BACKWARD STOCHASTIC DIFFERENTIAL EQUATIONS AND THEIR APPLICATIONS TO STOCHASTIC CONTROL PROBLEMS

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF APPLIED MATHEMATICS
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

HANİFE SEVDA NALBANT

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
FINANCIAL MATHEMATICS

MAY 2013

Approval of the thesis:

BACKWARD STOCHASTIC DIFFERENTIAL EQUATIONS AND THEIR APPLICATIONS TO STOCHASTIC CONTROL PROBLEMS

submitted by **HANİFE SEVDA NALBANT** in partial fulfillment of the requirements for the degree of **Master of Science in Department of Financial Mathematics, Middle East Technical University** by,

Prof. Dr. Bülent Karasözen
Director, Graduate School of **Applied Mathematics** __________

Assoc. Prof. Dr. Sevtap Kestel
Head of Department, **Financial Mathematics** __________

Assist. Prof. Dr. Yeliz Yolcu Okur
Supervisor, **Financial Mathematics** __________

Assoc. Prof. Dr. Azize Hayfavi
Co-supervisor, **Financial Mathematics** __________

Examining Committee Members:

Prof. Dr. Gerhard Wilhelm Weber
Institute of Applied Mathematics, METU __________

Assist. Prof. Dr. Yeliz Yolcu Okur
Institute of Applied Mathematics, METU __________

Assoc. Prof. Dr. Azize Hayfavi
Institute of Applied Mathematics, METU __________

Prof. Dr. Bülent Karasözen
Institute of Applied Mathematics, METU __________

Assoc. Prof. Dr. Ömür Uğur
Institute of Applied Mathematics, METU __________

Date: __________

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name: HANİFE SEVDA NALBANT

Signature :

ABSTRACT

BACKWARD STOCHASTIC DIFFERENTIAL EQUATIONS AND THEIR APPLICATIONS TO STOCHASTIC CONTROL PROBLEMS

Nalbant, Hanife Sevda
M.Sc., Department of Financial Mathematics
Supervisor : Assist. Prof. Dr. Yeliz Yolcu Okur
Co-Supervisor : Assoc. Prof. Dr. Azize Hayfavi

May 2013, 64 pages

Backward stochastic differential equations (BSDE) were firstly introduced by Bismut in 1973. Following decades, it has been great interest all over the world and appeared in numerious areas such as pricing and hedging claims, utility theory and optimal control theory. In 1997, El Karoui, Peng and Quenez brought together their brilliant studies in the article *Backward Stochastic Differential Equations in Finance*. They considered an adapted solution pair (Y, Z) of the following BSDE: $-dY_t = f(t, Y_t, Z_t)dt - Z_t^* dW_t$ with the terminal value $Y_T = \xi$. Here Z^* corresponds to the transponse of the $n \times n$ matrix Z, f is called the generator and ξ is the terminal condition. In this thesis, we study some chapter of this paper in detail. We prove the fundamental theorems of backward stochastic differential equations and associate them with stochastic control problems. After we prove the existence of unique solution using *a Priori estimates* under some restrictions, we show how to choose the optimal stochastic control that achieves the best utility or the least cost. At the end of the thesis, we offer an optimal choice for the solution of the BSDE in the cases of the standard generator f is concave or convex. An application for the model with consumption and an application for hedging claims with higher interest rate for borrowing are provided.

Keywords: backward stochastic differential equation, pricing, hedging portfolios,stochastic optimal control

ÖZ

GERİYE DOĞRU STOKASTİK DİFERANSİYEL DENKLEMLER VE STOKASTİK KONTROL PROBLEMLERİNE UYGULANMASI

Nalbant, Hanife Sevda
Yüksek Lisans, Finansal Matematik Bölümü
Tez Yöneticisi : Yrd. Doç. Dr. Yeliz Yolcu Okur
Ortak Tez Yöneticisi : Doç. Dr. Azize Hayfavi

Mayıs 2013, 64 sayfa

Geriye doğru stokastik diferansiyel denklemler (GSDD), ilk olarak 1973 yılında Bismut tarafından takdim edilmiştir. İlerleyen yıllarda, dünya çapında büyük ilgi uyandırarak, fiyatlama ve riskten korunma, fayda teorisi, optimal kontrol teorisi gibi bir çok alanda uygulanmaya başlamıştır. 1997 yılında ise El Karoui, Peng ve Quenez, bu alandaki çalışmalarını *Finans Alanında Geriye Doğru Stokastik Diferansiyel Denklemler* adlı makalelerinde bir araya getirmişlerdir. Bu çalışmada, $Y_T = \xi$ son değerine sahip $-dY_t = f(t, Y_t, Z_t)dt - Z_t^* dW_t$ biçimindeki GSDD' nin (Y, Z) uyarlanmış çözüm çifti incelenmiştir. Burada Z^*, $n \times n$ boyutundaki Z matrisinin transpozuna karşılık gelmektedir; f' ye standart üreten, ξ' ye ise son değer koşulu denilmektedir. Bu tezde, söz konusu makalenin bazı bölümlerini detaylı bir şekilde çalıştık. Geriye doğru stokastik diferansiyel denklemlerin temel teoremlerini ispat ettik ve stokastik kontrol problemleriyle ilişkilendirdik. *A Priori estimates* yöntemini kullanarak çözümün varlığını ve tekliğini belli koşullar altında ispatladıktan sonra, en iyi fayda veya en az maliyete tekabül eden optimal stokastik kontrol değerini nasıl seçeceğimizi gösterdik. Tezin sonunda, standart üreten olarak adlandırılan f' nin içbükey veya dışbükey olması durumunda, GSDD' nin çözümü için optimal seçim önerisinde bulunduk. Bununla ilgili olarak, tüketim sürecini göz önünde bulunduran model ve yüksek faizden borçlanmayla riskten korunma örnekleri için birer uygulama yaptık.

Anahtar Kelimeler: geriye doğru stokastik diferansiyel denklemler, fiyatlama, risten korunma portföyleri, stokastik optimal kontrol

To my love Efe and my beloved friend Burcu

ACKNOWLEDGMENTS

I would like to express my appreciation to all people who encouraged me to complete this thesis and get the Master of Science degree.

I am very greatful to my supervisor Assist. Prof. Dr. Yeliz Yolcu Okur and co-supervisor Assoc. Prof. Dr. Azize Hayfavi for their priceless patience, guidance and contribution throuht not only this thesis but also the whole three years in Financial Mathematics.

I would like to thank my mother Tülay, my unique love Efe, my beloved friend Burcu, my special friend Olcay, Cansu, Alev and my little cat Eda for their endless belief in me and motivation during the preparation of this thesis.

Special thanks to Prof. Dr. Ralf Korn for his valuable contribution.

I wish to acknowledge the support and kindness provided by Prof. Dr. Bülent Karasözen, Prof. Dr. Gerhard Wilhelm Weber, Assoc. Prof. Dr. Ömür Uğur and Res. Assist. Cansu Bilgir Evcin.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

It has been a great interest to study with backward stochastic differential equations in various areas in the last decades. Modelling financial assets such as stocks, interest rate processes, pricing and hedging claims and solving stochastic control problems could be counted as remarkable examples. Contrary to backward and forward equations which seem very similar to each other, a stochastic differential equation with terminal value has an anticipating solution generally.
Backward stochastic differential equations (BSDE in short) were firstly introduced by Bismut [12] in 1973 in the linear case as the equation for the adjoint process which of the form

$$dx = f(\omega, t, x, u)dt + \sigma(\omega, t, x, u)dw, \qquad x(0) = x_0,$$

where w is an m-dimensional Brownian motion, x belongs to an n-dimensional vector space, $\omega \in \Omega$, $t \in [0, \infty)$ and u comes from the completion of the σ-algebra of $\Omega \times [0, \infty)$ for the measure $d\mathbb{P} \otimes dt$.
The general case of backward stochastic differential equations was studied by Pardoux and Peng [4] in 1990. Both cases are related with our study and considered in the following manner:
The BSDE which has the form

$$-dY_t = f(t, Y_t, Z_t)dt - Z_t^* dW_t, \qquad Y_T = \xi \tag{1.1}$$

has an adapted solution pair (Y, Z) on the probability space of n-valued Brownian Motion W, where Y belongs to $\mathbb{R}$ and Z is considerable as an $n \times n$ matrix. In addition, f is called generator and ξ is called the terminal condition. The integral form can be written by

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s)ds - \int_t^T Z_s^* dW_s,$$

which is equivalent to (1.1).
Fundamental proporties of the backward stochastic differential equations were aggregated in the article *Backward Stochastic Differential Equations in Finance* written by El Karoui et al. [16] in 1997. In this thesis, we closely follow some chapters of this article and prove basic theorems, for instance the existence and uniqueness of the solution and the comparison theorem.

One of our problem is assumed to be constructed in a complete market. The main target is valuation of a contract which pays an amount $\xi \geq 0$ at the end of the time period T. It is so-called pricing the contingent claim ξ and used by firstly Black and Scholes [6] and then Merton [21, 22] and Karatzas et al. [8]. According to El Karoui et al. [16], with a linear generator f, a replicating portfolio Y associated with a corresponding hedging portfolio Z can be easily constructed which solves the BSDE (1.1) and attains the value ξ at maturity. It is a natural outcome to think that the price at time t coincides with the value at time t of the hedging portfolio. On the other hand, there are infinitely many replicating portfolios associated with the contingent claim ξ which implies that the price is not well-defined. Nevertheless, it can be possible to be well-posed with some restrictions on the integrability of the hedging portfolios, including to change the existed probability into risk adjusted probability measure. Hence existence and uniqueness of price and hedging portfolio make senses with the assumption that admissible strategies are restricted to be square-integrable under the first probability measure.
Following the articles [16] and [17], the solution of the BSDE (1.1) is considered as supremum (or infimum) of some related controlled processes. According to optimal stochastic control theory, the optimality of the solution could be determined in cases of concavity or convexity. As a consequence, constraints for an incomplete market were extended to convex constraints on the portfolios, in Cvitanic and Karatzas [10]. An example of a non-linear BSDE was studied by those authors which solves the hedging problem allows a higher interest rate for borrowing. It will be associated with our stochastic control problem.

In this thesis, we deeply prove the fundamental theorems in BSDE theory which are stated in the article [16] and relate them with optimal stochastic control theory following [13].
We start with a brief introduction to the structure of BSDE theory, in the second chapter. We prove the existence and uniqueness theorem using *a Priori Estimates* of the spread between the solution of two BSDEs. Besides, we derive a unique solution for the linear BSDE which has the form

$$-dY_t = (\varphi_t + Y_t\beta_t + Z_t^*\gamma_t)\, dt - Z_t^*\, dW_t, \qquad Y_T = \xi,$$

where some constraints are assumed for φ, β and γ. Then, we mention and prove a very useful comparison theorem which is efficiently used in the whole study.
In Chapter 3, we consider a complete market which contains n risky assets (for instance stocks) and a risk-free asset. After that, we explain the concept of hedging strategies (superstrategies) and define the fair price (upper price). The fair price (upper price) is related with a solution to a LBSDE by using a deflator. The standard work on this subject can be found in [16].
In Chapter 4, we introduce the structure of optimal stochastic control problems. This construction is adapted from Björk [26] and was motivated by Quenez [13]. The main idea in this chapter is to obtain an optimal solution. As we closely follow the main article [16], a state process X is considered as controlled by control parameters u which are assumed to be chosen from a Polish space U. We minimize an objective function, that contains a running cost and a terminal

cost, over all admissible controls. Hence, we derive the well-known Hamilton-Jacobi-Bellman (HJB) equation, which is modified from [13]. We conclude it by giving a clear example. We also derive HJB equation in order to maximize the objective function involving a utility function. In the following, we can write the standard parameter (f, ξ) of the BSDE as an essential infimum of controlled standard parameters (f^u, ξ^u) over all admissible control processes $u \in U$. A verification theorem is naturally arised and proved using measurable selection theorem (see [27]). For direct constructions along more classical lines, we refer to the reader [13]. Finally, we explain how to determine optimization (either infimum or supremum) in the cases of standard generator. If it is concave, then the least cost is achieved by taking infimum. On the other hand, if it is convex, then the best utility is achieved by taking supremum.
In Chapter 5, two applications are provided. At first, we study in a complete market with a model concerning consumption (see [20]), which is assumed to be concave. Carroll and Kimball [2] showed that consumption could be concave if stochastic income is included in the wealth process. We offer an optimal solution to the model that achieves the least cost. Secondly, we consider a model with a convex standard generator, in an incomplete market. In literature, the model is called hedging contingent claims with higher interest rate for borrowing. (see Cvitanic and Karatzas [10] and Korn [19] for further details). After we explain the idea of the model, we suggest an optimal solution to the problem that achieves the best utility.
At the end of the thesis, we conclude all work in Chapter 6.

CHAPTER 2

LINEAR AND GENERAL BACKWARD STOCHASTIC DIFFERENTIAL EQUATIONS

The history of backward stochastic differential equations (BSDE in short) begin with Bismut [12] in 1973 and continue following decades. They are preferable in matematical finance since they have an anticipating solution in general. In particular, pricing and hedging a contingent claim is modeled in terms of a linear BSDE. Numerious studies have been performed in the theory of BSDE by Pardoux and Peng [4] and El Karoui et al. [17]. Our main article El Karoui et al. [16] is a great reference for the eminent results of the theory.
In this chapter, we will mention and prove some outstanding results for backward stochastic differential equations, which are already stated in the main article [16]. For the proofs, we make all steps clear. At the beginning of the chapter, the terminology will be introduced and then the existence and uniqueness theorem is proven using *a Priori Estimates*. In particular, the linear BSDEs are examined in the next section. Finally, the comparison theorem is given as a consequence.

2.1 Terminology

Consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and let W be $\mathbb{R}^n$-valued Brownian motion, where the filtration $\{\mathcal{F}_t\}_{0\leq t\leq T}$ is generated by the Brownian motion W and assumed to be augmented. Denote the σ-field of predictable sets of $\Omega \times [0,T]$ by $\mathcal{P}$. We agree on some usual notations: For a d-dimesional vector $x \in \mathbb{R}^d$, $|x|$ denotes its Euclidian norm and the inner product of two d-dimesional vectors x, v is denoted by $< x, v >$. For an $n \times d$ matrix $y \in \mathbb{R}^{n\times d}$, its Euclidian norm $|y|$ calculated by $|y| := \sqrt{trace(yy^*)}$, where * denotes the tranpose, and the inner product of two $n \times d$ matrices y, z is calculated by $< y, z >= trace(yz^*)$.
The additional notations are given as follows:

- $\mathbb{L}^2_T(\mathbb{R}^d)$, the space of all $\mathcal{F}_T$ - measurable random variables $X : \Omega \to \mathbb{R}^d$ such that $\|X\|^2 := \mathbb{E}(|X|^2) < +\infty$.
- $\mathbb{H}^2_T(\mathbb{R}^d)$, the space of all predictable process $\phi : \Omega \times [0,T] \to \mathbb{R}^d$ such that $\|\phi\|^2 := \mathbb{E}\left[\int_0^T |\phi_t|^2 dt\right] < +\infty$.

- $\mathbb{H}_T^1(\mathbb{R}^d)$, the space of all predictable process $\phi : \Omega \times [0,T] \to \mathbb{R}^d$ such that $\mathbb{E}\left[\sqrt{\int_0^T |\phi_t|^2 dt}\right] < +\infty$.

- For $\beta > 0$ and $\phi \in \mathbb{H}_T^2(\mathbb{R}^d)$, $\|\phi\|_\beta^2$ denotes $\mathbb{E}\left[\int_0^T e^{\beta t}|\phi_t|^2 dt\right]$. $\mathbb{H}_{T,\beta}^2(\mathbb{R}^d)$ denotes the space $\mathbb{H}_T^2(\mathbb{R}^d)$ endowed with the norm $\|\cdot\|_\beta$.

After we compromised with the terminology, we can introduce the concept of the BSDE and its associated standard parameter.

2.2 A Priori Estimates

The BSDE appears in the following form

$$-dY_t = f(t, Y_t, Z_t)\, dt - Z_t^* dW_t, \qquad Y_T = \xi \tag{2.1}$$

or, equivalently, in the integral form

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s) ds - \int_t^T Z_s^* dW_s. \tag{2.2}$$

Here, $f : \Omega \times \mathbb{R}^+ \times \mathbb{R}^d \times \mathbb{R}^{n \times d} \to \mathbb{R}^d$ is called generator (or sometimes called driver) and assumed to be $\mathcal{P} \otimes \mathcal{B}^d \otimes \mathcal{B}^{n \times d}$- measurable. ξ is the terminal value and $\mathcal{F}_T$-measurable random variable such that $\xi : \Omega \to \mathbb{R}^d$. The following definition gives the condition for the pair (f, ξ) to be a standard parameter (see [16]).

Definition 2.1. The pair (f, ξ) is said to be *a standard parameter* for the BSDE in (2.1) if the following conditions hold:

- $\xi \in \mathbb{L}_T^2(\mathbb{R}^d)$.
- $f(\cdot, 0, 0) \in \mathbb{H}_T^2(\mathbb{R}^d)$.
- f is uniformly Lipschitz, i.e., for all (y_1, z_1), $(y_2, z_2) \in \mathbb{R}^d \times \mathbb{R}^{n \times d}$, there exists a constant $C > 0$ such that

$$|f(\omega, t, y_1, z_1) - f(\omega, t, y_2, z_2)| \leq C(|y_1 - y_2| + |z_1 - z_2|).$$

On the other hand, a solution of the BSDE is an ordered pair (Y, Z), where $\{Y_t\}_{0 \leq t \leq T}$ is a continuous, $\mathbb{R}^d$-valued, adapted process and $\{Z_t\}_{0 \leq t \leq T}$ is an $\mathbb{R}^{n \times d}$-valued predictable process and satisfies the integrability condition
$\int_0^T |Z_s|^2 ds < \infty$.
Consider that a standard parameter (f, ξ) is given. Our aim is to show the existence and uniqueness of the solution pair (Y, Z) associated with this standard parameter. Pardoux and Peng [4] firstly proved the existence of a unique solution, but then El Karoui et al. [16] offered a simplier way which is called *a Priori Estimates*. We follow *a Priori Estimates* approach to prove the existence and uniqueness theorem. Hence, we first examine the approach in the following proposition.

Proposition 2.1 (a Priori Estimates)**.** *Let $((f^i, \xi^i); i = 1, 2)$ be two standard parameters of the BSDE (2.1) and $((Y^i, Z^i); i = 1, 2)$ be the related square-integrable solutions. Put $\delta_2 f_t = f^1(t, Y_t^2, Z_t^2) - f^2(t, Y_t^2, Z_t^2)$, $\delta Y_t = Y_t^1 - Y_t^2$ and $\delta Z_t = Z_t^1 - Z_t^2$. Let C be the Lipschitz constant for f^1. For any triple (λ, μ, β) such that $\mu > 0, \lambda^2 > C$ and $\beta \geq C(2 + \lambda^2) + \mu^2$, the following inequalities hold:*

$$\|\delta Y\|_\beta^2 \leq T\left[e^{\beta T}\mathbb{E}(|\delta Y_T|^2) + \frac{1}{\mu}\|\delta_2 f\|_\beta^2\right],$$

$$\|\delta Z\|_\beta^2 \leq \frac{\lambda^2}{\lambda^2 - C}\left[e^{\beta T}\mathbb{E}(|\delta Y_T|^2) + \frac{1}{\mu}\|\delta_2 f\|_\beta^2\right].$$

Proof. Let $(Y, Z) \in \mathbb{H}_T^2(\mathbb{R}^d) \times \mathbb{H}_T^2(\mathbb{R}^{n\times d})$ be a solution of (2.1). Our first aim is to show that $\sup_{0\leq t\leq T}|Y_t| \in \mathbb{L}_T^2(\mathbb{R})$. Taking the absolute value of the equation (2.2) and using the triangle inequality, we obtain

$$\begin{aligned}
|Y_t| &\leq |\xi| + \left|\int_t^T f(s, Y_s, Z_s)ds\right| + \left|\int_t^T Z_s^* dW_s\right| \\
&\leq |\xi| + \int_t^T |f(s, Y_s, Z_s)|\, ds + \left|\int_t^T Z_s^* dW_s\right| \\
&\leq |\xi| + \int_0^T |f(s, Y_s, Z_s)|\, ds + \left|\int_t^T Z_s^* dW_s\right|.
\end{aligned}$$

Taking supremum of the inequality over $t \in [0, T]$, then we get

$$\sup_{0\leq t\leq T}|Y_t| \leq |\xi| + \int_0^T |f(s, Y_s, Z_s)|\, ds + \sup_{0\leq t\leq T}\left|\int_t^T Z_s^* dW_s\right|.$$

Since (f, ξ) is the standard parameter, by definition $\xi \in \mathbb{L}_T^2(\mathbb{R}^d)$, which means $\mathbb{E}\left(|\xi|^2\right) < \infty$. Then $|\xi| \in \mathbb{L}_T^2(\mathbb{R})$. Again by definition, $f(\cdot, 0, 0) \in \mathbb{H}_T^2(\mathbb{R}^d)$. The assumption that (Y, Z) is a solution, i.e., $(Y, Z) \in \mathbb{H}_T^2(\mathbb{R}^d) \times \mathbb{H}_T^2(\mathbb{R}^{n\times d})$ allows us to write $\mathbb{E}\left[\int_0^T |f(t, Y_t, Z_t)|^2 dt\right] < \infty$. Therefore,

$$\mathbb{E}\left[\left|\int_0^T |f(s, Y_s, Z_s)|\, ds\right|^2\right] \leq \mathbb{E}\left[\int_0^T |f(s, Y_s, Z_s)|^2\, ds\right] < \infty,$$

which implies $\int_0^T |f(s, Y_s, Z_s)|\, ds \in \mathbb{L}_T^2(\mathbb{R})$. It remains to show that

$\sup_{0\le t\le T}\left|\int_t^T Z_s^* dW_s\right|$ belongs to $\mathbb{L}_T^2(\mathbb{R})$:

$$\begin{aligned}
\mathbb{E}\left[\sup_{0\le t\le T}\left|\int_t^T Z_s^* dW_s\right|^2\right] &= \mathbb{E}\left[\sup_{0\le t\le T}\left|\int_0^T Z_s^* dW_s - \int_0^t Z_s^* dW_s\right|^2\right] \\
&\le \mathbb{E}\left[\sup_{0\le t\le T} 2\left(\left|\int_0^T Z_s^* dW_s\right|^2 + \left|\int_0^t Z_s^* dW_s\right|^2\right)\right] \\
&= 2\mathbb{E}\left[\left|\int_0^T Z_s^* dW_s\right|^2\right] + 2\mathbb{E}\left[\sup_{0\le t\le T}\left|\int_0^t Z_s^* dW_s\right|^2\right] \\
&\le 2\mathbb{E}\left[\left|\int_0^T Z_s^* dW_s\right|^2\right] + 2\mathbb{E}\left[\left(\sup_{0\le t\le T}\left|\int_0^t Z_s^* dW_s\right|\right)^2\right].
\end{aligned}$$

By Burkholder-Davis-Gundy inequalities in [9], for some constant K_1,

$$\begin{aligned}
&\le 2\mathbb{E}\left[\left|\int_0^T Z_s^* dW_s\right|^2\right] + 2K_1\mathbb{E}\left[\int_0^T |Z_s|^2 ds\right] \qquad (\text{ by isometry}) \\
&\le 2\mathbb{E}\left[\int_0^T |Z_s|^2 ds\right] + 2K_1\mathbb{E}\left[\int_0^T |Z_s|^2 ds\right] \\
&= (2+2K_1)\mathbb{E}\left[\int_0^T |Z_s|^2 ds\right] < \infty.
\end{aligned}$$

