

MÁSTERES de la UAM

Facultad de Ciencias / 16-17

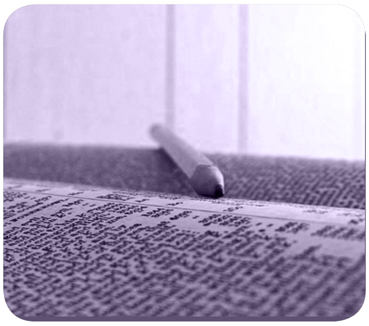
Matemáticas y Aplicaciones

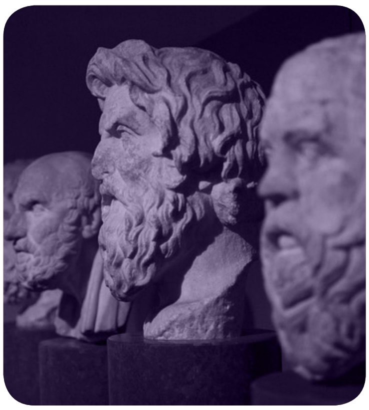
UNIVERSIDAD AUTONOMA DE MADRID

Campus Internacional excelencia UAM CSIC+

The Itô Integral and Anticipating Generalizations

Joan Carles Bastons Garcia



UAM EDICIONES

Trabajo Fin de Máster

Máster en Matemáticas y Aplicaciones

Facultad de Ciencias

Universidad Autónoma de Madrid

The Itô Integral and Anticipating Generalizations

Autor: Joan Carles Bastons Garcia

Director: Dr. Carlos Escudero Liébana
Departmento: Matemáticas

Madrid, 11 de septiembre de 2017

Abstract

Itô integration is one of the most important mathematical theories of the 20th century due to its applications in Mathematics and other fields. We study the construction of the Itô integral along with its applicability in the theory of stochastic differential equations. The second part of this dissertation consists in understanding the shortcomings of Itô integration related to anticipating calculus and examining stochastic integrals which overcome them. We construct the Skorohod integral, which was the first extension of the Itô integral, and we introduce the Ayed-Kuo integral, which is a new integral introduced in 2008. Some remarkable aspects of this integral are the facts that it deals with anticipating calculus and generalizes many of the properties of the Itô integral. Some open questions about the Ayed-Kuo integral are discussed and the insider problem in Finance is considered as a motivation for the study of anticipating stochastic calculus.

Acknowledgments

I would like to express my gratitude to Dr. Carlos Escudero for his guidance and advice. He has been always receptive to my interests and problems. I am also grateful to Dr. Pedro Balodis for his interest in what I was doing. I also feel thankful to the Department of Mathematics of Univesidad Autónoma de Madrid for the scholarship provided.

Agraeixo els ànims i l'afecte rebuts als meus familiars i amics durant tot el temps que m'ha ocupat aquest treball. En especial, als meus pares i a la meva germana.

También quiero expresar mi agradecimiento a todas las amistades que he hecho en Madrid porque me han hecho sentir menos lejos de casa. En especial, mi gratitud hacia Andoni y Chan Woo.

Contents

Introduction

The Itô integral was introduced in 1944 by Japanese mathematician Kiyosi Itô with the aim of obtaining, through stochastic differential equations, a way of generating diffusions. At that time, the unbounded variation of Brownian motion made very restrictive the existing attempts of integrating with respect to Brownian motion. Thus, Itô integration was a breakthrough in both pure and applied Mathematics. One of the most popular applications of Itô theory has been the Black-Scholes formula, proposed by Fisher Black and Myron Scholes in 1973, for pricing European options.

Although the Itô integral has a wide rage of applications, there are some shortcomings associated with measurability. Itô integration requires that the integrand does not possess any information about the future, we say that the stochastic process must be adapted or non-anticipating. Imagine that a trader had privileged information about future prices, then any trade decision would be ill-posed in Itô theory. These problems of insider trading are a motivation for the study of anticipating calculus. The origin of anticipating integrals was a theoretical motivation and dates back to the decade of 1970.

The aim of this dissertation is to understand how the Itô integral is constructed in order to identify its restrictions and provide generalizations which overcome them. The theory of stochastic differential equations and some financial problems are the main motivation to introduce new concepts, although they are not our main interest. This project also focuses on two anticipating generalizations of the Itô integral. The Skorohod integral, constructed by Anatoliy Skorohod in 1975, was the first anticipating integral and it was later closely related to Malliavin Calculus. The other one is the Ayed-Kuo integral, proposed by Wided Ayed and Hui-Hsiung Kuo in 2008. This integral keeps some analogies with the Itô integral and its relation with Skorohod integration is not clear yet. Moreover, many aspects of the Ayed-Kuo integral are still an area of research and it has barely been applied to financial problems. Therefore, some of the contents of this dissertation lead to open problems and a further development of new theories with an applied motivation.

This dissertation is composed of five chapters and three appendices. The first chapter consists in an introduction to Brownian motion because we will always integrate with respect to it. Some of its main characteristics are introduced and studied because of their implications in stochastic integration. Among them, the quadratic variation, the martingale property, the fact that it is not of bounded variation and its path regularity. In Chapter 2, the Itô integral is constructed through step processes and a density argument as in the construction of the Lebesgue integral. The notion of indefinite integral is introduced along with the properties of being a martingale or having continuous trajectories. Another general strategy when constructing integrals is to consider Riemann sums. Therefore, we show that the Itô integral can also be understood in terms of Riemann sums by taking partitions and evaluating the integrand at the left endpoints of the partition. This is one main difference with Calculus because the choice of the evaluation points matters. The third chapter focuses on the existence and uniqueness of a solution for stochastic differential equations and some path properties of the solutions.

The part concerning anticipating calculus consists of chapters 4 and 5. The main goal is to give some sense to integrals like

$$\int_0^T B(T)dB_t,\ T > 0,$$

which is clearly anticipating. In the fourth chapter, the Skorohod integral is constructed following Skorohod initial ideas. To this purpose, the Wiener-Itô chaos expansion is studied and also the iterated Itô integrals. Some examples, which would be ill-posed in the Itô theory, are computed. The final chapter is devoted to examine the Ayed-Kuo integral, which allows to integrate products of an adapted stochastic process and an instantly independent one with respect to Brownian motion. The definition of this integral is given in terms of Riemann sums and occurrences of the adapted process are evaluated at the left endpoints while for the other process the occurrences are evaluated at the right endpoints of the partition. We examine the zero-mean property, the near-martingale property, an isometry and a new Itô formula for this new integral. This dissertation finishes with these new properties of the Ayed-Kuo integral and a description of unsolved questions.

The appendixes include a summary on conditional expectation, whose properties are essential in many proofs, the Doob submartingale inequality and the Itô representation theorem.

Chapter 1

A Summary on Brownian Motion

The concept of Brownian motion appeared for the first time at the beginning of the 19th century and has become fundamental in the theory of stochastic processes and modern Probability. Nowadays, its applications go beyond Mathematics, and it is used in Physics, Economics and Medicine, among other disciplines. Such a wide applicability is due to the fact that Brownian motion models quite properly problems which have arisen in these fields. Take for instance, the trajectories of gas molecules in Physical Statistics, asset pricing in Finance or the expansion of some diseases in Medicine.

To our purpose, Brownian motion is fundamental in the theory of stochastic integration because its properties, as we will show, made impossible to consider Riemann-Stieltjes integration and entailed the development of new integration theories. The goal of this chapter is to define it and examine some properties which will be crucial for the study of stochastic integrals in consecutive chapters.

1.1 Stochastic Processes and Martingales

In this sections we recall elemental probabilistic concepts such as random variable, measurability with respect to a σ-field and the concepts of stochastic process and martingale.
We fix a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a σ-field $\mathcal{G}$ such that $\mathcal{G} \subset \mathcal{F}$.

Definition 1.1.1. A function $X : \Omega \to \mathbb{R}$ is *measurable* if and only if $X^{-1}(B) \in \mathcal{F}$ for any $B \in \mathcal{B}(\mathbb{R})$.

Definition 1.1.2. We say that $X : \Omega \to \mathbb{R}$ is a *random variable* if X is measurable.

Definition 1.1.3. A random variable $X : \Omega \to \mathbb{R}$ is *measurable with respect to* $\mathcal{G}$, or $\mathcal{G}$*-measurable*, if and only if $X^{-1}(B) \in \mathcal{G}$ for any $B \in \mathcal{B}(\mathbb{R})$.

Remark 1.1.4. Given a function $X : \Omega \to \mathbb{R}$, we will denote as $\sigma(X)$ the smallest σ-field for which X is measurable.

Proposition 1.1.5 (Doob-Dynkin Lemma)**.** *Let $X, Y : \Omega \to \mathbb{R}$ be two random variables. Then, Y is $\sigma(X)$-measurable if and only if there exists a Borel measurable function $f : \mathbb{R} \to \mathbb{R}$ such that $Y = f(X)$.*

Definition 1.1.6. A random variable $X : \Omega \to \mathbb{R}$ is *independent of* $\mathcal{G}$ if and only if the events $\{\omega : X(\omega) \in B\}$ and $G \in \mathcal{G}$ are independent for any $B \in \mathcal{B}(\mathbb{R})$ and any $G \in \mathcal{G}$.

Definition 1.1.7. A *stochastic process* with state space S is a family $\{X_i : i \in I\}$ of random variables $X_i : \Omega \to S$ indexed by a set I.

Definition 1.1.8. The *sample paths* of a stochastic process $\{X_t, t \in I\}$ are the family of functions indexed by $\omega \in \Omega$, $X(\omega) : I \to S$, defined by $X(\omega)(t) := X_t(\omega)$.

In this dissertation, we will take $S = \mathbb{R}$ and either $I = [0, \infty)$ or $I = [0, T]$ with $T > 0$.

Definition 1.1.9. A stochastic process $\{\widetilde{X}_t,\ t \geq 0\}$ is a *version* of the stochastic process $\{X_t,\ t \geq 0\}$ if for all $t \geq 0$

$$\mathbb{P}(X_t = \widetilde{X}_t) = 1.$$

Definition 1.1.10. A *filtration* is an increasing family $\{\mathcal{F}_t : t \geq 0\}$ of sub-σ-fields of $\mathcal{F}$.

Definition 1.1.11. A stochastic process $\{X_t, t \geq 0\}$ is *a martingale with respect to the filtration* $\{\mathcal{F}_t : t \geq 0\}$ if and only if

(i) $E\,|X_t| < \infty$ for all $t \geq 0$,

(ii) X_t is $\mathcal{F}_t$-measurable for any $t \geq 0$, and

(iii) $\mathbb{E}\left(X_t | \mathcal{F}_s\right) = X_s$ almost surely for all $0 \leq s \leq t$.

Remark 1.1.12. We adopt the notation $\sigma\left(X_s,\ 0 \leq s \leq t\right)$ to denote the smallest σ-field containing all the σ-fields generated by X_s, $0 \leq s \leq t$.

Remark 1.1.13. A martingale satisfies what is known as the fair game property. By Theorem $A.3$ and the definition of martingale, we have that

$$\mathbb{E}(X_t) = \mathbb{E}\left[\mathbb{E}\left(X_t | \mathcal{F}_0\right)\right] = \mathbb{E}(X_0)$$

for all $t \geq 0$. The name of this property is referred to bets in games. If X_0 is the initial wealth at the beginning of the game, then we expect to possess, in average, the same quantity when we quit it.

Definition 1.1.14. Let $\{X_t, t \geq 0\}$ be a stochastic process and $\{\mathcal{F}_t, t \geq 0\}$ a filtration. Assume that the stochastic process satisfies conditions $(i), (ii)$ of Definition 1.1.11.

(i) The process $\{X_t, t \geq 0\}$ is a *submartingale* with respect to the filtration if and only if $\mathbb{E}\left(X_t | \mathcal{F}_s\right) \geq X_s$ almost surely for all $0 \leq s \leq t$.

(ii) The process $\{X_t, t \geq 0\}$ is a *supermartingale* with respect to the filtration if and only if $\mathbb{E}\left(X_t | \mathcal{F}_s\right) \leq X_s$ almost surely for all $0 \leq s \leq t$.

Next, we establish a relation between martingales and submartingales.

Proposition 1.1.15. *Let $\{X_t, t \in [0, T]\}$ be a martingale and $\varphi : \mathbb{R} \to \mathbb{R}$ a convex function. If $\mathbb{E}\,|\varphi(X_t)|$ is finite for all $t \in [0, T]$, then $\{\varphi(X_t), t \in [0, T]\}$ is a submartingale.*

Proof. By the conditional Jensen inequality, for $0 \leq s \leq t$,

$$\mathbb{E}\left[\varphi(X_t) | \mathcal{F}_s\right] \geq \varphi\left(\mathbb{E}\left[X_t | \mathcal{F}_s\right]\right) = \varphi(X_s).$$

□

1.2 A Historical Perspective on Brownian Motion

In 1828, British botanist Robert Brown published a paper in which he described the irregular movements of pollen particles suspended in a fluid. He noticed that trajectories were irregular and random. The contributions of Robert Brown account for the name Brownian motion.

In 1900, Louis Bachelier used this new concept in order to model the evolution of stock prices and interest rates. It was not until 1905 that Albert Einstein provided a successful explanation to the remarks made by Robert Brown. The well-known physicist argued that pollen particles are permanently bombarded by fluid molecules, which account for the sharp paths, and he obtained the equations for Brownian motion as a diffusion.

The first mathematical construction of Brownian motion dates back to 1931, when Norbert Wiener made the first rigorous construction. Since then, Brownian motion has been deeply studied in the theory of stochastic processes and it has a wide range of applications in many disciplines such as Physics, Economics and Medicine.

1.3 Definition and First Properties

We begin defining a one-dimensional Brownian motion and giving some elementary properties.

Definition 1.3.1. A stochastic process $\{B_t, t \geq 0\}$ is a one-dimensional *Brownian motion* or a *Wiener Process* if

(i) $B_0 = 0$ almost surely.

(ii) $B_t - B_s \sim N(0, t-s)$ for all $0 \leq s \leq t$.

(iii) For any positive integer $n \in \mathbb{N}$ and any $0 \leq t_0 < t_1 < \ldots < t_n$, the random variables

$$B_{t_n} - B_{t_{n-1}}, B_{t_{n-1}} - B_{t_{n-2}}, \ldots, B_{t_1} - B_{t_0}$$

are independent.

Remark 1.3.2. Let $\{B_t, t \geq 0\}$ be a Brownian motion. Then,

(i) $\mathbb{E}(B_t) = \mathbb{E}(B_t - B_0) = 0$.

(ii) $\mathbb{E}(B_t^2) = \mathbb{E}\left[(B_t - B_0)^2\right] = \text{Var}\,(B_t - B_0) + (\mathbb{E}\,(B_t - B_0))^2 = t - 0 = t$.

Therefore, $B_t \sim N(0, t)$.

Proposition 1.3.3. *If* $\{B_t, t \geq 0\}$ *is a Brownian motion, then*

$$Cov(B_t, B_s) = \mathbb{E}(B_t B_s) = t \wedge s := \min\{s, t\}.$$

Proof. By Remark 1.3.2, $\text{Cov}(B_t, B_s) = \mathbb{E}(B_t B_s)$. Assume $s \leq t$,

$$\begin{aligned}\mathbb{E}(B_t B_s) &= \mathbb{E}\left[(B_t - B_s + B_s)\, B_s\right] = \mathbb{E}\left[(B_t - B_s)\, B_s\right] + \mathbb{E}\left(B_s^2\right) \\ &= \mathbb{E}(B_t - B_s)\mathbb{E}(B_s) + \mathbb{E}(B_s^2) = s = \min\{t, s\}.\end{aligned}$$

□

Remark 1.3.4. In some manuals, instead of defining Brownian motion as in Definition 1.3.1, it is defined as a Gaussian stochastic process with zero mean and covariance matrix as in Proposition 1.3.3. That is, for any integer $N \geq 1$ and any $0 \leq t_1 < t_2 < \ldots < t_N$, the random vector

$$(B_{t_1}, \ldots, B_{t_N})$$

has a multidimensional normal distribution with vector mean

$$\mu = (\mathbb{E}(B_{t_1}), \ldots, \mathbb{E}(B_{t_N})) = \vec{0}$$

and covariance matrix Σ defined by

$$\Sigma(i, j) = \min\{t_i, t_j\}$$

for any $1 \leq i \leq j \leq N$.

Next, we show that both definitions are equivalent.

Proposition 1.3.5. *A stochastic process $\{X_t, t \geq 0\}$ is a Brownian motion if and only if it is Gaussian with zero mean vector and covariance function given by $\Sigma(s, t) = \min(s, t)$.*

Proof. We first assume that the stochastic process is a Brownian motion. For any positive integer n and any $0 \leq t_1 < t_2 < \ldots < t_n$, we need to check that $(X_{t_1}, \ldots, X_{t_n})$ is a normal random vector with the right mean and covariance. Since the vector can be expressed as a linear transformation of random vectors with Gaussian independent components, the vector is also Gaussian. It is immediate to see that it has zero mean. For $0 \leq s \leq t$, we have that

$$\mathbb{E}(X_t X_s) = \mathbb{E}\left((X_t - X_s + X_s) X_s\right) = \mathbb{E}(X_s^2) = s = \min(s, t).$$

Now assume that $\{X_t, t \geq 0\}$ is a Gaussian stochastic process as in the statement. Then, $\mathbb{E}(X_0^2) = 0$, which implies that $X_0 = 0$ almost surely. For any $0 \leq s \leq t$ and any $r \geq 0$,

$$\mathbb{E}(X_s(X_{t+r} - X_t) = \mathbb{E}(X_s X_{t+r}) - \mathbb{E}(X_s X_t) = s - s = 0. \tag{1.3.1}$$

Note that $X_{t+r} - X_t$ is Gaussian because it is a linear transformation of Gaussian random variables. Since linear combinations of Gaussian random variables are Gaussian, we have that $X_t - X_s$ is Gaussian for any $0 \leq s \leq t$. Then, Equation (1.3.1) implies independent disjoint increments because the random variables are Gaussian and uncorrelated.
Moreover,

$$\mathbb{E}(X_t - X_s) = \mathbb{E}(X_t) - \mathbb{E}(X_s) = 0$$

and

$$\mathbb{E}\left[(X_t - X_s)^2\right] = \mathbb{E}(X_t^2) + \mathbb{E}(X_s^2) + 2\mathbb{E}(X_t X_s) = t + s - 2ts = t - s. \qquad \square$$

We show in the following theorem that Brownian motion is a martingale.

Theorem 1.3.6. *Brownian motion is a martingale.*

Proof. It is clear that B_t is integrable and $\mathcal{F}_t$-measurable. Since Brownian motion has independent increments and the increments have zero mean, given $0 \leq s \leq t$,

$$\mathbb{E}\left[B_t - B_s \mid \mathcal{F}_s\right] = \mathbb{E}\left[B_t - B_s\right] = 0.$$

Then,

$$\mathbb{E}\left[B_t \mid \mathcal{F}_s\right] = \mathbb{E}\left[(B_t - B_s) + B_s \mid \mathcal{F}_s\right] = \mathbb{E}\left[B_t - B_s \mid \mathcal{F}_s\right] + \mathbb{E}\left[B_s \mid \mathcal{F}_s\right] = B_s.$$

□

It should be checked that Brownian motion actually exists. Although there are many ways of showing existence, we suggest two. The first one is through the Kolmogorov extension theorem and the other one is the Lévy-Ciesielski construction.
For the first approach, one needs to check the conditions of the Kolmogorov extension theorem, basically, the consistency condition. In Section 3.3 of [14], a full construction is given. For the second idea, one constructs explicitly the process and takes the Haar wavelet as an orthonormal basis of $L^2([0,1])$. A full proof is available in Section 3.1 of [22].

1.4 Path Regularity of Brownian Motion

The goal of this section is to show that Brownian motion always has a version with Hölder continuous sample paths with exponent $\gamma \in (0, 1/2)$. To this purpose, we examine the Kolmogorov continuity theorem. We also study that Brownian motion has nowhere differentiable sample paths.

1.4.1 Hölder Continuity

Theorem 1.4.1 (Kolmogorov continuity theorem)**.** *Let $\{X_t,\ t \in [0,T]\}$ be a stochastic process indexed in a bounded interval $[0,T]$. If there exist three constants $C, \alpha, \beta > 0$ such that*

$$\mathbb{E}\left[\left|X_t - X_s\right|^\beta\right] \leq C \left|t - s\right|^{1+\alpha}$$

for all $s, t \in [0,T]$, then there exists a version $\{\widetilde{X}_t,\ t \geq 0\}$ of the process with almost surely Hölder continuous paths with exponent $\gamma \in \left(0, \frac{\alpha}{\beta}\right)$.

Proof. Without loss of generality, we assume $T = 1$. We pick $0 < \gamma < \frac{\alpha}{\beta}$ and for $n \geq 1$ we define the set

$$A_n = \left\{\omega \in \Omega : \left|X\left(\frac{i+1}{2^n}, \omega\right) - X\left(\frac{i}{2^n}, \omega\right)\right| > \frac{1}{2^{n\gamma}},\ \text{for some } 0 \leq i < 2^n\right\}.$$

Then,

$$\mathbb{P}(A_n) \leq \sum_{i=0}^{2^n-1} \mathbb{P}\left(\left|X\left(\frac{i+1}{2^n}\right) - X\left(\frac{i}{2^n}\right)\right| > \frac{1}{2^{n\gamma}}\right)$$

$$
\begin{aligned}
&\le \sum_{i=0}^{2^n-1} \mathbb{E}\left[\left|X\left(\frac{i+1}{2^n}\right) - X\left(\frac{i}{2^n}\right)\right|^{\beta}\right]\left(\frac{1}{2^{n\gamma}}\right)^{-\beta} \\
&\le C\sum_{i=0}^{2^n-1}\left(\frac{1}{2^n}\right)^{1+\alpha} 2^{n\gamma\beta} \\
&= C2^{n(\gamma\beta-\alpha)}.
\end{aligned}
$$

The choice of γ assures that $\gamma\beta - \alpha < 0$ and, therefore,

$$
\sum_{n=1}^{\infty} \mathbb{P}(A_n) \le C\sum_{n=1}^{\infty} 2^{n(\gamma\beta-\alpha)} < \infty.
$$

By the Borel-Cantelli Lemma,

$$
\mathbb{P}(\liminf_{n\to\infty} A_n^c) = 1.
$$

Hence, for almost all ω, there exists $n_0(\omega) \in \mathbb{N}$ such that for all $n \ge n_0(\omega)$ and all $i \in \{0, \dots, 2^n - 1\}$

$$
\left|X\left(\frac{i+1}{2^n}, \omega\right) - X\left(\frac{i}{2^n}, \omega\right)\right| \le \frac{1}{2^{n\gamma}}.
$$

For each $\omega \in \Omega$, we define the constant

$$
K_\gamma(\omega) = \sup_{n\ge1}\left(\sup_{0\le i\le 2^n-1}\left|X\left(\frac{i+1}{2^n}, \omega\right) - X\left(\frac{i}{2^n}, \omega\right)\right| 2^{n\gamma}\right),
$$

which is finite almost surely. Indeed, for all $n \ge n_0(\omega)$ the supremum on i is bounded by 1 a.s., and for $n \le n_0(\omega)$ there are a finite number of elements. We conclude that for almost every ω there exists a constant $K_\gamma(\omega)$ such that

$$
\left|X\left(\frac{i+1}{2^n}, \omega\right) - X\left(\frac{i}{2^n}, \omega\right)\right| \le K_\gamma(\omega)\frac{1}{2^{n\gamma}}. \tag{1.4.1}
$$

for all $n \ge 1$ and all $i = 0, \dots, 2^{n-1}$. Next, we want to extend the inequality in (1.4.1) to all dyadic numbers in $[0, 1]$.
We pick $\omega \in \{K_\gamma(\omega) < \infty\}$. Let t, s be two dyadic numbers such that $0 < s < t < 1$, and let $n \in \mathbb{N}$ be the smallest positive integer such that

$$
2^{-n} \le t - s < 2^{-n+1}.
$$

Then, we can write

$$
\begin{aligned}
s &= \frac{i}{2^n} - \sum_{r=1}^{k}\frac{1}{2^{p_r}} \quad \text{with } n < p_1 < \ldots < p_k, \\
t &= \frac{j}{2^n} + \sum_{r=1}^{l}\frac{1}{2^{q_r}} \quad \text{with } n < q_1 < \ldots < q_l,
\end{aligned}
$$

for some integers i, j satisfying

$$s \leq \frac{i}{2^n} \leq \frac{j}{2^n} \leq t.$$

Since

$$\frac{j-i}{2^n} \leq t - s < \frac{1}{2^{n-1}},$$

either $j = i$ or $j = i + 1$. Taking into account (1.4.1), it follows

$$\left| X\left(\frac{i}{2^n}, \omega\right) - X\left(\frac{j}{2^n}, \omega\right) \right| \leq K_\gamma(\omega) \left| \frac{i-j}{2^n} \right|^\gamma \leq K_\gamma(\omega) \left| t - s \right|^\gamma. \tag{1.4.2}$$

By (1.4.1), it follows that

$$\left| X\left(\frac{i}{2^n} - \frac{1}{2^{p_1}} - \ldots - \frac{1}{2^{p_r}}, \omega\right) - X\left(\frac{i}{2^n} - \frac{1}{2^{p_1}} - \ldots - \frac{1}{2^{p_{r-1}}}, \omega\right) \right| \leq K_\gamma(\omega) 2^{-p_r \gamma},$$

for all $1 \leq r \leq k$. Then,

$$\begin{aligned}
&\left| X\left(s, \omega\right) - X\left(\frac{i}{2^n}, \omega\right) \right| \\
&\leq \sum_{r=2}^{k} \left| X\left(\frac{i}{2^n} - \frac{1}{2^{p_1}} - \ldots - \frac{1}{2^{p_r}}, \omega\right) - X\left(\frac{i}{2^n} - \frac{1}{2^{p_1}} - \ldots - \frac{1}{2^{p_{r-1}}}, \omega\right) \right| \\
&\quad + \left| X\left(\frac{i}{2^n} - \frac{1}{2^{p_1}}, \omega\right) - X\left(\frac{i}{2^n}, \omega\right) \right| \\
&\leq K_\gamma(\omega) \sum_{r=1}^{k} \left(\frac{1}{2^{p_r}}\right)^\gamma \\
&\leq K_\gamma(\omega) \sum_{r=1}^{k} \left(\frac{1}{2^{n+r}}\right)^\gamma \\
&\leq K_\gamma(\omega) \frac{1}{2^{n\gamma}} \sum_{r=1}^{\infty} \left(\frac{1}{2^r}\right)^\gamma \\
&\leq C_\gamma(\omega) \frac{1}{2^{n\gamma}} \\
&\leq C_\gamma(\omega) \left| t - s \right|^\gamma.
\end{aligned} \tag{1.4.3}$$

Likewise, we obtain

$$\left| X\left(t, \omega\right) - X\left(\frac{j}{2^n}, \omega\right) \right| \leq C_\gamma(\omega) \left| t - s \right|^\gamma. \tag{1.4.4}$$

By the triangle inequality and combining (1.4.2), (1.4.3) and (1.4.4),

$$\left| X\left(t, \omega\right) - X\left(s, \omega\right) \right| \leq C_\gamma(\omega) \left| t - s \right|^\gamma,$$

for all $s,t \in [0,1]$ dyadic numbers. Since dyadic numbers, $\mathbb{D}$, are dense in $[0,1]$, we define

$$\tilde{X}_t(\omega) := \begin{cases} \lim_{\substack{s\to t\\ s\in\mathbb{D}}} X_s(\omega), & \text{if } \omega \in \{C_\gamma(\omega) < \infty\}, \\ 0, & \text{otherwise.} \end{cases}$$

Next, we prove that $t \mapsto \tilde{X}_t(\omega)$ is a.s. Hölder continuous. Let $t,s \in [0,1]$, we assume $s < t$ and $\omega \in \{K_\gamma(\omega) < \infty\}$. Given $\varepsilon > 0$,

$$\begin{aligned} \left|\tilde{X}_t(\omega) - \tilde{X}_s(\omega)\right| &\leq \left|\tilde{X}_t(\omega) - \tilde{X}_{d_1}(\omega)\right| + \left|\tilde{X}_{d_1}(\omega) - \tilde{X}_{d_2}(\omega)\right| + \left|\tilde{X}_{d_2}(\omega) - \tilde{X}_s(\omega)\right| \\ &< \frac{\varepsilon}{2} + C_\gamma(\omega)\,|d_1 - d_2|^\gamma + \frac{\varepsilon}{2} \\ &\leq C_\gamma(\omega)\,|t-s|^\gamma + \varepsilon, \end{aligned}$$

where we have picked $d_1, d_2 \in [s,t]$ two dyadic numbers close enough to the border of the interval. This proves that $t \mapsto \tilde{X}_t(\omega)$ is a.s. Hölder continuous with exponent γ. Showing that $\tilde{X}_t$ is a version of the process X_t will conclude the proof. Let $\{s_k\}_k \subset \mathbb{D}$ be such that $t = \lim_{k\to\infty} s_k$. The assumptions of the theorem entail that for all $\varepsilon > 0$

$$\begin{aligned} \mathbb{P}\left(\left|X_t - \tilde{X}_t\right| > \varepsilon\right) &\leq \frac{\mathbb{E}\left[\left|X_t - \tilde{X}_t\right|^\beta\right]}{\varepsilon^\beta} = \frac{\mathbb{E}\left[\liminf_{k\to\infty} |X_t - X_{s_k}|^\beta\right]}{\varepsilon^\beta} \\ &\leq \liminf_{k\to\infty} \frac{\mathbb{E}\left[|X_t - X_{s_k}|^\beta\right]}{\varepsilon^\beta} \leq \liminf_{k\to\infty} \frac{C\,|t-s_k|^{1+\alpha}}{\varepsilon^\beta} = 0, \end{aligned}$$

where we have used the Chebyshev inequality and Fatou Lemma. We conclude that for all $t \in [0,1]$

$$\mathbb{P}(X_t = \tilde{X}_t) = 1.$$

□

Corollary 1.4.2. *Let $\{B_t, t \geq 0\}$ be a Brownian motion. There exists a version $\{\tilde{B}_t, t \geq 0\}$ of the process such that for almost all $\omega \in \Omega$ and all $T > 0$, the sample paths $t \mapsto B_t(\omega)$ are Hölder continuous on $[0,T]$ with exponent $0 < \gamma < 1/2$.*

Proof. It suffices to check that any Brownian motion satisfies the hypothesis of Theorem 1.4.1. Since $B_t - B_s \sim N(0, t-s)$,

$$\begin{aligned} \mathbb{E}\left[|B_t - B_s|^{2m}\right] &= \frac{1}{\sqrt{2\pi(t-s)}} \int_{\mathbb{R}} x^{2m} e^{-\frac{x^2}{2(t-s)}}\,dx \\ &= \frac{1}{\sqrt{2\pi(t-s)}} \int_{\mathbb{R}} y^{2m}(t-s)^m e^{-\frac{y^2}{2}} \sqrt{t-s}\,dy \\ &= C(t-s)^m, \end{aligned}$$

because the normal distribution has all moments finite. By Theorem 1.4.1, there exists a version of the process with Hölder continuous sample paths with exponent

$$0 < \gamma < \frac{m-1}{2m} = \frac{1}{2} - \frac{1}{2m}.$$

Taking $m \to \infty$,

$$0 < \gamma < \frac{1}{2}.$$

□

Remark 1.4.3. Although the definition of Brownian motion in Definition 1.3.1 did not include the continuity of sample paths, from now on whenever we refer to Brownian motion, we will be always considering a version with continuous sample paths.

1.4.2 Nowhere Differentiability

Our next goal is to prove that Brownian motion has almost surely nowhere differentiable sample paths. To do so, we begin proving that almost all sample paths are nowhere Hölder continuous with exponent $1/2 < \gamma \leq 1$.

Theorem 1.4.4. *For any $1/2 < \gamma \leq 1$, almost all sample paths of Brownian motion are nowhere Hölder continuous with exponent γ.*

Proof. We assume $T = 1$ without loss of generality. Assume that $t \mapsto B_t(\omega)$ is γ-Hölder continuous at some point $s \in [0, 1)$. That is,

$$|B_t(\omega) - B_s(\omega)| \leq C\,|t - s|^{\gamma}\,,$$

for all $t \in [0, 1]$ and $C > 0$ is a constant.
Let N be a positive integer so large that

$$N\left(\gamma - \frac{1}{2}\right) > 1.$$

We pick a large integer n and define $i = [ns] + 1$. For all $j = i, i + 1, \ldots, i + N - 1$, it holds that

$$\begin{aligned}\left|B\left(\frac{j}{n}, \omega\right) - B\left(\frac{j+1}{n}, \omega\right)\right| &\leq \left|B\left(\frac{j}{n}, \omega\right) - B\left(s, \omega\right)\right| + \left|B\left(s, \omega\right) - B\left(\frac{j+1}{n}, \omega\right)\right| \\ &\leq C\left(\left|s - \frac{j}{n}\right|^{\gamma} + \left|s - \frac{j+1}{n}\right|^{\gamma}\right),\end{aligned}$$

because $j/n, (j + 1)/n \leq 1$ whenever n is chosen large enough. Since

$$\left|s - \frac{j+1}{n}\right| \leq \left|\frac{ns - (i+N)}{n}\right| = \left|\frac{ns - ([ns] + 1 + N)}{n}\right| \leq \left|\frac{ns - ([ns] + 1 + N)}{n}\right| \leq \frac{N}{n},$$

$$\left|B\left(\frac{j}{n}, \omega\right) - B\left(\frac{j+1}{n}, \omega\right)\right| \leq 2C\left(\frac{N}{n}\right)^{\gamma} \leq \frac{M}{n^{\gamma}}$$

for some integer M. We define the set

$$A^i_{M,n} = \left\{\omega : \left|B\left(\frac{j}{n}, \omega\right) - B\left(\frac{j+1}{n}, \omega\right)\right| \leq \frac{M}{n^{\gamma}},\ \text{for all } j = i, i + 1, \ldots, i + N - 1\right\}.$$

So far, we have proved that if ω is such that $t \mapsto B_t(\omega)$ is γ-Hölder continuous at some point $s \in [0,1)$, then there exists an integer $M \in \mathbb{N}$ such that for all n big enough there exists a positive integer i such that for all $j = i, i+1, \dots, i+N-1$

$$|B_t(\omega) - B_s(\omega)| \leq \frac{M}{n^\gamma}.$$

Therefore,

$$\bigcup_{M=1}^{\infty} \bigcup_{k=1}^{\infty} \bigcap_{n=k}^{\infty} \bigcup_{i=1}^{n} A_{M,n}^{i}$$

contains all ω whose sample paths are γ-Hölder continuous at s. We will prove that the set above has null probability. Indeed, for any $M, k \in \mathbb{N}$,

$$\begin{aligned}
\mathbb{P}\left(\bigcap_{n=k}^{\infty} \bigcup_{i=1}^{n} A_{M,n}^{i}\right) &\leq \liminf_{n\to\infty} \mathbb{P}\left(\bigcup_{i=1}^{n} A_{M,n}^{i}\right) \\
&\leq \liminf_{n\to\infty} \sum_{i=1}^{n} \mathbb{P}(A_{M,n}^{i}) \\
&\leq \liminf_{n\to\infty} n \left[\mathbb{P}\left(\left|B_{\frac{1}{n}}\right| \leq \frac{M}{n^\gamma}\right)\right]^N,
\end{aligned}$$

where in the last inequality we have taken into account that $B_{\frac{j+1}{n}} - B_{\frac{j}{n}} \sim N\left(0, \frac{1}{n}\right)$ and the fact that these increments are independent. Next, we calculate the probability,

$$\mathbb{P}\left(\left|B_{\frac{1}{n}}\right| \leq \frac{M}{n^\gamma}\right) = \sqrt{\frac{n}{2\pi}} \int_{-Mn^{-\gamma}}^{Mn^{-\gamma}} e^{-n\frac{x^2}{2}} dx = \sqrt{\frac{n}{2\pi}} \int_{-Mn^{-\gamma+\frac{1}{2}}}^{Mn^{-\gamma+\frac{1}{2}}} e^{-\frac{y^2}{2}} \frac{dy}{\sqrt{n}} \leq Cn^{-\gamma+\frac{1}{2}}.$$

Then,

$$\mathbb{P}\left(\bigcap_{n=k}^{\infty} \bigcup_{i=1}^{n} A_{M,n}^{i}\right) \leq \liminf_{n\to\infty} nCn^{\left(-\gamma+\frac{1}{2}\right)N} = 0,$$

since N was chosen so that

$$\left(-\gamma + \frac{1}{2}\right) N < -1.$$

As this holds for any $M, k \in \mathbb{N}$,

$$\mathbb{P}\left(\bigcup_{M=1}^{\infty} \bigcup_{k=1}^{\infty} \bigcap_{n=k}^{\infty} \bigcup_{i=1}^{n} A_{M,n}^{i}\right) = 0.$$

We conclude that

$$\mathbb{P}\left(\{\omega : t \mapsto B_t(\omega) \text{ is } \gamma\text{-Hölder continuous at } s \in [0,1)\}\right) = 0.$$

□

Corollary 1.4.5. *Any Brownian motion has almost surely nowhere differentiable sample paths.*

Proof. If sample paths were differentiable at some point s, then they would be Hölder continuous with exponent $\gamma = 1$ at s. The previous theorem assures that this can only happen for a set of null probability. Hence, for almost all ω, $t \mapsto B_t(\omega)$ is nowhere differentiable. $\square$

The previous theorems do not allow us to say whether the Hölder continuity holds or not when $\gamma = \frac{1}{2}$. The answer is negative (see $[22, p.\ 158 - 161]$).

