People's Democratic Republic of Algeria

Ministry of Higher Education and Scientific Research

MOHAMED KHIDER UNIVERSITY OF BISKRA

FACULTY OF EXACT AND NATURAL SCIENCES

MATHEMATICS DEPARTMENT

Thesis presented in view of obtaining the diploma :

MASTERS in Mathematics

Option: **Probability**

By

HAMDIKEN Iskander Ouassim

Title :

Stochastic Maximum Principle in Regime Switching Stochastic Systems

Members of jury :

Pr.	KHALFALLAH Nabil	UMKB	President
Pr.	CHIGHOUB Farid	UMKB	Supervisor
Dr.	ROMEILI Nacira	UMKB	Examinator

June 26th, 2022

Dedication

To my dear and beloved parents, my brother Anis, and all of my family for their unlimited support and help during every stage of my life.

To my grandfather, Mohamed, and my uncles, Hichame, Azzedine, and Yacine.

To all my teachers in the mathematics department.

To my high school mathematics teacher, Mrs. Boungueb.

To Pr. Redjel Nadjah.

To the souls of my grandfather, Ahmed the clockmaker, my uncle, Dr. Belgacem Hamdiken, my teacher, Pr. Ghezal Said, and Pr. Dehici Abdelkader. May your souls rest in peace.

Acknowledgement

All praise belongs to Allah The Lord of the worlds, Most Gracious, Most Merciful.

I would like to thank my supervisor, **Pr. CHIGHOUB Farid**, who read my numerous revisions and helped make some sense of the confusion, and for his helpful advice and detailed explanations.

I would like to thank the jury members, **Pr. KHALFALLAH Nabil** and **Dr. ROMEILI Nacira**, who kindly accepted to examine my thesis, as well as all the professors of the mathematics department of Biskra for their help throughout my years in the university.

I would like to gratefully acknowledge the guidance, support and encouragement of **Pr. YOUKANA Amar**, who, throughout the years, clarified many things for me in mathematics, and kept me motivated. You were right Pr., it has always been about asking the right question.

Many thanks to all my friends and colleagues of the mathematics department and outside of it, without any exception : **BERROUSSI Zohir, IKHENACHE Wail, LESLOUS Aymen**, **REMAIGUI Brahim**, and **MOUFFOK Aymen** for their constant help and motivation.

Contents

Dedication **i**

Acknowledgement **ii**

Table of Contents **iii**

List of Figures **v**

List of Tables **1**

Introduction **1**

1 Stochastic Calculus Preliminaries **4**

1.1 Stochastic Processes 4

1.2 Brownian Motion 5

1.3 Stochastic Integral 6

1.3.1 Wiener Integral 7

1.3.2 Itô's Integral 8

1.4 Stochastic Differential Equation 10

1.4.1 Existence and Uniqueness of SDE Solutions 11

1.5 Backward Stochastic Differential Equation 12

1.5.1 Existence and Uniqueness Theorem 13

1.6 Continuous-time Markov Chains 14

2 Stochastic Maximum Principle **17**

2.1 Optimization Problem Formulation 17

2.2 Stochastic Maximum Principle 18

2.3 Near-optimal Controls 21

2.4 Regime Switching Stochastic System 23

2.4.1 Existence and Uniqueness theorem 25

3 Regime Switching Stochastic Systems : Application in Viral Models **27**

3.1 Stochastic SIRS Model 27

3.2 Hypothesis 30

3.3 Sufficient Conditions for Near-optimal Controls 31

3.3.1 Estimates On The Parameters 31

3.3.2 Sufficient Conditions for Near-optimal Controls 33

3.4 Necessary Conditions For Near-optimal Controls 37

3.4.1 Estimates On The Parameters 37

3.4.2 Necessary Conditions for Near-optimal Controls 38

Conclusion **41**

Appendix A : Theorems **43**

Appendix B: Abbreviations and Notations **45**

List of Figures

1.1 Irreductible Markov chain. 15

List of Tables

3.1 Notations used in the model 28

Introduction

In mathematics, optimization theory represents an important tool used by mathematicians in order to solve maximization and minimization problems, both theoretical and practical ones. It helps us determine extreme values, study the behaviour of systems, and mainly, control these systems in a way that helps us take the best decisions.

As a motivation, let us recall a well known concept in mathematics. Given a function y, a differential equation is a mathematical equation that depends on y and its derivatives. The order depends on the "highest" derivative in a sense : A first order differential equation for example depends on y and its first derivative.

This type of differential equations is called Ordinary differential equations (ODEs for short), and it is of deterministic variation (it only depends on time).

As of the 50's, along with the major advancements in mathematics as well as other scientific fields, random systems (also known as dynamic systems) emerged, and more mathematicians started studying this new type of random systems, also called "stochastic". The differential equation here depends also on a second component which is stochastic, the equation becomes a Stochastic differential equations (SDE for short). The new stochastic part is called the diffusion and hence, ODEs represent a special case of SDE where the diffusion part is equal to zero. Itô's Calculus, introduced by Kiyosi Itô, is considered to be one of the first wokrs in the field of stochastic calculus. We mention also Pardoux and Peng's[8] important works on BSDE's, and many other works that are on the same scale of importance.

Optimization theory, few years later, got included into this new field of stochastic calculus and dynamical systems, allowing the extention and the developpment of numerous theories, as well as the generalization of older concepts.

Among the main concepts in optimization theory is the optimal control. That is, a system provided with a control function that helps us take decisions optimally.

The idea is simple : We suppose a stochastic system $dX(t) = b(t, X(t))dt + \sigma(t, X(t))dB(t)$. We then define a control function u which will be implemented in the system in order to optimally control it, according to our cost function J which we want maximized (or minimized). New optimization methods emerged, such as Bellman's dynamical programming principle, and Pontryagin's principle.

Among the various types of stochastic systems, we are interested in thesis in studying a type called the *Regime switching stochastic systems*. These systems have a special property : They switch their behaviour abruptly, leading to a new system that no longer depends on the previous one.

This property is somewhat similar to a famous mathematical concept. The Markov property. This is why regime switching systems usually depend on a Markov chain that plays this important role in sudden change of its behaviour.

Our aim now is to study the near-optimal controls of such systems using the stochastic maximum principle. And for that we propose the following apporach :

In Chapter One : We recall some of the main mathematical preliminaries concerning the stochastic calculus, as well as the stochastic differential equations, and Markov chains.

In Chapter Two : We give the formulation of an optimization problem, then we recall the stochastic maximum principle, along with the near-optimal controls and Ekeland's principle. Later on, we explain the regime switching stochastic systems and how they differ from other dynamical systems.

In Chapter Three : This chapter is more of an application. We study the necessary and sufficient conditions of near-optimality in a viral SIRS model. We introduce a Markov chain, in the system, allowing it to be of regime switching behaviour.

Chapter 1

Stochastic Calculus Preliminaries

1.1 Stochastic Processes

We consider a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, which is a probability space, provided with an increasing sequence of sub-σ-algebras of $\mathcal{F}$ called filtration.

Definition 1.1.1 *(Stochastic process) We call stochastic process $X = (X(t))_{t\in\mathbb{R}}$ every sequence of real-valued random variables*

$$X(t,.) : (\Omega, \mathcal{F}) \longrightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$$
$$\omega \longrightarrow X(t,\omega)$$

Proposition 1.1.2 *For each fixed $t \in \mathbb{R}_+$, the stochastic process X becomes a random variable, while for each fixed $\omega \in (\Omega, \mathcal{F})$, the stochastic process X becomes a real function, usually called trajectory.*

Definition 1.1.3 *(Adaptability) We say that a stochastic process X is adapted to the filtration $\mathbb{F}$ ($\mathbb{F}$-adapted for simplicity) if and only if $\forall t \in \mathbb{R}_+$, the random variable $X(t)$ is $\mathcal{F}_t$-measurable.*

Remark 1.1.4 *It is obvious that every stochastic process X is adapted to its natural filtration $(\mathcal{F}_t^X)_{t\in\mathbb{R}_+}$, that is $\mathcal{F}_t^X = \sigma(X(s), 0 \le s \le t), \forall t \in \mathbb{R}_+$.*

We define the following class $\mathcal{P}$

$$\mathcal{P} = \left\{ A \times \Omega / A \subseteq \mathbb{R}_+, \forall t \geq 0, (A \cap [0,t]) \times \Omega \in \mathcal{B}[0,t] \times \mathcal{F}_t \right\}.$$

The class $\mathcal{P}$ is a σ-algebra and is called the class of progressively measurable sets.

Definition 1.1.5 *(Progressive measurability) Let $X = (X(t))_{t\geq 0}$ be a stochastic process from $(\Omega, \mathcal{F})$ in $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Then X is said to be progressively measurable if $X : [0,t] \times \Omega \to \mathbb{R}$ is measurable from $\mathcal{B}([0,t]) \times \mathcal{F}_t$ in $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.*

1.2 Brownian Motion

Definition 1.2.1 *A brownian motion (Wiener process) $B = (B(t))_{t\geq 0}$ is characterized by four properties*

1. *$B(0) = 0$ $\mathbb{P} - a.s.$*
2. *For every $\omega \in \Omega$, the function $t \to B(t, \omega)$ is $\mathbb{P} - a.s$ continuous.*
3. *For every $s \in [0,t]$, the random variable $B(t) - B(s)$ is a centered gaussian random variable with variance $t - s$ ($B(t) - B(s) \sim \mathcal{N}(0, t-s)$) and is independent of the σ-algebra $(\mathcal{F}_u^B)_{t\geq 0}$. We say that the brownian motion has independent and stationary increments.*
4. *For every $t > 0$, $B(t) \sim \mathcal{N}(0,t)$.*

Property 1.2.2

- *A brownian motion $B = (B(t))_{t\geq 0}$ is a gaussian process.*
- *A brownian motion is nowhere differentiable.*
- *For every couple $(s,t) \in \mathbb{R}_+ \times \mathbb{R}_+$, $Cov(B(t), B(s)) = t \wedge s$.*
- *A brownian motion has an infinite variation, but its quadratic variation is finite.*

Example 1.2.3 *Let $B = (B(t))_{t\geq 0}$ be a standard brownian motion. Then the process $(\tilde{B}(t))_{t\geq 0}$ defined $\forall t > 0$ by*

$$\tilde{B}(t) = tB\left(\frac{1}{t}\right), \qquad \tilde{B}(0) = 0$$

is also a brownian motion (Time inverted brownian motion).

Definition 1.2.4 *(Martingale) We say that a stochastic process $X = (X(t))_{t\geq 0}$ is a martingale if it is*

1. *$\mathbb{F}$-adapted.*
2. *Integrable : $\mathbb{E}\,|X(t)| < \infty, \forall t \in \mathbb{R}_+$.*
3. *It verifies the following property $\forall t \geq s$*

$$\mathbb{E}[X(t)\big|\mathcal{F}_s] = X(s).$$

Definition 1.2.5 *Let X be a stochastic process verifying 1 and 2. Then X is*

- *Supermartingale, if $\mathbb{E}[X(t)\big|\mathcal{F}_s] \leq X(s)$.*
- *Submartingale, if $\mathbb{E}[X(t)\big|\mathcal{F}_s] \geq X(s)$.*

Remark 1.2.6 *A martingale is both a submartingale and a supermartingale at the same time.*

Example 1.2.7 *A brownian motion B is a martingale with respect to its natural filtration $(\mathcal{F}^B_t)_{t\geq 0}$.*

1.3 Stochastic Integral

The Lebesgue-Stieltjes integral represents an extension or, better said, a generalization of the Riemann-Stieltjes and Lebesgue integrals. Given an integral $\int_a^b f(x)dg(x)$, it is defined when $f : [a, b] \to \mathbb{R}$ is Borel-measurable and bounded, while $g : [a, b] \to \mathbb{R}$ is of bounded variation in $[a, b]$, or when f is non-negative and g is monotone and right-continuous (càd). If we take a closer look at these conditions on the function g, we will find that non of them apply on the brownian motion, which raises the question on how we will evaluate an integral of the following form

$$I_T(\theta) = \int_0^T \theta(s)dB(s), \tag{1.1}$$

where $\theta = (\theta(t))_{t\geq 0}$ is some process, and $B = (B(t))_{t\geq 0}$ is a brownian motion.

1.3.1 Wiener Integral

We note

$$L^2\left([0,T],\mathbb{R}\right) = \{f : [0,T] \to \mathbb{R} : \int_0^T |f(s)|^2 ds < +\infty\}.$$

Definition 1.3.1 *We define the Wiener integral*

$$\int_0^T f(s)dB(s),$$

where f is a deterministic function.

Remark 1.3.2 *We define a scalar product on $L^2\left([0,T],\mathbb{R}\right)$ by*

$$\langle f, g\rangle \to \int_0^T f(s)g(s)ds. \tag{1.2}$$

Under the scalar product 1.2, $L^2\left([0,T],\mathbb{R}\right)$ is a Hilbert space.

We define now a sequence of deterministic step functions $(f_n)_{n\in\mathbb{N}} \subset L^2\left([0,T],\mathbb{R}\right)$

$$f_n(t) = \sum_{i=0}^{n} \alpha_i \mathbf{1}_{]t_i^{(n)}, t_{i+1}^{(n)}[}(t),$$

where $(t_i^{(n)})$ is an increasing sequence in $[0,T]$.

