EVANGELIDES_KONSANTINOS_01507686

by Konstantinos Evangelides

Submission date: 05-Sep-2022 08:22PM (UTC+0100)
Submission ID: 185722801
File name: EVANGELIDES_KONSTANTINOS_01507686.pdf (645.05K)
Word count: 13006
Character count: 53741

Imperial College London

Imperial College London

Department of Mathematics

Backward Stochastic Differential Equations

Author:
Konstantinos Evangelides (CID: 01507686)

Supervisor:
Dr. Camilo *Hernández*

A thesis submitted for the degree of
MSc in Mathematics and Finance, 2021-2022.

September 6, 2022

"The work contained in this thesis is my own work unless otherwise stated"

Signature and date:

September 6th 2021

Konstantinos Evangelides

Acknowledgements

I would first of all like to express my gratitude to my supervisor, Dr. Camilo *Hernández* for his ongoing engagement and continuous support throughout the completion of this project. I would like to thank him for his advice, patience and insightful suggestions and whose contribution towards this Thesis has been proven to be critical and invaluable.

Moreover, I would also like to thank all my lecturers at Imperial College London for all the knowledge and support they provided me with throughout this academic year.

Finally, I dedicate this project to my parents, whose ongoing encouragement and care gave me the opportunity to complete this masters and allowed me to pursue my goals.

Abstract

This Thesis studies the Backward Stochastic Differential Equations (BSDEs). We begin by presenting two proofs of the Martingale Represenation Theorem (MRT), the first one is a proof by Zhang and Jianfeng and the second one by Karatzas and Shreve. We then proceed to compare their approaches in proving the result and the different assumptions they make. Then, we show how the BSDE is derived from the MRT and present the first theorems of existence and uniqueness of solutions to the BSDE. We also present linear BSDEs and some applications in finance. Then, we move on to address stochastic optimal control theory through BSDEs, and in particular, we study the connection between "regular" control problems and BSDEs. Finally, we present some extensions of this theory by studying the equivalence of the above connection for "singular" control problems. The motivation is through an example of the optimal dividend problem where we derive the relevant BSDE.

Contents

Chapter 1

Introduction

1.1 Background

In mathematics, a differential equation describes the relationship between unknown functions and their derivatives. The functions usually represent physical quantities and the derivatives represent their rates of change. Differential equations have a wide range of applications such as engineering, economics and biology.

Their history begins back in 1671 when they first came into existence with the invention of calculus by Isaac Newton. During his research on "fluxional equations", what we would now call differential equations, he introduced three kinds of differential equations. Not long after Newton's research, the great mathematician and philosopher Gottfried Wilhelm Leibniz came up with a solution to linear differential equations of first order. In 1695, Jacob Bernoulli introduced an ordinary differential equation, namely the Bernoulli differential equation for which Leibniz found a solution a year later.

A huge part in the evolution of differential equations was played by the Swiss mathematician Leonard Euler, who popularized the use of power series, finding an integrating factor to derive differential equations that were integrable in finite form and the theory of linear equations of arbitrary order.

Stochastic differential equations originated in the work of Einstein and Smoluchowski through the theory of Brownian motion and their theory was further developed in the 1940s by the Japanese mathematician Kiyosi Itô who popularized the concept of the stochastic integral. In particular, SDEs and BSDEs (which we will define later) can only be understood in terms of integrals as there is no concept of a derivative. They both characterize the behaviour of a continuous time stochastic process as the sum of a Lebesgue integral and an Itô integral.

Finally, Backward Stochastic Differential Equations (BSDEs in short) were first introduced by Bismut [2] in 1973 for the linear case and were then generalized by Pardoux and Peng [12] to the Lipschitz case. The theory of BSDEs attracted a lot of attention in the last couple of years in the field of mathematical finance due to their connections with the pricing of contingent claims and optimal stochastic control problems.

1.2 Notation and definitions

Throughout this Thesis, we will be working in a finite horizon setting with final time T where $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ is a filtered probability space supporting a n-dimensional standard brownian motion W_t, where $\mathbb{F} = \{\mathcal{F}_t, 0 \leq t \leq T\}$ is the natural filtration generated by W, satisfying the usual conditions of right continuity and augmented with all the $\mathbb{P}$-null sets, so that $\mathbb{F}$ is complete. The filtration $\mathbb{F}$ can be thought of as a family of sub-σ-fields of the σ-field $\mathcal{F}$ that are ordered nondecreasingly, i.e for any $0 \leq t_1 < t_2 \leq T,\ \mathcal{F}_{t_1} \subset \mathcal{F}_{t_2}$.

Now, let's define some useful spaces of random variables or processes that we will be using throughout this thesis:

- $\mathcal{F}$ is a σ-field on Ω and $\mathcal{F}^X$ is a sub-σ-field of $\mathcal{F}$ and is called the σ-field generated by X.
- $\mathbb{H}^2_T$ is the space of all $\mathbb{F}$-progressively measurable processes $\phi : \Omega \times [0,T] \to \mathbb{R}^d$ such that $\|\phi\|_{\mathbb{H}^2} := \mathbb{E}[\int_0^T |\phi_t|^2 dt] < \infty$.
- $\mathbb{L}^0(\mathcal{F}, \mathbb{R}^d)$ is the space of all $\mathcal{F}$-measurable random variables $X : \Omega \to \mathbb{R}^d$ where X satisfies $X^{-1}(A) := \{X \in A\} := \{\omega \in \Omega : X(\omega) \in A\} \in \mathcal{F}$ for all $A \in \mathcal{B}(\mathbb{R})$.
- For $p \geq 1$, $\mathbb{L}^p(\mathcal{F}, \mathbb{R}^d)$ is the space of all $\mathcal{F}$-measurable random variables $X : \Omega \to \mathbb{R}^d$ where $X \in \mathbb{L}^0(\mathcal{F})$ such that $\|X\|^p := \mathbb{E}(|X|^p) < \infty$.
- If $X \in \mathbb{L}^0(\mathbb{F})$ we say that X is $\mathbb{F}$-progressively measurable, if the restriction of X on $[0,t]$ is $\mathcal{B}([0,t]) \times \mathcal{F}_t$-measurable for any $t \in [0,T]$. We have that $X \in \mathbb{L}^0(\mathcal{F})$ implies that X is a random variable whereas $X \in \mathbb{L}^0(\mathbb{F})$ implies it is a process.
- $\mathbb{S}^2_T$ is the space of all progressively measurable processes $\phi : \Omega \times [0,T] \to \mathbb{R}^d$ such that $\mathbb{E}[\sup_{t\in[0,T]} |\phi_t|^2] < \infty$.
- $\mathcal{K}^2$ is the set of processes (K_t) in $\mathbb{S}^2$ with non-decreasing components starting at 0.
- $\mathcal{M}^{c,loc}$ the space of continuous local martingales.

1.2.1 Martingales

Definition 1.2.1. *Any $M \in \mathbb{L}^0(\mathbb{F})$ is called a martingale if it satisfies:*

- $\mathbb{E}[|M_t|] < \infty$ *for all* $t \in [0,T]$.
- $\mathbb{E}[M_t|\mathcal{F}_s] = M_s,\ \mathbb{P}\text{-a.s.}$ *for all* $0 \leq s < t \leq T$.

For any martingale M we have that $\mathbb{E}[M_t] = \mathbb{E}[M_0]$ for all t $\in [0,T]$.

Definition 1.2.2. *For a process $M \in \mathbb{L}^0(\mathbb{F})$, we say that M is a local martingale if there exists a sequence of stopping times $(\tau_n)_{n\geq 1}$ such that:*

- $\tau_n \nearrow T$ *a.s.*
- $\tau_n < T$ *a.s. on* $T > 0$.
- $\lim_{n\to\infty} \tau_n = T$ *a.s.*
- $M_{\tau_n \wedge .}$ *is a martingale for all* $n \geq 1$.

A martingale is always a local martingale, however, if $M \in \mathbb{L}^0(\mathbb{F})$ is uniformly integrable, then M is a martingale if and only if it is a local martingale.

1.2.2 Stochastic Differential Equations

Recall that a Stochastic Differential Equation (SDE in short) takes the form:

$$X_t = X_0 + \int_0^t b_s(X_s)ds + \int_0^t \sigma_s(X_s)dW_s, \quad 0 \leq t \leq T \tag{1.1}$$

where T>0 is a known maturity date and the functions $b_t(x)$ and $\sigma_t(x)$ for $\{0 \leq t \leq T\}$ are $\mathbb{F}$-progressively measurable.

Definition 1.2.3. *A strong solution of (2.1) is an $\mathbb{F}$-progressively measurable process $X \in \mathbb{S}_T^2$ such that $\int_0^T (|b_t(X_t)| + |\sigma_t(X_t)|^2)dt < \infty$, a.s, and*

$$X_t = X_0 + \int_0^t b_s(X_s)ds + \int_0^t \sigma_s(X_s)dW_s, \quad 0 \le t \le T \ .$$

A few notes to make here:

- SDEs are the non-linear extension of the stochastic integral.
- SDEs evolve forward in time.
- X is adapted, i.e X_t is $\mathcal{F}_t$-measurable for each $t \in T$.
- $X_0 = x_0$ is a given initial value.

Chapter 2

Martingale Representation Theorem

The Martingale Representation Theorem is considered to be a fundamental part for the theory of BSDEs and SDEs in general. In this chapter, we proceed to present and analyse two separate proofs of this result. Throughout this chapter, let the process $\mathbf{W} : [0,T] \times \Omega \to \mathbb{R}$ define a standard Brownian motion.

The first proof of this result, is presented by Zhang [13] in a one-dimensional setting, for notational simplicity, but he urges the reader to extend all arguments to multidimensional cases straightforwardly. Therefore, for this part, we assume that d=1. Zhang, uses a PDE approach which we will discuss later. We first state two Lemmas that we will need later on.

Lemma 1. *Let $\eta \in C^\infty(\mathbb{R}^d)$ such that: $\eta \geq 0$, $\eta(x) = 0$ for $|x| \geq 1$ and $\int\limits_{\mathbb{R}^d} \eta(x)dx = 1$. For any Borel measurable function $g : \mathbb{R}^d \to \mathbb{R}$ with $\int\limits_{\mathbb{R}^d} |g(y)|dy < \infty$, define its smooth modifier:*

$$g_n(x) = \int\limits_{\mathbb{R}^d} g(x - ny)\eta(y)dy = n^{-d} \int\limits_{\mathbb{R}^d} g(y)\eta(\tfrac{x-y}{n})dy, \quad \text{for all } n > 0.$$

Then, $g_n \in C^\infty(\mathbb{R}^d)$ with bounded derivatives, $\lim_{n\to\infty} g_n(x) = g(x)$, for Lebesgue-a. e. *$x \in \mathbb{R}^d$ where the convergence holds for all $x \in \mathbb{R}^d$ if g is continuous and if $|g| \leq C$, then $|g_n| \leq C$.*

Lemma 2. *Let $\eta_n \in \mathbb{L}^2(\mathcal{F}_0), \sigma^n \in \mathbb{L}^2(\mathbb{F}, \mathbb{R}^d)$, and denote $\xi_n := \eta_n + \int\limits_0^T \sigma_t^n \cdot dW_t$, $n \geq 1$. Assume $\lim_{n\to\infty} \mathbb{E}[|\xi_n - \xi|^2] = 0$ for some $\xi \in \mathbb{L}^2(\mathcal{F}_T)$. Then there exists unique $\sigma \in \mathbb{L}^2(\mathbb{F}, \mathbb{R}^d)$ such that $\xi = \mathbb{E}[\xi|\mathcal{F}_0] + \int\limits_0^T \sigma_t \cdot dW_t$, and $\lim_{n\to\infty} \mathbb{E}\Big[|\eta_n - \mathbb{E}[\xi|\mathcal{F}_0]|^2 + \int\limits_0^T |\sigma_t^n - \sigma_t|^2 dt\Big] = 0$.*

Proof. We have that

$$\lim_{n\to\infty} \mathbb{E}\left[|\eta_n - \mathbb{E}[\xi|\mathcal{F}_0]|^2 + \int_0^T |\sigma_t^n - \sigma_t|^2 dt\right] \leq \lim_{n\to\infty} \mathbb{E}\left[|\eta_n - \mathbb{E}[\xi|\mathcal{F}_0]|^2 + |\int_0^T \sigma_t^n - \sigma_t dt|^2\right] \tag{2.1}$$

$$= \lim_{n\to\infty} \mathbb{E}\left[|\eta_n - \mathbb{E}[\xi|\mathcal{F}_0]|^2 + |\int_0^T \sigma_t^n - \sigma_t|^2 dW_t\right]$$

$$\leq \lim_{n\to\infty} \mathbb{E}\left[|\eta_n - \mathbb{E}[\xi|\mathcal{F}_0]| + \int_0^T \sigma_t^n - \sigma_t dW_t\right]$$

$$\leq \lim_{n\to\infty} \mathbb{E}\left[\eta_n - \mathbb{E}[\xi|\mathcal{F}_0] + \int_0^T \sigma_t^n - \sigma_t dW_t\right]$$

$$= \lim_{n\to\infty} \mathbb{E}\left[\eta_n + \int_0^T \sigma_t^n \cdot dW_t - \left[\mathbb{E}[\xi|\mathcal{F}_0] + \int_0^T \sigma_t \cdot dW_t\right]\right.$$

$$= \lim_{n\to\infty} \mathbb{E}[\xi_n - \xi]$$

$$\leq \lim_{n\to\infty} \mathbb{E}[|\xi_n - \xi|^2] = 0$$

□

Now, we state the result, which appears as Theorem 2.5.2 in [13].

