April 2022

Anticipating Stochastic Integrals and Related Linear Stochastic Differential Equations

Sudip Sinha
Louisiana State University

ANTICIPATING STOCHASTIC INTEGRALS AND RELATED LINEAR STOCHASTIC DIFFERENTIAL EQUATIONS

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
Sudip Sinha
Bachelor of Engineering, Birla Institute of Technology and Science, Pilani, 2012
Master of Science, MathMods, 2015
Master of Science, Louisiana State University, 2018
May 2022

This dissertation is dedicated to Professor Fabio Antonelli. He is instrumental in leading me into the intriguing universe of mathematics. Had I not met him, I would probably not have taken up research in this discipline.

In precisely built mathematical structures, mathematicians find the same sort of beauty others find in enchanting pieces of music, or in magnificent architecture. There is, however, one great difference between the beauty of mathematical structures and that of great art. Music by Mozart, for instance, impresses greatly even those who do not know musical theory; the cathedral in Cologne overwhelms spectators even if they know nothing about Christianity. The beauty in mathematical structures, however, cannot be appreciated without understanding of a group of numerical formulae that express laws of logic. Only mathematicians can read "musical scores" containing many numerical formulae, and play that "music" in their hearts. Accordingly, I once believed that without numerical formulae, I could never communicate the sweet melody played in my heart. Stochastic differential equations, called "Itô Formula", are currently in wide use for describing phenomena of random fluctuations over time. When I first set forth stochastic differential equations, however, my paper did not attract attention. It was over ten years after my paper that other mathematicians began reading my "musical scores" and playing my "music" with their "instruments". By developing my "original musical scores" into more elaborate "music", these researchers have contributed greatly to developing "Itô Formula". In recent years, I find that my "music" is played in various fields, in addition to mathematics. Never did I expect that my "music" would be found in such various fields, its echo benefiting the practical world, as well as adding abstract beauty to the field of mathematics.

— Kiyosi Itô
My Sixty Years along the Path of Probability Theory

Acknowledgments

This dissertation results from a culmination of years of effort, help from many people, a robust support system, and love for knowledge. I would like to express heartfelt gratitude to my advisors — Professors Hui-Hsiung Kuo and Padmanabhan Sundar. Without their timely guidance, encouragement, patience, kindness, empathy, and unwavering support, I would not be where I am today. I shall strive to uphold the strong work ethic that I have witnessed in them. I would also like to thank Professors Ambar Sengupta and Gerry Knapp for serving on my advisory committee.

I am honored that Pujan Shrestha chose to work with me throughout the doctoral journey. Having a colleague with whom I could work in tandem made me cherish my time here. I am thankful to him for being patient with my mathematical ramblings, research discussions, and bringing me out of procrastination loops.

I am indebted to my master's thesis advisor Professor Fabio Antonelli. It is from him that I started to understand and appreciate mathematics for what it is. His teaching and advice motivated me to pursue research in stochastic analysis. I also want to thank Professors Davide Gabrielli, Ida Germana Minelli, and Prasanna Kumar for teaching me and supporting my endeavors. I am grateful to Professor Jeffrey Roland for helping me explore the foundations and philosophies of mathematics and to Professor James Oxley for helping me refine my communication techniques. I thank Professors Shipman, Adkins, Dasbach, and the other professors and graduate students for creating a lively and healthy environment conducive to research.

None of this would have been possible without a sturdy support system. I am grateful to my family for their unconditional love and support. I would like to thank Bree, Kaju, Jessica, Rachel, Steven, Irfan, Liem, Huy, and all my friends who made living in Baton Rouge an enjoyable experience. My heartfelt gratitude to Ankan and Anirudh for lively discussions on mathematics, the universe, and everything. I am grateful to Tuppu for being a constant source of love, support, and inspiration.

Finally, my work builds on the mathematical framework laid by mathematicians and philosophers over millennia of painstaking effort. Quoting Isaac Newton, "If I have seen further, it is by standing on the shoulders of giants."

Contents

Abstract

Itô's stochastic calculus revolutionized the field of stochastic analysis and has found numerous applications in a wide variety of disciplines. Itô's theory, even though quite general, cannot handle anticipating stochastic processes as integrands. There have been considerable efforts within the mathematical community to extend Itô's calculus to account for anticipation. The Ayed–Kuo integral — introduced in 2008 — is one of the most recent developments. It is arguably the most accessible among the theories extending Itô's calculus — relying solely on probabilistic methods. In this dissertation, we look at the recent advances in this area, highlighting our contributions. First, we extend the class of linear stochastic differential equations with anticipating initial conditions that have closed-form solutions in this theory. Then we prove an extension of Itô's isometry for the Ayed–Kuo integral using purely probabilistic tools and exploiting the intrinsic nature of the integral. We also prove an optional stopping theorem for near-martingales — the counterpart of martingales in the Ayed–Kuo theory. We study the behavior of conditioned processes corresponding to the solution of a linear stochastic differential equation with anticipating initial conditions. We analyze a particular class of linear stochastic differential equations with anticipating coefficients in the drift term and derive the solution using two methods: (1) by guessing an ansatz and using the general differential formula in the Ayed–Kuo theory, and (2) by introducing a novel "braiding" technique that allows us to construct the solution by an iterative process. We derive explicit solutions in each theory independently and show that both methods yield the same solution under identical conditions. We establish a Freidlin–Wentzell type large deviations principle for solutions of this specific class of anticipating linear stochastic differential equations. Finally, we highlight a few areas of research in this field.

Chapter 1. Introduction

In 1827, while examining pollen grains suspended in water under a microscope, the British botanist Robert Brown observed that minute particles ejected from the pollens executed a continuous jittery motion[Bro28]. Since he observed the same kind of motion in inanimate objects like dust particles, he concluded that the motion was not related to life. He did not provide a theory to explain the phenomenon. However, this seemingly random motion came to be known as Brownian motion.

Surprisingly, Brownian motion did not get a lot of attention for more than half a century. In 1900, Louis Bachelier used Brownian motion in his PhD thesis “The theory of speculation”[Bac00], where he presented a stochastic analysis of stock and option markets. The phenomenon gained a lot of attention again in 1905 when Albert Einstein presented an indirect method to verify the existence of atoms and molecules in one of his papers[Ein05]. Around the same time, in 1902, Henri Lebesgue laid the foundation of measure theory and abstract integration theory[Leb02].

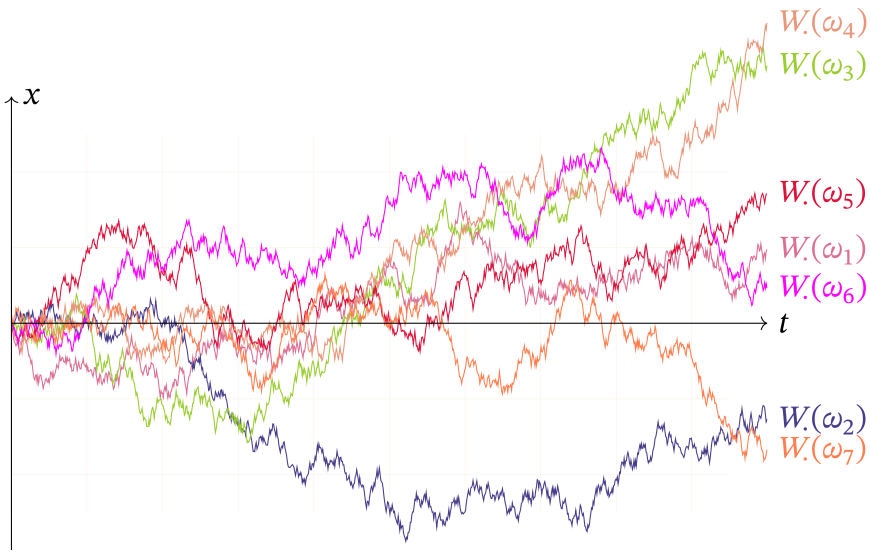

Figure 1.1. Some paths of a Wiener process.

However, there was no mathematical description of this phenomenon until this point. In 1923, the American mathematician Norbert Wiener showed how Brownian motions can be constructed in a mathematical sense by putting an appropriate measure on the space of continuous functions[Wie23], which is arguably the first really significant application of abstract measure theory. Note that this was a decade before Andreĭ Nikolaevich Kolmogorov laid the rigorous foundations of probability theory[Kol19]. Therefore, the stochastic process corresponding to Brownian motion process is popularly known as Wiener process in mathematics. Wiener also constructed an integral with respect to Wiener processes. We shall explore this idea in section 1.4.

Wiener’s integral allowed only integration of deterministic functions. In 1944, Kiyosi Itô published his seminal paper “Stochastic Integral”[Itô44], in which he demonstrated how to

integrate adapted stochastic processes. In layperson's terms, adapted processes are processes that can utilize information only until the present. Itô calculus forms the counterpart to the Leibniz–Newton calculus for random processes. Itô combined the ideas of the Riemann–Stieltjes integral (referring to the integrator) and the Lebesgue integral (referring to the integrand). This will be the focus of much of chapter 1, starting from section 1.5.

Since its inception, Itô's theory has been used in numerous scientific pursuits. One of the biggest achievements is the Black–Scholes–Merton model[BS73; Mer74], which is an application of Itô's calculus to model the dynamics of financial markets. Merton and Scholes received the 1997 Nobel Memorial Prize in Economic Sciences. Due to his death in 1995, Black was mentioned as a contributor but was not named as an awardee.

Itô's theory, even though quite general and useful, is not applicable to certain circumstances. For example, one cannot integrate "anticipating" processes, or processes that can have access to the future. Therefore, one cannot model insider trading using Itô's calculus. Similarly, if one knows the probability distribution of a road network being repaired at a certain future date and wants to model the logistics of delivering goods, the resulting model would fall outside the scope of Itô's theory.

There have been many attempts at enlarging the space of integrable stochastic processes. Some classic approaches are (1) enlargement of filtration due to Itô himself, (2) the Skorokhod integral and Malliavin calculus, (3) white noise distribution theory, and (4) the more recent Ayed–Kuo integral. In this dissertation, we mainly focus on the Ayed–Kuo integral. We look at the current state of research in this area. We also prove large deviation principles for anticipating linear stochastic differential equations defined in the Ayed–Kuo sense.

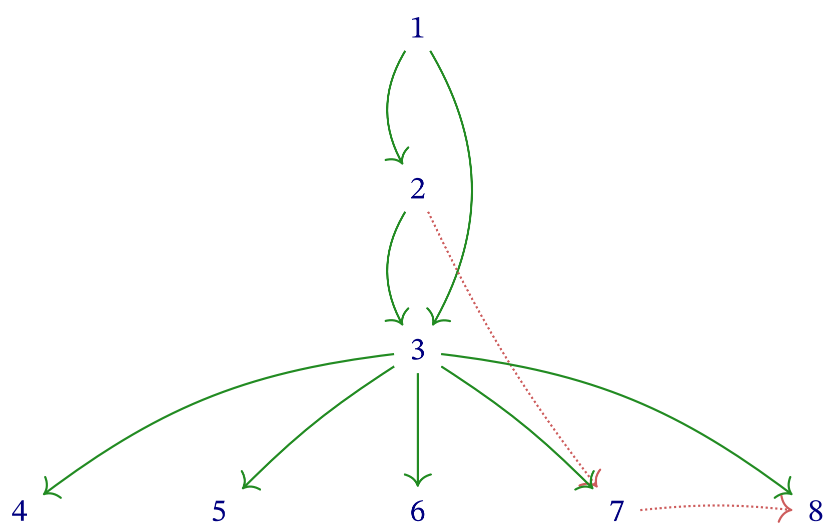

Figure 1.2. Chapter dependency graph. Dotted lines represent partial dependency.

This document is organized as follows. Chapter 1 introduces elementary ideas, martingales, stochastic integration, and stochastic differential equations. Chapter 2 discusses some mathematical motivations to extend Itô's calculus to a anticipating integrands, and a brief overview of the various approaches that have been taken in this direction. Chapter 3 introduces the

Ayed–Kuo integral, the central theme of this dissertation. In chapter 4, we prove an extension of Itô's isometry for Ayed–Kuo integrals. In chapter 5, we introduce near-martingales, and prove an optional stopping theorem that we use in chapter 8. In chapter 6, we discuss how conditionals of solutions of stochastic differential equations behave. In chapter 7, we focus on a particular class of stochastic differential equations and contrast different techniques to solve them. We also discuss existence and uniqueness results. In chapter 8, we derive Freidlin–Wentzell type large deviation principles for the class of stochastic differential equations discussed in chapter 7. Finally, in chapter 9, we conclude by looking at useful examples and open problems that need to be addressed.

In general, the document is designed to be read linearly. Chapter 2 can be skipped for the Ayed–Kuo part starting on chapter 3. However, the ideas of enlargement of filtration and Skorokhod integral (section 2.5) of chapter 2 are required for the corresponding sections in chapter 7. Chapters 4 to 6 are independent and can be read after chapter 3. The dependency graph of the chapters is shown in figure 1.2. Chapter 8 uses results from chapter 7, but is independent in essence. Chapter 9 is independent in spirit, though it refers to various results shown in the other chapters for context. See figure 1.2 for a dependency graph.

1.1. Elementary ideas and notation

A *probability space* is a triple $(\Omega, \Sigma, \mathbb{P})$ such that

1. Ω is a set.
2. Σ is a *σ-algebra* on Ω. This means $\Sigma \subseteq 2^{\Omega}$ such that
   (a) $\varnothing \in \Sigma$,
   (b) $E \in \Sigma$ implies $E^{\complement} \in \Sigma$, and
   (c) for every countable family of sets $(E_n)_{n=1}^{\infty}$ in Σ, the union $\bigcup E_n \in \Sigma$.
3. $\mathbb{P} : \Sigma \to [0,1]$ is a *probability measure*, that is
   (a) $\mathbb{P}(\Omega) = 1$, and
   (b) if $(E_n)_{n=1}^{\infty} \subseteq \Sigma$ is a mutually disjoint collection of sets, then $\mathbb{P}\big(\bigsqcup E_n\big) = \sum \mathbb{P}(E_n)$.

We usually think of a σ-algebra as the available information regarding the system of concern. A σ-algebra is called *complete* if all subsets of zero-measure sets are also measurable. In this document, we assume that all σ-algebras are complete. The σ-algebra generated by a topology is called *Borel σ-algebra*. Sometimes we suppress the filtration and write the probability space as $(\Omega, \mathbb{P})$ if the σ-algebra is the Borel σ-algebra. Elements of Σ are called *events*. We say that an event E occurs *almost surely* if $\mathbb{P}(E) = 1$.

A *random variable* X is a *measurable* function from a probability space $(\Omega, \Sigma, \mathbb{P})$ to the reals embedded with the Borel σ-algebra. That is, for every $B \in \mathcal{B}$, the set $\{X \in B\} = X^{-1}(B) \in \Sigma$. A random variable is called *integrable* if $\int |X| \,\mathrm{d}\mathbb{P} < \infty$. The *expectation* of an integrable random variable X is defined as the integral $\mathbb{E}(X) = \int_{\Omega} X \,\mathrm{d}\mathbb{P}$.

In order to model random phenomenon evolving in time, we define stochastic processes. A *stochastic process* is a function $X : \mathbb{T} \times \Omega \to \mathbb{R}$ such that

1. $X(t, \cdot) : \Omega \to \mathbb{R}$ is a random variable for each $t \in \mathbb{T}$, and

2. $X(\cdot, \omega) : \mathbb{T} \to \mathbb{R}$ is measurable for each $\omega \in \Omega$.

Stochastic processes are essentially families of random variables indexed by $t \in \mathbb{T}$. We think of the parameter t as representing time. We usually suppress the dependence on $\omega \in \Omega$ and simply write $X(t)$ or X_t. The space $(\mathbb{R}, \mathcal{B})$ is called the *state space* and $\mathbb{T}$ the *index set*. For a fixed ω, the function $X(\cdot, \omega) : \mathbb{T} \to \mathbb{R}$ is called a *path* of the process. We only consider the case where the $\mathbb{T} = [a, b]$, where $a, b \in \mathbb{R}$ and $a < b$. A stochastic process X is said to be *integrable* if the random variable X_t is integrable for each t.

We now present an axiomatic definition of a Wiener process — which shall be the focal point of this document.

Definition 1.1.1. A stochastic process $W = (W_t)_{t \geq 0}$ is called a *Wiener process* or *Brownian motion* if

1. $W_0 = 0$ a.s.
2. W has independent increments: for every $0 \leq t_1 \leq \cdots \leq t_n$, the random variables $W_{t_1}, W_{t_2} - W_{t_1}, \ldots, W_{t_n} - W_{t_{n-1}}$ are independent.
3. W has Gaussian increments: $W_t - W_s \sim N(0, |t - s|)$ for any s and t.
4. The paths of W are almost surely continuous in t.

Theorem 1.1.2 ([KS14, theorem 2.2.5]). *The paths of a Wiener process are almost surely nowhere differentiable.*

Often, we want to know the expected value of a process given restricted information. This is achieved by using conditional expectation. Suppose $\mathcal{F} \subseteq \Sigma$ is a sub-σ-algebra, and X is an integrable random variable. Then the *conditional expectation* of X given $\mathcal{F}$, denoted $\mathbb{E}(X \mid \mathcal{F})$ is a $\mathcal{F}$-measurable random variable Y such that $\int_E Y \, \mathrm{d}\mathbb{P} = \int_E X \, \mathrm{d}\mathbb{P}$ for any $E \in \mathcal{F}$.

Theorem 1.1.3 (properties of conditional expectation). *Let X, Y be random variables, a, b scalars, and $\mathcal{G} \subseteq \mathcal{F} \subseteq \Sigma$ σ-algebras. Then under suitable integrability conditions, the following hold almost surely.*

1. *Mean-preserving:* $\mathbb{E}(\mathbb{E}(X \mid \mathcal{F})) = \mathbb{E}(X)$.
2. *Linearity:* $\mathbb{E}(aX + bY \mid \mathcal{F}) = a\,\mathbb{E}(X \mid \mathcal{F}) + b\,\mathbb{E}(Y \mid \mathcal{F})$.
3. *Monotonicity: If* $X \geq 0$*, then* $\mathbb{E}(X \mid \mathcal{F}) \geq 0$.
4. *Taking out the known: If Y is $\mathcal{F}$-measurable, then* $\mathbb{E}(XY \mid \mathcal{F}) = Y\mathbb{E}(X \mid \mathcal{F})$.
5. *Independence: If $\mathcal{G}$ is independent of $\sigma\{\sigma(X), \mathcal{F}\}$, then* $\mathbb{E}(X \mid \sigma(\mathcal{F}, \mathcal{G})) = \mathbb{E}(X \mid \mathcal{F})$. *In particular, if X is independent of $\mathcal{H}$, then* $\mathbb{E}(X \mid \mathcal{G}) = \mathbb{E}(X)$.
6. *Tower property:* $\mathbb{E}(\mathbb{E}(X \mid \mathcal{F}) \mid \mathcal{G}) = \mathbb{E}(X \mid \mathcal{G})$.

For the proof of the above, see [Wil91, section 9.7].

An non-decreasing indexed family of σ-algebras $\mathcal{F} = (\mathcal{F}_t)_{t \in \mathbb{T}}$, is called a *filtration*. A probability space embedded with a filtration is called a *filtered probability space* and is denoted by $(\Omega, \Sigma, \mathcal{F}, \mathbb{P})$. A filtration is said to be *right-continuous* if $\mathcal{F}_{t^+} = \mathcal{F}_t$, where $\mathcal{F}_{t^+} = \bigcap_{u>t} \mathcal{F}_u$. A filtered probability space is said to satisfy the *usual conditions* if the associated filtration is complete and right-continuous. In the information interpretation, a filtered probability space represents the temporal evolution of knowledge about the system. We refer to the filtration generated by the relevant Wiener process as the *natural filtration*.

Definition 1.1.4. A stochastic process A is called *$\mathcal{F}$-adapted* if A_t is $\mathcal{F}_t$-measurable for every t. If the σ-algebra is obvious, we simply say *adapted*.

A Wiener process is trivially adapted to its natural filtration.

Now, we define stopping times.

Definition 1.1.5. A stochastic process τ is called a *stopping time* if $\{\tau \leq t\} \in \mathcal{F}_t$ for every t. The σ-algebra at a stopping time τ is defined as

$$\mathcal{F}_\tau = \{E \in \Sigma \mid E \cap \{\tau \leq t\} \in \mathcal{F}_t \text{ for all } t\}.$$

A process "stopped" using a stopping time has a constant value after it is stopped.

Definition 1.1.6. Suppose X is a stochastic process and τ is a stopping time. Then the *stopped process X^τ* is defined by

$$X_t^\tau(\omega) = X_{\tau(\omega)\wedge t}(\omega),$$

where we use the notation $a \wedge b = \min\{a, b\}$.

Conventions In this document, unless specified, we fix the following.

1. An underlying filtered space $(\Omega, \Sigma, \mathcal{F}, \mathbb{P})$ satisfying the usual conditions.
2. The interval of interest of the parameter t is $\mathbb{T} = [a, b]$, where $0 \leq a \leq b < \infty$.
3. A Wiener process $W = (W_t)_{t\geq 0}$.
4. The natural filtration as the reference filtration.

1.2. Martingales

Definition 1.2.1. An integrable adapted stochastic process M is called a *martingale* if $\mathbb{E}(M_t \mid \mathcal{F}_s) = M_s$ for every $s \leq t$. It is called a *submartingale* if we replace the condition with $\mathbb{E}(M_t \mid \mathcal{F}_s) \geq M_s$, and is called a *supermartingale* when $\mathbb{E}(M_t \mid \mathcal{F}_s) \leq M_s$.

From a modeling perspective, the best predictor of the future value of a martingale is its present value. Therefore, martingales are used to model fair games and find numerous application in pricing financial products.

A process that is both a submartingale and a supermartingale is necessarily a martingale. Moreover, any result that is true for a submartingale can be suitably modified for a supermartingale, and subsequently for a martingale. Therefore, in what follows, we only show results for submartingales.

Let us look at some examples of martingales.

1. Suppose $(X_n)_{n=1}^\infty$ is a sequence of integrable independent zero-mean random variables. Let $S_n = \sum_{i=1}^n X_i$. Then S is a martingale.
2. Suppose $(X_n)_{n=1}^\infty$ is a sequence of integrable, independent and identically distributed (i.i.d.) random variables with zero mean and unit variance, and $S_n = \sum_{i=1}^n X_i$. Then $S_n^2 - n$ is a martingale.

3. If ξ is integrable, then the process $M_t = \mathbb{E}(\xi \mid \mathcal{F}_t)$ is a martingale.
4. A Wiener process W is a martingale.
   Proof. Suppose $s \leq t$. We write $W_t = W_s + (W_t - W_s)$. By independence of increments, we know $W_t - W_s$ is independent of $\mathcal{F}_s$. Moreover, $\mathbb{E}(W_t - W_s) = 0$. Using theorem 1.1.3, we get $\mathbb{E}(W_t \mid \mathcal{F}_s) = \mathbb{E}(W_s \mid \mathcal{F}_s) + \mathbb{E}(W_t - W_s \mid \mathcal{F}_s) = W_s + \cancelto{0}{\mathbb{E}(W_t - W_s)} = W_s$. □
5. The processes $W_t^2 - t$ and $W_t^3 - 3tW_t$ are martingales.
6. $\exp\left(W_t^2 - \frac{1}{2}t\right)$ is a martingale.

Since the expected value of a submartingale increases, it is natural to ask if we can decompose it into a martingale and an increasing process. The following theorem answers this question.

Theorem 1.2.2 (Doob–Meyer decomposition)**.** *Suppose X is a integrable adapted process. Then there is a unique decomposition*

$$X_t = M_t + A_t \quad \textit{for all } t,$$

where M is a martingale and A is a adapted integrable process.

Theorem 1.2.3 (Doob–Meyer decomposition for submartingales)**.** *Suppose X is a submartingale. Then there is a unique decomposition*

$$X_t = M_t + A_t \quad \textit{for all } t,$$

where M is a martingale and A is a adapted integrable increasing process starting at 0*.*

The power of martingales come from martingale transforms. We think of this as follows. Imagine a discrete-time game of chance, where the difference $X_n - X_{n-1}$ is our winning per unit stake at time $n \geq 1$. At time 0, we place a random stake A_0 using our prior intuition. Our winning at time 1 is then $A_0(X_1 - X_0)$. At time 1, we update our prior to the present and place stake A_1 to get a winning of $A_1(X_2 - X_1)$ at time 2. Therefore, the net gain until time 2 is $\sum_{i=1}^{2} A_{n-1}(X_n - X_{n-1})$. In general, at time n, our portfolio value would be $\sum_{i=1}^{n} A_{n-1}(X_n - X_{n-1})$. This is the motivation behind the following definition.

Definition 1.2.4. Let $(A_n)_{n=0}^{\infty}$ be an adapted process and $(X_n)_{n=0}^{\infty}$ a submartingale. Then the processes $(Y_n)_{n=0}^{\infty}$, where $Y_0 = 0$ and

$$Y_n = (A \bullet X)_n = \sum_{i=1}^{n} A_{n-1}(X_n - X_{n-1})$$

is called the *martingale transform* of X by A.

Note that we required A to be adapted so that the bets are fair in the sense that one does not know the result of the lottery at any future time. Since we have no information of the future, if the underlying game is fair (martingale), no matter how we bet, we should neither win nor lose in expectation. This is what we now show.

Proposition 1.2.5.

1. *If X is a submartingale and A is a bounded non-negative adapted process, then $(A \bullet X)$ is a submartingale.*

2. *If X is a martingale and A is a bounded adapted process, then $(A \bullet X)$ is a martingale.*
3. *If X and A are both square integrable, then we do not require the boundedness condition in items 1 and 2.*

Proof. We only prove item 1 because the others follow the same process. Let X be a submartingale and $Y = (A \bullet X)$. Suppose n is an arbitrary time. Note that $Y_n - Y_{n-1} = A_{n-1}(X_n - X_{n-1})$, which is integrable since A is bounded. Using the adaptedness of A, we get

$$\mathbb{E}(Y_n - Y_{n-1} \mid \mathcal{F}_{n-1}) = \mathbb{E}(A_{n-1}(X_n - X_{n-1}) \mid \mathcal{F}_{n-1}) = A_{n-1}\,\mathbb{E}(X_n - X_{n-1} \mid \mathcal{F}_{n-1}) \geq 0,$$

where the last inequality holds since A is non-negative. □

Given a martingale, one can expect that the martingale property is preserved when sampling on stopping times. This always holds [for the discrete case, see Wil91, page 99]. However, it is not always true that the expected value of the stopped process is the same as the expected value at the initial time. For example, let X be a simple random walk on the natural numbers (including 0) starting at 0 and let $\tau = \inf\{X_n = 1\}$. Then X is a martingale and it is well-known that $\mathbb{P}(T < \infty) = 1$. However, even though $\mathbb{E}(X_{\tau\wedge n}) = \mathbb{E}(X_0)$ for every n, we have $1 = \mathbb{E}(X_\tau) \neq \mathbb{E}(X_0) = 0$. Doob's optional stopping theorem states that this is true under reasonable conditions.

Theorem 1.2.6 ([KS14, theorem 3.3.1])**.** *Let M be a uniformly integrable submartingale with right-continuous sample paths. Suppose σ and τ are two bounded stopping times with $\sigma \leq \tau$. Then M_σ and M_τ are integrable, and $\mathbb{E}(X_\tau \mid \mathcal{F}_\sigma) \geq X_\sigma$ almost surely.*

1.3. Naively approaching stochastic integration

Our goal in this section is to define integrals with respect to Wiener processes. That is, we want to define $\int_a^b f(t)\,\mathrm{d}W_t$. Can we do so in the usual Riemann–Stieltjes sense?

A set $\Pi_n = \{t_0, t_1, \dots, t_n\}$, is called a *partition* of $[a, b]$ if $a = t_0 < t_1 < \cdots < t_n = b$. We define the norm of a partition as $\|\Pi_n\| = \sup\{t_i - t_{i-1} \mid i \in [n]\}$. In this document, we always assume $\|\Pi_n\| \to 0$ as $n \to \infty$.

We first introduce the idea of variations of a real-valued function f.

Definition 1.3.1. The *linear variation* of a function $f : [a, b] \to \mathbb{R}$ is the quantity

$$L_a^b(f) = \sup_{\Pi_n} \sum_{i=0}^{n} |f(x_i) - f(x_{i-1})|\,,$$

where the supremum is taken over all partitions of $[a, b]$.

Similarly, the *quadratic variation* of a function $f : [a, b] \to \mathbb{R}$ is the quantity

$$Q_a^b(f) = \sup_{\Pi_n} \sum_{i=0}^{n} |f(x_i) - f(x_{i-1})|^2\,.$$

It can easily be checked that the quadratic variation of a Wiener process is given by

$$Q_a^b(W) = b - a, \tag{1.3.2}$$

which we symbolically denote by $(\Delta W_i)^2 \simeq \Delta t_i$.

Now, suppose $\|\Pi_n\| \to 0$ as $n \to \infty$. Then the Riemann–Stieltjes integral for f with respect to g is defined as

$$\int_a^b f(t)\,\mathrm{d}g(t) = \lim_{n\to\infty} \sum_{i=1}^{n} f(t_i^*)\left(g(t_i) - g(t_{i-1})\right),$$

where $t_i^* \in [t_{i-1}, t_i]$ for $i \in [n] = \{1, 2, \dots, n\}$. It can be shown that continuous functions are Riemann–Stieltjes integrable with respect to any function of finite linear variation.

Can we use this idea to define $\int_a^b f(t)\,\mathrm{d}W_t$ for a continuous function f? Since the linear variation of a Wiener process over any interval is infinite, we cannot define path-wise a stochastic integral of the form $\int_a^b f(t)\,\mathrm{d}W_t(\omega)$. In other words, naive integration with respect to a Wiener process is not possible.

1.4. Wiener's integral

As we remarked in equation (1.3.2), Wiener processes have finite quadratic variations. Can we exploit this fact to define an integral with respect to Wiener processes?

Wiener[Wie23] did exactly this for square-integrable deterministic functions $f \in L^2[a, b]$. His approach was to first define the integral on *step functions*. In particular, for $f(t) = \sum_{i=1}^n a_i \mathbb{1}_{[t_{i-1}, t_i)}(t)$, define the *Wiener integral* as

$$I_W(f) = \sum_{i=1}^{n} a_i\,\Delta W_i,$$

where $\Delta W_i = W_{t_i} - W_{t_{i-1}}$. In order to extend this idea to $L^2[a, b]$, we note that step functions are dense in the space of square integrable functions. Therefore, for any $f \in L^2[a, b]$, we can find a sequence of functions $\left(f^{(n)}\right)_{n=1}^{\infty}$ such that $f^{(n)} \to f$ in $L^2[a, b]$ as $n \to \infty$. It can be shown that $\left(I(X^{(n)})\right)_{n=1}^{\infty}$ is a Cauchy sequence in $L^2(\Omega)$. Since $L^2(\Omega)$ is a Hilbert space, the sequence $\left(I(f^{(n)})\right)_{n=1}^{\infty}$ converges in $L^2(\Omega)$. We call this limit as the Wiener integral of f, and write

$$I_W(f) = \int_a^b f(t)\,\mathrm{d}W_t = \lim_{n\to\infty} \int_a^b f^{(n)}(t)\,\mathrm{d}W_t.$$

The Wiener integral is linear and retains the Gaussian nature of Wiener processes. In particular,

$$I_W(f) \sim N(0, \|f\|),$$

where $\|f\|^2 = \int_a^b f(t)^2\,\mathrm{d}t$ is the canonical norm on $L^2[a, b]$. Moreover, the process X defined by $X_t = \int_a^t f(s)\,\mathrm{d}W_s$ is a continuous martingale.

1.5. Itô's integral

Itô generalized Wiener's integral to allow integrands to be adapted stochastic processes, thus expanding the class of integrable functions significantly [Itô44]. Itô's original motivation was to construct diffusion processes associated with infinitesimal operators directly in a probabilistic manner. However, we take a martingale viewpoint in this dissertation. To understand a motivation of his definition, we first try to compute the integral of a Wiener process with respect to itself. Writing as limits of Riemann sums,

$$\int_0^t W_s \,\mathrm{d}W_s = \lim_{n\to\infty} \sum_{i=1}^n W_{t_i^*} \,\Delta W_i,$$

where $t_i^* \in [t_{i-1}, t_i]$.

Now, define the following Riemann sums by taking the evaluation points as the left and right endpoints, respectively, of each subinterval:

$$L_n = \sum_{i=1}^n W_{t_{i-1}} \,\Delta W_i \qquad R_n = \sum_{i=1}^n W_{t_i} \,\Delta W_i.$$

Then we can easily see that

$$\begin{aligned} R_n + L_n &= W_t^2, \qquad \text{and} \\ R_n - L_n &= \sum_{i=1}^n \left(W_{t_{i-1}} - W_{t_i}\right)^2. \end{aligned}$$

Using the quadratic variation of Wiener processes and taking limit $n \to \infty$ in $L^2(\Omega)$, we get

$$\begin{aligned} L_n &= \frac{1}{2}\left(W_t^2 - \sum_{i=1}^n \left(W_{t_{i-1}} - W_{t_i}\right)^2\right) \to \frac{1}{2}\left(W_t^2 - t\right), \text{ and} \\ R_n &= \frac{1}{2}\left(W_t^2 + \sum_{i=1}^n \left(W_{t_{i-1}} - W_{t_i}\right)^2\right) \to \frac{1}{2}\left(W_t^2 + t\right). \end{aligned}$$

Therefore, we see that different choices for t_i^* give us different values of the integral. In table 1.1, we show the results of taking three representatives — the left, right, and midpoints.

Clearly, there is no single optimal choice. Therefore, one must choose according to one's convenience. There are two noticeable advantages of choosing the left-endpoint. First, if one thinks of the interval as time, then taking the left-endpoint signifies that the integrand does not "know" of the future behavior of the integrator. Therefore, it is expected that the left-endpoint approximation will give us a martingale, which it does. Therefore, Itô's integral can be thought of as a continuous martingale transform. Keeping this in mind, we give the formal definition of Itô's integral.

Table 1.1. Choice of endpoints for $\int_0^t W_s \,\mathrm{d}W_s$

t_j^*	$\int_0^t W_s \,\mathrm{d}W_s$	Intuitive?	$\mathbb{E}$	Martingale?	Theory
left	$\frac{1}{2}\left(W_t^2 - t\right)$	×	0	✓	Itô
mid	$\frac{1}{2}\left(W_t^2\right)$	✓	t	×	Stratonovich
right	$\frac{1}{2}\left(W_t^2 + t\right)$	×	t	×	

Just like Wiener's definition of the integral, Itô defines the stochastic integral in steps. First, consider *step processes* of the form $X_t(\omega) = \sum_{i=1}^n \xi_{i-1}(\omega)\mathbb{1}_{[t_{i-1},t_i)}(t)$, where ξ_{i-1} is $\mathcal{F}_{t_{i-1}}$-measurable for each i. The Itô integral for f is given by

$$I(X) = \sum_{i=1}^{n} \xi_{i-1}\,\Delta W_i.$$

Let $L^2_{\mathrm{ad}}([a,b] \times \Omega)$ denote the space of square-integrable adapted stochastic processes, that is

$$L^2_{\mathrm{ad}}([a,b] \times \Omega) = \left\{ X \,\middle|\, \int_a^b \mathbb{E}\left(|X_t|^2\right) \mathrm{d}t < \infty \right\}.$$

Note that $L^2_{\mathrm{ad}}([a,b] \times \Omega)$ is a Hilbert space with the inner product $\langle X, Y\rangle = \int_a^b \mathbb{E}(X_t Y_t)\,\mathrm{d}t$. Similarly, define

$$L^1_{\mathrm{ad}}([a,b] \times \Omega) = \left\{ X \,\middle|\, \int_a^b \mathbb{E}|X_t|\,\mathrm{d}t < \infty \right\}.$$

Again, $L^1_{\mathrm{ad}}([a,b] \times \Omega)$ is a Banach space with the norm $\|X\| = \int_a^b \mathbb{E}|X_t|\,\mathrm{d}t$.

