CHAPTER 1

Martingale Theory

We review basic facts from martingale theory. We start with discrete-time parameter martingales and proceed to explain what modifications are needed in order to extend the results from discrete-time to continuous-time. The Doob-Meyer decomposition theorem for continuous semimartingales is stated but the proof is omitted. At the end of the chapter we discuss the quadratic variation process of a local martingale, a key concept in martingale theory based stochastic analysis.

1. Conditional expectation and conditional probability

In this section, we review basic properties of conditional expectation.

Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space and $\mathscr{G}$ a σ-algebra of measurable events contained in $\mathscr{F}$. Suppose that $X \in L^1(\Omega, \mathscr{F}, \mathbb{P})$, an integrable random variable. There exists a unique random variable Y which have the following two properties:

(1) $Y \in L^1(\Omega, \mathscr{G}, \mathbb{P})$, i.e., Y is measurable with respect to the σ-algebra $\mathscr{G}$ and is integrable;

(2) for any $C \in \mathscr{G}$, we have

$$\mathbb{E}\{X; C\} = \mathbb{E}\{Y; C\}.$$

This random variable Y is called the conditional expectation of X with respect to $\mathscr{G}$ and is denoted by $\mathbb{E}\{X|\mathscr{G}\}$.

The existence and uniqueness of conditional expectation is an easy consequence of the Radon-Nikodym theorem in real analysis. Define two measures on $(\Omega, \mathscr{G})$ by

$$\mu\{C\} = \mathbb{E}\{X; C\}, \quad \nu\{C\} = \mathbb{P}\{C\}, \qquad C \in \mathscr{G}.$$

It is clear that μ is absolutely continuous with respect to ν. The conditional expectation $\mathbb{E}\{X|\mathscr{G}\}$ is precisely the Radon-Nikodym derivative $d\mu/d\nu$.

If Y is another random variable, then we denote $\mathbb{E}\{X|\sigma(Y)\}$ simply by $\mathbb{E}\{X|Y\}$. Here $\sigma(Y)$ is the σ-algebra generated by Y.

The conditional probability of an event A is defined by

$$\mathbb{P}\{A|\mathscr{G}\} = \mathbb{E}\{I_A|\mathscr{G}\}.$$

The following two examples are helpful.

EXAMPLE 1.1. Suppose that the σ-algebra $\mathscr{G}$ is generated by a partition:

$$\Omega = \cup_{i=1}^{\infty} A_i, \qquad A_i \cap A_j = \emptyset \text{ if } i \neq j,$$

and $\mathbb{P}\{A_i\} > 0$. Then $\mathbb{E}\{X|\mathscr{G}\}$ is constant on each A_i and is equal to the average of X on A_i, i.e.,

$$\mathbb{E}\{X|\mathscr{G}\}(\omega) = \frac{1}{\mathbb{P}\{A_i\}} \int_{A_i} X \, d\mathbb{P}, \quad \omega \in A_i.$$

EXAMPLE 1.2. The conditional expectation $\mathbb{E}\{X|Y\}$ is measurable with respect to $\sigma(Y)$. By a well known result, it must be a Borel function $f(Y)$ of Y. We usually write symbolically

$$f(y) = \mathbb{E}\{X|Y = y\}.$$

Suppose that (X, Y) has a joint density function $p(x, y)$ on $\mathbb{R}^2$. Then we can take

$$f(y) = \frac{\int_{\mathbb{R}} x p(x, y) \, dy}{\int_{\mathbb{R}} p(x, y) \, dx}.$$

The following three properties of conditional expectation are often used.

(1) If $X \in \mathscr{G}$, then $\mathbb{E}\{X|\mathscr{G}\} = X$; more generally, if $X \in \mathscr{G}$ then

$$\mathbb{E}\{XY|\mathscr{G}\} = X\mathbb{E}\{Y|\mathscr{G}\}.$$

(2) If $\mathscr{G}_1 \subseteq \mathscr{G}_2$, then

$$\mathbb{E}\{X|\mathscr{G}_1|\mathscr{G}_2\} = \mathbb{E}\{X|\mathscr{G}_2|\mathscr{G}_1\} = \mathbb{E}\{X|\mathscr{G}_1\}.$$

(3) If X is independent of $\mathscr{G}$, then

$$\mathbb{E}\{X|\mathscr{G}\} = \mathbb{E}\{X\}.$$

The monotone convergence theorem, the dominated convergence theorem, and Fatou's lemma are three basic convergence theorems in Lebesgue integration theory. They still hold when the usual expectation is replaced by the conditional expectation with respect to an aribtary σ-algebra.

2. Martingales

A sequence $\mathscr{F}_* = \{\mathscr{F}_n,\ n \in \mathbb{Z}_+\}$ of increasing σ-algebras on Ω is called a filtration (of σ-algebras). The quadruple $(\Omega, \mathscr{F}_*, \mathscr{F}, \mathbb{P})$ with $\mathscr{F}_n \subseteq \mathscr{F}$ is a filtered probability space. We usually assume that

$$\mathscr{F} = \mathscr{F}_\infty \stackrel{def}{=} \bigvee_{n=0}^{\infty} \mathscr{F}_n,$$

the smallest σ-algebra containing all $\mathscr{F}_n$. Intuitively $\mathscr{F}_n$ represents the information of an evolving random system under consideration accumulated up to time n.

A sequence of random variables $X = \{X_n\}$ is said to be adapted to the filtration $\mathscr{F}_*$ if X_n is measurable with respect to $\mathscr{F}_n$ for all n. The filtration $\mathscr{F}_*^X$ generated by the sequence X is

$$\mathscr{F}_n^X = \sigma\{X_i, i \le n\}.$$

It is the smallest filtration to which the sequence X is adapted.

DEFINITION 2.1. *A sequence of integrable random variables*

$$X = \{X_n,\ n \in \mathbb{Z}_+\}$$

on a filtered probability space $(\Omega, \mathscr{F}_*, \mathbb{P})$ *is called a martingale with respect to* $\mathscr{F}_*$ *if* X *is adapted to* $\mathscr{F}_*$ *and*

$$\mathbb{E}\{X_n | \mathscr{F}_{n-1}\} = X_n$$

for all n. *It is called a submartingale or supermartingale if in the above relation* $=$ *is replaced by* $\ge$ *or* $\le$, *respectively.*

If the reference filtration is not explicitly mentioned, it is usually understood what it should be from the context. In many situations, we simply take the reference filtration is the one generated by the sequence itself.

REMARK 2.2. If X is a submartingale with respect to some filtration $\mathscr{F}_*$, then it is also a martingale with respect to its own filtration $\mathscr{F}_*^X$.

The definition of a supermartingale is rather unfortunate, for $\mathbb{E}X_n \le \mathbb{E}X_{n-1}$, that is, the sequence of expected values is decreasing.

Intuitively a martingale represents a fair game. The defining property of a martingale can be written as

$$\mathbb{E}\{X_n - X_{n-1} | \mathscr{F}_{n-1}\} = 0.$$

We can regard the difference $X_n - X_{n-1}$ as a gambler's gain at the nth play of a game. The above equation says that even after applying all the information and knowledge he has accumulated up to time $n-1$, his expected gain is still zero.

EXAMPLE 2.3. If $\{X_n\}$ is a sequence of independent and integrable random variables with mean zero, then the partial sum

$$S_n = X_1 + X_2 + \cdots + X_n$$

is a martingale.

EXAMPLE 2.4. (1) If $\{X_n\}$ is a martingale and $f : \mathbb{R} \mapsto \mathbb{R}$ a convex function such that each $f(X_n)$ is integrable, then $\{f(X_n)\}$ is a submartingale. (2) $\{X_n\}$ is a submartingale and $f : \mathbb{R} \mapsto \mathbb{R}$ a convex and increasing function such that each $f(X_n)$ is integrable, then $\{f(X_n)\}$ is a submartingale.

EXAMPLE 2.5. (martingale transform) Let M be a martingale and $Z = \{Z_n\}$ adapted. Suppose further that each Z_n is uniformly bounded. Define

$$N_n = \sum_{i=1}^{n} Z_{i-1}(M_i - M_{i-1}).$$

Then N is a martingale. Note that in the general summand, the multiplicative factor Z_{i-1} is measurable with respect to the *left* time point of the martingale difference $M_i - M_{i-1}$.

EXAMPLE 2.6. (Reverse martingale) Suppose that $\{X_n\}$ is a sequence of i.i.d. integrable random variables and

$$Z_n = \frac{X_1 + X_2 + \cdots X_n}{n}.$$

Then $\{Z_n\}$ is a reverse martingale, which means that

$$\mathbb{E}\left\{Z_n | \mathscr{G}_{n+1}\right\} = Z_{n+1},$$

where

$$\mathscr{G}_n = \sigma\left\{S_i,\ i \geq n\right\}.$$

If we write $W_n = Z_{-n}$ and $\mathscr{H}_n = \mathscr{G}_{-n}$ for $n = -1, -2, \ldots$. Then we can write

$$\mathbb{E}\left\{W_n | \mathscr{H}_{n-1}\right\} = W_{n-1}.$$

3. Basic properties

Suppose that $X = \{X_n\}$ is a submartingale. We have $\mathbb{E}X_n \leq \mathbb{E}X_{n+1}$. Thus on average a submartingale is increasing. On the other hand, for a martingale we have $\mathbb{E}X_n = \mathbb{E}X_{n+1}$, which shows that it is purely noise. The Doob decomposition theorem claims that a submartingale can be decomposed uniquely into the sum of a martingale and an increasing sequence. The following example shows that the uniqueness question for the decomposition is not an entirely trivial matter.

EXAMPLE 3.1. Consider S_n, the sum of a sequence of independent and square integrable random variables with mean zero. Then $\left\{S_n^2\right\}$ is a submartingale. We have obviously

$$S_n^2 = S_n^2 - \mathbb{E}S_n^2 + \mathbb{E}S_n^2.$$

From $\mathbb{E}S_n^2 = \sum_{i=1}^n \mathbb{E}X_i^2$ it is easy to verify that $M_n = S_n^2 - \mathbb{E}S_n^2$ is a martingale. Therefore the above identity is a decomposition of the submartingale $\left\{S_n^2\right\}$ into the sum of a martingale and an increasing process. On the other hand,

$$S_n^2 = 2\sum_{i=1}^n S_{i-1}X_i + \sum_{i=1}^n X_i^2.$$

The first sum on the right side is a martingale (in the form of a martingale transform). Thus the above gives another such decomposition. In general these two decompositions are different.

The above example shows that in order to have a unique decomposition we need further restrictions.

DEFINITION 3.2. *A sequence $Z = \{Z_n\}$ is said to be predictable with respect to a filtration $\mathscr{F}_*$ if $\mathscr{F}_n$ if $Z_n \in \mathscr{F}_{n-1}$ for all $n \geq 0$.*

THEOREM 3.3. (Doob decomposition) *Let X be a submartingale. Then there is a unique increasing predictable process Z with $Z_0 = 0$ and a martingale M such that*

$$X_n = M_n + Z_n.$$

PROOF. Suppose that we have such a decomposition. Conditioning on $\mathscr{F}_{n-1}$ in

$$Z_n - Z_{n-1} = X_n - X_{n-1} - (M_n - M_{n-1}),$$

we have

$$Z_n - Z_{n-1} = \mathbb{E}\left\{X_n - X_{n-1}|\mathscr{F}_{n-1}\right\}.$$

The right side is nonnegative if X is a submartingale. This shows that a Doob decomposition, if exists, must be unique. It is now clear how to proceed to show the existence. We define

$$Z_n = \sum_{i=1}^{n} \mathbb{E}\left\{X_i - X_{i-1}|\mathscr{F}_{i-1}\right\}.$$

It is clear that X is increasing, predictable, and $Z_0 = 0$. Define $M_n = X_n - Z_n$. We have

$$M_n - M_{n-1} = X_n - X_{n-1} - \mathbb{E}\left\{X_n - X_{n-1}|\mathscr{F}_{n-1}\right\},$$

from which it is easy to see that $\mathbb{E}\left\{M_n - M_{n-1}|\mathscr{F}_{n-1}\right\} = 0$. This shows that M is a martingale. □

4. Optional sampling theorem

The concept of a martingale derives much of its power from the optional sampling theorem we will discuss in this section.

Most interesting events concerning a random sequence occurs not at a fixed constant time, but at a random time. The first time that the sequence reaches above a given level is a typical example.

DEFINITION 4.1. *A function $\tau : \Omega \to \mathbb{Z}_+$ on a filtered measurable space $(\omega, \mathscr{F}_*)$ is called a stopping time if $\{\tau \le n\} \in \mathscr{F}_n$ for all $n \ge 0$. Equivalently $\{\tau = n\} \in \mathscr{F}_n$ for all $n \ge 0$.*

Let X be a random sequence and τ is the first time such that $X_n > 100$. The event $\{\tau \le n\}$ means that the sequence has reached above 100 before or at time n. We can determine if $\{\tau \le n\}$ is true or false by looking at the sequence $X_1, X_2, \ldots, X_n$. We need to look at the sequence beyond time n. Therefore τ is a stopping time. On the other hand, let σ be the last time that $X_n > 100$. Knowing only the first n terms of the sequence will not determine the event $\{\sigma \le n\}$. We need to look beyond time n. Thus σ is not a stopping time.

EXAMPLE 4.2. Many years I was invited to give a talk by a fellow probabilist working in a town I had never been. When asking for the directions to the mathematics department, I was instructed to turn at the last traffic light on a street.

EXAMPLE 4.3. Let σ_n be stopping times. Then $\sigma_1 + \sigma_2$, $\sup_n \sigma_n$, and $\inf_n \sigma_n$ are all stopping times.

We have mentioned that $\mathscr{F}_n$ should be regarded as the information accumulated up to time n. We need the corresponding concept for a stopping time τ. Define

$$\mathscr{F}_\tau = \{C \in \mathscr{F}_\infty : C \cap \{\tau \le n\} \in \mathscr{F}_n \quad \text{for all} \quad n \ge 0\}.$$

It is easy to show that $\mathscr{F}_\tau$ is a σ-algebra. But more importantly, we have the following facts:

(1) $\mathscr{F}_\tau = \mathscr{F}_n$ if $\tau = n$;
(2) if $\sigma \le \tau$, then $\mathscr{F}_\sigma \subseteq \mathscr{F}_\tau$;
(3) τ is $\mathscr{F}_\tau$-measurable.
(3) if $\{X_n\}$ is adapted to $\mathscr{F}_*$, then X_τ is $\mathscr{F}_\tau$-measurable.

Note that the random variable X_τ is defined as $X_\tau(\omega) = X_{\tau(\omega)}(\omega)$.

We say that a stopping time τ is bounded if there is an integer N such that $\mathbb{P}\{\tau \le N\} = 1$. If X is a sequence of integrable random variables and τ a uniformly bounded stopping time, then X_τ is also integrable. Keep this technical point in mind when studying the following Doob's optional sampling theorem.

THEOREM 4.4. *Let X be a submartingale. Let σ and τ be two uniformly bounded stopping times such that $\sigma \le \tau$. Then*

$$\mathbb{E}\{X_\tau | \mathscr{F}_\sigma\} \ge X_\sigma.$$

PROOF. Let $C \in \mathscr{F}_\sigma$. It is enough to show that

$$\mathbb{E}\{X_\tau; C\} \ge \mathbb{E}\{X_\sigma; C\}.$$

This is implied by

$$\mathbb{E}\{X_\tau; C_n\} \ge \mathbb{E}\{X_n; C_n\}, \tag{4.1}$$

where $C_n = C \cap \{\sigma = n\}$.

For $k \ge n$, we have obviously

$$\begin{aligned}&\mathbb{E}\{X_\tau; C_n \cap (\tau = k\})\\ &\qquad = \mathbb{E}\{X_k; C_n \cap (\tau \ge k)\} - \mathbb{E}\{X_k; C_n \cap (\tau \ge k+1)\}.\end{aligned}$$

In the last term, the random variable X_k is integrated on the set $C_n(\tau \ge k+1)$. From $k \ge n$ we have $C_n \in \mathscr{F}_n \subseteq \mathscr{F}_k$, hence

$$C_n \cap (\tau \ge k+1) = C_n \cap (\tau \le k)^c \in \mathscr{F}_k.$$

Using this and the fact that X is a submartingale we have

$$\mathbb{E}\{X_k; C_n \cap (\tau \ge k+1)\} \le \mathbb{E}\{X_{k+1}; C_n \cap (\tau \ge k+1)\}.$$

It follows that

$$\begin{aligned}&\mathbb{E}\{X_\tau; C_n \cap \{\tau = k\}\}\\ &\qquad \ge \mathbb{E}\{X_k; C_n \cap (\tau \ge k)\} - \mathbb{E}\{X_{k+1}; C_n \cap (\tau \ge k+1)\}.\end{aligned}$$

We now sum over $k \geq n$. On the right side we have a telescoping sum and only the first term because τ is bounded. We also have $C_n \cap \{\tau \geq n\} = C_n$ because $\sigma \leq \tau$. It follows that $\mathbb{E}\{X_\tau; C_n\} \geq \mathbb{E}\{X_n; C_n\}$. □

COROLLARY 4.5. *Let X be a submartingale. Let $\sigma \leq \tau$ be two bounded stopping times such that $\sigma \leq \tau$. Then $\mathbb{E}\, X_\sigma \leq \mathbb{E}\, X_\tau$.*

COROLLARY 4.6. (optional sampling theorem) *Let X be a submartingale. Let $\tau_0 \leq \tau_1 \leq \tau_2 \leq \cdots$ be an increasing sequence of bounded stopping times. Then $\{X_{\tau_n}\}$ is a submartingale with respect to the filtration $\{\mathscr{F}_{\tau_n}\}$.*

COROLLARY 4.7. *Let X be a submartingale with respect to a filtration $\mathscr{F}_*$ and τ a stopping time. Then the stopped process $\{X_{n\wedge\tau}\}$ is a submartingale with respect to the original filtration $\mathscr{F}_*$.*

5. Submartingale inequalities

The L^p-norm of a random variable X is

$$\|X\|_p = \{\mathbb{E}|X|^p\}^{1/p}.$$

The set of random variables such that $\|X\|_p < \infty$ is denoted by $L^p(\Omega, \mathscr{F}, \mathbb{P})$ (or an abbreviated variant of it).

Let $X_n^* = \max_{1\leq i\leq n} X_i$. We first show that the tail probability of X_n^* is controled in some sense by the tail probability of the last element X_n.

LEMMA 5.1. (Doob's submartingale inequality) *Let X be a nonnegative submartingale. Then for any $\lambda > 0$ we have*

$$\mathbb{P}\{X_n^* \geq \lambda\} \leq \frac{1}{\lambda}\mathbb{E}\{X_n; X_n^* \geq \lambda\}.$$

PROOF. Let $\tau = \inf\{i : X_i \geq \lambda\}$ with the convention that $\inf \emptyset = n+1$. It is clear that τ is a stopping time; hence $\{\tau \leq n\} \in \mathscr{F}_\tau$ and by the optional sampling theorem we have

$$\mathbb{E}\{X_n; \tau \leq n\} \geq \mathbb{E}\{X_\tau; \tau \leq n\} \geq \lambda \mathbb{P}\{\tau \leq n\}.$$

In the last step we have used the fact that $X_\tau \geq \lambda$ if $\tau \leq n$. From the above inequality and the fact that $\{X_n^* \geq \lambda\} = \{\tau \leq n\}$ we have the desired inequality. □

COROLLARY 5.2. *Let X be a nonnegative submartingale. Then*

$$\mathbb{P}\{X_n^* \geq \lambda\} \leq \frac{\mathbb{E}X_n}{\lambda}.$$

EXAMPLE 5.3. The classical Kolmogorov inequality is a special case of Doob's submartingale inequality. Let $\{X_n, 1 \leq n \leq N\}$ be a sequence of independent, square integrable random variables with mean zero and $S_n = X_1 + \cdots + X_n$. Then $|S_n|^2$ is a nonnegative submartingale and we have

$$\mathbb{P}\left\{\max_{1\leq n\leq N} |S_n| \geq \lambda\right\} \leq \frac{\mathbb{E}|S_n|^2}{\lambda^2}.$$

To convert the above inequality about probability to an inequality about moments, we need the following fact about nonnegative random variables.

PROPOSITION 5.4. *Let X and Y be two nonnegative random variables. Suppose that they satisfy*

$$\mathbb{P}\left\{Y \geq \lambda\right\} \leq \frac{1}{\lambda}\mathbb{E}\left\{X; Y \geq \lambda\right\}$$

for all $\lambda > 0$. Then for any $p > 1$,

$$\|Y\|_p \leq \frac{p}{p-1}\|X\|_p.$$

For the case $p = 1$ we have

$$\|Y\|_1 \leq \frac{e}{e-1}\|1 + X\ln^+ X\|_1.$$

Here $\ln^+ x = \max\left\{\ln x, 0\right\}$ for $x \geq 0$.

PROOF. We need to truncate the random variable Y in order to deal with the possibility that the moment of Y may be infinite. Otherwise the the truncation is unnecessary and the proof may be easier to read. Let $YN = \min\left\{Y, N\right\}$. For $p > 0$, we have by Fubini's theorem,

$$\mathbb{E}Y_N^p = p\mathbb{E}\int_0^{Y_N}\lambda^{p-1}\,d\lambda = p\mathbb{E}\int_0^{\infty}\lambda^{p-1}I_{\{\lambda\leq Y_N\}} = p\int_0^N\lambda^{p-1}\mathbb{P}\left\{Y \geq \lambda\right\}\,d\lambda.$$

Now if $p > 1$, then

$$\begin{aligned}\mathbb{E}Y_N^p &= p\int_0^N\lambda^{p-1}\mathbb{P}\left\{Y \geq \lambda\right\}\,d\lambda\\ &\leq p\int_0^N\lambda^{p-2}\mathbb{E}\left\{X; Y \geq \lambda\right\}\,d\lambda\\ &= \frac{p}{p-1}\mathbb{E}\left\{XY_N^{p-1}\right\}.\end{aligned}$$

Using Hölder's inequality we have

$$\mathbb{E}Y_N^p \leq \frac{p}{p-1}\|X\|_p\,\|Y_N\|_p^{p-1}.$$

Now $\mathbb{E}Y_N^p$ is finite, hence we have above inequality

$$\|Y_N\|_p \leq \frac{p}{p1}\|X\|_p.$$

Letting $N \to \infty$ and using the monotone convergence theorem, we obtain the desired inequality follows immediately. For the proof of the case $p = 1$, see EXERCISE **??**. □

The following moment inequalities for a nonnegative submartingale are often useful.

THEOREM 5.5. *Let X be a nonnegative submartingale. Then*

$$\|X_n^*\|_p \leq \frac{p}{p-1} \|X_n\|_p, \qquad p > 1;$$
$$\|X_n^*\|_1 \leq \frac{e}{e-1} \|1 + X_n \ln^+ X_n\|_1.$$

Here $\ln^+ x = \max\{\ln x, 0\}$.

PROOF. These inequalities follow immediately from PROPOSITION 5.4 and LEMMA 5.1. □

6. Convergence theorems

Let $X = \{X_n\}$ be a submartingale. Fix two real numbers $a < b$ and let $U_N^X[a,b]$ be the number of upcrossings of the sequence $X_1, X_2, \cdots, X_N$ from below a to above b. The precise definition of $U_N^X[a,b]$ will be clear from the proof of the next lemma.

LEMMA 6.1. (Upcrossing inequality) *We have*

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

PROOF. Let $Y_n = (X_n - a)^+$. Then Y is a positive submartingale. An upcrossing of X from below a to above b is an upcrossing of Y from 0 to $b - a$ and vice versa, so we may consider Y instead of X. Starting from $\tau_0 = 1$ we define the following sequence of stopping times

$$\sigma_i = \inf\{n \geq \tau_{i-1} : Y_n = 0\},$$
$$\tau_i = \inf\{n \geq \sigma_i : Y_n \geq b - a\}$$

with the convention that $\inf \emptyset = N$. If $X_{\sigma_i} \leq a$ and $X_{\tau_i} \geq b$, that is, if $[\sigma_i, \tau_i]$ is an interval of a completed upcrossing from below a to above b, then we have $Y_{\tau_i} - Y_{\sigma_i} \geq b - a$. Even if $[\sigma_i, \tau_i]$ is not an interval of a completed upcrossing, we always have $Y_{\tau_i} - Y_{\sigma_i} \geq 0$. Therefore,

$$\sum_{i=1}^{N} (Y_{\tau_i} - Y_{\sigma_i}) \geq (b - a) U_n^X[a,b]. \tag{6.1}$$

On the other hand, since X is a submartingale and $\tau_{i-1} \leq \sigma_i$, we have

$$\mathbb{E}\left[Y_{\sigma_i} - Y_{\tau_{i-1}}\right] \geq 0.$$

Now we have

$$Y_n - Y_0 = \sum_{i=1}^{N} [Y_{\tau_i} - Y_{\sigma_i}] + \sum_{i=1}^{N} \left[Y_{\sigma_i} - Y_{\tau_{i-1}}\right].$$

Taking the expected value we have

$$\mathbb{E}\,[Y_N - Y_0] \geq (b - a)\mathbb{E} U_N^X[a,b].$$

□

The next result is the basic convergence theorem in martingale theory.

THEOREM 6.2. *Let $\{X_n\}$ be a submartingale such that $\sup_{n\geq 1} \mathbb{E}X_n^+ \leq C$ for some constant C. Then the sequence converges with probability one to an integrable random variable X_∞.*

PROOF. We use the following easy fact from analysis: a sequence $x = \{x_n\}$ of real numbers converges to a finite number or $\pm\infty$ if and only if the number of upcrossings from a to b is finite for every pair of rational numbers $a < b$. Thus it is enough to show that

$$\mathbb{P}\left\{U_\infty^X[a,b] < \infty \text{ for all rational } a < b\right\} = 1.$$

Since the set of intervals with rational endpoints is countable, we need only to show that

$$\mathbb{P}\left\{U_\infty^X[a,b] < \infty\right\} = 1$$

for fixed $a < b$. We will show the stronger statement that $\mathbb{E}U_\infty^X[a,b] < \infty$.

From the upcrossing inequality we have

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

for all N. From

$$(X_n - a)^+ \leq |X_n| + |a|$$

we have

$$\mathbb{E}U_N^X[a,b] \leq \frac{C + |a|}{b-a}.$$

We have $U_N^X[a,b] \uparrow U_\infty^X[a,b]$ as $N \uparrow \infty$, hence by the monotone convergence theorem, $\mathbb{E}U_N^X[a,b] \uparrow \mathbb{E}U_\infty^X[a,b]$. It follows that

$$\mathbb{E}\, U_\infty^X[a,b] \leq \frac{C + |a|}{b-a} < \infty.$$

It follows that the limit $X_\infty = \lim_{n\to} X_n$ exists with probability one.

From $|X_n| = 2X_n^+ - X_n$,

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

By Fatou's lemma we have $\mathbb{E}|X_\infty| \leq 2C - \mathbb{E}X_1$, which shows that X_∞ is integrable. □

COROLLARY 6.3. *Every nonnegative supermartingale converges to an integrable random variable with probability one.*

PROOF. If X is a nonnegative supermartingale, then $-X$ is a nonpositive submartingale. Now apply the basic convergence THEOREM 6.2. □

7. Uniformly integrable martingales

The basic convergence THEOREM 6.2 does not claim that $\mathbb{E}X_n \to \mathbb{E}X_\infty$ or $\mathbb{E}|X_n - X_\infty| \to 0$. This does not hold in general. In many applications it is desirable to show that this convergence is in the sense of L^1. To claim the convergence in this sense, we need more conditions. One such condition that is used very often is that of uniform integrability. We recall this concept and the relevant convergence theorem.

DEFINITION 7.1. *A sequence $\{X_n\}$ is called uniformly integrable if*

$$\lim_{C\to\infty} \sup \sup \mathbb{E}\left\{|X_n|; |X_n| \geq C\right\} = 0.$$

Equivalently, for any $\epsilon > 0$, there is a constant C such that

$$\mathbb{E}\left\{|X_n|; |X_n| \geq C\right\} \leq \epsilon$$

for all n; i.e., the sequence $\{X_n\}$ has uniformly small L^1-tails.

The following equivalent condition justifies the terminology uniform integrability from another point of view.

PROPOSITION 7.2. *A sequence $\{X_n\}$ is uniformly integrable if and only if for any positive ϵ, there exists a positive δ such that*

$$\mathbb{E}\left\{|X_n|; C\right\} \leq \epsilon$$

for all n and all sets C such that $\mathbb{P}\left\{C\right\} \leq \epsilon$.

The concept of uniform integrability is useful mainly because of the following so-called Vitali convergence theorem , which supplements the three more well-known convergence theorems (Fatou, monotone, and dominated).

THEOREM 7.3. *If $X_n \to X$ in probability and $\{X_n\}$ is uniformly integrable, then X is integrable and $\mathbb{E}|X_n - X| \to 0$.*

PROOF. If is clear that the uniform integrability implies that $\mathbb{E}|X_n|$ is uniformly bounded, say $\mathbb{E}|X_n| \leq C$. There is a subsequence of $\{X_n\}$ converging to X with probability 1, hence by Fatou's lemma we have $\mathbb{E}|X| \leq C$. This shows that X is integrable. Let $Y_n = |X_n - X|$. Then $\{Y_n\}$ is also uniformly integrable and $Y_n \to 0$ in probability. We need to show that $\mathbb{E}Y_n \to 0$. For an $\epsilon > 0$ and let $C_n = \{Y_n \geq \epsilon\}$. Then $\mathbb{P}\left\{C_n\right\} \to 0$. We have

$$\mathbb{E}Y_n = \mathbb{E}\left\{Y_n; Y_n < \epsilon\right\} + \mathbb{E}\left\{Y_n; Y_n \geq \epsilon\right\}.$$

The first term is bounded by ϵ, and the second term goes to zero as $n \to \infty$ by PROPOISITION 7.2. This shows that $\mathbb{E}Y_n \to 0$. □

The following criterion for uniform integrability is often useful.

PROPOSITION 7.4. *If there is $p > 1$ and C such that $\mathbb{E}|X_n|^p < \infty$ for all n, then $\{X_n\}$ is uniformly integrable.*

PROOF. EXERCISE **??** □

EXAMPLE 7.5. Let X be an integrable random variable and $\{\mathscr{F}_n\}$ be a family of σ-algebras. Then the family of random variables $X_n = \mathbb{E}\{X_n|\mathscr{F}_n\}$ is uniformly integrable. This can be seen as follows. We may assume that X is nonnegative. Let $C_n = \{X_n \geq C\}$. Then $C_n \in \mathscr{F}_n$. We have

$$\mathbb{E}\{X_n; X_n \geq C\} = \mathbb{E}\{X; C_n\}.$$

Considering the two cases $X < K$ and $X \geq K$, we have

$$\mathbb{E}\{X_n; X_n \geq C\} \leq K\mathbb{P}\{C_n\} + \mathbb{E}\{X; X \geq K\}.$$

The second term can be made arbitrarily small for sufficiently large K. For a fixed K, the first term can be made arbitrarily small if C is sufficiently large because

$$\mathbb{P}\{C_n\} \leq \frac{\mathbb{E}X_n}{C} = \frac{\mathbb{E}X}{C}.$$

THEOREM 7.6. *If $\{X_n\}$ is a uniformly integrable submartingale, then it converges with probability 1 to an integrable random variable X_∞ and $\mathbb{E}|X_n - X_\infty| \to 0$.*

PROOF. From uniform integrability we have $\mathbb{E}|X_n| \leq C$ for some constant C; hence the submartingale converges. The L^1-convergence follows from THEOREM 7.3. □

The following theorem gives the general form of a uniformly integrable martingale.

THEOREM 7.7. *Every uniformly integrable martingale $\{X_n\}$ has the form $X_n = \mathbb{E}\{X|\mathscr{F}_n\}$ for some integrable random variable X.*

PROOF. From EXAMPLE 7.5) we see that a martingale of the indicated form is uniformly integrable. On the other hand, if the martingale $\{X_n\}$ is uniformly integrable, then $X_n \to X$ with probability 1 for an integrable random variable X. Furthermore, $\mathbb{E}|X_n - X_\infty| \to 0$. Taking the limit as $m \to \infty$ in $X_n = \mathbb{E}\{X_m|\mathscr{F}_n\}$ we have $X_n = \mathbb{E}\{X|\mathscr{F}_n\}$. □

The following theorem gives the "last term" of the uniformly integrable martingale $X_n = \mathbb{E}\{X|\mathscr{F}_n\}$.

THEOREM 7.8. *Suppose that $\mathscr{F}_*$ is a filtration of σ-algebras and that X is an integrable random variable. Then with probability 1,*

$$\mathbb{E}\{X|\mathscr{F}_n\} \to \mathbb{E}\{X|\mathscr{F}_\infty\}$$

where $\mathscr{F}_\infty = \sigma\{\mathscr{F}_n, n \geq 0\}$, the smallest σ-algebra containing all $\mathscr{F}_n, n \geq 0$.

PROOF. Let $X_n = \mathbb{E}\{X|\mathscr{F}_n\}$. Then $\{X_n, n \geq 0\}$ is a uniformly integrable martingale (see EXERCISE **??** below). Hence the limit

$$X_\infty = \lim_{n\to\infty} X_n$$

exists and the convergence is in $L^1(\Omega, \mathscr{F}, \mathbb{P})$. We need to show that $X_\infty = \{X|\mathscr{F}_\infty\}$.

It is clear that $X_\infty \in \mathscr{F}_\infty$. Letting $m \to \infty$ in $X_n = \mathbb{E}\{X_m|\mathscr{F}_n\}$, we have

$$\mathbb{E}\{X|\mathscr{F}_n\} = X_n = \mathbb{E}\{X_\infty|\mathscr{F}_n\}.$$

Now consider the collection $\mathscr{G}$ of sets $G \in \mathscr{F}_\infty$ such that

$$\mathbb{E}\{X_\infty; G\} = \mathbb{E}\{X; G\}.$$

If we can show that $\mathscr{G} = \mathscr{F}_\infty$, then by the definition of conditional expectations we have immediately $X_\infty = \mathbb{E}\{X|\mathscr{F}_\infty\}$.

The following two facts are clear:

1. $\mathscr{G}$ is a monotone class;
2. It contains every $\mathscr{F}_n$.

Therefore $\mathscr{G}$ contains the field $\cup_{n\geq 1}\mathscr{F}_n$. By the monotone class theorem (see EXERCISE **??**, $\mathscr{G}$ contains the smallest σ-algebra generated by this field. This shows that $\mathscr{F}_\infty = \mathscr{F}$, and the proof is completed. □

8. Continuous time parameter martingales

Brownian motion $B = \{B_t\}$ is the most important continuous time parameter martingale. From this we will derive other examples such as $B_t^2 - t$ and $\exp[B_t - t/2]$. We will introduce Brownian motion in the next chapter. We assume that the reader has some familiarity with Brownian motion so that the topics of the next three sections are not as dry as it would have been.

The definitions of martingales, submartingales, and supermartingales extend in an obvious way to the case of continuous-time parameters. For example, a real-valued stochastic process $X = \{X_t\}$ is a submartingale with respect to a filtration $\mathscr{F}_* = \{\mathscr{F}_t\}$ if $X_t \in L^1(\Omega, \mathscr{F}_*, \mathbb{P})$ and

$$\mathbb{E}\{X_t|\mathscr{F}_s\} \geq X_s, \qquad s \leq t.$$

It is clear that for any increasing sequence $\{t_n\}$ of times the sequence $\{X_{t_n}\}$ is a submartingale with respect to $\{\mathscr{F}_{t_n}\}$. For this reason, the theory of continuous-parameter submartingales is by and large parallel to that of discrete-parameter martingales. Most of the results (except Doob's decomposition theorem) in the previous sections on discrete parameter martingales have their counterparts in continuous parameter martingales and the proofs there work in the present situation after obvious necessary modifications. However, there are two issues we need to deal with. The first one is a somewhat technical measurability problem. In order to use discrete approximations, we need to make sure that every quantity we deal with is measurable and can be approximated in some sense by the corresponding discrete counterpart. This is not true in general. For example, an expression such as

$$X_t^* = \sup_{0\leq s\leq t} X_s \tag{8.1}$$

is in general not measurable, hence it cannot be approximated this way. The second issue is to find a continuous-time version of the Doob decomposition, which plays a central role in stochastic analysis. In Doob's theorem, we need to use the concept of predictability to ensure the uniqueness. This concept does not seem to have a straightforward generalization in the continuous-time situation. We will resolve these issues by restricting our attention to a class of submartingales satisfying some technical conditions. This class turns out to be wide enough for our applications.

The first restriction is on the filtration. We will show that this class is sufficiently rich for our applications. Suppose that $(\Omega, \mathscr{F}, \mathbb{P})$ is a probability space. A set is a null set if it is contained in a set of measure zero. Recall that a σ-algebra $\mathscr{G}$ is said to be complete (relative to $\mathscr{F}$) with respect to the probability measure $\mathbb{P}$ if it contains all null sets.

DEFINITION 8.1. *A filtration of σ-algebras $\mathscr{F}_* = \{\mathscr{F}_t,\ t \in \mathbb{R}_+\}$ on a probability space $(\Omega, \mathscr{F}_\infty, \mathbb{P})$ is said to satisfy the usual conditions if*

(1) $\mathscr{F}_0$, hence every $\mathscr{F}_t$, is complete relative to $\mathscr{F}_\infty$ with respect to $\mathbb{P}$;

(2) It is right-continuous, i.e., $\mathscr{F}_t = \mathscr{F}_{t+}$ for all $t \geq 0$, where

$$\mathscr{F}_{t+} \stackrel{\text{def}}{=} \bigcup_{s>t} \mathscr{F}_s.$$

The right side is usually denoted by $\mathscr{F}_{t+}$.

CONDITION (1) can usually be satisfied by completing every $\mathscr{F}_t$ by adding subsets of sets of measure zero. For our purpose, the most important example is the filtration generated by Brownian motion. We will show that after completion CONDITION (2) is satisfied in this case. In general these conditions have to be stated as technical assumptions. *Unless otherwise stated, from now on we will assume that all filtrations satisfy the usual condition.*

Our second restriction is on sample paths of stochastic processes.