2.1 The Martingale Representation Theorem

Theorem 2.1. *Let $\xi \in \mathbb{L}^2(\mathcal{F}_T^W)$, i.e ξ is an $\mathcal{F}_T^W$-measurable random variable in the σ-field generated by the Brownian motion W. Then there exists a unique process $\sigma \in \mathbb{L}^2(\mathbb{F}, \mathbb{R}^d)$ such that:*

$$\xi = \mathbb{E}[\xi] + \int_0^T \sigma_t \cdot dW_t. \tag{2.2}$$

Consequently, for any $\mathcal{F}^W$-martingale M such that $\mathbb{E}[|M_T|^2] < \infty$, there exists a unique process $\sigma \in \mathbb{L}^2(\mathbb{F}, \mathbb{R}^d)$ such that:

$$M_t = M_0 + \int_0^t \sigma_s \cdot dW_s. \tag{2.3}$$

Proof. We first prove that (2.3) follows directly from (2.2). Indeed, by (2.2), for any $\mathbb{F}^W$-martingale such that $\mathbb{E}[|M_T|^2] < \infty$ there exists a unique process $\sigma \in \mathbb{L}^2(\mathbb{F}, \mathbb{R}^d)$ such that:

$$M_T = \mathbb{E}[M_T] + \int_0^T \sigma_t \cdot dW_t.$$

If we denote

$$\tilde{M}_t := \mathbb{E}[M_T] + \int_0^t \sigma_s \cdot dW_s$$

then by Lemma 2.2.3 in [13] and by the fact that $\mathbb{E}[M_T]$ is a constant we have that $\tilde{M}_t$ is an $\mathbb{F}^W$-martingale and $\tilde{M}_T = M_T$. Now, using the properties of martingales we have that

$$M_t = \mathbb{E}[M_T|\mathcal{F}_t^W] = \mathbb{E}[\tilde{M}_T|\mathcal{F}_t^W] = \tilde{M}_t \qquad \text{and} \qquad \mathbb{E}[M_T] = M_0 = \tilde{M}_0$$

Therefore, we have

$$M_t = \tilde{M}_t = \mathbb{E}[M_T] + \int_0^t \sigma_s \cdot dW_s = M_0 + \int_0^t \sigma_s \cdot dW_s$$

which implies (2.3) immediately.

Uniqueness of σ:

Assume there is another $\tilde{\sigma} \in \mathbb{L}^2(\mathbb{F}, \mathbb{R}^d)$ satisfying (2.2). We then have

$$\int_0^T \sigma_t \cdot dW_t = \int_0^T \tilde{\sigma}_t \cdot dW_t$$

$$\int_0^T (\sigma_t - \tilde{\sigma}_t) \cdot dW_t = 0$$

Squaring both sides and taking expectations we get

$$\mathbb{E}\Big[\int_0^T |\sigma_t - \tilde{\sigma}_t|^2 dt\Big] = 0.$$

Since the integrand is always positive and the above expression is equal to zero then it must be that $\sigma_t - \tilde{\sigma}_t = 0$, that is $\sigma_t = \tilde{\sigma}_t, \quad dt \times d\mathbb{P}$- a.s.

Existence in (2.2):

We know need to prove the existence in (2.2). In this part of the proof, Zhang considers different classes of $\mathcal{F}$-measurable random variables and proves the result for each one of them. He starts by considering the most "restrictive" or "smallest" class which is the one of all bounded and twice-continuously differentiable functions that take values in $\mathbb{R}$ and then uses the result for this class to prove the existence in the next "larger" class and so on until he gets to the general case.

Step 1. Assume that $\xi = g(B_T), g \in C_b^2(\mathbb{R})$. Now define

$$u(t,x) = \mathbb{E}\Big[g(x + B_{T-t})\Big] = \int_{\mathbb{R}} g(y)p(T-t, y-x)dy, \quad where \;\; p(t,x) := \frac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}} \tag{2.4}$$

We first compute the partial derivatives of the function $p(t,x)$:

$$\partial_t p(t,x) = \tfrac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2t}} \Big[-\tfrac{1}{2} t^{-\frac{3}{2}} + -\tfrac{x^2}{2} t^{-\frac{5}{2}} \Big]$$
$$\partial_x p(t,x) = \tfrac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}} (-\tfrac{x}{t})$$
$$\partial_{xx} p(t,x) = \tfrac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}} \Big[\tfrac{x^2}{t^2} - \tfrac{1}{t}\Big]$$

Then, we have

$$\partial_t p(t,x) - \tfrac{1}{2}\partial_{xx} p(t,x) = 0$$

We also note that all the partial derivatives above exist and are continuous in $[0,T] \times \mathbb{R}$ since the exponential function is continuous and the terms in the brackets are continuous so the product of continuous functions is continuous. Therefore, $u(t,x)$ is also continuous by the same argument. Additionally, since g is bounded and $p(t,x)$ is a decreasing function both as t goes to infinity and as x goes to infinity, thus it is also bounded, we get that $u(t,x)$ is bounded for $[0,T] \times \mathbb{R}$.

We also note that

$$\partial_t u(t,x) = \int_{\mathbb{R}} g(y)\partial_t p(T-t, y-x)dy$$

$$\partial_x u(t,x) = \int_{\mathbb{R}} g(y)\partial_x p(T-t, y-x)dy$$

and

$$\partial_t p(T-t, y-x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(y-x)^2}{2(T-t)}} \left[\frac{1}{2}(T-t)^{-\frac{3}{2}} - \frac{(y-x)^2}{2}(T-t)^{-\frac{5}{2}}\right]$$
$$\partial_x p(T-t, y-x) = \frac{1}{\sqrt{2\pi(T-t)}} e^{-\frac{(y-x)^2}{2(T-t)}} \left(\frac{(y-x)}{(T-t)}\right)$$
$$\partial_{xx} p(T-t, y-x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(y-x)^2}{2(T-t)}} \left[\frac{(y-x)^2}{(T-t)^{\frac{5}{2}}} - \frac{1}{(T-t)^{\frac{3}{2}}}\right]$$

So one can easily see that

$$\partial_t u(t,x) + \frac{1}{2}\partial_{xx} u(t,x) = 0, \quad u(T,x) = g(x) \tag{2.5}$$

since $u(T,x) = \mathbb{E}\Big[g(x+B_0)\Big] = \mathbb{E}\Big[g(x)\Big] = g(x)$.

Now define

$$M_t := u(t, W_t) \;\; and \;\; \sigma_t := \partial_x u(t, W_t). \tag{2.6}$$

Then by Itô's formula we get

$$du(t,W_t) = \partial_t u(t,W_t)dt + \partial_x u(t,W_t)dW_t + \tfrac{1}{2}\partial_{xx} u(t,W_t)dt = \partial_x u(t,W_t)dW_t = \sigma_t dW_t$$

where we have used that the quadratic variation of a brownian motion is $d < W >_t = dt$ and equation (2.5).

Integrating the above and using equations (2.4) and (2.5) we get

$$g(W_T) = u(T, W_T) = u(0,0) + \int_0^T \sigma_t dW_t = \mathbb{E}[g(W_T)] + \int_0^T \sigma_t dW_t.$$

Since $\partial_x u(t, W_t)$ is bounded then by (2.6) we have that $\sigma \in \mathbb{L}^2(\mathbb{F}^W)$ and thus (2.2) holds.

Step 2. Now, assume $\xi = g(W_T)$ where $g : \mathbb{R} \to \mathbb{R}$ is any Borel measurable and bounded function. For this part we will need Lemmas (1) and (2).

Now, using Lemma (1), we let g_n be a smooth modifier of g and we have that $g_n \in C_b^2(\mathbb{R}), \lim_{n\to\infty} g_n(x) = g(x)$ for dx a.e. x and $|g_n| \leq C$ for all n. Since W_T has density, then W_T does not belong to a Lebesque null set a.s., so that $\lim_{n\to\infty} g_n(W_T(\omega)) = g(W_T(\omega))$ a.s. and using the Dominated Convergence Theorem (A.1) we have that $\lim_{n\to\infty} \mathbb{E}[|g_n(W_T(\omega)) - g(W_T(\omega))|^2] = 0$. Since $g_n \in C_b^2(\mathbb{R})$ for all $n > 0$, then by Step 1 we have that for each n there exists $\sigma^n \in \mathbb{L}^2(\mathbb{F})$ such that $g_n(W_T) = \mathbb{E}[g_n(W_T(\omega))] + \int_0^T \sigma_t^n dW_t$. Finally, using Lemma (2) we see that (2.2) holds true.

Step 3. For this step, we consider a partition $0 < t_1 < t_2 < .... < t_{n-1} < t_n \leq T$ of the interval $[0, T]$ and we let $\xi = g(W_{t_1}, ..., W_{t_n})$ where $g : \mathbb{R}^n \to \mathbb{R}$ is Borel measurable and bounded. Then denote $g_n(x_1, ..., x_n) := g(x_1, ..., x_n)$ and apply Step 2 on $[t_{n-1}, t_n]$. Therefore, we know that there exists $\sigma^n \in \mathbb{L}^2(\mathbb{F}^W)$ such that

$$g_n(W_{t_1}, ..., W_{t_n}) = \mathbb{E}\Big[g(W_{t_1}, ..., W_{t_n})|\mathcal{F}^W_{t_{n-1}}\Big] + \int\limits_{t_{n-1}}^{t_n} \sigma^n_t dW_t =$$

$$g_{n-1}(W_{t_1}, ..., W_{t_{n-1}}) + \int\limits_{t_{n-1}}^{t_n} \sigma^n_t dW_t.$$

Now, since the Brownian motion has independent increment we have that

$$g_{n-1}(x_1, ..., x_{n-1}) = \mathbb{E}\Big[g_n(x_1, ..., x_{n-1}, x_{n-1} + W_{t_{n-1}, t_n})\Big]$$

is also Borel measurable and bounded and repeating the same argument for all other intervals backwardly we get that

$$g_{i+1}(W_{t_1}, ..., W_{t_{i+1}}) = g_i(W_{t_1}, ..., W_{t_i}) + \int\limits_{t_i}^{t_{i+1}} \sigma^{i+1}_t dW_t$$

where

$$g_i(x_1, ..., x_i) = \mathbb{E}\Big[g_n(x_1, ..., x_i, x_i + W_{t_i, t_{i+1}})\Big]$$

Now that we have σ^i for all $1 \leq i \leq n$ we define

$$\sigma := \textstyle\sum_{i=1}^n \sigma^i \mathbb{I}_{[t_{i-1}, t_i)}.$$

Then, since all $\sigma^i \in \mathbb{L}^2(\mathbb{F}^W)$ we have that their sum $\sigma \in \mathbb{L}^2(\mathbb{F})$ and (2.2) follows.

Step 4. Here, we go one step further and we consider partitions of the interval $[0, T]$ of varying size which in turn give us an increasing sequence of σ-fields generated by the Brownian motion. We assume $\xi \in \mathbb{L}^\infty(\mathcal{F}^W_T)$ and denote for each n, $t^n_i := \frac{iT}{2^n}, i = 0, ..., 2^n$. We let $\mathcal{F}^n_T$ be the σ-field generated by $\{W_{t^n_i}, 0 \leq i \leq 2^n\}$ and define $\xi_n := \mathbb{E}[\xi|\mathcal{F}^n_T]$. At this point we notice that as n increases we have larger and larger generated σ-fields. Using the Doob-Dynkin lemma (4) we get

$$\xi_n = g_n(W_{t^n_1}, ..., W_{t^n_{2^n}}) \qquad \text{for some Borel measurable function } g_n.$$

Since ξ is bounded then ξ_n is bounded and so is g_n. Using Step 3 we have

$$\xi_n = \mathbb{E}[\xi_n] + \int\limits_0^T \sigma^n_t dW_t \quad \text{for some } \sigma^n \in \mathbb{L}^2(\mathbb{F}^W).$$

Since W is continuous we have that $\mathcal{F}^W_T = \bigvee_n \mathcal{F}^n_T$ is the σ-field generated by $\bigcup_n \mathcal{F}^n_T$. Also, since the random variable ξ belongs to the sigma field generated by the Brownian motion up to time T we know that $\mathbb{E}[\xi|\mathcal{F}^W_T] = \xi$. We now have an increasing sequence of σ-fields $\mathcal{F}^n_T \subset \mathcal{F}^W_T$ and $\xi \in \mathbb{L}^1(\mathcal{F}^W_T)$ which imply that $\mathbb{E}[\xi|\mathcal{F}^n_T] \to \mathbb{E}[\xi|\mathcal{F}^W_T]$ in $\mathbb{L}^1(\mathcal{F}^W_T)$. Using this together with the Dominated Convergence (A.1) Theorem we have

$$\lim_{n\to\infty} \mathbb{E}\Big[|\xi_n - \xi|^2\Big] = 0.$$

Now (2.2) follows using Lemma (2) again.