For every $n \in \mathbb{N}$, the Wiener integral of f_n is

$$I_T(f_n) = \int_0^T f_n(s)dB(s) = \sum_{i=0}^{n} \alpha_i(B(t_{i+1}^{(n)}) - B(t_i^{(n)})).$$

By the characteristics of the brownian motion, and the independence of the increments, we find that $I_T(f_n)$ is a centered gaussian random variable with variance

$$\begin{aligned} Var(I_T(f_n)) &= \sum_{i=0}^{n} \alpha_i^2 Var(B(t_{i+1}^{(n)}) - B(t_i^{(n)})) \\ &= \sum_{i=0}^{n} \alpha_i^2 \int_{t_i^{(n)}}^{t_{i+1}^{(n)}} ds \\ &= \int_0^T \sum_{i=0}^{n} \alpha_i^2 \mathbf{1}_{]t_i^{(n)}, t_{i+1}^{(n)}[}(t) ds \\ &= \int_0^T (f_n(s))^2 dt. \end{aligned}$$

Remark 1.3.3 *Let $f \in L^2([0,T],\mathbb{R})$. There exists a sequence of step functions $(f_n)_{n\in\mathbb{N}}$ that converges to f in $L^2([0,T],\mathbb{R})$. We can construct the Wiener Integral $I_T(f_n)$ of f_n, which are gaussian random variables and form a Cauchy sequence in $L^2([0,T],\mathbb{R})$ that is complete. This sequence then converges to a gaussian random variable that we note $I_T(f)$ that does not depend on the choice of $(f_n)_{n\in\mathbb{N}}$. This random variable is called Wiener's Integral of f with respect to the brownian motion B.*

1.3.2 Itô's Integral

After defining the integral 1.1 in the case where the integrated function is deterministic, we go a bit further now and we try to give a sense to the integral

$$I_T(\theta) = \int_0^T \theta(s) dB(s),$$

where θ is a random variable. The construction method is the same.

Consider a step stochastic process $(\theta_n)_{n\in\mathbb{N}}$ defined $\forall n \in \mathbb{N}$ by

$$\theta^{(n)}(t) = \sum_{i=0}^{n} \theta_i \mathbf{1}_{[t_i, t_{i+1}[}(t), \tag{1.3}$$

where $\theta_i \in L^2(\Omega, \mathcal{F}_{t_i}, \mathbb{P})$, $\forall i$. Let $\mathcal{D}$ be the space of all càglàd (left continuous, with right limit),

$\mathbb{F}$-adapted processes $\theta = (\theta(t))_{t\geq 0}$ such that

$$\mathbb{E}\left[\int_0^T |\theta(s)|^2 ds\right] < \infty.$$

It is clear that the step stochastic process $\theta^{(n)} \in \mathcal{D}$. For every stochastic process $\theta \in \mathcal{D}$ we can define its stochastic integral by approaching it by a sequence of step stochastic processes. Since the limit is in $L^2(\Omega, \mathcal{F}_T, \mathbb{P})$, the stochastic integral $I_T(\theta)$ is simply the limit of $I_T(\theta^{(n)})$ as n tends to infinity.

Using the same steps in the previous section, we find that

$$\mathbb{E}\left[I_T(\theta)\right] = 0,$$

and

$$Var(I_T(\theta)) = \mathbb{E}\left[\int_0^T \theta^2(s) ds\right].$$

Definition 1.3.4 *(Itô process) We call Itô process a real valued stochastic process $X = (X(t))_{t\geq 0}$ such that*

$$X(t) = X(0) + \int_0^t b(s)ds + \int_0^t \sigma(s)dB(s), \tag{1.4}$$

where $X(0)$ is $\mathcal{F}_0$-measurable, b and σ are both $\mathbb{F}$-adapted, σ is càglàd, and

$$\int_0^T |b(s)|ds < +\infty, \qquad \int_0^T |\sigma(s)|^2 ds < +\infty. \tag{1.5}$$

Remark 1.3.5 *In an Itô process, $b(s)$ is called the drift coefficient, while $\sigma(s)$ is called the diffusion coefficient (volatility also in finance).The term $X(0) + \int_0^t b(s)ds$ has a finite variation, and $\int_0^t \sigma(s)dB(s)$ is the martingale part of the Itô process, regarding its form (A well defined stochastic integral).*

Theorem 1.3.6 *(Itô's first formula) Let $X = (X(t))_{t\geq 0}$ be an Itô process, $f : \mathbb{R} \to \mathbb{R}$ be a C^2-function. Then we have*

$$f(X(t)) = f(X(0)) + \int_0^t D_x f(X(s))dX(s) + \frac{1}{2}\int_0^t D_{xx} f(X(s))\sigma^2(s)ds. \tag{1.6}$$

Theorem 1.3.7 *(Itô's second formula) Let $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, $f \in C^2(\mathbb{R} \times \mathbb{R})$. Let $X = (X(t))_{t\geq 0}$, $Y = (Y(t))_{t\geq 0}$ be two Itô processes. Then we have*

$$f(X(t), Y(t)) = f(X(0), Y(0)) + \int_0^t D_x f(X(s), Y(s)) dX(s) + \int_0^t D_y f(X(s), Y(s)) dY(s)$$
$$+ \frac{1}{2}\int_0^t D_{xx} f(X(s), Y(s)) d\langle X\rangle_t + \frac{1}{2}\int_0^t D_{yy} f(X(s), Y(s)) d\langle Y\rangle_t + \int_0^t D_{xy} f(X(s), Y(s)) d\langle X, Y\rangle_t, \tag{1.7}$$

where $\langle X\rangle_t$ denotes the quadratic variation of X.

Remark 1.3.8 *In the case where $Y(t) = t, \forall t \in \mathbb{R}_+$, we get the following formula*

$$f(t, X(t)) = f(0, X(0)) + \int_0^t D_t f(s, X(s)) ds + \int_0^t D_x f(s, X(s)) dX(s) + \frac{1}{2}\int_0^t D_{xx} f(s, X(s)) d\langle X\rangle_t. \tag{1.8}$$

1.4 Stochastic Differential Equation

Let $B = (B(t))_{t\geq 0}$ be a d-dimensional brownian motion over a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t^B)_{t\geq 0}, \mathbb{P})$ with $(\mathcal{F}_t^B)_{t\geq 0}$ being the natural filtration of the brownian motion B. Let $T > 0$, we consider two functions $\alpha : [0, T] \times \Omega \times \mathbb{R}^n \to \mathbb{R}^n$ and $\beta : [0, T] \times \Omega \times \mathbb{R}^n \to \mathbb{R}^{n\times d}$. Denote $\|\beta\| = \text{trace}(\beta\beta^\top)$. Our aim now is to solve the following stochastic differential equation

$$X(t) = X(0) + \int_0^t \alpha(s, X(s)) ds + \int_0^t \beta(s, X(s)) dB(s), \tag{1.9}$$

or in its differential form

$$dX(t) = \alpha(t, X(t)) dt + \beta(t, X(t)) dB(t). \tag{1.10}$$

Definition 1.4.1 *A (strong) solution to the stochastic differential equation 1.9 is a stochastic process $X = (X(t))_{t\geq 0}$ such that*

1. *X is progressively measurable.*

2. *$\int_0^T \left(|\alpha(s, X(s))| + \|\beta(s, X(s))\|^2\right) ds < +\infty \quad \mathbb{P} - a.s.$*

3. $X(t) = X(0) + \int_0^t \alpha(s, X(s))ds + \int_0^t \beta(s, X(s))dB(s), \quad \mathbb{P} - a.s.$

Example 1.4.2 *The stochastic process $X = (X(t))_{t\geq 0}$ defined for every $t \geq 0$ by $X(t) = e^{B(t)-\frac{t}{2}}$ is a solution to the SDE : $dX(t) = X(t)dB(t)$.*

1.4.1 Existence and Uniqueness of SDE Solutions

Theorem 1.4.3 *Consider the SDE 1.9. If α and β satisfy the following conditions*

- *Lipschitz continuity : There exists $C > 0$ such that $\forall (t, x, y) \in \mathbb{R}_+ \times \mathbb{R}^n \times \mathbb{R}^n$*

$$|\alpha(t, x) - \alpha(t, y)| \leq C|x - y|, \qquad ||\beta(t, x) - \beta(t, y)|| \leq C|x - y|.$$

- *For every $t \in \mathbb{R}_+$*

$$\mathbb{E}\left[\int_0^T \left(|\alpha(t, 0)|^2 + \|\beta(t, 0)\|^2\right) dt\right] < +\infty$$

then the SDE has a unique solution X. Plus, for every $t \geq 0$, the solution verifies

$$\mathbb{E}\left(\int_0^T |X(t)|^2 dt\right) < +\infty.$$

Proof. The proof[1] consists of two parts

- Uniqueness : In which we suppose the existence of two distinct solutions X_1, X_2 to the SDE 1.9 then we prove that $X_1 = X_2$ a.s.
- Existence : A famous way to prove the existence of a solution under the previous conditions in theorem 1.4.3 is Picard iteration.

■

Example 1.4.4 *The SDE*

$$dX(t) = (a - bX(t))dt + \sigma dB(t),$$

with $X(0) = x_0$ and $(a, b, \sigma) \in \mathbb{R}^3$ has a unique solution $X = (X(t))_{t\geq 0}$ called the Ornstein-Uhlenbeck process.

[1] For a fully detailed proof, we suggest seeing [3] p 261-266 and [10] p 68-70.

1.5 Backward Stochastic Differential Equation

We go now to the concept of Backward stochastic differential equations. We recall the martingal representation theorem. Denote $L^2(\Omega, \mathcal{F}_T^B; \mathbb{R}^n)$ as the set of $\mathcal{F}_T^B$-measurable, $\mathbb{R}^n$-valued random variables X such that $\mathbb{E}[|X|^2] < +\infty$, and denote $\mathcal{M}^2([0,T], \mathbb{R}^{n\times d})$ as the set of all adapted, $\mathbb{R}^{n\times d}$-valued processes X such that $E\left[\int_0^T \|X(t)\|^2 dt\right] < +\infty$.

The theorem states that for every $\xi \in L^2(\Omega, \mathcal{F}_T^B; \mathbb{R}^n)$, there exists a unique $\sigma \in \mathcal{M}^2([0,T], \mathbb{R}^{n\times d})$ such that

$$\xi = \mathbb{E}[\xi] + \int_0^T \sigma(t) dB(t),$$

which is generalized later for any $\mathbb{F}^B$-martingale.

The martingale representation theorem turns out to be a special case of backward stochastic differential equations where the generator is equal to zero.[2] The general case was studied and proven by Pardoux and Peng.[3]

Let $\mathbb{H}^2([0,T], \mathbb{R}^n)$ be the set of all $\mathbb{R}^n$-valued, progressively-measurable stochastic processes X such that

$$\mathbb{E}\left[\sup_{0\le t\le T} |X(t)|^2\right] < \infty.$$

Let $\xi \in L^2(\Omega, \mathcal{F}_T; \mathbb{R}^n)$ be an $\mathcal{F}_T$-measurable random variable, $f : [0,T] \times \mathbb{R}^n \times \mathbb{R}^{n\times d} \to \mathbb{R}^n$ such that for every $(x,z) \in \mathbb{R}^n \times \mathbb{R}^{n\times d}$, the process $(f(t,x,z))_{t\ge 0}$ is progressively measurable.

Definition 1.5.1 *(BSDE) We consider now the following equation*

$$-dX(t) = f(t, X(t), Z(t)) - Z(t)dB(t), \qquad X(T) = \xi, \tag{1.11}$$

or using the integral form

$$X(t) = \xi + \int_t^T f(s, X(s), Z(s))ds - \int_t^T Z(s)dB(s), \qquad 0 \le t \le T. \tag{1.12}$$

This equation is called Backward stochastic differential equation (BSDE in short), where f is the generator, and ξ is the terminal condition.

[2] See [10]
[3] See [8].

