SJÄLVSTÄNDIGA ARBETEN I MATEMATIK

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET

An Introduction to Forward Backward Stochastic Differential Equations

av

Linus Nyström

2009 - No 11

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET, 10691 STOCKHOLM

An Introduction to Forward Backward Stochastic Differential Equations

Linus Nyström

Självständigt arbete i matematik 30 högskolepoäng, grundnivå

Handledare: Yishao Zhou

2009

Abstract

The aim of this paper is to introduce the reader to the concept of Forward Backward Stochastic Differential Equations (FBSDEs). We begin with an overview of some theoretical preliminaries and a formal presentation of our problem. We then proceed to study the solvability of FBSDEs through the use of mathematical control theory. This will subsequently lead us to a method for explicitly solving FBSDEs. We investigate when this method is applicable and what restrictions it brings. Finally, we conclude the text with two examples of applications of our theory.

[0]Contact: linus.nystrom@gmail.com

Contents

Chapter 1

Introduction

There are a number of reasons for studying mathematics. Some people see it only as a powerful tool to study and model real world phenomena, while others are fascinated by the elegance present in its pure theoretical form. When I first came into contact with stochastic analysis, I was struck by what an interesting combination of the practical and the theoretical the subject presented. There is a plethora of applications for the theory in physics, engineering and finance, to name a few. But there is also an amazing depth to the underlying mathematical theory. By combining elements from a variety of different areas of mathematics, such as real and functional analysis, probability theory and the theory of PDEs, we can reach some interesting and powerful results. In this particular paper, we will look at some rather recent developments in the theory of stochastic differential equations. This new theory will allow us to modify and verify some existent models in mathematical finance. However, there are other applications not mentioned in the text, in such fields as mathematical control theory.

Stochastic analysis is considered a technical subject. I have tried to be as clear as possible, without extending the scope of the paper too far. There is a great deal of PDE theory in the subject, but I have deliberately left it for the interested reader to explore outside of the text. Focus lies on the stochastic methods and analysis. I have also tried to present proper references, so that any reader wishing to delve deeper into a particular subject may do so. Unfortunately, this text is not self contained, some experience of stochastic analysis is required. Standard methods such as Itô calculus is frequently used, and I recommend [1] as primer and companion to anyone interested in learning the basics. I wish to point out that it is not necessary to understand all of the technicalities in order understand the results presented, some of the more exotic definitions and conditions are presented merely for completeness sake.

Finally, I would like thank my supervisor Yishao Zhou for her great patience and support.

Chapter 2

Theoretical Preliminaries

The reader is assumed to be familiar with basic stochastic analysis such as the creation of the Itô integral and the derivation of the Itô calculus. A recommended reference is [1] which gives a welcoming introduction to the subject. The theory of stochastic analysis is deeply linked to martingale theory, and as a service to the reader we will give a brief recap of some central concepts and definitions.

2.1 Filtrations and martingales

Although the reader will most likely be familiar with the concept of a σ-algebra, we start by introducing the formal definition for the sake of completeness.

Definition 2.1. *Let ω be a set. A collection of subsets $\mathcal{F}$ of Ω is called a σ-algebra if the following holds*

i) $\mathcal{F}$ *is non empty.*

ii) $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$.

iii) $\{A_i\}_{i=0}^{\infty} \in \mathcal{F} \Rightarrow \bigcup_{i=0}^{\infty} A_i \in \mathcal{F}$.

In martingale theory, information about the past is the key to predicting the future. Formally, a carrier of such information is called a *filtration.*

Definition 2.2. *Let $\mathcal{F}$ be a σ-algebra on a set Ω and let $\mathbf{P}$ be a probability measure defined on $(\Omega, \mathcal{F})$. A* filtration *is a collection of σ-algebras $\{\mathcal{F}_t\}_{t\geq 0}$ such that $\mathcal{F}_t \subset \mathcal{F}$ and*

$$0 \leq s < t \Rightarrow \mathcal{F}_s \subset \mathcal{F}_t \tag{2.1}$$

$\{\mathcal{F}_t\}_{t\geq 0}$ *is also called a nested family or an increasing sequence of σ-algebras.*

Let $X(t)$ be an n-dimensional stochastic process on the probability space $(\Omega, \mathcal{F}, \mathbf{P})$. We say that $X(t)$ is martingale *with respect to a filtration $\{\mathcal{F}_t\}_{t\geq 0}$ if*

- *$X(t)$ is $\mathcal{F}_t$-measurable for all t.*

- $E[|X(t)|] < \infty$ *for all* t.
- $E[X(t)|X(s)] = X(s)$, *for all* $s \leq t$.

If a stochastic process $X(t)$ is $\mathcal{F}_t$-measurable for all t, where $\{\mathcal{F}_t\}_{t\geq 0}$ is a filtration, we say that $X(t)$ is *adapted* to $\{\mathcal{F}_t\}_{t\geq 0}$. Adaptedness is often heuristically explained by saying that the filtration $\{\mathcal{F}_t\}_{t\geq 0}$ contains the necessary information to determine the process $X(t)$ at the time t.

Definition 2.3. *Let W(t) be a 1-dimensional stochastic process on a probability space* $(\Omega, \mathcal{F}, \mathbf{P})$. *We say that W(t) is a standard Brownian motion if*

i) W(0) = 0,

ii) W(t) is a.s. continuous,

iii) W(t) has independent increments with distribution $W(t) - W(s) \sim N(0, t - s)$ *for* $0 \leq s < t$.

If $W(t) = (W_1(t), \ldots, W_n(t))$ *is an n-dimensional stochastic process such that each of the* $W_i(t)$ *is a standard Brownian motion, we say that W(t) is an n-dimensional standard Brownian motion.*

Standard Brownian motion is often referred to simply as Brownian motion. In harmony with standard notation, $N(\mu, \sigma^2)$ is the normal distribution with mean μ and variance σ^2.

Definition 2.4. *Let* $X : \Omega \to \mathbf{R}^n$ *be any function. By* $\mathcal{F}_X$ *we denote the smallest* σ*-algebra in respect to which* X *is measurable. We call* $\mathcal{F}_X$ *the* σ-algebra generated by X.

Let $\mathcal{F}_{W(t)}$ be the σ-algebra generated by $\{W(s) : s \leq t\}$ where $W(t)$ is n-dimensional Brownian motion on some probability space $(\Omega, \mathcal{F}, \mathbf{P})$. Clearly, the family $\{\mathcal{F}_{W(t)}\}_{t\geq 0}$ is a filtration, and we call it the filtration *generated* by $W(t)$, or the *natural* filtration of $W(t)$. It is a well known fact that the process $W(t)$ is a martingale with respect to its natural filtration.

Martingales have many convenient properties. There is one in particular we will use in this paper, namely the martingale representation theorem

Theorem 2.5. (The martingale representation theorem). *Let* $(\Omega, \mathcal{F}, \mathbf{P})$ *be a probability space associated with an n-dimensional Brownian motion* $W(t)$. *Let* $\{\mathcal{F}_t\}_{t\geq 0}$ *be the natural filtration of* $W(t)$. *If* $X(t)$ *is an n-dimensional martingale with respect to* $\{\mathcal{F}_t\}_{t\geq 0}$, *then there exists a unique adapted square integrable stochastic process* $Z(t)$ *such that*

$$X(t) = E[X(0)] + \int_0^t Z(t)\,\mathrm{d}W(t), \quad \forall t \geq 0, \ \textit{a.s.} \tag{2.2}$$

If $\{\mathcal{F}_t\}_{t\geq 0}$ is a filtration of the measure space $(\Omega, \mathcal{F})$ with probability measure $\mathbf{P}$, we call the quadruple $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbf{P})$ a *filtered* probability space.

If every subset of a $\mathbf{P}$-null set is measurable, we say that the probability space is *complete*.

Throughout rest of this paper, unless otherwise stated, we are always working in a complete, filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbf{P})$ associated with an n-dimensional Brownian motion $W(t)$. $\{\mathcal{F}_t\}_{t\geq 0}$ is the natural filtration of $W(t)$ augmented with all $\mathbf{P}$-null sets. Since the Brownian filtration is continuous by construction, our standard probability space satisfies what is normally called the 'usual conditions' in stochastic analysis; it is complete, filtered and the filtration is right continuous.

2.2 Notation

Throughout this text, $\mathbf{R}^n$ is the usual n - dimensional Euclidean space with the usual Euclidean norm $|\cdot|$ and Euclidean inner product $\langle \cdot, \cdot \rangle$. $\mathbf{R}^{n\times m}$ denotes the Hilbert space consisting of all $n \times m$ matrices with inner product given by $\langle A, B\rangle \triangleq \text{tr}\{AB^T\}$, $\forall A, B \in \mathbf{R}^{n\times m}$.

In this text we work with processes and functions from a number of different function spaces. Below is a formal presentation of the notation used. It may initially give a very technical impression, but the conditions imposed by each space upon its members are standard boundedness and regularity conditions, such as Lipschitz continuity. There is further no need to memorise exactly what each space entails, as it suffices to remember that membership of a certain space generally ensures that the function or process in question is well behaved and regular enough.

Assume that $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbf{P})$ is a complete filtered probability space. For any fixed $T > 0$, we denote

- for any sub-σ-algebra $\mathcal{G}$ of $\mathcal{F}$, $L^2_{\mathcal{G}}(\Omega; \mathbf{R}^n)$ is the set of all $\mathcal{G}$-measurable random variables X, taking values in $\mathbf{R}^n$, such that $\left(\int_\Omega |X(\omega)|^2 \,\mathrm{d}\mathbf{P}(\omega)\right)^{1/2} < \infty$.

- $L^2_{\mathcal{F}}(\Omega; L^2(0, T; \mathbf{R}^n))$ is the set of all $\{\mathcal{F}_t\}_{t\geq 0}$-progressively measurable processes X, taking values in $\mathbf{R}^n$, such that $\int_0^T E|X(t)|^2 \,\mathrm{d}t < \infty$. This space is often denoted by $L^2_{\mathcal{F}}(0, T; \mathbf{R}^n)$ when there is no risk of confusion.

- $L^2_{\mathcal{F}}(\Omega; C([0, T]; \mathbf{R}^n))$ is the set of all $\{\mathcal{F}_t\}_{t\geq 0}$-progressively measurable continuous processes X, taking values in $\mathbf{R}^n$, such that $E\{\sup_{t\in[0,T]} |X(t)|^2\} < \infty$.

- If N and M are any Euclidean spaces, we denote by $L^2_{\mathcal{F}}(0, T; W^{1,\infty}(M; N))$ the set of all functions $f : [0, T] \times M \times \Omega \to N$, such that for any fixed $\theta \in M$, $(t, \omega) \mapsto f(t, \theta; \omega)$ is $\{\mathcal{F}_t\}_{t\geq 0}$-progressively measurable with $f(t, 0; \omega) \in L^2_{\mathcal{F}}(0, T; N)$, and there exists a constant $L > 0$ such that

$$|f(t, \theta; \omega) - f(t, \bar{\theta}; \omega)| \leq L|\theta - \bar{\theta}|, \quad \forall \theta,\, \bar{\theta} \in M, \text{ a.e. } t \in [0, T], \text{ a.s.} \tag{2.3}$$

- Similarly, $L^2_{\mathcal{F}}(\Omega; W^{1,\infty}(\mathbf{R}^n;\mathbf{R}^m))$ is the set of all functions $g : \mathbf{R}^n \times \Omega \to \mathbf{R}^m$, such that $\omega \mapsto g(x;\omega)$ is $\mathcal{F}_T$-measurable for all $x \in \mathbf{R}^n$ and $x \mapsto g(x;\omega)$ is uniformly Lipschitz in $x \in \mathbf{R}^n$ and $g(0;\omega) \in L_{\mathcal{F}}(\Omega;\mathbf{R}^m)$

The reader familiar with functional analysis will likely recognise $W^{1,\infty}(M;N)$ as the standard notation for the Sobolev space containing all functions $f : M \to N$ which are Lipschitz continuous.

We conclude this section by introducing the space

$$\mathcal{M}[0,T] \triangleq L^2_{\mathcal{F}}(\Omega; C([0,T];\mathbf{R}^n)) \times L^2_{\mathcal{F}}(\Omega; C([0,T];\mathbf{R}^m)) \times L^2_{\mathcal{F}}(0,T;\mathbf{R}^{m\times d}) \tag{2.4}$$

which has the norm

$$\|(X,Y,Z)\| = \left\{E \sup_{t\in[0,T]} |X(t)|^2 + E \sup_{t\in[0,T]} |Y(t)|^2 + E\int_0^T |Z(t)|^2 \,\mathrm{d}t\right\}^{1/2}, \tag{2.5}$$

for all $(X,Y,Z) \in \mathcal{M}[0,T]$. The interested reader may note that $\mathcal{M}[0,T]$ is a Banach space under the presented norm.

2.3 Stochastic Differential Equations

2.3.1 Forward Stochastic Differential Equations

Most readers will be familiar with the following type of stochastic differential equation (SDE)

$$\begin{cases} \mathrm{d}X(t) &= b(t,X(t))\,\mathrm{d}t + \sigma(t,X(t))\,\mathrm{d}W(t) \\ X(0) &= x \end{cases} \tag{2.6}$$

which is a representation of

$$X(t) = x + \int_0^t b(s,X(s))\,\mathrm{d}s + \int_0^t \sigma(s,X(s))\,\mathrm{d}W(s) \tag{2.7}$$

where $\int \mathrm{d}W(s)$ denotes the standard Itô integral. Given some standard regularity conditions on the coefficient functions b and σ, there will exist a unique adapted solution X such that $\int_0^T E|X(t)|^2\,\mathrm{d}t < \infty$ (see [1]). In other words, $X \in L^2_{\mathcal{F}}(0,T;\mathbf{R}^n)$. We shall refer to this type of SDE as a forward stochastic differential equation (FSDE) since it is a stochastic initial value problem. We move forward in time from the initial state x at $t = 0$. The theory of FSDEs is well explored and there are a number of excellent books on the subject, ranging from introductory [1] to advanced [2].