It can be shown that the set of step processes is dense in the set $L^2_{\mathrm{ad}}([a,b] \times \Omega)$. Therefore, for every $X \in L^2_{\mathrm{ad}}([a,b] \times \Omega)$, there is a sequence of step processes $\left(X^{(n)}\right)_{n=1}^{\infty}$ such that $X^{(n)} \to X$ in $L^2_{\mathrm{ad}}([a,b] \times \Omega)$ as $n \to \infty$. It can be shown that $\left(I(X^{(n)})\right)_{n=1}^{\infty}$ is a Cauchy sequence in $L^2(\Omega)$. From the completeness of $L^2(\Omega)$, we are assured the convergence of $\left(I(X^{(n)})\right)_{n=1}^{\infty}$ in $L^2(\Omega)$. We call this limit as the Itô integral of X, and write

$$I(X) = \int_a^b X_t \,\mathrm{d}W_t = \lim_{n\to\infty} \int_a^b X_t^{(n)} \,\mathrm{d}W_t.$$

The salient idea that made this process work is the isometry between $L^2_{\mathrm{ad}}([a,b] \times \Omega)$ and $L^2(\Omega)$.

Theorem 1.5.1 (Itô's isometry). *For every $X \in L^2_{ad}([a,b] \times \Omega)$, we have $\mathbb{E}\left(\int_a^b X_t \,\mathrm{d}W_t\right) = 0$, and*

$$\mathbb{E}\left[\left(\int_a^b X_t \,\mathrm{d}W_t\right)^2\right] = \int_a^b \mathbb{E}\left(X_t^2\right) \mathrm{d}t.$$

The Itô integral is linear. Moreover, the process Y defined by $Y_t = \int_a^t X_s \,\mathrm{d}W_s$ is a continuous martingale.

An *Itô process* is an adapted stochastic process that can be expressed as the sum of an integral with respect to time (called the *drift*) and an Itô integral (called the *diffusion*). That is, a process X is an Itô process if and only if there exists an $\mathcal{F}_0$-measurable random variable X_0 and stochastic processes $m \in L^1_{\mathrm{ad}}([a,b]\times\Omega)$ and $\sigma \in L^2_{\mathrm{ad}}([a,b]\times\Omega)$ such that

$$X_t = X_0 + \int_a^t m_s \,\mathrm{d}s + \int_a^t \sigma_s \,\mathrm{d}W_s.$$

It is customary to write this in the symbolic differential form

$$\mathrm{d}X_t = m_t \,\mathrm{d}t + \sigma_t \,\mathrm{d}W_t.$$

Itô's lemma, also known as *Itô's formula*, gives a method to obtain Itô processes from existing Itô processes.

Theorem 1.5.2 (Itô's formula)**.** *Suppose X be an Itô process. Let $f(t,x) : [a,b]\times\mathbb{R} \to \mathbb{R}$ be C^1 in the t-coordinate and C^2 in the x-coordinate with partial derivatives f_t, f_x, and f_{xx}. Then*

$$Y_t = f(t, X_t)$$

is also an Itô process, and in the symbolic differential form, we have

$$\mathrm{d}Y_t = f_t \,\mathrm{d}t + f_x \,\mathrm{d}X_t + \frac{1}{2} f_{xx} \left(\mathrm{d}X_t\right)^2.$$

where $\left(\mathrm{d}X_t\right)^2 = \mathrm{d}X_t \cdot \mathrm{d}X_t$ is calculated using the rules

$$\mathrm{d}W_t \cdot \mathrm{d}W_t = \mathrm{d}t, \text{ and } \mathrm{d}t \cdot \mathrm{d}t = \mathrm{d}t \cdot \mathrm{d}W_t = \mathrm{d}W_t \cdot \mathrm{d}t = 0.$$

Example 1.5.3. Towards the beginning of this section, we saw a special case of the Itô integral

$$\int_a^t W_s \,\mathrm{d}W_s = \frac{1}{2}\left(({W_t}^2 - {W_a}^2) - (t-a)\right).$$

Here we show another method of deriving this using theorem 1.5.2.

Recall that ordinary calculus gives $\int_a^t x \,\mathrm{d}x = x^2 - a^2$. Therefore, we expect the Itô integral to give a similar term. So we guess $f(x) = x^2$, which means $f'(x) = 2x$ and $f''(x) = 2$. Using theorem 1.5.2 on $f(W_t)$, we get

$$\begin{aligned}\mathrm{d}\left({W_t}^2\right) &= 2W_t \,\mathrm{d}W_t + \frac{1}{2}\cdot 2\left(\mathrm{d}W_t\right)^2 \\ &= 2W_t \,\mathrm{d}W_t + \mathrm{d}t.\end{aligned}$$

Rearranging and writing the integral form, we get

$$\int_a^t W_s \,\mathrm{d}W_s = \frac{1}{2}\left(({W_t}^2 - {W_a}^2) - (t-a)\right),$$

which is exactly what we expected.

1.6. Stochastic differential equations

Ordinary differential equations are ubiquitous in mathematical modeling of physical phenomena. For example, suppose we want to model the value S of a fixed-return asset (such as a bond) whose growth rate is proportional to its value at any time t. We can write this as

$$\frac{\mathrm{d}S_t}{\mathrm{d}t} = p_t S_t.$$

If we try to do the same thing for a variable-return asset (such as a stock), we have to incorporate randomness into the equation. We can do this by adding a *noise* term $\dot{W}_t$ in the coefficient. That is, we can write

$$p_t = m_t + \dot{W}_t$$

We only know the probabilistic behavior of the noise term, not its exact behavior.

In general, we would like to solve equations of the form

$$\frac{\mathrm{d}X_t}{\mathrm{d}t} = m(t, X_t) + \sigma(t, X_t)\dot{W}_t. \tag{1.6.1}$$

Based on our experience of physical phenomenon, we expect $\dot{W}$ to have the following properties:

1. $\dot{W}_{t_1}$ and $\dot{W}_{t_2}$ are independent if $t_1 \neq t_2$.
2. $\dot{W}$ is stationary, that is the joint distribution of $\{\dot{W}_{t_1+t}, \dots, \dot{W}_{t_k+t}\}$ does not depend on t for any $k \in \mathbb{N}$.
3. $\mathbb{E}(\dot{W}_t) = 0$ for all t.

However, it is well-known that there does not exist a stochastic process satisfying the above properties in the usual sense. It is possible to represent $\dot{W}$ as a *generalized* stochastic process called a *white noise process*. This means that it can be constructed as a probability measure on the larger space $\mathcal{S}'$ of tempered distributions on $[0, \infty)$. Compare this to the usual stochastic processes for which the probability measure is on the smaller space $\mathbb{R}^{[0,\infty)}$. This construction is mathematically involved and we refer the reader to [Kuo96] for more details.

So we are left with the question of whether we can still give a meaning to equation (1.6.1) in our usual framework. Now, the properties of $\dot{W}$ are similar to what we would expect the derivative of a Wiener process to be. But we know from theorem 1.1.2 that the paths of Wiener processes are almost surely nowhere differentiable. Nevertheless, this gives us an idea of how to move forward.

Note that if we "multiply" both sides of the equation (1.6.1) by $\mathrm{d}t$, we get

$$\mathrm{d}X_t = m(t, X_t)\,\mathrm{d}t + \sigma(t, X_t)\left(\dot{W}_t\,\mathrm{d}t\right).$$

Now, if we replace $\dot{W}_t\,\mathrm{d}t$ with $\mathrm{d}W_t$, where W is the usual Wiener process, we get the *stochastic differential equation* (SDE)

$$\mathrm{d}X_t = m(t, X_t)\,\mathrm{d}t + \sigma(t, X_t)\,\mathrm{d}W_t \tag{1.6.2}$$

Note that $\mathrm{d}W_t$ itself is meaningless since the paths of Wiener processes are almost surely nowhere differentiable. However, if we write the integral version of this, we can interpret the last term as the usual Itô's integral. Therefore, we have the *stochastic integral equation*

$$X_t = X_0 + \int_a^t m(s, X_s)\,\mathrm{d}s + \int_a^t \sigma(s, X_s)\,\mathrm{d}W_s, \tag{1.6.3}$$

which is meaningful. Since equation (1.6.2) is easier to read than equation (1.6.3), we define equation (1.6.2) to be a symbolic representation of equation (1.6.3). Whenever we write a stochastic differential equation, we shall keep in mind that we are referring to the associated stochastic integral equation.

A stochastic differential equation is called linear if in equation (1.6.2), we have $m(t, X_t) = \gamma_t X_t + \rho_t$ and $\sigma(t, X_t) = \alpha_t X_t + \beta_t$ for some adapted processes $\gamma, \rho, \alpha, \beta$. In this document, we shall only look at linear SDEs (LSDEs).

Let us now look at a two classic examples of linear stochastic differential equations and their solutions.

Example 1.6.4. Given $t \in [0, \infty)$, consider the equation $\frac{\mathrm{d}X_t}{\mathrm{d}t} = -\theta X_t + \sigma \dot{W}_t$, where $\theta, \sigma > 0$. This is known as the *Langevin equation* in physics. Since $\dot{W}$ does not exist, this equation is strictly symbolic. However, it can be made meaningful by writing

$$\begin{cases} \mathrm{d}X_t = -\theta X_t\,\mathrm{d}t + \sigma\,\mathrm{d}W_t, \\ X_0 = x_0 \in \mathbb{R}. \end{cases}$$

To solve this, we use Itô's formula along with integration by parts. The integrating factor is $e^{\int_0^t \theta\,\mathrm{d}s} = e^{\theta t}$. Therefore, let $f(t, x) = e^{\theta t}x$. Then the derivatives are $f_t(t, x) = \theta e^{\theta t}x = \theta f(t, x)$, $f_x(t, x) = e^{\theta t}$, and $f_{xx}(t, x) = 0$. By Itô's formula

$$\begin{aligned} \mathrm{d}\big(e^{\theta t}X_t\big) &= \theta e^{\theta t}X_t\,\mathrm{d}t + e^{\theta t}\,\mathrm{d}X_t \\ &= \cancel{\theta e^{\theta t}X_t\,\mathrm{d}t} + e^{\theta t}\left(\cancel{-\theta X_t\,\mathrm{d}t} + \sigma\,\mathrm{d}W_t\right) \\ &= e^{\theta t}\sigma\,\mathrm{d}W_t. \end{aligned}$$

Using this, the stochastic differential equation now becomes

$$\mathrm{d}\big(e^{\theta t}X_t\big) = e^{\theta t}\sigma\,\mathrm{d}W_t.$$

Converting into the integral form, we get

$$e^{\theta t}X_t - x_0 = \int_0^t e^{\theta s}\sigma\,\mathrm{d}W_s,$$

which gives us

$$X_t = e^{-\theta t}x_0 + \sigma\int_0^t e^{-\theta(t-s)}\,\mathrm{d}W_s.$$

The solution is known as the *Ornstein–Uhlenbeck process*. A minor variation of this is known as the *Vasicek model* in mathematical finance and is used to model the evolution of interest rates.

Example 1.6.5. Given $t \in [a, b]$, consider the stochastic differential equation

$$\begin{cases} \mathrm{d}X_t = m(t)X_t\,\mathrm{d}t + \sigma(t)X_t\,\mathrm{d}W_t \\ X_a = 1, \end{cases}$$

where m and σ are deterministic functions.

To solve this, we let $\theta(x) = \log(x)$, so $\theta'(x) = \frac{1}{x}$ and $\theta''(x) = -\frac{1}{x^2}$. Using Itô's formula, we get

$$\mathrm{d}(\log X_t) = \frac{1}{X_t}\,\mathrm{d}X_t - \frac{1}{2}\frac{1}{X_t^2}\,(\mathrm{d}X_t)^2.$$

Now, $(\mathrm{d}X_t)^2 = \sigma(t)^2X_t^2\,\mathrm{d}t$ since $(\mathrm{d}W_t)^2 = \mathrm{d}t$, the rest of the terms being zero. Therefore,

$$\begin{aligned} \mathrm{d}(\log X_t) &= \frac{1}{\cancel{X_t}}\big(m(t)\cancel{X_t}\,\mathrm{d}t + \sigma(t)\cancel{X_t}\,\mathrm{d}W_t\big) - \frac{1}{2}\frac{1}{\cancel{X_t^2}}\sigma(t)^2\cancel{X_t^2}\,\mathrm{d}t \\ &= \Big(m(t) - \frac{1}{2}\sigma(t)^2\Big)\,\mathrm{d}t + \sigma(t)\,\mathrm{d}W_t. \end{aligned}$$

Integrating from a to t, we get

$$\log X_t - \log X_a = \int_a^t \Big(m(t) - \frac{1}{2}\sigma(t)^2\Big)\,\mathrm{d}t + \int_a^t \sigma(t)\,\mathrm{d}W_t.$$

Since $\log X_a = \log 1 = 0$, we get

$$X_t = \exp\left(\int_a^t \Big(m(t) - \frac{1}{2}\sigma(t)^2\Big)\,\mathrm{d}t + \int_a^t \sigma(t)\,\mathrm{d}W_t\right).$$

Example 1.6.5 is an example of an important class of processes called exponential processes.

Definition 1.6.6. The *exponential process* associated with adapted stochastic processes m and σ is defined as

$$\mathcal{E}_t^{(\sigma,m)} = \exp\left(\int_a^t \sigma_s\,\mathrm{d}W_s + \int_a^t \Big(m_s - \frac{1}{2}\sigma_s^2\Big)\,\mathrm{d}s\right).$$

If $m \equiv 0$, then we write

$$\mathcal{E}_t^{(\sigma)} = \exp\left(\int_a^t \sigma_s\,\mathrm{d}W_s - \frac{1}{2}\int_a^t \sigma_s^2\,\mathrm{d}s\right).$$

Remark 1.6.7. The exponential process $\mathcal{E}^{(\sigma,m)}$ is an Itô process satisfying the stochastic differential equation

$$\begin{cases} \mathrm{d}\mathcal{E}_t^{(\sigma,m)} = \sigma_t \mathcal{E}_t^{(\sigma,m)}\,\mathrm{d}W_t + m_t \mathcal{E}_t^{(\sigma,m)}\,\mathrm{d}t, \\ \ \mathcal{E}_a^{(\sigma,m)} = 1. \end{cases}$$

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

$$\begin{cases} \mathrm{d}\mathcal{E}_t^{(\sigma)} = \sigma_t \mathcal{E}_t^{(\sigma)}\,\mathrm{d}W_t, \\ \ \mathcal{E}_a^{(\sigma)} = 1. \end{cases}$$

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

The following theorem states the existence and uniqueness of the solution for stochastic differential equations.

Theorem 1.6.8 ([Kuo06, section 10.3])**.** *Let $m(t,x)$ and $\sigma(t,x)$ be real-valued measurable functions on $[a,b] \times \mathbb{R}$ such that there is a $M > 0$ for which the following conditions are satisfied:*

1. *(linear growth in x)* $|m(t,x)| + |\sigma(t,x)| \leq M(1 + |x|)$,
2. *(Lipschitz condition in x)* $|m(t,x) - m(t,y)| + |\sigma(t,x) - \sigma(t,y)| \leq M\,|x - y|$.

Suppose also that the initial condition is a square-integrable $\mathcal{F}_a$-measurable random variable. Then the stochastic differential equation (1.6.2) *has a unique continuous solution.*

Girsanov's theorem Girsanov's theorem tells us that if a Wiener process is translated in certain directions (in the functional space), the resultant process is also a Wiener process with respect to an equivalent probability measure.

Theorem 1.6.9 ([Gir60])**.** *Let $X \in L^2_{ad}[0,T]$ for some $T \in [0,\infty)$ such that $\mathbb{E}_{\mathbb{P}}\left(\mathcal{E}_t^{(X)}\right) = 1$ for all $t \in [0,T]$. Then the stochastic process $\widetilde{W}$ given by*

$$\widetilde{W}_t = W_t - \int_0^T X_s\,\mathrm{d}s, \quad t \in [0,T],$$

is a Wiener process with respect to the probability measure $\mathbb{Q}$ defined by the Radon–Nikodym derivative $\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} = \mathcal{E}_T^{(X)}$.

Chapter 2. Towards a General Theory of Stochastic Integration

2.1. Motivation

Itô's calculus, even though being very useful, cannot handle anticipating integrands. We motivate this issue with a few examples.

Example 2.1.1. For $t \in [0,1]$, consider

$$\begin{cases} \mathrm{d}Z(t) = Z(t)\,\mathrm{d}W_t \\ Z(0) = W_1. \end{cases}$$

This is a stochastic differential equation with *anticipating* initial condition. If we try to apply the Picard iteration method to obtain a solution of this stochastic integral equation, we quickly run into the issue of having to assign a meaning to the expression $\int_0^t W_1\,\mathrm{d}W_s$, which has no meaning within Itô's theory.

Example 2.1.2. Instead of the initial condition being anticipating, we can also consider the diffusion coefficient to be anticipating.

$$\begin{cases} \mathrm{d}Z(t) = W_1 Z(t)\,\mathrm{d}W_t \\ Z(0) = 1. \end{cases}$$

We run into the exact same issue in this case.

We face the same problem when trying to define iterated Itô integrals. For example, note that

$$\int_0^1 \int_0^1 1\,\mathrm{d}W_v\,\mathrm{d}W_u = \int_0^1 W_1\,\mathrm{d}W_u,$$

which is outside the realm of Itô's theory. What is the way forward?

2.2. Multiple Wiener–Itô integrals

A naive approach would be to treat W_1 as a constant. Then we can pull it out of the integral and we would obtain

$$X_t \triangleq \int_0^t \int_0^t 1\,\mathrm{d}W_t\,\mathrm{d}W_u = W_t \int_0^t \mathrm{d}W_u = W_t \cdot W_t = W_t^2.$$

In fact, Itô himself posed a similar question in 1976 at the International Symposium on Stochastic Differential Equations in Kyoto. However, there is a drawback of this approach.

Recall that we emphasized a key aspect of Itô's integrals — their martingale nature. So we naturally ask, is the process associated with such an iterated integral a martingale? We know that the process W_t^2 is not a martingale. On the other hand, $W_t^2 - t$ is a martingale. Can we modify the definition of the integral so that we get this martingale as the result? For deterministic integrands, the answer is yes. We detail this below.

Under $L^2(\Omega)$-limits, we expect that

$$X_t = \lim_{n\to\infty} \sum_{i=1}^{n} \sum_{j=1}^{n} 1 \, \Delta W_j \, \Delta W_i.$$

Let $S_n = \sum_{i=1}^{n} \sum_{j=1}^{n} 1 \, \Delta W_j \, \Delta W_i$. Then breaking up S_n into off-diagonal and diagonal parts, we get

$$S_n = \sum_{i=1}^{n} \sum_{j=1}^{i-1} 1 \, \Delta W_j \, \Delta W_i + \sum_{i=1}^{n} \sum_{j=i+1}^{n} 1 \, \Delta W_j \, \Delta W_i + \sum_{i=1}^{n} 1 \, (\Delta W_i)^2 .$$

Since 1 is symmetric about u and v (and therefore with respect to i and j), we simplify this to

$$S_n = 2 \sum_{i=1}^{n} \sum_{j=1}^{i-1} \Delta W_j \, \Delta W_i + \sum_{i=1}^{n} (\Delta W_i)^2 = 2 \sum_{i=1}^{n} W_{t_{i-1}} \, \Delta W_i + \sum_{i=1}^{n} (\Delta W_i)^2 .$$

Taking L^2-limits, and recalling that $(\mathrm{d}W_t)^2 \to \mathrm{d}t$, we get

$$X_t = 2 \int_0^t W_u \, \mathrm{d}W_u + \int_0^t \mathrm{d}u = \left(W_t^2 - t\right) + t,$$

where the first integral is within Itô's theory (see example 1.5.3).

There are some key takeaways from the above computation.

1. We explicitly used the symmetry of the integrand to convert the inside integral into an adapted process.
2. The symmetrization introduced a multiplicative factor of 2!.
3. The off-diagonal terms gave rise to an Itô integral, which is a martingale and is exactly what we want the result of our iterated integral to be.
4. The diagonal terms gave rise to a function that kills the martingale nature of the Itô integral. This is due to the quadratic variation of Wiener processes.

This motivated Itô to define multiple integrals by first symmetrizing the function and then removing the diagonal terms[Itô51]. We define *multiple Wiener–Itô integrals* in two steps.

1. An *off-diagonal step function* is a function of the form

   $$f = \sum_{1 \le i_1, \ldots, i_n \le k} a_{i_1,\ldots,i_n} \mathbb{1}_{[t_{i_1-1}, t_{i_1}) \times \cdots \times [t_{i_n-1}, t_{i_n})},$$

   where $a = t_0 < \cdots < t_k = b$ and

   $$a_{i_1,\ldots,i_n} = 0 \text{ if } i_p = i_q \text{ for some } p \neq q.$$

   For an off-diagonal step function f as above, define

   $$I_n(f) = \sum_{1 \le i_1, \ldots, i_n \le k} a_{i_1,\ldots,i_n} \, \Delta W_{i_1} \cdots \Delta W_{i_n}.$$

2. For any $f \in L^2([a,b]^n)$, there exists a sequence of off-diagonal step functions $(f_k)_{k=1}^{\infty}$ such that $f_k \to f$ in $L^2([a,b]^n)$ (see [Kuo06, lemma 9.6.4]). Then we define the multiple Wiener–Itô integral of f as the $L^2(\Omega)$-limit

$$I_n(f) = \lim_{k\to\infty} I_n(f_k).$$

For $f \in L^2([a,b]^n)$, define the *symmetrization* $\hat{f}$ of f by

$$\widehat{f}(t_1, \dots, t_n) = \frac{1}{n!}\sum_{\sigma} f\left(t_{\sigma(1)}, \dots, t_{\sigma(n)}\right),$$

where the summation is over all permutations σ of the set $[n] = \{1, 2, \dots, n\}$. Let $L^2_{sym}([a,b]^n)$ be the Hilbert space of all square integrable functions. Then we have the following results.

Theorem 2.2.1 ([Kuo06, theorem 9.6.6]). *If $f \in L^2([a,b]^n)$, then*

1. $I_n(f) = I_n\left(\widehat{f}\right)$,
2. $\mathbb{E}(I_n(f)) = 0$, *and*
3. $\mathbb{E}\big(I_n(f)^2\big) = n!\left\|\widehat{f}\right\|^2$.

Therefore, we can expect to recover the integral from the symmetrization. This is the content of the following theorem.

Theorem 2.2.2 ([Kuo06, theorem 9.6.7]). *We can define the multiple Wiener–Itô integral of $f \in L^2([a,b]^n)$ for $n \geq 2$ as*

$$I_n(f) = n!\int_a^b \cdots \int_a^{t_{n-2}} \left(\int_a^{t_{n-1}} \widehat{f}(t_1, \dots, t_n)\,\mathrm{d}W_{t_n}\right)\mathrm{d}W_{t_{n-1}} \cdots \mathrm{d}W_{t_1}.$$

2.3. Homogeneous chaos

Recall that scaled Hermite polynomials form an orthonormal basis for the space $L^2(\mathbb{R}, \gamma)$, where γ is the Gaussian measure with mean 0 and variance t. Therefore, for any function $f \in L^2(\mathbb{R}, \gamma)$, we have

$$f(x) = \sum_{n=0}^{\infty} a_n \frac{H_n(x;t)}{\sqrt{n!t}},$$

where the coefficients a_n are given by

$$a_n = \frac{1}{\sqrt{n!t^n}}\int_{-\infty}^{\infty} f(x)H_n(x;t)\gamma(\mathrm{d}x).$$

Let $\mathcal{F}^W = \sigma(W_t \mid t \in [a,b])$. Denote the Hilbert space of $\mathbb{P}$-square integrable functionals that are $\mathcal{F}^W$-measurable — called square-integrable *Wiener functionals* — by $L^2_W(\Omega)$. Can we obtain an orthogonal decomposition for the space $L^2_W(\Omega)$ similar to that of $L^2(\mathbb{R}, \gamma)$? Wiener

came across this question when he wanted to develop the theory of Fourier analysis on Wiener functionals. In his work[Wie38], Wiener showed that there exists an orthogonal decomposition of the space $L^2_W(\Omega)$. These orthogonal subspaces are called *homogeneous chaos*es and denoted by K_n with $K_0 = \mathbb{R}$. That is,

$$L^2_W(\Omega) = \bigoplus_{n=0}^{\infty} K_n.$$

Therefore, any $\phi \in L^2_W(\Omega)$ can be uniquely expressed as

$$\phi = \sum_{n=0}^{\infty} \phi_n,$$

where $\phi_n \in K_n$, and

$$\|\phi\|^2 = \sum_{n=0}^{\infty} \|\phi_n\|^2 .$$

Homogeneous chaoses or different orders are orthogonal. In particular, any homogeneous chaos of order $n \geq 1$ has expectation zero. Note that K_1 is exactly the set of Wiener integrals (which are also martingales). Homogeneous chaoses can be thought of as extensions of the idea of martingales generated from Wiener integrals.

Wiener did not give a probabilistic interpretation of the homogeneous chaoses. Itô questioned whether such an interpretation is possible. This motivated him to define the multiple Wiener–Itô integrals $I_n(f)$ that we detailed in section 2.2. In fact, he showed the following relation between Hermite polynomials of Wiener integrals and multiple Wiener–Itô integrals.

Theorem 2.3.1 ([Kuo06, theorem 9.6.9])**.** *For any $f \in L^2[a, b]$, define the tensor product $f^{\otimes n}(t_1, \dots, t_n) = f(t_1) \cdots f(t_n)$. Then*

$$I_n\left(f^{\otimes n}\right) = H_n\left(I(f), \|f\|^2\right).$$

The last theorem, along with the orthogonality of $(I_n(f))_{n=0}^{\infty}$, gives us the result that K_n is exactly the set of multiple Wiener–Itô integrals of order n.

Theorem 2.3.2 ([Kuo06, theorem 9.7.1])**.** *If $f \in L^2([a, b]^n)$, then $I_n(f) \in K_n$. Conversely, any $\phi \in K_n$ has a unique representation $\phi = I_n(f)$, where $f \in L^2_{sym}([a, b]^n)$.*

This implies that $K_n = \left\{I_n(f) \mid f \in L^2_{sym}([a, b]^n)\right\}$.

All of the above culminates in the decomposition theorem for $L^2_W(\Omega)$.

Theorem 2.3.3 (Wiener–Itô [Kuo06, theorem 9.7.3])**.** *The space $L^2_W(\Omega)$ can be decomposed into the orthogonal direct sum*

$$L^2_W(\Omega) = \bigoplus_{n=0}^{\infty} K_n,$$

where K_n is the set of multiple Wiener–Itô integrals of order n. Therefore, any function in $\phi \in L^2_W(\Omega)$ can be represented by

$$\phi = \sum_{n=0}^{\infty} I_n(f_n), \quad f_n \in L^2_{sym}\left([a,b]^n\right),$$

along with the equality

$$\|\phi\|^2 = \sum_{n=0}^{\infty} n!\, \|f_n\|^2 .$$

The Wiener–Itô decomposition theorem provides a deep connection between infinite-dimensional analysis and stochastic analysis. However, we still have not been able to define multiple integrals for stochastic processes. There is also the question of how to define the derivative of a random variable with respect to ω. This gave rise to some theories, which we look at briefly in the following sections.

2.4. Itô's idea of enlargement of filtration

At the 1976 Kyoto International Symposium, Itô[Itô78] proposed enlarging the filtration to include W_1, that is,

$$\tilde{\mathcal{F}}_t = \sigma\left(W_s, W_1 \mid s \in [0,t]\right)$$

However, W is not a Wiener process with respect to $\tilde{\mathcal{F}}$. Therefore, he decomposed the Wiener process as $W = M + A$, where

$$\begin{aligned} M_t &= W_1 - \int_0^t \frac{W_1 - W_u}{1-u}\, \mathrm{d}u, \text{ and} \\ A_t &= \int_0^t \frac{W_1 - W_u}{1-u}\, \mathrm{d}u. \end{aligned}$$

Here M is a quasimartingale. Using this method, W_1 becomes $\tilde{\mathcal{F}}$-measurable, and for any $t \in [0,1]$, we have

$$\int_0^t W_1\, \mathrm{d}W_s = W_1 W_t.$$

$$\begin{aligned} \int_0^t W_1\, \mathrm{d}W_s &= \int_0^t W_1\, \mathrm{d}M_u + \int_0^t W_1\, \mathrm{d}A_u \\ &= W_1(M_t - M_0) + W_1(A_t - A_0) \\ &= W_1(W_t - W_0) \\ &= W_1 W_t. \end{aligned}$$

However, note that by the quadratic variation of Wiener process and the independence of increments, we have

$$\mathbb{E}(W_1 W_t) = \mathbb{E}\big((W_1 - W_t)W_t + {W_t}^2\big) = 0 + t \neq 0.$$

Therefore, the integral is not a martingale, which is undesirable.

2.5. Malliavin calculus and Hitsuda–Skorokhod integral

The original motivation of *Malliavin calculus* was to study Wiener functionals. In particular, its goals were to study the solutions of stochastic differential equations driven by Brownian noise. The presentation follows [Nua06, sections 1.2–1.3].

Let $\mathcal{C}_0$ denote the space of continuous functions $f : [0,1] \to \mathbb{R}$ such that $f(0) = 0$. Recall that $\left(\mathcal{C}_0, \|\cdot\|_\infty\right)$ is a Banach space, where $\|\cdot\|_\infty$ is the supremum norm. Define

$$\mathcal{H}^1 = \left\{ f \in \mathcal{C}_0 \mid f' \text{ exists and } f' \in L^2[0,1] \right\}.$$

Clearly, $\mathcal{H}^1 \subset \mathcal{C}_0$. Embedded with the inner product given by

$$\langle f, g \rangle_{\mathcal{H}^1} = \int_0^1 f'(t)\, g'(t)\, \mathrm{d}t,$$

$\mathcal{H}^1$ is a Hilbert space, and is known as the *Cameron–Martin space*. $\mathcal{H}^1$ is densely embedded in $\mathcal{C}_0$ with the canonical injection $i : \mathcal{H}^1 \to \mathcal{C}_0$.

We consider $\Omega = \mathcal{C}_0$ and Σ to be the Borel σ-algebra generated by the topology induced by the norm. This σ-algebra coincides with the σ-algebra generated by the cylinder sets. The measurable space (Ω, Σ) is induced with the *Wiener measure* defined on the cylinder sets by

$$\gamma\left\{W_{t_1} \in E_1, \dots, W_{t_n} \in E_n\right\} = \int_{E_1 \times \cdots \times E_n} \prod_{i=1}^n \left(\frac{1}{\sqrt{2\pi\, \Delta t_i}} \exp\left[-\frac{(\Delta u_i)^2}{2\, \Delta t_i} \right] \right) \mathrm{d}u_1 \cdots \mathrm{d}u_n,$$

where $\Delta u_i = u_i - u_{i-1}$, $\Delta t_i = t_i - t_{i-1}$, $t_0 = 0$ and $u_0 = 0$. The measure space $(\mathcal{C}_0, \gamma)$ is called the *Wiener space* because we can regard each path $t \mapsto W_t(\omega)$ of the Wiener process W as an element $\omega \in \mathcal{C}_0$. That is, we can identify $W_t(\omega)$ with the value $\omega(t)$ at time t and $\omega \in \mathcal{C}_0$ as $W_t(\omega) = \omega(t)$.

2.5.1. The derivative operator

For $h \in L^2[0,1]$, let $I_W(h) = \int_0^1 h(t)\, \mathrm{d}t$ denote the Wiener integral. Therefore, $W_t = I_W(\mathbb{1}_{[0,t]})$ for any t. Let $C_p^\infty(\mathbb{R}^n)$ be the set of all infinitely continuously differentiable functions $f : \mathbb{R}^n \to \mathbb{R}$ such that f and all of its partial derivatives have polynomial growth. Let $\mathcal{S}$ denote the class of *smooth random variables* such that a random variable $F \in \mathcal{S}$ has the form

$$F = f\left(I_W(h_1), \dots, I_W(h_n)\right),$$

where $f \in C_p^\infty(\mathbb{R}^n)$, and $h_1, \dots, h_n \in L^2[0,1]$ for any natural number n.

The *stochastic derivative* of a smooth random variable $F \in \mathcal{S}$ is given by

$$\mathrm{D}F = \mathrm{D}_\cdot F = \sum_{i=1}^n \frac{\partial f}{\partial x_i} \left(I_W(h_1), \dots, I_W(h_n)\right) h_i.$$

For example, $\mathrm{D}I_W(h) = h$, $\mathrm{D}(I_W(h))^2 = 2I_W(h)\,h$, and $\mathrm{D}W_{t_1} = \mathrm{D}I_W(\mathbb{1}_{[0,t_1]}) = \mathbb{1}_{[0,t_1]}$. Note that $\mathrm{D}F \in L^2([0,1]\times\Omega) \simeq L^2\left(\Omega; L^2[0,1]\right)$. So D is an unbounded linear operator from $\mathcal{S} \subset L^2(\Omega)$ into $L^2\left(\Omega; L^2[0,1]\right)$.

We can interpret the stochastic derivative as the directional derivative in a direction $\int_0^{\cdot} h(t)\,\mathrm{d}t$ of the Cameron–Martin space, since for any $h \in L^2[0,1]$, we get

$$\begin{aligned}\langle \mathrm{D}F, h\rangle &= \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}\left(I_W(\mathbb{1}_{[0,t_1]}), \dots, I_W(\mathbb{1}_{[0,t_n]})\right)\left\langle \mathbb{1}_{[0,t_i]}, h\right\rangle \\ &= \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}\left(I_W(\mathbb{1}_{[0,t_1]}), \dots, I_W(\mathbb{1}_{[0,t_n]})\right)\int_0^{t_i} h(s)\,\mathrm{d}s \\ &= \frac{\mathrm{d}}{\mathrm{d}\epsilon}\left[F\left(\omega + \epsilon\int_0^{\cdot} h(s)\,\mathrm{d}s\right)\right]\Bigg|_{\epsilon=0},\end{aligned}$$

where all inner products are in $L^2[0,1]$.

For example, if $F = W_{t_1}$, then

$$F\left(\omega + \epsilon\int_0^{\cdot} h(s)\,\mathrm{d}s\right) = \omega(t_1) + \epsilon\int_0^{t_1} h(s)\,\mathrm{d}s,$$

giving $\langle \mathrm{D}F, h\rangle = \int_0^{\cdot} h(s)\,\mathrm{d}s = \int_0^1 \mathbb{1}_{[0,t_1]}h(s)\,\mathrm{d}s$, so $\mathrm{D}F = \mathbb{1}_{[0,t_1]}$.

Now we state the integration-by-parts formula.

Lemma 2.5.1 ([Nua06, lemma 1.2.1])**.** *Suppose F is a smooth functional and $h \in L^2[0,1]$. Then*

$$\mathbb{E}(\langle \mathrm{D}F, h\rangle) = \mathbb{E}(F I_W(h))\,.$$

Applying the previous result to a product FG, we get the following consequence.

Lemma 2.5.2 ([Nua06, lemma 1.2.2])**.** *Suppose F and G are smooth functionals and $h \in L^2[0,1]$. Then*

$$\mathbb{E}(G\langle \mathrm{D}F, h\rangle) = \mathbb{E}(-F\langle \mathrm{D}G, h\rangle) + \mathbb{E}(FGI_W(h))\,.$$

As a consequence of the last lemma, D is closable as an operator from $L^p(\Omega)$ to $L^p\left(\Omega; L^2[0,1]\right)$ for any $p \geq 1$. Denote the domain of D in $L^p(\Omega)$ by $\mathbb{D}^{1,p}$, which represents the closure of $\mathcal{S}$ with respect to the norm

$$\|F\|_{1,p} = \left[\mathbb{E}\left(|F|^p\right) + \mathbb{E}\left(\|\mathrm{D}F\|_2^2\right)\right]^{\frac{1}{p}}.$$

For $p = 2$, the space $\mathbb{D}^{1,2}$ is a Hilbert space with the inner product

$$\langle F, G\rangle_{1,2} = \mathbb{E}(FG) + \mathbb{E}\left(\langle \mathrm{D}F, \mathrm{D}G\rangle_2\right).$$

Note that $\mathbb{D}^{1,2}$ is dense in $L^2(\Omega)$.

