Abstract

Backward Stochastic Differential Equations (BSDEs) are the stochastic counterpart of semi-linear parabolic differential equations. Since the discovery of the generalized Feynman-Kac formula, BSDEs gained a wide range of applications in economics and more generally in optimal control. As it is not possible to solve BSDEs analytically, it is logical to ask for numerical methods approximating the unique solution of this type of equations. This thesis presents the basic theory of BSDEs where the generator satisfies the Lipschitz condition. Moreover, a system of a BSDE associated with a state process satisfying some forward SDE results in a so-called Forward-Backward SDE, satisfying the Markovian property, will be introduced. Also, a Fourier-Cosine method for computing the solutions of BSDEs, which proved to be very efficient, is performed where a computational example is discussed.
Acknowledgements

This thesis marks one of the final steps in obtaining the Master of Science degree in Stochastics and Financial Mathematics at the University of Amsterdam (UvA). The research was conducted during an internship at RiskQuest.

First of all, I would like to thank RiskQuest for this opportunity as well as all my colleagues for their support. I would like to thank my supervisor, Asma Kheder, for her guidance, especially during the last months of my thesis. Without her support I would not have been able to go through this very challenging master thesis. Measure theory and Stochastic Integration are one of the most abstract and theoretical courses in my masters and therefore the thesis was a big challenge for me.

I would like to thank also my friend Jenya for the late working hours in the library, as well as for motivating me to go swimming. Sometimes presence is all what you need. Finally, my last thanks goes to my godmother, Justine, who supported me during my studies and who made it possible for me to study abroad for so many years. I am grateful for all the different places, experiences and people I met, it definitely made my life more diverse as well as complete.

Dla mojego ojca, który nigdy nie przestał być tatą.
# Contents

**Introduction** 6

1. **Backward Stochastic Differential Equations** 8  
   1.1. Motivation and preliminaries ................................. 8  
   1.2. Existence and Uniqueness of a solution to Backward Stochastic Differential Equation ........................................... 10  
   1.3. Linear BSDE .................................................................. 19  
   1.4. Comparison Theorem .................................................... 22  
   1.5. Applications of BSDEs in finance ............................... 23  
      1.5.1. Stochastic Optimal Control Problem ....................... 23  
      1.5.2. Hedging in the Black-Scholes model ....................... 25  

2. **FBSDEs and Feynman-Kac formula** 28  
   2.1. Forward-Backward Stochastic Differential Equations ....... 28  
      2.1.1. Regularity Properties of Solutions ......................... 29  
      2.1.2. Markov Properties of Solutions .............................. 30  
   2.2. The link with PDEs: Feynman-Kac formula ................. 30  

3. **Numerical Methods for BSDEs** 34  
   3.1. Discretization of the SDE ............................................ 35  
      3.1.1. Euler-Maruyama scheme ........................................ 35  
      3.1.2. The Milstein Scheme ............................................ 36  
      3.1.3. The weak Taylor Scheme of order 2.0 ....................... 36  
   3.2. Discretization of the BSDE ........................................ 39  

4. **The Fourier-Cosine method for Backward Stochastic Differential Equations** 42  
   4.1. Fourier Transform & Fourier Cosine series ................. 42  
   4.2. Approximations of the conditional expectations .......... 46  
      4.2.1. COS approximation of the function $Z_i^\Pi$ .............. 49  
      4.2.2. COS approximation of the function $Y_i^\Pi$ .............. 49  
      4.2.3. Fourier-cosine coefficients .................................. 50  
      4.2.4. Discretization of the Fourier-cosine coefficient ...... 50  
   4.3. BCOS summarized .................................................... 51  

5. **Numerical experiment: Hedging in the Black-Scholes model** 52  

A. **Appendix** 54
| Conclusion | 57 |
| Popular summary | 58 |
Introduction

Differential equations first came into existence in 1671, where Isaac Newton solved differential equations using infinite series and discussed the non-uniqueness of solutions. They are used in describing different phenomena mathematically, for instance, some can be modelled by ordinary differential equations (ODE)

\[ \frac{dx(t)}{dt} = f(t, x), \]

which is applied in different branches of science. However, noise, which is a sort of randomness cannot be omitted. So a better approximation to the reality is

\[ \frac{dx(t)}{dt} = f(t, x) + \text{'white noise'}, \]  \hspace{1cm} (0.1)

where 'white' means that the noise contains frequencies analogously to white light. Thus this leads us to the mathematical concept of stochastic differential equations (SDE).

In 1827, Robert Brown discovered the random movement of particles suspended in a fluid: Brownian motion. The mathematics behind Brownian motion were studied by many great mathematicians, among others Norbert Wiener after whom Wiener process is named. This stochastic process is denoted \( W_t \) where \( t \geq 0 \). The sample paths of a Brownian motion are continuous but proved to be nowhere differentiable and even of unbounded variation, so classic calculus fails to define integration with respect to it. Nevertheless for describing white noise it is the best candidate.

In 1940, Kiyosi Itô defined Itô stochastic integral and proved the Itô isometry (Lemma A.7). So the above equation (0.1) can be expressed as the following SDE

\[ dx(t, \omega) = f(t, x)dt + dW_t. \]

The notion of time is of big importance for SDEs, in contradiction to ODEs, where initial value problem and terminal value problem are treated in the same way.

Solving an SDE with known terminal value implies the manipulation of a system to achieve a prescribed aim in a randomly disturbed circumstance. A typical SDE on the time interval [0, T] has the form

\[ \begin{aligned}
X_t &= b(t, X_t)dt + \sigma(t, X_t)dW_t, \\
X_0 &= x.
\end{aligned} \]
This SDE can be solved by Picard iteration and Banach fixed point theorem and there exists a unique adapted solution under certain conditions.

However, this routine fails to give adapted solutions when the SDE is defined via the terminal value $X_T = \xi \in \mathcal{F}_T$, known as the Backward stochastic differential equation (BSDE).

In 1973, the BSDEs were first introduced by J.M. Bismut [2] in the case where $f$ is linear w.r.t. $(Y, Z)$. He used BSDEs to study stochastic optimal control problems in the stochastic version of the Pontryagin’s maximum principle.

In 1990, the theory of BSDEs was a main focus for many academic researchers and a large number of publications have been made. However the most famous authors are Pardoux and Peng [19], who studied the BSDEs of the following form,

$$
\begin{aligned}
-dY_t &= f(t, Y_t, Z_t)dt - Z_t dW_t, \\
Y_T &= \xi,
\end{aligned}
$$

and proved the existence and uniqueness of a solution to the latter BSDE where the generator satisfies the Lipschitz condition. In 1992, Peng [23] introduced the comparison theorem which provides a sufficient condition for the wealth process to be nonnegative. In [20], Pardou and Peng showed that the solution of the BSDE in the Markovian case corresponds to a probabilistic solution of a non-linear PDE, and gave a generalization of the Feynman-Kac formula. Since the discovery of the generalized Feynman-Kac formula, applications in mathematical finance and economics started to gain a lot of attention. Moreover in [22], [25], [24], Peng stated the connection between the BSDEs associated with a state process satisfying some forward classical SDEs, and PDEs in the Markovian cases. The theories behind can be used for European option pricing in the constrained Markovian case.

In 1995, in [9] the theory of contingent claim valuation, particularly the Black-Scholes-Merton model for option pricing is expressed in terms of a solution to a BSDE. In 2000, Kohlmann and Zhou [15], developed Peng’s stochastic optimal control theory and interpreted BSDEs as equivalent to stochastic control problems.

Nowadays, Backward Stochastic Differential Equations (BSDEs) remain an interesting field attracting lots of well-known researchers’ investigation. Therefore, it is interesting to know how the theory of BSDE and its applications is being developed. This thesis is organized as follows. In Chapter 1 we provide a general introduction to BSDEs and discuss uniqueness and existence results as well as the Comparison Theorem. In Chapter 2 Forward-Backward SDEs with their Markovian property are introduced and the generalized Feynman-Kac theorem is proven in both directions. This gives us an important link between BSDEs and PDEs. To obtain a numerical solution to a FB-SDE we present in Chapter 3 discretization methods for SDEs and introduce the theta-discretization for BSDEs which results in expressions with conditional expectations. To compute these conditional expectations, we develop, in Chapter 4, a Fourier-Cosine method, called BCOS method. Finally, in Chapter 5 a computational example is given.
1. Backward Stochastic Differential Equations

The aim of this chapter is to get familiar with the notion of Backward Stochastic Differential Equations (BSDEs). After introducing the notation and the main definitions of BSDEs, we will prove two important theorems: the existence and uniqueness of a solution for a BSDE satisfying the Lipschitz condition and the Comparison Theorem. We will finish this chapter by providing some examples. The following references [9], [18] are used.

1.1. Motivation and preliminaries

Let \((\Omega, \mathcal{F}, \mathbb{P})\) be a complete probability space, where the \(\mathbb{R}\)-valued standard Brownian motion \(W\) is defined with respect to the natural filtration \(\mathcal{F}_t\). We consider a simple ordinary differential equation (ODE)

\[
dY_t = 0, \quad 0 \leq t \leq T,
\]

(1.1)

where \(T > 0\) is a given terminal time. For any \(\xi \in \mathbb{R}\) we can suppose either \(Y_0 = \xi\) or \(Y_T = \xi\), such that the above ODE has a unique solution \(Y_t = \xi\). However, if the equation (1.1) is considered as a Stochastic Differential Equation (SDE), we are in a different setting. The solution of (1.1) seen in a stochastic sense should be adapted to the filtration \(\{\mathcal{F}_t\}_{t \geq 0}\). Therefore specifying \(Y_T\) or \(Y_0\) does a big difference.

Consider the ODE (1.1) with the following terminal condition:

\[
\begin{cases}
  dY_t = 0 & 0 \leq t \leq T, \\
  Y_T = \xi,
\end{cases}
\]

(1.2)

where \(\xi\) is a random \(\mathcal{F}_T\)-measurable and square-integrable variable. Since (1.2) is an ODE with a unique solution given by \(Y_t = \xi\), which is not necessarily adapted to the filtration \(\mathcal{F}_t\), unless \(\xi\) is constant, the equation (1.2) viewed as an SDE, does not have a solution in general. Note that in the case where an initial condition is given, the solution is adapted.

To deal with this terminal value problem we will reformulate (1.2) in such a way that we can ensure the adaptability of the solution to \(\mathcal{F}_t\). We introduce the following conditional expectation,

\[
Y_t := E(\xi|\mathcal{F}_t) \quad 0 \leq t \leq T.
\]

(1.3)
Since $\xi$ is $\mathcal{F}_T$-adapted we have $Y_T = \xi$. As the process $(Y_t)_{t \geq 0}$ defined by (1.3) is square-integrable $\mathcal{F}_t$-martingale, we get by the Martingale Representation Theorem (A.2),

$$Y_t = Y_0 + \int_0^t Z_s dW_s, \quad 0 \leq t \leq T \quad \text{a.s.,}$$

(1.4)

where $Z_s$ is a $\mathcal{F}_t$-adapted square integrable process. From (1.3) and (1.4) we get,

$$
\begin{cases} 
\text{d}Y_t = Z_t dW_t & 0 \leq t \leq T \\
Y_T = \xi.
\end{cases}
$$

(1.5)

In other words, we reformulated (1.2) by (1.5) and what is of big importance is that instead of seeking only for one $\mathcal{F}_t$-adapted stochastic process $Y$ as a solution to the SDE, we are looking for a pair $(Y, Z)$. This method by adding the extra component $Z$ to the solution makes it possible to find an $\mathcal{F}_t$-adapted solution.

We will rewrite the backward SDE (1.5) in an integral form. To do this, the SDE is first evaluated in (1.4) with the terminal condition $Y_T = \xi$ and solved for $Y_0$.

$$Y_0 = Y_T - \int_0^T Z_s dW_s = \xi - \int_0^T Z_s dW_s.$$

(1.6)

Plugging (1.6) into (1.4) we get

$$Y_t = \xi - \int_t^T Z_s dW_s \quad 0 \leq t \leq T.$$ 

(1.7)

Hence we have established the idea of how Backward SDEs appeared and that they can be represented by (1.5) or (1.7).

