Stochastic Calculus

David Prömel*
Universität Mannheim

August 17, 2023

Abstract

This are the lecture notes for the course "Stochastic Calculus". Please be careful when using it and let me know if you find mistakes or have suggestions for improvements.

*The present lecture notes are written jointly with Mathias Trabs, Karlsruher Institut für Technologie.

Contents

Recommended literature

- Kuo, H.-H., *Introduction to Stochastic Integration*, Springer, 2006.
  [Main recommendation - very accessible textbook on a similar level as the lecture.]
- Steele, J.M., *Stochastic Calculus and Financial Applications*, Springer, 2001.
  [Main recommendation - nice textbook on a similar level as the lecture.]
- Karatzas, I., and Shreve S.E. *Brownian Motion and Stochastic Calculus*, Springer, 1998.
  [Classical textbook also covering the material of this lecture course.]
- Protter, P.E., *Stochastic Integration and Differential Equations*, Springer, 2005.
  [More general classical textbook about stochastic calculus.]
- Revuz, D., and Yor, M., *Continuous Martingales and Brownian Motion*, Springer, 1999.
  [More general classical textbook about stochastic calculus.]
- Øksendal, B., Stochastic Differential Equations, Springer, 2003.
  [Good textbook treating further applications of stochastic calculus.]
- Klenke, A., *Probability Theory.* Springer-Verlag, 2006.
  [Very good textbook for the necessary foundation in probability theory. There is a German version of the book called "*Wahrscheinlichkeitstheorie*".]

1 Introduction

Lecture 1

Stochastic calculus provides the mathematical theory required for probabilistic modeling of real-world phenomena in continuous time, as they appear in various areas like mathematical finance, engineering and physics. Let us start by briefly discussing, on a heuristic level, two areas of applications where stochastic calculus is naturally required and which may serve us as motivation to develop it. Of course, there is a long list of further applications of stochastic calculus like stochastic control, stochastic filtering and optimal stopping, just to name a few.

Application 1: mathematical finance in continuous time. Let us consider a very simple financial market consisting of a risky asset and a risk-free asset ("bank account"), which both can be traded in continuous time. To capture the unpredictability and randomness of future prices, the price process is modeled by a one-dimensional stochastic process $(S_t)_{t\in[0,T]}$ and the price evolution of the risk-free asset $(B_t)_{t\in[0,T]}$ is given $B_t := 1$ for $t \in [0,T]$, i.e. we assume that the interest rate is $r = 0$.

For simplicity we restrict trading on this financial market to self-financing trading strategies $\varphi = (\varphi_t^0, \varphi_t^1)_{t\in[0,T]}$ of the form:

- $\varphi_t^1 = f(S_t)$, for some $f \in C(\mathbb{R};\mathbb{R})$, stands for the numbers of shares of risky assets hold at time t,
- φ_t^0 stands for the numbers of risk-free assets held at time t, chosen such that $\varphi = (\varphi_t^0, \varphi_t^1)_{t\in[0,T]}$ is self-financing.

Hence, the capital process $(V_t(\varphi))_{t\in[0,T]}$ generated by trading according to $\varphi = (\varphi_t^0, \varphi_t^1)_{t\in[0,T]}$ satisfies

$$\begin{aligned} V_t(\varphi) &= \int_0^t \varphi_s^0 \,\mathrm{d}B_s + \int_0^t \varphi_s^1 \,\mathrm{d}S_s \qquad \left[\approx \sum_{i=0}^{N-1} \varphi_{t_i}^0 (B_{t_{i+1}} - B_{t_i}) + \sum_{i=0}^{N-1} \varphi_{t_i}^1 (S_{t_{i+1}} - S_{t_i})\right] \\ &= \int_0^t f(S_s)\,\mathrm{d}S_s, \end{aligned}$$

for $0 \le t_0 \le \dots \le t_N \le T$ and $N \in \mathbb{N}$.

To keep our life simple, let us choose the stochastic process $(S_t)_{t\in[0,T]}$ with sufficiently smooth sample paths so that classical analysis can be applied to the sample paths of the stochastic process $(S_t)_{t\in[0,T]}$. In particular, this would allow us to use the fundamental theorem of calculus, i.e.

$$F(S_T) - F(S_0) = \int_0^T f(S_s) S_s' \,\mathrm{d}s = \int_0^T f(S_s)\,\mathrm{d}S_s$$

for any function $F\colon \mathbb{R} \to \mathbb{R}$ such that $F'(x) = f(x) \in C(\mathbb{R};\mathbb{R})$, and the integral $\int_0^T f(S_s)\,\mathrm{d}S_s$ is well-defined.

Now, we can take, for instance, the trading strategy $\varphi = (\varphi_t^0, \varphi_t^1)_{t\in[0,T]}$ with

$$\varphi_t^1 := 2(S_t - S_0), \quad \text{i.e.} \quad f(x) = 2(x - S_0) \quad \text{and} \quad F(x) := (x - S_0)^2,$$

and obtain that the corresponding capital process $(V_t(\varphi))_{t\in[0,T]}$ satisfies

$$V_T(\varphi) = \int_0^T 2(S_s - S_0)\,\mathrm{d}S_s = (S_T - S_0)^2 \ge 0 \quad \text{and} \quad V_0(\varphi) = 0,$$

which is an arbitrage opportunity as soon as $\mathbb{P}(S_T \neq S_0) > 0$. Recall that realistic probabilistic models for financial markets should exclude arbitrage opportunities and thus we see that we want to model financial markets in continuous time based on stochastic processes with rather irregular sample paths, like for instance the so-called Brownian motion, which we will define later (Definiton 2.1). However, before that we need to find good answers to the following questions:

- What are suitable stochastic processes to develop stochastic calculus?
- How can we define the integral $\int_0^T f(S_s)\,\mathrm{d}S_s$ for such stochastic processes?
- How can we replace the fundamental theorem of calculus?

Application 2: probabilistic models – stochastic differential equations. In the modeling of real-world dynamics, as they appear, e.g., in economics, physics, and engineering, (ordinary) differential equations are fundamental and omnipresent mathematical objects. In their maybe simplest form an ordinary differential equation reads as

$$\frac{\mathrm{d}}{\mathrm{d}t}X_t = \mu(X_t), \quad t \in [0,T],$$

for a fixed (Lipschitz) continuous function $\mu\colon \mathbb{R} \to \mathbb{R}$. Intuitively, this ordinary differential equations says that the infinitesimal change of the function X with respect to time t is given by $\mu(X_t)$, that is

$$X_{t+h} - X_t \approx \mu(X_t)h \quad \text{for small } h > 0. \tag{1.1}$$

Many real-world dynamics are subject to random perturbations like, e.g., the evolution of stock prices, growth dynamics of populations and physical systems subject to thermal fluctuations. In order to model such random phenomena, one wants to perturb the system (1.1) by a random noise. For this purpose we take a Brownian motion $(B_t)_{t\in[0,T]}$ and recall that $B_{t+h} - B_t \sim \mathcal{N}(0,h)$. Adding to the system (1.1) a random perturbation leads to

$$X_{t+h} - X_t \approx \mu(X_t)h + \sigma(X_t)(B_{t+h} - B_t) \quad \text{for small } h > 0 \tag{1.2}$$

for a fixed (Lipschitz) continuous function $\sigma\colon \mathbb{R} \to \mathbb{R}$. Naively, we would like to divide equation (1.2) and send $h \to 0$, i.e.

$$\frac{X_{t+h} - X_t}{h} \approx \mu(X_t) + \sigma(X_t)\frac{(B_{t+h} - B_h)}{h} \xrightarrow{h\to 0} \frac{\mathrm{d}}{\mathrm{d}t}X_t = \mu(X_t) + \sigma(X_t)\frac{\mathrm{d}}{\mathrm{d}t}B_t.$$

However, the derivative $\frac{\mathrm{d}}{\mathrm{d}t}B_t$ does <u>not</u> exist since the sample paths of a Brownian motion are almost surely nowhere differentiable! Instead, sending $h \to 0$ in equation (1.2) leads to a so-called *stochastic differential equation (SDE)*

$$\mathrm{d}X_t = \mu(X_t)\,\mathrm{d}t + \sigma(X_t)\,\mathrm{d}B_t, \quad t \in [0,T],$$

where the symbol d represents an infinitesimal difference. This brings us to the questions:

- How can we give a mathematical meaning to stochastic differential equations?
- When does there exist a unique solution to a stochastic differential equation?

The scope of this course is:

(i) Brownian motion and (local) martingales,

(ii) stochastic integration and Itô's formula,

(iii) stochastic differential equations,

(iv) further topics (martingale representation, Girsanov theorem),

(v) some first application in mathematical finance.

2 Brownian motion and local martingales

In this chapter we introduce and study continuous-time stochastic processes allowing us later to develop an associated stochastic integration theory. From previous lecture courses we might already know (discrete-time) martingales and Brownian motion, which is the most frequently appearing prototypical example of a continuous-time stochastic process. Throughout the course, we always suppose:

Assumption. *We work on a suitable complete probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *and* $T \in (0, \infty]$.

Let us agree on the following conventions:

- $[0,T] := [0,\infty)$ for $T = \infty$,
- $X = Y$ for two random variables X, Y means $X = Y$ a.s. (w.r.t. $\mathbb{P}$).

Furthermore, let us introduce (or recall) some terminology and definitions:

- A family of $\mathbb{R}^d$-valued random variables $(X_t)_{t\in[0,T]}$ is called **stochastic process.**
- For every state of the world $\omega \in \Omega$, the mapping

$$X_\cdot(\omega)\colon [0,T] \to \mathbb{R}^d, \quad t \mapsto X_t(\omega),$$

  is called **sample paths** of $(X_t)_{t\in[0,T]}$.
- We say $(X_t)_{t\in[0,T]}$ is **continuous** (or **right-continuous**) **process** if the sample paths of $(X_t)_{t\in[0,T]}$ are almost surely continuous (or right-continuous).
- A family of σ-algebras $(\mathcal{F}_t)_{t\in[0,T]} \subseteq \mathcal{F}$ is called **filtration** if $\mathcal{F}_s \subseteq \mathcal{F}_t$ for $s,t \in [0,T]$ with $s \le t$.

Given a $(\mathcal{F}_t)_{t\in[0,T]}$ be a filtration.

- A random variable τ with values in $[0,T] \cup \{\infty\}$ is called a $(\mathcal{F}_t)$**-stopping time** if

$$\{\tau \le t\} \in \mathcal{F}_t \quad \text{for any } t \in [0,T].$$

- A stochastic process $(X_t)_{t\in[0,T]}$ is called $(\mathcal{F}_t)$**-adapted** (or **adapted**) if X_t is $\mathcal{F}_t$-measurable.

Lecture 2

2.1 Brownian motion

Let us start by recalling the definition of a Brownian motion.

Definition 2.1. A real-valued stochastic process $B = (B_t)_{t\in[0,T]}$ is called a (standard one-dimensional) **Brownian motion** if:

(i) $B_0 = 0$ a.s.

(ii) B has independent increments, i.e., for all $n \in \mathbb{N}$ and $0 \le t_0 < t_1 < \cdots < t_n \le T$,

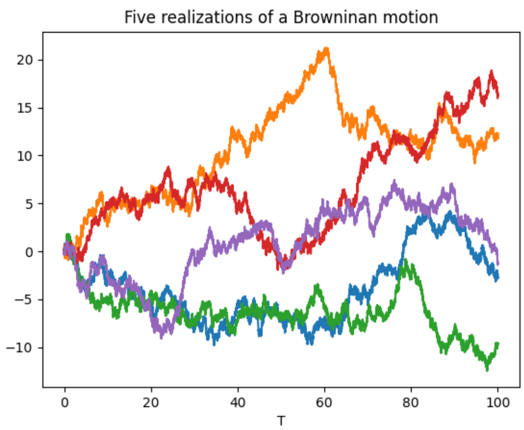
$$B_{t_0} - B_0, B_{t_1} - B_{t_0}, \ldots, B_{t_n} - B_{t_{n-1}} \quad \text{are independent random variables.}$$

(iii) The increments of B are stationary and normally distributed, i.e.,

$$B_t - B_s \sim \mathcal{N}(0, |t-s|), \quad s, t \in [0,T].$$

(iv) B has almost surely continuous sample paths, i.e., the map $t \mapsto B_t(\omega)$ is continuous for almost all $\omega \in \Omega$.

Remark 2.2. Note that there are various equivalent ways to define what it means for a stochastic process to be a “Brownian motion”. A construction of a Brownian motion and the study of its property is part of the lecture course “Wahrscheinlichkeitstheorie I”. The first construction of a Brownian motion $(B_t)_{t\in[0,T]}$ on a suitable probability space $(\Omega, \mathcal{F}, \mathbb{P})$ was provided by Norbert Wiener in 1923. In his honor, Brownian motion is also called **Wiener process**. A Brownian motion can be considered as a “functional version” of the normal distribution and can be constructed as a scaling limit of a normalized random walk, as stated by Donsker’s theorem.



Let $(X_t)_{t\in[0,T]}$ be a stochastic process, e.g. a Brownian motion. As we know, every stochastic process $(X_t)_{t\in[0,T]}$ generates a filtration $(\mathcal{F}_t^X)_{t\in[0,T]}$ by setting

$$\mathcal{F}_t^X := \sigma(X_s : s \le t) := \sigma(X_s^{-1}(A) : A \in \mathcal{B}(\mathbb{R}^d),\ s \in [0,t]), \quad t \in [0,T],$$

where $\mathcal{B}(\mathbb{R})$ denotes the Borel σ-algebra on $\mathbb{R}$. The filtration $(\mathcal{F}_t^X)_{t\in[0,T]}$ is called **natural filtration** of $(X_t)_{t\in[0,T]}$.

In case of a Brownian motion $(B_t)_{t\in[0,T]}$ and its natural filtration of $(\mathcal{F}_t^B)_{t\in[0,T]}$, it follows, thanks to its independent increments, that the future increments are independent of the past natural filtration. More precisely, we have:

Fact. *A Brownian motion $(B_t)_{t\in[0,T]}$ has the (simple) Markov property. In particular, we know that*

$$B_t - B_s \quad \textit{is independent of } \mathcal{F}_s^B,$$

for all $s,t \in [0,T]$ with $s<t$.

The Markov property is proven in the lecture course "Wahrscheinlichkeitstheorie I"

While the natural filtration of a stochastic process $(X_t)_{t\in[0,T]}$ is certainly a natural choice, it sometimes lacks some desirable regularity properties.

Example 2.3. For $A \in \mathcal{F} \setminus \{\emptyset, \Omega\}$ we define the stochastic process $(X_t)_{t\in[0,2]}$ by

$$X_t(\omega) := \begin{cases} (t-1)\mathbb{1}_{\{t>1\}} & \text{if } \omega \in A \\ (1-t)\mathbb{1}_{\{t>1\}} & \text{if } \omega \in A^c \end{cases}, \quad t \in [0,2].$$

In this case $(X_t)_{t\in[0,2]}$ is a continuous stochastic process and its natural filtration $(\mathcal{F}_t^X)_{t\in[0,2]}$ is given by

$$\mathcal{F}_t^X = \begin{cases} \{\emptyset, \Omega\} & \text{if } t \in [0,1] \\ \{\emptyset, A, A^c, \Omega\} & \text{if } t \in (1,2] \end{cases}.$$

However, the hitting time $\tau := \inf\{t \in [0,2] \,:\, X_t < 0\} \wedge 2$ is not a stopping time with respect to $(\mathcal{F}_t^X)_{t\in[0,2]}$ since

$$\{\tau \leq 1\} = \{\tau = 1\} = A^c \notin \mathcal{F}_1^X.$$

To ensure that hitting times are always stopping times as well as for other purposes, one often works in probability theory under the assumption that the underlying filtration satisfies the so-called "usual conditions". Let us denote the $\mathbb{P}$-null sets by

$$\mathcal{N} := \{A \in \mathcal{F} \,:\, \mathbb{P}(A) = 0\}.$$

Definition 2.4. The filtration $(\mathcal{F}_t)_{t\in[0,T]}$ satisfies the **usual conditions** if

- $\mathcal{F}_0$ contains all $\mathbb{P}$-null sets $\mathcal{N}$ ("*completeness*"),
- $\mathcal{F}_t = \mathcal{F}_{t+} := \bigcap_{s>t} \mathcal{F}_s$ for $t \in [0,T)$ (*"right-continuity"*).

For instance, the completeness implies especially that any modification of an adapted stochastic processes is again adapted. In case a filtration is generated by a Brownian motion, the completeness already implies right-continuity. However, in general, this implication is not true.

Proposition 2.5. *Let $(B_t)_{t\in[0,T]}$ be a Brownian motion. The completed natural filtration $(\mathcal{F}_t)_{t\in[0,T]}$ of a Brownian motion $(B_t)_{t\in[0,T]}$, defined by*

$$\mathcal{F}_t := \sigma\big(\mathcal{F}_t^B, \mathcal{N}\big) = \sigma\big(\{B_s^{-1}(A) : A \in \mathcal{B}(\mathbb{R}),\, s \in [0,t]\}, \mathcal{N}\big), \quad \textit{for} \quad t \in [0,T],$$

is right-continuous, i.e. $\mathcal{F}_t = \mathcal{F}_{t+}$ for $t \in [0,T)$.

Proof. Since $\mathcal{F}_t \subseteq \mathcal{F}_{t+} := \bigcap_{s>t} \mathcal{F}_s$ is obvious, it is sufficient to show that $\mathcal{F}_{t+} \subseteq \mathcal{F}_t$ for every $t \in [0,T)$. This means, we want to show that:

$$A \in \mathcal{F}_{t+} \quad \Rightarrow \quad A \in \mathcal{F}_t.$$

For this purpose, we first show, for $d \in \mathbb{N}$, $0 \leqslant t_1 < t_2 < \cdots < t_d$ with $t_1, \ldots, t_d \in [0,T]$, $t \in [0,T)$ and any continuous and bounded $f\colon \mathbb{R}^d \to \mathbb{R}$, that

$$\mathbb{E}[f(B_{t_1}, \ldots, B_{t_d})|\mathcal{F}_{t+}] \text{ is } \mathcal{F}_t\text{-measurable.} \tag{2.1}$$

Let $k \in \{1, \ldots, d-1\}$ be such that $t_k \leqslant t < t_{k+1}$. For $n \in \mathbb{N}$ sufficiently large such that $t + \frac{1}{n} < t_{k+1}$, we have

$$\begin{aligned}
&\mathbb{E}\big[f(B_{t_1}, \ldots, B_{t_d})|\mathcal{F}_{t+1/n}\big] \\
&\quad = \mathbb{E}\big[f(B_{t_1}, \ldots, B_{t_k}, B_{t+\frac{1}{n}} + (B_{t_{k+1}} - B_{t+\frac{1}{n}}), \ldots, B_{t+\frac{1}{n}} + (B_{t_d} - B_{t+\frac{1}{n}}))\big|\mathcal{F}_{t+1/n}\big] \\
&\quad = \int_{\mathbb{R}^{d-k}} f(B_{t_1}, \ldots, B_{t_k}, B_{t+\frac{1}{n}} + x_1, \ldots, B_{t+\frac{1}{n}} + x_{d-k})\, p_n(x_1, \ldots, x_{d-k})\, \mathrm{d}x_1 \cdots \mathrm{d}x_{d-k}
\end{aligned}$$

where p_n denotes the probability density of $(B_{t_{k+1}} - B_{t+\frac{1}{n}}, \ldots, B_{t_d} - B_{t+\frac{1}{n}})$. By the properties of the normal probability density, the normal distribution densities $(p_n)_{n\in\mathbb{N}}$ converge pointwise to the density p of $(B_{t_{k+1}} - B_t, \ldots, B_{t_d} - B_t)$ and there is some integrable function dominating all p_n. Together with (a.s.) continuity of B and f we conclude that $\mathbb{E}[f(B_{t_1}, \ldots, B_{t_d})|\mathcal{F}_{t+1/n}]$ converges almost surely to

$$\int_{\mathbb{R}^{d-k}} f(B_{t_1}, \ldots, B_{t_k}, B_t + x_1, \ldots, B_t + x_{d-k}) p(x_1, \ldots, x_{d-k})\, \mathrm{d}x_1 \cdots \mathrm{d}x_{d-k}$$

for $n \to \infty$ and this limit is $\mathcal{F}_t$-measurable. On the other hand $(\mathbb{E}[f(B_{t_1}, \ldots, B_{t_d})|\mathcal{F}_{t+1/n}])_{n\in\mathbb{N}}$ is a backward martingale and thus converges almost surely to $\mathbb{E}[f(B_{t_1}, \ldots, B_{t_d})|\mathcal{F}_{t+}]$ by the convergence theorem for backwards martingales (Theorem A.17). Since $\mathcal{N} \subseteq \mathcal{F}_t$, this implies (2.1).

By dominated convergence theorem, (2.1) generalizes to all bounded measurable functions f and especially we can replace $f(B_{t_1}, \ldots, B_{t_d})$ with indicator functions $\mathbb{1}_A$ for $A \in \sigma(B_{t_1}, \ldots, B_{t_d})$. By a standard argument using Dynkin systems, we conclude that $\mathbb{E}[\mathbb{1}_A|\mathcal{F}_{t+}]$ is $\mathcal{F}_t$-measurable for all $A \in \mathcal{F}_T^B := \sigma(B_t, t \in [0,T])$ and thus for all $A \in \sigma(\mathcal{F}_T^B, \mathcal{N})$. Therefore, for any $A \in \mathcal{F}_{t+}$ we have that $\mathbb{1}_A = \mathbb{E}[\mathbb{1}_A|\mathcal{F}_{t+}]$ is $\mathcal{F}_t$-measurable, i.e. $A \in \mathcal{F}_t$. □

Luckily, the simple Markov property still holds for the completed natural filtration of a Brownian motion.

Lemma 2.6. *Let $(B_t)_{t\in[0,T]}$ be a Brownian motion and let $(\mathcal{F}_t)_{t\in[0,T]}$ be its completed natural filtration. Then, for every $s, t \in [0,T]$ with $s < t$, the Brownian increment $B_t - B_s$ is independent of $\mathcal{F}_s$.*

Proof. To prove that $B_t - B_s$ is independent of $\mathcal{F}_s$, we need to show that

$$\mathbb{E}[\mathbb{1}_E(B_t - B_s)\mathbb{1}_F] = \mathbb{E}[\mathbb{1}_E(B_t - B_s)]\mathbb{E}[\mathbb{1}_F] \quad \text{for all } E \in \mathcal{B}(\mathbb{R}),\, F \in \mathcal{F}_s.$$

By the definition of $\mathcal{F}_t := \sigma(\mathcal{F}_t^B, \mathcal{N})$, for every $F \in \mathcal{F}_s$ there exists a set $F' \in \mathcal{F}_s^B$ such that the symmetric difference $F \setminus F' \cup F' \setminus F \in \mathcal{N}$. Hence, we see that

$$\begin{aligned}
\mathbb{E}[\mathbb{1}_E(B_t - B_s)\mathbb{1}_F] &= \mathbb{E}[\mathbb{1}_E(B_t - B_s)\mathbb{1}_{F'}] \\
&= \mathbb{E}[\mathbb{1}_E(B_t - B_s)]\mathbb{E}[\mathbb{1}_{F'}] = \mathbb{E}[\mathbb{1}_E(B_t - B_s)]\mathbb{E}[\mathbb{1}_F]
\end{aligned}$$

since $B_t - B_s$ is independent of $\mathcal{F}_s^B$.

(i) The completed natural filtration $(\mathcal{F}_t)_{t\in[0,T]}$ of a Brownian motion $(B_t)_{t\in[0,T]}$ is often called *Brownian standard filtration.*

(ii) Alternatively, one can modify the natural filtration $(\mathcal{F}^B_{t\in[0,T]})$ of a Brownian motion $(B_t)_{t\in[0,T]}$ to a right-continuous filtration $(\mathcal{F}^+_t)_{t\in[0,T)}$ by setting

$$\mathcal{F}^+_t := \bigcap_{s>t} \mathcal{F}^B_s \quad \text{for } t \in [0,T).$$

One still can verify that the Brownian increment $B_t - B_s$ is independent of $\mathcal{F}^+_s$, for every $s,t \in [0,T)$ with $s < t$. The difference between $(\mathcal{F}^B_t)_{t\in[0,T)}$ and $(\mathcal{F}^+_t)_{t\in[0,T)}$ are essentially the $\mathbb{P}$-null sets by Blumenthal's $0-1$ law.

□

2.2 Martingales

While the Brownian motion is a fundamental stochastic process in stochastic calculus, a more general class of continuous-time stochastic processes are so-called martingales and their variants. They often serve as models for fair or unfair games.

Suppose our underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is equipped with a filtration $(\mathcal{F}_t)_{t\in[0,T]}$.

Definition 2.7. Let $(X_t)_{t\in[0,T]}$ be a real-valued $(\mathcal{F}_t)$-adapted stochastic process such that $\mathbb{E}[|X_t|] < \infty$ for all $t \in [0,T]$. $(X_t)_{t\in[0,T]}$ is called a

- $(\mathcal{F}_t)$**-martingale** if $\mathbb{E}[X_t|\mathcal{F}_s] = X_s$,
- $(\mathcal{F}_t)$**-sub-martingale** if $\mathbb{E}[X_t|\mathcal{F}_s] \geq X_s$,
- $(\mathcal{F}_t)$**-super-martingale** if $\mathbb{E}[X_t|\mathcal{F}_s] \leq X_s$,

for all $s,t \in [0,T]$ with $s \leq t$.

Remark. We will often drop the prefix $(\mathcal{F}_t)$ in "$(\mathcal{F}_t)$-martingale" and just say "martingale" if the involved filtration is obvious from the context or it is just the associated natural filtration. The same applies to sub- and super-martingales.

Example 2.8. Let $(B_t)_{t\in[0,T]}$ be a Brownian motion and $(\mathcal{F}_t)_{t\in[0,T]}$ its complete natural filtration. The following processes are $(\mathcal{F}_t)$-martingales:

- $(X^1_t)_{t\in[0,T]}$, defined by $X^1_t := B_t$,
- $(X^2_t)_{t\in[0,T]}$, defined by $X^2_t := B^2_t - t$,
- $(X^3_t)_{t\in[0,T]}$, defined by $X^3_t := \exp(\sigma B_t - \frac{\sigma^2}{2}t)$ for all $\sigma > 0$.

It is a good exercise to check this yourself.

Lecture 3

The first thing we would like to investigate is how martingales behave with respect to stopping times.

Definition 2.9. Let $(\mathcal{F}_t)_{t\in[0,T]}$ be a filtration. A random variable τ with values in $[0,T]\cup\{\infty\}$ is called a $(\mathcal{F}_t)$**-stopping time** if

$$\{\tau \le t\} \in \mathcal{F}_t \quad \text{for any } t \in [0,T].$$

Let τ be a stopping time. The σ**-algebra of** τ**-past** is defined as

$$\mathcal{F}_\tau := \big\{A \in \mathcal{F} \,:\, A \cap \{\tau \le t\} \in \mathcal{F}_t \text{ for any } t \in [0,T]\big\}.$$

Remark. For every stopping time τ, the σ-algebra of the τ-past $\mathcal{F}_\tau$ is indeed a σ-algebra and τ is an $\mathcal{F}_\tau$-measurable random variable. Furthermore, if $(X_t)_{t\in[0,T]}$ is a right-continuous $(\mathcal{F}_t)$-adapted process, then X_τ is an $\mathcal{F}_\tau$-measurable random variable. (Please check yourself!)

Based on the notion of stopping times, we arrive at an equivalent characterization of martingales.

Theorem 2.10. *Let $(\mathcal{F}_t)_{t\in[0,T]}$ be a filtration and let $(X_t)_{t\in[0,T]}$ be a right-continuous $(\mathcal{F}_t)$-adapted process. Then, the following statements are equivalent:*

(i) *$(X_t)_{t\in[0,T]}$ is an $(\mathcal{F}_t)$-martingale.*

(ii) *For all bounded $(\mathcal{F}_t)$-stopping times τ, we have $X_\tau \in L^1$ and $\mathbb{E}[X_\tau] = \mathbb{E}[X_0]$.*

(iii) ***(Optional sampling)*** *For all bounded $(\mathcal{F}_t)$-stopping times σ, τ with $\sigma \leqslant \tau$, we have $X_\tau \in L^1$ and $\mathbb{E}[X_\tau|\mathcal{F}_\sigma] = X_\sigma$.*

(iv) ***(Optional stopping)*** *For all $(\mathcal{F}_t)$-stopping times τ, the stopped process $(X_t^\tau)_{t\in[0,T]} := (X_{\tau\wedge t})_{t\in[0,T]}$ is an $(\mathcal{F}_t)$-martingale.*

Proof. *(i)* $\Rightarrow$ *(ii):* Let $T' \in (0,T]$ such that $\tau \leqslant T'$ and define the stopping times

$$\tau_n := \sum_{k=1}^{2^n} s_{n,k} \mathbb{1}_{(s_{n,k-1}, s_{n,k}]}(\tau) \qquad \text{for} \qquad s_{n,k} := T' \frac{k}{2^n}.$$

In particular, $\tau_n(\omega) \downarrow \tau(\omega)$ for almost all $\omega \in \Omega$. Consider the discrete-time martingale

$$(X_{s_{n,k}})_{k=0}^{2^n} \qquad \text{w.r.t.} \qquad (\mathcal{F}_{s_{n,k}})_{k=0}^{2^n}.$$

The discrete-time optional sampling theorem (Theorem A.13) yields $X_{\tau_n} \in L^1$ and

$$X_{\tau_n} = \mathbb{E}[X_{T'}|\mathcal{F}_{\tau_n}] \qquad \text{for all } n \in \mathbb{N}.$$

Note that $(X_{\tau_n})_{n\in\mathbb{N}}$ is uniformly integrable (Please check yourself, cf. problem sheets.).