Definition 1.5.2 *(BSDE Solution) A couple $(X(t), Z(t))_{t\geq 0}$ is said to be a solution to the BSDE 1.12 if and only if*

1. *X and Z are progressively measurable.*
2. *$Z \in \mathbb{H}^2([0,T], \mathbb{R}^{n\times d})$ i.e $\mathbb{E}\big[\int_0^T \|Z(t)\|^2 dt\big] < \infty$.*
3. *We have $\mathbb{P}$-a.s*

$$X_t = \xi + \int_t^T f(s, X(s), Z(s))ds - \int_t^T Z(s)dB(s), \qquad 0 \leq t \leq T. \tag{1.13}$$

1.5.1 Existence and Uniqueness Theorem

Consider the BSDE 1.12. We suppose that the following properties hold

1. f is Lipschitz-continuous in (x, z) : For all $(t, x, x', z, z') \in [0,T] \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^{n\times d} \times \mathbb{R}^{n\times d}$

$$|f(t, x, z) - f(t, x', z')| \leq K(|x - x'| + \|z - z'\|). \tag{1.14}$$

2. The integrability condition

$$\mathbb{E}\left[|\xi^2| + \int_0^T |f(r, 0, 0)|^2 dr\right] < +\infty. \tag{1.15}$$

Then we have the following theorem

Theorem 1.5.3 *(Pardoux-Peng) Under conditions 1.14,1.15, the BSDE 1.12 has a unique solution.*

Proof. See [1]. ■

Property 1.5.4 *Consider the following BSDE*

$$X(t) = \xi + \int_t^T f(s, X(s), Z(s))ds - \int_t^T Z(s)dB(s), \qquad 0 \leq t \leq T, \tag{1.16}$$

then

$$X(t) = \mathbb{E}\left[\xi + \int_t^T f(s, X(s), Z(s))ds \,\middle|\, \mathcal{F}_t\right], \qquad 0 \leq t \leq T. \tag{1.17}$$

Proof. Using conditional expectation

$$
\begin{aligned}
\mathbb{E}[X(t)|\mathcal{F}_t] &= \mathbb{E}\left[\xi + \int_t^T f(s, X(s), Z(s))ds - \int_t^T Z(s)dB(s)\middle|\mathcal{F}_t\right] \\
&= \mathbb{E}\left[\xi + \int_t^T f(s, X(s), Z(s))ds - \int_0^T Z(s)dB(s) + \int_0^t Z(s)dB(s)\middle|\mathcal{F}_t\right] \\
&= \mathbb{E}\left[\xi + \int_t^T f(s, X(s), Z(s))ds\middle|\mathcal{F}_t\right] - \mathbb{E}\left[\int_0^T Z(s)dB(s) - \int_0^t Z(s)dB(s)\middle|\mathcal{F}_t\right] \\
&= \mathbb{E}\left[\xi + \int_t^T f(s, X(s), Z(s))ds\middle|\mathcal{F}_t\right] - \int_0^t Z(s)dB(s) + \int_0^t Z(s)dB(s) \\
&= \mathbb{E}\left[\xi + \int_t^T f(s, X(s), Z(s))ds\middle|\mathcal{F}_t\right].
\end{aligned}
\tag{1.18}
$$

■

1.6 Continuous-time Markov Chains

Markov chains, named after russian mathematician Andrey Markov (1856-1922) are stochastic processes with discrete state spaces that verify the Markovian property, meaning that the probability of an event defined by this process in the future is independent of its past, and depends only on its current state, and the jumps between states are countable.[4]

Definition 1.6.1 *(Memoryless property) We say that a random variable X is memoryless if it verifies the following property*

$$\mathbb{P}\left[X > t + s\middle|X > t\right] = \mathbb{P}\left[X > s\right], \quad \forall (s,t) \in \mathbb{R}_+ \times \mathbb{R}_+.$$

Definition 1.6.2 *(Continuous-time Markov chain) Let $X = (X(t))_{t\geq 0}$ be a continuous-time stochastic process with state space $\mathbb{S} = \{1, 2, ...\}$.*

[4] See [5].

We say that X is a continuous-time Markov chain if

$$\mathbb{P}\left[X(t+s)=j \middle| X(s)=i, X(r)=x_r, 0 \leq r < s\right] = \mathbb{P}\left[X(t+s)=j \middle| X(s)=i\right]$$
$$= P_{i,j}(t),$$
$$\forall (s,t) \in \mathbb{R}_+ \times \mathbb{R}_+, \text{ and } \forall (i,j,x_r) \in \mathbb{S}^3.$$

Remark 1.6.3

- *Continuous-time Markov chains are also known as Markov jump processes.*
- *The function $P_{i,j}$ is called the transition function (or Markovian kernel) of the continuous-time Markov chain.*
- *If there exists $t \geq 0$ for which $P_{i,j}(t) > 0$, and $t^* \geq 0$ for which $P_{j,i}(t^*) > 0$, then we say that states i and j communicate.*
- *A continuous-time Markov chain is said to be irreductible if all states communicate.*
- *We can write that $\sum_{j=1}^{+\infty} P_{i,j}(t) = 1, \ \forall i \in \mathbb{S}$.*

Example 1.6.4 *An irreductible Markov chain with three states $\mathbb{S} = \{A, B, C\}$*

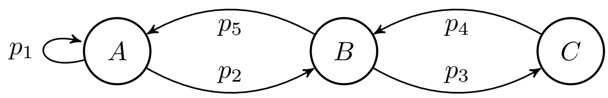

Figure 1.1: Irreductible Markov chain.

Proposition 1.6.5 *(Chapman-Kolmogorov Equation) For all $(s,t) \in \mathbb{R}_+ \times \mathbb{R}_+$, we have*

$$P_{i,j}(t+s) = \sum_{j=0}^{+\infty} P_{i,j}(t) P_{j,k}(s) = \sum_{j=0}^{+\infty} P_{i,j}(s) P_{j,k}(t).$$

Definition 1.6.6 *(Markov chain generator) Let $X = (X_t)_{t \geq 0}$ be a continuous-time Markov chain defined on a finite discrete state space $\mathbb{S} = \{1, 2, ..., N\}$. The Markov chain generator*

$Q = (q_{ij})_{N\times N}$ *is defined by*

$$P_{i,j}(h) = \mathbb{P}\left[X(t+h) = j | X(t) = i\right] = \begin{cases} q_{ij}h + o(h), & \text{if } i \neq j, \\ 1 + q_{ij}h + o(h), & \text{if } i = j. \end{cases}$$

The entries q_{ij} define the rate with which the chain leaves the state i for the state j.

Property 1.6.7 *The entries q_{ij} verify $\forall i \in \mathbb{S}$*

$$q_{ii} = -\sum_{j\neq i} q_{ij}.$$

Chapter 2

Stochastic Maximum Principle

In an attempt to understand and to control stochastic systems, many articles and works were published, proposing new methods to approach these problems. We mention among them Bellman's Dynamical programming principle [1950's] and Pontryagin's stochastic maximum principle, which will be our topic in this chapter.

2.1 Optimization Problem Formulation

We consider a d-dimensional brownian motion B over a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$, where $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$, is the natural filtration of the brownian motion B.

We define a control function $u : [0, T] \times \Omega \to \Gamma$. The control function u is usually referred to as a "decision" function. The space Γ represents the control constraint, which is usually a set to determine the image of the control based on the optimization problem (The amount of money spent in a month should not overpass the monthly income).

Consider now the following problem

$$\begin{cases} dX(t) = b(t, X(t), u(t))dt + \sigma(t, X(t), u(t))dB(t), \\ X(0) = x_0, \end{cases} \tag{2.1}$$

where $b : [0, T] \times \Omega \times \mathbb{R}^n \times \Gamma \to \mathbb{R}^n, \quad \sigma : [0, T] \times \Omega \times \mathbb{R}^n \times \Gamma \to \mathbb{R}^{n \times d}$.

We define the cost functional J

$$J(0,X(0),u(t)) = \mathbb{E}\left[\int_0^T f(t,X(t),u(t))dt + h(X(T))\right]. \tag{2.2}$$

Definition 2.1.1 *(Feasible control) The set of all feasible controls is defined by*

$$\mathcal{U}[0,T] = \{u : [0,T]\times\Omega \to \Gamma/u(.)\textit{is measurable}\}.$$

Remark 2.1.2 *The term "feasible" literally means "that can be done" which means that a feasible control u is a decision that can be taken at the time t regardless of its consequences. A feasible control u can be also interpreted this way : At any given time t, we have enough information to take a decision $u(t)$ ($\mathbb{F}$-adaptability), but that doesn't mean it's the best decision to be taken.*

Definition 2.1.3 *(Admissible control) A control u is called an admissible control, and $(X(.),u(.))$ an admissible pair if*

- $u(.) \in \mathcal{U}[0,T]$.
- *$X(.)$ is the unique solution to the equation 2.1.*
- $L(.,X(.),u(.)) \in L^1_{\mathbb{F}}([0,T],\mathbb{R}), \quad h(X(T)) \in L^1(\Omega,\mathcal{F}_T;\mathbb{R})$.

Remark 2.1.4 *We denote $\mathcal{U}_{ad}[0,T]$ the set of all admissible controls u over the time horizon $[0,T]$.*

2.2 Stochastic Maximum Principle

We suppose a finite-horizon stochastic control problem

$$dX(t) = b(t,X(t),u(t))dt + \sigma(s,X(t),u(t))dB(t), \tag{2.3}$$

with cost functional

$$J(0,X(0),u(t)) = \mathbb{E}\Bigg[\int_0^T L(s,X(s),u(s))ds + g(X(T))\Bigg], \tag{2.4}$$

where $L : [0,T] \times \Omega \times \mathbb{R}^n \times \Gamma \to \mathbb{R}$ is a continuous function in (t,x) for every $u \in \mathcal{U}_{ad}$, $g : \mathbb{R}^n \to \mathbb{R}$ is a C^1-convex function, and both f, g are of quadratic growth with respect to x.

Definition 2.2.1 *(Generalized Hamiltonian) The Generalized Hamiltonian is given by*

$$
\begin{aligned}
\mathcal{H} :&[0,T] \times \mathbb{R}^n \times \Gamma \times \mathbb{R}^n \times \mathbb{R}^{n\times d} \to \mathbb{R}, \\
\mathcal{H}(t,x,u,p,q) &= b(t,x,u).p + trace(\sigma'(t,x,u).q) + L(t,x,u).
\end{aligned}
$$

Definition 2.2.2 *(Adjoint equation) We call Adjoint Equation the following BSDE*

$$
-dp(t) = D_x H(t, X(t), u(t), p(t), q(t))dt - q(t)dB(t), \qquad Y_T = D_x g(X_T). \tag{2.5}
$$

Theorem 2.2.3 *(Verification theorem) Let $\tilde{u} \in \mathcal{U}_{ad}$ and let $\tilde{X}$ be the controlled diffusion. Suppose that there exists a solution $(\tilde{p}, \tilde{q})$ to the corresponding adjoint equation 2.5 such that a.s*

$$
\mathcal{H}(t, \tilde{X}(t), \tilde{u}(t), \tilde{p}(t), \tilde{q}(t)) = \min_{u \in \mathcal{U}_{ad}} \mathcal{H}(t, \tilde{X}(t), u(t), \tilde{p}(t), \tilde{q}(t)), \quad 0 \le t \le T, \tag{2.6}
$$

and suppose that

$$
(x,u) \longrightarrow \mathcal{H}(t, x, u, \tilde{p}(t), \tilde{q}(t)),
$$

is a convex function $\forall t \in [0,T]$. Then $\tilde{u}$ is an optimal control

$$
J(0, X(0), \tilde{u}(t)) = \min_{u \in \mathcal{U}_{ad}} J(0, X(0), u(t)).
$$

Proof. We have $\forall u \in \mathcal{U}_{ad}$:

$$
\begin{aligned}
&J(0, X(0), \tilde{u}(t)) - J(0, X(0), u(t)) \\
&\qquad = \mathbb{E}\bigg[\int_0^T \Big(L(t, \tilde{X}(t), \tilde{u}(t) - L(t, X(t), u(t) \Big) \, dt + g(\tilde{X}(T)) - g(X(T)) \bigg],
\end{aligned} \tag{2.7}
$$

by the convexity of g

$$\begin{aligned}
\mathbb{E}\bigg[g(\tilde{X}(T)) - g(X(T))\bigg] &\leq \mathbb{E}\left[\left(\tilde{X}_T - X(T)\right) D_x g(\tilde{X}(T))\right] = \mathbb{E}\bigg[(\tilde{X}(T) - X(T))\tilde{p}(T)\bigg] \\
&= \mathbb{E}\bigg[\int_0^T \left[\left(\tilde{X}(t) - X(t)\right) d\tilde{p}(t) + \tilde{p}(t)\left(d\tilde{X}(t) - dX(t)\right)\right]\bigg] \\
&+ \mathbb{E}\bigg[\int_0^T \mathrm{trace}\bigg[\left(\sigma(t, \tilde{X}(t), \tilde{u}(t) - \sigma(t, X(t), u(t))\right)^\top \tilde{q}(t)\bigg] dt\bigg] \\
&= \mathbb{E}\bigg[\int_0^T \left(\tilde{X}(t) - X(t)\right)\left(-D_x\mathcal{H}(t, \tilde{X}(t), \tilde{u}(t), \tilde{p}(t), \tilde{q}(t))\right) dt\bigg] \\
&+ \mathbb{E}\bigg[\int_0^T \tilde{p}(t)\left(b(t, \tilde{X}(t), \tilde{u}(t)) - b(t, X(t), u(t))\right) dt \\
&+ \mathbb{E}\bigg[\int_0^T \mathrm{trace}\bigg[\left(\sigma(t, \tilde{X}(t), \tilde{u}(t)) - \sigma(t, X(t), u(t))\right)^\top \tilde{q}(t)\bigg] dt\bigg].
\end{aligned} \tag{2.8}$$

On the other hand we have

$$\begin{aligned}
&\mathbb{E}\bigg[\int_0^T \left(L(t, \tilde{X}(t), \tilde{u}(t)) - L(t, X(t), u(t))\right) dt\bigg] \\
&= \mathbb{E}\bigg[\int_0^T \Big(\mathcal{H}(t, \tilde{X}(t), \tilde{u}(t), \tilde{p}(t), \tilde{q}(t)) - \mathcal{H}(t, X(t), u(t), p(t), q(t))\Big) dt\bigg] \\
&- \mathbb{E}\bigg[\int_0^T \tilde{p}(t)\left(b(t, \tilde{X}(t), \tilde{u}(t)) - b(t, X(t), u(t))\right) dt\bigg] \\
&- \mathbb{E}\bigg[\int_0^T \mathrm{trace}\bigg[\left(\sigma(t, \tilde{X}(t), \tilde{u}(t)) - \sigma(t, X(t), u(t))\right)^\top \tilde{q}(t)\bigg] dt\bigg].
\end{aligned} \tag{2.9}$$

Using 2.8 and 2.9 in 2.7 yields

$$\begin{aligned}
J(0, X(0), &\tilde{u}(t)) - J(0, X(0), u(t)) \\
&\leq \mathbb{E}\bigg[\int_0^T \Big(\mathcal{H}(t, \tilde{X}(t), \tilde{u}(t), \tilde{p}(t), \tilde{q}(t)) - \mathcal{H}(t, X(t), u(t), p(t), q(t))\Big) dt\bigg] \\
&+ \mathbb{E}\bigg[\int_0^T \left(\tilde{X}(t) - X(t)\right)\left(D_x H(t, \tilde{X}(t), \tilde{u}(t), \tilde{p}(t), \tilde{q}(t))\right) dt\bigg] \leq 0,
\end{aligned}$$

which implies that $J(0, X(0), \tilde{u}(t)) \leq J(0, X(0), u(t)), \quad \forall u \in \mathcal{U}_{ad}$.