DEFINITION 8.2. *A function $\phi : \mathbb{R}_+ \to \mathbb{R}^1$ is called a* Skorokhod *function (path) if at every point $t \in \mathbb{R}_+$ it is right continuous $\lim_{s\downarrow} \phi_s = \phi_t$ and has a finite left limit $\phi_{t-} = \lim_{s<t, s\uparrow t} \phi_s$. The space of Skorokhod functions is denote by $\mathscr{D}(\mathbb{R}_+)$ or simply $\mathscr{D}$. A stochastic process whose sample paths are Skorokhod with probability 1 is called a regular process.*

REMARK 8.3. $\mathscr{D}(\mathbb{R}_+)$ is a metric space under the so-called Skorohod metric ; see Billingsley [**2**].

The next theorem shows that under certain not very restrictive technical conditions, every submartingale has a regular version with Skorokhod sample paths. First we clarify the meaning of a version.

DEFINITION 8.4. *We say that $\{Y_t,\ t \geq 0\}$ is a version of $\{X_t,\ t \geq 0\}$ if for all $t \geq 0$,*

$$\mathbb{P}\{X_t = Y_t\} = 1$$

It is clear from definition that if two processes are versions of each other, then they have the same family of finite-dimensional marginal distributions. For this reason, they are identical for all practical purposes. In stochastic analysis it is rarely necessary to distinguish a stochastic process from a version of it.

The following result shows that having Skorokhod paths is not a very restrictive condition for submartingales.

THEOREM 8.5. (Path regularity of submartingales) *Let that* $\{X_t,\ t \geq 0\}$ *is a submartingale with respect to a filtration* $\mathscr{F}_*$ *which satisfies the usual conditions. If the function* $t \mapsto \mathbb{E}X_t$ *is right-continuous. Then* $\{X_t,\ t \geq 0\}$ *has a regular version, i.e., there is a version of* X *almost all of whose paths are Skorokhod.*

PROOF. See Meyer [**9**] , page 95-96. □

Once the process X has Skorokhod paths, a quantity such as X_t^* defined in (8.1) becomes measurable because the maximum can be taken over a dense countable set (e.g., the set of rational or dyadic numbers). We give a proof of Doob's submartingale inequality as an example of passing from discrete to continuous-time sumartingales under the assumption of sample path regularity.

THEOREM 8.6. (Submartingale inequality) *Suppose that* X *is a nonnegative submartingale with regular sample paths and let* $X_T^* = \max_{0 \leq t \leq T} X_t$*. Then for any* $\lambda > 0$,

$$\lambda \mathbb{P}\left\{X_T^* \geq \lambda\right\} \leq \mathbb{E}\left\{X_T; X_T^* \geq \lambda\right\}.$$

PROOF. Without loss of generality we assume $T = 1$. Let

$$Z_N = \max_{0 \leq i \leq 2^N} X_{i/2^N}$$

for simplicity. Since the paths are regular we have $Z_N \uparrow X_1^*$. For any positive ϵ, we have for all sufficiently large N,

$$\lambda \mathbb{P}\left\{X_1^* \geq \lambda\right\} \leq \lambda \mathbb{P}\left\{Z_N \geq \lambda - \epsilon\right\} \leq \frac{\lambda}{\lambda - \epsilon} \mathbb{E}\left\{X_1; X_1 \geq \lambda - \epsilon\right\}.$$

In the last step we have used the submartingale inequality for discrete submartingales and the fact that $X_1^* \geq Z_N$. Letting $\epsilon \downarrow 0$, we obtain the desired inequality. □

THEOREM 8.7. (Submartingale convergence theorem) *Let* $X = \{X_t\}$ *be a regular submartingale. If* $\sup_{t \geq 0} \mathbb{E}X_t^+ < \infty$*, then the limit* $X_\infty = \lim_{t \to \infty} X_t$ *exists and is integrable. If* $\{X_t\}$ *is uniformly integrable, then* $\mathbb{E}|X_t - X_\infty| \to 0$.

PROOF. EXERCISE **??**. □

THEOREM 8.8. *Under the same conditions as in the preceding theorem, if in addition* $\{X_t, t \geq 0\}$ *is uniformly integrable, then the limit* X_∞ *exists a.s. Furthermore,* $\mathbb{E}|X_t - X_\infty| \to 0$ *and* $X_t = \mathbb{E}\left\{X_\infty | \mathscr{F}_t\right\}$.

PROOF. Same as in the discrete-parameter case, *mutatis mutandis*. □

9. Doob–Meyer decomposition theorem

In this section we discuss the Doob-Meyer decomposition theorem, the analogue of Doob's decomposition for continuous-time submartingales. It states, roughly speaking, that a submartingale X can be uniquely written a sum of a martingale M and an increasing process Z, i.e.,

$$X_t = M_t + Z_t.$$

This result plays a central role in stochastic analysis. There are several proofs for this theorem and they all start naturally with approximating a continuous-time submartingale by discrete-time submartingales. There are two difficulties we need to overcome. First, we need to show that these discrete-time approximations converge to the given submartingale in an appropriate sense. Second, in view of the fact in the discrete-time case, the uniqueness holds only after we assume that the increasing part of the decomposition is predictable, we need to formulate an appropriate analogue of predictability in the continuous-time case. Neither of these two problems are easy to overcome. A complete proof of this result can be found in Meyer **[9]**. The proof can be somewhat simplified under the hypothesis of continuous sample paths and the problem of predictability disappears under the usual conditions for the filtration; see Bass **[1]**. Fortunately, the Doob-Meyer theorem is one of those results for which we gain very little in going through its proof, and understanding its meaning is quite adequate for our later study. We will restrict ourselves to a form of this decomposition theorem that is sufficient for our purpose.

We always assume that the filtration $\mathscr{F}_* = \{\mathscr{F}_t,\ t \in \mathbb{R}_+\}$ satisfies the usual conditions. The definition of a martingale requires that each X_t is integrable. A local martingale is a more flexible concept.

DEFINITION 9.1. *A continuous, $\mathscr{F}_*$-adapted process $\{X_t\}$ is called a local martingale if there exists an increasng sequence of stopping times $\tau_n \uparrow \infty$ with probability one such that for each n the stopped process $X^{\tau_n} = \{X_{t\wedge\tau_n},\ t \in \mathbb{R}_+\}$ is an $\mathscr{F}_*$-martingale.*

Similarly we can define a local martingale or a local supermartingale.

EXAMPLE 9.2. Let $\{B_t\}$ be a Brownian motion in $\mathbb{R}^3$ and $a \neq 0$. We will show later that

$$X_t = \frac{1}{|B_t + a|}$$

is a local martingale.

REMARK 9.3. A word of warning is in order. For a local martingale $X = \{X_t\}$, the condition $\mathbb{E}|X_t| < \infty$ for all t does not guarantee that it is a martingale.

The following fact often useful.

PROPOSITION 9.4. *A nonnegative local martingale* $\{M_t\}$ *is a supermartingale. In particular* $\mathbb{E}M_t \le \mathbb{E}M_0$.

PROOF. Let $\{\tau_n\}$ be the nondecreasing stopping times going to infinity such that the stopped processes M^{τ_n} are martingale. Then in particular $M_0 = M_0^{\tau_n}$ is integrable by definition. We have

$$\mathbb{E}\{M_{t\wedge\tau_n}|\mathscr{F}_s\} \le M_{s\wedge\tau_n}.$$

We let $n \to \infty$. On the left side, we use Fatou's lemma. The limit on the right side is M_s because $\tau_n \to \infty$. Hence

$$\mathbb{E}\{M_t|\mathscr{F}_s\} \le M_s.$$

It follows that every M_t is integrable and $\{M_t\}$ is a supermartingale. □

If with probability 1 the sample path $t \mapsto Z_t$ is an increasing function, we call Z an increasing process. If $A = Z^1 - Z^2$ for two increasing processes, then A is called a process of bounded variation.

REMARK 9.5. We will often drop the cumbersome phrase "with probability 1" when it is obviously understood from the context.

We are ready to state the important decomposition theorem.

THEOREM 9.6. (Doob-Meyer decomposition) *Suppose that the filtration* $\mathscr{F}_*$ *satisfies the usual conditions. Let* $X = \{X_t,\ t \in \mathbb{R}_+\}$ *be an* $\mathscr{F}_*$*-adapted local submartingale with continuous sample paths. There is a unique continuous* $\mathscr{F}_*$*-adapted local martingale* $\{M_t,\ t \in \mathbb{R}_+\}$ *and a continuous* $\mathscr{F}_*$*-adapted increasing process* $\{Z_t,\ t \in \mathbb{R}_+\}$ *with* $Z_0 = 0$ *such that*

$$X_t = M_t + Z_t, \quad t \in \mathbb{R}_+.$$

PROOF. See Karatzas and Shreve [**7**], page 21-30. Bass [**1**] contains a simple proof for continuous submartingales. □

In the theorem we stated that the martingale part M of a local submartingale is a local martingale. In general we cannot claim that it is a martingale even if X is a submartingale. In many applications, it is important to know that the martingale part of a submartingale is in fact a martingale so that we can claim the equality

$$\mathbb{E}X_t = \mathbb{E}M_t + \mathbb{E}Z_t.$$

We now discuss a condition for a local martingale to be a martingale, which can be useful in certain situations.

DEFINITION 9.7. *Let* $\mathscr{S}_T$ *be the set of stopping times bounded by* T*. A continuous process* $X = \{X_t\}$ *is said of* (DL)*-class if* $\{X_\sigma,\ \sigma \in \mathscr{F}_T\}$ *is uniformly integrable for all* $T \ge 0$.

This definition is useful for the following reason.

PROPOSITION 9.8. *A local martingale is a martingale if and only if it is of* (DL)*-class.*

PROOF. If M is a martingale, then

$$M_\sigma = \mathbb{E}\{M_T|\mathscr{F}_\sigma\} \quad \text{for all} \quad \sigma \in \mathscr{S}_T.$$

Hence $\{M_\sigma, \sigma \in \mathscr{S}_T\}$ is uniformly integrable (see EXAMPLE 7.5). Thus M is of (DL)-class.

Conversely, suppose that M is a (DL)-class local martingale and let $\{\tau_n\}$ be a sequence of stopping times going to infinity such that the stopped processes M^{τ_n} are martingales. We have

$$\mathbb{E}\{M_{t\wedge\tau_n}|\mathscr{F}_s\} = M_{s\wedge\tau_n}, \quad s \le t.$$

As $n \to \infty$, the limit on the right side is M_s. On the left side $\{M_{t\wedge\tau_n}\}$ is uniformly integrable because $t \wedge \tau_n \in \mathscr{S}_t$. Hence we can pass to the limit on the left side and obtain $\mathbb{E}\{M_t|\mathscr{F}_s\} = M_s$. This shows that M is a martingale. □

THEOREM 9.9. *Let X be a local submartingale X and $X = M + Z$ be its Doob-Meyer decomposition. Then the following two conditions are equivalent:*

(1) *M is a martingale and Z is integrable;*
(2) *X is of (DL)-class.*

PROOF. Suppose that Z is increasing and $Z_0 = 0$. If it is integrable, i.e., $\mathbb{E}Z_t < \infty$ for all $t \ge 0$, then it is clearly of (DL)-class. On the other hand, every martingale M is of (DL)-class because the family

$$M_\sigma = \mathbb{E}\{M_T|\mathscr{F}_\sigma\}, \ \sigma \in \mathscr{S}_T$$

is uniformly integrable. Thus if (1) holds, then both M and Z are of (DL)-class, hence the sum $X = M + Z$ is also of (DL)-class. This proves that (1) implies (2).

Suppose now that the submartingale X is of (DL)-class and $X = M + Z$ is its Doob-Meyer decomposition. Let $\{\tau_n\}$ be a sequence of stopping times tending to infinity such that the stopped processes M^{τ_n} are martingales. We have

$$X_{t\wedge\tau_n} = M_{t\wedge\tau_n} + Z_{t\wedge\tau_n}.$$

Taking the expectation we have

$$\mathbb{E}Z_{t\wedge\tau_n} = \mathbb{E}X_{t\wedge\tau_n} - \mathbb{E}X_0.$$

We have $Z_{t\wedge\tau_n} \uparrow Z_t$, hence $\mathbb{E}Z_{t\wedge\tau_n} \uparrow \mathbb{E}Z_t$ by the monotone convergence theorem. On the other hand, the assumption that X is of (DL)-class implies that $\mathbb{E}X_{t\wedge\tau_n} \to \mathbb{E}X_t$. Hence,

$$\mathbb{E}Z_t = \mathbb{E}X_t - \mathbb{E}X_0 < \infty.$$

Thus $\mathbb{E}Z_t$ is integrable. Since Z is also nonnegative and increasing, we see that Z is of (DL)-class. Now that both X and Z are of (DL)-class, the difference $M = X - Z$ is also of (DL)-class (see EXERCISE **??**) and is therefore a martingale. □

In these notes the broadest class of stochastic processes we will deal with is that of semimartingales.

DEFINITION 9.10. *A stochastic process is called a semimartingale (with respect to a given filtration of σ-algebras if it is the sum of a continuous local martingale and a continuous process of bounded variation.*

In real analysis it is well known that a function of bounded variation is the sum of two increasing functions. Using this result, we can show easily that a semimartingale X has the form $X = M + Z^1 - Z^2$, where M is a local martingale and Z^i are two incrasing processes.

The significance of this class of stochastic processes lies in the fact that it is closed under most common operations we usually perform on them. In particular, Itô's formula shows that a smooth function of a semimartingale is still a semimartingale.

For a semimartingale, the Doob-Meyer decomposition is still unique.

PROPOSITION 9.11. *A semimartingale X can be uniquely decomposed into the sum $X = M + Z$ of a continuous local martingale M and a continuous process of bounded variation Z with $Z_0 = 0$.*

PROOF. It is enough to show that if $M + Z = 0$ then $M = 0$ and $Z = 0$. Let $Z_t = Z_t^1 - Z_t^2$, where Z^i are increasing processes with $Z_0^1 = Z_0^2 = 0$. We have $M_t + Z_t^1 = Z_t^2$. By the uniqueness of Doob-Meyer decompositions we have $M_t = 0$ and $Z_t^1 = Z_t^2$, hence $Z_t = 0$. $\square$

10. Quadratic variation of a continuous martingale

The quadratic variation is an important characterization of a continuous martingale. In fact Lévy's criterion shows that Brownian motion B is a martingale completely characterized by its quadratic variation $\langle B \rangle_t = t$.

Let $M = \{M_t\}$ be a continuous local martingale. Then M^2 is a continuous local submartingale. By the Doob-Meyer decomposition theorem, there is a continuous increasing process, which we will denote by $\langle M, M \rangle = \{\langle M, M \rangle_t\}$ or simply $\langle M \rangle = \{\langle M \rangle_t\}$, with $\langle M, M \rangle_0 = 0$, such that $M^2 - \langle M, M \rangle$ is a continuous local martingale. This increasing process $\langle M, M \rangle$ is uniquely determined by M and is called the quadratic variation process of the local martingale M.

EXAMPLE 10.1. The most important example is $\langle B \rangle_t = t$. This means that $B_t^2 - t$ is a (local) martingale. This can be verified directly from the definition of Brownian motion. We will show that the quadratic variation process of the martingale $B_t^2 - t$ is

$$4 \int_0^t B_s^2 \, ds.$$

Let's look at the quadratic variation process from another point of view. This also gives a good motivation for the term itself. Consider a partition of the time interval $[0t]$:

$$\Delta : \quad 0 = t_0 < t_1 < \cdots < t_n = t.$$

The mesh of the partition is the lengths of its largest interval:

$$|\Delta| = \sup_{l\geq 1} |t_l - t_{l-1}|.$$

Now let M be a continuous local martingale. We have the identity

$$M_t^2 = M_0^2 + 2\sum_{i=1}^{n} M_{t_{i-1}}(M_{t_i} - M_{t_{i-1}}) + \sum_{i=1}^{n}(M_{t_i} - M_{t_{i-1}})^2.$$

The second term on the right side looks like a Riemann sum and the third term is the quadratic variation of the process M along the partition. In CHAPTER 3 we will show by stochastic integration theory that the Riemann sum converges (in probability) to the stochastic integral

$$\int_0^t M_s\, dM_s$$

as $\Delta| \to 0$. The stochastic integral above is a local martingale. It follows that the quadratic variation along Δ also converges, and by the uniqueness of Doob-Meyer decompositions we have in probability,

$$\lim_{|\Delta|\to 0} \sum_{i=1}^{n}(M_{t_i} - M_{t_{i-1}})^2 = \langle M, M\rangle_t.$$

This relation justifies the name "quadratic variation process" for $\langle M, M\rangle$. We also have

$$M_t^2 = M_0^2 + 2\int_0^t M_s\, dM_s + \langle M, M\rangle_t.$$

This gives an explicit Doob-Meyer decomposition of the local submartingale M. It is also a simplest case of Itô's formula.

For a semimartingale $X = M + Z$ in its Doob-Meyer decomposition, we define $\langle Z, Z\rangle$ to be the quadratic variation of its martingale part, i.e.,

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

Note that we still have

$$\lim_{|\Delta|\to 0} \sum_{i=1}^{n}(X_{t_i} - X_{t_{i-1}})^2 = \langle X, X\rangle_t. \tag{10.1}$$

To see this, letting $\Delta X_i = X_{t_i} - X_{t_{i-1}}$, we have

$$(\Delta X_i)^2 = (\Delta M_i)^2 + (2\Delta M_i + \Delta Z_i)\Delta Z_i.$$

By sample path continuity, we have

$$C(\Delta) \stackrel{\text{def}}{=} \max_{1\leq i\leq n} |2\Delta M_i + \Delta Z_i| \to 0$$

as $|\Delta| \to 0$. From this we have

$$\left|\sum_{i=1}^{n}(2\Delta M_i + \Delta Z_i)Z_i\right| \leq C(\Delta)\sum_{i=1}^{n}|\Delta Z_i| \leq C(\Delta)|Z|_t \to 0.$$

Here $|Z|_t$ denotes the total variation of $\{Z_s,\ 0 \le s \le t\}$. It follows that

$$\lim_{\Delta| \to 0} \sum_{i=1}^{n} (\Delta X_i)^2 = \lim_{\Delta| \to 0} \sum_{i=1}^{n} (\Delta M_i)^2 = \langle M, M \rangle_t.$$

This shows that (10.1) holds.

Finally we show that the product of two local martingales M and N is a semimartingale. The covariation process $\langle M, N \rangle$ is defined by polarization:

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

It is obviously a process of bounded variation.

PROPOSITION 10.2. *Let M and N be two local martingales. The process $MN - \langle M, N \rangle_t$ is a local martingale.*

PROOF. By the definition of covariation it is easy to verify that $MN - \langle M, N \rangle$ is equal to

$$\frac{(M+N)^2 - \langle M+N, M+N \rangle}{4} - \frac{(M-N)^2 - \langle M-N, M-N \rangle}{4}.$$

Both processes on the right side are local martingales, hence $MN - \langle M, N \rangle$ is also a local martingale. ☐

We will show later that the Doob-Meyer decomposition of MN is

$$M_t N_t = M_0 N_0 + \int_0^t M_s \, dN_s + \int_0^t N_s \, dM_s + \langle M, N \rangle_t.$$

11. First assignment

EXERCISE 1.1. Consider the Hilbert space $L^2(\Omega, \mathscr{F}, \mathbb{P})$ of square integrable random variables. The space $L^2(\Omega, \mathscr{G}, \mathbb{P})$ of $\mathscr{G}$-measurable square integrable random variables is a closed subspace. If $X \in L^2(\Omega, \mathscr{F}, \mathbb{P})$. Then $\mathbb{E}\{X|\mathscr{G}\}$ is the projection of X to $L^2(\Omega, \mathscr{G}, \mathbb{P})$.

EXERCISE 1.2. Conditioning reduces variance, i.e.,

$$\mathrm{Var}(X) \ge \mathrm{Var}\left(\mathbb{E}\{X|\mathscr{G}\}\right).$$

EXERCISE 1.3. Suppose that $X = \{X_n,\ n \ge 0\}$ is an i.i.d. integrable sequence. Let $S_n = X_1 + X_2 + \cdots X_n$. Then $\{S_n/n, n \ge 1\}$ is a reversed martingale with respect to the reversed filtration

$$\mathscr{F}^n = \sigma\{S_m,\ m \ge n\}, \quad n \in \mathbb{Z}_+.$$

EXERCISE 1.4. If a sequence X is adapted to $\mathscr{F}_*$ and τ is a stopping time, then $X_\tau \in \mathscr{F}_\tau$.

EXERCISE 1.5. Let X be a submartingale and σ and τ two bounded stopping times. Then

$$\mathbb{E}\{X_\tau|\mathscr{F}_\sigma\} \ge X_{\sigma \wedge \tau}.$$

EXERCISE 1.6. Let $\{X_n\}$ be a submartingale and $X_N^\dagger = \min_{0\le n\le N} X_n$. Then for any $\lambda > 0$, we have

$$\lambda \mathbb{P}\left[X_N^\dagger \le -\lambda\right] \le \mathbb{E}[X_N - X_1] - \mathbb{E}\left[X_N; X_N^\dagger \le -\lambda\right].$$

We also have for any $\lambda > 0$

$$\lambda \mathbb{E}\left[\max_{0\le n\le N} |X_n| \ge \lambda\right] \le \mathbb{E}|X_0| + 2\mathbb{E}|X_N|.$$

EXERCISE 1.7. If there is a C and $p > 1$ such that $\mathbb{E}|X_n|^p \le C$, then $\{X_n\}$ is uniformly integrable.

EXERCISE 1.8. Find a submartingale sequence such that $\mathbb{E}X_n^+ \le C$, but $\mathbb{E}X_n \to \mathbb{E}X_\infty$ does not hold.

EXERCISE 1.9. Let $\{X_i,\ i \in I\}$ be a uniformly integrable family of random variables. Then the family

$$\left\{X_i + X_j,\ (i,j) \in I \times I\right\}$$

is also uniformly integrable.

EXERCISE 1.10. Let M be a square integrable martingale. For any partition $\Delta = \{0 = t_0 < t_1 < \cdots < t_n = t\}$ of the interval $[0,t]$ we have

$$\mathbb{E}\sum_{j=1}^n \left[M_{t_j} - M_{t_{j-1}}\right]^2 = \mathbb{E}\langle M, M\rangle_t.$$

CHAPTER 2

Brownian Motion

In stochastic analysis, we deal with two important classes of stochastic processes: Markov processes and martingales. Brownian motion is the most important example for both classes, and is also the most thoroughly studied stochastic process. In this chapter we discuss Brownian motion only to the extent of what is needed for these lectures.

1. Stochastic processes

In this section we make some remarks about stochastic processes in general. We define a stochastic process X as a collection of random variables $\{X_t,\ t \geq 0\}$ on a measurable space $(\Omega, \mathscr{F})$. If we have a probability space $(\Omega, \mathscr{F}, \mathbb{P})$, then for each finite set of time points $T = \{t_1, \cdots, t_n\} \subset \mathbb{R}_+$, we have a Borel measure μ_T^X defined on $\mathbb{R}^n$ by

$$\mu_T^X(C) = \mathbb{P}\left\{(X_{t_1}, \cdots, X_{t_n}) \in C\right\}.$$

Each μ_T^X is called a finite dimensional marginal distribution of the process X. A theorem in measure theory (more precisely, Kolmogorov's extension theorem) says that the set of finite dimensional distributions

$$\left\{\mu_T^X :\ T \quad \text{finite subsets of} \quad \mathbb{R}_+\right\}$$

determines a unique probability measure μ^X on the product measurable space $\left(\mathbb{R}^{\mathbb{R}_+}, \mathscr{B}^{\mathbb{R}_+}\right)$ such that

$$\mu^X\left(\pi_T^{-1} C\right) = \mu_T^X(C),$$

where $C \in \mathscr{B}(\mathbb{R}^n)$ and $\pi_T : \mathbb{R}^{\mathbb{R}_+} \to \mathbb{R}^n$ is the projection

$$\pi_T(x) = \left\{x_{t_1}, \ldots, x_{t_n}\right\}, \quad x = \{x_t, t \in \mathbb{R}_+\} \in \mathbb{R}^{\mathbb{R}_+}.$$

Equivalently, μ^X is characterized by the relation

$$\mu^X\left\{x :\ (x_{t_1}, \ldots, x_{t_n}) \in C\right\} = \mathbb{P}\left\{\omega : (X_{t_1}(\omega), \ldots, X_{t_n}(\omega)) \in C\right\}.$$

The probability measure μ^X is called the law of the stochastic process.

We can regard a stochastic process X as a measurable map

$$X : (\Omega, \mathscr{F}) \to \left(\mathbb{R}^{\mathbb{R}_+}, \mathscr{B}(\mathbb{R})^{\mathbb{R}_+}\right)$$

defined by $\omega \mapsto \{X_t(\omega)\}$. Then the law of X is simply the induced measure $\mu^X = \mathbb{P} \circ X^{-1}$ defined by

$$\mu^X(A) = \mathbb{P}\left\{\omega : \ X(\omega) \in A\right\}, \quad A \in \mathscr{B}(\mathbb{R})^{\mathbb{R}_+}.$$

The sample path space $\left(\mathbb{R}^{\mathbb{R}_+}, \mathscr{B}(\mathbb{R})^{\mathbb{R}_+}\right)$ is too big to have any practical use. If the process X has certain special sample path properties, we can greatly reduce the sample path space. We say that X has continuous sample paths or X is a continuous process if

$$\mathbb{P}\left\{\omega : t \mapsto X_t(\omega) \text{ is continuous}\right\} = 1.$$

If this is the case, then the process can be regarded as a map

$$X : \Omega \to C(\mathbb{R}_+, \mathbb{R})$$

from the underlying space Ω to the space $W(\mathbb{R}) = C(\mathbb{R}_+, \mathbb{R})$ of continuous functions from $\mathbb{R}_+ = [0, \infty)$ to $\mathbb{R}$. The space $W(\mathbb{R})$ is a metric space under the distance function

$$\|f - g\| = \sum_{N=1}^{\infty} \frac{1}{2^N} \frac{\|f - g\|_{\infty,N}}{1 + \|f - g\|_{\infty,N}},$$

where

$$\|f\|_{\infty,N} = \max_{0 \leq t \leq N} |f(t)|.$$

Under this metric $f_n \to f$ if and only if $f_n(t) \to f(t)$ uniformly on every bounded interval. We denote the Borel σ-algebra on $W(\mathbb{R})$ by $\mathscr{B}(W(\mathbb{R}))$.

It is often convenient to consider the so-called coordinate process

$$\mathit{\Pi} = \{\mathit{\Pi}_t, \, t \geq 0\}$$

on $(W(\mathbb{R}), \mathscr{B}(W(\mathbb{R})))$. This is defined simply as $\mathit{\Pi}_t(w) = w_t$ for all $w \in W(\mathbb{R})$. This process generates a standard filtration

$$\mathscr{B}_*(W(\mathbb{R})) = \left\{\mathscr{B}_t(W(\mathbb{R})), \, t \geq 0\right\}.$$

We have the following basic fact.

PROPOSITION 1.1. *We have*

$$\mathscr{B}(W(\mathbb{R})) = \bigvee_{t \geq 0} \mathscr{B}_t(W(\mathbb{R})),$$

that is, the smallest σ-algebra generated by the coordinate process is precisely the Borel σ-algebra of the sample path space $W(\mathbb{R})$.

PROOF. Exercise. □

Suppose that X is a continuous stochastic process on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$. Note that by definition $X_t : \Omega \to \mathbb{R}$ is a random variable for each t. Then from the above PROPOSITION we see that

$$X : \, (\Omega, \mathscr{F}) \to (W(\mathbb{R}), \mathscr{B}(W(\mathbb{R}))$$

is a measurable map, or equivalently, X is a $W(\mathbb{R})$-valued random variable. The law $\mu^X = \mathbb{P} \circ X^{-1}$ is a probability measure on $C(\mathbb{R}_+, \mathbb{R})$. It is clear that the stochastic process X and the coordinate process Π have the same marginal distributions. In this sense Π on $(W(\mathbb{R}), \mathscr{B}(W(\mathbb{R})), \mu^X)$ is a standard copy of X, and for all practical purpose, we can regard X and Π as the same process.

Regarding a continuous stochastic process X as a measurable map from the underlying probability space to the sample path space $W(\mathbb{R})$ is a very convenient and useful point of view. A nice feature of $W(\mathbb{R})$ is that it has a natural shift operator $\theta_t : W(\mathbb{R}) \to W(\mathbb{R})$ defined by $(\theta_t w)_s = w_{s+t}$. .

REMARK 1.2. It is a curious fact that the space of continuous functions $C(\mathbb{R}_+, \mathbb{R})$ is *not* a measurable subset of the product measurable space $(\mathbb{R}^{\mathbb{R}_+}, \mathscr{B}(\mathbb{R})^{\mathbb{R}_+})$.

2. Brownian motion and its basic properties

DEFINITION 2.1. *A stochastic process $B = \{B_t, t \in \mathbb{R}_+\}$ defined on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ is called a Brownian motion if it has the following two properties: (1) B has independent increments, i.e., for any finite set of increasing nonnegative numbers $0 < t_1 < \cdots < t_n$, the random variables*

$$B_0, B_{t_1} - B_0, B_{t_2} - B_{t_1}, \ldots, B_{t_n} - B_{t_{n-1}} \tag{2.1}$$

are independent;

(2) for any $t > s \geq 0$, the distribution of the increment $B_t - B_s$ is $N(0, t-s)$, the Gaussian distribution with mean zero and variance $t - s$.

We can combine the two properties by saying that the joint distribution of the increments in (2.1) is the n-dimensional Gaussian distribution with zero mean vector and the diagonal covariance matrix

$$\operatorname{diag}(t_1 - t_0, t_2 - t_1, \cdots, t_n - t_{n-1}).$$

The distribution of B_0 is called the *initial distribution* of the Brownian motion B. If $\mathbb{P}\{B_0 = x\} = 1$ for a point $x \in \mathbb{R}^1$, we say that the Brownian motion B starts from x. It is clear that if B is a Brownian motion starting from 0, then $B + x = \{B_t + x, \ t \geq 0\}$ is a Brownian motion starting from x.

The above definition is a description on the finite dimensional marginal distributions of the Brownian motion except that the initial distribution is not explicitly given. Given a distribution (a probability measure) μ on $\mathbb{R}^1$, Brownian motion with initial distribution μ is unique in the sense that any two such Brownian motions have the same finite-dimensional distributions. In fact, the finite-dimensional marginal probability

$$\mathbb{P}\{B_0 \in A_0, \ B_{t_1} \in A_1, \ B_{t_n} \in A_n\}$$

is given by

$$\int_{A_0} \mu(dx_0) \int_{A_1} p(t_1 - t_0, x_0, x_1)\, dx_1 \cdots \int_{A_n} p(t_n - t_{n-1}, x_{n-1}, x_n)\, dx_n.$$

where

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

By Kolmogorov's extension theorem, the existence of a Brownian motion with any given initial distribution is immediate.

Depending on one's taste, one can add more properties into the definition of a Brownian motion. One can require that $B_0 = 0$. This makes Brownian motion into a Gaussian process characterized uniquely by the covariance function

$$\mathbb{E}\left\{B_s B_t\right\} = \min\left\{s, t\right\}.$$

Let B be a Brownian motion on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ such that $B_0 = 0$. Let μ be a probability distribution on $\mathbb{R}$. Then on the product probability space $(\mathbb{R} \times \Omega, \mathscr{B}(\mathbb{R}) \times \mathscr{F}, \mu \times \mathbb{P})$, the random variable $X(x, \omega) = x$ has the distribution μ and is independent of B. It is easy to see that $X + B = \{X + B_t,\ t \geq 0\}$ is a Brownian motion with initial distribution μ.

Some authors also require that a Brownian motion has continuous sample paths. This requirement calls immediately into the question the existence of Brownian motion with continuous sample paths. We will show in the next section that such Brownian motion indeed exists (Wiener's theorem) A continuous Brownian motion with initial value zero is usually called a standard Brownian motion.

We will also use a slight more general notion of a Brownian motion with respect to a filtration $\mathscr{F}_* = \{\mathscr{F}_t,\ t \geq 0\}$.

DEFINITION 2.2. *A stochastic process $B = \{B_t, t \in \mathbb{R}_+\}$ defined on a filtered probability space $(\Omega, \mathscr{F}_*, \mathbb{P})$ is called a Brownian motion with respect to the filtration $\mathscr{F}_*$ if it has the following two properties:*

(1) B is adapted to $\mathscr{F}_*$ and $B_t - B_s$ is independent of $\mathscr{F}_s$ for all $t \geq s \geq 0$;

(2) for any $t > s \geq 0$, the distribution of the increment $B_t - B_s$ is $N(0, t-s)$, the Gaussian distribution with mean zero and variance $t - s$.

It is clear that each Brownian motion B is a Brownian motion with respect to its own filtration $\mathscr{F}^B_*$. The following example shows why we need this slightly enlarged concept of Brownian motion.

EXAMPLE 2.3. An n-dimensional Brownian motion B is defined as $B_t = (B^1_t,\ B^2_t,\ B^n_t)$, where B^i are n independent Brownian motions. Let $\mathscr{F}^B_*$ be the filtration generated by B. Then each component B^i is a Brownian motion with respect to $\mathscr{F}^B_*$.

In the rest of this section we discuss a few basic and useful properties of Brownian motion.

PROPOSITION 2.4. *Let $B = \{B_t,\ t \geq 0\}$ be a Brownian motion starting from zero.*

1. MARKOV PROPERTY. *The process ${}^sB = \{B_{t+s} - B_s, t \geq 0\}$ is a Brownian motion independent of the σ-field $\mathcal{F}^B_s = \sigma\left\{B_u, 0 \leq u \leq s\right\}$;*

(2) SCALING PROPERTY. *For a nonzero c, the process $\{cB_{c^{-2}t}, t \geq 0\}$ is also a Brownian motion; in particular $-B = \{-B_t, t \geq 0\}$ is a Brownian motion;*

(3) TIME REVERSAL. *If B is a Brownian motion 0, then*

$$W = \{tB_{t^{-1}}, t \geq 0\}$$

(define $W_0 = 0$) is also a Brownian motion from 0.

PROOF. We prove the time-reversal property. Brownian motion B starting from zero can be uniquely characterized as a Gaussian process with mean zero and the covariance function $R(s,t) = \min\{s,t\}$. It is clear that the process $W_t = tB_{1/t}$ is a Gaussian process. We have

$$\mathbb{E}\{W_s W_t\} = st\mathbb{E}\{B_{1/s}B_{1/t}\} = st\min\{1/s, 1/t\} = \min\{s,t\}.$$

Hence W is a Gaussian process with the same covariance function as B. It therefore must also be a Brownian motion. □

PROPOSITION 2.5. *Let B be a Brownian motion with respect to a filtration $\mathscr{F}_*$.*

(1) B is a martingale with respect to $\mathscr{F}_*$ whose quadratic variation process is $\langle B, B\rangle_t = t$.

(2) For any real λ, the process

$$\exp\left[\lambda B_t - \frac{\lambda^2}{2}t\right], \quad t \geq 0,$$

is a martingale.

PROOF. Since $B_t - B_s$ independent of $\mathscr{F}_s$ with the distribution $N(0, t-s)$ we have

$$\mathbb{E}\{B_t - B_s | \mathscr{F}_s\} = \mathbb{E}(B_t - B_s) = 0.$$

Hence B is a martingale. Likewise, from

$$B_t^2 - B_s^2 = (B_t - B_s)^2 + 2(B_t - B_s)B_s,$$

we have

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

This shows that $B_t^2 - t$ is a martingale, which implies that the quadratic variation process is $\langle B, B\rangle_t = t$.

From $e^{\lambda B_t} = e^{\lambda B_s} e^{\lambda(B_t - B_s)}$ we have

$$\mathbb{E}\left\{e^{\lambda B_t} | \mathscr{F}_s\right\} = e^{\lambda B_s}\mathbb{E}e^{\lambda(B_t - B_s)} = e^{\lambda B_s} e^{\lambda^2(t-s)/2}.$$

Therefore $\exp\left[\lambda B_t - \lambda^2 t/2\right]$ is a martingale. □

3. Construction of Brownian motion

The goal of this section and the next section is to prove the existence of a Brownian motion with continuous sample paths. If $\{B^n_t,\, 0 \le t \le 1\}$ are independent copies of Brownian motions with time interval $[0,1]$, then the process defined by

$$B_t = B_n + B^{n+1}_{t-n}, \qquad n \le t < n+1,$$

is a Brownian motion with time interval $\mathbb{R}_+ = [0,\infty)$. Therefore it is e-nough to construct a Brownian motion on the time interval $[0,1]$.

Without sample path continuity, the existence of Brownian motion itself is guaranteed by the Kolmogorov extension theorem. We can also construct Brownian motion on $[0,1]$ more directly by using Fourier series. Suppose that B is a Brownian motion from 0. Let $W_t = B_t - tB_1$. The process W is called a Brownian bridge. Note that $W_0 = W_1 = 0$. Let us expand the sample paths of W into a Fourier sine series on $[0,1]$:

$$W_t = \sum_{n=1}^{\infty} X_n \sin n\pi t.$$

Of course the Fourier coefficients X_n are random variables and we have a formula for them:

$$X_n = 2\int_0^1 W_t \sin n\pi t\, dt.$$

From this formula it is clear that the random variables $\{X_n\}$ form a Gaussian system. All X_n have zero mean. Let us compute the covariance matrix. We have

$$\mathbb{E}\,[W_s W_t] = \min\{s,t\} - st.$$

Using this we have

$$\begin{aligned}
\mathbb{E}\,[X_m X_n] &= 4\mathbb{E}\left[\int_0^1 W_s \sin m\pi s\, ds \int_0^1 W_t \sin n\pi t\, dt\right] \\
&= 4\int_0^1\int_0^1 \mathbb{E}\,[W_s W_t] \sin m\pi s \sin n\pi t\, dsdt \\
&= 4\int_0^1\int_0^1 [\min\{s,t\} - st] \sin m\pi s \sin n\pi t\, dsdt.
\end{aligned}$$

After calculating the last integral we obtain

$$\mathbb{E}\,[X_m X_n] = \frac{2}{\pi^2 n^2}\delta_{mn}.$$

It follows from the above relation that $Z_n = n\pi X_n/\sqrt{2}$ is a sequence of i.i.d. random variables with the standard Gaussian distribution $N(0,1)$. Let $Z_0 = B_1$. Now we can write

$$B_t = tZ_0 + \frac{\sqrt{2}}{\pi}\sum_{n=1}^{\infty} \frac{Z_n}{n} \sin n\pi t, \qquad 0 \le t \le 1. \tag{3.1}$$

The covariance of Z_0 and Z_n is

$$\mathbb{E}\left[Z_0 Z_n\right] = \frac{n\pi}{\sqrt{2}} \int_0^1 \mathbb{E}\left[B_1 W_s\right] \sin n\pi s\, ds = 0.$$

This is because

$$\mathbb{E}\left[B_1 W_s\right] = \mathbb{E}[B_1(B_s - sB_1] = s - s = 0.$$

It follows that Z_0 is independent of Z_n. There in the Fourier expansion (3.1) of Brownian motion $\{Z_n, n \geq 0\}$ is an i.i.d. sequence with the normal distribution $N(0,1)$.

