Lecture notes in stochastic analysis

Taught to master's students at the University of Porto in 2010-2011

Contents

1. Overview of the measure theory and probability

This chapter reviews some basic topics and concepts of the measure theory and probability which are important for understanding the rest of the course.

1.1 The definition of measure

Definition 1.1. *Let Ω be a set. A σ-**algebra** $\mathcal{F}$ is a family of subsets of Ω possessing the following properties:*

(1) The empty set $\varnothing$ belongs to $\mathcal{F}$.

(2) If $A \in \mathcal{F}$, then $A^c \in \mathcal{F}$, where $A^c = \Omega \diagdown A$ is the complement of A in Ω.

(3) If $A_1, A_2, \ldots, A_n, \ldots \in \mathcal{F}$, then $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$.

Definition 1.2. *A **measure** μ on $\mathcal{F}$ is a function $\mathcal{F} \to [0, \infty)$ such that for any sequence of subsets $\{A_i\}_{i=1}^{\infty}$, $A_i \in \mathcal{F}$, with $A_i \cap A_j = \varnothing$, $i \neq j$,*

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i).$$

Definition 1.3. *A **signed measure** on $\mathcal{F}$ is a function $\mathcal{F} \to (-\infty, \infty)$ where the rest of the properties is the same as for the measure.*

Definition 1.4. *We say that some property A holds **almost everywhere** (a.e.) if the set of those point $x \in E$ where the property A does not hold is zero.*

Definition 1.5. *A **probability measure** P on $\mathcal{F}$ is a function $\mathcal{F} \to [0, 1]$ satisfying the following conditions:*

(1) $P(\varnothing) = 0$, $P(\Omega) = 1$.

(2) If a sequence of subsets $\{A_i\}_{i=1}^{\infty}$, $A_i \in \mathcal{F}$, is such that $A_i \cap A_j = \varnothing$, $i \neq j$, then

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i).$$

Sometimes, if a σ-algebra $\mathcal{F}$ of subsets of Ω is specified, we may use the expression "a probability measure on Ω" keeping in mind that it is actually a probability measure on the σ-algebra $\mathcal{F}$ of subsets of Ω. The $\mathcal{F}$-measurable subsets of Ω are often thought as *events*, while $P(A)$ is thought as the probability that the event A occurs. If $P(A) = 1$, we say that the event A occurs *almost surely (a.s.)*.

The triple $(\Omega, \mathcal{F}, P)$ is called a *probability space*.

1.2 Measurable functions and random variables

Let (E, μ) be a measurable space, i.e. a set E with the measure μ on a σ-algebra $\mathcal{G}$ of subsets of E. A function $Y : E \to \mathbb{R}^n$ is called *$\mathcal{F}$-measurable* if for any Borel subset $U \subset \mathbb{R}^n$,

$$Y^{-1}(U) = \{\omega \in \Omega : Y(\omega) \in U\} \in \mathcal{G}.$$

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

Definition 1.6. *An $\mathcal{F}$-measurable function $X : \Omega \to \mathbb{R}^n$ is called a* ***random variable****.*

Let $\mathcal{B}$ be the Borel σ-algebra of $\mathbb{R}^n$. Every random variable X generates a σ-algebra $\mathcal{H}_X$:

$$\mathcal{H}_X = \{X^{-1}(B), B \in \mathcal{B}\}.$$

Clearly, X is $\mathcal{H}_X$-measurable. Moreover, $\mathcal{H}_X$ is the smallest σ-algebra among σ-algebras with the property that X is measurable with respect to this σ-algebra.

Let I be a finite or infinite index set.

Definition 1.7.

(1) Two subsets $A \in \mathcal{F}$ and $B \in \mathcal{F}$ are called ***independent*** *if*

$$P(A \cap B) = P(A) \cdot P(B).$$

(2) σ-algebras $\{\mathcal{H}_i\}$, $i \in I$, are called ***independent*** *if for any choice of sets $H_{i_1} \in \mathcal{H}_{i_1}$, ..., $H_{i_k} \in \mathcal{H}_{i_k}$*

$$P(H_{i_1} \cap \cdots \cap H_{i_k}) = P(H_{i_1}) \cdots P(H_{i_k}).$$

(3) Random variables $\{X_i\}$, $i \in I$, are called ***independent*** *if the σ-algebras $\mathcal{H}_{X_i}$ generated by X_i are independent.*

Theorem 1.8. *If $f : E \to \mathbb{R}^n$ is a measurable function, and $g : \mathbb{R}^n \to \mathbb{R}^m$ is another measurable function, then the composition $g \circ f$ is a measurable function.*

Remark. We assume that on $\mathbb{R}^n$ we are give the Borel σ-algebra.
Assumption. Below we assume that the σ-algebra $\mathcal{G}$ of subsets of the measurable space E is augmented with all subsets of zero μ-measure sets.

Definition 1.9. *A measure μ defined on a σ-algebra augmented with all subsets of null sets is said to be* ***complete****.*

Theorem 1.10. *A funtion $f : E \to \mathbb{R}$ is measurable if and only if the set $\{x \in E : f(x) < c\}$ is measurable for every $c \in \mathbb{R}$.*

Theorem 1.11. *The sum, the difference, and the product of two measurable functions f and g are measurable. If $g(x) \neq 0$ for all $x \in E$, then $\frac{f}{g}$ is also measurable.*

Definition 1.12. *Let $f_n : E \to \mathbb{R}$, and $f : E \to \mathbb{R}$. We say that $f_n(x)$ converges to $f(x)$ **almost everywhere** (a.e.) on E if the measure of those $x \in E$ where the convergence does not hold is zero.*

Theorem 1.13. *If a sequence of measurable functions f_n converges to f almost everywhere, then f is also measurable.*

Definition 1.14. *A sequence of functions $f_n : E \to \mathbb{R}$ is said to converge **in measure** to a function $f : E \to \mathbb{R}$ if for every $\varepsilon > 0$,*

$$\lim_{n\to\infty} \mu\{x : |f_n(x) - f(x)| \geqslant \varepsilon\} = 0.$$

Theorem 1.15. *If a sequence of measurable functions $\{f_n\}$ converges to a function f almost everywhere, then it converges to the same functions in measure.*

Theorem 1.16. *If a sequence of measurable functions $\{f_n\}$ converges to a function f in measure, then there exists a subsequence $\{f_{n_k}\}$ of $\{f_n\}$ that converges to f almost everywhere.*

1.3 Integration and expectations

Below we assume that a measure μ is given on a complete σ-algebra of subsets of E.

Definition 1.17. *A function f on (E, μ) is called **simple** if it takes no more than countable number of values.*

Theorem 1.18. *A function f taking no more than countable number of different values*

$$y_1, \dots, y_n, \dots,$$

is measurable if and only if the sets

$$A_n = \{x : f(x) = y_n\}$$

are measurable.

Theorem 1.19. *A function f is measurable if and only if one can represent it as a limit of a uniformly convergent sequence of simple measurable functions.*

Let $f = \sum_{i=1}^{\infty} y_i \, \mathbb{I}_{A_i}$ where $y_i \neq y_j$ if $i \neq j$ and the sets A_i are μ-measurable.

Definition 1.20. *The integral of f over a μ-measurable set A with respect to the measure μ is defined as follows:*

$$\int_A f(x)\mu(dx) = \sum_{i=1}^{\infty} y_i \mu(A_i \cap A).$$

*If the series on the right-hand side converges, then the function f is called **integrable** on the set A.*

Below we list some properties of the integral for simple functions.

1. Additivity:

$$\int_A [f(x)+g(x)]\mu(dx) = \int_A f(x)\mu(dx) + \int_A g(x)\mu(dx),$$

moreover, the existence of the integrals on the right-hand side implies the existence of the integral on the left-hand side.

2. Multiplicativity: If $k \in \mathbb{R}$, then

$$\int_A kf(x)\mu(dx) = k\int_A f(x)\mu(dx),$$

moreover, the existence of the integrals on the right-hand side implies the existence of the integral on the left-hand side.

3. If f is bounded on A by a constant M, then

$$\Big|\int_A f(x)\mu(dx)\Big| \leqslant M\mu(A).$$

Definition 1.21. *A function f is called integrable on A if there exists a sequence of simple functions $\{f_n\}$ which converges to f uniformly. The limit*

$$\lim_{n\to\infty}\int_A f_n(x)\mu(dx) \tag{1}$$

is denoted $\int_A f(x)\mu(dx)$ and called the ***integral*** *of f over A with respect to μ.*

The integral of f over A with respect to μ is well defined. Indeed, limit (1) exists for any uniformly convergent system of simple functions since

$$\Big|\int_A f_n(x)\mu(dx) - \int_A f_m(x)\mu(dx)\Big| \leqslant \mu(A)\sup_{x\in A}|f_n(x)-f_m(x)|.$$

The above inequality also implies that limit (1) does not depend on the choice of a sequence of simple functions that uniformly converges to f. Properties 1, 2, and 3 easily follow for the integral defined for an arbitrary integrable function. Below we will list some additional properties.

4. If $f(x) \geqslant g(x)$, then

$$\int_A f(x)\mu(dx) \geqslant \int_A g(x)\mu(dx).$$

5. If $\mu(A)=0$, then $\int_A f(x)\mu(dx) = 0$.

6. If $f(x) = g(x)$ almost everywhere on A, then
$$\int_A f(x)\mu(dx) = \int_A g(x)\mu(dx).$$

7. σ-additivity: If $A = \cup_n A_n$ where $A_i \cap A_j = \emptyset$ $(i \neq j)$, then
$$\int_A f(x)\mu(dx) = \sum_n \int_{A_n} f(x)\mu(dx).$$
Moreover, from the existence of the integral on the left-hand side it follows that each integral over A_n on the right-hand side exists, and the series converges.

8. If $\mu(A) \neq 0$, and
$$\int_A |f(x)|\mu(dx) = 0,$$
then $f(x) = 0$ on A a.e..

9. Chebyshev's inequality: if $f(x) \geqslant 0$ and $c > 0$, then
$$\mu\{x \in A : f(x) > c\} \leqslant \frac{1}{c} \int_A f(x)\mu(dx).$$

10. Absolute continuity of the integral: If f is integrable on A, then for every $\varepsilon > 0$, there exists a $\delta > 0$ such that for every μ-measurable subset $C \subset A$ with $\mu(C) < \delta$,
$$\left| \int_C f(x)\mu(dx) \right| < \varepsilon.$$

Passing to the limit under the integral sign

Theorem 1.22 (Lebesgue's [or dominated convergence] theorem)**.** *Let f_n converge to f on A a.e. and for all n*
$$|f_n(x)| \leqslant \phi(x) \text{ a.e.}$$
where ϕ is integrable on A. Then, the function f is integrable on A, and
$$\lim_{n\to\infty} \int_A f_n(x)\mu(dx) = \int_A f(x)\mu(dx).$$

Theorem 1.23 (Beppo Levy's [or monotone convergence] theorem)**.** *Let $f_1(x) \leqslant f_2(x) \leqslant \cdots \leqslant f_n(x) \leqslant \cdots$ on A, moreover, all f_n are integrable and their integrals are bounded:*
$$\int_A f_n(x)\mu(dx) \leqslant K.$$

Then, almost everywhere on A there exists a limit

$$\lim_{n\to\infty} f_n(x) = f(x).$$

Moreover, the function $f(x)$ is integrable on A, and

$$\lim_{n\to\infty} \int_A f_n(x)\mu(dx) = \int_A f(x)\mu(dx).$$

Theorem 1.24 (Fatou Lemma)**.** *Let a sequence of integrable non-negative functions f_n converge to f a.e. on A, and*

$$\int_A f_n(x)\mu(dx) \leqslant K.$$

Then f is integrable on A, and

$$\int_A f(x)\mu(dx) \leqslant K.$$

Fubini's theorem

Let $A \subset U \times V$. Define

$$\begin{aligned} A_x &= \{y \in Y : (x, y) \in A\} \quad (x \text{ is fixed}) \\ A_y &= \{x \in X : (x, y) \in A\} \quad (y \text{ is fixed}). \end{aligned}$$

Theorem 1.25. *Let the measures μ_U and μ_V be defined on U and V, respectively, σ-additive, and complete. Further let $\mu = \mu_U \otimes \mu_V$ and the function $f(x, y)$ be integrable on the set $A \subset U \times V$. Then*

$$\int_A f(x)\mu(dx) = \int_U \mu_U(dx) \int_{A_x} f(x, y)\mu_V(dy) = \int_V \mu_V(dy) \int_{A_y} f(x, y)\mu_U(dx). \tag{2}$$

Corollary 1.26. *Let one the integrals*

$$\begin{aligned} &\int_U \mu_U(dx) \int_{A_x} |f(x, y)| \, \mu_V(dy) \quad \textit{or} \\ &\int_V \mu_V(dy) \int_{A_y} |f(x, y)| \, \mu_U(dx). \end{aligned}$$

exists. Then f is integrable on A and (2) holds.

Distributions, expectations, characteristic functions

Every random variable X induces a probability distribution μ_X on $\mathbb{R}^n$ defined on the σ-algebra of Borel subsets of $\mathbb{R}^n$ by the formula:

$$\mu_X(B) = P\big(X^{-1}(B)\big). \tag{3}$$

Definition 1.27. *The measure μ_X defined by (3) is called the* ***distribution*** *of X.*

Definition 1.28. *The number*

$$E[X] = \int_\Omega X(\omega)P(d\omega) = \int_{\mathbb{R}^n} x\,\mu_X(dx)$$

is called the ***expectation*** *of X.*

Definition 1.29. *The* ***characteristic function*** *of a random variable $X : \Omega \to \mathbb{R}^n$ is the functio $\varphi_X : \mathbb{R}^n \to \mathbb{C}$ (where $\mathbb{C}$ denotes the complex numbers) defined by*

$$\varphi_X(y_1, \dots, y_n) = \mathbb{E}[\exp\{i\,(y_1X_1 + \cdots + y_nX_n)\}] = \int_{\mathbb{R}^n} e^{i(y,x)_{\mathbb{R}^n}}\,(P \circ X^{-1})(dx)$$

where $X_1, \dots, X_n$ are the coordinates of X in $\mathbb{R}^n$, $(y,x)_{\mathbb{R}^n}$ is the scalar product in $\mathbb{R}^n$.

Theorem 1.30. *The characteristic function of X determines the distribution $P \circ X^{-1}$ of X uniquely.*

Proof. The characteristic function is actually the Fourier transform of the measure $P \circ X^{-1}$, and therefore defines this measure uniquely. □

Conditional expectations

Definition 1.31. *Let a measure μ and a signed measure ν be defined on a σ-algebra $\mathcal{G}$ of the space E. The signed measure ν is said to be* ***absolutely continuous*** *with respect to the measure μ if $\nu(A) = 0$ for any set $A \in \mathcal{G}$ with $\mu(A) = 0$.*

Theorem 1.32 (Radon-Nikodym's theorem)**.** *Let μ be a measure on a σ-algebra $\mathcal{G}$ of the space E, and let the measure ν, defined on the same σ-algebra $\mathcal{G}$, be absolutely continuous with respect to μ. Then, there exists an integrable function f on E such that for any $A \in \mathcal{G}$*

$$\nu(A) = \int_A f(x)\mu(dx).$$

The function f is called the derivative of ν with respect to μ and denoted $f = \frac{d\nu}{d\mu}$.

Now let $(\Omega, \mathcal{F}, P)$ be a probability space, X be a random variable and $\mathcal{B} \subset \mathcal{F}$ be a sub-σ-algebra.

Definition 1.33 (Conditional expectation)**.** *A $\mathcal{B}$-measurable random variable Y is called the* ***conditional expectation*** *of X with respect to $\mathcal{B}$, and denoted $Y = \mathbb{E}[X|\mathcal{B}]$, if for any $\mathcal{B}$-measurable set A, it holds that*

$$\int_A Y(\omega)P(d\omega) = \int_A X(\omega)P(d\omega).$$

Conditional expectation always exists by Radon-Nikodym's theorem. Indeed, consider the restriction of the measure P to $\mathcal{B}$. We denote this restriction by the same symbol P. Define the signed measure $\nu(A) = \int_A X(\omega)P(d\omega)$ on the σ-algebra $\mathcal{B}$. Clearly, $\nu(A)$ is absolutely continuous with respect to P. By Radon-Nikodym's theorem, there exists a $\mathcal{B}$-measurable function Y such that $\nu(A) = \int_A Y(\omega)P(d\omega)$. The latter implies that $Y = \mathbb{E}[X|\mathcal{B}]$.

Theorem 1.34 (Properties of conditional expectations)**.** *Let, as before, $\mathcal{B}$ be a sub-σ-algebra of $\mathcal{F}$.*

1. *If a random variable X is $\mathcal{B}$-measurable, then $\mathbb{E}[X|\mathcal{B}] = X$.*
2. *For any square integrable $\mathcal{B}$-measurable Z, $\mathbb{E}[ZX] = \mathbb{E}[Z\,\mathbb{E}[X|\mathcal{B}]]$.*
3. *$\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X|\mathcal{B}]]$.*
4. - *Additivity: $\mathbb{E}[X + Y|\mathcal{B}] = \mathbb{E}[X|\mathcal{B}] + \mathbb{E}[Y|\mathcal{B}]$.*
   - *Linearity: $\mathbb{E}[cX + d|\mathcal{B}] = c\,\mathbb{E}[X|\mathcal{B}] + d$, where $c, d \in \mathbb{R}$.*
5. *If Z is $\mathcal{B}$-measurable, then $\mathbb{E}[ZX|\mathcal{B}] = Z\mathbb{E}[X|\mathcal{B}]$ a.s..*
6. *If $\mathcal{H} \subset \mathcal{B}$ is a sub-σ-algebra, then $\mathbb{E}[\mathbb{E}[X|\mathcal{B}]|\mathcal{H}] = \mathbb{E}[X|\mathcal{H}]$.*
7. *If $X \leqslant Y$ where Y is another random variable, then $\mathbb{E}[X|\mathcal{B}] \leqslant \mathbb{E}[Y|\mathcal{B}]$ a.s..*

1.4 Examples

Example 1.(Wiener's probability space). Let $\Omega = [0, 1]$. Define a σ-algebra $\mathcal{F}$ as the σ-algebra of all Lebesgue-measurable subsets of $[0, 1]$, and let P be the Lebesgue measure on $[0, 1]$. The triple $(\Omega, \mathcal{F}, P)$ builds a probability space. This probability space is called *Wiener's probability space*.
Example 2.(Gaussian measure, Gaussian random variable). A probability measure γ on the real line is called *Gaussian* if its density is given by the formula:

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\Big(-\frac{(x-a)^2}{2\sigma^2}\Big).$$

This means that the measure γ of a Borel subset $A \subset \mathbb{R}$ is defined by

$$\gamma(A) = \int_A p(x)\,dx.$$

The numbers a and σ^2 are called the *mean* and resp. the *variance*. A real-valued random X variable is called *Gaussian* if its distribution is a Gaussian measure. If $a = 0$ and $\sigma^2 = 1$ the random variable X is called *standard Gaussian* random variable. A Gaussian random variable is also called a normally distributed random variable.

EXAMPLE 3.(Multi-normal distribution). An $\mathbb{R}^n$-valued random variable X is called **multi-normal** of its distribution has the density:

$$p_X(x_1,\ldots,x_n)=\frac{\sqrt{|A|}}{(2\pi)^{\frac{n}{2}}}\,\exp\Big(-\frac{1}{2}\sum_{j,k=1}^{n}(x_j-m_j)a_{jk}(x_k-m_k)\Big) \tag{4}$$

where $m=(m_1,\ldots,m_n)\in\mathbb{R}^n$ and $A=\{a_{jk}\}\in\mathbb{R}^{n\times n}$ is a symmetric positive definite matrix. The vector m is called the *mean*, and the matrix $C=\{c_{jk}\}=A^{-1}$ is called the *covariance matrix* of X. One can explicitly compute the Fourier transform of the distribution of the random variable X. Indeed,

$$\begin{aligned}\varphi_X(y_1,\ldots,y_n)=\int_{\mathbb{R}^n}e^{i(y,x)}\,(P\circ X^{-1})(dx)&=\int_{\mathbb{R}^n}e^{i(y,x)_{\mathbb{R}^n}}p_X(x)\,dx\\&=\exp\Big(-\frac{1}{2}\sum_{j,k}y_jc_{jk}y_k+i\sum_j y_jm_j\Big).\end{aligned} \tag{5}$$

EXAMPLE 4.(Discrete random variable, Poisson distribution, Poisson random variable). A random variable X is called a *discrete random variable* if it can be written in the from:

$$X=\sum_i u_i\,\mathbb{I}_{\Omega_i},$$

where $\mathbb{I}_{\Omega_i}$ is the indicator function of Ω_i, Ω_i are disjoint subsets of Ω such that $\cup_i\Omega_i=\Omega$, and u_i are the values that the random variable X takes with a non-zero probability. Namely,

$$P(\Omega_i)=P(\omega:X(\omega)=u_i)>0.$$

A *Poisson random variable* X is a discrete random variable taking non-negative integer values $k=0,1,2,\ldots$ and having a *Poisson probability distribution* (with some parameter λ), i.e.

$$P(X=k)=\frac{e^{\lambda}\,\lambda^k}{k!}.$$

2. Stochastic processes

Definition 2.1. *Let $(\Omega,\mathcal{F},P)$ be a probability space. A* ***stochastic process*** *is a parametrized collection of $\mathbb{R}^n$-valued random variables:*

$$\{X_t\}_{t\in T}$$

defined on the probability space $(\Omega,\mathcal{F},P)$.

T here is a parameter set. Usually it is the halfline $[0, \infty)$, or an interval $[a, b] \subset [0, \infty)$. It can also be the set of non-negative integers. Let us emphasize that in the definition of a stochastic process, for each fixed $t \in T$,

$$\omega \mapsto X_t(\omega)$$

is a random variable. Let us fix $\omega_0 \in \Omega$ and consider the function

$$t \mapsto X_t(\omega_0).$$

This function is called a *path* of the process X_t.

Definition 2.2. *A non-decreasing family $\{\mathcal{F}_t\}$ of sub-σ-algebras of $\mathcal{F}$ is called a **filtration**, i.e. for all $0 \leqslant s < t < \infty$,*

$$\mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F}.$$

Definition 2.3. *A **filtered probability space** $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ is a probability space with a filtration $\{\mathcal{F}_t\}$ of the σ-algebra $\mathcal{F}$.*

Definition 2.4. *A stochastic process X_t is called **adapted** with respect to the filtration $\mathcal{F}_t$ if for every $t \geqslant 0$ the random variable X_t is $\mathcal{F}_t$-measurable.*

Definition 2.5. *The **finite-dimensional distributions** of a process X_t are the measures $\mu_{t_1,\dots,t_k}$ on $\mathbb{R}^{nk}$, $k = 1, 2, \dots$, defined by:*

$$\mu_{t_1,\dots,t_k}(A_1 \times \cdots \times A_k) = P(X_{t_1} \in A_1, \cdots, X_{t_k} \in A_k),$$

where $A_1, \dots, A_k$ are Borel subsets in $\mathbb{R}^n$.

Let stochastic processes X_t and Y_t be defined on probability spaces $(\Omega, \mathcal{F}, P)$ and resp. on $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{P})$.

Definition 2.6. *The processes X_t and Y_t are said to have the same finite-dimensional distributions if for any $k \in \mathbb{N}$, for any choice of $t_1, \dots, t_k \in T$, the finite-dimensional distributions $\mu^X_{t_1,\dots,t_k}$ and $\mu^Y_{t_1,\dots,t_k}$ of X_t and resp. Y_t coinside.*

Let us consider now a converse problem. Suppose we are given a family $\{\nu_{t_1,\dots,t_k}, k \in \mathbb{N}, t_i \in T\}$ of probability measures on $\mathbb{R}^{nk}$. We would like to construct a stochastic process X_t whose finite-dimenstional distributions coincide with $\nu_{t_1,\dots t_k}$.

Theorem 2.7 (Kolmogorov's extension theorem)**.** *For all $t_1, \dots t_k$, $k \in \mathbb{N}$, let $\nu_{t_1,\dots,t_k}$ be probability measures on $\mathbb{R}^{nk}$ satisfying the two following properties (consistency conditions):*

1. $\nu_{t_{\sigma(1)},\dots,t_{\sigma(k)}}(F_1 \times \cdots \times F_k) = \nu_{t_1,\dots t_k}(F_{\sigma^{-1}(1)} \times \cdots \times F_{\sigma^{-1}(k)})$;
   for all permutations σ of $\{1, 2, \dots, k\}$ and for all Borel subsets $F_i \subset \mathbb{R}^n$.

2. $\nu_{t_1,\dots t_k}(F_1 \times \cdots \times F_k) = \nu_{t_1,\dots,t_k,t_{k+1},\dots,t_{k+m}}(F_1 \times \cdots \times F_k \times \underbrace{\mathbb{R}^n \times \cdots \times \mathbb{R}^n}_{m})$

   for all $m \in \mathbb{N}$ and for all Borel subsets $F_i \subset \mathbb{R}^n$.

Then there exists a probability space $(\Omega, \mathcal{F}, P)$ and an $\mathbb{R}^n$-valued stochastic process X_t on it, such that

$$\nu_{t_1,\dots,t_k}(F_1 \times \cdots \times F_k) = P(X_{t_1} \in F_1, \cdots, X_{t_k} \in F_k)$$

for all $t_i \in T$, $k \in \mathbb{N}$, and all Borell subsets $F_i \subset \mathbb{R}^n$.