2.3.2 Backward Stochastic Differential Equations

As the title of this paper suggests, we are also interested in a different kind of SDE, namely the backward stochastic differential equation (BSDE). It has the form

$$\begin{cases} \mathrm{d}Y(t) &= h(t, Y(t), Z(t))\,\mathrm{d}t + Z(t)\,\mathrm{d}W(t) \\ Y(T) &= y, \quad t \in [0, T] \end{cases} \tag{2.8}$$

which is a representation of

$$Y(t) = y - \int_t^T h(s, Y(s), Z(s))\,\mathrm{d}s - \int_t^T Z(s)\,\mathrm{d}W(s). \tag{2.9}$$

We call it a backward equation because its terminal value is known and we are moving towards it. Please note that we are still using the Itô integral based on forward increments, although we say that we are going backwards. Instead of looking for a solution in the form of a single adapted process, we are now looking for an adapted pair (Y, Z) as the solution of (2.9).

We will illustrate some motivation for the introduction of the process Z through a very basic example. Consider the stochastic terminal value problem

$$\begin{cases} \mathrm{d}Y(t) &= 0 \\ Y(T) &= \eta. \end{cases} \tag{2.10}$$

We require the solution to be adapted to our usual filtration $\{\mathcal{F}_t\}_{t\geq 0}$. An easy way to find an adapted solution is to consider $Y(t) = E[\eta|\mathcal{F}_t]$. Clearly, our choice of solution is adapted, and it is also a martingale. Hence we may use the martingale representation theorem to write

$$Y(t) = Y(0) + \int_0^t Z(s)\,\mathrm{d}W(s), \quad \forall t \in [0, T], \text{ a.s.} \tag{2.11}$$

Combining (2.10) and $Y(t) = E[\eta|\mathcal{F}_t]$, we may reformulate (2.10) as

$$\begin{cases} \mathrm{d}Y(t) &= Z(t)\,\mathrm{d}W(t), \quad t \in [0, T], \\ Y(T) &= \eta \end{cases} \tag{2.12}$$

and look for an $\{\mathcal{F}_t\}_{t\geq 0}$-adapted solution in the form of a pair (Y, Z). The additional process Z is what makes it possible to find such a solution and it can be shown that (2.12) is the correct way to reformulate our initial problem. For the integral representation we have

$$Y(t) = Y(0) + \int_0^t Z(s)\,\mathrm{d}W(s) = \eta - \int_t^T Z(s)\,\mathrm{d}W(s), \quad \forall t \in [0, T] \tag{2.13}$$

since

$$Y(T) = Y(0) + \int_0^T Z(s)\,\mathrm{d}W(s) \Rightarrow Y(0) = \eta - \int_0^T W(s)\,\mathrm{d}W(s) \tag{2.14}$$

where we use the fact that $Y(T) = \eta$.

2.3.3 Forward Backward Stochastic Differential Equations

A forward backward stochastic differential equation is system of SDEs consisting of a FSDE and a BSDE. To illustrate this, we revisit a familiar example, the *contingent claim valuation problem.*

The problem lies in finding the fair price of a contingent claim, such as a European call option, given its strike price p at time T. In order to do this, we consider a market model consisting of a risk free asset, a bond, with the price dynamics

$$\begin{cases} \mathrm{d}B(t) &= r(t)B(t)\,\mathrm{d}t \\ B(0) &= 1 \end{cases} \tag{2.15}$$

and a risky asset, a stock, with the price dynamics

$$\begin{cases} \mathrm{d}S(t) &= b(t)S(t)\,\mathrm{d}t + \sigma(t)S(t)\,\mathrm{d}W(t) \\ S(0) &= s. \end{cases} \tag{2.16}$$

In the case of the European call, the payoff to the owner of the option is $(S(T)-p)^+$ at the time T. Following standard procedure, we let $Y(t)$ denote the wealth of the investor at time t, wealth which he may choose to invest in bonds, stocks or consumption. At each time t, the investor puts $\pi(t)$ into the stock, $Y(t) - \pi(t)$ into the bond, and has an accumulated consumption of $C(t)$. The dynamics of the wealth $Y(t)$ and the portfolio/consumption pair $(\pi(t), C(t))$ follows an SDE given by

$$\begin{cases} \mathrm{d}Y(t) &= [r(t)Y(t) + Z(t)\theta(t)]\,\mathrm{d}t + Z(t)\,\mathrm{d}W(t) - \,\mathrm{d}C(t) \\ Y(0) &= y \end{cases} \tag{2.17}$$

where $\theta(t) \triangleq \sigma^{-1}(t)[b(t) - r(t)]$ is the *risk premium* and $Z(t) = \pi(t)\sigma(t)$.

In order to find the fair price of the European call option, we look for an portfolio/consumption pair $(\pi(t), C(t))$ such that $Y(T) = (S(T) - p)^+$ If we consider an investor who refrains from consumption ($C(t) \equiv 0$), we can find the appropriate portfolio by solving

$$\begin{cases} \mathrm{d}S(t) &= b(t)S(t)\,\mathrm{d}t + \sigma(t)S(t)\,\mathrm{d}W(t) \\ \mathrm{d}Y(t) &= [r(t)Y(t) + Z(t)\theta(t)]\,\mathrm{d}t + Z(t)\,\mathrm{d}W(t) \\ S(0) &= s, \quad Y(T) \;=\; (S(T)-p)^+, \quad t \in [0,T] \end{cases} \tag{2.18}$$

which is a *forward backward stochastic differential equation.* It is forward in the stock price $S(t)$, and backward in the wealth process $Y(t)$. This example exhibits a *decoupled* FBSDE, as the coefficients of the forward equation are independent of $Y(t)$ and $Z(t)$.

We will however look at more general examples, where the equations are fully *coupled*, i.e. there is explicit dependence on $Y(t)$ and $Z(t)$ in the forward equation, and there is explicit dependence on $X(t)$ in the backward equation. Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbf{P})$ be a filtered probability space satisfying our usual conditions. We are mainly interested in problems of the form

$$\begin{cases} \mathrm{d}X(t) &= b(t, X(t), Y(t), Z(t))\,\mathrm{d}t + \sigma(t, X(t), Y(t), Z(t))\,\mathrm{d}W(t) \\ \mathrm{d}Y(t) &= h(t, X(t), Y(t), Z(t))\,\mathrm{d}t + Z(t)\,\mathrm{d}W(t) \\ X(0) &= x, \quad Y(T) = g(X(T)), \quad t \in [0, T] \end{cases} \tag{2.19}$$

where b, σ, h and g are functions of suitable dimensions satisfying what we will refer to as our *standing assumptions*

$$\begin{cases} b \in L^2_{\mathcal{F}}(0, T; W^{1,\infty}(M; \mathbf{R}^n)), \quad \sigma \in L^2_{\mathcal{F}}(0, T; W^{1,\infty}(M; \mathbf{R}^{n\times d})) \\ h \in L^2_{\mathcal{F}}(0, T; W^{1,\infty}(M; \mathbf{R}^m)), \quad g \in L^2_{\mathcal{F}}(\Omega; W^{1,\infty}(\mathbf{R}^n; \mathbf{R}^m)) \end{cases} \tag{2.20}$$

where $M = \mathbf{R}^n \times \mathbf{R}^m \times \mathbf{R}^{m\times d}$. As mentioned before, our standing assumptions need not be memorised. It is sufficient to remember that they assure us that the functions we are working with are suitably well behaved.

It is possible to generalise our problem further to get an FBSDE in the form

$$\begin{cases} \mathrm{d}X(t) &= b(t, X(t), Y(t), Z(t))\,\mathrm{d}t + \sigma(t, X(t), Y(t), Z(t))\,\mathrm{d}W(t) \\ \mathrm{d}Y(t) &= h(t, X(t), Y(t), Z(t))\,\mathrm{d}t + \hat{\sigma}(t, X(t), Y(t), Z(t))\,\mathrm{d}W(t) \\ X(0) &= x, \quad Y(T) = g(X(T)), \quad t \in [0, T] \end{cases} \tag{2.21}$$

but for (2.21) to have an adapted solution, we require

$$\{\hat{\sigma}(0, x, y, z) | z \in \mathbf{R}^l\} = \mathbf{R}^{m\times d}, \quad \forall (x, y) \in \mathbf{R}^n \times \mathbf{R}^m, \text{ a.s.} \tag{2.22}$$

See [3] for further discussion on the subject. We will accept this as a fact and solve it in pleasingly simple way through setting

$$\hat{\sigma}(t, x, y, z) \equiv z, \quad \forall (t, x, y) \in [0, T] \times \mathbf{R}^n \times \mathbf{R}^m, \text{ a.s.} \tag{2.23}$$

with $z \in \mathbf{R}^{m\times d}$. We hence sacrifice generality for simplicity.

To end the section we present the solution to our equation.

Definition 2.6. *A triple* $(X, Y, Z) \in \mathcal{M}[0, T]$ *is called an* adapted solution *of* (2.19) *if the following holds for any* $t \in [0, T]$*, a.s.*

$$\begin{cases} X(t) &= x + \int_0^T b(s, X(s), Y(s), Z(s))\,\mathrm{d}s \\ & \quad + \int_0^T \sigma(s, X(s), Y(s), Z(s))\,\mathrm{d}W(s) \\ Y(t) &= g(X(T)) - \int_t^T h(s, X(s), Y(s), Z(s))\,\mathrm{d}s \\ & \quad - \int_t^T Z(s)\,\mathrm{d}W(s) \end{cases} \tag{2.24}$$

If (2.19) has an adapted solution, we say that it is *solvable*, otherwise we say that it is *non solvable*.

Chapter 3

A brief word on solvability

Although this paper will focus mainly on how to solve FBSDEs, it is of course important and interesting to know when a FBSDE is solvable. Also, the key to understanding the heuristic derivation of our future algorithm for solving FBSDEs lies within this chapter. We are interested in solving FBSDEs of the form

$$\begin{cases} \mathrm{d}X(t) &= b(t, X(t), Y(t), Z(t))\,\mathrm{d}t + \sigma(t, X(t), Y(t), Z(t))\,\mathrm{d}W(t) \\ \mathrm{d}Y(t) &= h(t, X(t), Y(t), Z(t))\,\mathrm{d}t + Z(t)\,\mathrm{d}W(t) \\ X(0) &= x,\ Y(T) = g(X(T)), \quad t \in [0, T] \end{cases} \tag{3.1}$$

where the triple $(X, Y, Z) \in \mathcal{M}[0, T]$ and the functions b, σ, h and g are deterministic satisfying our normal regularity and smoothness conditions. We are interested in the question of solvabiliy of (3.1) over any finite time duration $[0, T]$. To tackle this problem, we shall introduce a method using optimal control theory. In agreement with the title of the chapter, the treatment will be brief. Hopefully, it will however give us a heuristic understanding of the solvability of our chosen problem.

3.1 Method of optimal control

Consider the following Forward Stochastic Differential Equation (FSDE)

$$\begin{cases} \mathrm{d}X(t) &= b(t, X(t), Y(t), Z(t))\,\mathrm{d}t + \sigma(t, X(t), Y(t), Z(t))\,\mathrm{d}W(t) \\ \mathrm{d}Y(t) &= h(t, X(t), Y(t), Z(t))\,\mathrm{d}t + Z(t)\,\mathrm{d}W(t) \\ X(0) &= x,\ Y(0) = y, \quad t \in [0, T] \end{cases} \tag{3.2}$$

We shall impose the following condition upon our coefficient functions

(H1) Functions $b(t, x, y, z)$, $\sigma(t, x, y, z)$, $h(t, x, y, z)$ and $g(x)$ are all continuous and there exists a constant $L > 0$ such that for g and $\phi = b$, σ, h it holds that

$$\begin{cases} |\phi(t,x,y,z) - \phi(t,\bar{x},\bar{y},\bar{z})| \;\leq\; L(|x-\bar{x}| + |y-\bar{y}| + |z-\bar{z}|) \\ |\phi(t,0,0,0)|, |\sigma(t,x,y,0)|, |g(x)| \;\leq\; L \\ \forall t \in [0,T],\ x,\bar{x} \in \mathbf{R}^n,\ y,\bar{y} \in \mathbf{R}^m,\ z,\bar{z} \in \mathbf{R}^{n\times m}. \end{cases} \tag{3.3}$$

If we assume (H1) and that $Z \in \mathcal{Z}[0,T] = L^2_{\mathcal{F}}(0,T;\mathbf{R}^{m\times d})$ is given, equation (3.2) has a strong unique solution (X,Y) (see [1]). If we can somehow choose y and Z such that $Y(T) = g(X(T))$, (X,Y,Z) will be an adapted solution to our initial FBSDE (3.1). Hence, (3.1) being solvable is equivalent to the existence of $y \in \mathbf{R}^m$ and $Z \in \mathcal{Z}[0,T]$ such that (3.2) has a strong solution (X,Y) satisfying $Y(T) = g(X(T))$.

We will now rephrase our problem in the language of stochastic optimal control. We call (3.2) a *stochastic control system*, with *state process* (X,Y) and *control process* Z. We start in *initial state* $(x,y) \in \mathbf{R}^n \times \mathbf{R}^m$ and have the *target* $\mathcal{T} = \{(x,g(x))|x \in \mathbf{R}^n\}$.

The question of solvability of our initial FBSDE has now been rephrased to finding a $y \in \mathbf{R}^m$ such that (X,Y) can be steered from the initial state (x,y) at $t=0$ to the target at $t=T$, i.e. $(X(T),Y(T)) \in \mathcal{T}$ a.s., by a suitable control process $Z \in \mathcal{Z}[0,T]$. This is referred to as a *controllability problem*. Unfortunately, the controllability problem is very hard to solve if the coefficient functions are non-linear. Fortunately, we can use our controllability problem to formulate an *optimal control problem*, which decomposes the difficult task above into smaller, relatively easier parts.