Step 5. In the most general case, we let $\xi_n := (-n) \vee \xi \wedge n$. This means that $|\xi_n| \leq n$ and that $\xi_n := \xi$ only if $\xi \in [-n, n]$ which implies that as n tends to infinity $\xi_n \to \xi$ and that $|\xi_n| \leq |\xi|$. By Step 4 we have that there exists $\sigma^n \in \mathbb{L}^2(\mathbb{F}^{\mathbb{W}})$ such that

$$\xi_n = \mathbb{E}[\xi_n] + \int_0^T \sigma_t^n dW_t$$

By the Dominated Convergence Theorem (A.1) we then have

$$\lim_{n\to\infty} \mathbb{E}\Big[|\xi_n - \xi|^2\Big] = 0.$$

and (2.2) follows using Lemma (2) again.

□

Remark 2.2. *Note that, in Theorem (2.1), we have assumed that M is an $\mathbb{F}^W$-martingale,i.e. it is a martingale with respect to a filtration generated by the Brownian motion W. However, the result holds true without this condition only if the measure generated by the quadratic variation process $d < M >_t$ is equivalent to the Lebesgue measure dt. If this is not the case, but at least $d < M >_t$ is absolutely continuous with respect to the Lebesgue measure, then the result still holds true by extending the probability space. To see that without absolute continuity the result does not hold we could consider the following example:*

Let $\mathcal{C}_t^i$ be copies of the Cantor function on the unit interval $[0,1)$, for $1 \leq i < \infty$ and define the extension to the positive halfline by

$$K_t = \sum_{i=0}^{\infty} i + \mathcal{C}_{t-i}^i \mathbb{I}_{[I,I+1)}(T)$$

Now, define M by $M_t = W_{K_t}$ where W is a Brownian motion which has quadratic variation $< M >_t = K_t$. However, applying the Martingale Representation theorem we could write M as $M_t = \int_0^t \sigma_s dW_s$ so that $< M_t > = \int_0^t (\sigma_s)^2 ds$. This is a contradiction since the latter representation of the quadratic variation of M is an absolutely continuous function but K is not.

2.2 Representation of Continuous Local Martingales as Stochastic Integrals with Respect to Brownian Motion

The second proof, presented by Karatzas Ioannis and Shreve Steven in [8] follows a pure probabilistic argument where the two authors begin with a d-dimensional vector and break it down into d separate, one-dimensional cases. The result appears as theorem 4.2 in [8] and is as follows:

Theorem 2.3. *Suppose $M = \{M_t = (M_t^{(1)}, ..., M_t^{(d)}), \mathcal{F}_t; 0 \leq t < \infty\}$ is defined on $(\Omega, \mathcal{F}, \mathbb{P})$ with $M^{(i)} \in \mathcal{M}^{c,loc}, 1 \leq i \leq d$. Suppose also that for $1 \leq i, j \leq d$, the cross-variation $< M^{(i)}, M^{(j)} >_t (\omega)$ is an absolutely continuous function of t for $\mathbb{P}$-almost every ω. Then there is an extension $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ of $(\Omega, \mathcal{F}, \mathbb{P})$ on which is defined a d-dimensional Brownian motion $W = \{W_t = (W_t^{(1)}, ..., W_t^{(d)}), \tilde{\mathcal{F}}_t; 0 \leq t < \infty\}$, and a matrix $X = \{(X_t^{(i,k)})_{i,k=1}^d, \tilde{\mathcal{F}}_t; 0 \leq t < \infty\}$ of measurable,adapted processes with*

$$\tilde{\mathbb{P}}\left[\int_0^t (X_s^{(i,k)})^2 ds < \infty\right] = 1; \quad 1 \leq i, k \leq d; \quad 0 \leq t < \infty, \tag{2.7}$$

such that we have, $\tilde{\mathbb{P}}$-a.s., the representations

$$M_t^{(i)} = \sum_{k=1}^{d} \int_0^t X_s^{(i,k)} dW_s^{(k)}; \quad 1 \le i \le d; \quad 0 \le t < \infty, \tag{2.8}$$

$$< M^{(i)}, M^{(j)} >_t = \sum_{k=1}^{d} \int_0^t X_s^{(i,k)} X_s^{(j,k)} ds; \quad 1 \le i,j \le d; \quad 0 \le t < \infty. \tag{2.9}$$

Proof. We begin by defining

$$z_t^{i,j} = z_t^{j,i} = \frac{d}{dt} < M^{(i)}, M^{(j)} >_t = \lim_{n\to\infty} n[< M^{(i)}, M^{(j)} >_t - < M^{(i)}, M^{(j)} >_{(t-(\frac{1}{n}))^+}] \tag{2.10}$$

We know that the cross variation between two variables is symmetric and therefore their derivative is symmetric. Additionally, we know that the point-wise limit of measurable functions is measurable and since any progressively measurable function is measurable then we have that the point-wise limit of progressively measurable functions is progressively measurable.Therefore, the matrix-valued process $Z = \{Z_t = (z_t^{i,j})_{i,j=1}^d, \mathcal{F}_t; 0 \le t < \infty\}$ is symmetric and progressively measurable. Now, for any vector $\alpha = (\alpha_1, ..., \alpha_d) \in \mathbb{R}^d$, we have

$$\textstyle\sum_{i=1}^d \sum_{j=1}^d \alpha_i z_t^{i,j} \alpha_j = \frac{d}{dt} \Big\langle \sum_{i=1}^d \alpha_i M^{(i)} \Big\rangle_t \ge 0,$$

which shows that Z_t is positive-semidefinite for Lebesque-almost every t, $\mathbb{P}$- a.s. The last equality comes from the fact that we work with a continuous local martingale with associated increasing process $< M >$, so that its derivative is non-negative.
Now, any symmetric, positive-semidefinite matrix Z can be diagonalised by an orthogonal matrix Q such that $Q^{-1}ZQ = \Lambda$ with $Q^{-1} = Q^T$, and Λ a diagonal matrix with the non-negative eigenvalues of Z as its diagonal elements. There are many algorithms which compute Q and Λ from Z, such that Q and Λ are Borel-measurable functions of Z. Therefore, since Z is a symmetric, progressively measurable and positive-semidefinite matrix process then there exists processes $\{Q_t(\omega) = (q_t^{i,j}(\omega))_{i,j=1}^d; \mathcal{F}_t; 0 \le t < \infty\}$ and $\{\Lambda_t(\omega) = (\delta_{ij}\lambda_t^i(\omega))_{i,j=1}^d, \mathcal{F}_t; 0 \le t < \infty\}$ such that for Lebesgue-almost every t, we have

$$\sum_{k=1}^{d} q_t^{k,i} q_t^{k,j} = \sum_{k=1}^{d} q_t^{i,k} q_t^{j,k} = \delta_{ij}; \quad 0 \le i,j \le d, \tag{2.11}$$

$$\sum_{k=1}^{d} \sum_{l=1}^{d} q_t^{k,i} z_t^{k,l} q_t^{l,j} = \delta_{ij} \lambda_t^i \ge 0; \quad 0 \le i,j \le d, \tag{2.12}$$

$\mathbb{P}$-a.s.From (2.11) with i=j, we have that since δ_{ij} is the Kronecker delta and takes the value of 1 only if i=j and otherwise zero, we see that $(q_t^{k,i})^2 \le 1$, so

$$\int_0^t (q_t^{k,i})^2 d < M^{(k)} >_s \le \int_0^t d < M^{(k)} >_s = < M^{(k)} >_t \ < \infty$$

The last inequality follows from Theorem 6.13 of [7].We can now define continuous local martingales by

$$N_t^{(i)} := \sum_{k=1}^{d} \int_0^t q_s^{k,i} dM_s^{(k)}; \quad 1 \le i \le d, \quad 0 \le t < \infty. \tag{2.13}$$

Using equations (2.10) and (2.12) we have

$$< N^{(i)}, N^{(j)} >_t = \sum_{k=1}^{d}\sum_{l=1}^{d}\int\limits_0^t q_s^{k,i} q_s^{l,j} d < M^{(k)}, M^{(l)} >_s \tag{2.14}$$

$$= \sum_{k=1}^{d}\sum_{l=1}^{d}\int\limits_0^t q_s^{k,i} z_s^{k,l} q_s^{l,j} ds$$

$$= \delta_{ij}\int\limits_0^t \lambda_s^i ds$$

where again when i=j we have $\delta_{ij} = 1$ and

$$\int\limits_0^t \lambda_s^i ds =< N^{(i)} >_t < \infty \tag{2.15}$$

At this point we have constructed a vector of local martingales $N = \{(N_t^{(1)}, ..., N_t^{(d)}), \mathcal{F}_t; 0 \leq t < \infty\}$ and we will use the extended probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ on which is defined a d-dimensional Brownian motion $W = \{W_t = (W_t^{(1)}, ..., W_t^{(d)}), \tilde{\mathcal{F}}_t; 0 \leq t < \infty\}$ which is independent of N to represent this vector as a vector of stochastic integrals. At this point we should note that our vector of local martingales is not necessarily continuous so we can use (2.15) to show that

$$\int\limits_0^t \mathbb{I}_{\{\lambda_s^i > 0\}} \frac{1}{\lambda_s^i} d < N_s^{(i)} >_s = \int\limits_0^t \mathbb{I}_{\{\lambda_s^i > 0\}} ds \leq t.$$

and define continuous, local martingales

$$B_t^{(i)} \triangleq \int\limits_0^t \mathbb{I}_{\{\lambda_s^i > 0\}} \frac{1}{\sqrt{\lambda_s^i}} dN_s^{(i)} + \int\limits_0^t \mathbb{I}_{\{\lambda_s^i = 0\}} dW_s^{(i)}; \;\; 1 \leq i \leq d. \tag{2.16}$$

Now using (2.14) and the fact that $< N^{(i)}, N^{(j)} >= 0$ when $i \neq j$ we get

$$< B^{(i)}, B^{(j)} >_t = \delta_{ij} t, \quad 1 \leq i, j \leq d; \quad 0 \leq t < \infty,$$

so, according to P.Lévy [10], $B = \{(B_t^{(1)}, ..., B_t^{(d)}), \tilde{\mathcal{F}}_t; 0 \leq t < \infty\}$ is a d-dimensional Brownian motion.

Now using again (2.14) we have that $< N^{(i)}, N^{(i)} >_t =< N^{(i)} >_t = \int\limits_0^t \lambda_s^i ds$ which implies that $d < N^{(i)} >_t = \lambda_s^i ds$ we have that the martingale $\int\limits_0^t \mathbb{I}_{\{\lambda_s^i = 0\}} dN_s^{(i)}$ has quadratic variation which equals zero, i.e

$$\int\limits_0^t \mathbb{I}_{\{\lambda_s^i = 0\}} d < N_s^{(i)} >= \int\limits_0^t \mathbb{I}_{\{\lambda_s^i = 0\}} \lambda_s^i ds = 0$$

so that

$$\int\limits_0^t \sqrt{\lambda_s^i} dB_s^{(i)} = \int\limits_0^t \mathbb{I}_{\{\lambda_s^i > 0\}} dN_s^{(i)} = N_t^{(i)} \quad 1 \leq i \leq d; \quad 0 \leq t < \infty \tag{2.17}$$

To first prove (2.7) we note that for $1 \leq i, k \leq d$, and using (2.15) we have

$$\int\limits_0^t (q_s^{i,k})^2 \lambda_s^k ds \leq \int\limits_0^t \lambda_s^k ds < \infty, \quad 0 \leq t < \infty$$

We have now constructed a stochastic integral representation (2.17) for N in terms of the d-dimensional Brownian motion B so we need to invert the rotation of coordinates (2.13) to obtain a representation for M. This is the point where Karatzas and Shreve reduce the problem from a d-dimensional problem into d separate, one-dimensional cases. We let $X_t^{(i,k)} = q_t^{i,k}\sqrt{\lambda_t^k}$ in the above inequality at the bottom of page 17 to show that (2.7) holds. Moreover, using (2.17), (2.13) and (2.11) we have that

$$\begin{aligned}\sum_{k=1}^{d}\int_0^t X_s^{(i,k)}dB_s^{(k)} &= \sum_{k=1}^{d}\int_0^t q_s^{i,k}dN_s^{(k)} \\ &= \sum_{j=1}^{d}\sum_{k=1}^{d}\int_0^t q_s^{i,k}q_s^{j,k}dM_s^{(j)} \\ &= \sum_{j=1}^{d}\delta_{ij}\int_0^t dM_s^{(j)} = M_t^{(i)}\end{aligned} \tag{2.18}$$

which proves (2.8) and using rules of cross variation and that $< B >_t = t$ we can prove (2.9) using (2.8). □

Remark 2.4. *We notice that in both proofs, even if $\mathbb{P}$ is completed it may still not satisfy the usual hypothesis of right continuity and augmentation and since the Brownian motion W is continuous a.s., we need to extend it to a new probability space which includes the $\mathbb{P}$-null sets. However, in Karatza's and Shreve's proof, if for $\mathbb{P}$-a.e. $\omega \in \Omega$, the matrix-valued process $Z = \{Z_t = (z_t^{i,j})_{i,j=1}^d, \mathcal{F}_t; 0 \le t < \infty\}$ has constant rank r such that $1 \le r \le d$, for Lebesgue-almost every t, we can choose an r-dimensional Brownian motion W in (2.8) which allow us not to extend the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. This can be done by taking $\lambda_t^1, ..., \lambda_t^r$ to be the r strictly positive eigenvalues of Z_t (we know from Corollary 13 of [11] that If A is an $n\times n$ real and symmetric matrix, then rank(A) = the total number of nonzero eigenvalues of A and together with the fact that a semi-positive and symmetric matrix has non-negative eigenvalues then if Z has rank r then it must have r strictly positive eigenvalues). Then we replace (2.16) by $B_t^{(i)} = \int_0^t \frac{1}{\sqrt{\lambda_s^i}}dN_s^{(i)}$ for $1 \le r \le r$. Now, using (2.15) we have that for all $r+1 \le i \le d, 0 \le t < \infty$ (since $\lambda_t^{r+1}, ..., \lambda_t^d$ are all zero) $N_t^{(i)} = 0$ so that (2.18) becomes*

$$\sum_{k=1}^{r}\int_0^t X_s^{(i,k)}dB_s^{(k)} = \sum_{k=1}^{d}\int_0^t q_s^{i,k}dN_s^{(k)} = M_t^{(i)}, \quad 1 \le i \le d.$$

Since our new notation for $B_t^{(i)}$, $1 \le i \le r$ does not involve the Brownian motion W, there is no need to extend the original probability space.