We introduce the following notations which will be used throughout the thesis:

- $T > 0$, the so-called terminal time.
- $\{\mathcal{F}_t; t \in [0, T]\}$, the filtration generated by the Brownian motion $W$ and augmented by all the $\mathbb{P}$-null sets.
- $\mathcal{P}$ the $\sigma$-field of predictable sets on $\Omega \times [0, T]$.
- $\mathcal{B} = \mathcal{B}(\mathbb{R})$, the Borel $\sigma$-algebra and the Lebesgue measure on $\mathcal{B}$ is denoted by $\lambda$.
- $L^2_T(\mathbb{R})$, the space of all $\mathcal{F}_T$-measurable random variables $X : \Omega \to \mathbb{R}$ satisfying 
$$\|X\|^2 = \mathbb{E}|X|^2 < \infty.$$
- $H^2_T(\mathbb{R})$, the space of all predictable processes $\psi : \Omega \times [0, T] \to \mathbb{R}$ such that
$$\mathbb{E} \left( \int_0^T |\psi_t|^2 \text{d}t \right)^\alpha < \infty,$$
where $\alpha > 0$.

- For fixed $\beta \in \mathbb{R}^+$ and $\psi \in H^2_T(\mathbb{R})$, $||\psi||^2_{\beta} := \mathbb{E} \left( \int_0^T e^{\beta t} |\psi_t|^2 \text{d}t \right)$. $H^2_T(\mathbb{R})$ denotes the space $H^2_T(\mathbb{R})$ endowed with the norm $|| \cdot ||_{\beta}$. 

9
For notational simplicity we sometimes use: $L^2_T(\mathbb{R}) = L^2_T$, $\mathbb{H}^0_T(\mathbb{R}) = \mathbb{H}^0_T$, $\mathbb{H}^2_{T,\beta}(\mathbb{R}) = \mathbb{H}^2_{T,\beta}$ and for $f(\omega, t, Y_t, Z_t) = f(t, Y_t, Z_t)$.

**Definition 1.1.** Let $\xi : \Omega \to \mathbb{R}$ be an $\mathcal{F}_T$-measurable random variable and let $f : \Omega \times [0, T] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be a $\mathcal{P} \otimes \mathcal{B} \otimes \mathcal{B}$-measurable mapping. Then,

$$
\begin{cases}
-dY_t = f(\omega, t, Y_t, Z_t)dt - Z_tdW_t; & t \in [0, T], \\
Y_T = \xi,
\end{cases}
$$

(1.8)

or, equivalently,

$$Y_t = \xi + \int_t^T f(\omega, s, Y_s, Z_s)ds - \int_t^T Z_s dW_s \quad t \in [0, T]
$$

(1.9)

is called a Backward Stochastic Differential Equation (BSDE) with terminal value $\xi$ and generator $f$.

Properties of different BSDEs, such as, *linear BSDEs*, *Lipschitz BSDEs*, *Markovian BSDEs*, *quadratic BSDEs*, etc. are properties of the generator $f$. In other words a linear BSDE represents a BSDE with a linear generator, i.e. $f(t, Y_t, Z_t) = \alpha_t Y_t + \beta_t Z_t$. In this thesis we will only consider Lipschitz BSDEs, unless stated otherwise.

**Definition 1.2.** Suppose that $\xi \in L^2_T(\mathbb{R})$, $f(\cdot, 0, 0) \in \mathbb{H}^2_T(\mathbb{R})$ and assume $f$ is uniformly Lipschitz continuous; i.e. there exists constant $C > 0$ such that $d\mathbb{P} \otimes dt$-a.s.

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

holds true $\forall (y_1, z_1), (y_2, z_2) \in \mathbb{R} \times \mathbb{R}$. Then the pair $(f, \xi)$ is said to be a set of standard parameters of the BSDE (1.8).

**Definition 1.3.** A solution to the BSDE (1.8) is a pair of value $(Y, Z)$ such that $\{Y_t; \ t \in [0, T]\}$ is a continuous $\mathcal{F}_t$-adapted process and $\{Z_t; \ t \in [0, T]\}$ is an $\mathbb{R}$-valued predictable process satisfying $\int_0^T |Z_s|^2 ds < \infty$, $\mathbb{P}$-a.s..

Note that a solution $(Y, Z) \in \mathbb{H}^2_T \times \mathbb{H}^2_T$, when it exists, is often referred to a square-integrable solution.

### 1.2. Existence and Uniqueness of a solution to Backward Stochastic Differential Equation

In this section we present in detail the existence and uniqueness of a solution to the BSDE (1.8). Following the derivations in [9] we first introduce a priori estimates that we need to prove the main Theorem 1.5.
Lemma 1.4. [3] Let \((f^i, \xi^i)\) \(i = 1, 2\) be two standard parameters of the BSDE (1.8) and \(((Y^i, Z^i); i = 1, 2)\) be two square-integrable solutions to the corresponding standard parameters. Consider \(C\) be the Lipschitz constant of \(f^1\) and denote

\[
\begin{align*}
\delta Y_t &= Y^1_t - Y^2_t \\
\delta Z_t &= Z^1_t - Z^2_t \\
\delta_2 f_t &= f^1(t, Y^2_t, Z^2_t) - f^2(t, Y^1_t, Z^1_t).
\end{align*}
\]

Then for any \((\lambda, \mu, \beta)\) such that \(\mu > 0, \lambda^2 > C \) and \(\beta \geq C(2 + \lambda^2) + \mu^2\), it follows that

\[
\begin{align*}
\|\delta Y\|^2_\beta &\leq T \left[ e^{2T} \mathbb{E}(|\delta Y_T|^2) + \frac{1}{\mu^2} \|\delta_2 f\|^2_\beta \right], \\
\|\delta Z\|^2_\beta &\leq \frac{\lambda^2}{\lambda^2 - C} \left[ e^{2T} \mathbb{E}(|\delta Y_T|^2) + \frac{1}{\mu^2} \|\delta_2 f\|^2_\beta \right].
\end{align*}
\] (1.10) (1.11)

Proof. We divide the proof into four steps.

1. Let \((Y, Z) \in \mathbb{H}_T^2 \times \mathbb{H}_T^2\) be a solution of the BSDE where \((f, \xi)\) are the standard parameters,

\[
Y_t = \xi + \int_t^T f(s, Y_s, Z_s)ds - \int_t^T Z_s dW_s \quad t \in [0, T].
\] (1.12)

In this step we want to show that \(\sup_{0 \leq t \leq T} |Y_t| \leq L_T\). To show this we will take the absolute value on both sides as well as the supremum and show that the right hand side of equation (1.12) is in \(L_T^2\). Taking the absolute value on both sides of equation (1.12), we have

\[
\begin{align*}
|Y_t| &= \left| \xi + \int_t^T f(s, Y_s, Z_s)ds - \int_t^T Z_s dW_s \right| \\
|Y_t| &\leq \left| \xi \right| + \left| \int_t^T f(s, Y_s, Z_s)ds \right| + \left| \int_t^T Z_s dW_s \right| \quad \text{(by the triangle inequality)}
\end{align*}
\]

\[
\begin{align*}
|Y_t| &\leq \left| \xi \right| + \int_t^T |f(s, Y_s, Z_s)| ds + \left| \int_t^T Z_s dW_s \right| \quad \text{(by Jensen inequality)}
\end{align*}
\]

\[
\sup_{0 \leq t \leq T} |Y_t| \leq \left| \xi \right| + \int_0^T |f(s, Y_s, Z_s)| ds + \sup_{0 \leq t \leq T} \left| \int_t^T Z_s dW_s \right|.
\] (1.13)

By assumption, \(\xi \in L_T^2\), i.e. \(\mathbb{E}(|\xi|^2) < \infty\) so we have that \(\left| \xi \right| \in L_T^2\). For the second term in (1.13) we will use the Lipschitz continuity of \(f\), we obtain,

\[
|f(s, Y_s, Z_s)| \leq |f(s, Y_s, Z_s) - f(s, 0, 0)| + |f(s, 0, 0)| \\
\leq C(|Y_s| + |Z_s|) + |f(s, 0, 0)|. \quad (*)
\]
Moreover, we have,

\[
\mathbb{E}\left( \int_0^T |f(s, Y_s, Z_s)| \, ds \right)^2
\leq T \mathbb{E}\left( \int_0^T |f(s, Y_s, Z_s)|^2 \, ds \right) \quad \text{(by Cauchy-Schwarz inequality)}
\leq T \mathbb{E}\left( \int_0^T (C(|Y_s| + |Z_s|) + |f(s, 0, 0)|)^2 \, ds \right) \quad \text{(by (**))}
\leq 4TC^2 \mathbb{E}\left( \int_0^T |Y_s|^2 \, ds \right) + 4TC^2 \mathbb{E}\left( \int_0^T |Z_s|^2 \, ds \right)
\quad + 2T \mathbb{E}\left( \int_0^T |f(s, Y_s, Z_s)|^2 \, ds \right) \quad \text{(by Lemma A.5)}
< \infty.
\]

Hence \( \int_0^T |f(s, Y_s, Z_s)| \, ds \in L_2^2 \). For the third term in (1.13), we get

\[
\mathbb{E}\left( \sup_{0 \leq t \leq T} \left| \int_t^T Z_s dW_s \right|^2 \right)
= \mathbb{E}\left( \sup_{0 \leq t \leq T} \left| \int_t^T Z_s dW_s - \int_0^t Z_s dW_s \right|^2 \right)
\leq 2 \mathbb{E}\left( \left| \int_0^T Z_s dW_s \right|^2 \right) + 2 \mathbb{E}\left( \sup_{0 \leq t \leq T} \left| \int_0^T Z_s dW_s \right|^2 \right) \quad \text{(by Lemma A.5)}
\leq 2 \mathbb{E}\left( \left| \int_0^t Z_s dW_s \right|^2 \right) + 2 \mathbb{E}\left( \sup_{0 \leq t \leq T} \left| \int_0^t Z_s dW_s \right|^2 \right).
\]

By Itô’s isometry (A.7) we have:

\[
\mathbb{E}\left( \left| \int_0^t Z_s dW_s \right|^2 \right) = \mathbb{E}\left( \int_0^t |Z_s|^2 \, ds \right).
\]

By Burkholder-Davis-Gundy inequality (A.1), there exists a constant \( K > 0 \) such that

\[
\mathbb{E}\left( \sup_{0 \leq t \leq T} \left| \int_0^t Z_s dW_s \right|^2 \right) \leq K \mathbb{E}\left( \int_0^t |Z_s|^2 \, ds \right).
\]

So for the third term in (1.13), we get

\[
\mathbb{E}\left( \sup_{0 \leq t \leq T} \left| \int_t^T Z_s dW_s \right|^2 \right) \leq 2(1 + K) \mathbb{E}\left( \int_0^t |Z_s|^2 \, ds \right) < \infty.
\]

Thus we showed \( \sup_{0 \leq t \leq T} \left| \int_t^T Z_s dW_s \right| \in L_2^2 \), which leads to \( \sup_{0 \leq t \leq T} |Y_t| \in L_2^2 \).
2. Consider the two solutions \((Y^1, Z^1)\) and \((Y^2, Z^2)\) of the BSDE (1.8) with the standard parameters \((f^1, \xi^1)\) and \((f^2, \xi^2)\), respectively. By Itô formula applied to \(e^{\beta s} |\delta Y_s|^2\), one obtains
\[
d(e^{\beta s} |\delta Y_s|^2) = \beta e^{\beta s} |\delta Y_s|^2 \, ds + 2e^{\beta s} \delta Y_s d\delta Y_s + e^{\beta s} d\delta Y_s d\delta Y_s.
\]
Integrating the above equation from \(t\) to \(T\), we have
\[
e^{\beta t} |\delta Y_t|^2 + \beta \int_t^T e^{\beta s} |\delta Y_s|^2 \, ds + \int_t^T e^{\beta s} |\delta Z_s|^2 \, ds
\]
\[
= e^{\beta T} |\delta Y_T|^2 + 2 \int_t^T e^{\beta s} \delta Y_s (f^1(s, Y^1_s, Z^1_s) - f^2(s, Y^2_s, Z^2_s)) \, ds
\]
\[
- 2 \int_t^T e^{\beta s} \delta Y_s \delta Z_s dW_s.
\]

