4-9-2022

General Stochastic Calculus and Applications

Pujan Shrestha
Louisiana State University, Baton Rouge, LA 70803, USA

GENERAL STOCHASTIC CALCULUS AND APPLICATIONS

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The Department of Mathematics

by
Pujan Shrestha
B.S., Randolph College, 2015
M.S., Louisiana State University, 2018
May 2022

This dissertation is dedicated to my parents, Dev Bahadur and Rohini Shrestha, my brother Roshan, my sister Deepa, my sister-in-law Sonam and my lovely nephews Rivaan and Samik. I would not be here without your everlasting love and support.

Acknowledgments

I want to thank my advisors and my academic fathers' Professor Kuo and Professor Sundar for my growth as a mathematician and as a human being. I am eternally grateful for their guiding advice through this graduate experience.

I would also like to thank the other members of my dissertation committee — Professor Ganguly and Professor Launey for their time and patience. I appreciate them for being part of both my general exam and my dissertation defense committee.

I would like thank Professor Ordower for lighting the spark of mathematics within me. I was blessed to be in his classes for majority of my undergraduate years at Randolph College. I hope to be as good as a teacher as he is.

I want to thank Sudip Sinha for his helpful advice, intuition, rigor and his friendship. I am glad to have gain a brother in this arduous journey through graduate school.

In addition, I want to thank the Department of Mathematics at Louisiana State University for providing an ideal environment to grow as a mathematician.

Finally, I would like to thank my friends and family in Baton Rouge, in California and in Nepal for their continued love and support. I would like to thank Hannah Lichtl for her insightful advice, moral support and her wonderful dog, Luna. I would be a lonely island without you all.

Table of Contents

List of Illustrations

Tables

Figures

Abstract

In 1942, K. Itô published his pioneering paper on stochastic integration with respect to Brownian motion. This work led to the framework for Itô calculus. Note that, Itô calculus is limited in working with knowledge from the future. There have been many generalizations of the stochastic integral in being able to do so. In 2008, W. Ayed and H.-H. Kuo introduced a new stochastic integral by splitting the integrand into the adaptive part and the counterpart called instantly independent. In this doctoral work, we conduct deeper research into the Ayed–Kuo stochastic integral and corresponding anticipating stochastic calculus.

We provide a new proof for the extension of Itô isometry for the Ayed–Kuo stochastic integral which clearly demonstrates the intrinsic nature of the construction of the general integral. Furthermore, we extend classical Itô theory results for martingales to their Ayed–Kuo stochastic integral analogue, near-martingale. We show the near-martingale property of Ayed–Kuo stochastic integral and optional stopping theorem for near-martingales with bounded stopping times.

Using the general Itô formula for the Ayed–Kuo stochastic integral, we find explicit solutions for linear stochastic differential equations with anticipation. We show existence of solutions for certain classes linear stochastic differential equations with anticipation coming from initial condition as well as from the drift. We present a Trotter inspired product formula to construct the solution. In the process, we also show the uniqueness of the solution. While we mainly rely on the Ayed–Kuo formalism, other theories are used minimally and out of necessity. Using the explicit solution, we show the relation between a solution of an anticipating stochastic differential equations and its Itô projection. Fur-

thermore, we establish Wentzell–Friedlin type large deviation principle for the solution of a class of linear stochastic differential equation with an anticipating drift and non-adapted initial condition.

Introduction

In 1902, H. Lebesgue revolutionized mathematics with his introduction of the Lebesgue integral and subsequently, measure theory [23]. N. Wiener used these measure-theoretic tools in a mathematical model of Brownian motion, $B(t)$ [31]. In 1942, K. Itô introduced a new stochastic integral while studying for a probabilistic method to construct diffusion processes from infinitesimal generators [12] [16]. This Itô integral can be seen as a generalization of Wiener integral. Itô theory has a measurability requirement which leads to an inability to work knowledge from the future. There have been many generalization include future information in the stochastic integral. Wiener's work provided motivation for L. Gross's work into Abstract Wiener space for infinite dimensional analysis [7]. In 1975, Hida build up white noise distribution theory in his analysis of Brownian functionals [8]. In 1976, P. Malliavin proved the existence of transition probabilities via probabilistic methods in [24] (See also [25]) leading to the current theory of Malliavin calculus. In the white noise distribution theory, the white noise integral provides a generalization for Itô integrals.

In 2008, W. Ayed and H.-H. Kuo introduced a new anticipating stochastic integral to provide a natural and simplistic tool to analyze anticipating stochastic processes [1]. Within this framework, this doctoral work is interested in extending Itô's theory of stochastic integrals to a general setting via the Ayed-Kuo stochastic integral.

The following diagram shows the relationships between the four areas of stochastic analysis and where the Ayed–Kuo stochastic integral lies when considering the other areas.

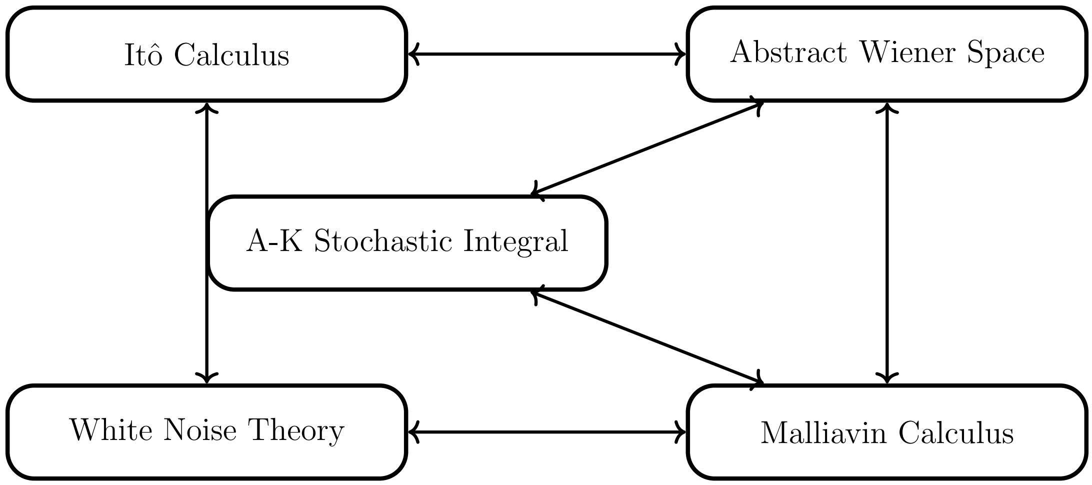

Figure I.1. Relationships between the different areas of stochastic analysis.

This doctoral work is laid out as follows. In chapter 1, we introduce some fundamental ideas in classical Itô theory. We follow that with some introduction and results into the analysis of rare events via large deviation. In chapter 3, we present the main focus of this research. We define the Ayed–Kuo integral and state important theorems and properties. In chapter 4, we apply the new integral in solving a class of linear stochastic differential equation with anticipation from the initial condition as well as from the drift term. Finally, in chapter 5, we combine the tools obtained in the Ayed–Kuo stochastic integral to obtain large deviation results for a particular class of anticipating linear stochastic differential equation. While this doctoral work greatly relies on the formalism of the Ayed–Kuo stochastic integral, other theories are used minimally and out of necessity.

Chapter 1. Classical Itô Theory

1.1. Background

In 1827, English botanist R. Brown observed an "peculiar" phenomena when studying pollen seeds suspended on the water [3]. He observed that the particles were moving in an irregular pattern despite no outside forces. In a series of papers he published regarding the phenomena [3] he writes, "I have formerly stated my belief that these motions of the particles neither arose from currents in the fluid containing them, nor depended on that intestine motion which may be supposed to accompany its evaporation." This irregular movement was named *Brownian motion* after Robert Brown.

It was not until much later in 1905 that renowned physicist Albert Einstein described this irregular motion as the resulting diffusion due to the pollen being battered by water molecules [6]. Einstein's explanation of Brownian motion, via normal distribution, is similar to its modern mathematical definition. Interestingly, Einstein did not know of Brown's empirical observation and deduced the theory from molecular kinetic theory of heat [6]. Alternately, in 1900, L. Bachelier applied Brownian motion to analyse stock price fluctuations in his doctoral work [2]. In 1923, N. Wiener provided the first mathematical model of Brownian motion, $B(t)$ [31]. In it, he defined the basic probabilities as values of a Gaussian measure defined on cylinder sets in the space of continuous functions.

In his pioneering paper "Stochastic Integral", K. Itô introduced a stochastic integral and a formula, known as the Itô's formula [12]. The practical applications of Itô's integral are diverse. One of the most prominent applications is to the Black-Scholes-Merton theory of derivative pricing in finance. Itô calculus has been useful in the study of other

scientific fields like statistical physics, biological systems.

In the following sections, we review the definitions of the preliminary mathematical concepts and their properties. We refer to [15] for details.

1.2. Stochastic Processes and Brownian Motion

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space where Ω is the sample space, $\mathcal{F}$ a sigma-field and $\mathbb{P}$ a probability measure.

Definition 1.2.1. *A stochastic process is a measurable function, $X(t, \omega)$, defined on $[0, \infty) \times \Omega$ with the following properties:*

- *for each t, $X(t, \cdot)$ is a random variable,*
- *for each ω, $X(\cdot, \omega)$ is a measurable function (also called sample paths).*

Definition 1.2.2. *A stochastic process, $B(t, \omega)$, is called a Brownian motion or a Wiener process if it satisfies the following conditions:*

1. $\mathbb{P}\{\, \omega \,|\, B(0, \omega) = 0 \,\} = 1$,

2. *For any $0 \le s \le t$, the random variable $B(t, \omega) - B(s, \omega)$ is normally distributed with mean 0 and variance $t - s$,*

3. *$B(t, \omega)$ has independent increments,*

4. $\mathbb{P}\{\, \omega \,|\, B(\cdot, \omega)$ *is continuous* $\} = 1$.

From the second condition, we can informally think of a Brownian increment as $|B_t - B_s| \approx \sqrt{|t - s|}$ for $s, t \in [a, b]$. Suppose $s < t$. Heuristically, when s approaches t,

$$\frac{|B_t - B_s|}{t - s} \approx \frac{\sqrt{t - s}}{t - s} = \frac{1}{\sqrt{t - s}} \to \infty.$$

This allows us to get an intuition on why Brownian motion paths are nowhere differentiable. While the Brownian motion does not have finite variation, its quadratic variation is finite. This well known result is a fundamental property in Itô's theory of stochastic integration.

Theorem 1.2.3 (Quadratic Variation of Brownian Motion)**.** *Consider a partition of the finite interval $[a, b]$ given by $\Delta_n = \{a = t_0 \leq t_1 \leq t_2 \cdots \leq t_n = b\}$. Then*

$$\sum_{i=1}^{n} \left(B_{t_i} - B_{t_{i-1}}\right)^2 \to b - a \qquad \text{in } L^2(\Omega),$$

as $\|\Delta_n\| = max_{1\leq i\leq n}(t_i - t_{i-1}) \to 0$.

1.3. Conditional Expectation and Martingales

Now we introduce some probabilistic concepts that will be used in much of this text. We first review conditional expectation and some of its properties.

Definition 1.3.1. *Suppose $\mathcal{F}$ is a sigma field such that $\mathcal{G} \subset \mathcal{F}$. Let X be a random variable with finite mean. We define the conditional expectation of X given $\mathcal{F}$, $\mathbb{E}[X \mid \mathcal{F}]$, as the unique random variable satisfying:*

1. *$\mathbb{E}[X \mid \mathcal{F}]$ is $\mathcal{F}$-measurable,*
2. *$\int_A X \, d\mathbb{P} = \int_A \mathbb{E}[X \mid \mathcal{F}] \, d\mathbb{P}$ for all $A \in \mathcal{F}$.*

Theorem 1.3.2. *Consider existing setup as in Definition 1.3.1. In addition consider another random variable Y. Then all the following equalities hold almost surely.*

1. *$\mathbb{E}[X + Y \mid \mathcal{F}] = \mathbb{E}[X|\mathcal{F}] + \mathbb{E}[Y|\mathcal{F}]$.*
2. *$\mathbb{E}(\mathbb{E}[X \mid \mathcal{F}]) = \mathbb{E}[X]$.*
3. *If X is independent of $\mathcal{F}$ then, $\mathbb{E}[X \mid \mathcal{F}] = \mathbb{E}[X]$.*
4. *If X is $\mathcal{F}$-measurable then, $\mathbb{E}[X \mid \mathcal{F}] = X$.*

5. *If $\mathcal{G} \subset \mathcal{F}$ then,* $\mathbb{E}[X \,|\, \mathcal{G}] = \mathbb{E}(\mathbb{E}[X \,|\, \mathcal{F}] \,|\, \mathcal{G})$.

6. *If X is $\mathcal{F}$-measurable and* $\mathbb{E}[XY] < \infty$ *then,* $\mathbb{E}[XY \,|\, \mathcal{F}] = X\mathbb{E}[Y \,|\, \mathcal{F}]$.

Definition 1.3.3. *A filtration on an interval, $[a, b] \in [0, \infty)$, is an increasing family of sigma-fields and is defined as $\{\mathcal{F}_t\} = \{\mathcal{F}_t \,|\, t \in [a, b]\}$.*

For the ease of notation, we suppress the ω in our stochastic processes for the ease of calculation and clarity. For example, $X_t = X(t, \omega)$ and $B_t = B(t, \omega)$.

Definition 1.3.4. *A stochastic process, X_t, is said to be adapted to the filtration $\{\mathcal{F}_t\}$ if for each $t \in [a, b]$, the random variable X_t is $\mathcal{F}_t$-measurable.*

Example 1.3.5. *B_t is adapted with respect to $\{\mathcal{F}_t\} = \sigma\{B_s \,|\, 0 \le s \le t\}$. We refer to $\{\mathcal{F}_t\}$ as the natural Brownian filtration.*

We introduce an extremely important class of processes that are used to model fair games. Due to this feature, they find applications not only in probability theory, but also in mathematical finance and numerous other fields.

Definition 1.3.6. *Let X_t be a stochastic process adapted to the filtration $\{\mathcal{F}_t\}$ with $\mathbb{E}|X_t| < \infty$ for all $t \in [a, b]$. Then, X_t is a martingale with respect to a filtration $\{\mathcal{F}_t\}$ if for any $s \le t$ in $[a, b]$,*

$$\mathbb{E}[X_t \,|\, \mathcal{F}_s] = X_s \;\; a.s\,(almost\,surely).$$

Example 1.3.7. *B_t is a martingale with respect to its natural Brownian filtration. Using the properties of conditional expectation (Theorem 1.3.2) and the independent increments*

of Brownian motion (Definition 1.2.2), whenever $s \leq t$,

$$\begin{aligned}\mathbb{E}\left[B_t \mid \mathcal{F}_s\right] &= \mathbb{E}\left[B_t - B_s + B_s \mid \mathcal{F}_s\right] \\ &= \mathbb{E}\left[B_t - B_s \mid \mathcal{F}_s\right] + \mathbb{E}\left[B_s \mid \mathcal{F}_s\right] \\ &= \mathbb{E}\left[B_t - B_s\right] + B_s \\ &= B_s.\end{aligned}$$

An important theorem to understand the evolution of the evolution of a martingale is the *Optional Stopping Theorem*. This theorem says that in expectation, one would gain nothing by stopping a stochastic process without looking into the future. As asset prices are modelled as martingales, this result is important in asset pricing in mathematical finance. We first define the concept of stopping times and then proceed to the theorem.

Definition 1.3.8. *A random variable $\tau : \Omega \to [a, b]$ is called a stopping time with respect to the filtration $\{\mathcal{F}_t; a \leq t \leq b\}$ if $\{\omega; \tau(\omega) \leq t\} \in \mathcal{F}_t$ for all $t \in [a, b]$.*

Theorem 1.3.9. *Let M_t, $a \leq t \leq b$ be a martingale with respect to the filtration $\{\mathcal{F}_t\}$ and τ a stopping time. Then, $\mathbb{E}|M_{t\wedge\tau}| < \infty$ and*

$$\mathbb{E}[M_\tau] = \mathbb{E}[M_a]$$

if any of the following conditions are true:

1. *τ is bounded almost surely,*
2. *There exists is a positive number, K, such that $|M_t| < K$ for all ω and t with τ almost surely bounded.*

1.4. Classical Ito Theory

From hereon, consider the probability space $(\Omega, \{\mathcal{F}_t\}, \mathbb{P})$ with $t \in [a,b]$ and $\{\mathcal{F}_t\}$ the natural Brownian filtration. Define $L^2_{ad}([a,b] \times \Omega)$ as the space of all stochastic processes, $f(t,\omega)$ with $t \in [a,b]$ and $\omega \in \Omega$ such that,

- $f(t,\omega)$ is adapted to the filtration $\{\mathcal{F}_t\}$,
- $\int_a^b \mathbb{E}\left[|f(t)|^2\right] \, dt < \infty$.

We start off the exposition by first introducing the Wiener integral.

Definition 1.4.1. *Let $t \in [a,b]$, B_t be Brownian motion and $f(t) \in L^2[a,b]$. The Wiener integral of $f(t)$ is denoted by*

$$\int_a^b f(t) \, dB_t,$$

and is defined in two steps.

Step 1*: Consider a partition on $[a,b]$ with $a = t_0 \leq t_1 \leq \cdots \leq t_n = b$. Define the step function given by*

$$f(t) = \sum_{i=1}^n a_i 1_{[t_{i-1},t_i)}(t), \quad t \in [a,b].$$

We define its Wiener integral as

$$\int_a^b f(t) \, dB_t = \sum_{i=1}^n a_i (B_{t_i} - B_{t_{i-1}}).$$

Step 2*: For any $f \in L^2[a,b]$, we choose a sequence of step functions $f_n \in L^2[a,b]$ such that $f_n \to f$. Via step 1, $\left\{\int_a^b f_n(t) dB(t)\right\}$ is Cauchy in $L^2(\Omega)$. As such we define*

$$\int_a^b f(t) \, dB_t = \lim_{n\to\infty} \int_a^b f_n(t) \, dB_t, \quad \textit{in } L^2(\Omega).$$

The Itô integral can be seen as a generalization of the Wiener integral. While the

construction is similar, the definition of the Itô integral has three steps and a key approximation lemma for step 2. We refer to [15] for details.

Definition 1.4.2. *Let $t \in [a, b]$, B_t be Brownian motion and $f(t)$ be an adapted mean square integrable stochastic process. The Itô integral of $f(t)$ is denoted by*

$$\int_a^b f(t)\, dB_t,$$

and is defined in three steps.

Step 1*: Consider a partition on $[a, b]$ with $a = t_0 \leq t_1 \leq \cdots \leq t_n = b$. Define the step stochastic process given by*

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

where ξ_{i-1} is $\mathcal{F}_{i-1}$ measurable and $\mathbb{E}\left[\xi_{i-1}^2\right] < \infty$. We define its Itô integral as

$$\int_a^b f(t)\, dB_t = \sum_{i=1}^{n} \xi_{i-1}(B_{t_i} - B_{t_{i-1}}).$$

Step 2*: We can show that for any $f \in L^2_{ad}([a, b] \times \Omega)$, there exists a sequence of adapted step stochastic functions f_n that converges to f in $L^2_{ad}([a, b] \times \Omega)$.*

Step 3: *For a general square integrable adapted function f, we define its Itô integral as*

$$\int_a^b f(t)\, dB_t = \lim_{n \to \infty} \int_a^b f_n(t)\, dB_t, \quad \text{in } L^2(\Omega).$$

This definition is well defined as shown in [15]. While the Itô integral does not follow the usual "partition - evaluation - summation - limit" strategy for Riemann integrals, the left hand evaluation of the step functions have an intuitive similarity to Riemann integrals. We state this property below.

Theorem 1.4.3. *Let $f(t)$ be a square integrable adapted stochastic process and suppose $\mathbb{E}[f(s)f(t)]$ is a continuous function of t and s . Then*

$$\int_a^b f(t)\,dB_t = \lim_{\|\Delta_n\|\to 0} \sum_{i=1}^{n} f(t_{i-1})\left(B_{t_i} - B_{t_{i-1}}\right) \quad in\ L^2(\Omega),$$

where Δ_n's are partitions of $[a,b]$.

We view this via a simple example.

Example 1.4.4. *[15] Let $t \in [0,1]$. We want to evaluate $\int_0^1 B_t dB_t$. From Example 1.3.5, we have that the integrand is adapted and as such, we are looking at an Itô integral. Furthermore, $\mathbb{E}[B_s.B_t] = min(t,s)$ which is continuous for both t and s. Hence, we can use Theorem 1.4.3 the above integral as a Reimann-like sum. Namely,*

$$\begin{aligned}
\int_a^b B_t\,dB_t &\approx \sum_{i=1}^{n} B_{t_{i-1}}\left(B_{t_i} - B_{t_{i-1}}\right) \\
&= \sum_{i=1}^{n}\left[\frac{1}{2}\left(B_{t_{i-1}} + B_{t_{i-1}} - \left(B_{t_i} - B_{t_{i-1}}\right)\right)\right]\left(B_{t_i} - B_{t_{i-1}}\right) \\
&= \frac{1}{2}\sum_{i=1}^{n}\left[\left(B_{t_i}\right)^2 - \left(B_{t_{i-1}}\right)^2\right] + \frac{1}{2}\sum_{i=1}^{n}\left(B_{t_i} - B_{t_{i-1}}\right)^2.
\end{aligned}$$

The first summation term is an alternating series summation which leaves only the first and the last term. The second summation term is the quadratic variation of Brownian motion and as such, we use Theorem 1.2.3 to get,

$$\int_a^b B_t\,dB_t = \frac{1}{2}\left(B_b^2 - B_a^2 - (b-a)\right).$$

Now that we have defined the Itô integral, we look into some of its fundamental properties. For their proofs, refer to [15].

Definition 1.4.5. *Let $f \in L^2_{ad}([a,b]\times\Omega)$ then the Itô integral, $\int_a^b f(t)\,dB_t$, is a random*

variable with mean 0 *and variance*

$$\mathbb{E}\left[\left(\int_a^b f(t)\,dB_t\right)^2\right] = \int_a^b \mathbb{E}\left[|f(t)|^2\right]\,dt. \tag{1.1}$$

In addition, if f is deterministic then its' Itô integral is normally distributed.

From equation 1.1, we can see that the mapping $f \to \int_a^b f(t)\,dB(t)$ is an isometry from $L^2_{ad}([a,b]\times\Omega)$ to $L^2(\Omega)$.

The stochastic integral performs a continuous time martingale transform of the integrand. As such, the Itô integral is also martingale as shown in the theorem below.

Theorem 1.4.6. *Let X_t be a stochastic process such that*

$$X_t = \int_a^t f(s)\,dB_s\,,\; f \in L^2_{ad}([a,b]\times\Omega),$$

then X_t is a continuous martingale.

We look at the diffusion processes that motivated K. Itô to introduce the Itô integral.

Definition 1.4.7. *An Itô process is a stochastic process of the form*

$$X_t = X_a + \int_a^t f(t)\,dB_t + \int_a^t g(t)\,dt,$$

with X_a - $\mathcal{F}_a$ measurable, $f \in L^2_{ad}([a,b]\times\Omega)$ and, $g \in L^1_{ad}([a,b]\times\Omega)$.

In the above definition of the Itô process, the first integral is an Itô integral of a mean square integrable adapted process and by Theorem 1.4.6, it is continuous. The second integral is a Riemann integral for each ω. As such, the Itô process is continuous.

An integral part of Itô theory is the Ito formula. It can be seen as the stochastic analog of the "chain rule". It is useful in showing existence and uniqueness as well as in

creating new stochastic processes from existing one. Furthermore, it helps in understanding stability results for stochastic differential equations among many other applications. Known classical results would be hard to obtain without the following *Itô formula.*

Theorem 1.4.8. *Let X_t be an Itô process given by*

$$X_t = X_a + \int_a^t f(t)\,dB_t + \int_a^t g(t)\,dt,$$

with X_a - $\mathcal{F}_a$ measurable, $f \in L^2_{ad}([a,b]\times\Omega)$ and, $g \in L^1_{ad}([a,b]\times\Omega)$. Suppose $\Theta(t,x)$ is a continuous function with continuous partial derivatives. Then $\Theta(t,X_t)$ is also an Itô process and is given by

$$\begin{aligned}\Theta(t,X_t) = \Theta(a,X_a) &+ \int_a^t \frac{\partial\Theta}{\partial x}(s,X_s)\,dB_s \\ &+ \int_a^t \left[\frac{\partial\Theta}{\partial t}(s,X_s) + \frac{\partial\Theta}{\partial x}(s,X_s)g(s) + \frac{1}{2}\frac{\partial^2\Theta}{(\partial x)^2}(s,X_s)\right] ds.\end{aligned}$$

The multiplication of the differentials follow the rule shown Itô table in Table 1.1:

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

Table 1.1. Itô table

For the ease of computation and calculation, it is convenient to express stochastic integral equations in their stochastic differential form. However, this is just mathematical convention as dB_t/dt by itself does not make sense as almost all Brownian paths are nowhere differentiable. As such, the stochastic differential equations are to be interpreted as stochastic integral equations. For example, the differential form of the Itô process, X_t, in Definition 1.4.7 is given by

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

As mentioned earlier, we can use the Itô formula to make new stochastic processes or to evaluate stochastic integrals without resorting to the definitions. We show this feature with an example.

Example 1.4.9. *Consider the function $f(x) = x^2$. Then $f_x = 2x$ and $f_{xx} = 2$. We use the Itô formula in Theorem 1.4.8 to get*

$$\begin{aligned} df(B_t) &= f_x(B_t)dB_t + \frac{1}{2}f_{xx}(B_t)dt, \\ d(B_t^2) &= 2B_t dB_t + dt. \end{aligned}$$

Hence, we have

$$B_t^2 - B_a^2 = 2\int_a^t B_s dB_s + (t-a).$$

We solve for the integral to get

$$\int_a^t B_s dB_s = \frac{1}{2}\left((B_t^2 - B_a^2) - (t-a)\right).$$

Note that when $t = b$ in the above example, we would get the same result as in Example 1.4.4.

1.5. Exponential Process and Girsanov Theorem

Exponential functions are important in ordinary differential equation theory due to their in variance with the derivative operator. We review a stochastic process with similar significance with respect to the stochastic differential equation.

Example 1.5.1. *Let $f \in L^2_{ad}([a,b] \times \Omega)$ and define the stochastic process,*

$$\mathcal{E}_f(t) = \exp\left[\int_a^t f(s)\, dB_s - \frac{1}{2}\int_a^t f^2(s)ds\right].$$

Using the Itô formula in Theorem 1.4.8, we have $d\mathcal{E}_f(t) = f(t)\mathcal{E}_f(t)dB_t$. Hence, $\mathcal{E}_f(t)$ is a martingale.

Exponential processes are important in the theory of stochastic integration as it is the direct analogue of the exponential function in ordinary differential equations. It is particularly useful in showing that translations of Brownian motion remains Brownian motion under an equivalent probability measure. This result is called the Girsanov Theorem and is listed below.

Theorem 1.5.2. *Let* $f \in L^2_{ad}([a,b] \times \Omega)$ *and let* $\mathbb{E}_{\mathbb{P}}[\mathcal{E}_f(t)] = 1$ *for all* t. *Then*

$$W_t = B_t - \int_a^t f(s)ds \ \ , \ t \in [a,b],$$

is Brownian motion with respect to the probability measure $d\mathbb{Q} = \mathcal{E}_f(b)d\mathbb{P}$.

Chapter 2. Rare Events and Large Deviation Principle

2.1. Introduction

The chance of buying a winning power ball ticket is astronomically low. However, people still buy lotto tickets in large numbers despite knowing that winning is a rare event. They do so because the potential win would completely change their life. Rare events have the potential to leave a significant impact. As such, understanding how these rare events happen is an interesting field of study.

This concept first appeared in mathematical history though the works of Swedish mathematician Harald Cramér in modelling an insurance business[4]. If one is operating an insurance company, they would have to consider both earning and claims when looking at profitability for the company. We can assume that the earning come as premium at a constant daily rate, x and the claims, X_i, arrive at a random rate. For any fixed earning period, if the total claims is more than the total earning then the profit margins for the company would decrease. Then what is the premium they have to charge their customers in order to have the average future claims to be less than the average daily future earnings?