Using the definition of the minimum, we get

$$J(0, X(0), \tilde{u}(t)) = \min_{u \in \mathcal{U}_{ad}} J(0, X(0), u(t)).$$

■

2.3 Near-optimal Controls

In some optimization problems, finding the optimal control is not usually easy. This is why we tend to find a family of controls that optimizes the problem.

This approach is called finding the *near-optimal* control for the optimization problem. The optimal control can be found, and sometimes can not be found,a and it all depends on the optimization problem's from.

One of the main methods used for near-optimal controls theory is the Ekeland's principle, which we will examine in this section.

Definition 2.3.1 *(Near-optimal control) We suppose the same optimization problem 2.1 and cost functional J as in 2.4. A family of admissible controls $(u^{\varepsilon})_{\varepsilon} \subset \mathcal{U}_{ad}$ is called near-optimal if the inequality*

$$|J(0, X(0), u^{\varepsilon}(t) - \min_{u \in \mathcal{U}_{ad}} J(0, X(0), u(t)| \leq \delta(\varepsilon) \tag{2.10}$$

is verified for a sufficiently small $\epsilon > 0$, where δ is a function of ε such that $\delta(\varepsilon) \to 0$ as $\varepsilon \to 0$.

Remark 2.3.2 *If $\delta(\varepsilon) = C\varepsilon^k$ for some $C > 0$, $k > 0$, $u^{\varepsilon}(.)$ is called near-optimal with order ε^k,*

Theorem 2.3.3 *(Ekeland's variational principle) Let (E, d) be a complete metric space and define a proper, semicontinuous function $f : E \longrightarrow \mathbb{R} \cup \{+\infty\}$ that is bounded from below. Let $x_0 \in Dom(f)$ and fix $\lambda > 0$. Then there exists $\tilde{x} \in E$ such that*

$$\begin{aligned} &f(\tilde{x}) + \lambda d(\tilde{x}, x_0) \leq f(x_0), \\ &f(\tilde{x}) < f(x) + \lambda d(\tilde{x}, x), \qquad \forall x \neq \tilde{x} \end{aligned} \tag{2.11}$$

Corollary 2.3.4 *Suppose the assumptions in 2.3.3 hold. Let* $\varepsilon > 0$*, and* $x_0 \in E$ *such that*

$$f(x_0) \leq \inf_{x \in E} f(x) + \varepsilon. \tag{2.12}$$

Then there exists $x_\varepsilon \in E$ *such that*

$$f(x_\varepsilon) \leq f(x_0), \qquad d(x_\varepsilon, x_0) \leq \sqrt{\varepsilon}, \tag{2.13}$$

and for all $x \in E$

$$f(x_\varepsilon) \leq f(x) + \sqrt{\varepsilon} d(x_\varepsilon, x). \tag{2.14}$$

Proof. We take $\lambda = \sqrt{\varepsilon}$. By applying Ekeland's principle, there exists $x_\varepsilon \in E$ such that

$$f(x_\varepsilon) \leq f(x_\varepsilon) + \sqrt{\varepsilon} d(x_\varepsilon, x_0) \leq f(x_0) \leq \inf_{x \in E} f(x) + \varepsilon \leq f(x_\varepsilon) + \varepsilon,$$

then

$$f(x_\varepsilon) + \sqrt{\varepsilon} d(x_\varepsilon, x_0) \leq f(x_\varepsilon) + \varepsilon,$$

we find that $d(x_\varepsilon, x_0) \leq \sqrt{\varepsilon}$.

Finally, 2.14 is a direct result of 2.11

$$f(x_\varepsilon) < f(x) + \sqrt{\varepsilon} d(x_\varepsilon, x), \quad \forall x \neq x_\varepsilon.$$

∎

Theorem 2.3.5 *The function*

$$\begin{aligned} d : \mathcal{U}_{ad} \times \mathcal{U}_{ad} &\longrightarrow \mathbb{R}_+, \\ d(u, u') &= \mathbb{E}\Big[\lambda \left\{ t \in [0, T] / u(t) \neq u'(t) \right\} \Big], \end{aligned}$$

defines a metric on $\mathcal{U}_{ad}$*, where* λ *represents the Lebesgue measure.*

Proof. The axioms of the metric d can be deduced from the properties of the Lebesgue measure.

∎

Theorem 2.3.6 *Under the metric d, $\mathcal{U}_{ad}$ is a complete meric space.*

Proof. Let $(u_n)_{n\in\mathbb{N}}$ be a Cauchy sequence in $(\mathcal{U}_{ad}, d)$. Then there exists a converging subsequence $(u_{n_k})_{k\in\mathbb{N}}$ such that

$$d(u_{n_k}, u_{n_{k+1}}) \leq 2^{-k}, \qquad \forall k \geq 2.$$

We define the following sets

$$\begin{cases} E_{nm} = \{(t,\omega) \in [0,T] \times \Omega / u_n(t,\omega) \neq u_m(t,\omega)\}, & , n \geq 1, m \geq 1, \\ A_k = \bigcup_{p\geq k} E_{n_p, n_{p+1}} \quad , k \geq 2, \end{cases}$$

the sequence $(A_k)_{k\geq 2}$ is decreasing $(A_{k+1} \subset A_k), \forall k \geq 1$, plus

$$|A_k| \leq \sum_{p=k}^{+\infty} 2^{-p} = 2^{1-k}, \qquad k \geq 2,$$

which leads to $\left|\bigcup_{k\geq 1} A_k^c\right| = T$. Now we define

$$\tilde{u}(t,\omega) = u_{n_k}(t,\omega), \qquad t \in A_k^c, \quad k \geq 2.$$

The control $\tilde{u}$ is well defined and is an admissible control. As a result

$$d(u_{n_k}, \tilde{u}) \leq |A_k| \leq 2^{1-k} \longrightarrow 0,$$

therefore

$$d(u_n, \tilde{u}) \longrightarrow 0,$$

concluding the proof. ■

2.4 Regime Switching Stochastic System

A stochastic system is said to be *regime switching* if at any given time t, it changes its behaviour in an "abrupt" way. In other words, the system tends to totally change its state in a given time horizon $[0, T]$.

The stock market is a good example where regime switching takes place : It often exhibits dramatic breaks in their behaviour, associated with events such as financial crises, or abrupt changes in government policy.[1]

Justifying the use of regime switching property in the medical field is subject to the following explanation : During an epidemic, the virus can behave differently depending on the weather condition. Any suddent change in the weather can lead to a critical change in many other variables such as the number of deaths (caused by other diseases such as flu in winter) or the transmission rate that becomes either high or low.

Notice that these changes, once they happen, the system no longer depends on its past state, it only depends on the present, which leads us to having a "Memoryless" system, from which we deduce the Markov property.

We can say that the stochastic system, in this case, depends also on a Markov chain that is responsible for its regime switchings.

Definition 2.4.1 *Let $\xi = (\xi(t))_{t\geq 0}$ be a Continuous-time Markov chain with a finite state space $\mathbb{S}$.*

We define a regime switching stochastic system as follows

$$\begin{cases} dX(t) = b(t, X(t), \xi(t))dt + \sigma(t, X(t), \xi(t))dB(t), \\ X(0) = x_0 \in \mathbb{R}^n, \xi_0 = \overline{w} \in \mathbb{S}, \end{cases} \tag{2.15}$$

where B is a d-dimensional brownian motion, and

$$b : [0, T] \times \Omega \times \mathbb{R}^n \times \mathbb{S} \longrightarrow \mathbb{R}^n, \qquad \sigma : [0, T] \times \Omega \times \mathbb{R}^n \times \mathbb{S} \longrightarrow \mathbb{R}^{n\times d}.$$

[1] See [2].

For any function $f \in C^2(\mathbb{R}^n \times \mathbb{S})$, the associated infinitesimal operator is given by the following formula

$$\mathcal{L}f(x,k) = \sum_{j=1}^{n} b_j(x,i)\frac{\partial f(x,i)}{\partial x_j} + \frac{1}{2}\sum_{j=1}^{n}\sum_{k=1}^{n} a_{jk}(x,i)\frac{\partial^2 f(x,i)}{\partial x_i \partial x_j} + \sum_{\substack{(j,k)\in\mathbb{S}^2 \\ j\neq k}}^{n} q_{kj}(f(x,j) - f(x,k)),$$

where $a(x,i) = \sigma(x,i)\sigma^\top(x,i)$.

2.4.1 Existence and Uniqueness theorem

In what follows, we will state some basic results on regime switching SDEs.

Let $t \in [0,T]$, denote by $\mathfrak{B}$ the $\mathcal{F}_t$-predictable σ-field on $[0,T] \times \mathcal{F}$. For any given $s \in [0,T]$. We denote $S^2_{\mathcal{F}}([t,T];\mathbb{R}^n)$ the set of all $(\mathcal{F}_s)_{s\in[t,T]}$-adapted, càdlag processes X such that

$$\mathbb{E}\left[\sup_{s\in[t,T]} |X(s)|^2\right] < +\infty.$$

Consider now the following SDE

$$X(t) = \alpha + \int_s^t b(r, X(r), \xi(r))dr + \int_s^t \sigma(r, X(r), \xi(r))dB(r), \tag{2.16}$$

where $s \leq t \leq T$. Here the coefficients (α, b, σ) are given mappings $\alpha : \Omega \longrightarrow \mathbb{R}^n, b : [0,T] \times \Omega \times \mathbb{R}^n \times \mathbb{S} \longrightarrow \mathbb{R}^n, \sigma : [0,T] \times \Omega \times \mathbb{R}^n \times \mathbb{S} \longrightarrow \mathbb{R}^{n\times d}$, satisfying the assumptions below

(H1) $\alpha \in \mathbb{L}^2(\Omega, \mathcal{F}_t; \mathbb{R}^n)$ and the coefficients b, σ are $\mathfrak{B} \otimes \mathcal{B}(\mathbb{R}^n) \otimes \mathcal{B}(\mathbb{S})$ measurable with: for all $w_i \in \mathbb{S}$

$$\mathbb{E}\left[\int_0^T (b(t,0,w_i) + \sigma(t,0,w_i))\, dt\right] < \infty.$$

(H2) b, σ are uniformly Lipschitz continuous with respect to x, that is, there exists a constant $C > 0$ such that for all $(t, x, \bar{x}, w_i) \in [0,T] \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{S}$ and a.s. $\omega \in \Omega$

$$|b(t,x,w_i) - b(t,\bar{x},w_i)|^2 + \|\sigma(t,x,w_i) - \sigma(t,\bar{x},w_i)\|^2 \leqslant C|x - \bar{x}|^2.$$

Theorem 2.4.2 *If the coefficients (α, b, σ) satisfy the assumptions* ***(H1)-(H2)****, then the SDE*

2.16 has a unique solution $X(\cdot) \in S^2_{\mathcal{F}}(s,T;\mathbb{R}^n)$*. Moreover,* $\exists K > 0$ *such that*

$$\mathbb{E}\left[\sup_{s\leq t\leq T}|X(s)|^2\right] \leq K\left(1+\mathbb{E}\left[|\alpha|^2\right]\right)$$

Proof.

Let $0 = \tau_0 < \tau_1 < \tau_2 < \ldots, < \tau_n < \ldots$ be the jump times of the Markov chain $\xi(\cdot)$, and let $w_1 \in \mathbb{S}$ be the starting state. Thus $\xi(t) = w_1$ on $[\tau_0, \tau_1)$, and the system 2.16 for $t \in [\tau_0, \tau_1[$ has the following form:

$$dX(t) = b(t, X(t), w_1)dt + \sigma(t, X(t), w_1)dB(t),$$

By theorem 1.4.3, the above SDE has the unique solution $X(\cdot)$ on the space $\mathcal{S}^2_{\mathcal{F}}([\tau_0, \tau_1[;\mathbb{R}^n)$, and by continuity for $t = \tau_1$, as well.

Now, by considering $\xi(\tau_1) = w_2$, the system for $t \in [\tau_1, \tau_2[$ becomes

$$dX(t) = b(t, X(t), w_2)dt + \sigma(t, X(t), w_2)dB(t), \tag{2.17}$$

By the same theorem 1.4.3, the SDE 2.17 has a unique solution $X(\cdot) \in \mathcal{S}^2_{\mathcal{F}}([\tau_1, \tau_2[;\mathbb{R}^n)$, and by continuity for $t = \tau_2$, as well. Repeating this process continuously, we get the same result : The solution $X(\cdot)$ of system 2.16 remains in $\mathcal{S}^2_{\mathcal{F}}(0,T;\mathbb{R}^n)$ with probability one.

