People's Democratic Republic of Algeria
Ministry of Higher Education and Scientific Research
Higher Normal School of Technological Education-Skikda

Department of Mathematics & Computer Sciences

Dissertation

Presented to obtain a degree in Mathematics as a teacher of middle school

Entitled

Stochastic Differential Inclusions

Presented by:

Afifi Houda Bouzenount Abir

Board of Examiners:

President: Seghier Fatma Zohra	MCA	ENSET SKIKDA
Supervisor: Ferrag Azouz	MCB	ENSET SKIKDA
Examiner: Khochemane Houssem Eddine	MCA	ENSET SKIKDA
Examiner: Mansouri bouzid	MCA	ENSET SKIKDA

June’s session 2024

Dedication

We dedicate this modest work:

To our parents, may God protect them, for their love, support, and encouragement.

To our brothers and sisters.

To all members of our family, young and old.

To all friends and colleagues.

Acknowledgments

We praise Allah Almighty who enabled us to accomplish this work.

Our deepest gratitude to our supervisor, Mr. Azouz Ferrag, who proposed the topic of this dissertation, for his advice, guidance, and support to us from the beginning of

this work until its completion.

We thank the committee members who accepted to discuss this work.

We thank all our family members and friends who helped and encouraged us to complete this dissertation.
We would like to express our sincere gratitude to our friend, Professor Aoufi Amani, who was a light that God sent us from where we did not expect, we are forever indebted to her for her continuous support and what she has done for us. We give you all our love, and no matter how many words of thanks we say about you, we will not give you what you deserve.

Thanks to all.

Abstract

In this dissertation, we introduce stochastic calculus, stochastic differential equations and stochastic differential inclusions. We begin by presenting the basic concepts of probability. Then, we study some of Itô's calculus, such as the Itô integral and Itô's formula. After studying Itô's calculus for solving stochastic differential equations, we briefly discuss the existence and uniqueness of solutions for a special kind of these equations. Finally, we provide a brief overview of stochastic differential inclusions.

Key words: Stochastic process, Brownian motion, Itô integral, Itô's formula, Stochastic differential equation, Stochastic differential inclusion.

Résumé

Dans ce mémoire, nous présentons le calcul stochastique, les équations différentielles stochastiques et les inclusions différentielles stochastiques. Nous commençons par exposer les concepts de base des probabilités. Ensuite, nous étudions une partie du calcul d'Itô, comme l'intégrale d'Itô et la formule d'Itô. Après avoir étudié le calcul d'Itô pour résoudre les équations différentielles stochastiques, nous discutons brièvement de l'existence et de l'unicité des solutions pour un type particulier de ces équations. Enfin, nous donnons un bref aperçu des inclusions différentielles stochastiques.

Mots clés: Processus stochastique, mouvement Brownien, intégrale d'Itô, formule d'Itô, équation différentielle stochastique, inclusion différentielle stochastique.

Contents

Basic Notation

Some notations will be used throughout this final dissertation that we list below:

- $\emptyset$: The empty set.
- $\mathbb{R}$: The set of real numbers.
- $\mathbb{R}^n$: n-dimensional real space.
- $\mathbb{Z}$: The set of integer numbers.
- $\mathbb{Z}^+$: The set of positive integer numbers.
- $\mathbb{N}$: The set of nature numbers.
- ω: Event, outcome of a random experiment.
- Ω: The set of outcomes.
- A: A subset in Ω.
- A^c: The complement of A in Ω.
- $\mathcal{F}$: The Sigma-field.
- $\mathcal{B}$: The Borel sigma-field.
- $\mathcal{B}(E)$: The Borel sigma-field generated by E.
- $\mathbb{P}$: Probability measure.
- $\mathbb{P}(A)$: probability of the event A.
- $\mathbb{P}(A \cap B)$: Probability of A and B.
- $\mathbb{P}(A|B)$: The conditional probability of B given A.
- $(\Omega, \mathcal{F})$: Mesurable space .
- $(\Omega, \mathcal{F}, \mathbb{P})$: Probability space.
- X or $X(\omega)$: Random variable.
- $f(x)$: Probability density.
- $\mathbb{E}(X)$: Expectation of X.
- $\mathcal{G}$:A sigma-field contained in $\mathcal{F}$.
- $\mathbb{E}(X|\mathcal{G})$: Conditional expectation of the random variable X given $\mathcal{G}$.

- $var(X)$: Variance of X.
- $cov(X, Y)$: Covariance of X and Y.
- $\{X_n\}_{n\geq 1}$: Sequence of random variables.
- a.s : Almoste sure limit.
- m.s : Mean square limit.
- X_t or $X(t, \omega)$: Stochastic process.
- iid : independent and identically distributed.
- T : Index set of the stochastic process.
- S : State space of the stochastic space.
- E_i : State.
- p_{ij}: The probability of a transition from state E_i to state E_j.
- $[p_{ij}]$: Transition matrix.
- B_t : Brownian motion.
- Π : Partition of the interval $[0, t]$.
- $FV_{[0,t]}(f)$: the first variation of f over $[0, t]$.
- $Q_{[0,t]}(f)$: The quadratic variation of f over $[0, t]$.
- w.p.1 : With probability 1.
- $\{\mathcal{F}_k\}_{k=0}^{n}$: Filtration.
- $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$: Filtred probability space.
- $\mathcal{P}(Y)$: The family of all nonempty subsets of Y.
- $cl(\mathbb{R}^n)$: The closed subsets of $\mathbb{R}^n$.
- SDE : Stochastic Differential Equation.
- SDE_s: Stochastic Differential Equations.
- ODE: Ordinary Differential Equation.
- ODE_s: Ordinary Differential Equations.
- DI : Differential Inclusion.
- DI_s : Differential Inclusions.
- SDI : Stochastic Differential Inclusion.
- SDI_s : Stochastic Differential Inclusions.

Introduction

The theory of probability is closely linked to reality (with terms such as: chance, random, event, probability), in the sense that it allows for the construction of a mathematical model for random experiments, which can then be better understood and studied.

Before defining a very general framework for modeling as many random experiments as possible, numerous principles for calculating the probabilities of events were put forward. The works of Pierre de Fermat, Blaise Pascal and Christian Huygens (17th century), then Pierre-Simon de Laplace, Abraham de Moivre, Jacques Bernoulli, and Denis Siméon Poisson (18th century) and Carl Friedrich Gauss and Henri Poincaré (19th century).

Stochastic calculus plays an important role in different branches of sciences, physics, social sciences and technology. The need for modeling to best manage risks, to automate tasks, in short to "industrialize" an activity that was initially very artisanal, found highly sophisticated means of calculation in the theory developed from the 1930s by Kolmogorov, Itō and many other mathematicians. In finance for example, stochastic calculus took on an important place since the mid-1970s and pushed mathematicians to further develop some of their theories for the better.

At the heart of stochastic calculus lies the concept of stochastic differential equations (SDEs). These equations provide a mathematical framework for modeling and analyzing systems that are subject to random fluctuations or noise. SDEs have proven invaluable in capturing the dynamics of phenomena as varied as stock price movements, population dynamics, and turbulent fluid flows.

However, in many real-world scenarios, the underlying dynamics may not be fully described by a single stochastic differential equation. Instead, the system's behavior may be better characterized by a set of inequalities or constraints, known as a stochastic differential inclusion (SDI). SDIs offer a more flexible and general approach, allowing for the incorporation of additional constraints, uncertainties, or nonlinearities that may better reflect the complexity of the system under investigation.

This work is divided into three chapters. The first chapter is devoted to the elementary notions of probability and random variables space, conditional probability, types of random variables, independence, expectation and limits of sequences of random variables.

The second chapter contains four sections, in the first two, we understood the

mening of the term "stochastic" by charifying the difference between deterministic and stochastic models, then, we give four important examples of stochastic process, Bernoulli process, Markov chain, martingales and Brownian motion. The third section explains the stochastic integration with several illustrative examples. The fourth section, potential methodes to solve stochastic differential equation are studied and the notion of stochatic differential equation is given.
The third chapter contains three sections. The first one presents some definitions of multivalued analysis, the second section gives the relation between differential equations and differential inclusions. In the third section, a result on the existence of solution for stochastic nonlocal random functional integral inclusion is given as an application that finalize our work.

I

General Notions and Definitions about Probability

Probability theory is one of the most important branches of mathematics, playing a central role in many fields such as statistics, finance, engineering, physics, biology, and others. Probability theory deals with the study of random phenomena and events that cannot be predicted with certainty, but whose likelihood of occurrence can be calculated.

Probability theory relies on a set of fundamental concepts and definitions that form the building blocks for understanding and applying this science. These concepts include: the sample space, events, probabilities, basic probability rules, independence, and random variable. Having a clear understanding of these concepts and definitions in crucial for mastering probability theory and its related applications.

In this chapter of the dissertation, we will provide a detailed explanation of the general concepts and basic definitions in probability theory, with relevant examples and applications to enhance understanding. These concepts will be covered in a systematic and organized manner to pave the way for more advanced topics in probability theory and statistics.

I.1 Sample Space

Definition I.1.1 *[33] Sample space it is a set, usually denoted Ω, whose elements correspond to all possible outcomes of the random experiment that we are trying to model. It is also called the observation space.*

Example I.1.1

1. *A coin toss (heads or tails):* $\Omega = \{H, T\}$,
2. *Two coin tosses (heads or tails):* $\Omega = \{HH, HT, TH, TT\}$,
3. *Height of a person:* $\Omega = \mathbb{R}^+$.

I.2 Sigma-Field

Let Ω be a set.

Definition I.2.1 *[24] A distinguished collection $\mathcal{F}$ of subsets of Ω is called a sigma-field if the following axioms are satisfied:*

(i) *If A in $\mathcal{F}$, then $A^c \in \mathcal{F}$ where $A^c := \Omega \cap A^c$ is the complement of $A \in \Omega$,*

(ii) *If $A_1, A_2, \ldots$ is a countable family of sets in $\mathcal{F}$, then their union $\bigcup_{i=1}^{\infty} A_i$ is also in $\mathcal{F}$,*

(iii) $\Omega \in \mathcal{F}$.

Remark I.2.1 *The pair $(\Omega, \mathcal{F})$ is called a measurable space, and the elements of $\mathcal{F}$ are called events or measurable sets.*

I.2.1 Sub-Sigma-Field

Let $\mathcal{F}$ be a sigma-field.

Definition I.2.2 *[25] A sigma-field $\mathcal{G}$ is a sub-sigma-field of $\mathcal{F}$ if $F \in \mathcal{F}$ for every $F \in \mathcal{G}$.*

I.2.2 Generated Sigma-Field

Definition I.2.3 *[34] Let X be a set of subsets of Ω. Then $\sigma(X)$, the sigma-field generated by X, is the smallest sigma-field Σ on Ω such that $X \subseteq \Sigma$. It is the intersection of all sigma-fields on Ω which have X as a subset.*

Example I.2.1 *[34] Let Ω be a topological space. $\mathcal{B}(\Omega)$, is the sigma-field generated by the family of open subsets of Ω.*

I.3 Probability Measure

Definition I.3.1 *[28]A probability measure $\mathbb{P}$ on a measurable $(\Omega, \mathcal{F})$ is a function $\mathbb{P} : \mathcal{F} \to [0, 1]$ such that:*

(i) $\mathbb{P}(\varnothing) = 0, \mathbb{P}(\Omega) = 1;$

(ii) *If $A_1, A_2 \ldots \in \mathcal{F}$ and $\{A_i\}_{i=1}^{\infty}$ is disjoint (i.e. $A_i \cap A_j = \varnothing$ if $i \neq j$), then $\mathbb{P}(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)$.*

I.4 Probability Space

The modern theory of probability stems from the work of A. N. Kolmogorov published in 1933. Kolmogorov associates a random experiment with probability space, $(\Omega, \mathcal{F}, \mathbb{P})$, which consists of:

(i) Ω, a nonempty set, called the sample space, which contains all possible outcomes of some random experiment,

(ii) $\mathcal{F}$, a sigma-field of subsets of Ω,

(iii) $\mathbb{P}$, a probability measure on $(\Omega, \mathcal{F})$. See[5]

Definition I.4.1 *[28] The function $Y : \Omega \to \mathbb{R}^n$ defined on $(\Omega, \mathcal{F}, \mathbb{P})$ is called $\mathcal{F}$-measurable if:*

$$Y^{-1}(B) := \{\omega \in \Omega; Y(\omega) \in B\} \in \mathcal{F}, \text{ for all open sets } B \in \mathbb{R}^n.$$

I.5 Conditional Probability

Definition I.5.1 *[14] Let A and B be two events, and suppose that $\mathbb{P}(B) > 0$. The conditional probability of A given B is defined as:*

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

The conditional probability thus measures the probability of B given that we know that A has occurred.

I.6 Random Variables

The standard situation in the modeling of a random phenomenon is that the quantities of interest, rather than being defined on the underlying probability space, are functions from the probability space to some other (measurable) space. These functions are called random variables. Strictly speaking, one uses the term random variable when they are functions from the probability space to $\mathbb{R}$. If the image is in $\mathbb{R}^n$ for some $n \geq 2$ one talks about n-dimensional random variables or simply random vectors. If the image space is a general abstract one, one talks

about random elements.
Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space.

