Elementary Introduction to Stochastic Calculus

A thesis submitted by

Ken Noël Wada-Shiozaki

In partial fulfillment for the degree of

Master of Science
in
Mathematics

TUFTS UNIVERSITY

August 2011

©2011, Ken N. Wada-Shiozaki

Advisor: Marjorie G. Hahn

Abstract

This paper attempts to provide an elementary background for stochastic calculus via construction of the Itô integrals. In addition, the Itô formula, an essential tool in the study of stochastic calculus, is proved in detail for the Brownian integrators. A relatively rough proof of the Itô formula in the case of continuous square integrable martingales is also discussed. Finally, Lévy's characterization theorem is proved as one of the most significant consequences of the Itô formula.

Acknowledgment

My first thanks goes to my advisor Professor Marjorie Hahn who has been generously patient throughout my often-troublesome thesis process. I also thank Professor Sabir Umarov for agreeing to be on my thesis committee despite the unusual circumstance.

Finally, I would like to thank Tufts University and all of its members, including faculty, staff, and students, for providing me with three years of an unparalleled experience.

Contents

Elementary Introduction to Stochastic Calculus

The development and careful construction and analysis of probability theory, lead mainly by the Russian school and the French school, allowed an interesting extension of the notion of "calculus." In particular, stochastic calculus (mainly of the Itô form) has contributed in many areas, most notably in development of the Black-Scholes equations. Despite its celebration throughout academic and non-academically oriented communities, the full generalization of stochastic calculus involves a deep understanding of functional analysis including abstract harmonic analysis, operator algebras, and functional spaces. They are involved in very subtle yet significant ways in the development and construction of the stochastic calculus.

In this paper, the author's aim is to introduce how Itô's integration is constructively defined, and the significance of the "Brownian integrators." From the view point of acquiring the fundamental mathematics, the methodology presented in this paper is rather inadequate and deceiving. In order to keep the focus of the paper simple, the author had to make an unfortunate choice of minimizing the explicit use of functional analysis in the construction. The approach presented in this paper is similar to that appeared in **[6]**.

After the preliminary section, the readers are introduced to how K. Itô had developed his version of stochastic integration; first with the Brownian integrators, and then using a special class of martingales as the integrators. In general, the integrators need not be continuous in order to define the Itô integral. The topics that are covered in this paper are only meant to be the most basic introduction to the theory of stochastic calculus. Because of the complexity of the material, even in just defining Itô's integral, the author is not going to present other important ideas such as stochastic differential equations.

1. Preliminaries

Throughout this paper, $(\Omega, \mathcal{S}, \mathsf{P})$ is used to denote the default probability space, and $\mathcal{B}(\mathbb{R}) =: \mathcal{B}$ is used to denote the Borel σ-algebra of $\mathbb{R}$.

Definition 1.0.1. For $\mathtt{I} \subset \mathbb{R}$, a ***stochastic process*** is a map $X\colon \mathtt{I} \times \Omega \to \mathbb{R}$ such that

(i) $X(t, \cdot\,)$ is a random variable for every $t \in \mathtt{I}$;

(ii) $X(\,\cdot\,, \omega)$ is a $\mathcal{B}$-measurable function for every $\omega \in \Omega$.

For each fixed $\omega \in \Omega$, a function $X(\,\cdot\,, \omega)$ is called a ***sample path*** of X.

Although the range space can be further generalized to a separable topological space with the topology given by a complete metric, the author only considers $\mathbb{R}$ for simplicity. For typographical convenience, X_t is sometimes used to denote $X(t, \cdot\,)$ (and $X(t, \omega)$ when the suppressed argument ω is clear in the context).

Definition 1.0.2. A ***filtration*** $\{\mathcal{F}_t\}_{t\in\mathtt{I}}$ is an increasing sequence of σ-algebras. A stochastic process $\{X_t\}_{t\in\mathtt{I}}$ is said to be ***adapted*** to $\{\mathcal{F}_t\}_{t\in\mathtt{I}}$ if the random variable X_t is $\mathcal{F}_t$-measurable for every $t \in \mathtt{I}$. A filtration $\{\mathcal{F}_t\}$ is said to satisfy ***the usual conditions*** if both of the following conditions are satisfied:

(i) (***completeness***) $\cap_{t\in \mathtt{I}}\mathcal{F}_t$ contains all P-null sets;

(ii) (***right-continuity***) $\mathcal{F}_t = \cap_{s>t}\mathcal{F}_s$ for $s, t \in \mathtt{I}$.

In order to avoid the tenacious technical problems, all the filtrations introduced in this paper are assumed to satisfy the usual conditions, unless specified otherwise.

Definition 1.0.3. A stochastic process $\widetilde{X}$ is said to be a ***realization*** of X if there exists a subset $\widetilde{\Omega} \subset \Omega$ such that $\mathsf{P}\big(\widetilde{\Omega}\big) = 1$ and, for each $\tilde{\omega} \in \widetilde{\Omega}$, $\widetilde{X}(t, \tilde{\omega}) = X(t, \tilde{\omega})$ for all $t \in \mathtt{I}$.

Definition 1.0.4. A stochastic process X_t is said to be a ***continuous process*** (or a.s.-continuous process) if the sample paths of X_t are continuous for ω-a.s.

Definition 1.0.5. Let $\{X_t\}_{t\in\mathtt{I}}$ be a stochastic process adapted to a filtration $\{\mathcal{F}_t\}_{t\in\mathtt{I}}$ with $\mathbb{E}\big(|X_t|\big) < \infty$ for all $t \in \mathtt{I}$, and satisfying either of the following conditions (not necessarily exclusive): for $s \le t$,

(i) $\mathbb{E}\big(X_t \mid \mathcal{F}_s\big) \ge X_s$ a.s.

(ii) $\mathbb{E}\big(X_t \mid \mathcal{F}_s\big) \le X_s$ a.s.

If X satisfies (i) then X is called a ***submartingale*** with respect to $\{\mathcal{F}_t\}_{t\in\mathtt{I}}$, and if the condition (ii) is satisfied then X is called a ***supermartingale*** with respect to $\{\mathcal{F}_t\}_{t\in\mathtt{I}}$. When X satisfies both (i) and (ii), X is said to be a ***martingale*** with respect to $\{\mathcal{F}_t\}_{t\in\mathtt{I}}$. The acronym a.s. is reserved exclusively to mean almost everywhere in Ω. When necessary, the acronym a.e. will be used for other circumstances instead.

1.1. **Basics of Brownian Motion.**

Definition 1.1.1. A ***Brownian motion*** (or ***Wiener process***) $B(t, \omega)$ is a stochastic process satisfying the following conditions:

(i) $\mathsf{P}\big\{\omega \in \Omega\colon B(0, \omega) = 0\big\} = 1$;

(ii) for any $0 \le s < t$, the increment $(B_t - B_s)$ has a normal distribution with mean 0 and variance $(t - s)$: for $a < b$

$$\mathsf{P}\big\{a \le B_t - B_s \le b\big\} = \frac{1}{\sqrt{2\pi(t-s)}} \int_a^b \exp\Big\{-\frac{x^2}{2(t-s)}\Big\}\, dx;$$

(iii) for $0 \le t_0 < t_1 < t_2 < \cdots < t_n$, the random variables $(B_{t_1} - B_{t_0}), \ldots, (B_{t_n} - B_{t_{n-1}})$ are independent;

(iv) $\mathsf{P}\big\{\omega \in \Omega\colon B(\,\cdot\,, \omega)$ is continuous$\big\} = 1$.

Proposition 1.1.1. *Let B be a Brownian motion (in $\mathbb{R}^1$). Then the following statements hold:*

(a) for $s, t \in \mathbb{R}^+$, $\mathbb{E}\big(B_s\, B_t\big) = s \wedge t$;

(b) for a fixed $s_0 \ge 0$, the process given by $\widetilde{B}_t := B_{t+s_0} - B_{s_0}$ is also a Brownian motion;

(c) for $\alpha > 0$, the process $\widehat{B}_t := \alpha^{-1/2} B_{\alpha t}$ is also a Brownian motion.

Proof.

(a) Assume that $s < t$. Then from Definition 1.1.1(ii) and (iii), the following computation establishes the result:

$$\mathbb{E}\big(B_s\, B_t\big) = \mathbb{E}\big(B_s(B_t - B_s) + B_s^2\big) = \mathbb{E}\big(B_s(B_t - B_s)\big) + \mathbb{E}\big(B_s^2\big) = 0 + s = s.$$

(b) It suffices to verify that conditions (i)-(iv) in Definition 1.1.1 are satisfied for $\widetilde{B}_t$. Trivially, $\widetilde{B}_t$ satisfies (i) and (iv). Since $\mathbb{E}(B_x) = 0$ for all $x \geq 0$, $\mathbb{E}(\widetilde{B}_t - \widetilde{B}_{t_0}) = \mathbb{E}(B_{t+s_0}) - \mathbb{E}(B_{t_0+s_0}) = 0$ for any $0 \leq t_0 < t$. Assuming $0 \leq s < t$, the result from (a) yields the variance of $(\widetilde{B}_t - \widetilde{B}_s)$:

$$\begin{aligned}\mathbb{E}\big((\widetilde{B}_t - \widetilde{B}_s)^2\big) &= \mathbb{E}(B_{t+s_0}^2) - 2\,\mathbb{E}(B_{t+s_0}\, B_{s+s_0}) + \mathbb{E}(B_{s+s_0}^2)\\ &= (t + s_0) - 2(s + s_0) + (s + s_0)\\ &= t - s.\end{aligned}$$

Hence, $\widetilde{B}_t$ satisfies (ii). For a given partition $0 \leq t_0 < t_1 < \cdots < t_n$, assuming $s_0 > 0$, consider the shifted partition $s_0 \leq (t_0 + s_0) < (t_1 + s_0) < (t_2 + s_0) < \cdots < (t_n + s_0)$. By the condition (iii) of B_t, it follows that the increments $\big\{B_{t_k+s_0} - B_{t_{k-1}+s_0} = \widetilde{B}_{t_k} - \widetilde{B}_{t_{k-1}} : 0 \leq k \leq n\big\}$ are independent. Thus, $\widetilde{B}_t$ is a Brownian motion.

(c) Conditions (i), (iii), and (iv) in Definition 1.1.1 easily follow from B_t. To verify the condition (ii), note that for any $0 \leq s < t$, $\widehat{B}_t - \widehat{B}_s = \alpha^{-1/2}(B_{\alpha t} - B_{\alpha s})$. Hence, $\mathbb{E}(\widehat{B}_t - \widehat{B}_s) = 0$ and $\mathbb{E}\big((\widehat{B}_t - \widehat{B}_s)^2\big) = \alpha^{-1}(\alpha t - \alpha s) = t - s$; therefore, $\widehat{B}_t$ is a Brownian motion.