Thus, we have shown $\sup_{0\le t\le T}|Y_s| \in \mathbb{L}_T^2(\mathbb{R})$.
Next, assume that (Y^1, Z^1) and (Y^2, Z^2) are two solutions of two standard parameters (f^1, ξ^1) and (f^2, ξ^2), respectively. Consider a function $g : [0, T]\times\mathbb{R} \mapsto \mathbb{R}$, defined by $g(s, x) = e^{\beta s}x^2$, where $x = |\delta Y_s| = (<\delta Y_s, \delta Y_s>)^{1/2}$. We can write the partial derivatives of g as follows:

$$\frac{\partial g}{\partial s}(s, x) = \beta e^{\beta s}x^2, \qquad \frac{\partial g}{\partial x}(s, x) = 2e^{\beta s}x, \qquad \frac{\partial^2 g}{\partial x^2}(s, x) = 2e^{\beta s},$$

or, equivalently,

$$\frac{\partial g}{\partial s}(s, |\delta Y_s|) = \beta e^{\beta s}|\delta Y_s|^2, \qquad \frac{\partial g}{\partial x}(s, |\delta Y_s|) = 2e^{\beta s} < \delta Y_s, d\delta Y_s >,$$

$$\frac{\partial^2 g}{\partial x^2}(s, |\delta Y_s|) = 2e^{\beta s}.$$

Let us apply the Itô's formula to the function $g(s, x)$ and integrate from $s = t$ to $s = T$, then we obtain

$$\begin{aligned}
e^{\beta T}|\delta Y_T^2| = e^{\beta t}|\delta Y_t^2| &+ \int_t^T \beta e^{\beta s}|\delta Y_s|^2 ds + 2\int_t^T e^{\beta s}|\delta Y_s| d(|\delta Y_s|) \\
&+ \frac{1}{2}.2\int_t^T e^{\beta s} d\left\langle |\delta Y_s|, |\delta Y_s|\right\rangle. \qquad (2.3)
\end{aligned}$$

Here, $d\langle|\delta Y_s|, |\delta Y_s|\rangle$ denotes the quadratic variation of $|\delta Y_s|$ and it equals to

$$d\langle|\delta Y_s|, |\delta Y_s|\rangle = |\delta Z_s|^2\, ds,$$

and, in addition, we can rewrite the term $2e^{\beta s} < \delta Y_s, d\delta Y_s >$ as follows,

$$\begin{aligned}2e^{\beta s} < \delta Y_s, d\delta Y_s > &= 2e^{\beta s} < \delta Y_s, (-f^1(s, Y_s^1, Z_s^1) + f^2(s, Y_s^2, Z_s^2))ds + \delta Z_s^* dW_s > \\ &= 2e^{\beta s} < \delta Y_s, (-f^1(s, Y_s^1, Z_s^1) + f^2(s, Y_s^2, Z_s^2))ds > \\ &\quad + 2e^{\beta s} < \delta Y_s, \delta Z_s^* dW_s > .\end{aligned}$$

Plugging these equailities into equation (2.3),

$$\begin{aligned}e^{\beta T}|\delta Y_T^2| = e^{\beta t}|\delta Y_t^2| &+ \int_t^T \beta e^{\beta s}|\delta Y_s|^2 ds \\ &+ 2\int_t^T e^{\beta s} < \delta Y_s, -f^1(s, Y_s^1, Z_s^1) + f^2(s, Y_s^2, Z_s^2) > ds \\ &+ 2\int_t^T e^{\beta s} < \delta Y_s, \delta Z_s^* dW_s > + \int_t^T e^{\beta s}|\delta Z_s|^2 ds\end{aligned}$$

or, equivalently,

$$\begin{aligned}e^{\beta t}|\delta Y_t^2| &+ \int_t^T \beta e^{\beta s}|\delta Y_s|^2 ds + \int_t^T e^{\beta s}|\delta Z_s|^2 ds \\ &= e^{\beta T}|\delta Y_T^2| + 2\int_t^T e^{\beta s} < \delta Y_s, f^1(s, Y_s^1, Z_s^1) - f^2(s, Y_s^2, Z_s^2) > ds \\ &\quad - 2\int_t^T e^{\beta s} < \delta Y_s, \delta Z_s^* dW_s > . \qquad (2.4)\end{aligned}$$

Let us calculate $f^1(s, Y_s^1, Z_s^1) - f^2(s, Y_s^2, Z_s^2)$ which we need in equality (2.4):

$$\begin{aligned}f^1(s, Y_s^1, Z_s^1) - f^2(s, Y_s^2, Z_s^2) &\leq \left|f^1(s, Y_s^1, Z_s^1) - f^2(s, Y_s^2, Z_s^2)\right| \\ &= \left|f^1(s, Y_s^1, Z_s^1) - f^2(s, Y_s^2, Z_s^2) \pm f^1(s, Y_s^2, Z_s^2)\right| \\ &\leq \left|f^1(s, Y_s^1, Z_s^1) - f^1(s, Y_s^2, Z_s^2)\right| \\ &\quad + \left|f^1(s, Y_s^2, Z_s^2) - f^2(s, Y_s^2, Z_s^2)\right| \\ \text{(since } f^1 \text{ is Lipschitz)} \qquad &\leq C(\left|Y^1 - Y^2\right| + \left|Z^1 - Z^2\right|) + |\delta_2 f_s| \\ &\leq C(|\delta Y_s| + |\delta Z_s|) + |\delta_2 f_s|\,.\end{aligned}$$

Then, the equation (2.4) becomes

$$
\begin{aligned}
e^{\beta t}|\delta Y_t^2| &+ \int_t^T \beta e^{\beta s}|\delta Y_s|^2 ds + \int_t^T e^{\beta s}|\delta Z_s|^2 ds \\
&= e^{\beta T}|\delta Y_T^2| + 2\int_t^T e^{\beta s}|\delta Y_s|\left(C(|\delta Y_s| + |\delta Z_s|) + |\delta_2 f_s|\right) ds \\
&\quad - 2\int_t^T e^{\beta s} < \delta Y_s, \delta Z_s^* dW_s > \\
&= e^{\beta T}|\delta Y_T^2| + 2C\int_t^T e^{\beta s}|\delta Y_s|^2 ds + \int_t^T e^{\beta s} 2|\delta Y_s|\left(C|\delta Z_s| + |\delta_2 f_s|\right) ds \\
&\quad - 2\int_t^T e^{\beta s} < \delta Y_s, \delta Z_s^* dW_s > . \qquad (2.5)
\end{aligned}
$$

Besides, we must show that the following inequality holds for all $C \in \mathbb{R}^+$, y, z, t, λ, $\mu \in \mathbb{R}$, with $\lambda \neq 0$ and $\mu \neq 0$:

$$2y(Cz + t) \leq C\frac{z^2}{\lambda^2} + Cy^2\lambda^2 + \frac{t^2}{\mu^2} + y^2\mu^2.$$

Indeed,

$$
\begin{aligned}
2y(Cz+t) &= \underbrace{2yCz} + 2yt \\
&\quad \left[0 \leq (\frac{z}{\lambda} - y\lambda)^2 = \frac{z^2}{\lambda^2} - 2yz + y^2\lambda^2 \right. \\
&\quad \left. \Rightarrow 2yz \leq \frac{z^2}{\lambda^2} + y^2\lambda^2 \Rightarrow 2Cyz \leq C\frac{z^2}{\lambda^2} + Cy^2\lambda^2 \text{ (since } C \text{ is positive)} \right] \\
&\leq C\frac{z^2}{\lambda^2} + Cy^2\lambda^2 + \underbrace{2yt} \\
&\quad \left[0 \leq (\frac{t}{\mu} - y\mu)^2 = \frac{t^2}{\mu^2} - 2ty + y^2\mu^2 \Rightarrow 2ty \leq \frac{t^2}{\mu^2} + y^2\mu^2 \right] \\
&\leq C\frac{z^2}{\lambda^2} + Cy^2\lambda^2 + \frac{t^2}{\mu^2} + y^2\mu^2.
\end{aligned}
$$

Setting $C = C$, $y = |\delta Y_s|$, $z = |\delta Z_s|$, $t = |\delta_2 f_s|$, we deduce the following inequality from equality (2.5)

$$
\begin{aligned}
e^{\beta t}|\delta Y_t^2| &+ \int_t^T \beta e^{\beta s}|\delta Y_s|^2 ds + \int_t^T e^{\beta s}|\delta Z_s|^2 ds \\
&\leq e^{\beta T}|\delta Y_T^2| + 2C\int_t^T e^{\beta s}|\delta Y_s|^2 ds \\
&\quad + \int_t^T e^{\beta s}\left(C\frac{|\delta Z_s|^2}{\lambda^2} + C|\delta Y_s|^2\lambda^2 + \frac{|\delta_2 f_s|^{2\,2}}{\mu} + |\delta Y_s|^2\mu^2 \right) ds \\
&\quad - 2\int_t^T e^{\beta s} < \delta Y_s, \delta Z_s^* dW_s > . \qquad (2.6)
\end{aligned}
$$

Note that $\sup_{0\leq s\leq T}|\delta Y_s| \in \mathbb{L}^2_T(\mathbb{R})$. For this reason, $e^{\beta s}\delta Z_s \delta Y_s$ belongs to $\mathbb{H}^1_T(\mathbb{R}^n)$. It implies that the stochastic integral $\int_t^T e^{\beta s} < \delta Y_s, \delta Z_s^* dW_s >$ is $\mathbb{P}$-integrable and expectation of it is equal to zero. Now, taking expectation of both sides in inequality (2.6) yields

$$\begin{aligned}
&\mathbb{E}\left[e^{\beta t}|\delta Y_t|^2\right] + \mathbb{E}\left[\int_t^T \beta e^{\beta s}|\delta Y_s|^2 ds\right] + \mathbb{E}\left[\int_t^T e^{\beta s}|\delta Z_s|^2 ds\right] \leq \mathbb{E}\left[e^{\beta T}|\delta Y_T|^2\right] \\
&\quad + \mathbb{E}\left[\int_t^T e^{\beta s}\left(2C|\delta Y_s|^2 + C\frac{|\delta Z_s|^2}{\lambda^2} + C|\delta Y_s|^2\lambda^2 + \frac{|\delta_2 f_s|^2}{\mu^2} + |\delta Y_s|^2\mu^2\right) ds\right] \\
&\quad = \mathbb{E}\left[e^{\beta T}|\delta Y_T|^2\right] + \left(C(2+\lambda^2)+\mu^2\right)\mathbb{E}\left[\int_t^T e^{\beta s}|\delta Y_s|^2 ds\right] \\
&\qquad + \frac{C}{\lambda^2}\mathbb{E}\left[\int_t^T e^{\beta s}|\delta Z_s|^2 ds\right] + \frac{1}{\mu^2}\mathbb{E}\left[\int_t^T e^{\beta s}|\delta_2 f_s|^2 ds\right]. \qquad (2.7)
\end{aligned}$$

By the assumptions on the coefficients, $C \leq \lambda^2$ and $(C(2+\lambda^2)+\mu^2) \leq \beta$, we finally get

$$\begin{aligned}
&\mathbb{E}\left[e^{\beta t}|\delta Y_t|^2\right] + \mathbb{E}\left[\int_t^T \beta e^{\beta s}|\delta Y_s|^2 ds\right] + \mathbb{E}\left[\int_t^T e^{\beta s}|\delta Z_s|^2 ds\right] \\
&\quad \leq \mathbb{E}\left[e^{\beta T}|\delta Y_T|^2\right] + \beta\mathbb{E}\left[\int_t^T e^{\beta s}|\delta Y_s|^2 ds\right] \\
&\qquad + \mathbb{E}\left[\int_t^T e^{\beta s}|\delta Z_s|^2 ds\right] + \frac{1}{\mu^2}\mathbb{E}\left[\int_t^T e^{\beta s}|\delta_2 f_s|^2 ds\right].
\end{aligned}$$

The same terms cancel each other, then

$$\mathbb{E}\left[e^{\beta t}|\delta Y_t|^2\right] \leq \mathbb{E}\left[e^{\beta T}|\delta Y_T|^2\right] + \frac{1}{\mu^2}\mathbb{E}\left[\int_t^T e^{\beta s}|\delta_2 f_s|^2 ds\right]. \qquad (2.8)$$

In this way, we can derive an upper bound for the β-norm of δY:

$$\begin{aligned}
\|\delta Y\|^2_\beta = \mathbb{E}\left[\int_0^T e^{\beta t}|\delta Y_t|^2 dt\right] &\leq \mathbb{E}\left[T\max_{0\leq t\leq T} e^{\beta t}|\delta Y_t|^2\right] \\
&\text{(assume } t_{max} \text{ maximizes } e^{\beta t}|\delta Y_t|^2) \\
&= T\,\mathbb{E}\left[e^{\beta t_{max}}|\delta Y_{t_{max}}|^2\right] \qquad \text{(by (2.8))} \\
&\leq T\left[\mathbb{E}\left(e^{\beta T}|\delta Y_T|\right) + \frac{1}{\mu^2}\int_{t_{max}}^T e^{\beta s}|\delta_2 f_s|^2 ds\right] \\
&\leq T\left[\mathbb{E}\left(e^{\beta T}|\delta Y_T|\right) + \frac{1}{\mu^2}\int_0^T e^{\beta s}|\delta_2 f_s|^2 ds\right].
\end{aligned}$$

Now, our aim is to find an upper bound for $\|\delta Z\|^2_\beta$. In order to do this, we need to turn back to inequality (2.7).

$$\begin{aligned}
\mathbb{E}\left[e^{\beta t}\left|\delta Y_t\right|^2\right] &+ \mathbb{E}\left[\int_t^T \beta e^{\beta s}\left|\delta Y_s\right|^2 ds\right] + \mathbb{E}\left[\int_t^T e^{\beta s}\left|\delta Z_s\right|^2 ds\right] \\
&\leq \mathbb{E}\left[e^{\beta T}\left|\delta Y_T\right|^2\right] + \big(C(2+\lambda^2)+\mu^2\big)\,\mathbb{E}\left[\int_t^T e^{\beta s}\left|\delta Y_s\right|^2 ds\right] \\
&\quad + \frac{C}{\lambda^2}\mathbb{E}\left[\int_t^T e^{\beta s}\left|\delta Z_s\right|^2 ds\right] + \frac{1}{\mu^2}\mathbb{E}\left[\int_t^T e^{\beta s}\left|\delta_2 f_s\right|^2 ds\right] \\
&\big(\text{by assumption, } \big(C(2+\lambda^2)+\mu^2\big) \leq \beta\big) \\
&\leq \mathbb{E}\left[e^{\beta T}\left|\delta Y_T\right|^2\right] + \beta\mathbb{E}\left[\int_t^T e^{\beta s}\left|\delta Y_s\right|^2 ds\right] \\
&\quad + \frac{C}{\lambda^2}\mathbb{E}\left[\int_t^T e^{\beta s}\left|\delta Z_s\right|^2 ds\right] + \frac{1}{\mu^2}\mathbb{E}\left[\int_t^T e^{\beta s}\left|\delta_2 f_s\right|^2 ds\right].
\end{aligned}$$

The same terms cancel each other, it yields,

$$\begin{aligned}
\mathbb{E}\left[e^{\beta t}\left|\delta Y_t\right|^2\right] + \mathbb{E}\left[\int_t^T e^{\beta s}\left|\delta Z_s\right|^2 ds\right] &\leq \mathbb{E}\left[e^{\beta T}\left|\delta Y_T\right|^2\right] + \frac{C}{\lambda^2}\mathbb{E}\left[\int_t^T e^{\beta s}\left|\delta Z_s\right|^2 ds\right] \\
&\quad + \frac{1}{\mu^2}\mathbb{E}\left[\int_t^T e^{\beta s}\left|\delta_2 f_s\right|^2 ds\right]
\end{aligned}$$

or, equivalently,

$$\begin{aligned}
(1-\frac{C}{\lambda^2})\,\mathbb{E}\left[\int_t^T e^{\beta s}\left|\delta Z_s\right|^2 ds\right] &\leq \mathbb{E}\left[e^{\beta T}\left|\delta Y_T\right|^2\right] - \mathbb{E}\left[e^{\beta t}\left|\delta Y_t\right|^2\right] \\
&\quad + \frac{1}{\mu^2}\mathbb{E}\left[\int_t^T e^{\beta s}\left|\delta_2 f_s\right|^2 ds\right] \\
&\leq \mathbb{E}\left[e^{\beta T}\left|\delta Y_T\right|^2\right] + \frac{1}{\mu^2}\mathbb{E}\left[\int_t^T e^{\beta s}\left|\delta_2 f_s\right|^2 ds\right]
\end{aligned}$$

which implies that

$$\mathbb{E}\left[\int_t^T e^{\beta s}\left|\delta Z_s\right|^2 ds\right] \leq \frac{\lambda^2}{\lambda^2 - C}\,\left[\mathbb{E}\left[e^{\beta T}\left|\delta Y_T\right|^2\right] + \frac{1}{\mu^2}\mathbb{E}\left(\int_t^T e^{\beta s}\left|\delta_2 f_s\right|^2 ds\right)\right]. \tag{2.9}$$

Inequality (2.9) is satisfied for all $t \in [0,T]$. In particular, for $t=0$, we obtain the final inequality:

$$\mathbb{E}\left[\int_0^T e^{\beta s}\left|\delta Z_s\right|^2 ds\right] \leq \frac{\lambda^2}{\lambda^2 - C}\,\left[\mathbb{E}\left[e^{\beta T}\left|\delta Y_T\right|^2\right] + \frac{1}{\mu^2}\mathbb{E}\left(\int_0^T e^{\beta s}\left|\delta_2 f_s\right|^2 ds\right)\right]$$

i.e.,

$$\|\delta Z\|_\beta^2 \leq \frac{\lambda^2}{\lambda^2 - C}\,\left[\mathbb{E}\left[e^{\beta T}\left|\delta Y_T\right|^2\right] + \frac{1}{\mu^2}\left\|\delta_2 f\right\|_\beta^2\right].$$

□

Hence, we have shown the following inequalities holds:

$$\begin{cases} \|\delta Y\|_\beta^2 & \leq T \left[\mathbb{E}\left[e^{\beta T} |\delta Y_T|^2 \right] + \dfrac{1}{\mu^2} \|\delta_2 f\|_\beta^2 \right] \\ \|\delta Z\|_\beta^2 & \leq \dfrac{\lambda^2}{\lambda^2 - C} \left[\mathbb{E}\left[e^{\beta T} |\delta Y_T|^2 \right] + \dfrac{1}{\mu^2} \|\delta_2 f\|_\beta^2 \right] \end{cases}$$

which are vital for the proof of the existence and uniqueness of the solution.

2.3 The Existence and Uniqueness Theorem

As we mention before, El Karoui et al. [16] offered a shorter proof to the theorem of existence and uniqueness of the solution. We prove the theorem using a Priori estimates (Proposition 2.1) and show all steps of the theorem in detail.

Theorem 2.2. *Given a standard parameter (f, ξ) there exists a unique pair $(Y, Z) \in \mathbb{H}_T^2(\mathbb{R}^d) \times \mathbb{H}_T^2(\mathbb{R}^{n \times d})$ which solves the BSDE (2.1).*

Proof. Consider a mapping $\Psi : \mathbb{H}_T^2(\mathbb{R}^d) \times \mathbb{H}_T^2(\mathbb{R}^{n \times d}) \to \mathbb{H}_T^2(\mathbb{R}^d) \times \mathbb{H}_T^2(\mathbb{R}^{n \times d})$. Our target is to use Banach's fixed point theorem (see Appendix, [5]) in order to find a fixed point of the contraction mapping Ψ. The mapping Ψ is defined in the following manner: $\Psi(y, z) = (Y, Z)$, where (Y, Z) is the solution of the BSDE associated with the generator $f(t, y_t, z_t)$ and the terminal condition ξ, which is equivalent to

$$Y_t = \xi + \int_t^T f(s, y_s, z_s) ds - \int_t^T Z_s^* dW_s.$$

The solution (Y, Z) is defined by considering the square-integrable continuous martingale M_t,

$$M_t = \mathbb{E}\left(\int_0^T f(s, y_s, z_s) ds + \xi \,|\, \mathcal{F}_t \right).$$

By Martingale representation theorem for the functionals of Brownian motion, there exists a unique integrable process $Z \in \mathbb{H}_T^2(\mathbb{R}^{n \times d})$, such that

$$\begin{aligned} M_t &= \mathbb{E}(M_t) + \int_0^t Z_s^* dW_s \\ &= \mathbb{E}\left(\int_0^T f(s, y_s, z_s) ds + \xi \right) + \int_0^t Z_s^* dW_s \\ &= M_0 + \int_0^t Z_s^* dW_s. \end{aligned}$$

Let us define the adapted and continouos process Y by

$$\begin{aligned} Y_t &= M_t - \int_0^t f(s, y_s, z_s) ds \\ \Rightarrow Y_t &= \mathbb{E}\left(\int_t^T f(s, y_s, z_s) ds + \xi \,|\, \mathcal{F}_t \right). \end{aligned}$$

Let (y^1, z^1) and (y^2, z^2) be two elements of $\mathbb{H}^2_{T,\beta}(\mathbb{R}^d) \times \mathbb{H}^2_{T,\beta}(\mathbb{R}^d)$, and let (Y^1, Z^1) and (Y^2, Z^2) be the images of them, i.e., the associated solutions, respectively. By Proposition 2.1 with assuming $C = 0$ and $\mu^2 = \beta$, we obtain

$$\begin{aligned}
\|\delta Y\|^2_\beta &\leq T \left(\mathbb{E}\left(e^{\beta T} |\delta Y_T|^2\right) + \frac{1}{\beta} \|\delta_2 f_s\|^2_\beta \right) \\
&= \frac{T}{\beta} \|\delta_2 f_s\|^2_\beta, \qquad (\text{since } \delta Y_T = Y^1_T - Y^2_T = \xi - \xi = 0) \\
&= \frac{T}{\beta} \mathbb{E}\left(\int_0^T e^{\beta s} \left| f(s, y^1_s, z^1_s) - f(s, y^2_s, z^2_s) \right|^2 ds \right).
\end{aligned}$$

Similarly,

$$\begin{aligned}
\|\delta Z\|^2_\beta &\leq \left(\mathbb{E}\left(e^{\beta T} |\delta Y_T|^2\right) + \frac{1}{\beta} \|\delta_2 f_s\|^2_\beta \right) \\
&= \|\delta_2 f_s\|^2_\beta \\
&\leq \frac{1}{\beta} \mathbb{E}\left(\int_0^T e^{\beta s} \left| f(s, y^1_s, z^1_s) - f(s, y^2_s, z^2_s) \right|^2 ds \right).
\end{aligned}$$

Combining these two inequalities, using the inequality $(a + b)^2 \leq 2(a^2 + b^2)$ and the Lipschitz property of f, we have

$$\begin{aligned}
\|\delta Y\|^2_\beta + \|\delta Z\|^2_\beta &\leq \frac{T+1}{\beta} \mathbb{E}\left(\int_0^T e^{\beta s} \left| f(s, y^1_s, z^1_s) - f(s, y^2_s, z^2_s) \right|^2 ds \right) \\
&\leq \frac{T+1}{\beta} \mathbb{E}\left(\int_0^T C e^{\beta s} \left(|\delta y| + |\delta z|\right)^2 ds \right) \\
&\leq \frac{2C(T+1)}{\beta} \mathbb{E}\left(\int_0^T e^{\beta s} \left(|\delta y|^2 + |\delta z|^2\right) ds \right) \\
&= \frac{2C(T+1)}{\beta} \left(\mathbb{E}\left(\int_0^T e^{\beta s} |\delta y|^2 ds \right) + \mathbb{E}\left(\int_0^T e^{\beta s} |\delta z|^2 ds \right) \right) \\
&= \frac{2C(T+1)}{\beta} \left(\|\delta y\|^2_\beta + \|\delta z\|^2_\beta \right)
\end{aligned}$$

$$\Rightarrow \|\delta Y\|^2_\beta + \|\delta Z\|^2_\beta \leq \frac{2C(T+1)}{\beta} \left(\|\delta y\|^2_\beta + \|\delta z\|^2_\beta \right). \tag{2.10}$$

Setting $2C(1 + T) < \beta$, we conclude that Ψ is a contraction mapping from $\mathbb{H}^2_{T,\beta}(\mathbb{R}^d) \times \mathbb{H}^2_{T,\beta}(\mathbb{R}^{n\times d})$ onto itself. By Banach's fixed point theorem in [5] (see Appendix), there exists a fixed point, which is the unique continuous solution of the BSDE in (2.1). □

In the following, we show a related corollary which is stated in [16].

