Theory of Stochastic Calculus*

BY PAWEŁ DUCH

Email: pawel.duch@epf.ch

Table of contents

1 Informal introduction . . . 2

2 Some elements of probability theory . . . 2

3 Definition of Brownian motion and basic properties . . . 3

4 Construction of Brownian motion . . . 7

5 Conditional expectation . . . 9

6 Martingales . . . 11

7 Integration with respect to bounded variation processes . . . 15

8 Stochastic integral . . . 18

8.1 Integral for simple predictable processes . . . 18
8.2 Isometric Itô integral . . . 20
8.3 Localization . . . 23

9 Integral with respect to continuous local martingale . . . 25

9.1 Doob-Meyer decomposition . . . 25
9.2 Integral for elementary processes . . . 26
9.3 Isometric Itô integral . . . 27
9.4 Localization . . . 28

10 Quadratic covariation . . . 29

11 Further properties of the stochastic integral . . . 30

11.1 Dominated convergence for stochastic integrals . . . 30
11.2 Integration by substitution . . . 31

12 Continuous semimartingales . . . 34

13 Itô formula . . . 35

14 Stochastic differential equations . . . 39

15 Girsanov theorem . . . 43

16 Martingale representation theorem . . . 47

17 PDE problems and diffusion processes . . . 48

17.1 Dirichlet–Poisson problem . . . 49
17.2 Cauchy problem for parabolic equations . . . 51

18 Markov property . . . 52

Bibliography . . . 55

*. These notes are under construction and will be updated regularly. If you find errors or typos, please let me know.

1 Informal introduction

Many physical phenomena can be modeled as stochastic processes that satisfy certain equations involving random terms or coefficients. One example of such an equation is

$$\frac{\mathrm{d}X_t}{\mathrm{d}t} = b(X_t) + \sigma(X_t)\xi_t,$$

where $(\xi_t)_{t\geqslant 0}$ is a certain collection of random variables (a stochastic process) called the **white noise** and b, σ are deterministic functions. Imagine that X_t is the price of an asset at time t. The first term on the RHS of the above equation models intrinsic predictable trends of price change. The second term represents unpredictable changes of price due to influences of some random events. It is natural to assume that these events are *independent*. We would like to define the white noise $(\xi_t)_{t\in\mathbb{R}}$ as a collection of Gaussian random variables such that their mean vanishes, ξ_s and ξ_t are *independent* whenever $s\neq t$ and $\int_0^t \xi_s\,\mathrm{d}s \neq 0$ for $t\geqslant 0$. Unfortunately, one shows that there is no collection $(\xi_t)_{t\in\mathbb{R}}$ satisfying the above conditions. Nevertheless, it is possible to give meaning to equations like (1). The idea is to first construct a stochastic process $(B_t)_{t\geqslant 0}$ that, at least *formally*, solves the equation

$$\frac{\mathrm{d}B_t}{\mathrm{d}t} = \xi_t.$$

We call such a process $(B_t)_{t\geqslant 0}$ a **Brownian motion**. As we will see, $(B_t)_{t\geqslant 0}$ is very irregular as a function of time and, in particular, is not differentiable. In order to make sense of (1) we rewrite it as

$$\mathrm{d}X_t = b(X_t)\,\mathrm{d}t + \sigma(X_t)\mathrm{d}B_t.$$

We say that $(X_t)_{t\geqslant 0}$ is a solution of the above **stochastic differential equation** if

$$X_t = X_0 + \int_0^t b(X_s)\mathrm{d}s + \int_0^t \sigma(X_s)\,\mathrm{d}B_s.$$

The last equation above is meaningful provided we can define an integral

$$\int_0^T Y_s\,\mathrm{d}B_s$$

of a sufficiently generic stochastic process $(Y_t)_{t\geqslant 0}$ with respect to a Brownian motion $(B_t)_{t\geqslant 0}$. We call the above integral the **stochastic integral**. Because of very irregular nature of Brownian motion the construction of the stochastic integral is quite nontrivial and will be one of the subjects of this course.

Content of the course:

- Brownian motion and martingales.
- Construction of the stochastic integral and its properties.
- Stochastic differential equations.
- Link between partial differential equations and stochastic processes.

2 Some elements of probability theory

Definition 2.1. *A triple $(\Omega, \mathcal{F}, \mathbb{P})$ is called a **probability space** provided Ω is a nonempty set, $\mathcal{F}$ is a σ-algebra of subsets of Ω, and $\mathbb{P}:\mathcal{F}\to[0,1]$ is a probability measure. A set $E\in\mathcal{F}$ is called an **event**, a point $\omega\in\Omega$ is called a **sample point** and $\mathbb{P}(E)$ is the probability of the event E.*

The smallest σ-algebra containing all the open subsets of $\mathbb{R}^d$ is called the **Borel σ-algebra**. Note that elements of this algebra, called **Borel sets**, can be formed from open sets through the operations of countable union, countable intersection, and relative complement.

Definition 2.2. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space.*

- *A map $X : \Omega \to \mathbb{R}^d$ is called a **random variable** if $X^{-1}(B) \in \mathcal{F}$ for all Borel sets $B \subset \mathbb{R}^d$.*
- *We write $\mathbb{P}(X \in B)$ for $\mathbb{P}(X^{-1}(B))$, that is, the probability that X takes values in B.*
- *We call $\mathbb{E}(X) := \int_\Omega X(\omega)\, \mathbb{P}(\mathrm{d}\omega)$ the **expected value** of X.*
- *Random variables $X_1, \ldots, X_n$ are **independent** if*

$$\mathbb{P}(X_1 \in B_1, \ldots, X_n \in B_n) = \mathbb{P}(X_1 \in B_1) \ldots \mathbb{P}(X_n \in B_n)$$

  for all Borel sets $B_1, \ldots, B_n$.
- *A collection $(X_t)_{t \in I}$ of random variables indexed by elements of some set I is called a **stochastic process**. For each sample point $\omega \in \Omega$ the map $t \mapsto X_t(\omega)$ is the corresponding **sample path**.*
- *We say that a process $(X_t)_{t \in I}$ is **continuous** (resp. **a.s. continuous**) if its sample paths are continuous (resp. a.s. continuous), that is the map $t \mapsto X_t(\omega)$ is continuous for all (resp. almost all) $\omega \in \Omega$.*

We say that a random variable $X : \Omega \to \mathbb{R}$ has **Gaussian** (or **normal**) distribution with mean m and variance σ^2, and write X is $\mathcal{N}(m, \sigma^2)$, if

$$\mathbb{P}(X < a) = \int_{-\infty}^{a} f(x)\, \mathrm{d}x, \qquad f(x) \equiv p(x; m, \sigma^2) := \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{|x - m|^2}{2\sigma^2}\right).$$

The map $a \mapsto \mathbb{P}(X < a)$ is called the **distribution function** of a random variable X and the function f related to $\mathbb{P}(X < a)$ by the formula above is called the **density function** of X (note that not every random variable has a density).

3 Definition of Brownian motion and basic properties

Definition 3.1. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A stochastic process $(B_t)_{t \geqslant 0}$ is called a **Brownian motion** (or a **Wiener process**) if it has the following properties:*

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

(ii). The increments of $(B_t)_{t \geqslant 0}$ are independent, that is, for every finite set of times $0 \leqslant t_1 < t_2 < \cdots < t_n < \infty$ the random variables

$$B_{t_2} - B_{t_1}, B_{t_3} - B_{t_2}, \ldots, B_{t_n} - B_{t_{n-1}}$$

are independent.

(iii). For any $0 \leqslant s < t < \infty$ the increment $B_t - B_s$ is $\mathcal{N}(0, t - s)$.

(iv). $(B_t)_{t \geqslant 0}$ is a.s. continuous.

Remark 3.2. A Brownian motion with initial point x is a stochastic process $(B_t)_{t\geqslant 0}$ such that $B_0 = x$ a.s. and the conditions (ii)-(iv) introduced above are satisfied.

Remark 3.3. One of the many reasons that Brownian motion is important in probability theory is that it can be obtained as the **continuous-time limit** of a simple symmetric random walk when the step size and time step shrink appropriately. A **simple symmetric random walk** starts at zero and at each step moves $+1$ or -1 with equal probability. More precisely, it is a discrete stochastic process $(S_n)_{n\in\mathbb{N}_0}$ defined by the conditions and

$$S_0 = 0, \qquad S_n = \sum_{i=1}^{n} X_i, \quad n \in \mathbb{N}_+,$$

where $X_1, X_2, X_3, \ldots$ are i.i.d. random variables taking values $+1$ or -1 with equal probability. Let $\varepsilon > 0$ and define

$$B_t^{(\varepsilon)} := \varepsilon S_{t/\varepsilon^2}, \quad t \in \{0, \varepsilon^2, 2\varepsilon^2, \ldots\}.$$

Note that $B_t^{(\varepsilon)}$ describes a walk that starts at zero and at each time $t \in \{0, \varepsilon^2, 2\varepsilon^2, \ldots\}$ moves $+\varepsilon$ or $-\varepsilon$ with equal probability. By the central limit theorem we know that the distribution of $\frac{S_n}{\sqrt{n}}$ converges to $\mathcal{N}(0,1)$ as $n \to \infty$. Hence, the distribution of

$$B_t^{(\varepsilon)} = \sqrt{t}\,\frac{S_{t/\varepsilon^2}}{\sqrt{t/\varepsilon^2}}$$

converges to $\mathcal{N}(0,t)$ as $\varepsilon \searrow 0$. Since increments of $B_t^{(\varepsilon)}$ are independent, this suggest that the process $\big(B_t^{(\varepsilon)}\big)_{t\geqslant 0}$ (obtained from $\big(B_t^{(\varepsilon)}\big)_{t\in\{0,\varepsilon^2,2\varepsilon^2,\ldots\}}$ by linear interpolation) should converge to a Brownian motion.

Remark 3.4. The history of the Brownian motion began in 1827 when a botanist **Robert Brown** looked through a microscope at pollen grains suspended in water and discovered the pollen was moving in a random fashion. He noted that the path of a given particle is very irregular and the motions of two distinct particles appear to be independent. It wasn't until later that scientists realized the true cause of this motion was not biological, but rather physical. The motion was due to the random collisions between the pollen particles and the much smaller water molecules, which were in constant, chaotic motion. This phenomenon, now known as **Brownian motion**, is a type of random movement that is a key concept in both physics and mathematics. In 1905 **Albert Einstein** provided a mathematical explanation of Brownian motion. He suggested that the random movement of particles was a result of thermal fluctuations at the molecular level, leading to what we now call diffusion. In 1900 the French mathematician **Louis Bachelier** applied a random process model to describe stock prices, which shares similarities with the randomness seen in Brownian motion. A rigorous construction of Brownian motion was given by **Norbert Wiener** in 1923.

Definition 3.5. *If a stochastic process $(X_t)_{t\geqslant 0}$ has the property that for every $0 \leqslant t_1 < t_2 < \ldots < t_n < \infty$ the vector $(X_{t_1}, \ldots, X_{t_n})$ has a multivariate Gaussian distribution, then $(X_t)_{t\geqslant 0}$ is called a* ***Gaussian process****.*

Lemma 3.6. *A process $(B_t)_{t\geqslant 0}$ is a Brownian motion iff it is Gaussian, a.s. continuous and for all $s,t\geqslant 0$ it holds that*

$$\mathbb{E}B_t=0, \qquad \mathbb{E}(B_sB_t)=s\wedge t.$$

Notation 3.7. *We set $s\wedge t:=\min(s,t)$ and $s\vee t:=\max(s,t)$.*

Proof. Suppose that $(B_t)_{t\geqslant 0}$ is a Brownian motion. Then for every $0\leqslant t_1<t_2<\ldots<t_n<\infty$ the vector

$$(B_{t_1}, B_{t_2}-B_{t_1}, B_{t_3}-B_{t_2}, \ldots, B_{t_n}-B_{t_{n-1}}) \tag{3.1}$$

has a multivariate Gaussian distribution. It follows that $(B_{t_1},\ldots,B_{t_n})$ also has a multivariate Gaussian distribution. In consequence, $(B_t)_{t\geqslant 0}$ is Gaussian. It is evident that $\mathbb{E}B_t=0$. Suppose that $s<t$. Since B_s and B_t-B_s are independent, we have $\mathbb{E}(B_s(B_t-B_s))=\mathbb{E}(B_s)\mathbb{E}(B_t-B_s)=0$. As a result, $\mathbb{E}(B_tB_s)=\mathbb{E}(B_s^2)+\mathbb{E}(B_s(B_t-B_s))=s$. This proves that $\mathbb{E}(B_tB_s)=t\wedge s$.

Now let us prove the reverse implication. We have to show that $(B_t)_{t\geqslant 0}$ satisfies the conditions (i)-(iii) stated in Def. 3.1 (the condition (iv) is satisfied by assumption). Since $\mathbb{E}(B_0^2)=0$, $B_0=0$ a.s. and (i) holds true. To prove (ii) we have to demonstrate that the vector (3.1) has a diagonal covariance matrix. To confirm the vanishing of the off-diagonal terms we use $\mathbb{E}B_t=0$ and $\mathbb{E}(B_sB_t)=s\wedge t$ to show that for $i<j$ it holds that

$$\begin{aligned}\mathrm{Cov}(B_{t_i}-B_{t_{i-1}}, B_{t_j}-B_{t_{j-1}}) &= \mathbb{E}((B_{t_i}-B_{t_{i-1}})(B_{t_j}-B_{t_{j-1}}))\\ &= \mathbb{E}(B_{t_i}B_{t_j})-\mathbb{E}(B_{t_i}B_{t_{j-1}})-\mathbb{E}(B_{t_{i-1}}B_{t_j})+\mathbb{E}(B_{t_{i-1}}B_{t_{j-1}})\\ &= t_i-t_i-t_{i-1}+t_{i-1}=0.\end{aligned}$$

By a similar computation we obtain $\mathbb{E}(B_t-B_s)^2=t-s$. Since $\mathbb{E}(B_t-B_s)=0$, we conclude that B_t-B_s is $N(0,t-s)$. Hence, the condition (iii) is satisfied. This finishes the proof of the lemma. ☐

Lemma 3.8. *Let $(B_t)_{t\geqslant 0}$ be a Brownian motion.*

(i). (Invariance under translations in time). Let $a>0$. The process $(X_t)_{t\geqslant 0}=(B_{t+a}-B_a)_{t\geqslant 0}$ is a Brownian motion.

(ii). (Scaling property). Let $a>0$. The process $(Y_t)_{t\geqslant 0}=\left(\frac{1}{a}B_{a^2t}\right)_{t\geqslant 0}$ is a Brownian motion.

(iii). (Time inversion). Let $Z_0=0$ and $Z_t=tB_{1/t}$ for $t>0$. Then the process $(Z_t)_{t\geqslant 0}$ is a Brownian motion.

Proof. We use the characterization of Brownian motion given in Lemma 3.6. It is evident that all of the processes defined in the statement are Gaussian with mean zero. A simple computation using $\mathbb{E}B_sB_t=s\wedge t$ shows that $\mathbb{E}X_sX_t=s\wedge t$, $\mathbb{E}Y_sY_t=s\wedge t$ and $\mathbb{E}Z_sZ_t=s\wedge t$. For example, we have

$$\mathbb{E}Z_sZ_t=st\,\mathbb{E}(B_{1/s}B_{1/t})=st\,(1/s\wedge 1/t)=s\wedge t.$$

This shows that the covariance condition stated in Lemma 3.6 is fulfilled. It is also clear that $(X_t)_{t\geqslant 0}$ and $(Y_t)_{t\geqslant 0}$ are a.s. continuous and $(Z_t)_{t\geqslant 0}$ is a.s. continuous away from zero.

To complete the proof of the lemma it remains to establish the continuity of $(Z_t)_{t\geqslant 0}$ at zero. To this end, consider two events

$$A = \bigcap_{m\in\mathbb{N}_+} \bigcup_{n\in\mathbb{N}_+} \bigcap_{k\in\mathbb{N}_+} \left\{|B_{t_{n,k}}| \leqslant \frac{1}{m}\right\} = \left\{\lim_{t\searrow 0} B_t = 0\right\},$$

$$B = \bigcap_{m\in\mathbb{N}_+} \bigcup_{n\in\mathbb{N}_+} \bigcap_{k\in\mathbb{N}_+} \left\{|Z_{t_{n,k}}| \leqslant \frac{1}{m}\right\} = \left\{\lim_{t\searrow 0} Z_t = 0\right\}.$$

Here for every $n\in\mathbb{N}_+$, $(t_{n,k})_{k\in\mathbb{N}_+}$ is a arbitrarily fixed sequence of all elements of $\mathbb{Q}\cap(0,1/n)$, where $\mathbb{Q}$ is the set of rational numbers. We used above the fact that if $f:[0,\infty)\to\mathbb{R}$ is continuous away from the origin, then $\lim_{t\searrow 0} f(t)=0$ iff

$$\forall_{m\in\mathbb{N}_+}\exists_{n\in\mathbb{N}_+}\forall_{t\in\mathbb{Q}\cap(0,1/n)}|f(t)| \leqslant \frac{1}{m} \quad\Leftrightarrow\quad \forall_{m\in\mathbb{N}_+}\exists_{n\in\mathbb{N}_+}\forall_{k\in\mathbb{N}_+}|f(t_{n,k})| \leqslant \frac{1}{m}.$$

Since for all $0\leqslant t_1 < t_2 < \ldots < t_n < \infty$ the vectors $(B_{t_1},\ldots,B_{t_n})$ and $(Z_{t_1},\ldots,Z_{t_n})$ have the same distribution (both vectors have Gaussian distribution with mean zero and the same covariance), by the continuity of the probability measure from below and above we conclude that the events A and B have the same probability. Since $\mathbb{P}(A)=1$ by the condition (iv) of Brownian motion, we have $\mathbb{P}(B)=1$. This proves that Z_t is a.s. continuous at $t=0$. $\square$

Lemma 3.9. *Let $(B_t)_{t\geqslant 0}$ be a Brownian motion. For all $0<t_1<t_2<\ldots<t_n<\infty$ the joint distribution of $(B_{t_1},\ldots,B_{t_n})$ is given by*

$$\begin{aligned}&\mathbb{P}(a_1\leqslant B_{t_1}\leqslant b_1, a_2\leqslant B_{t_2}\leqslant b_2,\ldots,a_n\leqslant B_{t_n}\leqslant b_n)\\ &\quad = \int_{a_1}^{b_1}\ldots\int_{a_n}^{b_n} p(x_1;t_1)p(x_2-x_1;t_2-t_1)\ldots p(x_n-x_{n-1};t_n-t_{n-1})\,\mathrm{d}x_1\ldots\mathrm{d}x_n,\end{aligned}$$

where $p(x;t):=\frac{1}{\sqrt{2\pi t}}\exp(-|x|^2/2t)$.

Proof. Set $t_0=0$ and define the following random vectors

$$\vec{X} = \begin{pmatrix} B_{t_1}\\ B_{t_2}\\ \vdots\\ B_{t_n}\end{pmatrix}, \qquad \vec{Y} = \begin{pmatrix} B_{t_1}\\ B_{t_2}-B_{t_1}\\ \vdots\\ B_{t_n}-B_{t_{n-1}}\end{pmatrix}.$$

By the conditions (i) and (ii) of Brownian motion the components of $\vec{Y}$ are independent. By condition (iii) the random variable $B_{t_i}-B_{t_{i-1}}$ is normally distributed with mean zero and variance t_i-t_{i-1}. Hence, the joint density $f_{\vec{Y}}$ of $\vec{Y}$ satisfies the equation

$$f_{\vec{Y}}(x_1,\ldots,x_n) = p(x_1;t_1)p(x_2;t_2-t_1)\ldots p(x_n;t_n-t_{n-1}).$$

Observe that $\vec{Y}=M\vec{X}$, where the matrix M is defined by the equation

$$M\begin{pmatrix} x_1\\ x_2\\ \vdots\\ x_n\end{pmatrix} = \begin{pmatrix} x_1\\ x_2-x_1\\ \vdots\\ x_n-x_{n-1}\end{pmatrix}.$$

As a result, the joint density $f_{\vec{X}}$ of $\vec{X}$ satisfies the equation

$$\begin{aligned} f_{\vec{X}}(x_1,\ldots,x_n) &= f_{\vec{X}}(\vec{x}) = f_{\vec{Y}}(M\vec{x})\,|\det M| \\ &= f_{\vec{Y}}(M\vec{x}) = f_{\vec{Y}}(x_1, x_2 - x_1, \ldots, x_n - x_{n-1}) \\ &= p(x_1;t_1)p(x_2 - x_1; t_2 - t_1)\ldots p(x_n - x_{n-1}; t_n - t_{n-1}), \end{aligned}$$

where we used the fact that $|\det M| = 1$ (in particular, M is invertible). This completes the proof. $\square$

4 Construction of Brownian motion

In this section, we present the construction of Brownian motion given originally in 1934 by Poley and Wiener. We only construct a Brownian motion B_t for $t \in [0,\pi]$. The construction for all $t \geqslant 0$ requires an additional step (see Exercise 1, Sheet 2). Let $(X_n)_{n\in\mathbb{N}_0}$ be a sequence of i.i.d. $\mathcal{N}(0,1)$ random variables in a probability space $(\Omega,\mathcal{F},\mathbb{P})$. Define

$$B_t^{(m)} = \frac{t}{\sqrt{\pi}} X_0 + \sum_{n=1}^{2^m-1} \sqrt{\frac{2}{\pi}} \frac{\sin(\mathrm{nt})}{n} X_n, \qquad t \in [0,\pi], \quad m \in \mathbb{N}_0.$$

Observe that for any fixed $m \in \mathbb{N}_0$ and $\omega \in \Omega$ the function $t \to B_t^{(m)}(\omega)$ is smooth. It is also easy to see that for any fixed $t \in [0,\pi]$ the series converges in mean square. Indeed, $\mathbb{E}(X_n X_l) = \delta_{n,l}$ by independence of $(X_n)_{n\in\mathbb{N}_0}$ and $\mathbb{E}X_n = 0$, $\mathbb{E}X_n^2 = 1$. Hence,

$$\mathbb{E}\big|B_t^{(m)} - B_t^{(l)}\big|^2 = \frac{2}{\pi}\mathbb{E}\left|\sum_{n=2^k-1}^{2^m-1} X_n \frac{\sin(\mathrm{nt})}{n}\right|^2 = \frac{2}{\pi}\sum_{n=2^k-1}^{2^m-1} \mathbb{E}(X_n^2)\frac{\sin^2(\mathrm{nt})}{n^2} \leqslant \frac{2}{\pi}\sum_{n=2^k-1}^{2^m-1}\frac{1}{n^2},$$

which implies that the series defined by partial sums is Cauchy. Hence, we can define in the probability space $(\Omega,\mathcal{F},\mathbb{P})$ a stochastic process $(B_t)_{t\in[0,\pi]}$ by the equality

$$B_t = \frac{t}{\sqrt{\pi}} X_0 + \sum_{n=1}^{\infty} \sqrt{\frac{2}{\pi}} \frac{\sin(\mathrm{nt})}{n} X_n, \qquad t \in [0,\pi].$$

We shall prove that $(B_t)_{t\in[0,\pi]}$ is a Brownian motion. To this end, we will use the characterization of Brownian motion given in Lemma 3.6. It is evident that $(B_t)_{t\in[0,\pi]}$ is Gaussian, $B_0 = 0$ and $\mathbb{E}B_t = 0$. To compute the covariance observe that

$$\mathbb{E}\big(B_s^{(m)} B_t^{(m)}\big) = \int_0^s \int_0^t \mathbb{E}\big(\partial_u B_u^{(m)}\, \partial_w B_w^{(m)}\big)\,\mathrm{d}u\mathrm{d}w \tag{4.1}$$

and

$$\partial_t B_t^{(m)} = \frac{1}{\sqrt{\pi}} X_0 + \sum_{n=1}^{2^m-1} \sqrt{\frac{2}{\pi}} \cos(\mathrm{nt}) X_n = \sum_{n=0}^{2^m-1} X_n\, e_n(t),$$

where

$$e_0(t) := \frac{1}{\sqrt{\pi}}, \qquad e_n(t) := \sqrt{\frac{2}{\pi}}\cos(nt), \quad n \in \mathbb{N}_+,$$

is an orthonormal basis of $L^2([0,\pi])$. Using $\mathbb{E}(X_n X_l) = \delta_{n,l}$ we obtain

$$\mathbb{E}\big(\partial_u B_u^{(m)}\, \partial_w B_w^{(m)}\big) = \sum_{n=0}^{2^m-1} \mathbb{E}(X_n^2)\, e_n(u) e_n(w) = \sum_{n=0}^{2^m-1} e_n(u) e_n(w).$$

Hence, for all $f, g \in L^2([0,\pi])$ we have

$$\begin{aligned}
&\lim_{m\to\infty} \int_0^\pi \int_0^\pi \mathbb{E}\big(\partial_u B_u^{(m)}\, \partial_w B_w^{(m)}\big)\, f(u) g(w)\, \mathrm{d}u \mathrm{d}w \\
&\quad = \lim_{m\to\infty} \sum_{n=0}^{2^m-1} (f, e_n)_{L^2([0,\pi])} (e_n, g)_{L^2([0,\pi])} = (f, g)_{L^2([0,\pi])},
\end{aligned} \tag{4.2}$$

where $(f, g)_{L^2([0,\pi])}$ denotes the scalar product. Let $\mathbf{1}_I$ be the characteristic function of a set I. Since $B_t^{(m)}$ converges to B_t in mean square for all $t \in [0,\pi]$, by (4.1) and (4.2) applied with $f = \mathbf{1}_{[0,t]}$ and $g = \mathbf{1}_{[0,t]}$ we obtain

$$\mathbb{E}(B_s B_t) = \lim_{m\to\infty} \mathbb{E}\big(B_s^{(m)} B_t^{(m)}\big) = (f, g)_{L^2([0,\pi])} = \int_0^\pi \mathbf{1}_{[0,s]}(u) \mathbf{1}_{[0,t]}(u)\, \mathrm{d}u = s \wedge t.$$

In view of Lemma 3.6 in order to complete the proof that $(B_t)_{t\in[0,\pi]}$ is a Brownian motion it remains to show that $(B_t)_{t\in[0,\pi]}$ has a.s. continuous paths. The idea is to prove that $\big(B_t^{(m)}\big)_{t\in[0,\pi]}$ converges a.s. to $(B_t)_{t\in[0,\pi]}$ as $m \to \infty$ uniformly in $t \in [0,\pi]$. Since $\big(B_t^{(m)}\big)_{t\in[0,\pi]}$ has continuous sample paths and uniform convergence preserves continuity this would imply that the sample paths of $(B_t)_{t\in[0,\pi]}$ are a.s. continuous.