Let use review some standard probabilistic results to analyse the question.

Theorem 2.1.1 (Law of Large Numbers)**.**

$$S_n \to \mu, \quad as\ n \to \infty,$$

where $S_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ *and the limit is in sense of convergence in probability in the weak sense and almost surely in the strong sense*

Theorem 2.1.2 (Central Limit Theorem)**.**

$$\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(X_i-\mu)\right)\rightarrow N(o,\sigma^2),$$

where the limit is in distribution

Theorem 2.1.1 tells us that as the number of claims increases, asymptotically the average value of a claim would approach a number μ while Theorem 2.1.2 tells us what the limiting distribution would be normal around the number μ. However, we do not know at what rate the average claim would approach μ. Let us assume that x is the premium earned per day and $S_n = \sum_{i=1}^{n} X_i$ is the average claim on the nth day. We want to know how the probability of the event that the average future claim is more that the average daily earnings depends on the premium charged. Namely we want to know how $P(|S_n \geq x|)$ is dependent of the value of x. We would want this probability to not only be small but exponentially small as to guarantee that that the chance of losing profitability for the company is exceedingly small. Large deviation principle is this study of rare events via analysis of probabilities of rare events that are exponentially small.

Heuristically, consider a topological space $\mathcal{X}$ and a complete Borel σ-field $\mathcal{B}$. Large deviation principle characterizes the asymptotic behavior of a family of probability measures, $\{\mu_\epsilon\}$, on $(\mathcal{X},\mathcal{B})$ as ϵ decreases to zero via a *rate function*. It does so via the asymptotic exponential bounds on the open and closed sets in the topological space.

In the following sections, we review some preliminary definitions and explore some known results in the theory. We follow the exposition given by Dembo and Zeitouni in [5].

2.2. Large Deviation Principle and Some Properties

We start off with the definition of the rate function and Large Deviation Principle.

Definition 2.2.1. *A function, I, is called a rate function on $\mathcal{X}$ if I maps $\mathcal{X}$ to $\mathbb{R}$ and if the level sets of I are compact subsets of $\mathcal{X}$*

Definition 2.2.2. *A family of probability measures, $\{\mu_\epsilon\}$, satisfies the large deviation principle with a unique rate function, $I(\bullet)$ if the following inequalities hold:*

- *(Uppper Bound)*
$$\limsup_{\epsilon\to 0} \epsilon \log \mu_\epsilon(F) \le -\inf_{x\in F} I(x),$$
for all closed sets F.

- *(Lower Bound)*
$$\liminf_{\epsilon\to 0} \epsilon \log \mu_\epsilon(O) \ge -\inf_{x\in O} I(x),$$
for all open sets O.

Now that we have defined what large deviation principle is, we look into some useful results in the theory. The following result is called *Contraction principle* and states that the continuous image of random variables satisfying large deviation principle also satisfies large deviation principle.

Theorem 2.2.3 (Theorem 4.2.1 of [5])**.** *Let $\mathcal{X}$ and $\mathcal{Y}$ be two separable metrizable topological spaces, I a rate function on $\mathcal{X}$, and f a continuous function mapping $\mathcal{X}$ to $\mathcal{Y}$. Then the following conclusions hold.*

1. *For each $y \in \mathcal{Y}$,*
$$J(y) \triangleq \inf\left\{I(x) \,|\, x \in f^{-1}(y)\right\}$$
is a rate function on $\mathcal{Y}$,

2. *If $\{X_n\}$ satisfies large deviation principle on $\mathcal{X}$ with rate function I, then $\{f(X_n)\}$ satisfies large deviation principle on $\mathcal{Y}$ with rate function J.*

Large deviation principle is also preserved between sequences of random variables that are super-exponentially close to each other as shown by the following theorem.

Theorem 2.2.4 (Theorem 4.2.16 of [5]). *For $n \in \mathbb{N}$, let X_n and Y_n be random variables on $(\Omega, \mathcal{F}, P)$ and take values in $\mathcal{X}$. Given that $\{X_n\}$ satisfies large deviation principle on $\mathcal{X}$ with rate function I and that $\{Y_n\}$ is super-exponentially close to $\{X_n\}$ i.e,*

$$\limsup_{n\to\infty} \frac{1}{n} \log P\{d(X_n, Y_n) > \delta\} = -\infty.$$

2.3. Large Deviation Principle for Random Variables

We return to the mathematical setting that Cramér investigated the insurance problem as introduced in Section 2.1.

Theorem 2.3.1 (Theorem 2.2.3 of [5]). *Let $X_n \in \mathbb{R}$ be i.i.d random vectors with X_1 distributed according to a d-dimensional probability measure μ. Let Λ be the logarithmic moment generating function associated with μ, namely,*

$$\Lambda(\lambda) \triangleq \log \mathbb{E}[e^{\langle \lambda, X_1 \rangle}].$$

Then, the sequence of probability measures $\{\mu_n\}$ satisfies large deviation principle with the convex rate function $\Lambda^*(\cdot)$, namely:

1. *For any closed $F \subset \mathbb{R}$,*
$$\limsup_{n\to\infty} \frac{1}{n} \log \mu_n(F) \leq -\inf_{x\in F} \Lambda^*(x).$$

2. *For any open $O \subset \mathbb{R}$,*
$$\limsup_{n\to\infty} \frac{1}{n} \log \mu_n(O) \geq -\inf_{x\in O} \Lambda^*(x).$$

where μ_n is the law of $S_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ and $\Lambda^*(x) \triangleq \sup_{\lambda\in\mathbb{R}^d} \{\langle \lambda\,,\, x\rangle - \Lambda(x)\}$ is the Frechet-Legendre transform of Λ.

We refer the reader to the Garter-Ellis Theorem for non-i.i.d case in [5].

2.4. Large Deviation Principle for Stochastic Processes

Let us explore some results to familiarize ourselves to some large deviation results on the sample paths. We start off with some results for the building blocks of stochastic calculus. From Theorem 2.3.1, we have an understanding about the large deviation behavior of the sample mean. Given a sequence of random variables, the following theorem provides insights into the large deviation behavior of the whole set of random variables indexed by t. For this section, we fix $t \in [0, 1]$.

Given a sequence of i.i.d random variables $\{X_n\}$ with $\Lambda(\lambda) \triangleq \log \mathbb{E}\left(e^{\langle \lambda, X_1 \rangle}\right) < \infty$ for all $\lambda \in \mathbb{R}$, we consider the random walk given by,

$$Z_n(t) = \frac{1}{n} \sum_{i=1}^{\lfloor nt \rfloor} X_i, \quad 0 \leq t \leq 1.$$

Theorem 2.4.1 (Theorem 5.2.1 of [5]). *Let μ_n is the law of $Z_n(\cdot)$ in $L_\infty([0,1])$. Then the measures, μ_n satisfy large deviation principle in $L_\infty([0,1])$ with the rate function,*

$$I(\psi) = \begin{cases} \displaystyle\int_0^1 \Lambda^*(\dot{\psi}(t))\, dt, & \text{if } \psi \in \mathcal{AC}, \\ \infty, & \text{otherwise}, \end{cases}$$

where $\mathcal{AC}$ denotes the space of absolutely continuous functions and $\Lambda * (x)$ is the Frechet-Legendre transform of $\Lambda(\bullet)$.

We proceed to Schilder's Theorem which establishes sample path deviations for Brownian motion. Let B_t, $t \in [0, 1]$, be Brownian motion starting at 0 and for all $\epsilon > 0$, let p_ϵ be the probability measure induced by $\sqrt{\epsilon} B_\bullet$ on $C_0([0,1])$. Then as $\epsilon \to 0$, $p_\epsilon \implies \delta_0$. As such,

Theorem 2.4.2 (Theorem 5.2.3 of [5]). *The sequence of probability measure $\{p_\epsilon\}$ as $\epsilon \to 0$ follows large deviation principle on $C_0([0,1])$ with rate function $I(f)$ where*

$$I(f) = \begin{cases} \frac{1}{2}\int_0^1 |f'(t)|^2 dt, & \textit{if } f \in \mathbb{H}^1, \\ \\ \infty, & \textit{Otherwise,} \end{cases}$$

where $\mathbb{H}^1 = \{f \in C_0\left([0,1]\right) \ | \, f(t) = \int_0^t f'(s)ds \ \textit{and} \ f' \in L^2[0,1]\}$.

Let us look at how we can use an existing large deviation principle results to obtain new ones.

Example 2.4.3. *Let* $Y_t^\epsilon = \sqrt{\epsilon}(B_1 - B_t)$ *with* $t \in [0,1]$*. Then,* $\{Y_t^\epsilon\}$ *follows large deviation principle on* $C_{1,0}[0,1]$ *with the rate function* J *given by*

$$J(\psi) = \begin{cases} \dfrac{1}{2}\displaystyle\int_0^1 |f(t)|^2 dt, & \textit{if } \psi(t) = \displaystyle\int_0^{1-t} f(s)\, ds \ \textit{with } f \in L^2[0,1], \\ \\ \infty, \quad \textit{otherwise,} \end{cases}$$

where $C_{1,0}[0,1] = \{f \,|\ f \textit{ is continuous and } f(1) = 0\}$.

Proof. Define $Z_t = (B_1 - B_{1-t})$ for $t \in [0,1]$. Then we have the following properties:

1. $Z_0 = B_1 - B_1 = 0$.

2. Let $0 \le u \le v \le s \le t \le 1$ then, we have that $1 \ge 1-u \ge 1-v \ge 1-s \ge 1-t \ge 0$. As such, $Z_t - Z_s = B_{1-s} - B_{1-t}$ is independent of $Z_v - Z_u = B_{1-u} - B_{1-v}$ via the independent increments of Brownian motion.

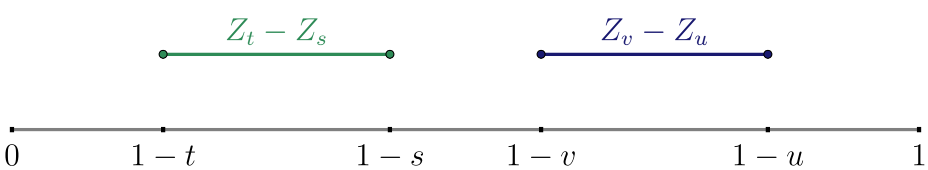

Figure 2.1. A t-dependence plot of the disjoint increments of $Z_\bullet$.

3. $Z_t - Z_s = B_{1-s} - B_{1-t} \sim N(0, t-s)$ for $0 \le s \le t \le 1$.

4. Z_t is continuous almost surely as a sum of continuous functions.

From Definition 1.2.2, we have that Z_t is a Brownian motion. Consider the function $\Lambda(\psi_t) = \psi_{1-t}$ for f continuous. Then, $\Lambda : C_0[0,1] \to C_{1,0}[0,1]$ is a continuous function that is its own inverse. Furthermore, $\Lambda(\sqrt{\epsilon}Z_t) = \Lambda(\sqrt{\epsilon}(B_1 - B_{1-t})) = \sqrt{\epsilon}(B_1 - B_t) = Y_t^\epsilon$. Using the contraction mapping property in Theorem 2.2.3 along side the large deviation principle results in Theorem 2.4.2, Y_t^ϵ follows large deviation principle with the rate function given by

$$
\begin{aligned}
J(\psi) &= I \circ \Lambda^{-1}(\psi_\bullet) = I(\psi(1-\bullet)) \\
&= \begin{cases} \dfrac{1}{2}\displaystyle\int_0^1 |f(t)|^2 dt, & \text{if } \psi(1-t) = \displaystyle\int_0^t f(s)\, ds \text{ with } f \in L^2[0,1], \\ \infty, & \text{otherwise,} \end{cases} \\
&= \begin{cases} \dfrac{1}{2}\displaystyle\int_0^1 |f(t)|^2 dt, & \text{if } \psi(t) = \displaystyle\int_0^{1-t} f(s)\, ds \text{ with } f \in L^2[0,1], \\ \infty, & \text{otherwise,} \end{cases}
\end{aligned}
$$

where we performed a change of variables for the last equality. □

Via Schilder's Theorem and an application of the super-exponential approximation given in Theorem 2.2.4, we can obtain large deviation results for a class of solutions of stochastic differential equations, namely :

Theorem 2.4.4 (Theorem 5.6.3 of [5]). *Given a stochastic differential equation with state driven noise:*

$$X_\epsilon(t) = X_\epsilon(0) + \sqrt{\epsilon}\int_0^t \alpha(X_\epsilon(s))dB(s) + \int_0^t \beta(X_\epsilon(s))ds, \quad t \in [0,1],$$

with $\{X_\epsilon(t)\} \sim \mu_\epsilon$ and $X(0) = 0$. The sequence of probability measures $\{\mu_\epsilon\}$ satisfies large

deviation principle with rate function I given by,

$$I(f) = \begin{cases} \frac{1}{2}\int_0^1 \left\langle f'(t) - \beta(f(t)) \ , \ A^{-1}(t)\left(f'(t) - \beta\left(f\left(t\right)\right)\right)\right\rangle dt, & \textit{if } f \in \mathbb{H}^1, \\ \infty, & \textit{otherwise,} \end{cases}$$

where $A(t) = \alpha\alpha^*(t)$.

We are interested in cases where the stochastic differential equations need not necessarily be adapted. From the definition of the Itô integral in Definition 1.4.2, we see that the stochastic integral is not defined when the integrand is not adapted. As such, we first switch our focus to the Ayed–Kuo integral which provides an intuitive and simplistic way to approach anticipating integrals.

Chapter 3. General Theory of Stochastic Analysis

3.1. Introduction

The following question was posed by Itô in the 1976 International Symposium of Stochastic Differential Equations in Kyoto [13]. Is

$$\int_0^1 B_1 dB_t = B_1 \int_0^1 dB_t = B_1^2 \quad ? \tag{3.1}$$

He added that B_1 is not adapted to the natural Brownian filtration $\mathcal{F}_t$ and as such, $\int_0^1 B_1 dB_t$ can not be defined as an Itô integral. In working with this limitation, he proposed an initial enlargement of filtration by taking $\mathcal{G}_t = \sigma\{B_1, B_s \,|\, 0 \leq s \leq t\}$ and decomposed the Brownian motion as

$$B_t = \left(B_t - \int_0^t \frac{B_1 - B_u}{1-u}\, du \right) + \int_0^t \frac{B_1 - B_u}{1-u}\, du.$$

In this formulation, B_t is quasi-martingale with respect to the new filtration while B_1 is adapted to $\mathcal{G}_t$. As such, Equation (3.1) holds true in the enlarged filtration.

The Ayed–Kuo stochastic integral takes inspiration from Itô. Instead of keeping the integrand B_1 while simultaneously changing the filtration and decomposing the integrator B_t, the new stochastic integral maintains both the Brownian motion and the filtration while decomposing the integrand into adapted part and the counter-part called instantly independent. In this chapter, we will review the Ayed–Kuo Stochastic integral and prove some new results.

Section 3.3 and Section 3.4 of this chapter previously appeared, substantially unchanged, in the open access *Journal of Stochasic Analysis* article: Kuo, H.-H., Shrestha, P., and Sinha, S.,“An Intrinsic Proof of an Extension of Itô’s Isometry for Anticipating Stochastic Integrals,” `https://digitalcommons.lsu.edu/josa/vol2/iss4/8/` and in the open-access *arXiv* article: Kuo, H.-H., Shrestha, P., Sinha, S., and Sundar, P., “On near-martingales and a class of anticipating linear SDEs,” `https://arxiv.org/abs/2204.01932` respectively.

3.2. Instantly Independent and Counter Filtration

Let us first define the notion of instantly independence and the filtration space we operate on for instantly independent processes.

Definition 3.2.1. *A stochastic process, $f(t)$, is said to be instantly independent with respect to the filtration $\{\mathcal{F}_t\}$ if $f(t)$ is independent to $\mathcal{F}_t$ for each $t \in [a, b]$.*

Definition 3.2.2. *A family, $\{\mathcal{G}^{(t)}\}$, of complete sigma fields is called a Counter-filtration of $\{\mathcal{F}_t\}$ if*

1. *For each $t \in [a.b]$, $\mathcal{G}^{(t)}$ is independent of $\mathcal{F}_t$,*
2. *for each $a \leq s < t \leq b$, $\mathcal{G}^{(s)} \supset \mathcal{G}^{(t)}$.*

We define the counter-filtration process space for our instantly independent processes.

Definition 3.2.3. *Define $L^2_{ct}([a, b] \times \Omega)$ as the space of all stochastic processes, $g(t, \omega)$, $t \in [a, b]$, $\omega \in \Omega$, satisfying,*

- *$g(t, \omega)$ is adapted to the natural Brownian counter-filtration, $\mathcal{G}^{(t)}$,*
- *$\int_a^b E\left[(\psi_t)^2\right] ds < \infty$.*

Remark 3.2.4. *$L^2_{ct}([a, b] \times \Omega)$ is a subspace orthogonal to $L^2_{ad}([a, b] \times \Omega)$.*

Example 3.2.5. *Define $\mathcal{F}_t = \sigma\{B_s - B_a | a \leq s \leq t \leq b\}$ and $\mathcal{G}^{(t)} = \sigma\{B_b - B_s | a \leq t \leq s \leq b\}$. Then, $\{\mathcal{G}^{(t)}\}$ is a counter-filtration of $\{\mathcal{F}_t\}$. We will refer these filtrations as the Natural Brownian filtration and Natural Brownian counter-filtration.*

The action of the stochastic integral on processes adapted to the natural Brownian counter-filtration is very similar to the action of the Ayed–Kuo integral on processes adapted to the natural Brownian filtration. Namely, the integral acts as an isometry be-

tween two spaces. As such, we have the following result.

Theorem 3.2.6 (Proposition 2.1.3 and Theorem 2.1.8 of [33])**.** *Let ψ_t be a square integrable stochastic process adapted to the natural Brownian counter-filtration, $\mathcal{G}^{(t)}$ or alternately $\psi \in L^2_{ct}([a,b] \times \Omega)$. Then*

1. *ψ_t is instantly independent to the natural Brownian filtration $\{\mathcal{F}_t\}$,*
2. $\mathbb{E}\left[\int_a^b \psi_s \, dB_s\right] = 0,$
3. $\mathbb{E}\left[\left(\int_a^b \psi_s \, dB_s\right)^2\right] = \int_a^b \mathbb{E}\left[(\psi_t)^2\right] ds.$

It is natural to ask what if a function is both instantly independent and adapted? Then the function would be deterministic. Indeed if a function $f(t)$, $t \in [a,b]$ is $\mathcal{F}_t$-adapted and instantly independent for each t then,

$$f(t) = \mathbb{E}\left[f(t) \,|\, \mathcal{F}_t\right] = \mathbb{E}\left[f(t)\right], \quad t \in [a,b],$$

where the first equality is via adaptedness and the second equality is via instantly independence. Thus, $f(t) = \mathbb{E}\left[f(t)\right]$ for all t in $[a,b]$. This means that f is deterministic.

3.3. Ayed–Kuo Stochastic Integral

As mentioned earlier, for the Ayed–Kuo stochastic integral, we decompose the integrand into adapted and instantly independent parts. Let us view it via an illuminatory example.

Example 3.3.1 (Example 2.4 of [10])**.** *Consider $\int_a^b B_b \, dB_t$. Since the integrand, B_b, is not $\{\mathcal{F}_t\}$-adapted, the integral is not defined within Itô theory. Note that for each t, we can decompose*

$$B_b = B_t + (B_b - B_t),$$

where B_t is $\mathcal{F}_t$-measurable, and $B_b - B_t$ is independent of $\mathcal{F}_t$ due to the independence of increments of Brownian motion. Thus, we have decomposed the anticipating integrand into an adapted and an instantly independent process.

Motivated by Itô's original construction, we take left end point approximation for the adapted parts and right end point approximation for the instantly independent part. This way we can "define" the integral

$$\begin{aligned}
\int_a^b B_b \, dB_t = &\int_a^b \left[B_t + (B_b - B_t)\right] \, dB_t \\
&= \lim_{\|\Delta_n\| \to 0} \sum_{j=1}^{n} \left[B_{t_{j-1}} + \left(B_b - B_{t_j}\right)\right] \left(B_{t_j} - B_{t_{j-1}}\right) \\
&= \lim_{\|\Delta_n\| \to 0} \sum_{j=1}^{n} \left[B_b - \left(B_{t_j} - B_{t_{j-1}}\right)\right] \left(B_{t_j} - B_{t_{j-1}}\right) \\
&= B_b \lim_{\|\Delta_n\| \to 0} \sum_{j=1}^{n} (B_{t_j} - B_{t_{j-1}}) - \lim_{\|\Delta_n\| \to 0} \sum_{j=1}^{n} (B_{t_j} - B_{t_{j-1}})^2 \\
&= B_b \lim_{\|\Delta_n\| \to 0} \sum_{j=1}^{n} (B_{t_j} - B_{t_{j-1}}) - \lim_{\|\Delta_n\| \to 0} \sum_{j=1}^{n} (t_j - t_{j-1}) \\
&= B_b \left(B_b - B_a\right) - (b - a),
\end{aligned}$$

where in the last equality, we have used the quadratic variation of Brownian motion.

With this example to illuminate the path ahead, we introduce the Ayed–Kuo stochastic integral.

Definition 3.3.2. *The Ayed–Kuo stochastic integral of a stochastic process $\Phi(t)$ introduced in [1] is defined in the following three steps.*

1. *Suppose $f(t)$ is an $\mathcal{F}_t$-adapted continuous stochastic process and $\phi(t)$ be an continuous stochastic processes that is instantly independent with respect to $\mathcal{F}_t$. Then the*

stochastic integral of $\Phi(t) = f(t)\phi(t)$ *is defined by*

$$\int_a^b f(t)\phi(t)\,dB(t) = \lim_{\|\Delta_n\|\to 0} \sum_{j=1}^{n} f(t_{j-1})\phi(t_j)(B(t_j) - B(t_{j-1})),$$

provided that the limit exists in probability.

2. *For a process of the form* $\Phi(t) = \sum_{i=1}^{n} f_i(t)\phi_i(t)$, *the stochastic integral is defined by*

$$\int_a^b \Phi(t)\,dB(t) = \sum_{i=1}^{n} \int_a^b f_i(t)\phi_i(t)\,dB(t).$$

3. *Let* $\Phi(t)$ *be a stochastic process such that there is a sequence* $\{\Phi_n(t)\}_{n=1}^{\infty}$ *of stochastic processes of the form in step 2 satisfying*

   (a) $\int_a^b |\Phi_n(t) - \Phi(t)|^2\,dt \to 0$ *almost surely as* $n \to \infty$, *and*

   (b) $\int_a^b \Phi_n(t)\,dB(t)$ *converges in probability as* $n \to \infty$.

   Then the stochastic integral of $\Phi(t)$ *is defined by*

$$\int_a^b \Phi(t)\,dB(t) = \lim_{n\to\infty} \int_a^b \Phi_n(t)\,dB(t) \quad \text{in probability.}$$

This integral is well defined, as demonstrated by the following lemma.

Lemma 3.3.3 (Lemma 2.1 of [10])**.** *Let* $f_i(t), 1 \le i \le m$ *and* $g_j(t), 1 \le j \le n$ *be* $\{\mathcal{F}_t\}$-*adapted stochastic processes and let* $\phi_i(t), 1 \le i \le m$ *and* $\psi_j(t), 1 \le j \le n$ *be instantly independent with respect to* $\{\mathcal{F}_t\}$. *Suppose the stochastic integrals* $\int_a^b f_i(s)\phi_i(s)\,dB(s)$ *and* $\int_a^b g_j(s)\psi_j(s)\,dB(s)$ *exist for all i, j. Assume that,*

$$\sum_{i=1}^{m} f_i(t)\phi_i(t) = \sum_{j=1}^{n} g_j(t)\psi_j(t).$$

Then,

$$\sum_{i=1}^{m} \int_a^b f_i(s)\phi_i(s)\,dB(s) = \sum_{j=1}^{n} \int_a^b g_j(s)\psi_j(s)\,dB(s).$$

This technique of decomposing the integrand into adapted and instantly independent processes easily extends to case when the integrand is a product of adapted and anticipating processes, as the following example demonstrates.

Example 3.3.4. *Using the decomposition* $B_bB_t = (B_b - B_t)B_t + (B_t)^2$ *for* $t \in [a, b]$, *we write the integrand as a sum of products of adapted and instantly independent parts. As such, we can use the definition to get*

$$\begin{aligned}\int_a^b B_bB_t \; dB_t = &\int_a^b \left(B_t(B_b - B_t) + B_t^2\right) \; dB_t \\ &= B_b \int_a^b B_t \; dB_t - \int_a^b B_t \; dt - \cancel{\int_a^b B_t^2 \; dB_t} + \cancel{\int_a^b B_t^2 \; dB_t} \\ &= \frac{1}{2} B_b \left(B_b^2 - B_a^2 - b + a\right) - \int_a^b B_t \; dt.\end{aligned}$$

3.3.1. Mean of the Ayed–Kuo Stochastic Integral

We first look at the mean of the Ayed–Kuo stochastic integral. From Definition 1.4.5, we have that the Itô integral is a mean zero process. We expect the same for the Ayed–Kuo stochastic integral. The following result shows that it is indeed the case for a certain class of integrands.

Theorem 3.3.5. *Let* $\Phi(x, y)$ *be continuous on both variable such that* $\Phi(B_t, B_b - B_t) \in L^2([a, b] \times \Omega)$. *Furthermore, assume that the partial sums*

$$\sum_{i=1}^{n} \Phi(B_{t_{i-1}}, B_1 - B_{t_i}) \left(B_{t_i} - B_{t_{i-1}}\right)$$

are uniformly integrable. Then

$$\mathbb{E}\left[\int_a^b \Phi(B_t, B_b - B_t) \; dB_t\right] = 0.$$

Proof. Using the definition of the Ayed–Kuo stochastic integral in Definition 3.3.2 and the uniform integrability condition on the partial sums to interchange the limit and the

expectation, we have

$$\mathbb{E}\left[\int_a^b \Phi(B_t, B_b - B_t)\ dB_t\right] = \lim_{\|\Delta\|_n \to 0} \sum_{i=1}^n \mathbb{E}\left[\Phi(B_{t_{i-1}}, B_b - B_{t_i})\left(B_{t_i} - B_{t_{i-1}}\right)\right] \tag{3.2}$$

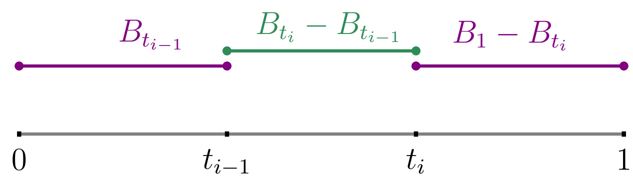

Figure 3.1. A t-dependence plot of the disjoint increments of $B_\bullet$.

Note that, as independent increments of Brownian motion, both $B_{t_{i-1}}$ and $B_b - B_{t_i}$ are independent of $\left(B_{t_i} - B_{t_{i-1}}\right)$. As such, continuity of Φ implies that $\Phi(B_{t_{i-1}}, B_b - B_{t_i})$ is independent of $B_{t_i} - B_{t_{i-1}}$. Hence,

$$\mathbb{E}\left[\Phi(B_{t_{i-1}}, B_b - B_{t_i})\left(B_{t_i} - B_{t_{i-1}}\right)\right] = \mathbb{E}\left[\Phi(B_{t_{i-1}}, B_b - B_{t_i})\right]\mathbb{E}\left[B_{t_i} - B_{t_{i-1}}\right] = 0. \tag{3.3}$$

where the last equality is due to the zero mean of the Brownian interval. We combine the results of Equation 3.2 and Equation 3.3 to complete the proof. □

3.3.2. Extension of Itô's Isometry (Variance)

Suppose f and ϕ are continuous functions on $\mathbb{R}$. It is proved in Theorem 3.1 of [18] that

$$\begin{aligned}\mathbb{E}\left[\left(\int_a^b f(B_t)\phi(B_b - B_t)\ dB_t\right)^2\right] &= \int_a^b \mathbb{E}\left[f(B_t)^2\phi(B_b - B_t)^2\right]\ dt \\ &+ 2\int_a^b \int_a^t \mathbb{E}\left[f(B_s)\phi'(B_b - B_s)f'(B_t)\phi(B_b - B_t)\right]\ ds\ dt.\end{aligned} \tag{3.4}$$

The result is motivated by Theorem 13.16 in the book [17] where it is shown via white noise methods. The proof in [18] is lengthy, imposes stronger conditions on the integrand, and involves tedious computations via the binomial expansion. More importantly,

it doesn't showcase the crucial feature of the Ayed–Kuo theory of stochastic integration - the left end point and the right endpoint evaluation for the adapted and instantly independent processes. In what follows, we provide intrinsic proof of the formula in Equation (3.4) for a more general case.