Remark 2.5. *Theorem (2.3) provides a more accurate representation of continuous martingales with respect to a Brownian motion W than Theorem (2.1). The reason is that in Theorem (2.3), Karatzas and Shreve begin with the class of continuous local martingales, whereas Zhang provides a proof only for martingales. Therefore, as the class of martingales is included in the class of local martingales there might be some rare situations that the proof provided by Zhang does not hold true for continuous local martingales that are not martingales.*

Chapter 3

Applications of Martingale Representation Theorem

As mentioned in the previous chapter, in order to apply the result of the Martingale Representation Theorem we need to assume that $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ is a filtered probability space supporting a n-dimensional standard brownian motion W_t, where $\mathbb{F} = \{\mathcal{F}_t, t \geq 0\}$ is the natural filtration generated by W_t, satisfying the usual conditions of right continuity and augmented with all the $\mathbb{P}$-null sets, so that $\mathbb{F}$ is complete. Now let $\xi \in \mathbb{L}^2(\mathcal{F}_T^W)$,i.e ξ is an $\mathcal{F}_T$-measurable and square integrable and consider the martingale $M_t = \mathbb{E}[\xi|\mathcal{F}_t]$. Note that $M_T = \xi$ since ξ is $\mathcal{F}_T$-measurable. Then by the results of the previous chapter there exists a unique process $Z \in \mathbb{L}^2(\mathbb{F}, \mathbb{R}^d)$ such that:

$$M_t = M_0 + \int_0^t Z_s \cdot dW_s = \mathbb{E}[\xi] + \int_0^t Z_s \cdot dW_s \tag{3.1}$$

and

$$M_T = M_0 + \int_0^T Z_s \cdot dW_s \tag{3.2}$$

Combining (3.1) and (3.2) we have:

$$M_t - M_T = -\int_t^T Z_s \cdot dW_s, \quad \textit{or equivalently}, \quad M_t = \xi - \int_t^T Z_s \cdot dW_s \tag{3.3}$$

Differentiating both sides of (3.3) we get

$$dM_t = Z_s dW_s, \qquad M_T = \xi \tag{3.4}$$

Equation (3.4) is a linear SDE with terminal condition $M_T = \xi$ and is what we call Backward Stochastic Differential Equation (BSDE).While SDEs are the non-linear extension of the stochastic integral and evolve forward in time, BSDEs are the non-linear extension of the Martingale Representation Theorem and they evolve backward in time. Now, the question is whether there is a $\mathbb{F}$- adapted process that satisfy the dynamics of (3.4).

In this chapter, we should consider the BSDE

$$dY_t = -f(t, Y_t, Z_t)dt + Z_t dW_t, \quad Y_T = \xi \tag{3.5}$$

or equivalently, re-writing the BSDE (3.5) in the integrated form:

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s)ds - \int_t^T Z_s dW_s, \quad t \leq T \tag{3.6}$$

where:

- The map $f : (\omega, s, y, z) \to f(\omega, s, y, z)$ which is measurable with respect to $\mathcal{P} \times \mathcal{B}(\mathbb{R}^n) \times \mathcal{B}(\mathbb{R}^{n\times d})$ where $\mathcal{P}, \mathcal{B}(\mathbb{R}^n), \mathcal{B}(\mathbb{R}^{n\times d})$ are respectively the predictable σ-algebra over $\Omega \times [0, T]$, the Borel σ-algebra on $\mathbb{R}^n$ and the Borel σ-algebra on $\mathbb{R}^{n\times d}$.

- f is uniformly Lipschitz continuous in (y, z), i.e there exists L such that

$$|f(\omega, t, y_1, z_1) - f(\omega, t, y_2, z_2)| \leq L(|y_1 - y_2| + |z_1 - z_2|) \quad \forall (y_1, z_1), (y_2, z_2) \quad dt \times d\mathbb{P}\text{- a.s.}$$

- ξ is an $\mathcal{F}_T$-measurable random variable with values in $\mathbb{R}^n$ and is called the terminal value and $f^0 := f(0, 0) \in \mathbb{L}^{1,2}(\mathbb{F}, \mathbb{R}^n)$.

Definition 3.0.1. *A solution to the BSDE (3.6) is a pair* $(Y, Z) \in \mathbb{S}^2_T(\mathbb{R}^n) \times \mathbb{H}^2_T(\mathbb{R}^{n\times d})$ *of* $\mathbb{F}$*-adapted processes such that (3.6) holds a.s..*

Remark 3.1. *From (3.4) and as we will see more clearly later on, the process* Z_t *is essentially the derivative of* M *with respect to the Brownian motion* W *so it is uniquely determined by* M *and* W*. At this point, it is also important to note that* Z *is needed to ensure the* $\mathbb{F}$*-measurability of* M*. Indeed, consider the following ordinary differential equation (ODE)*

$$dY_t = 0, \qquad 0 \leq t \leq T \tag{3.7}$$

where the terminal time T *is known. Suppose, for instance, that for every* $\xi \in \mathbb{R}$ *we have* $Y(0) = \xi$ *or* $Y(T) = \xi$*, such that (3.7) has a unique solution* $Y(t) = \xi$*. This is true because we have considered equation (3.7) to be an ODE. However, if we consider (3.7) as a SDE it makes a huge difference in the solution on whether we specify* $Y(0)$ *or* $Y(T)$*. The reason is that, since we are now in a filtered complete probability space* $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ *supporting a n-dimensional standard brownian motion* W_t*, where* $\mathbb{F} = \{\mathcal{F}_t, t \geq 0\}$ *is the natural filtration generated by* W_t*, we should note that the solution to (3.7) must be adapted to the filtration* $\mathbb{F}$*. To see this, consider the SDE*

$$dY_t = \sigma_t(Y_t)dB_t, \quad \text{with terminal condition} \quad Y_T = \xi \tag{3.8}$$

The problem here is that (3.8) in general has no $\mathbb{F}$*-measurable solution* Y*. For example, if we take* $\sigma = 0$*, then equation (3.8) becomes an ODE with unique solution* $Y(t) = \xi$ *which is not* $\mathbb{F}$*-measurable, i.e it is not adapted to the filtration* $\mathbb{F}$ *unless* ξ *is constant. In order to tackle this problem and to ensure the adaptability of the solution to* $\mathbb{F}$ *is by reformulating (3.8) as in (3.4), and now we are seeking two processes* (Y, Z)*.*

3.1 Linear BSDEs

In this section we study the case where f is linear and more specifically the one dimensional case where $n = 1$. Therefore, we should consider the general linear BSDE:

$$Y_t = \xi + \int_t^T [Y_s\alpha_s + Z_s\beta_s + f^0_s]ds - \int_t^T Z_s dB_s \tag{3.9}$$

Proposition 1. *Assume* $\xi \in \mathbb{L}^2(\mathcal{F}_T, \mathbb{R})$*, and* $f^0 \in \mathbb{L}^{1,2}(\mathbb{F}, \mathbb{R})$*. Let* (α, β) *be a bounded* $(\mathbb{R}, \mathbb{R}^d)$*-valued progressively measurable process. If* (Y, Z) *satisfy the linear BSDE (3.9), then the pair* (Y, Z) *is the unique solution and* Y *is given by*

$$Y_t = \mathbb{E}\left[\xi\Gamma_{t,T} + \int_t^T \Gamma_{t,s} f^0_s ds \,\middle|\, \mathcal{F}_t\right], \tag{3.10}$$

where $(\Gamma_{t,s})_{s\geq t}$ *is defined as*

$$\Gamma_{t,s} = \int_t^s \Gamma_{t,r}(\alpha_r dr + \beta_r dB_r); \quad \Gamma_{t,t} = 1 \tag{3.11}$$

where $\Gamma_{t,s}\Gamma_{s,u} = \Gamma_{t,u}$ *for all* $t \leq s \leq u$.

Remark 3.2. *From (3.10) we can see that if both* ξ *and* f *are non-negative then* Y_t *is non-negative as well. Additionally, if* $Y_t = 0$ *on any set* $G \in \mathcal{F}_t$ *then it follows that* $Y_s = 0, \xi = 0$ *and* $Z_s = 0, f_s^0 = 0$ $\mathbb{P}$-a.s *for all* $s \geq t$ *on* G.

3.2 Well-Posedness of BSDEs

Further to their study on BSDEs, Pardoux and Peng gave the first existence and uniqueness results for n-dimensional BSDEs [9]. Although there are a lot of proofs for the existence and uniqueness of the solution of BSDEs, the conditions under which a solution exists depends entirely on the form of the generator f. In this section, we will use the most general theorem which assumes that the generator f is a function that depends on (Y, Z) and satisfies the Lipschitz condition for (Y, Z). We first provide the following lemma which will prove useful when proving the uniqueness of the solution.

Lemma 3 (Gronwall's Lemma)**.** *Let* $g(s)$ *be an integrable and non-negative function and* $f(s)$ *be a measurable non-negative function both taking values in* $[0, T]$. *Assume also that* $T > 0$ *and* $L \geq 0$. *If*

$$f(t) \leq L + \int_0^t g(s)f(s)ds$$

then

$$f(t) \leq L \exp\left(\int_0^t g(s)ds\right) \qquad \forall t \in [0, T].$$

Theorem 3.3. *Assume that* $\mathbb{F} = \mathbb{F}^W$ *where* W *is a Brownian motion. Then, the BSDE (3.6) has a unique solution* $(Y, Z) \in \mathbb{L}^2(\mathbb{F}, \mathbb{R}^n) \times \mathbb{L}^2(\mathbb{F}, \mathbb{R}^{n\times d})$.

Proof. Uniqueness of (Y, Z):

Assume that (Y^1, Z^1) and (Y^2, Z^2) are two solutions of (3.6). By Itô's formula we have

$$d|Y_t|^2 = 2Y_t dY_t + |Z_t|^2 dt = -2Y_t f_t(t, Y_t, Z_t)dt + 2Y_t Z_t dW_t + |Z_t|^2 dt \tag{3.12}$$

which leads to

$$|Y_t|^2 = |\xi|^2 + 2\int_t^T Y_s f_s(s, Y_s, Z_s)ds + 2\int_t^T Y_s Z_s dW_s - \int_t^T |Z_s|^2 ds \tag{3.13}$$

Now, using (3.13) and Itô's formula again we have

$$|Y_t^1 - Y_t^2|^2 + \int_t^T |Z_s^1 - Z_s^2|^2 ds = -2\int_t^T \left(Y_s^1 - Y_s^2, f_s(s, Y_s^1, Z_s^1) - f_s(s, Y_s^2, Z_s^2)\right) ds - 2\int_t^T \left(Y_s^1 - Y_s^2, Z_s^1 - Z_s^2\right) dW_s \tag{3.14}$$

Now, using the formula $(x + y)^2 = x^2 + 2xy + y^2$ we get that $-2xy = x^2 + y^2 - (x + y)^2$ so letting $x = Y_s^1 - Y_s^2$ and $y = f_s(s, Y_s^1, Z_s^1) - f_s(s, Y_s^2, Z_s^2)$ we get

$$
\begin{aligned}
-2(x,y) &= -2(x,y)\cdot\frac{\sqrt{2}L}{\sqrt{2}L} \qquad (3.15)\\
&= -2(\sqrt{2}Lx, \frac{y}{\sqrt{2}L}) \quad and letting \sqrt{2}Lx = \alpha, \frac{y}{\sqrt{2}L} = \beta\\
&= -2(\alpha,\beta)\\
&= \alpha^2+\beta^2-(\alpha+\beta)^2\\
&= 2L^2x^2+\frac{y^2}{2L^2}-(\sqrt{2}Lx+\frac{y}{\sqrt{2}L})^2\\
&\le 2L^2x^2+\frac{y^2}{2L^2}\\
&\le 2L^2|x|^2+\frac{|y|^2}{2L^2}\\
&= 2L^2|Y_s^1-Y_s^2|^2+\frac{|f_s(s,Y_s^1,Z_s^1)-f_s(s,Y_s^2,Z_s^2)|^2}{2L^2}\\
&\le 2L^2|Y_s^1-Y_s^2|^2+\frac{|Y_s^1-Y_s^2|^2}{2}+\frac{|Z_s^1-Z_s^2|^2}{2}\\
&\le 2L^2|Y_s^1-Y_s^2|^2+\frac{|Z_s^1-Z_s^2|^2}{2}
\end{aligned}
$$

where in the last step we used that f is uniformly Lipschitz in (Y,Z).