Definition I.6.1 *[14] A random variable X is a measurable function from the sample space Ω to $\mathbb{R}$;*

$$X : \Omega \to \mathbb{R},$$

that is, the inverse image of any Borel set is $\mathcal{F}$ - measurable:

$$X^{-1}(A) = \{\omega : X(\omega) \in A\} \in \mathcal{F}, \text{ for all } A \in \mathbb{R}.$$

Example I.6.1 *In the experiment of flipping two coin twice, the sample space is the set $\Omega = \{HH, TH, HT, TT\}$.*
Let us consider the sigma-field $\mathcal{F} = \{\varnothing, \{HH, HT\}, \{TT, TH\}, \Omega\}$.
We can see that the function $X : \Omega \to \mathbb{R}$ that is worth 1 if the result is head on the first coin and 0 otherwise is a random variable on $(\Omega, \mathcal{F})$, while the function $Y : \Omega \to \mathbb{R}$ that is worth 1 if the result is head on the second coin and 0 otherwise. Indeed

$$\{\omega \in \Omega, X(\omega) = x\} = \begin{cases} \{HH, HT\} \in \mathcal{F}, & \text{if } x = 1; \\ \{TT, TH\} \in \mathcal{F}, & \text{if } x = 0; \\ \varnothing \in \mathcal{F}, & \text{otherwise.} \end{cases}$$

$$\{\omega \in \Omega, Y(\omega) = x\} = \begin{cases} \{HH, TH\} \notin \mathcal{F}, & \text{if } x = 1; \\ \{HT, TT\} \notin \mathcal{F}, & \text{if } x = 0; \\ \varnothing \in \mathcal{F}, & \text{otherwise.} \end{cases}$$

In this example X is $\mathcal{F}$-measurable while Y is not.

I.6.1 Types of Random Variables

Random variables are considered one of the fundamental concepts in probability theory. These variables are classified into two main categories: discrete variables and continuous variables. The distinguishing property that determines this classification is the set of possible values that the variable can take. Discrete variables take only specific and separate values, while continuous variables can take any value within a certain range of continuous values. This classification implies the difference in the methods used to find the probability distribution functions for each type.

Discrete Random Variables [13]

A random variable that takes values in a finite or countably infinite set is called a discrete random variable. As a function of x, the function $\mathbb{P}(X = x)$ is the probability mass function (pmf) of X. The pmf describes the distribution of a discrete random variable. For $R \subseteq \mathbb{R}$,

$$\mathbb{P}(X \in R) = \sum_{x \in \mathbb{R}} \mathbb{P}(X = x).$$

Examples I.6.1

1. ***Bernoulli Random Variable[13]*** *A Bernoulli random variables takes values 1 and 0, with probabilities p and $1 - p$, respectively. It is common to refer to the dichotomous values of a Bernoulli variable success and failure.*
2. ***Poisson Random Variable[5]*** *A discrete random variable X is said to have a Poisson probability distribution if:*

   $$\mathbb{P}(X = k) = \frac{\lambda^k}{k!} e^{-\lambda},\ k = 0, 1, 2, ...$$

   with $\lambda > 0$ parameter.

Continuous Random Variables [13]

A continuous random variable takes values in an uncountable set, most commonly $\mathbb{R}$, $(0, \infty)$ or (a, b), with $a < b$.
For continuous random variables $\mathbb{P}(X = x) = 0$ for all x, and probabilities are computed by integrating the probability density function. The density function plays a role analogous to the probability mass function for discrete variables for computing probabilities.

Definition I.6.2 *[13] A function f is a probability density function of X if:*

1. *$f(x) \geq 0$, for all x,*
2. *$\int_{-\infty}^{\infty} f(x)dx = 1$.*

Examples I.6.2

1. ***Normal Random Variable[5]*** *A random variable X is said to have a normal distribution if its probability density function is given by:*

   $$\mathbb{P}(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\dfrac{1}{2}\left(\dfrac{x-\mu}{\sigma}\right)^2}$$

   with μ and $\sigma > 0$ constant parameters.

2. ***Expononential Random Variable[29]*** *A continuous random variable whose probability density function is given, for some $\lambda > 0$, by:*

   $$f(x) = \begin{cases} \lambda e^{-\lambda x}, & \text{if } x \geq 0 \\ 0\ , & \text{if } x < 0 \end{cases}$$

   is said to be an exponential random variable with parameter λ.

I.7 Independence

Definition I.7.1 *[4]Two events $A, B \in \mathcal{F}$ are called independent if:*

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B).$$

In general, we say that n events $A_1, ..., A_n \in \mathcal{F}$ are independent if :

$$\mathbb{P}(\bigcap_{i=1}^{n} A_i) = \mathbb{P}(A_1)...\mathbb{P}(A_n).$$

I.8 Mathematical Expectation

The definitions of mathematical expectation are given separately for the two kinds of random variables : discrete and continuous.

I.8.1 Discrete random variables

Definition I.8.1 *[9] Mathematical expectation (mean value) of a discrete random variable X is the number $\mathbb{E}[X]$ defined by:*

$$\mathbb{E}[X] = \sum x\mathbb{P}(x).$$

Example I.8.1 *(Expectation of a Bernoulli Random Variable)*
Calculate $\mathbb{E}[X]$ when X is a Bernoulli random variable with parameter p.
Since $\mathbb{P}(0) = 1 - p$, $\mathbb{P}(1) = p$, we have:

$$\mathbb{E}[X] = 0(1 - p) + 1(p) = p.$$

I.8.2 Continuous random variables

Definition I.8.2 *[9] Mathematical expectation (mean value) of a continuous random variable X is the number $\mathbb{E}[X]$ defined by:*

$$\mathbb{E}(x) = \int_{-\infty}^{\infty} xf(x)dx.$$

Example I.8.2 *(Expectation of an Exponential Random Variable)*
Let X be exponentially distributed with parameter λ. Calculate $\mathbb{E}[X]$.

$$\begin{aligned}\mathbb{E}[X] &= \int_0^\infty x\lambda e^{-\lambda x} dx \\ &= -xe^{-\lambda x}\Big]_0^\infty + \int_0^\infty e^{-\lambda x} dx \\ &= 0 - \frac{e^{-\lambda x}}{\lambda}\Big]_0^\infty \\ &= \frac{1}{\lambda}.\end{aligned}$$

I.9 Variance and Covariance

Definition I.9.1 *[10] Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $X : \Omega \to \mathbb{R}$ be a random variable. Assume that the mathematical expectation $\mathbb{E}[X]$ exists and is finite. The variance is defined as the mean of the square of the difference $X - \mathbb{E}[X]$ i.e.*

$$Var(X) = \mathbb{E}[(X - \mathbb{E}[X])^2].$$

The standard deviation is defined as:

$$\sigma = \sigma(X) = \sqrt{Var(X)}.$$

Definition I.9.2 *[29] The covariance of any two random variables X and Y, denoted by $Cov(X, Y)$, is defined by:*

$$\begin{aligned}Cov(X,Y) &= \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])] \\ &= \mathbb{E}[XY] - Y\mathbb{E}[X] - X\mathbb{E}[Y] + \mathbb{E}[X]\mathbb{E}[Y] \\ &= \mathbb{E}[XY] - \mathbb{E}[Y]\mathbb{E}[X] - \mathbb{E}[X]\mathbb{E}[Y] + \mathbb{E}[X]\mathbb{E}[Y] \\ &= \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y].\end{aligned}$$

I.10 Limits of Sequences of Random Variables

Let $X_1, X_2, ...$ be a sequence of random variables on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. There are several ways of making sense of the limit expression $X = \lim_{n\to\infty} X_n$, some of them will be discussed in the following sections.[16, 23]

I.10.1 Almost Sure Limit

We say that a sequence of random variables $\{X_n\}_{n\geq 1}$ converges almost surely to a random variable X if:

$$\mathbb{P}\{\omega : \lim_{n\to\infty} X_n(\omega) \neq X(\omega)\} = 0.$$

We usually abbreviate almost sure convergence by writing:

$$\lim_{n\to\infty} X_n(\omega) = X(\omega) \; a.s.$$

Example I.10.1 *[23] An urn contains 2 black balls and 2 white balls. At time n a ball is selected at random from the urn, and the color is noted .If the number of balls of this color is greater than or equal to the number of balls of the other color, then the ball is put back in the urn ; otherwise, the ball is left out. let X_n be the number of black balls in the urn after the nth draw. Does this sequence of random variables converge ?*
The first draw is the critical draw. Suppose the first draw is black, then the black ball that is selected will be left out. Thereafter, each time a white ball is selected it will be put back in, and when the remaining black ball is selected it will be left out. Thus with probability one, the black ball will eventually be selected, and X_n will converge to zero. On the other hand, if a white ball is selected in the first draw, then eventually the remaining white ball will be removed, and hence with probability one X_n will convergo to 2. Thus X_n is equally likely to eventually converge to 0 or 2, that is,

$$X_n \to X \; as \; n \to \infty \; almost \; surely.Where;$$

$$\mathbb{P}[X = 0] = \frac{1}{2} = \mathbb{P}[X = 2].$$

I.10.2 Mean Square Limit

The sequance of random variables $\{X_n\}_{n\geq 1}$ converges in the mean square sense to the random variable X if:

$$\mathbb{E}[(X_n - X)^2] \to 0 \; as \; n \to \infty.$$

We usually abbreviate mean square convergence by writing:

$$\lim_{x\to\infty} X_n(\omega) = X(\omega) \; m.s.$$

Example I.10.2 *[5] Consider a sequence X_n of random variables such that there is a constant k with $\mathbb{E}[X_n] \to k$ and $Var(X_n) \to 0$ as $n \to \infty$.*

Show that $ms - \lim_{n\to\infty} X_n = k$.

$$\begin{aligned}\mathbb{E}[(X_n - k)^2] &= \mathbb{E}[X_n^2 - 2kX_n + k^2] \\ &= \mathbb{E}[X_n^2] - 2k\mathbb{E}[X_n] + k^2 \\ &= (\mathbb{E}[X_n^2] - \mathbb{E}[X_n]^2) + (\mathbb{E}[X_n]^2 - 2k\mathbb{E}[X_n] + k^2) \\ &= Var(X_n) + (\mathbb{E}[X_n] - k)^2;\end{aligned}$$

then, $\lim_{n\to\infty} \mathbb{E}[(X_n - k)^2] = 0$ *as* $n \to \infty$. *So,*

$$ms - \lim_{n\to\infty} X_n = k.$$

I.10.3 Central Limit Theorem

Theorem I.10.1 *[23] Let* S_n *be the sum of n iid random variables with finite mean* $\mathbb{E}[X] = \mu$ *and finite variance* σ^2, *and let* Z_n *be the zero-mean, unit-variance random variable defined by*

$$Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}}$$

then;

$$\lim_{n\to\infty} \mathbb{P}[Z_n \le z] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{\frac{-x^2}{2}} dx.$$

II Stochastic Calculus

Stochastic calculus is a branch of mathematics that deals with the study of stochastic processes, which are processes that involve randomness or uncertainty. It is an extension of classical calculus and is particularly useful in fields such as finance, physics, and engineering, where randomness plays a significant role. This chapter will provide an introduction to stochastic calculus. We will begin by defining stochastic processes and discussing their basic properties. We will then introduce the concepts of stochastic integration and stochastic differentiation. It will also cover stochastic differential equations, which are used to model various random phenomena.

II.1 Stochastic Processes

II.1.1 Deterministic and Stochastic Models

Probability theory, the mathematical science of uncertainty, plays an ever growing role in how we understand the world around us. The word "stochastic" comes from the Greek stokhazesthai, which means to aim at, or guess at. A stochastic process, also called a random process, is simply one in which outcomes are uncertain. By contrast, in a deterministic system there is no randomness. In a deterministic system, the same output is always produced from a given input. Functions and differential equations are typically used to describe deterministic processes. Random variables and probability distributions are the building blocks for stochastic systems. See [13]
The first use of the term "stochastic" in probability theory can be traced back to the Russian economist and statistician Ladislaus Bortkiewicz (1868- 1931). In his paper "Die Iterationen" published in 1917, he defines this term as follows: "The investigation of empirical varieties, which is based on probability theory, and, therefore, on the law of the large numbers, may be denoted as stochastic. But stochastic is not simply probability theory, but above all probability theory and its applications".

Definition II.1.1 *[31] A stochastic process is a collection of random variables* $\{X(t), t \in T\}$ *where* t *is a parameter that runs over an index set* T*. In general we call* t *the time-*

parameter (or simply the time) and $T \subseteq \mathbb{R}$*. Each* $X(t)$ *takes values in some set* $S \subseteq \mathbb{R}$ *called the state space, then* $X(t)$ *is the state of the process at time* t*. Stochastic processes are characterized by three principal properties: the state space, the parameter set and the dependence relations between the various random variables X(t). The parameter may be discrete, for example,* $T = \mathbb{Z}^+$ *or* $T = \mathbb{Z}$*, in which case we customarily use n to denote the time and call* $X(n)$ *a discrete-time process. When* T *is an interval in the real line, typically* $T = [0, \infty)$*, then* $X(t)$ *is called a continuous-time process.*

Remarks II.1.1

(i) *Stochastic processes are also often called random processes;*

(ii) *The function defined on the index set T and taking values in* $\mathbb{R}$:

$$t \to X_t(\omega),$$

is called the sample path (or the realization, or the trajectory) of the stochastic process X corresponding to the outcome ω*;*

(iii) *The state space S is called a discrete-state process if it is discrete, often referred to as a chain;*

(iv) $\{X(t), t \in T\}$ *is a discrete-time process if the index set T of the random process is discrete. It is also called a random sequence and is denoted by* $\{X(n), n = 1, 2, ...\}$ *or* $\{X_n, n = 1, 2, ...\}$.

II.2 Examples of Stochastic Processes

II.2.1 Bernoulli Process

One of the simplest stochastic processes is the Bernoulli process, which is a sequence of independent and identically distributed (iid) random variables, where each random variable takes either the value 1 *or* 0*, say* 1 *with probability p and* 0 *with probability* $1 - p$*. This process can be linked to repeatedly flipping a coin, where the probability of obtaining a head is p and its value is* 1*, while the value of a tail is* 0*. In other words, a Bernoulli process is a sequence of iid Bernoulli random variables, where each coin flip is an example of a Bernoulli trial.*

II.2.2 Markov chain

Markov chain represents a class of stochastic processes in which the future does not depend on the past but only on the present. The algorithm was first proposed by a Russian mathematician Andrei Markov. It has come to find its place in various practical applications like stock markets, weather prediction, spread of influenzas, ...