Since $(X_t)_{t\in[0,T]}$ is right-continuous, we have

$$X_{\tau_n} \to X_\tau \quad \text{as } n \to \infty \quad \mathbb{P}\text{-a.s.} \qquad \Longrightarrow \qquad X_{\tau_n} \xrightarrow{\mathbb{P}} X_\tau \quad \text{as } n \to \infty.$$

Together with the uniform integrability, this implies $X_{\tau_n} \xrightarrow{L^1} X_\tau$. Therefore,

$$\begin{aligned}
\mathbb{E}[X_\tau] &= \lim_{n\to\infty} \mathbb{E}[X_{\tau_n}] \\
&= \lim_{n\to\infty} \mathbb{E}\big[\mathbb{E}[X_{T'}|\mathcal{F}_{\tau_n}]\big] \\
&= \mathbb{E}[X_{T'}] \\
&= \mathbb{E}[X_0].
\end{aligned}$$

(ii) $\Rightarrow$ *(iii)* By the definition of conditional expectation, we need to show that

$$\mathbb{E}\big[X_\sigma \mathbb{1}_A\big] = \mathbb{E}\big[X_\tau \mathbb{1}_A\big] \qquad \forall A \in \mathcal{F}_\sigma.$$

For this purpose, we take $A \in \mathcal{F}_\sigma$ and consider

$$\widetilde{\sigma}(\omega) := \begin{cases} \sigma(\omega), & \omega \in A, \\ \tau(\omega), & \omega \in A^c. \end{cases}$$

Note that $\widetilde{\sigma}$ is a bounded stopping time, since

$$\{\widetilde{\sigma} \leqslant t\} = (\underbrace{\{\sigma \leqslant t\} \cap A}_{\in \mathcal{F}_t}) \cup (\underbrace{\{\tau \leqslant t\} \cap A^c}_{\in \mathcal{F}_t \text{ since } A^c \in \mathcal{F}_\sigma \subseteq \mathcal{F}_\tau}) \in \mathcal{F}_t \quad \text{for every } t \in [0,T].$$

Applying *(ii)* to $\widetilde{\sigma}$ and τ yields

$$\mathbb{E}\big[X_\tau(\mathbb{1}_A + \mathbb{1}_{A^c})\big] = \mathbb{E}[X_0] = \mathbb{E}[X_{\widetilde{\sigma}}] = \mathbb{E}\big[X_\sigma \mathbb{1}_A\big] + \mathbb{E}\big[X_\tau \mathbb{1}_{A^c}\big].$$

Hence,

$$\mathbb{E}\big[X_\sigma \mathbb{1}_A\big] = \mathbb{E}\big[X_\tau \mathbb{1}_A\big] \qquad \forall A \in \mathcal{F}_\sigma.$$

(iii) $\Rightarrow$ *(iv)* Let τ be a stopping time. Since $\tau \wedge t$ is a bounded stopping time, we have $X_{\tau\wedge t} \in L^1$ and $X_{\tau \wedge t}$ is $\mathcal{F}_{\tau\wedge t} \subseteq \mathcal{F}_t$-measurable. Finally for any $A \in \mathcal{F}_s$, $s \leqslant t$ *(iii)* yields

$$\begin{aligned}
\mathbb{E}[X_{\tau\wedge t}\mathbb{1}_A] &= \mathbb{E}[X_{\tau\wedge t}\underbrace{\mathbb{1}_A \mathbb{1}_{\{s\leq\tau\}}}_{\mathcal{F}_{\tau\wedge s}\text{-mb.}}] + \mathbb{E}[X_{\tau\wedge t}\underbrace{\mathbb{1}_A \mathbb{1}_{\{s>\tau\}}}_{\tau\leqslant s\leqslant t}] \\
&= \mathbb{E}\big[\mathbb{E}[X_{\tau\wedge t}|\mathcal{F}_{\tau\wedge s}]\mathbb{1}_A \mathbb{1}_{\{s\leq\tau\}}\big] + \mathbb{E}[X_{\tau\wedge s}\mathbb{1}_A \mathbb{1}_{\{s>\tau\}}] \\
&= \mathbb{E}\big[X_{\tau\wedge s}\mathbb{1}_A\big].
\end{aligned}$$

(iv) $\Rightarrow$ *(i)* To verify $X_t \in L^1$, measurability and $\mathbb{E}[X_t|\mathcal{F}_s] = X_s$ choose $\tau = t$. □

A very handy property of martingales is that, roughly speaking, the running maximum of a martingale can be controlled by its terminal value in a probabilistic manner. This is the content of Doob's martingale inequalities. We will need this inequalities later for the construct of a stochastic integral with respect to a Brownian motion.

Proposition 2.11. *Let $(X_t)_{t\in[0,T]}$ be a right-continuous martingale or a right-continuous non-negative sub-martingale and $T < \infty$.*

(i) ***Doob's maximal inequality*** *holds $p \geq 1$ and for all $\lambda > 0$:*

$$\mathbb{P}\Big(\sup_{0\leqslant t\leqslant T} |X_t| \geqslant \lambda\Big) \leqslant \frac{1}{\lambda^p}\mathbb{E}[|X_T|^p].$$

(ii) *For $p > 1$ and supposing $X_T \in L^p$, we have* ***Doob's L^p-inequality***

$$\mathbb{E}\Big[\sup_{0\leqslant t\leqslant T} |X_t|^p\Big] \leqslant \Big(\frac{p}{p-1}\Big)^p \mathbb{E}[|X_T|^p].$$

Proof. The proof follows by approximating the right-continuous stochastic process $(X_t)_{t\in[0,T]}$ by discretizing the time interval $[0,T]$ and applying the discrete-time version of Doob's maximal and L^p-inequality, which you can find in Proposition A.16. □

When one develops classical integration like Riemann or Lebesgue-Stieltjes integration, one usually relies (in a more or less direct manner) on the concept of *bounded variation.*

Definition 2.12. A **partition** Π of $[0,T]$ is a family of disjoint intervals $J = (s_J, t_J]$ which cover $[0,T]$. The **mesh size** of Π is given by

$$|\Pi| := \sup_{J\in\Pi} |t_J - s_J|.$$

A sequence of partitions $(\Pi_n)_{n\in\mathbb{N}}$ of $[0,T]$ is called a **zero-sequence of partitions** if $|\Pi_n| \to 0$ as $n \to \infty$, $\Pi_n \subseteq \Pi_{n+1}$ (meaning that each interval $J \in \Pi_n$ is a finite union of intervals in Π_{n+1}) and $\{J : J \in \Pi_n \text{ s.t. } J \subseteq [0,t]\}$ is finite for each $t \in (0,T]$ and $n \in \mathbb{N}$.

Let $(X_t)_{t\in[0,T]}$ be a right-continuous stochastic process. The **variation process** $(|X|_t)_{t\in[0,T]}$ of $(X_t)_{t\in[0,T]}$ is defined by

$$|X|_t := \sup_{\Pi} \sum_{J\in\Pi} |\Delta_{J\cap[0,t]} X|, \qquad t \in [0,T], \qquad \text{writing } \Delta_{(u,v]\cap[0,t]} X := X_{t\wedge u} - X_{v\wedge t},$$

where the supremum is taken over all partitions Π of $[0,T]$ into intervals of the form $J = (s_J, t_J]$ such that $\Pi \cap [0,t]$ is composed of finitely many intervals for any $t \in (0,T]$.

Remark 2.13. Let us briefly discuss the variation of a deterministic function $g\colon [0,T] \to \mathbb{R}$ and the associated Lebesgue-Stieltjes integration, both without giving proofs. Suppose $g\colon [0,T] \to \mathbb{R}$ is continuous and of finite variation, that is

$$|g|_t := \sup_{\Pi} \sum_{J\in\Pi} |\Delta_{J\cap[0,t]} g| = \lim_{n\to\infty} \sum_{J\in\Pi_n} |\Delta_{J\cap[0,t]} g| < \infty, \quad t \in [0,T],$$

where the limit holds along any zero-sequence of partitions $(\Pi_n)_{n\in\mathbb{N}}$. (Note that the equality in the previous equation actually needs to be proven.) Now it can be shown that the function g is of finite variation if and only if g can be decomposed into a difference of two non-negative and non-decreasing functions $g^{\uparrow}, g^{\downarrow}\colon [0,T] \to [0,\infty)$, i.e.

$$g(t) = g^{\uparrow}(t) - g^{\downarrow}(t) \quad \text{for } t \in [0,T].$$

In this case, we can identify $g^{\uparrow}$ and $g^{\downarrow}$ with the two measures $\mu_{g^{\uparrow}}$ and $\mu_{g^{\downarrow}}$ by setting

$$\mu_{g^{\uparrow}}((s,t]) := g^{\uparrow}(t) - g^{\uparrow}(s) \quad \text{and} \quad \mu_{g^{\downarrow}}((s,t]) := g^{\downarrow}(t) - g^{\downarrow}(s), \quad \text{for } (s,t] \subseteq [0,T].$$

If we now want to define the integral $\int_0^T f(s)\,\mathrm{d}g(s)$ for a continuous function $f\colon [0,T] \to \mathbb{R}$, we can simply define

$$\int_0^T f(s)\,\mathrm{d}g(s) := \int_0^T f(s)\,\mathrm{d}\mu_{g^{\uparrow}}(s) - \int_0^T f(s)\,\mathrm{d}\mu_{g^{\downarrow}}(s),$$

where $\int_0^T f(s)\,\mathrm{d}\mu_{g^{\uparrow}}(s)$ and $\int_0^T f(s)\,\mathrm{d}\mu_{g^{\downarrow}}(s)$ can be understood as Lebesgue-Stieltjes integrals (or as Riemann integrals).

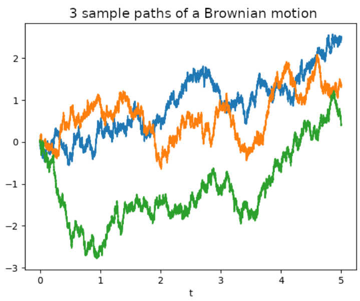
Unfortunately, classical integration theories do not apply to martingales (and thus to a Brownian motion).

Lemma 2.14. *Let $(X_t)_{t\in[0,T]}$ be a continuous martingale of locally finite variation, i.e., $|X|_t < \infty$ $\mathbb{P}$-a.s. for all $t \in [0,T]$. Then,*

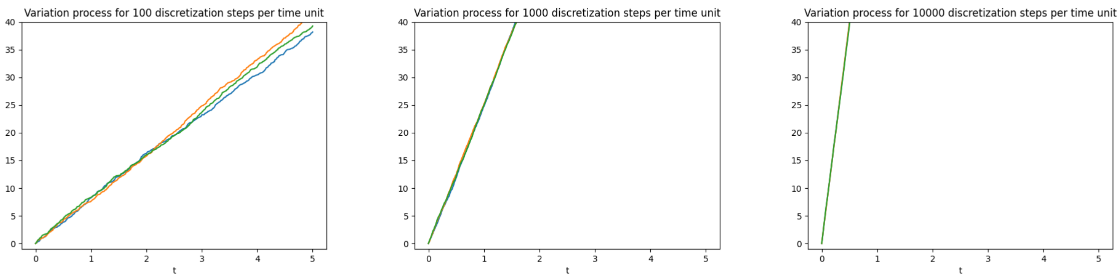
$$\mathbb{P}(\{\omega \in \Omega : X_t(\omega) = X_0(\omega) \text{ for } t \in [0,T]\}) = 1.$$

Proof. Please see the problem sheets. □

As an immediate consequence of Lemma 2.14, we know there are no non-trivial martingales $(X_t)_{t\in[0,T]}$ of locally finite variation. In particular, a Brownian motion $(B_t)_{t\in[0,T]}$ cannot be of finite variation as $B_t \sim \mathcal{N}(0,t)$. Simulating three sample paths of a Brownian motion leads to a figure like the following one.



When we numerically approximate the corresponding variation process of the three above sample paths of a Brownian motion, we see that, indeed, the variation processes directly explode to infinite if the mesh size of the partitions goes to zero, that is, as expected, we have $|B|_t = \infty$ for every $t \in (0,T]$.



Lecture 4

To replace (in a wider sense) the variation processes $(|X|_t)_{t\in[0,T]}$ for a martingale $(X_t)_{t\in[0,T]}$, we shall consider instead the sum of square increments

$$\sum_{J\in\Pi_n} (\Delta_{J\cap[0,t]} X)^2$$

along a zero-sequence of partitions $(\Pi_n)_{n\in\mathbb{N}}$.

Proposition 2.15. *Let $X = (X_t)_{t\in[0,T]}$ be a continuous and bounded (i.e. $|X_t(\omega)| \leqslant C$ for some $C > 0$, all $t \in [0,T]$ and a.e. $\omega \in \Omega$) martingale. Then, there exists a continuous stochastic process $\langle X\rangle = (\langle X\rangle_t)_{t\in[0,T]}$, given by*

$$\langle X\rangle_t := \lim_{n\to\infty} \sum_{J\in\Pi_n} (\Delta_{J\cap[0,t]} X)^2, \quad \text{uniformly on } [0,T] \text{ in } L^2,$$

for any zero-sequence of partitions $(\Pi_n)_{n\in\mathbb{N}}$, *with the following properties:*

(i) $\langle X\rangle_0 = 0$ *and* $\langle X\rangle$ *is non-decreasing, and*

(ii) $(X_t^2 - \langle X\rangle_t)_{t\in[0,T]}$ *is a martingale.*

The stochastic process $\langle X\rangle = (\langle X\rangle_t)_{t\in[0,T]}$ *is called* ***quadratic variation*** *of* $X = (X_t)_{t\in[0,T]}$.

Later we will see that the quadratic variation process $\langle X\rangle$ is uniquely determined by the properties (i) and (ii) of Proposition 2.15, that is any process satisfying (i) and (ii) coincides with the quadratic variation process.

Proof. Without loss of generality we may assume $X_0 = 0$ (If $X_0 \neq 0$, just consider the process $\widetilde{X}_t := X_t - X_0$.) and $T < \infty$. For a zero-sequence of partitions $(\Pi_n)_{n\in\mathbb{N}}$, we define

$$A_t^n := \sum_{J\in\Pi_n} (\Delta_{J\cap[0,t]}X)^2, \quad \text{for} \quad t\in[0,T],\, n\in\mathbb{N}.$$

Step 1: Show that $N^n := X^2 - A^n$ *is a martingale for all* $n\in\mathbb{N}$.
Using a telescoping sum argument, we have

$$\begin{aligned}
X_t^2 &= \sum_{J\in\Pi_n} (X_{t_J\wedge t}^2 - X_{s_J\wedge t}^2)\\
&= \sum_{J\in\Pi_n} \Big(2X_{s_J\wedge t}(X_{t_J\wedge t} - X_{s_J\wedge t}) + (X_{t_J\wedge t} - X_{s_J\wedge t})^2\Big)\\
&= 2\sum_{J\in\Pi_n} X_{s_J\wedge t}(\Delta_{J\cap[0,t]}X) + A_t^n.
\end{aligned}$$

Therefore, $N_t^n = 2\sum_{J\in\Pi_n} X_{s_J\wedge t}(\Delta_{J\cap[0,t]}X)$ is a martingale, since

$$\begin{aligned}
\mathbb{E}[N_t^n - N_s^n|\mathcal{F}_s] &= 2\sum_{J\in\Pi_n: t_J\geqslant s} \mathbb{E}\big[X_{s_J}(X_{t_J\wedge t} - X_{(s_J\vee s)\wedge t})|\mathcal{F}_s\big]\\
&= 2\sum_{J\in\Pi_n: t_J\geqslant s} \mathbb{E}\big[X_{s_J}\underbrace{\mathbb{E}[X_{t_J\wedge t} - X_{(s_J\vee s)\wedge t}|\mathcal{F}_{s_J\vee s}]}_{=0 \text{ as } X \text{ is a martingale}}|\mathcal{F}_s\big] = 0,
\end{aligned}$$

for all $s,t\in[0,T]$ with $s<t$.

Step 2: Show that $\lim_{m\to\infty}\sup_{n\geqslant m}\mathbb{E}\big[\sup_{0\leqslant t\leqslant T}|N_t^m - N_t^n|^2\big] = 0$ *for* $T>0$.

W.l.o.g. let T be a partition point of Π_m and abbreviate $J_n := \{J\in\Pi_n : J\subseteq[0,T]\}$. For $n,m\in\mathbb{N}$ with $n\geqslant m$, Doob's L^p-inequality (Proposition 2.11) and the martingale property of X yield

$$\begin{aligned}
\mathbb{E}\Big[\sup_{0\leqslant t\leqslant T}|N_t^m - N_t^n|^2\Big] &\leqslant 4\mathbb{E}[|N_T^m - N_T^n|^2]\\
&= 16\mathbb{E}\Big[\Big(\sum_{J\in J_m}\sum_{J\supseteq K\in J_n}(X_{s_J} - X_{s_K})\Delta_K X\Big)^2\Big]\\
&= 16\mathbb{E}\Big[\sum_{J\in J_m}\sum_{J\supseteq K\in J_n}(X_{s_J} - X_{s_K})^2(\Delta_K X)^2\Big]\\
&\leqslant 16\mathbb{E}\Big[\sup_{J\in J_m}\sup_{J\supseteq K\in J_n}(X_{s_J} - X_{s_K})^4\Big]^{1/2}\mathbb{E}\Big[\Big(\sum_{K\in J_n}(\Delta_K X)^2\Big)^2\Big]^{1/2} \quad (2.2)
\end{aligned}$$

with the Cauchy-Schwarz inequality in the last step. Now we treat these two terms in (2.2) separately.

The first term converges to 0 as $|J_m| \to 0$ for $m \to \infty$ due to continuity and boundedness of X and the dominated convergence theorem:

$$\lim_{m\to\infty} \mathbb{E}\Big[\sup_{J\in J_m} \sup_{J\supseteq K\in J_n} (X_{s_J} - X_{s_K})^4\Big]^{1/2} = \mathbb{E}\Big[\lim_{m\to\infty} \sup_{J\in J_m} \sup_{J\supseteq K\in J_n} (X_{s_J} - X_{s_K})^4\Big]^{1/2} = 0.$$

Let X be bounded by $C > 0$. Then the second term can be estimated by

$$\begin{aligned}\mathbb{E}\Big[\Big(\sum_{K\in J_n} (\Delta_K X)^2\Big)^2\Big] &= \mathbb{E}\Big[\sum_{K\in J_n} (\Delta_K X)^4\Big] + 2\mathbb{E}\Big[\sum_{J\in J_n} (\Delta_J X)^2 \sum_{K\in J_n: s_K\geqslant t_J} (\Delta_K X)^2\Big] \\ &\leqslant 4C^2\mathbb{E}\Big[\sum_{K\in J_n} (\Delta_K X)^2\Big] + 2\mathbb{E}\Big[\sum_{J\in J_n} (\Delta_J X)^2 \mathbb{E}\Big[\sum_{K\in J_n: s_K\geqslant t_J} (\Delta_K X)^2 \Big| \mathcal{F}_{t_J}\Big]\Big].\end{aligned}$$

Since the martingale property of X implies

$$\begin{aligned}\mathbb{E}\Big[(\Delta_K X)^2 \Big| \mathcal{F}_{t_J}\Big] &= \mathbb{E}\big[X_{t_K}^2 - 2X_{s_K}X_{t_K} + X_{s_K}^2 | \mathcal{F}_{t_J}\big] \\ &= \mathbb{E}\big[X_{t_K}^2 - X_{s_K}^2 | \mathcal{F}_{t_J}\big],\end{aligned}$$

we conclude

$$\begin{aligned}\mathbb{E}\Big[\Big(\sum_{K\in J_n} (\Delta_K X)^2\Big)^2\Big] &\leqslant 4C^2 \underbrace{\sum_{K\in J_n} \big(\mathbb{E}[X_{t_K}^2] - \mathbb{E}[X_{s_K}^2]\big)}_{=\mathbb{E}[X_T^2]} + 2\mathbb{E}\Big[\sum_{J\in J_n} (\Delta_J X)^2 \mathbb{E}[X_T^2 - X_{t_J}^2 | \mathcal{F}_{t_J}]\Big] \\ &\leqslant 4C^2\mathbb{E}[X_T^2] + 4C^2\mathbb{E}\Big[\sum_{J\in J_n} (\Delta_J X)^2\Big] \\ &= 8C^2\mathbb{E}[X_T^2]. \end{aligned} \tag{2.3}$$

Therefore, the second term in (2.2) is bounded and we have completed Step 2.

Step 3: Existence of $\langle X\rangle$ and conclude its properties.

Due to Step 2 there is a process $N = \lim_{n\to\infty} N^n$ uniformly on $[0,T]$ in L^2. Hence, there exists a sub-sequence $(N^{n_k})_{k\in\mathbb{N}}$ such that $N^{n_k} \to N$ uniformly on $[0,T]$, almost surely, as $k \to \infty$. Hence, N is continuous as it is a uniform limit of continuous stochastic processes $(N^{n_k})_{k\in\mathbb{N}}$.

Next we need to verify that N is a martingale. To show that $\mathbb{E}[N_t|\mathcal{F}_s] = N_s$ for $s,t \in [0,T]$ with $s < t$, it is sufficient to show that

$$\mathbb{E}[(N_t - N_s)\mathbb{1}_A] = 0 \quad \text{for all } A \in \mathcal{F}_s,$$

by the definition of conditional expectation. Indeed, using Fatou's lemma, the martingale properties of the stochastic processes $(N^{n_k})_{k\in\mathbb{N}}$ and the uniform convergence of $(N^{n_k})_{k\in\mathbb{N}}$ to N, we have

$$\begin{aligned}|\mathbb{E}[(N_t - N_s)\mathbb{1}_A]| &= |\mathbb{E}[(N_t - N_t^{n_k})\mathbb{1}_A] + \mathbb{E}[(N_t^{n_k} - N_s^{n_k})\mathbb{1}_A] - \mathbb{E}[(N_s - N_s^{n_k})\mathbb{1}_A] \\ &\leq |\mathbb{E}[(N_t - N_t^{n_k})\mathbb{1}_A]| + |\mathbb{E}[(N_s - N_s^{n_k})\mathbb{1}_A]| \\ &\leq \mathbb{E}[|N_t - N_t^{n_k}|\mathbb{1}_A] + \mathbb{E}[|N_s - N_s^{n_k}|\mathbb{1}_A] \\ &\to 0 \quad \text{as } k \to \infty.\end{aligned}$$

We now define

$$\langle X \rangle := X^2 - N.$$

Then, $\langle X \rangle$ inherits continuity and adaptedness from X^2 and N, $\langle X \rangle_0 = 0$ and $N = X^2 - \langle X \rangle$ is martingale. Moreover, $A_t^n \to \langle X \rangle_t$ uniformly on $[0,T]$ in L^2. Hence,

$$\langle X \rangle_t = \lim_{n\to\infty} \sum_{J\in\Pi_n} (\Delta_{J\cap[0,t]} X)^2$$

uniformly on $[0,T]$ in L^2.

Finally, to verify that $\langle X \rangle$ is non-decreasing, for $s,t \in [0,T]$ with $s < t$ we observe that

$$\begin{aligned}\langle X \rangle_t &= \lim_{n\to\infty} \sum_{J\in\Pi_n} (\Delta_{J\cap[0,t]} X)^2 \\ &= \lim_{n\to\infty} \sum_{J\in\Pi_n} (\Delta_{J\cap[0,s]} X)^2 + \lim_{n\to\infty} \sum_{J\in\Pi_n} (\Delta_{J\cap[s,t]} X)^2 \\ &\geqslant \langle X \rangle_s,\end{aligned}$$

where $\Delta_{J\cap[s,t]} X := X_{(t_J\vee s)\wedge t} - X_{(s_J\vee s)\wedge t}$ for $J = (s_J, t_J]$. □

The quadratic variation process of a martingale plays a central role in the construction of stochastic integration. However, before coming to stochastic integration, we need to introduce so-called local martingales.

2.3 Local martingales

While martingales, like the Brownian motion, are a good starting point to develop a stochastic integration theory, it turns out that this class of stochastic processes is still not large enough. As we will see later, the stochastic integral with respect to a martingale might not be a martingale anymore. Therefore, we need to extend the class of martingales a bit further.

Let $(\mathcal{F}_t)_{t\in[0,T]}$ be a filtration satisfying the usual conditions, that is, $(\mathcal{F}_t)_{t\in[0,T]}$ right-continuous and complete.

Definition 2.16. An $(\mathcal{F}_t)$-adapted process $(X_t)_{t\in[0,T]}$ is called **local martingale** if there is an increasing sequence $(\tau_n)_{n\in\mathbb{N}}$ of $(\mathcal{F}_t)$-stopping times with $\tau_n \uparrow T$ $\mathbb{P}$-a.s. and $(X_t^n)_{t\in[0,T]} := (X_{t\wedge\tau_n} - X_0)_{t\in[0,T]}$ is an $(\mathcal{F}_t)$-martingale for every $n \in \mathbb{N}$. The sequence $(\tau_n)_{n\in\mathbb{N}}$ is called **localizing sequence** for $(X_t)_{t\in[0,T]}$. **Local sub-martingales** and **local super-martingales** are defined analogously.

Note that there are really local martingales which are not martingales.

Example 2.17. Let $(B_t)_{t\in[0,\infty)}$ be a Brownian motion and $T_{-1} := \inf\{t \geq 0 : B_t = -1\}$. The stochastic process

$$X_t := \begin{cases} B_{\frac{t}{1-t}\wedge T_{-1}} & \text{for } 0 \leq t < 1 \\ -1 & \text{for } t = 1 \end{cases}, \quad t \in [0,1],$$

is a continuous local martingale but not a martingale. Indeed, $(X_t)_{t\in[0,1]}$ is not a martingale since

$$\mathbb{E}[X_t] = \begin{cases} 0 & \text{for } 0 \le t < 1 \\ -1 & \text{for } t = 1 \end{cases}$$

as the stopped process $(B_{t\wedge T_{-1}})_{t\in[0,\infty)}$ is a martingale. However, using the localizing sequence $(\tau_n)_{n\in\mathbb{N}}$ given by

$$\tau_n := \inf\{t \in [0,1] \,:\, X_t = n\} \wedge 1, \quad n \in \mathbb{N},$$

$(X_t)_{t\in[0,1]}$ is a local martingale with localizing sequence $(\tau_n)_{n\in\mathbb{N}}$. The proof is based on the dominated convergence theorem:

$$\mathbb{E}[X_{1\wedge\tau_n}] = \mathbb{E}\left[\lim_{k\to\infty} X_{\frac{k}{k+1}\wedge\tau_n}\right] = \lim_{k\to\infty} \mathbb{E}\left[X_{\frac{k}{k+1}\wedge\tau_n}\right] = 0.$$

The same arguments work for the conditional expectation in order to prove the martingale property of $(X_{t\wedge\tau_n})_{t\in[0,1]}$.

The next proposition summarizes some important properties of continuous local martingales:

Proposition 2.18. *Let $X = (X_t)_{t\in[0,T]}$ be a continuous local martingale with respect to $(\mathcal{F}_t)_{t\in[0,T]}$.*

(i) For any $(\mathcal{F}_t)$-stopping time τ the process $Y := X_{\cdot\wedge\tau}$ is a local martingale.

(ii) If $|X| \leqslant B$ for some constant $B > 0$, then X is a martingale.

(iii) If $X = (X_t)_{t\in[0,T]}$ is non-negative with $\mathbb{E}[X_0] < \infty$, then X is a super-martingale.

Proof. W.l.o.g we assume $X_0 = 0$. Let $(\tau_n)_n$ be a localizing sequence for X.

(i) We have

$$Y_{\cdot\wedge\tau_n} = (X_{\cdot\wedge\tau})_{\cdot\wedge\tau_n} = X_{\cdot\wedge(\tau\wedge\tau_n)} = (X_{\cdot\wedge\tau_n})_{\cdot\wedge\tau}$$

Since $X_{\cdot\wedge\tau_n}$ is a martingale, $(X_{\cdot\wedge\tau_n})_{\cdot\wedge\tau}$ is a martingale by optional stopping.

(ii) Integrability of $(X_t)_{t\in[0,T]}$ follows from boundedness. For all $n \in \mathbb{N}$ and $s,t \in [0,T]$ with $s < t$, we have $\mathbb{E}[X_{t\wedge\tau_n}|\mathcal{F}_s] = X_{s\wedge\tau_n}$. Then, the dominated convergence theorem implies

$$\begin{aligned}\mathbb{E}[X_t \mathbb{1}_A] &= \mathbb{E}\big[\lim_{n\to\infty} X_{t\wedge\tau_n}\mathbb{1}_A\big] = \lim_{n\to\infty}\mathbb{E}\big[X_{t\wedge\tau_n}\mathbb{1}_A\big] \\ &= \lim_{n\to\infty}\mathbb{E}\big[X_{s\wedge\tau_n}\mathbb{1}_A\big] = \mathbb{E}\big[X_s\mathbb{1}_A\big],\end{aligned}$$

for all $A \in \mathcal{F}_s$.

(iii) See problem sheets. □

Lemma 2.19. *Let $(X_t)_{t\in[0,T]}$ be a continuous local martingale with finite variation, i.e., $|X|_t < \infty$ $\mathbb{P}$-a.s. for all $t \in [0,T]$, and $X_0 = 0$. Then $X_t = 0$ for all $t \in [0,T]$ $\mathbb{P}$-a.s.*

Proof. The proof is left as an exercise: Use a suitable localizing sequence for $(X_t)_{t\in[0,T]}$ and Lemma 2.14. □

Lecture 5

In the next theorem we generalize the proposition (Proposition 2.15) regarding the quadratic variation of continuous martingales to the more general class of continuous local martingales.

Theorem 2.20. *Let $X = (X_t)_{t\in[0,T]}$ be a continuous local martingale. Then there exists a unique continuous process $\langle X\rangle = (\langle X\rangle_t)_{t\in[0,T]}$ with the following properties*

(i) $\langle X\rangle_0 = 0$ *and* $\langle X\rangle$ *is non-decreasing and*

(ii) $(X_t^2 - \langle X\rangle_t)_{t\in[0,T]}$ *is a continuous local martingale.*

Furthermore, $(\langle X\rangle_t)_{t\in[0,T]}$ *satisfies*

$$\langle X\rangle_t = \lim_{n\to\infty} \sum_{J\in\Pi_n} (\Delta_{J\cap[0,t]}X)^2 \quad \text{in probability for any } t \in [0,T],$$

where the limit is taken along any zero-sequence of partitions $(\Pi_n)_{n\in\mathbb{N}}$.

Proof. Existence and claimed properties: Apply Proposition 2.15 together with localizing sequence $(\tau_n)_{n\in\mathbb{N}}$, defined by

$$\tau_n := \inf\{t \in [0,T] \,:\, |X_t| \geq n\} \wedge T, \quad n \in \mathbb{N},$$

cf. Problem sheet 2.

Uniqueness: W.l.o.g. we may assume $X_0 = 0$. Let Y, Z be continuous processes satisfying (i) and (ii). Then

$$Y - Z = (X^2 - Z) - (X^2 - Y)$$

is a local martingale with $(Y - Z)_0 = 0$ and of locally finite variation

$$|Y - Z|_t \leqslant Y_t + Z_t < \infty \qquad \mathbb{P}\text{-a.s.}$$

since Y and Z are non-decreasing. Therefore, Lemma 2.19 yields $(Y - Z)_t = 0$ for all $t \in [0,T]$, $\mathbb{P}$-a.s. $\square$

Thanks to Theorem 2.20, we can make the following definition.

Definition 2.21. Let $X = (X_t)_{t\in[0,T]}$ be a continuous local martingale. The process $\langle X\rangle = (\langle X\rangle_t)_{t\in[0,T]}$, given by

$$\langle X\rangle_t = \lim_{n\to\infty} \sum_{J\in\Pi_n} (\Delta_{J\cap[0,t]}X)^2 \quad \text{in probability for any } t \in [0,T],$$

where the limit is taken along any zero-sequence of partitions $(\Pi_n)_{n\in\mathbb{N}}$, is called **quadratic variation** of $(X_t)_{t\in[0,T]}$.

Remark 2.22. Let $(X_t)_{t\in[0,T]}$ be a continuous local martingale with quadratic variation $(\langle X\rangle_t)_{t\in[0,T]}$. If τ is an a.s. finite stopping time, then the stopped process $X^\tau := (X_{t\wedge\tau})_{t\in[0,T]}$ has the quadratic variation $\langle X^\tau\rangle_t = \langle X\rangle_{t\wedge\tau}$ for $t \in [0,T]$. Indeed, $(\langle X\rangle_{t\wedge\tau})_{t\in[0,T]}$ satisfies (i) and $\big((X_{t\wedge\tau}^2 - \langle X\rangle)_{t\wedge\tau}\big)_{t\in[0,T]}$ is a local martingale thanks to optional stopping. Hence, by the uniqueness results of Theorem 2.20 , we get $\langle X^\tau\rangle_t = \langle X\rangle_{t\wedge\tau}$ for $t \in [0,T]$.

