Stochastic Differential Equations

Z. Qian
and Mathematical Institute, University of Oxford

January 16, 2011

Contents

Introduction

A stochastic differential equation is simply a differential equation perturbed by random noise (its intensity $\sigma(t, X_t)$ depends on time t and the position X_t), and therefore has the following form

$$\frac{dX_t}{dt} = A(t, X_t) + \sigma(t, X_t)\dot{W}_t \ .$$

The central limit theorem in probability theory suggests that $\dot{W}_t$ should have normal distribution, and for the sack of simplicity $(\dot{W}_t)_{t\geq 0}$ should be independent. Such random noise can be ideally modelled by Brownian motion $(W_t)_{t\geq 0}$, a mathematical model describing random movements of pollen particles in a liquid, observed by R. Brown in 1827. The mathematical model for Brownian motion and the description of its distribution were derived by Albert Einstein in a short paper "On the motion of small particles suspended in liquids at rest required by the molecular-kinetic theory of heat" published in 1905, *Annalen der Physik* 17, 549-560. About the same time, in 1900, L. Bachelier submitted his Ph. D. thesis in which he used Brownian motion to model stock markets. His results were published in a paper titled "Théorie de la spéculation" in *Ann. Sci. Ecole Norm. sup.*, 17 (1900), 21-86, which is the first paper devoted to applications of Brownian motion to finance.

On the other hand, the first mathematical construction of Brownian motion was achieved in 1923, N. Wiener published his article "Differential space", *J. Math. Phys.* 2, 132-174. Fruitful results and many unusual features of Brownian motions were revealed mainly by Paul Lévy in 30's - 40's. Among of them, P. Lévy showed that almost surely $t \to W_t$ is nonwhere differentiable, and therefore the time-derivative of Brownian motion, $\dot{W}_t$, does not exist. It is thus necessary to rewrite the previous differential equation in differential form

$$\mathrm{d}X_t = A(t, X_t)\mathrm{d}t + \sigma(t, X_t)\mathrm{d}W_t$$

which in turn has to be interpreted as an integral equation

$$X_t - X_0 = \int_0^t A(s, X_s)\mathrm{d}s + \int_0^t \sigma(s, X_s)\mathrm{d}W_s$$

It requires thus to define integral like

$$\int \sigma(t, X_t)\mathrm{d}W_t$$

which does not exist in the ordinary sense. It was Itô in 1940's who first established an integration theory for Brownian motion, and therefore the theory of stochastic differential equations. Among of the manifold applications and connections with PDE, one of the most remarkable applications of Itô's theory is in the theory of finance. Although Itô's theory never catch the attention of the Fields medals committee, while it was brought worldwide recognition by awarding H. Markowitz, W. Sharpe and M. Miller the 1990 Nobel Prize, and F. Black and M. Scholes the 1997 Nobel Prize, both in Economics.

This course covers the core part of Itô's calculus: it provides the necessary background in stochastic analysis for those who are interested in stochastic modelings and their applications including the theory of finance, stochastic control and filtering etc. Students who are majoring in (pure and applied) analysis, differential geometry, functional analysis, harmonic analysis, mathematical physics and PDEs will find this course relevant to their interests.

References:

Below is the list of a collection of text books and monographs on stochastic differential equations and related topics. Items 3 and 4 are recommended readings for this course.

1. N. Ikeda and S. Watanabe: Stochastic Differential Equations and Diffusion Processes. North-Holland (1981).

2. I. Karatzas and S. E. Shreve: Brownian Motion and Stochastic Calculus. Graduate Texts in Mathematics, Springer-Verlag (1988).

3. F. C. Klebaner: Introduction to Stochastic Calculus with Applications. Imperial College Press (1998).

4. B. Oksendal: Stochastic Differential Equations. 6th Edition, Universitetext. Springer (2003).

5. D. Revuz and M. Yor: Continuous Martingales and Brownian Motion. Springer-Verlag (1991).

6. L. C. G. Rogers and D. Williams: Diffusions, Markov Processes and Martingales, Volume 2 Itô Calculus. Cambridge Mathematical Library. Cambridge University Press 2000.

7. S. E. Shreve: Stochastic Calculus Finance II Continuous-Time Models. Springer Finance Textbook. Springer (2004).

Chapter 1

Preliminaries

1.1 Toolbox

1.1.1 The monotone class theorem

A collection $\mathcal{L}$ of subsets of Ω is called a π-system, if it is closed under finite intersections. By the monotone class theorem, we mean the following lemma or a version of its variations.

Lemma 1.1.1 *Let $\mathcal{L}$ be a π-system, and $\mathcal{F} = \sigma\{\mathcal{L}\}$ the smallest σ-algebra. Let $\mathcal{H}$ be a family of real-valued functions on Ω satisfying the following two conditions:*

1. *$1 \in \mathcal{H}$ and $1_A \in \mathcal{H}$ for every $A \in \mathcal{L}$.*

2. *If $f_n \in \mathcal{H}$, each f_n is non-negative, $f_n \uparrow$ (in n), and $\sup_n f_n < +\infty$, then $\sup_n f_n \in \mathcal{H}$.*

Then $\mathcal{H}$ contains all bounded, real-valued and $\mathcal{F}$-measurable functions on Ω.

1.2 Probability spaces

A probability space is a triple $(\Omega, \mathcal{F}, P)$ consisting of a sample space Ω of basic events (also called sample points), a σ-algebra $\mathcal{F}$ of events, and a probability measure P. The probability measure P is a function on $\mathcal{F}$ taking values in $[0, 1]$, which satisfies the following conditions:

1. $P(\Omega) = 1$, $P(\emptyset) = 0$ (where $\emptyset$ is an empty set representing impossible event), and $P(A) \geq 0$ for any event $A \in \mathcal{F}$.

2. Countably additive: If $\{A_i\}_{i=1,\cdots}$ is a countable family of mutually disjoint events, i.e. $A_i \in \mathcal{F}$ and $A_i \cap A_j = \emptyset$, then

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

By a random *variable* on $(\Omega, \mathcal{F}, P)$ valued in $\mathbb{R}^d$ (or an $\mathbb{R}^d$-valued random variable), we mean a measurable (vector-valued) *function* on $(\Omega, \mathcal{F})$. Recall that mapping $X : \Omega \to \mathbb{R}^d$ is measurable, if for every Borel subset B of $\mathbb{R}^d$, the pre-image of B under the map X

$$X^{-1}(B) = \{\omega : X(\omega) \in B\}$$

belongs to $\mathcal{F}$. Loosely speaking, a *random variable* is such a *function* X on Ω so that we may be able to talk about, for example, what is the probability of the event that X lies in a ball B: $\{\omega \in \Omega : X(\omega) \in B\}$.

If

$$E|X| \equiv \int_{\Omega} |X(\omega)| P(d\omega) < \infty$$

then we say X is *integrable*, denoted by $X \in L^1(\Omega, \mathcal{F}, P)$. In this case, the expectation of X, denoted by $E(X)$, is the integral of X against the probability measure P:

$$E(X) \equiv \int_{\Omega} X(\omega) P(d\omega) \ .$$

In general, we say a random variable X is in the L^p-space, write as $X \in L^p(\Omega, \mathcal{F}, P)$ for $p \geq 0$, if

$$\int_{\Omega} |X(\omega)|^p P(d\omega) < \infty \ .$$

In this case, we also say X is p-th integrable. For $p \geq 1$, the space $L^p(\Omega, \mathcal{F}, P)$ of all p-th integrable random variables X is a Banach space under the usual algebraic operations for functions and the L^p-norm

$$||X||_{L^p} \equiv \left(\int_{\Omega} |X(\omega)|^p P(d\omega)\right)^{1/p} \ .$$

Remark 1.2.1 *If $p \geq q$, then $L^p(\Omega, \mathcal{F}, P) \subset L^q(\Omega, \mathcal{F}, P)$ and $||X||_q \leq ||X||_p$. Therefore $p \to ||X||_p$ is increasing in $p \in (0, \infty]$. Indeed, by a simple use of Hölder's inequality we have*

$$\begin{aligned} ||X||_q^q &= \int_\Omega |X(\omega)|^q P(d\omega) \\ &\leq \left(\int_\Omega |X(\omega)|^{q\frac{p}{q}} P(d\omega) \right)^{q/p} \\ &= \left(\int_\Omega |X(\omega)|^p P(d\omega) \right)^{q/p} \end{aligned}$$

Stochastic processes are mathematical models which are used to describe random phenomena evolving in time. We thus need to have a set $\mathbf{T}$ of the time-parameter. In these lectures, $\mathbf{T}$ is either the set of non-negative integers $\mathbb{Z}^+$ or $[0, +\infty)$. $\mathbf{T}$ is thus an ordered set endowed with the natural topology.

Definition 1.2.2 A stochastic process *is a parameterized family $X = (X_t)_{t \in \mathbf{T}}$ of random variables valued in a topological space S. In this book, unless otherwise specified, S will be the real line $\mathbb{R}$, or the Euclidean space $\mathbb{R}^d$ of dimension d.*

A stochastic process $X = (X_t)_{t \in \mathbf{T}}$ may be considered as a function from $\mathbf{T} \times \Omega \to \mathbb{R}^d$, which is the reason why a stochastic process is also called a random function.

For each sample point $\omega \in \Omega$, function $t \to X_t(\omega)$ from $\mathbf{T}$ to S is called a *sample path* (or a trajectory, or a sample function). Naturally, a stochastic process $X = (X_t)_{t \in \mathbf{T}}$ is continuous (resp. right-continuous, right-continuous with left-limits) if sample paths $t \to X_t(\omega)$ are continuous (resp. right-continuous, right-continuous with left-limits) for almost all $\omega \in \Omega$.

Remark 1.2.3 *A function $f : (a, b) \to \mathbb{R}^d$ is right-continuous at $t_0 \in (a, b)$ if its right-limit at t_0 exists and equals $f(t_0)$. Similarly, f is right-continuous with left-limit at t_0, if f is right-continuous at t_0 and its left-limit at t_0 exists. For example, any monotone function on an interval has right- and left-limits.*

Example 1.2.4 *(Poisson process) Let (ξ_n) be a sequence of independent identically distributed (i.i.d.) random variables with Poisson distribution of intensity $\lambda > 0$. Let*

$$T_0 = 0; \quad T_n = \sum_{j=1}^{n} \xi_j$$

and, for every $t \geq 0$ define

$$X_t = n \quad \text{if} \quad T_n \leq t < T_{n+1} \ .$$

Then for every sample point ω, $t \to X_t(\omega)$ is a step function, constant on each interval (T_n, T_{n+1}), with jump 1 at (random time) T_n, and is right-continuous with left limit $n-1$ at T_n.

If $X = (X_t)_{t \geq 0}$ is a stochastic process taking values in $\mathbb{R}^d$, and $0 \leq t_1 < t_2 < \cdots < t_n$, the joint distribution of random variables $(X_{t_1}, \cdots, X_{t_n})$ given as

$$\mu_{t_1, t_2, \cdots, t_n}(\mathrm{d}x_1, \cdots, \mathrm{d}x_n) = P\left(X_{t_1} \in \mathrm{d}x_1, \cdots, X_{t_n} \in \mathrm{d}x_n\right) \ ,$$

a probability measure on $\mathbb{R}^d \times \cdots \times \mathbb{R}^d$, is called a finite-dimensional distribution of $X = (X_t)_{t \geq 0}$. If $d = 1$ and $(X_t)_{t \geq 0}$ is a real stochastic process, the distribution $\mu_{t_1, t_2, \cdots, t_n}$ is determined via its distribution function

$$F_{t_1, t_2, \cdots, t_n}(x_1, \cdots, x_n) = P\left(X_{t_1} \leq x_1, \cdots, X_{t_n} \leq x_n\right) \ .$$

We need to overcome some technical difficulties when we deal with stochastic processes in continuous-time. For example, a subset of Ω like

$$\{\omega \in \Omega : X_t(\omega) \in B \text{ for all } t \in [0, 1]\}$$

may be not measurable, i.e. not an event, so that

$$P\left(\omega \in \Omega : X_t(\omega) \in B \text{ for all } t \in [0, 1]\right)$$

may not make sense, unless additional conditions on $(X_t)_{t \geq 0}$ are imposed. Similarly, a function like $\sup_{t \in K} X_t$ may be not a random variable, which will be very inconvenient.

Exercise 1.2.5 *Let $(X_t)_{t \geq 0}$ be a stochastic process in $\mathbb{R}^d$ on $(\Omega, \mathcal{F}, \mathbb{P})$, and let B be a Borel measurable subset. If F is a finite or countable subset of $[0, +\infty)$, then*

$$\{\omega : X_t(\omega) \in B \text{ for any } t \in F\}$$

and

$$\sup_{t \in F} |X_t|$$

are measurable.

To avoid such technical difficulties, a common condition, which is good enough to include a large class of interesting stochastic processes, is that a process X involved is right-continuous almost surely, and the probability space $(\Omega, \mathcal{F}, P)$ is complete in the sense that any trivial subsets of probability null are events.

The main task of stochastic analysis is to study the probability (or distribution) properties of random functions determined by their families of finite-dimensional distributions.

Definition 1.2.6 *Two stochastic processes $X = (X_t)_{t\geq 0}$ and $Y = (Y_t)_{t\geq 0}$ are equivalent (distinguishable) if for every $t \geq 0$ we have*

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

In this case, $(Y_t)_{t\geq 0}$ is a version of $(X_t)_{t\geq 0}$.

By definition, the family of finite-dimensional distributions of a stochastic process $X = (X_t)_{t\geq 0}$ is unique up to equivalence of processes.

On the other hand, in many practical situations, we are given a collection of *compatible* finite dimensional distributions $\mathcal{D} = \{\mu_{t_1,\cdots,t_n}$, for $t_1 < \cdots < t_n$, $t_j \in \mathbf{T}\}$, we would like to construct a stochastic model $(X_t)_{t\in\mathbf{T}}$ on some probability space $(\Omega, \mathcal{F}, P)$ so that the family of finite dimensional distributions determined by $(X_t)_{t\in\mathbf{T}}$ coincides with the family $\mathcal{D}$ of distributions. In this case, $(X_t)_{t\in\mathbf{T}}$ is called a *realization* of $\mathcal{D}$.

1.3 Conditional expectations

The main concepts in the probability theory, including independence, martingale property and Markov property, are stated in terms of conditional expectations (and conditional probability). We follow the formulation by J. L. Doob.

Let X be an integrable or non-negative random variable on a probability space $(\Omega, \mathcal{F}, P)$, and let $\mathcal{G}$ be a sub σ-algebra of $\mathcal{F}$. The *conditional expectation $E(X|\mathcal{G})$* of X given $\mathcal{G}$ is a random variable (unique up to almost surely) which satisfies the following two conditions.

1. $E(X|\mathcal{G})$ is measurable with respect to $\mathcal{G}$; and

2. For any $A \in \mathcal{G}$ we have

$$E\{E(X|\mathcal{G})1_A\} = E\{X1_A\} \ .$$

The conditional expectation $E(X|\mathcal{G})$ is the best prediction of the random variable X based on available information $\mathcal{G}$. By a monotone class argument, it follows that

$$E\left\{E(X|\mathcal{G})Y\right\} = E\left(XY\right)$$

as long as both sides make sense.

For simplicity, $E(X|Y)$ stands for $E(X|\sigma(Y))$, where $\sigma(Y)$ is the smallest σ-algebra for which Y is measurable. It can be shown that $E(X|Y)$ is a measurable function of Y, that is, there is a function F such that

$$E(X|Y) = F(Y) \ .$$

By definition, if Y is $\mathcal{G}$-measurable, then

$$E(YX|\mathcal{G}) = YE(X|\mathcal{G}) \ .$$

If X and $\mathcal{G}$ are independent, then

$$E(X|\mathcal{G}) = E\left(X\right) \ .$$

Indeed X is independent of σ-algebra $\mathcal{G}$ if and only if for any bounded Borel measurable function f

$$E(f(X)|\mathcal{G}) = E\left(f(X)\right) \ .$$

1.4 The uniform integrability

The *uniform integrability* for a family of integrable random variables has been formulated to handle the convergence of random variables in $L^1(\Omega, \mathcal{F}, P)$. In spirit, it is very close to uniform convergence, uniform continuity that you have learned in the analysis course.

If ξ is integrable, i.e. $\xi \in L^1(\Omega, \mathcal{F}, P)$, then

$$\lim_{N\to\infty} \int_{\{|\xi|\geq N\}} |\xi| \mathrm{d}P = 0 \ .$$

Definition 1.4.1 *Let $\mathcal{A}$ be a family of integrable random variables on $(\Omega, \mathcal{F}, P)$. $\mathcal{A}$ is uniformly integrable if*

$$\lim_{N\to\infty} \sup_{\xi\in\mathcal{A}} \int_{\{|\xi|\geq N\}} |\xi| dP = 0 \ .$$

That is, $E\left\{1_{\{|\xi|\geq N\}}|\xi|\right\}$ tends to zero uniformly on $\mathcal{A}$ as $N \to \infty$.

In terms of ε-δ language, $\mathcal{A}$ is uniformly integrable, if for any $\varepsilon > 0$ there is a $N > 0$ depending only on ε such that

$$\int_{\{|\xi|\geq N\}} |\xi| \mathrm{d}P < \varepsilon$$

for all $\xi \in \mathcal{A}$.

According to the definition, we have the followings.

1. Any finite family of integrable random variables is uniformly integrable.

2. Let $\mathcal{A} \subset L^1(\Omega, \mathcal{F}, P)$ be a family of integrable random variables. If there is an integrable random variable η such that $|\xi| \leq \eta$ for every $\xi \in \mathcal{A}$, then $\mathcal{A}$ is uniformly integrable. In fact

$$\sup_{\xi\in\mathcal{A}} \int_{\{|\xi|\geq N\}} |\xi| \mathrm{d}P \leq \int_{\{\eta\geq N\}} \eta \mathrm{d}P \to 0 \quad \text{as } N \to \infty \ .$$

3. If $\mathcal{A} \subset L^p(\Omega, \mathcal{F}, P)$ for some $p > 1$ and

$$\sup_{\xi\in\mathcal{A}} E|\xi|^p < \infty \ ,$$

then $\mathcal{A}$ is uniformly integrable. That is, a bounded subset of $L^p(\Omega, \mathcal{F}, P)$ for $p > 1$ is uniformly integrable. Indeed

$$\begin{aligned} \sup_{\xi\in\mathcal{A}} \int_{\{|\xi|\geq N\}} |\xi| \mathrm{d}P &\leq \sup_{\xi\in\mathcal{A}} \int_{\{|\xi|\geq N\}} \frac{1}{N^{p-1}} |\xi|^p \mathrm{d}P \\ &\leq \frac{1}{N^{p-1}} \sup_{\xi\in\mathcal{A}} E|\xi|^p \to 0 \end{aligned}$$

as $N \to \infty$.

4. If $\xi \in L^1(\Omega, \mathcal{F}, P)$, and $\{\mathcal{G}_\alpha\}_{\alpha\in A}$ is a collection of sub σ-algebras of $\mathcal{F}$, then the family

$$\mathcal{A} = \{E\left(\xi|\mathcal{G}_\alpha\right) : \alpha \in A\}$$

is uniformly integrable. Let $\xi_\alpha = E\left(\xi|\mathcal{G}_\alpha\right)$. Since $\{\xi_\alpha \geq N\}$ and

$\{\xi_\alpha \le -N\}$ are $\mathcal{G}_\alpha$-measurable,

$$\begin{aligned}\int_{\{|\xi_\alpha|\ge N\}} |\xi_\alpha| \mathrm{d}P &= \int_{\{\xi_\alpha \ge N\}} \xi_\alpha \mathrm{d}P - \int_{\{\xi_\alpha \le -N\}} \xi_\alpha \mathrm{d}P \\ &= \int_{\{\xi_\alpha \ge N\}} \xi \mathrm{d}P - \int_{\{\xi_\alpha \le -N\}} \xi \mathrm{d}P \\ &= \int_{\{|\xi_\alpha|\ge N\}} |\xi| \mathrm{d}P \\ &\le \int_{\{|\xi|\ge N\}} |\xi| \mathrm{d}P\end{aligned}$$

which proves the claim.

Theorem 1.4.2 *Let $\mathcal{A} \subset L^1(\Omega, \mathcal{F}, P)$. Then $\mathcal{A}$ is uniformly integrable if and only if*

1. *$\mathcal{A}$ is a bounded subset of $L^1(\Omega, \mathcal{F}, P)$, that is, $\sup_{\xi\in\mathcal{A}} E|\xi| < \infty$.*
2. *For any $\varepsilon > 0$ there is a $\delta > 0$ such that*

$$\int_A |\xi| dP \le \varepsilon$$

whenever $A \in \mathcal{F}$ and $P(A) \le \delta$.

Proof. Necessity. For any $A \in \mathcal{F}$ and $N > 0$

$$\begin{aligned}\int_A |\xi| \mathrm{d}P &= \int_{A\cap\{|\xi|<N\}} |\xi| \mathrm{d}P + \int_{A\cap\{|\xi|\ge N\}} |\xi| \mathrm{d}P \\ &\le NP(A) + \int_{\{|\xi|\ge N\}} |\xi| \mathrm{d}P \ .\end{aligned}$$

Given $\varepsilon > 0$, choose $N > 0$ such that

$$\sup_{\xi\in\mathcal{A}} \int_{\{|\xi|\ge N\}} |\xi| \mathrm{d}P \le \frac{\varepsilon}{2} \ .$$

Then

$$\sup_{\xi\in\mathcal{A}} \int_A |\xi| \mathrm{d}P \le NP(A) + \frac{\varepsilon}{2}$$

for any $A \in \mathcal{F}$. In particular

$$\sup_{\xi\in\mathcal{A}} \int_\Omega |\xi| \mathrm{d}P \le N + \frac{\varepsilon}{2} \ ,$$

and by setting $\delta = \varepsilon/(2N)$ we also have

$$\sup_{\xi \in \mathcal{A}} \int_A |\xi| \mathrm{d}P \leq \varepsilon$$

as long as $P(A) \leq \delta$.

Sufficiency. Let $\beta = \sup_{\xi \in \mathcal{A}} E|\xi|$. By the Markov inequality

$$P(|\xi| \geq N) \leq \frac{\beta}{N}$$

for any $N > 0$. For any $\varepsilon > 0$, there is a $\delta > 0$ such that the inequality in 2 holds. Choose $N = \beta/\delta$. Then $P(|\xi| \geq N) \leq \delta$ so that

$$\int_{\{|\xi| \geq N\}} |\xi| \mathrm{d}P \leq \varepsilon$$

for any $\xi \in \mathcal{A}$. ∎

Corollary 1.4.3 *Let $\mathcal{A} \subset L^1(\Omega, \mathcal{F}, P)$ and $\eta \in L^1(\Omega, \mathcal{F}, P)$ such that for any $D \in \mathcal{F}$*

$$E\left(1_D |\xi|\right) \leq E\left(1_D |\eta|\right) \ ; \qquad \forall \xi \in \mathcal{A} \ .$$

Then $\mathcal{A}$ is uniformly integrable.

The following theorem demonstrates the importance of uniform integrability.

Theorem 1.4.4 *Let $\{X_n\}_{n \in \mathbb{Z}^+}$ be a sequence of integrable random variables on $(\Omega, \mathcal{F},\ P)$. Then $X_n \to X$ in $L^1(\Omega, \mathcal{F}, P)$ for some random variable as $n \to \infty$:*

$$\int_\Omega |X_n - X| dP \to 0 \quad as \quad n \to \infty \ ,$$

if and only if $\{X_n\}_{n \in \mathbb{Z}^+}$ is uniformly integrable and $X_n \to X$ in probability as $n \to \infty$.

Proof. Necessity. For any $\varepsilon > 0$ there is a natural number m such that

$$\int_\Omega |X_n - X| \mathrm{d}P \leq \frac{\varepsilon}{2} \quad \text{for all} \quad n > m \ .$$

Therefore for every measurable subset A

$$\int_A |X_n| \mathrm{d}P \leq \int_A |X| \mathrm{d}P + \int_\Omega |X_n - X| \mathrm{d}P$$

so that
$$\sup_n \int_A |X_n| \mathrm{d}P \le \int_A |X| \mathrm{d}P + \sup_{k\le m} \int_A |X_k| \mathrm{d}P + \frac{\varepsilon}{2} \ .$$
In particular
$$\sup_n E|X_n| \le E|X| + \sup_{k\le m} E|X_k| + \frac{\varepsilon}{2}$$
i.e. $\{X_n : n \ge 1\}$ is bounded in $L^1(\Omega, \mathcal{F}, P)$. Moreover, since $X, X_1, \cdots, X_m$ belong to L^1, so that there is $\delta > 0$ such that, if $P(A) \le \delta$, then
$$\int_A |X| \mathrm{d}P + \sum_{k=1}^m \int_A |X_k| \mathrm{d}P \le \frac{\varepsilon}{2}$$
and therefore
$$\sup_n \int_A |X_n| \mathrm{d}P \le \varepsilon$$
as long as $P(A) \le \delta$.

Sufficiency. By Fatou's lemma
$$\int_\Omega |X| \mathrm{d}P \le \sup_n \int_\Omega |X_n| \mathrm{d}P < +\infty$$
so that $X \in L^1(\Omega, \mathcal{F}, P)$. Therefore $\{X_n - X : n \ge 1\}$ is uniformly integrable, thus, by Theorem 1.4.2, for any $\varepsilon > 0$ there is $\delta > 0$ such that
$$\int_A |X_n - X| \mathrm{d}P < \varepsilon$$
for any $A \in \mathcal{F}$ such that $P(A) \le \delta$. Since $X_n \to X$ in probability, there is an $N > 0$ such that
$$P\left(|X_n - X| \ge \varepsilon\right) \le \delta \qquad \forall n \ge N \ .$$
Therefore for $n \ge N$ we have
$$\begin{aligned} \int_\Omega |X_n - X| \mathrm{d}P &\le \int_{\{|X_n - X| \ge \varepsilon\}} |X_n - X| \mathrm{d}P + \varepsilon P\left(X_n - X| < \varepsilon\right) \\ &\le \varepsilon + \varepsilon P\left(X_n - X| < \varepsilon\right) \\ &\le 2\varepsilon \ . \end{aligned}$$
Hence
$$\lim_{n\to\infty} E|X_n - X| = 0 \ .$$
■

Chapter 2

Elements in martingale theory

In this chapter we collect fundamental results about martingales, including Doob's martingale inequalities and the convergence theorem for martingales.

2.1 Martingales in discrete-time

In the probability theory, we study properties of random variables determined by their distributions. Let $(\Omega, \mathcal{F}, P)$ be a probability space, and $\mathbb{Z}_+$ denote the set of non-negative integers. An increasing family $\{\mathcal{F}_n\}_{n\in\mathbb{Z}_+}$ of sub σ-algebras of $\mathcal{F}$ is called a filtration. A probability space $(\Omega, \mathcal{F}, P)$ together with a filtration $\{\mathcal{F}_n\}_{n\in\mathbb{Z}_+}$ is called a *filtered probability space*, denoted by $(\Omega, \mathcal{F}_n, \mathcal{F}, P)$. Given a sequence $X \equiv \{X_n\}_{n\in\mathbb{Z}^+}$ of random variables on $(\Omega, \mathcal{F}, P)$, the filtration generated by the sequence $\{X_n\}$ is defined to be $\mathcal{F}_n^X = \sigma\{X_m : m \leq n\}$ which is the smallest σ-algebra with respect to which $X_0, \cdots, X_n$ are measurable. If $\{X_n\}$ represents the process of a random phenomenon evolving in discrete-time, then $\mathcal{F}_n^X$ is the information up to time n.

Definition 2.1.1 *A sequence $\{X_n : n \in \mathbb{Z}_+\}$ of random variables on $(\Omega, \mathcal{F}, P)$ is adapted to $\{\mathcal{F}_n\}$ if for every $n \in \mathbb{Z}_+$, X_n is $\mathcal{F}_n$-measurable. In this case we say $\{X_n\}$ is an adapted sequence, or adapted process with respect to $\{\mathcal{F}_n\}$.*

If X_n is $\mathcal{F}_{n-1}$-measurable for any $n \in \mathbb{N}$ and $X_0 \in \mathcal{F}_0$, then we say $\{X_n\}$ is predictable.

Definition 2.1.2 *Let $\{\mathcal{F}_n : n \in \mathbb{Z}_+\}$ be a filtration on a probability space $(\Omega, \mathcal{F}, P)$. Then a measurable function $T : \Omega \to \mathbb{Z}_+ \cup \{+\infty\}$ is called a* stopping time *(or a random time) with respect to the filtration* $\{\mathcal{F}_n\}$ *if* $\{\omega : T(\omega) = n\} \in \mathcal{F}_n$ *for every* n.

Remark 2.1.3 *1) By definition, if* T *is a stopping time, then* $\{\omega : T(\omega) = +\infty\} \in \mathcal{F}$.

2) A random variable $T : \Omega \to \mathbb{Z}_+ \cup \{+\infty\}$ *is a stopping time if and only if* $\{\omega : T(\omega) \leq n\} \in \mathcal{F}_n$ *for every* n. *Indeed*

$$\{\omega : T(\omega) \leq n\} = \cup_{k=0}^{n} \{\omega : T(\omega) = k\} \ .$$

Of course, a constant time $T = n$ for some $n \in \mathbb{N}$ or $+\infty$ is a stopping time.

Example 2.1.4 *A basic example of stopping times is the following. Let* $\{X_n\}$ *be an adapted process on a filtered probability space* $(\Omega, \mathcal{F}_n, \mathcal{F}, P)$, *and let* $B \in \mathcal{B}(\mathbb{R})$. *Then the first time* T *that the process* $\{X_n\}$ *hits a domain* B:

$$T(\omega) = \inf\{n \geq 0 : X_n(\omega) \in B\}$$

(with the convention that $\inf \Phi = +\infty$*) is a stopping time with respect to the filtration* $\{\mathcal{F}_n\}$. T *is called a hitting time.*

Observe that

$$\{T = n\} = \cap_{k=0}^{n-1} \{X_k \in B^c\} \cap \{X_n \in B\} \ .$$

which belongs to $\mathcal{F}_n$ for every n. IN fact, since $\{X_n\}$ is adapted, and therefore $\{X_k \in B^c\} \in \mathcal{F}_k$ and $\{X_n \in B\} \in \mathcal{F}_n$. Hence $\{T = n\} \in \mathcal{F}_n$ and therefore T is a stopping time.

Given a stopping time T on the filtered probability space $(\Omega, \mathcal{F}_n, \mathcal{F}, P)$, the σ-algebra that represents information available up to random time T:

$$\mathcal{F}_T = \{A \in \mathcal{F} : \text{ such that } A \cap \{T \leq n\} \in \mathcal{F}_n \text{ for any } n \in \mathbb{Z}_+\} \ .$$

It is obvious that $\mathcal{F}_T = \mathcal{F}_n$ if $T = n$ is a constant time n.