Definition II.2.1 *[32] A random process $\{X_n\}_{n\geq 0}$, where each X_n takes values in a finite set S, is called a Markov chain if:*

$$\mathbb{P}(X_{n+1} = x_{n+1} | X_n = x_n, X_{n-1} = x_{n-1}, ..., X_0 = x_0) = \mathbb{P}(X_{n+1} = x_{n+1} | X_n = x_n),$$

for all $n \geq 0$ and $x_0, x_1, ..., x_{n+1} \in S$.

Transition probabilities [17]

For a finite Markov chain with m states $E_1, E_2, ..., E_m$, introduce the notation

$$p_{ij} = \mathbb{P}(X_{n+1} = j | X_n = i),$$

where $i, j = 1, 2, ..., m$ to represent the probability of a transition from state E_i to E_j. The numbers p_{ij} are known as the transition probabilities of the chain, and must satisfy:

$$p_{ij} \geq 0 \text{ and } \sum_{j=1}^{m} p_{ij} = 1,$$

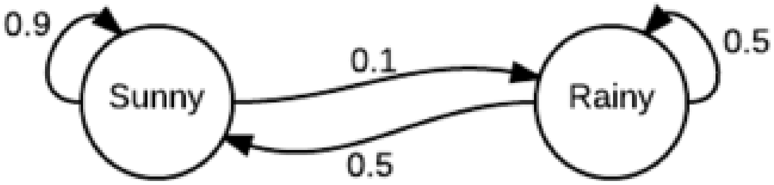
for each $i = 1, 2, ..., m$.
Transition probabilities form an $m \times m$ array which can be assembled into a transition matrix T, where:

$$T = [p_{ij}] = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1m} \\ p_{21} & p_{22} & \cdots & p_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ p_{m1} & p_{m2} & \cdots & p_{mm} \end{pmatrix}.$$

Remark II.2.1 *Note that i can equal j, so that transitions within the same state may be possible.*

Example II.2.1 *(A simple weather model)*
The probabilities of weather conditions (modeled as either rainy or sunny), given the weather on the preceding day, can be represented by a transition matrix:

$$P = \begin{pmatrix} 0.9 & 0.1 \\ 0.5 & 0.5 \end{pmatrix}.$$



Transition diagram

The matrix P represents the weather model in which a sunny day is 90% likely to be followed by another sunny day, and a rainy day is 50% likely to be followed by another

rainy day. The columns can be labelled "sunny" and "rainy", and the rows can be labelled in the same order. p_{ij} *is the probability that, if a given day is of type i, it will be followed by a day of type j. Notice that the rows of P sum to* 1*: this is because P is a stochastic matrix.*
The weather on day 0 *(today) is known to be sunny. This is represented by an initial state vector in which the "sunny" entry is* 100%*, and the "rainy" entry is* 0%*:*

$$x^{(0)} = \begin{pmatrix} 1 & 0 \end{pmatrix}.$$

The weather on day 1 *(tomorrow) can be predicted by multiplying the state vector from day* 0 *by the transition matrix:*

$$x^{(1)} = x^{(0)}P = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} 0.9 & 0.1 \\ 0.5 & 0.5 \end{pmatrix} = \begin{pmatrix} 0.9 & 0.1 \end{pmatrix},$$

Thus, there is a 90% *chance that day* 1 *will also be sunny. The weather on day* 2 *(the day after tomorrow) can be predicted in the same way, from the state vector we computed for day* 1*:*

$$x^{(2)} = x^{(1)}P = x^{(0)}P^2 = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} 0.9 & 0.1 \\ 0.5 & 0.5 \end{pmatrix}^2 = \begin{pmatrix} 0.86 & 0.14 \end{pmatrix},$$

or,

$$x^{(2)} = x^{(1)}P = \begin{pmatrix} 0.9 & 0.1 \end{pmatrix} \begin{pmatrix} 0.9 & 0.1 \\ 0.5 & 0.5 \end{pmatrix} = \begin{pmatrix} 0.86 & 0.14 \end{pmatrix}.$$

General rules for day n are:

$$x^{(n)} = x^{(n-1)}P \text{ or } x(n) = x^{(0)}P^n.$$

II.2.3 Martingales

The concept of a martingale has its origin in gambling, namely, it describes a fair game of chance. Similarly, the notions of submartingale and supermartingale defined below are related to favourable and unfavourable games of chance. Some aspects of gambling are inherent in the mathematics of finance, in particular, the theory of financial derivatives such as options. Not surprisingly, martingales play a crucial role there. In fact, martingales reach well beyond game theory and appear in various areas of modern probability and stochastic analysis. First of all, let us introduce some basic definitions and properties.See[4]

Filtration and Adapted Process

Let $(\Omega, \mathcal{F}, \mathbb{P})$ *a probability space.*

Definition II.2.2 *[21] A filtration* $\{\mathcal{F}_t,\ t \geq 0\}$ *is an increasing family of sub-sigma-fields of* $\mathcal{F}$ *indexed by* $t \geq 0$*, i.e., for each* $s,\ t \geq 0$ *such that* $s < t$*, we have* $\mathcal{F}_s \subset \mathcal{F}_t$ *with* $\mathcal{F}_0 = \{\Omega, \varnothing\}$*. To each process* $\{X(t),\ t \geq 0\}$ *and for each t, we can associate a sigma-field denoted by* $\mathcal{F}_t = \sigma(X(s),\ 0 \leq s \leq t)$*, which is the sigma-field generated by*

the process X up to time t. It is the smallest set of subsets of Ω that makes it possible to assign probabilities to all the events related to the process X up to time t.
Given a stochastic process $\{X_t,\ t \geq 0\}$ and a filtration $\{\mathcal{F}_t,\ t \geq 0\}$ (not necessarily the one generated by X), the process X is said to be adapted to $\{\mathcal{F}_t,\ t \geq 0\}$ if for every $t \geq 0,\ X(t)$ is $\mathcal{F}_t$-measurable.
We also say that $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ is a filtered probability space.

Remark II.2.2 *We also introduce $\mathcal{F}_\infty = \sigma\{\bigcup_{n=0}^{\infty} \mathcal{F}_n\}$. Recall that the union of a sequence of sigma-fields is not necessarily a sigma-field.*

Example II.2.2 *[4] For a sequence $X_1, X_2, ..., X_n$ of coin tosses we take $\mathcal{F}_n$ to be the sigma-field generated by $X_1, X_2, ..., X_n$,*

$$\mathcal{F}_n = \sigma(X_1, X_2, ..., X_n).$$

Let,

$$A = \{the\ first\ 5\ tosses\ produce\ at\ least\ 2\ heads\}.$$

At discrete time $n = 5$, i.e. once the coin has heen tossed five times, it will be possible to decide whether A has occurred or not. This means that $A \in \mathcal{F}_5$. However, at $n = 4$ it is not always possible to tell if A has occurred or not. If the outcomes of the first four tosses are, say,

$$tail, tail, head, tail,$$

then, the event A remains undecided. We will have to toss the coin once more to see what happens. Therefore $A \notin \mathcal{F}_4$.
This example illustrates another relevant issue. Suppose that the outcomes of the first four coin tosses are:

$$tail, head, tail, head.$$

In this case it is possible to tell that A has occurred already at $n = 4$, whatever the outcome of the fifth toss will be. It does not mean, however, that A belongs to $\mathcal{F}_4$. The point is that for A to belong $\mathcal{F}_4$ it must be possible to tell whether A has occurred or not after the first four tosses, no matter what the first four outcomes are.

Conditional Expectation

Conditional expectation is a crucial tool in the study of stochastic processes. It is therefore important to develop the necessary intuition behind this notion, the definition of which may appear some what abstract at first.

Definition II.2.3 *[4] Let X be an integrable random variable on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and let $\mathcal{G}$ be a σ-field contained in $\mathcal{F}$. Then conditional expectation of X given $\mathcal{G}$ is defined to be a random variable $\mathbb{E}(X|\mathcal{G})$ such that:*

1. *$\mathbb{E}(X|\mathcal{G})$ is $\mathcal{G}$-measurable,*

2. *For any $A \in \mathcal{G}$:*

$$\int_A \mathbb{E}(X|\mathcal{G})d\mathbb{P} = \int_A Xd\mathbb{P}.$$

Proposition II.2.1 *[4] Conditional expectation has the following properties:*

1) $\mathbb{E}[aX + bY|\mathcal{G}] = a\mathbb{E}[X|\mathcal{G}] + b\mathbb{E}[Y|\mathcal{G}]$*(Linearity);*
2) $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]] = \mathbb{E}(X)$;
3) $\mathbb{E}[XY|\mathcal{G}] = X\mathbb{E}[Y|\mathcal{G}]$ *if X is $\mathcal{G}$-measurable (Taking out what is known);*
4) $\mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X]$ *if X is independent of $\mathcal{G}$ (An independent condition drops out);*
5) $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]|\mathcal{H}] = \mathbb{E}[X|\mathcal{H}]$ *if $\mathcal{H} \subset \mathcal{G}$ (Tower property);*
6) *If $X \geq 0$, then $\mathbb{E}[X|\mathcal{G}] \geq 0$ (Positivity);*
7) *(Jensen's inequality) If $g(x)$ is a convex function on I, that is, for all $x, y \in I$ and $\lambda \in (0,1)$:*

   $$g(\lambda x + (1-\lambda)y) \leq \lambda g(x) + (1-\lambda)g(y),$$

   and X is a random variable with range I, then

   $$g(\mathbb{E}[X|\mathcal{G}]) \leq \mathbb{E}[g(X)|\mathcal{G}].$$

   In particular, with $g(x) = |x|$:

   $$|\mathbb{E}[X|\mathcal{G}]| \leq \mathbb{E}[|X||\mathcal{G}].$$

Here a, b are arbitrary real numbers, X, Y are integrable random variables on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and $\mathcal{G}, \mathcal{H}$ sigma-fields on Ω contained in $\mathcal{F}$. In 3 we also assume that the product XY is integrable.

Martingale, Supermartingale, Submartingale

Definition II.2.4 *[34] A process X is called a martingale (relative to $(\{\mathcal{F}_n\}, \mathbb{P})$) if:*

(i) *X is adapted;*
(ii) $\mathbb{E}(| X_n |) < \infty, \forall n$;
(iii) $\mathbb{E}(X_n|\mathcal{F}_{n-1}) = X_{n-1}$ *a.s.* $(n \geq 1)$.

A supermartingale is defined similarly, except that (iii) is replaced by:

$$\mathbb{E}(X_n|\mathcal{F}_{n-1}) \leq X_{n-1} \ a.s. (n \geq 1),$$

and a submartingale is defined with (iii) replaced by:

$$\mathbb{E}(X_n|\mathcal{F}_{n-1}) \geq X_{n-1} \ a.s. (n \geq 1).$$

A supermartingale decreases on average and a submartingale increases on average.

Example II.2.3 *Let X be a random variable with $\mathbb{E}[|X|] < \infty$. Suppose $(\mathcal{F}_n)_n$ is a filtration, i.e. $\mathcal{F}_n \subset \mathcal{F}_{n+1}, \forall n \geq 0$. Define $X_n = \mathbb{E}[X|\mathcal{F}_n]$, then $\{X_0, X_1, ...\}$ is a martingale, namely Doob martingale, with respect to filtration $(\mathcal{F}_n)_n$.*
To see this, note that:

1) $X_n = \mathbb{E}[X|\mathcal{F}_n]$ *is* $\mathcal{F}_n$*-measurable. (According to Kolmogorov (1933), the conditional expectation of X given a sigma-field $\mathcal{G}$, is itself a random variable. More precisely, $\mathbb{E}[X|\mathcal{G}]$ is a $\mathcal{G}$-measurable);*
2) $\mathbb{E}[|X_n|] = \mathbb{E}[|\mathbb{E}[X|\mathcal{F}_n]|] \leq \mathbb{E}[\mathbb{E}[|X||\mathcal{F}_n]] = \mathbb{E}[|X|] < \infty$. *(According to Jensen's inequality);*
3) $\mathbb{E}[X_{n+1}|\mathcal{F}_n] = \mathbb{E}[\mathbb{E}[X|\mathcal{F}_{n+1}]|\mathcal{F}_n] = \mathbb{E}[X|\mathcal{F}_n] = X_n$. *(As $\mathcal{F}_n \subset \mathcal{F}_{n+1}$).*

Stopping Time

In a game of chance, a stopping time is a time at which the player decides to stop playing, according to a criterion that depends only on the past and present. For example, he may decides to stop playing as soon as he has spent all of his capital, as soon as he has won a certain amount of money, as soon as he has won a certain number of times in a row, or according to any combination of these criteria. Stopping times therefore have two important properties: they are random, since they depend on the previous course of the game, and they cannot depend on the future, since the player must be able to decide at any time whether or not to stop.