□

Proposition 1.1.2. *Let B_t be a Brownian motion, and define a* natural filtration $\mathcal{F}_t^B := \sigma\{B_s : s \leq t\}$, *for $t \in [0, \infty)$. Then, B_t is a martingale with respect to $\{\mathcal{F}_t^B\}$.*

Proof. First, one should note here that because (Ω, P) is a finite measure space, $L^2(\mathsf{P}) \subset L^1(\mathsf{P})$ as a consequence of Jensen's inequality. Thus, $\mathbb{E}(|B_t|) < \infty$. Let $0 \leq s \leq t$. By a standard property of conditional expectations, $\mathbb{E}(B_s \mid \mathcal{F}_s^B) = B_s$ since B_s is clearly $\mathcal{F}_s^B$-measurable. By the condition (iii) in Definition 1.1.1, $(B_t - B_s)$ and B_s are independent; so it follows that $\mathbb{E}(B_t - B_s \mid \mathcal{F}_s^B) = \mathbb{E}(B_t - B_s)$. By the condition (ii), $\mathbb{E}(B_t - B_s) = 0$. Putting together the results,

$$\mathbb{E}(B_t \mid \mathcal{F}_s^B) = \mathbb{E}(B_t - B_s \mid \mathcal{F}_s^B) + \mathbb{E}(B_s \mid \mathcal{F}_s^B) = 0 + B_s = B_s.$$

Hence, B_t is a martingale with respect to $\{\mathcal{F}_t^B\}$. □

The computation of $\mathbb{E}(B_t^{2n})$, for $n \in \mathbb{N}$, will be useful later. Setting $0 \le s < t$, the random variables $(B_t - B_s)$ and $B_1(t-s)^{1/2}$ both have a normal distribution with mean 0 and variance $(t-s)$. Note that, for every $n \in \mathbb{N}$,

$$\mathbb{E}\big((B_t - B_s)^{2n}\big) = \mathbb{E}\big((B_1(t-s)^{1/2})^{2n}\big) = \mathbb{E}(B_1^{2n})(t-s)^n = \frac{(2n)!}{2^n n!}(t-s)^n. \tag{1.1}$$

1.2. **Integration Revisited.**

The aim of this section is to briefly summarize the basic facts of integration that are necessary for the development of elementary stochastic integration theory. Let $[a,b]$ be a finite closed interval with $a < b$, $\Delta_n := \{a = t_0, t_1, \ldots, t_{n-1}, t_n = b\}$ be a refining partition of $[a,b]$ such that $t_i < t_j$ for $0 \le i < j \le n$, and $\|\Delta_n\| := \max_{1 \le i \le n}(t_i - t_{i-1})$. A bounded function f defined on $[a,b]$ is said to be *Riemann integrable* if the following limit exists:

$$\int_a^b f(t)\,dt = \lim_{\|\Delta_n\| \to 0} \sum_{i=1}^n f(t_i^*)(t_i - t_{i-1}),$$

where t_j^* is some evaluation point in $[t_{j-1}, t_j]$ for every $1 \le j \le n$. Each bounded function on $[a,b]$ is Riemann integrable if and only if it is almost-everywhere continuous with respect to Lebesgue measure.

With the same notation as in the previous paragraph, let g be a monotonically increasing function defined on the same interval $[a,b]$. A function f is *Riemann-Stieltjes integrable* with respect to g if the limit

$$\int_a^b f(t)\,dg(t) = \lim_{\|\Delta_n\| \to 0} \sum_{i=1}^n f(t_i^*)\big(g(t_i) - g(t_{i-1})\big) \tag{1.2}$$

exists. By a well-known fact from basic real analysis, every continuous function (defined on $[a,b]$) is Riemann-Stieltjes integrable with respect to any monotonically increasing function on $[a,b]$. Now let $\tilde{f}$ be a continuous monotonically increasing function and $\tilde{g}$ be a continuous function (both defined on $[a,b]$). One can then apply integration by parts to define

$$\int_a^b \tilde{f}(t)\,d\tilde{g}(t) := \tilde{f}(t)\tilde{g}(t)\Big]_a^b - \int_a^b \tilde{g}(t)\,d\tilde{f}(t),$$

where the right-side integral is evaluated as an R-S integral of $\tilde{g}$ with respect to $\tilde{f}$.

Example 1.2.1. The two cases presented below illustrate when an integral can be expressed as a R-S integral in (1.2).

Let $f\colon [a,b] \to \mathbb{R}$ be such that $f' := \frac{df}{dt}$ is a continuous function (i.e. $f \in C^1$). For each $n \in \mathbb{N}$, define

$$(1.3) \qquad L_n f := \sum_{i=1}^{n} f(t_{i-1})\big(f(t_i) - f(t_{i-1})\big) \quad \text{and} \quad R_n f := \sum_{i=1}^{n} f(t_i)\big(f(t_i) - f(t_{i-1})\big).$$

Denote $\|\cdot\|_\infty$ to be the usual supremum norm. By the mean-value theorem, one can deduce that

$$(1.4) \qquad |R_n f - L_n f| = \sum_{i=1}^{n} \big(f'(t_i^*)\,(t_i - t_{i-1})\big)^2 \le \|f'\|_\infty^2 \sum_{i=1}^{n} (t_i - t_{i-1})^2 \le \|f'\|_\infty^2 \,\|\Delta_n\| \sum_{i=1}^{n} (t_i - t_{i-1})$$
$$= \|f'\|_\infty^2 \|\Delta_n\| (b-a),$$

where $t_{i-1}^* \in (t_{i-1}, t_i)$. Then as $\|\Delta_n\| \to 0$, the upper-bound converges to 0. Simple arithmetic manipulations of the equations in (1.3) yield that

$$R_n f = \frac{1}{2}\Big(f(b)^2 - f(a)^2 + \sum_{i=1}^{n} \big(f(t_i) - f(t_{i-1})\big)^2\Big),$$
$$L_n f = \frac{1}{2}\Big(f(b)^2 - f(a)^2 - \sum_{i=1}^{n} \big(f(t_i) - f(t_{i-1})\big)^2\Big).$$

Since the inequality (1.4) shows that $\sum_{i=1}^{n} \big(f(t_i) - f(t_{i-1})\big)^2$ converges to 0 as $\|\Delta_n\| \to 0$, it follows from the above new expressions that

$$\lim_{\|\Delta_n\| \to 0} L_n f = \frac{f(b)^2 - f(a)^2}{2} = \lim_{\|\Delta_n\| \to 0} R_n f.$$

Since $f \in C^1$, one may set the integrator $df(t) := f'(t)dt$ to define the integral $\int_a^b f(t)\, df(t)$ as a R-S integral:

$$\int_a^b f(t)\, df(t) := \int_a^b f(t) f'(t)\, dt = \frac{f(b)^2 - f(a)^2}{2}.$$

On the other hand, let g to be a continuous function such that

$$|g(t) - g(s)| \approx |t - s|^{1/2}.$$

Define $L_n g$ and $R_n g$ as in (1.3). Then it easily follows that

$$0 \le R_n g - L_n g = \sum_{i=1}^{n} \big(g(t_i) - g(t_{i-1})\big)^2 \approx \sum_{i=1}^{n} |t_i - t_{i-1}| = |b-a|.$$

Since $a \ne b$, $\lim_{\|\Delta_n\| \to 0} R_n g \ne \lim_{\|\Delta_n\| \to 0} L_n g$. Hence, the limit in (1.2) fails to exist and $\int_a^b g(t)\, dg(t)$ cannot be defined as a R-S integral.

2. Itô's Integral

Let $0 \leq a < b < +\infty$. Throughout this section, let $\{B_t\}_{t\in[a,b]}$ be a Brownian motion together with a filtration $\{\mathcal{F}_t\}_{t\in[a,b]}$ satisfying the usual conditions.

The aim in this section is to constructively define a stochastic integral $\int_a^b f(B_s)dB_s$. For application purposes, one would like to define the integral so that the stochastic process $\left\{\int_a^t f(B_s)\, dB_s\,;\, t \in [a,b]\right\}$ is a (continuous) martingale with respect to $\{\mathcal{F}_t\}$, for some suitable choice of a process f. This is a motivation of importance in the study of stochastic differential equations. The primary goal of this section is to present how one can derive such a nice definition. An essential tool here is *quadratic variation*.

Proposition 2.0.1. *Let* $\Delta_n := \{a = t_0, t_1, \ldots, t_{n-1}, t_n = b\}$ *be a partition of a finite closed interval* $[a,b]$*. Then, as* $\|\Delta_n\| := \max_{1\leq i\leq n}|t_i - t_{i-1}| \to 0$,

$$\sum_{i=1}^{n}(B_{t_i} - B_{t_{i-1}})^2 \xrightarrow{L^2(\mathsf{P})} b - a. \tag{2.1}$$

Proof. By the definition of Brownian motion and the equation (1.1) in Section 1, $\mathbb{E}\big((B_t - B_s)^2\big) = (t-s)$ and $\mathbb{E}\big((B_t - B_s)^4\big) = 3(t-s)^2$. Define

$$\sum_{i=1}^{n} X_i := \sum_{i=1}^{n}\Big((B_{t_i} - B_{t_{i-1}})^2 - (t_i - t_{i-1})\Big). \tag{2.2}$$

For $i \neq j$, since B has independent increments, it is clear that $\mathbb{E}(X_i\, X_j) = 0$. For $i = j$, one can compute that

$$\begin{aligned} \mathbb{E}(X_i^2) &= \mathbb{E}\Big((B_{t_i} - B_{t_{i-1}})^4 - 2(t_i - t_{i-1})\,(B_{t_i} - B_{t_{i-1}})^2 + (t_i - t_{i-1})^2\Big) \\ &= 3(t_i - t_{i-1})^2 - 2(t_i - t_{i-1})^2 + (t_i - t_{i-1})^2 \\ &= 2(t_i - t_{i-1})^2. \end{aligned} \tag{2.3}$$

Now using the fact that $(b-a) = \sum_{i=1}^{n}(t_i - t_{i-1})$, as $\|\Delta_n\| \to 0$,

$$\mathbb{E}\Big(\sum_{i,j=1}^{n} X_i\, X_j\Big) = \sum_{i=1}^{n} 2(t_i - t_{i-1})^2 \leq 2\|\Delta_n\| \sum_{i=1}^{n}(t_i - t_{i-1}) = 2\,\|\Delta_n\|(b-a) \longrightarrow 0.$$

Hence, by the above calculations, one can see that, as $\|\Delta_n\| \to 0$, the sum (2.2) converges to 0 in $L^2(\mathsf{P})$, which shows the desired result indeed holds. □

Remark. An important fact to mention here is that, if a sequence of partitions $\{\Delta_n\}$ is increasing (i.e. $\Delta_0 \subset \Delta_1 \subset \cdots \subset \Delta_n \subset \cdots$), then the convergence in (2.1) holds almost surely. For a full proof of this fact, see Theorem 2.4 and its corollary in **[2]**.

Definition 2.0.1. The ***quadratic variation*** of a stochastic process X is defined by

$$[X]_t := \lim_{\|\Delta_n\|\to 0} \sum_{i=1}^{n} \big(X_{t_i} - X_{t_{i-1}}\big)^2, \tag{2.4}$$

whenever the limit exists in probability.