Corollary 2.3. *Assume that β satisfies the inequality that $2C(1 + T) < \beta$. Let (Y^k, Z^k) be the sequence recursively defined with the initial value $(Y_0 = 0; Z_0 = 0)$ and*

$$-dY^{k+1}_t = f(t, Y^k_t, Z^k_t)dt - (Z^{k+1}_t)^* dW_t, \qquad Y^{k+1}_T = \xi.$$

Then the sequence (Y^k, Z^k) *converges to* (Y, Z) *in* $\mathbb{H}^2_{T,\beta}(\mathbb{R}^d) \times \mathbb{H}^2_{T,\beta}(\mathbb{R}^{n\times d})$ *as* $k \to \infty$.

Proof. If (Y^k, Z^k) defined as above, then by (2.10),

$$\begin{aligned}\left\|Y^{k+1} - Y^k\right\|^2_\beta + \left\|Z^{k+1} - Z^k\right\|^2_\beta &\le \frac{2C(T+1)}{\beta}\left(\left\|Y^k - Y^{k-1}\right\|^2_\beta + \left\|Z^k - Z^{k-1}\right\|^2_\beta\right)\\ &\le \left(\frac{2C(T+1)}{\beta}\right)^2\left(\left\|Y^{k-1} - Y^{k-2}\right\|^2_\beta + \left\|Z^{k-1} - Z^{k-2}\right\|^2_\beta\right)\\ &\le \cdots \le \underbrace{\left(\frac{2C(T+1)}{\beta}\right)^k}_{<1}\underbrace{\left(\left\|Y^1 - Y^0\right\|^2_\beta + \left\|Z^1 - Z^0\right\|^2_\beta\right)}_{<\infty}\end{aligned}$$

We have shown that the general term of the sum below is going to zero as $k \to \infty$. For this reason,

$$\sum_{k=0}^{\infty}\left\|Y^{k+1} - Y^k\right\|^2_\beta + \sum_{k=0}^{\infty}\left\|Z^{k+1} - Z^k\right\|^2_\beta < \infty$$

which means $\{Y^k\}_{k=0}^\infty$ and $\{Z^k\}_{k=0}^\infty$ are Cauchy sequences. Recall that, for all k, $Y^k \in \mathbb{R}^d$ and $Z^k \in \mathbb{R}^{n\times d}$, where $\mathbb{R}^d$ and $\mathbb{R}^{n\times d}$ are Banach spaces. Hence, $\{Y^k\}_{k=0}^\infty$ and $\{Z^k\}_{k=0}^\infty$ becomes convergent and converges to Y, Z, respectively (to put it in other words, they converge to the unique solution (Y, Z)). □

2.4 The Solution of the Linear Backward Stochastic Differential Equations

A BSDE which has a linear standard generator is called a linear backward stochastic differential equation and abbreviated by LBSDE. LBSDEs were first appeared in [12]. The existence and uniqueness theorem could be applicable to LBSDE, as well (see the following proposition). For direct constructions along more classical lines, we refer to the reader [16]. In addition, this section can be considerable as a preparation for pricing and hedging problem in Chapter 3.

Proposition 2.4. *Let* (β, γ) *be a bounded* $(\mathbb{R}, \mathbb{R}^n)$*-valued predictable process and* φ *be an element of* $\mathbb{L}^2_T(\mathbb{R})$. *Then, the LBSDE*

$$-dY_t = (\varphi_t + Y_t\beta_t + Z_t^*\gamma_t)\, dt - Z_t^*\, dW_t, \qquad Y_T = \xi \tag{2.11}$$

has a unique solution (Y, Z) *in* $\mathbb{H}^2_{T,\beta}(\mathbb{R}) \times \mathbb{H}^2_{T,\beta}(\mathbb{R}^n)$ *and* Y_t *is given by the closed formula*

$$Y_t = \mathbb{E}\left(\xi\,\Gamma^t_T + \int_t^T \Gamma^t_s\varphi_s\, ds \,|\, \mathcal{F}_t\right), \quad \mathbb{P}\text{-a.s.},$$

where Γ^t_s *is the adjoint process defined for* $s \ge t$ *by the forward linear stochastic differential equation (LSDE)*

$$d\Gamma^t_s = \Gamma^t_s\left(\beta_s\, ds + \gamma^*_s\, dW_s\right), \qquad \Gamma^t_t = 1. \tag{2.12}$$

In particular, if ξ and φ are nonnegative, then the process Y is nonnegative. Moreover, if $Y_0 = 0$, then $Y_t = 0$ a.s., $\xi = 0$ a.s. and $\varphi_t = 0$ $d\mathbb{P} \otimes dt$-a.s., for all t.

Proof. We first show that (f, ξ) is a standard parameter:

$$\begin{aligned} |f(w,t,y_1,z_1) - f(w,t,y_2,z_2)| &= |\varphi_t + \beta_t y_1 + \gamma_t^* z_1 - \varphi_t - \beta_t y_2 - \gamma_t^* z_2| \\ &\leq |\beta_t| \, |y_1 - y_2| + |\gamma_t^*| \, |z_1 - z_2| \,. \end{aligned}$$

By the assumption that β and γ are bounded, say the boundaries K_1 and K_2, respectively. Then,

$$\begin{aligned} |f(w,t,y_1,z_1) - f(w,t,y_2,z_2)| &\leq K_1 \, |y_1 - y_2| + K_2 \, |z_1 - z_2| \\ &\leq \max\{K_1, K_2\} \left(|y_1 - y_2| + |z_1 - z_2|\right). \end{aligned}$$

Setting $C = max\{K_1, K_2\}$, which is constant, the generator f becomes Lipschitz with coefficient C. $f(\cdot, 0, 0)$ belongs to $\mathbb{H}_T^2(\mathbb{R})$, since $\varphi \in \mathbb{H}_T^2(\mathbb{R})$. ξ is already an element of $\mathbb{L}_T^2(\mathbb{R}^d)$. Therefore (f, ξ) is a standard parameter. By Theorem 2.2, there exists a unique solution (Y, Z), which solves the BSDE (2.11).
Now, let us derive the closed formula of Y_t. Our aim is to apply Itò product rule to $Y\Gamma$, where Y is defined as in (2.11) and Γ is given by (2.12):

$$\begin{aligned} d(Y_s \Gamma_s^t) &= Y_s \, d\Gamma_s^t + \Gamma_s^t \, dY_s + d\left\langle Y_s, \Gamma_s^t \right\rangle \\ &= Y_s \Gamma_s^t \left(\beta_s \, ds + \gamma_s^* \, dW_s\right) + \Gamma_s^t \left(\left(-\varphi_s - Y_s\beta_s - Z_s^*\gamma_s\right) ds + Z_s^* \, dW_s\right) \\ &\quad + \Gamma_s^t \gamma_s^* Z_s \, ds \\ &= Y_s \Gamma_s^t \beta_s \, ds + Y_s \Gamma_s^t \gamma_s^* \, dW_s - \Gamma_s^t \varphi_s \, ds - \Gamma_s^t Y_s \beta_s \, ds - \Gamma_s^t Z_s^* \gamma_s \, ds \\ &\quad + \Gamma_s^t Z_s^* \, dW_s + \Gamma_s^t \gamma_s^* Z_s \, ds \\ &= -\Gamma_s^t \varphi_s \, ds + \Gamma_s^t (Y_s \gamma_s + Z_s)^* \, dW_s. \end{aligned}$$

Integrating from $s = t$ to $s = T$, we obtain

$$\underbrace{Y_T}_{=\xi} \Gamma_T^t = Y_t \underbrace{\Gamma_t^t}_{=1} - \int_t^T \Gamma_s^t \varphi_s \, ds + \int_t^T \Gamma_s^t (Y_s \gamma_s + Z_s)^* \, dW_s$$

$$\Rightarrow Y_t = \xi \, \Gamma_T^t + \int_t^T \Gamma_s^t \varphi_s \, ds - \int_t^T \Gamma_s^t (Y_s \gamma_s + Z_s)^* \, dW_s.$$

Taking the conditional expectation with respect to the filtration $\mathcal{F}_t$,

$$\mathbb{E}(Y_t \,|\, \mathcal{F}_t) = \mathbb{E}\left(\xi \, \Gamma_T^t + \int_t^T \Gamma_s^t \varphi_s \, ds \,|\, \mathcal{F}_t\right) - \mathbb{E}\left(\int_t^T \Gamma_s^t (Y_s \gamma_s + Z_s)^* \, dW_s \,|\, \mathcal{F}_t\right).$$

Y_t is $\mathcal{F}_t$-measurable and the stochastic integral is independent from the filtration $\mathcal{F}_t$, that is, $\mathbb{E}\left(\int_t^T \Gamma_s^t (Y_s \gamma_s^* + Z_s^*) \, dW_s \,|\, \mathcal{F}_t\right) = \mathbb{E}\left(\int_t^T \Gamma_s^t (Y_s \gamma_s^* + Z_s^*) \, dW_s\right)$, which is equal to zero. Therefore,

$$Y_t = \mathbb{E}\left(\xi \, \Gamma_T^t + \int_t^T \Gamma_s^t \varphi_s \, ds \,|\, \mathcal{F}_t\right).$$

Moreover, Γ is nonnegative since the dynamics of it is a Geometric SDE, i.e., its solution is exponential. In addition, if ξ and φ are nonnegative, then the process Y becomes nonnegative, since the structure of Y has nonnegative components. Furthermore, if additionally $Y_0 = 0$ then conditional expectation becomes normal expectation

$$\begin{aligned} Y_0 &= \mathbb{E}\left(\xi\,\Gamma_T^0 + \int_0^T \Gamma_s^0 \varphi_s\, ds \,|\, \mathcal{F}_0\right) \\ &= \mathbb{E}\left(\xi\,\Gamma_T^0 + \int_0^T \Gamma_s^0 \varphi_s\, ds\right). \end{aligned}$$

Expectation of nonnegative variable $\xi\,\Gamma_T^0 + \int_0^T \Gamma_s^0 \varphi_s\, ds$ is equal to zero. It implies that $\xi = 0$, $\mathbb{P}$-a.s. and $\varphi = 0$, $d\mathbb{P} \otimes dt$ -a.s., since Γ is exponential and never equals to zero. Hence, $Y_t = 0$ for any t. □

2.5 The Comparison Theorem

The comparison theorem is obtained as a consequence of the Proposition 2.4 in the previous section. The theorem was firstly introduced by Peng [24]. We follow the main article [16] to state the theorem and show all steps of the proof in detail. It will be efficiently used in many parts of the thesis.

Theorem 2.5 (The Comparison Theorem)**.** *Let (f^1, ξ^1) and (f^2, ξ^2) be two standard parameters of BSDEs, and let (Y^1, Z^1) and (Y^2, Z^2) be the associated square-integrable solutions. Suppose that*

- $\xi^1 \geq \xi^2, \quad \mathbb{P} - a.s..$
- $\delta_2 f_t = f^1(t, Y_t^2, Z_t^2) - f^2(t, Y_t^2, Z_t^2) \geq 0, \quad d\mathbb{P} \otimes dt - a.s..$

Then we have $Y_t^1 \geq Y_t^2$, almost surely for any time t.
Moreover the comparison is strict; that is, if, in addition, $Y_0^1 = Y_0^2$, then $\xi^1 = \xi^2$, $f^1(t, Y_t^2, Z_t^2) = f^2(t, Y_t^2, Z_t^2)$, $d\mathbb{P} \otimes dt$ -a.s., and $Y^1 = Y^2$ a.s. More generally, if $Y_t^1 = Y_t^2$ on a set $A \in \mathcal{F}_t$, then $Y_s^1 = Y_s^2$ almost surely on $[t, T] \times A$, $\xi^1 = \xi^2$ a.s. on A, and $f^1(t, Y_t^2, Z_t^2) = f^2(t, Y_t^2, Z_t^2)$ on $A \times [t, T]$, $d\mathbb{P} \otimes ds$-a.s.

Proof. Let $\delta Y = Y^1 - Y^2$, $\delta Z = Z^1 - Z^2$ and $\delta_2 f_t = f^1(t, Y_t^2, Z_t^2) - f^2(t, Y_t^2, Z_t^2)$. In order to obtain $(\delta Y, \delta Z)$, we use the solutions (Y^1, Z^1) and (Y^2, Z^2) of the following BSDEs:

$$\begin{aligned} -dY_t^1 &= f^1(t, Y_t^1, Z_t^1)\, dt - Z_t^{1*}\, dW_t, \qquad Y_T^1 = \xi^1; \\ -dY_t^2 &= f^2(t, Y_t^2, Z_t^2)\, dt - Z_t^{2*}\, dW_t, \qquad Y_T^2 = \xi^2. \end{aligned}$$

Derive the BSDE of δY as follows:

$$
\begin{aligned}
-d\delta Y_t &= -d(Y_t^1 - Y_t^2) = -dY_t^1 + dY_t^2 \\
&= f^1(t, Y_t^1, Z_t^1)\,dt - Z_t^{1*}\,dW_t - f^2(t, Y_t^2, Z_t^2)\,dt + Z_t^{2*}\,dW_t \\
&= \left(f^1(t, Y_t^1, Z_t^1) - f^2(t, Y_t^2, Z_t^2)\right)\,dt - \delta Z_t^*\,dW_t \\
&\quad \pm f^1(t, Y_t^2, Z_t^1)\,dt \\
&= \left(f^1(t, Y_t^1, Z_t^1) - f^1(t, Y_t^2, Z_t^1)\right)\,dt + \left(f^1(t, Y_t^2, Z_t^1) - f^2(t, Y_t^2, Z_t^2)\right)\,dt \\
&\quad - \delta Z_t^*\,dW_t \pm f^1(t, Y_t^2, Z_t^2)\,dt \\
&= \left(f^1(t, Y_t^1, Z_t^1) - f^1(t, Y_t^2, Z_t^1)\right)\,dt + \left(f^1(t, Y_t^2, Z_t^1) - f^1(t, Y_t^2, Z_t^2)\right)\,dt \\
&\quad + \left(f^1(t, Y_t^2, Z_t^2) - f^2(t, Y_t^2, Z_t^2)\right)\,dt - \delta Z_t^*\,dW_t \\
&= \left(f^1(t, Y_t^1, Z_t^1) - f^1(t, Y_t^2, Z_t^1)\right)\,dt + \left(f^1(t, Y_t^2, Z_t^1) - f^1(t, Y_t^2, Z_t^2)\right)\,dt \\
&\quad + \delta_2 f_t\,dt - \delta Z_t^*\,dW_t.
\end{aligned}
$$

Set

$$
\begin{aligned}
\Delta_y f^1(t) &= \frac{f^1(t, Y_t^1, Z_t^1) - f^1(t, Y_t^2, Z_t^1)}{Y_t^1 - Y_t^2}, \\
\Delta_z f^{1,i}(t) &= \frac{f^1(t, Y_t^2, \tilde{Z}_t^{i-1}) - f^1(t, Y_t^2, \tilde{Z}_t^i)}{Z_t^1 - Z_t^2},
\end{aligned}
$$

where $\tilde{Z}_t^i$ is a vector such that first i components come from the components of Z_t^2 and the remainder $n-i$ ones come from Z_t^1,i.e., $\tilde{Z}_t^i = (Z_t^{2,1}, \cdots, Z_t^{2,i}, Z_t^{1,i+1}, \cdots, Z_t^{1,n})$. Therefore, the equation becomes

$$
\begin{aligned}
-d\delta Y_t &= \left(\Delta_y f^1(t)\delta Y_t + \Delta_z f^1(t)^* \delta Z_t + \delta_2 f_t\right)\,dt - \delta Z_t^*\,dW_t, \\
\delta Y_T &= \xi^1 - \xi^2.
\end{aligned}
\tag{2.13}
$$

Because of the fact that f^1 is Lipschitz, $\Delta_y f^1(t)$ and $\Delta_z f^1(t)$ are bounded. Indeed,

$$
\begin{aligned}
\left|\Delta_y f^1(t)\right| &= \left|\frac{f^1(t, Y_t^1, Z_t^1) - f^1(t, Y_t^2, Z_t^1)}{Y_t^1 - Y_t^2}\right| \\
&\le \frac{1}{|Y_t^1 - Y_t^2|}\, C\left(\left|Y_t^1 - Y_t^2\right| + \left|Z_t^1 - Z_t^1\right|\right) = C.
\end{aligned}
$$

Note that $\Delta_z f^1(t)$ is bounded by C in a similar manner. As we proved in Proposition 2.4, there exists a unique solution to the LBSDE (2.13) as follows:

$$
\delta Y_t = \mathbb{E}\left((\xi^1 - \xi^2)\Gamma_T + \int_t^T \Gamma_s \delta_2 f_s\,ds | \mathcal{F}_t\right),
$$

where Γ is adjoint process satisfies the forward LSDE

$$
d\Gamma_s = \Gamma_s\left(\Delta_y f^1(s)\,ds + \Delta_z f^1(s)^*\,dW_s\right).
$$

Remember that Γ is Geometric Brownian Motion (GBM) and has nonnegative solution. By hypothesis, we have $\xi^1 - \xi^2 \ge 0$ and $\delta_2 f_s \ge 0$ and additonally Γ is

GBM then δY_t must be greater than or equal to zero almost surely ,i.e., $Y^1 \geq Y^2$ a.s for any time t.
Moreover, if $\delta Y_0 = Y_0^1 - Y_0^2 = 0$, then it implies that $\xi^1 - \xi^2 = 0$ and $\delta_2 f_t = 0$ which exactly means that $\xi^1 = \xi^2$ and $f^1(t, Y_t^2, Z_t^2) = f^2(t, Y_t^2, Z_t^2)$, $\mathrm{d}\mathbb{P} \otimes dt$ -a.s., and $Y^1 = Y^2$ a.s. by the Proposition 2.4. More generally, one can say that if $Y_t^1 = Y_t^2$ on a set $A \in \mathcal{F}_t$, then $Y_s^1 = Y_s^2$ almost surely on $[t, T] \times A$, $\xi^1 = \xi^2$ a.s. on A, and $f^1(t, Y_t^2, Z_t^2) = f^2(t, Y_t^2, Z_t^2)$ on $A \times [t, T]$, $d\mathbb{P} \otimes ds-$a.s. $\square$

After we prove outstanding results of the backward stochastic differential equations theory, we explain the concept of pricing and hedging contingent claims in a complete market and obtain the fair price as a unique solution to a linear backward stochastic differential equation in the following chapter.

CHAPTER 3

PRICING AND HEDGING CLAIMS WITH THE LINEAR BACKWARD STOCHASTIC DIFFERENTIAL EQUATIONS

Backward stochastic differential equations have been efficiently used in mathematical finance since Black and Scholes [6]. For instance, the theory of pricing and hedging a contingent claim could be modelled by linear BSDEs. Consumption did not take place at the earlies of the theory. In the following years, Merton [21, 22] considered a consumption process which is increasing, right-continuous and adapted with null at zero. In this chapter, we closely follow the main article [16] in order to define the structure of the fair (upper) price and hedging strategy (superstrategy) for a nonnegative contingent claim in a complete market.

3.1 Terminology

We first introduce the terminology that we use in our model for the financial market. Assume that the market consists of $n+1$ assets. One of them is the riskless asset, for instance treasury bond, with price per unit P^0 satisfying the equation

$$dP_t^0 = P_t^0 r_t dt; \qquad P_0^0 = 1, \tag{3.1}$$

with a short interest rate r_t for all $t \in [0,T]$. The remainder n assets are risky securities (or the stocks) which are allowed to be traded continuously. P^i denotes the ith stock price process per share and modeled by the linear stochastic differential equation as follows:

$$dP_t^i = P_t^i \left(b_t^i \, dt + \sum_{j=1}^{n} \sigma_t^{i,j} dW_t^j \right),$$

where n-dimensional column vector $W = (W^1, \dots, W^n)^*$ is a standard Brownian motion with values in $\mathbb{R}^n$, defined on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0\le t\le T}, \mathbb{P})$. Since we have n risky assets, we model with an n-dimensional standard Brownian motion in order to be in a complete market. Moreover, the probability measure $\mathbb{P}$ is assumed to be objective. The information up to time t is given by the augmented right- continuous filtration $\{\mathcal{F}_t\}_{0\le t\le T}$. $\{\mathcal{F}_t\}$ is considered as

σ-algebra generated by the Brownian motion $W = (W^1, \ldots, W^n)^*$. In addition, the following hypothesis are assumed throught this chapter:

- The short rate r is a predictable, bounded and nonnegative process.
- The drift terms (or the stock appreciation rates) $b = (b^1, \ldots, b^n)^*$ is an n-dimensional column vector process supposed to be predictable and bounded.
- The volatility matrix $\sigma = (\sigma^{i,j})$ is an $n \times n$ matrix process and assumed to be predictable and bounded.
- There exists an n dimensional predictable and bounded-valued process vector θ, called a risk premium, such that

$$b_t - r_t \mathbf{1} = \sigma_t \theta_t, \qquad d\mathbb{P} \otimes dt \quad \text{a.s},$$

  where $\mathbf{1}$ is an n-dimensional column vector with all components 1.

The market becomes dinamically complete under these assumptions.

3.2 Concept of Self-Financing Strategies

Assuming the all hypothesis stated in the previous section, consider a small investor who has wealth V_t at any time $t \in [0, T]$ and cannot effect the market prices. Let π_t^i denote the amount of money invested in the ith stock, $i = 1, \ldots, n$. His decision can only be based on current information $\mathcal{F}_t$, i.e., $\pi = (\pi^1, \ldots, \pi^n)^*$ and $\pi^0 = V - \sum_{i=1}^n \pi^i$ are adapted. Here, π^0 corresponds to the amount of money that invested in riskless asset with price per unit P^0.
Harrison and Pliska [14] introduce the concept of self-financing strategy as follows: The strategy (V, π) satisfies the following equation

$$V_t = V_0 + \int_0^t \sum_{i=0}^n \pi_t^i \frac{dP_t^i}{P_t^i} \tag{3.2}$$

is said to be a self-financing strategy. Using the equalities given before:

$$\frac{dP_t^0}{P_t^0} = r_t\, dt; \qquad \frac{dP_t^i}{P_t^i} = b_t^i\, dt + \sum_{j=1}^n \sigma_t^{i,j} dW_t^j, \; i = 1, \ldots, n$$

the linear stochastic differential equation in (3.2) is equal to

$$\begin{aligned}
dV_t &= \sum_{i=0}^{n} \pi_t^i \frac{dP_t^i}{P_t^i} \\
&= \pi_t^0 \frac{dP_t^0}{P_t^0} + \sum_{i=1}^{n} \pi_t^i \frac{dP_t^i}{P_t^i} \\
&= \pi_t^0 r_t \, dt + \sum_{i=1}^{n} \pi_t^i \left(b_t^i \, dt + \sum_{j=1}^{n} \sigma_t^{i,j} dW_t^j \right) \\
&= \left(V_t - \sum_{i=1}^{n} \pi_t^i \right) r_t \, dt + \sum_{i=1}^{n} \pi_t^i \left(b_t^i \, dt + \sum_{j=1}^{n} \sigma_t^{i,j} dW_t^j \right) \\
&= r_t V_t \, dt + \sum_{i=1}^{n} \pi_t^i \left(b_t^i - r_t \mathbf{1} \right) \, dt + \sum_{i=1}^{n} \pi_t^i \sum_{j=1}^{n} \sigma_t^{i,j} dW_t^j \\
&= r_t V_t \, dt + \pi_t^* (b_t - r_t \mathbf{1}) \, dt + \pi_t^* \sigma_t \, dW_t \\
&= r_t V_t \, dt + \pi_t^* \sigma_t \theta_t \, dt + \pi_t^* \sigma_t \, dW_t \\
&= r_t V_t \, dt + \pi_t^* \sigma_t (\theta_t \, dt + dW_t).
\end{aligned}$$

Note that, in the model proposed by Merton [20], he considers a consumption process $C_t = \int_0^t c_t \, dt$, where c_t is predictable. In this way, the market value process becomes

$$dV_t = r_t V_t \, dt + \pi_t^* \sigma_t (\theta_t \, dt + dW_t) - c_t \, dt.$$

Let us summarize the concept of self-financing strategies in the sense of [16].