3. In this step, we want to show
\[
\mathbb{E} \left( e^{\beta t} |\delta Y_t|^2 \right) \leq \mathbb{E} \left( e^{\beta T} |\delta Y_T|^2 \right) + \frac{1}{\mu^2} \mathbb{E} \left( \int_t^T e^{\beta s} |\delta_{2f s}|^2 \, ds \right) \quad t \in [0, T].
\]
From step 1, we know \(\sup_{0 \leq t \leq T} |Y_t| \in \mathbb{L}^2_T\), which leads to
\[
\mathbb{E} \left( \sqrt{\int_0^T e^{\beta s} \delta Z_s \delta Y_s \, ds} \right)
\]
\[
= \mathbb{E} \left( \sqrt{\int_0^T e^{2\beta s} |\delta Z_s|^2 |\delta Y_s|^2 \, ds} \right)
\]
\[
\leq \mathbb{E} \left( \sup_{0 \leq t \leq T} |\delta Y_t| \sqrt{\int_0^T e^{2\beta s} |\delta Z_s|^2 \, ds} \right)
\]
\[
\leq \mathbb{E} \left( \sup_{0 \leq t \leq T} |\delta Y_t|^2 \right) \sqrt{\mathbb{E} \left( \int_0^T e^{2\beta s} |\delta Z_s|^2 \, ds \right)} \quad \text{(by Cauchy Schwarz inequality)}
\]
\[
< \infty.
\]
We showed that \(e^{\beta s} \delta Z_s \delta Y_s\) belongs to \(\mathbb{L}^1_T\), which implies that the stochastic integral \(\int_t^T e^{\beta s} \delta Y_s \delta Z_s dW_s\) becomes \(\mathbb{P}\)-integrable and has zero expectation. Moreover,
\[
|f^1(s, Y^1_s, Z^1_s) - f^2(s, Y^2_s, Z^2_s)|
\]
\[
= |f^1(s, Y^1_s, Z^1_s) - f^1(s, Y^2_s, Z^2_s) + f^1(s, Y^2_s, Z^2_s) - f^2(s, Y^2_s, Z^2_s)|
\]
\[
\leq |f^1(s, Y^1_s, Z^1_s) - f^1(s, Y^2_s, Z^2_s)| + |f^1(s, Y^2_s, Z^2_s) - f^2(s, Y^2_s, Z^2_s)|
\]
\[
\leq C (|\delta Y_s| + |\delta Z_s|) + |\delta_{2f s}|
\]
\[ 2 \left| \langle \delta Y_s, f^1(s, Y^1_s, Z^1_s) - f^2(s, Y^2_s, Z^2_s) \rangle \right| \]
\[ \leq 2 |\delta Y_s| \left| f^1(s, Y^1_s, Z^1_s) - f^2(s, Y^2_s, Z^2_s) \right| \]
\[ \leq 2C |\delta Y_s|^2 + 2 |\delta Y_s| (C |\delta Z_s| + |\delta_2 f_s|), \]

where \( C \) is the Lipschitz constant. Taking the expectation in (1.14) we get:

\[
\mathbb{E} \left( e^{\beta t} |\delta Y_t|^2 \right) + \beta \mathbb{E} \left( \int_t^T e^{\beta s} |\delta Y_s|^2 \, ds \right) + \mathbb{E} \left( \int_t^T e^{\beta s} |\delta Z_s|^2 \, ds \right) \\
= \mathbb{E} \left( e^{\beta T} |\delta Y_T|^2 \right) + 2 \mathbb{E} \left( \int_t^T e^{\beta s} \delta Y_s (f^1(s, Y^1_s, Z^1_s) - f^2(s, Y^2_s, Z^2_s)) \, ds \right). \tag{**} 
\]

We take the last term in the above equation and we rewrite as follows,

\[
\mathbb{E} \left( \int_t^T 2e^{\beta s} \delta Y_s (f^1(s, Y^1_s, Z^1_s) - f^2(s, Y^2_s, Z^2_s)) \, ds \right) \\
\leq \mathbb{E} \left( \int_t^T 2e^{\beta s} |\delta Y_s| ||f^1(s, Y^1_s, Z^1_s) - f^2(s, Y^2_s, Z^2_s)|| \, ds \right) \\
\leq \mathbb{E} \left( \int_t^T 2e^{\beta s} C |\delta Y_s|^2 + 2e^{\beta s} |\delta Y_s| (C |\delta Z_s| + |\delta_2 f_s|) \, ds \right) \\
\leq \mathbb{E} \left( \int_t^T 2e^{\beta s} C |\delta Y_s|^2 + e^{\beta s} \left( C \frac{|\delta Z_s|^2}{\lambda^2} + \frac{|\delta_2 f_s|^2}{\mu^2} + |\delta Y_s|^2 (\mu^2 + C\lambda^2) \right) \, ds \right) \\
= (C(2 + \lambda^2) + \mu^2) \mathbb{E} \left( \int_t^T e^{\beta s} |\delta Y_s|^2 \, ds \right) + C \frac{1}{\lambda^2} \mathbb{E} \left( \int_t^T e^{\beta s} |\delta Z_s|^2 \, ds \right) \\
+ \frac{1}{\mu^2} \mathbb{E} \left( \int_t^T e^{\beta s} |\delta_2 f_s|^2 \, ds \right). 
\]

Now equation (**) becomes

\[
\mathbb{E} \left( e^{\beta t} |\delta Y_t|^2 \right) + \beta \mathbb{E} \left( \int_t^T e^{\beta s} |\delta Y_s|^2 \, ds \right) + \mathbb{E} \left( \int_t^T e^{\beta s} |\delta Z_s|^2 \, ds \right) \\
\leq \mathbb{E} \left( e^{\beta T} |\delta Y_T|^2 \right) + (C(2 + \lambda^2) + \mu^2) \mathbb{E} \left( \int_t^T e^{\beta s} |\delta Y_s|^2 \, ds \right) \tag{1.15} \\
+ \frac{C}{\lambda^2} \mathbb{E} \left( \int_t^T e^{\beta s} |\delta Z_s|^2 \, ds \right) + \frac{1}{\mu^2} \mathbb{E} \left( \int_t^T e^{\beta s} |\delta_2 f_s|^2 \, ds \right). 
\]

By rearranging the terms in (1.15) we get

\[
\mathbb{E} \left( e^{\beta t} |\delta Y_t|^2 \right) \leq \mathbb{E} \left( e^{\beta T} |\delta Y_T|^2 \right) + (C(2 + \lambda^2) + \mu^2 - \beta) \mathbb{E} \left( \int_t^T e^{\beta s} |\delta Y_s|^2 \, ds \right) \\
+ \left( \frac{C}{\lambda^2} - 1 \right) \mathbb{E} \left( \int_t^T e^{\beta s} |\delta Z_s|^2 \, ds \right) + \frac{1}{\mu^2} \mathbb{E} \left( \int_t^T e^{\beta s} |\delta_2 f_s|^2 \, ds \right). 
\]
The above inequality with \( \beta \geq C(2 + \lambda^2) + \mu^2 \) and \( C < \lambda^2 \), becomes
\[
E \left( e^{\beta t} |\delta Y_t|^2 \right) \leq E \left( e^{\beta T} |\delta Y_T|^2 \right) + \frac{1}{\mu^2} E \left( \int_t^T e^{\beta s} |\delta f_s|^2 \, ds \right).
\] (1.16)

4. In this step we integrate (1.16) on both sides from 0 to \( T \) and we use Fubini’s theorem to change the order of the integrals, we get
\[
E \left( \int_0^T e^{\beta t} |\delta Y_t|^2 \, dt \right) \leq T E \left( e^{\beta T} |\delta Y_T|^2 \right) + \frac{1}{\mu^2} E \left( \int_0^T e^{\beta t} |\delta Y_t|^2 \, dt \right).
\]

This is equivalent to
\[
\|\delta Y\|_{\beta}^2 \leq T \left[ e^{\beta T} E |\delta Y_T|^2 + \frac{1}{\mu^2} \|\delta f\|_{\beta}^2 \right].
\]

By rearranging the terms in (1.15), we can also derive
\[
\left( \frac{\lambda^2 - C}{\lambda^2} \right) E \int_t^T e^{\beta s} |\delta Z_s|^2 \, ds \leq E \left( e^{\beta T} |\delta Y_T|^2 \right) - E \left( e^{\beta t} |\delta Y_t|^2 \right)
\]
\[+ \left( C(2 + \lambda^2) + \mu^2 - \beta \right) E \left( \int_t^T e^{\beta s} |\delta Y_s|^2 \, ds \right) \]
\[+ \frac{1}{\mu^2} E \left( \int_t^T e^{\beta s} |\delta f_s|^2 \, ds \right).
\]

Now, the above inequality with \( \beta \geq C(2 + \lambda^2) + \mu^2 \) and \( C < \lambda^2 \), becomes
\[
\left( \frac{\lambda^2 - C}{\lambda^2} \right) E \int_t^T e^{\beta s} |\delta Z_s|^2 \, ds \leq E \left( e^{\beta T} |\delta Y_T|^2 \right) + \frac{1}{\mu^2} E \left( \int_t^T e^{\beta s} |\delta f_s|^2 \, ds \right).
\]

Notice that, we can take \( \beta > 1 \) and \( t = 0 \), which leads to
\[
\|\delta Z\|_{\beta}^2 \leq \frac{\lambda^2}{\lambda^2 - C} \left[ e^{\beta T} E(\|\delta Y_T\|^2) + \frac{1}{\mu^2} \|\delta f\|_{\beta}^2 \right].
\]

\[\square\]

**Theorem 1.5.** \([3]\) Given standard parameters \((f, \xi)\) there exists a unique pair \((Y, Z)\) ∈ \(\mathbb{H}^2_T(\mathbb{R}) \times \mathbb{H}^2_T(\mathbb{R})\) which solves the BSDE (1.8).

**Proof.** We will use the Banach fixed-point theorem for the following mapping:
\[
\Phi : \mathbb{H}^2_T(\mathbb{R}) \times \mathbb{H}^2_T(\mathbb{R}) \to \mathbb{H}^2_T(\mathbb{R}) \times \mathbb{H}^2_T(\mathbb{R})
\]
\[\Phi(U, V) \mapsto (Y, Z)\]

15
where \((Y, Z)\) is a solution of the BSDE (1.8) with generator \(f(s, U_s, V_s)\):

\[
Y_t = \xi + \int_t^T f(s, U_s, V_s) ds - \int_t^T Z_s dW_s \quad 0 \leq t \leq T. \tag{1.17}
\]

We divide the proof into steps.