■

Chapter 3

Regime Switching Stochastic Systems : Application in Viral Models

In order to show the important role of the optimal control theory in epidemic models, we propose an approach to a SIRS model.

3.1 Stochastic SIRS Model

A SIRS (Suspected-Infected-Recovered-Suspected) model can be explained as follows : We divide a population $N(t)$ into three groups : Suspected S, Infected I, and Recovered R. Every suspected person who catches the virus becomes infected and thus, gets vaccinated and goes into treatment. When treated successfully, the infected person recovers from the virus, but not for a long time before they come back to being suspected (Healthy carrier).

The stochastic SIRS model is given by the following equation

$$\begin{cases} dS(t) = \left((1-p)b - \mu_1 S(t) - \dfrac{\beta S(t)I(t)}{\varphi(I(t))} + \gamma R(t)\right)dt - \sigma_1 S(t)dB_1(t) - \sigma_4 \dfrac{S(t)I(t)}{\varphi(I(t))}dB_4(t) \\ dI(t) = \left(-(\mu_2 + c + \alpha)I(t) + \dfrac{\beta S(t)I(t)}{\varphi(I(t))}\right)dt - \sigma_2 I(t)dB_2(t) + \sigma_4 \dfrac{S(t)I(t)}{\varphi(I(t))}dB_4(t) \\ dR(t) = (pb - (\mu_3 + \gamma)R(t) + \alpha I(t))dt - \sigma_3 R(t)dB_3(t), \end{cases} \tag{3.1}$$

where $B = (B_1, B_2, B_3, B_4)$ is a 4-dimensional standard brownian motion, and the rest of the parameters are explained in Table 3.1

Notations	Signification
p	Propotion of vaccinated population
b	Birth rate of the population
β	Per capita transmission rate
$\mu_i(i = 1, 2, 3)$	Natural death rate of every class
γ	Per capita immunity loss rate
c	Per capita disease-induced death rate
α	Per capita recovery rate
$\sigma_i^2(i = 1, 2, 3, 4)$	Intensities of the white noises
$S(t)$	The suspectible class
$I(t)$	The infected class
$R(t)$	The recovered class

Table 3.1: Notations used in the model

In order to control this system, we need to introduce a control function to the system. Let $u(t) = (u_1(t), u_2(t))$ be the control function of vaccination and treatment respectively. We define the treatment function T as follows

$$T(u_2(t), I(t)) = \frac{mu_2(t)I(t)}{1 + \eta I(t)},$$

where $m \geq 0$ is the cure rate, $\eta \geq 0$ is the delay in treatment.

We obtain a new controlled system, that is

$$\begin{cases} dS(t) = \left((1-p)b - \mu_1 S(t) - \dfrac{\beta S(t)I(t)}{\varphi(I(t))} + \gamma R(t)\right)dt - \sigma_1 S(t)dB_1(t) - \sigma_4 \dfrac{S(t)I(t)}{\varphi(I(t))}dB_4(t) \\ dI(t) = \left(-(\mu_2 + c + \alpha)I(t) + \dfrac{\beta S(t)I(t)}{\varphi(I(t))} - \dfrac{mu_2(t)I(t)}{1+\eta I(t)}\right)dt - \sigma_2 I(t)dB_2(t) \\ \qquad +\sigma_4 \dfrac{S(t)I(t)}{\varphi(I(t))}dB_4(t) \\ dR(t) = \left(pb - (\mu_3 + \gamma)R(t) + \alpha I(t) + u_1(t)S(t) + \dfrac{mu_2(t)I(t)}{1+\eta I(t)}\right)dt - \sigma_3 R(t)dB_3(t), \end{cases}$$

where $\dfrac{\beta S(t)I(t)}{\varphi(I(t))}$ is the incident rate, defined as the ratio of a population, yet unaffected by a disease, that develops it, becomes infected, or dies during a limited time horizon.

The positive function φ verifies $\varphi(0) = 1$ and $\varphi'(I) > 0$.

The goal now is to minimize the suspected and the infected classes using minimal control efforts.

Some logical restrictions can be added to the control function such as having $u_1(t) \in [0, 1[$ $\forall t \in [0, T]$, since we can't get everyone vaccinated at once, plus having $u_2(t) \in [0, 1]$, $\forall t \in [0, T]$ where $u_2(t) = 0$ refers to a total absence of treatment, and $u_2(t) = 1$ refers to a fully effective treatment.

We can then easily consider our control function to be $u : [0, T] \longrightarrow \Gamma = [0, 1[\times[0, 1]$.

The objective function associated to the system is

$$J(0, S(0), I(0), R(0), u(t)) = \mathbb{E}\left[\int_0^T L(t, S(t), I(t), R(t), u(t))dt + h(x(T))\right], \tag{3.2}$$

where

$$L(t, S(t), I(t), R(t), u(t) = A_1 S(t) + A_2 I(t) + \frac{1}{2}(\tau_1 u_1^2(t) + \tau_2 u_2^2(t)),$$
$$h(x(T)) = (0, I(T), 0),$$

with $\tau_1 \geq 0$,$\tau_2 \geq 0$, and A_1, A_2 are two positive constants to keep balance between the suspected and the infected classes.

We then introduce the continuous-time Markov chain $\xi = (\xi(t))_{t\geq 0}$ on a finite state space $\mathbb{S}$.

To simplify the writing, let

$$x(t) = (x_1(t), x_2(t), x_3(t)) \triangleq (S(t), I(t), R(t)),$$

and let

$$\varphi(x) = 1 + x^2.$$

It's clear that φ is positive, with positive derivative, and $\varphi(0) = 1$.

We obtain the system

$$\begin{cases} dx_1(t) = & \left((1 - p(\xi(t)))b(\xi(t)) - \mu_1(\xi(t))x_1(t) - \dfrac{\beta(\xi(t))x_1(t)x_2(t)}{1 + (x_2^2(t))} + \gamma(\xi(t))x_3(t)\right)dt \\ & -\sigma_1(\xi(t))x_1(t)dB_1(t) - \sigma_4(\xi(t))\dfrac{x_1(t)x_2(t)}{1 + (x_2^2(t))}dB_4(t) \\ & = f_1(t, x(t), u(t))dt + \sigma_{1,4}(x(t))dB(t) \\ dx_2(t) = & \left(-\big(\mu_2(\xi(t)) + c(\xi(t)) + \alpha(\xi(t))\big)x_2(t) + \dfrac{\beta(\xi(t))x_1(t)x_2(t)}{1 + (x_2^2(t))} - \dfrac{m(\xi(t))u_2(t)x_2(t)}{1 + \eta(\xi(t))x_2(t)}\right)dt \\ & -\sigma_2(\xi(t))x_2(t)dB_2(t) + \sigma_4(\xi(t))\dfrac{x_1(t)x_2(t)}{1 + (x_2^2(t))}dB_4(t) \\ & = f_2(t, x(t), u(t))dt + \sigma_{2,4}(x(t))dB(t) \\ dx_3(t) = & \left(p(\xi(t))b(\xi(t)) - \big(\mu_3(\xi(t)) + \gamma(\xi(t))\big)x_3(t) + \alpha(\xi(t))x_2(t) + u_1(t)x_1(t)\right. \\ & \left. + \dfrac{m(\xi(t))u_2(t)x_2(t)}{1 + \eta(\xi(t))x_2(t)}\right)dt - \sigma_3(\xi(t))x_3(t)dB_3(t) \\ & = f_3(t, x(t), u(t))dt + \sigma_{3,4}(x(t))dB(t). \end{cases} \tag{3.3}$$

3.2 Hypothesis

We define the functions

$$f(t, x(t), u(t)) = \begin{pmatrix} f_1(t, x(t), u(t)) \\ f_2(t, x(t), u(t)) \\ f_3(t, x(t), u(t)) \end{pmatrix}, \quad \sigma_*(x(t)) = \begin{pmatrix} \sigma_{1,4}(x(t)) \\ \sigma_{2,4}(x(t)) \\ \sigma_{3,4}(x(t)) \end{pmatrix}.$$

- **(H1)** The functions

$$f:[0,T]\times\Omega\times\mathbb{R}^3\times\Gamma\to\mathbb{R}^3,\quad \sigma_*:[0,T]\times\Omega\times\mathbb{R}^3\times\Gamma\to\mathbb{R}^3\times\mathbb{R}^4,\quad L:[0,T]\times\Omega\times\mathbb{R}^3\times\Gamma\to\mathbb{R},$$

  are measurable in (t,x,u), twice continuously differentiable in x for every (t,u).

- **(H2)** The function $h:\mathbb{R}^3\to\mathbb{R}$ is twice continuously differentiable, and there exists $C>0$ such that $\forall(x,x')\in\mathbb{R}^3\times\mathbb{R}^3$

$$|h(x)|\le C(1+|x|),$$
$$|h(x)-h(x')|+|D_xh(x)-D_xh(x')|\le C|x-x'|.$$

- **(H3)** The set of admissible controls $\mathcal{U}_{ad}$ is convex.

- **(H4)** We suppose that $\forall t\in[0,T],\forall k\in\mathbb{S}$

$$\begin{aligned}K=&\frac{(1-p(k))b(k)}{x_1(t)}-\frac{\beta(k)x_2(t)}{1+x_2^2(t)}+\frac{\gamma(k)x_3(k)}{x_1(t)}-u_1(t)-\frac{1}{2}\sigma_4^2(t)\frac{x_2^2(t)}{(1+x_2^2(t))^2}\\&+\frac{\beta(k)x_1(t)}{1+x_2^2(t)}-\frac{1}{2}\sigma_4^2(t)\frac{x_1^2(t)}{(1+x_2^2(t))^2}+\frac{p(k)b(k)}{x_3(t)}-\frac{m(k)u_2(t)}{1+\eta(k)x_2(t)}\\&+\frac{u_1(t)x_1(t)}{x_3(t)}+\frac{m(k)u_2(t)x_2(t)}{(1+\eta(k)x_2(t)x_3(t)}+\frac{\alpha(k)x_2(t)}{x_3(t)}<0.\end{aligned}$$

- **(H5)** We suppose

$$\Pi=\sum_{k\in\mathbb{S}}\pi_k\left(\mu_1(k)+\frac{1}{2}\sigma_1^2(k)+\mu_2(k)+\alpha(k)+\frac{1}{2}\sigma_2^2(k)+\mu_3(k)+\gamma(k)+\frac{1}{2}\sigma_3^2(k)\right)>0.$$

3.3 Sufficient Conditions for Near-optimal Controls

3.3.1 Estimates On The Parameters

Theorem 3.3.1 *Suppose $\theta>0$, then $\forall t\in[0,T]$, $\exists C>0$, such that:*

$$\mathbb{E}\left[\sup_{0\le t\le T}|x_i(t)|^\theta\right]\le C,\qquad i=1,2,3. \tag{3.4}$$

Proof. See [6] ■

We introduce now the following adjoint equation

$$\begin{cases} dp_1(t) & = -b_1(x(t),u(t),p(t),q(t))dt + q_1(t)dB(t) \\ dp_2(t) & = -b_2(x(t),u(t),p(t),q(t))dt + q_2(t)dB(t) \\ dp_3(t) & = -b_3(x(t),u(t),p(t),q(t))dt + q_3(t)dB(t) \\ p_i(T) & = D_{x_i}h(x(T)), \quad i=1,2,3, \end{cases} \tag{3.5}$$

where

$$\begin{aligned} & b_1(x(t),u(t),p(t),q(t)) \\ & \quad = -\left(\mu_1(\xi(t)) + \frac{\beta(\xi(t))x_2(t)}{1+x_2^2(t)} + u_1(t)\right)p_1(t) + \frac{\beta(\xi(t))x_2(t)}{1+x_2^2(t)}p_2(t) + u_1(t)p_3(t) \\ & \quad - \left(\frac{\sigma_4(\xi(t))x_2(t)}{1+x_2^2(t)}\right)q_1(t) + \frac{\sigma_4(\xi(t))x_2(t)}{1+x_2^2(t)}q_2(t) + A_1 \\ & b_2(x(t),u(t),p(t),q(t)) \\ & \quad = -\left(\frac{\beta(\xi(t))x_1(t)(1-x_2^2(t))}{(1+x_2^2(t))^2}\right)p_1(t) + \left(\frac{\beta(\xi(t))x_1(t)(1-x_2^2(t))}{(1+x_2^2(t))^2} - (\mu_2(\xi(t)) + c(\xi(t)) + \alpha(\xi(t)))\right. \\ & \quad \left. - \frac{m(\xi(t))u_2(t)}{(1+\eta(\xi(t))x_2(t))^2}\right)p_2(t) + \left(\alpha(\xi(t)) + \frac{m(\xi(t))u_2(t)}{(1+\eta(\xi(t))x_2(t))^2}\right)p_3(t) \\ & \quad - \frac{\sigma_4(\xi(t))x_1(t)(1-x_2^2(t))}{(1+x_2^2(t))^2}q_1(t) - \left(\sigma_2(\xi(t)) - \frac{\sigma_4(\xi(t))x_1(t)(1-x_2^2(t))}{(1+x_2^2(t))^2}\right)q_2(t) + A_2 \\ & b_3(x(t),u(t),p(t),q(t)) \\ & \quad = \gamma(\xi(t))p_1(t) - (\mu_3(\xi(t)) + \gamma(\xi(t))p_3(t) - \sigma_3(\xi(t))q_3(t). \end{aligned}$$

The adjoint function represents a special case of BSDEs with nonlinear coefficients. Thus we need some estimates on the pair (p,q).

