Julius-Maximilians-
UNIVERSITÄT
WÜRZBURG

Lecture notes, summer semester 2021

Stochastic processes

Prof. Dr. Markus Bibinger
Institute of Mathematics
Chair of Mathematics VIII (Applied Stochastics)

Michael Sonntag
Tutorials

Homepage
www.mathematik.uni-wuerzburg.de/appliedstochastics

Julius-Maximilians-Universität Würzburg
Lehrstuhl für Angewandte Stochastik (Mathematik VIII)
Fakultät für Mathematik und Informatik, Institut für Mathematik

Emil-Fischer-Straße 30, D-97074 Würzburg
Tel. +49 931 31-87610
markus.bibinger@mathematik.uni-wuerzburg.de
michael.sonntag@mathematik.uni-wuerzburg.de
www.mathematik.uni-wuerzburg.de/appliedstochastics

This version: July 5, 2021

Contents

Vorbemerkung:

Die vorliegenden Lecture Notes sind begleitend zur 4+2 SWS Veranstaltung "Stochastic processes" ("Stochastische Prozesse") im Sommersemester 2021 an der Universität Würzburg. Es werden Kenntnisse der Vorlesung Stochastik 1 vorausgesetzt. Mit der bedingten Erwartung wird eine wichtige Grundlage aus der Wahrscheinlichkeitstheorie zu Beginn explizit behandelt. Hyperlinks und Links sind blau markiert, besonders wichtige Begriffe im Text werden rotbraun hervorgehoben und in Sätzen und Definitionen durch **breite Schrift**. Es handelt sich um in erster Version neu zusammengestellte Lecture Notes. Für Hinweise zu möglichen Tippfehlern und Inkonsistenzen wäre ich dankbar.

1 First important processes

1.1 The Poisson process

Stochastic processes

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and $(E, \mathcal{E})$ some measurable space. We call a family of random variables $(X_t)_{t\in\mathcal{T}}$, with $X_t : \Omega \to E$, a stochastic process with state space $(E, \mathcal{E})$. The index set $\mathcal{T}$ is arbitrary, but usually a subset of $\mathbb{R}$.

If $\mathcal{T} = \mathbb{N}, \mathbb{N}_0$ or $\mathbb{Z}$, we say that the process is in discrete time, for $\mathcal{T} = [0,T]$ or $\mathcal{T} = [0,\infty)$, we say that the process is in continuous time. For $(E, \mathcal{E}) = (\mathbb{R}, \mathcal{B})$, with the Borel σ-field $\mathcal{B}$, we say the process is real-valued.

Examples 1.1. A sequence of random variables is a stochastic process in discrete time, e.g.
a. $(X_n)_{n\in\mathbb{N}}$ independent and identically distributed (i.i.d.) with $X_n \sim \mathcal{N}(0,1)$ for all n.
b. For (X_n) as in a., set $S_n = X_1 + \ldots + X_n$. $(S_n)_{n\geq 1}$ is a discrete-time stochastic process. •

Formally, a stochastic process is considered as a mapping

$$(X_t) : \Omega \times \mathcal{T} \to E, \qquad (\omega, t) \mapsto X_t(\omega)\,.$$

For fix $t \in \mathcal{T}$, $X_t(\omega)$ is a random variable. The mappings

$$\mathcal{T} \to E, \quad t \mapsto X_t(\omega)$$

are the paths of the stochastic process. The realization of the process for some ω thus yields a path, i.e. a function of $t \in \mathcal{T}$. For instance, if (X_t) is real-valued and $\mathcal{T} = \mathbb{N}$, the path $(X_n(\omega))_{n\in\mathbb{N}}$ yields a sequence of real numbers. Paths a more interesting for processes in continuous time. Then, the path is a function for each ω, whose properties can be investigated, for instance, continuity and differentiability.

A process in continuous time is called (right-)continuous if the paths are (almost surely) (right-)continuous. Almost surely (right-)continuous paths means that there exists $A \in \mathcal{A}$, with $\mathbb{P}(A) = 1$, such that for all $\omega \in A$, the path $t \mapsto X_t(\omega)$, $t \in \mathcal{T}$ is (right-)continuous. We introduce next the first example of a continuous-time stochastic process.

Definition 1.2. *A $\mathbb{N}_0$-valued stochastic process $(N_t)_{t\geq 0}$ is a **Poisson process** with **intensity** $\lambda > 0$, if*

1. *(N_t) has almost surely right-continuous and monotone increasing paths.*
2. *(N_t) has **independent increments**:*
   For all $0 \leq t_1 < t_2 < \ldots < t_n$, $n \in \mathbb{N}$, the random variables $\left(N_{t_i} - N_{t_{i-1}}\right)_{1\leq i\leq n}$ are independent.
3. *$N_t \sim$ Poi(λt), $t \geq 0$, with $N_0 = 0$, almost surely.*

Recall that $N_t \sim \text{Poi}(\lambda t)$ refers to the Poisson distribution, i.e.

$$\mathbb{P}(N_t = k) = e^{-\lambda t}\,\frac{(\lambda t)^k}{k!}\,,\ k \in \mathbb{N}_0\,. \tag{1.1}$$

A $\mathbb{N}_0$-valued process with increasing paths is called a counting process. The Poisson process is thus a counting process. The definition readily implies that (N_t) has stationary increments: $(N_t - N_s) \sim$

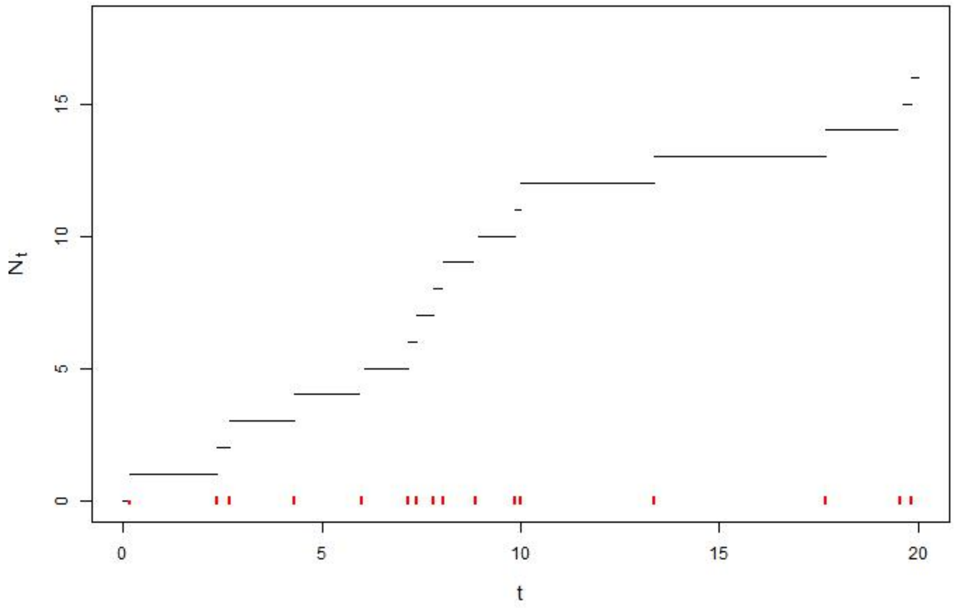


Figure 1.1: A path of a Poisson process with its jump times.

$\mathrm{Poi}(\lambda(t-s))$, $s<t$. This follows, since $N_t = N_s + (N_t - N_s)$, and denoting $\varphi_{N_t-N_s}$ the characteristic function of $N_t - N_s$, the independence of the increments implies that

$$\varphi_{N_t-N_s}(u) = \frac{\varphi_{N_t}(u)}{\varphi_{N_s}(u)} = \frac{e^{\lambda t(e^{iu}-1)}}{e^{\lambda s(e^{iu}-1)}} = e^{\lambda(t-s)(e^{iu}-1)},$$

using the characteristic function of the Poisson distribution. Since characteristic functions uniquely determine a distribution, the claim follows.

In Figure 1.1 one simulated path of a Poisson process is illustrated. A Poisson process is a process with increasing right-continuous, but obviously not continuous, paths and independent, stationary, Poisson distributed increments. The Poisson process is a central object in insurance mathematics. It can be used to model arriving claims caused by damages at some indemnity insurance. For instance, in the classical Cramér-Lundberg model [1] the "claim number process" is a Poisson process.

The Poisson distribution is not as arbitrarily chosen as it might seem. Some qualitative statements which are reasonable in many situations lead to the Poisson distribution. This universality comes from Poisson convergence. Korollar 3.31 from Bibinger and Holzmann [2019] yields the following result.

Proposition 1.3. *Let $(N_t)_{t\geq 0}$ be a stochastic process with values in $\mathbb{N}_0$, $N_0 = 0$, with independent, stationary increments and with increasing right-continuous paths. (N_t) is a Poisson process with intensity $\lambda > 0$,* ***if and only if (iff)***

$$\mathbb{P}(N_h = 1) = \lambda h + o(h) \quad \textit{and} \quad \mathbb{P}(N_h \geq 2) = o(h), \quad h \to 0. \tag{1.2}$$

That the Poisson process defined in Definition 1.2 satisfies these identities is seen by a first-order series expansion of the exponential function in the counting density (1.1). The crucial backwards direction is based on Poisson convergence on which we shall not elaborate here.

[1] wiki: ruin theory

We can write a process with a telescoping sum as a sum of the increments: $N_n = N_1 + (N_2 - N_1) + \ldots + (N_n - N_{n-1})$, $n \in \mathbb{N}$, and we conclude for a Poisson process that $N_t \uparrow \infty$, almost surely, by the strong law of large numbers.

Lemma 1.4. *$(N_t)_{t\geq 0}$ has almost surely only jumps of size 1.*

Proof. This is based on an expansion of the exponential function. It holds for any $T \in \mathbb{R}_+ = [0,\infty)$, that

$$\begin{aligned}\mathbb{P}\Big(\max_{1\leq k\leq 2^n T} N_{\frac{k}{2^n}} - N_{\frac{k-1}{2^n}} \geq 2\Big) &= 1 - \Big(\mathbb{P}\Big(\max_{1\leq k\leq 2^n T} N_{\frac{k}{2^n}} - N_{\frac{k-1}{2^n}} \leq 1\Big)\Big)^{2^n T} \\ &= 1 - \Big(\exp(-\lambda/2^n) + \frac{\lambda}{2^n}\exp(-\lambda/2^n)\Big)^{2^n T} \\ &= 1 - e^{-\lambda T}\Big(1+\frac{\lambda}{2^n}\Big)^{2^n T},\end{aligned}$$

which tends to zero as $n \to \infty$ for any T. ■

Set $T_0 = 0$, and $T_n = \min\{t > 0 : \quad N_t = n\}$. Then, it holds that

$$\{N_t = n\} = \{T_n \leq t < T_{n+1}\}, \text{ and } N_t = \sum_{n=1}^{\infty} \mathbf{1}_{\{T_n \leq t\}}.$$

T_n is the waiting time until the nth jump of the process. Denote by $W_n = T_n - T_{n-1}$ the waiting times between consecutive jumps. The next result shows that the Poisson process is characterized by the waiting times between its jumps.

Theorem 1.5. *$(N_t)_{t\geq 0}$ is a Poisson process with intensity λ, iff the waiting times $(W_n)_{n\geq 1}$ are independent, Exp(λ) distributed random variables.*

Recall that the exponential distribution, Exp(λ), is determined by its density $f(x) = \lambda e^{-\lambda x}\mathbf{1}_{[0,\infty)}(x)$.

Proof. a. Since $T_n \to \infty$ almost surely, the paths are almost surely $\mathbb{N}_0$-valued.
b. N_t is a random variable for all $t \geq 0$. Since

$$\{N_t = n\} = \{T_n \leq t < T_{n+1}\},$$

the measurability of N_t follows.
c. (N_t) has right-continuous paths: If $N_t(\omega) = n$, such that $T_n(\omega) \leq t < T_{n+1}(\omega)$, we have for small δ that $T_n(\omega) \leq t + \delta < T_{n+1}(\omega)$, also $N_{t+\delta}(\omega) = n$.
d. The main point is to prove the distributional property. It suffices to show for $0 \leq s < t$, $n,m \in \mathbb{N}_0$, that

$$\mathbb{P}(N_s = n, N_t - N_s = m) = \Big(e^{-\lambda s}\frac{(\lambda s)^n}{n!}\Big)\Big(e^{-\lambda(t-s)}\frac{(\lambda(t-s))^m}{m!}\Big). \tag{1.3}$$

Summation over m yields that $N_s \sim Poi(\lambda s)$, and the independence of N_s and $N_t - N_s$. In order to prove (1.3) we require the identity

$$\int_0^\infty \cdots \int_0^\infty \mathbf{1}_{\{x_1+\ldots+x_n\leq u\}}\, dx_1 \ldots dx_n = \frac{u^n}{n!}, \quad u \geq 0, \tag{1.4}$$

what can be proved by induction. We prove (1.3) for $m \geq 1$. Since $N_t - N_s$ takes values in $\mathbb{N}_0$, the result for $m = 0$ is then implied by the normalization of the probability measure $\mathbb{P}$. Let $x = (x_1,\ldots,x_n)^T \in \mathbb{R}^n$,

$l \le n$, and set $s_l(x) = x_1 + \ldots + x_l$. The joint distribution of $(W_1, \ldots, W_{n+m+1})$ by independence has a product density

$$x = (x_1, \ldots, x_{n+m+1})^T \mapsto \prod_{j=1}^{n+m+1} \left(\lambda e^{-\lambda x_j}\right) = \lambda^{n+m+1} e^{-\lambda s_{n+m+1}(x)}.$$

With the transformation formula, we compute

$$\begin{aligned}
& \mathbb{P}(N_s = n, N_t - N_s = m) \\
= \; & \mathbb{P}(N_s = n, N_t = m+n) = \mathbb{P}(T_n \le s < T_{n+1},\; T_{n+m} \le t < T_{n+m+1}) \\
= \; & \mathbb{P}(W_1 + \ldots + W_n \le s < W_1 + \ldots + W_{n+1},\; W_1 + \ldots + W_{n+m} \le t < W_1 + \ldots + W_{n+m+1}) \\
= \; & \int_0^\infty \cdots \int_0^\infty \lambda^{n+m+1} e^{-\lambda s_{n+m+1}(x)} \mathbf{1}_{\{s_n(x) \le s < s_{n+1}(x)\}} \mathbf{1}_{\{s_{n+m}(x) \le t < s_{n+m+1}(x)\}} \, dx_1 \ldots dx_{n+m+1}.
\end{aligned}$$

Substituting $y_{n+m+1} = s_{n+m+1}(x)$, $y_{n+1} = s_{n+1}(x) - s$, $y_i = x_i$, else, the determinant of the substitution is 1, and the integral yields with (1.4) that

$$\begin{aligned}
& \lambda^{n+m} \int_0^\infty \cdots \int_0^\infty \left(\int_t^\infty \lambda e^{-\lambda y_{n+m+1}} \, dy_{n+m+1} \right) \Bigg(\int_0^\infty \cdots \int_0^\infty \mathbf{1}_{\{0 < y_{n+1}\}} \\
& \mathbf{1}_{\{y_{n+1} + \ldots + y_{n+m} \le t-s\}} \, dy_{n+1} \ldots dy_{n+m} \Bigg) \mathbf{1}_{\{y_1 + \ldots + y_n \le s\}} \, dy_1 \ldots dy_n \\
= \; & \lambda^{n+m} \int_0^\infty \cdots \int_0^\infty e^{-\lambda t} \frac{(t-s)^m}{m!} \mathbf{1}_{\{y_1 + \ldots + y_n \le s\}} \, dy_1 \ldots dy_n \\
= \; & \lambda^{n+m} e^{-\lambda t} \frac{(t-s)^m}{m!} \frac{s^n}{n!}.
\end{aligned}$$

We conclude (1.3). ■

For the simulation of the Poisson process I used another characterization from the following proposition based on the conditional distribution of the jump times. We will, however, only establish this conditional distribution and not that it uniquely characterizes the Poisson process.

Proposition 1.6. *Let $(N_t)_{t \ge 0}$ be a Poisson process. Given $N_t = n$, the jump times $T_1, \ldots, T_n$, are jointly distributed as the order statistic of n i.i.d. on $(0,t)$ uniformly distributed random variables.*

Proof. For all $0 = s_0 < t_1 < \ldots < t_n \le t$, it holds that

$$\begin{aligned}
& \mathbb{P}(T_1 \le t_1, \ldots, T_n \le t_n | N_t = n) \\
& = \int_0^{t_1} \ldots \int_{t_{n-1}}^{t_n} f_{W_1, \ldots, W_n}(s_1, s_2 - s_1, \ldots, s_n - s_{n-1}) \frac{\mathbb{P}(W_{n+1} > t - s_n)}{\mathbb{P}(N_t = n)} \, ds_1 \ldots ds_n \\
& = \int_0^{t_1} \ldots \int_{t_{n-1}}^{t_n} \left(\prod_{i=1}^n \lambda e^{-\lambda(s_i - s_{i-1})} \right) \frac{e^{-\lambda(t - s_n)}}{(\lambda t)^n e^{-\lambda t}/n!} \, ds_1 \ldots ds_n \\
& = \int_0^{t_1} \ldots \int_{t_{n-1}}^{t_n} \frac{n!}{t^n} \, ds_1 \ldots ds_n,
\end{aligned}$$

coinciding with the integral over the density of the order statistic of independent $U((0,t))$-distributed random variables. ■

1.2 Markov chains

Definition 1.7. *Let $\mathcal{T} = \mathbb{N}_0$ (discrete time), or $\mathcal{T} = \mathbb{R}_+$ (continuous time) and $(X_t)_{t \in \mathcal{T}}$ a stochastic process with a countable state space E. $(X_t)_{t \in \mathcal{T}}$ is called a* ***Markov chain*** *if the* ***Markov property*** *is*

satisfied:
For all $n \in \mathbb{N}$, $t_1 < t_2 < \ldots < t_{n+1}$ and $s_i \in E, 1 \le i \le n+1$, with $\mathbb{P}(X_{t_1} = s_1, \ldots, X_{t_n} = s_n) > 0$, it holds that

$$\mathbb{P}(X_{t_{n+1}} = s_{n+1} | X_{t_1} = s_1, \ldots, X_{t_n} = s_n) = \mathbb{P}(X_{t_{n+1}} = s_{n+1} | X_{t_n} = s_n) . \tag{1.5}$$

Definition 1.8. *For a Markov chain (X_t) and $t_1 \le t_2$ and $k, l \in E$, the **transition probability** to reach state k at time t_2 from state l at time t_1 is defined by*

$$p_{lk}(t_1, t_2) := \mathbb{P}(X_{t_2} = k | X_{t_1} = l) . \tag{1.6}$$

*The matrix $(p_{lk}(t_1,t_2))_{l,k\in E} =: P(t_1,t_2)$ defines the **transition matrix**. The Markov chain (X_t) and its transition matrix $P(t_1,t_2)$ are called **(time-)homogeneous**, if for all $t_1 \le t_2$, it holds that*

$$P(t_1,t_2) = P(0, t_2 - t_1) =: P(t_2 - t_1) .$$

Proposition 1.9. *The transition matrix of a Markov chain (X_t) satisfies the **Chapman-Kolmogorov equation**:*

$$\forall t_1 \le t_2 \le t_3 : \; P(t_1,t_3) = P(t_1,t_2) \cdot P(t_2,t_3) , \tag{1.7}$$

*i.e. we have a matrix product right-hand side. In the homogeneous case, this yields the **semigroup property**:*

$$\forall t, s \in \mathcal{T} : \; P(t+s) = P(t) \cdot P(s) . \tag{1.8}$$

In particular, $P(n) = (P(1))^n$, for $n \in \mathbb{N}$.

Proof. It holds that

$$\begin{aligned}
(P(t_1,t_3))_{lk} &= \mathbb{P}(X_{t_3} = k | X_{t_1} = l) \\
&= \sum_{m\in E} \mathbb{P}(X_{t_3} = k, X_{t_2} = m | X_{t_1} = l) \\
&= \sum_{m\in E} \mathbb{P}(X_{t_3} = k | X_{t_2} = m, X_{t_1} = l) \, \mathbb{P}(X_{t_2} = m | X_{t_1} = l) \\
&= \sum_{m\in E} \mathbb{P}(X_{t_3} = k | X_{t_2} = m) \, \mathbb{P}(X_{t_2} = m | X_{t_1} = l) \\
&= \sum_{m\in E} (P(t_2,t_3))_{mk} \, (P(t_1,t_2))_{lm} \\
&= \big(P(t_1,t_2) \cdot P(t_2,t_3)\big)_{lk} .
\end{aligned}$$

For the fourth equality, we have used the Markov property. In the homogeneous case, this reduces to

$$P(t_3 - t_1) = P(t_2 - t_1) \cdot P(t_3 - t_2)$$

and setting $t = t_2 - t_1$, and $s = t_3 - t_2$ yields the result. ■

Example 1.10. Let $E = \{0,1\}$, $\mathcal{T} = \mathbb{N}_0$, and consider the homogeneous transition matrix

$$P(1) = \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix} ,$$

with $p, q \in (0,1)$.
Assume that Lufthansa in Germany has data that their Business Class customers switch for their next flight to a different airline with probability 15%. Customers who had a flight with another airline switch

to Lufthansa with a probability of 10%. We can model the situation with a two-state Markov chain as above with $p = 0.15$, and $q = 0.10$. Assume that on a connection, Lufthansa has currently 30% of all Business Class customers and that the customers fly four times a year. What can we forecast for their share of Business Class customers to be one year later?

We have with the initial state given by $\pi_0 = (0.30, 0.70)$, and the transition matrix $P(1)$, the state one year later given by

$$\begin{aligned}
\pi_4 := \pi_0 \left(P(1)\right)^4 &= (0.30, 0.70) \begin{pmatrix} 0.85 & 0.15 \\ 0.10 & 0.90 \end{pmatrix}^4 \\
&= (0.30, 0.70) \begin{pmatrix} 0.589844 & 0.410156 \\ 0.273438 & 0.726563 \end{pmatrix} \\
&= (0.368, 0.632) \,.
\end{aligned}$$

The fourth power is taken, since Business Class customers have 4 flights per year. Thus, after one year Lufthansa's share of the Business Class market increased to 36.8%. •

2 Discrete-time martingales

2.1 Conditional expectation

Existence in L_1 and basic properties

Definition 2.1. *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and $X : \Omega \to \mathbb{R}$ a real-valued random variable with $\mathbb{E}[|X|] < \infty$. Denote by $\mathcal{F} \subseteq \mathcal{A}$ a sub-σ-field of $\mathcal{A}$. A* ***conditional expectation*** *of X given $\mathcal{F}$ is an integrable random variable $Z : \Omega \to \mathbb{R}$ which satisfied*

1. Z is $(\mathcal{F} - \mathcal{B})$ measurable, where $\mathcal{B}$ denotes the Borel-σ-field.
2. For all $F \in \mathcal{F}$, it holds true that

$$\int_F Z \, d\mathbb{P} = \int_F X \, d\mathbb{P}\,. \tag{2.1}$$

The condition 1. means that Y is not only $(\mathcal{A} - \mathcal{B})$-measurable, but already $(\mathcal{F} - \mathcal{B})$-measurable which is a stronger condition, since $\mathcal{F} \subseteq \mathcal{A}$.

Lemma 2.2. *If Y and Y' are conditional expectations of X given $\mathcal{F}$, it holds that $Y = Y'$ $\mathbb{P}$-almost surely.*

Proof. We prove that $Y \geq Y'$ a.s., and exchanging the roles yields the result. For $n \in \mathbb{N}$ we have according to 1. that $\{Y' - Y \geq 1/n\} \in \mathcal{F}$ and thus 2. yields that

$$\mathbb{P}(Y' - Y \geq 1/n) \leq n \int_{\{Y'-Y \geq 1/n\}} (Y' - Y) \, d\mathbb{P} = 0\,.$$

We obtain that

$$\mathbb{P}\Big(\bigcup_{n \geq 1} \{Y' - Y \geq 1/n\} \Big) = 0,$$

and on the complement it holds that $Y \geq Y'$. ■

By this uniqueness result, we can speak of the conditional expectation of X given $\mathcal{F}$, $\mathbb{E}[X|\mathcal{F}]$, which is however only almost surely unique. By (2.1), it is characterized by the identity

$$\mathbb{E}\left[\mathbf{1}_F \, \mathbb{E}[X|\mathcal{F}]\right] = \mathbb{E}\left[\mathbf{1}_F X\right] \;\; \forall\, F \in \mathcal{F}\,. \tag{2.2}$$

Theorem 2.3. *It holds that $\mathbb{E}[\mathbb{E}[X|\mathcal{F}]] = \mathbb{E}[X]$, and $\mathbb{E}\left[|\mathbb{E}[X|\mathcal{F}]|\right] \leq \mathbb{E}\left[|X|\right]$.*

Proof. Define $Y = \mathbb{E}[X|\mathcal{F}]$. Condition 2. with $F = \Omega$ implies that $\mathbb{E}[X] = \mathbb{E}[Y]$. It is $A = \{Y > 0\} \in \mathcal{F}$, such that

$$\int_A Y \, d\mathbb{P} = \int_A X \, d\mathbb{P}, \qquad -\int_{A^\complement} Y \, d\mathbb{P} = -\int_{A^\complement} X \, d\mathbb{P}\,.$$

We conclude that

$$\begin{aligned} \mathbb{E}\left[|Y|\right] &= \int_A Y \, d\mathbb{P} - \int_{A^\complement} Y \, d\mathbb{P} \\ &= \int_A X \, d\mathbb{P} - \int_{A^\complement} X \, d\mathbb{P} \end{aligned}$$

$$
\begin{aligned}
&= \int_A X \, d\mathbb{P} + \int_{A^\complement} (-X) \, d\mathbb{P} \\
&\leq \int_A |X| \, d\mathbb{P} + \int_{A^\complement} |X| \, d\mathbb{P} = \mathbb{E}\left[|X|\right] < \infty .
\end{aligned}
$$
■

The first identity is known as the tower rule or the law of total expectation.
Next, we prove existence of conditional expectations based on the characterizing property (2.2). This builds upon the Radon-Nikodym theorem from measure and integration theory. [1] We repeat this theorem without proof here.
Let $(\Omega, \mathcal{F})$ be a measurable space and μ, ν measures on $(\Omega, \mathcal{F})$. If for $A \in \mathcal{F}$, $\mu(A) = 0$ always implies that $\nu(A) = 0$, we call ν absolutely continuous with respect to μ, notation $\nu \ll \mu$.

Theorem 2.4 (Radon-Nikodym)**.** *Let μ, ν be σ-finite measures on $(\Omega, \mathcal{A})$. Then, $\nu \ll \mu$ iff there exists a non-negative measurable function f on $(\Omega, \mathcal{A})$ taking values in $\bar{\mathbb{R}} = \mathbb{R} \cup \{-\infty\} \cup \{+\infty\}$, with*

$$
\nu(A) = \int_A f \, d\mu \quad \forall \, A \in \mathcal{A} .
$$

This function f is μ-almost everywhere uniquely determined. We call f a ***(Radon-Nikodym) density*** *or $f = d\nu/d\mu$ a* ***Radon-Nikodym derivative****.* ▪

Theorem 2.5 (Existence of conditional expectation)**.** *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and $X : \Omega \to \mathbb{R}$ a random variable with $\mathbb{E}[|X|] < \infty$, and $\mathcal{F} \subseteq \mathcal{A}$ a sub-σ-field. There exists a conditional expectation $\mathbb{E}[X|\mathcal{F}]$ of X given $\mathcal{F}$.* ▪

Proof. 1. $X \geq 0$: We define

$$
\nu(F) := \int_F X \, d\mathbb{P}, \qquad F \in \mathcal{F},
$$

which is a σ-finite measure on $(\Omega, \mathcal{F})$, and absolutely continuous w.r.t. $\mathbb{P}$ (restricted to $(\Omega, \mathcal{F})$). By Radon-Nikodym there exists a $(\mathcal{F} - \mathcal{B})$-measurable $Z : \Omega \to [0, \infty)$, with

$$
\nu(F) = \int_F Z \, d\mathbb{P}, \qquad A \in \mathcal{F} .
$$

This random variable Z satisfies the characterizing properties of conditional expectation.
2. For general $X = X^+ - X^-$, we show that $Z = \mathbb{E}[X^+|\mathcal{F}] - \mathbb{E}[X^-|\mathcal{F}]$ is a conditional expectation of X given $\mathcal{F}$. The measurability property is clear, since it holds for $\mathbb{E}[X^+|\mathcal{F}]$ and $\mathbb{E}[X^-|\mathcal{F}]$. For $F \in \mathcal{F}$, we obtain that

$$
\int_F Z \, d\mathbb{P} = \int_F \mathbb{E}[X^+|\mathcal{F}] \, d\mathbb{P} - \int_F \mathbb{E}[X^-|\mathcal{F}] \, d\mathbb{P} = \int_F X^+ \, d\mathbb{P} - \int_F X^- \, d\mathbb{P} = \int_F X \, d\mathbb{P} .
$$
■

For $Y : \Omega \to \Omega'$ a random variable, define

$$
\mathbb{E}[X|Y] = \mathbb{E}\left[X|\sigma(Y)\right] .
$$

[1] [Falk, 2019, Satz 8.21], there without the notion of absolute continuity of measures, $\nu \ll \mu$.

First examples

Proposition 2.6. *1. If $\sigma(X) \subseteq \mathcal{F}$, then $\mathbb{E}[X|\mathcal{F}] = X$ a.s.*
2. If X is independent of $\mathcal{F}$, then $\mathbb{E}[X|\mathcal{F}] = \mathbb{E}[X]$ a.s.

Proof. 1. readily follows from (2.2). Apparently, the constant random variable $\mathbb{E}[X]$ is $\mathcal{F}$-measurable. For $F \in \mathcal{F}$, we have in case of the independence that

$$\int_F X \, d\mathbb{P} = \mathbb{E}[\mathbf{1}_F X] = \mathbb{E}[\mathbf{1}_F]\,\mathbb{E}[X] = \mathbb{P}(F)\,\mathbb{E}[X] = \int_F \mathbb{E}[X]\, d\mathbb{P}\,. \qquad \blacksquare$$

For $X = \mathbf{1}_A$, $A \in \mathcal{A}$, we put

$$\mathbb{P}(A|\mathcal{F}) = \mathbb{E}[\mathbf{1}_A|\mathcal{F}]\,,$$

which defines the conditional probability of A given $\mathcal{F}$.

If Z and W are independent random variables with distributions P_Z and P_W, and $h(z,w)$ a measurable function with $\mathbb{E}\left[|h(Z,W)|\right] < \infty$, then

$$\mathbb{E}\left[h(Z,W)|Z\right] = g(Z) \quad \text{with} \quad g(z) = \int h(z,w)\, dP_W(w)\,. \tag{2.3}$$

To prove this, observe that the function g is measurable by Fubini, and thus $g(Z)$ is $\sigma(Z)$-measurable. Moreover, for $F = \{Z \in B\}$:

$$\mathbb{E}\left[\mathbf{1}_F\, g(Z)\right] = \int \mathbf{1}_B(z) g(z)\, dP_Z(z) = \int \mathbf{1}_B(z) \int h(z,w)\, dP_W(w)\, dP_Z(w) = \mathbb{E}\left[\mathbf{1}_F\, h(Z,W)\right],$$

since the joint distribution of Z and W is given by the product measure.

Computation based on a conditional density

Let $X : \Omega \to \mathbb{R}$ be a random variable and $Z : \Omega \to \mathbb{R}^k$ a random vector. Assume that the joint distribution $P_{X,Z}$ admits a density $f_{X,Z}$ w.r.t. the product measure $\mu \otimes \nu$ of two σ-finite measures (on $\mathbb{R}$ or $\mathbb{R}^k$). Then, Z has a marginal density w.r.t. ν

$$f_Z(z) = \int_{\mathbb{R}} f_{X,Z}(x,z)\, d\mu(x), \quad z \in \mathbb{R}^k\,.$$

Define the conditional density of X given the event $\{Z = z\}$ by $f_{X|Z}(x|z) = f_{X,Z}(x,z)/f_Z(z)$, if $f_Z(z) > 0$. Assume that

$$\mathbb{E}\left[|X|\right] = \int_{\mathbb{R}} \int_{\mathbb{R}^k} |x|\, f_{X,Z}(x,z)\, d\mu(x)\, d\nu(z) < \infty.$$

By Fubini's theorem there exists up to a ν-null set the function

$$g(z) = \int_{\mathbb{R}} x\, f_{X|Z}(x|z)\, d\mu(x),$$

and we set $g(z)$ constant, e.g. $= 0$, else. By Fubini g is measurable, and we claim that

$$\mathbb{E}[X|Z] = g(Z).$$

The measurability property is clear, since g is measurable. For $F = \{Z \in B\}$, we have that

$$\begin{aligned}
\mathbb{E}\left[\mathbf{1}_F\, g(Z)\right] &= \int_B g(z)\, f_Z(z)\, d\nu(z) \\
&= \int_B \left(\int_{\mathbb{R}} x\, f_{X|Z}(x|z)\, d\mu(x)\right) f_Z(z)\, d\nu(z) \\
&= \int_{\mathbb{R}} \int_B x\, f_{X,Z}(x,z)\, d\nu(z)\, d\mu(x) = \mathbb{E}\left[\mathbf{1}_F X\right].
\end{aligned}$$

Properties of conditional expectation

Proposition 2.7. *Let X, Y be integrable random variables, $\mathcal{F} \subseteq \mathcal{A}$ a sub-σ-field and $\lambda \in \mathbb{R}$. Then, it holds that*
1. $\mathbb{E}\left[\mathbb{E}[X|\mathcal{F}]\right] = \mathbb{E}[X]$.
2. $\mathbb{E}\left[\lambda X + Y|\mathcal{F}\right] = \lambda\,\mathbb{E}[X|\mathcal{F}] + \mathbb{E}[Y|\mathcal{F}]$ *a.s.* *(linearity)*
3. If $X \leq Y$ a.s., then $\mathbb{E}[X|\mathcal{F}] \leq \mathbb{E}[Y|\mathcal{F}]$ a.s. *(monotonicity)*
4. $|\mathbb{E}[X|\mathcal{F}]| \leq \mathbb{E}\left[|X|\,|\mathcal{F}\right]$ *a.s.*

Proof. 1.: This is a repetition from Proposition 2.3.
2.: It is obvious that $\lambda\,\mathbb{E}[X|\mathcal{F}] + \mathbb{E}[Y|\mathcal{F}]$ is $\mathcal{F}$-measurable. For $F \in \mathcal{F}$, we have that

$$\begin{aligned}\int_F \left(\lambda\,\mathbb{E}[X|\mathcal{F}] + \mathbb{E}[Y|\mathcal{F}]\right) d\mathbb{P} &= \lambda \int_F \mathbb{E}[X|\mathcal{F}]\,d\mathbb{P} + \int_F \mathbb{E}[Y|\mathcal{F}]\big)\,d\mathbb{P} \\ &= \lambda \int_F X\,d\mathbb{P} + \int_F Y\,d\mathbb{P} = \int_F \left(\lambda X + Y\right) d\mathbb{P}\,.\end{aligned}$$

3.: By linearity it suffices to show for $Z = Y - X \geq 0$, that $\mathbb{E}[Z|\mathcal{F}] \geq 0$. Since for all $F \in \mathcal{F}$, it holds that

$$\int_F \mathbb{E}[Z|\mathcal{F}]\,d\mathbb{P} = \int_F Z\,d\mathbb{P} \geq 0\,,$$

and $\mathbb{E}[Z|\mathcal{F}]$ is $\mathcal{F}$-measurable, we obtain that $\mathbb{E}[Z|\mathcal{F}] \geq 0$ a.s.
4.: Since $X \leq |X|$, and $-X \leq |X|$, this is implied by monotonicity and linearity. ∎

Proposition 2.8 (Jensen's inequality for conditional expectation)**.** *Let X be a random variable, and φ a convex function with $\mathbb{E}\left[|X|\right] < \infty$, $\mathbb{E}\left[|\varphi(X)|\right] < \infty$. Then, it holds that*

$$\varphi\left(\mathbb{E}[X|\mathcal{F}]\right) \leq \mathbb{E}\left[\varphi(X)|\mathcal{F}\right].$$

This is included in the general formulation of Jensen's inequality as in Satz T.3.8.5 from the lecture Stochastik 2 by Rainer Göb in the summer semester 2020. For a proof, let me refer to [Bibinger and Holzmann, 2019, Satz 4.8].

Examples 2.9. a. $\left(\mathbb{E}[X|\mathcal{F}]\right)^2 \leq \mathbb{E}[X^2|\mathcal{F}]$,
b. $\exp\left(\mathbb{E}[X|\mathcal{F}]\right) \leq \mathbb{E}\left[\exp(X)|\mathcal{F}\right]$. •

Corollary 2.10. If X is a random variable, with $\mathbb{E}\left[|X|^p\right] < \infty$, for some $p \geq 1$, this implies that $\mathbb{E}\left[\left|\mathbb{E}[X|\mathcal{F}]\right|^p\right] < \infty$, and it holds true that

$$\|\mathbb{E}[X|\mathcal{F}]\|_p \leq \|X\|_p := \left(\mathbb{E}[|X|^p]\right)^{1/p}.$$

Proof. The function $\varphi(x) = |x|^p$ is convex and thus $|\mathbb{E}[X|\mathcal{F}]|^p \leq \mathbb{E}\left[|X|^p|\mathcal{F}\right]$. This yields that

$$\mathbb{E}\left[\left|\mathbb{E}[X|\mathcal{F}]\right|^p\right] \leq \mathbb{E}\left[\mathbb{E}\left[|X|^p|\mathcal{F}\right]\right] = \mathbb{E}\left[|X|^p\right].$$ ∎

Proposition 2.11. *Let X, Y be random variables with $\mathbb{E}\left[|Y|\right] < \infty$, $\mathbb{E}\left[|XY|\right] < \infty$, and let X be $(\mathcal{F} - \mathcal{B})$-measurable. Then, it holds true that*

$$\mathbb{E}[XY|\mathcal{F}] = X\,\mathbb{E}[Y|\mathcal{F}]\,.$$

Proof. The $\mathcal{F}$-measurability of $X\,\mathbb{E}[Y|\mathcal{F}]$ is obvious. Concerning the characterizing property, observe that
a. If $X = \mathbf{1}_F$, $F \in \mathcal{F}$, this is readily implied by (2.2) for $\mathbb{E}[X|\mathcal{F}]$.
b. We can generalize the identity to step functions X, based on linearity.
c. For $X \geq 0, Y \geq 0$, we approximate X in a monotone way by a sequence of step functions, and use monotone convergence.
d. In general, we are left to decompose X and Y in their positive and negative parts. ■

Analogously to the poofs for expected values one can establish Hölder, Cauchy-Schwarz, and Minkowski inequalities for conditional expectations.