1. We will show that \(\Psi\) is indeed well defined, i.e. there exists a solution \((Y, Z)\) to (1.17). The assumption \((f, \xi)\) being standard parameters implies that \(f(\cdot, U, V) \in \mathbb{H}^2_T\)

\[
\mathbb{E} \left( \int_0^T |f(s, U_s, V_s)|^2 ds \right) \leq 4C^2 \mathbb{E} \left( \int_0^T |U_s|^2 ds \right) + 4C^2 \mathbb{E} \left( \int_0^T |V_s|^2 ds \right) + 2 \mathbb{E} \left( \int_0^T |f(s, 0, 0)|^2 ds \right) < \infty.
\]

For \(t \in [0, T]\) we have,

\[
\mathbb{E} \left( \left| \int_t^T f(s, U_s, V_s) ds \right|^2 \right) \leq (T - t) \mathbb{E} \left( \int_t^T |f(s, U_s, V_s)|^2 ds \right) \leq T \mathbb{E} \left( \int_0^T |f(s, U_s, V_s)|^2 ds \right) < \infty.
\]

Hence we showed that \(\int_t^T f(s, U_s, V_s) ds \in L^2_T \quad \forall t \in [0, T]\). Consider

\[
M_t := \mathbb{E} \left( \int_0^T f(s, U_s, V_s) ds + \xi \bigg| \mathcal{F}_t \right),
\]

which is a square-integrable martingale. Indeed, notice that since \(\forall s \leq t, \mathcal{F}_s \subset \mathcal{F}_t\), we have,

\[
\mathbb{E} (M_t | \mathcal{F}_s) = \mathbb{E} \left( \mathbb{E} \left( \int_0^T f(s, U_s, V_s) ds + \xi \bigg| \mathcal{F}_t \right) \bigg| \mathcal{F}_s \right) \quad \text{(by the tower property)}
\]

\[
= \mathbb{E} \left( \int_0^T f(s, U_s, V_s) ds + \xi \bigg| \mathcal{F}_s \right) = M_s.
\]

By the Martingale Representation Theorem (A.2), we know there exists a unique integrable process \(Z \in \mathbb{H}^2_T\) such that

\[
M_t = M_0 + \int_0^t Z_s dW_s, \quad t \in [0, T]. \tag{1.18}
\]
Next we define the adapted and continuous process $Y$ as

$$Y_t := M_t - \int_0^t f(s, U_s, V_s)ds.$$  

We obtain

$$Y_t = E \left( \int_0^T f(s, U_s, V_s)ds + \xi \bigg| \mathcal{F}_t \right) - \int_0^t f(s, U_s, V_s)ds$$

$$= E \left( \int_t^T f(s, U_s, V_s)ds + \xi \bigg| \mathcal{F}_t \right).$$

Now it is easy to show that $Y \in \mathbb{H}_T^2$,

$$E \left( \int_0^T |Y_t|^2 dt \right) = E \left( \int_0^T E \left( \int_t^T f(s, U_s, V_s)ds + \xi \bigg| \mathcal{F}_t \right)^2 dt \right)$$

$$= \int_0^T E \left( \left| \int_t^T f(s, U_s, V_s)ds + \xi \right|^2 \bigg| \mathcal{F}_t \right) dt$$

(by Fubini’s theorem)

$$\leq \int_0^T E \left( E \left( \int_t^T f(s, U_s, V_s)ds + \xi \bigg| \mathcal{F}_t \right)^2 \right) dt$$

$$\leq \int_0^T E \left( \left| \int_t^T f(s, U_s, V_s)ds + \xi \right|^2 \right) dt$$

$$\leq 2 \int_0^T E \left( \left| \int_t^T f(s, U_s, V_s)ds \right|^2 + |\xi|^2 \right) dt$$

(by Lemma A.5)

$$< \infty.$$  

In order to conclude this step, we need to clarify that the unique determined pair $(Y, Z)$ solves the equation (1.17). But

$$\xi = Y_T = M_T - \int_0^T f(s, U_s, V_s)ds$$

$$= M_0 + \int_0^T Z_s dW_s - \int_0^T f(s, U_s, V_s)ds.$$  

(1.19)

Using (1.18) and (1.19), we deduce

$$Y_t = M_t - \int_0^t f(s, U_s, V_s)ds$$

$$= M_0 + \int_0^t Z_s dW_s - \int_0^t f(s, U_s, V_s)ds$$

(by using 1.18)
\[\begin{align*}
&= \left( \xi - \int_0^T Z_s dW_s + \int_0^T f(s, U_s, V_s) ds \right) + \int_t^T Z_s dW_s - \int_t^T f(s, U_s, V_s) ds \\
&= \xi + \int_t^T f(s, U_s, V_s) ds - \int_t^T Z_s dW_s \\
&= 0 + \int_t^T f(s, U_s, V_s) ds - \int_t^T Z_s dW_s
\end{align*}\]

(by using 1.19)

and hence we get the claim.

2. Let \((U^1, V^1), (U^2, V^2)\) be two elements of \(\mathbb{H}^2_{T,\beta} \times \mathbb{H}^2_{T,\beta}\), and consider their images

\[\Psi(U^1, V^1) = (Y^1, Z^1)\] and \[\Psi(U^2, V^2) = (Y^2, Z^2)\]. We use the following notation

\[\delta_2 f_s = f(s, U^1_s, V^1_s) - f(s, U^2_s, V^2_s)\]

By Lemma (1.4) applied with \(C = 0\) and \(\beta = \mu^2\), we have

\[
\|\delta Y\|_\beta^2 \leq \frac{T}{\beta} \mathbb{E} \int_0^T e^{\beta s} |f(s, U^1_s, V^1_s) - f(s, U^2_s, V^2_s)|^2 ds
\]

\[
\leq \frac{T}{\beta} C^2 \mathbb{E} \int_0^T e^{\beta s} (|\delta U_s| + |\delta V_s|)^2 ds
\]

\[
\leq 2 \frac{T}{\beta} C^2 (\|\delta U\|_\beta^2 + \|\delta V\|_\beta^2)
\]

and

\[
\|\delta Z\|_\beta^2 \leq \frac{1}{\beta} \mathbb{E} \int_0^T e^{\beta s} |f(s, U^1_s, V^1_s) - f(s, U^2_s, V^2_s)|^2 ds
\]

\[
\leq \frac{2}{\beta} C^2 (\|\delta U\|_\beta^2 + \|\delta V\|_\beta^2)
\]

We deduce that

\[
\|\delta Y\|_\beta^2 + \|\delta Z\|_\beta^2 \leq \frac{2(1 + T)C^2}{\beta} (\|\delta U\|_\beta^2 + \|\delta V\|_\beta^2).
\]

(1.20)

Now, let \(\beta > 2(1 + T)C^2\), we see that the mapping \(\Psi\) is a contraction from \(\mathbb{H}^2_{T,\beta} \times \mathbb{H}^2_{T,\beta}\) onto itself and that there exists a unique fixed point \((\bar{Y}, \bar{Z}) \in \mathbb{H}^2_{T,\beta} \times \mathbb{H}^2_{T,\beta}\) satisfying

\[\bar{Y}_t = \xi + \int_t^T f(s, \bar{Y}_s, \bar{Z}_s) ds - \int_t^T \bar{Z}_s dW_s \quad 0 \leq t \leq T.\]

We can choose the continuous version \(Y\) defined by

\[Y_t := \mathbb{E} \left( \int_t^T f(s, Y_s, Z_s) ds + \xi \mid \mathcal{F}_t \right) = \mathbb{E} \left( \int_t^T f(s, Y_s, Z_s) ds + \xi \mid \mathcal{F}_t \right),\]

i.e. \((Y, \bar{Z})\) is the unique continuous solution of the BSDE.
1.3. Linear BSDE

Linear BSDEs have a very useful property as it has an explicitly solution. In financial mathematics the solution of a linear BSDE is related to the pricing and hedging problem of a contingent claim.

**Proposition 1.6.** [21] Let \((\beta, \mu)\) be a bounded \((\mathbb{R}, \mathbb{R})\)-valued predictable process, \(\varphi\) an element in \(\mathbb{H}^2_T\) and \(\xi \in \mathbb{L}^2_T\). We consider the following linear BSDE:

\[ Y_s = \xi + \int_s^T (\varphi_u + \beta_u Y_u + \mu_u Z_u)du - \int_s^T Z_u dW_u, \quad 0 \leq t \leq T \quad (1.21) \]

a) The equation (1.21) has a unique solution \((Y,Z)\) in \(\mathbb{H}^2_T \times \mathbb{H}^2_T\), where \(Y\) is given by the following explicit formula

\[ Y_t = \mathbb{E} \left[ \xi \Gamma^T_t + \int_t^T \Gamma^s_t \varphi_s ds \mid \mathcal{F}_t \right], \quad \mathbb{P} - a.s. \]

where \((\Gamma^s_t)_{s \geq t}\) is the adjoint process defined by the forward linear SDE

\[ d\Gamma^s_t = \Gamma^s_t (\beta_s ds + \mu_s dW_s), \quad \Gamma^s_t = 1. \quad (1.22) \]

b) If \(\xi, \varphi\) are non negative, then the process \((Y_t)_{t \leq T}\) is non negative. If in addition \(Y_0 = 0\), then for any \(s \geq t\), \(Y_s = 0\), \(\xi = 0\) and \(\varphi_s = 0\), \(d\mathbb{P} \otimes ds - a.s.

**Proof.** The first step is to show that \((f,\xi)\) are the standard parameters. \(\xi \in \mathbb{L}^2_T\) is given and \(f\) is uniformly Lipschitz by the below inequality,

\[ |f(t, y_1, z_1) - f(t, y_2, z_2)| = |\varphi_t + \beta_t y_1 + \mu_t z_1 - \varphi_t - \beta_t y_2 - \mu_t z_2| \]

\[ = |\beta_t (y_1 - y_2) + \mu_t (z_1 - z_2)| \]

\[ \leq C |(y_1 - y_2) + (z_1 - z_2)|, \]

where \(|\beta_t| < K_1, |\mu_t| < K_2\) and \(C = \max(K_1, K_2)\). Since \(\beta\) and \(\mu\) are bounded processes and the pair \((f, \xi)\) are standard parameters as well as the linear generator \(f(t, y, z) = \varphi_t + \beta_t y_t + \mu_t z_t\) satisfies the Lipschitz condition, by Theorem (1.5) there exists a unique square-integrable solution \((Y, Z)\) of the BSDE (1.21).

As \(\beta\) and \(\mu\) are bounded processes we also have the existence and uniqueness result for the SDE (1.22) (i.e. see section 6.2. of [28] ).

In the proof of Theorem (1.5) we have already shown that \(\mathbb{E} \left[ \sup_{0 \leq s \leq T} |Y_s| \right] < \infty\),
in other words that is $\sup_{0 \leq s \leq T} |Y_s| \in \mathbb{L}^2_T$. We will show that $\mathbb{E} \left[ \sup_{0 \leq s \leq T} |\Gamma^t_s|^2 \right] < \infty$, 

\[
\mathbb{E} \left( \sup_{0 \leq s \leq T} |\Gamma^t_s|^2 \right) = \mathbb{E} \left( \sup_{0 \leq s \leq T} \left| \Gamma^t_s + \int_0^t \beta_s \Gamma^t_s \, ds + \int_0^t \mu_s \Gamma^t_s \, dW_s \right|^2 \right) \\
\leq 3 \mathbb{E} \left( \sup_{0 \leq s \leq T} |\Gamma^t_s|^2 \right) + 3 \mathbb{E} \left( \sup_{0 \leq s \leq T} \left| \int_0^t \beta_s \Gamma^t_s \, ds \right|^2 \right) \\
+ 3 \mathbb{E} \left( \sup_{0 \leq s \leq T} \left| \int_0^t \mu_s \Gamma^t_s \, dW_s \right|^2 \right) \\
\leq 3 + 3 \mathbb{E} \left( \int_0^T |\beta_s|^2 |\Gamma^t_s|^2 \, ds \right) + 3 \mathbb{E} \left( \sup_{0 \leq s \leq T} \left| \int_0^t \mu_s \Gamma^t_s \, dW_s \right|^2 \right) \\
\quad (\text{by Jensen inequality}) \\
\leq 3 + 3 \mathbb{E} \left( \int_0^T |\beta_s|^2 |\Gamma^t_s|^2 \, ds \right) + 3 C_p \mathbb{E} \left( \int_0^T |\mu_s|^2 |\Gamma^t_s|^2 \, ds \right) \\
\quad (\text{by BDG inequality (A.1)}) \\
\leq 3 + 3 \mathbb{E} \left( \int_0^T |\beta_s|^2 |\Gamma^t_s|^2 \, ds \right) + 3 C_p \mathbb{E} \left( \int_0^T |\mu_s|^2 |\Gamma^t_s|^2 \, ds \right) \\
\leq 3 + 3 C (1 + C_p) \mathbb{E} \left( \int_0^T |\Gamma^t_s|^2 \, ds \right) \\
< \infty. \\
\quad (\text{as } |\Gamma^t_s| \text{ is square-integrable})
\]