Definition II.2.5 *[8] Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let $\{\mathcal{F}_k\}_{k=0}^n$ be a filtration. A stopping time is a random variable $\tau : \Omega \to \{0, 1, 2, ..., n\} \cup \{\infty\}$ with the property that :*

$$\{\omega \in \Omega, \tau(\omega) = k\} \in \mathcal{F}_k, \forall k = 0, 1, 2, ..., n, \infty.$$

Example II.2.4 *Suppose that $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$ and $T = \{0, 1, 2, 3\}$. The stochastic process $X = \{X_t,\ t \in \{0, 1, 2, 3\}\}$ represents the evolution of stock price, $X_t =$ stock as soon as the stock market is closed at the t^{th} day, instant $t = 0$ representing today.*

$\omega \setminus X_t(\omega)$	$X_0(\omega)$	$X_1(\omega)$	$X_2(\omega)$	$X_3(\omega)$
ω_1	1	0.5	1	0.5
ω_2	1	0.5	1	0.5
ω_3	1	2	1	1
ω_4	1	2	2	2

We had determined that the filtration containing the information revealed by the process at each moment is :

$$\mathcal{F}_0 = \sigma\{X_0\} = \{\varnothing, \Omega\};$$

$$\mathcal{F}_1 = \sigma\{X_0, X_1\} = \sigma\{\{\omega_1, \omega_2\}, \{\omega_3, \omega_4\}\};$$

$$\mathcal{F}_2 = \sigma\{X_0, X_1, X_2\} = \sigma\{\{\omega_1, \omega_2\}, \{\omega_3\}, \{\omega_4\}\};$$

$$\mathcal{F}_3 = \sigma\{X_0, X_1, X_2, X_3\} = \sigma\{\{\omega_1\}, \{\omega_2\}, \{\omega_3\}, \{\omega_4\}\}.$$

a) We are not selling our shares today ($t = 0$) *but we will sell then as soon as the price is greater than or equal to 1.*

- *The random time representing this situation is:* $\tau(\omega_1) = 2, \tau(\omega_2) = 2, \tau(\omega_3) = 1, \tau(\omega_4) = 1$.
  This random variable is indeed a stopping time since:

$$\{\omega \in \Omega; \tau(\omega) = 0\} = \varnothing \in \mathcal{F}_0;$$

$$\{\omega \in \Omega; \tau(\omega) = 1\} = \{\omega_3, \omega_4\} \in \mathcal{F}_1;$$

$$\{\omega \in \Omega; \tau(\omega) = 2\} = \{\omega_1, \omega_2\} \in \mathcal{F}_2;$$

$$\{\omega \in \Omega; \tau(\omega) = 3\} = \varnothing \in \mathcal{F}_3.$$

b) Now consider random time τ^* *modeling the following situation: we will buy shares as soon as we are able to make a profit later.*

- *This random variable takes the values:* $\tau^*(\omega_1) = 1, \tau^*(\omega_2) = 1, \tau^*(\omega_3) = 0, \tau^*(\omega_4) = 0$ *is not a stopping time since:*

$$\{\omega \in \Omega; \tau^*(\omega) = 0\} = \{\omega_3, \omega_4\} \notin \mathcal{F}_0.$$

Proposition II.2.2 *Properties of stopping times[22]*

- *Let* T *be a stopping time. Then* T *is* $\mathcal{F}_t$*-measurable;*
- *Let* S *and* T *be two stopping times such that* $S \leq T$. *Then,* $\mathcal{F}_S \subset \mathcal{F}_T$;
- *Let* S *and* T *be two stopping times. Then,* $S \wedge T$ *and* $S \vee T$ *are also stopping times.*

II.2.4 Brownian Motion

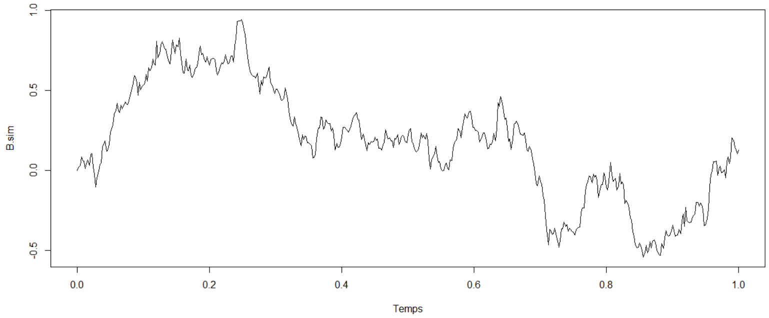
The term Brownian motion comes from the name of the botanist R. Brown, who described the irregular motion of minute particles suspended in water, while the water itself remained seemingly still. It is now known that this motion is due to the cumulative effect of water molecules hitting the particle at various angles. The rigorous definition and the first mathematical proof of the existence of Brownian motion are due to N. Wiener, who studied Brownian motion in the 1920s, almost a century after it was observed by R. Brown. Wiener process is another term for Brownian motion, both terms being used equally often. Brownian motion and more general diffusion processes are extremely important in physics, economics, finance, and many branches of mathematics beyond probability theory.

Definition II.2.6 *[6] A Brownian motion process is a stochastic process* $\{B_t, t \geq 0\}$, *which satisfies:*

1. *The process starts at the origin,* $B_0 = 0$;
2. B_t *has independent increments, i.e.* $\forall n, \forall t_i, 0 \leq t_0 \leq t_1 \leq ... \leq t_n$, *the random variables* $B_{t_n} - B_{t_{n-1}}, ..., B_{t_1} - B_{t_0}, B_{t_0}$ *are independent;*
3. *The process* B_t *is continuous in* t;

4. *The increments $B_t - B_s$ are normally distributed with mean zero and variance $t - s$,*

$$B_t - B_s \sim N(0, t - s).$$



Continuous-time Brownian motion.

Remarks II.2.1

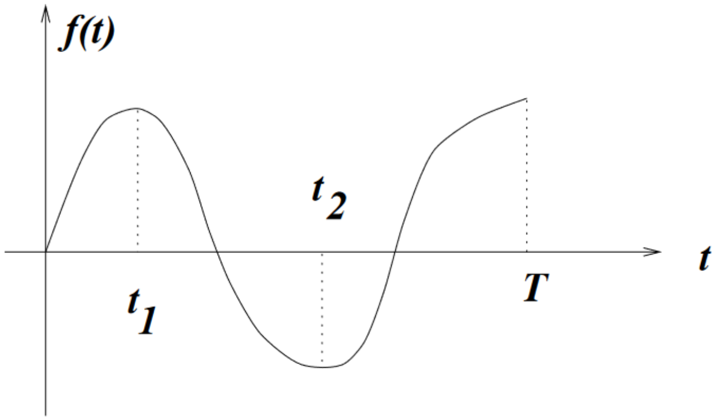
(i) From condition 4 we get that B_t is normally distributed with mean $\mathbb{E}[B_t] = 0$ and $Var[B_t] = t$,

$$B_t \sim N(0, t);$$

(ii) It is worth noting that even if B_t is continuous, it is nowhere differentiable.

Quadratic variation

Quadratic variation is a measure of volatility. First we will consider first variation, $FV(f)$, of a function $f(t)$.



Example function f(t)

For the function pictured, the first variation over the interval $[0, T]$ is given by:

$$\begin{aligned} FV_{[0,T]}(f) &= [f(t_1) - f(t_0)] - [f(t_2) - f(t_1)] + [f(T) - f(t_2)] \\ &= \int_0^{t_1} f^{'}(t)dt + \int_{t_1}^{t_2} (-f^{'}(t))dt + \int_{t_2}^{T} f^{'}(t)dt \\ &= \int_0^{T} |f^{'}(t)|dt. \end{aligned}$$

Thus, first variation measures the total amount of up and down motion of the path. The general definition of first variation is as follows:

Definition II.2.7 *[8](First variation) Let $\Pi = \{t_0, t_1, ..., t_n\}$ be a partition of [0, t], i.e.,*

$$0 = t_0 \leq t_1 \leq ... \leq t_n = t.$$

The mesh of the partition is defined to be :

$$\|\Pi\| = \max_k (t_{k+1} - t_k).$$

We then define :

$$FV_{[0,t]}(f) = \lim_{\|\Pi\| \to 0} \sum_{k=0}^{n-1} |f(t_{k+1}) - f(t_k)|.$$

Definition II.2.8 *[8](Quadratic variation) The quadratic variation of a function f on an interval $[0, t]$ is :*

$$Q_{[0,t]}(f) = \lim_{\|\Pi\| \to 0} \sum_{k=0}^{n-1} (f(t_{k+1}) - f(t_k))^2.$$

Proposition II.2.3 *With probability 1, the paths of Brownian motion $B(t)$ are not of bounded variation :*

$$\mathbb{P}(FV_{[0,t]}(B) = \lim_{\|\Pi\| \to 0} \sum_{k=0}^{n-1} |B(t_{k+1}) - B(t_k)| = \infty) = 1,$$

for all fixed t > 0.

We will prove Proposition (II.2.3) *in the next section after we introduce the quadratic variation of a Brownian motion, and prove that it has finite non-zero quadratic variation.*

Proposition II.2.4 *[8] With probability 1, the paths of Brownian motion $B(t)$ satisfy :*

$$Q_{[0,t]}(B) = \lim_{\|\Pi\| \to 0} \sum_{k=0}^{n-1} (B(t_{k+1}) - B(t_k))^2 = t.$$

Proof II.2.1 *[8] Let $\Pi = \{t_0, t_1, ..., t_n\}$ be a partition of $[0, t]$. To simplify notation, set $D_k = B(t_{k+1}) - B(t_k)$. Define the sample quadratic variation :*

$$Q_\Pi = \sum_{k=0}^{n-1} D_k^2.$$

Then,

$$Q_\Pi - t = \sum_{k=0}^{n-1} [D_k^2 - (t_{k+1} - t_k)].$$

We want to show that :

$$\lim_{\|\Pi\| \to 0} (Q_\Pi - t) = 0.$$

Consider an individual summand,

$$D_k^2 - (t_{k+1} - t_k) = [B(t_{k+1}) - B(t_k)]^2 - (t_{k+1} - t_k).$$

This has expectation 0, so,

$$\mathbb{E}(Q_\Pi - t) = \mathbb{E} \sum_{k=0}^{n-1} [D_k^2 - (t_{k+1} - t_k)] = 0.$$

For $j \neq k$, the terms,

$$D_j^2 - (t_{j+1} - t_j) \;\; and \;\; D_k^2 - (t_{k+1} - t_k)$$

are independent, so,

$$\begin{aligned} Var(Q_\Pi - t) &= \sum_{k=0}^{n-1} Var[D_k^2 - (t_{k+1} - t_k)] \\ &= \sum_{k=0}^{n-1} \mathbb{E}[D_K^4 - 2(t_{k+1} - t_k) D_k^2 + (t_{k+1} - t_k)^2] \\ &= \sum_{k=0}^{n-1} [3(t_{k+1} - t_k)^2 - 2(t_{k+1} - t_k)^2 + (t_{k+1} - t_k)^2] \\ &= \sum_{k=0}^{n-1} 2(t_{k+1} - t_k)^2 \\ &\leq 2\|\Pi\| \sum_{k=0}^{n-1} (t_{k+1} - t_k) \\ &= 2\|\Pi\| t. \end{aligned}$$

Thus we have,

$$\mathbb{E}(Q_\Pi - t) = 0,$$
$$Var(Q_\Pi - t) \leq 2\|\Pi\| t.$$

As $\|\Pi\| \to 0, Var(Q_\Pi - t) \to 0$, *so,*

$$\lim_{\|\Pi\| \to 0} (Q_\Pi - t) = 0.$$

□

Proof of Proposition(II.2.3): $FV_{[0,t]}(B) = \infty$, *w.p.1, for all* $t > 0$.

Proof II.2.2

$$\begin{aligned} Q_\Pi(B)[0,t] &= \sum_{k=0}^{n-1} (B(t_{k+1}) - B(t_k))^2 \\ &\leq \max_k |(B(t_{k+1}) - B(t_k))| \sum_{k=0}^{n-1} |(B(t_{k+1}) - B(t_k))| \\ &= MFV_\Pi(B)[0,t]. \end{aligned}$$

Where $M = M(\Pi) = \max_k |(B(t_{k+1}) - B(t_k))| \to 0$, *w.p.1, as* $\|\Pi\| \to 0$ *since the paths of Brownian motion are continuous, hence uniformly continuous over any interval* $[0,t]$. *If* $FV_{[0,t]}(B) \leq \infty$, *w.p.1, for all* $t > 0$, *then we would conclude that* $MFV_\Pi(B)[0,t] \to 0$, *w.p.1, as* $\|\Pi\| \to 0$ *and hence conclude that* $Q_{[0,t]}(B) = 0$, *w.p.1, for all* $t > 0$.
Hence indeed $FV_{[0,t]}(B) = \infty$, *w.p.1, for all* $t > 0$. □

II.3 Stochastic Integration

This section deals with one of the most useful stochastic integrals of the form $\int_0^t X_s dB_s$, *where* B_t *is Brownian motion and* X_t *is a stochastic process. This type of integral called the Itô integral which was introduced in 1944 by the Japanese mathematician K. Itô.*

II.3.1 The Itô Integral

We previously found in proposition (II.2.3) *that Brownian motion paths have infinite variation, therefore we cannot apply the Riemann-Stieltjes theory to them. This led to the concept of Itô's integral, which allows us to study the random nature of Brownian motion. So, we can say that Riemann's integral deals with simple processes with finite variation, while Itô's integral deals with stochastic processes with infinite variation.*
Consider $0 \leq a < b$ *and let* $X_t = f(t, B_t)$ *be a nonanticipating process with*

$$\mathbb{E}\Big[\int_a^b X_t^2 dt\Big] < \infty.$$

Divide the interval $[a, b]$ into n subintervals using the partition points

$$a = t_0 < t_1 < ... < t_{n-1} < t_n = b,$$

and consider the partial sums

$$S_n = \sum_{i=0}^{n-1} X_{t_i}(B_{t_{i+1}} - B_{t_i}).$$

We emphasize that the intermediate points are the left endpoints of each interval, and this is the way they should be always chosen. Since the process X_t is nonanticipative, the random variables X_{t_i} and $B_{t_{i+1}} - B_{t_i}$ are independent, this is an important feature in the definition of the Itô integral.
The Itô integral is the limit of the partial sums S_n

$$ms - \lim_{n\to\infty} S_n = \int_a^b X_t dB_t,$$

provided the limit exists. It can be shown that the choice of partition does not influence the value of the Itô integral. This is the reason why, for practical purposes, it suffices to assume the intervals equidistant, i.e.

$$t_{i+1} - t_i = \frac{(b-a)}{n}, i = 0, 1, ..., n-1.$$

The previous convergence is in the mean square sense, i.e.

$$\lim_{n\to\infty} \mathbb{E}\Big[\Big(S_n - \int_a^b X_t dB_t\Big)^2\Big] = 0.$$

Remark II.3.1 *A process X_t is called a nonanticipating process if X_t is independent of any future increment $B_s - B_t$ for any t and s with $t < s$.[5]*

Existence of the Itô integral[5]

It is known that the Itô stochastic integral $\int_a^b X_t dB_t$ exists if the process $X_t = f(t, B_t)$ satisfies the following properties:

(i) The paths $t \to X_t(\omega)$ are continuous on $[a, b]$ for any state of the world $\omega \in \Omega$;

(ii) The process X_t is nonanticipating for $t \in [a, b]$;

(iii) $\mathbb{E}\Big[\int_a^b X_t^2 dt\Big] < \infty$.