Projection properties

Theorem 2.12 (General tower rule). *Let X be an integrable random variable and $\mathcal{F}_1 \subseteq \mathcal{F}_2$ a sub-σ-field of $\mathcal{A}$. It holds true that*
a. $\mathbb{E}\left[\mathbb{E}[X|\mathcal{F}_1]|\mathcal{F}_2\right] = \mathbb{E}[X|\mathcal{F}_1]$,
b. $\mathbb{E}\left[\mathbb{E}[X|\mathcal{F}_2]|\mathcal{F}_1\right] = \mathbb{E}[X|\mathcal{F}_1]$.

Proof. a. $\mathbb{E}[X|\mathcal{F}_1]$ is apparently $\mathcal{F}_2$-measurable, such that it equals a.s. the conditional expectation.
b. The $\mathcal{F}_1$-measurability of $\mathbb{E}\left[\mathbb{E}[X|\mathcal{F}_2]|\mathcal{F}_1\right]$ is obvious. For $F \in \mathcal{F}_1$, we have that $F \in \mathcal{F}_2$, and thus

$$\int_F \mathbb{E}\left[\mathbb{E}[X|\mathcal{F}_2]|\mathcal{F}_1\right] d\mathbb{P} = \int_F \mathbb{E}[X|\mathcal{F}_2]\, d\mathbb{P} = \int_F X\, d\mathbb{P} = \int_F \mathbb{E}[X|\mathcal{F}_1]\, d\mathbb{P}\,.$$ ■

We call the identities from the theorem as well iterated conditional expectations.

By Proposition 2.11 we have for L_2 random variables that the definition of conditional expectation $\mathbb{E}[X|Y]$ coincides with the orthogonal projection from $L_2(\Omega,\mathcal{A},\mathbb{P})$ on $L_2(\Omega,\sigma(Y),\mathbb{P})$. In particular, for a $\sigma(Y)$-measurable, integrable random variable Z, we obtain that

$$\mathbb{E}\left[Z\,\mathbb{E}[X|Y]\right] = \mathbb{E}[ZX] \;\Leftrightarrow\; \mathbb{E}\left[Z\big(X - \mathbb{E}[X|Y]\big)\right] = 0\,, \tag{2.4}$$

and hence that

$$\begin{aligned}
\mathbb{E}[(X-Z)^2] &= \mathbb{E}\left[\big(X - \mathbb{E}[X|Y] + \mathbb{E}[X|Y] - Z\big)^2\right] \\
&= \mathbb{E}\left[\big(X - \mathbb{E}[X|Y]\big)^2\right] + \mathbb{E}\left[\big(\mathbb{E}[X|Y] - Z\big)^2\right] + 2\,\mathbb{E}\left[\big(X - \mathbb{E}[X|Y]\big)\big(\mathbb{E}[X|Y] - Z\big)\right] \\
&\geq \mathbb{E}\left[\big(X - \mathbb{E}[X|Y]\big)^2\right].
\end{aligned}$$

This shows that $\mathbb{E}[X|Y]$ minimizes the functional $\mathbb{E}[(X-g(Y))]^2$ within the class of $\sigma(Y)$-measurable random variables almost surely uniquely. This motivates $\mathbb{E}[X|Y]$ as the best approximation or prediction of X, given the random variable Y.

2.2 Definition, examples and first results

Definition 2.13. *Let $(\Omega,\mathcal{A})$ be a measurable space. An increasing sequence of sub-σ-fields $(\mathcal{F}_n)_{n\in\mathbb{N}}$ is called a* ***filtration****. A tuple $(\Omega,\mathcal{A},(\mathcal{F}_n)_{n\in\mathbb{N}},\mathbb{P})$, with a probability space $(\Omega,\mathcal{A},\mathbb{P})$ and a filtration $(\mathcal{F}_n)_{n\in\mathbb{N}}$ is called a* ***filtered probability space****. A stochastic process $X = (X_n)_{n\in\mathbb{N}}$, taking values in some measurable space $(E,\mathcal{E})$ is* ***adapted*** *with respect to a filtration $(\mathcal{F}_n)_{n\mathbb{N}}$ (short $(\mathcal{F}_n)$-adapted), if for all $n \in \mathbb{N}$, the random variable X_n is $(\mathcal{F}_n - \mathcal{E})$-measurable.*

Looking at filtrations $(\mathcal{F}_n)_{n\in\mathbb{N}}$, we always assume in the sequel that the sub-σ-fields are defined on some measurable space $(\Omega,\mathcal{A})$. If we state that some process is $(\mathcal{F}_n)$-adapted, we assume in particular that $(\Omega,\mathcal{A},(\mathcal{F}_n)_{n\in\mathbb{N}},\mathbb{P})$ is a filtered probability space. We write $(E,\mathcal{E})$ in this chapter for the state space of considered stochastic processes.

Example 2.14. 1. For a stochastic process $(X_n)_{n\in\mathbb{N}}$, setting

$$\mathcal{F}_n^X = \sigma(X_1,\dots,X_n)\,,$$

$(\mathcal{F}_n^X)_{n\in\mathbb{N}}$ is a filtration. The filtration $(\mathcal{F}_n^X)_{n\in\mathbb{N}}$ is the natural filtration generated by $(X_n)_{n\in\mathbb{N}}$. This is the smallest filtration with respect to which (X_n) is adapted.
2. Let $X=(X_n)_{n\in\mathbb{N}}$ be a $(\mathcal{F}_n)$-adapted process. For measurable mappings $F_n : E^n \to E'$, $Y=(Y_n)_{n\in\mathbb{N}}$, defined by $Y_n = F_n(X_1,\dots,X_n)$, is a $(\mathcal{F}_n)$-adapted process taking values in E'. •

A real-valued stochastic process $(X_t)_{t\in T}$ is called integrable, if $\mathbb{E}[|X_t|]<\infty$, $\forall t\in T$. Analogously, $(X_t)_{t\in T}$ is square integrable, if $\mathbb{E}[X_t^2]<\infty$, $\forall t\in T$.

Definition 2.15. *Let $(X_n)_{n\in\mathbb{N}}$ be an integrable $(\mathcal{F}_n)$-adapted process.*
*1. (X_n) is a $(\mathcal{F}_n)$-**martingale**, if for all $n,m\in\mathbb{N}$, with $n\le m$, it holds that*

$$\mathbb{E}[X_m|\mathcal{F}_n] = X_n, \quad \textit{almost surely}.$$

*2. (X_n) is a $(\mathcal{F}_n)$-**submartingale**, if for all $n,m\in\mathbb{N}$, with $n\le m$, it holds that*

$$\mathbb{E}[X_m|\mathcal{F}_n] \ge X_n, \quad \textit{almost surely}.$$

*3. (X_n) is a $(\mathcal{F}_n)$-**supermartingale**, if $(-X_n)$ is a $(\mathcal{F}_n)$-submartingale.*

The definition will be generalized to processes with some index set $T\subset\mathbb{R}$ below. We can use the definition analogously for the index set $\mathbb{N}_0$. In this chapter we consider discrete-time martingales with T countable. Later we shall be interested in continuous-time processes, e.g. $T=\mathbb{R}_+$. Since the definition uses an expectation, it hinges on the probability measure $\mathbb{P}$ also.

To show that an integrable adapted process $(X_n)_{n\in\mathbb{N}}$ is a (sub-)martingale it suffices to verify that

$$\mathbb{E}[X_{n+1}|\mathcal{F}_n] \overset{(\ge)}{=} X_n \quad \text{a.s. for all } n\in\mathbb{N}\,. \tag{2.5}$$

By an induction over k, Theorem 2.12 yields that almost surely

$$\mathbb{E}[X_{n+k}|\mathcal{F}_n] = \mathbb{E}\left[\mathbb{E}[X_{n+k}|\mathcal{F}_{n+k-1}]|\mathcal{F}_n\right] \overset{(\ge)}{=} \mathbb{E}[X_{n+k-1}|\mathcal{F}_n] \overset{(\ge)}{=} X_n\,.$$

The condition (2.5) is equivalent to

$$\int_F X_{n+1}\,d\mathbb{P} \overset{(\ge)}{=} \int_F X_n\,d\mathbb{P}\,, \qquad \forall\, F\in\mathcal{F}_n\,. \tag{2.6}$$

Any $(\mathcal{F}_n)$-(sub-)martingale is as well a $(\mathcal{F}_n^X)$-(sub-)martingale, since

$$\mathbb{E}[X_m|\mathcal{F}_n^X] = \mathbb{E}[\mathbb{E}[X_m|\mathcal{F}_n]|\mathcal{F}_n^X] \overset{(\ge)}{=} \mathbb{E}[X_n|\mathcal{F}_n^X] = X_n \text{ a.s.}$$

Examples 2.16. 1. Let $(\xi_n)_{n\in\mathbb{N}}$ be a sequence of independent, real-valued, integrable random variables with $\mathbb{E}[\xi_n]=0$, for all $n\in\mathbb{N}$. Then, $(X_n)_{n\in\mathbb{N}}$ defined by

$$X_n = \sum_{k=1}^{n} \xi_k\,,$$

is a $(\mathcal{F}_n^\xi)$-martingale, since (X_n) is $(\mathcal{F}_n^\xi)$-adapted and

$$\mathbb{E}[|X_n|] = \mathbb{E}[|\xi_1 + \cdots + \xi_n|] \leq \mathbb{E}[|\xi_1|] + \cdots + \mathbb{E}[|\xi_n|] < \infty,$$

such that (X_n) is integrable, and

$$\mathbb{E}[X_{n+1}|\mathcal{F}_n^\xi] = \mathbb{E}[X_n + \xi_{n+1}|\mathcal{F}_n^\xi] = X_n + \mathbb{E}[\xi_{n+1}] = X_n\,,$$

almost surely by the independence of ξ_{n+1} and $(\xi_1, \ldots, \xi_n)$. If $\mathbb{E}[\xi_n] \geq 0$ instead, (X_n) is a $(\mathcal{F}_n^\xi)$-submartingale.

2. For X some real-valued, integrable random variable and $(\mathcal{F}_n)$ some filtration, $X_n = \mathbb{E}[X|\mathcal{F}_n]$ defines a $(\mathcal{F}_n)$-martingale by the properties of conditional expectations.

3. Let $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n\in\mathbb{N}}, \mathbb{P})$ be a filtered probability space and let $\mathbb{Q}$ be a finite measure on $(\Omega, \mathcal{F})$, which is absolutely continuous with respect to $\mathbb{P}$, $\mathbb{Q} \ll \mathbb{P}$. Then, $(Z_n)_{n\in\mathbb{N}}$ defined by the Radon-Nikodym densities

$$Z_n := \frac{d\,\mathbb{Q}}{d\,\mathbb{P}}\bigg|_{\mathcal{F}_n} := \frac{d\,\mathbb{Q}\big|_{\mathcal{F}_n}}{d\,\mathbb{P}\big|_{\mathcal{F}_n}}$$

is a $(\mathcal{F}_n)$-martingale w.r.t. $\mathbb{P}$ on $(\Omega, \mathcal{F})$.
By construction, (Z_n) is $\mathcal{F}_n$-adapted, and for $A \in \mathcal{F}_n$, what implies that $A \in \mathcal{F}_{n+1}$, we have that

$$\int_A Z_{n+1}\, d\,\mathbb{P} = \mathbb{Q}(A) = \int_A Z_n\, d\,\mathbb{P}\,.$$

4. Martingales play a key role for modelling fair gambling and price models at financial markets. Efficient markets do not allow for systematic gains without taking any risk. One says there is no arbitrage in the market. If the price of a financial asset is a martingale the expected profit, given the information at the current time, is always zero. Now, let $(X_n)_{n\geq 0}$ be a $(\mathcal{F}_n)$-martingale which models the price of some financial asset. The sequence of increments, $(Y_n)_{n\geq 1}$, with $Y_n = X_n - X_{n-1}$, give the price changes (returns) from times $n-1$ to times n. This sequence is $(\mathcal{F}_n)_{n\geq 1}$-adapted and for all n: $\mathbb{E}[Y_n|\mathcal{F}_{n-1}] = 0$. We call such a sequence a martingale difference sequence. The sum of a martingale difference sequence yields a martingale. Assume at discrete time n, e.g. n is a day, we decide to buy H assets of a stock with price process $(X_n)_{n\geq 0}$. This decision needs to be taken based on the information available up to time n, modelled by $\mathcal{F}_n$. The investment with H assets is then fixed over the period from time n up to time $n+1$. The return of this investment in this period is $HY_{n+1} = H(X_{n+1} - X_n)$. If we index $H = H_{n+1}$, for the investment up to time $n+1$, we obtain some process (H_n), where H_n is $\mathcal{F}_{n-1}$-measurable. Such a process is called predictable w.r.t. $(\mathcal{F}_n)_{n\geq 0}$. The gains of the total investment strategy at time n are given by

$$M_n = \sum_{k=1}^n H_k Y_k = \sum_{k=1}^n H_k (X_k - X_{k-1})\,. \tag{2.7}$$

The process $(M_n)_{n\geq 1}$ is $(\mathcal{F}_n)$-adapted, and since $(H_n)_{n\geq 1}$ are bounded random variables, (M_n) is integrable. Moreover,

$$\mathbb{E}\left[M_n|\mathcal{F}_{n-1}\right] = \mathbb{E}\left[M_{n-1} + H_n Y_n|\mathcal{F}_{n-1}\right] = M_{n-1} + H_n\,\mathbb{E}[Y_n|\mathcal{F}_{n-1}] = M_{n-1}\,.$$

Any such investment strategy thus yields a process of gains which is a martingale with $\mathbb{E}[M_n] = \mathbb{E}[M_1] = 0$. •

Proposition 2.17. *Let $(X_n)_{n\geq 0}$ be a $(\mathcal{F}_n)$-adapted, integrable process. $(X_n)_{n\geq 0}$ is a $(\mathcal{F}_n)$-martingale, iff for any $(\mathcal{F}_n)$-predictable process $(H_n)_{n\geq 1}$, with H_n bounded for all n and for $(M_n)_{n\geq 1}$ defined as in (2.7) it holds that $\mathbb{E}[M_n] = \mathbb{E}[M_1] = 0$, for all $n \geq 1$.*

Proof. If (X_n) is a martingale, we have shown that (M_n) is as well a martingale and it holds that $\mathbb{E}[M_n] = 0$, for all $n \geq 1$.

If $F \in \mathcal{F}_{n-1}$, set $H_n = \mathbf{1}_F$, and all other $H_k = 0$. This yields $0 = \mathbb{E}[M_n] = \mathbb{E}\left[\mathbf{1}_F(X_n - X_{n-1})\right]$, what proves (2.6) and thus (2.5). ∎

We conclude an analogous result for submartingales.

Corollary 2.18. Let $(X_n)_{n\geq 0}$ be a $(\mathcal{F}_n)_{n\geq 0}$-adapted, integrable process. $(X_n)_{n\geq 0}$ is a $(\mathcal{F}_n)$-submartingale, iff for any $(\mathcal{F}_n)$-predictable process $(H_n)_{n\geq 1}$, with $H_n \geq 0$ and H_n bounded for all n and for $(M_n)_{n\geq 1}$ defined as in (2.7) it holds that $\mathbb{E}[M_n] \geq 0$, for all $n \geq 1$. In this case $(M_n)_{n\geq 1}$ is as well a $(\mathcal{F}_n)$-submartingale.

Proof. If (X_n) is a submartingale, we obtain that

$$\mathbb{E}\left[M_n|\mathcal{F}_{n-1}\right] = M_{n-1} + H_n\left(\mathbb{E}[X_n|\mathcal{F}_{n-1}] - X_{n-1}\right) \geq M_{n-1}\,,$$

and in particular $\mathbb{E}[M_n] \geq \mathbb{E}[M_1] = 0$. For the reverse claim we start as in the proposition and obtain that $0 \leq \mathbb{E}[M_n] = \mathbb{E}\left[\mathbf{1}_F(X_n - X_{n-1})\right]$, what proves the submartingale property. ∎

Theorem 2.19 (*Doob decomposition*)**.** *Let (X_n) be a $\mathcal{F}_n$-adapted, integrable process. There exists (almost surely) a unique decomposition $X_n = M_n + A_n$, where (M_n) is a martingale and (A_n) is predictable with $A_1 = 0$. (X_n) is a submartingale, iff (A_n) is monotone increasing, i.e. $A_n \leq A_{n+1}$.* ▪

Proof. We prove uniqueness first. We have $A_1 = 0$, and thus $M_1 = X_1$. Since $M_n = X_n - A_n$, with (A_n) predictable and (M_n) a martingale, taking conditional expectations we get that $M_{n-1} = \mathbb{E}[X_n|\mathcal{F}_{n-1}] - A_n$, and solving w.r.t. A_n by induction

$$M_n = X_n - \mathbb{E}[X_n|\mathcal{F}_{n-1}] + M_{n-1} = X_1 + \sum_{k=2}^{n}\left(X_k - \mathbb{E}[X_k|\mathcal{F}_{k-1}]\right).$$

Then, we obtain that

$$\begin{aligned} A_n - A_{n-1} &= \mathbb{E}\left[A_n - A_{n-1}|\mathcal{F}_{n-1}\right] \\ &= \mathbb{E}\left[X_n - M_n - (X_{n-1} - M_{n-1})|\mathcal{F}_{n-1}\right] = \mathbb{E}[X_n|\mathcal{F}_{n-1}] - X_{n-1}\,, \end{aligned}$$

and by induction

$$A_n = \sum_{k=2}^{n}\left(\mathbb{E}[X_k|\mathcal{F}_{k-1}] - X_{k-1}\right).$$

To prove the existence, we show that these choices of M_n and A_n satisfy the properties in the theorem. We have that

$$\mathbb{E}[X_{n+1}|\mathcal{F}_n] - X_n = A_{n+1} - A_n \geq 0 \quad \Leftrightarrow \quad A_{n+1} \geq A_n\,.$$ ∎

The last topic of this section is to study (sub-)martingales under convex transformations.

Proposition 2.20.

a) *Let (X_n) be a $(\mathcal{F}_n)$-martingale, and $\varphi : \mathbb{R} \to \mathbb{R}$ a convex mapping and $\mathbb{E}[|\varphi(X_n)|] < \infty$, for all $n \in \mathbb{N}$. Then, $\varphi((X_n)) := (\varphi(X_n))_{n\in\mathbb{N}}$ is a $(\mathcal{F}_n)$-submartingale.*

b) *Let (X_n) be a $(\mathcal{F}_n)$-submartingale, and $\varphi : \mathbb{R} \to \mathbb{R}$ a convex mapping and monotone increasing and $\mathbb{E}[|\varphi(X_n)|] < \infty$, for all $n \in \mathbb{N}$. Then, $\varphi((X_n))$ is a $(\mathcal{F}_n)$-submartingale.*

Proof. With the Jensen inequality for conditional expectations, we obtain in both cases for $n \in \mathbb{N}$:

$$\mathbb{E}[\varphi(X_{n+1})|\mathcal{F}_n] \geq \varphi(\mathbb{E}[X_{n+1}|\mathcal{F}_n])\,.$$

In case a) the last expression equals $\varphi(X_n)$. In case b), it is $\geq \varphi(X_n)$, since φ is monotone increasing. ∎

Examples 2.21. a. Let $(X_n)_{n\geq 0}$ be a submartingale and $a \in \mathbb{R}$. Then, $(X_n - a)^+$, $n \geq 0$, is a submartingale. This holds since the function $\varphi(x) = (x-a)^+$ is convex and monotone increasing and $(X_n - a)^+ \leq X_n^+ + |a|$ is integrable. This example is crucial for the price of a European call option.

b. Let $(\xi_n)_{n\in\mathbb{N}}$ be i.i.d. real-valued random variables with $\mathbb{E}[\xi_n] = 0$, $\mathbb{E}[\xi_n^2] = 1$. Here, $X_n = \xi_1 + \ldots + \xi_n$ defines a martingale (X_n), and thus (X_n^2) is a submartingale. Since $\mathbb{E}\left[X_n^2 | \mathcal{F}_{n-1}^{\xi}\right] = X_{n-1}^2 + 1$, $M_n = X_n^2 - n$ is a martingale and $X_n^2 = M_n + n$ the Doob decomposition. •

2.3 Stopping times and optional stopping

Define $\bar{\mathbb{N}} := \mathbb{N} \cup \{\infty\}$ and consider the measurable space $(\bar{\mathbb{N}}, \mathcal{P}(\bar{\mathbb{N}}))$, with the power set $\mathcal{P}(\bar{\mathbb{N}})$ as the standard σ-field on $\bar{\mathbb{N}}$.

Definition 2.22. *Let $(\mathcal{F}_n)_{n\in\mathbb{N}}$ be a filtration. A mapping $T : \Omega \to \bar{\mathbb{N}}$ is a $(\mathcal{F}_n)$* ***stopping time****, if for all $n \in \mathbb{N}$:*

$$\{T \leq n\} \in \mathcal{F}_n\,.$$

Let $T : \Omega \to \bar{\mathbb{N}}$ be a $(\mathcal{F}_n)$ stopping time. If we define for the given filtration $(\mathcal{F}_n)$:

$$\mathcal{F}_\infty = \sigma\Big(\mathcal{F}_n,\, n \geq 1\Big),$$

the random variable T is $\mathcal{F}_\infty$-measurable, since

$$\{T = \infty\} = \Big(\bigcup_{k\in\mathbb{N}} \underbrace{\{T \leq k\}}_{\in \mathcal{F}_k \subset \mathcal{F}_\infty}\Big)^{\complement} \in \mathcal{F}_\infty\,.$$

$T : \Omega \to \bar{\mathbb{N}}$ is a $(\mathcal{F}_n)$ stopping time iff

$$\{T = n\} \in \mathcal{F}_n \quad \forall n \in \mathbb{N}\,. \tag{2.8}$$

If T is a stopping time, it holds that

$$\{T = n\} = \{T \leq n\} \cap \{T \leq n-1\}^{\complement} \in \mathcal{F}_n\,.$$

Since $\{T \leq n\} = \bigcup_{k=1}^n \{T = k\} \in \mathcal{F}_n$, (2.8) ensures that T is a stopping time.

Examples 2.23 (*Entry times*)**.** Let $(X_n)_{n\in\mathbb{N}}$ be a $(E, \mathcal{E})$-valued $(\mathcal{F}_n)$-adapted stochastic process.

1. For $A \in \mathcal{E}$, $T := T_A := \inf\{n \in \mathbb{N} : X_n \in A\}$, with the convention that $\inf \emptyset = \infty$, is a $(\mathcal{F}_n)$ stopping time, since

$$\{T \leq n\} = \bigcup_{k=1}^{n} \underbrace{\{X_k \in A\}}_{\in \mathcal{F}_k} \in \mathcal{F}_n\,.$$

2. For another stopping time S and $A \in \mathcal{E}$, $T := \inf\{n > S : X_n \in A\}$ is a stopping time, since

$$\{T \leq n\} = \bigcup_{1\leq k<l\leq n} \Big(\{S \leq k\} \cap \{X_l \in A\}\Big) \in \mathcal{F}_n\,.$$

Therefore, first and second entry times and by induction kth entry times are stopping times. •

Proposition 2.24. *For a sequence $(T_m)_{m\in\mathbb{N}}$ of $(\mathcal{F}_n)$ stopping times, $\sup_{m\in\mathbb{N}} T_m$ and $\inf_{m\in\mathbb{N}} T_m$ are $(\mathcal{F}_n)$ stopping times.*

Proof.
$$\{\sup_m T_m \le n\} = \bigcap_{m=1}^{\infty} \{T_m \le n\} \in \mathcal{F}_n, \qquad \{\inf_m T_m \le n\} = \bigcup_{m=1}^{\infty} \{T_m \le n\} \in \mathcal{F}_n.$$
■

For events A and B, $\min(T_A, T_B) = T_{A\cup B}$. Let (X_n) be an adapted process, and T a stopping time. The at time T stopped process (X_n^T) is defined by

$$X_n^T = X_{T\wedge n}, \quad n \ge 1 \qquad \big(\omega \mapsto X_{T(\omega)\wedge n}(\omega)\big).$$

Since for $A \in \mathcal{E}$, it holds that

$$\{X_n^T \in A\} = \big(\{T \ge n\} \cap \{X_n \in A\}\big) \cup \bigcup_{k=1}^{n-1} \big(\{T = k\} \cap \{X_k \in A\}\big),$$

(X_n^T) is $(\mathcal{F}_n)$-adapted.

For a finite stopping time S, in general X_S $(\omega \mapsto X_{S(\omega)}(\omega))$ is a random variable, what we prove below.

If (X_n) is a real-valued, integrable process, this holds true as well for (X_n^T), since $|X_{T\wedge n}| \le |X_1| + \ldots + |X_n|$. The next theorem gives the main result about stopped martingales.

Theorem 2.25. *Let $(X_n)_{n\in\mathbb{N}}$ be a $(\mathcal{F}_n)$-(sub-)martingale and T a stopping time. Then, the stopped process $(X_{T\wedge n})_{n\in\mathbb{N}}$ is as well a $(\mathcal{F}_n)$-(sub)-martingale. In the martingale case, it holds that $\mathbb{E}[X_{T\wedge n}] = \mathbb{E}[X_1]$, for all n.*

Proof. It is already clear that the process is adapted and integrable. We have that $X_n^T - X_{n-1}^T = (X_n - X_{n-1})\mathbf{1}_{T\ge n}$. Since $\{T \ge n\} = \{T \le n-1\}^{\complement} \in \mathcal{F}_{n-1}$, it holds that

$$\mathbb{E}[X_n^T|\mathcal{F}_{n-1}] - X_{n-1}^T = \mathbf{1}_{\{T\ge n\}}\big(\mathbb{E}[X_n|\mathcal{F}_{n-1}] - X_{n-1}\big) \overset{(\ge)}{=} 0.$$

Since $X_1^T = X_1$, the equality for the expected values follows. ■

Example 2.26 (*Simple symmetric random walk*)**.** Let $\xi_1, \xi_2, \ldots$ be independent random variables with $\mathbb{P}(\xi_n = 1) = \mathbb{P}(\xi_n = -1) = 1/2$, and $X_n = \xi_1 + \ldots + \xi_n$. For $a, b \in \mathbb{N}$, consider the stopping time $T_{\{a,-b\}}$, defined by

$$T_{\{a,-b\}} = \inf\{n \ge 1 : X_n = a \quad \text{or} \quad X_n = -b\}.$$

We show that $\mathbb{P}(T_{\{a,-b\}} < \infty) = 1$. Thereto, set

$$E_k = \big\{\xi_{k(a+b)+1} = 1, \ldots, \xi_{(k+1)(a+b)} = 1\big\}, \qquad k = 0, 1, \ldots.$$

The events $E_0, E_1, \ldots$ are independent with

$$\mathbb{P}(E_k) = 1/2^{a+b} =: p.$$

Since a run of $a+b$ positive realizations in a row always results in reaching the upper bound a, if the boundaries were not crossed before, we have for $0 \le k \le n-1$:

$$E_k \cap \{T_{\{a,-b\}} \ge k(a+b)+1\} \subseteq \{T_{\{a,-b\}} \le (k+1)(a+b)\} \subseteq \{T_{\{a,-b\}} \le n(a+b)\}.$$

Since

$$E_k \cap \{T_{\{a,-b\}} \le k(a+b)\} \subseteq \{T_{\{a,-b\}} \le (k+1)(a+b)\} \subseteq \{T_{\{a,-b\}} \le n(a+b)\},$$

we obtain that

$$E_0 \cup \ldots \cup E_{n-1} \subseteq \{T_{\{a,-b\}} \le n(a+b)\}.$$

Thereby and by the independence of the (E_j), we obtain that

$$\mathbb{P}\left(T_{\{a,-b\}} > n(a+b)\right) \leq \mathbb{P}\left(\bigcap_{k=0}^{n-1} E_k^{\complement}\right) = (1-p)^n \,,$$

and thus

$$\mathbb{P}\left(T_{\{a,-b\}} = \infty\right) = \lim_{n\to\infty} \mathbb{P}\left(T_{\{a,-b\}} > n(a+b)\right) = \lim_{n\to\infty} (1-p)^n = 0\,.$$

We even conclude the stronger claim that

$$\begin{aligned}\mathbb{E}\left[T_{\{a,-b\}}\right] &= \sum_{n\geq 1} \mathbb{P}(T_{\{a,-b\}} \geq n) \leq (a+b)\sum_{n=0}^{\infty} \mathbb{P}(T_{\{a,-b\}} \geq n(a+b)) \\ &\leq (a+b)\sum_{n=0}^{\infty}(1-p)^n = (a+b)/p < \infty\,.\end{aligned}$$

We want to compute $\mathbb{P}(X_{T_{\{a,-b\}}} = a)$ and $\mathbb{P}(X_{T_{\{a,-b\}}} = -b) = 1 - \mathbb{P}(X_{T_{\{a,-b\}}} = a)$. Since $T_{\{a,-b\}} < \infty$ almost surely, the random variable $X_{T_{\{a,-b\}}}$ takes values a or $-b$. We deduce that

$$\mathbb{E}\left[X_{T_{\{a,-b\}}}\right] = a\mathbb{P}\left(X_{T_{\{a,-b\}}} = a\right) - b\big(1 - \mathbb{P}\left(X_{T_{\{a,-b\}}} = a\right)\big)\,. \tag{2.9}$$

We apply Theorem 2.25. It holds that $0 = \mathbb{E}\left[X_{T_{\{a,-b\}}\wedge n}\right]$, for all n. Since $X_{T_{\{a,-b\}}\wedge n} \to X_{T_{\{a,-b\}}}$, $n\to\infty$, almost surely, and $\left|X_{T_{\{a,-b\}}\wedge n}\right| \leq \max(a,b)$, by dominated convergence and Theorem 2.25 we obtain that

$$0 = \mathbb{E}\left[X_{T_{\{a,-b\}}\wedge n}\right] \overset{n\to\infty}{\longrightarrow} \mathbb{E}\left[X_{T_{\{a,-b\}}}\right], \qquad \text{and thus} \quad \mathbb{E}\left[X_{T_{\{a,-b\}}}\right] = 0\,.$$

With (2.9) we conclude that

$$\mathbb{P}\left(X_{T_{\{a,-b\}}} = a\right) = \frac{b}{a+b}, \qquad \mathbb{P}\left(X_{T_{\{a,-b\}}} = -b\right) = \frac{a}{a+b}\,.$$

Next, we show that $\mathbb{P}(T_a < \infty) = 1$, for $a \in \mathbb{N}$. For any $b \in \mathbb{N}$, the probability to reach a is larger than the probability of reaching a before $-b$, such that

$$\mathbb{P}(T_a < \infty) \geq \mathbb{P}\left(X_{T_{\{a,-b\}}} = a\right) = \frac{b}{a+b} \to 1, \qquad b\to\infty\,,$$

and thus $\mathbb{P}(T_a < \infty) = 1$.

We prove that $\mathbb{E}\left[T_{\{a,-b\}}\right] = ab$. Thereto, consider the martingale $M_n = X_n^2 - n$ (cf. Example 2.21). Since

$$\left|M_{T_{\{a,-b\}}\wedge n}\right| \leq \big(\max(a,b)\big)^2 + T_{\{a,-b\}}\,,$$

and we have shown that $\mathbb{E}\left[T_{\{a,-b\}}\right] < \infty$, the process has an integrable random variable as a majorant. Since, furthermore $\left|M_{T_{\{a,-b\}}\wedge n}\right| \to M_{T_{\{a,-b\}}}$ a.s., we obtain that

$$0 = \mathbb{E}\left[M_{T_{\{a,-b\}}\wedge n}\right] \overset{n\to\infty}{\longrightarrow} \mathbb{E}\left[M_{T_{\{a,-b\}}}\right],$$

and thus

$$\begin{aligned}0 &= a^2\,\mathbb{P}\left(X_{T_{\{a,-b\}}} = a\right) + b^2\,\mathbb{P}\left(X_{T_{\{a,-b\}}} = -b\right) - \mathbb{E}[T_{\{a,-b\}}] \\ &= a^2\frac{b}{a+b} + b^2\frac{a}{a+b} - \mathbb{E}[T_{\{a,-b\}}]\,.\end{aligned}$$

This implies that

$$\mathbb{E}[T_{\{a,-b\}}] = ab\,.$$

Finally, we show that $\mathbb{E}[T_a] = \infty$. It holds that $T_a \geq T_{\{a,-b\}}$, for all $b \in \mathbb{N}$, such that

$$\infty = \sup_b \mathbb{E}[T_{\{a,-b\}}] \leq \mathbb{E}[T_a].$$

•

We introduce a notion to model the information up to some stopping time.

Definition 2.27. *For some $(\mathcal{F}_n)$ stopping time T, denote by*

$$\mathcal{F}_T = \left\{A \in \mathcal{F}_\infty : A \cap \{T \leq n\} \in \mathcal{F}_n \ \ \forall n \in \mathbb{N}\right\}$$

*the σ-field generated by T, also known as **σ-field of T-past**.*

We show that this is well-defined by proving that this is a sub-σ-field of $\mathcal{F}_\infty$. For $A \in \mathcal{F}_T$, we have for all n that

$$A^\complement \cap \{T \leq n\} = \{T \leq n\} \setminus \left(A \cap \{T \leq n\}\right) \in \mathcal{F}_n,$$

such that $A^\complement \in \mathcal{F}_T$. For a $(\mathcal{F}_n)$ stopping time T, $A \in \mathcal{F}_\infty$ satisfies $A \in \mathcal{F}_T$, iff

$$\{T = n\} \cap A \in \mathcal{F}_n \ \ \forall n \in \mathbb{N}.$$

Examples 2.28.

- *Constant stopping times:* If $T(\omega) = m \in \mathbb{N}$, $\forall \omega \in \Omega$, it holds that

$$\{T \leq n\} = \begin{cases} \emptyset, & n < m \\ \Omega, & n \geq m \end{cases},$$

  and thus T is a stopping times w.r.t. to any arbitrary filtration. Furthermore, for some given filtration $(\mathcal{F}_n)$:

$$\mathcal{F}_T = \{A \in \mathcal{F}_\infty : A \cap \{T \leq n\} \in \mathcal{F}_n \ \ \forall n \in \mathbb{N}\} = \mathcal{F}_m.$$

- For $(X_n)_{n\geq 1}$ a stochastic process with natural filtration $(\mathcal{F}_n^X)$, and T a finite $(\mathcal{F}_n^X)$ stopping time, we interpret $\mathcal{F}_T$ as the information provided by $X_1, \ldots, X_T$:
  $A \in \mathcal{F}_T$ iff $\{T = n\} \cap A \in \sigma(X_1, \ldots, X_n)$, such that with Borel sets $B_n \in \mathcal{B}_n$

$$A = \bigcup_{n\geq 1} \left(\{T = n\} \cap \{(X_1, \ldots, X_n) \in B_n\}\right).$$

•

Lemma 2.29. *Let S, T be two stopping times.*
1. The inequality $S \leq T$, implies that $\mathcal{F}_S \subseteq \mathcal{F}_T$.
2. If $(X_n)_{n\in\mathbb{N}}$ is adapted and X_∞ a $\mathcal{F}_\infty$-measurable random variable, or T takes values in $\mathbb{N}$, respectively, then X_T is $\mathcal{F}_T$-measurable.

Proof. 1.: $A \in \mathcal{F}_S \Rightarrow A \in \mathcal{F}_\infty$, and $A \cap \{T \leq n\} = \underbrace{A \cap \{S \leq n\}}_{\in \mathcal{F}_n} \cap \underbrace{\{T \leq n\}}_{\in \mathcal{F}_n} \in \mathcal{F}_n \, \forall n \in \mathbb{N}$.

2.: If $B \in \mathcal{B}$ is a Borel set, we have that

$$\{X_T \in B\} \cap \{T = n\} = \{X_n \in B\} \cap \{T = n\} \in \mathcal{F}_n.$$ ■

Recall that Theorem 2.25 does in general not allow to conclude that $\mathbb{E}[X_T] = \mathbb{E}[X_1]$, for a martingale (X_n), and some stopping time T. In Example 2.26 we have used Theorem 2.25 combined with dominated convergence for such conclusions.

Theorem 2.30 (*Doob's optional stopping theorem*)**.** *Let $S \le T$ be bounded stopping times, i.e. $T \le N$, for some $N \in \mathbb{N}$, and let (X_n) be a submartingale. Then it holds true that*

$$\mathbb{E}[X_T|\mathcal{F}_S] \ge X_S\,,$$

with equality in the case that (X_n) is a martingale. In particular,

$$\mathbb{E}[X_S] \le \mathbb{E}[X_T],$$

with equality, in the case that (X_n) is a martingale.

First we prove this result in case that $T = N$ is a constant stopping time.