Example 1.2.1 in the previous section shows that the stochastic integral $\int_a^b B_t\, dB_t$ cannot be defined as a R-S integral. However, there is a remarkably useful observation to be made. For $\{\Delta_n\}$ a sequence of partitions of $[a,b] \subset \mathbb{R}^+$, define

$$L_n B := \sum_{i=1}^{n} B_{t_{i-1}}\big(B_{t_i} - B_{t_{i-1}}\big) \quad \text{and} \quad R_n B := \sum_{i=1}^{n} B_{t_i}\big(B_{t_i} - B_{t_{i-1}}\big).$$

With this expression and Proposition 2.0.1, it follows that

$$\begin{aligned} R_nB + L_nB &= \sum_{i=1}^{n}\big(B_{t_i}^2 - B_{t_{i-1}}^2\big) = B_b^2 - B_a^2, \\ R_nB - L_nB &= \sum_{i=1}^{n}\big(B_{t_i} - B_{t_{i-1}}\big)^2 \xrightarrow{L^2(\mathsf{P})} b - a, \quad \text{as } \|\Delta_n\| \to 0. \end{aligned}$$

By simple manipulations,

$$\lim_{\|\Delta_n\|\to 0} L_nB = \frac{1}{2}\Big(B_b^2 - B_a^2 - (b-a)\Big) \quad \text{and} \quad \lim_{\|\Delta_n\|\to 0} R_nB = \frac{1}{2}\Big(B_b^2 - B_a^2 + (b-a)\Big),$$

where both limits are defined in $L^2(\mathsf{P})$. Set $a = 0$ and $b = t$. Using the above limits, define the stochastic processes

$$L_B(t) := \frac{B_t^2 - t}{2} \quad \text{and} \quad R_B(t) := \frac{B_t^2 + t}{2}.$$

It is clear that $R_B(t)$ is not a martingale since $\mathbb{E}\big(R_B(t)\big) = t$ is not a constant. On the other hand, for any $s \le t$,

$$\mathbb{E}\big(L_B(t)\big|\mathcal{F}_s\big) = \frac{1}{2}\Big(\mathbb{E}\big(B_t^2\big|\mathcal{F}_s\big) - t\Big). \tag{2.5}$$

To compute the conditional expectation on the right-hand side above, note that, for all $u \le s$, $(B_t - B_s)$ and B_u are independent. Hence, $(B_t - B_s)$ and $\mathcal{F}_s$ are independent. Then it follows that

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

Plugging the result back into the original equation (2.5),

$$\mathbb{E}\big(L_B(t)\big|\mathcal{F}_s\big) = L_B(s)$$

for all $s \leq t$. Therefore, $L_B(t)$ is a martingale. The fact that $L_B(t)$ is a martingale suggests that if the integral is evaluated at the left-end-points, the desirable martingale property arises.

2.1. **From Step processes.**

Let $L^2_{adpt}\big([a,b] \times \Omega\big)$ denote the space of all stochastic processes satisfying the following conditions: for all $t \in [a,b]$ and $\omega \in \Omega$,

(i) $f(t,\omega)$ is an $\{\mathcal{F}_t\}$-adapted stochastic process;
(ii) $\int_a^b \mathbb{E}\big(|f(t,\omega)|^2\big)\,dt < \infty$.
(iii) For each $\omega \in \Omega$, $f(t,\omega)$ is left-continuous.

The focus of this section is to define the stochastic integral $\int_a^b f(t)\,dB_t$ for $f \in L^2_{adpt}\big([a,b] \times \Omega\big)$.

To begin, first investigate the restricted subset of $L^2_{adpt}\big([a,b] \times \Omega\big)$ which consists of *step stochastic processes.*

Definition 2.1.1. If a process $f \in L^2_{adpt}\big([a,b] \times \Omega\big)$ can be written as

$$f(t,\omega) := \sum_{i=1}^{n} \varphi_{i-1}(\omega)\mathbb{I}_{[t_{i-1},\,t_i)}(t),$$

where φ_{i-1} is $\mathcal{F}_{t_{i-1}}$-measurable and $\mathbb{E}(\varphi^2_{i-1}) < \infty$, then f is called a ***step stochastic process***.

For step stochastic processes, define a map $\mathcal{I}_{step}$ given by

$$\mathcal{I}_{step}(f) := \sum_{i=1}^{m} \varphi_{i-1}(B_{t_i} - B_{t_{i-1}}).$$

Observe here that $\mathcal{I}_{step}(\alpha f + \beta g) = \alpha \mathcal{I}_{step}(f) + \beta \mathcal{I}_{step}(g)$ for every $\alpha, \beta \in \mathbb{R}$ and step stochastic processes f and g. Furthermore, by taking the expectation of $\mathcal{I}_{step}\big(\varphi_{i-1}\,(B_{t_i} - B_{t_{i-1}})\big)$ conditional on $\mathcal{F}_{t_{i-1}}$ for every $1 \leq i \leq m$, it follows that $\mathbb{E}\big(\mathcal{I}_{step}(f)\big) = 0$. On the other hand, the following cases gives the computation of $\mathbb{E}\big(|f(t)|^2\big)$

(1) For $i < j$,

$$\begin{aligned} &\mathbb{E}\Big(\varphi_{i-1}\varphi_{j-1}\,\mathbb{E}\big((B_{t_i} - B_{t_{i-1}})\,(B_{t_j} - B_{t_{j-1}})\big|\mathcal{F}_{t_{j-1}}\big)\Big) \\ =\; &\mathbb{E}\Big(\varphi_{i-1}\varphi_{j-1}\,(B_{t_i} - B_{t_{i-1}})\,\mathbb{E}\big(B_{t_j} - B_{t_{j-1}}\big|\mathcal{F}_{t_{j-1}}\big)\Big) = 0. \end{aligned}$$

(2) For $i = j$,

$$\mathbb{E}\Big(\varphi_{i-1}^2 (B_{t_i} - B_{t_{i-1}})^2\Big) = \mathbb{E}\Big(\varphi_{i-1}^2\, \mathbb{E}\big((B_{t_i} - B_{t_{i-1}})^2 \big| \mathcal{F}_{t_{i-1}}\big)\Big) = \mathbb{E}\Big(\varphi_{i-1}^2 \mathbb{E}\big((B_{t_i} - B_{t_{i-1}})^2\big)\Big)$$
$$= (t_i - t_{i-1})\, \mathbb{E}(\varphi_{i-1}^2).$$

Moreover, notice that $|\mathcal{I}_{step}(f)|^2$ can be expressed as

$$|\mathcal{I}_{step}(f)|^2 = \sum_{i,j=1}^{m} \varphi_{i-1}\varphi_{j-1}\, (B_{t_i} - B_{t_{i-1}})\, (B_{t_j} - B_{t_{j-1}}),$$

and combining the above results shows that

$$\mathbb{E}\big(|\mathcal{I}_{step}(f)|^2\big) = \int_a^b \mathbb{E}\big(|f(t)|^2\big)\, dt. \tag{2.6}$$

2.2. **Approximation by Step processes.**

The idea presented here is critical in properly defining the stochastic integral for a general stochastic process $f \in L^2_{adpt}([a,b] \times \Omega)$. In a nutshell, step stochastic processes are dense in $L^2_{adpt}([a,b] \times \Omega)$ so that one can use (with appropriate modifications) approximation by step processes for general elements in $L^2_{adpt}([a,b] \times \Omega)$. Before stating the desired proposition, two lemmas need to be proved.

Lemma 2.2.1. *Let $f \in L^2_{adpt}([a,b] \times \Omega)$ be bounded, and define a sequence of stochastic processes $\{g_n\}$ by*

$$g_n(t,\omega) := \int_0^{n(t-a)} f\big(t - \tfrac{\tau}{n}, \omega\big) e^{-\tau}\, d\tau, \qquad \text{for } n \in \mathbb{N}.$$

Then the following hold:

(a) $\mathbb{E}\big(g_n(t)\, g_n(s)\big)$ *is a continuous function of* (t,s) *for all* $n \in \mathbb{N}$.
(b) $\lim_{n\to\infty} \int_a^b \mathbb{E}\big(|f(t) - g_n(t)|^2\big) = 0$.

Proof. First notice that, for all $n \in \mathbb{N}$, g_n is $\{\mathcal{F}_t\}$-adapted and $\int_a^b \mathbb{E}\big(|g_n(t,\omega)|^2\big)\, dt < \infty$. To prove (a), one only needs to rewrite g_n by putting $u := t - \frac{\tau}{n}$ to obtain

$$g_n(t,\omega) = \int_a^t f(u,\omega)\, ne^{-n(t-u)}\, du.$$

With the expression above, it follows that $\lim_{t\to s} \mathbb{E}\big(|g_n(t) - g_n(s)|^2\big) = 0$.

To prove (b), one can set $f(t) \equiv 0$ for all $t < a$. Then

$$f(t) - g_n(t) = \int_0^\infty \big(f(t) - f(t - \tfrac{\tau}{n})\big) e^{-\tau}\, d\tau.$$

Note that $e^{-\tau}\,d\tau$ is a probability measure on $[0,\infty)$. So applying the Schwarz inequality yields

$$|f(t)-g_n(t)|^2 \le \int_0^\infty |f(t)-f(t-\tfrac{\tau}{n})|^2 e^{-\tau}\,d\tau.$$

Furthermore,

$$\begin{aligned}\int_a^b \mathbb{E}\big(|f(t)-g_n(t)|^2\big)\,dt &\le \int_a^b\int_0^\infty \mathbb{E}\big(|f(t)-f(t-\tfrac{\tau}{n})|^2\big)e^{-\tau}\,d\tau\,dt\\ &= \int_0^\infty \left\{\int_a^b \mathbb{E}\big(|f(t)-f(t-\tfrac{\tau}{n})|^2\big)\,dt\right\} e^{-\tau}\,d\tau\\ &= \int_0^\infty \mathbb{E}\bigg(\int_a^b |f(t)-f(t-\tfrac{\tau}{n})|^2\,dt\bigg)\,e^{-\tau}\,d\tau.\end{aligned}$$

Because f is bounded, as $n\to\infty$,

$$\int_a^b |f(t)-f(t-\tfrac{\tau}{n})|^2\,dt \xrightarrow{\text{a.s.}} 0.$$

Thus, the conclusion of (b) follows. □

Lemma 2.2.2. *Let $f\in L^2_{adpt}([a,b]\times\Omega)$, and assume that $\mathbb{E}\big(\,f(t)f(s)\,\big)$ is a continuous function of $(t,s)\in[a,b]\times[a,b]$. Then there exists a sequence of stochastic processes $\{f_n\}\subset L^2_{adpt}\big([a,b]\times\Omega\big)$ such that, as $n\to\infty$,*

$$\int_a^b \mathbb{E}\big(|f(t)-f_n(t)|^2\big)\,dt \longrightarrow 0.$$

Proof. As before, let $\Delta_n := \{a=t_0,t_1,\ldots,t_{n-1},t_n=b\}$ be a partition of $[a,b]$ with $\|\Delta_n\|\to 0$. For each $n\in\mathbb{N}$, define $f_n(t,\omega):=f(t_{i-1},\omega)$, where $t\in(t_{i-1},t_i]$; so $\{f_n\}$ is a sequence of $\{\mathcal{F}_t\}$-adapted step processes. By the continuity of $\mathbb{E}\big(\,f(t)f(s)\,\big)$, for each $t\in[a,b]$, $\lim_{n\to\infty}\mathbb{E}\big(|f(t)-f_n(t)|^2\big)=0$. Now use the well-known inequality $|x\pm y|^2\le 2(|x|^2+|y|^2)$ for $x,y\in\mathbb{R}$, to deduce that

$$|f(t)-f_n(t)|^2 \le 2\big(|f(t)|^2+|f_n(t)|^2\big).$$

One can then compute the bound for $\mathbb{E}\big(|f(t)-f_n(t)|^2\big)$ as follows: for all $t\in[a,b]$,

$$\mathbb{E}\big(|f(t)-f_n(t)|^2\big)\le 2\Big(\mathbb{E}\big(|f(t)|^2\big)+\mathbb{E}\big(|f_n(t)|^2\big)\Big)\le 4\sup_{s\in[a,b]}\Big\{\mathbb{E}\big(|f(s)|^2\big)\Big\}.$$

From the work shown above, the dominated convergence theorem shows that the conclusion holds. □