This abstract formulation, though general, is difficult to use in computations. However, if we have a homogeneous chaos decomposition of F, we can compute the stochastic derivative explicitly. This is the content of the following theorem.

Theorem 2.5.3 ([Nua06, proposition 1.2.1]). *Suppose F is an square integrable random variable admitting a homogeneous chaos decomposition of the form*

$$F = \sum_{n=0}^{\infty} I_n(f_n), \quad f_n \in L^2_{sym}([a,b]^n).$$

Then $F \in \mathbb{D}^{1,2}$ if and only if

$$\sum_{n=1}^{\infty} n\, n! \, \|f_n\|^2 < \infty.$$

In that case, we have

$$\mathrm{D}_t F = \sum_{n=1}^{\infty} n I_{n-1}(f_n(\cdot, t)),$$

and

$$\|\mathrm{D}F\|_2^2 = \sum_{n=1}^{\infty} n\, n! \, \|f_n\|^2 < \infty.$$

2.5.2. The Hitsuda–Skorokhod integral

The *Hitsuda–Skorokhod stochastic integral* or *divergence operator* δ is defined as the $L^2(\Omega)$-adjoint of the stochastic derivative operator D.

Definition 2.5.4 ([Nua06, definition 1.3.1]). Let δ denote the adjoint of D. That means δ is an unbounded operator on $L^2([0,1] \times \Omega)$ with values in $L^2(\Omega)$ such that the following hold:

1. The domain of δ, denoted by $\mathrm{dom}(\delta)$, is the set of processes $u \in L^2([0,1] \times \Omega)$ such that
$$\mathbb{E}(\langle \mathrm{D}F, u \rangle) \le C_u \|F\|_2$$
for all $F \in \mathbb{D}^{1,2}$, where C_u is some constant depending on u.
2. If $u \in \mathrm{dom}(\delta)$, then $\delta(u)$ is the element of $L^2(\Omega)$ characterized by
$$\mathbb{E}(F\delta(u)) = \mathbb{E}(\langle \mathrm{D}F, u \rangle)$$
for any $F \in \mathbb{D}^{1,2}$.

We look at the action of the Hitsuda–Skorokhod integral on homogeneous chaos expansions. First we state a lemma.

Lemma 2.5.5 ([Nua06, lemma 1.3.1]). *Let $u \in L^2([0,1] \times \Omega)$. There exists a family of deterministic, measurable, and square integrable kernels $f_n(t_1, \dots, t_n, t)$ with $n \in \mathbb{N}_0$, such that every f_n is symmetric in the first n variables, and*

$$u(t) = \sum_{n=0}^{\infty} I_n(f_n(\cdot, t)),$$

where the convergence is in $L^2([0,1] \times \Omega)$. Moreover,

$$\|u\|^2 = \sum_{n=0}^{\infty} n! \, \|f_n\|^2_{L^2[0,1]^{n+1}}.$$

In this case, $u \in \mathrm{dom}(\delta)$ if an only if the series

$$\delta(u) = \sum_{n=0}^{\infty} n! I_{n+1}(f_n)$$

converges in $L^2(\Omega)$.

Recall that the stochastic derivative was first defined on the class of smooth random variables because it gave us simple expressions. Similarly, the Hitsuda–Skorokhod integral has a simple expression for *smooth elementary processes*. A smooth elementary processes is a process of the form

$$u(t) = \sum_{i=1}^{n} F_i \, h_i(t),$$

where $F_i \in \mathcal{S}$ and $h_i \in L^2[0,1]$ for all $i \in [n]$. By the integration-by-parts formula in lemma 2.5.2 above, we see that u is Skorokhod integrable, and

$$\delta(u) = \sum_{i=1}^{n} F_i \, I_W(h_i(t)) - \sum_{i=1}^{n} (\mathrm{D}_t F_i) \, h_i(t) \, \mathrm{d}t.$$

The reason the adjoint of the derivative is called an "integral" is because it reduces to Itô's integral for $u \in L^2_{\mathrm{ad}}([0,1] \times \Omega)$.

Theorem 2.5.6 ([Nua06, proposition 1.3.4])**.** *The space $L^2_{ad}([0,1] \times \Omega)$ is contained in* $\mathrm{dom}(\delta)$*, and for any $u \in L^2_{ad}([0,1] \times \Omega)$, we have*

$$\delta(u) = \int_0^1 u_t \, \mathrm{d}W_t,$$

where the right hand integral is in the sense of Itô.

We shall see applications of these ideas in section 7.3.

Chapter 3. The Ayed–Kuo Integral

3.1. The definition of the integral

We first define processes that can be thought of as purely anticipating in nature.

Definition 3.1.1. A stochastic process Y is called *instantly-independent* with respect to the filtration $\mathcal{F}$ if the random variable Y^t and the σ-algebra $\mathcal{F}_t$ are independent for each t.

Remark 3.1.2. To denote time for adapted processes, we use subscripts, for example X_t. For instantly-independent processes, we use superscripts, like Y^t. For deterministic functions or general processes, we use parenthesis, for example $Z(t)$.

The *Ayed–Kuo integral* of a stochastic process $Z(t)$ introduced in [AK08] is defined in the following two steps. In the entire process, we assume $(\Pi_n)_{n=1}^{\infty}$ is a sequence of partitions of $[a, b]$ such that $\|\Pi_n\| \to 0$ as $n \to \infty$. We also assume that the reader is familiar with modes of convergence of random variables.

Step 1 Suppose X is an $\mathcal{F}$-adapted continuous stochastic process and Y be an continuous stochastic processes that is instantly-independent with respect to $\mathcal{F}$. Then the stochastic integral of the product XY is defined by the limit in probability

$$\int_a^b X_t\, Y^t \,\mathrm{d}W_t = \lim_{n\to\infty} \sum_{j=1}^{n} X_{t_{j-1}} Y^{t_j}\, \Delta W_{t_j},$$

provided that the limit exists in probability.

For a process of the form $Z(t) = \sum_{i=1}^{m} X_t^{(i)} Y_{(i)}^t$, the stochastic integral is defined by

$$\int_a^b Z(t)\,\mathrm{d}W_t = \sum_{i=1}^{m} \int_a^b X_t^{(i)} Y_{(i)}^t \,\mathrm{d}W_t. \tag{3.1.3}$$

Example 3.1.4 ([AK08, equation 1.6]). We first decompose the integrand into adapted and instantly-independent parts as $W_b = W_t + (W_b - W_t)$. Using the definition of the integral under L^2-limits, we get

Parts of section 3.4 previously appeared in the following open-access journal articles:

1. Hui-Hsiung Kuo, Sudip Sinha, and Jiayu Zhai. “Stochastic Differential Equations with Anticipating Initial Conditions”. In: *Communications on Stochastic Analysis* 12.4 (2018). DOI: 10.31390/cosa.12.4.06
2. Hui-Hsiung Kuo, Pujan Shrestha, and Sudip Sinha. “Anticipating Linear Stochastic Differential Equations with Adapted Coefficients”. In: *Journal of Stochastic Analysis* 2.2 (2021). DOI: 10.31390/josa.2.2.05

$$\begin{aligned}
\int_a^b W_b \, \mathrm{d}W_t &= \int_a^b [W_t + (W_b - W_t)] \, \mathrm{d}W_t \\
&= \lim_{n\to\infty} \sum_{i=1}^{n} \left[W_{t_{i-1}} + (W_b - W_{t_i})\right] \Delta W_i \\
&= \lim_{n\to\infty} \sum_{i=1}^{n} (W_b - \Delta W_i) \, \Delta W_i \\
&= W_b \cdot \lim_{n\to\infty} \sum_{i=1}^{n} \Delta W_i - \lim_{n\to\infty} \sum_{i=1}^{n} (\Delta W_i)^2 \\
&= W_b(W_b - W_a) - (b - a).
\end{aligned}$$

Example 3.1.5 ([AK08, equation 1.6])**.** Following the exact same procedure as in example 3.1.4, we get

$$\int_a^t W_b \, \mathrm{d}W_s = W_b(W_t - W_a) - (t - a).$$

We give another way of computing this integral later in example 5.1.1.

Step 2 Let Z be a stochastic process such that there is a sequence $(Z_n(t))_{n=1}^{\infty}$ of stochastic processes each of the form in equation (3.1.3) satisfying

1. $\int_a^b |Z_n(t) - Z(t)|^2 \, \mathrm{d}t \to 0$ as $n \to \infty$ almost surely, and
2. $\int_a^b Z_n(t) \, \mathrm{d}W_t$ converges in probability as $n \to \infty$.

Then the stochastic integral of Z is defined by

$$\int_a^b Z(t) \, \mathrm{d}W_t = \lim_{n\to\infty} \int_a^b Z_n(t) \, \mathrm{d}W_t \quad \text{in } \mathbb{P}.$$

In this document, we assume the stronger condition of convergence in $L^2(\Omega)$ instead of convergence in probability, since the former implies the latter and is easier to work with.

Let us look at a few more examples of Ayed–Kuo integrals.

Example 3.1.6 ([HKS19, example 2.1])**.** Using L^2-limits,

$$\begin{aligned}
\int_a^t W_s(W_b - W_s)\,\mathrm{d}W_s &= \lim_{n\to\infty}\sum_{i=1}^n W_{t_{i-1}}(W_b - W_{t_i})\,\Delta W_i \\
&= \lim_{n\to\infty}\sum_{i=1}^n W_{t_{i-1}}(W_b - \Delta W_i - W_{t_{i-1}})\,\Delta W_i \\
&= W_b \lim_{n\to\infty}\sum_{i=1}^n W_{t_{i-1}}\,\Delta W_i - \lim_{n\to\infty}\sum_{i=1}^n W_{t_{i-1}}(\Delta W_i)^2 - \lim_{n\to\infty}\sum_{i=1}^n W_{t_{i-1}}^2\,\Delta W_i \\
&= W_b \int_a^t W_s\,\mathrm{d}W_s - \lim_{n\to\infty}\sum_{i=1}^n W_{t_{i-1}}(\Delta t_i) - \int_a^t W_s{}^2\,\mathrm{d}W_s \\
&= W_b \int_a^t W_s\,\mathrm{d}W_s - \int_a^t W_s\,\mathrm{d}s - \int_a^t W_s{}^2\,\mathrm{d}W_s.
\end{aligned}$$

From Itô's theory, we know

$$\begin{aligned}
\int_a^t W_s\,\mathrm{d}W_s &= \frac{1}{2}\left((W_t{}^2 - W_a{}^2) - (t-a)\right), \text{ and} \\
\int_a^t W_s{}^2\,\mathrm{d}W_s &= \frac{1}{3}\left((W_t{}^3 - W_a{}^3) - \int_a^t W_s\,\mathrm{d}s\right).
\end{aligned}$$

Putting these together, we get

$$\int_a^t W_s(W_b - W_s)\,\mathrm{d}W_s = \frac{1}{2}W_b\left((W_t{}^2 - W_a{}^2) - (t-a)\right) - \frac{1}{3}\left(W_t{}^3 - W_a{}^3\right).$$

Proposition 3.1.7. *For $p \in \mathbb{N}$, we have $\int_a^t W_b^p\,\mathrm{d}W_t = W_b^p\,(W_t - W_a) - pW_b^{p-1}(t-a)$.*

Proof. Taking L^2 limits,

$$\begin{aligned}
\int_a^t W_T^p\,\mathrm{d}W_s &= \int_a^t (W_s + (W_T - W_u))^p\,\mathrm{d}W_s \\
&= \lim_{n\to 0}\sum_{i=1}^n \left(W_{t_{i-1}} + \left(W_T - W_{t_i}\right)\right)^p(\Delta W_i) \\
&= \lim_{n\to 0}\sum_{i=1}^n (W_T - \Delta W_i)^p\,(\Delta W_i) \\
&= \lim_{n\to 0}\sum_{i=1}^n \sum_{k=0}^p \binom{p}{k}(-1)^k W_T^{p-k}\,(\Delta W_i)^{k+1} \\
&= \sum_{k=0}^p \left[\binom{p}{k}(-1)^k W_T^{p-k}\cdot \lim_{n\to 0}\sum_{i=1}^n (\Delta W_i)^{k+1}\right].
\end{aligned}$$

Since $(\Delta W_i)^{k+1} \to 0$ in L^2 as $n \to \infty$ for all $k \geq 2$, we have only have the $k = 0$ and $k = 1$ terms remaining. Moreover, $(\Delta W_i)^2 \to \Delta t_i$ in L^2 as $n \to \infty$. Therefore,

$$\begin{aligned}\int_0^t W_T^p \, \mathrm{d}W_s &= W_T^p \cdot \lim_{n\to 0} \sum_{i=1}^n \Delta W_i - pW_T^{p-1} \cdot \lim_{n\to 0} \sum_{i=1}^n \Delta t_i \\ &= W_T^p (W_t - W_a) - pW_T^{p-1}(t-a).\end{aligned}$$

□

3.2. Differential formula

Recall that Itô's formula (theorem 1.5.2) allowed us to create new Itô processes from old ones. Here we look at an extension of Itô's formula that can also account for instantly-independent processes. Let X and Y be stochastic processes of the form

$$X_t = X_a + \int_a^t m_s \, \mathrm{d}s + \int_a^t \sigma_s \, \mathrm{d}W_s, \qquad \text{and} \tag{3.2.1}$$

$$Y^t = Y^b + \int_t^b \eta^s \, \mathrm{d}s + \int_t^b \xi^s \, \mathrm{d}W_s, \tag{3.2.2}$$

where σ and m are adapted (so X is an Itô process), and ξ and η are instantly-independent such that Y is also instantly-independent.

Theorem 3.2.3 ([Hwa+16, theorem 3.2]). *Suppose $\left(X^{(i)}_\cdot\right)_{i=1}^n$ and $\left(Y^\cdot_{(j)}\right)_{j=1}^m$ are stochastic processes of the form given by equations equation (3.2.1) and equation (3.2.2), respectively. Suppose $\theta(t, x_1, \dots, x_n, y_1, \dots, y_m)$ is a real-valued function that is C^1 in t and C^2 in other variables. Then the stochastic differential of $\theta(t, X_t^{(1)}, \dots, X_t^{(n)}, Y_1^{(t)}, \dots, Y^t_{(m)})$ is given by*

$$\begin{aligned}&\mathrm{d}\theta\left(t, X_t^{(1)}, \dots, X_t^{(n)}, Y^t_{(1)}, \dots, Y^t_{(m)}\right) \\ &= \theta_t \, \mathrm{d}t + \sum_{i=1}^n \theta_{x_i} \, \mathrm{d}X_t^{(i)} + \sum_{j=1}^m \theta_{y_j} \, \mathrm{d}Y^t_{(j)} \\ &\quad + \frac{1}{2} \sum_{i,k=1}^n \theta_{x_i x_k} \, \mathrm{d}X_t^{(i)} \, \mathrm{d}X_t^{(k)} - \frac{1}{2} \sum_{j,l=1}^m \theta_{y_j y_l} \, \mathrm{d}Y^t_{(j)} \, \mathrm{d}Y^t_{(l)}.\end{aligned}$$

Corollary 3.2.4 ([HKS19, corollary 2.11]). *Suppose X_t is an Itô process and $\psi(t, x, z)$ is a real-valued function that is C^1 in in t and C^2 in x and z. Then the stochastic differential of $\psi(t, X_t, W_b$ is given by*

$$\psi(t, X_t, W_b) = \psi_t \, \mathrm{d}t + \psi_x \, \mathrm{d}X_t + \frac{1}{2}\psi_{xx}(\mathrm{d}X_t)^2 + \psi_{xz} \, \mathrm{d}X_t \, \mathrm{d}W_t.$$

Proof. Since $W_b = W_t + (W_b - W_t)$ for any t, define a function $\theta(t, x_1, x_2, y) = \psi(t, x_1, x_2 + y)$ and set $X_t^{(1)} = X_t$, $X_t^{(2)} = W_t$, and $Y^t = W_b - W_t$. Applying theorem 3.2.3 gives us the required result. □

Example 3.2.5. It is much simpler to calculate the result of proposition 3.1.7 using corollary 3.2.4. We guess the solution of the integral to be of the form $W_b^p W_t$. Therefore, let $\psi(x,z) = z^p x$. So $\psi_x(x,z) = z^p$ and $\psi_{xz}(x,z) = pz^{p-1}$, all other terms being zero. Putting this together and using $(\mathrm{d}W_t)^2 = \mathrm{d}t$ we get

$$\mathrm{d}\big(W_b^p W_t\big) = \mathrm{d}\psi\left(W_t, W_b^p\right) = W_b^p\,\mathrm{d}W_t + pW_b^{p-1}\,\mathrm{d}t.$$

Therefore,

$$\int_a^t W_b^p\,\mathrm{d}W_s = W_b^p\,(W_t - W_a) - pW_b^{p-1}(t-a).$$

Just as in Itô's theory, the differential formula plays a very important role in the study of anticipating stochastic differential equations in the Ayed–Kuo theory. However, our experience of stochastic differential equations from Itô's theory does not translate over completely; the solutions of anticipating stochastic differential equations has quite a few surprises in store. In the following two sections, we focus on building examples that demonstrate this non-trivial nature. We follow up the examples with theorems that generalize the examples.

3.3. LSDEs with anticipating coefficients

In this section and the next, we fix $t \in [0,1]$ unless stated otherwise.

In the following examples, we progressively increase the complexity of the diffusion coefficient of the stochastic differential equation and see its effect on the solution. This will help us to develop our intuition about the non-trivial nature of the results related to anticipating coefficients. The first three examples follow directly from remark 1.6.7.

Example 3.3.1. Let σ be a constant. The process

$$\mathcal{E}_t^{(\sigma)} = \exp\left[\sigma W_t - \frac{1}{2}\sigma^2 t\right]$$

is a solution of the stochastic differential equation

$$\begin{cases} d\mathcal{E}_t^{(\sigma)} = \sigma\,\mathcal{E}_t^{(\sigma)}\,\mathrm{d}W_t, \\ \mathcal{E}_0^{(\sigma)} = 1. \end{cases}$$

Example 3.3.2. Suppose $\sigma(t)$ is a deterministic function. The process

$$\mathcal{E}_t^{(\sigma)} = \exp\left[\int_0^t \sigma(s)\,\mathrm{d}W_s - \frac{1}{2}\int_0^t \sigma(s)^2\,\mathrm{d}s\right]$$

is a solution of the stochastic differential equation

$$\begin{cases} d\mathcal{E}_t^{(\sigma)} = \sigma_t\,\mathcal{E}_t^{(\sigma)}\,\mathrm{d}W_t, \\ \mathcal{E}_0^{(\sigma)} = 1. \end{cases}$$

Example 3.3.3. Consider the adapted coefficient $\sigma_t = W_t$. The process

$$X_t = \exp\left[\frac{1}{2}\left(W_t^2 - t - \int_0^t W_s^2 \, \mathrm{d}s\right)\right]$$

is a solution of the stochastic differential equation

$$\begin{cases} \mathrm{d}X_t = W_t \, X_t \, \mathrm{d}W_t, \\ X_0 = 1. \end{cases}$$

From the above examples and example 3.1.4, one might guess that the process

$$\begin{aligned} Z(t) &= \exp\left[\int_0^t W_1 \, \mathrm{d}W_s - \frac{1}{2}\int_0^t W_1^2 \, \mathrm{d}s\right] \\ &= \exp\left[W_1 \, W_t - t - \frac{1}{2} W_1^2 \, t\right] \end{aligned}$$

is a solution of the stochastic differential equation

$$\begin{cases} \mathrm{d}Z(t) = W_1 \, Z(t) \, \mathrm{d}W_t, \\ Z(0) = 1. \end{cases}$$

But this is not true. In fact, we can apply the generalized Itô formula to derive the following result.

Theorem 3.3.4 ([HKS19, theorem 3.3]). *The stochastic process*

$$Z(t) = \exp\left[W_1 \, W_t - t - \frac{1}{2} W_1^2 \, t\right]$$

is a solution of

$$\begin{cases} \mathrm{d}Z(t) = W_1 \, Z(t) \, \mathrm{d}W_t + W_1 \left(W_t - t W_1\right) Z(t) \, \mathrm{d}t, \\ Z(0) = 1. \end{cases}$$

Then what is the solution of the following stochastic differential equation?

$$\begin{cases} \mathrm{d}Z(t) = W_1 \, Z(t) \, \mathrm{d}W_t, \\ Z(0) = 1. \end{cases}$$

The answer is given by the following theorem.

Theorem 3.3.5 ([HKS19, theorem 3.1]). *The process*

$$Z(t) = \exp\left[W_1 \int_0^t e^{-(t-s)} \, \mathrm{d}W_s - t - \frac{1}{4} W_1^2 \left(1 - e^{-2t}\right)\right]$$

is a solution of the stochastic differential equation

$$\begin{cases} \mathrm{d}Z(t) = W_1 \, Z(t) \, \mathrm{d}W_t, \\ Z(0) = 1. \end{cases}$$

Therefore, we see that solutions of stochastic differential equations with anticipating coefficients are quite non-trivial.

3.4. LSDEs with anticipating initial conditions

We start our discussion on stochastic differential equations with anticipating initial conditions with the following example. We assume $t \in [0,1]$ in this section.

Example 3.4.1 ([AK08, examples 4.1–3]).

$$\begin{cases} \mathrm{d}X_t = X_t \,\mathrm{d}W_t, \\ X_0 = x \in \mathbb{R}. \end{cases} \tag{3.4.2}$$

It is well known that the solution to equation (3.4.2) is

$$X_t = xe^{W_t - \frac{1}{2}t}.$$

However, if we take this solution and replace x with W_1 and apply the generalized Itô formula to the resultant expression, we obtain a different stochastic differential equation. In particular,

$$Y_t = W_1 e^{W_t - \frac{1}{2}t} \tag{3.4.3}$$

is a solution of

$$\begin{cases} \mathrm{d}Y_t = Y_t \,\mathrm{d}W_t + \dfrac{1}{W_1} Y_t \,\mathrm{d}t, \\ Y_0 = W_1. \end{cases} \tag{3.4.4}$$

Here, the initial condition is outside the classical theory of Itô calculus since W_1 is not $\mathcal{F}_0$-measurable. We can use the generalized 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 W_1 terms in equation (3.4.4) with $x \in \mathbb{R}$ then we obtain the following stochastic differential equation

$$\begin{cases} \mathrm{d}Z_t = Z_t \,\mathrm{d}W_t + \dfrac{1}{x} Z_t \,\mathrm{d}t, \\ Z_0 = x \in \mathbb{R}. \end{cases}$$

with its solution

$$Z_t = xe^{W_t - \frac{1}{2}t + \frac{1}{x}t}.$$

The differences in equation (3.4.2) and equation (3.4.4) demonstrates that replacing the non anticipating term in the solution with an anticipating term yields an extra drift term in the SDE. Furthermore, by replacing all the anticipating terms in equation (3.4.4) with a real number, we obtained an extra drift factor in equation (3.4.3). These examples highlights some of the differences and interesting patterns between adapted linear stochastic differential equations and non-adapted ones.

Example 3.4.5 ([Kha+13, section 3]). Consider the following motivational example:

$$\begin{cases} \mathrm{d}X_t = X_t\,\mathrm{d}W_t, \\ X_0 = W_1. \end{cases}$$

Equation (3.4.2) would suggest that our solution would be equation (3.4.3). However, that is not the case. We have an extra drift term as demonstrated by equation (3.4.4). With that in mind, we "guess" that the solution has the form

$$X_t = (W_1 - \xi_t)\, e^{W_t - \frac{1}{2}t}$$

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

$$\mathrm{d}X_t = (W_1 - \xi_t)e^{W_t - \frac{1}{2}t}\,\mathrm{d}W_t + \left(e^{W_t - \frac{1}{2}t} - \xi_t' e^{W_t - \frac{1}{2}t}\right)\mathrm{d}t.$$

The $\mathrm{d}t$ term in the above equation must be zero for X_t to be a solution. Therefore, by solving the following ordinary differential equation

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

we get our solution

$$X_t = (W_1 - t)\, e^{W_t - \frac{1}{2}t}.$$

Motivated by the above results, we proved the following theorem in 2018 that provides solutions for a class of stochastic differential equations with anticipating initial conditions. Note that in this case, the coefficients α and β are deterministic.

Theorem 3.4.6 ([KSZ18, theorem 5.1]). *Let $\alpha, h \in L^2[0,1]$, $\beta \in L^1[0,1]$ and $\psi \in C^2(\mathbb{R})$. Then the solution of the stochastic differential equation*

$$\begin{cases} \mathrm{d}Z(t) = \alpha(t)Z(t)\,\mathrm{d}W_t + \beta(t)Z(t)\,\mathrm{d}t, \\ Z(0) = \psi\Big(\displaystyle\int_0^1 h(s)\,\mathrm{d}W_s\Big), \end{cases} \tag{3.4.7}$$

is given by

$$Z(t) = \psi\Big(\int_0^1 h(s)\,\mathrm{d}W_s - \int_0^t \alpha(s)h(s)\,\mathrm{d}s\Big)\,\mathcal{E}_t^{(\alpha,\beta)}.$$

Proof. Suppose $Z(t) = \psi\big(\int_0^1 h(s)\,\mathrm{d}W_s - u(t)\big)\mathcal{E}_t^{(\alpha,\beta)}$ is our ansatz solution, where the unknown function $u(t)$ has to be determined. In order to apply theorem 3.2.3, we write

$$Z(t) = \psi\Big(\int_0^t h(s)\,\mathrm{d}W_s + \int_t^1 h(s)\,\mathrm{d}W_s - u(t)\Big)\mathcal{E}_t^{(\alpha,\beta)}. \tag{3.4.8}$$

Motivated by this, we define

$$
\begin{cases}
X_t^{(1)} = \displaystyle\int_0^t h(s)\,\mathrm{d}W_s, \\
X_t^{(2)} = \mathcal{E}_t^{(\alpha,\beta)}, \\
\quad Y^t = \displaystyle\int_t^1 h(s)\,\mathrm{d}W_s, \text{ and}
\end{cases}
\tag{3.4.9}
$$
$$
\theta(t, x_1, x_2, y) = \psi(x_1 + y - u(t))x_2,
$$

so that $Z(t) = \theta(t, X_t^{(1)}, X_t^{(2)}, Y^t)$. From the definition of θ, we get the partial derivatives

$$
\begin{aligned}
\theta_t &= -\psi' u'(t)x_2, & \theta_{x_1} &= \psi' x_2, & \theta_{x_2} &= \psi, \\
\theta_{x_1x_1} &= \psi'' x_2, & \theta_{x_2x_2} &= 0, & \theta_{x_1x_2} &= \psi', \\
\theta_y &= \psi' x_2, & \theta_{yy} &= \psi'' x_2.
\end{aligned}
$$

From the definitions in equation (3.4.9), we have

$$
\begin{aligned}
\mathrm{d}X_t^{(1)} &= h(t)\,\mathrm{d}W_t, & \mathrm{d}X_t^{(2)} &= \alpha(t)\mathcal{E}_t^{(\alpha,\beta)}\,\mathrm{d}W_t + \beta(t)\mathcal{E}_t^{(\alpha,\beta)}\,\mathrm{d}t, \\
\left(\mathrm{d}X_t^{(1)}\right)^2 &= h(t)^2\,\mathrm{d}t, & \left(\mathrm{d}X_t^{(2)}\right)^2 &= \alpha(t)^2\left(\mathcal{E}_t^{(\alpha,\beta)}\right)^2\mathrm{d}t, \\
\mathrm{d}X_t^{(1)}\,\mathrm{d}X_t^{(2)} &= h(t)\alpha(t)\mathcal{E}_t^{(\alpha,\beta)}\,\mathrm{d}t, \\
\mathrm{d}Y^t &= -h(t)\,\mathrm{d}W_t, & \left(\mathrm{d}Y^t\right)^2 &= h(t)^2\,\mathrm{d}t.
\end{aligned}
$$

Applying the differential formula and putting everything together, we have

$$
\begin{aligned}
\mathrm{d}Z(t) &= \mathrm{d}\theta(t, X_t^{(1)}, X_t^{(2)}, Y^t) \\
&= \theta_t\,\mathrm{d}t + \theta_{x_1} h(t)\,\mathrm{d}W_t + \theta_{x_2}\,\mathrm{d}X_t^{(2)} \\
&\quad + \frac{1}{2}\theta_{x_1x_1}(\mathrm{d}X_t^{(1)})^2 + \frac{1}{2}\theta_{x_2x_2}(\mathrm{d}X_t^{(2)})^2 + \theta_{x_1x_2}(\mathrm{d}X_t^{(1)})(\mathrm{d}X_t^{(1)}) \\
&\quad + \theta_y\,\mathrm{d}Y^t - \frac{1}{2}\theta_{yy}\,\mathrm{d}Y^t \\
&= -\psi' u'(t)X_t^{(2)}\,\mathrm{d}t + \cancel{\psi' X_t^{(2)} h(t)}\,\mathrm{d}W_t + \psi\cdot\left[\alpha(t)\mathcal{E}_t^{(\alpha,\beta)}\,\mathrm{d}W_t + \beta(t)\mathcal{E}_t^{(\alpha,\beta)}\,\mathrm{d}t\right] \\
&\quad + \cancel{\frac{1}{2}\psi'' X_t^{(2)} h(t)^2}\,\mathrm{d}t + \frac{1}{2}\cdot 0\cdot\left\{\alpha(t)^2\left(\mathcal{E}_t^{(\alpha,\beta)}\right)^2\mathrm{d}t\right\} + \psi'\,h(t)\,\alpha(t)\,\mathcal{E}_t^{(\alpha,\beta)}\,\mathrm{d}t \\
&\quad - \cancel{\psi' X_t^{(2)} h(t)}\,\mathrm{d}W_t - \cancel{\frac{1}{2}\psi'' X_t^{(2)} h(t)^2}\,\mathrm{d}t \\
&= -\psi'\,u'(t)\,X_t^{(2)}\,\mathrm{d}t + \alpha(t)Z(t)\,\mathrm{d}W_t + \beta(t)Z(t)\,\mathrm{d}t + \psi'\,h(t)\,\alpha(t)\,\mathcal{E}_t^{(\alpha,\beta)}\,\mathrm{d}t \\
&= \alpha(t)Z(t)\,\mathrm{d}W_t + \beta(t)\,Z(t)\,\mathrm{d}t + \big(\alpha(t)\,h(t) - u'(t)\big)\psi'\mathcal{E}_t^{(\alpha,\beta)}\,\mathrm{d}t,
\end{aligned}
$$

where in the fourth equality we used $Z(t) = \psi\cdot\mathcal{E}_t^{(\alpha,\beta)}$.

Therefore, in order for $Z(t)$ to be the solution of equation (3.4.7), we need the condition $u'(t) = \alpha(t)h(t)$. On the other hand, if we put $t = 0$ in equation (3.4.8), we get $X_0 = \psi\big(\int_0^1 h(s)\,\mathrm{d}W_s - u(0)\big)$. Since $Z(t)$ is the solution of equation (3.4.7), comparing this with the initial condition gives us $u(0) = 0$. Thus we have the following ordinary differential equation for $t \in [0,1]$

$$\begin{cases} u'(t) = \alpha(t)h(t), \\ u(0) = 0, \end{cases}$$

whose solution is $u(t) = \int_0^t \alpha(s)h(s)\,\mathrm{d}s$. Therefore

$$Z(t) = \psi\Big(\int_0^1 h(s)\,\mathrm{d}W_s - \int_0^t \alpha(s)h(s)\,\mathrm{d}s\Big)\mathcal{E}_t^{(\alpha,\beta)}.$$

□

Example 3.4.10 ([KSZ18, example 5.2])**.** Consider the stochastic differential equation

$$\begin{cases} \mathrm{d}Z(t) = Z(t)\,\mathrm{d}W_t, \\ \quad Z_0 = \displaystyle\int_0^1 W_s\,\mathrm{d}s. \end{cases}$$

To solve this, we reformulate the initial condition as

$$\int_0^1 W_s\,\mathrm{d}s = -W_s(1-s)\Big|_0^1 + \int_0^1 (1-s)\,\mathrm{d}W_s = \int_0^1 (1-s)\,\mathrm{d}W_s.$$

Thus we have a special case of Theorem 3.4.6 with $\alpha(t) \equiv 1$, $\beta \equiv 0$, $h(t) = 1-t$, and $\psi(x) = x$. Therefore, the solution is given by

$$Z(t) = \Big(\int_0^1 W_s\,\mathrm{d}s - \big(t - \frac{1}{2}t^2\big)\Big)e^{W_t - \frac{1}{2}t}.$$

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

In 2021, we generalized theorem 3.4.6 to allow for adapted processes as coefficients.

