mathbb{N}}$, then gives

$$\mathbb{E}[X_\tau | \mathcal{F}_\sigma] = \lim_{k\to\infty} \mathbb{E}[X_\tau^k | \mathcal{F}_\sigma] \geq \lim_{k\to\infty} X_\sigma^k = X_\sigma,$$

hence the claim. □

The chapter ends with helpful maximal inequalities for martingales.

Definition 2.30. Let $T = \mathbb{N}_0$ or $T = [0, \infty)$, and let $(X_t)_{t\in T}$ be a stochastic process. Then

$$X_t^* = \sup_{s\in T, s\leq t} |X_s|, \quad t \in T,$$

is called the *running absolute supremum* until t.

Remark 2.31. By Markov inequality one obtains the trivial bounds

$$\mathbb{P}(X_t^* \geq \lambda) \leq \frac{\mathbb{E}[X_t^*]}{\lambda} \quad \text{and} \quad \mathbb{P}(X_t^* \geq \lambda) \leq \frac{\mathbb{E}[(X_t^*)^p]}{\lambda^p}$$

for every $\lambda > 0$ and $p > 0$.

Theorem 2.32. (Doob's inequalities) *Let $T = \mathbb{N}_0$ or $T = [0, \infty)$, and let $(X_t)_{t\in T}$ be a (right continuous) martingale or a non-negative (right continuous) submartingale. Then:*

(i) $\mathbb{P}(X_t^* \geq \lambda) \leq \frac{\mathbb{E}[|X_t|]}{\lambda}$ *for all $\lambda > 0$ and all $t \in T$.*

(ii) $\mathbb{P}(X_t^* \geq \lambda) \leq \frac{\mathbb{E}[|X_t|^p]}{\lambda^p}$ *for all $\lambda > 0$, all $p > 1$ and all $t \in T$.*

(iii) $\mathbb{E}[(X_t^*)^p] \leq \left(\frac{p}{p-1}\right)^p \mathbb{E}[|X_t|^p]$ *for all* $p > 1$ *and all* $t \in T$.

Proof: Theorem 2.3 proves that $(|X_t|)_{t\in T}$ is a non-negative submartingale always, and in case of the existence of the pth moment this also holds for $(|X_t|^p)_{t\in T}$.

(i) From Remark 1.25 we have that

$$\tau = \inf\{t \in T \mid |X_t| \geq \lambda\} \tag{2.2}$$

is a stopping time, even with right continuity only, and we have $X_t^* \geq \lambda$ if and only if $\tau \leq t$. In this case $|X_{\tau\wedge t}| \geq \lambda$ follows by right continuity. Then

$$\mathbb{P}(X_t^* \geq \lambda) = \mathbb{P}(\tau \leq t) = \mathbb{P}(\tau \leq t, |X_{\tau\wedge t}| \geq \lambda) \leq \frac{1}{\lambda}\mathbb{E}[|X_{\tau\wedge t}|1_{\{\tau\leq t\}}]. \tag{2.3}$$

By definition of a stopping time $\{\tau \leq t\} \in \mathcal{F}_t$, and Lemma 1.28(ii) gives $\{\tau \leq t\} \in \mathcal{F}_\tau$ as well. An exercise proves $\mathcal{F}_\tau \cap \mathcal{F}_t = \mathcal{F}_{\tau\wedge t}$, thus $\{\tau \leq t\} \in \mathcal{F}_{\tau\wedge t}$. Theorem 2.11 and Theorem 2.29 finally give $|X_{\tau\wedge t}| \leq \mathbb{E}[|X_t||\mathcal{F}_{\tau\wedge t}]$, so using iterated expectation

$$\mathbb{E}[|X_{\tau\wedge t}|1_{\{\tau\leq t\}}] \leq \mathbb{E}[|X_t|1_{\{\tau\leq t\}}] \leq \mathbb{E}[|X_t|]. \tag{2.4}$$

(ii) The proof works in the same way as (i) if $\mathbb{E}[|X_t|^p] < \infty$. Otherwise the inequality holds trivially.

(iii) For every random variable $Y \geq 0$ we have

$$\int_0^\infty px^{p-1}\mathbb{P}(Y > x)dx = \int_0^\infty \int px^{p-1}1_{\{Y>x\}}d\mathbb{P}dx = \int \int_0^\infty px^{p-1}1_{\{Y>x\}}dxd\mathbb{P} \tag{2.5}$$

using Fubini's theorem. Because of

$$\int_0^\infty px^{p-1}1_{\{Y>x\}}dx = \int_0^Y px^{p-1}dx = Y^p$$

the right hand side of (2.5) equals $\mathbb{E}[Y^p]$. With $\tau = \tau_\lambda$ as in (2.2)

$$\begin{aligned}\frac{1}{p}\mathbb{E}[(X_t^*)^p] &= \int_0^\infty \lambda^{p-1}\mathbb{P}(X_t^* > \lambda)d\lambda \leq \int_0^\infty \lambda^{p-2}\mathbb{E}[|X_t|1_{\{\tau_\lambda\leq t\}}]d\lambda \\ &= \int_0^\infty \lambda^{p-2}\mathbb{E}[|X_t|1_{\{X_t^*\geq\lambda\}}]d\lambda,\end{aligned}$$

using (2.3) and (2.4). Fubini's theorem again shows that the right hand side is

$$\begin{aligned}\int\int_0^\infty \lambda^{p-2}|X_t|1_{\{X_t^*\geq\lambda\}}d\lambda d\mathbb{P} &= \int |X_t| \int_0^\infty \lambda^{p-2}1_{\{X_t^*\geq\lambda\}}d\lambda d\mathbb{P} \\ &= \frac{1}{p-1}\mathbb{E}[|X_t|(X_t^*)^{p-1}].\end{aligned}$$

Hölder inequality finally shows

$$\mathbb{E}[|X_t|(X_t^*)^{p-1}] \leq \mathbb{E}[|X_t|^p]^{1/p}\mathbb{E}[(X_t^*)^p]^{(p-1)/p},$$

and this gives the claim. □

Chapter 3

Donsker's invariance principle

This chapter deals with Donsker's invariance principle which proves that Brownian motion naturally arises as the limit of the partial sum process. This result can be seen as a process-valued analogue to the central limit theorem and thus explains the role of Brownian motion as the prototypical process in continuous time.

Theorem 3.1. (Wald's first lemma) *Let $(B_t)_{t\geq 0}$ be a Brownian motion and let τ be a stopping time with $\mathbb{E}[\tau] < \infty$. Then*

$$\mathbb{E}[B_\tau] = 0.$$

Proof: We first prove the claim under the additional assumption that there exists some $Z \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ such that $|B_{t\wedge\tau}| \leq Z$ for all $t \geq 0$. Then

$$\tau_m = \tau \wedge m$$

is a bounded stopping time, and Theorem 2.29 gives

$$\mathbb{E}[B_{\tau_m}] = \mathbb{E}[B_0] = 0$$

because B is a martingale starting at zero. We also have $\lim_{m\to\infty} B_{\tau_m} = B_\tau$ almost surely because τ is almost surely finite. The additional $|B_{\tau_m}| \leq Z$ then allows us to deduce the claim from dominated convergence.

To prove that there is such a Z which dominates $(B_{t\wedge\tau})_{t\geq 0}$ we set

$$Z_k = \max_{0\leq t\leq 1} |B_{t+k} - B_k| \quad \text{and} \quad Z = \sum_{k=0}^{\lceil\tau\rceil-1} Z_k,$$

with $x \mapsto \lceil x \rceil$ denoting the ceiling function. By construction we have $|B_{t\wedge\tau}| \leq Z$, and obviously

$$\mathbb{E}[Z] = \sum_{k=0}^{\infty} \mathbb{E}[Z_k 1_{\{\tau>k\}}]$$

holds as well. Since $\{\tau > k\}$ is $\mathcal{F}_k$-measurable, Theorem 1.12 gives

$$\mathbb{E}[Z] = \sum_{k=0}^{\infty} \mathbb{E}[Z_k 1_{\{\tau>k\}}] = \mathbb{E}[Z_0] \sum_{k=0}^{\infty} \mathbb{P}(\tau > k) \leq \mathbb{E}[Z_0] \sum_{k=0}^{\infty} \mathbb{P}(\lceil\tau\rceil > k) = \mathbb{E}[Z_0]\mathbb{E}[\lceil\tau\rceil]$$

using

$$\mathbb{E}[N] = \sum_{k=0}^{\infty} \mathbb{P}(N > k) = \sum_{k=1}^{\infty} \mathbb{P}(N \geq k)$$

for any random variable N with values in $\mathbb{N}$. As $\lceil \tau \rceil \leq \tau + 1$ holds, the claim follows from

$$\begin{aligned}\mathbb{E}[Z_0] &= \int_0^\infty \mathbb{P}\Big(\max_{0\leq t\leq 1} |B_t| > x\Big)dx \leq 2\int_0^\infty \mathbb{P}\Big(\max_{0\leq t\leq 1} B_t > x\Big)dx = 4\int_0^\infty \mathbb{P}(B_1 > x)dx \\ &\leq 4\Big(\int_0^1 \mathbb{P}(B_1 > 0)dx + \int_1^\infty \mathbb{P}(B_1 > x)dx\Big) \leq 2 + \int_1^\infty \frac{4}{x\sqrt{2\pi}} \exp(-x^2/2)dx\end{aligned}$$

where we have used Fubini's theorem again (see also the proof of Theorem 2.32), Theorem 1.34 and Remark 1.35. The latter integral is clearly finite. □

Lemma 3.2. *Let $(B_t)_{t\geq 0}$ be a Brownian motion. Then the process $(Y_t)_{t\geq 0}$ with*

$$Y_t = B_t^2 - t$$

is a martingale with respect to the natural filtration.

Proof: The process is adapted, we have $\mathbb{E}[|Y_t|] < \infty$ by properties of the normal distribution, and we have

$$\begin{aligned}\mathbb{E}[B_t^2 - t|\mathcal{F}_s] &= \mathbb{E}[(B_t - B_s)^2|\mathcal{F}_s] + 2\mathbb{E}[B_tB_s|\mathcal{F}_s] - B_s^2 - t \\ &= (t-s) + 2B_s^2 - B_s^2 - t = B_s^2 - s,\end{aligned}$$

where we have used Theorem 1.12 and the martingale property of Brownian motion. □

Lemma 3.3. *Let $(B_t)_{t\geq 0}$ be a Brownian motion and let $\sigma \leq \tau$ be stopping times with $\mathbb{E}[\tau] < \infty$. Then*

$$\mathbb{E}[B_\tau^2] = \mathbb{E}[B_\sigma^2] + \mathbb{E}[(B_\tau - B_\sigma)^2].$$

Proof: This claim is a consequence of Theorem 3.1 and Theorem 1.30. The formal proof will be discussed as an exercise. □

Theorem 3.4. (Wald's second lemma) *Let $(B_t)_{t\geq 0}$ be a Brownian motion and let τ be a stopping time with $\mathbb{E}[\tau] < \infty$. Then*

$$\mathbb{E}[B_\tau^2] = \mathbb{E}[\tau].$$

Proof: First we discuss the martingale Y with $Y_t = B_t^2 - t$. For every $m \in \mathbb{N}$ we again have that

$$\tau_m = \tau \wedge m$$

is a bounded stopping time, and Theorem 2.29 proves

$$\mathbb{E}[Y_{\tau_m}] = \mathbb{E}[Y_0] = 0, \quad \text{so} \quad \mathbb{E}[B_{\tau_m}^2] = \mathbb{E}[\tau_m].$$

Now we know from Lemma 3.3 that $\mathbb{E}[\tau] < \infty$ implies $\mathbb{E}[B_\tau^2] \geq \mathbb{E}[B_{\tau_m}^2]$. Thus

$$\mathbb{E}[B_\tau^2] \geq \lim_{m\to\infty} \mathbb{E}[B_{\tau_m}^2] = \lim_{m\to\infty} \mathbb{E}[\tau_m] = \mathbb{E}[\tau]$$

where we have used monotone convergence because τ is almost surely finite. On the other hand,

$$\mathbb{E}[B_\tau^2] = \mathbb{E}[\lim_{m\to\infty} B_{\tau_m}^2] \le \liminf_{m\to\infty} \mathbb{E}[B_{\tau_m}^2] = \liminf_{m\to\infty} \mathbb{E}[\tau_m] = \mathbb{E}[\tau]$$

from Fatou's lemma. □

Example 3.5. Let B be a Brownian motion and $a < 0 < b$, and we discuss the stopping time

$$\tau = \inf\{t \ge 0 \mid B_t \notin [a,b]\}.$$

We show first $\mathbb{E}[\tau] < \infty$ in order to apply Wald's lemmata. We have

$$\mathbb{E}[\tau] = \int_0^\infty \mathbb{P}(\tau > t)dt = \int_0^\infty \mathbb{P}(B_s \in (a,b) \text{ for all } s \in [0,t])dt \tag{3.1}$$

using Theorem 1.19, and we set

$$\varrho = \sup_{x\in[a,b]} \mathbb{P}_x(B_s \in (a,b) \text{ for all } s \in [0,1]),$$

where $\mathbb{P}_x(\cdot)$ denotes the distribution of Brownian motion with start in x. Obviously $\varrho < 1$, and using Theorem 1.12, for any $t \ge k \in \mathbb{N}$, we have

$$\mathbb{P}(B_s \in (a,b) \text{ for all } s \in [0,t]) \le \varrho^k.$$

Now the integrand in (3.1) becomes exponentially small and the integral exists.

Theorem 3.1 now proves

$$0 = \mathbb{E}[B_\tau] = a\mathbb{P}(B_\tau = a) + b\mathbb{P}(B_\tau = b).$$

Because $\mathbb{P}(B_\tau = a) + \mathbb{P}(B_\tau = b) = 1$ this means

$$\mathbb{P}(B_\tau = a) = \frac{b}{|a|+b} \quad \text{and} \quad \mathbb{P}(B_\tau = b) = \frac{|a|}{|a|+b}.$$

Furthermore,

$$\mathbb{E}[\tau] = \mathbb{E}[B_\tau^2] = \frac{a^2 b}{|a|+b} + \frac{b^2|a|}{|a|+b} = |a|b$$

where we have used Theorem 3.4.

Remark 3.6. Wald's first and second lemma give a claim about the first two moments of B_τ, but not about its distribution. We will show in the following that every distribution with a vanishing first and a finite second moment is possible.

Definition 3.7. A martingale $(X_n)_{n\in\mathbb{N}_0}$ is called *binary splitting* if for every

$$A(x_0,\dots,x_n) = \{X_0 = x_0,\dots,X_n = x_n\}$$

with a positive probability, the conditional distribution of X_{n+1} given $A(x_0,\dots,x_n)$ only has two possible values.

Lemma 3.8. *Let X be a random variable with $\mathbb{E}[X^2] < \infty$. Then there exists a binary splitting martingale $(X_n)_{n\in\mathbb{N}_0}$ with*

$$X_n \to X \text{ almost surely and in } L^2(\Omega,\mathcal{F},\mathbb{P}).$$

Proof: We define both the martingale and the filtration recursively. First we set $X_0 = \mathbb{E}[X]$ and $\mathcal{G}_0 = \{\emptyset, \Omega\}$, and we let

$$\xi_0 = \begin{cases} 1, & \text{if } X \geq X_0, \\ -1, & \text{if } X < X_0. \end{cases}$$

Then we set recursively

$$\mathcal{G}_n = \sigma(\xi_0, \dots, \xi_{n-1}), \quad X_n = \mathbb{E}[X|\mathcal{G}_n]$$

and

$$\xi_n = \begin{cases} 1, & \text{if } X \geq X_n, \\ -1, & \text{if } X < X_n. \end{cases}$$

Because of Example 2.4 X_n clearly is a martingale. Now, for every $n \geq 1$, $\mathcal{G}_n$ is generated by a partition $\mathcal{P}_n$ of the probability space into 2^n sets of the form

$$A(x_0, \dots, x_{n-1}) = \{\xi_0 = x_0, \dots, \xi_{n-1} = x_{n-1}\},$$

where $x_i \in \{1, -1\}$ holds. Every element of $\mathcal{P}_n$ additionally is a union of two elements of $\mathcal{P}_{n+1}$ which belong to $\xi_n = 1$ and $\xi_n = -1$, respectively. As X_{n+1} is $\mathcal{G}_{n+1}$-measurable, with knowledge of $A(x_0, \dots, x_{n-1})$ there are only two possible values, so the martingale is binary splitting.

Because of $\mathbb{E}[X|\mathcal{G}_n] = X_n$, we have from iterated expectation

$$\mathbb{E}[(X - X_n)X_n] = \mathbb{E}[\mathbb{E}[X - X_n|\mathcal{G}_n]X_n] = 0$$

as well. Therefore

$$\mathbb{E}[X^2] = \mathbb{E}[(X - X_n)^2] + \mathbb{E}[X_n^2] \geq \mathbb{E}[X_n^2]$$

and $\sup_n \mathbb{E}[X_n^2]$ is bounded. By Remark 2.16 we obtain convergence of $(X_n)_{n\geq 0}$ almost surely and in L^2 with limit X_∞.

We now prove the auxiliary

$$\lim_{n\to\infty} \xi_n(X - X_{n+1}) = |X - X_\infty| \tag{3.2}$$

almost surely. For $X = X_\infty$ both sides of the equation are obviously zero, and if $X(\omega) < X_\infty(\omega)$, then $X(\omega) < X_n(\omega)$ for all $n \geq m$ with $m \in \mathbb{N}$ large enough (depending on ω). By definition, $\xi_n = -1$ for all $n \geq m$, and again (3.2) follows. Analogously the case $X(\omega) > X_\infty(\omega)$ can be discussed.

Finally, we use that if a sequence $(Y_n)_n$ satisfies $Y_n \to Y$ almost surely and $\mathbb{E}[Y_n^2] \leq \mathbb{E}[Z^2]$ for some $Z \in L^2(\Omega, \mathcal{F}, \mathbb{P})$, then uniform integrability yields $\mathbb{E}[Y_n] \to \mathbb{E}[Y]$ as well. We apply this rule to $(\xi_n(X - X_{n+1}))_n$. Then

$$\mathbb{E}[\xi_n(X - X_{n+1})] = \mathbb{E}[\xi_n \mathbb{E}[X - X_{n+1}|\mathcal{G}_{n+1}]] = 0$$

gives $\mathbb{E}[|X - X_\infty|] = 0$ from (3.2). □

Theorem 3.9. (Skorokhod embedding) *Let $(B_t)_{t\geq 0}$ be a Brownian motion and X a random variable with $\mathbb{E}[X] = 0$ and $\mathbb{E}[X^2] < \infty$. Then there exists a stopping time τ with respect to the filtration generated by B, so that B_τ has the same distribution as X and satisfies $\mathbb{E}[\tau] = \mathbb{E}[X^2]$.*

Proof: From Lemma 3.8 we know that there exists a binary splitting martingale $(X_n)_{n\geq 0}$ with $X_n \to X$ almost surely and in L^2.

The key to the proof is now that Example 3.5 gives the solution for every Y which takes values in $\{a, b\}$ and satisfies $\mathbb{E}[Y] = 0$. In this case the stopping time is given by

$$\tau(a, b) = \inf\{t \geq 0 \mid B_t \notin [a, b]\}.$$

This fact opens the door for a recursive definition of the stopping times for which we are following the construction given in the proof of Lemma 3.8. X_1 only takes two possible values a_1 and b_1, and because of $\mathbb{E}[X] = 0$ one is non-negative and the other is non-positive. Therefore there exists a stopping time $\tau_1 = \tau(a_1, b_1)$ with $X_1 = B_{\tau_1}$ in distribution, and Theorem 3.4 gives $\mathbb{E}[X_1^2] = \mathbb{E}[\tau_1]$.

The recursive construction now works as follows: Given $X_1 = a_1$, we only have two possible values $d(a_1)$ and $u(a_1)$ for X_2, and again by construction $d(a_1) \leq a_1 \leq u(a_1)$ because the conditional expectation of X_2 given $X_1 = a_1$ equals a_1 by the martingale property. Similarly for $X_1 = b_1$. Therefore one sets

$$\tau_2 = \inf\{t \geq \tau_1 \mid B_t \notin [d(X_1), u(X_1)]\},$$

and by Example 3.5 and Theorem 1.30 we have $X_2 = B_{\tau_2}$ in distribution as well as $\mathbb{E}[X_2^2] = \mathbb{E}[\tau_2]$ from Theorem 3.4. Inductively one can prove the existence of a sequence $(\tau_n)_n$ of stopping times with $X_n = B_{\tau_n}$ in distribution and $\mathbb{E}[X_n^2] = \mathbb{E}[\tau_n]$.

Summarizing, the sequence $(\tau_n)_n$ of stopping times converges increasingly to a random variable τ. Because of

$$\{\tau \leq t\} = \bigcap_{n\geq 1} \{\tau_n \leq t\}$$

we know that τ itself is a stopping time, and monotone convergence in connection with the L^2-convergence of X_n to X imply

$$\mathbb{E}[\tau] = \lim_{n\to\infty} \mathbb{E}[\tau_n] = \lim_{n\to\infty} \mathbb{E}[X_n^2] = \mathbb{E}[X^2] < \infty.$$

Finally we also have $B_\tau = X$ in distribution, because the sequence $(B_{\tau_n})_n$ is distributed as $(X_n)_n$ and therefore converges in distribution to X on one hand, and by continuity of the sample paths of B it converges almost surely to B_τ on the other hand because τ is almost surely finite. $\square$

Definition 3.10. Let $(X_n)_{n\geq 1}$ be a sequence of i.i.d. random variables with $\mathbb{E}[X_i] = 0$ and $\mathrm{Var}(X_i) = 1$. The sequence $(S_n)_{n\geq 0}$ with $S_0 = 0$ and

$$S_n = \sum_{i=1}^{n} X_i$$

is called a *random walk*, and via

$$t \mapsto S(t) = S_{\lfloor t \rfloor} + (t - \lfloor t \rfloor)(S_{\lfloor t \rfloor + 1} - S_{\lfloor t \rfloor}), \quad t \geq 0,$$

where $\lfloor t \rfloor$ denotes the floor function, and

$$t \mapsto S_n^*(t) = \frac{S(nt)}{\sqrt{n}}, \quad t \in [0, 1],$$

two new stochastic processes can be defined on $C[0, \infty]$ and $C[0, 1]$, respectively.

Remark 3.11. For every fixed $t \in [0,1]$ we know from the central limit theorem that the convergence in distribution

$$S_n^*(t) \xrightarrow{\mathcal{L}} \mathcal{N}(0,t) \sim B(t)$$

holds. With a little additional work one can also prove convergence in distribution of the finite-dimensional distributions, i.e.

$$(S_n^*(t_1), \ldots, S_n^*(t_\ell))^T \xrightarrow{\mathcal{L}} (B(t_1), \ldots, B(t_\ell))^T$$

for any fixed $0 \leq t_1 \leq \ldots \leq t_\ell \leq 1$. Donsker's invariance principle now deals with an extension towards a *functional* central limit theorem.

Definition 3.12. Let (S,d) be a metric space.

(i) A sequence (μ_n) of measures on S (with its Borel-σ-field) *converges weakly* to a measure μ on S if

$$\int f d\mu_n \to \int f d\mu$$

for all bounded and continuous functions $f : S \to \mathbb{R}$. The notation is: $\mu_n \xrightarrow{w} \mu$.

(ii) A sequence (X_n) of random variables on $(\Omega, \mathcal{F}, \mathbb{P})$ with values in S *converges in distribution* to X if the induced distributions satisfy $\mathbb{P}^{X_n} \xrightarrow{w} \mathbb{P}^X$. The notation is: $X_n \xrightarrow{\mathcal{L}} X$.

Theorem 3.13. (Donsker's invariance principle) *The sequence $(S_n^*(t))_{n\geq 1}$ converges on $(C[0,1], ||\cdot||_\infty)$ weakly to a Brownian motion $(B_t)_{t\in[0,1]}$.*

Remark 3.14. Theorem 3.13 is called an invariance principle because (just as for the classical central limit theorem) the convergence in distribution holds independently of the choice of the sequence $(X_n)_{n\geq 1}$, as long as two moment conditions are satisfied.

Before we come to the proof of Theorem 3.13 we will deal with a helpful auxiliary result.

Lemma 3.15. *Let B be a Brownian motion and let X be a random variable with $\mathbb{E}[X] = 0$ and $\mathrm{Var}(X) = 1$. Then there exists an increasing sequence $(\tau_n)_n$ of stopping times with respect to the natural filtration and with $\tau_0 = 0$ such that*

(i) the sequence $(B(\tau_n))_n$ has the same distribution as the random walk whose increments have the same distribution as X;

(ii) the sequence $(S_n^*)_n$ which can be constructed from the random walk with the increments $B(\tau_n) - B(\tau_{n-1})$ satisfies

$$\lim_{n\to\infty} \mathbb{P}\Big(\sup_{t\in[0,1]} \Big| \frac{B(nt)}{\sqrt{n}} - S_n^*(t) \Big| > \varepsilon \Big) = 0$$

for all $\varepsilon > 0$.

Proof:

(i) From Theorem 3.9 we know that there exists a stopping time τ_1 with $\mathbb{E}[\tau_1] = 1$ and $B(\tau_1) = X$ in distribution. Now Theorem 1.30 gives that

$$\{B_2(t) \mid t \geq 0\} = \{B(\tau_1 + t) - B(\tau_1) \mid t \geq 0\}$$

is again a Brownian motion, independent of $(\tau_1, B(\tau_1))$. As a consequence there exists a stopping time τ_2' with respect to B_2 with $\mathbb{E}[\tau_2'] = 1$ and $B_2(\tau_2') = X$ in distribution. If we now set $\tau_2 = \tau_1 + \tau_2'$, then $B(\tau_2)$ is equal in distribution to the sum of two independent copies of X, and clearly $\mathbb{E}[\tau_2] = 2$. The claim follows by induction.

(ii) Let $W_n(t) = B(nt)/\sqrt{n}$ and

$$A_n = \left\{ \sup_{t\in[0,1]} \Big| W_n(t) - S_n^*(t) \Big| > \varepsilon \right\}$$

for a fixed $\varepsilon > 0$. By the scaling property of the normal distribution it is clear that the process $(W_n(t))_{t\geq 0}$ is again a Brownian motion. Since S_n^* is piecewise linear, we also have

$$\begin{aligned} A_n \subset\, &\{\text{there exists } t \in [0,1) \text{ with } |S_k/\sqrt{n} - W_n(t)| > \varepsilon\} \\ &\cup \{\text{there exists } t \in [0,1) \text{ with } |S_{k-1}/\sqrt{n} - W_n(t)| > \varepsilon\} \\ =\, &\{\text{there exists } t \in [0,1) \text{ with } |W_n(\tau_k/n) - W_n(t)| > \varepsilon\} \\ &\cup \{\text{there exists } t \in [0,1) \text{ with } |W_n(\tau_{k-1}/n) - W_n(t)| > \varepsilon\}, \end{aligned}$$

where $k = k(t)$ denotes the integer with $(k-1)/n \leq t < k/n$ and where we have used $S_k = B(\tau_k) = \sqrt{n} W_n(\tau_k/n)$. For every $0 < \delta < 1$ we can then deduce

$$\begin{aligned} A_n \subset\, &\{\text{there exist } s, t \in [0,2] \text{ with } |s-t| < \delta \text{ and } |W_n(s) - W_n(t)| > \varepsilon\} \\ &\cup \{\text{there exists } t \in [0,1) \text{ with } |\tau_k/n - t| \vee |\tau_{k-1}/n - t| \geq \delta\}. \end{aligned}$$

Now let $\eta > 0$ be arbitrary. A Brownian motion has continuous paths, so it is uniformly continuous over $[0,2]$. It follows that

$$\mathbb{P}(\text{there exist } s, t \in [0,2] \text{ with } |s-t| < \delta \text{ and } |W_n(s) - W_n(t)| > \varepsilon)$$

is independent of n, because W_n is distributed as B, and becomes smaller than η for an adequate choice of δ, which can easily be shown by contradiction. Therefore it remains to prove

$$\lim_{n\to\infty} \mathbb{P}(\text{there exists } t \in [0,1) \text{ with } |\tau_k/n - t| \vee |\tau_{k-1}/n - t| > \delta) = 0 \tag{3.3}$$

for every fixed $\delta > 0$ in order to deduce $\limsup_{n\to\infty} \mathbb{P}(A_n) < \eta$. The claim then follows immediately.