We can construct a Brownian motion by reversing the above argument. More precisely, suppose that $\{Z_n,\ n \geq 0\}$ is a sequence of i.i.d. random variables $\{Z_n, n \geq 0\}$ with the distribution $N(0,1)$ on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ (e.g., a product probability space). Then we define a stochastic process B on this probability space by the formula (3.1). For each fixed t, the series converges in $L^2(\Omega, \mathscr{F}, \mathbb{P})$ to a random variable B_t, and $B = \{B_t,\ 0 \leq t \leq 1\}$ is a Gaussian process. It can be verified directly that its mean is zero and its covariance function is given by $\mathbb{E}\left\{B_s B_t\right\} = \min\left\{s, t\right\}$. This shows that B is indeed a Brownian motion.

4. Continuity of Brownian motion

In the last section we have shown the existence of Brownian motion. In this section we show that there is a Brownian motion with continuous sample paths. There are many proofs of this result, but the classical proof presented here is still the best. It should be pointed out that this result only claims that there is a Brownian motion with continuous sample paths; it does not claim that every Brownian motion as defined before has continuous sample paths with probability one.

DEFINITION 4.1. *Let X and Y be two stochastic processes on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$. We say that X is a version of Y (and vice versa) if*

$$\mathbb{P}\left\{X_t = Y_t\right\} = 1 \quad \textit{for all} \quad t \geq 0.$$

Note that $X_t = Y_t$ with probability one for each fixed t. Since the time set $\mathbb{R}_+$ is uncountable, we cannot conclude that $\mathbb{P}\left\{X_t = Y_t \quad \text{for all} \quad t\right\} = 1$. Thus X and Y may have very different sample paths properties. However, X and Y have the same finite dimensional marginal distributions. If one of them is a B Brownian motion, so will be the other. We will prove the following result.

THEOREM 4.2. *(Wiener's theorem) Every Brownian motion has a continuous version with continuous sample paths.*

It is helpful to understand the above theorem from another point of view. Recall the path space

$$W(\mathbb{R}) = \text{the space of real-valued continuous functions on } \mathbb{R}_+ = [0, \infty).$$

For any fixed N let $\|f-g\|_{\infty,N} = \sup_{0\le t\le N}|f(t)-g(t)|$ for f and g in $W(\mathbb{R})$. If we equip $W(\mathbb{R})$ with the distance function

$$d(f,g) = \sum_{N=1}^{\infty} \frac{1}{2^n} \frac{\|f-g\|_{\infty,N}}{1+\|f-g\|_{\infty,N}},$$

then $W(\mathbb{R})$ becomes a metric space and the convergence in this metric is the uniform convergence on every finite interval. As a metric space $W(\mathbb{R})$ has its canonical Borel σ-field $\mathscr{B}(W(\mathbb{R}))$. Recall that the coordinate process X on $W(\mathbb{R})$ is defined by $X_t(w) = w_t$ for each $w \in W(\mathbb{R})$. The existence of a Brownian motion with continuous sample paths is equivalent to the following existence theorem.

THEOREM 4.3. *There is a probability measure μ on the measurable space $(W(\mathbb{R}), \mathscr{B}(W(\mathbb{R})))$ such that the coordinate process is a Brownian motion.*

The measure μ, whose existence is claimed in the above theorem, is called the Wiener measure on the space $(W(\mathbb{R}), \mathscr{B}(W(\mathbb{R})))$. In the final analysis, we are claiming the existence of a probability measure in the space $W(\mathbb{R}) = C(\mathbb{R}_+, \mathbb{R})$ of continuous functions such that

$$\mu\left\{w \in W(\mathbb{R}) : w_{t_1} \in A_1, w_{t_2} \in A_2, \ldots, w_{t_n} \in A_n\right\} =$$

$$\int_{A_1} p(t_1, x_1)\, dx_1 \int_{A_2} p(t_2 - t_1, x_2 - x_1)\, dx_2 \cdots \int_{A_n} p(t_n - t_{n-1}, x_n - x_{n-1})\, dx_n.$$

From this point of view, we see that this is a highly nontrivial result.

The rest of this section is devoted to the proof of THEOREM 4.2. We start from a Brownian motion $B = \{B_t,\ 0 \le t \le 1\}$ some probability space $(\Omega, \mathscr{F}, \mathbb{P})$ and proceed to construct a version $\tilde{B}$ on the same probability space such that $\mathbb{P}\left\{t \mapsto \tilde{B}_t \text{ is continuous for } 0 \le t \le 1\right\} = 1$.

Let

$$D = \{k/2^n : n = 0, 1, \ldots, k = 0, 1, \ldots 2^n\}$$

be the set of dyadic rational numbers in [0,1]. D is a countable set. For the construction of a continuous Brownian motion, we only need B_t for $t \in D$. We will show that with probability one, B is uniformly continuous on D. In fact, we prove more: with probability one, B is Hölder continuous on D with any exponent $\alpha < 1/2$.

PROPOSITION 4.4. *Fix $0 < \alpha < 1/2$. There is a set $\Omega_0 \subset \Omega$ of full measure $\mathbb{P}(\Omega_0) = 1$ with the following property. For every $\omega \in \Omega_0$, there is a constant $C(\omega)$ such that*

$$|B_t(\omega) - B_s(\omega)| \le C(\omega)|t-s|^\alpha \quad \textit{for all } s, t \in D. \tag{4.1}$$

PROOF. Consider the event:

$$A_n = \bigcup_{k=1}^{2^n} \left\{ \left|B_{k/2^n} - B_{(k-1)/2^n}\right| \ge 2^{-\alpha n} \right\}.$$

We estimate the probability $\mathbb{P}\left[A_n\right]$:

$$\begin{aligned}
\mathbb{P}[A_n] &\leq \sum_{k=1}^{2^n} \mathbb{P}\left[\left|B_{k/2^n} - B_{(k-1)/2^n}\right| \geq 2^{-\alpha n}\right] \\
&= \sum_{k=1}^{2^n} \mathbb{P}\left[\left|B_{1/2^n}\right| \geq 2^{-\alpha n}\right] \\
&= 2^n \mathbb{P}\left[\left|B_1\right| \geq 2^{(1-2\alpha)n/2}\right] \\
&\leq 2^n \cdot \frac{2}{\sqrt{2\pi}} 2^{-(1-2\alpha)n/2} \exp\left[-2^{(1-2\alpha)n} - 1\right].
\end{aligned}$$

Here we have used the time-homogeneity in the second step, the scaling property in the third step, and the elementary inequality

$$\int_x^\infty e^{-t^2/2} dt \leq \frac{1}{x} e^{-x^2/2}$$

in the fourth step. It follows from the above bound for $\mathbb{P}[A_n]$ and the hypothesis $0 < \alpha < 1/2$ that $\sum_{n=1}^\infty \mathbb{P}\left[A_n\right] < \infty$. By the Borel-Cantelli lemma, we have $\mathbb{P}\left\{A_n, \text{i.o.}\right\} = 0$. Let

$$\Omega_0 = \left\{A_n \text{ i.o. }\right\}^c = \bigcup_{N=1}^\infty \bigcap_{n=N}^\infty \bigcap_{k=1}^{2^n} \left\{\left|B_{k/2^n} - B_{(k-1)/2^n}\right| < 2^{-\alpha n}\right\}.$$

Then $\mathbb{P}\left\{\Omega_0\right\} = 1$.

Suppose that $\omega \in \Omega_0$. The sequence of the events $\{A_n\}$ do not happen infinitely often. There is an N_0 depending on ω such that $\omega \notin A_n$ for all $n \geq N_0$. This means that

$$\left|B_{k/2^n}(\omega) - B_{(k-1)/2^n}(\omega)\right| < 2^{-\alpha n} \qquad \text{for all } n \geq N_0 \text{ and } 1 \leq k \leq 2^n.$$

It is enough to show the inequality (4.1) for dyadic points $s < t$ such that $t - s$ is less than a prefixed positive number. Suppose that $0 < t - s < 2^{-N_0}$. The difference $t - s$ is also a dyadic number and we have

$$t - s = \sum_{n=N_1}^\infty \frac{k_n}{2^n}$$

with $k_n = 0$ or 1 and $k_{N_1} = 1$. Since $0 < t - s < 2^{-N_0}$ we must have $N_1 \geq N_0$. Let

$$s_n = s + \sum_{j=N_1}^n k_j/2^j.$$

It is a finite sequence of dyadic points that leads from s to t. The difference of successive time points $s_n - s_{n-1} = 0$ or 2^{-n} with $n \geq N_0$, hence we have

$|B_{s_n}(\omega) - B_{s_{n-1}}(\omega)| \le 2^{-\alpha n}$ for all $n \ge N_1$. Now we have

$$\begin{aligned} \left|B_t(\omega) - B_s(\omega)\right| &\le \sum_{n=N_1}^{\infty} \left|B_{s_n} - B_{s_{n-1}}\right| \\ &\le \sum_{n=N_1}^{\infty} 2^{-\alpha n} \\ &\le \frac{2^{-\alpha N_1}}{1-2^{-\alpha}}. \end{aligned}$$

On the other hand $|t-s| \ge 2^{-N_1}$. It follows that

$$|B_t(\omega) - B_s(\omega) \le \frac{|t-s|^{\alpha}}{1-2^{-\alpha}}$$

for all dyadic points s and t such that $|t-s| < 2^{-N_0}$. □

We can now show the existence of a continuous Brownian motion. For each $\omega \in \Omega_0$ and $t \in [0,1]$ we define

$$\tilde{B}_t(\omega) = \lim_{D \ni r \to t} B_r(\omega).$$

The above proposition shows that $\tilde{B}_t(\omega)$ is well defined and $t \mapsto B_t(\omega)$ is continuous. To complete the proof, it is enough to show that $\tilde{B}$ is a version of B. From $\mathbb{E}|B_s - B_t|^2 = |s-t|$, we have

$$\mathbb{E}|\tilde{B}_t - B_t|^2 \le \liminf_{D \ni r \to t} \mathbb{E}|B_r - B_t|^2 = 0.$$

Therefore $\mathbb{P}\left\{\tilde{B}_t = B_t\right\} = 1$ for all $t \ge 0$ and $\tilde{B}$ is a version of B.

5. Strong Markov property

Recall the Markov property of Brownian motion. For any fixed time $s \ge 0$, the process ${}^sB$ defined by ${}^sB_t = B_{t+s} - B_s, t \ge 0$ is a Brownian motion from zero independent of the σ-field $\mathscr{F}_s^B = \sigma\{B_u; 0 \le u \le s\}$. The strong Markov property is more general. It asserts that the above Markov property holds for a class of random times called stopping times. In many applications, the possibility of using stopping times is the source of power of stochastic methods

Let us first discuss briefly random times. A *random time* is simply a nonnegative measurable random variable $\tau : \Omega \to [0,\infty]$. Sometimes the value ∞ is allowed. If $\mathbb{P}\{\tau < \infty\} = 1$, we say that τ is a finite random time. A typical random time is the first passage time τ_a of a point of $a \in \mathbb{R}$ defined by

$$\tau_a = \inf\{t \ge 0 : B_t(\omega) = a\}.$$

We can show that τ_a is a finite random time. Note that $B_{\tau_a} = a$. In general if τ is a random time, then B_τ stands for the random variable $B_\tau(\omega) = B_{\tau(\omega)}(\omega)$. Note that B_τ is indeed a random variable.

The stochastic process ${}^{\tau_a}B = \{B_{t+\tau_a} - a, t \geq 0\}$ is a continuous process starting from 0. We can ask the natural questions: Is ${}^{\tau_a}B$ a Brownian motion independent of the Brownian motion up to time τ_a? To answer this question, we must first know how to define $\mathscr{F}_\tau^B$ for a random time τ. Intuitively this σ-algebra represents the information up to time τ. Once this is properly done the answer to the above question is affirmative and follows from the usual Markov property of Brownian motion.

Before we state the strong Markov property, it is perhaps helpful to point out by a counterexample that this strong Markov property is not shared by all random times. Let λ_0 be the last time the Brownian path is at the point $z = 0$ before time 1:

$$\lambda_0 = \sup\{t \in [0,1] : B_t = 0\}.$$

Since $B_1 \neq 0$ with probability 1, we have $\mathbb{P}\{\lambda_0 < 1\} = 1$. The process ${}^{\lambda_0}B = \{B_{t+\lambda_0};\, t \geq 0\}$ is not a Brownian motion from 0. This is because $t = 0$ is an isolated zero, whereas for a Brownian motion $t = 0$ cannot be an isolated zero.

REMARK 5.1. The law of λ_0 is known:

$$\mathbb{P}\{\lambda_0 \leq t\} = \frac{2}{\pi} \arcsin \sqrt{t}.$$

This is one of Lévy's arcsine laws.

The above counterexample shows that random times at which Brownian motion has the Markov property are special. The problem with the random time λ_0 is that if we follow a Brownian traveler from time 0 to a fixed time $t < 1$, we cannot tell if the event $\{\lambda_0 \leq t\}$ has occurred or not, since to find this out we have to look beyond time t; in fact we have to look at the path all the way to time 1. On the other hand, the first passage time τ_a is a time which, so to speak, does not depend on the future.

We now introduce a class of random times for which we will prove the strong Markov property.

DEFINITION 5.2. *A nonnegative random variable $\tau : \Omega \to [0,\infty]$ on a filtered measurable space $(\Omega, \mathscr{F}_*)$ is called a stopping time (with respect to the filtration $\mathscr{F}_*$) if $\{\tau \leq t\} \in \mathscr{F}_t$ for all $t \geq 0$. If $\mathscr{F}_* = \mathscr{F}_*^X$ is the filtration generated by a stochastic process, a stopping time with respect to $\mathscr{F}_*^X$ is usually called a stopping time of the process X.*

The intuitive meaning of the above definition is as follows. Suppose that τ is the time a certain event associated to a process X takes place. Then τ is a stopping time if the question of whether the event takes place or not before time t can be determined by looking at the path of the process $\{X_s,\, 0 \leq s \leq t\}$ up to time t.

In the next proposition we list below a few properties of stopping times.

PROPOSITION 5.3. *The following statements hold.*

(1) *If σ and τ are stopping times, then $\sigma + \tau$, $\sigma \wedge \tau$, and $\sigma \vee \tau$ are stopping times;*
(2) *If $\sigma_n, n \geq 1$ is an increasing sequence of stopping times then the limit $\tau_\infty = \lim_{n\to\infty} \sigma_n$ is a stopping time;*
(3) *If $\sigma_n, n \geq 1$ is a decreasing sequence of stopping times, and for each $\omega \in \Omega$ there is an $N(\omega)$ such that $\sigma_n(\omega) = \sigma_{n+1}(\omega)$ for all $n \geq N(\omega)$, then the limit $\sigma_\infty = \lim_{n\to\infty} \sigma_n$ is a stopping time.*

PROOF. These assertions follow more or less directly from the definition of a stopping time. □

EXAMPLE 5.4. When proving a statement involving a stopping time we often approximate it by a sequence of discrete stopping times and prove the statement on each set on which a discrete stopping time takes a fixed constant value. Let τ be a stopping time and define

$$\tau_n = \frac{[2^n \tau] + 1}{2^n} = \frac{k}{2^n}, \quad \text{if } \frac{k-1}{2^n} \leq \tau < \frac{k}{2^n}.$$

Here $[x]$ denotes the largest integer not exceeding x. We can verify that $\{\tau_n\}$ is a decreasing sequence of stopping times and $\tau_n \downarrow \tau$.

We now define $\mathscr{F}_\tau$, the σ-algebra up to a stopping time τ. If $\mathscr{F}_* = \mathscr{F}_*^X$ is generated by a stochastic process X, then it is natural to define

$$\mathscr{F}_\tau^X = \sigma\left\{X_{t\wedge\tau},\, t \geq 0\right\},$$

the σ-algebra generated by $X^\tau = \{X_{t\wedge\tau}, t \geq 0\}$, the process X stopped at time τ. In general, when the filtration is not necessarily generated by a stochastic process, we use the following definition.

DEFINITION 5.5. *Let τ be a stopping time on a filtered measurable space $(\Omega, \mathscr{F}_*)$. The σ-algebra $\mathscr{F}_\tau$ is the collection of events $A \in \mathscr{F}_\infty$ such that*

$$A \cap \{\tau \leq t\} \in \mathscr{F}_t \quad \textit{for all } t \geq 0.$$

We can verify that $\mathscr{F}_\tau$ is indeed a σ-algebra and $\mathscr{F}_\tau = \mathscr{F}_t$ if τ is the constant stopping time $\tau = t$ (EXERCISE **??**). For $\mathscr{F}_*^X$, the filtration generated by a process X, the σ-algebra $\mathscr{F}_\tau^X$ defined above coincides with the σ-algebra generated by the stopped process X^τ (see EXERCISE **??**).

Here are some properties of $\mathscr{F}_\tau$ we will need in the future.

PROPOSITION 5.6. *Let σ, τ be $\mathscr{F}_t$-stopping times.*

(1) *If $\sigma \leq \tau$, then $\mathscr{F}_\sigma \subset \mathscr{F}_\tau$;*
(2) *$\tau \in \mathscr{F}_\tau$;*
(3) *If $A \in \mathscr{F}_\tau$, then $A \cap \{\sigma \leq \tau\} \in \mathscr{F}_\tau$ an $A \cap \{\sigma = \tau\} \in \mathscr{F}_\tau$; in particular, $\{\sigma \leq \tau\} \in \mathscr{F}_\sigma \cap \mathscr{F}_\tau$;*

PROOF. These properties follow easily from the definition of stopping times. □

PROPOSITION 5.7. *Suppose that $\mathscr{F}_*$ satisfies the usual conditions. If X is a continuous stochastic process adapted to $\mathscr{F}_*$ and τ a finite $\mathscr{F}_*$-stopping time, then $X_\tau \in \mathscr{F}_\tau$.*

PROOF. We consider the case where X is a real-valued process. Let

$$\tau_n = \frac{[2^n \tau] + 1}{2^n}$$

be the discrete approximation of τ in EXAMPLE 5.4. By considering each set on which τ_n is constant, we can show easily that $X_{\tau_n} \in \mathscr{F}_{\tau_n}$. In particular, for $a \in \mathbb{R}$ and $t > 0$ we have

$$\{X_{\tau_n} \le a\} \cap \{\tau_n < t\} \in \mathscr{F}_t.$$

Using the sample path continuity of X and the fact that $\tau_n \downarrow \tau$ we can verify that

$$\{X_\tau \le a\} \cap \{\tau < t\} = \bigcap_{l=1}^{\infty} \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} \left\{ X_{\tau_k} \le a + \frac{1}{l} \right\} \cap \{\tau_k < t\}.$$

This implies that $\{X_\tau \le a\} \cap \{\tau < t\} \in \mathscr{F}_t$, hence by the usual conditions

$$\{X_\tau \le a\} \cap \{\tau \le t\} \in \mathscr{F}_{t+} = \mathscr{F}_t.$$

From this relation and the definition of $\mathscr{F}_\tau$ we see that $\{X_\tau \le a\} \in \mathscr{F}_\tau$ for all a, which implies immediately that X_τ is measurable with respect to $\mathscr{F}_\tau$. □

Now we are in a position to state and prove the strong Markov property of Brownian motion. Let B be a Brownian motion with respect to a filtration $\mathscr{F}_*$ and τ a finite stopping time with respect to the same filtration. The strong Markov property is the assertion that the shifted process ${}^\tau B = \{B_{t+\tau} - B_\tau,\ t \ge 0\}$ is a Brownian motion independent of the σ-algebra $\mathscr{F}_\tau$. There are two assertions in this theorem: (1 the process ${}^\tau B$ is a Brownian motion (2) this Brownian motion is independent of $\mathscr{F}^B_\tau$. They are immediate consequences of the equality in the next theorem.

PROPOSITION 5.8. *Let $0 \le t_1 < t_2 < \ldots < t_m$ be m time points. Let $f : \mathbb{R}^m \to \mathbb{R}$ be a bounded continuous function on $\mathbb{R}^m$. Let $C \in \mathscr{F}_\tau$. Then*

$$\mathbb{E}\left[f\left({}^\tau B_{t_1}, \cdots, {}^\tau B_{t_m}\right); C \right] = \mathbb{P}[C] \mathbb{E}\left[f\left(B_{t_1}, \cdots, B_{t_m}\right) \right]. \tag{5.1}$$

PROOF. We first prove the equality for discrete stopping times and then approximate a general finite stopping time by discrete stopping times as in EXAMPLE 5.4.

Assume that τ is a discrete stopping time whose possible values are s_i. Since B is a Brownian motion with respect to $\mathscr{F}_*$, we have the usual Markov property for a constant time: ${}^s B = \{B_{t+s} - B_s; t \ge 0\}$ is a Brownian motion independent of $\mathscr{F}_s$. Assume that $C \in \mathscr{F}_\tau$ and let $C_i = C \cap \{\tau = s_i\}$. Then $C_i \in \mathscr{F}_{s_i}$ by PROPOSITION 5.6 (iii). We have

$$\begin{aligned}
\mathbb{E}\left[f\left({}^\tau B_{t_1}, \cdots, {}^\tau B_{t_m}\right); C\right] &= \sum_{i=1}^{\infty} \mathbb{E}\left[f\left(B_{t_1+s_1} - B_{s_1}, \cdots, B_{t_m+s_i} - B_{s_i}\right); C_i\right] \\
&= \sum_{i=1}^{\infty} \mathbb{P}\left[C_i\right] \mathbb{E}\left[f\left(B_{t_1+s_1} - B_{s_1}, \cdots, B_{t_m+s_i} - B_{s_i}\right)\right] \\
&= \sum_{i=1}^{\infty} \mathbb{P}\left[C_i\right] \mathbb{E}\left[f\left(B_{t_1}, \cdots, B_{t_m}\right)\right] \\
&= \mathbb{P}[C]\mathbb{E}\left[f\left(B_{t_1}, \cdots, B_{t_m}\right)\right].
\end{aligned}$$

Thus (5.1) holds for discrete stopping times.

For a general finite stopping time τ, we approximate it by the decreasing sequence of stopping times

$$\tau_n = \frac{[2^n \tau] + 1}{2^n}.$$

The strong Markov property holds for each τ_n. Now suppose that $C \in \mathscr{F}_\tau$. Since $\tau \leq \tau_n$, we have $\mathscr{F}_\tau \subset \mathscr{F}_{\tau_n}$ (PROPOSITION 5.6 (i)), hence $C \in \mathscr{F}_{\tau_n}$. We have

$$\mathbb{E}\left[f\left({}^\tau B_{t_1}, \cdots, {}^\tau B_{\tau_m}\right); C\right] = \mathbb{P}[C]\mathbb{E}\left[f\left(B_{t_1}, \cdots, B_{t_m}\right)\right].$$

We have ${}^{\tau_n} B_{t_j} \to^\tau B_{t_j}$ because Brownian motion has continuous sample paths. Therefore we can use the the dominated convergence theorem to take limit on the left side and obtained the desired equality (5.1). □

The strong Markov property of Brownian motion now follows easily from the equality proved in the above proposition.

THEOREM 5.9. *Let B be a Brownian motion with respect to a filtration $\mathscr{F}_*$ and τ a finite stopping time with respect to the same filtration. Then the shifted process ${}^\tau B = \{B_{t+\tau} - B_\tau,\ t \geq 0\}$ is a Brownian motion independent of the σ-algebra $\mathscr{F}_\tau$.*

PROOF. Everything is contained in the equality (5.1). Random variables of the form $f({}^\tau B_{t_1}, \cdots, {}^\tau B_{t_m})$ is dense in the space of integrable random variables measurable with respect to the process ${}^\tau B$. The fact that the right side of the equality is a product shows that $\mathscr{F}_\tau$ is independent of ${}^\tau B$. If we take $C = \Omega$ then the equality says that ${}^\tau B$ has the same finite dimensional marginal distributions as B, which implies that ${}^\tau B$ is a Brownian motion. □

REMARK 5.10. An equivalent statement of the strong Markov property is as follows. Let B be a Brownian motion and τ a finite stopping time, both with respect to a given filtration $\mathscr{F}_*$. Then the process $\{B_{t+\tau},\ t \geq 0\}$ is a Brownian motion with respect to the filtration $\mathscr{F}_{\tau+*} = \{\mathscr{F}_{t+\tau},\ t \geq 0\}$. This is a very useful restatement in view of future generalization of the strong

Markov property to a general diffusion process obtained from solving a stochastic differential equation.

6. Applications of the strong Markov property

In this section we discuss several properties of Brownian motion which can be proved by strong Markov property. Recall that this property claims that ${}^{\tau}B = \{B_{t+\tau} - B_\tau,\ t \geq 0\}$ is a Brownian motion independent of $\mathscr{F}_\tau$. This property is often used in the following form. Let X and Y be two random variables measurable with respect to ${}^{\tau}B$ and $\mathscr{F}_\tau$, respectively. Then the joint distribution of (X, Y) must be the product measure of the marginal distributions of X and Y. This allows us to use Fubini's theorem to calculate the expected value of a function of the two variables:

$$\begin{aligned}\mathbb{E}f(X,Y) &= \int_{E_1\times E_2} f(x,y)\,\mathbb{P}(X\in dx)\mathbb{P}(Y\in dy)\\ &= \mathbb{E}\left[\mathbb{E}[f(x,Y)]|_{x=X}\right]\\ &= \mathbb{E}\left[\mathbb{E}[f(X,y)]|_{y=Y}\right].\end{aligned}$$

PROPOSITION 6.1. *Let $M_t = \max_{0\leq s\leq t} B_s$ be the maximum process of a Brownian motion from 0. Then M_t and $|B_t|$ have the same distribution.*

PROOF. Let $b > 0$ and

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

be the first passage time of Brownian motion to b. It is a finite stopping time. We compute the probability $\mathbb{P}\left[B_t \geq b, \tau_b \leq t\right]$ by the strong Markov property. First note that $\{B_t = b\}$ implies $\{\tau_b \leq t\}$, therefore this probability is just $\mathbb{P}\left\{B_t \geq b\right\}$. Now let

$$W_t = B_{t+\tau_b} - B_{\tau_b} = B_{t+\tau_b} - b$$

the shifted process. Then the probability we wanted to compute can be written as $\mathbb{P}\left[W_{t-\tau_b} \geq 0, \tau_b \leq t\right]$. It is in the form of the expected value of a function of two random variables W and $\tau_b \in \mathscr{F}_{\tau_b}$, which are independent by the strong Markov property. By Fubini's theorem, we can compute this expectation by assuming one of the variables to be fixed first. In the present case we can regard τ_b as a constant and compute the probability that $W_{t-s} \geq 0$, which is of course equal to $1/2$ because W is a Brownian motion. It follows that

$$\mathbb{P}[B_t \geq b] = \mathbb{P}\left[B_t \geq b, \tau_b \leq t\right] = \frac{1}{2}\mathbb{P}\left[\tau_b \leq t\right].$$

On the other hand, we have $\{\tau_b \leq t\} = \{M_t \geq b\}$. Therefore,

$$\mathbb{P}\left[M_t \geq b\right] = 2\mathbb{P}\left[B_t \geq b\right] = \mathbb{P}\left[|B_t| \geq b\right].$$

This shows that M_t and $|B_t|$ have the same distribution. $\square$

We can calculate the density function of the first passage time τ_b. From the above result we have

$$\mathbb{P}\{\tau_b \le t\} = \frac{2}{\sqrt{2\pi t}} \int_b^\infty e^{-x^2/2t}\, dx.$$

Differentiating with respect to t and integrating by parts we obtain the density function

$$p_{\tau_b}(t) = \frac{b}{\sqrt{2\pi t^3}} e^{-b^2/2\,t} dt.$$

THEOREM 6.2. *Let B be a Brownian motion from 0 and τ a stopping time of B, both with respect to a given filtration $\mathscr{F}_*$. Let W be another Brownian motion from 0 independent of $\mathscr{F}_\tau$. Define a new process Z by*

$$Z_t = \begin{cases} B_t & \text{if } t \le \tau, \\ W_{t-\tau} + B_\tau & \text{if } t > \tau\,. \end{cases}$$

Then Z is a Brownian motion from 0.

PROOF. By the strong Markov property ${}^\tau B$ is a Brownian motion independent of (B^τ, τ), where $B^\tau = \{B_{t\wedge\tau},\, t \ge 0\}$ is the Brownian motion stopped at τ. By the assumption the triples (W, B^τ, τ) and $({}^\tau B, B^\tau, \tau)$ have the same distribution. The processes Z and B are constructed from the two triples in the same way. More precisely, there is a measurable function F such that $Z = F(W, B^\tau, \tau)$ and $B = F({}^\tau B, B^\tau, \tau)$. This shows that Z and B must have the same distribution, which implies that Z is a Brownian motion. □

COROLLARY 6.3. *(André's reflection principle) Let τ be a finite stopping time of a Brownian motion B from 0. Let Z be the process obtained from B by reflecting with respect to B_τ after the time τ:*

$$Z_t = \begin{cases} B_t & \text{if } t \le \tau, \\ 2B_\tau - B_t & \text{if } t > \tau. \end{cases}$$

Then Z is a Brownian motion.

PROOF. $-{}^\tau B$, like ${}^\tau B$ itself, is also independent of $\mathscr{F}_\tau$. We can then take $W = -{}^\tau B$ in the theorem. □

We use the reflection principle to compute the joint distribution of B_t and $M_t = \max_{0\le s\le t} B_s$.

PROPOSITION 6.4. *The joint distribution of M_t and B_t is given by*

$$\mathbb{P}\left[B_t \in da, M_t \in db\right] = \left(\frac{2}{\pi t^3}\right)^{1/2} (2b-a) e^{-(2b-a)^2/2\,t} da\, db$$

for $a < b$ and $b > 0$.

PROOF. Let Z be the Brownian motion obtained B by reflecting its path after the first passage time τ_b. For clarity we introduce the notations τ_b^B and τ_b^Z for the first passage times to b of the Brownian motions B and Z. Of course these two times are identical. For $a < b$ and $b > 0$, we have

$$\begin{aligned}\mathbb{P}\left[B_t \le a, M_t \ge b\right] =& \mathbb{P}\left[B_t \le a, \tau_b^B \le t\right] \\ =& \mathbb{P}\left[Z_t \le a, \tau_b^Z \le t\right] \\ =& \mathbb{P}\left[B_t \ge 2b - a, \tau_b^B \le t\right] \\ =& \mathbb{P}\left[B_t \ge 2b - a\right].\end{aligned}$$

Let's explain these steps. In the first step we have used the identity

$$\{M_t \ge b\} = \left\{\tau_b^B \le t\right\}.$$

In the second step we have simply replaced B by Z because they are both Brownian motions by the reflection principle. In the third step we have replaced τ_b^Z by τ_b^B because they are identical, and we have used $\{Z_t < a\} = \{B_t > 2b - a\}$. This holds because $Z_t = 2b - B_t$ if $t > \tau_b^B$ by the definition of Z. In the fourth step we have used the implication

$$\{B_t > 2b - a\} \subset \{B_t > b\} \subset \left\{\tau_t^B \le t\right\},$$

which follows from the assumption that $b > a$. Now we have

$$\mathbb{P}\left[B_t \le a, M_t \ge b\right] = \mathbb{P}\left[B_t \ge 2b - a\right] = \frac{1}{\sqrt{2\pi t}} \int_{2b-a}^{\infty} e^{-x^2/2t} dx.$$

Differentiating with respect to a and then with respect to b we obtain the joint density function of B_t and M_t. □

COROLLARY 6.5. *$M_t - B_t$ and $|B_t|$ has the same distribution for each fixed t.*

PROOF. Integrate the joint density function $p_{M_t,B_t}(b,a)$ on the region $b - a \ge c$. □

REMARK 6.6. It is a perhaps a curious observation that $M_t - B_t$ and M_t have the same distribution, since both are distributed as $|B_t|$. The process $|B| = \{|B_t|, t \ge 0\}$ is called a reflecting Brownian motion. Reflecting Brownian motion will be discussed in great detail later. Among other things we will show the much stronger result that the processes $\{M_t - B_t, t \ge 0\}$ and $\{|B_t|, t \ge 0\}$ have the same distribution.

7. Quadratic variation of Brownian motion

In this section we study Brownian motion from the point of view of martingale theory. We know that Brownian motion has continuous sample paths; in fact we have shown that sample paths are Hölder continuous for any exponent $\alpha < 1/2$. On the other hand, it can be shown that the sample paths are nowhere differentiable. Thus they cannot be functions of bounded variation. Nevertheless they have finite quadratic variations. We have shown that both B_t and $B_t^2 - t$ are continuous martingales. This implies that the quadratic variation process $\langle B, B\rangle_t = t$.

Fix a time $t > 0$ and consider a partition

$$\Delta:\ 0 = t_0 < t_1 < \cdots < t_{n-1} < t_n = t$$

of the interval $[0, t]$. Denote $|\Delta| = \max_{1\le i\le n}(t_i - t_{i-1})$, the length of the longest intervals. We write $\Delta_1 \subset \Delta_2$ and say that Δ_2 is a refinement of Δ_1 if every partition point in Δ_1 is also a partition point of Δ_2, or equivalently, every subinterval in Δ_1 is a union of subintervals in Δ_2.

The quadratic variation of Brownian motion B along a partition Δ is defined to be

$$Q(B;\Delta) = \sum_{i=1}^{n} |B_{t_i} - B_{t_{i-1}}|^2.$$

THEOREM 7.1. *We have*

$$\lim_{|\Delta|\to 0} \mathbb{E}\left|\sum_{i=1}^{n} |B_{t_i} - B_{t_{i-1}}|^2 - t\right| = 0.$$

PROOF. We have

$$\mathbb{E}\,|B_t|^2 = t \quad \text{and} \quad \mathbb{E}\big|B_t^2 - t\big|^2 = 2t^2.$$

Let $\Delta B_i = B_{t_i} - B_{t_{i-1}}$ and $\Delta t_i = t_i - t_{i-1}$. Then the expression under the limit is equal to

$$\sum_{i,j=1}^{n} \mathbb{E}\left[\left(|\Delta B_i|^2 - \Delta t_i\right)\left(|\Delta B_j|^2 - \Delta t_j\right)\right].$$

In the last sum, the off diagonal terms vanish because the increments are independent and each factor has zero expectation. For the diagonal term $i = j$ the expectation is equal to $2|\Delta t_i|^2$. Hence

$$\mathbb{E}\left|\sum_{i=1}^{n} |B_{t_i} - B_{t_{i-1}}|^2 - t\right| \le 2t|\Delta|.$$

The theorem follows immediately. □

REMARK 7.2. The existence of a bounded quadratic variation of a continuous function implies that it cannot be of bounded variation itself. Therefore almost surely Brownian motion paths have unbounded variation on any time interval.

REMARK 7.3. The above argument can be carried without much change to a general continuous martingale. Furthermore, if $|\Delta| \to 0$ sufficiently fast, then the quadratic variation along the partition Δ converges almost surely to t. This can be proved by using the Borel-Cantelli lemma.

Finally we state another much deeper almost sure version of the above theorem. See Freedman [**4**] for a proof.

THEOREM 7.4. *The quadratic variations $Q(B;\Delta_n)$ of Brownian motion B along a refining sequence of partitions $\Delta_1 \subset \Delta_2 \subset \cdots$ is a reverse martingale and hence converges almost surely. If $|\Delta_n| \to 0$, then $\lim_{n\to\infty} Q(B;\Delta_n) = t$ almost surely.*

PROOF. See Freedman [**4**]. □

8. Second assignment

EXERCISE 2.1. Let B be a Brownian motion. For any $0 \le r < s < t$ and Borel sets C, D on $\mathbb{R}^1$,

$$\mathbb{P}\left[B_r \in C, B_t \in D | B_s\right] = \mathbb{P}\left[B_r \in C | B_s\right] \mathbb{P}\left[B_t \in D | B_s\right].$$

This says that given the present, the past is independent of the future.

EXERCISE 2.2. Let $\{Z_n, n \ge 0\}$ be an i.i.d. sequence with the standard distribution $N(0,1)$. Define

$$B_t = tZ_0 + \frac{\sqrt{2}}{\pi}\sum_{n=1}^{\infty}\frac{Z_n}{n}\sin n\pi t, \qquad 0 \le t \le 1.$$

Show that B is a Brownian motion.

EXERCISE 2.3. Let $a \in \mathbb{R}$ and $\tau_a = \inf\{t \ge 0 : B_t = a\}$ be the first passage time of a. Show that τ_a is a finite stopping time, that is, $\mathbb{P}\{\tau_a < \infty\} = 1$.

EXERCISE 2.4. Let $Z = \{t \in \mathbb{R}_+ : B_t = 0\}$ be the zero set of the Brownian motion B. Show that with probability one the point $t = 0$ is not an isolated point of Z.

EXERCISE 2.5. Use the joint density function of $M_t = \max_{0\le s\le t} B_s$ and B_t to show that $M_t - B_t$ and $|B_t|$ has the same distribution.

EXERCISE 2.6. Let B be a Brownian motion. Show that

$$M_t = B_t^3 - 3\int_0^t B_s\, ds$$

is a martingale.