Theorem 4.1. *The sequence of stochastic processes $\big(B_t^{(m)}\big)_{t\in[0,\pi]}$ converges as $m \to \infty$ a.s. uniformly in $t \in [0,\pi]$ to $(B_t)_{t\in[0,\pi]}$, that is*

$$\mathbb{P}\bigg(\lim_{m\to\infty} \sup_{t\in[0,\pi]} \big|B_t^{(m)} - B_t\big| = 0\bigg) = 1.$$

The process $(B_t)_{t\in[0,\pi]}$ defined by (4) is a Brownian motion.

Remark 4.2. The following result is known as the **Weierstrass M-test**. Let $(f_n)_{n\in\mathbb{N}_0}$ be a sequence of functions defined on a set E. Suppose there exists a sequence of non-negative constants $(M_n)_{n\in\mathbb{N}_0}$ such that $|f_n(x)| \leqslant M_n$ for all $x \in E$ and all $n \in \mathbb{N}_0$ and $\sum_{n=0}^\infty M_n < \infty$. Then the series $\sum_{n=0}^\infty f_n(x)$ converges uniformly on E.

Lemma 4.3. *For all $l, p \in \mathbb{N}_0$, $p > l$, we have*

$$\mathbb{E}(T_{l,p}^2) \leqslant \frac{p-l}{l^2} + \frac{2(p-l)^{3/2}}{l^2}, \qquad T_{l,p} := \sup_{t\in[0,\pi]} \left|\sum_{n=l}^{p-1} X_n \frac{\sin(nt)}{n}\right|.$$

Proof. See Exercise 2, Sheet 1. It is crucial to use independence of $(X_n)_{n\in\mathbb{N}_0}$. □

Proof of Theorem 4.1. By the argument presented above the statement of the theorem, it is enough to show the uniform convergence. We have

$$B_t^{(m)} = \frac{t}{\sqrt{\pi}} x_0 + \sqrt{\frac{2}{\pi}} \sum_{i=0}^{m-1} f_i(t), \qquad f_i(t) = \sum_{n=2^i}^{2^{i+1}-1} X_n \frac{\sin(nt)}{n}.$$

Note that $|f_i(t)| \leqslant M_i := T_{2^i,2^{i+1}}$. It suffices to show that $\mathbb{P}(\sum_{i=0}^{\infty} M_i < \infty) = 1$ and apply the Weierstrass M-test. Observe that by the Cauchy-Schwarz inequality and Lemma 4.3 we have

$$(\mathbb{E}T_{l,2l})^2 \leqslant \mathbb{E}(T_{l,2l}^2) \leqslant \frac{1}{l} + \frac{2}{l^{1/2}} \leqslant \frac{4}{l^{1/2}}.$$

Consequently,

$$\mathbb{E}M_i = \mathbb{E}T_{2^i,2^{i+1}} \leqslant \frac{2}{2^{i/4}}, \qquad \mathbb{E}\left(\sum_{i=0}^{\infty} M_k\right) = \sum_{i=0}^{\infty} \mathbb{E}(M_k) \leqslant \sum_{i=0}^{\infty} \frac{2}{2^{i/4}} < \infty.$$

Since the random variable $\sum_{i=0}^{\infty} M_k \in [0, \infty]$ has finite expected value, it has to a.s. take finite values, that is $\mathbb{P}(\sum_{i=0}^{\infty} M_i < \infty) = 1$. This finishes the proof. □

5 Conditional expectation

The **conditional probability** of an event A given an event B is defined by

$$\mathbb{P}(A|\, B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} \quad \text{provided } \mathbb{P}(B) > 0.$$

Note that we can think of $B \subset \Omega$ as a new probability space equipped with the probability measure $\mathbb{P}(\cdot|\, B)$. Thus, it is natural to define the **conditional expected value** of a random variable X given an event B as

$$\mathbb{E}(X|\, B) = \int_{\Omega} X(\omega)\, \mathbb{P}(\mathrm{d}\omega|\, B) = \frac{1}{\mathbb{P}(B)} \int_B X(\omega) \mathbb{P}(\mathrm{d}\omega) \quad \text{provided } \mathbb{P}(B) > 0.$$

Assume we are given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a random variable Y such that

$$Y = \begin{cases} a_1 & \text{on } A_1, \\ a_2 & \text{on } A_2, \\ & \vdots \\ a_m & \text{on } A_m, \end{cases}$$

for distinct real numbers $a_1, a_2, \ldots, a_m$ and disjoint events $A_1, A_2, \ldots, A_m$, each of positive probability, whose union is Ω. Define a random variable

$$\mathbb{E}(X|\, Y) := \begin{cases} \mathbb{E}(X|A_1) & \text{on } A_1, \\ \mathbb{E}(X|A_2) & \text{on } A_2, \\ & \vdots \\ \mathbb{E}(X|A_m) & \text{on } A_m. \end{cases}$$

We call the random variable $\mathbb{E}(X|\, Y)$ the conditional expected value of X given Y. Note that (see the definition below):

- $\mathbb{E}(X|\, Y)$ is $\sigma(Y)$-measurable.
- $\int_A X(\omega)\, \mathbb{P}(\mathrm{d}\omega) = \int_A \mathbb{E}(X|\, Y)(\omega) \mathbb{P}(\mathrm{d}\omega)$ for all $A \in \sigma(Y)$.

In what follows, we generalize the above definition of $\mathbb{E}(X|\, Y)$ to arbitrary random variables Y.

Definition 5.1. *Let* $(\Omega, \mathcal{F}, \mathbb{P})$ *be a probability space,* $\mathcal{G}$ *be a sub-*σ*-algebra of* $\mathcal{F}$ *and* $X : \Omega \to \mathbb{R}^d$ *be a random variable. We say that* X *is* $\mathcal{G}$*-measurable if* $X^{-1}(B) \in \mathcal{G}$ *for all Borel sets* $B \subset \mathbb{R}^d$*. The sub-*σ*-algebra of* $\mathcal{F}$ *defined by*

$$\sigma(X) := \{X^{-1}(B) |\, B - \text{Borel subset of } \mathbb{R}^d\}$$

is called the σ*-algebra generated by* X*.*

Remark 5.2. $\sigma(X)$ is the smallest sub-σ-algebra of $\mathcal{F}$ with respect to which X is measurable. $\sigma(X)$ contains all the events that can be expressed in terms of X.

Remark 5.3. If a random variable Y is a Borel function of X, that is, if $Y = f(X)$ for some Borel function f, then Y is $\sigma(X)$-measurable. Conversely, suppose that a random variable Y is $\sigma(X)$-measurable. Then there exists a Borel function f such that $Y = f(X)$.

Definition 5.4. *A random variable* X *is* ***integrable*** *(resp.* ***square-integrable****) if* $\mathbb{E}|X| < \infty$ *(resp.* $\mathbb{E}X^2 < \infty$*). A random variable* X *is* ***bounded*** *if* $|X| < C$ *a.s. for some deterministic* $C > 0$*.*

Theorem 5.5. *Let* $(\Omega, \mathcal{F}, \mathbb{P})$ *be a probability space,* X *be an integrable random variable and* $\mathcal{G}$ *be a sub-*σ*-algebra of* $\mathcal{F}$*. There exists a* $\mathcal{G}$*-measurable random variable* Z *such that*

$$\int_A X(\omega)\, \mathbb{P}(\mathrm{d}\omega) = \int_A Z(\omega) \mathbb{P}(\mathrm{d}\omega) \qquad \text{for all} \quad A \in \mathcal{G}. \tag{5.1}$$

A random variable Z *satisfying the above properties is unique up to* $\mathbb{P}$*-equivalence. We denote by* $\mathbb{E}(X|\mathcal{G})$ *any representative of this equivalence class and call it the conditional expectation of* X *with respect to* $\mathcal{G}$*.*

Proof. See e.g. Sec. 4.2 of [Bal17]. □

Definition 5.6. *Let* $(\Omega, \mathcal{F}, \mathbb{P})$ *be a probability space and* X *and* Y *be random variables such that* $\mathbb{E}|X| < \infty$*. The conditional expectation of* X *given* Y *is defined by* $\mathbb{E}(X|Y) := \mathbb{E}(X|\sigma(Y))$*.*

Remark 5.7. Using an approximation argument (see e.g. Prop. 1.11 in [Bal17]) one shows that the condition (5.1) is equivalent to $\mathbb{E}(XW) = \mathbb{E}(ZW)$ for all bounded $\mathcal{G}$-measurable random variables W. If $\mathcal{G} = \sigma(Y)$, then by Remark 5.3 the above condition is equivalent to $\mathbb{E}(Xg(Y)) = \mathbb{E}(Zg(Y))$ for all bounded Borel functions g.

Example 5.8. Let X and Y be random variables with the joint density $f_{X,Y} \in L^1(\mathbb{R} \times \mathbb{R})$. If X is integrable, then we claim that

$$\mathbb{E}(X|Y) \equiv \mathbb{E}(X|\sigma(Y)) = \int x f_X(x|Y)\, \mathrm{d}x.$$

The conditional density $f_X(x|Y)$ of X given Y is defined by

$$f_X(x|y) = \begin{cases} \frac{f_{X,Y}(x,y)}{f_Y(y)} & \text{if } f_Y(y) \neq 0, \\ 0 & \text{if } f_Y(y) = 0, \end{cases}$$

where $f_Y(y) = \int f_{X,Y}(x,y)\,\mathrm{d}x$ is the density of Y. Let us verify the above claim about $\mathbb{E}(X|Y)$. Since $y \mapsto \int x f_X(x|y)\,\mathrm{d}x$ is a Borel function, the random variable $\int x f_X(x|Y)\,\mathrm{d}x$ is $\sigma(Y)$-measurable. Moreover, for all bounded Borel functions g we have

$$\begin{aligned}\mathbb{E}(g(Y)\,\mathbb{E}(X|Y)) &= \mathbb{E}\left(g(Y)\int x f_X(x|Y)\,\mathrm{d}x\right)\\ &= \int g(y)\left(\int x f_X(x|y)\,\mathrm{d}x\right) f_Y(y)\,\mathrm{d}y\\ &= \int g(y)\,x\,f_{X,Y}(x,y)\,\mathrm{d}x\mathrm{d}y\\ &= \mathbb{E}(g(Y)X).\end{aligned}$$

Lemma 5.9. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, $\mathcal{G}$, $\mathcal{H}$ be a sub-σ-algebra of $\mathcal{F}$ and X, Y, Z, $X_1, X_2, \ldots$ be integrable random variables.*

(a) Linearity: If $\alpha, \beta \in \mathbb{R}$, $\mathbb{E}(\alpha X + \beta Y|\mathcal{G}) = \alpha\mathbb{E}(X|\mathcal{G}) + \beta\mathbb{E}(Y|\mathcal{G})$ a.s.

(b) Monotonicity: If $X \leqslant Y$ a.s., $\mathbb{E}(X|\mathcal{G}) \leqslant \mathbb{E}(Y|\mathcal{G})$ a.s.

(c) Monotone convergence: If $X_n \geqslant 0$ and $X_n \nearrow X$ a.s., then $\mathbb{E}(X_n|\mathcal{G}) \nearrow \mathbb{E}(X|\mathcal{G})$ a.s.

(d) Jensen's inequality: If $\phi: \mathbb{R} \to \mathbb{R}$ is convex and $\mathbb{E}|\phi(X)| < \infty$, $\phi(\mathbb{E}(X|\mathcal{G})) \leqslant \mathbb{E}(\phi(X)|\mathcal{G})$ a.s.

(e) Expectation: $\mathbb{E}(\mathbb{E}(X|\mathcal{G})) = \mathbb{E}X$.

(f) Iteration: If $\mathcal{G} \subset \mathcal{H}$, $\mathbb{E}(\mathbb{E}(X|\mathcal{H})|\mathcal{G}) = \mathbb{E}(X|\mathcal{G})$ a.s.

(g) If Z is $\mathcal{G}$-measurable and $\mathbb{E}(|XZ|) < \infty$, $\mathbb{E}(Z|\mathcal{G}) = Z$ and $\mathbb{E}(XZ|\mathcal{G}) = Z\mathbb{E}(X|\mathcal{G})$ a.s.

(h) If X is independent of $\mathcal{G}$, then $\mathbb{E}(X|\mathcal{G}) = \mathbb{E}X$ a.s.

Proof. See Exercises 2 and 3, Sheet 2. □

Remark 5.10. A random variable X is independent of a σ-algebra $\mathcal{G}$ if $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$ for all $A \in \sigma(X)$ and $B \in \mathcal{G}$.

6 Martingales

Suppose $Y_1, Y_2, \ldots$ are independent random variables with mean zero and define the stochastic process $(S_n)_{n\in\mathbb{N}_+}$ by $S_n := Y_1 + \cdots + Y_n$. We have

$$\begin{aligned}\mathbb{E}(S_{n+k}|S_1, \ldots, S_n) &= \mathbb{E}(S_n|S_1, \ldots, S_n) + \mathbb{E}(Y_{n+1} + \cdots + Y_{n+k}|S_1, \ldots, S_n)\\ &= S_n + \mathbb{E}(Y_{n+1} + \cdots + Y_{n+k}) = S_n.\end{aligned}$$

Thus, the best estimate of the future value of the stochastic process $(S_n)_{n\in\mathbb{N}_+}$ given the history up to time n, is just S_n. The process $(S_n)_{n\in\mathbb{N}_+}$ is an example of a martingale. Martingales are stochastic processes that are meant to capture the notion of a fair game in the context of gambling. If we interpret Y_i as the payoff of a game at time i and S_n as the total winnings at time n, the condition $\mathbb{E}(S_{n+k}|S_1, \ldots, S_n) = S_n$ says that at any time the future expected winnings, given the winnings to date, is just the current amount of money.

Definition 6.1. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A* ***filtration*** *is a family $(\mathcal{F}_t)_{t \geqslant 0}$ of sub-σ-algebras of $\mathcal{F}$ such that if $s \leqslant t$, then $\mathcal{F}_s \subseteq \mathcal{F}_t$. A stochastic process $(X_t)_{t \geqslant 0}$ is* ***adapted*** *to $(\mathcal{F}_t)_{t \geqslant 0}$ if X_t is $\mathcal{F}_t$-measurable for every $t \geqslant 0$.*

Remark 6.2. You should think of $\mathcal{F}_t$ as the σ-algebra of the events for which at time t we can say whether they are satisfied or not.

Example 6.3. Given a process $(X_t)_{t \geqslant 0}$ define $\mathcal{F}_t^X$ to be the smallest sub-σ-algebra of $\mathcal{F}$ containing the σ-algebras generated by X_s for $s \in [0, t]$. We call $(\mathcal{F}_t^X)_{t \geqslant 0}$ the **natural filtration** of $(X_t)_{t \geqslant 0}$. The σ-algebra $\mathcal{F}_t^X$ contains all events that can be expressed in terms of $(X_s)_{s \in [0,t]}$. Every stochastic process is adapted to its natural filtration.

Definition 6.4. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $(\mathcal{F}_t)_{t \geqslant 0}$ a filtration. A stochastic process $(M_t)_{t \geqslant 0}$ is a* ***martingale*** *(resp. a* ***supermartingale****, a* ***submartingale****) if:*

(a) $\mathbb{E}|M_t| < \infty$ *for all* $t \geqslant 0$.

(b) $\mathbb{E}(M_t | \mathcal{F}_s) = M_s$ *a.s. (reps.* $\leqslant$, $\geqslant$*) if* $s < t$.

(c) M_t *is* $\mathcal{F}_t$*-measurable for* $t \geqslant 0$.

Example 6.5. A Brownian motion $(B_t)_{t \geqslant 0}$ is a martingale with respect to its natural filtration $(\mathcal{F}_t^B)_{t \geqslant 0}$. Indeed, by Cauchy-Schwarz inequality $\mathbb{E}|B_t| \leqslant (\mathbb{E}(B_t^2))^{1/2} = t^{1/2} < \infty$. Moreover, if $t \geqslant s$, then

$$\mathbb{E}(B_t | \mathcal{F}_s^B) = \mathbb{E}(B_t - B_s + B_s | \mathcal{F}_s^B) = \mathbb{E}(B_t - B_s | \mathcal{F}_s^B) + \mathbb{E}(B_s | \mathcal{F}_s^B) = \mathbb{E}(B_t - B_s) + B_s = B_s.$$

The second equality follows from the property (a) and the third from the properties (g) and (h) of the conditional expectation stated in Lemma 5.9.

Example 6.6. Let X be an integrable random variable and $(\mathcal{F}_t)_{t \geqslant 0}$ be a filtration. For $t \geqslant 0$ define M_t to be (any representative of the equivalence class) $\mathbb{E}(X | \mathcal{F}_t)$. Then $(M_t)_{t \geqslant 0}$ is a martingale with respect to the filtration $(\mathcal{F}_t)_{t \geqslant 0}$ called the **Doob martingale**. See Exercise 3, Sheet 3 for a proof.

Proposition 6.7. *If $(M_t)_{t \geqslant 0}$ is a martingale and $\varphi: \mathbb{R} \to \mathbb{R}$ is convex and satisfies $\mathbb{E}|\varphi(M_t)| < \infty$ for $t \geqslant 0$, then $(\varphi(M_t))_{t \geqslant 0}$ is a submartingale. In particular, $(|M_t|)_{t \geqslant 0}$ is a submartingale and $(M_t^2)_{t \geqslant 0}$ is a submartingale if $\mathbb{E}M_t^2 < \infty$ for all $t \geqslant 0$.*

Proof. The conditions (a) and (c) hold true. By property (d) of Lemma 5.9 we have for $s \leqslant t$

$$\mathbb{E}(\varphi(M_t) | \mathcal{F}_s) \geqslant \varphi(\mathbb{E}(M_t | \mathcal{F}_s)) = \varphi(M_s). \qquad \square$$

Definition 6.8. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $(\mathcal{F}_t)_{t \geqslant 0}$ a filtration. A* ***stopping time*** *is a random variable $\tau: \Omega \to [0, \infty]$ such that $\{\tau \leqslant t\} \in \mathcal{F}_t$ for all $t \geqslant 0$. Given a stopping time τ we define the* ***σ-algebra of events prior to τ*** *by*

$$\mathcal{F}_\tau = \{A \in \mathcal{F} \mid A \cap \{\tau \leqslant t\} \in \mathcal{F}_t \text{ for all } t \geqslant 0\}.$$

Remark 6.9. A stopping time can take infinite values. Intuitively, the condition $\{\tau \leqslant t\} \in \mathcal{F}_t$ means that at time t we should be able to say whether $\tau \leqslant t$ or not.

Remark 6.10. If $(X_t)_{t \geqslant 0}$ is a continuous adapted stochastic process and A is an open set, then the exit time from A defined by

$$\tau_A := \inf \{t \geqslant 0 \,|\, X_t \notin A\} \in [0, \infty]$$

is a stopping time. For a proof, see Prop. 3.7 in [Bal17].

Proposition 6.11. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $(\mathcal{F}_t)_{t \geqslant 0}$ a filtration.*

(i). If τ_1, τ_2 are stopping times, then $\tau_1 \wedge \tau_2$ and $\tau_1 \vee \tau_2$ are also stopping times.

(ii). If $(X_t)_{t \geqslant 0}$ is adapted and continuous and τ is an a.s. finite stopping time, then

$$X_\tau : \omega \mapsto X_{\tau(\omega)}(\omega) \mathbb{1}_{\tau < \infty}(\omega)$$

is an $\mathcal{F}_\tau$-measurable random variable.

Proof. (i) See Prop. 3.5 in [Bal17]. (ii) See Prop. 3.6 and Prop. 2.1 in [Bal17]. □

Theorem 6.12. (Optional Stopping Theorem) *Let $(M_t)_{t \geqslant 0}$ be a continuous martingale (resp. supermartingale, submartingale) and let τ_1, τ_2 be two stopping times such that $\tau_1 \leqslant \tau_2$ and τ_2 is bounded a.s. Then*

$$\mathbb{E}(M_{\tau_2} | \mathcal{F}_{\tau_1}) = M_{\tau_1} \quad (\textit{resp. } \leqslant, \geqslant).$$

Proof. See e.g. Theorem 5.13 and Theorem 5.2 in [Bal17]. □

Remark 6.13. Using the above theorem it is possible to prove that if $(M_t)_{t \geqslant 0}$ is a continuous martingale and τ is a stopping time, then $M^\tau := (M_{t \wedge \tau})_{t \geqslant 0}$ is a martingale. See Prop. 5.6 in [Bal17].

Theorem 6.14. (Doob's Inequalities) *Let $(M_t)_{t \geqslant 0}$ be a continuous martingale. Define*

$$M_t^* := \sup_{s \in [0,t]} |M_s|.$$

Then:

(a) $\lambda \mathbb{P}(M_t^* \geqslant \lambda) \leqslant \mathbb{E}|M_t|$ for all $t \geqslant 0$ and $\lambda > 0$.

(b) For all $t \geqslant 0$, if $\mathbb{E}(M_t^2) < \infty$, then $\mathbb{E}((M_t^*)^2) \leqslant 4 \mathbb{E}(M_t^2)$.

Proof. (a) Let $\tau_1 = \inf \{s \geqslant 0 \,|\, |M_s| \notin (-\infty, \lambda)\} \wedge t$ and $\tau_2 = t$. Then τ_1 and τ_2 are stopping times such that $\tau_1 \leqslant \tau_2 \leqslant t < \infty$. Applying the optional stopping theorem to the submartingale $(|M_s|)_{s \geqslant 0}$ we obtain

$$|M_{\tau_1}| \leqslant \mathbb{E}(|M_{\tau_2}| \,|\mathcal{F}_{\tau_1}) = \mathbb{E}(|M_t| \,|\mathcal{F}_{\tau_1}).$$

Thus,

$$|M_{\tau_1}|\mathbb{1}_{\{|M_{\tau_1}|\geqslant\lambda\}}\leqslant\mathbb{E}(|M_t|\,|\mathcal{F}_{\tau_1})\mathbb{1}_{\{|M_{\tau_1}|\geqslant\lambda\}}.$$

Since $\mathbb{1}_{\{|M_{\tau_1}|\geqslant\lambda\}}$ is $\mathcal{F}_{\tau_1}$-measurable, by Lemma 5.9 (g) we obtain

$$|M_{\tau_1}|\mathbb{1}_{\{|M_{\tau_1}|\geqslant\lambda\}}\leqslant\mathbb{E}(|M_t|\,\mathbb{1}_{\{|M_{\tau_1}|\geqslant\lambda\}}|\mathcal{F}_{\tau_1}).$$

Consequently,

$$\mathbb{E}(|M_{\tau_1}|\mathbb{1}_{\{|M_{\tau_1}|\geqslant\lambda\}})\leqslant\mathbb{E}(|M_t|\mathbb{1}_{\{|M_{\tau_1}|\geqslant\lambda\}}). \tag{6.1}$$

Note that

$$\{|M_{\tau_1}|\geqslant\lambda\}=\{|M_s|\notin(-\infty,\lambda)\,\text{for some}\,s\in[0,t]\}=\{M_t^*\geqslant\lambda\}.$$

Hence,

$$\begin{aligned}\lambda\mathbb{P}(M_t^*\geqslant\lambda)\ &=\lambda\mathbb{P}(|M_{\tau_1}|\geqslant\lambda)=\mathbb{E}(\lambda\mathbb{1}_{\{|M_{\tau_1}|\geqslant\lambda\}})\leqslant\mathbb{E}(|M_{\tau_1}|\mathbb{1}_{\{|M_{\tau_1}|\geqslant\lambda\}})\\ &\leqslant\mathbb{E}(|M_t|\mathbb{1}_{\{|M_{\tau_1}|\geqslant\lambda\}})=\mathbb{E}(|M_t|\mathbb{1}_{\{M_t^*\geqslant\lambda\}}),\end{aligned}$$

where the second bound follows from (6.1). We conclude that

$$\lambda\mathbb{P}(M_t^*\geqslant\lambda)\leqslant\mathbb{E}(|M_t|\mathbb{1}_{\{M_t^*\geqslant\lambda\}}), \tag{6.2}$$

which implies the claim (a).

(b) Recall that if $X\geqslant 0$ is a random variable and a deterministic constant $T>0$, then for $p>0$, we have

$$\mathbb{E}(X\wedge T)^p=\mathbb{E}\int_0^T p\lambda^{p-1}\mathbb{1}_{\{\lambda\leqslant X\}}\mathrm{d}\lambda=\int_0^T p\lambda^{p-1}\mathbb{P}(X\geqslant\lambda)\mathrm{d}\lambda.$$

Using this fact with $p=2$, the estimate (6.2), the Fubini theorem and the Cauchy–Schwartz inequality we obtain

$$\begin{aligned}\mathbb{E}((M_t^*\wedge T)^2)\ &=\int_0^T 2\lambda\mathbb{P}(M_t^*\geqslant\lambda)\mathrm{d}\lambda\\ &\leqslant\int_0^T 2\mathbb{E}(|M_t|\mathbb{1}_{\{M_t^*\geqslant\lambda\}})\,\mathrm{d}\lambda\\ &=2\mathbb{E}\bigg(|M_t|\int_0^T\mathbb{1}_{\{M_t^*\geqslant\lambda\}}\mathrm{d}\lambda\bigg)\\ &=2\mathbb{E}(|M_t|\,(M_t^*\wedge T))\\ &\leqslant 2\sqrt{\mathbb{E}(M_t^2)}\sqrt{\mathbb{E}((M_t^*\wedge T)^2)}.\end{aligned}$$

If $M_t^*=0$, then the statement is clearly true. Otherwise, $\mathbb{E}((M_t^*\wedge T)^2)\in(0,T^2]$ and

$$\sqrt{\mathbb{E}((M_t^*\wedge T)^2)}\leqslant 2\sqrt{\mathbb{E}(M_t^2)}.$$

The claim (b) follows by taking the limit $T\to\infty$ of both sides of the the above bound and invoking the monotone convergence theorem. □

7 Integration with respect to bounded variation processes

The fundamental problem we will address in the upcoming lectures is to rigorously define the integrals

$$\int_0^t H_s \mathrm{d}X_s,$$

where $(X_t)_{t\geqslant 0}$ and $(H_t)_{t\geqslant 0}$ are processes enjoying certain properties to be specified. The simplest approach would be to define the integral separately for each path, that is, to study

$$\int_0^t H_s(\omega)\mathrm{d}X_s(\omega) \tag{7.1}$$

for all sample points $\omega \in \Omega$. Such a construction is provided by the Stieltjes integral. As we shall see, this construction does not work if $(X_t)_{t\geqslant 0}$ is a martingale.