The proof of (3.3) utilizes

$$\lim_{n\to\infty} \frac{\tau_n}{n} = \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} (\tau_k - \tau_{k-1}) = 1$$

almost surely, using the strong law of large numbers; compare with the proof of (i). Also we know that for every deterministic sequence $(a_n)_n$ the implication

$$\lim_{n\to\infty} \frac{a_n}{n} = 1 \Longrightarrow \lim_{n\to\infty} \sup_{k\le n} \frac{|a_k - k|}{n} = 0$$

holds. Then, because of $t \in [(k-1)/n, k/n]$, a case distinction yields

$$\begin{aligned} &\mathbb{P}(\text{there exists } t \in [0,1) \text{ with } |\tau_k/n - t| \vee |\tau_{k-1}/n - t| \ge \delta) \\ \le\ &\mathbb{P}\Big(\sup_{k\le n} \frac{(\tau_k - (k-1)) \vee (k - \tau_{k-1})}{n} \ge \delta\Big). \end{aligned}$$

If one finally chooses $n > 2/\delta$ the latter term can be bounded from above by

$$\mathbb{P}\Big(\sup_{k\le n} \frac{|\tau_k - k|}{n} \ge \delta/2\Big) + \mathbb{P}\Big(\sup_{k\le n} \frac{|\tau_{k-1} - (k-1)|}{n} \ge \delta/2\Big).$$

Both terms converge to zero as prepared. $\square$

Lemma 3.16. (Portmanteau theorem) *Let (S,d) be a metric space. Then the following claims are equivalent:*

(i) $X_n \xrightarrow{\mathcal{L}} X$.

(ii) For every closed set $K \subset S$ one has

$$\limsup_{n\to\infty} \mathbb{P}(X_n \in K) \le \mathbb{P}(X \in K).$$

(iii) For every open set $G \subset S$ one has

$$\liminf_{n\to\infty} \mathbb{P}(X_n \in G) \ge \mathbb{P}(X \in G).$$

(iv) For every measurable set $A \subset S$ with $\mathbb{P}(X \in \partial A) = 0$ one has

$$\lim_{n\to\infty} \mathbb{P}(X_n \in A) = \mathbb{P}(X \in A).$$

Proof: see e.g. Theorem 12.6 in Mörters and Peres (2010). $\square$

We will now finally prove Theorem 3.13, and we assume to be in the situation of Lemma 3.15(ii). So, let $K \subset C[0,1]$ be closed, and we define

$$K(\varepsilon) = \{f \in C[0,1] \mid ||f - g||_\infty \le \varepsilon \text{ for some } g \in K\}$$

for every $\varepsilon > 0$. Now let $\eta > 0$ be arbitrary. Using again that $W_n(t) = B(nt)/\sqrt{n}$ defines a Brownian motion, we obtain

$$\lim_{\varepsilon\to 0} \mathbb{P}(W_n \in K(\varepsilon)) = \lim_{\varepsilon\to 0} \mathbb{P}(B \in K(\varepsilon)) = \mathbb{P}\Big(B \in \bigcap_{\varepsilon>0} K(\varepsilon)\Big) = \mathbb{P}(B \in K)$$

using continuity from above and that K is closed. In particular, we can choose $\varepsilon > 0$ in such a way that

$$\limsup_{n\to\infty} \mathbb{P}(W_n \in K(\varepsilon)) \le \mathbb{P}(B \in K) + \eta$$

holds. On the other hand, for this choice of $\varepsilon > 0$ we have the inequality

$$\mathbb{P}(S_n^* \in K) \leq \mathbb{P}(W_n \in K(\varepsilon)) + \mathbb{P}(||S_n^* - W_n||_\infty > \varepsilon).$$

By Lemma 3.15(ii) we can show $\mathbb{P}(||S_n^* - W_n||_\infty > \varepsilon) \to 0$ for $n \to \infty$. Thus

$$\limsup_{n\to\infty} \mathbb{P}(S_n^* \in K) \leq \mathbb{P}(B \in K) + \eta$$

for all $\eta > 0$. The invariance principle now follows from Lemma 3.16(ii). $\square$

Chapter 4

Lebesgue-Stieltjes integrals

In this chapter we will give the definition of Lebesgue-Stieltjes integrals. They provide a first idea how to define integrals with respect to stochastic processes but do not help when one is interested in integrating with respect to a Brownian motion.

Definition 4.1. Let $A = (A_t)_{t\geq 0}$ be a stochastic process with (almost surely) càdlàg paths. A is called *of bounded variation* if for every compact interval $[a, b] \subset [0, \infty)$

$$V^A_{[a,b]}(\omega) = \sup_{\pi\in\mathcal{P}} \sum_{t_i\in\pi} |A_{t_{i+1}} - A_{t_i}|(\omega) < \infty$$

holds $\mathbb{P}$-almost surely where $\mathcal{P}$ runs through all finite partitions of the form $\pi = \{a = t_0 < \ldots < t_k = b\}$, so with an arbitrary $k \in \mathbb{N}$.

Remark 4.2. We call a sequence $(\pi_n)_n$ of partitions *nested* if $t \in \pi_m$ implies $t \in \pi_n$ for any $m \leq n$. For every ω such that $t \mapsto A_t(\omega)$ is right continuous one can then show

$$V^A_{[a,b]}(\omega) = \lim_{n\to\infty} \sum_{t_i\in\pi_n} |A_{t_{i+1}} - A_{t_i}|(\omega)$$

where $(\pi_n)_n$ is any sequence of nested partitions with

$$\delta(\pi_n) = \max\{t_{i+1} - t_i \mid \{t_i, t_{i+1}\} \subset \pi_n\} \to 0. \tag{4.1}$$

Remark 4.3.

(i) Let A be (almost surely) right continuous and increasing, so there exists some Ω_0 with $\mathbb{P}(\Omega_0) = 1$ such that $t \mapsto A_t(\omega)$ is right continuous and increasing for every $\omega \in \Omega_0$. In particular, for every such ω one can define a (random) measure $\mu_A(\omega, ds)$ on $[0, \infty)$ via

$$\mu_A(\omega, (a, b]) = A_b(\omega) - A_a(\omega)$$

and with the usual extension to $\mathcal{B}|_{[0,\infty)}$. If then $f : [0, \infty) \to \mathbb{R}$ is bounded and measurable, then the Lebesgue integral

$$\int_0^t f(s)dA_s(\omega) = \begin{cases} \int_{[0,t]} f(s)\mu_A(\omega, ds), & \text{if } \omega \in \Omega_0, \\ 0, & \text{else,} \end{cases}$$

is well-defined. It is called the *(Lebesgue-)Stieltjes integral of f with respect to A*.

(ii) If A is (almost surely) right continuous and increasing and $F : \Omega \times [0, \infty) \to \mathbb{R}$ is bounded and measurable (with respect to $\mathcal{F} \otimes \mathcal{B}|_{[0,\infty)}$) then one can define analogously

$$Y(\omega, t) = \int_0^t F(\omega, s) dA_s(\omega) = \begin{cases} \int_{[0,t]} F(\omega, s)\mu_A(\omega, ds), & \text{if } \omega \in \Omega_0, \\ 0, & \text{else.} \end{cases}$$

$Y : \Omega \times [0, \infty) \to \mathbb{R}$ is measurable as well and because of the right continuity of A (almost surely) right continuous in t itself.

Definition 4.4. Let A be of bounded variation. Then the process

$$t \mapsto |A|_t(\omega) = V^A_{[0,t]}(\omega)$$

is called the *total variation process of* A.

Remark 4.5. Let $s < t$. Since every partition of $[0, s]$ can be extended to a partition of $[0, t]$ by adding the point t, it is clear that $t \mapsto |A|_t$ is increasing.

Theorem 4.6. *A process A is of bounded variation if and only if it can be written as the difference of two (almost surely) increasing processes.*

Proof:

$\Longleftarrow$ Let $A = A^+ - A^-$ with two increasing processes. Then for every partition π of $[a, b]$ we (almost surely) have

$$\begin{aligned}\sum_{t_i \in \pi} |A_{t_{i+1}} - A_{t_i}| &\le \sum_{t_i \in \pi} |A^+_{t_{i+1}} - A^+_{t_i}| + \sum_{t_i \in \pi} |A^-_{t_{i+1}} - A^-_{t_i}| \\ &= A^+_b - A^+_a + A^-_b - A^-_a < \infty.\end{aligned}$$

$\Longrightarrow$ Clearly $A^+_t = \frac{1}{2}(|A|_t + A_t)$ and $A^-_t = \frac{1}{2}(|A|_t - A_t)$ are well-defined, and we have $A_t = A^+_t - A^-_t$. Now let $s < t$. By definition, and using telescoping sums,

$$(|A|_t - |A|_s) + (A_t - A_s) = \sup_{\pi \in \mathcal{P}} \sum_{u_i \in \pi} \big(|A_{u_{i+1}} - A_{u_i}| + (A_{u_{i+1}} - A_{u_i})\big)$$

where $\mathcal{P}$ runs through all partitions of $[s, t]$. Since all summands are non-negative, A^+ is increasing. The same proof works for A^-. $\square$

Definition 4.7. Let $F : \Omega \times [0, \infty) \to \mathbb{R}$ be (almost surely) bounded and measurable, and let A be of bounded variation and right continuous. Then we call

$$Y(\omega, t) = \int_0^t F(\omega, s) dA_s(\omega) = \int_0^t F(\omega, s) dA^+_s(\omega) - \int_0^t F(\omega, s) dA^-_s(\omega)$$

the (almost surely unique) *(Lebesgue-)Stieltjes integral of F with respect to A.*

Remark 4.8.

(i) One could have defined the Stieltjes integral in an analogous way to Remark 4.3 as well, but

$$\mu_A(\omega, (a, b]) = A_b(\omega) - A_a(\omega)$$

now is in general only a signed measure, i.e. one which can also take negative values.

(ii) If $H : \Omega \times [0,\infty) \to \mathbb{R}$ is measurable and $t \mapsto H(\omega, t)$ is (almost surely) continuous, then the integral $\int_0^t H_s dA_s$ can also be defined as a pointwise Riemann-Stieltjes integral. This means that for every sequence $(\pi_n)_n$ of partitions of $[0,t]$ which satisfies (4.1) the convergence

$$\lim_{n\to\infty} \sum_{t_i \in \pi_n} H_{s_i}(A_{t_{i+1}} - A_{t_i}) = \int_0^t H_s dA_s$$

holds almost surely, with $t_i \leq s_i \leq t_{i+1}$ being arbitrary.

Theorem 4.9. (Change of variables) *Let A be of bounded variation and (almost surely) continuous. If $f : \mathbb{R} \to \mathbb{R}$ is continuously differentiable, then $(f(A_t))_{t\geq 0}$ is of bounded variation as well, and we have*

$$f(A_t) - f(A_0) = \int_0^t f'(A_s) dA_s.$$

Proof: The mapping

$$\begin{aligned} \Omega \times [0,\infty) &\to \mathbb{R} \\ (\omega, s) &\mapsto f'(A_s)(\omega) \end{aligned}$$

can shown to be measurable, and for every ω for which $s \mapsto f'(A_s)(\omega)$ is continuous we have that the mapping is bounded on $[0,t]$ for every $t > 0$. Therefore

$$Y_t = \int_0^t f'(A_s) dA_s$$

exists by Definition 4.7, and Remark 4.8(ii) proves because of the pointwise boundedness of $f'(A_s)$ that Y is of bounded variation as well.

Finally let $(\pi_n)_n$ be a sequence of partitions which satisfies (4.1). Then by the mean value theorem

$$f(A_t) - f(A_0) = \sum_{u_i \in \pi_n} \big(f(A_{u_{i+1}}) - f(A_{u_i})\big) = \sum_{u_i \in \pi_n} f'(A_{s_i})\big(A_{u_{i+1}} - A_{u_i}\big)$$

for some intermediate s_i, pointwise for all ω for which $s \mapsto f'(A_s)(\omega)$ is continuous. The claim then follows from Remark 4.8(ii). □

Theorem 4.10. *Let B be a Brownian motion, $[a,b]$ an interval and $(\pi_n)_n$ a sequence of nested partitions which satisfies (4.1). Then for*

$$\pi_n B = \sum_{t_i \in \pi_n} (B_{t_i} - B_{t_{i-1}})^2$$

we have $\lim_{n\to\infty} \pi_n B = b - a$ almost surely and in L^2.

Proof: Let

$$\pi_n B - (b-a) = \sum_{t_i \in \pi_n} \big((B_{t_{i+1}} - B_{t_i})^2 - (t_{i+1} - t_i)\big) = \sum_{t_i \in \pi_n} U_i,$$

where the U_i are, by definition of Brownian motion, independent and centered. Thus,

$$\mathbb{E}[(\pi_n B - (b-a))^2] = \sum_{t_i \in \pi_n} \mathbb{E}[U_i^2] = \sum_{t_i \in \pi_n} \mathbb{E}[(Z_i^2 - 1)^2](t_{i+1} - t_i)^2,$$

where we have set

$$Z_i = \frac{B_{t_{i+1}} - B_{t_i}}{\sqrt{t_{i+1} - t_i}} \sim \mathcal{N}(0,1).$$

In particular, $\mathbb{E}[(Z_i^2 - 1)^2] = 2$ holds independently of i. An argument using telescoping sums then gives

$$\begin{aligned}\mathbb{E}[(\pi_n B - (b-a))^2] &= 2 \sum_{t_i \in \pi_n} (t_{i+1} - t_i)^2 \\ &\le 2 \max\{t_{i+1} - t_i \mid \{t_i, t_{i+1}\} \subset \pi_n\}(b-a) \to 0\end{aligned}$$

by (4.1). Thus convergence in L^2 is shown.

Almost sure convergence follows by proving first (as an exercise) that

$$N_n = \sum_{t_i \in \pi_{-n}} (B_{t_{i+1}} - B_{t_i})^2, \quad n = -1, -2, \ldots,$$

is a backwards martingale with respect to $\mathcal{G}_n = \sigma(N_k \mid k \le n)$. Then $\pi_n B = N_{-n}$ converges almost surely by Theorem 2.18, and the limit equals the limit in L^2. $\square$

Corollary 4.11. *Let B be a Brownian motion. Then, for almost all ω, the paths $t \mapsto B_t(\omega)$ are of unbounded variation over every interval $[a,b]$ with $a < b$.*

Proof: Let $a < b$ be arbitrary. Using Theorem 4.10 we have for every sequence of nested partitions satisfying (4.1) that

$$b - a = \lim_{n\to\infty} \sum_{t_i \in \pi_n} (B_{t_{i+1}} - B_{t_i})^2 \le \lim_{n\to\infty} \sup_{t_i \in \pi_n} |B_{t_{i+1}} - B_{t_i}| \sum_{t_i \in \pi_n} |B_{t_{i+1}} - B_{t_i}|$$

holds almost surely. Now, by uniform continuity of Brownian motion over $[a,b]$,

$$\lim_{n\to\infty} \sup_{t_i \in \pi_n} |B_{t_{i+1}} - B_{t_i}| = 0$$

almost surely. On the other hand,

$$V_{[a,b]}^B = \lim_{n\to\infty} \sum_{t_i \in \pi_n} |B_{t_{i+1}} - B_{t_i}|$$

from Remark 4.2. Thus $V_{[a,b]}^B = \infty$ almost surely, and the claim follows since it is sufficient to prove the statement for the countably many intervals with rational end points. $\square$

Remark 4.12. The integral for processes of bounded variation was introduced in Remark 4.3 and Definition 4.7 in a purely analytical way, i.e. we have used results from calculus, pointwise in ω, which grant the existence of

$$t \mapsto Y(\omega, t) = \int_0^t F(\omega, s) dA_s(\omega).$$

Even though the processes F and A are in general random, the actual notion of a stochastic integral thus remains analytical. Such an approach is in general not possible anymore if we are interested in integrals with respect to martingales which (as seen for Brownian motion) are in general not of bounded variation.

Chapter 5

Integration with respect to L^2 martingales

We will introduce a novel definition of a stochastic integral which works for martingales in L^2, and thus for Brownian motion as well.

Definition 5.1.

(i) Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ be a filtered probability space. We call

$$\mathcal{R} = \{F \times \{0\} \mid F \in \mathcal{F}_0\} \cup \{F \times (s,t] \mid F \in \mathcal{F}_s, s \leq t\}$$

and

$$\mathcal{P} = \sigma(\mathcal{R})$$

the *system of the predictable rectangles* and the *σ-field of the predictable sets* on $\Omega \times [0,\infty)$, respectively.

(ii) A $\mathcal{P}$-measurable process $(H_t)_{t\geq 0}$ is called *predictable*. In particular, we call

$$\mathcal{E} = \Big\{ \sum_{j=1}^{n} a_j 1_{R_j} \mid n \in \mathbb{N}, a_1, \ldots, a_n \in \mathbb{R}, R_1, \ldots, R_n \in \mathcal{R}, R_i \cap R_j = \emptyset \text{ for } i \neq j \Big\}$$

the *set of the simple predictable processes.*

Remark 5.2. By construction, predictable processes are those for which, at every point in time, the value in the near future is already known. They are therefore prime candidates to serve as reasonable models for trading strategies.

Theorem 5.3.

(i) Predictable processes are adapted and progressively measurable.

(ii) Adapted left continuous processes are predictable.

Proof:

(i) Only progressive measurability is to be shown, and by the usual measure-theoretic (or algebraic) induction it is sufficient to prove the claim for processes of the form $H = 1_P$ with $P \in \mathcal{P}$. Also, it is straight forward to prove that

$$\mathcal{D} = \{C \in \mathcal{F} \otimes \mathcal{B}_{[0,\infty)} \mid 1_C \text{ is progressively measurable}\}$$

is a Dynkin system.

Now let $F \in \mathcal{F}_s$ and $G \in \mathcal{F}_u$. Because of

$$(F \times (s,t]) \cap (G \times (u,v]) = (F \cap G) \times (s \vee u, t \wedge v]$$

and $F \cap G \in \mathcal{F}_{s\vee u}$ we can conclude that $\mathcal{R}$ is stable under intersection. Since 1_R is progressively measurable for every $R \in \mathcal{R}$, which we have already used in the proof of Lemma 2.27, it follows from the usual π-λ theorem that

$$\mathcal{P} = \sigma(\mathcal{R}) \subset \mathcal{D}.$$

(ii) A left continuous process H can be written as the pointwise limit of the sequence

$$H_s^n = H_0 1_{\{0\}}(s) + \sum_{k=1}^{2^n} H_{\frac{(k-1)t}{2^n}} 1_{\left(\frac{(k-1)t}{2^n}, \frac{kt}{2^n}\right]}(s),$$

and these processes are $\mathcal{P}$-measurable by construction. □

Definition 5.4. Let σ and τ be stopping times. Then

$$[\sigma,\tau] = \{(\omega,t) \mid \sigma(\omega) \le t \le \tau(\omega)\} \subset \Omega \times [0,\infty)$$

denotes the closed random interval between σ and τ. The definition of (σ,τ), $(\sigma,\tau]$ and $[\sigma,\tau)$ works analogously.

Theorem 5.5. *For stopping times σ and τ both $[0,\tau]$ and $(\sigma,\tau]$ are elements of $\mathcal{P}$.*

Proof: Because of $(\sigma,\tau] = [0,\tau]\backslash[0,\sigma]$ it is sufficient to prove that $[0,\tau]$ is in $\mathcal{P}$. Here we choose the usual approximation by

$$\tau_n = (m+1)2^{-n}, \quad \text{if } m2^{-n} \le \tau < (m+1)2^{-n},\ m \in \mathbb{N}_0,$$

so that every τ_n is a stopping time and $\tau_n \searrow \tau$ holds. Because of $[0,\tau] = \bigcap_{n\in\mathbb{N}}[0,\tau_n]$ and

$$[0,\tau_n] = (\Omega \times \{0\}) \cup \left(\bigcup_{m\in\mathbb{N}_0} \left(\{\omega \in \Omega \mid \tau(\omega) \ge \frac{m}{2^n}\} \times (\frac{m}{2^n}, \frac{m+1}{2^n}] \right) \right) \in \mathcal{P}$$

the claim follows. □

Definition 5.6. Let $X = (X_t)_{t\ge 0}$ be a process with $\mathbb{E}[|X_t|] < \infty$ for all $t \ge 0$. Then we define $\nu_X : \mathcal{R} \to \mathbb{R}$ via $\nu_X(F \times \{0\}) = 0$, $F \in \mathcal{F}_0$, and

$$\nu_X(F \times (s,t]) = \int_F (X_t - X_s) d\mathbb{P}, \quad F \in \mathcal{F}_s,\ s \le t.$$

Remark 5.7.

(i) For a martingale X it is clear that $\nu_X = 0$, since for every $s \le t$ and every $F \in \mathcal{F}_s$

$$\int_F (X_t - X_s) d\mathbb{P} = \int_F (X_s - X_s) d\mathbb{P} = 0$$

holds by definition of the conditional expectation.

(ii) Analogously it can be shown that we have $\nu_X \geq 0$ for submartingales.

Theorem 5.8. *Let X be an L^2 martingale. Then*

$$\nu_{X^2}(F \times (s,t]) = \int_F (X_t - X_s)^2 d\mathbb{P}, \quad F \in \mathcal{F}_s, \ s \leq t.$$

Proof: We have

$$\begin{aligned}
\int_F (X_t - X_s)^2 d\mathbb{P} &= \int_F X_t^2 d\mathbb{P} - 2\mathbb{E}[1_F X_s X_t] + \int_F X_s^2 d\mathbb{P} \\
&= \int_F X_t^2 d\mathbb{P} - 2\mathbb{E}[1_F X_s \mathbb{E}[X_t|\mathcal{F}_s]] + \int_F X_s^2 d\mathbb{P} \\
&= \int_F X_t^2 d\mathbb{P} - \int_F X_s^2 d\mathbb{P}
\end{aligned}$$

using iterated expectation and the definition of a martingale. □

Definition 5.9. Let X be a stochastic process.

(i) For $R \in \mathcal{R}$ we set

$$\int 1_R dX = \begin{cases} 0, & R = F \times \{0\} \text{ for some } F \in \mathcal{F}_0, \\ 1_F(X_t - X_s), & R = F \times (s,t] \text{ for some } F \in \mathcal{F}_s. \end{cases}$$

(ii) For $H \in \mathcal{E}$ of the form

$$H = \sum_{j=1}^n a_j 1_{R_j}$$

we define

$$\int H dX = \sum_{j=1}^n a_j \int 1_{R_j} dX.$$

Remark 5.10.

(i) The mapping

$$\begin{aligned} I : \mathcal{E} &\to \mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P}) \\ H &\mapsto \int H dX \end{aligned}$$

is well-defined and linear. Both properties hold since $\mathcal{R}$ is stable under intersection, because then two processes H_1 and H_2 from $\mathcal{E}$ can be written as

$$H_1 = \sum_{j=1}^n a_j 1_{R_j} \quad \text{and} \quad H_2 = \sum_{j=1}^n b_j 1_{R_j}$$

with respect to a joint decomposition.

(ii) Our goal in the following is to extend the notion of a stochastic integral to more general processes H. Clearly it seems to be reasonable to work with predictable integrands, at least in a first step. On the other hand the class of integrators will be restricted to L^2 martingales.

(iii) In contrast to Stieltjes integrals it will not be possible to allow all bounded and $\mathcal{F} \otimes \mathcal{B}_{[0,\infty)}$-measurables processes H as integrands.

Theorem 5.11. *Let $X \geq 0$ be a (right continuous) submartingale. Then there exists a unique extension of ν_X to a measure on $\mathcal{P}$.*

Proof: We will start with showing the existence of the measure for which we define

$$\widetilde{\mathcal{R}} = \big\{ \bigcup_{j=1}^{n} R_j \mid n \in \mathbb{N}, R_1, \dots, R_n \in \mathcal{R}, R_i \cap R_j = \emptyset \text{ for all } i \neq j \big\}$$

and

$$\begin{aligned} \nu_X : \widetilde{\mathcal{R}} &\to [0, \infty) \\ \bigcup_{j=1}^{n} R_j &\mapsto \sum_{j=1}^{n} \nu_X(R_j). \end{aligned}$$

First, one needs to check that $\widetilde{\mathcal{R}}$ is a ring which by definition of $\widetilde{\mathcal{R}}$ amounts to proving that $A, B \in \mathcal{R}$ implies $A \backslash B \in \mathcal{R}$ and $A \cup B \in \mathcal{R}$. The proof is not difficult but involves several case distinctions, and it is left as an exercise. Also, one can show similarly to Remark 5.10(i) that ν_X is a well-defined content.

Since $\mathcal{P} = \sigma(\widetilde{\mathcal{R}})$ holds clearly as well, the only thing which remains to be proven in order to be able to apply Carathéodory's extension theorem is to show that ν_X defines a pre-measure on $\widetilde{\mathcal{R}}$ as well. An equivalent statement goes as follows: For every sequence $(A_n)_n \in \widetilde{\mathcal{R}}$ with $A_n \searrow \emptyset$ we have $\nu_X(A_n) \to 0$. By definition of $\widetilde{\mathcal{R}}$ and as the sequence is decreasing, we have $A_n \subset \Omega \times (0, T)$ for some $T > 0$ and all n.

We show first that for every $\varepsilon > 0$ there exist $B_n \in \widetilde{\mathcal{R}}$ and $C_n \subset \Omega \times (0, T)$ such that for all $n \in \mathbb{N}$ we have $C_n(\omega) \subset (0, T)$ being compact for all $\omega \in \Omega$, $B_n \subset C_n \subset A_n$ and

$$\nu_X(A_n \backslash B_n) \leq \varepsilon 2^{-n}. \tag{5.1}$$

To this end let $R = F \times (s, t] \in \mathcal{R}$ be arbitrary, and we set

$$R'_m = F \times \big(s + \frac{1}{m}, t\big] \quad \text{and} \quad R''_m = F \times \big[s + \frac{1}{m}, t\big]$$

for $m \in \mathbb{N}$, with the obvious convention of $(a, b]$ being empty for $a \geq b$. For $s + 1/m < t$ we have

$$0 \leq \nu_X(R \backslash R'_m) = \nu_X\big(F \times \big(s, s + \frac{1}{m}\big]\big) = \int_F (X_{s+1/m} - X_s) d\mathbb{P},$$

and by right continuity $X_{s+\frac{1}{m}} \to X_s$ almost surely. From

$$0 \leq X_{s+\frac{1}{m}} \leq \mathbb{E}[X_{s+1} | \mathcal{F}_{s+\frac{1}{m}}]$$

and Remark 2.16(i) it is clear that $(X_{s+\frac{1}{m}} - X_s)_m$ is uniformly integrable as well. This gives convergence in L^1, so $\nu_X(R \backslash R'_m) \to 0$. (5.1) then follows easily because A_n is a finite union of sets from $\mathcal{R}$. If we now define

$$\widehat{B}_n = \bigcap_{k \leq n} B_k, \quad \widehat{C}_n = \bigcap_{k \leq n} C_k$$

then clearly $\widehat{B}_n \subset \widehat{C}_n \subset A_n$ by monotonicity, and then

$$\nu_X\big(A_n \backslash \widehat{B}_n\big) \leq \nu_X\Big(\bigcup_{k \leq n} (A_k \backslash B_k)\Big) \leq \sum_{k \leq n} \nu_X(A_k \backslash B_k) \leq \varepsilon$$

by (5.1). It remains to prove

$$\nu_X(\widehat{B}_n) \to 0. \tag{5.2}$$

We use $\widehat{C}_n(\omega) \searrow \emptyset$ first. Since every $\widehat{C}_n(\omega)$ is compact it can be shown that there exists, for every $\omega \in \Omega$, some $n_0 = n_0(\omega) \in \mathbb{N}$ with $\widehat{C}_{n_0}(\omega) = \emptyset$. Furthermore,

$$\tau_n = \inf\{t \geq 0 \mid (\omega, t) \in \widehat{B}_n\}, \quad \inf \emptyset = \infty,$$

defines a sequence of stopping times, and then $\widehat{B}_n \subset \widehat{C}_n$ gives $\tau_n \to \infty$. In particular $X_{\tau_n \wedge T} \to X_T$ almost surely, and again

$$0 \leq X_{\tau_n \wedge T} \leq \mathbb{E}[X_T | \mathcal{F}_{\tau_n \wedge T}]$$

implies uniform integrability and $X_{\tau_n \wedge T} \to X_T$ in L^1. Also $\widehat{B}_n \subset (\tau_n \wedge T, T]$.