Theorem 3.4.11 ([KSS21b, theorem 4.2])**.** *Let $\alpha \in L^2_{ad}([a,b]\times\Omega)$, $\beta \in L^1_{ad}([a,b]\times\Omega)$ be stochastic processes. Suppose $h \in L^2[a,b]$ and $\psi \in C^2(\mathbb{R})$ are deterministic functions. Then the solution of the stochastic differential equation*

$$\begin{cases} \mathrm{d}Z(t) = \alpha_t\,Z(t)\,\mathrm{d}W_t + \beta_t\,Z(t)\,\mathrm{d}t, \\ \quad Z(a) = \psi\Big(\displaystyle\int_a^b h(s)\,\mathrm{d}W_s\Big), \end{cases} \tag{3.4.12}$$

is given by

$$Z(t) = \psi\Big(\int_a^b h(s)\,\mathrm{d}W_s - \int_a^t h(s)\,\alpha_s\,\mathrm{d}s\Big)\,\mathcal{E}_t^{(\alpha,\beta)}. \tag{3.4.13}$$

Proof. Suppose $Z(t) = \psi\big(\int_a^b h(s)\,\mathrm{d}W_s - U_t\big)\mathcal{E}_t^{(\alpha,\beta)}$ is our ansatz solution. We need to determine the unknown Itô process U_t with $U_a = 0$. In order to apply theorem 3.2.3, we write

$$Z(t) = \psi\Big(\int_a^t h(s)\,\mathrm{d}W_s - U_t + \int_t^b h(s)\,\mathrm{d}W_s\Big)\,\mathcal{E}_t^{(\alpha,\beta)}. \tag{3.4.14}$$

Define the instantly-independent process $Y^t = \int_t^b h(s)\,\mathrm{d}W_s$ and the adapted processes

$$X_t^{(1)} = \mathcal{E}_t^{(\alpha,\beta)} \quad \text{and} \quad X_t^{(2)} = \int_a^t h(s)\,\mathrm{d}W_s - U_t.$$

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

$$\begin{aligned}
\mathrm{d}X_t^{(1)} &= \alpha_t X_t^{(1)}\,\mathrm{d}W_t + \beta_t X_t^{(1)}\,\mathrm{d}t,\\
\mathrm{d}X_t^{(2)} &= h(t)\,\mathrm{d}W_t - \mathrm{d}U_t,\\
(\mathrm{d}X_t^{(1)})^2 &= \alpha_t^2 (X_t^{(1)})^2\,\mathrm{d}t,\\
(\mathrm{d}X_t^{(2)})^2 &= h(t)^2\,\mathrm{d}t - 2h(t)\,\mathrm{d}W_t\,\mathrm{d}U_t + (\mathrm{d}U_t)^2,\\
\mathrm{d}X_t^{(1)}\,\mathrm{d}X_t^{(2)} &= h(t)\alpha_t X_t^{(1)}\,\mathrm{d}t - \alpha_t X_t^{(1)}\,\mathrm{d}W_t\,\mathrm{d}U_t,\\
\mathrm{d}Y^t &= -h(t)\,\mathrm{d}W_t,\\
(\mathrm{d}Y^t)^2 &= h(t)^2\,\mathrm{d}t.
\end{aligned}$$

Now, define $\theta(x_1, x_2, y) = \psi(x_2 + y)x_1$, so that $Z(t) = \theta\Big(X_t^{(1)}, X_t^{(2)}, Y^t\Big)$. 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 3.2.3 and putting everything together, we can easily find the stochastic

differential of $Z(t)$:

$$\begin{aligned}
\mathrm{d}Z(t) &= \mathrm{d}\theta(X_t^{(1)}, X_t^{(2)}, Y^t) \\
&= \theta_{x_1}\,\mathrm{d}X_t^{(1)} + \theta_{x_2}\,\mathrm{d}X_t^{(2)} \\
&\quad + \frac{1}{2}\theta_{x_1x_1}(\mathrm{d}X_t^{(1)})^2 + \frac{1}{2}\theta_{x_2x_2}(\mathrm{d}X_t^{(2)})^2 \\
&\quad + \theta_{x_1x_2}(\mathrm{d}X_t^{(1)})(\mathrm{d}X_t^{(2)}) \\
&\quad + \theta_y\,\mathrm{d}Y^t - \frac{1}{2}\theta_{yy}(\mathrm{d}Y^t)^2 \\
&= \psi\cdot\Big(\alpha_t X_t^{(1)}\,\mathrm{d}W_t + \beta_t X_t^{(1)}\,\mathrm{d}t\Big) + \psi'\cdot X_t^{(1)}\,(\cancel{h(t)\,\mathrm{d}W_t} - \mathrm{d}U_t) \\
&\quad + 0 + \frac{1}{2}\psi''\cdot X_t^{(1)}\,\big(\cancel{h(t)^2\,\mathrm{d}t} - 2h(t)\,\mathrm{d}W_t\,\mathrm{d}U_t + (\mathrm{d}U_t)^2\big) \\
&\quad + \psi'\,\Big(h(t)\alpha_t X_t^{(1)}\,\mathrm{d}t - \alpha_t X_t^{(1)}\,\mathrm{d}W_t\,\mathrm{d}U_t\Big) \\
&\quad - \cancel{\psi'\cdot X_t^{(1)}h(t)\,\mathrm{d}W_t} - \cancel{\frac{1}{2}\psi''\cdot X_t^{(1)}h(t)^2\,\mathrm{d}t} \\
&= \psi\cdot\Big(\alpha_t X_t^{(1)}\,\mathrm{d}W_t + \beta_t X_t^{(1)}\,\mathrm{d}t\Big) \\
&\quad + \psi'\cdot X_t^{(1)}\cdot(-\,\mathrm{d}U_t + h(t)\alpha_t\,\mathrm{d}t - \alpha_t\,\mathrm{d}W_t\,\mathrm{d}U_t) \\
&\quad + \frac{1}{2}\psi''\cdot X_t^{(1)}\cdot\big(-2h(t)\,\mathrm{d}W_t\,\mathrm{d}U_t + (\mathrm{d}U_t)^2\big).
\end{aligned}$$

Therefore, in order for $Z(t)$ to be the solution of equation (3.4.12), we need to satisfy the following conditions

$$\mathrm{d}U_t = h(t)\alpha_t\,\mathrm{d}t - \alpha_t\,\mathrm{d}W_t\,\mathrm{d}U_t \tag{3.4.15}$$

$$(\mathrm{d}U_t)^2 = 2h(t)\,\mathrm{d}W_t\,\mathrm{d}U_t \tag{3.4.16}$$

From equation (3.4.15), we see that if $\mathrm{d}U_t$ contains only a $\mathrm{d}t$ term (no $\mathrm{d}W_t$ term), then $\mathrm{d}U_t\,\mathrm{d}W_t = 0$. On the other hand, if $\mathrm{d}U_t$ contains a $\mathrm{d}W_t$ term, then $\mathrm{d}U_t\,\mathrm{d}W_t = \gamma_t\,\mathrm{d}t$ for some γ_t. Then we have $\mathrm{d}U_t = (h(t) - \gamma_t)\alpha_t\,\mathrm{d}t$, which again gives $\mathrm{d}U_t\,\mathrm{d}W_t = 0$. Therefore, in either case, $\mathrm{d}U_t = h(t)\alpha_t\,\mathrm{d}t$. Note that this also agrees with equation (3.4.16).

Imposing the initial condition $U_a = 0$, we get that $U_t = \int_a^t h(s)\,\alpha_s\,\mathrm{d}s$. Putting this in the assumed form of the solution, we get our result. □

Now we look at a specific case of theorem 3.4.11 where $h(t) \equiv 1$.

Corollary 3.4.17 ([KSS21b, corollary 4.3])**.** *Under the same assumptions for α_t, β_t and ψ as in theorem 3.4.11, the solution of the stochastic differential equation*

$$\begin{cases} \mathrm{d}Z(t) = \alpha_t Z(t)\,\mathrm{d}W_t + \beta_t Z(t)\,\mathrm{d}t, & t \in [a,b], \\ Z(a) = \psi\Big(W_b - W_a\Big), \end{cases}$$

is given by

$$Z(t) = \psi\Big(W_b - W_a - \int_a^t \alpha_s\,\mathrm{d}s\Big)\mathcal{E}_t^{(\alpha,\beta)}.$$

Remark 3.4.18. Corollary 3.4.17 extends [Kha+13, theorem 4.1] to include adapted coefficients for the anticipating stochastic differential equation.

We apply these new results to obtain solutions for some examples of 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 $X_0 = W_1$. The second example demonstrates a case where the initial condition is a Riemann integral of a Brownian motion.

Example 3.4.19 ([KSS21b, example 4.5])**.** Consider the stochastic differential equation

$$\begin{cases} \mathrm{d}X(t) = W_t\, X(t)\,\mathrm{d}W_t + X(t)\,\mathrm{d}t, \quad t \in [0,1], \\ \quad X_0 = W_1. \end{cases}$$

Here $\alpha_t = W_t$, $\beta_t \equiv 1$, $h(t) \equiv 1$, and $\psi(x) = x$. Thus, by corollary 3.4.17, we have the solution

$$X(t) = \left(W_1 - \int_0^t W_s\,\mathrm{d}s\right)\exp\left[\frac{1}{2}\left(W_t^2 + t - \int_0^t W_s^2\,\mathrm{d}s\right)\right].$$

Example 3.4.20 ([KSS21b, example 4.6])**.** Consider the stochastic differential equation

$$\begin{cases} \mathrm{d}X(t) = W_t\, X(t)\,\mathrm{d}W_t, \quad t \in [0,1], \\ \quad X_0 = \displaystyle\int_0^1 W_s\,\mathrm{d}s. \end{cases}$$

As in example 3.4.10, we use stochastic integration by parts to modify the initial condition. Namely,

$$\int_0^1 W_s\,\mathrm{d}s = \int_0^1 (1-s)\,\mathrm{d}W_s.$$

Hence with $\alpha_t = W_t$, $\beta_t \equiv 0$, $h(t) = 1 - t$, and $\psi(x) = x$ in theorem 3.4.11, we have the solution

$$X(t) = \left(\int_0^1 W_s\,\mathrm{d}s - \int_0^t (1-s)W_s\,\mathrm{d}s\right)\exp\left[\frac{1}{2}\left(W_t^2 - t - \int_0^t W_s^2\,\mathrm{d}s\right)\right].$$

Chapter 4. Extension of Itô's Isometry

4.1. Motivation

We remarked in section 1.5 that Itô's isometry forms the backbone of Itô's integral. Therefore, for any theory that extends Itô's theory, it is important to know if the isometry holds. The focus of this chapter will be to derive an identity in Ayed–Kuo theory that reduces to Itô isometry for adapted processes.

Should we expect Itô's isometry to hold identically for the Ayed–Kuo integral? From white noise distribution theory, we expect otherwise. We take a brief detour to explore this idea.

Recall that the white nose approach to anticipating stochastic integral is given by the Hitsuda–Skorokhod integral

$$\int_a^b \partial_t^* Z(t)\,\mathrm{d}t,$$

where Z is a Wiener functional and ∂_t^* is the adjoint of the white noise differential operator ∂_t [see Kuo96, page 107]. (For a detailed description of the Hitsuda–Skorokhod integral, see [Kuo96, chapter 5].) Essentially, [Kuo96, theorem 13.16] asserts that under certain conditions on Z, the white noise integral $\int_a^b \partial_t^* Z(t)\,\mathrm{d}t$ is a Hitsuda–Skorokhod integral, and we have

$$\mathbb{E}\left[\left(\int_a^b \partial_t^* Z(t)\,\mathrm{d}t\right)^2\right] = \int_a^b \mathbb{E}\big[Z(t)^2\big]\,\mathrm{d}t + \int_a^b \int_a^b \mathbb{E}[\partial_t Z(s)\,\partial_s Z(t)]\,\mathrm{d}s\,\mathrm{d}t. \tag{4.1.1}$$

Moreover, from white noise analysis, we have the symbolic expression

$$\partial_t W_s = \mathbb{1}_{\{t<s\}}. \tag{4.1.2}$$

We can verify this expression using the last line on [Kuo96, page 103] and the representation of Wiener process in [Kuo96, page 254].

Now, let f and ϕ be C^1-functions on $\mathbb{R}$ and consider the Wiener functional

$$Z(t) = f(W_t)\,\phi(W_b - W_t), \quad t \in [a, b] \tag{4.1.3}$$

Suppose $s < t$. Then by the chain rule and equation (4.1.2), we have

$$\begin{aligned} \partial_t Z(s) &= f'(W_s)\;(\partial_t W_s)\;\phi(W_b - W_s) + f(W_s)\,\phi'(W_b - W_s)\;(\partial_t(W_b - W_s)) \\ &= f(W_s)\,\phi'(W_b - W_s). \end{aligned} \tag{4.1.4}$$

This chapter previously appeared in the following open-access journal article: Hui-Hsiung Kuo, Pujan Shrestha, and Sudip Sinha. "An Intrinsic Proof of an Extension of Itô's Isometry for Anticipating Stochastic Integrals". In: *Journal of Stochastic Analysis* 2.4 (2021). DOI: 10.31390/josa.2.4.08.

Similarly,

$$\partial_s Z(t) = f(W_t)\,\phi'(W_b - W_t). \tag{4.1.5}$$

Putting equations (4.1.1) and (4.1.3) to (4.1.5) together, we get

$$\begin{aligned}\mathbb{E}\left[\left(\int_a^b \partial_t^*\,(f(W_t)\,\phi(W_b - W_t))\,\mathrm{d}t\right)^2\right] &= \int_a^b \mathbb{E}\big[f(W_t)^2\,\phi(W_b - W_t)^2\big]\,\mathrm{d}t \\ &+ 2\int_a^b \int_a^t \mathbb{E}[f(W_s)\,\phi'(W_b - W_s)\,f'(W_t)\,\phi(W_b - W_t)]\,\mathrm{d}s\,\mathrm{d}t. \end{aligned}\tag{4.1.6}$$

But for a Wiener functional $Z(t) = f(W_t)\,\phi(W_b - W_t)$, its Hitsuda–Skorokhod integral turns out to be the same as the Ayed–Kuo integral, namely

$$\int_a^b \partial_t^*\,(f(W_t)\,\phi(W_b - W_t))\,\mathrm{d}t = \int_a^b f(W_t)\,\phi(W_b - W_t)\,\mathrm{d}W_t. \tag{4.1.7}$$

Therefore, equations (4.1.6) and (4.1.7)) yields the following identity

$$\begin{aligned}\mathbb{E}\left[\left(\int_a^b f(W_t)\,\phi(W_b - W_t)\,\mathrm{d}W_t\right)^2\right] &= \int_a^b \mathbb{E}\big[f(W_t)^2\,\phi(W_b - W_t)^2\big]\,\mathrm{d}t \\ &+ 2\int_a^b \int_a^t \mathbb{E}[f(W_s)\,\phi'(W_b - W_s)\,f'(W_t)\,\phi(W_b - W_t)]\,\mathrm{d}s\,\mathrm{d}t. \end{aligned}\tag{4.1.8}$$

which is what we expect for the Ayed–Kuo integral.

4.2. Previous attempts

It is important to note that this is not the first attempt to formulate such a result. In fact, in [KSS13, theorem 3.1], the authors demonstrated that equation (4.1.8) holds under strong assumptions of f and ϕ being analytic on the entire real line and having bounded derivatives.

However, the earlier result had significant drawbacks. First, the assumptions on f and ϕ are quite restrictive as noted. Second, the proof of the result is quite lengthy, and involves tedious computations using binomial expansion of the function f and ϕ. Third, the proof did not utilize the core idea of the Ayed–Kuo integral, that is, to evaluate adapted integrands at the left-endpoints and the instantly-independent processes at the right-endpoints of the intervals. Our new proof exploits this nature of the integral and resolves the stated shortcomings. We also do not imposing any restrictions on f and ϕ outside of what is required for equation (4.1.8) to be well-defined.

4.3. Identity for the simple case

In the proof of theorem 4.3.1 below, for $a \leq s \leq t \leq b$, we shall use the σ-algebras

$$\begin{aligned} \mathcal{F}_s &= \sigma\{W_u \mid a \leq u \leq s\}, \\ \mathcal{G}^t &= \sigma\{W_b - W_v | t \leq v \leq b\}, \text{ and} \\ \mathcal{H}_s^t &= \sigma\left(\mathcal{F}_s \cup \mathcal{G}^t\right). \end{aligned}$$

We call $\{\mathcal{F}_s : s \in [a,b]\}$ the *forward*-filtration and $\{\mathcal{G}^t : t \in [a,b]\}$ the *backward* or *counter*-filtration generated by the Wiener process. Taking conditional expectation judiciously with respect to the *separation σ-algebra* $\mathcal{H}_s^t$ plays a crucial part in the proof of the following theorem.

Theorem 4.3.1. *Suppose $f, \phi \in C^1(\mathbb{R})$ such that $f(W_t)\,\phi(W_b - W_t), f(W_t)\,\phi'(W_b - W_t), f'(W_t)\,\phi(W_b - W_t) \in L^2([a,b] \times \Omega)$. Moreover, assume the stronger condition of convergence in $L^2(\Omega)$ instead of a convergence in probability for the definition of the Ayed–Kuo integral. Then $\mathbb{E}\left[\int_a^b f(W_t)\,\phi(W_b - W_t)\,\mathrm{d}W_t\right] = 0$, and*

$$\begin{aligned} \mathbb{E}\left[\left(\int_a^b f(W_t)\,\phi(W_b - W_t)\,\mathrm{d}W_t\right)^2\right] &= \int_a^b \mathbb{E}\left[f(W_t)^2\,\phi(W_b - W_t)^2\right]\mathrm{d}t \\ &+ 2\int_a^b \int_a^t \mathbb{E}[f(W_s)\,\phi'(W_b - W_s)\,f'(W_t)\,\phi(W_b - W_t)]\,\mathrm{d}s\,\mathrm{d}t. \end{aligned} \tag{4.3.2}$$

Remark 4.3.3. For the right-hand side of equation (4.3.2) to be well-defined, we need the two integrals to be individually well-defined.

For the first integral, we directly see that the integral is well-defined if $f(W_t)\,\phi(W_b - W_t) \in L^2([a,b] \times \Omega)$. For conciseness, we write $f_t = f(W_t)$, $\phi_t = \phi(W_b - W_t)$, and similarly their corresponding derivatives. Using this notation, for the second integral, we can use Cauchy–Schwarz inequality to get

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

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 $f(W_t)\,\phi(W_b - W_t), f(W_t)\,\phi'(W_b - W_t), f'(W_t)\,\phi(W_b - W_t) \in L^2([a,b] \times \Omega)$.

Remark 4.3.4. In the proof of Itô's isometry, one typically takes conditional expectation with respect to the σ-algebra $\mathcal{F}_s$ in a simple way. On the other hand, our proof requires conditioning with respect to the σ-algebra $\mathcal{H}_s^t$ in a very specific manner.

Proof. For notational convenience, let

$$\begin{aligned}
\Delta W_k &= W_{t_k} - W_{t_{k-1}}, \\
\Delta t_k &= t_k - t_{k-1}, \\
f_{k-1} &= f(W_{t_{k-1}}), \text{ and} \\
\phi_k &= \phi(W_b - W_{t_k}).
\end{aligned}$$

Note that in this proof, we use ϕ_k (subscript k) to represent the time even though ϕ is instantly-independent. We use superscripts only for powers. We also assume that $\|\Pi_n\| \to 0$ as $n \to 0$ as usual.

By the definition of the anticipating stochastic integral, we get

$$\int_a^b f(W_t)\,\phi(W_b - W_t)\,\mathrm{d}W_t = \lim_{n\to\infty} \sum_{i=1}^n f_{i-1}\phi_i\,\Delta W_i,$$

where the limit is in $L^2(\Omega)$. Therefore, for the mean, we have

$$\begin{aligned}
\mathbb{E}\left[\int_a^b f(W_t)\,\phi(W_b - W_t)\,\mathrm{d}W_t\right] &= \lim_{n\to\infty} \sum_{i=1}^n \mathbb{E}[f_{i-1}\,\phi_i\,\Delta W_i] \\
&= \lim_{n\to\infty} \sum_{i=1}^n \mathbb{E}\Big[\mathbb{E}\Big(f_{i-1}\,\phi_i\,\Delta W_i \;\Big|\; \mathcal{H}_{t_{j-1}}^{t_j}\Big)\Big] \\
&= \lim_{n\to\infty} \sum_{i=1}^n \mathbb{E}\Big[f_{i-1}\,\phi_i\,\cancelto{0}{\mathbb{E}(\Delta W_i)}\Big] \\
&= 0,
\end{aligned}$$

where we used the fact that f_{i-1} and ϕ_i are $\mathcal{H}_{t_{j-1}}^{t_j}$-measurable, and ΔW_i is independent of $\mathcal{H}_{t_{j-1}}^{t_j}$.

For the variance,

$$\begin{aligned}
&\mathbb{E}\left[\left(\int_a^b f(W_t)\,\phi(W_b - W_t)\,\mathrm{d}W_t\right)^2\right] \\
&= \lim_{n\to\infty} \sum_{i=1}^n \sum_{j=1}^n \mathbb{E}\big[f_{i-1}\,\phi_i f_{j-1}\,\phi_j\,\Delta W_i\,\Delta W_j\big] \\
&= \lim_{n\to\infty} \sum_{i=1}^n \mathbb{E}\Big[f_{i-1}^2\,\phi_i^2\,(\Delta W_i)^2\Big] + 2\lim_{n\to\infty} \sum_{j=1}^n \sum_{i=1}^j \mathbb{E}\big[f_{i-1}\,\phi_i f_{j-1}\,\phi_j\,\Delta W_i\,\Delta W_j\big] \\
&=: D_0 + 2D_1,
\end{aligned}$$

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 ΔW_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

$$\begin{aligned}\mathbb{E}\Big[f_{i-1}^2\,\phi_i^2\,(\Delta W_i)^2\Big] &= \mathbb{E}\Big[\mathbb{E}\Big(f_{i-1}^2\,\phi_i^2\,(\Delta W_i)^2 \;\Big|\; \mathcal{F}_{t_{i-1}}\Big)\Big]\\ &= \mathbb{E}\Big[f_{i-1}^2\;\mathbb{E}\Big(\phi_i^2\,(\Delta W_i)^2 \;\Big|\; \mathcal{F}_{t_{i-1}}\Big)\Big]\\ &= \mathbb{E}\big[f_{i-1}^2\big]\;\mathbb{E}\Big[\phi_i^2\,(\Delta W_i)^2\Big].\end{aligned}$$

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

$$\begin{aligned}\mathbb{E}\Big[\phi_i^2\,(\Delta W_i)^2\Big] &= \mathbb{E}\Big[\mathbb{E}\Big(\phi_i^2\,(\Delta W_i)^2 \;\Big|\; \mathcal{G}^{t_i}\Big)\Big]\\ &= \mathbb{E}\Big[\phi_i^2\;\mathbb{E}\Big((\Delta W_i)^2 \;\Big|\; \mathcal{G}^{t_i}\Big)\Big]\\ &= \mathbb{E}\big[\phi_i^2\big]\;\mathbb{E}\Big[(\Delta W_i)^2\Big].\end{aligned}$$

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

$$\mathbb{E}\Big[f_{i-1}^2\,\phi_i^2\,(\Delta W_i)^2\Big] = \mathbb{E}\big[f_{i-1}^2\big]\,\mathbb{E}\big[\phi_i^2\big]\,\Delta t_i = \mathbb{E}\big[f(W_t)^2\phi(W_b - W_t)^2\big]\,\Delta t_i,$$

where we used the independence of increments of Wiener process in the last equality. Summing over i and taking limit of $n \to \infty$, we get the first term on the right side of equation (4.3.2) as

$$D_0 = \int_a^b \mathbb{E}\big[f(W_t)^2\,\phi(W_b - W_t)^2\big]\,\mathrm{d}t.$$

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

Trick 1 Note that ΔW_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

$$\begin{aligned}\mathbb{E}\Big(\Delta W_i \;\Big|\; \mathcal{H}_{t_{i-1}}^{t_i}\Big) &= \mathbb{E}[\Delta W_i] = 0, \text{ and}\\ \mathbb{E}\Big((\Delta W_i)^2 \;\Big|\; \mathcal{H}_{t_{i-1}}^{t_i}\Big) &= \mathbb{E}\big[(\Delta W_i)^2\big] = \Delta t_i.\end{aligned}$$

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

$$\begin{aligned}&\mathbb{E}\big[f_{i-1}\,\phi(W_b - W_{t_i} - \Delta W_j)\,f_{j-1}\,\phi_j\,\Delta W_i\,\Delta W_j\big]\\ &= \mathbb{E}\Big[\mathbb{E}\Big(f_{i-1}\,\phi(W_b - W_{t_i} - \Delta W_j)\,f_{j-1}\,\phi_j\,\Delta W_i\,\Delta W_j \;\Big|\; \mathcal{H}_{t_{j-1}}^{t_j}\Big)\Big]\\ &= \mathbb{E}\Big[f_{i-1}\,\phi(W_b - W_{t_i} - \Delta W_j)\,f_{j-1}\,\phi_j\,\Delta W_i\;\mathbb{E}\Big(\Delta W_j \;\Big|\; \mathcal{H}_{t_{j-1}}^{t_j}\Big)\Big]\\ &= 0.\end{aligned} \tag{4.3.5}$$

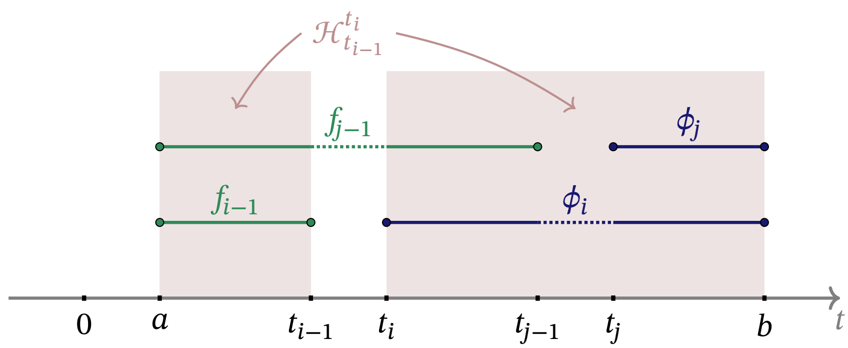

Figure 4.1. A t-dependence plot of the various processes. The dotted regions are removed.

Therefore, if we subtract the term $\mathbb{E}\big[f_{i-1}\,\phi(W_b - W_{t_i} - \Delta W_j)\,f_{j-1}\,\phi_j\,\Delta W_i\,\Delta W_j\big]$ from the expression $\mathbb{E}\big[f_{i-1}\,\phi_i f_{j-1}\,\phi_j\,\Delta W_i\,\Delta W_j\big]$, nothing changes. This allows us to remove the dependence of ϕ_i on $\big\{W_t\,:\,t \in (t_{j-1}, t_j)\big\}$. This is illustrated in figure 4.1 by the dotted region of ϕ_i.

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

$$\phi(W_b - W_{t_i}) - \phi(W_b - W_{t_i} - \Delta W_j) \simeq \phi'(W_b - W_{t_i} - \Delta W_j)\,\Delta W_j.$$

For brevity, we write $\Theta = \phi'(W_b - W_{t_i} - \Delta W_j)$. Note that Θ is $\mathcal{H}_{t_{i-1}}^{t_i}$-measurable.

Putting these together, we see that

$$\begin{aligned}
&\mathbb{E}\big[f_{i-1}\phi_i f_{j-1}\phi_j\,\Delta W_i\,\Delta W_j\big]\\
&= \mathbb{E}\big[f_{i-1}\left(\phi(W_b - W_{t_i}) - \phi(W_b - W_{t_i} - \Delta W_j)\right) f_{j-1}\phi_j\,\Delta W_i\,\Delta W_j\big]\\
&\simeq \mathbb{E}\big[f_{i-1}\Theta f_{j-1}\phi_j\,\Delta W_i(\Delta W_j)^2\big]\\
&= \mathbb{E}\Big[\mathbb{E}\Big(f_{i-1}\Theta f_{j-1}\phi_j\,\Delta W_i(\Delta W_j)^2\;\Big|\;\mathcal{H}_{t_{j-1}}^{t_j}\Big)\Big]\\
&= \mathbb{E}\Big[f_{i-1}\Theta f_{j-1}\phi_j\,\Delta W_i\;\mathbb{E}\big(\Delta W_j\big)^2\Big]\\
&= \mathbb{E}\big[f_{i-1}\Theta f_{j-1}\phi_j\,\Delta W_i\big]\,\Delta t_j.
\end{aligned} \tag{4.3.6}$$

We repeat Trick 2 on $f(W_{t_{j-1}} - \Delta W_i)$ just as we did for $\phi(W_b - W_{t_i} - \Delta W_j)$ to derive equation (4.3.5). This allows us to remove the dependence of f_{j-1} on $\{W_t\,:\,t \in (t_{i-1}, t_i)\}$. This is illustrated in figure 4.1 by the dotted region of f_{j-1}. Therefore,

$$\mathbb{E}\big[f_{i-1}\Theta f_{j-1}\phi_j\,\Delta W_i\,\Delta W_j\big] = 0,$$

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

$$f(W_{t_{j-1}}) - f(W_{t_{j-1}} - \Delta W_i) \simeq f'(W_{t_{j-1}} - \Delta W_i)\,\Delta W_i = F\,\Delta W_i,$$

where we write $F = f'(W_{t_{j-1}} - \Delta W_i)$ for brevity.

Continuing from equation (4.3.6),

$$\begin{aligned}
&\mathbb{E}\big[f_{i-1}\,\phi_i\,f_{j-1}\,\phi_j\,\Delta W_i\,\Delta W_j\big]\\
&= \mathbb{E}\big[f_{i-1}\,\Theta\,f_{j-1}\,\phi_j\,\Delta W_i\big]\,\Delta t_j\\
&= \mathbb{E}\Big[f_{i-1}\,\Theta\Big(f(W_{t_{j-1}}) - f(W_{t_{j-1}} - \Delta W_i)\Big)\,\phi_j\,\Delta W_i\Big]\,\Delta t_j\\
&\simeq \mathbb{E}\big[f_{i-1}\,\Theta\,F\,\phi_j(\Delta W_i)^2\big]\,\Delta t_j\\
&= \mathbb{E}\Big[\mathbb{E}\Big(f_{i-1}\,\Theta\,F\phi_j(\Delta W_i)^2\;\Big|\;\mathcal{H}_{t_{i-1}}^{t_i}\Big)\Big]\,\Delta t_j\\
&= \mathbb{E}\Big[f_{i-1}\,\Theta\,F\phi_j\;\mathbb{E}\Big((\Delta W_i)^2\;\Big|\;\mathcal{H}_{t_{i-1}}^{t_i}\Big)\Big]\,\Delta t_j\\
&= \mathbb{E}\big[f_{i-1}\,\Theta\,F\phi_j\big]\,\Delta t_i\,\Delta t_j.
\end{aligned} \tag{4.3.7}$$

By the continuity of W_t, we see that as $\|\Pi_n\| \to 0$, so does ΔW_i and ΔW_j. Moreover, by the continuity of f' and ϕ', we can conclude that as $\|\Pi_n\| \to 0$,

$$\begin{aligned}
F &= f'(W_{t_{i-1}} - \Delta W_i) \to f'(W_{t_{j-1}}) = f'_{j-1}, \text{ and}\\
\Theta &= \phi'(W_b - W_{t_i} - \Delta W_j) \to \phi'(W_b - W_{t_i}) = \phi'_i.
\end{aligned}$$

Finally, summing equation (4.3.7) over $i < j$ and taking limit, we get

$$D_1 = \int_a^b \int_a^t \mathbb{E}[f(W_s)\,\phi'(W_b - W_s)\,f'(W_t)\,\phi(W_b - W_t)]\,\mathrm{d}s\,\mathrm{d}t.$$

This concludes the proof. □

One of the features of theorem 4.3.1 is that the result enables us to evaluate the second moment of these anticipating integrals without having to evaluate the integral itself. This is advantageous as explicitly evaluating the integral via the definition can get very complicated. We demonstrate that particular feature with an example.

Example 4.3.8. Let $f(x) = x$ and $\phi(y) = y$. Then from theorem 4.3.1,

$$\begin{aligned}
&\mathbb{E}\left[\left(\int_a^b W_t\,(W_b - W_t)\,\mathrm{d}W_t\right)^2\right]\\
&= \int_a^b \mathbb{E}\big[W_t^2\,(W_b - W_t)^2\big]\,\mathrm{d}t + 2\int_a^b \int_a^t \mathbb{E}[W_s\,(W_b - W_t)]\,\mathrm{d}s\,\mathrm{d}t\\
&= \int_a^b \mathbb{E}\big[W_t^2\big]\;\mathbb{E}\big[(W_b - W_t)^2\big]\,\mathrm{d}t + 2\int_a^b \int_a^t \mathbb{E}[W_s]\;\mathbb{E}[W_b - W_t]\,\mathrm{d}s\,\mathrm{d}t\\
&= \int_a^b t(b-t)\,\mathrm{d}t\\
&= \frac{1}{6}(b^3 - 3a^2 b + 2a^3).
\end{aligned} \tag{4.3.9}$$

Note that this result can also be obtained by explicitly calculating the integral and then taking the expectation of the square. However, this process is extremely tedious and is impractical since we do not always have a closed-form representation of a stochastic integral. We briefly highlight this route mentioning only the key steps.

Let $I = \int_a^b W_t\,(W_b - W_t)\,\mathrm{d}W_t$. From example 3.1.6, we get

$$I = \frac{1}{2} W_b \left((W_b^2 - W_a^2) - (b-a) \right) - \frac{1}{3} \left(W_b^3 - W_a^3 \right).$$

For brevity, we write $\Delta_W = W_b - W_a$, so $W_b = (W_b - W_a) + W_a = \Delta_W + W_a$. Performing algebraic simplification, we get

$$I = \frac{1}{6} \left(\Delta_W^3 + 3 W_a \Delta_W^2 - 3(b-a)\Delta_W - 3(b-a) W_a \right).$$

Note that W_a and Δ_W are independent with $W_a \sim N(0, a)$ and $\Delta_W \sim N(0, (b-a))$. Therefore, any odd moment of either of W_a or Δ_W is zero. Using this, we get

$$\begin{aligned} \mathbb{E}[I^2] &= \frac{1}{36}\ \mathbb{E}\big[\Delta_W^6 + 9 W_a^2 \Delta_W^4 + 9(b-a)^2 \Delta_W^2 + 9(b-a)^2 W_a^2 \\ &\qquad\qquad -6(b-a)\Delta_W^4 - 18(b-a) W_a^2 \Delta_W^2\big] \\ &= \frac{1}{6} \left(b^3 - 3a^2 b + 2a^3 \right), \end{aligned}$$

which is exactly what we obtained in equation (4.3.9).

4.4. Identity for a more general case

We can use the same arguments as those in the proof of theorem 4.3.1 to get the following general theorem.

Theorem 4.4.1. *Let $\Theta(x, y) \in C^1(\mathbb{R}^2)$ and assume that*

$$\Theta(W_t, W_b - W_t),\ \Theta_x(W_t, W_b - W_t),\ \Theta_y(W_t, W_b - W_t) \in L^2([a, b] \times \Omega).$$

Then $\mathbb{E}\Big(\int_a^b \Theta(W_t, W_b - W_t)\,\mathrm{d}W_t\Big) = 0$, and

$$\begin{aligned} \mathbb{E}\left[\left(\int_a^b \Theta(W_t, W_b - W_t)\,\mathrm{d}W_t\right)^2\right] &= \int_a^b \mathbb{E}\big[\Theta(W_t, W_b - W_t)^2\big]\,\mathrm{d}t \\ &+ 2\int_a^b \int_a^t \mathbb{E}\Big[\Theta_y(W_s, W_b - W_s)\,\Theta_x(W_t, W_b - W_t)\Big]\,\mathrm{d}s\,\mathrm{d}t. \end{aligned} \tag{4.4.2}$$

We can extend this result to the product of two integrals.

Theorem 4.4.3. *Let $\Theta(x, y), \Lambda(x, y) \in C^1(\mathbb{R}^2)$ and assume that*

1. $\Theta(W_t, W_b - W_t),\ \Theta_x(W_t, W_b - W_t),\ \Theta_y(W_t, W_b - W_t) \in L^2([a, b] \times \Omega)$*, and*

2. $\Lambda(W_t, W_b - W_t)$, $\Lambda_x(W_t, W_b - W_t)$, $\Lambda_y(W_t, W_b - W_t) \in L^2([a,b] \times \Omega)$.

Then

$$\begin{aligned}
&\mathbb{E}\left[\left(\int_a^b \Theta(W_t, W_b - W_t)\,\mathrm{d}W_t\right)\left(\int_a^b \Lambda(W_t, W_b - W_t)\,\mathrm{d}W_t\right)\right] \\
&= \int_a^b \mathbb{E}[\Theta(W_t, W_b - W_t)\,\Lambda(W_t, W_b - W_t)]\,\mathrm{d}t \\
&\quad + \int_a^b \int_a^t \mathbb{E}\Big[\Theta_y(W_s, W_b - W_s)\,\Lambda_x(W_t, W_b - W_t) \\
&\qquad\qquad + \Theta_x(W_t, W_b - W_t)\,\Lambda_y(W_t, W_b - W_t)\Big]\,\mathrm{d}s\,\mathrm{d}t. \qquad (4.4.4)
\end{aligned}$$

Proof. For this proof, we write $F(t) = \Theta(W_t, W_b - W_t)$, $G(t) = \Lambda(W_t, W_b - W_t)$, and let $H = F + G$. Moreover, for brevity, we write $F_x(t) = \Theta_x(W_t, W_b - W_t)$, $F_y(t) = \Theta_y(W_t, W_b - W_t)$ and corresponding notation for G and H.