Since a Brownian motion $(B_t)_{t\in[0,T]}$ is a martingale, we can apply the results about quadratic variation to it.

Corollary 2.23. *Let $(B_t)_{t\in[0,T]}$ be a Brownian motion. Then, $(\langle B\rangle_t)_{t\in[0,T]}$ satisfies*

$$\langle B\rangle_t = t = \lim_{n\to\infty} \sum_{J\in\Pi_n} (\Delta_{J\cap[0,t]}B)^2 \quad \textit{in probability for any } t\in[0,T],$$

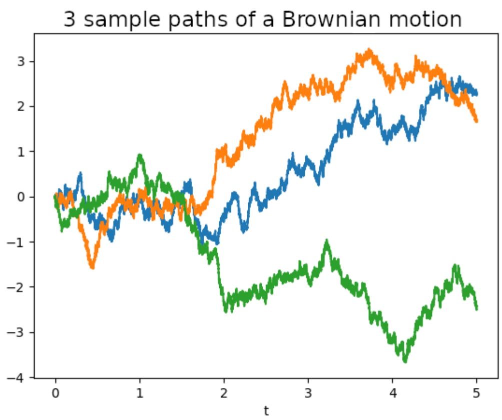
where the limit is taken along any zero-sequence of partitions $(\Pi_n)_{n\in\mathbb{N}}$.

Proof. We know that $(B_t)_{t\in[0,T]}$ and $(B_t^2 - t)_{t\in[0,T]}$ are continuous martingales, see Example 2.8, and the map $f\colon [0,T]\to[0,T]$, $t\mapsto f(t) := t$, is non-decreasing with $f(0)=0$. Hence, by Theorem 2.20 the quadratic variation of a Brownian motion is given by $\langle B\rangle_t = t$ and fulfills

$$\langle B\rangle_t = t = \lim_{n\to\infty} \sum_{J\in\Pi_n} (\Delta_{J\cap[0,t]}B)^2 \quad \text{in probability for any} \quad t\in[0,T],$$

where the limit is taken along any zero-sequence of partitions $(\Pi_n)_{n\in\mathbb{N}}$. □

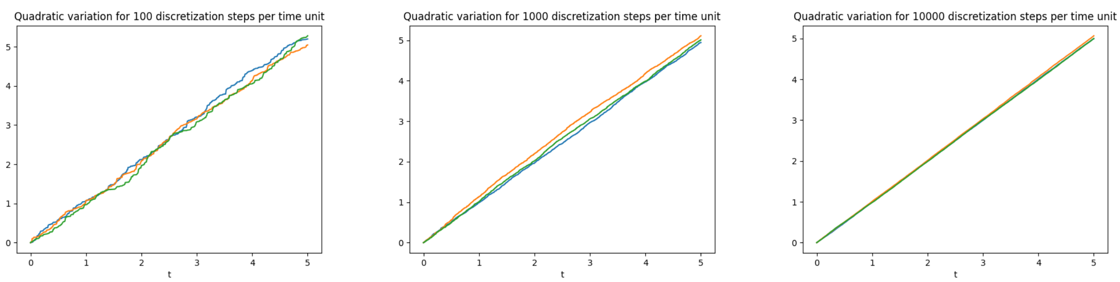
Let us illustrate Corollary 2.23 numerically. Simulating again three sample paths of a Brownian motion $(B_t)_{t\in[0,T]}$, we obtain the following figure:



When we numerically approximate the corresponding quadratic variation processes of the three illustrated sample paths of a Brownian motion, we see that,

$$\sum_{J\in\Pi_n} (\Delta_{J\cap[0,t]}B)^2 \to t, \quad t\in[0,T],$$

if the mesh size of the partitions Π_n goes to zero, as numerically represented in the following figures.



3 Stochastic Itô integration

In this chapter we shall develop stochastic integration with respect to a Brownian motion $(B_t)_{t\in[0,T]}$ and, in particular, introduce integrals of the form

$$\int_0^T f(t)\,\mathrm{d}B_t,$$

for a suitable class of stochastic processes $(f_t)_{t\in[0,T]}$. The presented stochastic integration theory was introduced by Kiyosi Itô in the 1940's. Recall that classical Lebesgue-Stieltjes integration does not apply to a Brownian motion $(B_t)_{t\in[0,T]}$ as its sample paths are not of finite variation, cf. Remark 2.13 and Lemma 2.14.

Let $(\Omega,\mathcal{F},\mathbb{P})$ be a complete probability space with a Brownian motion $B=(B_t)_{t\in[0,T]}$ and $(\mathcal{F}_t)_{t\in[0,T]}$ the associated Brownian standard filtration, i.e. $\mathcal{F}_t=\sigma(\mathcal{F}_t^B,\mathcal{N})$, $t\in[0,T]$, where $\mathcal{F}_t^B=\sigma(B_s, s\in[0,t])$ and $\mathcal{N}=\{A\in\mathcal{F}\,:\,\mathbb{P}(A)=0\}$. Recall that the filtration $(\mathcal{F}_t)_{t\in[0,T]}$ satisfies the usual conditions thanks to Proposition 2.5. Furthermore, <u>we fix $T\in(0,\infty)$.</u>

3.1 Construction of the Itô integral

As a first step in the construction of a stochastic integral with respect to a Brownian motion, we define the stochastic integral for "simple" integrands.

Definition 3.1.

(i) The set $\mathscr{H}_0^2$ of simple functions is defined as

$$\mathscr{H}_0^2:=\left\{f\colon\Omega\times[0,T]\to\mathbb{R}\;:\;\begin{array}{c}f(\omega,s)=\sum_{i=0}^{n-1}a_i(\omega)\mathbb{1}_{(t_i,t_{i+1}]}(s),\\ a_i \text{ is a } \mathcal{F}_{t_i}\text{-mb. r.v., } \mathbb{E}[a_i^2]<\infty, \text{ for}\\ 0=t_0<t_1<\dots<t_n=T \text{ and } 0\leqslant i\leqslant n-1, n\in\mathbb{N}\end{array}\right\}.$$

(ii) The **stochastic integral** of a simple function $f=\sum_{i=0}^{n-1}a_i\mathbb{1}_{(t_i,t_{i+1}]}\in\mathscr{H}_0^2$ is given by

$$I(f):=\sum_{i=0}^{n-1}a_i(B_{t_{i+1}}-B_{t_i}):=\int_0^T f(\cdot,s)\,\mathrm{d}B_s.$$

We consider the space $\mathscr{H}_0^2$ as a subspace of $L^2(\Omega\times[0,T],\mathbb{P}\otimes\lambda)$ with the norm

$$\|f\|_{\mathscr{H}^2}:=\|f\|_{L^2(\mathbb{P}\otimes\lambda)}:=\mathbb{E}\Big[\int_0^T f^2(\cdot,s)\,\mathrm{d}s\Big]^{1/2}\quad\text{for } f\in L^2(\Omega\times[0,T],\mathbb{P}\otimes\lambda).$$

On this space of simple functions, the integral operator $I\colon\mathscr{H}_0^2\to L^2(\Omega,\mathbb{P})$ turns out to be an isometry.

Lemma 3.2 (Itô's isometry - simple version)**.** *For $f\in\mathscr{H}_0^2$ we have*

$$\|f\|_{\mathscr{H}^2}=\|I(f)\|_{L^2}.$$

Proof. Let $f = \sum_{i=0}^{n-1} a_i \mathbb{1}_{(t_i,t_{i+1}]} \in \mathcal{H}_0^2$. Then

$$\|f\|_{\mathscr{H}^2}^2 = \mathbb{E}\Big[\int_0^T f^2(\cdot,s)\,\mathrm{d}s\Big] = \sum_{i=0}^{n-1} \mathbb{E}[a_i^2](t_{i+1}-t_i).$$

On the other hand by the independent and stationary increments of $(B_t)_{t\in[0,T]}$ and a_i being $\mathcal{F}_{t_i}$-measurable, we obtain

$$\begin{aligned}
\|I(f)\|_{L^2}^2 &= \mathbb{E}\big[I(f)^2\big] \\
&= \mathbb{E}\Big[\Big(\sum_{i=0}^{n-1} a_i(B_{t_{i+1}} - B_{t_i})\Big)^2\Big] \\
&= \mathbb{E}\Big[\sum_{i,j=0}^{n-1} a_i a_j (B_{t_{i+1}} - B_{t_i})(B_{t_{j+1}} - B_{t_j})\Big] \\
&= \sum_{i=0}^{n-1} \mathbb{E}[a_i^2 \mathbb{E}[(B_{t_{i+1}} - B_{t_i})^2|\mathcal{F}_{t_i}]] + 2\sum_{i,j=0,i<j}^{n-1} \mathbb{E}\Big[a_i a_j (B_{t_{i+1}} - B_{t_i})\mathbb{E}[(B_{t_{j+1}} - B_{t_j})|\mathcal{F}_{t_j}]\Big] \\
&= \sum_{i=0}^{n-1} \mathbb{E}[a_i^2]\mathbb{E}\big[(B_{t_{i+1}} - B_{t_i})^2\big] \\
&= \sum_{i=0}^{n-1} \mathbb{E}[a_i^2](t_{i+1} - t_i)
\end{aligned}$$

since $\mathbb{E}[(B_{t_{j+1}} - B_{t_j})|\mathcal{F}_{t_j}\big] = 0$ and $\mathbb{E}\big[(B_{t_{i+1}} - B_{t_i})^2\big] = t_{i+1} - t_i$. □

Remark. In Lemma 3.2 we actually have shown that

$$\mathbb{E}\Big[\int_0^T f^2(\cdot,s)\,\mathrm{d}\langle B\rangle_s\Big] = \mathbb{E}\Big[\Big(\int_0^T f(\cdot,s)\,\mathrm{d}B_s\Big)^2\Big] \quad \text{for } f \in \mathscr{H}_0^2$$

since $\langle B\rangle_s = s$ for $s \in [0,T]$. Replacing the Brownian motion $(B_t)_{t\in[0,T]}$ with a continuous martingale $(M_t)_{t\in[0,T]}$ would lead to

$$\mathbb{E}\Big[\int_0^T f^2(\cdot,s)\,\mathrm{d}\langle M\rangle_s\Big] = \mathbb{E}\Big[\Big(\int_0^T f(\cdot,s)\,\mathrm{d}M_s\Big)^2\Big]$$

assuming $\mathbb{E}\Big[\int_0^T f^2(\cdot,s)\,\mathrm{d}\langle M\rangle_s\Big] < \infty$ and $\sup_{t\in[0,T]} \mathbb{E}[M_t^2] < \infty$.

As a next step we want to extend the (continuous) integral operator to a wider class of integrands.

Definition 3.3.

- $f\colon \Omega \times [0,T] \to \mathbb{R}$ is called **measurable** if f is $(\mathcal{F} \otimes \mathcal{B}([0,T]), \mathcal{B}(\mathbb{R}))$-measurable.
- $f\colon \Omega \times [0,T] \to \mathbb{R}$ is **adapted** if $f(\cdot,t)$ is $\mathcal{F}_t$-measurable for all $t \in [0,T]$.
- The set of all adapted functions in $L^2(\Omega \times [0,T], \mathbb{P} \otimes \lambda)$ is given by

$$\mathscr{H}^2 := \Big\{f\colon \Omega \times [0,T] \to \mathbb{R} \,:\, f \text{ is measurable, adapted and } \mathbb{E}\Big[\int_0^T f(\cdot,s)^2\,\mathrm{d}s\Big] < \infty\Big\}.$$

Remark 3.4. To be precise, $\mathscr{H}^2$ is also considered as a subspace of $L^2(\Omega \times [0,T], \mathbb{P} \otimes \lambda)$ and thus is defined as the set of equivalence classes of all measurable and adapted processes f satisfying $\mathbb{E}\Big[\int_0^T f(\cdot,s)^2 \,\mathrm{d}s\Big] < \infty$. The other way around, each equivalence class in $\mathscr{H}^2$ contains an adapted and measurable representative.

Lecture 6

Proposition 3.5. *For every $f \in \mathscr{H}^2$ there exists a sequence $(f_n)_{n\in\mathbb{N}} \subseteq \mathscr{H}_0^2$ such that*

$$\|f_n - f\|_{\mathscr{H}^2} \to 0 \quad \text{as } n \to \infty.$$

Proof. Step 1: W.l.o.g. we may assume that f is bounded.

Set $f_n := -n \vee (f \wedge n)$ for $n \in \mathbb{N}$ and the dominated convergence theorem yields

$$\|f_n - f\|_{\mathscr{H}^2} \to 0 \qquad \text{as} \quad n \to \infty.$$

Note that f_n is bounded, measurable and adapted for every $n \in \mathbb{N}$.

Step 2: There is a sequence $(f_n)_{n\in\mathbb{N}}$ of bounded, measurable, adapted and continuous (but not necessarily simple) functions with

$$\|f_n - f\|_{\mathscr{H}^2} \to 0 \qquad \text{as} \quad n \to \infty.$$

For $t \in [0,T]$ and $n \in \mathbb{N}$ we define

$$f_n(\cdot,t) := n \int_{(t-1/n)_+}^t f(\cdot,s)\,\mathrm{d}s,$$

where $x_+ := 0 \vee x$. Then, f_n is bounded, measurable, adapted and continuous (but not necessarily simple) for $n \in \mathbb{N}$. For $\omega \in \Omega$ set

$$\begin{aligned} A &:= \big\{(\omega,t) \in \Omega \times [0,T] : \lim_{n\to\infty} f_n(\omega,t) \neq f(\omega,t)\big\}, \\ A_\omega &:= \big\{t \in [0,T] : (\omega,t) \in A\big\}. \end{aligned}$$

By the fundamental theorem of calculus, we have for $\omega \in \Omega$ that $\lambda(A_\omega) = 0$. Hence, $\mathbb{P}\otimes\lambda(A) = 0$. Dominated convergence implies $\|f_n - f\|_{\mathscr{H}^2} \to 0$.

Step 3: We have

$$\lim_{h\downarrow 0} \mathbb{E}\Big[\int_0^T \big|f(\cdot,t) - f(\cdot,(t-h)_+)\big|^2 \,\mathrm{d}t\Big] = 0.$$

Indeed, we have with the sequence $(f_n)_{n\in\mathbb{N}}$ from Step 2 and with the triangular inequality that

$$\begin{aligned} &\mathbb{E}\Big[\int_0^T \big|f(\cdot,t) - f(\cdot,(t-h)_+)\big|^2 \,\mathrm{d}t\Big]^{1/2} \\ &\quad\leqslant \|f - f_n\|_{\mathscr{H}^2} + \mathbb{E}\Big[\int_0^T \big|f_n(\cdot,(t-h)_+) - f(\cdot,(t-h)_+)\big|^2 \,\mathrm{d}t\Big]^{1/2} \\ &\qquad + \mathbb{E}\Big[\int_0^T \big|f_n(\cdot,t) - f_n(\cdot,(t-h)_+)\big|^2 \,\mathrm{d}t\Big]^{1/2} \\ &\quad\leqslant 2\underbrace{\|f - f_n\|_{\mathscr{H}^2}}_{\to 0,\, n\to\infty} + \underbrace{\sqrt{h}|f_n(\cdot,0) - f(\cdot,0)|}_{\to 0,\, h\downarrow 0} + \underbrace{\mathbb{E}\Big[\int_0^T \big|f_n(\cdot,t) - f_n(\cdot,(t-h)_+)\big|^2 \,\mathrm{d}t\Big]^{1/2}}_{\to 0,\, h\downarrow 0, \text{ by dom. conv., cont. of } f_n} \\ &\quad\to 0 \qquad \text{as} \quad h \downarrow 0 \quad \text{and} \quad n \to \infty, \end{aligned}$$

where we used Step 2 in the last line.

Step 4: There exists $(f_n)_{n\in\mathbb{N}} \subseteq \mathscr{H}_0^2$ *such that* $\|f_n - f\|_{\mathscr{H}^2} \to 0$ *as* $n \to \infty$.

Due to Step 1 and 2, it is sufficient to show the claim for f being continuous and bounded. For $n \in \mathbb{N}$ we introduce

$$\varphi_n \colon \mathbb{R} \to \Big\{\frac{j}{2^n} : j \in \mathbb{Z}\Big\}, \quad u \mapsto \sum_{j\in\mathbb{Z}} \frac{j-1}{2^n} \mathbb{1}_{(\frac{j-1}{2^n},\frac{j}{2^n}]}(u),$$

(as discretization of the time interval) and, for $n \in \mathbb{N}$, $s \in [0,1]$, we introduce

$$f_{n,s}(\omega,t) := f(\omega,(s+\varphi_n(t-s))_+) \quad \text{for } \omega\in\Omega,\ t\in[0,T].$$

By construction $-2^{-n} \le s+\varphi_n(t-s) \le T$ and $f_{n,s} \in \mathscr{H}_0^2$ for every $s \in [0,1]$. Moreover, we obtain

$$\begin{aligned}
&\mathbb{E}\Big[\int_0^T \int_0^1 |f_{n,s}(\cdot,t) - f(\cdot,t)|^2 \,\mathrm{d}s\,\mathrm{d}t\Big]\\
&= \mathbb{E}\Big[\int_0^T \int_0^1 |f(\cdot,(s+\varphi_n(t-s))_+) - f(\cdot,t)|^2 \,\mathrm{d}s\,\mathrm{d}t\Big]\\
&= \sum_{j\in\mathbb{Z}} \mathbb{E}\Big[\int_0^T \int_{\left[t-\frac{j}{2^n},t-\frac{j-1}{2^n}\right)\cap[0,1]} |f(\cdot,s+\tfrac{j-1}{2^n}) - f(\cdot,t)|^2 \,\mathrm{d}s\,\mathrm{d}t\Big]\\
&\leqslant (2^n+1)2^{-n} \int_{(0,1]} \underbrace{\mathbb{E}\Big[\int_0^T |f(\cdot,t-2^{-n}h) - f(\cdot,t)|^2 \,\mathrm{d}t\Big]}_{\to 0, n\to\infty \text{ (by Step 3)}} \mathrm{d}h\\
&\to 0 \quad \text{as } n\to\infty,
\end{aligned}$$

using the substitution $s+\frac{j-1}{2^n} = t-2^{-n}h$ and Fubini's theorem. Therefore, there is a sequence $(n_k)_{k\in\mathbb{N}} \subseteq \mathbb{N}$ such that $(\omega,s,t) \mapsto f_{n,s}(\omega,t)$ converges to $f(\omega,t)$ for $\mathbb{P}\otimes\lambda\otimes\lambda$-a.e. $(\omega,s,t) \in \Omega\times[0,1]\times[0,T]$. Fubini's theorem implies that there is some $s_0 \in [0,1]$ such that

$$f_{n_k,s_0}(\omega,t) \to f(\omega,t) \qquad \text{for } \mathbb{P}\otimes\lambda\text{-a.e. } (\omega,t)\in\Omega\times[0,T].$$

Finally, the dominated convergence theorem yields $\|f_{n_k,s_0} - f\|_{\mathcal{H}^2} \to 0$ for $k \to \infty$. □

Since $\mathscr{H}_0^2 \subseteq \mathscr{H}^2$ is dense (by Proposition 3.5) and $I \colon \mathscr{H}_0^2 \to L^2$ is a continuous operator (by Lemma 3.2), we can continuously extend the stochastic integral I from $\mathscr{H}_0^2$ to $\mathscr{H}^2$: For $f \in \mathscr{H}^2$ we can define

$$I\colon \mathscr{H}^2 \to L^2, \quad \text{via} \quad I(f) := \lim_{n\to\infty} I(f_n)$$

where the convergence of $I(f_n)$ takes place in L^2 for any sequence $(f_n)_{n\in\mathbb{N}} \subseteq \mathscr{H}_0^2$ such that $\|f_n - f\|_{\mathscr{H}^2} \to 0$ as $n\to\infty$.

Theorem 3.6 (Itô's isometry)**.** *For any* $f \in \mathscr{H}^2$ *we have* $\|f\|_{\mathscr{H}^2} = \|I(f)\|_{L^2}$.

Proof. Let $(f_n)_{n\in\mathbb{N}} \subseteq \mathscr{H}_0^2$ be such that $\|f-f_n\|_{\mathscr{H}^2} \to 0$ as $n\to\infty$. Then $\|I(f)-I(f_n)\|_{L^2} \to 0$ as $n\to\infty$. Therefore, by the simple version of Itô's isometry (by Lemma 3.2) we get

$$\|f\|_{\mathscr{H}^2} = \lim_{n\to\infty} \|f_n\|_{\mathscr{H}^2} = \lim_{n\to\infty} \|I(f_n)\|_{L^2} = \|I(f)\|_{L^2}.$$ □

Based on the integral operator I on the time interval $[0,T]$, we now want to study the corresponding integral process by integrating up to time $t \in [0,T]$. Note for $f \in \mathscr{H}^2$ we also have $f\mathbb{1}_{[0,t]} \in \mathscr{H}^2$.

Theorem 3.7. *For any $f \in \mathscr{H}^2$ there is a continuous martingale $X = (X_t)_{t\in[0,T]}$ with respect to $(\mathcal{F}_t)_{t\in[0,T]}$ such that for all $t \in [0,T]$:*

$$X_t = I(f\mathbb{1}_{[0,t]}) \qquad \mathbb{P}\text{-}a.s.$$

Proof. Step 1: The simple case.

Let $(f_n)_{n\in\mathbb{N}} \subseteq \mathscr{H}_0^2$ with $\|f - f_n\|_{\mathscr{H}^2} \to 0$ as $n \to \infty$, and

$$f_n = \sum_{i=0}^{m_n-1} a_i^n \mathbb{1}_{(t_i^n, t_{i+1}^n]}, \qquad n \in \mathbb{N}.$$

Define

$$X_t^n := I(f_n\mathbb{1}_{[0,t]}), \qquad t \in [0,T].$$

Then, for $t \in (t_k^n, t_{k+1}^n]$ with $k \in \{0,\dots,m_n-1\}$ we see that

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

Hence, for any $n \in \mathbb{N}$, $(X_t^n)_{t\in[0,T]}$ is a continuous martingale (check yourself!) and, for all $n,m \in \mathbb{N}$, $n \geqslant m$, $\varepsilon > 0$, Doob's maximal inequality and Itô isometry (Lemma 3.2) yield

$$\mathbb{P}\Big(\sup_{0\leqslant t\leqslant T} |X_t^n - X_t^m| \geqslant \varepsilon\Big) \leqslant \varepsilon^{-2}\mathbb{E}\big[|X_T^n - X_T^m|^2\big] = \varepsilon^{-2}\|f_n - f_m\|_{\mathscr{H}^2}^2.$$

Step 2: Show that there is a subsequence $(X^{n_k})_{k\in\mathbb{N}}$ such that X^{n_k} converges uniformly on $[0,T]$ to a continuous process X for $\mathbb{P}$-a.e. $\omega \in \Omega$.

Choose $(n_k)_{k\in\mathbb{N}}$ such that

$$\sup_{n\geqslant n_k} \|f_n - f_{n_k}\|_{\mathscr{H}^2}^2 \leqslant 2^{-3k}.$$

Then, for all $k \in \mathbb{N}$:

$$\mathbb{P}\Big(\sup_{0\leqslant t\leqslant T} |X_t^{n_{k+1}} - X_t^{n_k}| \geqslant 2^{-k}\Big) \leqslant 2^{2k}\|f_{n_{k+1}} - f_{n_k}\|_{\mathscr{H}^2}^2 \leqslant 2^{-k}.$$

By Borel-Cantelli's lemma there is an event $\Omega_0 \in \mathcal{F}$ with $\mathbb{P}(\Omega_0) = 1$ and there is a random variable $C > 0$ such that

$$\sup_{0\leqslant t\leqslant T} |X_t^{n_{k+1}} - X_t^{n_k}| \leqslant 2^{-k} \qquad \text{for all } k \geqslant C \text{ on } \Omega_0.$$

Therefore, for all $\omega \in \Omega_0$ we have that $(X^{n_k}(\omega))_{k\in\mathbb{N}}$ is a Cauchy sequence in $(C([0,T]), \|\cdot\|_\infty)$. Consequently, there is a continuous limit $(X_t)_{t\in[0,T]}$. On Ω_0^c we set $X \equiv 0$.

Step 3: Show that $(X_t)_{t\in[0,T]}$ is a $(\mathcal{F}_t)$-martingale.

Since $(X^{n_k})_{k\in\mathbb{N}}$ are adapted stochastic processes and $(\mathcal{F}_t)_{t\in[0,T]}$ is complete, $(X_t)_{t\in[0,T]}$ is adapted. For $t\in[0,T]$ we observe that

$$\begin{aligned}
\mathbb{E}\big[|X_t - X_t^{n_k}|^2\big] &\leq \liminf_{l\to\infty} \mathbb{E}\big[|X_t^{n_l} - X_t^{n_k}|^2\big] && \text{(Fatou's lemma)}\\
&= \lim_{l\to\infty} \|f_{n_l} - f_{n_k}\|^2_{\mathscr{H}^2} && \text{(Itô isometry)}\\
&= \|f - f_{n_k}\|^2_{\mathscr{H}^2} \to 0 \qquad \text{as } k\to\infty
\end{aligned}$$

and, in particular, $\mathbb{E}[|X_t|^2]<\infty$ for $t\in[0,T]$. For $0\leqslant s\leqslant t$, $A\in\mathcal{F}_s$ and $k\in\mathbb{N}$ we have

$$\begin{aligned}
\big|\mathbb{E}[(X_t - X_s)\mathbb{1}_A]\big| &\leqslant \big|\mathbb{E}[(X_t - X_t^{n_k})\mathbb{1}_A]\big| + \big|\underbrace{\mathbb{E}[(X_t^{n_k} - X_s^{n_k})\mathbb{1}_A]}_{=0 \text{ since } X^{n_k} \text{ is mart.}}\big| + \big|\mathbb{E}[(X_s^{n_k} - X_s)\mathbb{1}_A]\big|\\
&\leq \mathbb{E}\big[|X_t - X_t^{n_k}|^2\big]^{1/2} + \mathbb{E}\big[|X_s^{n_k} - X_s|^2\big]^{1/2} && \text{(Jensen's ineq.)}\\
&\leqslant 2\|f - f_{n_k}\|_{\mathscr{H}^2} \to 0 \quad \text{as } k\to\infty,
\end{aligned}$$

that is $\mathbb{E}[X_t|\mathcal{F}_s] = X_s$.

Step 4: Show that for $t\in[0,T]$ we have $X_t = I(f\mathbb{1}_{[0,t]})$ a.s.

For $k\in\mathbb{N}$ we have $X_t^{n_k} = I(f_{n_k}\mathbb{1}_{[0,t]})$ and

$$f_{n_k}\mathbb{1}_{[0,t]} \overset{\mathcal{H}^2}{\to} f\mathbb{1}_{[0,t]} \qquad \text{and} \qquad X_t^{n_k} \overset{L^2}{\to} X_t.$$

Therefore, Itô isometry reveals

$$X_t = \lim_{k\to\infty} X_t^{n_k} = \lim_{k\to\infty} I(f_{n_k}\mathbb{1}_{[0,t]}) = I(f\mathbb{1}_{[0,t]}), \quad \text{a.s.}$$

□

Due to the continuity, the process $(X_t)_{t\in[0,T]}$, with $X_t = I(f\mathbb{1}_{[0,t]})$, is not only adapted, but also $(\mathcal{F}\otimes\mathcal{B}([0,T]))$-measurable.

Definition 3.8. For any $f\in\mathscr{H}^2$ the **Itô integral** (or the **integral process)** is defined by

$$\int_0^t f(\cdot,s)\,\mathrm{d}B_s := X_t = I(f\mathbb{1}_{[0,t]}), \quad \mathbb{P}\text{-a.s.}, \quad t\in[0,T],$$

where $(X_t)_{t\in[0,T]}$ is defined as in Theorem 3.7.

Remark 3.9. The stochastic integral $\int_0^t f(\cdot,s)\,\mathrm{d}B_s$ is a linear operator on $\mathscr{H}^2$ since it is a linear operator on $\mathscr{H}_0^2$.

Another useful property of Itô integrals is that they preserve equality in the following sense:

Proposition 3.10. *Let $f\in\mathscr{H}^2$ and ν be a stopping time satisfying $f\mathbb{1}_{[0,\nu]}=0$. The integral process $X=(X_t)_{t\in[0,T]}$, with $X_t=\int_0^t f(\cdot,s)\,\mathrm{d}B_s$, then fulfills $X\mathbb{1}_{[0,\nu]}=0$. In particular, for two functions $f,g\in\mathscr{H}^2$ with $f\mathbb{1}_{[0,\nu]}=g\mathbb{1}_{[0,\nu]}$ the integral processes coincide on $[0,\nu]$.*

Proof. [The following proof is left as an exercise for your self-study.] First, consider $f \in \mathscr{H}_0^2$ and let X denote the integral process of f. We show $X\mathbb{1}_{[0,\nu]} = Y\mathbb{1}_{[0,\nu]}$ where Y denotes the integral process of $f\mathbb{1}_{[0,\nu]}$. Note that by linearity of the integral it suffices to consider

$$f = a\mathbb{1}_{(r,s]} \text{ for } 0 \le r < s \le T \text{ and } a \text{ an } \mathcal{F}_r\text{-measurable r.v. with } \mathbb{E}[a^2] < \infty.$$

We discretize the stopping time ν as follows:

$$\begin{aligned} s_{i,n} &:= r + (s-r)\frac{i}{2^n}, \quad i = 0, 1, \dots, 2^n, \\ \nu^n &:= \sum_{i=0}^{2^n-1} s_{i+1,n}\mathbb{1}_{(s_{i,n}, s_{i+1,n}]}(\nu). \end{aligned}$$

Then,

$$\begin{aligned} f\mathbb{1}_{[0,\nu^n]} &= f - f\mathbb{1}_{[\nu^n,T]} \\ &= f - f\sum_{i=0}^{2^n-1} \mathbb{1}_{(s_{i,n}, s_{i+1,n}]}(\nu)\mathbb{1}_{(s_{i+1,n},T]} \quad \in \mathscr{H}_0^2 \end{aligned}$$

and

$$Y_t^n := \int_0^t f(\cdot, s)\mathbb{1}_{[0,\nu^n]}(u)\,\mathrm{d}B_u = a(B_{s\wedge\nu^n\wedge t} - B_{r\wedge\nu^n\wedge t}).$$

Since B is continuous, it follows $Y_t = \lim_{n\to\infty} Y_t^n = a(B_{s\wedge\nu\wedge t} - B_{r\wedge\nu\wedge t})$. On the other hand, it holds $X_t = a(B_{s\wedge t} - B_{r\wedge t})$ which implies

$$X\mathbb{1}_{[0,\nu]} = Y\mathbb{1}_{[0,\nu]}. \tag{3.1}$$

□

Now, let $f \in \mathscr{H}^2$ and $(f_n)_{n\in\mathbb{N}} \subseteq \mathscr{H}_0^2$ such that $\|f_n - f\|_{\mathscr{H}^2} \to 0$ as $n \to \infty$ with $f_n\mathbb{1}_{[0,\nu]} = 0$, $n \in \mathbb{N}$. If X^n denotes the integral process of f_n and Y^n the integral process of $f_n\mathbb{1}_{[0,\nu]}$, it then holds

$$X\mathbb{1}_{[0,\nu]} = \lim_{n\to\infty} X^n\mathbb{1}_{[0,\nu]} \overset{(3.1)}{=} \lim_{n\to\infty} Y^n\mathbb{1}_{[0,\nu]} = 0. \tag{3.2}$$

This proves the first claim. Further, let $f, g \in \mathscr{H}^2$ with $f\mathbb{1}_{[0,\nu]} = g\mathbb{1}_{[0,\nu]}$ and note that $f - g \in \mathscr{H}^2$ with $(f-g)\mathbb{1}_{[0,\nu]} = 0$. We apply (3.2) to $f - g$ and conclude

$$\int_0^\cdot f\,\mathrm{d}B_s\mathbb{1}_{[0,\nu]} = \int_0^\cdot g\,\mathrm{d}B_s\mathbb{1}_{[0,\nu]},$$

i.e., the integral processes coincide on $\mathbb{1}_{[0,\nu]}$. Lecture 7

3.2 Extension of Itô integration via localization

So far we defined the stochastic integrals with respect to a Brownian motion for integrands in the space $\mathscr{H}^2$. Ideally, we would like to construct the stochastic integral

$$\int_0^t g(B_s)\,\mathrm{d}B_s, \quad t \in [0,T],$$

for <u>any</u> continuous function $g\colon \mathbb{R} \to \mathbb{R}$. However, choosing for instance $g(x) = e^{x^2/2}$, we see that $g(B_s) \notin \mathscr{H}^2$ since

$$\begin{aligned}\mathbb{E}[g(B_t)^2] &= \int_{\mathbb{R}} g(y)^2 \frac{1}{\sqrt{2\pi t}} \exp\big(-\tfrac{y^2}{2t}\big)\,\mathrm{d}y\\ &= \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi t}} \exp\big((1-\tfrac{1}{2t})y^2\big)\,\mathrm{d}y = \infty\end{aligned}$$

if $t \geq 1/2$. To ensure that $g(B_s)$ is always an admissible integrand, we extend the space $\mathscr{H}^2$ in the following definition.