Example II.3.1 *[5](The case $X_t = c$, constant)*
In this case the partial sums can be computed as follows:

$$S_n = \sum_{i=0}^{n-1} X_{t_i}(B_{t_{i+1}} - B_{t_i}) = \sum_{i=0}^{n-1} c(B_{t_{i+1}} - B_{t_i}) = c(B_b - B_a),$$

and since the answer does not depend on n, we have:

$$\int_a^b c dB_t = c(B_b - B_a).$$

In particular, taking $c = 1, a = 0$, and $b = T$, since the Brownian motion starts at 0, we have the following formula:

$$\int_0^T dB_t = B_T.$$

Example II.3.2 *[5](The case $X_t = B_t$)*
We shall integrate the process B_t between 0 and T. Considering an equidistant partition, we take $t_i = \dfrac{iT}{n}, i = 0, 1, ..., n-1$. The partial sums are given by:

$$\sum_{i=0}^{n-1} B_{t_i}(B_{t_{i+1}} - B_{t_i}).$$

Since

$$xy = \frac{1}{2}\Big[(x+y)^2 - x^2 - y^2\Big],$$

letting $x = B_{t_i}$ and $y = B_{t_{i+1}} - B_{t_i}$ yields:

$$B_{t_i}(B_{t_{i+1}} - B_{t_i}) = \frac{1}{2}\Big[B^2_{t_{i+1}} - B^2_{t_i} - (B_{t_{i+1}} - B_{t_i})^2\Big].$$

Then after pair cancelations the sum becomes:

$$\begin{aligned} S_n &= \frac{1}{2}\sum_{i=0}^{n-1} B^2_{t_{i+1}} - \frac{1}{2}\sum_{i=0}^{n-1} B^2_{t_i} - \frac{1}{2}\sum_{i=0}^{n-1}(B_{t_{i+1}} - B_{t_i})^2 \\ &= \frac{1}{2}B^2_{t_n} - \frac{1}{2}\sum_{i=0}^{n-1}(B_{t_{i+1}} - B_{t_i})^2. \end{aligned}$$

Using $t_n = T$, we get:

$$S_n = \frac{1}{2}B^2_T - \frac{1}{2}\sum_{i=0}^{n-1}(B_{t_{i+1}} - B_{t_i})^2.$$

Since the first term on the right side is independent of n, using Proposition(II.2.4), *we have:*

$$\begin{aligned} ms - \lim_{n\to\infty} S_n &= \frac{1}{2}B^2_T - \Big[ms - \lim_{n\to\infty}\frac{1}{2}\sum_{i=0}^{n-1}(B_{t_{i+1}} - B_{t_i})^2\Big] \\ &= \frac{1}{2}B^2_T - \frac{1}{2}T. \end{aligned}$$

We have now obtained the following explicit formula of a stochastic integral:

$$\int_0^T B_t dB_t = \frac{1}{2}B_T^2 - \frac{1}{2}T.$$

In a similar way one can obtain:

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

Properties of the Itô Integral

Some of the most important properties of the Itô integral are given by the following result:

Proposition II.3.1 *[5] Let $f(t, B_t), g(t, B_t)$ be nonanticipating processes and $c \in \mathbb{R}$. Then we have:*

1. *Additivity:*

$$\int_0^T [f(t,B_t) + g(t,B_t)]dB_t = \int_0^T f(t,B_t)dB_t + \int_0^T g(t,B_t)dB_t.$$

2. *Homogeneity:*

$$\int_0^T cf(t,B_t)dB_t = c\int_0^T f(t,B_t)dB_t.$$

3. *Partition property:*

$$\int_0^T f(t,B_t)dB_t = \int_0^u f(t,B_t)dB_t + \int_u^T f(t,B_t)dB_t, \forall\, 0 < u < T.$$

4. *Predictability:* $\int_0^T f(t,B_t)dB_t$ *is* $\mathcal{F}_T$*-measurable.*
5. *Zero mean:*

$$\mathbb{E}\Big[\int_a^b f(t,B_t)dB_t\Big] = 0.$$

6. *Isometry:*

$$\mathbb{E}\Big[\Big(\int_a^b f(t,B_t)dB_t\Big)^2\Big] = \mathbb{E}\Big[\int_a^b f(t,B_t)^2 dt\Big].$$

7. *Covariance:*

$$\mathbb{E}\Big[\Big(\int_a^b f(t,B_t)dB_t\Big)\Big(\int_a^b g(t,B_t)dB_t\Big)\Big] = \mathbb{E}\Big[\int_a^b f(t,B_t)g(t,B_t)dt\Big].$$

8. *The Itô integral* $\int_0^t f(s,B_s)dB_s$ *is* $\mathcal{F}_t$*-martingale.*

II.3.2 Itô's Formula

In order to explain Itô's formula, we begin by recalling Taylor's theorem. That is, if $f : \mathbb{R} \to \mathbb{R}$ *is infinitely differentiable, then* f *can be expressed as an infinite polynomial expanded around* $a \in \mathbb{R}$ *as:[19]*

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \frac{f'''(a)}{3!}(x-a)^3 + ...$$

We now let $x = t + \Delta t$ *and* $a = t$*, so that:*

$$f(t+\Delta t) = f(t) + f'(t)(\Delta t) + \frac{f''(t)}{2!}(\Delta t)^2 + \frac{f'''(t)}{3!}(\Delta t)^3 + ...$$

which we can write as:

$$\frac{f(t+\Delta t) - f(t)}{\Delta t} = f'(t) + \frac{f''(t)}{2!}\Delta t + \frac{f'''(t)}{3!}(\Delta t)^2 + ...$$

At this point we see that if $\Delta t \to 0$*, then:*

$$\lim_{\Delta t \to 0} \frac{f(t+\Delta t) - f(t)}{\Delta t} = \lim_{\Delta t \to 0} \left[f'(t) + \frac{f''(t)}{2!}\Delta t + \frac{f'''(t)}{3!}(\Delta t)^2 + ... \right] = f'(t),$$

which is exactly the definition of derivative.
The same argument can be used to prove the chain rule. That is, suppose that f *and* g *are infinitely differentiable. Let* $x = g(t) + \Delta g(t)$ *and* $a = g(t)$*, so that Taylor's theorem takes the form:*

$$f(g(t)+\Delta g(t)) = f(g(t)) + f'(g(t))\Delta g(t) + \frac{f''(g(t))}{2!}(\Delta g(t))^2 + \frac{f'''(g(t))}{3!}(\Delta g(t))^3 + ...$$

We now write $\Delta g(t) = g(t+\Delta t) - g(t)$*, so that:*

$$\begin{aligned} f(g(t+\Delta t)) - f(g(t)) = f'(g(t)) \cdot (g(t+\Delta t) - g(t)) &+ \frac{f''(g(t))}{2!}(g(t+\Delta t) - g(t))^2 \\ &+ \frac{f'''(g(t))}{3!}(g(t+\Delta t) - g(t))^3 + ... \end{aligned}$$

Dividing both sides by Δt *implies:*

$$\begin{aligned} \frac{f(g(t+\Delta t)) - f(g(t))}{\Delta t} = f'(g(t)) \cdot \frac{g(t+\Delta t) - g(t)}{\Delta t} &+ \frac{f''(g(t))}{2!} \cdot \frac{(g(t+\Delta t) - g(t))^2}{\Delta t} \\ &+ \frac{f'''(g(t))}{3!} \cdot \frac{(g(t+\Delta t) - g(t))^3}{\Delta t} + ... \end{aligned}$$

When $\Delta t \to 0$*, notice that the limit of the left-side of the previous equation is:*

$$\lim_{\Delta t \to 0} \frac{f(g(t+\Delta t)) - f(g(t))}{\Delta t} = \lim_{\Delta t \to 0} \frac{(f \circ g)(t+\Delta t) - (f \circ g)(t)}{\Delta t} = \frac{d}{dt}(f \circ g)(t).$$

As for the right-side, we find for the first term that:

$$\lim_{\Delta t\to 0}\Big[f'(g(t))\frac{g(t+\Delta t)-g(t)}{\Delta t}\Big] = f'(g(t))\cdot \lim_{\Delta t\to 0}\frac{g(t+\Delta t)-g(t)}{\Delta t} = f'(g(t))\cdot g'(t).$$

For the second term, however, we have:

$$\begin{aligned}\lim_{\Delta t\to 0}\Big[\frac{f''(g(t))}{2!}\cdot\frac{(g(t+\Delta t)-g(t))^2}{\Delta t}\Big] &= \frac{f''(g(t))}{2!}\cdot \lim_{\Delta t\to 0}\frac{g(t+\Delta t)-g(t)}{\Delta t}\lim_{\Delta t\to 0}[g(t+\Delta t)-g(t)]\\ &= \frac{f''(g(t))}{2!}\cdot g'(t)\cdot 0 = 0.\end{aligned}$$

which follows since g is differentiable (and therefore continuous). Similarly, the higher order terms all approach 0 in the $\Delta t \to 0$ limit. Combining everything gives:

$$\frac{d}{dt}(f\circ g)(t) = f'(g(t))\cdot g'(t).$$

In fact, this proof of the chain rule illustrates precisely why the fundamental theorem of calculus fails for Itô integrals. Brownian motion is nowhere differentiable, and so the step of the proof of the chain rule where the second order term vanishes as Δt is not valid. Indeed, if we take $g(t) = B_t$ and divide by Δt, then we find:

$$\begin{aligned}\frac{\Delta f(B_t)}{\Delta t} &= \frac{f(B_{t+\Delta t}) - f(B(t))}{\Delta t}\\ &= f'(B_t)\cdot\frac{\Delta B_t}{\Delta t} + \frac{f''(B(t))}{2!}\cdot\frac{(\Delta B(t))^2}{\Delta t} + \frac{f'''(B(t))}{3!}\cdot\frac{(\Delta B(t))^3}{\Delta t} + ...\end{aligned}$$

In the limit as $\Delta t \to 0$, the left-side of the previous equation is:

$$\lim_{\Delta t\to 0}\frac{\Delta f(B(t))}{\Delta t} = \frac{d}{dt}f(B_t).$$

As for the right-side, we are tempted to say that the first term approaches:

$$\lim_{\Delta t\to 0}\Big[f'(B_t)\cdot\frac{\Delta B_t}{\Delta t}\Big] = f'(B_t)\cdot\frac{dB_t}{dt},$$

so that:

$$\frac{d}{dt}f(B_t) = f'(B_t)\cdot\frac{dB_t}{dt} + \frac{f''(B_t)}{2!}\cdot\Big[\lim_{\Delta t\to 0}\frac{(\Delta B_t)^2}{\Delta t}\Big] + \frac{f'''(B_t)}{3!}\cdot\Big[\lim_{\Delta t\to 0}\frac{(\Delta B_t)^3}{\Delta t}\Big] + ...$$

(Even though Brownian motion is nowhere differentiable so that $\frac{dB_t}{dt}$ does not exist, bear with us.) We know that $\Delta B_t = B_{t+\Delta t} - B_t \sim \mathcal{N}(0,\Delta t)$, so that:

$$Var(\Delta B_t) = \mathbb{E}[(\Delta B_t)^2] = \Delta t,$$

or, approximately,

$$(\Delta B_t)^2 \approx \Delta t.$$

This suggests that:

$$\lim_{\Delta t\to 0}\frac{(\Delta B_t)^2}{\Delta t}=\lim_{\Delta t\to 0}\frac{\Delta t}{\Delta t}=1,$$

but if $k\geq 3$, then

$$\lim_{\Delta t\to 0}\frac{(\Delta B_t)^k}{\Delta t}=\lim_{\Delta t\to 0}\frac{(\sqrt{\Delta t})^k}{\Delta t}=\lim_{\Delta t\to 0}\frac{\Delta t(\sqrt{\Delta t})^{k-2}}{\Delta t}=0.$$

Hence, we conclude that:

$$\frac{d}{dt}f(B_t)=f'(B_t)\cdot\frac{dB_t}{dt}+\frac{f''(B_t)}{2!}.$$

Multiplying through by dt gives:

$$df(B_t)=f'(B_t)dB_t+\frac{f''(B_t)}{2}dt.$$

We shall present a general formula for the stochastic environment. In this case, the Brownian motion is replaced by a stochastic process X_t. The composition between the differentiable function f and the process X_t is denoted by $F_t=f(X_t)$, so:

$$dF_t=f'(X_t)dX_t+\frac{1}{2}f''(X_t)(dX_t)^2. \tag{II.1}$$

In the computation of dX_t we may take into the account stochastic relations such as: $(dB_t)^2=dt$, or $dtdB_t=0$.[5]

Itô's formula for diffusions

The previous formula is a general case of Itô's formula. However, in most cases the increments dX_t are given by some particular relations. An important case is when the increment is given by:

$$dX_t=b_tdt+\sigma_tdB_t.$$

A process X_t satisfying this relation is called an Itô diffusion.