From the definition of H, we see that

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

Now, applying theorem 4.4.1 for F, we get

$$\mathbb{E}\left[\left(\int_a^b F(t)\,\mathrm{d}W_t\right)^2\right] = \int_a^b \mathbb{E}\big[F(t)^2\big]\,\mathrm{d}t + 2\int_a^b \int_a^t \mathbb{E}\big[F_y(s)\,F_x(t)\big]\,\mathrm{d}s\,\mathrm{d}t.$$

We can obtain an equivalent result for G. Putting all this together, we get

$$\begin{aligned}
\mathbb{E}\left[\left(\int_a^b H(t)\,\mathrm{d}W_t\right)^2\right] &= \int_a^b \mathbb{E}\big[F(t)^2\big]\,\mathrm{d}t + 2\int_a^b \int_a^t \mathbb{E}\big[F_y(s)\,F_x(t)\big]\,\mathrm{d}s\,\mathrm{d}t \\
&\quad + \int_a^b \mathbb{E}\big[G(t)^2\big]\,\mathrm{d}t + 2\int_a^b \int_a^t \mathbb{E}\big[G_y(s)\,G_x(t)\big]\,\mathrm{d}s\,\mathrm{d}t \\
&\quad + 2\,\mathbb{E}\left[\left(\int_a^b F(t)\,\mathrm{d}W_t\right)\left(\int_a^b G(t)\,\mathrm{d}W_t\right)\right]. \qquad (4.4.5)
\end{aligned}$$

On the other hand, first applying theorem 4.4.1 and then using the definition of H, we get

$$\begin{aligned}
&\mathbb{E}\left[\left(\int_a^b H(t)\,\mathrm{d}W_t\right)^2\right] \\
&= \int_a^b \mathbb{E}\big[H(t)^2\big]\,\mathrm{d}t + 2\int_a^b \int_a^t \mathbb{E}\big[H_y(s)\,H_x(t)\big]\,\mathrm{d}s\,\mathrm{d}t \\
&= \int_a^b \mathbb{E}\big[F(t)^2\big]\,\mathrm{d}t + \int_a^b \mathbb{E}\big[G(t)^2\big]\,\mathrm{d}t + 2\int_a^b \mathbb{E}[F(t)G(t)]\,\mathrm{d}t \\
&\quad + 2\int_a^b \int_a^t \mathbb{E}\big[(F_y(s)+G_y(s))(F_x(t)+G_x(t))\big]\,\mathrm{d}s\,\mathrm{d}t \\
&= \int_a^b \mathbb{E}\big[F(t)^2\big]\,\mathrm{d}t + \int_a^b \mathbb{E}\big[G(t)^2\big]\,\mathrm{d}t + 2\int_a^b \mathbb{E}[F(t)G(t)]\,\mathrm{d}t \\
&\quad + 2\int_a^b \int_a^t \mathbb{E}\Big[F_y(s)F_x(t) + F_y(s)G_x(t) + G_y(s)F_x(t) + G_y(s)G_x(t)\Big]\,\mathrm{d}s\,\mathrm{d}t. \qquad (4.4.6)
\end{aligned}$$

Finally, equation (4.4.5) and equation (4.4.6) imply that

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

which is exactly the desired result. □

If $\Theta(x,y) = f(x)$ and $\Lambda(x,y) = \phi(y)$, we have $\Theta_y \equiv 0$ and $\Lambda_x \equiv 0$. Therefore, we obtain the following corollary.

Corollary 4.4.7. *Let $f, \phi \in C^1(\mathbb{R})$ and assume that*

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

Then

$$\begin{aligned}
&\mathbb{E}\left[\left(\int_a^b f(W_t)\,\mathrm{d}W_t\right)\left(\int_a^b \phi(W_b - W_t)\,\mathrm{d}W_t\right)\right] \\
&= \int_a^b \mathbb{E}[f(W_t)\,\phi(W_b - W_t)]\,\mathrm{d}t + \int_a^b \int_a^t \mathbb{E}[\phi'(W_b - W_s)\,f'(W_t)]\,\mathrm{d}s\,\mathrm{d}t.
\end{aligned}$$

Similar to theorem 4.3.1, Corollary corollary 4.4.7 enables us to evaluate the covariance between anticipating and adapted integrals without explicitly calculating the integral itself. This is illustrated in the following example.

Example 4.4.8. Let $f(x) = x$ and $\phi(y) = y$. Using Corollary corollary 4.4.7, we get

$$\begin{aligned}
&\mathbb{E}\left[\left(\int_a^b W_t \,\mathrm{d}W_t\right)\left(\int_a^b (W_b - W_t)\,\mathrm{d}W_t\right)\right] \\
&= \int_a^b \mathbb{E}(W_t\,(W_b - W_t))\,\mathrm{d}t + \int_a^b \int_a^t \mathbb{E}(1)\,\mathrm{d}s\,\mathrm{d}t \\
&= \int_a^b \mathbb{E}(W_t)\,\mathbb{E}(W_b - W_t)\,\mathrm{d}t + \int_a^b (t-a)\,\mathrm{d}t \\
&= \frac{1}{2}(b-a)^2.
\end{aligned}$$

Finally, we want to point our that the double integral in equation (4.4.2) 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 4.4.9. Consider the case $\Theta(x, y) = px + y$ in theorem 4.4.1, where $p \in \mathbb{R}$. Then $\Theta_x = p$ and $\Theta_y = 1$. Therefore, we can directly evaluate the double integral in equation (4.4.2) as

$$2\int_a^b \int_a^t \mathbb{E}\big[\Theta_y(W_s, W_b - W_s)\,\Theta_x(W_t, W_b - W_t)\big]\,\mathrm{d}s\,\mathrm{d}t = 2\int_a^b \int_a^t p\,\mathrm{d}s\,\mathrm{d}t = p(b-a)^2.$$

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

Chapter 5. Near-martingales

5.1. Motivation

We now come to a central motivation behind the definition of the Ayed–Kuo integral — martingales. Recall that a martingale M is an integrable adapted process such that $\mathbb{E}(M_t \mid \mathcal{F}_s) = M_s$ almost surely for every $s \leq t$ (see section 1.2). Moreover, the process $M_t = \int_a^t X_s \,\mathrm{d}W_s$ generated by the Itô's integral of an adapted process X_s is a martingale. Can we say the same about processes associated with Ayed–Kuo integrals?

Example 5.1.1. Consider the process $N(t) = \int_a^t W_b \,\mathrm{d}W_t$. Using the linearity of the integral and example 3.1.4, we can write

$$\begin{aligned} N(t) &= \int_a^t W_b \,\mathrm{d}W_s \\ &= \int_a^b W_b \,\mathrm{d}W_s - \int_t^b W_b \,\mathrm{d}W_s \\ &= [W_b(W_b - W_a) - (b-a)] - [W_b(W_b - W_t) - (b-t)] \\ &= W_b(W_t - W_a) - (t-a). \end{aligned}$$

We write $W_b = (W_b - W_t) + (W_t - W_s) + W_s$ and $W_t - W_a = (W_t - W_s) + (W_s - W_a)$. Using the independence of increments of Wiener process, we get

$$\mathbb{E}(N(t) \mid \mathcal{F}_s) = W_s(W_s - W_a) - (s-a) \neq N(s).$$

So the process N is not a martingale. However, note that $\mathbb{E}(N(s) \mid \mathcal{F}_s) = W_s(W_s - W_a) - (s-a) = \mathbb{E}(N(t) \mid \mathcal{F}_s)$, or equivalently, $\mathbb{E}(N(t) - N(s) \mid \mathcal{F}_s) = 0$.

The above example motivates the following definition.

Definition 5.1.2 ([Hwa+17, definition 2.1])**.** An integrable stochastic process N is called a *near-martingale* if $\mathbb{E}(N(t) - N(s) \mid \mathcal{F}_s) = 0$ almost surely for every $s \leq t$. It is called a *near-submartingale* if we replace the second condition with $\mathbb{E}(N(t) - N(s) \mid \mathcal{F}_s) \geq 0$, and is called a *near-supermartingale* when $\mathbb{E}(N(t) - N(s) \mid \mathcal{F}_s) \leq 0$.

Clearly, a process that is both a near-submartingale and a near-supermartingale is a near-martingale. Moreover, any result that is true for a near-submartingale can be suitably modified for a near-supermartingale, and subsequently for a near-martingale. Therefore, in what follows, we only show results for submartingales.

For any near-submartingale N, we will called the process $M_t = \mathbb{E}(N(t) \mid \mathcal{F}_t)$ as the *conditioned process*.

Parts of this chapter previously appeared in the following open-access article: Hui-Hsiung Kuo, Pujan Shrestha, Sudip Sinha, and Padmanabhan Sundar. *On near-martingales and a class of anticipating linear SDEs*. 2022. arXiv: `2204.01932 [math.PR]`.

It is evident from the definition that every submartingale is a near-submartingale. More generally, a near-submartingale and its conditioned process are related by the following fundamental result.

Proposition 5.1.3 ([Hwa+17, theorem 2.5])**.** *A process N is a near-submartingale if and only if the conditioned process M given by $M_t = \mathbb{E}(N(t) \mid \mathcal{F}_t)$ is a submartingale.*

Proof. Note that the integrability condition is trivially satisfied. We prove the result for the *sub*-case. The other cases can be similarly derived. In what follows, we assume $s \le t$ and hence $\mathcal{F}_s \subseteq \mathcal{F}_t$.

For one direction, assume that N is a near-submartingale. Then

$$\mathbb{E}(M_t \mid \mathcal{F}_s) = \mathbb{E}(\mathbb{E}(N(t) \mid \mathcal{F}_t) \mid \mathcal{F}_s) = \mathbb{E}(N(t) \mid \mathcal{F}_s) \ge \mathbb{E}(N(s) \mid \mathcal{F}_s) = M_s,$$

so M is a submartingale.

For the other direction, assume M is a submartingale. Then

$$\mathbb{E}(N(t) \mid \mathcal{F}_s) = \mathbb{E}(\mathbb{E}(N(t) \mid \mathcal{F}_t) \mid \mathcal{F}_s) = \mathbb{E}(M_t \mid \mathcal{F}_s) \ge M_s = \mathbb{E}(N(s) \mid \mathcal{F}_s),$$

so N is a near-submartingale. □

5.2. Ayed–Kuo integrals are near-martingales

The following theorem shows that Ayed–Kuo integrals are near-martingales.

Theorem 5.2.1 ([Kuo+22, theorem 3.3])**.** *Suppose Θ is a real-valued measurable function on $\mathbb{R}^2$. Assume $L^2(\Omega)$-convergence for the Ayed–Kuo integral. Then $N(t) = \int_a^t \Theta(W_u, W_b - W_u)\,\mathrm{d}W_u$ is a near-martingale.*

Proof. The definition of the Ayed–Kuo integral under $L^2(\Omega)$-limits implies that for any $s \le t$,

$$\begin{aligned}
\mathbb{E}[N(t) - N(s) \mid \mathcal{F}_s] &= \mathbb{E}\left[\int_s^t \Theta(W_u, W_b - W_u)\,\mathrm{d}W_u \,\middle|\, \mathcal{F}_s\right] \\
&= \mathbb{E}\left[\lim_{n\to\infty} \sum_{i=1}^{\infty} \Theta(W_{t_{i-1}}, W_b - W_{t_i})\,\Delta W_i \,\middle|\, \mathcal{F}_s\right] \\
&= \lim_{n\to\infty} \sum_{i=1}^{\infty} \mathbb{E}\big[\Theta(W_{t_{i-1}}, W_b - W_{t_i})\,\Delta W_i \mid \mathcal{F}_s\big].
\end{aligned}$$

Using the tower property of conditional expectation (theorem 1.1.3) and the inclusion $\mathcal{H}_{t_{i-1}}^{t_i} \supseteq \mathcal{F}_s$ (see figure 5.1), we get

$$\mathbb{E}\big[\Theta(W_{t_{i-1}}, W_b - W_{t_i})\,\Delta W_i \mid \mathcal{F}_s\big] = \mathbb{E}\Big[\mathbb{E}\Big(\Theta(W_{t_{i-1}}, W_b - W_{t_i})\,\Delta W_i \,\Big|\, \mathcal{H}_{t_{i-1}}^{t_i}\Big) \,\Big|\, \mathcal{F}_s\Big].$$

Since $\Theta(W_{t_{i-1}}, W_b - W_{t_i})$ is $\mathcal{H}_{t_{i-1}}^{t_i}$-measurable and ΔW_i is independent to $\mathcal{H}_{t_{i-1}}^{t_i}$, we get

$$\mathbb{E}\big[\Theta(W_{t_{i-1}}, W_b - W_{t_i})\,\Delta W_i \mid \mathcal{F}_s\big] = \mathbb{E}\Big[\Theta(W_{t_{i-1}}, W_b - W_{t_i})\cancelto{0}{\mathbb{E}(\Delta W_i)} \,\Big|\, \mathcal{F}_s\Big] = 0.$$

By the continuity of limits, we get our desired result $\mathbb{E}[N(t) - N(s) \mid \mathcal{F}_s] = 0$. □

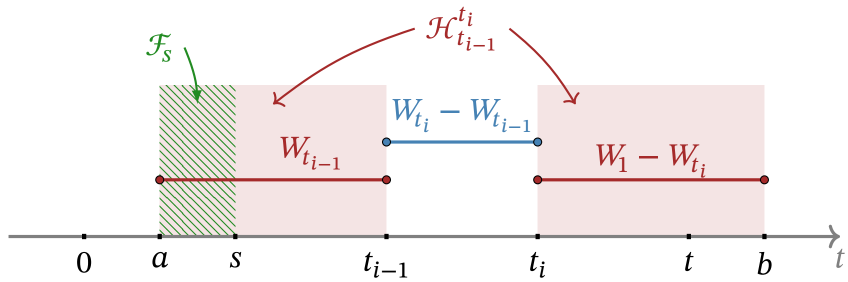

Figure 5.1. Inclusion of σ-algebras along with disjoint increments of W.

Remark 5.2.2. Note that the above proof also implies $\mathbb{E}\big[N(t) - N(s) \mid G^t\big] = 0$ for any $s \leq t$. This shows that Ayed–Kuo integrals are also *backward near-martingales*. Similarly, the process $\widetilde{N}(t) = \int_t^b \Theta(W_u, W_b - W_u)\,\mathrm{d}W_u$ is a near-martingale and a backward near-martingales by the same logic.

Let us look at an example of the above proposition.

Example 5.2.3 ([HKS19, example 2.7])**.** In example 3.1.6, we showed that

$$Z(t) = \int_a^t W_s(W_b - W_s)\,\mathrm{d}W_s = \frac{1}{2}W_b\left((W_t^2 - W_a^2) - (t-a)\right) - \frac{1}{3}\left(W_t^3 - W_a^3\right).$$

From theorem 5.2.1, we conclude that $Z(t)$ is a near-martingale. On the other hand, we can verify this using the conditional expectation $M_t = \mathbb{E}(Z(t) \mid \mathcal{F}_t)$ and proposition 5.1.3. Now,

$$\begin{aligned} M_t &= \mathbb{E}(Z(t) \mid \mathcal{F}_t) \\ &= \frac{1}{2}\,\mathbb{E}(W_b \mid \mathcal{F}_t)\left((W_t^2 - W_a^2) - (t-a)\right) - \frac{1}{3}\left(W_t^3 - W_a^3\right) \\ &= \frac{1}{2}W_t\left((W_t^2 - W_a^2) - (t-a)\right) - \frac{1}{3}\left(W_t^3 - W_a^3\right) \\ &= \frac{1}{6}\left((W_t^3 - 3tW_t) - 3(W_a^2 - a)W_t + 2W_a^2\right). \end{aligned}$$

We can easily check that W_t and $W_t^3 - 3tW_t$ are martingales. Therefore M is also a martingale, as expected.

The product of a martingale and an instantly-independent process is a near-martingale if and only if the instantly-independent process has constant mean.

Proposition 5.2.4 ([HKS19, theorem 2.9])**.** *Suppose M is a submartingale and Y an instantly-independent process such that both M_t and Y^t are square-integrable for each t. Then the process N given by $N(t) = M_tY^t$ is a near-submartingale if and only if $\mathbb{E}(Y^t)$ is a constant.*

Proof. Note that

$$\mathbb{E}(N(t) \mid \mathcal{F}_t) = \mathbb{E}\big(M_tY^t \mid \mathcal{F}_t\big) = M_t\,\mathbb{E}\big(Y^t \mid \mathcal{F}_t\big) = M_t\,\mathbb{E}\big(Y^t\big).$$

Therefore, $\mathbb{E}(N(t) \mid \mathcal{F}_t)$ is a submartingale if and only if $\mathbb{E}(Y^t)$ is a constant. By proposition 5.1.3, $N(t)$ is a near-submartingale if and only if $\mathbb{E}(Y^t)$ is a constant. □

5.3. Stopped near-martingales

In this section, we show that stopped near-martingales are near-martingales. Moreover, we give a optional stopping theorem for near-martingales on the lines of Doob's optional stopping theorem (theorem 1.2.6).

Definition 5.3.1. Let $(A_n)_{n=0}^{\infty}$ be an adapted process and $(X_n)_{n=0}^{\infty}$ a discrete time near-submartingale. Then the processes $(Y_n)_{n=0}^{\infty}$, where $Y_0 = 0$ and

$$Y_n = (A \bullet X)_n = \sum_{i=1}^{n} A_{n-1}(X_n - X_{n-1})$$

is called the *near-martingale transform* of X by A.

Near-martingale transforms retain the near-martingale property. This is a generalization proposition 1.2.5 to near-martingales.

Proposition 5.3.2.

1. *If X is a near-submartingale and A is a bounded non-negative adapted process, then $(A \bullet X)$ is a near-submartingale.*
2. *If X is a near-martingale and A is a bounded adapted process, then $(A \bullet X)$ is a near-martingale.*
3. *If X and A are both square integrable, then we do not require the boundedness condition in items 1 and 2.*

Proof. We only prove item 1 because the rest follow the same process. Let X be a near-submartingale and $Y = (A \bullet X)$. Suppose n is an arbitrary time. Note that $Y_n - Y_{n-1} = A_{n-1}(X_n - X_{n-1})$, which is integrable since A is bounded. Using the adaptedness of A, we get

$$\mathbb{E}(Y_n - Y_{n-1} \mid \mathcal{F}_{n-1}) = \mathbb{E}(A_{n-1}(X_n - X_{n-1}) \mid \mathcal{F}_{n-1}) = A_{n-1}\,\mathbb{E}(X_n - X_{n-1} \mid \mathcal{F}_{n-1}) \geq 0,$$

where the last inequality holds since A is non-negative. □

The following result says that stopped near-submartingales are near-submartingales.

Theorem 5.3.3. *Suppose X is a discrete time near-submartingale and τ a stopping time. Then the stopped process X^{τ} defined by $X_n^{\tau} = X_{\tau\wedge n}$ is a (discrete time) near-submartingale.*

Proof. Let $A_n = \mathbb{1}_{\{n\leq\tau\}}$. Clearly, the process A is bounded, non-negative, and adapted. Now, note that $X_n^{\tau} - X_0 = X_{\tau\wedge n} - X_0 = (A \bullet X)_n$. Therefore, by proposition 5.3.2, we get that X^{τ} is a near-submartingale. □

Now, we show the equivalent result of Doob's optional stopping theorem (theorem 1.2.6) for discrete time near-submartingales.

Theorem 5.3.4. *Let X be a discrete time near-submartingale. Suppose σ and τ are two bounded stopping times with $\sigma \leq \tau$. Then X_{σ} and X_{τ} are integrable, and $\mathbb{E}(X_{\tau} - X_{\sigma} \mid \mathcal{F}_{\sigma}) \geq 0$ almost surely.*

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

$$\int_{E\cap\{\sigma=n\}} (Y_N - Y_\sigma)\,\mathrm{d}\mathbb{P} = \int_{E\cap\{\sigma=n\}} (Y_N - Y_n)\,\mathrm{d}\mathbb{P} \geq 0.$$

Summing over n, we get $\int_E (Y_N - Y_\sigma)\,\mathrm{d}\mathbb{P} \geq 0$, and so $\mathbb{E}(Y_N - Y_\sigma \mid \mathcal{F}_\sigma) \geq 0$. Finally, let $Y_n = X_n^\tau$ to get

$$\mathbb{E}\big(X_N^\tau - X_\sigma^\tau \mid \mathcal{F}_\sigma\big) = \mathbb{E}(X_\tau - X_\sigma \mid \mathcal{F}_\sigma) \geq 0.$$

□

We need the following definition and lemma to prove the result in continuous time.

Definition 5.3.5. Let $(\mathcal{F}_n)_{n=1}^\infty$ be a decreasing sequence of σ-algebras, and let $X = (X_n)_{n=1}^\infty$ be a stochastic process. Then the pair $(X_n, \mathcal{F}_n)_{n=1}^\infty$ is called a *backward near-submartingale* if for every n,

1. X_n is integrable and $\mathcal{F}_n$-measurable, and
2. $\mathbb{E}(X_n - X_{n+1} \mid \mathcal{F}_{n+1}) \geq 0$.

Lemma 5.3.6. *Let $(X_n, \mathcal{F}_n)_{n=1}^\infty$ be a backward near-submartingale with $\mathbb{E}(X_n) > -\infty$. If X is non-negative for every n, then X is uniformly integrable.*

Proof. As $n \nearrow \infty$, we have $\mathbb{E}(X_n) \searrow \lim_{n\to\infty} \mathbb{E}(X_n) = \inf_n \mathbb{E}(X_n) > -\infty$. Fix $\epsilon > 0$. By the definition of infimum, there exists a $N > 0$ such that for any $n \geq N$, we have $\mathbb{E}(X_N) - \lim_{n\to\infty} \mathbb{E}(X_n) < \epsilon$.

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

$$\mathbb{E}\big(|X_k|\,\mathbb{1}_{\{|X_k|>\delta\}}\big) = \mathbb{E}\big(X_k \mathbb{1}_{\{X_k>\delta\}}\big) + \mathbb{E}\big(X_k \mathbb{1}_{\{X_k\geq-\delta\}}\big) - \mathbb{E}(X_k)\,.$$

Moreover, since X is a backward near-submartingale, $\mathbb{E}\big(X_k \mathbb{1}_{\{X_k\geq\delta\}}\big) \leq \mathbb{E}\big(X_n \mathbb{1}_{\{X_k\geq\delta\}}\big)$. Therefore,

$$\begin{aligned}\mathbb{E}\big(|X_k|\,\mathbb{1}_{\{|X_k|>\delta\}}\big) &\leq \mathbb{E}\big(X_n \mathbb{1}_{\{X_k>\delta\}}\big) + \mathbb{E}\big(X_n \mathbb{1}_{\{X_k\geq-\delta\}}\big) - (\mathbb{E}(X_n) - \epsilon) \\ &\leq \mathbb{E}\big(|X_n|\,\mathbb{1}_{\{|X_k|>\delta\}}\big) + \epsilon.\end{aligned}$$

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

$$\mathbb{P}\{|X_k| > \delta\} \leq \frac{1}{\delta}\,\mathbb{E}|X_k| = \frac{1}{\delta}\,\mathbb{E}(X_k) \leq \frac{1}{\delta}\,\mathbb{E}(X_0) \to 0$$

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

We are now ready to prove the near-martingale optional stopping theorem in continuous time.

Theorem 5.3.7. *Let N be a near-submartingale 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 $N(\sigma)$ and $N(\tau)$ are integrable, and*

$$\mathbb{E}(N(\tau) - N(\sigma) \mid \mathcal{F}_\sigma) \geq 0 \textit{ almost surely.}$$

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

$$f_n : [0, \infty) \to \left\{ \frac{k}{n} : k = 0, \dots, n \right\} : x \mapsto \frac{\lfloor 2^n x \rfloor + 1}{2^n} \wedge T, \tag{5.3.8}$$

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 the discrete time near-submartingale optional stopping theorem theorem 5.3.4, we get $N(\sigma_n)$ and $N(\tau_n)$ are integrable, and $\mathbb{E}\big(N(\tau_n) - N(\sigma_n) \bigm| \mathcal{F}_{\sigma_n}\big) \geq 0$ 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, $\mathbb{E}(N(\tau_n) - N(\sigma_n) \mid \mathcal{F}_\sigma) \geq 0$ almost surely.

If N is non-negative, by construction, $\big(N_{\sigma_n}, \mathcal{F}_{\sigma_n}\big)_{n=1}^\infty$ is a backward near-submartingale such that $N_{\sigma_n} \geq 0$ for every n. Therefore, $\mathbb{E}(N(\sigma_n)) \searrow \mathbb{E}(N(\sigma)) > -\infty$ as $n \nearrow \infty$. Using lemma 5.3.6, $(N(\sigma_n))_{n=1}^\infty$ 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 $N(\sigma)$ and $N(\tau)$ are integrable and that the convergence is also in L^1, giving us $\mathbb{E}(N(\tau) - N(\sigma) \mid \mathcal{F}_\sigma) \geq 0$ almost surely. □

We highlight the special case of theorem 5.3.7.

Corollary 5.3.9. *Let N be a non-negative near-martingale with right-continuous sample paths and τ a bounded stopping time. Then $N(\tau)$ is integrable, and $\mathbb{E}(N(\tau)) = \mathbb{E}(N(0))$ almost surely.*

Chapter 6. Conditionals of LSDEs

6.1. Motivation

Given an anticipating stochastic process $Z \in L^2([a, b] \times \Omega)$, its conditional expectation with respect to the σ-algebra $\mathcal{F}_t$ projects it to the space of adapted functions. That is, $\mathbb{E}(\cdot \mid \mathcal{F}_t) : L^2([a, b] \times \Omega) \to L^2_{\mathrm{ad}}([a, b] \times \Omega)$ is a projection operator. This is particularly interesting since it prunes a general stochastic process and brings it within the realm of Itô's theory. We call the process $X_t = \mathbb{E}(Z(t) \mid \mathcal{F}_t)$ as the *conditioned process of Z*.

Some questions arise immediately. Suppose Z is a solution of anticipating stochastic differential equation. Then is it true that the conditioned process X is the solution of adapted stochastic differential equation, where the coefficients and initial condition are also projected using the conditional expectation? What is the relationship between Z and X in general? We try to answer these questions in this chapter. We assume $t \in [a, b]$ throughout this chapter.

6.2. Conditional expectation of solutions of LSDEs

Theorem 6.2.1. *Assume $\alpha \in L^2_{ad}([a, b] \times \Omega)$, $\beta \in L^1_{ad}([a, b] \times \Omega)$ are stochastic processes. Suppose $h \in L^2[a, b]$ and ψ is an analytic function on the reals with derivative ψ'. Furthermore, let $Z(t)$ be the solution of the linear stochastic differential equations*

$$\begin{cases} \mathrm{d}Z(t) = \alpha_t \, Z(t) \, \mathrm{d}W_t + \beta_t \, Z(t) \, \mathrm{d}t \\ Z(a) = \psi\Big(\displaystyle\int_a^b h(s) \, \mathrm{d}W_s \Big), \end{cases} \tag{6.2.2}$$

and $X_t = \mathbb{E}(Z(t) \mid \mathcal{F}_t)$ is the conditioned process. Then X satisfies the linear stochastic differential equation

$$\begin{cases} \mathrm{d}X_t = \alpha_t \, X_t \, \mathrm{d}W_t + \beta_t \, X_t \, \mathrm{d}t + h(t) \, \widetilde{X}_t \, \mathrm{d}W_t \\ X_a = \mathbb{E}\Big[\psi\Big(\displaystyle\int_a^b h(s) \, \mathrm{d}W_s \Big) \Big], \end{cases} \tag{6.2.3}$$

where $\widetilde{X}_t = \mathbb{E}\big(\widetilde{Z}(t) \mid \mathcal{F}_t\big)$, and $\widetilde{Z}$ is the solution of the linear stochastic differential equation

$$\begin{cases} \mathrm{d}\widetilde{Z}(t) = \alpha_t \, \widetilde{Z}(t) \, \mathrm{d}W_t + \beta_t \, \widetilde{Z}(t) \, \mathrm{d}t \\ \widetilde{Z}(a) = \psi'\Big(\displaystyle\int_a^b h(s) \, \mathrm{d}W_s \Big). \end{cases} \tag{6.2.4}$$

Remark 6.2.5. In [KSZ18, theorem 4.1], we proved a similar result for the special case where α is deterministic, β is adapted, and $h \equiv 1$.

This chapter previously appeared in the following open-access journal article: Hui-Hsiung Kuo, Pujan Shrestha, and Sudip Sinha. "Anticipating Linear Stochastic Differential Equations with Adapted Coefficients". In: *Journal of Stochastic Analysis* 2.2 (2021). DOI: 10.31390/josa.2.2.05.

Proof. By the assumption and theorem 3.4.11, the solution processes $Z(t)$ can be written as

$$\begin{aligned} Z(t) &= \mathcal{E}_t^{(\alpha,\beta)} \cdot \psi\left(\Big(\int_a^t h(s)\,\mathrm{d}W_s - \int_a^t h(s)\alpha_s\,\mathrm{d}s\Big) + \int_t^b h(s)\,\mathrm{d}W_s\right) \\ &= \mathcal{E}_t^{(\alpha,\beta)} \cdot \sum_{k=0}^{\infty} \frac{1}{k!}\psi^{(k)}\Big(\int_a^t h(s)\,\mathrm{d}W_s - \int_a^t h(s)\alpha_s\,\mathrm{d}s\Big)\Big(\int_t^b h(s)\,\mathrm{d}W_s\Big)^k, \end{aligned}$$

where $\psi^{(k)}$ denotes the kth derivative of ψ.

For brevity, we henceforth denote

$$\psi_t^{(k)} = \psi^{(k)}\Big(\int_a^t h(s)\,\mathrm{d}W_s - \int_a^t h(s)\alpha_s\,\mathrm{d}s\Big). \tag{6.2.6}$$

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

$$Z(t) = \mathcal{E}_t^{(\alpha,\beta)} \cdot \sum_{k=0}^{\infty} \frac{1}{k!}\psi_t^{(k)} \cdot \Big(\int_t^b h(s)\,\mathrm{d}W_s\Big)^k.$$

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

$$V_t = \int_t^b h(s)^2\,\mathrm{d}s. \tag{6.2.7}$$

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

$$\mathbb{E}\left[\Big(\int_t^b h(s)\,\mathrm{d}W_s\Big)^{2k}\right] = V_t^k(2k-1)!!,$$

where !! denotes the double factorial defined as

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

for any natural number n.

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

$$\begin{aligned} X_t &= \mathcal{E}_t^{(\alpha,\beta)} \cdot \sum_{k=0}^{\infty} \frac{1}{(2k)!}\psi_t^{(2k)}\ \mathbb{E}\left[\Big(\int_t^b h(s)\,\mathrm{d}W_s\Big)^{2k}\right] \\ &= \mathcal{E}_t^{(\alpha,\beta)} \cdot \sum_{k=0}^{\infty} \frac{1}{(2k)!}\psi_t^{(2k)} V_t^k(2k-1)!! \\ &= \mathcal{E}_t^{(\alpha,\beta)} \cdot \sum_{k=0}^{\infty} \frac{1}{(2k)!!}\psi_t^{(2k)} V_t^k, \end{aligned} \tag{6.2.8}$$

and similarly,

$$\widetilde{X}_t = \mathcal{E}_t^{(\alpha,\beta)} \cdot \sum_{k=0}^{\infty} \frac{1}{(2k)!!} \psi_t^{(2k+1)} V_t^k, \tag{6.2.9}$$

Now we look at the differentials. From equation (6.2.6) and equation (6.2.7), we get

$$\mathrm{d}\big(V_t^k\big) = kV_t^{k-1}\,(-h(t)^2\,\mathrm{d}t),$$

and

$$\begin{aligned}
\mathrm{d}\Big(\psi_t^{(2k)}\Big) &= \psi_t^{(2k+1)} \cdot (h(t)\,\mathrm{d}W_t - h(t)\alpha_t\,\mathrm{d}t) + \frac{1}{2}\psi_t^{(2k+2)} \cdot \big(h(t)^2\,\mathrm{d}t\big) \\
&= \psi_t^{(2k+1)} h(t)\;\mathrm{d}W_t + \Big(\frac{1}{2}\psi_t^{(2k+2)} h(t)^2 - \psi_t^{(2k+1)} h(t)\alpha_t\Big)\,\mathrm{d}t.
\end{aligned}$$

Using the expressions for $\mathrm{d}\big(V_t^k\big)$ and $\mathrm{d}\psi_t^{(2k)}$, and remark 1.6.7, we get

$$\begin{aligned}
&\mathrm{d}\Big(\mathcal{E}_t^{(\alpha,\beta)}\psi_t^{(2k)}V_t^k\Big) \\
&= \psi_t^{(2k)}V_t^k\,\mathrm{d}\mathcal{E}_t^{(\alpha,\beta)} + \mathcal{E}_t^{(\alpha,\beta)}V_t^k\,\mathrm{d}\psi_t^{(2k)} + \mathcal{E}_t^{(\alpha,\beta)}\psi_t^{(2k)}(\mathrm{d}V_t)^k \\
&\quad + \mathcal{E}_t^{(\alpha,\beta)}\,\mathrm{d}\psi_t^{(2k)} \cdot (\mathrm{d}V_t)^k + \psi_t^{(2k)}\,\mathrm{d}\mathcal{E}_t^{(\alpha,\beta)} \cdot (\mathrm{d}V_t)^k + V_t^k\,\mathrm{d}\mathcal{E}_t^{(\alpha,\beta)} \cdot \mathrm{d}\psi_t^{(2k)} \\
&= \psi_t^{(2k)}V_t^k\left(\alpha_t\mathcal{E}_t^{(\alpha,\beta)}\,\mathrm{d}W_t + \beta_t\mathcal{E}_{\alpha,\beta}(t)\,\mathrm{d}t\right) \\
&\quad + \mathcal{E}_t^{(\alpha,\beta)}V_t^k\left(\psi_t^{(2k+1)}h(t)\,\mathrm{d}W_t + \Big(\frac{1}{2}\psi_t^{(2k+2)}h(t)^2 - \cancel{\psi_t^{(2k+1)}h(t)\alpha_t}\Big)dt\right) \\
&\quad + \mathcal{E}_t^{(\alpha,\beta)}\psi_t^{(2k)}\left(-kV_t^{k-1}h(t)^2\,\mathrm{d}t\right) \\
&\quad + 0 + 0 + V_t^k\left(\cancel{\mathcal{E}_t^{(\alpha,\beta)}\psi_t^{(2k+1)}\alpha_t h(t)}\right)\mathrm{d}t \\
&= \mathcal{E}_t^{(\alpha,\beta)}V_t^k\left(\psi_t^{(2k)}\alpha_t + \psi_t^{(2k+1)}h(t)\right)\mathrm{d}W_t \\
&\quad + \mathcal{E}_t^{(\alpha,\beta)}V_t^{k-1}\Big(\psi_t^{(2k)}V_t\beta_t + \frac{1}{2}\psi_t^{(2k+2)}V_t h(t)^2 - k\psi_t^{(2k)}h(t)^2\Big)\,\mathrm{d}t.
\end{aligned}$$

At this point, we note that

$$\begin{aligned}
\sum_{k=0}^{\infty}\frac{1}{(2k)!!}k\psi_t^{(2k)} &= \sum_{k=1}^{\infty}\frac{1}{(2k)(2k-2)!!}k\psi_t^{(2k)} \\
&= \frac{1}{2}\sum_{k-1=0}^{\infty}\frac{1}{(2(k-1))!!}\psi_t^{(2(k-1)+2)} \\
&= \frac{1}{2}\sum_{k=0}^{\infty}\frac{1}{(2k)!!}\psi_t^{(2k+2)}.
\end{aligned} \tag{6.2.10}$$

Now, since $X_t = \sum_{k=0}^{\infty} \frac{1}{(2k)!!} \mathcal{E}_t^{(\alpha,\beta)} \psi_t^{(2k)} V_t^k$ (see equation (6.2.8)), we get

$$
\begin{aligned}
\mathrm{d}X_t &= \sum_{k=0}^{\infty} \frac{1}{(2k)!!}\, \mathrm{d}\Big(\mathcal{E}_t^{(\alpha,\beta)} \psi_t^{(2k)} V_t^k\Big) \\
&= \sum_{k=0}^{\infty} \frac{1}{(2k)!!} \mathcal{E}_t^{(\alpha,\beta)} V_t^k \psi_t^{(2k)} \alpha_t \,\mathrm{d}W_t \\
&\quad + \sum_{k=0}^{\infty} \frac{1}{(2k)!!} \mathcal{E}_t^{(\alpha,\beta)} V_t^k \psi_t^{(2k+1)} h(t) \,\mathrm{d}W_t \\
&\quad + \sum_{k=0}^{\infty} \frac{1}{(2k)!!} \mathcal{E}_t^{(\alpha,\beta)} V_t^k \psi_t^{(2k)} \beta_t \,\mathrm{d}t \\
&\quad + \cancel{\sum_{k=0}^{\infty} \frac{1}{(2k)!!} \frac{1}{2} \mathcal{E}_t^{(\alpha,\beta)} V_t^k \psi_t^{(2k+2)} h(t)^2 \,\mathrm{d}t} \\
&\quad - \cancel{\sum_{k=0}^{\infty} \frac{1}{(2k)!!} k \mathcal{E}_t^{(\alpha,\beta)} V_t^{k-1} \psi_t^{(2k)} h(t)^2 \,\mathrm{d}t} \\
&= \alpha_t X_t \,\mathrm{d}W_t + h(t) \widetilde{X}_t \,\mathrm{d}W_t + \beta_t X_t \,\mathrm{d}W_t,
\end{aligned}
$$

where, in the second step, we used the result of equation (6.2.10). This completes the proof of the theorem. □

The presence of the extra term in the conditional stochastic differential equation in equation (6.2.3) poses an interesting question. Note that the stochastic differential equation for X_t is defined via $\widetilde{X}_t$. However, $\widetilde{X}_t$ is defined in equation (6.2.9) as an infinite series and a closed form is not guaranteed. It is important to note that $\widetilde{X}_t$ arose from taking the first derivative of ψ as the initial condition. Similarly, we can use the second derivative of ψ as the initial condition to arrive at the next step, ad infinitum. Therefore, a closed form solution is elusive in this method. Nevertheless, since we know that the derivative of the exponential function is itself, we have the following corollary.