EXERCISE 2.7. Suppose that $\{\tau_n\}$ is a decreasing sequence of stopping times on a filtered measurable space $(\Omega, \mathscr{F}_*)$ such that for each $\omega \in \omega$ there is a natural number $N(\omega)$ such that $\tau_n(\omega) = \tau_{n+1}(\omega)$ for all $n \ge N(\omega)$. Show that $\tau = \lim_{n\to\infty}\tau_n$ is a stopping time.

EXERCISE 2.8. Suppose that the filtration $\mathscr{F}_*$ is right continuous, i.e., $\mathscr{F}_t = \mathscr{F}_{t+}$. Let τ be a finite stopping time. Let $\{\tau_n\}$ be a decreasing sequence of finite stopping times such that $\tau_n \downarrow \tau$. Show that $\bigcap_{n=1}^{\infty}\mathscr{F}_{\tau_n} = \mathscr{F}_\tau$.

EXERCISE 2.9. Let B be a standard Brownian motion and $\mathscr{F}_*^B$ its filtration. Define a new filtration $\mathscr{G}_*$ by $\mathscr{G}_t = \sigma\left\{\mathscr{F}_t^B, B_1\right\}$. Then B is no longer a martingale respect to $\mathscr{G}_*$. Let $X_t = B_t - tB_1$. Show that

$$W_t = X_t + \int_0^t \frac{X_s}{1-s}\, ds$$

is a martingale with respect to $\mathscr{G}_*$. Thus the Doob-Meyer decomposition of B with respect to the filtration $\mathscr{G}_*$ is

$$B_t = W_t + tB_1 - \int_0^t \frac{B_s - sB_1}{1-s}\, ds.$$

We can show later by Lévy's criterion that W is in fact a Brownian motion.

EXERCISE 2.10. For each $x \in \mathbb{R}$ we use $\mathbb{P}_x$ to denote the law of Brownian motion starting from x. Let $a < x < b$ and let τ_a and τ_b be the first passage times of Brownian motion to a and b respectively. Show that

$$\mathbb{P}_x\left\{\tau_a < \tau_b\right\} = \frac{b-x}{b-a}.$$

CHAPTER 3

Stochastic Integration and Ito's Formula

In this chapter we discuss Itô's theory of stochastic integration. This is a vast subject. However, our goal is rather modest: we will develop this theory only generally enough for later applications. We will discuss stochastic integrals with respect to a Brownian motion and more generally with respect to a continuous local martingale. Instead of attempting to describe the largest possible class of integrand processes, we will only single out a class of integrand processes sufficiently large for our later applications. For example, in the case of a continuous local martingale as an integrator, we only restrict ourselves to continuous adapted integrand processes, a class of processes sufficiently large for most applications. This of course does not mean that integrands which are not continuous cannot be integrated with respect to a continuous martingale. It is usually the case that when dealing with a discontinuous integrand, we can quickly decide then and there whether the integral has a meaning. Since many excellent textbooks on stochastic integration are available (McKean [**8**], Ikeda and Watanabe [**6**], Chung and Williams [**3**], Oksendal [**10**], Karatzas and Shreve [**7**], to cite just a few), there is little motivation on the part of the author to go beyond what will be presented in this chapter.

1. Introduction

If A is a process of bounded variation and $f : \mathbb{R} \times \mathbb{R}$ is function such that $s \mapsto f(s, \omega)$ is Borel measurable function, then

$$\int_0^t f(s)\, dA_s$$

can be interpreted as a pathwise integral if proper integrability conditions are assumed and its value at $\omega \in \Omega$ is the usual Lebesques–Stieljes integral

$$\int_0^t f(s,\omega)\, dA_s(\omega).$$

In theoretical and applied problems, there is an obvious need to make sense out of an integral of the form

$$\int_0^t f(s)\, dB_s,$$

where B is a Brownian motion. As we know, Brownian motion sample paths are not functions of bounded variation (see REMARK 7.2). For this reason in general there is no easy and direct pathwise interpretation of the above integral. However, in some special situation, a simple interpretation is possible. For example, if $s \mapsto f(s,\omega)$ itself has bounded variation for each ω, we can define the above integral by integration by parts:

$$\int_0^t f(s)\, dB_s = f(t)B_t - \int_0^t B_s\, df(s).$$

Such stochastic integrals are rather limited in its scope of application. Itô's theory of stochastic integration greatly expands the class of integrand processes, thus making the theory into a powerful tool in pure and applied mathematics.

We first define the integration of a step (and deterministic) process with respect to a Brownian motion. Let

$$\Delta:\ 0 = t_0 < t_1 \cdots < t_n = t$$

be a partition of $[0,t]$. If f is a step process:

$$f(s) = f_{j-1}, \qquad t_{j-1} \le t < t_j,$$

then the natural definition of its stochastic integral with respect to Brownian motion B is

$$\int_0^t f(s)\, dB_s = \sum_{j=1}^{n-1} f_{j-1}\left[B_{t_j} - B_{t_{j-1}}\right].$$

Using the property of independent increments for Brownian motion, its second moment is given by

$$\begin{aligned}
\mathbb{E}\left[\int_0^t f(s)\, dB_s\right]^2 &= \sum_{i,j=1}^{n} f_{i1} f_{j-1} \mathbb{E}\left[\left(B_{t_i} - B_{t_{i-1}}\right)\left(B_{t_j} - B_{t_{j-1}}\right)\right] \\
&= \sum_{j=1}^{n} |f_{j-1}|^2 \mathbb{E}\left[\left(B_{t_j} - B_{t_{j-1}}\right)^2\right] \\
&= \sum_{j=1}^{n} |f_{j-1}|^2 \left(t_j - t_{j-1}\right).
\end{aligned}$$

This gives

$$\mathbb{E}\left|\int_0^1 f(s)\, dB_s\right|^2 = \sum_{j=1}^{n} |f_{j-1}|^2 (t_j - t_{j-1}) = \int_0^t |f(s)|^2\, ds. \tag{1.1}$$

If f and g are two step functions, then

$$\mathbb{E}\left|\int_0^t f(s)\, dB_s - \int_0^t g(s)\, dB_s\right|^2 = \int_0^t |f(s) - g(s)|^2\, ds = \|f - g\|_2^2.$$

For an arbitrary function $f : [0,t] \to \mathbb{R}$, it is now clear that as long as there is a sequence of step functions f_n such that $\|f_n - f\|_2 \to 0$, we can define the stochastic integral as the limit

$$\int_0^t f(s)\,dB_s = \lim_{n\to\infty}\int_0^t f_n(s)\,dB_s.$$

It is well know from real analysis that this approximation property is shared by all Borel measurable functions f such that $\|f\|_2 < \infty$. We have thus greatly enlarged the space of deterministic functions which can be integrated with respect to a Brownian motion.

The key observation in Itô's theory of stochastic integration is that we can pass to a random step process as long as $f_j \in \mathscr{F}_{t_j}$. Note that now the integrand has the form

$$f(s) = f_{j-1}, \quad t_{j-1} \le s < t_j, \quad f_{j-1} \in \mathscr{F}_{t_{j-1}},$$

Namely, in the time interval $[t_{j-1}, t_j]$, the integrand is measurable with respect to the σ-algebra at the *left* endpoint of the time interval. For this reason above argument still applies as long as we replace the equality (1.1) by the more general equality

$$\mathbb{E}\left|\int_0^1 f(s)\,dB_s\right|^2 = \mathbb{E}\sum_{j=1}^n |f_{j-1}|^2(t_j - t_{j-1}) = \mathbb{E}\int_0^t |f(s)|^2\,ds.$$

The key step in establishing this equality is the vanishing of the off diagonal term

$$\mathbb{E}\left[f_{i-1}f_{j-1}(B_{t_i} - B_{t_{i1}})(B_{t_j} - B_{t_{j-1}})\right] = 0.$$

This holds because if $i < j$, then by conditioning on $\mathscr{F}_{t_{j-1}}$, we have

$$\begin{aligned} & \mathbb{E}\left[f_{i-1}f_{j-1}(B_{t_i} - B_{t_{i1}})(B_{t_j} - B_{t_{j-1}})|\mathscr{F}_{t_{j-1}}\right] \\ = \ & f_{i-1}f_{j-1}(B_{t_i} - B_{t_{i1}})\mathbb{E}\left[B_{t_j} - B_{t_{j-1}}|\mathscr{F}_{t_{j-1}}\right] \\ = \ & 0. \end{aligned}$$

For the diagonal term, by conditioning on $\mathscr{F}tj - 1$ again, we have

$$\mathbb{E}\left[|f_{i-1}|^2(B_{t_i} - B_{t_{i-1}})^2\right] = \mathbb{E}|f_{i-1}|^2(t_i - t_{i-1}).$$

An important property of stochastic integrals is that

$$M_t = \int_0^t f(s)\,dB_s$$

is a continuous martingale. Let us see why this is so if the integrand process is a step process

$$f(s) = f_{j-1}, \qquad t_{j-1} < s \le t_j, \quad f_{j-1} \in L^2(\Omega, \mathscr{F}_{t_{j-1}}, \mathbb{P}).$$

We can write

$$M_t = \int_0^t f(s)\,dB_s = \sum_n f_{j-1}\left(Bt_j \wedge t - B_{t_{j-1}\wedge t}\right).$$

Note that sum has only finitely many terms for each fixed t. The process M has continuous sample paths because Brownian motion does. Suppose that $s < t$. If s and t are not among the points t_j we may simply insert them into the sequence. Now s and t are among the sequence we have

$$M_t - M_s = \sum_{s \le t_{j-1} \le t} f_{j-1} \left(B_{t_j} - B_{t_{j-1}} \right).$$

The general term in the sum vanishes after we condition it on $\mathscr{F}_{t_{j-1}}$, hence it also vanishes after we condition on $\mathscr{F}_s$ because $\mathscr{F}_s \subset \mathscr{F}_{t_{j-1}}$. It follows that

$$\mathbb{E}\left[M_t - M_s | \mathscr{F}_s \right] = 0$$

and M_t is a continuous martingale.

We now calculate the quadratic variation process for the continuous martingale

$$M_t = \int_0^t f(s)\, dB_s.$$

We claim that

$$\langle M, M \rangle_t = \int_0^t |f(s)|^2\, ds.$$

We again verify this for a step process f. As before, we assume that $s < t$ are among the points t_j. We have

$$M_t^2 - M_s^2 = \sum f_{i-1} f_{j-1} (B_{t_i} - B_{t_{i-1}})(B_{t_j} - B_{t_{j-1}}),$$

where the sum is over those i and j such that $t_i \le t$, $t_j \le t$ and at least one of t_{i-1} and t_{j-1} is greater or equal o s. For an off-diagonal term, say, $i < j$, the term vanishes after conditioning on $\mathscr{F}_{t_{j-1}}$, hence also vanishes after conditioning on $\mathscr{F}_s$. For a diagonal term $i = j$, since $t_{j-1} \ge s$, by first conditioning on $\mathscr{F}_{t_{j-1}}$ and then conditioning on $\mathscr{F}_s$ we have

$$\mathbb{E}\left[|f_{j-1}|^2 (B_{t_j} - B_{t_{j-1}})^2 | \mathscr{F}_s \right] = \mathbb{E}\left[|f_{j-1}|^2 (t_j - t_{j-1}) | \mathscr{F}_s \right].$$

It follows that

$$\mathbb{E}\left[M_t^2 - M_s^2 | \mathscr{F}_s \right] = \mathbb{E}\left[\int_s^t |f(s)|^2\, ds \middle| \mathscr{F}_s \right].$$

This shows that

$$M_t^2 - \int_0^t |f(s)|^2\, ds$$

is a martingale.

2. Stochastic integrals with respect to Brownian motion

The general setting is as follow. We have a Brownian motion B with respect to a filtration $\mathscr{F}_*$ defined on a filtered probability space $(\Omega, \mathscr{F}_*, \mathbb{P})$. We assume that the filtration $\mathscr{F}_*$ satisfies the usual conditions.

DEFINITION 2.1. *A real-valued function $f : \mathbb{R}_+ \times \Omega \to \mathbb{R}$ is called a step process if there exists a nondecreasing sequence $0 = t_0 < t_1 < t_2 < \cdots$ increasing to infinity and a sequence of square-integrable random variables $f_{j-1} \in \mathscr{F}_{t_{j-1}}$ such that*

$$f_t = f_{j-1}, \qquad t_{j-1} \le t < t_j.$$

The space of step processes is denoted by $\mathscr{S}$. The space of step processes on $[0,t]$ is denoted by $\mathscr{S}_t$.

Note that on an interval $[t_{j-1}, t_j)$ where f is constant, it is measurable with respect to the σ-algebra on the left endpoint of the interval, i.e., $f = f_{j-1} \in \mathscr{F}_{t_{j-1}}$.

Let f be a step process as above. The stochastic integral of f with respect to Brownian motion B is the process

$$I(f)_t = \int_0^t f(s)\, dB_s = \sum_{j=1}^{\infty} f_{j-1}\left(B_{t_j \wedge t} - B_{t_{j-1}\wedge t}\right).$$

Of course this is a finite sum for each fixed t. If $t_i \le t < t_{i+1}$, then

$$I(f)_t = \sum_{j=1}^{i} f_{j-1}(B_{t_j} - B_{t_{j-1}}) + f_i(B_t - B_{t_i}).$$

It is easy to see that the definition of $I(f)_t$ is independent of the partition $\{t_j\}$. As we have shown in the last section $I(f)$ is a continuous martingale with the quadratic variation

$$\langle I(f), I(f)\rangle_t = \int_0^t |f(s)|^2 ds.$$

In particular we have

$$\mathbb{E}|I(f)_t|^2 = \mathbb{E}\left[\int_0^t |f(s)|^2\, ds\right].$$

This relation allows us to extend the definition of stochastic integrals to more general integrands by a limit procedure. Since the quadratic variation process of a Brownian motion corresponds to the Lebesgue measure on $\mathbb{R}_+$, a very wide class of processes can be used as integrand processes.

DEFINITION 2.2. *A function $f : \mathbb{R}_+ \times \Omega \to \mathbb{R}$ is called progressively measurable with respect to the filtration $\mathscr{F}_*$ if for each fixed t the restriction $f : [0,t] \times \Omega \to \mathbb{R}$ is measurable with respect to the product σ-algebra $\mathscr{B}[0,t] \times \mathscr{F}_t$.*

EXAMPLE 2.3. Here are some examples of progressively measurable processes:

(1) A step process is progressively measurable.

(2) A continuous and adapted process is progressively measurable.

(3) Let τ be a finite stopping time. The the process

$$f(t,\omega) = I_{[0,\tau(\omega)]}(s)$$

is progressively measurable. For a discrete τ, this can be verified directly. For a general τ, let

$$\tau_n = ([2^n\tau] + 1)/2^n.$$

Then $\tau_n \downarrow \tau$ and $I_{[0,\tau_n(\omega)]}(s) \to I_{[0,\tau(\omega)]}(s)$ for all $(s,\omega) \in \mathbb{R}_+ \times \Omega$.

For a progressively measurable process f, we define for each fixed T,

$$\|f\|_{2,T}^2 = \mathbb{E}\left[\int_0^T |f(s)|^2 ds\right].$$

The progressive measurability assures that the integral on the right side has a meaning. We use $\mathscr{L}_T^2$ to denote the space of progressively measurable processes f on $[0,T] \times \Omega$ with $\|f\|_{2,T} < \infty$. The norm $\|\cdot\|_{2,T}$ makes $\mathscr{L}_T^2$ into a complete Hilbert space. We use $\mathscr{L}^2$ to denote the space of progressively measurable processes f such that $\|f\|_{2,T}$ is finite for all T. It can be made into a metric space by introducing the distance function

$$d(f,g) = \sum_{n=1}^{\infty} \frac{\|f-g\|_{2,n}}{1 + \|f-g\|_{2,n}}.$$

Note that $\mathscr{S}_T \subset \mathscr{L}^2$ and $\mathscr{S} \subset \mathscr{L}^2$, and there is an obvious restriction map $\mathscr{L}^2 \to \mathscr{L}_t^2$.

We now show how to define the stochastic integral

$$I(f)_t = \int_0^t f(s)\, dB_s$$

for every process in $\mathscr{L}^2$. We need the following approximation result.

LEMMA 2.4. *$\mathscr{S}_T$ is dense in $\mathscr{L}_T^2$. In other words, for any square integrable progressively measurable process $f \in \mathscr{L}_T^2$, there exists a sequence $\{f_n\}$ of step processes such that $\|f_n - f\|_{2,T} \to 0$.*

PROOF. First note that the set of uniformly bounded elements in $\mathscr{L}_T^2$ is dense. Suppose that f is a progressively measurable process uniformly bounded on $[0,T] \times \Omega$. Define

$$f_h(t,\omega) = \frac{1}{h}\int_{t-h}^t f(s,\omega)\, ds.$$

Each f_h is continuous and adapted. From real analysis we also know that

$$\lim_{h\to 0}\int_0^T |f_h(t,\omega) - f(t,\omega)|^2\, ds = 0.$$

It follows that $\|f_h - f\|_{2,T} \to 0$ as $h \to 0$. This shows that the set of uniformly bounded continuous and adapted processes is dense in $\mathscr{L}_T^2$.

Finally for a continuous, adapted, and uniformly bounded process f on $[0, T]$ we define

$$f_n(t) = f\left(\frac{j-1}{n}\right), \qquad \frac{j-1}{n} \le t < \frac{j}{n}.$$

Each $f_n \in \mathscr{S}_T$ and $f_n(t,\omega) \to f(t,\omega)$ for all $(t,\omega) \in [0,T] \times \Omega$ by continuity. Hence $\|f_n - f\|_{2,T} \to 0$ by the dominated convergence theorem. The lemma is proved. $\square$

We are now ready to define the stochastic integral of a progressively measurable and square integrable process with respect to a Brownian motion. Suppose that $f \in \mathscr{L}^2$. By restricting to $[0, t]$ we can regard it as an element in $\mathscr{L}_t^2$ for any t. Let f_n be a sequence of step processes on $[0, T]$ such that $\|f_n - f\|_{2,T} \to 0$. The stochastic integral $I(f_n)_t$ is well defined for $0 \le t \le T$ as a Riemannan sum and is a square integrable martingale. We have

$$I(f_m)_t - I(f_n)_t = \int_0^t [f_m(s) - f_n(s)]\, dB_s.$$

By Doob's martingale inequality,

$$\begin{aligned}
\mathbb{E}\left[\max_{0 \le t \le T} |I(f_m)_t - I(f_n)_t|^2\right] &\le 4\mathbb{E}|I(f_m - f_n)_T|^2 \\
&= 4\mathbb{E}\left[\int_0^T |f_m(s) - f_n(s)|^2 ds\right] \\
&= 4\|f_m - f_n\|_{2,T}^2.
\end{aligned}$$

This shows that $I(f_n)_t$ is a Cauchy sequence in $L^2(\Omega, \mathbb{P})$. By the completeness of $L^2(\Omega, \mathbb{P})$, the limit $I(f)_t = \lim_{n\to\infty} I(f_n)_t$ exists and we define

$$\int_0^t f(s)\, dB_s = \lim_{n\to\infty} \int_0^t f_n(s)\, dB_s.$$

Since $\|f_n - f\|_{2,T_1} \to 0$ implies $\|f_n - f\|_{2,T_2}$ for $T_2 \le T_1$, it is easy to see that $I(f)_t$ thus defined is independent of T and the choice of the approximating sequence $\{f_n\}$.

We have only defined $\{I(f)_t,\ t \ge 0\}$ as a collection of random variables. We can do more.

THEOREM 2.5. *Suppose that $f \in \mathscr{L}^2$ is a square integrable, progressively measurable process. Then the stochastic integral*

$$I(f)_t = \int_0^t f(s)\, dB_s$$

is a continuous martingale. Its quadratic variation process is

$$\langle I(f), I(f) \rangle_t = \int_0^t f(s)^2\, ds.$$

PROOF. The martingale property of $I(f)$ is inherited from the same property of $I(f_n)$ because this property is preserved when passing to the limit in $L^2(\Omega, \mathscr{F}*, \mathbb{P})$ as $n \to \infty$. We show that it has continuous sample paths. From

$$\mathbb{E}\left[\max_{0\le t\le T} |I(f_m)_t - I(f_n)_t|^2\right] \le 4\|f_m - f_n\|_{2,T}^2$$

we have by Chebyshev's inequality,

$$\mathbb{P}\left[\max_{0\le t\le T} |I(f_m)_t - I(f_n)_t| \ge \epsilon\right] \le \frac{4}{\epsilon^2}\|f_m - f_n\|_{2,t}^2.$$

Since $\|f_m - f_n\|_{2,T} \to 0$ as $m, n \to \infty$, by choosing a subsequence if necessary, we may assume that $\|f_n - f_{n+1}\|_{2,T} \le 1/n^3$, hence

$$\mathbb{P}\left\{\max_{0\le t\le T} |I(f_n)_t - I(f_{n+1})_t| \ge \frac{1}{n^2}\right\} \le \frac{4}{n^2}.$$

By the Borel–Cantelli lemma, there is a set Ω_0 with $\mathbb{P}\{\Omega_0\} = 1$ with the following property: for any $\omega \in \Omega_0$, there is an $n(\omega)$ such that

$$\max_{0\le t\le T} |I(f_n)_t(\omega) - I(f_{n+1})_t(\omega)| \le \frac{1}{n^2}$$

for all $n \ge n(\omega)$. It follows that $I(f_n)_t(\omega)$ converges uniformly on $[0, T]$. Since each $I(f_n)_t(\omega)$ is continuous in t, the limit, which necessarily coincide with $I(f)_t(\omega)$ with probability 1 for all t, must also be continuous. We therefore have shown that the stochastic integral $I(f)_t$ has a version with continuous sample paths.

We have shown before that for the step process f_n the process

$$I(f_n)_t^2 - \int_0^t f_n(s)^2\, ds$$

is a martingale. For each fixed t, as $n \to \infty$, the above random variable converges in $L^1(\Omega, \mathbb{P})$ to

$$I(f)_t^2 - \int_0^t f(s)^2\, ds.$$

Therefore the martingale property can be passed to the limiting process, which identifies the quadratic variation process of the stochastic integral as stated in the proposition. □

COROLLARY 2.6. *Suppose that $f, g \in \mathscr{L}^2$. Then*

$$\langle I(f), I(g)\rangle_t = \int_0^t f(s)g(s)\, ds.$$

PROOF. This follows from the equality

$$\langle M, N\rangle_t = \frac{\langle M+N, M+N\rangle_t - \langle M-N, M-N\rangle_t}{4}$$

for two continuous martingales M and N. □

If $Z \in \mathscr{F}_s$, we have

$$Z\int_s^t f_u\,dB_u = \int_s^t Zf_u\,dB_u.$$

This can be considered to be self-evident, even though it requires an approximation argument to verify. The same remark can be said of the identity

$$\int_s^t f_u\,dB_u = \int_0^t f_u\,dB_u - \int_0^s f_u\,dB_u.$$

Either we define the left side by the right side or we consider the left side as the stochastic integral with respect to $\{B_{s+u},\,u \geq 0\}$, which is a Brownian motion with respect to the filtration $\mathscr{F}_{s+*} = \{\mathscr{F}_{s+u},\,u \geq 0\}$, the end result is the same. However, if we replace s or t by a random time, the meaning of the two and other similar equalities may become ambiguous, especially when the random time is not a stopping time. In this respect, if we restrict ourselves to stopping times, the following results may help us clarify the situation.

PROPOSITION 2.7. *Suppose that $f \in \mathscr{L}_T^2$ and $\tau \leq T$ a stopping time.*
(1) We have

$$\int_0^\tau f_s\,dB_s = \int_0^T I_{[0,\tau]}(s)f_s\,dB_s.$$

(2) If $Z \in \mathscr{F}_\tau$ is a bounded bounded random variable, then

$$Z\int_\tau^T f\,dB_s = \int_0^T ZI_{[\tau,\infty)}(s)f_s\,dB_s.$$

PROOF. We will verify the equalities for discrete stopping times and step processes. In general we can approximate the stopping times σ and τ by discrete stopping times from above and the integrand process f by step processes in the $\|\cdot\|_{2,T}$ norm. The equalities are clearly preserved when passing to the limit under such approximations.

Let $0 = t_0 < t_1 < \cdots < t_n = T$ be a partition such that $f = f_{j-1} \in \mathscr{F}_{t_{j-1}}$ for $t_{j-1} \leq t < t_j$ and τ takes values only in the sequence $\{t_j\}$.

(1) Note that $I_{\{s<\tau\}}$ is a step process. We have

$$
\begin{aligned}
I(f)_\tau &= \sum_{i=1}^n I_{\{\tau=t_i\}} \int_0^{t_i} f_s\, dB_s \\
&= \sum_{i=1}^n I_{\left\{\tau=t_j\right\}} \sum_{j=1}^i f_{j-1}(B_{t_j} - B_{t_{j-1}}) \\
&= \sum_{j=1}^n f_{j-1}(B_{t_j} - B_{t_{j-1}}) \sum_{i=j}^n I_{\{\tau=t_i\}} \\
&= \sum_{j=1}^n I_{\left\{t_{j-1}<\tau\right\}} f_{j-1}(B_{t_j} - B_{t_{j-1}}) \\
&= \int_0^T I_{[0,\tau)}(s) f_s\, dB_s.
\end{aligned}
$$

(2) Note that $ZI_{\{\tau\le s\}}$ is a step process and belongs to $\mathscr{F}_s$ for fixed s. We have

$$
\begin{aligned}
Z \int_\sigma^T f_s\, dB_s Z \sum_{i=1}^n Z I_{\{\tau=t_i\}} \int_0^{t_i} f_s dB_s \\
&= \sum_{i=1}^n I_{\{\tau=t_i\}} \sum_{j=i+1}^n Z f_{j-1}(B_{t_j} - B_{t_{j-1}}) \\
&= \sum_{j=1}^n Z f_{j-1}(B_{t_j} - B_{t_{j-1}}) \sum_{i=1}^{j-1} I_{\{\tau=t_i\}} \\
&= \sum_{j=1}^n Z I_{\left\{\tau\le t_{j-1}\right\}} f_{j-1}(B_{t_j} - B_{t_{j-1}}) \\
&= \int_0^T Z I_{[\tau,\infty)}(s) f_s\, dB_s.
\end{aligned}
$$

□

EXAMPLE 2.8. Consider the stochastic integral

$$
\int_0^t B_s\, dB_s = \lim \sum_{i=1}^n B_{t_{j-1}}(B_{t_j} - B_{t_{j-1}}).
$$

Using the identity $2a(b-a) = b^2 - a^2 + (b-a)^2$ we have

$$
2 \sum_{i=1}^n B_{t_{j-1}}(B_{t_j} - B_{t_{j-1}}) = B_t^2 + \sum_{i=1}^n (B_{t_j} - B_{t_{j-1}})^2.
$$

The last sum converges to the quadratic variation $\langle B, B\rangle_t = t$. Hence we have

$$
2 \int_0^t B_s\, dB_s = B_t^2 - t,
$$

which is indeed a martingale.

EXAMPLE 2.9. One may wonder what will happen if in the approximating sum of a stochastic integral we do not take the value of the integrand at the left endpoint of a partition interval. For example, what is the sum

$$\sum_{j=1}^{n} B_{t_j}(B_{t_j} - B_{t_{j-1}})?$$

This is not hard to figure out. The difference between this sum and the usual sum taking the left endpoint value of the integrand is just the quadratic variation along the partition

$$\sum_{j=1}^{n}(B_{t_j} - B_{t_{j-1}})^2,$$

hence we ave

$$\lim_{|\Delta|\to 0}\sum_{j=1}^{n} B_{t_j}(B_{t_j} - B_{t_{j-1}}) = \int_0^t B_s\,dB_s + t.$$

This example shows where to take the value of the integrand on each partition interval really matters.

3. Extension to more general integrands

We have defined stochastic integral with respect to a Brownian motion for integrand process $f \in \mathscr{L}^2$. These are progressively measurable processes f such that

$$\mathbb{E}\left[\int_0^t f_s^2\,ds\right] < \infty$$

for all t. This integrability is too restrictive. Without too much effort we can define a much wider class of integrand processes.

DEFINITION 3.1. *We use $\mathscr{L}^2$ to denote the space of progressively measurable processes f such that*

$$\mathbb{P}\left[\int_0^t f_s^2\,ds < \infty\right] = 1$$

for all t.

This is a class of process wider than $\mathscr{L}^2$ and is sufficient for our applications. For example, all adapted continuous processes belong to $\mathscr{L}^2_{loc}$. We will extend the definition of stochastic integrals to processes in $\mathscr{L}^2_{loc}$. The resulting stochastic integral process is no longer a square integrable martingale but a local martingale.

Recall the definition of a local martingale.

DEFINITION 3.2. *We say that $M = \{M_t,\, t \geq 0\}$ is a local martingale if there exists a sequence of stopping times $\tau_n \uparrow \infty$ such that the stopped processes $M^{\tau} = \{M_{t\wedge\tau_n},\, t \geq 0\}$ are martingales.*

For continuous local martingale M with $M_0 = 0$, we can always take

$$\tau_n = \inf\{t \geq 0 : |M_t| \geq n\}.$$

Suppose that $f \in \mathscr{L}^2_{loc}$. Define the stopping time

$$\tau_n = \inf\left\{t > 0 : \int_0^t |f(s)|^2 ds \geq n\right\}.$$

the assumption that $f \in \mathscr{L}^2_{loc}$ implies that $\tau_n \uparrow \infty$ (with probability one, if one wants to be very precise). Now consider the process

$$f_n(s) = f(s) I_{\{s \leq \tau_n\}}.$$

It is clear that $f \in \mathscr{L}^2$, in fact $\|f_n\|_{2,T} \leq n$ for all T. The stochastic integral $I(f_n)$ is well-defined and is a square–integrable continuous martingale. We now define the stochastic integral $I(f)_t$ as follows:

$$I(f)_t = I(f_n)_t, \qquad t \leq \tau_n.$$

In order that this definition make sense, we have to verify the consistency: if $m \leq n$, then

$$I(f_n)_t = I(f_m)_t, \qquad t \leq \tau_m.$$

This is immediate. Indeed,

$$I(f_n)_{t\wedge\tau_m} = \int_0^t f_n I_{\{s\leq\tau_m\}}\, dB_s = \int_0^t f_m dB_s = I(f_m)_t.$$

From definition we have

$$I(f)_{t\wedge\tau_n} = I(f_n)_t,$$

which shows that $I(f)$ is a continuous local martingale.

EXAMPLE 3.3. The process $Be^{B^2} \notin \mathscr{L}^2$ because $\mathbb{E}\left[|B_t|^2 e^{B_t^2}\right]$ is infinite for $t \geq 1/2$. However, since it is continuous it belongs to $\mathscr{L}^2_{loc}$ and its stochastic integral with respect to Brownian motion is well defined. In fact, Itô's formula will show that

$$2\int_0^t B_s e^{B_s^2}\, dB_s = e^{B_t^2} - \int_0^t \left[1 + 2B_s^2\right] e^{B_s^2}\, ds.$$

4. Stochastic integration with respect to continuous local martingales

We have discussed stochastic integration with respect to Brownian motion so that only the martingale properties of Brownian motion is used. For this reason not much needs to be changed if we replace Brownian motion by a general continuous martingale. All we need to do is to replace the quadratic variation process of Brownian motion $\langle B, B\rangle_t = t$ by $\langle M.M\rangle_t$, which is a continuous increasing process. More specifically, the theory of stochastic integration with respect to Brownian motion is based on several properties of Brownian motion:

(1) B is a continuous square integrable martingale;
(2) The quadratic variation process is $\langle B, B\rangle_t = t$.

(3) If f is a step process, then

$$\mathbb{E}\left[\left(\int_0^t f\, dB_s\right)^2\right] = \mathbb{E}\left[\int_0^t |f_s|^2\, ds\right].$$

(4) The space of step processes $\mathscr{S}_T$ is dense in the space of progressively measurable and square integrable processes $\mathscr{L}^2_T$.

For a theory of stochastic integration with respect to a general continuous martingale M, we make the following observations:

(1) By the usual stopping time argument, we can restrict ourselves initially to uniformly bounded continuous martingales.

(2) The quadratic variation process $\langle M, M\rangle_t$ is continuous, adapted, and increasing. Therefore integration of a progressively measurable process with respect to $\langle M, M\rangle$ is well defined.

(3) For a bounded step process f we have

$$\mathbb{E}\left|\int_0^t f_s\, dM_s\right|^2 = \mathbb{E}\left[\int_0^t f_s^2\, d\langle M, M\rangle_s\right].$$

(4) We can introduce the norm

$$\|f\|_{2,T;M} = \sqrt{\mathbb{E}\left[\int_0^t f_s^2\, d\langle M, M\rangle_s\right]}.$$

There does not seem to be a good description of the closure of the space of $\mathscr{S}_T$ under this norm that is good for all M, but at least all progressively measurable and continuous processes with finite norm can be approximated by step processes.

It is also possible develop a theory of stochastic integration with respect to a general, not necessarily continuous martingale. Since we will mostly only deal with continuous martingales, we will restrict ourselves to continuous local martingales.

Suppose that M is a continuous local martingale and $f \in \mathscr{S}$ is a step process. The stochastic integral of f with respect to the martingale M is

$$I(f)_t = \int_0^t f\, dM_s = \sum_j f_{j-1}\left(M_{t_{j-1}\wedge t} - M_{t_j\wedge t}\right).$$

Just as in the case of Brownian motion we have

$$\mathbb{E}\left|\int_0^t f_s\, dM_s\right|^2 = \mathbb{E}\left[\int_0^t |f_s|^2 d\langle M, M\rangle\right].$$

Let $\mathscr{L}^2(M)$ be the space of progressively measurable processes f such that

$$\|f\|^2_{2,T;M} = \mathbb{E}\left[\int_0^T |f_s|^2 d\langle M, M\rangle_s\right] < \infty$$

for all $t \geq 0$. It can be shown that $\mathscr{S}_T$ is dense in $\mathscr{L}^2(M)_T$ (see Karatzas and Shreve **[7]**). Thus by the usual approximation argument, we see that the stochastic integral

$$I(f)_t = \int_0^t f_s \, dM_s$$

is well defined for all $f \in \mathscr{L}^2(M)$ and the quadratic variation process is

$$\langle I(f), I(f)\rangle_t = \int_0^t |f_s|^2 \, d\langle M, M\rangle_s.$$

Although it is not easy to find a good description of $\mathscr{L}^2(M)$ that fits all continuous martingales, it does contain all continuous adapted processes such that $\|f\|_{2,T;M}$ is finite for all T.

By the standard stopping time argument we have used in the case of Brownian motion, the stochastic integral

$$I(f)_t = \int_0^t f_s \, dM_s$$

can be defined for a continuous local martingale M and a progressively measurable process f such that

$$\mathbb{P}\left[\int_0^t |f_s|^2 \, d\langle M, M\rangle_s < \infty\right] = 1$$

for all t. Under these conditions, the stochastic integral $I(f)$ is a continuous local martingale.

EXAMPLE 4.1. Let M be a uniformly bounded continuous martingale. As in the case of Brownian motion we have

$$M_t^2 - M_0^2 = 2\sum_{i=1}^n M_{t_{j-1}}(M_{t_j} - M_{t_{j-1}}) + \sum_{i=1}^n (M_{t_j} - M_{t_{j-1}})^2.$$

The sum after the equal sign converges to the stochastic integral $\displaystyle\int_0^t M_s \, dM_s$. Therefore the second sum must also converge. It is clear that the limit of the second sum is an continuous, increasing, and adapted process. Hence by the Doob-Meyer decomposition theorem, the limit must be the quadratic variation process, i.e.,

$$\lim_{|\Delta|\to 0} \sum_{i=1}^n (M_{t_j} - M_{t_{j-1}})^2 = \langle M, M\rangle_t$$

and

$$M_t^2 = M_0^2 + 2\int_0^t M_s \, dM_s + \langle M, M\rangle_t.$$

The convergence of the quadratic variation process takes place in $L^2(\Omega, \mathbb{P})$. For a general local martingale, it can be shown by routine argument that the convergence takes place at least in probability.

5. Itô's formula

Itô's formula is the fundamental theorem for stochastic calculus. Let us recall that the fundamental theorem of calculus states that if F is a continuously differentiable function then

$$F(t) - F(0) = \int_0^t F'(s)\, ds.$$

Let $0 = t_0 < t_1 < \cdots < t_n = t$ be a partition of the interval $[a, b]$. We have

$$F(t) - F(0) = \sum_{i=1}^n \left[F(t_j) - F(t_{j-1})\right].$$

Since F' is continuous, by the mean value theorem, there is a point $\xi_i \in [t_{j-1}, t_j]$ such that

$$F(t_j) - F(t_{j-1}) = F(\xi_j)(t_j - t_{j-1}).$$

Hence, by the definition of Riemann integrals we have as $|\Delta| \to 0$,

$$F(b) - F(a) = \sum_{j=1}^n F'(\xi_j)(t_j - t_{j-1}) \to \int_0^t F'(s)\, ds.$$

What happens to this proof if we replace t by B_t. In this case ξ_j will be a point between $B_{t_{j-1}}$ and B_{t_j}. We have to deal with the sum

$$\sum_{i=1}^n F'(\xi_j)(B_{t_j} - B_{t_{j-1}}).$$

Here the argument has to depart from what we have done above. As EXAMPLE 2.9 shows, the place where we take the value of the integrand on each partition interval can change the limit of the Riemann sum. Therefore we should not expect that the above sum will converges to the stochastic integral $\int_0^t F'(B_s)\, dB_s$, for which the left endpoint value of the integrand is used in each partition interval. The method to remedy this situation is to take one more term of the Taylor expansion

$$F(B_{t_j}) - F(B_{t_{j-1}}) = F'(B_{t_{j-1}})(B_{t_j} - B_{t_{j-1}}) + \frac{1}{2}F''(\xi_j)(B_{t_j} - B_{t_{j-1}})^2.$$

The sum of the first term on the right side will converge to the stochastic integral as usual. The crucial observation is that because Brownian motion has finite quadratic variation, the sum of the second term on the right side

$$\sum_{i=1}^n F''(\xi_j)(B_{t_j} - B_{t_{j-1}})^2 \to \int_0^t F''(B_s)\, ds$$

as $|\Delta| \to 0$. The final result is the well know Itô's formula

$$F(B_t) = F(B_0) + \int_0^t F'(B_s)\, dB_s + \frac{1}{2}\int_0^t F''(B_s)\, ds.$$

On the right side, the first term is a continuous martingale, and the second term is a process of bounded variation. Thus if F is continuous and uniformly bounded together with its first and second derivatives, then the composition $F(B_t)$ is a semimartingale and Itô's formula gives an explicit expression for its Doob-Meyer decomposition.