The last step in the proof of the existence is to use $\widehat{B}_n \in \widetilde{\mathcal{R}}$ which follows in the same way as the proof of the ring property of $\widetilde{\mathcal{R}}$. For $\widehat{B}_n = F_n \times (u_n, v_n]$ we have $\tau_n \in \{u_n, \infty\}$, and as $\widetilde{\mathcal{R}}$ contains only finite unions of such sets it is clear that every τ_n and every $\tau_n \wedge T$ only takes values from a finite set D_n. Then (5.2) follows from

$$\begin{aligned}
\nu_X(\widehat{B}_n) \leq \nu_X((\tau_n \wedge T, T]) &= \nu_X\Big(\sum_{s \in D_n} \{\omega : \tau_n(\omega) \wedge T = s\} \times (s, T]\Big) \\
&= \sum_{s \in D_n} \nu_X(\{\omega : \tau_n(\omega) \wedge T = s\} \times (s, T]) \\
&= \sum_{s \in D_n} \mathbb{E}[1_{\{\omega : \tau_n(\omega) \wedge T = s\}}(X_T - X_s)] = \mathbb{E}[X_T - X_{\tau_n \wedge T}] \to 0,
\end{aligned}$$

where we have used that ν_X is a content on $\widetilde{\mathcal{R}}$.

Finally uniqueness follows by the uniqueness theorem for measures, because $\mathcal{R}$ is stable under intersection, ν_X is σ-finite over $\mathcal{R}$, and $\mathcal{P} = \sigma(\mathcal{R})$. $\square$

Definition 5.12. Let X be a (right continuous) L^2 martingale. The unique measure

$$\mu_X : \mathcal{P} \to [0, \infty) \text{ with } \mu_X = \nu_{X^2} \text{ on } \mathcal{R}$$

is called the *Doléans measure*.

Remark 5.13. For a Brownian motion B we have $\mu_B = \mathbb{P} \otimes \lambda$ (restricted to $\mathcal{P}$) because

$$\mu_B(F \times (s, t]) = \int_F (B_t - B_s)^2 d\mathbb{P} = \mathbb{P}(F)(t - s),$$

using Theorem 5.8 and Theorem 1.12, with λ being the Lebesgue measure on $\mathbb{R}$.

Lemma 5.14. *Let E and F be Banach spaces, $\mathcal{E} \subset E$ dense and $J : \mathcal{E} \to F$ be a linear isometry. Then there exists a unique linear isometry $I : E \to F$ with $I|_{\mathcal{E}} = J$.*

Proof: Let $e \in E$ and $(e_n)_n$ a sequence in $\mathcal{E}$ with $e_n \to e$. Then

$$||J(e_n) - J(e_m)|| = ||J(e_n - e_m)|| = ||e_n - e_m||,$$

and $(J(e_n))_n$ is a Cauchy sequence in F. In particular it converges to some $f \in F$. We then set $I(e) = f$.

Clearly this mapping is well-defined, because for every other sequence $(e'_n)_n$ which converges to e

$$||J(e_n) - J(e'_n)|| = ||e_n - e'_n|| \to 0.$$

By continuity of the norm it also follows that I is a linear isometry. □

Theorem 5.15. *Let X be a (right continuous) L^2 martingale. Then there exists a unique linear isometry*

$$I : \mathcal{L}^2(\Omega \times [0,\infty), \mathcal{P}, \mu_X) \to \mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P})$$

with

$$I(H) = \int H dX \text{ for all } H \in \mathcal{E}.$$

We will set in the following $\mathcal{L}^2(X) = \mathcal{L}^2(\Omega \times [0,\infty), \mathcal{P}, \mu_X)$ and also write $\int H dX$ instead of $I(H)$ for $H \in \mathcal{L}^2(X)$.

Proof: In order to use Lemma 5.14 we have to prove that the set $\mathcal{E}$ of all simple predictable processes is dense in $\mathcal{L}^2(X)$ and that $I(H) = \int H dX$ for $H \in \mathcal{E}$ defines an isometry. Linearity of the latter mapping was already discussed in Remark 5.10(i).

In order to prove that $\mathcal{E}$ lies dense we use $\mathcal{P} = \sigma(\widetilde{\mathcal{R}})$ as in the proof of Theorem 5.11. Using Theorem 1.65 in Klenke (2006) it follows that for every $A \in \mathcal{P}$ mit $\mu_X(A) < \infty$ and every $\varepsilon > 0$ there exists some $B \in \widetilde{\mathcal{R}}$ such that

$$\mu_X(A \Delta B) < \varepsilon.$$

Because of

$$\mu_X(A \Delta B) = \int (1_A - 1_B)^2 d\mu_X$$

it follows that $\{1_A \mid A \in \mathcal{P},\ \mu_X(A) < \infty\} \subset \overline{\mathcal{E}}$, and this property directly translates to simple processes of the form

$$H = \sum_{j=1}^{m} \alpha_j 1_{P_j},\ P_j \in \mathcal{P},\ \mu_X(P_j) < \infty.$$

For a non-negative $H \in \mathcal{L}^2(X)$ now let $\varepsilon > 0$ and $(H_n)_n$ be a sequence of simple processes with $H_n \nearrow H$ pointwise, and let further be $(K_n)_n$ a sequence from $\mathcal{E}$ with

$$\int (H_n - K_n)^2 d\mu_X < \varepsilon$$

for all $n \in \mathbb{N}$. Dominated convergence now yields

$$\limsup_{n\to\infty} \int (H - K_n)^2 d\mu_X \le \limsup_{n\to\infty} 2\Big(\int (H - H_n)^2 d\mu_X + \int (H_n - K_n)^2 d\mu_X\Big) \le \varepsilon.$$

Density of $\mathcal{E}$ finally follows with $H = H^+ - H^-$.

To finish the proof we have to show that

$$\mathbb{E}\Big[\Big(\int HdX\Big)^2\Big] = \int H^2 d\mu_X$$

holds for every $H \in \mathcal{E}$. To this end let without loss of generality

$$H = \sum_{j=0}^{n} a_j 1_{R_j}$$

for pairwise disjoint sets $R_j \in \mathcal{R}_j$, where $R_0 = F_0 \times \{0\}$ and $R_j = F_j \times (s_j, t_j]$ for $F_0 \in \mathcal{F}_0$ and $F_j \in \mathcal{F}_{s_j}$. Obviously $R_j \cap R_k = \emptyset$ means that

$$F_j \cap F_k = \emptyset \quad \text{or} \quad (s_j, t_j] \cap (s_k, t_k] = \emptyset$$

holds. In the second case we conclude, without restriction by assuming $t_j \leq s_k$,

$$\mathbb{E}[1_{F_j \cap F_k}(X_{t_j} - X_{s_j})(X_{t_k} - X_{s_k})] = \mathbb{E}[1_{F_j \cap F_k}(X_{t_j} - X_{s_j})\mathbb{E}[X_{t_k} - X_{s_k}|\mathcal{F}_{s_k}]] = 0,$$

since X is a martingale. This property trivially also holds in the first case. If we then set $s_0 = t_0 = 0$, we obtain

$$\begin{aligned}\mathbb{E}\Big[\Big(\int HdX\Big)^2\Big] &= \mathbb{E}\Big[\Big(\sum_{j=0}^{n} a_j 1_{F_j}(X_{t_j} - X_{s_j})\Big)^2\Big] = \sum_{j=0}^{n} a_j^2 \mathbb{E}[1_{F_j}(X_{t_j} - X_{s_j})^2] \\ &= \sum_{j=0}^{n} a_j^2 \mu_X(F_j \times (s_j, t_j]) = \int H^2 d\mu_X\end{aligned}$$

using Theorem 5.8. □

Remark 5.16. For stopping times σ and τ and $H \in \mathcal{L}^2(X)$ we have by Theorem 5.5 that

$$H1_{[0,\tau]} \in \mathcal{L}^2(X) \quad \text{and} \quad H1_{(\sigma,\tau]} \in \mathcal{L}^2(X)$$

holds as well. In particular, this applies for deterministic times $\sigma = s$ and $\tau = t$.

Theorem 5.17. *Let X be a (right continuous) L^2 martingale and $H \in \mathcal{L}^2(X)$. Then:*

(i) $\mathbb{E}\big[\int HdX \big| \mathcal{F}_t\big] = \int 1_{[0,t]} HdX$.

(ii) The process $(Y_t)_{t \geq 0}$ with

$$Y_t = \int 1_{[0,t]} HdX$$

is an L^2 martingale.

Proof: Using Example 2.4(iii) it is sufficient to prove the first claim. So let first

$$H = \sum_{j=0}^{n} a_j 1_{R_j} \in \mathcal{E},$$

with $R_0 = F_0 \times \{0\}$ and $R_j = F_j \times (s_j, t_j]$ for $F_0 \in \mathcal{F}_0$ and $F_j \in \mathcal{F}_{s_j}$. A case distinction gives

$$\begin{aligned}\mathbb{E}\Big[\int H dX \Big| \mathcal{F}_t\Big] &= \sum_{j=0}^{n} a_j \mathbb{E}[1_{F_j}(X_{t_j} - X_{s_j})|\mathcal{F}_t] \\ &= \sum_{j=0}^{n} a_j 1_{F_j}(X_{t_j \wedge t} - X_{s_j \wedge t}) = \int H 1_{[0,t]} dX.\end{aligned}$$

For a general $H \in \mathcal{L}^2(X)$ let $(H^n)_n$ be a sequence in $\mathcal{E}$ with $H^n \to H$ in $\mathcal{L}^2(X)$. In particular, we also have $H^n 1_{[0,t]} \to H 1_{[0,t]}$ in $\mathcal{L}^2(X)$. By the isometry property of the stochastic integral we can conclude

$$\begin{aligned}Y^n = \int H^n dX &\to \int H dX, \\ Y_t^n = \int H^n 1_{[0,t]} dX &\to \int H 1_{[0,t]} dX,\end{aligned}$$

both as convergence in L^2.

Now, on one hand, we have by preparation for $H_n \in \mathcal{E}$ the convergence in $\mathcal{L}^2$

$$\mathbb{E}[Y^n|\mathcal{F}_t] = Y_t^n \to \int H 1_{[0,t]} dX.$$

On the other hand,

$$\begin{aligned}\mathbb{E}\Big[\Big(\mathbb{E}[Y^n|\mathcal{F}_t] - \mathbb{E}\Big[\int H dX \Big| \mathcal{F}_t\Big]\Big)^2\Big] &\leq \mathbb{E}\Big[\mathbb{E}\Big[\Big(Y^n - \int H dX\Big)^2 \Big| \mathcal{F}_t\Big]\Big] \\ &= \mathbb{E}\Big[\Big(Y^n - \int H dX\Big)^2\Big] \to 0\end{aligned}$$

using the Jensen inequality. The claim then follows by almost sure uniqueness of the L^2 limit. □

Remark 5.18. From Theorem 2.23 we know that not only X is right continuous, but the martingale Y with

$$Y_t = \int 1_{[0,t]} H dX$$

is as well. It is called the *stochastic integral of H with respect to X*, and we write

$$Y_t = \int_0^t H dX.$$

Theorem 5.19. *Let X be a continuous L^2 martingale. Then Y with*

$$Y_t = \int_0^t H dX$$

is continuous as well.

Proof: The claim is clear for $H \in \mathcal{E}$ by

$$\int H 1_{[0,t]} dX = \sum_{j=0}^{n} a_j 1_{F_j}(X_{t_j \wedge t} - X_{s_j \wedge t}),$$

where we have used the notation from Theorem 5.17. Therefore, let $H \in \mathcal{L}^2(X)$, and for the approximating sequences we again have

$$Y^n = \int H^n dX \to \int H dX,$$
$$Y^n_t = \int H^n 1_{[0,t]} dX \to \int H 1_{[0,t]} dX.$$

Obviously $(Y^m_t - Y^n_t)_{t\geq 0}$ is a continuous L^2 martingale as well. Then monotone convergence, Theorem 2.32 and Theorem 2.29 imply

$$\begin{aligned}\mathbb{P}\big(\sup_{t\geq 0}|Y^m_t - Y^n_t| \geq 2^{-k}\big) &= \lim_{j\to\infty}\mathbb{P}\big(\sup_{0\leq t\leq j}|Y^m_t - Y^n_t| \geq 2^{-k}\big)\\ &\leq \limsup_{j\to\infty} 2^{2k}\mathbb{E}[|Y^m_j - Y^n_j|^2] \leq 2^{2k}\mathbb{E}[|Y^m - Y^n|^2].\end{aligned}$$

We first assume

$$\mathbb{E}[|Y^{n+1} - Y^n|^2] \leq 2^{-3n} \tag{5.3}$$

for all $n \in \mathbb{N}$. Then

$$\sum_{n=1}^{\infty}\mathbb{P}\big(\sup_{t\geq 0}|Y^{n+1}_t - Y^n_t| \geq 2^{-n}\big) \leq \sum_{n=1}^{\infty} 2^{-n} < \infty,$$

and the Borel-Cantelli lemma gives

$$\mathbb{P}\big(\sup_{t\geq 0}|Y^{n+1}_t - Y^n_t| < 2^{-n} \text{ for almost all } n\big) = 1.$$

Therefore $(Y^n_t(\omega))_{t\geq 0}$ is for almost all ω a Cauchy sequence with respect to the supremum norm. It converges uniformly so that the limit process $(Y_t(\omega))_{t\geq 0}$ is almost surely continuous as well. Because the limit in probability is unique we get

$$Y_t = \int H 1_{[0,t]} dX.$$

Finally note that we can assume (5.3) without restriction because the sequence $(Y^n)_n$ converges in L^2. It is therefore in particular a Cauchy sequence, and so (5.3) holds at least along a subsequence. Along this subsequence the previous reasoning can be reproduced. □

Lemma 5.20. *Let X be a (right continuous) L^2 martingale, and let further $H \in \mathcal{L}^2(X)$, $s < t$ and Z being bounded and $\mathcal{F}_s$-measurable. Then*

$$\int Z 1_{(s,t]} H dX = Z \int 1_{(s,t]} H dX.$$

Proof: For $Z = 1_F$ with $F \in \mathcal{F}_s$ and $H \in \mathcal{E}$ the claim holds by definition. In the general case we use measure-theoretic (or algebraic) induction plus density of $\mathcal{E}$ in $\mathcal{L}^2(X)$. □

Theorem 5.21. *Let X be a (right continuous) L^2 martingale, let $H \in \mathcal{L}^2(X)$ and let Y be the L^2 martingale with*

$$Y_t = \int_0^t H dX.$$

(i) *We have*

$$\frac{d\mu_Y}{d\mu_X} = H^2, \text{ i.e. } \mu_Y(A) = \int_A H^2 d\mu_X \text{ for all } A \in \mathcal{P}.$$

(ii) *If $K \in \mathcal{L}^2(Y)$ then $KH \in \mathcal{L}^2(X)$ as well, and we have*

$$Z_t := \int_0^t K dY = \int_0^t KH dX.$$

Proof:

(i) By measure uniqueness, and since $\mathcal{R}$ is stable under intersection and generates $\mathcal{P}$, it is sufficient to prove the claim for $A = F \times (s,t]$ with $F \in \mathcal{F}_s$. Then

$$\begin{aligned}\mu_Y(A) &= \mathbb{E}[1_F(Y_t - Y_s)^2] = \mathbb{E}\Big[1_F\Big(\int 1_{(s,t]} H dX\Big)^2\Big] \\ &= \mathbb{E}\Big[\Big(\int 1_{F\times(s,t]} H dX\Big)^2\Big] = \int 1_{F\times(s,t]} H^2 d\mu_X\end{aligned}$$

where we have used Theorem 5.8, Lemma 5.20 and the isometry property of the stochastic integral.

(ii) We have

$$\int K^2 d\mu_Y = \int K^2 H^2 d\mu_X$$

from (i), so the first claim follows immediately.

For the second claim, by Theorem 5.17 only

$$\int K dY = \int KH dX$$

needs to be shown. To this end let first

$$K = 1_{F\times(s,t]} \text{ for some } F \in \mathcal{F}_s.$$

Then

$$\begin{aligned}\int K dY &= 1_F(Y_t - Y_s) = 1_F \int 1_{(s,t]} H dX \\ &= \int 1_{F\times(s,t]} H dX = \int KH dX\end{aligned}$$

where we have used Lemma 5.20 again. Linearity of the integral gives the claim already for all $K \in \mathcal{E}$.

So let finally be $K \in \mathcal{L}^2(Y)$ and let $(K_n)_n$ be a sequence in $\mathcal{E}$ with $K_n \to K$ in $\mathcal{L}^2(Y)$. The isometry property gives in particular $\int K_n dY \to \int K dY$ in L^2. On the other hand, from $K_n \to K$ in $\mathcal{L}^2(Y)$ and (i) we also obtain $K_n H \to KH$ in $\mathcal{L}^2(X)$. In particular,

$$\int K_n H dX \to \int KH dX.$$

Because $K_n \in \mathcal{E}$ already implies

$$\int K_n dY = \int K_n H dX$$

the claim follows. $\square$

Chapter 6

Localisation

In the following chapter we will extend the class of integrators to more than just L^2 martingales. This comes with a cost, namely that we have to further restrict the class of possible integrands.

Theorem 6.1. *Let X be a (right continuous) L^2 martingale, $H \in \mathcal{L}^2(X)$, τ an (almost surely) finite stopping time and Y the process with*

$$Y_t = \int_0^t H dX.$$

Then

$$Y_\tau = \int 1_{[0,\tau]} H dX.$$

Proof: We first assume that τ is bounded and approximate it as usual via

$$\tau_n = (m+1)2^{-n}, \quad \text{if } m2^{-n} \leq \tau < (m+1)2^{-n},\ m \in \mathbb{N}_0.$$

In this case

$$[0,\tau_n] = (\Omega \times \{0\}) \cup \Big(\bigcup_{m \in \mathbb{N}_0} \Big(\{\omega \in \Omega \mid \tau(\omega) \geq \frac{m}{2^n}\} \times (\frac{m}{2^n}, \frac{m+1}{2^n}] \Big) \Big),$$

as we have already seen in the proof of Theorem 5.5. $Y_0 = 0$ then gives

$$\begin{aligned} Y_{\tau_n} &= \sum_{m=0}^{\infty} 1_{\{\frac{m}{2^n} \leq \tau < \frac{m+1}{2^n}\}} Y_{\frac{m+1}{2^n}} = \sum_{m=0}^{\infty} 1_{\{\frac{m}{2^n} \leq \tau\}} (Y_{\frac{m+1}{2^n}} - Y_{\frac{m}{2^n}}) \\ &= \sum_{m=0}^{\infty} 1_{\{\frac{m}{2^n} \leq \tau\}} \int 1_{(\frac{m}{2^n}, \frac{m+1}{2^n}]} H dX. \end{aligned}$$

As τ is bounded the sum is actually finite. Lemma 5.20 thus yields

$$Y_{\tau_n} = \int \sum_{m=0}^{\infty} 1_{\{\frac{m}{2^n} \leq \tau\}} 1_{(\frac{m}{2^n}, \frac{m+1}{2^n}]} H dX = \int 1_{[0,\tau_n]} H dX.$$

Now, on one hand, $\tau_n \searrow \tau$ almost surely gives

$$Y_{\tau_n} \to Y_\tau$$

almost surely as well because Y is almost surely right continuous as an L^2 martingale. On the other hand, we clearly have

$$1_{[0,\tau_n]}H \to 1_{[0,\tau]}H$$

in $\mathcal{L}^2(X)$ as well (use dominated convergence), and by the isometry property we obtain

$$\int 1_{[0,\tau_n]}HdX \to \int 1_{[0,\tau]}HdX$$

in L^2. This proves the claim for bounded stopping times.

In the unbounded case we use that the property holds for the bounded stopping time $\tau \wedge m$. If we then let $m \to \infty$ the finiteness of τ implies both

$$Y_{\tau\wedge m} \to Y_\tau$$

almost surely (here one uses that one has $\tau(\omega) \wedge m = \tau(\omega)$ for almost all ω if $m = m(\omega)$ is chosen large enough) as well as

$$\int 1_{[0,\tau\wedge m]}HdX \to \int 1_{[0,\tau]}HdX$$

in L^2 by dominated convergence again. □

Corollary 6.2. *Let X be a (right continuous) L^2 martingale and τ be a bounded stopping time. Then*

$$\int 1_{[0,\tau]}dX = X_\tau - X_0.$$

Proof: For $\tau \leq k$ we set $H = 1_{[0,k]}$. Then

$$Y_t = \int_0^t HdX = \int 1_{[0,t\wedge k]}dX = X_{t\wedge k} - X_0,$$

so $Y_\tau = X_\tau - X_0$. Theorem 6.1 then gives the claim because $1_{[0,\tau]}H = 1_{[0,\tau]}$. □

Definition 6.3. Let $\mathcal{C}$ be a class of stochastic processes. We say that a process X is *local of class* $\mathcal{C}$ if there exists a sequence $(\tau_n)_n$ of stopping times with $\tau_n \nearrow \infty$ almost surely such that the *stopped processes*

$$X^{\tau_n} = (X_{\tau_n\wedge t})_{t\geq 0}$$

are elements of $\mathcal{C}$ for all n. $(\tau_n)_n$ is called a *localising sequence.*

Remark 6.4. We usually only deal with classes $\mathcal{C}$ which are *stable under stopping*, i.e. for which $X \in \mathcal{C}$ implies $X^\tau \in \mathcal{C}$ whenever τ is a stopping time.

Typical examples can be shown to be

(i) $\mathcal{C} = \{X \mid X \text{ is a martingale}\}$, giving the class of *local martingales*;

(ii) $\mathcal{C} = \{X \mid X \text{ is an } L^2 \text{ martingale}\}$, giving the class of *local L^2 martingales*;

(iii) $\mathcal{C} = \{X \mid X \text{ is bounded}\}$, giving the class of *locally bounded processes.*

Lemma 6.5. *Let $\mathcal{C}$ be stable under stopping and let X be locally of class $\mathcal{C}$ with a localising sequence $(\tau_n)_n$. If $\sigma_n \nearrow \infty$ is another sequence of stopping times diverging to infinity, then $(\sigma_n \wedge \tau_n)_n$ is a localising sequence as well.*

Proof: We have

$$X_t^{\sigma_n \wedge \tau_n} = X_{\sigma_n \wedge \tau_n \wedge t} = (X^{\tau_n})_t^{\sigma_n}.$$

The claim then follows because X^{τ_n} is of class $\mathcal{C}$ and $\mathcal{C}$ is stable under stopping. □

In the following we assume always that X_0 is bounded.

Theorem 6.6. *Let X be a continuous local martingale. Then X is locally bounded and a local L^2 martingale.*

Proof: We have $|X_0| \leq k$ for an appropriate k, and we let $(\tau_n)_n$ be a localising sequence for X. We set further

$$\sigma_n = \inf\{t \geq 0 \mid |X_t| \geq n\}.$$

As X is continuous we have $\sigma_n \nearrow \infty$, and we clearly also have $|X_t^{\sigma_n}| \leq n + k$ for all $t \geq 0$. It remains to show that X^{σ_n} is a martingale for each n.

As for Theorem 2.9, which is stated for discrete processes only but holds in our situation as well and with the same proof, it is enough to prove

$$\mathbb{E}[X_\sigma^{\sigma_n}] = \mathbb{E}[X_{\sigma_n \wedge \sigma}] = \mathbb{E}[X_0]$$

for all bounded stopping times σ. In particular, $\sigma_n \wedge \sigma$ is bounded as well, and we also have $|X_{t \wedge (\sigma_n \wedge \sigma)}| \leq n + k$ for all $t \geq 0$ as above. Then, using first dominated convergence and afterwards Theorem 2.29 for the martingale X^{τ_m}, we obtain

$$\mathbb{E}[X_{\sigma_n \wedge \sigma}] = \lim_{m\to\infty} \mathbb{E}[X_{\tau_m \wedge (\sigma_n \wedge \sigma)}] = \lim_{m\to\infty} \mathbb{E}[X_{\sigma_n \wedge \sigma}^{\tau_m}] = \lim_{m\to\infty} \mathbb{E}[X_0] = \mathbb{E}[X_0],$$

hence the claim. □

Corollary 6.7. *Martingales are stable under stopping.*

Proof: Let X be martingale and τ a stopping time. Then

$$\mathbb{E}[X_\sigma^\tau] = \mathbb{E}[X_{\tau \wedge \sigma}] = \mathbb{E}[X_0]$$

for every bounded stopping time σ. We conclude as in the proof of Theorem 6.6. □

Theorem 6.8. *Let X be a (right continuous) L^2 martingale, $H \in \mathcal{L}^2(X)$ and let τ be a stopping time. Then we have*

$$1_{[0,\tau]} H \in \mathcal{L}^2(X^\tau)$$

and

$$\int 1_{[0,\tau]} H dX = \int 1_{[0,\tau]} H dX^\tau.$$

Proof: Let first $R = F \times (u, v]$ for some $F \in \mathcal{F}_u$ and set

$$Y_t = \int 1_{R \cap [0,t]} dX = 1_F (X_{t \wedge v} - X_{t \wedge u}).$$

This process is clearly constant after v. Using the isometry property and Theorem 6.1 applied to the finite stopping time $\tau \wedge v$ we then obtain

$$\mu_X(R \cap [0,\tau]) = \mu_X(R \cap [0,\tau \wedge v]) = \mathbb{E}\Big[\Big(\int 1_{R\cap[0,\tau\wedge v]} dX\Big)^2\Big] = \mathbb{E}[Y_{\tau\wedge v}^2] = \mathbb{E}[Y_\tau^2].$$

Also,

$$Y_\tau = 1_F(X_{\tau\wedge v} - X_{\tau\wedge u}) = 1_F(X^\tau_{\tau\wedge v} - X^\tau_{\tau\wedge u}), \tag{6.1}$$

and we obtain

$$\mu_X(R \cap [0,\tau]) = \mathbb{E}[Y_\tau^2] = \mathbb{E}\Big[\Big(\int 1_{R\cap[0,\tau]} dX^\tau\Big)^2\Big] = \mu_{X^\tau}(R \cap [0,\tau]).$$

If we now set $K = 1_P$ for $P \in \mathcal{P}$ with $\mu_X(P) < \infty$ we can apply Theorem 1.65 in Klenke (2006) as in the proof of Theorem 5.15 and obtain

$$\int 1_{[0,\tau]} K d\mu_X = \mu_X(P \cap [0,\tau]) = \mu_{X^\tau}(P \cap [0,\tau]) = \int 1_{[0,\tau]} K d\mu_{X^\tau}. \tag{6.2}$$

Using measure-theoretic (or algebraic) induction this identity holds for all non-negative $\mathcal{P}$-measurable K with $\int K d\mu_X < \infty$. In particular, Remark 5.16 gives

$$H \in \mathcal{L}^2(X) \Longrightarrow 1_{[0,\tau]} H \in \mathcal{L}^2(X) \Longrightarrow 1_{[0,\tau]} H \in \mathcal{L}^2(X^\tau),$$

by setting $K = H^2$.

Using Theorem 6.1, the second claim is already shown for $H = 1_R$ in (6.1). For $H \in \mathcal{E}$

$$\int 1_{[0,\tau]} H dX = \int 1_{[0,\tau]} H dX^\tau$$

follows by linearity, and in general let $H \in \mathcal{L}^2(X)$ and let $(H_n)_n$ be a sequence in $\mathcal{E}$ with $H_n \to H$ in $\mathcal{L}^2(X)$ which exists according to the proof of Theorem 5.15. In particular $1_{[0,\tau]} H_n \to 1_{[0,\tau]} H$ in $\mathcal{L}^2(X)$ as well, and by (6.2) we find that the convergence holds in $\mathcal{L}^2(X^\tau)$ also because the measures μ^X and μ^{X^τ} coincide on the random interval $[0,\tau]$. From the isometry property we conclude

$$\int 1_{[0,\tau]} H dX^\tau = \lim_{n\to\infty} \int 1_{[0,\tau]} H_n dX^\tau = \lim_{n\to\infty} \int 1_{[0,\tau]} H_n dX = \int 1_{[0,\tau]} H dX$$

in L^2. □

Definition 6.9. Let X be a (right continuous) local L^2 martingale. We denote with $\mathcal{L}(X)$ the set of all predictable processes H such that a localising sequence $(\tau_n)_n$ for X exists which satisfies

$$1_{[0,\tau_n]} H \in \mathcal{L}^2(X^{\tau_n}) \text{ for all } n \in \mathbb{N}.$$

Pointwise in ω we then set

$$\int_0^t H dX = \int_0^t 1_{[0,\tau_n]} H dX^{\tau_n}$$

if $t \in [0,\tau_n]$, $n \in \mathbb{N}$.