Definition 3.1. A self financing trading strategy is a pair (V, π) where V is the market value and $\pi = (\pi^1, \dots, \pi^n)^*$ is the portfolio process satisfying

$$dV_t = r_t V_t \, dt + \pi_t^* \sigma_t (\theta_t \, dt + dW_t), \qquad \int_0^T |\sigma_t^* \, \pi_t|^2 \, dt < \infty, \quad \mathbb{P}\text{-a.s.}$$

Definition 3.2. A self financing superstrategy is a vector process (V, π, C) where V is the market value, $\pi = (\pi^1, \dots, \pi^n)^*$ is the portfolio process and C is the cumulative consumption process satisfying

$$dV_t = r_t V_t \, dt + \pi_t^* \sigma_t (\theta_t \, dt + dW_t) - dC_t, \qquad \int_0^T |\sigma_t^* \, \pi_t|^2 \, dt < \infty, \quad \mathbb{P}\text{-a.s.},$$

where C is an increasing, right-continuous, adapted process with the initial value $C_0 = 0$.

Definition 3.3. If the nonnegative wealth constraint holds, i.e.,

$$V_t \geq 0, \quad t \in [0, T],$$

then the strategy (either self-financing strategy or superstrategy) is called *feasible*.

Now, we are ready to define hedging strategies (superstrategies) and the fair price (upper price).

3.3 Hedging Strategies and Pricing Contingent Claims

Suppose that we have a contingent claim ξ which is nonnegative and $\mathcal{F}_T$-measurable random variable. In the sense of arbitrage-free market, we invest the amount of money as an initial endowment in the $n+1$ assets and then the hedging portfolio must pay at least ξ at maturity. In this section, we give a definition to hedging strategies (hedging superstrategies) against ξ. We then classify them to explain the structures of the fair price (upper price). We closely follow Karatzas and Shreve [9] for the following definitions.

Definition 3.4. A hedging strategy against a contingent claim $\xi \geq 0$ is a feasible self-financing strategy (V, π) such that $V_T = \xi$. Let $\mathcal{H}(\xi)$ denote the class of all hedging strategies against ξ, then the *fair price* X_0 is defined as follows

$$X_0 = \inf\{x \geq 0 : \quad \exists (V, \pi) \in \mathcal{H}(\xi) \text{ such that } V_0 = x\}.$$

Definition 3.5. A hedging superstrategy against a contingent claim $\xi \geq 0$ is a feasible self-financing superstrategy (V, π, C) such that $V_T = \xi$. Let $\mathcal{H}'(\xi)$ denote the class of all hedging superstrategies against ξ, then the *upper price* X_0' is defined as follows

$$X_0{}' = \inf\{x \geq 0 : \quad \exists (V, \pi, C) \in \mathcal{H}'(\xi) \text{ such that } V_0 = x\}.$$

Moreover,

$$X_0{}' \geq e^{-rT}\mathbb{E}^{\mathbb{Q}}(\xi),$$

where $\mathbb{Q}$ is the risk neutral probability measure.

We deduce that for any square-integrable nonnegative contingent claim ξ, $\mathcal{H}(\xi) \neq \emptyset$. Hence the market is complete.
According to the main article [16], we show that the unique solution of a given linear BSDE coincides the fair price in the following theorem.

Theorem 3.1. *Let ξ be a nonnegative square-integrable contingent claim. Then there exists a hedging strategy (X, π) against ξ satisfying the LBSDE*

$$dX_t = r_t X_t\, dt + \pi_t^* \sigma_t \theta_t\, dt + \pi_t^* \sigma_t\, dW_t, \qquad X_T = \xi, \tag{3.3}$$

such that the market value X is the fair price and the upper price of the claim. Let $(H_s^t; s \geq t)$ be the deflator started at time t, that is,

$$dH_s^t = -H_s^t \left(r_s\, ds + \theta_s^*\, dW_s\right), \qquad H_t^t = 1. \tag{3.4}$$

Then

$$X_t = \mathbb{E}\left(H_T^t\, \xi | \mathcal{F}_t\right), \quad a.s.. \tag{3.5}$$

Proof. Recall that X_s and H_s^t, or more simply say H_s satisfy the following dynamics:

$$\begin{aligned} dX_s &= r_s X_s\, ds + \pi_s^* \sigma_s \theta_s\, ds + \pi_s^* \sigma_s\, dW_s \\ dH_s &= -H_s \left(r_s\, ds + \theta_s^*\, dW_s\right) \end{aligned}$$

for all $t \leq s \leq T$. Using Itô' s lemma, let us first calculate the product of the market value process X_s and the associated deflator H_s, $d(H_s X_s)$, as follows:

$$\begin{aligned}
d(H_s X_s) &= X_s\, dH_s + H_s\, dX_s + d\left\langle X_s, H_s \right\rangle \\
&= -X_s H_s \left(r_s\, ds + \theta_s^*\, dW_s\right) + H_t \left(r_s X_s\, ds + \pi_s^* \sigma_s \theta_s\, ds + \pi_s^* \sigma_s\, dW_s\right) \\
&\quad - H_s \pi_s^* \sigma_s \theta_s\, ds \\
&= -X_s H_s r_s\, ds - X_s H_s \theta_s^*\, dW_s + H_s X_s r_s\, ds + H_s \pi_s^* \sigma_s \theta_s\, ds \\
&\quad + H_s \pi_s^* \sigma_s\, dW_s - H_s \pi_s^* \sigma_s \theta_s\, ds \\
&= -X_s H_s \theta_s^*\, dW_s + H_s \pi_s^* \sigma_s\, dW_s \\
&= H_s \left(-X_s \theta_s + \sigma_s^* \pi_s\right)^*\, dW_s.
\end{aligned}$$

Define $U_s = H_s \left(-X_s \theta_s + \sigma_s^* \pi_s\right)$ and integrate from t to T. Then we obtain

$$H_T X_T = H_t X_t + \int_t^T U_s^*\, dW_s.$$

Taking the conditional expectation with respect to the filtration $\mathcal{F}_t$ yields

$$\mathbb{E}\left(H_T X_T \,|\, \mathcal{F}_t\right) = \mathbb{E}\Big(H_t\, X_t \,|\, \mathcal{F}_t\Big) + \mathbb{E}\left(\int_t^T U_s^*\, dW_s \,|\, \mathcal{F}_t\right).$$

Since $H_t\, X_t \in \mathcal{F}_t$, $\int_t^T U_s^* dW_s$ is independent of the history $\mathcal{F}_t$ and the conditional expectation of stochastic integral equals to zero, we have

$$H_t X_t = \mathbb{E}\left(H_T X_T \,|\, \mathcal{F}_t\right) = \mathbb{E}\left(H_T \xi \,|\, \mathcal{F}_t\right).$$

Hence, $H_t X_t$ is the continuous version of the uniformly integrable nonnegative martingale $\mathbb{E}\left(H_T \xi \,|\, \mathcal{F}_t\right)$. By the Martingale representation property for the functional of Brownian motion (see Appendix), there exists a predictable and bounded process $\{V_t\}_{0 \leq t \leq T}$ such that

$$\begin{aligned}
H_t X_t &= \mathbb{E}\left(\mathbb{E}\left(H_T \xi \,|\, \mathcal{F}_t\right)\right) + \int_0^t V_s^*\, dW_s \\
&= \mathbb{E}\left(H_T \xi\right) + \int_0^t V_s^*\, dW_s, \qquad \int_0^T |V_t|^2\, dt < \infty.
\end{aligned}$$

Here V_t is exactly equal to U_t, since the SDE of the last equation is equal to

$$d(H_s X_s) = V_s^* dW_s = U_s^* dW_s.$$

Put $\pi_s = (\sigma_s^*)^{-1}\left(H_s^{-1}U_s + X_s\theta_s\right)$, then we get

$$\begin{aligned}
H_tX_t &= \mathbb{E}\left(H_T\xi\right) + \int_0^t V_s^*\, dW_s \\
&= \mathbb{E}\left(H_T\xi\right) + \int_0^t U_s^*\, dW_s \\
&= \mathbb{E}\left(H_T\xi\right) + \int_0^t H_s\left(-X_s\theta_s + \sigma_s^*\pi_s\right)^*\, dW_s \\
&= \mathbb{E}\left(H_T\xi\right) + \int_0^t H_s\left(-X_s\theta_s + \sigma_s^*(\sigma_s^*)^{-1}\left(H_s^{-1}U_s + X_s\theta_s\right)\right)^*\, dW_s \\
&= \mathbb{E}\left(H_T\xi\right) + \int_0^t H_s\left(-X_s\theta_s + H_s^{-1}U_s + X_s\theta_s\right)^*\, dW_s \\
&= \mathbb{E}\left(H_T\xi\right) + \int_0^t H_sH_s^{-1}U_s^*\, dW_s \\
&= \mathbb{E}\left(H_T\xi\right) + \int_0^t U_s^*\, dW_s.
\end{aligned}$$

By Itô's lemma, (X, π) satisfies the linear BSDE (3.3). Moreover, we have to show that the constraint $\int_0^T |\sigma_t^*\pi_t|^2\, dt < \infty$ holds. After some arrangements of the choice of π we have:

$$\begin{aligned}
\sigma_t^*\pi_t &= \left(H_t^{-1}U_t + X_t\theta_t\right) \\
|\sigma_t^*\pi_t|^2 &= |H_t^{-1}U_t + X_t\theta_t|^2 \leq 2(|H_t^{-1}U_t|^2 + |X_t\theta_t|^2) \\
\int_0^T |\sigma_t^*\pi_t|^2\, dt &\leq 2\left(\int_0^T |H_t^{-1}U_t|^2\, dt + \int_0^T |X_t\theta_t|^2\, dt\right).
\end{aligned}$$

Since the continuity of the processes H and X, and the boundness of θ and U, $\int_0^T |\sigma_t^*\pi_t|^2\, dt < \infty$, which shows (X, π) is a hedging strategy against ξ with the initial endowment $X_0 = \mathbb{E}\left(H_T\xi\right)$.
Furthermore, let (V, φ, C) be a superhedging strategy against ξ. The proof is similar to the hegding strategy case. We first use Itô' s lemma for the product of the cádlág semimartingale V and the continuous semimartingale H.

$$\begin{aligned}
d(V_sH_s) &= V_sdH_s + H_sdV_s + d\left\langle V_s, H_s\right\rangle \\
&= V_s(-H_s\left(r_s\, ds + \theta_s^*\, dW_s\right)) + H_s\left(r_sV_s\, ds - dC_s + \pi_s^*\sigma_s(\theta_s\, ds + dW_s)\right) \\
&\quad - H_s\pi_s^*\sigma_s\theta_s\, ds \\
&= -V_sH_sr_s\, ds - V_sH_s\theta_s^*\, dW_s + H_sr_sV_s\, ds - H_sdC_s + H_s\pi_s^*\sigma_s\theta_s\, ds \\
&\quad + H_s\pi_s^*\sigma_sdW_s - H_s\pi_s^*\sigma_s\theta_s\, ds \\
&= -V_sH_s\theta_s^*\, dW_s + H_s\pi_s^*\sigma_sdW_s - H_sdC_s \\
&= H_s(-V_s\theta_s + \sigma_s^*\pi_s)^*\, dW_s - H_sdC_s.
\end{aligned}$$

Define $U_s^V = H_s(-V_s\theta_s + \sigma_s^*\pi_s)$ and integrate from t to T,

$$V_TH_T = V_tH_t + \int_t^T (U_s^V)^*dW_s - \int_t^T H_sdC_s.$$

Take the conditional expectation with respect to the filtration $\mathcal{F}_t$,

$$\mathbb{E}\left(V_T H_T|\mathcal{F}_t\right) = \mathbb{E}\left(V_t H_t|\mathcal{F}_t\right) + \mathbb{E}\left(\int_t^T (U_s^V)^* dW_s|\mathcal{F}_t\right) - \mathbb{E}\left(\int_t^T H_s dC_s|\mathcal{F}_t\right).$$

Here, $V_t H_t \in \mathcal{F}_t$, $\int_t^T (U_s^V)^* dW_s$ is independent of the history $\mathcal{F}_t$ and the expectation of stochastic integral equals to zero. Therefore,

$$\begin{aligned} V_t H_t &= \mathbb{E}\left(V_T H_T|\mathcal{F}_t\right) + \mathbb{E}\left(\int_t^T H_s dC_s|\mathcal{F}_t\right) \\ &\leq \mathbb{E}\left(V_T H_T|\mathcal{F}_t\right) = X_t H_t. \end{aligned}$$

Hence, $V_t H_t$ becomes a submartingale and X_t is the upper price, that is,

$$V_t \leq X_t, \qquad t \in [0, T].$$

Particularly, for $t = 0$, $V_0 \leq \mathbb{E}\left(H_T \xi\right) = X_0$. □

Remark 3.1. The representation of X_t in terms of H_t and ξ in (3.5) is associated with the well-known property that the fair price of a contingent claim ξ is equal to the expectation of the discounted value of the claim under the risk neutral probability measure $\mathbb{Q}$:

$$X_t = \mathbb{E}_{\mathbb{Q}}\left(e^{-\int_t^T r_s\, ds}\xi|\mathcal{F}_t\right).$$

$\mathbb{Q}$ is called the risk neutral probability measure with Radon-Nikodym derivative with respect to $\mathbb{P}$ on $\mathcal{F}_T$, given by

$$\frac{d\,\mathbb{Q}}{d\,\mathbb{P}} = exp\left(-\int_0^T \theta_s^*\, dW_s - \int_0^T |\theta_s|^2\, ds\right)$$

and $\mathbb{Q}$ is martingale measure, that is, the discounted wealth processes are $\mathbb{Q}$-martingales.

Up to this point, we explain the theory of backward stochastic differential equations and pricing and hedging a nonnegative contingent claim modeled by a linear backward stochastic differential equation. From now on, we introduce the optimal control theory, then we deal with stochastic control problems and associate them to the backward stochastic differential equation theory.

CHAPTER 4

STOCHASTIC CONTROL AND BACKWARD STOCHASTIC DIFFERENTIAL EQUATIONS

It is remarkable fact that backward stochastic differential equations are considerable as solving stochastic control problems (see for instance the main article [16]). In this chapter, we give a brief introduction to the optimal control theory following [26]. Besides, we derive the well-known Hamilton-Jacobi-Bellman equations. For direct constructions along more classical lines, we refer to the reader [24] and [25]. After that, we introduce the stochastic control problems. The standard work on this subject can be found in [13]. Last but not least, we analyse the optimization of solution in cases of concavity or convexity of the standard generator.

4.1 Preliminaries

We study with a system such that the state process $\{X_s\}_{t\leq s\leq T}$ is finite n-dimensional Markovian diffusion and satisfies the following dynamics:

$$dX_s = b(s, X_s, u_s)ds + \sigma(s, X_s, u_s)dW_s, \qquad X_t = x, \tag{4.1}$$

where

- u is a k-dimensional control process.
- $b = (b_1, b_2, \ldots, b_n)$ is an n-dimensional generator such that

$$b : \mathbb{R}^+ \times \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R}^n.$$

- σ is an $n \times n$ matrix function defined by

$$\sigma : \mathbb{R}^+ \times \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R}^{n\times d}.$$

- b and σ are supposed to be uniformly Lipschitz with respect to x and u.
- $b(s, 0, 0)$ and $\sigma(s, 0, 0)$ are uniformly bounded deterministic functions.

- W is an n-dimensional Brownian motion.

The control processes $\{u_s\}_{t\leq s\leq T}$ are supposed to be in a Polish space U. We will restrict ourselves to a class of control processes which has the form:

$$u_s = u(s, X_s).$$

In other words, u has a connection with the past values of state process X. As a consequence, u is called *feedback* control.

Definition 4.1. A control process is said to be admissible if the SDE in (4.1) has a unique solution. $\mathcal{U}$ denotes the class of all admissible control processes.

For a given control process u, a criterion or objective function (more common of usage) is defined by

$$Y(t,x,u) = \mathbb{E}_{t,x}\left[\int_t^T L(s, X_s, u_s)ds + \Psi(T, X_T)\right],$$

where $\mathbb{E}_{t,x}$ denotes the conditional expectation with respect to the time t and the state at time t $X_t = x$. The deterministic function $L(s,x,u)$ defines the running cost accociated with control process u and state x, while $\Psi(T,x)$ corresponds to the terminal cost. L and Ψ are defined by

$$\begin{aligned} &L : \mathbb{R}^+ \times \mathbb{R}^n \times \mathbb{R}^k \to \mathbb{R}, \\ &\Psi : \mathbb{R}^n \to \mathbb{R}. \end{aligned}$$

The problem is to minimize the objective function $Y(t,x,u)$ over all the feedback admissible control processes and to find (if there exists) a feedback control process u^0 which achieves the minimization. The value function $\bar{Y}(t,x)$ is defined by

$$\bar{Y}(t,x) = \min_{u\in\mathcal{U}} Y(t,x,u).$$

If for any $(t,x) \in \mathbb{R}^+ \times \mathbb{R}^n$, there exists a control u^0 such that

$$\bar{Y}(t,x) = Y(t,x,u^0),$$

then u^0 is called *optimal feedback control process.*

4.2 Hamilton-Jacobi-Bellman Equations

After we give a brief introduction to the optimal control theory, we derive the Dynamic Programming Principle and the Hamilton-Jacobi-Bellman (HJB) equation which is stated in the following theorem. The theorem is modified from the article [13].

Theorem 4.1 (Hamilton-Jacobi-Bellman Equation)**.** *Suppose that there exists an optimal control process u^0 and the optimal value function $\bar{Y}$ is regular in the sense that $\bar{Y} \in C^{1,2}$. Then the followings hold:*

1. $\bar{Y}$ *satisfies the Hamilton-Jacobi-Bellman equation*

$$\begin{cases} \partial_t \bar{Y}(t, X_t) + \inf_{u \in \mathcal{U}} \left\{ L(t, x, u) + \mathcal{L}^u \bar{Y}(t, x) \right\} &= 0, \qquad \forall (t, x) \in [0, T] \times \mathbb{R}^n \\ \bar{Y}(T, x) &= \Psi(x), \qquad \forall x \in \mathbb{R}^n, \end{cases}$$

where the operator $\mathcal{L}^u$ *is defined by*

$$\mathcal{L}^u_{t,x} = \sum_{i=1}^n b_i(t, x, u) \partial_{x_i} + \frac{1}{2} \sum_{i,j=1}^n [\sigma\sigma^*]_{ij}(t, x, u) \partial^2_{x_i x_j}.$$

2. *For each* $(t, x) \in [0, T] \times \mathbb{R}^n$ *the infimum in the HJB equation above is attained by* $u = u^0(t, x)$.

Proof. Suppose that a controller generates a strategy (say Strategy I) which uses a pairwise control $u_1(r, x)$ defined by

$$u_1(r, x) = \begin{cases} u(r, x), & \text{if } t \le r \le t + h, \\ u^0(r, x), & \text{if } t + h < r \le T, \end{cases}$$

where u^0 is optimal. In other words, the controller uses $u(r, x)$ for $r \in [t, t + h]$; and $u^0(r, x)$ for $r \in [t + h, T]$. In this way,

$$\begin{aligned} Y(t, x, u_1) &= \mathbb{E}_{t,x} \left[\int_t^T L(s, X_s, u_s^1) ds + \Psi(T, X_T) \right] \\ &= \mathbb{E}_{t,x} \left[\int_t^{t+h} L(s, X_s, u_s) ds \right] + \mathbb{E}_{t,x} \left[\int_{t+h}^T L(s, X_s, u_s^0) ds + \Psi(T, X_T) \right] \\ &= \mathbb{E}_{t,x} \left[\int_t^{t+h} L(s, X_s, u_s) ds \right] + Y(t + h, x, u^0) \\ &= \mathbb{E}_{t,x} \left[\int_t^{t+h} L(s, X_s, u_s) ds + Y(t + h, x, u^0) \right]. \end{aligned}$$

Here, we can also write $Y(t + h, X_{t+h}, u^0) = \bar{Y}(t + h, X_{t+h})$, since u^0 is optimal for range $[t + h, T]$.
Let us call the way that the controller choose the optimal control for the whole interval $[0, T]$ as Strategy II. Then by definition, Strategy II is the optimal one. His expected performance criterion cannot be less than $\bar{Y}(t, x)$, that is, $\bar{Y}(t, x) \le Y(t, x, u_1)$. For this reason,

$$\begin{aligned} \bar{Y}(t, x) \le Y(t, x, u_1) &= \mathbb{E}_{t,x} \left[\int_t^{t+h} L(s, X_s, u_s) ds + Y(t + h, x, u^0) \right] \\ &= \mathbb{E}_{t,x} \left[\int_t^{t+h} L(s, X_s, u_s) ds + \bar{Y}(t + h, X_{t+h}) \right]. \qquad (4.2) \end{aligned}$$

Equality holds for $u = u^0$. We have just derived the Dynamic Programming Principle in [13] as follows:

$$\bar{Y}(t, x) = \inf_{u \in \mathcal{U}} \mathbb{E}_{t,x} \left[\int_t^{t+h} L(s, X_s, u_s) ds + \bar{Y}(t + h, X_{t+h}) \right].$$

Continue from inequality (4.2),

$$0 \leq \mathbb{E}_{t,x}\left[\int_t^{t+h} L(s, X_s, u_s)ds + \bar{Y}(t+h, X_{t+h}) - \bar{Y}(t,x)\right]. \tag{4.3}$$

Apply Itô' s formula to $\bar{Y}(s, X_s)$ integrate from $s = t$ to $s = t + h$,

$$\begin{aligned}
\bar{Y}(t+h, X_{t+h}) &= \bar{Y}(t,x) + \int_t^{t+h} \partial_s \bar{Y}(s, X_s)ds + \sum_{i=1}^{n} \int_t^{t+h} \partial_{x_i} \bar{Y}(s, X_s)dX_i \\
&\quad + \frac{1}{2}\sum_{i,j=1}^{n} \int_t^{t+h} \partial^2_{x_i x_j} \bar{Y}(s, X_s) < dX_i, dX_j > \\
&= \bar{Y}(t,x) + \int_t^{t+h} \partial_s \bar{Y}(s, X_s)ds + \sum_{i=1}^{n} \int_t^{t+h} \partial_{x_i} \bar{Y}(s, X_s) b_i(s,x,u)ds \\
&\quad + \sum_{i=1}^{n} \int_t^{t+h} \partial_{x_i} \bar{Y}(s, X_s)\sigma_i(s,x,u)dW_s \\
&\quad + \frac{1}{2}\sum_{i,j=1}^{n} \int_t^{t+h} \partial^2_{x_i x_j} \bar{Y}(s, X_s)[\sigma\sigma^*]_{ij}ds,
\end{aligned}$$

where ∂_s is the partial derivative of b with respect to s, similarly ∂_{x_i} is the partial derivative of b with respect to x_i and $\partial^2_{x_i x_j}$ is the partial second derivative of b with respect to x_i and x_j. σ_i denotes the first row of $n \times n$ matrix σ. Finally, $[\sigma\sigma^*]_{ij}$ defines the ij-th component of the matrix $[\sigma\sigma^*]$.
Plug the equation derived from Itô' s formula into inequality (4.3),

$$\begin{aligned}
0 &\leq \mathbb{E}_{t,x}\left[\int_t^{t+h} L(s, X_s, u_s)ds + \bar{Y}(t+h, X_{t+h}) - \bar{Y}(t,x)\right] \\
&= \mathbb{E}_{t,x}\Bigg[\int_t^{t+h} L(s, X_s, u_s)ds + \bar{Y}(t,x) + \int_t^{t+h} \partial_s \bar{Y}(s, X_s)ds \\
&\quad + \sum_{i=1}^{n} \int_t^{t+h} \partial_{x_i} \bar{Y}(s, X_s) b_i(s,x,u)ds + \sum_{i=1}^{n} \int_t^{t+h} \partial_{x_i} \bar{Y}(s, X_s)\sigma_i(s,x,u)dW_s \\
&\quad + \frac{1}{2}\sum_{i,j=1}^{n} \int_t^{t+h} \partial^2_{x_i x_j} \bar{Y}(s, X_s)[\sigma\sigma^*]_{ij}ds - \bar{Y}(t,x)\Bigg].
\end{aligned}$$

Note that the expectation of the term with Brownian motion is equal to zero and $\bar{Y}(t,x)$ cancels each other. Hence,

$$\begin{aligned}
0 &\leq \mathbb{E}_{t,x}\Bigg[\int_t^{t+h} L(s, X_s, u_s)ds + \int_t^{t+h} \partial_s \bar{Y}(s, X_s)ds \\
&\quad + \sum_{i=1}^{n} \int_t^{t+h} \partial_{x_i} \bar{Y}(s, X_s) b_i(s,x,u)ds + \frac{1}{2}\sum_{i,j=1}^{n} \int_t^{t+h} \partial^2_{x_i x_j} \bar{Y}(s, X_s)[\sigma\sigma^*]_{ij}ds\Bigg]
\end{aligned}$$

or, more precisely,

$$0 \le \mathbb{E}_{t,x}\left[\int_t^{t+h} \left(L(s,X_s,u_s) + \partial_s \bar{Y}(s,X_s) + \mathcal{L}^u_{s,x}\bar{Y}(s,X_s)\right) ds\right]$$

where $\mathcal{L}^u_{t,x}$ is an differential operator defined by

$$\mathcal{L}^u_{t,x} = \sum_{i=1}^n b_i(t,x,u)\partial_{x_i} + \frac{1}{2}\sum_{i,j=1}^n [\sigma\sigma^*]_{ij}(t,x,u)\partial^2_{x_i x_j}.$$

Divide both sides by h and take the limit as $h \to 0$, we obtain

$$0 \le L(t,x,u) + \partial_t \bar{Y}(t,X_t) + \mathcal{L}^u \bar{Y}(t,x).$$

As we mention before, equality holds for an optimal control $u = u^0$, which derives the HJB equation:

$$0 = \partial_t \bar{Y}(t,X_t) + \inf_{u\in\mathcal{U}} \left\{L(t,x,u) + \mathcal{L}^u \bar{Y}(t,x)\right\}.$$

It is an easy consequence that $\bar{Y}(T,x) = \Psi(x)$, $\forall x \in \mathbb{R}^n$. □

Example 4.1. The cost functions are supposed to be propotional to the state process, which means $L(t,x,u) = xk(t,u)$ and $\Psi(T,x) = xK(T)$, where the state process X is denoted by H as a one-dimensional process and satisfies the SDE:

$$dH_s = H_s\left(d(s,u_s)ds + n(s,u_s)^* dW_s\right).$$

Then the value function $\bar{Y}(t,x)$ becomes also propotional to x which means $\bar{Y}(t,x) = x\bar{Y}(t) = x\bar{Y}(t,1)$. Indeed,

$$\begin{aligned}\bar{Y}(t,x) &= \inf_{u\in\mathcal{U}} Y(t,x,u) = \inf_{u\in\mathcal{U}} \mathbb{E}_{t,x}\left[\int_t^T xk(s,u)ds + xK(T)\right] \\ &= x \inf_{u\in\mathcal{U}} \mathbb{E}_t\left[\int_t^T k(s,u)ds + K(T)\right] = x\bar{Y}(t,1)\end{aligned}$$

since $L(s,1,u_s) = 1k(s,u)$ and $\Psi(T,1) = 1K(T)$. Moreover HJB equation can be written as

$$\bar{Y}'(t) + \inf_{u\in\mathcal{U}}\left\{k(t,u) + d(t,u_t)\bar{Y}(t)\right\} = 0.$$

We derive the HJB equation for the cost minimization. On the other hand, HJB equation could be also derived for the best utility. In this case, the objective function contains utility functions. We derive the associated HJB equation in the following remark.