Theorem 2.1.5 *Let* $\{X_n\}$ *be an adapted process on* $(\Omega, \mathcal{F}_n, \mathcal{F}, P)$, *let* ξ *be a random variable, and set* $X_\infty = \xi$. *Let* T *be a stopping time with respect to* $\{\mathcal{F}_n\}$, *and define* $X_T(\omega) = X_{T(\omega)}(\omega)$ *for any* $\omega \in \Omega$. *Then* X_T *is* $\mathcal{F}_T$*-measurable. In particular,* $X_T 1_{\{T<\infty\}}$ *is* $\mathcal{F}_T$*-measurable.*

Proof. For any $r \in \mathbb{R}$, we have

$$\{X_T \le r\} \cap \{T \le n\} = \cup_{k=0}^{n} \{X_k \le r\} \cap \{T = k\}$$

which belongs to $\mathcal{F}_n$ as $\{X_k \le r\} \cap \{T = k\} \in \mathcal{F}_k$, $k = 0, 1, \cdots, n$, so that X_T is $\mathcal{F}_T$-measurable. ■

The concept of a martingale originated from a fair game which once (perhaps still) was popular, in which regardless of the whims of chance in deciding the outcomes to the past and present, a gamer's fortune is exactly the gamer's current capital.

Definition 2.1.6 *Let $\{X_n\}$ be an adapted process on a filtered probability space $(\Omega, \mathcal{F}_n, \mathcal{F}, P)$. Suppose each $X_n \in L^1(\Omega, \mathcal{F}, P)$ (i.e. X_n is integrable).*

1. *$\{X_n\}$ is a martingale, if $E(X_{n+1}|\mathcal{F}_n) = X_n$ for all n.*

2. *$\{X_n\}$ is a supermartingale (resp. a submartingale), if $E(X_{n+1}|\mathcal{F}_n) \le X_n$ (resp. $E(X_{n+1}|\mathcal{F}_n) \ge X_n$) for every n.*

Example 2.1.7 *(Martingale transform) Let $\{H_n\}$ be a predictable process and $\{X_n\}$ be a martingale, and let*

$$(H.X)_n = \sum_{k=1}^{n} H_k(X_k - X_{k-1}), \ (H.X)_0 = 0 \ .$$

Then $\{(H.X)_n\}$ is a martingale.

Recall Jensen's inequality for conditional expectation: if $\varphi : \mathbb{R} \to \mathbb{R}$ is a convex function, ξ, $\varphi(\xi) \in L^1(\Omega, \mathcal{F}, P)$, and $\mathcal{G}$ is a sub σ-field of $\mathcal{F}$, then

$$\varphi(E(\xi|\mathcal{G})) \le E(\varphi(\xi)|\mathcal{G}) \ .$$

$\varphi(t) = (t \ln t)\, 1_{(1,\infty)}(t)$, $t1_{(0,\infty)}$ and $|t|^p$ (for $p \ge 1$) are all convex functions.

Theorem 2.1.8 *1) Let $\{X_n\}$ be a martingale, and let $\varphi : \mathbb{R} \to \mathbb{R}$ be a convex function. If $\varphi(X_n)$ is integrable for any n, then $\{\varphi(X_n)\}$ is a submartingale.*

2) If $\{X_n\}$ is a submartingale, and $\varphi : \mathbb{R} \to \mathbb{R}$ is a increasing and convex function. If $\varphi(X_n)$ is integrable for any n, then $\{\varphi(X_n)\}$ is a submartingale.

Proof. For example, let us prove the first claim. Indeed

$$\begin{aligned} \varphi(X_n) &= \varphi(E(X_{n+1}|\mathcal{F}_n)) \quad \text{(martingale property)} \\ &\le E(\varphi(X_{n+1})|\mathcal{F}_n) \quad \text{(Jensen's inequality).} \end{aligned}$$

■

Corollary 2.1.9 *If $X = (X_n)$ is a sub-martingale, so is (X_n^+). If, in addition, each $X_n \log^+ X_n$ is integrable, then $(X_n \log^+ X_n)$ is a sub-martingale, where* $\log^+ x = 1_{\{x \geq 1\}} \log x$.

2.2 Doob's inequalities

The most important result about martingales is the following Doob's optional sampling theorem, which says the (super-, sub-) martingale property holds at random times.

Theorem 2.2.1 *Let $\{X_n\}$ be a martingale (resp. super-martingale), and let S, T be two* bounded *stopping times. Suppose $S \leq T$. Then $E(X_T|\mathcal{F}_S) = X_S$ (resp. $E(X_T|\mathcal{F}_S) \leq X_S$).*

Proof. We prove this theorem for the case that $\{X_n\}$ is a supermartingale. Let $T \leq n$ (as it is bounded by our assumption). Then

$$\begin{aligned} E|X_T| &= \sum_{j=0}^n E\left(|X_T| : T = j\right) \\ &= \sum_{j=0}^n E\left(|X_j| : T = j\right) \\ &\leq \sum_{j=0}^n E|X_j| \ , \end{aligned}$$

which implies that X_T is integrable. Similarly $X_S \in L^1(\Omega, \mathcal{F}_S, P)$.

Let $A \in \mathcal{F}_S$, $j \geq 0$. Then $A \cap \{S = j\} \in \mathcal{F}_j$ and $\{T > j\} \in \mathcal{F}_j$ as S and T are stopping times. We consider several cases.

1. If $0 \leq T - S \leq 1$, then

$$\begin{aligned} \int_A (X_S - X_T) \mathrm{d}P &= \sum_{j=0}^n \int_{A \cap \{S=j\} \cap \{T>j\}} (X_j - X_{j+1}) \mathrm{d}P \\ &\geq 0 \end{aligned}$$

as each term in the above sum is non-negative.

2. General case. Set $R_j = T \wedge (S + j)$, $j = 1, \cdots, n$. Then each R_j is a stopping time,

$$S \leq R_1 \leq \cdots \leq R_n = T$$

and

$$R_1 - S \leq 1, \quad R_{j+1} - R_j \leq 1, \quad \text{for } 1 \leq j \leq n-1 \ .$$

Let $A \in \mathcal{F}_S$. Then $A \in \mathcal{F}_{R_j}$ (as $S \leq R_j$). Therefore apply the first case to R_j we have

$$\int_A X_S \mathrm{d}P \geq \int_A X_{R_1} \mathrm{d}P \geq \cdots \geq \int_A X_T \mathrm{d}P$$

so that

$$E\left(1_A X_S\right) \geq E\left(1_A X_T\right) \qquad \text{for any } A \in \mathcal{F}_S \ .$$

Since $X_S \in \mathcal{F}_S$ we may thus conclude that

$$X_S \geq E(X_T | \mathcal{F}_S) \ .$$

Thus we have proved the theorem. ∎

Corollary 2.2.2 *Let $\{X_n\}$ be a submartingale, and let T be a stopping time. Then*

$$E|X_{T\wedge k}| \leq E\left(X_0\right) + 2E\left(X_k^-\right) \quad \textit{for any } k = 0, 1, 2, \cdots$$

and therefore

$$E(|X_T| 1_{\{T<\infty\}}) \leq 3 \sup_n E|X_n| \ .$$

Proof. We know that $\{X_n^-\}$ be a supermartingale, so by the previous theorem we have

$$\begin{aligned} E|X_{T\wedge k}| &= EX_{T\wedge k} + 2E\left(X_{T\wedge k}^-\right) \\ &\leq EX_0 + 2E\left(X_k^-\right) \ . \end{aligned}$$

This is the first inequality. While

$$\begin{aligned} E(|X_{T\wedge k}| 1_{\{T<\infty\}}) &\leq EX_0 + 2E\left(X_k^-\right) \\ &\leq 3 \sup_n E\left(|X_n|\right) \end{aligned}$$

and the second inequality thus follows from the Fatou lemma. ∎

Theorem 2.2.3 *(Doob's maximal inequality). Let $\{X_n\}$ be a super-martingale, and let $n \geq 0$. Then for any $\lambda > 0$, we have*

$$\begin{aligned} \lambda P\left\{\sup_{k\leq n} X_k \geq \lambda\right\} &\leq EX_0 - \int_{\left\{\sup_{k\leq n} X_k < \lambda\right\}} X_n dP \\ &= EX_0 - E\left\{X_n \colon \sup_{k\leq n} X_k < \lambda\right\} \ ; \end{aligned}$$

$$\begin{aligned}\lambda P\left\{\inf_{k\le n} X_k \le -\lambda\right\} &\le \int_{\{\inf_{k\le n} X_k\le -\lambda\}} -X_n\,dP \\ &= -E\left\{X_n\colon \inf_{k\le n} X_k \le -\lambda\right\}\end{aligned}$$

and

$$\lambda P\left\{\sup_{k\le n}|X_k| \ge \lambda\right\} \le EX_0 + 2E\left(X_n^-\right) \ .$$

Proof. Let us prove the first inequality. Let $R = \inf\{k \ge 0 : X_k \ge \lambda\}$ and $T = R \wedge n$. Then T is a bounded stopping time. By definition,

$$X_R \ge \lambda, \qquad \text{on } \ \{R < \infty\},$$

so that

$$\begin{aligned}\{\sup_{k\le n} X_k &\ge \lambda\} \subseteq \{X_T \ge \lambda\}, \\ \{\sup_{k\le n} X_k &< \lambda\} \subseteq \{T = n\}.\end{aligned}$$

By Doob's optional sampling theorem,

$$\begin{aligned}EX_0 &\ge EX_T \\ &= \int_{\{\sup_{k\le n} X_k\ge\lambda\}} X_T \mathrm{d}P + \int_{\{\sup_{k\le n} X_k<\lambda\}} X_T \mathrm{d}P \\ &\ge \lambda P\left\{\sup_{k\le n} X_k \ge \lambda\right\} + \int_{\{\sup_{k\le n} X_k<\lambda\}} X_n \mathrm{d}P \\ &= \lambda P\left\{\sup_{k\le n} X_k \ge \lambda\right\} + E\left\{X_n; \sup_{k\le n} X_k < \lambda\right\} \ .\end{aligned}$$

In order to prove the second inequality, we set $Y_k = -X_k$. Then $\{Y_n\}$ is a submartingale. Define

$$R = \inf\{k \ge 0 : Y_k \ge \lambda\}, \qquad T = R \wedge n.$$

Then T is a stopping time and $T \le n$. Again we have

$$\begin{aligned}\{\sup_{k\le n} Y_k &\ge \lambda\} \subseteq \{Y_T \ge \lambda\} \ ; \\ \{\sup_{k\le n} Y_k &< \lambda\} \subseteq \{T = n\} \ .\end{aligned}$$

Therefore by applying Doob's stopping theorem to Y we have

$$\begin{aligned} EY_n &\geq EY_T \\ &= \int_{\{\sup_{k\leq n} Y_k \geq \lambda\}} Y_T \mathrm{d}P + \int_{\{\sup_{k\leq n} Y_k < \lambda\}} Y_T \mathrm{d}P \\ &\geq \lambda P\left\{\sup_{k\leq n} Y_k \geq \lambda\right\} + \int_{\{\sup_{k\leq n} Y_k < \lambda\}} Y_n \mathrm{d}P \\ &= \lambda P\left\{\sup_{k\leq n} -X_k \geq \lambda\right\} + \int_{\{\sup_{k\leq n} Y_k < \lambda\}} Y_n \mathrm{d}P. \end{aligned}$$

Therefore

$$\begin{aligned} \lambda P\left\{\sup_{k\leq n} -X_k \geq \lambda\right\} &= \lambda P\left\{\inf_{k\leq n} X_k \leq -\lambda\right\} \\ &\leq EY_n - \int_{\{\sup_{k\leq n} Y_k < \lambda\}} Y_n \mathrm{d}P \\ &= \int_{\{\sup_{k\leq n} Y_k \geq \lambda\}} Y_n \mathrm{d}P \\ &= -\int_{\{\inf_{k\leq n} X_k \leq -\lambda\}} X_n \mathrm{d}P. \end{aligned}$$

The third inequality follows from the first two inequalities. ■

As a consequence we have

Theorem 2.2.4 *(Kolmogorov's inequality) Let* $\{X_n\}$ *be a martingale and* $X_n \in L^2(\Omega, \mathcal{F}, P)$*. Then for any* $\lambda > 0$*,*

$$P\left\{\sup_{k\leq n} |X_k| \geq \lambda\right\} \leq \frac{1}{\lambda^2} E\left(X_n^2\right) \ .$$

Proof. By Jensen's inequality, for any $k \leq n$ we have

$$\begin{aligned} E\left(X_k^2\right) &= E\left(E(X_n|\mathcal{F}_k)\right)^2 \\ &\leq E\left(X_n^2\right) < \infty \ . \end{aligned}$$

Therefore $(-X_k^2)$ $(k = 0, 1, \cdots, n)$ is a supermartingale. By the second inequality in the above theorem, we have

$$\lambda^2 P\left\{\inf_{k\leq n} -X_k^2 \leq -\lambda^2\right\} \leq \int_{\left\{\inf_{k\leq n} -X_k^2 \leq -\lambda^2\right\}} X_n^2 \mathrm{d}P$$

and therefore

$$\begin{aligned}\lambda^2 P\left\{\sup_{k\le n} X_k^2 \ge \lambda^2\right\} &\le \int_{\{\inf_{k\le n} -X_k^2 \le -\lambda^2\}} X_n^2 \mathrm{d}P \\ &\le \int_\Omega X_n^2 \mathrm{d}P = E\left(X_n^2\right) .\end{aligned}$$

■

Next we establish Doob's L^p-inequality. Let $X_n^* = \max_{k\le n} X_k$. If $\Phi : \mathbb{R}_+ \to [0,\infty)$ is a continuous and increasing function such that $\Phi(0) = 0$, then

$$E\Phi(X_n^*) = E\int_0^{X_n^*} \mathrm{d}\Phi(\lambda) = \int_{\Omega\times[0,X_n^*]} \mathrm{d}\Phi(\lambda)\mathrm{d}P .$$

Theorem 2.2.5 *(Doob's L^p-inequality) 1) If (X_n) is a sub-martingale, then for any $p > 1$*

$$E\left(\max_{k\le n} X_k^+\right)^p \le \left(\frac{p}{p-1}\right)^p E|X_n^+|^p .$$

2) If (X_n) is a martingale, then for any $p > 1$,

$$E\left(\max_{k\le n} |X_k|^p\right) \le \left(\frac{p}{p-1}\right)^p E|X_n|^p .$$

Proof. If (X_n) is a martingale, then $(|X_n|)$ is a submartingale, so 2) follows 1). Let us prove the first conclusion. By replace (X_n) by (X_n^+), we may, without lose of the generality, assume that (X_n) is a non-negative sub-martingale. By Fubin's theorem

$$\begin{aligned}E\Phi(X_n^*) &= E\left\{\int_0^{X_n^*} \mathrm{d}\Phi(\lambda)\right\} \\ &= \int_0^\infty P(X_n^* \ge \lambda)\mathrm{d}\Phi(\lambda) \\ &\le \int_0^\infty \frac{1}{\lambda} E(X_n; X_n^* \ge \lambda)\mathrm{d}\Phi(\lambda)\end{aligned}$$

together with Doob's maximal inequality

$$P(X_n^* \ge \lambda) \le \frac{1}{\lambda} E\left\{X_n : X_n^* \ge \lambda\right\}$$

we thus obtain

$$\begin{aligned} E\Phi(X_n^*) &\leq \int_0^\infty \frac{1}{\lambda} E\left\{X_n : X_n^* \geq \lambda\right\} \mathrm{d}\Phi(\lambda) \\ &= \int_0^\infty \frac{1}{\lambda} \int_{\{X_n^* \geq \lambda\}} X_n \mathrm{d}P \mathrm{d}\Phi(\lambda) \\ &= E\left\{X_n \left(\int_0^{X_n^*} \frac{1}{\lambda} \mathrm{d}\Phi(\lambda)\right)\right\} . \end{aligned} \tag{2.1}$$

Choose $\Phi(\lambda) = \lambda^p$, then $\Phi'(\lambda) = p\lambda^{p-1}$, and therefore

$$\begin{aligned} E|X_n^*|^p &\leq E\left\{X_n \left(\int_0^{X_n^*} \frac{1}{\lambda} p\lambda^{p-1} \mathrm{d}\lambda\right)\right\} \\ &= E\left\{\frac{p}{p-1} X_n \left(X_n^*\right)^{p-1}\right\} \\ &= \frac{p}{p-1} E\left(X_n \left(X_n^*\right)^{p-1}\right) \\ &\leq \frac{p}{p-1} (|EX_n|^p)^{\frac{1}{p}} (E|X_n^*|^p)^{\frac{1}{q}} \end{aligned}$$

the lase equality follows from Hölder inequality. ■

Exercise 2.2.6 *Prove* $\log x \leq x/e$ *for all* $x > 0$*, hence prove that*

$$a \log^+ b \leq a \log^+ a + \frac{b}{e}.$$

Theorem 2.2.7 *(Doob's inequality) Let* (X_n) *be a non-negative sub-martingale. Then*

$$E\left\{\max_{k \leq n} X_k\right\} \leq \frac{e}{e-1}\left\{1 + \max_{k \leq n} E\left(X_k \log^+ X_k\right)\right\} .$$

Proof. We may use the same argument as in the proof of the previous theorem, but with the choice that $\Phi(\lambda) = (\lambda - 1)^+$. We thus obtain (by (2.1))

$$\begin{aligned} E\left\{\Phi(X_n^*)\right\} &\leq E\left\{X_n \left(\int_0^{X_n^*} \frac{1}{\lambda} \mathrm{d}\Phi(\lambda)\right)\right\} \\ &= E\left\{X_n \left(1_{\{X_n^* \geq 1\}} \int_1^{X_n^*} \frac{1}{\lambda} \mathrm{d}\lambda\right)\right\} \\ &= E\left(X_n \log^+ X_n^*\right) , \end{aligned}$$

which thus implies that

$$E\left(X_n^*-1\right) \leq E\left(X_n^*-1\right)^+ \leq E\left\{X_n \log^+ X_n^*\right\} \ .$$

Together with the inequality (see the previous exercise)

$$X_n \log^+ X_n^* \leq X_n \log^+ X_n + \frac{1}{e} X_n^*$$

it follows thus that

$$\begin{aligned} E\left(X_n^*-1\right) &\leq E\left\{X_n \log^+ X_n^*\right\} \\ &\leq E\left\{X_n \log^+ X_n\right\} + \frac{1}{e} E X_n^* \end{aligned}$$

so that

$$E X_n^* \leq \frac{1}{1-1/e} E\left\{X_n \log^+ X_n\right\} \ .$$

■

2.3 The convergence theorem

Let $\{X_n : n \in \mathbb{Z}_+\}$ be an adapted sequence of random variables, and $[a,b]$ be a closed interval. Define

$$\begin{aligned} T_0 &= \inf\{n \geq 0 : X_n \leq a\} \ ; \\ T_1 &= \inf\{n > T_0 : X_n \geq b\} \ ; \\ & \cdots\cdots \\ T_{2j} &= \inf\{n > T_{2j-1} : X_n \leq a\} \ ; \\ T_{2j+1} &= \inf\{n > T_{2j} : X_n \geq b\} \ . \end{aligned}$$

Then $\{T_k\}$ be a sequence of stopping times, which is increasing $T_k \uparrow$. If $T_{2j-1}(\omega) < \infty$, then the sequence

$$X_0(\omega), \cdots, X_{T_{2j-1}}(\omega)$$

upcrosses the interval $[a,b]$ j times. Denote by $U_a^b(X;n)$ the number of upcrossing $[a,b]$ by $\{X_k\}$ up to time n. Then

$$\{U_a^b(X;n) = j\} = \{T_{2j-1} \leq n < T_{2j+1}\} \in \mathcal{F}_n.$$

According to definition

$$\begin{aligned} X_{T_{2j}} &\leq a, \quad \text{on } \{T_{2j} < \infty\} \ ; \\ X_{T_{2j+1}} &\geq b, \quad \text{on } \{T_{2j+1} < \infty\} \ . \end{aligned}$$

Theorem 2.3.1 *(Doob's upcrosssing theorem) 1) If $X = \{X_n\}$ is a supermartingale, then for any $n \geq 1$, $k \geq 0$, we have*

$$P\left\{U_a^b(X;n) \geq k\right\} \leq \frac{1}{b-a} E\left\{(X_n - a)^- : U_a^b(X;n) = k\right\}$$

and

$$EU_a^b(X;n) \leq \frac{1}{b-a} E(X_n - a)^- \ .$$

2. Similarly, if $X = \{X_n\}$ is a submartingale, then

$$P\left\{U_a^b(X;n) \geq k\right\} \leq \frac{1}{b-a} E\left\{(X_n - a)^+ : U_a^b(X;n) = k\right\}$$

and

$$EU_a^b(X;n) \leq \frac{1}{b-a} E(X_n - a)^+.$$

Proof. We first prove the inequalities for a supermartingale. Since X is a supermartingale, by Doob's optional sampling theorem,

$$\begin{aligned} 0 &\geq E\left(X_{T_{2k+1}\wedge n} - X_{T_{2k}\wedge n}\right) \\ &= E\left(X_{T_{2k+1}\wedge n} - X_{T_{2k}\wedge n}\right) 1_{\{T_{2k} \leq n < T_{2k+1}\}} \\ &\quad + E\left(X_{T_{2k+1}\wedge n} - X_{T_{2k}\wedge n}\right) 1_{\{T_{2k+1} \leq n\}} \\ &\geq E\left(X_n - a\right) 1_{\{T_{2k} \leq n < T_{2k+1}\}} \quad (\text{ as } X_{T_{2k}\wedge n} = X_{T_{2k}} \leq a) \\ &\quad + E\left(b-a\right) 1_{\{T_{2k+1} \leq n\}} \quad (\text{as } X_{T_{2k}+1\wedge n} = X_{T_{2k+1}} \geq b). \end{aligned}$$

However

$$\{U_a^b(X;n) \geq k\} \subset \{T_{2k-1} \leq n\},$$

$$\{U_a^b(X;n) = k\} = \{T_{2k-1} \leq n < T_{2k}\}$$

so that

$$\begin{aligned} 0 &\geq E\left(X_n - a\right) 1_{\{U_a^b(X;n)=k\}} \\ &\quad + E\left(b-a\right) 1_{\{U_a^b(X;n)\geq k\}} \\ &= E\left(X_n - a\right) 1_{\{U_a^b(X;n)=k\}} + (b-a)\, P\{U_a^b(X;n) \geq k\} \end{aligned}$$

which yields the first inequality. By adding up over all $k \geq 0$ we get the second inequality.

Now we prove the inequalities for a submartingale X. The argument is very similar. Again by Doob's stopping theorem,

$$\begin{aligned} 0 &\geq E\left(X_{T_{2k-1}\wedge n} - X_{T_{2k}\wedge n}\right) \\ &= E\left(X_{T_{2k-1}\wedge n} - X_{T_{2k}\wedge n}\right) 1_{\{T_{2k-1}\leq n<T_{2k}\}} \\ &\quad +E\left(X_{T_{2k-1}\wedge n} - X_{T_{2k}\wedge n}\right) 1_{\{T_{2k}\leq n\}} \\ &\geq E\left(b - X_n\right) 1_{\{T_{2k-1}\leq n<T_{2k}\}} + E\left(b-a\right) 1_{\{T_{2k}\leq n\}} \\ &= E\left(a - X_n\right) 1_{\{T_{2k-1}\leq n<T_{2k}\}} + E\left(b-a\right) 1_{\{T_{2k-1}\leq n\}} \end{aligned}$$

which yields the wanted inequality. ■

Theorem 2.3.2 *(The martingale convergence theorem). Let $\{X_n\}$ be a supermartingale. If $\sup_n E|X_n| < +\infty$, then*

$$X_n \to X_\infty \text{ exists almost surely.}$$

Moreover if in addition $\{X_n\}$ is non-negative, then

$$E(X_\infty|\mathcal{F}_n) \leq X_n \text{ for any } n.$$

Proof. For any rationales $a, b \in \mathbb{Q}, a < b$ we set

$$U_a^b(X) = \lim_{n\to\infty} U_a^b(X; n).$$

Then by the Fatou lemma

$$\begin{aligned} EU_a^b(X) &\leq \frac{1}{b-a}\sup_n E(X_n - a)^- \\ &\leq \frac{|a|}{b-a} + \frac{1}{b-a}\sup_n E|X_n| < \infty. \end{aligned}$$

Therefore

$$U_a^b(X) < \infty, \quad \text{almost surely.}$$

Let

$$W_{(a,b)} = \{\mathrm{liminf}_{n\to\infty} X_n < a,\ \mathrm{limsup}_{n\to\infty} X_n > b\}$$

and

$$W = \cup_{(a,b)} W_{(a,b)}$$

the union over all rational pairs (a, b), $a < b$, which is a countably union. Clearly

$$W_{(a,b)} \subset \{U_a^b(X) = \infty\}$$

so that

$$P(W_{(a,b)}) = 0.$$

Hence $P(W) = 0$. However if $\omega \notin W$, then $\lim_{n\to\infty} X_n(\omega)$ exists, and we denote it by $X_\infty(\omega)$ and on W we let $X_\infty(\omega) = 0$. Then we have $X_n \to X_\infty$ almost surely. Moreover by the Fatou lemma,

$$E|X_\infty| \leq \sup_n E|X_n| < \infty,$$

i.e. $X_\infty \in L^1(\Omega, \mathcal{F}, P)$.

If in addition $\{X_n\}$ is non-negative, then

$$E(X_m|\mathcal{F}_n) = X_n, \text{ for any } m \geq n,$$

by letting $m \to \infty$, the Fatou lemma then yields that

$$E(X_\infty|\mathcal{F}_n) \leq X_n \ .$$

■

2.4 Martingales in continuous-time

The concept of martingales (super- and submartingales) and Doob's fundamental inequalities in discrete-time may be extended to martingales in continuous time.

In this section, we present the regularity theory for martingales, which does not appear in the discrete-time case.

Let $(\Omega, \mathcal{G}, \mathcal{G}_t, P)$ be a probability space with a filtration $(\mathcal{G}_t)_{t\geq 0}$ which is an increasing family of σ-algebras $\mathcal{G}_t \subset \mathcal{G}$ for all $t \geq 0$. A $(\mathcal{G}_t)$-adapted (real valued) process $(X_t)_{t\geq 0}$ is called a martingale (resp. supermartingale; resp. submartingale), if for any $t \geq s$, almost surely $E\left(X_t|\mathcal{G}_s\right) = X_s$ (resp. $E\left(X_t|\mathcal{G}_s\right) \leq X_s$; resp. $E\left(X_t|\mathcal{G}_s\right) \geq X_s$). Similarly, the concept of stopping times can be stated in this setting as well, namely, a function $T : \Omega \to [0, +\infty]$ is a $(\mathcal{G}_t)$-stopping time if for every $t \geq 0$, the event $\{T \leq t\}$ belongs to $\mathcal{G}_t$. A new kind of stopping times called predictable times which has no interest in discrete-time case will play a role if the underlying stochastic processes have jumps. A stopping time $T : \Omega \to [0, +\infty]$ is predictable if there is an increasing sequence $\{T_n\}$ of $(\mathcal{G}_t)$-stopping times such that for each n, $T_n < T$ and $\lim_{n\to\infty} T_n = T$.

Let

$$\mathcal{G}_T = \{A \in \mathcal{G} : \text{ for any } t \geq 0, (T \leq t) \cap A \in \mathcal{G}_t\}$$

be the σ-algebra representing the information available up to the random time T, and let

$$\mathcal{G}_{T-} = \{A \in \mathcal{G} : \text{ for any } t \geq 0, \, (T < t) \cap A \in \mathcal{G}_t\}$$

which represents information known strictly before time T.

The following lemma can be used to generalize many results about martingales in discrete-time to continuous-time setting.

Lemma 2.4.1 *Let $T : \Omega \to [0, +\infty]$ be a $(\mathcal{G}_t)$-stopping time. For every n let*

$$T^{(n)} = \sum_{k=1}^{\infty} \frac{k}{2^n} 1_{\left\{\frac{k-1}{2^n} \leq T < \frac{k}{2^n}\right\}} + (+\infty) 1_{\{T=+\infty\}} \ .$$

Then $T^{(n)} \geq T$ are $(\mathcal{G}_t)$-stopping times and $T^{(n)} \downarrow T$ as $n \to \infty$.

Proof. For any n and $t \geq 0$ we have

$$\begin{aligned}
\left\{T^{(n)} \leq t\right\} &= \bigcup_{k=1}^{\infty} \left\{T^{(n)} \leq t\right\} \cap \left\{\frac{k-1}{2^n} \leq T < \frac{k}{2^n}\right\} \\
&= \bigcup_{k/2^n \leq t} \left\{T^{(n)} \leq t\right\} \cap \left\{\frac{k-1}{2^n} \leq T < \frac{k}{2^n}\right\} \\
&\in \bigvee_{k/2^n \leq t} \mathcal{G}_{\frac{k}{2^n}} \subset \mathcal{G}_t \ .
\end{aligned}$$

■

For each $t \geq 0$ let $\mathcal{G}_{t+} = \cap_{s>t} \mathcal{G}_s$. Then $(\mathcal{G}_{t+})$ is again a filtration on the measurable space $(\Omega, \mathcal{G})$ and obviously $\mathcal{G}_{t+} \supseteq \mathcal{G}_t$ for every t. If $T : \Omega \to [0, +\infty]$ is a $(\mathcal{G}_{t+})$-stopping time then

$$\mathcal{G}_{T+} = \{A \in \mathcal{G} : \text{ for any } t \geq 0, \, (T \leq t) \cap A \in \mathcal{G}_{t+}\} \ .$$

A filtration (G_t) is said to be right-continuous if $\mathcal{G}_{t+} = \mathcal{G}_t$ for each $t \geq 0$. According to the definition, $(\mathcal{G}_{t+})$ is right-continuous. Similarly, for $t > 0$ we define $\mathcal{G}_{t-} = \sigma\{\mathcal{G}_s : s < t\}$.

Theorem 2.4.2 *If $(X_t)_{t\geq 0}$ is a martingale (resp. supermartingale, resp. submartingale) on $(\Omega, \mathcal{G}, \mathcal{G}_t, P)$ with right-continuous sample paths almost surely, then $(X_t)_{t\geq 0}$ is a martingale (resp. supermartingale, resp. submartingale) on $(\Omega, \mathcal{G}, \mathcal{G}_{t+}, P)$.*

Proof. Let us prove the supermartingale case. Since $(X_t)_{t\geq 0}$ is adapted to $(\mathcal{G}_{t+})_{t\geq 0}$ so we need to prove

$$E\left(X_t|\mathcal{G}_{s+}\right) \leq X_s \qquad P\text{-a.s.} \tag{2.2}$$

for every $t > s$. For any u between s and t

$$E\left(X_t|\mathcal{G}_u\right) \leq X_u \qquad P\text{-a.s.}$$

so that for any $A \in \mathcal{G}_{s+} \subset \mathcal{G}_u$

$$E\left(1_A X_t\right) \leq E\left(1_A X_u\right) \ .$$

Letting $u \downarrow s$, as $\lim_{u\downarrow s} X_u = X_s$, we thus obtain

$$E\left(1_A X_t\right) \leq E\left(1_A X_s\right)$$

for any $A \in \mathcal{G}_{s+}$, which is equivalent to (2.2). ■

Corollary 2.4.3 *(Doob's optional sampling) Let $(X_t)_{t\geq 0}$ be a supermartingale on $(\Omega, \mathcal{G}, \mathcal{G}_t, P)$ with almost all right-continuous sample paths, and let T be a $(\mathcal{G}_t)$-stopping time. Then*

$$E\left(X_{s+T}1_{\{T<+\infty\}}|\mathcal{G}_{T+}\right) \leq X_T 1_{\{T<+\infty\}} \qquad P\text{-a.s.}$$

for any $s \geq 0$.