Theorem II.3.1 (Itô's formula, first version) *If X_t is an Itô diffusion, and $F_t=f(X_t)$, then:*

$$dF_t=\Big[b_tf'(X_t)+\frac{\sigma_t^2}{2}f''(X_t)\Big]dt+\sigma_tf'(X_t)dB_t.$$

Proof II.3.1 *We shall provide an informal proof. Using relations $dB_t^2=dt$ and $dt^2=dB_tdt=0$, we have:*

$$\begin{aligned}(dX_t)^2&=(b_tdt+\sigma_tdB_t)^2\\&=b_t^2dt^2+2b_t\sigma_tdB_tdt+\sigma_t^2dB_t^2\\&=\sigma_t^2dt.\end{aligned}$$

Substituting into (II.1) *yields*

$$\begin{aligned} dF_t &= f'(X_t)dX_t + \frac{1}{2}f''(X_t)(dX_t)^2 \\ &= f'(X_t)(b_t dt + \sigma_t dB_t) + \frac{1}{2}f''(X_t)\sigma_t^2 dt \\ &= \Big[b_t f'(X_t) + \frac{\sigma_t^2}{2}f''(X_t)\Big]dt + \sigma_t f'(X_t)dB_t. \end{aligned}$$

☐

Example II.3.3 *Let $X_t = B_t$. If $f(x) = e^{kx}$, with k constant, then*

$$f'(x) = ke^{kx} \text{ and } f''(x) = k^2 e^{kx}.$$

Therefore,

$$d(e^{kB_t}) = ke^{kB_t}dB_t + \frac{1}{2}k^2 e^{kB_t}dt.$$

Example II.3.4 *Let $X_t = B_t$. If $f(x) = x^\alpha$, with α constant, then*

$$f'(x) = \alpha x^{\alpha-1} \text{ and } f''(x) = \alpha(\alpha-1)x^{\alpha-2}.$$

Then,

$$d[(B_t)^\alpha] = \alpha(B_t)^{\alpha-1}dB_t + \frac{1}{2}\alpha(\alpha-1)(B_t)^{\alpha-2}dt.$$

Theorem II.3.2 (Itô's formula, second version) *Let $f(t,x)$ be a smooth function of two variables, and let X_t be a stochastic process satisfying $dX_t = b_t dt + \sigma_t dB_t$ for a Brownian motion B_t. Then,*

$$df(t, X_t) = \Big[\frac{\partial f}{\partial t} + b_t\frac{\partial f}{\partial X_t} + \frac{1}{2}\sigma_t^2\frac{\partial^2 f}{\partial (X_t)^2}\Big]dt + \sigma_t\frac{\partial f}{\partial X_t}dB_t.$$

Proof II.3.2 *We have*

$$df(t, X_t) = \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial X_t}dX_t + \frac{1}{2}\Big[\frac{\partial^2 f}{\partial X_t^2}dX_t^2 + \frac{\partial^2 f}{\partial t^2}dt^2 + 2\frac{\partial^2 f}{\partial X_t \partial t}d_t dX_t\Big] + ...$$

We can ignore the terms $d_t dX_t$ and $(d_t)^2$

$$\begin{aligned} df(t, X_t) &= \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial X_t}dX_t + \frac{1}{2}\frac{\partial^2 f}{\partial X_t^2}(dX_t)^2 \\ &= \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial X_t}(b_t dt + \sigma_t dB_t) + \frac{1}{2}\frac{\partial^2 f}{\partial X_t^2}(\sigma_t^2 dt) \\ &= \Big[\frac{\partial f}{\partial t} + b_t\frac{\partial f}{\partial X_t} + \frac{1}{2}\sigma_t^2\frac{\partial^2 f}{\partial X_t^2}\Big]dt + \sigma_t\frac{\partial f}{\partial X_t}dB_t. \end{aligned}$$

☐

Example II.3.5 *Let* $f(t,X_t) = f(t,B_t)$*,i.e.* $b_t = 0$ *and* $\sigma_t = 1$*. If* $f(t,x) = e^{\alpha t+\beta x}$*, then:*

$$\frac{\partial f}{\partial t} = \alpha f, \frac{\partial f}{\partial x} = \beta f, \frac{\partial^2 f}{\partial x^2} = \beta^2 f,$$

so,

$$\begin{aligned} df(t,B_t) &= \left[\alpha f + \frac{1}{2}\beta^2 f\right] dt + \beta f dB_t \\ &= \left(\alpha + \frac{1}{2}\beta^2\right) f dt + \beta f dB_t. \end{aligned}$$

II.4 Stochastic Differential Equations

Stochastic differential equations (SDEs), are a type of differential equations that include a random component. This makes them useful for modeling phenomena that are subject to uncertainty, such as stock prices, weather patterns and the behavior of biological systems. Before we dive into SDEs, let's briefly review ordinary differential equations (ODEs).
If $x(t)$ *is a differentiable function defined for* $t \geq 0, \mu(x,t)$ *is a function of* x *and* t *and the following relation is satisfied for all* $t, 0 \leq t \leq T$*:*

$$\frac{dx(t)}{dt} = x'(t) = \mu\Big(x(t),t\Big) \text{ and } x(0) = x_0;$$

then $x(t)$ *is a solution of the ODE with the initial condition* x_0*. Usually, the requirement that* $x'(t)$ *is continuous is added. The above equation can be written in other forms as follows:*

$$dx(t) = \mu\Big(x(t),t\Big)dt$$

and (by continuity of $x'(t)$*):*

$$x(t) = x(0) + \int_0^t \mu\Big(x(s),s\Big)ds.$$

II.4.1 Stochastic Differential Equations

Let $\{B_t, t \geq 0\}$*, be a Brownian motion process. An equation of the form*

$$dX_t = \mu(t,X_t)dt + \sigma(t,X_t)dB_t, \tag{II.2}$$

where functions $\mu(t,x)$ *and* $\sigma(t,x)$ *are given and* X_t *is the unknown process, is called a stochastic differential equation (SDE) driven by Brownian motion. The functions* $\mu(t,x)$ *and* $\sigma(t,x)$ *are called respectively the drift and the diffusion coefficient.*

A process X_t is called a solution of the SDE (II.2) *if for all $t \geq 0$ the integrals $\int_0^t \mu(s, X_s)ds$ and $\int_0^t \sigma(s, X_s)dB_s$ exist, with the second being an Itô integral and*

$$X_t = X(0) + \int_0^t \mu(s, X_s)ds + \int_0^t \sigma(s, X_s)dB_s.$$

Remark II.4.1 *When $\sigma = 0$, the SDE becomes an ODE.*

II.4.2 Existence and Uniqueness of Solutions

Let X_t satisfy

$$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dB_t, \ \ X_0 = c. \tag{II.3}$$

Theorem II.4.1 (Existence and Uniqueness) *[5] If the following conditions are satisfied:*

1. *Coefficients satisfying the Lipschitz condition in x, that is, there is a constant $K > 0$, such that for all $x, y \in \mathbb{R}$ and all $0 \leq t \leq T$:*

   $$|\mu(t, x) - \mu(t, y)| + |\sigma(t, x) - \sigma(t, y)| \leq K|x - y|,$$

2. *Coefficients satisfying the linear growth condition in x, that is, there exists a constant $C > 0$ such that:*

   $$|\mu(t, x)| + |\sigma(t, x)| \leq C(1 + |x|), \forall x \in \mathbb{R}, 0 \leq t \leq T.$$

Then there exists a unique solution X_t of the SDE (II.3)*. X_t has continuous paths, moreover,*

$$\mathbb{E}\Big[\int_0^T X_t^2 dt\Big] < \infty.$$

II.4.3 Examples of Solution Methods

1. ***Exact Stochastic Equations****[5]*
   The stochastic differential equation

   $$dX_t = \mu(t, B_t)dt + \sigma(t, B_t)dB_t \tag{II.4}$$

   is called exact if there is a differentiable function $f(t, x)$ such that

   $$\mu(t, x) = \frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}, \tag{II.5}$$

   $$\sigma(t, x) = \frac{\partial f}{\partial x}. \tag{II.6}$$

Assume the equation is exact. Then substituting in (II.4) *yields*

$$dX_t = \Big(\frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}\Big)dt + \frac{\partial f}{\partial x}dB_t.$$

Applying Itô's formula, the previous equations becomes

$$dX_t = d\Big(f(t, B_t)\Big),$$

which implies $X_t = f(t, B_t) + c$, *with* c *constant.*
Solving the partial differential equations system (II.5) *and* (II.6) *requires the following steps:*

1. *Integrating partially with respect to* x *in the second equation to obtain* $f(t,x)$ *up to an additive function* $g(t)$,
2. *Substitute into the first equation and determine the function* $g(t)$,
3. *The solution is* $X_t = f(t, B_t) + c$, *with* c *determined from the initial condition on* X_t.

Example II.4.1 *[5] Find the solution of*

$$dX_t = (2tB_t^3 + 3t^2(1 + B_t))dt + (3t^2B_t^2 + 1)dB_t, X_0 = 0.$$

The coefficient functions are $\mu(t,x) = 2tx^3 + 3t^2(1+x)$ *and* $\sigma(t,x) = 3t^2x^2 + 1$. *The associated system is given by*

$$2tx^3 + 3t^2(1+x) = \frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial^2 f}{\partial x^2},$$

$$3t^2x^2 + 1 = \frac{\partial f}{\partial x}.$$

Integrating partially in the second equation yields

$$f(t,x) = \int (3t^2x^2 + 1)dx = t^2x^3 + x + g(t).$$

Then $\dfrac{\partial f}{\partial t} = 2tx^3 + g'(t)$ *and* $\dfrac{\partial^2 f}{\partial x^2} = 6t^2x$, *and substituting into the first equation we get*

$$2tx^3 + 3t^2(1+x) = 2tx^3 + g'(t) + \frac{1}{2}6t^2x.$$

After cancellations we get $g'(t) = 3t^2$, *so* $g(t) = t^3 + c$. *Then*

$$f(t,x) = t^2x^3 + x + t^3 + c = t^2(x^3 + t) + x + c.$$

The solution process is given by $X_t = f(t, B_t) = t^2(B_t^3 + t) + B_t + c$. *Using* $X_0 = 0$ *we get* $c = 0$. *Hence the solution is* $X_t = t^2(B_t^3 + t) + B_t$.

2. ***Linear Stochastic Differential Equations****[5]*

Consider the stochastic differential equation with drift term linear in X_t:

$$dX_t = (\alpha(t)X_t + \beta(t))dt + \sigma(t, B_t)dB_t, t \geq 0.$$

This also can be written as

$$dX_t - \alpha(t)X_t dt = \beta(t)dt + \sigma(t, B_t)dB_t.$$

Let $A(t) = \int_0^t \alpha(s)ds$. Multiplying by the integrating factor $e^{-A(t)}$, the left side of the previous equation becomes an exact expression

$$e^{-A(t)}\Big(dX_t - \alpha(t)X_t dt\Big) = e^{-A(t)}\beta(t)dt + e^{-A(t)}\sigma(t, B_t)dB_t,$$

$$d(e^{-A(t)}X_t) = e^{-A(t)}\beta(t)dt + e^{-A(t)}\sigma(t, B_t)dB_t.$$

Integrating yields

$$e^{-A(t)}X_t = X_0 + \int_0^t e^{-A(s)}\beta(s)ds + \int_0^t e^{-A(s)}\sigma(s, B_s)dB_s,$$

$$X_t = X_0 e^{A(t)} + e^{A(t)}\Big(\int_0^t e^{-A(s)}\beta(s)ds + \int_0^t e^{-A(s)}\sigma(s, B_s)dB_s\Big).$$

The first integral within the previous parentheses is a Riemann integral, and the latter one is an Itô stochastic integral.

Example II.4.2 *[5] Solve the linear stochastic differential equation*

$$dX_t = (2X_t + 1)dt + e^{2t}dB_t.$$

Write the equation as

$$dX_t - 2X_t dt = dt + e^{2t}dB_t$$

and multiply by the integrating factor e^{-2t} to get

$$d(e^{-2t}X_t) = e^{-2t}dt + dB_t.$$

Integrate between 0 and t and multiply by e^{2t} and obtain

$$X_t = X_0 e^{2t} + e^{2t}\int_0^t e^{-2s}ds + e^{2t}\int_0^t dB_s = X_0 e^{2t} + \frac{1}{2}(e^{2t} - 1) + e^{2t}B_t.$$

3. ***Integration by Inspection***
   When solving a stochastic differential equation by inspection we look for opportunities to apply the product formula:

   $$d(f(t)Y_t) = f(t)dY_t + Y_t df(t).$$

   For instance, if a stochastic differential equation can be written as

   $$dX_t = f'(t)B_t dt + f(t)dB_t,$$

the product rule brings the equation into the exact form

$$dX_t = d\Big(f(t)B_t\Big),$$

which after integration leads to the solution

$$X_t = X_0 + f(t)B_t.$$

Example II.4.3 *Solve the stochastic differential equation*

$$dX_t = (B_t + 3t^2)dt + tdB_t.$$

If we rewrite the equation as

$$dX_t = 3t^2dt + (B_tdt + tdB_t),$$

we note the exact expression formed by the last two terms $B_tdt + tdB_t = d(tB_t)$. *Then*

$$dX_t = d(t^3) + d(tB_t),$$

which is equivalent with $d(X_t) = d(t^3 + tB_t)$. *Hence* $X_t = t^3 + tB_t + c,\ c \in \mathbb{R}$.