1.5 Quadratic Variation

In this section, we show that almost all trajectories of a Brownian motion are not of bounded variation and we derive its quadratic variation. These results are deeply related to stochastic integration. The first one makes impossible to integrate with respect to Brownian motion in a Lebesgue-Stieltjes sense, while the second one is responsible of a different change of variables rule in Itô integration.

Theorem 1.5.1. *Let $T > 0$ and $\Pi_n = \{0 = t_0 < t_1 < \ldots < t_n = T\}$ be a partition of $[0, T]$. Then,*

$$\lim_{|\Pi_n| \to 0} \sum_{k=1}^{n} \left(B_{t_k} - B_{t_{k-1}}\right)^2 = T \tag{1.5.1}$$

in $L^2(\Omega)$.

Proof. For the sake of simplicity, we denote $\Delta B_k := B_{t_k} - B_{t_{k-1}}$ and $\Delta t_k = t_k - t_{k-1}$. Then,

$$\begin{aligned} \mathbb{E}\left[\left(\sum_{k=1}^{n} \Delta B_k{}^2 - T\right)^2\right] &= \mathbb{E}\left[\left(\sum_{k=1}^{n} \Delta B_k{}^2 - \sum_{k=1}^{n} \Delta t_k\right)^2\right] \\ &= \mathbb{E}\left[\left(\sum_{k=1}^{n} \left(\Delta B_k{}^2 - \Delta t_k\right)\right)^2\right]. \end{aligned} \tag{1.5.2}$$

Note that for $1 \leq l < k \leq n$,

$$\begin{aligned} \mathbb{E}\left[\left(\Delta B_k{}^2 - \Delta t_k\right)\left(\Delta B_l{}^2 - \Delta t_l\right)\right] &= \mathbb{E}\left(\Delta B_k{}^2 \Delta B_l{}^2\right) \\ &\quad - \Delta t_l \mathbb{E}(\Delta B_k{}^2) - \Delta t_k \mathbb{E}(\Delta B_l{}^2) + \Delta t_k \Delta_l \\ &= \mathbb{E}\left(\Delta B_k{}^2\right) \mathbb{E}\left(\Delta B_l{}^2\right) - \Delta t_k \Delta t_l \\ &= 0, \end{aligned} \tag{1.5.3}$$

where we have used that $\Delta B_k{}^2 \sim N(0, \Delta t_k)$ and that Brownian motion has independent increments in disjoint intervals. Applying (1.5.3) into (1.5.2), one gets

$$\mathbb{E}\left[\left(\sum_{k=1}^{n} \Delta B_k{}^2 - T\right)^2\right] = \mathbb{E}\left[\sum_{k=1}^{n} \left(\Delta B_k{}^2 - \Delta t_k\right)^2\right]$$

$$
\begin{aligned}
&= \sum_{k=1}^{n} \left(\mathbb{E}\left(\Delta B_k{}^4\right) - 2\Delta t_k \mathbb{E}\left(\Delta B_k{}^2\right) + \Delta t_k{}^2 \right) \\
&= \sum_{k=1}^{n} \left(3\Delta t_k{}^2 - 2\Delta t_k{}^2 + \Delta t_k{}^2 \right) \\
&= 2 \sum_{k=1}^{n} \Delta t_k{}^2 \\
&\leq 2 \max_{1 \leq k \leq n} \Delta t_k \sum_{k=1}^{n} \Delta t_k \\
&= 2 \left|\Pi_n\right| T \to 0 \qquad \text{as } \left|\Pi_n\right| \to 0.
\end{aligned}
$$
□

Next, we prove that Brownian motion is not of bounded variation.

Theorem 1.5.2. *Almost all sample paths of Brownian motion are not of bounded variation.*

Proof. Assume the opposite, that is,

$$V := \sup_{\Pi} \sum_{i=1}^{n} |B(t_i) - B(t_{i-1})| < \infty,$$

where the supremum is taken over all the partitions of the interval $[0, T]$. Then,

$$
\begin{aligned}
\sum_{i=1}^{n} \left(B(t_i) - B(t_{i-1})\right)^2 &\leq \sup_{1 \leq i \leq n} |B(t_i) - B(t_{i-1})| \sum_{i=1}^{n} |B(t_i) - B(t_{i-1})| \\
&\leq V \sup_{1 \leq i \leq n} |B(t_i) - B(t_{i-1})| \to 0 \qquad (1.5.4)
\end{aligned}
$$

as $|\Pi| \to 0$ because Brownian motion has continuous sample paths with probability one. But (1.5.4) contradicts Theorem 1.5.1. □

Chapter 2

The Itô Integral

In this chapter, we study the Itô integral by giving its construction and examining remarkable properties such as the zero mean property, the martingale property and the Itô isometry. It was introduced in 1944 by Kiyosi Itô and, at that time, there was not any rigorous integral which could be used in order to solve stochastic differential equations. The main problem was the necessity of integrating with respect to Brownian motion, but, as we have seen in Chapter 1, it is not of bounded variation and, therefore, Riemman-Stieltjes integration fails.

The lack of an integration theory for this type of functions inspired Itô to construct a stochastic integral which became central to the theory of stochastic processes.

2.1 Introduction

The first attempt to construct a stochastic integral is due to Norbert Wiener. He considered a deterministic function $f : \mathbb{R} \to \mathbb{R}$ and wanted to give some sense to

$$\int_0^T f(t) dB_t(\omega).$$

His idea was to use Riemann-Stieltjes integration for each $\omega \in \Omega$ as follows

$$\int_0^T f(t) dB_t(\omega) := f(t)B(t,\omega)|_0^T - \int_0^T B_t(\omega) df(t) \tag{2.1.1}$$

The first term in (2.1.1) has sense if we assume that f is a continuous function. According to Riemann Stieltjes integration theory, the integrator must be of bounded variation. Note that as B_t has continuous sample paths almost surely, no further conditions are required.

Although Equation (2.1.1) provides a stochastic integral, we see that the kind of functions f that we can integrate with respect to Brownian motion is rather limited as we must assume that f is a deterministic continuous function of bounded variation. In general, we aim to be able to handle stochastic differential equations in which we integrate stochastic processes whose paths are not of bounded variation. For instance, we might be interested in giving some sense to

$$\int_0^T B_t(\omega) dB_t(\omega).$$

The Itô integral will allow us to integrate a wide range of stochastic processes and we will see that it has very useful properties in stochastic calculus. Another goal of this chapter is to highlight that measurability is fundamental in order to be able to integrate in Itô sense. In further chapters, we will see that this is the main shortcoming of Itô integration and it will be the main motivation to generalize the Itô integral.

2.2 The Itô Integral for Step Processes

In this chapter, $\{B_t, t \geq 0\}$ is a Brownian motion defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We fix a filtration $\{\mathcal{F}_t, t \geq 0\}$ satisfying the following properties

(i) For all $t \geq 0$, B_t is $\mathcal{F}_t$-measurable.

(ii) For all $0 \leq s \leq t$, $B_t - B_s$ is independent of $\mathcal{F}_s$.

We also fix a time horizon $T > 0$.

Definition 2.2.1. A stochastic process $X_t : [0,T] \times \Omega \to \mathbb{R}$ is $\{\mathcal{F}_t\}$-*adapted* or *non-anticipating* if for all $t \in [0,T]$ the function

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

is $\mathcal{F}_t$-measurable.

Next, we defined a class of stochastic processes for which we will first construct the Itô stochastic integral.

Definition 2.2.2. We define $L^2_{ad}([0,T] \times \Omega)$ as the set of stochastic processes X satisfying

(i) $(t,\omega) \mapsto X(t,\omega)$ is $\mathcal{B}([0,T]) \times \mathcal{F}$-measurable.

(ii) X is adapted to the filtration $\{\mathcal{F}_t, t \in [0,T]\}$.

(iii) $\int_0^T \mathbb{E}(X_t^2)\, dt < \infty$.

Remark 2.2.3. $L^2_{ad}([0,T] \times \Omega, \mathcal{B}([0,T]) \times \mathcal{F}, d\mathbb{P} \times dm)$, where m denotes the Lebesgue measure, is a Banach space with the norm

$$\|X\|_{L^2_{ad}} = \left[\int_0^T \mathbb{E}\left(X_t^2\right) dt\right]^{\frac{1}{2}}.$$

For the sake of simplicity, we will denote L^2_{ad} instead of $L^2_{ad}([0,T] \times \Omega)$.

In order to define the Itô stochastic integral in L^2_{ad}, we are going to construct it for a more restrictive type of processes and, then, extend it. We begin defining what a step process is.

Definition 2.2.4. A stochastic process $X \in L^2_{ad}$ is *a step process* if it is of the form

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

with $0 = t_0 < t_1 < \ldots < t_n = T$.

Remark 2.2.5. Note that in Definition 2.2.4, as $X \in L^2_{ad}$, for all $1 \leq i \leq n$, $\xi_{t_{i-1}}$ has to be a $\mathcal{F}_{t_{i-1}}-$measurable random variable. Indeed, given $t \in [t_{i-1}, t_i)$, $X_t = \xi_{i-1}$ is $\mathcal{F}_t-$measurable, in particular, for $t = t_{i-1}$.

Definition 2.2.6. Let $X \in L^2_{ad}$ be a step process of the form

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

with $0 = t_0 < t_1 < \ldots < t_n = T$. We define the *Itô stochastic integral* as

$$\int_0^T X_t(\omega) dB_t(\omega) := \sum_{i=1}^{n} \xi_{i-1}(\omega) \left(B_{t_i}(\omega) - B_{t_{i-1}}(\omega)\right).$$

Remark 2.2.7. We outline that the definition above is a $\mathcal{F}_T$-measurable random variable.

Proposition 2.2.8. *For all $a, b \in \mathbb{R}$ and all $X, Y \in L^2_{ad}$ two step processes, the following properties hold*

(i) $\int_0^T (aX_t + bY_t)\, dB_t = a \int_0^T X_t dB_t + b \int_0^T Y_t dB_t$ *(linearity).*

(ii) $\mathbb{E}\left(\int_0^T X_t dB_t\right) = 0$ *(zero mean).*

(iii) $\mathbb{E}\left[\left(\int_0^T X_t dB_t\right)^2\right] = \int_0^T \mathbb{E}\left(X_t^2\right) dt$ *(Itô isometry).*

Proof. (i) We first note that any linear combination of two step process is a step process. Indeed, given

$$X_t(\omega) = \sum_{i=1}^{n} \xi_{i-1} \mathbb{1}_{[t_{i-1},t_i)}(t) \text{ and } Y_t(\omega) = \sum_{j=1}^{m} \eta_{j-1} \mathbb{1}_{[s_{j-1},s_j)}(t)$$

with $0 = t_0 < t_1 < \ldots < t_n = T$ and $0 = s_0 < s_1 < \ldots < s_m = T$, we can consider a thinner partition of the interval, the union of both partitions. Then,

$$aX_t + bY_t = \sum_{k=1}^{L} (a\ \xi_{k-1} + b\ \eta_{k-1})\, \mathbb{1}_{[r_{k-1},r_k)}(t)$$

with $0 = r_0 < r_1 < \ldots < r_L = T$, where $L \leq m + n - 1$, and with the conventions that $\xi_{k-1} = \xi_{i-1}$ if $t_{i-1} \leq r_{k-1} < r_k \leq t_i$ and $\eta_{k-1} = \eta_{j-1}$ if $s_{j-1} \leq r_{k-1} < r_k \leq s_j$. Hence, $aX + bY$ is a step process. Now, we prove the linearity of the integral.

$$\begin{aligned}
\int_0^T (aX_t + bY_t)\, dB_t &:= \sum_{k=1}^{L} (a\ \xi_{k-1} + b\ \eta_{k-1}) \left(B_{r_k} - B_{r_{k-1}}\right) \\
&= a\sum_{k=1}^{L}\ \xi_{k-1} \left(B_{r_k} - B_{r_{k-1}}\right) + b\sum_{k=1}^{L} \eta_{k-1} \left(B_{r_k} - B_{r_{k-1}}\right) \\
&= a\sum_{i=1}^{n}\ \xi_{i-1} \left(B_{t_i} - B_{t_{i-1}}\right) + b\sum_{j=1}^{m} \eta_{j-1} \left(B_{s_j} - B_{s_{j-1}}\right) \\
&=: a\int_0^T X_t\ dB_t + b\int_0^T Y_t\ dB_t.
\end{aligned}$$

(ii) Since ξ_{i-1} is $\mathcal{F}_{t_{i-1}}$-measurable and $B_{t_i} - B_{t_{i-1}}$ is independent of $\mathcal{F}_{t_{i-1}}$,

$$\begin{aligned}
\mathbb{E}\left(\int_0^T X_t\ dB_t\right) &:= \mathbb{E}\left(\sum_{i=1}^{n} \xi_{i-1} \left(B_{t_i} - B_{t_{i-1}}\right)\right) \\
&= \sum_{i=1}^{n} \mathbb{E}\left(\xi_{i-1} \left(B_{t_i} - B_{t_{i-1}}\right)\right) \\
&= \sum_{i=1}^{n} \mathbb{E}\left[\mathbb{E}\left[\xi_{i-1} \left(B_{t_i} - B_{t_{i-1}}\right) \middle| \mathcal{F}_{t_{i-1}}\right]\right] \\
&= \sum_{i=1}^{n} \mathbb{E}\left[\xi_{i-1}\mathbb{E}\left[\left(B_{t_i} - B_{t_{i-1}}\right) \middle| \mathcal{F}_{t_{i-1}}\right]\right] \\
&= \sum_{i=1}^{n} \mathbb{E}\left[\xi_{i-1}\mathbb{E}\left[\left(B_{t_i} - B_{t_{i-1}}\right)\right]\right] = 0.
\end{aligned}$$

(iii) By the definition of the Itô integral and the linearity of the expectation,

$$\mathbb{E}\left[\left(\int_0^T X_t dB_t\right)^2\right] = \sum_{i,j=1}^{n} \mathbb{E}\left[\xi_{i-1}\xi_{j-1} \left(B_{t_i} - B_{t_{i-1}}\right) \left(B_{t_j} - B_{t_{j-1}}\right)\right].$$

Firstly, consider $i \neq j$, suppose $i < j$, we have

$$\begin{aligned}
&\mathbb{E}\left[\xi_{i-1}\xi_{j-1} \left(B_{t_i} - B_{t_{i-1}}\right) \left(B_{t_j} - B_{t_{j-1}}\right)\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[\xi_{i-1}\xi_{j-1} \left(B_{t_i} - B_{t_{i-1}}\right) \left(B_{t_j} - B_{t_{j-1}}\right) \middle| F_{t_{j-1}}\right]\right] \\
&= \mathbb{E}\left[\xi_{i-1}\xi_{j-1} \left(B_{t_i} - B_{t_{i-1}}\right) \mathbb{E}\left[\left(B_{t_j} - B_{t_{j-1}}\right) \middle| F_{t_{j-1}}\right]\right] \\
&= \mathbb{E}\left[\xi_{i-1}\xi_{j-1} \left(B_{t_i} - B_{t_{i-1}}\right) \mathbb{E}\left[B_{t_j} - B_{t_{j-1}}\right]\right] \\
&= 0,
\end{aligned}$$

where we have used that ξ_{i-1},ξ_{j-1} and $(B_{t_i} - B_{t_{i-1}})$ are $\mathcal{F}_{t_{i-1}}$−measurable. Hence,

$$
\begin{aligned}
\mathbb{E}\left[\left(\int_0^T X_t dB_t\right)^2\right] &= \sum_{i=1}^n \mathbb{E}\left[\xi_{i-1}^2 \left(B_{t_i} - B_{t_{i-1}}\right)^2\right] \\
&= \sum_{i=1}^n \mathbb{E}\left[\mathbb{E}\left[\xi_{i-1}^2 \left(B_{t_i} - B_{t_{i-1}}\right)^2 \middle| \mathcal{F}_{t_{i-1}}\right]\right] \\
&= \sum_{i=1}^n \mathbb{E}\left[\xi_{i-1}^2 \mathbb{E}\left[\left(B_{t_i} - B_{t_{i-1}}\right)^2 \middle| \mathcal{F}_{t_{i-1}}\right]\right] \\
&= \sum_{i=1}^n \mathbb{E}\left[\xi_{i-1}^2 \mathbb{E}\left[\left(B_{t_i} - B_{t_{i-1}}\right)^2\right]\right] \\
&= \sum_{i=1}^n \mathbb{E}\left[\xi_{i-1}^2\right](t_i - t_{i-1}) \\
&= \sum_{i=1}^n \int_{t_{i-1}}^{t_i} \mathbb{E}\left[\xi_{i-1}^2\right] dt \\
&= \int_0^T \mathbb{E}\left[\sum_{i=1}^n \xi_{i-1}^2 \mathbb{1}_{[t_{i-1},t_i)}(t)\right] dt \\
&= \int_0^T \mathbb{E}\left[X_t^2\right] dt. \qquad \square
\end{aligned}
$$

2.3 The Itô Integral in L^2_{ad}

The aim of this section is to define the Itô stochastic integral in L^2_{ad}. To this purpose, we need to prove that step processes, introduced in the previous section, are dense in L^2_{ad}.

Lemma 2.3.1. *For any $X \in L^2_{ad}$, there exists a sequence of step processes $\{X^{(n)}\}_n$ in L^2_{ad} such that*

$$\lim_{n\to\infty} \int_0^T \mathbb{E}\left[(X_t^{(n)} - X_t)^2\right] dt = 0.$$

Proof. We divide the proof into three steps.

(1) We assume that there exists a positive constant $C > 0$ such that $|X(t,\omega)| \leq C$ for all $(t,\omega) \in [0,T] \times \Omega$ and $X(\cdot,\omega)$ is continuous for almost all $\omega \in \Omega$. We want to show that such a process can be approximated by step processes in L^2_{ad}. Indeed, we define

$$X^{(n)}(t,\omega) := \sum_{k=0}^{[nT]} X\left(\frac{k}{n},\omega\right) \mathbb{1}_{\left[\frac{k}{n},\frac{k+1}{n}\wedge T\right)}(t).$$

$X^{(n)}$ is adapted and $\mathcal{F} \times \beta([0,T])$-measurable because $X \in L^2_{ad}$ and it is evaluated at the left endpoints of the intervals. Moreover, $X^{(n)}$ is also bounded, which ensures that the L^2_{ad}-norm is finite. Hence, $X^{(n)} \in L^2_{ad}$. Next, we see that $\{X^{(n)}\}_n \subset L^2_{ad}$ converges to X. In order to apply the dominated convergence theorem,

$$\lim_{n\to\infty} \int_0^T \left|X^{(n)}(t) - X(t)\right|^2 dt = \lim_{n\to\infty} \sum_{k=0}^{[nT]} \int_{\frac{k}{n}}^{\frac{k+1}{n}\wedge T} \left|X\left(\frac{k}{n}\right) - X(t)\right|^2 dt$$
$$\leq \lim_{n\to\infty} T \sup_{0\leq k\leq [nT]} \sup_{t\in\left[\frac{k}{n},\frac{k+1}{n}\wedge T\right]} \left|X\left(\frac{k}{n}\right) - X(t)\right|^2 = 0,$$

almost surely, where we have used the fact that X has continuous sample paths, almost surely. By the dominated convergence theorem (dominating by a constant),

$$\lim_{n\to\infty} \mathbb{E}\left(\int_0^T \left|X^{(n)}(t) - X(t)\right|^2 dt\right) = \mathbb{E}\left(\lim_{n\to\infty} \int_0^T \left|X^{(n)}(t) - X(t)\right|^2 dt\right) = 0.$$

(2) Now we assume that X is bounded and we want to prove that it can be approximated by bounded L^2_{ad} processes with almost surely continuous sample paths. If we prove so, we are also proving that X can be approximated by step processes. The family of functions $f_n(t) = n\mathbb{1}_{[0,1/n]}(t)$ is an approximation of the identity. Indeed, it is clear that they are positive, they integrate one and, for a given $\delta > 0$, there exists a positive integer $n > \frac{1}{\delta}$ so that

$$\int_{t>\delta} f_n(t)dt = 0.$$

For each positive integer n, we define

$$X^{(n)}(t,\omega) = f_n * X(t,\omega) = \int_{-\infty}^{\infty} n\mathbb{1}_{[0,1/n]}(t-s)X(s,\omega)ds = n\int_{t-\frac{1}{n}}^{t} X(s,\omega)ds.$$

(i) Since X is adapted and we are integrating up to time t, $X^{(n)}$ is adapted and $\mathcal{F}\times\mathcal{B}([0,T])$-measurable.

(ii) Since X is bounded by a constant $C > 0$, we have that $|X^{(n)}(t,\omega)| \leq C$.

(iii) $X^{(n)}$ has continuous sample paths. Given $0 \leq t_1 \leq t_2 \leq T$,

$$\left|X^{(n)}(t_1,\omega) - X^{(n)}(t_2,\omega)\right| = n\left|\int_{t_1-\frac{1}{n}}^{t_1} X(s,\omega)ds - \int_{t_2-\frac{1}{n}}^{t_2} X(s,\omega)ds\right|$$
$$\leq 2Cn\left|t_2 - t_1\right|.$$

(iv) $X^{(n)}$ has finite L^2_{ad}-norm,

$$\int_0^T \mathbb{E}[(X^{(n)})^2]dt \leq C^2 T < \infty.$$

Taking into account that $\{f_n\}_n$ is an approximation of the identity, it holds that

$$\lim_{n\to\infty} \int_0^T \left|X^{(n)}(t,\omega) - X(t,\omega)\right|^2 dt = 0 \quad (a.s.).$$

By the dominated convergence theorem,

$$\lim_{n\to\infty} \mathbb{E}\left[\int_0^T \left|X^{(n)}(t) - X(t)\right|^2 dt\right] = \mathbb{E}\left[\lim_{n\to\infty} \int_0^T \left|X^{(n)}(t,\omega) - X(t,\omega)\right|^2 dt\right] = 0.$$

(iii) Finally, we assume that $X \in L^2_{ad}$. We aim to show that X can be approximated by bounded processes belonging to L^2_{ad}. If we prove so, then X can be approximated by L^2_{ad}-step processes according to what we have proved so far. For each positive integer n, we define

$$X^{(n)}(t,\omega) = \begin{cases} 0, & |X(t,\omega)| > n, \\ X(t,\omega), & |X(t,\omega)| \leq n. \end{cases}$$

Since $X \in L^2_{ad}$, it is clear that $X^{(n)} \in L^2_{ad}$ and, by construction, is bounded. By the dominated convergence theorem,

$$\lim_{n\to\infty} \mathbb{E}\left[\int_0^T \left|X^{(n)}(t) - X(t)\right|^2 dt\right] = \mathbb{E}\left[\lim_{n\to\infty} \int_0^T |X(t)|^2 \mathbb{1}_{\{|X_t|>n\}} dt\right] = 0. \quad \square$$

Once we have a definition of the Itô stochastic integral for step process and we have shown that they are dense in L^2_{ad}, we now aim to extend the integral to L^2_{ad}.

Definition 2.3.2. The *Itô stochastic integral* of a process $X \in L^2_{ad}$ is defined as

$$\int_0^T X_t dB_t := L^2(\Omega) - \lim_{n\to\infty} \int_0^T X_t^{(n)} dB_t,$$

where $\{X^{(n)}\}_n$ is a sequence of step processes in L^2_{ad} which approximates X, i.e.,

$$\lim_{n\to\infty} \int_0^T \mathbb{E}\left[\left(X_t - X_t^{(n)}\right)^2\right] dt = 0.$$

Remark 2.3.3. We need to check that such a limit exists. We will show that the sequence is Cauchy in $L^2(\Omega)$. Indeed,

$$\begin{aligned} \lim_{m,n\to\infty} \mathbb{E}\left[\left(\int_0^T X_t^{(m)} dB_t - \int_0^T X_t^{(n)} dB_t\right)^2\right] &= \lim_{m,n\to\infty} \mathbb{E}\left[\left(\int_0^T \left(X_t^{(m)} - X_t^{(n)}\right) dB_t\right)^2\right] \\ &= \lim_{m,n\to\infty} \int_0^T \mathbb{E}\left[\left(X_t^{(m)} - X_t^{(n)}\right)^2\right] dt \\ &= 0, \end{aligned}$$

where we have used the Itô isometry for step processes and the fact that $\{X^{(n)}\}_n$ is convergent in L^2_{ad} and, hence, Cauchy. Since $L^2(\Omega)$ is complete, the sequence of integrals converges. Moreover, the limit is $\mathcal{F}_T$-measurable.

Remark 2.3.4. The definition does not depend on the approximating sequence. Consider two sequences $\{X^{(n)}\}_n$ and $\{Y^{(n)}\}_n$ of step processes in L^2_{ad} which approximate $X \in L^2_{ad}$. That is,

$$\lim_{n\to\infty}\int_0^T \mathbb{E}\left[\left(X_t - X_t^{(n)}\right)^2\right]dt = 0 \qquad \text{and} \qquad \lim_{n\to\infty}\int_0^T \mathbb{E}\left[\left(X_t - Y_t^{(n)}\right)^2\right]dt = 0.$$

We denote

$$I_1 = L^2(\Omega) - \lim_{n\to\infty}\int_0^T X_t^{(n)}dB_t \qquad \text{and} \qquad I_2 = L^2(\Omega) - \lim_{n\to\infty}\int_0^T Y_t^{(n)}dB_t.$$

Then,

$$\begin{aligned}\|I_1 - I_2\|_{L^2(\Omega)} \le &\left\|I_1 - \int_0^T X_t^{(n)}dB_t\right\|_{L^2(\Omega)} + \left\|\int_0^T X_t^{(n)}dB_t - \int_0^T Y_t^{(n)}dB_t\right\|_{L^2(\Omega)} \\ &+ \left\|I_2 - \int_0^T Y_t^{(n)}dB_t\right\|_{L^2(\Omega)}.\end{aligned}$$

As $n \to \infty$, the first and last terms vanish by definition. We estimate the second one,

$$\begin{aligned}&\lim_{n\to\infty}\left\|\int_0^T X_t^{(n)}dB_t - \int_0^T Y_t^{(n)}dB_t\right\|^2_{L^2(\Omega)} = \lim_{n\to\infty}\mathbb{E}\left[\left(\int_0^T \left(X_t^{(n)} - Y_t^{(n)}\right)dB_t\right)^2\right] \\ &= \lim_{n\to\infty}\int_0^T \mathbb{E}\left[\left(X_t^{(n)} - Y_t^{(n)}\right)^2\right]dt = \lim_{n\to\infty}\int_0^T \mathbb{E}\left[\left(X_t^{(n)} - X_t + X_t - Y_t^{(n)}\right)^2\right]dt \\ &\le 2\lim_{n\to\infty}\int_0^T \mathbb{E}\left[\left(X_t^{(n)} - X_t\right)^2\right]dt + 2\lim_{n\to\infty}\int_0^T \mathbb{E}\left[\left(Y_t^{(n)} - X_t\right)^2\right]dt = 0.\end{aligned}$$

Hence, $\|I_1 - I_2\|_{L^2(\Omega)} = 0$, which means that $I_1 = I_2$ almost surely. Thus, the limit does not depend on the approximating sequence.

As we did for step process, our next goal is to extend the linearity, the zero mean property and the Itô isometry for L^2_{ad} processes.

Theorem 2.3.5. *Let X, Y be two stochastic processes in L^2_{ad}. The following holds:*

(i) For any $a, b \in \mathbb{R}$, $\int_0^T (aX_t + bY_t)dB_t = a\int_0^T X_t dB_t + b\int_0^T Y_t dB_t$ (linearity).

(ii) $\mathbb{E}\left(\int_0^T X_t dB_t\right) = 0$ (zero mean).

(iii) $\mathbb{E}\left[\left(\int_0^T X_t dB_t\right)^2\right] = \int_0^T \mathbb{E}\left(X_t^2\right)dt$ (Itô isometry).

(iv) $\mathbb{E}\left[\left(\int_0^T X_t dB_t\right)\left(\int_0^T Y_t dB_t\right)\right] = \int_0^T \mathbb{E}\left(X_t Y_t\right)dt$ (Itô isometry).

Proof. We consider $\{X^{(n)}\}_n$ and $\{Y^{(n)}\}_n$ two sequences of step processes in L^2_{ad} so that they approximate X and Y, respectively.

(i) Taking into account that the sum of two step processes is a step process and Proposition 2.2.8,

$$\begin{aligned}\int_0^T (aX_t + bY_t)dB_t &:= L^2(\Omega) - \lim_{n\to\infty} \int_0^T (aX_t^{(n)} + bY_t^{(n)})dB_t \\ &= a\ L^2(\Omega) - \lim_{n\to\infty} \int_0^T X_t^{(n)} dB_t + b\ L^2(\Omega) - \lim_{n\to\infty} \int_0^T Y_t^{(n)} dB_t \\ &=: a \int_0^T X_t dB_t + b \int_0^T Y_t dB_t.\end{aligned}$$

(ii) By the Hölder inequality,

$$\lim_{n\to\infty} \left[\mathbb{E} \left| \int_0^T (X_t - X_t^{(n)})dB_t \right| \right] \leq \lim_{n\to\infty} \mathbb{E} \left[\left| \int_0^T (X_t - X_t^{(n)})dB_t \right|^2 \right]^{1/2} = 0. \quad (2.3.1)$$

Then, by Proposition (2.2.8),

$$\mathbb{E}\left(\int_0^T X_t dB_t \right) = \lim_{n\to\infty} \mathbb{E}\left(\int_0^T X_t^{(n)} dB_t \right) = 0.$$

(iii) Taking into account that the property holds for step processes and (2.3.1), we have

$$\begin{aligned}\mathbb{E}\left[\left(\int_0^T X_t dB_t \right)^2 \right] &= \lim_{n\to\infty} \mathbb{E}\left[\left(\int_0^T X_t^{(n)} dB_t \right)^2 \right] = \lim_{n\to\infty} \int_0^T \mathbb{E}\left[\left(X_t^{(n)} \right)^2 \right] dt \\ &= \int_0^T \mathbb{E}\left[(X_t)^2 \right] dt.\end{aligned}$$

(iv) Using that for any $a, b \in \mathbb{R}$ $2ab = (a+b)^2 - a^2 - b^2$, (i) and (iii), we have

$$\begin{aligned}&2\mathbb{E}\left[\int_0^T X_t dB_t \int_0^T Y_t dB_t \right] \\ &= \mathbb{E}\left[\left(\int_0^T X_t dB_t + \int_0^T Y_t dB_t \right)^2 - \left(\int_0^T X_t dB_t \right)^2 - \left(\int_0^T Y_t dB_t \right)^2 \right] \\ &= \mathbb{E}\left[\left(\int_0^T (X_t + Y_t) dB_t \right)^2 \right] - \mathbb{E}\left[\left(\int_0^T X_t dB_t \right)^2 \right] - \mathbb{E}\left[\left(\int_0^T Y_t dB_t \right)^2 \right] \\ &= \int_0^T \mathbb{E}\left[(X_t + Y_t)^2 \right] dt - \int_0^T \mathbb{E}\left[(X_t)^2 \right] dt - \int_0^T \mathbb{E}\left[(Y_t)^2 \right] dt \\ &= 2 \int_0^T \mathbb{E}\left[X_t Y_t \right] dt.\end{aligned}$$

□

2.4 The Itô Integral as a Stochastic Process

So far, we have defined the Itô integral as a random variable. Now, we would like to define it as a stochastic process and we will study two main properties, the martingale property and the existence of a version with continuous sample paths almost surely.

Definition 2.4.1. Let $X \in L^2_{ad}$, we define its *indefinite Itô stochastic integral* as

$$\int_0^t X_s dB_s := \int_0^T X_s \mathbb{1}_{[0,t]}(s) dB_s.$$

for all $0 \leq t \leq T$.

Remark 2.4.2. We need to assure that the right-hand side makes sense, that is, to check that $\{X_s \mathbb{1}_{[0,t]}(s), s \in [0,T]\} \in L^2_{ad}$. Since $X \in L^2_{ad}$ and $\mathbb{1}_{[0,t]}$ is $\mathcal{B}([0,T]) \times \mathcal{F}$-measurable,

$$(s, \omega) \mapsto X(s, \omega)\mathbb{1}_{[0,t]}(s)$$

is $\mathcal{B}([0,T]) \times \mathcal{F}$-measurable. Likewise, $\omega \mapsto X_s(\omega)\mathbb{1}_{[0,t]}(s)$ is $\mathcal{F}_s$-measurable. Finally,

$$\int_0^T \mathbb{E}\left[\left|X_s \mathbb{1}_{[0,t]}\right|^2\right] ds \leq \int_0^T \mathbb{E}\left[|X_s|^2\right] ds < \infty.$$

Therefore, $\{X_s \mathbb{1}_{[0,t]}(s), s \in [0,T]\} \in L^2_{ad}$ and the definition is meaningful.

Lemma 2.4.3. *Let $X \in L^2_{ad}$ be a step process and $0 \leq s \leq t \leq T$. Then,*

$$\int_0^t X_r dB_r = \int_0^s X_r dB_r + \int_s^t X_r dB_r.$$

Proof. Assume that X is of the form,

$$X_r(\omega) = \sum_{i=1}^n \xi_{i-1}(\omega)\mathbb{1}_{[t_{i-1},t_i)}(r)$$

with $0 = t_0 < t_1 < \ldots < t_n = T$. Then, $s \in [t_{k-1}, t_k]$ and $t \in [t_{l-1}, t_l]$ for some integers $0 \leq k \leq l \leq n$. Note that $X\mathbb{1}_{[0,t]}$ is again a step process of the form

$$X_r(\omega)\mathbb{1}_{[0,t]}(r) = \sum_{i=1}^{l-1} \xi_{i-1}(\omega)\mathbb{1}_{[t_{i-1},t_i)}(r) + \xi_{l-1}(\omega)\mathbb{1}_{[t_{l-1},t]}(r).$$

Therefore,

$$\begin{aligned}
\int_0^t X_r dB_r &:= \int_0^T X_r \mathbb{1}_{[0,t]}(r) dB_r \\
&= \sum_{i=1}^{l-1} \xi_{i-1}\left(B(t_i) - B(t_{i-1})\right) + \xi_{l-1}(B(t) - B(t_{l-1}))
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{i=1}^{k-1} \xi_{i-1} \left(B(t_i) - B(t_{i-1}) \right) + \xi_{k-1} \left[B(t_k) - B(s) + B(s) - B(t_{k-1}) \right] \\
&\quad + \sum_{i=k}^{l-1} \xi_{i-1} \left(B(t_i) - B(t_{i-1}) + \xi_{l-1} (B(t) - B(t_{l-1})) \right. \\
&= \int_0^T X_r \mathbb{1}_{[0,s]}(r) dB_r + \int_0^T X_r \mathbb{1}_{[0,t]}(r) dB_r \\
&=: \int_0^s X_r dB_r + \int_s^t X_r dB_r.
\end{aligned}
$$

□

Theorem 2.4.4. *Let $X \in L^2_{ad}$. Then, the stochastic process*

$$
I_t = \int_0^t X_s dB_s, \qquad t \in [0, T],
$$

is a martingale with respect to the filtration $\{\mathcal{F}_t, t \geq 0\}$.