Proof. Without proof. A proof can be found, for example, in the book by Koralov, L., Sinai, Y. "Theory of probability and random processes", 2007, p. 167. □

Let X_t and Y_t be stochastic proesses given on the same probability space $(\Omega, \mathcal{F}, P)$.

Definition 2.8. *We say that the process Y_t is a **version** or **modification** of X_t if for all $t \in T$*

$$P\{\omega : X_t(\omega) = Y_t(\omega)\} = 1.$$

Note that the finite-dimensional distributions of X_t and Y_t coincide. However, their path properties can be different. If the process Y_t in Definition 2.8 has continuous paths, then it is called a *continuous path modification* of X_t.

Theorem 2.9 (Kolmogorov's continuity theorem)**.** *Suppose that the process X_t, $t \geqslant 0$, satisfies the following assumption: for all $\Lambda > 0$ there exist positive constants M, α, and β such that*

$$\mathbb{E}[|X_s - X_t|^\alpha] \leqslant M|s-t|^{1+\beta},$$

for all $0 \leqslant s, t \leqslant \Lambda$. Then, there exists a continuous path modification of X_t.

Proof. Without proof. A proof can be found in the book by Kunita, H. "Stochastic flows and stochastic differential equations", 1997, p. 31. □

Examples of basic stochastic processes

Brownian motion. Let $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ be a filtered probability space.

Definition 2.10. *A **(standard, one-dimesional) Brownian motion** is a continuous $\mathcal{F}_t$-adapted stochastic process B_t, $t \geqslant 0$, defined on $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ and possessing the following properties:*

1. *$B_0 = 0$ a.s.*

2. *For all $0 \leqslant s < t$, the increment $B_t - B_s$ is independent of $\mathcal{F}_s$.*

3. *For all $0 \leqslant s < t$, the increment $B_t - B_s$ is a Gaussian random variable with mean zero and variance $t - s$*

Poisson point process. A Poisson process N_t on the interval $[0, \infty)$ counts a number of times some premitive event has occured during the time interval $[0, t]$. More precisely, let $(\Omega, \mathcal{F}, P)$ be a probability space. A *Poisson process* with the parameter λ is a process with the following properties:

1. $N_0 = 0$ a.s.

2. N_t is a process with independent increments, i.e. for all $0 \leqslant t_1 \leqslant \cdots \leqslant t_k$, the random variables N_{t_1}, $N_{t_2} - N_{t_1}$, ..., $N_{t_k} - N_{t_{k-1}}$ are independent.

3. For any $0 \leqslant s < t < \infty$, the random variable $N_t - N_s$ has the Poisson distribution with the parameter $\lambda(t-s)$, i.e.

$$P(N_t - N_s = k) = \frac{e^{-\lambda(t-s)}(\lambda(t-s))^k}{k!}.$$

3. Brownian Motion

3.1 Existence of a Brownian motion

A Brownian motion is a stochastic process named after the Scottish botanist Robert Brown who observed that pollen grains suspended in liquid permormed an irregular motion. To describe this motion mathematically it is natural to introduce a stochastic process $B_t(\omega)$ and interpret it as the position of the pollen grain ω at time t.

We intend to specify a family $\{\nu_{t_1,\ldots,t_k}\}$ of probability measures satusfying the consitency conditions (1) and (2) of the Kolmogorov entension theorem. The latter theorem will imply the existence of a stochstic processes with the finite-dimensional distributions. $\{\nu_{t_1,\ldots,t_k}\}$.

Fix an $x \in \mathbb{R}^n$, and for all $y \in \mathbb{R}^n$ and $t > 0$ define a function:

$$p(t, x, y) = \frac{1}{(2\pi t)^{-\frac{n}{2}}} \exp\Big(-\frac{|x-y|^2}{2t}\Big).$$

If $t = 0$ we set $p(0, x, y) = \delta_x(y)$. The generalized function $\delta_x(y)$ is defined on continuous functions as follows: $\int_{\mathbb{R}^n} f(x)\delta(x)\,dx = f(0)$.

Let $0 \leqslant t_1 \leqslant t_2 \leqslant \cdots \leqslant t_k$. We define a measure $\nu_{t_1,\ldots,t_k}$ on $\mathbb{R}^{nk}$ by

$$\begin{aligned}&\nu_{t_1,\ldots,t_k}(F_1 \times \cdots \times F_k) \\ &\quad = \int_{F_1} p(t_1, x, x_1)\,dx_1 \int_{F_2} p(t_2 - t_1, x_1, x_2)\,dx_2 \ldots \int_{F_k} p(t_k - t_{k-1}, x_{k-1}, x_k)\,dx_k.\end{aligned}$$

If the sequence of $\{t_1, t_2, \ldots, t_k\}$ is not arranged in the ascending order, then we find a permutation σ that puts $\{t_1, t_2, \ldots, t_k\}$ in the the ascending order, and define $\nu_{t_1,\ldots,t_k}$ by formula (1) of the Kolmogorov extension theorem. Therefore by construction, the system of measures $\nu_{t_1,\ldots,t_k}$ satisfies consistency condition (1) of the latter theorem. Since $\int_{\mathbb{R}^n} p(t, x, y)\,dy = 1$, then condition (2) of the Kolmogorov extension theorem holds as well. Now the Kolmogorov extension theorem implies the existence of a probability space $(\Omega, \mathcal{F}, P^x)$ and a stochastic process B_t, $t \geqslant 0$, on it whose finite-dimensional distributions are the measures $\nu_{t_1,\dots,t_k}$. Specifically, if $0 \leqslant t_1 \leqslant t_2 \leqslant \dots \leqslant t_k$, then

$$\begin{aligned} &P(B_{t_1} \in F_1, \dots, B_{t_k} \in F_k) \\ &= \int_{F_1} p(t_1, x, x_1)\, dx_1 \int_{F_2} p(t_2 - t_1, x_1, x_2)\, dx_2 \dots \int_{F_k} p(t_k - t_{k-1}, x_{k-1}, x_k)\, dx_k. \end{aligned} \tag{6}$$

Note that $P^x(B_0 = x) = 1$. The process B_t constructed above is not unique. There exists several quadruples $(B_t, \Omega, \mathcal{F}, P^x)$ such that relation (6) holds. However, it is possible to choose a version of the process B_t with a.s. continuous paths. For this purpose we have to verify Kolmogorov's continuity theorem. We are going to use the formula

$$\mathbb{E}^x[|B_t - B_s|^4] = n(n+2)|t-s|^2 \tag{7}$$

which follows from some properties of the process B_t that we will prove below. The symbol $\mathbb{E}^x$ denotes the expectation with respect to P^x. Kolmogorov's continuity theorem and (7) imply that there is a version of B_t with a.s. continuous paths.

Definition 3.1. *A version of B_t with a.s. continuous paths is called a **Brownian motion** starting at x.*

We note once again that formula (7) is a consequence of the properties of a Brownian motion proved in Section (3.2). Our arguments can be summarized in the theorem below:

Theorem 3.2. *A Brownian motion exists.*

3.2 Some properties of a Brownian motion

Proposition 3.3. *B_t is a Gaussian process, i.e. for all $0 \leqslant t_1, \leqslant \dots \leqslant t_k$, the $\mathbb{R}^{nk}$-valued random variable $Z = (B_{t_1}, \dots, B_{t_k})$ has a multi-normal distribution.*

Proof. To show that the random variable Z has a multi-normal distribution it suffices to prove that there exists a vector $M \in \mathbb{R}^{nk}$ and a non-negative matrix $C = \{c_{jm}\} \in \mathbb{R}^{nk \times nk}$ auch that

$$\mathbb{E}^x\Big[\exp\Big(i\sum_{j=1}^{nk} u_j Z_j\Big)\Big] = \exp\Big(-\frac{1}{2}\sum_{j,m=1}^{nk} u_j c_{jm} u_m + i\sum_{j=1}^{nk} u_j M_j\Big) \tag{8}$$

for all $u = (u_1, \dots, u_{nk}) \in \mathbb{R}^{nk}$. $\mathbb{E}^x$ denotes the expectation with respect to P^x. As we have shown in Example 3 of Section 1 (Formula (5)), the right hand side of (8) is actually the Fourier transform of a density function p_Z of nk variables defined by (4). The Fourier tranform determines this function uniquely, and therefore the distribution of the random variable Z is given by the density function (4) where the number of variables is nk. $\square$

Proposition 3.4. *B_t possesses the following properties:*

$$\begin{aligned}
&\mathbb{E}^x[B_t] = x,\\
&\mathbb{E}^x[\,|B_t - x|^2] = nt,\\
&\mathbb{E}^x[\,(B_t - x, B_s - x)_{\mathbb{R}^n}] = n\min\{s,t\},\\
&\mathbb{E}^x[\,|B_t - B_s|^2] = n|t-s|.
\end{aligned}$$

Proof. It suffices to prove the above properties for the one-dimensional case. Therefore, without loss of generality we assume that the process B_t is one-dimensional. In the following computation we assume that $s < t$. We obtain:

$$\begin{aligned}
&\mathbb{E}^x[B_t] = \frac{1}{\sqrt{2\pi t}}\int_{-\infty}^{\infty} y\, e^{-\frac{(y-x)^2}{2t}}\,dy = x,\\
&\mathbb{E}^x[\,(B_t - x)^2] = \frac{1}{\sqrt{2\pi t}}\int_{-\infty}^{\infty} (y-x)^2\, e^{-\frac{(y-x)^2}{2t}}\,dy = t,\\
&\mathbb{E}^x[\,(B_t - x)(B_s - x)] = \frac{1}{\sqrt{2\pi s}}\int_{-\infty}^{\infty} (y_1-x)\, e^{-\frac{(y_1-x)^2}{2s}}\,dy_1\\
&\frac{1}{\sqrt{2\pi(t-s)}}\int_{-\infty}^{\infty} (y_2-x)\, e^{-\frac{(y_2-y_1)^2}{2(t-s)}}\,dy_2\\
&= \frac{1}{\sqrt{2\pi s}}\int_{-\infty}^{\infty} (y_1-x)^2\, e^{-\frac{(y_1-x)^2}{2s}}\,dy_1 = s,\\
&\mathbb{E}^x[\,(B_t - B_s)^2] = \mathbb{E}^x[(B_t-x)^2] + \mathbb{E}^x[\,(B_s-x)^2] - 2\,\mathbb{E}^x[\,(B_t-x)(B_s-x)] = t - s.
\end{aligned}$$

□

Definition 3.5. *A stochastic process X_t is called* ***stationary****, if for every $t > 0$, the process $\{X_{t+h}\}_{h>0}$, has the same distrbution.*

Proposition 3.6. *B_t has* ***stationary increments****, i.e. for any fixed t, the processes $\{B_{t+h} - B_t\}_{h>0}$, have the same distibution as $B_t - x$.*

Proof. Left as an exercise. □

For the subsequent property of a Brownian motion we will need the theorem below.

Theorem 3.7. *Two random variables X and Y are independent if and only if for any λ and μ*

$$\phi_{(X,Y)}(\lambda,\mu) = \phi_X(\lambda)\,\phi_Y(\mu)$$

where ϕ_X, ϕ_Y and $\phi_{(X,Y)}$ are the characteristic functions of the random variables X, Y, and (X,Y), respectively.

Proof. Without proof. □

Proposition 3.8. *B_t has* ***independent increments****, i.e. for all $0 \leqslant t_1 \leqslant \cdots \leqslant t_k$ the random variables*

$$B_{t_1} - x, B_{t_2} - B_{t_1}, \ldots, B_{t_k} - B_{t_{k-1}} \tag{9}$$

are independent.

Proof. Let $0 \leqslant t_1 < t_2 < \cdots < t_n$ be a partition. For any real numbers λ_k, $1 \leqslant k \leqslant n$, we have:

$$\begin{aligned}
\mathbb{E} \exp\big(\sum_{k=1}^{n} \lambda_k (B_{t_k} - B_{t_{k-1}})\big) &= \mathbb{E}[\mathbb{E} \exp\big(\sum_{k=1}^{n} \lambda_k (B_{t_k} - B_{t_{k-1}})\big) | \mathcal{F}_{t_{n-1}}] \\
&= \mathbb{E}\Big[\exp\big(\sum_{k=1}^{n-1} \lambda_k (B_{t_k} - B_{t_{k-1}})\big) \mathbb{E}\big[\exp\big(\lambda_n (B_{t_n} - B_{t_{n-1}})\big) | \mathcal{F}_{t_{n-1}}\big]\Big] \\
&= \mathbb{E}\big[\exp\big(\lambda_n (B_{t_n} - B_{t_{n-1}})\big)\big] \mathbb{E} \exp\big(\sum_{k=1}^{n-1} \lambda_k (B_{t_k} - B_{t_{k-1}})\big).
\end{aligned}$$

Here we used the properties of the conditional expectation. Namely, that for $k \leqslant n-1$, the random variables $B_{t_k} - B_{t_{k-1}}$ are $\mathcal{F}_{t_{n-1}}$-measurable, and that $B_{t_n} - B_{t_{n-1}}$ is independent of $\mathcal{F}_{t_{n-1}}$. We repeat the above argument inductively to conclude

$$\mathbb{E} \exp\big(\sum_{k=1}^{n} \lambda_k (B_{t_k} - B_{t_{k-1}})\big) == \prod_{k=1}^{\infty} \mathbb{E}\big[\exp\big(\lambda_k (B_{t_k} - B_{t_{k-1}})\big)\big].$$

By Theorem 3.7, the random variables $B_{t_k} - B_{t_{k-1}}$ are independent. □

Note that Propositions 3.3–3.8 were proved without any assumption on the continuity of paths of B_t. In Section 4.6 we left formula (7) without proof. Now we prove this formula.

Proposition 3.9. *Equality (7) holds.*

Proof. Let $t > s \geqslant 0$, and let $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$. Note that the processes $B_t - B_s$ and $\tilde{B}_{t-s} = B_{s+(t-s)} - B_s$, starting at 0 at time s, have the same probability distribution, i.e. $\tilde{B}_{t-s}$, $t > s$, is a Brownian motion (Proposition 3.6). Hence,

$$\begin{aligned}
\mathbb{E}^x[|B_t - B_s|^4] = \mathbb{E}^0[|\tilde{B}_{t-s}|^4] &= \frac{1}{(2\pi(t-s))^{\frac{n}{2}}} \int_{\mathbb{R}^n} |x|^4 \, e^{-\frac{|x|^2}{2(t-s)}} \, dx \\
&= \sum_{i=1}^{n} \frac{1}{(2\pi(t-s))^{\frac{1}{2}}} \int_{-\infty}^{\infty} x_i^4 \, e^{-\frac{x_i^2}{2(t-s)}} \, dx_i \\
+ \sum_{i \neq j} \frac{1}{(2\pi(t-s))^{\frac{1}{2}}} \int_{-\infty}^{\infty} x_i^2 \, e^{-\frac{x_i^2}{2(t-s)}} \, dx_i & \frac{1}{(2\pi(t-s))^{\frac{1}{2}}} \int_{-\infty}^{\infty} x_j^2 \, e^{-\frac{x_j^2}{2(t-s)}} \, dx_j \\
&= 3n(t-s)^2 + n(n-1)(t-s)^2 = n(n+2)(t-s)^2.
\end{aligned}$$

We just used the following facts: $|x|^4 = (x_1^2 + \cdots + x_n^2)^2 = \sum_{i=1}^n x_i^4 + \sum_{i \neq j} x_i^2 x_j^2$, $\frac{1}{(2\pi(t-s))^{\frac{1}{2}}} \int_{-\infty}^{\infty} x_i^2 \, e^{-\frac{x_i^2}{2(t-s)}} \, dx_i = (t-s)$, as well as the computation of the integral below:

$$\begin{aligned}
\frac{1}{(2\pi(t-s))^{\frac{1}{2}}} \int_{-\infty}^{\infty} x_i^4 \, e^{-\frac{x_i^2}{2(t-s)}} \, dx_i &= -\frac{1}{(2\pi(t-s))^{\frac{1}{2}}} \, (t-s) \int_{-\infty}^{\infty} x_i^3 d\Big(e^{-\frac{x_i^2}{2(t-s)}}\Big) \\
= -\frac{1}{(2\pi(t-s))^{\frac{1}{2}}} \, (t-s) \, x_i^3 \, e^{-\frac{x_i^2}{2(t-s)}} \Big|_{-\infty}^{\infty} &+ 3(t-s) \, \frac{1}{(2\pi(t-s))^{\frac{1}{2}}} \int_{-\infty}^{\infty} x_i^2 \, e^{-\frac{x_i^2}{2(t-s)}} \, dx_i \\
&= 3(t-s)^2.
\end{aligned}$$

□

3.3 Local properties of a Brownian path

Here we list some properties of a Brownian path without proof. The proofs can be found in the books of Karatzas and Shreve (1991), Revuz and Yor (2001).

Hölder continuity. For almost all $\omega \in \Omega$, the function $t \mapsto B_t$ is locally Hölder continuous of order α for every $\alpha < \frac{1}{2}$. In other words, for all $T > 0$, $0 < \alpha < \frac{1}{2}$, and for almost all $\omega \in \Omega$ there exists a constant $C_{T,\alpha}(\omega)$ such that for all $s, t \in [0, T)$

$$|B_t(\omega) - B_s(\omega)| \leqslant C_{T,\alpha}(\omega) \, |t-s|^{\alpha}.$$

Modulus of continuity. For Brownian paths

$$\overline{\lim}_{\delta \to 0} \sup_{\substack{s,t<T: \\ |t-s|<\delta}} \frac{|B_t - B_s|}{\sqrt{2\delta \ln(1/\delta)}} = 1 \quad \text{a.s.}$$

In general the *modulus of continuity* of a continuous function f on $[0, T]$ is a function $\varepsilon_f(\delta)$ defined by the formula

$$\varepsilon_f(\delta) = \sup\{|f(t) - f(s)| : t, s \in [0, T), \, |t-s| < \delta\}$$

for sufficiently small δ. Thus, for almost every Brownian path $B_\bullet(\omega)$

$$\varepsilon_{B_\bullet(\omega)}(\delta) \leqslant \sqrt{2\delta \ln(1/\delta)}$$

for all $\delta < \delta_0$ where δ_0 is sufficiently small.

Nowhere differentiability. Brownian paths are a.s. nowhere locally Hölder continuous of order $\alpha > \frac{1}{2}$. In particular, Brownian paths are nowhere differentiable.

Infinite variation. Brownian paths are of infinite variation on any interval $[s,t]$ a.s., i.e. a.s.

$$\sup \sum_{i=1}^{n} |B_{t_i} - B_{t_{i-1}}| = \infty$$

where the supremum is taken over all subdivisions $s \leqslant t_1 \leqslant \cdots \leqslant t_n \leqslant t$ of the interval $[s,t]$.

4. Markov Processes

Intuitevely speaking, the process X_t is Markov if, to make a prediction at time t on what is going to happen in the future, it is useless to know the whole past up to time t but only the present state X_t at time t.

4.1 Continuous-time Markov processes

Definition 4.1. *Let $\mathfrak{B}$ be the Borel σ-algebra of subsets of $\mathbb{R}^n$. A function $\pi : \mathbb{R}^n \times \mathfrak{B} \to [0,1]$ is called a **transition probability** if*

1. *for every $x \in \mathbb{R}^n$, the map $A \mapsto \pi(x, A)$ is a probability measure on $\mathbb{R}^n$;*
2. *for every $A \in \mathfrak{B}$, the map $x \mapsto \pi(x, A)$ is $\mathfrak{B}$-measurable.*

Definition 4.2. *A **transition function** on $(\mathbb{R}^n, \mathfrak{B})$ is a family $P_{s,t}$, $0 \leqslant s < t$, of transition probabilities on $(\mathbb{R}^n, \mathfrak{B})$ such that for every three numbers $s < r < t$,*

$$\int_{\mathbb{R}^n} P_{s,r}(x, dy) P_{r,t}(y, A) = P_{s,t}(x, A) \tag{10}$$

for every $x \in \mathbb{R}^n$ and for every $A \in \mathfrak{B}$.

Relation (10) is called the *Chapman-Kolmogorov* equation. The transition function is said to be *homogeneous* if $P_{s,t}$ depends on and s and t only through the difference $t-s$, i.e. $P_{s,t} = P_{s-t}$, where P_t is the notation for $P_{0,t}$. Equation (10) takes the form:

$$P_{s+t}(x, A) = \int P_s(x, dy) P_t(y, A) \tag{11}$$

for every $s, t \geqslant 0$.

Definition 4.3. *Let $(\Omega, \mathcal{F}, P)$ be a probability space. A process X_t is a Markov process with transition function $P_{s,t}$ if for any Borel-measurable function $f : \mathbb{R}^n \to \mathbb{R}$,*

$$\mathbb{E}[f(X_t) \,|\, \sigma(X_u, u \leqslant s)] = (P_{s,t} f)(X_s) \quad P-a.s.$$

In the above definition, $P_{s,t}f$ is a function $\mathbb{R}^n \to \mathbb{R}$, such that for any $x \in \mathbb{R}^n$ it is defined as follows:

$$(P_{s,t}f)(x) = \int_{\mathbb{R}^n} f(y)\, P_{s,t}(x, dy). \tag{12}$$

Definition 4.4. *The process X_t is said to be homogeneous if its transition function is homogeneous.*

For a homogeneous Markov process

$$\mathbb{E}[f(X_t)\,|\,\sigma(X_u, u \leqslant s)] = (P_{t-s}f)(X_s) \quad P-\text{a.s.}$$

In the following, let $\nu = P \circ X_0^{-1}$ denote the initial distribution of the process X_t.

Theorem 4.5. *A process X_t is a Markov process with transition function $P_{s,t}$ and initial measure μ if and only if for any sequence $0 = t_0 < t_1 < \cdots t_k$ and and for any Borel-measurable functions $f_i : \mathbb{R}^n \to \mathbb{R}$, $0 \leqslant i \leqslant k$,*

$$\begin{aligned}
&\mathbb{E}\Big[\prod_{i=1}^{k} f_i(X_{t_i})\Big] \\
&= \int_{\mathbb{R}^n} \nu(dx_0) f_0(x_0) \int_{\mathbb{R}^n} P_{0,t_1}(x_0, dx_1) f_1(x_1) \ldots \int_{\mathbb{R}^n} P_{t_{k-1},t_k}(x_{k-1}, dx_k) f_k(x_k).
\end{aligned} \tag{13}$$

Proof. We will only prove the "if" statement, i.e. we prove that if the process X_t is Markov, then formula (13) holds. The converse statement, i.e. if formula (13) holds then the process X_t is Markov, will be left without proof. We have:

$$\begin{aligned}
\mathbb{E}\left[\prod_{i=0}^{k} f_i(X_{t_i})\right] &= \mathbb{E}\left[\prod_{i=0}^{k-1} f_i(X_{t_i})\, \mathbb{E}\big[f_k(X_{t_k})\,|\,\sigma(X_s, s \leqslant t_{k-1})\big]\right] \\
&= \mathbb{E}\left[\prod_{i=0}^{k-1} f_i(X_{t_i})\, (P_{t_{k-1},t_k} f_k)(X_{t_{k-1}})\right].
\end{aligned} \tag{14}$$

When passing from the first to the second expectation in (14) we used the fact that $F(X_{t_i})$, $0 \leqslant i \leqslant k-1$, are $\sigma(X_s, s \leqslant t_{k-1})$-measurable. The latter expectation in (14) is of the product of $k-1$ functions but has the same form as the first expectation of the product of k functions. The $(k-1)$th function in the latter product is $f_{k-1}\, P_{t_{k-1},t_k} f_k$. We can proceed the same way to come to a product of $k-2$ functions, etc. After the last step we obtain:

$$\mathbb{E}\left[\prod_{i=0}^{k} f_i(X_{t_i})\right] = \mathbb{E}\left[f(X_0) P_{0,t_1}(f_1 P_{t_1,t_2}(f_2 P_{t_2,t_3} \ldots f_{k-1}(P_{t_{k-1}t_k} f_k) \ldots))(X_0)\right]. \tag{15}$$

Using equality (12) as well as the fact that for any Borel-measurable function $f : \mathbb{R}^n \to \mathbb{R}$

$$\mathbb{E}[f(X_0)] = \int_{\mathbb{R}^n} f(x)\, \nu(dx),$$

we conclude that the right-hand side of (15) equals to the right-hand side of (13). The proof of the converse statement, i.e. if formula (13) holds then the process X_t is Markov, will be left without proof. A proof can be found in the book by D. Revuz and M. Yor "Continuous martingales and Brownian motion", Chapter III. $\square$

Note that identity (13) shows that if we know the transition function of a Markov process, we can construct its finite-dimensional distributions.