In the proof of Theorem 3.3.6 and Theorem 3.4.2 below, we shall use the σ-fields

$$\begin{aligned}
\mathcal{F}_s &= \sigma\{B_u;\ a \le u \le s\}, \qquad a \le s \le b, \\
\mathcal{G}^{(t)} &= \sigma\{B_b - B_v;\ t \le v \le b\}, \qquad a \le t \le b, \text{ and} \\
\mathcal{H}_s^{(t)} &= \sigma\left(\mathcal{F}_s \cup \mathcal{G}^{(t)}\right), \qquad a \le s \le t \le b.
\end{aligned}$$

Here $\{\mathcal{F}_s : s \in [a,b]\}$ is the *natural Brownian* filtration and $\{\mathcal{G}^{(t)} : t \in [a,b]\}$ the *natural Brownian counter*-filtration generated by the Brownian motion. We shall call $\mathcal{H}_s^{(t)}$ as the *separation* σ-field. We conjecture on the importance of the joint σ-field $\mathcal{H}_s^{(t)}$. Taking conditional expectation judiciously with respect to the separation σ-field plays a significant role in the proof the following theorem.

Theorem 3.3.6 (Theorem 3.1 of [20])**.** *Suppose* $f, \phi \in C^1(\mathbb{R})$ *such that* $f(B_t)\phi(B_b - B_t), f(B_t)\phi'(B_b - B_t), f'(B_t)\phi(B_b - B_t) \in L^2([a,b] \times \Omega)$. *Then*

$$\begin{aligned}
\mathbb{E}\left[\left(\int_a^b f(B_t)\phi(B_b - B_t)\ dB_t\right)^2\right] &= \int_a^b \mathbb{E}\left[f(B_t)^2\phi(B_b - B_t)^2\right]\ dt \\
&+ 2\int_a^b \int_a^t \mathbb{E}\left[f(B_s)\phi'(B_b - B_s)f'(B_t)\phi(B_b - B_t)\right]\ ds\ dt. \qquad (3.5)
\end{aligned}$$

Remark 3.3.7 (Remark 3.2 of [20])**.** *For the right-hand side of* (3.5) *to be well-defined, we need the well-definedness of the two integrals. For the first integral, we directly see that the integral is well-defined if* $f(B_t)\phi(B_b - B_t) \in L^2([a,b] \times \Omega)$. *For conciseness, we write* $f_t = f(B_t)$, $\phi_t = \phi(B_b - B_t)$, *and similarly their corresponding derivatives. Using this*

notation, for the second integral, we can use Cauchy–Schwarz inequality to get

$$\begin{aligned}
&\int_a^b \int_a^t \mathbb{E}\left[f_s\phi_s' f_t'\phi_t\right] ds\ dt \\
&\le \int_a^b \int_a^t \left(\mathbb{E}\left[|f_s\phi_s'|^2\right]\right)^{\frac{1}{2}} \left(\mathbb{E}\left[|f_t'\phi_t|^2\right]\right)^{\frac{1}{2}} ds\ dt \\
&\le \int_a^b \left(\mathbb{E}\left[|f_s\phi_s'|^2\right]\right)^{\frac{1}{2}} ds \int_a^b \left(\mathbb{E}\left[|f_t'\phi_t|^2\right]\right)^{\frac{1}{2}} dt \\
&\le (b-a)\left(\int_a^b \mathbb{E}\left[|f_s\phi_s'|^2\right] ds\right)^{\frac{1}{2}} \left(\int_a^b \mathbb{E}\left[|f_t'\phi_t|^2\right] dt\right)^{\frac{1}{2}},
\end{aligned}$$

where we used the Schwarz's inequality in the last step.

Combining these results, we see that a sufficient condition for the second integral to exist is $f(B_t)\phi(B_b - B_t), f(B_t)\phi'(B_b - B_t), f'(B_t)\phi(B_b - B_t) \in L^2([a,b] \times \Omega)$.

Remark 3.3.8 (Remark 3.3 of [20])**.** *In the proof of Itô's isometry, one typically takes conditional expectation with respect to the* σ*-field* $\mathcal{F}_s$ *in a simple manner. On the other hand, our proof requires conditioning with respect to the* σ*-field* $\mathcal{H}_s^{(t)}$ *in a very specific manner.*

Proof. For notational convenience, let

$$\begin{aligned}
\Delta B_k &= B_{t_k} - B_{t_{k-1}}, \\
\Delta t_k &= t_k - t_{k-1}, \\
f_{k-1} &= f(B_{t_{k-1}}), \\
\phi_k &= \phi(B_b - B_{t_k}).
\end{aligned}$$

Then by the definition of the Ayed–Kuo stochastic integral, we get

$$\int_a^b f(B_t)\phi(B_b - B_t)\ dB_t = \lim_{\|\Delta_n\|\to 0} \sum_{i=1}^n f_{i-1}\phi_i \Delta B_i.$$

By taking a subsequence, if necessary, we may assume that the convergence is in $L^2(\Omega)$. Therefore,

$$
\begin{aligned}
&\mathbb{E}\left[\left(\int_a^b f(B_t)\phi(B_b - B_t)\ dB_t\right)^2\right] \\
&= \lim_{\|\Delta_n\|\to 0} \sum_{i=1}^{n}\sum_{j=1}^{n} \mathbb{E}\left[f_{i-1}\phi_i f_{j-1}\phi_j \Delta B_i \Delta B_j\right] \\
&= \lim_{\|\Delta_n\|\to 0} \sum_{i=1}^{n} \mathbb{E}\left[f_{i-1}^2\phi_i^2\left(\Delta B_i\right)^2\right] + 2\lim_{\|\Delta_n\|\to 0} \sum_{j=1}^{n}\sum_{i=1}^{j-1} \mathbb{E}\left[f_{i-1}\phi_i f_{j-1}\phi_j \Delta B_i \Delta B_j\right] \\
&=: D_0 + 2D_1,
\end{aligned}
$$

where we separated the sum into diagonal and off-diagonal elements in the penultimate step and used the symmetry of $i < j$ and $i > j$.

First we focus on the diagonal elements. Note that ΔB_i is independent of both $\mathcal{F}_{t_{i-1}}$ and $\mathcal{G}^{(t_i)}$. Moreover, f_{i-1} is $\mathcal{F}_{t_{i-1}}$-measurable and independent of $\mathcal{G}^{(t_i)}$. Similarly ϕ_i is $\mathcal{G}^{(t_i)}$-measurable and independent of $\mathcal{F}_{t_{i-1}}$. Therefore, by taking conditional expectation with respect to $\mathcal{F}_{t_{i-1}}$, we get

$$
\begin{aligned}
\mathbb{E}\left[f_{i-1}^2\phi_i^2\left(\Delta B_i\right)^2\right] &= \mathbb{E}\left[\mathbb{E}\left(f_{i-1}^2\phi_i^2\left(\Delta B_i\right)^2 \mid \mathcal{F}_{t_{i-1}}\right)\right] \\
&= \mathbb{E}\left[f_{i-1}^2\ \mathbb{E}\left(\phi_i^2\left(\Delta B_i\right)^2 \mid \mathcal{F}_{t_{i-1}}\right)\right] \\
&= \mathbb{E}\left[f_{i-1}^2\right]\mathbb{E}\left[\phi_i^2\left(\Delta B_i\right)^2\right].
\end{aligned}
$$

Similarly, taking conditional expectation with respect to $\mathcal{G}^{(t_i)}$ gives us

$$
\begin{aligned}
\mathbb{E}\left[\phi_i^2\left(\Delta B_i\right)^2\right] &= \mathbb{E}\left[\mathbb{E}\left(\phi_i^2\left(\Delta B_i\right)^2 \mid \mathcal{G}^{(t_i)}\right)\right] \\
&= \mathbb{E}\left[\phi_i^2\ \mathbb{E}\left(\left(\Delta B_i\right)^2 \mid \mathcal{G}^{(t_i)}\right)\right] \\
&= \mathbb{E}\left[\phi_i^2\right]\mathbb{E}\left[\left(\Delta B_i\right)^2\right].
\end{aligned}
$$

Putting it all together along with the fact that $\mathbb{E}\left[(\Delta B_i)^2\right] = \Delta t_i$, we get

$$\mathbb{E}\left[f_{i-1}^2 \phi_i^2 (\Delta B_i)^2\right] = \mathbb{E}\left[f_{i-1}^2\right] \mathbb{E}\left[\phi_i^2\right] \Delta t_i = \mathbb{E}\left[f(B_t)^2 \phi(B_b - B_t)^2\right] \Delta t_i,$$

where we used the independence of increments of Brownian motion in the last equality.

Summing over i and taking limits, we get

$$D_0 = \int_a^b \mathbb{E}\left[f(B_t)^2 \phi(B_b - B_t)^2\right] \, dt.$$

The method for the off-diagonal elements is not so direct, and we highlight the key tricks.

Trick 1 Note that ΔB_i is independent of both $\mathcal{F}_{t_{i-1}}$ and $\mathcal{G}^{(t_i)}$, and is therefore independent of $\mathcal{H}_{t_{i-1}}^{(t_i)}$. So conditioning with respect to $\mathcal{H}_{t_{i-1}}^{(t_i)}$ gives us

$$\begin{aligned} \mathbb{E}\left(\Delta B_i \mid \mathcal{H}_{t_{i-1}}^{(t_i)}\right) &= \mathbb{E}\left[\Delta B_i\right] = 0, \\ \mathbb{E}\left((\Delta B_i)^2 \mid \mathcal{H}_{t_{i-1}}^{(t_i)}\right) &= \mathbb{E}\left[(\Delta B_i)^2\right] = \Delta t_i. \end{aligned}$$

Trick 2 Consider $B_b - B_{t_i} - \Delta B_j = (B_b - B_{t_j}) + (B_{t_{j-1}} - B_{t_i})$. Since $B_b - B_{t_j}$ is $\mathcal{G}^{(t_j)}$-measurable and $B_{t_{j-1}} - B_{t_i}$ is $\mathcal{F}_{t_{j-1}}$-measurable, the sum $B_b - B_{t_i} - \Delta B_j$ is $\mathcal{H}_{t_{j-1}}^{(t_j)}$-measurable. By continuity of ϕ, we see that $\phi(B_b - B_{t_i} - \Delta B_j)$ is also $\mathcal{H}_{t_{j-1}}^{(t_j)}$-measurable. This allows us to conclude that

$$\begin{aligned} &\mathbb{E}\left[f_{i-1}\phi(B_b - B_{t_i} - \Delta B_j) f_{j-1}\phi_j \; \Delta B_i \; \Delta B_j\right] \\ &= \mathbb{E}\left[\mathbb{E}\left(f_{i-1}\phi(B_b - B_{t_i} - \Delta B_j) f_{j-1}\phi_j \; \Delta B_i \; \Delta B_j \mid \mathcal{H}_{t_{j-1}}^{(t_j)}\right)\right] \\ &= \mathbb{E}\left[f_{i-1}\phi(B_b - B_{t_i} - \Delta B_j) f_{j-1}\phi_j \; \Delta B_i \; \mathbb{E}\left(\Delta B_j \mid \mathcal{H}_{t_{j-1}}^{(t_j)}\right)\right] \\ &= 0. \end{aligned} \tag{3.6}$$

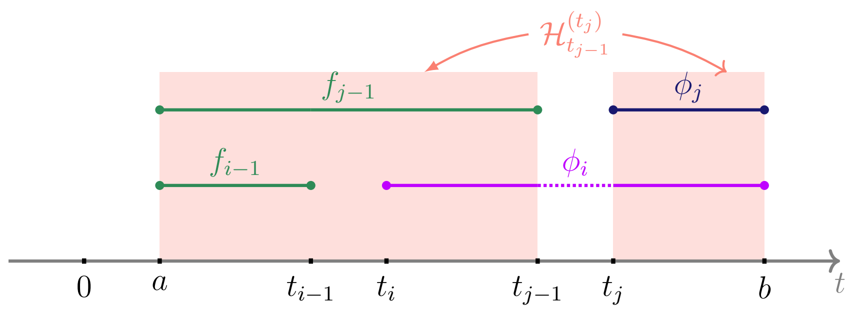

Figure 3.2. A t-dependence plot of the various processes. The dotted regions are removed. Shaded regions represent the separation σ-field.

Therefore, subtracting $\mathbb{E}\left[f_{i-1}\phi(B_b - B_{t_i} - \Delta B_j)f_{j-1}\phi_j\ \Delta B_i\ \Delta B_j\right]$ from the term $\mathbb{E}\left[f_{i-1}\phi_i f_{j-1}\phi_j \Delta B_i \Delta B_j\right]$ does not change anything. This allows us to remove the dependence of ϕ_i on $\{B_t : t \in (t_{j-1}, t_j)\}$. This is illustrated in Figure 3.2 by the purple dotted region of ϕ_i.

Trick 3 Using the assumption $\phi \in C^1(\mathbb{R})$ and considering the fact that B_t is continuous and so $\Delta B_j \to 0$ as $\|\Delta_n\| \to 0$, we can approximate

$$\phi(B_b - B_{t_i}) - \phi(B_b - B_{t_i} - \Delta B_j) \simeq \phi'(B_b - B_{t_i} - \Delta B_j)\ \Delta B_j.$$

For brevity, we write $\Phi_{ij} = \phi'(B_b - B_{t_i} - \Delta B_j)$, $i < j$. Note that Φ_{ij} is $\mathcal{H}_{t_{i-1}}^{(t_i)}$-measurable.

Putting these together, we see that

$$\begin{aligned}
&\mathbb{E}\left[f_{i-1}\phi_i f_{j-1}\phi_j\ \Delta B_i\ \Delta B_j\right] \\
&= \mathbb{E}\left[f_{i-1}\left(\phi(B_b - B_{t_i}) - \phi(B_b - B_{t_i} - \Delta B_j)\right) f_{j-1}\phi_j\ \Delta B_i\ \Delta B_j\right] \\
&\simeq \mathbb{E}\left[f_{i-1}\Phi_{ij} f_{j-1}\phi_j\ \Delta B_i\ (\Delta B_j)^2\right]
\end{aligned}$$

Conditioning with respect to the seperation σ - field, we have

$$
\begin{aligned}
&\mathbb{E}\left[f_{i-1}\phi_i f_{j-1}\phi_j\ \Delta B_i\ \Delta B_j\right] \\
&= \mathbb{E}\left[\mathbb{E}\left(f_{i-1}\Phi_{ij} f_{j-1}\phi_j\ \Delta B_i\ (\Delta B_j)^2 \mid \mathcal{H}^{(t_j)}_{t_{j-1}}\right)\right] \\
&= \mathbb{E}\left[f_{i-1}\Phi_{ij} f_{j-1}\phi_j\ \Delta B_i\ \mathbb{E}(\Delta B_j)^2\right] \\
&= \mathbb{E}\left[f_{i-1}\Phi_{ij} f_{j-1}\phi_j\ \Delta B_i\right] \Delta t_j. \qquad (3.7)
\end{aligned}
$$

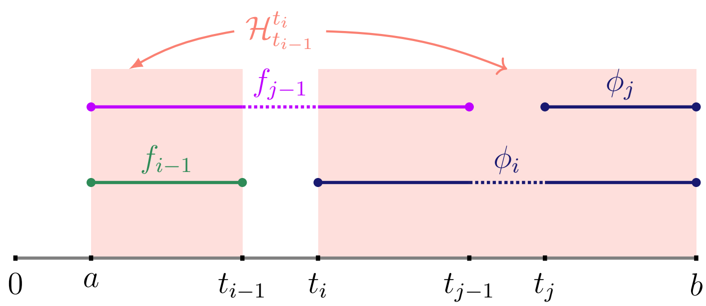

Figure 3.3. A t-dependence plot of the various processes. The dotted regions are removed. Shaded regions represent the separation σ-field.

We repeat Trick 2 on $f(B_{t_{j-1}} - \Delta B_i)$ just as we did for $\phi(B_b - B_{t_i} - \Delta B_j)$ to derive (3.6). This allows us to remove the dependence of f_{j-1} on $\{B_t : t \in (t_{i-1}, t_i)\}$. This is illustrated in Figure 3.3 by the purple dotted region of f_{j-1}. Therefore,

$$= \mathbb{E}\left[f_{i-1}\Phi_{ij} f_{j-1}\phi_j\ \Delta B_i\ \Delta B_j\right] = 0,$$

where we used the tower property with respect to the σ-field $\mathcal{H}^{(t_i)}_{t_{i-1}}$ in this case. As before, we get

$$f(B_{t_{j-1}}) - f(B_{t_{j-1}} - \Delta B_i) \simeq f'(B_{t_{j-1}} - \Delta B_i)\ \Delta B_i.$$

Continuing from (3.7),

$$
\begin{aligned}
&\mathbb{E}\left[f_{i-1}\phi_i f_{j-1}\phi_j\ \Delta B_i\ \Delta B_j\right] \\
&= \mathbb{E}\left[f_{i-1}\Phi_{ij} f_{j-1}\phi_j\ \Delta B_i\right]\Delta t_j \\
&= \mathbb{E}\left[f_{i-1}\Phi_{ij}\left(f(B_{t_{j-1}}) - f(B_{t_{j-1}} - \Delta B_i)\right)\phi_j\ \Delta B_i\right]\Delta t_j \\
&\simeq \mathbb{E}\left[f_{i-1}\Phi_{ij} f'(B_{t_{j-1}} - \Delta B_i)\phi_j\ (\Delta B_i)^2\right]\Delta t_j \\
&= \mathbb{E}\left[\mathbb{E}\left(f_{i-1}\Phi_{ij} f'(B_{t_{j-1}} - \Delta B_i)\phi_j\ (\Delta B_i)^2 \mid \mathcal{H}^{(t_i)}_{t_{i-1}}\right)\right]\Delta t_j \\
&= \mathbb{E}\left[f_{i-1}\Phi_{ij} f'(B_{t_{j-1}} - \Delta B_i)\phi_j \mathbb{E}\left((\Delta B_i)^2 \mid \mathcal{H}^{(t_i)}_{t_{i-1}}\right)\right]\Delta t_j \\
&= \mathbb{E}\left[f_{i-1}\Phi_{ij} f'(B_{t_{j-1}} - \Delta B_i)\phi_j\right]\ \Delta t_i\ \Delta t_j. \qquad (3.8)
\end{aligned}
$$

By the continuity of B_t, we see that as $\|\Delta_n\| \to 0$, so does ΔB_i and ΔB_j. Moreover, by the continuity of f' and ϕ', we can conclude that as $\|\Delta_n\| \to 0$,

$$
\begin{aligned}
f'(B_{t_{i-1}} - \Delta B_i) &\longrightarrow f'(B_{t_{j-1}}) = f'_{j-1}, \\
\Phi_{ij} = \phi'(B_b - B_{t_i} - \Delta B_j) &\longrightarrow \phi'(B_b - B_{t_i}) = \phi'_i.
\end{aligned}
$$

Finally, summing up (3.8) over $i < j$ and taking limit, we get

$$
D_1 = \int_a^b \int_a^t \mathbb{E}\left[f(B_s)\phi'(B_b - B_s) f'(B_t)\phi(B_b - B_t)\right]\ ds\ dt.
$$

This concludes the proof. □

Theorem 3.3.6 also serves as a new tool in evaluating the second moment of anticipating integrals. This is advantageous as explicitly evaluating the integral via the definition can get very tedious and complicated. We demonstrate this point of view with an example.

Example 3.3.9 (Example 3.4 of [20]). *Apply Theorem 3.3.6 to the case with $f(x) = x$ and $\phi(y) = y$. Then we have*

$$
\begin{aligned}
&\mathbb{E}\left[\left(\int_a^b B_t(B_b - B_t)\ dB_t\right)^2\right] \\
&= \int_a^b \mathbb{E}\left[B_t^2(B_b - B_t)^2\right]\ dt + 2\int_a^b \int_a^t \mathbb{E}\left[B_s(B_b - B_t)\right]\ ds\ dt \\
&= \int_a^b \mathbb{E}\left[B_t^2\right]\mathbb{E}\left[(B_b - B_t)^2\right]\ dt + 2\int_a^b \int_a^t \mathbb{E}\left[B_s\right]\mathbb{E}\left[B_b - B_t\right]\ ds\ dt \\
&= \int_a^b t(b-t)\ dt \\
&= \frac{1}{6}(b^3 - 3a^2 b + 2a^3).
\end{aligned} \tag{3.9}
$$

On the other hand, let us evaluate the stochastic integral $\int_a^b B_t(B_b - B_t)\ dB_t$ and then use it to compute its second moment. By equation (2.5) in [9], we have

$$
\int_0^t B_s(B_T - B_s)\, dB_s = \frac{1}{2}B_T(B_t^2 - t) - \frac{1}{3}B_t^3, \quad 0 \le t \le T,
$$

which immediately yields the following stochastic integral

$$
\int_a^b B_t(B_b - B_t)\ dB_t = \frac{1}{2}B_b\Big[(B_b^2 - B_a^2) - (b-a)\Big] - \frac{1}{3}\left(B_b^3 - B_a^3\right).
$$

For brevity, we write $\Delta_B = B_b - B_a$, so $B_b = (B_b - B_a) + B_a = \Delta_B + B_a$. Performing algebraic simplification, we get

$$
\int_a^b B_t(B_b - B_t)\ dB_t = \frac{1}{6}\left(\Delta_B^3 + 3B_a\Delta_B^2 - 3(b-a)\Delta_B - 3(b-a)B_a\right).
$$

Note that B_a and Δ_B are independent with $B_a \sim N(0,a)$ and $\Delta_B \sim N(0, b-a)$. Therefore,

any odd moment of either of B_a or Δ_B is zero. Using this, we get

$$\begin{aligned}
&\mathbb{E}\left[\left(\int_a^b B_t(B_b - B_t)\ dB_t\right)^2\right] \\
&= \frac{1}{36}\ \mathbb{E}\left[\Delta_B^6 + 9B_a^2\Delta_B^4 + 9(b-a)^2\Delta_B^2 + 9(b-a)^2 B_a^2\right. \\
&\qquad\qquad \left. -6(b-a)\Delta_B^4 - 18(b-a)B_a^2\Delta_B^2\right] \\
&= \frac{1}{6}\left(b^3 - 3a^2 b + 2a^3\right),
\end{aligned}$$

which is exactly what we obtained in equation (3.9)*. But, obviously, here the computation is more tedious and complicated.*

The arguments used in the proof of Theorem 3.3.6 can also be applied to show the following general results.

Theorem 3.3.10 (Theorem 3.5 of [20])**.** *Let $\Phi(x,y) \in C^1(\mathbb{R}^2)$ and assume that*

$$\Phi(B_t, B_b - B_t), \Phi_x(B_t, B_b - B_t),\ \Phi_y(B_t, B_b - B_t) \in L^2([a,b] \times \Omega).$$

Then

$$\begin{aligned}
\mathbb{E}\left[\left(\int_a^b \Phi(B_t, B_b - B_t)\ dB_t\right)^2\right] &= \int_a^b \mathbb{E}\left[\Phi(B_t, B_b - B_t)^2\right]\ dt \\
&+ 2\int_a^b \int_a^t \mathbb{E}\Big[\Phi_y(B_s, B_b - B_s)\ \Phi_x(B_t, B_b - B_t)\Big]\ ds\ dt. \qquad (3.10)
\end{aligned}$$

We use this general result to obtain the covariance between two Ayed–Kuo stochastic integrals.

Theorem 3.3.11 (Theorem 3.6 of [20])**.** *Let $\Phi(x,y), \Psi(x,y) \in C^1(\mathbb{R}^2)$ and assume that*

1. $\Phi(B_t, B_b - B_t), \Phi_x(B_t, B_b - B_t),\ \Phi_y(B_t, B_b - B_t) \in L^2([a,b] \times \Omega)$*, and*
2. $\Psi(B_t, B_b - B_t), \Psi_x(B_t, B_b - B_t),\ \Psi_y(B_t, B_b - B_t) \in L^2([a,b] \times \Omega)$*.*

Then

$$\begin{aligned}
&\mathbb{E}\left[\left(\int_a^b \Phi(B_t, B_b - B_t)\ dB_t\right)\left(\int_a^b \Psi(B_t, B_b - B_t)\ dB_t\right)\right] \\
&= \int_a^b \mathbb{E}\left[\Phi(B_t, B_b - B_t)\Psi(B_t, B_b - B_t)\right]\ dt \\
&\quad + \int_a^b \int_a^t \mathbb{E}\Big[\Phi_y(B_s, B_b - B_s)\Psi_x(B_t, B_b - B_t) \\
&\qquad\qquad + \Phi_x(B_t, B_b - B_t)\Psi_y(B_t, B_b - B_t)\Big]\ ds\ dt.
\end{aligned}$$

Proof. For this proof, we write

$$F(t) = \Phi(B_t, B_b - B_t),$$

$$G(t) = \Psi(B_t, B_b - B_t),$$

$$H(t) = F(t) + G(t).$$

Moreover, for brevity, we write $F_x(t) = \Phi_x(B_t, B_b - B_t)$, $F_y(t) = \Phi_y(B_t, B_b - B_t)$ and corresponding notations for $G(t)$ and $H(t)$.

From the definition of $H(t)$, we see that

$$\begin{aligned}
\mathbb{E}\left[\left(\int_a^b H(t)\ dB_t\right)^2\right] &= \mathbb{E}\left[\left(\int_a^b F(t)\ dB_t + \int_a^b G(t)\ dB_t\right)^2\right] \\
&= \mathbb{E}\left[\left(\int_a^b F(t)\ dB_t\right)^2\right] + \mathbb{E}\left[\left(\int_a^b G(t)\ dB_t\right)^2\right] \\
&\quad + 2\ \mathbb{E}\left[\left(\int_a^b F(t)\ dB_t\right)\left(\int_a^b G(t)\ dB_t\right)\right].
\end{aligned}$$

Applying Theorem 3.3.10 for $F(t)$, we get

$$\mathbb{E}\left[\left(\int_a^b F(t)\ dB_t\right)^2\right] = \int_a^b \mathbb{E}\left[F(t)^2\right]\ dt + 2\int_a^b \int_a^t \mathbb{E}\left[F_y(s)\ F_x(t)\right]\ ds\ dt.$$

We can obtain a similar equality for $G(t)$. Putting all this together, we get

$$\begin{aligned}\mathbb{E}\left[\left(\int_a^b H(t)\ dB_t\right)^2\right] &= \int_a^b \mathbb{E}\left[F(t)^2\right]\ dt + 2\int_a^b\int_a^t \mathbb{E}\left[F_y(s)\ F_x(t)\right]\ ds\ dt \\ &\quad + \int_a^b \mathbb{E}\left[G(t)^2\right]\ dt + 2\int_a^b\int_a^t \mathbb{E}\left[G_y(s)\ G_x(t)\right]\ ds\ dt \\ &\quad + 2\ \mathbb{E}\left[\left(\int_a^b F(t)\ dB_t\right)\left(\int_a^b G(t)\ dB_t\right)\right]. \end{aligned} \tag{3.11}$$

On the other hand, first applying Theorem 3.3.10 and then using the definition of $H(t)$, we get

$$\begin{aligned}&\mathbb{E}\left[\left(\int_a^b H(t)\ dB_t\right)^2\right] \\ &= \int_a^b \mathbb{E}\left[H(t)^2\right] dt + 2\int_a^b\int_a^t \mathbb{E}\left[H_y(s)\ H_x(t)\right]\ ds\ dt \\ &= \int_a^b \mathbb{E}\left[F(t)^2\right] dt + \int_a^b \mathbb{E}\left[G(t)^2\right] dt + 2\int_a^b \mathbb{E}\left[F(t)G(t)\right] dt \\ &\quad + 2\int_a^b\int_a^t \mathbb{E}\left[(F_y(s)+G_y(s))(F_x(t)+G_x(t))\right]\ ds\ dt \\ &= \int_a^b \mathbb{E}\left[F(t)^2\right] dt + \int_a^b \mathbb{E}\left[G(t)^2\right] dt + 2\int_a^b \mathbb{E}\left[F(t)G(t)\right] dt \\ &\quad + 2\int_a^b\int_a^t \mathbb{E}\Big[F_y(s)F_x(t) + F_y(s)G_x(t) + G_y(s)F_x(t) + G_y(s)G_x(t)\Big]\ ds\ dt. \end{aligned} \tag{3.12}$$

Finally, equations (3.11) and (3.12) imply that

$$\begin{aligned}&\mathbb{E}\left[\left(\int_a^b F(t)\ dB_t\right)\left(\int_a^b G(t)\ dB_t\right)\right] \\ &= \int_a^b \mathbb{E}\left[F(t)G(t)\right]\ dt + \int_a^b\int_a^t \mathbb{E}\left[F_y(s)G_x(t) + G_y(s)F_x(t)\right]\ ds\ dt,\end{aligned}$$

which is exactly the desired result. □

If $\Phi(x,y) = f(x)$ and $\Psi(x,y) = \phi(y)$, we have $\Phi_y \equiv 0$ and $\Psi_x \equiv 0$. Therefore, we obtain the following corollary.