Remark 4.1. Consider the same SDE in (4.1). Optimization idea can be expressed as the supremum of admissible control processes. In this case, the objective function is written by

$$V(t,x,u) = \mathbb{E}_{t,x}\left[\int_t^T F(s,X_s,u_s)ds + \phi(T,X_T)\right],$$

where F is utility function and ϕ measures the utility of having some money left at the end of the period. Then the value function appears as

$$\bar{V}(t,x) = \sup_{u \in \mathcal{U}} V(t,x,u).$$

If the same strategical approach is followed, then the HJB equation becomes

$$\frac{\partial \bar{V}}{\partial t}(t,x) + \sup_{u \in \mathcal{U}} \left\{ F(t,x,u) + \mathcal{V}^u_{t,x} \bar{V}(t,x) \right\} = 0,$$

where the differential operator $\mathcal{V}^u_{t,x}$ is the same as $\mathcal{L}^u_{t,x}$:

$$\mathcal{V}^u_{t,x} = \sum_{i=1}^{n} b_i(t,x,u)\partial_{x_i} + \frac{1}{2}\sum_{i,j=1}^{n} [\sigma\sigma^*]_{ij}(t,x,u)\partial^2_{x_i x_j}.$$

We examine HJB equations in the cases of the least cost or the best utility. Besides, we construct stochastic control problems and deal with them using optimal stochastic control theory and backward stochastic differential equations theory.

4.3 Stochastic Control Problems

In this section, we closely follow the work of Quenez [13] in the context of stochastic control problems. A filtered probability space $(\Omega, \mathcal{F}, \mathbb{F} = \{\mathcal{F}_t\}_{0 \le t \le T}, \mathbb{P})$ with usual assumptions models the uncertainity of the controlled state process. The usual assumptions are given by

- $\mathcal{F}_0$ contains all the $\mathbb{P}$-null sets of $\mathbb{F}$.
- $\mathcal{F}_t = \bigcap_{u>t} \mathcal{F}_u$, $\forall t, 0 \le t \le \infty$.
- The filtration $\{\mathcal{F}_t\}_{0 \le t \le T}$ is supposed to be generated by the n-dimensional Brownian Motion W.

The laws of controlled process belong to a family of equivalent measures whose densities are given by

$$dH^u_t = H^u_t \left(d(t,u_t)dt + n(t,u_t)^* dW_t \right), \qquad H^u_t = 1, \tag{4.4}$$

where $d(t,u_t)$ and $n(t,u_t)$ are predictable processes and uniformly bounded by δ_t and υ_t respectively. We could pick a feasible control $u = \{u_t\}_{0 \le t \le T}$ valued in $\mathcal{U}$. Suppose that we have a running cost $k(.,t,u_t)$ associated with the control process u and $K(.,u_T)$ as the terminal condition. The problem is to minimize the following objective function over all feasible control processes u,

$$J(u) = \mathbb{E}\left[\int_0^T H^u_t k(t,u_t)dt + H^u_T K(u_T) \right].$$

Here, the processes $\{k(w,t,u_t)\}_{0\le t\le T}$ (respectively the terminal contitions $\{K(w,u_T)\}_{0\le t\le T})$ are assumed to be measurable with respect to $\mathcal{P}\otimes\mathcal{B}(U)$ (respectively $\mathcal{F}_t\otimes\mathcal{B}(U)$), where $\mathcal{B}(U)$ is the Borelian σ-algebra on U. Moreover, $\{k(w,t,u_t)\}_{0\le t\le T}$ (respectively $\{K(w,u_T)\}_{0\le t\le T})$ is supposed to be uniformly bounded by a square-integrable process $\{k_t\}_{0\le t\le T}$ (respectively by a square-integrable variable χ).
The control processes act magnificently on the discount factor with bounded rate $d(s,u_s)$ in the following manner:

$$dD^u_s = D^u_s d(s,u_s)ds$$

and change of equivalent probability measures with Radon-Nikodym derivatives given by

$$dL^u_s = L^u_s n(s,u_s)^* dW_s.$$

Therefore we obtain $H^u_t = D^u_t L^u_t$, indeed

$$\begin{aligned}
d(D^u_t L^u_t) &= L^u_t dD^u_t + D^u_t dL^u_t + d < D^u_t, L^u_t > \\
&= L^u_t D^u_s d(s,u_s)ds + D^u_t L^u_s n(s,u_s)^* dW_s; \quad d < D^u_t, L^u_t >= 0 \\
&\text{and if } \quad D^u_t L^u_t = H^u_t \quad \text{then} \\
dH^u_t &= H^u_t \left(d(t,u_t)dt + n(t,u_t)^* dW_t\right)
\end{aligned}$$

which is exactly the same SDE as the equation (4.4).
The objective function could be written in terms of new probability measure $\mathbb{Q}^u$ with density L^u_T on $\mathcal{F}_T$ as follows:

$$J(u) = \mathbb{E}_{\mathbb{Q}^u}\left[\int_0^T D^u_t k(t,u_t)dt + D^u_T K(u_T)\right].$$

One can easily show that $J(u) = Y^u_0$, where (Y^u, Z^u) is the solution of the linear BSDE associated with the standard parameter (f^u, ξ^u) defined as

$$f^u(t,y,z) = k(t,u_t) + d(t,u_t)y + n(t,u_t)^* z; \qquad \xi^u = K(u_T).$$

Indeed, if f^u is a linear standard generator, then $H^u_{t,s}$ corresponds to the adjoint process beginning at time t (assumed to be equal to 1 at the beginning time t) and the solution Y^u_t could be written as

$$Y^u_t = \mathbb{E}\left[\xi^u H^u_{t,T} + \int_t^T H^u_{t,s} k(s,u_s)ds \mid \mathcal{F}_t\right]; \qquad 0\le t\le s\le T.$$

For $t=0$,

$$\begin{aligned}
Y^u_0 &= \mathbb{E}\left[\xi^u H^u_{0,T} + \int_0^T H^u_{0,s} k(s,u_s)ds \mid \mathcal{F}_0\right] \\
&= \mathbb{E}\left[H^u_{0,T} K(u_T) + \int_0^T H^u_{0,s} k(s,u_s)ds\right] = J(u).
\end{aligned}$$

For any control u, Y_t^u becomes objective function at time t as above. The main idea is to minimize the objective function over the feasible control processes. Let $\bar{Y}_t$ be the value function at time t, we obtain

$$\bar{Y}_t = \operatorname*{ess\,inf}_{u\in\mathcal{U}} Y_t^u, \qquad t \in [0,T].$$

If there exists an optimal control $u^0 \in \mathcal{U}$, it achieves the equality that

$$\bar{Y}_t = Y_t^{u^0}, \qquad t \in [0,T].$$

Following the same procedure in the previous section, the Dynamic Programming Principle becomes

$$\bar{Y}_t = \operatorname*{ess\,inf}_{u\in\mathcal{U}} \mathbb{E}\left[\int_t^{t+h} H_{t,s}^u k(s,u_s)ds + H_{t,t+h}^u \bar{Y}_{t+h} \mid \mathcal{F}_t\right]; \qquad 0 \le t \le t+h \le T.$$

In Markovian case, it could be written that

$$0 \le \mathbb{E}\left[\int_t^{t+h} H_{t,s}^u k(s,u_s)ds + H_{t,t+h}^u \bar{Y}_{t+h} - \bar{Y}_t \mid \mathcal{F}_t\right]. \tag{4.5}$$

We aggregate backward stochastic differential equation theory and optimal stochastic control theory in the following proposition.

Proposition 4.2. *Let $\bar{Y}$ be a value function such that*

$$\bar{Y}_t = \operatorname*{ess\,inf}_{u\in\mathcal{U}} Y_t^u = \operatorname*{ess\,inf}_{u\in\mathcal{U}} \mathbb{E}\left[\xi^u H_{t,T}^u + \int_t^T H_{t,s}^u k(s,u_s)ds \mid \mathcal{F}_t\right]; \qquad 0 \le t \le s \le T$$

with a deflator H given by

$$dH_t^u = H_t^u\left(d(t,u_t)dt + n(t,u_t)^* dW_t\right), \qquad H_t^u = 1.$$

Then $(\bar{Y}, Z)$ be the solution of the following BSDE

$$-d\bar{Y}_t = f(t,\bar{Y}_t,Z_t)dt - Z_t^* dW_t; \qquad \bar{Y}_T = \xi$$

and corresponds to the standard parameter (f,ξ):

$$f(t,y,z) = \operatorname*{ess\,inf}_{u\in\mathcal{U}} f^u(t,y,z), \qquad \xi = \operatorname*{ess\,inf}_{u\in\mathcal{U}} \xi^u,$$

where $f^u(t,y,z) = k(t,u_t) + d(t,u_t)y + n(t,u_t)^* z$ and $\xi^u = K(u_T)$.

Proof. Assume that $\bar{Y}$ is a semi-martingale of the form

$$d\bar{Y}_t = -f(t)dt + Z_t^* dW_t.$$

Applying Itô product rule to $H_{t,s}^u \bar{Y}_s$ and integrating from t to $t+h$, we obtain

$$\begin{aligned} H_{t,t+h}^u \bar{Y}_{t+h} &= H_{t,t}^u \bar{Y}_t + \int_t^{t+h} \bar{Y}_s dH_{t,s}^u + \int_t^{t+h} H_{t,s}^u d\bar{Y}_s + \int_t^{t+h} d < H_{t,s}^u, \bar{Y}_s > \\ &= H_{t,t}^u \bar{Y}_t + \int_t^{t+h} \bar{Y}_s H_{t,s}^u \left(d(s,u_s)ds + n(s,u_s)^* dW_s\right) \\ &\quad + \int_t^{t+h} H_{t,s}^u \left(-f(s)ds + Z_s^* dW_s\right) + \int_t^{t+h} H_{t,s}^u n(s,u_s)^* Z_s ds. \end{aligned}$$

Plugging this equality into (4.5), we have

$$
\begin{aligned}
0 &\leq \mathbb{E}\left[\int_t^{t+h} H^u_{t,s} k(s,u_s)ds + H^u_{t,t+h}\bar{Y}_{t+h} - \bar{Y}_t \mid \mathcal{F}_t\right] \\
&= \mathbb{E}\Bigg[\int_t^{t+h} H^u_{t,s} k(s,u_s)ds + H^u_{t,t}\bar{Y}_t + \int_t^{t+h} \bar{Y}_s H^u_{t,s}\left(d(s,u_s)ds + n(s,u_s)^* dW_s\right) \\
&\quad + \int_t^{t+h} H^u_{t,s}\left(-f(s)ds + Z_s^* dW_s\right) + \int_t^{t+h} H^u_{t,s} n(s,u_s)^* Z_s ds - \bar{Y}_t \mid \mathcal{F}_t\Bigg].
\end{aligned}
$$

Conditional expectations of the terms with Brownian Motion is equal to zero. Thus,

$$
\begin{aligned}
0 &\leq \mathbb{E}\Bigg[\int_t^{t+h} H^u_{t,s} k(s,u_s)ds + \underbrace{H^u_{t,t}}_{=1}\bar{Y}_t + \int_t^{t+h} \bar{Y}_s H^u_{t,s} d(s,u_s)ds \\
&\quad - \int_t^{t+h} H^u_{t,s} f(s)ds + \int_t^{t+h} H^u_{t,s} n(s,u_s)^* Z_s ds - \bar{Y}_t \mid \mathcal{F}_t\Bigg] \\
&= \mathbb{E}\Bigg[\int_t^{t+h} H^u_{t,s} k(s,u_s)ds + \int_t^{t+h} \bar{Y}_s H^u_{t,s} d(s,u_s)ds - \int_t^{t+h} H^u_{t,s} f(s)ds \\
&\quad + \int_t^{t+h} H^u_{t,s} n(s,u_s)^* Z_s ds \mid \mathcal{F}_t\Bigg] \\
&= \mathbb{E}\left[\int_t^{t+h} H^u_{t,s}\left(k(s,u_s) + d(s,u_s)\bar{Y}_s + n(s,u_s)^* Z_s - f(s)\right)ds \mid \mathcal{F}_t\right].
\end{aligned}
$$

Divide the inequality by h and take the limit as $h \to 0$, $H^u_{t,s}$ tends to $H^u_{t,t} = 1$ and we finally obtain $dt \otimes d\mathbb{P}$-a.s,

$$0 \leq k(t,u_t) + d(t,u_t)\bar{Y}_t + n(t,u_t)^* Z_t - f(t).$$

As we mention before, equality holds for an optimal control $u^0 \in \mathcal{U}$, i.e,

$$
\begin{aligned}
& 0 = k(t,u^0_t) + d(t,u^0_t)\bar{Y}_t + n(t,u^0_t)^* Z_t - f(t) \\
\Rightarrow \quad & f(t) = k(t,u^0_t) + d(t,u^0_t)\bar{Y}_t + n(t,u^0_t)^* Z_t \\
\Rightarrow \quad & f(t) = \operatorname{ess}\inf_{u\in\mathcal{U}} f^u(t,\bar{Y}_t,Z_t).
\end{aligned}
$$

Therefore, $(\bar{Y}, Z)$ is the solution of the following BSDE

$$-d\bar{Y}_t = f(t,\bar{Y}_t,Z_t)dt - Z_t^* dW_t, \qquad \bar{Y}_T = \xi$$

and corresponds to the standard parameter (f,ξ):

$$f(t,y,z) = \operatorname{ess}\inf_{u\in\mathcal{U}} f^u(t,y,z), \qquad \xi = \operatorname{ess}\inf_{u\in\mathcal{U}} \xi^u.$$

□

We extend the previous proposition to the BSDE with consumption.

Proposition 4.3. *Assume that* (Y, Z) *is the solution of the BSDE associated with the standard parameter* (f, ξ)*, where*

$$f(t, y, z) = \operatorname{ess}\inf_{u\in\mathcal{U}} f^u(t, y, z), \qquad \xi = \operatorname{ess}\inf_{u\in\mathcal{U}} \xi^u.$$

Then, there exists a predictable increasing consumption process C^u *such that*

$$C_t^u = \int_0^t (f^u(s, Y_s, Z_s) - f(s, Y_s, Z_s))\, ds + \mathbf{1}_{t=T}(\xi^u - \xi)$$

and (Y, Z, C^u) *is the subsolution of the BSDE associated with the standard parameter* (f^u, ξ^u)*, that is,*

$$-dY_t = f^u(t, Y_t, Z_t)dt - dC_t^u - Z_t^* dW_t, \qquad Y_T = \xi^u.$$

Moreover, $H_t^u Y_t + \int_0^t H_s^u k(s, u_s)ds$ *is a uniformly integrable submartingale, for all* $u \in \mathcal{U}$*.*

Proof. For a given admissible control process $u \in \mathcal{U}$, C^u is obtained in the following manner: If (Y, Z) is solution of the BSDE associated with standard parameter (f, ξ), then

$$-dY_t = f(t, Y_t, Z_t)dt - Z_t^* dW_t, \qquad Y_T = \xi.$$

Add and substract the term $f^u(t, Y_t, Z_t)dt$, we get

$$\begin{aligned} -dY_t &= f(t, Y_t, Z_t)dt - Z_t^* dW_t \pm f^u(t, Y_t, Z_t)dt \\ &= f^u(t, Y_t, Z_t)dt + (f(t, Y_t, Z_t) - f^u(t, Y_t, Z_t))\, dt - Z_t^* dW_t. \end{aligned}$$

Now, we integrate from 0 to t:

$$-Y_t + Y_0 = \int_0^t f^u(s, Y_s, Z_s)ds + \int_0^t (f(s, Y_s, Z_s) - f^u(s, Y_s, Z_s))\, ds - \int_0^t Z_s^* dW_s.$$

The equation above holds for all $t \in [0, T]$. In particular, if $t = T$, then the equation becomes

$$\begin{aligned} Y_0 &= \int_0^T f^u(s, Y_s, Z_s)ds + \int_0^T (f(s, Y_s, Z_s) - f^u(s, Y_s, Z_s))\, ds - \int_0^T Z_s^* dW_s + \xi \pm \xi^u \\ &= \int_0^T f^u(s, Y_s, Z_s)ds + \int_0^T (f(s, Y_s, Z_s) - f^u(s, Y_s, Z_s))\, ds + (\xi - \xi^u) + \xi^u \\ &\quad - \int_0^T Z_s^* dW_s. \end{aligned}$$

Putting $C_t^u = \int_0^t (f^u(s, Y_s, Z_s) - f(s, Y_s, Z_s))\, ds + \mathbf{1}_{t=T}(\xi^u - \xi)$, then (Y, Z, C^u) becomes a subsolution of the BSDE

$$-dY_t = f^u(t, Y_t, Z_t)dt - dC_t^u - Z_t^* dW_t, \qquad Y_T = \xi^u$$

associated with the standard parameter (f^u, ξ^u).
C^u is a predictable and increasing process: Predictability is obvious since all the components of C^u are predictable. Moreover, $f(t, Y_t, Z_t)$ (respectively ξ) is defined as essinf of $f^u(t, Y_t, Z_t)$ (respectively ξ^u) over the feasible control processes. Therefore $f^u(t, Y_t, Z_t) - f(t, Y_t, Z_t) \geq 0$ (respectively, $\xi^u - \xi \geq 0$) which implies that C^u is increasing.
Furthermore, $\forall u \in \mathcal{U}$, $H_t^u Y_t + \int_0^t H_s^u k(s, u_s) ds$ is a uniformly integrable submartingale. Indeed, apply Itô product rule to $H_t^u Y_t$ and integrate from 0 to T:

$$\begin{aligned}
d(H_t^u Y_t) &= H_t^u dY_t + Y_t dH_t^u + < dH_t^u, dY_t > \\
&= H_t^u \left(-f^u(t, Y_t, Z_t)dt + dC_t^u + Z_t^* dW_t\right) + Y_t H_t^u \left(d(t, u_t)dt + n(t, u_t)^* dW_t\right) \\
&\quad + H_t^u n(t, u_t) Z_t dt \\
&= H_t^u \left(-\left(k(t, u_t) + d(t, u_t)Y_t + n(t, u_t)^* Z_t\right) dt + dC_t^u + Z_t^* dW_t\right) \\
&\quad + Y_t H_t^u \left(d(t, u_t)dt + n(t, u_t)^* dW_t\right) + H_t^u n(t, u_t) Z_t dt \\
&= H_t^u Y_t n(t, u_t)^* dW_t - H_t^u k(t, u_t)dt + H_t^u Z_t^* dW_t + H_t^u dC_t \\
H_T^u Y_T &= H_0^u Y_0 + \int_0^T H_t^u \left(Y_t n(t, u_t) + Z_t\right)^* dW_t - \int_0^T H_t^u k(t, u_t)dt + \int_0^T H_t^u dC_t.
\end{aligned}$$

Take the conditional expectation with respect to $\mathcal{F}_0$ of both sides, we have

$$\begin{aligned}
\mathbb{E}\left[H_T^u Y_T + \int_0^T H_t^u k(t, u_t)dt \mid \mathcal{F}_0\right] &= Y_0 + \mathbb{E}\left[\int_0^T H_t^u \left(Y_t n(t, u_t) + Z_t\right)^* dW_t \mid \mathcal{F}_0\right] \\
&\quad + \mathbb{E}\left[\int_0^T H_t^u dC_t \mid \mathcal{F}_0\right] \\
&= Y_0 + \mathbb{E}\left[\int_0^T H_t^u dC_t\right] \\
\mathbb{E}\left[H_T^u Y_T + \int_0^T H_t^u k(t, u_t)dt \mid \mathcal{F}_0\right] &\geq Y_0 = \mathbb{E}\left[H_0^u Y_0 + \int_0^0 H_t^u k(t, u_t)dt\right]
\end{aligned}$$

which implies that $H_t^u Y_t + \int_0^t H_s^u k(s, u_s) ds$ is a uniformly integrable submartingale, $\forall u \in \mathcal{U}$. □

Let us think on the opposite side. If a standard generator f and a terminal condition ξ are given by essential infimum of some standard generators and terminal conditions, respectively, then we examine the solution of BSDE associated with the standard parameter (f, ξ) in the sense of optimality.

4.4 The Verification Theorem

We follow the work of Quenez [13] with assuming that (f, ξ) is a standard parameter. In this section, we show that f (respectively ξ) can be obtained as an essential infimum of standard generators f^α (respectively of terminal conditions ξ^α). For a given α, let (Y^α, Z^α) be the solution of the BSDE associated with the standard parameter (f^α, ξ^α). Then by the Comparison Theorem 2.5, the solution of BSDE associated with the standard parameter (f, ξ) is less than or equal to (Y^α, Z^α). The following proposition shows how the equality holds.