Proof. The only thing needed here is the fact that $\mathcal{G}_{T+}$ is the σ-algebra at random time T with respect to the right-continuous filtration $(\mathcal{G}_{t+})_{t\geq 0}$, Corollary follows from Doob's optional sampling and the above theorem. ■

Similar conclusions hold for martingales and submartingales.

One would ask when a martingale (supermartingale) has right-continuous sample paths almost surely. The question can be answered via Doob's convergence theorem for supermartingales.

Let $X = (X_t)_{t\geq 0}$ be a real valued stochastic process, and let $a < b$. If

$$F = \{0 \leq t_1 < t_2 < \cdots < t_N\}$$

is a finite subset of $[0, +\infty)$, then $U_a^b(X, F)$ denotes the number of upcrossings by $\{X_{t_1}, \cdots, X_{t_N}\}$, and if $D \subset [0, +\infty)$, then $U_a^b(X, D)$ denotes the superemum of $U_a^b(X, F)$ over F being subset of D. Obviously $D \to U_a^b(X, D)$ is increasing with respect to the inclusion $\subset$. In particular, if $X = (X_t)_{t\geq 0}$ is a $(\mathcal{G}_t)$-adapted process and if D is a countable subset of $[0, +\infty)$ then for every $t \geq 0$, $U_a^b(X, D\cap[0, t])$ is measurable with respect to $\mathcal{G}_t$. Since we may apply Doob's upcrossing number inequality to $(X_t)_{t\in F}$ where F is a finite subset, therefore establish the following

Theorem 2.4.4 *(Doob's upcrossing number inequality). If $X = (X_t)_{t\geq 0}$ is a supermartingale on $(\Omega, \mathcal{G}, \mathcal{G}_t, P)$, then for any $a < b$, $t > 0$ and any countable subset D of $[0, t]$*

$$EU_a^b(X, D) \leq \frac{1}{b-a} E\left(X_t - a\right)^-$$

where $x^- = (-x) \vee 0$.

It follows thus the following version of the supermartingale convergence theorem.

Corollary 2.4.5 *Let $X = (X_t)_{t\geq 0}$ be a supermartingale on $(\Omega, \mathcal{G}, \mathcal{G}_t, P)$, and let D be a countable dense subset of $[0, +\infty)$. Then for almost all $w \in \Omega$, the right limit of $(X_t)_{t\geq 0}$ along the countable dense set D at t*

$$\lim_{s\in D, s>t, s\downarrow t} X_s \quad \textit{exists} \ .$$

Similarly for almost all $w \in \Omega$, for each $t > 0$, the left limt of X_t along D at time t,

$$\lim_{s\in D, s<t, s\uparrow t} X_s \quad \textit{exists} \ .$$

We are now in a position to prove the following fundamental theorem, in literature, which is called Föllmer's lemma.

Theorem 2.4.6 *Let $(X_t)_{t\geq 0}$ be a supermartingale (resp. martingale) on $(\Omega, \mathcal{G}, \mathcal{G}_t, P)$, and let D be a countable dense subset in $[0, +\infty)$.*

1. *For almost all $w \in \Omega$*

$$Z_t(w) = \lim_{s\in D, s>t, s\downarrow t} X_s(w)$$

   exists for all $t \geq 0$, and Z_t is $\mathcal{G}_{t+}$-measurable,

2. *For almost all $w \in \Omega$ and for all $t > 0$ the following left limit*

$$Z_{t-}(w) = \lim_{s<t, s\uparrow} Z_s(w) \ ,$$

   and therefore $(Z_t)_{t\geq 0}$ is a $(\mathcal{G}_{t+})$-adapted process with right-continuous sample paths and left limits.

3. *For any* $t \geq 0$

$$E\left(Z_t|\mathcal{G}_t\right) \leq X_t \ , \quad \textit{resp. } E\left(Z_t|\mathcal{G}_t\right) = X_t \qquad P\textit{-a.s.}$$

4. $(Z_t)_{t\geq 0}$ *is a supermartingale (resp. martingale) on* $(\Omega, \mathcal{G}, \mathcal{G}_{t+}, P)$.

Proof. We only need to prove conclusions 3 and 4. Firstly we show the third statement. For any $r > t$ $r \in D$

$$E\left(X_r|\mathcal{G}_t\right) \leq X_t$$

so that for any $A \in \mathcal{G}_t$

$$E\left(1_A X_r\right) \leq E(1_A X_t) \ .$$

Letting $r \downarrow t$ along D

$$E\left(1_A Z_t\right) \leq E(1_A X_t) \ .$$

which is equivalent to the inequality in 3. Similarly, if $t > s$, $u > t > r > s$ and $u, r \in D$ then

$$E\left(X_u|\mathcal{G}_r\right) \leq X_r$$

In particular, for any $A \in \mathcal{G}_{s+} \subset \mathcal{G}_r$

$$E\left(1_A X_u\right) \leq E(1_A X_r) \ .$$

Letting $u \in D \downarrow t$ and $r \in D \downarrow s$ we obtain that

$$E\left(1_A X_t\right) \leq E(1_A X_s)$$

for every $A \in \mathcal{G}_{s+}$, which implies the statement 4. ■

We however can not in general conclude that $(Z_t)_{t\geq 0}$ is a version of $(X_t)_{t\geq 0}$. The two processes can be very different.

Corollary 2.4.7 *Under the same assumptions and notations as in Theorem 2.4.6. Assume that* $(\mathcal{G}_t)_{t\geq 0}$ *is right-continuous. Then* $(Z_t)_{t\geq 0}$ *is a version of* $(X_t)_{t\geq 0}$*, that is, for each* $t \geq 0$, $Z_t = X_t$ *almost surely, if and only if* $t \to E(X_t)$ *is right-continuous.*

Proof. Since $Z_t \in \mathcal{G}_t = \mathcal{G}_{t+}$ so that, according 3 of Theorem 2.4.6

$$Z_t = E\left(Z_t|\mathcal{G}_t\right) \leq X_t \ .$$

While, by the first conclusion in the same theorem,

$$E(Z_t) = \lim_{s \in D, s>t, s \downarrow t} E(X_s) \ .$$

Therefore, in order to have the equality $Z_t = X_t$, the necessary and sufficient condition is that

$$E(Z_t) = \lim_{s \in D, s>t, s \downarrow t} E(X_s) = E(X_t) \ .$$

However $s \to E(X_s)$ is decreasing, so the above equality is equivalent to

$$\lim_{s>t, s \downarrow t} E(X_s) = E(X_t)$$

i.e. $t \to E(X_t)$ is right-continuous. ■

Corollary 2.4.8 *Under the same assumptions and notations as in Theorem 2.4.6. If $(\mathcal{G}_t)_{t \geq 0}$ is right continuous, and if $(X_t)_{t \geq 0}$ is a martingale on $(\Omega, \mathcal{G}, \mathcal{G}_t, P)$, then the process $(Z_t)_{t \geq 0}$ defined in 2.4.6 is a version of $(X_t)_{t \geq 0}$.*

This is because for a martingale $(X_t)_{t \geq 0}$, $t \to E(X_t) = E(X_0)$ is a constant.

There is a previsible version of the optional sampling theorem. Let $(\mathcal{G}_t)_{t \geq 0}$ be a right-continuous filtration. We say a $(\mathcal{G}_t)$-stopping time $T : \Omega \to [0, +\infty]$ is predictable if there is a sequence $\{T_n\}$ of $(\mathcal{G}_t)$-stopping times such that $T_n < T$ for every n and $\lim_{n \to \infty} T_n = T$. The filtration $(\mathcal{G}_t)_{t \geq 0}$ is quasi-left continuous if for every predictable stopping time T we have $\mathcal{G}_{T-} = \mathcal{G}_T$.

Theorem 2.4.9 *(Doob's stopping time theorem) Let $(X_t)_{t \in [0, \infty]}$ be a supermartingale on $(\Omega, \mathcal{G}, \mathcal{G}_t, P)$ which is right-continuous with left-limits, where $(\mathcal{G}_t)_{t \geq 0}$ is a right-continuous filtration. Then for any predictable stopping time T and $s \geq 0$*

$$E\left(X_{s+T} 1_{\{T<+\infty\}} | \mathcal{G}_{T-} \right) \leq X_{T-} 1_{\{T<+\infty\}} \qquad a.s.$$

2.5 Local martingales

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

2.5.1 Stopping times

Recall that a random variable $T : \Omega \to [0, +\infty]$ (note that the value $+\infty$ is allowed) is called a *stopping time* (a random time) if for each $t \geq 0$ the event

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

If T is a stopping time, then

$$\mathcal{F}_T = \{A \in \mathcal{F} : A \cap \{T \leq t\} \in \mathcal{F}_t \text{ for all } t \geq 0\}$$

which represents the information available up to random time T. For technical reasons, we will require the following conditions to be satisfied, unless otherwise specified.

1. $(\Omega, \mathcal{F}, P)$ is a complete probability space.

2. The filtration $(\mathcal{F}_t)_{t \geq 0}$ is right-continuous, that is, for each $t \geq 0$

$$\mathcal{F}_t = \mathcal{F}_{t+} \equiv \cap_{s>t} \mathcal{F}_s \ .$$

3. Each $\mathcal{F}_t$ contains all null sets in $\mathcal{F}$.

In this case, we say the filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ satisfies *the usual conditions*.

Remark 2.5.1 *If $X = (X_t)_{t \geq 0}$ is a* ***right-continuous*** *stochastic process on a complete probability space $(\Omega, \mathcal{F}, P)$, then its natural filtration $(\mathcal{F}_t)_{t \geq 0}$ satisfies the usual conditions.*

The following is a result we will not prove in this course.

Theorem 2.5.2 *If $X = (X_t)_{t \geq 0}$ is a right-continuous stochastic process adapted to $(\mathcal{F}_t)_{t \geq 0}$ (recall that our filtration $(\mathcal{F}_t)_{t \geq 0}$ satisfies the usual conditions), and if $T : \Omega \to [0, +\infty]$ is a stopping time, then the random variable $X_T 1_{\{T<\infty\}}$ is measurable with respect to σ-algebra $\mathcal{F}_T$, where*

$$\begin{aligned} X_T 1_{\{T<\infty\}}(\omega) &= X_{T(\omega)}(\omega) 1_{\{\omega : T(\omega)<\infty\}}(\omega) \\ &= \begin{cases} X_{T(\omega)}(\omega) \ ; & \text{if } \ T(\omega) < +\infty \ , \\ 0; & \text{if } \ T(\omega) = +\infty \ . \end{cases} \end{aligned}$$

Remark 2.5.3 *If $X=(X_n)_{n\in\mathbb{Z}^+}$ and $T:\Omega\to\mathbb{Z}^+\cup\{+\infty\}$, then*

$$\begin{aligned} X_T 1_{\{T<\infty\}} &= \sum_{n\in\mathbb{Z}^+} X_n 1_{\{T=n\}} \\ &= \sum_{n=0}^{\infty} X_n 1_{\{T=n\}} \ . \end{aligned}$$

Therefore, if X is adapted to $\{\mathcal{F}_n\}_{n\in\mathbb{Z}^+}$ and T is a stopping time, then for any $n\in\mathbb{Z}^+$,

$$X_T 1_{\{T<\infty\}} 1_{\{T\le n\}} = \sum_{k=0}^{n} X_k 1_{\{T=k\}}$$

which is measurable with respect to $\mathcal{F}_n$, thus by definition $X_T 1_{\{T<\infty\}}$ is $\mathcal{F}_T$-measurable.

The following theorem provides us with a class of interesting stopping times.

Theorem 2.5.4 *Let $X=(X_t)_{t\ge 0}$ be an $\mathbb{R}^d$-valued, adapted stochastic process that is* right-continuous and has left-limits*. Then for any Borel subset $D\subset\mathbb{R}^d$ and $t_0\ge 0$*

$$T=\inf\{t\ge t_0 : X_t\in D\}$$

is a stopping time, where $\inf\emptyset=+\infty$*. T is called the* hitting time *of D by the process X.*

Remark 2.5.5 *Let us look at the discrete-time case. If $X=(X_n)_{n\in\mathbb{Z}^+}$ is adapted to $\{\mathcal{F}_n\}_{n\in\mathbb{Z}^+}$ taking values in $\mathbb{R}^d$. Then for a Borel subset $D\subset\mathbb{R}^d$, and $k\in\mathbb{Z}^+$*

$$T=\inf\{n\ge k : X_n\in D\}$$

is a stopping time. Indeed, if $n\le k-1$ then $\{T=n\}=\emptyset$ and for $n\ge k$ we have

$$\{T=n\}=\bigcap_{j=k}^{n-1}\{X_j\in D^c\}\bigcap\{X_n\in D\}$$

which belongs to $\mathcal{F}_n$.

Example 2.5.6 *If $X=(X_t)_{t\ge 0}$ is an adapted, continuous process on $(\Omega,\mathcal{F},\mathcal{F}_t,\mathbb{P})$ and if $D\in\mathbb{R}^d$ is a bounded closed subset of $\mathbb{R}^d$, then*

$$T=\inf\{t\ge 0 : X_t\in D\}$$

is a stopping time. If $X_0 \in D^c$*, then*

$$X_T 1_{\{T<+\infty\}} \in \partial D \ .$$

In particular, if $d = 1$ *and* b *is a real number, then*

$$T_b = \inf\{t \geq 0 : X_t = b\}$$

is a stopping time. In this case $\sup_{t\in[0,N]} X_t$ *is a random variable,*

$$\left\{ \sup_{t\in[0,N]} X_t < b \right\} = \{T_b > N\}$$

and

$$\left\{ \sup_{t\in[0,N]} X_t \geq b \right\} = \{T_b \leq N\} \ .$$

2.5.2 Technique of localization

The concept of stopping times provides us with a means of "localizing" quantities. Suppose $(X_t)_{t\geq 0}$ is a stochastic process, and T is a stopping time, then $X^T = (X_{t\wedge T})_{t\geq 0}$ is a stochastic process stopped at (random) time T, where

$$X_{t\wedge T}(\omega) = \begin{cases} X_t(\omega) & \text{if } \ t \leq T(\omega) \ ; \\ X_{T(\omega)}(\omega) & \text{if } \ t \geq T(\omega) \ . \end{cases}$$

Another interesting stopped process at random time T associated with X is the process $X1_{[0,T]}$ which is by definition

$$\begin{aligned} \left(X1_{[0,T]}\right)_t(\omega) &= X_t 1_{\{t\leq T\}}(\omega) \\ &= \begin{cases} X_t(\omega) & \text{if } \ t \leq T(\omega) \ ; \\ 0 & \text{if } \ t > T(\omega) \ . \end{cases} \end{aligned}$$

It is obvious that

$$X_t^T = X_t 1_{\{t\leq T\}} + X_T 1_{\{t>T\}} \ .$$

If $(X_t)_{t\geq 0}$ is adapted to the filtration $(\mathcal{F}_t)_{t\geq 0}$, so are the process $(X_{t\wedge T})_{t\geq 0}$ stopped at stopping time T and $X_t 1_{\{t\leq T\}}$.

Definition 2.5.7 *An adapted stochastic process* $X = (X_t)_{t\geq 0}$ *on the filtered probability space* $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ *is called a* local martingale *if there is an increasing family* $\{T_n\}$ *of finite stopping times such that*

$$T_n \uparrow +\infty \quad \text{as } n \to +\infty$$

and such that for each n*,* $(X_{t\wedge T_n})_{t\geq 0}$ *is a martingale.*

Similarly, we may define local super- or sub-martingales etc.

Chapter 3

Brownian motion

In this

3.1 Construction of Brownian motion

Brownian motion is a mathematical model of random movements observed by botanist Robert Brown.

Definition 3.1.1 *A stochastic process $B = (B_t)_{t\geq 0}$ on a probability space $(\Omega, \mathcal{F}, P)$ taking values in $\mathbb{R}^d$ is called a Brownian motion (BM) in $\mathbb{R}^d$, if*

1. *$(B_t)_{t\geq 0}$ possesses independent increments: for any $0 \leq t_0 < t_1 < \cdots < t_n$ random variables*

$$B_{t_0}, \;\; B_{t_1} - B_{t_0}, \;\cdots\;, \;\; B_{t_n} - B_{t_{n-1}}$$

   are independent.

2. *For any $t > s \geq 0$, random variable $B_t - B_s$ has a normal distribution $N(0, t-s)$, that is, $B_t - B_s$ has pdf (probability density function)*

$$p(t-s, x) = \frac{1}{(2\pi(t-s))^{d/2}} e^{-\frac{|x|^2}{2(t-s)}} \;\; ; \quad x \in \mathbb{R}^d \; .$$

   In other words

$$P\{B_t - B_s \in dx\} = p(t-s, x) dx \; .$$

3. *Almost all sample paths of $(B_t)_{t\geq 0}$ are continuous.*

If, in addition, $P\{B_0 = x\} = 1$ where $x \in \mathbb{R}^d$, then we say $(B_t)_{t\geq 0}$ is a Brownian motion starting at x. $P\{B_0 = 0\} = 1$ where 0 is the origin of $\mathbb{R}^d$, then we say $(B_t)_{t\geq 0}$ is a standard Brownian motion.

We will see that the condition 3) is not nontrivial, which ensure that many interesting functionals of Brownian motion are indeed random variables.

Let $p(t,x,y) = p(t, x-y)$, and define foe every $t > 0$

$$P_t f(x) = \int_{\mathbb{R}^d} f(y)p(t,x,y)\mathrm{d}y \qquad \forall f \in C_b(\mathbb{R}^d) \ .$$

Since

$$p(t+s,x,y) = \int_{\mathbb{R}^d} p(t,x,z)p(s,z,y)dz$$

therefore $(P_t)_{t\geq 0}$ is a semigroup on $C_b(\mathbb{R}^d)$. $(P_t)_{t\geq 0}$ is called the heat semigroup in $\mathbb{R}^d$: if $f \in C_b^2(\mathbb{R}^d)$, then $u(t,x) = (P_t f)(x)$ solves the heat equation

$$\left(\frac{1}{2}\Delta + \frac{\partial}{\partial t}\right) u(t,x) = 0 \ ; \quad u(0,\cdot) = f \ ,$$

where $\Delta = \sum_i \frac{\partial^2}{\partial x_i^2}$ is the Laplace operator.

The connection between Brownian motion and the Laplace operator Δ (hence the harmonic analysis) is demonstrated through the following identity:

$$\begin{aligned} \left(P_t f\right)(x) &= E\left(f(B_t + x)\right) \\ &= \frac{1}{(2\pi t)^{d/2}} \int_{\mathbb{R}^d} f(y) \mathrm{e}^{-\frac{|y-x|^2}{2t}} \mathrm{d}y \end{aligned}$$

where B_t is a standard Brownian motion.

Example 3.1.2 *If $B = (B_t)_{t\geq 0}$ is a BM in $\mathbb{R}$, then*

$$E|B_t - B_s|^p = c_p|t-s|^{p/2} \quad \textit{for all } s,t \geq 0 \tag{3.1}$$

for $p \geq 0$, where c_p is a constant depending only on p. Indeed

$$E|B_t - B_s|^p = \frac{1}{\sqrt{2\pi|t-s|}} \int_{\mathbb{R}} |x|^p \exp\left(-\frac{|x|^2}{2|t-s|}\right) dx \ .$$

Making change of variable

$$\frac{x}{\sqrt{|t-s|}} = y \ ; \quad dx = \sqrt{|t-s|}dy$$

we thus have

$$\begin{aligned} E|B_t - B_s|^p &= \frac{(\sqrt{|t-s|})^p}{\sqrt{2\pi}} \int_{\mathbb{R}} |x|^p \exp\left(-\frac{|x|^2}{2}\right) dx \\ &= c_p |t-s|^{p/2} \end{aligned}$$

where

$$c_p = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}^d} |x|^p \exp\left(-\frac{|x|^2}{2}\right) dx \ .$$

(3.1) remains true for BM in $\mathbb{R}^d$ with a constant c_p depending on p and d.

Remark 3.1.3 *Since $B_t - B_s \sim N(0, t-s)$, it is an easy exercise to show that for every $n \in \mathbb{Z}^+$*

$$E(B_t - B_s)^{2n} = \frac{(2n)!}{2^n n!} |t-s|^n \ .$$

Let $B = (B_t)_{t \geq 0}$ be a standard BM in $\mathbb{R}$. Then B is a centered Gaussian process with co-variance function $C(s,t) = s \wedge t$. Indeed, any finite-dimensional distribution of B is Gaussian [Exercise], so that B is a centered Gaussian process, and its co-variance function (if $s < t$)

$$\begin{aligned} E(B_t B_s) &= E((B_t - B_s)B_s + B_s^2) \\ &= E((B_t - B_s)B_s) + EB_s^2 \\ &= E(B_t - B_s)EB_s + EB_s^2 \\ &= s \ . \end{aligned}$$

Theorem 3.1.4 *(N. Wiener) There is a standard Brownian motion in $\mathbb{R}^d$.*

Proof. We may assume that $d = 1$, the proof in higher dimension is similar. Observe that a BM (B_t) must be a Gaussian process (i.e. a process whose finite-dimensional distributions are Gaussian distributions) with mean zero and variance function $E(B_t B_s) = s \wedge t$. Therefore we may first construct a Gaussian process (X_t) such that $EX_t = 0$ and $E(X_t X_s) = s \wedge t$ on some completed probability space $(\Omega, \mathcal{F}, P)$. It can be verified that $(X_t)_{t \geq 0}$ satisfies all conditions in the definition of BM, except the continuity of its sample paths. The Gaussian process (X_t) may be not continuous, we thus need to modify the construction of X_t to make it to be continuous. Let $D = \{\frac{j}{2^n} : j \in \mathbb{Z}^+, \, n \in \mathbb{N}\}$ the dyadic real numbers. The important fact is that D is dense in $\mathbb{R}^+$. Define

$$H = \bigcup_{N=1}^{\infty} \bigcap_{l=1}^{\infty} \bigcup_{n=l}^{\infty} \bigcup_{j=1}^{N2^n} \left(\left| X_{\frac{j}{2^n}} - X_{\frac{j-1}{2^n}} \right| \geq \frac{1}{2^{n/8}} \right) \ .$$

Let, for fixed N,

$$A_l = \bigcup_{n=l}^{\infty} \bigcup_{j=1}^{N2^n} \left(\left| X_{\frac{j}{2^n}} - X_{\frac{j-1}{2^n}} \right| \geq \frac{1}{2^{n/8}} \right) .$$

We are going to show that each $\bigcap_{l=1}^{\infty} A_l$ has probability zero, and therefore as a sum of countable many events with probability zero, $P(H) = 0$. Since

$$\begin{aligned}
& P\left\{ \bigcup_{j=1}^{N2^n} \left(\left| X_{\frac{j}{2^n}} - X_{\frac{j-1}{2^n}} \right| \geq \frac{1}{2^{n/8}} \right) \right\} \\
\leq \; & \sum_{j=1}^{N2^n} P\left(\left| X_{\frac{j}{2^n}} - X_{\frac{j-1}{2^n}} \right| \geq \frac{1}{2^{n/8}} \right) \\
= \; & N2^n P\left(\left| X_{\frac{1}{2^n}} \right| \geq \frac{1}{2^{n/8}} \right) \\
\leq \; & N2^n \left(2^{n/8}\right)^4 E\left| X_{\frac{1}{2^n}} \right|^4 \\
= \; & \left(2^{n/8}\right)^4 N2^n 3 \left(\frac{1}{2^n}\right)^2 \\
= \; & 3N \frac{1}{2^{n/2}}
\end{aligned}$$

so that

$$\begin{aligned}
P(A_l) \; & \leq \; \sum_{n=l}^{\infty} P\left\{ \bigcup_{j=1}^{N2^n} \left(\left| X_{\frac{j}{2^n}} - X_{\frac{j-1}{2^n}} \right| \geq \frac{1}{2^{n/8}} \right) \right\} \\
& \leq \; 3N \sum_{n=l}^{\infty} \frac{1}{2^{n/2}} \\
& = \; \frac{3N\sqrt{2}}{\sqrt{2}-1} \frac{1}{\left(\sqrt{2}\right)^l} .
\end{aligned}$$

Therefore

$$\begin{aligned}
P\left(\bigcap_{l=1}^{\infty} A_l \right) \; & = \; \lim_{n\to\infty} P\left\{A_l\right\} \\
& \leq \; \frac{3N\sqrt{2}}{\sqrt{2}-1} \lim_{n\to\infty} \frac{1}{\left(\sqrt{2}\right)^l} \\
& = \; 0 .
\end{aligned}$$

It follows that $P(H) = 0$, thus $P(H^c) = 1$. On the other hand, by De Morgan law

$$H^c = \bigcap_{N=1}^{\infty} \bigcup_{l=1}^{\infty} \bigcap_{n=l}^{\infty} \bigcap_{j=1}^{N2^n} \left\{ \omega : \left| X_{\frac{j}{2^n}}(\omega) - X_{\frac{j-1}{2^n}}(\omega) \right| < \frac{1}{2^{n/8}} \right\}$$

and thus, if $\omega \in H^c$, then for any N, there is an l such that for any $n > l$ and for all $j = 1, \cdots, N2^n$ we have

$$\left| X_{\frac{j}{2^n}}(\omega) - X_{\frac{j-1}{2^n}}(\omega) \right| < \frac{1}{2^{n/8}} .$$

We may show thus that for any $\omega \in H^c$ and $t \geq 0$ the limit of $X_s(\omega)$ exists as $s \to t$ along the dyadic numbers, i.e. as $s \to t$ and $s \in D$. Moreover, D is dense in $[0, \infty)$, thus for any $t \in [0, \infty)$ we may define

$$B_t(\omega) = \lim_{s \in D \to t} X_s(\omega) \quad \text{if } \omega \in H^c$$

otherwise if $\omega \in H$ we set $B_t(\omega) = 0$. By definition, $(B_t)_{t \geq 0}$ is a continuous process which coincides with X_t on H^c when $t \in D$. It remains to verify that $(B_t)_{t \geq 0}$ is a Brownian motion in $\mathbb{R}$ as an exercise. ∎

3.1.1 Scaling properties

Let $B = (B_t)_{t \geq 0}$ be a standard BM in $\mathbb{R}^d$. By definition, the distribution of BM $B = (B_t)_{t \geq 0}$ is stationary, so that for any fixed time S, $\tilde{B}_t = B_{t+S} - B_S$ is again a standard Brownian motion. This statement is true indeed for any finite stopping time S, see section 1.2.3.

Lemma 3.1.5 *(Scaling invariance, self-similarity) For any real number* $\lambda \neq 0$

$$M_t \equiv \lambda B_{t/\lambda^2}$$

is a standard BM in $\mathbb{R}^d$.

This statement follows directly from the definition of BM. In particular, $(-B_t)_{t \geq 0}$ is also a standard BM, so that $(-B_t)_{t \geq 0}$ and $(B_t)_{t \geq 0}$ have the same distribution.

Lemma 3.1.6 *If U is an $d \times d$ orthonormal matrix, then $UB = (UB_t)_{t \geq 0}$ is a standard BM in $\mathbb{R}^d$. That is, BM is invariant under the action of orthogonal group of $\mathbb{R}^d$.*

This lemma is an easy corollary of the invariance property of Gaussian distributions under the orthogonal group action.

Lemma 3.1.7 *Let $B = (B_t)_{t\geq 0}$ be a standard BM in $\mathbb{R}$, and define*

$$M_0 = 0 \ , \quad M_t = tB_{1/t} \quad \text{for } t > 0$$

is a standard BM in $\mathbb{R}$.

Proof. Obviously M_t possesses normal distribution with mean zero, and

$$\begin{aligned} E\left(M_t M_s\right) &= tsE\left(B_{1/t}B_{1/s}\right) \\ &= ts\left(\frac{1}{t}\wedge\frac{1}{s}\right) = s\wedge t \end{aligned}$$

so that (M_t) is a centered Gaussian process with co-variance function $s \wedge t$. Moreover $t \to M_t$ is continuous for $t > 0$. To see the continuity of M_t at $t = 0$, we use the fact that

$$\lim_{t\to\infty} \frac{B_t}{t} = 0$$

which is the law of large numbers for BM. We will not prove this here, but see the remark below. ■

Remark 3.1.8 *To convince yourself why the law of large numbers for BM is true, we may look at a special way $t \to \infty$ through natural numbers, namely*

$$\lim_{n\to\infty} \frac{B_n}{n} = \lim_{n\to\infty} \frac{X_1 + \cdots + X_n}{n}$$

where $X_i = B_i - B_{i-1}$. Notice that (X_i) is a sequence of independent random variables with identical distribution $N(0,1)$, so that by the strong law of large numbers

$$\frac{X_1 + \cdots + X_n}{n} \to EX_1 = 0 \quad \textit{almost surely.}$$

In order to handle the general case $t \geq 0$, we may write $t = [t] + r_t$ where $[t]$ is the integer part of t and $r_t \in [0,1)$. Then

$$\frac{B_t}{t} = \frac{B_t - B_{[t]}}{t} + \frac{[t]}{t}\frac{B_{[t]}}{[t]}$$

the second term tends to 0 since as $t \to \infty$, $\frac{[t]}{t} \to 1$ and $\frac{B_{[t]}}{[t]} \to 0$. To see why

$$\frac{B_t - B_{[t]}}{t} \to 0$$

as $t \to \infty$, we need the following Gaussian tail estimate for BM (see section 1.2.3 below)

$$\begin{aligned} P\left\{\omega : \sup_{t\in[0,T]} |B_t(\omega)| \geq R\right\} &= 2\sqrt{\frac{2}{\pi}}\int_{R/\sqrt{T}}^{\infty} e^{-x^2/2}dx \\ &\leq 2\exp\left(-\frac{R^2}{2T}\right) \qquad \textit{for all } R > 0\ . \end{aligned}$$

It follows that for any $\varepsilon > 0$

$$\sum_{n=0}^{\infty} P\left\{\omega : \sup_{t\in[n,n+1]} \frac{|B_t(\omega) - B_n(\omega)|}{n} \geq \varepsilon\right\} < \infty$$

and thus by the Borel-Cantelli lemma

$$\overline{\lim_{n\to\infty}} \sup_{t\in[n,n+1]} \left|\frac{B_t}{t} - \frac{B_n}{n}\right| = 0 \qquad \textit{almost surely.}$$

For more detail, see D. Stroock: Probability Theory: An Analytic View, page 180-181.

3.1.2 Markov property and finite-dimensional distributions

Let $X = (X_t)_{t\geq 0}$ be a stochastic process on $(\Omega, \mathcal{F}, P)$. For every $t \geq 0$, set

$$\mathcal{F}_t^0 = \sigma\{X_s : s \leq t\}$$

which is the smallest σ-algebra with to which every X_s (where $s \leq t$) are measurable. In particular, for each $t \geq 0$, $X_t \in \mathcal{F}_t^0$ and in this sense we say $(X_t)_{t\geq 0}$ is adapted to the filtration $\{\mathcal{F}_t^0\}$. $\{\mathcal{F}_t^0\}_{t\geq 0}$ is called the filtration generated by $X = (X_t)_{t\geq 0}$.

In this section, we use $(\mathcal{F}_t^0)_{t\geq 0}$ to denote the filtration generated by a standard Brownian motion $(B_t)_{t\geq 0}$, and let $\mathcal{F}_\infty^0 = \cup_{t\geq 0}\mathcal{F}_t^0$.