Corollary 3.3.12 (Corollary 3.7 of [20]). *Let $f, \phi \in C^1(\mathbb{R})$ and assume that*

1. $f(B_t), \phi(B_b - B_t) \in L^2([a,b] \times \Omega)$, *and*
2. $f'(B_t), \phi'(B_b - B_t) \in L^2([a,b] \times \Omega)$.

Then

$$\begin{aligned}&\mathbb{E}\left[\left(\int_a^b f(B_t)\ dB_t\right)\left(\int_a^b \phi(B_b - B_t)\ dB_t\right)\right] \\ &= \int_a^b \mathbb{E}\left[f(B_t)\phi(B_b - B_t)\right]\ dt + \int_a^b \int_a^t \mathbb{E}\left[\phi'(B_b - B_s) f'(B_t)\right] ds\ dt.\end{aligned}$$

Corollary 3.3.12 provides the same power as the extension of the isometry in that it allows us to explicitly calculate the covariance between anticipating and adapted integrals without having to calculate the integral itself. We demonstrate that fact with an example

Example 3.3.13 (Example 3.8 of [20]). *Let $f(x) = x$ and $\phi(y) = y$. Using Corollary 3.3.12, we get*

$$\begin{aligned}&\mathbb{E}\left[\left(\int_a^b B_t\ dB_t\right)\left(\int_a^b (B_b - B_t)\ dB_t\right)\right] \\ &= \int_a^b \mathbb{E}\left[B_t(B_b - B_t)\right]\ dt + \int_a^b \int_a^t \mathbb{E}\left[1\right] ds\ dt \\ &= \int_a^b \mathbb{E}\left[B_t\right] \mathbb{E}\left[B_b - B_t\right]\ dt + \int_a^b (t-a)\ dt \\ &= \frac{1}{2}(b-a)^2.\end{aligned}$$

Finally, we want to point our that the double integral in equation (3.10) can be regarded as a correction term when we extend Itô's theory to anticipating stochastic integration. This correction term can be positive or negative, as illustrated in the next example.

Example 3.3.14 (Example 3.9 of [20]). *Consider the case $\Phi(x,y) = px + y$ in Theorem 3.3.10, where $p \in \mathbb{R}$. Then $\Phi_x = p$ and $\Phi_y = 1$. Therefore, we can directly evaluate the*

double integral in equation (3.10) *as*

$$2\int_a^b \int_a^t \mathbb{E}\Big[\Phi_y(B_s, B_b - B_s)\ \Phi_x(B_t, B_b - B_t)\Big]\ ds\ dt = 2\int_a^b \int_a^t p\ ds\ dt = p(b-a)^2.$$

Therefore, the final term will be positive or negative depending on the sign of p.

3.4. Near-martingales

Let us consider the following stochastic process $X_t = \int_0^t B_1\, dB_s,\ t \in [0,1]$. Using the definition of the Ayed–Kuo stochastic integral in 3.3.2 and following the steps in example 3.3.1 we have

$$X_t = \int_0^t B_1\, dB_s = B_1 B_t - t, \quad t \in [0,1].$$

It is easy to see that X_t is not a martingale as B_1 is not adapted to the natural Brownian filtration $\{\mathcal{F}_t\}$ for $t < 1$. Let $s \le t$ for $s, t \in [0,1]$. We evaluate the conditional expectation of X_t and X_s with respect to $\mathcal{F}_s$. Hence

$$\begin{aligned}\mathbb{E}\left[X_s \,|\, \mathcal{F}_s\right] &= \mathbb{E}\left[B_1 B_s - s \,|\, \mathcal{F}_s\right] \\ &= B_s \mathbb{E}\left[B_1 \,|\, \mathcal{F}_s\right] - s \\ &= B_s^2 - s.\end{aligned}$$

$$
\begin{aligned}
\mathbb{E}\left[X_t \,|\, \mathcal{F}_s\right] &= \mathbb{E}\left[B_1 B_t - t \,|\, \mathcal{F}_s\right] \\
&= \mathbb{E}\left[(B_1 - B_t + B_t - B_s + B_s)(B_t - B_s + B_s) \,|\, \mathcal{F}_s\right] - t \\
&= \mathbb{E}\Bigg[(B_1 - B_t)(B_t - B_s) + (B_t - B_s)^2 \\
&\qquad\quad + 2(B_t - B_s) B_s + (B_1 - B_t) B_s + B_s^2 \,|\, \mathcal{F}_s \Bigg] - t \\
&= \mathbb{E}\left[(B_1 - B_t)(B_t - B_s)\right] + \mathbb{E}\left[(B_t - B_s)^2\right] \\
&\qquad\quad + B_s \mathbb{E}\left[B_1 - B_s\right] + 2 B_s \mathbb{E}\left[B_t - B_s\right] + B_s^2 - t \\
&= B_s^2 - s,
\end{aligned}
$$

where the last equality is due to the properties of Brownian motion. Notice that $\mathbb{E}\left[X_s \,|\, \mathcal{F}_s\right] = \mathbb{E}\left[X_t \,|\, \mathcal{F}_s\right]$. This property is satisfied by many other stochastic processes. In view of Definition 1.3.6, a martingale would trivially satisfy this relationship. In essence, such a relationship can be seen as a generalization of the martingale property. This discussion serves as the motivation for *near-martingales* defined below.

Definition 3.4.1 (Near-martingale)**.** *A stochastic process, N_t, $a \leq t \leq b$, with $\mathbb{E}|\, N_t \,| < \infty$ for all t is called a near-martingale with respect to the filtration$\{\mathcal{F}_t\}$ if for any $a \leq s \leq t \leq b$, $\mathbb{E}[\, N_t - N_s \,|\, \mathcal{F}_s \,] = 0$, a.s.*

From Theorem 1.4.6, we know that martingales occur naturally within Itô theory. Near-martingales and Ayed–Kuo integrals are true generalization of both of martingales and Itô integrals respectively. It is natural to ask, should we expect something similar? In what follows, we show that near-martingales do occur naturally within the Ayed–Kuo theory.

Theorem 3.4.2 (Theorem 3.3 of [21]). *Let $\Phi(x,y)$ be a function that is continuous in both variables such that the stochastic integral,*

$$N_t = \int_a^t \Phi(B_s, B_b - B_s)\, dB_s, \qquad a \leq t \leq b,$$

exists and $\mathbb{E}|\, N_t \,| < \infty$ for each t in $[a,b]$. Furthermore, assume that the partial sums

$$\sum_{i=1}^{n} \Phi(B_{t_{i-1}}, B_b - B_{t_i})\left(B_{t_i} - B_{t_{i-1}}\right)$$

are uniformly integrable. Then N_t, $a \leq t \leq b$, is a near-martingale with respect to the filtration generated by Brownian motion given by $\{\mathcal{F}_t\}$.

Remark 3.4.3. *This result is shown in Theorem 3.5 of [18] for the case when $\Phi(x,y) = f(x)\phi(y)$.*

Proof. Let $t > s$ and consider a partition, Δ_n, of $[s,t]$ with $t_0 = s$ and $t_n = t$. Via the definition of the Ayed–Kuo Stochastic Integral in conjunction with the uniform integrability condition on the partial sums, we have

$$\begin{aligned}
\mathbb{E}\left[N_t - N_s \,|\, \mathcal{F}_s\right] =& \mathbb{E}\left[\int_s^t \Phi(B_v, B_b - B_v)\, dB_v \,|\, \mathcal{F}_s\right] \\
=& \mathbb{E}\left[\lim_{n\to\infty} \sum_{k=1}^{n} \Phi(B_{k-1}, B_b - B_k)\Delta B_k \,|\, \mathcal{F}_s\right] \\
=& \lim_{n\to\infty} \sum_{k=1}^{n} \mathbb{E}\left[\Phi(B_{k-1}, B_b - B_k)\Delta B_k \,|\, \mathcal{F}_s\right].
\end{aligned} \tag{3.13}$$

Consider, the separation σ-field $\mathcal{H}_a^{(b)} = \sigma(\mathcal{F}_a \cup \mathcal{G}^{(b)})$. Then $\mathcal{F}_s \subseteq \mathcal{F}_{k-1} \subseteq \mathcal{H}_{k-1}^{(k)}$. Using this fact alongside the continuity of Φ in both variables, we have that $\Phi(B_{k-1}, B_b - B_k)$ is $\mathcal{H}_{k-1}^{(k)}$ - measurable. Furthermore, via the independence of the Brownian increments, ΔB_k is independent of $\mathcal{H}_{k-1}^{(k)}$.

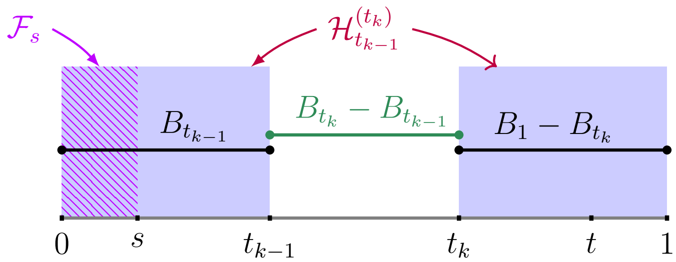

Figure 3.4. A t-dependence plot of the disjoint increments of $B_\bullet$. The shaded regions represents the forward and separation σ-field.

Thus,

$$\begin{aligned}
&\mathbb{E}\left[\Phi(B_{k-1}, B_b - B_k)\Delta B_k \,|\, \mathcal{F}_s\right] \\
=&\mathbb{E}\left[\mathbb{E}\left[\Phi(B_{k-1}, B_b - B_k)\Delta B_k \,|\, \mathcal{H}_{k-1}^{(k)}\right] \,|\, \mathcal{F}_s\right] \\
=&\mathbb{E}\left[\Phi(B_{k-1}, B_b - B_k)\mathbb{E}\left[\Delta B_k\right] \,|\, \mathcal{F}_s\right] \\
=&0.
\end{aligned}$$

Using this result for each k in equation (3.13), we have

$$\mathbb{E}\left[N_t - N_s \,|\, \mathcal{F}_s\right] = 0.$$

Thus, N_t is a near-martingale. □

Example 3.4.4. *Consider $N_t = \int_a^t B_b \, dB_s$. Using Theorem 3.4.2 with $\Phi(x, y) = x + y$, we have that N_t is a near-martingale for $a \le t \le b$.*

This theorem can be extended for the anticipating case as well. Namely,

Theorem 3.4.5. *Let $\Phi(x, y)$ be a function that is continuous in both variables such that the stochastic integral,*

$$N_t = \int_t^b \Phi(B_s, B_b - B_s) \, dB_s, \qquad a \le t \le b,$$

exists and $\mathbb{E}|\, N_t \,| < \infty$ *for each* t *in* $[a, b]$. *Then* N_t, $a \le t \le b$, *is a near-martingale with respect to the filtration generated by Brownian motion given by* $\{\mathcal{F}_t\}$.

Remark 3.4.6. *This result is shown in Theorem 3.6 of [18] for the case when* $\Phi(x, y) = f(x)\phi(y)$.

These two results, Theorem 3.4.2 and Theorem 3.4.5, show that the near-martingale property is an analogue of martingale property for the Ayed–Kuo stochastic integral. This relation is further solidified by the following result that shows the intrinsic relation between a near-martingale and a martingale.

Theorem 3.4.7 (Theorem 2.11 of [11])**.** *Let* N_t, $a \le t \le b$ *be a stochastic process with* $\mathbb{E}|\, N_t \,| < \infty$ *for each* $t \in [a, b]$ *and let* $M_t = \mathbb{E}\,[N_t \,|\, \mathcal{F}_t]$. *Then*

$$N_t \text{ is a near-martingale} \iff M_t \text{ is a martingale.}$$

We use this result in obtaining a near-martingale optional stopping theorem for the Ayed–Kuo stochastic integral.

3.4.1. Near-martingale Optional Stopping Theorem

We build up the optional stopping theorem using the strategy used for the martingale case in [32]. Namely, we prove near-martingale version of the optional stopping theorem for discrete stopping times and then extend it for continuous stopping times. For this section, fix $t \in [0, 1]$.

First we prove a result that shows that a stopped near-martingale is a near-martingale. We use this result to prove the discrete case of the theorem.

Theorem 3.4.8 (Theorem 3.6 of [21])**.** *Let* N_t, $t \in [0, 1]$, *be a discrete-time near-martingale and* τ *a stopping time. Then* $N_{\tau\wedge t}$ *is a near-martingale.*

Proof. Let t_n be an arbitrary time and consider a partition $0 = t_0 < t_1 < \cdots < t_n$ and a forward filtration $\mathcal{F}_\bullet$ associated with N_t. We define

$$Y_n \triangleq \sum_{i=1}^{n} \mathbb{1}_{\{t_{i-1} \leq \tau\}} \left(N_{t_i} - N_{t_{i-1}}\right) = N_{\tau \wedge t_n} - N_0. \tag{3.14}$$

Assume $m < n$, then

$$\begin{aligned}\mathbb{E}\left[Y_n - Y_m \,|\, \mathcal{F}_m\right] &= \mathbb{E}\left[\sum_{i=m+1}^{n} \mathbb{1}_{\{t_{i-1} \leq \tau\}} \left(N_{t_i} - N_{t_{i-1}}\right) \,|\, \mathcal{F}_m\right] \\ &= \sum_{i=m+1}^{n} \mathbb{E}\left[\mathbb{1}_{\{t_{i-1} \leq \tau\}} \left(N_{t_i} - N_{t_{i-1}}\right) \,|\, \mathcal{F}_m\right].\end{aligned}$$

Here $t_i > t_m$. In addition, $\mathbb{1}_{\{t_i \leq \tau\}}$ is F_i- measurable for each i. As such, taking conditional expectation judiciously with respect to $\mathcal{F}_{i-1}$ we get,

$$\begin{aligned}\mathbb{E}\left[Y_n - Y_m \,|\, \mathcal{F}_m\right] &= \sum_{i=m+1}^{n} \mathbb{E}\left[\mathbb{E}\left[\mathbb{1}_{\{t_{i-1} \leq \tau\}} \left(N_{t_i} - N_{t_{i-1}}\right) \,|\, \mathcal{F}_{i-1}\right] \,|\, \mathcal{F}_m\right] \\ &= \sum_{i=m+1}^{n} \mathbb{E}\left[\mathbb{1}_{\{t_{i-1} \leq \tau\}} \mathbb{E}\left[N_{t_i} - N_{t_{i-1}} \,|\, \mathcal{F}_{i-1}\right] \,|\, \mathcal{F}_m\right] \\ &= 0.\end{aligned}$$

where the last equality is due to the near-martingale property of N_t. Thus, Y_n is a near-martingale. We use this fact alongside equation (3.14) to get

$$\mathbb{E}\left[N_{\tau \wedge t_n} - N_{\tau \wedge t_m} \,|\, \mathcal{F}_m\right] = \mathbb{E}\left[Y_n - Y_m \,|\, \mathcal{F}_m\right] = 0.$$

□

Thus we have shown that a stopped near-martingale is a near-martingale. We use this result to prove a version of the near-martingale optional stopping theorem for discrete time near-martingales.

Theorem 3.4.9 (Theorem 3.7 of [21]). *Let N_t be a discrete time near-martingale. Suppose σ and τ are two bounded stopping times with $\sigma \leq \tau$. Then N_σ and N_τ are integrable, and $\mathbb{E}\left[N_\tau - N_\sigma \,|\, \mathcal{F}_\sigma\right] = 0$ almost surely.*

Proof. Since σ and τ are bounded, there exists $K < \infty$ such that $\sigma \leq \tau \leq K$. Let $X_\bullet$ be any near-martingale. Clearly, X_σ is integrable. Suppose $B \in \mathcal{F}_\sigma$. Then for any $n \leq K$, we have $B \cap \{\sigma = n\} \in \mathcal{F}_n$, and so

$$\int_{B\cap\{\sigma=n\}} (X_K - X_\sigma)\, d\mathbb{P} \;=\; \int_{B\cap\{\sigma=n\}} (X_K - X_n)\, d\mathbb{P} = 0.$$

Summing over n, we get $\int_B (X_K - X_\sigma)\, d\mathbb{P} = 0$, and so $\mathbb{E}\left[X_K - X_\sigma \,|\, \mathcal{F}_\sigma\right] = 0$. Finally, let $X_n = N_{\tau\wedge n}$ to get

$$\mathbb{E}\left[N_{\tau\wedge K} - N_{\tau\wedge\sigma} \;\,|\, \mathcal{F}_\sigma\right] = \mathbb{E}\left[N_\tau - N_\sigma \;\,|\, \mathcal{F}_\sigma\right] = 0.$$

□

Before we proceed to continuous time near-martingales, we will need the concept of backward near-martingales.

Definition 3.4.10 (Definition 3.8 of [21]). *Let $(\mathcal{F}_n)_{n=1}^\infty$ be a decreasing sequence of σ-algebras, and let $N_\bullet = (N_n)_{n=1}^\infty$ be a stochastic process. Then the pair $(N_n, \mathcal{F}_n)_{n=1}^\infty$ is called a* backward near-martingale *if for every n,*

1. *N_n is integrable and $\mathcal{F}_n$-measurable, and*
2. *$\mathbb{E}\left[N_n - N_{n+1} \,|\, \mathcal{F}_{n+1}\right] = 0$.*

Lemma 3.4.11 (Lemma 3.9 of [21]). *Let $(N_n, \mathcal{F}_n)_{n=1}^\infty$ be a backward near-martingale with $\lim_{n\to\infty} \mathbb{E}\left[N_n\right] > -\infty$. If $N_\bullet$ is non-negative for every n, then $N_\bullet$ is uniformly integrable.*

Proof. As $n \nearrow \infty$, we have $\mathbb{E}\left[N_n\right] \searrow \lim_{n\to\infty} \mathbb{E}\left[N_n\right] = \inf_n \mathbb{E}\left[N_n\right] > -\infty$. Fix $\epsilon > 0$.

By the definition of infimum, there exists a $K > 0$ such that for any $n \geq K$, we have $\mathbb{E}[N_k] - \lim_{n\to\infty} \mathbb{E}[N_n] < \epsilon$.

For any $k \geq n$ and $\delta > 0$, we have

$$\mathbb{E}\left[|N_k| \mathbb{1}_{\{|N_k|>\delta\}}\right] = \mathbb{E}\left[N_k \mathbb{1}_{\{N_k>\delta\}}\right] + \mathbb{E}\left[N_k \mathbb{1}_{\{N_k\geq -\delta\}}\right] - \mathbb{E}[N_k].$$

Moreover, since $N_\bullet$ is a backward near-martingale, $\mathbb{E}\left[N_k \mathbb{1}_{\{N_k\geq\delta\}}\right] = \mathbb{E}\left[N_n \mathbb{1}_{\{N_k\geq\delta\}}\right]$. Therefore,

$$\begin{aligned}\mathbb{E}\left[|N_k| \mathbb{1}_{\{|N_k|>\delta\}}\right] &\leq \mathbb{E}\left[N_n \mathbb{1}_{\{N_k>\delta\}}\right] + \mathbb{E}\left[N_n \mathbb{1}_{\{N_k\geq -\delta\}}\right] - (\mathbb{E}[N_n] - \epsilon) \\ &= \mathbb{E}\left[|N_n| \mathbb{1}_{\{|N_k|>\delta\}}\right] + \epsilon.\end{aligned}$$

By Markov's inequality and the non-negativity of X,

$$\mathbb{P}\{|N_k| > \delta\} \leq \frac{1}{\delta}\mathbb{E}|N_k| = \frac{1}{\delta}\mathbb{E}[N_k] \leq \frac{1}{\delta}\mathbb{E}[N_1] \to 0$$

as $\delta \to \infty$. This concludes the proof. $\square$

We use these backward near-martingales and Lemma 3.4.11 to extend the near-martingale optional stopping theorem to the continuous case. Namely,

Theorem 3.4.12 (Theorem 3.10 of [21])**.** *Let N be a near-martingale with right-continuous sample paths. Suppose σ and τ are two bounded stopping times with $\sigma \leq \tau$. If N is non-negative or uniformly integrable, then N_σ and N_τ are integrable, and*

$$\mathbb{E}[N_\tau - N_\sigma \mid \mathcal{F}_\sigma] = 0 \textit{ almost surely.}$$

Proof. We use a discretization argument to prove the result. Let $T > 0$ be a bound for τ. For every $n \in \mathbb{N}$, define the discretization function

$$f_n : [0,\infty) \to \left\{\frac{k}{n} : k = 0, \ldots, n\right\} : x \mapsto \frac{\lfloor 2^n x\rfloor + 1}{2^n} \wedge T,$$

and let $\sigma_n = f_n(\sigma)$ and $\tau_n = f_n(\tau)$.

Now, for any n and t,

$$\{\tau_n \leq t\} = \{f_n(\tau) \in [0,t]\} = \left\{\tau \in f_n^{-1}[0,t]\right\} = \left\{\tau \in f_n^{-1}\left[0, \frac{\lfloor 2^n t \rfloor}{2^n}\right]\right\} \in \mathcal{F}_{\frac{\lfloor 2^n t \rfloor}{2^n}} \subseteq \mathcal{F}_t,$$

so τ_n is a stopping time. Similarly, σ_n is a stopping time. Moreover, it can be easily seen that $\sigma_n \leq \tau_n$ for every n, and $\sigma_n \searrow \sigma$ and $\tau_n \searrow \tau$ as $n \nearrow \infty$. Therefore, by Theorem 3.4.9, we get N_{σ_n} and N_{τ_n} are integrable, and $\mathbb{E}\left[N_{\tau_n} - N_{\sigma_n} \,|\, \mathcal{F}_{\sigma_n}\right] = 0$ almost surely. Furthermore, it is easy to see that $\mathcal{F}_\sigma = \bigcap_{n=1}^{\infty} \mathcal{F}_{\sigma_n} \subseteq \mathcal{F}_{\sigma_n}$ for any n. Therefore, $\mathbb{E}\left[N_{\tau_n} - N_{\sigma_n} \,|\, \mathcal{F}_\sigma\right] = 0$ almost surely for any n.

If N is non-negative, by construction, $(N_{\sigma_n}, \mathcal{F}_{\sigma_n})_{n=1}^{\infty}$ is a backward near-martingale such that $N_{\sigma_n} \geq 0$ for every n. Therefore, $\mathbb{E}\left[N_{\sigma_n}\right] \searrow \mathbb{E}\left[N_\sigma\right] > -\infty$ as $n \nearrow \infty$. Using Lemma 3.4.11, $(N_{\sigma_n})_{n=1}^{\infty}$ is uniformly integrable. Similarly, $(N_{\tau_n})_{n=1}^{\infty}$ is also uniformly integrable. On the other hand, if N is uniformly integrable, this is trivial.

Using the right continuity of N and the boundedness assumption of σ and τ, we get $\lim_{n\to\infty} N_{\sigma_n} = N_\sigma$ and $\lim_{n\to\infty} N_{\tau_n} = N_\tau$ almost surely. Furthermore, the uniform integrability of $(N_{\sigma_n})_{n=1}^{\infty}$ and $(N_{\tau_n})_{n=1}^{\infty}$ allows us to conclude that N_σ and N_τ are integrable and that the convergence is also in L^1, giving us $\mathbb{E}\left[N_\tau - N_\sigma \,|\, \mathcal{F}_\sigma\right] = 0$ almost surely. □

We highlight the special case of Theorem 3.4.12.

Corollary 3.4.13 (Corollary 3.11 of [21]). *Let $N_\bullet$ be a non-negative near-martingale with right-continuous sample paths and τ a bounded stopping time. Then N_τ is integrable, and $\mathbb{E}\left[N_\tau\right] = \mathbb{E}\left(N_0\right)$ almost surely.*

Chapter 4. General Itô Formula and Applications of the General Ito integral

The Itô formula is a fundamental aspect of Itô theory. For any theory to extend Itô theory, it is also very useful to formulate the Itô formula in that setting. There are several such extensions in literature [17] [25]. The Ayed–Kuo stochastic integral provides a simplistic and intuitive extension that accounts for both instantly independent and adapted processes. We use the general Itô formula for the results in this chapter. Let $t \in [a,b]$. First, let X_t and $Y^{(t)}$ be stochastic processes of the form

$$X_t = X_a + \int_a^t g(s)\, dB(s) + \int_a^t h(s)\, ds, \tag{4.1}$$

$$Y^{(t)} = Y^{(b)} + \int_t^b \xi(s)\, dB(s) + \int_t^b \eta(s)\, ds, \tag{4.2}$$

where $g(t), h(t)$ are adapted (so X_t is an Itô process), and $\xi(t), \eta(t)$ are instantly independent such that $Y^{(t)}$ is also instantly independent. Then we have the following *general Itô formula*

Theorem 4.0.1 (Theorem 3.2 of [10]). *Suppose* $\left\{X_t^{(i)}\right\}_{i=1}^n$ *and* $\left\{Y_j^{(t)}\right\}_{j=1}^m$ *are stochastic processes of the form given by equations* (4.1) *and* (4.2)*, respectively. Suppose* $\theta(t, x_1, \ldots, x_n, y_1, \ldots, y_m)$ *is a real-valued function that is* C^1 *in* t *and* C^2 *in other variables. Then the stochastic differential of* $\theta(t, X_t^{(1)}, \ldots, X_t^{(n)}, Y_1^{(t)}, \ldots, Y_m^{(t)})$ *is given*

Sections 4.1 - 4.3 of this chapter previously appeared, substantially unchanged, in the open access *Journal of Stochasic Analysis* article: Kuo, H.-H., Shrestha, P., and Sinha, S., “Anticipating Linear Stochastic Differential Equations with Adapted Coefficients”, `https://digitalcommons.lsu.edu/josa/vol2/iss2/5/`.

by

$$
\begin{aligned}
&d\theta(t, X_t^{(1)}, \ldots, X_t^{(n)}, Y_1^{(t)}, \ldots, Y_m^{(t)}) \\
&= \theta_t\, dt + \sum_{i=1}^{n} \theta_{x_i} dX_t^{(i)} + \sum_{j=1}^{m} \theta_{y_j} dY_j^{(t)} \\
&\qquad + \frac{1}{2} \sum_{i,k=1}^{n} \theta_{x_i x_k} dX_t^{(i)} dX_t^{(k)} - \frac{1}{2} \sum_{j,l=1}^{m} \theta_{y_j y_l} dY_j^{(t)} dY_l^{(t)}.
\end{aligned}
$$

Corollary 4.0.2 (Corollary 2.11 of [9])**.** *Suppose X_t is an Itô process and $\psi(t, x, y)$ is a real-valued function that is C^1 in in t and C^2 in x and y. Then the stochastic differential of $\theta(t, X_t, B(b))$ is given by*

$$
\theta(t, X_t, B(b)) = \theta_t\, dt + \theta_x\, dX_t + \frac{1}{2}\theta_{xx}(dX_t)^2 + \theta_{xy}\, dX_t\, dY^{(t)}.
$$

Now, we come to an important class of processes that occur ubiquitously in solutions of stochastic differential equations.