Theorem 4.6 (Existence of a Markov process)**.** *Given a transition function $P_{s,t}$ on $(\mathbb{R}^n, \mathfrak{B})$ and a probability measure μ on $(\mathbb{R}^n, \mathfrak{B})$, there exists a probability space $(\Omega, \mathcal{F}, P_\mu)$ and a stochastic process X_t on it which is Markov with transition function $P_{s,t}$ and initial measure μ.*

Proof. Let us define the finite-dimensional distributions. Fix a sequence $0 < t_1 < t_2 < t_k$, and define:

$$\nu_{t_1,\dots,t_k}(F_1 \times \dots \times F_k) = \int_{F_1} P_{0,t_1}(x, dx_1) \int_{F_2} P_{t_1,t_2}(x_1, dx_2) \dots \int_{F_k} P_{t_{k-1},t_k}(x_{k-1}, dx_k).$$

$$\nu_{0,t_1,\dots,t_k}(F_0 \times F_1 \times \dots \times F_k) = \int_{F_0} \mu(dx_0) \int_{F_1} P_{0,t_1}(x_0, dx_1) \int_{F_2} P_{t_1,t_2}(x_1, dx_2) \\ \dots \int_{F_k} P_{t_{k-1},t_k}(x_{k-1}, dx_k).$$

Let $\{t_1, t_2, \dots, t_k\}$ be an arbitrary sequence (not necessary ascending). Let σ be the permutation of $\{1, \dots, k\}$ such that $t_{\sigma(1)} < t_{\sigma(2)} < \dots < t_{\sigma(k)}$. Define

$$\nu_{t_1,\dots,t_k}(F_1 \times \dots \times F_k) = \nu_{t_{\sigma(1)},\dots,t_{\sigma(k)}}(F_{\sigma(1)} \times \dots \times F_{\sigma(k)}).$$

The latter definition implies consistency condition 1 of the Kolmogorov extension theorem. Further let $\tilde{\sigma}$ be the permutation of $\{1, \dots, k, k+1, \dots, k+m\}$ such that $t_{\tilde{\sigma}(1)} < \cdots t_{\tilde{\sigma}(k)} < t_{\tilde{\sigma}(k+1)} < \dots < t_{\tilde{\sigma}(k+m)}$. Define $F_i = \mathbb{R}^n$ for $k+1 \leqslant i \leqslant k+m$. We obtain:

$$\begin{aligned}
&\nu_{t_1,\dots,t_k,t_{k+1},\dots,t_{k+m}}(F_1 \times \dots \times F_k \times \underbrace{\mathbb{R}^n \times \dots \times \mathbb{R}^n}_{m}) = \nu_{t_1,\dots,t_k,t_{k+1},\dots,t_{k+m}}(F_1 \times \dots \times F_{k+m}) \\
&= \nu_{t_{\tilde{\sigma}(1)},\dots,t_{\tilde{\sigma}(k)},t_{\tilde{\sigma}(k+1)},\dots,t_{\tilde{\sigma}(k+m)}}(F_{\tilde{\sigma}(1)} \times \dots \times F_{\tilde{\sigma}(k+m)}) = \nu_{t_{\sigma(1)},\dots,t_{\sigma(k)}}(F_{\sigma(1)} \times \dots \times F_{\sigma(k)}) \quad (16) \\
&= \nu_{t_1,\dots,t_k}(F_1 \times \dots \times F_k).
\end{aligned}$$

The third equality in (16) holds by the Chapman-Kolmogorov equations. Indeed, some of the sets $F_{\tilde{\sigma}(i)}$, $1 \leqslant i \leqslant k+m$, equal to $\mathbb{R}^n$ (namely, those that do not coincide with one of the sets $F_{\sigma(i)}$, $1 \leqslant i \leqslant k$), and therefore integration over these sets can be excluded. Specifically, we exclude integration over those $F_{\tilde{\sigma}(i)}$ which are equal to $\mathbb{R}^n$ as follows: for any three numbers $t_{i-1} < t_i < t_{i+1}$, for any function $f : \mathbb{R}^n \to \mathbb{R}$, for any Borel subset $F \subset \mathbb{R}^n$,

$$\begin{aligned}
&\int_{\mathbb{R}^n} P_{t_{i-1},t_i}(x_{i-1}, tdx_i) \int_F P_{t_i,t_{i+1}}(x_i, dx_{i+1}) f(x_{i+1}) \\
&= \int_F f(x_{i+1}) \int_{\mathbb{R}^n} P_{t_{i-1},t_i}(x_{i-1}, dx_i) P_{t_i,t_{i+1}}(x_i, dx_{i+1}) = \int_F f(x_{i+1}) P_{t_{i-1},t_{i+1}}(x_{i-1}, dx_{i+1})
\end{aligned}$$

and therefore, time t_i and the integration over $\mathbb{R}^n$ are excluded. This proves (16) and consistency condition 2 of the Kolmogorov extension theorem. The latter theorem implies the existence of a probability space and a stochastic processes X_t on it whose finite-dimensional distributions are the measures $\nu_{t_1,\dots,t_k}$. Let us prove that X_t is a Markov process. We will apply Theorem 4.5. Indeed, let $f_i : \mathbb{R}^n \to \mathbb{R}$ be Borel-measurable functions. We obtain:

$$
\begin{aligned}
\mathbb{E}\Big[\prod_{i=0}^{k} f_i(X_{t_i})\Big] &= \int_{\mathbb{R}^{nk}} \prod_{i=0}^{k} f_i(x_i)\, \nu_{0,t_1,\dots,t_k}(dx_0 dx_1 \cdots dx_k) \\
&= \int_{\mathbb{R}^n} f_0(x_0)\mu(dx_0) \int_{\mathbb{R}^n} f_1(x_1) P_{0,t_1}(x_0, dx_1) \dots \int_{\mathbb{R}^n} f_k(x_k)\, P_{t_{k-1},t_k}(x_{k-1}, dx_k).
\end{aligned}
$$

By Theorem 4.5 the process X_t is Markov. $\square$

4.2 Markov property

Let $(\mathbb{R}^n, \mathfrak{B})$, where $\mathfrak{B}$ is the Borel σ-algebra of subsets of $\mathbb{R}^n$, be a measurable space. Further let $\Omega = (\mathbb{R}^n)^{[0,\infty)} = \{\varphi : [0,\infty) \to \mathbb{R}^n\}$, let $\pi_t : \Omega \to \mathbb{R}^n$, $\pi_t(\varphi) = \varphi(t)$ be the evaluation map, and let the σ-algebra of subsets of Ω be defined as follows: $\mathcal{F} = \big(\mathfrak{B}\big)^{[0,\infty)} = \sigma\{F \subset \Omega : \pi_t(F) \in \mathfrak{B} \ \forall t \in [0,\infty)\}$. Finally, let X_t be the *coordinate process*, i.e. $X_t(\omega) = \omega(t)$. The theorem below is similar to Theorem 4.6 but states the existence of a probability measure P and a Markov process on a certain probability space which is $(\Omega, \mathcal{F}, P)$.

Theorem 4.7. *Given a transition function $P_{s,t}$ on $(\mathbb{R}^n, \mathfrak{B})$, for any probability measure on $(\mathbb{R}^n, \mathfrak{B})$, there exists a unique probability measure P_μ on $(\Omega, \mathcal{F})$ such that the coordinate process X_t is Markov with transition function $P_{s,t}$ and initial measure μ.*

Proof. Without proof. A proof can be found in the book by D. Revuz and M. Yor "Continuous martingales and Brownian motion", Chapter III. $\square$

Let $(\Omega, \mathcal{F})$ be a measurable space. Note that in Theorem 4.6 to an initial measure μ we associated a probability measure P_μ. For the probability measure P_{δ_x} which is associated to initial measure δ_x we will use the notation P_x. Furthermore, the symbols $\mathbb{E}_\mu$ and $\mathbb{E}_x$ will be used for the expectations relative to the measures P_μ and resp. P_x.

Proposition 4.8. *Let Z be an $\mathcal{F}$-measurable random variable which is either positive or bounded. Then,*

$$\mathbb{E}_\mu[Z] = \int_{\mathbb{R}^n} \mu(dx)\, \mathbb{E}_x[Z].$$

Proof. Without proof. A proof can be found in the book by D. Revuz and M. Yor "Continuous martingales and Brownian motion", Chapter III. $\square$

Below, $(\Omega, \mathcal{F}, P_\mu)$ with the objects defined above is fixed as a probability space. *From now on, unless otherwise is stated, we will consider only homogeneous transition functions and associated homogeneous Markov processes. In this case we have:*

$$
\begin{aligned}
&P_\mu(X_0 \in A_0, X_{t_1} \in A_1, \ldots, X_{t_n} \in A_n) \\
&\qquad = \int_{A_0} \mu(dx) \int_{A_1} P_{t_1}(x, dx_1) \int_{A_2} P_{t_2 - t_1}(x_1, dx_2) \ldots \int_{A_n} P_{t_n - t_{n-1}}(x_{n-1}, dx_n).
\end{aligned}
$$

For every $t > 0$, we define the *shift operator* $\theta_t : \Omega \to \Omega$ as follows:

$$
\theta_t(\omega) = \omega(t + \,\cdot\,), \quad \text{i.e. } \big(\theta_t(\omega)\big)(s) = \omega(t + s) \quad \forall s > 0.
$$

Theorem 4.9 (Markov property)**.** *Let Z be an $\mathcal{F}$-measurable random variable such that it is either positive or bounded. Then*

$$
\mathbb{E}_\nu[Z \circ \theta_t \,|\, \sigma(X_s, s \leqslant t)] = \mathbb{E}_{X_t}[Z].
$$

The expectation of the right-hand side of this formula is understood as $\mathbb{E}_x[Z]$ with X_t substituted for x. In other words one can say that it is the composition of two maps: $\omega \mapsto X_t(\omega)$ and $x \mapsto \mathbb{E}_x[Z]$.

Idea of proof. By the definition of conditional probability, we have to show that for any subset $A \in \sigma(X_s, s \leqslant t)$,

$$
\int_A (Z \circ \theta_t)(\omega)\, P_\mu(d\omega) = \int_A E_{X_t}[Z](\omega)\, P_\mu(d\omega).
$$

A more general statement would be to prove

$$
\mathbb{E}_\mu[(Z \circ \theta_t)\, Y] = \mathbb{E}_\mu[E_{X_t}[Z]\, Y]
$$

for any $\sigma(X_s, s \leqslant t)$-measurable and positive random variable Y. Consider the case when $Y = \prod_{i=0}^k f_i(X_{t_i})$ and $Z = \prod_{j=1}^m g_j(X_{\tau_j})$, where f_i and g_j are positive Borel-measurable functions and $0 = t_0 < t_1 < \cdots < t_k = t$, $0 < \tau_1 < \cdots < \tau_m$. Without loss of generality we can assume that $t_0 = 0$ and $t_k = t$ multiplying Y by functions identically equal to 1 if necessary. We obtain:

$$
\begin{aligned}
&\mathbb{E}_\mu\Big[Y\,(Z \circ \theta_t)\Big] = \mathbb{E}_\mu\Big[\prod_{i=0}^k f_i(X_{t_i}) \prod_{j=0}^m g_j(X_{t+\tau_j})\Big] \\
&\quad = \int_{\mathbb{R}^n} \mu(dx_0) f_0(x_0) \int_{\mathbb{R}^n} P_{t_1}(x_0, dx_1) f_1(x_1) \ldots \int_{\mathbb{R}^n} P_{t_k - t_{k-1}}(x_{k-1}, dx_k) f_k(x_k) \times \\
&\qquad \times \int_{\mathbb{R}^n} P_{t_{k+1} - t}(x_k, dx_{k+1}) g_0(x_{k+1}) \ldots \int_{\mathbb{R}^n} P_{\tau_m - \tau_{m-1}}(x_{m-1}, dx_m) g_m(x_m) \\
&\qquad\qquad = \int_{\mathbb{R}^n} \mu(dx_0) f_0(x_0) \int_{\mathbb{R}^n} P_{t_1}(x_0, dx_1) f_1(x_1) \ldots \\
&\ldots \int_{\mathbb{R}^n} P_{t - t_{k-1}}(x_{k-1}, dx_k) f_k(x_k)\, \mathbb{E}_{x_k}\Big[\prod_{j=1}^m g(X_{\tau_j})\Big] = \mathbb{E}_\mu\Big[\prod_{i=0}^k f_i(X_{t_i})\, \mathbb{E}_{X_t} \prod_{j=0}^m g_j(X_{\tau_j})\Big] \\
&\qquad\qquad\qquad = \mathbb{E}_\mu\Big[Y\, \mathbb{E}_{X_t}[Z]\Big].
\end{aligned}
$$

The general case that deals an arbitrary $\mathcal{F}$-measurable random variable Z and a $\sigma(X_s, s \leqslant t)$-measurable random variable Y we leave without proof. A proof can be found in the book by D. Revuz and M. Yor "Continuous martingales and Brownian motion", Chapter III. $\square$

5. Martingales

5.1 Filtrations and Stopping times

Filtrations play a fundamental role in the theory of stochastic processes, namely, in the definition of the basic objects of our consideration - martingales. In Section 2 we introduced the concept of a *filtration* and defined processed *adapted* with respect to filtration. Let us introduce further definitions. With every filtration $\mathcal{F}_t$ we associate two other filtrations:

$$\mathcal{F}_{t^-} = \sigma\Big(\bigcup_{s<t} \mathcal{F}_s\Big), \qquad \mathcal{F}_{t^+} = \bigcap_{\varepsilon>0} \mathcal{F}_{t+\varepsilon}.$$

The filtration $\mathcal{F}_\infty$ denotes

$$\mathcal{F}_\infty = \sigma\Big(\bigcup_{t<\infty} \mathcal{F}_t\Big)$$

Note that

$$\mathcal{F}_\infty = \sigma\Big(\bigcup_{t<\infty} \mathcal{F}_{t^-}\Big) = \sigma\Big(\bigcup_{t<\infty} \mathcal{F}_{t^+}\Big).$$

By convention, $\mathcal{F}_{0^-} = \mathcal{F}_0$. We always have the inclusion $\mathcal{F}_{t^-} \subseteq \mathcal{F}_t \subseteq \mathcal{F}_{t^+}$.

Definition 5.1. *If $\mathcal{F}_t = \mathcal{F}_{t^+}$, the filtration $\mathcal{F}_t$ is called **right-continuous**.*

Let $\mathcal{F}_t$ be a filtration.

Definition 5.2. *A **stopping time** relative to the filtration $\mathcal{F}_t$ is a random variable with values in $[0,\infty)$ such that for every $t \geqslant 0$*

$$\{\omega : T(\omega) \leqslant t\} \in \mathcal{F}_t.$$

Definition 5.3. *Let T be a stopping time. The class of sets $A \in \mathcal{F}_\infty$ such that*

$$A \cap \{T \leqslant t\} \in \mathcal{F}_t \quad \textit{for all}\, t$$

is a σ-algebra which is denoted by $\mathcal{F}_T$.

Let us give examples of stopping times.

Proposition 5.4. *Let $A \subset \mathbb{R}^n$ be an closed set, and let the process X_t have continuous paths. Define*

$$T_A(\omega) = \inf\{t \geqslant 0,\ X_t(\omega) \notin A\}.$$

Then T_A is a stopping time relative to the natural filtration $\mathcal{F}^0_t = \sigma(X_s, s \leqslant t)$.

Remark. The stopping time T_A is called the *exit time* of the process X_t from the set A.

Proof. We have to prove that for every $t \geqslant 0$, $\{T_A \leqslant t\} \in \mathcal{F}_t^0$, or, which is the same, that $\{T_A > t\} \in \mathcal{F}_t^0$. But

$$\{\omega : T_A(\omega) > t\} = \bigcap_{s \in \mathbb{Q},\, s \leqslant t} \{\omega : X_s(\omega) \in A\} \in \mathcal{F}_t^0.$$

□

Proposition 5.5. *Let $A \subset \mathbb{R}^n$ be an open set, and let the process X_t have continuous paths. Define*

$$\tilde{T}_A(\omega) = \inf\{t \geqslant 0,\ X_t(\omega) \in A\}.$$

Then $\tilde{T}_A$ is a stopping time relative to the natural filtration $\mathcal{F}_t^0 = \sigma(X_s, s \leqslant t)$.

Remark. The stopping time $\tilde{T}_A$ is called the *hitting time* of A.

Definition 5.6. *Let X_t be a process and T be a stopping time. The process $X_t^T = X_{t\wedge T}$ is called the **stopped** process.*

Remark. In the above definition $t \wedge T$ denotes $\min\{t, T\}$.

5.2 Definition and examples of martingales

Let $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ be a filtered probability space.

Definition 5.7. *An $\mathcal{F}_t$-adapted process X_t is called a **martingale** if $\mathbb{E}|X_t| < \infty$ and*

$$\mathbb{E}\big[X_t \,|\, \mathcal{F}_s\big] = X_s \quad \textit{a.s. for all}\ \ s \leqslant t.$$

Theorem 5.8. *Let B_t be a one-dimensional Brownian motion. Then, the following processes are martingales with respect to the natural filtration $\mathcal{F}_t^0 = \sigma(B_s, s \leqslant t)$.*

1. *B_t itself;*
2. *$B_t^2 - t$;*
3. *$M_t^\alpha = \exp\Big(\alpha B_t - \frac{\alpha^2}{2}\, t\Big)$ for all $\alpha \in \mathbb{R}$.*

Proof. 1. Left as an exercise.
2. Note that $\mathbb{E}\big|B_t^2 - t\big| \leqslant 2t$. Next,

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

We used the fact that B_t is an $\mathcal{F}_t^0$-martingale. The latter equality implies that a.s.

$$\mathbb{E}\big[B_t^2 - t \,|\, \mathcal{F}_s^0\big] = B_s^2 - s.$$

3. Note that

$$\mathbb{E}\big[\exp(\alpha B_t)\big] = \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{\infty} e^{\alpha x}\, e^{-\frac{x^2}{2t}}\, dx = \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{\infty} e^{-\frac{(x-\alpha t)^2}{2t}}\, e^{\frac{\alpha^2 t}{2}}\, dx = e^{\frac{\alpha^2 t}{2}}. \quad (17)$$

We have:

$$\begin{aligned}\mathbb{E}\big[\exp(\alpha B_t) \,|\, \mathcal{F}_s^0\big] &= \exp(\alpha B_s)\, \mathbb{E}\big[\exp(\alpha(B_t - B_s)) \,|\, \mathcal{F}_s^0\big] \\ &= \exp(\alpha B_s)\, \mathbb{E}\big[\exp(\alpha(B_t - B_s))\big] = \exp(\alpha B_s) \exp\Big(\frac{\alpha^2(t-s)}{2}\Big).\end{aligned}$$

We used relation (17) and the fact that $B_t - B_s$ does not depend on the σ-algebra $\mathcal{F}_s^0$. The latter relation implies:

$$\mathbb{E}\Big[\exp\big(\alpha B_t - \frac{\alpha^2 t}{2}\big) \,|\, \mathcal{F}_s^0\Big] = \exp\big(\alpha B_s - \frac{\alpha^2 s}{2}\big).$$

$\square$

5.3 Discrete-time martingales

In the definition of a martingale one can replace continuous time processes and filtrations with discrete time processes and filtrations. Namely, a discrete time process X_n, $n \in \mathbb{Z}_+$ §($\mathbb{Z}_+ = \mathbb{N} \cup \{0\}$), is called *adapted* with respect to a discrete-time filtration $\mathcal{F}_n$, if for any $n \in \mathbb{Z}_+$, the random variable X_n is $\mathcal{F}_n$-measurable. An adapted discrete-time process X_n is called a *martingale* if $\mathbb{E}|X_n| < \infty$ for any $n \in \mathbb{Z}_+$ and $\mathbb{E}[X_n \,|\, \mathcal{F}_{n-1}] = X_{n-1}$.

Proposition 5.9. *Let X_n, $n \in \mathbb{Z}_+$, be a martingale with respect to a discrete filtration $\mathcal{F}_n$, and let H_n, $n \in \mathbb{Z}_+$, be a positive bounded process such that H_n is $\mathcal{F}_{n-1}$-measurable for $n > 1$. Define the process Y_n:*

$$Y_0 = X_0, \qquad Y_n = Y_{n-1} + H_n(X_n - X_{n-1}).$$

Then, the process Y_n is a martingale. In particular, if T is an integer-valued stopping time, then the stopped process X_n^T is a martingale.

Proof. The first sentence (stating that Y_n is a martingale) can be verified immediately. Let us prove that the stopped process $X_n^T = X_{n\wedge T}$ is a martingale. Take $H_n = \mathbb{I}_{n\leqslant T}$. One can easily verify that if Y_n is constructed with the help of $H_n = \mathbb{I}_{n\leqslant T}$ is exactly the stopped process X_n^T. Note that $H_n = 1 - \mathbb{I}_{\{T\leqslant n-1\}}$, and therefore it is $\mathcal{F}_{n-1}$-measurable. This proves that X_n^T is a martingale. $\square$

Definition 5.10. *The process Y_n defined in Proposition 5.9 will be denoted by $(H \cdot X)_n$.*

Theorem 5.11 (A discrete-time version of the optional stopping theorem)**.** *Let X_n be a martingale and let S and T be two stopping times such that for every ω*

$$S(\omega) \leqslant T(\omega) \leqslant M < \infty,$$

where $M > 0$ is a constant. Then

$$X_S = \mathbb{E}[X_T \,|\, \mathcal{F}_S] \qquad a.s. \tag{18}$$

Proof. Take $H_n = \mathbb{I}_{\{n \leqslant T\}} - \mathbb{I}_{\{n \leqslant S\}}$. If $n > M$, then

$$(H \cdot X)_n - X_0 = X_T - X_S. \tag{19}$$

Indeed, let $H'_n = \mathbb{I}_{\{n \leqslant T\}}$ and let $H''_n = \mathbb{I}_{\{n \leqslant S\}}$. Then, if $n > M$, $(H' \cdot X)_n = X_T$ and $(H'' \cdot X)_n = X_S$. On the other hand, $(H \cdot X)_n - X_0 = (H' \cdot X)_n - (H'' \cdot X)_n$ which proves (19). Next, since $(H \cdot X)_n$ is a martingale, $\mathbb{E}[(H \cdot X)_n] = \mathbb{E}[X_0]$. Hence,

$$\mathbb{E}[X_T] = \mathbb{E}[X_S]. \tag{20}$$

Let us apply equality (20) to the pair of stopping times $S^B = S\,\mathbb{I}_B + M\,\mathbb{I}_{B^c}$ and $T^B = T\,\mathbb{I}_B + M\,\mathbb{I}_{B^c}$, where $B \in \mathcal{F}_S$. The fact that S^B and T^B are stopping times is to be proved as an exercise. Note that $X_{T^B} = X_T\,\mathbb{I}_B + X_M\,\mathbb{I}_{B^c}$ and $X_{S^B} = X_S\,\mathbb{I}_B + X_M\,\mathbb{I}_{B^c}$. Equality (20) implies:

$$\mathbb{E}[X_T\,\mathbb{I}_B + X_M\,\mathbb{I}_{B^c}] = \mathbb{E}[X_S\,\mathbb{I}_B + X_M\,\mathbb{I}_{B^c}],$$

and therefore, for any $\mathcal{F}_S$-measurable set B,

$$\int_B X_T(\omega)\, P(d\omega) = \int_B X_S(\omega)\, P(d\omega)$$

which is equivalent to (18). □

5.4 The optional stopping theorem

Lemma 5.12. *Let T be a stopping time. Define $T_k = +\infty$ if $T \geqslant k$, and $T_k = q2^{-k}$, if $(q-1)2^{-k} \leqslant T < q2^{-k}$, $q < 2^k k$. Then, T_k is a sequence of stopping times such that the stopping time T is the decreasing limit of T_k as $k \to \infty$.*

Proof. As k increases by one, each interval $[(q-1)2^{-k}, q2^{-k}]$ gets divided into two intervals of equal length. This shows that the limit $\{T_k\}$ is decreasing to T as $k \to \infty$. Now note that the set $\{T < \tau\} \in \mathcal{F}_\tau$. Indeed,

$$\{T < \tau\} = \bigcup_{n=1}^{\infty} \Big\{ T \leqslant \tau - \frac{1}{n} \Big\}$$

But $\{T \leqslant \tau - \frac{1}{n}\} \in \mathcal{F}_{\tau - \frac{1}{n}} \subset \mathcal{F}_\tau$, and therefore $\{T < \tau\} \in \mathcal{F}_\tau$. Note that every T_k can be represented as $T_k = ([T\,2^k] + 1)\,2^{-k}$ where $[\,\cdot\,]$ denotes the integer part. Analogously, we define $t_k = ([t\,2^k] + 1)\,2^{-k}$. Clearly, $t_{k-1} \leqslant t < t_k$. We have:

$$\{2^k T_k < 2^k t\} = \{2^k T_k \leqslant 2^k t_{k-1}\} = \{2^k T < 2^k t_{k-1}\} = \{T < t_{k-1}\} \subset \mathcal{F}_{t_{k-1}} \subset F_t$$

which shows that T_k is a stopping time.

□

Theorem 5.13 (Doob's martingale convergence theorem)**.** *Let X_t be a right-continuous $\mathcal{F}_t$-martingale with the property that*

$$\sup_{t>0} \mathbb{E}[X_t^-] < \infty,$$

where $X_t^- = \max(-X_t, 0)$. Then, the pointwise limit

$$X_\infty(\omega) = \lim_{t\to\infty} X_t(\omega)$$

exists a.s., and

$$X_t = \mathbb{E}[X_\infty \,|\, \mathcal{F}_t].$$

Proof. Without proof. A proof can be found in the book by Z. Brzezniak, T. Zastawniak "Basic stochastic processes: a course through exercises", Chapter 4. □

Theorem 5.14. *Suppose $X_n \to X$ where X_n and X are random variables such that a.s. $|X_n| \leqslant Z$ for all n and $\mathbb{E}|Z| < \infty$. Let $\mathcal{F}_n$ be a decreasing (increasing) family of σ-algebras whose intersection is a σ-algebra $\mathcal{F}$ (resp. whose union generates $\mathcal{F}$). Then a.s.*

$$\lim_{n\to\infty} \mathbb{E}[X_n \,|\, \mathcal{F}_n] = \mathbb{E}[X \,|\, \mathcal{F}].$$

Proof. Without proof. A proof can be found in the book by D. Revuz and M. Yor "Continuous martingales and Brownian motion", Chapter II. □

Definition 5.15. *A martingale X_t, $t \geqslant 0$, is called **uniformly integrable** if for any $\varepsilon > 0$ there exists a constant $K > 0$ such that*

$$\sup_{t\geqslant 0} \int_{\{\omega : |X_t(\omega)| > K\}} |X_t(\omega)|\, P(d\omega) < \varepsilon.$$

Theorem 5.16 (Optional stopping theorem)**.** *Let X_t, $t \geqslant 0$, be a continuous martingale, and let $S \leqslant T$ be stopping times such that the stopped process X_t^T is uniformly integrable. Then,*

$$\mathbb{E}[X_T \,|\, \mathcal{F}_S] = X_S. \tag{21}$$

In particular, the stopped process $X_t^T = X_{t\wedge T}$ is a martingale with respect to $\mathcal{F}_{t\wedge T}$.