Now, combining (3.14) and (3.15) we get

$$
|Y_t^1-Y_t^2|^2+\int_t^T|Z_s^1-Z_s^2|^2ds \le \int_t^T 2L^2|Y_s^1-Y_s^2|^2+\frac{|Z_s^1-Z_s^2|^2}{2}ds\ -2\int_t^T\Big(Y_s^1-Y_s^2, Z_s^1-Z_s^2\Big)dW_s \qquad (3.16)
$$

Taking expectation we have

$$
\mathbb{E}|Y_t^1-Y_t^2|^2+\mathbb{E}\int_t^T|Z_s^1-Z_s^2|^2ds \le 2L^2\mathbb{E}\int_t^T|Y_s^1-Y_s^2|^2ds+\frac{1}{2}\mathbb{E}\int_t^T|Z_s^1-Z_s^2|^2ds \qquad (3.17)
$$

which becomes

$$
\mathbb{E}|Y_t^1-Y_t^2|^2 \le 2L^2\mathbb{E}\int_t^T|Y_s^1-Y_s^2|^2ds-\frac{1}{2}\mathbb{E}\int_t^T|Z_s^1-Z_s^2|^2ds \qquad (3.18)
$$

Applying Gronwall's Lemma (3) to (3.18) we get

$$
0\le \mathbb{E}|Y_t^1-Y_t^2|^2 \le -\frac{1}{2}\mathbb{E}\int_t^T|Z_s^1-Z_s^2|^2ds\ exp\Big(2L^2(T-t)\Big)\le 0 \qquad (3.19)
$$

Therefore we must have that both $\mathbb{E}|Y_t^1-Y_t^2|^2=0$ and $\mathbb{E}\int_t^T|Z_s^1-Z_s^2|^2ds=0$ since the exponential function is always positive which implies that $Y_t^1=Y_t^2$ and $Z_t^1=Z_t^2$.

Existence of (Y,Z):

To prove the existence of the solution to the BSDE we will use Picard iteration [3].Denote $Y_t^0 := 0, Z_t^0 := 0$ and let $\{(Y_t^n,Z_t^n) : t\in[0,T]\}_{n\ge1}$ be a sequence in $\mathbb{H}_T^2(\mathbb{R}^n)\times\mathbb{H}_T^2(\mathbb{R}^{n\times d})$ which is defined recursively by

$$Y_t^n = \xi + \int_t^T f(s, Y_s^{n-1}, Z_s^{n-1})ds - \int_t^T Z_s^n dW_s \tag{3.20}$$

We will use induction to prove that the above sequence is well defined. Assume that $Y^{n-1}, Z^{n-1} \in \mathbb{L}^2(\mathbb{F}) \times \mathbb{L}^2(\mathbb{F})$ and using the linear growth condition in [13] we have

$$|f_s(Y_s^{n-1}, Z_s^{n-1})| \leq L[|f_s^0| + |Y_s^{n-1}| + |Z_s^{n-1}|] \tag{3.21}$$

so that $f_s(Y_s^{n-1}, Z_s^{n-1}) \in \mathbb{L}^{1,2}(\mathbb{F})$. By the previous part of the proof we have that $(Y^n, Z^n) \in \mathbb{L}^2(\mathbb{F}) \times \mathbb{L}^2(\mathbb{F})$ is uniquely determined by (3.20) so we have that $(Y^n, Z^n) \in \mathbb{S}^2(\mathbb{F}) \times \mathbb{L}^2(\mathbb{F})$ and by induction we get $(Y^n, Z^n) \in \mathbb{S}^2(\mathbb{F}) \times \mathbb{L}^2(\mathbb{F})\ \ \forall n \geq 1$.

Now, using Itô's formula to $|Y_s^{n+1} - Y_s^n|$ as before and using again that f is uniformly Lipschitz we get

$$\mathbb{E}|Y_t^{n+1} - Y_t^n|^2 + \frac{1}{2}\mathbb{E}\int_t^T |Z_s^{n+1} - Z_s^n|^2 ds \leq L\mathbb{E}\int_t^T |Y_s^{n+1} - Y_s^n|^2 ds + L\mathbb{E}\int_t^T |Y_s^n - Y_s^{n-1}|^2 ds \tag{3.22}$$

which implies that

$$-\frac{d}{dt}\mathbb{E}\int_t^T |Y_s^{n+1} - Y_s^n|^2 ds - L\mathbb{E}\int_t^T |Y_s^{n+1} - Y_s^n|^2 ds \leq L\mathbb{E}\int_t^T |Y_s^n - Y_s^{n-1}|^2 ds \tag{3.23}$$

Now define $\nu_n(t) := \mathbb{E}\int_t^T |Y_s^n - Y_s^{n-1}|^2 ds$ and note that $\nu_{n+1}(T) = 0$. Integrating (3.23) we get,

$$\nu_{n+1}(t) \leq L\int_t^T \frac{1}{2}e^{L(s-t)}\nu_n(s)ds \tag{3.24}$$

Iterating (3.24) we get

$$\nu_{n+1}(0) \leq L\frac{(Le^{LT})^n}{n!}\nu_1(0) \tag{3.25}$$

Therefore, by letting $n \to \infty$ we can see that $\nu_{n+1}(0)$ converges. A similar argument can be done for $\mathbb{E}\int_t^T |Z_s^{n+1} - Z_s^n|^2 ds$ so that we can conclude that $\{(Y_t^n, Z_t^n) : t \in [0,T]\}_{n\geq 1}$ is a Cauchy sequence, and taking limits as $n \to \infty$ we have that by construction $Y := lim_{n\to\infty} Y_n$ and $Z := lim_{n\to\infty} Z_n$ so that the pair (Y, Z) is a solution of (3.6).

□

3.3 Comparison Theorem

In Mathematics, comparison theorems provide a statement which allows us to compare different mathematical objects of the same type. For example, in the theory of Stochastic Differential Equations when we talk about comparison theorems we mean when one process is greater than or equal to another, with probability one. These kind of theorems are very helpful, and in particular, they are used to prove many results in mathematical theory. These can range from asymptotic behavior to the uniqueness and existence of solution of SDEs.

Theorem 3.4. *Suppose we have two BSDEs with associated parameters* (f^1, ξ^1) *and* (f^2, ξ^2) *and let* (Y^1, Z^1) *and* (Y^2, Z^2) *be the associated solutions. Assume that* (ξ^i, f^i) *satisfy the points on the top of page 20 and* $(Y^i, Z^i) \in \mathbb{S}^2(\mathbb{F}) \times \mathbb{L}^2(\mathbb{F})$. *If* $\xi^1 \leq \xi^2$ $\mathbb{P}$*-a.s. and* $f^1(t, Y_t^2, Z_t^2) \leq f^2(t, Y_t^2, Z_t^2)$, $dt \times d\mathbb{P}$*-a.s. , then*

$$Y_t^1 \le Y_t^2, \quad 0 \le t \le T, \quad \mathbb{P}\text{-a.s.} \tag{3.26}$$

Proof. Using the fact that f^1 is uniformly Lipschitz continuous we have that

$$f^1(t, Y_t^1, Z_t^1) - f^1(t, Y_t^2, Z_t^2) \le |f^1(t, Y_t^1, Z_t^1) - f^1(t, Y_t^2, Z_t^2)| \le L\Big(|Y_t^1 - Y_t^2| + |Z_t^1 - Z_t^2|\Big) \tag{3.27}$$

Since we want to prove that $Y_t^1 \le Y_t^2$ it is convenient that we use $e^{\gamma t}\left[\Big(Y_s^1 - Y_s^2\Big)^+\right]^2$ such that $\gamma > 0$. Now, using Itô's formula for semi-martingales (5) and Tanaka's formula (A.3) on $e^{\gamma t}\left[\Big(Y_s^1 - Y_s^2\Big)^+\right]^2$ such that $\gamma > 0$ we get

$$\begin{aligned} e^{\gamma t}\left[\Big(Y_s^1 - Y_s^2\Big)^+\right]^2 = e^{\gamma T}\left[\Big(\xi^1 - \xi^2\Big)^+\right]^2 &- \int_t^T \gamma e^{\gamma t}\left[\Big(Y_s^1 - Y_s^2\Big)^+\right]^2 ds \\ &- 2\int_t^T e^{\gamma t}\Big(Y_s^1 - Y_s^2\Big)^+ d\Big(Y_s^1 - Y_s^2\Big)^+ \\ &- \frac{1}{2}\int_t^T e^{\gamma s}\left[\mathbb{I}_{(Y_s^1 - Y_s^2)^+}\right] d\Big\langle Y^1 - Y^2\Big\rangle_s \end{aligned} \tag{3.28}$$

Using the assumption that $\xi^1 - \xi^2 \le 0$ $\mathbb{P}$-a.s. (3.28) becomes

$$\begin{aligned} e^{\gamma t}\left[\Big(Y_t^1 - Y_t^2\Big)^+\right]^2 &= -2\int_t^T e^{\gamma t}\Big(Y_s^1 - Y_s^2\Big)^+\Big(Z_s^1 - Z_s^2\Big) dW_s \\ &+ \int_t^T e^{\gamma s}\left\{-\gamma\left[(Y_s^1 - Y_s^2)^+\right]^2 - \mathbb{I}_{\{Y_s^1 - Y_s^2 > 0\}}|Z_s^1 - Z_s^2|^2\right\} ds \\ &+ \int_t^T e^{\gamma s}\left\{2(Y_s^1 - Y_s^2)^+[f_s^1(Y_s^1, Z_s^1) - f_s^2(Y_s^2, Z_s^2)]\right\} ds \\ &\le -2\int_t^T e^{\gamma t}\Big(Y_s^1 - Y_s^2\Big)^+\Big(Z_s^1 - Z_s^2\Big) dW_s \\ &+ \int_t^T e^{\gamma s}\left\{-\gamma\left[(Y_s^1 - Y_s^2)^+\right]^2 - \mathbb{I}_{\{Y_s^1 - Y_s^2 > 0\}}|Z_s^1 - Z_s^2|^2\right\} ds \\ &+ \int_t^T e^{\gamma s}\left\{2L(Y_s^1 - Y_s^2)^+\left[|Y_t^1 - Y_t^2| + |Z_t^1 - Z_t^2|\right]\right\} ds \end{aligned} \tag{3.29}$$

where we have used (3.27) for the last inequality. Now, for simplicity, let $m = Y_s^1 - Y_s^2$ and $n = Z_s^1 - Z_s^2$ so that (3.29) becomes

$$
\begin{aligned}
e^{\gamma t}\Big[\Big(Y_t^1 - Y_t^2\Big)^+\Big]^2 \le -2\int_t^T e^{\gamma t}\ m^+ n dW_s \\
+\int_t^T e^{\gamma s}\Big\{-\gamma\Big[m^+\Big]^2 - \mathbb{I}_{\{m>0\}}|n|^2\Big\}ds \\
+\int_t^T e^{\gamma s}\Big\{2L(m)^+\Big[|m|+|n|\Big]\Big\}ds
\end{aligned}
\tag{3.30}
$$

Using the identity $-\gamma m^2 + 2Lmn = -\gamma(m - \frac{L}{m}n)^2 + \frac{L^2}{m}n^2 \le \frac{L^2}{m}n^2$ we get the following:

$$
\begin{aligned}
-\gamma(m^+)^2 - \mathbb{I}_{\{m>0\}}n^2 + 2Lm^+(|m|+|n|) &= -\mathbb{I}_{\{m>0\}}\Big[-\gamma|m|^2 - |n|^2 + 2L|m|(|m|+|n|)\Big] \\
&= \mathbb{I}_{\{m>0\}}\Big[(L^2+2L-\gamma)|m|^2 - (|n| - L|m|)^2\Big] \le 0
\end{aligned}
\tag{3.31}
$$

for all $L^2 + 2L \le \gamma$. Therefore, (3.30) becomes

$$
e^{\gamma t}\Big[\Big(Y_t^1 - Y_t^2\Big)^+\Big]^2 \le -2\int_t^T e^{\gamma s}\Big(Y_s^1 - Y_s^2\Big)^+\Big(Z_s^1 - Z_s^2\Big)dW_s \quad \mathbb{P}-a.s. \tag{3.32}
$$

Taking the expected value on both sides of (3.32) and using the fact that W_s is a Brownian motion we get

$$
\mathbb{E}\Big[e^{\gamma t}\Big[\Big(Y_t^1 - Y_t^2\Big)^+\Big]^2\Big] \le 0 \quad \mathbb{P}-a.s. \tag{3.33}
$$

which is equivalent to saying that $Y_t^1 \le Y_t^2$ which concludes the proof.