Lemma 2.31. *Let (X_n) be a (sub-)martingale and S a stopping time with $S \le N$. Then, it holds true that*

$$X_S \overset{(\le)}{=} \mathbb{E}[X_N|\mathcal{F}_S], \qquad \text{and in particular} \quad \mathbb{E}[X_S] \overset{(\le)}{=} \mathbb{E}[X_N] \quad (= \mathbb{E}[X_1] \text{ in the martingale case}). \tag{2.10}$$

Proof. According to Lemma 2.29, X_S is $\mathcal{F}_S$-measurable. For $A \in \mathcal{F}_S$, i.e. $A \cap \{S = n\} \in \mathcal{F}_n$, by the martingale property it holds that

$$\begin{aligned}\int_A X_S\,d\mathbb{P} &= \sum_{n=1}^N \int_{A\cap\{S=n\}} X_n\,d\mathbb{P}\\ &\overset{(\le)}{=} \sum_{n=1}^N \int_{A\cap\{S=n\}} X_N\,d\mathbb{P} = \int_A X_N\,d\mathbb{P}\,.\end{aligned}$$

∎

Proof of Theorem 2.30. Consider first the martingale case. Since $\mathcal{F}_S \subseteq \mathcal{F}_T$, applying the lemma twice and the general tower rule for conditional expectations yield that

$$\mathbb{E}[X_T|\mathcal{F}_S] = \mathbb{E}\left[\mathbb{E}[X_N|\mathcal{F}_T]|\mathcal{F}_S\right] = \mathbb{E}[X_N|\mathcal{F}_S] = X_S \text{ a.s.}$$

For the submartingale case we consider the Doob decomposition $X_n = M_n + A_n$, using Theorem 2.19, where (M_n) is a martingale and (A_n) predictable and monotone increasing, $A_n \le A_{n+1}$, with $A_1 = 0$. Since $S \le T$, it holds that $A_S \le A_T$, and since A_S is $\mathcal{F}_S$-measurable, by the monotonicity of conditional expectations we obtain that

$$\mathbb{E}[A_T|\mathcal{F}_S] \ge A_S.$$

With the above result for the martingale case, we obtain that

$$\begin{aligned}\mathbb{E}[X_T|\mathcal{F}_S] &= \mathbb{E}[M_T|\mathcal{F}_S] + \mathbb{E}[A_T|\mathcal{F}_S]\\ &= M_S + \mathbb{E}[A_T|\mathcal{F}_S] \ge M_S + A_S = X_S \text{ a.s.}\end{aligned}$$

Finally, we conclude that

$$\mathbb{E}[X_T] = \mathbb{E}[\mathbb{E}[X_T|\mathcal{F}_S]] \ge \mathbb{E}[X_S]\,.$$

∎

Example 2.32. Reconsider the simple symmetric random walk. For $a \in \mathbb{N}$, set

$$T_a = \inf\{n \ge 1 : S_n = a\}\,.$$

We have shown that $T_a < \infty$, almost surely. Thereby, $X_{T_a} = a$, almost surely. Since $\mathbb{E}[X_n] = 0$, (2.10) is violated. •

Corollary 2.33 (*Characterization of martingales with stopping times*)**.** Let (X_n) be an integrable $(\mathcal{F}_n)$-adapted process. The following statements are equivalent:

(i) (X_n) is a $(\mathcal{F}_n)$-martingale.

(ii) $\mathbb{E}[X_T] = \mathbb{E}[X_1]$ for any bounded $(\mathcal{F}_n)$ stopping time T.

Proof. (i) $\Rightarrow$ (ii): This is implied by Lemma 2.31.
(ii) $\Rightarrow$ (i): Let $n \in \mathbb{N}$, $A \in \mathcal{F}_n$, and $T(\omega) := n\mathbf{1}_{A^\complement}(\omega) + (n+1)\mathbf{1}_A(\omega)$. T is a stopping time, since $\{T \leq n\} = A^\complement \in \mathcal{F}_n$. It holds that

$$\mathbb{E}[X_n \mathbf{1}_{A^\complement}] + \mathbb{E}[X_{n+1}\mathbf{1}_A] = \mathbb{E}[X_T] = \mathbb{E}[X_1] = \mathbb{E}[X_n] = \mathbb{E}[X_n \mathbf{1}_{A^\complement}] + \mathbb{E}[X_n \mathbf{1}_A] \,.$$

We conclude that

$$\mathbb{E}[X_{n+1}\mathbf{1}_A] = \mathbb{E}[X_n \mathbf{1}_A] \,,$$

for all $A \in \mathcal{F}_n$, and hence that $\mathbb{E}[X_{n+1}|\mathcal{F}_n] = X_n$, almost surely. ∎

We consider another counterexample where the identity $\mathbb{E}[X_T] = \mathbb{E}[X_1]$, for a martingale (X_n), and some stopping time T, is violated.

Example 2.34 (St. Petersburg game)**.** Consider the roulette strategy “Doubling Up” with which you always win. Bet on black or red. These bets have the maximum odds of winning (almost 50%, not exactly 50% since the 0 is neither black nor red). The payout will be the double bet if the bet is correct, i.e. you win the same amount of money you bet for the spin. Start with a small amount, say the table minimum, and keep betting the same amount until you lose. When this happens, double the size of your bet for the next spin. In this way, in case of a win, you recover the money you lost on the previous round and win something extra on top. If you keep on losing, keep on doubling your bet. As soon as you win, you should restart and bet the smallest amount for the next spin. So, one round of the strategy stops with the first win. By the geometric series

$$\sum_{i=0}^{n} 2^i = 2^{n+1} - 1 \,,$$

the payout of each round is always 1 (times the minimum bet). So, is there any problem about this simple success strategy?

To apply martingale theory, we consider a more rewarding, fair game with success probability 1/2 in each round. Let $(X_n)_{n\in\mathbb{N}}$ be i.i.d. random variables with

$$\mathbb{P}(X_n = 1) = \mathbb{P}(X_n = -1) = 1/2 \,.$$

If we start the game with 0 capital, and allow for negative capital during the game, we can model the gains of the strategy after n rounds by

$$Z_n = \sum_{i=1}^{n} 2^{i-1} X_i \,,$$

which is a martingale. The strategy is now to stop at time T, when $Z_T = +1$, with

$$T = \inf\{n \in \mathbb{N} : Z_n = +1\} \,,$$

and we easily conclude that $\mathbb{P}(T > k) = 2^{-k}$, such that $\mathbb{P}(T < \infty) = 1$. Apparently, it holds that

$$\mathbb{E}[Z_T] = 1 \neq \mathbb{E}[Z_1] = 0 \,.$$

This already points at some problem of the strategy, since dominated convergence cannot be applied to yield such an identity. The problem is, that the expected loss up to time T is unbounded. In fact, you require infinite capital or credit limit to make the strategy work. It holds that

$$\mathbb{E}[-Z_{T-1}] = \mathbb{E}\Big[\sum_{k=1}^{T-1} 2^{k-1}\Big] = \mathbb{E}\Big[\sum_{k=1}^{\infty} 2^{k-1}\mathbf{1}_{\{T>k\}}\Big]$$
$$= \sum_{k=1}^{\infty} 2^{k-1}\,\mathbb{P}\big(T>k\big) = \sum_{k=1}^{\infty}\frac{1}{2} = +\infty\,.$$

•

2.4 Martingale convergence theorems

For martingales limit results in asymptotic stochastics can be proved which are more general than the standard law of large numbers or the standard central limit theorem. Here, we shall establish two of the most important results. We begin with a useful inequality.

Upcrossing inequality

Consider for some real-valued stochastic process $(X_n)_{n\geq 1}$ the number of upcrossings of an interval $[a,b]$. For given real numbers a,b with $a<b$, we consider the stopping times

$$\tau_1 = \inf\{t\in\mathbb{N} : X_t \leq a\}\,,$$

and for $k=1,2,\ldots$, let

$$\sigma_k = \inf\{t>\tau_k : X_t \geq b\} \;\text{ and }\; \tau_{k+1} = \inf\{t>\sigma_k : X_t \leq a\}\,,$$

with the convention that the infimum over the empty set is ∞. Denote by

$$U_X(a,b,N) = \max\{k : \sigma_k \leq N\} = \sum_{k=1}^{\infty}\mathbf{1}_{\{\sigma_k\leq N\}}\,,$$

the number of upcrossings of the interval $[a,b]$ by (X_n), up to time N. We set $\max\emptyset$ and as well an empty sum $=0$. Denote by

$$U_X(a,b) = \lim_{N\to\infty} U_X(a,b,N)$$

the (total) number of upcrossings of the interval $[a,b]$ by (X_n). Apparently, $U_X(a,b,N)$ and $U_X(a,b)$ are random variables.

These upcrossings are of interest to characterize the convergence of (X_n) due to the following relation.

Lemma 2.35. *Let $(x_n)_{n\geq 1}$ be a sequence of real numbers. The sequence (x_n) is convergent with a limit in $\bar{\mathbb{R}}$, if for all intervals $[a,b]$, with $a,b\in\mathbb{Q}$, it holds that $U_x(a,b)<\infty$.*

Proof. Consider

$$l = \liminf_{n\to\infty} x_n \leq \limsup_{n\to\infty} x_n = r, \qquad l,r\in\mathbb{R}\,.$$

If $l<r$, for intervals $[a,b]$, with $l<a<b<r$, the number of upcrossings is $U_x(a,b)=\infty$. If $l=r$, and $a<b$, we can choose for $\varepsilon=(b-a)/3$ an integer n_0, such that $|x_n-l|<\varepsilon$, $n\geq n_0$. Then, after n_0 there is no further upcrossing of $[a,b]$, such that $U_x(a,b)<\infty$.

■

The expected number of upcrossings up to time N by a submartingale can be bounded by the following inequality.

Theorem 2.36 (*Doob's upcrossing inequality*)**.** *Let (X_n) be a $(\mathcal{F}_n)$-submartingale and $N \in \mathbb{N}$. Then, for $a < b$, it holds true that*

$$\mathbb{E}[U_X(a,b,N)] \leq \frac{\mathbb{E}[(X_N - a)^+] - \mathbb{E}[(X_1 - a)^+]}{b-a} \leq \frac{\mathbb{E}[(X_N - a)^+]}{b-a}.$$

For the proof, we begin with a lemma.

Lemma 2.37. *The processes (H_m) and (K_m) defined by*

$$H_m = \sum_{j=1}^{\infty} \mathbf{1}_{\{\tau_j < m \leq \sigma_j\}}, \qquad K_m = 1 - H_m, \qquad m \geq 2, \tag{2.11}$$

are $(\mathcal{F}_n)$-predictable (and bounded), and therefore it holds true that

$$\mathbb{E}\left[\sum_{k=2}^{N} K_k \left(X_k - X_{k-1}\right)\right] \geq 0. \tag{2.12}$$

H_m can be interpreted as a strategy to buy (=1), if the price goes below a, and sell (=0), if the price exceeds b. Thus, K_m can be interpreted as a strategy to buy if the price exceeds b and sell if the price goes below a. If you stop the strategy at some fix time, the strategy (K_n) results for a submartingale in a non-negative expected profit.

Proof. Since

$$\{\tau_j < m \leq \sigma_j\} = \{\tau_j \leq m-1\} \cap \{\sigma_j \leq m-1\}^{\complement} \in \mathcal{F}_{m-1},$$

the claim follows by Corollary 2.18. ∎

Proof of Theorem 2.36. We pass to $(a + (X_n - a)^+)_{n\in\mathbb{N}} = (Y_n)_{n\in\mathbb{N}}$. (Y_n) is as well a submartingale and has the same upcrossings of the interval $[a,b]$ as (X_n), $U_Y(a,b,N) = U_X(a,b,N)$, and also the same stopping times σ_j, τ_j.

The claim is implied by the following inequality:

$$(b-a)U_Y(a,b,N) \leq \sum_{k=2}^{N} H_k\left(Y_k - Y_{k-1}\right). \tag{2.13}$$

This is intuitively clear: Up to time N, there are $U_Y(a,b,N)$ upcrossings of $[a,b]$, and by the weights H_k we measure these upcrossings. The last and not completed upcrossing yields a non-negative term, since we stop X_k at a. Since

$$Y_N - Y_1 = \sum_{k=2}^{N}(Y_k - Y_{k-1}) = \sum_{k=2}^{N} H_k\left(Y_k - Y_{k-1}\right) + \sum_{k=2}^{N} K_k\left(Y_k - Y_{k-1}\right)$$

we obtain with (2.12) by (2.13) that

$$\begin{aligned}(b-a)\,\mathbb{E}\left[U_X(a,b,N)\right] &\leq \mathbb{E}\left[\sum_{k=2}^{N} H_k\left(Y_k - Y_{k-1}\right)\right] \\ &\leq \mathbb{E}\left[\sum_{k=2}^{N} H_k\left(Y_k - Y_{k-1}\right)\right] + \mathbb{E}\left[\sum_{k=2}^{N} K_k\left(Y_k - Y_{k-1}\right)\right] \\ &= \mathbb{E}[Y_N] - \mathbb{E}[Y_1],\end{aligned}$$

what yields the claim.

It remains to establish (2.13). For some ω with $\sigma_j(\omega) = \sigma_j < \infty$, we have that $Y_{\tau_j} = a$, and $Y_{\sigma_j} \geq b$. This yields that

$$(b-a) \leq Y_{\sigma_j} - Y_{\tau_j} = \sum_{k=\tau_j+1}^{\sigma_j} H_k\big(Y_k - Y_{k-1}\big)\,, \tag{2.14}$$

and $H_k = 0$, $\sigma_{j-1}+1 \leq k \leq \tau_j$.

By definition $U_Y(a,b,N) = \max\{k : \sigma_k \leq N\}$, and thus by (2.14) for $U_Y(a,b,N) \geq 1$:

$$\begin{aligned}
(b-a)U_Y(a,b,N) &\leq \sum_{j=1}^{U_Y(a,b,N)} \sum_{k=\tau_j+1}^{\sigma_j} H_k\big(Y_k - Y_{k-1}\big) \\
&= \sum_{k=2}^{\sigma_{U_Y(a,b,N)}} H_k\big(Y_k - Y_{k-1}\big) \\
&= \sum_{k=2}^{\tau_{U_Y(a,b,N)+1}\wedge N} H_k\big(Y_k - Y_{k-1}\big)\,.
\end{aligned}$$

Since $Y_n \geq a$, and $Y_{\tau_j} = a$, for $\tau_j < \infty$, we obtain for $\tau_{U_Y(a,b,N)+1} < N (< \sigma_{U_Y(a,b,N)+1})$ (the last, not completed upcrossing), that

$$\sum_{k=\tau_{U_Y(a,b,N)+1}+1}^{N} H_k\big(Y_k - Y_{k-1}\big) \geq 0\,,$$

and hence (2.13). If $U_Y(a,b,N) = 0$, such that $\sigma_1 > N$, if $\tau_1 > N$, all summands are $= 0$. Else, the sum is $Y_N - Y_{\tau_1} = Y_N - a \geq 0$, and we obtain (2.13). ∎

Martingale convergence for L_1-bounded submartingales

Theorem 2.38. *Let $(X_n)_{n\in\mathbb{N}}$ be a $(\mathcal{F}_n)$-submartingale with $\sup_{n\in\mathbb{N}} \mathbb{E}\left[|X_n|\right] < \infty$. There exists a $\mathcal{F}_\infty$-measurable, integrable random variable X_∞, such that*

$$X_\infty = \lim_{n\to\infty} X_n\,, \quad \textit{almost surely.}$$

Moreover, it holds true that

$$\mathbb{E}\left[|X_\infty|\right] \leq \liminf_{n\to\infty} \mathbb{E}\left[|X_n|\right]\,.$$

For a $(\mathcal{F}_n)$-submartingale, $n \mapsto \mathbb{E}[X_n]$ is monotone increasing. Hence, it holds that

$$\mathbb{E}\left[|X_n|\right] = 2\,\mathbb{E}\left[X_n^+\right] - \mathbb{E}[X_n] \leq 2\,\mathbb{E}\left[X_n^+\right] - \mathbb{E}[X_1]\,.$$

Furthermore, it holds that

$$\mathbb{E}\left[|X_n|\right] \geq \mathbb{E}\left[X_n^+\right]\,.$$

Therefore, the condition $\sup_{n\in\mathbb{N}} \mathbb{E}\left[X_n^+\right] < \infty$ is equivalent to $\sup_{n\in\mathbb{N}} \mathbb{E}\left[|X_n|\right] < \infty$.

Proof of Theorem 2.38. Let $a < b$. It holds that $(X_n - a)^+ \leq X_n^+ + |a|$. By Doob's upcrossing inequality we obtain with monotone convergence for the total number of upcrossings $U_X(a,b)$ of $[a,b]$:

$$\mathbb{E}[U_X(a,b)] = \lim_{N\to\infty} \mathbb{E}[U_X(a,b,N)] \leq \sup_{N\in\mathbb{N}} \frac{\mathbb{E}[(X_N - a)^+]}{b-a} < \infty\,.$$

Therefore, $U_X(a,b) < \infty$ holds almost surely for all $a < b$, and thus for all $a < b,\ a,b \in \mathbb{Q}$. By Lemma 2.35 (X_n) converges almost surely to some random variable X_∞. With Fatou's lemma we obtain that

$$\mathbb{E}\left[|X_\infty|\right] \leq \liminf_{n\to\infty} \mathbb{E}\left[|X_n|\right] \leq \sup_n \mathbb{E}\left[|X_n|\right] < \infty .$$

In particular, X_∞ is integrable and almost surely finite. Then, a real-valued, $\mathcal{F}_\infty$-measurable version of the limit is given by

$$X_\infty(\omega) := \begin{cases} \lim_{n\to\infty} X_n(\omega) & \text{if the limit exists in } \mathbb{R} . \\ 0 & \text{else} \end{cases}$$

The event $A = \{\omega : \lim_{n\to\infty} X_n(\omega) \text{ exists in } \mathbb{R}\}$ is in $\sigma(X_n; n \geq 1) \subseteq \mathcal{F}_\infty$, and $X_n \mathbf{1}_A \to X_\infty$ pointwise, what implies the $\mathcal{F}_\infty$-measurability of X_∞. ∎

Remark. If (X_n) is a submartingale with $X_n \leq C$ a.s. for some $C > 0$ and all $n \geq 1$, then it holds that $\sup_{n\in\mathbb{N}} \mathbb{E}\left[X_n^+\right] < \infty$. Therefore, (X_n) converges. This applies in particular to non-negative or non-positive martingales ($(-X_n)$ is as well a martingale).

Examples 2.39.

- In the situation of Theorem 2.38, we do in general not have L_1-convergence of (X_n) to X_∞. For a counterexample, consider the simple symmetric random walk $X_n = \xi_1 + \ldots + \xi_n$, and its in $a \in \mathbb{N}$ stopped version $(X_{T_a \wedge n})$. Since $X_{T_a\wedge n} \leq a$, we have a martingale with an upper bound, such that almost sure convergence holds true. Since $\mathbb{P}(T_a < \infty)$, we have $X_{T_a\wedge n} \to a$ a.s., but since $0 = \mathbb{E}\left[X_{T_a\wedge n}\right]$, convergence in L_1 does not hold true.
- *Pólya's urn scheme.* Consider a urn with r red and g green balls. We draw a ball, put it back and add c additional balls with the colour of the drawn one. We iterate this procedure. After n draws there are $r+g+n\cdot c$ balls in the urn, among them $r, r+c, \ldots$ or $r+n\cdot c$ red balls. Denote by X_n the fraction of red balls after n draws. We can define X_n formally as follows: Let $(U_n)_{n\geq 1}$ be a sequence of independent, $U(0,1)$ uniformly distributed random variables. Set $X_0 = r/(r+g)$, and

  $$X_{n+1} = \mathbf{1}\big(U_{n+1} \leq X_n\big)\left(\frac{(r+g+nc)X_n + c}{r+g+(n+1)c}\right) + \mathbf{1}\big(U_{n+1} > X_n\big)\left(\frac{(r+g+nc)X_n}{r+g+(n+1)c}\right), \qquad n \geq 0 .$$

  Then, we have that

  $$\begin{aligned}\mathbb{E}\left[X_n | X_{n-1}, \ldots, X_1\right] &= X_{n-1}\left(\frac{(r+g+nc)X_{n-1}+c}{r+g+(n+1)c}\right) + (1-X_{n-1})\left(\frac{(r+g+nc)X_{n-1}}{r+g+(n+1)c}\right) \\ &= X_{n-1} ,\end{aligned}$$

  such that (X_n) is a martingale with $0 \leq X_n \leq 1$, which converges almost surely to some X_∞. •

Convergence of L_2-martingales

Theorem 2.40. *Let $(X_n)_{n\in\mathbb{N}}$ be a $(\mathcal{F}_n)$-martingale with*

$$\sup_{n\in\mathbb{N}} \mathbb{E}[X_n^2] < \infty.$$

Then, there exists a $\mathcal{F}_\infty$-measurable random variable X_∞ with $\mathbb{E}[X_\infty^2] < \infty$, such that

$$X_\infty = \lim_{n\to\infty} X_n, \quad \textit{almost surely and in } L_2 .$$

Proof. Since $\mathbb{E}\left[|X_n|\right] \leq 1 + \mathbb{E}[X_n^2]$, our first martingale convergence theorem implies the almost sure convergence to X_∞. We show that (X_n) converges also in L_2. The L_2 limit is necessarily the same, and the result follows. Consider the square integrable martingale differences $\xi_1 = X_1$, and $\xi_n = X_n - X_{n-1}$, for $n \geq 2$, for which $\mathbb{E}[\xi_n|\mathcal{F}_{n-1}] = 0$, for $n \geq 2$ holds. For $1 \leq n < m$, we have that

$$\mathbb{E}[\xi_n\xi_m] = \mathbb{E}\left[\mathbb{E}[\xi_n\xi_m|\mathcal{F}_n]\right] = \mathbb{E}\left[\xi_n\, \mathbb{E}[\xi_m|\mathcal{F}_n]\right] = 0\,.$$

We conclude that martingale increments are uncorrelated, thus orthogonal in L_2. We deduce that

$$\mathbb{E}[X_n^2] = \sum_{k=1}^{n} \mathbb{E}[\xi_k^2]\,.$$

We obtain that

$$\sup_{n\in\mathbb{N}} \mathbb{E}[X_n^2] = \lim_{n\to\infty} \mathbb{E}[X_n^2] = \sum_{n=1}^{\infty} \mathbb{E}[\xi_n^2] < \infty\,.$$

For $1 \leq n < m$, we obtain that

$$\mathbb{E}\left[(X_n - X_m)^2\right] = \sum_{k=n+1}^{m} \mathbb{E}[\xi_k^2] \to 0, \qquad n,m \to \infty,$$

which is the Cauchy property for (X_n). Since L_2 is complete (Satz von Riesz-Fischer, see Satz 7.18 in Klenke [2008]), we conclude the claim. ■

3 Brownian motion

3.1 Definition and first properties

Brownian motion is named after the botanist Robert Brown, who discovered in the first half of the 19th century the thermal chaotic movements of microscopic particles. As a stochastic process it was first used by the French mathematician Louis Bachelier at the beginning of the 20th century to model stock prices in efficient markets. The first formal construction, and thus the mathematical proof of existence, was by Norbert Wiener (1923). For this reason it is as well called the Wiener process in stochastics. This process is the starting and central point of stochastic analysis and the theory of Markov processes and takes a universal role in continuous-time martingale theory. In fact, the theory of stochastic processes is mainly build upon the study of Brownian motion and its properties. Most classes of stochastic processes as Gaussian processes, Markov processes, self-similar processes and Lévy processes are defined as processes having one of the fundamental properties of Brownian motion. Moreover, Brownian motion is still crucial for mathematical models of prices of financial assets and thermodynamics.

Motivation and Definition

Let $X_1, X_2, \ldots$ be i.i.d. real-valued random variables with $\mathbb{E}[X_i] = 0$, $\mathrm{Var}(X_i) = 1$. In Chapter 2 we have considered an associated random walk $S_0 = 0$,

$$S_n = X_1 + \ldots + X_n, \quad n \geq 1.$$

Interpolating between the discrete time points yields a continuous-time stochastic process

$$S(t) = S_{\lfloor t \rfloor} + \big(t - \lfloor t \rfloor\big) X_{\lfloor t \rfloor + 1}, \qquad t \geq 0.$$

Now, we let the time distances between the interpolated points tend to 0, by rescaling

$$Z_n(t) = S(nt)/\sqrt{n}, \qquad t \geq 0. \tag{3.1}$$

In the limit, as $n \to \infty$, this sequences converges in distribution to a Brownian motion (what is not easy to prove) which we define below. For a motivation, observe first that for fix $\varepsilon > 0$, $T \in \mathbb{N}$, it holds as $n \to \infty$ that

$$\begin{aligned}
\mathbb{P}\Big(\sup_{0 \leq t \leq T} \big|Z_n(t) - S_{\lfloor nt \rfloor}/\sqrt{n}\big| \geq \varepsilon\Big) &= \mathbb{P}\Big(\max_{1 \leq k \leq nT} |X_k| \geq \sqrt{n}\,\varepsilon\Big) \\
&\leq nT\, \mathbb{P}\big(|X_1| \geq \sqrt{n}\,\varepsilon\big) \\
&\leq nT \frac{1}{n\varepsilon^2} \mathbb{E}\Big[|X_1|^2 1_{\{|X_1| \geq \sqrt{n}\,\varepsilon\}}\Big] \to 0\,.
\end{aligned}$$

Therefore, the differences between $Z_n(t)$ and $S_{\lfloor nt \rfloor}/\sqrt{n}$ are asymptotically negligible. Furthermore, we have that

$$S_{\lfloor nt \rfloor}/\sqrt{n} = \Big(\frac{\lfloor nt \rfloor}{n}\Big)^{1/2} S_{\lfloor nt \rfloor}/\sqrt{\lfloor nt \rfloor} \xrightarrow{d} \mathcal{N}(0,t)\,,$$

and for $0 \leq s < t$, the increments $\big(S_{\lfloor nt \rfloor} - S_{\lfloor ns \rfloor}\big)/\sqrt{n}$ and $S_{\lfloor ns \rfloor}/\sqrt{n}$ are independent. If the continuity of $Z_n(t)$ is preserved, the following form of the limiting process appears to be plausible.

Definition 3.1. *A stochastic process $(B_t)_{t \geq 0}$ is called a* ***(standard) Brownian motion****, or* ***Wiener process****, if it satisfies the following properties. 1. $(B_t)_{t \geq 0}$ has almost surely continuous paths.*

2. $(B_t)_{t\geq 0}$ has independent increments: For $n \geq 1$, and $0 \leq t_1 < \ldots < t_n$, the random variables

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

are independent.
3. For all $t \geq 0$, $B_t \sim \mathcal{N}(0,t)$ is normally distributed, in particular $B_0 = 0$.

As for the Poisson process the property 2. implies the stronger third property
3.' For all $0 \leq s < t$, $B_t - B_s \sim \mathcal{N}(0, t-s)$ and $B_0 = 0$ a.s.
Since $B_t = B_s + (B_t - B_s)$, and $B_s, (B_t - B_s)$ are independent, the characteristic functions satisfy

$$\varphi_{B_t - B_s}(x) = \frac{\varphi_{B_t}(x)}{\varphi_{B_s}(x)} = \frac{e^{-tx^2/2}}{e^{-sx^2/2}} = e^{-(t-s)x^2/2}.$$

The notion of filtration, martingales and stopping times are extended from discrete to continuous time analogously. We shall work with these concepts in the next chapter. The Wiener process is a stochastic process in continuous time $(B_t, t \geq 0)$ and is defined on some complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and $(\mathcal{F}_t)$-adapted w.r.t. a filtration $(\mathcal{F}_t)_{t\geq 0}$, where we want $(\mathcal{F}_t)_{t\geq 0}$ to satisfy the usual conditions:

- $\mathcal{F}_s = \bigcap_{t>s} \mathcal{F}_t$, for all $s \geq 0$ (right-continuity);
- all $A \in \mathcal{F}$ with $\mathbb{P}(A) = 0$ are contained in $\mathcal{F}_0$ (completeness).

For a $(\mathcal{F}_t)$-Brownian motion $(B_t)_{t\geq 0}$ we even have that $(B_t - B_s)$ is independent of $\mathcal{F}_s$, for all $0 \leq s \leq t$ (later proved in Theorem 3.11). Denote by $\mathcal{B}_+$ the Borel σ-field on $[0,\infty)$ and $\mathcal{B}$ the Borel σ-field over $\mathbb{R}$. A continuous-time real-valued stochastic process $(X_t)_{t\geq 0}$ is measurable, if the mapping $(t,\omega) \mapsto X(t,\omega)$ is $((\mathcal{B}_+ \otimes \mathcal{F}) - \mathcal{B})$-measurable. One can prove that having right-continuous paths is sufficient for measurability, see Proposition 1.13 in Chapter 1 of Karatzas and Shreve [1991].

Characterization as a Gaussian process

Theorem 3.2. *Let $(X_t)_{t\geq 0}$ be a stochastic process with almost surely continuous paths. $(X_t)_{t\geq 0}$ is a Brownian motion iff*
4. For $n \geq 1$, $0 \leq t_1 < \ldots < t_n$, it holds that

$$\left(X_{t_1}, X_{t_2}, \ldots, X_{t_n}\right)^\top \sim \mathcal{N}(0,\Sigma), \quad \Sigma_{j,k} = \mathrm{Cov}(X_{t_j}, X_{t_k}) = \min(t_j, t_k).$$

Proof. 2. + 3. $\Rightarrow$ 4.: By 2. and 3.

$$\left(X_{t_1}, X_{t_2} - X_{t_1}, \ldots, X_{t_n} - X_{t_{n-1}}\right)^\top \sim \mathcal{N}\big(0, \mathrm{diag}(t_1, t_2 - t_1, \ldots, t_n - t_{n-1})\big). \tag{3.2}$$

Since

$$\begin{pmatrix} X_{t_1} \\ X_{t_2} \\ \vdots \\ X_{t_n} \end{pmatrix} = A \begin{pmatrix} X_{t_1} \\ X_{t_2} - X_{t_1} \\ \vdots \\ X_{t_n} - X_{t_{n-1}} \end{pmatrix}, \qquad A = \begin{pmatrix} 1 & 0 & \ldots & 0 & 0 \\ 1 & 1 & 0 & \ldots & 0 \\ \vdots & \ldots & \vdots & \ldots & \vdots \\ 1 & 1 & 1 & \ldots & 1 \end{pmatrix},$$

it holds that

$$\left(X_{t_1}, X_{t_2}, \ldots, X_{t_n}\right)^\top \sim \mathcal{N}\big(0,\Sigma\big), \qquad \Sigma = A\,\mathrm{diag}(t_1, t_2 - t_1, \ldots, t_n - t_{n-1})A^\top.$$

Instead of computing Σ as a matrix product, we directly determine the covariance of X_s and X_t. For $s < t$, this yields

$$\mathrm{Cov}(X_s, X_t) = \mathrm{Cov}(X_s, X_t - X_s) + \mathrm{Cov}(X_s, X_s) = s.$$

4. $\Rightarrow$ 2. and 3.: We have to show (3.2). Since

$$\begin{pmatrix} X_{t_1} \\ X_{t_2}-X_{t_1} \\ \vdots \\ X_{t_n}-X_{t_{n-1}} \end{pmatrix} = A^{-1} \begin{pmatrix} X_{t_1} \\ X_{t_2} \\ \vdots \\ X_{t_n} \end{pmatrix}, \qquad A^{-1} = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 \\ -1 & 1 & 0 & \dots & 0 \\ \vdots & \dots & \vdots & \dots & \vdots \\ 0 & \dots & 0 & -1 & 1 \end{pmatrix},$$

it holds that

$$\left(X_{t_1}, X_{t_2}-X_{t_1}, \dots, X_{t_n}-X_{t_{n-1}}\right)^\top \sim \mathcal{N}\left(0, A^{-1}\Sigma A^{-\top}\right) = \mathcal{N}\left(0, \operatorname{diag}(t_1, t_2-t_1, \dots, t_n-t_{n-1})\right). \qquad \blacksquare$$

Remark. We will study Gaussian processes defined as stochastic processes $(X_t)_{t\geq 0}$, for which for all $n\geq 1$, $0\leq t_1 < \ldots < t_n$, it holds that $\left(X_{t_1}, X_{t_2}, \dots, X_{t_n}\right)^\top$ is multivariate normally distributed. This means that the finite-dimensional distributions are multivariate normal. In this case, the finite-dimensional distributions are determined by the expected value function $\mu_t = \mathbb{E}[X_t]$, and the covariance function $\gamma(s,t) = \operatorname{Cov}(X_s, X_t)$.
A Brownian motion is thus uniquely characterized as a Gaussian process with almost surely continuous paths, expected value function $\mu_t = 0$, and the covariance function $\gamma(s,t) = \min(s,t)$.

Transformation properties

Proposition 3.3. *Let $(B_t)_{t\geq 0}$ be a standard Brownian motion. The following processes are as well standard Brownian motions*
1. ***scaling:*** *$(B_t^{(1)})_{t\geq 0}$, with $B_t^{(1)} = a^{-1/2}B_{at}$, $a>0$,*
2. ***shifting:*** *$(B_t^{(2)})_{t\geq 0}$, with $B_t^{(2)} = B_{t+t_0} - B_{t_0}$, $t_0 > 0$,*
3. ***time inversion:*** *$(B_t^{(3)})_{t\geq 0}$, with $B_t^{(3)} = tB_{1/t}$, $B_0^{(3)} = 0$,*
4. ***reflection:*** *$(B_t^{(4)})_{t\geq 0}$, with $B_t^{(4)} = -B_t$.*

Proof. The processes $(B_t^{(j)})_{t\geq 0}, 1\leq j\leq 4$, are centered Gaussian processes where centered means that $\mathbb{E}[B_t^{(j)}] = 0$. We determine the covariance functions. For $0\leq s<t$:
$j=1$: $\operatorname{Cov}\left(B_s^{(1)}, B_t^{(1)}\right) = \frac{1}{a}\operatorname{Cov}\left(B_{as}, B_{at}\right) = \frac{1}{a}as = s$.
$j=3$: $\operatorname{Cov}\left(B_s^{(3)}, B_t^{(3)}\right) = st\operatorname{Cov}\left(B_{1/s}, B_{1/t}\right) = st/t = s$.
$j=2,4$: obvious. Since the processes $(B_t^{(j)})_{t\geq 0}$, $j=1,2,4$, have almost surely continuous paths we conclude that they are Brownian motions by Theorem 3.2. For $(B_t^{(3)})_{t\geq 0}$, we should check the almost sure continuity in 0:

$$\lim_{t\to 0} tB_{1/t} = 0\,, \text{ a.s.}$$

This is left as an exercise. $\blacksquare$

The reflected process at time $s\geq 0$ in B_s, denote it by (B_t^*), has up to time s the same path as (B_t), and afterwards has increments $-(B_t - B_s)$, and values $B_s - (B_t - B_s) = 2B_s - B_t$. The following neat result has some important implications.

Theorem 3.4. *Let $(B_t)_{t\geq 0}$ be a Brownian motion and $s\geq 0$. Then the reflected process (B_t^*) at time s in B_s with*

$$B_t^* = \begin{cases} B_t, & t\leq s, \\ 2B_s - B_t, & t>s, \end{cases}$$

is a Brownian motion. $\blacksquare$

Proof. It is clear that we have a centered Gaussian process with almost surely continuous paths. We consider the covariance function
1. If $0 \le t_1 < t_2 \le s$: obvious.
2. If $0 \le t_1 \le s < t_2$: $\mathrm{Cov}(B^*_{t_1}, B^*_{t_2}) = \mathrm{Cov}(B_{t_1}, 2B_s - B_{t_2}) = 2t_1 - t_1 = t_1$.
3. If $s \le t_1 < t_2$: $\mathrm{Cov}(B^*_{t_1}, B^*_{t_2}) = \mathrm{Cov}(2B_s - B_{t_1}, 2B_s - B_{t_2}) = 4s - 2s - 2s + t_1 = t_1$.

■

The theorem gives an extension of the reflection property from Proposition 3.3. This becomes really useful when we extend the result later to stopping times.

Path properties

For an interval $[a,b] \subset [0,\infty)$, consider partitions

$$P = \{a = t_0 < t_1 < \ldots < t_n = b\}$$

with mesh

$$\pi(P) = \max_{k=1,\ldots,n} (t_k - t_{k-1}).$$

For a Brownian motion (B_t) and $p > 0$, define the p-variation, or pth power variation:

$$V(P, p, (B_t)) = \sum_{k=1}^{n} |B_{t_k} - B_{t_{k-1}}|^p.$$

Proposition 3.5. *Let (P_n) be a sequence of partitions of the interval $[a,b]$ with meshs $\pi(P_n) \to 0$. Then it holds true that*

$$V(P_n, p, (B_t)) \xrightarrow{\mathbb{P}} \begin{cases} b-a, & p = 2, \\ \infty, & p < 2, \\ 0, & p > 2. \end{cases}$$

For $p = 2$, we have as well convergence in L_2.