Definition 4.0.3. *The* exponential process *associated with adapted stochastic processes $\alpha(t)$ and $\beta(t)$ is defined as*

$$
\mathcal{E}_{\alpha,\beta}(t) = \exp\left[\int_a^t \alpha(s)\, dB(s) + \int_a^t \left(\beta(s) - \frac{1}{2}\alpha(s)^2\right) ds\right]. \tag{4.3}
$$

If $\beta \equiv 0$, then we write

$$
\mathcal{E}_{\alpha}(t) = \exp\left[\int_a^t \alpha(s)\, dB(s) - \frac{1}{2}\int_a^t \alpha(s)^2\, ds\right].
$$

Remark 4.0.4. *The exponential process $\mathcal{E}_{\alpha,\beta}(t)$ is an Itô process satisfying the stochastic differential equation*

$$
\begin{cases}
d\mathcal{E}_{\alpha,\beta}(t) = \alpha(t)\mathcal{E}_{\alpha,\beta}(t)\, dB(t) + \beta(t)\mathcal{E}_{\alpha,\beta}(t)\, dt, \quad t \in [a, b], \\
\mathcal{E}_{\alpha,\beta}(a) = 1.
\end{cases} \tag{4.4}
$$

Similarly, the exponential process $\mathcal{E}_\alpha(t)$ is an Itô process satisfying the stochastic differential equation

$$\begin{cases} d\mathcal{E}_\alpha(t) = \alpha(t)\mathcal{E}_\alpha(t)\, dB(t), \quad t \in [a,b], \\ \mathcal{E}_\alpha(a) = 1. \end{cases}$$

The proof of the result follows from a direct application of Itô's formula.

4.1. Motivating Examples of Anticipating SDEs

We begin with some examples to show the non-trivial nature of the extension of the stochastic integral regardless of the origin of the anticipation. These serve as motivations for our main results. In this section, we fix $t \in [0,1]$.

4.1.1. Anticipation due to coefficients

We progressively increase the complexity of the diffusion coefficient of the stochastic differential equation and observe how it affects the solution. In the first case, we take the diffusion coefficient to be a constant.

Example 4.1.1 (Example 3.1 of [19]). *Let α be a constant. The process*

$$\mathcal{E}_\alpha(t) = \exp\left[\alpha B(t) - \frac{1}{2}\alpha^2 t\right], \quad t \in [0,1]$$

is a solution of the stochastic differential equation

$$\begin{cases} d\mathcal{E}_\alpha(t) = \alpha\mathcal{E}_\alpha(t)dB(t), \quad t \in [0,1], \\ \mathcal{E}_\alpha(0) = 1. \end{cases}$$

We proceed to upgrade our diffusion coefficient to be a deterministic function.

Example 4.1.2 (Example 3.2 of [19]). *Suppose $\alpha(t)$ is a deterministic function. The process*

$$\mathcal{E}_\alpha(t) = \exp\left[\int_0^t \alpha(s)\, dB(s) - \frac{1}{2}\int_0^t \alpha(s)^2\, ds\right], \quad t \in [0,1]$$

is a solution of the stochastic differential equation

$$\begin{cases} d\mathcal{E}_\alpha(t) = \alpha(t)\mathcal{E}_\alpha(t)dB(t), \quad t \in [0,1], \\ \\ \mathcal{E}_\alpha(0) = 1. \end{cases}$$

We now consider the case then f is adapted.

Example 4.1.3 (Example 3.3 of [19]). *Consider the adapted coefficient $\alpha(t) = B(t)$. The process*

$$X(t) = \exp\left[\frac{1}{2}\left(B(t)^2 - t - \int_0^t B(s)^2\, ds\right)\right], \quad t \in [0,1]$$

is a solution of the stochastic differential equation

$$\begin{cases} dX(t) = B(t)X(t)\, dB(t), \quad t \in [0,1], \\ \\ X(0) = 1. \end{cases}$$

From the example 3.3.1 and the examples presented above in this section, a reasonable guess is that

$$\begin{aligned} Z(t) &= \exp\left[\int_0^t B(1)\, dB(s) - \frac{1}{2}\int_0^t B(1)^2\, ds\right] \\ &= \exp\left[B(1)B(t) - t - \frac{1}{2}B(1)^2 t\right], \qquad t \in [0,1] \end{aligned}$$

is a solution of the stochastic differential equation

$$\begin{cases} dZ(t) = B(1)Z(t)\, dB(t), \quad t \in [0,1], \\ \\ Z(0) = 1. \end{cases}$$

However, this is false. Using the general Itô formula, we have the following result.

Theorem 4.1.4 (Theorem 3.3 of [9]). *The stochastic process*

$$Z(t) = \exp\left[B(1)B(t) - t - \frac{1}{2}B(1)^2 t\right]$$

is a solution of

$$\begin{cases} dZ(t) = B(1)Z(t)\, dB(t) + B(1)(B(t) - tB(1))Z(t)dt, \quad t \in [0,1], \\ Z(0) = 1. \end{cases}$$

We are then left with a question. What is the solution of the stochastic differential equation given by

$$\begin{cases} dZ(t) = B(1)Z(t)\, dB(t), \quad t \in [0,1], \\ Z(0) = 1. \end{cases}$$

We have the following result.

Theorem 4.1.5 (Theorem 3.1 of [9]). *The process*

$$Z(t) = \exp\left[B(1)\int_0^t e^{-(t-s)}\, dB(s) - \frac{1}{4}(B(1))^2(1-e^{-2t}) - t\right], \quad t \in [0,1]$$

is a solution of the stochastic differential equation

$$\begin{cases} dZ(t) = B(1)Z(t)\, dB(t), \quad t \in [0,1], \\ Z(0) = 1. \end{cases}$$

The above examples demonstrate the non-trivial nature of anticipating coefficients.

4.1.2. Anticipation due to initial condition

Consider the following stochastic differential equations with anticipating initial conditions.

Example 4.1.6 (Examples 4.1-3 of [1]).

$$\begin{cases} dX(t) = X(t)dB(t), & t \in [0,1], \\ X(0) = x, & x \in \mathbb{R}. \end{cases} \tag{4.5}$$

It is well known that the solution to (4.5) *is*

$$X(t) = xe^{B(t)-\frac{1}{2}t}$$

However, if we take this solution and replace x with $B(1)$ and apply the general Itô formula to the resultant expression, we obtain a different stochastic differential equation. In particular,

$$Y(t) = B(1)e^{B(t)-\frac{1}{2}t} \tag{4.6}$$

is a solution of

$$\begin{cases} dY(t) = Y(t)dB(t) + \dfrac{1}{B(1)}Y(t)dt, & t \in [0,1], \\ Y(0) = B(1). \end{cases} \tag{4.7}$$

Here, the initial condition is outside the classical theory of Itô calculus since B(1) is not $\mathcal{F}_0$-measurable. We can use the general Itô formula along with the Picard iteration method to show that $Y(t)$ is indeed the unique solution.

On the other hand, if we replace all the $B(1)$ terms in (4.7) with $x \in \mathbb{R}$ then we obtain the following stochastic differential equation

$$\begin{cases} dZ(t) = Z(t)dB(t) + \dfrac{1}{x}Z(t)dt, & t \in [0,1], \\ Z(0) = x, & x \in \mathbb{R}. \end{cases}$$

with its solution

$$Z(t) = xe^{B(t)-\frac{1}{2}t+\frac{1}{x}t}$$

In comparing the two equations (4.5) and (4.7), we see that when we replace the non-anticipating term with an anticipating term, an extra term appears in the drift term of the stochastic differential equation. Furthermore, by replacing all the anticipating terms in (4.7) with a real number, we obtained an extra drift factor in (4.6). Through these examples, we can observe interesting patterns with linear stochastic differential equations with respect to anticipation.

Example 4.1.7 (Section 3 of [14]). *Consider the following motivational example:*

$$\begin{cases} dX(t) = X(t)dB(t), \quad t \in [0,1], \\ X(0) = B(1). \end{cases}$$

Equation (4.5) *would suggest that our solution would be* (4.6)*. However, that is not the case. We have an extra drift term as demonstrated by* (4.7)*. With that in mind, we "guess" that the solution has the form*

$$X(t) = (B(1) - \xi(t))\, e^{B(t) - \frac{1}{2}t}$$

with ξ being a deterministic function that needs to be determined. Via a simple application of the general Itô formula to the function $\theta(t,x,y) = (y - \xi(t))e^{x - \frac{1}{2}t}$, we get that

$$dX(t) = (B(1) - \xi(t))e^{B(t)-\frac{1}{2}t}\, dB(t) + \left[e^{B(t)-\frac{1}{2}t} - \xi'(t)e^{B(t)-\frac{1}{2}t}\right] dt.$$

The dt term in the above equation must be zero for $X(t)$ to be a solution. Therefore, by solving the following ordinary differential equation

$$\begin{cases} \xi'(t) = 1\,, \quad t \in [0,1], \\ \xi(0) = 0, \end{cases}$$

we get our solution

$$X(t) = (B(1) - t)e^{B(t)-\frac{1}{2}t}.$$

We use this example as inspiration for the following theorem that provides solutions for a class of stochastic differential equations with anticipating initial conditions.

Theorem 4.1.8 (Theorem 5.1 of [22]). *Let $\alpha(t)$, $h(t) \in L^2[0,1]$, $\beta(t) \in L^1[0,1]$. Assume that $\psi(t)$ is a C^2 function. Then the unique solution of the stochastic differential equation*

$$\begin{cases} dX(t) = \alpha(t)X(t)dB(t) + \beta(t)X(t)\, dt, \quad t \in [0,1], \\ X(0) = \psi\left(\int_0^1 h(s)dB(s)\right), \end{cases}$$

is given by the equation

$$X(t) = \psi\left(\int_0^1 h(s)dB(s) - \int_0^t \alpha(s)h(s)ds\right)\mathcal{E}_{\alpha,\beta}(t),$$

where $\mathcal{E}_{\alpha,\beta}(t)$ *is the stochastic process defined in equation* (4.3)*.*

Remark 4.1.9 (Remark 3.9 of [19])**.** *In Theorem 4.1 of [14], the authors proved a similar result for the particular case where* $h(t) \equiv 1$ *and* ψ *is a function on* $\mathbb{R}$ *having power series expansion at* $t = 0$ *with infinite radius of convergence. In Theorem 4.1.8 and in [14],* $\alpha(t)$ *is assumed to be deterministic.*

Example 4.1.10 (Example 5.2 of [22])**.** *Consider the stochastic differential equation*

$$\begin{cases} dX(t) = X(t)dB(t), \quad t \in [0,1], \\ X(0) = \displaystyle\int_0^1 B(s)ds. \end{cases}$$

We can use stochastic integration by parts and the results of Theorem 4.1.8 to obtain the solution,

$$X(t) = \left(\int_0^1 B(s)ds - (t - \frac{1}{2}t^2)\right)e^{B(t)-\frac{1}{2}t}.$$

Thus we have solutions for a class of linear stochastic differential equations with deterministic coefficients.

4.2. Anticipating Stochastic Differential Equations

In Theorem 4.1.8 , we had assumed that $\alpha(t) \in L^2[0,1]$ and $\beta(t) \in L^1[0,1]$. In the following theorem, we generalize that condition to allow both $\alpha(t)$ and $\beta(t)$ to be adapted to the filtration generated by the Brownian motion.

Hypothesis 4.2.1. *Assume that* $\alpha(t)$*,* $\beta(t)$ *and* $h(t)$*, where* $t \in [a,b]$*, have the following properties:*

1. $\alpha(t)$ *is an adapted process with* $\mathbb{E}\left(\int_a^b |\alpha(t)|^2 \, dt\right) < \infty$,

2. $\beta(t)$ *is an adapted process with* $\mathbb{E}\left(\int_a^b |\beta(t)| \, dt\right) < \infty$,

3. $h(t) \in L^2[a,b]$ *is a deterministic function.*

Theorem 4.2.2 (Theorem 4.2 of [19])**.** *Let* $\alpha(t)$, $\beta(t)$, *and* $h(t)$ *satisfy Hypothesis 4.2.1, and* $\psi \in C^2(\mathbb{R})$. *Then the solution of the stochastic differential equation*

$$\begin{cases} dZ(t) = \alpha(t)Z(t)\, dB(t) + \beta(t)Z(t)\, dt, \quad t \in [a,b], \\ Z(0) = \psi\Big(\displaystyle\int_a^b h(s)\, dB(s)\Big), \end{cases} \tag{4.8}$$

is given by

$$Z(t) = \psi\Big(\int_a^b h(s)\, dB(s) - \int_a^t h(s)\alpha(s)\, ds\Big)\mathcal{E}_{\alpha,\beta}(t).$$

Proof. Suppose $Z(t) = \psi\big(\int_a^b h(s)\, dB(s) - Q(t)\big)\mathcal{E}_{\alpha,\beta}(t)$. We need to determine the Itô process $Q(t)$ with $Q(a) = 0$. In order to apply the general Itô formula, we write

$$Z(t) = \psi\Big(\int_a^t h(s)\, dB(s) - Q(t) + \int_t^b h(s)\, dB(s)\Big)\mathcal{E}_{\alpha,\beta}(t).$$

We define the instantly independent process $Y^{(t)} = \int_t^b h(s)\, dB(s)$ and the following adapted processes

$$X_t^{(1)} = \mathcal{E}_{\alpha,\beta}(t), \quad X_t^{(2)} = \int_a^t h(s)\, dB(s) - Q(t).$$

From the definitions of $X_t^{(1)}$, $X_t^{(2)}$, and $Y^{(t)}$ above, we get the differentials

$$\begin{aligned}
dX_t^{(1)} &= \alpha(t)X_t^{(1)}\,dB(t) + \beta(t)X_t^{(1)}\,dt,\\
dX_t^{(2)} &= h(t)\,dB(t) - dQ(t),\\
(dX_t^{(1)})^2 &= \alpha(t)^2(X_t^{(1)})^2\,dt,\\
(dX_t^{(2)})^2 &= h(t)^2\,dt - 2h(t)\,dB(t)dQ(t) + (dQ(t))^2,\\
dX_t^{(1)}dX_t^{(2)} &= h(t)\alpha(t)X_t^{(1)}\,dt - \alpha(t)X_t^{(1)}\,dB(t)dQ(t),\\
dY^{(t)} &= -h(t)\,dB(t),\\
(dY^{(t)})^2 &= h(t)^2\,dt.
\end{aligned}$$

Now, define $\theta(x_1, x_2, y) = \psi(x_2 + y)x_1$, so that $Z(t) = \theta\left(X_t^{(1)}, X_t^{(2)}, Y^{(t)}\right)$. From this, we get the partial derivatives

$$\begin{aligned}
\theta_{x_1} &= \psi, & \theta_{x_1x_1} &= 0,\\
\theta_{x_2} &= \psi' x_1, & \theta_{x_2x_2} &= \psi'' x_1,\\
\theta_y &= \psi' x_1, & \theta_{x_1x_2} &= \psi',\\
\theta_{yy} &= \psi'' x_1.
\end{aligned}$$

Applying Theorem 4.0.1 and putting everything together, we can easily find the

stochastic differential of $Z(t)$:

$$
\begin{aligned}
dZ(t) &= d\theta(X_t^{(1)}, X_t^{(2)}, Y^{(t)}) \\
&= \theta_{x_1} dX_t^{(1)} + \theta_{x_2} dX_t^{(2)} \\
&\quad + \frac{1}{2}\theta_{x_1x_1}(dX_t^{(1)})^2 + \frac{1}{2}\theta_{x_2x_2}(dX_t^{(2)})^2 \\
&\quad + \theta_{x_1x_2}(dX_t^{(1)})(dX_t^{(2)}) \\
&\quad + \theta_y\, dY^{(t)} - \frac{1}{2}\theta_{yy}(dY^{(t)})^2 \\
&= \psi \cdot \Big(\alpha(t)X_t^{(1)}\, dB(t) + \beta(t)X_t^{(1)}\, dt\Big) + \psi' \cdot X_t^{(1)}\, (\cancel{h(t)\, dB(t)} - dQ(t)) \\
&\quad + 0 + \frac{1}{2}\psi'' \cdot X_t^{(1)}\, \big(\cancel{h(t)^2\, dt} - 2h(t)\, dB(t)dQ(t) + (dQ(t))^2\big) \\
&\quad + \psi' \Big(h(t)\alpha(t)X_t^{(1)}\, dt - \alpha(t)X_t^{(1)}\, dB(t)dQ(t)\Big) \\
&\quad - \cancel{\psi' \cdot X_t^{(1)} h(t)\, dB(t)} - \cancel{\frac{1}{2}\psi'' \cdot X_t^{(1)} h(t)^2\, dt} \\
&= \psi \cdot \Big(\alpha(t)X_t^{(1)}\, dB(t) + \beta(t)X_t^{(1)}\, dt\Big) \\
&\quad + \psi' \cdot X_t^{(1)} \cdot (-dQ(t) + h(t)\alpha(t)\, dt - \alpha(t)\, dB(t)dQ(t)) \\
&\quad + \frac{1}{2}\psi'' \cdot X_t^{(1)} \cdot \big(-2h(t)\, dB(t)dQ(t) + (dQ(t))^2\big).
\end{aligned}
$$

Therefore, in order for $Z(t)$ to be the solution of equation (4.8), we need to satisfy the following conditions

$$dQ(t) = h(t)\alpha(t)\, dt - \alpha(t)\, dB(t)dQ(t) \tag{4.9}$$

$$(dQ(t))^2 = 2h(t)\, dB(t)dQ(t) \tag{4.10}$$

From equation (4.9), we see that if $dQ(t)$ contains only a dt term (no $dB(t)$ term), then $dQ(t)\, dB(t) = 0$. On the other hand, if $dQ(t)$ contains a $dB(t)$ term, then $dQ(t)\, dB(t) = \gamma(t)\, dt$ for some $\gamma(t)$. Then we have $dQ(t) = (h(t) - \gamma(t))\alpha(t)\, dt$, which

again gives $dQ(t)\,dB(t) = 0$. Therefore, in either case, $dQ(t) = h(t)\alpha(t)\,dt$. Note that this also agrees with equation (4.10).

Imposing the initial condition $Q(a) = 0$, we get that $Q(t) = \int_a^t h(t)\alpha(t)\,dt$. Putting this in the assumed form of the solution, we get our result. □

Now we look at a specific case of Theorem 4.2.2 where $h(t) \equiv 1$.

Corollary 4.2.3 (Corollary 4.3 of [19])**.** *Under the same assumptions for $\alpha(t)$, $\beta(t)$ and ψ as in Theorem 4.2.2, the solution of the stochastic differential equation*

$$\begin{cases} dZ(t) = \alpha(t)Z(t)\,dB(t) + \beta(t)Z(t)\,dt, \quad t \in [a,b], \\ Z(0) = \psi\Big(B(b) - B(a)\Big), \end{cases}$$

is given by

$$Z(t) = \psi\Big(B(b) - B(a) - \int_a^t \alpha(s)\,ds\Big)\mathcal{E}_{\alpha,\beta}(t).$$

Remark 4.2.4 (Remark 4.4 of [19])**.** *This corollary extends Theorem 4.1 of [14] to include adapted coefficients for the anticipating stochastic differential equation.*

These new results helps us obtain solutions for stochastic differential equations with anticipating initial conditions and adapted coefficients. In the first example, the diffusion and drift terms are adapted while the anticipation comes from $X(0) = B(1)$. The second example demonstrates a case where the initial condition is a Riemann integral of a Brownian motion.

Example 4.2.5 (Example 4.5 of [19])**.** *Consider the stochastic differential equation*

$$\begin{cases} dX(t) = B(t)X(t)dB(t) + X(t)\,dt, \quad t \in [0,1], \\ X(0) = B(1). \end{cases}$$

Here $\alpha(t) = B(t)$, $\beta(t) \equiv 1, h(t) \equiv 1$, and $\psi(x) = x$. Thus, by Corollary 4.2.3, we have the solution

$$X(t) = \left(B(1) - \int_0^t B(s)\,ds\right) \exp\left[\frac{1}{2}\left(B^2(t) + t - \int_0^t B^2(s)\,ds\right)\right].$$

Example 4.2.6 (Example 4.6 of [19]). *Consider the stochastic differential equation*

$$\begin{cases} dX(t) = B(t)X(t)\,dB(t), \quad t \in [0,1], \\ X(0) = \int_0^1 B(s)\,ds. \end{cases}$$

As in Example 4.1.10, we use stochastic integration by parts to modify the initial condition. Namely,

$$\int_0^1 B(s)\,ds = \int_0^1 (1-s)\,dB(s).$$

Hence with $\alpha(t) = B(t)$, $\beta(t) \equiv 0$, $h(t) = 1 - t$, and $\psi(x) = x$ in Theorem 4.2.2, we have the solution

$$X(t) = \left(\int_0^1 B(s)\,ds - \int_0^t (1-s)B(s)\,ds\right) \exp\left[\frac{1}{2}\left(B(t)^2 - t - \int_0^t B(s)^2\,ds\right)\right].$$

4.3. Conditional Expectation of Solutions of SDEs

Given a stochastic process $Z(t)$, the conditional expectation of the solution is an interesting property to study. By taking the conditional expectation with respect to the natural Brownian filtation, we can project our anticipating stochastic differential equation into the realm of classical Itô theory. Keeping in mind the near-martingale property in Definition 3.4.1, analysis of $X(t) = E\left(Z(t)|\mathcal{F}_t\right)$ provides an avenue to study the anticipatory nature of the stochastic process. As such, we ask which stochastic differential equation would $X(t)$ satisfy? Are $dX(t)$ and $dZ(t)$ the same? or are they different? With that motivation, we show the following result.

Theorem 4.3.1 (Theorem 5.1 of [19]). *Let $\alpha(t)$, $\beta(t)$, and $h(t)$ satisfy Hypothesis 4.2.1, and ψ an analytic function on the reals. Suppose that $Z_1(t)$ and $Z_2(t)$ are the solutions of the linear stochastic differential equations*

$$\begin{cases} dZ_1(t) = \alpha(t)Z_1(t)\,dB(t) + \beta(t)Z_1(t)\,dt, \quad t \in [a,b], \\ Z_1(a) = \psi\Big(\int_a^b h(s)\,dB(s)\Big), \end{cases}$$

and

$$\begin{cases} dZ_2(t) = \alpha(t)Z_2(t)\,dB(t) + \beta(t)Z_2(t)\,dt, \quad t \in [a,b], \\ Z_2(a) = \psi'\Big(\int_a^b h(s)\,dB(s)\Big), \end{cases}$$

respectively. Let $X_1(t) = \mathbb{E}\left(Z_1(t)|\mathcal{F}_t\right)$ and $X_2(t) = \mathbb{E}\left(Z_2(t)|\mathcal{F}_t\right)$. Then $X_1(t)$ satisfies the stochastic differential equation

$$\begin{cases} dX_1(t) = \alpha(t)X_1(t)\,dB(t) + \beta(t)X_1(t)\,dt + h(t)X_2(t)\,dB(t), \quad t \in [a,b], \\ X_1(a) = \mathbb{E}\Big[\psi\Big(\int_a^b h(s)\,dB(s)\Big)\Big]. \end{cases} \tag{4.11}$$

Remark 4.3.2 (Remark 5.2 of [19]). *In Theorem 4.1 of [22], the authors proved a similar result for the special case where α is deterministic, β is adapted, and $h \equiv 1$.*

Proof. By the assumption and Theorem 4.2.2, the solution processes $Z_1(t)$ can be written as

$$\begin{aligned} Z_1(t) &= \mathcal{E}_{\alpha,\beta}(t) \cdot \psi\left(\Big(\int_a^t h(s)\,dB(s) - \int_a^t h(s)\alpha(s)\,ds\Big) + \int_t^b h(s)\,dB(s)\right) \\ &= \mathcal{E}_{\alpha,\beta}(t) \cdot \sum_{k=0}^{\infty} \frac{1}{k!}\psi^{(k)}\Big(\int_a^t h(s)\,dB(s) - \int_a^t h(s)\alpha(s)\,ds\Big)\Big(\int_t^b h(s)\,dB(s)\Big)^k. \end{aligned}$$

For brevity, we henceforth denote

$$\psi_k(t) = \psi^{(k)}\Big(\int_a^t h(s)\,dB(s) - \int_a^t h(s)\alpha(s)\,ds\Big) \tag{4.12}$$

In this notation, the expression for $Z_1(t)$ becomes

$$Z_1(t) = \mathcal{E}_{\alpha,\beta}(t) \cdot \sum_{k=0}^{\infty} \frac{1}{k!} \psi_k(t) \Big(\int_t^b h(s)\, dB(s) \Big)^k.$$

Note that $\mathcal{E}_{\alpha,\beta}(t)$ and $\psi_k(t)$ are adapted for all k. Moreover, since $h(t)$ is deterministic, $\int_t^b h(s)\, dB(s)$ is a Wiener integral, and therefore, $\int_t^b h(s)\, dB(s)$ has the Gaussian distribution with mean 0 and variance

$$Y(t) = \int_t^b h(s)^2\, ds. \tag{4.13}$$

Therefore, for any k, we have $\mathbb{E}\left[\Big(\int_t^b h(s)\, dB(s) \Big)^{2k+1} \right] = 0$ and

$$\mathbb{E}\left[\Big(\int_t^b h(s)\, dB(s) \Big)^{2k} \right] = Y(t)^k (2k-1)!!,$$

where !! denotes the double factorial defined as

$$n!! = \prod_{k=0}^{\lceil \frac{n}{2} \rceil} (n - 2k)$$

for any natural number n.