Proof. Let $S_n = ([S\,2^n] + 1)\,2^{-n}$, $T_n = ([T\,2^n] + 1)\,2^{-n}$ (as in Lemma 5.12). Fix an $n \in \mathbb{N}$ and define a uniformly integrable martingale

$$Y_m = X_{T_n \wedge m 2^{-n}}$$

with respect the filtration $\mathcal{G}_m = \mathcal{F}_{m2^{-n}}$, since $X_{t2^{-n}}$ is an $\mathcal{G}_t = \mathcal{F}_{t2^{-n}}$-martingale. Note that $\tilde{S} = [S\,2^n] + 1$ is a $\mathcal{G}_m$-stopping time since

$$\{\tilde{S} \leqslant m\} = \{S_n \leqslant m2^{-n}\} \in \mathcal{F}_{m2^{-n}} = \mathcal{G}_m.$$

We apply Theorem 5.11 to the pair of stopping times $\tilde{S}$ and $\tilde{T} = [T\,2^n] + 1$. We obtain:

$$\mathbb{E}[X_{T_n} \,|\, \mathcal{F}_{S_n}] = \mathbb{E}[Y_{\tilde{T}} \,|\, \mathcal{G}_{\tilde{S}}] = Y_{\tilde{S}} = X_{S_n}. \tag{22}$$

The first equality in (22) holds because $X_T = Y_{\tilde{T}}$ and $\mathcal{F}_{S_n} = \mathcal{G}_{\tilde{S}}$ which follows directly from the definitions of the process Y_m, the filtration $\mathcal{G}_m$, and the stopping time $\tilde{S}$. The second equality in (22) holds by Theorem 5.11. Finally, the third equality in (22) follows again from the definition of the process Y_m. Taking the limit of the both parts of (22) as $n \to \infty$ and applying Theorem 5.14 we obtain:

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

Finally we conclude that the stopped process X_t^T is a martingale by applying identity (21) to the stopping times $T \wedge s$ and $T \wedge t$ where $s < t$. $\square$

5.5 Local martingales

Definition 5.17. *Let X_t be an adapted process with respect to the filtration $\mathcal{F}_t$. The process X_t is called a* ***local martingale*** *with respect to the filtration $\mathcal{F}_t$ if there exists a sequence of stopping times*

$$0 = T_0 \leqslant T_1 \leqslant \cdots T_n \leqslant \cdots$$

such that $T_n \to \infty$ as $n \to \infty$ and the process $X_t^{T_n}$ is an $\mathcal{F}_t$-martingale.

The sequence T_n is called a *reducing system* of stopping times. Every martingale is a local martingale since every sequence of stopping time increasing to infinity is reducing by Theorem 5.16.

Remark 1. Equivalently in the definition of a local martingale we could require that $X_t^{T_n}$ is a martingale with respect to $\mathcal{F}_{t \wedge T_n}$ instead of $\mathcal{F}_t$.

Proof of the remark. Let $X_t^{T_n}$ be an $\mathcal{F}_t$ martingale, then, by the Optional stopping theorem, $X_t^{T_n}$ is also an $\mathcal{F}_{T_n \wedge t}$-martingale.

Let us prove that if $X_t^{T_n}$ is an $\mathcal{F}_{T_n \wedge t}$-martingale, then it is an $\mathcal{F}_t$-martingale for every n. For simplicity of notations, let $T_n = S$, i.e. X_t^S is an $\mathcal{F}_{S \wedge t}$-martingale. Take $r < t$, and let $A \in \mathcal{F}_r$. We claim that $A \cap \{S > r\} \in \mathcal{F}_{S \wedge r}$. Let us prove this claim. Take a $\xi \in \mathbb{R}$. We have to prove that

$$(A \cap \{S > r\}) \cap \{S \wedge r < \xi\} \in \mathcal{F}_\xi.$$

Suppose $\xi \leqslant r$. Then, $\{r < S\} \cap \{S \wedge r < \xi\} = \emptyset$. Now let $\xi > r$. Then, $\{S \wedge r < \xi\} = \Omega$. Now since $A \in \mathcal{F}_r$ and $\{S > r\} \in \mathcal{F}_r$, then

$$(A \cap \{S > r\}) \cap \{S \wedge r < \xi\} \in \mathcal{F}_r \subset \mathcal{F}_\xi, \tag{23}$$

which proves the claim.

Next, since X_t^S is an $\mathcal{F}_{S\wedge t}$-martingale, we obtain:

$$\int_{A\cap\{S>r\}} X_t^S P(d\omega) = \int_{A\cap\{S>r\}} X_r^S P(d\omega).$$

On the other hand, taking into account that on the set $\{S \leqslant r\}$, one has $S \leqslant r < t$, we have:

$$\begin{aligned}\int_{A\cap\{S\leqslant r\}} X_t^S P(d\omega) &= \int_{A\cap\{S\leqslant r\}} X_{S\wedge t} P(d\omega) = \int_{A\cap\{S\leqslant r\}} \mathbb{I}_{S>0}\, X_S P(d\omega) \\ &= \int_{A\cap\{S\leqslant r\}} X_{S\wedge r} P(d\omega) = \int_{A\cap\{S\leqslant r\}} X_r^S P(d\omega) \quad (24)\end{aligned}$$

Adding (23) to (24), we obtain that

$$\int_A X_t^S P(d\omega) = \int_A X_r^S P(d\omega)$$

which proves that $\mathbb{E}[X_t^S|\mathcal{F}_r] = X_r^S$. □

Theorem 5.18. *Let X_t, $t \geqslant 0$, be a continuous local martingale. Then the sequence of stopping times*

$$T_n = \inf\{t \geqslant 0 : |X_t| > n\}$$

is always reducing. In particular, we can find a sequence of stopping times that reduces X_t to a bounded martingale.

Proof. Note that by Proposition 5.4 every T_n is a stopping time. Let $0 < s < t$, and let S_n be a reducing sequence of stopping times. We apply the optional stopping theorem (Theorem 5.16) to each martingale $X_t^{S_n}$ and conclude that for every $m > 0$ the process $X_{T_m\wedge t}^{S_n}$ is a martingale. This implies:

$$\mathbb{E}[X_{t\wedge T_m\wedge S_n} \,|\, \mathcal{F}_s] = X_{s\wedge T_m\wedge S_n}.$$

By Theorem 5.14 we can pass to the limit as $n \to \infty$ in the both parts of the above equality. We obtain that a.s.

$$\mathbb{E}[X_{t\wedge T_m} \,|\, \mathcal{F}_s] = X_{s\wedge T_m}.$$

Hence, $X_t^{T_m}$ is an $\mathcal{F}_t$-martingale. Note that $|X_t^{T_n}| \leqslant n$, and therefore, $X_t^{T_n}$ is a bounded martingale. □

Remark. Instead of the sequence $\{T_n\}$ we could consider any sequence of stopping times $T_n' \leqslant T_n$ such that $\lim_{n\to\infty} T_n' = \infty$. Clearly, $X_t^{T_n'}$ would be a bounded maringale.

Definition 5.19. *Let $\tau > 0$ be a random time. A process X_t, $t \geqslant 0$, is called a local martingale on $[0, \tau)$ if there is a sequence of stopping time $T_n \uparrow \tau$ such that $X_t^{T_n}$ is a martingale.*

Theorem 5.20. *Suppose X_t, $t \geqslant 0$, is a local martingale, and for every $t > 0$,*

$$\mathbb{E}\Big[\sup_{0\leqslant s\leqslant t} |X_s|\Big] < \infty. \tag{25}$$

Then, X_t is a martingale.

Proof. Clearly, $\mathbb{E}|X_t| < \infty$. Now let $\{T_n\}$ be a reducing sequence of stopping times so that

$$\mathbb{E}[X_t^{T_n} \,|\, \mathcal{F}_s] = X_s^{T_n} \qquad \text{a.s.} \tag{26}$$

Assumption (25) allows us to apply Theorem 5.14 since $|X_t^{T_n}| \leqslant \sup_{0\leqslant s\leqslant t} |X_s|$ and the latter function is integrable. Taking the limit in (26) as $n \to \infty$, we obtain:

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

This proves that X_t is a martingale. □

Corollary 5.21. *A bounded local martingale is a martingale.*

5.6 The variance process

Let $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ be a filtered probability space.

Definition 5.22. *An $\mathcal{F}_t$-adapted process X_t is called a submartingale if $\mathbb{E}|X_t| < \infty$ and*

$$\mathbb{E}(X_t|\mathcal{F}_s) \geqslant X_s \quad \text{a.s. for all} \quad s \leqslant t.$$

Theorem 5.23 (Doob–Meyer's theorem)**.** *Let X_t be a continuous submartingale such that for every constant $a > 0$, the family of rundom variables*

$$\{X_\tau,\ \tau \text{ is a stopping time with } \tau \leqslant a\}$$

is uniformly integrable. Then there exists a unique decomposition

$$X_t = M_t + A_t$$

where M_t is a continuous martingale and A_t is a continuous non-decreasing $\mathcal{F}_t$-adapted process starting at zero and such that $\mathbb{E}(A_t) < \infty$.

Proof. Without proof. □

Lemma 5.24. *Let X_t be a continuous martingale. Then X_t^2 is a continuous submartingale satisfying the assumptions of Theorem 5.23.*

Proof. Let τ be a stopping time with $\tau \leqslant a$ where a is a constant. By the Optional stopping theorem, we obtain:

$$\mathbb{E}[(X_a - X_\tau)^2|\mathcal{F}_\tau] = \mathbb{E}[X_a^2|\mathcal{F}_\tau] - 2X_\tau\mathbb{E}[X_a|\mathcal{F}_\tau] + X_\tau^2 = \mathbb{E}[X_a^2|\mathcal{F}_\tau] - X_\tau^2.$$

Since the left-hand side is positive, we obtain that

$$\mathbb{E}[X_a^2|\mathcal{F}_\tau] \geqslant X_\tau^2. \tag{27}$$

The above inequality, applied to two times $s < t$, implies, in particular, that X_t^2 is a submartingale. Fix an arbitrary $\varepsilon > 0$, and choose a constant $K > 0$ so that $\frac{1}{K}\,\mathbb{E}|X_a|^2 < \varepsilon$. Applying Chebyshev's inequality, we obtain:

$$\mathbb{P}(|X_\tau| > K) \leqslant \frac{1}{K^2}\,\mathbb{E}|X_\tau|^2 \leqslant \frac{1}{K^2}\,\mathbb{E}|X_a|^2.$$

We applied (27) after taking expectations of the both parts. Again, by Chebyshev's inequality,

$$\begin{aligned}\int_\Omega \mathbb{I}_{\{|X_\tau|>K\}}\,|X_\tau(\omega)|\mathbb{P}(d\omega) &\leqslant \Big(\int_\Omega \mathbb{I}_{\{|X_\tau|>K\}}\,\mathbb{P}(d\omega)\Big)^{\frac{1}{2}}\Big(\int_\Omega |X_t(\omega)|^2\mathbb{P}(d\omega)\Big)^{\frac{1}{2}} \\ &\leqslant \mathbb{P}(|X_\tau| > K)^{\frac{1}{2}}\left(\mathbb{E}|X_a|^2\right)^{\frac{1}{2}} \leqslant \frac{1}{K}\,\mathbb{E}|X_a|^2 < \varepsilon.\end{aligned}$$

□

In general, if X_t is a martingale, X_t^2 fails to be a martingale.

Theorem 5.25. *Let X_t be a square integrable continuous $\mathcal{F}_t$-martingale. Then, there exists a process $\langle X\rangle_t$, $t \geqslant 0$, which is continuous, non-decreasing, and such that the process $X_t^2 - \langle X\rangle_t$ is a continuous martingale. The proces $\langle X\rangle_t$ is a.s. unique.*

Proof. By Lemma 5.24, X_t^2 is a submartingale sarisfying the assumptions of Doob-Meyer's theorem. Therefore, there exists a continuous martingale M_t and a non-decreasing $\mathcal{F}_t$ continuous apadted process, denoted by $\langle X\rangle_t$, starting at zero and such that

$$X_t^2 = M_t + \langle X\rangle_t.$$

This also implies that $X_t^2 - \langle X\rangle_t$ is a continuous martingale. By the uniqueness of the Doob-Meyer's decomposition, the process $\langle X\rangle_t$ is unique. □

Definition 5.26. *The process $\langle X\rangle_t$ defined in Theorem 5.25 is called the **variance process**.*

Theorem 5.27. *Every continuous $\mathcal{F}_t$-local martingale starting at zero which is of bounded variation on every compact interval is constantly equal to zero.*

Proof. Let

$$V_t = \sup\Big\{\sum_{k=1}^{n} |X_{t_k} - X_{t_{k-1}}| : 0 = t_0 < t_1 < \cdots < t_n = t\Big\}$$

be the variation of X_t on $[0,t]$. Since X_t is continuous, and V_t is bounded on every compact interval, V_t is continuous. Let $S_k = \inf\{s : V_s > k\}$. Note that if $t \leqslant S$, then $|X_t| \leqslant k$. By Theorem 5.18, S_k is a reducing sequence of stopping times, and the process $Y_t = X_t^{S_k}$ is a bounded martingale with respect to $\mathcal{F}_t$. Note that:

$$\mathbb{E}[(Y_t - Y_s)^2 \,|\, \mathcal{F}_s] = \mathbb{E}[Y_t^2 \,|\, \mathcal{F}_s] - 2Y_s\,\mathbb{E}[Y_t \,|\, \mathcal{F}_s] + Y_s^2 = \mathbb{E}[Y_t^2 - Y_s^2 \,|\, \mathcal{F}_s]. \tag{28}$$

Take a partition of $\{0 = t_0 < t_1 < \cdots < t_n = t\}$ of the interval $[0,t]$. We have:

$$\begin{aligned}\mathbb{E}[Y_t^2] = \mathbb{E}\Big[\sum_{i=1}^{n} Y_{t_i}^2 - Y_{t_{i-1}}^2\Big] = \mathbb{E}\Big[\sum_{i=1}^{n}(Y_{t_i} - Y_{t_{i-1}})^2\Big] &\leqslant \mathbb{E}\Big[\sup_{i\geqslant 1}|Y_{t_i} - Y_{t_{i-1}}|\, V(t\wedge S)\Big] \\ &\leqslant k\,\mathbb{E}\Big[\sup_{i\geqslant 1}|Y_{t_i} - Y_{t_{i-1}}|\Big].\end{aligned}$$

Let the mesh $\delta_n = \max_{i\geqslant 1}|t_i - t_{i-1}|$ tend to zero as $n \to \infty$, and note that since $|Y_{t_i}| = |X_{S_k\wedge t_i}| \leqslant k$ for all $1 \leqslant i \leqslant n$, then $\sup_{i\geqslant 1}\Big[|Y_{t_i} - Y_{t_{i-1}}|\Big] \leqslant 2k$. By Lebesgue's theorem we can pass to the limit under the expectation sign. This implies that

$$\lim_{n\to\infty} \mathbb{E}\sup_{i\geqslant 1}\Big[|Y_{t_i} - Y_{t_{i-1}}|\Big] = 0.$$

Therefore $\mathbb{E}[Y_t^2] = 0$, and hence, $Y_t = X_{S_k\wedge t} = 0$ a.s. This proves that there exists a set $\Omega' \subset \Omega$ of full P-measure such that $X_t^{S_k} = 0$ for all rational $t \geqslant 0$ and for all $\omega \in \Omega'$. By continuity of paths, $X_t^{S_k} = 0$ for all $t \geqslant 0$ and for all $\omega \in \Omega'$. Finally we can find another set $\Omega'' \subset \Omega'$ of full P-measure such that $X_t = 0$ for all $t \geqslant 0$ and for all $\omega \in \Omega''$. □

Lemma 5.28. *Let X_t be a martingale and T be a stopping time. Then $\langle X^T\rangle_t = \langle X\rangle_t^T$.*

Proof. By what was proved $X_t^2 - \langle X\rangle_t$ is a martingale. By the optional stopping theorem $(X_t^T)^2 - \langle X\rangle_t^T$ is also a martingale. By uniqueness of the variance process for a martingale, $\langle X^T\rangle_t = \langle X\rangle_t^T$. □

Suppose X_t and Y_t, $t \geqslant 0$, are martingales. Then, in general, X_tY_t fails to be a martingale.

Theorem 5.29. *Let X_t and Y_t be martingales which are continuous and start at zero. Then there exists a unique, continuous process $\langle X, Y\rangle_t$ that has bounded variation on every compact interval and starts at zero, such that $X_tY_t - \langle X, Y\rangle_t$ is a martingale.*

Proof. Define

$$\langle X, Y\rangle_t = \frac{1}{4}\big(\langle X+Y\rangle_t - \langle X-Y\rangle_t\big).$$

Clearly, $\langle X, Y\rangle_t$ is continuous, predictable, and of bounded variation. The latter property holds since $\langle X, Y\rangle_t$ is defined as a difference of two increasing processes. Finally,

$$X_tY_t - \langle X, Y\rangle_t = \frac{1}{4}\Big((X_t+Y_t)^2 - \langle X+Y\rangle_t - \big((X_t-Y_t)^2 - \langle X-Y\rangle_t\big)\Big).$$

As a difference of two martingales, $X_tY_t - \langle X, Y\rangle_t$ is a martingale. Let us prove the uniqueness of $\langle X, Y\rangle_t$. Suppose there are two processes A_t and B_t satisfying requirements of the theorem. Then

$$A_t - B_t = (X_tY_t - A_t) - (X_tY_t - B_t)$$

is a continuous martingale which has bounded variation on every compact interval. By Theorem 5.27, $A_t - B_t$ is constantly equal to zero. This implies the uniqueness of the process $\langle X, Y\rangle_t$. □

5.7 Semimartingales

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

Definition 5.30. *A stochastic process X_t, $t \geqslant 0$, is called a* ***semimartingale*** *if there is an $\mathcal{F}_t$-local martingale M_t, $t \geqslant 0$, and a cádlág (the paths are right-continuous with left limits) $\mathcal{F}_t$-adapted process A_t, $t \geqslant 0$, of locally bounded variation and starting at 0 such that $X_t = M_t + A_t$ for all $t \geqslant 0$.*

Theorem 5.31. *If X_t is a continuous semimartingale, then the decomposition $X_t = M_t + A_t$ is unique.*

Proof. Suppose $X = M + A = M' + A'$. Then $A - A' = M' - M$ is a continuous local martingale locally of bounded variation and starting at zero. By Theorem 5.27, it is constantly equal to zero. □

6. Stochastic Integral

Let $(\Omega, \mathcal{F}, P)$ be a probability space. Let B_t be an n-dimensional Brownian motion, and let $\mathcal{F}_t = \sigma(B_s, s \leqslant t)$ be the natural filtration.

6.1 Construction of the stochastic integral and Itô's isometry

The class of integrands

Definition 6.1. *Let $\Phi = \Phi(S, T)$ be the class of functions*

$$f(t, \omega) : \,[0, \infty) \times \Omega \to \mathbb{R}$$

such that

1. *$(t,\omega) \mapsto f(t,\omega)$ is $\mathfrak{B} \times \mathcal{F}$-measurable, where $\mathfrak{B}$ denotes the σ-algebra of Borel subsets of $[0,\infty)$.*

2. *$f(t,\cdot)$ is $\mathcal{F}_t$-adapted.*

3. *$\mathbb{E}\Big[\int\limits_S^T f^2(t,\omega)\,dt\Big] < \infty$.*

Now we are going to show how to define the stochastic integral

$$I[f](\omega) = \int_S^T f(t,\omega)\,dB_t(\omega)$$

for functions $f \in \Phi$, where B_t is a one-dimensional Brownian motion. The idea is the following: first we define the stochastic integral for simple functions from Φ. Then we show that every function $f \in \Phi$ can be represented as a limit of simple functions ϕ_n in some sense. Then we define the stochastic integral $\int f dB_t$ as a limit $\int \phi_n dB_t$.

Definition of the Stochastic integral for simple functions

A function $\phi \in \Phi$ is called **simple** if it has the form:

$$\phi(t,\omega) = \sum_{j=0}^{N-1} e_j(\omega) \cdot \mathbb{I}_{[t_j,t_{j+1})} \tag{29}$$

where $S = t_0 < t_1 < \cdots t_N = T$ is a partition. Note that since $\phi \in \Phi$, it is $\mathcal{F}_t$-adapted. Therefore, e_j is $\mathcal{F}_{t_j}$-measurable. For functions of form (29) we define the stochastic integral by the formula:

$$I[\phi] = \int_S^T \phi(t,\omega)\,dB_t(\omega) = \sum_{j=0}^{N-1} e_j(\omega)(B_{t_{j+1}} - B_{t_j})(\omega) \tag{30}$$

Itô's isometry

Theorem 6.2 (Itô's isometry for simple functions)**.** *If $\phi \in \Phi$ is bounded and simple then*

$$\mathbb{E}\Big[\Big(\int_S^T \phi(t,\omega)dB_t(\omega)\Big)^2\Big] = \mathbb{E}\Big[\int_S^T \phi^2(t,\omega)\,dt\Big]. \tag{31}$$

Proof. Define $\Delta B_j = B_{t_{j+1}} - B_{t_j}$. Then,

$$\mathbb{E}[e_i e_j \Delta B_i \Delta B_j] = \begin{cases} 0, & \text{if } i \neq j, \\ \mathbb{E}[e_j^2]\,(t_{j+1} - t_j), & \text{if } i = j. \end{cases}$$

Indeed, if $i < j$ then the random variables $e_i\, e_j\, \Delta B_i$ and ΔB_j are independent. Indeed, $e_i(B_{t_{i+1}} - B_{t_i})$ is $\mathcal{F}_{t_{i+1}}$-measurable and therefore is $\mathcal{F}_{t_j}$-measurable $(i+1 \leqslant j)$,

e_j is $\mathcal{F}_{t_j}$-measurable. Therefore, the product $e_i\, e_j\, \Delta B_i$ is $\mathcal{F}_{t_j}$-measurable. On the other hand, ΔB_{t_j} is independent of $\mathcal{F}_{t_j}$. Hence, $e_i\, e_j\, \Delta B_i$ and ΔB_j are independent for $i < j$. This implies that

$$\mathbb{E}[e_i e_j \Delta B_i\, \Delta B_j | \mathcal{F}_{t_i}] = e_i e_j \Delta B_i\, \mathbb{E}[\Delta B_j | \mathcal{F}_{t_j}] = 0, \quad \text{and}$$
$$\mathbb{E}[e_i^2 \Delta B_i^2 | \mathcal{F}_{t_i}] = e_i^2\, \mathbb{E}[(B_{t_{i+1}} - B_{t_i})^2] = e_i^2 (t_{i+1} - t_i).$$

Hence,

$$\mathbb{E}\Big[\Big(\int_S^T \phi(t,\omega) dB_t(\omega)\Big)^2\Big] = \sum_{i,j=0}^{N-1} \mathbb{E}[e_i e_j\, \Delta B_i \Delta B_j] = \sum_{j=1}^{N-1} \mathbb{E}[e_j^2]\,(t_{j+1} - t_j)$$
$$= \mathbb{E}\Big[\int_S^T \phi^2(t,\omega)\, dt\Big].$$

□

Approximation of the integrands from Φ by simple functions

Lemma 6.3. *Let $g \in \Phi$ be bounded and $g(\cdot, \omega)$ be continuous for each ω. Then, there exist simple functions $\phi_n \in \Phi$ such that*

$$\lim_{n\to\infty} \mathbb{E}\Big[\int_S^T (g - \phi_n)^2 dt\Big] = 0.$$

Proof. Define $\phi_n = \sum_{j=0}^{N-1} g(t_j, \omega)\, \mathbb{I}_{[t_j, t_{j+1})}(t)$. Since $g \in \Phi$, $g(t_j, \cdot)$ is $\mathcal{F}_{t_j}$-measurable, and therefore ϕ_n is $\mathcal{F}_t$-adapted. Clearly, since every indicator function $\mathbb{I}_{[t_j, t_{j+1})}(t)$ is $\mathfrak{B}$-measurable, and every $g(t_j, \omega)$ is $\mathcal{F}_{t_j}$-measurable (and therefore $\mathcal{F}$-measurable), ϕ_n is $\mathfrak{B} \times \mathcal{F}$-measurable. Hence $\phi_n \in \Phi$. Since $g(\cdot, \omega)$ is continuous for every ω, then

$$\lim_{n\to\infty} \int_S^T (g - \phi_n)^2 dt = 0.$$

Since g is bounded, ϕ_n is bounded as well, and therefore, by Lebesgue's theorem,

$$\lim_{n\to\infty} \mathbb{E}\Big[\int_S^T (g - \phi_n)^2 dt\Big] = 0.$$

□

Definition 6.4. *Let K_n, $n \in \mathbb{N}$, be a sequence of Lebesgue-integrable functions on $\mathbb{R}$ possessing the properties:*

1. $K_n \geqslant 0$;
2. $\int_{-\infty}^{\infty} K_n(x) dx = 1$;
3. $\lim_{n\to\infty} \sup_{x \neq I} |K_n(x)| = 0$.