Remark 6.10.

(i) Let $(\tau_n)_n$ be a localising sequence of X and let $m \leq n$. Then we have for every $0 \leq t \leq \tau_m$ with an application of Theorem 6.8

$$\int_0^t 1_{[0,\tau_n]} H dX^{\tau_n} = \int 1_{[0,\tau_n \wedge t]} H dX^{\tau_n} = \int 1_{[0,\tau_n \wedge t \wedge \tau_m]} H dX^{\tau_n}$$
$$= \int 1_{[0,\tau_n \wedge t \wedge \tau_m]} H dX^{\tau_n \wedge \tau_m} = \int 1_{[0,\tau_m \wedge t]} H dX^{\tau_m} = \int_0^t 1_{[0,\tau_m]} H dX^{\tau_m}.$$

We conclude that $\int_0^t H dX$ is well-defined, at least given a particular localising sequence. A similar argument, however, proves that the definition is independent of the choice of this localising sequence.

(ii) Theorem 5.17(ii) proves that

$$t \mapsto \int_0^t 1_{[0,\tau_n]} H dX^{\tau_n}$$

is a (right continuous) L^2 martingale for each n. Thus

$$t \mapsto \int_0^t H dX$$

is by definition a local (right continuous) L^2 martingale.

Most properties of the classical stochastic integral from Theorem 5.15 translate after localisation to the stochastic integral in Definition 6.9.

Theorem 6.11. *Let X be a (right continuous) local L^2 martingale, $H \in \mathcal{L}(X)$ and let Y be the local L^2 martingale with*

$$Y_t = \int_0^t H dX.$$

If $K \in \mathcal{L}(Y)$ then $KH \in \mathcal{L}(X)$ as well, and we have

$$Z_t := \int_0^t K dY = \int_0^t KH dX.$$

Proof: By Lemma 6.5 there exists a joint localising sequence $(\tau_n)_n$ for X and Y such that $1_{[0,\tau_n]} H \in \mathcal{L}^2(X^{\tau_n})$ and $1_{[0,\tau_n]} K \in \mathcal{L}^2(Y^{\tau_n})$ for all $n \in \mathbb{N}$. Note also that

$$Y_t^{\tau_n} = Y_t = \int_0^t 1_{[0,\tau_n]} H dX^{\tau_n}$$

holds for $t \in [0, \tau_n]$ by definition. Thus we can apply Theorem 5.21(ii) and obtain

$$1_{[0,\tau_n]} KH \in \mathcal{L}^2(X^{\tau_n})$$

and

$$\int_0^t K dY = \int_0^t 1_{[0,\tau_n]} K dY^{\tau_n} = \int_0^t 1_{[0,\tau_n]} KH dX^{\tau_n} = \int_0^t KH dX$$

for $t \in [0, \tau_n]$. The claim follows. $\square$

Theorem 6.12. *Let X be a continuous local martingale and let H be a continuous adapted process with a bounded H_0. Then $H \in \mathcal{L}(X)$, and*

$$t \mapsto \int_0^t H dX$$

is a continuous local martingale.

Proof: Let $|X_0| + |H_0| \leq k$. Using Lemma 6.5 we can assume without loss of generality that

$$\tau_n = \inf\{t \geq 0 \mid |X_t| + |H_t| \geq n\},\ n \in \mathbb{N},$$

is a localising sequence with respect to which X is an L^2 martingale and with respect to which $|H|$ and $|X|$ are bounded by $n + k$ each. From Corollary 6.2 we obtain

$$\begin{aligned}\int 1_{[0,\tau_n]} H^2 d\mu_{X^{\tau_n}} \leq (n+k)^2 \int 1_{[0,\tau_n]} d\mu_{X^{\tau_n}} &= (n+k)^2 \mathbb{E}\Big[\Big(\int 1_{[0,\tau_n]} dX^{\tau_n}\Big)^2\Big] \\ &= (n+k)^2 \mathbb{E}[(X_{\tau_n} - X_0)^2] \leq 2(n+k)^4\end{aligned}$$

for every $n \in \mathbb{N}$. In particular,

$$1_{[0,\tau_n]} H \in \mathcal{L}^2(X^{\tau_n}),$$

so $H \in \mathcal{L}(X)$.

Finally we have to prove that

$$t \mapsto \int_0^t H dX$$

is a continuous local martingale. This property is a consequence of Theorem 5.17 and Theorem 5.19, however, because every

$$t \mapsto \int_0^t 1_{[0,\tau_n]} H dX^{\tau_n}$$

is a continuous martingale. $\square$

Chapter 7

Quadratic variation and the Itō formula

This chapter deals with the Itō formula which can be understood as a version of the change of variables formula for integrals with respect to martingales. As a central term in this context we will also introduce the quadratic variation of a stochastic process. Throughout this chapter we will mostly deal with continuous local martingales and only touch briefly on generalizations.

Remark 7.1. We have seen in Theorem 4.9 that the identity

$$f(A_t) - f(A_0) = \int_0^t f'(A_s)dA_s$$

holds for continuously differentiable functions $f : \mathbb{R} \to \mathbb{R}$ and a continuous process $(A_t)_{t\geq 0}$ of bounded variation. It is clear that such a formula cannot be true in general for continuous (local) martingales X because

$$t \mapsto f(X_0) + \int_0^t f'(X)dX$$

is a continuous (local) martingale according to Theorem 6.12 whereas $(f(X_t))_{t\geq 0}$ in general is not.

Definition 7.2. Let X be a continuous local martingale. The process $([X]_t)_{t\geq 0}$ with

$$[X]_t = X_t^2 - X_0^2 - 2\int_0^t X dX$$

is called the *quadratic variation of* X.

Remark 7.3.

(i) From Theorem 6.12 we have that $([X]_t)_t$ is well-defined, adapted and continuous.

(ii) For a continuous process A of bounded variation Theorem 4.9 gives

$$[A]_t = A_t^2 - A_0^2 - 2\int_0^t A dA = 0.$$

Theorem 7.4. *Let X be a continuous local martingale, $t \geq 0$, and let $\pi_n = \{0 = t_0^n < \ldots < t_{k_n}^n = t\}$ be a partition of $[0,t]$ for each n with*

$$\delta(\pi_n) = \max\{t_i^n - t_{i-1}^n \mid i = 1, \ldots, k_n\} \to 0.$$

Then

$$S_t^n = \sum_{i=1}^{k_n} (X_{t_i} - X_{t_{i-1}})^2$$

converges in probability to $[X]_t$.

Proof: Assume first that X is bounded. In this case we set

$$H^n = \sum_{i=1}^{k_n} X_{t_{i-1}^n} 1_{(t_{i-1}^n, t_i^n]}$$

and obtain

$$S_t^n = \sum_{i=1}^{k_n} \left(X_{t_i^n}^2 - X_{t_{i-1}^n}^2 - 2X_{t_{i-1}^n}(X_{t_i^n} - X_{t_{i-1}^n}) \right) = X_t^2 - X_0^2 - 2\int H^n dX.$$

As X is continuous, it is clear that H^n converges pointwise to $X_{-}1_{[0,t]} = X1_{[0,t]}$ almost surely. Boundedness of X yields convergence in $\mathcal{L}^2(X)$ as well. The isometry property then gives

$$\int H^n dX \to \int_0^t X dX$$

in L^2, so in particular in probability as well.

In the general case we assume that $(\tau_m)_m$ is a localising sequence so that X^{τ_m} is bounded. Now, let $t \leq \tau_m$. We have already shown the L^2 convergence

$$S_t^n = S_{t\wedge\tau_m}^n = \sum_{i=1}^{k_n} (X_{t_i^n}^{\tau_m} - X_{t_{i-1}^n}^{\tau_m})^2 \to [X^{\tau_m}]_t,$$

and by

$$\begin{aligned} \int_0^{t\wedge\tau_m} X^{\tau_m} dX^{\tau_m} &= \int 1_{[0,t\wedge\tau_m]} X^{\tau_m} dX^{\tau_m} = \int 1_{[0,t\wedge\tau_m]} X dX^{\tau_m} \\ &= \int 1_{[0,t\wedge\tau_m]} X dX = \int_0^{t\wedge\tau_m} X dX, \end{aligned}$$

which is a consequence of Theorem 6.8 and Theorem 6.1, we obtain

$$[X^{\tau_m}]_{t\wedge\tau_m} = X_{t\wedge\tau_m}^2 - X_0^2 - 2\int_0^{t\wedge\tau_m} X dX = [X]_{t\wedge\tau_m}. \tag{7.1}$$

So let finally be $\eta, \varepsilon > 0$ arbitrary and m large enough so that $\mathbb{P}(\tau_m < t) < \varepsilon$. Then

$$\begin{aligned} \mathbb{P}(|S_t^n - [X]_t| \geq \eta) &\leq \mathbb{P}(|S_t^n - [X]_t| \geq \eta, \tau_m \geq t) + \mathbb{P}(\tau_m < t) \\ &\leq \mathbb{P}(|S_t^n - [X]_t| 1_{\{\tau_m \geq t\}} \geq \eta) + \varepsilon, \end{aligned}$$

and the first term becomes arbitrary small as prepared. □

Remark 7.5. The quadratic variation can be defined in the case of a (càdlàg) local martingale X as well. In this case on sets

$$[X]_t = X_t^2 - X_0^2 - 2\int_0^t X_- dX$$

where X_- is the left continuous version of X. The same proof as above then gives Theorem 7.4 in this situation as well.

Corollary 7.6. *Let X be a (càdlàg) local martingale. Then the process $([X]_t)_{t\geq 0}$ is increasing.*

Theorem 7.7. *Let B be a Brownian motion. Then $[B]_t = t$ almost surely and*

$$\int_0^t BdB = \frac{1}{2}(B_t^2 - t).$$

Proof: We apply Theorem 4.10 and obtain $[B]_t = t$ almost surely. The second claim then follows by definition of the quadratic variation. □

Definition 7.8. Two stochastic processes X and Y on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ are called *indistinguishable* if for almost all ω the identity $X_t(\omega) = Y_t(\omega)$ holds for all $t \geq 0$. In this case we write $X = Y$ almost surely.

Theorem 7.9. *Let X be a continuous local martingale of bounded variation which satisfies $X_0 = 0$. Then $X = 0$ almost surely.*

Proof: Since X is continuous its values are determined by all rational time points. Therefore it is sufficient to prove $\mathbb{P}(X_t = 0) = 1$ for every fixed $t \geq 0$, i.e. it is sufficient to prove that the zero process is a modification of X in the sense of Definition 2.22. To this end let $(\tau_k)_k$ be a localising sequence. If $\varepsilon > 0$ is now arbitrary, there exists some k with $\mathbb{P}(t > \tau_k) < \varepsilon$. Thus

$$\mathbb{P}(X_t \neq 0) \leq \mathbb{P}(X_t \neq 0, t \leq \tau_k) + \mathbb{P}(t > \tau_k) \leq \mathbb{P}(X_{t\wedge\tau_k} \neq 0) + \varepsilon.$$

If we now apply $X_0 = 0$ and (7.1) to get

$$\mathbb{E}[(X_t^{\tau_k})^2] = 2\mathbb{E}\Big[\int_0^t X^{\tau_k} dX^{\tau_k}\Big] + \mathbb{E}[[X]_{t\wedge\tau_k}]$$

for all $t \geq 0$, then we obtain $\mathbb{E}[(X_t^{\tau_k})^2] = 0$ because the first summand is the expectation of a martingale with start in 0 whereas we can use Remark 7.3(ii) for the second summand. $\mathbb{P}(X_{t\wedge\tau_k} \neq 0) = 0$ then follows. □

Corollary 7.10. *Let M and N be continuous local martingales with $M_0 = N_0 = 0$ and let A and V be processes of bounded variation with $A_0 = V_0 = 0$. If*

$$X_0 + M + A = X_0 + N + V$$

almost surely (as an identity of processes), then $M = N$ and $A = V$ almost surely.

Proof: Use $M - N = V - A$ and apply Theorem 7.9. □

Definition 7.11.

(i) Let A be of bounded variation and let H be an adapted stochastic process. Then we call $H \in \mathcal{L}(A)$ if for almost all ω

$$\int_0^t |H|_s d|A|_s < \infty$$

holds for all $t \geq 0$, where $(|A|_t)_{t\geq 0}$ denotes the total variation process. In this case $t \mapsto \int_0^t HdA$ is well-defined and again of bounded variation. Also, if A is continuous so is the integral.

(ii) Let M be a continuous local martingale and let A be a continuous process of bounded variation. Then $X = X_0 + M + A$ is called a *continuous semimartingale*.

(iii) Let X be a continuous semimartingale and $H \in \mathcal{L}(X) = \mathcal{L}(M) \cap \mathcal{L}(A)$. Then we set

$$\int HdX = \int HdM + \int HdA.$$

(iv) We call $[X] = [M]$ the *quadratic variation of* X.

Remark 7.12. It is clear from the definition of bounded variation processes A that a version of Theorem 7.4 is also true for continuous processes of bounded variation, i.e. we have

$$\sum_{i=1}^{k_n} (A^n_{t_i} - A^n_{t_{i-1}})^2 \xrightarrow{\mathbb{P}} 0$$

for such processes. An application of the binomial formula and the Cauchy-Schwarz inequality (plus Slutsky's lemma) then proves

$$\sum_{i=1}^{k_n} (X^n_{t_i} - X^n_{t_{i-1}})^2 \xrightarrow{\mathbb{P}} [M]_t$$

for continuous semimartingales which explains Definition 7.11(iv).

Example 7.13. Let M be a continuous local martingale. From

$$M_t^2 = M_0^2 + 2\int_0^t MdM + [M]_t$$

it is clear that $(M_t^2)_{t\geq 0}$ is a continuous semimartingale, and we have

$$\int_0^t HdM^2 = 2\int_0^t HMdM + \int_0^t Hd[M]_t$$

from Theorem 6.11 as long as $H \in \mathcal{L}(Z) \cap \mathcal{L}([M])$ with $Z = \int MdM$.

Theorem 7.14. *Let M be a continuous L^2 martingale. Then*

(i) $\mathbb{E}[[M]_t] < \infty$ *for all* $t \geq 0$.

(ii) $Y_t = \int_0^t MdM$ *is a martingale.*

(iii) The Doléans measure satisfies

$$\mu_M(A) = \mathbb{E}\Big[\int 1_A d[M]\Big]$$

for all $A \in \mathcal{P}$.

Proof:

(i) We have

$$\mathbb{E}[S_t^n] = \sum_{i=1}^{k_n} \mathbb{E}\Big[\big(M_{t_i^n}^2 - M_{t_{i-1}^n}^2 - 2M_{t_{i-1}^n}(M_{t_i^n} - M_{t_{i-1}^n})\big)\Big] = \mathbb{E}[M_t^2] - \mathbb{E}[M_0^2] < \infty.$$

From $S_t^n \to [M]_t$ in probability we obtain $S_t^{n_\ell} \to [M]_t$ almost surely along a subsequence. Fatou's lemma then gives

$$\mathbb{E}[[M]_t] \leq \liminf_{\ell\to\infty} \mathbb{E}[S_t^{n_\ell}] < \infty.$$

(ii) We have to prove $\mathbb{E}[Y_\tau] = \mathbb{E}[Y_0]$ for all bounded stopping times τ, and we can apply Theorem 6.12 and Theorem 6.6 to deduce that $(Y_t)_{t\geq 0}$ is a continuous local L^2 martingale. Let $(\tau_n)_n$ be a localising sequence. For $\tau \leq k$ we clearly have

$$2|Y_{\tau\wedge\tau_n}| = |M_{\tau\wedge\tau_n}^2 - M_0^2 - [M]_{\tau\wedge\tau_n}| \leq M_{\tau\wedge\tau_n}^2 + M_0^2 + [M]_k,$$

and from Theorem 2.29 applied to the submartingale $(M^{\tau_n})^2$ and part (i) it follows that $(Y_{\tau\wedge\tau_n})_n$ is uniformly integrable. Together with $Y_{\tau\wedge\tau_n} \to Y_\tau$ almost surely (and again Theorem 2.29) we obtain

$$\mathbb{E}[Y_0] = \mathbb{E}[Y_{\tau\wedge\tau_n}] \to \mathbb{E}[Y_\tau].$$

(iii) Let first $A \in \mathcal{R}$, so without loss of generality $A = F \times (s,t]$ with $F \in \mathcal{F}_s$. Then

$$\begin{aligned}\mu_M(A) &= \mathbb{E}[1_F(M_t^2 - M_s^2)] = \mathbb{E}\Big[1_F\Big(2\int_s^t MdM + ([M]_t - [M]_s)\Big)\Big] \\ &= \mathbb{E}[1_F([M]_t - [M]_s)] = \mathbb{E}\Big[\int 1_{F\times(s,t]} d[M]\Big]\end{aligned}$$

using (ii). The previous computation in combination with (i) proves that $A \mapsto \mathbb{E}\big[\int 1_A d[M]\big]$ is a σ-finite measure, and the claim then follows from the uniqueness theorem for measures. $\square$

Theorem 7.15. *Let M be a continuous local martingale, $H \in \mathcal{L}(M)$ and*

$$Y_t = \int_0^t HdM.$$

Then

$$[Y]_t = \int_0^t H^2 d[M].$$

Proof: Let $(\tau_n)_n$ be a localising sequence for M. From Theorem 6.12 we know that Y^{τ_n} is a continuous martingale as well, and once the claim is proven for Y^{τ_n} we have

$$[Y]_t = \lim_{n\to\infty} [Y]_{t\wedge\tau_n} = \lim_{n\to\infty} [Y^{\tau_n}]_t = \lim_{n\to\infty} \int_0^t H^2 1_{[0,\tau_n]} d[M]$$

almost surely, which can be shown using (7.1). Monotone convergence then gives

$$\lim_{n\to\infty} \int_0^t H^2 1_{[0,\tau_n]} d[M] = \int_0^t H^2 d[M]$$

almost surely.

Thus let M and in turn Y with $Y_t = \int_0^t H dM$ be continuous martingales. Then

$$\mathbb{E}\bigg[\int 1_A H^2 d[M]\bigg] = \int_A H^2 d\mu_M = \mu_Y(A) = \mathbb{E}\bigg[\int 1_A d[Y]\bigg]$$

follows for all $A \in \mathcal{P}$ where we have used Theorem 7.14(iii) and measure-theoretic (algebraic) induction, Theorem 5.21 and again Theorem 7.14(iii). Then in particular

$$\mathbb{E}\bigg[\int_0^\tau H^2 d[M]\bigg] = \mathbb{E}\bigg[\int_0^\tau d[Y]\bigg]$$

for every bounded stopping time τ, and from Corollary 6.2 we get $\mathbb{E}[Z_\tau] = 0$ for

$$Z_t = \int_0^t H^2 d[M] - [Y]_t.$$

This yields that $Z = (Z_t)_{t\geq 0}$ is a martingale which is of bounded variation with continuous paths (use Remark 7.3(i)) and which satisfies $Z_0 = 0$. Theorem 7.9 gives $Z_t = 0$. □

Theorem 7.16. *Let M be a continuous local martingale. Then*

$$\mathcal{L}(M) = \bigg\{ H \text{ predictable} \;\bigg|\; \int_0^t H^2 d[M] < \infty \text{ almost surely for all } t \geq 0 \bigg\}.$$

Proof: Let $H \in \mathcal{L}(M)$ and $Y_t = \int_0^t H dM$. From

$$[Y]_t = \int_0^t H^2 d[M]$$

using Theorem 7.15 the right hand side is almost surely finite.

On the other hand, let $\int_0^t H^2 d[M]$ be almost surely finite. By the continuity of M_t and $[M]_t$ we can assume that

$$\tau_n = \inf\bigg\{ t \geq 0 \mid |M_t| + \Big|\int_0^t H^2 d[M]\Big| \geq n \bigg\}$$

defines a localising sequence such that M^{τ_n} is a bounded martingale and with Theorem 7.14(iii) and (7.1) we have

$$\int H^2 1_{[0,\tau_n]} d\mu_{M^{\tau_n}} = \mathbb{E}\bigg[\int_0^{\tau_n} H^2 d[M^{\tau_n}]\bigg] = \mathbb{E}\bigg[\int_0^{\tau_n} H^2 d[M]\bigg] \leq n,$$

so $H 1_{[0,\tau_n]} \in \mathcal{L}^2(M^{\tau_n})$. □

We have seen in Example 7.13 that for a local martingale M we hardly ever have that M^2 is a local martingale as well. What remains is the semimartingale property, however.

Theorem 7.17. **(Itō formula)** *Let X be a continuous semimartingale and $f \in C^2(\mathbb{R}, \mathbb{R})$. Then $f(X)$ is a continuous semimartingale as well, and we have*

$$f(X_t) - f(X_0) = \int_0^t f'(X)dX + \frac{1}{2}\int_0^t f''(X)d[X]$$

as an identity in the sense of indistinguishability.

Proof: As $f(X)$ is continuous it is enough to prove the result pointwise in t. Also after localisation with

$$\tau_m = \inf\{t \geq 0 \mid |X_t| \geq m\}$$

we can assume that X is bounded. Then we have for every partition $\pi_n = \{0 = t_0^n < \ldots < t_{k_n}^n = t\}$ of $[0, t]$ using Taylor's formula

$$f(X_t) - f(X_0) = \sum_{i=1}^{k_n}(f(X_{t_i^n}) - f(X_{t_{i-1}^n}))$$

$$= \sum_{i=1}^{k_n}\big(f'(X_{t_{i-1}^n})(X_{t_i^n} - X_{t_{i-1}^n}) + \frac{1}{2}f''(X_{t_{i-1}^n})(X_{t_i^n} - X_{t_{i-1}^n})^2 + R(X_{t_{i-1}^n}, X_{t_i^n})\big),$$

where

$$|R(x, y)| \leq v(|y - x|)(y - x)^2$$

for an increasing function $v : \mathbb{R} \to \mathbb{R}$ with $v(u) \to 0$ for $u \to 0$. (The latter property holds as f'' is uniformly continuous over a bounded interval.) If we choose π_n such that

$$\delta(\pi_n) = \max\{t_i^n - t_{i-1}^n \mid i = 1, \ldots, k_n\} \to 0$$

holds then

$$\sum_{i=1}^{k_n} f'(X_{t_{i-1}^n})(X_{t_i^n} - X_{t_{i-1}^n}) \to \int_0^t f'(X)dX$$

in probability, using Theorem 4.9 for the bounded variation part and the definition of the stochastic integral with respect to martingales otherwise. Analogously,

$$\sum_{i=1}^{k_n} f''(X_{t_{i-1}})(X_{t_i^n} - X_{t_{i-1}^n})^2 \to \int_0^t f''(X)d[X]$$

in probability using Remark 7.12 (and a.s. convergence along a subsequence). Finally,

$$\sum_{i=1}^{k_n} R(X_{t_{i-1}^n}, X_{t_i}^n) \leq \max\{v(|X_{t_i}^n - X_{t_{i-1}^n}|) \mid i = 1, \ldots, k_n\} \sum_{i=1}^{k_n}(X_{t_i^n} - X_{t_{i-1}^n})^2$$

Uniform continuity of X proves a.s. convergence of the first factor to 0 while the second one converges in probability to $[X]_t$. The claim follows from uniqueness of the limit in probability. $\square$

Definition 7.18. Let X and Y be continuous semimartingales. Then

$$[X, Y] = \frac{1}{4}\big([X + Y] - [X - Y]\big)$$

is called the *quadratic covariation of X and Y*.

Corollary 7.19.

(i) *For two continuous semimartingales X and Y we have that $[X,Y]$ is of bounded variation.*

(ii) *For a continuous semimartingale X we have*

$$[X,X] = [X].$$

(iii) *If X is a continuous semimartingale and Y continuous and of bounded variation, then*

$$[X,Y] = 0.$$

(iv) *For two continuous semimartingales X and Y we have the integration by parts rule*

$$X_tY_t = X_0Y_0 + \int_0^t X\,dY + \int_0^t Y\,dX + [X,Y]_t.$$

Proof:

(i) $[X,Y]$ is the difference of two increasing processes.

(ii) We have

$$[X,X] = \frac{1}{4}[2X] = \frac{1}{4}4[X] = [X].$$

(iii) By Definition 7.11(iv)

$$[X,Y] = \frac{1}{4}\big([X+Y] - [X-Y]\big) = \frac{1}{4}\big([X] - [X]\big) = 0.$$

(iv) It holds

$$\begin{aligned}
[X,Y] &= \frac{1}{4}\big([X+Y] - [X-Y]\big) \\
&= \frac{1}{4}\Big((X+Y)_t^2 - (X+Y)_0^2 - 2\int (X+Y)d(X+Y)\Big) \\
&\quad - \frac{1}{4}\Big((X-Y)_t^2 - (X-Y)_0^2 - 2\int (X-Y)d(X-Y)\Big) \\
&= X_tY_t - X_0Y_0 - \int X\,dY - \int Y\,dX,
\end{aligned}$$

where we have used

$$\int (X+Y)d(X+Y) = \int X\,dX + \int Y\,dX + \int X\,dY + \int Y\,dY$$

and similarly for $X - Y$. □

Notation 7.20. There are several alternative notations for a stochastic integral. In general we write

$$X_t = X_0 + \int_0^t H\,dM = X_0 + \int_0^t H_{s-}\,dM_s$$

as well as

(i) $dX_t = H_{t-}dM_t$,

(ii) $X = X_0 + H_- \bullet M$.

Using these notations the Itō formula and the integration by parts rule become

$$d(XY)_t = X_t dY_t + Y_t dX_t + d[X,Y]_t \quad \text{and}$$
$$df(X)_t = f'(X_t)dt + \frac{1}{2}f''(X_t)d[X]_t$$

in (i) and

$$XY = X_0Y_0 + X \bullet Y + Y \bullet X + [X,Y] \quad \text{and}$$
$$f(X) = f(X_0) + f'(X) \bullet X + \frac{1}{2}f''(X) \bullet [X]$$

in (ii). Note that we have used continuity of X and Y here.

Remark 7.21.

(i) There is a multivariate Itō formula as well: For continuous semimartingales $X^1, \ldots, X^k$ and $f \in C_2(\mathbb{R}^k, \mathbb{R})$ we have

$$f(X_t) - f(X_0) = \sum_{i=1}^{k} \int_0^t D_i f(X) dX^i + \frac{1}{2} \sum_{i,j=1}^{k} \int_0^t D_{ij} f(X) d[X^i, X^j]$$

with $X_t = (X_t^1, \ldots, X_t^k)^*$ and the partial derivatives D_i and D_{ij}, respectively, where x^* denotes the transpose of a vector x.

(ii) An Itō formula for semimartingales with jumps can e.g. be found as Theorem 32 of Chapter II in Protter (2004).

Theorem 7.22. *Let M and N be continuous local martingales. Then $MN - [M,N]$ is a continuous local martingale as well.*

Proof: We have

$$MN - [M,N] = \frac{1}{4}\big(((M+N)^2 - [M+N]) - ((M-N)^2 - [M-N])\big),$$

and for example

$$(M_t + N_t)^2 - [M+N]_t = (M_0 + N_0)^2 + 2\int_0^t (M+N)d(M+N)$$

is a continuous local martingale by Theorem 6.12. □

Remark 7.23. Let X and Y be two continuous semimartingales and let $H \in \mathcal{L}(X)$ and $K \in \mathcal{L}(Y)$. Then with a similar reasoning as in Theorem 7.15 we have

$$\left[\int H dX, \int K dY\right]_t = \int_0^t KH d[X,Y].$$

Definition 7.24. Let B be a Brownian motion and $I_t = t$ the identity process. Then X is called an *Itō process* if

$$X_t = X_0 + \int_0^t \mu_s ds + \int_0^t \sigma_s dB_s$$

with processes $\mu \in \mathcal{L}(I)$ and $\sigma \in \mathcal{L}(B)$.