Moreover, $\int_t^b h(s)\, dB(s)$ is independent of $\mathcal{F}_t$ for every t. Using all of these information, we get

$$\begin{aligned} X_1(t) &= \mathcal{E}_{\alpha,\beta}(t) \cdot \sum_{k=0}^{\infty} \frac{1}{(2k)!} \psi_{2k}(t)\ \mathbb{E}\left[\Big(\int_t^b h(s)\, dB(s) \Big)^{2k} \right] \\ &= \mathcal{E}_{\alpha,\beta}(t) \cdot \sum_{k=0}^{\infty} \frac{1}{(2k)!} \psi_{2k}(t) Y(t)^k (2k-1)!! \\ &= \mathcal{E}_{\alpha,\beta}(t) \cdot \sum_{k=0}^{\infty} \frac{1}{(2k)!!} \psi_{2k}(t) Y(t)^k, \end{aligned} \tag{4.14}$$

and similarly,

$$X_2(t) = \mathcal{E}_{\alpha,\beta}(t) \cdot \sum_{k=0}^{\infty} \frac{1}{(2k)!!} \psi_{2k+1}(t) Y(t)^k, \tag{4.15}$$

Now we look at the differentials. From equations (4.12) and (4.13), we get

$$d\left(Y(t)^k\right) = kY(t)^{k-1}\,(-h(t)^2\,dt),$$

and

$$\begin{aligned} d\psi_{2k}(t) &= \psi_{2k+1}(t)(h(t)\,dB(t) - h(t)\alpha(t)\,dt) + \frac{1}{2}\psi_{2k+2}(t)(h(t)^2\,dt) \\ &= \psi_{2k+1}(t)h(t)\,dB(t) + \left(\frac{1}{2}\psi_{2k+2}(t)h(t)^2 - \psi_{2k+1}(t)h(t)\alpha(t)\right)dt. \end{aligned}$$

Using the expressions for $d\left(Y(t)^k\right)$ and $d\psi_{2k}(t)$, and Remark 4.0.4, we get

$$\begin{aligned} &d\left(\mathcal{E}_{\alpha,\beta}(t)\psi_{2k}(t)Y(t)^k\right) \\ &= \psi_{2k}(t)Y(t)^k\,d\mathcal{E}_{\alpha,\beta}(t) + \mathcal{E}_{\alpha,\beta}(t)Y(t)^k\,d\psi_{2k}(t) + \mathcal{E}_{\alpha,\beta}(t)\psi_{2k}(t)\,dY(t)^k \\ &\quad + \mathcal{E}_{\alpha,\beta}(t)\,d\psi_{2k}(t)\cdot dY(t)^k + \psi_{2k}(t)\,d\mathcal{E}_{\alpha,\beta}(t)\cdot dY(t)^k + Y(t)^k\,d\mathcal{E}_{\alpha,\beta}(t)\cdot d\psi_{2k}(t) \\ &= \psi_{2k}(t)Y(t)^k\Big(\alpha(t)\mathcal{E}_{\alpha,\beta}(t)\,dB(t) + \beta(t)\mathcal{E}_{\alpha,\beta}(t)\,dt\Big) \\ &\quad + \mathcal{E}_{\alpha,\beta}(t)Y(t)^k\left[\psi_{2k+1}(t)h(t)\,dB(t) + \left(\frac{1}{2}\psi_{2k+2}(t)h(t)^2 - \cancel{\psi_{2k+1}(t)h(t)\alpha(t)}\right)dt\right] \\ &\quad + \mathcal{E}_{\alpha,\beta}(t)\psi_{2k}(t)\left(-kY(t)^{k-1}h(t)^2\,dt\right) \\ &\quad + 0 + 0 + Y(t)^k\left(\cancel{\mathcal{E}_{\alpha,\beta}(t)\psi_{2k+1}(t)\alpha(t)h(t)}\right)\,dt \\ &= \mathcal{E}_{\alpha,\beta}(t)Y(t)^k\left(\psi_{2k}(t)\alpha(t) + \psi_{2k+1}(t)h(t)\right)\,dB(t) \\ &\quad + \mathcal{E}_{\alpha,\beta}(t)Y(t)^{k-1}\left(\psi_{2k}(t)Y(t)\beta(t) + \frac{1}{2}\psi_{2k+2}(t)Y(t)h(t)^2 - k\psi_{2k}(t)h(t)^2\right)dt. \end{aligned}$$

At this point, we note that

$$\begin{aligned} \sum_{k=0}^{\infty}\frac{1}{(2k)!!}k\psi_{2k}(t) &= \sum_{k=1}^{\infty}\frac{1}{(2k)(2k-2)!!}k\psi_{2k}(t) \\ &= \frac{1}{2}\sum_{k-1=0}^{\infty}\frac{1}{(2(k-1))!!}\psi_{2(k-1)+2}(t) \\ &= \frac{1}{2}\sum_{k=0}^{\infty}\frac{1}{(2k)!!}\psi_{2k+2}(t). \end{aligned} \tag{4.16}$$

Now, since $X_1(t) = \sum_{k=0}^{\infty} \frac{1}{(2k)!!} \mathcal{E}_{\alpha,\beta}(t)\psi_{2k}(t)Y(t)^k$ (see equation (4.14)), we get

$$
\begin{aligned}
dX_1(t) &= \sum_{k=0}^{\infty} \frac{1}{(2k)!!} d\left(\mathcal{E}_{\alpha,\beta}(t)\psi_{2k}(t)Y(t)^k\right) \\
&= \sum_{k=0}^{\infty} \frac{1}{(2k)!!} \mathcal{E}_{\alpha,\beta}(t)Y(t)^k\psi_{2k}(t)\alpha(t)\, dB(t) \\
&\quad + \sum_{k=0}^{\infty} \frac{1}{(2k)!!} \mathcal{E}_{\alpha,\beta}(t)Y(t)^k\psi_{2k+1}(t)h(t)\, dB(t) \\
&\quad + \sum_{k=0}^{\infty} \frac{1}{(2k)!!} \mathcal{E}_{\alpha,\beta}(t)Y(t)^k\psi_{2k}(t)\beta(t)\, dt \\
&\quad + \cancel{\sum_{k=0}^{\infty} \frac{1}{(2k)!!} \frac{1}{2} \mathcal{E}_{\alpha,\beta}(t)Y(t)^k\psi_{2k+2}(t)h(t)^2\, dt} \\
&\quad - \cancel{\sum_{k=0}^{\infty} \frac{1}{(2k)!!} k\mathcal{E}_{\alpha,\beta}(t)Y(t)^{k-1}\psi_{2k}(t)h(t)^2\, dt} \\
&= \alpha(t)X_1(t)\, dB(t) + h(t)X_2(t)\, dB(t) + \beta(t)X_1(t)\, dB(t),
\end{aligned}
$$

where, in the second step, we used the result of equation (4.16). This completes the proof of the theorem. $\square$

The extra term in the conditional stochastic differential equation in (4.11) presents an interesting feature. We can see that the stochastic differential equation for $X_1(t)$ is defined via $X_2(t)$. However, $X_2(t)$ is defined in equation (4.15) as an infinite series and a closed form is not guaranteed. Similar to how $X_2(t)$ arose from taking the first derivative of ψ as the initial condition, we can use the second derivative of ψ as the initial condition to arrive at $X_3(t)$. We can then repeat the use of Theorem 4.3.1 to provide the link between $dX_2(t)$ to $dX_3(t)$. Suppose ψ was analytic then we could form an infinite chain of relations by using the infinite derivatives of ψ as initial conditions. However, there is no guarantee of a "nice" closed form. Despite this issue, we exploit the fact that the derivative of the exponential function is itself for the following example.

Example 4.3.3 (Example 5.3 of [19]). *Let $\alpha(t)$, $\beta(t)$, and $h(t)$ satisfy Hypothesis 4.2.1, and let $\psi(x) = e^x$. In this case, $\psi \equiv \psi'$, so $Z_1(t) \equiv Z_2(t)$. Consequently, $X_1(t) = X_2(t)$, which we call $X(t)$ for convenience. Then by Theorem 4.3.1,*

$$X(t) = \mathcal{E}_{\alpha,\beta}(t) \exp\Big(\int_a^t h(s)\, dB(s) - \int_a^t h(s)\alpha(s)\, ds \Big), \quad t \in [a, b],$$

and $X(t)$ satisfies the stochastic differential equation

$$\begin{cases} dX(t) = (\alpha(t) + h(t))X(t)\, dB(t) + \beta(t)X(t)\, dt, \\ X(a) = 1. \end{cases}$$

In general, the absence of a closed form is not uncommon and we look at alternate ways to analyse these anticipating stochastic differential equations. Recall that the scaled Hermite polynomials $\left\{ \frac{1}{\sqrt{n!\rho^n}} H_n(x;\rho) \right\}$ form an orthonormal basis for the space $L^2(\mathbb{R}, \gamma)$, where γ is the Gaussian measure with mean 0 and variance ρ. Therefore, if we are able to arrive at a closed form reformulation of Theorem 4.3.1 for Hermite polynomials, we can use this to state the result for conditional expectation of the solution when the initial condition is any $L^2(\mathbb{R}, \gamma)$-function of a Wiener integral. However, before we delve into the derivation, let us review some facts about Hermite polynomials.

The Hermite polynomial of degree n with parameter ρ defined by

$$H_n(x;\rho) = (-\rho)^n e^{\frac{x^2}{2\rho}} D_x^n e^{-\frac{x^2}{2\rho}},$$

where D_x is the differentiation operator with respect to the variable x. From page 334 of

[17], we have the following identities:

$$D_x H_n(x;\rho) = nH_{n-1}(x;\rho) \tag{4.17}$$

$$D_\rho H_n(x;\rho) = -\frac{1}{2} D_x^2 H_n(x;\rho) \tag{4.18}$$

$$H_n(x+y;\rho) = \sum_{k=0}^{n} \binom{n}{k} H_{n-k}(x;\rho) y^k \tag{4.19}$$

We use these facts to prove the following lemma.

Lemma 4.3.4 (Lemma 5.4 of [19]). *The stochastic process* $X(t) = H_n\Big(\int_a^t h(s)\,dB(s); \int_a^t h(s)^2\,ds\Big)$ *with* $h(t) \in L^2[a,b]$ *is a martingale with respect to the filtration generated by the Brownian motion* $B(t)$ *and*

$$dX(t) = nH_{n-1}\Big(\int_a^t h(s)\,dB(s); \int_a^t h(s)^2\,ds\Big) h(t)\,dB(t) \tag{4.20}$$

Proof. Here $x = \int_a^t h(s)\,dB(s)$ and $\rho = \int_a^t h(s)^2\,ds$. So we have $dx = h(t)\,dB(t)$ and $d\rho = h(t)^2\,dt$, and $(dx)^2 = d\rho$. Using Itô's formula, we get

$$\begin{aligned} dX(t) &= D_x H_n(x;\rho)dx + \cancel{\frac{1}{2} D_x^2 H_n(x;\rho)(dx)^2} + \cancel{D_\rho H_n(x;\rho) d\rho} \\ &= nH_{n-1}\Big(\int_a^t h(s)\,dB(s); \int_a^t h(s)^2\,ds\Big) h(t)\,dB(t), \end{aligned}$$

where we used equation (4.18) for the cancellation and (4.17) to get the final term. □

This leads to the following result.

Theorem 4.3.5 (Theorem 5.5 of [19]). *Let* $\alpha(t)$, $\beta(t)$, *and* $h(t)$ *satisfy Hypothesis 4.2.1. For a fixed* $n \geq 1$, *suppose* $Z(t)$ *is the solution of the linear stochastic differential equation*

$$\begin{cases} dZ(t) = \alpha(t)Z(t)\,dB(t) + \beta(t)Z(t)\,dt, & t \in [a,b], \\ Z(a) = H_n\Big(\int_a^b h(s)\,dB(s); \int_a^b h(s)^2\,ds\Big). \end{cases} \tag{4.21}$$

Then $X(t) = \mathbb{E}\left(Z(t)|\mathcal{F}_t\right)$ *is given by*

$$X(t) = H_n\Big(\int_a^t h(s)\,dB(s) - \int_a^t h(s)\alpha(s)\,ds; \int_a^t h(s)^2\,ds\Big)\mathcal{E}_{\alpha,\beta}(t), \quad t \in [a,b]. \tag{4.22}$$

Moreover, $X(t)$ *satisfies the following stochastic differential equation*

$$\begin{cases} dX(t) = \alpha(t)X(t)\,dB(t) + \beta(t)X(t)\,dt \\ \qquad + nH_{n-1}\Big(\int_a^t h(s)\,dB(s) - \int_a^t h(s)\alpha(s)\,ds; \int_a^t h(s)^2\,ds\Big)\mathcal{E}_{\alpha,\beta}(t)h(t)\,dB(t) \\ X(a) = 0. \end{cases} \tag{4.23}$$

Remark 4.3.6 (Remark 5.6 of [19])**.** *For any* x *and* ρ*, we have* $H_0(x;\rho) = 1$*. Hence the stochastic differential equation (4.21) is identically equation* (4.4)*.*

Proof. We first prove equation (4.22). Using Theorem 4.2.2 and equation (4.19), we can write

$$\begin{aligned} Z(t) &= \mathcal{E}_{\alpha,\beta}(t)H_n\Big(\int_a^b h(s)\,dB(s) - \int_a^t h(s)\alpha(s)\,ds; \int_a^b h(s)^2\,ds\Big) \\ &= \mathcal{E}_{\alpha,\beta}(t)\sum_{k=0}^n \binom{n}{k} H_{n-k}\Big(\int_a^b h(s)\,dB(s); \int_a^b h(s)^2\,ds\Big)\Big(-\int_a^t h(s)\alpha(s)\,ds\Big)^k \\ &= \mathcal{E}_{\alpha,\beta}(t)\sum_{k=0}^n \binom{n}{k} J_{n-k}(b)\Big(-\int_a^t h(s)\alpha(s)\,ds\Big)^k, \end{aligned}$$

where we used the notation

$$J_n(t) = H_n\Big(\int_a^t h(s)\,dB(s); \int_a^t h(s)^2\,ds\Big).$$

Using Lemma 4.3.4, we get $\mathbb{E}\left(J_{n-k}(b)|\mathcal{F}_t\right) = J_{n-k}(t)$. Taking the conditional expectation with the knowledge that $\mathcal{E}_{\alpha,\beta}(t)$ is adapted and that stochastic integrals of adapted

processes are adapted,

$$
\begin{aligned}
X(t) &= \mathcal{E}_{\alpha,\beta}(t) \sum_{k=0}^{n} \binom{n}{k} \mathbb{E}\left(J_{n-k}(b) | \mathcal{F}_t\right) \Big(- \int_a^t h(s)\alpha(s)\, ds \Big)^k \\
&= \mathcal{E}_{\alpha,\beta}(t) \sum_{k=0}^{n} \binom{n}{k} J_{n-k}(t) \Big(- \int_a^t h(s)\alpha(s)\, ds \Big)^k \\
&= \mathcal{E}_{\alpha,\beta}(t) H_n \Big(\int_a^t h(s)\, dB(s) - \int_a^t h(s)\alpha(s)\, ds; \int_a^t h(s)^2\, ds \Big),
\end{aligned}
$$

which proves equation (4.22).

Since $H_n(0;0) = 0$, we see that $X(a) = 0$. Using Itô's formula and equation (4.20),

$$
\begin{aligned}
dH_n &= dH_n \Big(\int_a^t h(s)\, dB(s) - \int_a^t h(s)\alpha(s)\, ds; \int_a^t h(s)^2\, ds \Big) \\
&= D_x H_n \cdot (h(t)\, dB(t) - h(t)\alpha(t)\, dt) \\
&\qquad + \cancel{\frac{1}{2} D_x^2 H_n \cdot (h(t)^2\, dt)} + \cancel{D_\rho H_n \cdot (h(t)^2\, dt)} \\
&= nH_{n-1} \cdot h(t)(dB(t) - \alpha(t)\, dt).
\end{aligned}
$$

Finally, using equation (4.22), we get

$$
\begin{aligned}
dX(t) &= H_n\, \mathcal{E}_{\alpha,\beta}(t) + \mathcal{E}_{\alpha,\beta}(t)\, dH_n + d\mathcal{E}_{\alpha,\beta}(t) \cdot dH_n \\
&= H_n \mathcal{E}_{\alpha,\beta}(t)(\alpha(t)\, dB(t) + \beta(t)\, dt) \\
&\quad + \mathcal{E}_{\alpha,\beta}(t) nH_{n-1} \cdot h(t)(dB(t) - \cancel{\alpha(t)}\, dt) + \cancel{\mathcal{E}_{\alpha,\beta}(t)\alpha(t)nH_{n-1}h(t)}\, dt. \\
&= \alpha(t)X(t)\, dB(t) + \beta(t)X(t)\, dt + nH_{n-1}h(t)\mathcal{E}_{\alpha,\beta}(t)\, dB(t),
\end{aligned}
$$

which gives us equation (4.23). □

In Equation (4.23), we specify an explicit form of the extra term in the stochastic differential equation for the conditioned process $X(t)$. We use this result in the following examples.

Example 4.3.7 (Example 5.7 of [19]). *Consider the stochastic differential equation*

$$\begin{cases} dZ(t) = B(t)Z(t)dB(t), \quad t \in [0,1], \\ Z(0) = B(1). \end{cases}$$

Here $\alpha(t) = B(t)$, $\beta(t) \equiv 0$, $h \equiv 1$, *and* $B(1) = H_1(B(1);1)$. *From Theorem 4.3.5,*

$$X(t) = E\left(Z(t)|\mathcal{F}_t\right) = \left(B(t) - \int_0^t B(s)ds\right)\exp\left[\frac{1}{2}\left(B(t)^2 - t - \int_0^t B(s)^2\,ds\right)\right]$$

and $X(t)$ *satisfies the following stochastic differential equation*

$$\begin{cases} dX(t) = \left\{B(t)X(t) + \exp\left[\frac{1}{2}\left(B(t)^2 - t - \int_0^t B(s)^2\,ds\right)\right]\right\}dB(t), \quad t \in [0,1], \\ X(0) = 0. \end{cases}$$

Example 4.3.8 (Example 5.8 of [19]). *Consider the stochastic differential equation*

$$\begin{cases} dZ(t) = B(t)Z(t)dB(t), \quad t \in [0,1], \\ Z(0) = B(1)^2 - 1. \end{cases}$$

From Theorem 4.3.5,

$$X(t) = \left[\left(B(t) - \int_0^t B(s)ds\right)^2 - t\right]\exp\left[\frac{1}{2}\left(B(t)^2 - t - \int_0^t B(s)^2\,ds\right)\right],$$

and for $t \in [0,1]$, $X(t)$ *satisfies the following stochastic differential equation.*

$$\begin{cases} dX(t) = \left\{B(t)X(t) + 2\exp\left[\frac{1}{2}\left(B(t)^2 - t - \int_0^t B(s)^2\,ds\right)\right]\right\}dB(t), \\ X(0) = 0. \end{cases}$$

Chapter 5. LDP Results for a Class of Anticipating Linear Differential Equations

We turn our focus back into our main goal - large deviation results for a class of anticipating linear stochastic differential equations. Consider the following stochastic differential equation

$$\begin{cases} dX_t = \sigma_t X_t \, dB_t + f\left(\int_0^1 \gamma_s dB_s\right) X_t dt, \quad t \in [0,1], \\ X_0 = \xi, \end{cases} \tag{5.1}$$

where ξ is a random variable. Using the intuition gained in finding the solution of the stochastic differential equations with anticipation in Chapter 4, we first "guess" the solution given by equation (5.1) and then use the general Itô formula to show that our guess was indeed the solution. However, is the solution obtained unique? We explore this question via interpreting the stochastic differential equation briefly in the Hitsuda–Skorohod sense. We briefly introduce the theory and present a Trotter inspired product formula to construct the solution and show that the two solutions from both strategies coincide. In the process, we also show the uniqueness of the solution as well. We mainly rely on the Ayed–Kuo formalism, while the Hitsuda–Skorohod perspective is used minimally and out of necessity. We use exploit the explicit form of the solution to find large deviation results for the case when the initial condition is a constant. We use this result to extend to the case when the initial condition is a random variable super-exponentially close to a constant in expectation. Without loss of generality, let us first fix the interval to be $[0, 1]$.

Sections 5.1 - 5.4 of this chapter previously appeared, substantially unchanged, in the open-access *arXiv* article: Kuo, H.-H., Shrestha, P., Sinha, S., and Sundar, P., "On near-martingales and a class of anticipating linear SDEs," `https://arxiv.org/abs/2204.01932`.

5.1. Existence in the Ayed–Kuo sense

We show a general result for existence with respect to the Ayed–Kuo stochastic integral.

Theorem 5.1.1 (Theorem 4.2 of [21]). *Let ξ be a square mean integrable random variable independent of the Brownian motion $B_\bullet$. Consider the linear stochastic differential equation*

$$\begin{cases} dX_t = \sigma_t X_t \, dB_t + f\left(\int_0^1 \gamma_s dB_s\right) X_t dt, \quad t \in [0,1], \\ X_0 = \xi, \end{cases} \tag{5.2}$$

where $\sigma_t, \gamma_t \in L^2[0,1]$, $f \in C^2(\mathbb{R}) \cap L^1(\mathbb{R})$. Then the solution in the Ayed–Kuo sense is given by

$$X_t = \xi \exp\left[\int_0^t \sigma_s dB_s - \frac{1}{2}\int_0^t \sigma_s^2 ds + \int_0^t f\left(\int_0^1 \gamma_v dB_v - \int_s^t \sigma_v \gamma_v dv\right) ds\right]. \tag{5.3}$$

Proof. Consider

$$K(t, x_1, x_2, y) = \xi \exp\left[x_1 - \frac{1}{2}\int_0^t \sigma_s^2 ds + \int_0^t f\left(x_2 + y - \int_s^t \sigma_v \gamma_v dv\right) ds\right].$$

Then, $K(t, X_1, X_2, Y) = X_t$ given

$$\begin{aligned} X_1(t) &= \int_0^t \sigma_s dB_s, & dX_1 &= \sigma_t dB_t, \\ X_2(t) &= \int_0^t \gamma_s dB_s, & dX_2 &= \gamma_t dB_t, \\ Y(t) &= \int_t^1 \gamma_s dB_s, & dY &= -\gamma_t dB_t, \end{aligned}$$

where we write the differential forms in the right-hand column. In order to use the general Itô formula, we need the following derivatives of K.

1. $K_{x_1} = K_{x_1 x_1} = K$,

2. $K_{x_2} = K \cdot \left(\int_0^t F' ds\right)$,

3. $K_{x_1 x_2} = K_{x_2}$,

4. $K_{x_2 x_2} = K \cdot \left(\int_0^t F' ds\right)^2 + K \cdot \left(\int_0^t F'' ds\right)$,

5. $K_y = K \cdot \left(\int_0^t F' ds\right)$,

6. $K_{yy} = K \cdot \left(\int_0^t F' ds\right)^2 + K \cdot \left(\int_0^t F'' ds\right)$,

7. $K_t = K \cdot \left(-\frac{\sigma_t^2}{2} + f\left(x_2 + y\right) + \left(-\sigma_t \gamma_t\right) \int_0^t F' ds\right)$,

where we write $F = f\left(\int_0^1 \gamma_t dB_t - \int_s^t \sigma_v \gamma_v dv\right)$ for brevity. From these calculations, we find the following relationships,

1. $K_y = K_{x_2}$,
2. $K_{yy} = K_{x_2 x_2}$.

Since ξ is independent of the Brownian motion, by the general Itô formula in Theorem 4.0.1,

$$
\begin{aligned}
dK = K_{x_1} dX_1 + \frac{1}{2} K_{x_1 x_1} (dX_1)^2 + K_{x_2} dX_1 + \frac{1}{2} K_{x_2 x_2} (dX_2)^2 + K_y dY \\
- \frac{1}{2} K_{yy} (dY)^2 + K_{x_1 x_2} dX_1 dX_2 + K_t dt.
\end{aligned}
$$

Using from our calculations above,

$$\begin{aligned}
dK =&K\sigma_t dB_t + \frac{1}{2}K(\sigma_t dB_t)^2 + K_{x_2}\gamma_t dB_t + \frac{1}{2}K_{x_2x_2}(\gamma_t dB_t)^2 \\
&+ K_{x_2}(-\gamma_t dB_t) - \frac{1}{2}K_{x_2x_2}(-\gamma_t dB_t)^2 + K_{x_2}(\sigma_t dB_t)(\gamma_t dB_t) + K_t dt \\
=&K\sigma_t dB_t + \frac{1}{2}K\sigma_t^2 dt + K_{x_2}(\sigma_t\gamma_t)dt + K_t dt \\
=&K\sigma_t dB_t + \frac{1}{2}K\sigma_t^2 dt + K\cdot\left(\int_0^t F' ds\right)(\sigma_t\gamma_t)dt \\
&+ \left(K\cdot\left(-\frac{\sigma_t^2}{2} + f\left(x_2+y\right) + (-\sigma_t\gamma_t)\int_0^t F' ds\right)\right)dt \\
=&K\sigma_t dB_t + f(x_2+y)K dt.
\end{aligned}$$

Since $X_t = K(t, X_1, X_2, Y)$, we get

$$dX_t = \sigma_t X_t dB_t + f\left(\int_0^1 \gamma_s dB_s\right) X_t dt,$$

which is exactly (5.2). $\square$

The ease of use of the Ayed–Kuo general Itô formula makes it an good way to analyze anticipating processes. We show an example by finding the stochastic differential equation which X_t^2 itself is a solution of. We will use the results from this example in the last section.

Theorem 5.1.2 (Theorem 4.3 of [21])**.** *Given the stochastic differential equation given by equation* (5.2)*, we have shown that we can obtain an explicit solution. Taking the square of both sides of equation* (5.3)*, we get*

$$X_t^2 = \xi^2 \times \exp\left[\int_0^t 2\sigma_s dB_s - \int_0^t \sigma_s^2 ds + \int_0^t 2f\left(\int_0^1 \gamma_v dB_v - \int_s^t \sigma_v\gamma_v dv\right)ds\right].$$

Consider the function

$$\psi(t, x_1, x_2, y) = \xi^2 \exp\left[x_1 - \int_0^t \sigma_s^2 ds + \int_0^t 2f\left(x_2 + y - \int_s^t \sigma_v\gamma_v dv\right)ds\right].$$

Then, $\psi(t, X_1, X_2, Y) = X_t^2$ *where,*

$$
\begin{aligned}
X_1(t) &= \int_0^t 2\sigma_s dB_s, & dX_1 &= 2\sigma_t dB_t, \\
X_2(t) &= \int_0^t \gamma_s dB_s, & dX_2 &= \gamma_t dB_t, \\
Y(t) &= \int_t^1 \gamma_s dB_s, & dY &= -\gamma_t dB_t.
\end{aligned}
$$

For ease of calculation, we define $L(s,t) = f'(x_2 + y - \int_s^t \sigma_v \gamma_v dv)$. *Here,* ψ *is continuous. Looking at the partials of* ψ *we have,*

1. $\psi_{x_1} = \psi = \psi_{x_1x_1}$,
2. $\psi_{x_2x_2} = \psi_{yy}$,
3. $\psi_{x_2} = \psi_{x_1x_2} = \psi_y = \left[\int_0^t 2L(s,t)\, ds\right] \psi$,
4. $\psi_t = \left(-\sigma_t^2 + 2f(x_2 + y) - 2\int_0^t L(s,t)\, \sigma_t\, \gamma_t\, ds\right) \psi$.

We use the general Itô formula in Theorem 4.0.1 to obtain

$$
\begin{aligned}
d\psi = \psi_{x_1} dX_1 + \frac{1}{2}\psi_{x_1x_1}(dX_1)^2 + \psi_{x_2} dX_2 + \frac{1}{2}\psi_{x_2x_2}(dX_2)^2 + \psi_y dY \\
- \frac{1}{2}\psi_{yy}(dY)^2 + \psi_{x_1x_2} dX_1 dX_2 + \psi_t dt.
\end{aligned}
$$

From our earlier calculations, we have some terms that cancel out,

$$
\begin{aligned}
d\psi = &2\sigma_t \psi dB_t + 2\psi\sigma_t^2 dt + \cancel{\psi_{x_2}\gamma_t dB_t} + \cancel{\frac{1}{2}\psi_{x_2x_2}\gamma_t^2} dt - \cancel{\psi_{x_2}\gamma_t dB_t} \\
&- \cancel{\frac{1}{2}\psi_{x_2x_2}\gamma_t^2} dt + 2\psi_{x_2}\sigma_t\gamma_t dt + \psi_t dt.
\end{aligned}
$$

Plugging in the expressions for ψ_{x_2} and ψ_t, we get,

$$\begin{aligned} d\psi =& 2\sigma_t\psi\, dB_t + 2\psi\sigma_t^2\, dt + 2\left[\int_0^t 2L(s,t)\, ds\right]\psi\sigma_t\gamma_t\, dt \\ &- \sigma^2\psi dt + 2f(x_2+y)\psi\, dt - 2\sigma_t\gamma_t\psi\int_0^t L(s,t)\, ds\, dt \\ =& 2\sigma_t\psi\, dB_t + \sigma^2\psi\, dt + 2f(x_2+y)\psi\, dt + 2\sigma_t\gamma_t\psi\int_0^t L(s,t)\, ds\, dt. \end{aligned}$$

From this calculation, we have that X_t^2 is a solution of the anticipating linear stochastic differential equation given by

$$\begin{cases} dY_t = 2\sigma_t Y_t\, dB_t + \Bigg[\sigma_t^2 + f\left(\int_0^1 \gamma_s dB_s\right) \\ \qquad\qquad + 2\sigma_t\gamma_t\int_0^t f'\left(\int_0^1 \gamma_v dB_v - \int_s^t \sigma_v\gamma_v dv\right) ds\Bigg] Y_t dt, \\ X_0 = \xi^2, \end{cases}$$

for $t \in [0,1]$. *Note that we can observe an interesting feature when we look at the stochastic differential equation we obtained. Namely, we obtain the derivative of the function f as an extra term.*

5.2. A Product Formula for Existence of Solutions

For this section, we temporarily leave the Ayed–Kuo formalism and introduce the Gross–Malliavin theory of stochastic calculus via the Gross–Malliavin derivative and its adjoint, the Hitsuda–Skorohod integral. Before we explore the product formula, let us set up the mathematical framework for this section. We define the construction of the derivative similar to the derivative operators in Sobolev spaces. Namely, we define the operations on a dense space first and then extend to a bigger space. For more details, refer to [26] or [27]. We use this theory in obtaining a existence and uniqueness result for the solution of a simple stochastic differential equation with no drift and anticipation from the

initial condition. We will then use this result in showing the existence and uniqueness of the anticipating linear stochastic differential equation given by Equation (5.1). Here, we fix the time interval $t \in [0, 1] \subset \mathbb{R}_+$.