Lemma 3.1.9 *For any $t > s \geq 0$, the increment $B_t - B_s$ is independent of $\mathcal{F}_s^0$.*

Recall that

$$p(t,x) = \frac{1}{(2\pi t)^{d/2}}\mathrm{e}^{-\frac{|x|^2}{2t}}$$

in $\mathbb{R}^d$, and $(P_t)_{t\geq 0}$ the heat semigroup

$$P_t f(x) = \int_{\mathbb{R}^d} f(y)p(t, x-y)\mathrm{d}y$$

for every $t > 0$.

Lemma 3.1.10 *If $t > s$, then the joint distribution of B_s and B_t is given by*

$$P\{B_s \in dx, \ B_t \in dy\} = p(s,x)p(t-s, y-x)dxdy \ .$$

Indeed, since B_s and $B_t - B_s$ are independent, so that $(B_s, B_t - B_s)$ has a pdf

$$p(s, x_1)p(t-s, x_2)$$

thus, for any bounded Borel measurable function f

$$\begin{aligned} Ef(B_s, B_t) &= Ef(B_s, B_t - B_s + B_s) \\ &= \iint f(x_1, x_2 + x_1)p(s, x_1)p(t-s, x_2)\mathrm{d}x_1\mathrm{d}x_2 \ . \end{aligned}$$

Making change of variables $x_1 = x$ and $x_2 + x_1 = y$ in the last double integral, the induced Jacobi is 1 so that $\mathrm{d}x_1\mathrm{d}x_2 = \mathrm{d}x\mathrm{d}y$ (as measures), and therefore

$$Ef(B_s, B_t) = \iint f(x,y)p(s,x)p(t-s, y-x)\mathrm{d}x\mathrm{d}y$$

which implies that the pdf of (B_s, B_t) is $p(s,x)p(t-s, y-x)$.

Theorem 3.1.11 *Let $t > s$, and f a bounded Borel measurable function. Then*

$$E\left\{f(B_t)|\mathcal{F}_s^0\right\} = P_{t-s}f(B_s) \qquad a.s. \tag{3.2}$$

where $(P_t)_{t>0}$ is the heat semigroup. In particular

$$E\left\{f(B_t)|\mathcal{F}_s^0\right\} = E\left\{f(B_t)|B_s\right\}$$

which is called Markov property, and $E\left\{f(B_t)|\mathcal{F}_s^0\right\}$ equals $F(B_s)$ where

$$F(x) = P_{t-s}f(x) \equiv \frac{1}{(2\pi(t-s))^{d/2}} \int_{\mathbb{R}^d} f(y) e^{-\frac{|x-y|^2}{2(t-s)}} dy \ .$$

Proof. First we show that

$$E\left\{f(B_t)|\mathcal{F}_s^0\right\} = E\left\{f(B_t)|B_s\right\}$$

which is called Markov property of $(B_t)_{t\geq 0}$. Clearly we only need to prove this for bounded continuous (and smooth) function f. For such a function, we can show that

$$f(x+y) = \lim_{n\to\infty} \sum_{k=1}^{N_n} f_{n_k}(x) g_{n_k}(y) \ .$$

for some functions f_{n_k}, g_{n_k} (for example, taking the Taylor expansion of $f(x+y)$). Hence

$$\begin{aligned}
E\left\{f(B_t)|\mathcal{F}_s^0\right\} &= E\left\{f(B_s+B_t-B_s)|\mathcal{F}_s^0\right\} \\
&= \lim_{n\to\infty}\sum_{k=1}^{N_n} E\left\{f_{n_k}(B_s)g_{n_k}(B_t-B_s)|\mathcal{F}_s^0\right\} \\
&= \lim_{n\to\infty}\sum_{k=1}^{N_n} E\left\{g_{n_k}(B_t-B_s)|\mathcal{F}_s^0\right\} f_{n_k}(B_s) \\
&= \lim_{n\to\infty}\sum_{k=1}^{N_n} f_{n_k}(B_s)E\left\{f_{n_k}(B_t-B_s)\right\} \\
&= \lim_{n\to\infty}\sum_{k=1}^{N_n} f_{n_k}(B_s)\int_{\mathbb{R}^d} f_{n_k}(z)p(t-s,z)dz
\end{aligned}$$

which depends only on B_s, denoted by $F(B_s)$. In particular

$$E\left\{f(B_t)|\mathcal{F}_s^0\right\} = E\left\{f(B_t)|B_s\right\} = F(B_s)\ .$$

To compute the conditional expectation $E\{f(B_t)|B_s\}$, we use the fact that the pdf of (B_s,B_t) is

$$p(s,x)p(t-s,y-x)$$

so that

$$\begin{aligned}
E\left\{1_A(B_s)f(B_t)\right\} &= \int\int 1_A(x)f(y)p(s,x)p(t-s,y-x)\mathrm{d}x\mathrm{d}y \\
&= \int 1_A(x)P_{t-s}f(x)p(s,x)\mathrm{d}x \\
&= E\left\{1_A(B_s)P_{t-s}f(B_s)\right\}
\end{aligned}$$

as

$$P_{t-s}f(x) = \int f(y)p(t-s,y-x)\mathrm{d}y\ .$$

Since $P_{t-s}f(B_s)$ is a function of B_s so that

$$E\left(f(B_t)|B_s\right) = P_{t-s}f(B_s)\ .$$

■

The family of finite-dimensional distributions of BM can be calculated in terms of the Gaussian density function $p(t,x)$.

Proposition 3.1.12 *For any $0 < t_1 < t_2 < \cdots < t_n$, the ($\mathbb{R}^{n\times d}$-valued) random variable $(B_{t_1}, \cdots, B_{t_n})$ has a pdf*

$$p(t_1, x_1)p(t_2 - t_1, x_2 - x_1)\cdots p(t_n - t_{n-1}, x_n - x_{n-1})$$

where

$$p(t,x) = \frac{1}{(2\pi t)^{d/2}} e^{-|x|^2/(2t)}$$

is a standard Gaussian pdf in $\mathbb{R}^d$. That is, the joint distribution of $(B_{t_1}, \cdots, B_{t_n})$ is given by

$$\begin{aligned} & P\{B_{t_1} \in dx_1, \ \cdots, \ \ B_{t_n} \in dx_n\} \\ = \ & p(t_1, x_1)p(t_2 - t_1, x_2 - x_1)\cdots p(t_n - t_{n-1}, x_n - x_{n-1})dx_1 \cdots dx_n \end{aligned} \tag{3.3}$$

Proof. Let f be a bounded, continuous function. We want to calculate

$$E\left(f(B_{t_1}, \cdots, B_{t_n})\right) \ .$$

One can use the fact that $B_{t_1}, B_{t_2} - B_{t_1}, \cdots, B_{t_n} - B_{t_{n-1}}$ are independent, and has the joint distribution with pdf

$$p(t_1, z_1)p(t_2 - t_1, z_2)\cdots p(t_n - t_{n-1}, z_n) \ .$$

(3.3) follows after change of variables. Below we present an induction argument which uses only the Markov property. Indeed, by the Markov property

$$\begin{aligned} & E\left(f_1(B_{t_1})\cdots f_n(B_{t_n})\right) \\ = \ & E\left\{E\left(f_1(B_{t_1})\cdots f_n(B_{t_n})|\mathcal{F}^0_{t_{n-1}}\right)\right\} \\ = \ & E\left\{f_1(B_{t_1})\cdots f_{n-1}(B_{t_{n-1}})E\left(f_n(B_{t_n})|\mathcal{F}^0_{t_{n-1}}\right)\right\} \\ = \ & E\left\{f_1(B_{t_1})\cdots f_{n-1}(B_{t_{n-1}})\left(P_{t_n - t_{n-1}} f_n\right)(B_{t_{n-1}})\right\} \\ = \ & E\left\{f_1(B_{t_1})\cdots f_{n-2}(B_{t_{n-2}})\left(f_{n-1}P_{t_n - t_{n-1}} f_n\right)(B_{t_{n-1}})\right\} \end{aligned}$$

which reduces the number of times t_i to $n-1$, so the conclusion follows from the induction immediately. ■

Corollary 3.1.13 *Let $B_t = (B^1_t, \cdots, B^d_t)$ be a d-dimensional standard Brownian motion. Then for each j, B^j_t is a standard BM in $\mathbb{R}$, and $(B^j_t)_{t\geq 0}$ ($j = 1, \cdots, d$) are mutually independent.*

Therefore a d-dimensional BM is d independent copies of BM in $\mathbb{R}$.

3.1.3 The reflection principle

Brownian motion starts afresh at a stopping time, i.e. the Markov property for Brownian motion remains true at stopping times. Therefore Brownian motion possesses the *strong Markov property*, a very important property which had been used by Paul Lévy in the form of the *reflection principle*, long before the concept of strong Markov property had been properly defined. We will exhibit this principle by computing the distribution of the running maximum of a Brownian motion.

In many applications, especially in statistics, we would like to estimate distributions of running maxima of a stochastic process. For Brownian motion $B = (B_t)_{t\geq 0}$, the distribution of $\sup_{s\in[0,t]} B_s$ can be derived by means of the reflection principle which will be made rigorous in Chapter .

Let $B = (B_t)_{t\geq 0}$ be a standard Brownian motion on $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ in $\mathbb{R}$. Let $b > 0$ and $b > a$, and let

$$T_b = \inf\{t > 0 : B_t = b\} \ .$$

Then T_b is a stopping time, and the Brownian motion starts afresh as a standard Brownian motion after hitting b, and therefore

$$\begin{aligned} P\left\{\sup_{s\in[0,t]} B_s \geq b, B_t \leq a\right\} &= P\left\{\sup_{s\in[0,t]} B_s \geq b, B_t \geq 2b - a\right\} \\ &= P\left\{B_t \geq 2b - a\right\} \end{aligned}$$

where the first equality follows from the "fact" that the Brownian motion starting at T_b (in position b): $B_{T_b} = b$, runs afresh like a Brownian motion, so that it moves with equal probability about the line $y = b$. The second equality follows from $2b - a = b + (b - a) > b$.

The above equation may be written as

$$\begin{aligned} P\left\{T_b \leq t, B_t \leq a\right\} &= P\left\{T_b \leq t, B_t \geq 2b - a\right\} \\ &= P\left\{B_t \geq 2b - a\right\} \ , \end{aligned}$$

which can be justified by the *strong Markov property* of Brownian motion, a topic that will not pursue here. Therefore

$$P\left\{\sup_{s\in[0,t]} B_s \geq b, B_t \leq a\right\} = \frac{1}{\sqrt{2\pi t}} \int_{2b-a}^{+\infty} e^{-\frac{x^2}{2t}} \mathrm{d}t \ ,$$

which gives us the joint distribution of a Brownian motion and its maximum at a fixed time t. By differentiating in a and in b we conclude the following

Theorem 3.1.14 *Let $B = (B_t)_{t\geq 0}$ be a standard BM in $\mathbb{R}$, and let $t > 0$. Then the pdf of the joint distribution of random variables $(M_t = \sup_{s\in[0,t]} B_s, B_t)$ is given as*

$$P\{M_t \in db, B_t \in da\} = \frac{2(2b-a)}{\sqrt{2\pi t^3}} \exp\left\{-\frac{(2b-a)^2}{2t}\right\} da\, db$$

over the region $\{(b,a) : a \leq b, b \geq 0\}$ in $\mathbb{R}^2$.

In particular, for any $b > 0$,

$$\begin{aligned} P\left\{\sup_{s\in[0,t]} B_s \geq b\right\} &= P\{T_b \leq t\} \\ &= \frac{2}{\sqrt{2\pi t^3}} \int\int_{\{a\leq c, c\geq b\}} (2c-a)\exp\left\{-\frac{(2c-a)^2}{2t}\right\} \mathrm{d}a\mathrm{d}c \\ &= \frac{2}{\sqrt{2\pi t^3}} \int_b^{+\infty}\int_{-\infty}^{c} (2c-a)\exp\left\{-\frac{(2c-a)^2}{2t}\right\} \mathrm{d}a\mathrm{d}c \\ &= \frac{2}{\sqrt{2\pi t^3}} \int_b^{+\infty}\int_{c}^{+\infty} x\exp\left\{-\frac{x^2}{2t}\right\} \mathrm{d}x\mathrm{d}c \\ &= \frac{2}{\sqrt{2\pi t}} \int_b^{+\infty} \exp\left(-\frac{x^2}{2t}\right) \mathrm{d}x \end{aligned}$$

which is the exact distribution function of $\sup_{s\in[0,t]} B_s$ (the stopping time T_b) and leads to an exact formula for the tail probability of the Brownian motion.

3.1.4 Martingale property

Let $X = (X_t)_{t\geq 0}$ be a stochastic process adapted to a filtration $(\mathcal{F}_t)_{t\geq 0}$ such that each random variable X_t is integrable. Then $X = (X_t)_{t\geq 0}$ is a martingale (resp. super-martingale; resp. sub-martingale) if for any $t > s$ we have

$$E(X_t|\mathcal{F}_s) = X_s$$

resp.

$$E(X_t|\mathcal{F}_s) \leq X_s \qquad \text{(super-martingale)}$$

resp.

$$E(X_t|\mathcal{F}_s) \geq X_s \qquad \text{(sub-martingale)} \quad .$$

If $X = (X_t)_{t\geq 0}$ is a martingale (resp. super-martingale; resp. sub-martingale), then $t \to E(X_t)$ is a constant function $E(X_0)$ (resp. decreasing function; resp. increasing function).

Let $B = (B_t^i)_{t\geq 0}$ $(i = 1, \cdots, d)$ be a standard BM in $\mathbb{R}^d$, with its generated filtration $(\mathcal{F}_t^0)_{t\geq 0}$.

Proposition 3.1.15 *1) Each B_t is p-th integrable for any $p > 0$, and for $t > s$*

$$E(|B_t - B_s|^p) = c_{p,d}|t-s|^{p/2} \ . \tag{3.4}$$

2) $(B_t)_{t\geq 0}$ is a continuous, square-integrable martingale.

3) For each pair i, j, $M_t = B_t^i B_t^j - \delta_{ij} t$ is a continuous martingale.

Proof. The first part was proved before. Since $B_t - B_s$ is independent of $\mathcal{F}_s^0$ when $t > s$ we thus have

$$E(B_t - B_s|\mathcal{F}_s^0) = E(B_t - B_s) = 0$$

so that

$$E(B_t|\mathcal{F}_s^0) = E(B_s|\mathcal{F}_s^0) = B_s$$

so that $(B_t)_{t\geq 0}$ is a continuous martingale.

Obviously we only need to show 3) for BM in $\mathbb{R}$. In this case

$$\begin{aligned} E(B_t^2 - B_s^2|\mathcal{F}_s^0) &= E((B_t - B_s)^2 |\mathcal{F}_s^0) \\ &\quad + E(2B_s (B_t - B_s) |\mathcal{F}_s^0) \\ &= E((B_t - B_s)^2) + 2B_s E((B_t - B_s) |\mathcal{F}_s^0) \\ &= E (B_t - B_s)^2 \\ &= t - s \end{aligned}$$

so that

$$\begin{aligned} E(B_t^2 - t|\mathcal{F}_s^0) &= \mathbb{E}(B_s^2 - s|\mathcal{F}_s^0) \\ &= B_s^2 - s \end{aligned}$$

which shows that $B_t^2 - t$ is a martingale. ■

Theorem 3.1.16 *Let $B = (B_t)_{t\geq 0}$ be a continuous stochastic process in $\mathbb{R}$ such that $B_0 = 0$. Then $(B_t)_{t\geq 0}$ is a standard BM in $\mathbb{R}$, if and only if for any $\xi \in \mathbb{R}$ and $t > s$*

$$E\left\{\exp\left(i\langle \xi, B_t - B_s\rangle\right) |\mathcal{F}_s^0\right\} = \exp\left(-\frac{(t-s)|\xi|^2}{2}\right) \ . \tag{3.5}$$

Proof. We observe that (3.5) implies $B_t - B_s$ is independent of $\mathcal{F}_s^0$ and has normal distribution with variance $t-s$. Conversely, if $(B_t)_{t\geq 0}$ is a standard BM in $\mathbb{R}$, then $B_t - B_s$ is independent of $\mathcal{F}_s^0$, and $B_t - B_s$ has a normal distribution of mean zero and variance $(t-s)$, so that

$$\begin{aligned}
& E\left\{\exp\left(i\langle\xi, B_t - B_s\rangle\right)|\mathcal{F}_s^0\right\} \\
= \; & E\left\{\exp\left(i\langle\xi, B_t - B_s\rangle\right)\right\} \\
= \; & \frac{1}{\sqrt{2\pi(t-s)}}\int_{\mathbb{R}} e^{i\langle\xi,x\rangle - \frac{|x|^2}{2(t-s)}}\mathrm{d}x \\
= \; & \exp\left(-\frac{(t-s)|\xi|^2}{2}\right) .
\end{aligned}$$

■

Corollary 3.1.17 *Let (B_t) be a standard BM in $\mathbb{R}$. If $\xi \in \mathbb{R}$, then*

$$M_t \equiv \exp\left(i\langle\xi, B_t\rangle + \frac{|\xi|^2}{2}t\right)$$

is a martingale.

Remark 3.1.18 *Note that both sides of (3.5) are analytic in ξ so that the identity continues to hold for any complex vector ξ. In particular, by replacing ξ by $-i\xi$ we obtain that*

$$E\left\{\exp\left(\langle\xi, B_t - B_s\rangle\right)|\mathcal{F}_s^0\right\} = \exp\left(\frac{(t-s)|\xi|^2}{2}\right)$$

so that for any vector ξ

$$\exp\left(\langle\xi, B_t\rangle - \frac{|\xi|^2}{2}t\right)$$

is a continuous martingale. This statement will be extended to vector fields ξ in $\mathbb{R}$. The resulted identity is called Cameron-Martin formula.

BM is the basic example of Lévy processes: right continuous stochastic processes in $\mathbb{R}^d$ which possess stationary independent increments, and (3.5) is the Lévy-Khinchin formula for BM. In general if (X_t) is a Lévy process in $\mathbb{R}^d$, then

$$E\left\{\exp\left(i\langle\xi, X_t - X_s\rangle\right)|\mathcal{F}_s^0\right\} = \exp\left(\psi(\xi)(t-s)\right)$$

for $t > s$ and $\xi \in \mathbb{R}^d$, where

$$\begin{aligned}\psi(\xi) &= -\frac{1}{2}\langle AA^T\xi, \xi\rangle + i\langle b, \xi\rangle \\ &\quad + \int_{\mathbb{R}^d\setminus\{0\}} \left(e^{i\langle \xi, x\rangle} - 1 - i1_{\{|x|<1\}}\langle \xi, x\rangle\right) \nu(dx)\end{aligned}$$

for some $d \times r$ matrix A, vector b and Lévy measure $\nu(dx)$ of (X_t) which is a σ-finite measure on $\mathbb{R}^d \setminus \{0\}$ satisfying the following integrable condition

$$\int_{\mathbb{R}^d\setminus\{0\}} \frac{|x|^2}{1+|x|^2}\nu(dx) < +\infty \ . \tag{3.6}$$

3.2 Quadratic variational processes

As we have seen that, both (one-dimensional) Brownian motion B_t and $M_t \equiv B_t^2 - t$ are martingales, and thus

$$B_t^2 = M_t + A_t$$

where of course $A_t = t$. Therefore, the continuous sub-martingale B_t^2 is a *sum of a martingale and an adapted increasing process.* We will see this decomposition for B_t^2 is the key to establish Itô's integration theory.

Lemma 3.2.1 *Let*

$$D = \{0 = t_0 < t_1 < \cdots < t_n = t\}$$

be a finite partition of the interval $[0, t]$*, and let*

$$V_D = \sum_{l=1}^n |B_{t_l} - B_{t_{l-1}}|^2$$

the quadratic variation of B *over the partition* D*, which is a non-negative random variable. Then*

$$EV_D = t$$

and the variance of V_D

$$E\left\{(V_D - EV_D)^2\right\} = 2\sum_{l=1}^n (t_l - t_{l-1})^2 \ .$$

Proof. Indeed

$$
\begin{aligned}
EV_D &= \sum_{l=1}^{n} E|B_{t_l} - B_{t_{l-1}}|^2 \\
&= \sum_{l=1}^{n} (t_l - t_{l-1}) \\
&= t \ .
\end{aligned}
$$

To prove the second formula we proceed as the following

$$
\begin{aligned}
& E\left\{ (V_D - EV_D)^2 \right\} \\
= \ & E\left\{ \left(\sum_{l=1}^{n} |B_{t_l} - B_{t_{l-1}}|^2 - t \right)^2 \right\} \\
= \ & E\left\{ \left(\sum_{l=1}^{n} \left(|B_{t_l} - B_{t_{l-1}}|^2 - (t_l - t_{l-1}) \right) \right)^2 \right\} \\
= \ & \sum_{k,l=1}^{n} E\left\{ \left(|B_{t_k} - B_{t_{k-1}}|^2 - (t_k - t_{k-1}) \right) \left(|B_{t_l} - B_{t_{l-1}}|^2 - (t_l - t_{l-1}) \right) \right\} \\
= \ & \sum_{l=1}^{n} E\left\{ \left(|B_{t_l} - B_{t_{l-1}}|^2 - (t_l - t_{l-1}) \right)^2 \right\} \\
& + \sum_{k \neq l}^{n} E\left\{ \left(|B_{t_k} - B_{t_{k-1}}|^2 - (t_k - t_{k-1}) \right) \left(|B_{t_l} - B_{t_{l-1}}|^2 - (t_l - t_{l-1}) \right) \right\} \ .
\end{aligned}
$$

Since the increments over different intervals are independent, so that the expectation of each product in the last sum on the right-hand side equals the product of their expectations, which gives contribution zero, therefore

$$\begin{aligned}
& E\left\{(V_D - EV_D)^2\right\} \\
= & \sum_{l=1}^n E\left\{\left(|B_{t_l} - B_{t_{l-1}}|^2 - (t_l - t_{l-1})\right)^2\right\} \\
= & \sum_{l=1}^n E\left\{|B_{t_l} - B_{t_{l-1}}|^4 - 2(t_l - t_{l-1})|B_{t_l} - B_{t_{l-1}}|^2 + (t_l - t_{l-1})^2\right\} \\
= & \sum_{l=1}^n \left\{E|B_{t_l} - B_{t_{l-1}}|^4 - 2(t_l - t_{l-1})E|B_{t_l} - B_{t_{l-1}}|^2 + (t_l - t_{l-1})^2\right\} \\
= & 2\sum_{l=1}^n (t_l - t_{l-1})^2
\end{aligned}$$

where we have used the integral

$$E|B_{t_l} - B_{t_{l-1}}|^4 = 3(t_l - t_{l-1})^2 \ .$$

■

We are now in a position to prove the following

Theorem 3.2.2 *Let $B = (B_t)_{t\geq 0}$ be a standard BM in $\mathbb{R}$. Then*

$$\lim_{m(D)\to 0} \sum_l |B_{t_l} - B_{t_{l-1}}|^2 = t \qquad \text{in } L^2(\Omega, P)$$

for any t, where D runs over all finite partitions of interval $[0,t]$, and

$$m(D) = \max_l |t_l - t_{l-1}| \ .$$

Therefore

$$\lim_{m(D)\to 0} \sum_l |B_{t_l} - B_{t_{l-1}}|^2 = t \qquad \text{in probability.}$$

Proof. According to the previous lemma we have

$$\begin{aligned}
E\left|\sum_l |B_{t_l} - B_{t_{l-1}}|^2 - t\right|^2 &= E\left|V_D - E\left(V_D\right)\right|^2 \\
&= 2\sum_{l=1}^n (t_l - t_{l-1})^2 \\
&\leq 2m(D)\sum_{l=1}^n (t_l - t_{l-1}) \\
&= 2tm(D)
\end{aligned}$$

and therefore

$$\lim_{m(D)\to 0} E\left|\sum_l |B_{t_l} - B_{t_{l-1}}|^2 - t\right|^2 = 0 \ .$$

■

For good partitions the convergence in the above theorem takes place almost surely. To this end, we recall the Borel-Cantelli lemma: if $\sum_n^\infty P(A_n) < \infty$, then $\mathrm{limsup}_n A_n = 0$, where

$$\begin{aligned} \mathrm{limsup}_n A_n &= \bigcap_{m=1}^\infty \bigcup_{n=m}^\infty A_n \\ &= \{\omega \text{ belongs to infinitely many } A_n\} \ . \end{aligned}$$

If in addition $\{A_n\}$ are independent, then $\sum_n^\infty P(A_n) = \infty$ if and only if

$$P\left(\mathrm{limsup}_n A_n\right) = 1 \ .$$

Proposition 3.2.3 *Let $(B_t)_{t\geq 0}$ be a standard BM in $\mathbb{R}$. Then for any $t > 0$ we have*

$$\sum_{j=1}^{2^n} \left| B_{\frac{j}{2^n}t} - B_{\frac{j-1}{2^n}t}\right|^2 \to t \quad a.s. \tag{3.7}$$

as $n \to \infty$.

Proof. Let D_n be the dyadic partition of $[0, t]$

$$D_n = \{0 = \frac{0}{2^n}t < \frac{1}{2^n}t < \cdots < \frac{2^n}{2^n}t = t\} \ .$$

and V_n denote V_{D_n}. Then, according to Lemma 3.2.1, $EV_n = t$ and

$$\begin{aligned} E\left|V_n - EV_n\right|^2 &= 2\sum_{l=1}^{2^n}\left(\frac{l}{2^n}t - \frac{l-1}{2^n}t\right)^2 \\ &= 22^n\left(\frac{1}{2^n}t\right)^2 \\ &= \frac{1}{2^{n-1}}t^2 \ . \end{aligned}$$

Therefore, by Markov's inequality,

$$\begin{aligned} P\left\{|V_n - EV_n| \geq \frac{1}{n}\right\} &\leq n^2 E\left|V_n - EV_n\right|^2 \\ &= \frac{n^2}{2^{n-1}}t^2 \end{aligned}$$

so that

$$\sum_{n=1}^{\infty} P\left\{|V_n - EV_n| \geq \frac{1}{n}\right\} = t^2 \sum_{n=1}^{\infty} \frac{n^2}{2^{n-1}} < +\infty \ .$$

By the Borel-Cantelli lemma, it follows that $V_n \to t$ almost surely. ■

Remark 3.2.4 *Indeed the conclusion is true for monotone partitions. More precisely, for each n let*

$$D_n = \{0 = t_{n,0} < t_{1,n} < \cdots < t_{n_k,n} = t\}$$

be a finite partition of $[0,t]$. Suppose $D_{n+1} \supset D_n$ and

$$\lim_{n\to\infty} m(D_n) = \lim_{n\to\infty} \max |t_{n_i,n} - t_{n_{i-1},n}| = 0 \ .$$

Then

$$\sum_{i=1}^{n_k} \left| B_{t_{n_i,n}} - B_{t_{n_{i-1},n}} \right|^2 \to t \quad a.s. \tag{3.8}$$

as $n \to \infty$. Indeed, in this case, if we denote by M_n the right-hand side of (3.8), then $(M_n)_{n\geq 1}$ is a non-negative martingale with respect to the filtration $\mathcal{G}_n = \sigma\left\{B_{t_{n_i,n}} : i = 0, \cdots, n_k\right\}$, so (3.8) follows from the martingale convergence theorem.

It can be shown (not easy) that

$$\sup_D \sum_l |B_{t_l} - B_{t_{l-1}}|^p < \infty \quad \text{a.s.}$$

if $p > 2$, where sup is taken over all finite partitions of $[0,1]$, and

$$\sup_D \sum_l |B_{t_l} - B_{t_{l-1}}|^2 = \infty \quad \text{a.s.}$$

That is to say, Brownian motion has finite p-variation for any $p > 2$. Indeed almost all Brownian motion sample paths are α-Hölder continuous for any $\alpha < 1/2$ but not for $\alpha = 1/2$. It follows that almost all Brownian motion paths are nowhere differentiable. We will not go into a deep study about the sample paths of BM, which are not needed in order to develop Itô's calculus for Brownian motion.

Definition 3.2.5 *Let $p > 0$ be a constant. A path $f(t)$ in $\mathbb{R}^d$ [a function on $[0,T]$ valued in $\mathbb{R}^d$] is said to have finite p-variation on $[0,T]$, if*

$$\sup_D \sum_l |f(t_i) - f(t_{i-1})|^p < +\infty$$

where D runs over all finite partitions of $[0,T]$. $f(t)$ in $\mathbb{R}^d$ has finite (total) variation if it has finite 1-variation.

A function with finite variation must be a difference of two increasing functions. It particular, it has at most countably many discontinuous points.

A stochastic process $V = (V_t)_{t\geq 0}$ is called a *variational process*, if for almost all $\omega \in \Omega$, the sample path $t \to V_t(\omega)$ possesses finite variation on any finite interval. A Brownian motion is not a variational process.

If $M = (M_t)_{t\geq 0}$ is a continuous, square-integrable martingale, then $(M_t^2)_{t\geq 0}$ is no longer a martingale but a sub-martingale, except for the trivial case. As in the case of Brownian motion, the following limit

$$\langle M \rangle_t = \lim_{m(D)\to 0} \sum_l \left| M_{t_l} - M_{t_{l-1}} \right|^2$$

exists both in probability and in the L^2-norm, where the limit takes over all finite partitions D of the interval $[0,t]$. $\{\langle M \rangle_t\}_{t\geq 0}$ is called the (quadratic) variational process of $(M_t)_{t\geq 0}$, or simply *the bracket process* of $(M_t)_{t\geq 0}$. The quadratic variational process $t \to \langle M \rangle_t$ is an adapted, continuous, increasing stochastic process [and therefore has finite variation] with initial zero. The following theorem demonstrates the importance of $\langle M \rangle_t$.

Theorem 3.2.6 (The quadratic variational process) *Let $M = (M_t)_{t\geq 0}$ be a continuous, square-integrable martingale. Then $\langle M \rangle_t$ is the unique continuous, adapted and increasing process with initial zero, such that $M_t^2 - \langle M \rangle_t$ is a martingale.*

The process $\langle M \rangle$ is called the *quadratic variational process* associated with the martingale M. The theorem is a special case of the Doob-Meyer decomposition for sub-martingales: any sub-martingale can be decomposed into a sum of a martingale and a predictable, increasing process with initial value zero. The decomposition was conjectured by L. Doob, and proved by P. A. Meyer in the 60's, which opened the new era of stochastic calculus.

Remark 3.2.7 *If $M = (M_t)_{t\geq 0}$ is a continuous martingale, and $A = (A_t)_{t\geq 0}$ is an adapted, continuous and integrable increasing process, then*

$X = M + A$ is a continuous sub-martingale. The reverse statement is also true, that is the context of Doob-Meyer's decomposition theorem. Consider a sub-martingale in discrete-time: $X = (X_n)_{n\in\mathbb{Z}^+}$ with respect to the filtration $(\mathcal{F}_n)_{n\in\mathbb{Z}^+}$. An increasing sequence $(A_n)_{n\in\mathbb{Z}^+}$ may be defined by

$$\begin{aligned} A_0 &= 0\ ; \\ A_n &= A_{n-1} + E(X_n - X_{n-1}|\mathcal{F}_{n-1}), \quad n = 1, 2, \cdots . \end{aligned}$$

Then

1. *$E(X_n - X_{n-1}|\mathcal{F}_{n-1}) = E(X_n|\mathcal{F}_{n-1}) - X_{n-1} \geq 0$, $(A_n)_{n\in\mathbf{Z}^+}$ is increasing.*
2. *$A_n \in \mathcal{F}_{n-1}$, so that $(A_n)_{n\in\mathbf{Z}^+}$ is* predictable*!*
3. *By definition*
$$E(X_n - A_n|\mathcal{F}_{n-1}) = X_{n-1} - A_{n-1}, \quad n = 1, 2, \cdots ,$$
therefore $M_n = X_n - A_n$ is a martingale.