The proof of Itô's formula for for a general semimartingale Z will not be a lot more difficult than the case of Brownian motion. For this reason, in this section we will prove it in this generality.

THEOREM 5.1. *Suppose that $F \in C^2(\mathbb{R})$ and Z is a semimartingale. Then*

$$F(Z_t) = F(Z_0) + \int_0^t F'(Z_s)\, dZ_s + \frac{1}{2}\int_0^t F''(Z_s)\, d\langle Z, Z\rangle_s. \tag{5.1}$$

Recall that a semimartingale has the form

$$Z_t = M_t + A_t,$$

where M is a local martingale and A a process of bounded variation. We first use a stopping time argument to reduce the proof to the case where Z_0, M, $\langle M, M\rangle$, and A are all uniformly bounded and F together with its first and second derivatives are uniformly bounded and uniformly continuous. First of all, let $\Omega_N = \{|Z_0| \le N\}$ and define

$$\mathbb{P}_N(C) = \frac{\mathbb{P}(C \cap \Omega_N)}{\mathbb{P}(\Omega_0)}.$$

It is clear that under the $\mathbb{P}_N$, the process Z is still a semimartingale with the same decomposition $Z = M + A$ and $\mathbb{P}_N\{|Z_0| \le N\} = 1$, i.e., Z_0 is bounded. Next, define the stopping time

$$\tau_N = \inf\{t : |M_t - M_0| + \langle M, M\rangle_t + |A|_t \ge N\}.$$

The stopped processes $Z^{\tau_N} = M^{\tau_N} + A^{\tau_N}$ and $\langle M^{\tau_N}, M^{\tau_N}\rangle_t = \langle M, M\rangle_t^{\tau_N}$ have the bounded properties we desired. Define a function $F_N \in C^2(\mathbb{R})$ such that it coincides with F on $[-2N, 2N]$ and vanishes outside $[-3N, 3N]$. On the probability space $(\Omega_N, \mathscr{F}_* \cap \Omega_N, \mathscr{P}_N)$ suppose that we have the formula

$$F_N(Z^{\tau_N}) = F_N(Z^{\tau_N}) + \int_0^t F_N'(Z_s^{\tau_N})\, dZ_s + \frac{1}{2}\int_0^t F_N''(Z_s^{\tau_N})\, d\langle Z^{\tau_N}, Z^{\tau_N}\rangle_s.$$

Replacing t there by $t \wedge \tau_N$ we see that

$$F(Z_{t\wedge\tau_N}) = F(Z_0) + \int_0^{t\wedge\tau_N} F'(Z_s)\, dZ_s + \frac{1}{2}\int_0^{t\wedge\tau_N} F''(Z_s)\, d\langle Z, Z\rangle_s.$$

Here we have used the definition of stochastic integral with respect to a local martingale. This shows that the formula (5.1) holds on Ω_N and $t < \tau_N$. Finally from $\mathbb{P}\{\Omega_N\} \uparrow 1$ and $\mathbb{P}\{\tau_N \uparrow \infty\} = 1$, we see that Itô's formula holds without any extra conditions.

After these reductions, we start the proof of Itô's formula itself. By Taylor's formula with remainder we have

$$F(Z_{t_j}) - F(Z_{t_{j-1}}) = F'(Z_{t_{j-1}})(Z_{t_j} - Z_{t_{j-1}}) + \frac{1}{2}F''(\xi_i)(Z_{t_j} - Z_{t_{j-1}})^2,$$

where ξ is a point between $Z_{t_{j-1}}$ and Z_{t_j}. For simplicity let us denote

$$\Delta Z_j = Z_{t_j} - Z_{t_{j-1}}, \quad \Delta M_j = M_{t_j} - M_{t_{j-1}}, \quad \Delta\langle M\rangle_j = \langle M\rangle_{t_j} - \langle M\rangle_{t_{j-1}}.$$

With these notations we can write

$$F(Z_t) - F(Z_0) = \sum_{i=1}^n F'(Z_{t_{j-1}})\Delta Z_j + \frac{1}{2}\sum_{i=1}^n F''(\xi_i)(\Delta Z_j)^2.$$

The first sum converges to the stochastic integral in $L^2(\Omega, \mathbb{P})$:

$$\sum_{i=1}^n F'(Z_{t_{j-1}})\Delta Z_j \to \int_0^t F'(Z_s)\,dZ_s.$$

The real work starts with the proof of

$$\sum_{i=1}^n F''(\xi_i)(\Delta Z_j)^2 \to \int_0^t F''(Z_s)\,d\langle Z, Z\rangle_s. \tag{5.2}$$

We will prove this through a series of three replacements

$$F''(\xi_i)(\Delta Z_j)^2 \Rightarrow F''(\xi_i)(\Delta M_j)^2 \Rightarrow F''(Z_{t_{j-1}})(\Delta M_j)^2 \Rightarrow F''(Z_{t_{j-1}})\Delta\langle M\rangle_j.$$

We will show that each replacement produces an error which will vanish in probability as the mesh of the partition $\Delta| \to 0$. After these replacements we will have

$$\sum_{i=1}^n F''(Z_{t_{j-1}})\Delta\langle M\rangle_j \to \int_0^t F''(Z_s)\,d\langle M, M\rangle_s,$$

which will complete the proof.

(1) From $(\Delta Z_j)^2 - (\Delta M_j)^2 = \Delta A_j(\Delta Z_j + \Delta M_j)$ we have

$$\left|\sum_{i=1}^n F''(\xi_i)\left\{(\Delta Z_j)^2 - (\Delta M_j)^2\right\}\right| \le \|F''\|_\infty |A|_t \max_{1\le i\le n} |\Delta Z_j + \Delta M_j|.$$

By sample path continuity we have

$$\lim_{|\Delta|\to 0}\max_{1\le i\le n} |\Delta Z_j + \Delta M_j| = 0$$

almost surely. Hence the error produced by the first replacement vanishes in probability as $|\Delta| \to 0$.

(2) We have

$$\left|\sum_{i=1}^n \left[F''(\xi_i) - F''(Z_{t_{j-1}})\right](\Delta M_j)^2\right| \le \max_{1\le i\le n} |F''(\xi_i) - F''(Z_{t_{j-1}})| \sum_{i=1}^n (\Delta M_j)^2.$$

Again by sample path continuity we have

$$\max_{1\le i\le n} |F''(\xi_i) - F''(Z_{t_{j-1}})| \to 0$$

almost surely. On the other hand, by EXAMPLE 4.1,

$$\lim_{|\Delta|\to 0} \sum_{i=1}^{n} (\Delta M_j)^2 = \langle M, M\rangle_t$$

in $L^2(\Omega, \mathbb{P})$. Therefore the error produced by this replacement vanishes in probability as $|\Delta| \to 0$.

(3) This is the most delicate replacement of the three. The error is

$$E_n = \sum_{i=1}^{n} F''(Z_{t_{j-1}}) \left\{ (\Delta M_j)^2 - \Delta\langle M\rangle_j \right\}.$$

The square $|E_n|^2$ is a sum of n^2 terms. From EXAMPLE 4.1

$$(\Delta M_j)^2 - \Delta\langle M\rangle_j = 2\int_{t_{j-1}}^{t_j} (M_s - M_{t_{j-1}})\, dM_s.$$

This is a stochastic integral with respect to a martingale, hence the conditional expectation of this expression with respect to $\mathscr{F}_{t_{j-1}}$ is clearly zero. For an off-diagonal term, say $i < j$, in $|E_n|^2$, its conditional expectation with respect to $\mathscr{F}_{t_{j-1}}$ is therefore zero. Hence all off-diagonal term has expectation value zero. For the diagonal term we have

$$\mathbb{E}\left[\left(\int_{t_{j-1}}^{t_j} (M_s - M_{t_{j-1}})\, dM_s\right)^2 \middle| \mathscr{F}_{t_{j-1}}\right] = \mathbb{E}\left[\int_{t_{j-1}}^{t_j} (M_s - M_{t_{j-1}})^2\, d\langle M, M\rangle_s\right].$$

Therefore the expected value of this off-diagonal term is bounded by

$$\|F''\|_\infty^2\, \mathbb{E}\left[\Delta\langle M\rangle_j \max_{t_{j-1}\le s\le t_j} |M_s - M_{t_{j-1}}|\right].$$

Adding the diagonal terms together we obtain

$$\mathbb{E}|E_n|^2 \le \|F''\|_\infty^2\, \mathbb{E}\left[\langle M\rangle_t \max_{1\le j\le n} \max_{t_{j-1}\le s\le t_j} |M_s - M_{t_{j-1}}|\right].$$

Note that under our assumption, the expression under the expectation is uniformly bounded. This expectation converges to zero by sample path continuity and the dominated convergence theorem.

With this final replacement, we have completed the proof of (5.2) and also the proof of Itô's formula.

We end this section with the statement of Itô's formula for multidimensional semimartingales. The basic ingredients of its proof have already explained in great detail in the proof of the one dimensional case. Since nothing will be learned from its tedious proof, we will justifiably omit it.

THEOREM 5.2. *Suppose that $Z = (Z^1, \dots, Z^n)$ be an $\mathbb{R}^n$ valued semimartingale. Then for any $F \in C^2(\mathbb{R}^n)$.*

$$F(Z_t) = F(Z_0) + \sum_{i=1}^{n} \int_0^t F_{x_i}(Z_s)\, dZ_s^i + \frac{1}{2} \sum_{i,j=1}^{n} \int_0^t F_{x_i x_j}(Z_s)\, d\langle Z^i, Z^j \rangle_s.$$

6. Differential notation and Stratonovich integrals

This section does not contain new results. We introduce differential notation for stochastic calculus and Stratonovich integrals, a form of stochastic integral more restrictive than Itô integrals.

We often work with the following classes of processes:

$\mathscr{M}$ = the space of continuous local martingales;
$\mathscr{A}$ = the space of continuous processes of bounded variation;
$\mathscr{I}$ = the space of increasing processes;
$\mathscr{Q}$= the space of semimartingales
$\mathscr{H}$ = the space of integrand processes.

Every process of bounded variation is the difference of two increasing processes: $\mathscr{A} = \mathscr{I} - \mathscr{I}$. A semimartingale is the sum of a local continuous martingale and a process of bounded variation: $\mathscr{Q} = \mathscr{M} + \mathscr{A}$

It is often convenient to write equalities among these processes in a differential form. We have defined the meaning of an expression such as $\displaystyle\int_0^t H_s dX_s$ for $H \in \mathscr{H}$ and $X \in \mathscr{Q}$. What exactly is the meaning a differential expression such as dX? We should consider dX as a symbol for the equivalence class of semimartingales Y such that

$$X_t - X_s = Y_t - Y_s$$

for all $s \le t$. The equivalent classes of semimartingale differentials is denoted by $d\mathscr{Q}$. We can carry out some purely symbolic calculations. To start with we can set

$$\int_s^t dX_u = X_t - X_s.$$

We can define a multiplication in $\mathscr{Q}$ by setting

$$dX \cdot dY = d\langle X, Y \rangle.$$

We also define the multiplication of an element from $\mathscr{Q}$ by an element from $\mathscr{H}$: HdX is the equivalence class containing the semimartingale $\displaystyle\int_0^t H_s\, dX_s^t$.

Here are some properties of these operations:

(1) $d\mathscr{Q} \cdot d\mathscr{Q} \subset d\mathscr{A}$;
(2) $d\mathscr{Q} \cdot d\mathscr{A} = 0$;
(3) $H(dX \cdot dY) = (H \cdot dX) \cdot dY$;
(4) $H_1(H_2 \cdot dX) = (H_1 H_2) dX$.

With these formal symbolic notations, the multi-dimensional Itô's formula can be written as

$$d\left\{f(Z_t)\right\} = \sum_{j=1}^{n} f_{x_i}(Z_t)\, dZ_t^i + \frac{1}{2}\sum_{i,j=1}^{n} f_{x_i x_j}(Z_t)\, dX^i \cdot dX^j.$$

If we are willing to adopt Einstein's summation convention, we may even drop the two summation signs.

Suppose that X and Y are two semimartingales. We define a new semimartingale differential by

$$X \circ dY = X \cdot dY + \frac{1}{2} dX \cdot dY.$$

This means that

$$\int_0^t X_s \circ dY_s = \int_0^t X_s\, dY_s + \frac{1}{2}\langle X, Y\rangle_t.$$

This is called the Stratonovich integral. Unlike Itô integral, the Stratonovich integral requires that the integrand process is also a continuous semimartingale, for in its definition we need the quadratic covariation of X and Y. In this sense Stratonovich integral is a weaker form of stochastic integral and its use is much more limited than Itô integrals. The main advantage of S-tratonovich integrals is that Itô's formula takes a form similar to that of the fundamental theorem of calculus.

PROPOSITION 6.1. *Let $F \in C^3(\mathbb{R}^n)$ and $Z = (Z^1, \ldots, Z^n)$ an n-dimensional semimartingale. Then*

$$d\left\{f(Z_t)\right\} = \sum_{i=1}^{n} F_{x_i}(Z_t) \circ dZ_t^i.$$

PROOF. The right side is equal to

$$F_{x_i}(Z_t) \cdot dZ_t^j + \frac{1}{2} F_{x_i x_j}(Z_t) \cdot dZ_t^j. \tag{6.1}$$

Using It&o's formula we have

$$dF_{x_i}(Z_t) = F_{x_i x_j}(Z_t) \cdot dZ_t^i + \frac{1}{2} F_{x_i x_j x_k}(Z_t) \cdot dZ^j \cdot dZ^k.$$

Thus the last term in (6.1) becomes

$$\frac{1}{2} F_{x_i x_j}(Z_t) \cdot dZ_t^i \cdot dZ_t^j + \frac{1}{4} F_{x_i x_j x_k}(Z_t) \cdot dZ_t^i \cdot dZ_t^j \cdot dZ_t^k.$$

The last triple sum is equal to zero because $dZ_t^i \cdot dZ_t^j \cdot dZ_t^k = 0$. Therefore the equality we wanted to prove is just Itô's formula. □

Finally we mention that Stratonovich integral can also be approximated by a Riemann sums. It is instructive to compare this Riemann sum with the one that approximates the corresponding Itô integral. The difference is subtle but crucial in making Itô integration such a successful theory.

THEOREM 6.2. *Let X and Y be continuous semimartingales. Then*

$$\int_0^t X_s \circ dY_s = \lim_{|\Delta|\to 0} \sum_{j=1}^n \frac{X_{t_{j-1}} + X_{t_j}}{2} \left[Y_{t_j} - Y_{t_{j-1}} \right],$$

where the convergence is in probability.

PROOF. Rewrite the summation as

$$\sum_{j=1}^n X_{t_{j-1}} \left[Y_{t_j} - Y_{t_{j-1}} \right] + \frac{1}{2} \left[X_{t_j} - X_{t_{j-1}} \right] \left[Y_{t_j} - Y_{t_{j-1}} \right].$$

The limit of the first term is the Itô integral $\int_0^t X_s dY_s$, that of the second term is exactly the quadratic covariation $\langle X, Y \rangle_t / 2$. $\square$

7. Third Assignment

EXERCISE 3.1. Let f be a (nonrandom) function of bound variation. Show that

$$\int_0^t f_s \, dB_s = f_t B_t - \int_0^t B_s \, df_s,$$

where the last integral is understood to be a Lebesgue-Stieljes integral.

EXERCISE 3.2. Let F be an entire function in the complex plane $\mathbb{C}$ and $Z = X + iY$ be the complex Brownian motion (meaning that X and Y are independent standard Brownian motion). We have

$$\int_0^t F'(Z_s) \, dZ_s = \int_0^t F'(Z_s) \circ dZ_s = F(Z_t) - F(Z_0).$$

EXERCISE 3.3. The price S_t of a stock with average rate of return μ and volatility σ is usually described by the stochastic differential equation

$$dS_t = S_t \left(\mu \, dt + \sigma \, dB_t \right).$$

Using Itô's formula to show that the solution of the above stochastic differential equation is

$$S_t = S_0 \exp \left[\sigma B_t + \left(\mu - \frac{\sigma^2}{2} \right) t \right].$$

EXERCISE 3.4. Let $B = (B^1, B^2, B^3)$ be a 3-dimensional Brownian motion which does not start from zero. Using Itô's formula to show that $1/|B_t|$ is a local martingale.

EXERCISE 3.5. Let B be an n-dimensional Brownian motion starting from zero and

$$X_t = |B_t| = \sqrt{\sum_{i=1}^n |B_t^i|^2}$$

be the radial process. Show that X satisfies the following Itô type stochastic differential equation:

$$X_t = W_t + \frac{n-1}{2}\int_0^t \frac{ds}{X_s},$$

where W is a one-dimensional Brownian motion.

EXERCISE 3.6. Let M and N be two continuous local martingales. Show that

$$|\langle M, N\rangle_t| \le \sqrt{\langle M\rangle_t \langle N\rangle_t}.$$

EXERCISE 3.7. Suppose that X, Y and Z be three semimartingales. Are the following relations true?
(1) $X(Y\,dZ) = (XY)\,dZ$;
(2) $X \circ (Y \circ Z) = (XY) \circ dZ$.

EXERCISE 3.8. A Brownian bridge is defined by

$$dX_t = dB_t - \frac{X_t\,dt}{1-t}, \qquad X_0 = 0.$$

Show that X is a Gaussian process with mean zero and

$$\mathbb{E}\left[X_s X_t\right] = \min\{s, t\} - st.$$

Show that $\{X_t,\ 0 \le t \le 1\}$ and the reversed process $\{X_{1-t},\ 0 \le t \le 1\}$ have the same law.

EXERCISE 3.9. Suppose that M and N are two bounded continuous martingales which are independent. Show that MN is a continuous martingale with respect to the filtration $\mathscr{F}_*^{M,N} = \sigma\{M_s,\ N_s;\ s \le t\}$ generated by M and N.

EXERCISE 3.10. Suppose that M is a strictly positive local martingale. Show that there is a local martingale N such that

$$M_t = M_0 \exp\left[N_t - \frac{1}{2}\langle N, N\rangle_t\right].$$

CHAPTER 4

Applications of Ito's Formula

In this chapter, we discuss several basic theorems in stochastic analysis. Their proofs are good examples of applications of Itô's formula.

1. Lévy's martingale characterization of Brownian motion

Recall that B is a Brownian motion with respect to a filtration $\mathscr{F}_*$ if the increment $B_t - B_s$ has the normal distribution $N(0, t-s)$ independent of $\mathscr{F}_s$. A consequence of this definition is that B is continuous martingale with quadratic variation process $\langle B, B\rangle_t = t$. The following theorem shows that this property characterizes Brownian motion completely.

THEOREM 1.1. *(Lévy's characterization of Brownian motion) Suppose that B is a continuous local martingale with respect to a filtration $\mathscr{F}_*$ whose quadratic variation process is $\langle B, B\rangle_t = t$ (or equivalently, the process $\left\{B_t^2 - t,\, t \geq 0\right\}$ is also a local martingale), then it is a Brownian motion with respect to $\mathscr{F}_*$.*

PROOF. We will establish the following: for any $0 \leq s < t$ and any bounded $\mathscr{F}_s$-measurable random variable Z,

$$\mathbb{E}\left[Z \exp\left\{ia(B_t - B_s)\right\}\right] = \exp\left[-\frac{|a|^2}{2}(t-s)\right]\mathbb{E}Z. \tag{1.1}$$

Letting $Z = 1$, we see that $B_t - B_s$ is distributed as $N(0, t-s)$. Letting $Z = e^{ibY}$ for $Y \in \mathscr{F}_s$, we infer by the uniqueness of two-dimensional characteristic functions that $B_t - B_s$ is independent of any $Y \in \mathscr{F}_s$, which shows that $B_t - B_s$ is independent of $\mathscr{F}_s$. Therefore it is sufficient to show (1.1)

Denote the left side of (1.1) by $F(t)$. We first use Itô's formula to obtain

$$\begin{aligned}\exp\left\{ia\cdot(B_t - B_s)\right\} =& 1 + ia\cdot\int_s^t \exp\left\{ia\cdot(B_u - B_s)\right\} dB_u \\ &- \frac{|a|^2}{2}\int_s^t \exp\left\{ia\cdot(B_u - B_s)\right\} du.\end{aligned}$$

The stochastic integral is uniformly bounded because the other terms in this equality are uniformly bounded. Therefore the stochastic integral is not only a local martingale, but a martingale. Multiply both sides by Z and take the expectation. Noting that $Z \in \mathscr{F}_s$, we have

$$F(t) = \mathbb{E}Z - \frac{|a|^2}{2}\int_s^t F(u)\, du.$$

Solving for $F(t)$, we obtain

$$F(t) = \exp\left[-\frac{|a|^2}{2}(t-s)\right]\mathbb{E}Z.$$

This is what we wanted to prove. □

A very useful corollary of Lévy's criterion is that every continuous local martingale M is a time change of Brownian motion. This can be seen as follows. We know that the quadratic variation process $\langle M, M\rangle$ is a continuous increasing process. Let us assume that as $t \to \infty$,

$$\langle M, M\rangle_t \to \infty$$

with probability 1. Let $\tau = \{\tau_t\}$ be the right inverse of $\langle M, M\rangle$ defined by

$$\tau_t = \inf\{s \geq 0 : \langle M, M\rangle_s > t\}.$$

It is easy to show that $t \mapsto \tau_t$ is a right-continuous, increasing process and $\langle M, M\rangle_{\tau_t} = t$ and $\tau_{\langle M,M\rangle_t} = t$. Furthermore, τ_t is a stopping time for each fixed t. Consider the time-changed process $B_t = M_{\tau_t}$. Let

$$\sigma_n = \inf\{t : |M_t| \geq n\}.$$

Since M is a continuous local martingale, we have $\sigma_n \uparrow \infty$ and each stopped process M^{σ_n} is a square integrable martingale. From $\langle M, M\rangle_t = t$, we have by Fatou's lemma

$$\mathbb{E}\left[B_t^2\right] \leq \liminf_{n\to\infty}\mathbb{E}\left[M^2_{\sigma_n\wedge\tau_t}\right] = \liminf_{n\to\infty}\mathbb{E}\left[\langle M, M\rangle_{\sigma_n\wedge\tau_t}\right] \leq \mathbb{E}\left[\langle M, M\rangle_{\tau_t}\right] = t.$$

By the optional sampling theorem, B is a martingale.

We now show that B is continuous. There exists a sequence of partitions $\{\Delta_n\}$ of the time set $\mathbb{R}_+$ such that $|\Delta_n| \to 0$ and that with probability 1,

$$\forall t \geq 0 : \sum_j \left[M_{t_j^n\wedge t} - M_{t_{j-1}^n\wedge t}\right]^2 = \langle M, M\rangle_t$$

for all $t \geq 0$. Now for any fixed t, the jump

$$(B_t - B_{t-})^2 = (M_{\tau_t} - M_{\tau_{t-}})^2$$

is bounded by the quadratic variation of M in the time interval $[\tau_{t-\delta}, \tau_t]$ for any $\delta > 0$. Hence

$$|B_t - B_{t-}|^2 \leq \lim_{n\to\infty}\sum_{t_j^n \geq \tau_{t-\delta}}\left[M_{t_j^n\wedge\tau_t} - M_{t_{j-1}^n\wedge\tau_t}\right]^2 = \langle M, M\rangle_{\tau_t} - \langle M, M\rangle_{\tau_{t-\delta}} = \delta.$$

Since δ is arbitrary, we see that $B_t = B_{t-}$. This shows that B is continuous with probability 1.

So far we have proved that B is a continuous local martingale. We compute the quadratic variation of B. The process $M^2 - \langle M, M\rangle$ is a continuous local martingale. By the optional sampling theorem, we see that

$$M^2_{\tau_t} - \langle M, M\rangle_{\tau_t} = B_t^2 - t$$

is a continuous local martingale. It follows that the quadratic variation process of B is just $\langle B, B\rangle_t = t$. Now we can use Lévy's characterization to conclude that $B_t = M_{\tau_t}$ is a Brownian motion. Therefore we have shown that every continuous local martingale can be transformed into a Brownian motion by a time change.

From $B_t = M_{\tau_t}$ and $\tau_{\langle M,M\rangle_t} = t$ we have $M_t = B_{\langle M,M\rangle_t}$. In this sense we say that every continuous local martingale is the time change of a Brownian motion.

2. Exponential martingale

We want to find an analog of the exponential function e^x. The defining property of the exponential function is the differential equation

$$\frac{df(x)}{dx} = f(x), \quad f(0) = 1,$$

or equivalently

$$f(x) = 1 + \int_0^x f(t)\, dt.$$

So we define an*exponential martingale* E_t by the stochastic integral equation

$$E_t = 1 + \int_0^t E_s\, dM_s,$$

where M is a continuous local martingale. This equation can be solved explicitly. Instead of writing down the formula and verify it, let us discover the formula. Since $E_0 = 1$ we can take the logarithm of E_t at least for small time t. Let $C_t = \log E_t$. By Itô's formula we have

$$C_t = \int_0^t E_s^{-1}\, dE_s - \frac{1}{2}\int_0^t E_s^{-2}\, d\langle E, E\rangle_s.$$

Since $dE_s = E_s dM_s$, we have $d\langle E, E\rangle_s = E_s^2 d\langle M, M\rangle_s$. Hence

$$C_t = M_t - \frac{1}{2}\langle M, M\rangle_t.$$

Therefore the formula for exponential martingale is

$$E_t = \exp\left[M_t - \frac{1}{2}\langle M, M\rangle_t\right].$$

Now it is easy to verify directly that this process satisfies the defining equation for E_t.

Every strictly positive continuous local martingale can be written in the form of an exponential martingale:

$$E_t = \exp\left[M_t - \frac{1}{2}\langle M, M\rangle_t\right],$$

where the local martingale M can be expressed in terms of E by

$$M_t = \int_0^t E_s^{-1}\, dE_s.$$

We can say that an exponential martingale is nothing but a positive continuous local martingale.

Exponential martingale is related to iterated stochastic integrals defined by $I_0(t) = 1$ and

$$I_n(t) = \int_0^t I_{n-1}(s)\, dM_s.$$

By iterating the defining equation of the exponential martingale we have the expansion

$$E_t = \exp\left[M_t - \frac{1}{2}\langle M, M\rangle_t\right] = \sum_{n=0}^{\infty} I_n(t)$$

This formula can be verified rigorously, namely, it can be shown that the infinite series converges and the remainder from the iteration tends to zero. To continue our discussion let us introduce a parameter λ by replacing M with λM and obtain

$$\exp\left[\lambda M_t - \frac{\lambda^2}{2}\langle M, M\rangle_t\right] = \sum_{n=0}^{\infty} \lambda^n I_n(t). \tag{2.1}$$

If we set

$$x = \frac{M_t}{\sqrt{2\langle M, M\rangle_t}} \quad \text{and} \quad \theta = \lambda\sqrt{\frac{\langle M, M\rangle_t}{2}},$$

then the left side of (2.1) becomes $\exp\left[2x\theta - \theta^2\right]$. The coefficients of its Taylor expansion in θ are called Hermite polynomials

$$e^{2x\theta - \theta^2} = \sum_{n=0}^{\infty} \frac{H_n(x)}{n!}\, \theta^n.$$

It can be shown that

$$H_n(x) = e^{-x^2/2}\left(\frac{d}{dx}\right)^n e^{-x^2/2}.$$

and $\{H_n, n \geq 0\}$ is the orthogonal basis of $L^2(\mathbb{R}, e^{-x^2/2}dx/\sqrt{2\pi})$ obtained by the Gramm-Schmidt procedure from the complete system $\{x^n, n \geq 0\}$.

Now we can write

$$\exp\left[\lambda M_t - \frac{\lambda^2}{2}\langle M, M\rangle_t\right] = \sum_{n=0}^{\infty} \frac{\lambda^n}{n!}\left(\frac{\langle M, M\rangle_t}{2}\right)^{n/2} H_n\left(\frac{M_t}{\sqrt{2\langle M, M\rangle_t}}\right).$$

It follows that the iterated stochastic integrals can be expressed in terms of M_t and $\langle M, M\rangle_t$ via Hermite polynomials as follows:

$$I_n(t) = \frac{1}{n!}\left(\frac{\langle M, M\rangle_t}{2}\right)^{n/2} H_n\left(\frac{M_t}{\sqrt{2\langle M, M\rangle_t}}\right).$$

Here are the first few iterated stochastic integrals:

$$\begin{aligned}
I_0(t) &= 1, \\
I_1(t) &= M_t, \\
I_2(t) &= \frac{1}{2}\left[M_t^2 - \langle M\rangle_t\right], \\
I_3(t) &= \frac{1}{6}\left[M_t^3 - \langle M\rangle_t M_t\right].
\end{aligned}$$

Further discussion on this topic can be found in Hida [**5**].

3. Uniformly integrable exponential martingales

In stochastic analysis exponential martingales (positive martingales) often appear in the following context. Suppose that $(\Omega, \mathscr{F}_*, \mathbb{P})$ is a filtered probability space. Let $\mathbb{Q}$ be another probability measure which is absolutely continuous with respect to $\mathbb{P}$ on the σ-algebra $\mathscr{F}_T$. Then $\mathbb{Q}$ is also absolutely continuous with respect to $\mathbb{P}$ on $\mathscr{F}_t$ for all $t \leq T$. Let E_t be the Radon-Nikodym derivative of $\mathbb{Q}$ with respect to $\mathbb{P}$ on $\mathscr{F}_t$:

$$E_t = \left.\frac{d\mathbb{Q}}{d\mathbb{P}}\right|_{\mathscr{F}_t}.$$

Then it is easy to verify that

$$E_t = \mathbb{E}\left\{E_T | \mathscr{F}_t\right\}.$$

This shows that $\{E_t,\, t \leq T\}$ is a positive martingale. If it is continuous, then it is an exponential martingale. Not only that, it is a uniformly integrable martingale. Conversely, if $\{E_t,\, t \leq T\}$ is a uniformly integrable exponential martingale, then it defines a change of measure on the filtered probability space $(\Omega, \mathscr{F}_*, \mathbb{P})$ up to time T.

It often happens that we know the local martingale M and wish that the exponential local martingale

$$E_t = \exp\left[M_t - \frac{1}{2}\langle M, M\rangle_t\right]$$

defines a change of probability measures by $d\mathbb{Q}/d\mathbb{P} = E_T$. For this it is necessary that $\mathbb{E}E_T = 1$. However, in general, we only know that $\{E_t\}$ is a positive local martingale. It is therefore a supermartingale and $\mathbb{E}E_T \leq 1$. Therefore the requirement $\mathbb{E}E_T = 1$ is not automatic and has to be proved by imposing further conditions on the local martingale M.

PROPOSITION 3.1. *Let M be a continuous local martingale and*

$$E_t = \exp\left[M_t - \frac{1}{2}\langle M, M\rangle_t\right].$$

Then $\mathbb{E}E_T \leq 1$ and $\{E_t,\, 0 \leq t \leq T\}$ is a uniformly integrable martingale if and only if $\mathbb{E}E_T = 1$.

PROOF. We know that E is a local martingale, hence there is a sequence of stopping times $\tau_n \uparrow \infty$ such that $E_{t\wedge\tau_n}$ is a martingale, hence

$$\mathbb{E}E_{T\wedge\tau_n} = \mathbb{E}E_0 = 1.$$

Letting $n \to \infty$ and using Fatou's lemma we have $\mathbb{E}E_T \le 1$. This inequality also follows from the fact that a nonnegative local martingale is always a supermartingale.

Suppose that $\mathbb{E}E_T = 1$. Since E is a supermartingale we have $E_t \ge \mathbb{E}\left\{E_T \mathscr{F}_t\right\}$. Taking expected value gives

$$1 \ge \mathbb{E}E_t \ge \mathbb{E}E_T = 1.$$

This shows that the equality must hold throughout and we have $E_t = \mathbb{E}\left\{E_T | \mathscr{F}_t\right\}$, which means that $\left\{E_t, 0 \le t \le T\right\}$ is a uniformly integrable martingale. $\square$

We need to impose conditions on the local martingale M to ensure that the local exponential martingale E is uniformly integrable. We have the following result due to Kamazaki

THEOREM 3.2. *Suppose that M is a martingale and $\mathbb{E}e^{M_T/2}$ is finite. Then*

$$\left\{\exp\left[M_t - \frac{1}{2}\langle M\rangle_t\right],\, 0 \le t \le T\right\}$$

is uniformly integrable.

PROOF. This is a very interesting proof. Let

$$E(\lambda)_t = \exp\left[\lambda M_t - \frac{\lambda^2}{2}\langle M\rangle_t\right].$$

We need to show that $E(1)$ is uniformly integrable, but first we take a slight step back and prove that $E(\lambda)$ is uniformly integrable for all $0 < \lambda < 1$. We achieve this by proving that there is an $r > 1$ and C such that $\mathbb{E}E(\lambda)_\sigma^r \le C$ for all stopping time $\sigma \le T$. We have

$$E(\lambda)_\sigma^r = \exp\left[\left(\lambda r - \sqrt{\lambda^3 r}\right) M_\sigma\right] \exp\left[\sqrt{\lambda^3 r} M_\sigma - \frac{\lambda^2 r}{2}\langle M\rangle_\sigma\right].$$

Using Hölder's inequality with the exponents $1 - \lambda + \lambda = 1$ we see that $\mathbb{E}E(\lambda)_\sigma^r$ is bounded by

$$\left\{\mathbb{E}\exp\left[\left(\frac{\lambda r - \sqrt{\lambda^3 r}}{1-\lambda}\right) M_\sigma\right]\right\}^{1-\lambda} \left\{\mathbb{E}\exp\left[\sqrt{\lambda^2 r} M_\sigma - \frac{\lambda^2 r}{2}\langle M\rangle_\sigma\right]\right\}^{\lambda}.$$

The second expectation on the right side does not exceed 1 (see PROPOSITION 3.1). For the first factor we claim that σ can be replaced by T and the coefficient can be replaced by $1/2$ if $r > 1$ is sufficiently close to 1. When $r = 1$ the coefficient is

$$\frac{\lambda - \sqrt{\lambda^3}}{1-\lambda} = \frac{\lambda}{1+\sqrt{\lambda}} < \frac{1}{2}$$

because $\lambda < 1$. Hence the coefficient is still less than $1/2$ if $r > 1$ but sufficiently close to 1. On the other hand, we have assumed that M is a martingale, hence $M_\sigma = \mathbb{E}\left\{M_T|\mathscr{F}_\sigma\right\}$. By Jensen's inequalty we have

$$e^{M_\sigma/2} \le \mathbb{E}\left\{e^{M_T/2}|\mathscr{F}_\sigma\right\}.$$

It follows that

$$\mathbb{E}\exp\left[\left(\frac{\lambda r - \sqrt{\lambda^3 r}}{1-\lambda}\right)M_\sigma\right] \le \mathbb{E}e^{M_\sigma/2} \le \mathbb{E}e^{M_T/2}.$$

We therefore have shown that for any $\lambda < 1$, there is an $r > 1$ such that

$$\mathbb{E}E(\lambda)_\sigma^r \le \mathbb{E}e^{M_T/2}$$

for all stopping times $\sigma \le T$. This shows that $E(\lambda)$ is a uniformly integrable martingale, which implies that $\mathbb{E}E(\lambda)_T = 1$ for all $\lambda < 1$.

We now use the same trick again to show $\mathbb{E}E(1)_T = 1$. We have

$$E(\lambda)_T = \exp\left[\lambda^2 M_T - \frac{\lambda^2}{2}\langle M\rangle_T\right]\exp\left[(\lambda-\lambda^2)M_T\right].$$

Using Hölder's inequality with the exponents $\lambda^2 + 1 - \lambda^2 = 1$ we have

$$1 = \mathbb{E}E(\lambda)_T \le \left\{\mathbb{E}\exp\left[M_T - \frac{1}{2}\langle M\rangle_T\right]\right\}^{\lambda^2}\left\{\mathbb{E}\exp\left(\frac{\lambda}{1+\lambda}M_T\right)\right\}^{1-\lambda^2}.$$

Because $\lambda/(1+\lambda) \le 1/2$, the second expectation on the right side can be replaced by $\mathbb{E}\exp[M_T/2]$. Letting $\lambda \downarrow 0$ we obtain

$$\mathbb{E}\exp\left[M_T - \frac{1}{2}\langle M\rangle_T\right] = 1.$$

□

The condition $\mathbb{E}\exp[M_T/2] < \infty$ is not easy to verify because we usually know the quadratic variation $\langle M\rangle_T$ much better than M_T itself. The following weaker criterion can often be used directly.

COROLLARY 3.3. *(Novikov's criterion) Suppose that M is a martingale. If $\mathbb{E}exp\left[\langle M\rangle_T/2\right]$ is finite, then*

$$\mathbb{E}\exp\left[M_T - \frac{1}{2}\langle M\rangle_T\right] = 1.$$

PROOF. We have

$$\exp\left[\frac{1}{2}M_T\right] = \exp\left[\frac{1}{2}M_T - \frac{1}{4}\langle M\rangle_T\right]\exp\left[\frac{1}{4}\langle M\rangle_T\right].$$

By the Cauchy-Schwarz inequality we have

$$\mathbb{E}\exp\left[\frac{1}{2}M_T\right] \le \sqrt{\mathbb{E}\exp\left[M_T - \frac{1}{2}\langle M\rangle_T\right]}\sqrt{\mathbb{E}\exp\left[\frac{1}{2}\langle M\rangle_T\right]}.$$

The factor factor on the right side does not exceed 1. Therefore

$$\mathbb{E}\exp\left[\frac{1}{2}M_T\right] \le \sqrt{\mathbb{E}\exp\left[\frac{1}{2}\langle M\rangle_T\right]}.$$

Therefore Novikov's condition implies Kamazaki's condition. □

4. Girsanov and Cameron-Martin-Maruyama theorems

In stochastic analysis, we often need to change the base probability measure from a given $\mathbb{P}$ to another measure $\mathbb{Q}$. Suppose that B is a Brownian motion under $\mathbb{P}$. In general it will no longer be Brownian motion under $\mathbb{Q}$. The Girsanov theorem describes the decomposition of B as the same of a martingale (necessarily a Brownian motion) and a process of bounded variation under a class of change of measures.