The sequence of functions K_n is called an ***approximate identity****.*

Proposition 6.5. *Let $\{K_n\}$ be an approximate identity, and let $f \in L_1$. Then $\{f * K_n\}$, where $(f * K_n)(t) = \int_{-\infty}^{\infty} K_n(s-t)\, f(s)\, ds$, converges to f in the L_p-norm, $1 \leqslant p < \infty$.*

Proof. Without proof. A proof can be found in the book by Hoffman "Banach spaces of analytic functions", 1962, p. 22. □

Lemma 6.6. *Let $h \in \Phi$ be bounded. Then there exist bounded functions $g_n \in \Phi$ such that $g_n(\cdot, \omega)$ is continuous for every ω and n, and*

$$\lim_{n\to\infty} \mathbb{E}\Big[\int_S^T (h - g_n)^2\, dt\Big] = 0.$$

Proof. Suppose M is a constant with the property $|h(t,\omega)| \leqslant M$ for all (t,ω). For each $n \in \mathbb{N}$, let ψ_n be a non-negative, continuous function on $\mathbb{R}$ such that

1. $\psi_n(x) = 0 \quad \text{for} \quad x \leqslant -\frac{1}{n}$ and $x \geqslant 0$,
2. and $\int_{-\infty}^{\infty} \psi_n(x) dx = 1$.

The sequence ψ_n is an approximate identity. Define

$$g_n(t,\omega) = \int_0^t \psi_n(s-t)\, h(s,\omega)\, ds \tag{32}$$

and note that the following properties of these functions. If $s > t$, then $\psi_n(s-t) = 0$, and, therefore, the integration in (32) can be extended to $[0,\infty)$. Furthermore, if $t \geqslant \frac{1}{n}$ and $s < 0$, then $s - t \leqslant -\frac{1}{n}$, and hence, $\psi_n(s-t) = 0$. This implies that for $t \geqslant \frac{1}{n}$, we can extend the integration in (32) to $(-\infty,\infty)$. Thus, if $t \geqslant \frac{1}{n}$, then $g_n(t,\omega) = (h(\cdot,\omega) * \psi_n)(t)$. As a convolution, g_n is t-continuous for each ω and n, and $g_n(t,\omega) \leqslant M$. Since $h(s,\omega)$ is $\mathfrak{B} \times \mathcal{F}$-measurable, so is the product $\psi_n(s-t)\, h(s,\omega)$, and therefore the integral on the right-hand side of (32) is a limit of $\mathfrak{B} \times \mathcal{F}$-measurable functions. Next, $h(s,\omega)$ is $\mathcal{F}_s$-measurable, and therefore $\mathcal{F}_t$-measurable since $s \leqslant t$. Again, the integral on the right-hand side of (32) can be represented as a limit of $\mathcal{F}_t$-measurable functions and therefore $g_n(t,\omega)$ is $\mathcal{F}_t$-measurable for all $t \geqslant 0$. Note that,

$$\mathbb{E}\Big[\int_S^T g_n(s,\omega)^2\, ds\Big] < \infty$$

since g_n is bounded. Hence, $g_n \in \Phi$. By Proposition (6.5)

$$\begin{aligned}\int_S^T (g_n(s,\omega) - h(s,\omega))^2\, ds &\leqslant 2\int_S^T \big(g_n(s,\omega) - (h(\cdot,\omega) * \psi_n)(s)\big)^2\, ds \\ &\quad + 2\int_S^T ((h(\cdot,\omega) * \psi_n)(t) - h(s,\omega))^2\, ds \to 0 \quad \text{as } n \to \infty.\end{aligned}$$

Since both functions g_n and h are bounded, by Lebesgue's theorem,

$$\lim_{n\to\infty} \mathbb{E}\Big[\int_S^T (g_n(s,\omega) - h(s,\omega))^2 = 0\Big].$$

$\square$

Lemma 6.7. *Let $f \in \Phi$. Then there exists a sequence $\{h_n\}$ such that it is bounded and*

$$\lim_{n\to\infty} \mathbb{E}\Big[\int_S^T (f - h_n)^2 \, ds\Big] = 0. \tag{33}$$

Proof. Define

$$h_n(t,\omega) = \begin{cases} -n, & \text{if } f(t,\omega) < -n, \\ f(t,\omega) & \text{if } -n \leqslant f(t,\omega) \leqslant n, \\ n, & \text{if } f(t,\omega) > n. \end{cases}$$

Note that $|h_n(t,\omega)| \leqslant |f(t,\omega)|$. Hence $\mathbb{E}\big[\int_S^T h_n(s,\omega)^2 ds\big] < \mathbb{E}\big[\int_S^T f(s,\omega)^2 ds\big] < \infty$. Clearly h_n possesses also other properties of the class Φ by its construction. Therefore, $h_n \in \Phi$. Next, $h_n \to f$ for each $\omega \in \Omega$ and each $t \in [S,T]$. By Lebesgue's theorem, in (33) we can pass to the limit under the expectation and the integral signs which will imply (33). $\square$

Summarizing Lemmas 6.3, 6.6, and 6.7, we obtain the following:

Lemma 6.8. *Let $f \in \Phi$, then there exists a system ψ_n of simple functions such that*

$$\lim_{n\to\infty} \mathbb{E}\Big[\int_S^T (f - \psi_n)^2 \, ds\Big] = 0.$$

Proof. Proof follows from Lemmas 6.3, 6.6, and 6.7. $\square$

Definition of the stochastic integral and Itô's isometry

Theorem 6.9. *Let $f \in \Phi$, and let $\psi_n \in \Phi$ be a sequence of simple functions such that*

$$\lim_{n\to\infty} \mathbb{E}\Big[\int_S^T (f - \psi_n)^2 \, dt\Big] = 0. \tag{34}$$

Then there exist a random variable $I[f](\omega)$ such that

$$\lim_{n\to\infty} \mathbb{E}\Big(I[f] - I[\psi_n] \Big)^2 = 0$$

where $I[\psi_n]$ is the stochastic integral of the simple function ψ_n defined by (30). Moreover $I[f]$ does not depend on the choice of a sequence ψ_n converging to f in the $L_2(\Omega \times [S,T])$-norm (as in (34)).

Definition 6.10 (Stochastic integral)**.** *The random variable $I[f]$ defined in the theorem above is called the* ***stochastic integral*** *of f with respect to a Brownian motion. It is denoted by the symbol*

$$I[f] = \int_S^T f(t, \omega)\, dB_t(\omega).$$

Proof of Theorem 6.9. By Theorem 6.2 (Itô's isometry for simple functions) we obtain:

$$\mathbb{E}\Big[\int_S^T (\psi_m - \psi_n)^2\, dt\Big] = \mathbb{E}\Big(\, I[\psi_n] - I[\psi_m]\,\Big)^2. \tag{35}$$

Since ψ_n converges to f in the $L_2(\Omega \times [S,T])$-norm, it is a Cauchy sequence with respect to this norm. Therefore, the sequence $I[\psi_n]$ on the right-hand side of (35) is a Cauchy sequence in the $L_2(\Omega)$-norm. But the $L_2(\Omega)$-space is complete. Hence, there exists a limit of $I[\psi_n]$ in this space, i.e. there exists a random variable $I[f]$ such that

$$\lim_{n\to\infty} \mathbb{E}\Big(\, I[f] - I[\psi_n]\,\Big)^2 = 0.$$

Suppose ψ'_n is another sequence of simple functions that converges to f with respect to the $L_2(\Omega \times [S,T])$-norm. Then, Itô's isometry implies:

$$\mathbb{E}\Big[\int_S^T (\psi'_n - \psi_n)^2\, dt\Big] = \mathbb{E}\Big(\, I[\psi_n] - I[\psi'_n]\,\Big)^2.$$

The left-hand side of this equality converges to zero since both sequences ψ_n and ψ'_n converge to f. Therefore the right-hand side converges to zero too. This proves that

$$\lim_{n\to\infty} I[\psi_n] = \lim_{n\to\infty} I[\psi'_n]$$

in the $L_2(\Omega)$-norm. Hence the stochastic integral $I[f] = \int_S^T f(t,\omega)dB_t(\omega)$ does not depend on the choice of the approximating system ψ_n. □

Corollary 6.11 (Itô's isometry)**.** *For any $f \in \Phi$,*

$$\Big[\Big(\int_S^T f(t,\omega)dB_t\Big)^2\Big] = \mathbb{E}\Big[\int_S^T f(t,\omega)^2\, dt\Big] \tag{36}$$

Proof. For any simple function ψ_n by Theorem 6.2 it holds that

$$\mathbb{E}\Big[\Big(\int_S^T \psi_n(t,\omega)dB_t\Big)^2\Big] = \mathbb{E}\Big[\int_S^T \psi_n(t,\omega)^2\, dt\Big]$$

Passing to the limit as $n \to \infty$ in the both parts of the above idenity we obtain (36). Namely, the convergence of expectations holds by the simple norm inequality: $\big|\|f\| - \|g\|\big| \leqslant \|f - g\|$ which was applied to the L_2-norms above. □

6.2 Properties of the stochastic integral

Theorem 6.12. *Let $f, g \in \Phi(0,T)$ and let $0 \leqslant S < U < T$. Then*

(i) $\int\limits_S^T f\,dB_t = \int\limits_S^U f\,dB_t + \int\limits_U^T f\,dB_t$ *a.s.*

(ii) $\int\limits_S^T (cf+g)dB_t = c \cdot \int\limits_S^T f\,dB_t + \int\limits_S^T g\,dB_t$ *a.s., where c is a constant*

(iii) $\mathbb{E}\big[\int\limits_S^t f\,dB_s\big] = 0$ *for all $S \leqslant t \leqslant T$.*

(iv) $\int_S^t f\,dB_s$ *is $\mathcal{F}_t$-measurable for all $S \leqslant t \leqslant T$.*

Proof. Properties (i) and (ii) clearly hold for simple functions. If $f \in \Phi$ is arbitrary we take a sequence of simple functions ψ_n converging to f as $n \to \infty$. Taking limits from the both sides in (i) and (ii) in the $L_2(\Omega)$-norm we obtain (i) and (ii) for f. Let us prove (iii). Again, (iii) clearly hold when f is a simple function of form (29), i.e. when the stochastic integral $\int\limits_S^t f\,dB_s$ is given by (30). Indeed, for each summand in (30) we obtain:

$$\mathbb{E}\big[e_j(\omega)(B_{t_{j+1}} - B_{t_j})\big] = \mathbb{E}[e_j]\,\mathbb{E}[B_{t_{j+1}} - B_{t_j}] = 0.$$

The equality holds since e_j is $\mathcal{F}_{t_j}$-measurable and $B_{t_{j+1}} - B_{t_j}$ is independent of $\mathcal{F}_{t_j}$. Next if $\psi_n \to f$ as $n \to \infty$ in the $L_2(\Omega \times [S,t])$-norm, then

$$\begin{aligned}\Big|\mathbb{E}\Big[\int_S^t f\,dB_s\Big] - \mathbb{E}\Big[\int_S^t \psi_n\,dB_s\Big]\Big| &\leqslant \mathbb{E}\Big|\int_S^t f\,dB_s - \int_S^t \psi_n\,dB_s\Big| \\ &\leqslant \Big(\mathbb{E}\Big(\int_S^t f\,dB_s - \int_S^t \psi_n\,dB_s\Big)^2\Big)^{\frac{1}{2}} \to 0 \quad \text{as } n \to \infty.\end{aligned}$$

To show (iv), note that if $\psi \in \Phi$ is simple, then

$$\int_S^t \psi(s,\omega)\,dB_s = \sum_{j=0}^{N-1} e_j(\omega)\,(B_{t_{j+1}} - B_{t_j}),$$

where $t_{j_N} = t$, is an $\mathcal{F}_t$-measurable function. If $\psi_n \to f$ as $n \to \infty$ with respect to the $L_2(\Omega \times [S,s])$-norm, then $\int_S^t \psi_n\,dB_s \to \int_S^t f\,dB_s$ with respect to the $L_2(\Omega)$-norm, and therefore the limit $\int_S^t f\,dB_s$ is $\mathcal{F}_t$-measurable as a limit of $\mathcal{F}_t$-measurable functions. □

Note that the stochastic integral $\int\limits_0^t f\,dB_s$ can be also regarded as a stochastic process where t is the time. We are going to show its continuity in t and that it is a martingale. We will need the following lemma.

Lemma 6.13. *Suppose for each n, $Z_t^{(n)}$, $t \geqslant 0$, is a martingale with respect to the filtration $\mathcal{F}_t$, and for each $t \geqslant 0$, $Z_t^{(n)} \to Z_t$ as $n \to \infty$ with respect to the L_p-norm, $p \geqslant 1$. Then Z_t, $t \geqslant 0$, is a martingale.*

Proof. It suffices to prove the statement for the case $p = 1$ since the L_1-convergence is the weakest. We obtain:

$$\mathbb{E}\Big|\mathbb{E}[Z_t^{(n)} \,|\, \mathcal{F}_s] - \mathbb{E}[Z_t \,|\, \mathcal{F}_s]\Big| = \mathbb{E}\Big|\mathbb{E}[Z_t^{(n)} - Z_t \,|\, \mathcal{F}_s]\Big| \leqslant \mathbb{E}\Big[\mathbb{E}[|Z_t^{(n)} - Z_t| \,|\, \mathcal{F}_s]\Big]$$
$$= \mathbb{E}|Z_t^{(n)} - Z_t| \to 0 \quad \text{as } n \to \infty.$$

Since $\mathbb{E}[Z_t^{(n)} \,|\, \mathcal{F}_s] = Z_s^{(n)}$, we obtain from the above inequalities that

$$\lim_{n\to\infty} Z_s^{(n)} = \mathbb{E}[Z_t \,|\, \mathcal{F}_s]$$

On the other hand

$$\lim_{n\to\infty} Z_s^{(n)} = Z_s$$

by assumption. Therefore $\mathbb{E}[Z_t \,|\, \mathcal{F}_s] = Z_s$. □

Theorem 6.14. *The stochastic integral integral $\int\limits_0^t f\, dB_s$ is a martingale.*

Proof. First we prove the statement when f is a simple function. Let us prove that if $s \leqslant t$, then

$$\mathbb{E}\Big[\int_s^t \psi\, dB_r \,|\, \mathcal{F}_s\Big] = 0.$$

Indeed, if $\psi = \sum_{j=0}^{N-1} e_j\, \mathbb{I}_{[t_{j+1},t_j)}$, then $\int_s^t \psi\, dB_r = \sum_{j=p}^{N-1} e_j\,(B_{t_{j+1}} - B_{t_j})$ where $t_p = s$. Compute the conditional expectation with respect to $\mathcal{F}_s$ of each summand of the stochastic integral. We obtain:

$$\mathbb{E}[e_j\,(B_{t_{j+1}} - B_{t_j}) \,|\, \mathcal{F}_s] = \mathbb{E}\big[\,\mathbb{E}\,[e_j\,(B_{t_{j+1}} - B_{t_j}) \,|\, \mathcal{F}_{t_j} \,|\, \mathcal{F}_s]\big]$$
$$= \mathbb{E}\big[e_j\, \mathbb{E}[(B_{t_{j+1}} - B_{t_j}) \,|\, \mathcal{F}_{t_j}] \,|\, \mathcal{F}_s\big] = 0.$$

Clealry, $\mathbb{E}[(B_{t_{j+1}} - B_{t_j}) \,|\, \mathcal{F}_{t_j}] = 0$. Moreover, e_j is $\mathcal{F}_{t_j}$-measurable, and therefore was written outside the conditional expectation sign. By additivity, $\mathbb{E}\big[\int_s^t \psi\, dB_r \,|\, \mathcal{F}_s\big] = 0$. Now again applying the additivity property of the stochastic integral (property (i)), we obtain:

$$\mathbb{E}[\int_0^t \psi\, dB_r \,|\, \mathcal{F}_s\Big] = \mathbb{E}\Big[\int_0^s \psi\, dB_r \,|\, \mathcal{F}_s\Big] + \mathbb{E}\Big[\int_s^t \psi\, dB_r \,|\, \mathcal{F}_s\Big]$$
$$= \mathbb{E}\Big[\int_0^s \psi\, dB_r \,|\, \mathcal{F}_s\Big] = \int_0^s \psi\, dB_r.$$

By what was proved $\mathbb{E}\Big[\int_s^t \psi\, dB_r \,|\, \mathcal{F}_s\Big] = 0$. Moreover, Property (iv) of Theorem 6.12 implies that $\int_0^s \psi\, dB_r$ is an $\mathcal{F}_s$-measurable random variable. Note that I_n converge to I with respect to the L_2-norm. Let us apply Lemma 6.13 which says that in this case $I(t, \cdot)$ is a martingale. □

Lemma 6.15. *Let M_t be a continuous martingale, and let $0 < T < \infty$. Then for any $\varepsilon > 0$,*

$$P\Big(\sup_{[0,T]} |M_t| \geqslant \varepsilon\Big) \leqslant \frac{1}{\varepsilon^p}\mathbb{E}\big[|M_T|^p\big].$$

Proof. Without proof. □

Theorem 6.16. *Let $f \in \Phi(0,T)$, then there exist a t-continuous version of the stochastic integral $\int\limits_0^t f\,dB_s$.*

Remark. In other words, we have to prove that there exists a continuous process J_t such that for every $0 \leqslant t \leqslant T$

$$J_t = \int_0^t f\,dB_s \quad \text{a.s.}$$

Proof. Let ψ_n be a sequence of simple functions converging to f, and let $I_n(t,\omega) = \int_0^t \psi_n(s,\omega)\,dB_s$, $I(t,\omega) = \int_0^t f(s,\omega)\,dB_s$. Note that $I_n(t,\omega) = \sum_{j=0}^{N-2} e_j(\omega)(B_{t_{j+1}} - B_{t_j}) + (B_t - B_{t_{N-1}})$, and therefore it is a continuous martingale. Then $I_n(t,\cdot) - I_m(t,\cdot)$ is also a continuous martingale. By Lemma 6.15,

$$P\Big[\sup_{[0,T]} |I_n(t,\omega) - I_m(t,\omega)| > \varepsilon\Big] \leqslant \frac{1}{\varepsilon^2}\,\mathbb{E}[(I_n - I_m)^2] = \frac{1}{\varepsilon^2}\,\mathbb{E}\Big[\int_0^T (\psi_n - \psi_m)^2 ds\Big] \to 0$$

as $n, m \to \infty$. Hence, we can choose a subsequence $n_k \uparrow \infty$ such that

$$P\Big[\sup_{[0,T]} |I_{n_k}(t,\omega) - I_{n_{k+1}}(t,\omega)| > 2^{-k}\Big] < 2^{-k}.$$

By Borelli-Cantelli lemma,

$$P\Big[\forall\, n \geqslant 0,\, \exists\, k \geqslant n : \sup_{t\in[0,T]} |I_{n_{k+1}} - I_{n_k}| > 2^{-k}\Big] = 0,$$

since if $E_k = \{\omega : \sup_{t\in[0,T]} |I_{n_{k+1}} - I_{n_k}| > 2^{-k}\}$, then $P\Big(\bigcap_{n=1}^\infty \bigcup_{k=n}^\infty E_k\Big) = 0$. Hence,

$$P\Big[\exists\, n \geqslant 0,\, \forall\, k \geqslant n : \sup_{t\in[0,T]} |I_{n_{k+1}} - I_{n_k}| \leqslant 2^{-k}\Big] = 1.$$

This implies that for almost all $\omega \in \Omega$, I_{n_k} converges uniformly in $t \in [0,T]$. Let J_t be the limit of I_{n_k}. Clearly J_t is continuous in t a.s. Next, since $I_{n_k}(t,\omega) \to I(t,\omega)$ as $k \to \infty$ for all t in the $L_2(\Omega)$-norm, then $I(t,\cdot) = J_t$ a.s. for all $t \in [0,T]$. The theorem is proved. □

Everywhere below the stochastic integral $\int_0^t f\,dB_s$ means its continuous version.

6.3 An extension of the stochastic intergral and multi-dimensional stochastic intergrals

The stochastic integral can be defined for a larger class of integrands. For this, we relax the measurability Condition 2 of Definition 6.1. Let us assume that there exists a larger filtration $\mathcal{G}_t$, i.e. such that $\mathcal{F}_t^0 \subset \mathcal{G}_t$, $t \geqslant 0$, such that the following replacement for Condition 2 of Definition 6.1

2'. a) B_t is a martingale with respect to $\mathcal{G}_t$;

b) $f(t, \cdot)$ is $\mathcal{G}_t$-adapted.

To emphasize the fact that Condition 2' is used instead of Condition 2, we will use the notation $\Phi(S, T, \mathcal{G}_t)$ for the corresponding class of integrands.

We are going to apply this definition to define a multi-dimensional stochastic integral. Let B_t^k be the k'th coordinate of an n-dimensional Brownian motion B_t, and let

$$\mathcal{F}_t^n = \sigma(B_{s_1}^1, \dots, B_{s_n}^n; s_1, \dots, s_n \leqslant t). \tag{37}$$

Then every Brownian motion B_t^k is a martingale with respect to $\mathcal{F}_t^n$.

Let $\Phi' = \Phi(S, T, \mathcal{F}_t^n)$.

Definition 6.17. *Let $B_t = (B_t^1, \dots B_t^n)$ be an n-dimensional Brownian motion, and let $v = v_{ij}(t, \omega)$ be an $n \times m$ matrix such that $v_{ij} \in \Phi'$ for all $1 \leqslant i \leqslant m$, $1 \leqslant j \leqslant 1$. Let us define the stochastic integral as follows:*

$$\int_S^T v\, dB_t = \int_S^T \begin{pmatrix} v_{11} & \cdots & v_{1n} \\ \vdots & & \vdots \\ v_{n1} & \cdots & v_{mn} \end{pmatrix} \begin{pmatrix} dB_t^1 \\ \vdots \\ dB_t^n \end{pmatrix}$$

where the right-hand side of this formula is an m-dimensional vector whose ith component equals to

$$\sum_{j=1}^n \int_S^T v_{ij}\, dB_t^j.$$

7. Itô's formula

7.1 The one-dimensional Itô formula

Definition 7.1 (1-dimensional Itô process.)**.** *Let B_t be a 1-dimensional $\mathcal{G}_t$-Brownian motion on $(\Omega, \mathcal{F}, P)$. A 1-dimensional Itô process is a stochastic process X_t on $(\Omega, \mathcal{F}, P)$ of the form:*

$$X_t = X_0 + \int_0^t u(s, \omega) ds + \int_0^t v(s, \omega) dB_s \tag{38}$$

where $v \in \Phi(0, T, \mathcal{G}_t)$ for all $T > 0$, and $u \in L_2([0, \infty) \times \Omega)$ is $\mathcal{G}_t$-adapted.

Note that, an Itô process X_t is a semimartingale. For the theorem below, we will need the concept of the *covariance process* of two semimartingales.

Definition 7.2. *Let $X_t = A_t + M_t$ and $Y_t = C_t + N_t$ be two semimartingales where M_t and N_t are their martingale parts. The covariance process $\langle X, Y\rangle_t$ is defined as follows:*

$$\langle X, Y\rangle_t = \langle M, N\rangle_t.$$

Theorem 7.3. *Let X_t be an Itô process given by (38), and let $g(t,x) \in \mathrm{C}^2([0,\infty)\times\mathbb{R})$, i.e. g is twice continuously differentiable on $[0,\infty)\times\mathbb{R})$. Then*

$$Y_t = g(t, X_t)$$

is again an Itô process, and

$$Y_t = g(0, X_0) + \int_0^t \frac{\partial g}{\partial t}(s, X_s)\, ds + \int_0^t \frac{\partial g}{\partial x}(s, X_s)\, dX_s + \frac{1}{2}\int_0^t \frac{\partial^2 g}{\partial x^2}(s, X_s) d\langle X\rangle_s. \tag{39}$$

Remark. In (39), $dX_t = u(t,\omega)\, dt + v(t,\omega)\, dB_t$ according to (38). Also, the last integral is understood as a Riemann-Stieltjes integral with respect to the increasing funcion $\langle X\rangle_s$.

Lemma 7.4. *Let X_t be an Itô process given by (38). Then*

$$\langle X\rangle_t = \Big\langle \int_0^\bullet v(s,\cdot)\, dB_s \Big\rangle_t = \int_0^t v^2(s,\cdot)\, ds.$$

Proof. Left as an exercise. □

Remark. Equivalently (39) can be written as follows:

$$g(t, X_t) = g(0, X_0) + \int_0^t \Big(\frac{\partial g}{\partial s}(s, X_s) + u(s,\cdot)\,\frac{\partial g}{\partial x}(s, X_s) + \frac{1}{2}\, v(s,\cdot)^2\, \frac{\partial^2 g}{\partial x^2}(s, X_s)\Big) ds + \int_0^t v(s,\cdot)\, \frac{\partial g}{\partial x}\, dB_s. \tag{40}$$

Proof of Theorem 7.3. Let

$$\tau_n = \begin{cases} 0, \text{ if } |X_0| > n, \\ \inf\{t:\ \max\big(|\int_0^t v\, dB_s|, \int_0^t |u| ds, \int_0^t v^2 ds\big) > n\}, \text{ if } |X_0| \leqslant n. \end{cases}$$

Clearly $\tau_n \uparrow \infty$ as $n \to \infty$ a.s. Therefore, if we prove (40) for $X_{\tau_n \wedge t}$ on the set $\{\tau_n > 0\}$, then, letting $n \to \infty$, we prove (40) for the general case. Hence, we can assume that X_0, $\int_0^t v\, dB_s$, $\int_0^t |v| ds$, and $\int_0^t v^2 ds$ are bounded in (t,ω), and that the function g has a compact support, and therefore, bounded.