Theorem 3.3.2 *Under (H1)*

$$\sum_{i=1}^{3}\mathbb{E}\left[\sup_{0\le t\le T}|p_i(t)|^2\right] + \sum_{i=1}^{3}\mathbb{E}\left[\int_0^T |q_i(t)|^2 dt\right] \le C. \tag{3.6}$$

Proof. See [6] ■

3.3.2 Sufficient Conditions for Near-optimal Controls

We define the Hamiltonian function $\mathcal{H}$ as follows

$$\mathcal{H}(t,x(t),u(t),p(t),q(t)) \triangleq f^{\top}(t,x(t),u(t))p(t) + \sigma_*(x(t))q(t)) + L(t,x(t),u(t)),$$

Theorem 3.3.3 *Let $(x^\varepsilon, u^\varepsilon)$ be an admissible pair, $(p^\varepsilon, q^\varepsilon)$ be a solution to the adjoint equation. Assume the hamiltonian $\mathcal{H}$ is convex. If for some $\varepsilon > 0$*

$$\mathbb{E}\left[\int_0^T \mathcal{H}(t,x^\varepsilon(t),u(t),p^\varepsilon(t),q^\varepsilon(t))dt\right] \geq \sup_{u^\varepsilon \in \mathcal{U}_{ad}[0,T]} \mathbb{E}\left[\int_0^T \mathcal{H}(t,x^\varepsilon(t),u^\varepsilon(t),p^\varepsilon(t),q^\varepsilon(t))dt\right] - \varepsilon, \tag{3.7}$$

$$\begin{aligned}
&\mathbb{E}\Bigg[\int_0^T \Bigg(u_1(t)x_1^\varepsilon(t)(p_3^\varepsilon(t) - p_1^\varepsilon(t)) + \frac{m(\xi(t))u_2(t)x_2^\varepsilon(t)}{1+\eta(\xi(t))x_2^\varepsilon(t)}(p_3^\varepsilon(t) - p_2^\varepsilon(t)) + \frac{1}{2}(\tau_1 u_1^2(t) + \tau_2 u_2^2(t))\Bigg)dt\Bigg] \\
&\qquad \geq \sup_{u^\varepsilon \in \mathcal{U}_{ad}[0,T]} \mathbb{E}\Bigg[\int_0^T \Bigg(u_1^\varepsilon(t)x_1^\varepsilon(t)(p_3^\varepsilon(t) - p_1^\varepsilon(t)) + \frac{m(\xi(t))u_2^\varepsilon(t)x_2^\varepsilon(t)}{1+\eta(\xi(t))x_2^\varepsilon(t)}(p_3^\varepsilon(t) - p_2^\varepsilon(t)) \\
&\qquad + \frac{1}{2}(\tau_1 (u_1^\varepsilon)^2(t) + \tau_2 (u_2^\varepsilon)^2(t))\Bigg)dt\Bigg] - \varepsilon,
\end{aligned}$$

then

$$J(0,x(0),u^\varepsilon(t)) \leq \inf_{u \in \mathcal{U}_{ad}[0,T]} J(0,x(0),u(t)) + C\varepsilon^{\frac{1}{2}}. \tag{3.8}$$

Proof. We define the metric $\tilde{d}$ on $\mathcal{U}_{ad}$: For any $\varepsilon > 0$ and for any $(u,v) \in \mathcal{U}_{ad} \times \mathcal{U}_{ad}$

$$\tilde{d}(u,v) = \mathbb{E}\left[\int_0^T y^\varepsilon(t)|u(t) - v(t)|dt\right], \tag{3.9}$$

where

$$y^\varepsilon(t) = 1 + \sum_{i=1}^{3} |p_i^\varepsilon(t)| + \sum_{i=1}^{3} |q_i^\varepsilon(t)|.$$

Using the definitions of both the Hamiltonian function and the cost function, we can find the following decomposition:

$$J(0,x^{\varepsilon}(0),u^{\varepsilon}(t)) - J(0,x(0),u(t)) = I_1 + I_2 - I_3,$$

where

$$\begin{aligned}
I_1 &= \mathbb{E}\bigg[\int_0^T \mathcal{H}(t,x^{\varepsilon}(t),u^{\varepsilon}(t),p^{\varepsilon}(t),q^{\varepsilon}(t)) - \mathcal{H}(t,x(t),u(t),p^{\varepsilon}(t),q^{\varepsilon}(t))\bigg]dt,\\
I_2 &= \mathbb{E}\bigg[h(x^{\varepsilon}(T)) - h(x(T))\bigg],\\
I_3 &= \mathbb{E}\bigg[\int_0^T \Big[\big(f^{\top}(t,x^{\varepsilon}(t),u^{\varepsilon}(t)) - f^{\top}(t,x(t),u(t))\big)p^{\varepsilon}(t) + \big(\sigma_*^{\top}(x^{\varepsilon}(t)) - \sigma_*^{\top}(x(t))\big)q^{\varepsilon}(t)\Big]dt.
\end{aligned}$$

Using the convexity of the hamiltonian function and h we get:

$$I_2 \leq \sum_{i=1}^{3}\mathbb{E}\Big[D_x h(x^{\varepsilon}(T))(x_i^{\varepsilon}(T) - x_i(T)\Big],$$

$$\begin{aligned}
I_1 \leq &\sum_{i=1}^{3}\mathbb{E}\bigg[\int_0^T D_{x_i}\mathcal{H}(t,x^{\varepsilon}(t),u^{\varepsilon}(t),p^{\varepsilon}(t),q^{\varepsilon}(t))(x_i^{\varepsilon}(t) - x_i(t))dt\bigg]\\
&+\sum_{i=1}^{2}\mathbb{E}\bigg[\int_0^T D_{u_i}\mathcal{H}(t,x^{\varepsilon}(t),u^{\varepsilon}(t),p^{\varepsilon}(t),q^{\varepsilon}(t))(u_i^{\varepsilon}(t) - u_i(t))dt\bigg].
\end{aligned}$$

We define the following function $\forall t \in [0,T], \forall k \in \mathbb{S}$

$$\begin{aligned}
V(x(t),p(t),q(t),k) &= \sum_{i=1}^{3} p_i^{\varepsilon}(t)(x_i^{\varepsilon}(t) - x_i(t)) + \sum_{i=1}^{3}\ln x_i(t) + (\bar{w}_k + |\bar{w}|)\\
&= V_1(x(t),p(t),q(t) + V_2(x(t)) + V_3(k).
\end{aligned}$$

Applying the linear operator $\mathcal{L}$ on V, we get

$$\begin{aligned}
\mathcal{L}V_1(x(t),p(t),q(t)) &= \sum_{i=1}^{3} D_{x_i}\mathcal{H}(t,x^{\varepsilon}(t),u^{\varepsilon}(t),p^{\varepsilon}(t),q^{\varepsilon}(t))(x_i^{\varepsilon}(t) - x_i(t))\\
&+\sum_{i=1}^{3} p_i^{\varepsilon}(t)\left|f_i(t,x^{\varepsilon}(t),u^{\varepsilon}(t)) - f_i(t,x(t),u(t))\right|\\
&+\sum_{i=1}^{3} q_i^{\varepsilon}(t)\left|\sigma_{i4}(x^{\varepsilon}(t)) - \sigma_{i4}(x(t))\right|.
\end{aligned} \tag{3.10}$$

$$\begin{aligned}
\mathcal{L}V_2(x(t)) &= \frac{(1-p(k))b(k)}{x_1(t)} - \frac{\beta(k)x_2(t)}{1+x_2^2(t)} + \frac{\gamma(k)x_3(k)}{x_1(t)} - u_1(t) - \frac{1}{2}\sigma_4^2(t)\frac{x_2^2(t)}{(1+x_2^2(t))^2} + \frac{\beta(k)x_1(t)}{1+x_2^2(t)} \\
&- \frac{1}{2}\sigma_4^2(t)\frac{x_1^2(t)}{(1+x_2^2(t))^2} + \frac{p(k)b(k)}{x_3(t)} - \frac{m(k)u_2(t)}{1+\eta(k)x_2(t)} + \frac{u_1(t)x_1(t)}{x_3(t)} + \frac{m(k)u_2(t)x_2(t)}{(1+\eta(k)x_2(t)x_3(t)} \\
&+ \frac{\alpha(k)x_2(t)}{x_3(t)} - \bigg(\mu_1(k) + \frac{1}{2}\sigma_1^2(k) + \mu_2(k) + \alpha(k) + \frac{1}{2}\sigma_2^2(k) + \mu_3(k) + \gamma(k) + \frac{1}{2}\sigma_3^2(k)\bigg) \\
&= K - \bigg(\mu_1(k) + \frac{1}{2}\sigma_1^2(k) + \mu_2(k) + \alpha(k) + \frac{1}{2}\sigma_2^2(k) + \mu_3(k) + \gamma(k) + \frac{1}{2}\sigma_3^2(k)\bigg).
\end{aligned} \tag{3.11}$$

$$\mathcal{L}V_3(k) = \sum_{l\in\mathbb{S}} q_{kl}\bar{w}_l \tag{3.12}$$

We then have

$$\sum_{l\in\mathbb{S}} q_{kl}\bar{w}_l - \bigg(\mu_1(k) + \frac{1}{2}\sigma_1^2(k) + \mu_2(k) + \alpha(k) + \frac{1}{2}\sigma_2^2(k) + \mu_3(k) + \gamma(k) + \frac{1}{2}\sigma_3^2(k)\bigg) = -\Pi < 0$$

We finally obtain

$$\begin{aligned}
\mathcal{L}V(x(t),p(t),q(t),k) &= \mathcal{L}V_1(x(t),p(t),q(t),k) + \mathcal{L}V_2(x(t)) + \mathcal{L}V_3(k) \\
&= -\Pi + K - \sum_{i=1}^{3} D_{x_i}\mathcal{H}(t,x^\varepsilon(t),u^\varepsilon(t),p^\varepsilon(t),q^\varepsilon(t))(x_i^\varepsilon(t) - x_i(t)) \\
&+ \sum_{i=1}^{3} p_i^\varepsilon(t)\left|f_i(t,x^\varepsilon(t),u^\varepsilon(t)) - f_i(t,x(t),u(t))\right|. \\
&+ \sum_{i=1}^{3} q_i^\varepsilon(t)\left|\sigma_{i4}(x^\varepsilon(t)) - \sigma_{i4}(x(t))\right|
\end{aligned}$$

Integrating both sides on $[0,T]$ and using the expectation, we get

$$\begin{aligned}
&\sum_{i=1}^{3}\mathbb{E}\bigg[D_x h(x^\varepsilon(T))(x_i^\varepsilon(T) - x_i(T))\bigg] \\
&\quad\le -\sum_{i=1}^{3}\mathbb{E}\bigg[\int_0^T D_{x_i}\mathcal{H}(t,x^\varepsilon(t),u^\varepsilon(t),p^\varepsilon(t),q^\varepsilon(t))(x_i^\varepsilon(t) - x_i(t))dt\bigg] \\
&\quad+ \sum_{i=1}^{3}\mathbb{E}\bigg[\int_0^T p_i^\varepsilon(t)|f_i(t,x^\varepsilon(t),u^\varepsilon(t)) - f_i(t,x(t),u(t))|dt\bigg] \\
&\quad+ \sum_{i=1}^{3}\mathbb{E}\bigg[\int_0^T q_i^\varepsilon(t)|\sigma_{i4}(x^\varepsilon(t)) - \sigma_{i4}(x(t))|dt\bigg].
\end{aligned}$$

We finally get that

$$J(0, x^{\varepsilon}(0), u^{\varepsilon}(t)) - J(0, x(0), u(t)) \leq \sum\nolimits_{i=1}^{2} \mathbb{E}\left[\int_0^T \tau u_i^{\varepsilon}(t)(u_i^{\varepsilon}(t) - u_i(t))dt\right].$$

We define now a function $F : \mathcal{U}_{ad} \to \mathbb{R}$

$$F(u(t)) = \mathbb{E}\left[\int_0^T \mathcal{H}(t, x^{\varepsilon}(t), u(t), p^{\varepsilon}(t), q^{\varepsilon}(t))dt\right]. \tag{3.13}$$

Since F is continuous on $\mathcal{U}_{ad}$, and using Ekeland's principle 2.3.3, we find that if there exists $\tilde{u} \in \mathcal{U}_{ad}$, then $\forall u \in \mathcal{U}_{ad}$

$$\tilde{d}(u^{\varepsilon}, \tilde{u}^{\varepsilon}) \leq \varepsilon^{\frac{1}{2}} \text{ and } F(\tilde{u}^{\varepsilon}(t)) \leq F(u(t)) + \varepsilon^{\frac{1}{2}}\tilde{d}(u(t), \tilde{u}^{\varepsilon}(t)), \tag{3.14}$$

which yields

$$\mathcal{H}(t, x^{\varepsilon}(t), \tilde{u}^{\varepsilon}(t), p^{\varepsilon}(t), q^{\varepsilon}(t)) = \min_{u \in \mathcal{U}_{ad}}\left[\mathcal{H}(t, x^{\varepsilon}(t)u(t), p^{\varepsilon}(t), q^{\varepsilon}(t)) + \varepsilon^{\frac{1}{2}} y^{\varepsilon}(t)|u(t) - \tilde{u}^{\varepsilon}(t)|\right].$$