Proposition 4.4. *Let (f, ξ) and (f^α, ξ^α) be a family of standard parameters, and (Y, Z) and (Y^α, Z^α) be the solutions of BSDE's associated with given standard parameters, respectively. Suppose that there exists a parameter $\bar{\alpha}$ such that*

$$f(t, Y_t, Z_t) = \operatorname*{ess\,inf}_\alpha f^\alpha(t, Y_t, Z_t) = f^{\bar{\alpha}}(t, Y_t, Z_t), \qquad d\mathbb{P} \otimes dt \ \ a.s.$$
$$\xi = \operatorname*{ess\,inf}_\alpha \xi^\alpha = \xi^{\bar{\alpha}} \qquad d\mathbb{P} \ \ a.s.. \tag{4.6}$$

Then, the process Y and Y^α satisfy:

$$Y_t = \operatorname*{ess\,inf}_\alpha Y_t^\alpha = Y_t^{\bar{\alpha}}, \forall t \in [0, T], d\mathbb{P} \ \ a.s.. \tag{4.7}$$

Proof. By hypothesis, one can easily write the following inequalities:

$$f(t, Y_t, Z_t) \leq f^\alpha(t, Y_t, Z_t), \qquad d\mathbb{P} \otimes dt \text{ a.s, and}$$
$$\xi \leq \xi^\alpha, \qquad d\mathbb{P} \text{ a.s..}$$

By the first part of the comparison theorem (Theorem 2.5),

$$Y_t \leq Y_t^\alpha, \qquad \forall t \in [0, T].$$

It implies that Y is a lower bound for the family $\{Y^\alpha\}$. By definition of essential infimum,

$$Y_t \leq \operatorname*{ess\,inf}_\alpha Y_t^\alpha, \qquad \forall t \in [0, T], \qquad d\mathbb{P} \text{ a.s.}$$

Let $Y^{\bar{\alpha}}$ be the solution of BSDE associated with $(f^{\bar{\alpha}}, \xi^{\bar{\alpha}})$. From the hypothesis equalities

$$f(t, Y_t, Z_t) = f^{\bar{\alpha}}(t, Y_t, Z_t)$$
$$\xi = \xi^{\bar{\alpha}}$$

and the uniqueness of the solution, we have $Y_t = Y_t^{\bar{\alpha}}, \forall t \in [0, T]$. Hence, it is obtained that

$$\operatorname*{ess\,inf}_\alpha Y_t^\alpha \geq Y_t = Y_t^{\bar{\alpha}} \geq \operatorname*{ess\,inf}_\alpha Y_t^\alpha$$

which exactly shows what we claim:

$$Y_t = \operatorname*{ess\,inf}_\alpha Y_t^\alpha = Y_t^{\bar{\alpha}}, \qquad \forall t \in [0, T], \qquad d\mathbb{P} \text{ a.s..}$$

□

We have a related corollary that will be used in the proof of the verification theorem (see [13] for further details).

Corollary 4.5. *Assume that the standard generators f^α are equi-Lipschitz with the same constant C and for each $\varepsilon > 0$ the following inequalities hold:*

$$f(t, Y_t, Z_t) = \operatorname*{ess\,inf}_\alpha f^\alpha(t, Y_t, Z_t) \geq f^{\alpha^\varepsilon}(t, Y_t, Z_t) - \varepsilon, \qquad d\mathbb{P} \otimes dt \ a.s.;$$
$$\xi = \operatorname*{ess\,inf}_\alpha \xi^\alpha \geq \xi^{\alpha^\varepsilon} - \varepsilon \qquad d\mathbb{P} \ a.s..$$

Then,

$$Y_t = \operatorname*{ess\,inf}_\alpha Y_t^\alpha \geq Y_t^{\alpha^\varepsilon} - \varepsilon, \qquad \forall t \in [0, T], d\mathbb{P} \ a.s..$$

Proof. Define $\delta f_t^\varepsilon = f(t, Y_t, Z_t) - f^{\alpha^\varepsilon}(t, Y_t, Z_t)$, $\delta Y_t = Y_t - Y_t^{\alpha^\varepsilon}$ and $\delta Z_t = Z_t - Z_t^{\alpha^\varepsilon}$ and derive $d\delta Y_t$:

$$\begin{aligned}
-d\delta Y_t &= -d(Y_t - Y_t^{\alpha^\varepsilon}) = -dY_t + dY_t^{\alpha^\varepsilon} \\
&= f(t, Y_t, Z_t)dt - Z_t^* dW_t - f^{\alpha^\varepsilon}(t, Y_t^{\alpha^\varepsilon}, Z_t^{\alpha^\varepsilon})dt + Z_t^{\alpha^\varepsilon *} dW_t \\
&\quad \pm f^{\alpha^\varepsilon}(t, Y_t, Z_t)dt \\
&= \left[f^{\alpha^\varepsilon}(t, Y_t, Z_t) - f^{\alpha^\varepsilon}(t, Y_t^{\alpha^\varepsilon}, Z_t^{\alpha^\varepsilon})\right] dt - \delta Z_t^* dW_t + \delta f_t^\varepsilon dt \\
&\quad \pm f^{\alpha^\varepsilon}(t, Y_t^{\alpha^\varepsilon}, Z_t) \\
&= \left[\Delta_y f(t) \delta Y_t + \Delta_z f(t) \delta Z_t + \delta f_t^\varepsilon\right] dt - \delta Z_t^* dW_t,
\end{aligned}$$

where

$$\begin{aligned}
\Delta_y f(t) &= \frac{f^{\alpha^\varepsilon}(t, Y_t, Z_t) - f^{\alpha^\varepsilon}(t, Y_t^{\alpha^\varepsilon}, Z_t)}{Y_t - Y_t^{\alpha^\varepsilon}}, \\
\Delta_z f(t) &= \frac{f^{\alpha^\varepsilon}(t, Y_t^{\alpha^\varepsilon}, Z_t) - f^{\alpha^\varepsilon}(t, Y_t^{\alpha^\varepsilon}, Z_t^{\alpha^\varepsilon})}{Z_t - Z_t^{\alpha^\varepsilon}}.
\end{aligned}$$

Here, $\Delta_y f(t)$ and $\Delta_z f(t)$ are predictable processes and it remains to show that they are bounded by the Lipschitz coefficient C of f:

$$\begin{aligned}
|\Delta_y f(t)| &= \left| \frac{f^{\alpha^\varepsilon}(t, Y_t, Z_t) - f^{\alpha^\varepsilon}(t, Y_t^{\alpha^\varepsilon}, Z_t)}{Y_t - Y_t^{\alpha^\varepsilon}} \right| \\
&\leq \frac{1}{\delta Y_t} C \left(|Y_t - Y_t^{\alpha^\varepsilon}| + |Z_t - Z_t|\right) = C.
\end{aligned}$$

Similarly, $\Delta_z f(t)$ is bounded by C of f. Therefore, $(\delta Y, \delta Z)$ becomes the solution of the following LBSDE:

$$\begin{aligned}
-d\delta Y_t &= \left(\Delta_y f(t) \delta Y_t + \Delta_z f(t) \delta Z_t + \delta f_t^\varepsilon\right) dt - \delta Z_t^* dW_t, \\
\delta Y_T &= \xi - \xi^{\alpha^\varepsilon}.
\end{aligned} \tag{4.8}$$

The standard generator of the LBSDE appears as

$$f(t, \delta Y_t, \delta Z_t) = \Delta_y f(t) \delta Y_t + \Delta_z f(t) \delta Z_t + \delta f_t^\varepsilon.$$

We can write the associated adjoint process Γ beginning at time t in the following manner:

$$d\Gamma_s^t = \Gamma_s^t \left(\Delta_y f(s) ds + \Delta_z f(s) dW_s\right), \qquad \Gamma_t^t = 1,$$

and the solution δY can be written as

$$\delta Y_t = \mathbb{E}\left[\delta Y_T \Gamma_T^t + \int_t^T \Gamma_s^t \delta f_s^\varepsilon ds \mid \mathcal{F}_t\right]. \tag{4.9}$$

The solution of Γ process is

$$\Gamma_T^t = \exp\left(\int_t^T \left(\Delta_y f(s) - \frac{1}{2}\Delta_z f(s)^2\right) ds + \int_t^T \Delta_z f(s) dW_s\right).$$

We then calculate the expectations of Γ_T^t and $\int_t^T \Gamma_s^t ds$:

$$\begin{aligned}
\mathbb{E}\left[\Gamma_T^t\right] &= \mathbb{E}\left[e^{\int_t^T \left(\Delta_y f(s) - \frac{1}{2}\Delta_z f(s)^2\right)ds + \int_t^T \Delta_z f(s) dW_s}\right] \\
&= e^{\int_t^T \left(\Delta_y f(s) - \frac{1}{2}\Delta_z f(s)^2\right)ds}\mathbb{E}\left[e^{\int_t^T \Delta_z f(s) dW_s}\right] \\
&= e^{\int_t^T \left(\Delta_y f(s) - \frac{1}{2}\Delta_z f(s)^2\right)ds} e^{\frac{1}{2}\int_t^T \Delta_z f(s)^2 ds} \\
&= e^{\int_t^T (\Delta_y f(s) ds)} \\
&\leq e^{\int_0^T (\Delta_y f(s) ds)} \leq e^{CT}.
\end{aligned}$$

For any $\varepsilon > 0$,

$$-\varepsilon\mathbb{E}\left[\Gamma_T^t\right] \geq -\varepsilon e^{CT}. \tag{4.10}$$

Since the fact that Γ is a positive and increasing process, we can write

$$\mathbb{E}\left[\int_t^T \Gamma_s^t ds\right] \leq \mathbb{E}\left[\int_0^T \Gamma_s^t ds\right] \leq \mathbb{E}\left[\int_0^T \Gamma_T^t ds\right] \leq \mathbb{E}\left[Te^{CT}\right] = Te^{CT}.$$

For any $\varepsilon > 0$,

$$-\varepsilon\mathbb{E}\left[\int_t^T \Gamma_s^t ds\right] \geq -\varepsilon T e^{CT}. \tag{4.11}$$

Then equation (4.9) turns out to be

$$\begin{aligned}
\delta Y_t &= \mathbb{E}\left[\delta Y_T \Gamma_T^t + \int_t^T \Gamma_s^t \delta f_s^\varepsilon ds \mid \mathcal{F}_t\right] \\
&= \mathbb{E}\left[\underbrace{\delta Y_T}_{\geq\varepsilon} \Gamma_T^t + \int_t^T \Gamma_s^t \underbrace{\delta f_s^\varepsilon}_{\geq\varepsilon} ds\right] \qquad \text{(since } \mathcal{F}_t \text{ measurability)} \\
&\geq -\varepsilon\mathbb{E}\left[\int_t^T \Gamma_s^t ds + \Gamma_T^t\right] \qquad \text{by (4.10) - (4.11)} \\
&\geq -\varepsilon e^{CT} - \varepsilon T e^{CT} = -\varepsilon(T+1)e^{CT} = -\varepsilon_1.
\end{aligned}$$

Since $\delta Y_t = Y_t - Y_t^{\alpha^\varepsilon}$, we have shown that

$$Y_t \geq Y_t^{\alpha^\varepsilon} - \varepsilon_1, \qquad \forall t \in [0,T], \qquad d\mathbb{P} \text{ a.s.}.$$

□

Recall that for a given control process $u \in U$, $d(t, u_t)$, $n(t, u_t)$, $k(\cdot, t, u_t)$ and $K(\cdot, u_T)$ is bounded by δ_t, υ_t, k_t and χ respectively. Let δ, υ, k, χ be bounded processes. The verification theorem leads a sufficent condition for a process to be the value function. Before that, we must prove the following lemma which is stated in [13].

Lemma 4.6. *For each $\varepsilon > 0$, there exists a feasible control u^ε such that*

$$f(t, Y_t, Z_t) = \operatorname{ess}\inf_{u \in \mathcal{U}} f^u(t, Y_t, Z_t) \geq f^{u^\varepsilon}(t, Y_t, Z_t) - \varepsilon, \qquad d\mathbb{P} \otimes dt \ \ a.s.,$$
$$\xi = \operatorname{ess}\inf_{u \in \mathcal{U}} \xi^u \geq \xi^{u^\varepsilon} - \varepsilon \qquad d\mathbb{P} \ \ a.s.. \tag{4.12}$$

Proof. For each $(\omega, t) \in \Omega \otimes [0, T]$, the sets defined by $\{u \in U | \xi(\omega) \geq K(\omega, u) - \varepsilon\}$ and $\{u \in U | f(t, Y_t(\omega), Z_t(\omega)) \geq k(t, \omega, u) + d(t, \omega, u) Y_t(\omega) + n(t, \omega, u) Z_t(\omega) - \varepsilon\}$are non-empty, since U is a Polish space. Here, Y and Z are predictable processes and k, d, n, K are measurable. Therefore, by the measurable selection theorem in [27] (see Appendix), there exists a predictable process $u^\varepsilon \in U$ such that the inequalities in (4.12) holds. Control process u^ε becomes an admissible control process, since $(Y^{u^\varepsilon}, Z^{u^\varepsilon})$ solves the BSDE associated with the standard parameter $(f^{u^\varepsilon}, \xi^{u^\varepsilon})$. □

Now, we are ready to prove the verification theorem. The theorem is stated in [13], but we offer a different way to prove.

Theorem 4.7 (Verification Theorem)**.** *Assume that (f^u, ξ^u) are standard parameters. The parameter (f, ξ) defined by*

$$f(t, y, z) = \operatorname{ess}\inf_{u \in \mathcal{U}} f^u(t, y, z) \qquad \xi = \operatorname{ess}\inf_{u \in \mathcal{U}} \xi^u$$

becomes standard parameter, as well. Let (Y, Z) be the solution of the BSDE associated with (f, ξ), then Y is the value function $\bar{Y}$ of the control problem, that is, $\forall t \in [0, T]$,

$$Y_t = \bar{Y}_t = \operatorname{ess}\inf_{u \in \mathcal{U}} Y_t^u.$$

Proof. By hypothesis, $f(t, y, z) = \operatorname{ess}\inf_{u \in \mathcal{U}} f^u(t, y, z)$, where

$$f^u(t, y, z) = k(t, u_t) + d(t, u_t) y + n(t, u_t)^* z$$

are linear standard generators. The claim is that f is also a standard generator; measurability, to be an element of $\mathbb{H}_T^2(\mathbb{R}^d)$ and Lipschitz condition are satisfied. First of all it must be shown that f is measurable. For given (ω, t), consider a countable dense family$\{(y_n, z_n)\}$. For each n, define a measurable process $f(\omega, t, y_n, z_n)$ and $d\mathbb{P} \otimes dt$-null set N such that for $(\omega, t) \in N^c$,

$$f(\omega, t, y_n, z_n) = \inf_{u \in \mathcal{U}} \{k(\omega, t, u) + d(\omega, t, u) y_n + n(\omega, t, u)^* z_n\}.$$

For $(\omega, t) \in N^c$, $\{f(\omega, t, y_n, z_n)\}$ is a Cauchy sequence. Indeed, for a given control process u, define a function $g(\omega, t, y_n, z_n) := k(\omega, t, u) + d(\omega, t, u)y_n + n(\omega, t, u)^* z_n$. For $(\omega, t) \in N^c$ and m, n big enough,

$$\begin{aligned} |g(\omega, t, y_n, z_n) - g(\omega, t, y_n, z_n)| &= |k(\omega, t, u) + d(\omega, t, u)y_n + n(\omega, t, u)^* z_n \\ &\quad - k(\omega, t, u) - d(\omega, t, u)y_m - n(\omega, t, u)^* z_m| \\ &\le |d(\omega, t, u)||y_n - y_m| + \|n(\omega, t, u)\|\|z_n - z_m\| < \varepsilon, \end{aligned}$$

since d and n are assumed to be bounded. We have shown that g is Cauchy, for the pair $(\omega, t) \in N^c$. Therefore, $f(\omega, t, y_n, z_n) = \inf_{u \in \mathcal{U}} g(\omega, t, y_n, z_n)$ is Cauchy, for the pair $(\omega, t) \in N^c$. $\mathbb{R} \otimes \mathbb{R}^n$ is complete space, so we can define $f(\omega, t, y, z)$ as the limit of the Cauchy sequence $f(\omega, t, y_n, z_n)$ as $(y_n, z_n) \to (y, z)$. Since the fact that limit of measurable functions are also measurable, f is measurable. Secondly, define a set in the following manner:

$$\bar{\mathcal{U}} := \left\{ \bar{u} \in \mathcal{U} \mid \mathbb{P}\{f^u(t, y, z) < f^{\bar{u}}(t, y, z)\} = 0, \forall u \in \mathcal{U}, \forall (y, z) \in \mathbb{R} \otimes \mathbb{R}^n, \forall t \in [0, T] \right\}.$$

Then, we can write

$$f(t, y, z) = \operatorname*{ess\,inf}_{u \in \mathcal{U}} f^u(t, y, z) = \sup_{\bar{u} \in \bar{\mathcal{U}}} f^{\bar{u}}(t, y, z).$$

We have to show the remainder two properties of f in order to be a standard generator:

1. Is $f(\cdot, 0, 0)$ a member of $\mathbb{H}_T^2(\mathbb{R}^d)$?

$$\begin{aligned} \mathbb{E}\left[\int_0^T |f(t, 0, 0)|^2 dt\right] &= \mathbb{E}\left[\int_0^T |\sup_{\bar{u} \in \bar{\mathcal{U}}} f^{\bar{u}}(t, 0, 0)|^2 dt\right] \\ &= \mathbb{E}\left[\int_0^T |\sup_{\bar{u} \in \bar{\mathcal{U}}} \{k(t, u_t)\}|^2 dt\right] < \infty \end{aligned}$$

since k is bounded by assumption.

2. Is f Lipschitz? For any $(y_1, z_1), (y_2, z_2) \in \mathbb{R} \otimes \mathbb{R}^n$

$$\begin{aligned} |f(t, y_1, z_1) - f(t, y_2, z_2)| &= |\sup_{\bar{u} \in \bar{\mathcal{U}}} f^{\bar{u}}(t, y_1, z_1) - \sup_{\bar{u} \in \bar{\mathcal{U}}} f^{\bar{u}}(t, y_2, z_2)| \\ &\le |\sup_{\bar{u} \in \bar{\mathcal{U}}} \{f^{\bar{u}}(t, y_1, z_1) - f^{\bar{u}}(t, y_2, z_2)\}| \\ &= |\sup_{\bar{u} \in \bar{\mathcal{U}}} \{k(t, \bar{u}_t) + d(t, \bar{u}_t)y_1 + n(t, \bar{u}_t)^* z_1 \\ &\qquad - k(t, \bar{u}_t) - d(t, \bar{u}_t)y_2 - n(t, \bar{u}_t)^* z_2\}| \\ &= |\sup_{\bar{u} \in \bar{\mathcal{U}}} \{\underbrace{d(t, \bar{u}_t)}_{\le \delta}(y_1 - y_2) + \underbrace{n(t, \bar{u}_t)^*}_{\le \upsilon}(z_1 - z_2)\}| \\ &\le \max\{\delta, \upsilon\} \left(|y_1 - y_2| + |z_1 - z_2|\right). \end{aligned}$$

$\max\{\delta, \upsilon\}$ becomes Lipschitz coefficient of f.

ξ must be a standard parameter as a terminal condition. Again, by hypothesis,

$$\xi = \operatorname{ess}\inf_{u\in\mathcal{U}} \xi^u$$

Define a set $\tilde{\mathcal{U}}$ in a similar way,

$$\tilde{\mathcal{U}} := \left\{ \tilde{u} \in \mathcal{U} \,|\, \mathbb{P}\{\xi^u < \xi^{\tilde{u}}\} = 0, \forall u \in \mathcal{U} \right\}.$$

Then

$$\xi = \operatorname{ess}\inf_{u\in\mathcal{U}} \xi^u = \sup_{\tilde{u}\in\tilde{\mathcal{U}}} \xi^{\tilde{u}}.$$

Calculate $\mathbb{E}\left(|\xi|^2\right)$:

$$\mathbb{E}\left(|\xi|^2\right) = \mathbb{E}\left(|\sup_{\tilde{u}\in\tilde{\mathcal{U}}} \xi^{\tilde{u}}|^2 \right) \leq \mathbb{E}\left(\sup_{\tilde{u}\in\tilde{\mathcal{U}}} |\xi^{\tilde{u}}|^2 \right) \leq \sup_{\tilde{u}\in\tilde{\mathcal{U}}} \mathbb{E}\left(|\xi^{\tilde{u}}|^2\right) < \infty,$$

since $\xi^{\tilde{u}}$ is standard parameter as terminal condition for $\tilde{u} \in \mathcal{U}$ satisfying $\mathbb{E}\left(|\xi^{\tilde{u}}|^2\right) < \infty$. From Lemma 4.6, for each $\varepsilon > 0$, there exists a feasible control u^ε such that

$$f(t, Y_t, Z_t) = \operatorname{ess}\inf_{u\in\mathcal{U}} f^u(t, Y_t, Z_t) \geq f^{u^\varepsilon}(t, Y_t, Z_t) - \varepsilon, \qquad d\mathbb{P} \otimes dt \text{ a.s.},$$

$$\xi = \operatorname{ess}\inf_{u\in\mathcal{U}} \xi^u \geq \xi^{u^\varepsilon} - \varepsilon \qquad d\mathbb{P} \text{ a.s..}$$

It implies that, by the Corollary 4.5,

$$Y_t = \operatorname{ess}\inf_{u\in\mathcal{U}} Y_t^u \geq Y_t^{u^\varepsilon} - \varepsilon, \qquad \forall t \in [0, T], d\mathbb{P} \text{ a.s..}$$

By definition of essential infimum, $Y^{u^\varepsilon} \geq \operatorname{ess}\inf_{u\in\mathcal{U}} Y^u$, since $u^\varepsilon \in \mathcal{U}$. Combine the last two inequalities, it yields

$$Y_t^{u^\varepsilon} \geq \operatorname{ess}\inf_{u\in\mathcal{U}} Y_t^u \geq Y_t^{u^\varepsilon} - \varepsilon.$$

As $\varepsilon \to 0$, we obtain

$$Y^{u^\varepsilon} = \operatorname{ess}\inf_{u\in\mathcal{U}} Y^u.$$

Y^{u^ε} becomes a value function and u^ε becomes an optimal control. □

Remark 4.2. Actually, the theorem still holds if the boundness assumption on the coefficients d, n, k, χ and the Brownian filtration assumption is relaxed to weaker assumptions which ensure the uniqueness for the BSDE associated with (f, ξ) and the comparison theorem. In particular, by the results stated by El Karoui and Huang [15], the verification theorem still holds under the weaker assumptions:

$$\mathbb{E}\left(\int_0^T e^{\beta A_s} \frac{k_s^2}{\alpha_s^2} ds \right) < \infty; \qquad \mathbb{E}\left(e^{\beta A_T} \xi^2\right) < \infty,$$

where $\alpha_t^2 = \delta_t + v_t^2$ and $A_t = \int_0^t \alpha_s^2 ds$.

The main purpose of this chapter is to find a (0)-optimal control which is defined as follows:

Definition 4.2. (0)-optimal control u^0 satisfies the value function at time $t = 0$, that is,

$$\bar{Y}_0 = \operatorname*{ess\,inf}_{u \in \mathcal{U}} Y_0^u = Y_0^{u^0}.$$

Furthermore, the comparison theorem (Theorem 2.5) gives a criterion to find (0)-optimal controls in the following corollary (see [13]).

Corollary 4.8. *A control $\{u_s^0\}_{0 \le s \le T}$ is (0)-optimal if and only if*

$$f(s, Y_s, Z_s) = f^{u^0}(s, Y_s, Z_s), \qquad d\mathbb{P} \otimes dt \ \ a.s. \tag{4.13}$$

$$\xi = \xi^{u^0}, \qquad d\mathbb{P} \ \ a.s.. \tag{4.14}$$

In this case, u^0 is also optimal for the problem starting at time t, that is, $\bar{Y}_t = Y_t^{u^0}$.