Our goal is to find a natural sufficient condition for sample paths of $(X_t)_{t\geqslant 0}$ that allows to construct the integral (7.1). The remark below suggests a possible but unnecessarily restrictive sufficient condition.

Remark 7.1. Recall that if μ is a finite positive measure on $(0,T]$, then $t\mapsto\mu((0,t])$ is a right-continuous non-decreasing function vanishing at zero. Conversely, given right-continuous non-decreasing function g on $[0,T]$ there is a unique associated finite positive measure μ_g on $(0,T]$ such that $\mu((0,t]) = g(t) - g(0)$.

Thus, if $s\mapsto X_s(\omega)$ is a right-continuous non-decreasing function, then (7.1) can be defined as the integral of $H(\omega)$ with respect to the positive measure $\mu_{X(\omega)}$. Let us try to extend this definition to the situation where $X_s(\omega)$ is not monotonic.

Definition 7.2. *A* ***signed measure*** *μ on $(0,T]$ is the difference of two finite positive measures on $(0,T]$.*

It turns out that if μ is a sign measure on $(0,T]$, then $t\mapsto\mu((0,t])$ is a right-continuous function of bounded variation vanishing at zero.

Definition 7.3. *The* ***variation*** *of a function $g\colon[0,T]\to\mathbb{R}$ on an interval $[0,t]$ is defined by*

$$\mathcal{V}_t(g) := \sup\left\{ \sum_{i=1}^n |g(t_i) - g(t_{i-1})| \;\middle|\; 0 = t_0 \leqslant t_1 \leqslant \cdots \leqslant t_{n-1} \leqslant t_n = t, n\in\mathbb{N}_+ \right\}.$$

We say that g is of ***bounded variation*** *if its* ***total variation*** *$\mathcal{V}_T(g)$ is finite.*

Remark 7.4. The variation $\mathcal{V}_t(g)$ is a non-decreasing function of t. If g is non-decreasing, then $\mathcal{V}_t(g) = g(t)$. If g is Lipschitz continuous with a Lipschitz constant L, then $\mathcal{V}_t(g) \leqslant Lt$.

Theorem 7.5. *Let g be a right-continuous function of bounded variation. There exist unique non-decreasing right-continuous functions g_+, g_- such that*

$$g = g_+ - g_- \qquad \text{and} \qquad \mathcal{V}(g) = g_+ + g_-.$$

Idea of proof. Define $g_+(t) = \frac{1}{2}(\mathcal{V}_t(g) + g(t))$ and $g_-(t) = \frac{1}{2}(\mathcal{V}_t(g) - g(t))$. □

Theorem 7.6. *For every right-continuous function g of bounded variation on $[0,T]$ there is a unique associated sign measure μ_g on $(0,T]$ such that $\mu_g((0,t]) = g(t) - g(0)$.*

Idea of proof. Let g_+, g_- be as in the previous theorem and let μ_{g_+}, μ_{g_-} be the finite positive measures associated to g_+, g_-. Define $\mu_g = \mu_{g_+} - \mu_{g_-}$. Note that

$$\mu((0,t]) = \mu_{g_+}((0,t]) - \mu_{g_-}((0,t]) = (g_+(t) - g_+(0)) - (g_-(t) - g_-(0)) = g(t) - g(0).$$ □

Example 7.7. If $g \in C^1$, then $\mu(\mathrm{d}t) = g'(t)\,\mathrm{d}t$, where $\mathrm{d}t$ is the Lebesgue measure. If $g = \mathbf{1}_{[a,\infty)}$, then $\mu = \delta_a$ is the Dirac delta at a. If $g = \mathbf{1}_{[a,b)} = \mathbf{1}_{[a,\infty)} - \mathbf{1}_{[b,\infty)}$, then $\mu = \delta_a - \delta_b$.

Definition 7.8. *Let $g\colon [0,T] \to \mathbb{R}$ be right-continuous and of bounded variation with associated signed measure μ_g. The positive measure $|\mu_g| := \mu_{g_+} + \mu_{g_-}$ on $(0,T]$ associated to the non-decreasing right-continuous function $t \mapsto \mathcal{V}_t(g)$ is called the **variation** of μ_g. For $f \in L^1([0,T], |\mu_g|)$ and $t \in [0,T]$ we define*

$$(f \cdot g)_t \equiv \int_0^t f(s)\,\mathrm{d}g(s) := \int \mathbf{1}_{(0,t]}(s) f(s) \mu_g(\mathrm{d}s),$$

where the integral with respect to a sign measure is defined by

$$\int f(s)\mu_g(\mathrm{d}s) := \int f(s)\mu_{g_+}(\mathrm{d}s) - \int f(s)\mu_{g_-}(\mathrm{d}s).$$

*We call $(f \cdot g)_t$ as above the **Lebesgue–Stieltjes integral** of f with respect to g.*

Remark 7.9. One shows that $t \mapsto (f \cdot g)_t$ is of bounded variation for f, g as in the definition above.

Proposition 7.10. *Let $g\colon [0,T] \to \mathbb{R}$ be right-continuous and of bounded variation and $f\colon [0,T] \to \mathbb{R}$ be continuous. Then the Lebesgue-Stieltjes integral coincides with the **Riemann–Stieltjes integral**, that is,*

$$(f \cdot g)_t = \lim_{n \to \infty} \sum_{i=1}^{n} f(s_i^{(n)})(g(t_i^{(n)}) - g(t_{i-1}^{(n)})),$$

where $0 = t_0^{(n)} < t_1^{(n)} < \ldots < t_n^{(n)} = t$ and $s_i^{(n)} \in [t_{i-1}^{(n)}, t_i^{(n)}]$ are arbitrary such that

$$\lim_{n \to \infty} \max_{i \in \{1,\ldots,n\}} (t_i^{(n)} - t_{i-1}^{(n)}) = 0.$$

Definition 7.11. *We say that stochastic processes $(X)_{t \geqslant 0}$ and $(Y)_{t \geqslant 0}$ are **indistinguishable** if their sample paths coincide a.s., that is, $\mathbb{P}(\forall_{t \geqslant 0} X_t = Y_t) = 1$.*

Definition 7.12. (Integral with respect to a process of bounded variation) *Let $X = (X_t)_{t\in[0,T]}$ be a.s. of bounded variation and $H = (H_t)_{t\in[0,T]} \in L^1([0,T], |\mu_X|)$ a.s. The integral of H with respect to X is the equivalence class of indistinguishable stochastic processes*

$$(H\cdot X)_{t\in[0,T]} \equiv \left(\int_0^t H_s\,\mathrm{d}X_s\right)_{t\in[0,T]}$$

of bounded variation such that

$$(H\cdot X)_t(\omega) := (H(\omega)\cdot X(\omega))_t$$

for all $t\in[0,T]$ and all $\omega\in\Omega$ for which $(H(\omega)\cdot X(\omega))_t$ is well-defined as the Lebesgue–Stieltjes integral.

Imagine that H_s is the quantity of an asset held by an investor at time s and X_s is the price of the asset at time s. Then the integral $\int_0^t H_s \mathrm{d}X_s$ represents the gain realized in the time interval $[0,t]$. The following proposition shows that the construction of $\int_0^t H_s \mathrm{d}X_s$ presented above cannot be apply in the situation when the price of the asset is modeled by a Brownian motion or, more generally, a martingale.

Proposition 7.13. *A continuous martingale $(M_t)_{t\geqslant 0}$ is of bounded variation iff it is a.s. constant.*

Remark 7.14. The above proposition implies that Brownian motion is a.s. not of finite variation. In particular, it is a.s. not differentiable.

Proof. We may suppose that $M_0 = 0$ and prove that $M = (M_t)_{t\geqslant 0}$ is identically zero if it is of bounded variation. Let $\mathcal{V}_t(M)$ be the variation of M on $[0,t]$. For $K > 0$ define

$$\tau_K(\omega) := \inf\{s\geqslant 0 \,|\, \mathcal{V}_s(M(\omega)) \geqslant K\}, \qquad \omega\in\Omega.$$

By Remark 6.10 the random variable τ_K is a stopping time and by Remark 6.13 $(M_t^{\tau_K})_{t\geqslant 0} := (M_{t\wedge\tau_K})_{t\geqslant 0}$ is a martingale. By the above definitions $(\tilde{M}_t)_{t\geqslant 0}$ has the variation bounded by K. In particular, we have $|M_t^{\tau_K}| \leqslant K$ and $\mathbb{E}((M_t^{\tau_K})^2) \leqslant K^2$. Moreover, since $t\mapsto M_t(\omega)$ is of bounded variation for every $s\geqslant 0$ there is $K>0$ such that $\mathcal{V}_s(M(\omega)) \leqslant K$ and $\tau_K(\omega)\geqslant s$. Hence, $\lim_{K\to\infty}\tau_k = \infty$ a.s.

For $0 = t_0 < t_1 < \ldots < t_k = t$ we obtain

$$\mathbb{E}((M_t^{\tau_K})^2) = \mathbb{E}\left(\sum_{i=1}^k \left((M_{t_i}^{\tau_K})^2 - (M_{t_{i-1}}^{\tau_K})^2\right)\right) = \mathbb{E}\left(\sum_{i=1}^k (M_{t_i}^{\tau_K} - M_{t_{i-1}}^{\tau_K})^2\right).$$

The last equality follows from $\mathbb{E}(M_{t_i}^{\tau_K} M_{t_{i-1}}^{\tau_K}) = \mathbb{E}(\mathbb{E}(M_{t_i}^{\tau_K}|\mathcal{F}_{t_i-1})M_{t_{i-1}}^{\tau_K}) = \mathbb{E}({M^{\tau_K}}_{t_{i-1}}^2)$ since M^{τ_K} is a martingale. As a result,

$$\mathbb{E}((M_t^{\tau_K})^2) \leqslant E\Big[\mathcal{V}_t(M^{\tau_K}) \max_i |M_{t_i}^{\tau_K} - M_{t_{i-1}}^{\tau_K}|\Big] \leqslant K\mathbb{E}\Big[\max_i |M_{t_i}^{\tau_K} - M_{t_{i-1}}^{\tau_K}|\Big].$$

When $\max_i |t_i - t_{i-1}|$ goes to zero, $\max_i |M_{t_i}^{\tau_K} - M_{t_{i-1}}^{\tau_K}|$ goes to zero since M^{τ_K} is continuous (and hence uniformly continuous) on $[0,T]$. Thus, by by the dominated convergence theorem and the bound $\max_i |M_{t_i}^{\tau_K} - M_{t_{i-1}}^{\tau_K}| \leqslant K$ we infer that $\mathbb{E}((M_t^{\tau_K})^2) = 0$. This shows that $M_t^{\tau_K} = M_{t\wedge\tau_K} = 0$ a.s. for all $t\geqslant 0$.

Since $\lim_{K\to\infty}\tau_k = \infty$ a.s., we have $M_t = \lim_{K\to\infty} M_{t\wedge\tau_K} = 0$ a.s. for all $t \geqslant 0$. Thus, for every $t \geqslant 0$ there is an event $E_t \subset \Omega$ such that $\mathbb{P}(E_t) = 1$ and $M_t(\omega) = 0$ for $\omega \in E_t$. Let $E = \bigcap_{t\in\mathbb{Q}\cap[0,\infty)} E_t$, where $\mathbb{Q}$ is the set of rational numbers. Then E is an event such that $\mathbb{P}(E) = 1$ and $M_t(\omega) = 0$ for $\omega \in E$ and $t \in \mathbb{Q}\cap[0,\infty)$. The statement follows now from the assumed a.s. continuity of M. $\square$

8 Stochastic integral

Our goal is to define an integral

$$\int_0^t H_s\,\mathrm{d}B_s \tag{8.1}$$

of a sufficiently generic stochastic process $H = (H_t)_{t\geqslant 0}$ with respect to a Brownian motion $B = (B_t)_{t\geqslant 0}$. Since Brownian motion is a non-zero martingale, its variation is a.s. unbounded and the construction of an integral presented in the previous section does not apply.

Throughout this section we assume that $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space, $(\mathcal{F}_t)_{t\geqslant 0}$ is a filtration such that $\mathcal{F}_0$ contains all the events of zero probability and $(B_t)_{t\geqslant 0}$ is a continuous Brownian motion adapted to $(\mathcal{F}_t)_{t\geqslant 0}$ and such that $(B_{s+t} - B_t)_{s\geqslant 0}$ is independent of $\mathcal{F}_t$ for all $t \geqslant 0$. For simplicity, we fix a finite time horizon $T > 0$ and construct the stochastic integral (8.1) for $t \in [0, T]$.

Remark 8.1. The assumptions that $\mathcal{F}_0$ contains all the events of zero probability is of technical nature. Note that, for example, it guarantees that an a.s. limit of an adapted process is adapted.

Remark 8.2. Let $(\mathcal{F}_t^B)_{t\geqslant 0}$ be the natural filtration of a Brownian motion $(B_t)_{t\geqslant 0}$ and $\mathcal{N} = \{A \in \mathcal{F} \,|\, \mathbb{P}(A) = 0\}$. We can define $\mathcal{F}_t$ to be the σ-algebra generated by $\mathcal{F}_t^B$ and $\mathcal{N}$, that is, the smallest sub-σ-algebra of $\mathcal{F}$ containing all the events from $\mathcal{F}_t^B$ and all the events of zero probability. One checks that $B_{s+t} - B_t$ is independent of $\mathcal{F}_t$ for all $s, t \geqslant 0$.

8.1 Integral for simple predictable processes

Definition 8.3. *We say that $H = (H_t)_{t\in[0,T]}$ is a **simple predictable process** if*

$$H_t = \sum_{i=1}^n X_i\,\mathbf{1}_{(t_{i-1},t_i]}(t) \tag{8.2}$$

for some $n \in \mathbb{N}_+$, $0 = t_0 < t_1 < \ldots < t_n = T$ and random variables $X_1, \ldots, X_n$ such that $\mathbb{E}X_i^2 < \infty$ and X_i is a $\mathcal{F}_{t_{i-1}}$-measurable for all $i \in \{1, \ldots, n\}$. Let $\mathcal{E}_T$ denote the vector space of simple predictable processes.

Remark 8.4. Every $H \in \mathcal{E}_T$ is adapted and left-continuous.

Definition 8.5. *The integral of $H \in \mathcal{E}_T$ of the form (8.2) with respect to the Brownian motion B is defined by*

$$(H\cdot B)_t \equiv \int_0^t H_s\,\mathrm{d}B_s := \sum_{i=1}^n X_i(B_{t_i\wedge t} - B_{t_{i-1}\wedge t}), \qquad t \in [0, T].$$

The value of $(H\cdot B)_t$ does not depend on the representation of H as an element of $\mathcal{E}_T$

Remark 8.6. If $t \in (t_{k-1}, t_k]$, then

$$(H\cdot B)_t = \sum_{i=1}^{k-1} X_i(B_{t_i} - B_{t_{i-1}}) + X_k(B_t - B_{t_{k-1}}).$$

Proposition 8.7. *Let $H, K \in \mathcal{E}_T$ and $a, b \in \mathbb{R}$. Then:*

(i). $((aH + bK) \cdot B)_t = a\,(H \cdot B)_t + b\,(K \cdot B)_t$.

(ii). $\mathbb{E}(H \cdot B)_t = 0$.

(iii). $\mathbb{E}((H \cdot B)_t^2) = \mathbb{E}(\int_0^t H_s^2\,\mathrm{d}s)$ *(Itô isometry).*

(iv). $((H \cdot B)_t)_{t \in [0,T]}$ *is a continuous martingale.*

(v). $(H \cdot B)_t = (H \mathbf{1}_{[0,t]} \cdot B)_T$.

Proof. To verify (i) we write H and K using the same partition. Let H be of the form (8.2) and set $\bar{t}_i = t_i \wedge t$. To prove (ii) note that

$$\begin{aligned}\mathbb{E}(H \cdot B)_t &= \sum_{i=1}^n \mathbb{E}(X_i(B_{\bar{t}_i} - B_{\bar{t}_{i-1}})) = \sum_{i=1}^n \mathbb{E}(\mathbb{E}(X_i(B_{\bar{t}_i} - B_{\bar{t}_{i-1}})|\mathcal{F}_{t_{i-1}})) \\ &= \sum_{i=1}^n \mathbb{E}(X_i \mathbb{E}(B_{\bar{t}_i} - B_{\bar{t}_{i-1}}|\mathcal{F}_{t_{i-1}})) = 0\end{aligned}$$

since

$$\mathbb{E}(B_{\bar{t}_i} - B_{\bar{t}_{i-1}}|\mathcal{F}_{t_{i-1}}) = \mathbb{E}(B_{\bar{t}_i} - B_{\bar{t}_{i-1}}) = \mathbb{E}(B_{t_i \wedge t} - B_{t_{i-1} \wedge t}) = 0.$$

Property (iii) follows from

$$\begin{aligned}\mathbb{E}((H \cdot B)_t)^2 = & \sum_{i=1}^n \sum_{j=1}^n \mathbb{E}(X_i X_j (B_{\bar{t}_i} - B_{\bar{t}_{i-1}})(B_{\bar{t}_j} - B_{\bar{t}_{j-1}})) \\ &= \sum_{i=1}^n \mathbb{E}(X_i^2 (B_{\bar{t}_i} - B_{\bar{t}_{i-1}})^2) + 2 \sum_{i=1}^{n-1} \sum_{j=i+1}^n \mathbb{E}(X_i X_j (B_{\bar{t}_i} - B_{\bar{t}_{i-1}})(B_{\bar{t}_j} - B_{\bar{t}_{j-1}})).\end{aligned}$$

For $i < j$ we have

$$\begin{aligned}\mathbb{E}(X_i X_j (B_{\bar{t}_i} - B_{\bar{t}_{i-1}})(B_{\bar{t}_j} - B_{\bar{t}_{j-1}})) &= \mathbb{E}(\mathbb{E}(X_i X_j (B_{\bar{t}_i} - B_{\bar{t}_{i-1}})(B_{\bar{t}_j} - B_{\bar{t}_{j-1}})|\mathcal{F}_{t_{j-1}})) \\ &= \mathbb{E}(X_i X_j (B_{\bar{t}_i} - B_{\bar{t}_{i-1}}))\,\mathbb{E}(B_{\bar{t}_j} - B_{\bar{t}_{j-1}}|\mathcal{F}_{t_{j-1}}) = 0.\end{aligned}$$

We also have

$$\mathbb{E}(X_i^2 (B_{\bar{t}_i} - B_{\bar{t}_{i-1}})^2) = \mathbb{E}(\mathbb{E}(X_i^2 (B_{\bar{t}_i} - B_{\bar{t}_{i-1}})^2|\mathcal{F}_{t_{i-1}})) = \mathbb{E}(X_i^2)\,\mathbb{E}((B_{\bar{t}_i} - B_{\bar{t}_{i-1}})^2|\mathcal{F}_{t_{i-1}}).$$

Since

$$\mathbb{E}((B_{\bar{t}_i} - B_{\bar{t}_{i-1}})^2|\mathcal{F}_{t_{i-1}}) = \mathbb{E}((B_{\bar{t}_i} - B_{\bar{t}_{i-1}})^2) = \bar{t}_i - \bar{t}_{i-1}, \tag{8.3}$$

we obtain

$$\mathbb{E}(X_i^2 (B_{\bar{t}_i} - B_{\bar{t}_{i-1}})^2) = (\mathbb{E}X_i^2)(\bar{t}_i - \bar{t}_{i-1}).$$

As a result,

$$\mathbb{E}((H\cdot B)_t)^2 = \sum_{i=1}^n (\mathbb{E}X_i^2)(\bar{t}_i - \bar{t}_{i-1}) = \sum_{i=1}^n \mathbb{E}\int_{\bar{t}_{i-1}}^{\bar{t}_i} X_i^2\,\mathrm{d}s = \sum_{i=1}^n \mathbb{E}\int_{\bar{t}_{i-1}}^{\bar{t}_i} H_s^2\,\mathrm{d}s = \mathbb{E}\int_0^t H_s^2\,\mathrm{d}s.$$

Let us turn to the proof of (iv). Continuity of $((H\cdot B)_t)_{t\in[0,T]}$ follows immediately from the definition and continuity of the Brownian motion B. The process $((H\cdot B)_t)_{t\in[0,T]}$ is clearly integrable since by (iii) it is square-integrable. It remains to check that $\mathbb{E}((H\cdot B)_t|\mathcal{F}_s) = (H\cdot B)_s$ if $s<t$. By Lemma 5.9 (f) it suffices to demonstrate this for $t_{k-1}\leqslant s<t\leqslant t_k$. For such s,t we have

$$\mathbb{E}((H\cdot B)_t - (H\cdot B)_s|\mathcal{F}_s) = \mathbb{E}(X_k(B_t - B_s)|\mathcal{F}_s) = X_k\mathbb{E}((B_t - B_s)|\mathcal{F}_s) = X_k\mathbb{E}(B_t - B_s) = 0$$

since X_k is $\mathcal{F}_{t_{k-1}}$-measurable and $\mathcal{F}_{t_{k-1}}\subset\mathcal{F}_s$.

To prove (v) we observe that if $H\in\mathcal{E}_T$ if of the form (8.2) and $t\in(t_{k-1},t_k]$, then

$$H\mathbf{1}_{[0,t]} = \sum_{i=1}^{k-1} X_i\mathbf{1}_{(t_{i-1},t_i]} + X_k\mathbf{1}_{(t_{k-1},t]}\in\mathcal{E}_T$$

and $(H\cdot B)_t = (H\mathbf{1}_{[0,t]}\cdot B)_T$. This finishes the proof. □

8.2 Isometric Itô integral

Definition 8.8. *Let $\mathcal{B}(A)$ denote the Borel σ-algebra of a topological space A. We say that a process $H=(H_t)_{t\in[0,T]}$* ***measurable*** *if the map*

$$([0,T]\times\Omega, \mathcal{B}([0,T])\otimes\mathcal{F}) \ni (t,\omega)\mapsto H_t(\omega)\in(\mathbb{R},\mathcal{B}(\mathbb{R}))$$

is measurable. The σ-algebra $\mathcal{G}_{\mathrm{pr}}$ over $[0,T]\times\Omega$ generated by $(s,t]\times A$ with $s<t$ and $A\in\mathcal{F}_s$ is called the ***predictable σ-algebra****. We say that a process $H=(H_t)_{t\in[0,T]}$* ***predictable*** *if the map*

$$([0,T]\times\Omega, \mathcal{G}_{\mathrm{pr}}) \ni (t,\omega)\mapsto H_t(\omega)\in(\mathbb{R},\mathcal{B}(\mathbb{R}))$$

is measurable.

Every predictable process is adapted. In practice, we often encounter adapted and continuous processes. The following lemma demonstrates that such processes are predictable.

Lemma 8.9. *If the process $X=(X_t)_{t\in[0,T]}$ is adapted and continuous, then X is predictable.*

Proof. Define

$$X_t^{(n)} := \sum_{k=0}^{n-1} X_{kT/n}\,\mathbf{1}_{(kT/n,(k+1)T/n]}(t), \qquad t\in[0,T], \quad n\in\mathbb{N}_+.$$

Then $X^{(n)}$ is predictable and from the continuity of X, it follows that $X_t{}^{(n)}(\omega)\to X_t(\omega)$ for all $t\in(0,T]$ and $\omega\in\Omega$. The predictability of X follows from the fact that the pointwise limit of a sequence of measurable functions is measurable. □

Definition 8.10. *We define* $\mathcal{H}_T := L^2([0,T]\times\Omega, \mathcal{G}_{\mathrm{pr}}, \lambda\otimes\mathbb{P})$*, where* λ *is the Lebesgue measure and* $\mathcal{G}_{\mathrm{pr}}$ *is the predictable* σ*-algebra. That is* $\mathcal{H}_T$ *is the set of predictable processes* $H=(H_t)_{t\in[0,T]}$ *such that*

$$\|H\|_{\mathcal{H}_T} := \sqrt{\mathbb{E}\left(\int_0^T H_s^2\,\mathrm{d}s\right)} < \infty.$$

Simple predictable processes are predictable and square integrable. Hence, $\mathcal{E}_T \subset \mathcal{H}_T$.

Lemma 8.11. $\mathcal{H}_T$ *coincides with the closure* $\bar{\mathcal{E}}_T$ *of* $\mathcal{E}_T$ *in* $L^2([0,T]\times\Omega, \mathcal{B}([0,T])\otimes\mathcal{F}, \lambda\otimes\mathbb{P})$.

Proof. Since $\mathcal{H}_T$ is closed we have $\bar{\mathcal{E}}_T\subset\mathcal{H}_T$. It remains to prove that $\bar{\mathcal{E}}_T$ is dense in $\mathcal{H}_T$. Denote by $\mathcal{G}_{\mathrm{pr}}^\circ$ the algebra generated by sets of the form $(s,t]\times A$ with $A\in\mathcal{F}_s$. Elements of $\mathcal{G}_{\mathrm{pr}}^\circ$ are of the form $((t_0,t_1]\times A_1)\cup\ldots\cup((t_{n-1},t_n]\times A_n)$ for some $0=t_0<t_1<\ldots<t_n=T$ and $A_i\in\mathcal{F}_{t_{i-1}}$. Observe that:

(i). $1_G\in\mathcal{E}_T\subset\bar{\mathcal{E}}_T$ for every set $G\in\mathcal{G}_{\mathrm{pr}}^\circ$.

(ii). $\bar{\mathcal{E}}_T$ is a vector space.

(iii). If $f_n\in\bar{\mathcal{E}}_T$ is a sequence of non-negative functions that increase to a bounded function f, then $f\in\bar{\mathcal{E}}_T$.

It follows from the monotone class theorem that $\bar{\mathcal{E}}_T$ contains all bounded functions that are measurable with respect to $\sigma(\mathcal{G}_{\mathrm{pr}}^\circ)=\mathcal{G}_{\mathrm{pr}}$. □

Definition 8.12. *Let* $\mathcal{M}_T$ *be the set of equivalence classes of indistinguishable continuous martingales* $M=(M_t)_{t\in[0,T]}$ *such that* $M_0=0$ *and*

$$\|M\|_{\mathcal{M}_T} := \sqrt{\mathbb{E}M_T^2} < \infty.$$

Lemma 8.13. $\mathcal{M}_T$ *is a Hilbert space and we have*

$$\mathbb{E}\left(\sup_{t\in[0,T]} M_t^2\right) \leqslant 4\|M\|_{\mathcal{M}_T}^2. \tag{8.4}$$

Proof. The bound (8.4) follows immediately from Doob's inequality stated in Theorem 6.14 (b). For the proof that $\mathcal{M}_T$ is a Hilbert space see Exercise 4, Sheet 3. □

Note that by Proposition 8.7 for all $H\in\mathcal{E}_T$, the process $(H\cdot B)_{t\in[0,T]}$ is a continuous martingale such that

$$\|H\cdot B\|_{\mathcal{M}_T}^2 = \mathbb{E}((H\cdot B)_T)^2 = \mathbb{E}\left(\int_0^T H_s^2\,\mathrm{d}s\right) = \|H\|_{\mathcal{H}_T}^2 < \infty.$$

Hence, the map

$$I^\circ : \mathcal{H}_T \supset \mathcal{E}_T \ni H \mapsto (H\cdot B)_{t\in[0,T]} \in \mathcal{M}_T$$

is well-defined and is an isometry. In particular, the map I° is bounded. Since $\mathcal{E}_T$ is dense in $\mathcal{H}_T$ the map $I^\circ\colon\mathcal{E}_T\to\mathcal{M}_T$ extends to the unique map $I\colon\mathcal{H}_T\to\mathcal{M}_T$. We have $I(H)=\lim_{n\to\infty}I^\circ(H^{(n)})$ for every sequence $(H^{(n)})_{n\in\mathbb{N}_+}$ of elements of $\mathcal{E}_T$ converging to $H\in\mathcal{H}_T$. The map $I\colon\mathcal{H}_T\to\mathcal{M}_T$ is called the **Itô isometry**.