Theorem 3.2.8 *Let $(M_t)_{t\geq 0}$ and $(N_t)_{t\geq 0}$ be two continuous, square-integrable martingales, and let*

$$\langle M, N\rangle_t = \frac{1}{4}\left(\langle M + N\rangle_t - \langle M - N\rangle_t\right)$$

called the bracket process of M and N. Then $\langle M, N\rangle_t$ is the unique adapted, continuous, variational process with initial zero, such that $M_tN_t - \langle M, N\rangle_t$ is a martingale. Moreover

$$\lim_{m(D)\to 0} \sum_{l=1}^{n} (M_{t_l} - M_{t_{l-1}})(N_{t_l} - N_{t_{l-1}}) = \langle M, N\rangle_t \ , \quad \textit{in prob.} \tag{3.9}$$

where $D = \{0 = t_0 < \cdots < t_n = t\}$ and $m(D) = \max_l(t_l - t_{l-1})$.

If $(B_t)_{t\geq 0}$ is a Brownian motion in $\mathbb{R}^d$, then for any $f \in C_b^2(\mathbb{R}^d)$

$$M_t^f \equiv f(B_t) - f(B_0) - \int_0^t \frac{1}{2}\Delta f(B_s)ds$$

is a continuous martingale (with respect to the natural filtration generated by the Brownian motion $(B_t)_{t\geq 0}$), and

$$\langle M^f, M^g\rangle_t = \int_0^t \langle \nabla f, \nabla g\rangle(B_s)\mathrm{d}s \ .$$

These claims will be proven below after we have established Itô's lemma for Brownian motion.

Chapter 4

Itô's calculus

In this part we develop Itô's integration theory in a traditional way, that is, we first define stochastic integral $\int_0^t F_s dB_s$ for adapted simple processes $(F_t)_{t\geq 0}$, then extend the definition to a large class of integrands by exploiting the martingale characterization of Itô's integrals.

4.1 Introduction

Let $B = (B_t)_{t\geq 0}$ be a standard Brownian motion in $\mathbb{R}$ on a complete probability space $(\Omega, \mathcal{F}, P)$ and let $(\mathcal{F}_t^0)_{t\geq 0}$ be the filtration generated by $(B_t)_{t\geq 0}$, called the Brownian filtration. That is, for each $t \geq 0$

$$\mathcal{F}_t^0 = \sigma\{B_s \text{ for } s \leq t\}$$

which represents the history of the Brownian motion $B = (B_t)_{t\geq 0}$ up to time t.

We are going to define Itô's integrals of the following form

$$\int_0^t F_s \mathrm{d}B_s \quad \text{for} \quad t \geq 0$$

as a continuous stochastic process, where the integrand $F = (F_t)_{t\geq 0}$ is a stochastic process satisfying certain conditions that will be described later. For example, we would like to define integrals like

$$\int_0^t f(B_s)\mathrm{d}B_s$$

for Borel measurable functions f.

Since, for almost all $\omega \in \Omega$, the sample path of Brownian motion $t \to B_t(\omega)$ is nowhere differentiable, the obvious definition via Riemann sums

$$\sum_i F_{t_i^*}(B_{t_i} - B_{t_{i-1}})$$

does not work: the limit of Riemann sums does not exist. The limit exists however in a probability sense, if for any finite partition we properly choose $t_i^* \in [t_{i-1}, t_i]$ and if the integrand process $(F_t)_{t\geq 0}$ is adapted to the Brownian filtration $(\mathcal{F}_t^0)_{t\geq 0}$. That is to say, for every $t \geq 0$, F_t is measurable with respect to $\mathcal{F}_t^0$. This approach works because both $(B_t)_{t\geq 0}$ and $(B_t^2 - t)_{t\geq 0}$ are continuous martingales.

In summary, Itô's integral $\int_0^t F_s \mathrm{d}B_s$ of *an adapted process* $F = (F_t)_{t\geq 0}$ [such that F is measurable with respect to the product σ-algebra $\mathcal{B}([0,\infty)) \otimes \mathcal{F}_\infty^0$, a condition you are advised to forget at the first reading] with respect to the Brownian motion $B = (B_t)_{t\geq 0}$ may be simply defined to be the limit of special sort of Riemann sums:

$$\int_0^t F_s \mathrm{d}B_s = \lim_{m(D)\to 0} \sum_i F_{t_{i-1}}(B_{t_i} - B_{t_{i-1}})$$

where the limit takes place in L^2-sense [with respect to the product measure $P(\mathrm{d}\omega)\otimes \mathrm{d}t$: you are forgiven to ignore its meaning at this stage], over finite partitions

$$D = \{0 = t_0 < t_1 < \cdots < t_n = t\}$$

of $[0,t]$ so that $m(D) = \max_i(t_i - t_{i-1}) \to 0$. The reason to choose $F_{t_{i-1}}(B_{t_i} - B_{t_{i-1}})$ is the following: only with this choice

$$E\left(F_{t_{i-1}}(B_{t_i} - B_{t_{i-1}})\right) = 0 \tag{4.1}$$

and

$$E\left(F_{t_{i-1}}^2(B_{t_i} - B_{t_{i-1}})^2 - F_{t_{i-1}}^2(t_i - t_{i-1})\right) = 0\ . \tag{4.2}$$

It will become clear that, it is these important features that this sort of Riemann sums converge to a martingale! (4.1, 4.2) imply that both the Itô integral

$$\int_0^t F_s \mathrm{d}B_s$$

and

$$\left(\int_0^t F_s \mathrm{d}B_s\right)^2 - \int_0^t F_s^2 \mathrm{d}s$$

are martingales.

Exercise 4.1.1 *Prove equations (4.1) and (4.2).*

Indeed, since $F_{t_{i-1}}$ is $\mathcal{F}^0_{t_{i-1}}$-measurable, so that

$$\begin{aligned} E\left(F_{t_{i-1}}(B_{t_i}-B_{t_{i-1}})\right) &= E\left\{E\left(F_{t_{i-1}}(B_{t_i}-B_{t_{i-1}})\middle|\,\mathcal{F}^0_{t_{i-1}}\right)\right\} \\ &= E\left\{F_{t_{i-1}}E\left((B_{t_i}-B_{t_{i-1}})\middle|\,\mathcal{F}^0_{t_{i-1}}\right)\right\} \\ &= E\left(F_{t_{i-1}}E\left(B_{t_i}-B_{t_{i-1}}\right)\right) \\ &= 0\,. \end{aligned}$$

Similarly, since $B_t^2 - t$ is a martingale

$$\begin{aligned} & E\left((B_{t_i}-B_{t_{i-1}})^2-(t_i-t_{i-1})\middle|\,\mathcal{F}^0_{t_{i-1}}\right) \\ = \; & E\left(B_{t_i}^2-t_i\middle|\,\mathcal{F}^0_{t_{i-1}}\right)-2E\left(B_{t_{i-1}}B_{t_i}\middle|\,\mathcal{F}^0_{t_{i-1}}\right) \\ & +B_{t_{i-1}}^2+t_{i-1} \\ = \; & 2B_{t_{i-1}}^2-2B_{t_{i-1}}E\left(B_{t_i}\middle|\,\mathcal{F}^0_{t_{i-1}}\right) \\ = \; & 0 \end{aligned}$$

and therefore

$$\begin{aligned} & E\left(F_{t_{i-1}}^2(B_{t_i}-B_{t_{i-1}})^2-F_{t_{i-1}}^2(t_i-t_{i-1})\right) \\ = \; & E\left\{E\left(F_{t_{i-1}}^2(B_{t_i}-B_{t_{i-1}})^2-F_{t_{i-1}}^2(t_i-t_{i-1})\middle|\,\mathcal{F}^0_{t_{i-1}}\right)\right\} \\ = \; & E\left\{F_{t_{i-1}}^2E\left((B_{t_i}-B_{t_{i-1}})^2-F_{t_{i-1}}^2(t_i-t_{i-1})\middle|\,\mathcal{F}^0_{t_{i-1}}\right)\right\} \\ = \; & 0\,. \end{aligned}$$

Itô's integration theory can be established for a continuous, square-integrable martingale by the same approach. In fact, if $M=(M_t)_{t\geq 0}$ is a continuous, square-integrable martingale on a filtered probability space $(\Omega,\mathcal{F},\mathcal{F}_t,P)$ and $F=(F_t)_{t\geq 0}$ is an adapted stochastic process, then

$$\int_0^t F_s \mathrm{d}M_s = \lim_{m(D)\to 0}\sum_{l=1}^n F_{t_{l-1}}(M_{t_l}-M_{t_{l-1}})$$

exists under certain integrable conditions.

The Stratonovich integral $\int_0^t F_s \circ \mathrm{d}M_s$, which was discovered later than Itô's, is defined on the other hand by

$$\int_0^t F_s\circ \mathrm{d}M_s = \lim_{m(D)\to 0}\sum_{l=1}^n \frac{F_{t_{l-1}}+F_{t_l}}{2}(M_{t_l}-M_{t_{l-1}})$$

which in general is different from the Itô integral $\int_0^t F_s \mathrm{d}M_s$.

According to definition,

$$\begin{aligned}
\int_0^t M_s \mathrm{d}M_s &= \lim_{m(D)\to 0} \sum_{l=1}^n M_{t_{l-1}}(M_{t_l} - M_{t_{l-1}}) \\
&= \lim_{m(D)\to 0} \sum_{l=1}^n \left\{ -\frac{1}{2}(M_{t_l} - M_{t_{l-1}})^2 + \frac{1}{2}\left(M_{t_l}^2 - M_{t_{l-1}}^2\right) \right\} \\
&= -\frac{1}{2} \lim_{m(D)\to 0} \sum_{l=1}^n (M_{t_l} - M_{t_{l-1}})^2 + \frac{1}{2} \lim_{m(D)\to 0} \sum_{l=1}^n \left(M_{t_l}^2 - M_{t_{l-1}}^2\right) \\
&= -\frac{1}{2}\langle M\rangle_t + \frac{1}{2}(M_t^2 - M_0^2) \ .
\end{aligned}$$

That is

$$M_t^2 - M_0^2 = 2\int_0^t M_s \mathrm{d}M_s + \langle M\rangle_t \ .$$

On the other hand

$$\begin{aligned}
\int_0^t M_s \circ \mathrm{d}M_s &= \lim_{m(D)\to 0} \sum_{l=1}^n \frac{M_{t_l} + M_{t_{l-1}}}{2}(M_{t_l} - M_{t_{l-1}}) \\
&= \frac{1}{2} \lim_{m(D)\to 0} \sum_{l=1}^n (M_{t_l} - M_{t_{l-1}})^2 + \lim_{m(D)\to 0} \sum_{l=1}^n M_{t_{l-1}}(M_{t_l} - M_{t_{l-1}}) \\
&= \frac{1}{2}\langle M\rangle_t + \int_0^t M_s \mathrm{d}M_s \\
&= \frac{1}{2}\left(M_t^2 - M_0^2\right)
\end{aligned}$$

so that

$$M_t^2 - M_0^2 = 2\int_0^t M_s \circ \mathrm{d}M_s$$

which coincides with the fundamental theorem in Calculus.

In general, we have

Lemma 4.1.2 *Let N, M be two continuous, square-integrable martingales, and $F_t = N_t + A_t$ where A_t is an adapted process with finite variation. Define Stratonovich's integral*

$$\int_0^t F_s \circ dM_s = \lim_{m(D)\to 0} \sum_l \frac{F_{t_l} + F_{t_{l-1}}}{2}(M_{t_l} - M_{t_{l-1}})$$

Then

$$\int_0^t F_s \circ dM_s = \int_0^t F_s dM_s + \frac{1}{2}\langle N, M\rangle_t$$

Indeed,

$$\begin{aligned}
\int_0^t F_s \circ \mathrm{d}M_s &= \lim_{m(D)\to 0}\sum_{l=1}^n \frac{N_{t_l}+N_{t_{l-1}}}{2}(M_{t_l}-M_{t_{l-1}}) \\
&\quad + \lim_{m(D)\to 0}\sum_{l=1}^n \frac{A_{t_l}+A_{t_{l-1}}}{2}(M_{t_l}-M_{t_{l-1}}) \\
&= \frac{1}{2}\lim_{m(D)\to 0}\sum_{l=1}^n \left(N_{t_l}-N_{t_{l-1}}\right)(M_{t_l}-M_{t_{l-1}}) \\
&\quad + \lim_{m(D)\to 0}\sum_{l=1}^n F_{t_{l-1}}(M_{t_l}-M_{t_{l-1}}) \\
&\quad + \frac{1}{2}\lim_{m(D)\to 0}\sum_{l=1}^n \left(A_{t_l}-A_{t_{l-1}}\right)(M_{t_l}-M_{t_{l-1}}) \\
&= \frac{1}{2}\langle M, N\rangle_t + \int_0^t F_s \mathrm{d}M_s \ .
\end{aligned}$$

In particular, if $F=(F_t)_{t\geq 0}$ is a process with finite variation, then

$$\int_0^t F_s \circ \mathrm{d}M_s = \int_0^t F_s \mathrm{d}M_s \ .$$

We will concentrate on Itô's integrals only.

4.2 Stochastic integrals for simple processes

An adapted stochastic process $F=(F_t)_{t\geq 0}$ is called a simple process, if it has a representation

$$F_t(\omega) = f(\omega)1_{\{0\}}(t) + \sum_{i=0}^{\infty} f_i(\omega)1_{(t_i,t_{i+1}]}(t) \tag{4.3}$$

where $0=t_0<t_1<\cdots<t_i\to\infty$, so that for any finite time $T\geq 0$, there are only finite many $t_i\in[0,T]$, each $f_i\in\mathcal{F}^0_{t_i}$ [i.e. f_i is measurable with respect to $\mathcal{F}^0_{t_i}$], $f_0\in\mathcal{F}^0_0$, and F is a bounded process. The space of all simple

(adapted) stochastic processes will be denoted by $\mathcal{L}_0$. If $F = (F_t)_{t\geq 0} \in \mathcal{L}_0$, then Itô's integral of F against Brownian motion $B = (B_t)_{t\geq 0}$ is defined as

$$I(F)_t \equiv \sum_{i=0}^{\infty} f_i(B_{t\wedge t_{i+1}} - B_{t\wedge t_i})$$

where the sum makes sense because only finite terms may not vanish. It is obvious that $I(F) = (I(F)_t)_{t\geq 0}$ is continuous, square-integrable, adapted to $\left(\mathcal{F}_t^0\right)_{t\geq 0}$.

Lemma 4.2.1 *Let $M = (M_t)_{t\geq 0}$ be a continuous, square-integrable martingale, and $s < t \leq u < v$, $f \in \mathcal{F}_s^0$, $g \in \mathcal{F}_t^0$. Then*

$$E\left(g(M_v - M_u)(M_t - M_s)|\mathcal{F}_s^0\right) = 0$$

and

$$E\left(f(M_t - M_s)^2|\mathcal{F}_s^0\right) = E\left(f\left(\langle M\rangle_t - \langle M\rangle_s\right)|\mathcal{F}_s^0\right) \ .$$

Proof. By the tower property of conditional expectations

$$\begin{aligned} & E\left(g(M_v - M_u)(M_t - M_s)|\mathcal{F}_s^0\right) \\ & = E\left\{E\left(g(M_v - M_u)(M_t - M_s)|\mathcal{F}_u^0\right)|\mathcal{F}_s^0\right\} \\ & = E\left\{g(M_t - M_s)E(M_v - M_u|\mathcal{F}_u^0)|\mathcal{F}_s^0\right\} \\ & = 0 \ . \end{aligned}$$

The second equality is trivial as $f \in \mathcal{F}_s^0$ that can be moved out from the conditional expectation. ∎

Lemma 4.2.2 $(I(F)_t)_{t\geq 0}$ *is a martingale*

$$E\left(I(F)_t - I(F)_s|\mathcal{F}_s^0\right) = 0 \ , \quad \forall t > s \ .$$

Proof. Assume that $t_j < t \leq t_{j+1}$, $t_k < s \leq t_{k+1}$ for some $k, j \in \mathbb{N}$. Then $k \leq j$ and

$$\begin{aligned} I(F)_t &= \sum_{i=0}^{j-1} f_i(B_{t_{i+1}} - B_{t_i}) + f_j(B_t - B_{t_j}) \ ; \\ I(F)_s &= \sum_{i=0}^{k-1} f_i(B_{t_{i+1}} - B_{t_i}) + f_k(B_s - B_{t_k}) \ . \end{aligned}$$

If $k < j-1$, then

$$\begin{aligned} I(F)_t - I(F)_s &= \sum_{i=k+1}^{j-1} f_i(B_{t_{i+1}} - B_{t_i}) \\ &\quad + f_j(B_t - B_{t_j}) + f_k(B_{t_{k+1}} - B_s) \ . \end{aligned} \tag{4.4}$$

If $k+1 \leq i \leq j-1$, $s \leq t_i$ so that $\mathcal{F}_s^0 \subset \mathcal{F}_{t_i}^0$. Hence

$$\begin{aligned} & E\left(f_i(B_{t_{i+1}} - B_{t_i})|\mathcal{F}_s^0\right) \\ & \quad = E\left\{E(\left\{f_i(B_{t_{i+1}} - B_{t_i})|\mathcal{F}_{t_i}^0\right\}|\mathcal{F}_s^0\right\} \\ & \quad = E\left\{f_i E\left\{B_{t_{i+1}} - B_{t_i}|\mathcal{F}_{t_i}^0\right\}|\mathcal{F}_s^0\right\} \\ & \quad = 0 \ . \end{aligned}$$

The first equality follows from $f_i \in \mathcal{F}_{t_i}^0$, and in the second equality follows from that (B_t) is a martingale. Similarly

$$\begin{aligned} & E\left(f_j(B_t - B_{t_j})|\mathcal{F}_s^0\right) = 0, \quad t > t_j \geq s, f_j \in \mathcal{F}_{t_j}^0, \\ & E\left(f_k(B_{t_{k+1}} - B_s)|\mathcal{F}_s^0\right) = 0, \quad t_{k+1} \geq s > t_k, f_k \in \mathcal{F}_{t_k}^0 \subset \mathcal{F}_s^0 \ . \end{aligned}$$

Putting these equations together we obtain

$$E\left(I(F)_t - I(F)_s|\mathcal{F}_s^0\right) = 0 \ .$$

If $k = j-1$, then $t_{j-1} < s \leq t_j < t \leq t_{j+1}$ and

$$I(F)_t - I(F)_s = f_{j-1}(B_{t_j} - B_s) + f_j(B_t - B_{t_j})$$

we thus again have

$$E\left(I(F)_t - I(F)_s|\mathcal{F}_s^0\right) = 0 \ .$$

■

Lemma 4.2.3 $\left(I(F)_t^2 - \int_0^t F_s^2 ds\right)_{t\geq 0}$ *is a martingale. Therefore* $I(F) \in \mathcal{M}_2^c$ *and*

$$\langle I(F)\rangle_t = \int_0^t F_s^2 ds \ .$$

Proof. We want to prove that for any $t \geq s$

$$E\left(I(F)_t^2 - \int_0^t F_u^2 du \middle| \mathcal{F}_s^0 \right) = I(F)_s^2 - \int_0^s F_u^2 du \ .$$

In other words, we have to prove that

$$E\left(I(F)_t^2 - I(F)_s^2 - \int_s^t F_u^2 du \middle| \mathcal{F}_s^0 \right) = 0 \ .$$

Obviously

$$\begin{aligned} I(F)_t^2 - I(F)_s^2 &= (I(F)_t - I(F)_s)^2 - 2I(F)_t I(F)_s - 2I(F)_s^2 \\ &= (I(F)_t - I(F)_s)^2 - 2(I(F)_t - I(F)_s)I(F)_s, \end{aligned}$$

and $(I(F)_t)_{t\geq 0}$ is a martingale, so that

$$E\left(I(F)_t - I(F)_s | \mathcal{F}_s^0\right) = 0 \ .$$

While, $I(F)_s \in \mathcal{F}_s^0$ so that

$$\begin{aligned} &E\left\{I(F)_s\left(I(F)_t - I(F)_s\right)|\mathcal{F}_s^0\right\} \\ &= I(F)_s E\left\{I(F)_t - I(F)_s|\mathcal{F}_s^0\right\} = 0 \ . \end{aligned}$$

We therefore only need to show

$$E\left\{ \left(I(F)_t - I(F)_s\right)^2 - \int_s^t F_u^2 du \middle| \mathcal{F}_s^0 \right\} = 0 \ .$$

Now we use the same notations as in the proof of Lemma 4.2.2.

It is clear from eqn 4.4 that if $k < j-1$, then

$$\begin{aligned} (I(F)_t - I(F)_s)^2 &= \sum_{i,l=k+1}^{j-1} f_i f_l (B_{t_{i+1}} - B_{t_i})(B_{t_{l+1}} - B_{t_l}) \\ &\quad + \sum_{i=1}^{j-1} f_i f_j (B_{t_{i+1}} - B_{t_i})(B_t - B_{t_j}) \\ &\quad + \sum_{i=1}^{j-1} f_i f_k (B_{t_{i+1}} - B_{t_i})(B_{t_{k+1}} - B_s) \\ &\quad + f_j^2 (B_t - B_{t_j})^2 + f_k^2 (B_{t_{k+1}} - B_s)^2 \\ &\quad + f_k f_j (B_t - B_{t_i})(B_{t_{k+1}} - B_s) \ . \end{aligned}$$

Using Lemma 4.2.1 below and the fact that both $(B_t)_{t\geq 0}$ and $(B_t^2 - t)_{t\geq 0}$ are martingales, we get

$$\begin{aligned} &E\left\{ \left(I(F)_t - I(F)_s\right)^2 \middle| \mathcal{F}_s^0 \right\} \\ &= E\left(\sum_{j=k+1}^{j-1} f_i^2 (t_{i+1} - t_i) + f_j^2 (t - t_j) + f_k^2 (t_{k+1} - s) \middle| \mathcal{F}_s^0 \right) \end{aligned}$$

so that

$$E\left\{(I(F)_t - I(F)_s)^2 | \mathcal{F}_s^0\right\} = E\left(\int_s^t F_u^2 du \middle| \mathcal{F}_s^0\right) .$$

■

Lemma 4.2.4 *$F \to I(F)$ is linear, and for any $T \geq 0$*

$$E\left(I(F)_T^2\right) = E\left(\int_0^T F_s^2 ds\right) .$$

4.3 Stochastic integrals for adapted processes

In this section we extend the definition of Itô's integrals to integrands which are limits of simple processes. Obviously we only need to define Itô's integrals $I(F)_t$ for $t \leq T$ for arbitrary positive number T. Thus, throughout this section, we are given an arbitrary but fixed time $T > 0$.

4.3.1 The space of square-integrable martingales

If F is a simple process, then the Ito integral $I(F)$ is a continuous, square-integrable martingale with initial zero, and its bracket process $\langle I(F)\rangle_t = \int_0^t F_s^2 \mathrm{d}s$. In particular we have the Itô isometry

$$E|I(F)_T|^2 = E\int_0^t F_s^2 ds$$

which allows us to extend the definition of Itô's integrals to a larger class of integrands.

Let $T > 0$ be a fixed but arbitrary number, and $\mathcal{M}_0^2$ be the vector space of all continuous, square-integrable martingales up to time T on a probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ (where $(\mathcal{F}_t)$ is a filtration $(\mathcal{F}_t)$) with *initial value zero*, endowed with the distance

$$d(M,N) = \sqrt{E|M_T - N_T|^2} \quad \text{for } M, N \in \mathcal{M}_0^2 .$$

By definition, a sequence of square-integrable martingales $(M(k)_t)_{t\geq 0}$ ($k = 1, \cdots,$) converges to M in $\mathcal{M}_0^2$, if and only if

$$M(k)_T \to M_T \quad \text{in } \ L^2(\Omega, \mathcal{F}, P)$$

as $k \to \infty$. The following maximal inequality, which is the "martingale version" of the Markov inequality, allows us to show that $(\mathcal{M}_0^2, d)$ indeed is complete.

Theorem 4.3.1 *(Kolmogorov's inequality) Let $M \in \mathcal{M}_0^2$. Then for any $\lambda > 0$*

$$P\left\{\sup_{0\leq t\leq T}|M_t| \geq \lambda\right\} \leq \frac{1}{\lambda^2}E\left(M_T^2\right) \ .$$

Proof. Since $(M_t)_{t\geq 0}$ is continuous

$$\sup_{0\leq t\leq T}|M_t| = \sup_{t\in D}|M_t|$$

for any countable dense subset D of $[0,T]$, $\sup_{0\leq t\leq T}|M_t|$ is a random variable. For each $n \in \mathbb{N}$, we may apply the Kolmogorov inequality to martingale in discrete-time $\{M_{Tk/2^n}; \mathcal{F}_{Tk/2^n}\}_{k\geq 0}$ to obtain

$$P\left\{\sup_{0\leq k\leq 2^n}|M_{Tk/2^n}| \geq \lambda\right\} \leq \frac{1}{\lambda^2}E\left(M_T^2\right) \ .$$

However, since $D = \{Tk/2^n : n, k \in \mathbb{N}\}$ is dense in $[0,T]$ so that

$$\sup_{0\leq k\leq 2^n-1}|M_{Tk/2^n}| \uparrow \sup_{0\leq t\leq T}|M_t|$$

as $n \to \infty$. Therefore

$$\begin{aligned} P\left\{\sup_{0\leq t\leq T}|M_t| \geq \lambda\right\} &= \lim_{n\to\infty} P\left\{\sup_{0\leq k\leq 2^n-1}|M_{Tk/2^n}| \geq \lambda\right\} \\ &\leq \frac{1}{\lambda^2}E\left(M_T^2\right) \ . \end{aligned}$$

■

Theorem 4.3.2 $(\mathcal{M}_0^2, d)$ *is a complete metric space.*

Proof. Let $M(k) \in \mathcal{M}_0^2$ ($k = 1, 2, \cdots$) be a Cauchy sequence in $\mathcal{M}_0^2$. Then

$$E|M(k)_T - M(l)_T|^2 \to 0 \ , \quad \text{as } k, l \to \infty \ .$$

According to Kolmogorov's inequality

$$P\left\{\sup_{0\leq t\leq T}|M(k)_t - M(l)_t| \geq \lambda\right\} \leq \frac{1}{\lambda^2}E|M(k)_T - M(l)_T|^2 \ ,$$

so that, $M(k)$ uniformly converges to a limit M on $[0,T]$ in probability. Therefore there exists a stochastic process $M \equiv (M_t)$ such that

$$\sup_{0\leq t\leq T}|M(k)_t - M_t| \to 0 \quad \text{ in prob.}$$

Obviously $(M_t)_{t\geq 0}$ is a continuous and square -integrable martingale (up to time T) as the uniform limit of a sequence of continuous martingales. ■

4.3.2 Stochastic integrals as martingales

If $F = (F_t)_{t\geq 0}$ is a limit of simple processes

$$E\int_0^T |F(n)_t - F_t|^2 dt \to 0$$

as $n \to \infty$ for a sequence of some simple processes $\{F(n) : n \in \mathbb{N}\}$, then we denote by $F \in \mathcal{L}^2$. Then the linearity of Ito's integral together with Ito's isometry imply that

$$\begin{aligned} d(I(F(n)), I(F(m))) &= E|I(F(n))_T - I(F(m))_T|^2 \\ &= E\int_0^T |F(n)_t - F(m)_t|^2 dt \to 0 \end{aligned}$$

as $n, m \to \infty$, i.e. $\{I(F(n)\}$ is a Cauchy sequence in $(\mathcal{M}_0^2, d)$. Since $(\mathcal{M}_0^2, d)$ is complete, so that $\lim_{n\to\infty} I(F(n))$ exists in $(\mathcal{M}_0^2, d)$. We naturally define $I(F) = \lim_{n\to\infty} I(F(n))$, which is called Itô's integral of (F_t) against the Brownian motion B. We often write $I(F)_t$ as $\int_0^t F_s dB_s$ or $F.B_t$.

Remark 4.3.3 *1) A process $F = (F_t)_{t\leq T}$ in $\mathcal{L}^2$ is* adapted *and*

$$E\int_0^T F_t^2 dt < +\infty \ .$$

2) The map $F \to I(F)$ is a linear isometry from $\mathcal{L}^2$ to $\mathcal{M}_0^2$, where $\mathcal{L}^2$ is endowed with the norm

$$||F|| = \sqrt{E\int_0^T F_t^2 dt} \ .$$

$\mathcal{M}_0^2$ is a Hilbert space with norm $||M|| = \sqrt{E(M_T^2)}$.

3) If $F \in \mathcal{L}^2$, then $I(F)$ is a continuous, square-integrable martingale with initial value zero (up to time T), and $\langle I(F)\rangle_t = \int_0^t F_s^2 ds$.

$\mathcal{L}^2$ is a very big space which includes many interesting stochastic processes. For example

Lemma 4.3.4 *Let $F = (F_t)_{t\geq 0}$ be an* adapted, left-continuous *stochastic process, satisfying*

$$E\int_0^T F_s^2 ds < +\infty \ . \tag{4.5}$$

Then $F \in \mathcal{L}^2$ *and*

$$I(F)_t = \lim_{m(D)\to 0} \sum_l F_{t_{l-1}} \left(B_{t_l} - B_{t_{l-1}}\right) \quad \textit{in probability}$$

where the limit takes over all finite partitions of $[0,t]$.