Proof.

$$p = 2: \quad \mathbb{E}\left[V(P_n, 2, (B_t))\right] = \sum_{k=1}^{r_n} \mathbb{E}\left[(B_{t_k} - B_{t_{k-1}})^2\right] = \sum_{k=1}^{r_n} (t_k - t_{k-1}) = b - a,$$

and for the variance we obtain with the independence of the increments that

$$\begin{aligned} \mathrm{Var}\left(V(P_n, 2, (B_t))\right) &= \sum_{k=1}^{r_n} \mathrm{Var}\left((B_{t_k} - B_{t_{k-1}})^2\right) \\ &= 2\sum_{k=1}^{r_n} (t_k - t_{k-1})^2 \\ &\le 2\pi(P_n)(b-a) \to 0, \qquad n \to \infty. \end{aligned}$$

We have used that for $X \sim \mathcal{N}(0, \sigma^2)$, it holds that $\mathrm{Var}(X^2) = 2\sigma^4$.

$$p > 2: \quad V(P_n, p, (B_t)) \le \sum_{k=1}^{r_n} (B_{t_k} - B_{t_{k-1}})^2 \Big(\max_{1 \le k \le r_n} |B_{t_k} - B_{t_{k-1}}| \Big)^{p-2}. \tag{3.3}$$

(B_t) has almost surely continuous paths on $[a,b]$, and thus uniformly continuous paths on the compact interval. Since $\pi(P_n) \to 0$, we obtain that

$$\max_{1 \le k \le r_n} |B_{t_k} - B_{t_{k-1}}| \to 0 \quad \text{a.s.} \tag{3.4}$$

With the convergence in probability in the case $p=2$, the claim follows by (3.3).

$$p<2: \quad V(P_n,p,(B_t)) \geq \sum_{k=1}^{r_n} (B_{t_k}-B_{t_{k-1}})^2 \Big(\max_{1\leq k\leq r_n} |B_{t_k}-B_{t_{k-1}}|\Big)^{p-2}. \tag{3.5}$$

With (3.4) and the convergence in probability for the case $p=2$, the claim follows. ■

Corollary 3.6. It holds almost surely for all intervals $[a,b]$ that

$$\sup_{P \text{ Part. } [a,b]} V(P,1,(B_t)) = \infty.$$

Proof. If suffices to prove the claim for all intervals $[a,b]$, with $a,b \in \mathbb{Q}$. We consider some arbitrary interval.
Since for a sequence which converges in probability an almost sure convergent subsequence exists, there exists a sequence of partitions P_n, such that

$$V(P_n,2,(B_t)) \to (b-a) \quad \text{a.s.}$$

With (3.5) and (3.4), we obtain that

$$V(P_n,1,(B_t)) \to \infty \quad \text{a.s.}$$

what yields the result. ■

It is clear that for all monotone and for all Lipschitz continuous functions, the variation ($p=1$) considered in the corollary is finite. We conclude that Brownian motion has almost surely paths which are neither Lipschitz continuous nor monotone. This motivates that although the paths are continuous, they are quite irregular. This becomes more clear by the intriguing next result which has been proved by Wiener in 1931. Our simpler proof has been supposed by the famous mathematician Paul Erdös. [1]

Theorem 3.7. *Let $(B_t)_{t\geq 0}$ be a Brownian motion. The paths $t \mapsto B_t(\omega)$ are almost surely nowhere differentiable.* ■

Proof. For $n \in \mathbb{N}$, set

$$A_n = \left\{\omega \in \Omega : (B_t(\omega)) \text{ is nowhere differentiable on } [0,n)\right\}.$$

For $f : \mathbb{R}_+ \to \mathbb{R}$ a in $t_0 \in [0,n)$ differentiable function, there exist $\delta > 0$ and $L > 0$, such that for all t with $|t-t_0| < \delta$, it holds that

$$|f(t) - f(t_0)| \leq L|t-t_0|.$$

Then, there exists some $m \in \mathbb{N}$, such that for all $k \geq m$, and with $j = \lceil kt_0 \rceil = \lfloor kt_0 \rfloor + 1$, it holds that $j/k, (j+1)/k, (j+2)/k, (j+3)/k \in U_\delta(t_0) = \{t : |t-t_0| < \delta\}$. Note that $t_0 \leq j/k$, and $t_0 \geq (j-1)/k$. It follows that for $i \in \{j+1, j+2, j+3\}$:

$$\begin{aligned}\left|f\left(\frac{i}{k}\right) - f\left(\frac{i-1}{k}\right)\right| &\leq \left|f\left(\frac{i}{k}\right) - f(t_0)\right| + \left|f(t_0) - f\left(\frac{i-1}{k}\right)\right| \\ &\leq L\left(\frac{4}{k} + \frac{3}{k}\right) \leq L\frac{7}{k}.\end{aligned}$$

Since we have that

$$A_n^\complement \subseteq \bigcup_{L=1}^{\infty} \bigcup_{m=1}^{\infty} D_m^L,$$

[1] wiki: Paul Erdös

with

$$D_m^L = \bigcap_{k=m}^{\infty} \bigcup_{j=1}^{kn} \bigcap_{i=j+1}^{j+3} \left\{ \left| B_{\frac{i}{k}} - B_{\frac{i-1}{k}} \right| \leq L \frac{7}{k} \right\},$$

we conclude that the event A_n is measurable. We will show that $\mathbb{P}(D_m^L) = 0$, for all $m, L \in \mathbb{N}$, what implies that $\mathbb{P}(A_n^{\complement}) = 0$, such that $\mathbb{P}(A_n) = 1$.

We are left to prove that $\mathbb{P}(D_m^L) = 0$, for all $m, L \in \mathbb{N}$. For all $k \geq m$, it holds that

$$\begin{aligned}
\mathbb{P}(D_m^L) &\leq \mathbb{P}\left(\bigcup_{j=1}^{kn} \bigcap_{i=j+1}^{j+3} \left\{ \left| B_{\frac{i}{k}} - B_{\frac{i-1}{k}} \right| \leq L \frac{7}{k} \right\} \right) \\
&\leq \sum_{j=1}^{kn} \mathbb{P}\left(\bigcap_{i=j+1}^{j+3} \left\{ \left| B_{\frac{i}{k}} - B_{\frac{i-1}{k}} \right| \leq L \frac{7}{k} \right\} \right) \\
&= \sum_{j=1}^{kn} \prod_{i=j+1}^{j+3} \mathbb{P}\left(\left| B_{\frac{i}{k}} - B_{\frac{i-1}{k}} \right| \leq L \frac{7}{k} \right) \\
&= kn\, \mathbb{P}\left(\left| B_{\frac{1}{k}} \right| \leq L \frac{7}{k} \right)^3 .
\end{aligned}$$

We have used subadditivity for the second inequality and the independence and the stationarity of the increments for the equalities. From Proposition 3.3, we obtain that $B_{kt}/\sqrt{k}$ is another Brownian motion, such that

$$\begin{aligned}
\mathbb{P}\left(\left| B_{\frac{1}{k}} \right| \leq L \frac{7}{k} \right) &= \mathbb{P}\left(\left| B_1 \right| \leq L \frac{7}{\sqrt{k}} \right) \\
&= \frac{1}{\sqrt{2\pi}} \int_{\frac{-7L}{\sqrt{k}}}^{\frac{7L}{\sqrt{k}}} \exp(-x^2/2)\, dx \\
&\leq \frac{c}{\sqrt{k}},
\end{aligned}$$

with some constant c. Therefore, for all $k \geq m$, we obtain that

$$\mathbb{P}(D_m^L) \leq kn \left(\frac{c}{\sqrt{k}} \right)^3 .$$

Since this bound tends to zero as $k \to \infty$, we finish the proof. ■

3.2 Construction of Brownian motion

In this section, we establish a construction of Brownian motion. Such a construction proves that a process with the properties from Definition 3.1 exists. Today several construction methods are known, see Chapters 3 and 4 of Schilling and Partzsch [2012]. We follow the interpolation construction by Paul Lévy. We construct a sequence of polygons which converges almost surely uniformly on compact sets to a limit process which we identify to be a Brownian motion.

Theorem 3.8. *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and $(\xi_n)_{n\geq 1}$ a sequence of i.i.d. $\mathcal{N}(0,1)$ distributed random variables on this space, $\xi_n : \Omega \to \mathbb{R}$. Then, there exists a Brownian motion on $(\Omega, \mathcal{A}, \mathbb{P})$.* ▪

For instance, one can show that on the probability space $\big((0,1], \mathcal{B}(0,1], \lambda_{(0,1]}\big)$, with λ the Lebesgue measure, such a sequence of i.i.d. $\mathcal{N}(0,1)$ distributed random variables exists.

Motivation for the construction

A first idea could be to use the polygons $Z_n(t)$ from (3.1), e.g. with $\mathcal{N}(0,1)$ distributed random variables (X_k), for the construction. As mentioned before, they converge to a Brownian motion, however, this convergence is only in distribution. Pointwise convergence does not hold true, since, for instance, we have that $\limsup_n Z_n(1) = \limsup_n S_n/\sqrt{n} = \infty$, and $\liminf_n Z_n(1) = -\infty$.

The idea of the following construction is the observation that, if (B_t) is a Brownian motion, and $0 < t-a < t < t+a$, we obtain for the conditional distribution that

$$B_t|(B_{t-a}, B_{t+a}) \sim \mathcal{N}\big((B_{t-a}+B_{t+a})/2, a/2\big)\,.$$

If we have a random variable $\xi \sim \mathcal{N}(0,1)$, which is independent of (B_{t-a}, B_{t+a}), then we have that

$$\big(B_{t-a}, B_t, B_{t+a}\big) \stackrel{d}{=} \big(B_{t-a}, (B_{t-a}+B_{t+a})/2 + (a/2)^{1/2}\,\xi, B_{t+a}\big)\,.$$

We use this approach for an iterative construction by interpolation. For a small, the increment $(a/2)^{1/2}\,\xi$ gets small, and we obtain almost surely uniform convergence.

Proof of Theorem 3.8. Set

$$\mathbb{D}_n = \Big\{\frac{k}{2^n}, k = 0,1,2,\dots\Big\},$$

with $\mathbb{D}_0 = \mathbb{N}_0$. Define $\mathbb{D} = \bigcup_n \mathbb{D}_n$, which is a countable set. Let $(Z_t)_{t\in\mathbb{D}}$ be independent, $\mathcal{N}(0,1)$ distributed random variables which exist on $(\Omega, \mathcal{A}, \mathbb{P})$.
We first construct (B_t), for $t \in \mathbb{D}$, as follows. Set $B_0 = 0$, $B_k = Z_1 + \dots + Z_k$, $k \in \mathbb{N}$, and then inductively for $t \in \mathbb{D}_n \setminus \mathbb{D}_{n-1}$,

$$B_t = \frac{1}{2}\big(B_{t+2^{-n}} + B_{t-2^{-n}}\big) + \frac{1}{2^{(n+1)/2}} Z_t\,, \tag{3.6}$$

where $t \pm 2^{-n} \in \mathbb{D}_{n-1}$. Set

$$X_t^{(n)} = \begin{cases} B_t, & t \in \mathbb{D}_n, \\ \text{linear} & \text{in between.} \end{cases}$$

We shall show that $X_t^{(n)}$ converges on any compact set almost surely uniformly and that the limit is a Brownian motion.

Step 1: $(X_t^{(n)})_{t\geq 0}$ has on $\mathbb{D}_n$ the same distribution as a Brownian motion.

Lemma 3.9. *For all $n \geq 0$, the increments*

$$\big\{\big(B_{k2^{-n}} - B_{(k-1)2^{-n}}\big), k = 1,2,\dots\big\}$$

are independent, $\mathcal{N}(0, 2^{-n})$ distributed and $\sigma\big(Z_t, t \in \mathbb{D}_n\big)$-measurable random variables.

Proof. Induction over n: For $n = 0$, we have that $B_k = Z_1 + \dots + Z_k$ and the claim is obvious.
Induction step from $n-1$ to n:
The random variables B_t are linear transformations of the (Z_t), $t \in \mathbb{D}_n$, and thus jointly normally distributed and $\sigma\big(Z_t, t \in \mathbb{D}_n\big)$-measurable. Therefore, it suffices to prove pairwise independence of the increments and to determine their expectations and variances.
Consider an increment $B_{k/2^n} - B_{(k-1)/2^n}$.
Case 1: $k = 2l$ is even, i.e. $k/2^n = l/2^{n-1} \in \mathbb{D}_{n-1}$. By (3.6), it holds that

$$\begin{aligned} B_{k/2^n} - B_{(k-1)/2^n} &= B_{k/2^n} - \Big(\frac{1}{2}\big(B_{k/2^n} + B_{(k-2)/2^n}\big) + \frac{1}{2^{(n+1)/2}} Z_{(k-1)/2^n}\Big) \\ &= \frac{1}{2}\big(B_{l/2^{n-1}} - B_{(l-1)/2^{n-1}}\big) - \frac{1}{2^{(n+1)/2}} Z_{(2l-1)/2^n}\,. \end{aligned}$$

Case 2: $k = 2l - 1$ is odd, i.e. $(k-1)/2^n = (l-1)/2^{n-1} \in \mathbb{D}_{n-1}$. By (3.6), it holds that

$$\begin{aligned} B_{k/2^n} - B_{(k-1)/2^n} &= \frac{1}{2}\big(B_{(k+1)/2^n} + B_{(k-1)/2^n}\big) + \frac{1}{2^{(n+1)/2}} Z_{k/2^n} - B_{(k-1)/2^n} \\ &= \frac{1}{2}\big(B_{l/2^{n-1}} - B_{(l-1)/2^{n-1}}\big) + \frac{1}{2^{(n+1)/2}} Z_{(2l-1)/2^n} \,. \end{aligned}$$

By the induction hypothesis the increments for $t \in \mathbb{D}_{n-1}$ are independent and $\sigma\big(Z_t\,,\, t \in \mathbb{D}_{n-1}\big)$-measurable. Thereby, with the independence of the Z_t and by the above identity we conclude the independence of two increments $B_{k/2^n} - B_{(k-1)/2^n}$ and $B_{k'/2^n} - B_{(k'-1)/2^n}$, for $k' < k$, unless k' is odd and $k = k'+1$, if both increments belong to the same $\mathbb{D}_{n-1}$-interval. In this case, we directly prove the independence showing that the increments are uncorrelated. For $k' = 2l-1$, and $k = 2l$, we obtain by the induction hypothesis that

$$\begin{aligned} & \operatorname{Cov}\Big(B_{k/2^n} - B_{(k-1)/2^n}, B_{k'/2^n} - B_{(k'-1)/2^n}\Big) \\ = \; & \operatorname{Cov}\Big(\frac{1}{2}\big(B_{l/2^{n-1}} - B_{(l-1)/2^{n-1}}\big) - \frac{1}{2^{(n+1)/2}} Z_{(2l-1)/2^n}, \\ & \qquad \frac{1}{2}\big(B_{l/2^{n-1}} - B_{(l-1)/2^{n-1}}\big) + \frac{1}{2^{(n+1)/2}} Z_{(2l-1)/2^n}\Big) \\ = \; & \frac{1}{4}\operatorname{Var}\Big(B_{l/2^{n-1}} - B_{(l-1)/2^{n-1}}\Big) - \frac{1}{2^{n+1}} = \frac{1}{4}\frac{1}{2^{n-1}} - \frac{1}{2^{n+1}} = 0\,. \end{aligned}$$

Since the (Z_t) have expectations zero, the same applies to the (B_t). For the variance we obtain by induction in both cases that

$$\begin{aligned} \operatorname{Var}\Big(B_{k/2^n} - B_{(k-1)/2^n}\Big) &= \frac{1}{4}\operatorname{Var}\Big(B_{l/2^{n-1}} - B_{(l-1)/2^{n-1}}\Big) + \frac{1}{2^{n+1}} \\ &= \frac{1}{4}\frac{1}{2^{n-1}} + \frac{1}{2^{n+1}} = \frac{1}{2^n}\,. \end{aligned}$$

■

Step 2: $(X_t^{(n)})_{t\geq 0}$ converges almost surely uniformly on compact sets.

Let $F_0(t) = X_0(t)$ and

$$F_n(t) = \begin{cases} \frac{1}{2^{(n+1)/2}} Z_t & t \in \mathbb{D}_n \setminus \mathbb{D}_{n-1}, \\ 0 & t \in \mathbb{D}_{n-1} \\ \text{linear} & \text{in between.} \end{cases}$$

Lemma 3.10. *It holds true that*

$$X_t^{(n)} = \sum_{j=0}^{n} F_j(t), \qquad n \geq 0\,.$$

Proof. Let $Y_t^{(n)} = \sum_{j=0}^n F_j(t)$. Then, $Y_t^{(n)}$ is continuous and piecewise linear between the points $t \in \mathbb{D}_n$, since this holds for all $F_j(t)$, $j \leq n$. Therefore, $Y_t^{(n)}$ is the linear interpolation between the points $Y_t^{(n)}$, $t \in \mathbb{D}_n$, and we are left to show that

$$Y_t^{(n)} = B_t, \quad t \in \mathbb{D}_n\,. \tag{3.7}$$

Induction: This holds true for $n = 0$. For the induction set, observe that $Y_t^{(m)} = Y_t^{(n)}$, $t \in \mathbb{D}_n$, $m \geq n$, directly from the definition of $F_j(t)$. If (3.7) holds for $n-1$, and if $t \in \mathbb{D}_n \setminus \mathbb{D}_{n-1}$, we have for the linear interpolation that

$$Y_t^{(n-1)} = \frac{1}{2}\big(B_{t+2^{-n}} + B_{t-2^{-n}}\big)\,.$$

This implies for $t \in \mathbb{D}_n \setminus \mathbb{D}_{n-1}$:

$$Y_t^{(n)} = Y_t^{(n-1)} + F_n(t) = \frac{1}{2}\big(B_{t+2^{-n}} + B_{t-2^{-n}}\big) + \frac{1}{2^{(n+1)/2}} Z_t = B_t\,.$$

■

In order to prove the almost sure uniform convergence of $\left(X_t^{(n)}\right)$ on $[0,T]$, it suffices to establish absolute convergence of $F_j(t)$.

$$\sum_{n=0}^{\infty} \|F_n(t)\|_{[0,T]} < \infty \qquad \text{, a.s.}$$

We show that for $c > \sqrt{2\log(2)}$, it holds true that

$$\mathbb{P}\left(\|F_n(t)\|_{[0,T]} \geq c\sqrt{n}2^{-(n+1)/2} \; i.o. \right) = 0, \tag{3.8}$$

where i.o. means “infinitely often”. By the convergence of the series of $\sqrt{n}2^{-(n+1)/2}$, this yields the result. Since $Z_t \sim \mathcal{N}(0,1)$, we can use the known bound for Gaussian tails

$$\mathbb{P}(|Z_t| \geq y) \leq e^{-y^2/2}, \qquad y > \sqrt{2/\pi}.$$

We obtain that (with $\mathbb{D}_{-1} = \{0\}$)

$$\begin{aligned}
\sum_{n=0}^{\infty} \mathbb{P}\left(\|F_n(t)\|_{[0,T]} \geq c\sqrt{n}2^{-(n+1)/2}\right) &= \sum_{n=0}^{\infty} \mathbb{P}\left(\exists t \in \left(\mathbb{D}_n \setminus \mathbb{D}_{n-1}\right) \cap [0,T] : |Z_t| \geq c\sqrt{n}\right) \\
&\leq \sum_{n=0}^{\infty} \operatorname{card}\{\left(\mathbb{D}_n \setminus \mathbb{D}_{n-1}\right) \cap [0,T]\}\, e^{-c^2 n/2} \\
&\leq T \sum_{n=0}^{\infty} 2^n e^{-c^2 n/2} = T \sum_{n=0}^{\infty} e^{(2\log(2)-c^2)n/2} < \infty.
\end{aligned}$$

This yields (3.8) with the first Borel-Cantelli lemma.

Step 3: The limit $B_t = \sum_{n=0}^{\infty} F_n(t)$, $t \geq 0$, is a Brownian motion.

(B_t) is almost surely continuous by the almost sure uniform convergence on compact intervals, and on $\mathbb{D}$ coincides with the definition (3.6). It remains to show the distributional properties of Brownian motion. Let $0 \leq t_1 < \ldots < t_d$ be given. If $t_j \in \mathbb{D}$, which means in $\mathbb{D}_m$ for sufficiently large m, we obtain with Lemma 3.9 by Step 1 the desired distributional properties.
For general t_j, select $t_j^{(n)} \in \mathbb{D}$, with $t_j^{(n)} \to t_j$. Then, by the almost sure continuity of (B_t), we obtain that

$$\left(B_{t_1^{(n)}}, B_{t_2^{(n)}} - B_{t_1^{(n)}}, \ldots, B_{t_d^{(n)}} - B_{t_{d-1}^{(n)}}\right)^\top \to \left(B_{t_1}, B_{t_2} - B_{t_1}, \ldots, B_{t_d} - B_{t_{d-1}}\right)^\top \quad \text{a.s.},$$

and in particular convergence in distribution holds true. Since

$$\operatorname{diag}\left(t_1^{(n)}, t_2^{(n)} - t_1^{(n)}, \ldots, t_d^{(n)} - t_{d-1}^{(n)}\right) \to \operatorname{diag}\left(t_1, t_2 - t_1, \ldots, t_d - t_{d-1}\right),$$

considering the characteristic functions, we obtain that

$$\left(B_{t_1}, B_{t_2} - B_{t_1}, \ldots, B_{t_d} - B_{t_{d-1}}\right)^\top \sim \mathcal{N}\big(0, \operatorname{diag}\left(t_1, t_2 - t_1, \ldots, t_d - t_{d-1}\right)\big). \qquad \blacksquare$$

3.3 Strong Markov property

Markov property

Let $(B_t)_{t\geq 0}$ be a Brownian motion with continuous paths. We define

$$\mathcal{F}_t^B = \mathcal{F}_t = \sigma\{B_s:\ s \leq t\},$$

the filtration generated by (B_t).

Theorem 3.11 (Markov property)**.** *For $t_0 \geq 0$, the process $(B_s^{(1)})_{s\geq 0}$, with $B_s^{(1)} = B_{s+t_0} - B_{t_0}$, is a Brownian motion and independent of $\mathcal{F}_{t_0}^B$, i.e. $\sigma\{B_s^{(1)}, s\geq 0\}$ and $\mathcal{F}_{t_0}^B$ are independent.*

Proof. From the transformation properties, we already know that $(B_s^{(1)})_{s\geq 0}$ is a Brownian motion. Let

$$0 \leq s_1 < \ldots < s_d, \qquad 0 \leq t_1 < \ldots < t_q \leq t_0 .$$

Since $B_{s_j}^{(1)} - B_{s_{j-1}}^{(1)} = B_{s_j+t_0} - B_{s_{j-1}+t_0}$, and by the independent increments of Brownian motion, the random vectors

$$\left(B_{s_1}^{(1)}, B_{s_2}^{(1)} - B_{s_1}^{(1)}, \ldots, B_{s_d}^{(1)} - B_{s_{d-1}}^{(1)}\right)^\top \quad \text{and} \quad \left(B_{t_1}, B_{t_2} - B_{t_1}, \ldots, B_{t_q} - B_{t_{q-1}}\right)^\top$$

are independent, and by linear transformation

$$\left(B_{s_1}^{(1)}, B_{s_2}^{(1)}, \ldots, B_{s_d}^{(1)}\right)^\top \quad \text{and} \quad \left(B_{t_1}, B_{t_2}, \ldots, B_{t_q}\right)^\top$$

are as well independent. The families of sets

$$\left\{\{\left(B_{s_1}^{(1)}, B_{s_2}^{(1)}, \ldots, B_{s_d}^{(1)}\right)^\top \in A_d\},\ A_d \in \mathcal{B}^d,\ d \geq 1,\ 0 \leq s_1 < \ldots < s_d\right\}$$
$$\left\{\{\left(B_{t_1}, B_{t_2}, \ldots, B_{t_q}\right)^\top \in A_q\},\ A_q \in \mathcal{B}^q,\ q \geq 1,\ 0 \leq t_1 < \ldots < t_q \leq t_0\right\}$$

generate $\sigma\{B_s^{(1)}, s \geq 0\}$ and $\mathcal{F}_{t_0}^B$, and are closed w.r.t. intersections such that the generated σ-fields are independent. ∎

Let $\mathcal{F} = (\mathcal{F}_t)_{t\in[0,\infty)}$ be a filtration and $\mathcal{F}_{t+} = \bigcap_{s>t} \mathcal{F}_s$, for $t \geq 0$. Then, $(\mathcal{F}_{t+})_{t\in[0,\infty)}$ is a filtration.

Lemma 3.12. *For a filtration $\mathcal{F} = (\mathcal{F}_t)_{t\in[0,\infty)}$, $(\mathcal{F}_{t+})_{t\in[0,\infty)}$ is right-continuous.*

Proof.

$$\bigcap_{s>t} \mathcal{F}_{s+} = \bigcap_{s>t} \bigcap_{n>s} \mathcal{F}_n = \bigcap_{n>t} \mathcal{F}_n = \mathcal{F}_{t+} .$$

∎

Theorem 3.13. *For all $s \geq 0$, it holds true that $(B_{t+s} - B_s)_{t\geq 0}$ and $\mathcal{F}_{s+}^B$ are independent.*

Proof. We show that for all $m \geq 1$ and $0 \leq t_1 < \ldots < t_m$, that $(B_{s+t_1} - B_s, \ldots, B_{s+t_m} - B_s)$ and $\mathcal{F}_{s+}^B$ are independent.

Let $s_n \searrow s$ ($s_n > s_{n+1} > s$), by the continuity of the paths (in fact right-continuity of $s \mapsto B_{t+s} - B_s$ is sufficient), that

$$(B_{s_n+t_1} - B_{s_n}, \dots, B_{s_n+t_m} - B_{s_n}) \to (B_{s+t_1} - B_s, \dots, B_{s+t_m} - B_s)$$

pointwise. Therefore, $(B_{s+t_1} - B_s, \dots, B_{s+t_m} - B_s)$ is measurable w.r.t.

$$G = \sigma\Big\{\bigcup_{n\geq 1} G_n\Big\}, \qquad G_n = \sigma\{B_{s_n+t} - B_{s_n},\, t \geq 0\}\,.$$

We are left to show that G and $\mathcal{F}^B_{s+}$ are independent. Since $s_n > s_{n+1}$, we obtain that

$$B_{s_n+t} - B_{s_n} = B_{s_{n+1}+t+s_n-s_{n+1}} - B_{s_{n+1}} - \big(B_{s_{n+1}+s_n-s_{n+1}} - B_{s_{n+1}}\big)\,,$$

what shows that $G_n \subseteq G_{n+1}$. Therefore, $\bigcup_{n\geq 1} G_n$ generates G and is closed w.r.t. intersections and it suffices to verify independence of G_n and $\mathcal{F}^B_{s+}$ for any n. Since $\mathcal{F}^B_{s+} = \bigcap_{n\geq 1} \mathcal{F}^B_{s_n}$, the independence of G_n and $\mathcal{F}^B_{s_n}$ for any n is sufficient. This is readily implied by the Markov property from Theorem 3.11. ■

Corollary 3.14 (Blumenthal's zero-one law)**.** Let $(B_t)_{t\geq 0}$ be a standard Brownian motion. Then, $\mathcal{F}^B_{0+}$ is trivial, i.e. for all $F \in \mathcal{F}^B_{0+}$ it holds that $\mathbb{P}(F) \in \{0,1\}$.

Proof. Since $B_0 = 0$, $B = (B_t)_{t\geq 0}$ and $\mathcal{F}^B_{0+} = \bigcap_{s>0} \mathcal{F}^B_s$ are independent by Theorem 3.13. Since $\mathcal{F}^B_{0+} \subseteq \sigma\{B_t,\, t \geq 0\}$, it follows that $\mathcal{F}^B_{0+}$ is independent of itself what implies the claim. ■

Example 3.15. Let $f : [0,\infty) \to [0,\infty)$ be a continuous function with $f(t) > 0$ for $t > 0$. Since $\sup\limits_{\substack{0\leq s\leq t\\ s\in\mathbb{Q}}} \frac{B_s}{f(s)}$ is $\mathcal{F}^B_t$-measurable, it holds that

$$\limsup_{t\to 0} \frac{B_t}{f(t)} = \lim_{t\searrow 0} \sup_{0\leq s\leq t} \frac{B_s}{f(s)} = \lim_{t\searrow 0} \sup_{\substack{0\leq s\leq t\\ s\in\mathbb{Q}}} \frac{B_s}{f(s)}$$

is $\mathcal{F}^B_{0+}$-measurable.

Therefore, it holds that $\limsup\limits_{t\to 0} \frac{B_t}{f(t)} = \text{const} \in [-\infty,\infty]$ almost surely. Since $(-B_t)_{t\geq 0}$ is another standard Brownian motion, it follows that

$$\limsup_{t\to 0} \frac{B_t}{f(t)} = \limsup_{t\to 0} \frac{-B_t}{f(t)} = -\liminf_{t\to 0} \frac{B_t}{f(t)}\,.$$

Therefore, and since $\liminf \leq \limsup$, we have that $\limsup\limits_{t\to 0} \frac{B_t}{f(t)} = \text{const} \in [0,\infty]$. The explicit constant hinges on the function f and its growth behaviour.

Let $f(t) = \sqrt{t}$, $t_n = \frac{1}{n}$, for $n \in \mathbb{N}$, and $K > 0$. On $A := \{\sqrt{n}B_{t_n} \geq K \text{ i.o.}\}$, it holds that $\limsup\limits_{t\to 0} \frac{B_t}{\sqrt{t}} \geq K$. We have that

$$\begin{aligned}
\mathbb{P}(A) = \mathbb{P}\left(\bigcap_{n\geq 1}\bigcup_{m\geq n} \{\sqrt{m}B_{t_m} \geq K\}\right) &= \lim_{n\to\infty} \mathbb{P}\left(\bigcup_{m\geq n} \{\sqrt{m}B_{t_m} \geq K\}\right)\\
&\geq \limsup_{n\to\infty} \mathbb{P}\big(\{\sqrt{n}B_{t_n} \geq K\}\big)\\
&= 1 - \Phi(K) > 0\,,
\end{aligned}$$

where Φ is the cumulative distribution function of the standard normal distribution. Since $\limsup\limits_{t\to 0} \frac{B_t}{\sqrt{t}}$ is almost surely a constant, which is on A larger than K and $\mathbb{P}(A) > 0$ gilt, and since K can be chosen arbitrarily, we obtain that $\limsup\limits_{t\to 0} \frac{B_t}{\sqrt{t}} = \infty$ a.s. •

Example 3.16. Since $X_t = tB_{1/t}$ is a Brownian motion, results about the limit as $t \to 0$ can be transferred to the limit as $t \to \infty$. Let $\tilde{\mathcal{F}}_t^B = \sigma\{B_s \mid s \geq t\}$, then $\bigcap_{t\geq 0} \tilde{\mathcal{F}}_t^B$ is trivial by Blumenthal's zero-one law for X_t.

We deduce the following limit result: $\limsup\limits_{t\to\infty} \frac{B_t}{\sqrt{t}} = \infty$, since

$$\infty = \limsup_{t\to 0} \frac{X_t}{\sqrt{t}} = \limsup_{t\to 0} \frac{B_{1/t}}{\sqrt{1/t}} = \limsup_{t\to\infty} \frac{B_t}{\sqrt{t}}\,.$$

•

The growth behaviour of Brownian motion is more precisely described by the law of iterated logarithm:

$$\limsup_{t\to\infty} \frac{B_t}{\sqrt{2t\,\log\log(t)}} = 1\ ,\ \text{a.s.}$$

Let me refer to Satz 22.1 in Klenke [2008] for a proof of this more accurate result.

Theorem 3.17 (*Strong Markov property*). *For $T \geq 0$ a stopping time which is almost surely finite, $(B_s^{(1)})_{s\geq 0}$, with $B_s^{(1)} = B_{s+T} - B_T$, is a Brownian motion and independent of $\mathcal{F}_{T+}^B$.*

Proof. Let (B_t) be a $(\mathcal{F}_t)$-Brownian motion, for instance with $\mathcal{F}_t = \mathcal{F}_t^B$. We approximate T by a monotone decreasing sequence of discrete stopping times

$$T_n = \frac{\lfloor 2^n T + 1 \rfloor}{2^n} \quad \downarrow T \text{ as } n \to \infty\,.$$

We define for some Borel set B:

$$A := \left\{(B_{t+T_n} - B_{T_n}) \in B\right\} \text{ and } A_k := \left\{(B_{t+\frac{k}{2^n}} - B_{\frac{k}{2^n}}) \in B\right\}.$$

For $E \in \mathcal{F}_{T_n+}$, it holds that

$$\mathbb{P}(A\cap E) = \sum_{k=0}^{\infty} \mathbb{P}\left(A_k \cap E \cap \{T_n = k2^{-n}\}\right).$$

By Proposition 3.13, $(B_{t+k2^{-n}} - B_{k2^{-n}})$ is independent of $\mathcal{F}_{k2^{-n}+}$. Therefore, by the Markov property in Theorem 3.11 we deduce that

$$\begin{aligned}
\mathbb{P}(A\cap E) &= \sum_{k=0}^{\infty} \mathbb{P}\left(A_k \cap E \cap \{T_n = k2^{-n}\}\right)\\
&= \sum_{k=0}^{\infty} \mathbb{P}(A_k)\,\mathbb{P}\left(E \cap \{T_n = k2^{-n}\}\right)\\
&= \mathbb{P}\left(B_t \in B\right) \sum_{k=0}^{\infty} \mathbb{P}\left(E \cap \{T_n = k2^{-n}\}\right)\\
&= \mathbb{P}\left(B_t \in B\right) \mathbb{P}(E)\,.
\end{aligned}$$

We conclude that $(B_{t+T_n} - B_{T_n})$ is independent of $\mathcal{F}_{T_n+} \supset \mathcal{F}_{T_+}$. The process defined by

$$\lim_{n\to\infty} \left(B_{t+T_n} - B_{T_n}\right)$$

satisfies the characteristic properties of Brownian motion such that we can apply Theorem 3.2). With the independence of $\mathcal{F}_{T+}$, we finish the proof. ∎

Corollary 3.18 (Reflection principle)**.** Let $(B_t)_{t\geq 0}$ be a Brownian motion and T an almost surely finite stopping time. Then we have that

$$B^*_t = \begin{cases} B_t, & t \leq T, \\ 2B_T - B_t, & t > T, \end{cases}$$

is another Brownian motion.

4 Continuous-time martingales and stopping times

4.1 Continuous-time stochastic processes and filtrations

From now on we shall always be working on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ where a filtration $(\mathcal{F}_t)_{t\geq 0}$, that is, a nested family of σ-fields $\mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F}$ for $s \leq t$, is defined that satisfies the usual conditions:

- $\mathcal{F}_s = \bigcap_{t>s} \mathcal{F}_t$ for all $s \geq 0$ (right-continuity);
- all $A \in \mathcal{F}$ with $\mathbb{P}(A) = 0$ are contained in $\mathcal{F}_0$.

A family $(X(t), t \geq 0)$ of $\mathbb{R}^d$-valued random variables on our probability space is called a (continuous-time) stochastic process and this process is $(\mathcal{F}_t)$-adapted if all $X(t)$ are $\mathcal{F}_t$-measurable. Denoting the Borel σ-field on $[0,\infty)$ by $\mathcal{B}_+$, this process X is measurable if $(t,\omega) \mapsto X(t,\omega)$ is a $\mathcal{B}_+ \otimes \mathcal{F}$-measurable mapping.
We can also consider for some fixed times $t_1 < \ldots < t_n$ the random vector $(X(t_1), \ldots, X(t_n))^\top$. The distribution of $(X(t_1), \ldots, X(t_n))$ is called a finite-dimensional (marginal) distribution of $X(t)$ (at times $(t_1, \ldots, t_n)$). To a given stochastic process we can associate the collection of all its finite-dimensional distributions. It already determines the distribution, considered as a random variable on $\mathbb{R}^{[0,\infty)}$, since it determines the measure on a generator of the product-σ-field. There is an important theorem in stochastic process theory, Kolmogorov's existence theorem, that shows that given finite-dimensional distributions with some basic sufficient properties, we can always find an associated continuous-time stochastic process. Here, we focus more on path properties than on this existence theory. In fact, the interplay between path properties and finite-dimensional distributions is often important to analyze stochastic processes.
In this chapter, we say that $(X(t), t \geq 0)$ is continuous if the paths $t \mapsto X(t,\omega)$ are continuous for all $\omega \in \Omega$. A process is measurable if it is (right-) continuous [Karatzas and Shreve, 1991, Prop. 1.13]. A stochastic process $(X(t), t \geq 0)$ is called càdlàg if the paths are right-continuous and left limits $X(t-) = \lim_{s\uparrow t} X(s,\omega)$ exist. We denote with $\Delta X(t) = X(t) - X(t-)$. Analogously, $(X(t), t \geq 0)$ is called càglàd if the paths are left-continuous and right limits $X(t+) = \lim_{s\downarrow t} X(s,\omega)$ exist. Such path properties of $(X(t), t \geq 0)$ as continuity are not determined by the finite-dimensional distributions. A (standard one-dimensional) Brownian motion with respect to the filtration $(\mathcal{F}_t)$ is a continuous $(\mathcal{F}_t)$-adapted real-valued process $(W(t), t \geq 0)$ satisfying the properties from Definition 3.1.

Remark. One often considers a larger filtration $(\mathcal{F}_t)$ than the canonical filtration $(\mathcal{F}_t^W)$ of Brownian motion. For instance, in order to include random initial conditions of some stochastic differential equation (SDE). Given a Brownian motion process W' on a probability space $(\Omega', \mathcal{F}', \mathbb{P}')$ with the canonical filtration $\mathcal{F}'_t = \sigma(W'(s), s \leq t)$ and the random variable X''_0 on a different space $(\Omega'', \mathcal{F}'', \mathbb{P}'')$, we can construct the product space with $\Omega = \Omega' \times \Omega''$, $\mathcal{F} = \mathcal{F}' \otimes \mathcal{F}''$, $\mathbb{P} = \mathbb{P}' \otimes \mathbb{P}''$ such that $W(t,\omega',\omega'') := W'(t,\omega')$ and $X_0(\omega',\omega'') := X''_0(\omega'')$ are independent and W is an $(\mathcal{F}_t)$-Brownian motion for $\mathcal{F}_t = \sigma(X_0; W(s), s \leq t)$. Note that X_0 is $\mathcal{F}_0$-measurable which always implies that X_0 and W are independent.

Notation: We write $(X_t)_{t\geq 0}$ for a stochastic process and usually hide the dependence on ω in the notation.

4.2 Equality of stochastic processes

There are three important types of equalities to distinguish for stochastic processes:

(a) (X_t) and (Y_t) have the same finite-dimensional marginal distributions, i.e.

$$\mathbb{P}\left(X_{t_1} \in B_1, \ldots, X_{t_n} \in B_n\right) = \mathbb{P}\left(Y_{t_1} \in B_1, \ldots, Y_{t_n} \in B_n\right)$$

for all $n \in \mathbb{N}, (t_1, \ldots, t_n) \in \mathbb{R}_+^n$ and $B_1, \ldots, B_n \in \mathcal{B}^d$, with $\mathcal{B}^d$ the Borel σ-field on $\mathbb{R}^d$.

(b) Two processes (X_t) and (Y_t) are modifications of each other if $\mathbb{P}(X_t = Y_t) = 1$ for all $t \geq 0$.

(c) Two processes (X_t) and (Y_t) are indistinguishable if $\mathbb{P}(X_t = Y_t, \forall t \geq 0) = 1$.

We have the following implications: (c) $\Rightarrow$ (b) $\Rightarrow$ (a). For (b) and (c), we require the processes to be defined on a joint probability space. The next example shows that there is a difference between (b) and (c).