We assume that B is an $\mathscr{F}_*$-Brownian motion and V a progressively measurable process such that

$$\exp\left[\int_0^t V_s\, dB_s - \frac{1}{2}\int_0^t |V_s|^2\, ds\right],\ 0 \le t \le T$$

is a uniformly integrable martingale. This exponential martingale therefore defines a change of measure on $\mathscr{F}_T$ by

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \exp\left[\int_0^T V_s\, dB_s - \frac{1}{2}\int_0^T |V_s|^2\, ds\right],\ 0 \le t \le T.$$

This is the class of changes of measures we will consider.

THEOREM 4.1. *Suppose that $\mathbb{Q}$ is the new measure described above. Consider the Brownian motion with a drift*

$$X_t = B_t - \int_0^t V_\tau d\tau, \quad 0 \le t \le T. \tag{4.1}$$

Then X is Brownian motion under the probability measure $\mathbb{Q}$.

PROOF. Let $\{e_s\}$ be the exponential martingale

$$e_s = \exp\left\{\int_0^s \langle V_\tau, dB_\tau\rangle - \frac{1}{2}\int_0^s |V_\tau|^2 d\tau\right\}.$$

Then the density function of $\mathbb{Q}$ with respect to $\mathbb{P}$ on $\mathscr{F}_s$ is just e_s. From this fact it is easy to verify that if Y is an adapted process then

$$\mathbb{E}^{\mathbb{Q}}\left\{Y_t|\mathscr{F}_s\right\} = e_s^{-1}\mathbb{E}^{\mathbb{P}}\left\{Y_t e_t|\mathcal{F}_s\right\}.$$

This means that Y is a (local) martingale under $\mathbb{Q}$ if and only if $eY = \{e_s Y_s,\ s \ge 0\}$ is a (local) martingale under $\mathbb{Q}$.

Now e is a martingale and $de_s = e_s V_s\, dB_s$. On the other hand, using Itô's formula we have

$$d(e_s X_s) = e_s\, dB_s + e_s X_s V_s\, dB_s.$$

This shows that eX is local martingale under $\mathbb{P}$, hence X is a local martingale under $\mathbb{Q}$. Since $\mathbb{Q}$ and $\mathbb{P}$ are mutually absolutely continuous, it should be clear that the quadratic variation process X under $\mathbb{P}$ and under $\mathbb{Q}$ should be the same, i.e., $\langle X\rangle_t = t$. We can also verify this fact directly by applying Itô's formula to $Z_s = e_s(X_s^2 - s)$. We have

$$dZ_s = e_s\, d(X_s^2 - s) + (X_s^2 - s)\, de_s + 2X_s d\langle e, X\rangle_s.$$

The second term on the right side is a martingale. We have

$$d(X_s^2 - s) = 2X_s\, dX_s - d\langle X, X\rangle_s - ds = 2X_s\, dB_s - 2X_s V_s\, ds.$$

On the other hand, from $de_s = e_s V_s\, dB_s$ we have

$$d\langle e, X\rangle_s = e_s V_s\, ds.$$

It follows that

$$dZ_s = 2e_s X_s\, dB_s + (X_s^2 - s)\, de_s,$$

which shows that Z is a local martingale. Now we have shown that both X_s and $X_s^2 - s$ are local martingales under $\mathbb{Q}$. By Lévy's criterion we conclude that X is a Brownian motion under $\mathbb{Q}$. □

The classical Cameron-Martin-Maruyama theorem is a special case of the Girsanov theorem but stated in a slightly different form. Let μ be the Wiener measure on the path space $W(\mathbb{R})$. Let $h \in W(\mathbb{R})$ and consider the shift in the path space $\xi_h w = w + h$. The shifted Wiener measure is $\mu^h = \mu \circ \xi_h^{-1}$. If X is the coordinate process on $W(\mathbb{R})$, then μ^h is just the law of $X + h$. We will prove the following dichotomy: either μ^h and μ are mutually absolutely continuous or they are mutually singular. In fact we have an explicit criterion for this dichotomy.

DEFINITION 4.2. *A path (function) $h \in W(\mathbb{R})$ is called Cameron-Martin path (function) if it is absolutely continuous and its derivative is square integrable. The Cameron-Martin norm is defined by*

$$|h|_{\mathscr{H}}^2 = \int_0^1 |\dot{h}_s|^2\, ds.$$

The space of Cameron-Martin paths is denoted by $\mathscr{H}$.

It is clear that $\mathscr{H}$ is a Hilbert space.

THEOREM 4.3. *(Cameron-Martin-Maruyama theorem) Suppose that $h \in \mathbb{H}$. Then the shifted Wiener measure μ^h and μ are mutually absolutely continuous and*

$$\frac{d\mu^h}{d\mu}(w) = \exp\left[\int_0^1 \dot{h}_s\, dw_s - \frac{1}{2}\int_0^1 |\dot{h}_s|^2\, ds\right].$$

PROOF. Denote the exponential martingale by $e_1(w)$. We need to show that for nonnegative measurable function F on $W(\mathbb{R})$,

$$\mathbb{E}^{\mu^h}(F) = \mathbb{E}^{\mu}(Fe_1).$$

Let X be the coordinate process on $W(\mathbb{R})$. Then the left side is simply $\mathbb{E}^{\mu}F(X+h)$. Introduce the measure ν by

$$\frac{d\nu}{d\mu} = \exp\left[-\int_0^1 \dot{h}_s\,dw_s - \frac{1}{2}\int_0^1 |\dot{h}_s|^2\,ds\right].$$

We have

$$\begin{aligned}\frac{d\mu}{d\nu} &= \exp\left[\int_0^1 \dot{h}_s\,dw_s + \frac{1}{2}\int_0^1 |\dot{h}_s|^2\,ds\right]\\ &= \exp\left[-\int_0^1 \dot{h}_s\,d(w_s+h) - \frac{1}{2}\int_0^1 |\dot{h}_s|^2\,ds\right].\end{aligned}$$

Hence we can write $d\mu/d\nu = e_1(X+h)$. It follows that

$$\mathbb{E}^{\mu^h}F = \mathbb{E}^{\mu}\left[F(X+h)\right] = \mathbb{E}^{\nu}\left[F(X+h)\frac{d\mu}{d\nu}\right] = \mathbb{E}^{\nu}\left[F(X+h)e_1(X+h)\right].$$

By Girsanov's theorem $X+h$ is a Brownian motion under ν. On the other hand, X is a Brownian motion under μ. Therefore on the right side of the above equality we can replace ν by μ and at the same time replace $X+h$ by X, hence

$$\mathbb{E}^{\mu^h}F = \mathbb{E}^{\mu}\left[F(X)e_1(X)\right] = \mathbb{E}^{\mu}(Fe_1).$$

□

THEOREM 4.4. *Let $h \in W(\mathbb{R})$. The shifted Wiener measure μ^h mutually absolutely continuous or mutually singular with respect to μ according as $h \in \mathscr{H}$ or $h \notin \mathscr{H}$.*

PROOF. We need to show that if $h \notin \mathscr{H}$, then μ^h is singular with respect to μ.

First we need to convert the condition $h \notin \mathscr{H}$ into a more convenient condition. A function $\dot{h}$ to be square integrable if and only if

$$\int_0^1 f_s\dot{h}_s\,ds \le C|f|_2$$

for some constant C. Therefore it is conceivable that if $h \notin \mathscr{H}$, then for any C, there is a step function f such that

$$|f|_2^2 = \sum_{i=1}^n |f_i|^2(s_i - s_{i-1}) = 1$$

and

$$\int_0^1 f_s\,dh_s = \sum_{i=1}^n f_i(h_{s_i} - h_{s_{i-1}}) \ge C.$$

This characterization of $h \notin \mathscr{H}$ can indeed be verified rigorously.

Second, the convenient characterization that μ^h is singular with respect to μ is the following: for any positive ϵ, there is a set A such that $\mu^h(A) \ge 1-\epsilon$ and $\mu(A) \le \epsilon$.

Consider the random variable

$$Z(w) = \int_0^1 f_s \, dw_s = \sum_{i=1}^n f_i(w_{s_i} - w_{s_{i-1}}).$$

It is a Gaussian random variable with mean zero and variance $|f|_2^1 = 1$. Therefore it has the standard Gaussian distribution $N(0,1)$. Let

$$A = \{Z \geq C/2\}.$$

We have $\mu(A) \leq \epsilon$ for sufficiently large C. On the other hand,

$$\mu^h(A) = \mu\left\{w \in W(\mathbb{R}) : Z(w+h) \geq \frac{C}{2}\right\}.$$

By the definition of $Z(w)$ we have

$$Z(w+h) = Z(w) + \sum_{i=1}^n f_i(h_{s_i} - h_{s_{i-1}}) \geq Z(w) + C.$$

Therefore

$$\mu^h(A) = \mu\left\{Z + C \geq \frac{C}{2}\right\} = \mu\left\{Z \geq -\frac{C}{2}\right\}$$

and $\mu^h(A) \geq 1 - \epsilon$ for sufficiently large C. Thus we have shown that μ^h and μ are mutually singular. $\square$

5. Moment inequalities for martingales

Let M be a continuous martingale and

$$M_t^* = \max_{0 \leq s \leq t} |M_s|.$$

The moment $\mathbb{E}\left[M_t^{*p}\right]$ can be bounded both from above and from below by $\mathbb{E}\left[\langle M, M\rangle_t^{p/2}\right]$.

THEOREM 5.1. *Let M be a continuous local martingale. For any $p > 0$, there are positive constants c_p, C_p such that*

$$c_p \mathbb{E}\left[\langle M, M\rangle^{p/2}\right] \leq \mathbb{E}\left[(M_t^*)^p\right] \leq C_p \mathbb{E}\left[\langle M, M\rangle_t^{p/2}\right].$$

PROOF. The case $p = 2$ is obvious. We only prove the case $p > 2$. The case $0 < p < 2$ is slightly more complicated, see Ikeda and Watanabe **[6]**.

By the usual stopping time argument we may assume without loss of generality that M is uniformly bounded, so there is no problem of integrability. We prove the upper bound first. We start with the Doob's submartingale inequality

$$\mathbb{E}\left[M_t^{*p}\right] \leq \left(\frac{p}{p-1}\right)^p \mathbb{E}\left[|M_t|^p\right]. \tag{5.1}$$

We use Itô's formula to $|M_t|^p$. Note that $x \mapsto |x|^p$ is twice continuously differentiable because $p > 2$. This gives

$$|M_t|^p = p\int_0^t |M_s|^{p-1}\mathrm{sgn}(M_s)\, dM_s + \frac{p(p-1)}{2}\int_0^t |M_s|^{p-2}\mathrm{sgn}(M_s)\, d\langle M\rangle_s.$$

Take expectation and using the obvious bound $|M_s| \le M_s^*$ we have

$$\mathbb{E}\left[|M_t|^p\right] \le \frac{p(p-1)}{2}\mathbb{E}\left[M_t^{*(p-2)}\langle M, M\rangle_t\right].$$

We use Hölder's inequality on the right side and (5.1) on the left side to obtain

$$\mathbb{E}\left[M_t^{*p}\right] \le C\left\{\mathbb{E}\left[M_t^{*p}\right]\right\}^{(p-2)/p}\left\{\mathbb{E}\left[\langle M, M\rangle_t^{p/2}\right]\right\}^{2/p},$$

where C is a constant depending on p. The upper bound follows immediately.

The lower bound is slightly trickier. Using Itô's formula we have

$$M_t\langle M, M\rangle_t^{(p-2)/4} = \int_0^t \langle M, M\rangle_s^{(p-2)/4} dM_s + \int_0^t M_s d\langle, M, M\rangle_s^{(p-2)/4}.$$

The first term on the right side is the term we are aiming at because its second moment is precisely the left side of the inequality we wanted to prove. This is the reason why choose the exponent $(p-2)/4$. We have

$$\left|\int_0^t \langle M, M\rangle_s^{(p-2)/4} dM_s\right| \le 2M_t^*\langle M, M\rangle_t^{(p-2)/4}.$$

Squaring the inequality and taking the expectation, we have after using Hölder's inequality,

$$\frac{2}{p}\mathbb{E}\left[\langle M, M\rangle_t^{p/2}\right] \le 4\left\{\mathbb{E}\left[M_t^{*p}\right]\right\}^{2/p}\left\{\mathbb{E}\left[\langle M, M\rangle_t^{p/2}\right]\right\}^{(p-2)/2}.$$

The lower bound follows immediately. □

6. Martingale representation theorem

Let B be a standard Brownian motion and $\mathscr{F}_t^B = \sigma\{B_s, s \le t\}$ be its associated filtration of σ–fields properly completed so that it satisfies the usual condition. Now suppose M is a square-integrable martingale with respect to this filtration. We will show it can always be represented as a stochastic integral with respect to the Brownian motion, namely, there exists a progressively measurable process H such that

$$M_t = \int_0^t H_s\, dB_s.$$

Thus every martingale with respect to the filtration generated by a Brownian motion is a stochastic integral with respect to this Brownian motion. This is a very important result in stochastic analysis. In this section we will give a proof of this result by an approach we believe is direct, short, and well motivated. It uses nothing more than Itô's formula in a very elementary way. In the next section we will discuss another approach to this useful theorem.

We make a few remarks before the proof. First of all, the martingale representation theorem is equivalent to the following representation theorem: if X is a square integrable random variable measurable with respect to $\mathscr{F}_T^B$, then it can be represented in the form

$$X = \mathbb{E}X + \int_0^T H_s\, dB_s,$$

for if we take $X = M_T$, then we have

$$M_T = \int_0^T H_s\, dB_s.$$

Since both sides are martingales, the equality must also holds if T is replaced by any $t \le T$.

Second, the representation is unique because if

$$\int_0^T H_s\, dB_s = \int_0^T G_s\, dB_s,$$

then from

$$\mathbb{E}\left|\int_0^T H_s\, dB_s - \int_0^T G_s\, dB_s\right|^2 = \mathbb{E}\int_0^T |H_s - G_s|^2\, ds$$

we have immediately $H = G$ on $[0, T] \times \Omega$ with respect to the measure $\mathbb{L} \times \mathbb{P}$, where $\mathbb{L}$ is the Lebesgue measure.

Third, if $X_n \to X$ in $L^2(\Omega, \mathscr{F}, \mathbb{P})$ and

$$X_n = \mathbb{E}X_n + \int_0^T H_s^n\, dB_s,$$

then $\mathbb{E}X_n \to \mathbb{E}X_n$ and

$$\int_0^T |H_s^m - H_s^n|^2\, ds = \mathbb{E}|X_m - X_n|^2 \to 0.$$

This shows that $\{H^n\}$ is a Cauchy sequence in $L^2([0, T] \times \Omega, \mathbb{L} \times \mathbb{P})$. It therefore converges to a process H and

$$X = \mathbb{E}X + \int_0^T H_s\, dB_s.$$

The point here is that it is enough to show the representation theorem for a dense subset of random variables.

Finally we observe a square integrable martingale with respect to the filtration generated by a Brownian motion is necessarily continuous because every stochastic integral is continuous with respect to its upper time limit.

We may assume without loss generality that $T = 1$. Our method starts with a simple case $X = f(B_1)$ for a smooth bounded function f. It is easy to prove the theorem in this case and find the explicit formula for the process H.

PROPOSITION 6.1. *For any bounded smooth function f,*

$$f(W_1) = \mathbb{E} f(W_1) + \int_0^1 \mathbb{E}\left\{f'(W_1)|\mathscr{F}_s\right\} dW_s.$$

PROOF. We have $f(W_1) = f(W_t + W_1 - W_t)$. We know that $W_1 - W_t$ has the normal distribution $N(0, 1-t)$ and is independent of $\mathscr{F}_t$. Hence,

$$\mathbb{E}\left\{f(W_1)|\mathscr{F}_t\right\} = \frac{1}{\sqrt{2\pi(1-t)}} \int_{\mathbb{R}^1} f(W_t + x) e^{-|x|^2/2(1-t)} dx. \tag{6.1}$$

Now we regard the right side as a function of B_t and t and apply Itô's formula. Since we know that it is a martingale, we only need to find out its martingale part, which is very easy: just differentiate with respect to B_t and integrate the derivative with respect to B_t. We have

$$f(W_1) = \mathbb{E} f(W_1) + \int_0^1 \left[\frac{1}{\sqrt{2\pi(1-t)}} \int_{\mathbb{R}^1} f'(W_t + x) e^{-|x|^2/2(1-t)} dx\right] dW_s.$$

The difference between the integrand and the right side of (6.1) is simply that f is replaced by its derivative f'. This shows that

$$f(W_1) = \mathbb{E} f(W_1) + \int_0^1 \mathbb{E}\left\{f'(W_1)\mathscr{F}_t\right\}\, dt.$$

□

The general case can be handled by an induction argument.

THEOREM 6.2. *Let $X \in L^2(\Omega, \mathscr{F}_1, \mathbb{P})$. Then there is a progressively measurable process H such that*

$$X = \mathbb{E}\, X + \int_0^1 H_s\, dW_s.$$

PROOF. It is enough to show the representation theorem for random variables of the form

$$X = f(W_{s_1}, W_{s_2} - W_{s_1}, \cdots, W_{s_n} - W_{s_1}),$$

where f is a bounded smooth function with bounded first derivatives because the random variables of this type form a dense subset of $L^2(\Omega, \mathscr{F}_1, \mathbb{P})$ (see the remarks at the beginning of this section). By the induction hypothesis applied to the Brownian motion $\{W_s - W_{s_1},\ 0 \le s \le 1 - s_1\}$, we have

$$f(x, W_{s_2} - W_{s_1}, \cdots, W_{s_n} - W_{s_1}) = h(x) + \int_{s_1}^1 H_s^x dW_s,$$

where

$$h(x) = \mathbb{E} f(x, W_{s_2} - W_{s_1}, \cdots, W_{s_n} - W_{s_1}),$$

and the dependence of H^x on x is smooth. Hence, replacing x by W_{s_1} we have

$$X = h(W_{s_1}) + \int_{s_1}^{1} H_s\, dW_s,$$

where $H_s = H_s^{W_{s_1}}$. We also have

$$h(W_{s_1}) = \mathbb{E}\, h(W_{s_1}) + \int_0^{s_1} H_s\, dW_s.$$

It is clear that $\mathbb{E}\, h(W_{s_1}) = \mathbb{E}\, X$, hence

$$X = \mathbb{E} X + \int_0^1 H_s\, dW_s.$$

$\square$

We now give a general formula for the integrand in the martingale representation theorem for a wide class of random variables. Recall the definition of the Cameron-Martin space

$$\mathscr{H} = \left\{ h \in C[0,1] : h(0) = 0,\ \dot{h} \in L^2[0,1] \right\}$$

and the Cameron-Martin norm

$$|h|_{\mathscr{H}} = \sqrt{\int_0^1 |\dot{h}_s|^2 ds}.$$

A function $F : C[0,1] \to \mathbb{R}$ is called a cylinder function if it has the form

$$F(w) = f(w_{s_1}, \ldots, w_{s_l}),$$

where $0 < s_1 < \cdots < s_l \leq 1$ and $f : \mathbb{R}^l \to \mathbb{R}$ is bounded smooth function. We also assume that f has bounded first derivatives and use the notation

$$F_{x_i}(w) = f_{x_i}(w_{s_1}, \ldots, w_{s_l}).$$

For a cylinder function F and $h \in C[0,1]$, the directional derivative along $h \in W(\mathbb{R})$ is defined by

$$D_h F(w) = \lim_{t \to 0} \frac{F(w + th) - F(w)}{t}.$$

If $h \in \mathscr{H}$, then it is easy to verify that

$$D_h F(w) = \langle DF(w), h \rangle_{\mathscr{H}},$$

where

$$DF(w)_s = \sum_{i=1}^{l} \min\{s, s_i\}\, F_{x_i}(w).$$

Note that the derivative

$$D_s F(w) = \frac{d\,\{DF(w)_s\}}{ds}$$

is given by

$$D_s F(w) = \sum_{i=1}^{l} F_{x_i}(w) I_{[0,s_i]}(s).$$

We have the following integration by parts formula.

THEOREM 6.3. *Let $H : \Omega \to \mathscr{H}$ be $\mathscr{F}_*$-adapted and $\mathbb{E} \exp\left\{|H|_{\mathscr{H}}^2/2\right\}$ is finite. Then the following integration by parts formula holds for a cylinder function F:*

$$\mathbb{E} D_H F = \mathbb{E}\langle DF, H\rangle = \mathbb{E}\left[F \int_0^1 \dot{H}_s\, dW_s\right].$$

PROOF. By Girsanov's theorem, under the probability $\mathbb{Q}$ defined by

$$\frac{d\,\mathbb{Q}}{d\,\mathbb{P}} = \exp\left[t \int_0^1 \dot{H}_s dW_s - \frac{t^2}{2}\int_0^1 |\dot{H}_s|^2 ds\right],$$

the process

$$X_s = W_s - tH_s, \qquad 0 \le s \le 1,$$

is a Brownian motion, hence $\mathbb{E}^{\mathbb{Q}} F(X) = \mathbb{E}\, F(W)$. This means that

$$\mathbb{E}^{\mathbb{Q}} F(X) = \mathbb{E}\left\{F(W - tH) \exp\left[t \int_0^1 \dot{H}_s\, dW_s - \frac{t^2}{2}\int_0^1 |\dot{H}_s|^2 ds\right]\right\}$$

is independent of t. Differentiating with respect to t and letting $t = 0$, we obtain the integration by parts formula. $\square$

The explicit martingale representation theorem is given by the following Clark-Ocone formula.

THEOREM 6.4. *Let F be a cylinder function on the path space $W(\mathbb{R})$. Then*

$$F(W) = \mathbb{E}\, F(W) + \int_0^1 \mathbb{E}\left\{D_s F(W) | \mathscr{F}_s\right\} dW_s.$$

PROOF. By the martingale representation theorem we have

$$F(W) = \mathbb{E}\, F(W) + \int_0^1 \dot{H}_s\, dW_s,$$

where $\dot{H}$ is $\mathscr{F}_*$-adapted $\dot{H}$. By definition,

$$\mathbb{E}\, D_G F = \mathbb{E}\langle DF, G\rangle_{\mathscr{H}}.$$

By the integration by parts formula, the left side is

$$\begin{aligned}\mathbb{E}\, D_G F &= \mathbb{E}\left[F \int_0^1 \dot{G}_s\, dW_s\right]\\ &= \mathbb{E}\left[\int_0^1 \dot{H}_s\, dW_s \int_0^1 \dot{G}_s\, dW_s\right]\\ &= \mathbb{E}\int_0^1 \dot{H}_s \dot{G}_s\, ds.\end{aligned}$$

The right side is

$$\mathbb{E}\langle DF, G\rangle_{\mathscr{H}} = \mathbb{E}\int \dot{G}_s D_s F\, ds = \mathbb{E}\int_0^1 \mathbb{E}\left\{D_s F|\mathscr{F}_s\right\} \dot{G}_s\, ds.$$

Hence for all $\mathscr{F}_*$-adapted process $\dot{G}$ we have

$$\mathbb{E}\int_0^1 \mathbb{E}\left\{D_s F|\mathscr{F}_s\right\} \dot{G}_s\, ds = \mathbb{E}\int_0^1 \dot{H}_s \dot{G}_s\, ds.$$

Since $\mathbb{E}\left\{D_s F|\mathscr{F}_s\right\} - \dot{H}_s$ is also adapted, we must have

$$\dot{H}_s = \mathbb{E}\left\{D_s F|\mathscr{F}_s\right\}.$$

□

7. Reflecting Brownian motion

Let B be a one-dimensional Brownian motion starting from zero. The process $X_t = |B_t|$ is called reflecting Brownian motion. We have $X_t = F(B_t)$, where $F(x) = |x|$. We want to apply Itô's formula to $F(B_t)$, but unfortunately F is not C^2. We approximate F by

$$F_\epsilon(x) = \frac{1}{\epsilon}\int_0^x du_1 \int_0^{u_1} I_{[-\epsilon,\epsilon]}(u_2)\, du_2.$$

The above function is still not C^2 because $F_\epsilon'' = I_{[-\epsilon,\epsilon]}/\epsilon$, which is not continuous, but it is clear that $F_\epsilon(x) \to |x|$ and $F_\epsilon'(x) \to \mathrm{sgn}(x)$ as $\epsilon \to 0$. Now let ϕ be a continuous function and define

$$F_\phi(x) = \int_0^x du_1 \int_0^{u_1} \phi(u_2)\, du_2.$$

Itô's formula can be applied to $F_\phi(B_t)$ and we obtain

$$F_\epsilon(B_t) = \int_0^t F_\phi'(B_s)\, dB_s + \frac{1}{2}\int_0^t \phi(B_s) ds.$$

Now for a fixed ϵ we let ϕ in the above formula to be the continuous function

$$\phi_n(x) = \begin{cases} 0, & \text{if } |x| \geq \epsilon + n^{-1}, \\ \epsilon^{-1}, & \text{if } |x| \leq \epsilon, \\ \text{linear}, & \text{in the two remaining intervals.} \end{cases}$$

Then $F_{\phi_n} \to F_\epsilon$, and $F_{\phi_n}' \to F'$. It follows that we have

$$F_\epsilon(B_t) = \int_0^t F'(B_s)\, dB_s + \frac{1}{2\epsilon}\int_0^t I_{[-\epsilon,\epsilon]}(B_s)\, ds.$$

In the last term,

$$\int_0^t I_{[-\epsilon,\epsilon]}(B_s)\, ds$$

is the amount of time Brownian paths spends in the interval $[-\epsilon, \epsilon]$ up to time t. Now let $\epsilon \to 0$ in the above identity for $F_\epsilon(B_t)$. The term on the left side converges to $|B_t|$ and the first term on the right side converges to the stochastic integral

$$\int_0^t \operatorname{sgn}(B_s)\, dB_s.$$

Hence the limit

$$L_t = \lim_{\epsilon \to 0} \frac{1}{\epsilon} \int_0^t I_{(-\epsilon,\epsilon)}(B_s)\, ds$$

must exist and we have

$$|B_t| = \int_0^t \operatorname{sgn}(B_s)\, dB_s + \frac{1}{2} L_t.$$

We see that L_t can be interpreted as the amount of time Brownian motion spends in the interval $(-\epsilon, \epsilon)$ properly normalized. It is called the local time of Brownian motion B at $x = 0$. Let

$$W_t = \int_0^t \operatorname{sgn}(B_s)\, dB_s.$$

Then W is a continuous martingale with quadratic variation process

$$\langle W, W \rangle_t = \int_0^t |\operatorname{sgn}(B_s)|^2 ds = t.$$

Note that Brownian motion spends zero amount of time at $x = 0$ because $\mathbb{E} I_{\{0\}}(B_s) = \mathbb{P}\{B_s = 0\} = 0$ and

$$\mathbb{E} \int_0^t I_{\{0\}}(B_s)\, ds = \int_0^t \mathbb{E} I_{\{0\}}(B_s)\, ds = 0.$$

We thus conclude that reflecting Brownian motion $|B_t|$ is submartingale with the decomposition

$$|B_t| = W_t + \frac{1}{2} L_t.$$

It is interesting to note that W can be expressed in terms of reflecting Brownian motion by

$$W_t = X_t - \frac{1}{2} L_t,$$

where

$$L_t = \lim_{\epsilon \to 0} \frac{1}{\epsilon} \int_0^t I_{[0,\epsilon]}(X_s)\, ds. \tag{7.1}$$

We now pose the question: Can X and L be expressed in terms of W? That the answer to this question is affirmative is the content of the so-called Skorokhod problem.

DEFINITION 7.1. *Given a continuous path $f : \mathbb{R}_+ \to \mathbb{R}$ such that $f(0) \geq 0$. A pair of functions (g, h) is the solution of the Skorokhod problem if*

(1) *$g(t) \geq 0$ for all $t \geq 0$;*
(2) *h is increasing from $h(0) = 0$ and increases only when $g = 0$;*
(3) *$g = f + h$.*

The main result is that the Skorokhod problem can be solved uniquely and explicitly.

THEOREM 7.2. *There exists a unique solution to the Skorokhod equation.*

PROOF. It is interesting that the solution can be written down explicitly:

$$h(t) = -\min_{0\le s\le t} f(s) \wedge 0, \qquad g(t) = f(t) - \min_{0\le s\le t} f(s)\wedge 0.$$

Let's assume that $f(0) = 0$ for simplicity. If $f(0) > 0$, then $h(t) = 0$ and $g(t) = f(t)$ before the first time f reaches 0 and after this time it is as if the path starts from 0. The explicit solution in this case is

$$h(t) = \min_{0\le s\le t} f(s), \qquad g(t) = f(t) - \min_{0\le s\le t} f(s).$$

It is clear that $g(t) \ge 0$ for all t and h increases starting from $h(0) = 0$. The equation $f = g + h$ is also obvious. We only need to show that h increases only when $g(t) = 0$. This means that as a Borel measure h only charges the zero set $\{t : g(t) = 0\}$. This requirement is often written as

$$h(t) = \int_0^t I_{\{0\}}(g(s))\, dh(s).$$

Equivalently, it is enough to show that for any t such that $g(t) > 0$ there is a neighborhood $(t-\delta, t+\delta)$ of t such that h is constant on there. This should be clear, for if $g(t) > 0$, then $f(t) > \min_{0\le s\le t} f(s)$, which means that the minimum must be achieved at a point $\xi \in [0, t)$ and $f(t) > f(\xi)$. By continuity a small change of t will not alter this situation, which means that $h = f(\xi)$ in a neighborhood of t. More precisely, from $g(t) = f(t) - \min_{0\le s\le t} f(s) > 0$ and the continuity of f, there is a positive δ such that

$$\min_{t-\delta\le s\le t+\delta} f(s) > \min_{0\le s\le t-\delta} f(s).$$

Hence

$$\min_{0\le s\le t+\delta} f(s) = \min\left\{\min_{0\le s\le t-\delta} f(s), \min_{t-\delta\le s\le t+\delta} f(s)\right\} = \min_{0\le s\le t-\delta} f(s).$$

This means that $h(t+\delta) = h(t-\delta)$, which means that h must constant on $(t-\delta, t+\delta)$ because h is increasing.

We now show that the solution to the Skorokhod problem is unique. Suppose that (g_1, h_1) and let $\xi = h - h_1$. It is continuous and of bounded variation, hence

$$\xi(t)^2 = 2\int_0^t \xi(s)\, d\xi(s).$$

On the other hand, $\xi(s) = g(s) - g_1(s)$, hence

$$\xi(t)^2 = 2\int_0^t \{g(s) - g_1(s)\}\, d\{h(s) - h_1(s)\}.$$

There are four terms on the right side: $g(s)\, dh(s) = g_1(s)\, dh_1(s) = 0$ because h increases on when $g = 0$ and h_1 increases only when $g_1 = 0$;

$g(s)\,dh_1(s) \geq 0$ and $g_1(s)\,dh(s) \geq 0$ because $g(s) \geq 0$ and $g_1(s) \geq 0$. Putting these observations together we have $\xi(t)^2 \leq 0$, which means that $\xi(t) = 0$. This proves the uniqueness. □

If we apply Skorokhod equation to Brownian motion by replacing f with Brownian paths, we obtain some interesting results. We have shown that

$$|B_t| = W_t + \frac{1}{2} L_t,$$

where W is a Brownian motion. From the solution of the Skorokhod problem we conclude that $|B|$ and L are determined by W:

$$|B_t| = W_t - \min_{0 \leq s \leq t} W_s, \quad \frac{1}{2} L_t = -\min_{0 \leq s \leq t} W_s. \tag{7.2}$$

THEOREM 7.3. *Let W be a Brownian motion. (1) The processes*

$$\left\{ \max_{0 \leq s \leq t} W_s - W_t,\ t \geq 0 \right\} \quad \textit{and} \quad \{|W_t|,\ t \geq 0\}$$

have the same law, i.e, that of a reflecting Brownian motion. (2) We have

$$\max_{0 \leq s \leq t} W_s = \lim_{\epsilon \to 0} \frac{1}{2\epsilon} \int_0^t I_{[0,\epsilon]}(\max_{0 \leq u \leq s} W_u - W_s)\, ds.$$

PROOF. The assertions follow immediately from (7.2) by replacing W with $-W$, which is also a Brownian motion; in the second assertion, we use the fact that L is the normalized occupation time of reflecting Brownian motion (7.1). □

REMARK 7.4. We have calculated the joint distribution of $\max_{0 \leq s \leq t} W_s$ and $|W_t$ for a fixed t and we know that $\max_{0 \leq s \leq t} W_s - W_s$ and $|W_t|$ have the same distribution for each fixed t. The above theorem claims much more: they have the same distribution as two stochastic processes.

8. Brownian bridge

If we condition a Brownian motion to return to a fixed point x at time $t = 1$ we obtain Brownian bridge from o to x with time horizon 1. Let

$$L_x(\mathbb{R}^n) = \{ w \in W_o(\mathbb{R}^n) : w_1 = x \}.$$

The law of a Brownian bridge from o to x in time 1 is a probability measure μ_x on $L_x(\mathbb{R}^n)$, which we will call the Wiener measure on $L_x(\mathbb{R}^n)$. Note that $L_x(\mathbb{R}^n)$ is a subspace of $W_o(\mathbb{R}^n)$, thus μ_x is also a measure on $W_o(\mathbb{R}^n)$. By definition, we can write intuitively

$$\mu_x(C) = \mu \{ C | w_1 = x \}.$$

Here μ is the Wiener measure on $W_o(\mathbb{R}^n)$. The meaning of this suggestive formula is as follows. If F is a nice function measurable with respect to $\mathscr{B}_s$ with $s < 1$ and f a measurable function on $\mathbb{R}^n$, then

$$\mathbb{E}^\mu \{ F f(X_1) \} = \mathbb{E}^\mu \{ \mathbb{E}^{\mu_{X_1}}(F) f(W_1) \},$$

where W denotes the coordinate process on $W_o(\mathbb{R}^n)$. Using the Markov property at time s we have

$$\mathbb{E}^{\mu}\left[F\int_{\mathbb{R}^n}p(1-s,W_s,y)f(y)\,dy\right]=\int_{\mathbb{R}^n}\mathbb{E}^{\mu_y}(F)p(1,o,y)f(y)dy,$$

where

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

is the transition density function of Brownian motion X. This being true for all measurable f, we have for all $F\in\mathscr{B}_s$,

$$\mathbb{E}^{\mu_x}F=\mathbb{E}^{\mu}\left[\frac{p(1-s,W_s,x)}{p(1,o,x)}F\right]. \tag{8.1}$$

Therefore μ_x is absolutely continuous with respect to μ on $\mathscr{F}_s$ for any $s<1$ and the Radon-Nikodym density is given by

$$\left.\frac{d\mu_x}{d\mu}\right|_{\mathscr{F}_s}(w)=\frac{p(1-s,w_s,x)}{p(1,o,x)}=e_s.$$

The process $\{e_s, 0\le s<1\}$ is a necessarily a positive (local) martingale under the probability μ. It therefore must have the form of an exponential martingale, which can be found explicitly by computing the differential of $\log e_s$. The density function $p(t,y,x)$ satisfies the heat equation

$$\frac{\partial p}{\partial t}=\frac{1}{2}\Delta_y p$$

in (t,y) for fixed x. This equation gives

$$\Delta_y\log p=\frac{\partial\log p}{\partial t}-|\nabla_y\log p|^2.$$

Using this fact and Iô's formula we find easily that

$$d\log e_s=\langle\nabla\log p(1-s,w_s,x),dw_s\rangle-\frac{1}{2}|\nabla p(1-s,w_s,x)|^2ds.$$

Hence e_s is an exponential martingale of the form

$$\left.\frac{d\mu_o}{d\mu}\right|_{\mathscr{B}_s}=\exp\left[\int_0^s\langle V_u,dw_u\rangle-\frac{1}{2}\int_0^s|V_u|^2du\right],$$

where

$$V_s=\nabla_y\log p(1-s,w_s,x).$$

By Girsanov's theorem, under probability μ_x, the process

$$B_s=W_s-\int_0^s\nabla\log p(1-\tau,W_\tau,x)d\tau,\qquad 0\le s<1$$

is a Brownian motion. The explicit formula for $p(t,y,x)$ gives

$$\nabla_y\log p(1-\tau,y,x)=-\frac{y-x}{1-\tau}.$$

Under the probability μ_x the coordinate process W is a reflecting Brownian motion from o to x in time 1. Therefore we have shown that reflecting Brownian motion is the solution to the following stochastic differential equation

$$dW_s = dB_s - \frac{W_s - x}{1-s}\, ds.$$

This simple equation can be solved explicitly. From the equation we have

$$d(W_s - x) + \frac{W_s - x}{1-s}\, ds = dB_s.$$

The left side after dividing by $1-s$ is the differential of $(W_s - x)/(1-s)$, hence

$$W_s = sx + (1-s)\int_0^s \frac{dB_u}{1-u}.$$

This formula shows that, like Brownian motion itself, Brownian bridge is a Gaussian procss.

The term "Brownian bridge" is often reserved specifically for Brownian bridge which returns to its starting point at the terminal time. In this case we have

$$dW_s = dB_s - \frac{W_s}{1-s}\, ds$$

and

$$W_s = (1-s)\int_0^s \frac{dB_u}{1-u}.$$

For dimension $n = 1$ it is easy verify that the covariance function is given by

$$\mathbb{E}\{W_s W_t\} = \min\{s,t\} - st.$$

Using this fact we obtain another representation of Brownian bridge.

THEOREM 8.1. *If $\{B_s\}$ is a Brownian motion, then the process*

$$W_s = B_s - sB_1$$

is a Brownian bridge.