Definition 3.11. For a fixed $T \in (0,\infty)$ we introduce

$$\mathscr{H}^2_{loc} := \Big\{ f\colon \Omega \times [0,T] \to \mathbb{R} : f \text{ is measurable, adapted and } \int_0^T f^2(\cdot,s)\,\mathrm{d}s < \infty\,\mathbb{P}\text{-a.s.}\Big\}.$$

A increasing sequence $(\nu_n)_{n\in\mathbb{N}}$ of $[0,T]$-valued stopping times is called **localizing sequence** for $f \in \mathscr{H}^2_{loc}$ if $f\mathbb{1}_{[0,\nu_n]} \in \mathscr{H}^2$ for all $n \in \mathbb{N}$ and

$$\mathbb{P}\Big(\bigcup_{n=1}^{\infty}\{\nu_n = T\}\Big) = 1.$$

Remark 3.12.

- $\mathscr{H}^2 \subseteq \mathscr{H}^2_{loc}$
- For any continuous $g\colon \mathbb{R} \to \mathbb{R}$ we have $f(\omega,t) = g(B_t(\omega)) \in \mathcal{H}^2_{loc}$ since B is a.s. pathwise bounded on $[0,T]$.

Proposition 3.13. *For every $f \in \mathscr{H}^2_{loc}$ there is a localizing sequence $(\nu_n)_{n\in\mathbb{N}}$.*

Proof. For $f \in \mathscr{H}^2_{loc}$ we define

$$\nu_n := \inf\Big\{t \in [0,T] \,:\, \int_0^t f^2(\cdot,s)\,\mathrm{d}s \geqslant n\Big\} \wedge T, \qquad n \in \mathbb{N},$$

Note, the sequence $(\nu_n)_{n\in\mathbb{N}}$ is increasing and $[0,T]$-valued.

Step 1: ν_n is a stopping time for every $n \in \mathbb{N}$.

Notice that ν_n is the hitting time of the stochastic process $(\int_0^t f^2(\cdot,s)\,\mathrm{d}s)_{t\in[0,T]}$ hitting the set $[0,\infty)$. Hence, to show that ν_n is a stopping time, it is sufficient to verify that $t \mapsto \int_0^t f^2(\cdot,s)\,\mathrm{d}s$ is adapted and continuous since we know that hitting times of adapted and continuous stochastic processes are stopping times.

For $f \in \mathscr{H}^2_0$ we can directly see that $t \mapsto \int_0^t f^2(\cdot,s)\,\mathrm{d}s$ is adapted and continuous and, thus, it is true for $f \in \mathscr{H}^2$ (by the uniform convergence of the associated integrals). Moreover, $t \mapsto \int_0^t f^2(\cdot,s)\,\mathrm{d}s$ is adapted and continuous for $f \in \mathscr{H}^2_{loc}$ as $f_m(\cdot,t) \to f(\cdot,t)$ as $m \to \infty$ for all $t \in [0,T]$, where $f_m := -m \vee (f \wedge m) \in \mathscr{H}^2$.

Step 2: ν_n is a localizing sequence for f.

For all $n \in \mathbb{N}$, we have by the continuity of $t \mapsto \int_0^t f^2(\cdot,s)\,\mathrm{d}s$ that

$$\|f\mathbb{1}_{[0,\nu_n]}\|^2_{\mathscr{H}^2} = \mathbb{E}\Big[\int_0^{\nu_n} f^2(\cdot,s)\,\mathrm{d}s\Big] \leqslant n,$$

that is $f\mathbb{1}_{[0,\nu_n]} \in \mathscr{H}^2$. Moreover,

$$\bigcup_{n\geqslant 1}\{\nu_n = T\} = \Big\{\int_0^T f^2(\cdot,s)\,\mathrm{d}s < \infty\Big\}$$

which has probability one. □

Definition 3.14. Let $f \in \mathscr{H}^2_{loc}$ and $(\nu_n)_{n\in\mathbb{N}}$ be a localizing sequence for f. The **Itô integral process** $(\int_0^t f(\cdot,s)\,\mathrm{d}B_s)_{t\in[0,T]}$ is defined as the continuous process $X = (X_t)_{t\in[0,T]}$ such that

$$\int_0^t f(\cdot,s)\,\mathrm{d}B_s := X_t = \lim_{n\to\infty}\int_0^t f(\cdot,s)\mathbb{1}_{[0,\nu_n]}(s)\,\mathrm{d}B_s \qquad \mathbb{P}\text{-a.s.} \qquad \text{for all } t\in[0,T],$$

where we recall that $f\mathbb{1}_{[0,\nu_n]} \in \mathscr{H}^2$.

We have to prove that the integral process $(\int_0^t f(\cdot,s)\,\mathrm{d}B_s)_{t\in[0,T]}$ is well-defined and has a continuous modification. Note first that Proposition 3.10 implies

$$\Big(\int_0^{\cdot} f(\cdot,s)\mathbb{1}_{[0,\nu_n]}(s)\,\mathrm{d}B_s\Big)\mathbb{1}_{[0,\nu_m]} = \Big(\int_0^{\cdot} f(\cdot,s)\mathbb{1}_{[0,\nu_m]}(s)\,\mathrm{d}B_s\Big)\mathbb{1}_{[0,\nu_m]} \quad \text{for } n\geqslant m.$$

Theorem 3.15. *For $f \in \mathscr{H}^2_{loc}$ there exists a continuous local martingale $(X_t)_{t\in[0,T]}$ such that for any localizing sequence $(\nu_n)_{n\in\mathbb{N}}$ of f it holds:*

$$\int_0^t f(\cdot,s)\mathbb{1}_{[0,\nu_n]}(s)\,\mathrm{d}B_s \to X_t \quad \text{as } n\to\infty \quad \mathbb{P}\text{-a.s.},$$

for $t\in[0,T]$. In particular, $(X_t)_{t\in[0,T]}$ does not depend on the choice of the localizing sequence $(\nu_n)_{n\in\mathbb{N}}$ and the Itô integral process $(\int_0^t f(\cdot,s)\,\mathrm{d}B_s)_{t\in[0,T]}$ is well-defined.

Proof. Let $f \in \mathscr{H}^2_{loc}$ and $(\nu_n)_{n\in\mathbb{N}}$ be a localizing sequence for f. For $n\in\mathbb{N}$ we define

$$X^n_t := \int_0^t f(\cdot,s)\mathbb{1}_{[0,\nu_n]}(s)\,\mathrm{d}B_s, \quad t\in[0,T].$$

Step 1: Existence of a continuous limit $(X_t)_{t\in[0,T]}$.
Let $N(\omega) := \min\{n\in\mathbb{N} : \nu_n(\omega) = T\}$. Note $N<\infty$ $\mathbb{P}$-a.s. Define

$$\Omega_0 := \big\{\omega\in\Omega : t\mapsto X^n_t(\omega) \text{ is continuous } \forall n\in\mathbb{N}\big\}$$

satisfying $\mathbb{P}(\Omega_0) = 1$. Set $\Omega_1 := \Omega_0 \cap \{N<\infty\}$ and

$$X_t(\omega) := \begin{cases} X^{N(\omega)}_t(\omega), & \omega\in\Omega_1,\\ 0, & \omega\notin\Omega_1,\end{cases} \qquad t\in[0,T].$$

Then, X is continuous and $X_t = \lim_{n\to\infty} X^n_t$ $\mathbb{P}$-a.s. for all $t\in[0,T]$.
Step 2: Independence of the localizing sequence.
Let $(\widetilde{\nu}_n)_{n\in\mathbb{N}}$ be another localizing sequence and write $\widetilde{X}^n_{\cdot} := \int_0^{\cdot} f(\cdot,s)\mathbb{1}_{[0,\widetilde{\nu}_n]}(s)\,\mathrm{d}B_s$. Setting $\tau_n := \nu_n \wedge \widetilde{\nu}_n$ for $n\in\mathbb{N}$, we have by Proposition 3.10:

$$X^n\mathbb{1}_{[0,\tau_m]} = \widetilde{X}^n\mathbb{1}_{[0,\tau_m]} \quad \text{for all } n\geqslant m.$$

Therefore, $\lim_{n\to\infty} X^n = \lim_{n\to\infty} \widetilde{X}^n$ on $[0, \tau_m]$ for any m. It remains to note that $\tau_m \uparrow T$.

Step 3: $(X_t)_{t\in[0,T]}$ is a local martingale.

It is sufficient to find a localizing sequence of stopping times. For this purpose, we introduce stopping times

$$\sigma_n := \inf\left\{t \in [0,T] \,:\, \int_0^t f^2(\cdot, s)\,\mathrm{d}s \geq n\right\} \wedge T \quad \text{for } n \in \mathbb{N}.$$

Note that $(\sigma_n)_{n\in\mathbb{N}}$ is a localizing sequence for $(X_t)_{t\in[0,T]}$. Indeed $\sigma_n \to T$ as $n \to \infty$ since $\int_0^T f^2(\cdot, s)\,\mathrm{d}s < \infty$ and

$$X_{t\wedge\sigma_n} = \int_0^t f(\cdot, s)\mathbb{1}_{[0,\sigma_n]}\,\mathrm{d}B_s, \quad t \in [0,T],$$

is a martingale since $f(\cdot, s)\mathbb{1}_{[0,\sigma_n]} \in \mathscr{H}^2$, see Theorem 3.7. □

We briefly generalize Proposition 3.10 regarding the persistence of identity of stochastic integrals.

Theorem 3.16 (Persistence of identity)**.** *Let $f, g \in \mathscr{H}^2_{loc}$ and ν be a stopping time such that $f\mathbb{1}_{[0,\nu]} = g\mathbb{1}_{[0,\nu]}$. Then*

$$\int_0^t f(\cdot, s)\,\mathrm{d}B_s\mathbb{1}_{[0,\nu]} = \int_0^t g(\cdot, s)\,\mathrm{d}B_s\mathbb{1}_{[0,\nu]}, \qquad \mathbb{P}-a.s.,$$

for $t \in [0,T]$.

Proof. For $n \in \mathbb{N}$ define the stopping time

$$\tau_n = \inf\left\{t \in [0,T] \,:\, \int_0^t f^2(\cdot, s)\,\mathrm{d}s \geqslant n \text{ or } \int_0^t g^2(\cdot, s)\,\mathrm{d}s \geqslant n\right\} \wedge T$$

and set

$$X^n := \int_0^\cdot f(\cdot, s)\mathbb{1}_{[0,\tau_n]}(s)dB_s \quad \text{and} \quad Y^n := \int_0^\cdot g(\cdot, s)\mathbb{1}_{[0,\tau_n]}(s)dB_s.$$

Then Proposition 3.10 implies that for all $n \in \mathbb{N}$

$$X^n\mathbb{1}_{[0,\nu]} = Y^n\mathbb{1}_{[0,\nu]}.$$

By Theorem 3.15, we conclude the assertion of the theorem. □

Coming back to the discussion of the beginning of the subsection, we want to investigate the Riemann sum approximation of the stochastic integral with integrands of the form $f(B_t) \in \mathscr{H}^2_{loc}$. Note that we have to use the left end point of the time intervals for the approximation of $f(B_t)$.

Theorem 3.17 (Riemann sum approximation)**.** *If $f\colon \mathbb{R} \to \mathbb{R}$ is a continuous function and $t_i := \frac{i}{n}T$, $0 \leqslant i \leqslant n$, then for $n \to \infty$ we have*

$$\sum_{i=1}^n f(B_{t_{i-1}})\big(B_{t_i} - B_{t_{i-1}}\big) \to \int_0^T f(B_s)\,\mathrm{d}B_s \qquad \textit{in probability.}$$

Proof. Step 1: Localization.

For $m \in \mathbb{N}$ we set

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

Then τ_m is a stopping time and $(\tau_m)_{m\in\mathbb{N}}$ is a localizing sequence for $f(B)$ because continuity of f yields $|f(B_{\cdot\wedge\tau_m})| \leqslant \sup_{|x|\leqslant m} |f(x)| < \infty$, that is $f(B) \in \mathscr{H}^2_{loc}$.

For all $m \in \mathbb{N}$ there is a continuous function f_m with compact support and $f_m|_{[-m,m]} = f|_{[-m,m]}$. Hence, we have

$$f(B) = f_m(B) \qquad \text{on } \{\tau_m = T\} \qquad \text{and} \qquad f_m(B) \in \mathscr{H}^2.$$

Step 2: Let $m \in \mathbb{N}$ be fixed. For $\varphi_n(\omega,s) := \sum_{i=1}^n f_m(B_{t_{i-1}}(\omega))\mathbb{1}_{(t_{i-1},t_i]}(s) \in \mathscr{H}^2_0$, $n \in \mathbb{N}$, we show $\varphi_n \to f_m(B)$ in $\mathscr{H}^2$ as $n \to \infty$.

We observe that

$$\begin{aligned}\mathbb{E}\Big[\int_0^T \big(\varphi_n(\cdot,s) - f_m(B_s)\big)^2\,\mathrm{d}s\Big] &= \sum_{i=1}^n \mathbb{E}\Big[\int_{t_{i-1}}^{t_i} \big(\underbrace{f_m(B_{t_{i-1}}) - f_m(B_s)}_{\leqslant \sup_{r\in[t_{i-1},t_i]} |f_m(B_{t_{i-1}}) - f_m(B_r)|}\big)^2\,\mathrm{d}s\Big] \\ &\leqslant \frac{T}{n}\sum_{i=1}^n \mathbb{E}\Big[\sup_{s\in[t_{i-1},t_i]} |f_m(B_{t_{i-1}}) - f_m(B_s)|^2\Big].\end{aligned}$$

For $h > 0$ we introduce the modulus of continuity of f_m:

$$\mu_{f_m}(h) := \sup\{|f_m(x) - f_m(y)| : x,y \in \mathbb{R} \text{ with } |x-y| \leqslant h\}. \tag{3.3}$$

Since f_m is uniformly continuous, we have $\mu_{f_m}(h) \to 0$ for $h \to 0$. Hence, we get

$$\begin{aligned}\mathbb{E}\Big[\int_0^T \big(\varphi_n(\cdot,s) - f_m(B_s)\big)^2\,\mathrm{d}s\Big] &\leqslant \frac{T}{n}\sum_{i=1}^n \mathbb{E}[\mu_{f_m}(\sup_{s\in[t_{i-1},t_i]} |B_s - B_{t_{i-1}}|)^2] \\ &\leqslant T\mathbb{E}[\mu_{f_m}(\underbrace{\sup_{s\in[t_{i-1},t_i],\, i=1,\dots,n} |B_s - B_{t_{i-1}}|}_{\to 0,\, n\to\infty})^2] \to 0, \quad n \to \infty,\end{aligned}$$

by dominated convergence (μ_{f_m} is bounded) and B is pathwise uniformly continuous on $[0,T]$. In particular, we have

$$\sum_{i=1}^n f_m(B_{t_{i-1}})\big(B_{t_i} - B_{t_{i-1}}\big) \to \int_0^T f_m(B_s)\,\mathrm{d}B_s \quad \text{in } L^2 \quad \text{as } n \to \infty.$$

Step 3: For all $\varepsilon > 0$ and

$$A_{n,\varepsilon} := \Big\{\Big|\sum_{i=1}^n f(B_{t_{i-1}})(B_{t_i} - B_{t_{i-1}}) - \int_0^T f(B_s)\,\mathrm{d}B_s\Big| \geqslant \varepsilon\Big\}$$

we show $\mathbb{P}(A_{n,\varepsilon}) \to 0$ for $n \to \infty$. [Keep in mind that the $(t_i)_{i=0}^n$ depend on n.]

On the set $\{\tau_m = T\}$ Theorem 3.16 yields

$$\int_0^T f(B_s)\,\mathrm{d}B_s = \int_0^T f_m(B_s)\,\mathrm{d}B_s.$$

Hence, we have

$$\begin{aligned}\mathbb{P}(A_{n,\varepsilon}) &= \mathbb{P}\big(A_{n,\varepsilon} \cap \{\tau_m < T\}\big) + \mathbb{P}\big(A_{n,\varepsilon} \cap \{\tau_m = T\}\big)\\ &\leqslant \underbrace{\mathbb{P}\big(\tau_m < T\big)}_{\to 0, m\to\infty \text{ (Step 1)}} + \underbrace{\mathbb{P}\big(A_{n,\varepsilon} \cap \{\tau_m = T\}\big)}_{\to 0, n\to\infty \text{ (Step 2).}}\\ &\to 0 \quad \text{as } n \to \infty,\end{aligned}$$

which completes the proof. □

Lecture 8

3.3 Itô formula for Brownian motion

As discussed in the Introduction, we expect that a stochastic integral might not satisfy the classical fundamental theorem of calculus. However, we still have a very much related fundamental theorem for the stochastic Itô integral available: *the Itô formula*.

Theorem 3.18 (Itô formula)**.** *For any twice continuously differentiable function $f\colon \mathbb{R} \to \mathbb{R}$ we have*

$$f(B_t) = f(0) + \int_0^t f'(B_s)\,\mathrm{d}B_s + \frac{1}{2}\int_0^t f''(B_s)\,\mathrm{d}s, \quad t \in [0,T], \quad \mathbb{P}\text{-}a.s.$$

Remark 3.19.

(i) Note the second integral which does not appear in the usual fundamental theorem of calculus. Thanks to the continuity of f and its derivatives and our extension of the integral domain to $\mathscr{H}^2_{loc}$ both integrals are well-defined.

(ii) Since the first integral has mean zero, the second integral captures all information about the drift of $(f(B_t))_{t\in[0,T]}$. On the other hand we will see that the first integral is most informative for the local variability of $(f(B_t))_{t\in[0,T]}$. Hence, we could understand Itô's formula as a decomposition of $f(B_t)$ into a "noise term" and a "signal".

(iii) Theorem 3.18 presents the simplest version of Itô's formula which we will generalize successively later.

Proof. W.l.o.g. we show Itô's formula for a fixed $t \in [0,T]$. Indeed, then Itô's formula holds for all $t \in [0,T] \cap \mathbb{Q}$, $\mathbb{P}$-a.s and, by continuity, also holds for all $t \in [0,T]$.

Step 1: Decomposition of $f(B_t) - f(0)$.

Fix $t \geqslant 0$ and set $t_i := \frac{i}{n}t$ for $i = 0,\dots,n$. Using a telescoping sum argument and a second order Taylor expansion at $B_{t_{i-1}}$ yields

$$\begin{aligned}f(B_t) - f(0) &= \sum_{i=1}^n \big(f(B_{t_i}) - f(B_{t_{i-1}})\big)\\ &= \underbrace{\sum_{i=1}^n f'(B_{t_{i-1}})(B_{t_i} - B_{t_{i-1}})}_{=:A_n} + \underbrace{\frac{1}{2}\sum_{i=1}^n f''(B_{t_{i-1}})(B_{t_i} - B_{t_{i-1}})^2}_{=:B_n} + \underbrace{\sum_{i=1}^n r(B_{t_{i-1}}, B_{t_i})}_{=:C_n},\end{aligned}$$

where $r(\cdot,\cdot)$ is the reminder of the Taylor expansion given by

$$\begin{aligned} r(x,y) &:= \int_x^y (y-u)(f''(u) - f''(x))\,\mathrm{d}u \\ &= (y-x)^2 \int_0^1 (1-t)\big(f''(x+t(y-x)) - f''(x)\big)\,\mathrm{d}t \end{aligned}$$

using the substitution $u = x + t(y-x)$ in the second equality.

Let us assume in the next steps (Step 2 to 4), that f has compact support. In this case we have

$$|r(x,y)| \leqslant |y-x|^2 |h(x,y)|$$

for some continuous, bounded function $h\colon \mathbb{R}^2 \to \mathbb{R}$ with $h(x,x) = 0$ for $x \in \mathbb{R}$ with compact support.

We will now consider the terms A_n, B_n and C_n separately and prove the convergence of A_n and B_n to the terms in Itô's formula and the convergence of C_n to 0 in probability. This implies that for all $t \in \mathbb{R}_+$ there is a subsequence $(n_k)_{k\in\mathbb{N}} \subseteq \mathbb{N}$ such that

$$(A_{n_k})_{k\in\mathbb{N}}, (B_{n_k})_{k\in\mathbb{N}}, (C_{n_k})_{k\in\mathbb{N}}$$

converge jointly with probability one.

Step 2: Show that $A_n \xrightarrow{\mathbb{P}} \int_0^t f'(B_s)\,\mathrm{d}B_s$ *as* $n \to \infty$.

Due to continuity of f, Theorem 3.17 yields the convergence.

Step 3: Show that $B_n \xrightarrow{\mathbb{P}} \frac{1}{2}\int_0^t f''(B_s)\,\mathrm{d}s$ *as* $n \to \infty$.

Recall our assumption that f has compact support. We write B_n as

$$\begin{aligned} B_n &= \frac{1}{2}\sum_{i=1}^n f''(B_{t_{i-1}})(t_i - t_{i-1}) + \frac{1}{2}\sum_{i=1}^n f''(B_{t_{i-1}})\big((B_{t_i} - B_{t_{i-1}})^2 - (t_i - t_{i-1})\big) \\ &=: B_{n,1} + B_{n,2}. \end{aligned}$$

Since f'' and B_t are (a.s.) continuous, we have

$$B_{n,1} \to \frac{1}{2}\int_0^t f''(B_s)\,\mathrm{d}s \quad \text{as } n \to \infty, \quad \mathbb{P}\text{-a.s.}$$

Moreover, since $B_{n,2}$ has mean 0 and by independence of Brownian increments, we have

$$\begin{aligned} \mathbb{E}\big[(B_{n,2})^2\big] &= \frac{1}{4}\sum_{i,j=1}^n \mathbb{E}\Big[f''(B_{t_{i-1}})\big((B_{t_i} - B_{t_{i-1}})^2 - (t_i - t_{i-1})\big) \\ &\qquad\qquad \times f''(B_{t_{j-1}})\big((B_{t_j} - B_{t_{j-1}})^2 - (t_j - t_{j-1})\big)\Big] \\ &= \frac{1}{4}\sum_{i=1}^n \mathbb{E}\Big[f''(B_{t_{i-1}})^2\big((B_{t_i} - B_{t_{i-1}})^2 - (t_i - t_{i-1})\big)^2\Big] \\ &\leqslant \frac{1}{4}\|f''\|_\infty^2 \sum_{i=1}^n \mathbb{E}\big[(B_{t_i} - B_{t_{i-1}})^2 - (t_i - t_{i-1})\big)^2\big] \\ &= \frac{1}{2}\|f''\|_\infty^2 \sum_{i=1}^n (t_i - t_{i-1})^2 \to 0 \quad \text{for } n \to \infty. \end{aligned}$$

where we note for the last line that $\mathbb{E}\big[(B_{t_i}-B_{t_{i-1}})^4\big]=3(t_i-t_{i-1})^2$, i.e. we have

$$\begin{aligned}
&\mathbb{E}\big[(B_{t_i}-B_{t_{i-1}})^2-2(B_{t_i}-B_{t_{i-1}})^2(t_i-t_{i-1})+(t_i-t_{i-1})^2\big]\\
&\quad=3(t_i-t_{i-1})^2-2(t_i-t_{i-1})^2+(t_i-t_{i-1})^2\\
&\quad=2(t_i-t_{i-1})^2.
\end{aligned}$$

Therefore, $B_{n,2}\to 0$ in L^2 and thus in probability, too.

Step 4: Show that $C_n \xrightarrow{\mathbb{P}} 0$ as $n\to\infty$.

Since h is uniformly continuous, for all $\varepsilon>0$ there is a $\delta>0$ such that

$$\forall x,y\in\mathbb{R} \text{ with } |x-y|\leqslant\delta: \qquad |h(x,y)|=|h(x,y)-h(x,x)|\leqslant\varepsilon.$$

Using Tschebychev's inequality, we obtain

$$\begin{aligned}
\mathbb{E}\big[h(B_{t_{i-1}},B_{t_i})^2\big]&=\mathbb{E}\big[h(B_{t_{i-1}},B_{t_i})^2\mathbb{1}_{\{|B_{t_i}-B_{t_{i-1}}|\leqslant\delta\}}\big]+\mathbb{E}\big[h(B_{t_{i-1}},B_{t_i})^2\mathbb{1}_{\{|B_{t_i}-B_{t_{i-1}}|>\delta\}}\big]\\
&\leqslant\varepsilon^2\mathbb{P}(|B_{t_i}-B_{t_{i-1}}|\leqslant\delta)+\|h\|_\infty^2\mathbb{P}(|B_{t_i}-B_{t_{i-1}}|>\delta)\\
&\leqslant\varepsilon^2+\|h\|_\infty^2\frac{\mathbb{E}[(B_{t_i}-B_{t_{i-1}})^2]}{\delta^2}\\
&=\varepsilon^2+\|h\|_\infty^2\frac{t_i-t_{i-1}}{\delta^2}\\
&=\varepsilon^2+\|h\|_\infty^2\frac{t}{\delta^2 n}.
\end{aligned}$$

Moreover, if f has compact support, then the Cauchy-Schwarz inequality yields

$$\begin{aligned}
\mathbb{E}[|C_n|]&\leqslant\mathbb{E}\Big[\sum_{i=1}^n|B_{t_i}-B_{t_{i-1}}|^2|h(B_{t_{i-1}},B_{t_i})|\Big]\\
&\leqslant\sum_{i=1}^n\mathbb{E}\big[(B_{t_i}-B_{t_{i-1}})^4\big]^{1/2}\mathbb{E}[h(B_{t_{i-1}},B_{t_i})^2]^{1/2}\\
&=\sqrt{3}\sum_{i=1}^n(t_i-t_{i-1})\mathbb{E}[h(B_{t_{i-1}},B_{t_i})^2]^{1/2}.
\end{aligned}$$

Therefore,

$$\mathbb{E}[|C_n|]\leqslant\sqrt{3}\sum_{i=1}^n(t_i-t_{i-1})\Big(\varepsilon^2+\|h\|_\infty^2\frac{t}{\delta^2 n}\Big)^{1/2}\to\sqrt{3}t\varepsilon \quad\text{as}\quad n\to\infty.$$

Since $\varepsilon>0$ was arbitrary, we have $C_n\to 0$ in L^1 and thus in probability.

Step 5: Localization.

Set

$$\tau_m:=\inf\big\{t\in[0,T]\,:\,|B_t|\geqslant m\big\}\wedge T, \qquad m\in\mathbb{N}.$$

Then τ_m is a stopping time and $\Omega=\bigcup_{m\geqslant1}\{\tau_m=T\}$. For all $m\in\mathbb{N}$ there is some $f_m\in C^2(\mathbb{R})$ with compact support and $f|_{[-m,m]}=f_m|_{[-m,m]}$. Then Itô's formula is true for f_m. Due to Theorem 3.16 the claim is true on $\{\tau_m=T\}$ (since we have $f_m(B_t)=f(B_t)$ on this event) and thus $\mathbb{P}$-a.s. on Ω. □

Example 3.20. Let us take for example $f(x) = \frac{x^2}{2}$. Then, Itô's formula yields

$$\begin{aligned}\frac{1}{2}B_t^2 &= f(B_t) - f(0)\\ &= \int_0^t f'(B_s)\,\mathrm{d}B_s + \frac{1}{2}\int_0^t f''(B_s)\,\mathrm{d}s\\ &= \int_0^t B_s\,\mathrm{d}B_s + \frac{t}{2}.\end{aligned}$$

In particular, $\frac{1}{2}(B_t^2 - t) = \int_0^t B_s\,\mathrm{d}B_s$ is a martingale, cf. problem sheets.

In order to take functions into account which depend on B_t and on t, we need the following generalization of Theorem 3.18. We denote by $C^{1,2}([0,T]\times\mathbb{R})$ the space of continuous functions $f\colon [0,T]\times\mathbb{R} \ni (t,x) \mapsto f(t,x) \in \mathbb{R}$ such that $f(t,x)$ is continuously differentiable in $t \in (0,T)$ and twice continuously differentiable in $x \in \mathbb{R}$.

Theorem 3.21 (Itô formula, space-time version)**.** *For any $f \in C^{1,2}([0,T]\times\mathbb{R})$ we have*

$$f(t,B_t) = f(0,0) + \int_0^t \frac{\partial f}{\partial t}(s,B_s)\,\mathrm{d}s + \int_0^t \frac{\partial f}{\partial x}(s,B_s)\,\mathrm{d}B_s + \frac{1}{2}\int_0^t \frac{\partial^2 f}{\partial x^2}(s,B_s)\,\mathrm{d}s,$$

for $t \in [0,T]$, $\mathbb{P}$-a.s.