3.1.1 The optimal cost function

Let us assume (H1) and $Z \in \mathcal{Z}[0,T]$. As mentioned above, (3.2) will have a strong solution $(X(t),Y(t))$ dependent on the choice of initial state (x,y) and control process Z. To emphasise this dependence, we shall write $(X(t;x,y,Z),Y(t;x,y,Z))$ for such a solution. We introduce the following function

$$f(x,y) = \sqrt{1+|y-g(x)|^2} - 1, \quad \forall (x,y) \in \mathbf{R}^n \times \mathbf{R}^m \tag{3.4}$$

The function f can be interpreted as it penalises any difference between y and $g(x)$. If the difference is large, so is f. Obviously, f satisfies

$$\begin{cases} f(x,y) \geq 0, \quad \forall (x,y) \in \mathbf{R}^n \times \mathbf{R}^m \\ f(x,y) = 0 \Longleftrightarrow y = g(x). \end{cases} \tag{3.5}$$

Given f, we define the *cost functional* as follows

$$J(x,y;Z) \triangleq E\left[f(X(T;x,y,Z),Y(T;x,y,Z))\right]. \tag{3.6}$$

We now present the optimal control problem associated with (3.1):

For any given intitial state $(x,y) \in \mathbf{R}^n \times \mathbf{R}^m$, find a $\bar{Z} \in \mathcal{Z}[0,T]$ s.t.

$$\bar{V}(x,y) \triangleq \inf_{Z\in\mathcal{Z}[0,T]} J(x,y;Z) = J(x,y;\bar{Z}). \tag{3.7}$$

In other words, we want to find a Z which minimises the cost functional for a given initial state. We call $\bar{V}(x,y)$ the *optimal cost function* and $\bar{Z}$ an *optimal control*. For each function $\bar{V}: \mathbf{R}^n \times \mathbf{R}^m \to \mathbf{R}^+$, we define the nodal set $\mathcal{N}(\bar{V})$ as the set of all initial states $(x,y) \in \mathbf{R}^n \times \mathbf{R}^m$ such that $\bar{V}(x,y) = 0$. In other words

$$\mathcal{N}(\bar{V}) = \{(x,y) \in \mathbf{R}^n \times \mathbf{R}^m | \bar{V}(x,y) = 0\}. \tag{3.8}$$

It is apparent that the solvability of (3.1) is deeply linked to the contents of the nodal set of the corresponding cost function. If for $x \in \mathbf{R}^n$ given in our initial FBSDE (3.1) we have

$$\mathcal{N}(\bar{V}) \cap [\{x\} \times \mathbf{R}^m] \neq \emptyset, \tag{3.9}$$

there will exist an optimal pair (X,Y) which can be steered to the target $(X(T), g(X(T)))$ by a control $Z \in \mathcal{Z}[0,T]$. In other words, $(X,Y,Z) \in \mathcal{M}[0,T]$ will be an adapted solution to our initial problem.

Here we see the first hint of how to actually solve the type of FBSDE problems we are interested in, given that the nodal set satisfies (3.9). We have decomposed our initial problem into three lesser ones:

i) Find the optimal cost function $\bar{V}(x,y)$

ii) Find the nodal set of $\bar{V}$, and restrict $x \in \mathbf{R}^n$ as given in the initial problem.

iii) For the given $x \in \mathbf{R}^n$, find a $y \in \mathbf{R}^m$ such that $(x,y) \in \mathcal{N}(\bar{V})$. Find an optimal control $Z \in \mathcal{Z}[0,T]$ with initial state (x,y) which minimises the cost function. $(X(t),Y(t),Z(t)) = (X(t;x,y,Z), Y(t;x,y,Z), Z(t)) \in \mathcal{M}[0,T]$ is thus an adapted solution to our initial FBSDE (3.1).

3.1.2 The value function

Consider the following (controlled) system

$$\begin{cases} \mathrm{d}X(t) &= b(t,X(t),Y(t),Z(t))\,\mathrm{d}t + \sigma(t,X(t),Y(t),Z(t))\,\mathrm{d}W(t) \\ \mathrm{d}Y(t) &= h(t,X(t),Y(t),Z(t))\,\mathrm{d}t + Z(t)\,\mathrm{d}W(t) \\ X(s) &= x,\ Y(s) = y, \quad t \in [s,T) \end{cases} \tag{3.10}$$

Under assumption (H1), for any given $(s,x,y) \in [0,t) \times \mathbf{R}^n \times \mathbf{R}^m$ and $Z \in \mathcal{Z}[s,T] \triangleq L^2_{\mathcal{F}}(s,T;\mathbf{R}^{m\times d})$, FSDE (3.10) admits a unique strong solution

$$(X(t),Y(t)) = (X(t;s,x,y,Z), Y(t;s,x,y,Z)). \tag{3.11}$$

In this case, the solution will also depend on the choice of initial time s. Similarly to our previous cost functional, we define

$$J(s,x,y;Z) \triangleq E[f(X(t;s,x,y,Z),Y(t;s,x,y,Z))]. \tag{3.12}$$

In this setting, we present the parametrised optimal control problem associated with (3.1):

For any given $(s,x,y) \in [s,T) \times \mathbf{R}^n \times \mathbf{R}^m$, find a $\bar{Z} \in \mathcal{Z}[s,T]$ such that

$$V(s,x,y) \triangleq \inf_{Z \in \mathcal{Z}[0,T]} J(s,x,y;Z) = J(s,x,y;\bar{Z}) \tag{3.13}$$

and define

$$V(T,x,y) \triangleq f(x,y), \quad \forall (x,y) \in \mathbf{R}^n \times \mathbf{R}^m \tag{3.14}$$

where f is as defined in (3.4). V is called the *value function* of the optimal control problem parametrised by s. Obviously

$$V(0,x,y) = \bar{V}(x,y), \quad \forall (x,y) \in \mathbf{R}^n \times \mathbf{R}^m. \tag{3.15}$$

This method of trying to find a control for each $s \in [0,T)$ is called the *dynamic programming method* due to the obvious fact that the control may change as the parameter s changes. If we could find a way of determining the value function, we would know a great deal about the optimal cost function and the solvability of our system. The good news is that the value function is the solution of a PDE, the Hamilton-Jacobi-Bellman (HJB) equation associated with our optimal control problem. Hence, we could, theoretically, find $V(s,x,y)$ by solving the associated HJB equation. The bad news is that this is generally very hard, and we will not delve deeper into the subject. For an introduction to the HJB equation, please see [1], chapter 11.

Apparently, we must find another way of finding information about $V(s,x,y)$ and $\bar{V}(x,y)$. To do this, we return to the nodal set.

As in our previous case with the optimal cost function, we define the nodal set of V:

$$\mathcal{N}(V) = \{(s,x,y) \in [0,T) \times \mathbf{R}^n \times \mathbf{R}^m | V(s,x,y) = 0\} \tag{3.16}$$

Clearly, we have

$$\mathcal{N}(V) \cap \{0\} \times \mathbf{R}^n \times \mathbf{R}^m = \{0\} \times \mathcal{N}(\bar{V}). \tag{3.17}$$

Hence, by studying the set $\mathcal{N}(V)$, we find information about $\mathcal{N}(\bar{V})$. In order to find information about the contents of $\mathcal{N}(V)$, suppose that there is exists a function $\theta : [0,T] \times \mathbf{R}^n \to \mathbf{R}^m$ such that

$$V(s,x,\theta(s,x)) = 0, \quad \forall (s,x) \in [0,T] \times \mathbf{R}^n. \tag{3.18}$$

We obviously have

$$\{(s,x,\theta(s,x)) | (s,x) \in [0,T] \times \mathbf{R}^n\} \subseteq \mathcal{N}(V) \tag{3.19}$$

and in particular

$$(x, \theta(0, x)) \in \mathcal{N}(\bar{V}), \quad \forall x \in \mathbf{R}^n. \tag{3.20}$$

The function θ plays an important role in understanding our problem. If it exists, it shows that $\mathcal{N}(\bar{V})$ is non-empty, and hence that our initial FBSDE is solvable. If we can explicitly find such a function, it will help us describe the nodal set and the relation between any pair (x, y) s.t. $y = \theta(0, x)$.

For further information on mathematical control theory, the reader may refer to [4]

Chapter 4

The four step scheme

4.1 An algorithmic approach to solving Forward Backward Stochastic Differential equations.

In this chapter we shall introduce an algorithm consisting of four steps which, when applicable, explicitly solves FBSDEs. We shall start out by giving a heuristic argument based our previous result on solvability and standard methods from stochastic analysis. After this, we proceed to investigate when this scheme is applicable, in other words when we can actually solve the type of systems we are interested in. As we shall see, this relies heavily on the theory of partial differential equations. This will come as no surprise to the reader familiar with stochastic analysis, since the link to partial differential equations seem to permeate the subject. We must unfortunately limit ourselves to FBSDEs with deterministic coefficients in order not to stray beyond the scope of this paper. We shall, however, briefly mention the more general case at the end of this section.

4.2 A heuristic derivation of the four step scheme.

We are interested in solving the following FBSDE:

$$\begin{cases} \mathrm{d}X(t) &=& b(t,X(t),Y(t),Z(t))\,\mathrm{d}t + \sigma(t,X(t),Y(t),Z(t))\,\mathrm{d}W(t) \\ \mathrm{d}Y(t) &=& h(t,X(t),Y(t),Z(t))\,\mathrm{d}t + Z(t)\,\mathrm{d}W(t) \\ X(0) &=& x,\ Y(T) \ = \ g(X(T)), \quad t \in [0,T]. \end{cases} \tag{4.1}$$

The functions b, σ, h and g are all reasonably well behaved and deterministic, $\{W(t)\}_{t\geq 0}$ is a standard d-dimensional Brownian motion, and (X,Y,Z) takes values in $\mathbf{R}^n \times \mathbf{R}^m \times \mathbf{R}^{m\times d}$. In order to simplify notation we will often suppress the explicit t-dependence of the processes X, Y and Z, writing $X(t) = X$ etc.

Based on our previous observations on solvability, we assume that $Y(t) = \theta(t,X)$. We shall see that θ is the solution to a parabolic system of PDEs. We further assume that

$\theta \in C^{1,2}([0,T], \mathbf{R}^n)$, in order to be able to apply the Itô formula. Please recall the Itô formula in the multidimensional setting:

$$\mathrm{d}Y^k(t) = d\theta^k(t,X) = \frac{\partial \theta^k(t,X)}{\partial t}\,\mathrm{d}t + \sum_{i=1}^{n} \frac{\partial \theta^k(t,X)}{\partial x_i}\,\mathrm{d}X^i + \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2 \theta^k(t,X)}{\partial x_i \partial x_j}\,\mathrm{d}X^i\,\mathrm{d}X^j \tag{4.2}$$

where k denotes the k-th element of the m-dimensional random vector $Y(t)$ etc. If we instead use subscripts to denote partial derivatives, and apply the multidimensional Itô formula to our problem, we have

$$\begin{aligned}
\mathrm{d}Y^k(t) &= \mathrm{d}\theta^k(t,X) \\
&= \Big[\theta^k_t(t,X) + \sum_{i=1}^{n} b_i(t,X,\theta(t,X),Z)\,\theta^k_{x_i}(t,X) \\
&\quad + \frac{1}{2}\sum_{i,j=1}^{n} \theta^k_{x_i x_j}(t,X) \sum_{k=1}^{d} \sigma_{ik}(t,X,\theta(t,X),Z)\sigma_{jk}(t,X,\theta(t,X),Z)\Big]\,\mathrm{d}t \\
&\quad + \sum_{i=1}^{n} \theta^k_{x_i}(t,X)\,\langle \sigma_i(t,X,\theta(t,X),Z),\,\mathrm{d}W(t)\rangle
\end{aligned} \tag{4.3}$$

where $\sigma_i(t,X,Y,Z)$ is the vector consisting of the i-th row of the matrix $\sigma(t,X,Y,Z) \in \mathbf{R}^{n\times d}$. Let us introduce the following notation; by $\theta^k_x(t,X)$ we denote the n-dimensional vector $\{\,\theta^k_{x_1}(t,X),\ldots,\theta^k_{x_n}(t,X)\}$, and by $\theta^k_{xx}(t,X)$ we denote the $n\times n$ matrix $(a_{ij}) = \theta^k_{x_i x_j}(t,X)$. We thus have a more convenient form for (4.3)

$$\begin{aligned}
\mathrm{d}Y^k(t) &= \mathrm{d}\theta^k(t,X)) \\
&= \Big\{\theta^k_t(t,X) + \big\langle \theta^k_x(t,X), b(t,X,\theta(t,X),Z)\big\rangle \\
&\quad + \frac{1}{2}\mathrm{tr}\left[\theta^k_{xx}(t,X)(\sigma\sigma^T)(t,X,\theta(t,X),Z)\right]\Big\}\,\mathrm{d}t \\
&\quad + \big\langle \theta^k_x(t,X), \sigma(t,X,\theta(t,X),Z)\,\mathrm{d}W\big\rangle.
\end{aligned} \tag{4.4}$$

Recall that $\langle\cdot\,,\cdot\rangle$ denotes ordinary vector inner product and tr denotes trace of a matrix.

Comparing the coefficients from (4.1) and (4.4), given that $Y(t) = \theta(t,X)$, we must have the following equation

$$\begin{cases}
h^k(t,X,\theta(t,X),Z) = \theta^k_t(t,X) + \big\langle \theta^k_x(t,X), b(t,X,\theta(t,X),Z)\big\rangle \\
\qquad\qquad + \frac{1}{2}\mathrm{tr}\left[\theta^k_{xx}(t,X)(\sigma\sigma^T)(t,X,\theta(t,X)),Z)\right] \\
\theta(T,X(T)) = g(X(T))
\end{cases} \tag{4.5}$$

and

$$\theta_x(t,X)\,\sigma(t,X,\theta(t,X),Z) = Z(t) \tag{4.6}$$

If we work our way backwards through the steps above, we should be able to find solution to our initial FBSDE. Inspired by this, we now introduce the formal algorithm.

4.3 The Four Step Scheme

Throughout this section we suppress the dependence of θ on t and x, in other words $\theta = \theta(t,x)$, $\theta_x = \theta_x(t,x)$ and $\theta_{xx} = \theta_{xx}(t,x)$.

Step 1. Find a function $z(t,x,y,p)$ that satisfies the following

$$\begin{aligned} z(t,x,y,p) = p\,\sigma(t,x,y,z(t,x,y,p)) \\ \forall (t,x,y,p) \in [0,T] \times \mathbf{R}^n \times \mathbf{R}^m \times \mathbf{R}^{n\times m}. \end{aligned} \tag{4.7}$$

Obviously, the function σ is known from our initial problem (4.1).