Example 4.1. Define $T \sim \mathrm{Exp}(1)$, a standard exponentially distributed random variable. Define $X_t(\omega) \equiv 0$ for all $(t, \omega) \in \mathbb{R}_+ \times \Omega$ and

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

Then, (X_t) and (Y_t) are modifications of each other but not indistinguishable. In particular, (X_t) is continuous and (Y_t) not. •

The difference between (b) and (c) is due to the non countable index set $\mathbb{R}_+$. For discrete-time stochastic processes (X_n) and (Y_n) with $n \in \mathbb{N}$ instead, the σ-additivity of $\mathbb{P}$ implies that

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

However, assuming certain path properties the two notions of equality coincide also for continuous-time processes.

Proposition 4.2. *Let (X_t) and (Y_t) be two right-continuous stochastic processes and modifications of each other. Then, (X_t) and (Y_t) are indistinguishable.*

Proof. Consider a countable dense subset of $\mathbb{R}_+$, for instance, $\mathbb{Q}_+ = \mathbb{Q} \cap \mathbb{R}_+$. Let $A = \{\omega \in \Omega | X_s = Y_s \,, \forall s \in \mathbb{Q}_+\}$. By the σ-sub-additivity of $\mathbb{P}$, we obtain that

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

since (Y_t) is a modification of (X_t). We conclude that $\mathbb{P}(A) = 1$. Since (X_t) and (Y_t) are right-continuous, there exists a set $B \in \mathcal{F}$ with $\mathbb{P}(B) = 1$, such that the paths $t \mapsto X_t(\omega)$ and $t \mapsto Y_t(\omega)$ are for all $\omega \in B$ right-continuous functions. For any $t \in \mathbb{R}_+$, there is a decreasing sequence $t_n \downarrow t$ of numbers in $\mathbb{Q}_+$ and it holds that

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

for all $\omega \in A \cap B$. The result follows with

$$\mathbb{P}(X_t = Y_t \,, \forall t \in \mathbb{R}_+) \geq \mathbb{P}(A \cap B) = 1 \,.$$ ■

4.3 Stopping times

Let $\mathcal{F} = (\mathcal{F}_t)_{t \geq 0}$ be a filtration. A map $T : \Omega \to [0, \infty]$ is called $\mathcal{F}$-stopping time, if for all $t \in [0, \infty)$ it holds that

$$\{T \leq t\} \in \mathcal{F}_t \,.$$

Proposition 4.3. *Let $\mathcal{F} = (\mathcal{F}_t)_{t \geq 0}$ be a right-continuous filtration and for $T : \Omega \to [0, \infty]$ assume that $\{T < t\} \in \mathcal{F}_t$ for all $t \geq 0$. Then T is a $\mathcal{F}$-stopping time.*

Proof. It holds that

$$\{T \leq t\} \subset \left\{T < t + \frac{1}{n}\right\} \in \mathcal{F}_{t+\frac{1}{n}} \quad \forall n \geq 1$$

and $\{T \leq t\} = \bigcap_{n\geq 1}\{T < t + \frac{1}{n}\}$. Hence, $\{T \leq t\} \in \bigcap_{n\geq 1} \mathcal{F}_{t+\frac{1}{n}} = \mathcal{F}_{t+} = \mathcal{F}_t$. ∎

Proposition 4.4. *Let $(T_n)_{n\geq 1}$ be a sequence of $\mathcal{F}$-stopping times.*

1. *If $T_n \nearrow T$ as $n \to \infty$, then T is a $\mathcal{F}$-stopping time.*
2. *If $T_n \searrow T$ as $n \to \infty$ and $\mathcal{F}$ is a right-continuous filtration, then T is a $\mathcal{F}$-stopping time.*

Proof. 1. It holds for all $t \geq 0$

$$\{T \leq t\} = \bigcap_{n\geq 1}\{T_n \leq t\} \in \mathcal{F}_t .$$

2. It holds for all $t \geq 0$

$$\{T \geq t\} = \bigcap_{n\geq 1}\{T_n \geq t\} .$$

We obtain that $\{T < t\} = \bigcup_{n\geq 1}\{T_n < t\} \in \mathcal{F}_t$. Since $\mathcal{F}$ is right-continuous, the claim follows with Proposition 4.3. ∎

For $X = (X_t)_{t\geq 0}$ an $\mathbb{R}^d$-valued stochastic process and $B \in \mathcal{B}^d$, with $\mathcal{B}^d$ the Borel σ-field on $\mathbb{R}^d$,

$$T_B(\omega) = \inf\{t \geq 0 \mid X_t \in B\} ,$$

where $\inf\emptyset = \infty$, defines the first entry time of X in B.

Proposition 4.5. *Let $X = (X_t)_{t\geq 0}$ be an $\mathcal{F}$-adapted stochastic process with right-continuous paths and let $\mathcal{F}$ be right-continuous. Then for all $U \subseteq \mathbb{R}^d$ open, T_U is a stopping time.*

Proof. If $T_U(\omega) \leq s$, there exists a sequence s_n with $s_n \geq s$ and $s_n \to s$ such that $X_{s_n}(\omega) \in U$. Thus, we have that

$$\{T_U < t\} = \{\exists s \in [0,t) : \; X_s \in U\} = \{\exists s \in [0,t) \cap \mathbb{Q} : \; X_s \in U\} = \bigcup_{s\in[0,t)\cap\mathbb{Q}} \{X_s \in U\} \in \mathcal{F}_t ,$$

where we use the right-continuity of X_s and that U is open. Since $\mathcal{F}$ ist right-continuous, Proposition 4.3 gives that T_U is a $\mathcal{F}$-stopping time. ∎

The same result holds for left-continuous $X = (X_t)_{t\geq 0}$.

Proposition 4.6. *Let $X = (X_t)_{t\geq 0}$ be an $\mathcal{F}$-adapted stochastic process with continuous paths and let $\mathcal{F}$ be right-continuous. Then for all $A \subseteq \mathbb{R}^d$ closed, T_A is a stopping time.*

Proof. Let $U_n = \{x \in \mathbb{R}^d \mid d(x,A) < \frac{1}{n}\}$, where $d(x,A) = \inf_{a\in A} \|x - a\|$. For all $n \geq 1$, U_n is open and $U_{n+1} \subseteq U_n$ and $A = \bigcap_{n\geq 1} U_n$, since A is closed. With Proposition 4.5 we conclude that T_{U_n} are $\mathcal{F}$-stopping times. By Proposition 4.4 the claim follows if we can show that $T_{U_n} \nearrow T_A$.
We are left to prove that $T_{U_n} \nearrow T_A$. It holds that T_{U_n} is increasing and $T_{U_n} \leq T_A$. Let $\lim_{n\to\infty} T_{U_n}(\omega) = t < \infty$ (the case $t = \infty$ is clear). By continuity of the paths, $X_{T_{U_n}(\omega)}(\omega) \in \overline{U_n}$ and thus $d(X_{T_{U_n}(\omega)}(\omega), A) \leq \frac{1}{n}$. For $n \to \infty$, the continuity of the paths implies that $d(X_{T_t}(\omega), A) = 0$ and thus $X_{T_t} \in A$, since A is closed. Since $T_A(\omega) \leq t$, we obtain the result. ∎

Recall the definition of the σ-field generated by a $\mathcal{F}$-stopping time T:

$$\mathcal{F}_T = \sigma\big(A \in \mathcal{F} : \; A \cap \{T \leq t\} \in \mathcal{F}_t \; \forall t \geq 0\big) . \tag{4.1}$$

Proposition 4.7. *Let $X = (X_t)_{t\geq 0}$ be an $\mathcal{F}$-adapted stochastic process with right-continuous paths and let $\mathcal{F}$ be right-continuous and T a $\mathcal{F}$-stopping time with $T < \infty$ almost surely. Then, X_T is a $\mathcal{F}_T$-measurable random variable.*

Proof. The maps $(t,\omega) \mapsto X(t,\omega)$ and $\omega \mapsto T(\omega)$ are measurable and thus $X_T(\omega) = X(T(\omega),\omega)$ is well-defined as a composition of measurable mappings and hence a random variable. We are left to show that

$$\{X_T \in A\} \cap \{T \leq t\} \in \mathcal{F}_t \;\forall t \geq 0 \text{ and } A \in \mathcal{B}^d\,.$$

Consider the dyadic partition $S(n) = \{k2^{-n}, k \in \mathbb{N}\}$ and $T_n(\omega) = \min\{s \in S(n) | s > T(\omega)\}$, $t_n = \min\{s \in S(n) | s > t\}$. Then we have that $\{T_n \leq t_n\} = \{T \leq t_n\} \in \mathcal{F}_{t_n}$ and

$$\{X_{T_n} \in A\} \cap \{T_n \leq t_n\} = \bigcup_{k\in\mathbb{N}} \left\{X_{k2^{-n}} \in A, k2^{-n} \leq t_n, T_n = k2^{-n}\right\} \in \mathcal{F}_{t_n}\,.$$

Taking intersections over all $n \in N$ on the left and right side yields that

$$\{X_T \in A\} \cap \{T \leq t\} \in \mathcal{F}_t\,.$$

■

Compared to discrete-time stopping times, considered in Section 2.3, the results for continuous-time stopping times are less simple. They motivate why we usually restrict to right-continuous filtrations.

4.4 Uniform integrability

Definition 4.8. *For a family of random variables $(X_i)_{i\in\mathcal{I}}$, set*

$$\rho(r) := \sup_{i\in\mathcal{I}} \mathbb{E}\left[|X_i| \mathbf{1}_{\{|X_i|>r\}}\right]\,, r \geq 0\,.$$

*$(X_i)_{i\in\mathcal{I}}$ is **uniformly integrable** if* $\lim_{r\to\infty} \rho(r) = 0$.

A uniformly integrable family of random variables $(X_i)_{i\in\mathcal{I}}$ is uniformly bounded in L_1, since

$$\mathbb{E}\left[|X_i|\right] \leq r + \rho(r) \;\forall i \in \mathcal{I}\,.$$

Recall (Satz 1.57 in Bibinger and Holzmann [2019]): $X_n \xrightarrow{\mathbb{P}} X$, $X \in L_1$, and (X_n) uniformly integrable implies that $X_n \xrightarrow{L_1} X$.

Example 4.9. Let (Z_n) be i.i.d. random variables with $\mathbb{P}(Z_n = 0) = 1 - 1/n$ and $\mathbb{P}(Z_n = n) = 1/n$. Then $\mathbb{E}[|Z_n|] = 1$ for all n, such that (Z_n) is uniformly bounded in L_1. It holds that $Z_n \to 0$ almost surely, however, $\lim_{n\to\infty} \mathbb{E}[|Z_n - 0|] = 1$, such that L_1 convergence to 0 does not hold true. In fact, $\rho(r) = 1$ for all $r \geq 0$. •

Corollary 4.10. Let $\mathcal{G} \subseteq \mathcal{F}$ be a sub-σ-field. For $Z_n \xrightarrow{\mathbb{P}} Z$, $Z \in L_1$, and (Z_n) uniformly integrable, it holds that

$$\mathbb{E}[Z_n|\mathcal{G}] \xrightarrow{L_1} \mathbb{E}[Z|\mathcal{G}]\,.$$

Proof. Since

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

the result follows from the L_1-convergence (Satz 1.57 in Bibinger and Holzmann [2019]). ■

Proposition 4.11 (criterion for uniform integrability)**.** *If for a measurable function $\Phi : \mathbb{R}_+ \to \mathbb{R}_+$, which satisfies* $\lim_{r\to\infty} \Phi(r)/r = \infty$*, it holds that*

$$\mathbb{E}\left[\Phi(|X_i|)\right] \le C < \infty \; \forall i \in \mathcal{I}\,,$$

then $(X_i)_{i\in\mathcal{I}}$ is uniformly integrable.

Proof. For $\varepsilon > 0$, there exists some $N \in \mathbb{N}$, such that

$$Z_i = \frac{|X_i|}{\Phi(|X_i|)} \mathbf{1}_{\{|X_i|\ge N\}} \le \frac{\varepsilon}{C}\,.$$

We obtain that

$$\mathbb{E}\left[|X_i|\mathbf{1}_{\{|X_i|\ge N\}}\right] = \mathbb{E}\left[\Phi(|X_i|)Z_i\right] \le \frac{\varepsilon}{C}\,\mathbb{E}\left[\Phi(|X_i|)\right] \le \varepsilon\,.$$

■

Example 4.12. If $\mathbb{E}\left[|X_i|^p\right] \le C$ for all $i \in \mathcal{I}$ with some $p > 1$, uniform integrability holds true. •

Proposition 4.13. *For a family of random variables $(X_i)_{i\in\mathcal{I}}$ are equivalent*

1. *$(X_i)_{i\in\mathcal{I}}$ is uniformly integrable,*
2. $\sup_{i\in\mathcal{I}} \mathbb{E}\left[|X_i|\right] = C < \infty$ *and for all $\varepsilon > 0$ there exists $\delta > 0$, such that for all $A \in \mathcal{F}$ with $\mathbb{P}(A) \le \delta$, it holds that $\mathbb{E}\left[|X_i|\mathbf{1}_A\right] < \varepsilon$.*

Proof. 1.$\Rightarrow$ 2.: It holds that

$$\begin{aligned}\mathbb{E}\left[|X_i|\mathbf{1}_A\right] &\le \mathbb{E}\left[|X_i|\mathbf{1}_{\{A\cap\{|X_i|\le r\}\}}\right] + \mathbb{E}\left[|X_i|\mathbf{1}_{\{A\cap\{|X_i|>r\}\}}\right]\\ &\le r\,\mathbb{P}(A) + \mathbb{E}\left[|X_i|\mathbf{1}_{\{|X_i|>r\}}\right].\end{aligned}$$

For $\varepsilon > 0$, we choose r such that

$$\mathbb{E}\left[|X_i|\mathbf{1}_{\{|X_i|>r\}}\right] < \frac{\varepsilon}{2}\,.$$

Then, we choose $\delta = \varepsilon/(2r)$.
2.$\Rightarrow$ 1.: For $\varepsilon > 0$, set $\delta = \varepsilon/(2r)$ with r chosen as above. By Markov's inequality:

$$\mathbb{P}\left(|X_i| \ge C/\delta\right) \le \frac{\delta}{C}\mathbb{E}[|X_i|] \le \delta\,,\; \forall i \in \mathcal{I}.$$

We conclude that

$$\mathbb{E}\left[|X_i|\mathbf{1}_{\{|X_i|\ge C/\delta\}}\right] < \varepsilon\,,\; \forall i \in \mathcal{I}.$$

■

Proposition 4.14. *Let $(X_i)_{i\in\mathcal{I}}$ be uniformly integrable on $(\Omega,\mathcal{F},\mathbb{P})$. Then*

$$\left\{\mathbb{E}[X_i|\mathcal{G}],\; i\in\mathcal{I},\, \mathcal{G}\subseteq\mathcal{F} \text{ some sub-}\sigma\text{-field}\right\}$$

is uniformly integrable.

Proof. It holds that $\mathbb{E}[|X_i|] \le C\; \forall i \in \mathcal{I}$, and by Proposition 4.13 for $\varepsilon > 0$, there exists $\delta > 0$, such that $\mathbb{E}[|X_i|\mathbf{1}_A] < \varepsilon$ for all $A \in \mathcal{F}$ with $\mathbb{P}(A) \le \delta$, and for all $i \in \mathcal{I}$. Applying Jensen's inequality, we obtain that

$$\left|\mathbb{E}[X_i|\mathcal{G}]\right| \le \mathbb{E}\left[|X_i|\,|\mathcal{G}\right].$$

An application of Markov's inequality yields that

$$\mathbb{P}\left(\left|\mathbb{E}[X_i|\mathcal{G}]\right| > \frac{C}{\delta}\right) \le \frac{\delta}{C}\mathbb{E}\left[\left|\mathbb{E}[X_i|\mathcal{G}]\right|\right]$$

$$\leq \frac{\delta}{C}\mathbb{E}\left[\mathbb{E}[|X_i|\,|\mathcal{G}]\right] = \frac{\delta}{C}\mathbb{E}[|X_i|] \leq \delta\,.$$

We conclude that

$$\mathbb{E}\left[\left|\,\mathbb{E}[X_i|\mathcal{G}]\right|\mathbf{1}_{\{|\mathbb{E}[X_i|\mathcal{G}]|>C/\delta\}}\right] \leq \mathbb{E}\left[|X_i|\mathbf{1}_{\{|\mathbb{E}[X_i|\mathcal{G}]|>C/\delta\}}\right] < \varepsilon\,.$$ ■

4.5 Optional stopping

Let $\mathcal{F}=(\mathcal{F}_t)_{t\geq 0}$ be a filtration. A stochastic process $X=(X_t)_{t\geq 0}$ is a $\mathcal{F}$-martingale, if it is $\mathcal{F}$-adapted, integrable and if for all $0\leq s\leq t$:

$$\mathbb{E}[X_t|\mathcal{F}_s] = X_s \text{ a.s.}$$

For $\mathbb{E}[X_t|\mathcal{F}_s]\geq X_s$, we call X a submartingale, for $\mathbb{E}[X_t|\mathcal{F}_s]\leq X_s$ a supermartingale.
In the sequel, we assume that $\mathcal{F}=(\mathcal{F}_t)_{t\geq 0}$ satisfies the usual conditions.

Theorem 4.15 (Optional stopping)**.** *Let $X=(X_t)_{t\geq 0}$ be a $\mathcal{F}$-(sub)martingale with right-continuous paths and S,T $\mathcal{F}$-stopping times with $S\leq T\leq M$ for some $M>0$. Then, we have that*

$$\mathbb{E}[X_T|\mathcal{F}_S] \overset{(\geq)}{=} X_S\,.$$

Proof. X_S is $\mathcal{F}_S$-measurable by Proposition 4.7. For $\mathcal{T}_n=\{\frac{k}{2^n}, k\in\mathbb{N}_0\}\subset[0,\infty)$, it holds that $\mathcal{T}_n\subset\mathcal{T}_{n+1}$ for all $n\geq 1$. Define

$$T_n = \sum_{k=1}^{\infty} 1_{(\frac{k-1}{2^n},\frac{k}{2^n}]}(T)\frac{k}{2^n}$$

and

$$S_n = \sum_{k=1}^{\infty} 1_{(\frac{k-1}{2^n},\frac{k}{2^n}]}(S)\frac{k}{2^n}\,.$$

It holds that

1. T_n takes values in $\mathcal{T}_n$.
2. T_n is a $\mathcal{F}$-stopping time: $\{T_n=0\}=\{T=0\}\in\mathcal{F}_0$ and for $t>0$, i.e. $t\in[\frac{k-1}{2^n},\frac{k}{2^n})$ for some $k\in\mathbb{N}$, we have by 1.
$$\{T_n\leq t\} = \{T_n\leq \frac{k-1}{2^n}\} = \{T\leq\frac{k-1}{2^n}\}\in\mathcal{F}_{\frac{k-1}{2^n}}\subseteq\mathcal{F}_t\,.$$
3. For all $\omega\in\Omega$, $T_n(\omega)\searrow T(\omega)$ since $T_n(\omega)-T(\omega)\leq\frac{1}{2^n}$ and T_n decreases as n increases.
4. $T_n\leq M+\frac{1}{2^n}\leq M+1$, since $T\leq M$.

These properties apply likewise to S_n. By 3. and the right-continuity of X, it holds that $X_{T_n}\to X_T$ and $X_{S_n}\to X_S$ pointwise. This convergence should hold in L_1 also, thus we require that $(X_{T_n})_{n\geq 1}$ resp. $(X_{S_n})_{n\geq 1}$ are uniformly integrable.

Case 1: Assume X is a martingale. T_{n+1} and T_n take values in $\mathcal{T}_{n+1}$. In particular $T_1,\ldots,T_{n+1}$ are stopping times with respect to $(\mathcal{F}_t)_{t\in\mathcal{T}_{n+1}}$. $(X_t)_{t\in\mathcal{T}_{n+1}}$ is a discrete-time martingale with respect to $(\mathcal{F}_t)_{t\in\mathcal{T}_{n+1}}$. Since $T_{n+1}\leq T_n$, discrete-time optional stopping from Theorem 2.30 implies that

$$\mathbb{E}[X_{T_1}|\mathcal{F}_{T_{n+1}}] = \mathbb{E}[X_{T_n}|\mathcal{F}_{T_{n+1}}] = X_{T_{n+1}}\ a.s.$$

Since $\{X_{T_n},\ n\geq 1\}=\{\mathbb{E}[X_{T_1}|\mathcal{F}_{T_n}],\ n\geq 1\}$, uniform integrability holds true by Proposition 4.14.

Case 2: Assume X is a submartingale. Analogously to Case 1, we have that $\mathbb{E}[X_{T_n}|\mathcal{F}_{T_{n+1}}] \geq X_{T_{n+1}}$. Applying Theorem 2.30 with $S=0$ and $T=T_n$, we obtain that $\mathbb{E}[X_{T_n}] \geq \mathbb{E}[X_0]$. Uniform integrability follows (see problem 3.11 in Karatzas and Shreve [1991]).

For S_n we conclude analogously. Therefore, L_1-convergence holds in both cases.

It holds that $S_n \leq T_n \leq M+1$ and S_n, T_n take values in $\mathcal{T}_n$. By optional stopping from Theorem 2.30 applied to $(X_t)_{t\in\mathcal{T}_n}$, we derive that

$$\mathbb{E}[X_{T_n}|\mathcal{F}_{S_n}] \overset{(\geq)}{=} X_{S_n} \ a.s.$$

Thus, for all $F \in \mathcal{F}_{S_n}$: $\mathbb{E}[X_{T_n}1_F] \overset{(\geq)}{=} \mathbb{E}[X_{S_n}1_F]$. Since $S \leq S_n$ for all n, $\mathcal{F}_S \subseteq \mathcal{F}_{S_n}$ for all n. We obtain that for all $n \geq 1$, $F \in \mathcal{F}_S$:

$$\mathbb{E}[X_{T_n}1_F] \overset{(\geq)}{=} \mathbb{E}[X_{S_n}1_F]$$

and with the L_1-convergence

$$\mathbb{E}[X_T 1_F] \overset{(\geq)}{=} \mathbb{E}[X_S 1_F].$$

■

Proposition 4.16. *Let $X=(X_t)_{t\geq 0}$ be a $\mathcal{F}$-(sub)martingale with right-continuous paths and T a $\mathcal{F}$ stopping time. Then it holds that $(X_{T\wedge t})_{t\geq 0}$ is a $\mathcal{F}$-(sub)martingale.*

Proof. Let $0 \leq s \leq t$. We show that for all $F \in \mathcal{F}_s$ $\mathbb{E}[X_{T\wedge t}1_F] \overset{(\geq)}{=} \mathbb{E}[X_{T\wedge s}1_F]$. Thereto, consider $\mathcal{T}_n = \{s + \frac{(t-s)k}{2^n},\ k \in \mathbb{Z}\} \cap [0,\infty)$. For $u = s + \frac{(t-s)k}{2^n} \in \mathcal{T}_n$, let $u_+ = u + \frac{(t-s)}{2^n} \in \mathcal{T}_n$ be the next grid point. We have that $\mathcal{T}_n \subset \mathcal{T}_{n+1}$ and $s,t \in \mathcal{T}_n$ for all $n \geq 1$. Define

$$T_n = \sum_{u\in\mathcal{T}_n} 1_{(u,u_+]}(T)u_+ + \infty 1_{\{T=\infty\}}\,,$$

which is a $\mathcal{F}$-stopping time. It holds that $T_n \wedge t \searrow T\wedge t$ and $T_n \wedge s \searrow T \wedge s$. With uniform integrability (analogously to the proof of Theorem 4.15), we have that $X_{T_n\wedge t} \to X_{T\wedge t}$ and $X_{T_n\wedge s} \to X_{T\wedge s}$ pointwise and in L_1.
The discrete-time version of the proposition, Theorem 2.25, applied to $(X_n)_{n\in\mathcal{T}_n}$ gives that

$$\mathbb{E}[X_{T_n\wedge t}|\mathcal{F}_s] \overset{(\geq)}{=} X_{T_n\wedge s} \ a.s.$$

We obtain that for all $F \in \mathcal{F}_s$ by the L_1-convergence

$$\mathbb{E}[X_{T\wedge t}1_F] \overset{(\geq)}{=} \mathbb{E}[X_{T\wedge s}1_F]\,.$$

■

Remark. By optional stopping for all $0 \leq s \leq t$, $S_1 := T\wedge s$, $S_2 := T\wedge t$, it holds that $S_1 \leq S_2 \leq t$ and $\mathbb{E}[X_{T\wedge t}|\mathcal{F}_{T\wedge s}] = X_{T\wedge s}$. Hence, $(X_{T\wedge t})_{t\geq 0}$ is a $(\mathcal{F}_{T\wedge t})_{t\geq 0}$-(sub)martingale. This is a smaller filtration compared to the one in Proposition 4.16.

Example 4.17. Denote by $(B_t)_{t\geq 0}$ a standard Brownian motion and $a,b>0$. Then

$$T_{-a,b} = \inf\{t \geq 0 \mid B_t = -a \text{ or } B_t = b\}$$

is a $(\mathcal{F}^B_{t+})$-stopping time.

1. It holds that $\mathbb{P}(B_1 > a+b) = \varepsilon > 0$. For $B_i - B_{i-1} > a+b$ it follows that $T_{-a,b} \leq i$. Thus, $\{T_{-a,b} > n\} \subseteq \bigcap_{i=1}^n \{B_i - B_{i-1} \leq a+b\}$ and $\mathbb{P}(T_{-a,b} > n) \leq (1-\varepsilon)^n$. We obtain that $\mathbb{P}(T_{-a,b} < \infty) = 1$ and thus all moments of $T_{-a,b}$ exist.

2. With Proposition 4.16 , $(B_{T_{-a,b}\wedge t})_{t\geq 0}$ is a martingale. It holds that $\mathbb{E}[B_{T_{-a,b}\wedge t}] = \mathbb{E}[B_0] = 0$ for all $t \in [0,\infty)$. For $t \to \infty$, it follows that

$$B_{T_{-a,b}\wedge t} \to B_{T_{-a,b}} \quad \text{a.s., since } T_{-a,b} < \infty \text{ a.s.}$$

By $|B_{T_{-a,b}\wedge t}| \leq \max\{a,b\}$, $0 = \mathbb{E}[B_{T_{-a,b}\wedge t}] \to \mathbb{E}[B_{T_{-a,b}}]$ with dominated convergence. We obtain that $0 = \mathbb{E}[B_{T_{-a,b}}] = -a\,\mathbb{P}(B_{T_{-a,b}} = -a) + b(1-\mathbb{P}(B_{T_{-a,b}} = -a))$ and thus $\mathbb{P}(B_{T_{-a,b}} = -a) = \frac{b}{a+b}$.

3. Stop $(B_t^2 - t)_{t\geq 0}$ with $T_{-a,b}$ and we have that $\mathbb{E}[T_{-a,b}] = ab$, since $\mathbb{E}[B_{T_{-a,b}}^2] - \mathbb{E}[T_{-a,b}] = 0$ and $\mathbb{E}[B_{T_{-a,b}}^2] = b^2\frac{a}{a+b} + a^2\frac{b}{a+b} = ab$.

4. Let $T_b = \inf\{t \geq 0 \,|B_t = b\}$. For all $a > 0$, $\{T_b < \infty\} \supseteq \{B_{T_{-a,b}} = b\}$ and thus

$$\mathbb{P}(T_b < \infty) \geq \mathbb{P}(B_{T_{-a,b}} = b) = \frac{a}{a+b} \to 1, \quad a \to \infty,$$

and $\mathbb{P}(T_b < \infty) = 1$.

5. Consider the martingale $X_t = \exp(\sigma B_t - \frac{\sigma^2}{2}t)$ for $\sigma > 0$. Stopping in T_b, we obtain for all $t \in [0,\infty)$ that

$$0 \leq X_{T_b \wedge t} \leq \exp(\sigma b)$$

and $\mathbb{E}[X_{T_b\wedge t}] = \mathbb{E}[X_0] = 1$. By dominated convergence for $t \to \infty$ it holds that $1 = \mathbb{E}[X_{T_b}] = \exp(\sigma b - \frac{\sigma^2}{2}T_b)$ resp. $\mathbb{E}[\exp(sT_b)] = \exp(b\sqrt{2s})$ with $s = \frac{\sigma^2}{2}$. This is the moment generating function of the normal distribution. For the density of T_b, it follows that

$$f_{T_b}(t) = \frac{b}{t^{3/2}}\varphi\Big(\frac{b}{t^{1/2}}\Big)1_{[0,\infty)}(t)\,.$$

•

4.6 Doob's Martingale Inequality

Theorem 4.18. *Suppose $(X_n,\mathcal{F}_n)_{0\leq n\leq N}$ is a discrete martingale. Then for every $p \geq 1$ and $\lambda > 0$*

$$\lambda^p\,\mathbb{P}\Big(\sup_{0\leq n\leq N}|X_n| \geq \lambda\Big) \leq \mathbb{E}[|X_N|^p]\,, \tag{4.2}$$

and for every $p > 1$

$$\mathbb{E}\Big[\sup_{0\leq n\leq N}|X_n|^p\Big] \leq \Big(\frac{p}{p-1}\Big)^p\,\mathbb{E}[|X_N|^p]\,. \tag{4.3}$$

▪

Proof. Introduce the stopping time $\tau := \inf\{n\,|\,|X_n| \geq \lambda\} \wedge N$. Since $(|X_n|^p)$ is a submartingale, the optional stopping theorem gives

$$\mathbb{E}[|X_N|^p] \geq \mathbb{E}[|X_\tau|^p] \geq \lambda^p\,\mathbb{P}\Big(\sup_n|X_n| \geq \lambda\Big) + \mathbb{E}[|X_N|^p\mathbf{1}_{\{\sup_n|X_n|<\lambda\}}],$$

which proves the first part. Moreover, we deduce from this inequality for any $K > 0$ and $p > 1$

$$\begin{aligned}\mathbb{E}\Big[\big(\sup_n|X_n| \wedge K\big)^p\Big] &= \mathbb{E}\Big[\int_0^K p\lambda^{p-1}\mathbf{1}_{\{\sup_n|X_n|\geq\lambda\}}\,d\lambda\Big]\\ &\leq \int_0^K p\lambda^{p-2}\,\mathbb{E}[|X_N|\mathbf{1}_{\{\sup_n|X_n|\geq\lambda\}}]\,d\lambda\end{aligned}$$

$$= p\,\mathbb{E}\Big[|X_N|\int_0^{\sup_n|X_n|\wedge K}\lambda^{p-2}\,d\lambda\Big]$$
$$= \frac{p}{p-1}\,\mathbb{E}\big[|X_N|(\sup_n|X_n|\wedge K)^{p-1}\big].$$

By Hölder's inequality,

$$\mathbb{E}\big[(\sup_n|X_n|\wedge K)^p\big] \le \frac{p}{p-1}\,\mathbb{E}\big[(\sup_n|X_n|\wedge K)^p\big]^{(p-1)/p}\,\mathbb{E}[|X_N|^p]^{1/p},$$

which after cancellation and taking the limit $K\to\infty$ yields the asserted moment bound. ■

The continuous-time optional stopping theorem readily allows to extend the first inequality to continuous-time martingales.

Corollary 4.19. (Doob's L_p-inequality) If $(X(t),\mathcal{F}_t)_{t\in I}$ is a right-continuous martingale indexed by a subinterval $I\subset\mathbb{R}$, then for any $p>1$

$$\mathbb{E}\big[\sup_{t\in I}|X(t)|^p\big]^{1/p} \le \frac{p}{p-1}\sup_{t\in I}\mathbb{E}[|X(t)|^p]^{1/p}.$$

Proof. By the right-continuity of X we can restrict the supremum on the left to a countable subset $D\subset I$. This countable set D can be exhausted by an increasing sequence of finite sets $D_n\subset D$ with $\bigcup_n D_n = D$. Then the supremum over D_n increases monotonically to the supremum over D, the preceding theorem applies for each D_n and the monotone convergence theorem yields the asserted inequality. ■

Be aware that Doob's L_p-inequality is different for $p=1$ [Revuz and Yor, 1999, p. 55]. By Theorem 4.15, the inequality (4.2) generalizes as well to continuous-time processes.

4.7 Local martingales

Definition 4.20. *Let $\mathcal{X}$ be some class of stochastic processes. We say that a stochastic process $(X_t)_{t\ge 0}$ is in $\mathcal{X}_{loc}$, if there exists an increasing sequence of stopping times (T_n), with $T_n\uparrow\infty$ almost surely (localizing sequence), such that*

$$\text{for } (X_t^{T_n}),\ X^{T_n} := X_{\bullet\wedge T_n}\,,\ \text{it holds that } (X_t^{T_n})\in\mathcal{X},\ \text{for all } n\in\mathbb{N}\,.$$

*A stochastic process (M_t) is a **local martingale**, if $(M_t^{T_n})$ are martingales for all n and a localizing sequence of stopping times (T_n).*

Similarly, we formalize that path properties hold "locally" in the same way. Note that $\mathcal{X}\subseteq\mathcal{X}_{loc}$ is always true.

Example 4.21. The example reconsiders the "martingale strategy". For (Y_i) i.i.d. random variables with

$$\mathbb{P}(Y_i=1)=\mathbb{P}(Y_i=-1)=\frac{1}{2}\,,$$

we set $X_i=\sum_{j=1}^i \xi_i Y_i$, for $\xi_1=1$ and $\xi_i=2^{i-1}\mathbf{1}_{\{Y_1=\ldots=Y_{i-1}=-1\}}$ for $i\ge 2$. Then, for $T=\min\{i|Y_i=1\}$, the strategy (double bet in case of loss and stop after first success) leads to $X_T=1$ almost surely. To construct a local martingale which is not a martingale, we transfer this game from rounds $n\in\mathbb{N}$ to the continuous interval $[0,1]$. Consider $(Z_t)_{0\le t\le 1}$ with

$$Z_t=\sum_{i=1}^\infty \xi_i Y_i \mathbf{1}_{\{1-2^{-i}\le t\}}\,.$$

Since $\mathbb{E}[Z_{1/2}] = 0$, but $Z_1 = 1$ almost surely, (Z_t) is not a martingale. The sequence

$$T_n = \begin{cases} 1 - 2^{-n}\,, & \text{on } \{Y_1 = \ldots = Y_n = -1\} \\ 1\,, & \text{else} \end{cases}\,,$$

is a localizing sequence and $(Z_t^{T_n})$ a martingale for all n, such that (Z_t) is a local martingale. •

Proposition 4.22. *Let (M_t) be a local martingale.*

(a) If $|M_t| \leq C < \infty$ for all $t \geq 0$ (bounded), then (M_t) is a martingale.

(b) If $M_t^- \leq C < \infty$ for all $t \geq 0$ (bounded from below), then (M_t) is a supermartingale.

Proof. (a) For all $n \in \mathbb{N}$ and $0 \leq s \leq t$ and all $A \in \mathcal{F}_s$, it holds that

$$\mathbb{E}\left[\mathbf{1}_A\left(M_t^{T_n} - M_s^{T_n}\right)\right] = 0\,.$$

Since the integrand is bounded and satisfies $\mathbf{1}_A\left(M_t^{T_n} - M_s^{T_n}\right) \to \mathbf{1}_A\left(M_t - M_s\right)$ almost surely, we obtain by dominated convergence:

$$\mathbb{E}\left[\mathbf{1}_A\left(M_t^{T_n} - M_s^{T_n}\right)\right] \to \mathbb{E}\left[\mathbf{1}_A\left(M_t - M_s\right)\right]\,.$$

By this L_1-convergence, we conclude that $\mathbb{E}\left[M_t - M_s | \mathcal{F}_s\right] = 0$ almost surely.

(b) From $\mathbb{E}\left[M_t^{T_n} | \mathcal{F}_s\right] = M_s^{T_n}$ almost surely and since $M_t - C \geq 0$, we conclude with Fatou's lemma:

$$\mathbb{E}\left[M_t | \mathcal{F}_s\right] \leq M_s \;\; a.s.$$

■

5 Gaussian processes

5.1 Definition and examples

The finite-dimensional distributions of Brownian motion are multivariate normal. Processes with this property are referred to as Gaussian processes.

Definition 5.1. *A real-valued stochastic process $(X_t)_{t\geq 0}$ is called a **Gaussian process** if for all $t_1,\dots,t_n$, $n\in\mathbb{N}$, it holds that $(X_{t_1},\dots,X_{t_n})$ is n-dimensional normally distributed. The process (X_t) is **centered** if $\mathbb{E}[X_t]=0$ for all $t\geq 0$. The function*

$$\Gamma:\mathbb{R}_+\times\mathbb{R}_+\to\mathbb{R},\ \ \Gamma(s,t)\mapsto \mathrm{Cov}(X_s,X_t)$$

*is called the **covariance function** of (X_t).*

The distribution of a Gaussian process is uniquely determined by its expectation and its covariance function. The distribution of a centered Gaussian process is uniquely determined by its covariance function. This is immediate, since the finite-dimensional distributions uniquely characterize the distribution of a process and the normal distribution is uniquely determined by its expectation and variance-covariance matrix. In particular, we have the following characterization from Theorem 3.2 of a standard Brownian motion as a Gaussian process.