Sketch of the proof. We generalize the proof of Theorem 3.18 using the following (first and second) Taylor expansion for the telescoping sum

$$\begin{aligned}f(t,B_t) - f(0,0) &= \sum_{i=1}^n f(t_i,B_{t_i}) - f(t_{i-1},B_{t_{i-1}})\\ &= \sum_{i=1}^n \big(f(t_i,B_{t_i}) - f(t_{i-1},B_{t_i})\big) + \sum_{i=1}^n \big(f(t_{i-1},B_{t_i}) - f(t_{i-1},B_{t_{i-1}})\big)\\ &= \sum_{i=1}^n \frac{\partial f}{\partial t}(t_{i-1},B_{t_i})(t_i - t_{i-1})\\ &\quad + \sum_{i=1}^n \frac{\partial f}{\partial x}(t_{i-1},B_{t_{i-1}})(B_{t_i} - B_{t_{i-1}}) + \frac{1}{2}\sum_{i=1}^n \frac{\partial^2 f}{\partial x^2}(t_{i-1},B_{t_{i-1}})(B_{t_i} - B_{t_{i-1}})^2\\ &\quad + \sum_{i=1}^n r(t_{i-1},t_i,B_{t_{i-1}},B_{t_i})\end{aligned}$$

with

$$r(s,t,x,y) = \int_s^t \Big(\frac{\partial f}{\partial t}(r,y) - \frac{\partial f}{\partial t}(s,y)\Big)\,\mathrm{d}r + \int_x^y (y-u)\Big(\frac{\partial^2 f}{\partial x^2}(s,u) - \frac{\partial^2 f}{\partial x^2}(s,x)\Big)\,\mathrm{d}u.$$

If f has compact support, we have

$$|r(s,t,x,y)| \leqslant |t-s||k(s,t,y)| + |y-x|^2|h(s,x,y)|,$$

where h and k are uniformly continuous, bounded functions on $\mathbb{R}_+^2 \times \mathbb{R}$ and $\mathbb{R}_+ \times \mathbb{R}^2$, respectively, which are 0 if $s = t$ and $x = y$, respectively. The rest of the proof is analogous to the proof of Itô formula without time dependency (Theorem 3.18). □

4 Itô processes

Lecture 9

So far, we considered stochastic integration and Itô's formula only with respect to a Brownian motion. The purpose of this chapter is to generalize Itô's theory to a larger class of stochastic processes as feasible integrators. As before, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space, $(B_t)_{t\in[0,T]}$ be a Brownian motion with its completed filtration $(\mathcal{F}_t)_{t\in[0,T]}$ and $T \in (0,\infty)$.

Definition 4.1. A stochastic process $X = (X_t)_{t\in[0,T]}$ is called **Itô process** if there is a ($\mathbb{P}$-a.s.) representation

$$X_t = X_0 + \int_0^t a(\cdot, s)\,\mathrm{d}s + \int_0^t b(\cdot, s)\,\mathrm{d}B_s, \qquad \text{for } t \in [0,T],$$

where $X_0 \in \mathbb{R}$ and $a, b\colon \Omega \times [0,T] \to \mathbb{R}$ are adapted, measurable processes satisfying the integrability conditions

$$\mathbb{P}\Big(\int_0^T |a(\omega, s)|\,\mathrm{d}s < \infty\Big) = 1 \qquad \text{and} \qquad \mathbb{P}\Big(\int_0^T |b(\omega, s)|^2\,\mathrm{d}s < \infty\Big) = 1.$$

Note that $b \in \mathscr{H}^2_{loc}$ and thus the stochastic integral in the representation of X is well-defined.

Remark. Let $a\colon \Omega \times [0,T] \to \mathbb{R}$ be an adapted, measurable process satisfying $\mathbb{P}(\int_0^T |a(\omega,s)|\,\mathrm{d}s < \infty) = 1$, as it is assumed in the definition of Itô processes. Then, the stochastic process $(\int_0^t a(\cdot,s)\,\mathrm{d}s)_{t\in[0,T]}$ is a continuous, adapted process of finite variation.

Furthermore, the representation of an Itô process is unique in the sense that the representing functions a and b are uniquely determined for the Itô process X, which is the content of the next lemma.

Lemma 4.2. *Let $X = (X_t)_{t\in[0,T]}$ be an Itô process with representations*

$$X_t = X_0 + \int_0^t a(\cdot, s)\,\mathrm{d}s + \int_0^t b(\cdot, s)\,\mathrm{d}B_s = \widetilde{X}_0 + \int_0^t \widetilde{a}(\cdot, s)\,\mathrm{d}s + \int_0^t \widetilde{b}(\cdot, s)\,\mathrm{d}B_s$$

for $t \in [0,T]$ and $X_0 = \widetilde{X}_0$. Then, we have $a = \widetilde{a}$ and $b = \widetilde{b}$, $\mathbb{P}\otimes\lambda$-a.s.

Proof. See problem sheets. □

4.1 Itô's integration for Itô processes

The quadratic variation plays a central role in Itô's integration theory, for instance, in Itô's formula for a Brownian motion we used it to treat the second order term in the Taylor expansion. Therefore, we start by extending the notation of quadratic variation to Itô processes and calculate it.

Proposition 4.3. *For any Itô process $X = (X_t)_{t\in[0,T]}$ with representation*

$$X_t = X_0 + \int_0^t a(\cdot, s)\,\mathrm{d}s + \int_0^t b(\cdot, s)\,\mathrm{d}B_s, \quad t \in [0,T],$$

the ***quadratic variation*** *of X is given by*

$$\langle X\rangle_t := \lim_{n\to\infty}\sum_{J\in\Pi_n}(\Delta_{J\cap[0,t]}X)^2 = \int_0^t b^2(\cdot,s)\,\mathrm{d}s, \quad t\in[0,T],$$

where the limit is taken in probability.

Proof. We decompose $X = M + A$ with

$$M_t := \int_0^t b(\cdot,s)\,\mathrm{d}B_s \qquad \text{and} \qquad A_t := X_0 + \int_0^t a(\cdot,s)\,\mathrm{d}s.$$

Note that M is a local martingale and A has finite variation.

Step 1: We show $\langle M\rangle = \int_0^\cdot b^2(\cdot,s)\,\mathrm{d}s$.

Take a localizing sequence $(\tau_n)_{n\in\mathbb{N}}$ such that $M_{\cdot\wedge\tau_n}$ is a martingale and $\int_0^{t\wedge\tau_n} b^2(\cdot,s)\,\mathrm{d}s$ is uniformly bounded in t. Since $\int_0^\cdot b^2(\cdot,s)\,\mathrm{d}s$ starts in 0 and is increasing, it suffices to verify that

$$\left(M_{t\wedge\tau_n}^2 - \int_0^{t\wedge\tau_n} b^2(\cdot,s)\,\mathrm{d}s\right)_{t\in[0,T]}$$

is a martingale, see Theorem 2.20. To that end, we need to show the martingale property (the rest is easy to see), i.e. for $0\leqslant s<t\leq T$ we need to show

$$\begin{aligned} 0 \overset{!}{=} &\mathbb{E}\Big[M_{t\wedge\tau_n}^2 - M_{s\wedge\tau_n}^2 - \int_{s\wedge\tau_n}^{t\wedge\tau_n} b^2(\cdot,s)\,\mathrm{d}s\Big|\mathcal{F}_s\Big] \\ =&\mathbb{E}\Big[(M_{t\wedge\tau_n} - M_{s\wedge\tau_n})^2 - \int_{s\wedge\tau_n}^{t\wedge\tau_n} b^2(\cdot,s)\,\mathrm{d}s\Big|\mathcal{F}_s\Big], \end{aligned}$$

using that $M_{\cdot\wedge\tau_n}$ is a martingale in last equality. For any $A\in\mathcal{F}_s$ we have

$$\begin{aligned} (M_{t\wedge\tau_n} - M_{s\wedge\tau_n})\mathbb{1}_A &= (M_{t\wedge\tau_n} - M_s)\mathbb{1}_A\mathbb{1}_{\{\tau_n\geqslant s\}} \\ &= \int_0^T \underbrace{\underbrace{\mathbb{1}_A\mathbb{1}_{\{\tau_n\geqslant s\}}\mathbb{1}_{(s,t]}(r)}_{\in\mathcal{H}_0^2}\, b(\cdot,r)\mathbb{1}_{[0,\tau_n]}(r)}_{\in\mathcal{H}^2}\ \mathrm{d}B_r. \end{aligned}$$

Therefore, Itô's isometry implies

$$\begin{aligned} \mathbb{E}\Big[(M_{t\wedge\tau_n} - M_{s\wedge\tau_n})^2\mathbb{1}_A\Big] &= \mathbb{E}\Big[\Big(\int_0^T \mathbb{1}_A\mathbb{1}_{(s\wedge\tau_n,t\wedge\tau_n]}(r)b(\cdot,r)\,\mathrm{d}B_r\Big)^2\Big] \\ &= \mathbb{E}\Big[\int_0^T \big(\mathbb{1}_A\mathbb{1}_{(s\wedge\tau_n,t\wedge\tau_n]}(r)b(\cdot,r)\big)^2\,\mathrm{d}r\Big] \\ &= \mathbb{E}\Big[\mathbb{1}_A\int_{s\wedge\tau_n}^{t\wedge\tau_n} b^2(\cdot,r)\,\mathrm{d}r\Big]. \end{aligned}$$

Step 2: We deduce $\langle X\rangle = \int_0^\cdot b^2(\cdot,s)\,\mathrm{d}s$.

Let $t>0$ and $(\Pi_n)_{n\in\mathbb{N}}$ be a zero-sequence of partitions. W.l.o.g. t is a partitioning point in Π_1 and $J_n = \{J\in\Pi_n : J\subseteq[0,t]\}$. Then it is sufficient to show that

$$\lim_{n\to\infty}\sum_{J\in J_n}(\Delta_J X)^2 = \int_0^t b^2(\cdot,s)\,\mathrm{d}s \qquad \text{in probability.}$$

We decompose

$$\sum_{J\in J_n}(\Delta_J X)^2=\sum_{J\in J_n}(\Delta_J M)^2+\sum_{J\in J_n}(\Delta_J A)^2+2\sum_{J\in J_n}(\Delta_J M)(\Delta_J A).$$

By Step 1 we know that

$$\lim_{n\to\infty}\sum_{J\in J_n}(\Delta_J M)^2=\int_0^t b^2(\cdot,s)\,\mathrm{d}s$$

in probability. Moreover,

$$\sum_{J\in J_n}(\Delta_J A)^2\leqslant\sup_{J\in J_n}|\Delta_J A|\sum_{J\in J_n}|\Delta_J A|\leqslant\sup_{J\in J_n}|\Delta_J A|\,|A|_t\to 0\qquad\mathbb{P}\text{-a.s.}$$

since A is of finite variation and continuous and $|\Pi_n|\to 0$. The Cauchy-Schwarz inequality yields finally

$$\sum_{J\in J_n}(\Delta_J M)(\Delta_J A)\leqslant\underbrace{\Big(\sum_{J\in J_n}(\Delta_J M)^2\Big)}_{\xrightarrow{\mathbb{P}}\langle M\rangle_t}{}^{1/2}\underbrace{\Big(\sum_{J\in J_n}(\Delta_J A)^2\Big)}_{\to 0\,\mathbb{P}\text{-a.s.}}{}^{1/2}\xrightarrow{\mathbb{P}}0.$$

□

To extend Itô integration with respect to a Brownian motion to the class of Itô processes, let us do a heuristic calculation. Let $f\colon\Omega\times[0,T]\to\mathbb{R}$ be a continuous, adapted, bounded function. Considering a Riemann sum approximation as in Theorem 3.17, we heuristically arrive at

$$\begin{aligned}&\sum_{i=1}^n f(\cdot,t_{i-1})(X_{t_i}-X_{t_{i-1}})\\&\quad=\sum_{i=1}^n f(\cdot,t_{i-1})\Big(\int_{t_{i-1}}^{t_i}a(\cdot,s)\,\mathrm{d}s+\int_{t_{i-1}}^{t_i}b(\cdot,s)\,\mathrm{d}B_s\Big)\\&\quad=\int_0^t\underbrace{\sum_{i=1}^n f(\cdot,t_{i-1})a(\cdot,s)\mathbb{1}_{(t_{i-1},t_i]}(s)}_{\to f(\cdot,s)a(\cdot,s)}\,\mathrm{d}s+\int_0^t\underbrace{\sum_{i=1}^n f(\cdot,t_{i-1})b(\cdot,s)\mathbb{1}_{(t_{i-1},t_i]}(s)}_{\to f(\cdot,s)b(\cdot,s)}\,\mathrm{d}B_s\end{aligned}$$

for $0=t_0<t_1<\dots<t_n=t$. This leads us to the following natural definition:

Definition 4.4. Let $X=(X_t)_{t\in[0,T]}$ be an Itô process with representation $X_t=X_0+\int_0^t a(\cdot,s)\,\mathrm{d}s+\int_0^t b(\cdot,s)\,\mathrm{d}B_s$ for $t\in[0,T]$. We write $\mathcal{L}(X)$ for all adapted, measurable functions $f\colon\Omega\times[0,T]\to\mathbb{R}$ satisfying

$$\int_0^T|f(\cdot,s)a(\cdot,s)|\,\mathrm{d}s<\infty\qquad\text{and}\qquad\int_0^T|f(\cdot,s)b(\cdot,s)|^2\,\mathrm{d}s<\infty\qquad\mathbb{P}\text{-a.s.}$$

For $f\in\mathcal{L}(X)$ we define the **stochastic Itô integral** by

$$\int_0^t f(\cdot,s)\,\mathrm{d}X_s:=\int_0^t f(\cdot,s)a(\cdot,s)\,\mathrm{d}s+\int_0^t f(\cdot,s)b(\cdot,s)\,\mathrm{d}B_s,\quad t\in[0,T].$$

Note that the stochastic Itô integral is well-defined as the $f(\cdot,s)b(\cdot,s)\in\mathscr{H}^2_{loc}$ and the representation of the Itô process X is unique (see Lemma 4.2).

Lemma 4.5. *Let $X = (X_t)_{t\in[0,T]}$ be an Itô process with representation $X_t = X_0 + \int_0^t b(\cdot,s)\,\mathrm{d}B_s$ for $t \in [0,T]$, that is $a = 0$, and $f \in \mathcal{L}(X)$.*

(i) The integral process $(\int_0^t f(\cdot,s)\,\mathrm{d}X_s)_{t\in[0,T]}$ is a continuous local martingale.

(ii) If $\mathbb{E}\big[\int_0^T f^2(\cdot,s)b^2(\cdot,s)\,\mathrm{d}s\big] < \infty$, then Itô isometry holds true:

$$\mathbb{E}\bigg[\Big(\int_0^t f(\cdot,s)\,\mathrm{d}X_s\Big)^2\bigg] = \mathbb{E}\bigg[\int_0^t f^2(\cdot,s)\,\mathrm{d}\langle X\rangle_s\bigg], \quad t \in [0,T].$$

Remark 4.6. Since $(\langle X\rangle_t)_{t\in[0,T]}$ is a non-decreasing process, $\int_0^t f^2(\cdot,s)\,\mathrm{d}\langle X\rangle_s$ is just a Lebesgue-Stieltjes integral and by the associativity of Lebesgue-Stieltjes integration we have

$$\int_0^t f^2(\cdot,s)\,\mathrm{d}\langle X\rangle_s = \int_0^t f^2(\cdot,s)b^2(\cdot,s)\,\mathrm{d}s, \quad t \in [0,T].$$

Proof. (i) For $f \in \mathcal{L}(X)$ we know that $f(\cdot,s)b(\cdot,s) \in \mathscr{H}^2_{loc}$ by the definition of $\mathcal{L}(X)$. Hence, $(\int_0^t f(\cdot,s)\,\mathrm{d}X_s)_{t\in[0,T]}$ is a continuous local martingale by Theorem 3.15.

(ii) If $\mathbb{E}\big[\int_0^T f^2(\cdot,s)b^2(\cdot,s)\,\mathrm{d}s\big] < \infty$, we can apply Itô isometry (Theorem 3.6) for $\mathscr{H}^2$ leading to

$$\begin{aligned}\mathbb{E}\bigg[\Big(\int_0^t f(\cdot,s)\,\mathrm{d}X_s\Big)^2\bigg] &= \mathbb{E}\bigg[\Big(\int_0^T f(\cdot,s)b(\cdot,s)\mathbb{1}_{[0,t]}\,\mathrm{d}B_s\Big)^2\bigg] \\ &= \mathbb{E}\bigg[\int_0^T f^2(\cdot,s)b^2(\cdot,s)\mathbb{1}_{[0,t]}\,\mathrm{d}s\bigg] \\ &= \mathbb{E}\bigg[\int_0^t f^2(\cdot,s)\,\mathrm{d}\langle X\rangle_s\bigg],\end{aligned}$$

where we used Proposition 4.3 and the associativity of Riemann integration. □

4.2 Itô's formula for Itô processes

Lecture 10

In this subchapter we want to derive an Itô formula for Itô processes. The derivation is based on the same ideas as in the simpler case of an Itô formula for a Brownian motion, cf. Theorem 3.18.

Let $f\colon \mathbb{R}^2 \to \mathbb{R}$ be a twice continuously differentiable function (i.e. $f \in C^2(\mathbb{R}^2)$) and X, Y be two Itô processes with representations

$$X_t = X_0 + \int_0^t a(\cdot,s)\,\mathrm{d}s + \int_0^t b(\cdot,s)\,\mathrm{d}B_s \quad \text{and} \quad Y_t = Y_0 + \int_0^t \alpha(\cdot,s)\,\mathrm{d}s + \int_0^t \beta(\cdot,s)\,\mathrm{d}B_s,$$

for $t \in [0,T]$. For a moment, we assume that $\int_0^T \big(|a(\cdot,s)| + |\alpha(\cdot,s)|\big)\,\mathrm{d}s$ and $\int_0^T \big(b^2(\cdot,s) + \beta^2(\cdot,s)\big)\,\mathrm{d}s$ are bounded (Using a localization argument, the general situation can always be reduced to this case.). As in the proof of Theorem 3.18, relying on a telescoping sum argument and a (multi-dimensional) Taylor expansion, $t_i := t\frac{i}{n}$, $i = 0, \dots, n$, we obtain

$$
\begin{aligned}
&f(X_t, Y_t) - f(X_0, Y_0) \\
&= \sum_{i=1}^{n} \big(f(X_{t_i}, Y_{t_i}) - f(X_{t_{i-1}}, Y_{t_{i-1}}) \big) \\
&= \sum_{i=1}^{n} f_x(X_{t_{i-1}}, Y_{t_{i-1}})(X_{t_i} - X_{t_{i-1}}) + \sum_{i=1}^{n} f_y(X_{t_{i-1}}, Y_{t_{i-1}})(Y_{t_i} - Y_{t_{i-1}}) \\
&\quad + \frac{1}{2} \sum_{i=1}^{n} f_{xx}(X_{t_{i-1}}, Y_{t_{i-1}})(X_{t_i} - X_{t_{i-1}})^2 + \frac{1}{2} \sum_{i=1}^{n} f_{yy}(X_{t_{i-1}}, Y_{t_{i-1}})(Y_{t_i} - Y_{t_{i-1}})^2 \\
&\quad + \sum_{i=1}^{n} f_{xy}(X_{t_{i-1}}, Y_{t_{i-1}})(X_{t_i} - X_{t_{i-1}})(Y_{t_i} - Y_{t_{i-1}}) + \sum_{i=1}^{n} r(X_{t_{i-1}}, X_{t_i}, Y_{t_{i-1}}, Y_{t_i}) \\
&=: A_n^1 + A_n^2 + B_n^1 + B_n^2 + C_n + R_n, \qquad (4.1)
\end{aligned}
$$

with $|r(x, y)| \leqslant |y - x|^2 |h(x, y)|$ for some bounded, continuous function $h \colon \mathbb{R}^4 \to \mathbb{R}$ with $h(x, x) = 0$ for all $x, y \in \mathbb{R}^2$, and

$$
f_x := \frac{\partial}{\partial x} f, \quad f_y := \frac{\partial}{\partial y} f, \quad f_{xx} := \frac{\partial^2}{\partial x^2} f, \quad f_{yy} := \frac{\partial^2}{\partial y^2} f, \quad \text{and} \quad f_{xy} := \frac{\partial^2}{\partial x \partial y} f.
$$

In the following we will investigate the convergence of the five terms A_n^1, A_n^2, B_n^1, B_n^2, C_n and R_n.

First we observe that Itô's isometry yields

$$
\begin{aligned}
A_n^1 &= \underbrace{\int_0^t \sum_{i=1}^{n} f_x(X_{t_{i-1}}, Y_{t_{i-1}}) a(\cdot, s) \mathbb{1}_{(t_{i-1}, t_i]}(s) \, \mathrm{d}s}_{\to \int_0^t f_x(X_s, Y_s) a(\cdot, s) \, \mathrm{d}s \, \mathbb{P}\text{-a.s. (dom. conv.)}} + \int_0^t \underbrace{\sum_{i=1}^{n} f_x(X_{t_{i-1}}, Y_{t_{i-1}}) b(\cdot, s) \mathbb{1}_{(t_{i-1}, t_i]}(s)}_{\to f_x(X_s, Y_s) b(\cdot, s) \text{ in } L^2(\Omega \times [0, t]) \supseteq \mathcal{H}^2} \, \mathrm{d}B_s \\
&\to \int_0^t f_x(X_s, Y_s) a(\cdot, s) \, \mathrm{d}s + \int_0^t f_x(X_s, Y_s) b(\cdot, s) \, \mathrm{d}B_s \quad \text{as } n \to \infty \\
&= \int_0^t f_x(X_s, Y_s) \, \mathrm{d}X_s,
\end{aligned}
$$

and the analogous convergence for A_n^2. Secondly, we expect the terms B_n^1, B_n^2 to converge in the following manner:

$$
B_n^1 \to \frac{1}{2} \int_0^t f_{xx}(X_s, Y_s) \, \mathrm{d}\langle X \rangle_s \quad \text{as} \quad n \to \infty.
$$

Note that the latter integral is well-defined in the Lebesgue-Stieltjes sense because $(\langle X \rangle_t)_{t \in [0,T]}$ is non-decreasing. Indeed, we have:

Lemma 4.7. *Let $g \in C(\mathbb{R}^2)$, $(Z_t)_{t \in [0,T]}$ be a continuous, adapted $\mathbb{R}^2$-valued stochastic process and $(X_t)_{t \in [0,T]}$ be an Itô process. For any $t \in [0, T]$ and $t_i := t\frac{i}{n}$, $i = 0, \dots, n$, we have*

$$
\sum_{i=1}^{n} g(Z_{t_{i-1}})(X_{t_i} - X_{t_{i-1}})^2 \xrightarrow{\mathbb{P}} \int_0^t g(Z_s) \, \mathrm{d}\langle X \rangle_s \qquad \textit{for } n \to \infty.
$$

Proof. Step 1: Assume g is bounded. We show that it is sufficient to assume that $(X_t)_{t\in[0,T]}$ *is a local martingale.*

Since $(X_t)_{t\in[0,T]}$ is an Itô process, it can be decomposed into a local martingale $(M_t)_{t\in[0,T]}$ and a stochastic process $(A_t)_{t\in[0,T]}$ of finite variation, that is

$$X_t = M_t + A_t, \quad t \in [0,T],$$

where $M_t := \int_0^t b(\cdot,s)\,\mathrm{d}B_s$ and $A_t := X_0 + \int_0^t a(\cdot,s)\,\mathrm{d}s$. By Proposition 4.3 we observe that

$$\langle X\rangle_t = \langle M\rangle_t, \quad t \in [0,T],$$

Furthermore, we get

$$\begin{aligned}\sum_{i=1}^n g(Z_{t_{i-1}})(X_{t_i} - X_{t_{i-1}})^2 &= \sum_{i=1}^n g(Z_{t_{i-1}})(M_{t_i} - M_{t_{i-1}})^2 + \sum_{i=1}^n g(Z_{t_{i-1}})(A_{t_i} - A_{t_{i-1}})^2 \\ &\quad + 2\sum_{i=1}^n g(Z_{t_{i-1}})(M_{t_i} - M_{t_{i-1}})(A_{t_i} - A_{t_{i-1}})\end{aligned}$$

and, since $(A_t)_{t\in[0,T]}$ is of finite variation, we get

$$\sum_{i=1}^n g(Z_{t_{i-1}})(A_{t_i} - A_{t_{i-1}})^2 \to 0 \quad \text{and} \quad 2\sum_{i=1}^n g(Z_{t_{i-1}})(M_{t_i} - M_{t_{i-1}})(A_{t_i} - A_{t_{i-1}}) \to 0,$$

as $n\to\infty$. Hence,

$$\lim_{n\to\infty}\sum_{i=1}^n g(Z_{t_{i-1}})(X_{t_i} - X_{t_{i-1}})^2 = \lim_{n\to\infty}\sum_{i=1}^n g(Z_{t_{i-1}})(M_{t_i} - M_{t_{i-1}})^2$$

if the limit exists.

Step 2: Assume first that X is a bounded martingale with bounded quadratic variation $\langle X\rangle$ and that g is bounded.

We have by dominated convergence

$$\begin{aligned}\int_0^t g(Z_s)\,\mathrm{d}\langle X\rangle_s &= \int_0^t \lim_{n\to\infty}\sum_{i=1}^n g(Z_{t_{i-1}})\mathbb{1}_{(t_{i-1},t_i]}(s)\,\mathrm{d}\langle X\rangle_s \\ &= \lim_{n\to\infty}\sum_{i=1}^n g(Z_{t_{i-1}})\big(\langle X\rangle_{t_i} - \langle X\rangle_{t_{i-1}}\big).\end{aligned}$$

Moreover,

$$\begin{aligned}&\mathbb{E}\Big[\Big(\sum_{i=1}^n g(Z_{t_{i-1}})(X_{t_i} - X_{t_{i-1}})^2 - \sum_{i=1}^n g(Z_{t_{i-1}})\big(\langle X\rangle_{t_i} - \langle X\rangle_{t_{i-1}}\big)\Big)^2\Big] \\ &= \mathbb{E}\Big[\Big(\sum_{i=1}^n g(Z_{t_{i-1}})\big((X_{t_i} - X_{t_{i-1}})^2 - (\langle X\rangle_{t_i} - \langle X\rangle_{t_{i-1}})\big)\Big)^2\Big] \\ &= \sum_{i,j=1}^n \mathbb{E}\Big[\Big(g(Z_{t_{i-1}})\big((X_{t_i} - X_{t_{i-1}})^2 - (\langle X\rangle_{t_i} - \langle X\rangle_{t_{i-1}})\big)\Big) \\ &\qquad\qquad \times \Big(g(Z_{t_{j-1}})\big((X_{t_j} - X_{t_{j-1}})^2 - (\langle X\rangle_{t_j} - \langle X\rangle_{t_{j-1}})\big)\Big)\Big].\end{aligned}$$

Since X and $X^2 - \langle X\rangle$ are martingales, we have

$$\mathbb{E}\big[(X_{t_j} - X_{t_{j-1}})^2 - (\langle X\rangle_{t_j} - \langle X\rangle_{t_{j-1}})|\mathcal{F}_{t_{j-1}}\big] = \mathbb{E}\big[(X_{t_j}^2 - X_{t_{j-1}}^2) - (\langle X\rangle_{t_j} - \langle X\rangle_{t_{j-1}})|\mathcal{F}_{t_{j-1}}\big] = 0.$$

Using $(a+b)^2 \leqslant 2a^2 + 2b^2$ and the Cauchy-Schwarz inequality and $C > 0$ such that $\sup_{t\in[0,T]} |X_t| \leq C$, we thus obtain

$$\begin{aligned}
&\mathbb{E}\Big[\Big(\sum_{i=1}^n g(Z_{t_{i-1}})(X_{t_i} - X_{t_{i-1}})^2 - \sum_{i=1}^n g(Z_{t_{i-1}})\big(\langle X\rangle_{t_i} - \langle X\rangle_{t_{i-1}}\big)\Big)^2\Big] \\
&= \sum_{i=1}^n \mathbb{E}\Big[g^2(Z_{t_{i-1}})\big((X_{t_i} - X_{t_{i-1}})^2 - (\langle X\rangle_{t_i} - \langle X\rangle_{t_{i-1}})\big)^2\Big] \\
&\leq 2\|g\|_\infty^2 \Big(\mathbb{E}\Big[\sum_{i=1}^n (X_{t_i} - X_{t_{i-1}})^4\Big] + \mathbb{E}\Big[\sum_{i=1}^n (\langle X\rangle_{t_i} - \langle X\rangle_{t_{i-1}})^2\Big]\Big) \\
&\leq 2\|g\|_\infty^2 \Big(\mathbb{E}\Big[\underbrace{\mu_X(1/n)^2 \sum_{i=1}^n (X_{t_i} - X_{t_{i-1}})^2}_{\to 0,\, n\to\infty}\Big] + \mathbb{E}\Big[\underbrace{\mu_{\langle X\rangle}(1/n) \sum_{i=1}^n |\langle X\rangle_{t_i} - \langle X\rangle_{t_{i-1}}|}_{=\mu_{\langle X\rangle}(1/n)\langle X\rangle_t \to 0,\, n\to\infty}\Big]\Big) \\
&\leq 2\|g\|_\infty^2 \Big(\mathbb{E}\Big[\underbrace{\mu_X(1/n)^4}_{\to 0,\, n\to\infty}\Big]^{1/2} \underbrace{\mathbb{E}\Big[\Big(\sum_{i=1}^n (X_{t_i} - X_{t_{i-1}})^2\Big)^2\Big]^{1/2}}_{\leqslant \sqrt{6}C\mathbb{E}[X_t^2]^{1/2} \text{ as in (2.3)}} \\
&\qquad + \mathbb{E}\Big[\underbrace{\mu_{\langle X\rangle}(1/n) \sum_{i=1}^n |\langle X\rangle_{t_i} - \langle X\rangle_{t_{i-1}}|}_{=\mu_{\langle X\rangle}(1/n)\langle X\rangle_t \to 0,\, n\to\infty}\Big]\Big) \\
&\to 0
\end{aligned}$$

as $n \to \infty$, recalling the modulus of continuity μ from (3.3) restricted to the time interval $[0, t]$ and using dominated convergence.

Step 3: Localization.

In the general case let

$$\tau_m := \inf\big\{t \in [0,T] \,:\, |X_t| \geqslant m \text{ or } |\langle X\rangle_t| \geqslant m \text{ or } |Z_t| \geq m\big\} \wedge T$$

and define $X^m := (X_{t\wedge\tau_m})_{t\in[0,T]}$ and $Z^m := (Z_{t\wedge\tau_m})_{t\in[0,T]}$, which coincides with X and Z on $\{\tau_m = T\}$, respectively. Then, by Step 1 we can assume that X^m is a (bounded) martingale and by Step 2 we have

$$\sum_{i=1}^n g(Z_{t_{i-1}}^m)(X_{t_i}^m - X_{t_{i-1}}^m)^2 \xrightarrow{L^2} \int_0^t g(Z_s^m)\, \mathrm{d}\langle X^m\rangle_s$$

where we can replace g by some bounded continuous function g_m with $g_m(x) = g(x)$ for all

$x \in \mathbb{R}$ with $|x| \leqslant m$. Then for any $\varepsilon > 0$ we have

$$\begin{aligned}
&\mathbb{P}\Big(\Big|\sum_{i=1}^{n} g(Z^m_{t_{i-1}})(X_{t_i} - X_{t_{i-1}})^2 - \int_0^t g(Z_s)\,\mathrm{d}\langle X\rangle_s\Big| \geqslant \varepsilon\Big)\\
&\leqslant \mathbb{P}\Big(\Big\{\Big|\sum_{i=1}^{n} g(Z_{t_{i-1}})(X_{t_i} - X_{t_{i-1}})^2 - \int_0^t g(Z_s)\,\mathrm{d}\langle X\rangle_s\Big| \geqslant \varepsilon\Big\} \cap \{\tau_m = T\}\Big) + \mathbb{P}(\tau_m < T)\\
&\leqslant \mathbb{P}\Big(\Big\{\Big|\sum_{i=1}^{n} g(Z^m_{t_{i-1}})(X^m_{t_i} - X^m_{t_{i-1}})^2 - \int_0^t g(Z^m_s)\,\mathrm{d}\langle X^m\rangle_s\Big| \geqslant \varepsilon\Big\}\Big) + \mathbb{P}(\tau_m < T)\\
&\to 0
\end{aligned}$$

as $n, m \to \infty$. $\square$

Finally, we have to treat the term C_n, which was given by

$$\sum_{i=1}^{n} f_{xy}(X_{t_{i-1}}, Y_{t_{i-1}})(X_{t_i} - X_{t_{i-1}})(Y_{t_i} - Y_{t_{i-1}}).$$

Since $ab = \frac{1}{4}\big((a+b)^2 - (a-b)^2\big)$, we can rewrite C_n as

$$\frac{1}{4}\sum_{i=1}^{n} f_{xy}(X_{t_{i-1}}, Y_{t_{i-1}})\Big(\big((X+Y)_{t_i} - (X+Y)_{t_{i-1}}\big)^2 - \big((X-Y)_{t_i} - (X-Y)_{t_{i-1}}\big)^2\Big).$$

Since $X+Y$ and $X-Y$ are again Itô processes, both terms can be treated as in Lemma 4.7. Therefore, we arrive at the Itô formula for Itô processes.