Step 2. Using the function z found in step 1, solve the parabolic system for $1 \le k \le m$

$$\begin{cases} \theta_t^k + \frac{1}{2}\mathrm{tr}[\theta_{xx}^k \sigma\sigma^T(t,x,\theta,z(t,x,\theta,\theta_x))] + \left\langle b(t,x,\theta,z(t,x,\theta,\theta_x)), \theta_x^k \right\rangle \\ \qquad -h^k(t,x,\theta,z(t,x,\theta,\theta_x)) = 0, \quad (t,x) \in [0,T] \times \mathbf{R}^n \\ \theta(T,x) = g(x), \quad x \in \mathbf{R}^n \end{cases} \tag{4.8}$$

Step 3. Again, using the functions z and θ obtained in steps 1 and 2, solve the following FSDE

$$\begin{cases} \mathrm{d}X(t) &= \tilde{b}(t,X)\,\mathrm{d}t + \tilde{\sigma}(t,X)\,\mathrm{d}W(t) \\ X(0) &= x, \quad t \in [0,T] \end{cases} \tag{4.9}$$

with

$$\begin{cases} \tilde{b}(t,x) &= b(t,x,\theta,z(t,x,\theta,\theta_x)) \\ \tilde{\sigma}(t,x) &= \sigma(t,x,\theta,z(t,x,\theta,\theta_x)) \end{cases} \tag{4.10}$$

For the final step, we combine our previous results.

Step 4. Set

$$\begin{aligned} Y(t) &= \theta(t,X) \\ Z(t) &= z(t,X,\theta(t,X),\theta_x(t,X)) \end{aligned} \tag{4.11}$$

Whenever we are able to realise all four of the steps above, X found in step 3 together with Y and Z from step 4 form a triple (X,Y,Z) which is an adapted solution to (4.1). We now present the main result of this chapter.

Theorem 4.1. *Let* (4.7) *admit a unique solution* $z(t,x,y,p)$ *which is uniformly Lipschitz continuous in* (x,y,p) *with* $z(t,0,0,0)$ *being bounded. Let* (4.8) *admit a classical solution* $\theta(t,x)$ *with bounded* θ_x *and* θ_{xx}. *Let functions* b *and* σ *be uniformly Lipschitz continuous in* (x,y,z) *with* $b(t,0,0,0)$ *and* $\sigma(t,0,0,0)$ *both being bounded. Then the process* (X,Y,Z) *determined by* (4.7) - (4.11) *is an adapted solution to* (4.1). *Moreover, if* h *is also uniformly Lipschitz continuous in* (x,y,z), σ *is bounded, and there exists a constant* $\beta \in (0,1)$, *such that*

$$\begin{aligned} \left|[\sigma(s,x,y,z)-\sigma(s,x,y,\tilde{z})]^T\theta_x(s,x)\right| &\le \beta\,|z-\tilde{z}| \\ \forall (s,x,y)\in[0,T]\times\mathbf{R}^n\times\mathbf{R}^m,\ \ &z,\tilde{z}\in\mathbf{R}^{m\times d}, \end{aligned} \tag{4.12}$$

then the adapted solution (X,Y,Z) *above is unique.*

Proof. To prove adaptedness of our solution, we turn to (4.9). The coefficient functions $\tilde{b}$ and $\tilde{\sigma}$ defined in (4.10) are both Lipschitz continuous in x, given the assumptions made in theorem 4.1. This implies that (4.9) has a unique strong solution (see [1]). Now define Y and Z according to *Step 4* in the algorithm above. We already know from the heuristic derivation of the four step scheme that any such triple (X,Y,Z) is a solution of our initial equation (4.1).

Left is to prove that the solution is unique. In order to do this we claim that every solution of (4.1) is of the form constructed in the four step scheme. We will show this for the case $m=1$, i.e. when $Y(t)$ is a one dimensional process. The result can be generalised for any $m\in\mathbf{N}$. To show this, let (X,Y,Z) be any adapted solution to (4.1). Now define new processes $\tilde{Y}$ and $\tilde{Z}$ by

$$\tilde{Y}(t)=\theta(t,X), \qquad \tilde{Z}(t)=z(t,X,\theta(t,X),\theta_x(t,X)). \tag{4.13}$$

By assumption, equation (4.7) admits a unique solution and we have

$$\tilde{Z}(t)=z(t,X,\theta(t,X),\theta_x(t,X))=\theta_x(t,X)\,\sigma(t,X,\tilde{Y},\tilde{Z}), \quad t\in[0,T],\ \text{a.s.} \tag{4.14}$$

If we apply Itô's formula to $\tilde{Y}(t)=\theta(t,X)$, we know from (4.4) that we will have the following dynamics

$$\begin{aligned} \mathrm{d}\tilde{Y}(t) &= \mathrm{d}\theta(t,X) \\ &= \Big\{\theta_t(t,X)+\langle\theta_x(t,X),b(t,X,Y,Z)\rangle \\ &\qquad +\frac{1}{2}\mathrm{tr}[\theta_{xx}(t,X)\sigma\sigma^T(t,X,Y,Z)]\Big\}\,\mathrm{d}t \\ &\quad +\ \langle\theta_x(t,X),\sigma(t,X,Y,Z)\,\mathrm{d}W(t)\rangle. \end{aligned} \tag{4.15}$$

With (4.13) in mind, (4.8) becomes

$$\begin{aligned} &\theta_t(t,X)+\frac{1}{2}\mathrm{tr}[\theta_{xx}(t,X)\sigma\sigma^T(t,X,\tilde{Y},\tilde{Z})] \\ &\qquad +\langle b(t,X,\tilde{Y},\tilde{Z}),\theta_x(t,X)\rangle-h(t,X,\tilde{Y},\tilde{Z})=0. \end{aligned} \tag{4.16}$$

Subtracting (4.16) multiplied by $\mathrm{d}t$ from (4.15) yields

$$\begin{aligned} \mathrm{d}\tilde{Y}(t) &= \Big\{\langle\theta_x(t,X), b(t,X,Y,Z) - b(t,X,\tilde{Y},\tilde{Z})\rangle \\ &+ \frac{1}{2}\mathrm{tr}[\theta_{xx}(t,X)\{\sigma\sigma^T(t,X,Y,Z) - \sigma\sigma^T(t,X,\tilde{Y},\tilde{Z})\}] \\ &+ h(t,X,\tilde{Y},\tilde{Z})\Big\}\,\mathrm{d}t + \langle\theta_x(t,X), \sigma(t,X,Y,Z)\,\mathrm{d}W(t)\rangle. \end{aligned} \tag{4.17}$$

Now apply the Itô formula to $|\tilde{Y}(t) - Y(t)|^2$

$$\mathrm{d}|\tilde{Y}(t) - Y(t)|^2 = 2(\tilde{Y} - Y)(\,\mathrm{d}\tilde{Y} - \,\mathrm{d}Y) + \frac{1}{2}(\,\mathrm{d}\tilde{Y} - \,\mathrm{d}Y)^2. \tag{4.18}$$

With the dynamics from (4.1) for $Y(t)$ and from (4.17) for $\tilde{Y}(t)$ we get

$$\begin{aligned} \mathrm{d}|\tilde{Y}(t) - Y(t)|^2 &= 2(\tilde{Y} - Y)\cdot\Big[\Big\{\langle\theta_x(t,X), b(t,X,Y,Z) - b(t,X,\tilde{Y},\tilde{Z})\rangle \\ &+ \frac{1}{2}\mathrm{tr}[\theta_{xx}(t,X)\{\sigma\sigma^T(t,X,Y,Z) - \sigma\sigma^T(t,X,\tilde{Y},\tilde{Z})\}] \\ &+ h(t,X,\tilde{Y},\tilde{Z}) - h(t,X,Y,Z)\Big\}\,\mathrm{d}t \\ &+ \langle\sigma^T(t,X,Y,Z)\theta_x(t,X) - Z(t),\,\mathrm{d}W(t)\rangle\Big] \\ &+ \Big|\sigma^T(t,X,Y,Z)\theta_x(t,X) - Z(t)\Big|^2\,\mathrm{d}t. \end{aligned} \tag{4.19}$$

The final term stems from

$$\begin{aligned} (\,\mathrm{d}\tilde{Y} - \,\mathrm{d}Y)^2 &= (\langle\theta_x(t,X), \sigma(t,X,Y,Z)\,\mathrm{d}W(t)\rangle - \langle Z(t),\,\mathrm{d}W(t)\rangle)^2 \\ &= (\langle\sigma^T(t,X,Y,Z)\theta_x(t,X) - Z(t),\,\mathrm{d}W(t)\rangle)^2 \\ &= \Big|\sigma^T(t,X,Y,Z)\theta_x(t,X) - Z(t)\Big|^2\,\mathrm{d}t. \end{aligned} \tag{4.20}$$

By utilising the standard technique of taking conditional expectation we get

$$\begin{aligned} E|\tilde{Y}(t) - Y(t)|^2 &= \tilde{Y}(T) - Y(T) + E\int_T^t \Big\{2(\tilde{Y} - Y)\cdot \\ &\Big[\langle\theta_x(s,X), b(s,X,Y,Z) - b(s,X,\tilde{Y},\tilde{Z})\rangle \\ &+ \frac{1}{2}\mathrm{tr}[\theta_{xx}(s,X)\{\sigma\sigma^T(s,X,Y,Z) - \sigma\sigma^T(s,X,\tilde{Y},\tilde{Z})\}] \\ &+ h(s,X,\tilde{Y},\tilde{Z}) - h(s,X,Y,Z)\Big] \\ &+ \Big|\sigma(s,X,Y,Z)^T\theta_x(s,X) - Z(s)\Big|^2\Big\}\,\mathrm{d}s. \end{aligned} \tag{4.21}$$

Since $\tilde{Y}(T) = \theta(T,X(T)) = g(X(T)) = Y(T)$, we are left with the following expression for the expectation

$$
\begin{aligned}
E|\tilde{Y}(t) - Y(t)|^2 \;=\; & -E\int_t^T \Big\{2(\tilde{Y} - Y)\cdot \\
& \Big[\langle \theta_x(s,X), b(s,X,Y,Z) - b(s,X,\tilde{Y},\tilde{Z})\rangle \\
& + \frac{1}{2}\mathrm{tr}[\theta_{xx}(s,X)\{\sigma\sigma^T(s,X,Y,Z) - \sigma\sigma^T(s,X,\tilde{Y},\tilde{Z})\}] \\
& + h(s,X,\tilde{Y},\tilde{Z}) - h(s,X,Y,Z)\Big] \\
& + \big|(\sigma(s,X,Y,Z)^T\theta_x(s,X) - Z(s)\big|^2\Big\}\,\mathrm{d}s.
\end{aligned}
\tag{4.22}
$$

We can approximate the final term by

$$
\begin{aligned}
& \Big|\{\sigma(s,X,Y,Z) - \sigma(s,X,\tilde{Y},\tilde{Z})\}^T\theta_x(s,X) + \tilde{Z}(s) - Z(s)\Big|^2 \\
& \geq (|\tilde{Z}(s) - Z(s)| - |\{\sigma(s,X,Y,Z) - \sigma(s,X,\tilde{Y},\tilde{Z}\}^T\theta_x(s,X)|)^2 \\
& \geq (1-\beta_1)|\tilde{Z}(s) - Z(s)|^2 - C_{\beta_1}|\{\sigma(s,X,Y,Z) - \sigma(s,X,\tilde{Y},\tilde{Z}\}^T\theta_x(s,X)|^2
\end{aligned}
\tag{4.23}
$$

where $C_{\beta_1} = (1/\beta_1 - 1) > 0$ for all $\beta_1 \in (0,1)$. In the first line we have used the relationship (4.14) to add the extra terms. For the first inequality, we have used the reverse triangle inequality, and for the second we have used Young's inequality with $\epsilon = \beta_1$.

Similarly, we have

$$
\begin{aligned}
\big|\{\sigma(s,X,Y,Z) - \sigma(s,X,\tilde{Y},\tilde{Z}\}^T\theta_x(s,X)\big| &\leq C_1|\tilde{Y} - Y| + \beta|\tilde{Z} - Z| \\
\big|\langle\theta_x(s,X), b(s,X,Y,Z) - b(s,X,\tilde{Y},\tilde{Z})\rangle\big| &\leq C_1|\tilde{Y} - Y| + C_2|\tilde{Z} - Z| \\
\mathrm{tr}[\theta_{xx}(s,X)\{\sigma\sigma^T(s,X,Y,Z) - \sigma\sigma^T(s,X,\tilde{Y},\tilde{Z})\}] &\leq C_1|\tilde{Y} - Y| + C_2|\tilde{Z} - Z| \\
\big|h(s,X,\tilde{Y},\tilde{Z}) - h(s,X,Y,Z)\big| &\leq C_1|\tilde{Y} - Y| + C_2|\tilde{Z} - Z|
\end{aligned}
\tag{4.24}
$$

where we have used the boundedness of θ_x and θ_{xx} together with the Lipschitz continuity of the functions b, σ and h, and the condition (4.12). Please note that C_1 and C_2 are generic constants which may vary from line to line. Combining the results above, we have the following approximation

$$
\begin{aligned}
& E|\tilde{Y}(t) - Y(t)|^2 + (1-\beta)\int_t^T E|\tilde{Z}(s) - Z(s)|^2\,\mathrm{d}s \\
\leq\; & C\int_t^T E\big[|\tilde{Y}(s) - Y(s)|(|\tilde{Y}(s) - Y(s)| + |\tilde{Z}(s) - Z(s)|)\big]\,\mathrm{d}s \\
& \qquad + C_{\beta_1}\int_t^T E[(C|\tilde{Y}(s) - Y(s)| + \beta|\tilde{Z}(s) - Z(s)|)^2]\,\mathrm{d}s \\
\leq\; & C_\epsilon\int_t^T E[|\tilde{Y}(s) - Y(s)|^2]\,\mathrm{d}s + (C_{\beta_1}\beta^2 + \epsilon)\int_t^T E[|\tilde{Z}(s) - Z(s)|^2]\,\mathrm{d}s
\end{aligned}
\tag{4.25}
$$

where we have used Young's inequality on $(C + 2C_{\beta_1}C\beta)|\tilde{Z}(s) - Z(s)||\tilde{Y}(s) - Y(s)|$. C is an arbitrary constant, C_ϵ is a constant dependent on the constant $\epsilon > 0$. If we choose the arbitrary constants β_1 and ϵ such that $C_{\beta_1}(\beta^2 + \epsilon) < 1 - \beta_1$, we have

$$E|\tilde{Y}(t) - Y(t)|^2 \leq C_\epsilon \int_t^T E|\tilde{Y}(s) - Y(s)|^2 \, \mathrm{d}s \tag{4.26}$$

and we can apply Gronwall's inequality to see that the left hand side equals zero. This further implies that

$$\int_t^T E|\tilde{Z}(s) - Z(s)|^2 \, \mathrm{d}s \leq \frac{C_{\beta_1}(\beta^2 + \epsilon)}{1 - \beta_1} \int_t^T E|\tilde{Z}(s) - Z(s)|^2 \, \mathrm{d}s. \tag{4.27}$$

Since the constant is less than one, we must have that the integral equals zero. We thus have

$$\tilde{Y}(t) = Y(t), \quad \tilde{Z}(t) = Z(t), \text{ a.s., a.e. } t \in [0, T]. \tag{4.28}$$

We have thus proven the claim that any solution of (4.1) has the required form.