Proof. We want to prove that for any $0 \leq s \leq t \leq T$, it holds $\mathbb{E}[I_t \mid \mathcal{F}_s] = I_s$ almost surely. We first assume that $X \in L^2_{ad}$ is a step process. By Lemma 2.4.3,

$$
I_t = I_s + \int_s^t X_r dB_r.
$$

By the linearity of the conditional expectation and the fact that I_s is $\mathcal{F}_s$-measurable, it is enough to prove that

$$
\mathbb{E}\left[\int_s^t X_r dB_r \mid \mathcal{F}_s \right] = 0 \text{ (a.s.)}.
$$

Note that X is of the form

$$
X_r(\omega) = \sum_{i=1}^n \xi_{i-1}(\omega) \mathbb{1}_{[t_{i-1}, t_i)}(r),
$$

where $s = t_0 < t_1 < \ldots < t_n = t$ and ξ_{i-1} is $\mathcal{F}_{t_{i-1}}$-measurable. Then,

$$
\begin{aligned}
\mathbb{E}\left[\int_s^t X_r dB_r \mid \mathcal{F}_s \right] &= \sum_{i=1}^n \mathbb{E}\left[\xi_{i-1} \left(B(t_i) - B(t_{i-1}) \right) \mid \mathcal{F}_s \right] \\
&= \sum_{i=1}^n \mathbb{E}\left[\mathbb{E}\left[\xi_{i-1} \left(B(t_i) - B(t_{i-1}) \right) \mid \mathcal{F}_{t_{i-1}} \right] \mid \mathcal{F}_s \right] \\
&= \sum_{i=1}^n \mathbb{E}\left[\xi_{i-1} \mathbb{E}\left[\left(B(t_i) - B(t_{i-1}) \right) \mid \mathcal{F}_{t_{i-1}} \right] \mid \mathcal{F}_s \right] \\
&= \sum_{i=1}^n \mathbb{E}\left[\xi_{i-1} \mathbb{E}\left[B(t_i) - B(t_{i-1}) \right] \mid \mathcal{F}_s \right] = 0,
\end{aligned}
$$

where we have used that ξ_{i-1} is $\mathcal{F}_{t_{i-1}}$-measurable and $B(t_i) - B(t_{i-1})$ is independent of $\mathcal{F}_{t_{i-1}}$.

Now, assume that $X \in L^2_{ad}$. By Lemma 2.3.1, there exists a sequence of step processes $\{X^{(n)}\}_n$ in L^2_{ad} such that

$$\lim_{n\to\infty} \int_0^T \mathbb{E}\left[\left|X_r - X_r^{(n)}\right|^2\right] dr = 0.$$

For each n, we define the following stochastic process

$$I_t^{(n)} = \int_0^t X_r^{(n)} dB_r, \qquad t \in [0,T].$$

For $0 \le s \le t \le T$,

$$I_t - I_s = I_t - I_t^{(n)} + I_t^{(n)} - I_s^{(n)} + I_s^{(n)} - I_s.$$

Taking conditional expectations on both sides, we have

$$\mathbb{E}\left[I_t - I_s \mid \mathcal{F}_s\right] = \mathbb{E}\left[I_t - I_t^{(n)} \mid \mathcal{F}_s\right] + \mathbb{E}\left[I_t^{(n)} - I_s^{(n)} \mid \mathcal{F}_s\right] + \mathbb{E}\left[I_s^{(n)} - I_s \mid \mathcal{F}_s\right].$$

According to the previous case, $\{I_t^{(n)}, 0 \le t \le T\}$ is a martingale for each $n \in \mathbb{N}$. Hence,

$$\mathbb{E}\left[I_t - I_s \mid \mathcal{F}_s\right] = \mathbb{E}\left[I_t - I_t^{(n)} \mid \mathcal{F}_s\right] + \mathbb{E}\left[I_s^{(n)} - I_s \mid \mathcal{F}_s\right]. \tag{2.4.1}$$

We take $L^2(\Omega)$-limits. For the first term on the right-hand side of (2.4.1), we have

$$\begin{aligned}
\lim_{n\to\infty} \mathbb{E}\left[\left(\mathbb{E}\left[I_t - I_t^{(n)}\middle| \mathcal{F}_s\right]\right)^2\right] &\le \lim_{n\to\infty} \mathbb{E}\left[\mathbb{E}\left[\left(I_t - I_t^{(n)}\right)^2\middle| \mathcal{F}_s\right]\right] \\
&= \lim_{n\to\infty} \mathbb{E}\left[\left(\int_0^t \left(X_r - X_r^{(n)}\right) dB_r\right)^2\right] \\
&= \lim_{n\to\infty} \int_0^t \mathbb{E}\left[\left|X_r - X_r^{(n)}\right|^2\right] dr \\
&\le \lim_{n\to\infty} \int_0^T \mathbb{E}\left[\left|X_r - X_r^{(n)}\right|^2\right] dr = 0,
\end{aligned}$$

where the first inequality is the conditional Jensen inequality.
So, $L^2(\Omega) - \lim_{n\to\infty} \mathbb{E}\left[I_t - I_t^{(n)} \mid \mathcal{F}_s\right] = 0$. As a result, there exists a subsequence such that

$$\lim_{k\to\infty} \mathbb{E}\left[I_t - I_t^{(n_k)} \mid \mathcal{F}_s\right] = 0$$

almost surely. Since the whole proof can be rewritten considering the subsequence instead of the sequence, we can assume that the subsequence coincides with the sequence. Likewise,

$$\lim_{n\to\infty} \mathbb{E}\left[I_s - I_s^{(n)} \mid \mathcal{F}_s\right] = 0$$

almost surely. Therefore, taking limits in (2.4.1), we obtain that $\mathbb{E}\left[I_t - I_s \middle| \mathcal{F}_s\right] = 0$ almost surely. □

Next, we show that an Itô process always has a version with continuous sample paths.

Theorem 2.4.5. *Let $X \in L^2_{ad}$. Then, for almost all $\omega \in \Omega$, the*

$$t \mapsto I(t,\omega) = \int_0^t X_s(\omega) dB_s(\omega)$$

is continuous.

Proof. Let $\{X^{(n)}\}_n$ be a sequence of step processes in L^2_{ad} such that

$$\lim_{n\to\infty} \int_0^T \mathbb{E}\left(\left|X_t^{(n)} - X_t\right|^2\right) dt = 0.$$

For each n,

$$I_t^{(n)} = \int_0^t X_s^{(n)} dB_s := \sum_{j=0}^{k-1} X_j^{(n)} \left(B(t_{j+1}^{(n)}) - B(t_j^{(n)})\right) + X_k^{(n)} \left(B(t) - B(t_k^{(n)})\right).$$

Since Brownian motion has almost all sample paths continuous, the mapping $t \mapsto I_t^{(n)}(\omega)$ is continuous for almost all $\omega \in \Omega$. Thus, we aim to prove that $I^{(n)}$ converges uniformly to I as $n \to \infty$. By Theorem 2.4.4, $\{I_t^{(n)}, t \geq 0\}$ is a martingale for each n. By Proposition 1.1.15, $\{|I^{(n)} - I^{(m)}|^2, t \geq 0\}$ is a submartingale for each pair of integers m, n because $f(x) = x^2$ is a convex function and $I^{(n)} \in L^2(\Omega)$. Moreover, by the continuous Doob submartingale inequality (Theorem $B.2$),

$$\begin{aligned}
\mathbb{P}\left(\sup_{0\leq t\leq T} \left|I_t^{(n)} - I_t^{(m)}\right| > \varepsilon\right) &= \mathbb{P}\left(\sup_{0\leq t\leq T} \left|I_t^{(n)} - I_t^{(m)}\right|^2 > \varepsilon^2\right) \\
&\leq \frac{1}{\varepsilon^2}\mathbb{E}\left(\left|I_T^{(n)} - I_T^{(m)}\right|^2\right) \\
&= \frac{1}{\varepsilon^2}\mathbb{E}\left(\left|\int_0^T \left(X_t^{(n)} - X_t^{(m)}\right) dB_t\right|^2\right) \\
&= \frac{1}{\varepsilon^2} \int_0^T \mathbb{E}\left(\left|X_t^{(n)} - X_t^{(m)}\right|^2\right) dt \to 0 \qquad \text{as } n, m \to \infty.
\end{aligned}$$

In particular, for $\varepsilon = \frac{1}{2^k}$, there exists n_k such that for all $m, n \geq n_k$,

$$\mathbb{P}\left(\sup_{0\leq t\leq T} \left|I_t^{(n)} - I_t^{(m)}\right| > \frac{1}{2^k}\right) \leq \frac{1}{k^2}. \tag{2.4.2}$$

Since the probability increases and the bound decreases with k, we can assume that $n_{k+1} \geq n_k \geq n_{k-1} \geq \ldots$ and that $n_k \to \infty$. We define for each positive integer k the set

$$A_k := \left\{\sup_{0\leq t\leq T} \left|I_t^{(n_k)} - I_t^{(n_{k+1})}\right| > \frac{1}{2^k}\right\}.$$

By (2.4.2), we get that

$$\sum_{k=1}^{\infty} \mathbb{P}(A_k) \le \sum_{k=1}^{\infty} \frac{1}{k^2} < \infty.$$

By the Borel-Cantelli lemma,

$$\mathbb{P}(\liminf_k A_k^c) = 1.$$

That is, for almost all $\omega \in \Omega$, there exists $k_0(\omega)$ such that for all $k \ge k_0(\omega)$,

$$\sup_{0 \le t \le T} \left| I^{(n_k)}(t,\omega) - I^{(n_{k+1})}(t,\omega) \right| \le \frac{1}{2^k}.$$

Therefore, for almost all $\omega \in \Omega$, $\{I^{(n_k)}(\cdot,\omega)\}_k$ is Cauchy uniformly, hence, converges uniformly on $[0,T]$. This implies that there exists

$$J_t(\omega) = \lim_k I_t^{(n_k)}(\omega), \ t \in [0,T],$$

where the limit is uniform and, hence, J_t is pathwise continuous almost surely. On the other hand, we know that

$$I_t := L^2(\Omega) - \lim_n \int_0^t X_s^n dB_s,$$

which ensures that there exists a subsequence converging almost surely. The uniqueness of the limit implies that

$$\mathbb{P}(I_t = J_t) = 1.$$

Hence, we have found a version of $I = \{I_t, 0 \le t \le T\}$ with continuous sample paths almost surely. □

2.5 The Itô Integral in $\mathcal{L}_{ad}$

In this section, we aim to extend the Itô integral to a wider class of integrands.
We continue considering a filtration $\{\mathcal{F}_t, t \ge 0\}$ satisfying the following properties

(i) For all $t \ge 0$, B_t is $\mathcal{F}_t$-measurable.

(ii) For all $0 \le s \le t$, $B_t - B_s$ is independent of $\mathcal{F}_s$.

Definition 2.5.1. We define $\mathcal{L}_{ad}\left(\Omega, L^2[0,T]\right)$ as the set of stochastic processes X satisfying

(i) $(t,\omega) \mapsto X(t,\omega)$ is $\mathcal{B}\left([0,T]\right) \times \mathcal{F}$-measurable.

(ii) X is adapted to the filtration $\mathcal{F}_t$.

(iii) $\mathbb{P}\left(\int_0^T X_t^2 dt < \infty\right) = 1$.

Observe that we are changing condition (iii) in comparison with the definition of L^2_{ad} and this change means that $X(\cdot,\omega) \in L^2([0,T])$ for almost all $\omega \in \Omega$.

Remark 2.5.2. $L^2_{ad} \subset \mathcal{L}_{ad}$. Indeed, given $X \in L^2_{ad}$, by definition,

$$\mathbb{E}\left[\int_0^T |X_t|^2 \, dt\right] = \int_0^T \mathbb{E}\left[|X_t|^2\right] dt < \infty.$$

Therefore,

$$\int_0^T |X_t|^2 \, dt < \infty$$

for almost all $\omega \in \Omega$. The converse is not true. For instance, we can pick $X_t = e^{B_t^2}$ with $0 \le t \le 1$. Then,

$$\mathbb{E}\left(|X_t|^2\right) = \mathbb{E}\left(e^{2B_t^2}\right) = \int_{\mathbb{R}} e^{2x^2} \frac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}} \, dx = \frac{2}{\sqrt{2\pi t}} \int_0^\infty e^{\left(2-\frac{1}{2t}\right)x^2} dx = \infty,$$

whenever $t > \frac{1}{4}$. Therefore, $\int_0^1 \mathbb{E}\left(|X_t|^2\right) dt = \infty$. We have that $X \notin L^2_{ad}$, but

$$\int_0^1 |X_t|^2 dt < \infty \quad \text{a.s.}$$

because X_t is a continuous function of Brownian motion and any Brownian motion has continuous sample paths almost surely.

Lemma 2.5.3. *Let $X \in \mathcal{L}_{ad}$. There exists a sequence of processes $\{X_n\}_n \subset L^2_{ad}$ such that*

$$\lim_{n\to\infty} \int_0^T |X_n(t) - X(t)|^2 \, dt = 0$$

for almost all $\omega \in \Omega$ and, hence, in probability.

Proof. For each positive integer n, we define

$$X_n(t,\omega) = \begin{cases} X(t,\omega), & \text{if } \int_0^t |X(s,\omega)|^2 \, ds \le n, \\ 0, & \text{otherwise.} \end{cases}$$

Clearly, X_n is adapted as $X \in \mathcal{L}_{ad}$. We consider the following stopping times,

$$\tau_n(\omega) = \sup\left\{t \in [0,T] : \int_0^t |X(s,\omega)|^2 \, ds \le n\right\}.$$

Then,

$$\int_0^T |X_n(t,\omega)|^2 \, dt = \int_0^{\tau_n(\omega)} |X(t,\omega)|^2 \, dt \le n.$$

Taking conditional expectations on both sides, $\int_0^T \mathbb{E}\,|X_n(t,\omega)|^2 \, dt \le n$. Hence, $X_n \in L^2_{ad}$. Now, we need to show that X_n approximates X. By assumption, there exists $N \subset \Omega$ such that $\mathbb{P}(N) = 0$ and that for all $\omega \in N^c$

$$\int_0^T |X(t,\omega)|^2 \, dt < \infty.$$

For a fixed $\omega \in N^c$, we can pick n so large that

$$\int_0^T |X(t,\omega)|^2 \, dt \leq n.$$

By construction, we have that $X_n(t,\omega) = X(t,\omega)$ for all $t \in [0,T]$. Since they coincide,

$$\lim_{n\to\infty} \int_0^T |X_n(t) - X(t)|^2 \, dt = 0.$$

As this holds for all $\omega \in N^c$, the lemma is proved. □

Our next lemma provides an inequality which will be useful from now on in order to extend the integral.

Lemma 2.5.4. *Let $\{X_t, 0 \leq t \leq T\}$ be a step process in L^2_{ad}. The following inequality*

$$\mathbb{P}\left(\left|\int_0^T X_t dB_t\right| > \varepsilon\right) \leq \frac{C}{\varepsilon^2} + \mathbb{P}\left(\int_0^T |X_t|^2 \, dt > C\right)$$

holds for any constants $\varepsilon, C > 0$.

Proof. For each constant $C > 0$, we define the process

$$X_C(t,\omega) = \begin{cases} X(t,\omega), & \text{if } \int_0^t |X(s,\omega)|^2 \, ds \leq C, \\ 0, & \text{otherwise.} \end{cases}$$

Since $X \in L^2_{ad}$, it is clear that $X_C \in L^2_{ad}$. We have that

$$\left\{\omega : \left|\int_0^T X(t) dB_t\right| > \varepsilon\right\} \subset \left\{\omega : \left|\int_0^T X_C(t) dB_t\right| > \varepsilon\right\} \cup \left\{\omega : \int_0^T X(t) dB_t \neq \int_0^T X_C(t) dB_t\right\}.$$

Moreover, since X is a step process,

$$\left\{\omega : \int_0^T X(t) dB_t \neq \int_0^T X_C(t) dB_t\right\} \subset \left\{\omega : \int_0^T |X(t,\omega)|^2 \, dt > C\right\}.$$

Therefore,

$$\begin{aligned}
&\mathbb{P}\left(\left\{\omega : \left|\int_0^T X(t) dB_t\right| > \varepsilon\right\}\right) \\
&\leq \mathbb{P}\left(\left\{\omega : \left|\int_0^T X_C(t) dB_t\right| > \varepsilon\right\}\right) + \mathbb{P}\left(\left\{\omega : \int_0^T X(t) dB_t \neq \int_0^T X_C(t) dB_t\right\}\right) \\
&\leq \mathbb{P}\left(\left\{\omega : \left|\int_0^T X_C(t) dB_t\right| > \varepsilon\right\}\right) + \mathbb{P}\left(\left\{\omega : \int_0^T |X(t,\omega)|^2 \, dt > C\right\}\right) \\
&\leq \frac{1}{\varepsilon^2}\mathbb{E}\left[\int_0^T |X_C(t)|^2 \, dt\right] + \mathbb{P}\left(\left\{\omega : \int_0^T |X(t,\omega)|^2 \, dt > C\right\}\right) \\
&\leq \frac{C}{\varepsilon^2} + \mathbb{P}\left(\left\{\omega : \int_0^T |X(t,\omega)|^2 \, dt > C\right\}\right).
\end{aligned}$$

□

Remark 2.5.5. For processes in L^2_{ad}, Lemma 2.5.4 is still true. See [22, p. 229].

Lemma 2.5.6. *Let $X \in \mathcal{L}_{ad}$. Then, there exists a sequence of step processes $\{X_n\}_n$ in L^2_{ad} such that*

$$\lim_{n\to\infty} \int_0^T |X_n(t) - X(t)|^2 \, dt = 0$$

in probability.

Proof. By Lemma 2.5.3, there exists a sequence $\{Y_n\}_n \subset L^2_{ad}$ such that

$$\lim_{n\to\infty} \int_0^T |X(t) - Y_n(t)|^2 \, dt = 0 \tag{2.5.1}$$

in probability. By Lemma 2.3.1, for each positive integer n there exists a step process $X_n \in L^2_{ad}$ such that

$$\mathbb{E}\left(\int_0^T |X_n(t) - Y_n(t)|^2 \, dt\right) < \frac{1}{n}. \tag{2.5.2}$$

Using that $(a+b)^2 \le 2(a^2+b^2)$, we have that

$$\int_0^T |X_n(t) - X(t)|^2 \, dt \le 2\int_0^T |X_n(t) - Y_n(t)|^2 \, dt + 2\int_0^T |Y_n(t) - X(t)|^2 \, dt.$$

Then, we have the following inclusion,

$$\begin{aligned}\left\{\omega : \int_0^T |X_n(t) - X(t)|^2 \, dt > \varepsilon\right\} \subset &\left\{\omega : \int_0^T |X_n(t) - Y_n(t)|^2 \, dt > \frac{\varepsilon}{4}\right\} \\ &\cup \left\{\omega : \int_0^T |Y_n(t) - X(t)|^2 \, dt > \frac{\varepsilon}{4}\right\}.\end{aligned}$$

Let $\varepsilon > 0$,

$$\begin{aligned}&\mathbb{P}\left(\int_0^T |X_n(t) - X(t)|^2 \, dt > \varepsilon\right) \\ &\le \mathbb{P}\left(\int_0^T |X_n(t) - Y_n(t)|^2 \, dt > \frac{\varepsilon}{4}\right) + \mathbb{P}\left(\int_0^T |Y_n(t) - X(t)|^2 \, dt > \frac{\varepsilon}{4}\right) \\ &\le \frac{4}{\varepsilon}\mathbb{E}\left(\int_0^T |X_n(t) - Y_n(t)|^2 \, dt\right) + \mathbb{P}\left(\int_0^T |Y_n(t) - X(t)|^2 \, dt > \frac{\varepsilon}{4}\right) \\ &\le \frac{4}{\varepsilon}\frac{1}{n} + \mathbb{P}\left(\int_0^T |Y_n(t) - X(t)|^2 \, dt > \frac{\varepsilon}{4}\right).\end{aligned}$$

Taking limits on both sides and taking into account (2.5.1), we get that

$$\lim_{n\to\infty} \mathbb{P}\left(\int_0^T |X_n(t) - X(t)|^2 \, dt > \varepsilon\right) = 0.$$

□

Now, we are ready to define the Itô integral in $\mathcal{L}_{ad}$. Let $X \in \mathcal{L}_{ad}$, by Lemma 2.5.6, there exists a sequence $\{X_n\}_n$ of step processes in L^2_{ad} such that

$$\mathbb{P} - \lim_{n\to\infty} \int_0^T |X_n(t) - X(t)|^2 \, dt = 0. \tag{2.5.3}$$

For each step process $X_n \in L^2_{ad}$, we have already defined its Itô stochastic integral as

$$I^{(n)} = \int_0^T X_n(t) dB_t \in L^2(\Omega).$$

Proposition 2.5.7. *The sequence of $L^2(\Omega)$-random variables $\{I^{(n)}\}_n$ is a Cauchy sequence in probability.*

Proof. Let $\varepsilon > 0$, we are going to apply Lemma 2.5.4 with $C = \varepsilon^3/2$,

$$\begin{aligned}\mathbb{P}\left(\left|I^{(n)} - I^{(m)}\right| > \varepsilon\right) &= \mathbb{P}\left(\left|\int_0^T (X_n(t) - X_m(t))\, dB_t\right| > \varepsilon\right) \\ &\le \frac{\varepsilon}{2} + \mathbb{P}\left(\int_0^T |X_n(t) - X_m(t)|^2 \, dt > \frac{\varepsilon^3}{2}\right).\end{aligned}$$

Using again that $(a+b)^2 \le 2(a^2+b^2)$, and by virtue of (2.5.3) we have that

$$\lim_{n,m\to\infty} \mathbb{P}\left(\int_0^T |X_n(t) - X_m(t)|^2 \, dt > \frac{\varepsilon^3}{2}\right) \le$$
$$\lim_{n\to\infty} \mathbb{P}\left(\int_0^T |X_n(t) - X(t)|^2 \, dt > \frac{\varepsilon^3}{8}\right) + \lim_{m\to\infty} \mathbb{P}\left(\int_0^T |X(t) - X_m(t)|^2 \, dt > \frac{\varepsilon^3}{8}\right) = 0.$$

Hence, there exists $N \in \mathbb{N}$ such that for all $m, n \ge N$,

$$\begin{aligned}\mathbb{P}\left(\left|I^{(n)} - I^{(m)}\right| > \varepsilon\right) &= \mathbb{P}\left(\left|\int_0^T (X_n(t) - X_m(t))\, dB_t\right| > \varepsilon\right) \\ &\le \frac{\varepsilon}{2} + \mathbb{P}\left(\int_0^T |X_n(t) - X_m(t)|^2 \, dt > \frac{\varepsilon^3}{2}\right) \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.\end{aligned}$$

□

We define an equivalence relation between random variables by $X \sim Y$ if and only if $\mathbb{P}(X = Y) = 1$.
We define $L^0(\Omega, \mathcal{F}, \mathbb{P}) = \{X : \Omega \to \mathbb{R} : \mathbb{P}(X < \infty) = 1\}/_\sim$. The convergence in probability is metrizable by the Ky Fan metric. The space $L^0(\Omega, \mathcal{F}, \mathbb{P})$ with this metric is complete. So, if a sequence in $L^0(\Omega, \mathcal{F}, \mathbb{P})$ is Cauchy in probability, then it converges in probability to some random variable in $L^0(\Omega, \mathcal{F}, \mathbb{P})$. A proof can be found in [25].

Definition 2.5.8. Given a stochastic process $X = \{X_t, 0 \le t \le T\} \in \mathcal{L}_{ad}$, we define *its stochastic Itô integral* as

$$\int_0^T X(t) dB_t := \lim_{n\to\infty} \int_0^T X_n(t) dB_t$$

in probability, where $\{X_n\}_n$ is a sequence of step processes in L^2_{ad} such that

$$\mathbb{P} - \lim_{n\to\infty} \int_0^T |X_n(t) - X(t)|^2 \, dt = 0.$$

Remark 2.5.9. We have already discussed that such a limit exists, it is unique and belongs to $L^0(\Omega, \mathcal{F}, \mathbb{P})$. We need to show that it is well-defined, meaning that it does not depend on the approximating sequence of step processes. Indeed, let $\{X_n\}_n, \{Y_n\}_n$ be two sequences of step processes in L^2_{ad} such that for any $\varepsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}\left(\int_0^T |X_n(t) - X(t)|^2 \, dt > \varepsilon\right) = 0 \tag{2.5.4}$$

$$\lim_{n\to\infty} \mathbb{P}\left(\int_0^T |Y_n(t) - X(t)|^2 \, dt > \varepsilon\right) = 0. \tag{2.5.5}$$

Then, we define

$$I_1 := \mathbb{P} - \lim_{n\to\infty} \int_0^T X_n(t) dB_t,$$

$$I_2 := \mathbb{P} - \lim_{n\to\infty} \int_0^T Y_n(t) dB_t.$$

Given $\varepsilon > 0$,

$$\mathbb{P}\left(|I_1 - I_2| > 3\varepsilon\right) \leq \mathbb{P}\left(\left|I_1 - \int_0^T X_n(t) dB_t\right| > \varepsilon\right) + \mathbb{P}\left(\left|I_2 - \int_0^T Y_n(t) dB_t\right| > \varepsilon\right) + \mathbb{P}\left(\left|\int_0^T \left(X_n(t) - Y_n(t)\right) dB_t\right| > \varepsilon\right).$$

As $n \to \infty$, it is clear that the two first terms tend to 0. We show that so does the last one,

$$\begin{aligned}
\mathbb{P}\left(\left|\int_0^T \left(X_n(t) - Y_n(t)\right) dB_t\right| > \varepsilon\right) &\leq \frac{\varepsilon}{2} + \mathbb{P}\left(\int_0^T |X_n(t) - Y_n(t)|^2 \, dt > \frac{\varepsilon^3}{2}\right) \\
&\leq \frac{\varepsilon}{2} + \mathbb{P}\left(\int_0^T |X_n(t) - X(t)|^2 \, dt > \frac{\varepsilon^3}{8}\right) \\
&\quad + \mathbb{P}\left(\int_0^T |X(t) - Y_n(t)|^2 \, dt > \frac{\varepsilon^3}{8}\right) \\
&< \varepsilon
\end{aligned}$$

for all n greater than some positive integer n_0.

Remark 2.5.10. In $\mathcal{L}_{ad}$, we do not have an isometry property as we had for L^2_{ad} because the mean might not be defined. Therefore, there is neither a martingale property, but locally it can be considered. As in L^2_{ad}, the trajectories are of an indefinite Itô integral are continuous almost surely. The proof is quite similar. For further reading, see [14].

2.6 The Itô Integral as Riemann Sums

As we mentioned at the beginning of this chapter, Itô integration can be understood in terms of Riemann sums by evaluating the integrand at the left endpoints of the intervals of the partition. We prove it in the following theorem.

Theorem 2.6.1. *Let $X \in \mathcal{L}_{ad}$ be pathwise continuous $\{\mathcal{F}_t\}$-adapted stochastic process. Then,*

$$\int_0^T X(t)dB_t = \lim_{|\Pi_n|\to 0} \sum_{i=1}^{n} X(t_{i-1})\left[B(t_i) - B(t_{i-1})\right] \tag{2.6.1}$$

in probability, where $\Pi_n = \{0 = t_0 < t_1 < \ldots < t_n = T\}$ *is a partition of* $[0, T]$.

Proof. For each positive integer n, we define the step process

$$X^{(n)}(t) = \sum_{i=1}^{n} X(t_{i-1})\mathbb{1}_{[t_{i-1},t_i)}(t)$$

for some partition $\Pi_n = \{0 = t_0 < t_1 < \ldots < t_n = T\}$ of $[0, T]$. By uniform continuity, given $\varepsilon > 0$, there exists $\delta > 0$ such that for all $s, t \in [0, T]$ with $|t - s| < \delta$, then

$$|X(t) - X(s)| < \frac{\varepsilon}{\sqrt{T}} \qquad a.s.$$

For n large enough, $|\Pi_n| < \delta$, and

$$\int_0^T \left|X^{(n)}(t) - X(t)\right|^2 dt = \int_0^T \left|\sum_{i=1}^{n} X(t_{i-1})\mathbb{1}_{[t_{i-1},t_i)}(t) - X(t)\right|^2 dt < \varepsilon^2 \qquad a.s.$$

We have proved that

$$\int_0^T \left|X^{(n)}(t) - X(t)\right|^2 dt \to 0$$

almost surely and, hence, in probability as $n \to \infty$. By definition,

$$\int_0^T X(t)dB_t := \mathbb{P} - \lim_{n\to\infty} \int_0^T X^{(n)}(t)dB_t = \mathbb{P} - \lim_{|\Pi_n|\to 0} \sum_{i=1}^{n} X(t_{i-1})\left[B(t_i) - B(t_{i-1})\right].$$

□

Chapter 3

Stochastic Differential Equations

As we have already mentioned, the main motivation for constructing the Itô integral is to solve stochastic differential equations. As a stochastic differential equation we understand an expression like

$$X_t = \xi + \int_0^t \sigma(s, X_s)dB_s + \int_0^t \mu(s, X_s)ds, \quad t \in [0, T],$$

in which we aim to find a "solution" X_t satisfying it.
More precisely, Itô realized that the solutions of stochastic differential equations have the Markov property and are diffusions. Therefore, stochastic differential equations are a way of generating diffusions. We will not deal with these properties, but a discussion and a derivation of the Kolmogorov equations can be found in [14, p. 211 − 228]. We will begin considering the Itô formula which plays the role of the change of variables rule in Itô calculus. We will study the existence and uniqueness of solutions as in the classical theory of differential equations. We will also focus on stability properties of solutions and the regularity of sample paths. Again, in this chapter we will outline that measurability is very important. We will consider always non-anticipating initial conditions. The main shortcoming of the theory exposed in this chapter is when the initial condition or the integrands are not adapted to the filtration. The chapter finishes with a financial example.

3.1 The Itô Formula

In this section, we recall the expression of the Itô formula. We will omit proofs as it is a well-known formula obtained by K. Itô in 1944 and it has too many technical details. A full proof can be found in [12, p. 149] or in [22, p. 233 − 242].

On some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, we consider a one-dimensional Brownian motion $\{B_t, t \in [0, T]\}$ along with a filtration $\{\mathcal{F}_t, t \in [0, T]\}$ satisfying that

(i) For all $t \in [0, T]$, B_t is $\mathcal{F}_t$-measurable.

(ii) For all $0 \leq s \leq t \leq T$, $B_t - B_s$ is independent of $\mathcal{F}_s$.

We will keep this setting for the rest of the chapter.

Definition 3.1.1. Let $\{f(t,\omega), t \in [0,T]\}$ be a $\{\mathcal{F}_t\}$-adapted stochastic process such that

$$\int_0^T |f(t,\omega)| < \infty$$

almost surely. Consider a stochastic process $\{g(t,\omega), t \in [0,T]\} \in \mathcal{L}_{ad}$ and X_0 a $\mathcal{F}_0$-measurable random variable. An *Itô process* $\{X_t, t \in [0,T]\}$, is a stochastic process of the form

$$X_t = X_0 + \int_0^t g(t,\omega)dB(t,\omega) + \int_0^t f(t,\omega)dt, \qquad t \in [0,T]. \tag{3.1.1}$$

Note that the conditions in Definition 3.1.1 assure that both integrals are well-defined.

Theorem 3.1.2 (Itô formula)**.** *Let $\{X_t, t \in [0,T]\}$ be an Itô process of the form*

$$X_t = X_0 + \int_0^t g(t,\omega)dB(t,\omega) + \int_0^t f(t,\omega)dt$$

with $t \in [0,T]$. Let $F : [0,T] \times \mathbb{R} \to \mathbb{R}$ be a $\mathcal{C}^{1,2}$ function. Then, $F(t,X_t)$ is an Itô process and

$$F(t,X_t) = F(0,X_0) + \int_0^t \partial_s F(s,X_s)ds + \int_0^t \partial_x F(s,X_s)dX_s + \frac{1}{2}\int_0^T \partial^2_{xx} F(s,X_s)g(s)^2 ds \tag{3.1.2}$$

holds almost surely.

We outline that in Equation (3.1.2) we have obtained an extra third term compared to the usual change of variables formula. This additional term is due to the quadratic variation of Brownian motion. In differential notation, the Itô process is denoted by

$$dX_t = f(t)dt + g(t)dB(t).$$

Then, the differential form of the Itô formula, if we denote $Y_t = F(t,X_t)$, is given by

$$dY_t = \partial_t F(t,X_t)dt + \partial_x F(t,X_t)dX_t + \frac{1}{2}\partial^2_{xx} F(t,X_t)dX_t^2.$$

The rules for the computation of $(dX_t)^2$ are as follows

$\times$	dB_t	dt
dB_t	dt	0
dt	0	0

Table 3.1: Itô rules for the differentials.

Note that with these rules we have that $dX_t^2 = g(s)^2 dt$ for the additional term.

3.2 Definitions and Examples

In this section, we introduce the concepts of solution of a stochastic differential equation and uniqueness of the solution. We also examine some examples so as to motivate the necessity of a theorem on existence and uniqueness of solutions.

Let $\sigma, \mu : [0,T] \times \mathbb{R} \to \mathbb{R}$ be two measurable functions. As a stochastic differential equation, we understand any expression of the form

$$dX_t = \sigma(t, X_t)dB_t + \mu(t, X_t)dt, \qquad X_0 = \xi, \tag{3.2.1}$$

where ξ is a $\mathcal{F}_0$-measurable random variable. More precisely, (3.2.1) means

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

Note that conditions on σ, μ, and X_t are required so as to assure that the integrals make sense. For this purpose, we begin defining what is a strong solution of a stochastic differential equation.

Definition 3.2.1. A measurable and $\{\mathcal{F}_t\}$-adapted stochastic process $\{X_t, t \in [0,T]\}$ is *a strong solution of* (3.2.1) if the following conditions are satisfied

(i) The process $\{\sigma(t, X_t), 0 \leq t \leq T\}$ belongs to $L^2_{ad}([0,T] \times \Omega)$.

(ii) Almost all sample paths of the process $\{\mu(t, X_t), 0 \leq t \leq T\}$ belong to $L^1[0,T]$.

(iii) For each $t \in [0,T]$, Equation (3.2.1) holds almost surely.

Remark 3.2.2. The previous definition is referred to *strong* solutions. Weak solutions consist in not only finding X_t, but also a Brownian motion and the filtration associated to it. A study of weak solutions can be found in [12] and [13].
From now on, whenever we refer to solutions of stochastic differential equations, we mean in the strong sense.

Definition 3.2.3. We say that the stochastic differential equation (3.2.1) has a *pathwise unique solution* if given two strong solutions X_1 and X_2 fulfilling the three conditions of Definition 3.2.1, they are indistinguishable, that is,

$$\mathbb{P}\left(X_1(t) = X_2(t), \text{ for all } t \in [0,T]\right) = 1.$$

Remark 3.2.4. As in the previous remark, the definition concerns uniqueness in the strong sense. In the weak sense, two weak solutions are unique if they are equal in distribution.

Example 3.2.5. As a first example, we propose a well-known stochastic differential equation, *Langevin equation*, given by

$$dX_t = \alpha dB_t - \beta X_t dt.$$

with initial condition $X_0 = x_0$, a constant.
This equation is used in some physical models where X_t is the velocity of a particle whose trajectory is a random variable which takes into account inertial effects superimposed on a Brownian motion. We consider the process $Y_t = e^{\beta t}X_t$ and we apply to it the Itô formula,

$$dY_t = \beta e^{\beta t}X_t dt + e^{\beta t}dX_t = \beta e^{\beta t}X_t dt + e^{\beta t}\left[\alpha dB_t - \beta X_t dt\right] = \alpha e^{\beta t}dB_t.$$

In integral form, it is expressed as

$$Y_t = Y_0 + \alpha \int_0^t e^{\beta s} dB_s.$$

Removing the change of variables,

$$e^{\beta t}X_t = x_0 + \alpha \int_0^t e^{\beta s} dB_s.$$

We get that

$$X_t = x_0 e^{-\beta t} + \alpha \int_0^t e^{-\beta(t-s)} dB_s.$$

It is clear that X_t is measurable and $\{\mathcal{F}_t\}$-adapted and the Itô integral makes sense. Therefore, X_t fulfills the conditions for being a solution.
The process X_t is called an *Ornstein-Uhlenbeck process.*

In the previous example, we have been able to find a solution of the stochastic differential equation. We are not sure whether the solution is unique in the strong sense. To this purpose, we need a result on existence and uniqueness of solutions. Next, we show an example of a stochastic differential equation with infinitely many strong solutions.

Example 3.2.6. Consider the following stochastic differential equation,

$$\begin{cases} dX_t = 3X_t^{2/3}dB_t + 3X_t^{1/3}dt \\ X_0 = 0. \end{cases}$$

In integral form, it is expressed as

$$X_t = 3\int_0^t X_s^{2/3} dB_s + 3\int_0^t X_s^{1/3} ds.$$

Let $a \geq 0$. We define the function $f_a(x) = (x-a)^3 \mathbb{1}_{\{x \geq a\}}$. The function is twice continuously differentiable, therefore, we apply the Itô formula,

$$d(f_a(B_t)) = 3f_a(B_t)^{2/3}dB_t + 3f_a(B_t)^{1/3}dt.$$

It is clear that $f_a(B_t)$ satisfies the equation. Hence, the family of processes

$$X_t = (B_t - a)^3 \mathbb{1}_{\{B_t \geq a\}}$$

are solutions for all $a \geq 0$. Note that X_t is adapted since it is a function of Brownian motion and all integrals make sense. We conclude that the stochastic differential equation has infinitely many solutions.

3.3 Existence and Uniqueness of Solutions

As in the theory of differential equations, we expect that for "good" enough functions $\mu(t,x), \sigma(t,x)$ there will exist a unique strong solution of a stochastic differential equation. We begin with two lemmas which are required for proving the existence and uniqueness theorem.