Theorem 4.8 (Itô's formula for Itô processes). *Let $f \in C^2(\mathbb{R}^2)$ and X, Y be two Itô processes with representations*

$$X_t = X_0 + \int_0^t a(\cdot,s)\,\mathrm{d}s + \int_0^t b(\cdot,s)\,\mathrm{d}B_s \quad \text{and} \quad Y_t = Y_0 + \int_0^t \alpha(\cdot,s)\,\mathrm{d}s + \int_0^t \beta(\cdot,s)\,\mathrm{d}B_s,$$

for $t \in [0,T]$, $\mathbb{P}$-a.s. Then, we have $\mathbb{P}$-a.s for $t \in [0,T]$ that

$$\begin{aligned}
f(X_t, Y_t) = f(X_0, Y_0) &+ \int_0^t f_x(X_s, Y_s)\,\mathrm{d}X_s + \int_0^t f_y(X_s, Y_s)\,\mathrm{d}Y_s\\
&+ \frac{1}{2}\int_0^t f_{xx}(X_s, Y_s)\,\mathrm{d}\langle X\rangle_s + \frac{1}{2}\int_0^t f_{yy}(X_s, Y_s)\,\mathrm{d}\langle Y\rangle_s\\
&+ \int_0^t f_{xy}(X_s, Y_s)\,\mathrm{d}\langle X, Y\rangle_s,
\end{aligned}$$

where

$$\langle X, Y\rangle_t := \frac{1}{4}\big(\langle X+Y\rangle_t - \langle X-Y\rangle_t\big), \qquad t \in [0,T],$$

is the ***co-variation process****.*

Sketch of the proof. We have all ingredients together to prove the Itô formula for Itô processes analogously to Theorem 3.18. Recall the decomposition (4.1) of $f(X_t, Y_t) - f(X_0, Y_0)$ and apply the convergence results for the terms A^1, A^2, B^1, B^1 and C. The remainder can be treated as in Theorem 3.18. $\square$

5 Stochastic differential equations

Lecture 11

Many continuous-time evolution in the real world can be modeled by stochastic differential equations. A stochastic differential equation (SDE) is a differential equation of the form

$$\mathrm{d}X_t = \mu(t, X_t)\,\mathrm{d}t + \sigma(t, X_t)\,\mathrm{d}B_t, \quad t \in [0,T], \qquad X_0 = x_0, \tag{5.1}$$

which is a short writing for the integral equation

$$X_t = x_0 + \int_0^t \mu(s, X_s)\,\mathrm{d}s + \int_0^t \sigma(s, X_s)\,\mathrm{d}B_s, \quad t \in [0,T].$$

where $\mu\colon [0,T]\times\mathbb{R} \to \mathbb{R}$ is the so-called **drift function** and $\sigma\colon [0,T]\times\mathbb{R} \to \mathbb{R}$ is the so-called **volatility** or **diffusion function** and $x_0 \in \mathbb{R}$ is the **initial value**. A stochastic process $X = (X_t)_{t\in[0,T]}$ is called **solution** of the SDE (5.1) if $X \in \mathcal{H}^2_{loc}$ is a continuous process satisfying (5.1). Stochastic differential equations are the stochastic counterpart to the initial value problems from the theory of ordinary differential equations.

5.1 Linear SDEs

Before we study some general conditions to ensure existence and uniqueness of solutions of stochastic differential equations, we want to investigate the case of linear SDEs, that is, we assume that the coefficient functions μ and σ are affine functions in the space variable. Let us start with two examples of linear SDE.

Example 5.1. *(i) Geometric Brownian motion:* Consider the SDE

$$\mathrm{d}X_t = \mu X_t\,\mathrm{d}t + \sigma X_t\,\mathrm{d}B_t, \quad t \in [0,T], \quad X_0 = x_0 > 0 \tag{5.2}$$

with drift coefficient $\mu \in \mathbb{R}$ and volatility parameter $\sigma > 0$.

In order to find the solution $(X_t)_{t\in[0,T]}$, we make the ansatz $X_t = f(t, B_t)$ for a suitable function $f\colon [0,T]\times\mathbb{R} \to \mathbb{R}$, which we shall determine in the following. We have to find conditions on f such that (5.2) is satisfied. Itô's formula yields

$$\begin{aligned}
\mathrm{d}X_t &= f_t(t,B_t)\,\mathrm{d}t + f_x(t,B_t)\,\mathrm{d}B_t + \frac{1}{2} f_{xx}(t,B_t)\,\mathrm{d}t \\
&= \big(f_t(t,B_t) + \frac{1}{2} f_{xx}(t,B_t)\big)\,\mathrm{d}t + f_x(t,B_t)\,\mathrm{d}B_t \\
&\overset{!}{=} \mu X_t\,\mathrm{d}t + \sigma X_t\,\mathrm{d}B_t \\
&= \mu f(t,B_t)\,\mathrm{d}t + \sigma f(t,B_t)\,\mathrm{d}B_t.
\end{aligned}$$

Since the coefficients of an Itô process are a.s. uniquely determined, matching the coefficients of the two representations yields

$$\mu f = f_t + \frac{1}{2} f_{xx} \qquad \text{and} \qquad \sigma f = f_x.$$

The second equation implies

$$f(t,x) = \exp\big(\sigma x + g(t)\big)$$

for some regular $g\colon [0,T] \to \mathbb{R}$. Therefore,

$$g' f + \frac{\sigma^2}{2} f = \mu f$$

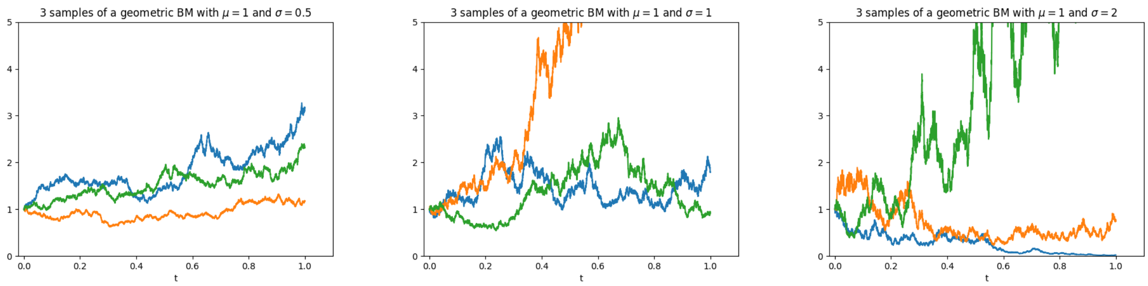
which is equivalent (as $f > 0$) to

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

Therefore, g is an affine function of the form $g(t) = (\mu - \frac{\sigma^2}{2})t + g_0$ for some $g_0 \in \mathbb{R}$ and thus

$$\begin{aligned} X_t = f(t, B_t) &= \exp\Big(\sigma B_t + \big(\mu - \frac{\sigma^2}{2}\big)t + g_0\Big) \\ &= x_0 \exp\Big(\sigma B_t + \big(\mu - \frac{\sigma^2}{2}\big)t\Big), \quad t \in [0,T], \end{aligned}$$

with $g_0 = \log x_0$. $(X_t)_{t\in[0,T]}$ is called **geometric Brownian motion**. In financial mathematics (5.2) is also referred to as Black-Scholes model for the dynamics of price evolutions on financial markets with trend μ and volatility σ.



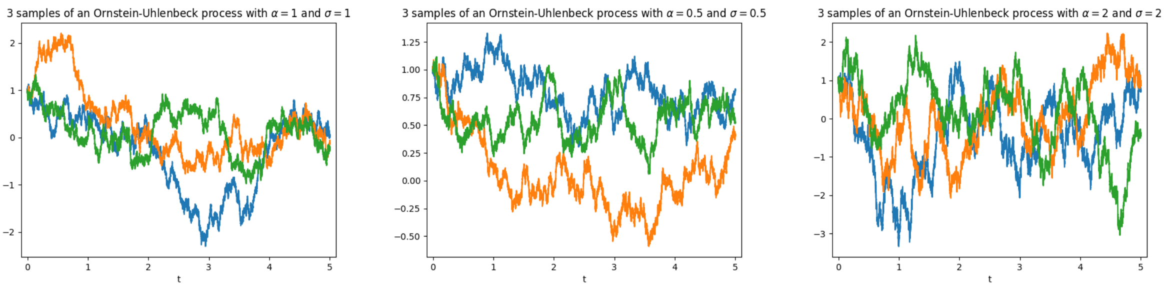
(ii) Ornstein-Uhlenbeck process: Consider the SDE

$$\mathrm{d}X_t = -\alpha X_t \,\mathrm{d}t + \sigma \,\mathrm{d}B_t, \quad t \in [0,T], \quad X_0 = x_0,$$

with a drift parameter $\alpha, \sigma > 0$ and $x_0 \in \mathbb{R}$. This SDE can also be explicitly solved. The solution $(X_t)_{t\in[0,T]}$ is given by

$$X_t = e^{-\alpha t} x_0 + \sigma \int_0^t e^{-\alpha(t-s)} \,\mathrm{d}B_s, \quad t \in [0,T],$$

and is called **Ornstein-Uhlenbeck processes**, cf. problem sheets.



In general, a **linear stochastic differential equations** is of the form

$$\mathrm{d}X_t = (\alpha(t)X_t + \beta(t))\,\mathrm{d}t + (\varphi(t)X_t + \vartheta(t))\,\mathrm{d}B_t, \quad t \in [0,T], \quad X_0 = x_0, \tag{5.3}$$

where $\alpha, \beta, \varphi, \vartheta \in \mathscr{H}^2$ and $x_0 \in \mathbb{R}$. It turns out that linear SDEs always possess an explicit solution.

Proposition 5.2. *A solution $(X_t)_{t\in[0,T]}$ of the linear SDE (5.3) is given by*

$$X_t := x_0 \exp(Y_t) + \int_0^t \exp(Y_t - Y_s)(\beta(s) - \vartheta(s)\varphi(s))\,\mathrm{d}s + \int_0^t \exp(Y_t - Y_s)\vartheta(s)\,\mathrm{d}B_s,$$

for $t \in [0,T]$, where

$$Y_t := \int_0^t \varphi(s)\,\mathrm{d}B_s + \int_0^t \big(\alpha(s) - \frac{1}{2}\varphi^2(s)\big)\,\mathrm{d}s.$$

Proof. Setting $H_t := \exp(Y_t)$ and $Z_t := x_0 + \int_0^t H_s^{-1}(\beta(s) - \vartheta(s)\varphi(s))\,\mathrm{d}s + \int_0^t H_s^{-1}\vartheta(s)\,\mathrm{d}B_s$, we observe that

$$X_t = H_t Z_t, \quad t \in [0,T].$$

Using Itô's formula for Itô processes (Theorem 4.8) and $\langle Y\rangle_s = \int_0^s \varphi^2(r)\,\mathrm{d}r$, we get

$$\begin{aligned} H_t := \exp(Y_t) &= 1 + \int_0^t \exp(Y_s)\,\mathrm{d}Y_s + \frac{1}{2}\int_0^t \exp(Y_s)\,\mathrm{d}\langle Y\rangle_s \\ &= 1 + \int_0^t \varphi(s)H_s\,\mathrm{d}B_s + \int_0^t \alpha(s)H_s\,\mathrm{d}s. \end{aligned}$$

Using again Itô's formula for Itô processes (or directly the product rule), we arrive at

$$\begin{aligned} X_t = H_t Z_t &= x_0 + \int_0^t H_s\,\mathrm{d}Z_s + \int_0^t Z_s\,\mathrm{d}H_s + \langle H, Z\rangle_t \\ &= x_0 + \int_0^t (\beta(s) - \vartheta(s)\varphi(s))\,\mathrm{d}s + \int_0^t \vartheta(s)\,\mathrm{d}B_s \\ &\quad + \int_0^t Z_s(H_s\varphi(s))\,\mathrm{d}B_s + \int_0^t Z_s(H_s\alpha(s))\,\mathrm{d}s + \int_0^t \vartheta(s)\varphi(s)\,\mathrm{d}s \\ &= x_0 + \int_0^t (\alpha(s)X_s + \beta(s))\,\mathrm{d}s + \int_0^t (\varphi(s)X_s + \vartheta(s))\,\mathrm{d}B_s \end{aligned}$$

as $\langle H, Z\rangle_t = \int_0^t \vartheta(s)\varphi(s)\,\mathrm{d}s$. □

5.2 SDEs with Lipschitz continuous coefficients

Most stochastic differential equations are actually not explicitly solvable. However, the assumption that the coefficients μ and σ are (uniformly) Lipschitz continuous, ensures that there always exists a unique solution to the associated SDEs. To calculate these solutions requires then numerical methods.

Theorem 5.3 (Existence and uniqueness of solution of SDEs). *Let $\mu, \sigma\colon [0,T] \times \mathbb{R} \to \mathbb{R}$ be two measurable functions satisfying for some constant $C > 0$*

$$\begin{aligned} |\mu(t,x) - \mu(t,y)|^2 + |\sigma(t,x) - \sigma(t,y)|^2 &\leqslant C|x-y|^2, \qquad t \in [0,T],\ x,y \in \mathbb{R}, \\ |\mu(t,x)|^2 + |\sigma(t,x)|^2 &\leqslant C(1+|x|^2), \qquad t \in [0,T],\ x \in \mathbb{R}, \end{aligned}$$

and let $x_0 \in \mathbb{R}$. Then, we have:

(i) There exists a solution $(X_t)_{t\in[0,T]}$ to stochastic differential equation

$$\mathrm{d}X_t = \mu(t, X_t)\,\mathrm{d}t + \sigma(t, X_t)\,\mathrm{d}B_t, \quad t \in [0,T], \quad X_0 = x_0. \tag{5.4}$$

(ii) Every solution $(X_t)_{t\in[0,T]}$ of (5.4) is uniformly bounded in L^2, i.e.,

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

(iii) The solution $(X_t)_{t\in[0,T]}$ of (5.4) is pathwise unique, i.e., if there is another solution $(Y_t)_{t\in[0,T]}$ of (5.4), we have $\mathbb{P}(\forall t \in [0,T] : X_t = Y_t) = 1$.

Remark 5.4.

(i) The Lipschitz condition in x on μ and σ is also imposed in the classical Picard-Lindelöf theorem which guarantees the existence of a unique solution of (deterministic) initial value problems. Indeed, our existence proof is based on a Picard iteration.

(ii) The second condition on μ and σ is a growth condition which is essential to guarantee a global solution: Let $\sigma = 0$ and $\mu(t,x) = \frac{1}{\beta-1}x^\beta$ for $\beta > 1$, $x \geqslant 0$. Then the differential equation

$$\mathrm{d}X_t = \frac{1}{\beta - 1} X_t^\beta\,\mathrm{d}t, \qquad X_0 = 1,$$

is solved by

$$X_t = (1-t)^{-1/(\beta-1)}.$$

Indeed: $\frac{\mathrm{d}}{\mathrm{d}t}X_t = \frac{1}{\beta-1}(1-t)^{-1/(\beta-1)-1} = \frac{1}{\beta-1}(1-t)^{-\beta/(\beta-1)} = \frac{1}{\beta-1}X_t^\beta$. The solution has no extension to $[0,T]$ for any $T > 1$.

In order to prove uniqueness of a solution we apply a classical auxiliary result:

Lemma 5.5 (Grönwall's lemma). *For $T \in (0,\infty)$ let $g\colon [0,T] \to \mathbb{R}$ be a bounded and measurable function. If there are constants $A, B \in \mathbb{R}$ such that*

$$g(t) \leqslant A + B\int_0^t g(s)\,\mathrm{d}s \qquad \text{for all } t \in [0,T],$$

then $g(t) \leqslant Ae^{Bt}$ for all $t \in [0,T]$.

Proof. The proof can be found in the Appendix A.5. □

Proof of Theorem 5.3 (uniqueness). Let $(X_t)_{t\in[0,T]}$ and $(Y_t)_{t\in[0,T]}$ be two solutions to the SDE (5.4). Then

$$X_t - Y_t = \int_0^t \big(\mu(s,X_s)-\mu(s,Y_s)\big)\,\mathrm{d}s + \int_0^t \big(\sigma(s,X_s)-\sigma(s,Y_s)\big)\,\mathrm{d}B_s, \quad t\in[0,T].$$

Since $(a+b)^2 \leqslant 2a^2+2b^2$, we have due to the Jensen's inequality, Itô's isometry and the Lipschitz condition:

$$\begin{aligned}\mathbb{E}[|X_t-Y_t|^2] &\leqslant 2\mathbb{E}\Big[\Big(\int_0^t \big(\mu(s,X_s)-\mu(s,Y_s)\big)\,\mathrm{d}s\Big)^2\Big] + 2\mathbb{E}\Big[\Big(\int_0^t \big(\sigma(s,X_s)-\sigma(s,Y_s)\big)\,\mathrm{d}B_s\Big)^2\Big]\\ &\leqslant 2t\mathbb{E}\Big[\int_0^t \big(\mu(s,X_s)-\mu(s,Y_s)\big)^2\,\mathrm{d}s\Big] + 2\mathbb{E}\Big[\int_0^t \big(\sigma(s,X_s)-\sigma(s,Y_s)\big)^2\,\mathrm{d}s\Big]\\ &\leqslant 2(t+1)C\mathbb{E}\Big[\int_0^t |X_s-Y_s|^2\,\mathrm{d}s\Big].\end{aligned}$$

Defining $g(t) := \mathbb{E}[|X_t-Y_t|^2]$, $t\in[0,T]$, the previous bound reads as

$$g(t) \leqslant 2(T+1)C\int_0^t g(s)\,\mathrm{d}s.$$

Since $(X_t)_{t\in[0,T]}$ and $(Y_t)_{t\in[0,T]}$ are uniformly bounded in L^2, we can apply Grönwall's lemma, which implies that $0\leqslant g(t)\leqslant 0$. Therefore, $\mathbb{P}(X_t=Y_t)=1$ for all $t\in[0,T]$ and by continuity of $(X_t)_{t\in[0,T]}$ and $(Y_t)_{t\in[0,T]}$ we conclude $\mathbb{P}(\forall t\in[0,T]: X_t=Y_t)=1$. □

Lecture 12

The proof of the existence of a solution to the SDE (5.1) relies on a Picard iteration: We define a map

$$\begin{aligned}F\colon \mathscr{H}^2([0,T]) &\to \mathscr{H}^2([0,T])\\ X &\mapsto F(X) := \Big(x_0 + \int_0^t \mu(s,X_s)\,\mathrm{d}s + \int_0^t \sigma(s,X_s)\,\mathrm{d}B_s\Big)_{t\in[0,T]}\end{aligned}$$

such that a solution of the SDE (5.4) is a fixed point of F. Moreover, we define the sequence

$$X^{(n+1)} := F(X^{(n)}) \quad \text{and} \quad X^{(0)} := (x_0)_{t\in[0,T]}.$$

We will show that $(X^{(n)})_{n\in\mathbb{N}}$ converges to a solution of (5.1) a.s. and in L^2. First, we have to verify that the iteration is well-defined.

Lemma 5.6. *Grant the assumptions of Theorem 5.3. Then:*

(i) For any $X\in\mathscr{H}^2([0,T])$ we have $\sigma(\cdot,X)\in\mathscr{H}^2([0,T])$ and $\mu(\cdot,X)\in L^2(\Omega\times[0,T])$.

(ii) If $X\in\mathscr{H}^2([0,T])$ is uniformly bounded in L^2, then $F(X)\in\mathscr{H}^2([0,T])$ is uniformly bounded in L^2.

Proof. *(i)* The growth condition on σ implies

$$\begin{aligned}\mathbb{E}\Big[\int_0^T \sigma^2(s,X_s)\,\mathrm{d}s\Big] &\leqslant C\mathbb{E}\Big[\int_0^T (1+|X_s|^2)\,\mathrm{d}s\Big]\\ &= CT + C\|X\|^2_{L^2(\Omega\times[0,T])}\end{aligned}$$

and analogously for $\mu(\cdot, X)$.

(ii) Let $B > 0$ be such that $\sup_{t\in[0,T]} \mathbb{E}[|X_t|^2] \leqslant B$. The Jensen's inequality, Itô's isometry and the growth condition yield

$$\begin{aligned}
\mathbb{E}[|F(X)_t|^2] &= \mathbb{E}\Big[\Big(x_0 + \int_0^t \mu(s, X_s)\,\mathrm{d}s + \int_0^t \sigma(s, X_s)\,\mathrm{d}B_s\Big)^2\Big] \\
&\leqslant 3x_0^2 + 3\mathbb{E}\Big[\Big(\int_0^t \mu(s, X_s)\,\mathrm{d}s\Big)^2\Big] + 3\mathbb{E}\Big[\Big(\int_0^t \sigma(s, X_s)\,\mathrm{d}B_s\Big)^2\Big] \\
&\leqslant 3x_0^2 + 3t\mathbb{E}\Big[\int_0^t \mu^2(s, X_s)\,\mathrm{d}s\Big] + 3\mathbb{E}\Big[\int_0^t \sigma^2(s, X_s)\,\mathrm{d}s\Big] \\
&\leqslant 3x_0^2 + 3C(t+1)\mathbb{E}\Big[\int_0^t (1 + |X_s|^2)\,\mathrm{d}s\Big] \\
&\leqslant 3x_0^2 + 3C(T+1)\int_0^t (1 + \mathbb{E}[|X_s|^2])\,\mathrm{d}s \\
&\leqslant 3x_0^2 + 3C(1+B)(T+1)T
\end{aligned}$$

for all $t \in [0, T]$. □

In particular, all sequence elements $X^{(n)} = F(X^{(n-1)})$, for $n \in \mathbb{N}$, are uniformly bounded in L^2. Moreover, we require the following a priori estimate:

Lemma 5.7. *Grant the assumptions of Theorem 5.3. Then there is some $D > 0$ such that*

$$\mathbb{E}\Big[\sup_{0\leqslant s\leqslant t} \big|X_s^{(n+1)} - X_s^{(n)}\big|^2\Big] \leqslant D \int_0^t \mathbb{E}\big[|X_s^{(n)} - X_s^{(n-1)}|^2\big]\,\mathrm{d}s, \quad t \in [0, T],$$

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

Proof. For $n \in \mathbb{N}$ and $t \in [0, T]$ we have

$$X_s^{(n+1)} - X_s^{(n)} = \int_0^s \big(\mu(r, X_r^{(n)}) - \mu(r, X_r^{(n-1)})\,\mathrm{d}r + \int_0^s \big(\sigma(r, X_r^{(n)}) - \sigma(r, X_r^{(n-1)})\,\mathrm{d}B_r.$$

Therefore, the Cauchy-Schwarz inequality, Doob's inequality and Itô's isometry imply

$$\begin{aligned}
\mathbb{E}\Big[\sup_{0\leqslant s\leqslant t} \big|X_s^{(n+1)} - X_s^{(n)}\big|^2\Big] &\leqslant 2\mathbb{E}\Big[\sup_{0\leqslant s\leqslant t} \Big(\int_0^s \big(\mu(r, X_r^{(n)}) - \mu(r, X_r^{(n-1)})\big)\,\mathrm{d}r\Big)^2\Big] \\
&\quad + 2\mathbb{E}\Big[\sup_{0\leqslant s\leqslant t} \Big(\int_0^s \big(\sigma(r, X_r^{(n)}) - \sigma(r, X_r^{(n-1)})\big)\,\mathrm{d}B_r\Big)^2\Big] \\
&\leqslant 2\mathbb{E}\Big[\sup_{0\leqslant s\leqslant t} s\int_0^s \big(\mu(r, X_r^{(n)}) - \mu(r, X_r^{(n-1)})\big)^2\,\mathrm{d}r\Big] \\
&\quad + 8\mathbb{E}\Big[\Big(\int_0^t \big(\sigma(r, X_r^{(n)}) - \sigma(r, X_r^{(n-1)})\big)\,\mathrm{d}B_r\Big)^2\Big] \\
&\leqslant 2t\mathbb{E}\Big[\int_0^t \big(\mu(r, X_r^{(n)}) - \mu(r, X_r^{(n-1)})\big)^2\,\mathrm{d}r\Big] \\
&\quad + 8\mathbb{E}\Big[\Big(\int_0^t \big(\sigma(r, X_r^{(n)}) - \sigma(r, X_r^{(n-1)})\big)^2\,\mathrm{d}r\Big] \\
&\leqslant \underbrace{C(2T+8)}_{=:D}\,\mathbb{E}\Big[\int_0^t |X_r^{(n)} - X_r^{(n-1)}|^2\,\mathrm{d}r\Big],
\end{aligned}$$

where we used the Lipschitz condition in the last estimate. □

Now we have all ingredients to prove the existence of a solution to (5.4).

Proof of Theorem 5.3 (Existence). Step 1: Almost sure convergence of $(X^{(n)})_{n\in\mathbb{N}}$ *to a continuous limit* X.

For $n \in \mathbb{N}_0$ and $t \in [0,T]$ define

$$g_n(t) := \mathbb{E}\Big[\sup_{0\leqslant s\leqslant t} |X_s^{(n+1)} - X_s^{(n)}|^2\Big].$$

Lemma 5.7 yields

$$g_n(t) \leqslant D\int_0^t g_{n-1}(s)\,\mathrm{d}s, \qquad n \in \mathbb{N}.$$

For $M := \sup_{0\leqslant t\leqslant T} g_0(t) < \infty$ (Lemma 5.6) we have

$$\begin{aligned} g_1(t) &\leqslant D\int_0^t g_0(s)\,\mathrm{d}s \leqslant MDt,\\ g_2(t) &\leqslant D\int_0^t g_1(s)\,\mathrm{d}s \leqslant M\frac{D^2t^2}{2}, \end{aligned}$$

and by induction we conclude

$$g_n(t) \leqslant M\frac{(Dt)^n}{n!}.$$

Therefore, Markov's inequality implies

$$\begin{aligned} \mathbb{P}\Big(\sup_{0\leqslant s\leqslant t} |X_s^{(n+1)} - X_s^{(n)}| \geqslant 2^{-n}\Big) &\leqslant 2^{2n}\mathbb{E}\Big[\sup_{0\leqslant s\leqslant t} |X_s^{(n+1)} - X_s^{(n)}|^2\Big]\\ &\leqslant M\frac{(4Dt)^n}{n!}. \end{aligned}$$

For $A_n := \{\sup_{0\leqslant s\leqslant t} |X_s^{(n+1)} - X_s^{(n)}| \geqslant 2^{-n}\}$ we obtain $\sum_{n\geqslant 0}\mathbb{P}(A_n) \leqslant Me^{4Dt} < \infty$. By the Borel-Cantelli lemma,

$$A := \limsup_{n\to\infty} A_n = \bigcap_{n\geqslant 0}\bigcup_{m\geqslant n} A_m$$

satisfies $\mathbb{P}(A) = 0$. Hence,

$$1 = \mathbb{P}(A^c) = \mathbb{P}\Big(\bigcup_{n\geqslant 0}\bigcap_{m\geqslant n} A_m^c\Big)$$

that means for $\mathbb{P}$-a.e. $\omega \in \Omega$ there is some $n(\omega) \in \mathbb{N}$ such that

$$\forall m \geqslant n(\omega): \qquad \sup_{0\leqslant s\leqslant t} |X_s^{(m+1)}(\omega) - X_s^{(m)}(\omega)| < 2^{-m}.$$

For $\mathbb{P}$-a.e. $\omega \in \Omega$ we conclude that $(X_s^{(n)}(\omega))_{n\in\mathbb{N}}$ is a Cauchy sequence uniformly in $t \in [0,T]$ since

$$\begin{aligned} \sup_{0\leqslant s\leqslant t} |X_s^{(k)}(\omega) - X_s^{(l)}(\omega)| &\leqslant \sum_{m=l}^{k-1} \sup_{0\leqslant s\leqslant t} |X_s^{(m+1)}(\omega) - X_s^{(m)}(\omega)|\\ &\leqslant \sum_{m\geqslant l} 2^{-m} = 2^{-l+1} \qquad \text{for all } k > l \geq n(\omega). \end{aligned}$$

Let $X(\omega)$ denote the continuous limit on A^c and set $X(\omega) = 0$ for $\omega \in A$.

Step 2: L^2-convergence and L^2-boundedness of the limit.

Step 1 and Fatou's lemma imply

$$\begin{aligned}
\mathbb{E}\Big[\sup_{0\leqslant t\leqslant T}|X_t - X_t^{(n)}|^2\Big]^{1/2} &= \mathbb{E}\Big[\lim_{m\to\infty}\sup_{0\leqslant t\leqslant T}|X_t^{(m)} - X_t^{(n)}|^2\Big]^{1/2} \\
&\leqslant \liminf_{m\to\infty}\mathbb{E}\Big[\sup_{0\leqslant t\leqslant T}|X_t^{(m)} - X_t^{(n)}|^2\Big]^{1/2} \\
&\leqslant \liminf_{m\to\infty}\sum_{k=n}^{m-1}\mathbb{E}\Big[\sup_{0\leqslant t\leqslant T}|X_t^{(k+1)} - X_t^{(k)}|^2\Big]^{1/2} \\
&\leqslant \sum_{k=n}^{\infty}\Big(M\frac{(Dt)^k}{k!}\Big)^{1/2} \\
&\to 0 \qquad \text{for } n\to\infty.
\end{aligned}$$

This shows the L^2-convergence and together with Lemma 5.6 also

$$\sup_{t\in[0,T]}\mathbb{E}[|X_t|^2] \leqslant \sup_{t\in[0,T]}\Big(2\mathbb{E}[|X_t - X_t^{(n)}|^2] + 2\mathbb{E}[|X_t^{(n)}|^2]\Big) < \infty.$$

Step 3: Show that X solves the SDE.