Definition 8.14. *The integral of $H \in \mathcal{H}_T$ with respect to the Brownian motion B is defined by*

$$(H \cdot B)_{t \in [0,T]} \equiv \left(\int_0^t H_s \, \mathrm{d}B_s \right)_{t \in [0,T]} := I(H)$$

and satisfies the properties formulated in Proposition 8.7.

Remark 8.15. Note that we cannot say that the value of the integral at ω depends only on the paths $t \mapsto H_t(\omega)$ and $t \mapsto B_t(\omega)$, as the integral is not defined pathwise.

Theorem 8.16. (Stopping of stochastic integral) *Let $\tau \geqslant 0$ be a stopping time and $H \in \mathcal{H}_T$. Then a.s.*

$$(H \cdot B)_{\tau \wedge t} = (\mathbf{1}_{[0,\tau]} H \cdot B)_t, \qquad t \in [0,T].$$

Proof. *Step 1.* Let $H \in \mathcal{E}_T$ and τ takes only finitely many values. By possibly extending the sequence of times we can assume that τ takes values $0 = t_0 \leqslant t_1 \leqslant \ldots \leqslant t_n \leqslant T$ and $H = \sum_{i=1}^n X_i \, \mathbf{1}_{(t_{i-1}, t_i]}$. Then we have

$$\mathbf{1}_{[0,\tau]}(t) H_t = \sum_{i=1}^n \mathbf{1}_{[0,\tau]}(t) X_i \, \mathbf{1}_{(t_{i-1}, t_i]}(t) = \sum_{i=1}^n \mathbb{1}_{\{\tau \geqslant t\}} X_i \, \mathbf{1}_{(t_{i-1}, t_i]}(t) = \sum_{i=1}^n \mathbb{1}_{\{\tau > t_{i-1}\}} X_i \, \mathbf{1}_{(t_{i-1}, t_i]}(t).$$

Since $\mathbb{1}_{\{\tau > t_{i-1}\}} X_i$ is $\mathcal{F}_{t_{i-1}}$-measurable, $\mathbf{1}_{[0,\tau]} H \in \mathcal{E}_T$. We compute

$$\begin{aligned}
(H \mathbf{1}_{[0,\tau]} \cdot B)_t &= \sum_{i=1}^n \mathbb{1}_{\{\tau > t_{i-1}\}} X_i (B_{t_i \wedge t} - B_{t_{i-1} \wedge t}) \\
&= \sum_{i=1}^n \sum_{j=i}^n \mathbb{1}_{\{\tau = t_j\}} X_i (B_{t_i \wedge t} - B_{t_{i-1} \wedge t}) \\
&= \sum_{j=1}^n \mathbb{1}_{\{\tau = t_j\}} \sum_{i=1}^j X_i (B_{t_i \wedge t} - B_{t_{i-1} \wedge t}) \\
&= \sum_{j=1}^n \mathbb{1}_{\{\tau = t_j\}} (H \cdot B)_{t_j \wedge t} = (H \cdot B)_{\tau \wedge t}
\end{aligned}$$

Step 2. Let $H \in \mathcal{E}_T$ and τ arbitrary stopping time. Take a sequence of stopping times τ_n taking finitely many values such that $\tau_n \searrow \tau$. By Step 1, $(H \cdot B)_{\tau_n \wedge t} = (\mathbf{1}_{[0,\tau_n]} H \cdot B)_t$. By the continuity of the stochastic integral, $(H \cdot B)_{\tau_n \wedge t} \to (H \cdot B)_{\tau \wedge t}$ a.s. On the other hand, by linearity of the integral and Itô isometry

$$\mathbb{E}((\mathbf{1}_{[0,\tau_n]} H \cdot B)_t - (\mathbf{1}_{[0,\tau]} H \cdot B)_t)^2 = \mathbb{E}((\mathbf{1}_{(\tau,\tau_n]} H \cdot B)_t)^2 = \mathbb{E} \int_0^t \mathbf{1}_{(\tau,\tau_n]}(s) H_s^2 \, \mathrm{d}s \to 0.$$

The convergence follows from the Lebesgue theorem, as the process $\mathbf{1}_{(\tau,\tau_n]}(s) H_s^2$ converges pointwise to zero and is dominated by H_s^2. Hence

$$(H \cdot B)_{\tau \wedge t} \xleftarrow{a.s.} (H \cdot B)_{\tau_n \wedge t} = (\mathbf{1}_{[0,\tau_n]} H \cdot B)_t \xrightarrow{L_2(\Omega)} (\mathbf{1}_{[0,\tau]} H \cdot B)_t.$$

That is, $(H \cdot B)_{\tau \wedge t} = (\mathbf{1}_{[0,\tau]} H \cdot B)_t$.

Step 3. Let $H \in \mathcal{H}_T$ and τ arbitrary stopping time. We take $H^{(n)} \in \mathcal{E}_T$ such that $H^{(n)} \to H$ in $\mathcal{H}_T$. From Step 2 we know that $(H^{(n)} \cdot B)_{\tau \wedge t} = (\mathbf{1}_{[0,\tau]} H^{(n)} \cdot B)_t$. We have

$$\mathbb{E}((H \cdot B)_{\tau \wedge t} - (H^{(n)} \cdot B)_{\tau \wedge t})^2 \leqslant 4\mathbb{E}((H - H^{(n)}) \cdot B)_T^2 = 4\mathbb{E}\int_0^T (H_s - H_s^{(n)})^2 \,\mathrm{d}s \to 0,$$

where the inequality follows from Doob's Theorem 6.14 applied to the martingale $((H - H^{(n)}) \cdot B)$. Moreover,

$$\mathbb{E}((\mathbf{1}_{[0,\tau]} H \cdot B)_t - (\mathbf{1}_{[0,\tau]} H^{(n)} \cdot B)_t)^2 = \mathbb{E}\int_0^t \mathbf{1}_{[0,\tau]}(s)\big(H_s - H_s^{(n)}\big)^2 \,\mathrm{d}s \leqslant \mathbb{E}\int_0^T (H_s - H_s^{(n)})^2 \,\mathrm{d}s \to 0.$$

In consequence,

$$(H \cdot B)_{\tau \wedge t} \xleftarrow{L_2(\Omega)} (H^{(n)} \cdot B)_{\tau \wedge t} = (\mathbf{1}_{[0,\tau]} H^{(n)} \cdot B)_t \xrightarrow{L_2(\Omega)} (\mathbf{1}_{[0,\tau]} H \cdot B)_t$$

Step 4. By the previous step we know that for every $t \geqslant 0$ there is an event $E_t \subset \Omega$ such that $\mathbb{P}(E_t) = 1$ and $(H \cdot B)_{\tau \wedge t} = (\mathbf{1}_{[0,\tau]} H \cdot B)_t$ on the event E_t. Let $E = \bigcap_{t \in \mathbb{Q} \cap [0,\infty)} E_t$, where $\mathbb{Q}$ is the set of rational numbers. Then E is an event such that $\mathbb{P}(E) = 1$ and $(H \cdot B)_{\tau \wedge t} = (\mathbf{1}_{[0,\tau]} H \cdot B)_t$ on the event E for all $t \in \mathbb{Q} \cap [0, \infty)$. The statement follows now from the continuity of the stochastic integral. $\square$

8.3 Localization

Suppose that f is a continuous function. Using Itô isometry we can define the stochastic integral

$$\left(\int_0^t f(B_s) \,\mathrm{d}B_s\right)_{t \in [0,T]} = I(f(B_s)_{s \in [0,T]})$$

only if $(f(B_s))_{s \in [0,T]} \in \mathcal{H}_T$. Since $(f(B_s))_{s \in [0,T]}$ is predictable, we only need to assume that $\mathbb{E}\big(\int_0^T f(B_s)^2 \,\mathrm{d}s\big) < \infty$, which is, unfortunately, a quite restrictive condition.

Example 8.17. For $f(x) = \exp(x^4)$ we have

$$\mathbb{E}\left(\int_0^T (f(B_s))^2 \,\mathrm{d}s\right) = \mathbb{E}\left(\int_0^T \mathrm{e}^{2B_s^4} \,\mathrm{d}s\right) = \int_0^T \mathbb{E}(\mathrm{e}^{2B_s^4})\mathrm{d}s = \int_0^T \left(\int_{\mathbb{R}} \mathrm{e}^{2x^4} \frac{\mathrm{e}^{-\frac{x^2}{2s}}}{\sqrt{2\pi s}} \,\mathrm{d}x\right) \mathrm{d}s = \infty$$

Thus, $(\mathrm{e}^{B_s^4})_{s \in [0,T]} \notin \mathcal{H}_T$ despite the fact that the function $x \mapsto \exp(x^4)$ is smooth.

Definition 8.18. *Let $\mathcal{H}_{T,\mathrm{loc}}$ be the space of equivalence classes of indistinguishable predictable processes $H = (H_t)_{t \in [0,T]}$ such that $\int_0^T H_s^2 \,\mathrm{d}s < \infty$ a.s.*

Definition 8.19. *A non-decreasing sequence of stopping times $(\tau_n)_{n \in \mathbb{N}_+}$ taking values in $[0, T]$ is a **localizing sequence** for $H \in \mathcal{H}_{T,\mathrm{loc}}$ if:*

(1) $H\mathbf{1}_{[0,\tau_n]} \in \mathcal{H}_T$ for every $n \in \mathbb{N}_+$ and

(2) $\mathbb{P}(\exists_{n \in \mathbb{N}_+} \tau_n = T) = 1$.

Proposition 8.20. *Let $H \in \mathcal{H}_{T,\mathrm{loc}}$ and define*

$$\tau_n = \inf\left\{ t \in [0,T] \,\middle|\, \int_0^t H_s^2 \,\mathrm{d}s \geqslant n \right\} \wedge T.$$

Then $(\tau_n)_{n\in\mathbb{N}_+}$ is a localizing sequence for H.

Proof. Since $(\int_0^t H_s^2 \,\mathrm{d}s)_{t\geqslant 0}$ is a continuous adapted process, by Remark 6.10 the random variable τ_n is a stopping time. It is evident that $\tau_n \leqslant \tau_{n+1}$. Moreover,

$$\mathbb{P}(\exists_{n\in\mathbb{N}_+} \tau_n = T) = \mathbb{P}\left(\int_0^T H_t^2 \,\mathrm{d}t < \infty \right) = 1.$$

Finally, $H\mathbf{1}_{[0,\tau_n]}$ is predictable and

$$\mathbb{E}\left(\int_0^T (H_t \,\mathbf{1}_{[0,\tau_n]}(t))^2 \mathrm{d}t \right) = \mathbb{E}\left(\int_0^{\tau_n} H_t^2 \,\mathrm{d}t \right) \leqslant n < \infty.$$

This finishes the proof. ☐

Lemma 8.21. *Suppose that $H \in \mathcal{H}_{T,\mathrm{loc}}$ and $(\tau_n)_{n\in\mathbb{N}_+}$ is a localizing sequence for H. There exists an event E of probability one such that on the event E we have*

$$(H\mathbf{1}_{[0,\tau_m]} \cdot B)_{\tau_n \wedge t} = (H\mathbf{1}_{[0,\tau_n]} \cdot B)_t,$$

for all $t \in [0,T]$, $n, m \in \mathbb{N}_+$ such that $n \leqslant m$.

Proof. By Theorem 8.16 and the identity $H\mathbf{1}_{[0,\tau_m]}\mathbf{1}_{[0,\tau_n]} = H\mathbf{1}_{[0,\tau_n]}$ for every $n, m \in \mathbb{N}_+$ such that $n \leqslant m$, there exists an event $E_{n,m}$ of probability one such that the stated equality holds true for all $t \in [0,T]$ and all sample points from $E_{n,m}$. To conclude the proof we set $E = \bigcap_{n,m\in\mathbb{N}_+, n\leqslant m} E_{n,m}$. ☐

Definition 8.22. *Let $H \in \mathcal{H}_{T,\mathrm{loc}}$ and $(\tau_n)_{n\in\mathbb{N}_+}$ be a localizing sequence for H. The integral of H with respect to the Brownian motion B is a continuous process $(H \cdot B)_{t\in[0,T]}$ such that*

$$(H \cdot B)_{t\wedge\tau_n} \equiv \int_0^{t\wedge\tau_n} H_s \,\mathrm{d}B_s = (H\mathbf{1}_{[0,\tau_n]} \cdot B)_t, \qquad t \in [0,T], \quad n \in \mathbb{N}_+, \tag{8.5}$$

holds true on an event of probability one.

Proposition 8.23. *The process $(H \cdot B)_{t\in[0,T]}$ satisfying the above conditions exists, is unique up to indistinguishability and does not depend on the choice of the localizing sequence.*

Proof. Let $(\tau_n)_{n\in\mathbb{N}_+}$ be a localizing sequence and E be the event from Lemma 8.21. On the event $A_n := E \cap \{\tau_n = T\}$ we define

$$(H \cdot B)_t := (H\mathbf{1}_{[0,\tau_n]} \cdot B)_t, \qquad t \in [0,T].$$

The above definition is consistent since by Lemma 8.21 for all $m \geqslant n$ we have

$$(H\mathbf{1}_{[0,\tau_n]} \cdot B)_t = (H\mathbf{1}_{[0,\tau_m]} \cdot B)_t$$

on the event $E\cap\{\tau_n=T\}$. Moreover, it guarantees that (8.5) is satisfied. Uniqueness is evident since $\mathbb{P}\big(\bigcup_{n\in\mathbb{N}_+}A_n\big)=1$ by Definition 8.19 and Lemma 8.21. Continuity of $(H\cdot B)_{t\in[0,T]}$ follows from continuity of $(H\mathbf{1}_{[0,\tau_n]}\cdot B)_{t\in[0,T]}$. The fact that $(H\cdot B)_{t\in[0,T]}$ does not depend on the choice of a localizing sequence is a consequence of Theorem 8.16. □

Example 8.24. The integral $X_t=\int_0^t f(B_s)\,\mathrm{d}B_s$ is well-defined for any continuous function $f\colon\mathbb{R}\to\mathbb{R}$ but need not be a martingale. For example,

$$X_t=\int_0^t \exp(B_s^4)\,\mathrm{d}B_s$$

is well-defined but is not a martingale. One shows that $\mathbb{E}|X_t|=\infty$ and $\mathbb{E}X_t$ is not defined.

Definition 8.25. *A process $M=(M_t)_{t\in[0,T]}$ is a* ***local martingale*** *if there exists a non-decreasing sequence $(\tau_n)_{n\in\mathbb{N}_+}$ of stopping times such that $\lim_{n\to\infty}\tau_n=T$ a.s and such that $(M_{t\wedge\tau_n})_{t\in[0,T]}$ is a martingale. We say that the sequence $(\tau_n)_{n\in\mathbb{N}_+}$ as above* ***reduces*** *the local martingale M. Let $\mathcal{M}_{T,\mathrm{loc}}$ be the set of equivalence classes of indistinguishable continuous local martingales $M=(M_t)_{t\in[0,T]}$ such that $M_0=0$.*

Proposition 8.26. *For all $H\in\mathcal{H}_{T,\mathrm{loc}}$ we have*

$$(H\cdot B)_{t\in[0,T]}\in\mathcal{M}_{T,\mathrm{loc}}.$$

Proof. Let $(\tau_n)_{n\in\mathbb{N}_+}$ be as in Definition 8.22. By Definition 8.14 $(H\mathbf{1}_{[0,\tau_n]}\cdot B)_{t\in[0,T]}\in\mathcal{M}_T$. Hence, the statement follows immediately from Definition 8.22. □

9 Integral with respect to continuous local martingale

In the previous section, we defined the integral $\int H_s\,\mathrm{d}B_s$. It turns out that without much difficulty, this definition can be generalized to $\int H_s\,\mathrm{d}M_s$, where $(M_t)_{t\geqslant0}$ is a continuous martingale (or even a continuous local martingale).

9.1 Doob-Meyer decomposition

The foundation of the stochastic integral construction with respect to a Brownian motion is that $(B_t)_{t\geqslant0}$ and $(B_t^2-t)_{t\geqslant0}$ are martingales. It turns out that for any square-integrable continuous martingale $(M_t)_{t\geqslant0}$, there exists a non-decreasing process $(Y_t)_{t\geqslant0}$ such that $(M_t^2-Y_t)_{t\geqslant0}$ is a martingale.

Theorem 9.1. (Doob-Meyer decomposition) *Let $M\in\mathcal{M}_T$. There exists a process $\langle M\rangle=(\langle M\rangle_t)_{t\in[0,T]}$ with continuous, non-decreasing paths such that $\langle M\rangle_0=0$ and $(M_t^2-\langle M\rangle_t)_{t\in[0,T]}$ is a martingale. Moreover, the process $\langle M\rangle$ is uniquely determined up to indistinguishability.*

Proof. We will only prove the uniqueness of the decomposition; the proof of existence can be found in Sec. IV.1 of [RY04]. Assume that the processes $(Y_t)_{t\in[0,T]}$ and $(Z_t)_{t\in[0,T]}$ are non-decreasing and $(M_t^2-Y_t)_{t\in[0,T]}$ and $(M_t^2-Z_t)_{t\in[0,T]}$ are martingales with continuous trajectories. The paths of the process Y_t-Z_t have finite variation, and furthermore, $Y_t-Z_t=(M_t^2-Z_t)-(M_t^2-Y_t)$ is a continuous martingale. Therefore, by Proposition 7.13, $Y_t-Z_t=0$ for all $t\in[0,T]$ a.s. □

Example 9.2. If B is a Brownian motion, then $\langle B\rangle_t = t$. Indeed,

$$\mathbb{E}(B_t^2 - t|\mathcal{F}_s) = \mathbb{E}(B_s^2 + 2(B_t - B_s)B_s + (B_t - B_s)^2 - t|\mathcal{F}_s) = B_s^2 + 0 + (t - s) - t = B_s^2 - s.$$

Remark 9.3. An integrable right-continuous adapted process $(X_t)_{t\geqslant 0}$ is a martingale if and only if, for every bounded stopping time τ we have $\mathbb{E}X_\tau = \mathbb{E}X_0$. See Problem 3, Sheet 5.

Proposition 9.4. *Let $H \in \mathcal{H}_T$. Then $\langle \int H_s \,\mathrm{d}B_s\rangle_t = \int_0^t H_s^2 \,\mathrm{d}s$.*

Proof. We have to prove that the process $M = (M_t)_{t\in[0,T]}$ defined by

$$M_t = \left(\int_0^t H_s \,\mathrm{d}B_s\right)^2 - \int_0^t H_s^2 \,\mathrm{d}s$$

is a martingale. We know that $(\int_0^t H_s \,\mathrm{d}B_s)_{t\in[0,T]}$ is continuous, adapted and square integrable. Hence, M is continuous, adapted and square integrable. For a bounded stopping time $\tau \in [0, T]$, by Theorem 8.16 and the Itô isometry we obtain

$$\mathbb{E}\left(\left(\int_0^\tau H_s \,\mathrm{d}B_s\right)^2\right) = \mathbb{E}\left(\left(\int_0^T \mathbf{1}_{[0,\tau]}(s)\, H_s \,\mathrm{d}B_s\right)^2\right) = \mathbb{E}\left(\int_0^T \mathbf{1}_{[0,\tau]}(s)\, H_s^2 \,\mathrm{d}s\right) = \mathbb{E}\left(\int_0^\tau H_s^2 \,\mathrm{d}s\right).$$

Therefore,

$$\mathbb{E}M_\tau = \mathbb{E}\left(\left(\int_0^\tau H_s \,\mathrm{d}B_s\right)^2 - \int_0^\tau H_s^2 \,\mathrm{d}s\right) = 0 = \mathbb{E}M_0.$$

The statement follows from Remark 9.3. □

Remark 9.5. If $M = (M_t)_{t\in[0,T]}$ is a continuous bounded martingale, then

$$\langle M\rangle_t = \lim_{n\to\infty} \sum_{i=1}^n |M_{t_i} - M_{t_{i-1}}|^2$$

in mean square, where $0 = t_0^{(n)} < t_1^{(n)} < \ldots < t_n^{(n)} = t \leqslant T$ are arbitrary such that

$$\lim_{n\to\infty} \max_{i\in\{1,\ldots,n\}} (t_i^{(n)} - t_{i-1}^{(n)}) = 0.$$

See Exercise 3, Sheet 7.

9.2 Integral for elementary processes

Definition 9.6. *The integral of a simple predictable process $H \in \mathcal{E}_T$ of the form*

$$H_t = \sum_{i=1}^n X_i \,\mathbf{1}_{(t_{i-1},t_i]}$$

with respect to $M \in \mathcal{M}_T$ is defined by

$$(H \cdot M)_t \equiv \int_0^t H_s \,\mathrm{d}M_s := \sum_{i=1}^n X_i(M_{t_i\wedge t} - M_{t_{i-1}\wedge t}), \qquad t \in [0, T].$$

Proposition 9.7. *Let $M\in\mathcal{M}_T$, $H,K\in\mathcal{E}_T$ and $a,b\in\mathbb{R}$. Then:*

(i). $((aH+bK)\cdot M)_t=a(H\cdot M)_t+b(K\cdot M)_t$.

(ii). $\mathbb{E}(H\cdot M)_t=0$.

(iii). $\mathbb{E}((H\cdot M)_t^2)=\mathbb{E}\int_0^t H_s^2\,\mathrm{d}\langle M\rangle_s$.

(iv). $((H\cdot M)_t)_{t\in[0,T]}$ *is a continuous martingale.*

(v). $(H\cdot M)_t=(H\mathbf{1}_{[0,t]}\cdot M)_T$.

Proof. The proof is almost identical to the proof of Proposition 8.7. To prove Item (iii) we use the following observation

$$\begin{aligned}\mathbb{E}((M_t-M_s)^2|\mathcal{F}_s) &= \mathbb{E}(M_t^2-\langle M\rangle_t|\mathcal{F}_s)+\mathbb{E}(\langle M\rangle_t|\mathcal{F}_s)-2M_s\,\mathbb{E}(M_t|\mathcal{F}_s)+M_s^2\\ &= M_s^2-\langle M\rangle_s+\mathbb{E}(\langle M\rangle_t|\mathcal{F}_s)-M_s^2=\mathbb{E}(\langle M\rangle_t-\langle M\rangle_s|\mathcal{F}_s),\end{aligned}$$

which is a generalization of (8.3). □

9.3 Isometric Itô integral

Definition 9.8. *For $M\in\mathcal{M}_T$ let $\mathcal{H}_T(M)$ be the space of predictable processes $H=(H_t)_{t\in[0,T]}$ such that*

$$\|H\|_{\mathcal{H}_T(M)}:=\sqrt{\mathbb{E}\bigg(\int_0^T H_s^2\,\mathrm{d}\langle M\rangle_s\bigg)}<\infty$$

Lemma 9.9. *$\mathcal{E}_T\subset\mathcal{H}_T(M)$ is dense.*

Proof. The statement can be proved along the lines of the proof of Lemma 8.11. □

Note that by Proposition 9.7 for all $M\in\mathcal{M}_T$ and $H\in\mathcal{E}_T$, the process $(H\cdot M)$ is a continuous martingale such that

$$\|H\cdot M\|_{\mathcal{M}_T}^2=\mathbb{E}((H\cdot M)_T)^2=\mathbb{E}\int_0^T H_s^2\,\mathrm{d}\langle M\rangle_s=\|H\|_{\mathcal{H}_T(M)}^2<\infty.$$

Hence, the map

$$I_M^\circ:\mathcal{H}_T(M)\supset\mathcal{E}_T\ni H\mapsto(H\cdot B)_{t\in[0,T]}\in\mathcal{M}_T$$

is well-defined and is an isometry. In particular, the map I° is bounded. Since $\mathcal{E}_T$ is dense in $\mathcal{H}_T$ the map $I_M^\circ:\mathcal{E}_T\to\mathcal{M}_T$ extends to the unique map $I_M:\mathcal{H}_T(M)\to\mathcal{M}_T$. We have $I_M(H)=\lim_{n\to\infty}I_M^\circ(H^n)$ for every sequence $(H^n)_{n\in\mathbb{N}_+}$ of elements of $\mathcal{E}_T$ converging to $H\in\mathcal{H}_T(M)$.

Definition 9.10. *The integral of $H \in \mathcal{H}_T(M)$ with respect to $M \in \mathcal{M}_T$ is defined by*

$$(H \cdot M)_{t \in [0,T]} \equiv \left(\int_0^t H_s \, \mathrm{d}M_s \right)_{t \in [0,T]} := I_M(H)$$

and satisfies the properties formulated in Proposition 9.7.

9.4 Localization

Notation 9.11. *For a stochastic process $X = (X_t)_{t \geqslant 0}$ and a stopping time $\tau \geqslant 0$ we denote by X^τ the stopped process $(X_{t \wedge \tau})_{t \geqslant 0}$.*

Lemma 9.12. *For all $M \in \mathcal{M}_T$ and all stopping times τ we have $M^\tau \in \mathcal{M}_T$ and $\langle M^\tau \rangle = \langle M \rangle^\tau$.*

Proof. We know that M^τ is a continuous martingale. By Theorem 6.12 and the Jensen inequality,

$$\mathbb{E}(M_t^\tau)^2 = \mathbb{E}M_{t\wedge\tau}^2 = \mathbb{E}(\mathbb{E}(M_T | \mathcal{F}_{t\wedge\tau})^2) \leqslant \mathbb{E}(\mathbb{E}(M_T^2 | \mathcal{F}_{t\wedge\tau})) = \mathbb{E}(M_T).$$

Thus, $M^\tau \in \mathcal{M}_T$. The process $\langle M \rangle^\tau$ starts from zero, has continuous, non-decreasing paths and

$$(M^\tau)^2 - \langle M \rangle^\tau = (M^2 - \langle M \rangle)^\tau$$

is a martingale, so $\langle M \rangle^\tau$ satisfies all the conditions of the definition of $\langle M^\tau \rangle$. □

We can generalize the Doob-Meyer decomposition to the case of continuous local martingales.