PROOF. ($n = 1$) Verify directly that W defined above has the correct covariance function. □

The following heuristic discussion of $W_s = B_s - sB_1$ is perhaps more instructive. Let $\mathscr{F}_*$ be the filtration of the Brownian motion B. We enlarge the filtration to

$$\mathscr{G}_s = \sigma\{\mathscr{F}_s, B_1\}.$$

We compute the Doob-Meyer decomposition of W with respect to $\mathscr{G}_*$. Of course the martingale part is a Brownian motion because its quadratic variation process will be the same has that of W. Denote this Brownian motion by Ω. Doob's explicit decomposition formula for a semimartingale suggest that

$$W_s = \Omega_s + \int_0^s \mathbb{E}\{dW_s | \mathscr{G}_s\}.$$

We have $dW_s = dB_s - B_1 ds$. The differential $dW_s = W_{s+ds} - W_s$ is forward differential. We need to project dW_s to the L^2-space generated by

$$\mathscr{G}_s = \sigma\left\{B_u,\, u \leq s; B_1\right\} = \left\{B_u,\, u \leq s; B_1 - B_s\right\}.$$

Note that $\sigma\left\{B_u,\, u \leq s\right\}$ is orthogonal to $B_1 - B_s$. The differential dB_s is orthogonal to the first part and its projection to the second part is

$$\frac{dB_s \cdot (B_1 - B_s)}{1-s} \cdot (B_1 - B_s) = \frac{B_1 - B_s}{1-s}\, ds.$$

The differential $B_1 ds$ is of course in the target space already, hence

$$\mathbb{E}\left\{dW_s | \mathscr{G}_s\right\} = \frac{B_1 - B_s}{1-s}\, ds - B_1\, ds = -\frac{B_s - sB_1}{1-s}\, ds = -\frac{W_s\, ds}{1-s}.$$

It follows that

$$W_s = \Omega_s - \int_0^s \frac{W_\tau}{1-\tau} d\tau,$$

which is exactly the stochastic differential equation for a reflecting Brownian motion.

9. Fourth assignment

EXERCISE 4.1. Let ϕ be a strictly convex function. If both N and $\phi(N)$ are continuous local martingales then N is trivial, i.e., there is a constant C such that $N_t = 1$ with probability 1 for all t.

EXERCISE 4.2. Let M be a continuous local martingale. Show that there is a sequence of partitions $\Delta_1 \subset \Delta_2 \subset \Delta_3 \subset \cdots$ such that $|\Delta_n| \to 0$ and with probability 1 the following holds: for all $t \geq 0$,

$$\lim_{n\to\infty} \sum_{i=1}^{\infty} \left(M_{t_i^n \wedge t} - M_{t_{i-1}^n \wedge t}\right)^2 = \langle M, M\rangle_t.$$

EXERCISE 4.3. Let B be the standard Brownian motion. Then the reflecting Brownian motion $X_t = |B_t|$ is a Markov process. This means

$$\mathbb{P}\left\{X_{t+s} \in C | \mathscr{F}_s^X\right\} = \mathbb{P}\left\{X_{t+s} \in C | X_s\right\}.$$

What is its transition density function

$$q(t, x, y) = \frac{\mathbb{P}\left\{X_{t+s} \in dy | X_s = x\right\}}{dy}?$$

EXERCISE 4.4. Let L_t be the local time of Brownian motion at $x = 0$. Show that

$$\mathbb{E} L_t = \sqrt{\frac{8t}{\pi}}.$$

EXERCISE 4.5. Show that Brownian bridge from o to x in time 1 is a Markov process with transition density is

$$q(s_1, y; s_2, z) = \frac{p(s_2 - s_1, y, z)p(1 - s_2, z, x)}{p(1 - s_1, y, x)}.$$

EXERCISE 4.6. The finite-dimensional marginal density for Brownian bridge from o to x in time 1 is

$$p(1,o,x)^{-1}\prod_{i=0}^{l} p(s_{i+1}-s_i, x_i, x_{i+1}).$$

[Convention: $x_0 = o, s_0 = 0$.]

EXERCISE 4.7. Let $\{W_s,\ 0 \le s \le 1\}$ be a Brownian bridge at o. Then the reversed process

$$\{W_{1-s}, 0 \le s \le 1\}$$

is also a Brownian bridge.

EXERCISE 4.8. For an $a > 0$ define the stopping time

$$\sigma_a = \inf\{t:\ B_t - t = -a\}.$$

Show that for $\mu \ge 0$,

$$\mathbb{E}e^{-\mu\sigma_a} = e^{-(\sqrt{1+2\mu}-1)a}.$$

EXERCISE 4.9. Use the result of the previous exercise to show that $\mathbb{E}e^{\mu\sigma_a}$ is infinite if $\mu > 1/2$.

EXERCISE 4.10. By the martingale representation theorem there is a process H such that

$$W_1^3 = \int_0^1 H_s\, dW_s.$$

Fin an explicit expression for H.

CHAPTER 5

Stochastic Differential Equations

1. Simplest stochastic differential equations

In this section we discuss a stochastic differential equation of a very simple type.

Let M be a martingale in and A a process of bounded variation. Let a and b be two real-valued functions and consider the following stochastic differential equation

$$dX_t = a(X_t)\,dM_t + b(X_t)\,dA_t, \quad X_0 = x.$$

The meaning of this equation is the following: we are looking for a semimartingale X such that

$$X_t = x + \int_0^t a(X_s)\,dM_s + \int_0^t b(X_s)\,dAs.$$

Thus the equation we want to solve is a stochastic differential equation question only in name; it is more appropriate to be called a stochastic integral equation, but nobody does that. Note that diffusion coefficient $a(X_s)$ and drift coefficient $b(X_s)$ are functions of the current position X_s of the solution rather than functions of the path X. This type of stochastic equation is often called an Itô type stochastic differential equation. In the most popular case M is a Brownian motion and A is the time itself, i.e.,

$$dX_t = a(X_t)\,dB_t + b(X_t)\,dt, \qquad X_0 = x.$$

This is a stochastic differential equation driven by the Brownian motion B with starting point x. The handling of the general form of equations we formulated above is only slightly more complicated (by a time change) than that of this special case.

The main theorem below concerns the existence and uniqueness of Itô type stochastic differential equation under global Lipschtiz condition on the coefficients.

THEOREM 1.1. *Assume that $a(x), b(x)$ satisfy the global Lipschitz condition: there exists a constant K such that*

$$|a(x) - a(y)| + |b(x) - b(y)| \le K|x - y|$$

for all x, y. Then the stochastic differential equation

$$dX_t = a(X_t)\,dM_t + b(X_t)\,dA_t, \quad X_0 = x$$

has a unique solution.

PROOF. The stochastic differential equation looks very much like an ordinary differential equation: $dx_t = b(x_t)\,dt$. In fact this is a special case of the general stochastic differential equation formulated above. Recall that ordinary differential equations of this type can be solved by Picard's iteration. The same method can be used to solve the stochastic differential equation.

As will be seen below, it will be a great convenience if the quadratic variation $\langle M, M\rangle$ and the total variation $|A|$ are absolutely continuous with respect to a fixed non-random increasing function. This is the case if M is a Brownian motion and A is the time. In general we can achieve this by a time change. Let

$$\phi(t) = \langle M, M\rangle_t + |A|_t + t.$$

The random function ϕ is continuous and strictly increasing, the latter property being one of the reasons we have added t in the definition of $\phi(t)$. We denote its (pathwise) inverse function by τ. It is easy to see that $\tau(t)$ is a stopping time for each fixed t. Furthermore, $t \mapsto \tau(t)$ is absolutely continuous with respect to t, which is the second reason we have added t in the definition of $\phi(t)$. We can write this property as $d\tau(t) \le dt$. We make a time change to all the processes involved; namely,

$$\tilde{M}_t = M_{\tau(t)}, \quad \tilde{A}_t = A_{\tau(t)}, \quad \tilde{X}_t = X_{\tau(t)}.$$

Then we have

$$d\langle \tilde{M}, \tilde{M}\rangle_t \le dt \quad \text{and} \quad d|A|_t \le dt.$$

The time changed equation is

$$d\tilde{X}_t = a(\tilde{X}_t)\,d\tilde{M}_t + b(\tilde{X}_t)\,d\tilde{A}_t, \quad \tilde{X}_0 = x.$$

If the existence and uniqueness hold for this equation for the semimartingale $\tilde{X}$ (with respect to the time changed filtration), then they also hold for the original equation of for X, and vice versa. Thus without loss of generality we may assume that

$$d\langle M, M\rangle_t \le dt, \quad \text{and} \quad d|A|_t \le dt \tag{1.1}$$

for the original equation. The advantage of these assumptions will be clear in the course of the proof.

We will prove the uniqueness first. Suppose that that Y is another solution,

$$dY_t = a(Y_t)\,dM_t + b(Y_t)\,dA_t, \quad Y_0 = x.$$

Subtracting the two equations, squaring the result, using the inequality $(a+b)^2 \le 2a^2 + 2b^2$ on the right side, and taking the expectation, we have

$$\begin{aligned}
E|Y_t - X_t|^2 \le & 2\mathbb{E}\left(\int_0^t \{a(Y_s) - a(X_s)\}\, dM_s\right)^2 \\
& + 2\mathbb{E}\left(\int_0^t \{b(Y_s) - b(X_s)\}\, dA_s\right)^2 \\
\le & 2\mathbb{E}\int_0^t |a(Y_s) - a(X_s)|^2 d\langle M, M\rangle_s \\
& + 2\mathbb{E}\left(|A|_t \int_0^t |b(Y_s) - b(X_s)|^2\, d|A|_s\right) \\
\le & 2K^2\mathbb{E}\int_0^t |Y_s - X_s|^2\, dt + 2K^2 t\mathbb{E}\int_0^t |Y_s - X_s|^2\, ds.
\end{aligned}$$

Note that we have taken the advantage of the assumptions (1.1) and the fact that $|A|_t \le t$. Now let $c(t) = \mathbb{E}\left[|Y_t - X_t|^2\right]$. For $t \le T$, we have

$$c(t) \le C_T \int_0^t c(s)\, ds.$$

with $C_T = 2K^2(1+T)$. From this inequality it follows immediately that $c(t) = 0$, which proves the uniqueness.

We now prove the existence by iteration. Let $X_t^0 = x$ and define

$$X_t^n = x + \int_0^t a(X_s^{n-1})\, dM_s + \int_0^t b(X_s^{n-1})\, dA_s. \tag{1.2}$$

We prove that X^n converges to a process which is a solution of the equation. Consider the difference $X_t^{n+1} - X_t^n$. We have as before

$$\begin{aligned}
\max_{0\le s\le t} |X_s^{n+1} - X_s^n| \le & 2\max_{0\le s\le t}\left(\int_0^s \left\{a(X_u^n) - a(X_u^{n-1})\right\}\, dM_u\right)^2 \\
& + 2TK^2 \int_0^t \max_{0\le u\le s} |X_u^n - X_u^{n-1}|^2\, ds.
\end{aligned}$$

It follows from Doob's martingale moment inequality we have for $t \le T$

$$\eta_{n+1}(t) \le D_T \int_0^t \eta_n(s)\, ds,$$

where

$$\eta_n(t) = \mathbb{E}\left[\max_{0\le s\le t} |X_s^n - X_s^{n-1}|^2\right]$$

and $D_T = 2K^2(4+T)$. Let $C = \eta_1(T)$. We have by the recursive inequality for $\eta_n(t)$,

$$\eta_n(T) \le \frac{C(TD_T)^{n-1}}{(n-1)!}.$$

From this we have by the Markov inequality,

$$\sum_{n=1}^{\infty} \mathbb{P}\left[\max_{0\le s\le T} |X_s^n - X_s^{n-1}| \ge 2^{-n}\right] < \infty.$$

By the easy part of the Borel–Cantelli lemma, we conclude that the limit $\lim_{n\to\infty} X_t^n = X_t$ exists and the limiting process X_t is continuous. Since the convergence is also in L^2 sense, we may take the limit in the recursive relation (1.2) and obtain

$$X_t = x + \int_0^t a(X_s)\, dM_s + \int_0^t b(X_s)\, dA_s.$$

Thus the limit process is a continuous semimartingale which satisfies the equation. The existence has been proved. □

Although we have proved the existence and uniqueness for one dimensional Itô type stochastic differential equations only, the theorem and the proof itself can be easily generalized to multi-dimensional setting. We point out that the dimensions of the driving process and the solution semimartingale are not necessarily the same. A general Itô type stochastic differential equation has the form

$$dX_t = \sigma(X_t)\, dZ_t,$$

where $\sigma : \mathbb{R}^n \to \mathscr{M}(n,m)$ (the space of $(n \times m)$ matrices, Z an $\mathbb{R}^m$-valued semimartingale, and the solution is an $\mathbb{R}^n$-valued semimartingale. If σ is globally Lipschitz, then the equation has a unique solution. For example, if we take $Z = (M, A)^T$ and $\sigma = (\sigma_1, b)$, then this general form reduces to the one dimensional case considered above. Note that for the proof, we need to separate the martingale and bounded variation part of the semimartingale Z and treat them differently, as we have done in the proof of the above theorem.

2. Locally Lipschitz coefficients. Explosion.

In applications the coefficients of a stochastic differential equation may only be locally Lipschitz. A function $\sigma : \mathbb{R}^n \to \mathbb{R}^m$ from a euclidean space to another is called locally Lipschitz if for any positive integer N there is a constant C_N such that

$$|\sigma(x) - \sigma(y)| \le C_N |x - y|$$

for all $|x| \le N$ and $|y| \le N$. A globally Lipschitz function grows at most linearly, but this is not so for a locally Lipschitz function. For this reason, we have to allow the possibility of explosion of a stochastic differential equation.

For a continuous function $x : [0, e) \to \mathbb{R}^d$ we say that e is the explosion time for x if either $e = \infty$ or $\lim_{t\uparrow e} |x_t| = \infty$. The explosion time of a path x is denoted by $e(x)$. Let τ be a stopping time and X a continuous process defined on the time interval $[0, \tau)$. If there exists a sequence of stopping times $\tau_n \uparrow \tau$ such that each stopped process $X_t^{\tau_n} = X_{t\wedge\tau_n}$ is a semimartingale for each n, then X is called a semimartingale up to time τ, or a semimartingale defined on $[0, \tau)$. Consider a stochastic differential equation

$$dX_t = \sigma(X_t)\, dZ_t,$$

where the semimartingale Z is defined up to a stopping time τ. A semimartingale X up to a stopping time τ is a solution of the stochastic differential equation if

$$X_t = X_0 + \int_0^t \sigma(X_s)\, dZ_s, \qquad t < \tau.$$

Note that if X is a semimartingale defined up to time τ, then the stochastic integral on the right side makes sense for all $t < \tau$. Equivalently, X is a solution if there exists a sequence of stopping times $\tau_n \uparrow \tau$ such that for all $t \geq 0$ and all n,

$$X_{t\wedge\tau_n} = X_0 + \int_0^{t\wedge\tau_n} \sigma(X_s) dZ_s.$$

It now makes sense to speak of a solution X of an Itô type stochastic differential equation up to its explosion time. The main theorem is as follows.

THEOREM 2.1. *Suppose that that σ is locally Lipschitz and Z is a semimartingale (defined for all time). Then there is a unique solution to the stochastic differential equation*

$$dX_t = \sigma(X_t)\, dZ_t$$

up to its explosion time.

PROOF. We divide the proof into several steps. We first assume that X_0 is uniformly bounded, say, by N_0. In the last step we remove this restriction.

(a) *Construction of a solution.* For a fixed positive integer $N \geq N_0$, let $\sigma^N : \mathbb{R}^d \to \mathbb{R}^{d\times l}$ be globally Lipschitz such that $\sigma^N(z) = \sigma(z)$ for $z \in B(0; N)$, the ball of radius N centered at the origin. Consider the equation

$$X_t^N = X_0 + \int_0^t \sigma^N \left(X_s^N\right)\, dZ_s. \tag{2.1}$$

This equation has a unique solution. We show that X^N and X^{N+1} coincides before either one of them wanders more than a distance of N from the origin. Let

$$\tau_N = \inf\left\{t \geq 0 : |X_t^N| \text{ or } |X_t^{N+1}| = N\right\}.$$

Since $\sigma^N = \sigma^{N+1}$ on $B(0; N)$, both X^{N,τ_N} and X^{N+1,τ_N} (the two processes X^N and X^{N+1} stopped at τ_N) are solutions of the equation

$$Y_t = X_0 + \int_0^t \sigma^N(Y_s)\, dZ_s^{\tau_N}. \tag{2.2}$$

By uniqueness we have $X^{N,\tau_N} = X^{N+1,\tau_N}$, hence $X_t^N = X_t^{N+1}$ for $0 \leq t \leq< \tau_N$, and τ_N is the first time the common process reaches a distance of N from the origin. It is clear that $\tau_N \leq \tau_{N+1}$.

Let $e = \lim_{N\uparrow\infty} \tau_N$, and define a semimartingale X on $[0, e)$ by $X_t = X_t^N$ for $0 \le t < \tau_N$. We have from (2.1)

$$X_t = X_0 + \int_0^t \sigma^N (X_s)\, dZ_s, \qquad 0 \le t < \tau_N.$$

Therefore X is a solution of the stochastic differential equation up to time e.

(b) *e is the explosion time for X.* This is the key step of the proof. This means that almost surely, either $e = \infty$ or $e < \infty$ and $\lim_{t\uparrow e} |X_t| = \infty$. Equivalently, if $e < \infty$, then for each fixed positive $R \ge N_0$, then there exists a $t_R < e$ such that $|X_t| \ge R$ for all $t_R \le t < e$.

The idea of the proof is as follows. Because on the ball $B(0; R+1)$ the coefficients of the equation are bounded, X needs to spend at least a fixed amount of time (in an appropriate probabilistic sense) when it crosses from $\partial B(0; R)$ to $\partial B(0; R+1)$. If $e < \infty$, this can happen only finitely many times. Thus after some time, X either never returns to $B(0; R)$ or stays inside $B(0; R+1)$ forever; but the second possibility contradicts the facts that $|X_{\tau_N}| = N$ and $\tau_N \uparrow e$ as $N \uparrow \infty$.

To proceed rigorously, define two sequences $\{\eta_n\}$ and $\{\zeta_n\}$ of stopping times inductively by

$$\begin{aligned}
\zeta_0 &= 0, \\
\eta_n &= \inf \{ t > \zeta_{n-1} : |X_t| = R \}, \\
\zeta_n &= \inf \{ t > \eta_n : |X_t| = R+1 \},
\end{aligned}$$

with the convention that $\inf \emptyset = e$. If $\zeta_n < e$, the difference $\zeta_n - \eta_n$ is the time X takes to cross from $\partial B(0; R)$ to $\partial B(0; R+1)$ for the nth time. By Itô's formula applied to the function $f(x) = |x|^2$ we have for $t < e$,

$$|X_t|^2 = |X_0|^2 + N_t + \int_0^t \Psi_s dQ_s, \tag{2.3}$$

where

$$\begin{aligned}
N_t &= 2 \int_0^t \sigma^i_\alpha (X_s) X^i_s \, dM^\alpha_s \\
\Psi_s &= 2\sigma^i_\alpha (X_s) X^i_s \frac{dA^\alpha_s}{dQ_s} + \sigma^i_\alpha (X_s) \sigma^i_\beta (X_s) \frac{d\langle M^\alpha, M^\beta \rangle_s}{dQ_s} \\
\langle N, N \rangle_t &= \int_0^t \Phi_s dQ_s \\
\Phi_s &= 4\sigma^i_\alpha (X_s) \sigma^j_\beta (X_s) X^i_s X^j_s \frac{d\langle M^\alpha, M^\beta \rangle_s}{dQ_s}.
\end{aligned}$$

Here we have assumed that $Z = M + A$ is the Dooby-Meyer decomposition of the semimartingale Z and By Lévy's criterion, there exists a one-dimensional Brownian motion W such that

$$N_{s+\eta_n} - N_{\eta_n} = W_{\langle N,N \rangle_{s+\eta_n} - \langle N,N \rangle_{\eta_n}}.$$

When $\eta_n \leq s \leq \zeta_n$ we have $|X_s| \leq R+1$, hence there is a constant C depending on R such that $|\Psi_s| \leq C$ and $|\Phi_s| \leq C$. From (2.3) we have

$$\begin{aligned} 1 &\leq |X_{\zeta_n}|^2 - |X_{\eta_n}|^2 \\ &= W_{\langle N,N\rangle_{\zeta_n} - \langle N,N\rangle_{\eta_n}} + \int_{\eta_n}^{\zeta_n} \Psi_s dQ_s \\ &\leq W^*_{\langle N,N\rangle_{\zeta_n} - \langle N,N\rangle_{\eta_n}} + C\left(Q_{\zeta_n} - Q_{\eta_n}\right), \end{aligned}$$

where $W_t^* = \max_{0\leq s\leq t} |W_s|$, and

$$\langle N, N\rangle_{\zeta_n} - \langle N, N\rangle_{\eta_n} \leq C(Q_{\zeta_n} - Q_{\eta_n}).$$

Now it is clear that $\zeta_n < e$ and $Q_{\zeta_n} - Q_{\eta_n} \leq (Cn)^{-1}$ imply that

$$W^*_{1/n} \geq 1 - \frac{1}{n} \geq \frac{1}{2},$$

and

$$\begin{aligned} P\left\{\zeta_n < e, Q_{\zeta_n} - Q_{\eta_n} \leq \frac{1}{Cn}\right\} &\leq P\left\{W^*_{1/n} \geq 1/2\right\} \\ &= \sqrt{\frac{2n}{\pi}} \int_{1/2}^{\infty} e^{-nu^2/2} du \leq \sqrt{\frac{8}{\pi n}} e^{-n^2/8}. \end{aligned}$$

By the Borel-Cantelli lemma, we have either $\zeta_n = e$ for some n or $\zeta_n < e$ and $Q_{\zeta_n} - Q_{\eta_n} \geq (Cn)^{-1}$ for all sufficiently large n. The second possibility implies that

$$Q_{\zeta_n} \geq \sum_{m=1}^{n-1} (Q_{\zeta_m} - Q_{\eta_m}) \to \infty.$$

This implies in turn that $\zeta_n \uparrow \infty$ and $e > \zeta_n \to \infty$. Thus if $e < \infty$, we must have $\zeta_n = e$ for some n. Let ζ_{n_0} be the last ζ_n strictly less than e. Then X never returns to $B(0; R)$ for $\zeta_{n_0} \leq t < e$. This shows that e is indeed the explosion time of X.

(c) *Uniqueness*. Suppose that Y is another solution up to its explosion time. Let τ_N be the first time either $|X_t|$ or $|Y_t|$ is equal to N. Then X and Y stopped at time τ_N are solutions of the equation (see (2.2)). By uniqueness we have $X_t = Y_t$ for $0 \leq t < \tau_N$ and τ_N is the time the common process first reaches a distance N from the origin. Hence

$$e(X) = e(Y) = \lim_{N\uparrow\infty} \tau_N$$

and $X_t = Y_t$ for all $0 \leq t < e(X)$.

(d) *General initial condition*. For a general X_0 let $\Omega_N = \{|X_0| \leq N\}$ and X^N the solution of the equation with the initial condition $X_0 I_{\Omega_N}$). Define a new probability measure by $\mathbb{Q}^N(C) = \mathbb{P}(C \cap \Omega_N)/\mathbb{P}(\Omega_N)$. Since $\Omega_N \in \mathscr{F}_0$, both X^N and X^{N+1} are solutions to the same equation under Q^N but now also with the same initial condition $X_0 I_{\Omega_N}$. Hence by Part (c) they must coincide, i.e., $X^N = X^{N+1}$ and $e(X^N) = e(X^{N+1})$ on Ω_N. In view of the fact that $\mathbb{P}(\Omega_N) \uparrow 1$ as $N \uparrow \infty$, we can define $X = X^N$ and $e = e(X^N)$ on Ω_N. Then it is clear that X is a semimartingale and satisfies the SDE up to its explosion time. The uniqueness follows from the observation that if Y is another solution, then it must be a solution to the SDE with the initial condition $X_0 I_{\Omega_N}$ under $\mathbb{Q}^N$. Thus it must coincide with X on the set Ω_N for all N. □

For future reference we need the following slightly more general form of uniqueness. We will leave its proof to the reader.

PROPOSITION 2.2. *Suppose that σ is locally Lipschitz. Let X and Y be two solutions of the stochastic differential equation*

$$dX_t = \sigma(X_t)\, dZ_t$$

up to stopping times τ and η, respectively, with the same initial condition. Then $X_t = Y_t$ for $0 \le t < \tau \wedge \eta$. In particular, if X is a solution up to its explosion time $\tau = e(X)$, then $\eta \le e(X)$ and $X_t = Y_t$ for $0 \le t < \eta$.

The following result gives a well known sufficient condition for non-explosion.

PROPOSITION 2.3. *If σ is locally Lipschitz and there is a constant C such that $|\sigma(x)| \le C(1 + |x|)$, then solutions of the stochastic differential equation of the type*

$$dX_t = \sigma(X_t)\, dZ_t$$

do not explode.

PROOF. We may assume that X_0 is uniformly bounded (see Part (d)) of the proof of THEOREM 2.1). Let η_t be defined as before and let $\tau_N = \inf\{t > 0 : |X_t| = N\}$. We have

$$X_t = X_0 + \int_0^t \sigma(X_s) dZ_s$$

we have

$$\begin{aligned}\mathbb{E}|X|^2_{\infty,\eta_T\wedge\tau_N} &\le 2\mathbb{E}|X_0|^2 + C\mathbb{E}\int_0^{\eta_T\wedge\tau_N} \left\{1 + |X_s|^2\right\} dQ_s \\ &\le 2\mathbb{E}|X_0|^2 + CT + C\int_0^T \mathbb{E}|X|^2_{\infty,\eta_s\wedge\tau_N} ds,\end{aligned}$$

Hence by Gronwall's inequality

$$\mathbb{E}|X|^2_{\infty,\eta_T\wedge\tau_N} \le \left(2\mathbb{E}|X_0|^2 + CT\right) e^{CT}.$$

Letting $N \uparrow \infty$ we see that $|X|_{\infty,\eta_T\wedge e(X)} < \infty$, a.s. This implies $\eta_T \wedge e(X) < e(X)$, hence $\eta_T < e(X)$. Now $e(X) = \infty$ follows from the fact that $\eta_T \uparrow \infty$ as $T \uparrow \infty$. □

EXAMPLE 2.4. *Accelerated Brownian motion.* According to PROPOSITION 2.3, if σ has at most linear growth, then no explosion is possible for any driving semimartingales Z and initial condition X_0. However, explosion or non-explosion is not a property of σ alone. It may happen that if Z is intrinsically slow, there will be no explosion no matter how fast σ grows. To see this consider the equation

$$dX_t = \sigma(X_t)dB_t, \quad X_0 = 0, \tag{2.4}$$

where B be a d-dimensional Brownian motion and σ a positive, locally Lipschitz function on $\mathbb{R}^d$. When $d = 1$ or 2, there will be no explosion regardless what σ is. The reason is, of course, that Brownian motion itself is recurrent for in these dimensions. To see this more clearly let

$$\phi_t = \int_0^t \sigma(X_s)^2 ds, \tag{2.5}$$

and $\tau : [0, \phi_\infty) \to [0, e(X))$ the inverse function of $\phi : [0, e(X)) \to [0, \phi_\infty)$. Then $W_t = X_{\tau_t}$ is a Brownian motion. If the original process $X_t = W_{\phi_t}$ explodes, then $|X_t| \to \infty$ as $t \uparrow e(X) < \infty$ and this cannot happen in dimensions 1 and 2 because W is recurrent in these dimensions. In dimensions 3 and higher, W is transient: $|W_t| \to \infty$ as $t \to \infty$. Thus X explodes if and only if ϕ explodes, and this happens only when τ_∞ is finite. By the law of iterated logarithm, $|W_t| \leq C\sqrt{t \log\log t}$ for large t. Thus if σ grows at most linearly, we have

$$\tau_t = \int_0^t \frac{ds}{a(W_s)^2} \geq C \int_{t_0}^t \frac{ds}{s \log\log s} \to \infty, \quad \text{as } t \to \infty$$

and X does not explode, in agreement with PROPOSITION 2.3.

So far we have not consider the possibility that the driving semimartingale Z may be defined only up to a stopping time τ.

THEOREM 2.5. *Let Z be a semimartingale defined up to a stopping time τ. Then the stochastic differential equation*

$$dX_t = \sigma(X_t)\, dZ_t$$

has a unique solution X up to the stopping time $e(X) \wedge \tau$. If Y is another solution up to a stopping time $\eta \leq \tau$, then $\eta \leq e(X) \wedge \tau$ and $X_t = Y_t$ for $0 \leq t < e(X) \wedge \eta$.

3. Weak solution and weak uniqueness

For simplicity, we will denote a general stochastic differential equation

$$dX_t = \sigma(X_t)\, dZ_t$$

with initial condition X_0 by $SDE(\sigma, Z, X_0)$. Stochastic differential equations of this type makes sense at least for continuous coefficients and can be solved for coefficients much more general than Lipschitz ones if they are interpreted appropriately. Uniqueness for such equations is in general a complicated question. In contrast to the pathwise uniqueness , the weak uniqueness (also called uniqueness in law) asserts roughly that if (Z, X_0) and $(\hat{Z}, \hat{X}_0)$ have the same law, then the solutions X and $\hat{X}$ of $SDE(\sigma, Z, X_0)$ and $SDE(\sigma, \hat{Z}, \hat{X}_0)$ (not necessarily on the same probability space or with the same filtration) also have the same law. This concept is most useful when we deal with classical stochastic differential equations of Itô type, those driven by multi-dimensional Brownian motion: $Z = (W, t), \sigma = (\tilde{\sigma}, \tilde{b})$, where W is an euclidean Brownian motion. In this case the equation can be written in the form

$$X_t = X_0 + \int_0^t \sigma(X_s) dW_s + \int_0^t b(X_s) ds. \tag{3.1}$$

We will need the weak uniqueness for this type of equations when we discuss the uniqueness of diffusion measures generated by second order elliptic operators. In the following $W(\mathbb{R}^d)$ denotes the space of continuous functions $x : [0, e(x)) \to \mathbb{R}^d$ such that $\lim_{t \uparrow e(x)} |x_t| = \infty$. Our main result is that for such equations weak uniqueness holds.

PROPOSITION 3.1. *Suppose that σ is locally Lipschitz. Then the weak uniqueness holds for the stochastic differential equation (3.1). More precisely, let $\hat{X}$ be the solution of*

$$\hat{X}_t = \hat{X}_0 + \int_0^t \sigma(\hat{X}_s) d\hat{W}_s + \int_0^t b(\hat{X}_s) ds, \tag{3.2}$$

where $\hat{W}$ is a Brownian motion, possibly on a different filtered probability space $(\hat{\Omega}, \hat{\mathscr{F}}_*, \tilde{\mathbb{P}})$ and $\hat{X}_0$ and X_0 have the same distribution, then $\hat{X}$ and X have the same law.

PROOF. The idea is to pass to the product probability space

$$(\mathbb{R}^d \times W(\mathbb{R}^l), \mathscr{B}(\mathbb{R}^d) \times \mathscr{B}(W(\mathbb{R}^l)), \mathscr{B}(\mathbb{R}^d) \times \mathscr{B}(W(\mathbb{R}^l))_*, \mu_0 \times \mu^W),$$

where μ_0 is the common distribution of X_0 and $\hat{X}_0$, and μ^W is the law of l-dimensional Brownian motion. A point in this space is denoted by (x_0, w). The stochastic differential equation on this space

$$\phi_t = x_0 + \int_0^t \sigma(\phi_s) dw_s + \int_0^t b(\phi_s) ds \tag{3.3}$$

has a unique solution $\phi = \phi(x_0, w)$ up to its explosion time. Now $\phi : \mathbb{R}^d \times W(\mathbb{R}^l) \to W(\mathbb{R}^d)$ is a measurable map defined $(\mu_0 \times \mu^W)$-almost everywhere. Since the law of (X_0, B) is $\mu_0 \times \mu^W$, the composition $\phi(X_0, W) : \Omega \to W(\mathbb{R}^d)$ is well defined. By (3.3), $\phi(X_0, W)$ as a stochastic process satisfies (3.1). It should be pointed out that although the stochastic integral on the right side is not defined path by path, the above passage from (3.3) to (3.1) is valid because the stochastic integral is the limit (in probability) of an appropriate sequence of discrete approximations, e.g.,

$$\sum_{j=0}^{m} \sigma(\phi_{jt/m}) \left\{ w_{(j+1)t/m} - w_{jt/m} \right\} \to \int_0^t \sigma(\phi_s) dw_s.$$

Now by the pathwise uniqueness for (3.1) we have $X = \phi(X_0, W)$. Likewise we have $\hat{X} = \phi(\hat{X}_0, \hat{W})$. Because (X_0, W) and $(\hat{X}_0, \hat{W})$ have the same law, we conclude that X and $\hat{X}$ must also have the same law. $\square$

4. Stratonovich formulation

We now turn to the Stratonovich formulation of stochastic differential equations. The advantage of this formulation is that Itô's formula appears in the same form as the fundamental theorem of calculus and stochastic calculus in this formulation takes a more familiar form (compare (4.2)). This is a very convenient feature when we study stochastic differential equations under different coordinate systems (e.g., on a differentiable manifold). However, it often happens that useful probabilistic and geometric information reveals itself only after we separate martingale and bounded variation components of the equation under consideration and its solutions.

Suppose that $V_\alpha, \alpha = 1, \ldots, l$ are smooth vector fields on $\mathbb{R}^d$. Each V_α can be regarded as a smooth function $V_\alpha : \mathbb{R}^d \to \mathbb{R}^d$ so that $V = (V_1, \ldots, V_l)$ is an $\mathscr{M}(d, 1)$-valued function on $\mathbb{R}^d$. Let Z be a semimartingale and consider a Stratonovich stochastic differential equation

$$X_t = X_0 + \int_0^t V(X_s) \circ dZ_s,$$

where the stochastic integral is in the Stratonovich sense. To emphasize the fact that V is a set of l vector fields, we rewrite the equation as

$$X_t = X_0 + \int_0^t V_\alpha(X_s) \circ dZ_s^\alpha. \tag{4.1}$$

Converting the Stratonovich integral into the equivalent Itô integral, we have

$$X_t = X_0 + \int_0^t V_\alpha(X_s)\, dZ_s^\alpha + \frac{1}{2} \int_0^t \nabla_{V_\beta} V_\alpha(X_s)\, d\langle Z^\alpha, Z^\beta \rangle_s.$$

Here $\nabla_\beta V_\alpha$ is the derivative of V_α along V_β. This is an Itô type stochastic differential equation we have studied before and is equivalent to the original equation (4.1). Itô's formula in this setting becomes the following. We leave its proof as an exercise.

PROPOSITION 4.1. *Let X be a solution to the SDE (4.1) and $f \in C^2(\mathbb{R}^d)$. Then*

$$f(X_t) = f(X_0) + \int_0^t V_\alpha f(X_s) \circ dZ_s, \qquad 0 \le s < e(X). \tag{4.2}$$

5. Stochastic differential equation with reflecting boundary condition

When the space where a solution of a stochastic differential equation lives in has a boundary, we in general have to specify its behavior at the boundary. Normal reflection is the most typical boundary condition. In this section we will consider the simplest case where the space is the half line $\mathbb{R}_+ = [0, \infty)$. A stochastic differential equation with normal reflection has the following form

$$dX_t = a(X_t)\, dM_t + b(X_t)\, dA_t + d\phi_t \quad X_0 = x \geq 0.$$

The meaning of the equation is as follows. As before, a is the diffusion coefficient and b is the drift coefficient. Given a continuous local martingale M and a process of bounded variation A, we look for a semimartingale X such that

$$X_t = x + \int_0^t \sigma(X_s)\, dM_s + \int_0^t b(X_s)\, dA_s + \phi_t,$$

where ϕ is an increasing process which increases only when $X_t = 0$. More precisely, (X_t, ϕ_t) is the solution of the Skorokhod equation for the process

$$\int_0^t a(X_s)\, dM_s + \int_0^t b(X_s)\, dA_s.$$

Note that it is boundary local time term ϕ that keeps the process in the positive half line. Also note that we are seeking not only the semimartingale X but also the boundary local time ϕ at the same time. The solution is the pair (X, ϕ), not just X alone. When $a = 1$, $b = 0$, and M is a Brownian motion, we have the Skorokhod problem discussed in SECTION 7 and the solution (X, ϕ) consists of a reflecting Brownian motion X and its local time ϕ.

The fundamental result of stochastic differential equations with reflection is as follows.

THEOREM 5.1. *Suppose that $a(x), b(x)$ satisfy a global Lipschitz condition, i.e., there exists a constant K such that for all x and y,*

$$|a(x) - a(y)| + |b(x) - b(y)| \leq K|x - y|.$$

Then the stochastic differential equation with reflecting boundary condition

$$dX_t = a(X_t)\, dM_t + b(X_t)\, dA_t + \phi_t, \quad X_0 = x \geq 0$$

has a unique solution (X, ϕ).

For the existence of the solution, we need the explicit solution of the Skorokhod problem. Recall that for a continuous function $f : \mathbb{R}_+ \to \mathbb{R}_+$, the Skorokhod equation is

$$x_t = f_t - \min_{0 \leq s \leq t} f_s \wedge 0.$$

Define the Skorokhod map γ from the space $W(\mathbb{R})$ of continuous paths on $\mathbb{R}^1$ to the space $W(\mathbb{R}_+)$ of continuous paths on $\mathbb{R}_+$ by

$$\gamma(f)_t = f_t - \min_{0 \leq s \leq t} f_s \wedge 0.$$

The map γ is Lipschitz continuous in the following sense

$$\max_{0\le s\le t} |\gamma(f)_s - \gamma(g)_s| \le \max_{0\le s\le t} |f_s - g_s|.$$

Consider the following stochastic differential equation:

$$dY_t = \sigma(\gamma(Y)_t)\, dM_t + b(\gamma(Y)_t)\, dA_t, \quad Y_0 = x. \tag{5.1}$$

Note that the boundary local time term has been removed from the original equation and the diffusion and drift coefficients are no longer functions of Y_t, the solution at time t. Instead, they depend on the whole past history of the solution. Thus we are not dealing with an Itô type stochastic differential equation. Nevertheless, because of the Lipschitz continuity of the map γ mentioned above, the equation can be solved in exactly the same way as we have done in SECTION 1, using the Lipschitz continuity

$$\max_{0\le s\le t} |\sigma(\gamma(f)_s) - \sigma(\gamma(g)_s)| \le K \max_{0\le s\le t} |f_s - g_s|$$

and

$$\max_{0\le s\le t} |b(\gamma(f)_s) - b(\gamma(g)_s)| \le K \max_{0\le s\le t} |f_s - g_s|.$$

The equation (5.1) can be solved uniquely.