□

Remark 3.5. *In the above theorem, if additionally we had $Y_0^1 = Y_0^2$ then the inequality becomes an equality, i.e $Y_t^1 = Y_t^2$ for $0 \le t \le T\mathbb{P}$-a.s. with $\xi^1 = \xi^2$. If on the other hand we had a strict inequality between either ξ or f, i.e if $\xi^1 < \xi^2$ or $f^1(t, Y_t^2, Z_t^2) < f^2(t, Y_t^2, Z_t^2)$, $dt \times d\mathbb{P}$-a.s. then $Y_t^1 < Y_t^2$.*

3.4 Applications in Finance

As discussed before, the theory of BSDEs has a variety of applications in the financial markets and Mathematical Finance in general. In this section, we will show two of these applications. In particular, the first example considers asset pricing and hedging theory where as the second example considers the concept of recursive utility. We consider again working in a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ with a one-dimensional $(\mathbb{P}, \mathbb{F})$-Brownian motion W.

3.4.1 European call option pricing and hedging theory

A European call option is a financial contract which gives its owner the right but not the obligation to buy the underlying asset at a specific agreed price K, called the strike, at a certain date T in the future, called the maturity. The underlying asset can be any financial asset, i.e it can be a stock, index or a bond. The holders of a European call option can only exercise their right to buy the underlying at the maturity time T. For this reason, we shall assume that the holder of such an option will decide whether or not to exercise his option depending on the price of the underlying at maturity $S(T)$ which means that the profit of the investor will equal to $(S_T - K)^+$ which is an $\mathcal{F}_T$-measurable random variable. For simplicity, we let our underlying asset to be a stock.

Assume that the financial market is complete and consists of two assets: a bank account (or a riskless bond) with constant interest rate r and an asset (a stock) with price S_t such that their respective differential equations are as summarised below:

$$\begin{aligned} dS_t &= \mu S_t dt + \sigma S_t dW_t \\ dB_t &= rB_t dt \end{aligned}$$

where r, μ and σ are constant, and represent the bond's risk-free rate, the stock's drift and diffusion rate respectively and W_t is a Brownian motion.Now let's define some state variables:

- X_t be the investor's wealth at time t
- $\pi(t)$ the proportion of wealth invested in the risky stock at time t
- $c(t)$ the rate of consumption of wealth at time t

Note that $c(t)$ is a non-decreasing $\mathcal{F}_t$-adapted process.Now, assume that the investor has sold the European option at price x at time $t - 0$.
The controlled portfolio satisfies:

$$dX_t = \left(\left(r + (\mu - r)\,\pi_t\right) X_t^{\pi,c} - c_t\right)\ dt + \sigma\pi_t X_t^{\pi,c}\ dW_t \quad , \quad X_0^{\pi,c} = x$$

The problem of the investor is to find an optimal hedging strategy (π, c) against any European option $G \in \mathbb{L}_T^2$ which satisfies $X_T^{\pi,c} \geq G$. The fair price of such a contract is equivalent to the smallest value of x such that an optimal hedging strategy (π, c) exists. Let us denote this by x^* which satisfies

$$x^* = inf\{X_0^{\pi,c} = x;\ there\ exists\ (\pi, c)\ and\ X_T^{\pi,c} \geq G\}. \tag{3.34}$$

A case where this problem could be written as a BSDE would be if the investor is very risk-averse and decides not to spend any of his wealth,i.e $c(t) = 0$. He would then be able to choose the proportion of wealth $\pi(t)$ invested in the risky stock such that $X_T^{\pi,c} = G$ where G is any European option that can be written as a function of the price of the stock at maturity time T, i.e $G = f(S_T)$. Therefore, the above dynamics change to the BSDE

$$\begin{aligned} dS_t &= \mu S_t dt + \sigma S_t dW_t \quad , \quad X_0^{\pi} = x \\ dX_t &= \left(\left(r + (\mu - r)\,\pi_t\right) X_t^{\pi}\right)\ dt + \sigma\pi_t X_t^{\pi}\ dW_t \\ X_T^{\pi} &= f(S_T) \end{aligned} \tag{3.35}$$

Remark 3.6. *If a solution (Y, Z) to (3.35) exists, then $\pi(t)$ is the optimal hedging strategy. Additionally, if a solution exists, then the correct price for the contract that the investor sells at $t = 0$ is equal to x.*

3.4.2 Stochastic Differential utility in continuous time

Stochastic Differential utility was introduced by Duffie and Epstein [5] in a continuous-time setting. When we talk about utility in this section we should mean recursive utility. This approach to model utility allows to separate the risk aversion and intertemporal substitution. Intertemporal substitution refers to the fact that someone substitutes his future consumption for current consumption due to some economic factors such as new policies. Modelling utility instead of human behavior directly allows us to use the same model when we deal with changing environments or circumstances.

Let us now consider again an investor whose consumption rate at time t is denoted as $(c(t)), 0 \leq t \leq T$. Let us define the utility at time t by $Y(t) := u(\{c(s) : 0 \leq t \leq s \leq T\})$ for some utility function u. This stochastic model assumes also that $Y(t)$ and $c(t)$ are $\mathcal{F}_t$-adapted stochastic processes. If we also have the utility of terminal consumption at the final time T, i.e $u(c(T)) = Y(T)$ then the utility at time t is equal to the expectation of the function of future consumption and utility f called the aggregator ,satisfying the Lipschitz condition, and is defined recursively by

$$Y(t) = \mathbb{E}\Big[\int_t^T f_s(c(s), Y(s), Z(s))ds + Y(T)|\mathcal{F}_t\Big], \quad 0 \leq t \leq T \tag{3.36}$$

In Finance, future information is very uncertain and random, so we work in a filtered probability space generated by a brownian motion W. This allows us to use results of the martingale representation theorem from chapter 2 to have a stochastic process $\sigma \in \mathbb{L}^2(\mathbb{F})$ such that

$$Y(t) = Y(T) + \int_t^T f_s(c(s), Y(s), Z(s))ds - \int_t^T \sigma(s)dW_s, \quad 0 \leq t \leq T \tag{3.37}$$

which is a BSDE.

Remark 3.7. *In a discrete-time setting we would have the backward recursive relation between the consumption rate $c(t)$ and utility $Y(t)$ as:*

$$Y_t = f(c(t), Y_{t+1}) \tag{3.38}$$

where the function f is the aggregator.

Chapter 4

Optimal Stochastic Control through BSDEs and extensions

In this section we will first try to establish a connection between "standard" stochastic control problems and BSDEs and then we will study the equivalence of that connection for "singular" control problems through an example. For the first part, we will mainly refer to the theory discussed in [13].

Consider a controlled SDE of the form

$$\begin{cases} dX_t = b_t(X_t, u_t)dt + \sigma_t(X_t, u_t)dW_t, & t \in [0,T], \ \mathbb{P}\text{-a.s.} \\ X_0 = x \end{cases} \tag{4.1}$$

Equation (4.1) can also be written as

$$X_t^u = x + \int_0^t b(s, X_s^u, u_s)ds + \int_0^t \sigma_s(s, X_s^u, u_s)dW_s, \quad t \in [0,T], \ \mathbb{P}\text{-a.s.} \tag{4.2}$$

In the above, W is an $\mathbb{R}^d$-valued Brownian motion defined on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with respect to the filtration $\mathbb{F} = \{\mathcal{F}_t, t \in [0,T]\}$ satisfying the usual conditions. The unknown process X_t is the state of the system at time t and is $\mathcal{F}_t$-adapted and square-integrable process taking values in $\mathbb{R}^{d_1}$; $u \in \mathcal{U}$ are admissible control which is also $\mathcal{F}_t$-adapted and square-integrable process taking values in a given non-empty closed set $\mathcal{U}$. The functions $b : [0,T] \times \mathbb{R}^{d_1} \to \mathbb{R}^{d_1}$ and $\sigma : [0,T] \times \mathbb{R}^{d_1} \to \mathbb{R}^{d_1,d}$ are deterministic, Borel-measurable and uniformly Lipschitz continuous with respect to x and u. Now, we would like to solve the following stochastic optimization problem

$$V_0 := \sup_{u \in \mathcal{U}} J(u), \quad with \quad J(u) := \mathbb{E}^{\mathbb{P}}\left[\Psi(X_T) + \int_0^T f(t, X_t^u, u_t)dt\right] \tag{4.3}$$

We should note here that the functions Ψ and f are one-dimensional which means that V_0 and $J(u)$ are scalar values. The problem that arises here, is that if we try and solve the above equation using the stochastic maximum principle and the Hamiltonian system we would come up with a forward-backward SDE which generally does not have a solution. Therefore, our approach is to make certain assumptions that will make our problem simpler.

Assumption 1. *The function $\sigma(t,x)$ does not depend on the control u and is uniformly Lipschitz in x;*

Assumption 2. *There exists a function $\theta(t,x,u)$ such as $b(t,x,u) = \sigma(t,x)\theta(t,x,u)$ which is bounded and takes values in $\mathbb{R}^d$.*

Our next step is to show that under the above assumptions the probability measures $\mathbb{P}^u$ and $\mathbb{P}$ are equivalent. To do this, assume that the following SDE does not involve any control u and has unique solution X:

$$X_t = x + \int_0^t \sigma(s, X_s) dW_s, \quad t \in [0,T], \;\; \mathbb{P}\text{-a. s.} \tag{4.4}$$

Now let's define some variables.Let

$$\theta_t^u := \theta(t, X_t, u_t), \quad W_t^u := W_t - \int_0^t \theta_s^u ds, \quad \mathbb{P}^u := \mathbb{P}^{\theta^u} \tag{4.5}$$

Assumption (2) comes in handy here because it allows us to use Girsanov's theorem and have that W^u is a $\mathbb{P}^u$- Brownian motion as the function θ^u is bounded. This implies that the two probability measures $\mathbb{P}^u$ and $\mathbb{P}$ are equivalent. Thus, using (4.4) we have

$$X_t = x + \int_0^t b(s, X_s, u_s) ds + \int_0^t \sigma_s(s, X_s) dW_s^u, \quad t \in [0,T], \;\; \mathbb{P}^u\text{-a. s} \tag{4.6}$$

Assumption (2) and equation (4.6) allow us to write (4.3) as

$$V_0 := \sup_{u \in \mathcal{U}} J(u), \quad with \quad J(u) := \mathbb{E}^{\mathbb{P}^u}\left[\Psi(X_T) + \int_0^T f(t, X_t, u_t) dt\right] \tag{4.7}$$

Equation (4.7) represents the stochastic optimization problem under weak formulation that we seek to solve since it contains a drift control only.

Remark 4.1. *If we had not used assumptions (1) and (2), the probability measure $\mathbb{P}$ would not be equivalent with any other measure and thus we would have to control the state process X^u instead. This is known as the strong formulation of the stochastic optimization problem. In our case, we fixed the state process X and tried to control its distribution by controlling the probability $\mathbb{P}^u$.*

From (4.7) we use the Martingale Representation theorem that guarantees us a process Z such as

$$Y_t^u = \Psi(X_T) + \int_t^T f(s, X_s, u_s) ds - \int_t^T Z_s^u dW_s^u, \quad \mathbb{P}^u\text{-a. s.} \tag{4.8}$$

From chapter 2 we know that equation (4.8) is a linear BSDE which has a unique solution (Y^u, Z^u) under the probability measure $\mathbb{P}^u$. We should also note here that (4.8) at $t = 0$ becomes $Y_0^u = \Psi(X_T) + \int_0^T f(s, X_s, u_s) ds$ which is equal to $J(u)$. By (4.5) and the fact that $\mathbb{P}$ and $\mathbb{P}^u$ are equivalent, (4.8) becomes

$$Y_t^u = \Psi(X_T) + \int_t^T \left[f(s, X_s, u_s) + Z_s^u \theta(s, X_s, u_s)\right] ds - \int_t^T Z_s^u dW_s, \quad \mathbb{P}\text{-a. s.} \tag{4.9}$$

We re-write (4.10) as

$$Y_t^u = \Psi(X_T) + \int_t^T H(s, X_s, Z_s^u, u_s) ds - \int_t^T Z_s^u dW_s, \quad \mathbb{P}\text{-a. s.} \tag{4.10}$$

where the Hamiltonian function H is defined as $H(t, x, z, u) := f(t, x, z) + z\theta(t, x, u)$. Since our optimization problem involves the supremum of the cost functional J we should also define the maximal Hamiltonian $H^*(t, x, z) := \sup_{u \in \mathcal{U}} H(t, x, z, u)$. Note that H^* does not depend on the control u since it is just the supremum value of H under all $u \in \mathcal{U}$. Using assumption (2) and the fact that the functions f, Ψ, σ, b are all deterministic, Borel-measurable and bounded, we have that H^* is uniformly Lipschitz continuous in z and bounded when $z = 0$. Therefore, we know again using the theory in Chapter 2 that the BSDE

$$Y_t^* = \Psi(X_T) + \int_t^T H^*(s, X_s, Z_s^*)ds - \int_t^T Z_s^* dW_s, \quad \mathbb{P}\text{-a. s.} \tag{4.11}$$

has a unique solution (Y^*, Z^*). Now, all the above lead us to the following theorem:

Theorem 4.2. *Let the functions f, Ψ, σ, b be deterministic, Borel-measurable and bounded and suppose that assumptions (1) and (2) hold. Then,*

$$Y_0^* = \sup_{u \in \mathcal{U}} J(u) \tag{4.12}$$

Additionally, if there exists a function $g : [0, T] \times \mathbb{R}^{d_1} \times \mathbb{R}^d \to \mathcal{U}$ that is Borel-measurable and

$$H^*(t, x, z) = H(t, x, z, g(t, x, z)) \tag{4.13}$$

we have that

$$u_t^* := g(t, X_t, Z_t^*) \tag{4.14}$$

is an optimal control.