By applying Itô’s formula to $\Gamma^t_s Y_s$, we get

\[
d(\Gamma^t_s Y_s) = \Gamma^t_s dY_s + Y_s d\Gamma^t_s + d\Gamma^t_s dY_s \\
= \Gamma^t_s (-\varphi_s ds - Y_s \beta_s ds - \mu_s Z_s ds + Z_s dW_s) + \Gamma^t_s Y_s (\beta_s ds + \mu_s dW_s) + \Gamma^t_s Z_s \mu_s ds \\
= -\varphi_s \Gamma^t_s ds + \Gamma^t_s (Z_s + Y_s \mu_s) dW_s.
\]

Which implies that the process $(\Gamma^t_s Y_s + \int_0^s \varphi_u \Gamma^t_u \, du)_{t \leq T}$ is a local martingale. If the latter is a martingale, we will get 

\[
\Gamma^t_s Y_t + \int_0^t \varphi_u \Gamma^t_u \, ds = \mathbb{E} \left( \Gamma^t_T \xi + \int_0^T \varphi_s \Gamma^t_s \, ds \mid \mathcal{F}_t \right) \\
\Gamma^t_s Y_t = \mathbb{E} \left( \Gamma^t_T \xi + \int_0^T \varphi_s \Gamma^t_s \, ds \mid \mathcal{F}_t \right),
\]

which implies

\[
Y_t = \mathbb{E} \left( \Gamma^t_T \xi + \int_0^T \varphi_s \Gamma^t_s \, ds \mid \mathcal{F}_t \right),
\]

20
and the claim a) of the proposition is proved. Thus we need to prove that \( \Gamma^t_s Y_s + \int_0^s \varphi_u \Gamma^t_u du \) is a martingale. To do that we prove \( \mathbb{E} \left[ \sup_{0 \leq s \leq T} \left| \Gamma^t_s Y_s + \int_0^s \varphi_u \Gamma^t_u du \right| \right] < \infty \). Notice that

\[
\mathbb{E} \left[ \sup_{0 \leq s \leq T} \left| \Gamma^t_s Y_s + \int_0^s \varphi_u \Gamma^t_u du \right| \right] \leq \mathbb{E} \left[ \sup_{0 \leq s \leq T} |\Gamma^t_s| \right] + \mathbb{E} \left[ \sup_{0 \leq s \leq T} \left| \int_0^s \varphi_u \Gamma^t_u du \right| \right].
\]

(1.23)

From (1.23) we will first show that the first and second terms are finite, we have

\[
\mathbb{E} \left[ \sup_{0 \leq s \leq T} |\Gamma^t_s| \right] \leq \mathbb{E} \left[ \sup_{0 \leq s \leq T} |\Gamma^t_s| \right] \leq \mathbb{E} \left[ \sup_{0 \leq s \leq T} |\Gamma^t_s|^2 \right]^{\frac{1}{2}} \mathbb{E} \left[ \sup_{0 \leq s \leq T} |Y_s|^2 \right]^{\frac{1}{2}} \quad \text{(by Hölder’s inequality)}
\]

< \infty.

\[
\mathbb{E} \left[ \sup_{0 \leq s \leq T} \left| \int_0^s \varphi_u \Gamma^t_u du \right| \right] \leq \mathbb{E} \left[ \int_0^T |\varphi_u|^2 du \right]^{\frac{1}{2}} \mathbb{E} \left[ \int_0^T |\Gamma^t_u|^2 du \right]^{\frac{1}{2}} \quad \text{(by Hölder’s inequality)}
\]

\[
\leq T^{\frac{1}{2}} \mathbb{E} \left[ \int_0^T |\varphi_u|^2 du \right]^{\frac{1}{2}} \mathbb{E} \left[ \sup_{0 \leq s \leq T} |\Gamma^t_s|^2 \right]^{\frac{1}{2}}
\]

< \infty.

Hence \( \Gamma^t_s Y_s + \int_0^s \varphi_u \Gamma^t_u du \) is a uniform-integrable local martingale which implies it is a martingale.

b) Consider \( Y_0 = 0 \) and use a), we get

\[
\Gamma^0_0 Y_0 + \int_0^0 \varphi_s \Gamma^0_s ds = \mathbb{E} \left( \Gamma^0_0 \xi + \int_0^T \varphi_s \Gamma^0_s ds \right | F_0)
\]

\[
0 = \mathbb{E} \left( \Gamma^0_0 \xi + \int_0^T \varphi_s \Gamma^0_s ds \right | F_0)
\]

then the nonnegative variable \( \Gamma^0_\xi + \int_0^T \varphi_s \Gamma^0_s ds \) has zero expectation. Therefore \( \xi = 0 \), \( \mathbb{P} - a.s. \), \( \varphi_s = 0 \), \( d\mathbb{P} \otimes dt - a.s. \) and \( Y = 0 \) a.s.. Particularly, if \( \xi \) and \( \varphi \) are nonnegative, \( Y_s \) is nonnegative.

\( \square \)
1.4. Comparison Theorem

This subsection is devoted to the comparison theorem which allows to compare the solutions of two BSDEs as soon as one can compare the terminal conditions and the generators.

**Theorem 1.7.** [9] Let \((Y, Z)\) and \((\bar{Y}, Z)\) be the solutions of two BSDEs with associated parameters \((f, \xi)\) and \((\bar{f}, \bar{\xi})\). We suppose that

\[
\xi \leq \bar{\xi}, \quad \mathbb{P}\text{-a.s.}
\]

and

\[
f(t, \bar{Y}_t, \bar{Z}_t) - \bar{f}(t, \bar{Y}_t, \bar{Z}_t) \leq 0, \quad dt \otimes d\mathbb{P} \text{-a.s.}
\]

Then we have that \(\mathbb{P}\)-almost surely for any time \(t \in [0, T]\),

\[
Y_t \leq \bar{Y}_t.
\]

Moreover the comparison is strict, i.e., if in addition \(Y_0 = \bar{Y}_0\), then \(\xi = \bar{\xi}, \ f(t, \bar{Y}_t, \bar{Z}_t) = \bar{f}(t, \bar{Y}_t, \bar{Z}_t)\) and \(Y_t = \bar{Y}_t, \ \forall t \in [0, T] \quad \mathbb{P} \text{-a.s.}
\]

In particular, whenever \(\mathbb{P}(\xi < \bar{\xi}) > 0\) or \(f(t, \bar{Y}_t, \bar{Z}_t) - \bar{f}(t, \bar{Y}_t, \bar{Z}_t) < 0\) on a set of positive \(dt \otimes d\mathbb{P}\)-measure, then \(Y_0 < \bar{Y}_0\).

**Proof.** We define:

\[
\begin{align*}
\varphi_t &= f(t, \bar{Y}_t, \bar{Z}_t) - \bar{f}(t, \bar{Y}_t, \bar{Z}_t) \\
Y' &= Y - \bar{Y} \\
Z' &= Z - \bar{Z} \\
\xi' &= \xi - \bar{\xi}.
\end{align*}
\]

For all \(t \geq 0\),

\[
Y'_t = \xi' + \int_t^T (f(t, \bar{Y}_t, \bar{Z}_t) - \bar{f}(t, \bar{Y}_t, \bar{Z}_t))ds - \int_t^T Z'_sdW_s
\]

We rewrite \(f\) as follows

\[
f(t, Y_t, Z_t) - \bar{f}(t, \bar{Y}_t, \bar{Z}_t) = f(t, Y_t, Z_t) - f(t, \bar{Y}_t, Z_t) + f(t, \bar{Y}_t, Z_t) - \bar{f}(t, \bar{Y}_t, \bar{Z}_t)
\]

\[
+ f(t, \bar{Y}_t, \bar{Z}_t) - \bar{f}(t, \bar{Y}_t, \bar{Z}_t).
\]

Now we introduce two processes \(\beta\) and \(\mu\) as follows

\[
\beta_t = \begin{cases} 
(Y_t - \bar{Y}_t)^{-1}(f(t, Y_t, Z_t) - f(t, \bar{Y}_t, Z_t)), & \text{if } Y_t \neq \bar{Y}_t, \\
0, & \text{otherwise},
\end{cases}
\]

\[
\mu_t = \begin{cases} 
(Z_t - \bar{Z}_t)^{-1}(f(t, Y_t, Z_t) - f(t, \bar{Y}_t, \bar{Z}_t)), & \text{if } Z_t \neq \bar{Z}_t, \\
0, & \text{otherwise}.
\end{cases}
\]
By writing $Y'_t$ in terms of $\beta$ and $\mu$, we get

$$Y'_t = \xi' + \int_t^T (Y'_s/\beta_s + Z'_s/\mu_s + \varphi_s)ds - \int_t^T Z'_s dW_s.$$ 

The couple $(Y'_t, Z'_t)$ is a solution of a linear BSDE. Note that $(\beta, \mu)$ satisfy the assumptions of Proposition 1.6. Indeed $\beta, \mu$ are bounded as $f$ is Lipschitz and the integrability condition of the process $\varphi$ can be deduced easily from the standards hypotheses and from the properties of $Y$ and $Z$. From Proposition 1.6 we get

$$Y_t = E\left[\xi' + \int_t^T \Gamma'_s dF_s|\mathcal{F}_t\right],$$

where $(\Gamma'_s)_{s\geq t}$ is the adjoint process defined by the following forward linear SDE

$$d\Gamma'_s = \Gamma'_s (\beta_s ds + \mu_s dW_s).$$

Hence $\xi' \leq 0$ and $\varphi_s \leq 0$, then $Y'_t \leq 0$ follows.

\section*{1.5. Applications of BSDEs in finance}

\subsection*{1.5.1. Stochastic Optimal Control Problem}

Stochastic optimal control is used to solve optimization problems in random systems (which can be controlled) that evolve over time and whose evolution can be influenced by external forces.

Optimal control theory appears in the 1950's, when engineers became interested in problems where the goal was to maximize the returns and minimize the costs of the operation, i.e. in aerospace problems a small improvement of optimal control had a large impact in reducing the costs. The solution of a stochastic optimal control problem can be found by solving a system of differential equations.

In this example, based on [18], we will focus on the following dynamic stochastic differential equation given by the controlled process:

$$\begin{align*}
    dX_t &= (aX_t + bu_t) dt + dW_t \quad t \in [0, T] \\
    X_0 &= x
\end{align*}$$

(1.24)

where $W$ is a Brownian motion, $X_t$, with $t \geq 0$, is called the state process taking values in $(S, \mathcal{B}(S))$, where $S$ is a Polish space, that is, a closed and bounded set of $\mathbb{R}$.

It is assumed that the system can be controlled, i.e. the system is equipped with controllers whose position dictates its future evolution. These controllers are characterized by points $u = (u^1, ..., u^n) \in \mathbb{R}^n$, the control variable.

The values that can be assumed by the control variables are restricted to a certain control region $\mathcal{U} \in \mathbb{R}^n$, which is bounded.

Every function $u_t$ defined on some interval $0 \leq t \leq T$ is called a control.
Definition 1.8. A continuous control $u_t$, defined on some time interval $0 \leq t \leq T$, with range in the control region $U$,

$$u_t \in U \ \forall t \in [0, T]$$

is said to be an admissible control.

We consider processes $(X, u)$ to be $(\mathcal{F}_t)_{t \geq 0}$-adapted and square-integrable. It is also necessary to specify a cost function for evaluating the performance of a system quantitatively. We define the cost function as follows

$$J : U[0, T] \to \mathbb{R}$$

$$J(u) = \frac{1}{2} \mathbb{E} \left[ \int_0^T (|X_t|^2 + |u_t|^2) dt + |X_T|^2 \right].$$

(1.25)

The optimal control problem is to minimize the value of the cost function (1.25) subject to the equation (1.24). There exists a control $u \in U$ that minimizes (1.25) and is unique a.s.