Lemma 3.3.1 (Gronwall lemma)**.** *Let $\Phi \in L^1[0,T]$ be such that*

$$\Phi(t) \le f(t) + \beta \int_0^t \Phi(s)ds, \tag{3.3.1}$$

where $f \in L^1[0,T]$ and $\beta > 0$ is a constant. Then,

$$\Phi(t) \le f(t) + \beta \int_0^t f(s)e^{\beta(t-s)}ds. \tag{3.3.2}$$

Proof. We define $g(t) = \beta \int_0^t \Phi(s)ds$. Since $\Phi \in L^1([0,T])$, by the Lebesgue differentiation theorem, $g'(t) = \beta\Phi(t)$ for a.e. $t \in [0,T]$. By assumption,

$$g'(t) = \beta\Phi(t) \le \beta f(t) + \beta g(t) \quad \text{for a.e. } t \in [0,T].$$

Multiplying by an exponential factor, we get

$$g'(t)e^{-\beta t} - \beta g(t)e^{-\beta t} \le \beta f(t)e^{-\beta t} \quad \text{for a.e. } t \in [0,T].$$

On the left-hand side, we have the derivative of a product. Hence,

$$\frac{d}{dt}\left(g(t)e^{-\beta t}\right) \le \beta f(t)e^{-\beta t} \quad \text{for a.e. } t \in [0,T].$$

Integrating on both sides,

$$g(t) \le \beta \int_0^t f(s)e^{\beta(t-s)}ds.$$

By (3.3.1), we conclude that

$$\Phi(t) \le f(t) + g(t) \le f(t) + \beta \int_0^t f(s)e^{\beta(t-s)}ds.$$

□

Lemma 3.3.2. *Let $\{f_n\}_{n=1}^{\infty}$ be a sequence of functions in $L^1([0,T])$ such that*

$$f_{n+1}(t) \le \Phi(t) + \beta \int_0^t f_n(s)ds, \quad \textit{for all } t \in [0,T], \tag{3.3.3}$$

where $\Phi \in L^1([0,T])$ is non-negative and $\beta \ge 0$ is a constant. Then,

$$f_{n+1}(t) \le \Phi(t) + \beta \int_0^t \Phi(s)e^{\beta(t-s)}ds + \beta^n \int_0^t \frac{(t-s)^{n-1}}{(n-1)!} f_1(s)ds \tag{3.3.4}$$

holds for all $n \ge 2$.

Proof. By assumption,

$$f_2(t) \le \Phi(t) + \beta \int_0^t f_1(s)ds. \tag{3.3.5}$$

For $n = 2$, we first apply (3.3.3) and secondly (3.3.5),

$$\begin{aligned}
f_3(t) &\le \Phi(t) + \beta \int_0^t f_2(s)ds \\
&\le \Phi(t) + \beta \int_0^t \left(\Phi(s) + \beta \int_0^s f_1(u)du \right) ds \\
&= \Phi(t) + \beta \int_0^t \Phi(s)ds + \beta^2 \int_0^t \left(\int_0^s f_1(u)du \right) ds \\
&= \Phi(t) + \beta \int_0^t \Phi(s)ds + \beta^2 \int_0^t \int_u^t ds f_1(u)du \\
&\le \Phi(t) + \beta \int_0^t \Phi(s)e^{\beta(t-s)}ds + \beta^2 \int_0^t (t-u)f_1(u)du.
\end{aligned}$$

Note that we have applied Fubini theorem since $f_1 \in L^1([0,T])$. Now, assume that (3.3.4) holds for f_n. Then, by induction,

$$\begin{aligned}
f_{n+1}(t) &\le \Phi(t) + \beta \int_0^t f_n(s)ds \\
&\le \Phi(t) + \beta \int_0^t \left(\Phi(s) + \beta \int_0^s \Phi(u)e^{\beta(s-u)}du + \beta^{n-1} \int_0^s \frac{(s-u)^{n-2}}{(n-2)!} f_1(u)du \right) ds \\
&= \Phi(t) + \beta \int_0^t \Phi(s)ds + \beta^2 \int_0^t \int_0^s \Phi(u)e^{\beta(s-u)}du\, ds \\
&\quad + \beta^n \int_0^t \int_0^s \frac{(s-u)^{n-2}}{(n-2)!} f_1(u)du\, ds \\
&= \Phi(t) + \beta \int_0^t \Phi(s)ds + \beta^2 \int_0^t \int_u^t e^{\beta(s-u)}ds\Phi(u)du \\
&\quad + \beta^n \int_0^t \int_u^t \frac{(s-u)^{n-2}}{(n-2)!} ds f_1(u)du \\
&= \Phi(t) + \beta \int_0^t \Phi(s)ds + \beta \int_0^t \Phi(u)e^{\beta(t-u)}du - \beta \int_0^t \Phi(u)du \\
&\quad + \beta^n \int_0^t \frac{(t-u)^{n-1}}{(n-1)!} f_1(u)du \\
&= \Phi(t) + \beta \int_0^t \Phi(u)e^{\beta(t-u)}du + \beta^n \int_0^t \frac{(t-u)^{n-1}}{(n-1)!} f_1(u)du.
\end{aligned}$$

□

Definition 3.3.3. A measurable function $f : [0,T] \times \mathbb{R} \to \mathbb{R}$ is *of linear growth in x* if there exists a constant $C > 0$ such that

$$|f(t,x)| \le C(1 + |x|)$$

for all $t \in [0,T]$ and all $x \in \mathbb{R}$.

Theorem 3.3.4 (Existence and uniqueness of solutions)**.** *Let* $\sigma, \mu : [0,T]\times\mathbb{R} \to \mathbb{R}$ *be two measurable Lipschitz functions of linear growth in* x*. Let* $\xi \in L^2(\Omega)$ *be a* $\mathcal{F}_0$*-measurable random variable. Then, the stochastic differential equation*

$$dX_t = \sigma(t, X_t)dB_t + \mu(t, X_t)dt, \qquad X_0 = \xi, \tag{3.3.6}$$

has a unique pathwise continuous solution.

Proof. We begin proving uniqueness. Assume that X_t and Y_t, $t \in [0,T]$, are two strong solutions of (3.3.6). We denote $Z_t = X_t - Y_t$, $t \in [0,T]$, a pathwise continuous stochastic process and our goal is to prove that

$$\mathbb{P}\left(Z_t = 0, \text{ for all } t \in [0,T]\right) = 1.$$

By definition,

$$Z_t = \int_0^t \left(\sigma(s, X_s) - \sigma(s, Y_s)\right) dB_s + \int_0^t \left(\mu(s, X_s) - \mu(s, Y_s)\right) ds.$$

We take on both sides squares and use that $(a+b)^2 \le 2(a^2+b^2)$,

$$Z_t^2 \le 2\left(\int_0^t \left(\sigma(s, X_s) - \sigma(s, Y_s)\right) dB_s\right)^2 + 2\left(\int_0^t \left(\mu(s, X_s) - \mu(s, Y_s)\right) ds\right)^2.$$

We take expectations on both sides,

$$\begin{aligned}\mathbb{E}(Z_t^2) \le\ & 2\mathbb{E}\left[\left(\int_0^t \left(\sigma(s, X_s) - \sigma(s, Y_s)\right) dB_s\right)^2\right] \\ & + 2\mathbb{E}\left[\left(\int_0^t \left|\mu(s, X_s) - \mu(s, Y_s)\right) ds\right)^2\right].\end{aligned} \tag{3.3.7}$$

By definition of strong solution, $\sigma(s, X_s) \in L^2_{ad}$ and we apply the Itô isometry and the Lipschitz condition,

$$\begin{aligned}\mathbb{E}\left[\left(\int_0^t \left(\sigma(s, X_s) - \sigma(s, Y_s)\right) dB_s\right)^2\right] &= \int_0^t \mathbb{E}\left(\left|\sigma(s, X_s) - \sigma(s, Y_s)\right|^2\right) ds \\ &\le C^2 \int_0^t \mathbb{E}\left(\left|X_s - Y_s\right|^2\right) ds \\ &= C^2 \int_0^T \mathbb{E}\left(Z_s^2\right) ds.\end{aligned} \tag{3.3.8}$$

For the Lebesgue integral in (3.3.7), we use the Hölder inequality, the Lipschitz condition and Fubini theorem,

$$\mathbb{E}\left[\left(\int_0^t \left|\mu(s, X_s) - \mu(s, Y_s)\right) ds\right)^2\right] \le t\mathbb{E}\left[\int_0^t \left|\mu(s, X_s) - \mu(s, Y_s)\right|^2 ds\right]$$

$$\leq TC^2 \int_0^t \mathbb{E}(Z_s^2)ds. \tag{3.3.9}$$

Substituting (3.3.8) and (3.3.9) into (3.3.7), we get

$$\mathbb{E}(Z_t^2) \leq 2C^2(1+T) \int_0^t \mathbb{E}(Z_s^2)ds. \tag{3.3.10}$$

Since $Z_s \in L^2_{ad}$, we apply Lemma 3.3.1 (identifying $f = 0$) to get

$$\mathbb{E}[Z_t^2] \leq 0.$$

Obviously, $\mathbb{E}[Z_t^2] = 0$ for all $t \in [0,T]$, which implies that for each $t \in [0,T]$ $Z_t(\omega) = 0$ for almost all $\omega \in \Omega$. We denote $\mathbb{Q}_T := \mathbb{Q} \cap [0,T] = \{q_n\}_n$. For each $q_n \in \mathbb{Q}_T$, there exists $\Omega_n \subset \Omega$ such that $\mathbb{P}(\Omega_n) = 1$ and for all $\omega \in \Omega_n$ $Z_{q_n}(\omega) = 0$.
Consider $\Omega' = \cap_{n=1}^{\infty}\Omega_n$, which has probability 1 and for all $\omega \in \Omega'$ and all $n \in \mathbb{N}$ $Z_{q_n}(\omega) = 0$. Since $t \mapsto Z(t,\omega)$ is continuous almost surely, there exists $\Omega'' \subset \Omega$ such that $\mathbb{P}(\Omega'') = 1$ and for all $\omega \in \Omega''$ $Z(\cdot,\omega)$ is a continuous function on $[0,T]$.
Consider $\Omega_0 = \Omega' \cap \Omega''$ with $\mathbb{P}(\Omega_0) = 1$ such that for all $\omega \in \Omega_0$ $Z(\cdot,\omega)$ is a continuous function which vanishes on $\mathbb{Q}_T$. Since $\mathbb{Q}_T$ is dense in $[0,T]$, $Z(\cdot,\omega)$ vanishes on $[0,T]$ for almost all $\omega \in \Omega_0$. We conclude that

$$1 \geq \mathbb{P}\left(\{\omega \in \Omega : Z_t(\omega) = 0 \text{ for all } t \in [0,T]\}\right) \geq \mathbb{P}(\Omega_0) = 1.$$

This proves the uniqueness. Now, we move to prove the existence of a solution of the stochastic differential equation (3.3.6).
We consider a sequence of stochastic processes $\{X_t^{(n)}\}_{n=1}^{\infty}$ defined by

$$X_t^{(n+1)} = \xi + \int_0^t \sigma\left(s, X_s^{(n)}\right) dB_s + \int_0^t \mu\left(s, X_s^{(n)}\right) ds. \tag{3.3.11}$$

For $n = 1$, $X_t^{(1)} = \xi$.

(1) *For all integers $n \geq 1$, $\{X_t^{(n)}, t \in [0,T]\}$ belongs to L^2_{ad} and has continuous sample paths almost surely.*
We will proceed by induction. For $n = 1$, we know by assumption that ξ is $\mathcal{F}_0$-measurable, hence, $\{\mathcal{F}_t\}$-adapted and

$$\int_0^T \mathbb{E}(\xi^2)dt = T\mathbb{E}(\xi^2) < \infty.$$

Assume that $X_t^{(n)} \in L^2_{ad}$ and has continuous sample paths. Note that

$$\begin{aligned}\mathbb{E}\left(\int_0^T \sigma\left(t, X_t^{(n)}\right)^2 dt\right) &\leq \mathbb{E}\left(\int_0^T 2C\left(1 + \left(X_t^{(n)}\right)^2\right) dt\right) \\ &= 2CT + 2C\int_0^T \mathbb{E}\left[\left(X_t^{(n)}\right)^2\right] dt < \infty. \end{aligned} \tag{3.3.12}$$

Besides, it is clear that $\sigma(t, X_t^{(n)})$ is $\mathcal{F}_t$-measurable, so $\sigma(t, X_t^{(n)}) \in L^2_{ad}$. Therefore, the Itô integral in (3.3.11) makes sense, is $\mathcal{F}_t$-measurable and has continuous sample paths.
For the Lebesgue integral in (3.3.11), note that

$$\int_0^t \left|\mu\left(s, X_s^{(n)}\right)\right| ds \leq \sqrt{2CT}\left(\int_0^t \left(1 + \left(X_s^{(n)}\right)^2\right) ds\right)^{1/2} < \infty \text{ (a.s.)}, \qquad (3.3.13)$$

since $X_s^{(n)}$ has continuous sample paths almost surely. Hence, the Lebesgue integral in (3.3.11) is continuous almost surely and is $\{\mathcal{F}_t\}$-adapted by assumption on $X_s^{(n)}$. We conclude that $X_t^{(n+1)}$ is a $\{\mathcal{F}_t\}$-adapted process with continuous sample paths almost surely. Using that $(a+b+c)^2 \leq 3(a^2+b^2+c^2)$, we get that

$$\left|X_t^{(n+1)}\right|^2 \leq 3\left(\xi^2 + \left(\int_0^t \sigma\left(s, X_s^{(n)}\right) dB_s\right)^2 + \left(\int_0^t \mu\left(s, X_s^{(n)}\right) ds\right)^2\right).$$

Taking expectations on both sides and integrating, one obtains

$$\begin{aligned}
\int_0^T \mathbb{E}\left(\left|X_t^{(n+1)}\right|^2\right) dt &= 3T\mathbb{E}[\xi^2] + 3\int_0^T \mathbb{E}\left[\left(\int_0^t \sigma\left(s, X_s^{(n)}\right) dB_s\right)^2\right] dt \\
&\quad + 3\int_0^T \mathbb{E}\left[\left(\int_0^t \mu(s, X_s) ds\right)^2\right] dt \\
&\leq 3T\mathbb{E}[\xi^2] + 3\int_0^T \int_0^t \mathbb{E}\left(\sigma\left(s, X_s^{(n)}\right)^2\right) ds\ dt \\
&\quad + 3\int_0^T \mathbb{E}\left[\left(\int_0^t \mu(s, X_s) ds\right)^2\right] dt < \infty
\end{aligned}$$

by the linear growth condition and the fact that $X_t^{(n)} \in L^2_{ad}$. Thus, $X_t^{(n+1)} \in L^2_{ad}$.

(2) *The sequence $\{\{X_t^{(n)}, t \in [0,T]\}\}_{n=1}^{\infty}$ converges uniformly on t almost surely.*
We define

$$Y_t^{(n+1)} = \int_0^t \sigma\left(s, X_s^{(n)}\right) dB_s \quad \text{and} \quad Z_t^{(n+1)} = \int_0^t \mu\left(s, X_s^{(n)}\right) ds$$

so that $X_t^{(n+1)} = \xi + Y_t^{(n+1)} + Z_t^{(n+1)}$. We are going to calculate some inequalities. By the Itô isometry and the Lipschitz condition,

$$\begin{aligned}
\mathbb{E}\left(\left|Y_t^{(n+1)} - Y_t^{(n)}\right|^2\right) &= \mathbb{E}\left[\left(\int_0^t \left(\sigma\left(s, X_s^{(n)}\right) - \sigma\left(s, X_s^{(n-1)}\right)\right) dB_s\right)^2\right] \\
&= \int_0^t \mathbb{E}\left|\sigma\left(s, X_s^{(n)}\right) - \sigma\left(s, X_s^{(n-1)}\right)\right|^2 ds \\
&\leq C^2 \int_0^t \mathbb{E}\left|X_s^{(n)} - X_s^{(n-1)}\right|^2 ds. \qquad (3.3.14)
\end{aligned}$$

Using the Hölder inequality and the Lipschitz property, it follows that

$$\begin{aligned}\left|Z_t^{(n+1)} - Z_t^{(n)}\right|^2 &= \left(\int_0^t \left(\mu\left(s, X_s^{(n)}\right) - \mu\left(s, X_s^{(n-1)}\right)\right) ds\right)^2 \\ &\leq TC^2 \int_0^t \left|X_s^{(n)} - X_s^{(n-1)}\right|^2 ds. \end{aligned} \tag{3.3.15}$$

By virtue of (3.3.14) and (3.3.15), we have that

$$\begin{aligned}\mathbb{E}\left(\left|X_t^{(n+1)} - X_t^{(n)}\right|^2\right) &= \mathbb{E}\left(\left|Y_t^{(n+1)} + Z_t^{(n+1)} - Y_t^{(n)} - Z_t^{(n)}\right|^2\right) \\ &\leq 2\mathbb{E}\left(\left|Y_t^{(n+1)} - Y_t^{(n)}\right|^2\right) + 2\mathbb{E}\left(\left|Z_t^{(n+1)} - Z_t^{(n)}\right|^2\right) \\ &\leq 2C^2(1+T)\int_0^t \mathbb{E}\left|X_s^{(n)} - X_s^{(n-1)}\right|^2 ds. \end{aligned} \tag{3.3.16}$$

By Lemma 3.3.2,

$$\mathbb{E}\left(\left|X_t^{(n+1)} - X_t^{(n)}\right|^2\right) \leq [2C^2(1+T)]^{n-1} \int_0^t \frac{(t-s)^{n-2}}{(n-2)!} \mathbb{E}\left|X_s^{(2)} - X_s^{(1)}\right|^2 ds. \tag{3.3.17}$$

Note that, arguing as before, we have that

$$\begin{aligned}\mathbb{E}\left|X_s^{(2)} - X_s^{(1)}\right|^2 &= \mathbb{E}\left[\left(\int_0^s \sigma(u,\xi)dB_u + \int_0^s \mu(u,\xi)du\right)^2\right] \\ &\leq 2\int_0^s \mathbb{E}\left(\sigma(u,\xi)^2\right) du + 2T\int_0^s \mathbb{E}\left(\mu(u,\xi)^2\right) du \\ &\leq 2C^2\int_0^s \left(1+\mathbb{E}\xi^2\right) du + 2T\int_0^s \left(1+\mathbb{E}\xi^2\right) du \\ &\leq 2C^2(1+T)(1+\mathbb{E}\xi^2)s. \end{aligned} \tag{3.3.18}$$

Using inequality (3.3.18) in (3.3.17), we get that

$$\begin{aligned}\mathbb{E}\left(\left|X_t^{(n+1)} - X_t^{(n)}\right|^2\right) &\leq (1+\mathbb{E}\xi^2)[2C^2(1+T)]^n \int_0^t \frac{(t-s)^{n-2}}{(n-2)!} s\, ds \\ &\leq A\cdot B^n\left(-\frac{s(t-s)^{n-1}}{(n-1)!}\Bigg|_0^t + \int_0^t \frac{(t-s)^{n-1}}{(n-1)!}ds\right) \\ &= A\cdot B^n\, \frac{t^n}{n!}, \end{aligned} \tag{3.3.19}$$

where $A = (1+\mathbb{E}\xi^2)$ and $B = 2C^2(1+T)$. This inequality will be important for what comes next. As

$$\left|X_t^{(n+1)} - X_t^{(n)}\right| \leq \left|Y_t^{(n+1)} - Y_t^{(n)}\right| + \left|Z_t^{(n+1)} - Z_t^{(n)}\right|,$$

it holds that

$$\sup_{0\le t\le T}\left|X_t^{(n+1)}-X_t^{(n)}\right|\le \sup_{0\le t\le T}\left|Y_t^{(n+1)}-Y_t^{(n)}\right|+\sup_{0\le t\le T}\left|Z_t^{(n+1)}-Z_t^{(n)}\right|.$$

This implies the following set inclusion,

$$\left\{\sup_{0\le t\le T}\left|X_t^{(n+1)}-X_t^{(n)}\right|>\frac{1}{n^2}\right\}\subset\left\{\sup_{0\le t\le T}\left|Y_t^{(n+1)}-Y_t^{(n)}\right|>\frac{1}{2n^2}\right\}$$
$$\cup\left\{\sup_{0\le t\le T}\left|Z_t^{(n+1)}-Z_t^{(n)}\right|>\frac{1}{2n^2}\right\}.$$

Taking probabilities on both sides,

$$\mathbb{P}\left\{\sup_{0\le t\le T}\left|X_t^{(n+1)}-X_t^{(n)}\right|>\frac{1}{n^2}\right\}\le\mathbb{P}\left\{\sup_{0\le t\le T}\left|Y_t^{(n+1)}-Y_t^{(n)}\right|>\frac{1}{2n^2}\right\}$$
$$+\mathbb{P}\left\{\sup_{0\le t\le T}\left|Z_t^{(n+1)}-Z_t^{(n)}\right|>\frac{1}{2n^2}\right\}. \quad (3.3.20)$$

Since $\left|Y_t^{(n+1)}-Y_t^{(n)}\right|$ is a submartingale, we use the Doob inequality to obtain

$$\begin{aligned}\mathbb{P}\left\{\sup_{0\le t\le T}\left|Y_t^{(n+1)}-Y_t^{(n)}\right|>\frac{1}{2n^2}\right\}&\le 4n^4\mathbb{E}\left(\left|Y_T^{(n+1)}-Y_T^{(n)}\right|^2\right)\\&\le 4n^4\mathbb{E}\left[\left(\int_0^T\left(\sigma\left(s,X_s^{(n)}\right)-\sigma\left(s,X_s^{(n-1)}\right)\right)dB_s\right)^2\right]\\&\le 4n^4C^2\int_0^T\mathbb{E}\left|X_s^{(n)}-X_s^{(n-1)}\right|^2ds\\&\le 4n^4C^2AB^{n-1}\int_0^T\frac{s^{n-1}}{(n-1)!}ds\\&=4n^4C^2AB^{n-1}\frac{T^n}{n!},\end{aligned}\quad (3.3.21)$$

where we have used inequality (3.3.19). Taking the supremum on both sides in (3.3.15), one gets

$$\sup_{0\le t\le T}\left|Z_t^{(n+1)}-Z_t^{(n)}\right|^2\le TC^2\int_0^T\left|X_s^{(n)}-X_s^{(n-1)}\right|^2ds. \quad (3.3.22)$$

By Chebyshev inequality, (3.3.19) and (3.3.22),

$$\begin{aligned}\mathbb{P}\left\{\sup_{0\le t\le T}\left|Z_t^{(n+1)}-Z_t^{(n)}\right|>\frac{1}{2n^2}\right\}&\le 4n^4\mathbb{E}\left[\left(\sup_{0\le t\le T}\left|Z_t^{(n+1)}-Z_t^{(n)}\right|\right)^2\right]\\&\le 4n^4C^2TAB^{n-1}\frac{T^n}{n!}.\end{aligned}\quad (3.3.23)$$

Substituting (3.3.21) and (3.3.23) into (3.3.20), we obtain

$$\mathbb{P}\left\{\sup_{0\le t\le T}\left|X_t^{(n+1)}-X_t^{(n)}\right|>\frac{1}{n^2}\right\}\le 4A\frac{n^4B^nT^n}{n!}. \tag{3.3.24}$$

Summing on both sides,

$$\sum_{n=1}^{\infty}\mathbb{P}\left\{\sup_{0\le t\le T}\left|X_t^{(n+1)}-X_t^{(n)}\right|>\frac{1}{n^2}\right\}\le 4A\sum_{n=1}^{\infty}\frac{n^4B^nT^n}{n!}<\infty \tag{3.3.25}$$

by the ratio test. By the Borel-Cantelli lemma, for almost all $\omega\in\Omega$, there exists $n_0(\omega)\in\mathbb{N}$ such that for all integers $n\ge n_0$

$$\sup_{0\le t\le T}\left|X_t^{(n+1)}(\omega)-X_t^{(n)}(\omega)\right|\le\frac{1}{n^2}. \tag{3.3.26}$$

Note that

$$X_t^{(n)}=\xi+\sum_{k=1}^{n-1}\left(X_t^{(k+1)}-X_t^{(k)}\right).$$

By (3.3.26), the series on the right-hand side converges uniformly on t for almost all $\omega\in\Omega$. Therefore the limit

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

exists and is uniform on $t\in[0,T]$ almost surely.

3) $\{X_t, t\in[0,T]\}$ *is the solution of the stochastic differential equation of the statement.* Since $X_t^{(n)}$ is adapted to the filtration and has continuous sample paths almost surely, by uniform convergence, X_t is also adapted to the filtration and has continuous sample paths almost surely. We check that $X_t\in L^2_{ad}$. Indeed,

$$\begin{aligned}\|X_t\|_{L^2(\Omega)}&=\left\|\xi+\sum_{k=1}^{\infty}\left(X_t^{(k+1)}-X_t^{(k)}\right)\right\|_{L^2(\Omega)}\\&\le\|\xi\|_{L^2(\Omega)}+\sum_{n=1}^{\infty}\left\|X_t^{(k+1)}-X_t^{(k)}\right\|_{L^2(\Omega)}\\&\le\|\xi\|_{L^2(\Omega)}+\sum_{n=1}^{\infty}\sqrt{A}\ \frac{B^{n/2}T^{n/2}}{\sqrt{n!}}<\infty.\end{aligned}$$

Hence,

$$\int_0^T\mathbb{E}\left(X_t^2\right)dt<\infty. \tag{3.3.27}$$

Now, we check that X_t fulfills the conditions of Definition 3.2.1. Note that

$$\mathbb{E}\int_0^T\sigma\left(t,X_t\right)^2dt\le C^2\int_0^T\left(1+\mathbb{E}X_t^2\right)dt\le C^2\left(T+\int_0^T\mathbb{E}X_t^2dt\right)<\infty$$

because of (3.3.27). Thus, $\sigma(t, X_t) \in L^2_{ad}$. Likewise, one gets that $\mu(t, X_t) \in L^1([0,T])$ almost surely. In order to prove the last condition in Definition 3.2.1, we want to take limits on the following expression

$$X_t^{(n)} = \xi + \int_0^t \sigma(s, X_s^{(n-1)})dB_s + \int_0^t \mu(s, X_s^{(n-1)})ds. \tag{3.3.28}$$

For the left-hand side, we know that

$$\lim_{n\to\infty} X_t^{(n)} = X_t \tag{3.3.29}$$

uniformly on $t \in [0,T]$ almost surely. For the Lebesgue integral, we have that

$$\begin{aligned} \left|\int_0^t \left[\mu\left(s, X_s^{(n-1)}\right) - \mu(s, X_s)\right] ds\right| &\le C \int_0^t \left|X_s^{(n-1)} - X_s\right| ds \\ &\le C \cdot T \sup_{0\le t\le T} \left|X_t^{(n)} - X_t\right| \to 0 \end{aligned} \tag{3.3.30}$$

as $n \to \infty$ almost surely. For the Itô integral, we have that for any $\varepsilon > 0$

$$\begin{aligned} &\mathbb{P}\left(\left|\int_0^t \left(\sigma\left(s, X_s^{(n)}\right) - \sigma\left(s, X_s\right)\right) dB_s\right| > \varepsilon\right) \\ &\le \mathbb{P}\left(\int_0^t \left|\sigma\left(s, X_s^{(n)}\right) - \sigma\left(s, X_s\right)\right|^2 ds > \varepsilon^3\right) + \varepsilon \\ &\le \mathbb{P}\left(C^2 \int_0^t \left|X_s^{(n)} - X_s\right|^2 ds > \varepsilon^3\right) + \varepsilon \\ &\le \mathbb{P}\left(C^2 t \sup_{0\le s\le T} \left|X_s^{(n)} - X_s\right|^2 ds > \varepsilon^3\right) + \varepsilon \\ &\le 2\varepsilon \end{aligned} \tag{3.3.31}$$

for n large enough. By taking a subsequence if necessary, we can assume that convergence is almost surely. Letting $n \to \infty$ in (3.3.28) and by virtue of (3.3.29), (3.3.30) and (3.3.31), we conclude that for any $t \in [0,T]$

$$X_t = \xi + \int_0^t \sigma(s, X_s)dB_s + \int_0^t \mu(s, X_s)ds$$

holds almost surely. □

3.4 Properties of the Solution

In this section, we study some properties of the solutions of stochastic differential equations. More precisely, we study an estimate for the moments of the solution, the continuous dependence with initial conditions and the path regularity. We begin stating as a lemma the Burkholder inequality because it will be extremely useful throughout this section. A proof of it can be found in [22, p. 266].

Lemma 3.4.1 (Burkholder inequality). *Let $\{X_t, t \in [0,T]\}$ be a stochastic process in L^2_{ad}. For all $p \geq 2$, it holds that*

$$\mathbb{E}\left[\sup_{0\leq t\leq T}\left|\int_0^t X_s dB_s\right|^p\right] \leq C(p)\mathbb{E}\left[\left(\int_0^T |X_s|^2\, ds\right)^{p/2}\right],$$

where $C(p)$ is a constant which depends on $p \geq 2$.

Theorem 3.4.2. *Let $\sigma, \mu : [0,T]\times\mathbb{R} \to \mathbb{R}$ be two measurable Lipschitz functions of linear growth in x with constant $L > 0$. Let $p \geq 2$ and $\xi \in L^p(\Omega)$ be a $\mathcal{F}_0$-measurable random variable. Then, the unique solution of the stochastic differential equation*

$$dX_t = \sigma(t, X_t)dB_t + \mu(t, X_t)dt, \quad X_0 = \xi,$$

satisfies that for any $t \in [0,T]$ there exists a positive constant $C = C(p,t,L)$ such that

$$\mathbb{E}\left(\sup_{0\leq s\leq t}|X_s|^p\right) \leq C\left(1 + \mathbb{E}\left(|\xi|^p\right)\right).$$

Proof. We begin noting the fact that $(a+b+c)^p \leq C(p)(a^p + b^p + c^p)$. Thus, using this inequality it follows that

$$\begin{aligned}\mathbb{E}\left(\sup_{0\leq s\leq t}|X_s|^p\right) \leq C(p)\Bigg(\mathbb{E}\,|\xi|^p + \mathbb{E}\left(\sup_{0\leq s\leq t}\left|\int_0^s \sigma(u, X_u)dB_u\right|^p\right) +\\ + \mathbb{E}\left(\sup_{0\leq s\leq t}\left|\int_0^s \mu(u, X_u)du\right|^p\right)\Bigg).\end{aligned} \tag{3.4.1}$$

For the Itô integral, we are going to use the Burkholder inequality, followed by the Hölder inequality and the linear growth condition. We have that

$$\begin{aligned}\mathbb{E}\left(\sup_{0\leq s\leq t}\left|\int_0^s \sigma(u, X_u)dB_u\right|^p\right) &\leq C(p)\mathbb{E}\left[\left(\int_0^t |\sigma(s, X_s)|^2\, ds\right)^{p/2}\right]\\ &\leq C(p,t)\mathbb{E}\left(\int_0^t |\sigma(s, X_s)|^p\, ds\right)\\ &\leq C(p,t,L)\mathbb{E}\left(\int_0^t (1 + |X_s|^p)\, ds\right).\end{aligned} \tag{3.4.2}$$

For the Lebesgue integral, we apply the Hölder inequality to get

$$\begin{aligned}\mathbb{E}\left(\sup_{0\leq s\leq t}\left|\int_0^s \mu(u, X_u)du\right|^p\right) &\leq \mathbb{E}\left(\sup_{0\leq s\leq t}\left|\int_0^s |\mu(u, X_u)|\, du\right|^p\right)\\ &= \mathbb{E}\left(\left(\int_0^t |\mu(u, X_u)|\, du\right)^p\right)\\ &\leq C(p,t)\mathbb{E}\left(\int_0^t |\mu(u, X_u)|^p\, du\right)\end{aligned}$$

$$\leq C(p,t,L)\mathbb{E}\left(\int_0^t (1+|X_s|^p)\,ds\right). \tag{3.4.3}$$

Substituting (3.4.2) and (3.4.3) into (3.4.1), we obtain that

$$\begin{aligned}\mathbb{E}\left(\sup_{0\leq s\leq t}|X_s|^p\right) &\leq \tilde{C}(p,t,L)\left(\mathbb{E}\,|\xi|^p + t + \int_0^t \mathbb{E}\,|X_s|^p\,ds\right)\\ &\leq K(p,t,L)\left(\mathbb{E}\,|\xi|^p + 1 + \int_0^t \mathbb{E}\left(\sup_{0\leq u\leq s}|X_u|^p\right)ds\right).\end{aligned}$$

We denote

$$\Phi(t) = \mathbb{E}\left(\sup_{0\leq s\leq t}|X_s|^p\right).$$

Note that Φ is well-defined and belongs to $L^1([0,T])$ as $\mathbb{E}\,|\xi|^p < \infty$. Indeed, it follows from the iteration scheme in the proof of Theorem 3.3.4 and using Burkholder inequality instead of Itô isometry. By Gronwall lemma (Lemma 3.3.1), one concludes that

$$\mathbb{E}\left(\sup_{0\leq s\leq t}|X_s|^p\right) \leq C(p,t,L)\left(1+\mathbb{E}\left(|\xi|^p\right)\right). \qquad \square$$

We outline that Theorem 3.4.2 provides bounds for the moments of the solution of a stochastic differential equation in terms of the moments of the initial condition. The next theorem shows a continuous dependence between the initial conditions and solutions.

Theorem 3.4.3. *Let $\sigma,\mu : [0,T]\times\mathbb{R}\to\mathbb{R}$ be two measurable Lipschitz functions of linear growth in x with constant $L>0$. Let $p\geq 2$ and $\xi,\eta\in L^p(\Omega)$ be two $\mathcal{F}_0$-measurable random variables. Then, the unique solutions of the stochastic differential equation*

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

with respective initial conditions $X_0=\xi$ and $X_0=\eta$ satisfy that for any $t\in[0,T]$ there exists a positive constant $C=C(p,t,L)$ such that

$$\mathbb{E}\left(\sup_{0\leq s\leq t}\left|X_s^\xi - X_s^\eta\right|^p\right) \leq C\mathbb{E}\left(|\xi-\eta|^p\right).$$

Proof. As the proof is very similar to the proof of Theorem 3.4.2, we may omit some details which have already been explained. Note that

$$X_t^\xi - X_t^\eta = \xi - \eta + \int_0^t \left(\sigma\left(s,X_s^\xi\right) - \sigma\left(s,X_s^\eta\right)\right)dB_s + \int_0^t \left(\mu\left(s,X_s^\xi\right) - \mu\left(s,X_s^\eta\right)\right)ds.$$

Taking supremums and expectations, we obtain that

$$\mathbb{E}\left(\sup_{0\leq s\leq t}\left|X_s^\xi - X_s^\eta\right|^p\right)$$

$$\leq C(p)\left(\mathbb{E}\left|\xi-\eta\right|^p+\mathbb{E}\left(\sup_{0\leq s\leq t}\left|\int_0^s\left(\sigma\left(u,X_u^\xi\right)-\sigma\left(u,X_u^\eta\right)\right)dB_u\right|^p\right)\right.$$
$$\left.+\,\mathbb{E}\left(\sup_{0\leq s\leq t}\left|\int_0^s\left(\mu\left(u,X_u^\xi\right)-\mu\left(u,X_u^\eta\right)\right)du\right|^p\right)\right). \tag{3.4.4}$$

From the Burkholder inequality, the Hölder inequality and the Lipschitz condition, we find for the Itô integral that

$$\begin{aligned}
&\mathbb{E}\left(\sup_{0\leq s\leq t}\left|\int_0^s\left(\sigma\left(u,X_u^\xi\right)-\sigma\left(u,X_u^\eta\right)\right)dB_u\right|^p\right)\\
&\leq C(p)\mathbb{E}\left[\left(\int_0^t\left|\sigma\left(s,X_s^\xi\right)-\sigma\left(s,X_s^\eta\right)\right|^2ds\right)^{p/2}\right]\\
&\leq C(p,t)\mathbb{E}\left(\int_0^t\left|\sigma\left(s,X_s^\xi\right)-\sigma\left(s,X_s^\eta\right)\right|^pds\right)\\
&\leq C(p,t,L)\mathbb{E}\left(\int_0^t\left|X_s^\xi-X_s^\eta\right|^pds\right).
\end{aligned}\tag{3.4.5}$$

For the Lebesgue integral, we use the Hölder inequality and the Lipschitz condition to get

$$\begin{aligned}
\mathbb{E}\left(\sup_{0\leq s\leq t}\left|\int_0^s\left(\mu\left(u,X_u^\xi\right)-\mu\left(u,X_u^\eta\right)\right)du\right|^p\right)&\leq\mathbb{E}\left(\left|\int_0^t\left|\mu\left(s,X_s^\xi\right)-\mu\left(s,X_s^\eta\right)\right|ds\right|^p\right)\\
&\leq C(p,t)\mathbb{E}\left(\int_0^t\left|\mu\left(s,X_s^\xi\right)-\mu\left(s,X_s^\eta\right)\right|^pds\right)\\
&\leq C(p,t,L)\mathbb{E}\left(\int_0^t\left|X_s^\xi-X_s^\eta\right|^pds\right).
\end{aligned}\tag{3.4.6}$$

Substituting (3.4.5) and (3.4.6) into (3.4.4), we get

$$\mathbb{E}\left(\sup_{0\leq s\leq t}\left|X_s^\xi-X_s^\eta\right|^p\right)\leq\tilde{C}(p,t,L)\left(\mathbb{E}\left|\xi-\eta\right|^p+\int_0^t\mathbb{E}\left(\sup_{0\leq u\leq s}\left|X_u^\xi-X_u^\eta\right|^p\right)ds\right). \tag{3.4.7}$$

By Gronwall Lemma (Lemma 3.3.1), we conclude that

$$\mathbb{E}\left(\sup_{0\leq s\leq t}\left|X_s^\xi-X_s^\eta\right|^p\right)\leq C(p,t,L)\mathbb{E}\left(\left|\xi-\eta\right|^p\right). \qquad \square$$

Our next aim is to prove that if the initial condition has all moments finite, or it is a constant, then the solution has the same regularity as Brownian motion.