Corollary 6.2.11. *Let α, β, h be as in theorem 6.2.1. Then*

$$
X_t = \mathcal{E}_t^{(\alpha,\beta)} \exp\Big(\int_a^t h(s) \,\mathrm{d}W_s - \int_a^t h(s)\, \alpha_s \,\mathrm{d}s \Big)
$$

satisfies the stochastic differential equation

$$
\begin{cases}
\mathrm{d}X_t = (\alpha_t + h(t))\, X_t \,\mathrm{d}W_t + \beta_t\, X_t \,\mathrm{d}t, \\
X_a = 1.
\end{cases}
$$

Proof. If $\psi(x) = e^x$, then $\psi \equiv \psi'$. Therefore $Z(t) \equiv \widetilde{Z}(t)$, and consequently $X_t = \widetilde{X}_t$. Then the result follows from theorem 6.2.1. □

6.3. Extending to a larger space

In general, the absence of a closed form for the associated conditional process does not pose significant problems. 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 6.2.1 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.

Recall that the Hermite polynomial of degree n with parameter ρ defined by

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

where D_x is the differentiation operator with respect to the variable x. From [Kuo96, page 334], we state the following identities:

$$\mathrm{D}_x H_n(x;\rho) = nH_{n-1}(x;\rho) \tag{6.3.1}$$

$$\mathrm{D}_\rho H_n(x;\rho) = -\frac{1}{2}\,\mathrm{D}_x^2\,H_n(x;\rho) \tag{6.3.2}$$

$$H_n(x+y;\rho) = \sum_{k=0}^{n}\binom{n}{k}H_{n-k}(x;\rho)y^k \tag{6.3.3}$$

We use these facts to prove the following lemma.

Lemma 6.3.4. *The stochastic process* $X_t = H_n\Big(\int_a^t h(s)\,\mathrm{d}W_s; \int_a^t h(s)^2\,\mathrm{d}s\Big)$ *with* $h \in L^2[a,b]$ *is a martingale with respect to the filtration generated by the Wiener process* W_t *and*

$$\mathrm{d}X_t = nH_{n-1}\Big(\int_a^t h(s)\,\mathrm{d}W_s; \int_a^t h(s)^2\,\mathrm{d}s\Big)h(t)\ \mathrm{d}W_t \tag{6.3.5}$$

Proof. Here $x = \int_a^t h(s)\,\mathrm{d}W_s$ and $\rho = \int_a^t h(s)^2\,\mathrm{d}s$. So we have $\mathrm{d}x = h(t)\,\mathrm{d}W_t$ and $\mathrm{d}\rho = h(t)^2\,\mathrm{d}t$, and $(\mathrm{d}x)^2 = \mathrm{d}\rho$. Using Itô's formula, we get

$$\begin{aligned}\mathrm{d}X_t &= \mathrm{D}_x H_n(x;\rho)\,\mathrm{d}x + \cancel{\frac{1}{2}\,\mathrm{D}_x^2\,H_n(x;\rho)(\mathrm{d}x)^2} + \cancel{\mathrm{D}_\rho H_n(x;\rho)\,\mathrm{d}\rho}\\ &= nH_{n-1}\Big(\int_a^t h(s)\,\mathrm{d}W_s; \int_a^t h(s)^2\,\mathrm{d}s\Big)h(t)\,\mathrm{d}W_t,\end{aligned}$$

where we used equation (6.3.2) for the cancellation and equation (6.3.1) to get the final term. □

This leads to the following result.

Theorem 6.3.6. *Assume $\alpha \in L^2_{ad}([a,b]\times\Omega)$, $\beta \in L^1_{ad}([a,b]\times\Omega)$ are stochastic processes and $h \in L^2[a,b]$ is a deterministic function. For a fixed $n \in \mathbb{N}$, suppose Z is the solution of the linear stochastic differential equation*

$$\begin{cases} \mathrm{d}Z(t) = \alpha_t Z(t)\,\mathrm{d}W_t + \beta_t Z(t)\,\mathrm{d}t, \\ Z(a) = H_n\Big(\int_a^b h(s)\,\mathrm{d}W_s; \int_a^b h(s)^2\,\mathrm{d}s\Big). \end{cases} \tag{6.3.7}$$

Then $X_t = \mathbb{E}(Z(t) \mid \mathcal{F}_t)$ is given by

$$X_t = H_n\Big(\int_a^t h(s)\,\mathrm{d}W_s - \int_a^t h(s)\,\alpha_s\,\mathrm{d}s; \int_a^t h(s)^2\,\mathrm{d}s\Big)\,\mathcal{E}_t^{(\alpha,\beta)}, \quad t \in [a,b]. \tag{6.3.8}$$

Moreover, X_t satisfies the following stochastic differential equation

$$\begin{cases} \mathrm{d}X_t = \alpha_t X_t\,\mathrm{d}W_t + \beta_t X_t\,\mathrm{d}t \\ \quad + nH_{n-1}\Big(\int_a^t h(s)\,\mathrm{d}W_s - \int_a^t h(s)\,\alpha_s\,\mathrm{d}s; \int_a^t h(s)^2\,\mathrm{d}s\Big)\,\mathcal{E}_t^{(\alpha,\beta)}\,h(t)\,\mathrm{d}W_t \\ X_a = 0. \end{cases} \tag{6.3.9}$$

Remark 6.3.10. For any x and ρ, we have $H_0(x;\rho) = 1$. Hence the stochastic differential equation (6.3.7) is identically the one in remark 1.6.7.

Proof. We first prove equation (6.3.8). Using theorem 3.4.11 and equation (6.3.3), we can write

$$\begin{aligned} Z(t) &= \mathcal{E}_t^{(\alpha,\beta)} H_n\left(\int_a^b h(s)\,\mathrm{d}W_s - \int_a^t h(s)\alpha_s\,\mathrm{d}s; \int_a^b h(s)^2\,\mathrm{d}s\right) \\ &= \mathcal{E}_t^{(\alpha,\beta)} \sum_{k=0}^n \binom{n}{k} H_{n-k}\left(\int_a^b h(s)\,\mathrm{d}W_s; \int_a^b h(s)^2\,\mathrm{d}s\right)\left(-\int_a^t h(s)\alpha_s\,\mathrm{d}s\right)^k \\ &= \mathcal{E}_t^{(\alpha,\beta)} \sum_{k=0}^n \binom{n}{k} J_{n-k}(b)\left(-\int_a^t h(s)\alpha_s\,\mathrm{d}s\right)^k, \end{aligned}$$

where we used the notation

$$J_n(t) = H_n\Big(\int_a^t h(s)\,\mathrm{d}W_s; \int_a^t h(s)^2\,\mathrm{d}s\Big).$$

Using lemma 6.3.4, we get $\mathbb{E}(J_{n-k}(b) \mid \mathcal{F}_t) = J_{n-k}(t)$. Taking the conditional expectation with the knowledge that $\mathcal{E}_t^{(\alpha,\beta)}$ is adapted and that stochastic integrals of adapted processes are

adapted,

$$
\begin{aligned}
X_t &= \mathcal{E}_t^{(\alpha,\beta)} \sum_{k=0}^{n} \binom{n}{k} \mathbb{E}(J_{n-k}(b) \mid \mathcal{F}_t) \left(- \int_a^t h(s)\alpha_s \, \mathrm{d}s \right)^k \\
&= \mathcal{E}_t^{(\alpha,\beta)} \sum_{k=0}^{n} \binom{n}{k} J_{n-k}(t) \left(- \int_a^t h(s)\alpha_s \, \mathrm{d}s \right)^k \\
&= \mathcal{E}_t^{(\alpha,\beta)} H_n \left(\int_a^t h(s) \, \mathrm{d}W_s - \int_a^t h(s)\alpha_s \, \mathrm{d}s; \int_a^t h(s)^2 \, \mathrm{d}s \right),
\end{aligned}
$$

which proves equation (6.3.8).

Since $H_n(0;0) = 0$, we see that $X_a = 0$. Using Itô's formula and equation (6.3.5),

$$
\begin{aligned}
\mathrm{d}H_n &= \mathrm{d}H_n \left(\int_a^t h(s) \, \mathrm{d}W_s - \int_a^t h(s)\alpha_s \, \mathrm{d}s; \int_a^t h(s)^2 \, \mathrm{d}s \right) \\
&= \mathrm{D}_x H_n \cdot (h(t) \, \mathrm{d}W_t - h(t)\alpha_t \, \mathrm{d}t) \\
&\qquad + \cancel{\frac{1}{2} \mathrm{D}_x{}^2 H_n \cdot (h(t)^2 \, \mathrm{d}t)} + \cancel{\mathrm{D}_\rho H_n \cdot (h(t)^2 \, \mathrm{d}t)} \\
&= n H_{n-1} \cdot h(t)(\mathrm{d}W_t - \alpha_t \, \mathrm{d}t).
\end{aligned}
$$

Finally, using equation (6.3.8), we get

$$
\begin{aligned}
\mathrm{d}X_t &= H_n \, \mathcal{E}_t^{(\alpha,\beta)} + \mathcal{E}_t^{(\alpha,\beta)} \, \mathrm{d}H_n + d\mathcal{E}_t^{(\alpha,\beta)} \cdot \mathrm{d}H_n \\
&= H_n \mathcal{E}_t^{(\alpha,\beta)} (\alpha_t \, \mathrm{d}W_t + \beta_t \, \mathrm{d}t) \\
&\quad + \mathcal{E}_t^{(\alpha,\beta)} n H_{n-1} \cdot h(t)(\mathrm{d}W_t - \cancel{\alpha_t} \, \mathrm{d}t) + \cancel{\mathcal{E}_t^{(\alpha,\beta)} \alpha_t n H_{n-1} h(t)} \, \mathrm{d}t. \\
&= \alpha_t X_t \, \mathrm{d}W_t + \beta_t X_t \, \mathrm{d}t + n H_{n-1} h(t) \mathcal{E}_t^{(\alpha,\beta)} \, \mathrm{d}W_t,
\end{aligned}
$$

which gives us equation (6.3.9). □

In equation (6.3.9), we specify an explicit form of the extra term in the stochastic differential equation for the conditioned process. We use this result in the following examples.

Example 6.3.11. Consider the stochastic differential equation

$$
\begin{cases}
\mathrm{d}Z(t) = W_t Z(t) \, \mathrm{d}W_t, \quad t \in [0,1], \\
\quad Z_0 = W_1.
\end{cases}
$$

Here $\alpha_t = W_t$, $\beta_t \equiv 0$, $h \equiv 1$, and $W_1 = H_1(W_1; 1)$. From theorem 6.3.6,

$$
X_t = \mathbb{E}(Z(t) \mid \mathcal{F}_t) = \left(W_t - \int_0^t W_s \, \mathrm{d}s \right) \exp\left[\frac{1}{2} \left(W_t{}^2 - t - \int_0^t W_s{}^2 \, \mathrm{d}s \right) \right]
$$

and X_t satisfies the following stochastic differential equation

$$
\begin{cases}
\mathrm{d}X_t = \left\{ W_t X_t + \exp\left[\frac{1}{2} \left(W_t{}^2 - t - \int_0^t W_s{}^2 \, \mathrm{d}s \right) \right] \right\} \mathrm{d}W_t, \quad t \in [0,1], \\
X_0 = 0.
\end{cases}
$$

Example 6.3.12. Consider the stochastic differential equation

$$\begin{cases} \mathrm{d}Z(t) = W_t Z(t)\,\mathrm{d}W_t, \quad t \in [0,1], \\ \quad Z_0 = W_1^2 - 1. \end{cases}$$

From theorem 6.3.6,

$$X_t = \left[\left(W_t - \int_0^t W_s\,\mathrm{d}s\right)^2 - t\right] \exp\left[\frac{1}{2}\left(W_t^2 - t - \int_0^t W_s^2\,\mathrm{d}s\right)\right],$$

and for $t \in [0,1]$, X_t satisfies the following stochastic differential equation.

$$\begin{cases} \mathrm{d}X_t = \left\{W_t X_t + 2\exp\left[\frac{1}{2}\left(W_t^2 - t - \int_0^t W_s^2\,\mathrm{d}s\right)\right]\right\}\mathrm{d}W_t, \\ X_0 = 0. \end{cases}$$

Chapter 7. Solving a Class of LSDEs

7.1. Motivation

In this chapter, we fix $t \in [0,1]$. Our goal is to find the solution of the linear stochastic differential equation

$$\begin{cases} \mathrm{d}Z(t) = f\left(\int_0^1 \gamma(s)\,\mathrm{d}W_s\right) Z(t)\,\mathrm{d}t + \sigma_t\, Z(t)\,\mathrm{d}W_t \\ Z(0) = \xi, \end{cases} \tag{7.1.1}$$

under reasonable conditions on γ, σ, f, and ξ. Note that the anticipation comes from the drift. We shall explore finding the solution using: (1) Ayed–Kuo theory by formulating an ansatz, and (2) Skorokhod integral and a novel braiding technique. We shall see that the Ayed–Kuo theory gives us a much simpler method to find the solution, albeit at the cost of requiring us to guess the form of the solution first. On the other hand, the braiding technique is constructive and is amenable to solving other types of equations. Both the methods yield identical solution under the same constraints.

7.2. The Ayed–Kuo solution

Theorem 7.2.1. *Suppose $\sigma \in L^2_{ad}([0,1] \times \Omega)$, $\gamma \in L^2[0,1]$, and ξ be a random variable independent of the Wiener process W. Moreover, suppose $f \in C^2(\mathbb{R})$ along with $f, f', f'' \in L^1(\mathbb{R})$. Then the solution of equation (7.1.1) in the Ayed–Kuo theory is given by*

$$Z(t) = \xi \exp\left[\int_0^t \sigma_s\,\mathrm{d}W_s - \frac{1}{2}\int_0^t \sigma_s^2\,\mathrm{d}s + \int_0^t f\left(\int_0^1 \gamma(u)\,\mathrm{d}W_u - \int_s^t \gamma(u)\,\sigma_u\,\mathrm{d}u\right)\mathrm{d}s\right]. \tag{7.2.2}$$

Proof. We show that equation (7.2.2) solves equation (7.1.1). The initial condition is trivially verified.

Note that equation (7.2.2) can be written as

$$\begin{aligned} Z(t) =\xi \exp\Bigg[&\int_0^t \sigma_s\,\mathrm{d}W_s - \frac{1}{2}\int_0^t \sigma_s^2\,\mathrm{d}s \\ &+ \int_0^t f\left(\int_0^t \gamma(u)\,\mathrm{d}W_u + \int_t^1 \gamma(u)\,\mathrm{d}W_u - \int_s^t \gamma(u)\,\sigma_u\,\mathrm{d}u\right)\mathrm{d}s\Bigg]. \end{aligned}$$

Motivated by this, we define

$$\theta(t, x_1, x_2, y) = \xi \exp\left[x_1 - \frac{1}{2}\int_0^t \sigma_s^2\,\mathrm{d}s + \int_0^t f\left(x_2 + y - \int_s^t \gamma(u)\,\sigma_u\,\mathrm{d}u\right)\mathrm{d}s\right].$$

This chapter previously appeared in the following open-access article: Hui-Hsiung Kuo, Pujan Shrestha, Sudip Sinha, and Padmanabhan Sundar. *On near-martingales and a class of anticipating linear SDEs.* 2022. arXiv: 2204.01932 [math.PR].

Moreover, let

$$\begin{aligned}
X_t^{(1)} &= \int_0^t \sigma_s \,\mathrm{d}W_s && (\text{so } \mathrm{d}X_t^{(1)} = \sigma_t \,\mathrm{d}W_t),\\
X_t^{(2)} &= \int_0^t \gamma(s) \,\mathrm{d}W_s && (\text{so } \mathrm{d}X_t^{(2)} = \gamma(t) \,\mathrm{d}W_t),\\
\text{and } Y^t &= \int_t^1 \gamma(s) \,\mathrm{d}W_s && (\text{so } \mathrm{d}Y^t = -\gamma(t) \,\mathrm{d}W_t).
\end{aligned}$$

Then we can write $Z(t) = \theta\left(t, X_t^{(1)}, X_t^{(2)}, Y^t\right)$.

For conciseness, we denote $F = f\left(\int_0^1 \gamma(t)\,\mathrm{d}W_t - \int_s^t \gamma(u)\,\sigma_u\,\mathrm{d}u\right)$, and similarly the derivatives $F' = f'\left(\int_0^1 \gamma(t)\,\mathrm{d}W_t - \int_s^t \gamma(u)\,\sigma_u\,\mathrm{d}u\right)$ and $F'' = f''\left(\int_0^1 \gamma(t)\,\mathrm{d}W_t - \int_s^t \gamma(u)\,\sigma_u\,\mathrm{d}u\right)$. Note that for the derivatives of θ, we have

$$\begin{aligned}
\theta_{x_1} &= \theta_{x_1x_1} = \theta,\\
\theta_{x_2} &= \theta_{x_1x_2} = \theta_y = \theta \cdot \int_0^t F'\,\mathrm{d}s,\\
\theta_{x_2x_2} &= \theta_{yy} = \theta \cdot \left(\int_0^t F'\,\mathrm{d}s\right)^2 + \theta \cdot \int_0^t F''\,\mathrm{d}s, \text{ and}\\
\theta_t &= -\frac{1}{2}\theta\sigma_t^2 + \theta f(x_2 + y) - \gamma(t)\sigma_t\theta_y,
\end{aligned}$$

where we used the Leibniz integral rule and the second line for the last identity.

Since ξ is independent of the Wiener process, by theorem 3.2.3, we get

$$\begin{aligned}
\mathrm{d}\theta =& \theta_t\,\mathrm{d}t + \theta_{x_1}\,\mathrm{d}X_t^{(1)} + \theta_{x_2}\,\mathrm{d}X_t^{(2)} + \theta_y\,\mathrm{d}Y^t\\
&+ \frac{1}{2}\theta_{x_1x_1}\left(\mathrm{d}X_t^{(1)}\right)^2 + \frac{1}{2}\theta_{x_2x_2}\left(\mathrm{d}X_t^{(2)}\right)^2 + \theta_{x_1x_2}\,\mathrm{d}X_t^{(1)}\,\mathrm{d}X_t^{(2)} - \frac{1}{2}\theta_{yy}(\mathrm{d}Y^t)^2.
\end{aligned}$$

Using the relationships between the derivatives of θ and its differential form, we get

$$\begin{aligned}
\mathrm{d}\theta =& \theta_t\,\mathrm{d}t + \theta\sigma_t\,\mathrm{d}W_t + \cancel{\theta_y\gamma(t)\,\mathrm{d}W_t} - \cancel{\theta\gamma(t)\,\mathrm{d}W_t}\\
&+ \frac{1}{2}\theta\sigma_t^2\,\mathrm{d}t + \cancel{\frac{1}{2}\theta_{yy}\gamma(t)^2\,\mathrm{d}t} + \theta_y\gamma(t)\sigma_t\,\mathrm{d}t - \cancel{\frac{1}{2}\theta_{yy}\gamma(t)^2\,\mathrm{d}t}\\
&\left(\theta_t + \frac{1}{2}\theta\sigma_t^2 + \theta_y\gamma(t)\sigma_t\right)\mathrm{d}t + \theta\sigma_t\,\mathrm{d}W_t.
\end{aligned}$$

Now,

$$\theta_t + \frac{1}{2}\theta\sigma_t^2 + \theta_y\gamma(t)\sigma_t = \left(-\cancel{\frac{1}{2}\theta\sigma_t^2} + \theta f(x_2 + y) - \cancel{\theta_y\gamma(t)\sigma_t}\right) + \cancel{\frac{1}{2}\theta\sigma_t^2} + \cancel{\theta_y\gamma(t)\sigma_t} = \theta f(x_2 + y),$$

and so

$$\mathrm{d}\theta = f(x_2 + y)\theta\,\mathrm{d}t + \sigma_t\theta\,\mathrm{d}W_t.$$

Since $Z(t) = \theta\left(t, X_t^{(1)}, X_t^{(2)}, Y^t\right)$, we get

$$\mathrm{d}Z(t) = f\left(\int_0^1 \gamma(s)\,\mathrm{d}W_s\right) Z(t)\,\mathrm{d}t + \sigma_t\, Z(t)\,\mathrm{d}W_t,$$

which is exactly equation (7.1.1). $\square$

The differential formula (Theorem 3.2.3) is an indispensable tool for analyzing anticipating processes. We show another example by finding the stochastic differential equation corresponding to the square of the above solution.

Theorem 7.2.3. *Under the condition of theorem 7.2.1, the stochastic differential equation*

$$\begin{cases} \mathrm{d}V(t) = \left[\sigma_t^2 + f\left(\int_0^1 \gamma(s)\,\mathrm{d}W_s\right) + 2\gamma(t)\,\sigma_t \int_0^t f'\left(\int_0^1 \gamma(u)\,\mathrm{d}W_u - \int_s^t \gamma(u)\,\sigma_u\,\mathrm{d}u\right)\mathrm{d}s\right] V(t)\,\mathrm{d}t \\ \qquad\qquad + 2\sigma_t\, V(t)\,\mathrm{d}W_t, \\ V(0) = \xi^2 \end{cases}$$

is solved by $Z(t)^2$, where Z is given by equation (7.2.2).

Remark 7.2.4. An interesting feature is that the derivative of f appears in the stochastic differential equation.

Proof. We follow the exact same strategy as the proof of theorem 7.2.1. The initial condition is trivially true. Let $V(t) = Z(t)^2$.

Taking the square of both sides of equation (7.2.2), we get

$$V(t) = \xi^2 \exp\left[\int_0^t 2\sigma_s\,\mathrm{d}W_s - \int_0^t \sigma_s^2\,\mathrm{d}s + \int_0^t 2f\left(\int_0^1 \gamma(u)\,\mathrm{d}W_u - \int_s^t \gamma(u)\,\sigma_u\,\mathrm{d}u\right)\mathrm{d}s\right]$$

We have $V(t) = \theta\left(t, X_t^{(1)}, X_t^{(2)}, Y^t\right)$, where

$$\theta(t, x_1, x_2, y) = \xi^2 \exp\left[x_1 - \int_0^t \sigma_s^2\,\mathrm{d}s + \int_0^t 2f\left(x_2 + y - \int_s^t \gamma(u)\,\sigma_u\,\mathrm{d}u\right)\mathrm{d}s\right],$$

and

$$\begin{aligned} X_t^{(1)} &= \int_0^t 2\sigma_s\,\mathrm{d}W_s && (\text{so } \mathrm{d}X_t^{(1)} = 2\sigma_t\,\mathrm{d}W_t), \\ X_t^{(2)} &= \int_0^t \gamma(s)\,\mathrm{d}W_s && (\text{so } \mathrm{d}X_t^{(2)} = \gamma(t)\,\mathrm{d}W_t), \\ \text{and } Y^t &= \int_t^1 \gamma(s)\,\mathrm{d}W_s && (\text{so } \mathrm{d}Y^t = -\gamma(t)\,\mathrm{d}W_t). \end{aligned}$$

As before, writing $F = f\left(\int_0^1 \gamma(t)\,\mathrm{d}W_t - \int_s^t \gamma(u)\,\sigma_u\,\mathrm{d}u\right)$, $F' = f'\left(\int_0^1 \gamma(t)\,\mathrm{d}W_t - \int_s^t \gamma(u)\,\sigma_u\,\mathrm{d}u\right)$, and $F'' = f''\left(\int_0^1 \gamma(t)\,\mathrm{d}W_t - \int_s^t \gamma(u)\,\sigma_u\,\mathrm{d}u\right)$, we get

$$\begin{aligned}
\theta_{x_1} &= \theta_{x_1x_1} = \theta, \\
\theta_{x_2} &= \theta_{x_1x_2} = \theta_y = 2\theta \cdot \int_0^t F'\,\mathrm{d}s, \\
\theta_{x_2x_2} &= \theta_{yy} = \theta \cdot \left(\int_0^t F'\,\mathrm{d}s\right)^2 + \theta \cdot \int_0^t F''\,\mathrm{d}s, \text{ and} \\
\theta_t &= -\theta\sigma_t^2 + 2\theta f(x_2 + y) - \gamma(t)\sigma_t\theta_y.
\end{aligned}$$

Using the above in theorem 3.2.3, we get

$$\begin{aligned}
\mathrm{d}\theta =& \theta_t\,\mathrm{d}t + \theta_{x_1}\,\mathrm{d}X_t^{(1)} + \theta_{x_2}\,\mathrm{d}X_t^{(2)} + \theta_y\,\mathrm{d}Y^t \\
&+ \frac{1}{2}\theta_{x_1x_1}\left(\mathrm{d}X_t^{(1)}\right)^2 + \frac{1}{2}\theta_{x_2x_2}\left(\mathrm{d}X_t^{(2)}\right)^2 + \theta_{x_1x_2}\,\mathrm{d}X_t^{(1)}\,\mathrm{d}X_t^{(2)} - \frac{1}{2}\theta_{yy}(\mathrm{d}Y^t)^2 \\
=& \theta_t\,\mathrm{d}t + 2\theta\sigma_t\,\mathrm{d}W_t + \cancel{\theta_{x_2}\gamma(t)\,\mathrm{d}W_t} - \cancel{\theta\gamma(t)\,\mathrm{d}W_t} \\
&+ 2\theta\sigma_t^2\,\mathrm{d}t + \cancel{\frac{1}{2}\theta_{x_2x_2}\gamma(t)^2\,\mathrm{d}t} + 2\theta_y\gamma(t)\sigma_t\,\mathrm{d}t - \cancel{\frac{1}{2}\theta_{yy}\gamma(t)^2\,\mathrm{d}t} \\
=& \left(\theta_t + 2\theta\sigma_t^2 + 2\theta_y\gamma(t)\sigma_t\right)\mathrm{d}t + 2\theta\sigma_t\,\mathrm{d}W_t \\
=& \left[\theta\sigma_t + 2\theta f(x_2 + y) + 2\gamma(t)\sigma_t\theta\int_0^t F'\,\mathrm{d}s\right]\mathrm{d}t + 2\theta\sigma_t\,\mathrm{d}W_t.
\end{aligned}$$

Finally, using $V(t) = \theta\left(t, X_t^{(1)}, X_t^{(2)}, Y^t\right)$, we get the stochastic differential equation. □

7.3. Solution in the Skorokhod sense

In the prior section, we showed the existence of the solution via the Ayed–Kuo differential formula. However, the procedure started with intelligently guessing an ansatz for the solution and applying the differential formula to it. Can a solution be found without this "guessing"? In this section, we use elementary ideas from Malliavin calculus to interpret the stochastic differential equation in the Skorokhod sense. Then we introduce an iterative "braiding" technique in the spirit of Trotter's product formula[Tro59] that allows us to construct the solution without needing to know the form of the solution. Note that we expect to arrive at the same solution as in section 7.2 since under the definition of the Ayed–Kuo integral using $L^2(\Omega)$ convergence, the Hitsuda–Skorokhod integral and the Ayed–Kuo integrals are equivalent, as shown in [Par17, theorem 2.3].

In this section, we use subscripts throughout for time. First, fix the family of translation $A_t : \mathcal{C}_0 \to \mathcal{C}_0$ in the Cameron–Martin direction given by

$$(A_t(\omega))_s = \omega_s - \int_0^{t\wedge s} \sigma(u)\,\mathrm{d}u \qquad \text{and} \qquad (T_t(\omega))_s = \omega_s + \int_0^{t\wedge s} \sigma(u)\,\mathrm{d}u.$$

We look at an existence result for stochastic differential equations in the Skorokhod sense.

Lemma 7.3.1. *Suppose $\sigma \in L^2[0,1]$ and $\xi \in L^p(\Omega)$ for some $p > 2$. Then the stochastic differential equation*

$$\begin{cases} \mathrm{d}Z(t) = \sigma(t)\, Z(t)\, \mathrm{d}W_t \\ Z(0) = \xi, \end{cases} \tag{7.3.2}$$

has the unique solution given by

$$Z(t) = (\xi \circ A_t)\, \mathcal{E}_t. \tag{7.3.3}$$

Proof. It is clear that the family $\{(\xi \circ A_t)\mathcal{E}_t \mid t \in [0,1]\}$ is $L^r(\Omega)$-bounded for all $r < p$ by Girsanov's theorem and Hölder 's inequality. Let G be any smooth random variable. Multiply both sides of equation (7.3.2) by G. With the process X given by equation (7.3.3),

$$\begin{aligned}
\mathbb{E}\left(G \int_0^t \sigma(s)\, Z(s)\, \mathrm{d}W_s\right) &= \mathbb{E}\left(\int_0^t \sigma(s)\, Z(s)\, \mathrm{D}_s G\, \mathrm{d}s\right) \\
&= \mathbb{E}\left(\xi \int_0^t \sigma(s)\, (\mathrm{D}_s G)(T_s)\, \mathrm{d}s\right) \qquad \text{(using Girsanov theorem)} \\
&= \mathbb{E}\left(\xi \int_0^t \frac{\mathrm{d}}{\mathrm{d}s} G(T_s)\, \mathrm{d}s\right) \\
&= \mathbb{E}(\xi(G(T_t) - G)) \\
&= \mathbb{E}(\xi(A_t)\, \mathcal{E}_t\, G) - \mathbb{E}(\xi\, G) \qquad \text{(again by Girsanov theorem)} \\
&= \mathbb{E}(Z(t)\, G) - \mathbb{E}(\xi\, G)\,.
\end{aligned}$$

Thus, a solution of the stochastic equation equation (7.3.2) is explicitly given by equation (7.3.3). Uniqueness follows since the solution of equation (7.3.2) started at $\xi \equiv 0$ is identically zero at all times. □

Now we introduce the braiding technique to generalize this to equations of the type equation (7.1.1), where $\gamma \in L^2[0,1]$ and $f : \mathbb{R} \to \mathbb{R}$. First, to simplify notation, define

$$\begin{aligned}
I_\gamma &= \int_0^1 \gamma_s\, \mathrm{d}W_s, \\
A_u^v(\omega_\cdot) &= \omega_\cdot - \int_u^{(\cdot\wedge v)\vee u} \sigma(s)\, \mathrm{d}s, \\
E_u^v &= \exp\left[\int_u^v \sigma(s)\, \mathrm{d}W_s - \frac{1}{2}\int_u^v \sigma(s)^2\, \mathrm{d}s\right], \text{ and} \\
g_u^v &= \exp\left[f(I_\gamma)\,(v-u)\right].
\end{aligned}$$

Directly from the definitions above, for any $u < v < w$, we get the compositions

$$\begin{aligned}
A_v^w \circ A_u^v &= A_u^w, \\
E_u^v \circ A_v^w &= E_u^v, \\
g_u^v \circ A_v^w &= \exp\left[f(I_\gamma \circ A_v^w)\,(v-u)\right],
\end{aligned}$$

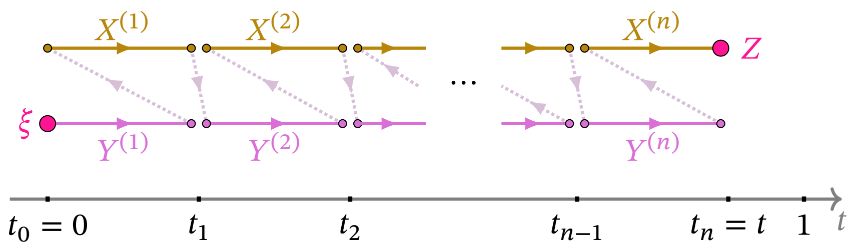

Figure 7.1. A graphical representation of the braiding technique.

and the products

$$\begin{aligned} E_u^v \cdot E_v^w &= E_u^w, \text{ and} \\ g_u^v \cdot g_v^w &= g_u^w. \end{aligned}$$

We suppress the dependence on ω for notational convenience.

Fix $t \in [0,1]$, and consider a sequence of partitions $\Pi_n = \{0 = t_0 < t_1 < \cdots < t_n = t\}$ of $[0,t]$ such that $\|\Pi_n\| = \sup\{t_i - t_{i-1} \mid i \in [n]\} \to 0$. On each subinterval, we

1. solve the equation having only the diffusion with the initial condition as the solution of the previous step, and
2. use the solution obtained in step 1 as the initial condition and solve the equation having only the drift.

For the first subinterval, the initial condition of step 1 is taken to be ξ. For a visual representation of the idea, see figure 7.1.

We explicitly demonstrate the process for the first two subintervals. An index (i) in the superscript refers to the ith subinterval.

First subinterval.

1. The stochastic differential equation that we want to solve is

$$\begin{cases} \mathrm{d}Y_u^{(1)} = \sigma(u) Y_u^{(1)} \, \mathrm{d}W_u, & u \in [0, t_1], \\ Y_0^{(1)} = \xi. \end{cases}$$

   Lemma 7.3.1 gave us the almost sure unique solution $Y_u^{(1)} = (\xi \circ A_0^u) \, E_0^u$, so

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

   on a set Ω_1, where $\mathbb{P}(\Omega_1) = 1$.
2. For each $\omega \in \Omega_1$, we want to solve the ordinary differential equation

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

By the existence and uniqueness theorem of ordinary differential equations, the unique solution is given by $X_u^{(1)} = Y_{t_1}^{(1)}\, g_0^u = (\xi \circ A_0^{t_1})\, E_0^{t_1}\, g_0^u$, and so

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

Second subinterval.