We suppose that $u$ is the optimal control and $X$ is the corresponding process state. For any admissible controls $v \in U$, we have

$$0 \leq \frac{J(u + \varepsilon v) - J(u)}{\varepsilon}$$

$$= \frac{1}{2} \mathbb{E} \left\{ \int_0^T (|\bar{X}_t|^2 + |u_t + \varepsilon v_t|^2) dt + |\bar{X}_T|^2 \right\} - \frac{1}{2} \mathbb{E} \left\{ \int_0^T (|X_t|^2 + |u_t|^2) dt + |X_T|^2 \right\},$$

(1.26)

where $\bar{X}$ has the following dynamic:

$$\begin{cases}
  d\bar{X}_t = a\bar{X}_t dt + b(u_t + \varepsilon v_t) dt + dW_t, & t \in [0, T], \\
  \bar{X}_0 = x.
\end{cases}$$

(1.27)

From (1.26) it follows:

$$0 \leq \frac{1}{2} \mathbb{E} \left\{ \int_0^T (|\bar{X}_t|^2 - |X_t|^2 + 2u_t \varepsilon v_t + \varepsilon^2 v_t^2) dt + |\bar{X}_T|^2 - |X_T|^2 \right\}$$

$$= \frac{1}{2} \mathbb{E} \left\{ \int_0^T \left( \frac{(\bar{X}_t + X_t)(\bar{X}_t - X_t)}{\varepsilon} + 2u_t v_t + \varepsilon v_t^2 \right) dt + (\bar{X}_T + X_T) \left( \frac{X_T - X_T}{\varepsilon} \right) \right\},$$

(1.28)

(1.29)

We define $\xi = \lim_{\varepsilon \to 0} \frac{(X_t - X_t)}{\varepsilon}$. From (1.24) and (1.27), it follows that $\xi$ satisfies:

$$\begin{cases}
  d\xi_t = (a\xi_t + b\xi_t) dt, & t \in [0, T], \\
  \xi_0 = 0.
\end{cases}$$

(1.30)
Letting $\varepsilon \to 0$, equation (1.29) becomes:

$$0 \leq \mathbb{E} \left\{ \int_0^T (X_t \xi_t + u_t v_t) dt + X_T \xi_T \right\}. \quad (1.31)$$

We introduce the following BSDE:

$$\begin{cases}
    dY_t = -(aY_t + X_t)dt + Z_t dW_t, & t \in [0,T], \\
    Y_T = X_T.
\end{cases} \quad (1.32)$$

We suppose that the BSDE (1.32) admits a unique adapted solution $(Y, Z)$. By applying Itô's formula to $Y_T \xi_t$, we get:

$$\mathbb{E}(X_T \xi_T) = \mathbb{E}(Y_T \xi_T)$$

$$= \mathbb{E} \int_0^T \{(-aY_t - X_t)\xi_t + Y_t(a\xi_t + b v_t)\} dt$$

$$= \mathbb{E} \int_0^T \{-X_t \xi_t + Y_t b v_t\} dt.$$

By (1.31) we get,

$$0 \leq \mathbb{E} \int_0^T (bY_t + u_t)v_t dt$$

As $v$ is arbitrary, we get the following

$$u_t = -bY_t \quad a.s. \quad \forall t \in [0,T]. \quad (1.33)$$

Now, since $Y$ is part of the solution of the BSDE (1.32) and is $\mathcal{F}_t$-adapted, so that $u$ is an admissible control. By substituting (1.33) in (1.24) we arrive at the following optimal system:

$$\begin{cases}
    dX_t = (aX_t - b^2 Y_t) dt + dW_t, & t \in [0,T], \\
    dY_t = -(aY_t + bX_t)dt + Z_t dW_t, & t \in [0,T], \\
    X_0 = x, \\
    Y_T = X_T.
\end{cases}$$

Finally, we have a system of one Forward Stochastic Differential Equation (which are introduced later in the thesis) as well as a Backward Stochastic Differential Equation, such a system is known as a Forward Backward Stochastic Differential Equation (FBSDE). If we can solve the FBSDE and find a unique adapted solution $(X, Y, Z)$ then $u$ in (1.32) is the optimal control that solves the original problem.

1.5.2. Hedging in the Black-Scholes model

We give an example for the application of BSDEs in risk hedging. In a financial market we have the following: for $t \in [0,T]$

Finally, we have a system of one Forward Stochastic Differential Equation (which are introduced later in the thesis) as well as a Backward Stochastic Differential Equation, such a system is known as a Forward Backward Stochastic Differential Equation (FBSDE). If we can solve the FBSDE and find a unique adapted solution $(X, Y, Z)$ then $u$ in (1.32) is the optimal control that solves the original problem.

1.5.2. Hedging in the Black-Scholes model

We give an example for the application of BSDEs in risk hedging. In a financial market we have the following: for $t \in [0,T]$
1. There exists a risk-free asset, whose price is modelled by:
\[
\begin{align*}
\frac{dB_t}{B_t} &= rdt, \\
B_0 &= y \in \mathbb{R},
\end{align*}
\]
where \( r \) denotes the interest rate.

2. We model a stock with a risky asset by the following forward SDE:
\[
\begin{align*}
\frac{dS_t}{S_t} &= \mu dt + \sigma dW_t, \\
S_0 &= x \in \mathbb{R},
\end{align*}
\]
where \( \mu \in \mathbb{R}, \sigma > 0 \), \( \mu \) is called drift and \( \sigma \) is the volatility.

The solutions of equations (1.34) and (1.35) are given by
\[
\begin{align*}
B_t &= ye^{rt}, \\
S_t &= x \exp \left( \left( \mu - \frac{1}{2} \sigma^2 \right) t + \sigma W_t \right), \quad t \geq 0.
\end{align*}
\]

A portfolio is a couple of adapted processes \((a_t, b_t)\) that represents the amount of investment in both assets at time \( t \). The wealth process is given by
\[
Y_t = a_t S_t + b_t B_t, \quad t \in [0, T].
\]

A main assumption is that \( Y \) is self-financing, which means that the wealth process satisfies the following
\[
dY_t = a_t dS_t + b_t dB_t = a_t S_t (\mu dt + \sigma dW_t) + rb_t dt.
\]

As \( b_t B_t = Y_t - a_t S_t \), then we have
\[
dY_t = (rY_t + a_t S_t (\mu - r)) dt + a_t S_t \sigma dW_t
\]
Now by putting \( Z_t = a_t S_t \sigma \), we get
\[
dY_t = rY_t dt + Z_t \frac{\mu - r}{\sigma} dt + Z_t dW_t.
\]

One interesting question in the financial market is: At which price \( v \) should one sell the European call option?

The answer is by replicating the portfolio: the seller sells the option at the price \( v \) and invests this sum in the market. The value of his portfolio is characterized by the SDE:
\[
\begin{align*}
\frac{dY_t}{Y_t} &= rY_t dt + \frac{\mu - r}{\sigma} Z_t dt + Z_t dW_t, \\
Y_0 &= v.
\end{align*}
\]
The problem is to find $v$ and $Z$ such that the solution of the previous SDE verifies $Y_T = \xi = (S_T - K)^+$,

$$
\begin{align*}
\begin{cases}
    dY_t &= rY_t dt + \frac{\mu - r}{\sigma} Z_t dt + Z_t dW_t \\
    Y_T &= \xi.
\end{cases}
\end{align*}
$$

In this case $v$ is called the fair price of the option and it suffices to sell the option at the price $v = V_0$. We see that pricing the option is to solve a linear BSDE.
2. FBSDEs and Feynman-Kac formula

In this chapter we present the Forward-Backward Stochastic Differential Equations (FBSDE). FBSDEs provide an intensively studied modelling tool for stochastic control problems as well as for financial mathematics. They appeared for the first time in Bismut’s paper [2] and then studied in a general way by Pardoux and Peng in [19].

At the end of this chapter the Feynman-Kac representation theorem is introduced which gives us a method to connect PDEs and stochastic processes. We can solve PDEs by making simulations for the stochastic process.

The books [4], [20] and the article [9] are used as the main references for this chapter.

2.1. Forward-Backward Stochastic Differential Equations

We consider the solution of certain BSDEs associated with some forward classical stochastic differential equations. We are going to suppose that the randomness of the parameters \((f, \xi)\) of the BSDE comes from the state of the forward equation.

We consider the following Forward Stochastic Differential Equation:

\[
\begin{aligned}
\mathrm{d}X_s &= b(s, X_s)\mathrm{d}s + \sigma(s, X_s)\mathrm{d}W_s \quad s \in [t, T], \\
X_t &= x,
\end{aligned}
\]  

(2.1)

where the coefficients

\[
\begin{align*}
b & : [0, T] \times \mathbb{R} \to \mathbb{R} , \\
\sigma & : [0, T] \times \mathbb{R} \to \mathbb{R},
\end{align*}
\]

satisfy the following assumptions with \(C > 0\)

\[
\begin{align*}
|\sigma(t, x) - \sigma(t, x')| + |b(t, x) - b(t, x')| & \leq C |x - x'|, \quad \text{H1} \\
|\sigma(t, x)| + |b(t, x)| & \leq C(1 + |x|). \quad \text{H2}
\end{align*}
\]

Under the assumptions H1 and H2, there exists a unique solution for (2.1) (see Theorem A.12) which we denote as \(X_s^{t,x}\), for \(s \in [0, T]\).

Now we consider the associated BSDE,

\[
\begin{aligned}
-\mathrm{d}Y_s &= f(s, X_s^{t,x}, Y_s, Z_s)\mathrm{d}s - Z_s\mathrm{d}W_s, \quad s \in [0, T], \\
Y_T &= g(X_T^{t,x}),
\end{aligned}
\]  

(2.2)

where the coefficients

\[
\begin{align*}
f & : [0, T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \to \mathbb{R} \\
g & : \mathbb{R} \to \mathbb{R}
\end{align*}
\]
satisfy the following assumptions with \( C > 0 \) such that
\[
|f(t, x, y, z) - f(t, x', y', z')| \leq C(|y - y'| + |z - z'| + |x - x'|) \quad \text{H3}
\]
\[
|g(x)| \leq C(1 + |x|). \quad \text{H4}
\]
From Theorem 1.5 we know that there exists a unique solution for (2.2) which we denote as \((Y_{s}^{t,x}, Z_{s}^{t,x})\) with \( s \in [0, T] \). Notice that we needed to introduce a new assumption H4 as the function \( g \) in the BSDE (2.2) is linked to the SDE (2.1).

We rewrite (2.1) and (2.2) in one system as follows
\[
\begin{aligned}
X_{s}^{t,x} &= x + \int_{t}^{s} b(r, X_{r}^{t,x})dr + \int_{t}^{s} \sigma(r, X_{r}^{t,x})dW_{r} \quad s \geq t \\
Y_{s}^{t,x} &= g(X_{T}^{t,x}) + \int_{s}^{T} f(r, X_{r}^{t,x}, Y_{r}^{t,x}, Z_{r}^{t,x})dr - \int_{s}^{T} Z_{r}^{t,x}dW_{r} \quad s \geq t
\end{aligned}
\tag{2.3}
\]

The system (2.3) is a combination of a forward SDE and a BSDE, where the terminal condition of the BSDE is now allowed to depend on the process \( X_{s}^{t,x} \). Such a type of equation is called a Forward-Backward Stochastic Differential Equation (FBSDE). \( X_{s}^{t,x} \) represents the forward component and \((Y_{s}^{t,x}, Z_{s}^{t,x})\) the backward component.

It is important to notice that the solution of the backward component does not appear in the coefficients of the forward SDE. We call the system (2.3) a decoupled FBSDE.

A more general FBSDE has the following form
\[
\begin{aligned}
X_{t} &= X_{0} + \int_{0}^{t} b(s, X_{s}, Y_{s}, Z_{s})ds + \int_{0}^{t} \sigma(s, X_{s}, Y_{s})dW_{s}, \\
Y_{t} &= g(X_{T}) + \int_{t}^{T} f(s, X_{s}, Y_{s}, Z_{s})ds - \int_{t}^{T} Z_{s}dW_{s}.
\end{aligned}
\]

We observe that the solution of the backward component appears in the coefficients of the forward component, we call this class of FBSDE, coupled FBSDE. This type of FBSDE is used in the application to carbon emissions markets. However the theory of the coupled BSDEs is out of scope for this thesis, for further reading we refer to the book [4].