Corollary 7.25. *Let X and Y be Itō processes with respect to μ, σ and μ', σ', respectively, and with the same Brownian motion B.*

(i) For $f \in C^2(\mathbb{R}^2, \mathbb{R})$ we have

$$\begin{aligned} f(X_t, t) = f(X_0, 0) &+ \int_0^t D_1 f(X_s, s)\sigma_s dB_s + \\ &+ \int_0^t \big(D_1 f(X_s, s)\mu_s + D_2 f(X_s, s) + \frac{1}{2} D_{11} f(X_s, s)\sigma_s^2\big) ds. \end{aligned}$$

(ii) It holds

$$[X, Y]_t = \int_0^t \sigma_s \sigma_s' ds.$$

Proof: Part (i) follows from the Itō formula in connection with Theorem 4.10 and Theorem 7.15 which are needed to compute the quadratic variation of the martingale parts. Remark 7.23 finally gives (ii). □

Chapter 8

The stochastic exponential and Girsanov's theorem

This chapter deals mostly with the stochastic exponential, an important process which allows to prove several key results like Lévy's characterisation theorem, the martingale representation theorem and Girsanov's theorem. The latter is of particular importance in mathematical finance as it discusses what happens to martingales in the case of a measure change.

Theorem 8.1. *Let X be a local martingale and assume that there exists some $Z \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ with $X_t \geq Z$ for all $t \geq 0$. If then there exists some $T > 0$ such that $\mathbb{E}[X_T] = \mathbb{E}[X_0]$ holds then the stopped process X^T is a martingale.*

Proof: Let $s < t$ and let $(\tau_n)_n$ be a localising sequence for X. Then

$$X_s = \lim_{n\to\infty} X_{s\wedge\tau_n} = \lim_{n\to\infty} \mathbb{E}[X_{t\wedge\tau_n}|\mathcal{F}_s] \geq \mathbb{E}[\liminf_{n\to\infty} X_{t\wedge\tau_n}|\mathcal{F}_s] = \mathbb{E}[X_t|\mathcal{F}_s]$$

almost surely where we have used the conditional version of Fatou's lemma which needs boundedness from below using Z. Let then τ be a finite stopping time. Then

$$\mathbb{E}[X_0] \geq \mathbb{E}[X_{T\wedge\tau}] \geq \mathbb{E}[X_T] = \mathbb{E}[X_0],$$

using Theorem 2.29 twice. $\mathbb{E}[X^T_\tau] = \mathbb{E}[X_0]$ then proves that X^T is a martingale. □

Corollary 8.2. *Every locally bounded (from below) local martingale is a supermartingale.*

Definition 8.3. Let X be a continuous semimartingale. Then the process

$$\mathcal{E}(X)_t = \exp\left(X_t - X_0 - \frac{1}{2}[X]_t\right)$$

is called the *stochastic exponential of* X.

Theorem 8.4.

(i) $\mathcal{E}(X)$ is a continuous semimartingale and solves

$$\mathcal{E}(X)_t = 1 + \int_0^t \mathcal{E}(X)_s dX_s.$$

(ii) If X is a continuous local martingale so is $\mathcal{E}(X)$.

(iii) Let X be a continuous local martingale. If for some $T < \infty$ the Novikov condition

$$\mathbb{E}\left[\exp\left(\frac{1}{2}[X]_T\right)\right] < \infty$$

holds then $\mathcal{E}(X)^T$ is a martingale.

Proof:

(i) Set $Y_t = X_t - X_0 - \frac{1}{2}[X]_t$ so that $\mathcal{E}(X) = \exp(Y)$ and $[Y] = [X]$. Theorem 7.17 then gives

$$\begin{aligned}\exp(Y_t) &= \exp(Y_0) + \int_0^t \exp(Y_s)dY_s + \frac{1}{2}\int_0^t \exp(Y_s)d[Y]_s \\ &= \exp(0) + \int_0^t \exp(Y_s)dX_s - \frac{1}{2}\int_0^t \exp(Y_s)d[X]_s + \frac{1}{2}\int_0^t \exp(Y_s)d[Y]_s \\ &= 1 + \int_0^t \mathcal{E}(X)_s dX_s.\end{aligned}$$

(ii) This claim is a direct consequence of (i).

(iii) We use Kazamaki's Criterion (Theorem 40 of Chapter III in Protter (2004)) from which it is sufficient to prove

$$\mathbb{E}\left[\exp\left(\frac{1}{2}(X_t - X_0)\right)\right] \le \mathbb{E}\left[\exp\left(\frac{1}{2}[X]_t\right)\right]^{\frac{1}{2}}$$

for every $t \le T$ in order to deduce uniform integrability of $\mathcal{E}(X)^T$. $\mathbb{E}[\mathcal{E}(X)_T] = \mathbb{E}[\mathcal{E}(X)_0]$ then follows as in the proof of Theorem 7.14(ii), and finally Theorem 8.1 can be applied. We use

$$\exp\big(\frac{1}{2}(X_t - X_0)\big) = (\mathcal{E}(X)_t)^{\frac{1}{2}} \exp(\frac{1}{2}[X]_t)^{\frac{1}{2}},$$

and since $\mathcal{E}(X)$ (as a non-negative local martingale) is a supermartingale by Corollary 8.2

$$\mathbb{E}[\mathcal{E}(X)_t] \le \mathbb{E}[\mathcal{E}(X)_0] = 1$$

and the Cauchy-Schwarz inequality finish the proof. □

Theorem 8.5. (Lévy's characterisation theorem) *Let X be a continuous local martingale with $X_0 = 0$ and $[X]_t = t$ for all $t \ge 0$. Then X is a Brownian motion.*

Proof: Clearly, $Z = (Z_t)_{t\ge 0}$ with $Z_t = iuX_t$ is a complex-valued local martingale with $[Z]_t = -u^2 t$. In particular

$$\mathcal{E}(Z)_t = \exp\left(Z_t - Z_0 - \frac{1}{2}[Z]_t\right) = \exp\left(iuX_t + \frac{1}{2}u^2 t\right),$$

and from Theorem 8.4(iii) we obtain that $\mathcal{E}(Z)_t$ is a martingale with $\mathbb{E}[\mathcal{E}(Z)_t] = 1$ on the entire interval $[0, \infty)$. Thus

$$\mathbb{E}\big[\exp(iu(X_t - X_s))\big|\mathcal{F}_s\big] = \exp\left(-\frac{1}{2}u^2(t-s)\right).$$

Thus the $\mathcal{F}_s$-conditional distribution of $X_t - X_s$ is independent of $\mathcal{F}_s$ and equals a normal distribution with variance $t - s$. It is then easy to deduce that X is a Brownian motion in the sense of Definition 1.3. □

The next result is the martingale representation theorem which has several important applications, in particular in mathematical finance. Even though its statement shows some resemblance to the one of Theorem 3.9, it is about almost sure identity rather than identity in distribution and thus needs a different proof.

Theorem 8.6. **(Martingale representation theorem)** *Let B be a Brownian motion with generated filtration $(\mathcal{F}_t)_{t\geq 0}$, and let $X \in L^2(\Omega, \mathcal{F}, \mathbb{P})$ with*

$$\mathcal{F} = \mathcal{F}_\infty = \sigma(\mathcal{F}_t \mid t \geq 0).$$

Then there exists a unique predictable process $H \in \mathcal{L}^2(B)$ with

$$X = \mathbb{E}[X] + \int H dB. \tag{8.1}$$

Proof: If H and K both satisfy (8.1) then

$$\int (H - K) dB = 0$$

$\mathbb{P}$-almost surely. The isometry property gives

$$\int (H - K)^2 d\mu_B = 0,$$

so with Remark 5.13 $H = K$ $\mathbb{P} \otimes \lambda$-almost surely.

To prove the existence of such an H note first that

$$\mathcal{U} = \{X \in L^2(\Omega, \mathcal{F}, \mathbb{P}) \mid \text{there exists some } H \in \mathcal{L}^2(B) \text{ with } (8.1)\}$$

is closed, for which we use the isometry property and that both $L^2(\Omega, \mathcal{F}, \mathbb{P})$ and $\mathcal{L}^2(B)$ are closed themselves. We will prove

$$\mathcal{U}^\perp = \{0\}$$

in $L^2(\Omega, \mathcal{F}, \mathbb{P})$. Then, since $\mathcal{U}$ is a closed linear subspace and $L^2(\Omega, \mathcal{F}, \mathbb{P})$ is a Hilbert space, one obtains (see e.g. Chapter 2 in Brockwell and Davis (1991), in particular Proposition 2.3.2)

$$L^2(\Omega, \mathcal{F}, \mathbb{P}) = \mathcal{U} \oplus \mathcal{U}^\perp = \mathcal{U}.$$

So let $Y \in \mathcal{U}^\perp$, and we need to show $\mathbb{E}[XY] = 0$ for all $X \in L^2(\Omega, \mathcal{F}, \mathbb{P})$. To this end, let first $H = \sum_{i=1}^n \alpha_i 1_{(t_{i-1}, t_i]}$. If we set

$$Z_t = \int_0^t H dB$$

then

$$\mathcal{E}(Z)_t = 1 + \int_0^t \mathcal{E}(Z)_s dZ_s = 1 + \int_0^t \mathcal{E}(Z)_s H_s dB_s \in \mathcal{U}$$

from Theorem 8.4. Thus

$$\exp\Big(\sum_{i=1}^n \alpha_i (B_{t_i} - B_{t_{i-1}}) - \sum_{i=1}^n \frac{1}{2}\alpha_i^2 (t_i - t_{i-1})\Big) \in \mathcal{U}$$

for all n, all $\alpha_1, \dots, \alpha_n \in \mathbb{R}$ and any choice of $0 = t_0 < t_1 < \dots < t_n$. By assumption $Y \in \mathcal{U}^\perp$, so

$$\mathbb{E}\Big[Y \exp\big(\sum_{i=1}^n \alpha_i (B_{t_i} - B_{t_{i-1}}) - \sum_{i=1}^n \frac{1}{2}\alpha_i^2 (t_i - t_{i-1})\big)\Big] = 0$$

follows. Since $\mathcal{U}$ contains all constants, $\mathbb{E}[Y] = 0$. Thus by forgetting the multiplicative constant and applying partial summation it is clear that $\mathbb{E}[XY] = 0$ for all X of the form

$$X = \exp\big(\sum_{i=1}^n \beta_i B_{t_i}\big).$$

Using methods from Fourier analysis it can be shown that the family of all such X lies dense in $L^2(\Omega, \mathcal{F}, \mathbb{P})$. For details see Theorem 4.15 and Proposition 4.18 of Chapter 3 in Karatzas and Shreve (1991). Then $\mathbb{E}[XY] = 0$ for all $X \in L^2(\Omega, \mathcal{F}, \mathbb{P})$ as required. □

Theorem 8.7. *Let $T > 0$ and let $\mathbb{Q}$ be a probability measure which is equivalent to $\mathbb{P}$ on $\mathcal{F}_T$, i.e. $\mathbb{Q}|_{\mathcal{F}_T} \sim \mathbb{P}|_{\mathcal{F}_T}$. Then its density*

$$L_T = \frac{d\mathbb{Q}}{d\mathbb{P}}\Big|_{\mathcal{F}_T}$$

defines a martingale with respect to $\mathbb{P}$ via

$$L_t = \mathbb{E}[L_T | \mathcal{F}_t], \quad 0 \le t \le T,$$

and it holds

$$L_t = \frac{d\mathbb{Q}}{d\mathbb{P}}\Big|_{\mathcal{F}_t}.$$

The resulting process $(L_t)_{0\le t\le T}$ is called the density process.

Proof: The martingale property follows from Example 2.4(iii). So let $A \in \mathcal{F}_t$, so in particular $A \in \mathcal{F}_T$. Then

$$\mathbb{Q}(A) = \int_A L_T d\mathbb{P} = \int_A L_t d\mathbb{P},$$

by definition of the conditional expectation, using that $(L_t)_{0\le t\le T}$ is a martingale. □

Condition 8.8. From now on let $\mathbb{Q}$ be an equivalent probability measure on $\mathcal{F}_T$ and assume that $L = (L_t)_{0\le t\le T}$ is càdlàg, positive and satisfies $L_0 = 1$, so $\mathbb{Q}|_{\mathcal{F}_0} = \mathbb{P}|_{\mathcal{F}_0}$. Terms which depend on the actual choice of $\mathbb{P}$ and $\mathbb{Q}$, like expectations, will be denoted with $\mathbb{E}_\mathbb{P}$ and $\mathbb{E}_\mathbb{Q}$, respectively.

Theorem 8.9. *Let X be a right continuous adapted process. Then the following are equivalent:*

(i) $(L_t X_t)_{0\le t\le T}$ is a (local) $\mathbb{P}$-martingale.

(ii) $(X_t)_{0\le t\le T}$ is a (local) $\mathbb{Q}$-martingale.

Proof: It is enough to prove the result for martingales, using localisation afterwards for the general case. So let $s \le t \le T$ and $F \in \mathcal{F}_s$. From

$$\begin{aligned}\mathbb{E}_{\mathbb{Q}}[1_F(X_t - X_s)] &= \mathbb{E}_{\mathbb{P}}[L_T 1_F(X_t - X_s)] = \mathbb{E}_{\mathbb{P}}[L_T 1_F X_t] - \mathbb{E}_{\mathbb{P}}[L_T 1_F X_s] \\ &= \mathbb{E}_{\mathbb{P}}[L_t 1_F X_t] - \mathbb{E}_{\mathbb{P}}[L_s 1_F X_s] = \mathbb{E}_{\mathbb{P}}[1_F(L_t X_t - L_s X_s)]\end{aligned}$$

both statements are equivalent. □

Theorem 8.10. (Girsanov's theorem) *Let X be a continuous local martingale with respect to $\mathbb{P}$ and assume that the density process L is continuous. Then $Z = (Z_t)_{0 \le t \le T}$ defined by*

$$Z_t = X_t - \int_0^t \frac{1}{L} d[L, X]$$

is a local martingale with respect to $\mathbb{Q}$, and it holds

$$[Z]^{\mathbb{Q}} = [X]^{\mathbb{Q}} = [X]^{\mathbb{P}}.$$

Proof: Using Theorem 8.9 we have to show that LZ is a local martingale with respect to $\mathbb{P}$. So let

$$D_t = \int_0^t \frac{1}{L} d[L, X].$$

Then

$$\begin{aligned} L_t Z_t = L_t X_t - L_t D_t &= L_0 X_0 + \int_0^t L dX + \int_0^t X dL + [L, X]_t \\ &\quad - \big(L_0 D_0 + \int_0^t L dD + \int_0^t D dL + [L, D]_t\big) \\ &= L_0 X_0 + \int_0^t L dX + \int_0^t X dL + [L, X]_t \\ &\quad - \int_0^t L \frac{1}{L} d[L, X] - \int_0^t D dL. \end{aligned}$$

Here we have used that D is of bounded variation and satisfies $D_0 = 0$, from which

$$dD = \frac{1}{L} d[L, X]$$

follows, and we have also applied Corollary 7.19(iii). Thus

$$L_t Z_t = L_0 X_0 + \int_0^t L dX + \int_0^t X dL - \int_0^t D dL$$

and Theorem 8.7 give the first claim.

For the second claim note that $[Z]^{\mathbb{Q}} = [X]^{\mathbb{Q}}$ holds as $[L, X]$ is of bounded variation using Corollary 7.19(i). Note also that the equivalence of $\mathbb{P}$ and $\mathbb{Q}$ implies that the sequence converges in probability with respect to $\mathbb{P}$ if and only if it converges in probability with respect to $\mathbb{Q}$, and with the same limit then. Thus $[X]^{\mathbb{P}} = [X]^{\mathbb{Q}}$ follows from Theorem 7.4. □

Corollary 8.11. *Let B be a Brownian motion, $H \in \mathcal{L}^2(B)$ and let $L_t = \mathcal{E}(H \bullet B)_t$ be a martingale for $0 \le t \le T$. Then*

$$Z_t = B_t - \int_0^t H_s ds, \quad 0 \le t \le T,$$

is a Brownian motion with respect to $\mathbb{Q}$.

Proof: Using Theorem 8.4 and Theorem 6.11 we have

$$L_t = \mathcal{E}(H \bullet B)_t = 1 + \int_0^t \mathcal{E}(H \bullet B)_s d(H \bullet B)_s = 1 + \int_0^t L_s H_s dB_s.$$

With Remark 7.23

$$[L, B]_t = [(LH) \bullet B, B]_t = \int_0^t L_s H_s ds$$

follows, so

$$\int_0^t \frac{1}{L_s} d[L, B]_s = \int_0^t H_s ds.$$

With Theorem 8.10 we then obtain first that Z is a continuous local $\mathbb{Q}$ martingale with $[Z]_t = [B]_t = t$. The claim then follows from Theorem 8.5. $\square$

Chapter 9

Modelling financial markets

In this chapter we are discussing a stock market with $d+1$ assets whose price processes are modelled by continuous semimartingales $S^0, \dots, S^d$. As a shorthand notation we write $S = (S^0, \dots, S^d)^*$, and for simplicity we assume that all $\mathcal{F}_0$-measurable random variables are deterministic.

Definition 9.1.

(i) An $\mathbb{R}^{d+1}$-valued predictable process $\varphi = (\varphi^0, \dots, \varphi^d)^*$ is called a *(trading) strategy* or a *portfolio*.

(ii) The process

$$V(\varphi) = \varphi^* S = \sum_{i=0}^{d} \varphi^i S^i$$

is called the *value* or *wealth process*.

(iii) A trading strategy is called *self-financing* if

$$\varphi \in \mathcal{L}(S) \quad \text{and} \quad V(\varphi) = V_0(\varphi) + \varphi^* \bullet S.$$

Here we have set

$$\varphi \in \mathcal{L}(S) \iff \varphi^i \in \mathcal{L}(S^i) \text{ for all } i$$

and

$$(\varphi^* \bullet S)_t = \sum_{i=0}^{d} (\varphi^i \bullet S^i)_t = \sum_{i=0}^{d} \int_0^t \varphi^i dS^i.$$

Remark 9.2.

(i) The random variable φ_t^i denotes the number of shares of type i held at time t. Using the discrete approximation

$$\varphi_n = \sum_{j=1}^{k_n} \varphi_{t_{j-1}^n} 1_{(t_{j-1}^n, t_j^n]}$$

of φ with a partition π_n of $[0,T]$, it can be seen that the stochastic integral $(\varphi^* \bullet S)_t$, which is approximated by

$$(\varphi_n^* \bullet S)_t = \sum_{i=0}^{d} \sum_{j=1}^{k_n} \varphi_{t_{j-1}^n}^i (S_{t_j^n \wedge t}^i - S_{t_{j-1}^n \wedge t}^i), \quad 0 \le t \le T,$$

stands for the cumulative gains and losses from trading with the portfolio φ over $[0,t]$.

(ii) If we assume that the portfolio is self-financing then this means that no money is put into or withdrawn from the wealth process after time 0. All changes in $V(\varphi)$ are thus due to changes in S.

In many situations it makes sense to discuss trading gains and strategies with respect to a reference asset and thus just to discuss relative changes. This role will be played by S^0 in the following, and we assume $S^0 > 0$.

Definition 9.3.

(i) The price process S^0 is called a *numeraire*.

(ii) The processes

$$\hat{S} = \frac{1}{S^0}S = \big(1, \frac{S^1}{S^0}, \dots, \frac{S^d}{S^0}\big)^*$$

and

$$\hat{V}(\varphi) = \frac{1}{S_0}V(\varphi) = \varphi^* \hat{S}$$

are called the *discounted price* and *wealth process* of φ, respectively.

Theorem 9.4. *A strategy φ is self-financing if and only if*

$$\varphi \in \mathcal{L}(\hat{S}) \quad \text{and} \quad \hat{V}(\varphi) = \hat{V}_0(\varphi) + \varphi^* \bullet \hat{S}.$$

Proof: First, let $\varphi \in \mathcal{L}(\hat{S})$ and $\hat{V}(\varphi) = \hat{V}_0(\varphi) + \varphi^* \bullet \hat{S}$. Using integration by parts (Corollary 7.19(iv)) we have

$$\begin{aligned} V(\varphi) = \varphi^* S = (\varphi^*\hat{S})S^0 &= \varphi_0^*\hat{S}_0 S_0^0 + (\varphi^*\hat{S}) \bullet S^0 + S^0 \bullet (\varphi^*\hat{S}) + [\varphi^*\hat{S}, S^0] \\ &= \varphi_0^* S_0 + (\varphi^*\hat{S}) \bullet S^0 + S^0 \bullet (\varphi_0^*\hat{S}_0^0 + \varphi^* \bullet \hat{S}) + [(\varphi_0^*\hat{S}_0^0 + \varphi^* \bullet \hat{S}), S^0]. \end{aligned}$$

Now, using (a slight generalization of) Theorem 6.11, Remark 7.23 and the fact that the addition of constants in an integrator does not change an integral we obtain

$$V(\varphi) = \varphi_0^* S_0 + (\varphi^*\hat{S}) \bullet S^0 + (\varphi^* S^0) \bullet \hat{S} + \varphi^* \bullet [\hat{S}, S^0].$$

Then with Theorem 6.11 again

$$\begin{aligned} V(\varphi) &= \varphi_0^* S_0 + \varphi^* \bullet (\hat{S} \bullet S^0 + S^0 \bullet \hat{S} + [\hat{S}, S^0]) \\ &= \varphi_0^* S_0 + \varphi^* \bullet (\hat{S}S^0 - \hat{S}_0 S_0^0) = V_0(\varphi) + \varphi^* \bullet S. \end{aligned}$$

The other direction works analogously with the roles interchanged. (Note that we have not provided the equivalence between $\varphi \in \mathcal{L}(\hat{S})$ and $\varphi \in \mathcal{L}(S)$.) □

Theorem 9.5. *For every predictable process $(\varphi^1, \dots, \varphi^d)^* \in \mathcal{L}((\hat{S}^1, \dots, \hat{S}^d)^*)$ and every $V_0 \in \mathbb{R}$ there exists a unique predictable process φ^0 such that $\varphi = (\varphi^0, \dots, \varphi^d)^*$ is self-financing with $V_0(\varphi) = V_0$.*

Proof: Using Theorem 9.4 we have that φ is self-financing if and only if

$$\begin{aligned}\varphi_t^0 \hat{S}_t^0 + \big((\varphi^1,\dots,\varphi^d)^*(\hat{S}^1,\dots,\hat{S}^d)\big)_t &= \hat{V}_t(\varphi) = \hat{V}_0(\varphi) + (\varphi^* \bullet \hat{S})_t \\ &= \hat{V}_0(\varphi) + \big((\varphi^1,\dots,\varphi^d)^* \bullet (\hat{S}^1,\dots,\hat{S}^d)\big)_t,\end{aligned}$$

where we have used in the last step that $\hat{S}^0 = 1$. Solving the equation this means that

$$\varphi_t^0 = \hat{V}_0(\varphi) + \big((\varphi^1,\dots,\varphi^d)^* \bullet (\hat{S}^1,\dots,\hat{S}^d)\big)_t - \big((\varphi^1,\dots,\varphi^d)^*(\hat{S}^1,\dots,\hat{S}^d)\big)_t$$

and $\hat{V}_0(\varphi) = V_0(\varphi)/S_0^0$ have to hold. □

Remark 9.6. The key idea in mathematical finance is that no riskless gains can be obtained, i.e. that there does not exist a self-financing strategy φ and a time $T < \infty$ such that

$$V_0(\varphi) = 0, \quad V_T(\varphi) \geq 0 \quad \text{and} \quad \mathbb{P}(V_T(\varphi) > 0) > 0$$

hold. This *absence of arbitrage* is the base for most non-trivial results in mathematical finance.

Example 9.7.

(i) In Example 2.10 we have investigated the (discrete time) martingale strategy which allows to obtain almost sure gains if the underlying asset is a martingale. Key to this example was the formulation of the strategy via a stopping time for which to happen one can wait in theory arbitrarily long. Thus this strategy does not prove the existence of an arbitrage opportunity in the sense of Remark 9.6 because T is fixed there.

(ii) On the other hand, if we are in continuous time we can trade in principle even in short time periods arbitrarily often. If we use the interpretation of a Brownian motion as a time continuous limit of a rescaled random walk following Theorem 3.13 it is not difficult to see that there exist arbitrage opportunities even in the simplest possible market with $S^0 = 1$ and a *geometric Brownian motion*

$$S_t^1 = \mathcal{E}(B)_t = \exp(B_t - t/2),$$

i.e. there exists a self-financing strategy $\varphi = (\varphi^0, \varphi^1)^*$ analogously to Example 2.10 so that $V_0(\varphi) = 0$ and $V_T(\varphi) = 1$ almost surely.

Remark 9.8. To still be able to exclude arbitrage opportunities one thus needs to include *admissibility restrictions* for the portfolios allowed. A first natural idea could be to allow only finitely many trades in finite time, but even though this restriction is realistic it excludes a lot of interesting strategies. Typically, one therefore works with restrictions on the size of the debts allowed, and we additionally assume $S^i \geq 0$ for all $i = 1, \dots, d$ from now on.

Definition 9.9. A self-financing strategy φ is called *admissible* if there exists some $c > 0$ such that

$$V(\varphi) \geq -c \sum_{i=0}^{d} S^i.$$

Remark 9.10. In the following we let $T > 0$ be fixed, and for simplicity we just discuss price processes and strategies over $[0, T]$ rather than $[0, \infty)$. In particular we set $\mathcal{F} = \mathcal{F}_T$. Then a market is called *free of arbitrage* if there does not exist a self-financing admissible strategy φ such that

$$V_0(\varphi) = 0, \quad V_T(\varphi) \geq 0 \quad \text{and} \quad \mathbb{P}(V_T(\varphi) > 0) > 0.$$

Theorem 9.11. *If there exists an* equivalent martingale measure $\mathbb{Q}$*, i.e. a measure* $\mathbb{Q} \sim \mathbb{P}$ *with respect to which* $\hat{S}$ *is a martingale, then the market is free of arbitrage.*

Proof: Let φ be an admissible self-financing strategy with $V_0(\varphi) = 0$ and let $\hat{S}$ be a martingale with respect to $\mathbb{Q}$. Then

$$\hat{V}(\varphi) = \varphi^* \bullet \hat{S}$$

is at least a local martingale with respect to $\mathbb{Q}$, and by choice of $c > 0$ large enough it is clear that

$$M = \hat{V}(\varphi) + c \sum_{i=1}^{d} \hat{S}^i \geq 0$$

is a non-negative local martingale. Thus from Corollary 8.2 we have that first M and then $\hat{V}(\varphi)$ is a supermartingale with respect to $\mathbb{Q}$. In particular

$$\mathbb{E}_{\mathbb{Q}}[\hat{V}_T(\varphi)] \leq \mathbb{E}_{\mathbb{Q}}[\hat{V}_0(\varphi)] = 0.$$

If we then have $\hat{V}_T(\varphi) \geq 0$ it is clear that $\hat{V}_T(\varphi) = 0$ has to hold $\mathbb{Q}$-almost surely. Because of $S_0 > 0$ then $V_T(\varphi) = 0$ $\mathbb{Q}$-almost surely as well, and the claim follows because of $\mathbb{P} \sim \mathbb{Q}$. $\square$

Corollary 9.12. *Let* φ *be admissible and* $\mathbb{Q}$ *an equivalent martingale measure for* $\hat{S}$*. Then*

$$\hat{V}_t(\varphi) = \hat{V}_0(\varphi) + (\varphi^* \bullet \hat{S})_t$$

is a supermartingale with respect to $\mathbb{Q}$*.*

Remark 9.13. In contrast to discrete mathematical finance freedom of arbitrage in a market is not sufficient for the existence of an equivalent martingale measure, i.e. there does not exist the opposite direction to Theorem 9.11. One therefore needs to assume a slightly stronger condition.

Definition 9.14. A random variable $X \geq 0$ with $\mathbb{P}(X > 0) > 0$ is called a *free lunch with vanishing risk* if there exists a sequence $(\varphi^{(n)})_n$ of admissible strategies and a zero sequence $(v_n)_n$ with $v_n \geq 0$ such that

$$V_0(\varphi^{(n)}) \leq v_n \quad \text{and} \quad V_T(\varphi^{(n)}) \geq X$$

for all $n \in \mathbb{N}$. A market satisfies *NFLVR* (*no free lunch with vanishing risk*) if no such X exists.