Given the validity of the claim, we can easily prove the uniqueness of the solution given the current assumptions. Let (X, Y, Z) and $(\tilde{X}, \tilde{Y}, \tilde{Z})$ be two solutions to (4.1). By the the above claim, it is apparent that we must have the following relations amongst the variables.

$$\begin{array}{rclrcl} Y(t) &=& \theta(t, X), & Z(t) &=& z(t, X, \theta, \theta_x) \\ \tilde{Y}(t) &=& \theta(t, \tilde{X}), & \tilde{Z}(t) &=& z(t, \tilde{X}, \tilde{\theta}, \tilde{\theta}_x) \end{array} \tag{4.29}$$

Where $\tilde{\theta} = \theta(t, \tilde{X})$, etc. It is easily seen that X and $\tilde{X}$ satisfy the same forward SDE given in Step 3, with the same initial state x. Hence we must have that

$$X(t) = \tilde{X}(t), \qquad \forall t \in [0, T], \text{ a.s.} \tag{4.30}$$

which by (4.29) implies

$$Y(t) = \tilde{Y}(t), \qquad Z(t) = \tilde{Z}(t), \qquad \forall t \in [0, T], \text{ a.s.} \tag{4.31}$$

This concludes our proof. □

We end this section with a couple of observations regarding the four step scheme. Any reader familiar with stochastic analysis will be familiar with Feynman-Kac formula and how it can be used to solve a PDE through the solution of a corresponding SDE. What we are doing here is basically a reversal of that. We are trying to solve a FBSDE through solving an associated PDE. This inevitably makes the four step scheme reliant on the solvability of (4.8). The subject of the solvability of parabolic systems of PDEs is a very technical subject, too sophisticated to be dealt with in this brief paper. We shall content ourselves with a presentation some of the technical conditions required for solvability, and move on knowing that somewhere someone has written an excellent book on the subject for us to refer to (see

[5]). Before we get to step 2, we have to get past step 1; we need to be able to find a unique solution to (4.7). There is a very comfortable way to ensure that (4.7) is solvable through the restriction of σ to be independent of z, $\sigma = \sigma(t, x, y)$ in other words. This is comfortable because it also ensures that (4.12) is trivially true and we do not need to bother with this seemingly ad-hoc condition to ensure uniqueness of our solution. Lastly, the regularity and boundedness conditions imposed in theorem (4.1) are restrictive. However, this reduction in generality is awarded with an elegant result which presents a remarkable relationship between the variables $(X(t), Y(t), Z(t)) = (X(t), \theta(t, X), \theta_x(t, X)\sigma(t, X, Y, Z))$.

4.3.1 A general case

We shall now focus our attention on when the four step scheme can be realised, since this implies that the FBSDE described in (4.1) can be solved. As already remarked above, this relies on the solvability of the equations in step 1 and step 2 of the algorithm. In order to insure solvability, we need to make a few more assumptions.

(A1) We assume that the Brownian motion is of the same dimension as the variable X, in other words $d = n$. Further, the functions b, h, σ and g are smooth functions taking values in $\mathbf{R}^n$, $\mathbf{R}^m$, $\mathbf{R}^{n\times n}$ and $\mathbf{R}^m$, respectively, and with first order derivatives in x, y, z being bounded by some constant $L > 0$.

(A2) The function σ is independent of z and there exists a positive continuous function υ and a constant $\mu > 0$ s.t. $\forall (t, x, y, z) \in [0, T] \times \mathbf{R}^n \times \mathbf{R}^m \times \mathbf{R}^{n\times n}$

$$\upsilon(|y|)I \leq \sigma(t, x, y)\sigma(t, x, y)^T \leq \mu I \tag{4.32}$$

and

$$|b(t, x, 0, 0)| + |h(t, x, 0, z)| \leq \mu \tag{4.33}$$

(A3) There exists constants $C > 0$ and $\alpha \in (0, 1)$ s.t. g is bounded in $C^{2+\alpha}(\mathbf{R}^m)$

Given these assumptions, the original FBSDE (4.1) takes the form

$$\begin{cases} \mathrm{d}X(t) &= b(t, X, Y, Z)\,\mathrm{d}t + \sigma(t, X, Y)\,\mathrm{d}W(t) \\ \mathrm{d}Y(t) &= h(t, X, Y, Z)\,\mathrm{d}t + Z(t)\,\mathrm{d}W(t) \\ X(0) &= x,\ Y(T) \ = \ g(X(T)), \quad t \in [0, T] \end{cases} \tag{4.34}$$

Thus, the four step scheme can be realised if we can find a classical solution to (4.8), which under our current setting has the following appearance for $1 \leq k \leq m$

$$\begin{cases} \theta^k_t + \dfrac{1}{2}\mathrm{tr}[\theta^k_{xx}\sigma\sigma^T(t, x, \theta)] + \langle b(t, x, \theta, z(t, x, \theta, \theta_x)), \theta^k_x\rangle \\ \qquad -h^k(t, x, \theta, z(t, x, \theta, \theta_x)) = 0, \quad (t, x) \in [0, T] \times \mathbf{R}^n \\ \theta(T, x) = g(x), \quad x \in \mathbf{R}^n. \end{cases} \tag{4.35}$$

We shall take the time to briefly comment on the assumptions made in (A1)-(A3). Collectively, they are made to ensure that (4.35) and (4.7) both are uniquely solvable. As mentioned above, we assume σ to be independent of the variable z in order to ensure solvability of (4.7). The further restriction made upon σ is to insure its non-degeneracy, and is related to the solvability of (4.35). The rather specific boundedness assumptions on b and h are there to insure that (4.35) fulfill some particular technical conditions which relate to its solvability. Our final assumption (A3), that g is bounded and in $C^{2+\alpha}(\mathbf{R}^m)$, implies that g is twice differential with its derivatives Hölder continuous with exponential α. Hölder spaces are common in the functional analysis present in the theory of PDEs. This particular condition will ensure that the solution θ to our PDE will have sufficiently bounded derivatives in line with the requirements present in our main theorem (4.1).

We now return our focus to (4.35). To rigorously show that a classical solution exists is beyond this paper. We will content ourselves with an explanation of the main ideas of such a proof:

Firstly, we show that the boundary value problem described by

$$\begin{cases} \theta_t^k + \sum_{i,j=1}^{n} a_{ij}(t,x,\theta)\theta_{x_i x_j} + \sum_{i=1}^{n} b_i(t,x,\theta,z(t,x,\theta,\theta_x))\theta_x^k \\ \qquad - h^k(t,x,\theta,z(t,x,\theta,\theta_x)) = 0, \\ \qquad\qquad (t,x) \in [0,T] \times B_R,\ 1 \le k \le m \\ \theta|_{\partial B_R} = g(x),\ |x| = R \\ \theta(T,x) = g(x),\ x \in B_R \end{cases} \tag{4.36}$$

where B_R is the ball with radius $R > 0$ centered at the origin and

$$\begin{cases} (a_{ij}(t,x,y)) = \frac{1}{2}\sigma(t,x,y)\sigma(t,x,y)^T \\ (b_1(t,x,y,z),\dots,b_n(t,x,y,z))^T = b(t,x,y,z) \\ (h_1(t,x,y,z),\dots,h_m(t,x,y,z))^T = h(t,x,y,z) \end{cases} \tag{4.37}$$

admits a unique classical solution under numerous technical conditions which all can be derived from assumptions (A1)-(A3). Using this result, we can create a (bounded) sequence of solutions $\theta(t,x;R)$, with $\theta_t(t,x)$, $\theta_x(t,x;R)$ and $\theta_{xx}(t,x;R)$ uniformly bounded. As mentioned above, the boundedness of θ and its derivatives comes from the assumption $g \in C^{2+\alpha}(\mathbf{R}^m)$. From this sequence, we can find a subsequence which converges uniformly to a bounded $\theta(t,x)$ as $R \to \infty$, with bounded derivatives as required. Using standard approximation techniques and Gronwall's inequality, we show that the solution is unique and the proof is finished.

Despite its short appearance here, the proof is long and very technical. For further information and all the details, please consult [5] and [3]. The reader may rest assured that no understanding of the actual proof is necessary to follow the rest of this text, as long as we are willing to accept that the PDE is uniquely solvable.

4.3.2 The case when h has linear growth in z

As mentioned in the section title, the case we are interested in is when the coefficient function h has linear growth in the variable z (this occurs in mathematical control theory). We cannot directly apply our previous theorem, as condition (4.33) of assumption (A2) is not fulfilled. However, if we sacrifice some generality of our problem and further simplify the FBSDE in question, we will be able to relax condition (4.33). Our equation will now have the form

$$\begin{cases} \mathrm{d}X(t) &= b(t,X,Y,Z)\,\mathrm{d}t + \sigma(t,X)\,\mathrm{d}W(t) \\ \mathrm{d}Y(t) &= h(t,X,Y,Z)\,\mathrm{d}t + Z(t)\,\mathrm{d}W(t) \\ X(0) &= x,\ Y(T) \ = \ g(X(T)) \quad t\in[0,T]. \end{cases} \tag{4.38}$$

In other words, we demand that σ is independent of both variables y and z. In exchange for this further restriction, the four step scheme will once again be realisable and we will be able to find an adapted solution to (4.38). The proof of this is, again, omitted, but the interested reader may refer to [3].

We end this section with a brief remark on a third case, namely when $m = 1$.

4.3.3 The case when $m = 1$

When $m = 1$, $Y(t)$ is a one dimensional process. This implies that step 2 of the four step scheme no longer requires to solve a parabolic system, since the function $\theta(t,x)$ will be scalar valued. Instead we have a quasilinear parabolic equation, for which the theory is much more complete. This will allow for much more complicated non-linearities in our FBSDE, with the four step scheme still applicable.

Example. Consider the following FBSDE

$$\begin{cases} \mathrm{d}X(t) &= \dfrac{X(t)}{(Z(t)-Y(t))^2+1}\,\mathrm{d}t + X(t)\,\mathrm{d}W(t) \\ \mathrm{d}Y(t) &= \dfrac{Z(t)}{(Z(t)-Y(t))^2+1}\,\mathrm{d}t + Z(t)\,\mathrm{d}W(t) \\ X(0) &= x, \quad Y(T) \ = \ X(T). \end{cases} \tag{4.39}$$

We will solve it using the four step scheme:

Step 1 $\sigma(t,x,\theta) = x \Rightarrow z = \theta_x\sigma = \theta_x x$.

Step 2 With $y = \theta$, $z = \theta_x x$, $b = x/((\theta_x x-\theta)^2+1)$ and $h = \theta_x x/((\theta_x x-\theta)^2+1)$ we have the following PDE for our problem

$$\begin{cases} \theta_t - x^2\theta_{xx} = 0 \\ \theta(T,x) = x, \quad x\in\mathbf{R} \end{cases} \tag{4.40}$$

since

$$\begin{aligned}
&\theta_t + \frac{1}{2}\theta_{xx}\sigma^2(t,x,\theta) + b(t,x,\theta,\theta_x)\theta_x - h(t,x,\theta,\theta_x) = 0 \\
\Rightarrow\ &\theta_t + \frac{1}{2}\theta_{xx}x^2 + \frac{\theta_x x}{(\theta_x x - \theta)^2 + 1} - \frac{\theta_x x}{(\theta_x x - \theta)^2 + 1} = \theta_t + x^2\theta_{xx} = 0.
\end{aligned} \tag{4.41}$$

This has the unique solution $\theta(t,x) = x$. We thus have $Y(t) = X(t)$ and $Z(t) = \theta_x(t,X)X(t) = X(t)$

Step 3 We now have to solve

$$\begin{cases} \mathrm{d}X(t) &= X(t)\,\mathrm{d}t + X(t)\,\mathrm{d}W(t) \\ X(0) &= x \end{cases} \tag{4.42}$$

which is the well recognised dynamics for the geometric Brownian motion $xe^{t/2+W(t)}$.

Step 4 Combining the above results, we have that $X(t) = Y(t) = Z(t) = xe^{t/2+W(t)}$. We have thus found a unique adapted solution to (4.39).

4.3.4 The infinite horizon case

The example presented in this section is of special interest to us since it will play major role in one of our later examples of the applications of the theory of FBSDEs. We shall concern ourselves with the following FBSDE

$$\begin{cases} \mathrm{d}X(t) &= b(X,Y)\,\mathrm{d}t + \sigma(X,Y)\,\mathrm{d}W(t) \\ \mathrm{d}Y(t) &= [h(X)Y(t) - 1]\,\mathrm{d}t + \langle Z(t),\, \mathrm{d}W(t)\rangle \\ X(0) &= x,\ Y(t) \text{ is bounded a.s., uniformly } \forall t \in [0,\infty). \end{cases} \tag{4.43}$$

As both the title and equation (4.43) implies, the time duration here is infinite. We will further allow Y to be only one dimensional. However, X still takes values in $\mathbf{R}^n$, and W is still n-dimensional Brownian motion. We also assume that there is no explicit t-dependence in our coefficient functions.