Corollary 5.2 (Brownian Motion as a Gaussian Process)**.** A real-valued stochastic process (W_t) is a standard Brownian motion with respect to its natural filtration if and only if (W_t) is a continuous centered Gaussian process with $W_0=0$, and with covariance function

$$\Gamma:\mathbb{R}_+\times\mathbb{R}_+\to\mathbb{R},\ \ \Gamma(s,t)\mapsto \mathrm{Cov}(W_s,W_t)=s\wedge t\,.$$

Example 5.3 (Brownian bridge)**.** Given a stochastic process $(X_t)_{0\leq t\leq T}$ with $X_0=x_0$, an associated bridge from x_0 to x_T is a process $(X_t^{br})_{0\leq t\leq T}$ whose finite-dimensional distributions coincide to those of $(X_t)|\{X_T=x_T\}$.
For (B_t) a standard Brownian motion, the finite-dimensional distributions are Gaussian and the conditional distributions are as well Gaussian. For the bridge (X_t) we thus have the (conditional) density of the finite-dimensional distribution

$$f_{X_{t_1},\dots,X_{t_n}}(x_1,\dots,x_n)=\Big(\prod_{k=1}^n\varphi_{t_k-t_{k-1}}(x_k-x_{k-1})\Big)\frac{\varphi_{T-t_n}(x_T-x_n)}{\varphi_T(x_T-x_0)}\,,$$

with $t_0=0$, where $\varphi_{\sigma^2}(x)$ denotes the normal density with expectation 0 and variance σ^2. $(X_t)_{0\leq t\leq T}$ is called Brownian bridge and $(X_t)_{0\leq t\leq 1}$ with $X_1=0$ standard Brownian bridge. The Brownian bridge can be defined by

$$X_t=x_0+B_t-\frac{t}{T}\big(B_T-(x_T-x_0)\big)\,,0\leq t\leq T\,. \tag{5.1}$$

We find that $\mathbb{E}[X_t]=x_0+t(x_T-x_0)/T$ and $\mathrm{Cov}(X_t,X_s)=s\wedge t-ts/T$. An equivalent definition is hence that $(X_t)_{0\leq t\leq T}$ is a continuous Gaussian process with this expectation and covariance. The Brownian bridge is very important in statistics for cusum and Kolmogorov-Smirnov tests. By Donsker's invariance principle the empirical distribution function F_n of random variables $(Z_i)_{1\leq i\leq n}\overset{iid}{\sim}U(0,1)$ satisfies as $n\to\infty$:

$$\sqrt{n}\big(F_n(t)-F(t)\big)\xrightarrow{d}(X_t)_{0\leq t\leq 1}\,,$$

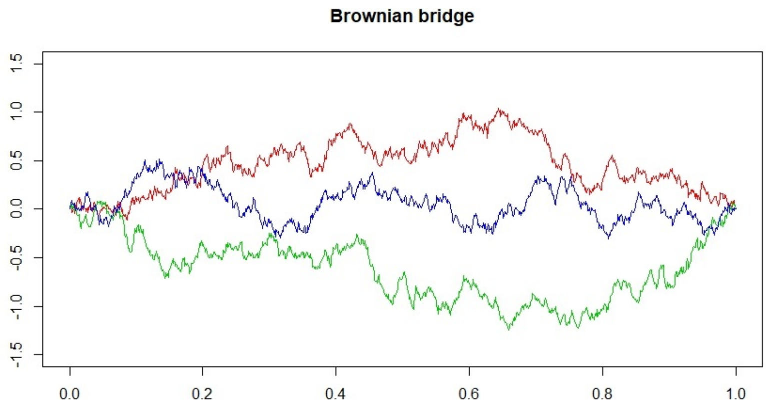


Figure 5.1: Three paths of a standard Brownian bridge.

with F the cdf of the uniform distribution $U(0,1)$ and (X_t) a standard Brownian bridge. This convergence is weak convergence of processes in the Skorokhod space of càdlàg functions. By continuous mapping it further holds as $n \to \infty$ that

$$\sqrt{n} \sup_t \left(F_n(t) - F(t)\right) \xrightarrow{d} \sup_t (X_t)\,.$$

A standard transformation allows to obtain a similar result for a general cdf F. •

Example 5.4 (Ornstein-Uhlenbeck process)**.** Let $\vartheta > 0$ and $\sigma > 0$ be given. The continuous centered Gaussian process (X_t) with covariance function

$$\Gamma : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}, \;\; \Gamma(s,t) \mapsto \mathrm{Cov}(X_s, X_t) = \frac{\sigma^2}{2\vartheta}\left(e^{-\vartheta|s-t|} - e^{-\vartheta(s+t)}\right).$$

is called Ornstein-Uhlenbeck process. The Ornstein-Uhlenbeck process has the feature that it tends to return to its mean. This feature is referred to as mean reversion. The larger the choice of ϑ, the stronger is this effect. The parameter σ on the other hand controls the volatility of the process. These features are exemplified in Figure 5.2. This process is not a martingale, but satisfies the Markov property. •

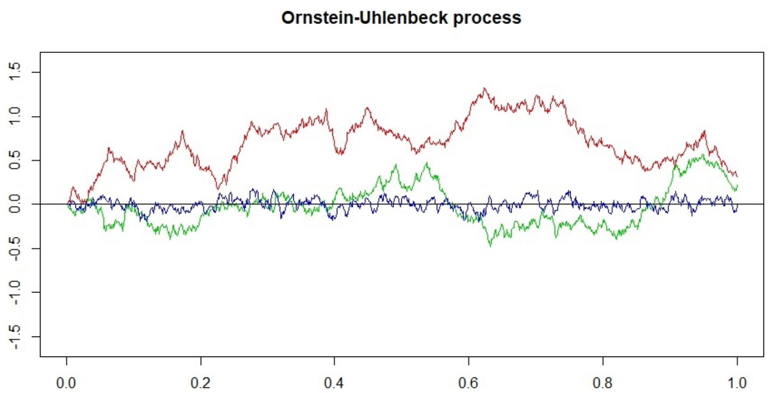


Figure 5.2: Three paths of Ornstein-Uhlenbeck processes with $\sigma = 1$ and $\vartheta = 1, 10, 100$, respectively.

5.2 Fractional Brownian motion and selfsimilarity

Example 5.5 (Hurst phenomenon)**.** Hurst [1] studied in the 1950s an R/S statistic of nile river flow data:

$$R/S(X_1,\ldots,X_n) = \frac{\max_{1\le i\le n}\left(S_i - \frac{i}{n}S_n\right) - \min_{1\le i\le n}\left(S_i - \frac{i}{n}S_n\right)}{\sqrt{\frac{1}{n}\sum_{i=1}^n \left(X_i - \frac{S_n}{n}\right)^2}}$$

with $S_i = \sum_{k=1}^i X_k$. For $(X_i)_{1\le i\le n}$ an i.i.d. sequence of real-valued L_2 random variables with variances σ^2, Donsker's invariance principle would yield that

$$\frac{1}{\sigma\sqrt{n}}\left(\max_{1\le i\le n}\left(S_i - \tfrac{i}{n}S_n\right) - \min_{1\le i\le n}\left(S_i - \tfrac{i}{n}S_n\right)\right) \xrightarrow{d} \sup_{0\le t\le 1}(W_t - tW_1) - \inf_{0\le t\le 1}(W_t - tW_1)$$

as $n\to\infty$ for a Wiener process (W_t). Hence, $n^{-1/2}R/S(X_1,\ldots,X_n)$ would converge as well in distribution to the right-hand side such that $R/S(X_1,\ldots,X_n) \propto \sqrt{n}$, and the growth rate of $\log(R/S(X_1,\ldots,X_n))/\log(n)$ should be 1/2. Instead of the expected growth rate 1/2, however, Hurst found empirically approx. 0.74! This questioned the standard methods at that time and asked for new stochastic models. •

Definition 5.6. *A stochastic process $(B_t^H)_{t\ge 0}$ is called a **fractional Brownian motion (fBm)** with **Hurst exponent** $H\in(0,1)$ if it is a centered Gaussian process with continuous paths and with*

$$\mathrm{Cov}(B_t^H, B_s^H) = \mathbb{E}[B_t^H B_s^H] = \tfrac{1}{2}\left(|t|^{2H} + |s|^{2H} - |t-s|^{2H}\right),\ s,t\in\mathbb{R}_+\,. \tag{5.2}$$

Remark. We readily obtain that $B_0^H = 0$ a.s. and that for $H = 1/2$, fBm is a Brownian motion.

Proposition 5.7 (Properties of fBm)**.** *Let $(B_t^H)_{t\ge 0}$ be a fBm with Hurst exponent $H\in(0,1)$. It holds that*

1. *$\forall t\ge 0$: $B_t^H \sim \mathcal{N}(0,|t|^{2H})$;*
2. *$\forall s,t\ge 0$: $B_t^H - B_s^H \sim \mathcal{N}(0,|t-s|^{2H})$;*
3. *$(B_t^H)_{t\ge 0}$ has stationary increments;*
4. *$(B_t^H)_{t\ge 0} \overset{d}{=} \left(t^{2H}B_{1/t}^H\right)_{t\ge 0}$ (time inversion).*

Proof. 1. Insert $s = t$ in (5.2).

2. It suffices to show that

$$\begin{aligned}\mathrm{Var}(B_t^H - B_s^H) &= \mathrm{Var}(B_t^H) + \mathrm{Var}(B_s^H) - 2\,\mathrm{Cov}(B_t^H, B_s^H)\\ &= |t|^{2H} + |s|^{2H} - \left(|t|^{2H} + |s|^{2H} - |t-s|^{2H}\right) = |t-s|^{2H}\,.\end{aligned}$$

3. By 2. we obtain that $B_{s+t}^H - B_s^H \sim \mathcal{N}(0,|t|^{2H})$ and by 1. we have that $B_t^H \sim \mathcal{N}(0,|t|^{2H})$ for all $s,t\in\mathbb{R}_+$, such that $B_{s+t}^H - B_s^H \overset{d}{=} B_t^H$.

4. The claim follows by

$$\mathbb{E}\left[t^{2H}B_{1/t}^H s^{2H}B_{1/s}^H\right] = \frac{|ts|^{2H}}{2}\left(\left|\frac{1}{t}\right|^{2H} + \left|\frac{1}{s}\right|^{2H} - \left|\frac{1}{t} - \frac{1}{s}\right|^{2H}\right) = \tfrac{1}{2}\left(|t|^{2H} + |s|^{2H} - |t-s|^{2H}\right).$$

■

[1] wiki: Edwin Hurst

Definition 5.8. *A stochastic process $(X_t)_{t\geq 0}$ is called H-**selfsimilar** (H-**ss**) if*

$$\forall a > 0: \quad \left(a^{-H} X_{at}\right) \overset{d}{=} X_t \, .$$

Theorem 5.9. *Let $(B_t^H)_{t\geq 0}$ be a continuous Gaussian process with* $\mathrm{Var}(B_1^H) = 1$. *$(B_t^H)_{t\geq 0}$ is a fBm with Hurst exponent H iff $(B_t^H)_{t\geq 0}$ is H-ss with stationary increments.*

Proof. "$\Rightarrow$": Using (5.2), we compute

$$\mathbb{E}\left[a^{-H} B_{at}^H a^{-H} B_{as}^H\right] = a^{-2H}\, \mathbb{E}[B_{at}^H B_{as}^H] = \tfrac{1}{2}\left(|t|^{2H} + |s|^{2H} - |t-s|^{2H}\right)$$

and we conclude by Definition 5.6.
"$\Leftarrow$": By stationarity of increments we have that $\mathbb{E}[B_{2t}^H] = 2\,\mathbb{E}[B_t^H]$ and by selfsimilarity $\mathbb{E}[B_{2t}^H] = 2^H\,\mathbb{E}[B_t^H]$. Since $H \in (0,1)$, $\mathbb{E}[B_t^H] = 0$ follows. By stationarity of increments and selfsimilarity we obtain that for all $s,t \in \mathbb{R}_+$ that

$$\begin{aligned}
\mathbb{E}[B_t^H B_s^H] &= \tfrac{1}{2}\,\mathbb{E}\left[\left(B_t^H\right)^2 + \left(B_s^H\right)^2 - \left(B_t^H - B_s^H\right)^2\right] \\
&= \tfrac{1}{2}\left(\mathbb{E}\left[\left(B_t^H\right)^2\right] + \mathbb{E}\left[\left(B_s^H\right)^2\right] - \mathbb{E}\left[\left(B_{|t-s|}^H\right)^2\right]\right) \\
&= \tfrac{1}{2}\,\mathbb{E}\left[\left(B_1^H\right)^2\right]\left(|t|^{2H} + |s|^{2H} - |t-s|^{2H}\right),
\end{aligned}$$

and we conclude by Definition 5.6. ∎

Remark. For $H \neq 1/2$, increments of a fBm $(B_t^H)_{t\geq 0}$ are correlated and not independent. In particular, $(B_t^H)_{t\geq 0}$ is not a martingale. The correlation of increments is positive for $H > 1/2$ and negative for $H < 1/2$.

5.3 Regularity and the Kolmogorov-Chentsov theorem

Definition 5.10. *A function $f : D \to \mathbb{R}$, $D \subseteq \mathbb{R}$, is called Hölder-continuous with regularity γ in $x \in D$, if for any $\varepsilon > 0$ there is a constant $C < \infty$, such that for all y with $|x-y| < \varepsilon$:*

$$|f(x) - f(y)| \leq C|x-y|^{\gamma} \, . \tag{5.3}$$

*If (5.3) holds for all $x, y \in D$, f is called **Hölder-continuous** with regularity γ. We say f is **locally Hölder-continuous** with regularity γ, if for any $x \in D$, there exist $\varepsilon > 0$ and $C(x,\varepsilon)$, such that for all y,z with* $\max(|x-z|, |x-y|) < \varepsilon$, *(5.3) holds true.*

The definition extends analogously to maps between metric spaces. For $\gamma = 1$, we know this property as Lipschitz continuity.

Theorem 5.11 (Kolmogorov-Chentsov)**.** *Let $(X_t)_{t\in T}$ be a real-valued random field with T a dense subset of $D \subseteq \mathbb{R}^d$, D a bounded set. Assume there exist $\alpha, \beta, C > 0$, such that*

$$\mathbb{E}\left[|X_t - X_s|^{\alpha}\right] \leq C\,|t-s|^{d+\beta} \; , \; \forall s,t \in T \, . \tag{5.4}$$

Then, there is a continuous modification $(\tilde{X}_t)_{t\in\overline{T}}$, which satisfies for $\gamma < \beta/\alpha$:

$$\max_{\substack{s\neq t \\ s,t\in K}} \frac{|\tilde{X}_t - \tilde{X}_s|}{|t-s|^{\gamma}} < \infty \;\; a.s. \tag{5.5}$$

for every compact set $K \subseteq \overline{T}$.

Proof. We consider w.l.o.g. $D = [0,1]^d$ and T the dyadic rationals.

1. For $f(t) : T \to \mathbb{R}$, $T = \bigcup_{m=1}^{\infty} T_m$, $T_m = \{x \in \mathbb{R}^d \,|x_j = k2^{-m}, 0 \leq k \leq 2^m\}$, with $|f(t) - f(s)| \leq C_1|t-s|^\gamma$, $\gamma > 0, C_1 < \infty$, for all neighboring pairs $s,t \in T_m$, $m \in \mathbb{N}$, there is a continuous modification $\tilde{f} : [0,1]^d \to \mathbb{R}$, with

$$\left|\tilde{f}(s) - \tilde{f}(t)\right| \leq \tilde{C}|t-s|^\gamma \text{ with a constant } \tilde{C} < \infty\,.$$

We prove **1.** by 'chaining'. For each $t \in [0,1]^d$, there is a sequence $(s_m(t))_{m\in\mathbb{N}} \subseteq T$, such that $s_m(t) \to t$. It holds that $|t - s_m(t)| \leq \sqrt{d}2^{-m}$. Since $s_{m+1}(t)$ can be reached starting from $s_m(t)$ via a short chain along neighbors with at most $2d$ steps, we derive that

$$\left|f(s_{m+1}(t)) - f(s_m(t))\right| \leq 2dC_1 2^{-\gamma m}\,.$$

Since the right-hand side is summable (in m), $(f(s_m(t)))_{m\in\mathbb{N}}$ forms a Cauchy sequence. Define $\tilde{f}(t) := \lim_{m\to\infty} f(s_m(t))$. Denote by C_d a generic constant (which may change from line to line) which depends on d. We have that

$$\left|\tilde{f}(t) - f(s_m(t))\right| \leq \sum_{k=m+1}^{\infty} \left|f(s_k(t)) - f(s_{k-1}(t))\right| \leq \sum_{k=m+1}^{\infty} C_d\, 2^{-\gamma(k-1)} \leq C_d 2^{-\gamma m}\,.$$

For some $s \neq t$, $s,t \in [0,1]^d$, there is some $m \in \mathbb{N}$, such that

$$2^{-(m+1)} \leq |s-t| < 2^{-m}\,,$$

and the triangle inequality yields that

$$\begin{aligned}\left|\tilde{f}(s) - \tilde{f}(t)\right| &\leq \left|\tilde{f}(s) - f(s_m(s))\right| + \left|f(s_m(s)) - f(s_m(t))\right| + \left|f(s_m(t)) - \tilde{f}(t)\right| \\ &\leq C_d 2^{-\gamma m}\,,\end{aligned}$$

for some constant C_d. This completes the proof of **1.**

2. By (5.4) and Markov's inequality, we obtain for s,t neighbors in T_m that:

$$\begin{aligned}\mathbb{P}\left(|X_t - X_s| \geq |t-s|^\gamma\right) &\leq \frac{\mathbb{E}\left[|X_t - X_s|^\alpha\right]}{|t-s|^{\gamma\alpha}} \\ &\leq C|t-s|^{d+\beta-\gamma\alpha}\,.\end{aligned}$$

Since $\#T_m = (2^m+1)^d \sim 2^{md}$, and for fix $x \in T_m$ it holds that $\#\{z \in T_m||x-z| \leq 2^{-m}\} \leq 3^d$, we deduce for

$$B_m := \left\{|X_t - X_s| \geq |t-s|^\gamma \text{ for some neighbors } s,t \in T_m\right\},$$

with subadditivity, that

$$\begin{aligned}\mathbb{P}\left(B_m\right) &\leq 3^d(2^m+1)^d C\, 2^{-m(d+\beta-\gamma\alpha)} \\ &\leq C_d\, 2^{-(\beta-\gamma\alpha)}\,,\end{aligned}$$

which is summable (in m) for $\gamma < \beta/\alpha$. The Borel-Cantelli lemma thus yields that

$$\mathbb{P}(B_m \text{ infinitely often}) = 0\,.$$

But then, almost surely, $|X_t - X_s| \leq C_1|t-s|^\gamma$, for all neighboring pairs $s,t \in T_m$, $m \in \mathbb{N}$, with some constant $C_1 < \infty$. Applying **1.** finishes the proof. ■

Corollary 5.12. Let $(X_t)_{t\geq 0}$ be a real-valued process. For any $T > 0$, there exist $\alpha, \beta, C > 0$, such that (5.4) holds true for all $s,t \in [0,T]$. Then there is a locally Hölder-continuous modification of $(X_t)_{t\geq 0}$

with regularity γ for all $\gamma < \beta/\alpha$.

Proof. By Theorem 5.11, for any $T > 0$, there is a Hölder-continuous modification of (X_t^T) with regularity $\gamma < \beta/\alpha$. For $S, T > 0$, (X_t^T) and (X_t^S) are indistinguishable on $[0, S \wedge T]$ by Proposition 4.2. Hence,

$$\mathbb{P}\Big(\bigcup_{S,T\in\mathbb{N}} \{\text{for some } t \in [0, S\wedge T] : X_t^T \neq X_t^S\}\Big) = 0\,.$$

For all $\omega \in \Omega$ in the complement of this null set, set $\tilde{X}_t(\omega) := X_t^T(\omega)$, then $(\tilde{X}_t)_{t\geq 0}$ is a locally Hölder-continuous modification of $(X_t)_{t\geq 0}$ with regularity $\gamma < \beta/\alpha$. ■

Example 5.13. 1. For (W_t) the one-dimensional standard Brownian motion, we obtain from $(W_t - W_s) \sim \mathcal{N}(0, t-s)$, $t \geq s$, and by the moments of the normal distribution:

$$\mathbb{E}\left[|W_t - W_s|^{2k}\right] = (2k-1)!!\,|t-s|^k,\, k \in \mathbb{N},\ \text{for all } t, s \geq 0\,.$$

For all $k \in \mathbb{N}$ (5.4) holds with $\alpha = 2k$ and $\beta = k-1$. Theorem 5.11 thus implies existence of a version of Brownian motion with continuous paths. Moreover, since

$$\frac{\beta}{\alpha} = \frac{k-1}{2k} \longrightarrow \frac{1}{2} \ \text{ as } k \to \infty\,,$$

by Corollary 5.12 the paths of W_t are locally Hölder-continuous for all $\gamma < 1/2$, we write $W_t \in C^{1/2-}$.

2. For (N_t) the Poisson process, $t \geq s$, it holds that

$$\mathbb{E}\left[|N_t - N_s|\right] = \lambda(t-s),\ \text{ for } \alpha = 1\,,$$

such that (5.4) holds for $\beta = 0$. $\beta > 0$ is, however, required for the claim and the process of course does not have continuous paths. Note that higher moments of the Poisson distribution contain as well a first power of the parameter. The above moment bound implies that $X_{t_n} \xrightarrow{\mathbb{P}} X_t$ for $t_n \to t$, i.e. the process is continuous in probability.

3. For (B_t^H) the fractional Brownian motion with Hurst exponent H, we have that $(B_t^H - B_s^H) \sim \mathcal{N}(0, |t-s|^{2H})$. We obtain that

$$\mathbb{E}\left[|B_t^H - B_s^H|^{2k}\right] = (2k-1)!!\,|t-s|^{2Hk},\, k \in \mathbb{N},\ \text{for all } t, s \geq 0\,.$$

With $\alpha = 2k$ and $\beta = 2Hk - 1$, by

$$\frac{\beta}{\alpha} = H - \frac{1}{k} \longrightarrow H \ \text{ as } k \to \infty\,,$$

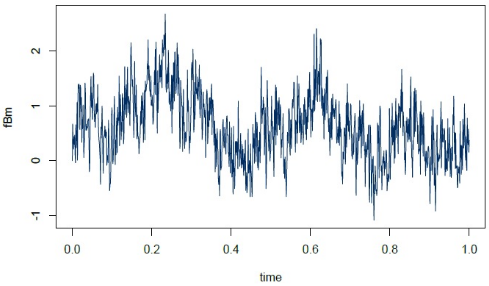
we conclude that $B_t^H \in C^{H-}$. The smoothness of the sample paths hinges on H, see the illustration in Figure 5.3.

•

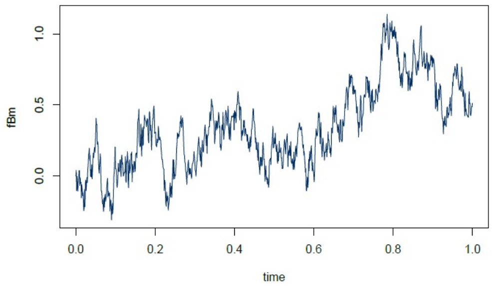
Proposition 5.14. *For any $\varepsilon > 0$, the paths of Brownian motion (W_t) are almost surely locally Hölder-continuous with regularity $1/2 - \varepsilon$. The paths of Brownian motion are, however, almost surely nowhere Hölder-continuous with regularity $1/2$.*

Proof. We only need to prove the second claim, since the first claim has been found in the preceding example. Assume that with some $\gamma > 1/2$ on some interval $[a, b] \subset \mathbb{R}_+$

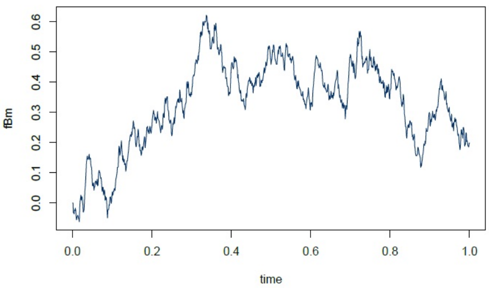
$$|W_t - W_s| \leq c(\omega, \gamma, [a,b])\,|t-s|^{\gamma}\,. \tag{5.6}$$



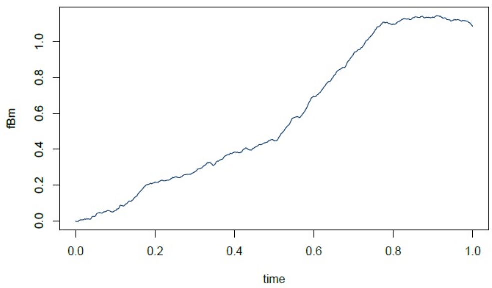
(a) H=0.1, n=1000



(b) H=0.4, n=1000



(c) H=0.6, n=1000



(d) H=0.9, n=1000

Figure 5.3: Simulated paths of a fBm with different Hurst exponents.

Then, for all partitions $\{t_j, j=1,\ldots,n\}$ of $[a,b]$ we have that

$$\sum_j (W_{t_j} - W_{t_{j-1}})^2 \leq c^2 \sum_j (t_j - t_{j-1})^{2\gamma} \leq c^2 \sup_j (t_j - t_{j-1})^{2\gamma-1} (b-a) \to 0\,.$$

This is in contrast to the result that $\sum_j (W_{t_j} - W_{t_{j-1}})^2 \xrightarrow{L_2} (b-a)$, compare Proposition 6.5. Hence, we obtain a contradiction such that (5.6) can only hold on a null set. ■

Analogously, a similar result can be proved for fractional Brownian motion.

6 Itô calculus

6.1 The Stieltjes integral

From measure and integration theory, one knows that there is a one-to-one correspondence between right-continuous functions $F_\mu : \mathbb{R} \to \mathbb{R}$ of bounded variation and finite signed measures on $(\mathbb{R}, \mathcal{B})$ via the relation $\mu((a,b]) = F_\mu(b) - F_\mu(a)$. The integral $\int f \, d\mu$ in this case is called the (Lebesgue-) Stieltjes integral $\int f \, dF_\mu$. A process (X_t) on a compact time interval $[0,T]$ is of bounded variation, if

$$\sup_\pi \sum_i \left| X(t_{i+1}) - X(t_i) \right| < \infty, \text{ a.s.},$$

for partitions $\pi = \{t_1, \ldots, t_n | t_1 < \ldots < t_n\}$ of $[0,T]$. A process $(X(t))_{t \in \mathbb{R}_+}$ is of finite variation if it is of bounded variation over any compact interval. Denote with $\mathcal{V}$ the set of adapted càdlàg processes of finite variation.
For $X(t,\omega) \in \mathcal{V}$ define pathwise for a measurable bounded process H:

$$\int_0^T H(s,\omega) \, dX(s,\omega) = \int_0^T H(s,\omega) \, \mu_X(\omega, ds)$$

with $\mu_X(\omega, [0,t]) = X(t,\omega)$ the associated Lebesgue-Stieltjes measure. If $s \mapsto H(s,\omega)$ is continuous, the above integral exists as Riemann-Stieltjes integral and

$$\sum_i H(\tau_i, \omega) \big(X(t_i, \omega) - X(t_{i-1}, \omega) \big) \to \int_0^T H(s,\omega) \, dX(s,\omega) \tag{6.1}$$

as $\sup_i |t_i - t_{i-1}| \to 0$ for any $t_{i-1} \le \tau_i \le t_i$. The Riemann-Stieltjes integral as the above limit exists for one continuous process and one process of bounded variation. We have the partial integration formula $\int_0^T f \, dF = f F \big|_0^T - \int_0^T F \, df$. Thereby, we obtain with a Wiener process W, for instance,

$$\int_0^t e^t \, dW(t,\omega) = e^t W(t,\omega) \big|_0^T - \int_0^T W(t,\omega) \, de^t = e^T W(T,\omega) - \int_0^T W(t,\omega) e^t \, dt \,.$$

Problem 6.1. *The pathwise integral can not be extended to integrators with infinite variation.*

Example 6.2. Consider as an example $\int_0^T W(t) \, dW(t)$ with a Wiener process W. Different discrete approximations are

$$\underline{W}(t,\omega) = \sum_{j=0}^{2^n T - 1} W(j/2^n, \omega) \mathbf{1}_{[j/2^n, (j+1)/2^n)}(t) \,, \tag{6.2a}$$

$$\overline{W}(t,\omega) = \sum_{j=0}^{2^n T - 1} W((j+1)/2^n, \omega) \mathbf{1}_{(j/2^n, (j+1)/2^n]}(t) \,. \tag{6.2b}$$

assuming for simplicity that $2^n T \in \mathbb{N}$. We derive for the two different Riemann-Stieltjes sum approximations

$$\begin{aligned}
&\mathbb{E}\Big[\sum_j W(j/2^n) \big(W((j+1)/2^n) - W(j/2^n) \big) \Big] \\
&\quad = \sum_j \mathbb{E}\left[W(j/2^n) \big(W((j+1)/2^n) - W(j/2^n) \big) \right] = 0 \,, \\
&\mathbb{E}\Big[\sum_j W((j+1)/2^n) \big(W((j+1)/2^n) - W(j/2^n) \big) \Big]
\end{aligned}$$

$$= \sum_j \mathbb{E}\left[\left(W((j+1)/2^n) - W(j/2^n)\right)^2\right] = T\,.$$

Further, one can show that both, (6.2a) and (6.2b), are tight. Thus, we conclude that the natural approximation of the integral by Riemann-Stieltjes sums of the form $\sum_j W(\tau_j)(W(t_j) - W(t_{j-1})$ are, contrarily to the Riemann-Stieltjes integral, not independent of the exact choice of the τ_j. •

6.2 Quadratic variation of Brownian motion

Throughout this chapter, we work with a filtration which satisfies the usual conditions.

Definition 6.3. *We say that a real-valued process $(X(t))_{t\in\mathbb{R}_+}$ has **finite quadratic variation**, if a finite process $\big([X,X](t)\big)_{t\in\mathbb{R}_+}$ exists, such that*

$$\sum_{i=1}^{n}\left(X(t_i) - X(t_{i-1})\right)^2 \xrightarrow{p} [X,X](t)\ , t\in\mathbb{R}_+\,, \tag{6.3}$$

as $|\Pi| \to 0$ where Π denotes a partition given by real numbers (t_i) with $t_0 = 0 < t_1 < \ldots < t_n = t$, and $|\Pi| = \max_i(t_{i+1} - t_i)$.
*We call $[X,X](t)$ the **quadratic variation (process)** of X. The **quadratic covariation** up to time t between two processes X and Y is given by*

$$[X,Y](t) = p - \lim_{|\Pi|\to 0}\sum_{t_i\in\Pi}(X(t_{i+1}) - X(t_i))(Y(t_{i+1}) - Y(t_i)) \quad \forall t \geq 0\,.$$

Lemma 6.4. *For a continuous real-valued process X of bounded variation it holds true that $[X,X](t) = 0$.*

Proof.

$$\sum_{i=1}^{n}\left(X(t_i) - X(t_{i-1})\right)^2 \leq \sup_{1\leq i\leq n}\left|X(t_i) - X(t_{i-1})\right| \sum_{i=1}^{n}|X(t_i) - X(t_{i-1})| \to 0$$

what implies the claim. ■

Proposition 6.5. *For a one-dimensional standard Brownian motion, we have $[B,B](t) = t$.*

This proof is similar to the proof of Proposition 3.5.

Proof. It is sufficient to prove that

$$\sum_{i=1}^{n}\left(B(t_i) - B(t_{i-1})\right)^2 \xrightarrow{L_2} t\,. \tag{6.4}$$

We have already seen that

$$\mathbb{E}\Big[\sum_{i=1}^{n}\left(B(t_i) - B(t_{i-1})\right)^2\Big] = t\,.$$

Set $Z_i = \big(B(t_i) - B(t_{i-1})\big)^2 - (t_i - t_{i-1}), i = 1,\ldots,n$. Then, we derive that $\mathbb{E}[Z_i] = 0$ for all $i = 1,\ldots,n$, and

$$\begin{aligned}\operatorname{Var}\Big(\sum_{i=1}^{n} Z_i\Big) &= \sum_{i=1}^{n}\operatorname{Var}(Z_i) = \sum_{i=1}^{n} 2(t_i - t_{i-1})^2 \\ &\leq 2\sup_i |t_i - t_{i-1}| \sum_{i=1}^{n}(t_i - t_{i-1}) \to 0\,.\end{aligned}$$

■

In fact, it is possible to show that the above convergence holds in the stronger almost sure sense, cf. page 29 of Revuz and Yor [1999]. Thus, we rediscover that the paths of Brownian motion are almost surely not of bounded variation, cf. Corollary 3.6. Consequently, we can not define integrals like $\int_0^t W(t)\,dW(t)$ in the Riemann-Stieltjes sense.

Example 6.6. Reconsider a Riemann-Stieltjes type approximating sum for $\int_0^T B(t)\,dB(t)$, with a standard Brownian motion B, again. First, we take left interval endpoints and find that

$$\begin{aligned}\sum_{i=1}^n B(t_{i-1})\big(B(t_i)-B(t_{i-1})\big) &= \sum_{i=1}^n \frac{B^2(t_i)-B^2(t_{i-1})-\big(B(t_i)-B(t_{i-1})\big)^2}{2}\\ &= \frac{1}{2}\sum_{i=1}^n \big(B^2(t_i)-B^2(t_{i-1})\big)-\frac{1}{2}\sum_{i=1}^n \big(B(t_i)-B(t_{i-1})\big)^2\\ &\xrightarrow{p} \frac{1}{2}\big(B_T^2-T\big).\end{aligned}$$

Using left endpoints was Itô's choice and the limit defines the specific Itô integral

$$\int_0^T B(t)\,dB(t)=\frac{1}{2}\big((B(T))^2-T\big). \tag{6.5}$$

Choosing other intermediate points τ_i than left endpoints renders different results. For $\tau_i=(1-\alpha)t_{i-1}+\alpha t_i, 0\le i\le n$, one obtains for $0\le\alpha\le 1$:

$$\sum_{i=1}^n B(\tau_i)\big(B(t_i)-B(t_{i-1})\big)\xrightarrow{p}\frac{(B(T))^2}{2}+\Big(\alpha-\frac{1}{2}\Big)T. \tag{6.6}$$

Hence, each $\alpha\in[0,1]$ induces a different notion of the integral. There are three important and meaningful cases:

- $\alpha=0$, this renders the Itô integral.
- $\alpha=1/2$, this leads to the Fisk-Stratonovich integral, the second important notion of stochastic integrals.
- $\alpha=1$ corresponds to the backward (Itô) integral.

•

Observe that for the Fisk-Stratonovich integral, which we denote by $\int_0^T B(t)\circ dB(t)=B_T^2/2$, the usual calculus well-known from Riemann integration theory applies. However, in this case $\int_0^T B(t)\circ dB(t)$ is not a martingale. Only Itô's choice renders martingales if we perform stochastic integration with martingales as integrators. Especially for financial applications and interpretation, this is a crucial aspect.

6.3 The integral for simple processes

Definition 6.7. *A stochastic process $(H(t))_{t\in\mathbb{R}_+}$ is called **simple predictable** if it is for some $n\in\mathbb{N}$ of the form*

$$H(t,\omega)=H_0\mathbf{1}_0(t)+\sum_{i=1}^n H_{i-1}(\omega)\mathbf{1}_{(T_{i-1},T_i]}(t,\omega),$$

with an increasing finite sequence $(T_i)_{0\le i\le n}$ of stopping times and $\mathcal{F}_{T_{i-1}}$-measurable random variables H_{i-1}.
Equipped with the norm $\|H\|=\sup_{t,\omega}|H(t,\omega)|$, we consider the normed space of simple predictable

processes $\mathcal{S}$. Let L^0 be the space of real-valued random variables topologized through convergence in probability, e.g., with the metric (cf. Theorem 3.5 of Dudley [1976])

$$\rho(X,Y) = \mathbb{E}\left[\frac{|X-Y|}{1+|X-Y|}\right].$$

*For some càdlàg processes X and $H \in \mathcal{S}$, we define **the stochastic integral** $I_X(H)$ as the random variable*

$$I_X(H) = \sum_{i=1}^{n} H_{i-1}\big(X(T_i) - X(T_{i-1})\big) \in L^0. \tag{6.7}$$

This is independent from the particular illustration. $I_X : \mathcal{S} \to L^0$ is a linear functional.

Remark. Our construction of stochastic integration follows Protter [1992], with one minor modification. In Protter [1992], $I_X(H)$ is defined with the additional summand H_0X_0. This is in contrast to the interpretation of the stochastic integral in finance and to other other books on stochastic calculus, for instance, [Karatzas and Shreve, 1991, Ch. 3, Def. 2.3]. Here, we modify Protter's definition accordingly, which means that we integrate rather with respect to $(X_t - X_0)$ instead of X_t. Since this discussion of the initial value is not particularly important, let us simply suppose that $X_0 = 0$, and we may set $H_0 = 0$ as well.

6.4 Extension of the integral

Definition 6.8. *Some adapted càdlàg process X is called a **total semimartingale (or good integrator)** if $I_X : \mathcal{S} \to L^0$ is continuous, that is,*

$$\text{for } (H^n)_{n\in\mathbb{N}} \subseteq \mathcal{S}, H \in \mathcal{S}: \; H^n \to H \;\Rightarrow\; I_X(H^n) \xrightarrow{p} I_X(H).$$

The last definition is equivalent to:

$$\text{For } (H^n)_{n\in\mathbb{N}} \subseteq \mathcal{S}, H \in \mathcal{S}: \; \|H^n\| \to 0 \;\Rightarrow\; \lim_{n\to\infty} \mathbb{P}\big(|I_X(H^n)| \geq \varepsilon\big) = 0 \;\forall \varepsilon > 0.$$

Definition 6.9. *A process X is called a semimartingale if the stopped process $(X(s\wedge t)_{s\in\mathbb{R}_+}$ is a total semimartingale for all $t \in \mathbb{R}_+$.*

Theorem 6.10 (Jacod's theorem)**.** *Let $\mathbb{P}, \mathbb{Q}$ be two probability measures with $\mathbb{Q} \ll \mathbb{P}$. Then, any $\mathbb{P}$-semimartingale is a $\mathbb{Q}$-semimartingale.* ▪

Proof. Since convergence in probability under $\mathbb{P}$ implies convergence in probability under $\mathbb{Q}$, the theorem readily follows from Definition 6.9. ■

Proposition 6.11. *Any process $X \in \mathcal{V}$, that is, an adapted càdlàg process of bounded variation, is a semimartingale.*

Proof. Denote with $\mathbf{V}(X)_T$ the variation of X on $[0,T]$. It holds that

$$\big|I_X(H)\big| \leq \|H\| \, \mathbf{V}(X)_T.$$

If $\|H^n\| \to 0$, it follows that $\big|I_X(H^n)\big| \xrightarrow{p} 0$. ■

Proposition 6.12. *Any L_2-martingale with càdlàg paths is a semimartingale.*

Proof. From Doob's optional stopping theorem we know that for a right-continuous martingale $(X(t))_{t\in\mathbb{R}_+}$ and $[0,T]$-valued stopping times τ_1,τ_2 with $\tau_1\le\tau_2$, $X(\tau_1)=\mathbb{E}\left[X(\tau_2)|\mathcal{F}_{\tau_1}\right]$, almost surely.
Jensen's inequality yields that $X^2(\tau_1)\le\mathbb{E}\left[X^2(\tau_2)|\mathcal{F}_{\tau_1}\right]$. Let $H\in\mathcal{S}$. By optional stopping and the martingale property, we conclude

$$\begin{aligned}\mathbb{E}\left[\left(I_X(H)\right)^2\right]&=\mathbb{E}\left[\left(\sum_{i=1}^n H_{i-1}\left(X(T_i)-X(T_{i-1})\right)\right)^2\right]\\&=\mathbb{E}\left[\sum_{i=1}^n H_{i-1}^2\left(X(T_i)-X(T_{i-1})\right)^2\right]\\&\le\|H\|^2\,\mathbb{E}\left[\sum_{i=1}^n\left(X(T_i)-X(T_{i-1})\right)^2\right]=\|H\|^2\,\mathbb{E}[X^2(T_n)-X^2(T_0)]\,.\end{aligned}$$

Therefore, when $\|H^n\|\to 0$, we have that $I_X(H^n)\to 0$ in $L_2(\mathbb{P})$. This implies convergence of $I_X(H^n)$ to zero in probability. ∎

Corollary 6.13. The standard Brownian motion is a semimartingale.