Proof. We first prove that there exists a unique solution in the BSDE (4.11) and then show that $Y_0^* = \sup_{u \in \mathcal{U}} J(u)$ and that u_t^* is an optimal control. Assumption (2) says that $\theta(t, x, u) = \sigma^{-1}(t, x)b(t, x, u)$ and $\Psi(X_T)$ are uniformly bounded and since $H(t, x, z, u) = f(t, x, z) + z\theta(t, x, u) = f(t, x, z) + z\sigma^{-1}(t, x)b(t, x, u)$ then $H(t, x, z, u)$ is a linear generator of a BSDE which satisfies the hypothesis of Proposition (1) in Chapter 3. This means that the BSDE

$$Y_t^u = \Psi(X_T) + \int_t^T H(t, x, Z_t^u, u_t)dt - \int_t^T Z_t^u dW_t \tag{4.15}$$

has a unique solution (Y^u, Z^u) $\mathbb{P}$-a. s.. Thus, we have

$$Y_0^u = \Psi(X_T) + \int_t^T H(s, X_s, Z_s^u, u_s)ds - \int_t^T Z_s^u dW_s. \tag{4.16}$$

Moreover, using equation (4.5) and assumption (2) we get

$$Y_0^u = \Psi(X_T) + \int_t^T \Big(f(s, X_s, Z_s^u) + Z_s^u \sigma^{-1}(s, X_s)b(s, X_s, u_s)\Big)ds - \int_t^T Z_s^u \Big[dW_t^u + \theta(s, X_s, u_s)dt\Big] \tag{4.17}$$

$$= \Psi(X_T) + \int_t^T \Big(f(s, X_s, Z_s^u) + Z_s^u \sigma^{-1}(s, X_s)b(s, X_s, u_s)\Big)ds - \int_t^T Z_s^u \Big[dW_t^u + \sigma^{-1}(s, X_s)b(s, X_s, u_s)dt\Big]$$

$$= \Psi(X_T) + \int_t^T f(s, X_s, u_s)ds - \int_t^T Z_s^u dW_s^u$$

Now taking expectations on both sides of (4.15) we get

$$Y_0^u = \mathbb{E}^{\mathbb{P}^u}\left[\Psi(X_T) + \int_t^T f(s, X_s, u_s)ds\right] = J(u) \tag{4.18}$$

To show that $(Y*, Z*)$ is a solution to (4.11) we just need to take $H^*(t, x, z)$ the generator of (4.11) as a supremum of uniformly Lipschitz coefficient which we can do using the Comparison theorem, i.e

$$\begin{aligned} |H^*(t,x,z) - H^*(t,x,\tilde{z})| &= |\sup_{u\in\mathcal{U}} H(t,x,z,u) - \sup_{u\in\mathcal{U}} H(t,x,\tilde{z},u)| \\ &\leq \sup_{u\in\mathcal{U}} |H(t,x,z,u) - H(t,x,z,u)| \\ &\leq C|z-\tilde{z}|. \end{aligned} \tag{4.19}$$

Now that we have shown that there exists a solution $(Y*, Z*)$ to (4.11) we need to show that $Y_0^* = \sup_{u\in\mathcal{U}} J(u)$. We will do that by showing that $Y_0^* \leq \sup_{u\in\mathcal{U}} J(u)$ and $Y_0^* \geq \sup_{u\in\mathcal{U}} J(u)$ which is equivalent to saying $Y_0^* = \sup_{u\in\mathcal{U}} = J(u)$.

Take any $\epsilon > 0$. Then you can find a Borel measurable function $Q^\epsilon : [0,T] \times \mathbb{R}^d \times \mathbb{R}^d \to \mathcal{U}$ such that $H^*(t,x,z) = \sup_{u\in\mathcal{U}} H(t,x,z,u) \leq H(t,x,z,Q^\epsilon(t,x,z)) + \epsilon$. Define $q_t^\epsilon := Q^\epsilon(t, X_t, Z_t^*)$. We have $H^*(t, X_t, Z_t^*) \leq H(t, X_t, Z_t^*, q_t^\epsilon) + \epsilon$ and

$$Y_t^{q^\epsilon} = \psi(X_T) + \int_t^T H(s, X_s, Z_s^{q^\epsilon}, q_s^\epsilon) - \int_t^T Z_s^{q^\epsilon} dW_s. \tag{4.20}$$

Consider

$$\begin{aligned} Y_t^* - Y_t^{q^\epsilon} &= \int_t^T \left[H^*(s,X_s,Z_s^*) - H(s,X_s,Z_s^*,q_s^\epsilon) + \left(Z_s^* - Z_s^{q^\epsilon}\right)\theta(s,X_s,q_s^\epsilon)\right] ds - \int_t^T \left(Z_s^* - Z_s^{q^\epsilon}\right) dW_s \qquad (4.21) \\ &= \int_t^T \left[H^*(s,X_s,Z_s^*) - H(s,X_s,Z_s^*,q_s^\epsilon)\right] ds - \int_t^T \left(Z_s^* - Z_s^{q^\epsilon}\right) dW_s^{q^\epsilon} \\ &\leq \epsilon(T-t) - \int_t^T \left(Z_s^* - Z_s^{q^\epsilon}\right) dW_s^{q^\epsilon} \end{aligned}$$

For $t = 0$ we have $Y_0^* - Y_0^{q^\epsilon} \leq \epsilon T$ and as $\epsilon > 0$ was chosen arbitrarily we get that $Y_0^* \leq \sup_{u\in\mathcal{U}} J(u)$. For the other inequality we have , by the Comparison theorem, that $Y_0^k \leq Y_0^* \;\; \forall u \in \mathcal{U}$ which implies that $\sup_{u\in\mathcal{U}} \leq Y_0^*$. □

4.1 Singular Stochastic Control Problems and BSDEs

In the previous section we studied the existence of "regular" optimal controls for the stochastic control problem through BSDEs where the state variable X_t satisfies the controlled SDE (4.2) . In this section we are set to study "singular" stochastic control problems and find a connection with BSDEs. This type of stochastic control problems differ from the typical control problems in finance as the possess some additional features not always present in the classical control problems. For example, as we will show later, the dividend optimization control problem can be seen as a consumption model with linear utility which involves risky assets which are governed by an arithmetic Brownian Motion whose growth is linear rather than exponential. This particular example involves a singular control which results in a cumulative impact of finite variation, i.e the cumulative amount of dividends paid out up to time t.

It is known in the literature, see [4], that for typical Markovian singular stochastic control problems, it can be shown using dynamic programming arguments that the optimal value is given by a viscosity solution of a Partial Differential Equation (PDE) with gradient constraint, which in turn can be shown that this type of PDE are related to BSDEs with Z-constraints. As the dividend optimization control problem involves only a singular control, we shall consider optimization problems of the form

$$I := \sup_{\beta} \mathbb{E}\Big[\int_0^T f(t, X^\beta)dt + \int_0^T g(t, X^\beta)d\beta_t + h(X^\beta) \Big] \tag{4.22}$$

where X_t satisfies

$$dX_t^\beta = \mu(t, X^\beta)dt + \nu(t, X^\beta)d\beta_t + \sigma(t, X^\beta)dW_t, \quad X_0 = x \in \mathbb{R}^d \tag{4.23}$$

where W is an n-dimensional Brownian Motion and (β_t) is an l-dimensional process with non-decreasing components known as the singular control.

We will need the following definition for the remaining of the section.

Definition 4.1.1. *A supersolution of the BSDE*

$$Y_t = h(X) + \int_t^T p(s, X, Z_s)ds - \int_t^T Z_s dW_s \;\; \textit{with constraint} \;\; q(t, X, Z_s) \in \mathbb{R}^l \tag{4.24}$$

consists of a triplet $(Y, Z, K) \in \mathbb{S}^2 \times \mathbb{H}^2 \times \mathcal{K}^2$ *such that*

$$Y_t = h(X) + \int_t^T p(s, X, Z_s)ds - \int_t^T Z_s dW_s + (K_T - K_t), \qquad \forall t \in [0, T] \tag{4.25}$$

Additionally, the triplet (Y, Z, K) *is a minimal supersolution of (4.25) if* $Y_t \leq \tilde{Y}_t, 0 \leq t \leq T$ *for any other supersolution* $(\tilde{Y}, \tilde{Z}, \tilde{K})$.

It can be shown that the optimal value of (4.22) is given by the initial value of the minimal supersolution of a BSDE

$$Y_t = h(X) + \int_t^T p(s, X, Z_s)ds - \int_t^T Z_s dW_s \tag{4.26}$$

subject to a constraint $q(t, X, Z_s) \in \mathbb{R}^l$, and (X_t) is the unique strong solution of an SDE

$$dX_t = \eta(t, X)dt + \sigma(t, X)dW_t, \quad X_0 = x. \tag{4.27}$$

To further understand the connection between PDEs with gradient constaints and BSDEs with Z-constaints we will consider a problem particularly important in actuarial science, namely the dividend optimal problem. This problem involves finding the optimal rate of a dividend pay-outs for an insurance company. As mentioned before, this type of problem possesses some additional features compared to "regular" control problems such as the singularity of the dividend distribution process. More specifically, the dividend pay-outs are subtracted from the current surplus of a portfolio which makes a control action to change the value of the underlying. As in the classical stochastic control problems, the dynamic programming principle is also satisfied for singular stochactic control problems and the Hamiltonian-Jacobi-Bellman (HJB) equation is a second-order variational inequality.

Let us now consider the reserve of an insurance company, denoted as R_t, which satisfies

$$R_t = x + pt - \sum_{i=1}^{N_t} Y_i \tag{4.28}$$

where x is the deterministic initial capital and p represents the constant intensity at which premiums are collected by the insurance company. We further assume that claims arrive at a Poison rate λ, where the size of the i-th claim is denoted as Y_i and we have that $\mathbb{E}(Y_i) := m, Var(Y_i) := s^2$. If we normalize the state space as: $r_t \to \frac{r_{nt}}{\sqrt{n}}$ then R_t satisfies

$$dR_t = \mu dt + \sigma dW_t \tag{4.29}$$

where W_t is a standard Brownian Motion and $\mu = p - \lambda m, \sigma^2 = \lambda(m^2 + s^2)$.

In most cases, (4.29) represents the dynamics of the reserve process when no control actions are taken. However, depending on the type of risk controls and constraints on the dividend pay-out rates, we get different types of diffusion control models, including the one we will study, the singular control model.

Considering now the dividend control model, the dynamics of R_t involve an extra term $-dL_t$ where L_t represents the cumulative amount of dividenda paid out up to time t, thus is subtracted from (4.29). we assume that L_t is non-negative, adapted to the filtration $\mathbb{F}$ and is right-continuous with left limits (cadlag) with $L_0 = 0$. We only look at cadlag processes because we assume that decisions have to be fixed in a predictable way.

Definition 4.1.2. *The bankruptcy time τ^L denotes the time of ruin of R_t and is defined as*

$$\tau^L := inf\{t \geq 0 : R_t \leq 0\}$$

The classical performance measure for a certain dividend strategy L is denoted as

$$J_x(L) := \mathbb{E}\Big[\int_0^{\tau^L} e^{-\delta t} dL_t + e^{-\delta(\tau\wedge\gamma)} V(R^L_{(\tau\wedge\gamma)})\Big] \tag{4.30}$$

i.e the expected value of discounted future dividend payments, where δ is the discount factor. The associated optimization problem consists of computing

$$V(x) = \sup_{L\in\Pi} J_x(L) \tag{4.31}$$

and an optimal admissible strategy L^* such that $V(x) = J_x(L^*)$, where V is the value function, Π is the set of all admissible controls and the functional L^* is the optimal control or optimal strategy.for simplicity reasons, we will consider a subclass of admissible control strategies that admit an adapted non-negative density process $L = l(s)_{s\geq 0}$ such that

$$l_t = \int_0^t l_s ds$$

and therefore we will write $J_x(l(\cdot))$ instead of $J_x(L)$. Now, the set of admissible controls Π will vary depending on the assumption one makes. In our case, we shall assume working with unbounded dividend rates and therefore we drop the assumption of absolute continuity of L so that the density process $l = l(s)_{s0}$ of a dividend policy L is not necessarily bounded.