For the Lipschitz condition and Step 2 imply

$$\begin{aligned}
\mathbb{E}\Big[\int_0^T \big|\sigma(s,X_s) - \sigma(s,X_s^{(n)})\big|^2\,\mathrm{d}s\Big] &\leqslant C\mathbb{E}\Big[\int_0^T |X_s - X_s^{(n)}|^2\,\mathrm{d}s\Big] \\
&\leqslant CT\mathbb{E}\Big[\sup_{0\leqslant t\leqslant T}|X_t - X_t^{(n)}|^2\Big] \to 0 \quad \text{as} \quad n\to\infty.
\end{aligned}$$

Analogously, we have

$$\mathbb{E}\Big[\int_0^T \big|\mu(s,X_s) - \mu(s,X_s^{(n)})\big|^2\,\mathrm{d}s\Big] \to 0.$$

Therefore, $\sigma(\cdot,X) \in \mathscr{H}^2$ and $\mu(\cdot,X) \in L^2(\Omega\times[0,T])$ and

$$\begin{aligned}
\int_0^{\cdot} \sigma(s,X_s^{(n)})\,\mathrm{d}B_s &\to \int_0^{\cdot}\sigma(s,X_s)\,\mathrm{d}B_s, \\
\int_0^{\cdot} \mu(s,X_s^{(n)})\,\mathrm{d}s &\to \int_0^{\cdot}\mu(s,X_s)\,\mathrm{d}s
\end{aligned}$$

in L^2 and a.s. for a subsequence. Now taking limits on both sides of the Picard iteration, we get

$$\begin{array}{ccl}
X_t^{(n+1)} = F(X_t^{(n)}) & = & x_0 + \int_0^t \mu(s,X_s^{(n)})\,\mathrm{d}s + \int_0^t \sigma(s,X_s^{(n)})\,\mathrm{d}B_s \\
\downarrow & & \downarrow \\
X_t & & x_0 + \int_0^t \mu(s,X_s)\,\mathrm{d}s + \int_0^t \sigma(s,X_s)\,\mathrm{d}B_s.
\end{array}$$

So for any $t\in[0,T]$ we have $X_t = x_0 + \int_0^t \mu(s,X_s)\,\mathrm{d}s + \int_0^t \sigma(s,X_s)\,\mathrm{d}B_s$ a.s. and thus the SDE is satisfied almost surely for all $t\in[0,T]\cap\mathbb{Q}$. By continuity, X then solves (5.1) on $[0,T]$ with probability one. □

6 Martingale representation

Lecture 13

As before, we work on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with a Brownian motion $B = (B_t)_{t\in[0,T]}$ and we consider the *Brownian standard filtration* $(\mathcal{F}_t)_{t\in[0,T]}$, i.e. $\mathcal{F}_t = \sigma(\mathcal{F}_t^B \cup \mathcal{N})$ with $\mathcal{F}_t^B := \sigma(B_s,\, s \leqslant t)$ and $\mathcal{N} := \{A \subseteq \Omega : \exists B \in \mathcal{F}, \mathbb{P}(B) = 0 : A \subseteq B\}$. We assume $T \in (0,\infty)$.

6.1 Martingale representation theorem

The aim of this chapter is to prove that all L^2-martingales with respect to the Brownian standard filtration $(\mathcal{F}_t)_{t\in[0,T]}$ can be represented by a stochastic integral process. To this end, we show first that every $\mathcal{F}_T$-measurable random variable in L^2 can be represented by a stochastic integral.

To obtain these results, we will study Fourier transform type functionals and thus we need $\mathbb{C}$-valued stochastic processes $X = (X_t)_{t\in[0,T]}$. Their (stochastic) integrals are defined by

$$\int_0^t X_s \,\mathrm{d}s := \int_0^t \mathrm{Re}\; X_s \,\mathrm{d}s + i\int_0^t \mathrm{Im}\, X_s \,\mathrm{d}s,$$
$$\int_0^t X_s \,\mathrm{d}B_s := \int_0^t \mathrm{Re}\; X_s \,\mathrm{d}B_s + i\int_0^t \mathrm{Im}\, X_s \,\mathrm{d}B_s,$$

for $t \in [0,T]$, and especially Itô's formula still applies to $\mathbb{C}$-valued integral processes.

Our starting point is the functional

$$\exp\Big(iuB_T + \frac{u^2}{2}T\Big) \qquad \text{for } u \in \mathbb{R}.$$

Itô's formula yields

$$\begin{aligned}
&\exp\Big(iuB_T + \frac{u^2}{2}T\Big)\\
&\quad = 1 + \int_0^T iu\exp(iuB_s + u^2 s/2)\,\mathrm{d}B_s + \frac{1}{2}\int_0^T (iu)^2\exp(iuB_s + u^2 s/2)\,\mathrm{d}s\\
&\qquad + \int_0^T \frac{u^2}{2}\exp(iuB_s + u^2 s/2)\,\mathrm{d}s\\
&\quad = 1 + \int_0^T \underbrace{iu\exp(iuB_s + u^2 s/2)}_{=:\varphi(s)\in\mathscr{H}^2}\,\mathrm{d}B_s.
\end{aligned}$$

So, we have a representation of $\exp(iuB_s)$ as a stochastic integral of the form

$$\begin{aligned}
\exp(iuB_T) &= \exp(-\frac{u^2}{2}T) + \int_0^T iu\exp(iuB_s + u^2(s-T)/2)\,\mathrm{d}B_s\\
&= \mathbb{E}\big[\exp(iuB_T)\big] + \int_0^T iu\exp(iuB_s + u^2(s-T)/2)\,\mathrm{d}B_s.
\end{aligned}$$

Using this calculation and the product rule, we arrive at the following lemma.

Lemma 6.1. *Let $n \in \mathbb{N}$, $0 = t_0 < t_1 < \dots < t_n = T$ and $u_1, \dots, u_n \in \mathbb{R}$. Then there exists some $\varphi \in \mathscr{H}^2$ such that*

$$\prod_{j=1}^{n} \exp\Big(iu_j(B_{t_j} - B_{t_{j-1}}) + \frac{u_j^2}{2}(t_j - t_{j-1})\Big) = 1 + \int_0^T \varphi(s)\,\mathrm{d}B_s.$$

Proof. See problem sheets. □

Therefore, we have integral representations for random variables in the class

$$\mathscr{S} := \operatorname{span}\Big\{\prod_{j=1}^{n} \exp\big(iu_j(B_{t_j} - B_{t_{j-1}})\big) : n \in \mathbb{N}, u_1, \dots, u_n \in \mathbb{R}, 0 = t_0 < t_1 < \dots t_n = T\Big\}.$$

By proving that $\mathscr{S}$ is dense in $L^2(\Omega, \mathcal{F}_T, \mathbb{P})$, we obtain the following representation result for $\mathcal{F}_T$-measurable random variables:

Proposition 6.2 ($\mathscr{H}^2$-representation)**.** *Let $X \in L^2(\Omega, \mathcal{F}, \mathbb{P})$ be $\mathcal{F}_T$-measurable. Then, there exists some $\varphi \in \mathscr{H}^2$ with*

$$X = \mathbb{E}[X] + \int_0^T \varphi(s)\,\mathrm{d}B_s.$$

If $\varphi, \psi \in \mathscr{H}^2$ satisfy

$$X = \mathbb{E}[X] + \int_0^T \varphi(s)\,\mathrm{d}B_s = \mathbb{E}[X] + \int_0^T \psi(s)\,\mathrm{d}B_s,$$

then $\varphi = \psi$, $\mathbb{P} \otimes \lambda$-a.s.

Proof. Uniqueness: Since $\varphi - \psi \in \mathscr{H}^2$, we have by Itô's isometry

$$0 = \mathbb{E}\Big[\Big(\int_0^T (\varphi(s) - \psi(s))\,\mathrm{d}B_s\Big)^2\Big] = \mathbb{E}\Big[\int_0^T (\varphi(s) - \psi(s))^2\,\mathrm{d}s\Big].$$

Therefore, $\varphi = \psi$, $\mathbb{P} \otimes \lambda$-a.s.
Existence: Step 1: Finite-dimensional approximation.
Define

$$\begin{aligned}\mathscr{D} &:= \left\{ f(B_{t_1} - B_{t_0}, \dots, B_{t_n} - B_{t_{n-1}}) \,:\, \begin{array}{c} f\colon \mathbb{R}^n \to \mathbb{R} \text{ bounded measurable,} \\ n \in \mathbb{N}, 0 = t_0 < t_1 < \dots t_n = T \end{array} \right\} \\ &= \left\{ g(B_{t_1}, \dots, B_{t_n}) \,:\, \begin{array}{c} g\colon \mathbb{R}^n \to \mathbb{R} \text{ bounded measurable,} \\ n \in \mathbb{N}, 0 = t_0 < t_1 < \dots t_n = T \end{array} \right\}\end{aligned}$$

where the equality follows by taking $g(x_1, \dots, x_n) := f(x_1, x_2 - x_1, \dots, x_n - x_{n-1})$. Set

$$\mathscr{C} := \{A \in \mathcal{F}_T^B : \mathbb{1}_A \in \overline{\mathscr{D}}\},$$

where $\overline{\mathscr{D}}$ denotes the closure of $\mathscr{D}$ with respect to $L^2(\Omega, \mathcal{F}_T, \mathbb{P})$. Then $\mathscr{C}$ is a Dynkin system containing $\mathcal{F}_t^B$ for all $t \in [0, T]$. Therefore,

$$\mathscr{C} = \sigma(\mathcal{F}_t^B, t \in [0, T]) = \mathcal{F}_T^B,$$

implying that $\mathscr{D}$ is dense in $L^2(\Omega,\mathcal{F}_T^B,\mathbb{P})$. Moreover, for all $X\in L^2(\Omega,\mathcal{F}_T,\mathbb{P})$ there is some $Y\in L^2(\Omega,\mathcal{F}_T^B,\mathbb{P})$ such that $\mathbb{P}(X\neq Y)=0$. Hence, $\mathscr{D}$ is dense in $L^2(\Omega,\mathcal{F}_T,\mathbb{P})$.

Step 2: Prove that $\mathscr{S}_n$ is dense in $\mathscr{D}_n$ where

$$\mathscr{D}_n := \mathscr{D}_{t_1,\ldots,t_n} := \big\{ f(B_{t_1}-B_{t_0},\ldots,B_{t_n}-B_{t_{n-1}}) \in L^2(\Omega,\mathcal{F}_T,\mathbb{P}) \,:\, f\colon \mathbb{R}^n\to\mathbb{C} \text{ measurable}\big\},$$
$$\mathscr{S}_n := \mathscr{S}_{t_1,\ldots,t_n} := \operatorname{span}\Big\{ \prod_{j=1}^n \exp\big(iu_j(B_{t_j}-B_{t_{j-1}})\big) \,:\, u_1,\ldots,u_n\in\mathbb{R}\Big\},$$

for fixed $n\in\mathbb{N}, 0=t_0<t_1<\ldots t_n=T$.

Note that $\mathscr{D}_n$ is a closed linear subspace of $L^2(\Omega,\mathcal{F}_T,\mathbb{P})$ containing $\mathscr{S}_n$. First, we show that

$$\mathscr{S}_n^\perp\cap\mathscr{D}_n=\{0\},\quad \text{i.e. } \forall X\in\mathscr{D}_n \text{ with } \mathbb{E}[XZ]=0 \text{ for all } Z\in\mathscr{S}_n \text{ we have } X=0. \tag{6.1}$$

Denoting the density of $(B_{t_1}-B_{t_0},\ldots,B_{t_n}-B_{t_{n-1}})$ by $p\colon\mathbb{R}^n\to(0,\infty)$, we have for any $X\in\mathscr{S}_n^\perp\cap\mathscr{D}_n$ with $X=f(B_{t_1}-B_{t_0},\ldots,B_{t_n}-B_{t_{n-1}})$ and $Z_u=\prod_{j=1}^n\exp\big(iu_j(B_{t_j}-B_{t_{j-1}})\big)$, $u\in\mathbb{R}^n$, that

$$0=\mathbb{E}[XZ_u]=\int_{\mathbb{R}^n} f(x)p(x)e^{i\langle u,x\rangle}\,\mathrm{d}x \qquad \text{for all } u\in\mathbb{R}^n.$$

Hence, the Fourier transform $\mathcal{F}[fp]$ of fp is everywhere 0. This implies $fp=0$ and since $p>0$, we obtain $f=0$. This gives (6.1).

Because $\mathscr{S}_n^\perp$ is the orthogonal complement of $\overline{\mathscr{S}_n}$, there is a unique decomposition for any $f\in\mathscr{D}_n$

$$f=g+h,\qquad g\in\overline{\mathscr{S}_n},\ h\in\mathscr{S}_n^\perp.$$

Then (6.1) implies $h=0$, i.e. $f=g\in\overline{\mathscr{S}_n}$. So indeed, $\mathscr{S}_n$ is dense in $\mathscr{D}_n$.

Step 3: Representation for all $X\in L^2(\Omega,\mathcal{F}_T,\mathbb{P})$.

By Step 2, $\mathscr{S}=\bigcup_{n\in\mathbb{N},0=t_0<t_1<\ldots t_n=T}\mathscr{S}_{t_1,\ldots,t_n}$ is dense in $\mathscr{D}=\bigcup_{n\in\mathbb{N},0=t_0<t_1<\ldots t_n=T}\mathscr{D}_{t_1,\ldots,t_n}$. Let $(X_n)_{n\in\mathbb{N}}\subseteq\mathscr{S}$ be an approximating sequence with $X_n\to X$ in L^2 as $n\to\infty$. In particular, $\mathbb{E}[X_n]\to\mathbb{E}[X]$. By Lemma 6.1 there are $\varphi_n\in\mathscr{H}^2$ such that

$$X_n=\mathbb{E}[X_n]+\int_0^T\varphi_n(s)\,\mathrm{d}B_s.$$

In view of

$$\|X_n-X_m\|_{L^2}^2=\big(\mathbb{E}[X_n]-\mathbb{E}[X_m]\big)^2+\underbrace{\mathbb{E}\Big[\Big(\int_0^T(\varphi_n-\varphi_m)(s)\,\mathrm{d}B_s\Big)^2\Big]}_{=\mathbb{E}[\int_0^T(\varphi_n-\varphi_m)^2(s)\,\mathrm{d}s]},$$

$(\varphi_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in $\mathscr{H}^2$ because $(X_n)_{n\in\mathbb{N}}$ is a Cauchy sequence in L^2. Therefore φ_n converges to some $\varphi\in\mathscr{H}^2$ and we obtain

$$\begin{aligned}X=\lim_{n\to\infty}X_n&=\lim_{n\to\infty}\Big(\mathbb{E}[X_n]+\int_0^T\varphi_n(s)\,\mathrm{d}B_s\Big)\\&=\mathbb{E}[X]+\int_0^T\varphi(s)\,\mathrm{d}B_s\end{aligned}$$

by Itô's isometry. □

Based on the previous representation theorem (Proposition 6.2), it is straightforward to deduce the martingale representation theorem.

Theorem 6.3 (Martingale representation theorem). *Let $(X_t)_{t\in[0,T]}$ be an $(\mathcal{F}_t)$-martingale with $\mathbb{E}[X_T^2] < \infty$. Then, there exists some $\varphi \in \mathscr{H}^2$ such that*

$$X_t = \mathbb{E}[X_0] + \int_0^t \varphi(s)\,\mathrm{d}B_s, \quad t \in [0,T].$$

This representation is $\mathbb{P}\otimes\lambda$-a.s. unique.

Proof. Proposition 6.2 applied to X_T gives some a.s. unique $\varphi \in \mathscr{H}^2$ such that

$$X_T = \mathbb{E}[X_T] + \int_0^T \varphi(s)\,\mathrm{d}B_s.$$

It remains to take the conditional expectation with respect to $\mathcal{F}_t$ on both sides taking into account that $\varphi \in \mathscr{H}^2$. □

6.2 Time-changes and Lévy's characterization of Brownian motion

Lecture 14

The martingale representation theorem allows to characterize the Brownian motion as a special continuous local martingale. To that end, we first show that every local martingale with respect to the Brownian standard filtration can be time-changed to be a Brownian motions.

Theorem 6.4 (Time-change representation). *Let $\varphi \in \mathscr{H}^2_{loc}$ for all $T > 0$ and $X := \int_0^\cdot \varphi(s)\,\mathrm{d}B_s$. Assume that $t \mapsto \int_0^t \varphi^2(s)\,\mathrm{d}s$ is a.s. strictly increasing with $\int_0^\infty \varphi^2(s)\,\mathrm{d}s = \infty$ $\mathbb{P}$-a.s. Then $(X_{\tau_t})_{t\in[0,\infty)}$ is a Brownian motion where*

$$\tau_t := \inf\Big\{u \in [0,\infty) \,:\, \int_0^u \varphi^2(s)\,\mathrm{d}s \geqslant t\Big\}.$$

Proof. Step 1: Prove for all $0 \leqslant s \leqslant t < \infty$ that

$$\mathbb{E}\big[\exp(iu(X_{\tau_t} - X_{\tau_s}))\big|\mathcal{F}_{\tau_s}\big] = \exp\Big(-\frac{u^2(t-s)}{2}\Big).$$

Set

$$Z_t := \exp(Y_t) \qquad \text{with} \qquad Y_t := iu\int_0^t \varphi(s)\,\mathrm{d}B_s + \frac{u^2}{2}\int_0^t \varphi^2(s)\,\mathrm{d}s.$$

Since Y is an $\mathbb{C}$-valued Itô process, Itô's formula yields

$$\begin{aligned} Z_t &= 1 + \int_0^t Z_s\,\mathrm{d}Y_s + \frac{1}{2}\int_0^t Z_s\,\mathrm{d}\langle Y\rangle_s \\ &= 1 + iu\int_0^t Z_s\varphi(s)\,\mathrm{d}B_s + \frac{u^2}{2}\int_0^t Z_s\varphi^2(s)\,\mathrm{d}s + \frac{1}{2}\int_0^t Z_s(iu\varphi(s))^2\,\mathrm{d}s \\ &= 1 + iu\int_0^t Z_s\varphi(s)\,\mathrm{d}B_s. \end{aligned}$$

Hence, Z is a $\mathbb{C}$-valued continuous local martingale by the definition of the stochastic integral $\mathbb{C}$-valued integrands. Hence, $(Z_{\tau_t})_{t\in[0,T]}$ is a $(\mathcal{F}_{\tau_t})$-local martingale, see problem sheets, and, since $|Z_{\tau_t}| \leqslant \exp(\frac{u^2}{2}t)$, it is even a proper martingale. Therefore,

$$\mathbb{E}\big[\exp\big(iu(X_{\tau_t} - X_{\tau_s}) + \frac{u^2}{2}(t-s)\big)|\mathcal{F}_{\tau_s}\big] = 1.$$

Step 2: Conclusion.
By Step 1 we have

$$\mathbb{E}\big[\exp\big(iu(X_{\tau_t} - X_{\tau_s})\big)\big] = \exp\big(-\frac{u^2}{2}(t-s)\big),$$

which means that $X_{\tau_t} - X_{\tau_s} \sim \mathcal{N}(0, t-s)$. Since for all $A \in \mathcal{F}_{\tau_s}$, using again Step 1 we have

$$\mathbb{E}\big[\exp\big(iu(X_{\tau_t} - X_{\tau_s})\big)\mathbb{1}_A\big] = \exp\big(-\frac{u^2}{2}(t-s)\big)\mathbb{P}(A),$$

we can conclude independence of $X_{\tau_t} - X_{\tau_s}$ and $\mathcal{F}_{\tau_s}$. □

We now combine the two main results from above to obtain Lévy's characterization of Brownian motion:

Theorem 6.5 (Lévy's characterization of Brownian motion)**.** *The following statements are equivalent:*

(i) $(B_t)_{t\in[0,\infty)}$ *is a Brownian motion.*

(ii) $(B_t)_{t\in[0,\infty)}$ *is a continuous local martingale with respect to the Brownian standard filtration* $(\mathcal{F}_t)_{t\in[0,\infty)}$ *and satisfies* $B_0 = 0$ *and* $\langle B\rangle_t = t$ *for* $t \in [0,\infty)$.

Proof. $(i) \Rightarrow (ii)$ is obvious.

$(ii) \Rightarrow (i)$ Since $\mathbb{E}[\langle B\rangle_t] = t < \infty$ for all $t \in [0,\infty)$, B is a continuous martingale, cf. Exercises. Then the martingale representation theorem implies that there is some $\varphi \in \mathscr{H}^2$ such that

$$B_t = \int_0^t \varphi(s)\,\mathrm{d}\widetilde{B}_s, \quad t \in [0,\infty),$$

from some Brownian motion $\widetilde{B}$ with respect to $(\mathcal{F}_t)_{t\in[0,\infty)}$. Therefore,

$$t = \langle B\rangle_t = \int_0^t \varphi^2(s)\,\mathrm{d}s.$$

Then the time-change representation (Theorem 6.4) is given by $\tau_t = t$, which shows that $B_{\tau_t} = B_t$ is a Brownian motion. □

7 Girsanov's theorem

In this chapter we study what happens with local martingales (with respect to the probability measure $\mathbb{P}$) if we consider a different equivalent probability measure Q on the underlying measurable space $(\Omega, \mathcal{F})$. Changing the probability measure is a common technique in mathematical finance and statistics.

Let Q and $\mathbb{P}$ be two equivalent probability measures on $(\Omega, \mathcal{F})$, that is

$$\forall A \in \mathcal{F}: \quad \mathbb{P}(A) = 0 \quad \Leftrightarrow \quad Q(A) = 0.$$

By the Radon-Nikodym theorem (Theorem A.20) there exists a strictly positive Radon-Nikodym density

$$L_T := \frac{\mathrm{d}Q}{\mathrm{d}\mathbb{P}} \quad \text{on } (\Omega, \mathcal{F}),$$

that is

$$\mathbb{E}^Q[\mathbb{1}_A] = \mathbb{E}[\mathbb{1}_A L_T] \quad \text{for all } A \in \mathcal{F}.$$

If $L_T \in L^1(\Omega, \mathcal{F}, \mathbb{P})$, we can define a martingale $(L_t)_{t\in[0,T]}$ by setting

$$L_t := \mathbb{E}[L_T|\mathcal{F}_t], \quad t \in [0, T],$$

which is often called **density process** of Q with respect to $\mathbb{P}$. Conversely, we may assume:

Assumption 7.1. *The stochastic process $(L_t)_{t\in[0,T]}$ satisfies:*

- *$(L_t)_{t\in[0,T]}$ is a continuous $\mathbb{P}$-martingale,*
- *$L_t > 0$ for all $t \in [0, T]$ $\mathbb{P}$-a.s. and $L_0 = 1$.*

If a stochastic process $(L_t)_{t\in[0,T]}$ satisfies Assumption 7.1, we define the associated probability measure Q by

$$Q(A) = \mathbb{E}^Q[\mathbb{1}_A] = \mathbb{E}[\mathbb{1}_A L_T] \quad \text{for all } A \in \mathcal{F},$$

Note that Q is equivalent to the probability measure $\mathbb{P}$ and the Radon-Nikodym density is $\frac{\mathrm{d}Q}{\mathrm{d}\mathbb{P}} = L_T$.

Based on density processes, we first investigate how (local) martingales behave under a change of measure.

Proposition 7.2. *Let $(L_t)_{t\in[0,T]}$ be as in Assumption 7.1 with associated probability measure Q and $(X_t)_{t\in[0,T]}$ be an adapted, continuous stochastic process. Then, the following conditions are equivalent:*

(i) *$(L_tX_t)_{t\in[0,T]}$ is a (local) $\mathbb{P}$-martingale.*

(ii) *$(X_t)_{t\in[0,T]}$ is a (local) Q-martingale.*

Proof. "$\Rightarrow$" We want to show: If $(L_tX_t)_{t\in[0,T]}$ is a $\mathbb{P}$-martingale, then $(X_t)_{t\in[0,T]}$ is a Q-martingale.

Notice, that $(X_t)_{t\in[0,T]}$ is adapted and Q-integrable as

$$\mathbb{E}^Q[|X_t|] = \mathbb{E}[L_T|X_t|] = \mathbb{E}[\mathbb{E}[L_T|X_t|\,|\mathcal{F}_t]] = \mathbb{E}[L_t|X_t|] = \mathbb{E}[|L_tX_t|] < \infty$$

because $(L_t)_{t\in[0,T]}$ and $(L_tX_t)_{t\in[0,T]}$ are $\mathbb{P}$-martingales.

In order to prove that $\mathbb{E}^Q[X_t|\mathcal{F}_s] = X_s$ for $s, t \in [0, T]$ with $s \le t$, it is by definition sufficient to check

$$\mathbb{E}^Q[\mathbb{1}_A(X_t - X_s)] = 0 \quad \text{for all } A \in \mathcal{F}_s.$$

Indeed, we have that

$$\begin{aligned}\mathbb{E}^Q[\mathbb{1}_A(X_t - X_s)] &= \mathbb{E}[L_T\mathbb{1}_A(X_t - X_s)] \\ &= \mathbb{E}[\mathbb{E}[L_T\mathbb{1}_A X_t|\mathcal{F}_t] - \mathbb{E}[L_T\mathbb{1}_A X_s|\mathcal{F}_s]] \\ &= \mathbb{E}[L_t\mathbb{1}_A X_t - L_s\mathbb{1}_A X_s] \\ &= \mathbb{E}[\mathbb{1}_A(L_tX_t - L_sX_s)] \\ &= 0,\end{aligned}$$

because $(L_t)_{t\in[0,T]}$ and $(L_tX_t)_{t\in[0,T]}$ are $\mathbb{P}$-martingales.

"$\Leftarrow$" The converse direction follows by the same arguments.

The statement for local martingales can be obtained by applying a localization argument. □

The fundamental theorem regarding change of measures is the so-called *Girsanov's theorem*, which is here presented in case of Brownian motion.

Theorem 7.3 (Girsanov's theorem). *If $L = (L_t)_{t\in[0,T]}$ where*

$$L_t = \exp\left(\int_0^t X_s \,\mathrm{d}B_s - \frac{1}{2}\int_0^t X_s^2 \,\mathrm{d}s\right), \quad t \in [0,T],$$

is a martingale on $(\Omega, \mathcal{F}_T, \mathbb{P})$ with respect to the Brownian standard filtration $(\mathcal{F}_t)_{t\in[0,T]}$, then

$$\widetilde{B}_t := B_t - \int_0^t X_s \,\mathrm{d}s, \quad t \in [0,T],$$

defines a Brownian motion $\widetilde{B} = (\widetilde{B}_t)_{t\in[0,T]}$ with respect to $(\Omega, \mathcal{F}_T, (\mathcal{F}_t)_{t\in[0,T]}, Q)$ where $\frac{\mathrm{d}Q}{\mathrm{d}\mathbb{P}} := L_T$.

Proof. By Lévy's characterization of Brownian motion (Theorem 6.5) we need to show that $(\widetilde{B}_t)_{t\in[0,T]}$ is a local Q-martingale with $\langle \widetilde{B}\rangle_t = t$ for $t \in [0,T]$.

Step 1: $(\widetilde{B}_t)_{t\in[0,T]}$ is a Q-martingale.

By Proposition 7.2 we need to show that $(\widetilde{B}_tL_t)_{t\in[0,T]}$ *is a continuous local* $\mathbb{P}$*-martingale.* The product formula yields

$$\begin{aligned}\widetilde{B}_tL_t &= \widetilde{B}_0L_0 + \int_0^t \widetilde{B}_s \,\mathrm{d}L_s + \int_0^t L_s \,\mathrm{d}\widetilde{B}_s + \langle \widetilde{B}, L\rangle_t \\ &= 0 + \int_0^t \widetilde{B}_sL_sX_s \,\mathrm{d}B_s + \int_0^t L_s \,\mathrm{d}B_s - \int_0^t L_sX_s \,\mathrm{d}s + \int_0^t 1\cdot L_sX_s \,\mathrm{d}s \\ &= \int_0^t \big(\widetilde{B}_sX_s + 1\big)L_s \,\mathrm{d}B_s,\end{aligned}$$

which is indeed a continuous local martingale with respect to $\mathbb{P}$.

Step 2: $(\widetilde{B}_t)_{t\in[0,T]}$ *is a continuous local Q-martingale.*

Let $((\widetilde{B}L)_{t\wedge\tau_n})_{,t\in[0,T]}$ be a $\mathbb{P}$-martingale for some suitable localizing stopping times $\tau_n \to T$. By Step 1 and Proposition 7.2 the process $(\widetilde{B}_{t\wedge\tau_n})_{t\in[0,T]}$ is a continuous $\mathbb{Q}$-martingale with respect to $(\mathcal{F}_t)_{t\in[0,T]}$ and thus $(\widetilde{B}_t)_{t\in[0,T]}$ is a continuous local $\mathbb{Q}$-martingale.

Step 3: $(\widetilde{B}_t)_{t\in[0,T]}$ *has quadratic variation* $\langle \widetilde{B}\rangle_t = t$, $t \in [0,T]$, *under* $\mathbb{Q}$.

To this end, we show that $(\widetilde{B}_t^2 - t)_{t\in[0,T]}$ is a continuous local martingale under $\mathbb{Q}$. Due to Proposition 7.2, it suffices to verify that

$$\big((\widetilde{B}_t^2 - t)L_t\big)_{t\in[0,T]}$$

is a continuous local $\mathbb{P}$-martingale. Under $\mathbb{P}$ we have

$$\langle \widetilde{B}\rangle_t = \Big\langle B - \int_0^{\cdot} X_s \,\mathrm{d}s\Big\rangle_t = \langle B\rangle_t = t.$$

Therefore,

$$\widetilde{B}_t^2 = \int_0^t 2\widetilde{B}_s \,\mathrm{d}\widetilde{B}_s + \langle \widetilde{B}\rangle_t = \int_0^t 2\widetilde{B}_s \,\mathrm{d}\widetilde{B}_s + t = \int_0^t 2\widetilde{B}_s \,\mathrm{d}B_s - \int_0^t 2\widetilde{B}_s X_s \,\mathrm{d}s + t$$

and the product formula implies that

$$\begin{aligned}(\widetilde{B}_t^2 - t)L_t &= \int_0^t (\widetilde{B}_s^2 - s)\,\mathrm{d}L_s + \int_0^t L_s \cdot 2\widetilde{B}_s \,\mathrm{d}\widetilde{B}_s + \int_0^t 2\widetilde{B}_s L_s X_s \,\mathrm{d}s \\ &= \int_0^t (\widetilde{B}_s^2 - s) L_s X_s \,\mathrm{d}B_s + \int_0^t 2L_s\widetilde{B}_s \,\mathrm{d}B_s - \int_0^t 2L_s\widetilde{B}_s X_s \,\mathrm{d}s + \int_0^t 2\widetilde{B}_s L_s X_s \,\mathrm{d}s \\ &= \int_0^t \big((\widetilde{B}_s^2 - s)X_s + 2\widetilde{B}_s\big) L_s \,\mathrm{d}B_s\end{aligned}$$

is a continuous local $\mathbb{P}$-martingale. □

8 Application: Bachelier model

Bonus material

Stochastic calculus has many applications in various areas of applied mathematics, engineering and related fields. For instance, continuous-time mathematical finance relies fundamentally on stochastic calculus. In this chapter, as a toy application, we shall briefly discuss the task of pricing and hedging of European options in the Bachelier model. Already in 1900 Louis Bachelier proposed to use the Brownian motion as a model for the price evolution of stocks and derived explicit formulas for the prices of European call and put options.

The **Bachelier model** assumes that the financial market $(S_t^0, S_t^1)_{t\in[0,T]}$ consists of two assets:

- The price process $(S_t^1)_{t\in[0,T]}$ of the *risky asset*, e.g. some stock, is given by

  $$S_t^1 = S_0^1 + \mu t + \sigma B_t, \quad t \in [0,T],$$

  where, $S_0^1 \in \mathbb{R}$, $\mu \in \mathbb{R}$ is the drift parameter, $\sigma > 0$ is the volatility parameter and $(B_t)_{t\in[0,T]}$ is a Brownian motion.