Proposition 2.2.1. *For $f \in L^2_{adpt}([a,b]\times\Omega)$, there exists a sequence of step stochastic processes $\{f_n(t)\}_{n\in\mathbb{N}}$ in $L^2_{adpt}([a,b]\times\Omega)$ such that*

$$\lim_{n\to\infty}\int_a^b \mathbb{E}\big(|f(t)-f_n(t)|^2\big)\,dt = 0.$$

Proof. First assume that $f \in L^2_{adpt}\big([a,b]\times\Omega\big)$ is bounded. By Lemma 2.2.1(a) and Lemma 2.2.2 applied to g_n, for each $n\in\mathbb{N}$ there exists an adapted step process $f_n(t,\omega)$ such that

$$\int_a^b \mathbb{E}\big(|g_n(t)-f_n(t)|^2\big)\,dt \le \frac{1}{n}.$$

Applying Lemma 2.2.1(b) together with the above inequality yield

$$\lim_{n\to\infty}\int_a^b \mathbb{E}\big(|f(t)-f_n(t)|^2\big)\,dt = 0.$$

Now consider a general $f \in L^2_{adpt}\big([a,b]\times\Omega\big)$. For each $n\in\mathbb{N}$, let g_n be an $\{\mathcal{F}_t\}$-adapted, bounded process given by

$$g_n(t,\omega) := \begin{cases} f(t,\omega) & \text{for } |f(t,\omega)|\le n,\\ 0 & \text{for } |f(t,\omega)| > n.\end{cases}$$

So by the dominated convergence theorem,

$$\lim_{n\to\infty}\int_a^b \mathbb{E}\big(|f(t)-g_n(t)|^2\big)\,dt = 0.$$

Then, by applying the result from the bounded case to g_n (for each $n\in\mathbb{N}$), one can pick an $\{\mathcal{F}_t\}$-adapted step process $f_n(t,\omega)$ such that

$$\int_a^b \mathbb{E}\big(|g_n(t)-f_n(t)|^2\big)\,dt \le \frac{1}{n}.$$

The statement is therefore proved. □

By Proposition 2.2.1, every $f \in L^2_{adpt}\big([a,b]\times\Omega\big)$ can be approximated, in $L^2(\mathsf{P})$, by a sequence of $\{\mathcal{F}_t\}$-adapted step stochastic processes $\{f_n\}$. With the results thus far, it is now possible to provide a suitable prototypical definition for the *Itô integral*.

Definition 2.2.1. Let $f \in L^2_{adpt}\big([a,b]\times\Omega\big)$ and $\{f_n\}$ be an approximating step sequence of f. The map $\mathcal{I}(f) := \lim_{n\to\infty}\mathcal{I}_{step}(f_n)$ is called the ***Itô integral*** of f.

From Definition 2.2.1, it is not clear whether the Itô integral defined above is independent of the choice of the sequence $\{f_n\}$. For completeness, it is important to verify that $\mathcal{I}(f)$ is well-defined. The argument here can be used again to show the well-definedness of the Itô integrals defined in the later sections.

Let $f \in L^2_{\text{adpt}}([a,b]\times\Omega)$ be given and $\{f_n\}$ be a sequence of step processes converging to f in $L^2(\mathsf{P})$. Using the $\mathbb{R}$-linearity of $\mathcal{I}$ and the equation (2.6)

$$\mathbb{E}\big(|\mathcal{I}_{\text{step}}(f_n) - \mathcal{I}_{step}(f_k)|^2\big) = \int_a^b \mathbb{E}\big(|f_n(t) - f_k(t)|^2\big)\,dt \longrightarrow 0, \quad \text{as } n,k \uparrow \infty,$$

which shows that $\{\mathcal{I}_{step}(f_n)\}$ is a Cauchy sequence in $L^2(\mathsf{P})$. Since $L^2(\mathsf{P})$ is a Hilbert space, if there exists another step sequence $\{g_m\}$ converging to f in $L^2(\mathsf{P})$, then $\lim_{m\to\infty}\mathcal{I}_{step}(g_m) = \mathcal{I}(f) = \lim_{n\to\infty}\mathcal{I}_{\text{step}}(f_n)$. So the choice of approximating sequence is irrelevant in the above definition; hence $\mathcal{I}(f)$ is well-defined. The well-definedness of $\mathcal{I}$ gives rise to the following analogue of (2.6).

Theorem 2.2.1. *The Itô Isometry*

Let $f \in L^2_{adpt}([a,b]\times\Omega)$. Denote the Itô integral by $\mathcal{I}(f) := \int_a^b f(t)\,dB(t)$. Then

$$\mathbb{E}\big(|\mathcal{I}(f)|^2\big) = \int_a^b \mathbb{E}\big(|f(t)|^2\big)\,dt.$$

2.3. **Verification of the Properties.**

Assume $f \in L^2_{adpt}\big([a,b]\times\Omega\big)$. For any $t \in [a,b]$, it is easy to see that $f \in L^2_{adpt}\big([a,t]\times\Omega\big)$ since

$$\int_a^t \mathbb{E}\big(|f_s|^2\big)\,ds \le \int_a^b \mathbb{E}\big(|f_s|^2\big)\,ds < \infty.$$

A stochastic process X_t associated with the integral is defined by

$$X_t := \int_a^t f_s\,dB_s, \quad t \in [a,b].$$

By the Itô isometry,

$$\mathbb{E}\big(|X_t|^2\big) = \int_a^t \mathbb{E}\big(|f_s|^2\big)\,ds < \infty.$$

So $\mathbb{E}\big(|X_t|\big) \le \big(\mathbb{E}\big(|X_t|^2\big)\big)^{1/2} < \infty$. Since X_t is integrable for every $t \in [a,b]$, conditional expectation of X_t with respect to a σ-algebra $\mathcal{F}_u$ from the filtration $\{\mathcal{F}_t\}_{t\in[a,b]}$ is defined. The next theorem verifies that the martingale property is satisfied.

Theorem 2.3.1. *Let $f \in L^2_{adpt}\big([a,b]\times\Omega\big)$, then the stochastic process X_t given by*

$$X_t := \int_a^t f_s\,dB_s, \quad \text{for} \quad t \in [a,b],$$

is a martingale with respect to the filtration $\{\mathcal{F}_t\}$.

Proof. For the moment, let f be a step stochastic process given by

$$f(u,\omega) := \sum_{i=1}^{n} \varphi_{i-1}(\omega)\, \mathbb{I}_{[t_{i-1},t_i)}(u),$$

where $a \le s = t_0 < t_1 < \cdots < t_n = t \le b$, and each φ_{i-1} is an $\mathcal{F}_{t_{i-1}}$-measurable function such that $\varphi_{t_{i-1}} \in L^2(\mathsf{P})$. From the definition of X, $X_t - X_s = \int_s^t f(u)\, dB_u$ where $a \le s < t \le b$. The integral in the equation can be written as

$$\int_s^t f(u)\, dB_u = \sum_{i=1}^{n} \varphi_{i-1}(B_{t_i} - B_{t_{i-1}}).$$

For every $i = 1, \dots, n$, the towering property of conditional expectations implies that

$$\begin{aligned}\mathbb{E}\big(\varphi_{i-1}(B_{t_i} - B_{t_{i-1}})\big|\mathcal{F}_s\big) &= \mathbb{E}\big(\mathbb{E}\big(\varphi_{i-1}(B_{t_i} - B_{t_{i-1}})\big|\mathcal{F}_{t_{i-1}}\big)\big|\mathcal{F}_s\big)\\ &= \mathbb{E}\big(\varphi_{i-1}\mathbb{E}\big(B_{t_i} - B_{t_{i-1}}\big|\mathcal{F}_{t_{i-1}}\big)\big|\mathcal{F}_s\big) = 0,\end{aligned}$$

where the last equality follows from $\mathbb{E}\big(B_{t_i} - B_{t_{i-1}} \mid \mathcal{F}_{t_{i-1}}\big) = 0$. Hence

$$\mathbb{E}\left(\int_s^t f(u)\, dB_u \middle| \mathcal{F}_s\right) = \mathbb{E}\big(X_t - X_s \mid \mathcal{F}_s\big) = 0.$$

Therefore, the statement is valid whenever f is a step stochastic process.

Now let f be a general element in $L^2_{adpt}\big([a,b]\times\Omega\big)$ and $\{f_n\}$ be a sequence of step stochastic processes in $L^2_{adpt}\big([a,b]\times\Omega\big)$ such that

$$\lim_{n\to\infty} \int_a^b \mathbb{E}\big(|f(u) - f_n(u)|^2\big)\, du = 0.$$

For each $n \in \mathbb{N}$, define a process $X_t^{(n)} := \int_a^t f_n(u)\, dB_u$. From the result above and the fact that f_n is a step process, $X_t^{(n)}$ must be a martingale. For $a \le s < t \le b$, the expectation of $X_t - X_s$ conditioned on $\mathcal{F}_s$ gives

$$\begin{aligned}\mathbb{E}\big(X_t - X_s\big|\mathcal{F}_t\big) &= \mathbb{E}\big(X_t - X_t^{(n)}\big|\mathcal{F}_s\big) + \mathbb{E}\big(X_t^{(n)} - X_s^{(n)}\big|\mathcal{F}_s\big) + \mathbb{E}\big(X_s^{(n)} - X_s\big|\mathcal{F}_s\big)\\ &= \mathbb{E}\big(X_t - X_t^{(n)}\big|\mathcal{F}_s\big) + \mathbb{E}\big(X_s^{(n)} - X_s\big|\mathcal{F}_s\big).\end{aligned}$$

Note here that

$$\mathbb{E}\big(|\mathbb{E}\big(X_t - X_t^{(n)} \mid \mathcal{F}_s\big)|^2\big) \le \mathbb{E}\big(\mathbb{E}\big(|X_t - X_t^{(n)}|^2\big|\mathcal{F}_s\big)\big) = \mathbb{E}\big(|X_t - X_t^{(n)}|^2\big). \tag{2.7}$$

Furthermore, applying the Itô isometry yields

$$\mathbb{E}\big(|X_t - X_t^{(n)}|^2\big) \le \int_a^t \mathbb{E}\big(|f(u) - f_n(u)|^2\big)\, du \le \int_a^b \mathbb{E}\big(|f(u) - f_n(u)|^2\big)\, du \longrightarrow 0, \quad \text{as } n \to \infty.$$

Combining the above result with the equation (2.7) (and taking a subsequence if necessary), $\mathbb{E}\big(X_t - X_t^{(n)} \big| \mathcal{F}_s\big)$ a.s.-converges 0. Likewise, $\mathbb{E}\big(X_s - X_s^{(n)} \big| \mathcal{F}_s\big)$ also converges almost surely to 0. Thus, $\mathbb{E}\big(X_t - X_s \mid \mathcal{F}_s\big)$ is almost surely equal to 0, thereby proving the statement. □

Finally, one needs to verify the continuity condition which is presented below.

Theorem 2.3.2. *Let $f \in L^2_{adpt}\big([a,b] \times \Omega\big)$, then the stochastic process*

$$X_t := \int_a^t f(s)\, dB_s, \qquad \text{for } t \in [a,b],$$

is continuous.

Proof. As usual, assume first that f is a step process given by

$$f(s,\omega) := \sum_{i=1}^{n} \varphi_{i-1}(\omega) \mathbb{I}_{(t_{i-1},t_i]}(s),$$

where φ_{i-1} is an $\mathcal{F}_{t_{i-1}}$-measurable function. For each $\omega \in \Omega$, the sample path of X_t is given by

$$X_t(\omega) = \sum_{i=1}^{k-1} \varphi_{i-1}(\omega)\big(B(t_i,\omega) - B(t_{i-1},\omega)\big) + \varphi_{k-1}(\omega)\big(B(t,\omega) - B(t_{k-1},\omega)\big),$$

for $t_{k-1} \le t < t_k$. Since B_t is a.s.-continuous on $[a,b]$, X_t must also be a.s.-continuous on $[a,b]$.