Proof. For $n > 0$, let

$$D_n \equiv \{0 = t_0^n < t_1^n < \cdots < t_{n_k}^n = T\}$$

be a sequence of finite partitions of $[0,T]$ such that

$$m(D_n) = \sup_j |t_j^n - t_{j-1}^n| \to 0 \quad \text{as } n \to \infty \ .$$

Let

$$F(n)_t = F_0 1_{\{0\}}(t) + \sum_{l=1}^{n_k} F_{t_{l-1}^n} 1_{(t_{l-1}^n, t_l^n]}(t) \ ; \quad \text{for } \ t \geq 0 \ . \tag{4.6}$$

Then each F_n is simple, and, since F is left-continuous $F(n)_t \to F_t$ for each t. Therefore

$$E\int_0^T |F(n)_s - F_s|^2 \mathrm{d}s \to 0 \ \text{ as } n \to \infty \ .$$

By definition, $F \in \mathcal{L}^2$. ■

Remark 4.3.5 *The condition that* $F = (F_t)_{t\geq 0}$ *is adapted to the filtration* $(\mathcal{F}_t)_{t\geq 0}$ *generated by the Brownian motion, i.e. each* F_t *is measurable with respect to* $\mathcal{F}_t$*, is essential in the definition of Itô's integrals. On the other hand, left-continuity of* $t \to F_t$ *is a technical one, which can be replaced by some sort of Borel measurability (e.g.; right-continuous, continuous, measurable in* (t,ω) *etc.). Left-continuity becomes a correct condition if one attempts to define stochastic integrals of* $F = (F_t)_{t\geq 0}$ *against martingales which may have jumps. The reason is that the left-limit of* F *at time* t *"happens" before time* t*, and if* $t \to F_t$ *is left-continuous, then, for any time* t*, the value* F_t *can be "predicted" by the values taking place strictly before time* t*:*

$$F_t = \lim_{s\uparrow t} F_s \ .$$

Remark 4.3.6 *We should point out that some kind of measurability of random function* $(t,\omega) \to F_t(\omega)$ *is necessary in order to ensure (4.5) make sense. Note that (4.5) may be written as*

$$\int_\Omega \int_0^T F_s(\omega)^2 ds P(d\omega) < +\infty$$

so the natural measurability condition should be that the function

$$F(t,\omega) \equiv F_t(\omega)$$

is measurable with respect to $\mathcal{B}([0,T]) \otimes \mathcal{F}_T$ *for any* $T > 0$*, where* $\mathcal{B}([0,T])$ *is the Borel* σ*-algebra generated by open subsets in* $[0,T]$*, and* $\mathcal{B}([0,T]) \otimes \mathcal{F}_T$ *is the product* σ*-algebra on* $[0,T] \times \Omega$*.*

If $X = (X_t)_{t\geq 0}$ is a continuous stochastic process adapted to $(\mathcal{F}_t)_{t\geq 0}$, f is a Borel function, and

$$E\int_0^T f(X_t)^2 \mathrm{d}t < \infty$$

then the stochastic process $(f(X_t))_{t\geq 0}$ belongs to $\mathcal{L}^2$. In particular, for any Borel measurable function f such that

$$E\int_0^T f(B_t)^2 \mathrm{d}t < \infty \tag{4.7}$$

then $(f(B_t))_{t\geq 0}$ is in $\mathcal{L}^2$. What does condition (4.7) mean? While

$$\begin{aligned} E\int_0^T f(B_t)^2 \mathrm{d}t &= \int_0^T Ef(B_t)^2 \mathrm{d}t \\ &= \int_0^T P_t(f^2)(0) \mathrm{d}t \end{aligned}$$

where

$$\begin{aligned} P_t(f^2)(0) &= \frac{1}{(2\pi t)^{d/2}} \int_{\mathbb{R}^d} f(x)^2 \mathrm{e}^{-|x|^2/2t} \mathrm{d}x \\ &= \frac{1}{(2\pi)^{d/2}} \int_{\mathbb{R}^d} f(\sqrt{t}x)^2 \mathrm{e}^{-|x|^2/2} \mathrm{d}x \ . \end{aligned}$$

Therefore, if f is a polynomial, then $f(B_t)$ is in $\mathcal{L}^2$, and for any constant α the process $(\mathrm{e}^{\alpha B_t})_{t\geq 0}$ belongs to $\mathcal{L}^2$ as well. How about the stochastic process $F_t = \mathrm{e}^{\alpha B_t^2}$? In this case

$$E\int_0^T F_t^2 \mathrm{d}t = \frac{1}{(2\pi)^{d/2}} \int_0^T \int_{\mathbb{R}^d} \mathrm{e}^{2\alpha t x^2} \mathrm{e}^{-|x|^2/2} \mathrm{d}x$$

and therefore

$$E\int_0^T F_t^2 \mathrm{d}t < \infty \qquad \text{if } \alpha \leq 0 \ .$$

In the case $\alpha > 0$, then

$$E\int_0^T F_t^2 \mathrm{d}t < \infty \ \text{ iff } \quad T < \frac{1}{4\alpha} \ .$$

4.3.3 Summary of main properties

If $F = (F_t)_{t\geq 0} \in \mathcal{L}_2$, then both

$$\int_0^t F_s \mathrm{d}B_s \qquad \text{and} \qquad \left(\int_0^t F_s \mathrm{d}B_s\right)^2 - \int_0^t F_s^2 \mathrm{d}s$$

are continuous martingales with initial zero, and therefore $\langle F.B\rangle_t = \int_0^t F_s^2 \mathrm{d}s$. In general (by the use of the polarization $xy = \frac{1}{4}\left((x+y)^2 - (x-y)^2\right)$)

$$\langle F.B, G.B\rangle_t = \int_0^t F_s G_s \mathrm{d}s \qquad \forall F, G \in \mathcal{L}^2 .$$

In particular

$$E\left[\int_0^T F_s \mathrm{d}B_s\right]^2 = E\left(\int_0^T F_s^2 \mathrm{d}s\right) .$$

and for any $t \geq s$,

$$E\left\{\left(\int_s^t F_u \mathrm{d}B_u\right)^2 \middle| \mathcal{F}_s\right\} = E\left\{\int_s^t F_u^2 \mathrm{d}u \middle| \mathcal{F}_s\right\} .$$

4.4 Itô's integration for semi-martingales

We may apply the same procedure of defining Itô's integrals along Brownian motion to any continuous, square-integrable martingales. Indeed, if $M \in \mathcal{M}_0^2$ if $F = (F_t)_{t\geq 0}$ is a bounded, adapted, simple process

$$F_t = f 1_{\{0\}}(t) + \sum_i f_i 1_{(t_i, t_{i+1}]}(t)$$

then define

$$I^M(F) = \sum_{i=0}^{\infty} f_i \cdot (M_{t\wedge t_{i+1}} - M_{t\wedge t_i}) .$$

As before, we have

1. $I^M(F) \in \mathcal{M}_0^2$.
2. The bracket process $\langle I^M(F)\rangle_t = \int_0^t F_s^2 \mathrm{d}\langle M\rangle_s$, i.e. $I^M(F)_t^2 - \int_0^t F_s \mathrm{d}\langle M\rangle_s$ is a martingale.

3. (Itô's isometry) For any $T > 0$, we have

$$E\left(\int_0^T F_t \mathrm{d}M_t\right)^2 = E\int_0^T F_t^2 \mathrm{d}\left\langle M\right\rangle_t \ .$$

Let $T > 0$ be a fixed time.

Definition 4.4.1 *A stochastic process $F = (F_t)_{t\geq 0} \in \mathcal{L}^2(M)$, if there is a sequence $\{F(n)\}$ of simple stochastic processes $(F(n))$ such that*

$$E\left\{\int_0^T F(n)_t^2 d\left\langle M\right\rangle_t\right\} < \infty$$

and

$$E\left\{\int_0^T |F(n)_t - F_t|^2 d\left\langle M\right\rangle_t\right\} \to 0 \text{ as } n \to \infty \ .$$

In other words, $\mathcal{L}^2(M)$ is the closure of all simple processes (up to time T) under the norm

$$||F|| = \sqrt{E\left\{\int_0^T F(n)_t^2 \mathrm{d}\left\langle M\right\rangle_t\right\}}$$

(this norm of course depends on the running time T and the martingale $M \in \mathcal{M}_0^2$), and thus $\mathcal{L}^2(M)$ is a Banach space. Indeed, the above norm is induced by a scalar product, so that $\mathcal{L}^2(M)$ is a Hilbert space. If $F \in \mathcal{L}^2(M)$, and $||F - F(n)|| \to 0$ for a sequence of simple processes, thanks to Ito's isometry

$$E\left\{I^M(F)_T^2\right\} = ||F|| \ ,$$

it follows that

$$I^M(F) \equiv \lim_{n\to\infty} I^M(F_n) \ , \qquad \text{in } \mathcal{M}_0^2$$

exists.We use either $F.M$ or $\int_0^t F_s \mathrm{d}M_s$ to denote $I^M(F)$. According to definition, $I^M(F) \in \mathcal{M}_0^2$ and $\langle I^M(F)\rangle_t = \int_0^t F_s^2 \mathrm{d}\langle M\rangle_s$. By the use of the polarization identity, if M, $N \in \mathcal{M}_0^2$ and $F \in \mathcal{L}^2(M)$, $G \in \mathcal{L}^2(N)$, then

$$\left\langle F.M, G.N\right\rangle_t = \int_0^t F_s G_s \mathrm{d}\left\langle M, N\right\rangle_s$$

and $F.(G.M) = (FG).M$, as far as these stochastic integrals make sense. That is,

$$\int_0^t F_s \mathrm{d}\left(\int_0^s G_u \mathrm{d}M_u\right)_s = \int_0^t F_s G_s \mathrm{d}M_s \ .$$

The Itô integration may be extended to local martingales (see next section for the concepts of local martingales etc.). Let me briefly describe the idea. Suppose $M = (M_t)_{t\geq 0}$ is a continuous, local martingale with initial zero, then we may choose a sequence $\{T_n\}$ of stopping times such that $T_n \uparrow \infty$ a.s. and for each n, $M^{T_n} = (M_{t\wedge T_n})_{t\geq 0}$ is a continuous, square-integrable martingale with initial zero. In this case we may define

$$\langle M\rangle_t = \langle M^{T_n}\rangle_t \quad \text{ if } \ t \leq T_n$$

which is an adapted, continuous, increasing process with initial zero such that

$$M_t^2 - \langle M\rangle_t$$

is a local martingale.

Let $F = (F_t)_{t\geq 0}$ be a left-continuous, adapted process such that for each $T > 0$

$$\int_0^T F_s^2 \mathrm{d}\langle M\rangle_s < \infty \qquad \text{a.s.} \tag{4.8}$$

and define

$$S_n = \inf\left\{t \geq 0 : \int_0^t F_s^2 \mathrm{d}\langle M\rangle_s \geq n\right\} \wedge n$$

which is a sequence of stopping times. Condition (4.8) ensures that $S_n \uparrow \infty$. Let $\tilde{T}_n = T_n \wedge S_n$. Then $\tilde{T}_n \uparrow \infty$ almost surely, and for each n, $M^{\tilde{T}_n} \in \mathcal{M}_0^2$. Let

$$F(n)_t = F_t 1_{\{t\leq \tilde{T}_n\}} \ .$$

Then

$$\int_0^\infty F(n)_s^2 \mathrm{d}\langle M\rangle_s = \int_0^{\tilde{T}_n} F_s^2 \mathrm{d}\langle M\rangle_s \leq n$$

so that $F(n) \in \mathcal{L}_2(M^{\tilde{T}_n})$. We may define

$$(F.M)_t = \int_0^t F(n)_s \mathrm{d}\left(M^{\tilde{T}_n}\right)_s \qquad \text{if } \ t \leq \tilde{T}_n \uparrow \infty$$

for $n = 1, 2, 3, \cdots$, called the Itô integral of F with respect to local martingale M. It can be shown that $F.M$ does not depend on the choice of stopping times T_n. By definition, both $F.M$ and

$$(F.M)_t^2 - \int_0^t F_s^2 \mathrm{d}\langle M\rangle_s$$

are continuous, local martingales with initial zero.

Finally let us extend the theory of stochastic integrals to the most useful class of (continuous) semimartingales. An adapted, continuous stochastic process $X = (X_t)_{t\geq 0}$ is a semimartingale if X possesses a decomposition

$$X_t = M_t + V_t$$

where $(M_t)_{t\geq 0}$ is a continuous local martingale, and $(V_t)_{t\geq 0}$ is stochastic processes with finite variation on any finite interval.

If $f(t)$ is a function on $[0,T]$ *having finite variation*:

$$\sup_D \sum_l |f(t_l) - f(t_{l-1})| < +\infty$$

where D runs over all finite partitions of $[0,t]$ (for any fixed t), then

$$\int_0^t g(s)\mathrm{d}f(s)$$

is understood as the Lebesgue-Stieltjes integral. If in addition $s \to f(s)$ is continuous, then

$$\int_0^t g(s)\mathrm{d}f(s) = \lim_{m(D)\to 0} \sum_l g(t_{l-1})(f(t_l) - f(t_{l-1})) \ .$$

Therefore, if $V = (V_t)_{t\geq 0}$ is a continuous stochastic process with finite variation, then

$$\int_0^t F_s \mathrm{d}V_s$$

is a stochastic process defined path-wisely as the Lebesgue-Stieltjes integral

$$\begin{aligned}\int_0^t F_s \mathrm{d}V_s(\omega) &\equiv \int_0^t F_s(\omega)\mathrm{d}V_s(\omega) \\ &= \lim_{m(D)\to 0} \sum_l F_{t_{l-1}}(\omega)(V_{t_l}(\omega) - V_{t_{l-1}}(\omega)) \ .\end{aligned}$$

The definition of stochastic integrals may be extended to any continuous *semi-martingale* in an obvious way, namely

$$\int_0^t F_s \mathrm{d}X_s = \int_0^t F_s \mathrm{d}M_s + \int_0^t F_s \mathrm{d}V_s$$

where, the first term on the right-hand side is the Itô's integral with respect to local martingale M defined in probability sense, which is again a local martingale, the second term is the usual Lebesgue-Stieltjes integral which is defined path-wisely. Moreover

$$\int_0^t F_s \mathrm{d}X_s = \lim_{m(D)\to 0} \sum_l F_{t_{l-1}} \left(X_{t_l} - X_{t_{l-1}} \right) \qquad \text{in probab.}$$

4.5 Ito's formula

Ito's formula was established by K. Itô in 1944. Since Itô's stated it as a lemma in his seminal paper [ ], Itô's formula is also refereed in literature as Itô's Lemma. Itô's Lemma is indeed the Fundamental Theorem in stochastic calculus.

We have used in many occasions the following elementary formula

$$X_{t_j}^2 - X_{t_{j-1}}^2 = \left(X_{t_j} - X_{t_{j-1}} \right)^2 + 2X_{t_{j-1}} \left(X_{t_j} - X_{t_{j-1}} \right) .$$

If in addition $(X_t)_{t\geq 0}$ is a continuous square-integrable martingale, then, by adding up the above identity over $j = 1, \cdots, n$, where $0 = t_0 < t_1 < \cdots < t_n = t$ is an arbitrary finite partition, one obtains

$$X_t^2 - X_0^2 = 2\sum_{j=1}^n X_{t_{j-1}} \left(X_{t_j} - X_{t_{j-1}} \right) + \sum_{j=1}^n \left(X_{t_j} - X_{t_{j-1}} \right)^2 .$$

Letting $m(D) \to 0$, we obtain

$$X_t^2 - X_0^2 = 2\int_0^t X_s dX_s + \langle X \rangle_t .$$

which is the Itô formula for the martingale $(X_t)_{t\geq 0}$ applying to $f(x) = x^2$. By using polarization and localization, we establish the following integration by parts formula.

Lemma 4.5.1 *If X, Y are two continuous semi-martingales: $X = M + A$ and $Y = N + B$, where M and N are two continuous local martingales, A and B are two adapted variational process, then*

$$X_t Y_t - X_0 Y_0 = \int_0^t Y_s dX_s + \int_0^t X_s dY_s + \langle M, N \rangle_t .$$

Corollary 4.5.2 *(Integration by parts) Let $X = M + A$ and $Y = N + B$ be a continuous semimartingale: M and N are continuous local martingales, and A, B are continuous, adapted processes with finite variations. Then*

$$X_t Y_t - X_0 Y_0 = \int_0^t X_s dY_s + \int_0^t Y_s dX_s + \langle M, N \rangle_t .$$

The following is the fundamental theorem in stochastic calculus.

Theorem 4.5.3 *(Itô's formula) Let $X = (X_t^1, \cdots, X_t^d)$ be a continuous semi-martingale in $\mathbb{R}^d$ with decompositions $X_t^i = M_t^i + A_t^i$: $M_t^1, \cdots, M_t^d$ are continuous local martingales, and $A_t^1, \cdots, A_t^d$ are continuous, locally integrable, adapted processes with finite variations. Let $f \in C^2(\mathbb{R}^d, \mathbb{R})$. Then*

$$\begin{aligned} f(X_t) - f(X_0) &= \sum_{i=1}^d \int_0^t \frac{\partial f}{\partial x_i}(X_s) dX_s^i \\ &\quad + \frac{1}{2} \sum_{i,j=1}^d \int_0^t \frac{\partial^2 f}{\partial x_i \partial x_j}(X_s) d\langle M^i, M^j \rangle_s . \end{aligned} \tag{4.9}$$

The first term on the right-hand side of (4.9) can be decomposed into

$$\sum_{j=1}^d \int_0^t \frac{\partial f}{\partial x_i}(X_s) dM_s^j + \sum_{j=1}^d \int_0^t \frac{\partial f}{\partial x_i}(X_s) dA_s^j$$

so that $f(X_t) - f(X_0)$ is again a semi-martingale with its martingale part given by

$$M_t^f = \sum_{j=1}^d \int_0^t \frac{\partial f}{\partial x_i}(X_s) dM_s^j .$$

It follows that

$$\langle M^f, M^g \rangle_t = \int_0^t \sum_{i,j=1}^d \frac{\partial f}{\partial x_i}(X_s) \frac{\partial g}{\partial x_j}(X_s) \mathrm{d}\langle M^i, M^j \rangle_s .$$

4.5.1 Itô's formula for BM

If $B = (B_t^1, \cdots, B_t^d)_{t \geq 0}$ is Brownian motion in $\mathbb{R}^d$, then, for $f \in C^2(\mathbb{R}^d, \mathbb{R})$

$$f(B_t) - f(B_0) = \int_0^t \nabla f(B_s) . \mathrm{d}B_s + \int_0^t \frac{1}{2} \Delta f(B_s) \mathrm{d}s .$$

Let

$$M_t^f = f(B_t) - f(B_0) - \int_0^t \frac{1}{2} \Delta f(B_s) \mathrm{d}s .$$

Then M^f is a local martingale and

$$\langle M^f, M^g \rangle_t = \int_0^t \langle \nabla f, \nabla g \rangle (B_s) \mathrm{d}s .$$

4.5.2 Proof of Itô's formula.

Let us prove the Itô formula for one-dimensional case. By using localization technique, we only need to prove it for a continuous, square-integrable martingale $M = (M_t)_{t\geq 0}$. Thus we need to show

$$f(M_t) - f(M_0) = \int_0^t f'(M_s)\mathrm{d}M_s + \frac{1}{2}\int_0^t f''(M_s)\mathrm{d}\langle M\rangle_s \ . \tag{4.10}$$

The formula is true for $f(x) = x^2$ ($f'(x) = 2x$ and $f''(x) = 2$) as we have seen

$$M_t^2 - M_0^2 = 2\int_0^t M_s\mathrm{d}M_s + \langle M\rangle_t \ .$$

Suppose (4.10) is true for $f(x) = x^n$:

$$M_t^n - M_0^n = n\int_0^t M_s^{n-1}\mathrm{d}M_s + \frac{n(n-1)}{2}\int_0^t M_s^{n-2}\mathrm{d}\langle M\rangle_s \ ,$$

by applying integration by parts formula to M^n and M, one obtains

$$\begin{aligned}
M_t^{n+1} - M_0^{n+1} &= \int_0^t M_s^n\mathrm{d}M_s + \int_0^t M_s\mathrm{d}M_s^n + \langle M, M^n\rangle_t \\
&= \int_0^t M_s^n\mathrm{d}M_s + \int_0^t M_s\mathrm{d}\left\{nM_s^{n-1}\mathrm{d}M_s + \frac{n(n-1)}{2}M_s^{n-2}\mathrm{d}\langle M\rangle_s\right\} \\
&\quad + \int_0^t nM_s^{n-1}\mathrm{d}\langle M\rangle_s \\
&= (n+1)\int_0^t M_s^n\mathrm{d}M_s + \frac{(n+1)n}{2}\int_0^t M_s^{n-1}\mathrm{d}\langle M\rangle_s
\end{aligned}$$

which implies that (4.10) for power function x^{n+1}. Itô's formula holds thus for any polynomial, so is it for any C^2 function f due to Taylor's expansions.

4.6 Selected applications of Itô's formula

In this section, we present several applications of Itô's lemma.

4.6.1 Lévy's characterization of Brownian motion

Our first application is Lévy's martingale characterization of Brownian motion. Let $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ be a filtered probability space satisfying the usual condition.

Theorem 4.6.1 *Let $M_t = (M_t^1, \cdots, M_t^d)$ be an adapted, continuous stochastic process on $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ taking values in $\mathbb{R}^d$ with initial zero. Then $(M_t)_{t\geq 0}$ is a Brownian motion if and only if*

1. *Each M_t^i is a continuous square-integrable martingale.*
2. *$M_t^i M_t^j - \delta_{ij} t$ is a martingale, that is, $\langle M^i, M^j \rangle_t = \delta_{ij} t$ for every pair (i,j).*

Proof. We only need to prove the sufficient part. Recall that, under the assumption, $(M_t)_{t\geq 0}$ is a Brownian motion if and only if

$$E\left(e^{\sqrt{-1}\langle \xi, M_t - M_s\rangle} \middle| \mathcal{F}_s \right) = \exp\left\{ -\frac{|\xi|^2}{2}(t-s) \right\} \tag{4.11}$$

for any $t > s$ and $\xi = (\xi_i) \in \mathbb{R}^d$. We thus consider the adapted process

$$Z_t = \exp\left(\sqrt{-1} \sum_{i=1}^d \xi_i M_t^i + \frac{|\xi|^2}{2} t \right)$$

and we show it is a martingale. To this end, we apply Itô's formula to $f(x) = e^x$ (in this case $f' = f'' = f$) and semi-martingale

$$X_t = \sqrt{-1} \sum_{i=1}^d \xi_i M_t^i + \frac{|\xi|^2}{2} t \ ,$$

and obtain

$$\begin{aligned}
Z_t &= Z_0 + \int_0^t Z_s \mathrm{d}\left(\sqrt{-1} \sum_{i=1}^d \xi_i M_s^i + \frac{|\xi|^2}{2} s \right) \\
&\quad + \frac{1}{2} \int_0^t Z_s \mathrm{d}\langle \sqrt{-1} \sum_{i=1}^d \xi_i M^i \rangle_s \\
&= 1 + \sqrt{-1} \sum_{i=1}^d \xi_i \int_0^t Z_s \mathrm{d}M_s^i + \frac{|\xi|^2}{2} \int_0^t Z_s \mathrm{d}s \\
&\quad - \frac{1}{2} \int_0^t \sum_{i,j=1}^d \xi_i \xi_j Z_s \mathrm{d}\langle M^i, M^j \rangle_s \\
&= 1 + \sqrt{-1} \sum_{i=1}^d \xi_i \int_0^t Z_s \mathrm{d}M_s^i
\end{aligned}$$

the last equality follows from

$$\frac{1}{2}\int_0^t \sum_{i,j=1}^d \xi_i\xi_j Z_s \mathrm{d}\langle M^i, M^j\rangle_s = \frac{1}{2}|\xi|^2\int_0^t Z_s \mathrm{d}s \ .$$

due to the assumption that $\langle M^i, M^j\rangle_s = \delta_{ij}s$. Since $|Z_s| = e^{|\xi|^2 s/2}$, so that for any $T > 0$

$$E\int_0^T |Z_s|^2 \mathrm{d}s = \int_0^T e^{|\xi|^2 s}\mathrm{d}s < +\infty$$

and therefore $(Z_t) \in \mathcal{L}^2(M^i)$ for $i = 1, \cdots, d$ as $\langle M^i\rangle_t = t$. It follows that

$$\int_0^t Z_s \mathrm{d}M_s^i \in \mathcal{M}_2^c \ .$$

That is, Z_s is a continuous, square-integrable martingale with initial value 1. (4.11) follows from the martingale property

$$E\left(e^{i\langle \xi, M_t\rangle + \frac{|\xi|^2}{2}t} \middle| \mathcal{F}_s \right) = e^{i\langle \xi, M_s\rangle + \frac{|\xi|^2}{2}s}$$

for $t > s$. ■

4.6.2 Time-changes of Brownian motion

Theorem 4.6.2 *(Dambis, Dubins and Schwarz) Let* $M = (M_t)_{t\geq 0}$ *be a continuous, local martingale on* $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ *with initial value zero satisfying* $\langle M\rangle_\infty = \infty$*, and let*

$$T_t = \inf\{s : \langle M\rangle_s > t\} \ .$$

Then T_t *is a stopping time for each* $t \geq 0$*,* $B_t = M_{T_t}$ *is an* $(\mathcal{F}_{T_t})$*-Brownian motion, and* $M_t = B_{\langle M\rangle_t}$*.*

Proof. The family $T = (T_t)_{t\geq 0}$ is called a time-change, because each T_t is a stopping time (*exercise*), and obviously $t \to T_t$ is increasing (*another exercise*). Each T_t is finite $\mathbb{P}$-a.e. because $\langle M\rangle_\infty = \infty$ P-a.e. (*exercise*). By continuity of $\langle M\rangle_t$

$$\langle M\rangle_{T_t} = t \quad P\text{-a.s.}$$

Applying Doob's stopping theorem for the square integrable martingale $(M_{s\wedge T_t})_{s\geq 0}$ and stopping times $T_t \geq T_s$ $(t \geq s)$, we obtain that

$$E\left(M_{T_t}|\mathcal{F}_{T_s}\right) = M_{T_s}$$

i.e. B_t is a $(\mathcal{F}_{T_t})$-local martingale. By the same argument but to the martingale $(M^2_{s\wedge T_t} - \langle M\rangle_{s\wedge T_t})_{s\geq 0}$ we have

$$E\left(M^2_{T_t} - \langle M\rangle_{T_t} | \mathcal{F}_{T_s}\right) = M^2_{T_s} - \langle M\rangle_{T_s}.$$

Hence $(B_t^2 - t)$ is an $(\mathcal{F}_{T_t})$-local martingale. We can prove that $t \to B_t$ is continuous, so that $B = (B_t)_{t\geq 0}$ is an $(\mathcal{F}_{T_t})$ Brownian motion. ■

4.6.3 Stochastic exponentials

In this section we consider a simple stochastic differential equation

$$dZ_t = Z_t dX_t \ , \quad Z_0 = 1 \tag{4.12}$$

where $X_t = M_t + A_t$ is a continuous semi-martingale. The solution of (4.12) is called the *stochastic exponential* of X. The equation (4.12) should be understood as an integral equation

$$Z_t = 1 + \int_0^t Z_s dX_s \tag{4.13}$$

where the integral is taken as Itô's integral. To find the solution to (4.13) we may try

$$Z_t = \exp(X_t + V_t)$$

where $(V_t)_{t\geq 0}$ to be determined as a "correction" term (which has finite variation) due to the Itô's integration. Applying Itô's formula we obtain

$$Z_t = 1 + \int_0^t Z_s \mathrm{d}(X_s + V_s) + \frac{1}{2}\int_0^t Z_s \mathrm{d}\langle M\rangle_s$$

and therefore, in order to match the equation (4.13) we must choose $V_t = -\frac{1}{2}\langle M\rangle_t$.

Lemma 4.6.3 *Let $X_t = M_t + A_t$ (where M is a continuous local martingale, A is an adapted continuous process with finite total variation) with $X_0 = 0$. Then*

$$\mathcal{E}(X)_t = \exp\left(X_t - \frac{1}{2}\langle M\rangle_t\right)$$

is the solution to (4.13).

$\mathcal{E}(X)$ is called the *stochastic exponential* of $X = (X_t)_{t\geq 0}$.

Proposition 4.6.4 *Let $(M_t)_{t\geq 0}$ be a continuous local martingale with $M_0 = 0$. Then the stochastic exponential $\mathcal{E}(M)$ is a continuous, non-negative local martingale.*

Remark 4.6.5 *According to definition of Itô's integrals, if $T > 0$ such that*

$$E\int_0^T e^{2M_t - \langle M\rangle_t} d\langle M\rangle_t < +\infty \tag{4.14}$$

then the stochastic exponential

$$\mathcal{E}(M)_t = \exp\left(M_t - \frac{1}{2}\langle M\rangle_t\right)$$

is a non-negative, continuous martingale.

The remarkable fact is that, although $\mathcal{E}(M)$ may fail to be a martingale, but it is nevertheless a super-martingale.

Lemma 4.6.6 *Let $X = (X_t)_{t\geq 0}$ be a* ***non-negative****, continuous local martingale. Then $X = (X_t)_{t\geq 0}$ is a super-martingale: $E(X_t|\mathcal{F}_s) \leq X_s$ for any $t < s$. In particular $t \to EX_t$ is decreasing, and therefore $EX_t \leq EX_0$ for any $t > 0$.*

Proof. Recall Fatou's lemma: if $\{f_n\}$ is a sequence of non-negative, integrable functions on a probability space $(\Omega, \mathcal{F}, P)$, such that

$$\underline{\lim}_{n\to\infty} E(f_n) < +\infty \ ,$$

then $\underline{\lim}_{n\to\infty} f_n$ is integrable and

$$E\left(\underline{\lim}_{n\to\infty} f_n|\mathcal{G}\right) \leq \underline{\lim}_{n\to\infty} E\left(f_n|\mathcal{G}\right)$$

for any sub σ-algebra $\mathcal{G}$ (see page 88, D. Williams: Probability with Martingales).

By definition, there is a sequence of finite stopping times $T_n \uparrow +\infty$ P-a.e. such that $X^{T_n} = (X_{t\wedge T_n})_{t\geq 0}$ is a martingale for each n. Hence

$$E\left(X_{t\wedge T_n}|\mathcal{F}_s\right) = X_{s\wedge T_n}, \quad \forall t \geq s, n = 1, 2, \cdots \ .$$

In particular

$$E\left(X_{t\wedge T_n}\right) = EX_0$$

so that, by Fatou's lemma, $X_t = \lim_{n\to\infty} X_{t\wedge T_n}$ is integrable. Applying Fatou's lemma to $X_{t\wedge T_n}$ and $\mathcal{G} = \mathcal{F}_s$ for $t > s$ we have

$$\begin{aligned} E(X_t|\mathcal{F}_s) &= E\left(\lim_{n\to\infty} X_{t\wedge T_n}|\mathcal{F}_s\right) \\ &\leq \underline{\lim}_{n\to\infty} E(X_{t\wedge T_n}|\mathcal{F}_s) \\ &= \underline{\lim}_{n\to\infty} X_{s\wedge T_n} \\ &= X_s \end{aligned}$$

According to definition, $X = (X_t)_{t\geq 0}$ is a super-martingale. ■

Corollary 4.6.7 *Let $M = (M_t)_{t\geq 0}$ be a continuous, local martingale with $M_0 = 0$. Then $\mathcal{E}(M)$ is a super-martingale. In particular,*

$$E\exp\left(M_t - \frac{1}{2}\langle M\rangle_t\right) \leq 1 \quad \textit{for all} \quad t \geq 0 \ .$$

Clearly, a continuous super-martingale $X = (X_t)_{t\geq 0}$ is a martingale if and only if its expectation $t \to E(X_t)$ is constant. Therefore

Corollary 4.6.8 *Let $M = (M_t)_{t\geq 0}$ be a continuous, local martingale with $M_0 = 0$. Then $\mathcal{E}(M)$ is a martingale up to time T, if and only if*

$$E\exp\left(M_T - \frac{1}{2}\langle M\rangle_T\right) = 1 \ . \tag{4.15}$$

Stochastic exponentials of local martingales play an important rôle in probability transformations. It is vital in many applications to know whether the stochastic exponential of a given martingale $M = (M_t)_{t\geq 0}$ is indeed a martingale. A simple sufficient condition to ensure (4.15) is the so-called Novikov's condition stated in Theorem 4.6.9 below (A. A. Novikov: On moment inequalities and identities for stochastic integrals, *Proc. second Japan-USSR Symp. Prob. Theor., Lecture Notes in Math.*, **330**, 333-339, Springer-Verlag, Berlin 1973).