Before tackling the problem, we should mention that the corresponding HJB equation for the singular control problem (for details see ([1])) with constant drift $\mu > 0$ and constant volatility $\sigma > 0$ is

$$\max\left\{\mu V_x + \frac{\sigma^2}{2}V_{xx} - \delta V, 1 - V_x\right\}, \quad V(0) = 0 \tag{4.32}$$

To derive (4.32) we used the the fact that we have to deal with the following equation from [1]

$$\sup_{l\geq 0}\left\{(1 - V_x(x))l + cV_x(x) - (\lambda+\delta)V(x) + \lambda\int_0^x V(x-y)dF_Y(y)\right\} = 0 \tag{4.33}$$

so that if we have $1 > V_x(x)$ for some $x \geq 0$ then the local maximizer $l^*(x)$ and (4.33) is unbounded which does not make sense. If on the other hand we have $V_x(x) > 1$ we get $l^*(x) = 0$ and

$$cV_x(x) - (\lambda+\delta)V(x) + \lambda\int_0^x V(x-y)dF_Y(y) = 0$$

so that restricting to $1 - V_x(x) \leq 0$ for all x greater or equal than zero we get the HJB (4.32). Now coming back to equations (4.22) and (4.23) we note that the functions g and ν are related to the constraint $1 - V_x(x) \leq 0$ since they involve the control L.

4.1.1 BSDE approach with a non-constant interest rate

Our goal in this final part of the project is to find an optimal strategy L for an insurance company whose reserve X_t is given by

$$X_t = x + \mu t + \sigma W_t \tag{4.34}$$

as in (4.29) where we have added the initial capital x. We also assume that the insurance companys pays out dividends, which represent a accumulated and non-negative process which is given by $L_t = \int_0^t l_s ds$. The dynamics of X then change to

$$X_t^l = x + \mu t + \sigma W_t - L_t \tag{4.35}$$

Let $\{B_t\}$ be a standard Brownian Motion which is independent of $\{W_t\}$, with corresponding filtration $\{\mathcal{F}_t^B\}$ which is complete and satisfies the usual conditions of right continuity and augmented with all the $\mathbb{P}$-null sets. Let also $\{\mathcal{F}_t\}$ be the filtration generated by the two Brownian Motions $\{B_t, W_t\}$ also satisfying the usual conditions. At this point, we should also assume that the dividends are discounted with rate δ which satisfies the SDE

$$d\delta_t = m(n - \delta_t)dt + rdB_t, \quad m, r > 0, n \in \mathbb{R} \tag{4.36}$$

Let's define also

$$\Delta_t^s := \int_t^s \delta_u du, \quad s \geq t \tag{4.37}$$

where $t \in [0,T]$ for some fixed horizon $T \in (0,\infty)$. It is important to mention that we should take into account only admissible strategies $l_t \in \Pi$ such that $l_t \in [0,\varsigma]$ where $\varsigma \in \mathbb{R}$. The corresponding forms of equations (4.30) and (4.31) for this situation becomes

$$V^l(t,\delta,x) = \mathbb{E}_{t,\delta,x}\Big[\int_t^{\tau^L\wedge T} e^{-\Delta_t^s} l_s ds + e^{-\Delta_t^{(\tau^L\wedge T)}} X^l_{(\tau^L\wedge T)}\Big] \tag{4.38}$$

and

$$V(t,\delta,x) = \sup_{l\in\Pi} V^l(t,\delta,x), \quad (t,\delta,x) \in [0, T\wedge\tau^L]\times\mathbb{R}\times\mathbb{R}, \quad V(T,\delta,x) = x \tag{4.39}$$

where $\mathbb{E}_{s,x}[\cdot] = \mathbb{E}[\cdot|X_s = x]$.

Our goal now is to find an optimal strategy l for (4.38) using a generalized HJB. Note that the dynamics of X^l is now

$$dX_s^l = (\mu - l_s)ds + \sigma dW_s + 0dB_s, \quad s \geq t, \quad X_t^l = x \tag{4.40}$$

Thus, the generalized Hamiltonian, (see [6] for details) $H : \Omega \times [0,T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}^{1\times 2} \to \mathbb{R}$ is of the form

$$H(s,x,l,y,z) = (\mu - l)y + tr\Big((\sigma \;\; 0)^T\Big) + e^{-\Delta_t^s} l \tag{4.41}$$

We notice that (4.41) does not depend on x. The maximum is attained at

$$l^* = arg\max_{l\in\Pi} H(s,x,l,y,z) = \begin{cases} \varsigma, & e^{-\Delta_t^s} - y > 0 \\ 0, & e^{-\Delta_t^s} - y \leq 0 \end{cases} \tag{4.42}$$

Then, with all the above we are in a position to write the corresponding BSDE

$$Y_t = h_x(X_T^{L^*}) + \int_t^T H_x(u, X_u^{L^*}, l_u^*, Y_u, Z_u)du - \int_t^T Z_u dW_u \tag{4.43}$$

where $h(x) := e^{-\Delta_t^T} x$.

Chapter 5

Conclusion

In this Thesis we started by giving two different proofs of the Martingale Representation Theorem where the first one uses a PDE approach in a one-dimensional setting and the second one follows a pure probabilistic argument which starts with a d-dimensional vector and then breaks the problem down into d separate, one-dimensional cases. The reason we started this Thesis in this way is because the Martingale Representation Theorem (MRT) is the cornerstone for the theory of BSDEs. The main result of MRT is that under the assumptions described in Chapter 2, it guarantees the existence and uniqueness of a process $\sigma \in \mathbb{L}^2(\mathbb{F})$ and basically states that a random variable that is measurable with respect to a filtration generated by a Brownian Motion W can be represented as an Itô integral with respect to this Brownian Motion.

Then, in Chapter 3, we went along and showed some financial applications of the Martingale Representation Theorem and of BSDEs in particular. More specifically, we derived the corresponding BSDEs for two problems in finance, namely finding an optimal hedging strategy against a European call option which satisfies a final condition at maturity time T and also modelling stochastic differential utility for an investor whose consumption rate c_t is considered to be an $\mathcal{F}$-t adapted stochastic process. Before tackling these two problems however, we proved the uniquness and existence of solutions to BSDEs, which consists of a pair (Y, Z) that satisfy equation (3.6). Moreover, we gave a proof of the Comparison theorem for BSDEs which states that if we have two BSDEs with associated solutions (Y^1, Z^1) and (Y^2, Z^2) with $\xi^1 \leq \xi^2$ $\mathbb{P}$-a.s and $f^1(t, y, z) \leq f^2(t, y, z), dt \times d\mathbb{P}$-a.s. for any pair (y, z) then we have that $Y_t^1 \leq Y_t^2$ $\mathbb{P}$-a.s.

In our final part of this Thesis, we tried to establish a connection between "classical" stochastic control problems and BSDEs where the MRT allowed us to use a process Z to represent the stochastic problem as a linear BSDE with the generator function being the maximal Hamiltonian and move on to compute the optimal control. For the last part of this chapter, we studied a very well-known problem in actuarial science, the dividend optimization problem. Through the dividend problem we were set to study singular stochastic control problems and find the connection with BSDEs. We found that these types of problems have optimal values which are given by the viscosity solutions of PDEs with gradient contraints and we showed that these type of PDEs are linked with BSDEs with Z-constraints.

All in all, we conclude that the theory of BSDEs can be very useful in solving problems in finance that involve stochastic processes and random variables that are measurable with respect to a filtration generated by a Brownian Motion. The theory of BSDEs is relatively new compared to other theories in Mathematics and is still under a lot of research.

Bibliography

[1] Hansjörg Albrecher and Stefan Thonhauser. Optimality results for dividend problems in insurance. *RACSAM-Revista de la Real Academia de Ciencias Exactas, Fisicas y Naturales. Serie A. Matematicas*, 103(2):295–320, 2009.

[2] Jean-Michel Bismut. *Théorie probabiliste du contrôle des diffusions*, volume 167. American Mathematical Soc., 1976.

[3] Tino Böhme. Bsde–existence and uniqueness. 2008.

[4] Bruno Bouchard, Patrick Cheridito, and Ying Hu. Bsde formulation of combined regular and singular stochastic control problems. *arXiv preprint arXiv:1801.03336*, 2018.

[5] Darrell Duffie and Larry G Epstein. Stochastic differential utility. *Econometrica: Journal of the Econometric Society*, pages 353–394, 1992.

[6] Julia Eisenberg, Stefan Kremsner, and Alexander Steinicke. Two approaches for a dividend maximization problem under an ornstein-uhlenbeck interest rate. *Mathematics*, 9(18):2257, 2021.

[7] Alison Etheridge. Continuous martingales and stochastic calculus. 2018.

[8] Ioannis Karatzas and Steven Shreve. *Brownian motion and stochastic calculus*, volume 113. Springer Science & Business Media, 2012.

[9] Magdalena Kobylanski. Backward stochastic differential equations and partial differential equations with quadratic growth. *the Annals of Probability*, 28(2):558–602, 2000.

[10] Paul Lévy. *Processus stochastiques et mouvement brownien*, volume 1948. Paris, 1948.

[11] The University of Utah. Some linear algebra.

[12] Etienne Pardoux and Shige Peng. Adapted solution of a backward stochastic differential equation. *Systems & control letters*, 14(1):55–61, 1990.

[13] Jianfeng Zhang. Backward stochastic differential equations. In *Backward Stochastic Differential Equations*, pages 42–46. Springer, 2017.

Appendix A

Stochastic Calculus

Theorem A.1 (Dominated Convergence Theorem)**.** *Assume $X, X^n \in \mathbb{L}^0(\mathbb{F}),\ \ n \geq 1$ and that $X^n \to X$ in $\mathbb{P}$. Let $1 \leq p < \infty$. If the random variables $\{|X^n|^p\}_{n\geq 1}$ are uniformly integrable, i.e if there exists $H \in \mathbb{L}^p(\mathbb{F})$ such as $|X^n| \leq H,\ \ n \geq 1$ then $X^n \to X$ in $\mathbb{L}^p(\mathbb{F})$.*

Remark A.2. *The Dominated Convergence Theorem is also true if you replace the filtration $\mathbb{F}$ with the σ-field $\mathcal{F}$ in $\mathbb{L}^p(\mathbb{F})$.*

Lemma 4 (Doob-Dynkin Lemma)**.** *Let $N, M \in \mathbb{L}^0(\mathcal{F}^N)$ if and only if ther exists a Borel-measurable function $g : \mathbb{R} \to \mathbb{R}$ such that $M = g(N)$.*

Lemma 5 (Itô's formula for semimartingales)**.** *Let X_t be a d-dimensional semimartingale and assume $f \in \mathcal{C}_2$. Then*

$$f(X_t) = f(X_0) + \int_0^t \sum_{i=1}^d f_i(X_s)\, dX_s^i + \frac{1}{2}\int_0^t \sum_{i,j=1}^d f_{i,j}(X_s)\, d\left\langle X^i, X^j \right\rangle_s \tag{A.1}$$

Lemma 6 (Itô's formula for BSDEs)**.** *Let X_t be a d-dimensional semimartingale and assume $f \in \mathcal{C}_2$. Define $\xi := X_T$. Then, for all $t \in [0, T]$,*

$$f(X_t) = f(\xi) - \int_t^T \sum_{i=1}^d f_i(X_s)\, dX_s^i - \frac{1}{2}\int_t^T \sum_{i,j=1}^d f_{i,j}(X_s)\, d\left\langle X^i, X^j \right\rangle_s \tag{A.2}$$

Proof. Since we assume that X_t is a d-dimensional semimartingale and $f \in \mathcal{C}_2$ then we can apply Lemma (5) on $f(\xi)$ to get

$$f(\xi) = f(X_0) + \int_0^T \sum_{i=1}^d f_i(X_s)\, dX_s^i + \frac{1}{2}\int_0^T \sum_{i,j=1}^d f_{i,j}(X_s)\, d\left\langle X^i, X^j \right\rangle_s \tag{A.3}$$

Now, subtracting (A.3) from (A.1) we get

$$f(X_t) = f(\xi) - \int_t^T \sum_{i=1}^d f_i(X_s)\, dX_s^i - \frac{1}{2}\int_t^T \sum_{i,j=1}^d f_{i,j}(X_s)\, d\left\langle X^i, X^j \right\rangle_s \tag{A.4}$$

□

Theorem A.3 (Tanaka's formula). *Given a Brownian Motion $W = \{W_t, t \geq 0\}$, Tanaka's formula can be written as*

$$|W_t| = \int_0^t sgn(W_s)dW_s + L_0^W(t) \tag{A.5}$$

where $L_0^W(t)$ is the local time of W at zero and is given by

$$L_0^W(t) = \mathbb{P}\, lim_{\epsilon \to 0} \frac{1}{2\epsilon} \int_0^t \mathbb{I}_{\{-\epsilon,\epsilon\}}(W(s))ds \tag{A.6}$$

and sgn denotes the sign function

$$sgn(x) = \begin{cases} 1 & \text{if } x \geq 0; \\ -1 & \text{if } x < 0. \end{cases} \tag{A.7}$$

Remark A.4. *The limit $L_0^W(t)$ exists in the L^2 sense and it is clearly adapted to the Brownian filtration. Also, we have that $L_0^W(s) \leq L_0^W(t)$ a.s. for $s < t$. In general, Tanaka's formula provides a decomposition of the submartingale $|W(t)|$.*