Now for $f \in L^2_{adpt}\big([a,b] \times \Omega\big)$, let $\{f_n\}$ be a sequence of step processes in $L^2_{adpt}\big([a,b] \times \Omega\big)$ such that

$$\lim_{n \to \infty} \int_a^b \mathbb{E}\big(|f(s) - f_n(s)|^2\big)\, ds = 0.$$

Upon taking a subsequence if necessary, one may assume without loss of generality that, for every $n \in \mathbb{N}$,

$$\int_a^b \mathbb{E}\big(|f(s) - f_n(s)|^2\big)\, ds \le n^{-6}.$$

For each n, let

$$X_t^{(n)} := \int_a^t f_n(s)\, dB_s, \quad t \in [a,b].$$

From the step process case, one can see that $X_t^{(n)}$ is continuous almost surely. X_t and $X_t^{(n)}$ are martingales by the previous theorem, so $X_t - X_t^{(n)}$ is also a martingale. For each $n \in \mathbb{N}$, let A_n be a subset of Ω such that

$$A_n := \Big\{ \sup_{t \in [a,b]} |X_t - X_t^{(n)}| \ge n^{-1} \Big\}.$$

By Doob's submartingale inequality, $\mathsf{P}\big(A_n\big) \leq n\,\mathbb{E}\big(|X_b - X_b^{(n)}|\big)$. To further evaluate the upper-bound,

$$\begin{aligned}\Big(\mathbb{E}\big(|X_b - X_b^{(n)}|\big)\Big)^2 &\leq \mathbb{E}\big(|X_b - X_b^{(n)}|^2\big) \quad \text{(the Schwarz inequality)}\\ &= \int_a^b \mathbb{E}\big(|f(s) - f_n(s)|^2\big)\, ds \quad \text{(the Itô isometry)}\\ &\leq n^{-6}.\end{aligned}$$

So it follows that $\mathsf{P}\big(A_n\big) \leq n^{-2}$ holds for all $n \in \mathbb{N}$. Since $\sum_{i=1}^{\infty} \mathsf{P}\big(A_n\big) \leq \sum_{i=1}^{\infty} n^{-2} < \infty$, by the Borel-Cantelli lemma, $\mathsf{P}\big(A_n \text{ finitely often}\big) = 1$. This shows the existence of $\Omega_0 \subset \Omega$ such that $\mathsf{P}\big(\Omega_0\big) = 1$ and for each $\omega_0 \in \Omega_0$, there exists $N_{\omega_0} \in \mathbb{N}$ satisfying

$$\sup_{t\in[a,b]} \big|X_t(\omega_0) - X_t^{(n)}(\omega_0)\big| < n^{-1}, \quad \text{for } n \geq N_{\omega_0}.$$

Thus for all $\omega_0 \in \Omega_0$, the sequence of functions $\{X_{(\cdot)}^{(n)}(\omega_0)\}_{n\in\mathbb{N}}$ converges uniformly to $X_{(\cdot)}(\omega_0)$ on $[a,b]$. Moreover, the process $X_t^{(n)}$ is continuous for every $n \in \mathbb{N}$, and hence there exists a subset Ω_n such that $\mathsf{P}\big(\Omega_n\big) = 1$ and the function $X_{(\cdot)}^{(n)}(\omega_n)$ is continuous for all $\omega_n \in \Omega_n$. Setting $\widetilde{\Omega} := \cup_{m=0}^{\infty}\Omega_m$, it follows that $\mathsf{P}\big(\widetilde{\Omega}\big) = 1$ and, for $\tilde{\omega} \in \widetilde{\Omega}$, the sequence $\{X_{(\cdot)}^{(n)}(\tilde{\omega})\}_{n\in\mathbb{N}}$ converges uniformly to $X_{(\cdot)}(\tilde{\omega})$ on $[a,b]$: hence, $X_{(\cdot)}(\tilde{\omega})$ is a continuous function for every $\tilde{\omega} \in \widetilde{\Omega}$. Thus, X_t is a continuous process. □

3. Extension of Integrands

Having established the definition of the Itô integral for $L^2_{adpt}\big([a,b]\times\Omega\big)$, it is now appropriate to discuss how to extend the Itô integration to processes that are not in $L^2_{adpt}\big([a,b]\times\Omega\big)$ in a manner consistent with the previous definition. To illustrate the limitation of the previous definition of the Itô integral, consider the following example.

Example 3.0.1. Define the process $f(t,\omega) := \exp\{\big(B_t(\omega)\big)^2\}$. One can deduce from a simple computation that

$$\mathbb{E}\big(|f_t|^2\big) = \mathbb{E}\Big(\exp\{2B_t^2\}\Big) = \begin{cases}(1-4t)^{-1/2} & \text{on } t \in [0, \frac{1}{4})\\ \infty & \text{on } t \geq \frac{1}{4}.\end{cases}$$

Directly from the computation, it is clear that $\int_0^1 \mathbb{E}\big(|f_t|^2\big)\, dt = \infty$; hence, $f \notin L^2_{adpt}\big([0,1]\times\Omega\big)$, and the stochastic integral $\int_0^1 \exp\{B_t^2\}\, dB_t$ cannot be defined as done in the previous section.

Motivated by the above example, define $\mathfrak{L}^2_{adpt}\big([a,b]\times\Omega\big)$ to be the space of all stochastic processes $g(t,\omega)$ satisfying the following conditions:

(i) g is $\{\mathcal{F}_t\}$-adapted,

(ii) $\mathsf{P}\Big\{\omega\colon \int_a^b |g(t,\omega)|^2\,dt < \infty\Big\} = 1.$

Set $[a,b] := [0,1]$ and $f(t,\omega) := \exp\{\big(B_t(\omega)\big)^2\}$ as in the example. Since $f(t,\omega)$ is continuous with respect to t, it follows that $\int_0^1 |f(t,\omega)|^2\,dt < \infty$ holds almost surely.; hence, $f \in \mathfrak{L}^2_{adpt}\big([0,1]\times\Omega\big)$.

Without going through an explicit calculation, it is not easy to determine whether a given $\{\mathcal{F}_t\}$-adapted process belongs to $L^2_{adpt}\big([a,b]\times\Omega\big)$. On the other hand, there is a relatively easy sufficient criterion for elements of $\mathfrak{L}^2_{adpt}\big([a,b]\times\Omega\big)$: namely, if an $\{\mathcal{F}_t\}$-adapted process is a.s.-continuous, then it is an element of $\mathfrak{L}^2_{adpt}\big([a,b]\times\Omega\big)$.

The basic approach to define the Itô integral of $\mathfrak{L}^2_{adpt}\big([a,b]\times\Omega\big)$ is quite similar to the approach used in the case for $L^2_{adpt}\big([a,b]\times\Omega\big)$. A key fact is that, for a given process $f \in \mathfrak{L}^2_{adpt}\big([a,b]\times\Omega\big)$, there exists a sequence of step processes $\{f_n\}$ in $L^2_{adpt}\big([a,b]\times\Omega\big)$ such that

$$\lim_{n\to\infty} \int_a^b |f - f_n|^2\,dB_t = 0, \tag{3.1}$$

where the limit is defined in probability. Let $\mathcal{I}$ be the Itô integral defined in Section 2, so that $\mathcal{I}\big(f_n\big) := \int_a^b f_n(t)\,dB_t$ for every $n \in \mathbb{N}$. For a given $\varepsilon > 0$, there exists $N_\varepsilon \in \mathbb{N}$ such that

$$\mathsf{P}\Big(\big\{|\mathcal{I}(f_n) - \mathcal{I}(f_m)| > \varepsilon\big\}\Big) < \varepsilon, \quad \forall\, n, m \geq N_\varepsilon. \tag{3.2}$$

By the convergence of $\big\{\mathcal{I}(f_n)\big\}_{n\in\mathbb{N}}$ in probability, the stochastic integral of $f \in \mathfrak{L}^2_{adpt}\big([a,b]\times\Omega\big)$ can be defined as

$$\int_a^b f(t)\,dB_t := \lim_{n\to\infty} \mathcal{I}(f_n), \tag{3.3}$$

in probability. By an argument similar to that used in Section 2, the above limit is well-defined.

Assume now that $f \in \mathfrak{L}^2_{adpt}\big([a,b]\times\Omega\big)$ is a.s.-continuous. For a partition Δ_n, define a sequence $\{f_n\}$ by

$$f_n(t) := \sum_{i=1}^{n} f(t_{i-1})\,\mathbb{I}_{[t_{i-1},t_i)}(t) \qquad \text{for } t \in [a,b]. \tag{3.4}$$

By the a.s.-continuity of f,

$$\lim_{n\to\infty} \int_a^b |f_n(t) - f(t)|^2\,dt = 0$$

holds almost surely, hence in probability as well. Using the equations (3.1), (3.2), and (3.3), it can be shown that, as $\|\Delta_n\| \to 0$,

$$\mathcal{I}(f_n) := \int_a^b f_n(t)\,dB_t \longrightarrow \int_a^b f(t)\,dB_t, \quad \text{in probability.}$$

It follows from the definition of $\mathcal{I}$ that, as $\|\Delta_n\| \to 0$,

$$\int_a^b f_n(t)\,dB_t = \sum_{i=1}^{n} f(t_{i-1})\big(B_{t_i} - B_{t_{i-1}}\big) \longrightarrow \int_a^b f(t)\,dB_t, \quad \text{in probability.} \tag{3.5}$$

4. Martingale Integrators

Just as done previously with the Brownian integrators, there exists a generalization of the Itô integral with respect to non-Brownian integrators. The full generalization is beyond the scope of this paper. This section considers the extension to a continuous, square integrable martingale.

Let M_t be a continuous, square integrable martingale with respect to $\{\mathcal{F}_t\}$. Though its proof is omitted here, one can assume that M_t^2 has the following unique decomposition:

$$M_t^2 = \ell_t + \langle M\rangle_t, \quad t \in [a,b], \tag{4.1}$$

where ℓ_t is a martingale with respect to $\{\mathcal{F}_t\}$, and $\langle M\rangle_t$ is a non-negative, increasing process such that $\langle M\rangle_a = 0$ and $\mathbb{E}\big(\langle M\rangle_t\big) < \infty$ for all $t \in [a,b]$. More detailed discussion is presented in **[9]**.

Let $L^2_{\langle M\rangle}\big([a,b]\times\Omega\big)$ be a class of $\{\mathcal{F}_t\}$-adapted processes such that

$$\mathbb{E}\Big(\int_a^b |f(t,\omega)|^2\, d\langle M\rangle_t(\omega)\Big) < \infty.$$

The aim here is to define the integral $\int_a^b f(t)\,dM_t$.