4. ***The Method of Variation of Parameters***
Let's start by considering the following stochastic equation

$$dX_t = \alpha X_t dB_t,$$

with α *constant. This is the equation which, in physics is known to model the linear noise. Dividing by* X_t *yields*

$$\frac{dX_t}{X_t} = \alpha dB_t.$$

Switch to the integral form

$$\int \frac{dX_t}{X_t} = \int \alpha dB_t,$$

and integrate "blindly" to get $\ln X_t = \alpha B_t + c$, *with c integration constant. This leads to the "pseudo-solution"*

$$X_t = e^{\alpha B_t + c}.$$

The nomination "pseudo" stands for the fact that X_t *does not satisfy the initial equation. We shall find a correct solution by letting the parameter c be a function of t. In other words, we are looking for a solution of the following type:*

$$X_t = e^{\alpha B_t + c(t)},$$

where the function $c(t)$ *is subject to be determined. Using Ito's formula we get*

$$dX_t = d(e^{\alpha B_t + c(t)}) = e^{\alpha B_t + c(t)}\Big(c'(t) + \frac{\alpha^2}{2}\Big)dt + \alpha e^{\alpha B_t + c(t)} dB_t$$
$$= X_t\Big(c'(t) + \frac{\alpha^2}{2}\Big)dt + \alpha X_t dB_t.$$

Substituting the last term from the initial equation yields

$$dX_t = X_t\Big(c'(t) + \frac{\alpha^2}{2}\Big)dt + dX_t,$$

which leads to the equation

$$c'(t) + \frac{\alpha^2}{2} = 0$$

with the solution $c(t) = -\frac{\alpha^2}{2}t + k$*. Subsitituting into yieldes*

$$X_t = e^{\alpha B_t - \frac{\alpha^2}{2}t + k}.$$

The value of the constant k *is determined by taking* $t = 0$*. This leads to* $X_0 = e^k$ *. Hence we have obtained the solution of the equation*

$$X_t = X_0 e^{\alpha B_t - \frac{\alpha^2}{2}t}.$$

Example II.4.4 *Use the method of variation of parameters to solve the equation*

$$dX_t = X_t B_t dB_t.$$

Dividing by X_t *and converting the differential equation into the equivalent integral form, we get*

$$\int \frac{1}{X_t} dX_t = \int B_t dB_t.$$

The right side is a well-known stochastic integral given by

$$\int B_t dB_t = \frac{B_t^2}{2} - \frac{t}{2} + C.$$

The left side will be integrated "blindly" according to the rules of elementary Calculus

$$\int \frac{1}{X_t} dX_t = \ln X_t + C.$$

Equating the last two relations and solving for X_t we obtain the "pseudo-solution"

$$X_t = e^{\frac{B_t^2}{2} - \frac{t}{2} + c},$$

with c constant. In order to get a correct solution, we let c to depend on t and B_t. We shall assume that $c(t, B_t) = a(t) + b(B_t)$, so we are looking for a solution of the form

$$X_t = e^{\frac{B_t^2}{2} - \frac{t}{2} + a(t) + b(B_t)}.$$

Applying Ito's formula, we have

$$dX_t = X_t\Big[-\frac{1}{2} + a'(t) + \frac{1}{2}(1 + b''(B_t))\Big]dt + X_t(B_t + b'(B_t))dB_t.$$

Subtracting the initial equation $dX_t = X_t B_t dB_t$ yields

$$0 = X_t\Big(a'(t) + \frac{1}{2}b''(B_t)\Big)dt + X_t b'(B_t)dB_t.$$

This equation is satisfied if we are able to choose the functions $a(t)$ and $b(B_t)$ such that the coefficients of dt and dB_t vanish

$$b'(B_t) = 0 \text{ and } a'(t) + \frac{1}{2}b''(B_t) = 0.$$

From the first equation b must be a constant. Substituting into the second equation it follows that a is also a constant. It turns out that the aforementioned"pseudo-solution" is in fact a solution. The constant $c = a + b$ is obtained letting $t = 0$. Hence the solution is given by

$$X_t = X_0 e^{\frac{B_t^2}{2} - \frac{t}{2}}.$$

III Stochastic Differential Inclusions

In our present era replete with comlexities, dynamic systems characterized by uncertainty, nonlinearity, and random behavior have become a widespread and ubiquitous phenomenon. Whether representing the fluctuating financial market dynamics, intricate biochemical interacions, or advanced industrial control systems, these systems do not readily conform to traditional mathematical modeling using ordinary differential equations or even simple stochastic differential equations.
As a result of these challenges, the theory of stochastic differential inclusions has emerged as an advanced mathematical approach that provides a more comprehensive and flexible framework for understanding and modeling such complex system. The strength of this theory lies in its ability to handle a multitude of possible solution for the system at any given point in time, rather than being confined to a singel, determined solution, allowing it to adapt better to the random fluctuations, abrupt changes, and uncertainty prevalent in many real-word systems. Moreover, this theory enables the study of irregular and discontinuous system trajectories, which may be impossible to annlyze using traditional models.
This chapter will provide an introduction to SDI_s*. We will start by giving some definitions of multivalued analysis, we will then explain how we can go from differential equation to differential inclusion and we conclude by an application.*

III.1 Some Definitions of Multivalued Analysis

The multivalued analysis is an important foundation in the theory of differential inclusions. From the word of "multivalued", it is obvious that the mapping maps a variable to a set. Hence, the image of one variable may have more than one elements. In order to distinguish from the multivalued mapping, the mapping defined where the image of a variable only has one element, is often called by single-valued mapping.

Definition III.1.1 *[2](Set-Valued Maps)*

Let X and Y be two sets. A set-valued map F from X to Y is a map that associates with any $x \in X$ a subset $F(x)$ of Y. The subsets $F(x)$ are called the images or the values of F. The subset

$$Dom(F) = \{x \in X | F(x) \neq \emptyset\}$$

is called the domain of F.
When $Dom(F) = X$, we say that the map F is stirct. It is very convenient to characterize a set-valued map by its graph: the graph of F is the subset of pairs (x, y) where $y \in F(x)$:

$$Graph(F) = \{(x, y) \in X \times Y | y \in F(x)\}.$$

Example III.1.1 *$F : \mathbb{R} \to \mathcal{P}(\mathbb{R})$ is a multivalued function defined by:*

$$F(x) = \begin{cases} \{1\}, \ x > 0 \\ \{-1, 1\}, x = 0 \\ \{-1\}, \ x < 0 \end{cases}$$

Example III.1.2 *$F : \mathbb{R} \to \mathcal{P}(\mathbb{R})$ is a multivalued function defined by:*

$$F(x) = \begin{cases} \{x + 1\}, \ x > 0 \\ [-1, 1], x = 0 \\ \{x - 1\}, \ x < 0 \end{cases}$$

Definition III.1.2 *[7](Selection)*
Let $F : X \to \mathcal{P}(Y)$. A function $f : X \to Y$ will be said to be a selection of F if $f(t) \in F(t)$ for every t.

Definition III.1.3 *[12] Let $F : X \to \mathcal{P}(Y)$ be a multi-function. We say that F is upper semi-continuous (u.s.c) at $x_0 \in X$ if for every open set U of Y with $F(x_0) \subseteq U$, there exists an open neighborhood V of x_0 such that for all $x \in V$, we have $F(x) \subseteq U$. Or equivalently, if for every closed set $A \subset Y$, the set*

$$F^{-1}(A) = \{x \in X : F(x) \cap A \neq \emptyset\}$$

is closed in X.

Definition III.1.4 *[12] Let $F : X \to \mathcal{P}(Y)$ be a multi-function. We say that F is lower semi-continuous (l.s.c) at the point $x_0 \in X$ if the set*

$$\{x \in X : F(x) \cap U \neq \emptyset\}$$

is open for every open set $U \in Y$. Or equivalently, if for every open set $A \subset Y$, the set $F^{-1}(A)$ is an open set in X.

Definition III.1.5 *[12] A multivalued function $F : X \to \mathcal{P}(Y)$ is said to be continuous if it is both lower semi-continuous and upper semi-continuous.*

Definition III.1.6 *[18](Set-valued stochastic process)*
A family of set-valued random variables taking values in $cl(\mathbb{R}^n)$ is called a set-valued stochastic process.

III.2 From Differential Equations To Differential Inclusions

The differential equation is a relation of the kind: $x'(t) = f(t, x(t))$, where f is a single-valued mapping. If the function $f(t, x(t))$ in the right side is replaced by a set-valued mapping F, the previous equation becomes: $x'(t) \in F(t, x(t))$, this relation is called differential inclusion.

Definition III.2.1 *A differential inclusion is a differential equation with a set-valued right-hand side. In a sufficiently general form, the differential inclusion can be written as $x'(t) \in F(t, x(t))$, where $t \in \mathbb{R}$ is the time parameter, $x(t)$ is a function defined on a certain subset of $\mathbb{R}$ with values in $\mathbb{R}^n$, and $F = F(t, x)$ is a set-valued map that associates to every $t \in \mathbb{R}$ and $x \in \mathbb{R}^n$ a nonempty closed subset $F(t, x) \subset \mathbb{R}^n$.*

From the definition of the differential inclusion, one may suggest a way to solve it. Firstly, we try to find a selection $f(t, x(t)) \in F(t, x(t))$, then solve the Cauchy problem of the differential equation:

$$x'(t) = f(t, x(t)), \ \ x(0) = x_0.$$

It is really a way that we used to prove many conclusions such as the existence of solutions; however, the scheme is almost impracticable to serve to solve the differential inclusions. At first, finding a satisfactory selection with an explicit description borders on a fantasy. Secondly, even if a simple selection has been obtained, we still lack a method to solve the Cauchy problem except some very special cases. Thirdly, even if a solution of the Cauchy problem is obtained, it still cannot analyze the properties of all solutions. Therefore such a way has been abandoned to solve the previous differential inclusion by almost all researchers. There are two main differences from the investigation of the differential equation. At first, the solution of DI is not unique; hence, we usually do not investigate its uniqueness of solution except some special form of DIs. Secondly, the solution $x(t)$ of a differential equation is required to be differentiable, but for the DI, the solution $x(t)$ is only required to be absolutely continuous. When $x(t)$ is not differntiable, the left side of DI becomes meaningless.[15]
Like DIs, stochastic differential inclusions represent an important generalization of the notion of stochastic differential equations. In the case of an SDE, one wants to find a stochastic process $x = x(t)$, whose stochastic differential $dX(t)$ is given by an equation:

$$dx(t) = \mu(t, x(t))dt + \sigma(t, x(t))dB_t,$$

with a drift term μ and diffusion term σ. The following SDI will be investigated with a multivalued coefficients μ and σ:

$$\begin{cases} dx(t) \in \mu(t, x(t))dt + \sigma(t, x(t))dB_t \\ x(0) = x_0. \end{cases}$$

Now we will present a set of definitions that we were unable to exploit in the example provided later due to time constraints, especially since we are dealing with a difficult topic. However, it is a good tool for us and researchers to use in results of existence and uniqueness of solutions to stochastic differential inclusions, upon which the most famous existence theory in stochastic processes is based, which we will present later.

Definition III.2.2 *[15](Convex set)*
$A \subset X$ is a convex set in X, if for any two points $x_1, x_2 \in A$, the segment $[x_1, x_2] \subset A$.

Definition III.2.3 *[15](Compact set)*
A set $A \subset X$ is said to be compact if for an open covering $\{O_a\}$, there exists a finite set $\{O_1, O_2, ..., O_n\} \subset \{O_a\}$ such that $A \subset O_1 \cup O_1 \cup O_2 \cup ... \cup O_n$.

Definition III.2.4 *[15](Closed set)*
A set $A \subset X$ is said to be closed if its complement A^c is an open set.

Theorem III.2.1 *[3] Let Y be a nonempty, closed convex subset of a Banach space X. If $F : Y \to 2^Y$ is such that*

- *a) $F(v)$ is nonempty and convex for each $v \in Y$;*
- *b) the graph of F, $graph(F) \subset Y \times Y$ is closed;*
- *c) $\cup\{F(v), v \in Y\}$ is contained in a sequentially compact set $C \subset X$, then the set-valued map F has a fixed point, that is, there exists a $v_0 \in Y$ such that $v_0 \in F(v_0)$.*

III.3 Application

In this section, we present an attempt to draft a paper about existence of solution for Stochastic nonlocal random functional integral inclusion, that we have sent to a specialized journal and we are looking forward for a positive response.

Existence of mild solution for stochastic nonlocal random functional integral inclusion

FERRAG AZOUZ, BOUZENOUNT ABIR, AFIFI HOUDA

Abstract

In this article, we use Banach's fixed point theorem to estabilish sufficient conditions which garantee the existence of the solution for a functional non local stochastic differential inclusions.

Key words: *Stochastic, existence, Banach's fixed point theorem stochastic integral inclusion.*

Introduction

Random differential and integral inclusions play important role in characterizing many social, physical, biological and engineering problem. Stochastic differential inclusions are important from the viewpoint of application, since they incorpo- rate(natural) randomness into the mathematical description of the phenomena and provide a more accurate description of them. Stochastic differential inclusions called multivalued stochastic differential equations because the drift and the diffusion terms are set-valued stochastic processes. SDI_s *have been studied by Kree [20], Ahmed [1], Da Prato, Iannelli and Tubaro [11], Michta [[26] [27]] and others.*
In this paper, we show that a solution to an stochastic differential inclusion exists under some conditions with the help of Banach's fixed point theorem. The paper is organized as follows: some preliminaries are presented, then we investigate the existence of solutions for SDI_s *by using Banach's fixed point theorem.*

Preliminaries and Main Result

Throughout this paper, we assume that $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in[0,T]}, \mathbb{P})$ *is a complete filtered probability space satisfying the usual conditions (i.e.* $\{\mathcal{F}_t\}_{t\in[0,T]}$ *is an increasing and right continuous family of sub-sigma-algebras of* $\mathcal{F}$ *and* $\mathcal{F}_0$ *contains all P-null sets). Let* $L^2(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{R}^d)$ *stand for the space of all* $\mathbb{R}^d$*-valued random variables* $\{X(t),\ t \in [0,T]\}$ *such that*

$$\mathbf{E}|X|^2 = \int_\Omega |X|^2 \, d\mathbb{P} < \infty.$$

For $X \in L^2(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{R}^d)$, *we let*

$$\|X\|_2 := \Big(\int_\Omega |X|^2 \, d\mathbb{P} \Big)^{1/2}.$$

Then $L^2(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{R}^d)$ *is a Hilbert space equipped with the norm* $\|\cdot\|_2$. *We consider the following nonlocal functional stochastic differential inclusion*

$$\begin{aligned} dx(t) \in & F(t, x(t), A(t)x(t))dt + G(t, x(t), C(t)x(t))dB(t),\ t \in [0,T] \\ x(0) + & \sum_{k=1}^{p} c_k x(t_k) = x_0 \end{aligned} \tag{III.1}$$

where $0 = t_0 < t_1 < ... < t_p \leq T$, c_k *are constants* $(k = 1, ..., p)$, $p \in \mathbb{N}$, $\{B_t\}$ *is a standard Brownian motion defined on the complete probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *adapted to the filtration* $\{\mathcal{F}_t\}_{t\in[0,T]}$. $x(0)$ *is an* $\mathcal{F}_0$*-mesurable random variable independent of* B *with finite second moment.*
$A(t), t \in [0,T]$ *and* $C(t), t \in [0,T]$ *are families of linear bounded operators defined on* $Y := \mathcal{C}([0,T], L^2(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{R}^d))$ *the space of all continuous stochastic processes defined from* $[0,T]$ *into* $L^2(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{R}^d)$. *The set-valued stochastic processes* F *and* G *are defined on* $[0,T] \times Y \times Y$ *with values in the space* $L^2(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{R}^d)$.
$A(t)$ *and* $C(t)$ *are families of bounded operators. So, there is a function* $\alpha(t) : [0,T] \to \mathbb{R}^+$ *and a function* $\gamma(t) : [0,T] \to \mathbb{R}^+$ *such that* $\|A(t)x\|_Y \leq \alpha(t)\|x\|_Y$ *and* $\|C(t)x\|_Y \leq \gamma(t)\|x\|_Y$.