Proposition 9.13. *For all $M \in \mathcal{M}_{T,\mathrm{loc}}$ there exists a process $\langle M \rangle = (\langle M \rangle_t)_{t \in [0,T]}$ with continuous, non-decreasing paths such that $\langle M \rangle_0 = 0$ and $(M_t^2 - \langle M \rangle_t)_{t \geqslant 0} \in \mathcal{M}_{T,\mathrm{loc}}$. Moreover, the process $\langle M \rangle$ is uniquely determined up to indistinguishability.*

Proof. Since $M \in \mathcal{M}_{T,\mathrm{loc}}$, there exists an increasing sequence of stopping times $\hat{\tau}_n$ converging to T such that $M^{\hat{\tau}_n}$ is a martingale. Define

$$\tilde{\tau}_n := \inf \{ t \geqslant 0 \,|\, |M_t^{\hat{\tau}_n}| \geqslant n \}.$$

Then $(M^{\hat{\tau}_n})^{\tilde{\tau}_n} = M^{\tau_n}$, where $\tau_n := \hat{\tau}_n \wedge \tilde{\tau}_n$, is a bounded martingale. Hence, $M^{\tau_n} \in \mathcal{M}_T$. Define $Y^{(n)} = \langle M^{\tau_n} \rangle$, then for $n \leqslant m$,

$$Y^{(n)} = \langle M^{\tau_n} \rangle = \langle (M^{\tau_m})^{\tau_n} \rangle = \langle M^{\tau_m} \rangle^{\tau_n} = (Y^{(m)})^{\tau_n}.$$

Hence, there exists a continuous process $Y = (Y_t)_{t \in [0,T]}$ satisfying $Y_t = Y_t^{(n)}$ on the events $\{t \leqslant \tau_n\}$, $n \in \mathbb{N}_+$. Obviously, $Y_0 = Y_0^{(n)} = 0$. Moreover, Y has non-decreasing paths and

$$(M^2 - Y)^{\tau_n} = (M^{\tau_n})^2 - Y^{\tau_n} = (M^{\tau_n})^2 - \langle M^{\tau_n} \rangle,$$

so $M^2 - Y$ is a continuous local martingale on $[0,T]$. This proves existence.

To show uniqueness, let Y and $\tilde{Y}$ be continuous processes with non-decreasing trajectories such that $Y_0 = \tilde{Y}_0 = 0$ and $M^2 - Y$ and $M^2 - \tilde{Y}$ are local martingales. There exists a sequence of stopping times $\tau_n \nearrow T$ such that $(M^2 - Y)^{\tau_n}$ and $(M^2 - \tilde{Y})^{\tau_n}$ are martingales. The process $(Y - \tilde{Y})^{\tau_n}$ is therefore a martingale with bounded variation. Thus, by Proposition 7.13 it is constant. Consequently, $Y^{\tau_n} = \tilde{Y}^{\tau_n}$. Upon taking the limit $n \to \infty$, we obtain $Y = \tilde{Y}$. $\square$

Along the lines of the proof of Theorem 8.16, one proves a stopping theorem for the stochastic integral with respect to square-integrable martingales.

Theorem 9.14. *Let $\tau \geqslant 0$ be a stopping time, $M \in \mathcal{M}_T$ and $H \in \mathcal{H}_T(M)$. Then a.s.*

$$(H \cdot M)_{\tau \wedge t} = (H \mathbf{1}_{[0,\tau]} \cdot M^{\tau})_t, \qquad t \in [0, T].$$

Definition 9.15. *For $M \in \mathcal{M}_{T,\mathrm{loc}}$ let $\mathcal{H}_{T,\mathrm{loc}}(M)$ be the space of equivalence classes of indistinguishable predictable processes $H = (H_t)_{t \in [0,T]}$ such that $\int_0^T H_s^2 \, \mathrm{d}\langle M \rangle_s < \infty$ a.s.*

Definition 9.16. *Let $M \in \mathcal{M}_{T,\mathrm{loc}}$, $H \in \mathcal{H}_{T,\mathrm{loc}}(M)$ and $(\tau_n)_{n \in \mathbb{N}_+}$ be an increasing sequence of stopping times converging to T such that $M^{\tau_n} \in \mathcal{M}_T$ and $\mathbf{1}_{[0,\tau_n]} H \in \mathcal{H}_T(M^{\tau_n})$. The integral of H with respect to M is a continuous process $(H \cdot M)_{t \in [0,T]}$ such that*

$$(H \cdot M)_{t \wedge \tau_n} \equiv \int_0^{t \wedge \tau_n} H_s \, \mathrm{d}M_s = (H \mathbf{1}_{[0,\tau_n]} \cdot M^{\tau_n})_t, \qquad t \in [0, T], \quad n \in \mathbb{N}_+,$$

holds true on an event of probability one.

Following the proof of Proposition 8.23 it is not difficult to show that the integral $(H \cdot M)_{t \in [0,T]}$ is well-defined, unique (up to indistinguishability) and does not depend on the choice of the sequence of stopping times τ_n. The following fact generalizing Theorem 9.14 is true.

Theorem 9.17. *Let $\tau \geqslant 0$ be a stopping time, $M \in \mathcal{M}_{T,\mathrm{loc}}$ and $H \in \mathcal{H}_{T,\mathrm{loc}}(M)$. Then a.s.*

$$(H \cdot M)_{\tau \wedge t} = (H \mathbf{1}_{[0,\tau]} \cdot M)_t = (H \mathbf{1}_{[0,\tau]} \cdot M^{\tau})_t = (H \cdot M^{\tau})_t \qquad t \in [0, T].$$

It is also possible to show that the constructions of the integral on $[0, T]$ and $[0, \tilde{T}]$ are consistent for arbitrary $T < \tilde{T}$, that is $(H \cdot M)_{t \in [0,T]}$ coincides with $(H \cdot M)_{t \in [0,\tilde{T}]}$ for all $t \in [0, T]$. This allows to define the process $(H \cdot M)_{t \geqslant 0}$ for all $H = (H_t)_{t \geqslant 0}$ and $M = (M_t)_{t \geqslant 0}$ such that $(H_t)_{t \in [0,T]} \in \mathcal{H}_{T,\mathrm{loc}}$ and $(M_t)_{t \in [0,T]} \in \mathcal{M}_{T,\mathrm{loc}}$ for all $T \in (0, \infty)$.

10 Quadratic covariation

The quadratic covariation is defined not only for a single martingale, but also for a pair of martingales.

Definition 10.1. *The quadratic covariation of two continuous local martingales M and N is the process $\langle M, N \rangle$ defined by the formula:*

$$\langle M, N \rangle = \frac{1}{4} \langle M + N \rangle - \frac{1}{4} \langle M - N \rangle.$$

Theorem 10.2. *Let $M, N \in \mathcal{M}_T$ (resp. $\mathcal{M}_{T,\text{loc}}$). Then $\langle M, N\rangle$ is the unique process on $[0, T]$ with continuous paths of bounded variation such that $\langle M, N\rangle_0 = 0$ and $MN - \langle M, N\rangle$ is a martingale (resp. a local martingale) on $[0, T]$.*

Proof. Uniqueness is proved as for $\langle M\rangle$, and the mentioned properties follow from the identity

$$MN - \langle M, N\rangle = \frac{1}{4}\big((M+N)^2 - \langle M+N\rangle\big) - \frac{1}{4}\big((M-N)^2 - \langle M-N\rangle\big). \qquad \square$$

Theorem 10.3. *Let $M, N \in \mathcal{M}_{T,\text{loc}}$. We have*

(a). $\langle M, M\rangle = \langle M\rangle = \langle -M\rangle$,

(b). $\langle M, N\rangle = \langle N, M\rangle$,

(c). $\langle M - M_0, N\rangle = \langle M, N - N_0\rangle = \langle M - M_0, N - N_0\rangle = \langle M, N\rangle$,

(d). $(N, M) \mapsto \langle M, N\rangle$ is a bilinear map,

(e). $\langle M^\tau, N^\tau\rangle = \langle M, N\rangle^\tau = \langle M, N^\tau\rangle = \langle M, N\rangle^\tau$ for every stopping time τ,

(f). If $H \in \mathcal{H}_{T,\text{loc}}(M)$ and $G \in \mathcal{H}_{T,\text{loc}}(N)$, then $\langle \int H\,\mathrm{d}M, \int G\,\mathrm{d}N\rangle = \int HG\,\mathrm{d}\langle M, N\rangle$.

Proof. See Exercise 1, Sheet 8. $\square$

11 Further properties of the stochastic integral

In this section, we will show a number of important properties of the stochastic integral, which will allow us later to prove the Itô formula.

11.1 Dominated convergence for stochastic integrals

Theorem 11.1. (Dominated convergence) *Let $M \in \mathcal{M}_{T,\text{loc}}$, $G \in \mathcal{H}_{T,\text{loc}}(M)$ and $H^{(n)} = \big(H_t^{(n)}\big)_{t \geqslant 0}$ be a sequence of predictable processes such that a.s. $\lim_{n\to\infty} H_t^{(n)} = H_t$ and $|H_t^{(n)}| \leq G_t$ for all $t \in [0, T]$. Then $H^{(n)}, H \in \mathcal{H}_{T,\text{loc}}(M)$ and for all $t \in [0, T]$*

$$\lim_{n\to\infty} \int_0^t H_s^{(n)} \mathrm{d}M_s = \int_0^t H_s\,\mathrm{d}M_s$$

in probability.

Proof. The process H is predictable as the limit of predictable processes. For $t \in [0, T]$ we have

$$\int_0^t H_s^2\,\mathrm{d}\langle M\rangle_s \vee \int_0^t \big(H_s^{(n)}\big)^2\,\mathrm{d}\langle M\rangle_s \leqslant \int_0^t G_s^2\,\mathrm{d}\langle M\rangle_s < \infty \qquad a.s.$$

Hence, $H^{(n)}, H \in \mathcal{H}_{T,\text{loc}}(M)$ and the integrals appearing in the statement of the theorem are well-defined.

Step 1. Suppose that $M\in\mathcal{M}_T$ and $G\in\mathcal{H}_T(M)$. We have

$$\lim_{n\to\infty} H_t^{(n)}=H_t, \qquad \big(H_t^{(n)}-H_t\big)^2\leqslant 4G_t^2, \qquad \|G\|_{\mathcal{H}_T(M)}=\mathbb{E}\int G_t^2\,\mathrm{d}\langle M\rangle_t<\infty.$$

Hence, by the Lebesgue dominated convergence theorem for the integral with respect to a function of a bounded variation we have $\lim_{n\to\infty} H^{(n)}\to H$ in $\mathcal{H}_T(M)$. Consequently, by the Itô isometry

$$\int_0^t H_s^{(n)}\mathrm{d}M_s \xrightarrow{L^2(\Omega)} \int_0^t H_s\mathrm{d}M_s$$

as $n\to\infty$. To conclude we note that convergence in mean square implies convergence in probability.

Step 2. Let $(\tau_k)_{k\in\mathbb{N}_+}$ be an increasing sequence of stopping times such that $\lim_{k\to\infty}\tau_k=T$ and

$$M^{\tau_k}\in\mathcal{M}_T, \quad \mathbf{1}_{[0,\tau_k]}G\in\mathcal{H}_T(M^{\tau_k}).$$

Since

$$\mathbf{1}_{[0,\tau_k]}H^{(n)}\leqslant\mathbf{1}_{[0,\tau_k]}G,$$

it follows that $\mathbf{1}_{[0,\tau_k]}H^{(n)}\in\mathcal{H}_T(M^{\tau_k})$. By Theorem 9.17 and Step 1 we obtain

$$\int_0^{t\wedge\tau_k} H_s^{(n)}\mathrm{d}M_s=\int_0^t \mathbf{1}_{[0,\tau_k]}(s)\,H_s^{(n)}\,\mathrm{d}M_s^{\tau_k} \xrightarrow{L^2(\Omega)} \int_0^t \mathbf{1}_{[0,\tau_k]}H_s\,\mathrm{d}M_s^{\tau_k}=\int_0^{t\wedge\tau_k} H_s\,\mathrm{d}M_s$$

as $n\to\infty$. Consequently, on the event $\{\tau_k\geqslant t\}$ we have

$$\int_0^t H_s^{(n)}\mathrm{d}M_s \xrightarrow{L^2(\Omega)} \int_0^t H_s\mathrm{d}M_s$$

as $n\to\infty$. Hence, for every $\delta>0$ and $k\in\mathbb{N}_+$,

$$\lim_{n\to\infty} A_{n,k}=0, \qquad A_{n,k}:=\mathbb{P}\bigg(\tau_k\geqslant t \text{ and } \bigg|\int_0^t H_s^{(n)}\mathrm{d}M_s-\int_0^t H_s\mathrm{d}M_s\bigg|\geqslant\delta\bigg).$$

Observe that

$$\mathbb{P}\bigg(\bigg|\int_0^t H_s^{(n)}\mathrm{d}M_s-\int_0^t H_s\mathrm{d}M_s\bigg|\geqslant\delta\bigg)\leqslant A_{n,k}+\mathbb{P}(\tau_k<t).$$

To conclude we choose $k\in\mathbb{N}_+$ big enough so that $\mathbb{P}(\tau_k<t)\leqslant\varepsilon/2$ and then $n\in\mathbb{N}_+$ big enough so that $A_{n,k}\leqslant\varepsilon/2$. □

11.2 Integration by substitution

Lemma 11.2. *Let $0\leqslant u\leqslant t\leqslant T$, $M\in\mathcal{M}_{T,\mathrm{loc}}$, $H\in\mathcal{H}_{T,\mathrm{loc}}(M)$ and Y be a bounded $\mathcal{F}_u$-measurable random variable. Then*

$$\int_u^t YH_s\,\mathrm{d}M_s=Y\int_u^t H_s\,\mathrm{d}M_s,$$

where $\int_u^t H_s\,\mathrm{d}M_s:=\int_0^t\mathbf{1}_{(u,t]}(s)\,H_s\,\mathrm{d}M_s$.

Proof. *Step 1.* Let $M\in\mathcal{M}_T$ and $H\in\mathcal{E}_T$. By possibly extending the sequence of times we can assume that $H=\sum_{i=1}^n X_i\,\mathbf{1}_{(t_{i-1},t_i]}$ and $u=t_k$. Since Y is $\mathcal{F}_u$-measurable we have $YH_s\in\mathcal{E}_T$ and

$$\int_u^t YH_s\,\mathrm{d}M_s=\sum_{i=k+1}^n YX_i(M_{t_i\wedge t}-M_{t_{i-1}\wedge t})=Y\sum_{i=k+1}^n X_i(M_{t_i\wedge t}-M_{t_{i-1}\wedge t})=Y\int_u^t H_s\,\mathrm{d}M_s.$$

Step 2. Let $M\in\mathcal{M}_T$ and $H\in\mathcal{H}_T(M)$. We take $H^{(n)}\in\mathcal{E}_T$ such that $H^{(n)}\to H$ in $\mathcal{H}_T(M)$. From Step 1 we know that $\int_u^t YH_s^{(n)}\,\mathrm{d}M_s=Y\int_u^t H_s^{(n)}\,\mathrm{d}M_s$. Let $C>0$ be a deterministic constant such that $|Y|\leqslant C$. By the Itô isometry

$$\mathbb{E}\left(\int_u^t Y\left(H_s-H_s^{(n)}\right)\mathrm{d}M_s\right)^2=\mathbb{E}\int_0^t Y^2\left(H_s-H_s^{(n)}\right)\mathrm{d}\langle M\rangle_s\leqslant C^2\,\|H-H^{(n)}\|^2_{\mathcal{H}_T(M)}\to 0.$$

Similarly,

$$\mathbb{E}\left(Y\int_u^t\left(H_s-H_s^{(n)}\right)\mathrm{d}M_s\right)^2\leqslant C^2\,\mathbb{E}\left(\int_u^t\left(H_s-H_s^{(n)}\right)\mathrm{d}M_s\right)^2\leqslant C^2\,\|H-H^{(n)}\|^2_{\mathcal{H}_T(M)}\to 0.$$

Hence, by the linearity of the stochastic integral we obtain

$$\int_u^t YH_s^{(n)}\,\mathrm{d}M_s\xleftarrow{L_2(\Omega)}\int_u^t YH_s^{(n)}\,\mathrm{d}M_s=Y\int_u^t H_s^{(n)}\,\mathrm{d}M_s\xrightarrow{L_2(\Omega)}Y\int_u^t H_s\,\mathrm{d}M_s.$$

Step 3. Let $M\in\mathcal{M}_{T,\mathrm{loc}}$ and $H\in\mathcal{H}_{T,\mathrm{loc}}(M)$. Let $(\tau_n)_{n\in\mathbb{N}_+}$ be an increasing sequence of stopping times converging to T a.s. such that $M^{\tau_n}\in\mathcal{M}_T$ and $\mathbf{1}_{[0,\tau_n]}H\in\mathcal{H}_T(M^{\tau_n})$. By Step 2 we have

$$\int_u^t\mathbf{1}_{[0,\tau_n]}(s)\,YH_s\,\mathrm{d}M_s^{\tau_n}=Y\int_u^t\mathbf{1}_{[0,\tau_n]}(s)\,H_s\,\mathrm{d}M_s^{\tau_n}.$$

By Theorem 9.17 we obtain

$$\int_u^{t\wedge\tau_n}YH_s\,\mathrm{d}M_s=Y\int_u^{t\wedge\tau_n}H_s\,\mathrm{d}M_s.$$

To complete the proof we take the limit $n\to\infty$. □

Definition 11.3. *A process $H=(H_t)_{t\in[0,T]}$ is **locally bounded** if there exists a sequence of stopping times $(\tau_n)_{n\in\mathbb{N}_+}$ such that $\tau_n\nearrow T$ a.s. and for all $n\in\mathbb{N}_+$ the process $(H_{t\wedge\tau_n}-H_0)_{t\geqslant 0}$ is bounded.*

Remark 11.4. Every continuous, adapted process is locally bounded since one can take

$$\tau_n=\inf\{t\in[0,T]\,|\,|H_t-H_0|\geqslant n\}\wedge T.$$

Theorem 11.5. (Integration by substitution) *(i) If $N\in\mathcal{M}_T$, $H\in\mathcal{H}_T(N)$, G is a bounded predictable process and $M=(H\cdot N)$, then $G\in\mathcal{H}_T(M)$, $HG\in\mathcal{H}_T(N)$ and $(G\cdot M)=(GH\cdot N)$.*

(ii) If $N\in\mathcal{M}_{T,\mathrm{loc}}$, $H\in\mathcal{H}_{T,\mathrm{loc}}(N)$, G is a locally bounded predictable process and $M=(H\cdot N)$, then $G\in\mathcal{H}_{T,\mathrm{loc}}(M)$, $HG\in\mathcal{H}_{T,\mathrm{loc}}(N)$ and $(GH\cdot N)=(G\cdot M)$.

Remark 11.6. Note that $(GH\cdot N)=(G\cdot M)$ is equivalent to $\int_0^t G_s\,H_s\,\mathrm{d}N_s=\int_0^t G_s\,\mathrm{d}M_s$, or more suggestively, $H_t\,\mathrm{d}N_t=\mathrm{d}(\int_0^t H_s\,\mathrm{d}N_s)=\mathrm{d}M_t$.

Proof. First, assume that G is a simple predictable process of the form $G=\sum_{i=1}^n X_i\,\mathbf{1}_{(t_{i-1},t_i]}$. Then

$$\begin{aligned}\int_0^t G_s\,\mathrm{d}M_s &= \sum_{i=1}^n X_i\,(M_{t_{i-1}\wedge t}-M_{t_i\wedge t})\\ &= \sum_{i=1}^n X_i\left(\int_0^t \mathbf{1}_{[0,t_{i-1}]}H_s\,\mathrm{d}N_s-\int_0^t 1_{[0,t_i]}H_s\,\mathrm{d}N_s\right)\\ &= \sum_{i=1}^n X_i\int_0^t \mathbf{1}_{(t_{i-1},t_i]}(s)H_s\,\mathrm{d}N_s\\ &= \sum_{i=1}^n \int_0^t X_i\,\mathbf{1}_{(t_{i-1},t_i]}(s)\,H_s\,\mathrm{d}N\\ &= \int_0^t\sum_{i=1}^n X_i\,\mathbf{1}_{(t_{i-1},t_i]}(s)\,H_s\,\mathrm{d}N_s=\int_0^t G_sH_s\,\mathrm{d}N_s,\end{aligned}$$

where the fourth equality above follows from Lemma 11.2.

a) Let G be a bounded predictable process and $C>0$ be a deterministic constant such that $|G|\leqslant C$. Then

$$\mathbb{E}\int_0^T G_s^2\,\mathrm{d}\langle M\rangle_s\leqslant C^2\,\mathbb{E}\int_0^T \mathrm{d}\langle M\rangle_s=C^2\,\mathbb{E}\langle M\rangle_T=C^2\,\mathbb{E}M_T^2<\infty.$$

Thus, $G\in\mathcal{H}_T(M)$. By a similar argument, $GH\in\mathcal{H}_T(N)$. We can find $G^{(n)}\in\mathcal{E}_T$ such that $\lim_{n\to\infty}G^{(n)}=G$ in $\mathcal{H}_T(M)$. Moreover, we can assume that $\|G^{(n)}\|_\infty\leqslant C$ (if $G^{(n)}$ does not satisfy this bound, take $(G^{(n)}\wedge C)\vee(-C)$, which still converges to G in $\mathcal{H}_T(M)$). Note that

$$\begin{aligned}\|GH-G^{(n)}H\|_{\mathcal{H}_T(N)}^2 &= \mathbb{E}\int_0^T\big(G_s\,H_s-G_s^{(n)}H_s\big)^2\,\mathrm{d}\langle N\rangle_s\\ &= \mathbb{E}\int_0^T\big(G_s-G_s^{(n)}\big)^2\,H_s^2\,\mathrm{d}\langle N\rangle_s\\ &= \mathbb{E}\int_0^T\big(G_s-G_s^{(n)}\big)^2\,\mathrm{d}\langle M\rangle_s=\|G_s-G_s^{(n)}\|_{\mathcal{H}_T(M)}^2\to 0,\end{aligned}$$

where we used the identity $\langle M\rangle_t=\int_0^t H_s^2\,\mathrm{d}\langle N\rangle_s$, which follows from Theorem 10.3 (f).

Hence, $\lim_{n\to\infty}G^{(n)}H=GH$ in $\mathcal{H}_T(N)$. As a result, we have

$$\int_0^t G_s\,H_s\,\mathrm{d}N_s\xleftarrow{L^2(\Omega)}\int_0^t G_s^{(n)}H_s\,\mathrm{d}N_s=\int G_s^{(n)}\,\mathrm{d}M_s\xrightarrow{L^2(\Omega)}\int_0^t G_s\,\mathrm{d}M_s.$$

b) First, note that

$$\int_0^t G_0\,\mathrm{d}M_s=G_0\int_0^t \mathrm{d}M_s=\int_0^t H_s\,\mathrm{d}N_s=\int_0^t G_0\,H_s\,\mathrm{d}N_s.$$

Hence, by considering $G-G_0$ instead of G we can assume without loss of generality that $G_0=0$. Let $\tau_n\nearrow T$ be such that G^{τ_n} is bounded, $N^{\tau_n}\in\mathcal{M}_T$, and $H\mathbf{1}_{[0,\tau_n]}\in\mathcal{H}_T(N^{\tau_n})$. Then by Theorem 9.17 we get

$$M_t^{\tau_n}=\int_0^{t\wedge\tau_n}H_s\,\mathrm{d}N_s=\int\mathbf{1}_{[0,\tau_n]}(s)\,H_s\,\mathrm{d}N_s^{\tau_n}.$$

Therefore, by Theorem 9.17 and part a) we have

$$\begin{aligned}\int_0^{t\wedge t_n} G_s\,\mathrm{d}M_s &= \int_0^t \mathbf{1}_{[0,\tau_n]}(s)\,G_s\,\mathrm{d}M_s^{\tau_n}\\ &= \int_0^t \mathbf{1}_{[0,\tau_n]}(s)\,G_s\,\mathbf{1}_{[0,\tau_n]}(s)\,H_s\,\mathrm{d}N_s^{\tau_n}\\ &= \int_0^t \mathbf{1}_{[0,\tau_n]}(s)\,G_s H_s\,\mathrm{d}N_s^{\tau_n} = \int_0^{t\wedge t_n} G_s H_s\,\mathrm{d}N_s.\end{aligned}$$

We complete the proof by taking the limit $n\to\infty$. $\square$

12 Continuous semimartingales

Definition 12.1. *Let $\mathcal{A}_T$ denote the space of continuous, adapted processes $A=(A_t)_{t\in[0,T]}$ of bounded variation such that $A_0=0$. Recall that $\mathcal{M}_{T,\mathrm{loc}}$ denotes the space of continuous local martingales $M=(M_t)_{t\in[0,T]}$ such that $M_0=0$. A process $Z=(Z_t)_{t\in[0,T]}$ is called a **continuous semimartingale** if*

$$Z=Z_0+M+A,$$

where Z_0 is an $\mathcal{F}_0$-measurable random variable, $M\in\mathcal{M}_{T,\mathrm{loc}}$ and $A\in\mathcal{A}_T$.

Remark 12.2. The decomposition of a semimartingale is unique (up to indistinguishability). If $Z=Z_0+M+A=Z_0+M'+A'$, then $M-M'=A'-A$ is a continuous local martingale starting from zero with bounded variation on $[0,t]$ so it is identically zero.

Example 12.3. An Itô process, i.e., a process of the form $Z_t=Z_0+\int_0^t H_s\,\mathrm{d}B_s+\int_0^t Y_s\,\mathrm{d}s$, is a semimartingale.

Example 12.4. Let N be a square integrable martingale. Then by Theorem 9.1 $M=N^2-\langle N\rangle$ is a martingale. Hence, $N^2=N_0^2+(M-N_0^2)+\langle N\rangle$ is a semimartingale.