1. The stochastic differential equation that we want to solve is

$$\begin{cases} \mathrm{d}Y_u^{(2)} = \sigma(u)\, Y_u^{(2)}\, \mathrm{d}W_u, \quad u \in [t_1, t_2], \\ Y_{t_1}^{(2)} = X_{t_1}^{(1)}. \end{cases}$$

Lemma 7.3.1 gives us the almost sure unique solution $Y_u^{(2)} = (X_{t_1}^{(1)} \circ A_{t_1}^u)\, E_{t_1}^u$. Now,

$$\begin{aligned} Y_u^{(2)} &= \left[\left((\xi \circ A_0^{t_1})\, E_0^{t_1}\, g_0^{t_1}\right) \circ A_{t_1}^u\right]\, E_{t_1}^u \\ &= (\xi \circ A_0^{t_1} \circ A_{t_1}^u)\, E_0^{t_1}\, E_{t_1}^u\, (g_0^{t_1} \circ A_{t_1}^u) \\ &= (\xi \circ A_0^u)\, E_0^u\, (g_0^{t_1} \circ A_{t_1}^u), \end{aligned}$$

where we used the fact that $E_0^{t_1}$ is invariant under $A_{t_1}^u$. This is because, by definition,

$$A_{t_1}^u(\omega_\cdot) = \omega_\cdot - \int_{t_1}^{(\cdot \wedge u) \vee t_1} \sigma(s)\, \mathrm{d}s.$$

Now, for $E_0^{t_1}$, we have $t \in [0, t_1]$. Therefore,

$$A_{t_1}^u(\omega_t) = \omega_t - \int_{t_1}^{t_1} \sigma(s)\, \mathrm{d}s = \omega_t,$$

showing the invariance. This gives the motivation behind why we define A as such, and is a key trick in the method.
Continuing, we get

$$Y_{t_2}^{(2)} = (\xi \circ A_0^{t_2})\, E_0^{t_2}\, (g_0^{t_1} \circ A_{t_1}^{t_2})$$

on a set $\Omega_2 \subseteq \Omega_1$, where $\mathbb{P}(\Omega_2) = 1$.
2. For each $\omega \in \Omega_2$, we have the ordinary differential equation

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

The unique solution is given by $X_u^{(2)} = Y_{t_1}^{(2)}\, g_{t_1}^u$. Using the definition of $Y_{t_1}^{(2)}$ and the fact that $A_{t_2}^{t_2}$ is the identity function,

$$\begin{aligned} X_{t_2}^{(2)} &= \left[(\xi \circ A_0^{t_2})\, E_0^{t_2}\, (g_0^{t_1} \circ A_{t_1}^{t_2})\right]\, (g_{t_1}^{t_2} \circ A_{t_2}^{t_2}) \\ &= (\xi \circ A_0^{t_2})\, E_0^{t_2}\, \prod_{i=1}^{2} (g_{t_1}^{t_2} \circ A_{t_2}^{t_2}) \end{aligned}$$

It should now become obvious what the pattern is. We prove this using induction in the following lemma.

Lemma 7.3.4. *Let $\xi \in L^p(\Omega)$ for some $p > 2$. Consider the kth subinterval $u \in [t_{k-1}, t_k]$ for any $k \in [n]$, and define*

1. *the stochastic differential equation*

$$\begin{cases} \mathrm{d}Y_u^{(k)} = \sigma(u)\, Y_u^{(k)}\, \mathrm{d}W_u, \quad u \in [t_{k-1}, t_k], \\ Y_{t_{k-1}}^{(k)} = X_{t_{k-1}}^{(k-1)}, \text{ and} \end{cases}$$

2. *the ordinary differential equation*

$$\begin{cases} \mathrm{d}X_u^{(k)} = f(I_\gamma)\, X_u^{(k)}\, \mathrm{d}u, \quad u \in [t_{k-1}, t_k], \\ X_{t_{k-1}}^{(k)} = Y_{t_k}^{(k)}. \end{cases}$$

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

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

Proof. *Base cases.* This is true for $k = 1$ and $k = 2$ as shown in the computations above.

Induction step. Assume that the result holds for $k = m - 1$. This means that there exists Ω_{m-1} with $\mathbb{P}(\Omega_{m-1}) = 1$ such that on Ω_{m-1}, we have

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

Using the ideas of computations on the second subinterval, we get that there exists Ω_m with $\mathbb{P}(\Omega_m) = 1$ such that on Ω_m, we have

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

Since $A_{t_m}^{t_m}$ is the identity function, on Ω_m, we have

$$\begin{aligned} X_{t_m}^{(m)} &= Y_{t_{m-1}}^{(m)}\, g_{t_{m-1}}^{t_m} \\ &= (\xi \circ A_0^{t_m})\, E_0^{t_m} \prod_{i=1}^{m} (g_{t_{i-1}}^{t_i} \circ A_{t_i}^{t_m}). \end{aligned}$$

The proof is now complete by mathematical induction. □

We are now able to derive a closed form solution of equation (7.1.1) in the Skorokhod sense. This is the main theorem of the section.

Theorem 7.3.5. *Suppose $\sigma, \gamma \in L^2[0,1]$, $f : \mathbb{R} \to \mathbb{R}$, and $\xi \in L^p(\Omega)$ for some $p > 2$. Then the unique solution of equation* (7.1.1) *in the Skorokhod sense is given by*

$$\begin{aligned} Z(t) = (\xi \circ A_0^t) \exp\Bigg[\int_0^t \sigma(s) \, \mathrm{d}W_s - \frac{1}{2} \int_0^t \sigma(s)^2 \, \mathrm{d}s \\ + \int_0^t f\left(\int_0^1 \gamma(u) \, \mathrm{d}W_u - \int_s^t \gamma(u)\, \sigma(u) \, \mathrm{d}u \right) \mathrm{d}s \Bigg]. \end{aligned}$$

Remark 7.3.6. Note that ξ may depend on the Wiener process.

Proof. Using lemma 7.3.4, for any $t \in [0,1]$, we have

$$X_t^{(n)} = (\xi \circ A_0^t) \, E_0^t \, \prod_{i=1}^k (g_{t_{i-1}}^{t_i} \circ A_{t_i}^t).$$

Now,

$$\begin{aligned} \prod_{i=1}^k (g_{t_{i-1}}^{t_i} \circ A_{t_i}^t) &= \prod_{i=1}^k \exp\left[f(I_\gamma \circ A_{t_i}^t) \, (t_i - t_{i-1}) \right] \\ &= \exp\left[\sum_{i=1}^k f\left(\int_0^1 \gamma(u) \, \mathrm{d}W_u - \int_0^t \gamma(u) \, \sigma(u) \, \mathrm{d}u \right) \Delta t_i \right]. \end{aligned}$$

Finally, taking $n \to \infty$, we get

$$\begin{aligned} Z(t) &= \lim_{n \to \infty} X_t^{(n)} \\ &= (\xi \circ A_0^t) \, E_0^t \, \exp\left[\int_0^t f\left(\int_0^1 \gamma(u) \, \mathrm{d}W_u - \int_0^t \gamma(u) \, \sigma(u) \, \mathrm{d}u \right) \mathrm{d}s \right], \end{aligned}$$

which exactly equals the proposed solution.

The solution exists almost surely, due to the continuity of the measure. Moreover, the solution is unique. For if not, there are two solutions which disagree for the first time on 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 these uniqueness results. □

Chapter 8. Probabilistic Behavior of a Class of LSDEs

8.1. Rare events and asymptotic analysis

In mathematics, particularly in analysis, we are frequently concerned with asymptotic behavior of mathematical objects. Elementary probability theory gives us some simple tools to know the asymptotic behavior of means of random variables. First, the law of large numbers says that if we repeat the same experiment a large number of times, we should expect the empirical mean of the result to be close to the theoretical mean.

To formalize this, assume we have a sequence of independent and identically distributed (i.i.d.) integrable random variables $(X_n)_{n=1}^{\infty}$ with such that $\mathbb{E}(X_1) = m$. Let $S_n = \sum_{i=1}^{n} X_i$ be the sum and $\overline{X}_n = \frac{S_n}{n}$ be the empirical mean of n such random variables.

Theorem 8.1.1.

$$\overline{X}_n \to m \text{ as } n \to \infty \quad \text{almost surely or in } \mathbb{P}.$$

This is called the strong law of large numbers *if the convergence is almost surely. If the convergence is in probability, then the theorem is called the* weak law of large numbers.

Since almost sure convergence implies convergence in probability, the strong law implies the weak law. Even though it is called a "law" for historical reasons, it is a theorem of probability theory. Using notation from asymptotic analysis, we can write this as $\overline{X}_n = m + o(1)$.

The law of large numbers do not give us any knowledge of how far down the sequence we have to go in order for our empirical means to come "close" to the theoretical mean. The central limit theorem helps in this regard.

Theorem 8.1.2 (Lindeberg–Lévy)**.** *Assume the setting above along with the condition that* $\mathbb{V}(X_1) = \sigma^2 < \infty$. *Let* $Z \sim N(0,1)$ *be a standard normal variable. Then*

$$\sqrt{n}\left(\overline{X}_n - m\right) \to \sigma Z \quad \text{in distribution.}$$

Alternatively, in terms of the empirical sum S_n, *we can write this as*

$$\frac{S_n - nm}{\sqrt{n}} \to \sigma Z \quad \text{in distribution.}$$

Therefore, the central limit theorem gives us one more level of detail. Asymptotically, it says that

$$\overline{X}_n \simeq m + \frac{1}{\sqrt{n}}\sigma Z + o\left(\frac{1}{\sqrt{n}}\right).$$

Sections 8.4 and 8.5 previously appeared in the following open-access article: Hui-Hsiung Kuo, Pujan Shrestha, Sudip Sinha, and Padmanabhan Sundar. *On near-martingales and a class of anticipating linear SDEs*. 2022. arXiv: 2204.01932 [math.PR].

The tradeoff here is that we now have to contend ourselves with the weaker sense of convergence is in distribution.

Can we do better? Let us look at a specific example. Suppose that now all our random variables are i.i.d. standard normal, that is $X_1 \sim N(0,1)$. Then the moment generating function of X_1 is given by $M(\theta) = \mathbb{E}(e^{\theta X_1}) = e^{\frac{1}{2}\theta^2}$. We write the moment generating function of S_n as

$$M_{S_n}(\theta) = \mathbb{E}(e^{\theta S_n}) = \mathbb{E}\left(e^{\theta \sum_{i=1}^n X_i}\right) = \mathbb{E}\left(\prod_{i=1}^n e^{\theta X_i}\right) = \prod_{i=1}^n \mathbb{E}(e^{\theta X_i}),$$

where we used the independence assumption in the last step. Using the identical distribution assumption, we get

$$M_{S_n}(\theta) = \left(\mathbb{E}(e^{\theta X_1})\right)^n = \left(e^{\frac{1}{2}\theta^2}\right)^n = e^{\frac{1}{2}\theta^2 n} = e^{o(n)}.$$

On the other hand, using law of large numbers, we get $S_n = n\overline{X}_n = o(n)$, and so

$$M_{S_n}(\theta) = \mathbb{E}(e^{\theta S_n}) = \mathbb{E}(e^{o(n)}),$$

which does not equal $e^{o(n)}$ that we obtained using the moment generating function approach.

Why does the moment generating function give us better results than the law of large numbers? For one, the moment generating function actually takes into account the distribution of the random variables and does not use any approximation. However, there is another simple explanation for this. When we compute the expectations of random variable that can take very large values with very small probabilities, we cannot simply ignore the contributions of such values. This is because the product pq, where p is large and q is small depends on the magnitude of p and q, and can be quite large.

These kinds of problems are the mainstay of the insurance industry. Actuaries have to compute the expectation of net payout given that probabilities of accidents are quite small. Consider the following real-world problem. Let the i.i.d. random variables X_n denote the value of claims received by an insurance company on the nth month. The steady income for the company per month is the premium x. The company wants to use a planning period of n months to set the value of x so that they make a profit on average with a certain degree of certainty. Essentially, the goal is to determine x such that $\mathbb{P}(\overline{X}_n \geq x) < \epsilon$, where ϵ is an acceptable error probability. Therefore, it is hardly surprising that the theory of large deviations was started by actuarial mathematicians, prime among them being the Swedish mathematician Harald Cramér. However, it was Varadhan[Var66] who formally defined large deviations and unified the various results into a coherent theory, for which he was awarded the Abel Prize in 2007.

To get an idea of what Cramér showed, we use Markov's inequality to obtain exponential tail bounds of $\overline{X}_n$. Suppose X_1 has a finite moment generating function M. For an arbitrary $\theta > 0$,

using the i.i.d. nature of (X_n), we get

$$\begin{aligned}\mathbb{P}\Big\{\overline{X}_n \geq x\Big\} &= \mathbb{P}\Big\{e^{\theta n \overline{X}_n} \geq e^{\theta n x}\Big\} \\ &\leq e^{-\theta n x}\, \mathbb{E}\Big(e^{\theta n \overline{X}_n}\Big) \\ &= e^{-\theta n x} M_{\overline{X}_n}(\theta n) \\ &= \left(e^{-\theta x}\right)^n M(\theta)^n \\ &= e^{-n(\theta x - \log M(\theta))}.\end{aligned}$$

Since $\theta > 0$ is arbitrary, we have

$$\mathbb{P}\Big\{\overline{X}_n \geq x\Big\} \leq \inf_{\theta} e^{-n(\theta x - \log M(\theta))} = e^{-n \sup_{\theta}(\theta x - \log M(\theta))}.$$

Letting $\Lambda^*(x) := \sup_{\theta>0}\{\theta x - \log M(\theta)\}$, moving all terms dependent on n to the left, and taking $\overline{\lim}$, we get

$$\overline{\lim_{n\to\infty}} \frac{1}{n} \log \mathbb{P}\Big\{\overline{X}_n \geq x\Big\} \leq \Lambda^*(x).$$

Now, instead of the set $[x, \infty)$, if we consider an arbitrary closed set F, we can expect

$$\overline{\lim_{n\to\infty}} \frac{1}{n} \log \mathbb{P}\Big\{\overline{X}_n \in F\Big\} \leq \inf_{x\in F} \Lambda^*(x).$$

This is called a *large deviation upper bound*.

In fact, for this case, we can also obtain a *large deviation lower bound*. This means that for any arbitrary open set G, we can expect

$$\underline{\lim_{n\to\infty}} \frac{1}{n} \log \mathbb{P}\Big\{\overline{X}_n \in G\Big\} \geq \inf_{x\in G} \Lambda^*(x).$$

This is the statement of the famous Cramér's theorem.

Theorem 8.1.3 ([Cra38]). *The following hold.*

1. *(upper bound) For every closed set F,*

$$\overline{\lim_{n\to\infty}} \frac{1}{n} \log \mathbb{P}\Big\{\overline{X}_n \in F\Big\} \leq -\inf_{x\in F} \Lambda^*(x).$$

2. *(lower bound) For every open set G,*

$$\underline{\lim_{n\to\infty}} \frac{1}{n} \log \mathbb{P}\Big\{\overline{X}_n \in G\Big\} \geq -\inf_{x\in G} \Lambda^*(x).$$

Putting the two bounds together, we say that $\left(\overline{X}_n\right)$ follows a *large deviation principle* with *rate function* Λ^*, and we informally write

$$\mathbb{P}\Big\{\overline{X}_n \in \mathrm{d}x\Big\} \asymp e^{-n\Lambda^*(x)}.$$

In essence, the theory of large deviation allow us to find probabilities of rare events that decay exponentially.

Even though the theory of large deviations originated in the works of Scandinavian actuaries, it has become one of the central topics in probability theory, having found numerous applications in the seemingly disparate fields of dynamical systems, statistical mechanics, information theory, and fluid mechanics, among others.

8.2. Large deviation principles

We now formally define large deviation principles in a general setting. Recall that a Polish space is any space that is homeomorphic to a complete metric space having a countable dense subset.

Definition 8.2.1. Let $(\mathcal{X}, d)$ be a Polish space and $(\mu^\epsilon)_{\epsilon>0}$ a sequence of Borel probability measures on $\mathcal{X}$. Suppose $I : \mathcal{X} \to [0, \infty]$ is a lower semicontinuous functional. Then the sequence $(\mu^\epsilon)_{\epsilon>0}$ is said to satisfy a *large deviation principle* (LDP) on $\mathcal{X}$ with *rate function I* if and only if for each Borel measurable $E \subseteq \mathcal{X}$,

$$-\inf_{x\in E^\circ} I(x) \leq \epsilon \underline{\lim}_{\epsilon\to 0} \log \mu^\epsilon(E) \leq \epsilon \overline{\lim}_{\epsilon\to 0} \log \mu^\epsilon(E) \leq -\inf_{x\in \overline{E}} I(x),$$

where E° and $\overline{E}$ are the interior and closure of E, respectively.

Equivalently, the sequence $(\mu^\epsilon)_{\epsilon>0}$ is said to satisfy a large deviation principle with rate function I if and only if

1. (upper bound) for every closed set $F \subseteq \mathcal{X}$,

$$\overline{\lim}_{\epsilon\to 0} \epsilon \log \mu^\epsilon(F) \leq -\inf_{x\in F} I(x),$$

2. (lower bound) and for every open set $G \subseteq \mathcal{X}$,

$$\underline{\lim}_{\epsilon\to 0} \epsilon \log \mu^\epsilon(G) \geq -\inf_{x\in G} I(x).$$

Recall that we used the second version to write Cramér's theorem in the previous section.

The next result states how LDPs are transferred under continuous transformations.

Theorem 8.2.2 (contraction/continuity principle [DZ98, theorem 4.2.1])**.** *Let $\mathcal{X}$ and $\mathcal{Y}$ be Hausdorff spaces and $f : \mathcal{X} \to \mathcal{Y}$ a continuous function. Suppose the sequence of probability measures $(\mu^\epsilon)_{\epsilon>0}$ on $\mathcal{X}$ satisfy a large deviation principle with rate function $I : \mathcal{X} \to [0, \infty]$. For each $y \in \mathcal{Y}$, define $J(y) = \inf\{(I \circ f^{-1})(y)\}$. Then the sequence of pushforward probability measures $\left(\mu^\epsilon \circ f^{-1}\right)_{\epsilon>0}$ on $\mathcal{Y}$ satisfy a large deviation principle with rate function $J : \mathcal{Y} \to [0, \infty]$.*

Are there situations where a large deviation principle is preserved? To answer this, we define exponential equivalence of measures.

Definition 8.2.3 ([DZ98, definition 4.2.10]). Let $(\mathcal{X}, d)$ be a Polish space with two families of measures $(\mu^\epsilon)_{\epsilon>0}$ and $(\nu^\epsilon)_{\epsilon>0}$ on it. Suppose there exists probability spaces $(\Omega, \Sigma_\epsilon, \mathbb{P}_\epsilon)$ and two families of $\mathcal{X}$-valued random variables $(X^\epsilon)_{\epsilon>0}$ and $(Y^\epsilon)_{\epsilon>0}$ with joint laws $(\mathbb{P}_\epsilon)_{\epsilon>0}$ and marginals $(\mu^\epsilon)_{\epsilon>0}$ and $(\nu^\epsilon)_{\epsilon>0}$, respectively. Moreover, for all $\delta > 0$, the set $\{\omega \mid (X^\epsilon, Y^\epsilon) \in E_\delta\}$ is Σ_ϵ-measurable, and $\overline{\lim}_{\epsilon\to 0} \epsilon \log \mathbb{P}_\epsilon(E_\delta) = -\infty$, where $E_\delta = \{(x, y) \mid d(x, y) > \delta\} \subseteq \mathcal{X}^2$. Then the families $(\mu^\epsilon)_{\epsilon>0}$ and $(\nu^\epsilon)_{\epsilon>0}$ (and correspondingly $(X^\epsilon)_{\epsilon>0}$ and $(Y^\epsilon)_{\epsilon>0}$) are called *exponentially equivalent*.

Exponentially equivalent families of measures induce the same large deviation principle. This is the content of the following theorem.

Theorem 8.2.4 ([DZ98, theorem 4.2.13]). *Suppose the families $(\mu^\epsilon)_{\epsilon>0}$ and $(\nu^\epsilon)_{\epsilon>0}$ of measures on $\mathcal{X}$ are exponentially equivalent. Then $(\mu^\epsilon)_{\epsilon>0}$ satisfies a large deviation principle with rate function I if and only if $(\nu^\epsilon)_{\epsilon>0}$ satisfies a large deviation principle with the same rate function I.*

8.3. Sample path large deviations

Just as in the previous chapter, we fix $t \in [0, 1]$. We state everything in one-dimension, even though all of the results of the section can be replaced by their higher-dimensional equivalents.

Recall that $\left(\mathcal{C}_0, \|\cdot\|_\infty\right)$ is the space of real-valued continuous functions on $[0, 1]$ and $\mathcal{H}^1$ the Cameron–Martin space. Let the translation of the set $\mathcal{C}_0$ by $\kappa \in \mathbb{R}$ be denoted using $\mathcal{C}_\kappa$. Similarly, $\mathcal{H}^1_\kappa$ denotes the linear translation of $\mathcal{H}^1$ by $\kappa \in \mathbb{R}$.

Consider the family of process $\left(\sqrt{\epsilon} W\right)_{\epsilon>0}$, where a Wiener process W is scaled down by a parameter $\sqrt{\epsilon}$. Since for every t, the random variable $W_t \sim N(0, t)$, and so $\sqrt{\epsilon} W_t \sim N(0, \epsilon t)$. Therefore, as $\epsilon \to 0$, the sequence $\sqrt{\epsilon} W_t \to 0$ in probability (also almost surely). Can we determine the rate of this convergence? In other words, can we estimate the probability that a scaled-down sample path of a Brownian motion will stray far from the mean path? This is answered by Schilder's theorem.

Theorem 8.3.1 ([Sch66]). *For every ϵ, let μ^ϵ be the law of $\sqrt{\epsilon} W$ on $\left(\mathcal{C}_0, \|\cdot\|_\infty\right)$. Then $(\mu^\epsilon)_{\epsilon>0}$ satisfies a large deviation principle with rate function*

$$I(\omega) = \begin{cases} \frac{1}{2} \int_0^1 |\omega'(t)|^2 \, \mathrm{d}t & \text{if } \omega \in \mathcal{H}^1, \\ \infty & \text{otherwise.} \end{cases}$$

Example 8.3.2. Define the *reverse Wiener process* as $V^t = W_1 - W_t$. For every ϵ, let ν^ϵ be the law of $\sqrt{\epsilon} V$ on the space of continuous functions $f : [0, 1] \to \mathbb{R}$ such that $f(1) = 0$ embedded with the supremum norm. Then $(\nu^\epsilon)_{\epsilon>0}$ satisfies a large deviation principle with rate function

$$J(\phi) = \begin{cases} \frac{1}{2} \int_0^1 |f(t)|^2 \, \mathrm{d}t & \text{if } \phi(t) = \int_0^{1-t} f(s) \, \mathrm{d}s \text{ for some } f \in L^2[0, 1], \\ \infty & \text{otherwise.} \end{cases}$$

Proof. Define $Z_t = W_1 - W_{1-t}$. Then we have the following properties:

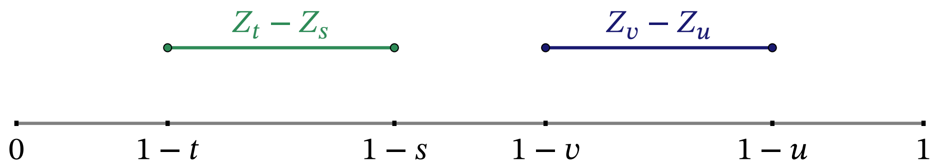

Figure 8.1. The disjoint increments of Z.

1. $Z_0 = W_1 - W_1 = 0$ almost surely.
2. For $0 \leq u \leq v \leq s \leq t \leq 1$, the increment $Z_t - Z_s = W_{1-s} - W_{1-t}$ is independent of $Z_v - Z_u = W_{1-u} - W_{1-v}$ due to the independence of increments of W (see figure 8.1).
3. $Z_t - Z_s = W_{1-s} - W_{1-t} \sim N(0, t-s)$ for any $0 \leq s \leq t \leq 1$.
4. The paths of Z are continuous almost surely since they are just a sum of continuous functions.

From definition 1.1.1, we see that Z is a Brownian motion. Therefore, theorem 8.3.1 applies for the family of processes $\left(\sqrt{\epsilon}Z\right)_{\epsilon>0}$.

Consider the function $\theta : \mathcal{C} \to \mathcal{C} : \phi(t) \mapsto \phi(1-t)$, where $\mathcal{C}$ is the space of continuous functions. Note that θ is an continuous involution on $\mathcal{C}$. Moreover,

$$\theta(\sqrt{\epsilon}Z_t) = \theta(\sqrt{\epsilon}(W_1 - W_{1-t})) = \sqrt{\epsilon}(W_1 - W_t) = \sqrt{\epsilon}V^t \sim \nu^\epsilon.$$

Using the continuity principle (theorem 8.2.2) along with the rate function I from theorem 8.3.1, $(\nu^\epsilon)_{\epsilon>0}$ follows a large deviation principle with the rate function given by

$$\begin{aligned}
J(\phi) &= I \circ \theta^{-1}(\phi(\cdot)) = I(\phi(1-\cdot)) \\
&= \begin{cases} \frac{1}{2}\int_0^1 |f(t)|^2 \,\mathrm{d}t & \text{if } \phi(1-t) = \int_0^t f(s)\,\mathrm{d}s \text{ for some } f \in L^2[0,1] \\ \infty & \text{otherwise} \end{cases} \\
&= \begin{cases} \frac{1}{2}\int_0^1 |f(t)|^2 \,\mathrm{d}t & \text{if } \phi(t) = \int_0^{1-t} f(s)\,\mathrm{d}s \text{ for some } f \in L^2[0,1] \\ \infty & \text{otherwise.} \end{cases}
\end{aligned}$$

□

The extension of Schilder's theorem to Itô diffusions is the content of the *Freidlin–Wentzell theory*. We start with a simple situation where we perturb an ordinary differential equation with a diffusion that goes to zero as $\epsilon \to 0$. Let X^ϵ denote the unique solution of the stochastic differential equation

$$\begin{cases} \mathrm{d}X_t^\epsilon = b(X_t^\epsilon)\,\mathrm{d}t + \sqrt{\epsilon}\,\mathrm{d}W_t, \\ \ \ X_0^\epsilon = 0, \end{cases}$$

where $b : \mathbb{R} \to \mathbb{R}$ is a uniformly Lipschitz continuous function. Let $(\nu^\epsilon)_{\epsilon>0}$ be the sequence of probability measures induced by $(X^\epsilon)_{\epsilon>0}$ on $\left(\mathcal{C}_0, \|\cdot\|_\infty\right)$.

Theorem 8.3.3 ([DZ98, theorem 5.6.3])**.** $(\nu^\epsilon)_{\epsilon>0}$ *satisfies a large deviation principle with rate function*

$$I(f) = \begin{cases} \frac{1}{2}\int_0^1 |f'(t) - b(f(t))|^2 \,\mathrm{d}t & \textit{if } f \in \mathcal{H}^1, \\ \infty & \textit{otherwise.} \end{cases}$$

The idea of the proof again rests on the continuity principle given in theorem 8.2.2. Let us elaborate further. Suppose $F : \mathcal{C}_0 \to \mathcal{C}_0$ be the deterministic map determined by $f = F(g)$, where f is the unique continuous solution of

$$f(t) = \int_0^t b(f(s))\,\mathrm{d}s + g(t).$$

Then $\nu^\epsilon = \mu^\epsilon \circ F^{-1}$, where μ^ϵ is the measure induced by $\sqrt{\epsilon}W$ on $\left(\mathcal{C}_0, \|\cdot\|_\infty\right)$. Now we can apply the continuity principle to arrive at the required large deviation principle result.

Finally, we look at the general case. Suppose $\kappa \in \mathbb{R}$, the functions $b : \mathbb{R} \to \mathbb{R}$ and $\sigma : \mathbb{R} \to \mathbb{R}$ are bounded uniformly Lipschitz continuous functions, and σ is uniformly bounded away from 0. Let X^ϵ be the diffusion process that is the unique solution of the stochastic differential equation

$$\begin{cases} \mathrm{d}Y_t^\epsilon = b(Y_t^\epsilon)\,\mathrm{d}t + \sqrt{\epsilon}\sigma(Y_t^\epsilon)\,\mathrm{d}W_t, \\ \quad Y_0^\epsilon = \kappa. \end{cases}$$

Theorem 8.3.4 (Freidlin–Wentzell [DZ98, theorem 5.6.7])**.** *The sequence of probability measures induced by* $(Y^\epsilon)_{\epsilon>0}$ *on* $\left(\mathcal{C}_0, \|\cdot\|_\infty\right)$ *satisfies a large deviation principle with rate function*

$$I_x(f) = \begin{cases} \frac{1}{2}\int_0^1 \left|\frac{1}{\sigma(f(t))}\left(f'(t) - b(f(t))\right)\right|^2 \mathrm{d}t & \textit{if } f \in \mathcal{H}_x^1, \\ \infty & \textit{otherwise.} \end{cases}$$

8.4. LSDEs with anticipating coefficients and constant initial condition

In what follows, we develop a Freidlin–Wentzell type result for linear stochastic differential equations with anticipating coefficients. In particular, we look at a class of linear stochastic differential equations of the type equation (7.1.1) that we analyzed in chapter 7. We start off with the case of constant initial conditions.

Suppose σ and γ are deterministic functions of bounded variation on $[0, 1]$. Moreover, suppose $f \in C^2(\mathbb{R})$ is Lipschitz continuous along with $f, f', f'' \in L^1(\mathbb{R})$. For a fixed $\kappa \in \mathbb{R}$, consider the family of linear stochastic differential equations with parameter $\epsilon > 0$ given by

$$\begin{cases} \mathrm{d}Z_\kappa^\epsilon(t) = f\left(\sqrt{\epsilon}\displaystyle\int_0^1 \gamma(s)\,\mathrm{d}W_s\right) Z_\kappa^\epsilon(t)\,\mathrm{d}t + \sqrt{\epsilon}\sigma(t)\, Z_\kappa^\epsilon(t)\,\mathrm{d}W_t \\ \\ Z_\kappa^\epsilon(0) = \kappa, \end{cases} \tag{8.4.1}$$

Using the results from chapter 7, the unique solutions to equation (8.4.1) are given by

$$Z_\kappa^\epsilon(t) = \kappa \exp\left[\sqrt{\epsilon}\int_0^t \sigma(s)\,\mathrm{d}W_s - \frac{\epsilon}{2}\int_0^t \sigma(s)^2\,\mathrm{d}s + \int_0^t f\left(\sqrt{\epsilon}\int_0^1 \gamma(u)\,\mathrm{d}W_u - \epsilon\int_s^t \gamma(u)\,\sigma(u)\,\mathrm{d}u\right)\mathrm{d}s\right]. \tag{8.4.2}$$

Our approach will be similar to how the Freidlin–Wentzell results in theorem 8.3.3 are derived from Schilder's theorem (theorem 8.3.1) using the continuity principle (theorem 8.2.2). In order to use the continuity principle, we need the following lemma.

Lemma 8.4.3. *The function $\theta : \mathcal{C}_0 \to \mathcal{C}_\kappa$ defined by*

$$\theta(x) = \kappa \exp\left[\int_0^t \sigma(s)\,\mathrm{d}x(s) - \frac{\epsilon}{2}\int_0^t \sigma(s)^2\,\mathrm{d}s + \int_0^t f\left(\int_0^1 \gamma(u)\,\mathrm{d}x(u) - \epsilon\int_s^t \gamma(u)\,\sigma(u)\,\mathrm{d}u\right)\mathrm{d}s\right]$$

is continuous in the topology induced by the canonical supremum norm.

Proof. We can write

$$\theta(x) = \kappa \exp\left[\phi(x) - \frac{\epsilon}{2}\int_0^t \sigma_s^2\,\mathrm{d}s + \psi(x)\right],$$

where $\phi, \psi : \mathcal{C}_0 \to \mathcal{C}_0$ is given by

$$\begin{aligned}\phi(x) &= \int_0^t \sigma(s)\,\mathrm{d}x(s) = \sigma(t)x(t) - \int_0^t x(s)\,\mathrm{d}\sigma(s), \text{ and}\\ \psi(x) &= \int_0^t f\left(\int_0^1 \gamma(u)\,\mathrm{d}x(u) - \epsilon\int_s^t \gamma(u)\,\sigma(u)\,\mathrm{d}u\right)\mathrm{d}s.\end{aligned}$$

Using integration by parts,

$$\begin{aligned}\phi(x) &= \sigma(t)x(t) - \int_0^t x(s)\,\mathrm{d}\sigma(s), \text{ and}\\ \psi(x) &= \int_0^t f\left(\gamma(1)x(1) - \int_0^1 x(s)\,\mathrm{d}\gamma(s) - \epsilon\int_s^t \gamma(u)\,\sigma(u)\,\mathrm{d}u\right)\mathrm{d}s.\end{aligned}$$

Since multiplication by $\kappa \exp\left(-\frac{\epsilon}{2}\int_0^t \sigma_s^2\,\mathrm{d}s\right)$ and exp are continuous transformations, continuity of θ is guaranteed if we prove continuity of ϕ and ψ. This is what we show below. For ϕ,

we have

$$
\begin{aligned}
\|\phi(x)-\phi(y)\|_\infty &= \left\|\left(\sigma(t)x(t)-\int_0^t x(s)\,\mathrm{d}\sigma(s)\right)-\left(\sigma(t)y(t)-\int_0^t y(s)\,\mathrm{d}\sigma(s)\right)\right\|_\infty \\
&\leq \|\sigma(t)\,(x(t)-y(t))\|_\infty + \left\|\int_0^t (x(s)-y(s))\,\mathrm{d}\sigma(s)\right\|_\infty \\
&\leq \|\sigma\|_\infty \|x-y\|_\infty + |\sigma(t)-\sigma(0)|\,\|x-y\|_\infty \\
&\leq 3\,\|\sigma\|_\infty \|x-y\|_\infty ,
\end{aligned}
$$

so ϕ is continuous.

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

$$
\begin{aligned}
\|\psi(x)-\psi(y)\|_\infty \leq &\left\|\int_0^t L_f\left[\left(\gamma(1)x(1)-\int_0^1 x(s)\,\mathrm{d}\gamma(s)-\cancel{\epsilon\int_s^t \gamma(u)\,\sigma_u\,\mathrm{d}u}\right)\right.\right. \\
&\qquad \left.\left.-\left(\gamma(1)y(1)-\int_0^1 y(s)\,\mathrm{d}\gamma(s)-\cancel{\epsilon\int_s^t \gamma(u)\,\sigma_u\,\mathrm{d}u}\right)\right]\mathrm{d}s\right\|_\infty \\
\leq & L_f\left\|\int_0^t \left(\gamma(1)\,(x(1)-y(1))-\int_0^1 (x(s)-y(s))\,\mathrm{d}\gamma(s)\right)\mathrm{d}s\right\|_\infty \\
\leq & L_f\left(\|\gamma\|_\infty \|x-y\|_\infty + |\gamma(1)-\gamma(0)|\,\|x-y\|\right)_\infty \\
= & 3L_f\,\|\gamma\|_\infty \|x-y\|_\infty ,
\end{aligned}
$$

which proves the continuity of ψ. □

The following result now follows directly from the continuity of θ (lemma 8.4.3), the continuity principle (theorem 8.2.2), and Schilder's theorem (theorem 8.3.1).

Theorem 8.4.4. *The laws of the solutions Z_κ^ϵ given by equation (8.4.2) of the family of stochastic differential equations given by equation (8.4.1) follow a large deviation principle on $\left(\mathcal{C}_\kappa, \|\cdot\|_\infty\right)$ with the rate function*

$$J(y) = \inf\left\{I \circ \theta^{-1}(y)\right\}, \tag{8.4.5}$$

where θ is as defined in lemma 8.4.3, and I is the Schilder's rate function given in theorem 8.3.1.

8.5. LSDEs with anticipating coefficients and random initial condition

Is it not necessary for the family of linear stochastic differential equations equation (8.4.1) to start at a constant point $\kappa \in \mathbb{R}$ in order for it to have a large deviation principle. In this section, we generalize theorem 8.4.4 and show that we can derive a similar result under a stronger version of exponential equivalence and more restrictive conditions on the functions f, σ, and γ.