2.1.1. Regularity Properties of Solutions

**Proposition 2.1.** [9] For each \( t \in [0, T] \) and \( x \in \mathbb{R} \), there exists \( C \geq 0 \) such that
\[
\mathbb{E} \left( \sup_{0 \leq s \leq T} |Y_{s}^{t,x}|^2 \right) + \mathbb{E} \left( \int_{0}^{T} |Z_{s}^{t,x}|^2 ds \right) \leq C(1 + |x|^2).
\]

**Proposition 2.2.** [9] For each \( t, t' \in [0, T] \), \( x, x' \in \mathbb{R} \) and if \( f, g \) are Lipschitz in \( x \), uniformly in \( t \) concerning \( f \) there exists \( C \geq 0 \) such that
\[
\begin{aligned}
&\mathbb{E} \left( \sup_{0 \leq s \leq T} |Y_{s}^{t,x} - Y_{s}^{t',x'}|^2 \right) + \mathbb{E} \left( \int_{0}^{T} |Z_{s}^{t,x} - Z_{s}^{t',x'}|^2 ds \right) \\
&\leq C(1 + |x|^2)(|x - x'|^2 + |t - t'|).
\end{aligned}
\]
2.1.2. Markov Properties of Solutions

In this subsection, we show that the Markov property of the solution of an SDE translates to the solution of the BSDE. This property is mainly used to study the link to PDEs. We introduce the following filtration $\mathcal{F}_s^t$ which denotes the future $\sigma$-algebra of $W$ after $t$, i.e.

$$\mathcal{F}_s^t = \sigma(W_u - W_t, \ t \leq u \leq s).$$

**Proposition 2.3.** [4] If $(t, x) \in [0, T] \times \mathbb{R}$, then $(X_s^{t,x}, Y_s^{t,x})$, where $t \leq s \leq T$ is adapted to the filtration $\mathcal{F}_s^t$. In particular $Y_s^{t,x}$ is deterministic. We can choose a version $Z_s^{t,x}$ where $t \leq s \leq T$ adapted to the filtration $\mathcal{F}_s^t$. Since $Y_s^{t,x}$ is deterministic we can define a function

$$u(t, x) := Y_t^{t,x},$$

for all $(t, x) \in [0, T] \times \mathbb{R}$.

**Proposition 2.4.** [4] The function $u$ verifies, for any $(t, x), (t', x')$

$$|u(t, x) - u(t', x')| \leq C(|x - x'| + |t - t'|^{1/2}(1 + |x|)),$$

for some $c \geq 0$.

By the help of the function $u$ we can establish the Markov property for the BSDEs.

**Proposition 2.5.** [4] Let $t \in [0, T]$ and $x$ be a square integrable random variable. Then $\mathbb{P}$-a.s.,

$$Y_s^{t,x} = u(s, X_s^{t,x}), \quad \forall s \in [t, T].$$

This property is called the Markov property.

2.2. The link with PDEs: Feynman-Kac formula

We introduce a semi-linear parabolic PDE, as follows:

$$\begin{cases}
u'(t, x) + Au(t, x) + f(t, x, u(t, x), Du(t, x)\sigma(t, x)) = 0, & (t, x) \in [0, T] \times \mathbb{R}, \\
u(T, x) = g(x),
\end{cases}$$

where $A$ represents the linear differential operator:

$$Au(t, x) = b(t, x)Du(t, x) + \frac{1}{2} \left(\sigma^2(t, x)D^2u(t, x)\right).$$

For what concerns the notation, when $u$ is a function of $t$ and $x$ we denote $u'$ the partial differential in time, $Du$ the partial differential in space of first order and $D^2u$ the partial differential in space of second order.

**Feynman-Kac**
The following proposition gives us a probabilistic representation formula for the solution of a semi-linear parabolic PDE. Hence by solving the PDE (2.4) we can deduce the solution of the FBSDE (2.3).

**Proposition 2.6.** [9] Let \( u \in C^2([0, T] \times \mathbb{R}) \) and suppose that there exists a constant \( C \) such that, for each \( (t, x) \in [0, T] \times \mathbb{R} \), it holds

\[
|u(s, x)| + |Du(s, x)\sigma(s, x)| \leq C(1 + |x|).
\]

Moreover let \( u \) be the solution of the semilinear parabolic PDE,

\[
\begin{aligned}
\left\{ 
&u'(t, x) + Au(t, x) + f(t, x, u(t, x), Du(t, x)\sigma(t, x)) = 0, \quad (t, x) \in [0, T] \times \mathbb{R}, \\
&u(T, x) = g(x),
\end{aligned}
\]

where \( A \) is defined as in (2.5). Then for any \( (t, x) \in [0, T] \times \mathbb{R} \),

\[ Y^{t,x}_t := u(t, x), \]

where \( \{(Y^{t,x}_s, Z^{t,x}_s), t \leq s \leq T\} \) is the unique solution of the BSDE in (2.3). Moreover \( (Y^{t,x}_s, Z^{t,x}_s) = \left( u(s, X^{t,x}_s), Du(t, X^{t,x}_s)\sigma(t, X^{t,x}_s) \right) \).

**Proof.** We apply Itô’s lemma to \( u(s, X^{t,x}_s) \), we get

\[
du(s, X^{t,x}_s) = u'(s, X^{t,x}_s)ds + Du(s, X^{t,x}_s)dX^{t,x}_s + \frac{1}{2}\sigma^2(s, X^{t,x}_s)D^2u(s, X^{t,x}_s)ds
\]

\[
= \left( u'(s, X^{t,x}_s) + \frac{1}{2}\sigma^2(s, X^{t,x}_s)D^2u(s, X^{t,x}_s) + Du(s, X^{t,x}_s)b(s, X^{t,x}_s) \right) ds
\]

\[
+ Du(s, X^{t,x}_s)\sigma(s, X^{t,x}_s)dW_s
\]

\[
= (u'(s, X^{t,x}_s) + Au(s, X^{t,x}_s)) ds + Du(s, X^{t,x}_s)\sigma(s, X^{t,x}_s)dW_s.
\]

(by definition of \( A (2.5) \))

As \( u \) is a solution of the PDE (2.4), we have \( u' + Au = -f \), we get

\[
du(s, X^{t,x}_s) = -f(s, X^{t,x}_s, u(s, X^{t,x}_s), Du\sigma(s, X^{t,x}_s)) ds + Du\sigma(s, X^{t,x}_s)dW_s,
\]

which proves the statement. \( \square \)

To get the reverse way, i.e. studying the BSDE and being able to deduce the construction of the solution of the PDE we need to define the following viscosity definition. The basic idea of viscosity solution (see for example [11]) is to replace the differential \( Du(x) \) at a point \( x \) where it does not exist (for example because of a kink in \( u \)) with the differential \( D\phi(x) \) of a smooth function \( \phi \) touching the graph of \( u \), from above for the sub solution condition and from below for the super solution one, at the point \( x \).
Define the stopping time, by continuity there exists $0 < \tau$ such that for each $\phi \in C^2([0, T] \times \mathbb{R})$, we have,
\[
\phi'(t, x) + \mathcal{A}\phi(t, x) + f(t, x, \phi(t, x), D\phi(t, x)\sigma(t, x)) \geq 0,
\]
$\forall (t, x) \in [0, T] \times \mathbb{R}$ whenever $u - \phi$ has a local maximum.

Suppose $u \in C([0, T] \times \mathbb{R}, \mathbb{R})$ satisfies $u(T, x) = g(x)$ s.t. $x \in \mathbb{R}$, then $u$ is called viscosity sub solution of the PDE (2.4), if for each $\phi \in C^2([0, T] \times \mathbb{R})$, we have,
\[
\phi'(t, x) + \mathcal{A}\phi(t, x) + f(t, x, \phi(t, x), D\phi(t, x)\sigma(t, x)) \leq 0,
\]
$\forall (t, x) \in [0, T] \times \mathbb{R}$ whenever $u - \phi$ has a local minimum.

Hence $u \in C([0, T] \times \mathbb{R}, \mathbb{R})$ is called a viscosity solution of (2.4) if it is both a viscosity sub- and super solution.

For a further reading on viscosity solution, we refer to the papers [6] and [8].

**Theorem 2.8.** [4] The function $u(t, x) := Y_{t,x}^{t, x}$, as defined in (2.3), continuous on $[0, T] \times \mathbb{R}$, is a viscosity solution of the PDE (2.4).

**Proof.** By construction $u$ is continuous and $u(T, \cdot) = g(\cdot)$ so we only need to prove is the viscosity sub solution property. Let $\phi$ in $C^2$ such that $u - \phi$ has at $(t_0, x_0)$ $(0 < t_0 < T)$ a local maximum equal to 0. We want to show that
\[
\phi'(t_0, x_0) + \mathcal{A}\phi(t_0, x_0) + f(t_0, x_0, u(t_0, x_0), D\phi(t_0, x_0)) \geq 0.
\]

We will proceed by finding a contradiction, so we assume that there exists $\delta > 0$ such that
\[
\phi'(t_0, x_0) + \mathcal{A}\phi(t_0, x_0) + f(t_0, x_0, u(t_0, x_0), D\phi(t_0, x_0)) = -\delta < 0,
\]
The function $u - \phi$ has a local maximum at $(t_0, x_0)$ equal to 0, this means that $\forall (t, x)$ close to $(t_0, x_0)$ we have,
\[
(u - \phi)(t, x) \leq (u - \phi)(t_0, x_0)
\]
$u(t, x) - \phi(t, x) \leq u(t_0, x_0) - \phi(t_0, x_0) = 0$
$u(t, x) \leq \phi(t, x)$.

By continuity there exists $0 < \alpha \leq t - t_0$ such that for all $t_0 \leq t \leq t_0 + \alpha$ and $|x - x_0| \leq \alpha$,
$u(t, x) \leq \phi(t, x)$
\[
\phi'(t, x) + \mathcal{A}\phi(t, x) + f(t, x, \phi(t, x), D\phi(t, x)) \leq -\frac{\delta}{2} < 0.
\]
Define the stopping time,
\[
\tau = \inf\{u \geq t_0, \quad |X_u^{t_0,x_0} - x_0| \geq \alpha\} \wedge (t_0 + \alpha).
\]
Since $X_{t_0,x_0}$ is a continuous process, we have $|X_{r,t_0,x_0} - x_0| \leq \alpha$.

By Itô formula applied to $\phi(r, X_{r,t_0,x_0})$ between $u \land \tau$ and $(t_0 + \alpha) \land \tau = \tau$ for $t_0 \leq u \leq t_0 + \alpha$, we obtain

$$
\phi(u \land \tau, X_{u \land \tau}^{t_0,x_0}) = \phi(\tau, X_{\tau}^{t_0,x_0}) - \int_{u \land \tau}^{\tau} (r + A\phi) (r, X_{r}^{t_0,x_0})dr - \int_{u \land \tau}^{\tau} D\phi \sigma (r, X_{r}^{t_0,x_0}) dW_r
$$

(2.6)

We consider the order two couples of processes.