Note that (40) is an Itô process in the sense of Definition 7.1. Let us assume first that u and v and bounded simple functions. Using Taylor's expansion we obtain:

$$g(t,X_t) = g(0,X_0) + \sum_j \Delta g(t_j, X_{t_j}) = g(0,X_0) + \sum_j \frac{\partial g}{\partial t}\,\Delta t_j + \sum_j \frac{\partial g}{\partial x}\,\Delta X_j$$
$$+ \frac{1}{2}\sum_j \frac{\partial^2 g}{\partial t^2}\,(\Delta t_j)^2 + \sum_j \frac{\partial^2 g}{\partial t \partial x}\,(\Delta t_j)(\Delta X_j) + \frac{1}{2}\sum_j \frac{\partial^2 g}{\partial x^2}\,(\Delta X_j)^2 + \sum_j R_j$$

where $\frac{\partial g}{\partial t}$, $\frac{\partial g}{\partial x}$, $\frac{\partial^2 g}{\partial t \partial x}$ etc. are evaluated at the points (t_j, X_{t_j}), $\Delta t_j = t_{j+1} - t_j$, $\Delta X_j = X_{t_{j+1}} - X_{t_j}$, $\Delta g(t_j, X_{t_j}) = g(t_{j+1}, X_{t_{j+1}}) - g(t_j, X_{t_j})$, and $R_j = o\big(|\Delta t_j|^2 + |\Delta X_{t_j}|^2\big)$ for all j. As $\Delta t_j \to 0$,

$$\sum_j \frac{\partial g}{\partial t}\Delta t_j = \sum_j \frac{\partial g}{\partial t}(t_j, X_{t_j})\Delta t_j \to \int_0^t \frac{\partial g}{\partial s}(s, X_s)\,ds$$
$$\sum_j \frac{\partial g}{\partial x}\Delta X_j = \sum_j \frac{\partial g}{\partial x}(t_j, X_{t_j})\Delta X_j \to \int_0^t \frac{\partial g}{\partial x}(s, X_s)\,dX_s.$$

Since we assumed that u and v are simple functions, we obtain:

$$\sum_j \frac{\partial^2 g}{\partial x^2}\,(\Delta X_j)^2 = \sum_j \frac{\partial^2 g}{\partial x^2}\,u_j^2\,(\Delta t_j)^2 + 2\sum_j \frac{\partial^2 g}{\partial x^2} u_j v_j\,(\Delta t_j)(\Delta B_j)$$
$$+ \sum_j \frac{\partial^2 g}{\partial x^2}\,v_j^2\,(\Delta B_j)^2. \quad (41)$$

where $u_j = u(t_j, \omega)$ and $v_j = v(t_j, \omega)$. The first term in (41) is bounded, and tends to zero pointwise as the mesh of the partition $\{0 = t_0 < \cdots < t_n = t\}$ tends to zero. For the second term we have:

$$\mathbb{E}\Big(\sum_j \frac{\partial^2 g}{\partial x^2} u_j v_j\,(\Delta t_j)(B_{t_{j+1}} - B_{t_j})\Big)^2 = \sum_j \mathbb{E}\Big(\frac{\partial^2 g}{\partial x^2} u_j v_j\Big)^2 (\Delta t_j)^3 \to 0 \quad \text{as } \Delta t_j \to 0$$

by boundedness of u, v, and $\frac{\partial^2 g}{\partial x^2}$, and by the fact that $\mathbb{E}[(B_{t_{j+1}} - B_{t_j})^2] = \Delta t_j$. Let us prove that the last term in (41) tends to $\frac{\partial^2 g}{\partial x^2} v^2$ with respect to the $L^2(\Omega)$-norm as the mesh $\max_j |\Delta t_j|$ tends to zero. Define $a(t) = \frac{\partial^2 g}{\partial x^2}(t, X_t)\,v^2(t,\omega)$, and $a_j = a(t_j)$. We have:

$$\mathbb{E}\Big(\sum_j a_j(\Delta B_j)^2 - \sum_j a_j \Delta t_j\Big)^2 = \sum_{i,j} \mathbb{E}\big[a_i a_j\big((\Delta B_i)^2 - \Delta t_i\big)\,\big((\Delta B_j)^2 - \Delta t_j\big)\big]. \quad (42)$$

If $i < j$ then $a_i\, a_j\big((\Delta B_i)^2 - \Delta t_i\big)$ and $\big((\Delta B_j)^2 - \Delta t_j\big)$ are independent, and therefore

$$\mathbb{E}\big[a_i a_j\big((\Delta B_i)^2 - \Delta t_i\big)\,\big((\Delta B_j)^2 - \Delta t_j\big)\big] = \mathbb{E}\big[a_i a_j((\Delta B_i)^2 - \Delta t_i)\,\mathbb{E}[\big((\Delta B_j)^2 - \Delta t_j\big)|\mathcal{F}_{t_j}]\big]$$
$$= 0.$$

Similarly for the case $i > j$. For the case $i = j$, we obtain:

$$\sum_j \mathbb{E}\big[a_j^2\,\big((\Delta B_j)^2 - \Delta t_j)^2\big] = \mathbb{E}[a_j^2]\,\mathbb{E}\big((\Delta B_j)^2 - \Delta t_j\big)^2$$

since a_j^2 is $\mathcal{F}_{t_j}$-measurable and $(\Delta B_j)^2$ is independent of $\mathcal{F}_{t_j}$. Further, we have:

$$\begin{aligned}\mathbb{E}\big((\Delta B_j)^2 - \Delta t_j\big)^2 &= \mathbb{E}\big[(\Delta B_j)^4 - 2(\Delta B_j)^2\Delta t_j + (\Delta t_j)^2\big] \\ &= 3(\Delta t_j)^2 - 2(\Delta t_j)^2 + (\Delta t_j)^2 = 2(\Delta t_j)^2.\end{aligned}$$

Therefore,

$$\sum_j \mathbb{E}\big[a_j^2\,\big((\Delta B_j)^2 - \Delta t_j)^2\big] = 2\sum_j \mathbb{E}[a_j^2]\,(\Delta t_j)^2$$

which tends to zero as the mesh $\max|\Delta t_j|$ goes to zero. This implies, in turn, that the right-hand side of (42) converges to zero as $\max|\Delta t_j| \to 0$. But

$$\sum_j a_j\Delta t_j \to \int_0^t a(s)\,ds, \quad \text{as } \max_j |\Delta t_j| \to 0,$$

and therefore,

$$\sum_j a_j(\Delta B_j)^2 \to \int_0^t a(s)\,ds \quad \text{as } \max_j |\Delta t_j| \to 0.$$

Clearly, $\sum_j R_j \to 0$ as $\max_j |\Delta t_j| \to \infty$. Indeed,

$$\mathbb{E}\Big[\sum_j o\big((\Delta t_j)^2 + (\Delta X_j)^2\big)^2\Big] = \mathbb{E}\,\Big[\sum_j o(\Delta t_j)^3\Big] \to 0 \quad \text{as } \max_j |\Delta t_j| \to \infty.$$

The latter equality holds since

$$\mathbb{E}(\Delta X_j)^2 = v_j^2\,\mathbb{E}(\Delta B_j)^2 = v_j^2\,\Delta t_j + u_j^2\Delta t_j^2,$$

and

$$\begin{aligned}\mathbb{E}(\Delta X_j)^4 &= u_j^4\Delta t_j^4 + 6u_j^2v_j^2\,(\Delta t_j)^2\,\mathbb{E}[(\Delta B_j)^2] + v_j^4 E(\Delta B_j)^4 \\ &= u_j^4\Delta t_j^4 + 6u_j^2v_j^2\,(\Delta t_j)^3 + 3v_j^4\,(\Delta t_j)^3.\end{aligned}$$

This proves Itô's formula for the case when u and v are bounded simple functions To prove the general case let us note that we can always approximate u and v by bounded simple functions with respect to the $L_2(\Omega)$-norm. Passing to the limit in (40) w.r.t the $L_2(\Omega)$-norm we obtain (40) for those u and v for which the Riemann and the stochastic integrals are well defined. This proves Itô's formula. □

7.2 The multi-dimensional Itô formula

Let $B_t = (B_t^1, \dots, B_t^m)$ be an m-dimensional Brownian motion, and let $\mathcal{F}_t^m$ be the filtration defined by (37).

Definition 7.5 (n-dimensional Itô process)**.** *A process X_t is called an n-dimensional **Itô process** if it can be written in the form:*

$$\begin{cases} X_t^1 = X_0^1 + \int_0^t u_1(s,\omega)\,ds + \sum_{j=1}^m \int_0^t v_{1j}(s,\omega)\,dB_s^j \\ \vdots \\ X_t^n = X_0^n + \int_0^t u_n(s,\omega)\,ds + \sum_{j=1}^m \int_0^t v_{nj}(s,\omega)\,dB_s^j \end{cases} \tag{43}$$

where $u_i \in L_2([0,\infty)\times\Omega)$, $1 \leqslant i \leqslant n$, are $\mathcal{F}_t^m$-adapted, and $v_{ij} \in \Phi(0,T,\mathcal{F}_t^m)$ for all $T > 0$, $1 \leqslant i \leqslant n$, $1 \leqslant j \leqslant m$.

In matrix notation we can rewrite (43) as follows:

$$dX_t = u_t dt + v_t dB_t, \tag{44}$$

where

$$X_t = \begin{pmatrix} X_t^1 \\ \vdots \\ X_t^2 \end{pmatrix}, \quad u_t = \begin{pmatrix} u_1(t,\cdot) \\ \vdots \\ u_n(t,\cdot) \end{pmatrix}, \quad v_t = \begin{pmatrix} v_{11} & \cdots & v_{1m} \\ \vdots & & \vdots \\ v_{n1} & \cdots & v_{nm} \end{pmatrix}, \quad dB_t = \begin{pmatrix} dB_t^1 \\ \vdots \\ dB_t^m \end{pmatrix}.$$

Theorem 7.6 (Itô's formula)**.** *Let X_t be an n-dimensional Itô's process given by (44), and let $g(t,x) = (g_1(t,x), \cdots, g_p(t,x)) \in \mathrm{C}^2([0,\infty)\times\mathbb{R}^n, \mathbb{R}^p)$. Then the process*

$$Y_t(\omega) = g(t, X_t)$$

is again an Itô process, whose k'th component Y_t^k is given by

$$Y_t^k = g^k(0,X_0) + \int_0^t \frac{\partial g^k}{\partial s}(s,X_s)\,ds + \sum_{i=1}^n \int_0^t \frac{\partial g^k}{\partial x_i}(s,X_s)\,dX_s^i \\ + \frac{1}{2}\sum_{i=1}^n \sum_{j=1}^n \int_0^t \frac{\partial^2 g^k}{\partial x_i \partial x_j}(s,X_s)\,d\langle X^i, X^j\rangle_s$$

or equivalently,

$$Y_t^k = g^k(0,X_0) + \int_0^t \frac{\partial g^k}{\partial s}(s,X_s)\,ds + \sum_{i=1}^n \int_0^t \frac{\partial g^k}{\partial x_i}(s,X_s)\,u_i(s,\cdot)\,ds \\ + \sum_{i=1}^n \sum_{j=1}^m \int_0^t \frac{\partial g^k}{\partial x_i}(s,X_s)\,v_{ij}(s,\cdot)\,dB_s^j + \frac{1}{2}\sum_{i=1}^n \sum_{j=1}^n \sum_{p=1}^m \int_0^t \frac{\partial^2 g^k}{\partial x_i \partial x_j}(s,X_s)\,(v_{ip}v_{jp})(s,\cdot)\,ds.$$

Lemma 7.7. *Let B_t^i and B_t^j be independent Brownian motions. Then*

$$\langle B^i, B^j\rangle_t = 0.$$

Proof. It suffices to prove that $B_t^i B_t^j$ is an $\mathcal{F}_t$-martingale ($\mathcal{F}_t = \mathcal{F}_t^m$). Indeed, taking into account that both B_t^i and B_t^j are $\mathcal{F}_t$-martingales, for $s < t$ we obtain:

$$0 = \mathbb{E}[(B_t^i - B_s^i)]\,\mathbb{E}[(B_t^j - B_s^j)] = \mathbb{E}[(B_t^i - B_s^i)(B_t^j - B_s^j)] = \mathbb{E}[(B_t^i - B_s^i)(B_t^j - B_s^j)\,|\,\mathcal{F}_s]$$
$$= \mathbb{E}[B_t^i B_t^j\,|\,\mathcal{F}_s] - B_s^i B_s^j.$$

□

Lemma 7.8. *Let X_t^i and X_t^j be two 1-dimensional Itô's processes with the martingale parts $\int_0^t v_i(s,\cdot)\,dB_s^i$ and resp. $\int_0^t v_j(s,\cdot)\,dB_s^j$. Then*

$$\langle X^i, X^j\rangle_t = \Big\langle \int_0^\bullet v_i(s,\cdot)\,dB_s^i\,, \int_0^\bullet v_j(s,\cdot)\,dB_s^j \Big\rangle_t = \int_0^t (v^i v^j)(s,\cdot)\,d\langle B^i, B^j\rangle_s$$
$$= \int_0^t (v^i v^j)(s,\cdot)\,\delta_{ij}\,ds$$

where $\delta_{ij} = 0$ if $i \neq j$ and $\delta_{ii} = 1$.

Proof. Left as an exercise. □

The proof of Theorem 7.6 is similar to the proof of Theorem 7.3, and therefore is omitted.

7.3 The Itô representation theorem

Let $B_t = (B_t^1, \ldots, B_t^n)$ be an n-dimensional Brownian motion and let $\mathcal{F}_t = \mathcal{F}_t^n$ be the filtration defined by (37).

Lemma 7.9. *The linear span of random variables of the type*

$$\exp\Big\{\int_0^T h(t)dB_t - \frac{1}{2}\int_0^T h^2(t)\,dt\Big\} \tag{45}$$

where $h \in L_2([0,T])$ is deterministic, is dense in $L_2(\mathcal{F}_T, \Omega)$.

Proof. Without proof. □

Theorem 7.10 (Itô's representation theorem)**.** *Let $B_t = (B_t^1, \ldots, B_t^n)$ be an n-dimensional Brownian motion, and let the filtration $\mathcal{F}_t^n$ be defined by (37). Let $F \in L_2(\mathcal{F}_T^n, \Omega)$. Then there exists a unique stochastic process $f(t,\omega) \in \Phi(0,T,\mathcal{F}_t^n)$ such that*

$$F(\omega) = \mathbb{E}[F] + \int_0^T f(t,\omega)\,dB_t. \tag{46}$$

Proof. For simplicity we consider the case $n = 1$. The proof in the general case is similar. First we assume that F has the form (45), i.e.

$$F(\omega) = \exp\Big(\int_0^T h(t)dB_t - \frac{1}{2}\int_0^T h^2(t)dt\Big) \tag{47}$$

for some $h(t) \in L_2[0,T]$. Define

$$Y_t = \exp\Big(\int_0^t h(s)dB_s - \frac{1}{2}\int_0^t h^2(s)ds\Big), \quad 0 \leqslant t \leqslant T. \tag{48}$$

By Itô's formula applied to $Y_t = \exp(X_t)$ where $X_t = \int_0^T h(t)dB_t - \frac{1}{2}\int_0^T h^2(t)dt$. Note that $\frac{d}{dx}\exp(X_t) = \frac{d^2}{dx^2}\exp(X_t) = \exp(X_t) = Y_t$. Also, note that by Lemma 7.8,

$$\langle X\rangle_t = \Big\langle \int_0^\bullet h(s)dB_s\Big\rangle_t = \int_0^t h^2(s)\,ds.$$

We obtain:

$$Y_t = Y_0 + \int_0^t Y_s\,\big(h(s)\,dB_s - \frac{1}{2}h^2(s)\,ds\big) + \frac{1}{2}\int_0^t Y_s\,h^2(s)\,ds = 1 + \int_0^t Y_s\,h(s)\,dB_s.$$

This implies that $\mathbb{E}[F] = 1$, and that formula (46) holds for the dase when F is given by (47). By linearity, (46) holds for linear combinations of funcions of form (45).

Now let $F \in L_2(\mathcal{F}_T, \Omega)$ be arbitrary. We approximate F by linear combinations F_n of functions of form (45). For every n we have:

$$F_n(\omega) = \mathbb{E}[F_n] + \int_0^T f_n(s,\omega)\,dB_s(\omega)$$

where $f_n \in \Phi(0,T)$. By Itô's isometry and by the fact that $\mathbb{E}\Big[\int_0^T (f_n - f_m)\,dB_s\Big] = 0$, we obtain:

$$\begin{aligned}
&\mathbb{E}[(F_n - F_m)^2] = \mathbb{E}\Big[\Big(\mathbb{E}[F_n - F_m] + \int_0^T (f_n - f_m)\,dB_s\Big)^2\Big] \\
&= \big(\mathbb{E}[F_n - F_m]\big)^2 + 2\,\mathbb{E}[F_n - F_m]\mathbb{E}\Big[\int_0^T (f_n - f_m)\,dB_s\Big] + \mathbb{E}\Big(\int_0^T (f_n - f_m)\,dB_s\Big)^2 \\
&\qquad = \big(\mathbb{E}[F_n - F_m]\big)^2 + \int_0^T \mathbb{E}[(f_n - f_m)^2]\,dt.
\end{aligned}$$

Therefore

$$\int_0^T \mathbb{E}[(f_n - f_m)^2]\,dt = \mathbb{E}[(F_n - F_m)^2] - \big(\mathbb{E}[F_n - F_m]\big)^2 \to 0, \quad \text{as } n,m \to \infty.$$

Hence f_n is a Cauchy sequence in $L_2(\Omega \times [0,T])$, and hence converges to some $f \in L_2(\Omega \times [0,T])$. This implies that $f \in \Phi(0,T)$. Indeed, there is a subsequence of $\{f_n\}$ that converges to f for almost all $(\omega,t) \in \Omega \times [0,T]$. Therefore, f is $\mathcal{F}_t$-measurable for almost all $t \in [0,T]$. By modifying f on a t-set of measure zero we obtain that $f(t,\omega)$ is $\mathcal{F}_t$- adapted. Similar, the map $(t,\omega) \mapsto f(t,\omega)$ is $\mathfrak{B} \times \mathcal{F}$-measurable. Clearly $\mathbb{E}\int_0^T f^2(t,\cdot)dt < \infty$. We obtain:

$$F = \lim_{n\to\infty} F_n = \lim_{n\to\infty}\Big(\mathbb{E}[F_n] + \int_0^T f_n\,dB_t\Big) = \mathbb{E}[F] + \int_0^T f\,dB_t$$

where the limit is taken in $L_2(\mathcal{F}_T, \Omega)$. Also, we used Itô's isometry to show the convergence of stochastic integrals.

The uniqueness follows from Itô's isometry. Suppose there exists two different functions $f_1, f_2 \in \Phi(0,T)$ such that

$$F(\omega) = \mathbb{E}[F] + \int_0^T f_1(t,\omega)\, dB_t = \mathbb{E}[F] + \int_0^T f_2(t,\omega)\, dB_t.$$

Then

$$\int_0^T (f_1 - f_2)\, dB_t = 0,$$

and therefore

$$\mathbb{E}\int_0^T (f_1 - f_2)^2 dt = \mathbb{E}\Big(\int_0^T (f_1 - f_2)\, dB_t\Big)^2 = 0.$$

This implies that $f_1(t,\omega) = f_2(t,\omega)$ for almost all $(t,\omega) \in [0,T] \times \Omega$. □

7.4 The martingale representation theorem

Theorem 7.11. *Let $B_t = (B_t^1, \dots, B_t^n)$ be an n-dimensional Brownian motion, and let M_t be an $\mathcal{F}_t^n$-martingale, where the filtration $\mathcal{F}_t^n$ is defined by (37), such that $M_t \in L_2(\Omega)$ for all $t \geqslant 0$. Then there exists a unique stochastic process $g(s,\omega)$ such that $g(s,\,\cdot\,) \in \Phi'(0,t)$ for all $t \geqslant 0$ with respect to $\mathcal{F}_s^n$, and*

$$M_t = \mathbb{E}[M_0] + \int_0^t g(s,\omega)\, dB_s \quad \textit{a.s., for all}\, t \geqslant 0.$$

Proof. Again, we prove the theorem for the case when $n = 1$ for simplicity. By Itô's representation theorem (Theorem 7.11), which we apply to $T = t$ and $F = M_t$, we obtain that for every $t \geqslant 0$, there exists a unique $f^t(s,\omega) \in L_2(\Omega, \mathcal{F}_t$ such that

$$M_t = \mathbb{E}[M_t] + \int_0^t f^t(s,\omega)\, dB_s = \mathbb{E}[M_0] + \int_0^t f^t(s,\omega)\, dB_s, \tag{49}$$

since for the martingale M_t it holds that $\mathbb{E}[M_t] = \mathbb{E}[M_0]$. Note that in the above formula, the integrand of the stochastic integral depends on t. Let us prove that $f^t(s,\omega)$ is actually does not depend on t. Assume $0 \leqslant t_1 < t_2$. Then,

$$\begin{aligned} M_{t_1} = \mathbb{E}[M_{t_2} \,|\, \mathcal{F}_{t_1}] = \mathbb{E}[M_0] + \mathbb{E}\Big[\int_0^{t_2} f^{t_2}(s,\omega)\, dB_s \,|\, \mathcal{F}_{t_1}\Big] \\ = \mathbb{E}[M_0] + \int_0^{t_1} f^{t_2}(s,\omega)\, dB_s. \end{aligned}$$

On the other hand, by (49),

$$M_{t_1} = \mathbb{E}[M_0] + \int_0^{t_1} f^{t_1}(s,\omega)\, dB_s.$$

By the uniqueness result of Itô's representation theorem (Theorem **??**), for almost all $s \in [0, t_1]$,

$$f^{t_1}(s, \omega) = f^{t_2}(s, \omega) \quad \text{a.s..}$$

Let us define $f(s, \omega)$ for almost all $s \in [0, \infty)$ by setting

$$f(s, \omega) = f^N(s, \omega) \quad \text{a.s. for almost all } s \in [0, N]$$

We obtain that for almost all $t \geqslant 0$,

$$M_t = \mathbb{E}[M_0] + \int_0^t f^t(s, \omega)\, dB_s = \mathbb{E}[M_0] + \int_0^t f(s, \omega)\, dB_s.$$

By t-continuity of the stochastic integral, the latter equality holds for all $t \geqslant 0$ a.s. □

8. Stochastic Differential Equations

Stochastic differential equations can be defined in several contexts that vary their generality. Informally, a stochastic differential equation (SDE) can be written in the form:

$$dX_t = b(t, X_t)\, dt + \sigma(t, X_t)\, dB_t.$$

Below we define weak and strong solutions of the above SDE and prove the existence and uniqueness theorem.

8.1 Formal definitions

Let $T > 0$, and let the functions $b : [0, T] \times \mathbb{R}^n \to \mathbb{R}^n$ and $\sigma : [0, T] \times \mathbb{R}^n \to \mathbb{R}^{n \times m}$ be measurable. Further let Z be a random variable, and let $\mathcal{F}_t^Z$ be the filtration generated by Z and B_s, $s \leqslant t$.

Definition 8.1. *A process X_t is called a **strong solution** to the SDE*

$$dX_t = b(t, X_t)\, dt + \sigma(t, X_t)\, dB_t \tag{50}$$

on $[0, T]$ with the initial condition $X_0 = Z$ if X_t is $\mathcal{F}_t^Z$-adapted and

$$X_t = Z + \int_0^t b(s, X_s)\, ds + \int_0^t \sigma(s, X_s)\, dB_s.$$

Definition 8.2. *A pair of processes $(\tilde{X}_t, \tilde{B}_t)$ is called a **weak solution** to SDE (50) on $[0, T]$ with the initial condition $X_0 = Z$ if there exists a filtered probability space $(\Omega, \mathcal{H}_t, \mathcal{H}, P)$ such that $\tilde{B}_t$ is an $\mathcal{H}_t$-Brownian motion and $\tilde{X}_t$ is $\mathcal{H}_t$-adapted and*

$$\tilde{X}_t = Z + \int_0^t b(s, \tilde{X}_s)\, ds + \int_0^t \sigma(s, \tilde{X}_s)\, d\tilde{B}_s.$$

Remark. We say that $\tilde{B}_t$ is an $\mathcal{H}_t$-Brownian motion if $\tilde{B}_t$ is a Brownian motion and an $\mathcal{H}_t$-martingale, i.e. $\mathbb{E}[\tilde{B}_{t+h} \,|\, \mathcal{H}_t] = \tilde{B}_t$ for all $t, h \geqslant 0$.

8.2 An existence and uniqueness theorem

Theorem 8.3 (Existence and uniqueness). *Let, as before, $T > 0$ and let the functions $b : [0,T] \times \mathbb{R}^n \to \mathbb{R}^n$ and $\sigma : [0,T] \times \mathbb{R}^n \to \mathbb{R}^{n\times m}$ be measurable and satisfying conditions (i) and (ii) below:*

(i) $|b(t,x)| + |\sigma(t,x)| \leqslant C\,(1+|x|)$, $x \in \mathbb{R}^n$, $t \in [0,T]$, *for some constant* $C > 0$*;*

(ii) $|b(t,x) - b(t,y)| + |\sigma(t,x) - \sigma(t,y)| \leqslant D|x-y|$, $x, y \in \mathbb{R}^n$, $t \in [0,T]$, *for some constant* $D > 0$.