Theorem 4.6.9 *(A. A. Novikov) Let $M = (M_t)_{t\geq 0}$ be a continuous local martingale with $M_0 = 0$. If*

$$E\exp\left(\frac{1}{2}\langle M\rangle_T\right) < +\infty \ , \tag{4.16}$$

then $\mathcal{E}(M)$ is a martingale up to time T.

Proof. The following proof is due to J. A. Yan: *Critères d'intégrabilité uniforme des martingales exponentielles, Acta. Math. Sinica* **23**, 311-318 (1980). The idea is the following, first show that, under the Novikov condition (4.16), for any $0 < \alpha < 1$

$$\mathcal{E}(\alpha M)_t \equiv \exp\left(\alpha M_t - \frac{1}{2}\alpha^2 \langle M \rangle_t\right)$$

is a uniformly integrable martingale up to time T. For any α, $\mathcal{E}(\alpha M)_t$ is the stochastic exponential of the local martingale αM_t, so that $\mathcal{E}(\alpha M)$ is a non-negative, continuous local martingale, $E\left(\mathcal{E}(\alpha M)_t\right) \leq 1$. We also have the following scaling property

$$\begin{aligned}
\mathcal{E}(\alpha M)_t &\equiv \exp\left\{\alpha\left(M_t - \frac{1}{2}\langle M \rangle_t\right) - \frac{1}{2}\alpha(\alpha-1)\langle M \rangle_t\right\} \\
&= \left(\mathcal{E}(M)_t\right)^\alpha \exp\left\{\frac{1}{2}\alpha(1-\alpha)\langle M \rangle_t\right\} .
\end{aligned}$$

For any finite stopping time $S \leq T$ and for any $A \in \mathcal{F}_T$

$$E\left(1_A \mathcal{E}(\alpha M)_S\right) = E\left\{1_A \left(\mathcal{E}(M)_S\right)^\alpha \exp\left[\frac{1}{2}\alpha(1-\alpha)\langle M \rangle_S\right]\right\} . \tag{4.17}$$

Using Hölder's inequality with $\frac{1}{\alpha} > 1$ and $\frac{1}{1-\alpha}$ in (4.17) one obtains

$$\begin{aligned}
E\left\{1_A \mathcal{E}(\alpha M)_S\right\} &= E\left\{\left(\mathcal{E}(M)_S\right)^\alpha \exp\left[\frac{1}{2}\alpha(1-\alpha)\langle M \rangle_S\right]\right\} \\
&\leq \left\{E\left(\mathcal{E}(M)_S\right)\right\}^\alpha \left\{E\left[1_A \exp\left(\frac{1}{2}\alpha\langle M \rangle_S\right)\right]\right\}^{1-\alpha} \\
&\leq \left\{E\left(\mathcal{E}(M)_T\right)\right\}^\alpha \left\{E\left[1_A \exp\left(\frac{1}{2}\alpha\langle M \rangle_T\right)\right]\right\}^{1-\alpha} \\
&\leq \left\{E\left[1_A \exp\left(\frac{1}{2}\alpha\langle M \rangle_T\right)\right]\right\}^{1-\alpha} \\
&\leq E\left\{1_A \exp\left(\frac{1}{2}\langle M \rangle_T\right)\right\} .
\end{aligned} \tag{4.18}$$

According to Corollary **??**

$$\left\{\mathcal{E}(\alpha M)_S : \text{any stopping times } S \leq T\right\}$$

is uniformly integrable, so that, according to Corollary **??**, $\mathcal{E}(\alpha M)$ must be a martingale on $[0, T]$. Therefore

$$E\left(\mathcal{E}(\alpha M)_T\right) = E\left(\mathcal{E}(\alpha M)_0\right) = 1, \quad \forall \alpha \in (0,1).$$

Set $A = \Omega$ and $S = t \leq T$ in (4.18), the first inequality of (4.18) becomes

$$\begin{aligned} 1 &= E\left(\mathcal{E}(\alpha M)_t\right) \\ &\leq \left(E\left(\mathcal{E}(M)_t\right)\right)^{\alpha} \left\{ E\left(\exp\left(\frac{1}{2} \langle M \rangle_T \right) \right) \right\}^{1-\alpha} \end{aligned}$$

for every $\alpha \in (0,1)$. Letting $\alpha \uparrow 1$ we thus obtain

$$E\left(\mathcal{E}(M)_t\right) \geq 1$$

so that $E\left(\mathcal{E}(M)_t\right) = 1$ for any $t \leq T$, it follows thus that $\mathcal{E}(M)_t$ is a martingale up to T. ■

Consider a standard Brownian motion $B = (B_t)$, and $F = (F_t)_{t\geq 0} \in \mathcal{L}_2$. If

$$E \exp\left[\frac{1}{2} \int_0^T F_t^2 \mathrm{d}t \right] < \infty$$

then

$$X_t = \exp\left\{ \int_0^t F_s \mathrm{d}B_s - \frac{1}{2} \int_0^t F_s^2 \mathrm{d}s \right\} \tag{4.19}$$

is a positive martingale on $[0,T]$. For example, for any bounded process $F = (F_t)_{t\geq 0} \in \mathcal{L}_2$: $|F_t(\omega)| \leq C$ (for all $t \leq T$ and $\omega \in \Omega$), where C is a constant, then

$$E\left\{ \exp\left(\frac{1}{2} \int_0^T F_t^2 \mathrm{d}t \right) \right\} \leq \exp\left(\frac{1}{2} C^2 T \right) < \infty$$

so that, in this case, $X = (X_t)$ defined by (4.19) is a martingale up to time T.

Novikov's condition is very nice, it is however not easy to verify in many interesting cases. For example, consider the stochastic exponential of the martingale $\int_0^t B_s \mathrm{d}B_s$, the Novikov condition requires to estimate the integral

$$E\left\{ \exp\left[\frac{1}{2} \int_0^T B_t^2 \mathrm{d}t \right] \right\}$$

which is already not an easy task.

4.6.4 Exponential inequality

We are going to present three significant applications of stochastic exponentials: a sharp improvement of Doob's maximal inequality for martingales, Girsanov's theorem, and the martingale representation theorem (in the next section). Additional applications will be discussed in the next chapter.

Recall that, according to Doob's maximal inequality, if $(X_t)_{t\geq 0}$ is a continuous super-martingale on $[0,T]$, then for any $\lambda > 0$

$$P\left\{\sup_{t\in[0,T]} |X_t| \geq \lambda\right\} \leq \frac{1}{\lambda}\left(E(X_0) + 2E(X_T^-)\right)$$

where $x^- = -x$ if $x < 0$ and $= 0$ if $x \geq 0$. In particular, if $(X_t)_{t\geq 0}$ is a *non-negative*, continuous super-martingale on $[0,T]$, then

$$P\left\{\sup_{t\in[0,T]} X_t \geq \lambda\right\} \leq \frac{1}{\lambda}E(X_0) \ . \tag{4.20}$$

This inequality has a significant improvement stated as follows.

Theorem 4.6.10 *Let $M = (M_t)_{t\geq 0}$ be a continuous square-integrable martingale with $M_0 = 0$. Suppose there is a (deterministic) continuous, increasing function $a = a(t)$ such that $a(0) = 0$, $\langle M\rangle_t \leq a(t)$ for all $t \in [0,T]$. Then*

$$P\left\{\sup_{t\in[0,T]} M_t \geq \lambda a(T)\right\} \leq e^{-\frac{\lambda^2}{2}a(T)} \ . \tag{4.21}$$

Proof. For every $\alpha > 0$ and $t \leq T$

$$\begin{aligned} \alpha M_t - \frac{\alpha^2}{2}\langle M\rangle_t &\geq \alpha M_t - \frac{\alpha^2}{2}\langle M\rangle_T \\ &\geq \alpha M_t - \frac{\alpha^2}{2}a(T) \end{aligned}$$

so that

$$\mathcal{E}(\alpha M)_t \geq e^{\alpha M_t - \frac{\alpha^2}{2}a(T)} \qquad \text{for } \alpha > 0 \ .$$

Hence, by applying Doob's maximal inequality to the non-negative super-martingale $\mathcal{E}(\alpha M)$ we obtain

$$\begin{aligned} P\left\{\sup_{t\in[0,T]} M_t \geq \lambda a(T)\right\} &\leq P\left\{\sup_{t\in[0,T]} \mathcal{E}(\alpha M)_t \geq e^{\alpha\lambda a(T) - \frac{\alpha^2}{2}a(T)}\right\} \\ &\leq e^{-\alpha\lambda a(T) + \frac{\alpha^2}{2}a(T)} E\left\{\mathcal{E}(\alpha M)_0\right\} \\ &= e^{-\alpha\lambda a(T) + \frac{\alpha^2}{2}a(T)} \end{aligned}$$

for any $\alpha > 0$. The exponential inequality follows by setting $\alpha = \lambda$. ■

In particular, by applying the exponential inequality to a standard Brownian motion $B = (B_t)_{t\geq 0}$,

$$P\left\{\sup_{t\in[0,T]} B_t \geq \lambda T\right\} \leq e^{-\frac{\lambda^2}{2}T} \ . \tag{4.22}$$

4.6.5 Girsanov's theorem

We are given a filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$. Let $T > 0$, and Q be a probability measure on $(\Omega, \mathcal{F}_T)$ such that

$$\left.\frac{dQ}{dP}\right|_{\mathcal{F}_T} = \xi$$

for some non-negative random variable $\xi \in L^1(\Omega, \mathcal{F}_T, P)$. By definition, for any bounded $\mathcal{F}_T$-measurable random variable X

$$\int_\Omega X(\omega)Q(d\omega) = \int_\Omega X(\omega)\xi(\omega)P(d\omega)$$

or simply written as

$$E^Q(X) = E^P(\xi X) \ .$$

If, however, X is $\mathcal{F}_t$-measurable, $t \leq T$, then

$$\begin{aligned} E^Q(X) &= E^P(E^P(\xi X|\mathcal{F}_t)) \\ &= E^P(E^P(\xi|\mathcal{F}_t) X) \ . \end{aligned}$$

That is, for every $t \leq T$

$$\left.\frac{dQ}{dP}\right|_{\mathcal{F}_t} = E^P(\xi|\mathcal{F}_t)$$

which is a non-negative martingale up to T under the probability P.

Conversely, if $T > 0$ and $Z = (Z_t)_{t\geq 0}$ is a continuous, positive martingale up to time T, with $Z_0 = 1$, on a filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$. We define a measure Q on $(\Omega, \mathcal{F}_T)$ by

$$Q(A) = P(Z_T A) \quad \text{if } A \in \mathcal{F}_T \ . \tag{4.23}$$

That is, $\left.\frac{dQ}{dP}\right|_{\mathcal{F}_T} = Z_T$. Q is a probability measure on $(\Omega, \mathcal{F}_T)$ as $E(Z_T) = 1$. Since $(Z_t)_{t\leq T}$ is a martingale up to time T, so that $\left.\frac{dQ}{dP}\right|_{\mathcal{F}t} = Z_t$ for all $t \leq T$.

If $(Z_t)_{t\geq 0}$ is a positive martingale with $Z_0 = 1$, then there is a probability measure Q on $(\Omega, \mathcal{F}_\infty)$, where $\mathcal{F}_\infty \equiv \sigma\{\mathcal{F}_t : t \geq 0\}$, such that $\left.\frac{dQ}{dP}\right|_{\mathcal{F}_t} = Z_t$ for all $t \geq 0$. In this case,

We are now in a position to prove Girsanov's theorem.

Theorem 4.6.11 *(Girsanov's theorem) Let* $(M_t)_{t\geq 0}$ *be a continuous local martingale on* $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ *up to time* T*. Then*

$$X_t = M_t - \int_0^t \frac{1}{Z_s} d\langle M, Z\rangle_s$$

is a continuous local martingale on $(\Omega, \mathcal{F}, \mathcal{F}_t, Q)$ *up to time* T*.*

Proof. Using localization technique, we may assume that $M, Z, 1/Z$ are all bounded. In this case M, Z are bounded martingales. We want to prove that X is a martingale under the probability Q:

$$Q\{X_t|F_s\} = X_s \quad \text{for all} \quad s < t \leq T \ ,$$

that is,

$$Q\{1_A(X_t - X_s)\} = 0 \quad \text{for all} \quad s < t \leq T \ , A \in \mathcal{F}_s \ .$$

By definition

$$Q\{1_A(X_t - X_s)\} = P\{(Z_tX_t - Z_sX_s)1_A\}$$

thus we only need to show that (Z_tX_t) is a martingale up to time T under probability measure P. By use of integration by parts, we have

$$\begin{aligned} Z_tX_t &= Z_0X_0 + \int_0^t Z_s \mathrm{d}X_s + \int_0^t X_s \mathrm{d}Z_s + \langle Z, X\rangle_t \\ &= Z_0X_0 + \int_0^t Z_s\left(\mathrm{d}M_s - \frac{1}{Z_s}\mathrm{d}\langle M, Z\rangle_s\right) \\ &\quad + \int_0^t X_s \mathrm{d}Z_s + \langle Z, X\rangle_t \\ &= Z_0X_0 + \int_0^t Z_s \mathrm{d}M_s + \int_0^t X_s \mathrm{d}Z_s \end{aligned}$$

which is a local martingale. ■

Since $Z_t > 0$ is a positive martingale up to time T, we may apply the Itô formula to $\log Z_t$, to obtain

$$\log Z_t - \log Z_0 = \int_0^t \frac{1}{Z_s} \mathrm{d}Z_s - \int_0^t \frac{1}{Z_s^2} \mathrm{d}\langle Z \rangle_s \ ,$$

that is, $Z_t = \mathcal{E}(N)_t$ with

$$N_t = \int_0^t \frac{1}{Z_s} \mathrm{d}Z_s$$

is a continuous local martingale. Hence $Z_t = \mathcal{E}(N)_t$ solves the Itô integral equation

$$Z_t = 1 + \int_0^t Z_s \mathrm{d}N_s \ ,$$

and therefore

$$\langle M, Z \rangle_t = \langle \int_0^t \mathrm{d}M_s, \int_0^t Z_s \mathrm{d}N_s \rangle = \int_0^t Z_s \mathrm{d}\langle N, M \rangle_s \ .$$

It follows thus that

$$\int_0^t \frac{1}{Z_s} \mathrm{d} \left\langle M, Z \right\rangle_s = \langle N, M \rangle_t \ .$$

Corollary 4.6.12 *Let N_t be a continuous local martingale on $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$, $N_0 = 0$, such that its stochastic exponential $\mathcal{E}(N)_t$ is a continuous martingale up to T. Define a probability measure Q on the measurable space $(\Omega, \mathcal{F}_T)$ by*

$$\left.\frac{dQ}{dP}\right|_{\mathcal{F}_t} = \mathcal{E}(N)_t \quad for \quad all \quad t \leq T.$$

If $M = (M_t)_{t \geq 0}$ is a continuous local martingale under the probability P, then

$$X_t = M_t - \langle N, M \rangle_t$$

is a continuous, local martingale under Q up to time T. (You should carefully define the concept of a local martingale up to time T).

4.7 The martingale representation theorem

The martingale representation theorem is a deep result about Brownian motion. There is a natural version for multi-dimensional Brownian motion, for simplicity of notations, we however concentrate on one-dimensional Brownian motion.

Let $B = (B_t)_{t\geq 0}$ be a standard Brownian motion in $\mathbb{R}$ on a complete probability space $(\Omega, \mathcal{F}, P)$, and $(\mathcal{F}_t^0)_{t\geq 0}$ (together with $\mathcal{F}_\infty^0 = \cup \mathcal{F}_t^0$) be the filtration generated by the Brownian motion $(B_t)_{t\geq 0}$. Let $\mathcal{F}_t$ be the completion, and $\mathcal{F}_\infty = \cup \mathcal{F}_t$. As a matter of fact, $(\mathcal{F}_t)_{t\geq 0}$ is continuous.

Theorem 4.7.1 *Let $M = (M_t)_{t\geq 0}$ be a square-integrable martingale on $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$. Then there is a stochastic process $F = (F_t)_{t\geq 0}$ in $\mathcal{L}_2$, such that*

$$M_t = E(M_0) + \int_0^t F_s dB_s \qquad a.s.$$

for any $t \geq 0$. In particular, any martingale with respect to the Brownian filtration $(\mathcal{F}_t)_{t\geq 0}$ has a continuous version.

The proof of this theorem relies on the following several lemmata. Let $T > 0$ be any fixed time.

Lemma 4.7.2 *The following collection of random variables on $(\Omega, \mathcal{F}_T, P)$*

$$\left\{ \phi(B_{t_1}, \cdots, B_{t_k}) : \forall k \in \mathbb{Z}_+, \ \ t_j \in [0,T] \ \textit{and} \ \phi \in C_0^\infty(\mathbb{R}^k) \right\}$$

is dense in $L^2(\Omega, \mathcal{F}_T, P)$.

Proof. If $X \in L^2(\Omega, \mathcal{F}_T, P)$, then, by definition, there is an $\mathcal{F}_T^0$-measurable function which equals X almost surely. Therefore, without losing generality, we may assume that $X \in L^2(\Omega, \mathcal{F}_T^0, P)$. According to definition, $\mathcal{F}_T^0 = \sigma\{B_t : t \leq T\}$. Let $D = \mathbb{Q} \cap [0,T]$ the set of all rational numbers in the interval $[0,T]$. Since D is dense in $[0,T]$, so that $\mathcal{F}_T^0 = \sigma\{B_t : t \in D\}$. Moreover D countable, so that we may write $D = \{t_1, \cdots, t_n, \cdots\}$. Let $D_n = \{t_1, \cdots, t_n\}$ for each n, and $\mathcal{G}_n = \sigma\{B_{t_1}, \cdots, B_{t_n}\}$. Then $\{\mathcal{G}_n\}$ is increasing, and $\mathcal{G}_n \uparrow \mathcal{F}_T^0$. Let $X_n = E(X|\mathcal{G}_n)$. Then $(X_n)_{n\geq 1}$ is square-integrable martingale, thus, according to the martingale convergence theorem

$$X_n \to X \quad \text{almost surely.}$$

Moreover $X_n \to X$ in L^2. While, for each n, X_n is measurable with respect to $\mathcal{G}_n$, so that

$$X_n = f_n(B_{t_1}, \cdots, B_{t_n})$$

for some Borel measurable function $f_n : \mathbb{R}^n \to \mathbb{R}$. Since $X_n \in L^2$, so that $f_n \in L^2(\mathbb{R}^n, \mu)$ where μ is a Gaussian measure such that

$$EX_n^2 = \int_{\mathbb{R}^n} f(x)^2 \mu(dx) \ .$$

Since $C_0^\infty(\mathbb{R}^n)$ is dense in $L^2(\mathbb{R}^n, \mu)$, for each n, there is a sequence $\{\phi_{nk}\}$ in $C_0^\infty(\mathbb{R}^n)$ such that $\phi_{nk} \to f_n$ in $L^2(\mathbb{R}^n, \mu)$. It follows that

$$\phi_{nn}(B_{t_1}, \cdots, B_{t_n}) \to X$$

in L^2. ■

If $I \subset \mathbb{R}$ is an interval, then we use $L^2(I)$ to denote the Hilbert space of all functions h on I which are square-integrable.

Lemma 4.7.3 *Let $T > 0$. For any $h \in L^2([0,T])$, we associate with an exponential martingale up to time T:*

$$M(h)_t = \exp\left\{ \int_0^t h(s) dB_s - \frac{1}{2} \int_0^t h(s)^2 ds \right\} \ ; \quad t \in [0,T]. \tag{4.24}$$

Then $\mathbb{L} =$span$\{M(h)_T : h \in L^2([0,T])\}$ is dense in $L^2(\Omega, \mathcal{F}_T, P)$.

Proof. The conclusion will follow if we can prove the following: if $H \in L^2(\Omega, \mathcal{F}_T, P)$ such that

$$\int_\Omega H \Phi \mathrm{d}P = 0 \qquad \text{for all } \ \Phi \in \mathbb{L},$$

then $H = 0$.

For any $0 = t_0 < t_1 < \cdots < t_n = T$ and $c_i \in \mathbb{R}$, consider a step function $h(t) = c_i$ for $t \in (t_i, t_{i+1}]$. Then

$$M(h)_T = \exp\left\{ \sum_i c_i (B_{t_{i+1}} - B_{t_i}) - \frac{1}{2} \sum_i c_i^2 (t_{i+1} - t_i) \right\} \ .$$

Since $\int_\Omega H \Phi \mathrm{d}P = 0$ for any $\Phi \in \mathbb{L}$, so that

$$\int_\Omega H \exp\left\{ \sum_i c_i (B_{t_{i+1}} - B_{t_i}) - \frac{1}{2} \sum_i c_i^2 (t_{i+1} - t_i) \right\} \mathrm{d}P = 0 \ .$$

The deterministic, positive term $e^{-\frac{1}{2}\sum_i c_i^2(t_{i+1}-t_i)}$ can be removed from the integrand, and it follows therefore that

$$\int_\Omega H \exp\left\{\sum_i c_i(B_{t_{i+1}} - B_{t_i})\right\} \mathrm{d}P = 0 \ .$$

Since c_i are arbitrary numbers, hence

$$\int_\Omega H \exp\left\{\sum_i c_i B_{t_i}\right\} \mathrm{d}P = 0$$

for any c_i and $t_i \in [0,T]$. Since the left-hand is analytic in c_i, so that the equality remains true for any complex numbers c_i. If $\phi \in C_0^\infty(\mathbb{R}^n)$, then

$$\phi(x) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \hat{\phi}(z) e^{i\langle z,x\rangle} \mathrm{d}z$$

where

$$\hat{\phi}(z) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \phi(x) e^{-i\langle z,x\rangle} \mathrm{d}x$$

is the Fourier transform of ϕ. Hence

$$\begin{aligned}
\int_\Omega H\phi(B_{t_1}, \cdots, B_{t_n})\mathrm{d}P &= \frac{1}{(2\pi)^{n/2}} \int_\Omega \left\{ H \int_{\mathbb{R}^n} \hat{\phi}(z) \exp\left(i \sum_j z_j B_{t_j}\right)\right\} \mathrm{d}z\mathrm{d}P \\
&= \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} \left\{ \hat{\phi}(z) \int_\Omega H \exp\left(i \sum_i z_i B_{t_i}\right) \mathrm{d}P \right\} \mathrm{d}z \\
&= 0 \ .
\end{aligned}$$

Therefore, for any $\phi \in C_0^\infty(\mathbb{R}^n)$,

$$\int_\Omega H\phi(B_{t_1}, \cdots, B_{t_n})\mathrm{d}P = 0 \ . \tag{4.25}$$

By Lemma 4.7.2, the collection of all functions like $\phi(B_{t_1}, \cdots, B_{t_n})$ is dense in $L^2(\Omega, \mathcal{F}_T, P)$, so that

$$\int_\Omega HG\mathrm{d}P = 0 \quad \text{for any} \quad G \in L^2(\Omega, \mathcal{F}_T, P) \ .$$

In particular, $\int_\Omega H^2 \mathrm{d}\mathbb{P} = 0$ so that $H = 0$. ∎

Theorem 4.7.4 *(Itô's representation theorem) Let* $\xi \in L^2(\Omega, \mathcal{F}_T, P)$*. Then there is a* $F = (F_t)_{t \geq 0} \in \mathcal{L}_2$*, such that*

$$\xi = E(\xi) + \int_0^T F_t dB_t \ .$$

Proof. By Lemma 4.7.3 we only need to show this lemma for $\xi = X(h)_T$ (where $h \in L^2([0,T])$) defined by (4.24). While, $X(h)_t$ is an exponential martingale so that it must satisfy the following integral equation

$$\begin{aligned} X(h)_T &= 1 + \int_0^T X(h)_t \mathrm{d}\left(\int_0^t h(s)\mathrm{d}B_s\right) \\ &= E(X(h)_T) + \int_0^T X(h)_t h(t)\mathrm{d}B_t \ . \end{aligned}$$

Therefore $F_t = X(h)_t h(t)$ will do. ■

The martingale representation theorem now follows easily from the martingale property and Itô's representation theorem.

Chapter 5

Stochastic differential equations

The main goal of the chapter is to establish the basic existence and uniqueness theorem for a class of stochastic differential equations which are most important in application.

5.1 Introduction

Stochastic differential equations (SDE) are ordinary differential equations perturbed by noises. We will consider a simple class of noises modelled by Brownian motion. Thus we consider the following type of equation

$$dX_t^j = \sum_{i=1}^n f_i^j(t, X_t)dB_t^i + f_0^j(t, X_t)dt \;, \quad j = 1, \cdots, N \tag{5.1}$$

where $B_t = (B_t^1, \cdots, B_t^n)_{t\geq 0}$ is a standard Brownian motion in $\mathbb{R}^n$ on a filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$, and

$$f_i^j : [0, +\infty) \times \mathbb{R}^N \to \mathbb{R}^N$$

are Borel measurable functions. Of course, differential equation (5.1) should be interpreted as an integral equation in terms of Itô's integration. More precisely, an adapted, continuous, $\mathbb{R}^N$-valued stochastic process $X_t \equiv (X_t^1, \cdots, X_t^N)$ is a solution of (5.1), if

$$X_t^j = X_0^j + \sum_{k=1}^n \int_0^t f_k^j(s, X_s)dB_s^k + \int_0^t f_0^j(s, X_s)ds \tag{5.2}$$

for $j = 1, \cdots, N$. Since we are concerned only with the distribution determined by the solution $(X_t)_{t\geq 0}$ of (5.1), we therefore expect that any solution of SDE (5.1) should have the same distribution for *any* Brownian motion $B = (B_t)_{t\geq 0}$. It thus leads to different concepts of solutions and uniqueness: strong solutions and weak solutions, path-wise uniqueness and uniqueness in law.

Definition 5.1.1 *1) An adapted, continuous, $\mathbb{R}^N$-valued stochastic process $X = (X_t)_{t\geq 0}$ on $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ is a (weak) solution of (5.1), if there is a Brownian motion $W = (W_t)_{t\geq 0}$ in $\mathbb{R}^n$, adapted to the filtration $(\mathcal{F}_t)$, such that*

$$X_t^j - X_0^j = \sum_{l=1}^n \int_0^t f_l^j(s, X_s) dW_s^l + \int_0^t f_0^j(s, X_s) ds, \quad j = 1, \cdots, N.$$

In this case we also call the pair (X, W) a (weak) solution of (5.1).

2) Given a standard Brownian motion $B = (B_t)_{t\geq 0}$ in $\mathbb{R}^n$ on $(\Omega, \mathcal{F}, P)$ with its natural filtration $(\mathcal{F}_t)_{t\geq 0}$, an adapted, continuous stochastic process $X = (X_t)_{t\geq 0}$ on $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ is a strong solution *of (5.1), if*

$$X_t^j - X_0^j = \sum_{i=1}^n \int_0^t f_i^j(s, X_s) dB_s^i + \int_0^t f_0^j(s, X_s) ds\ .$$

We also have different concepts of uniqueness.

Definition 5.1.2 *Consider SDE (5.1).*

1. *We say that the **path-wise uniqueness** holds for (5.1), if whenever (X, B) and $(\widetilde{X}, B)$ are two solutions defined on the same filtered space and same Brownian motion B, and $X_0 = \widetilde{X}_0$, then $X = \widetilde{X}$.*

2. *It is said that **uniqueness in law** holds for (5.1), if (X, B) and $(\widetilde{X}, \widetilde{B})$ are two solutions (with possibly different Brownian motions B and $\widetilde{B}$, even can be on different probability spaces), and X_0 and $\widetilde{X}_0$ possess the same distribution, then X and $\widetilde{X}$ have same distribution.*

Theorem 5.1.3 *(Yamada-Watanabe) Path-wise uniqueness implies uniqueness in law.*

The following is a simple example of SDE for which has no strong solution, but possesses weak solutions and uniqueness in law holds.

Example 5.1.4 *(H. Tanaka) Consider 1-dimensional stochastic differential equation:*

$$X_t = \int_0^t sgn(X_s) dB_s \ , \qquad 0 \le t < \infty$$

where $sgn(x) = 1$ *if* $x \ge 0$*, and equals* -1 *for negative value of* x*.*

1. Uniqueness in law holds, since X is a standard Brownian motion (Lévy's Theorem).

2. If (X, B) is a weak solution, then so is $(-X, B)$.

3. There is a weak solution. Let W_t be a one-dimensional Brownian motion, and let $B_t = \int_0^t \mathrm{sgn}(W_s) dW_s$. Then B is a one-dimensional Brownian motion, and

$$W_t = \int_0^t \mathrm{sgn}(W_s) dB_s \ ,$$

so that (W, B) is a solution.

4. Path-wise uniqueness does not hold.

5. There is no any strong solution.

5.2 Several examples

5.2.1 Linear-Gaussian diffusions

Linear stochastic differential equations can be solved explicitly. Consider

$$dX_t^j = \sum_{i=1}^n \sigma_i^j dB_t^i + \sum_{k=1}^N \beta_k^j X_t^k dt \tag{5.3}$$

$(j = 1, \cdots, N)$, where B is a Brownian motion in $\mathbb{R}^n$, $\sigma = (\sigma_i^j)$ a constant $N \times n$ matrix, and $\beta = (\beta_k^j)$ a constant $N \times N$ matrix. (5.3) may be written as

$$dX_t = \sigma dB_t + \beta X_t dt \ .$$

Let

$$e^{\beta t} = \sum_{k=0}^\infty \frac{t^k}{k!} \beta^k$$

be the exponential of the square matrix β. Using Itô's formula, we have

$$\begin{aligned} e^{-\beta t}X_t - X_0 &= \int_0^t e^{-\beta s}dX_s - \int_0^t e^{-\beta s}\beta X_s ds \\ &= \int_0^t e^{-\beta s}(dX_s - \beta X_s ds) \\ &= \int_0^t e^{-\beta s}\sigma dB_s \end{aligned}$$

so that

$$X_t = e^{\beta t}X_0 + \int_0^t e^{\beta(t-s)}\sigma \mathrm{d}B_s \ .$$

In particular, if $X_0 = x$, then X_t has a normal distribution with mean $e^{\beta t}x$. For example, if $n = N = 1$, then

$$X_t \sim N(e^{\beta t}x, \frac{\sigma^2}{2}\left(e^{2\beta t} - 1\right)) \ .$$

It can be shown that (X_t) is a diffusion process, and thus its distribution can be described by its transition probability $P_t(x, dz)$. According to definition

$$\begin{aligned} (P_t f)(x) &\equiv \int_{\mathbb{R}^N} f(z)P_t(x, dz) \\ &= E\left(f(X_t)|X_0 = x\right) \ , \end{aligned}$$

thus

$$\begin{aligned} (P_t f)(x) &= E\left(f(X_t)|X_0 = x\right) \\ &= \int_{\mathbb{R}} f(z)\frac{1}{\sqrt{2\pi\frac{\sigma^2}{2}\left(e^{2\beta t} - 1\right)}} \exp\left(-\frac{|z - e^{\beta t}x|^2}{\frac{\sigma^2}{2}\left(e^{2\beta t} - 1\right)}\right) dz \\ &= \int_{\mathbb{R}} f(e^{\beta t}x + \sqrt{\frac{\sigma^2}{2}\left(e^{2\beta t} - 1\right)}z)\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{|z|^2}{2}\right) dz \\ &= Ef(e^{\beta t}x + \sqrt{\frac{\sigma^2}{2}\left(e^{2\beta t} - 1\right)}\xi) \end{aligned}$$

where ξ has the standard normal distribution $N(0, 1)$. From the second line of the above formula, and compare to the definition of $P_t(x, dz)$, we can conclude that

$$P_t(x, dz) = p(t, x, z)dz$$

with

$$p(t,x,z) = \frac{1}{\sqrt{2\pi\frac{\sigma^2}{2}\left(e^{2\beta t}-1\right)}} \exp\left(-\frac{|z-e^{\beta t}x|^2}{\frac{\sigma^2}{2}\left(e^{2\beta t}-1\right)}\right) .$$

$p(t,x,z)$ is called the transition density of the diffusion process $(X_t)_{t\geq 0}$. We thus, again following from the above formula, have a probability representation

$$(P_t f)(x) = Ef(e^{\beta t}x + \sqrt{\frac{\sigma^2}{2}\left(e^{2\beta t}-1\right)}\xi)$$

which is quite useful in some computations.