The nodal solution

Definition 4.2. *A process $\{(X,Y,Z)\}_{t\geq 0}$ is called an adapted solution of (4.43) if for any $T > 0$, $(X,Y,Z)|_{[0,T]} \in \mathcal{M}[0,T]$ and*

$$\begin{cases} X(t) &= x + \int_0^t b(X(s),Y(s))\,\mathrm{d}s + \int_0^t \sigma(X(s),Y(s))\,\mathrm{d}W(s) \\ Y(t) &= Y(T) - \int_t^T [h(X(s))Y(s) - 1]\,\mathrm{d}s + \int_t^T \langle Z(s),\, \mathrm{d}W(s)\rangle \\ & \qquad 0 \leq t \leq T < \infty \end{cases} \tag{4.44}$$

such that $\exists\, M > 0$, $|Y(t)| \le M$, $\forall t$, *a.s. Moreover, if an adapted solution* (X, Y, Z) *is such that for some* $\theta \in C^2(\mathbf{R}^n) \cap C_b^1(\mathbf{R}^n)$, *the following relation holds*

$$\begin{aligned} Y(t) &= \theta(X) \\ Z(t) &= \sigma(X, \theta(X))^T \theta_x(X) \end{aligned} \tag{4.45}$$

then we call (X, Y, Z) *a nodal solution of* (4.43) *with the representing function* θ.

The reason it is called a nodal solution is because the function θ describes the nodal set of some function V, which is the solution of the HJB-equation associated with our FBSDE (cf. chapter 2).

In order to proceed, we need to make some assumptions regarding the coefficient functions present.

(H1) The functions σ, b and h are C^1 with bounded partial derivatives, and $\exists\, \lambda$, $\mu > 0$, and some continuous increasing function $\upsilon[0, \infty) \to [0, \infty)$ such that

$$\lambda I \le \sigma(x, y)\sigma(x, y)^T \le \mu I, \quad (x, y) \in \mathbf{R}^n \times \mathbf{R} \tag{4.46}$$

$$|b(x, y)| \le \upsilon(|y|), \quad (x, y) \in \mathbf{R}^n \times \mathbf{R} \tag{4.47}$$

$$\inf_{x \in \mathbf{R}^n} h(x) \equiv \delta > 0, \qquad \sup_{x \in \mathbf{R}^n} h(x) \equiv \gamma < \infty \tag{4.48}$$

This assumption allows the following important result:

Lemma 4.3. *Let (H1) hold. Then the following equation has a classical solution* $\theta \in C^{2+\alpha}(\mathbf{R}^n)$*:*

$$\frac{1}{2}\mathrm{tr}[\theta_{xx}\sigma(x, \theta)\sigma^T(x, \theta)] + \langle b(x, \theta), \theta_x \rangle - (h(x)\theta - 1) = 0 \tag{4.49}$$

such that

$$\frac{1}{\gamma} \le \theta(x) \le \frac{1}{\delta}, \qquad x \in \mathbf{R}^n \tag{4.50}$$

We will leave it to the interested reader to investigate the proof in [3] and accept the result as true. Again, we see technical assumptions introduced to ensure solvability of a PDE associated to our initial stochastic problem. It is becoming evident that the biggest obstacle in the four step scheme is its reliance upon the realisation of step 2 of the algorithm.

We are now ready to show the following result concerning the existence of nodal solutions of (4.43).

Theorem 4.4. *Let (H1) hold. Then there exists at least one nodal solution* (X, Y, Z) *of* (4.43)*, with representing function* θ *being the solution of* (4.49)*. Conversely, if* (X, Y, Z) *is a nodal solution of* (4.43) *with representing function* θ*, then* θ *is a solution of* (4.49).

Proof. By our previous lemma, we can find a classical solution $\theta \in C^{2+\alpha}(\mathbf{R}^n)$ of (4.49). Now consider the forward SDE

$$\begin{cases} \mathrm{d}X(t) &= b(X, \theta(X))\,\mathrm{d}t + \sigma(X, \theta(X))\,\mathrm{d}W, \qquad t > 0 \\ X(0) &= x \end{cases} \tag{4.51}$$

Since θ_x is bounded and b and σ are uniformly Lipschitz by assumption, (4.51) admits a unique strong solution X, $t \in [0, \infty)$. Now define

$$\begin{cases} Y(t) &= \theta(X) \\ Z(t) &= \sigma(X, \theta(X))^T \theta_x(X) \end{cases} \tag{4.52}$$

If we apply the Itô formula to $Y = \theta(X)$, and do our usual comparing of coefficients trick, we easily see that Y indeed has the required dynamics, and (X, Y, Z) is an adapted solution to (4.43). By definition, it is also a nodal solution.

Conversely, let (X, Y, Z) be a nodal solution of (4.43) with representing function θ. Since $\theta \in C^2$, we can apply Itô to $Y(t) = \theta(X)$ with the result

$$\begin{aligned} \mathrm{d}Y(t) &= \left\{ \langle b(X, \theta(X)), \theta_x(X) \rangle + \frac{1}{2}\mathrm{tr}[\theta_{xx}(X)\sigma\sigma^T(X, \theta(X))] \right\} \mathrm{d}t \\ &\quad + \langle \theta(X), \sigma(X, \theta(X))\,\mathrm{d}W(t) \rangle \end{aligned} \tag{4.53}$$

By application of our standard technique of comparison of coefficients we obtain

$$\langle b(X, \theta(X)), \theta_x(X) \rangle + \frac{1}{2}\mathrm{tr}[\theta_{xx}(X)\sigma\sigma^T(X, \theta(X))] = h(X)\theta(X) - 1, \ \forall t \geq 0, \text{ a.s.} \tag{4.54}$$

We now define a new continuous function $F : \mathbf{R}^n \to \mathbf{R}$ by

$$F(x) \triangleq \langle b(x, \theta(x)), \theta_x(x) \rangle + \frac{1}{2}\mathrm{tr}[\theta_{xx}(x)\sigma\sigma^T(x, \theta(x))] - h(x)\theta(x) + 1 \tag{4.55}$$

We are now faced with the task of showing that $F \equiv 0$. We content ourselves with a brief outline on how to do it.

The process X is a time homogeneous Markov process since it satisfies the forward SDE

$$\begin{cases} \mathrm{d}X(t) &= \tilde{b}(X)\,\mathrm{d}t + \tilde{\sigma}(X)\,\mathrm{d}W(t) \\ X(0) &= x \end{cases} \tag{4.56}$$

where $\tilde{b}(x) \triangleq b(x, \theta(x))$ and $\tilde{\sigma}(x) \triangleq \sigma(x, \theta(x))$. Hence the process X will posses a transition probability density $p(t, x, y)$, which will be positive everywhere. Together with the fact that by (4.49) $F(X) = 0$, $\forall t \geq 0$, a.s., we have

$$0 = E_{0,x}[F(X(t))^2] = \int_{\mathbf{R}^n} p(t, x, y) F(y)^2\,\mathrm{d}y, \quad \forall t > 0 \tag{4.57}$$

Since $p(t, x, y) > 0$, we must have $F(y) = 0$ a.e. under the Lebesgue measure in $\mathbf{R}^n$. Since F is continuous by construction, it must be that $F(y) = 0$ everywhere in $\mathbf{R}^n$. We have thus shown that $F \equiv 0$ and this concludes the proof. □

A brief reflection on the result of this section is in order. From what we have seen, the number of nodal solutions to (4.43) is exactly the same as the number of classical solutions to (4.49). Hence, if the solution to the parabolic equation is unique, so will the nodal solution of our main equation be. However, there might be other adapted, non - nodal solutions to our FBSDE. Under some rather technical assumptions on the coefficient functions, it is possible to show that there exists a unique adapted solution to (4.43), and that such a solution will be nodal. The proof is of a highly technical nature, and is per standard procedure omitted. There is a very straight forward way to ensure the existence of a unique adapted solutions, albeit a bit limiting. If we condition b and σ to depend on X only, there will be a unique nodal solution to (4.43).

4.3.5 The limit of equations in finite time duration

In this section we present a result on the convergence of finite time duration solution, in the sense below. By $(X^K, Y^K, Z^K) \in \mathcal{M}[0, K]$, we denote the adapted solution to the equation

$$\begin{cases} \mathrm{d}X(t) &= b(X,Y)\,\mathrm{d}t + \sigma(X,Y)\,\mathrm{d}W(t) \\ \mathrm{d}Y(t) &= [h(X)Y(t) - 1]\,\mathrm{d}t + \langle Z(t),\, \mathrm{d}W(t)\rangle \\ X(0) &= x,\ Y(t) \text{ is bounded a.s., uniformly } \forall t \in [0, K]. \end{cases} \tag{4.58}$$

with $Y^K(K) = g(X(K))$ for some smooth function g. We are interested in what happens if we let $K \to \infty$.

If we let (X, Y, Z) be a nodal solution of (4.1), it is possible to show that

$$\begin{aligned} E|X^K(t) - X(t)|^2 &\leq Ce^{-2\eta(K-t)} \\ E|Y^K(t) - Y(t)|^2 &\leq Ce^{-2\eta(K-t)} \\ E|Z^K(t) - Z(t)|^2 &\leq Ce^{-2\eta(K-t)} \end{aligned} \tag{4.59}$$

where C and η are some positive constants. From this it is easy to conclude that

$$\lim_{K\to\infty} E\left\{|X^K(t) - X(t)|^2 + |Y^K(t) - Y(t)|^2 + |Z^K(t) - Z(t)|^2\right\} = 0. \tag{4.60}$$

In other words, the limit of our finite duration problem corresponds with our infinite horizon problem.

4.4 FBSDEs with stochastic coefficients

We will end this chapter chapter with a comment on FBSDEs with stochastic coefficients. If we let the functions b, σ, h and g depend explicitly on $\omega \in \Omega$, as well as on (t, X, Y, Z), and try to apply the four step scheme we are faced with a different type of problem. Our parabolic PDE will become a (possibly degenerate non-linear) backward stochastic partial differential equation (BSPDE). The subject of BSPDEs is a sophisticated and technical matter, and will be left to the reader to explore. An introduction limited to a type of linear BSPDE can be found in chapter 5 of [3].

Chapter 5

Black's consol rate conjecture

5.1 Introduction

In this section we shall present and example of the application of our theory. It is based upon a conjecture by one of the pioneers in the field of financial mathematics, Fisher Black. In the late seventies Brennan and Schwartz [6] presented a paper where they introduced a two factor term structure model. Their choice of variables was the short rate r and the consol rate l, both taken as exogenously given. The model was very well received in the community, and used extensively in the pricing of assets depending on the mutual behaviour of the short and long (consol) rate, such as short rate/long rate swaps. In the early nineties, Hogan [7] showed that under the proposed solution to the Brennan - Schwartz model, the consol rate could explode in finite time under certain circumstances. Around this time, the major advancements in the field of FBSDEs had started to make themselves known, and perhaps Black saw the possibility of application when he presented his consol rate conjecture in a private communication to the authors of [8]. We begin with some definitions

Definition 5.1. *A consol is security that pays dividends continually and in perpetuity, also known as a perpetual annuity.*

In other words, a consol is a security that pays a continuous stream of dividends, forever. In the case of a constant short rate $r > 0$, we can easily calculate the present price of an annuity that pays 1 unit of currency at the end of each period. If we denote the price by Y, we have the following formula

$$Y = \sum_{k=1}^{\infty} \frac{1}{(1+r)^k} = \frac{1}{1+r} \times \frac{1}{1-\frac{1}{1+r}} = \frac{1}{r} \tag{5.1}$$

since the short rate is equal the long rate in this case, and the price should be the sum of the discounted values of future dividends. We thus have that the consol (long) rate is the reciprocal of the consol price in the case of a constant short rate. Inspired by this, we define

Definition 5.2. *The consol rate $l(t)$ is the reciprocal of the consol price $Y(t)$, $l(t) = 1/Y(t)$.*

Since we are not only interested in the case where the short rate is constant, we need to generalise the above idea to be applicable to any given short rate process $r = \{r(t) : t \geq 0\}$. In order to do so we introduce the expected discounted value formula to calculate the consol price process $Y = \{Y(t) : t \geq 0\}$:

$$Y(t) = E\left\{\int_t^\infty e^{-\int_t^s r(u)\,\mathrm{d}u}\,\mathrm{d}s\Big|\mathcal{F}_t\right\}, \quad t \geq 0. \tag{5.2}$$

This is consistent with our above formula for Y in the case with constant short rate. Note that we are working under the equivalent (risk neutral) martingale measure $\mathbf{P}$ from the onset, and hence the expectation is under said measure. Inspired by Brennan-Schwartz we consider a short rate process $r(t)$ dependent on $Y(t)$, with the following dynamics:

$$\mathrm{d}r(t) = \mu(r, Y)\,\mathrm{d}t + \alpha(r, Y)\,\mathrm{d}W(t) \tag{5.3}$$

where $W(t)$ is standard Brownian motion in $\mathbf{R}^2$, and the coefficients μ and α are measurable functions of appropriate dimensions. The question we now ask is whether there exists a pair of processes $(r(t), Y(t))$ such that the above equations are satisfied? If so, what are the dynamics of the consol rate process $Y(t)$?

If we assume such a pair exists, we can get an idea of the dynamics of the consol rate by applying the Itô formula to (5.2). Hence we have the following qualified guess

$$\mathrm{d}Y(t) = (r(t)Y(t) - 1)\,\mathrm{d}t + A(r, Y)\,\mathrm{d}W(t) \tag{5.4}$$

where $A(r, Y)$ is some function $(0, \infty) \times (0, \infty) \to \mathbf{R}^2$. The drift term is simply the t-derivate of $Y(t)$, and can be interpreted as that the expected return (under the equivalent martingale measure) should be equal to the short rate, minus the dividend.

5.2 Black's consol rate conjecture

Given the above preliminaries, we are now ready to present **Black's conjecture:**
Given a short rate r(t) with the dynamics presented above, we have that under at most technical conditions, for any pair $(\mu(r, Y), \alpha(r, Y))$, there is always a function $A : (0, \infty) \times (0, \infty) \to \mathbf{R}^2$ depending on μ and α, such that

$$\mathrm{d}Y(t) = (r(t)Y(t) - 1)\,\mathrm{d}t + A(r, Y)\,\mathrm{d}W(t). \tag{5.5}$$

The reader familiar with the work of Brennan and Schwartz will realise that this is basically a confirmation of their model. However, the reason the model could be shown to fail was because of a lack of insight into how to find A, and its relation to the pair (μ, α).