- The price process $(S_t^0)_{t\in[0,T]}$ of the *risk-free asset* with price process is given by $S_t^0 := 1$ for $t \in [0,T]$, i.e. the interest rate is $r = 0$.

The *information flow* $(\mathcal{F}_t)_{t\in[0,T]}$ on this financial market is supposed to be generated by the processes $(S^0_t, S^1_t)_{t\in[0,T]}$, that is, we take $(\mathcal{F}_t)_{t\in[0,T]}$ as the Brownian standard filtration. A *financial derivative* ξ is an $\mathcal{F}_T$-measurable random variable.

Before relying on the Bachelier model for any task in mathematical finance, we need to check that it is arbitrage-free, i.e., that there is no way to make money without any risk. So what does it mean to trad on a financial market modeled by the Bachelier model?

- A **trading strategy** is a $\mathbb{R}^2$-valued adapted process $\varphi = (\varphi^0_t, \varphi^1_t)_{t\in[0,T]}$ with $\varphi^1 \in L(S^1)$. The corresponding **capital process** $V(\varphi) = (V_t(\varphi))_{t\in[0,T]}$ is given by

$$V_t(\varphi) := \varphi_t^T S_t = \sum_{i=0}^{1} \varphi^i_t S^i_t, \quad t \in [0,T],$$

  where φ_t^T is the transpose of the random vector φ_t.

- A trading strategy φ is called **admissible** if

$$V_t(\varphi) \geq C(1 + |S^1_t|) \quad \text{and} \quad V_t(\varphi) = V_0(\varphi) + \int_0^t \varphi^0_r \,\mathrm{d}S^0_r + \int_0^t \varphi^1_r \,\mathrm{d}S^1_r, \quad t \in [0,T],$$

  for some constant $C > 0$.

Definition 8.1. An admissible trading strategy $\varphi = (\varphi_t)_{t\in[0,T]}$ is called **arbitrage opportunity** if

$$V_0(\varphi) = 0, \quad V_T(\varphi) \geqslant 0, \quad \mathbb{P}\text{-a.s} \quad \text{and} \quad \mathbb{P}(V_T(\varphi) > 0) > 0.$$

Lemma 8.2. *The Bachelier model $(S^0_t, S^1_t)_{t\in[0,T]}$ is*

(i) arbitrage-free, i.e. there exists no arbitrage opportunity.

(ii) complete, i.e. for every bounded financial derivative ξ there exists an admissible trading strategy φ such that $\xi = V_T(\varphi)$.

Proof. Introducing a new probability measure Q by the Radon-Nikodym density

$$\frac{\mathrm{d}Q}{\mathrm{d}\mathbb{P}} := \exp\Big(-\frac{\mu}{\sigma} B_T - \frac{1}{2}\Big(\frac{\mu}{\sigma}\Big)^2 T\Big),$$

we know, by Girsanov's theorem (Theorem 7.3), that $(\widetilde{B}_t)_{t\in[0,T]} := (B_t + \frac{\mu}{\sigma} t)_{t\in[0,T]}$ is a Brownian motion w.r.t. Q and $(S^1_t)_{t\in[0,T]}$ is a martingale w.r.t. Q, see the problem sheets.

(i) Suppose $\varphi = (\varphi^0_t, \varphi^1_t)_{t\in[0,T]}$ is an admissible trading strategy with $V_0(\varphi) = 0$ and $V_T(\varphi) \geqslant 0$, $\mathbb{P}$-a.s. Since φ is admissible, we get

$$\mathbb{E}^Q[V_T(\varphi)] = \mathbb{E}^Q[\int_0^T \varphi^1_r \,\mathrm{d}S^1_r] \leq 0.$$

Hence, since Q and $\mathbb{P}$ are equivalent probability measures and $V_T(\varphi) \geqslant 0$, we have $V_T(\varphi) = 0$. Consequently, there exists no arbitrage opportunity.

(ii) We work again with the probability measure Q: Since ξ is a bounded $\mathcal{F}_T$-valued random variable, the martingale representation theorem (Proposition 6.2) provides a $\psi \in \mathscr{H}^2$ such that

$$\xi = \mathbb{E}^Q[\xi] + \int_0^T \psi(r)\,\mathrm{d}\widetilde{B}_r, \quad Q\text{-a.s.},$$

since $(\widetilde{B}_t)_{t\in[0,T]}$ is a Brownian motion w.r.t. Q. Hence,

$$\xi = \mathbb{E}^Q[\xi] + \int_0^T \frac{\psi(r)}{\sigma}\,\mathrm{d}(\sigma B_r + \mu t) = \mathbb{E}^Q[\xi] + \int_0^T \frac{\psi(r)}{\sigma}\,\mathrm{d}S_r^1.$$

Setting $\varphi = (\varphi_t^0, \varphi_t^1)_{t\in[0,T]}$ with $\varphi_t^1 := \frac{\psi(t)}{\sigma}$ and $\varphi_t^0 := \mathbb{E}^Q[\xi] + \int_0^t \varphi_r^1 \mathrm{d}S_r^1 - \varphi_t^1 S_t^1$, we arrive at $\xi = V_T(\varphi)$. □

The aim is to find an arbitrage-free price and a hedging strategy for a European option

$$g(S_T^1), \quad \text{for } g \in C(\mathbb{R};\mathbb{R}),$$

for example,

- a call option $g(S_T^1) := (S_T^1 - K)^+ := \max\{0, S_T^1 - K\}$ with strike $K > 0$ and maturity T,
- a put option $g(S_T^1) := (K - S_T^1)^+$ with strike $K > 0$ and maturity T.

The idea is to determine the arbitrage-free price as the cost of a (dynamic) replicating strategy, that is, we are looking for a self-financing trading strategy $\varphi = (\varphi_t^0, \varphi_t^1)_{t\in[0,T]}$ such that

$$g(S_T^1) = V_0(\varphi) + \int_0^T \varphi_t^0\,\mathrm{d}S_t^0 + \int_0^T \varphi_t^1\,\mathrm{d}S_t^1 = V_0(\varphi) + \int_0^T \varphi_t^1\,\mathrm{d}S_t^1.$$

Then, the *law of one price* reveals that the arbitrage-free price $\pi(g(S_T^1))$ of the option $g(S_T^1)$ has to be $V_0(\varphi)$, i.e.

$$\pi(g(S_T^1)) = V_0(\varphi).$$

As an ansatz we apply Itô formula to $F(t, S_t^1)$ with

$$F \in C^{1,2}([0,T]\times\mathbb{R};\mathbb{R}) \quad \text{and} \quad F(T,x) = g(x), \quad x \in \mathbb{R},$$

which leads to

$$F(T, S_T^1) = F(0, S_0^1) + \int_0^T \frac{\partial}{\partial x}F(t, S_t^1)\,\mathrm{d}S_t^1 + \frac{1}{2}\sigma^2 \int_0^T \frac{\partial^2}{\partial x^2}F(t, S_t^1)\,\mathrm{d}t + \int_0^T \frac{\partial}{\partial t}F(t, S_t^1)\,\mathrm{d}t. \tag{8.1}$$

since $\langle S^1\rangle_t = \langle \sigma W\rangle_t = \sigma^2 t$. Hence, we are looking for a function F solving the partial differential equation (PDE)

$$\begin{cases} \frac{1}{2}\sigma^2 \frac{\partial^2}{\partial x^2}F(t,x) + \frac{\partial}{\partial t}F(t,x) & = 0, \quad (t,x) \in [0,T)\times\mathbb{R}, \\ F(T,x) & = g(x). \end{cases} \tag{8.2}$$

Indeed, if $F \in C^{1,2}([0,T]\times\mathbb{R};\mathbb{R})$ is a solution to the terminal value problem (8.2), then

$$g(S_T^1) = F(0, S_0) + \int_0^T \frac{\partial}{\partial x}F(t, S_t^1)\,\mathrm{d}S_t, \quad \mathbb{P}\text{-a.s.}$$

Luckily, the PDE in (8.2) can be explicitly solved. It is, loosely speaking, a time-reversed heat equation. Let us recall that the fundamental solution of the heat equation is given by

$$\varphi_t(x) := \frac{1}{\sqrt{2\pi t}} \exp\Big(-\frac{x^2}{2t}\Big), \quad t > 0, x \in \mathbb{R}.$$

Assuming the growth condition

$$|g(x)| \le C\big(1 + \exp(C|x|)\big)^2 \quad \text{for some } C > 0, \tag{8.3}$$

the function

$$F(t,x) := \int_{\mathbb{R}} g(y)\varphi_{\sigma^2(T-t)}(x-y)\,\mathrm{d}y, \quad (t,x) \in [0,T) \times \mathbb{R}, \tag{8.4}$$

belongs to $C^\infty([0,T) \times \mathbb{R};\mathbb{R})$, satisfies $\lim_{t\to T} F(t,x) = g(x)$ and solves the terminal value problem (8.2).

Corollary 8.3. *Suppose that $g\colon \mathbb{R} \to \mathbb{R}$ satisfies the growth condition (8.3) and that F is defined as in (8.4). Then, the arbitrage-free price of the European option $g(S_T^1)$ is*

$$\pi(g(S_T^1)) = F(0, S_0^1) = \int_{\mathbb{R}} g(y) \frac{1}{\sigma\sqrt{2\pi T}} \exp\Big(-\frac{(S_0^1 - y)^2}{2\sigma^2 T}\Big)\,\mathrm{d}y.$$

Remark 8.4. Since the solution of the heat equation is just the density of the normal distribution, we can also observe a more probabilistic representation for the price of the option $g(S_T^1)$. Recall, that

$$\begin{aligned}
\pi(g(S_T^1)) &= \int_{\mathbb{R}} g(y) \frac{1}{\sigma\sqrt{2\pi T}} \exp\Big(-\frac{(S_0^1 - y)^2}{2\sigma^2 T}\Big)\,\mathrm{d}y \\
&= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} g(S_0^1 + \sigma\sqrt{T}z) \exp\Big(-\frac{z^2}{2}\Big)\,\mathrm{d}z \qquad \big[\text{Substitution: } y = S_0 + \sigma\sqrt{T}z\big] \\
&= \mathbb{E}^Q\Big[g\big(S_T^1\big)\Big],
\end{aligned}$$

since S_T^1 is normally distributed with mean S_0^1 and variance σT, where Q is probability measure defined in the proof of Lemma 8.2. It is important to notice that the representation of the price $\pi(g(S_T^1))$ depends on the volatility parameter σ but does not depend on the drift parameter μ.

A Mathematical Foundation

The appendix collects definitions and results which are treated as known prior knowledge from other lecture course, like Mathematical Finance and Wahrscheinlichkeitstheorie I. If you want to learn about these results properly, please visit these lecture courses. All the results presented in the appendix can also be found in the book [Klenke, 2014].

A.1 Conditional expectation

Definition A.1. Let $X \in L^1$ and $\mathcal{G} \subseteq \mathcal{F}$ be a σ-algebra. A random variable Y is called **conditional expectation** of X given $\mathcal{G}$, denoted by $\mathbb{E}[X|\mathcal{G}] := Y$, if

(i) Y is $\mathcal{G}$-measurable;

(ii) for every $A \in \mathcal{G}$ one has $\mathbb{E}[X\mathbb{1}_A] = \mathbb{E}[Y\mathbb{1}_A]$.

If $X, Y \in L^1$, we set $\mathbb{E}[X|Y] := \mathbb{E}[X|\sigma(Y)]$.

Next, we summarize some properties of the conditional expectation.

Theorem A.2 (Properties of the conditional expectation)**.** *Let $\mathcal{H} \subseteq \mathcal{G} \subseteq \mathcal{F}$ be σ-algebras and $X, Y \in L^1(\Omega, \mathcal{F}, \mathbb{P})$. Then:*

(i) Linearity: *For $\lambda \in \mathbb{R}$ we have $\mathbb{E}[\lambda X + Y|\mathcal{G}] = \lambda\mathbb{E}[X|\mathcal{G}] + \mathbb{E}[Y|\mathcal{G}]$.*

(ii) Monotonicity: *If $X \geq Y$, then $\mathbb{E}[X|\mathcal{G}] \geq \mathbb{E}[Y|\mathcal{G}]$.*

(iii) If $\mathbb{E}[|XY|] < \infty$ and Y is measurable w.r.t. $\mathcal{G}$, then

$$\mathbb{E}[XY|\mathcal{G}] = Y\mathbb{E}[X|\mathcal{G}] \quad \textit{and} \quad \mathbb{E}[Y|\mathcal{G}] = \mathbb{E}[Y|\sigma(Y)] = Y.$$

(iv) Tower property: $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]|\mathcal{H}] = \mathbb{E}[X|\mathcal{H}] = \mathbb{E}[\mathbb{E}[X|\mathcal{H}]|\mathcal{G}]$.

(v) Triangle inequality: $|\mathbb{E}[X\,|\mathcal{G}]| \leq \mathbb{E}[|X|\,|\mathcal{G}]$.

(vi) Independence: *If $\sigma(X)$ and $\mathcal{G}$ are independent, then $\mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X]$.*

(vii) Fatou's lemma: *If the sequence of random variables $(X_n)_{n\in\mathbb{N}}$ such that $X_n \geq c$, then*

$$\mathbb{E}[\liminf_{n\to\infty} X_n|\mathcal{G}] \leq \liminf_{n\to\infty} \mathbb{E}[X_n|\mathcal{G}] \quad \mathbb{P}\textit{-a.s.}$$

(viii) Dominated Convergence: *If the sequence of random variables $(X_n)_{n\in\mathbb{N}}$ such that $|X_n| \leq Y$, then*

$$\lim_{n\to\infty} \mathbb{E}[X_n|\mathcal{G}] = \mathbb{E}[X|\mathcal{G}] \quad \mathbb{P}\textit{-a.s.} \quad \textit{and in } L^1(\mathbb{P}).$$

Proposition A.3 (Conditional Jensen's inequality)**.** *Let $I \subseteq \mathbb{R}$ be an interval, let $\varphi\colon I \to \mathbb{R}$ be convex and let X be an I-valued random variable on $(\Omega, \mathcal{F}, \mathbb{P})$. If $\mathbb{E}[|X|] < \infty$ and $\mathcal{G} \subseteq \mathcal{F}$ be a σ-algebra, then*

$$\varphi(\mathbb{E}[X|\mathcal{G}]) \leq \mathbb{E}[\varphi(X)|\mathcal{G}] \leq \infty.$$

A.2 Filtration, stochastic processes and stopping times

Let us fix an arbitrary set $\mathcal{I} \subseteq [0,\infty)$. We mostly care about $\mathcal{I} = \mathcal{T} := \{0,1,\ldots,N\}$, $\mathcal{I} = \mathbb{N}$, $\mathcal{I} = [0,T]$ and $\mathcal{I} = [0,\infty)$.

Definition A.4.

- A family of random variables $(X_t)_{t\in\mathcal{I}}$ (with values in $\mathbb{R}^d$) is called **stochastic process** with index set $\mathcal{I}$ and range $\mathbb{R}^d$.
- A family of σ-algebras $(\mathcal{F}_t)_{t\in\mathcal{I}} \subseteq \mathcal{F}$ is called **filtration** if $\mathcal{F}_s \subseteq \mathcal{F}_t$ for $s,t \in \mathcal{I}$ with $s \le t$.
- A stochastic process $(X_t)_{t\in\mathcal{I}}$ is called **adapted** to $(\mathcal{F}_t)_{t\in\mathcal{I}}$ if X_t is $\mathcal{F}_t$-measurable.

Remark A.5. A stochastic process $(X_t)_{t\in\mathcal{I}}$ is always adapted to the filtration $\mathcal{F}_t := \sigma(X_s : s \in \mathcal{I},\, s \le t)$, i.e., this is the smallest filtration to which the process $(X_t)_{t\in\mathcal{I}}$ is adapted.

As a probabilistic base we fix a filtered probability space $(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\in\mathcal{I}},\mathbb{P})$, i.e., a probability space $(\Omega,\mathcal{F},\mathbb{P})$ equipped with a filtration $(\mathcal{F}_t)_{t\in\mathcal{I}}$.

Definition A.6. A random variable τ with values in $\mathcal{I} \cup \{\infty\}$ is called a **stopping time** (with respect to $(\mathcal{F}_t)_{t\in\mathcal{I}}$) if

$$\{\tau \le t\} \in \mathcal{F}_t \quad \text{for any } t \in \mathcal{I}.$$

Lemma A.7. *Let σ and τ be stopping times. Then:*

(i) $\sigma \vee \tau := \max\{\sigma,\tau\}$ and $\sigma \wedge \tau := \min\{\sigma,\tau\}$ are stopping times.

(ii) If $\sigma,\tau \ge 0$ and $\mathcal{I} \subseteq [0,\infty)$ is closed under addition, then $\sigma+\tau$ is a stopping time.

Definition A.8. Let τ be a stopping time. The **σ-algebra of τ-past** is defined as

$$\mathcal{F}_\tau := \big\{A \in \mathcal{F} \,:\, A \cap \{\tau \le t\} \in \mathcal{F}_t \text{ for any } t \in \mathcal{I}\big\}.$$

Lemma A.9. *If σ and τ are stopping times with $\sigma \le \tau$, then $\mathcal{F}_\sigma \subseteq \mathcal{F}_\tau$.*

A.3 Martingales

Let us fix an arbitrary set $\mathcal{I} \subseteq [0,\infty]$, for instance, $\mathcal{I} = \mathcal{T} := \{0,1,\ldots,T\}$ and $\mathcal{I} = [0,T]$.

Definition A.10. Let $(X_t)_{t\in\mathcal{I}}$ be a real-valued $(\mathcal{F}_t)$-adapted stochastic process with $\mathbb{E}[|X_t|] < \infty$ for all $t \in \mathcal{I}$. X is called a

- **martingale** if $\mathbb{E}[X_t|\mathcal{F}_s] = X_s$,
- **sub-martingale** if $\mathbb{E}[X_t|\mathcal{F}_s] \ge X_s$,
- **super-martingale** if $\mathbb{E}[X_t|\mathcal{F}_s] \le X_s$,

for all $s,t \in \mathcal{I}$ with $s \le t$.

Remark A.11. Every martingale is also a sub- and a super-martingale. For a martingale $(X_t)_{t\in\mathcal{I}}$, the map $t \mapsto \mathbb{E}[X_t]$ is constant as

$$\mathbb{E}[X_t] = \mathbb{E}[\mathbb{E}[X_t|\mathcal{F}_0]] = \mathbb{E}[X_0], \quad t \in \mathcal{I}.$$

Proposition A.12. *Let $(X_t)_{t\in\mathcal{I}}$ be a martingale and $\varphi\colon \mathbb{R}\to\mathbb{R}$ be a convex function. If $\mathbb{E}[|\varphi(X_t)|]<\infty$ for all $t\in\mathcal{I}$, then $(\varphi(X_t))_{t\in\mathcal{I}}$ is a sub-martingale.*

In the following we focus on discrete-time martingales, i.e.,

$$\mathcal{I}=\mathcal{T}:=\{0,1,\dots,T\}$$

Theorem A.13 (Doob's optional sampling theorem)**.** *Let σ,τ be stopping times such that $\sigma\le\tau\le T$.*

(i) If $(X_t)_{t\in\mathcal{T}}$ is a martingale, then

$$\mathbb{E}[X_\tau|\mathcal{F}_\sigma]=X_\sigma \quad \textit{and thus} \quad \mathbb{E}[X_\tau]=X_0.$$

(ii) If $(X_t)_{t\in\mathcal{T}}$ is a sub-martingale (super-martingale), then

$$\mathbb{E}[X_\tau|\mathcal{F}_\sigma]\ge X_\sigma \quad \big(\mathbb{E}[X_\tau|\mathcal{F}_\sigma]\le X_\sigma\big).$$

As an immediate consequence of Doob's optional sampling theorem (Theorem A.13), we obtain the following corollary.

Corollary A.14 (Doob's stopping theorem)**.** *Let τ be a bounded stopping time and $(X_t)_{t\in\mathcal{T}}$ be a martingale. Then, one has*

$$\mathbb{E}[|X_\tau|]<\infty \quad \textit{and} \quad \mathbb{E}[X_\tau]=X_0.$$

The analog statement holds for sub- and super-martingales.

Let $(X_t)_{t\in\mathcal{T}}$ be an adapted stochastic process and τ be a stopping time. The stopped process $(X_t^\tau)_{t\in\mathcal{T}}$ is defined as

$$X_t^\tau := X_{t\wedge\tau} := \begin{cases} X_t(\omega) & \text{if } t\le\tau(\omega)\\ X_{\tau(\omega)}(\omega) & \text{if } t>\tau(\omega)\end{cases} \quad \text{for } \omega\in\Omega \quad \text{and for } t\in\mathcal{T}.$$

Proposition A.15. *Let $(X_t)_{t\in\mathcal{T}}$ be a (super-, sub-)martingale and τ be a stopping time. Then, the stopped process $(X_t^\tau)_{t\in\mathcal{T}}$ is a (super-, sub-)martingale.*

Another application of Doob's optional sampling theorem leads to the so-called Doob's L^p-inequality.

Proposition A.16 (Doob's L^p-inequality)**.** *Let $(X_t)_{t\in\mathcal{T}}$ be a martingale.*

(i) ***Doob's maximal inequality*** *holds for all $\lambda>0$:*

$$\mathbb{P}\Big(\sup_{t\in\mathcal{T}}|X_t|\geqslant\lambda\Big)\leqslant\frac{1}{\lambda}\mathbb{E}[|X_T|].$$

(ii) For $p>1$ and supposing $X_T\in L^p$, we have ***Doob's L^p-inequality***

$$\mathbb{E}\Big[\sup_{t\in\mathcal{T}}|X_t|^p\Big]\leqslant\Big(\frac{p}{p-1}\Big)^p\mathbb{E}[|X_T|^p].$$

A.4 Backward martingales

The concepts of filtration and martingale do not require the index set to be a subset of $[0, \infty)$. Indeed, one can consider more general index sets. Here we are interested in

$$\mathcal{I} = -\mathbb{N}_0 := \{0, -1, -2, \dots\}, \quad \text{where } \mathbb{N}_0 := \{0, 1, 2, \dots\}.$$

- Let $(\mathcal{F}_t)_{t\in -\mathbb{N}_0}$be a filtration, that is, $(\mathcal{F}_t)_{t\in -\mathbb{N}_0} \subseteq \mathcal{F}$ and $\mathcal{F}_s \subseteq \mathcal{F}_t$ for $s, t \in -\mathbb{N}_0$ with $s < t$.
- Let $(M_t)_{t\in -\mathbb{N}_0}$be a martingale with respect to $(\mathcal{F}_t)_{t\in -\mathbb{N}_0}$, that is, M_t is $\mathcal{F}_t$-measurable, $\mathbb{E}[|M_t|] < \infty$, and $\mathbb{E}[M_t|\mathcal{F}_t] = M_s$ for all $s, t \in -\mathbb{N}_0$ with $s < t$.

The stochastic process $(M_{-t})_{t\in \mathbb{N}_0}$ is called **backwards martingale**. For backwards martingales we have the following convergence theorem.

Theorem A.17. *Let $(M_{-t})_{t\in\mathbb{N}_0}$ be a martingale with respect to the filtration $(\mathcal{F}_t)_{t\in-\mathbb{N}_0}$. Then, there exists a random variable $M_{-\infty}$ such that*

$$\lim_{t\to\infty} M_{-t} = M_{-\infty} \quad \textit{almost surely and in } L^1.$$

Moreover,

$$M_{-\infty} = \mathbb{E}[M_0|\mathcal{F}_\infty], \quad \textit{where } \mathcal{F}_{-\infty} := \bigcap_{t=1}^{\infty} \mathcal{F}_{-t}.$$

A.5 Grönwall's lemma

In the study of differential equations Grönwall's lemma (also called Grönwall's inequality) is a frequently applied tool, in particular, to prove the uniqueness of solutions of differential equations.

Lemma A.18 (Grönwall's lemma)**.** *For $T \in (0, \infty)$ let $g\colon [0, T] \to \mathbb{R}$ be a bounded and measurable function. If there are constants $A, B \in \mathbb{R}$ such that*

$$g(t) \leqslant A + B \int_0^t g(s)\,\mathrm{d}s \qquad \textit{for all } t \in [0, T],$$

then $g(t) \leqslant Ae^{Bt}$.

Proof. Define the function

$$\varphi(t) := B \int_0^t g(s)\,\mathrm{d}s, \quad t \in [0, T].$$

By the fundamental theorem of calculus and the assumptions of the lemma we get

$$\varphi'(t) = Bg(t) \leq AB + B\varphi(t), \quad \text{a.e. } t \in [0, T],$$

which implies

$$\varphi'(t) - B\varphi(t) \leq AB.$$

Multiplying both side with $\exp(-Bt)$, we observe that

$$\frac{\mathrm{d}}{\mathrm{d}t}\big(\exp(-Bt)\varphi(t)\big) \leq \exp(-Bt)\big(\varphi'(t) - B\varphi(t)\big) \leq \exp(-Bt)AB$$

and integrating both sides leads to

$$\exp(-Bt)\varphi(t) \leq AB\int_0^t \exp(-Bs)\,\mathrm{d}s.$$

Therefore, we arrive at

$$g(t) \leq A + \varphi(t) \leq A + AB\int_0^t \exp(B(t-s))\,\mathrm{d}s = A\exp(Bt).$$

□

A.6 Radon-Nikodym theorem

This subsections collects results about (probability) measures, which are properly treated in the courses “Functional Analysis” and/or “Maßtheorie”.

Definition A.19. Let Q_1 and Q_2 be measures on a measurable space $(\Omega, \mathcal{F})$ (not necessarily probability measure).

- Q_1 is said to be **σ-finite** if there is a sequence $(A_n)_{n\in\mathbb{N}} \subseteq \mathcal{F}$ such that

$$\Omega = \bigcup_{n\in\mathbb{N}} A_n \quad \text{and} \quad Q_1(A_n) < \infty \quad \text{for } n \in \mathbb{N}.$$

- Q_1 is called **absolutely continuous** w.r.t. Q_2, denoted by $Q_1 \ll Q_2$, if

$$Q_2(A) = 0 \quad \text{for } A \in \mathcal{F} \quad \Rightarrow \quad Q_1(A) = 0.$$

- Q_1 and Q_2 are said to be **equivalent** (in symbols $Q_1 \sim Q_2$) if $Q_1 \ll Q_2$ and $Q_2 \ll Q_1$.
- Q_1 *has a* **density** *w.r.t.* Q_2 if there exists a measurable map $\frac{\mathrm{d}Q_1}{\mathrm{d}Q_2}\colon \Omega \to [0,\infty]$ such that

$$Q_1(A) = \int_\Omega \mathbb{1}_A \frac{\mathrm{d}Q_1}{\mathrm{d}Q_2}\,\mathrm{d}Q_2.$$

A fundamental theorem in functional analysis as well as measure theory is the so-called Radon-Nikodym theorem.

Theorem A.20 (Radon-Nikodym theorem)**.** *Let Q_1 and Q_2 be σ-finite measures on $(\Omega, \mathcal{F})$. Then,*

$$Q_1 \text{ has a density w.r.t. } Q_2 \quad \Leftrightarrow \quad Q_1 \ll Q_2.$$

In this case, the density $\frac{\mathrm{d}Q_1}{\mathrm{d}Q_2}$ is $\mathcal{F}$-measurable and $\mathbb{P}$-a.s. finite. $\frac{\mathrm{d}Q_1}{\mathrm{d}Q_2}$ is called the ***Radon-Nikodym density (also called Radon-Nikodym derivative)*** *of Q_1 w.r.t. Q_2.*

Proof. See see Corollary 7.34 in [Klenke, 2014]. □

Let us restrict our attention to two probability measures Q_1 and Q_2 on $(\Omega, \mathcal{F})$.

Lemma A.21. *If $Q_1 \ll Q_2$ on $(\Omega, \mathcal{F})$, then*

$$Q_1 \sim Q_2 \quad \Leftrightarrow \quad \frac{\mathrm{d}Q_1}{\mathrm{d}Q_2} > 0, \quad Q_2\text{-}a.s.$$

In this case, the Radon-Nikodym density of Q_2 w.r.t. Q_1 is given by

$$\frac{Q_2}{Q_1} = \left(\frac{\mathrm{d}Q_1}{\mathrm{d}Q_2}\right)^{-1}.$$

B Miscellaneous

B.1 Dictionary English-German

English	**German**
absolutely continuous	absolut stetig
adapted	adaptiert
almost sure convergence	fast sichere Konvergenz
bounded	beschränkt
contingent claim	Zahlungsanspruch
continuous	stetig
countable	abzählbar
conditional expectation	bedingter Erwartungswert
density	Dichte
density process	Dichteprozess
derivative	Ableitung
differentiable	differenzierbar
dominated convergence theorem	Satz von der majorisierten Konvergenz
expectation	Erwartungswert
equivalent	äquivalent
Fatou's lemma	Lemma von Fatou
filtered probability space	gefilterter Wahrscheinlichkeitsraum
filtration	Filtration, Filtrierung
(financial) derivative	Derivat
identically distributed	identisch verteilt
independent	unabhängig
indistinguishable	ununterscheidbar
inequality	Ungleichung
integers	ganze Zahlen
integrable	integrierbar
integration	Integration
integral	Integral
intermediate value theorem	Zwischenwertsatz
martingale	Martingal
maturity	Fälligkeit
measure	Maß
measurable space	Messraum, messbarer Raum
monotone convergence theorem	Satz von der monotonen Konvergenz
natural numbers	natürliche Zahlen
$\mathbb{P}$-almost surely	$\mathbb{P}$-fast sicher
ordinary differential equation	gewöhnliche Differentialgleichung
partial differential equation	partielle Differentialgleichung
power set	Potenzmenge
predictable	vorhersehbar
probability measure	Wahrscheinlichkeitsmaß
probability space	Wahrscheinlichkeitsraum

English	**German**
quadratic variation	quadratische Variation
Radon-Nikodym density	Radon-Nikodym-Dichte
random variable	Zufallsgröße, Zufallsvariable
rational numbers	rationale Zahlen
real numbers	reelle Zahlen
σ-algebra	σ-Algebra
sample path	Pfad
sample space	Ergebnisraum
set	Menge
stochastic process	stochastischer Prozess
trading strategy	Handelsstrategie
trajectory	Trajektorie
stopping time	Stoppzeit
triangle inequality	Dreiecksungleichung
tower property	Turmeigenschaft
uniformly integrable	gleichgradig integrierbar
variance	Varianz
volatility	Volatilität, Schwankungsanfälligkeit

B.2 English abbreviations

Abbreviation	Meaning
ad.	adapted
a.e.	almost everywhere
a.s.	almost surely
bdd	bounded
BM	Brownian motion
cts	continuous
eq.	equation
fct.	function
ineq.	inequality
iid	independent and identically distributed
iff	if and only if
loc.	local
mart.	martingale
mb.	measurable
ODE	ordinary differential equations
$\mathbb{P}$-a.s.	$\mathbb{P}$-almost surely
PDE	partial differential equations
pred.	predictable
prob.	probability
RN density	Radon-Nikodym density
r.v.	random variable
SDE	stochastic differential equation
s.t.	such that
stoch.	stochastic
trad.	trading
u.i.	uniformly integrable
w.l.o.g.	without lost of generality
w.r.t.	with respect to

References

[Klenke, 2014] Klenke, A. (2014). *Probability Theory*. Springer-Verlag London.