Definition 6.14. *A sequence of processes $(H^n)_{n\in\mathbb{N}}$ is said to converge* ***uniformly on compacts in probability (ucp)*** *to the process H, if for any $t>0$:*

$$\sup_{0\le s\le t}\left|H^n(s)-H(s)\right|\xrightarrow{P}0\,.$$

In the following, we write $\mathbb{D}$ and $\mathbb{L}$ for the space of all adapted càdlàg processes and the space of all adapted càglàd processes, respectively. Restricting to bounded processes, we write $\mathbf{b}\mathbb{D}$ and $\mathbf{b}\mathbb{L}$. Note that ucp-convergence is metrizable and we write $\mathbb{D}_{ucp}$ for the normed space.

Theorem 6.15. *$\mathbb{D}_{ucp}$ is a Banach space.* ▪

Proof. Let X^n be a Cauchy sequence with respect to the metric d in $\mathbb{D}_{ucp}$. Then, for sufficiently large n,m, it holds true that

$$\mathbb{P}\Big(\sup_{0\le s\le t}\left|X^n(s)-X^m(s)\right|>\varepsilon\Big)\le\varepsilon$$

for any given $\varepsilon>0$. We construct a sequence $(n_k)_{k\in\mathbb{N}}$ recursively with $n_0=0$ and n_k the smallest number, such that for all $n,m\ge n_k$:

$$\mathbb{P}\Big(\sup_{0\le s\le t}\left|X^n(s)-X^m(s)\right|>2^{-k}\Big)\le 2^{-k}\,.$$

With Borel-Cantelli, we deduce that

$$\underbrace{\mathbb{P}\left(\sup_{0\le t\le T}\left|X^{n_{k+1}}(t)-X^{n_k}(t)\right|>2^{-k}\text{ infinitely often}\right)}_{=A}=0\,.$$

For $\omega\in A^\complement$ the sequence $(X^{n_k})_{k\in\mathbb{N}}$ is a Cauchy sequence with respect to the sup-norm on $[0,T]$. In particular, for any fix $t\in\mathbb{R}_+$, $(X^{n_k}(t))_{k\in\mathbb{N}}$ forms a real Cauchy sequence. Set

$$X(t,\omega)=\begin{cases}\lim_{k\to\infty}X^{n_k}(t,\omega) & ,\omega\in A^\complement\\ 0 & ,\omega\in A\end{cases}\quad\forall\,t\in[0,T]\,.$$

X is adapted and càdlàg by uniform convergence. Moreover,

$$\mathbb{P}\Big(\sup_{0\le t\le T}\left|X^{n_k}(t)-X(t)\right|\to 0\Big)\ge\mathbb{P}(A^\complement)=1\,.$$

Convergence with respect to d of a subsequence implies for a Cauchy sequence convergence to the same limit, what completes the proof. ∎

Theorem 6.16. *$\mathcal{S}$ is dense in $\mathbb{L}_{ucp}$, that is, for all $Y \in \mathbb{L}$ there exists a sequence $(H^n) \subseteq \mathcal{S}$, such that $H^n \xrightarrow{ucp} Y$.*

Proof. Let $Y \in \mathbb{L}$ and $R_n = \inf\{t \,|\, |Y(t)| > n\}$. For the sequence of stopping times R_n, define $Y^n(t) = Y(t \wedge R_n)$, such that $Y^n \in \mathbf{b}\mathbb{L}$. The càglàd property ensures that

$$\mathbb{P}\big(\sup_{0\le t\le T} \big|Y^n(t) - Y(t)\big| \ge \varepsilon\big) \ge \mathbb{P}\big(\sup_{0\le t\le T} |Y(s)| \ge n\big) \to 0\,.$$

Thus $\mathbf{b}\mathbb{L}$ is dense in $\mathbb{L}$ and it remains to prove that $\mathcal{S}$ is dense in $\mathbf{b}\mathbb{L}$. Set $Z(t) = \lim_{u\downarrow t} Y(u)$ with $Z \in \mathbb{D}$ and $T_0^\varepsilon = 0, \dots, T_{n+1}^\varepsilon = \inf\{t \,|\, t > T_n^\varepsilon \text{ and } |Z(t) - Z(T_n^\varepsilon)| > \varepsilon\}$. Since $Z \in \mathbb{D}$, by Proposition 4.5 T_n^ε is a localizing sequence of stopping times increasing to infinity. It holds that

$$Z^\varepsilon = \sum_n Z(T_n^\varepsilon)\mathbf{1}_{[T_n^\varepsilon, T_{n+1}^\varepsilon)} \xrightarrow{ucp} Z \quad \text{as } \varepsilon \to 0\,.$$

Finally, define the simple processes

$$Y^{n,\varepsilon} = Y_0\mathbf{1}_{\{0\}} + \sum_{i=1}^n Z(T_i^\varepsilon)\mathbf{1}_{[T_i^\varepsilon \wedge n, T_{i+1}^\varepsilon \wedge n)} \xrightarrow{ucp} Y \quad \text{as } \varepsilon \to 0, n \to \infty\,.$$

∎

Definition 6.17. *For $X \in \mathbb{D}$ and $Y \in \mathcal{S}$, let $J_X : \mathcal{S} \to \mathbb{D}$ be the linear functional*

$$J_X(Y) = \sum_{i=1}^n Y_i\big(X(T_i \wedge \bullet) - X(T_{i-1} \wedge \bullet)\big) \tag{6.8}$$

*for $Y = \sum_{i=1}^n Y_i \mathbf{1}_{(T_{i-1}, T_i]}$ with an increasing finite sequence $(T_i)_{0\le i\le n}$ of stopping times and $\mathcal{F}_{T_i}$-measurable random variables Y_i. The image space is now $\mathbb{D}$ and the process $((J_X(Y))_t)_{t\in\mathbb{R}_+}$ is called **the stochastic integral** of Y with respect to X. A common short notation is $Y \bullet X$.*

Theorem 6.18. *For a semimartingale X, the mapping $J_X : \mathcal{S} \to \mathbb{D}_{ucp}$ is continuous, that is,*

$$H^n \xrightarrow{ucp} H \;\Rightarrow J_X(H^n) \xrightarrow{ucp} J_X(H)\,.$$

Proof. By linearity we can without loss of generality consider the case $H = 0$.
Step 1 For $\|H^n\| \to 0$ and H^n uniformly bounded $J_X(H^n) \to 0$ holds.
Let X be a total semimartingale and $\delta > 0$ fix. We construct a sequence of stopping times

$$T_n = \inf\big\{t \ge 0 \mid |(J_X(H^n))_t > \delta\big\}\,.$$

Then, $H^n\mathbf{1}_{[0,T_n]} \in \mathcal{S}$ with $\|H^n\mathbf{1}_{[0,T_n]}\| \to 0$. By right-continuity of $(J_X(H^n))_t$:

$$\begin{aligned}\mathbb{P}\big(\sup_{0\le s\le t} |(J_X(H^n))_s| > \delta\big) &\le \mathbb{P}\left(|(J_X(H^n))_{T_n\wedge t}| > \delta\right)\\ &= \mathbb{P}\left(|(J_X(H^n\mathbf{1}_{[0,T_n]}))_t| > \delta\right)\\ &= \mathbb{P}\left(|I_X(H^n\mathbf{1}_{[0,T_n\wedge t]})| > \delta\right) \to 0\,,\end{aligned}$$

such that $J_X(H^n) \to 0$ follows.
Step 2 For $H^n \xrightarrow{ucp} 0$, $J_X(H^n) \xrightarrow{ucp} 0$ holds.
Let $\|H^n\|_{ucp} \to 0$ and $\delta > 0, \eta > 0, t > 0$, $R_n = \inf\{s| \, |H^n(s)| > \eta\}$, $\tilde{H}^n = H^n \mathbf{1}_{[0,R_n]}$. Hence, $\tilde{H}^n \in \mathcal{S}$ and $\|\tilde{H}^n\| \le \eta$ by left-continuity. If $R_n \ge t$ we have that $\sup_{0\le s\le t} |(J_X(H^n))_s| = \sup_{0\le s\le t} |(J_X(\tilde{H}^n))_s|$. We conclude that

$$\begin{aligned}\mathbb{P}\big(\sup_{0\le s\le t} |(J_X(H^n))_s| > \delta\big) &\le \mathbb{P}\big(\sup_{0\le s\le t} |(J_X(\tilde{H}^n))_s| > \delta\big) + \mathbb{P}(R_n < t)\\ &\mathbb{P}\big(\sup_{0\le s\le t} |(J_X(\tilde{H}^n))_s| > \delta\big) + \mathbb{P}\big(\sup_{0\le s\le t} |H^n(s)| > \eta\big)\,.\end{aligned}$$

By step 1, for any $\varepsilon > 0$, we can choose η small enough to bound the probability in the first summand from above by $\varepsilon/2$. For large enough n the second probability can be bounded from above by $\varepsilon/2$ what yields the result. ∎

Since $\mathcal{S}_{ucp}$ is dense in $\mathbb{L}_{ucp}$ and $\mathbb{D}_{ucp}$ is a Banach space the result above that $J_X : \mathcal{S}_{ucp} \to \mathbb{D}_{ucp}$ is continuous extends to $J_X : \mathbb{L}_{ucp} \to \mathbb{D}_{ucp}$.

Definition 6.19. *Let X be a semimartingale. The continuous linear functional $J_X : \mathbb{L}_{ucp} \to \mathbb{D}_{ucp}$ is called the* ***stochastic integral****.*

Remark. If one restricts to integrators which are continuous L_2-martingales (like Brownian motion), the class of integrands can be taken even larger. Let V be the class of real-valued stochastic processes $(Y(t), t \ge 0)$ that are adapted, measurable and that satisfy

$$\|Y\|_V := \Big(\int_0^\infty \mathbb{E}\big[Y(t)^2\big]\,dt\Big)^{1/2} < \infty.$$

Let V^* be the class of real-valued stochastic processes $(Y(t), t \ge 0)$ that are adapted, measurable and that satisfy

$$\mathbb{P}\Big(\int_0^\infty Y(t)^2\,dt < \infty\Big) = 1.$$

For integrators which are continuous L_2-martingales, all elements of V are possible integrands.

6.5 Properties of stochastic integrals

Proposition 6.20. *Let X_1, X_2 be semimartingales and $Y_1, Y_2 \in \mathbb{L}$. It holds true that:*

(i) $$\begin{aligned}\int (Y_1 + Y_2)\,d(X_1 + X_2) &= \int Y_1\,d(X_1 + X_2) + \int Y_2\,d(X_1 + X_2)\\ &= \int Y_1\,dX_1 + \int Y_1\,dX_2 + \int Y_2\,dX_1 + \int Y_2\,dX_2\,.\end{aligned}$$

(ii) $\Delta\big(\int Y_1\,dX_1\big) = Y_1\,\Delta X_1$. [1]

(iii) $\int Y_1\,dX_1$ *is a semimartingale and*

$$\int Y_2\,d\Big(\int Y_1\,dX_1\Big) = \int Y_1 Y_2\,dX_1 \text{ (associativity).}$$

(iv) If X has paths of finite variation, $\int Y_1\,dX_1$ is indistinguishable from the Lebesgue-Stieltjes integral.

(v) $\mathbb{P} \ll \mathbb{Q}$. *Then,* $Y_1 \bullet_{\mathbb{P}} X_1 = Y_1 \bullet_{\mathbb{Q}} X_1$. [1]

(vi) If X_1 is an L_2-martingale and $Y_1 \in \mathbf{b}\mathbb{L}$, then $\int Y_1\,dX_1$ is a L_2-martingale.

[1] That means the processes are indistinguishable.

Proof. For (i)–(v) see Protter [1992], Chapter II.5. We shall prove (vi).
Let $|Y_1(t)| \leq L$, $H^n \subseteq \mathcal{S}$ with $H^n \xrightarrow{ucp} Y_1$. We can assume that $|H^n(t)| \leq L$. Then

$$\begin{aligned}
\mathbb{E}\left[\left(\left(\int H^n\,dX_1\right)(t)\right)^2\right] &= \mathbb{E}\left[\sum_{i=1}^n \left(H^n_{i-1}(X_1(T_i) - X_1(T_{i-1}))\right)^2\right] \\
&\leq L^2\,\mathbb{E}\left[\sum_{i=1}^n (X_1(T_i) - X_1(T_{i-1}))^2\right] \\
&= L^2\,\mathbb{E}\left[X_\infty^2\right] < \infty\,.
\end{aligned}$$

Since $\|H^n \bullet X_1 - Y_1 \bullet X_1\|_{ucp} \to 0$, there exists for fix t a subsequence (n_k) with

$$\sup_{0\leq s\leq t}\left(\left(H^{n_k} - Y_1\right)\bullet X_1\right)(s) \to 0 \ \text{ a.s.}$$

Therefore, by Fatou's lemma

$$\begin{aligned}
\mathbb{E}\left[\left((Y_1 \bullet X_1)(t)\right)^2\right] &= \mathbb{E}\left[\lim_{n_k}\left((H^{n_k} \bullet X_1)(t)\right)^2\right] \\
&\leq \liminf_{n_k}\mathbb{E}\left[\left((H^{n_k} \bullet X_1)(t)\right)^2\right] \\
&\leq L^2\,\mathbb{E}\left[X_\infty^2\right].
\end{aligned}$$

Thus, $Y_1 \bullet X_1$ is in L_2. It easily follows that $(H^n \bullet X_1)_{n\geq 1}$ are martingales.

$$\begin{aligned}
\mathbb{P}\left(\left|\mathbb{E}\left[(Y_1 \bullet X_1)(t)|\mathcal{F}_s\right] - (Y_1 \bullet X_1)(s)\right| \geq \varepsilon\right) &\leq \mathbb{P}\left(\mathbb{E}\left[((Y_1 - H^{n_k})\bullet X_1)(t)|\mathcal{F}_s\right] \geq \varepsilon/2\right) \\
&\quad + \mathbb{P}\left(((Y_1 - H^{n_k})\bullet X_1)(s) \geq \varepsilon/2\right).
\end{aligned}$$

The second term tends to zero and for the first term we deduce that

$$\begin{aligned}
\mathbb{E}\left[\mathbb{E}\left[|(Y_1 - H^{n_k})\bullet X_1|(t)|\mathcal{F}_s\right]\right] &\leq \left(\mathbb{E}\left[\left(|(Y_1 - H^{n_k})\bullet X_1|(t)\right)^2\right]\mathbb{P}\left(|(Y_1 - H^{n_k})\bullet X_1|(t) > 1\right)\right)^{1/2} \\
&\quad + \mathbb{E}\left[|(Y_1 - H^{n_k})\bullet X_1|(t)\mathbf{1}\{|(Y_1 - H^{n_k})\bullet X_1|(t) \leq 1\}\right]
\end{aligned}$$

and thus tends to zero by dominated convergence. We derive that

$$\mathbb{E}\left[|(Y_1 - H^{n_k})\bullet X_1|(t)|\mathcal{F}_s\right] \xrightarrow{p} 0\,.$$

what completes the proof. ∎

We write that a sequence of random partitions $\sigma_n : 0 = T_0^n \leq T_1^n \leq \ldots \leq T_{k_n}^n$ tends to identity if

- $\lim_{n\to\infty} T_{k_n}^n = \infty$ a.s.
- The mesh $\sup_k |T_k^n - T_{k-1}^n| \to 0$ a.s.

For some given process Y and a random partition σ, we define the simple predictable process

$$Y^\sigma = \sum_{i=1}^k Y(\omega, T_{i-1})\,\mathbf{1}_{(T_{i-1},T_i]}(\omega,t)\,. \tag{6.9}$$

Theorem 6.21. *Let X be a semimartingale and $Y \in \mathbb{L}$ or $Y \in \mathbb{D}$. Let σ_n be a sequence of random* *partitions which tends to identity. Then, the elementary integrals*

$$\int Y^{\sigma_n}(s)\,dX(s) = \sum_{i=1}^{k} Y(T_{i-1}^n)\big(X(T_i^n \wedge \bullet) - X(T_{i-1}^n \wedge \bullet)\big)$$

converge ucp to the stochastic integral $\int Y_-\,dX$.

Proof. Without loss of generality, let $X(0) = 0$ and $Y_- \in \mathbb{L}$. Suppose $(Y^k) \subseteq \mathcal{S}$, such that $Y^k \xrightarrow{ucp} Y_-$. We have that

$$\begin{aligned}\int \big(Y_- - Y^{\sigma_n}\big)(s)\,dX(s) &= \int \big(Y_- - Y^k\big)(s)\,dX(s) + \int \big(Y^k - (Y^k)^{\sigma_n}\big)(s)\,dX(s) \\ &\quad + \int \big((Y^k)^{\sigma_n} - Y^{\sigma_n}\big)(s)\,dX(s)\end{aligned}$$

with X a càdlàg version. The first and the third addends tend to zero and it remains to show that

$$\int \big(Y^k - (Y^k)^{\sigma_n}\big)(s)\,dX(s) \xrightarrow{ucp} 0\,.$$

For fix k and as $n \to \infty$, this is ensured by right-continuity of X and since $\sup|T_i^n - T_{i-1}^n| \to 0$. ■

6.6 Quadratic variation

Definition 6.22. *Let X and Y be semimartingales. The* ***quadratic variation process*** $[X,X](t)$ *of X is defined by*

$$[X,X] := X^2 - 2\int X_-\,dX - (X(0))^2\,. \tag{6.10a}$$

The ***quadratic covariation process*** $[X,Y](t)$ *of X and Y is defined by*

$$[X,Y] := XY - \int X_-\,dY - \int Y_-\,dX - X(0)Y(0)\,. \tag{6.10b}$$

Remark.

- Bilinearity and symmetry gives the polarization identity

$$[X,Y] = \frac{1}{2}\Big([X+Y,X+Y] - [X,X] - [Y,Y]\Big)\,. \tag{6.11}$$

- Definition 6.22 for semimartingales turns out to be equivalent to the general definition of quadratic variation used in Proposition 6.5, see below.

Example 6.23. For a standard Brownian motion B, we obtain

$$[B,B](t) = (B(t))^2 - 2\int B(t)\,dB(t) = t \quad \forall\, t \in \mathbb{R}_+\,.$$

•

Proposition 6.24. *For a semimartingale X, the quadratic variation process $[X,X]$ is an adapted, monotone increasing, càdlàg process that satisfies*

(i) $[X,X](t) - [X,X]_-(t) = \big(\Delta X(t)\big)^2$.

(ii) *For any sequence of random partitions* $0 = T_0^n \le T_1^n \le \ldots \le T_{k_n}^n$ *which tends to identity, it holds true that*

$$\sum_{i=1}^{k_n} \big(X(T_i^n \wedge \bullet) - X(T_{i-1}^n \wedge \bullet)\big)^2 \xrightarrow{ucp} [X,X]\,.$$

(iii) *For a stopping time* τ*, we have*

$$[X(\tau \wedge \bullet), X] = [X, X(\tau \wedge \bullet)] = [X,X](\tau \wedge \bullet)\,.$$

(iv) $[X,X]$ *is of bounded variation.*

Proof. (i)

$$\begin{aligned}[X,X](t) - [X,X]_-(t) &= \Delta(X^2(t)) - 2X(t-)\Delta X(t)\\ &= (X(t))^2 - (X(t-))^2 - 2X(t-)\Delta X(t)\\ &= \big(X(t) - X(t-)\big)^2 = \big(\Delta X(t)\big)^2\,.\end{aligned}$$

(ii)

$$\begin{aligned}\sum_{i=1}^{k_n} \big(X(T_i^n \wedge t) - X(T_{i-1}^n \wedge t)\big)^2 &= \sum_{i=1}^{k_n} \big((X(T_i^n \wedge t))^2 - (X(T_{i-1}^n \wedge t))^2\big)\\ &\quad - 2\sum_{i=1}^{k_n} X(T_{i-1}^n \wedge t)\big(X(T_i^n \wedge t) - X(T_{i-1}^n \wedge t)\big)\\ = (X(t))^2 - (X(0))^2 - 2\sum_{i=1}^{k_n} X(T_{i-1}^n \wedge t)\big(X(T_i^n \wedge t) - X(T_{i-1}^n \wedge t)\big)&\\ \xrightarrow{ucp} (X(t))^2 - (X(0))^2 - 2\int_0^t X_-\, dX\,.&\end{aligned}$$

(ii) implies (iii). (iv) follows by monotonicity, for the covariation with polarization. ■

For a semimartingale (X_t), we decompose the quadratic variation

$$[X,X](t) = [X,X]^c(t) + \sum_{0 \le s \le t} (\Delta X_s)^2\,, \tag{6.12}$$

in the quadratic variation of the continuous part and the sum of squared jumps.
We conclude that a continuous L_2-martingale of bounded variation is almost surely constant.

Corollary 6.25 (Partial integration)**.** For semimartingales X and Y, XY is a semimartingale and

$$XY - X_0Y_0 = \int X_-\, dY + \int Y_-\, dX + [X,Y]\,.$$

Proposition 6.26. *Let* X,Y *be semimartingales and* $H,K \in \mathbb{L}$*. It holds true that*

$$[H \bullet X, K \bullet Y](t) = \int_0^t H(s)\,K(s)\,d[X,Y](s)\,. \tag{6.13}$$

Proof. It suffices to show that

$$[H \bullet X, Z] = H \bullet [X,Z]$$

for a semimartingale Z. Without loss of generality, $X_0 = Y_0 = 0$. For $H(s) = \mathbf{1}_{[0,T]}(s)$ with a stopping time T, we have by Proposition 6.24 (iii)

$$[X(T \wedge \bullet), Z] = [X,Z](T \wedge \bullet)\,.$$

For $H(s) = U\,\mathbf{1}_{(S,T]}(s)$ with stopping times S,T and $U \in \mathcal{F}_S$, it holds that

$$\int H(s)\,dX(s) = U\big(X(T\wedge\bullet) - X(S\wedge\bullet)\big)$$

and

$$\begin{aligned}[H\bullet X, Z] &= [U\big(X(T\wedge\bullet) - X(S\wedge\bullet)\big), Z] \\ &= U\Big([X,Z](T\wedge\bullet) - [X,Z](S\wedge\bullet)\Big) \\ &= \int H(s)\,d[X,Z](s) = H\bullet[X,Z]\,.\end{aligned}$$

For $H \in \mathcal{S}$, the claim follows by linearity. For $H \in \mathbb{L}$ there exists a sequence $(H^n), H^n \in \mathcal{S}$ with $H^n \xrightarrow{ucp} H$. Thereby

$$\begin{aligned}[H^n\bullet X, Z] &= (H^n\bullet X)Z - \int Z_-\,d(H^n\bullet X) - \int (H^n\bullet X)_-\,dZ \\ &= (H^n\bullet X)Z - \int Z_- H^n{}_-\,dX - \int (H^n X)_-\,dZ\,,\end{aligned}$$

using partial integration. Now, $(H^nX)_- \to (HX)_-$ in $\mathbb{L}_{ucp}$, $H^n \to H$ in $\mathbb{L}_{ucp}$, such that

$$\lim_{n\to\infty}[H^n\bullet X, Z] = [H\bullet X, Z] = \lim_{n\to\infty} H^n\bullet[X,Z]\,,$$

and the proposition follows. ■

Corollary 6.27 (Itô isometry)**.** Let W be a standard Brownian motion and $H \in \mathbb{L}$. It holds that

$$[H\bullet W, H\bullet W](t) = \int_0^t H(s)^2\,ds$$

as well as

$$\mathbb{E}\left[\left(\int_0^t H(s)\,dW(s)\right)^2\right] = \int_0^t \mathbb{E}\left[H^2(s)\right]ds$$

with Fubini's theorem.

6.7 Multidimensional Case

Definition 6.28.

1. *An m-dimensional process is a semimartingale if each component is a one-dimensional semimartingale.*
2. *An $\mathbb{R}^m$-valued $(\mathcal{F}_t)$-adapted stochastic process $W(t) = (W_1(t),\dots,W_m(t))^T$ is an **m-dimensional Brownian motion** if each component W_i, $i = 1,\dots,m$, is a one-dimensional $(\mathcal{F}_t)$-Brownian motion and all components are independent.*
3. *If Y is an $\mathbb{R}^{d\times m}$-valued stochastic process such that each component Y_{ij}, $1\le i\le d$, $1\le j\le m$, is an element of V^* from the remark below Definition 6.19, then the multidimensional Itô integral $\int Y\,dW$ for m-dimensional Brownian motion W is an $\mathbb{R}^d$-valued random variable with components*
$$\left(\int_0^\infty Y(t)\,dW(t)\right)_i := \sum_{j=1}^m \int_0^\infty Y_{ij}(t)\,dW_j(t), \quad 1\le i\le d\,,$$
and analogously for m-dimensional semimartingales.

Proposition 6.29. *The Itô isometry extends to the multidimensional case such that for $\mathbb{R}^{d\times m}$-valued processes X, Y with components in V and m-dimensional Brownian motion W*

$$\mathbb{E}\Big[\Big\langle \int_0^\infty X(t)\,dW(t), \int_0^\infty Y(t)\,dW(t)\Big\rangle\Big] = \int_0^\infty \sum_{i=1}^d \sum_{j=1}^m \mathbb{E}[X_{ij}(t)Y_{ij}(t)]\,dt\,,$$

where $\langle \bullet, \bullet\rangle$ denotes the scalar product.

Proof. The term in the brackets on the left hand side is equal to

$$\sum_{i=1}^d \sum_{j=1}^m \sum_{k=1}^m \int_0^\infty X_{ij}(t)\,dW_j(t) \int_0^\infty Y_{ik}(t)\,dW_k(t)$$

and the result follows from the one-dimensional Itô isometry once the following claim has been proved: stochastic integrals with respect to independent Brownian motions are uncorrelated (attention: they may well be dependent).

For this, let us consider two independent Brownian motions W_1 and W_2 and two simple processes Y_1, Y_2 in V on the same filtered probability space with

$$Y_k(t) = \sum_{i=0}^\infty \eta_{ik}(\omega)\mathbf{1}_{[t_i,t_{i+1})}(t), \quad k \in \{1,2\}.$$

The common partition of the time axis can always be achieved by taking a common refinement of the two partitions. Then by the $\mathcal{F}_{t_i}$-measurability of η_{ik} we obtain

$$\begin{aligned}
&\mathbb{E}\Big[\int_0^\infty Y_1(t)\,dW_1(t) \int_0^\infty Y_2(t)\,dW_2(t)\Big] \\
&= \sum_{0\le i\le j<\infty} \mathbb{E}\big[\eta_{i1}\eta_{j2}(W_1(t_{i+1})-W_1(t_i))(W_2(t_{j+1})-W_2(t_j))\big] \\
&\quad + \sum_{0\le j<i<\infty} \mathbb{E}\big[\eta_{i1}\eta_{j2}(W_1(t_{i+1})-W_1(t_i))(W_2(t_{j+1})-W_2(t_j))\big] \\
&= \sum_{0\le i\le j<\infty} \mathbb{E}\big[\eta_{i1}\eta_{j2}(W_1(t_{i+1})-W_1(t_i))\big]\,\mathbb{E}\big[W_2(t_{j+1})-W_2(t_j)\big] \\
&\quad + \sum_{0\le j<i<\infty} \mathbb{E}\big[\eta_{i1}\eta_{j2}(W_2(t_{i+1})-W_2(t_i))\big]\,\mathbb{E}\big[W_1(t_{j+1})-W_1(t_j)\big] \\
&= 0\,.
\end{aligned}$$

For each process in V there exists a sequence of simple processes such that the corresponding stochastic integrals converge in $L_2(\mathbb{P})$, which implies that the respective covariances converge, too. This density argument proves the general case. ∎

6.8 The Itô formula

Theorem 6.30 (Itô's lemma). *Let X be a one-dimensional continuous semimartingale and $F \in \mathcal{C}^{2,1}(\mathbb{R}\times\mathbb{R}_+,\mathbb{R})$. Then, $F(X(t),t)$ is a continuous semimartingale with*

$$\begin{aligned}
F(X(t),t) = F(X(0),0) &+ \int_0^t \frac{\partial F}{\partial t}(X(s),s)\,ds + \int_0^t \frac{\partial F}{\partial x}(X(s),s)\,dX(s) \\
&+ \frac{1}{2}\int_0^t \frac{\partial^2 F}{\partial x^2}(X(s),s)\,d[X,X](s)\,.
\end{aligned} \tag{6.14}$$

For a d-dimensional continuous semimartingale and $F \in \mathcal{C}^{2,1}(\mathbb{R}^d\times\mathbb{R}_+,\mathbb{R})$, $F(X(t),t)$ is a continu-

ous semimartingale with

$$F(X(t),t) = F(X(0),0) + \int_0^t \frac{\partial F}{\partial t}(X(s),s)\,ds + \sum_{i=1}^d \int_0^t \frac{\partial F}{\partial x_i}(X(s),s)\,dX_i(s) \tag{6.15}$$
$$+ \frac{1}{2}\sum_{i,j=1}^d \int_0^t \frac{\partial^2 F}{\partial x_i \partial x_j}(X(s),s)\,d[X_i,X_j](s)\,.$$

Proof. The equations hold for constant F. We first restrict to $F(X(t))$ with $F \in \mathcal{C}^2(\mathbb{R}^d,\mathbb{R})$. Let F be a polynomial of degree n and $G(X(t)) = X_l(t)F(X(t))$, $1 \le l \le d$. By virtue of additivity it suffices to consider monomials. Partial integration yields

$$X_l(t)F(X(t)) = X_l(0)F(X(0)) + \int_0^t X_l(s)\,dF(X(s)) + [F(X),X_l](t) + \int_0^t F(X(s))\,dX_l(s).$$

We pursue an induction. Using associativity and by the induction hypothesis, we have that

$$\int_0^t X_l(s)\,dF(X(s)) = \sum_{i=1}^d \int_0^t X_l(s)\frac{\partial F}{\partial x_i}(X(s))\,dX_i(s) + \frac{1}{2}\sum_{i,j=1}^d \int_0^t X_l(s)\frac{\partial^2 F}{\partial x_i \partial x_j}(X(s))\,d[X_i,X_j](s).$$

By Corollary 6.25:

$$\begin{aligned}
[F(X),X_l](t) &= \sum_{i=1}^d \int_0^t \frac{\partial F}{\partial x_i}(X(s))\,d[X_i,X_l](s) \\
&\quad + \frac{1}{2}\sum_{i,j=1}^d \int_0^t \frac{\partial^2 F}{\partial x_i \partial x_j}(X(s))\,d[[X_i,X_j],X_l](s) \\
&= \sum_{i=1}^d \int_0^t \frac{\partial F}{\partial x_i}(X(s))\,d[X_i,X_l](s)\,.
\end{aligned}$$

We obtain that

$$\begin{aligned}
G(X(t)) = X_l(0)F(X(0)) &+ \sum_{i=1}^d \int_0^t X_l(s)\frac{\partial F}{\partial x_i}(X(s))\,dX_i(s) \\
&+ \sum_{i=1}^d \int_0^t \frac{\partial F}{\partial x_i}(X(s))\,d[X_i,X_l](s) \\
&+ \frac{1}{2}\sum_{i,j=1}^d \int_0^t X_l(s)\frac{\partial^2 F}{\partial x_i \partial x_j}(X(s))\,d[X_i,X_j](s) \\
&+ \int_0^t F(X(s))\,dX_l(s)\,,
\end{aligned}$$

such that

$$\frac{\partial G}{\partial x_i}(X(s)) = X_l(s)\frac{\partial F}{\partial x_i}(X(s)) + \delta_{il}F(X(s))\,,$$

with δ_{il} being Kronecker's delta. Furthermore, we derive that

$$\frac{\partial^2 G}{\partial x_i \partial x_j}(X(s)) = X_l(s)\frac{\partial^2 F}{\partial x_i \partial x_j}(X(s)) + (\delta_{il}+\delta_{jl})\frac{\partial F}{\partial x_l}(X(s))\,.$$

The above identities imply the formula (6.15).
By the Stone-Weierstrass-Theorem (cf. Klenke [2008], page 302) for any continuous function $F \in \mathcal{C}^2(\mathbb{R}^d,\mathbb{R})$ there exists a sequence of polynomials which uniformly converges to F. Continuity of the

stochastic integral thus ensures ucp-convergence. The general case for $F \in \mathcal{C}^{2,1}(\mathbb{R}\times\mathbb{R}_+,\mathbb{R})$ follows with continuity of F and the usual convergence to the Riemann integral. ■

We refer to Theorem II.32 on pages 78-79 of Protter [1992] for a more general version of Itô's lemma (resp. the Itô formula) including jumps.

6.9 Lévy's characterization of Brownian motion

As one important application of Itô's lemma we prove the following characterization of Brownian motion which is crucial in stochastic analysis.

Theorem 6.31. *A stochastic process $(X(t))_{t\geq 0}$, with $X(0)=0$, is a standard Brownian motion iff $(X(t))_{t\geq 0}$ is a continuous local martingale with quadratic variation $[X,X](t)=t$.*

Proof. If $(X(t))_{t\geq 0}$ is a standard Brownian motion, we already know the properties that it is a continuous local martingale and has quadratic variation $[X,X](t)=t$. Therefore, it suffices to prove the other implication. We consider a complex-valued transformation with i the imaginary unit. The Itô formula yields that

$$\begin{aligned}\exp\left(\mathrm{i}uX(t)+u^2t/2\right) &= e^{\mathrm{i}uX(0)}+\frac{u^2}{2}\int_0^t \exp\left(\mathrm{i}uX(s)+u^2s/2\right)ds+\mathrm{i}u\int_0^t \exp\left(\mathrm{i}uX(s)+u^2s/2\right)dX(s)\\ &\quad -\frac{u^2}{2}\int_0^t \exp\left(\mathrm{i}uX(s)+u^2s/2\right)d[X,X](s)\\ &= 1+\mathrm{i}u\int_0^t \exp\left(\mathrm{i}uX(s)+u^2s/2\right)dX(s)\,.\end{aligned}$$

In particular, if $(X(t))_{t\geq 0}$ is a local martingale, the process $\exp\left(\mathrm{i}uX(t)+u^2t/2\right)$ is as well a local martingale. Since for any fix t, it holds that

$$\sup_{s\leq t}\left|\exp\left(\mathrm{i}uX(s)+u^2s/2\right)\right| = \exp\left(u^2t/2\right)<\infty\,,$$

it is even a martingale. This yields that

$$\mathbb{E}\left[\exp\left(\mathrm{i}u(X(t)-X(s))\right)|\mathcal{F}_s\right] = \exp\left(-\frac{u^2}{2}(t-s)\right), \tag{6.16}$$

almost surely for $s\leq t$. $(\mathcal{F}_s)$ is a filtration with respect to which $X(s)$ is adapted, e.g. $\mathcal{F}_s=\mathcal{F}_s^X$. The function right-hand side in (6.16) is the characteristic function of $\mathcal{N}(0,t-s)$. With the next lemma this shows that

$$(X(t)-X(s))\sim\mathcal{N}(0,t-s)\,,\ \text{independent of } \mathcal{F}_s\,,$$

what completes the proof. ■

Lemma 6.32. *If $\mathbb{E}[\exp(\mathrm{i}tX)|\mathcal{F}]=\varphi(t)$, a.s. for some real-valued random variable X, a σ-field $\mathcal{F}$, and a characteristic function φ, then it holds that $X\sim\varphi$, and X is independent of $\mathcal{F}$.*

Proof. The distribution of X is characterized by its characteristic function $\varphi_X(t)=\mathbb{E}[\exp(\mathrm{i}tX)]$. Since by the tower rule, we have that

$$\varphi_X(t)=\mathbb{E}[\exp(\mathrm{i}tX)]=\mathbb{E}\left[\mathbb{E}[\exp(\mathrm{i}tX)|\mathcal{F}]\right]=\mathbb{E}\left[\varphi(t)\right]=\varphi(t)\,,$$

we obtain that $X \sim \varphi$. Note that for all $F \in \mathcal{F}$, it holds that

$$\int_F e^{\mathrm{i}tX}\, d\mathbb{P} = \int_F \varphi(t)\, d\mathbb{P} = \varphi(t)\,\mathbb{P}(F)\,.$$

For $Z = \mathbf{1}_F$, $F \in \mathcal{F}$, we derive for the joint distribution that

$$\begin{aligned}
\varphi_{X,Z}(t,u) &= \mathbb{E}\left[\exp(\mathrm{i}tX + \mathrm{i}uZ)\right] \\
&= e^{\mathrm{i}u} \int_F e^{\mathrm{i}tX}\, d\mathbb{P} + \int_{F^\complement} e^{\mathrm{i}tX}\, d\mathbb{P} \\
&= \varphi(t)\big(e^{\mathrm{i}u}\,\mathbb{P}(F) + \mathbb{P}(F^\complement)\big) \\
&= \varphi_X(t) \cdot \varphi_Z(u)\,.
\end{aligned}$$

This factorization yields the independence of X and $\mathcal{F}$. ■

7 Applications in Mathematical Finance

7.1 Girsanov's theorem

The Girsanov theorem is one of the main tools of stochastic analysis. In the theory of stochastic differential equations it often allows to extend results for a particular equation to other equations with more general drift coefficients. Moreover, the Girsanov theorem opens the door to the theory of equivalent martingale measures which is a main ingredient of mathematical finance. Abstractly seen, a Radon-Nikodym density for a new measure is obtained, under which the original process behaves differently. We only work in dimension one and start with a technical lemma on conditional Radon-Nikodym densities.