Theorem 9.15. *A market which satisfies NFLVR is free of arbitrage.*

Proof: If there exists an arbitrage opportunity then one can choose $\varphi^{(n)} = \varphi$, $X = V_T(\varphi)$ and $v_n = 0$. This sequence of strategies corresponds to a free lunch without any risk. $\square$

Theorem 9.16. (First fundamental theorem of asset pricing) *A market satisfies NFLVR if and only if there exists an equivalent martingale measure $\mathbb{Q}$ for $\hat{S}$.*

Proof:

$\Longleftarrow$ Suppose X is a FLVR. If we set

$$\hat{X} = \frac{X}{S_T^0}$$

then Corollary 9.12 gives

$$\mathbb{E}_{\mathbb{Q}}[\hat{X}] \leq \mathbb{E}_{\mathbb{Q}}\big[\hat{V}_T(\varphi^{(n)})\big] \leq \mathbb{E}_{\mathbb{Q}}\big[\hat{V}_0(\varphi^{(n)})\big] \leq \frac{v_n}{S_0^0} \to 0$$

where we have used that S_0^0 is deterministic and positive. Clearly $\hat{X} \geq 0$ implies $\hat{X} = 0$ $\mathbb{Q}$-almost surely, so in particular $\hat{X} = 0$ $\mathbb{P}$-almost surely. We conclude $X = 0$ $\mathbb{P}$-almost surely which gives the contradiction.

$\Longrightarrow$ The general proof for bounded discounted price processes is a consequence of the Hahn-Banach theorem and can be found in Delbaen and Schachermayer (1994). Here we only prove how the general case can be concluded: Instead of S^0 is can be shown that

$$S^\Sigma = \sum_{i=0}^{d} S^i$$

can be chosen as the numeraire as well. Boundedness of

$$\widetilde{S} = \frac{S}{S^\Sigma}$$

proves existence of an equivalent measure $\widetilde{\mathbb{Q}}$ with respect to which $\widetilde{S}$ is a martingale. Now let L denote the density process of $\widetilde{\mathbb{Q}}$ from Theorem 8.7. Theorem 8.9 proves that

$$\widetilde{S}^i L = \frac{S^i}{S^\Sigma} L$$

is a martingale with respect to $\mathbb{P}$ for every i. Due to Condition 8.8 the rescaled process

$$Z = \frac{S_0^\Sigma}{S_0^0} \frac{S^0}{S^\Sigma} L$$

is a positive martingale with respect to $\mathbb{P}$ with

$$\mathbb{E}_{\mathbb{P}}[Z_T] = \mathbb{E}_{\mathbb{P}}[Z_0] = \mathbb{E}_{\mathbb{P}}\big[\frac{S_0^\Sigma}{S_0^0} \frac{S_0^0}{S_0^\Sigma} L_0\big] = \mathbb{E}_{\mathbb{P}}[L_0] = 1$$

where we have used Theorem 2.29. Thus Z is the density process of another probability measure $\mathbb{Q}$, i.e.

$$\frac{d\mathbb{Q}}{d\mathbb{P}}\Big|_{\mathcal{F}_T} = Z_T,$$

and by positivity of S^0 and L (Condition 8.8) one can conclude that $\mathbb{P}$ and $\mathbb{Q}$ are equivalent. Finally, Since for every $i = 1, \ldots, d$

$$\hat{S}^i Z = \frac{S_0^\Sigma}{S_0^0} \frac{S^i}{S^\Sigma} L$$

is a martingale with respect to $\mathbb{P}$ as well we again obtain with Theorem 8.9 that $\hat{S}^i$ is a martingale with respect to $\mathbb{Q}$ for every $i = 1, \ldots, d$. $\square$

Remark 9.17. In discrete mathematical finance the *law of one price* holds which says that $V_T(\varphi)$ completely determines the price process $V(\varphi)$ if φ is self-financing and the market is free of arbitrage. The following example shows that this law does now even hold in continuous mathematical finance when we restrict to self-financing admissible strategies.

Example 9.18. Let φ denote the continuous version of the martingale strategy for which $V_0(\varphi) = 0$ and $V_T(\varphi) = 1$ hold in the setting of Example 9.7. Theorem 8.4(iii) and Theorem 9.16 show that the market is free of arbitrage but the strategy φ is not admissible as losses can become arbitrarily large. $-\varphi$ is admissible, however, and clearly $V_0(-\varphi) = 0$ and $V_T(-\varphi) = -1$. The trivial strategy $\widetilde{\varphi} = (-1, 0)^*$ also satisfies $V_T(\widetilde{\varphi}) = -1$ but $V_0(\widetilde{\varphi}) = -1$.

Definition 9.19.

(i) A self-financing strategy φ is called *double admissible* if φ and $-\varphi$ are both admissible.

(ii) A self-financing strategy φ is called *allowed* if $\hat{V}(\varphi)$ is a martingale with respect to every equivalent martingale measure $\mathbb{Q}$.

Corollary 9.20. *Every double admissible strategy is allowed.*

Proof: By Corollary 9.12 it follows that $\hat{V}(\varphi)$ is both a supermartingale and a submartingale with respect to $\mathbb{Q}$. □

Remark 9.21. While it is intuitively clear that the restriction to admissible strategy makes sense there is no reason to pass on strategies that allow for arbitrarily large gains. From a purely mathematical perspective it is interesting, however, that there is a law of one price for double admissible strategies.

Theorem 9.22. (Law of one price) *Assume that the market is free of arbitrage with respect to allowed strategies, i.e. there does not exists an allowed strategy δ with*

$$V_0(\delta) = 0, \quad V_T(\delta) \geq 0 \quad \text{and} \quad \mathbb{P}(V_T(\delta) > 0) > 0.$$

(i) If φ and ψ are allowed strategies with $V_T(\varphi) = V_T(\psi)$ then $V(\varphi) = V(\psi)$ as processes on $[0, T]$.

(ii) If φ is an allowed strategy and if $V_T(\varphi) = S_T^i$ holds for some $i = 0, \ldots, d$ then $V(\varphi) = S^i$ as processes on $[0, T]$.

Proof:

(i) Suppose there exists some $t \geq 0$ with $\mathbb{P}(V_t(\varphi) \neq V_t(\psi)) > 0$. Without loss of generality we can assume $\mathbb{P}(A) > 0$ with $A = \{V_t(\psi) > V_t(\varphi)\}$ and set

$$(\delta^1, \ldots, \delta^d)_s^* = \begin{cases} 0, & s \leq t, \\ ((\varphi^1, \ldots, \varphi^d)_s^* - (\psi^1, \ldots, \psi^d)_s^*) 1_A, & s > t. \end{cases}$$

From Theorem 9.5 it is clear that there exists a strategy δ^0 such that $\delta = (\delta^0, \ldots, \delta^d)^*$ is self-financing and satisfies $V_0(\delta) = 0$. In particular $\hat{V}_0(\delta) = 0$

and

$$\begin{aligned}\hat{V}_s(\delta) &= \big((\varphi-\psi)^* 1_{A\times(t,T]} \bullet \hat{S}\big)_s = 1_{A\times(t,T]} \bullet \big((\varphi-\psi)^* \bullet \hat{S}\big)_s \\ &= 1_A\big(((\varphi-\psi)^* \bullet \hat{S})_s - ((\varphi-\psi)^* \bullet \hat{S})_t\big) \\ &= 1_A(\hat{V}_s(\varphi) - \hat{V}_s(\psi) - (\hat{V}_t(\varphi) - \hat{V}_t(\psi)))\end{aligned}$$

for all $t < s < T$ where we have used Lemma 5.20. Therefore one can prove

$$\begin{aligned}\mathbb{E}_{\mathbb{Q}}[\hat{V}_T(\delta)|\mathcal{F}_s] &= 1_A\big(\mathbb{E}_{\mathbb{Q}}[\hat{V}_T(\varphi) - \hat{V}_T(\psi)|\mathcal{F}_s] - (\hat{V}_t(\varphi) - \hat{V}_t(\psi))\big) \\ &= 1_A\big(\hat{V}_s(\varphi) - \hat{V}_s(\psi) - (\hat{V}_t(\varphi) - \hat{V}_t(\psi))\big) = \hat{V}_s(\delta)\end{aligned}$$

for all $t < s < T$ and any equivalent martingale measure $\mathbb{Q}$ since φ and ψ are allowed strategies. It is then easy to conclude that $\hat{V}(\delta)$ is a martingale with respect to $\mathbb{Q}$ and δ thus allowed. Setting $s = T$ we obtain

$$\hat{V}_T(\delta) = 1_A(\hat{V}_T(\varphi) - \hat{V}_T(\psi) - (\hat{V}_t(\varphi) - \hat{V}_t(\psi))) = 1_A(\hat{V}_t(\psi) - \hat{V}_t(\varphi)),$$

and the term vanishes on A^c while it is positive on A. Thus δ is an arbitrage opportunity.

(ii) By definition, S^i corresponds to the wealth processes of an allowed strategy. □

Remark 9.23. With a similar proof one can show that

$$V_T(\varphi) \geq V_T(\psi) \implies V(\varphi) \geq V(\psi),$$

holds for allowed strategies φ and ψ if the market is free of arbitrage.

Chapter 10

Valuation and hedging of derivatives

As in discrete mathematical finance we are now interested in random payments at maturity, i.e. we are interested in $\mathcal{F}_T$-measurable random variables X. Our focus specifically is on two things: How can we determine the current price of such a payoff? And how can the risk of an investment be diminished? These questions will be answered within the market model from the previous chapter, and we are additionally assuming that the market is free of arbitrage, i.e. that the condition NFLVR is satisfied. Further let $\mathcal{F}_0$ be trivial (so that S_0 is deterministic), $S^0 > 0$ and let $S^i \geq 0$ for all $i = 1, \ldots, d$.

Definition 10.1.

(i) An $\mathcal{F}_T$-measurable random variable X is called a *payoff*, a *derivative* or an *option*.

(ii) The random variable

$$\hat{X} = \frac{X}{S_T^0}$$

is called the *discounted payoff*.

(iii) A payoff X is called *replicable* if there exists an allowed strategy φ such that $V_T(\varphi) = X$.

Theorem 10.2. *Let X be a non-negative replicable payoff with accompanying strategy φ. Then*

(i) There exists a unique price process S^{d+1} with $S_T^{d+1} = X$ such that the market $(S^0, \ldots, S^{d+1})^*$ is free of arbitrage and satisfies NFLVR, namely $S^{d+1} = V(\varphi)$.

(ii) If $\mathbb{Q}$ is an equivalent martingale measure for $(\hat{S}^0, \ldots, \hat{S}^{d+1})^*$ then $\hat{X}$ is integrable with respect to $\mathbb{Q}$ and we have

$$\hat{S}_t^{d+1} = \mathbb{E}_{\mathbb{Q}}[\hat{X}|\mathcal{F}_t].$$

Proof:

(i) Uniqueness is a consequence of Theorem 9.22.

To prove existence let $\mathbb{Q}$ be an equivalent martingale measure for $\hat{S} = (\hat{S}^0, \dots, \hat{S}^d)^*$. By definition of allowed strategies it is clear that $\hat{S}^{d+1} = \hat{V}(\varphi)$ is a martingale with respect to $\mathbb{Q}$ and $\mathbb{Q}$ thus is an equivalent martingale measure for $(\hat{S}^0, \dots, \hat{S}^{d+1})^*$. Theorem 9.11 proves that the new market is free of arbitrage and even satisfies NFLVR by Theorem 9.16. Finally Remark 9.23 proves that $S_T^{d+1} \geq 0$ implies that the entire process S^{d+1} is non-negative.

(ii) Since $\hat{S}^{d+1}$ is a martingale it follows that $\hat{X} = \hat{S}_T^{d+1}$ is integrable.

$$\hat{S}_t^{d+1} = \mathbb{E}_{\mathbb{Q}}[\hat{S}_T^{d+1} | \mathcal{F}_t] = \mathbb{E}_{\mathbb{Q}}[\hat{X} | \mathcal{F}_t]$$

then finishes the proof. □

Definition 10.3. Let X be a non-negative replicable payoff with accompanying strategy φ. Then S^{d+1} from Theorem 10.2 is called the *(fair) price process* of X and φ is called a *hedging* or *replicating strategy*.

Remark 10.4.

(i) Theorem 10.2 proves that the addition of a replicable option to a market bears additional possibilities for trading but does not lead to the arbitrage in the market.

(ii) A hedging strategy is used to protect against the risk connected with the investment in X. If a bank sells an option with payoff X at time zero and invests at the same time in the self-financing portfolio φ then the wealth at maturity T becomes

$$-X + V_T(\varphi) = 0,$$

i.e. there is no risk of losses.

Example 10.5. A classical example for an option is a *forward contract* where one person gets a certain asset at maturity T (like S^1) while the other gets the amount K (which is agreed upon at time 0) at maturity as well. From the view of the buyer the value of the contract at maturity is

$$X = S_T^1 - K.$$

Such an option can be hedged easily using the base asset S^1 and a bond with maturity T, i.e. with a riskless asset S^0 which has the value 1 at time T. To this end the choose a constant portfolio of the form

$$\varphi = (\varphi_0, \varphi_1)^* = (-K, 1)^*.$$

Then

$$V_T(\varphi) = -KS_T^0 + S_T^1 = X.$$

The fair price K of the forward contract is then chosen in such a way that the wealth process vanishes at time 0, i.e.

$$V_0(\varphi) = -KS_0^0 + S_0^1 = 0 \iff K = \frac{S_0^1}{S_0^0} = \hat{S}_0^1.$$

Otherwise arbitrage opportunities are present, as can e.g. be seen from the perspective of the seller: If the price K is larger than $\hat{S}_0^1$ then the seller buys the replicating portfolio at time 0 for a price of $\hat{S}_0^1$ while the remainder $K - \hat{S}_0^1$ is the profit of the seller. Thus $K = \hat{S}_0^1$ is the fair price of the contract.

One has to be careful, though, because the formal theory of Theorem 10.2 is not directly applicable in this simple example: On one hand, the payoff X is in general not non-negative, on the other hand, it is somewhat unintuitive that continuous trading in a forward contract takes place as suggested by Theorem 10.2.

Definition 10.6.

(i) A payoff X is called *(discountedly) bounded* if $\hat{X}$ is bounded.

(ii) A market is called *complete* if every discountedly bounded payoff is replicable.

Theorem 10.7. (Second fundamental theorem of asset pricing) *The following claims are equivalent:*

(i) The market satisfies the condition NFLVR and is complete.

(ii) There exists a unique equivalent martingale measure $\mathbb{Q}$*.*

In this case every claim X can be replicated which satisfies

$$|X| \leq c \sum_{i=0}^{d} S_T^i$$

for some $c > 0$*.*

Proof:

$\Longrightarrow$ An equivalent martingale measure $\mathbb{Q}$ exists according to Theorem 9.16. To prove uniqueness let $A \in \mathcal{F}$ and $\hat{X} = 1_A$. Clearly $\hat{X}$ is bounded and thus replicable. Theorem 10.2(ii) then proves

$$\mathbb{Q}(A) = \mathbb{E}_{\mathbb{Q}}[\hat{X}] = \mathbb{E}_{\mathbb{Q}}[\hat{X}|\mathcal{F}_0] = \hat{S}_0^{d+1}$$

where S^{d+1} denotes the unique price process. Thus $\mathbb{Q}$ is unique.

$\Longleftarrow$ As in the proof of Theorem 9.16 we choose

$$S^{\Sigma} = \sum_{i=0}^{d} S^i$$

and set

$$\widetilde{S} = \frac{S}{S^{\Sigma}} \quad \text{and} \quad \widetilde{V}(\varphi) = \frac{V(\varphi)}{S^{\Sigma}}.$$

We prove first that the uniqueness condition is independent of the choice of the numeraire: If $\widetilde{L}$ denotes the density process of a measure $\widetilde{\mathbb{Q}} \sim \mathbb{P}$ with respect to which $\widetilde{S}$ is a local martingale then we already know from the proof of Theorem 9.16 that the density process L of a measure $\mathbb{Q} \sim \mathbb{P}$ with respect to which $\hat{S}$ is a local martingale is given by

$$L = \frac{S_0^{\Sigma}}{S_0^0} \frac{S^0}{S^{\Sigma}} \widetilde{L}$$

because $\hat{S}L$ and $\widetilde{S}\widetilde{L}$ are then proportional and $L_0 = \widetilde{L}_0 = 1$ holds. Conversely, one can compute $\widetilde{L}$ starting with L. We thus conclude: If $\mathbb{Q}$ is unique so is $\widetilde{\mathbb{Q}}$, and vice versa. Note also, that if $\hat{X}$ is bounded then so is

$$\widetilde{X} = \frac{X}{S_T^\Sigma}.$$

Thus, let X be a random payoff such that $\widetilde{X}$ is bounded by $c \in \mathbb{R}$, and we set

$$M_t = \mathbb{E}_{\widetilde{\mathbb{Q}}}[\widetilde{X}|\mathcal{F}_t]$$

to be the $\widetilde{\mathbb{Q}}$ martingale generated by $\widetilde{X}$. We now apply the following result which we have not proven in this lecture:

There exists an extension of the martingale representation theorem (Theorem 8.6) as follows: As $\widetilde{X}$ is $\mathcal{F}_T$-measurable and as $\widetilde{\mathbb{Q}}$ is the only probability measure so that $\widetilde{S}$ is martingale with respect to $(\mathcal{F}_t)_{0\le t\le T}$, there exists some $H \in \mathcal{L}(\widetilde{S})$ such that

$$\widetilde{X} = M_T = M_0 + (H^* \bullet \widetilde{S})_T. \tag{10.1}$$

Since

$$\sum_{i=0}^d \widetilde{S}^i = \frac{\sum_{i=0}^d S^i}{\sum_{i=0}^d S^i} = 1$$

holds we have

$$(J^* \bullet \widetilde{S})_t = \sum_{i=0}^d (K \bullet \widetilde{S}^i)_t = (K \bullet \sum_{i=0}^d \widetilde{S}^i)_t = 0$$

for every integrable $J = (K, \dots, K)^*$. As a consequence $H + J$ then satisfies (10.1). Choosing specifically

$$K_t = M_0 + (H^* \bullet \widetilde{S})_t - H_t^* \widetilde{S}_t$$

we obtain analogously to the proof of Theorem 9.5 that the portfolio

$$\varphi = H + (K, \dots, K)^*$$

if self-financing according to

$$\varphi_t^* \widetilde{S}_t - (M_0 + (\varphi^* \bullet \widetilde{S})_t) = H_t^* \widetilde{S}_t + K_t - (M_0 + (H^* \bullet \widetilde{S})_t) = 0$$

and with $\widetilde{V}_0(\varphi) = M_0$. Thus

$$\widetilde{X} = \widetilde{V}_0(\varphi) + (\varphi^* \bullet \widetilde{S})_T = \widetilde{V}_T(\varphi) \quad \text{and} \quad X = \widetilde{X} S_T^\Sigma = \widetilde{V}_T(\varphi) S_T^\Sigma = V_T(\varphi).$$

Because of $|\widetilde{V}(\varphi)| = |M| \le c$ we finally obtain

$$V(\varphi) = \widetilde{V}(\varphi) S^\Sigma \ge -cS^\Sigma \quad \text{and} \quad V(-\varphi) = -\widetilde{V}(\varphi) S^\Sigma \ge -cS^\Sigma$$

so that φ is double admissible and allowed due to Corollary 9.20. Thus X is replicable and the claim follows from Theorem 9.16. □

Definition 10.8. Let I denote the identity process, let B denote a Brownian motion, and assume that there exist constants $S_0^1 > 0$, $\mu \in \mathbb{R}$, $\sigma \neq 0$ and $\widetilde{\mu} = \mu + \sigma^2/2$. The market $S = (S^0, S^1)^*$ with one *bond*

$$S_t^0 = \exp(rt), \ \ r \in \mathbb{R},$$

and one *stock*

$$S_t^1 = S_0^1 \exp(\mu t + \sigma B_t) = S_0^1 \mathcal{E}(\widetilde{\mu} I + \sigma B)_t$$

is called the *Black-Scholes model (with one stock).*

We will work with the Black-Scholes model in the following, and we assume additionally that the filtration $(\mathcal{F}_t)_{0 \leq t \leq T}$ is generated by S^1.

Theorem 10.9.

(i) An equivalent martingale measure $\mathbb{Q} \sim \mathbb{P}$ is defined via

$$\frac{d\mathbb{Q}}{d\mathbb{P}}\Big|_{\mathcal{F}_T} = \mathcal{E}\big(-\frac{\widetilde{\mu} - r}{\sigma} B\big)_T.$$

(ii) It holds

$$\hat{S}_t^1 = \hat{S}_0^1 \mathcal{E}(\sigma \widetilde{B})_t$$

where

$$\widetilde{B}_t = B_t + \frac{\widetilde{\mu} - r}{\sigma} t$$

is a Brownian motion with respect to $\mathbb{Q}$.

(iii) The market satisfies NFLVR and is complete.

Proof:

(i) and (ii) Using Theorem 8.4(iii) we know that

$$Z_t = \mathcal{E}\big(-\frac{\widetilde{\mu} - r}{\sigma} B\big)_t = \exp\Big(-\frac{\widetilde{\mu} - r}{\sigma} B_t - \frac{1}{2}\big(\frac{\widetilde{\mu} - r}{\sigma}\big)^2 t\Big)$$

is a martingale since the Novikov condition is satisfied. From $Z_0 = 1$ we obtain that Z is the density process of a probability measure $\mathbb{Q}$ with $\mathbb{Q} \sim \mathbb{P}$. Corollary 8.11 then proves that

$$\widetilde{B}_t = B_t + \frac{\widetilde{\mu} - r}{\sigma} t$$

is a Brownian motion with respect to $\mathbb{Q}$. Also

$$\begin{aligned} \hat{S}_t^1 &= \hat{S}_0^1 \exp((\mu - r)t + \sigma B_t) = \hat{S}_0^1 \exp((\mu - \widetilde{\mu})t + \sigma \widetilde{B}_t) \\ &= \hat{S}_0^1 \exp\big(\sigma \widetilde{B}_t - \frac{\sigma^2}{2} t\big) = \hat{S}_0^1 \mathcal{E}(\sigma \widetilde{B})_t. \end{aligned} \tag{10.2}$$

Since $\widetilde{B}$ is a Brownian motion with respect to $\mathbb{Q}$ we again have from Theorem 8.4(iii) that $\hat{S}_t^1$ is a martingale with respect to $\mathbb{Q}$.

(iii) Note first that NFLVR holds as a consequence of Theorem 9.16.

Furthermore, because of

$$S_t^1 = S_0^1 \exp\Big((r - \frac{\sigma^2}{2})t + \sigma \widetilde{B}_t\Big) \tag{10.3}$$

it is clear that the filtration $(\mathcal{F}_t)_{0\le t\le T}$ is generated by the process $\widetilde{B}$ as well. Let now X be a payoff with $|\hat{X}| \le c$. Then the martingale representation theorem (Theorem 8.6) proves the existence of some $H \in \mathcal{L}^2(\widetilde{B})$ with

$$\hat{X} = \mathbb{E}_{\mathbb{Q}}[\hat{X}] + (H \bullet \widetilde{B})_T.$$

In particular $H \bullet \widetilde{B}$ is a martingale. Thus

$$\hat{S}_t^2 = \mathbb{E}_{\mathbb{Q}}[\hat{X}|\mathcal{F}_t]$$

has the representation

$$\hat{S}_t^2 = \hat{S}_0^2 + (H \bullet \widetilde{B})_t$$

and satisfies $|\hat{S}^2| \le c$. We now set

$$\varphi^1 = \frac{H}{\sigma \hat{S}^1}.$$

Then

$$\hat{S}_0^2 + \varphi^1 \bullet \hat{S}^1 = \hat{S}_0^2 + \varphi^1 \hat{S}^1 \bullet (\sigma \widetilde{B}) = \hat{S}_0^2 + H \bullet \widetilde{B} = \hat{S}^2,$$

where we have used (10.2) and Theorem 8.4(i) first as well as Theorem 6.11 afterwards. If then φ is the self-financing strategy with φ^1 as above and starting capital $\hat{S}_0^2$, which exists according to Theorem 9.5, then

$$\hat{V}_T(\varphi) = \hat{S}_T^2 = \hat{X} \quad \text{with} \quad |\hat{V}_t(\varphi)| \le c \le c(\hat{S}_t^0 + \hat{S}_t^1)$$

is satisfied. We conclude that φ is double admissible, thus allowed. The market then is complete. □

Theorem 10.10. *Suppose that we are given a payoff of the form $X = g(S_T^1)$ where $g : [0,\infty) \to [0,\infty)$ satisfies $g(x) \le c(1+x)$ for some $c > 0$. Then*

(i) X is replicable.

(ii) The fair price process S^2 of X is given by $S_t^2 = f(t, S_t^1)$ with

$$f(t,x) = \exp(-r(T-t)) \int g(\exp(z))\varphi_{\log(x)+(r-\frac{\sigma^2}{2})(T-t),\sigma^2(T-t)}(z)dz$$

where φ_{μ,σ^2} denotes the density of a $\mathcal{N}(\mu,\sigma^2)$ distribution.

(iii) f is twice continuously differentiable on $(0,T)\times(0,\infty)$ and satisfies the partial differential equation

$$D_1 f(t,x) + rxD_2 f(t,x) + \frac{1}{2}\sigma^2 x^2 D_{22} f(t,x) - rf(t,x) = 0.$$

(iv) *The portfolio given by*

$$\varphi_t^0 = \exp(-rt)(f(t,S_t^1) - D_2 f(t,S_t^1)S_t^1),$$
$$\varphi_t^1 = D_2 f(t,S_t^1)$$

is a hedging strategy.

Proof:

(i) Because of

$$0 \le X \le c(1+S_T^1) \le C(S_T^0 + S_T^1)$$

with $C = c\max(\exp(-rT),1)$ we obtain that X is replicable due to Theorem 10.9 and Theorem 10.7.

(ii) We have

$$\begin{aligned}\hat S_t^2 &= \mathbb{E}_{\mathbb{Q}}[\hat X|\mathcal{F}_t] = \exp(-rT)\mathbb{E}_{\mathbb{Q}}[g(S_T^1)|\mathcal{F}_t]\\ &= \exp(-rT)\mathbb{E}_{\mathbb{Q}}\Big[g\big(S_t^1\exp\big((r-\frac{\sigma^2}{2})(T-t)+\sigma(\widetilde B_T-\widetilde B_t)\big)\big)|\mathcal{F}_t\Big]\end{aligned}$$

from (10.3). Since $\widetilde B_T - \widetilde B_t$ is a Brownian motion under $\mathbb{Q}$ and independent of $\mathcal{F}_t$ we obtain with the substitution

$$S_t^1\exp\big((r-\frac{\sigma^2}{2})(T-t)+\sigma y\big) = \exp(z)$$

that

$$\begin{aligned}\hat S_t^2 &= \exp(-rT)\int g\big(S_t^1\exp\big((r-\frac{\sigma^2}{2})(T-t)+\sigma y\big)\big)\varphi_{0,T-t}(y)dy\\ &= \exp(-rT)\int g(\exp(z))\varphi_{\log(S_t^1)+(r-\frac{\sigma^2}{2})(T-t),\sigma^2(T-t)}(z)dz\end{aligned}$$

holds. $S_t^2 = \exp(rt)\hat S_t^2$ gives the claim.

(iii) The growth condition on g along with dominated convergence shows that derivatives and integrals can be exchanged. The first claim then follows because the density of a normal distribution is infinitely often differentiable in its parameters.