Proposition 3.4.4. *Let $\sigma,\mu:[0,T]\times\mathbb{R}\to\mathbb{R}$ be two measurable Lipschitz functions of linear growth in x with constant $L>0$. Let $p\geq 2$ and $\xi\in L^p(\Omega)$ be a $\mathcal{F}_0$-measurable random variable. Then, the unique solution of the stochastic differential equation*

$$dX_t=\sigma(t,X_t)dB_t+\mu(t,X_t)dt,\quad X_0=\xi,$$

satisfies that there exists a positive constant $C = C(p,T,L)$ such that, for $0 \le s \le t \le T$,

$$\mathbb{E}\left(|X_t - X_s|^p\right) \le C(p,T,L)\left(1 + \mathbb{E}\left(|\xi|^p\right)\right)|t-s|^{\frac{p}{2}}.$$

Proof. Note that

$$\mathbb{E}\left(|X_t - X_s|^p\right) \le C(p)\left(\mathbb{E}\left(\left|\int_s^t \sigma(u,X_u)dB_u\right|^p\right) + \mathbb{E}\left(\left|\int_s^t \mu(u,X_u)du\right|^p\right)\right). \quad (3.4.8)$$

By the Burkholder inequality, the Hölder inequality and the linear growth conditions, we find that

$$\begin{aligned}
\mathbb{E}\left(\left|\int_s^t \sigma(u,X_u)dB_u\right|^p\right) &\le C(p)\mathbb{E}\left(\left(\int_s^t |\sigma(u,X_u)|^2 du\right)^{\frac{p}{2}}\right) \\
&\le C(p)\,|t-s|^{\frac{p}{2}-1}\,\mathbb{E}\left(\int_s^t |\sigma(u,X_u)|^p du\right) \\
&\le C(p,L)\,|t-s|^{\frac{p}{2}-1}\left(\int_s^t \left(1 + \mathbb{E}\,|X_u|^p\right) du\right) \\
&\le C(p,L)\,|t-s|^{\frac{p}{2}-1}\left(\int_s^t \left(1 + \mathbb{E}\left(\sup_{0\le u\le t}|X_u|^p\right)\right) du\right) \\
&\le C(p,L)\,|t-s|^{\frac{p}{2}-1}\,|t-s|\left(1 + \mathbb{E}\left(|\xi|^p\right)\right) \\
&\le C(p,L)\,|t-s|^{\frac{p}{2}}\left(1 + \mathbb{E}\left(|\xi|^p\right)\right). \quad (3.4.9)
\end{aligned}$$

Note that the moments of the solution have been estimated using Theorem 3.4.2. For the Lebesgue integral, we use the Hölder inequality and the linear growth condition so as to get

$$\begin{aligned}
\mathbb{E}\left(\left|\int_s^t \mu(u,X_u)du\right|^p\right) &\le |t-s|^{p-1}\int_s^t \mathbb{E}\,|\mu(u,X_u)|^p du \\
&\le C(p,L)\,|t-s|^{p-1}\int_s^t \left(1 + \mathbb{E}\,|X_u|^p\right) du \\
&\le C(p,L)\,|t-s|^p\left(1 + \mathbb{E}\,|\xi|^p\right). \quad (3.4.10)
\end{aligned}$$

Substituting (3.4.9) and (3.4.10) into (3.4.8), we obtain that

$$\begin{aligned}
\mathbb{E}\left(|X_t - X_s|^p\right) &\le \tilde{C}(p,L)(1 + \mathbb{E}\,|\xi|^p)\left(|t-s|^{\frac{p}{2}} + |t-s|^p\right) \\
&\le C(p,L,T)(1 + \mathbb{E}\,|\xi|^p)\,|t-s|^{\frac{p}{2}}.
\end{aligned}$$

□

Corollary 3.4.5. *Let $\sigma, \mu : [0,T] \times \mathbb{R} \to \mathbb{R}$ be two measurable Lipschitz functions of linear growth in x with constant $L > 0$. Assume that ξ is a $\mathcal{F}_0$-measurable random variable with all moments finite. Then, the unique solution of the stochastic differential equation*

$$dX_t = \sigma(t,X_t)dB_t + \mu(t,X_t)dt, \quad X_0 = \xi,$$

has Hölder-continuous sample paths of degree $\gamma \in (0, 1/2)$.

Proof. By virtue of Proposition 3.4.4, it holds that

$$\mathbb{E}\left(|X_t - X_s|^p\right) \leq C(p, T, L)\left(1 + \mathbb{E}\left(|\xi|^p\right)\right)|t - s|^{\frac{p}{2}}.$$

By Kolmogorov continuity theorem (Theorem 1.4.1), X_t has Hölder continuous sample paths of degree

$$0 < \gamma < \frac{\frac{p}{2} - 1}{p} = \frac{1}{2} - \frac{1}{p}.$$

Since this holds for all $p \geq 2$, letting $p \to \infty$, we obtain that $\gamma \in (0, 1/2)$. □

3.5 A Classical Financial Problem

In this section, we study a simple application of the theory of stochastic differential equations in Finance. Consider a free of risk asset (the bond) whose evolution is given by

$$\begin{cases} dS_0(t) = rS_0(t)dt, \\ S_0(0) = M_0, \end{cases} \tag{3.5.1}$$

where $r > 0$ is the interest rate, a constant, and M_0 is the initial amount of money we invest in the bond. We also consider a risky asset in the stock market modeled by the Black-Scholes model,

$$\begin{cases} dS_1(t) = \mu S_1(t)dt + \sigma S_1(t)dB(t), \\ S_1(0) = M_1, \end{cases} \tag{3.5.2}$$

where the positive constants μ, σ, M_1 are the appreciation rate of the stock (drift term), the volatility of the stock and the initial amount of money that we invest in the risky asset, respectively.

We assume that an investor has an initial amount of wealth $M > 0$ and wants to invest some of it in the free of risk asset and the rest of it in the risky asset in order to maximize his average payoff. So, $M = M_0 + M_1$. Moreover, we assume that $r < \mu$ because the risky asset must have a higher return in mean, otherwise it would be worthless to invest money in it.

In the next proposition, we solve the differential equation and the stochastic differential equation. According to the theory of this chapter, we know that this Black-Scholes stochastic differential equation will have a unique solution whose regularity will be the same as Brownian motion.

Proposition 3.5.1. *The solutions of* (3.5.1) *and* (3.5.2) *are*

$$\begin{aligned} S_0(t) &= M_0 e^{rt}, \\ S_1(t) &= M_1 \exp\left\{\left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma B_t\right\}. \end{aligned}$$

with $t \in [0, T]$.

Proof. To obtain $S_0(t)$ is immediate. To solve the stochastic differential equation, we use Itô formula. Let $Z(t) = \log S_1(t)$. Then,

$$\begin{aligned} dZ_t &= \frac{1}{S_1(t)} dS_1(t) + \frac{1}{2} \frac{-1}{S_1(t)^2} (dS_1(t))^2 \\ &= \mu dt + \sigma dB_t - \frac{1}{2}\sigma^2 dt \end{aligned}$$

Integrating in $[0, T]$, we get that

$$Z_t = Z_0 + \left(\mu - \frac{1}{2}\sigma^2\right) t + \sigma B_t.$$

Removing the change of variables, one gets that

$$S_1(t) = M_1 \exp\left\{ \left(\mu - \frac{1}{2}\sigma^2\right) t + \sigma B_t \right\}, \quad t \in [0, T].$$

We have already discussed that both solutions are unique. □

The total wealth at any time $t \in [0, T]$ is given by $S(t) = S_0(t) + S_1(t)$.

Theorem 3.5.2. *The expected value for the total wealth at time T is*

$$\mathbb{E}[S(T)] = M_0 e^{rt} + M_1 e^{\mu t}.$$

Proof. As $S_0(t)$ is a deterministic function, we only need to calculate the expected value of $S_1(T)$. Then,

$$\begin{aligned} \mathbb{E}[S_1(T)] &= M_1 e^{\left(\mu - \frac{1}{2}\sigma^2\right)T} \mathbb{E}\left[e^{\sigma\sqrt{T}B(1)}\right] \\ &= M_1 e^{\left(\mu - \frac{1}{2}\sigma^2\right)T} \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} e^{\sigma\sqrt{T}x} dx \\ &= M_1 e^{\mu T} \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-\frac{1}{2}\left(x^2 - \sigma\sqrt{T}\right)^2} dx \\ &= M_1 e^{\mu T}. \end{aligned}$$

Hence,

$$\mathbb{E}[S(T)] = \mathbb{E}[S_0(T)] + \mathbb{E}[S_1(T)] = M_0 e^{rT} + M_1 e^{\mu T}.$$ □

Since we have assumed that $r < \mu$, a trader should take the strategy $M_0 = 0$ and $M_1 = M$ in order to maximize the expected wealth at time T. That means investing all the initial wealth M in the risky asset.

Chapter 4

The Skorohod Integral

The Skorohod integral, also called Hitsuda-Skorohod integral, is a stochastic integral that dates back to the decade of 1970. The Skorohod integral was the first generalization of the Itô integral which allowed to integrate non-adapted stochastic processes.

There are two main ways to construct the integral. While Anatoliy Skorohod constructed the stochastic integral through the Wiener-Itô chaos expansion, Masuyuki Hitsuda used instead the white noise theory. We will follow Skorohod initial ideas published in [26].

The goal of this section is to construct the Skorohod integral, prove that the Itô integral is a particular case and compute some examples.

4.1 Introduction

In the 1976 International Symposium on Stochastic Differential Equations, Kiyosi Itô raised the question on how to define the stochastic integral

$$\int_0^t B(1)dB(s), \qquad 0 \leq t \leq 1.$$

Note that this integral cannot be computed within Itô theory because $B(1)$ is not adapted to the Brownian filtration $\{\mathcal{F}_t = \sigma(B_s, 0 \leq s \leq t), t \in [0,T]\}$. Itô proposed the idea of enlarging the filtration $\mathcal{F}_t$ by considering the σ-field generated by $\mathcal{F}_t$ and $B(1)$. Although the measurability problem disappears, $B(t)$ is no longer a Brownian motion with respect to the enlarged filtration. Then, Itô used the integration theory for semi-martingales to obtain that

$$\int_0^t B(1)dB(s) = B(1)B(t), \qquad 0 \leq t \leq 1. \tag{4.1.1}$$

In this chapter and in the next one, we will deal with this kind of integrals and related problems. As a financial motivation for the study of anticipating integrals, consider a modification of the example examined in Section 3.5. In that case, we had a trader who knew nothing about the future and we determined that the optimal strategy was to invest all his initial wealth in the risky asset. Imagine now that this trader has privileged information and knows at time $t = 0$ the value of the risky asset at maturity T, that is, $\bar{S}_1(T)$. In this situation, a trader would follow the following strategy:

(i) $M_0 = M\mathbb{1}_{\{\bar{S}_1(T)\leq\bar{S}_0(T)\}}$,

(ii) $M_1 = M\mathbb{1}_{\{\bar{S}_1(T)>\bar{S}_0(T)\}}$.

Note that $\bar{S}_0(T)$ is known as the bond and it is free of risk. So, the dishonest trader can compare the value of both assets at maturity and invest in the one with higher value at time T. The equations for this problem are

$$\begin{cases} dS_0(t) = rS_0(t)dt, \\ S_0(0) = M\mathbb{1}_{\{\bar{S}_1(T)\leq\bar{S}_0(T)\}}, \end{cases} \tag{4.1.2}$$

$$\begin{cases} dS_1(t) = \mu S_1(t)dt + \sigma S_1(t)dB(t), \\ S_1(0) = M\mathbb{1}_{\{\bar{S}_1(T)>\bar{S}_0(T)\}}. \end{cases} \tag{4.1.3}$$

Note that the stochastic differential equation (4.1.3) cannot be solved using the Itô theory as the initial condition is anticipating. With this example, we only pretend to motivate anticipating calculus from a financial point of view. In [7], the problem is solved using both Skorohod and Russo-Vallois (also called forward integral) integration and the conclusion is that the solution in Skorohod sense has no financial meaning as the expected wealth is smaller than in the case of the trader with no information on the future. By contrast, the forward integral gives a full financial sense because the expected wealth is higher than the one in the Itô theory.

4.2 Iterated Itô Integrals

In this section, we study the Itô iterated integrals and give some properties needed for the construction of the Skorohod integral.

4.2.1 Definition and Properties

We begin with some notation, which will often appear in this chapter.

Definition 4.2.1. A function $f : [0,T]^n \to \mathbb{R}$ is *symmetric* if and only if

$$f(x_1,\ldots,x_n) = f\left(x_{\sigma(1)},\ldots,x_{\sigma(n)}\right)$$

for any permutation σ of the set $\{1,\ldots,n\}$.

Remark 4.2.2. We denote by $\hat{L}^2([0,T])$ the space of symmetric functions f such that

$$\|f\|^2_{\hat{L}^2([0,T])} = \int_{[0,T]^n} |f(x_1,\ldots,x_n)|^2\, dx_1\ldots dx_n < \infty.$$

We also define the following subset of $[0,T]^n$,

$$S_n := \{(x_1,\ldots,x_n) : 0 \leq x_1 \leq x_2 \leq \ldots \leq x_n \leq T\}.$$

Definition 4.2.3. The *symmetrization* of a function $f : [0,T]^n \to \mathbb{R}$ is another real function $\hat{f} : [0,T]^n \to \mathbb{R}$ defined by

$$\hat{f}(x_1, \dots, x_n) = \frac{1}{n} \sum_{\sigma} f\left(x_{\sigma(1)}, \dots, x_{\sigma(n)}\right)$$

where we sum over all permutations σ of the set $\{1, \dots, n\}$.

Example 4.2.4. If we consider the function $f(x,y) = xy^2$, the symmetrization of f is

$$\hat{f}(x,y) = \frac{1}{2}\left(xy^2 + x^2 y\right).$$

Definition 4.2.5. Let $f : S_n \to \mathbb{R}$ be a deterministic function such that

$$\|f\|^2_{L^2(S_n)} := \int_{S_n} |f(t_1, \dots, t_n)|^2 \, dt_1 \dots dt_n < \infty.$$

Then, we define the *n-fold iterated Itô integral* as

$$J_n(f) := \int_0^T \int_0^{t_n} \dots \int_0^{t_3} \int_0^{t_2} f\left(t_1, \dots, t_n\right) dB_{t_1} \dots dB_{t_n}.$$

Remark 4.2.6. We outline that the iterated Itô integral defined above is well-defined. Note that each integral is a $\mathcal{F}_{t_i}$-measurable stochastic process and it is square integrable with respect to the measure $d\mathbb{P} \times dm$, where m stands for the Lebesgue measure. Therefore, all the integrals have Itô sense.

Proposition 4.2.7. *Let $f \in L^2(S_n)$ and $g \in L^2(S_m)$. Then,*

$$\mathbb{E}\left(J_n(f) J_m(g)\right) = \begin{cases} 0, & \text{if } n \neq m, \\ (f,g)_{L^2(S_n)}, & \text{if } n = m, \end{cases}$$

where

$$(f,g)_{L^2(S_n)} = \int_0^T \int_0^{t_n} \dots \int_0^{t_2} f(t_1, \dots, t_n) g(t_1, \dots, t_n) dt_1 \dots dt_n,$$

holds for any integers n,$m \geq 1$. In particular, for all integers $n \geq 1$,

$$\mathbb{E}\left[J_n(f)^2\right] = \|f\|^2_{L^2(S_n)}.$$

Proof. We assume that $m < n$. By the Itô isometry,

$$\begin{aligned} \mathbb{E}\left(J_n(f) J_m(g)\right) = \mathbb{E}\Bigg[& \left(\int_0^T \int_0^{s_m} \dots \int_0^{s_2} g\left(s_1, \dots, s_m\right) dB_{s_1} \dots dB_{s_m}\right) \\ & \left(\int_0^T \int_0^{s_m} \dots \int_0^{s_1} \int_0^{t_{n-m}} \dots \int_0^{t_2} f\left(t_1, \dots, s_m\right) dB_{t_1} \dots dB_{s_m}\right)\Bigg] \\ = \int_0^T \mathbb{E}\Bigg[& \left(\int_0^{s_m} \dots \int_0^{s_2} g\left(s_1, \dots, s_m\right) dB_{s_1} \dots dB_{s_{m-1}}\right) \end{aligned}$$

$$
\begin{aligned}
&\left(\int_0^{s_m} \dots \int_0^{t_2} f\left(t_1, \dots, s_m\right) dB_{t_1} \dots dB_{s_{m-1}}\right)\Big] \, ds_m \\
&= \dots = \int_0^T \int_0^{s_m} \dots \int_0^{s_2} \mathbb{E}\left[g\left(s_1, \dots, s_m\right)\right. \\
&\left.\left(\int_0^{s_1} \dots \int_0^{t_2} f\left(t_1, \dots, s_m\right) dB_{t_1} \dots dB_{t_{n-m}}\right)\right] ds_1 \dots ds_m \\
&= \int_0^T \int_0^{s_m} \dots \int_0^{s_2} g\left(s_1, \dots, s_m\right) \\
&\quad \mathbb{E}\left[\left(\int_0^{s_1} \dots \int_0^{t_2} f\left(t_1, \dots, s_m\right) dB_{t_1} \dots dB_{t_{n-m}}\right)\right] ds_1 \dots ds_m \\
&= 0.
\end{aligned}
$$

We have used that Itô integrals have zero mean. Now, if $n = m$, repeating the same arguments, we get that

$$
\begin{aligned}
\mathbb{E}\left(J_n(g) J_n(f)\right) &= \int_0^T \int_0^{s_n} \dots \int_0^{s_2} \mathbb{E}\left[g\left(s_1, \dots, s_n\right) f\left(s_1, \dots, s_n\right)\right] ds_1 \dots ds_n \\
&= \int_0^T \int_0^{s_n} \dots \int_0^{s_2} g\left(s_1, \dots, s_n\right) f\left(s_1, \dots, s_m\right) ds_1 \dots ds_n \\
&= (f, g)_{L^2(S_n)}.
\end{aligned}
$$

Finally, one gets the particular case taking $f = g$ when $n = m$. □

Definition 4.2.8. Let $f \in \hat{L}^2([0,T])$. We define

$$
I_n(f) := \int_{[0,T]^n} f\left(t_1, \dots, t_n\right) dB_{t_1} \dots dB_{t_n} := n! J_n(f).
$$

We also say that $I_n(f)$ is the *n-fold iterated Itô integral.*

Proposition 4.2.9. *Let $f \in \hat{L}^2([0,T])$. The following equality holds*

$$
\mathbb{E}\left(I_n(f)^2\right) = n! \left\|f\right\|_{L^2([0,T]^n)}
$$

for all integers $n \geq 1$.

Proof. Since f is symmetric, we note that

$$
\left\|f\right\|^2_{L^2([0,T]^n)} = n! \left\|f\right\|^2_{L^2(S_n)}.
$$

Then, by Proposition 4.2.7,

$$
\mathbb{E}\left(I_n(f)^2\right) = (n!)^2 \mathbb{E}\left(J_n(f)^2\right) = (n!)^2 \left\|f\right\|^2_{L^2(S_n)} = n! \left\|f\right\|^2_{L^2([0,T]^n)}.
$$

□

4.2.2 A relation with Hermite Polynomials

Definition 4.2.10. The *Hermite Polynomial* of degree n with parameter ρ is defined by

$$H_n(x;\rho) = \sum_{k=0}^{[n/2]} \binom{n}{2k} (2k-1)!!(-\rho)^k x^{n-2k}.$$

Example 4.2.11. We give the expression of some Hermite polynomials.

$$\begin{aligned} H_0(x;\rho) &= 1, \\ H_1(x;\rho) &= x, \\ H_2(x;\rho) &= x^2 - \rho, \\ H_3(x;\rho) &= x^3 - 3\rho x. \end{aligned}$$

Remark 4.2.12. We will denote $h_n(x) = H_n(x;1)$.

Remark 4.2.13. If $\rho = \alpha^2 > 0$, then

$$H_n(x,\alpha^2) = \alpha^n H_n\left(\frac{x}{\alpha};1\right) = \alpha^n h_n\left(\frac{x}{\alpha}\right).$$

Lemma 4.2.14. *The Hermite polynomials satisfy the following properties:*

(i) $\frac{\partial}{\partial x} H_n(x;\rho) = nH_{n-1}(x;\rho)$.

(ii) $\frac{\partial}{\partial \rho} H_n(x;\rho) = -\frac{1}{2}\frac{\partial^2}{\partial x^2} H_n(x;\rho)$.

(iii) Any monomial can be expressed as a linear combination of Hermite polynomials given by

$$x^n = \sum_{k=0}^{[n/2]} \binom{n}{2k} (2k-1)!!\rho^k H_{n-2k}(x;\rho).$$

These properties of Hermite polynomials and others can be found in [1].

Proposition 4.2.15. *Let $f \in L^2([0,T])$ be a nonzero deterministic function. Then, for any $n \geq 1$*

$$n! \int_0^T \int_0^{t_n} \cdots \int_0^{t_2} f(t_1)\dots f(t_n) dB_{t_1}\dots dB_{t_n} = H_n\left(\int_0^T f(t)dB_t; \int_0^T f(t)^2 dt\right). \quad (4.2.1)$$

Proof. For $n = 1$, we know that $H_1(x;\rho) = x$ and Equation (4.2.1) clearly holds. We assume that the property holds for n and we prove it for $(n+1)$. By induction,

$$\begin{aligned} &(n+1)! \int_0^T \int_0^{t_{n+1}} \cdots \int_0^{t_2} f(t_1)\dots f(t_{n+1}) dB_{t_1}\dots dB_{t_{n+1}} \\ &= (n+1) \int_0^T f(t_{n+1}) n! \left(\int_0^{t_{n+1}} \cdots \int_0^{t_2} f(t_1)\dots f(t_n) dB_{t_1}\dots dB_{t_n}\right) dB_{t_{n+1}} \end{aligned}$$

$$= (n+1)\int_0^T f(t_{n+1}) H_n\left(\int_0^{t_{n+1}} f(s)dB_s; \int_0^{t_{n+1}} f(s)^2 ds\right) dB_{t_{n+1}}$$
$$= (n+1)\int_0^T f(t) H_n\left(\int_0^{t} f(s)dB_s; \int_0^{t} f(s)^2 ds\right) dB_t. \tag{4.2.2}$$

On the other hand, we can apply the Itô formula to $H_{n+1}(X(t,\omega), \rho(t))$, where

$$X(t,\omega) = \int_0^t f(s)dB_s \qquad \text{and} \qquad \rho(t) = \int_0^t f(s)^2 ds.$$

We get that

$$\begin{aligned} dH_{n+1}(X_t; \rho(t)) &= \frac{\partial}{\partial x} H_{n+1}(X_t;\rho(t))dX_t + \frac{1}{2}\frac{\partial^2}{\partial x^2} H_{n+1}(X_t;\rho(t)) f(t)^2 dt \\ &\quad + \frac{\partial}{\partial \rho} H_{n+1}(X_t;\rho(t)) \frac{d\rho(t)}{dt} dt \\ &= (n+1)H_n(X_t;\rho(t)) f(t) dB_t + \frac{1}{2}\frac{\partial^2}{\partial x^2} H_{n+1}(X_t;\rho(t)) f(t)^2 dt \\ &\quad - \frac{1}{2}\frac{\partial^2}{\partial x^2} H_{n+1}(X_t;\rho(t)) f(t)^2 dt \\ &= (n+1)H_n(X_t;\rho(t)) f(t) dB_t \end{aligned}$$

Integrating on $[0,T]$, we get that

$$H_{n+1}\left(\int_0^T f(s)dB_s; \int_0^T f(s)^2 ds\right) = (n+1)\int_0^T f(t) H_n\left(\int_0^t f(s)dB_s; \int_0^t f(s)^2 ds\right) dB_t. \tag{4.2.3}$$

Comparing (4.2.2) with (4.2.3), we conclude that

$$\begin{aligned} &(n+1)! \int_0^T \int_0^{t_{n+1}} \cdots \int_0^{t_2} f(t_1)\dots f(t_{n+1}) dB_{t_1} \dots dB_{t_{n+1}} \\ &= H_{n+1}\left(\int_0^T f(s)dB_s; \int_0^T f(s)^2 ds\right). \end{aligned}$$

□

Remark 4.2.16. Recalling Remark 4.2.13, we have that

$$(n+1)! \int_0^T \int_0^{t_{n+1}} \cdots \int_0^{t_2} f(t_1)\dots f(t_{n+1}) dB_{t_1} \dots dB_{t_{n+1}} = \|f\|_{L^2}^n h_n\left(\frac{\int_0^T f(s)dB_s}{\|f\|_{L^2}}\right).$$

4.3 The Wienner-Itô Chaos Expansion

In this section, we study the Wiener-Itô chaos expansion. It is a fundamental result in the construction of the Skorohod integral and in Malliavin Calculus. The proof of the theorem strongly relies on the Itô representation theorem (see Appendix C).

Theorem 4.3.1 (Wienner-Itô chaos expansion theorem)**.** *Let Z be a $\mathcal{F}_T$-measurable random variable satisfying that*

$$\|Z\|_{L^2(\Omega)} := \mathbb{E}\left(Z^2\right) < \infty.$$

Then, there exists a unique sequence $\{f_n\}_{n=0}^{\infty}$ with $f_n \in \hat{L}^2([0,T]^n)$ such that

$$Z(\omega) = \sum_{n=0}^{\infty} I_n(f_n)$$

in $L^2(\Omega)$. Besides, it holds that

$$\|Z\|_{L^2(\Omega)} = \sum_{n=0}^{\infty} n! \, \|f_n\|^2_{L^2([0,T]^n)} \, .$$

Proof. By the Itô representation theorem (Theorem $C.3$), there exists a unique stochastic process $\phi_1 \in L^2_{ad}([0,T] \times \Omega)$ such that

$$Z(\omega) = \mathbb{E}[Z] + \int_0^T \phi_1(s_1) dB_{s_1}. \tag{4.3.1}$$

Moreover,

$$\begin{aligned}\int_0^T \mathbb{E}\left(\phi_1(s_1)^2\right) ds_1 &= \mathbb{E}\left[\left(\int_0^T \phi_1(s_1) dB_{s_1}\right)^2\right] \\ &= \mathbb{E}\left(Z^2\right) - \mathbb{E}(Z)^2 \\ &\le \mathbb{E}\left(Z^2\right) < \infty. \end{aligned} \tag{4.3.2}$$

We knew from the representation theorem that $\phi_1 \in L^2_{ad}([0,T] \times \Omega)$, but (4.3.2) gives a bound in terms of $\mathbb{E}(Z^2)$. Since

$$\mathbb{E}\left(\phi_1(s_1)^2\right) < \infty$$

for almost all $s_1 \in [0,T]$, by the Itô representation theorem, for almost all $s_1 \in [0,T]$ there exists a stochastic process $\{\phi_2(s_2, s_1, \omega), 0 \le s_2 \le s_1\} \in L^2_{ad}([0,s_1] \times \Omega)$, which depends on s_1, such that

$$\phi_1(s_1, \omega) = \mathbb{E}\left(\phi_1(s_1, \omega)\right) + \int_0^{s_1} \phi_2(s_2, s_1, \omega) dB_{s_2}. \tag{4.3.3}$$

Moreover, for almost all $s_1 \in [0,T]$,

$$\int_0^{s_1} \mathbb{E}\left(\phi_2(s_2, s_1)^2\right) ds_2 = \mathbb{E}\left[\left(\int_0^{s_1} \phi_2(s_2, s_1, \omega) dB_{s_2}\right)^2\right] \le \mathbb{E}\left(\phi_1(s_1)^2\right) < \infty.$$

Integrating on $[0,T]$, by (4.3.2),

$$\int_0^T \int_0^{s_1} \mathbb{E}\left(\phi_2(s_2, s_1)^2\right) ds_2 ds_1 \le \int_0^T \mathbb{E}\left(\phi_1(s_1)^2\right) ds_1 \le \mathbb{E}\left(Z^2\right) < \infty. \tag{4.3.4}$$

Substituting (4.3.3) into (4.3.1), we obtain that

$$Z(\omega) = \mathbb{E}[Z] + \int_0^T \mathbb{E}\left(\phi_1(s_1)\right) dB_{s_1} + \int_0^T \int_0^{s_1} \phi_2(s_2, s_1, \omega) dB_{s_2} dB_{s_1}. \tag{4.3.5}$$

We denote

$$J_0(g_0) := g_0 := \mathbb{E}(Z) \quad \text{and} \quad g_1(s_1) := \mathbb{E}\left(\phi_1(s_1)\right).$$

Note that g_0 is a constant and $g_1 \in L^2([0,T])$ is a deterministic function. With this notation,

$$Z(\omega) = J_0(g_0) + J_1(g_1) + \int_0^T \int_0^{s_1} \phi_2(s_2, s_1, \omega) dB_{s_2} dB_{s_1}. \tag{4.3.6}$$

By virtue of (4.3.4), we have that

$$\mathbb{E}\left(\phi_2(s_2, s_1)^2\right) < \infty$$

for almost all $0 \le s_2 \le s_1 \le T$. We apply again the Itô representation theorem, for almost all $0 \le s_2 \le s_1 \le T$, there exists a stochastic process $\{\phi_3(s_3, s_2, s_1, \omega), 0 \le s_3 \le s_2\}$ belonging to $L^2_{ad}([0, s_2] \times \Omega)$, which depends on s_1 and s_2, such that

$$\phi_2(s_2, s_1, \omega) = \mathbb{E}\left(\phi_2(s_2, s_1, \omega)\right) + \int_0^{s_2} \phi_3(s_3, s_2, s_1, \omega) dB_{s_3}. \tag{4.3.7}$$

Moreover, for almost all $(s_2, s_1) \in S_2$,

$$\int_0^{s_2} \mathbb{E}\left(\phi_3(s_3, s_2, s_1)^2\right) ds_3 = \mathbb{E}\left[\left(\int_0^{s_2} \phi_3(s_3, s_2, s_1, \omega) dB_{s_3}\right)^2\right] \le \mathbb{E}\left(\phi_2(s_2, s_1)^2\right) < \infty.$$

Integrating on S_2, by (4.3.4),

$$\int_0^T \int_0^{s_1} \int_0^{s_2} \mathbb{E}\left(\phi_3(s_3, s_2, s_1)^2\right) ds_3 ds_2 ds_1 \le \int_0^T \int_0^{s_1} \mathbb{E}\left(\phi_2(s_2, s_1)^2\right) ds_2 ds_1 \le \mathbb{E}\left(Z^2\right). \tag{4.3.8}$$

Substituting (4.3.7) into (4.3.6), we obtain that

$$\begin{aligned} Z(\omega) = {} & J_0(g_0) + J_1(g_1) + \int_0^T \int_0^{s_1} \mathbb{E}\left(\phi_2(s_2, s_1, \omega)\right) dB_{s_2} dB_{s_1} \\ & + \int_0^T \int_0^{s_1} \int_0^{s_2} \phi_3(s_3, s_2, s_1, \omega) dB_{s_3} dB_{s_2} dB_{s_1}. \end{aligned} \tag{4.3.9}$$

We define the deterministic function

$$g_2 = \mathbb{E}\left(\phi_2(s_2, s_1, \omega)\right) \in L^2(S_2).$$

With this notation, we have that

$$Z(\omega) = J_0(g_0) + J_1(g_1) + J_2(g_2) + \int_0^T \int_0^{s_1} \int_0^{s_2} \phi_3(s_3, s_2, s_1, \omega) dB_{s_3} dB_{s_2} dB_{s_1}. \tag{4.3.10}$$

By iterating this method, we get after $(n+1)$ steps that there exists a stochastic process

$$\{\phi_{n+1}(t_1,\dots,t_{n+1},\omega), 0 \le t_1 \le t_2 \le \dots \le t_n \le t_{n+1} \le T\} \in L^2_{ad}([0,t_2]\times\Omega),$$

which depends on $0 \le t_1 \le t_2 \le \dots \le t_n$, and $(n+1)$ deterministic functions $g_0, g_1, \dots, g_n$ with g_0 a constant and $g_k \in L^2(S_k)$, $1 \le k \le n$, such that

$$Z(\omega) = \sum_{k=0}^{n} J_k(g_k) + \int_0^T \int_0^{t_{n+1}} \dots \int_0^{t_2} \phi_{n+1}(t_1,\dots,t_{n+1},\omega) dB_{t_1} \dots dB_{t_{n+1}}. \tag{4.3.11}$$

Moreover,

$$\mathbb{E}\left[\left(\int_0^T \int_0^{t_{n+1}} \dots \int_0^{t_2} \phi_{n+1}(t_1,\dots,t_{n+1},\omega) dB_{t_1} \dots dB_{t_{n+1}}\right)^2\right] \le \mathbb{E}\left(Z^2\right). \tag{4.3.12}$$

In particular, the sequence of random variables $\{\psi_n\}_{n=1}^{\infty}$ defined by

$$\psi_{n+1} = \int_0^T \int_0^{t_{n+1}} \dots \int_0^{t_2} \phi_{n+1}(t_1,\dots,t_{n+1},\omega) dB_{t_1} \dots dB_{t_{n+1}}$$

is a sequence in $L^2(\Omega)$ because of (4.3.12). Furthermore, for all $k \le n$ and all deterministic functions $f_k \in L^2([0,T]^k)$, we have, by the Itô isometry, that

$$\begin{aligned}
\mathbb{E}\left(J_k(f_k)\psi_{n+1}\right) &= \int_0^T \int_0^{s_k} \dots \int_0^{s_2} \mathbb{E}\Bigg[f_k(s_1,\dots,s_k)\Bigg(\int_0^{s_1} \int_0^{t_{n+1-k}} \dots \\
&\quad \int_0^{t_2} \phi_{n+1}(t_1,\dots,t_{n+1-k},s_1,\dots,s_k)\, dB_{t_1} \dots dB_{t_{n+1-k}}\Bigg)\Bigg] ds_1 \dots ds_k \\
&= \int_0^T \int_0^{s_k} \dots \int_0^{s_2} f_k(s_1,\dots,s_k)\, \mathbb{E}\Bigg[\Bigg(\int_0^{s_1} \int_0^{t_{n+1-k}} \dots \\
&\quad \int_0^{t_2} \phi_{n+1}(t_1,\dots,t_{n+1-k},s_1,\dots,s_k)\, dB_{t_1} \dots dB_{t_{n+1-k}}\Bigg)\Bigg] ds_1 \dots ds_k = 0
\end{aligned}$$

because the Itô integral has zero mean. By Proposition 4.2.7, we also know that

$$\mathbb{E}\left((J_k(g_k)J_l(g_l)\right) = 0$$

whenever $l \ne k$. Applying Pythagoras theorem to

$$Z(\omega) = \sum_{k=0}^{n} J_k(g_k) + \psi_{n+1}, \tag{4.3.13}$$

we have that

$$\|Z\|^2_{L^2(\Omega)} = \sum_{k=0}^{n} \|J_k(g_k)\|^2_{L^2(\Omega)} + \|\psi_{n+1}\|^2_{L^2(\Omega)}. \tag{4.3.14}$$

Since $\|\psi_{n+1}\|^2_{L^2(\Omega)} \le \|Z\|^2_{L^2(\Omega)}$, we have that

$$\sum_{k=0}^{n} \|J_k(g_k)\|^2_{L^2(\Omega)} \le 2\,\|Z\|^2_{L^2(\Omega)}$$

for all $n \in \mathbb{N}$. Hence,

$$\sum_{k=0}^{\infty} \|J_k(g_k)\|^2_{L^2(\Omega)} \leq 2\,\|Z\|^2_{L^2(\Omega)} < \infty. \tag{4.3.15}$$