Remark 5.2.1 *It is easy to see from the above representation that*

$$\frac{d}{dx}(P_t f) = e^{\beta t} P_t\left(\frac{d}{dx}f\right) .$$

The distribution of (X_t) is determined by the transition density $p(t,x,z)$. Indeed, for any $0 < t_1 < \cdots < t_k$, the joint distribution of $(X_{t_1},\cdots,X_{t_k})$ is Gaussian, and indeed its pdf is

$$p(t_1,x,z_1)p(t_2-t_1,z_1,z_2)\cdots p(t_k-t_{k-1},z_{k-1},z_k) .$$

If $B=(B_t^1,\cdots,B_t^n)_{t\geq 0}$ is a Brownian motion in $\mathbb{R}^n$, then the solution X_t of the SDE:

$$dX_t = dB_t - (AX_t)\,dt$$

is called the Ornstein-Uhlenbeck process, where $A \geq 0$ is a $d\times d$ matrix called the drift matrix. Hence we have

$$X_t = e^{-At}X_0 + \int_0^t e^{-(t-s)A}dB_s .$$

Exercise 5.2.2 *If $X_0 = x \in \mathbb{R}^n$, compute $Ef(X_t)$, where X_t is the Ornstein-Uhlenbeck process with drift matrix A.*

5.2.2 Geometric Brownian motion

The Black-Scholes model satisfies the stochastic differential equation

$$dS_t = S_t\left(\mu dt + \sigma dB_t\right) \tag{5.4}$$

and so that the solution to (5.4) is the stochastic exponential of

$$\int_0^t \mu ds + \int_0^t \sigma dB_s .$$

Hence

$$S_t = S_0 \exp\left(\int_0^t \sigma dB_s + \int_0^t \left(\mu - \frac{1}{2}\sigma^2\right) ds\right) .$$

In the case σ and μ are constants, then

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

which is called the geometric Brownian motion. If $S_0 = x > 0$, then S_t remains positive, and

$$\log S_t = \log x + \sigma B_t + \left(\mu - \frac{1}{2}\sigma^2\right) t$$

has a normal distribution with mean $\log x + \left(\mu - \frac{1}{2}\sigma^2\right) t$ and variance σ^2. Again, as a solution to the stochastic differential equation (5.4), $(S_t)_{t\geq 0}$ is a diffusion process, its distribution is determined by its transition function $P_t(x, dz)$ (unfortunately we have to use the same notations as in the last sub-section), and according to definition

$$\begin{aligned}
\int_{\mathbb{R}} f(z) P_t(x, dz) &= E\left(f(X_t) | X_0 = x\right) \\
&= E\left(f(xe^{\sigma B_t + \left(\mu - \frac{1}{2}\sigma^2\right)t})\right) \\
&= \int_{\mathbb{R}} f(xe^{\sigma z + \left(\mu - \frac{1}{2}\sigma^2\right)t}) \frac{1}{\sqrt{2\pi t}} e^{-\frac{z^2}{2\pi t}} dz \\
&= \int_0^\infty f(y) \frac{1}{\sqrt{2\pi t}} \frac{1}{\sigma y} e^{-\frac{1}{2\pi t}\left(\frac{1}{\sigma}\log\frac{y}{x} - \left(\frac{\mu}{\sigma} - \frac{1}{2}\sigma\right)\right)^2} dy
\end{aligned}$$

where we assume that $\sigma > 0$ and have made the change of variable

$$xe^{\sigma z + \left(\mu - \frac{1}{2}\sigma^2\right)t} = y .$$

As usual, we define $(P_t f)(x) = \int_{\mathbb{R}} f(z) P_t(x, dz)$. By the third line of the previous formula

$$\begin{aligned}
(P_t f)(x) &= \int_{\mathbb{R}} f(xe^{\sigma z + \left(\mu - \frac{1}{2}\sigma^2\right)t}) \frac{1}{\sqrt{2\pi t}} e^{-\frac{z^2}{2\pi t}} dz \\
&= \int_{\mathbb{R}} f(xe^{\sigma\sqrt{t}y + \left(\mu - \frac{1}{2}\sigma^2\right)t}) \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2\pi}} dy \\
&= E\left(f(xe^{\sigma\sqrt{t}\xi + \left(\mu - \frac{1}{2}\sigma^2\right)t})\right)
\end{aligned}$$

[we have made a change variable z into $\sqrt{t}y$], where $\xi \sim N(0,1)$. Compare with the definition of $P_t(x,dy)$ we have

$$P_t(x,dy) = \frac{1}{\sqrt{2\pi t}}\frac{1}{\sigma y}e^{-\frac{1}{2\pi t}\left(\frac{1}{\sigma}\log\frac{y}{x}-\left(\frac{\mu}{\sigma}-\frac{1}{2}\sigma\right)\right)^2}dy \quad \text{on } (0,+\infty)$$

That is, (S_t) has the transition density

$$p(t,x,y) = \frac{1}{\sqrt{2\pi t}}\frac{1}{\sigma y}e^{-\frac{1}{2\pi t}\left(\frac{1}{\sigma}\log\frac{y}{x}-\left(\frac{\mu}{\sigma}-\frac{1}{2}\sigma\right)\right)^2} \quad \text{on } (0,+\infty)\ .$$

and, therefore, for geometric Brownian motion

$$(P_t f)(x) = \int_0^\infty \frac{1}{\sqrt{2\pi t}}\frac{1}{\sigma y}e^{-\frac{1}{2\pi t}\left(\frac{1}{\sigma}\log\frac{y}{x}-\left(\frac{\mu}{\sigma}-\frac{1}{2}\sigma\right)\right)^2}f(y)dy$$

for any $x > 0$.

5.2.3 Cameron-Martin's formula

Consider a simple stochastic differential equation

$$dX_t = dB_t + b(t,X_t)dt \tag{5.5}$$

where $b(t,x)$ is a bounded, Borel measurable function on $[0,+\infty)\times\mathbb{R}$. We may solve (5.5) by means of *change of probabilities.*

Let $(W_t)_{t\geq 0}$ be a standard Brownian motion on $(\Omega,\mathcal{F},\mathcal{F}_t,P)$, and define probability measure Q on $(\Omega,\mathcal{F}_\infty)$ by

$$\left.\frac{dQ}{dP}\right|_{\mathcal{F}_t} = \mathcal{E}(N)_t \qquad \text{for all} \quad t\geq 0$$

where $N_t = \int_0^t b(s,W_s)\mathrm{d}W_s$ is a martingale (under the probability P), with $\langle N\rangle_t = \int_0^t b(s,W_s)^2\mathrm{d}s$, which is bounded on any finite interval. Hence

$$\mathcal{E}(N)_t = \exp\left(\int_0^t b(s,W_s)\mathrm{d}W_s - \frac{1}{2}\int_0^t b(s,W_s)^2\mathrm{d}s\right)$$

is a martingale. According to Girsanov's theorem

$$B_t \equiv W_t - W_0 - \langle W,N\rangle_t$$

is a martingale under the new probability Q, and $\langle B\rangle_t = \langle W\rangle_t = t$. By Lévy's martingale characterization of Brownian motion, $(B_t)_{t\geq 0}$ is a Brownian motion. Moreover

$$\begin{aligned}\langle W, N\rangle_t &= \langle \int_0^t \mathrm{d}W_s, \int_0^t b(s,W_s)\mathrm{d}W_s\rangle \\ &= \int_0^t b(s,W_s)\mathrm{d}s\end{aligned}$$

and therefore

$$W_t - W_0 - \int_0^t b(s,W_s)\mathrm{d}s = B_t$$

is a standard Brownian motion on $(\Omega, \mathcal{F}, Q)$. Thus

$$W_t = W_0 + B_t + \int_0^t b(s,W_s)\mathrm{d}s \tag{5.6}$$

so that $(W_t)_{t\geq 0}$ on $(\Omega, \mathcal{F}_\infty, Q)$ is a solution of (5.5). The solution we have just constructed is a weak solution of SDE (5.5).

Theorem 5.2.3 *(Cameron-Martin's formula) Let $b(t,x) = (b^1(t,x), \cdots, b^n(t,x))$ be bounded, Borel measurable functions on $[0,+\infty)\times\mathbb{R}^n$. Let $W_t = (W^1, \cdots, W_t^n)$ be a standard Brownian motion on a filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$, and let $\mathcal{F}_\infty = \sigma\{\mathcal{F}_t, t\geq 0\}$. Define probability measure Q on $(\Omega, \mathcal{F}_\infty)$ by*

$$\left.\frac{dQ}{dP}\right|_{\mathcal{F}_t} = e^{\sum_{k=1}^n \int_0^t b^k(s,W_s)dB_s^k - \frac{1}{2}\sum_{k=1}^n \int_0^t \left|b^k(s,W_s)\right|^2 ds} \quad \textit{for} \quad t\geq 0\ .$$

Then $(W_t)_{t\geq 0}$ under the probability measure Q is a solution to

$$dX_t^j = dB_t^j + b^j(t,X_t)dt \tag{5.7}$$

for some Brownian motion $(B_t^1, \cdots, B_t^n)_{t\geq 0}$ under probability Q.

On the other hand, if (X_t) is a solution of SDE (5.7) on some probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ and define $\tilde{P}$

$$\left.\frac{d\tilde{P}}{dP}\right|_{\mathcal{F}_t} = \exp\left\{-\sum_{k=1}^n \int_0^t b^k(s,X_s)\mathrm{d}B_s^k - \frac{1}{2}\sum_{k=1}^n \int_0^t \left|b^k(s,X_s)\right|^2 \mathrm{d}s\right\} \quad \text{for} \quad t\geq 0$$

we may show that $(X_t)_{t\geq 0}$ under probability measure $\tilde{P}$ is a Brownian motion. Therefore solutions to SDE (5.7) is unique in law: all solutions have the same distribution.

5.3 Existence and uniqueness

In this section we present a fundamental result about the existence and uniqueness of strong solutions.

5.3.1 Strong solutions: existence and uniqueness

By definition, any strong solution is a weak solution. We next prove a basic existence and uniqueness theorem for a stochastic differential equation under a global Lipschitz condition. Our proof will rely on two inequalities: The Gronwall inequality and Doob's L^p-inequality (Theorem 2.2.5).

Lemma 5.3.1 *(The Gronwall inequality) If a non-negative function g satisfies the integral equation*

$$g(t) \leq h(t) + \alpha \int_0^t g(s)ds \ , \qquad 0 \leq t \leq T$$

where α is a constant and $h : [0,T] \to \mathbb{R}$ is an integrable function, then

$$g(t) \leq h(t) + \alpha \int_0^t e^{\alpha(t-s)} h(s)ds \quad , \qquad 0 \leq t \leq T \ .$$

Proof. Let $F(t) = \int_0^t g(s)\mathrm{d}s$. Then $F(0) = 0$ and

$$F'(t) \leq h(t) + \alpha F(t)$$

so that

$$\left(e^{-\alpha t}F(t)\right)' \leq e^{-\alpha t}h(t) \ .$$

Integrating the differential inequality we obtain

$$\int_0^t \left(e^{-\alpha s}F(s)\right)' \mathrm{d}s \leq \int_0^t e^{-\alpha s}h(s)\mathrm{d}s$$

and therefore

$$F(t) \leq \int_0^t e^{\alpha(t-s)}h(s)\mathrm{d}s$$

which yields Gronwall's inequality. ■

Consider the following stochastic differential equation

$$dX_t^j = \sum_{l=1}^n f_l^j(t, X_t)\mathrm{d}B_t^l + f_0^j(t, X_t)\mathrm{d}t \ ; \quad j = 1, \cdots, N \qquad (5.8)$$

where $f_k^j(t,x)$ are Borel measurable functions on $\mathbb{R}_+ \times \mathbb{R}^N$, which are bounded on any compact subset in $\mathbb{R}^N$. We are going to show the existence and uniqueness by Picard's iteration. The main ingredient in the proof is a special case of Doob's L^p- inequality: if $(M_t)_{t\geq 0}$ is a square-integrable, continuous martingale with $M_0 = 0$, then for any $t > 0$

$$E\left\{\sup_{s\leq t}|M_s|^2\right\} \leq 4\sup_{s\leq t} E\left(|M_s|^2\right) = 4E\langle M\rangle_t \ . \tag{5.9}$$

Lemma 5.3.2 *Let $(B_t)_{t\geq 0}$ be a standard BM in R on $(\Omega, \mathcal{F}_t, \mathcal{F}, P)$, and $(Z_t)_{t\geq 0}$ and $(\tilde{Z}_t)_{t\geq 0}$ be two continuous, adapted processes. Let $f(t,x)$ be a Lipschitz function*

$$|f(t,x) - f(t,y)| \leq C|x-y| \ ; \quad \forall t \geq 0, \ \ x, y \in \mathbb{R}$$

for some constant C.

1. *Let*

$$M_t = \int_0^t f(s, Z_s)\,dB_s - \int_0^t f(s, \tilde{Z}_s)\,dB_s \qquad \forall t \geq 0 \ .$$

Then

$$E\sup_{s\leq t}|M_s|^2 \leq 4C^2 \int_0^t E\left|Z_s - \tilde{Z}_s\right|^2 ds$$

for all $t \geq 0$.

2. *If*

$$N_t = \int_0^t f(s, Z_s)\,ds - \int_0^t f(s, \tilde{Z}_s)\,ds \qquad \forall t \geq 0$$

then

$$E\sup_{s\leq t}|N_s|^2 \leq C^2 t \int_0^t E\left|Z_s - \tilde{Z}_s\right|^2 ds \qquad \forall t \geq 0 \ .$$

Proof. To prove the first statement, we notice that

$$\sup_{s\leq t}|M_s|^2 = \sup_{s\leq t}\left|\int_0^s \left(f(u, Z_u) - f(u, \tilde{Z}_u)\right) \mathrm{d}B_u\right|^2$$

so that, by Doob's L^2-inequality

$$\begin{aligned}
E\sup_{s\leq t}|M_s|^2 &= E\sup_{s\leq t}\left|\int_0^s \Big(f(u,Z_u)-f(u,\tilde{Z}_u)\Big)\,\mathrm{d}B_u\right|^2 \\
&\leq 4E\left|\int_0^t \Big(f(s,Z_s)-f(s,\tilde{Z}_s)\Big)\,\mathrm{d}B_s\right|^2 \\
&= 4E\int_0^t \left|f(s,Z_s)-f(s,\tilde{Z}_s)\right|^2 \mathrm{d}s \\
&\leq 4C^2\int_0^t E\left|Z_s-\tilde{Z}_s\right|^2 \mathrm{d}s\ .
\end{aligned}$$

Next we prove the second claim. Indeed

$$\begin{aligned}
\sup_{s\leq t}|N_s|^2 &= \sup_{s\leq t}\left|\int_0^s \Big(f(u,Z_u)-f(u,\tilde{Z}_u)\Big)\,\mathrm{d}u\right|^2 \\
&\leq \left(\int_0^t \left|f(s,Z_s)-f(s,\tilde{Z}_s)\right|\mathrm{d}s\right)^2 \\
&\leq t\int_0^t \left|f(s,Z_s)-f(s,\tilde{Z}_s)\right|^2 \mathrm{d}s \\
&\leq C^2 t\int_0^t \left|Z_s-\tilde{Z}_s\right|^2 \mathrm{d}s
\end{aligned}$$

where the second inequality follows from the Schwartz inequality. ■

Theorem 5.3.3 *Consider SDE (5.8). Suppose that f_i^j satisfy the Lipschitz condition:*

$$\left|f_i^j(t,x)-f_i^j(t,y)\right|\leq C|x-y| \tag{5.10}$$

and the linear-growth condition that

$$\left|f_i^j(t,x)\right|\leq C(1+|x|) \tag{5.11}$$

for $t\in\mathbb{R}^+$ and $x,y\in\mathbb{R}^N$. Then for any $\eta\in L^2(\Omega,\mathcal{F}_0,P)$ and a standard Brownian motion $B_t=(B_t^i)$ in $\mathbb{R}^n$, there is a unique strong solution (X_t) of (5.8) with $X_0=\eta$.

Proof. For simplicity, let us prove a special case of this important theorem: the existence and uniqueness to the one-dimensional stochastic differential equation

$$dX_t = f(t,X_t)\mathrm{d}B_t\ ,\quad X_0=\eta\ ,$$

and leave the details of the proof for the general case to an appendix. Like for ODE, construct an approximation solutions via Picard's iteration:

$$Y_0(t) = \eta$$

and

$$Y_{n+1}(t) = \eta + \int_0^t f(s, Y_n(s)) \mathrm{d}B_s \ ,$$

where $n = 0, 1, 2, \cdots$. We are going to show that, for every $T > 0$, the sequence $\{Y_n(t)\}$ converges to a solution $Y(t)$ uniformly on $[0, T]$ almost surely. Note that, every Y_n is a continuous square-integrable martingale. Indeed

$$\begin{aligned} E \sup_{0 \le s \le t} |Y_1(s) - Y_0(s)|^2 &\le E \sup_{0 \le s \le t} \left(\int_0^s |f(\tau, \eta)| \mathrm{d}B_\tau \right)^2 \\ &\le 4E \int_0^t f(\tau, \eta)^2 \mathrm{d}s \\ &\le 8tC \left(1 + E\eta^2\right) \end{aligned}$$

and, for any $t \le T$,

$$\begin{aligned} E \sup_{s \le t} |Y_{n+1}(s) - Y_n(s)|^2 &= E \sup_{s \le t} \left| \int_0^s \left(f(r, Y_n(r)) - f(r, Y_{n-1}(r)) \right) \mathrm{d}B_r \right|^2 \\ &\le 4E \int_0^t \left(f(s, Y_n(s)) - f(s, Y_{n-1}(s)) \right)^2 \mathrm{d}s \end{aligned}$$

where the second inequality follows from the Kolmogorov's inequality. Since f is Lipschitz continuous, so that

$$\begin{aligned} & \int_0^t \left(f(s, Y_n(s)) - f(s, Y_{n-1}(s)) \right)^2 \mathrm{d}s \\ \le \ & C^2 \int_0^t |Y_n(s) - Y_{n-1}(s)|^2 \mathrm{d}s \\ \le \ & C^2 t \sup_{s \le t} |Y_n(t) - Y_{n-1}(t)|^2 \ . \end{aligned}$$

Combining the previous inequality, we obtain

$$E \sup_{s \le t} |Y_{n+1}(s) - Y_n(s)|^2 \le 4C^2 t E \sup_{s \le t} |Y_n(t) - Y_{n-1}(t)|^2$$

for any $t \leq T$, and therefore

$$E \sup_{s \leq t} |Y_{n+1}(s) - Y_n(s)|^2 \leq \frac{\left(4C^2\right)^n t^n}{n!} E \sup_{s \leq t} |Y_1(t) - Y_0(t)|^2$$

for all $t \leq T$. In particular

$$E \sup_{s \leq T} |Y_{n+1}(t) - Y_n(t)|^2 \leq \frac{\left(4C^2\right)^n T^n}{n!} E \sup_{s \leq T} |Y_1(t) - Y_0(t)|^2$$

so that

$$\begin{aligned}\sum_{n=0}^{\infty} E \sup_{s \leq T} |Y_{n+1}(t) - Y_n(t)|^2 &\leq \sum_{n=0}^{\infty} \frac{\left(4C^2\right)^n T^n}{n!} E \sup_{s \leq T} |Y_1(t) - Y_0(t)|^2 \\ &< \infty .\end{aligned}$$

Hence $\{Y_n : n \geq 1\}$ is a Cauchy sequence in $\mathcal{M}_c^2$, so that

$$Y_n(t) \to X_t \quad \text{uniformly on } [0, T] , \quad P\text{-a.s.}$$

It is easy to see that (X_t) is a strong solution of the stochastic differential equation.

Next we prove the uniqueness. Let Y and Z be two solutions with same Brownian motion B. Then

$$Y_t = \eta + \int_0^t f(s, Y_s) \mathrm{d}B_s$$

and

$$Z_t = \eta + \int_0^t f(s, Z_s) \mathrm{d}B_s \ .$$

Then, as in the proof of the existence,

$$E\left(|Y_t - Z_t|^2\right) \leq 4C^2 \int_0^t E|Y_s - Z_s|^2 \mathrm{d}s$$

The Gronwall inequality implies thus that

$$E\left(|Y_t - Z_t|^2\right) = 0 \ .$$

■

Remark 5.3.4 *The iteration Y_n constructed in the proof of Theorem 5.3.3 is a function of the Brownian motion B, and $Y_n(t)$ only depends on η and B_s, $0 \leq s \leq t$.*

5.3.2 Continuity in initial conditions

Theorem 5.3.5 *Under the same assumptions as in Theorem 5.3.3. Given a BM $B = (B_t)_{t\geq 0}$ in $\mathbb{R}^n$ on $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$, let $(X^x(t))_{t\geq 0}$ be the unique strong solution of (5.8). Then $x \to X^x$ is uniformly continuous almost surely on any finite interval $[0,T]$:*

$$\lim_{\delta\downarrow 0} \sup_{|x-y|<\delta} E\left\{ \sup_{0\leq t\leq T} |X^x(t) - X^y(t)|^2 \right\} = 0 \ . \tag{5.12}$$

Proof. Let us only consider 1-dimensional case. Thus

$$X^x(t) = x + \int_0^t f_1(s, X^x(s))\mathrm{d}B_s + \int_0^t f_0(s, X^x(s))\mathrm{d}s$$

and

$$X^y(t) = y + \int_0^t f_1(s, X^y(s))\mathrm{d}B_s + \int_0^t f_0(s, X^y(s))\mathrm{d}s \ .$$

Therefore, by Doob's maximal inequality,

$$\begin{aligned}
& E\left\{ \sup_{0\leq t\leq T} |X^x(t) - X^y(t)|^2 \right\} \leq 3|x-y|^2 \\
& +3E\left\{ \sup_{0\leq t\leq T} \left| \int_0^t (f_1(s, X^x(s)) - f_1(s, X^y(s)))\mathrm{d}B_s \right|^2 \right\} \\
& +3E\left\{ \sup_{0\leq t\leq T} \left| \int_0^t (f_0(s, X^x(s)) - f_0(s, X^y(s)))\mathrm{d}s \right|^2 \right\} \\
\leq\ & 3|x-y|^2 + 12E\left\{ \left| \int_0^T (f_1(s, X^x(s)) - f_1(s, X^y(s)))\, \mathrm{d}B_s \right|^2 \right\} \\
& +3TE\left\{ \int_0^T |f_0(X^x(s)) - f_0(X^y(s))|^2 \,\mathrm{d}s \right\} \\
\leq\ & 3|x-y|^2 + 12E\left\{ \int_0^T |f_1(s, X^x(s)) - f_1(s, X^y(s))|^2 \,\mathrm{d}s \right\} \\
& +3TC^2E\left\{ \int_0^T |X^x(s) - X^y(s)|^2 \, ds \right\} \\
\leq\ & 3|x-y|^2 + 3C^2(4+T) \int_0^T E\left(|X^x(t) - X^y(t)|^2\right) \mathrm{d}t.
\end{aligned}$$

Setting

$$\Delta(t) = E\left\{\sup_{0\le s\le t} |X^x(s) - X^y(s)|^2\right\} ,$$

then we have

$$\Delta(T) \le 3|x-y|^2 + 3C^2(4+T)\int_0^T \Delta(t)dt$$

and therefore by Gronwall's inequality

$$\Delta(T) \le 6|x-y|^2 \exp(12C^2 + 3TC^2)$$

which yields (5.12). ■

5.4 Martingales and weak solutions

For simplicity, let us consider the following one-dimensional, homogenous SDE

$$dX_t = \sigma(X_t)\mathrm{d}B_t + b(X_t)\mathrm{d}t \tag{5.13}$$

where $\sigma \in C^\infty(\mathbb{R})$ is a positive smooth function with at most linear growth, and $b \in C^\infty(\mathbb{R})$ has at most linear growth. Let $X = (X_t)_{\ge 0}$ be the strong solution with initial X_0 on a filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$. If $f \in C_b^2(\mathbb{R}^N, \mathbb{R})$, then by Itô's formula

$$\begin{aligned} f(X_t) - f(X_0) &= \int_0^t f'(X_s)\mathrm{d}X_s + \frac{1}{2}\int_0^t f''(X_s)\mathrm{d}\langle X\rangle_s \\ &= \int_0^t f'(X_s)\left(\sigma(X_s)\mathrm{d}B_s + b(X_s)\mathrm{d}s\right) \\ &\quad + \frac{1}{2}\int_0^t f''(X_s)\sigma^2(X_s)\mathrm{d}s \\ &= \int_0^t \sigma(X_s)f'(X_s)\mathrm{d}B_s + \int_0^t \left\{\frac{1}{2}\sigma^2 f'' + bf'\right\}(X_s)\mathrm{d}s \ . \end{aligned}$$

Let us introduce

$$L = \frac{1}{2}\sigma(x)^2\frac{d^2}{dx^2} + b(x)\frac{d}{dx} \tag{5.14}$$

which is an elliptic differential operator of second-order. Then the previous formula may be written as

$$f(X_t) - f(X_0) = \int_0^t \sigma(X_s)f'(X_s)\mathrm{d}B_s + \int_0^t (Lf)(X_s)\mathrm{d}s \ .$$

If we set

$$M_t^f = f(X_t) - f(X_0) - \int_0^t (Lf)(X_s)\mathrm{d}B_s \ ,$$

then

$$M_t^f = \int_0^t \sigma(X_s)f'(X_s)\mathrm{d}B_s$$

is a martingale on $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$, and

$$\langle M^f, M^g\rangle_t = \int_0^t (\sigma^2 f')(X_s)\mathrm{d}s \ .$$

Lemma 5.4.1 *If $(X_t)_{t\geq 0}$ is a strong solution to SED (5.13) on $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$ (with a given Brownian motion) then for any $f \in C_b^2(\mathbb{R})$*

$$M_t^f = f(X_t) - f(X_0) - \int_0^t (Lf)(X_s)ds$$

is a martingale under the probability, where L is defined by (5.14).

For example, if $\sigma = 1$ and $b = 0$ (in this case $L = \frac{1}{2}\frac{d^2}{dx^2} = \frac{1}{2}\Delta$), then $(B_t)_{t\geq 0}$ itself is a strong solution to

$$dX_t = dB_t$$

so that

$$M_t^f = f(B_t) - f(B_0) - \frac{1}{2}\int_0^t (\Delta f)(B_s)\mathrm{d}s$$

is a martingale under P. On the other hand, Lévy's martingale characterization shows that the previous property that

$$f(B_t) - f(B_0) - \frac{1}{2}\int_0^t (\Delta f)(B_s)\mathrm{d}s$$

is martingale, which in particular implies that X_t^j and $X_i^j X_t^i - \delta_{ij}t$ are martingales, completely characterizes Brownian motion. Therefore we may believe that the martingale property of all M^f should completely the distribution of a solution $(X_t)_{t\geq 0}$ to SDE (5.13), and hence those of weak solution of (5.13). Thus we give

Definition 5.4.2 *Let L be a linear operator on $C^\infty(\mathbb{R})$. Let $(X_t)_{t\geq 0}$ be) a stochastic process on a filtered space $(\Omega, \mathcal{F}, \mathcal{F}_t, P)$. Then we say that $(X_t)_{t\geq 0}$*

together with the probability P is a solution to the L-martingale problem, if for every $f \in C_b^\infty(\mathbb{R})$

$$M_t^f \equiv f(X_t) - f(X_0) - \int_0^t Lf(X_s)ds$$

is a local martingale under the probability P.

Therefore a strong solution $(X_t)_{t\geq 0}$ of SDE (5.13) on $(\Omega, \mathcal{F}, P)$ is a solution to L-martingale problem, where L is given by (5.14):

$$M_t^f = f(X_t) - f(X_0) - \int_0^t Lf(X_s)\mathrm{d}s$$

is a martingale under P. Moreover, since

$$L(fg) - f\,(Lg) - g\,(Lf) = \sigma^2 f'$$

we thus have

$$\langle M^f, M^g \rangle_t = \int_0^t \{L(fg) - f\,(Lg) - g\,(Lf)\}\,(X_s)\mathrm{d}s\ .$$

Conversely, we can show that any solution to the L-martingale problem is a weak solution to SDE.

Theorem 5.4.3 *Let b, σ be Borel measurable functions on $\mathbb{R}$ which are bounded on any compact subset, and let*

$$L = \frac{1}{2}\sigma(x)^2\frac{d^2}{dx^2} + b(x)\frac{d}{dx}.$$

If $(X_t)_{t\geq 0}$ on $(\Omega, \mathcal{F}, P)$ is a continuous process solving the L-martingale problem: for any $f \in C_b^2(\mathbb{R})$

$$M_t^f = f(X_t) - f(X_0) - \int_0^t Lf(X_s)ds$$

is a continuous local martingale, then $(X_t)_{t\geq 0}$ on $(\Omega, \mathcal{F}, P)$ is a weak solution to SDE

$$dX_t = \sigma(X_t)dB_t + b(X_t)dt\ . \tag{5.15}$$

Let us outline the proof only, a detailed proof will be given in next section that handles the multi-dimensional case. To show $(X_t)_{t\geq 0}$ on $(\Omega, \mathcal{F}, P)$ is a weak solution, we need to construct a Brownian motion $B = (B_t)_{t\geq 0}$ such that

$$X_t = X_0 + \int_0^t \sigma(X_s)\mathrm{d}B_s + \int_0^t b(X_s)\mathrm{d}s \ . \tag{5.16}$$

The key of the proof is to compute $\langle X \rangle_t$, and the result is

$$\begin{aligned} \left\langle M^f, M^g \right\rangle_t &= \int_0^t (L(fg) - fLg - gLf)(X_s)\mathrm{d}s \\ &= \int_0^t \left(\sigma^2 \frac{\partial f}{\partial x}\frac{\partial g}{\partial x}\right)(X_s)\mathrm{d}s \ . \end{aligned}$$

In particular, if we choose $f(x) = x$ the coordinate function (and write in this case M^f as M), then

$$\langle M \rangle_t = \int_0^t (\sigma(X_s))^2 \, \mathrm{d}s$$

so that

$$B_t = \int_0^t \frac{1}{\sigma(X_s)} \mathrm{d}M_s$$

is a Brownian motion (Lévy's martingale characterization for Brownian motion). It is then obvious that (X_t, B_t) satisfies the stochastic integral equation (5.16), so that $(X_t)_{t\geq 0}$ is a weak solution to (5.15).