Definition 12.5. *Let $Z=Z_0+M+A$ be a continuous semimartingale and $\mathcal{H}_{T,\mathrm{loc}}(Z)$ denote the space consisting of $H\in\mathcal{H}_{T,\mathrm{loc}}(M)$ such that $H\in L^1([0,T],|\mu_A|)$ a.s. The stochastic integral of $H\in\mathcal{H}_{T,\mathrm{loc}}(Z)$ with respect to Z is a continuous semimartingale $(H\cdot Z)$ defined by*

$$(H\cdot Z)_t\equiv\int_0^t H_s\,\mathrm{d}Z_s:=\int_0^t H_s\,\mathrm{d}M_s+\int_0^t H_s\,\mathrm{d}A_s,$$

where the first integral is a stochastic integral and the second is an integral with respect to a process of bounded variation.

Theorem 12.6. (Integration by substitution) *Let $Z=Z_0+M+A$ be a continuous semimartingale, $H\in\mathcal{H}_{T,\mathrm{loc}}(Z)$, G be a locally bounded predictable process and $Z'=(H\cdot Z)$. Then $G\in\mathcal{H}_{T,\mathrm{loc}}(Z')$, $HG\in\mathcal{H}_{T,\mathrm{loc}}(Z)$ and $(GH\cdot Z)=(G\cdot Z')$. Equivalently,*

$$\int_0^t G_s\,H_s\,\mathrm{d}Z_s=\int_0^t G_s\,\mathrm{d}Z_s',\qquad Z_t'=\int_0^t H_s\,\mathrm{d}Z_s.$$

Proof. We use the integration by substitution theorem for the stochastic integral and a similar result for the integral with respect to a process of bounded variation. $\square$

Theorem 12.7. (Integration by parts) *If $Z = Z_0 + M + A$ and $Z' = Z_0' + M' + A'$ are continuous semimartingales, then ZZ' is also a semimartingale and*

$$Z_t Z_t' = Z_0 Z_0' + \int_0^t Z_s \, \mathrm{d}Z_s' + \int_0^t Z_s' \, \mathrm{d}Z_s + \langle M, M' \rangle_t.$$

Proof. See Exercise 1, Sheet 8 for the proof for $Z = Z_0 + M$ and $Z' = Z_0' + M'$. □

Remark 12.8. For example, for a Brownian motion B we have

$$\int_0^t B_s \mathrm{d}B_s = \frac{1}{2}(B_t^2 - B_0^2) - \frac{1}{2}\langle B \rangle_t = \frac{1}{2}(B_t^2 - t).$$

On the other hand, for any $A \in \mathcal{A}_T$ we have $\int_0^t A_s \mathrm{d}A_s = \frac{1}{2}(A_t^2 - A_0^2)$. Moreover, if $H, G \in \mathcal{H}_{T,\mathrm{loc}}$, $M_t = \int_0^t H_s \, \mathrm{d}B_s$ and $N_t = \int_0^t G_s \, \mathrm{d}B_s$, then

$$M_t N_t = \int_0^t M_s \, \mathrm{d}N_s + \int_0^t N_s \, \mathrm{d}M_s + \langle M, N \rangle_t = \int_0^t M_s G_s \, \mathrm{d}B_s + \int_0^t N_s H_s \, \mathrm{d}B_s + \int_0^t H_s G_s \, \mathrm{d}s.$$

For $M \in \mathcal{M}_{T,\mathrm{loc}}$ and $A \in \mathcal{A}_T$ we have

$$M_t A_t = \int_0^t A_s \, \mathrm{d}M_s + \int_0^t M_s \, \mathrm{d}A_s$$

and for $A, K \in \mathcal{A}_T$ we have

$$A_t K_t = \int_0^t A_s \, \mathrm{d}K_s + \int_0^t K_s \, \mathrm{d}A_s.$$

The last integration by parts formula is a straightforward consequence of the definition of the Riemann-Stieltjes integral.

13 Itô formula

Computing the stochastic integral by following its definition is typically very cumbersome. This is similar to the usual approach used for the ordinary Riemann or Lebesgue integrals, where the integral is initially defined through approximations with step functions, but more efficient and intuitive computational techniques are later introduced. In this lecture, we will prove a fundamental theorem for stochastic analysis. It shows that the class of continuous semimartingales is closed under smooth functions.

Theorem 13.1. (Itô's formula) *Assume that $Z = Z_0 + M + A$ is a continuous semimartingale and $f \in C^2(\mathbb{R})$. Then $f(Z)$ is also a semimartingale and*

$$f(Z_t) = f(Z_0) + \int_0^t f'(Z_s) \, \mathrm{d}Z_s + \frac{1}{2}\int_0^t f''(Z_s) \, \mathrm{d}\langle M \rangle_s.$$

Remark 13.2. Let $(B_t)_{t \geqslant 0}$ be a Brownian motion and $f \in C^2(\mathbb{R})$. Since $\langle B \rangle_t = t$, for every $t \geqslant 0$ we have

$$f(B_t) - f(B_0) = \int_0^t f'(B_s) \, \mathrm{d}B_s + \frac{1}{2}\int_0^t f''(B_s) \, \mathrm{d}s,$$

where $\int_0^t f'(B_s) \mathrm{d}B_s$ is the stochastic integral and $\int f''(B_s) \mathrm{d}s$ is the Lebesgue or Riemann integral.

The term $\frac{1}{2}\int_0^t f''(B_s)\,\mathrm{d}s$ appears because the quadratic variation of the Brownian motion is not zero. Note that if A is of bounded variation, then

$$f(A_t) - f(A_0) = \int_0^t f'(A_s)\,\mathrm{d}A_s.$$

Example 13.3. Let $f(x) = \mathrm{e}^x$. Then Itô's formula gives

$$\mathrm{e}^{B_t} - \mathrm{e}^{B_0} = \int_0^t \mathrm{e}^{B_s}\,\mathrm{d}B_s + \frac{1}{2}\int_0^t \mathrm{e}^{B_s}\,\mathrm{d}s$$

Let $X_t = \mathrm{e}^{B_t}$, then $X_t = 1 + \int_0^t X_s\,\mathrm{d}B_s + \frac{1}{2}\int_0^t X_s\,\mathrm{d}s$, or equivalently

$$\begin{cases} \mathrm{d}X_t = X_t\,\mathrm{d}B_t + \frac{1}{2}X_t\,\mathrm{d}t \\ X_0 = 1. \end{cases}$$

Proof. The integrals in (13.1) are well-defined because the processes $f'(Z_s)$ and $f''(Z_s)$ are continuous, $f'(Z_s) \in \mathcal{H}_{T,\mathrm{loc}}(M)$ and $f''(Z_s) \in L^1([0,T], \langle M \rangle)$.

Step 1. Let Z be a bounded semimartingale and f be a polynomial. By linearity of both sides of (13.1), it suffices to consider the case when $f(x) = x^n$. We will show this formula by induction on n. For $n = 0$, the thesis is obvious. Assume that (13.1) holds for $f(x) = x^n$, we will show it for $g(x) = x\,f(x)$.

By induction hypothesis $f(Z_t)$ is a semimartingale with the decomposition

$$\begin{aligned} f(Z_t) &= f(Z_0) + \int_0^t f'(Z_s)\,\mathrm{d}Z_s + \frac{1}{2}\int_0^t f''(Z_s)\,\mathrm{d}\langle M \rangle_s \\ &= f(Z_0) + \int_0^t f'(Z_s)\,\mathrm{d}M_s + \left(\int_0^t f'(Z_s)\,\mathrm{d}A_s + \frac{1}{2}\int_0^t f''(Z_s)\,\mathrm{d}\langle M \rangle_s\right). \end{aligned}$$

Using integration by parts we obtain

$$g(Z_t) = Z_t\,f(Z_t) = Z_0\,f(Z_0) + \int_0^t Z_s\,\mathrm{d}f(Z_s) + \int_0^t f(Z_s)\,\mathrm{d}Z_s + \left\langle \int f'(Z_s)\,\mathrm{d}M_s, M \right\rangle_t.$$

Using integration by substitution we obtain

$$\int_0^t Z_s\,\mathrm{d}f(Z_s) = \int_0^t \left(Z_s\,f'(Z_s)\,\mathrm{d}Z_s + \frac{1}{2} Z_s\,f''(Z_s)\ \mathrm{d}\langle M \rangle_s \right).$$

By Theorem 10.3 (f) we have

$$\left\langle \int f'(Z_s)\,\mathrm{d}M_s, M \right\rangle_t = \int_0^t f'(Z_s)\ \mathrm{d}\langle M \rangle_s\,.$$

Applying the above identities and noting that $g'(x) = f(x) + x\,f'(x)$ and $g''(x) = 2\,f'(x) + x\,f''(x)$ we arrive at

$$g(Z_t) = g(Z_0) + \int_0^t g'(Z_t)\,\mathrm{d}Z_t + \frac{1}{2}\int_0^t g''(Z_t)\ \mathrm{d}\langle M \rangle_s.$$

This proves the induction step.

Step 2. Let Z be a bounded semimartingale and $f\in C^2(\mathbb{R})$ be arbitrary. Denote by $C>0$ a deterministic constant such that $|Z_t|\leqslant C$. There exists a sequence of polynomials f_n such that

$$|f_n(x)-f(x)|,|f_n'(x)-f'(x)|,|f_n''(x)-f''(x)|\leqslant\frac{1}{n}\quad\text{for } x\in[-C,C].$$

Moreover, letting $K=\sup_{x\in[-C,C]}(|f'(x)|\vee|f''(x)|)$, we have

$$|f_n'(Z_s)|\leqslant\sup_{x\in[-C,C]}|f_n'(x)|\leqslant 1+K,\qquad |f_n''(Z_s)|\leqslant\sup_{x\in[-C,C]}|f_n''(x)|\leqslant 1+K.$$

Hence, from Lebesgue's dominated convergence theorem (for ordinary and stochastic integrals),

$$\begin{aligned}f(Z_t)&=\lim_{n\to\infty}f_n(Z_t)\\&=\lim_{n\to\infty}\left(f_n(Z_0)+\int_0^t f_n'(Z_s)\,\mathrm{d}Z_s+\frac{1}{2}\int_0^t f_n''(Z_s)\,\mathrm{d}\langle M\rangle_s\right)\\&=f(Z_0)+\int_0^t f'(Z_s)\,\mathrm{d}Z_s+\frac{1}{2}\int_0^t f''(Z_s)\,\mathrm{d}\langle M\rangle_s.\end{aligned}$$

Step 3. Let $Z=Z_0+M+A$ be a continuous semimartingale such that Z_0 is bounded and $f\in C^2(\mathbb{R})$. In this case, we define

$$\tau_n:=\inf\{t>0\,|\,|Z_t|\geqslant n\}\wedge T.$$

Since Z_0 is bounded and Z is continuous, $\lim_{n\to\infty}\tau_n=T$ a.s. Moreover, $Z^{\tau_n}:=Z_0+M^{\tau_n}+A^{\tau_n}$ is a continuous bounded semimartingale and $\lim_{n\to\infty}Z_t^{\tau_n}=Z_t$ a.s. By Step 2, (13.1) holds for $Z^{(n)}$ and by Theorem 9.17 and Theorem 10.3 (e) we have

$$\begin{aligned}f(Z_t^{\tau_n})&=f(Z_0)+\int_0^t f'(Z_s^{\tau_n})\,\mathrm{d}Z_s^{\tau_n}+\frac{1}{2}\int_0^t f''(Z_s^{\tau_n})\,\mathrm{d}\langle M^{\tau_n}\rangle_s\\&=f(Z_0)+\int_0^t\mathbf{1}_{[0,\tau_n]}(s)f'(Z_s^{\tau_n})\,\mathrm{d}Z_s+\frac{1}{2}\int_0^t\mathbf{1}_{[0,\tau_n]}(s)f''(Z_s^{\tau_n})\,\mathrm{d}\langle M\rangle_s\\&=f(Z_0)+\int_0^t\mathbf{1}_{[0,\tau_n]}(s)f'(Z_s)\,\mathrm{d}Z_s+\frac{1}{2}\int_0^t\mathbf{1}_{[0,\tau_n]}(s)f''(Z_s)\,\mathrm{d}\langle M\rangle_s\\&=f(Z_0)+\int_0^{t\wedge\tau_n}f'(Z_s)\,\mathrm{d}Z_s+\frac{1}{2}\int_0^{t\wedge\tau_n}f''(Z_s)\,\mathrm{d}\langle M\rangle_s.\end{aligned}$$

Taking the limit $n\to\infty$ we obtained (13.1).

Step 4. To prove (13.1) in the general case let $Z_0^{(n)}:=(Z_0\wedge n)\vee(-n)$ and $Z^{(n)}:=Z_0^{(n)}+M+A$. Note that $\int_0^t X_s\mathrm{d}Z_s=\int_0^t X_s\mathrm{d}Z_s^{(n)}$. Since we already know that (13.1) holds when Z_0 is bounded,

$$f(Z_t^{(n)})=f(Z_0^{(n)})+\int_0^t f'(Z_s^{(n)})\,\mathrm{d}Z_s+\frac{1}{2}\int_0^t f''(Z_s^{(n)})\,\mathrm{d}\langle M\rangle_s.$$

We shall prove convergence as $n\to\infty$ of every term appearing in the above equation. It is evident that $\lim_{n\to\infty}f(Z_t^{(n)})=f(Z_t)$ a.s (and similarly for f' and f''). We observe that

$$\left|f'(Z_s^{(n)})\right|\leqslant\sup_{n\in\mathbb{N}_+}\left|f'(Z_s^{(n)})\right|=:Y_s.$$

The process Y is predictable as the supremum of predictable processes, and

$$\sup_{s\leqslant t}|Z_s^{(n)}|\leqslant|Z_0|+\sup_{s\leqslant t}|M_s|+\sup_{s\leqslant t}|A_s|<\infty\quad a.s.$$

Hence, from the continuity of f' we infer that $\sup_{s\leqslant t}|Y_s|<\infty$ a.s. and $Y\in\mathcal{H}_{T,\mathrm{loc}}(M)$. By the dominated convergence theorem for stochastic integrals,

$$\lim_{n\to\infty}\int_0^t f'(Z_s^{(n)})\,\mathrm{d}M_s=\int_0^t f'(Z_s)\,\mathrm{d}M_s\quad\text{in probability.}$$

Moreover, by the dominated convergence theorem for ordinary integrals,

$$\lim_{n\to\infty}\int_0^t f'(Z_s^{(n)})\,\mathrm{d}A_s=\int_0^t f'(Z_s)\,\mathrm{d}A_s\quad a.s.$$

Similarly, $\sup_n\sup_{s\leqslant t}|f''(Z_s^{(n)})|<\infty$ a.s., and again applying the dominated convergence theorem for ordinary integrals, we get

$$\lim_{n\to\infty}\int_0^t f''(Z_s^{(n)})\,\mathrm{d}\langle M\rangle_s=\int_0^t f''(Z_s)\,\mathrm{d}\langle M\rangle_s\quad\text{a.s.}$$

To complete the proof we pass to the limit $n\to\infty$ in (13.3). □

In a similar way as in the one-dimensional case, we can prove the multidimensional version of Itô's theorem.

Theorem 13.4. *Assume that $f:\mathbb{R}_\geqslant\times\mathbb{R}^\mathrm{d}\to\mathbb{R}$ is a function that is C^1 in $\mathbb{R}_\geqslant$ and C^2 in $\mathbb{R}^d$ and $Z=(Z^1,\ldots,Z^d)$, where $Z^i=Z_0^i+M^i+A^i$ are continuous semimartingales for $i\in\{1,\ldots,d\}$. Then $(f(t,Z_t))_{t\geqslant 0}$ is a semimartingale and*

$$\begin{aligned} f(t,Z_t) &= f(0,Z_0)+\int_0^t\frac{\partial f}{\partial s}(s,Z_s)\,\mathrm{d}s+\sum_{i=1}^d\int_0^t\frac{\partial f}{\partial x^i}(s,Z_s)\,\mathrm{d}Z_s^i\\ &\quad+\frac{1}{2}\sum_{i,j=1}^d\int_0^t\frac{\partial^2 f}{\partial x^i\partial x^j}(s,Z_s)\,\mathrm{d}\langle M^i,M^j\rangle_s.\end{aligned}$$

Example 13.5. Let B be a Brownian motion. The application of Itô's formula yields

$$\begin{aligned} X_t &= \exp\Big(B_t-\frac{t}{2}\Big)\\ &= \exp(B_0)-\frac{1}{2}\int_0^t\exp\Big(B_s-\frac{s}{2}\Big)\,\mathrm{d}s+\int_0^t\exp\Big(B_s-\frac{s}{2}\Big)\,\mathrm{d}B_s+\frac{1}{2}\int_0^t\exp\Big(B_s-\frac{s}{2}\Big)\,\mathrm{d}s\\ &= 1+\int_0^t\exp\Big(B_s-\frac{s}{2}\Big)\,\mathrm{d}B_s=\int_0^t X_s\,\mathrm{d}B_s.\end{aligned}$$

That is, X satisfies the following stochastic differential equation

$$\begin{cases}\mathrm{d}X_t=X_t\,\mathrm{d}B_t\\ X_0=1.\end{cases}$$

Note that for all $T>0$ we have

$$\mathbb{E}\Big(\int_0^T X_s^2\mathrm{d}s\Big)=\int_0^T\mathbb{E}(\exp(2B_s-s))\,\mathrm{d}s<\infty.$$

Hence, $X\in\mathcal{H}_T$ and since $X=\int X_s\,\mathrm{d}B_s$ we conclude that $X\in\mathcal{M}_T$ is a square integrable martingale. We call X the **exponential martingale** associated with B.

Example 13.6. Let $B=(B^1,\ldots,B^d)$ be a vector of d independent Brownian motions and $f\in C^2(\mathbb{R}^d)$. Note that for $i\neq j$ we have

$$\mathbb{E}(B_t^iB_t^j|\mathcal{F}_s)=\mathbb{E}(B_t^i|\mathcal{F}_s)\mathbb{E}(B_t^j|\mathcal{F}_s)=B_s^iB_s^j.$$

Thus, the process $(B_t^iB_t^j-t\,\delta_{ij})_{t\geqslant 0}$ is a martingale. This proves that $\langle B^i,B^j\rangle_t=t\,\delta_{ij}$ and by multidimensional Itô's formula

$$f(B_t)=f(B_0)+\sum_{i=1}^d\int_0^t\frac{\partial f}{\partial x_i}(B_s)\,\mathrm{d}B_s^i+\frac{1}{2}\int_0^t(\Delta f)(B_s)\,\mathrm{d}s,$$

where Δ is the Laplace operator.

14 Stochastic differential equations

Definition 14.1. *Let $(\Omega,\mathcal{F},P)$ be a probability space, $(\mathcal{F}_t)_{t\geqslant 0}$ a filtration such that $\mathcal{F}_0$ contains all null events and $(B_t)_{t\geqslant 0}$ a Brownian motion adapted to $(\mathcal{F}_t)_{t\geqslant 0}$ such that $(B_{t+s}-B_t)_{s\geqslant 0}$ is independent of $\mathcal{F}_t$ for all $t\geqslant 0$. Assume that $b,\sigma\colon[0,\infty)\times\mathbb{R}\to\mathbb{R}$ are continuous functions, and x is an $\mathcal{F}_0$-measurable random variable. We say that the process $X=(X_t)_{t\in[0,T]}$ is a **strong solution** of the **stochastic differential equation** (SDE)*

$$\mathrm{d}X_t=b(t,X_t)\,\mathrm{d}t+\sigma(t,X_t)\,\mathrm{d}B_t,\quad X_0=x, \tag{14.1}$$

if X is a continuous and adapted process such that

$$X_t=x+\int_0^tb(s,X_s)\,\mathrm{d}s+\int_0^t\sigma(s,X_s)\,\mathrm{d}B_s,\quad t\in[0,T]. \tag{14.2}$$

Remark 14.2. The assumption that b and σ are continuous functions is not necessary and it is possible to study more general stochastic differential equations. Note that the continuity of b and σ automatically implies the measurability and local boundedness of the processes $b(s,X_s)$ and $\sigma(s,X_s)$, which guarantees that the integrals appearing in (14.2) are well defined.

Remark 14.3. The process X solving equation (14.1) is called a **diffusion** with the **diffusion coefficient** σ and the **drift coefficient** b.

Remark 14.4. Note that the Itô formula can be equivalently written as an SDE

$$\mathrm{d}f(t,Z_t)=\frac{\partial f}{\partial t}(t,Z_t)\,\mathrm{d}t+\sum_{i=1}^d\frac{\partial f}{\partial x^i}(t,Z_t)\,\mathrm{d}Z_t^i+\frac{1}{2}\sum_{i,j=1}^d\frac{\partial^2 f}{\partial x^i\partial x^j}(t,Z_t)\,\mathrm{d}\langle M^i,M^j\rangle_t.$$

Example 14.5. The Black-Scholes SDE

$$\mathrm{d}X_t=\mu X_t\,\mathrm{d}t+\sigma X_t\,\mathrm{d}B_t$$

models the dynamics of a financial asset, such as a stock, under the assumptions of continuous trading and no arbitrage.

- X_t is the price of the asset at time t.
- $\mu\in\mathbb{R}$ is the drift coefficient (representing the expected return rate).
- $\sigma>0$ is the volatility coefficient (measuring the asset's randomness or risk).

Note that in financial markets it is natural to assume that the changes in price are proportional to the current price.

Assumption 14.6. *The functions b and σ (i) have sublinear growth at infinity and (ii) are globally Lipschitz, that is,*

$$\text{(i)}. \; |b(t,x)| \leqslant L(1+|x|), \qquad |\sigma(t,x)| \leqslant L(1+|x|),$$

$$\text{(ii)}. \; |b(t,x)-b(t,y)| \leqslant L|x-y|, \qquad |\sigma(t,x)-\sigma(t,y)| \leqslant L|x-y|$$

for all $x, y \in \mathbb{R}$, $t \in [0,T]$. In addition, (iii) x is a square-integrable random variable.

Example 14.7. Let $b(t,x) = 2\sqrt{x}$ and $\sigma(t,x) = 0$. The function b is not Lipschitz continuous at $x = 0$. The uniqueness of solutions fails since the functions

$$X_t = 0 \quad \text{and} \quad X_t = t^2$$

are both solutions of $\mathrm{d}X_t = 2\sqrt{|X_t|}\mathrm{d}t$ with $X_0 = 0$.

Example 14.8. Let $b(t,x) = x^2$ and $\sigma(t,x) = 0$. The function b is smooth but it is not globally Lipschitz and grows faster than linearly. For any initial value $x > 0$, the function

$$X_t = \frac{1}{\frac{1}{x} - t}$$

solves $\mathrm{d}X_t = X_t^2\,\mathrm{d}t$ with $X_0 = x$ and it can be shown that there is no other solution. However, the function X_t blows up as $t \to 1/x$, so the solution does not exist for all time $t > 0$.

Definition 14.9. *For continuous adapted processes X the process $M(X) = (M_t(X))_{t\in[0,T]}$ is defined by*

$$M_t(X) := x + \int_0^t b(s, X_s)\,\mathrm{d}s + \int_0^t \sigma(s, X_s)\,\mathrm{d}B_s.$$

Observe that a process X is a solution of the stochastic differential equation (14.1) if and only if it is a fixed point of the map M introduced above, that is, $M(X) = X$.

Definition 14.10. *Let $\mathcal{S}_T$ denote the Banach space of continuous and adapted processes $(X_t)_{t\in[0,T]}$ such that*

$$\|X\|_{\mathcal{S}_T} := \left(\mathbb{E}\left[\sup_{s\in[0,T]} X_s^2\right]\right)^{1/2} < \infty.$$

Remark 14.11. To prove that the space $\mathcal{S}_T$ is complete one uses the fact that uniform convergence preserves continuity.

Lemma 14.12. *The map $M: \mathcal{S}_T \to \mathcal{S}_T$ is well-defined. Moreover, for all $t \in [0,T]$ and $X, Y \in \mathcal{S}_T$ we have*

$$\|M(X) - M(Y)\|_{\mathcal{S}_t}^2 \leqslant C \int_0^t \|X - Y\|_{\mathcal{S}_u}^2\,\mathrm{d}u$$

with $C = 2L^2(T+4)$.