Proposition 4.0.1. *For $a \le s < t \le b$,*

$$\mathbb{E}\big((M_t - M_s)^2 \big| \mathcal{F}_s\big) = \mathbb{E}\big(\langle M\rangle_t - \langle M\rangle_s \big| \mathcal{F}_s\big).$$

Proof.

$$\begin{aligned}
\mathbb{E}\big((M_t - M_s)^2 \big| \mathcal{F}_s\big) &= \mathbb{E}\big(M_t^2 - 2M_tM_s + M_s^2 \big| \mathcal{F}_s\big) = \mathbb{E}\big(M_t^2 \big| \mathcal{F}_s\big) - M_s^2 \\
&= \mathbb{E}\big(\ell_t + \langle M\rangle_t \big| \mathcal{F}_s\big) - \big(\ell_s + \langle M\rangle_s\big) \qquad \text{by (4.1)} \\
&= \mathbb{E}\big(\langle M\rangle_t - \langle M\rangle_s \big| \mathcal{F}_s\big)
\end{aligned}$$

□

Following the procedures similar to those used in Section 2, let f be a step stochastic process given by

$$f(t,\omega) := \sum_{i=1}^{n} \xi_{i-1}(\omega)\mathbb{I}_{[t_{i-1},t_i)}(t),$$

and define $\mathfrak{I}_{step}(f)$ by

$$\mathfrak{I}_{step}(f) := \sum_{i=1}^{n} \xi_{i-1}(M_{t_i} - M_{t_{i-1}}). \tag{4.2}$$

Using Proposition 4.0.1, one can establish the following analogue of the equation (2.6) in Section 2.

Proposition 4.0.2. *Let* $\mathfrak{I}_{step}(f)$ *be as defined in* (4.2). *Then*

$$\mathbb{E}\big(|\mathfrak{I}_{step}(f)|^2\big) = \mathbb{E}\Big(\int_a^b |f(t)|^2 \, d\langle M\rangle_t\Big).$$

Proof. Since M_t is a martingale, $\mathbb{E}\big(M_{t_i} - M_{t_{i-1}} \big| \mathcal{F}_{t_{i-1}}\big) = 0$. Hence, for $i < j$, it follows that

$$\begin{aligned}
&\mathbb{E}\Big(\xi_{i-1}\big(M_{t_i} - M_{t_{i-1}}\big)\xi_{j-1}\big(M_{t_j} - M_{t_{j-1}}\big)\Big) \\
&= \mathbb{E}\Big(\xi_{i-1}\xi_{j-1}\big(M_{t_i} - M_{t_{i-1}}\big)\mathbb{E}\big(M_{t_j} - M_{t_{j-1}} \big| \mathcal{F}_{t_{j-1}}\big)\Big) = 0.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\mathbb{E}\big(|\mathfrak{I}_{step}(f)|^2\big) &= \sum_{i,j=1}^{n} \mathbb{E}\Big(\xi_{i-1}\big(M_{t_i} - M_{t_{i-1}}\big)\xi_{j-1}\big(M_{t_j} - M_{t_{j-1}}\big)\Big) \\
&= \sum_{i=1}^{n} \mathbb{E}\Big(\xi_{i-1}^2\big(M_{t_i} - M_{t_{i-1}}\big)^2\Big) \\
&= \sum_{i=1}^{n} \mathbb{E}\Big(\xi_{i-1}^2 \mathbb{E}\big(\big(M_{t_i} - M_{t_{i-1}}\big)^2 \big| \mathcal{F}_{t_{i-1}}\big)\Big).
\end{aligned}$$

Then, applying Proposition 4.0.1 yields

$$\begin{aligned}
\mathbb{E}\big(|\mathfrak{I}_{step}(f)|^2\big) &= \sum_{i=1}^{n} \mathbb{E}\big(\xi_{i-1}^2 \mathbb{E}(\langle M\rangle_{t_i} - \langle M\rangle_{t_{i-1}} \mid \mathcal{F}_{t_{i-1}})\big) \\
&= \sum_{i=1}^{n} \mathbb{E}\big(\xi_{i-1}^2(\langle M\rangle_{t_i} - \langle M\rangle_{t_{i-1}})\big) \\
&= \mathbb{E}\Big(\sum_{i=1}^{n} \xi_{i-1}^2(\langle M\rangle_{t_i} - \langle M\rangle_{t_{i-1}})\Big) \\
&= \mathbb{E}\Big(\int_a^b |f(t)|^2 \, d\langle M\rangle_t\Big).
\end{aligned}$$

□

Using the same argument as in Section 2, a general process $f \in L^2_{\langle M\rangle}\big([a,b] \times \Omega\big)$ can be approximated by a sequence of step processes $\{f_n\}$ so that

$$\lim_{n\to\infty} \mathbb{E}\Big(\int_a^b |f(t) - f_n(t)|^2 \, d\langle M\rangle_t\Big) = 0.$$

Thus, by Proposition 4.0.2, the sequence $\{\mathfrak{I}_{step}(f_n)\}$ is Cauchy in $L^2(\mathsf{P})$, and

$$\mathfrak{I}(f) := \lim_{n\to\infty} \mathfrak{I}_{step}(f_n) := \int_a^b f(t,\omega)\, dM_t, \tag{4.3}$$

as a limit in $L^2(\mathsf{P})$, defines the Itô integral of f with respect to M_t. By modifying the proof of Theorem 2.3.1 in Section 2, one can show that the following theorem holds.

Theorem 4.0.3. *Let $f \in L^2_{\langle M\rangle}\big([a,b]\times\Omega\big)$. Then the stochastic process X_t given by*

$$X_t := \int_a^t f_s\, dM_s, \quad t \in [a,b],$$

is a martingale.

5. The Itô Formula

In the ordinary calculus, the chain rule for differentiation gives a way to compute the derivative of a composite function. More specifically, if two functions f and g are differentiable, then the composite $f\big(g(t)\big)$ is differentiable with the derivative $\frac{d}{dt} f\big(g(t)\big) = f'\big(g(t)\big)g'(t)$. Alternatively, by the fundamental theorem of calculus, it can be re-expressed as

$$f\big(g(t)\big) - f\big(g(a)\big) = \int_a^t f'\big(g(s)\big)g'(s)\, ds.$$

On the other hand, attempting to establish the corresponding rule for the stochastic calculus encounters certain difficulties; in particular, almost all sample paths of B_t are nowhere differentiable (see **[4]**).

5.1. The Simplest Itô Formula.

The following two lemmas are critical pieces for proving the Itô formula.

Lemma 5.1.1. *Let g be a continuous function defined on $\mathbb{R}$ and $\Delta_n := \{a = t_0, t_1, \ldots, t_{n-1}, t_n = t\}$ be a refining partition of the interval $[a,t]$, for every $n \in \mathbb{N}$. Suppose that $\lambda_i \in (0,1)$ for every $1 \le i \le n$. Then, the sequence*

$$X_n := \sum_{i=1}^n \Big(g\big(B_{t_{i-1}} + \lambda_i(B_{t_i} - B_{t_{i-1}})\big) - g\big(B_{t_{i-1}}\big)\Big)\big(B_{t_i} - B_{t_{i-1}}\big)^2$$

contains a subsequence that, as $\|\Delta_n\| \to 0$, almost surely converges to 0.

Proof. For every $n \in \mathbb{N}$, let ξ_n be a random variable given by

$$\xi_n := \sup_{\substack{1\le i\le n\\ \lambda\in(0,1)}} \Big| g\big(B_{t_{i-1}} + \lambda(B_{t_i} - B_{t_{i-1}})\big) - g\big(B_{t_{i-1}}\big)\Big|.$$

Then,

$$|X_n| \leq \xi_n \sum_{i=1}^{n} (B_{t_i} - B_{t_{i-1}})^2.$$

From Proposition 2.0.1 from Section 2, the summation in the above inequality converges to $(t-a)$ in $L^2(\mathsf{P})$. Since the partitions $\{\Delta_n\}$ form a refining sequence, the summation converges a.s. to $(t-a)$. Moreover, by the continuity of $g(x)$ and B_t, ξ_n a.s.-converges to 0. So the right-hand side of the inequality has a subsequence that a.s.-converges to 0. Therefore, X_n converges almost surely to 0. □

Lemma 5.1.2. *Let g and Δ_n be the same as in the previous lemma. Then, as $\|\Delta_n\| \to 0$, the sequence*

$$S_n := \sum_{i=1}^{n} g\big(B_{t_{i-1}}\big)\Big((B_{t_i} - B_{t_{i-1}})^2 - (t_i - t_{i-1})\Big)$$

converges to 0 in probability.

Proof. Let $L > 0$ be given. For $n \in \mathbb{N}$, set

$$A_{i-1}^{(L)} := \big\{|B_{t_j}| \leq L \quad \text{for all } j \leq (i-1)\big\},$$

where $1 \leq i \leq n$. Then define a random variable $S_{(n,L)}$ by

$$\begin{aligned} S_{(n,L)} :&= \sum_{i=1}^{n} \mathbb{I}_{A_{i-1}^{(L)}} g\big(B_{t_{i-1}}\big)\Big(\big(B_{t_i} - B_{t_{i-1}}\big)^2 - (t_i - t_{i-1})\Big) \\ &= \sum_{i=1}^{n} \mathbb{I}_{A_{i-1}^{(L)}} g\big(B_{t_{i-1}}\big)Y_i = \sum_{i=1}^{n} Z_i, \end{aligned}$$

where $Y_i := \big(B_{t_i} - B_{t_{i-1}}\big)^2 - (t_i - t_{i-1})$ and $Z_i := \mathbb{I}_{A_{i-1}^{(L)}} g\big(B_{t_{i-1}}\big)Y_i$.

Now consider the filtration $\mathcal{F}_t^B := \sigma\{B_s; s \leq t\}$. Assume $i < j$. Since $(B_{t_i} - B_{t_{i-1}})$ is independent of $\mathcal{F}_{t_{i-1}}^B$, $\mathbb{E}\big((B_{t_i} - B_{t_{i-1}})^2 \big| \mathcal{F}_{t_{i-1}}^B\big) = \mathbb{E}\big((B_{t_i} - B_{t_{i-1}})^2\big) = t_i - t_{i-1}$.

$$\begin{aligned} \mathbb{E}\big(Z_i Z_j\big) = \mathbb{E}\Big(\mathbb{E}\big(Z_i Z_j \mid \mathcal{F}_{t_{j-1}}^B\big)\Big) &= \mathbb{E}\Big(Z_i \mathbb{I}_{A_{j-1}^{(L)}} g(B_{t_{j-1}}) \mathbb{E}\big(Y_j \mid \mathcal{F}_{t_{j-1}}^B\big)\Big) \\ &= \mathbb{E}\Big(Z_i \mathbb{I}_{A_{j-1}^{(L)}} g(B_{t_{j-1}}) \, \Big(\mathbb{E}\big((B_{t_j} - B_{t_{j-1}})^2\big) - (t_j - t_{j-1})\Big)\Big) = 0. \end{aligned}$$

From the definition of $A_{i-1}^{(L)}$, $Z_i^2 \leq \sup_{|x| \leq L}\{|g(x)|^2\} Y_i^2$. By the equation (2.3) in the proof of Proposition 2.0.1, $\mathbb{E}(Y_i^2) = 2(t_i - t_{i-1})^2$. Then it follows that

$$\begin{aligned} \mathbb{E}\big(S_{(n,L)}^2\big) = \sum_{i=1}^{n} \mathbb{E}\big(Z_i^2\big) &\leq \sup_{|x| \leq L} \big\{|g(x)|^2\big\} 2 \sum_{i=1}^{n} (t_i - t_{i-1})^2 \\ &\leq \sup_{|x| \leq L} \big\{|g(x)|^2\big\} 2 \|\Delta_n\| (t-a), \end{aligned}$$

which converges to 0 as $\|\mathsf{\Delta}_n\| \to 0$. Thus, for any $L > 0$, the sequence $\{S_{(n,L)}\}_{n\in\mathbb{N}}$ converges to 0 in $L^2(\mathsf{P})$ as $\|\mathsf{\Delta}_n\| \to 0$, hence in probability as well.