For the nonlocal stochastic differential inclusion III.1, we have the following result. It shows that under some sufficient conditions, there exists a solution.

Definition III.3.1 *A solution to stochastic differential inclusion III.1 is a continuous and nonanticipating stochastic process* $x : [0, T] \times \Omega \to \mathbb{R}^d$ *which has a representation*

$$x(t) = x_0 + \int_0^t f(s, x, A(s)x(s)ds + \int_0^t g(s, x, C(s)x(s)dB_s,$$

where $f(t, x(t), A(t)x(t)) \in F(t, x(t), A(t)x(t))$ *and* $g(t, x(t), C(t)x(t)) \in G(t, x(t), C(t)x(t))$.

Theorem III.3.1 *Assume that the following conditions holds*

(i) For all $x, y \in Y$ *and* $t \in [0, T]$, $F(t, x, y)$ *and* $G(t, x, y)$ *are nonanticipating set-valued processes;*

(ii) For all $x, y, x', y' \in Y$ *and* $t \in [0, T]$, *there exists a constant* $L > 0$ *such that*

$$\|F(t, x, y) - F(t, x', y')\|^2_{\mathbb{R}^d} \leq L\|y - y'\|^2_{\mathbb{R}^d},$$
$$\|G(t, x, y) - G(t, x', y')\|^2_{\mathbb{R}^d} \leq L\|y - y'\|^2_{\mathbb{R}^d};$$

(iii) for all $x, y \in Y$ *and* $t \in [0, T]$, *there exist* $m, m' \in L^2(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{R}^d)$ *such that*

$$\|F(t, x, y)\|_{\mathbb{R}^d} \leq m(t),$$

$$\|G(t, x, y)\|_{\mathbb{R}^d} \leq m'(t);$$

(iv) $\sum_{k=1}^{p} c_k \neq -1$

Then there exists a solution $x : [0, T] \times \Omega \to \mathbb{R}^d$ *to the nonlocal functional stochastic integral inclusion III.1.*

Proof of the Theorem

Assume that $\sum_{k=1}^{p} c_k \neq -1$, *we have*

$$\begin{aligned} x(t) = x(0) &+ \int_0^t f(s, x(s), A(s)x(s))ds \\ &+ \int_0^t g(s, x(s), C(s)x(s))dB(s) \end{aligned} \tag{III.2}$$

From III.2 we have

$$\begin{aligned} x(t_k) = x(0) &+ \int_0^{t_k} f(s, x(s), A(s)x(s))ds \\ &+ \int_0^{t_k} g(s, x(s), C(s)x(s))dB(s) \quad (k = 1, ..., p) \end{aligned} \tag{III.3}$$

By III.1 and III.3,

$$x(0) + \sum_{k=1}^{p} c_k \Big[x(0) + \int_0^{t_k} f(s, x(s), A(s)x(s))ds + \int_0^{t_k} g(s, x(s), C(s)x(s))dB(s) \Big] = x_0 \tag{III.4}$$

Since $\sum_{k=1}^{p} c_k \neq -1$, *then III.4 implies*

$$x(0) = \Big(x_0 - \sum_{k=1}^{p} c_k \Big[\int_0^{t_k} f(s, x(s), A(s)x(s))ds + \int_0^{t_k} g(s, x(s), C(s)x(s))dB(s) \Big] \Big) \Big/ \Big(1 + \sum_{k=1}^{p} c_k \Big) \tag{III.5}$$

Then

$$x(t) = \frac{x_0 - \sum_{k=1}^{p} c_k \Big[\int_0^{t_k} f(s, x(s), A(s)x(s))ds + \int_0^{t_k} g(s, x(s), C(s)x(s))dB(s) \Big]}{1 + \sum_{k=1}^{p} c_k} + \int_0^{t} f(s, x(s), A(s)x(s))ds + \int_0^{t} g(s, x(s), C(s)x(s))dB(s) \tag{III.6}$$

Define the following integral operator V by

$$Vx(t) = \frac{x_0 - \sum_{k=1}^{p} c_k \Big[\int_0^{t_k} f(s, x(s), A(s)x(s))ds + \int_0^{t_k} g(s, x(s), C(s)x(s))dB(s) \Big]}{1 + \sum_{k=1}^{p} c_k} + \int_0^{t} f(s, x(s), A(s)x(s))ds + \int_0^{t} g(s, x(s), C(s)x(s))dB(s) \tag{III.7}$$

Lemma III.3.1 *The operator V sends the space* $\mathcal{C}([0, T], L^2(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{R}^d))$ *into itself.*

Proof III.3.1 *Let* $0 \leq t_1 \leq t_2 \leq T$. *Then for,* $x \in \mathcal{C}([0, T], L^2(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{R}^d))$. *Apply-*

ing the Cauchy-Schawrz inequality and the condition (iii) yields:

$$\begin{aligned}\mathbf{E}\|Vx(t_2)-Vx(t_1)\|^2_{\mathbb{R}^d} &\leq 2\mathbb{E}\|\int_{t_1}^{t_2} f(s,x(s),A(s)x(s))ds\|^2_{\mathbb{R}^d}\\ &+2\mathbb{E}\|\int_{t_1}^{t_2} g(s,x(s),C(s)x(s))dB_s\|^2_{\mathbb{R}^d} \qquad \text{(III.8)}\\ &\leq 2(t_2-t_1)\mathbb{E}\int_{[0,T]} m^2 ds + 2\mathbb{E}\int_{t_1}^{t_2} m'^2(s)ds.\end{aligned}$$

Therefore

$$\lim_{t_1\mapsto t_2}\mathbf{E}\|Vx(t_2)-Vx(t_1)\|^2_{\mathbb{R}^d}=0.$$

Consequently, Vx is continuous in mean square on $[0,T]$.
But the function Vx is square integrable with respect to mesure probability, has a finite second moment, and adapted to the filtration $\{\mathcal{F}_t\}_{t\in[0,T]}$.This implies that V maps Y into itself. □
Now, we will show that V is a contraction on Y under the distance

$$D_{x,y}:=D(x,y):=\sup_{t\in[0,T]} e^{-LK^2(T+1)t}[\mathbb{E}\|x(t)-y(t)\|^2_{\mathbb{R}^d}]^{\frac{1}{2}}$$

in Y. Let $K=max\{\max_{t\in[0,T]}\alpha(t),\max_{t\in[0,T]}\gamma(t)\}$. Indeed, for $x,y\in Y$ we have

$$\begin{aligned}D^2_{V(x),V(y)} &\leq 2\sup_{t\in[0,T]} e^{-2LK^2(T+1)t}\Big[\mathbb{E}\|\int_0^t \Big(f(s,x(s),A(s)x(s))-(f(s,x(s),A(s)y(s))\Big)ds\|^2_{\mathbb{R}^d}\\ &+\mathbb{E}\|\int_0^t\Big(g(s,x(s),C(s)x(s))-(g(s,x(s),C(s)y(s))\Big)dB_s\|^2_{\mathbb{R}^d}\Big]\\ &\leq 2LK^2\sup_{t\in[0,T]} e^{-2LK^2(T+1)t}\int_0^t\mathbb{E}\|x(s)-y(s)\|^2_{\mathbb{R}^d}ds\\ &\leq 2LK^2(T+1)D^2(x,y)\sup_{t\in[0,T]} e^{-2LK^2(T+1)t}\int_0^t e^{2LK^2(T+1)s}ds\\ &\leq (1-e^{-2LK^2(T+1)T})D^2(x,y).\end{aligned}$$

Since (Y,D) is a complete metric space, applying Banach's fixed point theorem we infer that there exists a solution $x^\star\in Y$ such that

$$x^\star(t)=x_0+\int_0^t f(s,x^\star,A(s)x^\star(s))ds+\int_0^t g(s,x^\star,C(s)x^\star(s))dB_s,\ \ for\ t\in[0,T].$$

Conclusion

This work has provided only a brief introduction to the topic of stochastic differential inclusions. Before delving deeper into this subject, it is necessary to have a strong foundation in stochastic differential equations. We began by reviewing key concepts from probability theory as well as stochastic calculus tools like the Itō integral and Itō's formula. However, a comprehensive understanding of stochastic differential inclusions requires further study into areas such as set-valued analysis, and the properties of set-valued stochastic processes. While this work has laid the groundwork, the field of stochastic differential inclusions remains rich with opportunities for continued research into both its theoretical underpinnings and applications across disciplines like engineering, biology, economics and beyond. The study of such inclusive models allows us to better capture the inherent uncertainties present in complex dynamical systems.

III.0 Bibliography

[1] *Ahmed, N. U. (2020). Optimal relaxed controls for nonlinear infinite dimensional stochastic differential inclusions. In Optimal Control of Differential Equations (pp. 1-19). CRC Press.*

[2] *Aubin, J. P., Cellina, A., Aubin, J. P., Cellina, A. (1984). Set-valued maps (pp. 37-92). Springer Berlin Heidelberg.*

[3] *Balasubramaniam, P. (2002). Existence of solutions of functional stochastic differential inclusions. Tamkang Journal of Mathematics, 33(1), 25-34.*

[4] *Brzezniak, Z., &, T. (2000). Basic stochastic processes: a course through exercises. Springer & Business Media.*

[5] *Calin, O. (2012). An introduction to stochastic calculus with applications to finance. Ann Arbor.*

[6] *Calin, O. (2015). An informal introduction to stochastic calculus with applications.*

[7] *Castaing, C., Valadier, M., Castaing, C., Valadier, M. (1977). Measurable multifunctions (pp. 59-90). Springer Berlin Heidelberg.*

[8] *Chalasani, P., & Jha, S. (1997). Steven Shreve: Stochastic Calculus and Finance. Lecture Notes, October, 1-343.*

[9] *Chilingar, G. V., Khilyuk, L. F., Reike, H. H. (2012). Probability in petroleum and environmental engineering. Elsevier.*

[10] *Cyganowski, S., Kloeden, O., &, J. (2001). From elementary probability to stochastic differential equations with MAPLE®. Springer Science & Business Media.*

[11] *Da Prato, G., Iannelli, M., Tubaro, L. (1978). Dissipative functions and finite-dimensional stochastic differential equations. JOURNAL DE MATHÉMATIQUES PURES ET APPLIQUÉES, 1978(2), 173-180.*

[12] *Deimling, K. (1992). Multivalued differential equations, volume 1 of de Gruyter Ser. Nonlinear Anal. Appl. Walter de Gruyter Co., Berlin.*

[13] *Dobrow, R. P. (2016). Introduction to stochastic processes with R. John Wiley & Sons.*

[14] *Gut, A., & Gut, A. (2006). Probability: a graduate course (Vol. 200, No. 5). New York: Springer.*

[15] *Han, Z., Cai, X., Huang, J. (2016). Theory of control systems described by differential inclusions. Heidelberg: Springer.*

[16] *Jacod, J., &, P. (2004). Probability essentials. Springer Science & Business Media.*

[17] *Jones, P. W., & Smith, P. (2017). Stochastic processes: an introduction. CRC Press.*

[18] *Kisielewicz, M. (2013). Stochastic differential inclusions and applications (Vol. 80). New York: Springer.*

[19] *Kozdron, M. J. Lectures on Stochastic Calculus with Applications to Finance.*

[20] *Krée, P. (1982). Diffusion equation for multivalued stochastic differential equations. Journal of Functional Analysis, 49(1), 73-90.*

[21] *Lacu, S. M. (2008). Simulation and inference for stochastic differential equations: with R examples (Vol. 486). New York: Springer.*

[22] *Le Gall, J. F. (2016). Brownian motion, martingales, and stochastic calculus (Vol. 274). New York: Springer.*

[23] *Leon-Garcia, A. (1994). Probability and random processes for electrical engineering. Pearson Education India.*

[24] *Lieb, E. H., & Loss, M. (2001). Analysis (Vol. 14). American Mathematical Soc.*

[25] *Lord, G. J., Powell, C. E., & Shardlow, T. (2014). An introduction to computational stochastic PDEs (Vol. 50). Cambridge University Press.*

[26] *Michta, M. (2002). Stochastic inclusions with multivalued integrators.*

[27] *Michta, M., Motyl, J. E. R. Z. Y. (2007). Set valued Stratonovich integral and Stratonovich type stochastic inclusion. Dynamic Systems and Applications, 16(1), 141.*

[28] *Oksendal, B. (2013). Stochastic differential equations: an introduction with applications. Springer Science & Business Media.*

[29] *Ross, S. M. (1997). Introduction to probability models, San Diego. Cal: Academic Press.*

[30] *Shreve, S. E. (2004). Stochastic calculus for finance II: Continuous-time models (Vol. 11). New York: springer.*

[31] *Stirzaker, D. (2005). Stochastic processes and models. OUP Oxford.*

[32] *van Handel, R. (2016). Probability and Random Processes. ORF 309/MAT 380 Lecture Notes.*

[33] *Velenik, Y. (2011). Probabilités et statistique. Université de Geneve.*

[34] *Williams, D. (1991). Probability with martingales. Cambridge university press.*