Suppose σ and γ are deterministic functions of bounded variation on $[0,1]$. Moreover, suppose $f \in C^2(\mathbb{R})$ is Lipschitz continuous along with $f, f', f'' \in L^1(\mathbb{R})$. Consider the family of linear

stochastic differential equations with parameter $\epsilon > 0$ given by

$$\begin{cases} \mathrm{d}Z_\xi^\epsilon(t) = f\left(\sqrt{\epsilon}\int_0^1 \gamma(s)\,\mathrm{d}W_s\right) Z_\xi^\epsilon(t)\,\mathrm{d}t + \sqrt{\epsilon}\sigma_t\, Z_\xi^\epsilon(t)\,\mathrm{d}W_t \\ Z_\xi^\epsilon(0) = \xi^\epsilon, \end{cases} \tag{8.5.1}$$

where ξ^ϵs are random variables independent of the Wiener process W. Just as before, the unique solutions to equation (8.5.1) are given by

$$\begin{aligned} Z_\xi^\epsilon(t) = \xi^\epsilon \exp\Bigg[&\sqrt{\epsilon}\int_0^t \sigma(s)\,\mathrm{d}W_s - \frac{\epsilon}{2}\int_0^t \sigma(s)^2\,\mathrm{d}s \\ &+ \int_0^t f\left(\sqrt{\epsilon}\int_0^1 \gamma(u)\,\mathrm{d}W_u - \epsilon\int_s^t \gamma(u)\,\sigma(u)\,\mathrm{d}u\right)\mathrm{d}s\Bigg] \end{aligned} \tag{8.5.2}$$

We now state a more general large deviation principle.

Theorem 8.5.3. *Let $\kappa \in \mathbb{R}$ and consider the family of random variables ξ^ϵ such that the following hold*

$$\lim_{\epsilon\to 0} \epsilon \log \mathbb{E}\left[(\xi^\epsilon - \kappa)^2\right] = -\infty. \tag{8.5.4}$$

Moreover, assume that the functions f, f', σ, γ are all bounded. Then the laws of the solutions Z_ξ^ϵ given by equation (8.5.2) of the family of stochastic differential equations given by equation (8.5.1) follow a large deviation principle on $\left(\mathcal{C}_\kappa, \|\cdot\|_\infty\right)$ with the rate function given by equation (8.4.5), where θ is as defined in lemma 8.4.3, and I is the Schilder's rate function given in theorem 8.3.1.

Proof. Let $V^\epsilon = Z_\xi^\epsilon - Z_\kappa^\epsilon$. Then V^ϵ satisfies the stochastic differential equation

$$\begin{cases} \mathrm{d}V^\epsilon(t) = f\left(\sqrt{\epsilon}\int_0^1 \gamma(s)\,\mathrm{d}W_s\right) V^\epsilon(t)\,\mathrm{d}t + \sqrt{\epsilon}\sigma_t V^\epsilon(t)\,\mathrm{d}W_t \\ V^\epsilon(0) = \xi^\epsilon - \kappa, \end{cases} \tag{8.5.5}$$

whose solution is given by

$$\begin{aligned} V^\epsilon(t) = (\xi^\epsilon - \kappa)\exp\Bigg[&\sqrt{\epsilon}\int_0^t \sigma_s\,\mathrm{d}W_s - \frac{\epsilon}{2}\int_0^t \sigma_s^2\,\mathrm{d}s \\ &+ \int_0^t f\left(\sqrt{\epsilon}\int_0^1 \gamma(u)\,\mathrm{d}W_u - \epsilon\int_s^t \gamma(u)\,\sigma_u\,\mathrm{d}u\right)\mathrm{d}s\Bigg]. \end{aligned}$$

Let $\phi(z) = |z|^2$ and let $U^\epsilon = \phi(V^\epsilon)$. From theorem 7.2.3, U^ϵ satisfies the integral equation

$$
\begin{aligned}
U^\epsilon(t) = {} & (\xi^\epsilon - \kappa)^2 + 2\sqrt{\epsilon} \int_0^t \sigma_s \, U^\epsilon(s) \, \mathrm{d}W_s \\
& + \epsilon \int_0^t \sigma_s^2 \, U^\epsilon(s) \, \mathrm{d}s + f\left(\int_0^1 \sqrt{\epsilon} \, \gamma(s) \, \mathrm{d}W_s \right) \int_0^t U^\epsilon(s) \, \mathrm{d}s \\
& + 2\epsilon \int_0^t \gamma(s) \, \sigma_s \, U^\epsilon(s) \int_0^s f'\left(\int_0^1 \sqrt{\epsilon} \, \gamma(v) \, \mathrm{d}W_v - \epsilon \int_u^s \gamma(v) \, \sigma_v \, \mathrm{d}v \right) \mathrm{d}u \, \mathrm{d}s.
\end{aligned}
$$

Fix $\delta > 0$ and let $\tau = \inf\{t \in [0,1] : |V^\epsilon(t)| \geq \delta\}$. Taking expectation of the stopped process $U^\epsilon(t \wedge \tau)$, we get

$$
\begin{aligned}
& \mathbb{E}(U^\epsilon(t \wedge \tau)) \\
& = \mathbb{E}\left[(\xi^\epsilon - \kappa)^2 \right] + 2\sqrt{\epsilon} \, \mathbb{E}\left[\int_0^{t \wedge \tau} \sigma_s \, U^\epsilon(s \wedge \tau) \, \mathrm{d}W_s \right] \\
& \quad + \epsilon \, \mathbb{E}\left[\int_0^{t \wedge \tau} \sigma_s^2 \, U^\epsilon(s \wedge \tau) \, \mathrm{d}s \right] + \mathbb{E}\left[f\left(\int_0^1 \sqrt{\epsilon} \, \gamma(s) \, \mathrm{d}W_s \right) \int_0^{t \wedge \tau} U^\epsilon(s \wedge \tau) \, \mathrm{d}s \right] \\
& \quad + 2\epsilon \, \mathbb{E}\left[\int_0^{t \wedge \tau} \gamma(s) \, \sigma_s \, U^\epsilon(s \wedge \tau) \int_0^s f'\left(\int_0^1 \sqrt{\epsilon} \, \gamma(v) \, \mathrm{d}W_v - \epsilon \int_u^s \gamma(v) \, \sigma_v \, \mathrm{d}v \right) \mathrm{d}u \, \mathrm{d}s \right].
\end{aligned}
$$

The second integral on the right-hand side is a near-martingales by theorem 5.2.1. Suppose f, f', σ, γ are all bounded by some $M > 1$. Using non-negativity of U^ϵ and the near-martingale optional stopping theorem (corollary 5.3.9), we get

$$
\begin{aligned}
\mathbb{E}(U^\epsilon(t \wedge \tau)) \leq {} & \mathbb{E}\left[(\xi^\epsilon - \kappa)^2 \right] + 0 \\
& + \epsilon M^2 \, \mathbb{E}\left[\int_0^{t \wedge \tau} U^\epsilon(s \wedge \tau) \, \mathrm{d}s \right] + M \, \mathbb{E}\left[\int_0^{t \wedge \tau} U^\epsilon(s \wedge \tau) \, \mathrm{d}s \right] \\
& + 2\epsilon M^3 \, \mathbb{E}\left[\int_0^{t \wedge \tau} U^\epsilon(s \wedge \tau) \, \mathrm{d}s \right] \\
\leq {} & \mathbb{E}\left[(\xi^\epsilon - \kappa)^2 \right] + \left(M + 2\epsilon M^3 \right) \mathbb{E}\left[\int_0^{t \wedge \tau} U^\epsilon(s \wedge \tau) \, \mathrm{d}s \right].
\end{aligned}
$$

By Gronwall's inequality, we get

$$
\mathbb{E}(U^\epsilon(\tau)) = \mathbb{E}(U^\epsilon(1 \wedge \tau)) \leq \mathbb{E}\left[(\xi^\epsilon - \kappa)^2 \right] e^{M + 2\epsilon M^3}.
$$

Since $\phi(z)$ is a monotonically increasing nonnegative function in $|z|$, we use Markov's inequality to get

$$
\mathbb{P}\{|V^\epsilon(\tau)| \geq \delta\} = \mathbb{P}\{\phi(V^\epsilon(\tau)) \geq \phi(\delta)\} \leq \frac{\mathbb{E}(\phi(V^\epsilon(\tau)))}{\phi(\delta)} = \frac{\mathbb{E}(U^\epsilon(\tau))}{\delta^2} \leq \frac{\mathbb{E}\left[(\xi^\epsilon - \kappa)^2 \right]}{\delta^2} e^{M + 2\epsilon M^3}.
$$

Taking log and multiplying by ϵ, we get

$$\epsilon \log \mathbb{P}\{|V^\epsilon(\tau)| \geq \delta\} \leq \epsilon \log \mathbb{E}\Big[(\xi^\epsilon - \kappa)^2\Big] - 2\epsilon \log \delta + \epsilon(M + 2\epsilon M^3).$$

Finally, taking limit of $\epsilon \to 0$ and using equation (8.5.4),

$$\lim_{\epsilon \to 0} \epsilon \log \mathbb{P}\{|V^\epsilon(\tau)| \geq \delta\} = -\infty.$$

This result allows us to say that Z_ξ^ϵ and Z_κ^ϵ are exponentially equivalent. Since exponentially equivalent families have the same large deviation principle due to theorem 8.2.4, Z_ξ^ϵ follows a large deviation principle with the same rate function given by equation (8.4.5). $\square$

Chapter 9. Future Research Directions

Being a new approach to anticipating stochastic integrals, there are various avenues of research for the Ayed–Kuo stochastic calculus. In this chapter, we look at the current state of research and highlight some open problems. Unless specified otherwise, we take $t \in [a, b]$ for $0 \leq a \leq b < \infty$.

9.1. Definition and domain of the integral

One of the most important questions is to identify the largest class of stochastic processes for which the Ayed–Kuo integral exists. In chapter 4, whe showed that for $\Theta(x, y) \in C^1(\mathbb{R}^2)$, the integral $\int_a^b \Theta(W_t, W_b - W_t)\,\mathrm{d}W_t$ exists if

$$\Theta(W_t, W_b - W_t),\ \Theta_x(W_t, W_b - W_t),\ \Theta_y(W_t, W_b - W_t) \in L^2([a, b] \times \Omega).$$

However, it is unknown whether a general stochastic process can be integrated.

Note that it was shown in [Par17, theorem 2.3] that the Ayed–Kuo integral and the Hitsuda–Skorokhod integral are equivalent if the definition of the Ayed–Kuo integral uses $L^2(\Omega)$ convergence. However, since the original definition of the Ayed–Kuo integral uses convergence in probability, its domain is a strict superset of the domain of the Hitsuda–Skorokhod integral.

This is closely related to finding a decomposition result for general anticipating processes. For example, while calculating examples for the Ayed–Kuo integral, we used the decomposition $W_b = W_t + (W_b - W_t)$ of the anticipating integrand W_b into the adapted integrand W_t and the instantly-independent integrand $W_b - W_t$. Is it possible to decompose any stochastic process into adapted and anticipating parts (or their sums of products thereof)? If not, can we determine a class for which such a decomposition can be guaranteed? Such a result would greatly expand upon the class of integrands.

Another open problem is to generalize the extension of Itô's isometry to products of the form $X_t Y^t$, where X_t is adapted and Y^t is instantly-independent. This probably requires the definition of an equivalent of the white noise derivative operator.

9.2. Near-martingales

First, we highlight some results that have already been shown for near-martingales.

1. In theorem 5.2.1, we showed that Ayed–Kuo integrals are near-martingales.
2. Proposition 5.1.3 shows that process is a near-submartingale if and only if its conditioned process is a martingale.
3. Proposition 5.2.4 shows that the product of a submartingale and an instantly-independent process is a near-submartingale if and only if the instantly-independent process has constant mean.
4. In theorem 5.3.3, we proved that stopped near-martingales are near-martingales.
5. In theorem 5.3.7, we established a near-martingale optional stopping theorem.

6. A Doob–Mayer's decomposition for near-submartingales in the lines of theorem 1.2.3 was shown in [Hwa+17, section 3].
   Theorem 9.2.1. *Let X be a continuous near-submartingale. Then X has a unique decomposition*
   $$X(t) = N(t) + A_t,$$
   where N is a continuous near-martingale, and A is a continuous adapted process starting at 0 that is increasing in t almost surely.

There are a lot of open problems in this area. We look at some results that exist for martingales that we do not currently have near-martingale equivalents of.

1. Doob's upcrossing inequality
   Theorem 9.2.2 ([KS14, theorem 3.3.2])**.** *Let X_n be a discrete-time submartingale. Denote the number of time X crosses over $[p, q]$ (starting below p and going above q) by $U_{[p,q]}$. Then*
   $$\mathbb{E}\big(U_{[p,q]}\big) \leq \sup_n \frac{\mathbb{E}(X_n - p)^+}{q - p}.$$
2. Doob's martingale convergence theorem
   Theorem 9.2.3 ([KS14, theorem 3.3.4])**.** *Let X be an L^1-bounded submartingale with right-continuous paths. Then there exists an integrable random variable X_∞ such that $X_n \to X_\infty$ almost surely as $t \to \infty$.*
3. Doob's martingale inequalities
   Theorem 9.2.4 ([KS14, theorem 3.3.9])**.** *Let X be a right-continuous submartingale for $t \in [a, b]$, and $X^* = \sup_t X_t$. Then the following hold:*
   (a) For any $\lambda > 0$, we have
   $$\mathbb{P}(X^* > \lambda) \leq \frac{\mathbb{E}\big(X_b^+\big)}{\lambda}.$$
   (b) If $X_b \in L^p$ for some $p \in (1, \infty)$, then
   $$\left\|X^*\right\|_p \leq \frac{p}{p-1} \left\|X_b\right\|_p.$$
4. Burkhölder–Davis–Gundy inequality
   Theorem 9.2.5 ([KS14, theorem 5.6.3])**.** *Suppose M is a continuous locally square integrable martingale. Let $M^* = \sup_t M_t$ and $\langle M\rangle$ denote the quadratic variation process of M. Then, for any $p \in (0, \infty)$, there exists universal constants c_p and C_p such that*
   $$c_p\, \mathbb{E}\big[\langle M\rangle_t^p\big] \leq \mathbb{E}\Big[(M_t^*)^{2p}\Big] \leq C_p\, \mathbb{E}\big[\langle M\rangle_t^p\big].$$

9.3. Girsanov's theorem

Recall that Girsanov's theorem (theorem 1.6.9) tells us that if a Wiener process is can be translated in certain directions to obtain another Wiener process with respect to an equivalent probability measure. Special cases of Girsanov's theorem for the new stochastic integral was given in [KPS13].

Theorem 9.3.1 ([KPS13, theorems 5.1 and 5.2]). *Let X and Y be continuous square-integrable processes on $[0, T]$ such that X is adapted and Y is instantly-independent. Let*

$$\widetilde{W}_t = W_t + \int_0^t (X_s + Y^s)\,\mathrm{d}s, \qquad \text{and} \qquad \widehat{W}_t = W_T - W_t + \int_t^T (X_s + Y^s)\,\mathrm{d}s.$$

Then the stochastic processes $\widetilde{W}_t$ and $\widehat{W}_t{}^2 - (T - t)$ are near-martingales with respect to the probability measure $\mathbb{Q}$ defined by the Radon–Nikodym derivative $\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}} = \mathcal{E}_T^{(X+Y)}$.

We do not currently have a Girsanov's theorem where the translation is a product of adapted and instantly-independent processes, or other functions of them.

This raises another interesting question. Let (i, H, B) be an abstract Wiener space with the standard Gaussian measure μ on B, and $T : B \to B$ a nonlinear transformation such that $\mu \circ T^{-1}$ is absolutely continuous with respect to μ. Since the Radon-Nikodym derivative $\frac{\mathrm{d}(\mu\circ T^{-1})}{\mathrm{d}\mu}$ is related to the exponential process in Girsanov's theorem, can use the translation formula for an abstract Wiener space to derive the exponential process required for a general Girsanov's theorem?

9.4. Near-Markov property

It is well known that the solutions of stochastic differential equations in Itô's theory are Markov processes.

Theorem 9.4.1 ([Kuo06, theorem 10.6.1]). *Suppose $m(t, x), \sigma(t, x) : [a, b] \times \mathbb{R} \to \mathbb{R}$ are adapted and satisfy the Lipschitz and linear growth conditions on x. Let ξ be a $\mathcal{F}_a$-measurable square-integrable random variable. Then the unique solution X of the stochastic differential equation*

$$\begin{cases} \mathrm{d}X_t = m(t, X_t)\,\mathrm{d}t + \sigma(t, X_t)\,\mathrm{d}W_t \\ X_0 = \xi \end{cases}$$

is a Markov process. That is, for any $s \leq t$, we have $\mathbb{P}(X_t \in A \mid \mathcal{F}_s) = \mathbb{P}(X_t \in A \mid X_s)$.

In fact, the martingale property and the Markov properties were Itô's original motivations behind the definition of Itô's integral.

However, it is not true that solutions of stochastic differential equations with anticipating initial conditions are Markov processes. Therefore, we would like to know if there is an analogue of Markov processes that are satisfied by solutions of stochastic differential equations in the Ayed–Kuo theory, on the lines of how we defined near-martingales from martingales.

A follow-up question concerns Wiener martingales. An Wiener process X is called a *Wiener martingale* with respect to a filtration $\mathcal{F}$ if it is $\mathcal{F}$-adapted and $X_t - X_s$ is independent of $\mathcal{F}_s$ for any $s \leq t$. We know that Wiener martingales have the strong Markov property, as shown in the following theorem.

Theorem 9.4.2 ([KS14, theorem 3.2.3]). *Let X be a Wiener martingale with respect to the filtration $\mathcal{F}$, and let τ be a finite $\mathcal{F}$-stopping time. Then*

1. the process $Y_t = X_{\tau+t} - X_\tau$ for $t \geq 0$ is a Wiener process, and
2. the σ-algebra $\sigma\{Y_t \mid t \geq 0\}$ is independent of $\mathcal{F}_\tau$.

Therefore, it is natural to ask if we can get a similar "near-Markov" result for "Wiener near-martingales". This would establish a relationship between near-martingales and the "near-Markov" property.

9.5. Exponential processes.

Exponential processes are very important in Itô's theory of stochastic integration and has numerous applications in mathematical finance. In Itô calculus, there are three major ways to think of the map $\exp\left(\int_0^t h(s)\,\mathrm{d}W_s\right) \mapsto \mathcal{E}_t^{(h)}$ (see also [HKS19, section 1]). For this idea, we will consider $t \in [0,1]$.

1. *Multiplicative renormalization.* For a random variable X with non-zero expectation, the *multiplicative renormalization* of X is the random variable $X/\mathbb{E}(X)$. For any $h \in L^2[0,1]$, the Wiener integral $\int_0^t h(s)\,\mathrm{d}W_s$ is a Gaussian random variable with mean 0 and variance $\int_0^t h(s)^2\,\mathrm{d}s$. Using the idea of moment generating functions, we get

$$\mathbb{E}\left[\exp\left(\int_0^t h(s)\,\mathrm{d}W_s\right)\right] = \exp\left(\frac{1}{2}\int_0^t h(s)^2\,\mathrm{d}s\right).$$

   Therefore, $\mathcal{E}_t^{(h)}$ is the multiplicative renormalization of $\exp\left(\int_0^t h(s)\,\mathrm{d}W_s\right)$.
2. *Martingales.* Under the condition $\mathbb{E}\exp\left(\frac{1}{2}\int_0^1 h(s)^2\,\mathrm{d}s\right) < \infty$ (Novikov), $\mathcal{E}^{(h)}$ is a martingale. Therefore, for any deterministic h, the exponential process $\mathcal{E}^{(h)}$ is necessarily a martingale since the Novikov condition is trivially satisfied.
3. *Stochastic differential equations.* From remark 1.6.7, we can say that $\mathcal{E}^{(h)}$ is the solution of the stochastic differential equation

$$\begin{cases} \mathrm{d}\mathcal{E}_t^{(h)} = \sigma_t \mathcal{E}_t^{(h)}\,\mathrm{d}W_t \\ \mathcal{E}_a^{(h)} = 1. \end{cases}$$

These ideas coincide even when h is adapted. Therefore, a natural question is whether these ideas concur when h is anticipating.

The non-trivial nature of the stochastic exponential for anticipating processes is demonstrated by the following example.

Example 9.5.1 ([Item 5 of Kuo14, section 4.6])**.** Suppose we consider the near-martingale property to define the exponential process. Then for any $\theta \in C^2(\mathbb{R})$, we can use the differential formula to derive the exponential process associated with $\theta(W_1)$ as

$$X_t = \exp\left\{\int_0^t \theta(W_1)\,\mathrm{d}W_s - \int_0^t \left[\frac{1}{2}\theta(W_1)^2 + (\theta'(W_1)W_s - \theta''(W_1)s)\,\theta(W_1)\right]\mathrm{d}s\right\}.$$

9.6. LSDEs and conditioned processes

We covered linear stochastic differential equations with anticipation arising from three areas: (1) initial condition, (2) drift coefficient, and (3) diffusion coefficient. We have mostly focused on special classes for each type. Therefore, there is immense potential of expanding the scope of the individual classes, and show general results that cover combinations of different causes of anticipation. There is also a great scope of understanding probabilistic behavior for such equations, over and above the large deviations results we derived in chapter 8. Non-linear stochastic differential equations with anticipation is another exciting field of research.

We also extensively studied the conditioned processes and its stochastic differential equation in chapter 7 for linear stochastic differential equations with anticipating initial conditions. We can perform similar studies for linear stochastic differential equations with anticipating coefficients in both the diffusion and the drift terms.

9.7. Applications

The most well-known application of anticipating integrals is in modeling insider trading in the field of mathematical finance. This idea can also be extended to model prices of assets determined by initial public offerings, where information is asymmetric by nature. Similar is the case of venture capital investments, which are typically closed-off to retail investors. An extension of the Black–Scholes–Merton model was obtained by Zhai[Zha18, chapter 5] under restrictive conditions of deterministic coefficients for the asset dynamics. The author showed the completeness of the market, and derived call option pricing and hedging formulas in the presence of initial conditions that are functions of W_1. There remains much room to generalize these results.

Mathematical finance is not the only application area of anticipating integrands. These tools apply to any situation where there is an asymmetry of information about the future. As a hypothetical use case, consider a global logistics company that has access to information about the geopolitics of a particular region. Using this information, the organization can constantly update its delivery systems. Consider another hypothetical use case, where an environmental agency plans to release a certain amount of genetically modified mosquitoes into an environment one year down the line. However, it does not want to disclose its plans to the public ahead of time since it expects negative public perceptions of such a step. In this situation, anticipating calculus may be used to model the population dynamics of mosquitoes. Of course, the exact modeling procedure and application of anticipating stochastic calculus would depend on the specifics of the given problem.

Appendix. Copyright Information

Copyright information for Hui-Hsiung Kuo, Pujan Shrestha, Sudip Sinha, and Padmanabhan Sundar. *On near-martingales and a class of anticipating linear SDEs*. 2022. arXiv: 2204.01932 [math.PR] used in chapters 5, 7 and 8.

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.

Copyright information for Hui-Hsiung Kuo, Pujan Shrestha, and Sudip Sinha. "An Intrinsic Proof of an Extension of Itô's Isometry for Anticipating Stochastic Integrals". In: *Journal of Stochastic Analysis* 2.4 (2021). DOI: 10.31390/josa.2.4.08 used in 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:

$$\begin{aligned}\mathbb{E}\left[\left(\int_a^b f(B_t)\phi(B_b - B_t)\, dB_t\right)^2\right] &= \int_a^b \mathbb{E}\Big[f(B_t)^2\phi(B_b - B_t)^2\Big]\, dt \\ &\quad + 2\int_a^b \int_a^t \mathbb{E}\Big[f(B_s)\phi'(B_b - B_s)f'(B_t)\phi(B_b - B_t)\Big]\, ds\, dt,\end{aligned} \tag{1.2}$$

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 Hui-Hsiung Kuo, Pujan Shrestha, and Sudip Sinha. "Anticipating Linear Stochastic Differential Equations with Adapted Coefficients". In: *Journal of Stochastic Analysis* 2.2 (2021). DOI: 10.31390/josa.2.2.05 used in chapters 3 and 6.

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.

Copyright information for Hui-Hsiung Kuo, Sudip Sinha, and Jiayu Zhai. "Stochastic Differential Equations with Anticipating Initial Conditions". In: *Communications on Stochastic Analysis* 12.4 (2018). DOI: 10.31390/cosa.12.4.06 used in chapter 3.

Communications on Stochastic Analysis
Vol. 12, No. 4 (2018) 473-485

STOCHASTIC DIFFERENTIAL EQUATIONS WITH ANTICIPATING INITIAL CONDITIONS

HUI-HSIUNG KUO, SUDIP SINHA, AND JIAYU ZHAI*

Abstract. In this paper, we study the solutions of a stochastic differential equation with various anticipating initial conditions. We show that the conditional expectation of the solution of such a stochastic differential equation is not simply the solution of the corresponding stochastic differential equation with initial condition taken as the conditional expectation of the anticipating initial condition. We derive the conditional expectation of the solution in general, and apply it to the special case of anticipating initial condition given by Hermite polynomials. We also extend the class of initial conditions to functions of Wiener integrals.

1. Introduction

In 1942 [7], Kiyosi Itô published his pioneering paper on stochastic integration, which enabled integration of stochastic processes with respect to a Brownian motion. In 1944 [8], his efforts to model Markov processes led him to construct stochastic differential equations of the form $dX_t = \alpha(X_t)\, dB(t) + \beta(X_t)\, dt$, $X_0 = x$, which subsequently led him to publish what is now known as the Itô formula. In 1973, Black and Scholes [4], and Merton [12] used Itô's calculus to give a framework for option pricing, which rapidly expanded the interest of stochastic calculus to practitioners in other fields.

Even though it is extremely useful, the Itô calculus cannot handle *anticipating* conditions. For example, consider the following simple stochastic differential equation with anticipating initial condition

$$\begin{cases} dX_t = X_t\, dB(t), \quad t \in [0,1], \\ X_0 = B(1). \end{cases} \tag{1.1}$$

To solve the equation analytically, we have to assign a meaning to the integral

$$\int_0^t B(1)\, dB(s), \quad t \in [0,1], \tag{1.2}$$

which is outside the theory of Itô calculus since the integral is not adapted with respect to the filtration generated by the Brownian motion $B(t)$. This is the primary motivation for extending the Itô integral to anticipating integrands.

Received 2018-11-11; Communicated by the editors.

2010 *Mathematics Subject Classification.* Primary 60H05; Secondary 60H10, 60H20.

Key words and phrases. Stochastic integral, near-martingale, Itô formula, stochastic differential equations, conditional expectation.

* Corresponding author.

Bibliography

[AK08] Wided Ayed and Hui-Hsiung Kuo. "An extension of the Itô integral". In: *Communications on Stochastic Analysis* 2.3 (2008), pp. 323–333. DOI: 10.31390/cosa.2.3.05.

[Bac00] Louis Bachelier. "Théorie de la spéculation". In: *Annales scientifiques de l'École Normale Supérieure*. Vol. 17. 3. 1900, pp. 21–86. DOI: 10.24033/asens.476.

[BS73] Fischer Black and Myron Scholes. "The Pricing of Options and Corporate Liabilities". In: *Journal of Political Economy* 81.3 (1973), pp. 637–654. ISSN: 00223808, 1537534X. URL: http://www.jstor.org/stable/1831029.

[Bro28] Robert Brown. "XXVII. 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". In: *The Philosophical Magazine* 4.21 (1828), pp. 161–173. DOI: 10.1080/14786442808674769.

[Cra38] Harald Cramér. "Sur un nouveau théorème-limite de la théorie des probabilités". French. In: *Actualités Scientifiques et Industrielles* 736 (1938), pp. 5–23.

[DZ98] Amir Dembo and Ofer Zeitouni. *Large Deviations Techniques and Applications*. Applications of mathematics. Springer, 1998. ISBN: 9780387984063. DOI: 10.1007/978-3-642-03311-7.

[Ein05] Albert Einstein. "Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen". In: *Annalen der Physik* 322.8 (1905), pp. 549–560. DOI: 10.1002/andp.19053220806.

[Gir60] I. V. Girsanov. "On Transforming a Certain Class of Stochastic Processes by Absolutely Continuous Substitution of Measures". In: *Theory of Probability & Its Applications* 5 (3 1960), pp. 285–301. DOI: 10.1137/1105027.

[HKS19] Chii Ruey Hwang, Hui-Hsiung Kuo, and Kimiaki Saitô. "Anticipating Exponential Processes and Stochastic Differential Equations". In: *Communications on Stochastic Analysis* 13.3 (2019). DOI: 10.31390/cosa.13.3.09.

[Hwa+16] Chii-Ruey Hwang, Hui-Hsiung Kuo, Kimiaki Saitô, and Jiayu Zhai. "A general Itô formula for adapted and instantly independent stochastic processes". In: *Communications on Stochastic Analysis* 10.3 (2016). DOI: 10.31390/cosa.10.3.05.

[Hwa+17] Chii-Ruey Hwang, Hui-Hsiung Kuo, Kimiaki Saitô, and Jiayu Zhai. "Near-martingale Property of Anticipating Stochastic Integration". In: *Communications on Stochastic Analysis* 11.4 (2017). DOI: 10.31390/cosa.11.4.06.

[Itô44] Kiyosi Itô. "Stochastic integral". In: *Proc. Imp. Acad.* 20.8 (1944), pp. 519–524. DOI: 10.3792/pia/1195572786. URL: https://doi.org/10.3792/pia/1195572786.

[Itô51] Kiyosi Itô. "Multiple Wiener Integral". In: *J. Math. Soc. Japan* 3.1 (May 1951), pp. 157–169. DOI: 10.2969/jmsj/00310157. URL: https://doi.org/10.2969/jmsj/00310157.

[Itô78] Kiyosi Itô. "Extension of stochastic integrals". In: *Proceedings of the International Symposium on Stochastic Differential Equations*. Kinokuniya, 1978, pp. 95–109.

[KS14] Gopinath Kallianpur and P Sundar. *Stochastic Analysis and Diffusion Processes*. 2014. ISBN: 9780199657063. DOI: 10.1093/acprof:oso/9780199657063.001.0001.

[Kha+13] Narjess Khalifa, Hui-Hsiung Kuo, Habib Ouerdiane, and Benedykt Szozda. "Linear stochastic differential equations with anticipating initial conditions". In: *Communications on Stochastic Analysis* 7.2 (2013). DOI: 10.31390/cosa.7.2.05.

[Kol19] Andreĭ Nikolaevich Kolmogorov. *Foundations of the Theory of Probability*. Trans. by Nathan Morrison. AMS Chelsea Publishing. American Mathematical Society, 2019. ISBN: 9781470452995. URL: https://books.google.com/books?id=BdDBDwAAQBAJ.

[Kuo96] Hui-Hsiung Kuo. *White Noise Distribution Theory*. 1st ed. Universitext. CRC Press, 1996. ISBN: 978-0-367-57961-6.

[Kuo06] Hui-Hsiung Kuo. *Introduction to Stochastic Integration*. 1st ed. Universitext. New York: Springer-Verlag, 2006. ISBN: 978-0-387-31057-2. DOI: 10.1007/0-387-31057-6.

[Kuo14] Hui-Hsiung Kuo. "The Itô calculus and white noise theory: a brief survey toward general stochastic integration". In: *Communications on Stochastic Analysis* 8.1 (2014). DOI: 10.31390/cosa.8.1.07.

[KPS13] Hui-Hsiung Kuo, Yun Peng, and Benedykt Szozda. "Generalization of the anticipative Girsanov theorem". In: *Communications on Stochastic Analysis* 7.4 (2013). DOI: 10.31390/cosa.7.4.06.

[KSS13] Hui-Hsiung Kuo, Anuwat Sae-Tang, and Benedykt Szozda. "An isometry formula for a new stochastic integral". In: *Quantum Probability and Related Topics*. Ed.

by Luigi Accardi and Franco Fagnola. Vol. 29. QP-PQ: Quantum Probability and White Noise Analysis. World Scientific, 2013, pp. 222–232. DOI: 10.1142/9789814447546_0014.

[KSS21a] Hui-Hsiung Kuo, Pujan Shrestha, and Sudip Sinha. “An Intrinsic Proof of an Extension of Itô’s Isometry for Anticipating Stochastic Integrals”. In: *Journal of Stochastic Analysis* 2.4 (2021). DOI: 10.31390/josa.2.4.08.

[KSS21b] Hui-Hsiung Kuo, Pujan Shrestha, and Sudip Sinha. “Anticipating Linear Stochastic Differential Equations with Adapted Coefficients”. In: *Journal of Stochastic Analysis* 2.2 (2021). DOI: 10.31390/josa.2.2.05.

[Kuo+22] Hui-Hsiung Kuo, Pujan Shrestha, Sudip Sinha, and Padmanabhan Sundar. *On near-martingales and a class of anticipating linear SDEs*. 2022. arXiv: 2204.01932 [math.PR].

[KSZ18] Hui-Hsiung Kuo, Sudip Sinha, and Jiayu Zhai. “Stochastic Differential Equations with Anticipating Initial Conditions”. In: *Communications on Stochastic Analysis* 12.4 (2018). DOI: 10.31390/cosa.12.4.06.

[Leb02] Henri Lebesgue. “Intégrale, Longueur, aire”. PhD thesis. Université de Paris, 1902. URL: http://lsulibraries.worldcat.org/oclc/9433794.

[Mer74] Robert C. Merton. “On the Pricing of Corporate Debt: The Risk Structure of Interest Rates”. In: *The Journal of Finance* 29.2 (1974), pp. 449–470. ISSN: 00221082, 15406261. URL: https://www.jstor.org/stable/2978814.

[Nua06] David Nualart. *The Malliavin Calculus and Related Topics*. 2nd ed. Berlin, Heidelberg: Springer-Verlag, 2006. ISBN: 978-3-540-28329-4. DOI: 10.1007/3-540-28329-3.

[Par17] Peter Parczewski. “Extensions of the Hitsuda–Skorokhod Integral”. In: *Communications on Stochastic Analysis* 11.4 (2017). DOI: 10.31390/cosa.11.4.05.

[Sch66] M. Schilder. “Some Asymptotic Formulas for Wiener Integrals”. In: *Transactions of the American Mathematical Society* 125.1 (1966), pp. 63–85. ISSN: 00029947. DOI: 10.2307/1994588.

[Tro59] H. F. Trotter. “On the Product of Semi-Groups of Operators”. In: *Proceedings of the American Mathematical Society* 10.4 (1959), pp. 545–551. ISSN: 00029939, 10886826. DOI: 10.2307/2033649.

[Var66] S. R. S. Varadhan. “Asymptotic probabilities and differential equations”. In: *Communications on Pure and Applied Mathematics* 19 (1966), pp. 261–286. DOI: 10.1002/cpa.3160190303.

[Wie23] Norbert Wiener. “Differential-Space”. In: *Journal of Mathematics and Physics* 2.1–4 (1923), pp. 131–174. DOI: 10.1002/sapm192321131.

[Wie38] Norbert Wiener. “The Homogeneous Chaos”. In: *American Journal of Mathematics* 60.4 (1938), pp. 897–936. ISSN: 00029327, 10806377. DOI: 10.2307/2371268.

[Wil91] David Williams. *Probability with Martingales*. Cambridge University Press, 1991. DOI: 10.1017/CBO9780511813658.

[Zha18] Jiayu Zhai. “General Stochastic Integral and Itô Formula with Application to Stochastic Differential Equations and Mathematical Finance”. LSU Doctoral Dissertations. Louisiana State University, Agricultural, and Mechanical College, Mar. 2018. URL: https://digitalcommons.lsu.edu/gradschool_dissertations/4518.

Vita

Born in a picturesque Indian town at the foothills of the Himalayas, Sudip Sinha grew up in Kolkata. After obtaining his bachelor's degree in chemical engineering from Birla Institute of Technology & Science, Pilani in Goa, he worked for the Innovation Lab at Mu Sigma in Bengaluru. Subsequently, he joined the Erasmus Mundus master's program MathMods, studying mathematical modeling at the Università degli Studi dell'Aquila in Italy and Universität Hamburg in Germany.

He loves mathematics and is awestruck by its beauty and universality, and is inquisitive about its philosophical foundations. He has gradually developed a keen interest in teaching and writing expository articles, and desires to continue them throughout his life. After graduating, Sudip plans to join Amazon as an Applied Scientist.

Outside of academics, Sudip is an avid-reader, an amateur guitarist, and a self-confessed travel-addict who loves exploring various cultures and cuisines. Sudip finds peace in nature and hopes for an equitable and harmonious future for planet earth.