From their definitions, one can deduce that $\{A_m^{(L)}\}_{0\le m\le n-1}$ is decreasing and

$$\Big\{\sup_{s\in[a,t]}|B_s| \le L\Big\} \subset A_{n-1}^{(L)} \subset \big\{S_n = S_{(n,L)}\big\}.$$

By taking complements of the above relation,

$$\mathsf{P}\Big\{S_n \neq S_{(n,L)}\Big\} \le \mathsf{P}\Big\{\sup_{s\in[a,t]}|B_s| > L\Big\} \le \frac{\mathbb{E}\big(|B_t|\big)}{L} \tag{5.1}$$

where the upper-bound is obtained by Doob's submartingale inequality. Since B_1 is a normal random variable with mean 0 and variance 1, a simple computation yields

$$\mathbb{E}\big(|B_1|\big) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}|z|\exp\Big\{-\frac{z^2}{2}\Big\}\,dz = \sqrt{\frac{2}{\pi}}.$$

By (b) and (c) in Proposition 1.1.1, it follows that the upper-bound in (5.1) equals to $\sqrt{\frac{2t}{\pi L^2}}$.

Finally, for all $\varepsilon > 0$,

$$\begin{aligned} &\big\{|S_n| > \varepsilon\big\} \subset \big\{|S_{(n,L)}| > \varepsilon\big\} \cup \big\{S_n \neq S_{(n,L)}\big\} \\ \implies \quad &\mathsf{P}\big\{|S_n| > \varepsilon\big\} \le \mathsf{P}\big\{|S_{(n,L)}| > \varepsilon\big\} + \mathsf{P}\big\{S_n \neq S_{(n,L)}\big\}. \end{aligned} \tag{5.2}$$

Putting together (5.1) and (5.2),

$$\mathsf{P}\big\{|S_n| > \varepsilon\big\} \le \mathsf{P}\big\{|S_{(n,L)}| > \varepsilon\big\} + \Big(\frac{2t}{\pi L^2}\Big)^{1/2}.$$

For any $\varepsilon > 0$, $L > 0$ can be chosen such that $\big(\frac{2t}{\pi L^2}\big)^{1/2} < \varepsilon$. As $\|\mathsf{\Delta}_n\| \to 0$, $S_{(n,L)}$ converges to 0 for any $L > 0$, so it follows that there exists $N \ge 1$ such that $\mathsf{P}\big\{|S_{(n,L)}| > \varepsilon\big\} < \varepsilon/2$ for all $n \ge N$. Therefore, $\mathsf{P}\big\{|S_n| > \varepsilon\big\} < \varepsilon$ for all $n \ge N$, which establishes the desired result. □

As usual, let $\mathsf{\Delta}_n := \{a = t_0, t_1, \ldots, t_{n-1}, t_n = t\}$ be a refining partition of $[a,t]$. So trivially,

$$f\big(B_t\big) - f\big(B_a\big) = \sum_{i=1}^{n}\Big(f\big(B_{t_i}\big) - f\big(B_{t_{i-1}}\big)\Big). \tag{5.3}$$

Assume $f \in C^2(\mathbb{R})$. For a fixed point $x_0 \in [a,t]$, the first-order Taylor series expansion yields

$$f(x) = f(x_0) + f'(x_0)\big(x - x_0\big) + \mathfrak{R}_{x_0}^{(1)}(x), \tag{5.4}$$

where $\mathfrak{R}_{x_0}^{(1)}(x)$ denotes the remainder term. From elementary calculus, the remainder term can be written in the Cauchy form: that is, there exists $\lambda \in (0,1)$ such that

$$\mathfrak{R}_{x_0}^{(1)}(x) = \frac{1}{2} f''\big(x_0 + \lambda(x - x_0)\big)\,(x - x_0)^2.$$

Therefore, the equation (5.4) becomes

$$f(x) = f(x_0) + f'(x_0)\big(x - x_0\big) + \frac{1}{2} f''\big(x_0 + \lambda(x - x_0)\big)\,(x - x_0)^2. \tag{5.5}$$

Now, plugging (5.5) into the equation (5.3),

$$\begin{aligned} f\big(B_t\big) - f\big(B_a\big) = &\sum_{i=1}^{n} f'\big(B_{t_{i-1}}\big)\big(B_{t_i} - B_{t_{i-1}}\big) \\ &+ \frac{1}{2}\sum_{i=1}^{n} f''\big(B_{t_{i-1}} + \lambda_i(B_{t_i} - B_{t_{i-1}})\big)\big(B_{t_i} - B_{t_{i-1}}\big)^2, \end{aligned} \tag{5.6}$$

where $\lambda_i \in (0,1)$ for $1 \leq i \leq n$.

From the equation (3.5) in Section 3, it follows that as $\|\Delta_n\| \to 0$,

$$\sum_{i=1}^{n} f'(B_{t_{i-1}})(B_{t_i} - B_{t_{i-1}}) \longrightarrow \int_a^t f'(B_s)\,dB_s, \quad \text{in probability.}$$

To determine the remaining sum in (5.6), consider the expression

$$\begin{aligned} T_n :=& \sum_{i=1}^{n} f''\big(B_{t_{i-1}} + \lambda_i(B_{t_i} - B_{t_{i-1}})\big)(B_{t_i} - B_{t_{i-1}})^2 \\ =& \sum_{i=1}^{n}\Big(f''\big(B_{t_{i-1}} + \lambda_i(B_{t_i} - B_{t_{i-1}})\big) - f''(B_{t_{i-1}})\Big)(B_{t_i} - B_{t_{i-1}})^2 && (5.7) \\ &+ \sum_{i=1}^{n} f''(B_{t_{i-1}})\big((B_{t_i} - B_{t_{i-1}})^2 - (t_i - t_{i-1})\big) && (5.8) \\ &+ \sum_{i=1}^{n} f''(B_{t_{i-1}})(t_i - t_{i-1}). && (5.9) \end{aligned}$$

As $\|\Delta_n\| \to 0$, (5.7) converges almost surely to 0 by Lemma 5.1.1, and (5.8) converges to 0 in probability by Lemma 5.1.2. One can easily see that (5.9) a.s.-converges to the integral $\int_a^t f''(B_s)\,ds$, as $\|\Delta_n\| \to 0$. Hence, there exists a subsequence $\{T_{n_k}\} \subset \{T_n\}$ such that T_{n_k} converges to $\int_a^t f''(B_s)\,ds$ almost surely as $\|\Delta_{n_k}\| \to 0$. Thus the following theorem has been shown.

Theorem 5.1.1. *Let $f \in C^2(\mathbb{R})$. Then*

$$f(B_t) - f(B_a) = \int_a^t f'(B_s)\,dB_s + \frac{1}{2}\int_a^t f''(B_s)\,ds.$$

Let $f(t,x)$ be a function of (t,x). By putting $x := B_t$, one obtains a process $f(t, B_t)$. For typographical convenience, denote $\partial_x f := \frac{\partial f}{\partial x}$, $\partial_t f := \frac{\partial f}{\partial t}$, and $\partial_x^2 f := \frac{\partial^2 f}{\partial x^2}$.

Theorem 5.1.2. *Let $f(t,x)$ be a continuous function with continuous partial derivatives $\partial_x f$, $\partial_t f$, and $\partial_x^2 f$. Then*

$$f(t, B_t) - f(a, B_a) = \int_a^t \partial_x f(s, B_s)\, dB_s + \int_a^t \big(\partial_t f(s, B_s) + \frac{1}{2}\partial_x^2 f(s, B_s)\big)\, ds.$$

Proof. For each $n \in \mathbb{N}$, let Δ_n be a partition of the interval $[a,t]$. By the Taylor expansion, it follows that, for $\rho, \lambda \in (0,1)$,

$$\begin{aligned} f(t,x) - f(s,x_0) &= \big(f(t,x) - f(s,x)\big) + \big(f(s,x) - f(s,x_0)\big) \\ &= \partial_t f\big(s + \rho(t-s), x\big)(t-s) + \partial_x f(s,x_0)(x - x_0) + \frac{1}{2}\partial_x^2 f\big(s, x_0 + \lambda(x - x_0)\big)(x - x_0)^2. \end{aligned}$$

Mimicking the procedure used above, one can deduce that

$$\begin{aligned} f(t, B_t) - f(a, B_a) &= \sum_{i=1}^{n} \big(f(t_i, B_{t_i}) - f(t_{i-1}, B_{t_{i-1}})\big) \\ &= \sum_{i=1}^{n} \partial_t f\big(t_{i-1} + \rho(t_i - t_{i-1}), B_{t_{i-1}}\big)(t_i - t_{i-1}) && (5.10) \\ &+ \sum_{i=1}^{n} \partial_x f(t_{i-1}, B_{t_{i-1}})(B_{t_i} - B_{t_{i-1}}) && (5.11) \\ &+ \frac{1}{2}\sum_{i=1}^{n} \partial_x^2 f\big(t_{i-1}, B_{t_{i-1}} + \lambda(B_{t_i} - B_{t_{i-1}})\big)(B_{t_i} - B_{t_{i-1}})^2. && (5.12) \end{aligned}$$

Since B_t is a.s.-continuous and $\partial_t f$ is a continuous function, as $\|\Delta_n\| \to 0$, the summation (5.10) converges a.s. to $\int_a^t \partial_t f(s, B_s)\, ds$. Now, by modifying the similar argument used in the proof of Theorem 5.1.1, there exists a subsequence $\{\Delta_{n_k}\}$ such that

$$(5.11) \xrightarrow{a.s.} \int_a^t \partial_x f(s, B_s)\, dB_s \quad \text{and} \quad (5.12) \xrightarrow{a.s.} \frac{1}{2}\int_a^t \partial_x^2 f(s, B_s)\, ds,$$

which completes the proof. □

5.2. **The Itô Formula for Martingales.**

When M_t is a continuous, square integrable martingale, the proof of the Itô formula for M_t is very similar to the proof of Theorem 5.1.2; the difference in this case is to use the second-order Taylor expansion instead of the first-order expansion. As usual, let Δ_n be a partition of $[a,t]$ for every $n \in \mathbb{N}$, and let $F(x)$ be a C^2-function, so $F\big(M_t\big) - F\big(M_a\big) = \sum_{i=1}^{n}\big(F(M_{t_i}) - F(M_{t_{i-1}})\big)$. On each interval $[t_{i-1}, t_i]$, the second-order

Taylor expansion series gives

$$F(M_{t_i}) - F(M_{t_{i-1}}) = F'(M_{t_{i-1}})(M_{t_i} - M_{t_{i-1}}) + \frac{1}{2}F''(M_{t_{i-1}})(M_{t_i} - M_{t_{i-1}})^2 + \mathfrak{R}^{(2)}(M_{t_{i-1}}, M_{t_i}),$$

where $\mathfrak{R}^{(2)}(M_{t_{i-1}}, M_{t_i})$ denotes the corresponding remainder term. Moreover, Taylor's theorem implies the existence of a non-decreasing function $\mathfrak{r}^{(2)}\colon \mathbb{R}^+ \to \mathbb{R}^+$ such that $\lim_{s\downarrow 0} \mathfrak{r}^{(2)}(s) = 0$ and

$$\left|\mathfrak{R}^{(2)}(M_{t_{i-1}}, M_{t_i})\right| \le \mathfrak{r}^{(2)}\left(|M_{t_i} - M_{t_{i-1}}|\right)\left(M_{t_i} - M_{t_{i-1}}\right)^2. \tag{5.13}$$

For more details, see **[13]** and **[9]**. Summing all the partitions of $[a, t]$,

$$F(M_t) - F(M_a) = \sum_{i=1}^{n} F'(M_{t_{i-1}})(M_{t_i} - M_{t_{i-1}}) + \frac{1}{2}\sum_{i=1}^{n} F''(M_{t_{i-1}})(M_{t_i} - M_{t_{i-1}})^2 + \sum_{i=1}^{n} \mathfrak{R}^{(2)}(M_{t_{i-1}}, M_{t_i}). \tag{5.14}$$