Proof. The proof of this corollary can be found in [13]. For the sake of completeness, we also give the proof here. Assume that the equalities (4.13) and (4.14) hold. Then, by Proposition 4.4, $Y_s = Y_s^{u^0}, \forall s \in [0, T]$. So, we can say u^0 is an optimal control. In particular, choose $s = 0$, then $Y_0 = Y_0^{u^0}$ which means u^0 is (0)-optimal control.
On the other hand, let u^0 be (0)-optimal control. By definition, $Y_0 = Y_0^{u^0}$ (or $= \bar{Y}_0$). As an easy consequence of second part of the comparison theorem (Theorem 2.5), if $Y_0 = \bar{Y}_0$ on a set $\mathcal{A} \in \mathcal{F}_0$, then $Y_0 = \bar{Y}_0$ a.s. on $[0, T] \times \mathcal{A}$ and $\xi^{u^0} = \xi$ on $\mathcal{A}$ and $f(s, Y_s^{u^0}, Z_s^{u^0}) = f^{u^0}(s, Y_s^{u^0}, Z_s^{u^0})$ on $\mathcal{A} \times [0, T], d\mathbb{P} \otimes ds$ a.s. □

After we give the verification theorem, we use it to examine the optimization idea in the choices of standard generator.

4.5 Concavity or Convexity of Generators and Associated Optimality

Consider a standard generator of a given BSDE. In this section, we restrict ourselves to two cases of standard generators which are concave or convex. The approaches for two different situations are followed in detail. We refer to the reader [16] for the definitions, the lemma and the proposition for concave standard generator, but the rest of this section is firstly examined in this thesis. in this thesis and arranged in a clear manner. The arrangement is managed in two cases. In case 1, the optimal solution (the least cost) occurs as infimum of some related controlled state processes for the concave standard generator. On the other hand, the optimal solution (the best utility) occurs as supremum of some related controlled state processes for the convex standard generator in case 2.

Case 1: Assume that a standard generator denoted by $f(t,y,z)$ is concave with respect to (y,z). Then the polar process associated with f is defined by

$$F(\omega,t,\beta,\gamma) = \sup_{(y,z)\in\mathbb{R}\times\mathbb{R}^n} \left\{ f(\omega,t,y,z) - \beta y - \gamma^* z \right\}.$$

The effective domain of F occurs

$$\mathcal{D}_F = \left\{ (\omega,t,\beta,\gamma) \in \Omega \times [0,T] \times \mathbb{R} \times \mathbb{R}^n \mid F(\omega,t,\beta,\gamma) < +\infty \right\}.$$

Denote the (ω,t) section of $\mathcal{D}_F$ by $\mathcal{D}_F^{(\omega,t)}$. The following proposition gives a bound for $\mathcal{D}_F^{(\omega,t)}$.

Proposition 4.9. *$\mathcal{D}_F^{(\omega,t)}$ is bounded by the domain $K = [-C,C]^{n+1}$, where C is Lipschitz coefficient of f. More precisely, $\beta \in [-C,C]$ and $\gamma \in [-C,C]^n$.*

Proof. On the contrary, if $|\beta| > C$, then $\beta > C$ or $\beta < -C$. It follows that

$$f(\omega,t,y,z) - \beta y - \gamma^* z \geq f(\omega,t,y,z) - C|y| - \beta y - \gamma^* z.$$

Take supremum of only the components $-C|y| - \beta y$ over $y \in \mathbb{R}$. For $y \in \mathbb{R}^+$ and $\beta < -C$, it is obtained that

$$\sup_{y\in\mathbb{R}} \left\{ -C|y| - \beta y \right\} = \sup_{y\in\mathbb{R}} \left\{ -Cy - \beta y \right\} = \sup_{y\in\mathbb{R}} \Big\{ -\underbrace{y}_{>0}\underbrace{(C+\beta)}_{<0} \Big\} = \infty.$$

Therefore, we get the equality $F(\omega,t,\beta,\gamma) = \infty$ which contradicts with the assumption $(\beta,\gamma) \in \mathcal{D}_F$. In the similar way, $\gamma \in [-C,C]^n$. □

Let us intertwine f and F. F is derived from concave conjugate of f in the following manner:

$$f^{**}(t,\beta,\gamma) = \inf_{(y,z)\in\mathbb{R}\times\mathbb{R}^n} \left\{ \beta y + \gamma^* z - f(t,y,z) \right\}$$

is defined as concave conjugate of $f(t,y,z)$ (see [23]).
Notice that

$$-f^{**}(t,\beta,\gamma) = \sup_{(y,z)\in\mathbb{R}\times\mathbb{R}^n} \left\{ f(t,y,z) - \beta y - \gamma^* z \right\} = F(t,\beta,\gamma).$$

On the other hand, f is derived using concave conjugate of $-F$:

$$(-F)^{**}(t,y,z) = \inf_{(\beta,\gamma)\in\mathcal{D}_F^{(\omega,t)}} \left\{ \beta y + \gamma^* z + F(t,\beta,\gamma) \right\} = f(t,y,z).$$

Hence f and F are in such a conjugacy relation:

$$\begin{cases} F &= -f^{**}, \\ f &= (-F)^{**}. \end{cases}$$

Define a family of linear standard generators $f^{\beta,\gamma}$ as

$$f^{\beta,\gamma}(t,y,z) = F(t,\beta_t,\gamma_t) + \beta_t y + \gamma_t^* z,$$

where (β,γ) is called control parameters and predictable processes. Note that, f is infimum of $f^{\beta,\gamma}$. To guarantee that $f^{\beta,\gamma}$ is a standard generator, it is sufficient to assume that (β,γ) belongs to a set $\mathcal{A}_F$ which is defined as

$$\mathcal{A}_F = \left\{ (\beta,\gamma) \in \mathcal{P}, K\text{-valued} \mid \mathbb{E}\left[\int_0^T F(t,\beta_t,\gamma_t)^2 dt \right] < +\infty \right\}.$$

A member of $\mathcal{A}_F$ is said to be admissible control parameter.
Let (Y,Z) be the unique solution of the BSDE associated with the standard parameter (f,ξ), where f is concave and defined by

$$f(t,Y_t,Z_t) = \operatorname{ess\,inf} \left\{ f^{\beta,\gamma}(t,Y_t,Z_t), (\beta,\gamma) \in \mathcal{A}_F \right\}.$$

Then, the same logic of the lemma which states existence of an optimal control and proposition that shows the solution as an optimum of related controlled state processes occurs.

Lemma 4.10. *There exists an optimal control* $(\bar{\beta},\bar{\gamma}) \in \mathcal{A}_F$ *such that*

$$f(t,Y_t,Z_t) = \operatorname{ess\,inf} \left\{ f^{\beta,\gamma}(t,Y_t,Z_t), (\beta,\gamma) \in \mathcal{A}_F \right\} = f^{\bar{\beta},\bar{\gamma}}(t,Y_t,Z_t), \qquad d\mathbb{P}\otimes dt \text{ a.s..}$$

Proof. Recall that $f(\cdot,Y,Z)$, Y and Z are predictable processes. By the measurable selection theorem in [27] (see Appendix), there exists a pair of predictable (bounded) processes $(\bar{\beta},\bar{\gamma})$ which satisfies the equality

$$f(t,Y_t,Z_t) = f^{\bar{\beta},\bar{\gamma}}(t,Y_t,Z_t), \qquad d\mathbb{P}\otimes dt \text{ a.s..}$$

Moreover, $f(\cdot,Y,Z)$, Y and Z are square integrable by assumption, and we conclude that $\bar{\beta},\bar{\gamma}$ are also bounded. Then $\mathbb{E}\left[\int_0^T F(\cdot,\bar{\beta}_t,\bar{\gamma}_t)^2 dt\right] < +\infty$, which means that $(\bar{\beta},\bar{\gamma})$ belongs to $\mathcal{A}_F$. □

We extend the Proposition 4.4 for optimal control parameters $(\bar{\beta},\bar{\gamma})$, in the following proposition.

Proposition 4.11. *Let f be a concave standard generator and $f^{\beta,\gamma}$ be the associated linear standard generators. Suppose that f and $f^{\beta,\gamma}$ are linked by*

$$f(t,Y_t,Z_t) = \operatorname{ess\,inf} \left\{ f^{\beta,\gamma}(t,Y_t,Z_t), (\beta,\gamma) \in \mathcal{A}_F \right\}, \qquad d\mathbb{P}\otimes dt \text{ a.s..}$$

Then, for any time $t \in [0,T]$,

$$Y_t = \operatorname{ess\,inf} \left\{ Y_t^{\beta,\gamma} \mid (\beta,\gamma) \in \mathcal{A}_F \right\}, \qquad d\mathbb{P} \text{ a.s..}$$

Proof. Procedure of the proof is similar to the proof of the Proposition 4.4 in the previous section. □

The structure of the standard linear generators $f^{\beta,\gamma}$ allows us to write the adjoint process $(\Gamma^{\beta,\gamma}_{t,s}, t \leq s \leq T)$ which is the unique solution of the following forward SDE

$$d\Gamma_s = \Gamma_s \left(\beta_s ds + \gamma_s^* dW_s\right), \qquad \Gamma_t = 1.$$

By Proposition 2.4, the solution of the LBSDE can be written as

$$Y_t^{\beta,\gamma} = \mathbb{E}\left[\int_t^T \Gamma^{\beta,\gamma}_{t,s} F(s,\beta_s,\gamma_s) ds + \Gamma^{\beta,\gamma}_{t,s} \xi \mid \mathcal{F}_t\right].$$

$Y^{\beta,\gamma}$ turns out to be the objective function which is controlled by the control processes (β,γ). Here, $F(t,\beta,\gamma)$ is the running cost function and ξ is the the terminal cost.

Case 2: Suppose that a standard generator denoted by $b(t,x,z)$ is convex with respect to (x,z). Then, the polar process associated with b is defined by

$$B(t,\beta,\gamma) = \inf_{(x,z)\in\mathbb{R}\times\mathbb{R}^n} \left\{\beta x + \gamma^* z + b(t,x,z)\right\}.$$

Effective domain of B is

$$\mathcal{D}_B = \left\{(\omega,t,\beta,\gamma) \in \Omega \times [0,T] \times \mathbb{R} \times \mathbb{R}^n \mid B(\omega,t,\beta,\gamma) > -\infty\right\}.$$

$\mathcal{D}_B^{(\omega,t)}$ is defined similar to $\mathcal{D}_F^{(\omega,t)}$, that is, bounded by the domain $K = [-C,C]^{n+1}$ where C is Lipschitz coefficient of f.
We have a connection between b and B in such a way: B is derived from concave conjugate of $-b$

$$(-b)^{**}(t,\beta,\gamma) = \inf_{(x,z)\in\mathbb{R}\times\mathbb{R}^n} \left\{\beta x + \gamma^* z + b(t,x,z)\right\} = B(t,\beta,\gamma).$$

On the other hand, b is derived from the concave conjugate of B

$$\begin{aligned} B^{**}(t,x,z) &= \inf_{\mathcal{D}_B^{(\omega,t)}} \left\{\beta x + \gamma^* z - B(t,\beta,\gamma)\right\}, \\ -B^{**}(t,x,z) &= \sup_{\mathcal{D}_B^{(\omega,t)}} \left\{B(t,\beta,\gamma) - \beta x - \gamma^* z\right\} = b(t,x,z). \end{aligned}$$

Hence b and B are in such a conjugacy relation:

$$\begin{cases} B &= (-b)^{**}, \\ b &= -B^{**}. \end{cases}$$

Define a family of linear standard generators $b^{\beta,\gamma}$ as

$$b^{\beta,\gamma}(t,x,z) = B(t,\beta_t,\gamma_t) - \beta_t x - \gamma_t^* z,$$

where (β,γ) is predictable control parameters. To ensure that $b^{\beta,\gamma}$ is a standard generator, it is sufficient to assume that (β,γ) belongs to a set $\mathcal{A}_B$ which is defined as

$$\mathcal{A}_B = \left\{(\beta,\gamma) \in \mathcal{P}, K\text{-valued} \mid \mathbb{E}\left[\int_0^T B(t,\beta_t,\gamma_t)^2 dt\right] < +\infty\right\}.$$

A member of $\mathcal{A}_B$ is said to be admissible, as well. Let (X,Z) be the solution of BSDE associated with the standard parameter (b,ξ), where b is convex and defined by

$$b(t,X_t,Z_t) = \operatorname{ess\,sup}\left\{b^{\beta,\gamma}(t,X_t,Z_t),(\beta,\gamma)\in\mathcal{A}_B\right\}.$$

Then, the following lemma and proposition occur.

Lemma 4.12. *There exists an optimal control $(\bar\beta,\bar\gamma)\in\mathcal{A}_B$ such that*

$$b(t,X_t,Z_t) = \operatorname{ess\,sup}\left\{b^{\beta,\gamma}(t,Y_t,Z_t),(\beta,\gamma)\in\mathcal{A}_B\right\} = b^{\bar\beta,\bar\gamma}(t,X_t,Z_t), \qquad d\mathbb{P}\otimes dt \ \ a.s..$$

Proof. By the measurable selection theorem in [27] (see Appendix), there exists a pair of predictable (bounded) processes $\bar\beta,\bar\gamma$ such that

$$b(t,X_t,Z_t) = b^{\bar\beta,\bar\gamma}(t,X_t,Z_t), \qquad d\mathbb{P}\otimes dt \text{ a.s..}$$

Furthermore, $b(\cdot,X,Z),X,Z$ are square-integrable by assumption and we derive the fact that $\bar\beta,\bar\gamma$ are also bounded. Therefore, $\mathbb{E}\left[\int_0^T B(t,\bar\beta_t,\bar\gamma_t)^2dt\right] < +\infty$ which implies that $(\bar\beta,\bar\gamma)\in\mathcal{A}_B$. □

Proposition 4.13. *Let b be a convex standard generator and $b^{\beta,\gamma}$ be the associated linear standard generators. Suppose that b and $b^{\beta,\gamma}$ are related such as*

$$b(t,X_t,Z_t) = \operatorname{ess\,sup}\left\{b^{\beta,\gamma}(t,Y_t,Z_t),(\beta,\gamma)\in\mathcal{A}_B\right\}, \qquad d\mathbb{P}\otimes dt \ \ a.s..$$

Then, for any time $t\in[0,T]$,

$$X_t = \operatorname{ess\,sup}\left\{X_t^{\beta,\gamma} \mid (\beta,\gamma)\in\mathcal{A}_B\right\}, \qquad d\mathbb{P} \ \ a.s..$$

Proof. Steps of the proof is similar to the proof of the Proposition 4.4 in the previous section. □

The structure of the standard linear generators $b^{\beta,\gamma}$ implies to write the adjoint process $(\Theta_{t,s}^{\beta,\gamma}, t\le s\le T)$ which is the unique solution of the following forward SDE:

$$d\Theta_s = -\Theta_s\left(\beta_s ds + \gamma_s^* dW_s\right), \qquad \Theta_t = 1.$$

By Proposition 2.4, the solution of the LBSDE can be written as

$$X_t^{\beta,\gamma} = \mathbb{E}\left[\int_t^T \Theta_{t,s}^{\beta,\gamma} B(s,\beta_s,\gamma_s)ds + \Theta_{t,s}^{\beta,\gamma}\xi \mid \mathcal{F}_t\right].$$

$X^{\beta,\gamma}$ becomes the objective function controlled by the control processes (β,γ), where the utility function is $B(t,\beta,\gamma)$ and the utility of money left at maturity is ξ.
An application for each case is provided in the following chapter.

CHAPTER 5

APPLICATIONS

We provide two different applications in this chapter, according to studies examined in Chapter 4. The first application is given for the BSDE with concave standard generator in a complete market. We study with the model involving consumption (see [20]). The consumption rate is chosen as a concave function and this yields to a concave standard generator of the BSDE. We refer to the reader [2] for the idea of concave consumption rate. After some arrangements, we obtain the optimal solution as infimum of some related controlled standard generators. Next, the second application is provided for the BSDE with convex standard generator in an incomplete market. We study with the model called hedging contingent claims with a higher interest rate for borrowing. The model and the structure of related optimal solution are provided in [16], but all details are shown in this thesis.

5.1 Optimality with Concave Generator

In this section, we study with an application of a BSDE associated with a concave standard generator f. The model is the same as Merton's (see [20] as a further study) and it seems a linear BSDE with a consumption rate of the form:

$$dY_t = r_t Y_t \, dt + \pi_t^* \sigma_t \theta_t \, dt + \pi_t^* \sigma_t \, dW_t - c_t dt, \qquad Y_T = \xi,$$

where Y is the wealth process, $\sigma^* \pi = Z$ corresponds the hedging portfolio and c_t is the predictable consumption rate at time t. According to Carroll and Kimball [2], when the wealth process Y contains stochastic income, a concave consumption arises. Therefore, we work with a concave consumption rate defined by

$$c_t = c(t, Y_t) = \eta Y_t,$$

where η is nonnegative constant. Then, the linear BSDE becomes

$$dY_t = r_t Y_t \, dt + \pi_t^* \sigma_t \theta_t \, dt + \pi_t^* \sigma_t \, dW_t - \eta Y_t \, dt, \qquad Y_T = \xi,$$

or, equivalently,

$$dY_t = r_t Y_t \, dt + Z_t^* \theta_t \, dt + Z_t^* \, dW_t - \eta Y_t \, dt, \qquad Y_T = \xi.$$

The pair (f, ξ) is a standard parameter, where f is obtained by

$$f(t, y, z) = -r_t y - z^* \theta_t + \eta y.$$

The standard generator f is concave with respect to (y, z), since the consumption rate c is concave with respect to y. Now, we can write the polar process F as follows:

$$\begin{aligned} F(t, \beta, \gamma) &= \sup_{(y,z)\in\mathbb{R}\times\mathbb{R}^n} \{-r_t y - z^* \theta_t + \eta y - \beta y - z^* \gamma\} \\ &= \sup_{(y,z)\in\mathbb{R}\times\mathbb{R}^n} \{-(r_t + \beta - \eta) y - z^*(\theta_t + \gamma)\} \\ &= \begin{cases} 0, & \text{if } \beta > \eta - r_t \text{ and } \gamma = -\theta_t, \\ \infty, & \text{otherwise.} \end{cases} \end{aligned}$$

The effective domain $\mathcal{D}_F^{(\omega,t)}$ must be less then $+\infty$, therefore we only consider $F = 0$ for the case $\gamma = -\theta_t$ and $\beta > \eta - r_t$.
Remember that f can be written as essential infimum of the family of related standard generators $f^{\beta,\gamma}$, which is equal to

$$\begin{aligned} f^{\beta,\gamma}(t, y, z) &= F(t, \beta_t, \gamma_t) + \beta_t y + z^* \gamma_t = \beta_t y + z^* \gamma_t \\ &= \beta_t y + (\pi_t^\beta)^* \sigma_t \gamma_t = \beta_t y - (\pi_t^\beta)^* \sigma_t \theta_t = \beta_t y - z^* \theta_t = f^\beta(t, y, z). \end{aligned}$$

The standard generators f^β can be also written by

$$f^\beta(t, y, \sigma^* \pi_t^\beta) = \beta_t y - (\pi_t^\beta)^* \sigma_t \theta_t.$$

The BSDE associated with the standard parameter (f^β, ξ) is

$$\begin{aligned} -dY_t^\beta &= (-\beta_t Y_t + (\pi_t^\beta)^* \sigma_t \theta_t) dt - (\pi_t^\beta)^* \sigma_t dW_t, \\ Y_T &= \xi. \end{aligned} \tag{5.1}$$

Hence, the solution Y associated with the standard parameter (f, ξ) can be written as

$$Y_t = \operatorname{ess\,inf} \left\{ Y_t^\beta, \beta > \eta - r_t \right\},$$

where Y^β is the solution of the BSDE in (5.1).
Moreover, we rewrite the related standard generator f^β as follows:

$$f^\beta(t, y, z) = \beta_t y - \theta_t^* z.$$

By Proposition 2.4, the solution of linear BSDE in (5.1) can be written as

$$Y_t^\beta = \mathbb{E}\left[\int_t^T \Gamma_{t,s}^\beta F(s, \beta_s, \gamma_s) ds + \Gamma_{t,s}^\beta \xi \mid \mathcal{F}_t \right],$$

where $\{\Gamma_{t,s}^\beta\}_{t \le s \le T}$ is the adjoint process with the following dynamics

$$d\Gamma_{t,s}^\beta = \Gamma_{t,s}^\beta \left(\beta_s ds + \theta_s^* dW_s \right), \qquad \Gamma_{t,t}^\beta = 1.$$

Furthermore, $F(s, \beta_s, \gamma_s)$ corresponds the running cost while ξ is the the terminal cost.
Since the fact that $F(s, \beta_s, \gamma_s) = 0$, we can write

$$Y_t^\beta = \mathbb{E}\left[\Gamma_{t,s}^\beta \xi \mid \mathcal{F}_t\right]. \tag{5.2}$$

Therefore, the optimal solution can be represented by

$$Y_t = \operatorname{ess\,inf}\{Y_t^\beta, \beta > \eta - r_t\},$$

where Y^β is obtained from the equation (5.2).
We apply Theorem 4.7 (the verification theorem) for the Case 1 examined in Section 4.5 and offer an optimal solution which achieves the least cost. So far, we obtain the optimal solution in the case of concave standard generator. In the next section, we follow the similar procedure for the case of convex standard generator in order to get the optimal solution which achieves the best utility.