5.2.1. Brief Introduction to the Gross–Malliavin Derivative and the Hitsuda–Skorohod Integral

Let us first set up the spaces to operate on. We operate on the probability space $(\Omega, \mathcal{F}, P)$ where $\mathcal{F}$ is the σ-field generated by the Brownian motion, B_t, $t \in [0, 1]$. Heuristically, the Gross–Malliavin derivative operator makes precise the concept of differentiation with respect to $\omega \in \Omega$. Let us consider the following random variable,

$$B_{\frac{1}{2}} = \int_0^1 \mathbb{1}_{[0,\frac{1}{2}]}(t)\, dB_t.$$

We can view this random variable to be a function of the integrand. We generalize on this concept. Let $\mathbb{H} = L^2([0, 1])$ be the space of square integrable functions defined on $[0, 1]$. For any $h \in \mathbb{H}$, consider the Wiener integral,

$$B(h) = \int_0^1 h(t)\, dB_t.$$

From here on, we will suppress the time dependency for $h \in \mathbb{H}$ unless otherwise specified. This Hilbert space structure of $\mathbb{H}$ plays an important role in the definition of the derivative. Let $\mathfrak{G}$ be the class of smooth random variables such that $G \in \mathfrak{G}$ has the form

$$G = g\left(B(h_1), B(h_2) \ldots, B(h_n)\right), \quad h_i \in \mathbb{H}\, \forall\, i \in [1, n], \tag{5.4}$$

where g is a real valued n-dimensional smooth function whose derivatives grow at most polynomially.

Definition 5.2.1 (Definition 2.1 of [27]). *The Gross–Malliavin derivative of a smooth random variable $G \in \mathfrak{G}$ is the real valued random variable given by*

$$D_t G = \sum_{i=1}^{n} d_i g\left(B(h_1), B(h_2) \ldots, B(h_n)\right) h_i(t),$$

where d_i is the derivative with respect to the ith variable.

The definition above is well defined meaning that the derivative operator does not depend on the representation given in Equation (5.4). The Gross–Malliavin derivative of a random variable G is a random variable that takes values in $\mathbb{H}$. Furthermore, when considering the time variable, $(D_t G)$, $t \geq 0$ is a stochastic process on $L^2(\Omega \times [0,1])$. This derivative operator allows us to generalize the notion of the Brownian derivative in terms of ω.

Similarly, we define $\mathfrak{F}$ as the class of smooth cylindrical stochastic process $u = (u_t)_{t \geq 0}$ given by

$$u_t = \sum_{i=1}^{n} G_i h_i(t), \qquad \text{for } G_i \in \mathfrak{G}, h_i \in \mathbb{H}. \tag{5.5}$$

Definition 5.2.2 (Definition 2.2 of [27]). *We define the divergence of an element of the form given by Equation* (5.5) *as the random variable*

$$\delta(u) \triangleq \sum_{i=1}^{n} G_i B(h_i) - \sum_{i=1}^{n} \langle DG_i, h_i \rangle_{\mathbb{H}}.$$

We have that the divergence operator δ is the adjoint of the derivative operator $D_\bullet$ as shown in the following proposition.

Proposition 5.2.3 (Proposition 2.1 of [27]). *Let $G \in \mathfrak{G}$ and $u \in \mathfrak{F}$. Then*

$$\mathbb{E}\left[G\delta(u)\right] = \mathbb{E}\left[\langle DG, u \rangle_{\mathbb{H}}\right]$$

Using this adjointed-ness and the dense nature of $\mathfrak{G}$ and $\mathfrak{F}$ in $L^2(\Omega)$ and $L^2(\Omega \times [0,1])$ respectively, we have a closed extension of the derivative operator which we also refer to as D. We define $\mathbb{D}^{1,2}$ as the closure of $\mathfrak{G}$ with respect to the semi-norm $\|\bullet\|_{1,2}$ given by,

$$\|G\|_{1,2} = \left[E\left(|G|^2\right) + E\left[\|DG\|_{\mathbb{H}}^2\right]\right]$$

We define the divergence by extending the adjoint relationship with the derivative operator as far as possible. We follow Nualart in denoting the divergence operater by δ. Namely,

Definition 5.2.4 (Definition 1.3.1 of [26])**.** *We denote by δ the adjoint of the operator D. That is, δ is an unbounded operator on $L^2(\Omega;\mathbb{H})$ with values in $L^2(\Omega)$ such that:*

1. *The domain of δ, $Dom(\delta)$, is the set of $\mathbb{H}$-valued square integrable random variables $u \in L^2(\Omega;\mathbb{H})$ such that for any $F \in \mathbb{D}^{1,2}$, where c is some constant depending on u.*

$$\mathbb{E}\left(\langle DF, u\rangle_{\mathbb{H}}\right) \leq c\,\|F\|_2\,.$$

2. *If u belongs to the domain of δ, then $\delta(u)$ is the element of $L^2(\Omega)$ characterized by*

$$\mathbb{E}\left(F\delta u\right) = \mathbb{E}\left(\langle DF, u\rangle_{\mathbb{H}}\right).$$

for any $F \in \mathbb{D}^{1,2}$.

We call this divergence operator as a Hitsuda-Skorohod integral as it coincides with the anticipating integrals introduced by Skorohod when considering the Brownian motion case [29]. The divergence operator is considered an extension of the Itô integral as it coincides with the Itô integral when the integrand is adapted. Namely,

Theorem 5.2.5 (Proposition 1.3.18 of [26]). *Let u_t, $t \in [0,1]$ be a $\mathcal{F}_t$ adapted stochastic process such that $\mathbb{E}\left[\int_0^1 u_t^2\, dt\right] < \infty$. Then $u \in Dom(\delta)$ and its Hitsuda–Skorohod integral coincides with the Itô integral*

$$\int_0^1 u_t\, \delta B_t = \int_0^1 u_t\, dB_t.$$

It is natural to ask about the nature of the relationship between the Ayed–Kuo integral and the Hitsuda–Skorohod integrals. To that, we list the following result.

Theorem 5.2.6 (Theorem 2.3 of [28]). *Assume $[a,b] \subset [0,\infty)$. Let f be an adapted L^2-continuous stochastic process and ϕ be an instantly independent L^2-continuous stochastic process such that the sequence*

$$\sum_{i=1}^{n} f(t_{i-1})\phi(t_i)\left(B_{t_i} - B_{t_{i-1}}\right),$$

converges strongly in $L^2(\Omega)$ as the mesh $\|\Delta_n\|$ tends to zero. Then the limit $I(f\psi)$ equals the Hitsuda–Skorokhod integral $\delta(f\psi)$ in $Dom(\delta)$.

In Remark 2.5 of [28] , the author shows that $\Phi(B_t, B_1 - B_t) = e^{(B_t)^2+(B_1-B_t)^2}$ for $t \in [0,1]$ is not a Hitsuda–Skorohod integrable process. Since the definition of the Ayed–Kuo integral only needs continuity with respect to the arguments, we can obtain the stochastic integral of Φ in the Ayed–Kuo sense. This is but one of many examples of how the Ayed–Kuo stochastic integral is easier to work with when dealing with anticipation.

5.2.2. Existence and Uniqueness in the Hitsuda–Skorohod Sense

We apply the heavy machinery in the previous subsection to solve a simple stochastic differential equation that we shall use in the next subsection. Consider the family of

transformations $A_t, T_t : \Omega \to \Omega,\ t \in [0,1]$, given by

$$T_t(\omega)_s = \omega_s + \int_0^{t\wedge s} \sigma_u\, du,$$
$$A_t(\omega)_s = \omega_s - \int_0^{t\wedge s} \sigma_u\, du,$$

and define,

$$\mathcal{E}_t = e^{\int_0^t \sigma_s\, dB_s - \frac{1}{2}\int_0^t \sigma_s^2\, ds}.$$

Theorem 5.2.7 (Lemma 4.8 of [21])**.** *Consider the stochastic equation*

$$X_t = X_0 + \int_0^t \sigma_s X_s dB_s, \tag{5.6}$$

for $0 \le t \le 1$*. Here,* $\sigma \in L^2([0,1])$ *and* $X_0 \in L^p(\Omega)$ *for some* $p > 2$*. Then*

$$X_t = X_0(A_t)\mathcal{E}_t \tag{5.7}$$

is the unique solution of equation (5.6)*.*

Proof. It is clear that $\{X_0(A_t)\mathcal{E}_t : 0 \le t \le 1\}$ is $L^r(\Omega)$-bounded for all $r < p$ by the Girsanov theorem and Hölder inequality.

Let G be any smooth random variable. Multiply both sides of (5.6) by G. With the process X given by (5.7), we use the duality relationship given by Definition 5.2.4 to

write

$$
\begin{aligned}
\mathbb{E}\left[G\int_0^t \sigma_s X_s dB_s\right] &= \mathbb{E}\left[\int_0^t \sigma_s X_s D_s G ds\right] \\
&= \mathbb{E}\left[X_0\int_0^t \sigma_s\left(D_s G\right)(T_s) ds\right] \\
&= \mathbb{E}\left[X_0\int_0^t \frac{d}{ds} G(T_s) ds\right] \\
&= \mathbb{E}\left[X_0(G(T_t) - G)\right] \\
&= \mathbb{E}\left[X_0(A_t)\mathcal{E}_t\, G\right] - \mathbb{E}\left[X_0\, G\right] \\
&= \mathbb{E}\left[X_t\, G\right] - \mathbb{E}\left[X_0\, G\right],
\end{aligned}
$$

where the second and second to last equality are given by the Girsanov Theorem in Theorem 1.5.2. Thus a solution of the stochastic equation (5.6) is explicitly given by (5.7). Uniqueness follows since the solution of (5.6) started at $X_0 \equiv 0$ is identically zero at all times. □

5.2.3. Existence via Product formula

Now we come back to tacking the large deviation principle problem statement for anticipating stochastic differential equations. The procedure of finding the solution in the previous section started with "guessing" the form of the solution and applying the formula to it. In this section, we introduce an iterative "braiding" technique in the spirit of Trotter's product formula [30]. We will use existence and uniqueness of solutions to anticipating linear stochastic differential equations without drift that was obtained in Theorem 5.2.7. For the ease of computation, we define,

1. $I_\gamma = \int_0^1 \gamma_s\, dB_s$,
2. $A_u^v(\omega_\cdot) = \omega_\bullet - \int_u^{(\bullet\wedge v)\vee u} \sigma_t\, dt$,

3. $E_u^v = \exp\left\{\int_u^v \sigma_t \, dB_t - \frac{1}{2}\int_u^v \sigma_t^2 \, dt\right\}$,

4. $g_u^v = e^{f(I_\gamma)(v-u)}$.

Via the definition, we can also obtain the following properties for the functions. For $u < v < s$,

1. $A_u^v \circ A_v^s(\omega.) = A_u^s(\omega.)$,

2. $E_u^v \cdot E_v^s = E_u^s$,

3. $g_u^v \cdot g_v^s = g_u^s$.

We list some relevant interactions between these functions for $u < v < s$.

1. $E_u^v \circ A_v^s = E_u^v$.

2. $g_u^v \circ A_v^s = \exp\{f(I_\gamma \circ A_v^s)(v-u)\}$.

The first interaction equality is a statement regarding the invariance of an adapted process when considering shifts from a later time steps. We show the result via a simple example which can easily be extended to the exponential case above. For $u < v < t$,

$$(B_u) \circ A_v^t = B_u - \int_v^{(u \wedge v) \vee t} \sigma_s ds = B_u - \int_v^v \sigma_s ds = B_u.$$

For convenience, we will suppress the dependence on ω for the random variables.

Theorem 5.2.8 (Theorem 4.10 of [21]). *Consider the stochastic differential equation,*

$$\begin{cases} dX_t = \sigma_t X_t \, dB_t + f(\int_0^1 \gamma_s dB_s) X_t dt, \quad t \in [0,1], \\ \\ X_0 = \xi, \end{cases} \tag{5.8}$$

where $\sigma, \gamma \in L^2[0,1], f \in C^2(\mathbb{R}) \cap L^1(\mathbb{R})$, and $\xi \in L^p(\Omega)$ for some $p > 2$. Then the unique solution in the Hitsuda–Skorohod sense is given by

$$X_t = (\xi \circ A_0^t) \exp\left[\int_0^t \sigma_s dB_s - \frac{1}{2}\int_0^t \sigma_s^2 ds + \int_0^t f\left(\int_0^1 \gamma_v dB_v - \int_s^t \sigma_v \gamma_v dv\right) ds\right]. \tag{5.9}$$

Proof. Let $t \in [0,1]$ and let Δ_n be a partition of $[0,t]$.

First time step. Let $u \in [0, t_1]$.

a) Define the following stochastic integral equation.

$$\begin{cases} dY_u^{(1)} = \sigma_u Y_u^{(1)} \, dB_u, \quad u \in [0, t_1], \\ \\ Y_0^{(1)} = \xi. \end{cases}$$

Then, via Theorem 5.2.7 we have the unique solution,

$$Y_u^{(1)} = (\xi \circ A_0^u) E_0^u$$

is the unique solution almost surely. Then

$$Y_{t_1}^{(1)} = (\xi \circ A_0^{t_1}) E_0^{t_1}.$$

b) For each $\omega \in \Omega$, define the ordinary differential equation,

$$\begin{cases} dX_u^{(1)} = f(I_\gamma) X_u^{(1)} \, du, \quad u \in [0, t_1], \\ \\ X_0^{(1)} = Y_{t_1}^{(1)}. \end{cases}$$

Then, there exists a solution

$$X_u^{(1)} = Y_{t_1}^{(1)} \cdot g_0^u = (\xi \circ A_0^{t_1}) E_0^{t_1} \cdot g_0^u.$$

Thus,

$$X_{t_1}^{(1)} = (\xi \circ A_0^{t_1}) E_0^{t_1} \cdot g_0^{t_1}.$$

Define $\Omega_1 = \{\omega \,|\, X_{t_1}^{(1)} \text{ exists}\}$. Since $Y_t^{(1)}$ exists almost surely and $X_t^{(1)}$ is a solution of a ordinary differential equation for each ω, $\mathbb{P}(\Omega_1) = 1$. Thus, $X_{t_1}^{(1)}$ exists almost surely.

Second time step. Let $u \in [t_1, t_2]$.

a) Define the following stochastic integral equation.

$$\begin{cases} dY_u^{(2)} = \sigma_u Y_u^{(2)}\, dB_u, \quad u \in [t_1, t_2], \\ Y_{t_1}^{(2)} = X_{t_1}^{(1)}. \end{cases}$$

Then, via Theorem 5.2.7 we again have the unique solution,

$$Y_u^{(2)} = (X_{t_1}^{(1)} \circ A_{t_1}^u) E_{t_1}^u,$$

almost surely. Then

$$Y_{t_2}^{(2)} = (X_{t_1}^{(1)} \circ A_{t_1}^{t_2}) E_{t_1}^{t_2}.$$

b) For each $\omega \in \Omega$, define the ordinary differential equation,

$$\begin{cases} dX_u^{(2)} = f(I_\gamma) X_u^{(2)}\, du, \quad u \in [t_1, t_2], \\ X_{t_1}^{(2)} = Y_{t_2}^{(2)}. \end{cases}$$

Then, there exists a solution

$$X_u^{(2)} = Y_{t_2}^{(2)} \cdot g_0^u = (X_{t_1}^{(1)} \circ A_{t_1}^{t_2}) E_{t_1}^{t_2} \cdot g_{t_1}^u.$$

Thus,

$$\begin{aligned} X_{t_2}^{(2)} &= (X_{t_1}^{(1)} \circ A_{t_1}^{t_2}) E_{t_1}^{t_2} \cdot g_{t_1}^{t_2} \\ &= (((\xi \circ A_0^{t_1}) E_0^{t_1} \cdot g_0^{t_1}) \circ A_{t_1}^{t_2}) E_{t_1}^{t_2} \cdot g_{t_1}^{t_2} \\ &= (\xi \circ A_0^{t_1} \circ A_{t_1}^{t_2}) E_0^{t_1} \cdot E_{t_1}^{t_2} (g_0^{t_1} \circ A_{t_1}^{t_2}) g_{t_1}^{t_2} \\ &= (\xi \circ A_0^{t_2}) E_0^{t_2} \prod_{i=1}^{2} (g_{t_{i-1}}^{t_i} \circ A_{t_i}^{t_2}), \end{aligned}$$

where the last equality is by rewriting $g_{t_1}^{t_2} = g_{t_1}^{t_2} \circ A_{t_2}^{t_2}$ and condensing into a product form.

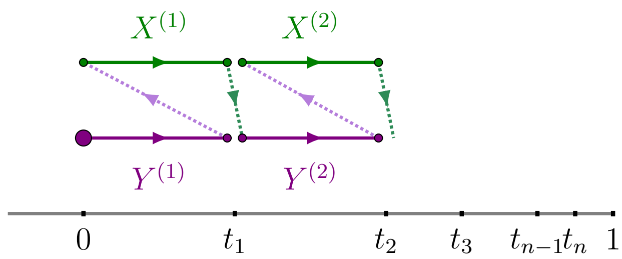

Figure 5.1. A t-dependence plot of the various constructed processes. Solid line represents the constructed process. Dotted line represents the final value of one differential equation within the time step being used as initial condition for the next one.

Define $\Omega_2 = \{\omega \,|\, X^{(2)}_{t_2} \text{ exists }\}$. Since $Y^{(2)}_t$ exists almost surely and $X^{(2)}_t$ is a solution of a ordinary differential equation for each ω, $\mathbb{P}(\Omega_2) = 1$. As an intersection of two sets of probability one, $\mathbb{P}(\Omega_1 \cap \Omega_2) = 1$. Thus, $X^{(2)}_{t_2}$ exists almost surely on $\Omega_1 \cap \Omega_2$.

Lemma 5.2.9 (Lemma 4.9 of [21]). *Let $Y^{(1)}_0 = \xi \in L^p(\Omega)$ for some $p > 2$. Consider the k-th sub-interval $u \in [t_{k-1}, t_k]$ for any $k \in [0, n]$, and define*

1. *the stochastic differential equation*

$$\begin{cases} dY^{(k)}_u = \sigma_u Y^{(k)}_u \, dB_u, \quad u \in [t_{k-1}, t_k], \\ Y^{(k)}_{t_{k-1}} = X^{(k-1)}_{t_{k-1}}, \textit{and} \end{cases}$$

2. *the ordinary differential equation*

$$\begin{cases} dX^{(k)}_u = f(I_\gamma) X^{(2)}_u \, du, \quad u \in [t_1, t_2], \\ X^{(k)}_{t_{k-1}} = Y^{(k)}_{t_k}. \end{cases}$$

Then there exists a set Ω_k with $\mathbb{P}(\Omega_k) = 1$ such that on Ω_k, we have

$$X^{(k)}_{t_k} = (\xi \circ A^{t_k}_0) E^{t_k}_0 \cdot \prod_{i=1}^{k} (g^{t_i}_{t_{i-1}} \circ A^{t_k}_{t_i}).$$

Proof. We prove this by induction.

Base case: This is true for $k = 1, 2$ as shown in the computation above.

Induction step: Assume true for $k = m - 1$. This means that, for all ω in $\cap_{l=1}^{m-1}\Omega_l$ with $\mathbb{P}(\cap_{l=1}^{m-1}\Omega_l) = 1$,

$$X_{t_{m-1}}^{(m-1)} = (\xi \circ A_0^{t_{m-1}})E_0^{t_{m-1}} \cdot \prod_{i=1}^{m-1}(g_{t_{i-1}}^{t_i} \circ A_{t_i}^{t_{m-1}}).$$

We define the next step stochastic differential equation and ordinary differential equation as described in the statement of the lemma with $k = m$. As in the case $k = 2$, we have that

$$\begin{aligned}
X_{t_m}^{(m)} &= (X_{t_{m-1}}^{(m-1)} \circ A_{t_{m-1}}^{t_m})E_{t_{m-1}}^{t_m} \cdot g_{t_{m-1}}^{t_m} \\
&= \left[\left((\xi \circ A_0^{t_{m-1}})E_0^{t_{m-1}} \cdot \prod_{i=1}^{m-1}(g_{t_{i-1}}^{t_i} \circ A_{t_i}^{t_{m-1}})\right) \circ A_{t_{m-1}}^{t_m}\right] E_{t_{m-1}}^{t_m} \cdot g_{t_{m-1}}^{t_m} \\
&= (\xi \circ A_0^{t_m})E_0^{t_m} \cdot \prod_{i=1}^{m-1}\left(g_{t_{i-1}}^{t_i} \circ A_{t_{m-1}}^{t_m} \circ A_{t_i}^{t_{m-1}}\right) \cdot g_{t_{m-1}}^{t_m} \\
&= (\xi \circ A_0^{t_m})E_0^{t_m} \cdot \prod_{i=1}^{m-1}\left(g_{t_{i-1}}^{t_i} \circ A_{t_i}^{t_m}\right) \cdot g_{t_{m-1}}^{t_m} \\
&= (\xi \circ A_0^{t_m})E_0^{t_m} \cdot \prod_{i=1}^{m}\left(g_{t_{i-1}}^{t_i} \circ A_{t_i}^{t_m}\right).
\end{aligned}$$

We can use the almost sure existence of the stochastic differential equation and everywhere existence of the ordinary differential equation to construct Ω_m with $\mathbb{P}(\Omega_m) = 1$. Thus $X_{t_m}^{(m)}$ exists for all $\omega \in \cap_{l=1}^{m}\Omega_l$ with $\mathbb{P}(\cap_{l=1}^{m}\Omega_l) = 1$. □

This allows us to obtain a closed form for the solution at time t using the partition Δ_n. Namely,

$$X_t^{(n)} = (\xi \circ A_0^t)E_0^t \cdot \prod_{i=1}^{n}\left(g_{t_{i-1}}^{t_i} \circ A_{t_i}^t\right). \tag{5.10}$$

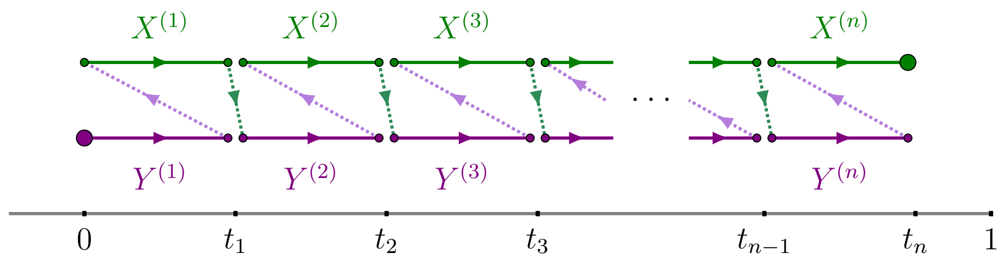

Figure 5.2. A t-dependence plot of the various constructed processes. Solid line represents the constructed process. Dotted line represents the final value of one differential equation within the time step being used as initial condition for the next one.

Let us compute these terms out. When $t_i < t < 1$,

$$I_\gamma(A^t_{t_{i-1}}) = \int_0^1 \gamma_s dB_s - \int_{t_{i-1}}^t \gamma_s \sigma_s ds.$$

We use the results above to evaluate the product term in equation (5.10).

$$\begin{aligned}
\prod_{i=1}^n (g^{t_i}_{t_{i-1}} \circ A^t_{t_i}) &= \prod_{i=1}^n e^{f(I_\gamma \circ A^t_{t_i})(t_i - t_{i-1})} \\
&= \exp\left\{ \sum_{i=1}^n f(\int_0^1 \gamma_v dB_v - \int_{t_{i-1}}^t \gamma_v \sigma_v dv)(t_i - t_{i-1}) \right\} \\
&\to \exp\left\{ \int_0^t f(\int_0^1 \gamma_v dB_v - \int_s^t \gamma_v \sigma_v dv) ds \right\},
\end{aligned}$$

as $n \to \infty$. Thus,

$$\begin{aligned}
X_t &= \lim_{n\to\infty} X_t^{(n)} \\
&= (\xi \circ A^t_0) E^t_0 \exp\left\{ \int_0^t f(\int_0^1 \gamma_s dB_s - \int_s^t \gamma_v \sigma_v dv) ds \right\} \\
&= (\xi \circ A^t_0) \exp\left[\int_0^t \sigma_s dB_s - \frac{1}{2}\int_0^t \sigma_s^2 ds + \int_0^t f\left(\int_0^1 \gamma_v dB_v - \int_s^t \sigma_v \gamma_v dv \right) ds \right].
\end{aligned}$$

Furthermore, X_t exists for all $\omega \in \bar{\Omega} \triangleq \cap_{l=1}^\infty \Omega_l$. As a countable intersection of probability one sets, $\mathbb{P}(\bar{\Omega}) = 1$. Thus X_t is a solution of (5.8) almost surely.

The solution is unique. For if not, there are two solutions which disagree for the first time in a particular interval, say the kth interval. Recall that the solutions obtained

using Malliavin calculus and also for ordinary differential equations are unique for each interval of time. Therefore, such a disagreement would violate this uniqueness. □

5.3. Rate Function for the Solution

We pick up from where we left up at the end of Chapter 2. We are interested in obtaining the large deviation results for the anticipating stochastic differential equation. Now that we have defined the anticipating integral, proved pertinent theorems, established uniqueness and existence of the solution, we proceed to finding the large deviation properties of the solution of the stochastic differential equation given below.

Given $\epsilon > 0$, Consider the stochastic differential equation

$$\begin{cases} dX_t^\epsilon = \sqrt{\epsilon}\sigma_t X_t^\epsilon \, dB_t + f\left(\int_0^1 \sqrt{\epsilon}\gamma_s dB_s\right) X_t^\epsilon dt, \quad t \in [0,1], \\ \\ X_0 = c, \end{cases}$$

where f is a Lipschitz function in $L^2(\mathbb{R})$. Furthermore, assume σ, γ are bounded deterministic functions of bounded variation on $[0,1]$. From our earlier discussion, Theorem 5.1.1 gives us our unique solution

$$\begin{aligned} X_t^\epsilon = c \exp\bigg[\int_0^t \sqrt{\epsilon}\sigma_t dB_s - \frac{\epsilon}{2}\int_0^t \sigma_s^2 ds \\ + \int_0^t f\left(\int_0^1 \sqrt{\epsilon}\gamma_v dB_v - \int_s^t \epsilon\sigma_v\gamma_v dv\right) ds \bigg]. \end{aligned}$$

In view of the Contraction Principle in Theorem 2.2.3, if we show that the solutions are continuous image something that already satisfies large deviation principle, then we are done. As such, we want to show that X_t^ϵ is a continuous functional of Brownian motion in order transfer the results from Schilder's Theorem to our case. We exploit the fact that we have the explicit form of the solution. Namely, let C_k be the space of continuous functions

starting from $k \in \mathbb{R}$ and consider the function $h : C_0 \to C_c$ defined by

$$h(x_\bullet) = c \exp\left[\int_0^t \sigma_s dx_s - \frac{\epsilon}{2}\int_0^t \sigma_s^2 ds \right.$$
$$\left. + \int_0^t f\left(\int_0^1 \gamma_v dx_v - \int_s^t \epsilon \sigma_v \gamma_v dv\right) ds\right]. \tag{5.11}$$

Then $X_t^\epsilon = h(\sqrt{\epsilon} B_\bullet)$. Now we show the continuity of h.