To confirm Black's conjecture we need to make a certain assumption on the short rate $r(t)$, that it is 'hidden Markovian'. This means that there exists an n-dimensional Markovian state process $X(t)$ s.t. $r(t) = h(X(t))$ a.s. $\forall\ t$, where h is some well behaved function. We assume the following dynamics for X

$$\begin{cases} \mathrm{d}X(t) &= b(X,Y)\,\mathrm{d}t + \sigma(X,Y)\,\mathrm{d}W(t) \\ X(0) &= x \quad t \in [0,T] \end{cases} \tag{5.6}$$

where b and σ are some appropriate functions on $\mathbf{R}^n \times \mathbf{R}$.

Since b and σ can be computed in terms of μ and α with the help of the Itô formula, we can recast our problem as follows:

The infinite horizon consol rate problem (IHCR for short): Find a pair of adapted, locally square integrable processes $(X(t), Y(t))$ s.t.

$$\begin{cases} \mathrm{d}X(t) &= b(X,Y)\,\mathrm{d}t + \sigma(X,Y)\,\mathrm{d}W(t) \\ \mathrm{d}Y(t) &= E\left\{\int_t^\infty e^{\int_t^s h(X(u))\,\mathrm{d}u}\,\mathrm{d}s \,\middle|\, \mathcal{F}_t\right\} \\ X(0) &= x \quad t \in [0,\infty) \end{cases} \tag{5.7}$$

As the name implies, the key to solving the IHCR problem is to exploit its relation to the infinite horizon problem studied in connection to the four step scheme. Following this path, we introduce the following definition:

Definition 5.3. *An adapted solution to the IHCR problem is called a nodal solution with representing function θ if there exists a bounded C^2 function θ, with bounded θ_x, s.t. $Y(t) = \theta(X(t))$.*

Please recall the system studied when we discussed the infinite horizon problem in connection to the four step scheme:

$$\begin{cases} \mathrm{d}X(t) &= b(X(t),Y(t))\ \mathrm{d}t + \sigma(X(t),Y(t))\ \mathrm{d}W(t) \\ \mathrm{d}Y(t) &= (h(X(t))Y(t) - 1)\,\mathrm{d}t - \langle Z(t),\, \mathrm{d}W(t)\rangle \\ X(0) &= x, \\ Y(t) &\text{is bounded a.s., uniformly in } t \in [0,T] \end{cases} \tag{5.8}$$

Please also recall the technical assumption made upon the coefficients σ and h.

(A2) The functions σ, b and h are C^1 with bounded partial derivatives. Further there exists constants λ, $\mu > 0$, and some increasing function $\upsilon[0,\infty) \to [0,\infty]$ s.t.

$$\begin{array}{ll} \lambda I \leq \sigma(x,y)\sigma(x,y)^T \leq \mu I & (x,y) \in \mathbf{R}^n \times \mathbf{R} \\ |b(x,y)| \leq \upsilon(|y|) & (x,y) \in \mathbf{R}^n \times \mathbf{R} \\ \inf_{x \in \mathbf{R}^n} h(x) = \delta > 0, & \sup_{x \in \mathbf{R}^n} h(x) = \gamma < \infty \end{array} \tag{5.9}$$

Theorem 5.4. *Assume (A2). If (X,Y,Z) is an adapted solution to (5.8), then (X,Y) is an adapted solution to our IHCR problem. Conversely, if (X,Y) is an adapted solution to the IHCR problem, there exists an adapted, $\mathbf{R}^2$-valued, locally square integrable process Z, s.t. (X,Y,Z) is an adapted solution of (5.8)*

Proof. Let (X,Y,Z) be an adapted solution to (5.8). By definition, X is as in the IHCR problem. We are left to show that the process Y has the required dynamics. In order to do so, let $\Gamma(t) = e^{-\int_0^t h(X(u))\,\mathrm{d}u}$, $t \in [0,T]$. Consider the function $f(Y,\Gamma) = \Gamma Y$; the Itô formula implies

$$\mathrm{d}f(Y,\Gamma) = \Gamma(t)\,\mathrm{d}Y(t) + Y(t)\,\mathrm{d}\Gamma(t) + \mathrm{d}Y(t)\,\mathrm{d}\Gamma(t) = \Gamma(t)\,\mathrm{d}Y(t) + Y(t)\,\mathrm{d}\Gamma(t) \tag{5.10}$$

since

$$\mathrm{d}\Gamma(t) = -h(X(t))e^{-\int_0^t h(X(u))\,\mathrm{d}u}\,\mathrm{d}t = -h(X(t))\Gamma(t)\,\mathrm{d}t. \tag{5.11}$$

We thus have

$$\begin{aligned} \mathrm{d}f(Y,\Gamma) &= \Gamma(t)\left\{(h(X)Y(t) - 1)\,\mathrm{d}t - \langle Z(t),\,\mathrm{d}W(t)\rangle\right\} - Y(t)h(X)\Gamma(t)\,\mathrm{d}t \\ &= -\Gamma(t)\,\mathrm{d}t - \langle \Gamma(t)Z(t),\,\mathrm{d}W(t)\rangle \end{aligned} \tag{5.12}$$

or

$$\begin{aligned} Y(T)\Gamma(T) &= Y(t)\Gamma(t) - \int_t^T \Gamma(s)\,\mathrm{d}s - \int_t^T \langle \Gamma(s)Z(s),\,\mathrm{d}W(s)\rangle \\ \Rightarrow \quad Y(t)\Gamma(t) &= Y(t)\Gamma(t) + \int_t^T \Gamma(s)\,\mathrm{d}s + \int_t^T \langle \Gamma(s)Z(s),\,\mathrm{d}W(s)\rangle \\ \Rightarrow \quad Y(t) &= \frac{Y(T)\Gamma(T)}{\Gamma(t)} + \int_t^T \frac{\Gamma(s)}{\Gamma(t)}\,\mathrm{d}s + \int_t^T \langle \frac{\Gamma(s)}{\Gamma(t)}Z(s),\,\mathrm{d}W(s)\rangle \\ \Rightarrow \quad Y(t) &= Y(T)e^{-\int_t^T h(X(u))\,\mathrm{d}u} + \int_t^T e^{-\int_t^s h(X(u))\,\mathrm{d}u}\,\mathrm{d}s \\ &\quad + \int_t^T \langle e^{-\int_t^s h(X(u))\,\mathrm{d}u} Z(s),\,\mathrm{d}W(s)\rangle. \end{aligned} \tag{5.13}$$

We now follow standard procedure and take expectation of both sides conditioned on $\mathcal{F}_t$ and let $T \to \infty$. Our last term above disappears under the expectation since $\Gamma(t)Z(t)$ is adapted by construction. Our first term is a.s. bounded since

$$Y(T)e^{-\int_t^T h(X(u))\,\mathrm{d}u} < Ce^{-(T-t)\delta}, \text{ a.s.} \tag{5.14}$$

by definition of Y and assumptions made on h in (A2). As the right hand side obviously tends to zero as T goes to infinity, we are left with

$$Y(t) = E\left\{\int_t^\infty e^{-\int_t^s r(u)\,\mathrm{d}u}\,\mathrm{d}s\middle|\,\mathcal{F}_t\right\},\quad t \geq 0. \tag{5.15}$$

For the converse, assume (X, Y) is an adapted solution to our IHCR problem. Define

$$U(t) = \int_t^\infty e^{\int_t^s h(X(u))\,\mathrm{d}u}\,\mathrm{d}s. \tag{5.16}$$

The technical conditions implied on the function h through (A2) assures that $U(t)$ is well defined. We further claim that $U(t)$ is the unique bounded solution to the ODE

$$\frac{\mathrm{d}U(t)}{\mathrm{d}t} = h(X)U(t) - 1. \tag{5.17}$$

It is easily verified by direct differentiation of $U(t)$ that it is indeed a solution to (5.17). From the definition of U we can also see that it is bounded. For uniqueness, let U be any bounded solution to (5.17) defined on $[0, \infty)$. By the method of integrating factor we have that for any $0 \leq t \leq T$

$$U(t) = e^{-\int_t^T h(X(u))\,\mathrm{d}u}U(T) + \int_t^T e^{-\int_t^s h(X(u))\,\mathrm{d}u}\,\mathrm{d}s. \tag{5.18}$$

Taking the limit as $T \to \infty$, we realise that the first term disappears as $U(T)$ is bounded for all $T > 0$, and $h(X)$ is bounded from below a.s. by assumption (A2). We have hence shown that $U(t)$ is the unique bounded solution to (5.17). We now define $Y(t) = E[U(t)|\,\mathcal{F}_t]$. Since the filtration $\mathcal{F}_t$ is generated by Brownian motion, which is continuous, $Y(t)$ is also continuous. This implies that for any adapted process $H(t)$

$$\begin{aligned}
E\left\{\int_t^T H(s)Y(s)\,\mathrm{d}s\middle|\mathcal{F}_t\right\} &= \int_t^T E[H(s)Y(s)|\mathcal{F}_t]\,\mathrm{d}s \\
&= \int_t^T H(s)E[Y(s)|\mathcal{F}_t]\,\mathrm{d}s = \int_t^T H(s)E[E[U(s)|\mathcal{F}_s]|\mathcal{F}_t]\,\mathrm{d}s \\
&= \int_t^T H(s)E[U(s)|\mathcal{F}_t]\,\mathrm{d}s = \int_t^T E[H(s)U(s)|\mathcal{F}_t]\,\mathrm{d}s \\
&= E\left\{\int_t^T H(s)U(s)\,\mathrm{d}s\middle|\mathcal{F}_t\right\}
\end{aligned} \tag{5.19}$$

where we have utilised that $H(t)$ is adapted and that $\mathcal{F}_t \subseteq \mathcal{F}_s$ for all $t \leq s$.

By integrating (5.17) we have

$$U(t) = U(T) + \int_T^t [h(X(s))U(s) - 1]\,\mathrm{d}s = U(T) - \int_t^T [h(X(s))U(s) - 1]\,\mathrm{d}s \tag{5.20}$$

and by combining this with (5.19), we have for $0 \le t \le T$

$$\begin{aligned} Y(t) &= E[U(t)|\mathcal{F}_t] = E\left\{U(T) - \int_t^T [h(X(s))U(s) - 1]\,\mathrm{d}s \Big| \mathcal{F}_t\right\} \\ &= E\left\{Y(T) - \int_t^T [h(X(s))Y(s) - 1]\,\mathrm{d}s \Big| \mathcal{F}_t\right\}. \end{aligned} \tag{5.21}$$

$Y(t)$ is by construction a martingale, and we can apply the martingale representation theorem to get the following expression

$$Y(t) = Y(T) - \int_t^T [h(X(s))Y(s) - 1]\,\mathrm{d}s + \int_t^T \langle Z^{(T)}(s), dW(s)\rangle. \tag{5.22}$$

Here $Z^{(T)}$ is a well behaved adapted process defined on [0,T]. This can be done for all $0 < T < \infty$, which gives us the following relationship between two different representations

$$\begin{aligned} Y(t) &= Y(T_1) - \int_t^{T_1} [h(X(s))Y(s) - 1]\,\mathrm{d}s + \int_t^{T_1} \langle Z^{(T_1)}(s), dW(s)\rangle \\ &= Y(T_2) - \int_t^{T_2} [h(X(s))Y(s) - 1]\,\mathrm{d}s + \int_t^{T_2} \langle Z^{(T_2)}(s), dW(s)\rangle \end{aligned} \tag{5.23}$$

where $0 < T_1 < T_2 < \infty$. This implies

$$\begin{aligned} Y(T_1) &= Y(T_2) - \int_{T_1}^{T_2} [h(X(s))Y(s) - 1]\,\mathrm{d}s + \int_{T_1}^{T_2} \langle Z^{(T_2)}(s), dW(s)\rangle \\ &= Y(T_2) - \int_{T_1}^{T_2} [h(X(s))Y(s) - 1]\,\mathrm{d}s + \int_{T_1}^{T_2} \langle Z^{(T_2)}(s), dW(s)\rangle \\ &+ \int_t^{T_1} \langle Z^{(T_2)}(s) - Z^{(T_1)}(s), dW(s)\rangle \end{aligned} \tag{5.24}$$

From this we conclude that

$$\int_t^{T_1} \langle Z^{(T_2)}(s) - Z^{(T_1)}(s), dW(s)\rangle = 0, \qquad \forall t \in [0, T_1] \tag{5.25}$$

which by the Itô isometry implies

$$E\left\{\int_0^{T_1} \left|Z^{(T_2)} - Z^{(T_1)}\right|^2 \mathrm{d}s\right\} = 0. \tag{5.26}$$

This is equivalent to $Z^{(T_1)} = Z^{(T_2)}$, $\mathrm{d}t \otimes \mathrm{d}\mathbf{P}$ - a.s. on $[0, T_1] \times \Omega$. This allows us to create a process $Z(t) = Z^{(N)}(t)$ for $t \in [0, N]$, where $N = 1, 2\ldots$ We can thus rewrite (5.22) in our desired form

$$Y(t) = Y(T) - \int_t^T [h(X(s))Y(s) - 1]\,\mathrm{d}s + \int_t^T \langle Z(s), dW(s)\rangle \tag{5.27}$$

and we conclude that (X, Y) indeed satisfies (5.8). Finally, we recall that $Y(t) = E[U(t)|\mathcal{F}_t]$, and that $U(t) \le \frac{1}{\delta}$, $\forall t \ge 0$, a.s., by definition. From these facts, it is apparent that the process Y is indeed bounded, and we have proven our theorem. □

By the above result we have almost confirmed Black's conjecture. However, to fully confirm it, we still need to show that the process $Z(t)$ in (5.8) can be written as a function $B(h(X(t)), Y(t)) = B(h^{-1}(r(t)), Y(t)) = A(r(t), Y(t))$ in our specific case. To do so, we utilise theorem 4.4 from our earlier chapter. The theorem can be modified to:

Theorem 5.5. *Assume (A2). Then there exists at least one nodal solution (X, Y) of problem IHCR. Moreover, the representing function θ satisfies*

(i) $\gamma^{-1} \leq \theta(x) \leq \delta^{-1}$, *for all x $x \in \mathbf{R}$*

(ii) θ *satisfies the following differential equation for $x \in \mathbf{R}^n$*

$$\frac{1}{2}tr\left(\theta_{xx}\sigma(x,\theta)\sigma^T(x,\theta)\right) + \langle b(t,\theta), \theta_x\rangle - h(x)\theta + 1 = 0 \tag{5.28}$$

Proof. This is a direct consequence of theorem 4.4 □

Finally, whenever a nodal solution exists, $Z(t) = \sigma^T(X,Y)\theta_x(X)$, and we have confirmed Black's conjecture with $A(r,y) = \sigma^T(h^{-1}(r), y)\theta_x(h^{-1}(r))$.