5.2 Optimality with Convex Generator

In this section, we have an application for the solution of a non-linear BSDE with convex standard generator. We study with the model called hedging contingent claims with higher interest rate for borrowing. The idea in this model is that interest rate for borrowing is higher than for investing. Cvitanic and Karatzas [10] and Korn [19] studied this model. Our aim is to apply Theorem 4.7 (the verification theorem) and obtain the optimal solution for the Case 2 examined in Section 4.5. In this thesis, we firstly construct the problem and explain the idea of the model. Then, we show all steps to obtain an optimal solution which achieves the best utility.
Suppose that the individual is allowed to borrow money at time t at a predictable and bounded interest rate $R_t > r_t$, where r_t is the bond rate (risk-free rate). Because of the reason that there is no point to borrow money and invest money in bond at the same time, we decide the amount borrowed at time t as follows:

$$(\pi_t^0)^- = \left(V_t - \sum_{i=1}^n \pi_t^i\right)^- = \max\left\{-\left(V_t - \sum_{i=1}^n \pi_t^i\right), 0\right\}.$$

To be more clear, if the total amount of money invested in n stocks does not exceed the wealth at the same time t, there is no point to borrow money and investor naturally invest the amount of money which equals to the difference $V_t - \sum_{i=1}^n \pi_t^i$ in the riskless asset at the risk-free rate r_t. Then the BSDE of the strategy (V, π) is the same as in Section 3.2:

$$dV_t = r_t V_t \, dt + \pi_t^* \sigma_t \theta_t \, dt + \pi_t^* \sigma_t \, dW_t, \qquad V_T = \xi. \tag{5.3}$$

On the other hand, if the total amount of money invested in n stocks exceeds the wealth at the same time t, investor has to borrow money which exactly equals to the difference $\sum_{i=1}^n \pi_t^i - V_t$ at a higher rate R_t. Then, the term came from borrowed money is negative and subtracted from the wealth SDE which is of the following form:

$$\begin{aligned}
dV_t &= \sum_{i=0}^n \pi_t^i \frac{dP_t^i}{P_t^i} \\
&= -(\pi_t^0)^- \frac{dP_t^0}{P_t^0} + \sum_{i=1}^n \pi_t^i \frac{dP_t^i}{P_t^i} \\
&= -\left(-\left(V_t - \sum_{i=1}^n \pi_t^i\right)\right) R_t\,dt + \sum_{i=1}^n \pi_t^i \left(b_t^i\,dt + \sum_{j=1}^n \sigma_t^{i,j} dW_t^j\right) \\
&= R_t V_t\,dt - \sum_{i=1}^n \pi_t^i R_t\,dt + \sum_{i=1}^n \pi_t^i b_t^i\,dt + \sum_{i=1}^n \pi_t^i \sum_{j=1}^n \sigma_t^{i,j} dW_t^j \pm \sum_{i=1}^n \pi_t^i r_t\,dt \\
&= R_t V_t\,dt - \sum_{i=1}^n \pi_t^i (R_t - r_t)\,dt + \sum_{i=1}^n \pi_t^i (b_t^i - r_t \mathbf{1})\,dt + \sum_{i=1}^n \pi_t^i \sum_{j=1}^n \sigma_t^{i,j} dW_t^j \\
&= R_t V_t\,dt - \sum_{i=1}^n \pi_t^i (R_t - r_t)\,dt + \pi_t^* \sigma_t \theta_t\,dt + \pi_t^* \sigma_t\,dW_t \pm r_t V_t\,dt \\
&= r_t V_t\,dt + (R_t - r_t) V_t\,dt - \sum_{i=1}^n \pi_t^i (R_t - r_t)\,dt + \pi_t^* \sigma_t \theta_t\,dt + \pi_t^* \sigma_t\,dW_t \\
&= r_t V_t\,dt + \pi_t^* \sigma_t \theta_t\,dt + \pi_t^* \sigma_t\,dW_t + (R_t - r_t)\left(V_t - \sum_{i=1}^n \pi_t^i\right) dt. \qquad (5.4)
\end{aligned}$$

From the equations (5.3) and (5.4), we obtain

$$dV_t = r_t V_t\,dt + \pi_t^* \sigma_t \theta_t\,dt + \pi_t^* \sigma_t\,dW_t - (R_t - r_t)\left(V_t - \sum_{i=1}^n \pi_t^i\right)^- dt.$$

The fair price of a contingent claim ξ is still defined as the minimal endowment which guarantees ξ at time T. According to the results of Pardoux and Peng [4] and El Karoui et al. [16], there exists a unique square integrable strategy (X, π) which solves the nonlinear BSDE. The nonlinear term depends on both wealth and portfolio;

$$dX_t = r_t X_t\,dt + \pi_t^* \sigma_t \theta_t\,dt + \pi_t^* \sigma_t\,dW_t - (R_t - r_t)(\pi_t^0)^-, \quad X_T = \xi, \qquad (5.5)$$

with $\mathbb{E}\left(\int_0^T |\pi_t^* \sigma_t|^2\,dt\right) < \infty$. Moreover, X_t is the fair price of ξ at time t.
Here, we consider the solution X as an supremum of some related controlled state processes $X^{\beta,\gamma}$ over the pair of control process (β, γ) which belong to a Polish space. The solution pair (X, π) is the hedging portfolio and the unique solution of the BSDE (5.5).
Defining $Z = \sigma^* \pi$, the equation (5.5) becomes

$$dX_t = r_t X_t\,dt + Z_t^* \theta_t\,dt + Z_t^*\,dW_t - (R_t - r_t)\left(X_t - \sum_{i=1}^n ((\sigma_t^*)^{-1} Z_t)^i\right)^-, \quad X_T = \xi.$$

The pair (b,ξ) is a standard parameter, where b is obtained by

$$b(t,x,z) = -r_t x - z^*\theta_t + (R_t - r_t)\left(x - ((\sigma_t^*)^{-1}z)^*\mathbf{1}\right)^- .$$

Here, $\mathbf{1}$ denotes the n dimensional column vector with all components is equal to 1.
The standard generator b is a convex function. Indeed, for any $\lambda \in [0,1]$ and $(x^1,z^1),(x^2,z^2) \in \mathbb{R}\times\mathbb{R}^n$,

$$\begin{aligned}\lambda\, b(t,x^1,z^1) + (1-\lambda)b(t,x^2,z^2) = \lambda\Big(& -r_t x^1 - z^{1^*}\theta_t \\ &+ (R_t - r_t)\left(x^1 - ((\sigma_t^*)^{-1}z^1)^*\mathbf{1}\right)^- \Big) \\ + (1-\lambda)\Big(& -r_t x^2 - z^{2^*}\theta_t \\ &+ (R_t - r_t)\left(x^2 - ((\sigma_t^*)^{-1}z^2)^*\mathbf{1}\right)^- \Big).\end{aligned}$$

Define a function g such that

$$g(x,z) = \left(x - ((\sigma_t^*)^{-1}z)^*\mathbf{1}\right),$$

where $g(x,z)^-$ is convex with respect to x and z. Therefore,

$$\begin{aligned}\lambda b(t,x^1,z^1) + (1-\lambda)b(t,x^2,z^2) &= -r_t(\lambda x^1 + (1-\lambda)x^2) - (\lambda z^1 + (1-\lambda)z^2)^*\theta_t \\ &\quad + (R_t - r_t)\left[\lambda g(x^1,z^1)^- + (1-\lambda)g(x^2,z^2)^-\right] \\ &\geq -r_t(\lambda x^1 + (1-\lambda)x^2) - (\lambda z^1 + (1-\lambda)z^2)^*\theta_t \\ &\quad + (R_t - r_t)\left[\lambda g(x^1,z^1) + (1-\lambda)g(x^2,z^2)\right]^- \\ &= b(t, \lambda x^1 + (1-\lambda)x^2, \lambda z^1 + (1-\lambda)z^2) \\ &= b(\lambda(t,x^1,z^1) + (1-\lambda)(t,x^2,z^2)).\end{aligned}$$

Recall that $z = \sigma^*\pi$, hence $\pi^* = z^*\sigma^{-1}$. For simplicity, let us rewrite b in the following manner:

$$\begin{aligned}b(t,x,z) &= -r_t x - z^*\theta_t + (R_t - r_t)\left(x - ((\sigma_t^*)^{-1}z)^*\mathbf{1}\right)^- \\ &= -r_t x - z^*\theta_t + (R_t - r_t)\left(x - z^*\sigma_t^{-1}\mathbf{1}\right)^- .\end{aligned}$$

Then, we can obtain the associated polar process B as follows:

$$\begin{aligned}B(t,\beta,\gamma) &= \inf_{(x,z)\in\mathbb{R}\times\mathbb{R}^n}\{b(t,x,z) + \beta x + \gamma^* z\} \\ &= \inf_{(x,z)\in\mathbb{R}\times\mathbb{R}^n}\left\{-r_t x - z^*\theta_t + (R_t - r_t)\left(x - z^*\sigma_t^{-1}\mathbf{1}\right)^- + \beta x + \gamma^* z\right\} \\ &= \inf_{(x,z)\in\mathbb{R}\times\mathbb{R}^n}\left\{x(\beta - r_t) + (\gamma - \theta_t)^* z + (R_t - r_t)\left(x - z^*\sigma_t^{-1}\mathbf{1}\right)^-\right\} \\ &= \inf_{(x,z)\in\mathbb{R}\times\mathbb{R}^n}\left\{x(\beta - r_t) + z^*(\gamma - \theta_t) + (R_t - r_t)\left(x - z^*\sigma_t^{-1}\mathbf{1}\right)^-\right\}.\end{aligned}$$

Put $\gamma = \theta_t - \sigma_t^{-1}(\beta - r_t)\mathbf{1}$, then it yields

$$B(t,\beta,\gamma) = \inf_{(x,z)\in\mathbb{R}\times\mathbb{R}^n} \left\{(\beta - r_t)\left(x - z^*\sigma_t^{-1}\mathbf{1}\right) + (R_t - r_t)\left(x - z^*\sigma_t^{-1}\mathbf{1}\right)^-\right\}.$$

If $x - z^*\sigma_t^{-1}\mathbf{1} < 0$, then $(x - z^*\sigma_t^{-1}\mathbf{1})^- = -(x - z^*\sigma_t^{-1}\mathbf{1})$ and the polar process becomes

$$\begin{aligned} B(t,\beta,\gamma) &= \inf_{(x,z)\in\mathbb{R}\times\mathbb{R}^n} \left\{(\beta - r_t)(x - z^*\sigma_t^{-1}\mathbf{1}) - (R_t - r_t)(x - z^*\sigma_t^{-1}\mathbf{1})\right\} \\ &= \inf_{(x,z)\in\mathbb{R}\times\mathbb{R}^n} \left\{(\beta - R_t)(x - z^*\sigma_t^{-1}\mathbf{1})\right\} \\ &= \begin{cases} -\infty, & \text{if } \beta > R_t, \\ 0, & \text{if } \beta \leq R_t. \end{cases} \end{aligned}$$

If $x - z^*\sigma_t^{-1}\mathbf{1} \geq 0$, then $(x - z^*\sigma_t^{-1}\mathbf{1})^- = 0$. By the $x - z^*\sigma_t^{-1}\mathbf{1} \geq 0$ assumption, the polar process becomes

$$\begin{aligned} B(t,\beta,\gamma) &= \inf_{(x,z)\in\mathbb{R}\times\mathbb{R}^n} \left\{(\beta - r_t)(x - z^*\sigma_t^{-1}\mathbf{1})\right\} \\ &= \begin{cases} -\infty, & \text{if } \beta < r_t, \\ 0, & \text{if } \beta \geq r_t. \end{cases} \end{aligned}$$

Hence ,

$$B(t,\beta,\gamma) = \begin{cases} 0, & \text{if } \gamma = \theta_t - \sigma_t^{-1}(\beta - r_t)\mathbf{1} \text{ and } r_t \leq \beta \leq R_t, \\ -\infty, & \text{otherwise.} \end{cases}$$

The effective domain $\mathcal{D}_B^{(\omega,t)}$ must be greater than $-\infty$, therefore, we only consider $B = 0$ for the case $\gamma = \theta_t - \sigma_t^{-1}(\beta - r_t)\mathbf{1}$ and $r_t \leq \beta \leq R_t$.
Recall that b can be written as essential supremum of the family of related standard generators which is equal to

$$\begin{aligned} b^{\beta,\gamma}(t,x,z) &= B(t,\beta_t,\gamma_t) - \beta_t x - \gamma_t^* z \\ &= -\beta_t x - z^*\gamma_t \\ &= -\beta_t x - z^*(\theta_t - \sigma_t^{-1}(\beta - r_t)\mathbf{1}), \qquad z^* = \pi^*\sigma, \\ &= -\beta_t x - \pi^*\sigma_t\theta_t + \pi^*(\beta - r_t)\mathbf{1} \\ &= b^\beta(t,x,\sigma^*\pi). \end{aligned}$$

The BSDE associated with the standard parameter (b^β,ξ) is

$$\begin{aligned} -dX_t^\beta &= \left(-\beta_t X_t^\beta - (\pi_t^\beta)^*\sigma_t\theta_t + (\pi_t^\beta)^*(\beta_t - r_t)\mathbf{1}\right)dt - (\pi_t^\beta)^*\sigma_t dW_t, \\ X_T^\beta &= \xi. \end{aligned} \tag{5.6}$$

Finally, the solution X associated with the standard parameter (b,ξ) can be written as

$$X_t = \operatorname{ess\,sup}\left\{X_t^\beta, r_t \leq \beta_t \leq R_t\right\},$$

where X^β is the solution of the BSDE in (5.6).
Moreover, we rewrite the related standard generator b^β as follows:

$$b^\beta(t,x,z) = -\beta_t x - (\theta_t - \sigma_t^{-1}(\beta_t - r_t)\mathbf{1})^* z.$$

By Proposition 2.4, the solution of linear BSDE in (5.6) can be written as

$$X_t^\beta = \mathbb{E}\left[\int_t^T \Theta_{t,s}^\beta B(s,\beta_s,\gamma_s)ds + \Theta_{t,s}^\beta \xi \mid \mathcal{F}_t\right],$$

where $\{\Theta_{t,s}^\beta\}_{t\leq s\leq T}$ is the adjoint process with the following dynamics

$$d\Theta_{t,s}^\beta = -\Theta_{t,s}^\beta\left(\beta_s ds + (\theta_t - \sigma_t^{-1}(\beta_t - r_t)\mathbf{1})^* dW_s\right), \qquad \Theta_{t,t}^\beta = 1.$$

Furthermore, $B(s,\beta_s,\gamma_s)$ corresponds the utility function while ξ is the utility of money left at maturity.
Since the fact that $B(s,\beta_s,\gamma_s) = 0$, we can write

$$X_t^\beta = \mathbb{E}\left[\Theta_{t,s}^\beta \xi \mid \mathcal{F}_t\right]. \tag{5.7}$$

Therefore, the optimal solution can be represented by

$$X_t = \operatorname{ess\,sup}\{X_t^\beta, r_t \leq \beta \leq R_t\},$$

where X^β is obtained from the equation (5.7).
Hence, we offer an optimal solution that achieves the best utility for the model called hedging contingent claims with a higher interest rate for borrowing.

CHAPTER 6

CONCLUSION

This thesis fundamentally deals with the theory of backward stochastic differential equations (BSDEs) and stochastic control problems. By this thesis, we aggregate these two theories. The idea is that we use BSDE in order to obtain an optimal solution to stochastic control problems. We study the article [16] for backward stochastic differential equations theory and the article [13] for stochastic control problems in detail.
First of all, we mention and prove the fundamental theorems of BSDE theory. Following [16], we show that there exists a unique solution pair (Y, Z) to the following BSDE

$$-dY_t = f(t, Y_t, Z_t)dt - Z_t^* dW_t, \qquad Y_T = \xi,$$

where (f, ξ) is a standard parameter. We use a method called *a Priori Estimates*. The existence and uniqueness theorem is also applied to a linear BSDE of the form

$$-dY_t = (\varphi_t + Y_t\beta_t + Z_t^*\gamma_t)\ dt - Z_t^*\ dW_t, \qquad Y_T = \xi,$$

under some restrictions on the coefficients φ, β and γ. In addition, we give the comparison theorem as a consequence.
In the following, we construct a complete market including n risky assets and a riskless asset. We define the fair price (upper price) with the help of hedging strategies (superstrategies). Thanks to the existence and uniqueness theorem, the fair price can be written as a unique solution of a linear BSDE associted with a given standard parameter.
We also closely study [13], [24] and [26]. After a brief introduction to optimal control theory, we derive the Hamilton-Jacobi-Bellman equation in the cases of the objective function. Optimization idea differs with respect to its components. If it contains cost functions, then it turns out to be minimization. On the other hand, if the components are utility functions, then maximization approach arises. To construct stochastic control problems, we follow the work of Quenez [13]. Given a standard parameter (f, ξ), we can write the standard generator f as essential infimum of some related f^α, so as the terminal condition ξ. Then, we show that the solution of BSDE associated with the standard parameter (f, ξ) can be written as essential infimum of the solution of BSDE associated with the standard parameter (f^α, ξ^α) by the verification theorem. As an application of the verification theorem, we first examine the optimization problem in two choices

of standard generator. If the standard generator is concave, we minimize the related controlled parameters which corresponds the least cost. However, if it is convex, we maximize the related controlled parameters which corresponds the best utility.
At the end of the thesis, an application is provided for each situation. In the first application, we consider the model with consumption. We assume that the consumption rate concave in order to obtain a concave standard generator, as well. It is possible since the fact that the wealth process including stochastic income yields concave consumption (see [2]). Therefore, we can write the optimal solution as the infimum of the solution associated with some related controlled standard parameters. Secondly, we consider a model with higher interest rate for borrowing. For direct results, we refer to the reader [16]. We explain the idea of that model and show all steps to get the solution. In this way, the standard generator is convex and the optimal solution can be represented by the supremum of the solution associated with some related controlled standard parameters.
More applications could be modeled in a similar way. Note that, the optimization problem can be also examined in terms of objective functions and the derivation of Hamilton-Jacobi-Bellman (HJB) equations could be constructed for the given cost or utility functions. As a result, the optimal solution can be found by two approaches: using the standard generator of BSDE and using the objective function. Combining these two approaches and comparing the results of them is left as a further study.

REFERENCES

[1] A.V. Arhangel'skii, *General Topology III*, Springer- Verlag Berlin Heidelberg, Volume 3, pp. 206, 1995.

[2] C.D. Carroll and M.S. Kimball, *On the Concavity of the Consumption Function*, Econometrica, 64, pp. 981-992, 1996.

[3] D. Lamberton, B. Lapeyre, *Introduction to Stochastic Calculus Applied to Finance*, Chapman & Hall/CRC, pp. 18,91, 2008.

[4] E. Pardoux, S. Peng, *Adapted Solution of a Backward Stochastic Differential Equation*, Systems Control Lett., 14, pp. 55-61, 1990.

[5] E. Kreyzig, *Introductory Functional Analysis with Applications*, John Wiley & Sons. Inc., pp. 299, 1978.

[6] F. Black, M. Scholes, *The Pricing of Options and Corporate Liabilities*, J. Political Econ., 3, pp. 637-654, 1973.

[7] H. Körezlioğlu, A.B. Hayfavi, *Elements of Probability Theory*, ODTÜ Basım İşliği, pp. 118, 2001.

[8] I. Karatzas, J.P. Lehoczky, S.E. Shreve and G.L. Xu, *Martingale and Duality Methods for Utility Maximisation in an Incomplete Market*, SIAM J. Control Optim., 29, pp. 702-730, 1989.

[9] I. Karatzas and S. Shreve, *Brownian Motion and Stochastic Calculus*, New York: Springer Verlag, 1987.

[10] J. Cvitanic and I. Karatzas, *Convex Duality in Constrained Portfolio Optimization*, Ann. App. Prob., 2, pp. 767-818, 1992.

[11] J. Cvitanic and I. Karatzas, *Hedging Contingent Claims with Constrained Portfolios*, Ann. App. Prob., 3, pp. 652-681, 1993.

[12] J.M. Bismut, *Conjugate Convex Functions in Optimal Stochastic Control*, J. Math. Anal. Appl., 44, pp. 384-404, 1973.

[13] M.C. Quenez, *Stochastic Control and Backward Stochastic Differential Equations*, Addison Wesley Longman Ltd., Backward Stochastic Differential Equations (Paris, 1995-1996), Pitman Res. Notes Math. Ser., 364, pp. 83-95, 1997.

[14] M. Harrison, S.R. Pliska, *Martingales and Stochastic Integrals in the Theory of Continuous Trading*, Stoch. Process. Appl., 11, pp. 215-260, 1981.

[15] N. El Karoui and S. Huang, *A General Result of Existence and Uniqueness of Backward Stochastic Differential Equations*, Addison Wesley Longman Ltd., Backward Stochastic Differential Equations (Paris, 1995-1996), Pitman Res. Notes Math. Ser., 364, pp. 27-36, 1997.

[16] N. El Karoui, S. Peng, M.C. Quenez, *Backward Stochastic Differential Equations in Finance*, Mathematical Finance, 7, pp. 1-10, 17-41, 1997.

[17] N. El Karoui, C. Kapoudjian, E. Pardoux, S. Peng, M.C. Quenez, *Reflected Solutions of Backward SDE's, and Related Obstacle Problems for PDE's*, Ann. Probab., 25, pp. 702-737, 1997.

[18] R.A. Silverman, *Calculus with Analytic Geometry*, Prentice-Hall Inc., 1985.

[19] R. Korn, *Contingent Claim Valuation in a Market with Different Interest Rates*, Mathematical Methods of Operations Research, 42, pp. 255-274, 1995.

[20] R. Merton, *Optimum Consumption and Portfolio Rules in a Continuous Time Model*, J. Econ. Theory, 3, pp. 373-413, 1971.

[21] R. Merton, *Theory of Rational Option Pricing*, Bell J. Econ. Manage. Sci., 4, pp. 141-183, 1973.

[22] R. Merton, *Continuous Time Finance*, Basil Blackwell, Revised Edition, 1991.

[23] S. Boyd and L. Vanderberghe, *Convex Optimization*, Cambridge University Press, 2004.

[24] S. Peng, *Stochastic Hamilton-Jacobi-Bellman Equations*, SIAM J. Control Optim., 30, pp. 284-304, 1992a.

[25] S. Peng, *A Generalized Dynamic Programming Principle and Hamilton-Jacobi-Bellman Equation*, Stochasitcs, 38, pp. 119-134, 1992b.

[26] T. Björk, *Arbitrage Theory in Continuous Time*, Oxford University Press Inc., pp. 198-206, 1998.

[27] V.I. Bogachev, *Measure Theory*, Springer- Verlag Berlin Heidelberg, Volume 2, pp. 33-36, 2007c.

APPENDIX A

DEFINITIONS AND THEOREMS

A.1 Definitions

Martingale Definitions:
An adapted sequence $(M_t)_{0\leq t\leq T}$ of real-valued random variables is

- a martingale if $\mathbb{E}(M_{t+1}|\mathcal{F}_t) = M_t$ for all $t \leq T-1$,
- a supermartingale if $\mathbb{E}(M_{t+1}|\mathcal{F}_t) \leq M_t$ for all $t \leq T-1$,
- a submartingale if $\mathbb{E}(M_{t+1}|\mathcal{F}_t) \geq M_t$ for all $t \leq T-1$.

(Reference: [3])

Concavity and Convexity:
Let X and Y be integrable real random variables and let f be continuous function on $\mathbb{R}^2$. Then,

- f is said to be concave if and only if for any $(X^1, Y^1), (X^2, Y^2) \in \mathbb{R}^2$ and $\lambda \in [0, 1]$,
$$\lambda f(X^1, Y^1) + (1-\lambda) f(X^2, Y^2) \leq f(\lambda(X^1, Y^1) + (1-\lambda)(X^2, Y^2)).$$
- f is said to be convex if and only if for any $(X^1, Y^1), (X^2, Y^2) \in \mathbb{R}^2$ and $\lambda \in [0, 1]$,
$$\lambda f(X^1, Y^1) + (1-\lambda) f(X^2, Y^2) \geq f(\lambda(X^1, Y^1) + (1-\lambda)(X^2, Y^2)).$$

(Reference: [7])

Polish Space:
According to descriptive set theory, a space is said to be *Polish space* if and only if it is seperable and complete metrizable.
(Reference: [1])

Let $\Psi : X \mapsto Y$ be some mapping. For every point $y \in \Psi(X)$, we can pick an element $x = \zeta(y) = \Psi^{-1}(y)$. Thus, we obtain a mapping ζ such that $\Psi \circ \zeta$ is the identity mapping on the range of Ψ. The mapping ζ is called a selection of the mapping Ψ. One can easily take for ζ a mapping with measurability properties. This is the content of the following measurable selection theorem.

A.2 Theorems

Theorem A.1 (Measurable Selection Theorem)**.** *Let X be a Polish space and let Ψ be a mapping on $(\Omega, \mathcal{B})$ with values in the set of nonempty closed subsets of X. Suppose that for every open set $U \subset X$, we have*

$$\hat{\Psi} := \{\omega \mid \Psi(\omega) \cap U \neq \emptyset\} \in \mathcal{B}.$$

Then, Ψ has a selection ζ that is measurable with respect to the pair of σ-algebras $\mathcal{B}$ and $\mathcal{B}(X)$.

(Reference: [27])

Theorem A.2 (Martingale Representation Theorem)**.** *Let $(M_t)_{0\le t\le T}$ be a square-integrable martingale, with respect to the filtration $(\mathcal{F}_t)_{0\le t\le T}$. There exists an adapted process $(Z_t)_{0\le t\le T}$ such that $\mathbb{E}\left(\int_0^T Z_s^2 ds\right) < +\infty$ and*

$$\forall t \in [0,T], \qquad M_t = M_0 + \int_0^t Z_s\, dW_s \;\; a.s..$$

It follows that if X is an $\mathcal{F}_t$-measurable, square-integrable random variable, it can be written as

$$X = \mathbb{E}(X) + \int_0^T Z_s\, dW_s \;\; a.s.,$$

where (Z_t) is an adapted process such that $\mathbb{E}\left(\int_0^T Z_s^2 ds\right) < +\infty$

(Referance: [3])

Theorem A.3 (Banach Fixed Point Theorem)**.** *Consider a metric space $X = (X,d)$, where $X \neq \emptyset$. Suppose that X is complete and $\Phi : X \mapsto X$ be a contraction on X. Then Φ has precisely one fixed point.*
(Contraction mapping: Let $X = (X,d)$ be a metric space. $\Phi : X \mapsto X$ is said to be contraction mapping on X if there is a positive real number $\alpha < 1$ such that for all $(x,y) \in X$, $d(\Phi(x),\Phi(y)) \le \alpha d(x,y)$.)

(Reference Number: [5])

Theorem A.4 (Burkholder-Davis-Gundy Inequalities)**.** *Assume that M is continuous random variable and local martingale. Let $M_t^* = \max_{0\le s\le t}|M_s|$. For every $m > 0$, there exists a universal positive constants k_m, K_m (depending only on m), such that*

$$k_m \mathbb{E}\left(< M >_T^m\right) \le \mathbb{E}\left((M_T^*)^{2m}\right) \le K_m \mathbb{E}\left(< M >_T^m\right).$$

(Reference: [9])