Lemma 5.3.1 (Lemma 5.6 of [21])**.** *The function $h : C_0 \to C_c$ defined by Equation* (5.11) *is continuous in the topology induced by the canonical supremum norm.*

Proof. We can write

$$h(x) = c \exp\left[\phi(x) - \frac{\epsilon}{2}\int_0^t \sigma_s^2 ds + \psi(x)\right],$$

where $\phi, \psi : C_0 \to C_0$ is given by

$$\phi(x) = \int_0^t \sigma_s dx_s = \sigma_t x_t - \int_0^t x_s d\sigma_s, \text{ and}$$
$$\psi(x) = \int_0^t f\left(\int_0^1 \gamma_u dx(u) - \epsilon \int_s^t \gamma_u \, \sigma_u du\right) ds.$$

Using integration by parts,

$$\phi(x) = \sigma_t x_t - \int_0^t x_s d\sigma_s, \text{ and}$$
$$\psi(x) = \int_0^t f\left(\gamma_1 x_1 - \int_0^1 x_s d\gamma_s - \epsilon \int_s^t \gamma_u \, \sigma_u du\right) ds.$$

Since multiplication by $c \exp\left(-\frac{\epsilon}{2}\int_0^t \sigma_s^2 ds\right)$ and exp are continuous transformations, continuity of h is guaranteed if we prove continuity of ϕ and ψ. This is what we show below. In what follows, $\|\bullet\|$ refers to the supremum norm.

Let $x, y \in C_0$. For ϕ, we have

$$
\begin{aligned}
\|\phi(x)-\phi(y)\| &= \left\| \left(\sigma_t x_t - \int_0^t x_s d\sigma_s\right) - \left(\sigma_t y_t - \int_0^t y_s d\sigma_s\right) \right\| \\
&\leq \|\sigma_t \left(x_t - y_t\right)\| + \left\| \int_0^t (x_s - y_s) d\sigma_s \right\| \\
&\leq \|\sigma\| \, \|x-y\| + \|x-y\| \, \|\sigma_t - \sigma_0\| \\
&\leq 3 \, \|\sigma\| \, \|x-y\| ,
\end{aligned}
$$

so ϕ is continuous.

For ψ, if L_f is the Lipschitz constant for f, we get

$$
\begin{aligned}
\|\psi(x)-\psi(y)\| \leq & \left\| \int_0^t L_f \left[\left(\gamma_1 x_1 - \int_0^1 x_s d\gamma_s - \epsilon \cancel{\int_s^t \gamma_u \, \sigma_u du} \right) \right. \right. \\
& \qquad \left. \left. - \left(\gamma_1 y_1 - \int_0^1 y_s d\gamma_s - \epsilon \cancel{\int_s^t \gamma_u \, \sigma_u du} \right) \right] ds \right\| \\
\leq & L_f \left\| \int_0^t \left(\gamma_1 \left(x_1 - y_1\right) - \int_0^1 \left(x_s - y_s\right) d\gamma_s \right) ds \right\| \\
\leq & L_f \left(\|\gamma\| \, \|x-y\| + 2 \, \|\gamma\| \, \|x-y\| \right) \\
= & 3 L_f \, \|\gamma\| \, \|x-y\| ,
\end{aligned}
$$

which proves the continuity of ψ. $\square$

Thus we have that h is continuous and $X_t^\epsilon = h(\sqrt{\epsilon} B_t)$. From Theorem 2.2.3 alongside Theorem 2.4.2, X_t^ϵ follows LDP with the rate function

$$
J(Y) = \inf \left\{ \frac{1}{2} \int_0^1 \left|v'(t)\right|^2 dt \right\} , \tag{5.12}
$$

where the infimum is over all $v \in C_0[0,1]$ that solves the control differential equation

$$
\begin{cases}
dV_t = \sigma_t V_t \, v'(t) dt + f\left(\int_0^1 \gamma_u v'(d) du \right) V_t dt, \quad t \in [0,1], \\
V_0 = c,
\end{cases} \tag{5.13}
$$

with $V_t = X_t^\epsilon$ as the solution. We state the above discussion as a theorem.

Theorem 5.3.2 (Theorem 5.7 of [21]). *Given $\epsilon > 0$, define the family of stochastic differential equations*

$$\begin{cases} dX_t^\epsilon = \sqrt{\epsilon}\sigma_t X_t^\epsilon \, dB_t + f\left(\int_0^1 \sqrt{\epsilon}\gamma_s dB_s\right) X_t^\epsilon dt, \quad t \in [0,1], \\ \\ X_0 = c \in \mathbb{R}, \end{cases}$$

where σ, γ are bounded deterministic functions of bounded variation on $[0,1]$. Moreover, consider $f \in L^2(\mathbb{R})$ is Lipschitz. Then the solutions $\{X_t^\epsilon\}$ follows large deviation principle with rate function $J(\bullet)$ given by equation (5.12).

5.4. Rate Function with Anticipating Initial Conditions

We have shown that the solution of the stochastic differential equation

$$\begin{cases} dX_t^\epsilon = \sqrt{\epsilon}\sigma_t X_t^\epsilon \; dB_t + f\left(\sqrt{\epsilon}\int_0^1 \gamma_u \; dB_u\right) X_t^\epsilon \; dt, \quad t \in [0,1], \\ \\ X_0^\epsilon = c, \quad c \in \mathbb{R}, \end{cases}$$

is given by

$$X_t^\epsilon = c\exp\Bigg[\int_0^t \sqrt{\epsilon}\sigma_t dB_s - \frac{\epsilon}{2}\int_0^t \sigma_s^2 ds + \int_0^t f\left(\int_0^1 \sqrt{\epsilon}\gamma_v dB_v - \int_s^t \epsilon\sigma_v\gamma_v dv\right) ds\Bigg].$$

using the definitions of both the Ayed–Kuo and the Hitsuda–Skorokhod integrals. For $\epsilon > 0$, we also derived the large deviation principle for such a family of stochastic differential equations.

Consider the family of stochastic differential equations

$$\begin{cases} dY_t^\epsilon = \sqrt{\epsilon}\sigma_t Y_t^\epsilon \; dB_t + f\left(\sqrt{\epsilon}\int_0^1 \gamma_u \; dB_u\right) Y_t^\epsilon \; dt, \quad t \in [0,1], \\ \\ Y_0^\epsilon = \xi^\epsilon, \end{cases} \tag{5.14}$$

where the initial conditions ξ^ϵ are random variables independent of the Brownian motion and super-exponentially close in expectation to a constant c. Can we derive a large deviation principle for this family? With that motivation, we state the following result.

Theorem 5.4.1 (Theorem 5.8 of [21]). *Given $\epsilon > 0$, let $\{X_t^\epsilon\}$ be a solutions for a family of anticipating stochastic differential equations given by Equation* (5.14)*, where σ, γ, and their first derivatives σ' and γ' are bounded. Moreover, consider $f \in L^2(\mathbb{R})$ is Lipschitz. Furthermore, let $c \in \mathbb{R}$ and ξ^ϵ be a family of random variables that are independent of the Brownian motion $B_\bullet$ such that*

$$\lim_{\epsilon \to 0} \epsilon \log \mathbb{E}\left[|\xi^\epsilon - c|^2\right] = -\infty. \tag{5.15}$$

Then the solutions $\{X_t^\epsilon\}$ follows large deviation principle with rate function $J(\bullet)$ given by Equation (5.12)*.*

Proof. We show this via super exponential approximation. Let $Z_t^\epsilon = Y_t^\epsilon - X_t^\epsilon$. Then Z_t^ϵ satisfies the equation

$$\begin{cases} dZ_t^\epsilon = \sqrt{\epsilon}\sigma_t Z_t^\epsilon \, dB_t + f\left(\sqrt{\epsilon}\displaystyle\int_0^1 \gamma_u \, dB_u\right) Z_t^\epsilon \, dt, \quad t \in [0,1], \\ Z_0^\epsilon = \xi^\epsilon - c. \end{cases}$$

From our earlier results, the solution is,

$$\begin{aligned} Z_t^\epsilon = (\xi^\epsilon - c) \exp\Bigg[&\int_0^t \sqrt{\epsilon}\sigma_t dB_s - \frac{\epsilon}{2}\int_0^t \sigma_s^2 ds \\ &+ \int_0^t f\left(\int_0^1 \sqrt{\epsilon}\gamma_v dB_v - \int_s^t \epsilon\sigma_v\gamma_v dv\right) ds\Bigg]. \end{aligned}$$

Let $\phi(z) = |z|^2$ and let $U_t^\epsilon = \phi(Z_t^\epsilon)$. From Theorem 5.1.2, we have that U_t^ϵ satisfies

the stochastic integral equation

$$
\begin{aligned}
U_t^\epsilon &= (\xi^\epsilon - c)^2 + 2\int_0^t \sqrt{\epsilon}\sigma_s \, U_s^\epsilon \, dB_s \\
&\quad + \int_0^t \epsilon\sigma_s^2 \, U_s^\epsilon \, ds + 2\int_0^t f\left(\int_0^1 \sqrt{\epsilon}\gamma_u \, dB_u\right) \, U_s^\epsilon \, ds \\
&\quad + 2\int_0^t \epsilon\sigma_s \, \gamma_s \, U_s^\epsilon \, \int_0^s f'\left(\int_0^1 \sqrt{\epsilon}\gamma_v \, dB_v - \int_u^s \epsilon\sigma_v\gamma_v \, dv\right) du \, ds.
\end{aligned} \tag{5.16}
$$

Fix $\delta > 0$ and let τ be a stopping time defined by $\tau = \inf\{t : |Z_t^\epsilon| \geq \delta\} \wedge 1$. Consider the stopped process $U_{t\wedge\tau}^\epsilon$. Using the stochastic equation given by equation (5.16) and taking expectations, we have

$$
\begin{aligned}
\mathbb{E}\left(U_{t\wedge\tau}^\epsilon\right) &= \mathbb{E}\left(U_0^\epsilon\right) + 2\mathbb{E}\left[\int_0^{t\wedge\tau} \sqrt{\epsilon}\sigma_s \, U_s^\epsilon \, dB_s\right] + \mathbb{E}\left[\int_0^{t\wedge\tau} \epsilon\sigma_s^2 \, U_{s\wedge\tau}^\epsilon \, ds\right] \\
&\quad + 2\mathbb{E}\left[\int_0^{t\wedge\tau} f\left(\int_0^1 \sqrt{\epsilon}\gamma_u \, dB_u\right) \, U_{s\wedge\tau}^\epsilon \, ds\right] \\
&\quad + 2\mathbb{E}\left[\int_0^t \epsilon\sigma_s \, \gamma_s \, U_s^\epsilon \, \int_0^s f'\left(\int_0^1 \sqrt{\epsilon}\gamma_v \, dB_v - \int_u^s \epsilon\sigma_v\gamma_v \, dv\right) du \, ds\right].
\end{aligned}
$$

By Theorem 3.4.2, the second integral gives us a near-martingale. We use our assumption that f, f', σ, γ are bounded by $M \geq 1$. Using the version of near-martingale optional stopping theorem in corollary 3.4.13 and the non-negativity of U_t^ϵ we get

$$
\begin{aligned}
\mathbb{E}\left(U_{t\wedge\tau}^\epsilon\right) &\leq \mathbb{E}\left(U_0^\epsilon\right) + 2\cdot 0 \\
&\quad + \int_0^{t\wedge\tau} \epsilon M^2 \mathbb{E}\left[U_{s\wedge\tau}^\epsilon\right] \, ds + 2\int_0^{t\wedge\tau} M\mathbb{E}\left[U_{s\wedge\tau}^\epsilon\right] \, ds \\
&\quad + 2\int_0^{t\wedge\tau} \epsilon M \cdot M \cdot \mathbb{E}\left[U_{s\wedge\tau}^\epsilon\right] \cdot M \cdot 1 \, ds \\
&\leq \mathbb{E}\left(U_0^\epsilon\right) + \left(2M + \epsilon M^2 + 2\epsilon M^3\right) \int_0^{t\wedge\tau} \mathbb{E}\left[U_{s\wedge\tau}^\epsilon\right] \, ds.
\end{aligned}
$$

We define $K_\epsilon = 2M + \epsilon M^2 + 2\epsilon M^3$. Using Grönwall's inequality, we get

$$
\mathbb{E}\left(U_\tau^\epsilon\right) = \mathbb{E}\left(U_{1\wedge\tau}^\epsilon\right) \leq \mathbb{E}\left(U_0^\epsilon\right) e^{K_\epsilon}.
$$

Since $\phi(z)$ is positive and monotone increasing in $|z|$, we use Markov's inequality to get

$$\mathbb{P}\left\{|Z_\tau^\epsilon| \geq \delta\right\} = \mathbb{P}\left\{\phi\left(|Z_\tau^\epsilon|\right) \geq \phi(\delta)\right\} \leq \frac{\mathbb{E}\left(\phi\left(|Z_\tau^\epsilon|\right)\right)}{\phi(\delta)} = \frac{\mathbb{E}\left(U_\tau^\epsilon\right)}{\delta^2} \leq \frac{\mathbb{E}\left(U_0^\epsilon\right)}{\delta^2} e^{K_\epsilon}.$$

We take log of both sides and multiply by ϵ to get,

$$\epsilon \log \mathbb{P}\left\{|Z_\tau^\epsilon| \geq \delta\right\} \leq \epsilon \log\left(\frac{1}{\delta^2} e^{K_\epsilon}\right) + \epsilon \log \mathbb{E}\left|\xi^\epsilon \;\; - c\right|^2.$$

Taking the limit superior as ϵ goes to zero, we have that the first term on the right goes to zero while we use our assumption from Equation (5.15) to get

$$\limsup_{\epsilon \to 0} \epsilon \log \mathbb{P}\left\{|Z_{t\wedge\tau}^\epsilon| > \delta\right\} = -\infty.$$

This result allows us to say that X_t^ϵ and Y_t^ϵ are super-exponentially close. Thus by the super-exponential approximation theorem in Theorem 2.2.4, we have that Y_t^ϵ follows large deviation principle with the same rate function given by Equation (5.12). Namely,

$$J(Y) = \inf_v \left\{\frac{1}{2}\int_0^1 \left|v'(t)\right|^2 dt\right\},$$

with the infimum over $v \in C_0[0,1]$ as described in (5.13). □

Appendix. Copyright Information

Copyright information for Chapter 3

Journal of Stochastic Analysis
Vol. 2, No. 2 (2021) Article 5 (18 pages)
DOI: 10.31390/josa.2.2.05

ANTICIPATING LINEAR STOCHASTIC DIFFERENTIAL EQUATIONS WITH ADAPTED COEFFICIENTS

HUI-HSIUNG KUO, PUJAN SHRESTHA*, AND SUDIP SINHA

Abstract. Stochastic differential equations with adapted integrands and initial conditions are well studied within Itô's theory. However, such a general theory is not known for corresponding equations with anticipation. We use examples to illustrate essential ideas of the Ayed–Kuo integral and techniques for dealing with anticipating stochastic differential equations. We prove the general form of the solution for a class of linear stochastic differential equations with adapted coefficients and anticipating initial condition, which in this case is an analytic function of a Wiener integral. We show that for such equations, the conditional expectation of the solution is not the same as the solution of the corresponding stochastic differential equation with the initial condition as the expectation of the original initial condition. In particular, we show that there is an extra term in the stochastic differential equation, and give the exact form of this term.

1. Introduction

Let $B(t)$, where $t \in [a,b]$, be a Brownian motion starting at 0 and let $\{\mathcal{F}_t\}$ be the filtration generated by $B(t)$, that is, $\mathcal{F}_t = \sigma\{B(s); a \leq s \leq t\}$. In the framework of Itô's calculus, a stochastic differential equation

$$\begin{cases} dX(t) = \alpha(t, X(t))\, dB(t) + \beta(t, X(t))\, dt, & t \in [a,b], \\ X(a) = \xi, \end{cases}$$

with the initial condition ξ being $\mathcal{F}_a$-measurable, is a symbolical representation of the stochastic integral equation

$$X(t) = \xi + \int_a^t b(s, X(s))\, ds + \int_a^t \sigma(s, X(s))\ dB(s), \quad t \in [a,b],$$

where $\int_a^t \sigma(s, X(s))\ dB(s)$ is defined as an Itô integral. In Itô's framework, we require both the coefficients $b(t,x,\omega)$ and $\sigma(t,x,\omega)$ to be adapted apart from usual integrability constraints, and and the initial condition ξ to be measurable with respect to the initial σ-algebra $\mathcal{F}_a$. The question of how the stochastic integral can be defined when any of these quantities are not adapted (called *anticipating*) has been an open question in the field of stochastic analysis for past decades.

Received 2021-4-17; Accepted 2021-5-6; Communicated by the editors.
2010 *Mathematics Subject Classification.* Primary 60H05; Secondary 60H10, 60H20.
Key words and phrases. anticipation, stochastic integral, stochastic integral, Itô formula, stochastic differential equations, conditional expectation, adapted coefficients.

* Corresponding author.

1

Copyright information for Chapter 4

Journal of Stochastic Analysis
Vol. 2, No. 4 (2021) Article 8 (13 pages)
DOI: 10.31390/josa.2.4.08

AN INTRINSIC PROOF OF AN EXTENSION OF ITÔ'S ISOMETRY FOR ANTICIPATING STOCHASTIC INTEGRALS

HUI-HSIUNG KUO, PUJAN SHRESTHA, AND SUDIP SINHA*

Abstract. Itô's isometry forms the cornerstone of the definition of Itô's integral and consequently the theory of stochastic calculus. Therefore, for any theory which extends Itô's theory, it is important to know if the isometry holds. In this paper, we use probabilistic arguments to demonstrate that the extension of the isometry formula contains an extra term for the anticipating stochastic integral defined by Ayed and Kuo. We give examples to illustrate the usage of this formula and to show that the extra term can be positive or negative.

1. Introduction

Let $B_t, t \geq 0$, be a Brownian motion and $[a, b]$ a fixed interval with $a \geq 0$. Suppose f and ϕ are continuous functions on $\mathbb{R}$. In [1] the following anticipating stochastic integral is defined as

$$\int_a^b f(B_t)\phi(B_b - B_t)\, dB_t = \lim_{\|\Delta_n\| \to 0} \sum_{i=1}^n f(B_{t_{i-1}})\phi(B_b - B_{t_i})\Delta B_i \tag{1.1}$$

provided that the limit exists in probability. Here $\Delta_n = \{a = t_0, t_1, t_2, \ldots, t_n = b\}$ is a partition of $[a, b]$ and $\Delta B_i = B_{t_i} - B_{t_{i-1}}$. Note that when $\phi \equiv 1$ this stochastic integral is an Itô integral (see Theorem 5.3.3 in [6].) It is proved in Theorem 3.1 [8] that when f and ϕ are C^1-functions we have the equality:

$$\mathbb{E}\left[\left(\int_a^b f(B_t)\phi(B_b - B_t)\, dB_t\right)^2\right] = \int_a^b \mathbb{E}\Big[f(B_t)^2\phi(B_b - B_t)^2\Big]\, dt$$
$$+ 2\int_a^b \int_a^t \mathbb{E}\Big[f(B_s)\phi'(B_b - B_s)f'(B_t)\phi(B_b - B_t)\Big]\, ds\, dt, \tag{1.2}$$

provided that the integrals in the right-hand side exist. In particular, when $\phi \equiv 1$, the equality in equation (1.2) is the well-known Itô isometry.

We need to point out that the proof of equation (1.2) in [8] is too lengthy and involves rather tedious computations by using the binomial expansion. Moreover,

Received 2021-11-7; Accepted 2021-12-5; Communicated by the editors.
2020 *Mathematics Subject Classification.* Primary 60H05; Secondary 60H40.
Key words and phrases. Anticipating stochastic integral, Itô's isometry, Hitsuda–Skorokhod integral, white noise differential operator, forward-filtration, backward-filtration, counter-filtration, conditional expectation.
* Corresponding author.

Copyright information for Chapter 3 and 5

ON NEAR-MARTINGALES AND A CLASS OF ANTICIPATING LINEAR SDES

HUI-HSIUNG KUO, PUJAN SHRESTHA*, SUDIP SINHA, AND PADMANABHAN SUNDAR

Abstract. The primary goal of this paper is to prove a near-martingale optional stopping theorem and establish solvability and large deviations for a class of anticipating linear stochastic differential equations. We prove the existence and uniqueness of solutions using two approaches: (1) Ayed–Kuo differential formula using an ansatz, and (2) a novel braiding technique by interpreting the integral in the Skorokhod sense. We establish a Freidlin–Wentzell type large deviations result for solution of such equations.

1. Introduction

Anticipating stochastic calculus has been an active and important research area for several years, and lies at the intersection of probability theory and infinite-dimensional analysis. Enlargement of filtration, Malliavin calculus, and white noise theory provide three distinct methodologies to incorporate anticipation (of future) into classical Itô theory of stochastic integration and differential equations.

It is to the credit of Itô who constructed an anticipating stochastic integral in 1976[6], and laid the foundation for the idea of enlargement of the underlying filtration. Ever since, the method was embraced by several researchers that led to many important works (see articles in [7]). The advent of an integral invented by Skorokhod resulted in an impressive edifice built by Malliavin on stochastic calculus of variations in order to prove Hörmander's hypoellipticity result by stochastic analysis. Malliavin calculus provided a natural basis for the development and study of anticipative stochastic analysis and differential equations. Around the same time, a systematic study of Hida distributions gave rise to white noise theory and a general framework for stochastic calculus.

Malliavin calculus and white noise theory have vast applicability to the theory of stochastic differential equations with anticipation. However, the results obtained by these theories are primarily abstract though general. A more tractable theory was envisaged by Kuo based on a concrete stochastic integral known as the Ayed–Kuo integral[1]. Under less generality, the latter allows one to obtain results under easily understood, verifiable hypotheses.

In this article, we prove some results about stopped near-martingales, which are generalizations of martingales. We then study existence, uniqueness and large deviation principle for linear stochastic differential equations with anticipating initial conditions and drifts. While we rely mostly on the Ayed–Kuo formalism, other theories are minimally used either out of necessity, or to compare and contrast the conclusions of certain results.

2020 *Mathematics Subject Classification.* Primary 60H10, 60F10, 60G48, 60G40; Secondary 60H05, 60H07, 60H20.

Key words and phrases. anticipating integral, stochastic integral, stochastic differential equation, near-martingale, optional stopping theorem, large deviation principles.

* Corresponding author.

Bibliography

[1] Wided Ayed and Hui-Hsiung Kuo. An extension of the itô integral. *Communications on Stochastic Analysis*, 2(3):5, 2008.

[2] Louis Bachelier. Theory of speculation: The origins of modern finance. *Francia: Gauthier-Villars*, 1900.

[3] Robert Brown. A brief account of microscopical observations made in the months of june, july and august 1827, on the particles contained in the pollen of plants, and on the general existence of active molecules in organic and inorganic bodies. *Edin. Phil. J*, 5(358):3, 1828.

[4] Harald Cramér. On a new limit theorem of the theory of probability. *Uspekhi Matematicheskikh Nauk*, pages 166–178, 1944.

[5] Amir Dembo and Ofer Zeitouni. *Large Deviations Techniques and Applications*. Stochastic Modelling and Applied Probability. Springer Berlin Heidelberg, 2009.

[6] Albert Einstein. Über die von der molekularkinetischen theorie der wärme geforderte bewegung von in ruhenden flüssigkeiten suspendierten teilchen. *Annalen der physik*, 4, 1905.

[7] Leonard Gross. Abstract wiener spaces. Technical report, Cornell University, Ithaca, United States, 1967.

[8] Takeyuki Hida. Analysis of brownian functionals, carleton mathematical lecture notes 13. *Princeton University book*, 1975.

[9] Chii Ruey Hwang, Hui-Hsiung Kuo, and Kimiaki Saitô. Anticipating exponential processes and stochastic differential equations. *Communications on Stochastic Analysis*, 13(3):9, 2019.

[10] Chii-Ruey Hwang, Hui-Hsiung Kuo, Kimiaki Saitô, and Jiayu Zhai. A general itô formula for adapted and instantly independent stochastic processes. *Communications on Stochastic Analysis*, 10(3):5, 2016.

[11] Chii Ruey Hwang, Hui-Hsiung Kuo, Kimiaki Saitô, and Jiayu Zhai. Near-martingale property of anticipating stochastic integration. *Communications on Stochastic Analysis*, 11(4):6, 2017.

[12] Kiyosi Ito. Differential equations determining Markov processes. *Zenkoku Shijo Sugaku Dan-wakai*, 244(1077):1352–1400, 1942.

[13] Kiyosi Itô. *Extension of stochastic integrals*. Proceedings of the International Sympo-

sium on Stochastic Differential Equations, Kyoto, 1976. Wiley, 1978.

[14] Narjess Khalifa, Hui-Hsiung Kuo, Habib Ouerdiane, and Benedykt Szozda. Linear stochastic differential equations with anticipating initial conditions. *Communications on Stochastic Analysis*, 7(2):5, 2013.

[15] Hui-Hsiung Kuo. *Introduction to Stochastic Integration*. Universitext. Springer New York, 2006.

[16] Hui-Hsiung Kuo. The itô calculus and white noise theory: A brief survey toward general stochastic integration. *Communications on Stochastic Analysis*, 8(1):8, 2014.

[17] Hui-Hsiung Kuo. *White noise distribution theory*. CRC press, 2018.

[18] Hui-Hsiung Kuo, Anuwat Sae-Tang, and Benedykt Szozda. An isometry formula for a new stochastic integral. In *Quantum Probability and Related Topics*, pages 222–232. World Scientific, 2013.

[19] Hui-Hsiung Kuo, Pujan Shrestha, and Sudip Sinha. Anticipating linear stochastic differential equations with adapted coefficients. *Journal of Stochastic Analysis*, 2(2), 2021.

[20] Hui-Hsiung Kuo, Pujan Shrestha, and Sudip Sinha. An intrinsic proof of an extension of itô's isometry for anticipating stochastic integrals. *Journal of Stochastic Analysis*, 2(4), 2021.

[21] Hui-Hsiung Kuo, Pujan Shrestha, Sudip Sinha, and Padmanabhan Sundar. On near-martingales and a class of anticipating linear SDEs. *arXiv preprint arXiv:2204.01932*, 2022.

[22] Hui-Hsiung Kuo, Sudip Sinha, and Jiayu Zhai. Stochastic differential equations with anticipating initial conditions. *Communications on Stochastic Analysis*, 12(4):6, 2018.

[23] Henri Lebesgue. Intégrale, longueur, aire. *Annali di Matematica Pura ed Applicata (1898-1922)*, 7(1):231–359, 1902.

[24] Paul Malliavin. Stochastic calculus of variations and hypoelliptic operators. In *Proc. Internat. Symposium on Stochastic Differential Equations, Kyoto Univ., Kyoto, 1976*. Wiley, 1978.

[25] Paul Malliavin. *Stochastic Analysis*. Grundlehren der mathematischen Wissenschaften. Springer Berlin Heidelberg, 2015.

[26] David Nualart. *The Malliavin Calculus and Related Topics*. Probability and Its Applications. Springer Berlin Heidelberg, 2006.

[27] David Nualart. Malliavin calculus and normal approximations. *Ensaios Matemáticos*, 34:1–74, 2019.

[28] Peter Parczewski. Extensions of the hitsuda-skorokhod integral. *arXiv preprint arXiv:1711.06091*, 2017.

[29] Anatolii Volodimirovich Skorokhod. On a generalization of stochastic integral. *Teoriya Veroyatnostei I Ee Primeneniya*, 20(2):223–238, 1975.

[30] Hale F Trotter. On the product of semi-groups of operators. *Proceedings of the American Mathematical Society*, 10(4):545–551, 1959.

[31] Norbert Wiener. Differential-space. *Journal of Mathematics and Physics*, 2(1-4):131–174, 1923.

[32] David Williams. *Probability with Martingales*. Cambridge mathematical textbooks. Cambridge University Press, 1991.

[33] Jiayu Zhai. *General Stochastic Integral and Itô Formula with Application to Stochastic Differential Equations and Mathematical Finance*. PhD thesis, LSU Doctoral Dissertations, 2018.

Vita

Pujan Shrestha was born in Kathmandu, Nepal. He finished his secondary schooling at Shuvatara School and his post secondary schooling at Rato Bangala School. After graduating with his B.S. in mathematics and physics from Randolph College, he worked as a data analyst for the self driving car team at Google. As his interest in mathematics grew, he decided to enroll at Louisiana State University for his doctorate. Upon completion of his doctorate, he plans to spend considerable time playing with his nephews.