5.2.1 Relation between the the limit of the long term rate and the consol rate

Normally we are interested in modeling not the consol rate, but the long term interest rate. Below, we shall see how this is done within our framework. A natural question arises; is the limit of the long rate equal to the consol rate? As indicated in the more general case in the previous chapter on the four step scheme, this is indeed the case. In our current setting, we shall consider

The Finite Horizon Valuation Problem (FHV):

$$\begin{cases} X(t) &= x + \int_0^t b(X(s),Y(s))\,\mathrm{d}s + \int_0^t \sigma(X(s),Y(s))\,\mathrm{d}W(s) \\ Y(t) &= E\left\{\Gamma_t^T g(X(T)) + \int_t^t \Gamma_t^s\,\mathrm{d}s \middle| \mathcal{F}_t\right\} \end{cases} \tag{5.29}$$

As before, $X(t)$ is the underlying state process determining the short rate $r(t)$, and the function h satisfies (A2). We further assume at time T we have the explicit relationship $Y(T) = g(X(T))$. Lastly, $\Gamma_t^s = e^{\int_t^s h(X(u))\,\mathrm{d}u}$. Here we choose to regard Y as the price of a long term bond, instead of a consol. Similarly to our IHCR problem, we call an adapted solution of (5.29) a nodal solution of problem FHV is there exists a function $\theta : [0,T] \times \mathbf{R}^n \to \mathbf{R}$ which is C^1 in t and C^2 in x, such that

$$Y(t) = \theta(t, X(t)), \quad t \in [0,T]. \tag{5.30}$$

Following the same approach as in the IHCR case, we can show that if (X, Y) is an adapted solution to (5.29), there exists a well behaved process Z such that the triple (X, Y, Z) is an adapted solution to

$$\begin{cases} \mathrm{d}X(t) &= b(X(t), Y(T))\,\mathrm{d}t + \sigma(X(t), Y(t))\,\mathrm{d}W(t) \\ \mathrm{d}Y(t) &= (H(X(t))Y(t) - 1)\,\mathrm{d}t - \langle Z(t),\, \mathrm{d}W(t)\rangle \\ X(0) &= x, \quad Y(T) \;=\; g(X(T)), \quad t \in [0, T] \end{cases} \tag{5.31}$$

It should come as no surprise that, just like in our previous case, if (X, Y, Z) is an adapted solution to (5.31), then Y satisfies (5.29). If we further impose regularity conditions on the function g, any adapted solution to (5.31) must be a nodal solution. Under these conditions, the limit of the FHV problem coincides with the IHCR problem. In other words, the long term interest rate converges to the consol rate (in expectation).

Chapter 6

Pricing a contingent claim with an illiquid underlying asset market

6.1 Introduction

This example is due to [9]. In the standard Black-Scholes market model one of the main assumptions is that the market is perfectly liquid and all market participants are price takers, i.e. they cannot influence the prices of the traded assets. In reality, this is usually not the case. Generally, whenever a market participant is trading outside of the quoted depth of a particular asset, there is a possibility that the price may change for the worse for the trading party initiating the trade. In this example we will introduce a market model which allows for *price impact* based on the trader's actions. With the help of the four step scheme, we will find a modified Black-Scholes PDE for the fair price of any replicable contingent claim. As always, we are working in our standard setting where $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{\geq 0}, \mathbf{P})$ is a complete filtered probability space, and $W(t)$ is a one dimensional Brownian motion. We shall consider a market consisting of a risk free asset with the price $B(t)$ and a risky asset, which we will call a stock, with the price $S(t)$. We assume the that the stock price is ex-dividend and that the dividend yield of the stock is equivalently equal to zero. The risk free asset price has the following dynamics

$$\mathrm{d}B(t) = r(t,S)B(t)\,\mathrm{d}t, \quad t \geq 0 \tag{6.1}$$

where we let the interest rate r depend on the stock price $S(t)$ as well as the time t. We also introduce the process $N(t)$ which describes the number of stocks in the portfolio at time t. Given this setting, the stock price follows

$$\mathrm{d}S(t) = \mu(t,S)S(t)\,\mathrm{d}t + \sigma(t,S)S(t)\,\mathrm{d}W(t) + \lambda(t,S)\,\mathrm{d}N(t), \quad t \geq 0 \tag{6.2}$$

where $\lambda(t,S) \geq 0$ is the price impact function, and the term $\lambda(t,S)\,\mathrm{d}N(t)$ is the price impact on the stock price. This implies that whenever the stock is sold, the price will go down, and conversely, when the stock is bought, the price will go up. We will have the following dynamics for the wealth process $Y(t)$

$$\begin{aligned} \mathrm{d}Y(t) &= r(t,S)Y(t)\,\mathrm{d}t + [\mu(t,S) - r(t,S)]S(t)N(t)\,\mathrm{d}t \\ &\quad + S(t)N(t)\sigma(t,S)\,\mathrm{d}W(t) + S(t)N(t)\lambda(t,S)\,\mathrm{d}N(t) \end{aligned} \tag{6.3}$$

where we have assumed that there is no consumption. If we assume that

$$\begin{cases} \mathrm{d}N(t) &= \eta(t)\,\mathrm{d}t + \zeta(t)\,\mathrm{d}W(t) \\ N(0) &= N_0 \end{cases} \tag{6.4}$$

where η and ζ are some processes to be determined, the dynamics for the stock price and the wealth become

$$\begin{cases} \mathrm{d}S(t) &= [\mu(t,S) + \lambda(t,S)\eta(t)]S(t)\,\mathrm{d}t + [\sigma(t,S) + \lambda(t,S)\zeta(t)]S(t)\,\mathrm{d}W(t) \\ S(0) &= S_0, \quad t \geq 0 \end{cases} \tag{6.5}$$

and

$$\begin{aligned} \mathrm{d}Y(t) &= r(t,S)Y(t)\,\mathrm{d}t + [\mu(t,S) - r(t,S) + \lambda(t,S)\eta(t)]S(t)N(t)\,\mathrm{d}t \\ &\quad + [\sigma(t,S) + \lambda(t,S)\zeta(t)]S(t)N(t)\,\mathrm{d}W(t), \quad t \geq 0. \end{aligned} \tag{6.6}$$

We assume that all trading is continuous, except for at $t = 0$ where a discrete trade is allowed. Thus N_0 is the number of shares and S_0 is the stock price after after the initial trade. If we want to replicate the contingent claim $g(S(T))$, we need to solve the FBSDE

$$\begin{cases} \mathrm{d}N(t) &= \eta(t)\,\mathrm{d}t + \zeta(t)\,\mathrm{d}W(t) \\ \mathrm{d}S(t) &= [\mu(t,S) + \lambda(t,S)\eta(t)]S(t)\,\mathrm{d}t + [\sigma(t,S) + \lambda(t,S)\zeta(t)]S(t)\,\mathrm{d}W(t) \\ \mathrm{d}Y(t) &= r(t,S)Y(t)\,\mathrm{d}t + [\mu(t,S) - r(t,S) + \lambda(t,S)\eta(t)]S(t)N(t)\,\mathrm{d}t \\ &\quad + [\sigma(t,S) + \lambda(t,S)\zeta(t)]S(t)N(t)\,\mathrm{d}W(t) \\ N(0) &= N_0, \quad S(0) = S_0, \quad Y(T) = g(S(T)), \quad t \in [0,T]. \end{cases} \tag{6.7}$$

6.2 A modified Black-Scholes equation

If we set $Z(t) = [\sigma(t, S(t) + \lambda(t,S)\zeta(t)]S(t)N(t)$, we can apply the four step scheme to (6.7) to find an adapted solution (S, Y, N).

Step 1. The coefficient of the diffusion term in the stock price is $[\sigma(t,S) + \lambda(t,S)\zeta(t)]S(t)$ which gives

$$\begin{aligned} &Z(t) = \theta_S(t,S)[\sigma(t,S) + \lambda(t,S)\zeta(t)]S(t) = [\sigma(t,S(t) + \lambda(t,S)\zeta(t)]S(t)N(t) \\ &\Rightarrow N(t) = \theta_S(t,S). \end{aligned} \tag{6.8}$$

Step 2. We suppress the dependence upon the variables t and S until further notice in order to save space and facilitate presentation. We have the following PDE for θ

$$\theta_t + \frac{1}{2}S^2[\sigma + \lambda\zeta]^2\theta_{SS} + S[\mu + \lambda\eta]\theta_S - (r\theta + S[\mu - r + \lambda\eta]\theta_S) = 0 \tag{6.9}$$

which implies

$$\begin{cases} \theta_t + \frac{1}{2}S^2[\sigma + \lambda\zeta]^2\theta_{SS} + rS\theta_S - r\theta = 0 \\ \theta(T, S) = g(S), \qquad (t, S) \in [0, T] \times (0, \infty) \end{cases} \tag{6.10}$$

In order to find the processes η and ζ, we utilise the relationship $N = \theta_S$

$$\begin{aligned} \eta\,\mathrm{d}t + \zeta\,\mathrm{d}W = \mathrm{d}N = \mathrm{d}(\theta_S) &= [\theta_{St} + S\theta_{SS}(\mu + \lambda\eta) + \tfrac{1}{2}S^2\theta_{SSS}(\sigma + \lambda\zeta)^2]\,\mathrm{d}t \\ &\quad + S\,\theta_{SS}(\sigma + \lambda\zeta)\,\mathrm{d}W \end{aligned} \tag{6.11}$$

and compare the diffusion coefficients to eliminate ζ

$$\zeta = S(\sigma + \lambda\zeta)\theta_{SS} \Rightarrow \zeta = \frac{\sigma S\theta_{SS}}{1 - \lambda S\theta_{SS}} \tag{6.12}$$

where we assume $1 - \lambda S\theta_{SS} \neq 0$. For η, we compare the drift coefficients

$$\begin{aligned} \eta &= \theta_{St} + S\theta_{SS}(\mu + \lambda\eta) + \tfrac{1}{2}S^2\theta_{SSS}(\sigma + \lambda\zeta)^2 \\ \Rightarrow \eta &= \frac{1}{1 - \lambda S\theta_{SS}}\left\{\theta_{St} + \mu S\theta_{SS} + \frac{\sigma^2 S^2\theta_{SSS}}{2(1 - \lambda S\theta_{SS})^2}\right\} \end{aligned} \tag{6.13}$$

Thus, we have the modified Black-Scholes PDE

$$\begin{cases} \theta_t + \dfrac{\sigma(t,S)^2 S^2\theta_{SS}}{2[1 - \lambda(t,S)S\theta_{SS}]^2} + r(t,S)S\theta_S - r(t,S)\theta = 0 \\ \theta(T, S) = g(S), \qquad (t, S) \in [0, T] \times (0, \infty) \end{cases} \tag{6.14}$$

Step 3 Substituting our expressions for η and ζ into (6.5), we have

$$\begin{cases} \mathrm{d}S(t) = \tilde{\mu}(t,S)S(t)\,\mathrm{d}t + \tilde{\sigma}(t,S)S(t)\,\mathrm{d}W(t) \\ S(0) = S_0 \end{cases} \tag{6.15}$$

where

$$\tilde{\mu} = \mu + \lambda\eta = \frac{\mu + \lambda\theta_{St}}{1 - \lambda S\theta_{SS}} + \frac{\lambda\sigma^2 S^2\theta_{SSS}}{2(1 - \lambda S\theta_{SS})^3} \tag{6.16}$$

and

$$\tilde{\sigma} = \sigma + \lambda\zeta = \frac{\sigma}{1 - \lambda S\theta_{SS}}. \tag{6.17}$$

Step 4. To conclude we have the adapted solution $(S, Y, N) = (S(t), \theta(t, S), \theta_S(t, S))$, where $Y(t)$ is the price of the contingent claim $g(S(T))$.

As always, we need to impose regularity conditions on the coefficient functions in order for our PDE to be uniquely solvable. In particular, as we saw in the general case of the four step scheme, we require the function g to be bounded. This excludes some of the most common contingent claims, such as European calls and puts. This restriction can be changed to include functions which are unbounded but regular, such as the mentioned options.

We can also see that the only price impact from the stock price after the initial trade is the stock volatility $\tilde{\sigma}$, just as in the traditional case. We can thus see that if the trader's actions only influence the instantaneous expected return $\tilde{\mu}$ of the stock, there will be no difference in price compared to the case with no price impact.

Bibliography

[1] B. Øksendal. *Stochastic differential equations, An introduction with applications.* Universitext. Springer-Verlag, Berlin Heidelberg New York, 5^{th} edition, 1998.

[2] I. Karatzas and S. E. Shreve. *Brownian motion and stochastic calculus*, volume 113 of *Grad. Texts in Math.* Springer-Verlag, New York, 2^{nd} edition, 1991.

[3] J. Ma and J. Yong. *Forward-backward stochastic differential equations and their applications*, volume 1702 of *Lecture Notes in Math.* Springer-Verlag, Berlin, 1999.

[4] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, and E. Mishchenko. *The mathematical theory of optimal processes.* International series of monographs in pure and applied mathematics. Interscience Publishers, London New York, 1962.

[5] O.A. Ladyzenskaja, V. A. Solonnikov, and N. N. Ural'ceva. *Linear and quasi-linear equations of parabolic type.* American Mathematical Society, Providence, RI., translation of mathematical monographs edition, 1968.

[6] Michael J. Brennan and Eduardo S. Schwartz. A continuous time approach to the pricing of bonds. *Journal of Banking & Finance*, 3(2):133–155, July 1979.

[7] Michael Hogan. Problems in certain two-factor term structure models. *The Annals of Applied Probability*, 3(2):576–581, May 1993.

[8] D. Duffie, J. Ma, and J. Yong. Black's console rate conjecture. *Ann. Appl. Probab.*, 5(2):356–382, 1995.

[9] Hong Liu and Jiongmin Yong. Option pricing with an illiquid underlying asset market. *Journal of Economic Dynamics and Control*, 29(12):2125–2156, December 2005.