Proof. Using the elementary inequality $(a+b)^2 \leqslant 2a^2+2b^2$ we obtain

$$\begin{aligned}(M_s(X)-M_s(Y))^2 &= \left(\int_0^s (b(r,X_r)-b(r,Y_r))\,\mathrm{d}r + \int_0^s (\sigma(r,X_r)-\sigma(r,Y_r))\,\mathrm{d}B_r\right)^2 \\ &\leqslant 2\left(\int_0^s (b(r,X_r)-b(r,Y_r))\,\mathrm{d}r\right)^2 + 2\left(\int_0^s (\sigma(r,X_r)-\sigma(r,Y_r))\,\mathrm{d}B_r\right)^2.\end{aligned}$$

Using the Cauchy-Schwartz inequality and Assumption 14.6 (ii) we estimate

$$\begin{aligned}\mathbb{E}\left[\sup_{s\in[0,t]}\left(\int_0^s (b(r,X_r)-b(r,Y_r))\,\mathrm{d}r\right)^2\right] &\leqslant t\,\mathbb{E}\left(\int_0^t |b(r,X_r)-b(r,Y_r)|^2\,\mathrm{d}r\right) \\ &\leqslant t L^2\,\mathbb{E}\left(\int_0^t |X_r-Y_r|^2\,\mathrm{d}r\right).\end{aligned}$$

Similarly, by the Doob inequality stated in Theorem 6.14 (b), the Itô isometry and Assumption 14.6 (ii) we obtain

$$\begin{aligned}\mathbb{E}\left[\sup_{s\in[0,t]}\left(\int_0^s (\sigma(r,X_r)-\sigma(r,Y_r))\,\mathrm{d}B_r\right)^2\right] &\leqslant 4\,\mathbb{E}\left(\int_0^t (\sigma(r,X_r)-\sigma(r,Y_r))\,\mathrm{d}B_r\right)^2 \\ &= 4\,\mathbb{E}\left(\int_0^t (\sigma(r,X_r)-\sigma(r,Y_r))^2\,\mathrm{d}r\right) \\ &\leqslant 4L^2\,\mathbb{E}\left(\int_0^t |X_r-Y_r|^2\,\mathrm{d}r\right).\end{aligned}$$

This proves

$$\mathbb{E}\left[\sup_{s\in[0,t]}(M_s(X)-M_s(Y))^2\right] \leqslant C\int_0^t \mathbb{E}\left[\sup_{s\in[0,u]}(X_s-Y_s)^2\right]\mathrm{d}u,$$

which implies the desired bound. ☐

Lemma 14.13. *In the setting of the above lemma we have*

$$A_t^{(n)} := \Bigg\|\underbrace{M\circ\ldots\circ M}_{n}(X) - \underbrace{M\circ\ldots\circ M}_{n}(Y)\Bigg\|_{\mathcal{S}_t}^2 \leqslant \frac{C^n t^n}{n!}\,\|X-Y\|_{\mathcal{S}_t}^2.$$

Proof. By Lemma 14.12 it holds that

$$\begin{aligned}A_{t_n}^{(n)} &\leqslant C\int_0^{t_n} A_{t_{n-1}}^{(n-1)}\,\mathrm{d}t_{n-1} \leqslant C^2\int_0^{t_n}\int_0^{t_{n-1}} A_{t_{n-2}}^{(n-2)}\,\mathrm{d}t_{n-2}\,\mathrm{d}t_{n-1} \\ &\leqslant C^n\int_0^{t_n}\int_0^{t_{n-1}}\ldots\int_0^{t_1} A_{t_0}^{(0)}\,\mathrm{d}t_0\ldots\mathrm{d}t_{n-2}\,\mathrm{d}t_{n-1} \\ &\leqslant A_{t_n}^{(0)} C^n\int_0^{t_n}\int_0^{t_{n-1}}\ldots\int_0^{t_1}\mathrm{d}t_0\ldots\mathrm{d}t_{n-2}\,\mathrm{d}t_{n-1} = A_{t_n}^{(0)}\frac{C^n\,t_n^n}{n!},\end{aligned}$$

which proves the claim. ☐

Theorem 14.14. *There exists exactly one (up to indistinguishability) solution X of the stochastic differential equation (14.1). Moreover, $X\in\mathcal{S}_T$.*

Proof of existence. We use Picard's iteration method, that is, we define recursively

$$X^{(0)} = 0, \qquad X^{(n+1)} = M(X^{(n)}), \quad n \in \mathbb{N}_0.$$

Note that

$$X_t^{(1)} = M_t(0) = x + \int_0^t b(s,0)\,\mathrm{d}s + \int_0^t \sigma(s,0)\,\mathrm{d}B_s.$$

Hence, by $(a+b+c)^2 \leqslant 3a^2 + 3b^2 + 3c^2$, the Doob inequality, the Itô isometry and Assumption 14.6 (i) and (iii) we obtain

$$\begin{aligned}\|X^{(1)} - X^{(0)}\|_{\mathcal{S}_T}^2 &\leqslant 3\mathbb{E}x^2 + 3\mathbb{E}\left[\sup_{t\in[0,T]}\left(\int_0^t b(s,0)\,\mathrm{d}s\right)^2\right] + 3\mathbb{E}\left[\sup_{t\in[0,T]}\left(\int_0^t \sigma(s,0)\,\mathrm{d}B_s\right)^2\right]\\ &\leqslant 3\mathbb{E}x^2 + 3L^2T^2 + 12\,\mathbb{E}\left(\int_0^T \sigma(s,0)^2\,\mathrm{d}s\right) \leqslant 3\mathbb{E}x^2 + 3L^2T(T+4) < \infty.\end{aligned}$$

By Lemma 14.13 we have

$$\sum_{n=0}^{\infty} \|X^{(n+1)} - X^{(n)}\|_{\mathcal{S}_T} \leqslant \|X^{(1)} - X^{(0)}\|_{\mathcal{S}_T} \sum_{n=0}^{\infty} \left(\frac{C^nT^n}{n!}\right)^{1/2} < \infty.$$

Hence, the sequence $(X^{(n)})_{n\in\mathbb{N}_0}$ is Cauchy in the norm $\|\cdot\|_{\mathcal{S}_T}$. We denote by $X \in \mathcal{S}_T$ its limit. By Lemma 14.12 we have

$$\|M(X) - M(Y)\|_{\mathcal{S}_T}^2 \leqslant CT\,\|X - Y\|_{\mathcal{S}_T}^2,$$

which implies that the map M is continuous. Hence,

$$M(X) = \lim_{n\to\infty} M(X^{(n)}) = \lim_{n\to\infty} X^{(n+1)} = X.$$

Consequently, $X \in \mathcal{S}_T$ is a solution of (14.1). □

Remark 14.15. If $X, Y \in \mathcal{S}_T$ are such that $M(X) = X$ and $M(Y) = Y$, then by Lemma 14.13 we have

$$\|X - Y\|_{\mathcal{S}_T}^2 \leqslant \frac{C^nT^n}{n!}\,\|X - Y\|_{\mathcal{S}_T}^2.$$

Taking $n \in \mathbb{N}_+$ big enough we conclude $\|X - Y\|_{\mathcal{S}_T}^2 \leqslant 0$, which implies that $X = Y$. This proves uniqueness of solutions of (14.1) in $\mathcal{S}_T$. However, the uniqueness claimed in Theorem 14.14 is more general.

Proof of uniqueness. Suppose that X, Y are continuous adapted processes solving (14.1). Define

$$\tau_m := \inf\{t \geqslant 0 \mid |X_t - x| \vee |Y_t - x| \geqslant m\}, \qquad m \in \mathbb{N}_+.$$

Note that $X^{\tau_m}, Y^{\tau_m} \in \mathcal{S}_T$ since $|X_t^{\tau_m} - x| \vee |Y_t^{\tau_m} - x| \leqslant m$ and $\mathbb{E}x^2 < \infty$ by Assumption 14.6 (iii). Define the map

$$M_t^{(m)}(X) := x + \int_0^t \mathbf{1}_{[0,\tau_m]}(s)\,b(s, X_s)\,\mathrm{d}s + \int_0^t \mathbf{1}_{[0,\tau_m]}(s)\,\sigma(s, X_s)\,\mathrm{d}B_s.$$

Using Theorem 9.17 and the fact that X and Y are solutions of (14.1) we obtain

$$\begin{aligned} M_t^{(m)}(X^{\tau_m}) &= x + \int_0^t \mathbf{1}_{[0,\tau_m]}(s)\, b(s, X_s)\, \mathrm{d}s + \int_0^t \mathbf{1}_{[0,\tau_m]}(s)\, \sigma(s, X_s)\, \mathrm{d}B_s \\ &= x + \int_0^{t\wedge\tau_m} b(s, X_s)\, \mathrm{d}s + \int_0^{t\wedge\tau_m} \sigma(s, X_s)\, \mathrm{d}B_s = X_t^{\tau_m} \end{aligned}$$

and the same for Y^{τ_m}. We conclude that $M^{(m)}(X^{\tau_m}) = X^{\tau_m}$ and $M^{(m)}(Y^{\tau_m}) = Y^{\tau_m}$. The map $M^{(m)}$ satisfies the bounds stated in Lemmas 14.12 and 14.13 as can be easily checked by following the proofs of these lemmas. Hence, by the argument from Remark 14.15 we have $X^{\tau_m} = Y^{\tau_m}$ for all $m \in \mathbb{N}_+$, which implies that $X = Y$. □

15 Girsanov theorem

We assume that $(\Omega, \mathcal{F}, \mathbb{P})$ is a fixed probability space. We will construct other probability measure on $(\Omega, \mathcal{F})$ under which a Brownian motion with drift has the same distribution as a standard Brownian motion.

Notation 15.1. *By $\mathbb{E}X$ we always mean the expectation with respect to $\mathbb{P}$, while the expectation of X with respect to another measure $\mathbb{Q}$ will be denoted by $\mathbb{E}_{\mathbb{Q}}X$. Note that if $\mathbb{Q}(A) = \mathbb{E}(\mathbb{1}_A Z)$ for some $Z \geqslant 0$ such that $\mathbb{E}Z = 1$, then $\mathbb{E}_{\mathbb{Q}}X = \mathbb{E}(XZ)$.*

We begin with the following motivating example.

Example 15.2. Let $X_1, X_2, \ldots, X_n$ be independent $\mathcal{N}(0,1)$-random variables and let $\mu_1, \ldots, \mu_n \in \mathbb{R}$ be deterministic. We define a new measure $\mathbb{Q}_n$ on $(\Omega, \mathcal{F})$ by

$$\mathbb{Q}_n(A) = \mathbb{E}(\mathbb{1}_A Z_n) \quad \text{for all} \quad A \in \mathcal{F}, \qquad \text{where} \qquad Z_n := \exp\left(\sum_{i=1}^n \mu_i X_i - \frac{1}{2}\sum_{i=1}^n \mu_i^2\right).$$

Note that

$$\mathbb{Q}_n(\Omega) = \mathbb{E}Z_n = \prod_{i=1}^n \mathbb{E}\left(\exp\left(\mu_i X_i - \frac{1}{2}\mu_i^2\right)\right) = 1.$$

Hence, $\mathbb{Q}_n$ is a probability measure on $(\Omega, \mathcal{F})$. Moreover, for any Borel set $A \in \mathcal{B}(\mathbb{R}^n)$ we have

$$\begin{aligned} \mathbb{Q}_n((X_1, \ldots, X_n) \in A) &= \frac{1}{(2\pi)^{n/2}} \int_A \exp\left(\sum_{i=1}^n \mu_i x_i - \frac{1}{2}\sum_{i=1}^n \mu_i^2\right) \exp\left(-\frac{1}{2}\sum_{i=1}^n x_i^2\right) \mathrm{d}x_1 \ldots \mathrm{d}x_n \\ &= \frac{1}{(2\pi)^{n/2}} \int_A \exp\left(-\frac{1}{2}\sum_{i=1}^n (x_i - \mu_i)^2\right) \mathrm{d}x_1 \ldots \mathrm{d}x_n. \end{aligned}$$

In consequence, with respect to $\mathbb{Q}_n$, $X_i - \mu_i$ are independent $\mathcal{N}(0,1)$-random variables. Defining $S_k = X_1 + \ldots + X_k$, we see that with respect to $\mathbb{Q}_n$, the random variables

$$\left(S_k - \sum_{i=1}^k \mu_i\right)_{k \in \{1, \ldots n\}}$$

are sums of independent standard normal variables, i.e., they have the same distribution as $(S_k)_{k \in \{1, \ldots n\}}$ with respect to $\mathbb{P}$. In what follows, we show a similar fact in the continuous case, where S_k is replaced by a Brownian motion, and the sums $\sum_{i=1}^k \mu_i$ are replaced by the integral $\int_0^t H_s \, \mathrm{d}s$.

Assume that $T<\infty$ and $H\in\mathcal{H}_{T,\mathrm{loc}}$. Then the process $M_t=\int H_s\,\mathrm{d}B_s$ is a local martingale and $\langle M\rangle_t=\int_0^t H_s^2\mathrm{d}s$. Let

$$Z_t:=\exp\biggl(M_t-\frac{1}{2}\langle M\rangle_t\biggr)=\exp\biggl(\int_0^t H_s\,\mathrm{d}B_s-\frac{1}{2}\int_0^t H_s^2\,\mathrm{d}s\biggr).$$

Then the process $Z=(Z_t)_{t\in[0,T]}$ is a local martingale. Actually, the following is true.

Lemma 15.3. *If $M\in\mathcal{M}_{T,\mathrm{loc}}$, then $Z=\bigl(\exp\bigl(M_t-\frac{1}{2}\langle M\rangle_t\bigr)\bigr)_{t\in[0,T]}\in\mathcal{M}_{T,\mathrm{loc}}$.*

Proof. Applying the Itô formula to the semimartingale $Y_t=M_t-\frac{1}{2}\langle M\rangle_t$, we obtain

$$Z_t=Z_0+\int_0^t Z_s\,\mathrm{d}Y_s+\frac{1}{2}\int_0^t Z_s\,\mathrm{d}\langle M\rangle_s=Z_0+\int_0^t Z_s\,\mathrm{d}M_s. \qquad \square$$

Lemma 15.4. *Let $M\in\mathcal{M}_{T,\mathrm{loc}}$. a) If M is bounded, then M is a martingale. b) If M is non-negative, then M is a supermartingale.*

Proof. Let $\tau_n\nearrow T$ be the reducing sequence for M. Fix $0\leqslant s<t\leqslant T$ and $A\in\mathcal{F}_s$.

a) If M is bounded, then by Lebesgue's dominated convergence theorem,

$$\mathbb{E}[M_t\mathbb{1}_A]=\mathbb{E}\Bigl[\lim_{n\to\infty}M_t^{\tau_n}\mathbb{1}_A\Bigr]=\lim_{n\to\infty}\mathbb{E}[M_t^{\tau_n}\mathbb{1}_A]=\lim_{n\to\infty}\mathbb{E}[M_s^{\tau_n}\mathbb{1}_A]=\mathbb{E}[M_s\mathbb{1}_A],$$

hence $M_s=\mathbb{E}(M_t|\mathcal{F}_s)$ and M is a martingale.

b) If M is non-negative, then

$$\begin{aligned}\mathbb{E}[M_s\mathbb{1}_A] &= \lim_{n\to\infty}\mathbb{E}[M_s\mathbb{1}_{A\cap\{\tau_n>s\}}]=\lim_{n\to\infty}\mathbb{E}[M_s^{\tau_n}\mathbb{1}_{A\cap\{\tau_n>s\}}]\\ &= \lim_{n\to\infty}\mathbb{E}[M_t^{\tau_n}\mathbb{1}_{A\cap\{\tau_n>s\}}]\geqslant\mathbb{E}\Bigl[\lim_{n\to\infty}M_t^{\tau_n}\mathbb{1}_{A\cap\{\tau_n>s\}}\Bigr]=\mathbb{E}[M_t\mathbb{1}_A],\end{aligned}$$

where we used Lebesgue's monotone convergence theorem, the fact that M^{τ_n} is a martingale, $A\cap\{\tau_n>s\}\in\mathcal{F}_s$ and Fatou's lemma. Hence, $M_s\geqslant\mathbb{E}(M_t|\mathcal{F}_s)$. $\square$

Lemma 15.5. *Let $M\in\mathcal{M}_{T,\mathrm{loc}}$. The process*

$$Z=(\exp(M_t-\langle M\rangle_t/2))_{t\in[0,T]}$$

is a martingale if and only if $\mathbb{E}Z_T=1$.

Proof. If Z is a martingale, then $\mathbb{E}Z_T=\mathbb{E}Z_0=1$. It remains to prove that if $\mathbb{E}Z_T=1$, then Z is a martingale. Since Z is a non-negative local martingale, by Lemma 15.4, it is a supermartingale. Thus, for all $0\leqslant s\leqslant t\leqslant T$ we have $Z_s\geqslant\mathbb{E}(Z_t|\mathcal{F}_s)$ a.s. Consequently, $1=\mathbb{E}Z_0\geqslant\mathbb{E}Z_t\geqslant\mathbb{E}Z_T=1$ and $\mathbb{E}Z_t=1$. As a result,

$$\mathbb{E}(Z_t-\mathbb{E}(Z_T|\mathcal{F}_t))=\mathbb{E}Z_t-\mathbb{E}Z_T=0,$$

and thus $Z_t=\mathbb{E}(Z_T|\mathcal{F}_t)$ almost surely. $\square$

Proposition 15.6. *Let* $(X_t)_{t\geqslant 0}$ *be a continuous process such that* $X_0 = 0$ *a.s. and for all* $\lambda \in \mathbb{R}$

$$U^\lambda := \left(\mathrm{e}^{\mathrm{i}\lambda X_t + \frac{1}{2}\lambda^2 t}\right)_{t\geqslant 0}$$

is a complex martingale. Then X *is a Brownian motion.*

Proof. The martingale relation

$$\mathbb{E}\left(\mathrm{e}^{\mathrm{i}\lambda X_t + \frac{1}{2}\lambda^2 t}\middle|\mathcal{F}_s\right) = \mathbb{E}(U_t^\lambda|\,\mathcal{F}_s) = U_s^\lambda = \mathrm{e}^{\mathrm{i}\lambda X_s + \frac{1}{2}\lambda^2 s}$$

implies that, for every $\lambda \in \mathbb{R}$,

$$\mathbb{E}(\mathrm{e}^{\mathrm{i}\lambda(X_t - X_s)}|\,\mathcal{F}_s) = \mathrm{e}^{-\frac{1}{2}\lambda^2(t-s)}.$$

Taking the expectation of both sides of the above equality we obtain that the increment $X_t - X_s$ is $\mathcal{N}(0, t-s)$. Moreover, by Lemma 15.7 the above equality implies that $X_t - X_s$ is independent of $\mathcal{F}_s$. Therefore, all conditions of Def. 3.1 are satisfied. □

Lemma 15.7. *Let* $\mathcal{G} \subset \mathcal{F}$ *be a sub-*σ*-algebra and* X *a random variable such that for every* $\lambda \in \mathbb{R}$

$$\mathbb{E}[\mathrm{e}^{\mathrm{i}\lambda X}|\,\mathcal{G}] = \mathbb{E}[\mathrm{e}^{\mathrm{i}\lambda X}] \quad a.s.$$

Then X *is independent of* $\mathcal{G}$.

Proof. We have to prove that every $\mathcal{G}$-measurable real random variable Y is independent of X. The characteristic function of the pair (X, Y) computed at $\theta = (\lambda, \mu)$ is equal to

$$\mathbb{E}[\mathrm{e}^{\mathrm{i}\lambda X}\mathrm{e}^{\mathrm{i}\mu Y}] = \mathbb{E}[E(\mathrm{e}^{\mathrm{i}\lambda X}|\mathcal{G})\mathrm{e}^{\mathrm{i}\mu Y}] = \mathbb{E}[\mathrm{e}^{\mathrm{i}\lambda X}]\mathbb{E}[\mathrm{e}^{\mathrm{i}\mu Y}].$$

This implies the claim. □

Theorem 15.8. (Girsanov) *Let* $0 < T < \infty$ *and* $H \in \mathcal{H}_{T,\mathrm{loc}}$*, that is,* H *is predictable and* $\int_0^T H_s^2 \mathrm{d}s < \infty$ *a.s. Define a stochastic process* $Z = (Z_t)_{t\in[0,T]}$ *by*

$$Z_t = \exp\left(\int_0^t H_s\,\mathrm{d}B_s - \frac{1}{2}\int_0^t H_s^2\,\mathrm{d}s\right).$$

If $\mathbb{E}Z_T = 1$*, that is,* Z *is a martingale, then the process*

$$\tilde{B} = \left(B_t - \int_0^t H_s\,\mathrm{d}s\right)_{t\in[0,T]}$$

is a Brownian motion in the modified probability space $(\Omega, \mathcal{F}, \mathbb{Q}_T)$*, where the measure* $\mathbb{Q}_T$ *is defined by* $\mathbb{Q}_T(A) = \mathbb{E}(\mathbb{1}_A Z_T)$ *for* $A \in \mathcal{F}$.

Proof. The random variable Z_T is non-negative and $\mathbb{E}Z_T = 1$. Thus, $\mathbb{Q}_T$ is a probability measure. Note also that if $\mathbb{P}(A) = 0$, then $\mathbb{Q}_T(A) = 0$. Hence, events that occur $\mathbb{P}$-almost surely also occur $\mathbb{Q}_T$-almost surely. The process $\tilde{B}$ is continuous, adapted to $\mathcal{F}_t$, and $\tilde{B}_0 = 0$. By Proposition 15.6 it is therefore sufficient to show that for all $\lambda \in \mathbb{R}$, the process

$$U = \exp\left(\mathrm{i}\lambda\tilde{B}_t + \frac{1}{2}\lambda^2 t\right)_{t\in[0,T]}$$

is a martingale with respect to $\mathbb{Q}_T$. Note that

$$\begin{aligned} U_t Z_t &= \exp\left(\mathrm{i}\lambda \tilde{B}_t + \frac{1}{2}\lambda^2 t\right)\exp\left(\int_0^t H_s\,\mathrm{d}B_s - \frac{1}{2}\int_0^t H_s^2\,\mathrm{d}s\right) \\ &= \exp\left(\mathrm{i}\lambda B_t + \int_0^t H_s\,\mathrm{d}B_s - \frac{1}{2}\int_0^t (2\mathrm{i}\lambda H_s - \lambda^2 + H_s^2)\,\mathrm{d}s\right) \\ &= \exp\left(\int_0^t (\mathrm{i}\lambda + H_s)\,\mathrm{d}B_s - \frac{1}{2}\int_0^t (\mathrm{i}\lambda + H_s)^2\,\mathrm{d}s\right) = \exp\left(N_t - \frac{1}{2}\langle N\rangle_t\right) \end{aligned}$$

where $N = \int (\mathrm{i}\lambda + H_s)\,\mathrm{d}B_s \in \mathcal{M}_{T,\mathrm{loc}}$. Thus, the process UZ is a local martingale with respect to $\mathbb{P}$. Hence, there exist stopping times $\tau_n \nearrow T$ such that $U^{\tau_n} Z^{\tau_n}$ is a martingale. For every $n \in \mathbb{N}_+$ and every bounded stopping time τ we have

$$\begin{aligned} \mathbb{E}_{\mathbb{Q}_T} U_0 &= \mathbb{E}(U_0 Z_T) = \mathbb{E}(U_0 \mathbb{E}(Z_T|\mathcal{F}_0)) = \mathbb{E}(U_0 Z_0) = \mathbb{E}(U_{\tau_n\wedge\tau} Z_{\tau_n\wedge\tau}) \\ &= \mathbb{E}(U_{\tau_n\wedge\tau}\mathbb{E}(Z_T|\mathcal{F}_{\tau_n\wedge\tau})) = \mathbb{E}(U_{\tau_n\wedge\tau} Z_T) = \mathbb{E}_{\mathbb{Q}_T} U_{\tau_n\wedge\tau}, \end{aligned}$$

where we used Remark 9.3 and the fact that Z and $U^{\tau_n} Z^{\tau_n}$ are martingales. By Remark 9.3 we conclude that U^{τ_n} is a martingale with respect to $\mathbb{Q}_T$. This implies that U is a $\mathbb{Q}_T$-local martingale. Since U is bounded by Lemma 15.4 it is a martingale. □

Theorem 15.9. *Let $H \in \mathcal{H}_{T,\mathrm{loc}}$. Define $M = (M_t)_{t\in[0,T]}$ and $Z = (Z_t)_{t\in[0,T]}$ by*

$$M_t = \int_0^t H_s\,\mathrm{d}B_s, \qquad Z_t = \exp\left(M_t - \frac{1}{2}\langle M_t\rangle_t\right), \qquad t \in [0,T].$$

The process Z is a martingale if any of the following conditions is satisfied:

(i). $\mathbb{E}\left(\mathrm{e}^{\frac{1}{2}\langle M\rangle_T}\right) = \mathbb{E}\left(\mathrm{e}^{\frac{1}{2}\int_0^T |H_s|^2\,\mathrm{d}s}\right) < +\infty$ *(the Novikov criterion).*

(ii). M is a martingale, $\sup_{t\in[0,T]}\mathbb{E}(M_t^2) < \infty$ *and* $\mathbb{E}\left(\mathrm{e}^{\frac{1}{2}M_T}\right) < +\infty$ *(the Kazamaki criterion).*

(iii). There exists $\mu > 0$ such that $\sup_{t\in[0,T]}\mathbb{E}(\mathrm{e}^{\mu|H_t|^2}) < \infty$.

Proof. See Theorem 12.2 and Corollary 12.1 in [Bal17]. □

Example 15.10. Let $X = (X_t)_{t\in[0,T]}$ be a Brownian motion, $\lambda \in \mathbb{R}$ and

$$Z_t = \exp\left(\lambda\int_0^t X_s\,\mathrm{d}X_s - \frac{\lambda^2}{2}\int_0^t X_s^2\,\mathrm{d}s\right).$$

Using the fact that X_t is $\mathcal{N}(0,t)$ it is easy to see that $\mathbb{E}(e^{aX_t^2}) = (1-2at)^{-1/2}$ for all $a < \frac{1}{2t}$. Hence,

$$\mathbb{E}(\mathrm{e}^{\mu\lambda^2 X_t^2}) = \frac{1}{1 - 2\mu\lambda^2 t} \leqslant \frac{1}{1 - 2\mu\lambda^2 T}$$

and Theorem 15.9 (iii) applied with $\mu \in \left(0, \frac{1}{2\lambda^2 T}\right)$ implies that Z is a martingale. We can therefore consider on $\mathcal{F}$ the probability $\mathbb{Q}_T$ with density Z_T with respect to $\mathbb{P}$. The Girsanov theorem states that

$$B_t = X_t - \lambda\int_0^t X_s\,\mathrm{d}s$$

is a Brownian motion in the probability space $(\Omega, \mathcal{F}, \mathbb{Q}_T)$. Thus, under $\mathbb{Q}_T$ the process $(X)_{t\in[0,T]}$ solves the following SDE

$$\mathrm{d}X_t = \lambda X_t \,\mathrm{d}t + \mathrm{d}B_t.$$

For $\lambda < 0$ the solution of the above equation is known as Ornstein-Uhlenbeck process.

16 Martingale representation theorem

Let B be a Brownian motion adapted to a filtration $(\mathcal{F}_t)_{t\geqslant 0}$ and such that $(B_{s+t} - B_t)_{s\geqslant 0}$ is independent of $\mathcal{F}_t$ for all $t \geqslant 0$. We have seen that the stochastic integral of $H \in \mathcal{H}_T$ with respect to B is a square integrable martingale. In this section, we prove that the converse is also true for a particular choice of the filtration.

Theorem 16.1. *Let B be a Brownian motion and $(\mathcal{F}_t)_{t\geqslant 0}$ be the natural filtration of B augmented with the events of probability zero, which was introduced in Remark 8.2. Then every square integrable random variable Z measurable with respect to $\mathcal{F}_T$ for some $T > 0$ is of the form*

$$Z = c + \int_0^T H_s \,\mathrm{d}B_s,$$

where $c \in \mathbb{R}$ and $H \in \mathcal{H}_T$. Moreover, the above representation is unique.

Proof. The uniqueness is obvious, as c is determined by $c = \mathbb{E}[Z]$ whereas, if $H^{(1)}$ and $H^{(2)}$ were two processes in $\mathcal{H}_T$ satisfying (16.1), then from the relation

$$\int_0^T \big(H_s^{(1)} - H_s^{(2)}\big) \,\mathrm{d}B_s = 0$$

we have immediately, by the isometry property of the stochastic integral, that

$$\mathbb{E}\bigg[\int_0^T |H_s^{(1)} - H_s^{(2)}|^2 \,\mathrm{d}s\bigg] = 0$$

and therefore $H_s^{(1)} = H_s^{(2)}$ for almost every $s \in [0, T]$ a.s. For the proof of the existence see Theorem 12.4 in [Bal17]. □

Theorem 16.2. *Let B and $(\mathcal{F}_t)_{t\geqslant 0}$ be as in Theorem 16.1 and $(M_t)_{t\in[0,T]}$ be a square integrable martingale adapted to $(\mathcal{F}_t)_{t\geqslant 0}$. Then there exist a unique process $H \in \mathcal{H}_T$ and a constant $c \in \mathbb{R}$ such that*

$$M_t = c + \int_0^t H_s \,\mathrm{d}B_s \quad a.s.$$

for all $t \in [0, T]$. In particular, $(M_t)_{t\in[0,T]}$ admits a continuous modification, that is there exists a continuous process $(\tilde{M}_t)_{t\in[0,T]}$ such that $M_t = \tilde{M}_t$ a.s. for each $t \in [0, T]$.