As $L^2(\Omega)$ is a Banach space, we conclude that the series $\sum_{k=0}^{\infty} J_k(g_k)$ converges in $L^2(\Omega)$. Then,

$$\psi := L^2(\Omega) - \lim_{n\to\infty} \psi_{n+1} = L^2(\Omega) - \lim_{n\to\infty} \left(Z - \sum_{k=0}^{n} J_k(g_k) \right) \tag{4.3.16}$$

exists. Moreover, for all positive integer k and all deterministic function $f_k \in L^2(S_k)$

$$\mathbb{E}\left[J_k(f_k)\psi\right] = 0. \tag{4.3.17}$$

Indeed, by the Hölder inequality, we have that

$$\begin{aligned}
|\mathbb{E}\left[J_k(f_k)\psi\right]| &= |\mathbb{E}\left[J_k(f_k)\left(\psi - \psi_{n+1} - \psi_{n+1}\right)\right]| \\
&\leq |\mathbb{E}\left[J_k(f_k)\left(\psi - \psi_{n+1}\right)\right]| + |\mathbb{E}\left[J_k(f_k)\psi_{n+1}\right]| \\
&\leq \mathbb{E}\left[|J_k(f_k)\left(\psi - \psi_{n+1}\right)|\right] \\
&\leq \|J_k(f_k)\|_{L^2(\Omega)} \|\psi - \psi_{n+1}\|_{L^2(\Omega)} \to 0 \quad \text{as } n \to \infty.
\end{aligned}$$

We consider $f \in L^2([0,T])$. For each positive integer k we can define $f_k : S_k \to \mathbb{R}$ by

$$f_k(t_1, \ldots, t_n) = f(t_1) \ldots f(t_k).$$

We denote $\theta = \int_0^T f(t)dB_t$. Note that θ is a Gaussian random variable. By Proposition 4.2.15,

$$\begin{aligned}
J_k(f_k) &= \int_0^T \int_0^{t_k} \cdots \int_0^{t_2} f_k(t_1, \ldots, t_k) dB_{t_1} \ldots dB_{t_k} \\
&= \int_0^T \int_0^{t_k} \cdots \int_0^{t_2} f(t_1) \ldots f(t_k) dB_{t_1} \ldots dB_{t_k} \\
&= H_k\left(\theta, \|f\|^2_{L^2([0,T])}\right) = \|f\|^k_{L^2} h_k\left(\frac{\theta}{\|f\|_{L^2}}\right).
\end{aligned} \tag{4.3.18}$$

We apply (4.3.18) to (4.3.17) with this choice of f_k and we have that

$$\mathbb{E}\left[h_k\left(\frac{\theta}{\|f\|_{L^2}}\right)\psi\right] = 0 \tag{4.3.19}$$

for all $k \geq 0$ and all $f \in L^2([0,T])$. By Lemma 4.2.14,

$$\mathbb{E}\left[\theta^k \psi\right] = 0 \tag{4.3.20}$$

for all $k \geq 0$ and all $f \in L^2([0,T)]$. As a result,

$$\mathbb{E}\left(e^{\theta}\psi\right) = \mathbb{E}\left(\sum_{k=1}^{\infty} \frac{\theta^k}{k!}\psi\right) = \sum_{k=1}^{\infty} \frac{1}{k!}\mathbb{E}\left(\theta^k\psi\right) = 0.$$

We need to check that the limit and the integral have been exchanged using the dominated convergence theorem. Indeed,

$$\mathbb{E}\left|\sum_{k=0}^{N}\frac{\theta^k}{k!}\psi\right| \leq E\left(e^{\theta}\left|\psi\right|\right) \leq \left\|e^{\theta}\right\|_{L^2(\Omega)}\left\|\psi\right\|_{L^2(\Omega)} < \infty$$

because θ is a Gaussian random variable. So far, we have proved that ψ is orthogonal to

$$\left\{e^{\int_0^T f(t)dB_t} : f \in L^2([0,T])\right\}.$$

By Lemma $C.2$, this set is dense in $L^2(\Omega)$. Therefore, $\psi = 0$ almost surely. Thus, we conclude that

$$Z(\omega) = \sum_{k=0}^{\infty} J_k(g_k)(\omega) \tag{4.3.21}$$

in $L^2(\Omega)$. By the Pythagorean theorem,

$$\|Z\|^2_{L^2(\Omega)} = \sum_{k=0}^{\infty}\|J_k(g_k)\|^2_{L^2(\Omega)}. \tag{4.3.22}$$

Recall that each g_n is a deterministic function in $L^2(S_n)$, though the statement refers to functions defined on $[0,T]^n$. Thus, we extend them as follows

$$g_n(t_1,\ldots,t_n) = 0 \quad \text{ if } (t_1,\ldots,t_n) \in [0,T]^n \setminus S_n.$$

The statement also refers to symmetric functions, so we consider for each $n \geq 0$, $f_n = \hat{g}_n$, the symmetrization of g_n. Observe that

$$\begin{aligned} I_n(f_n) = n!J_n(f_n) = n!J_n(\hat{g}_n) &= n!\int_0^T\int_0^{t_n}\cdots\int_0^{t_2}\frac{1}{n!}\sum_{\sigma} g_n(t_{\sigma(1)},\ldots,t_{\sigma(n)})dB_{t_1}\ldots dB_{t_n} \\ &= \int_0^T\int_0^{t_n}\cdots\int_0^{t_2} g_n(t_1,\ldots,g_n)dB_{t_1}\ldots dB_{t_n} \\ &= J_n(g_n). \end{aligned} \tag{4.3.23}$$

Substituting (4.3.23) into (4.3.21), we get that

$$Z(\omega) = \sum_{n=0}^{\infty} I_n(f_n)(\omega). \tag{4.3.24}$$

Substituting (4.3.23) into (4.3.22), we obtain that

$$\begin{aligned} \|Z\|^2_{L^2(\Omega)} &= \sum_{n=0}^{\infty}\|J_n(g_n)\|^2_{L^2(\Omega)} = \sum_{n=0}^{\infty}(n!)^2\|J_n(\hat{g}_n)\|^2_{L^2(\Omega)} \\ &= \sum_{n=0}^{\infty}(n!)^2\|f_n\|^2_{L^2(S_n)} \end{aligned}$$

$$= \sum_{n=0}^{\infty} n! \left\| f_n \right\|^2_{L^2([0,T]^n)} . \tag{4.3.25}$$

This proves the existence part of the statement, so we move to show the uniqueness part. Assume that there exists two sequences $\{f_n\}_{n=0}^{\infty}$,$\{h_n\}_{n=1}^{\infty}$ with $f_n, h_n \in \hat{L}^2([0,T])$ for all $n \geq 0$ such that

$$Z(\omega) = \sum_{n=0}^{\infty} I_n(f_n)(\omega) \quad \text{and} \quad Z(\omega) = \sum_{n=0}^{\infty} I_n(h_n)(\omega).$$

Subtracting both expressions we get that,

$$0 = \sum_{n=0}^{\infty} I_n(f_n) - \sum_{n=0}^{\infty} I_n(h_n) = \sum_{n=0}^{\infty} I_n(f_n - h_n). \tag{4.3.26}$$

Note that the linearity is obvious as the Itô integral is linear. By (4.3.25), we have

$$0 = \sum_{n=0}^{\infty} n! \left\| f_n - h_n \right\|^2_{L^2([0,T]^n)} .$$

This implies that for all $n \geq 0$,

$$\left\| f_n - h_n \right\|^2_{L^2([0,T]^n)} = 0.$$

Hence we conclude that, for all $n \geq 0$, $f_n = h_n$ almost everywhere with respect to Lebesgue measure. □

Example 4.3.2. Consider that $Z(\omega) = W^2(T,\omega)$. It is clear that Z is $\mathcal{F}_T$-measurable and

$$\mathbb{E}\left(Z^2\right) = \mathbb{E}\left(W^4(T,\omega)\right) = 3T^2 < \infty.$$

Hence, we know that Z has a Wiener-Itô chaos expansion. In this case, it is easy to compute it. Indeed,

$$2\int_0^T \int_0^{t_2} dW_{t_1} dW_{t_2} = 2\int_0^T W_{t_2} dW_{t_2} = W(T)^2 - T.$$

Hence,

$$W(T,\omega)^2 = T + I_2(1).$$

4.4 The Skorohod Integral

Let $\{X_t, t \in [0,T]\}$ be a (t,ω)-measurable stochastic process and $\{\mathcal{F}_t, 0 \leq t \leq T\}$ a Brownian filtration. Assume that

$$X_t \text{ is } \mathcal{F}_T - \text{measurable for all } t \in [0,T] \tag{4.4.1}$$

and

$$\mathbb{E}\left(X_t^2\right) < \infty \text{ for all } t \in [0,T]. \tag{4.4.2}$$

Note that for each $t \in [0,T]$ we can apply the Wiener-Itô chaos expansion to the random variable $X(t,\omega)$. Therefore, for each $t \in [0,T]$, there exists a sequence of functions $\{f_{n,t}\}_{n=0}^{\infty}$ with $f_{n,t} \in \hat{L}^2([0,T]^n)$ such that

$$X(t,\omega) = \sum_{n=0}^{\infty} I_n(f_{n,t})(\omega). \tag{4.4.3}$$

in $L^2(\Omega)$. We outline that the functions $f_{n,t}$, $n \in \mathbb{N}$, depend on $t \in [0,T]$. Thus we write

$$f_{n,t}(t_1,\ldots,t_n) = f_n(t_1,\ldots,t_n,t),$$

which is a function defined on $[0,T]^{n+1}$ and symmetric with respect to the n first variables. The symmetrization of $f_n(t_1,\ldots,t_n,t_{n+1})$ is given by

$$\hat{f}_n(t_1,\ldots,t_{n+1}) = \frac{1}{n+1}\left[f_n(t_1,\ldots,t_{n+1}) + f_n(t_{n+1},t_2,\ldots,t_1) + \ldots + f_n(t_1,\ldots,t_{n+1},t_n)\right] \tag{4.4.4}$$

because we only need to take into account the permutations which exchange the last variable with any other one.

Definition 4.4.1. Let $\{X_t, t \in [0,T]\}$ be a stochastic process satisfying (4.4.1), (4.4.2) and with Wiener-Itô chaos expansion

$$X(t,\omega) = \sum_{n=0}^{\infty} I_n(f_n(\cdot,t)). \tag{4.4.5}$$

Then, we define the *Skorohod integral* of X by

$$\delta(X) := \int_0^T X(t,\omega)\delta B(t) := \sum_{n=0}^{\infty} I_{n+1}(\hat{f}_n) \tag{4.4.6}$$

whenever the series converges in $L^2(\Omega)$ and $\hat{f}_n$ is the symmetrization of $f_n(t_1,\ldots,t_n,t)$ as a function of $(n+1)$ variables.

Definition 4.4.2. We say that X as in Definition 4.4.1 is *Skorohod integrable*, we denote $X \in \mathrm{Dom}(\delta)$, if the series in (4.4.6) converges in $L^2(\Omega)$.

Remark 4.4.3. A process X as in Definition 4.4.1 is Skorohod integrable if and only if

$$\mathbb{E}\left[\delta(X)^2\right] = \sum_{n=0}^{\infty} \left\| I_{n+1}(\hat{f}_n)\right\|^2_{L^2(\Omega)} = \sum_{n=0}^{\infty}(n+1)!||\hat{f}_n||^2_{L^2([0,T]^{n+1})} < \infty. \tag{4.4.7}$$

Example 4.4.4. Consider the following Skorohod integral

$$\int_0^T B(T,\omega)\delta B_t.$$

Clearly, the integrand is anticipating and, therefore, the integral would have no sense according to Itô theory. As $B(T)$ is $\mathcal{F}_T$-measurable and $\mathbb{E}\left(B(T)^2\right) = T < \infty$, we can calculate its Wiener-Itô chaos expansion. Note that

$$B(T,\omega) = \int_0^T 1 dW_t = I_1(1).$$

The process is obviously Skorohod integrable and the integral is given by

$$\int_0^T B(T)\delta B_t = I_2(\hat{1}) = 2\int_0^T \int_0^{t_2} dB_{t_1} dB_{t_2} = 2\int_0^T B_{t_2} dB_{t_2} = B(T)^2 - T.$$

Note that the result differs from (4.1.1), obtained by K. Itô in 1976.

Example 4.4.5. Consider the following integral

$$\int_0^T B(t)\left[B(T) - B(t)\right]\delta B_t.$$

Again it is clear that the integrand is $\mathcal{F}_T$-measurable, though it is not $\{\mathcal{F}_t\}$-adapted, and

$$\mathbb{E}\left[\left[(B(t)\left(B(T) - B(t)\right)\right]^2\right] = \mathbb{E}\left(B(t)^2\right)\mathbb{E}\left[\left(B(T) - B(t)\right)^2\right] = t(T-t) < \infty.$$

Thus, we can calculate the unique Wiener-Itô chaos expansion of the integrand. For a fixed $t \in [0,T]$,

$$\begin{aligned}\int_0^T \int_0^{t_2} \mathbb{1}_{\{t_1<t<t_2\}} dB_{t_1} dB_{t_2} &= \int_0^T B(t)\mathbb{1}_{\{t<t_2\}} dB_{t_2} \\ &= \int_t^T B(t) dB_{t_2} \\ &= B(t)\left[B(T) - B(t)\right].\end{aligned}$$

We denote $f_{2,t}(t_1,t_2) = \mathbb{1}_{\{t_1<t<t_2\}}$ and $\hat{f}_{2,t}(t_1,t_2) = \frac{1}{2}\mathbb{1}_{\{t_1<t<t_2\}} + \frac{1}{2}\mathbb{1}_{\{t_2<t<t_1\}}$. With this notation,

$$B(t)\left[B(T) - B(t)\right] = J_2(f_{2,t}) = 2J_2(\hat{f}_{2,t}) = I_2(\hat{f}_{2,t}).$$

We denote $g(t_1,t_2,t) = \hat{f}_{2,t}(t_1,t_2)$ a symmetric function with respect to the first two variables. Then,

$$\begin{aligned}\hat{g}(t_1,t_2,t) &= \frac{1}{6}\left[\mathbb{1}_{\{t_1<t<t_2\}} + \mathbb{1}_{\{t_2<t<t_1\}} + \mathbb{1}_{\{t<t_1<t_2\}} + \mathbb{1}_{\{t_2<t_1<t\}} + \mathbb{1}_{\{t_1<t_2<t\}} + \mathbb{1}_{\{t<t_2<t_1\}}\right] \\ &= \frac{1}{6}\end{aligned}$$

for almost every $(t_1,t_2,t_3) \in [0,T]^3$. We conclude that

$$\int_0^T B(t)\left[B(T) - B(t)\right]\delta B_t = I_3(\hat{g}) = 3!J_3(\hat{g}) = \int_0^T \int_0^{t_3} \int_0^{t_2} dB_{t_1} dB_{t_2} dB_{t_3}$$

$$= \int_0^T \frac{1}{2}\left(B_{t_3}^2 - t_3\right) dB_{t_3}$$
$$= \frac{1}{6}\left[B(T)^3 - 3TB(T)\right]$$

where we have used the Itô formula as follows

$$d\left(B(t)^3 - 3tB(t)\right) = -3B(t)dt + (3B(t)^2 - 3t)dB(t) + 3B(t)dt = 3(B(t)^2 - t)dB(t).$$

Example 4.4.6. Consider the following Skorohod integral

$$\int_0^T B_t \delta B_t.$$

This integral is Itô and Skorohod integrable. The aim of this example is to illustrate that both integrals coincide. As in the previous examples, we begin with the Wiener Itô chaos expansion of the integrand, for a fixed $t \in [0,T]$,

$$B(t) = \int_0^T \mathbb{1}_{\{s<t<T\}} dB_s = I_1(\mathbb{1}_{\{s<t<T\}}).$$

Symmetrizing,

$$\hat{f}(s,t) = \frac{1}{2}\left[\mathbb{1}_{\{s<t<T\}} + \mathbb{1}_{\{t<s<T\}}\right] = \frac{1}{2}$$

for almost all $(s,t) \in [0,T]^2$. Then,

$$\int_0^T B(t)\delta B(t) = I_2(\hat{f}) = 2\frac{1}{2}\int_0^T \int_0^t dB(s)dB(t) = \int_0^T B(t)dB(t).$$

We conclude that both integrals coincide.

4.5 The Skorohod Integral as an Extension of Itô Integration

From Example 4.4.6, it naturally arises the question whether Itô and Skorohod integrals coincide for adapted integrands. If so, that would mean that the Skorohod integral generalizes the Itô one. In this section we show that the answer is positive.

Lemma 4.5.1. *Let $X = \{X_t, t \in [0,T]\}$ be a (t,ω)-measurable stochastic process such that, for all $t \in [0,T]$, the random variable X_t is $\mathcal{F}_T$-measurable and $\mathbb{E}(X_t^2) < \infty$. Let*

$$X(t) = \sum_{n=0}^{\infty} I_n\left(f_n(\cdot, t)\right)$$

be its Wiener-Itô chaos expansion. Then, X is $\{\mathcal{F}_t\}$-adapted if and only if

$$f_n(t_1, \dots, t_n, t) = 0$$

for almost every $(t_1, \dots, t_n) \in \left\{(s_1, \dots, s_n) \in [0,T]^n : t < \max_{1 \le i \le n} s_i\right\}.$

Proof. Let $g \in \hat{L}^2([0,T])$. Since Itô integrals are martingales, for all $t \in [0,T]$, we have

$$\begin{aligned}
\mathbb{E}\left[I_n(g) \middle| \mathcal{F}_t \right] &= n!\mathbb{E}\left[\int_0^T \int_0^{t_n} \cdots \int_0^{t_2} g(t_1,\ldots,t_n) dB_{t_1} \ldots dB_{t_n} \middle| \mathcal{F}_t \right] \\
&= n! \int_0^t \int_0^{t_n} \cdots \int_0^{t_2} g(t_1,\ldots,t_n) dB_{t_1} \ldots dB_{t_n} \\
&= n! \int_0^T \mathbb{1}_{\{0 \le t_n < t\}} \int_0^{t_n} \cdots \int_0^{t_2} g(t_1,\ldots,t_n) dB_{t_1} \ldots dB_{t_n} \\
&= n! J_n \left(g(t_1,\ldots,t_n) \mathbb{1}_{\{\max_i t_i < t\}} \right) \\
&= I_n \left(g(t_1,\ldots,t_n) \mathbb{1}_{\{\max_i t_i < t\}} \right). \qquad (4.5.1)
\end{aligned}$$

The process X is $\{\mathcal{F}_t\}$-adapted if and only if, for all $n \in \mathbb{N}$, $\mathbb{E}\left(X_t | \mathcal{F}_t \right) = X_t$. In terms of the Wiener-Itô chaos expansion, X is $\{\mathcal{F}_t\}$-adapted if and only if

$$\mathbb{E}\left(\sum_{n=0}^{\infty} I_n\left(f_n(\cdot,t)\right) \big| \mathcal{F}_t \right) = \sum_{n=0}^{\infty} I_n\left(f_n(\cdot,t)\right).$$

Since the series converges in $L^2(\Omega)$, X is $\{\mathcal{F}_t\}$-adapted if and only if

$$\sum_{n=0}^{\infty} \mathbb{E}\left(I_n\left(f_n(\cdot,t)\right) \big| \mathcal{F}_t \right) = \sum_{n=0}^{\infty} I_n\left(f_n(\cdot,t)\right).$$

By (4.5.1), X is $\{\mathcal{F}_t\}$-adapted if and only if

$$\sum_{n=0}^{\infty} I_n\left(f_n(\cdot,t)\mathbb{1}_{\{\max_i t_i < t\}}\right) = \sum_{n=0}^{\infty} I_n\left(f_n(\cdot,t)\right) = X(t).$$

By uniqueness of the Wiener-Itô chaos expansion, X is $\{\mathcal{F}_t\}$-adapted if and only if for all $n \in \mathbb{N}$

$$f_n(t_1,\ldots,t_n,t)\mathbb{1}_{\{\max_i t_i < t\}} = f_n(t_1,\ldots,t_n,t)$$

for almost every $(t_1,\ldots,t_n) \in [0,T]^n$. This is equivalent to the statement of the lemma. □

Theorem 4.5.2. *Let $X = \{X_t, t \in [0,T]\}$ be a stochastic process in $L^2_{ad}([0,T] \times \Omega)$. Then, X is Skorohod integrable, $X \in Dom(\delta)$, and*

$$\int_0^T X(t)\delta B_t = \int_0^T X(t) dB_t.$$

Proof. For almost every $t \in [0,T]$, we have

$$X(t,\omega) = \sum_{n=0}^{\infty} I_n\left(f_n(\cdot,t)\right)$$

in $L^2(\Omega)$. For each $n \in \mathbb{N}$, $f_n(t_1, \dots, t_n, t_{n+1})$ is symmetric with respect to the first n variables. We have already seen that its symmetrization is given by

$$\hat{f}_n(t_1, \dots, t_{n+1}) = \frac{1}{n+1}\left[f_n(t_1, \dots, t_{n+1}) + f_n(t_{n+1}, t_2, \dots, t_1) + \dots + f_n(t_1, \dots, t_{n+1}, t_n)\right].$$

Since X is $\{\mathcal{F}_t\}$-adapted, by Lemma 4.5.1,

$$\hat{f}_n(t_1, \dots, t_{n+1}) = \frac{1}{n+1} f_n(x_1, \dots, x_{n+1}) \tag{4.5.2}$$

where $x_i \in \{t_1, \dots, t_{n+1}\}$ and $x_{n+1} = \max(t_1, \dots, t_{n+1})$. Then, we have that

$$\begin{aligned}
||\hat{f}_n||^2_{L^2([0,T]^{n+1})} &= (n+1)! \int_{S_{n+1}} \left[\hat{f}_n(t_1, \dots, t_{n+1})\right]^2 dt_1 \dots dt_{n+1} \\
&= \frac{(n+1)!}{(n+1)^2} \int_0^T \left(\int_0^t \int_0^{x_n} \dots \int_0^{x_2} \left[f_n(x_1, \dots, x_n, t)\right]^2 dx_1 \dots dx_n \right) dt \\
&= \frac{n!}{n+1} \int_0^T \left(\int_0^T \int_0^{x_n} \dots \int_0^{x_2} \left[f_n(x_1, \dots, x_n, t)\right]^2 dx_1 \dots dx_n \right) dt \\
&= \frac{n!}{n+1} \int_0^T ||\hat{f}_n(\cdot, t)||^2_{L^2(S_n)} dt \\
&= \frac{1}{n+1} \int_0^T ||\hat{f}_n(\cdot, t)||^2_{L^2([0,T]^n)} dt.
\end{aligned} \tag{4.5.3}$$

Then, using (4.5.3),

$$\begin{aligned}
\sum_{n=0}^{\infty} (n+1)! ||\hat{f}_n||^2_{L^2([0,T]^{n+1})} &= \sum_{n=0}^{\infty} (n+1)! \frac{1}{n+1} \int_0^T ||\hat{f}_n(\cdot, t)||^2_{L^2([0,T]^n)} dt \\
&= \sum_{n=0}^{\infty} n! \int_0^T ||\hat{f}_n(\cdot, t)||^2_{L^2([0,T]^n)} dt \\
&= \int_0^T \sum_{n=0}^{\infty} n! ||\hat{f}_n(\cdot, t)||^2_{L^2([0,T]^n)} dt \\
&= \int_0^T \mathbb{E}\left(X_t^2\right) dt < \infty.
\end{aligned}$$

The limit and the integral have been exchanged by the monotone convergence theorem. By Remark 4.4.3, X is Skorohod integrable. Our next goal is to prove that

$$\int_0^T X(t) dB_t = \sum_{n=0}^{\infty} \int_0^T I_n(f_n(\cdot, t)) dB_t. \tag{4.5.4}$$

Indeed, by the Itô isometry,

$$\left\| \int_0^T X(t) dB_t - \sum_{n=0}^{N} \int_0^T I_n(f_n(\cdot, t)) dB_t \right\|^2_{L^2(\Omega)} = \left\| \int_0^T \left[X(t) - \sum_{n=0}^{N} I_n(f_n(\cdot, t)) \right] dB_t \right\|^2_{L^2(\Omega)}$$

$$= \int_0^T \left\| X(t) - \sum_{n=0}^{N} I_n(f_n(\cdot,t)) \right\|_{L^2(\Omega)}^2 dt$$
$$\to 0$$

as $n \to \infty$ by the dominated convergence theorem. We dominate by

$$\begin{aligned}
\int_0^T \left\| X(t) - \sum_{n=0}^{N} I_n(f_n(\cdot,t)) \right\|_{L^2(\Omega)}^2 dt &\leq 2 \int_0^T \left[\|X(t)\|_{L^2(\Omega)}^2 + \sum_{n=0}^{N} \|I_n(f_n(\cdot,t))\|_{L^2(\Omega)}^2 \right] dt \\
&\leq 2 \int_0^T \left[\mathbb{E}(X_t^2) + \sum_{n=0}^{\infty} n! \, \|f_n(\cdot,t)\|_{L^2([0,T]^n)}^2 \right] dt \\
&= 4 \int_0^T \mathbb{E}(X_t^2) dt < \infty.
\end{aligned}$$

We conclude that

$$\begin{aligned}
\int_0^T X(t) dB_t &= \sum_{n=0}^{\infty} \int_0^T I_n(f_n(\cdot,t)) dB_t \\
&= \sum_{n=0}^{\infty} \int_0^T n! \int_0^{t_n} \dots \int_0^{t_2} f_n(t_1,\dots,t_n,t) dB_{t_1} \dots dB_{t_n} dB_t \\
&= \sum_{n=0}^{\infty} \int_0^T n!(n+1) \int_0^{t_n} \dots \int_0^{t_2} \hat{f}_n(t_1,\dots,t_n,t_{n+1}) dB_{t_1} \dots dB_{t_n} dB_{t_{n+1}} \\
&= \sum_{n=0}^{\infty} (n+1)! J_{n+1}(\hat{f}_n) \\
&= \sum_{n=0}^{\infty} I_{n+1}(\hat{f}_n) =: \int_0^T X(t) \delta B_t.
\end{aligned}$$

□

Chapter 5

A New Stochastic Integral: the Ayed-Kuo Integral

In this final chapter, we study a new stochastic integral which generalizes Itô integration in the sense that, as Skorohod integration, it deals with anticipating calculus.

This integral is due to W. Ayed and H.-H. Kuo, who introduced it for the first time in [3] in 2008. Their idea is inspired by Itô original construction of the integral based on Riemann sums, which consisted in evaluating the integrand at the left endpoints of the intervals of a partition. In the new setting, the integrand is assumed to be a product of an adapted stochastic process with respect to a Brownian filtration and an instantly independent process. Then, the adapted process will be evaluated at the left endpoints of a partition while the instantly independent process at the right endpoints. The integral is defined by taking limits.

We will examine analogous properties of the Itô integral for this integral, which have been studied from 2008 until current time. We also focus on a stochastic differential equation with an anticipating initial condition in order to show that stochastic differential equations are an important motivation and application of the development of new integrals.

We also aim to outline the simplicity of this integral compared to Skorohod integration and the fact that the probabilistic meaning is much more clear. Many questions about this integral are still open and summarized at the end of this chapter.

5.1 Definition and Examples

We consider a Brownian motion $\{B_t, t \geq 0\}$ and a filtration $\{\mathcal{F}_t, t \geq 0\}$ such that

(i) For all $t \geq 0$, B_t is $\mathcal{F}_t$-measurable.

(ii) For all $0 \leq s \leq t$, $B_t - B_s$ is independent of $\mathcal{F}_s$.

We fix a time horizon $T > 0$.

Definition 5.1.1. A stochastic process $\{\varphi(t),\, t \in [0,T]\}$, is *instantly independent* with respect to the filtration $\{\mathcal{F}_t, t \in [0,T]\}$ if and only if $\varphi(t)$ is independent of $\mathcal{F}_t$ for each $t \in [0,T]$.

Example 5.1.2. We provide some examples of instantly independent processes with respect to the Brownian filtration $\{\mathcal{F}_t, t \in [0,T]\}$.

(1) By the definition of Brownian motion, $B(T)-B(t)$ is independent of the σ-field $\mathcal{F}_t$ for all $t \in [0,T]$. Hence, $\varphi(t) = B(T)-B(t)$ is an instantly independent stochastic process.

(2) Likewise, $\varphi(t) = e^{B(T)-B(t)}$, $t \in [0,T]$, is an instantly independent process.

(3) For any deterministic function $h \in L^2([0,T])$,

$$\varphi(t) = \int_t^T h(s)dB_s, \qquad t \in [0,T],$$

is an instantly independent process.

(4) The process $\varphi(t) = B(T)$, $t \in [0,T]$, is not an instantly independent process.

Proposition 5.1.3. *If $\{\varphi(t), t \in [0,T]\}$ is an instantly independent stochastic process with respect to $\{\mathcal{F}_t, t \in [0,T]\}$ and also $\{\mathcal{F}_t\}$-adapted, then $\varphi(t)$ is a deterministic function.*

Proof. Since φ is $\{\mathcal{F}_t\}$-adapted,

$$\mathbb{E}\left(\varphi(t)|\,\mathcal{F}_t\right) = \varphi(t) \tag{5.1.1}$$

for all $t \in [0,T]$. Since the process is instantly independent,

$$\mathbb{E}\left(\varphi(t)|\,\mathcal{F}_t\right) = \mathbb{E}(\varphi(t)) \tag{5.1.2}$$

for all $t \in [0,T]$. Putting together (5.1.1) and (5.1.2), we conclude that $\mathbb{E}(\varphi(t)) = \varphi(t)$ for all $t \in [0,T]$. Hence, φ is a deterministic function. □

Definition 5.1.4. Let $\{f(t), t \in [0,T]\}$ be a $\{\mathcal{F}_t\}$-adapted stochastic process and consider $\{\varphi(t), t \in [0,T]\}$ an instantly independent stochastic process with respect to $\mathcal{F}_t$. We define the *Ayed-Kuo stochastic integral* of $f(t)\varphi(t)$ by

$$\int_0^T f(t)\varphi(t)dB_t := \lim_{|\Pi_n|\to 0} \sum_{i=1}^n f(t_{i-1})\varphi(t_i)\left[B(t_i) - B(t_{i-1})\right] \tag{5.1.3}$$

in probability, provided that the limit exists, where Π_n are partitions of the interval $[0,T]$.

When the Itô integral was built in Chapter 2, we showed that it could be understood in terms of Riemann sums by evaluating the integrand at the left endpoints of the intervals of the partition, see Theorem 2.6.1. In Definition 5.1.4, we follow the same idea with the adapted process. On the other hand, the instantly independent process is evaluated at the right endpoints of the intervals of the partition in order to take advantage of the independence property.

Remark 5.1.5. Note that if we consider $\varphi(t) = 1$ and $f(t) \in \mathcal{L}_{ad}$ pathwise continuous, then, by Theorem 2.6.1, the Ayed-Kuo integral coincides with the Itô integral.

One could argue that the integral strongly relies on the factorized structure of the integrand, but we will see that many anticipating integrals can be converted to a product of an adapted and an instantly stochastic processes.

Remark 5.1.6. Looking at Equation (5.1.3), the Ayed-Kuo integral is clearly linear.

Next, we provide some examples in order to show that the Ayed-Kuo allows us to compute some anticipating integrals that we have already computed in Chapter 4 and check that we obtain the same results.

Example 5.1.7. We want to calculate the integral $\int_0^T B(T)dB_t$ in Example 4.4.4. By linearity,

$$\int_0^T B(T)dB_t = \int_0^T (B(T) - B(t))\, dB_t + \int_0^T B(t)dB_t. \tag{5.1.4}$$

The second integral on the right-hand side of (5.1.4) has Itô sense, while the first one must be considered in Ayed-Kuo sense. Note that $\varphi(t) = B(T) - B(t)$ is an instantly independent stochastic process with respect to the filtration according to what we discussed in the first example of Example 5.1.2. Moreover, we are taking $f(t) = 1$. Therefore, we are going to calculate the following integral

$$\int_0^T (B(T) - B(t))\, dB_t := \mathbb{P} - \lim_{|\Pi_n| \to 0} \sum_{i=1}^n (B(T) - B(t_i))\,(B(t_i) - B(t_{i-1}))\,. \tag{5.1.5}$$

Note that

$$\begin{aligned}\sum_{i=1}^n (B(T) - B(t_i))\,(B(t_i) - B(t_{i-1})) &= B(T)\sum_{i=1}^n (B(t_i) - B(t_{i-1})) \\ &\quad - \sum_{i=1}^n B(t_i)(B(t_i) - B(t_{i-1})) \\ &= B(T)^2 - \left[\sum_{i=1}^n (B(t_i) - B(t_{i-1}))^2\right. \\ &\quad \left. + \sum_{i=1}^n B(t_{i-1})\,(B(t_i) - B(t_{i-1}))\right]\end{aligned}$$

By the quadratic variation of Brownian motion (Theorem 1.5.1) and Theorem 2.6.1, we obtain that

$$\mathbb{P} - \lim_{|\Pi_n| \to 0} \sum_{i=1}^n (B(T) - B(t_i))\,(B(t_i) - B(t_{i-1})) = B(T)^2 - T - \int_0^T B(t)dB_t. \tag{5.1.6}$$

Hence, substituting (5.1.6) into (5.1.5), we obtain

$$\int_0^T (B(T) - B(t))\, dB_t = B(T)^2 - T - \int_0^T B(t)dB_t. \tag{5.1.7}$$

Substituting (5.1.7) into (5.1.4), we get that

$$\int_0^T B(T)dB_t = B(T)^2 - T,$$

which coincides with $\int_0^T B(T)\delta B(t)$ calculated in Example 4.4.4.

Example 5.1.8. Consider the stochastic integral $I_t = \int_0^t (B(T) - B(s))B(s)dB_s$ with $0 \le t \le T$. We are going to calculate the integral with the same strategies as in Example 5.1.7.
For any $0 \le t \le T$,

$$\begin{aligned}
I_t &:= \lim_{|\Pi_n|\to 0} \sum_{i=1}^n (B(T) - B(s_i))\, B(s_{i-1})\,(B(s_i) - B(s_{i-1})) \\
&= \lim_{|\Pi_n|\to 0} \left[B(T) \sum_{i=1}^n B(s_{i-1})\,(B(s_i) - B(s_{i-1})) - \sum_{i=1}^n B(s_i)B(s_{i-1})\,(B(s_i) - B(s_{i-1})) \right] \\
&= B(T) \int_0^t B(s)dB(s) - \lim_{|\Pi_n|\to 0} \sum_{i=1}^n B(s_{i-1})\,(B(s_i) - B(s_{i-1}))^2 \\
&\quad - \lim_{|\Pi_n|\to 0} \sum_{i=1}^n B(s_{i-1})^2\,(B(s_i) - B(s_{i-1})) \\
&= B(T) \int_0^t B(s)dB(s) - \int_0^t B(s)ds - \int_0^t B(s)^2 dB_s \\
&= \frac{B(T)}{2}\left[B(t)^2 - t\right] - \frac{B(t)^3}{3}.
\end{aligned}$$

When $t = T$, we have that

$$\int_0^T (B(T) - B(s))B(s)dB_s = \frac{1}{6}\left[B(T)^3 - 3TB(T)\right],$$

which coincides with $\int_0^T (B(T) - B(s))B(s)\delta B_s$ obtained in Example 4.4.5.

5.2 Properties of the Ayed-Kuo Integral

In this section, we study some analogue properties of the ones studied in Itô integration. We show that the Ayed-Kuo integrals have zero mean, the integral stochastic process is a near-martingale and we give a more general version of the Itô isometry.

5.2.1 The Zero Mean Property

In Example (5.1.7) and Example (5.1.8), one can easily check that in both cases the integrals have zero mean. Thus, in the next theorem we prove that whenever the Ayed-Kuo integral exists, it has zero mean.