Also,

$$\hat S_t^2 = \hat f(t,\hat S_t^1) \quad \text{with} \quad \hat f(t,x) = \exp(-rt)f(t,x\exp(rt))$$

from (ii), and the Itō formula gives

$$\hat S_t^2 = \hat S_0^2 + \int_0^t D_1\hat f(s,\hat S_s^1)ds + \int_0^t D_2\hat f(s,\hat S_s^1)d\hat S_s^1 + \frac{1}{2}\int_0^t D_{22}\hat f(s,\hat S_s^1)d[\hat S^1]_s.$$

Because of $\hat S^1 = \mathcal{E}(\sigma\widetilde B)$ from Theorem 10.9(ii) and because of Theorem 8.4(i) we have

$$d[\hat S^1]_s = \sigma^2(\hat S_s^1)^2 d[\widetilde B]_s = \sigma^2(\hat S_s^1)^2 ds.$$

Thus

$$\hat S_t^2 = \hat S_0^2 + \int_0^t \big(D_1\hat f(s,\hat S_s^1) + \frac{1}{2}D_{22}\hat f(s,\hat S_s^1)\sigma^2(\hat S_s^1)^2\big)ds + \int_0^t D_2\hat f(s,\hat S_s^1)d\hat S_s^1.$$

Clearly, as both $\hat{S}^1$ and $\hat{S}^2$ are martingales under $\mathbb{Q}$ we have that

$$\int_0^t \big(D_1\hat{f}(s,\hat{S}^1_s) + \frac{1}{2}D_{22}\hat{f}(s,\hat{S}^1_s)\sigma^2(\hat{S}^1_s)^2\big)ds = \hat{S}^2_t - \hat{S}^2_0 - \int_0^t D_2\hat{f}(s,\hat{S}^1_s)d\hat{S}^1_s \tag{10.4}$$

is a continuous local martingale with respect to $\mathbb{Q}$ which is of bounded variation. Thus from Theorem 7.9

$$\int_0^t \big(D_1\hat{f}(s,\hat{S}^1_s) + \frac{1}{2}D_{22}\hat{f}(s,\hat{S}^1_s)\sigma^2(\hat{S}^1_s)^2\big)ds = 0$$

$\mathbb{Q}$-almost surely and then $\mathbb{P}$-almost surely, for all $t \geq 0$. In particular

$$D_1\hat{f}(s,\hat{S}^1_s) + \frac{1}{2}D_{22}\hat{f}(s,\hat{S}^1_s)\sigma^2(\hat{S}^1_s)^2 = 0$$

for λ-almost all $s \in [0,T]$, and by continuity of the term we obtain the identity actually for all $s \in [0,T]$. This means that for a fixed s we have

$$D_1\hat{f}(s,x) + \frac{1}{2}D_{22}\hat{f}(s,x)\sigma^2x^2 = 0$$

for $\mathbb{P}^{\hat{S}^1_s}$-almost all $x > 0$. Since $\hat{S}^1_s$ has a density, this identity first holds for λ-almost all $x > 0$ and then by continuity for all $x > 0$. The multidimensional chain rule then yields

$$\begin{aligned} D_1\hat{f}(s,x) &= -r\exp(-rs)f(s,x\exp(rs)) \\ &\quad + \exp(-rs)\big(D_1f(s,x\exp(rs)) + \exp(rs)rxD_2f(s,x\exp(rs))\big) \end{aligned}$$

and

$$D_2\hat{f}(s,x) = D_2f(s,x\exp(rs)) \quad \text{and} \quad D_{22}\hat{f}(s,x) = \exp(rs)D_{22}f(s,x\exp(rs)).$$

Setting $\tilde{x} = \exp(rs)x$ we finally obtain

$$\begin{aligned} 0 &= D_1\hat{f}(s,x) + \frac{1}{2}D_{22}\hat{f}(s,x)\sigma^2x^2 \\ &= \exp(-rs)\big(-rf(s,\hat{x}) + D_1f(s,\hat{x}) + r\hat{x}D_2f(s,\hat{x}) + \frac{1}{2}\sigma^2\hat{x}^2D_{22}f(s,\hat{x})\big). \end{aligned}$$

(iv) As a consequence of (10.4) we have

$$\hat{S}^2_t = \hat{S}^2_0 + (D_2\hat{f}(I,\hat{S}^1) \bullet \hat{S}^1)_t$$

where we have used $I(t) = t$ again. For $\varphi^1_t = D_2\hat{f}(t,\hat{S}^1_t)$ and initial capital S^2_0 we know from Theorem 9.5 that there exists a self-financing strategy with

$$\hat{S}^2 = \hat{S}^2_0 + \varphi^1 \bullet \hat{S}^1 = \hat{V}(\varphi).$$

The definition of S^2 then gives

$$0 \leq V(\varphi) = S^2 \leq C(S^0 + S^1),$$

so that φ is double admissible and thus a replicating strategy. With

$$\varphi^1_t = D_2\hat{f}(t,\hat{S}^1_t) = D_2f(t,S^1_t)$$

and

$$\varphi^0_t = \hat{V}_t(\varphi) - \varphi^1_t\hat{S}^1_t = \exp(-rt)(f(t,S^1_t) - D_2f(t,S^1_t)S^1_t)$$

the claim follows. □

Remark 10.11. In general, price process and hedging strategy cannot be stated explicitly as the integral can not always be solved and f is thus not always known. In special cases, as for call and put options, an explicit computation is possible.

Theorem 10.12. (Black-Scholes formula) *Let $X = (S_T^1 - K)^+$ for some $K > 0$ be a call option. Then the fair price process S^2 of X satisfies the identity*

$$S_t^2 = S_t^1 \Phi\Big(\frac{\log\big(\frac{S_t^1}{K}\big) + \big(r + \frac{\sigma^2}{2}\big)(T-t)}{\sigma\sqrt{T-t}}\Big)$$
$$- \exp(-r(T-t))K\Phi\Big(\frac{\log(\frac{S_t^1}{K}) + \big(r - \frac{\sigma^2}{2}\big)(T-t)}{\sigma\sqrt{T-t}}\Big)$$

where Φ denotes the distribution function of a standard normal distribution. As a hedging strategy one obtains

$$\varphi_t^0 = -\exp(-rT)K\Phi\Big(\frac{\log(\frac{S_t^1}{K}) + \big(r - \frac{\sigma^2}{2}\big)(T-t)}{\sigma\sqrt{T-t}}\Big),$$
$$\varphi_t^1 = \Phi\Big(\frac{\log\big(\frac{S_t^1}{K}\big) + \big(r + \frac{\sigma^2}{2}\big)(T-t)}{\sigma\sqrt{T-t}}\Big).$$

Proof: Set $g(x) = (x-K)^+$ in Theorem 10.10(ii). Then the substitution

$$z = \sigma\sqrt{T-t}y + \log(x) + \Big(r - \frac{\sigma^2}{2}\Big)(T-t)$$

gives

$$\begin{aligned}
f(t,x) &= \exp(-r(T-t)) \int g(\exp(z))\varphi_{\log(x)+(r-\frac{\sigma^2}{2})(T-t),\sigma^2(T-t)}(z)dz \\
&= \exp(-r\tau) \int g(x\exp(\widetilde{r}\tau + \sigma\sqrt{\tau}y))\varphi(y)dy \\
&= \exp(-r\tau) \int_w^\infty (x\exp(\widetilde{r}\tau + \sigma\sqrt{\tau}y) - K)\varphi(y)dy \\
&= x\exp\big(-\frac{\sigma^2}{2}\tau\big) \int_w^\infty \exp(\sigma\sqrt{\tau}y)\varphi(y)dy - \exp(-r\tau)K(1-\Phi(w)) \\
&= x\int_w^\infty \frac{1}{\sqrt{2\pi}}\exp\Big(-\frac{(y-\sigma\sqrt{\tau})^2}{2}\Big)dy - \exp(-r\tau)K\Phi(-w) \\
&= x\Phi(-w+\sigma\sqrt{\tau}) - \exp(-r\tau)K\Phi(-w)
\end{aligned}$$

where we have set $\tau = T-t$, $\widetilde{r} = r - \sigma^2/2$, $w = \frac{\log(\frac{K}{x}) - \widetilde{r}\tau}{\sigma\sqrt{\tau}}$ and $\varphi = \varphi_{0,1}$. Obviously,

$$-w + \sigma\sqrt{\tau} = \frac{\log(\frac{x}{K}) + (\widetilde{r} + \sigma^2)\tau}{\sigma\sqrt{\tau}} = \frac{\log(\frac{x}{K}) + \big(r + \frac{\sigma^2}{2}\big)(T-t)}{\sigma\sqrt{T-t}}.$$

We also get from Theorem 10.10(iv)

$$\begin{aligned}\varphi_t^1 = D_2 f(t, S_t^1) = \Phi\Bigg(&\frac{\log\left(\frac{S_t^1}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}}\Bigg)\\ &+ \frac{1}{\sigma\sqrt{T-t}}\Phi'\Bigg(\frac{\log\left(\frac{S_t^1}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}}\Bigg)\\ &- \frac{\exp(-r(T-t))K}{\sigma\sqrt{T-t}S_t^1}\Phi'\Bigg(\frac{\log\left(\frac{S_t^1}{K}\right) + \left(r - \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}}\Bigg).\end{aligned}$$

Because of

$$\Phi'(x) = \varphi(x) = \frac{1}{\sqrt{2\pi}}\exp(-x^2/2)$$

we have

$$\frac{\Phi'(a+b)}{\Phi'(a-b)} = \exp(-2ab).$$

Applying this identity for

$$a = \frac{\log\left(\frac{S_t^1}{K}\right) + r(T-t)}{\sigma\sqrt{T-t}} \quad \text{and} \quad b = \frac{\sigma\sqrt{T-t}}{2}$$

with

$$\exp(-2ab) = \frac{\exp(-r(T-t))K}{S_t^1}$$

the representation for φ^1 follows. Additionally,

$$\begin{aligned}\varphi_t^0 &= \exp(-rt)(f(t, S_t^1) - D_2 f(t, S_t^1)S_t^1)\\ &= -\exp(-rT)K\Phi\Bigg(\frac{\log\left(\frac{S_t^1}{K}\right) + \left(r - \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}}\Bigg).\end{aligned}$$

□

Corollary 10.13. *Let $X = (K - S_T^1)^+$ for some $K > 0$ be a put option. Then the fair price process S^2 of X satisfies*

$$\begin{aligned}S_t^2 = \exp(-r(T-t))K\Phi\Bigg(&\frac{\log\left(\frac{K}{S_t^1}\right) - \left(r - \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}}\Bigg)\\ &- S_t^1\Phi\Bigg(\frac{\log\left(\frac{K}{S_t^1}\right) - \left(r + \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}}\Bigg).\end{aligned}$$

As a hedging strategy one obtains

$$\begin{aligned}\varphi_t^0 &= \exp(-rT)K\Phi\Bigg(\frac{\log\left(\frac{K}{S_t^1}\right) - \left(r - \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}}\Bigg),\\ \varphi_t^1 &= -\Phi\Bigg(\frac{\log\left(\frac{K}{S_t^1}\right) - \left(r + \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}}\Bigg).\end{aligned}$$

Proof: The claim follows as a simple exercise using the *put call parity*

$$(K - S_T^1)^+ = (S_t^1 - K)^+ - S_T^1 + K\exp(-rT)S_T^0$$

and the fair valuation of a call option from Theorem 10.12. □

Remark 10.14.

(i) As a consequence of Theorem 10.12 and Corollary 10.13 it is only the *interest rate* r and the *volatility* σ that is included explicity in the pricing formulas for call and put options, and not the *drift rate* μ. This may appear unintuitive as μ has an enormous influence on whether the option has a value at maturity or not. Note, however, that the price of the option S^2 is always computed relatively to the current price of S^1. If S^1 grows, as one could expect for large values of μ, the price S^2 of a call option automatically grows as well. This phenomenon is visible for many other options as well.

(ii) If one ignores the interest rate, which is often negligible in short time periods, the main parameter in many applications is the volatility which can be estimated in many different ways. An elegant way is to determine the *implied volatility* which means that one uses the price of traded call options in the market to solve the Black-Scholes formula for σ. In this case one often sees that the implied volatility depends crucially on the choice of K and $T-t$. This observation is one reason why people believe that actual prices in reality do not coincide with the Black-Scholes model.

Definition 10.15. Let X be a payoff and let φ be a replicating portfolio with the wealth process

$$V_t(\varphi) = f(t, S_t^1, r, \sigma).$$

Then the partial derivatives

$$\begin{aligned}
\Theta &= D_1 f(t, S_t^1, r, \sigma), \\
\Delta &= D_2 f(t, S_t^1, r, \sigma), \\
\Gamma &= D_{22} f(t, S_t^1, r, \sigma), \\
\rho &= D_3 f(t, S_t^1, r, \sigma), \\
\nu &= D_4 f(t, S_t^1, r, \sigma)
\end{aligned}$$

are called the *greeks*.

Remark 10.16.

(i) The terms Θ, Δ and Γ appear in the *Black-Scholes differential equation* from Theorem 10.10(iii) and explain how much the price process of the option changes with time and with the stock price. Of special importance is in particular Δ which coincides, using Theorem 10.10(iv), with the amount of shares of the stock in the hedging portfolio. Ideally one then wants to have a small absolute value of Γ as otherwise one needs to change the portfolio often which leads to transaction costs in financial practice.

(ii) The terms ρ and ν (the latter often denoted as *vega*) explain how sensitive the price process is with respect to the (typically unknown) model parameters r and σ. ρ usually does not play an important role, for the reasons laid out before.

Finally we discuss some examples for options which are not of the form $X = g(S_T^1)$. To this end we need the following auxiliary result.

Theorem 10.17. *Let B be a Brownian motion, $\mu \in \mathbb{R}$ and $\sigma > 0$, and let $X_t = \mu t + \sigma B_t$ as well as $M_t = \max_{s\in[0,t]} X_s \geq 0$. Then, for any $t > 0$, $(X_t, M_t)^*$ has the density*

$$f(x,m) = \frac{2(2m-x)}{\sqrt{2\pi}\sigma^3 t^{3/2}} \exp\Big(-\frac{(x-2m)^2 - 2\mu x t + (\mu t)^2}{2\sigma^2 t}\Big) 1_{\{x\leq m\}},$$

and we have

$$\mathbb{P}(X_t \leq x, M_t \leq m) = \Phi\Big(\frac{x-\mu t}{\sigma\sqrt{t}}\Big) - \exp\big(\frac{2m\mu}{\sigma^2}\big)\Phi\Big(\frac{x-2m-\mu t}{\sigma\sqrt{t}}\Big)$$

for any $x \leq m$.

Proof: Let first $\mu = 0$ and $\sigma = 1$. Then the reflection principle (Theorem 1.33) proves that the process

$$\widetilde{B}_t = B_t 1_{\{t\leq\tau_a\}} + (2B_{\tau_a} - B_t)1_{\{t>\tau_a\}}$$

with $\tau_a = \inf\{t \geq 0 \mid B_t \geq a\}$ and $a > 0$ is a Brownian motion as well. Setting $S_t = \max_{s\in[0,t]} B_s$ and $\widetilde{\tau}_a = \inf\{t \geq 0 \mid \widetilde{B}_t \geq a\}$ we obtain

$$\{S_t \geq a\} = \{\tau_a \leq t\} = \{\widetilde{\tau}_a \leq t\}$$

and

$$\begin{aligned}
\mathbb{P}(B_t \leq a-y, S_t \geq a) &= \mathbb{P}(B_t \leq a-y, \tau_a \leq t) = \mathbb{P}(\widetilde{B}_t \leq a-y, \widetilde{\tau}_a \leq t)\\
&= \mathbb{P}(2a - B_t \leq a-y, \tau_a \leq t) = \mathbb{P}(B_t \geq a+y, \tau_a \leq t)\\
&= \mathbb{P}(B_t \geq a+y)
\end{aligned}$$

for any $y \geq 0$ where we have used in the second step that $(B,\tau_a)^*$ and $(\widetilde{B},\widetilde{\tau}_a)^*$ share the same distribution. Thus we obtain for any $x \leq y$

$$\mathbb{P}(B_t \leq x, S_t \geq y) = \mathbb{P}(B_t \geq 2y - x) = 1 - \Phi\Big(\frac{2y-x}{\sqrt{t}}\Big),$$

and by computation of the second derivative we obtain the density of $(B_t, S_t)^*$ via

$$\begin{aligned}
g(x,y) &= -\frac{\partial^2}{\partial y\partial x}\mathbb{P}(B_t \leq x, S_t \geq y) = -\frac{\partial}{\partial y}\frac{1}{\sqrt{2\pi t}}\exp\Big(-\frac{(2y-x)^2}{2t}\Big)\\
&= \sqrt{\frac{2}{\pi t^3}}(2y-x)\exp\Big(-\frac{(2y-x)^2}{2t}\Big)
\end{aligned}$$

for any pair $x \leq y$.

In the next step we still assume $\sigma = 1$ and define a probability measure $\mathbb{Q} \sim \mathbb{P}$ via

$$\frac{d\mathbb{Q}}{d\mathbb{P}}\Big|_{\mathcal{F}_T} = \mathcal{E}(-\mu B)_T = \exp\Big(-\mu B_T - \frac{1}{2}\mu^2 T\Big).$$

Then Corollary 8.11 proves that $X_t = \mu t + B_t$ is a Brownian motion with respect to

$\mathbb{Q}$, and from Theorem 8.7 we obtain

$$\begin{aligned}
&\mathbb{P}(X_t \leq x, M_t \leq m) = \mathbb{E}_{\mathbb{Q}}\Big[\frac{d\mathbb{P}}{d\mathbb{Q}}\big|_{\mathcal{F}_T} 1_{\{X_t \leq x, M_t \leq m\}}\Big]\\
&= \mathbb{E}_{\mathbb{Q}}\Big[\exp\Big(\mu X_t - \frac{1}{2}\mu^2 t\Big) 1_{\{X_t \leq x, M_t \leq m\}}\Big]\\
&= \int_{-\infty}^{x}\int_0^m \exp\Big(\mu z - \frac{1}{2}\mu^2 t\Big) g(z,y)\,dy\,dz\\
&= \frac{1}{\sqrt{2\pi t}}\int_{-\infty}^{x} \exp\Big(\mu z - \frac{1}{2}\mu^2 t\Big)\Big(\exp\Big(-\frac{z^2}{2t}\Big) - \exp\Big(-\frac{(2m-z)^2}{2t}\Big)\Big)dz\\
&= \frac{1}{\sqrt{2\pi t}}\int_{-\infty}^{x} \Big(\exp\Big(-\frac{(z-\mu t)^2}{2t}\Big) - \exp(2m\mu)\exp\Big(-\frac{(z-2m-\mu t)^2}{2t}\Big)\Big)dz\\
&= \Phi\Big(\frac{x-\mu t}{\sqrt{t}}\Big) - \exp(2m\mu)\Phi\Big(\frac{x-2m-\mu t}{\sqrt{t}}\Big)
\end{aligned}$$

for $x \leq m$.

We finally discuss the general case and use that the distribution of $X_t = \mu t + \sigma B_t$ and $M_t = \max_{s\in[0,t]} X_s$ coincides with the distribution of $(\overline{X}_{\sigma^2 t}, \overline{M}_{\sigma^2 t})^*$ where

$$\overline{X}_t = \frac{\mu}{\sigma^2}t + B_t \quad \text{and} \quad \overline{M}_t = \max_{s\in[0,t]} \overline{X}_s.$$

Thus

$$\mathbb{P}(X_t \leq x, M_t \leq m) = \Phi\Big(\frac{x-\mu t}{\sigma\sqrt{t}}\Big) - \exp\big(\frac{2m\mu}{\sigma^2}\big)\Phi\Big(\frac{x-2m-\mu t}{\sigma\sqrt{t}}\Big)$$

for $x \leq m$ as claimed. Taking second derivatives again we obtain

$$\begin{aligned}
&f(x,m) = \frac{\partial^2}{\partial x\partial m}\mathbb{P}(X_t \leq x, M_t \leq m)\\
&= \exp\big(\frac{2m\mu}{\sigma^2}\big)\frac{\partial}{\partial x}\Big(-\frac{2\mu}{\sigma^2}\Phi\Big(\frac{x-2m-\mu t}{\sigma\sqrt{t}}\Big) + \frac{2}{\sqrt{2\pi\sigma^2 t}}\exp\Big(-\frac{(x-2m-\mu t)^2}{2\sigma^2 t}\Big)\Big)\\
&= -\frac{2\mu}{\sigma^2}\frac{1}{\sqrt{2\pi\sigma^2 t}}\exp\big(\frac{2m\mu}{\sigma^2}\big)\exp\Big(-\frac{(x-2m-\mu t)^2}{2\sigma^2 t}\Big)\\
&\quad - \frac{2(x-2m-\mu t)}{\sqrt{2\pi\sigma^2 t}\sigma^2 t}\exp\big(\frac{2m\mu}{\sigma^2}\big)\exp\Big(-\frac{(x-2m-\mu t)^2}{2\sigma^2 t}\Big)\\
&= \frac{2(2m-x)}{\sqrt{2\pi}\sigma^3 t^{3/2}}\exp\Big(-\frac{(x-2m)^2 - 2\mu x t + (\mu t)^2}{2\sigma^2 t}\Big),
\end{aligned}$$

hence the claim. □

Example 10.18. *Barrier options* gain or lose their value once the underlying stock price reaches a certain threshold. In the following we will focus on a *down-and-out call* which loses its value once the stock price becomes too small. Analogously to the put call parity one obtains the value of similar barrier options (*down-and-in*, *up-and-out*, *up-and-in*) by symmetry arguments.

Theorem 10.19. *Let* $0 \leq H \leq K$ *and*

$$X = \begin{cases} (S_T^1 - K)^+, & \textit{if } S_t^1 > H \textit{ for all } t \in [0,T],\\ 0, & \textit{else,}\end{cases}$$

a down-and-out call. Then the fair price process S^2 of X takes the form

$$S_t^2 = f(t, S_t^1)1_{\{\min_{s\le t} S_s^1 \ge H\}}$$

with

$$\begin{aligned}
&f(t,x)\\
&= x\Big(\Phi\Big(\frac{\log\big(\frac{x}{K}\big) + \big(r + \frac{\sigma^2}{2}\big)(T-t)}{\sigma\sqrt{T-t}}\Big) - \Big(\frac{H}{x}\Big)^{1+\frac{2r}{\sigma^2}}\Phi\Big(\frac{\log\big(\frac{H^2}{Kx}\big) + \big(r + \frac{\sigma^2}{2}\big)(T-t)}{\sigma\sqrt{T-t}}\Big)\Big)\\
&\quad - K\exp(-r(T-t))\Big(\Phi\Big(\frac{\log\big(\frac{x}{K}\big) + \big(r - \frac{\sigma^2}{2}\big)(T-t)}{\sigma\sqrt{T-t}}\Big) - \Big(\frac{H}{x}\Big)^{\frac{2r}{\sigma^2}-1}\Phi\Big(\frac{\log\big(\frac{H^2}{Kx}\big) + \big(r - \frac{\sigma^2}{2}\big)(T-t)}{\sigma\sqrt{T-t}}\Big)\Big).
\end{aligned}$$

A hedging strategy is given by

$$\begin{aligned}
\varphi_t^0 &= \exp(-rt)(f(t,S_t^1) - D_2 f(t,S_t^1)S_t^1)1_{\{\min_{s\le t} S_s^1 \ge H\}},\\
\varphi_t^1 &= D_2 f(t,S_t^1)1_{\{\min_{s\le t} S_s^1 \ge H\}}.
\end{aligned}$$

Proof: Since the market satisfies NFLVR by Theorem 10.9(iii) and is complete as well we know that

$$X \le (S_T^1 - K)^+ \le S_T^1$$

is replicable by Theorem 10.7. Theorem 10.2(ii) then gives

$$\begin{aligned}
\hat{S}_t^2 &= \mathbb{E}_{\mathbb{Q}}[\hat{X}|\mathcal{F}_t] = \exp(-rT)\mathbb{E}_{\mathbb{Q}}\big[(S_T^1 - K)^+ 1_{\{\min_{s\le T} S_s^1 \ge H\}}|\mathcal{F}_t\big] \\
&= \exp(-rT)\mathbb{E}_{\mathbb{Q}}\big[S_T^1 1_{\{S_T^1 \ge K, \min_{s\le T} S_s^1 \ge H\}}|\mathcal{F}_t\big]\\
&\quad - \exp(-rT)K\mathbb{E}_{\mathbb{Q}}\big[1_{\{S_T^1 \ge K, \min_{s\le T} S_s^1 \ge H\}}|\mathcal{F}_t\big].
\end{aligned} \tag{10.5}$$

Let us first discuss the second term in (10.5). We set

$$X_t = \big(r - \frac{\sigma^2}{2}\big)t + \sigma\widetilde{B}_t$$

and use that $\widetilde{B}$ is a Brownian motion under $\mathbb{Q}$. From (10.3) we obtain

$$S_T^1 = S_t^1\exp(X_T - X_t)$$

so that Theorem 10.17 gives

$$\begin{aligned}
&\mathbb{E}_{\mathbb{Q}}\Big[1_{\{S_T^1 \ge K, \min_{s\le T} S_s^1 \ge H\}}\Big|\mathcal{F}_t\Big]\\
&= 1_{\{\min_{s\le t} S_s^1 \ge H\}}\mathbb{E}_{\mathbb{Q}}\big[1_{\big\{-(X_T - X_t)\le \log\big(\frac{S_t^1}{K}\big), \max_{t\le s\le T}(-(X_s - X_t))\le \log\big(\frac{S_t^1}{H}\big)\big\}}|\mathcal{F}_t\big]\\
&= 1_{\{\min_{s\le t} S_s^1 \ge H\}}\Big(\Phi\Big(\frac{\log\big(\frac{S_t^1}{K}\big) + \widetilde{r}\tau}{\sigma\sqrt{\tau}}\Big) - \Big(\frac{S_t^1}{H}\Big)^{1-\frac{2r}{\sigma^2}}\Phi\Big(\frac{\log\big(\frac{S_t^1}{K}\big) - 2\log\big(\frac{S_t^1}{H}\big) + \widetilde{r}\tau}{\sigma\sqrt{\tau}}\Big)\Big)\\
&= 1_{\{\min_{s\le t} S_s^1 \ge H\}}\Big(\Phi\Big(\frac{\log\big(\frac{S_t^1}{K}\big) + \widetilde{r}\tau}{\sigma\sqrt{\tau}}\Big) - \Big(\frac{H}{S_t^1}\Big)^{\frac{2r}{\sigma^2}-1}\Phi\Big(\frac{\log\big(\frac{H^2}{KS_t^1}\big) + \widetilde{r}\tau}{\sigma\sqrt{\tau}}\Big)\Big),
\end{aligned}$$

where we have again set $\tau = T - t$ and $\widetilde{r} = r - \sigma^2/2$.

For the first term in (10.5) we define $\mathbb{R} \sim \mathbb{Q}$ via

$$\frac{d\mathbb{R}}{d\mathbb{Q}}\Big|_{\mathcal{F}_T} = \mathcal{E}(\sigma\widetilde{B})_T = \exp\Big(\sigma\widetilde{B}_T - \frac{1}{2}\sigma^2 T\Big).$$

Corollary 8.11 then proves that

$$\overline{B}_t = \widetilde{B}_t - \sigma t$$

is a Brownian motion with respect to $\mathbb{R}$, and we have

$$X_t = \big(r + \frac{\sigma^2}{2}\big)t + \sigma \overline{B}_t.$$

Theorem 8.7 can be used to deduce how conditional expectations behave under a measure change. One obtains

$$\mathbb{E}_{\mathbb{Q}}\big[\frac{d\mathbb{R}}{d\mathbb{Q}}\big|_T Y\big|\mathcal{F}_t\big] = \frac{d\mathbb{R}}{d\mathbb{Q}}\big|_t \mathbb{E}_{\mathbb{R}}[Y|\mathcal{F}_t]$$

for any $\mathbb{R}$-integrable random variable Y. With

$$\frac{d\mathbb{Q}}{d\mathbb{R}}\Big|_T S^1_T = \exp\Big(-\sigma\overline{B}_T - \frac{1}{2}\sigma^2 T\Big)\exp\Big(\big(r + \frac{\sigma^2}{2}\big)T + \sigma\overline{B}_T\Big) = \exp(rT)$$

one gets

$$\begin{aligned}
&\mathbb{E}_{\mathbb{Q}}\big[S^1_T 1_{\{S^1_T \geq K, \min_{s\leq T} S^1_s \geq H\}}|\mathcal{F}_t\big]\\
&= \exp(rT)\mathbb{E}_{\mathbb{Q}}\big[\frac{d\mathbb{R}}{d\mathbb{Q}}\big|_T 1_{\{S^1_T \geq K, \min_{s\leq T} S^1_s \geq H\}}|\mathcal{F}_t\big]\\
&= \exp(r(T-t))1_{\{\min_{s\leq t} S^1_s \geq H\}} S^1_t\\
&\qquad \mathbb{E}_{\mathbb{R}}\big[1_{\big\{-(X_T - X_t)\leq \log\big(\frac{S^1_t}{K}\big), \max_{t\leq s\leq T}(-(X_s - X_t))\leq \log\big(\frac{S^1_t}{H}\big)\big\}}|\mathcal{F}_t\big]\\
&= \exp(r(T-t))1_{\{\min_{s\leq t} S^1_s \geq H\}} S^1_t\\
&\qquad \Big(\Phi\Big(\frac{\log\big(\frac{S^1_t}{K}\big) + \overline{r}\tau}{\sigma\sqrt{\tau}}\Big) - \Big(\frac{S^1_t}{H}\Big)^{-1-\frac{2r}{\sigma^2}}\Phi\Big(\frac{\log\big(\frac{S^1_t}{K}\big) - 2\log\big(\frac{S^1_t}{H}\big) + \overline{r}\tau}{\sigma\sqrt{\tau}}\Big)\Big)\\
&= \exp(r(T-t))1_{\{\min_{s\leq t} S^1_s \geq H\}} S^1_t\\
&\qquad \Big(\Phi\Big(\frac{\log\big(\frac{S^1_t}{K}\big) + \overline{r}\tau}{\sigma\sqrt{\tau}}\Big) - \Big(\frac{H}{S^1_t}\Big)^{1+\frac{2r}{\sigma^2}}\Phi\Big(\frac{\log\big(\frac{H^2}{KS^1_t}\big) + \overline{r}\tau}{\sigma\sqrt{\tau}}\Big)\Big)
\end{aligned}$$

where this time $\overline{r} = r + \frac{\sigma^2}{2}$ holds. The pricing formula follows.