Finally, from 3.3.2, and by using Clarke's generalized gradient A.7, we get

$$0 \in \partial_u \mathcal{H}(t, x^{\varepsilon}(t), \tilde{u}^{\varepsilon}(t), p^{\varepsilon}(t), q^{\varepsilon}(t)) \subset \partial_u \mathcal{H}(t, x^{\varepsilon}(t), \tilde{u}^{\varepsilon}(t), p^{\varepsilon}(t), q^{\varepsilon}(t)) + [-\varepsilon^{\frac{1}{2}} y^{\varepsilon}(t), \varepsilon^{\frac{1}{2}} y^{\varepsilon}(t)],$$

which means that if there exists $\lambda_1^{\varepsilon}(t) \in [-\varepsilon^{\frac{1}{2}} y^{\varepsilon}(t), \varepsilon^{\frac{1}{2}} y^{\varepsilon}(t)]$, then

$$\sum_{i=1}^{2} \tau u_i^{\varepsilon}(t) + \lambda_1^{\varepsilon}(t) = 0.$$

As a consequence, we obtain

$$\begin{aligned}
\left|D_u \mathcal{H}(t, x^{\varepsilon}(t), u^{\varepsilon}(t), p^{\varepsilon}(t), q^{\varepsilon}(t))\right| &\leq \left|D_u \mathcal{H}(t, x^{\varepsilon}(t), \tilde{u}^{\varepsilon}(t), p^{\varepsilon}(t), q^{\varepsilon}(t))\right| \\
&\quad + \left|D_u \mathcal{H}(t, x^{\varepsilon}(t), u^{\varepsilon}(t), p^{\varepsilon}(t), q^{\varepsilon}(t)) - D_u \mathcal{H}(t, x^{\varepsilon}(t), \tilde{u}^{\varepsilon}(t), p^{\varepsilon}(t), q^{\varepsilon}(t))\right| \\
&\leq C y^{\varepsilon}(t)|u^{\varepsilon}(t) - \tilde{u}^{\varepsilon}(t)| + \lambda_1^{\varepsilon}(t) \\
&\leq C y^{\varepsilon}(t)|u^{\varepsilon}(t) - \tilde{u}^{\varepsilon}(t)| + 2\varepsilon^{\frac{1}{2}} y^{\varepsilon}(t).
\end{aligned}$$

■

3.4 Necessary Conditions For Near-optimal Controls

3.4.1 Estimates On The Parameters

Lemma 3.4.1 *For all $\theta \geq 0$ and $0 < \kappa < 1$ such that $\kappa\theta < 1$, and for all $(u, u') \in \mathcal{U}_{ad}$, there exists $C = C(\theta, \kappa)$ such that*

$$\sum_{i=1}^{3} \mathbb{E}\left[\sup_{0\leq t\leq T} |x_i(t) - x_i'(t)|^{2\theta}\right] \leq C \sum_{i=1}^{2} d(u_i, u_i')^{\kappa\theta}. \tag{3.15}$$

Proof. We distinguish two cases

- $\theta \geq 1$

  For every $r > 0$, and by using Hölder's inequality, we get

$$\begin{aligned}\mathbb{E}\left[\sup_{0\leq t\leq r} |x_1(t) - x_1'(t)|^{2\theta}\right] &\leq C\mathbb{E}\left[\int_0^r \sum_{i=1}^{3} |x_i(t) - x_i'(t)|^{2\theta} dt\right] + C\mathbb{E}\left[\int_0^r \mathbf{1}_{\{u_1 \neq u_1'\}} dt\right]^{\kappa\theta} \\ &\leq C\mathbb{E}\left[\int_0^r \sum_{i=1}^{3} |x_i(t) - x_i'(t)|^{2\theta} dt\right] + C\mathbb{E}\left[d(u_1, u_1')^{\kappa\theta}\right].\end{aligned}$$

  We can get the same estimates aswell for $i = 2, 3$

$$\mathbb{E}\left[\sup_{0\leq t\leq r} |x_2(t) - x_2'(t)|^{2\theta}\right] \leq C\mathbb{E}\left[\int_0^r \sum_{i=1}^{2} |x_i(t) - x_i'(t)|^{2\theta} dt + d(u_2, u_2')^{\kappa\theta}\right],$$
$$\mathbb{E}\left[\sup_{0\leq t\leq r} |x_3(t) - x_3'(t)|^{2\theta}\right] \leq C\mathbb{E}\left[\int_0^r \sum_{i=1}^{3} |x_i(t) - x_i'(t)|^{2\theta} dt + \sum_{i=1}^{2} d(u_i, u_i')^{\kappa\theta}\right].$$

  Adding the equations yields

$$\sum_{i=1}^{3} \mathbb{E}\left[\sup_{0\leq t\leq r} |x_i(t) - x_i'(t)|^{2\theta}\right] \leq C\mathbb{E}\left[\int_0^r \sum_{i=1}^{3} \sup_{0\leq t\leq s} |x_i(t) - x_i'(t)|^{2\theta} ds + \sum_{i=1}^{2} d(u_i, u_i')^{\kappa\theta}\right].$$

Using the Gronwall inequality, we get the result

$$\sum_{i=1}^{3}\mathbb{E}\left[\sup_{0\le t\le r}|x_i(t)-x_i'(t)|^{2\theta}\right]\le C\left[\sum_{i=1}^{2}d(u_i,u_i')^{\kappa\theta}\right].$$

- $0\le\theta<1$

Using Hölder inequality, the previous result for $\theta\ge 1$, and Gronwall's inequality, we get

$$\begin{aligned}\sum_{i=1}^{3}\mathbb{E}\left[\sup_{0\le t\le r}|x_i(t)-x_i'(t)|^{2\theta}\right]&\le\sum_{i=1}^{3}\left[\mathbb{E}\left[\sup_{0\le t\le r}|x_i(t)-x_i'(t)|^{2}\right]\right]^{\theta}\\&\le C\left[\int_0^r\sum_{i=1}^{3}\mathbb{E}\left[\sup_{0\le t\le s}|x_i(t)-x_i'(t)|^{2}ds\right]+\sum_{i=1}^{2}d(u_i,u_i')^{\kappa}\right]^{\theta}\\&\le C\left[\sum_{i=1}^{2}d(u_i,u_i')^{\kappa\theta}\right],\end{aligned}$$

which completes the proof.

■

3.4.2 Necessary Conditions for Near-optimal Controls

Lemma 3.4.2 *Under (H3) and (H4), $\forall\kappa\in]0,1[,\forall\theta\in]0,2[$ satisfying $(1+\kappa)\theta<2$, and for every $(u,u')\in\mathcal{U}_{ad}\times\mathcal{U}_{ad}$, $(p,q),(p',q')$ solutions of the corresponding adjoint equation, there exists $C=C(\kappa,\theta)$ such that*

$$\sum_{i=1}^{3}\mathbb{E}\left[\int_0^T|p_i(t)-p_i'(t)|^{\theta}dt\right]+\sum_{i=1}^{3}\mathbb{E}\left[\int_0^T|q_i(t)-q_i'(t)|^{\theta}dt\right]\le C\sum_{i=1}^{3}d(u_i,u_i')^{\frac{\kappa\theta}{2}}. \tag{3.16}$$

Proof. See [6] ■

Theorem 3.4.3 *Let $(p^{\varepsilon},q^{\varepsilon})$ be the solution to the adjoint equation under the control u^{ε}. Then, under hypothesis (H1),(H2), there exists C such that $\forall\theta\in[0,1[,\forall\varepsilon>0$ and for any ε-optimal* *pair* $(x^{\varepsilon}, u^{\varepsilon})$*, we have*

$$\min_{u\in\mathcal{U}_{ad}[0,T]} \mathbb{E}\left[\int_0^T \mathcal{H}(t, x^{\varepsilon}(t), u(t), p^{\varepsilon}(t), q^{\varepsilon}(t))dt\right] + C\varepsilon^{\frac{\theta}{3}} \geq \mathbb{E}\left[\int_0^T \mathcal{H}(t, x^{\varepsilon}(t), u^{\varepsilon}(t), p^{\varepsilon}(t), q^{\varepsilon}(t))dt\right], \tag{3.17}$$

that is

$$\begin{aligned}
&\min_{u\in\mathcal{U}_{ad}[0,T]} \mathbb{E}\Bigg[\int_0^T \bigg(u_1(t)x_1^{\varepsilon}(t)(p_3^{\varepsilon}(t) - p_1^{\varepsilon}(t)) + \frac{m(\xi(t))u_2(t)x_2^{\varepsilon}(t)}{1+\eta(\xi(t))x_2^{\varepsilon}(t)}(p_3^{\varepsilon}(t) - p_2^{\varepsilon}(t)) \\
&\qquad + \frac{1}{2}(\tau_1 u_1^2(t) + \tau_2 u_2^2(t))\bigg)dt\Bigg] + C\varepsilon^{\frac{\theta}{3}} \\
&\geq \mathbb{E}\Bigg[\int_0^T \bigg(u_1^{\varepsilon}(t)x_1^{\varepsilon}(t)(p_3^{\varepsilon}(t) - p_1^{\varepsilon}(t)) + \frac{m(\xi(t))u_2^{\varepsilon}(t)x_2^{\varepsilon}(t)}{1+\eta(\xi(t))x_2^{\varepsilon}(t)}(p_3^{\varepsilon}(t) - p_2^{\varepsilon}(t)) \\
&\qquad + \frac{1}{2}(\tau_1 (u_1^{\varepsilon})^2(t) + \tau_2 (u_2^{\varepsilon})^2(t))\bigg)dt\Bigg].
\end{aligned}$$

Proof. We first define a new metric d

$$d : \mathcal{U}_{ad} \times \mathcal{U}_{ad} \longrightarrow \mathbb{R}_+,$$
$$d(u^{\varepsilon}, \tilde{u}^{\varepsilon}) \leq \varepsilon^{\frac{3}{2}},$$

and a new cost function

$$\tilde{J}(0, x(0), u(t)) = J(0, x(0), u(t)) + \varepsilon^{\frac{1}{3}} d(u(t), \tilde{u}^{\varepsilon}(t)).$$

We directly have

$$\tilde{J}(0, x(0), \tilde{u}^{\varepsilon}(t)) \leq \tilde{J}(0, x(0), u(t)). \tag{3.18}$$

If we take a look at 3.18, we can see that the pair $(\tilde{x}^{\varepsilon}(t), \tilde{u}^{\varepsilon}(t))$ is optimal for the system 3.3 with the cost function 3.2. If we consider the couple $(\tilde{p}^{\varepsilon}(t), \tilde{q}^{\varepsilon}(t))$ to be the solution to the adjoint equation 3.5 under $\tilde{u}^{\varepsilon}(t)$, then using the stochastic maximum principle $\forall t \in [0,T], \forall\theta \in [0,1[$ yields

$$\mathcal{H}(t, \tilde{x}^{\varepsilon}(t), \tilde{u}^{\varepsilon}(t), \tilde{p}^{\varepsilon}(t), \tilde{q}^{\varepsilon}(t)) = \min_{u^{\varepsilon}\in\mathcal{U}_{ad}[0,T]} \mathcal{H}(t, \tilde{x}^{\varepsilon}(t), u^{\varepsilon}(t), \tilde{p}^{\varepsilon}(t), \tilde{q}^{\varepsilon}(t)) + \varepsilon^{\frac{\theta}{3}}|u(t) - \tilde{u}^{\varepsilon}(t)|.$$

By denoting

$$H_1(t) = \mathcal{H}(t, \tilde{x}^\varepsilon(t), \tilde{u}^\varepsilon(t), \tilde{p}^\varepsilon(t), \tilde{q}^\varepsilon(t)) - \mathcal{H}(t, x^\varepsilon(t), u^\varepsilon(t), p^\varepsilon(t), q^\varepsilon(t))$$

$$H_2(t) = \mathcal{H}(t, \tilde{x}^\varepsilon(t), u(t), \tilde{p}^\varepsilon(t), \tilde{q}^\varepsilon(t)) - \mathcal{H}(t, x^\varepsilon(t), u^\varepsilon(t), p^\varepsilon(t), q^\varepsilon(t))$$

we get

$$\begin{aligned}
&\mathbb{E}\bigg[\int_0^T \mathcal{H}(t, x^\varepsilon(t), u^\varepsilon(t), p^\varepsilon(t), q^\varepsilon(t))dt\bigg] \\
&\leq \mathbb{E}\bigg[\int_0^T \mathcal{H}(t, \tilde{x}^\varepsilon(t), \tilde{u}^\varepsilon(t), \tilde{p}^\varepsilon(t), \tilde{q}^\varepsilon(t))dt\bigg] + \mathbb{E}\bigg[\int_0^T |H_1(t)|dt\bigg] \\
&\leq \mathbb{E}\bigg[\int_0^T \min_{u \in \mathcal{U}_{ad}} \Big(\mathcal{H}(t, \tilde{x}^\varepsilon(t), \tilde{u}^\varepsilon(t), \tilde{p}^\varepsilon(t), \tilde{q}^\varepsilon(t)) + \varepsilon^{\frac{\theta}{3}}|u(t) - \tilde{u}^\varepsilon(t)|\Big)dt\bigg] + \mathbb{E}\bigg[\int_0^T |H_1(t)|dt\bigg] \\
&\leq \mathbb{E}\bigg[\int_0^T \min_{u \in \mathcal{U}_{ad}} \mathcal{H}(t, \tilde{x}^\varepsilon(t), \tilde{u}^\varepsilon(t), \tilde{p}^\varepsilon(t), \tilde{q}^\varepsilon(t))dt + \mathbb{E}\bigg[\int_0^T |H_1(t)|dt\bigg] + C\varepsilon^{\frac{\theta}{3}} \\
&\leq \min_{u \in \mathcal{U}_{ad}} \mathbb{E}\bigg[\int_0^T \mathcal{H}(t, x^\varepsilon(t), u^\varepsilon(t), p^\varepsilon(t), q^\varepsilon(t))dt\bigg] + \min_{u \in \mathcal{U}_{ad}} \mathbb{E}\bigg[\int_0^T |H_2(t)|dt\bigg] \\
&\quad + \mathbb{E}\bigg[\int_0^T |H_1(t)|dt\bigg] + C\varepsilon^{\frac{\theta}{3}}.
\end{aligned}$$