Theorem 5.2.1. *Let* $\{f(t), t \in [0,T]\}$ *be an* $\{\mathcal{F}_t\}$*-adapted stochastic process and consider* $\{\varphi(t), t \in [0,T]\}$ *an instantly independent stochastic process with respect to the filtration. If the Ayed-Kuo integral* $\int_0^T f(t)\varphi(t)dB_t$ *exists,* $\mathbb{E}\,|f(t)| < \infty$ *and* $\mathbb{E}\,|\varphi(t)| < \infty$ *for all* $t \in [0,T]$*, then*

$$\mathbb{E}\int_0^T f(t)\varphi(t)dB_t = 0.$$

Proof. We consider a partition $\Pi_n = \{t_0 = 0 < t_1 < \ldots < t_n = T\}$. For the sake of simplicity, we denote $\Delta B_i = B(t_i) - B(t_{i-1})$. By the properties of the conditional expectation, we have

$$\begin{aligned}
\mathbb{E}\left[f(t_{i-1})\varphi(t_i)\Delta B_i\right] &= \mathbb{E}\left[\mathbb{E}\left[\left. f(t_{i-1})\varphi(t_i)\Delta B_i\right| \mathcal{F}_{t_i}\right]\right] \\
&= \mathbb{E}\left[f(t_{i-1})\Delta B_i \mathbb{E}\left[\left.\varphi(t_i)\right| \mathcal{F}_{t_i}\right]\right] \\
&= \mathbb{E}\left[\varphi(t_i)\right]\mathbb{E}\left[f(t_{i-1})\Delta B_i\right] \\
&= \mathbb{E}\left[\varphi(t_i)\right]\mathbb{E}\left[\mathbb{E}\left[\left. f(t_{i-1})\Delta B_i\right| \mathcal{F}_{t_{i-1}}\right]\right] \\
&= \mathbb{E}\left[\varphi(t_i)\right]\mathbb{E}\left[f(t_{i-1})\mathbb{E}\left[\left.\Delta B_i\right| \mathcal{F}_{t_{i-1}}\right]\right] \\
&= \mathbb{E}\left[\varphi(t_i)\right]\mathbb{E}\left[f(t_{i-1})\mathbb{E}\left[\Delta B_i\right]\right] \\
&= 0.
\end{aligned}$$

By definition of the integral,

$$\int_0^T f(t)\varphi(t)dB_t = \lim_{|\Pi_n|\to 0}\sum_{i=1}^{n} f(t_{i-1})\varphi(t_i)\Delta B_i.$$

in probability. We know that the limit also holds in $L^1(\Omega)$ by taking a subsequence if necessary. We denote by S_n the partial sums and by I the integral. We have already proved that $\mathbb{E}(S_n) = 0$. Then,

$$|\mathbb{E}(I)| = |\mathbb{E}(I - S_n) + \mathbb{E}(S_n)| \le \mathbb{E}\,|I - S_n| \to 0$$

as $n \to \infty$. □

5.2.2 The Near-martingale Property

One of the most important properties of the indefinite Itô integral is the martingale property. In this section, we will show that the Ayed-Kuo integral does not fulfill this property. However, we will introduce the notion of near-martingale and we will check that the Ayed-Kuo satisfies this condition. In the next example, we show that Ayed-Kuo integrals are not martingales.

Example 5.2.2. We consider Example 5.1.7 with a slightly modification. We aim to calculate the integral $\int_0^t B(1)dB(s)$ with $0 \le t \le 1$. By linearity,

$$\int_0^t B(1)dB(s) = \int_0^t \left(B(1) - B(s)\right)dB(s) + \int_0^t B(s)dB(s). \tag{5.2.1}$$

We only need to compute the first integral on the right-hand side.

$$\int_0^t (B(1) - B(s))\, dB_s := \mathbb{P} - \lim_{|\Pi_n| \to 0} \sum_{i=1}^n (B(1) - B(s_i))\,(B(s_i) - B(s_{i-1}))\,. \qquad (5.2.2)$$

Note that

$$\begin{aligned}\sum_{i=1}^n (B(1) - B(s_i))\,(B(s_i) - B(s_{i-1})) &= B(1)B(t) - \sum_{i=1}^n B(s_i)(B(s_i) - B(s_{i-1})) \\ &= B(1)B(t) - \Bigg[\sum_{i=1}^n (B(s_i) - B(s_{i-1}))^2 \\ &\qquad + \sum_{i=1}^n B(s_{i-1})\,(B(s_i) - B(s_{i-1}))\Bigg]\end{aligned}$$

By the quadratic variation of Brownian motion (Theorem 1.5.1) and Theorem 2.6.1, we obtain that

$$\mathbb{P} - \lim_{|\Pi_n| \to 0} \sum_{i=1}^n (B(1) - B(s_i))\,(B(s_i) - B(s_{i-1})) = B(1)B(t) - t - \int_0^t B(s)dB(s).$$

Hence, substituting into (5.2.2), we get that

$$\int_0^t (B(1) - B(s))\, dB_s = B(1)B(t) - t - \int_0^t B(s)dB(s).$$

Therefore,

$$\int_0^t B(1)dB(s) = B(1)B(t) - t$$

for any $0 \leq t \leq 1$. We define the stochastic process

$$X_t := \int_0^t B(1)dB_s = B(1)B(t) - t$$

with $t \in [0, 1]$. For any $s \in [0, t]$, we have that

$$\begin{aligned}\mathbb{E}\,(X_t |\, \mathcal{F}_s) &= \mathbb{E}\,(B(1)B(t) - t |\, \mathcal{F}_s) \\ &= \mathbb{E}\,[\,[(B(1) - B(t)) + (B(t) - B(s)) + B(s)]\,[(B(t) - B(s)) + B(s)]|\, \mathcal{F}_s] - t \\ &= t - s + B_s^2 - t \\ &= B_s^2 - s \neq X_s.\end{aligned}$$

Hence, X_t is not a martingale.

Although the Ayed-Kuo is not a martingale, we will show that it satisfies a similar property. To this purpose, we introduce the concept of near-martingale.

Definition 5.2.3. A stochastic process $\{X_t, t \geq 0\}$ is a *near-martingale* with respect to the filtration $\{\mathcal{F}_t, t \geq 0\}$ if

(i) For all $t \geq 0$, $\mathbb{E}\,|X_t| < \infty$.

(ii) For all $s \in [0,t]$, $\mathbb{E}(X_t|\,\mathcal{F}_s) = \mathbb{E}(X_s|\,\mathcal{F}_s)$ almost surely.

The definition of near-martingale means that having information up to current time, the prediction for future times is the same as for the current one. Note that when the process is adapted to the filtration, then we recover the martingale notion.

Example 5.2.4. In Example 5.2.2, we have shown that the process $X_t = B(1)B(t) - t$, $t \in [0,1]$, is not a martingale. We check that it is a near-martingale. We have already shown that $\mathbb{E}\,(X_t|\,\mathcal{F}_s) = B_s^2 - s$ for $0 \leq s \leq t \leq 1$. On the other hand,

$$\mathbb{E}\,(X_s|\,\mathcal{F}_s) := \mathbb{E}\,(B(1)B(s) - s|\,\mathcal{F}_s) = B(s)\mathbb{E}\,(B(1)|\,\mathcal{F}_s) - s = B(s)^2 - s.$$

Hence, the process $\{X_t, t \in [0,1]\}$ is a near-martingale.

Remark 5.2.5. One can easily check that a near-martingale has the fair-game property. By the properties of the conditional expectation, for all $0 \leq s \leq t$,

$$\mathbb{E}(X_t) = \mathbb{E}\,[\mathbb{E}(X_t|\,\mathcal{F}_s)] = \mathbb{E}\,[\mathbb{E}(X_s|\,\mathcal{F}_s)] = \mathbb{E}(X_s).$$

Hence, for all $t \geq 0$, $\mathbb{E}(X_t) = \mathbb{E}(X_0)$.

Theorem 5.2.6. *Let $\{f(t), t \in [0,T]\}$ be an $\{\mathcal{F}_t\}$-adapted stochastic process and consider $\{\varphi(t), t \in [0,T]\}$ an instantly independent stochastic process with respect to the filtration. Assume that $f(t)$ and $\varphi(t)$ have continuous sample paths with probability one. If*

$$\mathbb{E}\left|\int_0^T f(t)\varphi(t)dB(t)\right| < \infty,$$

then the stochastic process

$$X_t = \int_0^t f(s)\varphi(s)dB(s), \qquad t \in [0,T],$$

is a near-martingale with respect to the filtration.

Proof. It is enough to prove that $\mathbb{E}(X_t - X_s|\,\mathcal{F}_s) = 0$ almost surely. By the properties of Riemann sums, it is clear that

$$X_t - X_s = \int_s^t f(u)\varphi(u)dB(u)$$

for all $0 \leq s \leq t \leq T$. We consider partitions $\Pi_n = \{s = u_0 < u_1 < \ldots < u_n = t\}$. By the properties of the conditional expectation,

$$\mathbb{E}\,(f(u_{i-1})\varphi(u_i)\Delta B_i|\,\mathcal{F}_s) = \mathbb{E}\left[\mathbb{E}\left[f(u_{i-1})\varphi(u_i)\Delta B_i|\,\mathcal{F}_{u_{i-1}}\right]\big|\,\mathcal{F}_s\right]$$

$$
\begin{aligned}
&= \mathbb{E}\left[f(u_{i-1})\,\mathbb{E}\left[\varphi(u_i)\Delta B_i |\, \mathcal{F}_{u_{i-1}}\right] \big|\, \mathcal{F}_s\right] \\
&= \mathbb{E}\left[f(u_{i-1})\,\mathbb{E}\left[\mathbb{E}\left[\varphi(u_i)\Delta B_i |\, \mathcal{F}_{u_i}\right] |\, \mathcal{F}_{u_{i-1}}\right] \big|\, \mathcal{F}_s\right] \\
&= \mathbb{E}\left[f(u_{i-1})\,\mathbb{E}\left[\Delta B_i \mathbb{E}\left[\varphi(u_i) |\, \mathcal{F}_{u_i}\right] |\, \mathcal{F}_{u_{i-1}}\right] \big|\, \mathcal{F}_s\right] \\
&= \mathbb{E}\left[\varphi(u_i)\right] \mathbb{E}\left[f(u_{i-1})\,\mathbb{E}\left[\Delta B_i |\, \mathcal{F}_{u_{i-1}}\right] \big|\, \mathcal{F}_s\right] \\
&= \mathbb{E}\left[\varphi(u_i)\right] \mathbb{E}\left[\Delta B_i\right] \mathbb{E}\left[f(u_{i-1}) \,|\, \mathcal{F}_s\right] \\
&= 0.
\end{aligned}
$$

Then,

$$
\begin{aligned}
\mathbb{E}\left(X_t - X_s |\, \mathcal{F}_s\right) &= \mathbb{E}\left(\int_s^t f(u)\varphi(u)dB(u) \middle|\, \mathcal{F}_s\right) \\
&= \mathbb{E}\left(\lim_{|\Pi_n|\to 0} \sum_{i=1}^n f(u_{i-1})\varphi(u_i)\Delta B_i \middle|\, \mathcal{F}_s\right) \\
&= \lim_{|\Pi_n|\to 0} \sum_{i=1}^n \mathbb{E}\left(f(u_{i-1})\varphi(u_i)\Delta B_i |\, \mathcal{F}_s\right) \\
&= 0.
\end{aligned}
$$

The limit and the integral were exchanged as there is convergence in $L^1(\Omega)$ for some subsequence. □

5.2.3 The Isometry Property

In previous chapters, we have outlined that the Itô isometry is really useful for calculating second moments of Itô integrals. In this section, we give a similar new isometry for the Ayed-Kuo integral.

Theorem 5.2.7. *Let $f, \varphi \in \mathcal{C}^\infty(\mathbb{R})$ be two functions whose McLaurin expansion has infinite radius of convergence. Then,*

$$
\begin{aligned}
\mathbb{E}\left[\left(\int_0^T f(B_t)\varphi(B_T - B_t)dB_t\right)^2\right] &= \int_0^T \mathbb{E}\left[f(B_t)^2\varphi(B_t)^2\right] dt \\
&\quad + 2\int_0^T \int_0^t \mathbb{E}\left[f(B_s)\varphi'(B_T - B_s)f'(B_t)\varphi(B_T - B_t)\right] dt.
\end{aligned}
$$

We omit the proof as it is rather long and based on McLaurin expansions and properties of the conditional expectation. The proof can be found in [28].

Remark 5.2.8. Note that if $\varphi(x) = 1$, we recover the same isometry as in the Itô theory. If $f(x) = 1$, we also have the isometry of the Itô theory.

We also remark that the identity in Theorem 5.2.7 is for a specific type of stochastic processes. However, this includes polynomial and exponential functions of Brownian motion and, hence, all the examples we have considered so far and, in particular, solutions of a Black-Scholes equation.

We give some examples in order to show how the identity in Theorem 5.2.7 works.

Example 5.2.9. In Example 5.1.8, we showed that

$$\int_0^T B(t)[B(T) - B(t)]dB(t) = \frac{1}{6}B(T)^3 - \frac{1}{2}TB(T).$$

We will check the new isometry for a particular example. First, we begin with a direct derivation,

$$\begin{aligned}
\mathbb{E}\left[\left(\int_0^T B(t)[B(T) - B(t)]dB(t)\right)^2\right] &= \mathbb{E}\left[\left(\frac{1}{6}B(T)^3 - \frac{1}{2}TB(T)\right)^2\right] \\
&= \frac{1}{36}\mathbb{E}[B(T)^6] - \frac{1}{6}T\mathbb{E}[B(T)^4] + \frac{1}{4}T^2\mathbb{E}[B(T)^2] \\
&= \frac{1}{36}5!!T^3 - \frac{1}{6}T3!!T^2 + \frac{1}{4}T^2T \\
&= T^3\left(\frac{15}{36} - \frac{1}{2} + \frac{1}{4}\right) \\
&= \frac{1}{6}T^3.
\end{aligned}$$

Now, we use Theorem 5.2.7. In this case, $f(x) = \varphi(x) = x$ and, hence, $f'(x) = \varphi'(x) = 1$.

$$\begin{aligned}
\mathbb{E}\left[\left(\int_0^T B(t)[B(T) - B(t)]dB(t)\right)^2\right] &= \int_0^T \mathbb{E}\left[B(T)^2\left(B(T) - B(t)\right)^2\right]dt \\
&+ \int_0^T\int_0^t \mathbb{E}\left[B(s)\left(B(T) - B(t)\right)\right]dsdt
\end{aligned}$$

Note that $B(T) - B(t)$ is independent of $B(s)$ and, since Brownian motion has zero mean, we get that $\mathbb{E}\left[B(s)\left(B(T) - B(t)\right)\right] = 0$. Then,

$$\begin{aligned}
\mathbb{E}\left[\left(\int_0^T B(t)[B(T) - B(t)]dB(t)\right)^2\right] &= \int_0^T \mathbb{E}\left[B(t)^2\right]\mathbb{E}\left[(B(T) - B(t))^2\right]dt \\
&= \int_0^T t(T - t)dt \\
&= \frac{1}{2}T^3 - \frac{1}{3}T^3 \\
&= \frac{1}{6}T^3.
\end{aligned}$$

We conclude that the result is the same in both methods, as we expected.

Example 5.2.10. We will compute the second moment of the following integral

$$\int_0^T B(t)^2(B(T) - B(t))^2dB(t).$$

We note that in this case $f(x) = \varphi(x) = x^2$, which fulfill the conditions of Theorem 5.2.7. The derivatives are $f'(x) = \varphi'(x) = 2x$. Applying the new isometry, we get that

$$\mathbb{E}\left[\left(\int_0^T B(t)^2(B(T)-B(t))^2 dB(t)\right)^2\right] = \int_0^T \mathbb{E}\left[B(t)^4(B(T)-B(t))^4\right] dt$$
$$+ 2\int_0^T \int_0^t \mathbb{E}\left[B(s)^2 2(B(T)-B(s))2B(t)(B(T)-B(t))^2\right] dsdt$$

For the first term, we have that

$$\begin{aligned}\int_0^T \mathbb{E}\left[B(t)^4(B(T)-B(t))^4\right] dt &= \int_0^T \mathbb{E}\left[B(t)^4\right]\left[(B(T)-B(t))^4\right] dt\\ &= 9\int_0^T t^2(T-t)^2 dt\\ &= \frac{3}{10}T^5.\end{aligned}$$

For the second one, note that

$$\begin{aligned}&2\int_0^T \int_0^t \mathbb{E}\left[B(s)^2 2(B(T)-B(s))2B(t)(B(T)-B(t))^2\right] dsdt\\ &= 8\int_0^T \int_0^t \mathbb{E}\left[B(s)^2(B(T)-B(t)+B(t)-B(s))(B(t)-B(s)+B(s))\right.\\ &\qquad\left.(B(T)-B(t))^2\right] dsdt\\ &= 8\int_0^T \int_0^t \mathbb{E}[B(s)^2]\mathbb{E}[((B(t)-B(s))^2]\mathbb{E}[((B(T)-B(t))^2]dsdt\\ &= 8\int_0^T \int_0^t s(t-s)(T-t)dsdt\\ &= \frac{8}{6}\int_0^T t^3(T-t)dt\\ &= \frac{1}{15}T^5.\end{aligned}$$

We conclude that

$$\mathbb{E}\left[\left(\int_0^T B(t)^2(B(T)-B(t))^2 dB(t)\right)^2\right] = T^5\left(\frac{3}{10}+\frac{1}{15}\right) = \frac{11}{30}T^5.$$

5.3 An Itô Formula for the Ayed-Kuo Integral

In this section, we will informally derive an Itô formula proposed in [15]. Then, we will check that the integrals in Example 5.1.7 and in Example 5.1.8 can be easily obtained by using this new change of variables rule. We consider an Itô process of the form

$$X_t = X_0 + \int_0^t g(s,\omega)dB(s) + \int_0^t \gamma(s,\omega)ds \tag{5.3.1}$$

where X_0 is a $\mathcal{F}_0$-measurable random variable, $g \in L^2_{ad}(\Omega \times [0,T])$ and $\gamma \in L^1([0,T])$ almost surely. Consider the process

$$Y^{(t)} = Y_0 + \int_t^T h(s)dB(s) + \int_t^T \chi(s)ds, \tag{5.3.2}$$

where Y_0 is a random variable independent of $\mathcal{F}_T$, and $h \in L^2([0,T])$ and $\chi \in L^1([0,T])$ are two deterministic functions. Note that under these assumptions all integrals make sense. Moreover, h, χ could be stochastic process (see [10]), but, for the purposes of this dissertation, we do not need to consider it. Recall that $\mathcal{F}_t$ is the filtration generated by a Brownian motion up to time t.

From Itô theory, it is clear that X_t is adapted to the filtration. We will briefly discuss that Y_t is an instantly independent stochastic process in the following proposition. We will use sub-t and sup-t to respectively denote the integrals $\int_0^t$ and $\int_t^T$.

Proposition 5.3.1. *Let $h \in L^2([0,T])$, $\chi \in L^1([0,T])$ be two deterministic functions and Y_0 a random variable independent of $\mathcal{F}_T$. Then,*

$$Y^{(t)} = Y_0 + \int_t^T h(s)dB(s) + \int_t^T \chi(s)ds, \ t \in [0,T],$$

is an instantly independent stochastic process with respect to the Brownian filtration.

Proof. We need to check that, for all $t \in [0,T]$, $Y^{(t)}$ is independent of $\mathcal{F}_t$. Note that Y_0 is independent of $\mathcal{F}_t$ for all $t \in [0,T]$ by assumption and the Lebesgue integral is also independent of $\mathcal{F}_t$ because it is deterministic. Therefore, we only need to study the stochastic integral. By virtue of Definition 5.1.4,

$$\int_t^T h(s)dB(s) := \lim_{|\Pi_n|\to 0} \sum_{i=1}^n h(s_i)\left[B(s_i) - B(s_{i-1})\right],$$

where $\Pi_n = \{t = s_0 < s_1 < \ldots < s_n = T\}$ is a partition of $[t,T]$. Note that the evaluation points of h are not important as it is a deterministic function. By the properties of Brownian motion, $h(s_i)\left[B(s_i) - B(s_{i-1})\right]$ is independent of $\mathcal{F}_t$ for all $1 \le i \le n$ because $t \le s_{i-1}$. Then, the sum

$$\sum_{i=1}^n h(s_i)\left[B(t_i) - B(t_{i-1})\right]$$

is also independent of $\mathcal{F}_t$. We conclude that the integral is also independent of $\mathcal{F}_t$ as a limit of independent random variables. □

Theorem 5.3.2 (Kuo, Sae-Tang, Szozda, 2012)**.** *Let $F(x,y) = f(x)\varphi(y)$ be a function such that $f, \varphi \in \mathcal{C}^2(\mathbb{R})$. Let X_t and $Y^{(t)}$, $0 \le t \le T$, be two stochastic processes as in*

Equation (5.3.1) *and Equation* (5.3.2)*, respectively. Then, the following equality holds almost surely for* $0 \leq t \leq T$*,*

$$F\left(X_t, Y^{(t)}\right) = F\left(X_0, Y^{(0)}\right) + \int_0^t F_x\left(X_s, Y^{(s)}\right) dX_s + \frac{1}{2}\int_0^t F_{xx}\left(X_s, Y^{(s)}\right)(dX_s)^2 + \int_0^t F_y\left(X_s, Y^{(s)}\right) dY^{(s)} - \frac{1}{2}\int_0^t F_{yy}\left(X_s, Y^{(s)}\right)(dY^{(s)})^2. \tag{5.3.3}$$

In differential form,

$$dF\left(X_t, Y^{(t)}\right) = F_x\left(X_s, Y^{(s)}\right) dX_s + \frac{1}{2}F_{xx}\left(X_s, Y^{(s)}\right)(dX_s)^2 + F_y\left(X_s, Y^{(s)}\right) dY^{(s)} - \frac{1}{2}F_{yy}\left(X_s, Y^{(s)}\right)(dY^{(s)})^2. \tag{5.3.4}$$

Proof. We consider a partition $\Pi_n = \{0 = t_0 < t_1 < \ldots < t_n = t\}$ of the interval $[0, t]$. We use the notation $\Delta X_i = X_{t_i} - X_{t_{i-1}}$. We begin expressing $F\left(X_t, Y^{(t)}\right) - F\left(X_0, Y^{(0)}\right)$ as a telescoping sum,

$$\begin{aligned} F\left(X_t, Y^{(t)}\right) - F\left(X_0, Y^{(0)}\right) &= \sum_{i=1}^n \left[F\left(X_{t_i}, Y^{(t_i)}\right) - F\left(X_{t_{i-1}}, Y^{(t_{i-1})}\right)\right] \\ &= \sum_{i=1}^n \left[f\left(X_{t_i}\right)\varphi\left(Y^{(t_i)}\right) - f\left(X_{t_{i-1}}\right)\varphi\left(Y^{(t_{i-1})}\right)\right]. \end{aligned} \tag{5.3.5}$$

In (5.3.5), we have to take the left endpoints of the intervals $[t_{i-1}, t_i]$ to evaluate every occurrence of f and the right endpoints of the intervals to evaluate every occurrence of φ, in order to get an Ayed-Kuo integral.
As in the proof of the classical Itô formula, we will use Taylor expansions up to second order. The restriction to second order is enough because for $k > 2$, we know that

$$o((\Delta X_i)^k) > o(\Delta t) \quad \text{and} \quad o((\Delta Y_i)^k) > o(\Delta t)$$

and, therefore, both $(\Delta X_i)^k$ and $(\Delta Y_i)^k$ tend to zero as $|\Pi_n| \to 0$.
We expand $f(X_{t_i})$ around the point $X_{t_{i-1}}$ with $1 \leq i \leq n$,

$$f(X_{t_i}) \approx f(X_{t_{i-1}}) + f'(X_{t_{i-1}})\Delta X_i + \frac{1}{2}f''(X_{t_{i-1}})\left(\Delta X_i\right)^2. \tag{5.3.6}$$

We expand $\varphi(Y^{t_{i-1}})$ around the point $Y^{(t_i)}$ for $1 \leq i \leq n$,

$$\varphi\left(Y^{(t_{i-1})}\right) \approx \varphi\left(Y^{(t_i)}\right) + \varphi'\left(Y^{(t_i)}\right)(-\Delta Y_i) + \frac{1}{2}\varphi''\left(Y^{(t_i)}\right)(-\Delta Y_i)^2. \tag{5.3.7}$$

Substituting (5.3.6) and (5.3.7) into (5.3.5), we get that

$$F\left(X_t, Y^{(t)}\right) - F\left(X_0, Y^{(0)}\right) \approx \sum_{i=1}^n \left[\left(f(X_{t_{i-1}}) + f'(X_{t_{i-1}})\Delta X_i + \frac{1}{2}f''(X_{t_{i-1}})\left(\Delta X_i\right)^2\right)\varphi\left(Y^{(t_i)}\right)\right.$$

$$
\begin{aligned}
&-f\left(X_{t_{i-1}}\right)\left(\varphi\left(Y^{(t_i)}\right)-\varphi'\left(Y^{(t_i)}\right)\Delta Y_i+\frac{1}{2}\varphi''\left(Y^{(t_i)}\right)(\Delta Y_i)^2\right)\Bigg]\\
&=\sum_{i=1}^{n}\Bigg[f'(X_{t_{i-1}})\varphi\left(Y^{(t_i)}\right)\Delta X_i+\frac{1}{2}f''(X_{t_{i-1}})\varphi\left(Y^{(t_i)}\right)(\Delta X_i)^2\\
&\quad+f\left(X_{t_{i-1}}\right)\varphi'\left(Y^{(t_i)}\right)\Delta Y_i-\frac{1}{2}f\left(X_{t_{i-1}}\right)\varphi''\left(Y^{(t_i)}\right)(\Delta Y_i)^2\Bigg]\\
&\to\int_0^t F_x\left(X_s,Y^{(s)}\right)dX_s+\frac{1}{2}\int_0^t F_{xx}\left(X_s,Y^{(s)}\right)(dX_s)^2\\
&\quad+\int_0^t F_y\left(X_s,Y^{(s)}\right)dY^{(s)}-\frac{1}{2}\int_0^t F_{yy}\left(X_s,Y^{(s)}\right)(dY^{(s)})^2,
\end{aligned}
$$

as $|\Pi_n|\to 0$. □

Corollary 5.3.3. *Consider a function $F(t,x,y)=\tau(t)f(x)\varphi(y)$ such that $f,\varphi\in\mathcal{C}^2(\mathbb{R})$ and $\tau\in\mathcal{C}^1([0,T])$. Let X_t and $Y^{(t)}$, $0\le t\le T$, be two stochastic processes as in Equation* (5.3.1) *and Equation* (5.3.2), *respectively. Then, the following equality holds almost surely for $0\le t\le T$,*

$$
\begin{aligned}
F\left(t,X_t,Y^{(t)}\right)=&F\left(0,X_0,Y^{(0)}\right)+\int_0^t F_s\left(s,X_s,Y^{(s)}\right)ds\\
&+\int_0^t F_x\left(s,X_s,Y^{(s)}\right)dX_s+\frac{1}{2}\int_0^t F_{xx}\left(s,X_s,Y^{(s)}\right)(dX_s)^2\\
&+\int_0^t F_y\left(s,X_s,Y^{(s)}\right)dY^{(s)}-\frac{1}{2}\int_0^t F_{yy}\left(s,X_s,Y^{(s)}\right)(dY^{(s)})^2. \quad (5.3.8)
\end{aligned}
$$

In differential form,

$$
\begin{aligned}
dF\left(X_t,Y^{(t)}\right)=&F_s\left(s,X_s,Y^{(s)}\right)ds+F_x\left(s,X_s,Y^{(s)}\right)dX_s+\frac{1}{2}F_{xx}\left(s,X_s,Y^{(s)}\right)(dX_s)^2\\
&+F_y\left(s,X_s,Y^{(s)}\right)dY^{(s)}-\frac{1}{2}F_{yy}\left(s,X_s,Y^{(s)}\right)(dY^{(s)})^2. \quad (5.3.9)
\end{aligned}
$$

Note that the classical Itô formula can be regarded as a particular case of (5.3.9) when $\varphi(t)=1$. Since in all examples we consider functions of Brownian motion, we are going to derive a particular case of the formula in Corollary 5.3.3.

Corollary 5.3.4. *Consider a function $F(t,x,y)=\tau(t)f(x)\varphi(y)$ such that $f,\varphi\in\mathcal{C}^2(\mathbb{R})$ and $\tau\in\mathcal{C}^1([0,T])$. Then, it holds that*

$$
dF(t,B_t,B_T)=\left[F_t+\frac{1}{2}F_{xx}+F_{xy}\right]dt+F_x dB_t. \quad (5.3.10)
$$

Proof. Note that $X_t=B_t$ is an adapted process. However, $Y_t=B_T$ is not an instantly independent process. So, we express it as

$$
B(T)=B(T)-B(t)+B(t).
$$

We define a function G such that $G(t, x, y) = F(t, x, x + y)$. Then,

$$dG(t, B_t, B_T - B_t) = dF(t, B_t, B(T)).$$

Thus, $dG(t, B_t, B_T - B_t)$ can be calculated using Corollary 5.3.3. We calculate the derivatives,

$$G_t = F_1, \quad G_x = F_2 + F_3, \quad G_y = F_3, \quad G_{xx} = F_{22} + 2F_{23} + F_{33} \text{ and } G_{yy} = F_{33},$$

where the indexes $1, 2, 3$ are referred to derivatives with respect to the first, second and third variables of F, respectively. By Equation (5.3.9),

$$\begin{aligned} dF(t, B_t, B(T)) &= G_t dt + G_x dB_t + \frac{1}{2} G_{xx} (dB_t)^2 + G_y(-dB_t) - \frac{1}{2} G_{yy} (-dB_t)^2 \\ &= F_1 dt + (F_2 + F_3) dB_t + \frac{1}{2}(F_{22} + 2F_{23} + F_{33}) dt - F_3 dB_t - \frac{1}{2} F_{33} dt \\ &= F_1 dt + F_2 dB_t + \frac{1}{2} F_{22} dt + F_{23} dt \\ &= \left[F_t + \frac{1}{2} F_{xx} + F_{xy} \right] dt + F_x dB_t. \end{aligned}$$

□

Next, we are going to calculate the integrals in the examples of Section 5.1 using the new Itô formulas we have just derived.

Example 5.3.5. We are going to calculate $\int_0^T B(T) dB_t$ using Equation (5.3.10). The function $F(t, x, y)$ that we need to consider is such that

$$F_x(t, x, y) = y.$$

Hence, we take $F(t, x, y) = xy$. Then,

$$d\left(B(t)B(T)\right) = B(T) dB_t + dt,$$

Integrating on both sides,

$$B(T)^2 = \int_0^T B(T) dB_t + T.$$

We conclude that

$$\int_0^T B(T) dB_t = B(T)^2 - T,$$

which is the result we obtained in Example 5.1.7.

Example 5.3.6. We are going to calculate $\int_0^T B(t) \left[B(T) - B(t)\right] dB_t$ using the general expression in Equation (5.3.4). We identify

$$X_t = B_t, \qquad dX_t = dB_t, \qquad Y_t = B(T) - B(t) \quad \text{and} \quad dY_t = -dB_t.$$

We are looking for a function $F(x, y)$ such that $F_x = xy$. Hence, $F(x, y) = x^2y/2$. By Equation (5.3.3),

$$\begin{aligned} &d\left(\frac{1}{2}B(t)^2(B(T) - B(t))\right) \\ &= B(t)(B(T) - B(t))dB_t + \frac{1}{2}(B(T) - B(t))dt + \frac{1}{2}B(t)^2(-dB_t). \end{aligned}$$

Integrating on both sides from 0 to T, it follows that

$$0 = \int_0^T B(t)(B(T) - B(t))dB_t + \frac{1}{2}\int_0^T (B(T) - B(t))dt - \frac{1}{2}\int_0^T B(t)^2 dB_t.$$

Hence,

$$\begin{aligned} \int_0^T B(t)(B(T) - B(t))dB_t &= -\frac{1}{2}\int_0^T (B(T) - B(t))dt + \frac{1}{2}\int_0^T B(t)^2 dB_t. \\ &= \frac{1}{2}\int_0^T B(t)dt - \frac{1}{2}TB(T) + \frac{1}{2}\int_0^T B(t)^2 dB_t \qquad (5.3.11) \end{aligned}$$

By the classical Itô formula, we have that

$$\frac{1}{6}B(T)^3 = \frac{1}{2}\int_0^T B_t^2 dB_t + \frac{1}{2}\int_0^t B_t dB_t \qquad (5.3.12)$$

Substituting (5.3.12) into (5.3.11), we get that

$$\int_0^T B(t)(B(T) - B(t))dB_t = \frac{1}{6}\left[B(T)^3 - 3TB(T)\right],$$

which coincides with the result obtained in Example 5.1.8.

Example 5.3.7. We calculate $\int_0^T B(T)B(t)dB(t)$ using Equation (5.3.10). We need to consider a function $F(x, y)$ such that $F_x = xy$. Thus, $F(x, y) = x^2y/2$.

$$d\left(\frac{1}{2}B_t^2 B_T\right) = B_t B_T dB_t + \left(\frac{1}{2}B_T + B_t\right)dt$$

Hence,

$$\int_0^T B(t)B(T)dB(t) = \frac{1}{2}B(T)^3 - \frac{1}{2}TB_T - \int_0^T B(t)dt.$$

5.4 An Example of a Linear Stochastic Differential Equation

In this section, we study a Black-Scholes stochastic differential equation with an anticipating initial condition. We will solve it using the calculus developed so far in this chapter. Consider the following stochastic differential equation,

$$\begin{cases} dS_t = \mu S_t dt + \sigma S_t dB_t, & t \in [0,T], \\ S_0 = B_T. \end{cases} \tag{5.4.1}$$

In the Black-Scholes model, S_t is the evolution of the price of a stock, μ is the appreciation rate and σ is the volatility of the market. Our initial guess for a solution is to consider the solution that Itô calculus provides when the initial condition is adapted. So, we guess with

$$S(t) = B(T) e^{\left(\mu - \frac{\sigma^2}{2}\right)t} e^{\sigma B_t}. \tag{5.4.2}$$

Now, we check whether (5.4.2) is a solution or not of (5.4.1). By Equation (5.3.10), we have that

$$\begin{aligned} dS_t &= \left(\mu - \frac{\sigma^2}{2}\right) S_t dt + \sigma S_t dB_t + \frac{1}{2}\sigma^2 S_t dt + \sigma e^{\left(\mu - \frac{\sigma^2}{2}\right)t} e^{\sigma B_t} dt \\ &= \mu S_t dt + \sigma S_t dB_t + \sigma e^{\left(\mu - \frac{\sigma^2}{2}\right)t} e^{\sigma B_t} dt. \end{aligned} \tag{5.4.3}$$

Since we have obtained an extra third term in (5.4.3), it is clear that (5.4.2) is not a solution of the stochastic differential equation (5.4.1). However, we can try with a slight modification of (5.4.2) so as to cancel out the extra term we have obtained in (5.4.3). Since we have guessed with a solution of the Black-Scholes equation in the Itô theory, the failure must be related to the anticipating initial condition. Furthermore, note that the extra term is with a dt and, therefore, a deterministic term must be added to the factor $B(T)$.
As a result, we consider a function

$$F(t,x,y) = (y - h(t)) e^{\left(\mu - \frac{\sigma^2}{2}\right)t} e^{\sigma x},$$

where $h(t)$ is a deterministic function that we will determine by imposing the condition that S_t is a solution of (5.4.1). We calculate the derivatives of F involved in the new Itô formula,

$$\begin{aligned} F_t &= h'(t) e^{\left(\mu - \frac{\sigma^2}{2}\right)t} e^{\sigma x} + \left(\mu - \frac{\sigma^2}{2}\right) F, \\ F_x &= \sigma F, \\ F_{xx} &= \sigma^2 F, \\ F_{xy} &= \sigma e^{\left(\mu - \frac{\sigma^2}{2}\right)t} e^{\sigma x}. \end{aligned}$$

We consider

$$S_t = (B(T) - h(t))\, e^{\left(\mu - \frac{\sigma^2}{2}\right)t} e^{\sigma B_t}. \tag{5.4.4}$$

Hence, by Equation (5.3.10),

$$dS_t = \left[-h'(t) e^{\left(\mu - \frac{\sigma^2}{2}\right)t} e^{\sigma B_t} + \left(\mu - \frac{\sigma^2}{2}\right) S_t \right] dt + \sigma S_t dB_t + \frac{1}{2}\sigma^2 S_t dt + \sigma e^{\left(\mu - \frac{\sigma^2}{2}\right)t} e^{\sigma B_t} dt$$

$$= -h'(t)e^{\left(\mu-\frac{\sigma^2}{2}\right)t}e^{\sigma B_t}dt + \mu S_t dt + \sigma S_t dB_t + \sigma e^{\left(\mu-\frac{\sigma^2}{2}\right)t}e^{\sigma B_t}dt.$$

If we want S_t in (5.4.4) to be a solution of (5.4.1), $h(t)$ must fulfill the following conditions

$$\begin{cases} h'(t) = \sigma, \ t \in [0,T], \\ h(0) = 0. \end{cases}$$

Clearly, $h(t) = \sigma t$, $t \in [0,T]$. Substituting into (5.4.4), we obtain that a solution of the stochastic differential equation in (5.4.1) is

$$S_t = (B(T) - \sigma t)\, e^{\left(\mu-\frac{\sigma^2}{2}\right)t}e^{\sigma B_t}. \tag{5.4.5}$$

We conclude that the solution is similar to the one in Itô calculus because both solutions concern the same exponents. Unlike non-anticipating calculus, there is no existence and uniqueness theorem and, for this reason, we proved existence by a guess based on Itô calculus. We have left to check that S_t in Equation (5.4.5) is the unique solution of the stochastic differential equation in (5.4.1). Assume that X_t is another solution of (5.4.1). We have that

$$S_t = B(T) + \mu \int_0^t S_u du + \sigma \int_0^t S_u dB_u,$$
$$X_t = B(T) + \mu \int_0^t X_u du + \sigma \int_0^t X_u dB_u.$$

If we define $Z_t = S_t - X_t$, $t \in [0,T]$, we have that Z_t is a stochastic process satisfying that

$$\begin{cases} Z_t = \mu \int_0^t Z_u du + \sigma \int_0^t Z_u dB_u, \ t \in [0,T], \\ Z_0 = 0. \end{cases}$$

This linear stochastic differential equation is a non-anticipating Black-Scholes type of equation whose unique solution, see Proposition 3.5.1, is given by

$$Z_t = Z_0 e^{\left(\mu-\frac{\sigma^2}{2}\right)t}e^{\sigma B_t}.$$

Uniqueness is guaranteed by Theorem 3.3.4. Since $Z_0 = 0$, we conclude that

$$\mathbb{P}(Z_t = 0, \quad \text{for all } t \in [0,T]) = 1,$$

which entails that

$$\mathbb{P}(S_t = X_t, \quad \text{for all } t \in [0,T]) = 1.$$

Therefore, we conclude that S_t in (5.4.5) is the unique solution of the stochastic differential equation in (5.4.1).