Let Z be a random variable which is independent of the σ-algebra $\mathcal{F}^m_\infty$ and such that

$$\mathbb{E}|Z|^2 < \infty.$$

Then SDE (50) has a unique strong t-continuous solution X_t on $[0,T]$ with the initial condition $X_0 = Z$ and with the property that X_t is adapted with respect to the filtration $\mathcal{F}^Z_t$ and

$$\mathbb{E}\Big[\int_0^T |X_t|^2 dt\Big] < \infty. \tag{51}$$

Lemma 8.4 (Gronwall's lemma). *Let α, β, and u be real-valued functions defined on the interval $[a,b]$. Assume that β and u are continuous, β is non-negative, and α is integrable on $[a,b]$. Then if u satisfies the inequality*

$$u(t) \leqslant \alpha(t) + \int_a^t \beta(s)u(s)\,ds,$$

then

$$u(t) \leqslant \alpha(t) + \int_a^t \alpha(s)\beta(s)\big[\exp\int_s^t \beta(r)\,dr\big]\,ds.$$

If $\alpha(t)$ is a constant, say equal to α, then

$$u(t) \leqslant \alpha\,\exp\int_a^t \beta(s)\,ds.$$

If $\beta(t)$ is a constant, say equal to β, then

$$u(t) \leqslant \alpha\,e^{\beta\,(t-a)}.$$

Proof. Left as an exercise. □

Proposition 8.5 (Itô's isometry for a multi-dimensional stochastic integral). *Let $B_t = (B^1_t, \ldots, B^m_t)$ be an m-dimensional Brownian motion, and let the function $v : [S,T] \times \Omega \to \mathbb{R}^{n\times m}$ be such that $v_{ij} \in \Phi(S,T,\mathcal{F}^m_t)$, $1 \leqslant i \leqslant n$, $1 \leqslant j \leqslant m$, with respect to the filtration $\mathcal{F}^m_t$. Then,*

$$\mathbb{E}\Big|\int_S^T v(t,\omega)\,dB_t\Big|^2 = \mathbb{E}\int_S^T |v(t,\omega)|^2\,dt.$$

Proof. Left as an exercise. $\square$

Proof of Theorem 8.3. Uniqueness. The uniqueness follows from Itô's isometry and Lipschitz property (ii). Indeed, let X_t and $\hat{X}_t$ be two solutions with initial values Z and resp. $\hat{Z}$, i.e. $X_0 = Z$, $\hat{X}_0 = \hat{Z}$. In the proof of uniqueness we are only interested in case $Z = \hat{Z}$. The following more general estimate which also holds when $Z \neq \hat{Z}$ will be useful later. Define $a(s,\omega) = b(s, X_s) - b(s, \hat{X}_s)$ and $\gamma(s,\omega) = \sigma(s, X_s) - \sigma(s, \hat{X}_s)$. Then,

$$
\begin{aligned}
\mathbb{E}[|X_t - \hat{X}_t|^2] &= \mathbb{E}\Big[\Big(Z - \hat{Z} + \int_0^t a(s,\,\cdot\,)\,ds + \int_0^t \gamma(s,\,\cdot\,)\,dB_s\Big)^2\Big] \\
&\leqslant 3\,\mathbb{E}[|Z - \hat{Z}|^2] + 3\,\mathbb{E}\Big[\Big(\int_0^t a(s,\,\cdot\,)\,ds\Big)^2\Big] + 3\,\mathbb{E}\Big[\Big(\int_0^t \gamma(s,\,\cdot\,)\,dB_s\Big)^2\Big] \\
&\leqslant 3\,\mathbb{E}[|Z - \hat{Z}|^2] + 3t\,\mathbb{E}\Big[\int_0^t a^2(s,\,\cdot\,)\,ds\Big] + 3\,\mathbb{E}\Big[\int_0^t \gamma^2(s,\,\cdot\,)\,ds\Big] \qquad (52) \\
&\leqslant 3\,\mathbb{E}[|Z - \hat{Z}|^2] + 3(1+t)\,D^2 \int_0^t \mathbb{E}|X_s - \hat{X}_s|^2\,ds. \qquad (53)
\end{aligned}
$$

Denote $v(t) = \mathbb{E}|X_t - \hat{X}_t|^2$, $t \in [0,T]$. Let $F = 3\,\mathbb{E}[|Z - \hat{Z}|^2]$ and $A = 3(1+T)\,D^2$. In this notations the above inequality can be written as follows:

$$
v(t) \leqslant F + A \int_0^t v(s)\,ds.
$$

By Gronwall's lemma,

$$
v(t) \leqslant F \exp(A\,t).
$$

Now we assume that $Z = \hat{Z}$, and therefore $F = 0$. Hence $v(t) = 0$ for all $t \geqslant 0$. This implies that for all $t \geqslant 0$, $X_t = \hat{X}_t$ a.s. This means that for every $t \geqslant 0$, there exists a set Ω_t of full P-measure ($P(\Omega_t) = 1$) such that $X_t = \hat{X}_t$ everywhere on Ω_t. Now since the set of rational numbers is countable we conclude that there exists a set Ω' of full P-measure such that $X_t = \hat{X}_t$ for all $t \in \mathbb{Q} \cap [0,T]$ everywhere on Ω'. By continuity of the map $t \mapsto |X_t - \hat{X}_t|$,

$$
X_t = \hat{X}_t \quad \text{for all } t \in [0,T]
$$

everywhere on Ω'. This proves the uniqueness.

Existence. Define $Y_t^{(0)} = X_0$, and $Y_t^{(k)}$ inductively as follows:

$$
Y_t^{(k+1)} = X_0 + \int_0^t b(s, Y_s^{(k)})\,ds + \int_0^t \sigma(s, Y_s^{(k)})\,dB_s. \qquad (54)
$$

Hence,

$$
Y_t^{(k+1)} - Y_t^{(k)} = \int_0^t \big(b(s, Y_s^{(k)}) - b(s, Y_s^{(k-1)})\big)\,ds + \int_0^t \big(\sigma(s, Y_s^{(k)}) - \sigma(s, Y_s^{(k-1)})\big)\,dB_s.
$$

Performing a similar computation as in (53) but with $Y_t^{(k+1)}$ and $Y_t^{(k)}$ instead of X_t and resp. $\hat{X}_t$ on the left-hand side, and with $Y_t^{(k)}$ and $Y_t^{(k-1)}$ instead of X_t and resp. $\hat{X}_t$ on the right-hand side we obtain:

$$\mathbb{E}|Y_t^{(k+1)} - Y_t^{(k)}|^2 \leqslant 2\,\mathbb{E}\Big(\int_0^t \big(b(s,Y_s^{(k)}) - b(s,Y_s^{(k-1)})\big)\,ds\Big)^2$$
$$+ 2\,\mathbb{E}\Big(\int_0^t \big(\sigma(x,Y_s^{(k)}) - \sigma(x,Y_s^{(k-1)})\big)\,dB_s\Big)^2 \leqslant 2(1+T)\,D^2 \int_0^t \mathbb{E}|Y_s^{(k)} - Y_s^{(k-1)}|^2\,ds \tag{55}$$

for $k \geqslant 1$ and $t \leqslant T$. Now recalling that $Y^{(0)} = X_0$, and taking into account condition (i) of the theorem saying that the functions b and σ have no more than the linear growth we obtain:

$$\begin{aligned}\mathbb{E}|Y_t^{(1)} - Y_t^{(0)}|^2 &= \mathbb{E}\Big(\int_0^t b(s,X_0)\,ds + \int_0^t \sigma(s,X_0)\,dB_s\Big)^2\\ &\leqslant 2\,\mathbb{E}\Big(\int_0^t b(s,X_0)\,ds\Big)^2 + 2\,\mathbb{E}\Big(\int_0^t \sigma(s,X_0)\,dB_s\Big)^2\\ &\leqslant 4\,C^2\,t^2\,\big(1+\mathbb{E}|X_0^2|\big) + 4\,C^2\,t\,\big(1+\mathbb{E}|X_0^2|\big) \leqslant A_1 t\end{aligned}$$

where $A_1 = 4\,C^2\big(1+\mathbb{E}|X_0^2|\big)(1+T)$. By induction on k and inequality (55) we obtain that there exists a constant A_2 depending on C, D, T, and $\mathbb{E}|X_0|^2$ such that

$$\mathbb{E}|Y_t^{(k+1)} - Y_t^{(k)}|^2 \leqslant \frac{A_2^{k+1}\,t^{k+1}}{(k+1)!} \quad k \geqslant 0, t \in [0,T].$$

By Chebyshev's and Doob's maximal inequalities,

$$\begin{aligned}P\Big[\sup_{0\leqslant t\leqslant T} |Y_t^{(k+1)} - Y_t^{(k)}| > 2^{-k}\Big] &\leqslant 4^k \mathbb{E}\big[\sup_{0\leqslant t\leqslant T} |Y_t^{(k+1)} - Y_t^{k}|^2\big]\\ &\leqslant 4^{k+1}\,\mathbb{E}\,|Y_T^{(k+1)} - Y_T^{(k)}|^2 \leqslant \frac{(4A_2)^{k+1}\,T^{k+1}}{(k+1)!}.\end{aligned}$$

The latter inequality implies that the series

$$\sum_{k=0}^{\infty} P\Big[\sup_{0\leqslant t\leqslant T} |Y_t^{(k+1)} - Y_t^{(k)}| > 2^{-k}\Big]$$

converges. Recall the Borel-Cantelli lemma that says that if E_k are events such that

$$\sum_{k=0}^{\infty} P(E_k) < \infty,$$

then

$$P\Big(\bigcap_{n=0}^{\infty}\bigcup_{k=n}^{\infty} E_k\Big) = 0,$$

or, equivalently,

$$P\Big(\bigcup_{n=0}^{\infty}\bigcap_{k=n}^{\infty} E_k^c\Big) = 1,$$

where $E_k^c = \Omega \diagdown E_k$. In our context the latter means that

$$P\Big(\exists\, n \leqslant 0,\ \forall\, k \geqslant n,\ \sup_{0\leqslant t\leqslant T} |Y_t^{(k+1)} - Y_t^{(k)}| \leqslant 2^{-k}\Big) = 1.$$

This implies that the sequence

$$Y_t^{(n)} = Y_t^{(0)} + \sum_{k=0}^{n-1}(Y_t^{(k+1)} - Y_t^{(k)})$$

converges uniformly in $t \in [0,T]$ a.s. Let X_t be the uniform limit of this sequence, i.e.

$$X_t = \lim_{n\to\infty} Y_t^{(n)}$$

which exists almost surely. Then a.s. X_t is continuous in t as a uniform limit of the t-continuous processes $Y_t^{(n)}$. Moreover, note that X_t is $\mathcal{F}_t^Z$-measurable for all t since for all n, $Y_t^{(n)}$ is $\mathcal{F}_t^Z$-measurable for all t Since $Y_t^{(n)}$ converges to X_t uniformly in $t \in [0,T]$ a.s., then it also converges to X_t with respect to the $L_2(\Omega \times [0,T])$-norm and with respect to the $L_2(\Omega)$-norm for each $t \in [0,T]$. In particular, this implies (51).

It remains to show that X_t verifies SDE (50). For all n we have:

$$Y_t^{(n+1)} = X_0 + \int_0^t b(s, Y_s^{(n)})\, ds + \int_0^t \sigma(x, Y_s^{(n)})\, dB_s. \tag{56}$$

Note that $\lim_{n\to\infty} \mathbb{E}|Y_t^{(n+1)} - X_t|^2 = 0$. On the other hand, by Condition (ii) of the theorem

$$\begin{aligned}\mathbb{E}\Big|\int_0^t (b(s, Y_s^{(n)}) - b(s, X_s)\, ds\Big|^2 &\leqslant t\,\mathbb{E}\int_0^t |b(s, Y_s^{(n)}) - b(s, X_s)|^2\, ds \\ &\leqslant D^2\, t\,\mathbb{E}\int_0^t |Y_s^{(n)} - X_s|^2 \to 0 \quad \text{as}\, n \to \infty. \end{aligned}\tag{57}$$

Next by Itô's isometry and again Condition (ii) of the theorem, we obtain:

$$\begin{aligned}\mathbb{E}\Big|\int_0^t (\sigma(s, Y_s^{(n)}) - \sigma(s, X_s))\, dB_s\Big|^2 &= \mathbb{E}\int_0^t |\sigma(s, Y_s^{(n)}) - \sigma(s, X_s)^2|^2\, ds \\ &\leqslant D^2\,\mathbb{E}\int_0^t |Y_s^{(n)} - X_s|^2 \to 0 \quad \text{as}\, n \to \infty. \end{aligned}\tag{58}$$

Now (56), (57), and (58) imply that X_t is a (strong) solution to SDE (50). $\square$

Definition 8.6. *We say that a solution* (X_t, B_t) *to SDE* (50) *is* ***weakly unique*** *if for any other weak solution* $(\tilde{X}_t, \tilde{B}_t)$ *to* (50) *it holds that* (X_t, B_t) *and* $(\tilde{X}_t, \tilde{B}_t)$ *are identical in law.*

Theorem 8.7 (Weak uniqueness)**.** *Let* σ *and* b *satisfy assumptions of Theorem 8.3. Then a weak or strong solution to SDE* (50) *is weakly unique.*

Remark. Clearly, every strong solution to (50) is also a weak solution. When we say a strong solution X_t is "weakly unique" we mean that the solution (X_t, B_t) is weakly unique.

Sketch of proof. We prove the theorem for the case when $\sigma(t,x)$ and $b(t,x)$ are bounded and continuous for all $(t,x) \in [0,T] \times \mathbb{R}^n$. Let (X_t, B_t) and $(\tilde{X}_t, \tilde{B}_t)$ be two weak solutions on probability spaces $(\Omega, P, \mathcal{H}_t, \mathcal{H})$ and $(\tilde{\Omega}, \tilde{P}, \tilde{\mathcal{H}}_t, \tilde{\mathcal{H}})$, respectively. Further let Y_t and $\tilde{Y}_t$ be strong solutions on $(\Omega, P, \mathcal{H}_t, \mathcal{H})$ and $(\tilde{\Omega}, \tilde{P}, \tilde{\mathcal{H}}_t, \tilde{\mathcal{H}})$ constructed with the help of B_t and resp. $\tilde{B}_t$. By the uniqueness of a strong solution, $X_t = Y_t$ and $\tilde{X}_t = \tilde{Y}_t$ for all t, a.s. Therefore it suffices to prove that Y_t and $\tilde{Y}_t$ are identical in law. We prove that by induction using iteration procedure (55). Note that the processes $(Y_t^{(k)}, B_t)$ and $(\tilde{Y}_t^{(k)}, \tilde{B}_t)$ are identical in law for all k. Indeed, $(Y_t^{(0)}, B_t)$ and $(\tilde{Y}_t^{(0)}, \tilde{B}_t)$ are identical in low because $Y_t^{(0)} = \tilde{Y}_t^{(0)} = Z$, and on the other hand, B_t and $\tilde{B}_t$ are identical in law. Suppose we know that $Y_t^{(k-1)}$ and $\tilde{Y}_t^{(k-1)}$ are identical in law. Also, $Y_t^{(k-1)}$ and $\tilde{Y}_t^{(k-1)}$ are t-continuous. Then $\sigma(t, Y_t^{(k-1)})$, $b(t, Y_t^{(k-1)})$, $\sigma(t, \tilde{Y}_t^{(k-1)})$, and $b(t, \tilde{Y}_t^{(k-1)})$ are continuous in t. Therefore the integrals $\int_0^t \sigma(s, Y_s^{(k-1)})\, dB_s$ and $\int_0^t \sigma(s, \tilde{Y}_s^{(k-1)})\, d\tilde{B}_s$ can be approximated by the sums $\sum_{i=0}^{N-1} \sigma(t_i, Y_{t_i}^{(k-1)})(B_{t_{i+1}} - B_{t_i})$ resp. $\sum_{i=0}^{N-1} \sigma(t_i, \tilde{Y}_{t_i}^{(k-1)})(\tilde{B}_{t_{i+1}} - \tilde{B}_{t_i})$ which are clearly identically in law. Therefore their L_2-limits, i.e. the above stochastic integrals are identical in law as well. Analogously, the integrals $\int_0^t b(s, Y_s^{(k-1)})\, ds$ and $\int_0^t b(s, Y_s^{(k-1)})\, ds$ are identical in law. The above arguments and formula (55) imply that $Y_t^{(k)}$ and $\tilde{Y}_t^{(k)}$ are identical by law and t-continuous. By induction, $(Y_t^{(k)}, B_t)$ and $(\tilde{Y}_t^{(k)}, \tilde{B}_t)$ are identical in law for all k. Therefore their L_2 limits (Y_t, B_t) and $(\tilde{Y}_t, \tilde{B}_t)$ are also identical in law. The weak uniqueness is proved. $\square$

8.3 The Tanaka equation

There are stochastic differential equations which do not have *strong solutions* but have a *unique weak solution*. The example below discusses such an SDE. Consider the one-dimensional stochastic differential equation:

$$dX_t = \operatorname{sign}(X_t)\, dB_t, \quad X_0 = 0 \tag{59}$$

where

$$\operatorname{sign}(x) = \begin{cases} +1 & x \geqslant 0, \\ -1, & x < 0. \end{cases}$$

Note that here $\sigma(x,t) = \operatorname{sign}(x)$ does not satisfy the Lipschitz condition (Condition (ii) of Theorem 8.3), so Theorem 8.3 cannot be applied here.

Proposition 8.8. *SDE (59) has no strong solution.*

Lemma 8.9 (Tanaka's formula)**.** *Let B_t be a 1-dimensional Brownian motion. Then,*

$$|B_t| = |B_0| + \int_0^t \operatorname{sign}(B_s)\, dB_s + L_t(\omega), \tag{60}$$

where

$$L_t = \lim_{\varepsilon \to 0} \frac{1}{2\varepsilon} \lambda\{s \in [0,t] : B_s \in (-\varepsilon, \varepsilon)\}$$

and λ is the Lebesgue measure.

Remark. L_t is called the *local time* for a Brownian motion at 0. Formula (60) is known as *Tanaka's formula* for a Brownian motion.

Proof of Lemma 8.9. Without proof. □

Lemma 8.10 (How to recognize a Brownian motion)**.** *Let B_t be an m-dimensional Bronian motion, and $v : [0,\infty) \times \Omega \to \mathbb{R}^{n\times m}$ satisfy the assumptions (see Section 6.3) under which the stochastic integral is well defined. Further let*

$$Y_t = \int_0^t v(s,\omega)\, dB_s.$$

Then Y_t be a Brownian motion if and only if

$$v(t,\omega)v^T(t,\omega) = I_n$$

for almost all (t,ω). I_n in the above formula is the $n \times n$ identity matrix.

Proof. Without proof. □

Proof of Proposition 8.8. Let $\hat{B}_t$ be a Brownian motion and let $\hat{\mathcal{F}}_t$ be its natural filtration (i.e. $\sigma(\hat{B}_s, s \leqslant t)$). Define the process

$$Y_t = \int_0^t \operatorname{sign}(\hat{B}_s)\, d\hat{B}_s.$$

By Tanaka's formula (Lemma 8.9),

$$Y_t = |\hat{B}_t| - |\hat{B}_0| - \hat{L}_t,$$

where $\hat{L}_t$ is the local time for $\hat{B}_t$ at 0. Let $\mathcal{G}_t = \sigma(|\hat{B}_s|, s \leqslant t)$. Clearly $\mathcal{G}_t \subset \hat{\mathcal{F}}_t$, where the inclusion is strict. Hence, the σ-algebra $\mathcal{N}_t = \sigma(Y_s, s \leqslant t) \subset \hat{\mathcal{F}}_t$ where the inclusion is strict as well.

Now suppose X_t is a strong solution to (59). Then, since $\big(\operatorname{sign}(X_r)\big)^2 = 1$, by Lemma 8.10, X_t is a Brownian motion starting at 0. Multiplying both parts of (59) by $\operatorname{sign}(X_t)$ we obtain:

$$dB_t = \operatorname{sign}(X_t)\, dX_t.$$

Let us apply the above argument to $\hat{B}_t = X_t$ (since we proved that X_t is a Brownian motion) and $Y_t = \int_0^t \mathrm{sign}(X_r)\, dX_r = B_t$. We obtain that the natural filtration $\mathcal{F}_t$ of B_t is strictly contained in $\mathcal{M}_t = \sigma(X_s, s \leqslant t)$. This contradicts to the fact that X_t is a strong solution since a strong solution is $\mathcal{F}_t$-adapted. Thus, a stong solution to (59) does not exist. $\square$

Let us find a weak solution to (59).

Lemma 8.11. *SDE (59) has a weak solution.*

Proof. Choose X_t to be any Brownian motion $\hat{B}_t$. Let us define

$$\tilde{B}_t = \int_0^t \mathrm{sign}(\hat{B}_s)\, d\hat{B}_s = \int_0^t \mathrm{sign}(X_s)\, dX_s.$$

Again, by Lemma 8.10, $\tilde{B}_t$ is a Brownian motion. Finally the SDE

$$d\tilde{B}_t = \mathrm{sign}(X_t) dX_t$$

is equivalent to

$$dX_t = \mathrm{sign}(X_t)\, d\tilde{B}_t$$

which can be obtained by means of multiplication of the both parts by $\mathrm{sign}(X_t)$. Therefore, $(X_t, \tilde{B}_t)$ is a weak solution.

Let us prove the weak uniqueness. Let $(\bar{X}_t, \bar{B}_t)$ be another weak solution. Another application of Lemma 8.10 shows that $\bar{X}_t$ is a Brownian motion. Hence, all weak solutions are identical in law. $\square$

9. Application to Mathematical finance

9.1 Pricing and hedging financial options

A *call option* on stock is a contract that gives its holder the right to buy this stock in the future at the price K written in the contract, called the *exercise price* or the *strike price*.

Example. On April we are offered the opportunity to buy shares of a company **A**. Currently the shares are valued at €1 each. Suppose that except of buying shares, we can also buy an ‘option’. Specifically, for a cost of €0.20 we can buy a ticket that gives us the right to buy one share of company **A** for €1.20 on August 1st, irrespective of the actual market value of this share. Suppose we buy 1000 of these tickets. August 1st arrives, and the shares are now worth €1.80 each. We then exercise our option to buy 1000 shares at €1.20 each and sell them immediately at their market value to make a profit of €600 (€400 if you include the cost of the options). Alternatively, suppose that the shares drop to €0.70 each. In this case, we choose not to exercise the option to purchase the shares and throw all my tickets away. We make then an overall profit of €0 (or a loss of €200, if I include the cost of the tickets).

This is an example of a *European call option*. Specifically, A *European call option* allows the holder to exercise the contract (that is, to buy this stock at K) at a particular date T, called the *maturity* or the *expiration date*.

Now we introduce some general options and notations. Our market (a simplified model) consists of stock of a single type and a riskless investment such as a bank account. We model the value in time of a single unit of stock as a stochastic process $S = (S_t, t \geqslant 0)$ on some probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$. We will also require S_t to be $\mathcal{F}_t$-adapted. The cash at the bank account grows deterministically in accordance with the interest formula

$$A_t = A_0\, e^{rt}, \quad t \geqslant 0.$$

where $r > 0$ is the constant interest rate.

We are concerned with the *European call option*. One buys this option at time 0 to buy stock at the *expriration time* T at the *exrecise price* K. The value of the option at time T is the random variable

$$Z = \max\{S_T - k, 0\} = (S_T - k)^+.$$

Another key concept is the notion of *arbitrage*. Here is an example of the *arbitrage opportunity*. Suppose a stock sells in Frankfurt for €100 and in New York for \$100, and that the 1€= 1.34\$. Then one can buy for \$100 the stock in New York to sell immediately in Frankfurt for €100 having a \$34 profit. Thus, in this case, the disparity in pricing stocks in Germany and the USA has led to the availability of 'free money'.

Now suppose that a sum of money P is invested at a constant interest rate r. At time t it becomes $P\, e^{rt}$. Conversely, if we want to obtain a given sum of money Q at time t then we must invest $Q\, e^{rt}$ at time 0. The process of obtaining $Q\, e^{rt}$ from Q is called *discounting*. In particular, if S_t, $t \geqslant 0$ is is the stock price process, we define the *discounted process* $\tilde{S}_t = e^{-rt} S_t$, $t \geqslant 0$.

Definition 9.1. *Two probability measures P and Q are called equivalent if they have same null sets, i.e. for any set A with $P(A) = 0$, $Q(A) = 0$ and vice versa.*

In discrete time, we have the following result:

Theorem 9.2 (Fundamental theorem of asset pricing)**.** *If the market is free of arbitrage opportunities, then there exists a probability measure Q, which is equivalent to P, with respect to which the discounted process $\tilde{S}_t$ is a martingale.*

Under additional technical assumptions, the result holds in the continuous time setting.

9.1.1 Portfolios

The concept of a *portfolio.* An investor holds his investments as a combination of risky *stocks* and *cash* in the bank. Let α_t and β_t, $t \geqslant 0$, denote the amount of each of these, respectively, that the investor holds at time t. The pair of adapted processes

(α_t, β_t) is called a *portfolio* or *trading strategy*. The total value of all our investments at time t is

$$V_t = \alpha_t S_t + \beta_t A_t.$$

One of the aims of the Black–Scholes approach to option pricing is to be able to *hedge* the risk involved in selling options, i.e. to be able to construct a portfolio whose value V_T at the expiration time T equals to the value

$$Z = (S_T - k)^+ = \max\{(S_T - k), 0\}$$

of the option. A portfolio is said to be *replicating* if

$$V_T = Z.$$

A portfolio is said to be *self-financing* if any change in wealth V_t is due only to changes in the values of stocks and bank accounts and not to any injections of capital from outside. Assuming that the stock price process S_t is a semimartingale, we give the following precise definition.