Based on (H1), 3.2, 3.4.1, and 3.16, and the definition of the metric d, we get

$$\min_{u \in \mathcal{U}_{ad}} \mathbb{E}\bigg[\int_0^T |H_2(t)|dt\bigg] + \mathbb{E}\bigg[\int_0^T |H_1(t)|dt\bigg] \leq C\varepsilon^{\frac{\theta}{3}},$$

from which we get directly

$$\mathbb{E}\left[\int_0^T \mathcal{H}(t, x^\varepsilon(t), u(t), p^\varepsilon(t), q^\varepsilon(t))dt\right] \leq \min_{u^\varepsilon \in \mathcal{U}_{ad}[0,T]} \mathbb{E}\left[\int_0^T \mathcal{H}(t, x^\varepsilon(t), u^\varepsilon(t), p^\varepsilon(t), q^\varepsilon(t))dt\right] + C\varepsilon^{\frac{\theta}{3}}.$$

■

Conclusion

IN THIS THESIS, we studied the necessary and sufficient conditions of near-optimal controls in a regime switching stochastic system using the stochastic maximum principle.

As a start, basic concepts and preliminaries concerning stochastic calculus were provided. Then, after recalling the basics of optimal control theory, we defined the stochastic maximum principle, as well as Ekeland's principle. We later gave the detailed explanation of a regime switching stochastic system. After that, we introduced a stochastic medical SIRS model transformed later into a regime switching system.

Finally, we studied the necessary and sufficient conditions on near-optimal controls using the stochastic maximum principle.

Bibliography

[1] P. Briand. *Equations Différentielles Stochastiques Rétrogrades.* Polycopié (2001).

[2] J.D. Hamilton. *Regime Switching Models.* Palgrave Dictionary of Economics, 2nd Ed. Vol.2. Palgrave Macmillan Editions (2008).

[3] J.C. Laleuf. *Processus et intégrales stochastiques, cours et exercices.* Ellipses (2014).

[4] M. Lefebvre. *Applied stochastic processes.* Springer (2006).

[5] S. Lessard. *Processus stochastiques. Cours et exercices corrigés.* Ellipses (2014).

[6] X. Mu, Q. Zhang. *Optimal strategy of vaccination and treatment in a SIRS model with makrovian switching.* Wiley. Math Meth. Appli Sci. 2018;1-23.

[7] B. Øksendal, A.Sulem. *Applied Stochastic Control of Jump Diffusions.* Springer (2005).

[8] E. Pardoux, S. Peng. *Backward doubly stochastic differential equations and systems of quasilinear SPDEs.* Springer-Verlag. Probab. Theory Relat. Fields 98, 209-227 (1994)

[9] J. Yong, X. Zhou. *Stochastic Controls, Hamiltonian Systems, and HJB Equations.* New York: Springer (1999).

[10] J. Zhang. *Backward Stochastic Differential Equations, from linear to fully nonlinear theory.* Probability Theory and Stochastic Modelling Vol. 86. Springer (2017).

Appendix A : Theorems

Theorem A.1 *(Hölder's inequality) Let $(E, \mathcal{E}, \mu)$ be a measured space under the measure μ. Then for every two measurable functions f, g, and for every conjugate pair $(p, q) \in ([1, +\infty[)^2$ (i.e $\frac{1}{p} + \frac{1}{q} = 1$) we have*

$$\|fg\|_1 \leq \|f\|_p \|g\|_q, \tag{3.19}$$

where

$$\|f\|_p = \left(\int_E |f|^p d\mu \right)^{\frac{1}{p}}. \tag{3.20}$$

Remark A.2 *The special case where $p = q = \frac{1}{2}$ is called the Cauchy-Schwarz inequality*

$$\left(\int_E |fg| d\mu \right) \leq \left(\int_E |f|^2 d\mu \right)^{\frac{1}{2}} \left(\int_E |g|^2 d\mu \right)^{\frac{1}{2}}. \tag{3.21}$$

Definition A.3 *(Convex function) Let A be a convex set. We say that $f : A \longrightarrow \mathbb{R}$ is convex iff $\forall (x, y) \in A \times A$ and $\forall \alpha \in [0, 1]$*

$$f(\alpha x + (1 - \alpha) y) \leq \alpha f(x) + (1 - \alpha) f(y). \tag{3.22}$$

Example A.4 *For $x \in \mathbb{R}$, the functions $x \rightarrow |x|$, $x \rightarrow x^2$, $x \rightarrow e^x$ are convex functions.*

Property A.5 *Let $f : A \longrightarrow \mathbb{R}$ be a differentiable, convex function. Then $\forall (x, y) \in A \times A$ we have*

$$f(x) - f(y) \leq D_x f(x)(x - y), \tag{3.23}$$

where $D_x f$ denotes the derivative of f with respect to x.

Theorem A.6 *(Gronwall's inequality) Let $f \in L^1([0,T])$ be a C^1 function, and $a \geq 0$, $b \geq 0$. If $\forall t \geq 0$*

$$f(t) \leq a + b\int_0^t f(s)ds,$$

then

$$f(t) \leq ae^{bt}.$$

Definition A.7 *Let $\Omega \subset \mathbb{R}^n$ and $f : \Omega \longrightarrow \mathbb{R}^n$ be a locally Lipschitz continuous function. For any $x \in \Omega$, we define*

$$\partial f(x) \triangleq \left\{ \xi \in \mathbb{R}^n \middle| \langle \xi, y \rangle \leq \overline{\lim_{\substack{z \to x, z \in \Omega \\ t \downarrow 0}}} \frac{f(z+ty) - f(z)}{t} \right\}. \tag{3.24}$$

The set $\partial f(x)$ is called Clarke's generalized gradient, and it verifies

- *$\partial f(x)$ is a nonempty, convex, and compact set in $\mathbb{R}^n$.*
- *$0 \in \partial f(x)$ if f attains a local minimum (resp. maximum) in x.*

Theorem A.8 *(Burkholder-Davis-Gundy Inequality) Let $\sigma = (\sigma(t))_{t \geq 0}$ be a local martingale. For any $p > 0$, then there exist universal constants $C_p > c_p > 0$ depending only on p and d such that*

$$c_p \mathbb{E}\left[\left(\int_0^T |\sigma(t)|^2 dt \right)^{\frac{p}{2}} \right] \leq \mathbb{E}\big[| \sup_{0 \leq t \leq T} \sigma(t)|^p \big] \leq C_p \mathbb{E}\left[\left(\int_0^T |\sigma(t)|^2 dt \right)^{\frac{p}{2}} \right]. \tag{3.25}$$

Appendix B : Abbreviations and Notations

All used abbreviations are explained here :

$\mathbb{R}^n$	:	The space of n-dimensional, real-valued vectors.
$\mathbb{P}$	:	Probability measure.
$\|f\|_p$	:	p-norm defined by $\|f\|_p = \left(\int_E |f|^p d\mu\right)^{\frac{1}{p}}$.
Γ	:	The space of the control constraint.
B	:	Brownian motion.
$\mathbb{F}$	:	Filtration of some σ-algebra $\mathcal{F}$.
λ	:	Lebesgue measure.
$\mathfrak{B}$	:	The $\mathcal{F}_t$-predictable σ-field on $[0,T] \times \mathcal{F}$
$(\mathcal{F}_t^X)_{t\geq 0}$	:	Natural filtration of the stochastic process X defined by $\mathcal{F}_t^X = \sigma(X(s), 0 \leq s \leq t)$
$D_x f$	:	First derivative of f in x.
$D_{xx} f$	:	Second derivative of f in x.
$\mathcal{N}(\mu, \sigma^2)$	:	Normal distribution with expected value μ and variance σ^2.
$\mathcal{B}(A)$	:	Borel σ-algebra over the subset $A \subset \mathbb{R}$.
$L^p([0,T], \mathbb{R}^n)$	:	The set of all $\mathbb{R}^n$-valued functions f such that $\left(\int_0^T |f|^p dt\right)^{\frac{1}{p}} < +\infty$.
$L^p(\Omega, \mathcal{F}; \mathbb{R}^n)$	:	The set of all $\mathcal{F}$-measurable, $\mathbb{R}^n$-valued random variables such that $\mathbb{E}[|X|^2] < +\infty$.

$\mathcal{M}^p([0,T],\mathbb{R}^n)$	:	The set of all adapted, $\mathbb{R}^n$-valued random processes $(X(t))_{t\geq 0}$such that $\mathbb{E}\left[\int_0^T \lvert X(t)\rvert^p dt\right] < +\infty$.
$\partial f(x)$	:	Clarke's generalized gradient of f.
$\partial_{x_i} f(x)$	:	Clarke's generalized gradient of f with respect to the variable x_i.
$\mathbb{E}[X]$	:	The expectation of X.
$\mathbb{E}[X\lvert\mathcal{G}]$	:	The conditional expectation of X with respect to the σ-algebra $\mathcal{G}$.
$Var(X)$	:	Variance of the random variable X.
$Cov(X,Y)$	:	Covariance of the two random variables X and Y.
$\langle X\rangle_t$	:	The quadratic variation of the process X defined by the limit $\lim\limits_{n\to+\infty} \sum_{k=1}^n \left(X(t_{k+1}) - X(t_k)\right)^2$.
$C^k(E)$	:	The set of all continuous, k times differentiable functions $f : E \to \mathbb{R}$ with continuous derivatives.
$\mathbb{H}^p([0,T],\mathbb{R}^n)$	:	The set of all $\mathbb{R}^n$-valued, progressively-measurable processes X such that $\forall t \geq 0 : \mathbb{E}\left[\int_0^T \lvert X(t)\rvert^p dt\right] < +\infty$.
$S^p_{\mathcal{F}}([s,t];\mathbb{R}^n)$	:	The set of all$(\mathcal{F}_u)_{u\in[s,t]}$-adapted, càdlag processes X such that $\mathbb{E}\left[\sup\limits_{s\in[t,T]} \lvert X(s)\rvert^2\right] < +\infty$.
$\mathcal{U}[0,T]$	:	The set of all feasible controls defined by $\{u : [0,T]\times\Omega \in \Gamma\lvert u(.)$ is measurable$\}$.
$\mathcal{U}_{ad}[0,T]$	:	The set of all admissible controls u over the time horizon $[0,T]$.
$a \wedge b$	:	$\inf\{a,b\}$.
$\mathrm{Dom}(f)$	:	Domaine of f given by $f : E \to F,\ \ \mathrm{Dom}(f) = \{x \in E/\exists y \in F : y = f(x)\}$.
càdlàg/càglàd	:	Right-continuous, with left limits. / Left-continuous with right limits.

الملخص :

يتعلق هذا البحث بمبدأ بونترياغين العشوائي وتطبيقاته في الأنظمة العشوائية متغيرة النظام. نبدأ الفصل الأول بأهم التذكيرات والتعاريف المتعلقة بالتحليل العشوائي, المعادلات التفاضلية العشوائية وسلاسل ماركوف. نتطرق بعدها إلى الحديث عن التحكم المثالي في الأنظمة العشوائية وبعض التفاصيل الأخرى.
أخيرا يتعلق الفصل الثالث بتطبيق لمبدأ بونترياغين في الأنظمة العشوائية الطبية (إنتشار الفيروسات) ودراسة تطور هذا النظام وكذا التحكم المثالي به.

Résumé :

Dans ce travail, nous étudions le principe du maximum stochastique et ses applications dans les systèmes avec changement de régime. Dans le premier chapitre on commence par des généralités mathématiques du calcul stochastique. Dans le deuxième chapitre, on fournit des rappels sur la théorie du contrôle optimal et le principe du maximum stochastique. Finalement, on applique ce principe sur des systèmes viraux (SIRS) avec un changement de régime.

Summary :

In this thesis, we studied the stochastic maximum principle and its application in regime switching stochastic systems. As a start, we recall some mathematical preliminaries (Stochastic calculus). Then we provide some basic definitions regardin optimal control theory. Finally, we study a regime switching viral system (SIRS) using stochastic maximum principle.