For the first summation in the equation (5.14), it follows that, as $\|\Delta_n\| \to 0$,

$$\sum_{i=1}^{n} F'(M_{t_{i-1}})(M_{t_i} - M_{t_{i-1}}) \longrightarrow \int_a^t F'(M_s)\, dM_s, \quad \text{in probability.}$$

To determine the second summation term in (5.14), recall that

$$[M]_t := \lim_{\|\Delta_n\|\to 0} \sum_{i=1}^{n} \left(M_{t_i} - M_{t_{i-1}}\right)^2,$$

as a limit in probability. Therefore, as $\|\Delta_n\| \to 0$,

$$\sum_{i=1}^{n} F''(M_{t_{i-1}})(M_{t_i} - M_{t_{i-1}})^2 \longrightarrow \int_a^t F''(M_s)\, d[M]_s, \quad \text{in probability.}$$

Whenever M_t is a square integrable, continuous martingale, it is a fact that $\langle M\rangle_t = [M]_t$, so the integrator $d[M]_s$ can replaced by $d\langle M\rangle_s$ (see **[9]** or **[5]**). Finally, it follows from the inequality (5.13) that

$$\begin{aligned}\left|\sum_{i=1}^{n} \mathfrak{R}^{(2)}(M_{t_{i-1}}, M_{t_i})\right| \le \sum_{i=1}^{n} \left|\mathfrak{R}^{(2)}(M_{t_{i-1}}, M_{t_i})\right| &\le \sum_{i=1}^{n} \mathfrak{r}^{(2)}\left(|M_{t_i} - M_{t_{i-1}}|\right)\left(M_{t_i} - M_{t_{i-1}}\right)^2 \\ &\le \sup_{1\le j\le n}\left\{\mathfrak{r}^{(2)}\left(|M_{t_j} - M_{t_{j-1}}|\right)\right\} \sum_{i=1}^{n}\left(M_{t_i} - M_{t_{i-1}}\right)^2.\end{aligned}$$

From the a.s.-continuity of M_t, one can deduce that, as $\|\Delta_n\| \to 0$, the above inequality vanishes in probability. The following theorem summarizes the above argument.

Theorem 5.2.1. *Let M_t be a square integrable, continuous martingale, and $F(x)$ be a $C^2(\mathbb{R})$ function. Then*

$$F(M_t) - F(M_a) = \int_a^t F'(M_s)\, dM_s + \frac{1}{2}\int_a^t F''(M_s)\, d\langle M\rangle_s.$$

Moreover, following the similar procedure used in the proof of Theorem 5.1.2 yields the next theorem.

Theorem 5.2.2. *Let M_t be a continuous, square integrable martingale and let $F(t,x)$ be a continuous function with continuous partial derivatives $\frac{\partial F}{\partial t}$, $\frac{\partial F}{\partial x}$, and $\frac{\partial^2 F}{\partial x^2}$. Then*

$$F(t, M_t) - F(a, M_a) = \int_a^t \frac{\partial F}{\partial t}(s, M_s)\, ds + \int_a^t \frac{\partial F}{\partial x}(s, M_s)\, dM_s + \frac{1}{2}\int_a^t \frac{\partial^2 F}{\partial x^2}(s, M_s)\, d\langle M\rangle_s.$$

6. Lévy's Characterization of Brownian Motion

As one of the most remarkable applications of the Itô formula, this section aims to present an important alternative characterization of Brownian motion.

Theorem 6.0.3. *Lévy's Characterization Theorem*

A stochastic process W_t is a Brownian motion if and only if the following conditions are satisfied:

(a) $P\{\omega\colon W(0,\omega) = 0\} = 1$,

(b) for each $t \in [0,\infty)$, $\langle W\rangle_t = t$ almost surely,

(c) there exists a filtration $\{\mathcal{F}_t\}$ such that W_t is a continuous martingale with respect $\{\mathcal{F}_t\}$.

Proof. If W_t is a Brownian motion, then conditions (a)-(c) are trivially satisfied.

Now assume that W_t satisfies conditions (a)-(c). From hypotheses (a) and (c), conditions (i) and (iv) in Definition 1.1.1 are easily satisfied. Let $F(t,x) := \exp\{\mathrm{i}\lambda x + \frac{\lambda^2 t}{2}\}$. Applying the Itô formula to $F(t, W_t)$ and expressing it in the differential form gives

$$\begin{aligned} dF(t, W_t) &= \mathrm{i}\lambda F(t, M_t)\, dW_t + \frac{\lambda^2}{2}\big(F(t, W_t)\, dt - F(t, W_t)\, d\langle W\rangle_t\big) \\ &= \mathrm{i}\lambda F(t, W_t)\, dW_t, \qquad \text{by condition (b).} \end{aligned}$$

Converting to the integral form gives

$$F(t, W_t) - F(0, W_0) = \int_0^t \partial_x F(s, W_s)\, dW_s = \mathrm{i}\lambda \int_0^t F(s, W_s)\, dW_s.$$

Since $\mathbb{E}\Big(\int_0^t |F_s(W_s)|^2\,ds\Big) < \infty$, it follows from Theorem 4.0.3 that $F(t, W_t)$ is a martingale with respect to $\{\mathcal{F}_t\}$. Hence, for $s \in [0, t]$,

$$\begin{aligned}\mathbb{E}\big(F(t,W_t)\big|\mathcal{F}_s\big) = F(s,W_s) &\Longleftrightarrow \mathbb{E}\Big(\exp\big\{\mathrm{i}\lambda W_t + \tfrac{1}{2}\lambda^2 t\big\}\Big|\mathcal{F}_s\Big) = \exp\big\{\mathrm{i}\lambda W_s + \tfrac{1}{2}\lambda^2 s\big\} \\ &\Longleftrightarrow \mathbb{E}\Big(\exp\big\{\mathrm{i}\lambda(W_t - W_s)\big\}\Big|\mathcal{F}_s\Big) = \exp\big\{-\tfrac{1}{2}\lambda^2(t-s)\big\}. \qquad (6.1)\end{aligned}$$

By taking the additional expectation of both sides in (6.1),

$$\mathbb{E}\Big(\exp\big\{\mathrm{i}\lambda(W_t - W_s)\big\}\Big) = \exp\big\{-\tfrac{1}{2}\lambda^2(t-s)\big\} \qquad (6.2)$$

holds for all $\lambda \in \mathbb{R}$. Therefore, $W_t - W_s$ has a normal distribution with mean 0 and variance $(t - s)$, which confirms condition (ii) in Definition 1.1.1.

To verify Definition 1.1.1 (iii), the proof proceeds by induction. Let $0 \le t_1 < t_2$. For any $\lambda_1, \lambda_2 \in \mathbb{R}$,

$$\begin{aligned}\mathbb{E}\Big(\exp\big\{\mathrm{i}\lambda_1 W_{t_1} + \mathrm{i}\lambda_2(W_{t_2} - W_{t_1})\big\}\Big) &= \mathbb{E}\Big(\mathbb{E}\Big(\exp\big\{\mathrm{i}\lambda_1 W_{t_1} + \mathrm{i}\lambda_2(W_{t_2} - W_{t_1})\big\}\Big|\mathcal{F}_{t_1}\Big)\Big) \\ &= \mathbb{E}\Big(\exp\big\{\mathrm{i}(\lambda_1 - \lambda_2)W_{t_1}\big\}\mathbb{E}\Big(\exp\big\{\mathrm{i}\lambda_2 W_{t_2}\big\}\Big|\mathcal{F}_{t_1}\Big)\Big). \qquad (6.3)\end{aligned}$$

Setting $s := t_1$ and $t := t_2$ in (6.1),

$$\mathbb{E}\Big(\exp\big\{\mathrm{i}\lambda_2 W_{t_2}\big\}\Big|\mathcal{F}_{t_1}\Big) = \exp\big\{\mathrm{i}\lambda_2 W_{t_1} - \tfrac{1}{2}\lambda_2^2(t_2 - t_1)\big\}. \qquad (6.4)$$

Plugging (6.4) into (6.3), for all $\lambda_1, \lambda_2 \in \mathbb{R}$,

$$\begin{aligned}\mathbb{E}\Big(\exp\big\{\mathrm{i}\lambda_1 W_{t_1} + \mathrm{i}\lambda_2(W_{t_2} - W_{t_1})\big\}\Big) &= \exp\big\{-\tfrac{1}{2}\lambda_2^2(t_2 - t_1)\big\}\,\mathbb{E}\Big(\exp\big\{\mathrm{i}\lambda_1 W_{t_1}\big\}\Big) \\ &= \exp\big\{-\tfrac{1}{2}\lambda_2^2(t_2 - t_1)\big\}\exp\big\{-\tfrac{1}{2}\lambda_1^2 t_1\big\},\end{aligned}$$

where the last equality follows from (6.2). Hence W_{t_1} and $(W_{t_2} - W_{t_1})$ are independent and normally distributed. For a given partition $t_0 := 0 \le t_1 < \cdots < t_n \le 1$, one can inductively repeat the above argument to deduce the equality

$$\mathbb{E}\Big(\prod_{j=1}^{n} \exp\big\{\mathrm{i}\lambda_j(W_{t_j} - W_{t_{j-1}})\big\}\Big) = \exp\Big\{-\frac{1}{2}\sum_{j=1}^{n}\lambda_j^2(t_j - t_{j-1})\Big\}, \qquad (6.5)$$

for all $\lambda_j \in \mathbb{R}$ with $1 \le j \le n$. So it follows from (6.5) that the increments $\big\{(W_{t_j} - W_{t_{j-1}}); 1 \le j \le n\big\}$ are independent and normally distributed, which establishes the condition (iii) of Definition 1.1.1. Therefore, W_t is a Brownian motion. $\square$

References

[1] Folland, G.B.: *Real Analysis: Modern Techniques and Their Applications*, Second edition. John Wiley & Sons, 1999.

[2] Hida, T.: *Brownian Motion*. Springer-Verlag, 1980.

[3] Igari, S.: *Introduction to Real Analysis (in Japanese)*. Iwanami, 1996.

[4] Klenke, A.: *Probability Theory: A Comprehensive Course*. Springer-Verlag, 2008.

[5] Kopp, P.E.: *Martingales and Stochastic Integrals*. Cambridge University Press, 1984.

[6] Kuo, H.-H.: *Introduction to Stochastic Integration*. Springer-Verlag, 2006.

[7] Medvegyev, P.: *Stochastic Integration Theory*. Oxford University Press, 2007.

[8] Øksendal, B.: *Stochastic Differential Equations*, Sixth edition. Springer-Verlag, 2003.

[9] Protter, P. E.: *Stochastic Integration and Differential Equations*, Second edition. Springer-Verlag, 2005.

[10] Rogers, L. C. G. and Williams, D.: *Diffusions, Markov Processes and Martingales, vol. 1*, Second edition. Cambridge University Press, 2000.

[11] Rogers, L. C. G. and Williams, D.: *Diffusions, Markov Processes and Martingales, vol. 2*, Second edition. Cambridge University Press, 2000.

[12] Shiryaev, A. N.: *Probability*, Second edition. Springer-Verlag, 1989.

[13] Spivak, M.: *Calculus*, Third edition. Publish or Perish, 1994.

[14] Steele, J. M.: *Stochastic Calculus and Financial Applications*. Springer-Verlag, 2001.