Definition 9.3. *A portfolio (α_t, β_t) is called self-financing if*

$$V_t = \alpha_0 S_0 + \beta_0 A_0 + \int_0^t \alpha_s dS_s + \int_0^t \beta_s dA_s =$$

Taking into consideration the formula for the bank account growth, we obtain:

$$dV_t = \alpha_t dS_t + r\beta_t A_t dt.$$

A *contingent claim* with maturity date T, is a non-negative $\mathcal{F}_T$-measurable random variable. *European options* are examples of *contingent claims*. Another example of a contingent claim is the *American call option*, where stocks may be purchased at any time within the interval $[0, T]$, not only at the endpoint. A market is said to be *complete* if every contingent claim can be replicated by a self-financing portfolio.

9.2 The Black-Scholes model

The Black-Scholes Model is one of the most important concepts in modern financial theory. It was developed in 1973 by Fisher Black and Myron Scholes and is still widely used today. The stock price process S_t is assumed to satisfy the following SDE:

$$dS_t = \mu\, S_t\, dt + \sigma S_t dB_t, \tag{61}$$

where B_t is a standard Brownian motion, $\mu > 0$ and $\sigma > 0$ are constants where the latter is called *volatility* of the stock. The cash at the bank account is assumed to grow in accordance to

$$A_t = e^{rt}, \quad t \geqslant 0.$$

Theorem 9.4. *SDE (61) has a unique solution with the initial condition S_0. The solution is given by*

$$S_t = S_0 \exp\{\sigma B_t + (\mu t - \frac{1}{2}\sigma^2 t)\} \tag{62}$$

Proof. We have to apply Itô's formula to $u(t, B_t)$ where

$$u(t,x) = \exp\Big(\sigma\, x - \frac{\sigma^2\, t}{2} + \mu t\Big).$$

The uniqueness of solution follows from Theorem 8.3. □

We would like to prove the existence and uniqueness of a measure Q which is equivalent to P and possesses the property that the discounted process $\tilde{S}_t = A_t^{-1} S_t$ is a Q-martingale. We will call such a measure the *equivalent martingale measure.* Define the measure Q by its density with respect to P:

$$\Lambda(\omega) = \frac{dQ}{dP} = e^{b\, B_T - \frac{b^2}{2}\, T} \tag{63}$$

with $b = \frac{\mu - r}{\sigma}$.

Theorem 9.5 (Girsanov's theorem)**.** *Let $(\Omega, \mathcal{F}, P)$ be a probability space, and let $\mathcal{F}_t$ be the natural filtration of the Brownian motion B_t. Then, for each t, the measure Q, defined by (63) and restricted to $\mathcal{F}_t$, is equivalent to the measure P restricted to $\mathcal{F}_t$. Moreover, the process*

$$W_t = B_t + b\, t$$

is a Brownian motion under Q.

Proof. Without proof. □

Theorem 9.6. *The equivalent martingale measure Q is unique. The SDE for S_t under Q becomes*

$$dS_t = r\, S_t\, dt + \sigma S_t dW_t, \tag{64}$$

where W_t is Q-Brownian motion.

Proof. By Girsanov's theorem $W_t = B_t + \frac{\mu - r}{\sigma}$ is a Brownian motion under Q. We have:

$$\begin{aligned} d\tilde{S}_t = d(e^{-rt} S_t) = e^{-rt}(\mu\, S_t dt + \sigma S_t dB_t) - r\, e^{-rt} S_t = e^{-rt} S_t\big((\mu - r)dt + \sigma dB_t\big) \\ = \sigma \tilde{S}_t dW_t. \end{aligned}$$

The process $\tilde{S}_t$ satisfying this SDE has the form

$$\tilde{S}_t = e^{\sigma W_t - \frac{\sigma^2}{2}\, t} \tag{65}$$

and it is a Q-martingale. On the other hand,

$$\tilde{S}_t = S_0 \exp\{\sigma B_t + (\mu - r)\, t - \frac{\sigma^2}{2}\, t\}. \tag{66}$$

Therefore, the process $\tilde{S}_t$ is a P-martingale if and only if $\mu = r$ which implies that $P = Q$. Therefore, the measure Q is unique. Now, the SDE $d\tilde{S}_t = \sigma \tilde{S}_t dW_t$ becomes (64) if we substitute $\tilde{S}_t = e^{-rt} S_t$. The theorem is proved. □

9.2.1 The Black-Scholes portfolio

Now our goal is to construct the Black-Scholes portfolio. The Black-Scholes strategy is to construct a portfolio V_t which is both *self-financing* and *replicating*, and which effectively fixes the value of the option at each time t. Let E_Q be the expectation with respect to the measure Q. Define the Q-martingale Z_t, $0 \leqslant t \leqslant T$, by

$$Z_t = A_t^{-1}\mathbb{E}_Q[Z|\mathcal{F}_t]. \tag{67}$$

Then Z_t is an L_2-martingale, since, by Jensen's inequality, we have:

$$\mathbb{E}_Q\big[\mathbb{E}_Q[Z|\mathcal{F}_t]^2\big] \leqslant \mathbb{E}_Q\big[\mathbb{E}_Q[Z^2|\mathcal{F}_t]\big] = \mathbb{E}_Q[Z^2] < \infty.$$

We can apply the martingale representation theorem in the probability space $(\Omega, \mathcal{F}, Q)$ to conclude that there exists a square-integrable process δ_t, $0 \leqslant t \leqslant T$, such that

$$dZ_t = \delta_t\, dW_t = \gamma_t\, d\tilde{S}_t, \tag{68}$$

where $\gamma_t = \frac{\delta_t}{\sigma\tilde{S}_t}$, since $d\tilde{S}_t = \sigma\tilde{S}_t\, dW_t$. Define the portfolio (α_t, β_t) by the formulas:

$$\alpha_t = \gamma_t, \qquad \beta_t = Z_t - \gamma_t\tilde{S}_t, \tag{69}$$

for all $t \in [0, T]$. We call this the *Black-Scholes portfolio*. Its value is

$$V_t = \alpha_t S_t + \beta_t A_t = \gamma_t S_t + (Z_t - \gamma_t\tilde{S}_t)A_t$$

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

Theorem 9.7. *The Black-Scholes portfolio is self-financing and replicating.*

Proof. Since $\tilde{S}_t = A_t^{-1}S_t$, then

$$V_t = A_t\gamma_t\tilde{S}_t + (Z_t - \gamma_t\tilde{S}_t)A_t = Z_t A_t. \tag{70}$$

To see that this portfolio is replicating, we observe that

$$V_T = A_T Z_T = A_T A_T^{-1}\mathbb{E}_Q[Z|\mathcal{F}_T] = Z$$

since Z is $\mathcal{F}_T$-measurable. To see that the portfolio is self-financing, we apply formula (68) and Itô's formula:

$$dV_t = d(Z_t A_t) = Z_t dA_t + A_t dZ_t = A_t\gamma_t d\tilde{S}_t + Z_t dA_t.$$

By (69), $Z_t = \beta_t + \gamma_t\tilde{S}_t$. Therefore,

$$\begin{aligned} dV_t =& \gamma_t A_t d\tilde{S}_t + (\beta_t + \gamma_t\tilde{S}_t)dA_t \\ =& \beta_t dA_t + \gamma_t(A_t d\tilde{S}_t + \tilde{S}_t dA_t) = \beta_t dA_t + \gamma_t d(\tilde{S}_t A_t) \\ =& \beta_t dA_t + \alpha_t dS_t \end{aligned}$$

since $\alpha_t = \gamma_t$ by the definition. The theorem is proved. □

9.2.2 The Black–Scholes pricing formula

We can now derive the celebrated Black-Scholes pricing formula for a European option. Note that formula (70) and the definition of the martingale Z_t (see (67)) allows us to obtain the value of the portfolio at any time $t \in [0, T]$:

$$V_t = A_t Z_t = A_t A_T^{-1} \mathbb{E}_Q[Z|\mathcal{F}_t] = e^{-r(T-t)} \mathbb{E}_Q[Z|\mathcal{F}_t]. \tag{71}$$

In particular,

$$V_0 = e^{-rT} \mathbb{E}_Q[Z] = e^{-rT} \mathbb{E}_Q[(S_T - k)^+]. \tag{72}$$

Note that, by (65), at time T

$$\tilde{S}_T = S_0 \, e^{\sigma W_T - \frac{\sigma^2}{2} T}.$$

Hence,

$$S_T = A^{-1} \tilde{S}_T = S_0 \, e^{W_T + (r - \frac{\sigma^2}{2} T)}.$$

Since $W_T \sim N(0, T)$, then, by the scaling property of a Brownian motion,

$$\sigma W_T - \frac{\sigma^2 T}{2} \sim W_{\sigma^2 T} - \frac{\sigma^2 T}{2} \sim N(-\frac{\sigma^2 T}{2}, \sigma^2 T).$$

This and (72) imply that

$$V_0 = e^{-rT} \mathbb{E}_Q[(s e^{U + rt} - k)^+],$$

where $s = S_0$ and U is a random variable with the distribution $N(-\frac{\sigma^2 T}{2}, \sigma^2 T)$ Computing the expectation, we obtain:

$$\begin{aligned}
V_0 =& \frac{e^{-rT}}{\sigma\sqrt{2\pi T}} \int_{-\infty}^{\infty} (s \, e^{x+rT} - k)^+ \, e^{-\frac{x + \frac{\sigma^2 T}{2}}{2\sigma^2 T}} \, dx \\
=& \frac{1}{\sigma\sqrt{2\pi T}} \int_{s e^{x+rT} - k > 0} (s \, e^x - k \, e^{-rT}) e^{-\frac{x + \frac{\sigma^2 T}{2}}{2\sigma^2 T}} \, dx \\
=& \frac{1}{\sigma\sqrt{2\pi T}} \int_{-(\log(\frac{s}{k}) + rT)}^{\infty} (s \, e^x - k \, e^{-rT}) e^{-\frac{x + \frac{\sigma^2 T}{2}}{2\sigma^2 T}} \, dx & (73) \\
=& \frac{s}{\sigma\sqrt{2\pi T}} \int_{-(\log(\frac{s}{k}) + rT + \frac{\sigma^2 T}{2})}^{\infty} e^{-\frac{y^2}{2\sigma^2 T}} \, dy - \frac{k e^{-rT}}{\sigma\sqrt{2\pi T}} \int_{-(\log(\frac{s}{k}) + rT - \frac{\sigma^2 T}{2})}^{\infty} e^{-\frac{y^2}{2\sigma^2 T}} \, dy. & (74)
\end{aligned}$$

Passing from line (73) to line (74), in the first summand we rearranged the product of two exponential functions into one, and made the substitution $y = x - \frac{\sigma^2 T}{2}$, and in the second summand we made the substitution $y = x + \frac{\sigma^2 T}{2}$. Finally, taking into account that for any $a > 0$,

$$\frac{1}{\sigma\sqrt{2\pi T}} \int_{-a}^{\infty} e^{-\frac{y^2}{2\sigma^2 T}} \, dy.$$

where Φ is the distribution function of a standard normal random variable, we obtain the Black-Scholes pricing formula for European calls:

$$V_0 = s\Phi\Big(\frac{\log(\frac{s}{k}) + rT + \frac{\sigma^2 T}{2}}{\sigma\sqrt{T}}\Big) - k\, e^{-rT}\, \Phi\Big(\frac{\log(\frac{s}{k}) + rT - \frac{\sigma^2 T}{2}}{\sigma\sqrt{T}}\Big). \tag{75}$$

The Black-Scholes pricing formula allows one to make a profit by buying or selling options. Suppose that the option sells for a price $P > V_0$. Then one can sell options and invest the value V_0 in α_0 units of stock and in β_0 units of the bank account, where (α_t, β_t) is the Black-Scholes portfolio. Using the fact that the Black-Scholes portfolio is replicating, we know that at time T the value of the option will be V_T. Therefore, one makes a profit of $P - V_0$. Suppose now that $P < V_0$. Then one sells α_0 units of stock, borrows β_0 units of the bank account, and spends the amount P to buy options. At times T he gets the value of the option Z by means of buying stocks and selling them immediately. Since the portfolio is replicating, $Z = V_T$. He returns then the borrowed value β_T of units of the bank accounts and buys α_T units of stock. In this case, one makes a profit of $V_0 - P$.

9.2.3 The Black-Scholes PDE

Formula (71) implies that

$$V_t = e^{r(T-t)}\mathbb{E}_Q[(S_T k)^+|\mathcal{F}_t]$$

where $\mathcal{F}_t$ is the natural filtration of W_t. By (62) $\mathcal{F}_t$ coincides with the filtration $\mathcal{G}_t = \sigma(S_s, 0 \leqslant s \leqslant t)$. Remind that for a Markov process X_t with the transition function $P_{s,t}$ by the definition and Theorems 4.5 and 4.6 we have:

$$(P_{s,t}f)(X_s) = \mathbb{E}_Q[f(X_t)|\sigma(X_u, u \leqslant s)] \quad \text{and}$$
$$\mathbb{E}_x[f(X_t)] = (P_{s,t}f)(x) = \int_{\mathbb{R}^n} P_{s,t}(x, dy)\, f(y).$$

where E_x is the expectation with respect to the measure P_{δ_x} defined in Theorem 4.6. These formulas imply:

$$\mathbb{E}_Q[f(X_t)|\sigma(X_u, u \leqslant s)] = \mathbb{E}_{X_s}[f(X_t)].$$

Applying to our case, we obtain:

$$V_t = e^{-r(T-t)}\mathbb{E}_Q[(S_T - k)^+|\mathcal{G}_t] = e^{-r(T-t)}\mathbb{E}_{S_t}[(S_T - k)^+].$$

Define

$$C(x,t) = e^{-r(T-t)}\, \mathbb{E}_{S_t}[(S_T - k)^+]. \tag{76}$$

Note that $V_t = C(S_t, t)$. Therefore, we have:

$$dC(S_t, t) = \alpha_t dS_t + \beta_t dA_t = \alpha_t dS_t + r\beta_t A_t\, dt \tag{77}$$

where $A_t = e^{rt}$. On the other hand, assuming that $C(x,t)$ belongs to the class C_2, we can apply Itô's formula. Taking into account that, by (61), $d\langle S\rangle_t = \sigma^2 S_t^2$, we obtain:

$$dC(S_t,t) = \frac{\partial C(S_t,t)}{\partial t}\,dt + \frac{\partial C(S_t,t)}{\partial x}\,dS_t + \frac{1}{2}\,\frac{\partial^2 C(S_t,t)}{\partial x^2}\,\sigma^2\,S_t^2\,dt.$$

Comparing with (77), we obtain:

$$\Big(\alpha_t - \frac{\partial C(S_t,t)}{\partial x}\Big)dS_t = \Big(r\beta_t\,A_t - \frac{\partial C(S_t,t)}{\partial t} - \frac{1}{2}\,\frac{\partial^2 C(S_t,t)}{\partial x^2}\,\sigma^2 S_t^2\Big)\,dt.$$

Note that if $\alpha_t - \frac{\partial C(S_t,t)}{\partial x} \neq 0$, then the left-hand side has a positive quadratic variation while the right-hand side has zero quadratic variation. This implies that

$$\alpha_t = \frac{\partial C(S_t,t)}{\partial x}.$$

Consequently,

$$\beta_t\,A_t = r^{-1}\Big(\frac{\partial C(S_t,t)}{\partial t} + \frac{1}{2}\,\frac{\partial^2 C(S_t,t)}{\partial x^2}\,\sigma^2 S_t^2\Big).$$

Finally, we obtain:

$$C(S_t,t) = V_t = \alpha_t S_t + \beta_t A_t = \frac{\partial C(S_t,t)}{\partial t}\,S_t + r^{-1}\Big(\frac{\partial C(S_t,t)}{\partial t} + \frac{1}{2}\,\frac{\partial^2 C(S_t,t)}{\partial x^2}\,\sigma^2 S_t^2\Big).$$

In the above equation, we replace S_t with x to obtain the Black-Scholes PDE.

Theorem 9.8 (The Black-Sholes PDE)**.** *The function $C(x,t)$ satisfies the Black-Scholes PDE:*

$$-rC(x,t) + \frac{\partial C(x,t)}{\partial t} + rx\,\frac{\partial C(x,t)}{\partial x} + \frac{\sigma^2 x^2}{2}\,\frac{\partial^2 C(x,t)}{\partial x^2} = 0 \tag{78}$$

with the final condition $C(x,T) = (x-K)^+$. The function $C(x,t)$ solving the above SDE can be explicitely given by the formula:

$$C(x,t) = x\Phi\Big(\frac{\log(\frac{x}{k}) + (r+\frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{(T-t)}}\Big) - k\,e^{-r(T-t)}\,\Phi\Big(\frac{\log(\frac{x}{k}) + (r-\frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{(T-t)}}\Big), \tag{79}$$

where Φ is a function of standard normal distrbution.

Proof. Note that we can apply the argument that we used to obtain the Black-Scholes pricing formula (75) to $C(x,t)$ given by (76) with

$$S_T = x\,e^{\sigma(W_T - W_t) + (r-\frac{\sigma^2}{2})(T-t)}$$

since we replaced $S_t = S_0\,e^{\sigma W_t + (r-\frac{\sigma^2}{2})t}$ by x. We obtain then (79). ☐

9.3 Bonds and interest rates

To avoid arbitrage between bonds and savings account, a certain relation must hold between bonds and the spot rate:

$$P(t,T) = e^{-\int_t^T r(s)ds}.$$

Let $(\Omega, P, \mathcal{F})$ be a proability space with the filtration $\mathcal{F}_t$, $0 \leqslant t \leqslant T^*$, and that the process $t \mapsto P(t,T)$ is $\mathcal{F}_t$-adapted.

We make the assumption of the existence of the equivallent martingale measure, i.e. that there exists a probability measure Q which is equivalent to P and such that simultaneously for all $T < T^*$, the process $t \mapsto \frac{P(t,T)}{A_t}$, $0 \leqslant t \leqslant T$, is a martingale. The martingale property and the fact that $P(T,T) = 1$ imply that

$$\mathbb{E}_Q\Big(\frac{1}{A_T} \,|\, \mathcal{F}_t) = \mathbb{E}_Q\Big(\frac{P(T,T)}{A_T} \,|\, \mathcal{F}_t) = \frac{P(t,T)}{A_t}. \tag{80}$$

From this we obtain:

$$P(t,T) = \mathbb{E}_Q\Big(\frac{A_t}{A_T} \,|\, \mathcal{F}_t\Big) = \mathbb{E}_Q\Big(e^{-\int_t^T r_s ds} \,|\, \mathcal{F}_t\Big) \tag{81}$$

9.3.1 SDE for the bond price

Let $(\Omega, P, \mathcal{F})$ be a probability space with the filtration $\mathcal{F}_t$. Further let B_t be a Brownnian motion under P. We assume that the spot rate process r_t generates the filtration $\mathcal{F}_t$, and that for all $T < T^*$, the process $t \mapsto T$ is $\mathcal{F}_t$-adapted. By (80),

$$A_t^{-1} P(t,T) = \mathbb{E}_Q\Big(e^{-\int_0^T r_s ds} \,|\, \mathcal{F}_t\Big).$$

By the martingale representation theorem, there exists a process h_t so that

$$A_t^{-1} P(t,T) = \int_0^t h_t dW_t$$

where W_t is a Q-Brownian motion. Now let $\sigma(t,T) = h_t A_t P(t,T)^{-1}$. Then

$$d(P(t,T)A_t^{-1}) = \sigma(t,T) P(t,T) A_t^{-1} dW_t.$$

Note that since $A_t = e^{\int_0^t r_s ds}$, then $dA_t = A_t r_t dt$, and, therefore $d(A_t^{-1}) = -A_t^{-1} r_t dt$. This implies that

$$A_t^{-1} dP(t,T) = P(t,T) A_t^{-1} r_t dt + \sigma(t,T) P(t,T) A_t^{-1} dW_t.$$

Canceling the factor A_t^{-1} in the both sides of the above SDE, we obtain the SDE for the bond price $P(t,T)$ with respect to the equivalent martingale measure Q:

$$dP(t,T) = P(t,T) r_t dt + \sigma(t,T) P(t,T) dW_t. \tag{82}$$

To find the SDE for $P(t,T)$ with respect to the original measure P, we will need the following version of Girsanov's theorem:

Theorem 9.9. *Let P be the Wiener measure, B_t be a P-Brownian motion. Further let Q be equivalent to P. Then there exists a predictable process q_t, such that*

$$\frac{dP}{dQ} = e^{\int_0^T q_t dB_t - \frac{1}{2}\int_0^T q_t^2 dt}.$$

Moreover,

$$B_t = W_t + \int_0^t q_s \, ds \tag{83}$$

where W_t is a Q-Brownian motion.

Girsanov's theorem implies, therefore, the existence of a predictable process q_s such that relation (83) holds. After substituting $dW_t = dB_t = q_t \, dt$ into (82) we obtain the SDE for the bond price $P(t,T)$ under P:

$$dP(t,T) = P(t,T)(r_t - \sigma(t,T)q_t)\, dt + \sigma(t,T)P(t,T)dB_t.$$

The process $-q_t$ is known as the *market price of risk*.

9.3.2 Model for a spot rate

Let $(\Omega, P, \mathcal{F})$ be a probability space. The spot rate r_t is assumed to satisfy the SDE:

$$dr_t = m(r_t)dt + \sigma(r_t)\, dB_t \tag{84}$$

where B_t is a P-Brownian motion, m and σ are functions of a real variable. The bond price satisfies (81). Remark that SDE (84) is under measure P, but the expectation in (81) is under measure Q. Let us find the SDE for r_t under Q. By Girsanov's theorem, we have:

$$dW_t = dB_t - q_t dt.$$

From this, we obtain the SDE for r_t under Q.

$$dr_t = (m(r_t) + \sigma(r_t)q_t)\, dt + \sigma(r_t)\, dW_t. \tag{85}$$

Using the same argument that we used to derive the Black-Scholes PDE and taking into account that the filtration $\mathcal{F}_t$ is generated by r_t, we obtain that

$$\mathbb{E}_Q\Big(e^{-\int_t^T r_s\, ds} \,|\, \mathcal{F}_t\Big) = \mathbb{E}_{r_t}\big(e^{-\int_t^T r_s\, ds}\big).$$

Define

$$C(x,t) = \mathbb{E}_x\big(e^{-\int_t^T r_s\, ds}\big).$$

This function satisfies a certain PDE which can be obtained by applying Itô's formula to $C(r_t,t)$, substituting the right-hand side of (85) for dr_t and obtaining a coefficient at dt. On the other hand we know that $C(r_t,t) = P(t,T)$, and therefore, by (82), the dt-part of $dC(r_t,t) = dP(t,T)$ equals to $P(t,T)r_t dt = C(r_t,t)r_t dt$. Finally we substitute $r_t = x$ to obtain:

$$\frac{1}{2}\sigma^2(x)\frac{\partial^2 C}{\partial x^2}(x,t) + (m(x) + \sigma(x)q_t)\frac{\partial C}{\partial x}(x,t) + \frac{\partial C}{\partial t}(x,t) - xC(x,t) = 0.$$

Since $S(r_T,T) = P(T,T) = 1$, the boundary condition will be $C(x,T) = 1$.

9.3.3 Merton's model and Vasicek's model

Among well known models for the spot rate are

- Merton's model

$$dr_t = \mu\, dt + \sigma\, dB_t.$$

- Vasicek's model

$$dr_t = b(a - r_t)\, dt + \sigma\, dB_t.$$

The solution for Merton's model is

$$r_t = r_0 + \mu\, t + \sigma\, B_t.$$

The solution for Vasicek's model is

$$r_t = a - e^{-bt}(a - r_0) + \sigma \int_0^t e^{-b(t-s)} dB_s.$$

9.3.4 Heath-Jarrow-Morton (HJM) model

The model is based on modelling of *forward rates*. By definition, the forward rate $f(t,T)$, $t \leqslant T \leqslant T^*$, is defined by the relation:

$$P(t,T) = e^{-\int_t^T f(t,s)ds}.$$

This implies

$$f(t,T) = -\frac{\partial \log P(t,T)}{\partial T}.$$

One can say that the forward rate $f(t,T)$ is the rate at time T as seen from time t. The spot rate and the forward rate are related by $r_t = f(t,t)$. Consequently, the savings account A_t grows as

$$A_t = e^{\int_0^t f(s,s)ds}.$$

According to the HJM model, for a fixed time T, $f(t,T)$ is the solution of the SDE:

$$df(t,T) = \mu(t,T)dt + \xi(t,T)dB_t,$$

where B_t is a Brownian motion, $\mu(t,T)$ and $\xi(t,T)$ are adapted and continuous. For this assumption to be compatible with the assumption of the existence of martingale measures we need the following relation to hold:

$$\frac{dP(t,T)}{P(t,T)} = (r_t - \alpha(t,T)\theta_t)dt + \alpha(t,T)dB_t.$$

where $\alpha(t,T) = -\int_t^T \xi(t,s)ds$ and $\mu(t,T) = \xi(t,T)\big(\int_t^T \xi(t,s)ds - \theta_t\big)$. The latter equation is known as the no-arbitrage condition of the HJM model.