Proof. As M_T is square integrable, by Theorem 16.1 there exists a unique process $H \in \mathcal{H}_T$ such that

$$M_T = c + \int_0^T H_s \,\mathrm{d}B_s$$

and therefore

$$M_t = \mathbb{E}(M_T|\mathcal{F}_t) = c + \int_0^t H_s \, \mathrm{d}B_s \quad \text{a.s.},$$

which finishes the proof. □

17 PDE problems and diffusion processes

In this section, we discuss representations of solutions of elliptic and parabolic PDEs as expectations of functionals of diffusion processes.

Example 17.1. Suppose that u solves the heat equation

$$\begin{cases} \dfrac{\partial u}{\partial t} = \Delta u & \text{on} \quad \mathbb{R}_{\geqslant} \times \mathbb{R}^d, \\ u(0, \cdot) = f & \text{on} \quad \mathbb{R}^d. \end{cases}$$

Let $B = (B^1, \ldots, B^d)$ be a vector of independent Brownian motions. The probability density of B_t coincides with $x \mapsto \frac{1}{(2\pi t)^{d/2}} \exp\left(-\frac{|x|^2}{2t}\right)$. Thus, the probability density of $\sqrt{2} B_t$ coincides with the fundamental solution of the heat equation $y \mapsto K(t, y) = \frac{1}{(4\pi t)^{d/2}} \exp\left(-\frac{|y|^2}{4t}\right)$ and

$$u(t, x) = \int_{\mathbb{R}^d} f(x + y) \, K(t, y) \, \mathrm{d}y = \mathbb{E}(f(x + \sqrt{2} B_t)) = \mathbb{E}_x(f(\sqrt{2} B_t)).$$

We would like to find a relation between PDEs and stochastic processes solving a multidimensional stochastic differential equation

$$\mathrm{d}X_t^i = b^i(X_t) \, \mathrm{d}t + \sum_{j=1}^m \sigma^{ij}(X_t) \, \mathrm{d}B_t^j, \qquad i \in \{1, \ldots, m\},$$

with continuous $b \colon \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma \colon \mathbb{R}^d \to \mathbb{R}^{d \times m}$. We assume these functions do not depend on time. Using the matrix notation we write equivalently

$$\mathrm{d}X_t = b(X_t) \, \mathrm{d}t + \sigma(X_t) \, \mathrm{d}B_t. \tag{17.1}$$

Notation 17.2. *Given a solution X^x to (17.1) with a deterministic initial condition $x \in \mathbb{R}^d$ we write $\mathbb{E}_x(f(X)) := \mathbb{E}(f(X^x))$.*

Definition 17.3. *The **generator of the diffusion** govern by (17.1) is the differential operator*

$$L := \frac{1}{2} \sum_{i,j=1}^d a^{ij} \, \partial_i \partial_j + \sum_{i=1}^d b^i \partial_i,$$

where $a = \sigma \sigma^T$, that is $a^{ij} = \sum_{k=1}^m \sigma^{ik} \sigma^{jk}$.

Example 17.4. Let $m=d$, $b=0$ and $\sigma=\sqrt{2}\mathbb{1}$, where $\mathbb{1}$ is the identity matrix of size d. Then $L=\Delta$ is the Laplacian and $X=\sqrt{2}B$.

Proposition 17.5. *Let $x\in\mathbb{R}^d$ and $X=(X^1,\ldots,X^d)$ be a solution to (17.1). Then for every function $u{:}\,\mathbb{R}_{\geqslant}\times\mathbb{R}^d\to\mathbb{R}$ that is C^1 in $\mathbb{R}_{\geqslant}$ and C^2 in $\mathbb{R}^d$, the process*

$$t\mapsto M_t^u=u(t,X_t)-u(0,X_0)-\int_0^t\left(\frac{\partial u}{\partial s}+Lu\right)(s,X_s)\,\mathrm{d}s$$

is a continuous local martingale.

Proof. The result follows by a standard application of the Itô formula. □

Definition 17.6. (Uniformly elliptic) *Given a domain $U\subseteq\mathbb{R}^d$ we say that $a{:}\,\bar{U}\to\mathbb{R}^{d\times d}$ is uniformly elliptic if there is a constant $c>0$ such that for all $\xi\in\mathbb{R}^d$ and $x\in\bar{U}$, we have*

$$\xi^T a(x)\xi\geqslant c|\xi|^2,$$

where $\bar{U}$ denotes the closure.

Remark 17.7. Equivalently, the smallest eigenvalue of a is bounded away from 0.

In what follows, we assume that $b{:}\,\mathbb{R}^d\to\mathbb{R}^d$ and $\sigma{:}\,\mathbb{R}^d\to\mathbb{R}^{d\times m}$ are globally Lipschitz such that $a=\sigma\sigma^T$ is uniformly elliptic. Note that since b,σ do not depend on time, they also satisfy the sublinear growth condition from Assumption 14.6. Consequently, the SDE (17.1) has a unique solution.

17.1 Dirichlet–Poisson problem

We first study the *Dirichlet–Poisson problem*. Let $U\subseteq\mathbb{R}^d$ be a domain, $f\in C_b(U)$ and $g\in C_b(\partial U)$. We want to find $u\in C^2(U)\cap C(\bar{U})$ such that

$$\begin{cases}-Lu=f & \text{on } U,\\ \quad u=g & \text{on } \partial U.\end{cases}$$

If $f=0$, this is called the *Dirichlet problem* and if $g=0$, this is called the *Poisson problem*.

It is possible to prove existence of a solution to the Dirichlet-Poisson using the diffusion process govern by (17.1). However, we shall prove a slightly weaker result. Assuming that we have a solution of the PDE we will show that it is represented by a certain formula involving the diffusion process. We first note the following theorem without proof.

Theorem 17.8. *Let U be a non-empty, connected, bounded, open subset of $\mathbb{R}^d$ with a smooth boundary. Then for every Hölder continuous $f{:}\,\bar{U}\to\mathbb{R}$ and continuous $g{:}\,\partial U\to\mathbb{R}$, the Dirichlet-Poisson process has a solution $u\in C^2(U)\cap C(\bar{U})$.*

We now state the main theorem of this section.

Theorem 17.9. *Let U, f and g be as in the previous theorem, $u \in C^2(U) \cap C(\bar{U})$ be a solution to the Dirichlet-Poisson problem and X be a solution to (17.1). Define the stopping time*

$$\tau_U = \inf \{ t \geqslant 0 \mid X_t \notin U \}.$$

Then $\mathbb{E}\tau_U < \infty$ and

$$u(x) = \mathbb{E}_x\biggl(g(X_{\tau_U}) + \int_0^{\tau_U} f(X_s)\, \mathrm{d}s \biggr).$$

Proof. Proposition 17.5 applies to functions defined on all of $\mathbb{R}^n$, while u is just defined on U. To circumvent this problem, we define

$$U_n = \left\{ y \in U \,\middle|\, \mathrm{dist}(y, \partial U) > \frac{1}{n} \right\}, \qquad \tau_n = \inf \{ t \geqslant 0 \,| X_t \notin U_n \},$$

and pick $u_n \in C_b^2(\mathbb{R}^d)$ such that $u|_{U_n} = u_n|_{U_n}$. Recalling our previous notation, let

$$M_t^n = (M^{u_n})_t^{\tau_n} = u_n(X_{t \wedge \tau_n}) - u_n(X_0) - \int_0^{t \wedge \tau_n} L u_n(X_s)\, \mathrm{d}s.$$

The process M_t^n is a bounded continuous local martingale. Hence, it is a true martingale. For $x \in U$ and n large enough, the martingale property implies

$$u(x) = u_n(x) = \mathbb{E}\biggl(u(X_{t \wedge \tau_n}) - \int_0^{t \wedge \tau_n} L u(X_s)\, \mathrm{d}s \biggr) = \mathbb{E}\biggl(u(X_{t \wedge \tau_n}) + \int_0^{t \wedge \tau_n} f(X_s)\, \mathrm{d}s \biggr).$$

To complete the proof, it remains to demonstrate that the limit $n \to \infty$ can be taken on both sides of the identity above.

We first show that $\mathbb{E}\tau_U < \infty$. Note that $\mathbb{E}\tau_U$ depends only on the process X and does not involve f or g. So we can take $f = 1$ and $g = 0$, and let v be a solution of the corresponding Dirichlet–Poisson problem. Then we have

$$\mathbb{E}(t \wedge \tau_n) = \mathbb{E}\biggl(-\int_0^{t \wedge \tau_n} L v(X_s)\, \mathrm{d}s \biggr) = v(x) - \mathbb{E}(v(X_{t \wedge \tau_n})) \leqslant 2 \|v\|_{L^\infty}.$$

Since $t \wedge \tau_n \nearrow \tau_U$ a.s. as $n \to \infty$ and $t \to \infty$, by monotone convergence theorem we obtain

$$\mathbb{E}(\tau_U) = \lim_{t \to \infty} \lim_{n \to \infty} \mathbb{E}(t \wedge \tau_n) \leqslant 2 \|v\|_{L^\infty} < \infty.$$

Using

$$\mathbb{E}\biggl(\int_0^\infty \mathbf{1}_{[0, \tau_U]}(s)\, |f(X_s)|\, \mathrm{d}s \biggr) \leqslant \|f\|_{L^\infty}\, \mathbb{E}(\tau_U) < \infty,$$

by the dominated convergence theorem, we conclude that

$$\lim_{t \to \infty} \lim_{n \to \infty} \mathbb{E}\biggl(\int_0^{t \wedge \tau_n} f(X_s)\, \mathrm{d}s \biggr) = \lim_{t \to \infty} \lim_{n \to \infty} \mathbb{E}\biggl(\int_0^\infty \mathbf{1}_{[0, t \wedge \tau_n]}(s) f(X_s)\, \mathrm{d}s \biggr) = \mathbb{E}\biggl(\int_0^{\tau_U} f(X_s)\, \mathrm{d}s \biggr).$$

Since u is continuous on $\bar{U}$, we also have

$$\lim_{t \to \infty} \lim_{n \to \infty} \mathbb{E}(u(X_{t \wedge \tau_n})) = \mathbb{E}(u(X_{\tau_U})) = \mathbb{E}(g(X_{\tau_U})).$$

This finishes the proof. $\square$

17.2 Cauchy problem for parabolic equations

We can also use SDEs to solve the *Cauchy problem* for parabolic equations. For $f\in C_b^2(\mathbb{R}^d)$, we want to find $u:\mathbb{R}_\geqslant\times\mathbb{R}^d\to\mathbb{R}$ that is C^1 in $\mathbb{R}_\geqslant$ and C^2 in $\mathbb{R}^d$ such that

$$\begin{cases}\dfrac{\partial u}{\partial t} = Lu & \text{on} \quad \mathbb{R}_\geqslant\times\mathbb{R}^d,\\ u(0,\cdot) = f & \text{on} \quad \mathbb{R}^d.\end{cases}$$

Theorem 17.10. *For every $f\in C_b^2(\mathbb{R}^d)$, there exists a solution $u\in C_b^{1,2}(\mathbb{R}_\geqslant\times\mathbb{R}^d)$ to the Cauchy problem.*

Theorem 17.11. *Let $f\in C_b^2(\mathbb{R}^d)$ and $u\in C_b^{1,2}(\mathbb{R}_\geqslant\times\mathbb{R}^d)$ be a solution to the Cauchy problem and X be a solution to (17.1). Then for $0\leqslant s\leqslant t$ we have*

$$\mathbb{E}_x(f(X_t)|\mathcal{F}_s)=u(t-s,X_s).$$

In particular,

$$u(t,x)=\mathbb{E}_x(f(X_t)).$$

Proof. Let $g(s,x)=u(t-s,x)$. Then

$$\left(\frac{\partial g}{\partial s}+Lg\right)(s,x)=-\frac{\partial}{\partial t}u(t-s,x)+Lu(t-s,x)=0.$$

Hence, by Proposition 17.5, $g(s,X_s)$ is a bounded martingale, and

$$u(t-s,X_s)=g(s,X_s)=\mathbb{E}(g(t,X_t)|\mathcal{F}_s)=\mathbb{E}(u(0,X_t)|\mathcal{F}_s)=\mathbb{E}(f(X_t)|\mathcal{F}_s),$$

which proves the desired identity. $\square$

The following generalization of the identity from the previous theorem is known as the *Feynman–Kac formula*.

Theorem 17.12. (Feynman–Kac formula) *Let $f\in C_b^2(\mathbb{R}^d)$ and $V\in C_b(\mathbb{R}^d)$ and suppose that $u\in C_b^{1,2}(\mathbb{R}_\geqslant\times\mathbb{R}^d)$ satisfies*

$$\begin{cases}\dfrac{\partial u}{\partial t}=Lu+Vu & \text{on} \quad \mathbb{R}_\geqslant\times\mathbb{R}^d,\\ u(0,\cdot)=f & \text{on} \quad \mathbb{R}^d,\end{cases}$$

where $(Vu)(t,x)=V(x)u(t,x)$ for all $t\geqslant 0$ and $x\in\mathbb{R}^d$. Then for all $t\geqslant 0$ and $x\in\mathbb{R}^d$ we have

$$u(t,x)=\mathbb{E}_x\left(f(X_t)\exp\left(\int_0^t V(X_r)\,\mathrm{d}r\right)\right),$$

where X is the solution to (17.1).

Proof. Let $Z_s = \exp(\int_0^s V(X_r)\,\mathrm{d}r)$. For $s \in [0,t]$, set

$$M_s = u(t-s, X_s) Z_s = f(s, X_s, Z_s), \qquad f(s,x,z) = u(t-s,x)z.$$

Let us show that M is a martingale on $[0,t]$. Indeed, by the Itô formula we have

$$\begin{aligned}\mathrm{d}M_s &= \frac{\partial f}{\partial s}(s, X_s, Z_s)\,\mathrm{d}s + \sum_{i=1}^d \frac{\partial f}{\partial x^i}(s, X_s, Z_s)\,\mathrm{d}X_s^i + \frac{\partial f}{\partial z}(s, X_s, Z_s)\,\mathrm{d}Z_s \\ &\quad + \frac{1}{2}\sum_{i,j=1}^d \frac{\partial f}{\partial x^i \partial x^j}(s, X_s, Z_s)\,a^{ij}(X_s)\,\mathrm{d}s.\end{aligned}$$

Since $\mathrm{d}Z_s = V(X_s) Z_s\,\mathrm{d}s$ and $\partial_t u = Lu + Vu$ we arrive at

$$\begin{aligned}\mathrm{d}M_s &= (-\partial_t u + Lu)(t-s, X_s)\,Z_s\,\mathrm{d}s \\ &\quad + (\nabla u)(t-s, X_s)\,Z_s\,\sigma(X_s)\,\mathrm{d}B_s \\ &\quad + u(t-s, X_s) V(X_s) Z_s\,\mathrm{d}s \\ &= (\nabla u)(t-s, X_s) Z_s\,\sigma(X_s)\,\mathrm{d}B_s.\end{aligned}$$

This proves that M is a local martingale. Since u and V are bounded, M is also bounded. Consequently, M is a martingale and

$$u(t,x) = M_0 = \mathbb{E}M_t = \mathbb{E}[u(0, X_t) Z_t] = \mathbb{E}[f(X_t) Z_t],$$

which finishes the proof. □

18 Markov property

Definition 18.1. *Let $B(\mathbb{R}^d)$ be the Banach space of bounded Borel functions equipped with the norm $\|f\| = \sup_{x\in\mathbb{R}^d}|f(x)|$. A collection of bounded linear operators $(Q_t)_{t\geqslant 0}$ on $B(\mathbb{R}^d)$ is a transition semigroup if:*

(i). $Q_t f \geqslant 0$ a.e. if $f \geqslant 0$ a.e.,

(ii). $Q_t 1 = 1$ where $1(x) = 1$ for all $x \in \mathbb{R}^d$,

(iii). $\|Q_t\| \leqslant 1$,

(iv). $Q_{t+s} = Q_t Q_s$ for all $t, s \geqslant 0$ (semigroup property).

We say that a process $X = (X_t)_{t\geqslant 0}$ adapted to a filtration $(\mathcal{F}_t)_{t\geqslant 0}$ is a Markov process with transition a semigroup $(Q_t)_{t\geqslant 0}$ if

$$\mathbb{E}(f(X_{s+t})|\mathcal{F}_s) = (Q_t f)(X_s) \tag{18.1}$$

for all $s, t \geqslant 0$ and $f \in B(\mathbb{R}^d)$.

Remark 18.2. The probability that the Markov process X starting at $x \in \mathbb{R}^d$ at time $t \geqslant 0$ belongs to a Borel set A coincides with $\mathbb{P}_x(X_t \in A) = \mathbb{E}_x(\mathbf{1}_A(X_t)) = (Q_t \mathbf{1}_A)(x)$.

Theorem 18.3. *Let B be a Brownian motion and $(\mathcal{F}_t)_{t\geqslant 0}$ be the natural filtration of B augmented with the events of probability zero. Assume that $b\colon \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma\colon \mathbb{R}^d \to \mathbb{R}^{d\times m}$ are Lipschitz continuous functions and $X = X^x$ is a solution to*

$$\mathrm{d}X_t = b(X_t)\,\mathrm{d}t + \sigma(X_t)\,\mathrm{d}B_t \tag{18.2}$$

with the initial condition $X_0 = x \in \mathbb{R}^d$. Then X is a Markov process with the semigroup

$$(Q_t f)(x) = \mathbb{E}_x f(X_t).$$

Lemma 18.4. *Under the assumptions of Theorem 18.3 there exists a map $S\colon \mathbb{R}_\geqslant \times \mathbb{R} \times \Omega \to \mathbb{R}$ such that:*

(i). for every $t \in \mathbb{R}_\geqslant$ and $x \in \mathbb{R}^d$ we have $X_t^x(\omega) = S(t,x,\omega)$ for almost all $\omega \in \Omega$,

(ii). the map $(t,x) \mapsto S(t,x,\omega)$ is continuous for almost all $\omega \in \Omega$,

(iii). for every $t \in \mathbb{R}_\geqslant$ and $x \in \mathbb{R}^d$ the map $\omega \mapsto S(t,x,\omega)$ is $\mathcal{F}_t$-measurable.

In particular, for all $t \geqslant 0$ the map $(x,\omega) \mapsto S(t,x,\omega)$ is measurable with respect to the σ-algebra $\mathcal{B}(\mathbb{R}^d) \otimes \mathcal{F}_t$ and the solution X of (18.2) with random initial condition $X_0 = Y$ satisfies for all $t \geqslant 0$ the idenity $X_t(\omega) = S(t, Y(\omega), \omega)$ for almost all $\omega \in \Omega$.

Proof. See Sec. 9.8 of [Bal17]. □

Proof of Theorem 18.3. It is easy to check that $(Q_t)_{t\geqslant 0}$ satisfies the conditions (i)-(iii). The condition (iv) follows from (18.1) and the definition of Q_t, since

$$(Q_{t+s} f)(x) = \mathbb{E}_x f(X_{t+s}) = \mathbb{E}_x(\mathbb{E}_x(f(X_{t+s})|\mathcal{F}_s)) = \mathbb{E}_x((Q_t f)(X_s)) = (Q_s Q_t f)(x)$$

for all $x \in \mathbb{R}^d$. Thus, it remains to prove (18.1). We have

$$X_s = X_0 + \int_0^s b(X_u)\,\mathrm{d}u + \int_0^s \sigma(X_u)\,\mathrm{d}B_u$$

and

$$X_{t+s} = X_0 + \int_0^{t+s} b(X_u)\,\mathrm{d}u + \int_0^{t+s} \sigma(X_u)\,\mathrm{d}B_u.$$

Set $\tilde{X}_t = X_{t+s}$, $\tilde{\mathcal{F}}_t = \mathcal{F}_{t+s}$, and $\tilde{B}_t = B_{t+s} - B_s$. Then $\tilde{B}$ is a Brownian motion adapted to $(\tilde{\mathcal{F}}_t)_{t\geqslant 0}$ and we have

$$\tilde{X}_t = \tilde{X}_0 + \int_0^t b(\tilde{X}_u)\,\mathrm{d}u + \int_0^t \sigma(\tilde{X}_u)\,\mathrm{d}\tilde{B}_u.$$

Indeed, this follows from $\int_s^{t+s} \sigma(X_u)\,\mathrm{d}B_u = \int_0^t \sigma(\tilde{X}_u)\,\mathrm{d}\tilde{B}_u$, which can be seen by approximating both sides by sums. Thus $\tilde{X}$ solves (18.2) with $\tilde{X}_0 = X_s$ and $B = \tilde{B}$.

Define $F\colon \mathbb{R}^d\times\Omega\to\mathbb{R}$ by $F(x,\omega)=f(S(t,x,\omega))$, where S is the solution map from Lemma 18.4. Let $\mathcal{G}=\tilde{\mathcal{F}}_0=\mathcal{F}_s$ and $\mathcal{H}$ be the σ-algebra generated by $\tilde{B}_{u\in[0,t]}=(B_u-B_s)_{u\in[s,s+t]}$ augmented with zero probability events. Then $\mathcal{G}$ and $\mathcal{H}$ are independent, $\tilde{X}_0=X_s$ is $\mathcal{G}$-measurable and F is $\mathcal{B}(\mathbb{R}^n)\otimes\mathcal{H}$-measurable. Hence, the assumptions of the freezing lemma stated below are satisfied and we obtain

$$\mathbb{E}(F(\tilde{X}_0,\cdot)|\mathcal{G})=G(\tilde{X}_0),$$

where the function $G\colon\mathbb{R}^d\to\mathbb{R}$ satisfies the equation

$$G(x)=\mathbb{E}(f(F(x,\cdot)))=\mathbb{E}f(S(t,x,\cdot))=(Q_tf)(x)$$

for all $x\in\mathbb{R}^d$. Since we have $\tilde{X}_t(\omega)=S(t,\tilde{X}_0(\omega),\omega)$ for almost all $\omega\in\Omega$ we conclude that

$$\mathbb{E}(f(X_{s+t})|\mathcal{F}_s)=\mathbb{E}(f(\tilde{X}_t)|\tilde{\mathcal{F}}_0)=\mathbb{E}(F(\tilde{X}_0,\cdot)|\mathcal{G})=G(\tilde{X}_0)=(Q_tf)(x).$$

This shows (18.1) and completes the proof. □

Lemma 18.5. (Freezing lemma) *Let $(\Omega,\mathcal{F},P)$ be a probability space and $\mathcal{G}$ and $\mathcal{H}$ be independent sub-σ-algebras of $\mathcal{F}$. Suppose that X is a $\mathcal{G}$-measurable random variable taking values in $\mathbb{R}^d$ and $F\colon\mathbb{R}^d\times\Omega\to\mathbb{R}$ is a $\mathcal{B}(\mathbb{R}^d)\otimes\mathcal{H}$-measurable function such that the random variable $\omega\mapsto F(X(\omega),\omega)$, denoted by $F(X,\cdot)$, is integrable. Then, we have*

$$\mathbb{E}(F(X,\cdot)|\mathcal{G})=G(X),\tag{18.3}$$

where the function $G\colon\mathbb{R}^d\to\mathbb{R}$ is defined by $G(x):=\mathbb{E}(F(x,\cdot))$ for all $x\in\mathbb{R}^d$.

Proof. Let us assume first that F is of the form $F(x,\omega)=f(x)Z(\omega)$, where Z is $\mathcal{H}$-measurable. In this case, $G(x)=f(x)\mathbb{E}Z$ and

$$\mathbb{E}(F(X,\cdot)|\mathcal{G})=\mathbb{E}(f(X)Z|\mathcal{G})=f(X)\mathbb{E}(Z|\mathcal{G})=f(X)\mathbb{E}Z=G(X).$$

Therefore, the statement is true for linear combinations of F of the form described. One obtains the general case with the help of Theorem 1.5 from [Bal17]. □

When modeling systems with Markov processes, we are often interested in their long-term behavior. In particular, we seek to understand whether the system stabilizes and where it spends most of its time. This is where invariant measures come into play. They capture the statistical equilibrium of a Markov process — a distribution that remains unchanged as the process evolves.

Definition 18.6. *Let X be a random variable in $\mathbb{R}^d$. The probability measure*

$$\mathcal{B}(\mathbb{R}^d)\ni A\mapsto\mathbb{P}(X\in A)\in[0,1]$$

is called the law of X and denoted by $\mathrm{Law}(X)$.

Definition 18.7. *Let $(X_t)_{t\geqslant 0}$ be a Markov process. We say that a measure μ is invariant for X if the condition $\mathrm{Law}(X_0)=\mu$ implies $\mathrm{Law}(X_t)=\mu$ for all $t\geqslant 0$.*

Remark 18.8. Let $(Q_t)_{t\geqslant 0}$ be the transition semigroup of $(X_t)_{t\geqslant 0}$. A measure μ is invariant iff

$$\int f(x)\,\mu(\mathrm{d}x) = \int (Q_t f)(x)\,\mu(\mathrm{d}x)$$

for all $t\geqslant 0$ and $f\in B(\mathbb{R}^d)$.

Example 18.9. Let $\lambda>0$ and B be a Brownian motion. Consider the Ornstein-Uhlenbeck process

$$\mathrm{d}X_t = -\lambda X_t\,\mathrm{d}t + \mathrm{d}B_t$$

for all $t\geqslant 0$. We know that the solution is given by

$$X_t = X_0\,e^{-\lambda t} + \int_0^t e^{-\lambda(t-s)}\mathrm{d}B_s.$$

Let X_0 be $\mathcal{N}(0,1/(2\lambda))$ and independent of B. Then, X_t has Gaussian distribution with vanishing expectation and variance

$$\mathrm{Var}(X_t) = \frac{e^{-2\lambda t}}{2\lambda} + \frac{1-e^{-2\lambda t}}{2\lambda} = \frac{1}{2\lambda}.$$

Therefore, X_t is $\mathcal{N}(0,1/(2\lambda))$. Thus, the measure

$$\mu(\mathrm{d}x) = \sqrt{\frac{\lambda}{\pi}}e^{-\lambda x^2}\mathrm{d}x$$

is invariant for X.

Bibliography

[Bal17] P. Baldi. *Stochastic Calculus: An Introduction Through Theory and Exercises*. Springer, 2017.

[Eva12] L.C. Evans. *An Introduction to Stochastic Differential Equations*. American Mathematical Society, 2012.

[RY04] D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion*. Springer, 2004.

[Ste01] J.M. Steele. *Stochastic Calculus and Financial Applications*. Springer, 2001.