The obvious thing to do now is to set

$$\phi_t = -\min_{0\le s\le t} Y_s \wedge 0 \quad \text{and} \quad X_t = Y_t + \phi_t$$

and show that the pair (X, ϕ) is a solution to the original stochastic differential equation with reflection. This is a triviality, for $X_t = \gamma(Y)_t$ and

$$X_t = Y_t + \phi_t = x + \int_0^t \sigma(X_s)\, dM_s + \int_0^t b(X_s)\, dA_s + \phi_t.$$

We have therefore completed the proof of the existence part of THEOREM 5.1.

The proof of uniqueness is quite long, but it follows by and large the idea in the proof of the uniqueness for stochastic differential equation without boundary; see THEOREM 1.1. Let (Y, ψ) be another solution. As was done in SECTION 1, by a time change we may assume that

$$d\langle M, M\rangle_t \le t \quad \text{and} \quad d|A|_t \le dt.$$

We subtract the equation for Y from that for X. The result can be written as

$$X_t - Y_t = Z_t + \phi_t - \psi_t, \tag{5.2}$$

where

$$Z_t = \int_0^t \{a(X_s) - a(Y_s)\}\, dM_s + \int_0^t \{b(X_s) - b(Y_s)\}\, dA_s. \tag{5.3}$$

The first term in (5.2) can be estimated as in the case of stochastic differential equation with boundary condition, but the second local time term needs a slightly different treatment. Instead of squaring the difference, we take the fourth power; the reason for doing so will be clear in later in the proof. We use the inequality

$$(x+y)^4 \leq 2^4x^4 + 2^4y^4,$$

take the maximum over the time interval $[0,t]$, and then take the expectation. After these steps we have

$$\mathbb{E}\max_{0\leq s\leq t}|X_s - Y_s|^4 \leq 2^4\mathbb{E}\max_{0\leq s\leq t}|Z_s|^4 + 2^4\mathbb{E}\max_{0\leq s\leq t}|\phi_s - \psi_s|^4.$$

Now, separating the martingale part

$$N_t = \int_0^t \{a(X_s) - a(Y_s)\}\, dM_s$$

of Z from its boundary variation part, we have

$$\mathbb{E}\max_{0\leq s\leq t}|Z_s|^4 \leq 2^4\mathbb{E}\max_{0\leq s\leq t}|N_s|^4 + 2^4\mathbb{E}\left(\int_0^t |b(X_s) - b(Y_s)|\, d|A|_s\right)^4.$$

The first term can be bounded by the martingale moment inequality and we obtain

$$\begin{aligned}
\mathbb{E}\max_{0\leq s\leq t}|N_s|^4 \leq& 4\mathbb{E}\langle N, N\rangle_t^2 \\
=& 4\mathbb{E}\left(\int_0^t |a(X_s) - a(Y_s)|^2\, d\langle M, M\rangle_s\right)^2 \\
\leq& 4K^4\mathbb{E}\left(\int_0^t |X_s - Y_s|^2\, ds\right)^2 \\
\leq& 4K^4T\mathbb{E}\int_0^t |X_s - Y_s|^4\, ds.
\end{aligned}$$

The second term of $\mathbb{E}\max_{0\leq s\leq t}|Z_s|^4$ can be bounded easily by Hölder's inequality and we have

$$\mathbb{E}\left(\int_0^t |b(X_s) - b(Y_s)|\, d|A|_s\right)^4 \leq K^4T^3\mathbb{E}\int_0^t |X_s - Y_s|^4\, ds.$$

Combining the two parts, we have

$$\mathbb{E}\max_{0\leq s\leq t}|Z_s|^4 \leq C_T E\left[\int_0^t |X_s - Y_s|^4\, ds\right].$$

Note that we are assuming that $t \leq T$ and C_T is the notation for a generic constant depending on T, whose value may vary from one appearance to another. So far we have proved the following inequality

$$\mathbb{E}\max_{0\leq s\leq t}|X_s - Y_s|^4 \leq C_T\mathbb{E}\int_0^t |X_s - Y_s|^4\, ds + 2^4\mathbb{E}\max_{0\leq s\leq t}|\phi_s - \psi_s|^4. \tag{5.4}$$

We now bound the second term involving $\phi - \psi$. Since it is continuous with bounded variation, we have

$$2^{-1}|\phi_t - \psi_t|^2 = \int_0^t (\phi_s - \psi_s)\, d(\phi_s - \psi_s).$$

Using

$$\phi_s - \psi_s = X_s - Y_s - Z_s$$

in the integrand, we have

$$2^{-1}|\phi_t - \psi_t|^2 = \int_0^t (X_s - Y_s)\, d(\phi_s - \psi_s) - \int_0^t Z_s\, d(\phi_s - \psi_s). \tag{5.5}$$

The crucial step in this proof is that the first term on the right is negative. In order to see this, we write

$$\int_0^t (X_s - Y_s)\, d(\phi_s - \psi_s) = \int_0^t (X_s - Y_s)\, d\phi_s - \int_0^t (X_s - Y_s)\, d\psi_s.$$

Since ϕ increases only when $X_s = 0$ and Y_s is always nonnegative, we see that ϕ increases on when $X_s - Y_s \le 0$, hence the first term on the right side is always negative; likewise the second term is always positive. It follows that

$$\int_0^t (X_s - Y_s)\, d(\phi_s - \psi_s) \le 0.$$

We are allowed to integrate by parts in the second term in (5.5) and write have by Itô's formula

$$\int_0^t Z_s\, d(\phi_s - \psi_s) = Z_t(\phi_t - \psi_t) - \int_0^t (\phi_s - \psi_s)\, dZ_s$$

Splitting the integrating process Z in the last term according to (5.3) and using the inequality

$$(x + y + z)^2 \le 2^3(x^2 + y^2 + z^2)$$

we have

$$2^{-5}\mathbb{E} \max_{0 \le s \le t} |\phi_s - \psi_s|^4 = J_1 + J_2 + J_3 \tag{5.6}$$

where

$$\begin{aligned}
J_1 &= \mathbb{E} \max_{0 \le s \le t} |\phi_s - \psi_s|^2 \max_{0 \le s \le t} |Z_s|^2 \\
J_2 &= \mathbb{E} \max_{0 \le s \le t} \left| \int_0^t (\phi_s - \psi_s)[a(X_s) - a(Y_s)]\, dM_s \right|^2 \\
J_3 &= \mathbb{E} \left| \int_0^t (\phi_s - \psi_s)\, dA_s \right|^2.
\end{aligned}$$

It should be clear now why we have taken the fourth power at the beginning of the proof. If we took the second power, we would be at a loss for a way to estimate the expected value of the maximum of the absolute value of a martingale in J_2. Now Using the inequality

$$x^2y^2 \le \epsilon x^4 + \frac{1}{4\epsilon}y^4$$

we have ($\epsilon = 2^{-7}$)

$$\begin{aligned} J_1 \le & 2^{-7}\mathbb{E}\max_{0\le s\le t}|\phi_s - \psi_s|^4 + 2^5\mathbb{E}\max_{0\le s\le t}|Z_s|^4 \\ \le & 2^{-7}\mathbb{E}\max_{0\le s\le t}|\phi_s - \psi_s|^4 + C_T\mathbb{E}\int_0^t |X_s - Y_s|^4\, ds. \end{aligned}$$

For J_2 we have ($\epsilon = 2^{-9}K^{-2}$)

$$\begin{aligned} J_2 \le & 4\mathbb{E}\int_0^t |\phi_s - \psi_s|^2 |a(X_s) - a(Y_s)|^2\, d\langle M, M\rangle_s \\ \le & 4K^2\mathbb{E}\max_{0\le s\le t}|\phi_s - \psi_s|^2 \cdot \int_0^t |X_s - Y_s|^2\, ds \\ \le & 2^{-7}\mathbb{E}\max_{0\le s\le t}|\phi_s - \psi_s|^4 + C_T\mathbb{E}\int_0^t |X_s - Y_s|^4\, ds. \end{aligned}$$

For J_3 we have ($\epsilon = 2^{-7}$)

$$J_3 \le 2^{-7}\mathbb{E}\max_{0\le s\le t}|\phi_s - \psi_s|^4 + C_T\mathbb{E}\int_0^t |X_s - Y_s|^4\, ds.$$

These estimates for J_1, J_2, and J_3 are to be used in (5.6). The coefficient of $\mathbb{E}\max_{0\le s\le t}|\phi_s - \psi_s|^4$ on the right side is $3\cdot 2^{-7}$, which is less than the coefficient 2^{-5} on the left side, hence we have

$$\mathbb{E}\max_{0\le s\le t}|\phi_s - \psi_s|^4 \le C_T\mathbb{E}\int_0^t |X_s - Y_s|^4\, ds$$

for some constant C_T. Using this inequality in (5.4), we obtain the following inequality for all $t \le T$:

$$\mathbb{E}|X_s - Y_s|^4 \le C_T\int_0^t \mathbb{E}|X_s - Y_s|^4\, ds.$$

It follows that $X_t = Y_t$ for all t, which is the uniqueness, and with this, we have completed the proof of THEOREM 5.1.

6. Fifth assignment

EXERCISE 5.1. Let X be the solution of the stochastic differential equation

$$dX_t^x = \sigma(X_t^x)\, dB_t + b(X_t^x)\, dt, \qquad X_0^x = x.$$

We assume that a, b are continuous and uniformly bounded. If f is a twice continuously differential function and f, f', and f'' are uniformly bounded. Show that

$$F(x,t) = \mathbb{E}f\left(X_t^x\right)$$

satisfies the parabolic partial differential equation

$$\frac{\partial F}{\partial t} = \frac{1}{2}\sigma^2 \frac{\partial^2 F}{\partial x^2} + b\frac{\partial F}{\partial x}, \qquad (x,t) \in \mathbb{R} \times \mathbb{R}_+$$

and the initial condition $\lim_{t\to 0} F(x,t) = f(x)$.

EXERCISE 5.2. Prove the Itô's formula in the Stratonovich form stated in PROPOSITION 4.1.

EXERCISE 5.3. Consider the stochastic differential equation

$$dX_t = \operatorname{sgn}(X_t)\, dB_t, \qquad X_0 = 0.$$

Show that the uniqueness does not hold for this equation. Note that by convention $\operatorname{sgn} 0 = 0$.

EXERCISE 5.4. Let $B = (B^1, B^2, \cdots, B^n)$ be an n-dimensional Brownian motion. Derive an Itô type stochastic differential equation for its radial process $r = |B|$.

EXERCISE 5.5. Show that a two-dimensional Brownian motion B can be written in the form $B_t = r_t e^{i\Theta_t}$, where the angular process Θ is a time-changed one-dimensional Brownian motion.

EXERCISE 5.6. One dimensional stochastic differential equations are uniquely solvable for a Hölder continuous diffusion coefficient with exponent $1/2$. Consider the equation

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

If there is a constant C such that

$$|\sigma(x) - \sigma(y)|^2 + |b(x) - b(y)| \le K|x - y|,$$

then the equation has a unique solution.

EXERCISE 5.7. Consider the stochastic differential equation

$$dX_t = \sqrt{X_t \vee 0}\, dB_t + \alpha dt, \qquad X_0 = x \ge 0$$

with $\alpha > 0$. Show that it has a unique positive solution.

EXERCISE 5.8. Consider the stochastic differential equation

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

Suppose that the diffusion and drift coefficients are even functions are even. Show that $Y_t = |X_t|$ is the solution of a stochastic differential equation with reflection.

EXERCISE 5.9. The equation

$$dX_t = X_t^2\, dB_t$$

does not explode, whereas the equation

$$dX_t = X_t^2\, dB_t + X_t^2\, dt$$

explodes.

EXERCISE 5.10. Consider the stochastic differential equation

$$dX_t = dB_t - X_t\,dt, \qquad X_0 = x.$$

Its solution is called an Ornstein-Uhlenbeck process. Show that it is a Gaussian process and

$$\mathbb{E}f(X_t^x) = \int_{\mathbb{R}} f\left(e^{-t/2}x + \sqrt{1-e^{-t}}y\right)\mu(dy),$$

where μ is the standard Gaussian measure on $\mathbb{R}$. This is called Mehler's formula

EXERCISE 5.11. Consider the stochastic differential equation for a Brownian bridge

$$dX_t = dB_t - \frac{X_t}{1-t}\,dt, \qquad X_0 = 0.$$

Show that $\lim_{t\uparrow 1} X_t = 0$ almost surely.

CHAPTER 6

Financial Mathematics

1. Portfolio management

Let $\mathscr{F}_*$ be the filtration generated by the Brownian motion. It represents the information of the market up to time t. We consider the simplest case where the market consists of a risk-free investment called bond and a risky investment called stock. For a time being, a portfolio is a combination of the two investments. At time t it can be described by (ξ_t, η_t), where ξ_t is the number of shares of the bond, and η_t is the number of shares in the stock. The bond and the stock are described as follows:

The bond pays a variable interest at the rate r_t, which we assume is adapted to $\mathscr{F}_*$. If the bond price is $B_0 = 1$ at time 0, then its share price B_t at time t is given by the equation

$$dB_t = B_t r_t \, dt, \qquad B_0 = 1.$$

This can be solved as

$$B_t = \exp\left[\int_0^t r_s ds\right].$$

The bond is considered as a risk-free investment. It is completely characterized by the process B_t. We call B_t the growth factor and B_t^{-1} the discount factor.

Note that if we want to interpret the risk-free investment as a bank deposit, there is a slight confusion of terminology. For a cash deposit, it is customary to consider that the share price is always fixed at 1 per share and it is the number of shares that is changing, the extra shares being considered as dividend from the investment. In this case, B_t is interpreted as the number of shares instead of the price per share. In these notes, we considered B_t as the share price of the bond if the initial share price is $B_0 = 1$. Therefore the bond pays no dividend and has volatility zero (see below).

The stock is considered as a risky investment. It is characterized by the rate of return μ_t and the volatility σ_t. Its share price S_t is assumed to satisfy the equation

$$dS_t = S_t(\sigma_t \, dW_t + \mu_t \, dt).$$

Here W is a one-dimensional $\mathscr{F}_*$-Brownian motion. Whenever necessary we assume that $\mathscr{F}_* = \mathscr{F}_*^W$, i.e., the market filtration is generated by that of the Brownian motion. Note that we do not specify its initial price S_0 as we did in the bond price. Explicitly, we have

$$S_t = S_0 \exp\left[\int_0^t \sigma_s \, dW_s s + \int_0^t \left(\mu_s - \frac{\sigma_s^2}{2}\right) ds\right].$$

A portfolio is described at time t by a pair of adapted processes $\pi = (\xi, \eta)$, where ξ_t and η_t are the numbers of shares of the bond and the stock, respectively, at time t. The wealth process X is simply the total value of the portfolio, i.e.,

$$X_t = \xi_t B_t + \eta_t W_t.$$

Suppose that at time $t = 0$, the investor has the initial asset $X_0 = x$. He invests his asset in the market to achieve a certain goal G at a terminal time T. He also needs to consume at a given rate c_t. The pair $C = (c, G)$ is called a contingent claim. To achieve these ends, he needs to constantly readjust his portfolio $\pi_t = (\xi_t, \eta_t)$ using the information up to time t of the market, i.e., ξ and η have to be adapted to the filtration $\mathscr{F}_*$. For him, the portfolio pair $\pi_t = (\xi_t, \eta_t)$ becomes a trading strategy. The question we want to solve is: What is the minimum initial investment $X_0 = x$ he must have in order to achieve his consumption and investment goal without outside infusion of capital? This minimum, if any, is defined to be the value of the contingent claim C. .

If we do not impose some conditions on the trading strategy π besides adaptability, pathological situations may arise. Roughly speaking, these pathological situations can be avoided if he is not allowed to be wildly in debt. There are various versions of such condition that work well mathematically. Here we take an easy version of it, namely, he is not allowed to be in debt at all.

DEFINITION 1.1. *A trading strategy π is said to be admissible if the wealth process is nonnegative: $X_t = \xi_t B_t + \eta_t S_t \geq 0$.*

We also need to give a mathematical definition of the condition which we phrased as "without outside infusion of capital." Such a portfolio is called self-financing. An arbitrary portfolio $\pi = (\xi_t, \eta_t)$ may need outside capital to maintain its value $X_t = \xi_t B_t + \eta_t S_t$. Suppose that this portfolio consists only of the bond. We have $X_t = \xi_t B_t$. How can we tell that it is self-financing? This is the case precisely when ξ_t is constant, i.e., its increase in value comes solely from the increase of the bond price B_t. Therefore we have

$$dX_t = \xi_t \, dB_t = X_t r_t \, dt.$$

Now if the portfolio has only the stock in it, then $X_t = \eta_t S_t$ and it is self-financing only when η_t is constant, i.e.,

$$dX_t = \eta_t \, dS_t.$$

If the portfolio contains both the stock and the bond, then exchanging between these two are allowed. We make the following definition.

DEFINITION 1.2. *A portfolio (or a trading strategy) with consumption rate c is self-financing if its wealth process $X_t = \xi_t B_t + \eta_t S_t$ satisfies*

$$dX_t = \xi_t \, dB_t + \eta_t \, dS_t - c_t \, dt.$$

Now we can formulate the problem we want to solve as follows. For a given contingent claim $C = (c, G)$, what is the smallest initial capital $X_0 = x$ for that there is an admissible, self-financing trading strategy $\pi = (\xi, \eta)$ such that

$$\mathbb{P}\{X_T \geq C\} = 1.$$

It is intuitively clear that there will be no such admissible strategy if $X_0 = x$ is too small. Our task is to find the smallest x for the existence of an admissible strategy.

We will first assume that there is a trading strategy $\pi = (\xi, \eta)$ to satisfy the contingent claim and derive a lower bound for $X_0 = x$. It is clear that we need to eliminate π from the picture and find a lower bound in terms of the given parameters. We start with the two equations for the wealth process

$$X_t = \xi_t B_t + \eta_t S_t$$

and

$$dX_t = \xi_t \, dB_t + \eta_t \, dS_t - c_t dt.$$

From these two equations we can eliminate ξ_t. We have

$$dB_t = B_t r_t \, dt \quad \text{and} \quad dS_t = S_t(\sigma_t \, dW_t + \mu_t \, dt).$$

Hence

$$\xi_t \, dB_t = \xi_t B_t r_t \, dt = (X_t - \eta_t S_t) r_t \, dt$$

and we have an equation without ξ_t as follows:

$$dX_t - X_t r_t \, dt = \eta_t S_t \{\sigma_t \, dW_t + (\mu_t - r_t) \, dt\} - c_t \, dt.$$

Now we come to the key step in the argument. We need to eliminate η and solve for x. For this purpose we introduce the shifted Brownian motion

$$\tilde{W}_t = W_t - \int_0^t \theta_s \, ds,$$

where

$$\theta_t = \frac{r_t - \mu_t}{\sigma_t}.$$

The equation for X can be written

$$dX_t - r_t X_t \, dt = \eta_t S_t \sigma_t \, d\tilde{W}_t - c_t \, dt.$$

It is clear now that we can do away with η if integrate with respect to a measure under which $\tilde{W}$ is a Brownian motion. This measure is readily provided by the Girsanov theorem. We define the risk-neutral probability measure $\mathbb{P}^*$ by

$$\frac{d\mathbb{P}^*}{d\mathbb{P}} = \exp\left[\int_0^T \theta_t \, dW_t - \frac{1}{2}\int_0^T \theta_t^2 \, dt\right].$$

The process $\tilde{W}$ is a Brownian motion under $\mathbb{Q}$. Using B_t^{-1} as an integrating factor we can solve the equation for X in the form

$$B_t^{-1}X_t - x = M_t - \int_0^t B_s^{-1}c_s\,ds,$$

where

$$M_t = \int_0^t B_s^{-1}\eta_s S_s\sigma_s\,d\tilde{W}_s$$

is a local martingale under $\mathbb{Q}$. The wealth process is given by

$$X_t = xB_t + B_tM_t - B_t\int_0^t B_s^{-1}c_s\,ds.$$

In particular,

$$B_T^{-1}X_T - x = M_T - \int_0^T B_s^{-1}c_s\,ds$$

and

$$x = B_T^{-1}X_T + \int_0^T B_s^{-1}c_s\,ds - M_T.$$

It is clear that M is a local martingale with an integrable lower bound, hence using Fatou's lemma on

$$\mathbb{E}^*M_{\tau_n\wedge T} = 0,$$

we have $\mathbb{E}^*M_T \geq 0$. It follows that

$$x \geq \mathbb{E}^*\left[B_T^{-1}G + \int_0^T B_s^{-1}c_s\,ds\right].$$

Here $\mathbb{E}^*$ is the expectation with respect to the risk-neutral probability $\mathbb{P}^*$.

THEOREM 1.3. *The price of the contingent claim $C = (c, G)$ is given by*

$$V(c,G) = \mathbb{E}^*\left[B_T^{-1}G + \int_0^T B_s^{-1}c_s\,ds\right].$$

PROOF. We have shown that the price cannot be smaller than the quantity on the right side. Denote the right side by x. We need to construct a trading strategy π for our contingent claim. We do this by reversing the argument above. The existence of a trading strategy is achieved by using the martingale representation theorem.

In light of the argument given above, the stock component η of our trading strategy should satisfy the equality

$$\int_0^T B_s^{-1}\eta_s S_s\sigma_s\,d\tilde{W}_s = B_T^{-1}G + \int_0^T B_s^{-1}c_s\,ds - x.$$

By assumption the expected value of the right side is zero. If we let $H_s = \eta_s S_s\sigma_s$, then the above equality can be written as

$$B_T^{-1}G + \int_0^T B_s^{-1}c_s\,ds = x + \int_0^T H_s\,d\tilde{W}_s.$$

This equality looks very much like a martingale representation of the random variable on the left side. However, this random variable is measurable with respect to the filtration of the original Brownian motion W, whereas we want to express it as a stochastic integral with respect to the Brownian motion $\tilde{W}$ under $\mathbb{P}^*$. We know that martingales under $\mathbb{P}$ and $\mathbb{P}^*$ are related in a simple way as follows. Let

$$e_t = \exp\left[\int_0^t \theta_s\, dW_s - \frac{1}{2}\int_0^t \theta_s^2\, ds\right].$$

Then $d\mathbb{P}^*/d\mathbb{P} = e_T$ and M_t is a martingale under $\mathbb{P}^*$ if and only if $M_t e_t$ is a martingale under $\mathbb{P}$. By the martingale representation theorem, there is a unique progressively measurable process H such that

$$\left[B_T^{-1}G + \int_0^T B_s^{-1}c_s\, ds\right] e_T = x + \int_0^T \tilde{H}_s\, dW_s.$$

By Ito's formula it is easy to verify that there exists a progressively measurable process H such that

$$e_T^{-1}\left[x + \int_0^T H_s\, dW_s\right] = x + \int_0^T H_s\, d\tilde{W}_s.$$

In fact we have

$$H_t = \left[\tilde{H}_t - \left(x + \int_0^t \tilde{H}_s\, dW_s\right)\theta_t\right] e_t^{-1}.$$

It follows that

$$B_T^{-1}G + \int_0^T B_s^{-1}c_s\, ds = x + \int_0^T H_s\, d\tilde{W}_s.$$

Having found H by applying the martingale representation theorem, we can proceed by reversing the argument we have carried out before. Define the stock component of the trading strategy by

$$H_t = B_t^{-1}\eta_t S_t \sigma_t, \qquad \eta_t = \frac{B_t H_t}{S_t \sigma_t}.$$

The wealth process is defined by

$$X_t = xB_t + B_t M_t - B_t \int_0^t B_s^{-1} c_s\, ds,$$

where

$$M_t = \int_0^t B_s^{-1}\eta_s S_s \sigma_s\, d\tilde{W}_s.$$

It is clear then the bond component of the trading strategy should be

$$\xi_t = \frac{X_t - \eta_t S_t}{B_t}.$$

Now we have automatically

$$X_t = \xi_t B_t + \eta_t S_t$$

by the definition of ξ. It remains to show that the trading strategy $\pi = (\xi, \eta)$ is

(1) self-financing: $dX_t = \xi_t \, dB_t + \eta_t \, dS_t - c_t \, dt$;

(2) admissible: $X_t \geq 0$;

(3) having the correct payoff: $X_T = G$.

We first show that the trading strategy defined above is self-financing with consumption rate c. From the definition of X we have

$$B_t^{-1} X_t = x - \int_0^t B_s^{-1} c_s ds + \int_0^t B_s^{-1} \eta_s S_s \sigma_s \, d\tilde{W}_s.$$

Recall that previously we obtained this from its equivalent differential form

$$dX_t - X_t r_t \, dt = \eta_t S_t \sigma_t \, d\tilde{W}_t - c_t \, dt.$$

Using $X_t = \xi_t B_t + \eta_t S_t$ and

$$d\tilde{W}_t = dW_t - \left[\frac{r_t - \mu_t}{\sigma_t} \right] dt$$

we have

$$\begin{aligned} dX_t &= (\xi_t B_t + \eta_t S_t) r_t \, dt + \eta_t S_t \sigma_t \, dW_t - \eta_t S_t (r_t - \mu_t) \, dt - c_t \, dt \\ &= \xi_t \, dB_t + \eta_t S_t (\sigma_t \, dW_t + \mu_t \, dt) - c_t \, dt \\ &= \xi_t \, dB_t + \eta_t \, dS_t - c_t \, dt. \end{aligned}$$

This shows that the trading strategy $\pi = (\xi, \eta)$ is self-financing with consumption rate c.

We now show that the trading strategy is admissible and $X_T = G$. Recall that

$$B_t^{-1} X_t = x - \int_0^t B_s^{-1} c_s ds + \int_0^t \eta_s S_s \sigma_s \, d\tilde{W}_s.$$

By the definition of η we also have

$$B_T^{-1} G + \int_0^T B_s^{-1} c_s \, ds = x + \int_0^T B_s^{-1} \eta_s S_s \sigma_s \, d\tilde{W}_s.$$

Hence,

$$\mathbb{E}^* \left[B_T^{-1} G + \int_0^T B_s^{-1} c_s \, ds \middle| \mathscr{F}_t \right] = x + \int_0^t B_t^{-1} \eta_s S_s \sigma_s \, d\tilde{W}_s.$$

It follows that

$$B_t^{-1} X_t = \mathbb{E}^* \left[B_T^{-1} G + \int_t^T B_s^{-1} c_s \, ds \middle| \mathscr{F}_t \right].$$

Therefore the wealth process is given by

$$X_t = B_t \, \mathbb{E}^* \left[B_T^{-1} G + \int_t^T B_s^{-1} c_s \, ds \middle| \mathscr{F}_t \right]. \tag{1.1}$$

It is clear from this expression that $X_t \geq 0$ and $X_T = G$. □

REMARK 1.4. Let's take a closer look at the relation

$$B_t^{-1}X_t + \int_0^t B_s^{-1}c_s ds = x + \int_0^t B_s^{-1}\eta_s S_s \sigma_s \, d\tilde{W}_s.$$

Intuitively, this relation means that if the bond interest rate is equal to the mean stock return rate, which is true under the risk-neutral probability $\mathbb{P}^*$, then sum of the wealth and the total consumption properly discounted by the interest rate is a martingale (with respect to the risk-neutral probability $\mathbb{P}^*$) with mean value x. The local martingale on the right side has the quadratic variation

$$\epsilon_t = \int_0^t B_s^{-2}\eta_s^2 S_s^2 \sigma_s^2 \, ds.$$

Hence the local martingale has the form W_{ϕ_t} for some Brownian motion W and we have

$$W_{\epsilon_t} = B_t^{-1}X_t + \int_0^t B_s^{-1}c_s ds - x.$$

Since $X_t \geq 0$ and we have assumed that the interest rate and the consumption rate are uniformly bounded, there is a constant K such that

$$W_{\epsilon_t} \geq -K, \qquad 0 \leq t \leq T.$$

Now if $|\eta_t|S_t$ is bounded from below by a positive constant for $0 \leq t \leq T$, then $\epsilon_t \geq \epsilon t$ for some positive constant ϵ. This implies that the probability that the local martingale part is less than $-K$ somewhere on $[0, T]$ is strictly positive, which is impossible. Hence the trading strategy must vanish sometime, i.e., the investor has to invest all his asset in the bond.

REMARK 1.5. It is clear from the proof that the trading strategy $\pi = (\xi, \eta)$ is unique. The source of this strong conclusion is the assumption that the trading strategy is self-financing.

2. Black-Scholes equation and formula

We now consider a special case, where the market characteristics involved are functions of the stock price and time:

$$r_t = r(S_t, t), \quad \mu_t = \mu(S_t, t), \quad \sigma_t = \sigma(S_t, t).$$

We also assume that the contingent claim has the form

$$c_t = c(S_t, t), \quad G = G(S_T).$$

This case is generally considered as the Black-Scholes option price theory. We will see that it is intimately connected to the theory of partial differential equations. The price of a contingent claim is a function of time and stock price and it satisfies the so-called Black-Scholes partial differential equation . The case of constant interest rate and volatility and the call option $G(S_T) = (S_T - K)^+$ can be computed explicitly and gives the classical Black-Scholes formula.

The stock price satisfies the stochastic differential equation

$$dS_t = S_t(\sigma(S_t, t)\, dW_t + \mu(S_t, t)\, dt).$$

We assume that the time parameter t covers the range $\mathbb{R}_+ = [0, \infty)$. To create a more flexible theory, we sometimes consider the price history from a starting point s other than 0. In this case, if the initial stock price is $S_s = x$, then the stock price S_t is the solution of the following stochastic differential equation

$$dS_t = S_t(\sigma(S_t, t)\, dW_t + \mu(S_t, t)\, dt), \quad t \geq s; \quad S_s = x.$$

But what we really need is the equation under the risk-neutral probability, i.e.,

$$dS_t = S_t(\sigma(S_t, t)\, dW_t + r(S_t, t)\, dt), \quad t \geq s; \quad S_s = x.$$

We will assume that this equation has a unique solution. This assumption holds if we assume, for example, that the coefficients are globally Lipschitz. For the theory of stochastic differential equations, we know that the process $\{S_t,\ t \geq s\}$ is a strong Markov process. We will denote its law by $\mathbb{P}_{t,x}$.

We have shown that the price of the contingent claim $C = (c, G)$ is given by

$$V^C = \mathbb{E}^* \left[B_T^{-1} G(S_T) + \int_0^T B_s^{-1} c(S_s, s)\, ds \right].$$

Note that the stock price is the solution of the stochastic differential equation

$$dS_t = S_t(\sigma(S_t, t)\, dW_t + \mu(S_t, t)\, dt).$$

This can be written as

$$dS_t = S_t(\sigma(S_t, t)\, d\tilde{W}_t + r(S_t, t)\, dt).$$

Under the risk-neutral probability $\mathbb{P}^*$, the process $\tilde{W}$ is a Brownian motion. This is a diffusion process with the generator

$$\mathscr{L} = \frac{(\sigma S)^2}{2} \frac{\partial^2}{\partial x^2} + rS \frac{\partial}{\partial x}.$$

We also have

$$B_t^{-1} = \exp \left[-\int_0^t r(S_s, s)\, ds \right].$$

Recall that $\mathbb{P}_S = \mathbb{P}_{0,S}$ is the law of the stock price process $\{S_t,\ t \geq 0\}$ under the risk-neutral probability. We can now write

$$V(S; T) = \mathbb{E}_S \left[B_T^{-1} G(S_T) + \int_0^T B_s^{-1} c(S_s, s)\, ds \right].$$

Therefore the formula for V^C is precisely the probabilistic representation of an initial value problem. More precisely, the price $V^C = V(S; T)$ is a function of the expiration time T and the initial stock price S and satisfies the Black-Scholes partial differential equation

$$V_T = \frac{(\sigma S)^2}{2} V_{SS} + rSV_S - rV + c$$

with the initial condition $V(0, S) = G(S)$.

We now consider the price of the contingent claim at a general time $t \leq T$ instead of at time 0. We use $V(t, T; S)$ to denote the price of the contingent claim expiring at time T at the initial time $t \leq T$ when the stock price is S at time t. Thus the old price function is $V(T; S) = V(0, T; S)$. The only difference between the old and the new price function is that now the initial time point is shifted from 0 to t. Therefore the formula for $V(T; S)$ applies to the new price function after an obvious modification, namely,

$$V(t, S; T) = \mathbb{E}_{t,S}\left[B_{t,T}^{-1} G(S_T) + \int_t^T B_{t,s}^{-1} c(S_s, s)\, ds\right],$$

where

$$B_{t,s} = \exp\left[\int_t^s r(S_u, u)\, du\right].$$

As before for a fixed t, the price function $V(t, T; S)$ is the solution of the initial boundary value problem

$$V_T = \frac{(\sigma S)^2}{2} V_{SS} + rSV_S - rV + c, \quad V(t, S; t) = G(S).$$

Similarly, $V(t, S; T)$, for a fixed expiration time T, is the solution of the following terminal boundary value problem:

$$V_t + \frac{(\sigma S)^2}{2} V_{SS} + rSV_s - rV + c = 0, \quad V(T, S; S) = 0.$$

This is similar to the equation of $V(t, S; T)$ if t is fixed, the difference being the sign of the time derivative and the change from a initial value problem to a terminal value problem. The above equation is also referred to as the Black-Scholes equation. In general diffusion theory, this is an example of backward diffusion equations.

We now express the unique trading strategy $\pi_t = (\xi_t, \eta_t)$ in terms of the price function $V(t, T; S)$. Suppose that we are given a contingent claim as above. If we start with $X_0 = V(0, S_0)$, the value of the contingent claim at time 0 if the stock price is S_0. Then we expect that at time time t, the wealth process X_t should be equal to the value of the contingent claim when the stock price is S_t, i.e.,

$$X_t = V(t, S_t).$$

THEOREM 2.1. *For the trading strategy constructed in* THEOREM *1.3 we have* $X_t = V(t, S_t)$ *and*

$$\begin{aligned} \xi_t &= B_t^{-1}\left\{V(t, S_t) - S_t V_S(t, S_t)\right\}, \\ \eta_t &= V_S(t, S_t). \end{aligned}$$

PROOF. We have shown that the wealth process is given by

$$X_t = B_t\,\mathbb{E}^*\left[B_T^{-1}G(S_T) + \int_t^T B_s^{-1}c(S_s,s)\,ds\middle|\mathscr{F}_t\right].$$

See (1.1). By the Markov property of the stock process and $B_T = B_t B_{t,T}$, we can write X_t can be written as

$$X_t = \mathbb{E}_{t,S_t}\left[B_{t,T}^{-1}G(S_T) + \int_t^T B_{t,s}^{-1}c(S_s,s)\,ds\right].$$

This shows that

$$X_t = V(t,S_t).$$

Calculating the stochastic differential of X_t we have

$$\begin{aligned} dX_t &= V_t\,dt + V_S\,dS_t + \frac{1}{2}V_{SS}(\sigma S)^2\,dt \\ &= \left[V_t + \frac{(\sigma S)^2}{2}V_{SS}\right]\,dt + V_S\,dS_t. \end{aligned}$$

Comparing this with

$$\begin{aligned} dX_t &= \xi_t\,dB_t + \eta_t\,dS_t - c(S_t,t)\,dt \\ &= \{r(S_t,t)B_t\xi_t - c(S_t,t)\}\,dt + \eta_t\,dS_t \end{aligned}$$

we have immediately $\eta_t = V_S(t,S_t)$. Using the Black-Scholes equation we have

$$V_t + \frac{(\sigma S)^2}{2}V_{SS} = rV - rSV_S - c.$$

Hence

$$r(S_t,t)B_t\xi_t - c(S_t,t) = r(S_t,t)V(t,S_t) - r(S_t,t)S_tV_S(t,S_t) - c(S_t,t).$$

The formula for ξ_t follows immediately. □

We now prove the classical Black-Scholes formula for a call option.

PROPOSITION 2.2. *The call price is given by*

$$C(S;T) = SN(d_1) - Ke^{-rT}N(d_2),$$

where

$$N(z) = \frac{1}{\sqrt{2\pi}}\int_\infty^z e^{-y^2/2}dy$$

is the distribution of the standard normal distribution and

$$d_1 = \frac{\log(S/K) + \left(r + \dfrac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}}$$

and

$$d_2 = d_1 - \sigma\sqrt{T} = \frac{\log(S/K) + \left(r - \dfrac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}}$$

PROOF. Under the risk-neutral probability $\mathbb{P}^*$, the stock price has the lognormal distribution. More precisely, $\ln S_T$ is a normal random variable with mean $\left(r - \frac{\sigma^2}{2}\right) T + \ln S$ and standard deviation $\sigma\sqrt{T}$. Therefore

$$\ln S_T = \sigma\sqrt{T}X + \left(r - \frac{\sigma^2}{2}\right) T + \ln S,$$

where X is a standard normal random variable. The call payoff is $(S_T - K)^+$, which is positive if

$$\sigma\sqrt{T}X + \left(r - \frac{\sigma^2}{2}\right) T + \ln S \geq \ln K, \quad \text{or} \quad X \geq -d_2.$$

It follows that

$$\begin{aligned}
C(S;T) &= e^{-rT}\mathbb{E}\left\{\exp\left[\left(r - \frac{\sigma^2}{2}\right) T + \ln S + \sigma\sqrt{T}X\right]; X \geq -d_2\right\} \\
&\quad - e^{-rT}K\mathbb{P}\left\{X \geq -d_2\right\} \\
&= S\,\mathbb{E}\left\{\exp\left[\sigma\sqrt{T}X - \frac{\sigma^2}{2}T\right]; X - \sigma\sqrt{T} \geq -d_2 - \sigma\sqrt{T}\right\} \\
&\quad - Ke^{-rT}\left\{1 - N(-d_2)\right\} \\
&= S\,\mathbb{P}\left\{X \geq -d_1\right\} - Ke^{-rT}N(d_2) \\
&= SN(d_1) - Ke^{-rT}N(d_2).
\end{aligned}$$

This completes the proof. □

3. Sixth assignment