5.5 Open Problems

So far, we have studied the Ayed-Kuo integral as an extension of the Itô integral and we have shown that it has many properties which are analogies of some properties of the Itô integral. Furthermore, we have noticed that when $\varphi(t) = 1$, we recover all the notions and properties belonging to Itô theory. However, there are still some aspects in the Itô theory which have not been studied yet for the Ayed-Kuo integral. In this final section, we aim to expose them.

1) The definition of the Ayed-Kuo requires that the integrand has to be a product of an adapted process and an instantly independent one, and that the limit exists in probability. By contrast, the Itô integral is well-defined for stochastic processes belonging to $\mathcal{L}_{ad}(\Omega, L^2[0,T])$. Hence,
For which class of stochastic processes, the Ayed-Kuo is defined?
Some partial work has been done in [10] so as to obtain more general integrands. The idea is to consider a stochastic process of the form

$$X(t,\omega) = \sum_{i=1}^{n} f_i(t,\omega)\varphi_i(t,\omega), \tag{5.5.1}$$

where f_i, $1 \leq i \leq n$, are $\{\mathcal{F}_t\}$-adapted processes, while φ_i, $1 \leq i \leq n$, are instantly independent processes. Then,

$$\int_0^t X(s)dB(s) := \sum_{i=1}^{n} \int_0^t f_i(s,\omega)\varphi_i(s,\omega)dB(s) \tag{5.5.2}$$

provided that all the Ayed-Kuo integrals exist. In [10, Lemma 2.1], it is proved that the definition does not depend on the representation of X.

Definition 5.5.1. Let $\{X(t), 0 \leq t \leq T\}$ be a stochastic process. Assume that there exists a sequence $\{X^{(n)}\}_{n=1}^{\infty}$ of stochastic processes of the form in Equation 5.5.1 such that

(i) The limit $\lim_{n\to\infty} \int_0^T \left|X(t) - X^{(n)}(t)\right|^2 dt = 0$ holds almost surely.

(ii) The limit $\lim_{n\to\infty} \int_0^T X^{(n)}(t)dB(t)$ exists in probability.

Then, the *Ayed-Kuo integral* of $X(t)$ is defined by

$$\int_0^T X(t)dB(t) := \lim_{n\to\infty} \int_0^T X^{(n)}(t)dB(t)$$

in probability.

Note that Definition 5.5.1 enlarges the class of integrands for the Ayed-Kuo, but it does not provide a full characterization of the Ayed-Kuo integrable stochastic processes.

(2) In Chapter 4, we constructed the Skorohod integral and we proved that it is an extension of the Itô integral. In the case of the Ayed-Kuo integral, it was immediate to see that it extends the Itô integral. Then, it naturally arises the following question:
Which is the relation between the Ayed-Kuo and the Skorohod integrals?
In Example 5.1.7 and 5.1.8, we have shown that for those particular cases, both integrals coincide.
In [17], H.-H. Kuo conjectures that when the Skorohod and the Ayed-Kuo integrals exist, they coincide.

(3) In Chapter 2, we have shown that the Itô integral as a stochastic process has continuous sample paths with probability one. Moreover, it has the same path regularity as Brownian motion by virtue of Theorem 3.4.5. Consider the following stochastic process

$$X_t = \int_0^t f(s)\varphi(s)dB(s), \ t \in [0, T],$$

where f is an $\{\mathcal{F}_t\}$-adapted stochastic process and φ is an instantly independent stochastic process. Then, we may ask the following question:
Is there a version of X_t with continuous sample paths?
To answer this question, take into account that in Example 5.2.2, we have shown that the Ayed-Kuo integral does not have the martingale property. However, in Definition 5.2.3 we have introduced the notion of near-martingale. Likewise, we could introduce the concept of near-submartingale and near-supermatingale by changing the equal sign by "$\geq$" and "$\leq$", respectively.

For proving the continuity of the sample paths of the Itô integral as a stochastic process, we used the Doob submartingale inequality (Theorem B.2). So, a new inequality for near-submartingales might be needed.

(4) We have informally derived an Itô formula in order to deal with the kind of processes studied in this dissertation. In [10], a general Itô formula is given for the more general definition of the Ayed-Kuo integral given in Definition 5.5.1. The structure of the formula given in [10] is the same as Equation 5.3.8.
We think that more general versions will be need as long as more general definitions of the Ayed-Kuo integral appear.

(5) In Chapter 3, we studied existence and uniqueness of solutions of stochastic differential equations. In that chapter, we used techniques from the Itô theory because we assumed that the integrands and the initial conditions were non-anticipating.

The aim of the Ayed-Kuo integral is to solve stochastic differential equations with both anticipating integrands and initial conditions. We studied the existence and uniqueness in an example of anticipating stochastic differential equation. Then, it arises the following question:
Under which conditions can be derived a theorem on the existence and uniqueness of a solution of a general stochastic differential equation with the Ayed-Kuo integral?

There are results involving a specific type of linear stochastic differential equations in [10] and [28].

(6) In Section 4.1, we described a simplified version of the insider problem. The stochastic differential equations involved were

$$\begin{cases} dS_0(t) = rS_0(t)dt, \\ S_0(0) = M\mathbb{1}_{\{\bar{S}_1(T)\leq \bar{S}_0(T)\}}, \end{cases} \tag{5.5.3}$$

$$\begin{cases} dS_1(t) = \mu S_1(t)dt + \sigma S_1(t)dB(t), \\ S_1(0) = M\mathbb{1}_{\{\bar{S}_1(T)> \bar{S}_0(T)\}}. \end{cases} \tag{5.5.4}$$

In [10], there is a general solution for this type of equations when the initial condition is a continuous function of Brownian motion. Therefore, we think that through an approximation argument, we might obtain a solution for the insider trading problem with the Ayed-Kuo theory. Some partial work indicates that the results coincide with the Skorohod approach in [7]. However, we still need to rigorously solve the stochastic differential equation. This coincidence would strengthen the conjecture in point (2).

(7) In the Itô theory, the solutions of stochastic differential equations have the Markov property. We wonder which would be the analogue of the Markov property in this new stochastic integration theory. Furthermore, as we have already mentioned, the Itô solutions of stochastic differential equations are diffusions. Then, we also wonder which would be the analogue of diffusions in this context.

Conclusions

The development of this project has provided the necessary knowledge to understand some of the new research ideas in stochastic integration. We have shown that the Itô integral has many advantageous probabilistic properties which justify its success in Mathematics and other fields. As a result, it is a starting point for the study of new integrals with the aim of dealing with anticipating calculus and finding similar properties.

The restriction of only considering Brownian motion has not supposed any limitation for the study of the three integrals because we have been able to fully construct them and suggest applications. We have also revised some rather classical results such as the Kolmogorov continuity theorem or the quadratic variation of Brownian motion. Moreover, many features of the Itô integral have been examined such as its construction, the martingale property, the pathwise continuity and the Itô isometry. As a main application of it, we have considered stochastic differential equations, that is, the classical theorem on existence and uniqueness, and the path regularity of solutions. Regarding to financial applications, the insider trading problem has been a motivation for introducing anticipating calculus. In Chapter 4, the Skorohod integral has been presented as an extension of Itô integration and has entailed the introduction of fundamental theorems such as the Wiener-Itô chaos expansion and the Itô representation theorem. In the last chapter, the Ayed-Kuo integral has been taken into account along with properties such as the zero mean property, the near-martingale property, an isometry and a new change of variables formula. Moreover, we outline that this integration theory coincides with Itô theory in a non-anticipating setting.

Another achievement has been to develop the ability of consulting many references, being critic and extracting own conclusions. Furthermore, the fact of dealing with theories which are still open has contributed to learn how to look for recent papers and also judge future lines of research. One of the limitations of this dissertation has been the impossibility of including Malliavin calculus so as to have a deeper understanding of the Skorohod integral. By contrast, we have outlined that the Ayed-Kuo integral can be deeply examined with few mathematical prerequisites.

Further work could focus on the Russo-Vallois integral, another anticipating integral from 1990s, and comparing it with the studied ones. A study of integrals for fractional Brownian motion and Lévy processes could also be taken into consideration. Another line of research could consist in examining the open problems described in Section 5.5, which are related to the Ayed-Kuo integral, and solving insider trading problems with it.

Appendix A

Conditional Expectation

The concept of conditional expectation is central to modern Probability and the theory of stochastic processes. Our idea is to define and examine its properties without proofs. For details, we suggest reading [$6, p.$ $445 - 455$] or [11].

Theorem A.1. *Let X be a random variable defined on $(\Omega, \mathcal{F}, \mathbb{P})$ such that $\mathbb{E}\,|X| < \infty$ and $\mathcal{G} \subset \mathcal{F}$ is a σ-field. Then, there exists a random variable $Z \in L^1(\Omega)$ such that*

(i) Z is $\mathcal{G}$-measurable.

(ii) For all $A \in \mathcal{G}$,

$$\int_A Z(\omega) d\mathbb{P}(\omega) = \int_A X(\omega) d\mathbb{P}(\omega).$$

Moreover, Z is unique up to a set of probability zero.

The proof of this theorem is based on the Radon-Nykodym theorem.

Definition A.2. Let $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ be a random variable and $\mathcal{G}$ a sub-σ-field. *The conditional expected value* of X given $\mathcal{G}$, $\mathbb{E}\left(X|\,\mathcal{G}\right)$, is a random variable satisfying that

(i) $\mathbb{E}\left(X|\,\mathcal{G}\right)$ is $\mathcal{G}$-measurable.

(ii) For all $A \in \mathcal{G}$,

$$\int_A \mathbb{E}\left(X|\,\mathcal{G}\right) d\mathbb{P} = \int_A X d\mathbb{P}.$$

By Theorem A.1, the conditional expectation is well-defined and it is unique up to a set of probability zero. The next theorem provides some important properties of conditional expectation.

Theorem A.3. *Let X and Y be two integrable random variables and $\mathcal{G}$ a sub-σ-field. The following properties hold:*

(i) If X is $\mathcal{G}$-measurable, then $\mathbb{E}\left(X|\,\mathcal{G}\right) = X$.

(ii) $\mathbb{E}\left[\mathbb{E}\left(X|\,\mathcal{G}\right)\right] = \mathbb{E}(X)$.

(iii) *For any* $a, b \in \mathbb{R}$, $\mathbb{E}\left(aX + bY \mid \mathcal{G}\right) = a\mathbb{E}\left(X \mid \mathcal{G}\right) + b\mathbb{E}\left(Y \mid \mathcal{G}\right)$.

(iv) *If* $X(\omega) \leq Y(\omega)$ *for almost all* $\omega \in \Omega$, *then* $\mathbb{E}\left(X \mid \mathcal{G}\right) \leq \mathbb{E}\left(Y \mid \mathcal{G}\right)$ *almost surely.*

(v) *If* X *and* $\mathcal{G}$ *are independent, then* $\mathbb{E}\left(X \mid \mathcal{G}\right) = \mathbb{E}(X)$.

(vi) *If* Y *is* $\mathcal{G}$*-measurable and* $\mathbb{E}\left|XY\right| < \infty$, *then* $\mathbb{E}\left(XY \mid \mathcal{G}\right) = Y\mathbb{E}\left(X \mid \mathcal{G}\right)$.

(vii) *If* $\mathcal{H} \subset \mathcal{G}$ *is a* σ*-field, then*

$$\mathbb{E}\left[\mathbb{E}\left(X \mid \mathcal{G}\right) \mid \mathcal{H}\right] = \mathbb{E}\left(X \mid \mathcal{H}\right) = \mathbb{E}\left[\mathbb{E}\left(X \mid \mathcal{H}\right) \mid \mathcal{G}\right].$$

Next, we provide the versions of the monotone convergence theorem and the dominated convergence theorem.

Theorem A.4 (Monotone convergence theorem). *If* $\{X_n\}_{n\in\mathbb{N}}$ *is a sequence of* $L^1(\Omega)$*-random variables such that* $X_n \geq 0$ *for all* $n \in \mathbb{N}$ *and converges to a random variable* X *with probability one, then*

$$\mathbb{E}\left(X \mid \mathcal{G}\right) = \lim_{n\to\infty} \mathbb{E}\left(X_n \mid \mathcal{G}\right)$$

with probability one.

Theorem A.5 (Dominated convergence theorem). *Let* $\{X_n\}_{n\in\mathbb{N}}$ *be a sequence of* $L^1(\Omega)$*-random variables which converges to some* $X \in L^1(\Omega)$ *almost surely. If there exists a random variable* $Y \in L^1(\Omega)$ *such that* $|X_n| \leq Y$ *for all* $n \in \mathbb{N}$, *then*

$$\mathbb{E}\left(X \mid \mathcal{G}\right) = \lim_{n\to\infty} \mathbb{E}\left(X_n \mid \mathcal{G}\right)$$

with probability one.

We also give the Jensen inequality for conditional expectations.

Theorem A.6 (Jensen inequality). *Let* $f : \mathbb{R} \to \mathbb{R}$ *be a convex function and* $X \in L^1(\Omega)$ *a random variable. If* $\mathbb{E}\left|f(X)\right| < \infty$, *then*

$$f\left(\mathbb{E}\left(X \mid \mathcal{G}\right)\right) \leq \mathbb{E}\left(f\left(X\right) \mid \mathcal{G}\right)$$

with probability one.

Next, we give a convergence theorem for increasing families of σ-fields. A proof can be found in [18, p.303].

Theorem A.7 (Conditional monotone convergence theorem). *Let* $X \in L^1(\Omega)$ *be a random variable and* $\{\mathcal{G}_k\}_{k=1}^{\infty}$ *a family of* σ*-fields such that*

(i) $\mathcal{G}_k \subset \mathcal{G}_{k+1}$ *for all* $k \geq 1$.

(ii) $\mathcal{G}_k \subset \mathcal{F}$ *for all* $k \geq 1$.

If $\mathcal{F}_\infty$ *is the smallest* σ*-field which contains all* $\mathcal{F}_k$, $k \geq 1$, *then*

$$\mathbb{E}\left(X \mid \mathcal{F}_\infty\right) = \lim_{k\to\infty} \mathbb{E}\left(X \mid \mathcal{F}_k\right)$$

with probability one.

Appendix B

Doob submartingale Inequalities

In many proofs, we need to estimate the second moment of a martingale, which means dealing with the expectation of submartingales. A useful bound is provided by the Doob submartingale inequality.

Lemma B.1 (Discrete Doob submartingale inequality). *If $\{X_n\}_{n=1}^{\infty}$ is a submartingale, then*

$$\mathbb{P}\left(\max_{1\le k\le n} X_k \ge \lambda\right) \le \frac{1}{\lambda}\mathbb{E}\left[\max(X_n, 0)\right]$$

for all integers $n \ge 1$ and all $\lambda > 0$.

Proof. We fix $\lambda > 0$. For every $k = 1, \ldots, n$, we define the sets

$$A_k = \bigcap_{j=1}^{k-1} \{\omega : X_j(\omega) \le \lambda\} \cap \{\omega : X_k(\omega) > \lambda\}.$$

We consider the set

$$A = \left\{\omega : \max_{1\le k\le n} X_k(\omega) > \lambda\right\} = \biguplus_{k=1}^{n} A_k.$$

Since

$$\lambda\mathbb{P}(A) = \lambda\sum_{k=1}^{n}\mathbb{P}(A_k) = \sum_{k=1}^{n}\int_{A_k}\lambda\, d\mathbb{P} \le \sum_{k=1}^{n}\int_{\Omega} X_k \mathbb{1}_{A_k} d\mathbb{P} = \sum_{k=1}^{n}\mathbb{E}\left[X_k \mathbb{1}_{A_k}\right],$$

it follows

$$\begin{aligned}\mathbb{E}[\max(X_n, 0)] \ge \mathbb{E}[\max(X_n, 0)\mathbb{1}_A] &= \sum_{k=1}^{n}\mathbb{E}[\max(X_n, 0)\mathbb{1}_{A_k}] \\ &= \sum_{k=1}^{n}\mathbb{E}\left[\mathbb{E}[\max(X_n, 0)\mathbb{1}_{A_k} | \, X_1, \ldots, X_k]\right] \\ &= \sum_{k=1}^{n}\mathbb{E}\left[\mathbb{1}_{A_k}\mathbb{E}[\max(X_n, 0) | \, X_1, \ldots, X_k]\right]\end{aligned}$$

$$\geq \sum_{k=1}^{n} \mathbb{E}\left[\mathbb{1}_{A_k}\mathbb{E}[X_n|\, X_1,\ldots,X_k]\right]$$
$$\geq \sum_{k=1}^{n} \mathbb{E}[X_k \mathbb{1}_{A_k}]$$
$$\geq \lambda\mathbb{P}(A).$$

□

Proposition B.2 (Continuous Doob submartingale inequality)**.** *Let* $\{Y_t, 0 \leq t \leq T\}$ *be a submartingale with almost surely continuous sample paths. Then, for any* $\varepsilon > 0$*,*

$$\mathbb{P}\left(\sup_{0\leq t\leq T} Y_t \geq \varepsilon\right) \leq \frac{1}{\varepsilon}\mathbb{E}[\max(Y(T),0)].$$

Proof. For the sake of simplicity, we adopt the notation $Y(T)^+ = \max(Y(T), 0)$. Since rational numbers are countable, we express

$$\mathbb{Q}_T := \mathbb{Q} \cap [0,T] = \bigcup_{n=1}^{\infty}\{q_n\}.$$

We want to prove first that

$$\sup_{q\in\mathbb{Q}_T} Y(q) = \sup_{0\leq t\leq T} Y(t) \quad (a.s.). \tag{B.0.1}$$

Since $\mathbb{Q}_T$ is dense set in $[0,T]$ and Y has continuous paths a.s., it holds that

$$Y([0,T]) \subset \overline{Y(\mathbb{Q}_T)} \quad (a.s.).$$

Therefore,

$$\sup_{0\leq t\leq T} Y(t) \leq \sup \overline{Y(\mathbb{Q}_T)} = \sup Y(\mathbb{Q}_T) = \sup_{q\in\mathbb{Q}_T} Y(q) \quad (a.s.).$$

The converse inequality is clear. For each k, the set $\{q_1,\ldots,q_k\}$ is written in increasing order $\{q_1^{(k)} < q_2^{(k)} < \ldots < q_k^{(k)}\}$. For any $\varepsilon > 0$,

$$\left\{\sup_{q\in\mathbb{Q}_T} Y(q) > \varepsilon\right\} = \bigcap_{n=1}^{\infty}\bigcup_{k=1}^{\infty}\left\{\max_{1\leq l\leq k} Y\left(q_l^{(k)}\right) \geq \varepsilon - \frac{1}{n}\right\}. \tag{B.0.2}$$

By (B.0.1) and (B.0.2), we get that

$$\mathbb{P}\left(\sup_{0\leq t\leq T} Y(t) \geq \varepsilon\right) = \mathbb{P}\left(\sup_{q\in\mathbb{Q}_T} Y(q) \geq \varepsilon\right) = \lim_{n\to\infty}\lim_{k\to\infty}\mathbb{P}\left(\max_{1\leq l\leq k} Y(q_l^{(k)}) \geq \varepsilon - \frac{1}{n}\right). \tag{B.0.3}$$

Note that $Y(q_1^{(k)}), Y(q_2^{(k)}), \ldots, Y(q_k^{(k)})$ is a discrete submartingale. By Lemma $B.1$,

$$\mathbb{P}\left(\max_{1\leq l\leq k} Y(q_l^{(k)}) \leq \varepsilon - \frac{1}{n}\right) \leq \frac{1}{\varepsilon - \frac{1}{n}}\mathbb{E}\left[Y(q_k^{(k)})^+\right]. \tag{B.0.4}$$

Combining (B.0.3) and (B.0.4), we have

$$\mathbb{P}\left(\sup_{0\le t\le T} Y(t) \ge \varepsilon\right) \le \lim_{n\to\infty}\lim_{k\to\infty} \frac{1}{\varepsilon - \frac{1}{n}}\mathbb{E}\left[Y(q_k^{(k)})^+\right]$$
$$\le \lim_{n\to\infty}\lim_{k\to\infty} \frac{1}{\varepsilon - \frac{1}{n}}\mathbb{E}\left[Y(T)^+\right] = \frac{1}{\varepsilon}\mathbb{E}\left[Y(T)^+\right],$$

where we have used that $\mathbb{E}\left[Y(T)^+\right] \ge \mathbb{E}\left[Y(q_k^{(k)})^+\right]$. Indeed,

$$\mathbb{E}\left[Y(T)^+ \middle| \mathcal{F}_{q_k^{(k)}}\right] \ge \mathbb{E}\left[Y(T) \middle| \mathcal{F}_{q_k^{(k)}}\right] \ge Y(q_k^{(k)})$$

and

$$\mathbb{E}\left[Y(T)^+ \middle| \mathcal{F}_{q_k^{(k)}}\right] \ge \mathbb{E}\left[0 \middle| \mathcal{F}_{q_k^{(k)}}\right] = 0.$$

Taking expectations on both sides, we obtain the desired inequality. □

Appendix C

The Itô Representation Theorem

The goal of this section is to prove the Itô representation theorem, which states that any $L^2(\Omega)$-random variable can be expressed as an Itô integral of a L^2_{ad}-stochastic process. This result is extremely useful in the construction of the Skorohod integral. We begin with two technical lemmas before proving this theorem.

Lemma C.1. *Let $T > 0$. The set of random variables*

$$\{\varphi\left(B_{t_1}, \ldots, B_{t_n}\right) : t_i \in [0, T], \varphi \in \mathcal{C}_0^\infty(\mathbb{R}^n), n = 1, 2, \ldots\}$$

is dense in $L^2(\Omega, \mathcal{F}_T, \mathbb{P})$.

Proof. Let $\{t_i\}_{i=1}^\infty$ be a countable dense subset of $[0, T]$. For each $n = 1, 2, \ldots$, we consider

$$H_n = \sigma < B_{t_1}, \ldots, B_{t_n} >$$

the σ-field generated by $B_{t_1}, \ldots, B_{t_n}$. It is clear that

$$H_n \subset H_{n+1}$$

for all n and that $\mathcal{F}_T$ is the smallest σ-field containing all σ-fields H_n. We pick a random variable $X \in L^2(\Omega, \mathcal{F}_T, \mathbb{P})$. By Theorem $A.7$, we have that

$$X = \mathbb{E}[X|\,\mathcal{F}_T] = \lim_{n\to\infty} \mathbb{E}[X|\,H_n] \tag{C.0.1}$$

almost surely with respect to $\mathbb{P}$ and, hence, in $L^2(\Omega, \mathcal{F}_T, \mathbb{P})$. By the Doob-Dynkin Lemma (Proposition 1.1.5), we have that, for each n,

$$\mathbb{E}[X|\,H_n] = f_n(B_{t_1}, \ldots, B_{t_n}) \tag{C.0.2}$$

for some Borel measurable functions $f_n : \mathbb{R}^n \to \mathbb{R}$. Combining (C.0.1) and (C.0.2), we have that

$$X = \lim_{n\to\infty} f_n(B_{t_1}, \ldots, B_{t_n}) \tag{C.0.3}$$

almost surely with respect to $\mathbb{P}$ and in $L^2(\Omega, \mathcal{F}_T, \mathbb{P})$. Each function $f_n(B_{t_1}, \ldots, B_{t_n})$ can be approximated in $L^2(\Omega, \mathcal{F}_T, \mathbb{P})$ by functions $\varphi_n(B_{t_1}, \ldots, B_{t_n})$ with $\varphi_n \in \mathcal{C}_0^\infty(\mathbb{R}^n)$. Hence, the lemma follows. □

Lemma C.2. *The linear span of random variables of the type*

$$\exp\left\{\int_0^T h(t)dB_t - \frac{1}{2}\int_0^T h(t)^2 dt\right\}, \quad h \in L^2([0,T]), \tag{C.0.4}$$

is dense in $L^2(\Omega, \mathcal{F}_T, \mathbb{P})$.

Proof. Consider a random variable $X \in L^2(\Omega, \mathcal{F}_T, \mathbb{P})$ such that is orthogonal (in $L^2(\Omega)$ sense) to all random variables like in (C.0.4). In particular,

$$G(\lambda) := \int_\Omega \exp\left\{\lambda_1 B_{t_1}(\omega) + \ldots + \lambda_n B_{t_n}(\omega)\right\} X(\omega) d\mathbb{P}(\omega) = 0$$

for all $\lambda = (\lambda_1, \ldots, \lambda_n) \in \mathbb{R}^n$ and all $t_1, \ldots, t_n \in [0,T]$. Note that G is clearly analytic in $\mathbb{R}^n$ and, therefore, it has an analytic extension to $\mathbb{C}^n$ given by

$$G(z) := \int_\Omega \exp\left\{z_1 B_{t_1}(\omega) + \ldots + z_n B_{t_n}(\omega)\right\} X(\omega) d\mathbb{P}(\omega)$$

for all $z = (z_1, \ldots, z_n) \in \mathbb{C}^n$. As $G = 0$ on $\mathbb{R}^n$ and G is analytic, we conclude that $G = 0$ on $\mathbb{C}^n$. Then, in particular,

$$G(2\pi i y) := \int_\Omega \exp\left\{2\pi i \left[y_1 B_{t_1}(\omega) + \ldots + y_n B_{t_n}\right]\right\} X(\omega) d\mathbb{P}(\omega) = 0$$

for all $y = (y_1, \ldots, y_n) \in \mathbb{R}^n$. We pick a function $\varphi \in \mathcal{C}_0^\infty(\mathbb{R}^n) \subset \mathcal{S}(\mathbb{R}^n)$. By Fourier inversion formula and Fubini theorem,

$$\begin{aligned}
&\int_\Omega \varphi(B_{t_1}, \ldots, B_{t_n}) X(\omega) d\mathbb{P}(\omega) &\text{(C.0.5)}\\
&= \int_\Omega \left(\int_{R^n} \hat{\varphi}(B_{t_1}, \ldots, B_{t_n}) e^{2\pi i\left(y_1 B_{t_1} + \ldots + y_n B_{t_n}\right)} dy\right) X(\omega) d\mathbb{P}(\omega)\\
&= \int_{R^n} \left(\int_\Omega \hat{\varphi}(B_{t_1}, \ldots, B_{t_n}) e^{2\pi i\left(y_1 B_{t_1} + \ldots + y_n B_{t_n}\right)} X(\omega) d\mathbb{P}(\omega) dy\right)\\
&= 0. &\text{(C.0.6)}
\end{aligned}$$

By (C.0.5) and Lemma $C.1$, X is orthogonal to a dense subset of $L^2(\Omega, \mathcal{F}_T, \mathbb{P})$. Hence, $X = 0$ a.s. and we conclude that the linear span of random variables of the form in (C.0.4) are dense in $L^2(\Omega, \mathcal{F}_T, \mathbb{P})$. □

Theorem C.3 (Itô representation theorem)**.** *Let* $F \in L^2(\Omega, \mathcal{F}_T, \mathbb{P})$ *be a random variable. Then, there exists a unique stochastic process* $h(t,\omega) \in L^2_{ad}([0,T] \times \Omega)$ *such that*

$$F(\omega) = \mathbb{E}(F) + \int_0^T h(t,\omega) dB_t.$$

Proof. We first assume that F is of the form

$$F = e^{\int_0^T f(t)dB_t - \frac{1}{2}\int_0^T f(t)^2 dt}$$

where $f \in L^2([0,T])$ is a deterministic function. Define

$$Y_t(\omega) = \int_0^t f(s)dB_s - \frac{1}{2}\int_0^t f(s)^2 ds, \qquad 0 \le t \le T.$$

We also define

$$Z_t = e^{Y_t}, \qquad 0 \le t \le T,$$

and note that $Z_T = F$. By the Itô formula,

$$\begin{aligned} dZ_t &= e^{Y_t} dY_t + \frac{1}{2} e^{Y_t} f(t)^2 dt \\ &= Z_t \left(f(t)dB_t - \frac{1}{2} f(t)^2 dt \right) + \frac{1}{2} Z_t f(t)^2 dt \\ &= f(t) Z_t dB_t. \end{aligned}$$

As $Y_0 = 0$, we have that $Z_0 = 1$. Therefore,

$$Z_t = 1 + \int_0^t Z_s f(s) dB_s.$$

For $t = T$, we have that

$$F = 1 + \int_0^T Z_s f(s) dB_s.$$

Since $\mathbb{E}(F) = 1$,

$$F = \mathbb{E}(F) + \int_0^T Z_s f(s) dB_s.$$

Now, we assume that $F \in L^2(\Omega, \mathcal{F}_T, \mathbb{P})$. By Lemma $C.2$, there exists a sequence $\{F_n\}_{n=1}^\infty$ in $L^2(\Omega, \mathcal{F}_T, \mathbb{P})$, where F_n are linear combinations of random variables of the form in (C.0.4), such that

$$F(\omega) = L^2(\Omega) - \lim F_n(\omega).$$

For each integer $n \ge 1$, there exists a stochastic process $h_n(t,\omega) \in L^2_{ad}$ such that

$$F_n(\omega) = \mathbb{E}(F_n) + \int_0^T h_n(t,\omega) dB_t.$$

By the Itô isometry, note that

$$\begin{aligned} \mathbb{E}\left[(F_n - F_m)^2\right] &= \mathbb{E}\left[\left(\mathbb{E}(F_n - F_m) + \int_0^T (h_n - h_m) dB_t\right)^2\right] \\ &= [\mathbb{E}(F_n - F_m)]^2 + \int_0^T \mathbb{E}(h_n - h_m)^2 ds. \end{aligned}$$

Then,

$$\lim_{n,m\to\infty} \int_0^T \mathbb{E}(h_n - h_m)^2 ds = \lim_{n,m\to\infty} \mathbb{E}\left[(F_n - F_m)^2\right] - \lim_{n,m\to\infty} [\mathbb{E}(F_n - F_m)]^2 = 0.$$

This means that $\{h_n\}$ is a Cauchy sequence in $L^2_{ad}([0,T]\times\Omega)$. Hence, the sequence converges to some $h \in L^2_{ad}([0,T]\times\Omega)$. Moreover, a subsequence $\{h_{n_k}(t,\omega)\}_k$ converges to $h(t,\omega)$ for almost every $(t,\omega) \in [0,T]\times\Omega$. Therefore, $h(t,\cdot)$ is $\mathcal{F}_t$-measurable for almost every $t \in [0,T]$. By modifying $h(t,\omega)$ in a set of Lebesque measure zero, we obtain that $h(t,\omega)$ is $\{\mathcal{F}_t\}$-adapted and, hence, $h(t,\omega) \in L^2_{ad}([0,T]\times\Omega)$.
We conclude, using again the Itô isometry, that

$$F = L^2(\Omega) - \lim_{n\to\infty} F_n = L^2(\Omega) - \lim_{n\to\infty}\left[\mathbb{E}(F_n) + \int_0^T h_n(s)dB_s\right] = \mathbb{E}[F] + \int_0^T h(s)dB_s.$$

So far, we have proved the representation part. We have to show the uniqueness part of the statement. Consider $h_1(t,\omega), h_2(t,\omega) \in L^2_{ad}$ such that

$$F = \mathbb{E}[F] + \int_0^T h_1(t,\omega)dB_t = \mathbb{E}[F] + \int_0^T h_2(t,\omega)dB_t.$$

Subtracting, we get that

$$0 = \int_0^T \left(h_1(t,\omega) - h_2(t,\omega)\right) dB_t.$$

By the Itô isometry,

$$0 = \mathbb{E}\left[\left(\int_0^T (h_1(t,\omega) - h_2(t,\omega))dB_t\right)^2\right] = \int_0^T \mathbb{E}\left[(h_1(t,\omega) - h_2(t,\omega))^2\right] dt.$$

Hence, $h_1(t,\omega) = h_2(t,\omega)$ for almost all $(t,\omega) \in [0,T]\times\Omega$. $\square$

References

[1] Abramowitz, M. and Stegun, I.A.: *Handbook of Mathematical Functions*, Dover, New York, 1972.

[2] Arnold, L.: *Stochastic Differential Equations. Theory and Applications*, John Wiley & Sons, New York, 1974.

[3] Ayed, W. and Kuo, H.-H.: An extension of the Itô integral, *Communications on Stochastic Analysis*, Vol. 2, No. 3 (2008) 323-333.

[4] Ayed, W. and Kuo, H.-H.: An extension of the Itô integral: toward a general theory of stochastic integration, *Theory of Stochastic Processes*, Vol. 16, No. 1 (2010) 17-28.

[5] Bakstein, D. and Capasso, V.: *An Introduction to Continuous-Time Stochastic Processes. Theory, Models and Applications to Finance, Biology, and Medicine*, 1st edition, Birkhäuser, Boston, 2005.

[6] Billingsley, P.: *Probability and Measure*, 3rd edition, John Wiley & Sons, New York, 1995.

[7] Escudero, C.: A simple comparison between Skorokhod & Russo-Vallois integration for insider trading, *preprint*.
https://arxiv.org/pdf/1704.05014.pdf

[8] Evans, L.C.: *An Introduction to Stochastic Differential Equations*, 1st edition, AMS, Rhode Island, 2013.

[9] Hull, J.: *Options, Futures, and other derivative securities*, 2n edition, Prentice Hall International, New Jersey, 1993.

[10] Hwang, C.-H., Kuo, H.-H., Saitô, K. and Zhai, J.: A general Itô formula for adapted and instantly independent stochastic processes, *Communications on Stochastic Analysis*, Vol. 10, No. 3 (2016) 341-362.

[11] Ibarrola, P., Pardo, L. and Quesada, V.: *Teoría de la probabilidad*, Editorial Síntesis, Madrid, 1999.

[12] Karatzas, I. and Shreve, S.: *Brownian Motion and Stochastic Calculus*, Graduate Texts in Mathematics, series volume 113, 2nd edition, Springer-Verlag, New York, 1998.

[13] Klebaner, F.: *Introduction to Stochastic Calculus with Applications*, 2nd edition, Imperial College Press, London, 2005.

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

[15] Kuo, H.-H., Sae-Tang, A. and Szozda, S.: The Itô formula for a new stochastic integral, *Communications on Stochastic Analysis*, Vol. 6, No. 4 (2012) 603-614.

[16] Kuo, H.-H., Peng, Y. and Szozda, S.: Itô formula and Giranov theorem for anticipating stochastic integrals, *Communications on Stochastic Analysis*, Vol. 7, No. 3 (2013) 441-458.

[17] Kuo, H.-H.: The Itô calculus and white noise theory: A brief survey toward general stochastic integration, *Communications on Stochastic Analysis*, Vol. 8, No. 1 (2014) 111-139.

[18] Øksendal, B.: *Stochastic differential equations*, 6th edition, Springer-Verlag, Berlin, 2005.

[19] Nualart, D.: *Malliavin calculus and related topics*, 2nd edition, Springer-Verlag, Berlin, 2006.

[20] Nualart, D. and Pardoux, E.: Stochastic calculus with anticipating integrands, *Probability Theory and Related Fields* Vol. 78 (1988) 535–581.

[21] Nuno, G., Øksendal, B. and Proske, F.: *Malliavin Calculus for Lévy Processes with Applications to Finance*, Springer-Verlag, Berlin, 2009.

[22] Partzch, L. and Schilling, R.L.: *Brownian Motion: An Introduction to Stochastic Processes*, 1st edition, De Gruyter, Berlin, 2012.

[23] Sae-Tang, A.: *A new theory of stochastic integration*, Ph.D. thesis, Louisiana State University, 2011.

[24] Sanz-Solé, M.: *Probabilitats*, Edicions de la Universitat de Barcelona, 1999.

[25] Sanz-Solé, M.: *An Introduction to Stochastic Calculus*, manuals, 2015. `http://www.ub.edu/plie/Sanz-Sole/`

[26] Skorokhod, A.V.: On a generalization of a stochastic integral, *Theory of Probability and its Applications*, Vol. 20 (1975) 219–233.

[27] Steele, J.: *Stochastic Calculus and Financial Applications*, Springer-Verlag, New York, 2001.

[28] Szozda, B.: *The new stochastic integral and anticipating stochastic differential equations*, Ph.D. thesis, Louisiana State University, 2012.

Index