Lemma 7.1. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, $\mathcal{H} \subset \mathcal{F}$ be a sub-σ-field and $f \in L_1(\mathbb{P})$ be a density, that is nonnegative and integrating to one. Then a new probability measure $\mathbb{Q}$ on $\mathcal{F}$ is defined by $\mathbb{Q}(d\omega) = f(\omega)\,\mathbb{P}(d\omega)$ and for any $\mathcal{F}$-measurable random variable X with $\mathbb{E}_{\mathbb{Q}}[|X|] < \infty$ we obtain*

$$\mathbb{E}_{\mathbb{Q}}[X \,|\, \mathcal{H}]\, \mathbb{E}_{\mathbb{P}}[f \,|\, \mathcal{H}] = \mathbb{E}_{\mathbb{P}}[Xf \,|\, \mathcal{H}] \quad \mathbb{P}\text{-a.s.}$$

Remark. In the unconditional case we obviously have

$$\mathbb{E}_{\mathbb{Q}}[X] = \int X\, d\mathbb{Q} = \int Xf\, d\mathbb{P} = \mathbb{E}_{\mathbb{P}}[Xf].$$

Proof. We show that the left-hand side is a version of the conditional expectation on the right. Since it is obviously $\mathcal{H}$-measurable, it suffices to verify

$$\int_H \mathbb{E}_{\mathbb{Q}}[X \,|\, \mathcal{H}]\, \mathbb{E}_{\mathbb{P}}[f \,|\, \mathcal{H}]\, d\mathbb{P} = \int_H Xf\, d\mathbb{P} = \int_H X\, d\mathbb{Q} \quad \forall H \in \mathcal{H}.$$

By the projection property of conditional expectations we obtain

$$\mathbb{E}_{\mathbb{P}}[\mathbf{1}_H\, \mathbb{E}_{\mathbb{Q}}[X \,|\, \mathcal{H}]\, \mathbb{E}_{\mathbb{P}}[f \,|\, \mathcal{H}]] = \mathbb{E}_{\mathbb{P}}[\mathbf{1}_H\, \mathbb{E}_{\mathbb{Q}}[X \,|\, \mathcal{H}] f] = \mathbb{E}_{\mathbb{Q}}[\mathbf{1}_H\, \mathbb{E}_{\mathbb{Q}}[X \,|\, \mathcal{H}]] = \mathbb{E}_{\mathbb{Q}}[\mathbf{1}_H X],$$

which is the above identity. ■

Lemma 7.2. *Let $(\beta(t), 0 \le t \le T)$ be an $(\mathcal{F}_t)$-adapted process with $\beta \mathbf{1}_{t \le T} \in V^*$. Then*

$$M(t) := \exp\Big(-\int_0^t \beta(s)\, dW(s) - \frac{1}{2}\int_0^t \beta^2(s)\, ds\Big), \quad 0 \le t \le T,$$

is an $(\mathcal{F}_t)$-supermartingale. It is a martingale if and only if $\mathbb{E}[M(T)] = 1$ holds.

Proof. If we apply Itô's formula to M, we obtain

$$dM(t) = -\beta(t) M(t)\, dW(t), \quad 0 \le t \le T.$$

Hence, M is always a nonnegative local $\mathbb{P}$-martingale. By Proposition 4.22 we infer that M is a supermartingale and a proper martingale if and only if $\mathbb{E}_{\mathbb{P}}[M(T)] = \mathbb{E}_{\mathbb{P}}[M(0)] = 1$. ■

Lemma 7.3. *M is a martingale if β satisfies one of the following conditions (whereas the implications $1. \Rightarrow 2. \Rightarrow 3.$ hold):*

1. *β is uniformly bounded;*

2. *Novikov's condition:*
$$\mathbb{E}\Big[\exp\Big(\frac{1}{2}\int_0^T \beta^2(t)\,dt\Big)\Big] < \infty;$$
3. *Kazamaki's condition:*
$$\mathbb{E}\Big[\exp\Big(\frac{1}{2}\int_0^T \beta(t)\,dW(t)\Big)\Big] < \infty.$$

Proof. We prove $\beta M\mathbf{1}_{[0,T]} \in V$ by hand: If β is uniformly bounded by some $K > 0$, then we have for any $p > 0$ and any partition $0 = t_0 \le t_1 \le \cdots \le t_n = t$:

$$\begin{aligned}
&\mathbb{E}\Big[\exp\Big(p\sum_{i=1}^{n}\beta(t_{i-1})(W(t_i)-W(t_{i-1}))\Big)\Big]\\
&= \mathbb{E}\Big[\exp\Big(p\sum_{i=1}^{n-1}\beta(t_{i-1})(W(t_i)-W(t_{i-1}))\Big)\,\mathbb{E}[\exp(p\beta(t_{n-1})(W(t_n)-W(t_{n-1}))\,|\,\mathcal{F}_{t_{n-1}}]\Big]\\
&= \mathbb{E}\Big[\exp\Big(p\sum_{i=1}^{n-1}\beta(t_{i-1})(W(t_i)-W(t_{i-1}))\Big)\exp(p^2\beta(t_{n-1})^2(t_n-t_{n-1})\Big]\\
&\le \mathbb{E}\Big[\exp\Big(p\sum_{i=1}^{n-1}\beta(t_{i-1})(W(t_i)-W(t_{i-1}))\Big)\exp(p^2K^2(t_n-t_{n-1})\Big]\\
&\le \exp\Big(\sum_{i=1}^{n}p^2K^2(t_i-t_{i-1})\Big)\\
&= \exp(p^2K^2t).
\end{aligned}$$

This shows that the random variables $\exp\Big(\sum_{i=1}^n \beta(t_{i-1})(W(t_i)-W(t_{i-1}))\Big)$ are uniformly bounded in any $L^p(\mathbb{P})$-space and thus uniformly integrable. Since by taking finer partitions these random variables converge to $\exp(\int_0^t \beta(s)\,dW(s))$ in $\mathbb{P}$-probability, we infer that $M(t)$ has finite expectation and even moments of all orders. Consequently, $\int_0^T \mathbb{E}[(\beta(t)M(t))^2]\,dt$ is finite and M is a martingale.

For the sufficiency of Novikov's and Kazamaki's condition we refer to Liptser and Shiryaev [2001] and the references therein. ■

Theorem 7.4. *Let $(X(t), 0 \le t \le T)$ be a stochastic (Itô) process on $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ satisfying*
$$X(t) = \int_0^t \beta(s)\,ds + W(t), \quad 0 \le t \le T,$$
with a Brownian motion W and a process $\beta\mathbf{1}_{t\le T} \in V^*$. If β is such that M is a martingale, then $(X(t), 0 \le t \le T)$ is a Brownian motion under the measure $\mathbb{Q}$ on $(\Omega, \mathcal{F}, (\mathcal{F}_t))$ defined by $\mathbb{Q}(d\omega) = M(T,\omega)\,\mathbb{P}(d\omega)$.

Proof. We use Lévy's characterization of Brownian motion. Since M is a martingale, $M(T)$ is a density and $\mathbb{Q}$ is well-defined.

We put $Z(t) = M(t)X(t)$ and obtain by Itô's formula (or partial integration) that

$$\begin{aligned}
dZ(t) &= M(t)\,dX(t) + X(t)\,dM(t) + d[M,X]_t\\
&= M(t)\Big(\beta(t)\,dt + dW(t) - X(t)\beta(t)\,dW(t) - \beta(t)dt\Big)\\
&= M(t)(1 - X(t)\beta(t))\,dW(t).
\end{aligned}$$

This shows that Z is a local martingale. If Z is a martingale, then we accomplish the proof using the preceding lemma:

$$\mathbb{E}_{\mathbb{Q}}[X(t) \mid \mathcal{F}_s] = \frac{\mathbb{E}_{\mathbb{P}}[M(t)X(t) \mid \mathcal{F}_s]}{\mathbb{E}_{\mathbb{P}}[M(t) \mid \mathcal{F}_s]} = \frac{Z(s)}{M(s)} = X(s), \quad s \le t,$$

implies that X is a $\mathbb{Q}$-martingale which by its very definition has quadratic variation t. Hence, X is a Brownian motion under $\mathbb{Q}$.

If Z is only a local martingale with associated stopping times (τ_n), then the above relation holds for the stopped processes $X^{\tau_n}(t) = X(t \wedge \tau_n)$, which shows that X is a local $\mathbb{Q}$-martingale, and we can apply Lévy's characterization of Brownian motion. ∎

7.2 Stochastic market model

We denote by S_t^j, $j = 1, \dots, d, t \in [0,T]$, d stock price processes. One (usually risk-less) financial asset S_t^0 is used as a benchmark such that the other prices are expressed relative to this so-called numéraire. The discounted price processes are $\tilde{S}_t^j = S_t^j / S_t^0$, where we assume $S_t^0 > 0$ for all $t \ge 0$.

Definition 7.5. *A (time-continuous)* ***market*** *(model) is a $(d+1)$-dimensional stochastic process $(S_t)_{t\in[0,T]}$ on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$, where $(\mathcal{F}_t)$ satisfies the usual conditions. It holds that $\inf_{0\le t\le T} S_t^0 > 0$ $\mathbb{P}$-almost surely. The market is normalized if $S_t^0 \equiv 1$. A* ***trading strategy*** *or portfolio is a predictable $(d+1)$-dimensional process (ϑ_t), $\vartheta^j \in \mathbb{L}, 0 \le j \le d$. The value (process) of a portfolio (ϑ_t) is given by*

$$V_t(\vartheta) = \langle \vartheta_t, S_t \rangle = \sum_{j=0}^{d} \vartheta_t^j S_t^j .$$

The portfolio $\vartheta = (\vartheta(t))_{t\in[0,T]}$ is called ***self-financing*** *if for all $t \in [0,T]$:*

$$V_t(\vartheta) = V_0(\vartheta) + \int_0^t \sum_{j=0}^{d} \vartheta_s^j \, dS_s^j .$$

Restricting to continuous processes for price models, we only require that (ϑ_t) is adapted.

Lemma 7.6. *Given that ϑ is self-financing, we have that*

$$d\tilde{V}(\vartheta) = \sum_{j=1}^{d} \vartheta_-^j \, d\tilde{S}^j = \sum_{j=1}^{d} \vartheta^j \, d\tilde{S}^j .$$

Hence ϑ is also self-financing for the normalized market.

Proof. We apply the partial integration formula from Corollary 6.25 :

$$\begin{aligned}
d\tilde{V}(\vartheta) &= d\big((S^0)^{-1} V(\vartheta)\big) \\
&= (S_-^0)^{-1} dV(\vartheta) + V(\vartheta)_- d(S^0)^{-1} + d[V(\vartheta), (S^0)^{-1}] \\
&= (S_-^0)^{-1} \langle \vartheta, dS \rangle + \langle \vartheta, S_- \rangle d(S^0)^{-1} + \langle \vartheta, d[S, (S^0)^{-1}] \rangle \\
&= \langle \vartheta, (S_-^0)^{-1} dS + S_- d(S^0)^{-1} + d[S, (S^0)^{-1}] \rangle \\
&= \langle \vartheta, d\big((S^0)^{-1} S\big) \rangle = \langle \vartheta, d\tilde{S} \rangle .
\end{aligned}$$

∎

Definition 7.7. *The trading strategy ϑ is called **admissible**, when $\tilde{V}_t(\vartheta)$ is bounded from below, $\tilde{V}_t(\vartheta) \geq -K, 0 \leq t \leq T$, $\mathbb{P}$-almost surely. A self-financing trading strategy ϑ is called an **arbitrage**, if it is admissible and*

$$V_0(\vartheta) \leq 0, V_T(\vartheta) \geq 0 \ \mathbb{P}-a.s.\ , \mathbb{P}\left(V_T(\vartheta) > 0\right) > 0.$$

Theorem 7.8 (Fundamental theorem of asset pricing)**.** *If there exists an equivalent measure $\mathbb{Q} \sim \mathbb{P}$, such that $\tilde{S}$ is a square-integrable local martingale with respect to $\mathbb{Q}$, then the market has no arbitrage.*

Proof. Apparently, $\tilde{V}_t(\vartheta) = \tilde{V}_0(\vartheta) + \int_0^t \sum_{j=1}^d \vartheta_s^j \, d\tilde{S}_s^j$ is a local martingale. Since $\tilde{V}_t(\vartheta) \geq -K$, it is by Proposition 4.22 a supermartingale. Hence,

$$\mathbb{E}_{\mathbb{Q}}\left[\tilde{V}_T(\vartheta)\right] \leq \mathbb{E}_{\mathbb{Q}}\left[\tilde{V}_0(\vartheta)\right] \leq 0.$$

Thus, if $\tilde{V}_T(\vartheta) \geq 0$ $\mathbb{P}$-a.s., we conclude that $\tilde{V}_T(\vartheta) \geq 0$ $\mathbb{Q}$-a.s. The above requirements imply that $\mathbb{Q}\left(\tilde{V}_T(\vartheta) > 0\right) = 0$, which implies $\mathbb{P}\left(\tilde{V}_T(\vartheta) > 0\right) = 0$. We obtain the result that the market has no arbitrage. ∎

The measure $\mathbb{Q}$ is called an equivalent (local) martingale measure. Theorem 7.8 states the relevant direction that if an equivalent local martingale measure exists, then the market has no arbitrage. It can be shown that the market satisfies an even stronger condition of 'no free lunch with vanishing risk (NFLVR)'. Moreover, an equivalence holds true. If the market satisfies the NFLVR condition, then there exists an equivalent martingale measure, see Delbaen and Schachermayer [1994]. This result asks for a much deeper proof and is nowadays often referred to as the fundamental theorem of asset pricing.

Consider in the sequel the market model

$$\begin{aligned} dS_t^j &= b_t^j \, dt + \sum_{i=1}^d \sigma_t^{ji} \, dB_t^i \\ &= b_t^j \, dt + \langle \sigma_t^j \, dB_t \rangle \end{aligned}$$

with $dS_t^0 = \rho_t S_t^0 \, dt$, $S_0^0 = 1$, and with a $(d \times d)$ adapted matrix process σ and a d-dimensional standard Brownian motion B. Based on Theorem 7.4, we obtain that if a d-dimensional process u_t exists such that

$$\sigma_t u_t = b_t - \rho_t S_t^1 \text{ and } \mathbb{E}\left[\exp\left(\frac{1}{2}\int_0^T \langle u_t, u_t \rangle \, dt\right)\right] < \infty,$$

the market has no arbitrage.

Proposition 7.9. *Let $\vartheta^1, \ldots, \vartheta^j$ be chosen. Choosing a specific ϑ^0, one can always make the portfolio self-financing. In particular, this is possible for any $V_0(\vartheta)$.*

Proof. By virtue of the the self-financing property, it holds that

$$\begin{aligned} V_t(\vartheta) = \langle \vartheta_t, S_t \rangle = \vartheta_t^0 S_t^0 + \sum_{i=1}^d \vartheta_t^i S_t^i \\ = V_0(\vartheta) + \int_0^t \langle \vartheta, dS \rangle \\ = V_0(\vartheta) + \int_0^t \vartheta_s^0 \, dS_s^0 + \sum_{i=1}^d \int_0^t \vartheta_s^i \, dS_s^i. \end{aligned}$$

Thereby, we derive that

$$d(\vartheta_t^0 S_t^0) = \vartheta_t^0\, dS_t^0 + dA_t \text{ with } A_t = \sum_{i=1}^{d} \Big(\int_0^t \vartheta_s^i\, dS_s^i - \vartheta_t^i S_t^i \Big).$$

Since $S_t^0 = \exp\big(\int_0^t \rho_s\, ds\big)$,

$$d(\vartheta_t^0 S_t^0) = \rho_t \big(\vartheta_t^0 S_t^0\big)\, dt + dA_t\,.$$

This affine SDE admits by Itô's formula (or partial integration) the solution

$$\vartheta_t^0 S_t^0 = S_t^0 \vartheta_0^0 + S_t^0 \int_0^t \big(S_s^0\big)^{-1} dA_s\,.$$

Dividing by S_t^0 and applying partial integration yields

$$\begin{aligned}
\vartheta_t^0 &= \vartheta_0^0 + \int_0^t \big(S_s^0\big)^{-1} dA_s \\
&= \vartheta_0^0 + \big(S_t^0\big)^{-1} A_t - A_0 - \int_0^t A_s\, d\big(S_s^0\big)^{-1} \\
&= V_0(\vartheta) + \big(S_t^0\big)^{-1} A_t + \int_0^t \rho_s \big(S_s^0\big)^{-1} A_s\, ds\,.
\end{aligned} \tag{7.1}$$

Choosing ϑ^0 according to (7.1), the portfolio is self-financing. ■

7.3 Applications for the Black-Scholes model

Definition 7.10. *A (contingent) T-**claim** is a lower bounded $\mathcal{F}_T$-measurable random variable H_T. We say that the claim H_T is attainable if there exists an admissible self-financing strategy ϑ such that $V_T(\vartheta) = H_T$.*

Important examples of contingent claims are options and futures. The European call option gives the owner the right (but not the obligation) to buy one stock at the specified price K (strike price) at maturity time T. If the price of this stock satisfies $S_T^j(\omega) > K$, then the owner of the option obtains the payoff $S_T^j(\omega) - K$ at maturity time T, while $S_T^j(\omega) \le K$ results in the payoff 0 and the owner will not exercise his option in this case. Hence, the payoff is $C_T = (S_T^j(\omega) - K)^+$. The European put option gives the owner the right (but not the obligation) to sell one stock at strike price K at time T. This option has the payoff $P_T = (K - S_T^j(\omega))^+$.

More complicated financial instruments include American options, where the option can be executed up to time T at a fix strike price. Exotic options as barrier or Asian options depend on the path $(S_t^j)_{t\in[0,T]}$.

Definition 7.11. *For an attainable T-claim H_T, the smallest value of admissible self-financing strategies ϑ with $V_T(\vartheta) = H_T$,*

$$\pi_t(H_T) = \operatorname{essinf}_\vartheta V_t(\vartheta)$$

*is called **hedge price** (also fair price) of H_T at time t.*

Proposition 7.12. *Let H_T be a T-claim and $\mathbb{Q}$ an equivalent martingale measure. If there exists an attainable strategy ϑ for H_T, such that $\tilde{V}_t(\vartheta)$ is a $\mathbb{Q}$-martingale, then*

$$\pi_t(\tilde{H}_T) = \mathbb{E}_{\mathbb{Q}}\big[\tilde{H}_T | \mathcal{F}_t\big]\,.$$

*Such a portfolio/strategy is called a **replicating or hedging** portfolio/strategy for H_T.*

Proof. $\tilde{S}$ is a $\mathbb{Q}$-local martingale. We have that

$$\mathbb{E}_{\mathbb{Q}}\left[\tilde{H}_T|\mathcal{F}_t\right] = \mathbb{E}_{\mathbb{Q}}\left[\tilde{V}_T(\vartheta)|\mathcal{F}_t\right] \le \tilde{V}_t(\vartheta)\,.$$

Thus, $\mathbb{E}_{\mathbb{Q}}\left[\tilde{H}_T|\mathcal{F}_t\right] \le \pi_t(\tilde{H}_T)$. If $\tilde{V}(\vartheta)$ is a $\mathbb{Q}$-martingale, then

$$\tilde{V}_t(\vartheta) = \mathbb{E}_{\mathbb{Q}}\left[\tilde{H}_T|\mathcal{F}_t\right].$$

∎

If $\mathbb{Q}$ is not unique $\pi(\tilde{H}_T)$ is independent of $\mathbb{Q}$. If any T-claim $\tilde{H}_T \le 1$ has a replicating strategy, it follows that $\mathbb{Q}$ is unique on $\mathcal{F}_T$ (set $\tilde{H}_T = \mathbf{1}_A$ for $A \in \mathcal{F}_T$). The market is called complete if for every bounded T-claim there exists a hedging strategy. If $\mathbb{Q}$ is unique the market is complete. The last assertion, however, is not easy to establish in general, see, for instance, Chapter 6 of Duffie [2010]. Instead, we establish at the end of this section completeness of the Black-Scholes model directly.

In the sequel, we focus on the Black-Scholes model with two assets S^0 and S^1, where

$$dS_t^1 = S_t^1\mu\,dt + S_t^1\sigma\,dB_t \tag{7.2a}$$

with a one-dimensional standard Brownian motion B. S_t^0 is a risk-free asset with interest rate r such that

$$S_t^0 = \exp(rt)\;,\; S_t^1 = S_0^1\exp\left(\sigma B_t + \left(\mu t - \frac{\sigma^2 t}{2}\right)\right), 0 \le t \le T\,. \tag{7.2b}$$

Proposition 7.13. *In the Black-Scholes model, (7.2b), $\mathbb{Q}$ is unique and*

$$d\mathbb{Q} = \exp\left(\frac{r-\mu}{\sigma}B_T - \frac{1}{2}\left(\frac{r-\mu}{\sigma}\right)^2 T\right)d\mathbb{P}\,. \tag{7.3}$$

Proof. We illustrate $\tilde{S}_t^1$ as $\mathbb{Q}$-local martingale:

$$\tilde{S}_t^1 = S_0^1 + \int_0^t h_s\,dW_s$$

with $\mathbb{Q}$-Brownian motion W. If such an illustration exists, $\tilde{S}_t^1$ is a $\mathbb{Q}$-local martingale. We deduce with partial integration that

$$\begin{aligned} d\tilde{S}_t^1 &= d\left(e^{-rt}S_t^1\right) = S_t^1\,d\left(e^{-rt}\right) + e^{-rt}\,dS_t^1 + d[e^{-rt}, S_t^1] \\ &= -rS_t^1e^{-rt}\,dt + e^{-rt}S_t^1\mu\,dt + e^{-rt}S_t^1\sigma\,dB_t \\ &= \tilde{S}_t^1\left((\mu-r)\,dt + \sigma\,dB_t\right) \\ &= \sigma\tilde{S}_t^1\,dW_t \end{aligned} \tag{7.4}$$

with $W_t = \sigma^{-1}(\mu-r)t + B_t$. Theorem 7.4 grants that W is a $\mathbb{Q}$-Brownian motion with

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = M(T) = \exp\left(\frac{r-\mu}{\sigma}B_T - \frac{(r-\mu)^2}{2\sigma^2}T\right)$$

if $M(t) = \mathbb{E}_{\mathbb{P}}\left[M(T)|\mathcal{F}_t\right]$ is a martingale. This is ensured by the Novikov criterion. The solution of (7.4) is

$$\tilde{S}_t^1 = S_0^1\exp\left(\sigma W_t - \frac{1}{2}\sigma^2 t\right).$$

$\tilde{S}_t^1$ is a martingale. ∎

Corollary 7.14. The Black-Scholes model has no arbitrage.

Lemma 7.15. *In the Black-Scholes model it holds that $\pi_t(H_T) = V_t(\vartheta)$ for a T-claim $H_T = f(S_T^1)$, $f \geq 0$, $\mathbb{E}_{\mathbb{Q}}[H_T^2] < \infty$, with $V_t(\vartheta) = F(t, S_t^1)$ and*

$$F(t,x) = e^{-r(T-t)} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f\big(x \exp\big((r - \frac{\sigma^2}{2})(T-t) + \sigma u (T-t)^{1/2}\big)\big) e^{-u^2/2} \, du\,, \tag{7.5}$$

where $0 \leq t \leq T$ and $x > 0$.

Proof. We have that

$$\begin{aligned} S_T^1 &= \exp\big(\sigma B_T + (\mu - \frac{\sigma^2}{2}) T\big) \\ &= \exp\big(\sigma W_T + (r - \frac{\sigma^2}{2}) T\big) \\ &= e^{\sigma W_t + \left(r - \frac{\sigma^2}{2}\right) t} e^{\sigma (W_T - W_t) + \left(r - \frac{\sigma^2}{2}\right)(T-t)} \\ &= S_t^1 e^{\sigma (W_T - W_t) + \left(r - \frac{\sigma^2}{2}\right)(T-t)}\,. \end{aligned}$$

$\tilde{V}(\vartheta)$ is a $\mathbb{Q}$-martingale for $H_T = f(S_T^1)$, when

$$\mathbb{E}_{\mathbb{Q}}\left[f(\tilde{S}_T^1)|\mathcal{F}_t\right] = \mathbb{E}_{\mathbb{Q}}\left[e^{-rT} f(S_T^1)|\mathcal{F}_t\right] = \tilde{V}_t(\vartheta) = e^{-rt} V_t(\vartheta)\,.$$

We derive that

$$V_t(\vartheta) = \mathbb{E}_{\mathbb{Q}}\left[e^{-r(T-t)} f(S_T^1)|\mathcal{F}_t\right]$$

and that $V_t(\vartheta) = F(t, S_t^1)$ with

$$F(t,x) = \mathbb{E}_{\mathbb{Q}}\left[e^{-r(T-t)} f\big(x \exp\big((r - \frac{\sigma^2}{2})(T-t) + \sigma(W_T - W_t)\big)\big)|\mathcal{F}_t\right],$$

since $W_T - W_t$ is independent from $\mathcal{F}_t$. Using $W_T - W_t \sim \mathcal{N}(0, T-t)$ yields the result. ∎

In the following the cdf of the standard normal distribution is denoted by Φ.

Theorem 7.16 (Black-Scholes formula)**.** *In the Black-Scholes model, the hedge price of a European call-option with strike price K is given by*

$$C^{BS}(t, S_t^1) = S_t^1 \Phi(d_+) - K e^{-r(T-t)} \Phi(d_-)\,, \tag{7.6a}$$

$$\text{with } d_\pm = \frac{\log(S_t^1/K) + \left(r \pm \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}}\,. \tag{7.6b}$$

Proof. With $f(x) = (x-K)^+$ it holds that $H_T = f(S_T^1)$. Applying Lemma 7.15 yields

$$C^{BS}(t,x) = \mathbb{E}\left[\left(x \exp\left(-\frac{\sigma^2}{2}(T-t) + \sigma\sqrt{T-t}\, Z\right) - K e^{-r(T-t)}\right)^+\right],$$

with $Z \sim \mathcal{N}(0,1)$ a standard normal random variable. We rewrite this

$$C^{BS}(t,x) = \mathbb{E}\left[\left(x \exp\left(-\frac{\nu^2}{2} + \nu Z\right) - k\right)^+\right],$$

with $k = K e^{-r(T-t)}$ and $\nu = \sigma\sqrt{T-t}$. The integrand above is positive if

$$Z \geq \nu^{-1}\big(-\log(x/k) + \nu^2/2\big) = -d_-\,,$$

such that

$$C^{BS}(t,x) = \frac{x}{\sqrt{2\pi}} \int_{-d_-}^{\infty} \exp\Big(-\frac{\nu^2}{2} + \nu u - u^2/2\Big)\,du - \frac{k}{\sqrt{2\pi}} \int_{-d_-}^{\infty} e^{-u^2/2}\,du$$
$$= x\big(1-\Phi(-d_- - \nu)\big) - k\Phi(d_-) = x\Phi(d_+) - k\Phi(d_-)\,,$$

with

$$d_- = \frac{\log(x/k) - \frac{\nu^2}{2}}{\nu}\,,\ d_+ = d_- + \nu\,.$$ ■

In the Black-Scholes model, the hedge price of a European put-option with strike price K is determined by Theorem 7.16 and the put-call parity

$$C^{BS}(t,S_t^1) - P^{BS}(t,S_t^1) = S_t^1 - Ke^{-r(T-t)}\,.$$

Proposition 7.17 (Δ-hedge). *Let $H_T = f(S_T^1)$ be a T-claim with $\mathbb{E}_{\mathbb{Q}}[H_T^2] < \infty$ and $V_t(\vartheta) = F(t,S_t^1)$. If $F(t,x)$ is twice continuously differentiable with respect to x, then*

$$\vartheta_t^0 = e^{-rt}\big(F(t,S_t^1) - \vartheta_t^1 S_t^1\big)\,,$$
$$\vartheta_t^1 = \frac{\partial F}{\partial x}(t,S_t^1)$$

is a hedging strategy with continuous paths. $\Delta = \frac{\partial F}{\partial x} = \frac{\partial V}{\partial S^1}$ is called the option's delta and the above hedging strategy Δ-hedge.

Proof. It holds that

$$\tilde{V}_t(\vartheta) = \tilde{F}(t,\tilde{S}_t^1) = e^{-rt}F\big(t,\tilde{S}_t^1 e^{rt}\big) = e^{-rt}F(t,S_t^1)\,.$$

An application of Itô's formula gives

$$d\tilde{V}_t(\vartheta) = \Big(\frac{\partial \tilde{F}}{\partial t}(t,\tilde{S}_t^1) + \frac{1}{2}\frac{\partial^2 \tilde{F}}{\partial y^2}(t,\tilde{S}_t^1)\sigma^2(\tilde{S}_t^1)^2\Big)dt + \frac{\partial \tilde{F}}{\partial y}(t,\tilde{S}_t^1)\,d\tilde{S}_t^1\,.$$

Since $\tilde{V}(\vartheta)$ is a $\mathbb{Q}$-martingale and the second addend of the right-hand side is as well a local $\mathbb{Q}$-martingale, the integral over the first addend vanishes:

$$\int_0^t \Big(\frac{\partial \tilde{F}}{\partial s}(s,\tilde{S}_s^1) + \frac{1}{2}\frac{\partial^2 \tilde{F}}{\partial y^2}(s,\tilde{S}_s^1)\sigma^2(\tilde{S}_s^1)^2\Big)ds = 0\,.$$

Hence, we conclude that

$$d\tilde{V}_t(\vartheta) = \frac{\partial \tilde{F}}{\partial y}(t,\tilde{S}_t^1)\,d\tilde{S}_t^1\,.$$

Variables y and x are linked by $y = e^{-rt}x$. Putting

$$\vartheta_t^1 = \frac{\partial \tilde{F}}{\partial y}(s,\tilde{S}_t^1) = \frac{\partial F}{\partial x}(t,S_t^1)$$

and $\vartheta_t^0 = \tilde{F}(t,\tilde{S}_t^1) - \vartheta_t^1\tilde{S}_t^1$, ϑ is self-financing with $\tilde{V}_t(\vartheta) = \tilde{F}(t,\tilde{S}_t^1)$ and thus $V_T(\vartheta) = f(S_T^1)$. ■

Corollary 7.18 (Black-Scholes PDE)**.** Let $H_T = f(S_T^1)$ and ϑ be the Δ-hedge. Then $V_t(\vartheta) = F(t, S_t^1)$ with F satisfying the partial differential equation

$$\frac{\partial F}{\partial t} + \frac{1}{2}\sigma^2 x^2 \frac{\partial^2 F}{\partial x^2} + rx\frac{\partial F}{\partial x} = rF$$

and boundary condition $F(T,x) = f(x)$.

Proof. The preceding proposition implies that

$$\frac{\partial \tilde{F}}{\partial t} + \frac{1}{2}\sigma^2 y^2 \frac{\partial^2 \tilde{F}}{\partial y^2} = 0\,.$$

Inserting $\tilde{F} = e^{-rt}F(t, xe^{rt})$ and carefully considering the variable transform $y = e^{-rt}x$ yields the result. ■

Proposition 7.19. *Let $a > 0$. In the Black-Scholes model there exists a self-financing strategy ϑ such that $V_0(\vartheta) = 0$ and $\tilde{V}_T(\vartheta) = a$. This ϑ is not admissible.*

Proof. Set

$$Z_t^1 = \frac{1}{\sigma \tilde{S}_t^1 \sqrt{T-t}}\ ,\ 0 \le t < T\,.$$

Choose Z_t^0 such that $Z = (Z_t^0, Z_t^1)$ is self-financing and $V_0(\vartheta) = 0$. This is possible according to Proposition 7.9. $\tilde{S}_t^1$ is a $\mathbb{Q}$-local martingale and by (7.4):

$$d\tilde{S}_t^1 = \sigma \tilde{S}_t^1\, dW_t\,,$$

with $\mathbb{Q}$-Brownian motion W. Therefore

$$d\tilde{V}_t(Z) = Z_t^1\, d\tilde{S}_t^1 = \frac{1}{\sqrt{T-t}}\, dW_t\,.$$

Since $\tilde{V}_t(Z) = Z_t^0 + Z_t^1 \tilde{S}_t^1$, we obtain that

$$Z_t^0 = \int_0^t \frac{dW}{\sqrt{T-s}} - \frac{1}{\sigma\sqrt{T-t}}\ ,\ 0 \le t < T\,.$$

Z is self-financing. $\tilde{V}_t(Z)$ is a Gaussian process with

$$\mathrm{Cov}_{\mathbb{Q}}\left(\tilde{V}_s(Z), \tilde{V}_t(Z)\right) = \int_0^{t\wedge s} \frac{1}{T-u}\, du = -\log\left(1 - \frac{s\wedge t}{T}\right).$$

The time-shifted process

$$\tilde{W}_t = \tilde{V}_{T(1-e^{-t})}(Z)$$

is a Brownian motion, since it is a centered Gaussian process with

$$\mathrm{Cov}_{\mathbb{Q}}\left(\tilde{W}_s, \tilde{W}_t\right) = t \wedge s\,.$$

Define the stopping times

$$T_1 = \inf\left\{s \,|\, \tilde{V}_s(Z) = a\right\},$$
$$T_2 = \inf\left\{s \,|\, \tilde{W}_s = a\right\}.$$

They are related by $T_1 = T(1 - e^{-T_2})$. Since $\tilde{W}$ is a $\mathbb{Q}$-Brownian motion:

$$\mathbb{Q}(T_2 < \infty) = 1 \;\Rightarrow\; \mathbb{Q}(T_1 < T) = 1 \;\Rightarrow\; \mathbb{P}(T_1 < T) = 1\,.$$

The strategy that obeys $V_0(\vartheta) = 0$ and $\tilde{V}_T(\vartheta) = a$ is now defined as

$$\vartheta_t^0 = \begin{cases} Z_t^0 & , 0 \le t \le T_1 \\ a & , T_1 < t \le T \end{cases}, \quad \vartheta_t^1 = \begin{cases} Z_t^1 & , 0 \le t \le T_1 \\ 0 & , T_1 < t \le T \end{cases}.$$

Note however that $\tilde{V}(\vartheta)$ is not bounded from below. ∎

Proposition 7.20. *The Black-Scholes market model is complete.*

Proof.

$$M(t) = \mathbb{E}_{\mathbb{Q}}\left[\tilde{H}_T | \mathcal{F}_t\right] \ge 0$$

is a $\mathbb{Q}$-martingale. By the representation theorem for continuous local martingales [Kallenberg, 2002, Thm. 18.12] there exists a Brownian motion W such that

$$M(t) = M(0) + \int_0^t h(s)\, dW(s)$$

holds for all $t \ge 0$ $\mathbb{Q}$-almost surely. We derive that

$$dM(t) = h(t)\, dW(t) = h(t)\left(\sigma \tilde{S}_t^1\right)^{-1} d\tilde{S}_t^1$$

such that

$$\vartheta_t^1 = h(t)\left(\sigma \tilde{S}_t^1\right)^{-1}, \quad \vartheta_t^0 = M(t) - \vartheta_t^1 \tilde{S}_t^1$$

satisfies $\tilde{V}_t(\vartheta) = M(t) \ge 0$ and

$$d\tilde{V}_t(\vartheta) = dM(t) = \vartheta_t^1\, d\tilde{S}_t^1 .$$

such that ϑ is a hedge. ∎

Bibliography

Markus Bibinger and Hajo Holzmann. *Wahrscheinlichkeitstheorie WS 2018/19.* Lecture Notes, 2019.

Freddy Delbaen and Walter Schachermayer. A general version of the fundamental theorem of asset pricing. *Mathematische Annalen*, 300(1):463–520, 1994.

R. M. Dudley. *Probabilities and metrics*. Matematisk Institut, Aarhus Universitet, Aarhus, 1976. Convergence of laws on metric spaces, with a view to statistical testing, Lecture Notes Series, No. 45.

D. Duffie. *Dynamic Asset Pricing Theory*. Princeton Series in Finance. Princeton University Press, 2010. ISBN 9781400829200.

Michael Falk. Stochastik 1. Vorlesungsskript, 2019.

Olav Kallenberg. *Foundations of modern probability. 2nd ed.* Probability and Its Applications. New York, Springer., 2002.

Ioannis Karatzas and Steven E. Shreve. *Brownian motion and stochastic calculus. 2nd ed.* Springer, New York, 1991.

Achim Klenke. *Wahrscheinlichkeitstheorie*. Berlin: Springer, 2nd revised ed. edition, 2008. ISBN 978-3-540-76317-8/pbk; 978-3-540-77571-3/ebook.

Robert S. Liptser and Albert N. Shiryaev. *Statistics of random processes. 1: General theory. 2nd ed.* Springer, Berlin, 2001.

Philip Protter. *Stochastic integration and differential equations. A new approach. 2nd corr. printing.* Springer, Berlin, 1992.

Daniel Revuz and Marc Yor. *Continuous martingales and Brownian motion. 3rd ed.* Springer, Berlin, 1999.

René L. Schilling and Lothar Partzsch. *Brownian motion. An introduction to stochastic processes. With a chapter on simulation by Björn Böttcher*. Berlin: de Gruyter, 2012. ISBN 978-3-11-027889-7/pbk; 978-3-11-027898-9/ebook.