To obtain the replicating strategy we define the auxiliary stopping time

$$\tau = \inf\{t \geq 0 \mid S^1_t < H\} \wedge T.$$

Then

$$S^2 = Y^\tau \quad \text{with} \quad Y_t = f(t, S^1_t)$$

and the strategy follows analogously to the proof of Theorem 10.10(iv) with the help of Itō's formula. □

Example 10.20. Besides barrier options one also trades *lookback options* whose payoff depends on the minimum or the maximum of the underlying stock price until maturity. Here we discuss the *lookback call*

$$X = S^1_T - \min_{0\leq t\leq T} S^1_t$$

which intuitively resembles a call option whose strike price K equals the minimum of the stock price. A *lookback put* can be defined analogously.

Theorem 10.21. *Let*

$$X = S_T^1 - \min_{0 \le t \le T} S_t^1$$

be a lookback call. Then the fair price process S^2 of X takes the form

$$S_t^2 = f(t, S_t^1, \min_{0 \le s \le t} S_s^1)$$

with

$$\begin{aligned}
&f(t,x,m) \\
&= x\Phi\Big(\frac{\log\big(\frac{x}{m}\big) + \big(r + \frac{\sigma^2}{2}\big)(T-t)}{\sigma\sqrt{T-t}}\Big) - \exp(-r(T-t))m\Phi\Big(\frac{\log\big(\frac{x}{m}\big) + \big(r - \frac{\sigma^2}{2}\big)(T-t)}{\sigma\sqrt{T-t}}\Big) \\
&\quad - \frac{x\sigma^2}{2r}\Phi\Big(\frac{\log\big(\frac{m}{x}\big) - \big(r + \frac{\sigma^2}{2}\big)(T-t)}{\sigma\sqrt{T-t}}\Big) + \exp(-r(T-t))\frac{x\sigma^2}{2r}\big(\frac{m}{x}\big)^{\frac{2r}{\sigma^2}}\Phi\Big(\frac{\log\big(\frac{m}{x}\big) + \big(r - \frac{\sigma^2}{2}\big)(T-t)}{\sigma\sqrt{T-t}}\Big)\Big).
\end{aligned}$$

A hedging strategy is given by

$$\begin{aligned}
\varphi_t^0 &= \exp(-rt)(f(t, S_t^1, \min_{0 \le s \le t} S_s^1) - D_2 f(t, S_t^1, \min_{0 \le s \le t} S_s^1) S_t^1), \\
\varphi_t^1 &= D_2 f(t, S_t^1, \min_{0 \le s \le t} S_s^1).
\end{aligned}$$

Proof: We set again

$$X_t = \big(r - \frac{\sigma^2}{2}\big)t + \sigma \widetilde{B}_t$$

with $\widetilde{B}$ being a Brownian motion under $\mathbb{Q}$. From (10.3) we obtain

$$S_T^1 = S_t^1 \exp(X_T - X_t)$$

so that

$$\min_{0 \le s \le T} S_s^1 = \min_{0 \le s \le t} S_s^1 \wedge S_t^1 \exp(-\max_{t \le u \le T}(-(X_u - X_t))).$$

Using that X has independent and stationary increments, we have

$$\mathbb{E}_{\mathbb{Q}}[\min_{0 \le s \le T} S_s^1 | \mathcal{F}_t] = g(\min_{0 \le s \le t} S_s^1, S_t^1)$$

with

$$g(m,x) = \mathbb{E}_{\mathbb{Q}}[m \wedge x \exp(-M_{T-t})]$$

and $M_s = \max_{0 \le u \le s}(-(X_u))$. As usual let $\tau = T - t$, and we set $z = \log(x/m)$. We first compute

$$\begin{aligned}
g(m,x) - m &= \mathbb{E}_{\mathbb{Q}}[(x\exp(-M_\tau) - m)1_{\{M_\tau > z\}}] \\
&= x\mathbb{E}_{\mathbb{Q}}[(\exp(-M_\tau) - \exp(-z))1_{\{M_\tau > z\}}] \\
&= -x\mathbb{E}_{\mathbb{Q}}\Big[\int_z^\infty \exp(-y)1_{\{M_\tau > y\}}dy\Big] \\
&= -x\int_z^\infty \exp(-y)\mathbb{Q}(M_\tau > y)dy \qquad (10.6)
\end{aligned}$$

where we have used the fundamental theorem of calculus as well as Fubini's theorem. Theorem 10.17 now gives

$$\begin{aligned}\mathbb{Q}(M_\tau > m) &= 1 - \mathbb{Q}(M_\tau \le m) = 1 - \mathbb{Q}(-X_\tau \le m, M_\tau \le m)\\ &= 1 - \Phi\Big(\frac{m + (r - \frac{\sigma^2}{2})\tau}{\sigma\sqrt{\tau}}\Big) + \exp\big((1 - \frac{2r}{\sigma^2})m\big)\Phi\Big(\frac{-m + (r - \frac{\sigma^2}{2})\tau}{\sigma\sqrt{\tau}}\Big)\\ &= \mathbb{Q}(-X_\tau > m) + \exp\big((1 - \frac{2r}{\sigma^2})m\big)\mathbb{Q}(-X_\tau > m + \sigma^2\tau - 2r\tau)\end{aligned}$$

so that

$$\begin{aligned}g(m,x) - m = -x\int_z^\infty &\exp(-y)\mathbb{Q}(-X_\tau > y)dy \\ - x\int_z^\infty &\exp\big(-\frac{2r}{\sigma^2}y\big)\mathbb{Q}(-X_\tau > y + \sigma^2\tau - 2r\tau)dy.\end{aligned} \tag{10.7}$$

Both terms in (10.7) can again be computed using appropriate measure changes. For the first one we obtain as in (10.6)

$$\begin{aligned}&-x\int_z^\infty \exp(-y)\mathbb{Q}(-X_\tau > y)dy = x\mathbb{E}_{\mathbb{Q}}[(\exp(X_\tau) - \exp(-z))1_{\{-X_\tau > z\}}]\\ &= x\exp(r\tau)\mathbb{E}_{\mathbb{Q}}[\exp(\sigma\widetilde{B}_\tau - \frac{\sigma^2}{2}\tau)1_{\{-X_\tau>z\}}] - m\mathbb{Q}(-X_\tau > z)\\ &= x\exp(r\tau)\widetilde{\mathbb{Q}}(-X_\tau > z) - m\mathbb{Q}(-X_\tau > z)\end{aligned}$$

where $\widetilde{\mathbb{Q}}$ denotes the probability measure with density

$$\frac{d\widetilde{\mathbb{Q}}}{d\mathbb{Q}}\Big|_{\mathcal{F}_\tau} = \exp(\sigma\widetilde{B}_\tau - \frac{\sigma^2}{2}\tau) = \mathcal{E}(\sigma\widetilde{B})_\tau.$$

Corollary 8.11 gives

$$X_t = \big(r - \frac{\sigma^2}{2}\big)t + \sigma\widetilde{B}_t = \big(r + \frac{\sigma^2}{2}\big)t + \sigma\widehat{B}_t$$

where $\widehat{B}_t = \widetilde{B}_t - \sigma t$ is a $\widetilde{\mathbb{Q}}$-Brownian motion. Thus

$$\begin{aligned}&-x\int_z^\infty \exp(-y)\mathbb{Q}(-X_\tau > y)dy = x\exp(r\tau)\widetilde{\mathbb{Q}}(X_\tau < -z) - m\mathbb{Q}(X_\tau < -z)\\ &= x\exp(r\tau)\Phi\Big(\frac{\log\big(\frac{m}{x}\big) - \big(r + \frac{\sigma^2}{2}\big)\tau}{\sigma\sqrt{\tau}}\Big) - m\Phi\Big(\frac{\log\big(\frac{m}{x}\big) - \big(r - \frac{\sigma^2}{2}\big)\tau}{\sigma\sqrt{\tau}}\Big).\end{aligned}$$

The second term in (10.7) can be discussed analogously. We have

$$\begin{aligned}&-x\int_z^\infty \exp\big(-\frac{2r}{\sigma^2}y\big)\mathbb{Q}(-X_\tau > y + \sigma^2\tau - 2r\tau)dy\\ &= \frac{x\sigma^2}{2r}\mathbb{E}_{\mathbb{Q}}\big[\big(\exp\big(-\frac{2r}{\sigma^2}(-X_\tau - \sigma^2\tau + 2r\tau)\big) - \exp\big(-\frac{2r}{\sigma^2}z\big)\big)1_{\{-X_\tau - \sigma^2\tau + 2r\tau > z\}}\big]\\ &= \frac{x\sigma^2}{2r}\exp(r\tau)\mathbb{E}_{\mathbb{Q}}\big[\exp\big(\frac{2r}{\sigma}\widetilde{B}_\tau - \frac{2r^2}{\sigma^2}\tau\big)1_{\{-X_\tau > z + \sigma^2\tau - 2r\tau\}}\big]\\ &\quad - \frac{x\sigma^2}{2r}\exp\big(-\frac{2r}{\sigma^2}z\big)\mathbb{Q}(-X_\tau > z + \sigma^2\tau - 2r\tau)\\ &= \frac{x\sigma^2}{2r}\exp(r\tau)\check{\mathbb{Q}}(-X_\tau > z + \sigma^2\tau - 2r\tau)\\ &\quad - \frac{x\sigma^2}{2r}\exp\big(-\frac{2r}{\sigma^2}z\big)\mathbb{Q}(-X_\tau > z + \sigma^2\tau - 2r\tau)\end{aligned}$$

where this time $\check{\mathbb{Q}}$ denotes the probability measure with density

$$\frac{d\check{\mathbb{Q}}}{d\mathbb{Q}}\Big|_{\mathcal{F}_\tau} = \mathcal{E}(\frac{2r}{\sigma}\widetilde{B})_\tau.$$

Corollary 8.11 again gives

$$X_t = \big(r - \frac{\sigma^2}{2}\big)t + \sigma\widetilde{B}_t = \big(3r - \frac{\sigma^2}{2}\big)t + \sigma\check{B}_t$$

with $\check{B}_t = \widetilde{B}_t - \frac{2r}{\sigma}t$ being a $\check{\mathbb{Q}}$-Brownian motion. Thus

$$\begin{aligned}
&- x\int_z^\infty \exp\big(-\frac{2r}{\sigma^2}y\big)\mathbb{Q}(-X_\tau > y + \sigma^2\tau - 2r\tau)dy \\
&= \frac{x\sigma^2}{2r}\Big(\exp(r\tau)\Phi\Big(\frac{\log\big(\frac{m}{x}\big) - \big(r+\frac{\sigma^2}{2}\big)\tau}{\sigma\sqrt{\tau}}\Big) - \big(\frac{m}{x}\big)^{\frac{2r}{\sigma^2}}\Phi\Big(\frac{\log\big(\frac{m}{x}\big) + \big(r-\frac{\sigma^2}{2}\big)\tau}{\sigma\sqrt{\tau}}\Big)\Big).
\end{aligned}$$

The pricing formula can finally be derived quickly from Theorem 10.2 (ii) since

$$\begin{aligned}
\hat{S}_t^2 &= \mathbb{E}_{\mathbb{Q}}[\hat{X}|\mathcal{F}_t] = \mathbb{E}_{\mathbb{Q}}\big[\hat{S}_T^1|\mathcal{F}_t] - \exp(-rT)\mathbb{E}_{\mathbb{Q}}\big[\min_{0\le s\le T} S_s^1|\mathcal{F}_t\big] \\
&= \hat{S}_t^1 - \exp(-rT)\mathbb{E}_{\mathbb{Q}}\big[\min_{0\le s\le T} S_s^1|\mathcal{F}_t\big]
\end{aligned}$$

so

$$S_t^2 = S_t^1 - \exp(-r(T-t))(g(\min_{0\le s\le t} S_s^1, S_t^1) - \min_{0\le s\le t} S_s^1) - \exp(-r(T-t))\min_{0\le s\le t} S_s^1.$$

We finally set

$$U_t = \min_{0\le s\le t} S_s^1 \quad \text{and} \quad \hat{f}(t,x,u) = \exp(-rt)f(t,\exp(rt)x,u).$$

Then Itō's formula gives

$$\begin{aligned}
\hat{S}_t^2 = \hat{S}_0^2 &+ \int_0^t D_1\hat{f}(s,\hat{S}_s^1,U_s)ds + \int_0^t D_2\hat{f}(s,\hat{S}_s^1,U_s)d\hat{S}_s^1 \\
&+ \int_0^t D_3\hat{f}(s,\hat{S}_s^1,U_s)dU_s + \frac{1}{2}\int_0^t D_{22}\hat{f}(s,\hat{S}_s^1,U_s)d[\hat{S}^1]_s
\end{aligned}$$

because is U decreasing and continuous, hence of finite variation. The claim now follows analogously to the proof of Theorem 10.10(iv) by first obtaining

$$\hat{S}_t^2 = \hat{S}_0^2 + \int_0^t D_2\hat{f}(s,\hat{S}_s^1,U_s)d\hat{S}_s^1$$

and then choosing the corresponding self-financing strategy which exists due to Theorem 9.5. □

Chapter 11

Stochastic differential equations

This chapter deals with stochastic differential (or, rather, integral) equations which are generalizations of ordinary differential equations. We will give a few examples and state a general result on the existence and uniqueness of solutions to such equations.

Remark 11.1. An *ordinary differential equation* is of the form

$$x'(t) = f(t, x(t)) \quad \text{or} \quad \frac{dx(t)}{dt} = f(t, x(t))$$

with $x(0) = y$ and for some function $f : [0, T] \times \mathbb{R} \to \mathbb{R}$, and the goal is to find a unique global *solution*, i.e. a unique function $x : [0, T] \to \mathbb{R}$ which satisfies $x(0) = y$ and $x'(t) = f(t, x(t))$ for all $t \in [0, T]$.

Example 11.2.

(i) Let $T > 0$ be arbitrary. The ordinary differential equation

$$\frac{dx(t)}{dt} = x^2(t), \quad x(0) = 1,$$

has the unique solution

$$x(t) = \frac{1}{1-t}, \quad 0 \leq t < 1,$$

but it *explodes* in finite time. Thus a condition is needed which limits the growth of f in order to obtain a global solution on $[0, T]$.

(ii) The ordinary differential equation

$$\frac{dx(t)}{dt} = 3x^{2/3}(t), \quad x(0) = 0,$$

has more than one solution as every

$$x(t) = (t-a)^3 1_{\{t \geq a\}}$$

with $a > 0$ satisfies the differential equation. Hence a condition is needed which guarantees uniqueness of the solution.

Remark 11.3. A classical result is the Picard-Lindelöf theorem which guarantees existence and uniqueness of a global solution to an ordinary differential equation if $C, D > 0$ exist such that

$$|f(t,x)| \leq C(1+|x|), \quad t \in [0,T], \ x \in \mathbb{R},$$

and

$$|f(t,x) - f(t,y)| \leq D|x-y|, \quad t \in [0,T], \ x, y \in \mathbb{R},$$

hold.

Definition 11.4. Let $T > 0$ be arbitrary, let $b : [0,T] \times \mathbb{R} \to \mathbb{R}$ and $\sigma : [0,T] \times \mathbb{R} \to \mathbb{R}$ be measurable functions, let B be a Brownian motion and let Z be a random variable. An equation of the form

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dB_t, \quad X_0 = Z, \tag{11.1}$$

is called a *stochastic differential equation*. Any process $X = (X_t)_{0 \leq t \leq T}$ with

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

is called a *solution to the stochastic differential equation* (11.1).

Remark 11.5. One has to be careful about the exact formulation of a solution to a stochastic differential equation:

(i) Suppose that some probability space is given on which B and Z are defined, let $\mathcal{F} = (\mathcal{F}_t)_{0 \leq t \leq T}$ denote the filtration generated by B and suppose that Z is independent of $\mathcal{F}$. Call $\mathcal{G} = (\mathcal{G}_t)_{0 \leq t \leq T}$ the filtration where $\mathcal{G}_t$ is generated by Z and B_s with $s \leq t$. If there exists a solution X to the stochastic differential equation (11.1) which is adapted to $\mathcal{G}$ then X is called a *strong solution*.

(ii) If there exists a probability space $(\Omega, \mathcal{H}, \mathbb{P})$ with a filtration $(\mathcal{H}_t)_{0 \leq t \leq T}$, an $\mathcal{H}_0$-measurable random variable Z and adapted processes (X_t, B_t) such that B is a Brownian motion, independent of Z, and a martingale with respect to $(\mathcal{H}_t)_{0 \leq t \leq T}$, and such that the pair (X, B) satisfies (11.1) then (X, B) is called a *weak solution*.

Obviously, strong solutions are always weak solutions, but the converse is in general not true. (See e.g. Example 5.3.2 in Øksendal (2003).)

Lemma 11.6. (Gronwall's lemma) *Let $v : [0, \infty) \to [0, \infty)$ be a continuous function with*

$$v(t) \leq q + r \int_0^t v(s)ds$$

for some $q, r \geq 0$. Then $v(T) \leq q(1 + \exp(rT))$ for all $T > 0$.

Proof: Set

$$u(t) = \exp(-rt) \int_0^t v(s)ds.$$

Then

$$u'(t) = \exp(-rt)\left(v(t) - r\int_0^t v(s)ds\right) \le q\exp(-rt)$$

and

$$u(T) = u(T) - u(0) \le \int_0^T q\exp(-rt)dt \le \int_0^\infty q\exp(-rt)dt = \frac{q}{r}.$$

Clearly

$$v(T) \le q + r\int_0^T v(s)ds = q + r\exp(rT)u(T) \le q(1+\exp(rT))$$

ends the proof. □

Theorem 11.7. *Let $T > 0$ be arbitrary, let B be a Brownian motion, and let $b : [0,T]\times\mathbb{R} \to \mathbb{R}$ and $\sigma : [0,T]\times\mathbb{R} \to \mathbb{R}$ be measurable functions such that*

$$|b(t,x)| + |\sigma(t,x)| \le C(1+|x|), \quad t \in [0,T],\ x \in \mathbb{R}, \tag{11.2}$$

and

$$|b(t,x) - b(t,y)| + |\sigma(t,x) - \sigma(t,y)| \le D|x-y|, \quad t \in [0,T],\ x,y \in \mathbb{R}, \tag{11.3}$$

for some $C, D > 0$. Further, let $\mathcal{F} = (\mathcal{F}_t)_{0\le t\le T}$ denote the filtration generated by B and suppose that Z is independent of $\mathcal{F}$ with $\mathbb{E}[Z^2] < \infty$. Then the stochastic differential equation

$$dX_t = b(t,X_t)dt + \sigma(t,X_t)dB_t, \quad 0 \le t \le T,\ X_0 = Z, \tag{11.4}$$

has a unique (in the sense of indistinguishability over $[0,T]$) continuous strong solution which satisfies

$$\mathbb{E}[|X_T|^2] \le A(1+\mathbb{E}[Z^2])\exp(AT) \tag{11.5}$$

for some A which depends on C, T and $\mathbb{E}[Z^2]$.

Proof: To prove existence we define a new sequence of processes as follows: Set $Y_t^{(0)} = Z$ and

$$Y_t^{(k+1)} = Z + \int_0^t b(s,Y_s^{(k)})ds + \int_0^t \sigma(s,Y_s^{(k)})dB_s$$

for $k \ge 0$. We check first that the term above is well-defined for all $k \ge 0$, i.e. that

$$\mathbb{E}\Big[\int_0^T \big(|b(s,Y_s^{(k)})| + \sigma^2(s,Y_s^{(k)})\big)ds\Big] < \infty,$$

where we have used Theorem 7.16, and which implies that $\int_0^t \sigma(s,Y_s^{(k)})dB_s$ is an L^2 martingale over $[0,T]$. Using (11.2) the claim clearly holds for $k=0$, so we can work by induction. We obtain from the Cauchy-Schwarz inequality, Theorem 2.32(iii) and (11.2)

$$\begin{aligned}\mathbb{E}[\sup_{0\le t\le T}|Y_t^{(k+1)}|^2] &\le 3\mathbb{E}[Z^2] + 3C^2T\int_0^T \mathbb{E}[(1+|Y_s^{(k)}|)^2]ds + 12C^2\int_0^T \mathbb{E}[(1+|Y_s^{(k)}|)^2]ds \\ &\le A\left(1+\mathbb{E}[Z^2] + \int_0^T \mathbb{E}[|Y_s^{(k)}|^2]ds\right)\end{aligned}$$

for some $A > 0$ which only depends on C and T. Using (11.2) the claim then follows for $k+1$, and iteratively we also obtain

$$\mathbb{E}[|Y_T^{(k+1)}|^2] \leq A(1+\mathbb{E}[Z^2])\exp(AT). \tag{11.6}$$

Now, for any $k \geq 1$ we have $Y_t^{(k+1)} - Y_t^{(k)} = R_t^{(k)} + M_t^{(k)}$ with

$$R_t^{(k)} = \int_0^t \left(b(s, Y_s^{(k)}) - b(s, Y_s^{(k-1)})\right) ds \quad \text{and} \quad M_t = \int_0^t \left(\sigma(s, Y_s^{(k)}) - \sigma(s, Y_s^{(k-1)})\right) dB_s.$$

Clearly, the Cauchy-Schwarz inequality together with (11.3) gives

$$\begin{aligned}\mathbb{E}\left[\sup_{0\leq t\leq T}(R_t^{(k)})^2\right] &\leq \mathbb{E}\left[\sup_{0\leq t\leq T} t\int_0^t \left(b(s, Y_s^{(k)}) - b(s, Y_s^{(k-1)})\right)^2 ds\right] \\ &\leq TD^2 \int_0^T \mathbb{E}[|Y_s^{(k)} - Y_s^{(k-1)}|^2] ds,\end{aligned}$$

and as M is a martingale, Theorem 2.32(iii) and the isometry property give

$$\begin{aligned}\mathbb{E}\left[\sup_{0\leq t\leq T}(M_t^{(k)})^2\right] &\leq 4\mathbb{E}\left[\int_0^T \left(\sigma(s, Y_s^{(k)}) - \sigma(s, Y_s^{(k-1)})\right)^2 ds\right] \\ &\leq 4D^2 \int_0^T \mathbb{E}[|Y_s^{(k)} - Y_s^{(k-1)}|^2] ds.\end{aligned}$$

Thus

$$\mathbb{E}\left[\sup_{0\leq t\leq T}|Y_t^{(k+1)} - Y_t^{(k)}|^2\right] \leq L\int_0^T \mathbb{E}[|Y_s^{(k)} - Y_s^{(k-1)}|^2] ds$$

for any $k \geq 1$ and any $0 \leq t \leq T$, and where L depends only on D and T. Since (11.6) yields

$$A^* = \sup_{0\leq s\leq T} \mathbb{E}[|Y_s^{(1)} - Y_s^{(0)}|^2] < \infty$$

for some A^* which only depends on C, T and $\mathbb{E}[Z^2]$ we obtain

$$\mathbb{E}\left[\sup_{0\leq t\leq T}|Y_t^{(k+1)} - Y_t^{(k)}|^2\right] \leq A^* \frac{L^k T^k}{k!},$$

and

$$\mathbb{P}\left(\sup_{0\leq t\leq T}|Y_t^{(k+1)} - Y_t^{(k)}| > 2^{-(k+1)}\right) \leq 4A^* \frac{(4LT)^k}{k!}.$$

The Borel-Cantelli lemma then proves the existence of some Ω^* with $\mathbb{P}(\Omega^*) = 1$ such that for all $\omega \in \Omega^*$ there exists $N(\omega) \geq 1$ such that $\sup_{0\leq t\leq T}|Y_t^{(k+1)}(\omega) - Y_t^{(k)}(\omega)| \leq 2^{-(k+1)}$ for all $k \geq N(\omega)$. As a consequence of the triangle inequality $\sup_{0\leq t\leq T}|Y_t^{(k+m)}(\omega) - Y_t^{(k)}(\omega)| \leq 2^{-k}$ for all $m \geq k \geq N(\omega)$. Thus $Y^{(k)}$ converges almost surely uniformly to a continuous limit X, and the limit is measurable with respect to the filtration $\mathcal{G}$ because every $Y_t^{(k)}$ is. Also, (11.5) follows from (11.6) and Fatou's lemma.

The last step is to prove that X actually satisfies the stochastic differential equation for which by continuity for each $0 \leq t \leq T$

$$\mathbb{E}\left[\left(\int_0^t \left(b(s, X_s) - b(s, Y_s^{(k)})\right) ds\right)^2\right] \to 0$$

and

$$\mathbb{E}\left[\left(\int_0^t \left(\sigma(s, X_s) - \sigma(s, Y_s^{(k)})\right) dB_s\right)^2\right] \to 0$$

need to be shown. This claim follows from similar techniques as before.

Finally, suppose that X and Y are both solutions to (11.4) which satisfy (11.5). Again the local martingales are actually martingales, so with $a_s = b(s, X_s) - b(s, Y_s)$ and $\gamma_s = \sigma(s, X_s) - \sigma(s, Y_s)$ we have

$$\begin{aligned}\mathbb{E}[|X_t - Y_t|^2] = \mathbb{E}\left[\left(\int_0^t a_s ds + \int_0^t \gamma_s dB_s\right)^2\right] &\leq 2\mathbb{E}\left[\left(\int_0^t a_s ds\right)^2\right] + 2\mathbb{E}\left[\left(\int_0^t \gamma_s dB_s\right)^2\right] \\ &\leq 2t\mathbb{E}\left[\int_0^t a_s^2 ds\right] + 2\mathbb{E}\left[\int_0^t \gamma_s^2 ds\right] \leq 2(1+T)D^2 \int_0^t \mathbb{E}[|X_s - Y_s|^2] ds\end{aligned}$$

for any $0 \leq t \leq T$. Here we have used the Cauchy-Schwarz inequality, the isometry property and (11.3). From Lemma 11.6 with $q = 0$ and $r = 2(1+T)D^2$ we obtain $\mathbb{E}[|X_t - Y_t|^2] = 0$, i.e. $X_t = Y_t$ almost surely for all $0 \leq t \leq T$. Continuity of both processes gives indistinguishability. □

Example 11.8. The stochastic differential equation

$$dX_t = -\lambda X_t dt + \sigma dB_t, \quad X_0 = x,$$

has a unique strong solution over the entire $[0, \infty)$, because (11.2) and (11.3) are satisfied for all $T > 0$. Its solution is given by

$$X_t = x \exp(-\lambda t) + \sigma \int_0^t \exp(\lambda(s - t)) dB_s$$

and is called the *Ornstein-Uhlenbeck process.*

Bibliography

Brockwell, P. and R. Davis (1991). *Time series: theory and methods.* Springer, New York.

Delbaen, F. and W. Schachermayer (1994). A general version of the fundamental theorem of asset pricing. *Math. Ann. 300*(3), 463–520.

Karatzas, I. and S. Shreve (1991). *Brownian motion and stochastic calculus.* Springer, New York.

Klenke, A. (2006). *Probability theory.* Springer, Berlin.

Mörters, P. and Y. Peres (2010). *Brownian motion.* Cambridge University Press, Cambridge.

Øksendal, B. (2003). *Stochastic differential equations.* Springer, Berlin.

Protter, P. (2004). *Stochastic integration and differential equations.* Springer, Berlin.