On the one hand, for $t_0 \leq u \leq t_0 + \alpha$, we define

$$(Y_u', Z_u') := \left( \phi(u \land \tau, X_{u \land \tau}^{t_0,x_0}), \mathbb{1}_{u \leq \tau} D\phi \sigma (r, X_{r}^{t_0,x_0}) \right).$$

By replacing $Y_u'$ and $Z_u'$ in (2.6), we get

$$Y_u' = \phi(\tau, X_{\tau}^{t_0,x_0}) + \int_{u}^{t_0 + \alpha} \mathbb{1}_{\{r \leq \tau\}} (r + A\phi)(r, X_{r}^{t_0,x_0})dr - \int_{u}^{t_0 + \alpha} Z_{r}dW_r, \quad t_0 \leq u \leq t_0 + \alpha.
$$

On the other hand, for $t_0 \leq u \leq t_0 + \alpha$, we define

$$(Y_u, Z_u) = \left( Y_{u \land \tau}^{t_0,x_0}, \mathbb{1}_{u \leq \tau} Z_{u \land \tau}^{t_0,x_0} \right).$$

By replacing $Y_u$ and $Z_u$ in the backward SDE of the system (2.3), we get

$$Y_u = Y_{t_0 + \alpha} + \int_{u}^{t_0 + \alpha} \mathbb{1}_{\{r \leq \tau\}} f(r, X_{r}^{t_0,x_0}, Y_{r}, Z_{r})dr - \int_{u}^{t_0 + \alpha} Z_{r}dW_r, \quad t_0 \leq u \leq t_0 + \alpha.
$$

The Markov property (2.5) implies that $\mathbb{P}$-a.s. for all $t_0 \leq r \leq t_0 + \alpha$, $Y_{r}^{t_0,x_0} = u(r, X_{r}^{t_0,x_0})$, and thus $Y_{t_0 + \alpha} = Y_{\tau}^{t_0,x_0} = u(\tau, X_{\tau}^{t_0,x_0})$ and we get

$$Y_u = u(\tau, X_{\tau}^{t_0,x_0}) + \int_{u}^{t_0 + \alpha} \mathbb{1}_{\{r \leq \tau\}} f(r, X_{r}^{t_0,x_0}, u(r, X_{r}^{t_0,x_0}), Z_{r})dr - \int_{u}^{t_0 + \alpha} Z_{r}dW_r,$

where $t_0 \leq u \leq t_0 + \alpha$.

In the next step we will apply the Comparison Theorem (1.7) to $(Y_u', Z_u')$ and $(Y_u, Z_u)$. From the definition of $\tau$ we get,

$$u(\tau, X_{\tau}^{t_0,x_0}) \leq \phi(\tau, X_{\tau}^{t_0,x_0})$$

Moreover, we always have by definition of $\tau$,

$$
E \left( \int_{t_0}^{t_0 + \alpha} -\mathbb{1}_{\{r \leq \tau\}} (r + A\phi + f)(r, X_{r}^{t_0,x_0}, u(r, X_{r}^{t_0,x_0}), D\phi \sigma (r, X_{r}^{t_0,x_0})) dr \right) 
\geq E(\tau - t_0) \frac{\delta}{2} > 0
$$

This quantity is indeed strictly positif, as $\delta > 0$ and $\tau > t$ since $|X_{t_0,x_0} - x_0| = 0 < \alpha$.

We can apply the strict version of the Comparison Theorem (1.7) to obtain $u(t_0, x_0) = Y_0 < Y_0' = \phi(t_0, x_0)$. This shows the contradiction as we have $u(t_0, x_0) = \phi(t_0, x_0)$. \(\square\)
3. Numerical Methods for BSDEs

Typically the dynamics of stock prices and interest rates are driven by a continuous-time stochastic process. However, simulation is done in discrete time steps. Hence, the first step in any simulation scheme is to find a way to "discretize" a continuous-time process into a discrete time process. We will discretize the forward component by the simple Euler scheme and the backward component by a theta-scheme. The use of a theta-scheme introduces two parameters $\theta_1$ and $\theta_2$, which are used to generate multiple discretizations.

It is the Feynman–Kac theorem, proven in the previous section, that relates the conditional expectation of the value of a contract payoff function under the risk-neutral measure to the solution of a partial differential equation.

The existing numerical methods can be classified into 3 major groups:

- PDE methods
- Monte-Carlo simulation
- Numerical Integration Methods (often rely on Fourier)

Each of them have advantages and disadvantages and the last one is often used for Calibration purposes. We will use probabilistic numerical methods to solve BSDEs which will rely on time discretization of the stochastic process and approximations for the appearing conditional expectations.

In this chapter we will make use of the following defined FBSDE

$$
\begin{align*}
X_t &= X_0 + \int_0^t b(s, X_s)ds + \int_0^t \sigma(s, X_s)dW_s, \quad X_0 = x_0, \quad (3.1) \\
Y_t &= \xi + \int_t^T f(s, X_s, Y_s, Z_s)ds - \int_t^T Z_s dW_s, \quad \xi = g(X_T). \quad (3.2)
\end{align*}
$$

The equation (3.1) represents a forward SDE and (3.2) represents a BSDE. As seen in the previous chapter we know that this FBSDE satisfies the Markov property.

Let $\Pi$ be a partition of time points $0 = t_0 < t_1 < t_2 < t_3 < \ldots < t_i < \ldots < t_M = T$, $\Delta t := t_{i+1} - t_i$, be fixed time steps and $\Delta W_{ti} := W_{ti+1} - W_{ti}$, where $W_t$ is a Wiener process. Notice that the increments $\Delta W_{ti} \sim N(0, \sqrt{\Delta t})$.

For notational simplicity we use: $X_i = X_{t_i}$, $Y_i = Y_{t_i}$, $Z_i = Z_{t_i}$. To determine the values $X_{i+1}^{\Pi}$ for $i = 0, \ldots, M - 1$, we use three different Taylor schemes: Euler, Milstein and 2.0 weak Taylor.
3.1. Discretization of the SDE

In this section, based on the paper [29], we will approximate the numerical solution of the FSDE (3.1). Just as in the deterministic case, we can derive numerical schemes by looking at the Taylor expansions of certain functions. In the thesis we consider the stochastic setting. First we introduce the following definitions of strong and weak convergence.

**Definition 3.1.** Let $X$ be the solution of the FSDE (3.1) and $X_i^\Pi$ be the time-discretized approximation of $X$. A time-discretized approximation $X_i^\Pi$ converges to the stochastic process $X$ in the strong sense with order $\alpha_1$, if there exists a constant $C \in \mathbb{R}$ such that:

$$
\mathbb{E}[|X_i^\Pi - X_i|] \leq C \Delta t^{\alpha_1}.
$$

**Definition 3.2.** Let $X$ be the solution of the FSDE (3.1) and $X_i^\Pi$ be the time-discretized approximation of $X$. A time-discretized approximation $X_i^\Pi$ converges to the stochastic process $X$ in the weak sense, with order $\alpha_2$ if there exists a constant $C \in \mathbb{R}$ such that for every infinitely often differentiable function $\phi : \mathbb{R}^d \to \mathbb{R}$ with at most polynomially growing derivatives it holds that:

$$
|\mathbb{E}[\phi(X_i^\Pi)] - \mathbb{E}[\phi(X_i)]| \leq C \Delta t^{\alpha_2}.
$$

3.1.1. Euler-Maruyama scheme

Let $k : \mathbb{R} \to \mathbb{R}$ be a function that is twice differentiable and an Itô process $X_t$ with drift term $b(t, X_t) \in C^2$, and diffusion term $\sigma(t, X_t) \in C^2$.

Consider the forward SDE of (3.1),

$$
X_t = X_0 + \int_{t_0}^{t} b(s, X_s)ds + \int_{t_0}^{t} \sigma(s, X_s)dW_s, \quad \forall t \in [t_0, T].
$$

(3.3)

By Itô’s lemma we get:

$$
k(X_t) = k(X_0) + \int_{t_0}^{t} \left[ b(s, X_s)k'(X_s) + \frac{1}{2} \sigma(s, X_s)^2 k''(X_s) \right] ds + \int_{t_0}^{t} \sigma(s, X_s)k'(X_s)dW_s.
$$

(3.4)

Define:

$$
L^0k := bk' + \frac{1}{2} \sigma^2 k''
$$

$$
L^1k := \sigma k'.
$$

We can rewrite (3.4) as :

$$
k(X_t) = k(X_0) + \int_{t_0}^{t} L^0k(X_s)ds + \int_{t_0}^{t} L^1k(X_s)dW_s.
$$

(3.5)
Consider now the integral form of $X_t$ in (3.3) and substitute the functions $b$ and $\sigma$ for $k$ in the equation above (3.5), we get the so called Itô-Taylor expansion of first order of $X_t$:

$$X_t = X_{t_0} + \int_{t_0}^{t} \left( b(t, X_{t_0}) + \int_{t_0}^{s} L^0 b(t, X_{t_0}) dp + \int_{t_0}^{s} L^1 b(t, X_{t_0}) dW_p \right) ds$$

$$+ \int_{t_0}^{t} \left( \sigma(t, X_{t_0}) + \int_{t_0}^{s} L^0 \sigma(t, X_{t_0}) dp + \int_{t_0}^{s} L^1 \sigma(t, X_{t_0}) dW_p \right) dW_s$$

$$= X_{t_0} + b(t, X_{t_0}) \int_{t_0}^{t} ds + \sigma(t, X_{t_0}) \int_{t_0}^{t} dW_s + R_t$$

$$= X_{t_0} + b(t, X_{t_0})(t - t_0) + \sigma(t, X_{t_0})(W_t - W_{t_0}) + R_t,$$

where $R_t$ is the remaining term consisting of double integrals. This Itô-Taylor expansion gives us the following Euler-Maruyama scheme as an approximation for the FSDE in (3.3):

$$\begin{aligned}
X_{i+1}^\Pi &= X_i^\Pi + b(t_i, X_i^\Pi) \Delta t + \sigma(t_i, X_i^\Pi)\Delta W_i & \forall i = 0, ..., M - 1 \\
X_0^\Pi &= x_0.
\end{aligned}$$

**Proposition 3.3.** [30] Let $X$ be given by (3.3) and assume conditions H1 and H2 (from Chapter 2) such that there exists a unique solution. Then the Euler-Maruyama scheme (3.6) has order of strong convergence $\alpha_1 = \frac{1}{2}$ and order of weak convergence $\alpha_2 = 1$.

### 3.1.2. The Milstein Scheme

The Milstein approximation for the SDE in (3.3) has the form:

$$\begin{aligned}
X_{i+1}^\Pi &= X_i^\Pi + b(t_i, X_i^\Pi) \Delta t + \sigma(t_i, X_i^\Pi)\Delta W_i + \frac{1}{2} \sigma(t_i, X_i^\Pi) D\sigma(t_i, X_i^\Pi)(\Delta W_i^2 - \Delta t),
\end{aligned}$$

for $i = 0, ..., M - 1$

**Proposition 3.4.** [30] Let $X$ be given by (3.3) and assume conditions H1 and H2 such that there exists a unique solution. Then the Milstein scheme (3.7) has order of strong convergence $\alpha_1 = 1$ and order of weak convergence $\alpha_2 = 1$.

### 3.1.3. The weak Taylor Scheme of order 2.0

The weak Taylor scheme of order 2.0 for the SDE in (3.3), given $X_i^\Pi = x$, has the form:

$$\begin{aligned}
X_{i+1}^\Pi &= x_0 + b(t_i, x) \Delta t + \sigma(t_i, x) \Delta W_i + \frac{1}{2} \sigma(t_i, x) D\sigma(t_i, x)((\Delta W_i)^2 - \Delta t) \\
&\quad + D\sigma(t_i, x) \Delta Z_{i+1} + \frac{1}{2} \left( b(t_i, x) D\sigma(t_i, x) + \frac{1}{2} D^2 b(t_i, x) \sigma^2(x) \right) (\Delta t)^2 \\
&\quad + (b(t_i, x) D\sigma(t_i, x) + \frac{1}{2} D^2 \sigma(t_i, x) \sigma^2(x)) (\Delta W_i \Delta t - \Delta Z_{i+1}),
\end{aligned}$$

for $i = 0, ..., M - 1$ and where $\Delta Z_{i+1} = \frac{1}{2}(\Delta W_i \Delta t + \zeta_{i+1}(\Delta t)^{3/2})$, $\zeta_{i+1} \sim \mathcal{N}(0, 1/3)$.
**Proposition 3.5.** [30] Let $X$ be given by (3.3) and assume conditions H1 and H2. Then the weak Taylor scheme of order 2.0 (3.7) has order of strong convergence $\alpha_1 = 1$ and order of weak convergence $\alpha_2 = 2$.

We observe the following,

$$\mathbb{E}[\Delta Z_{i+1}] = 0, \quad \text{Var}(\Delta Z_{i+1}) = \frac{1}{3}(\Delta t)^3 \quad \text{and} \quad \text{Cov}(\Delta W_i, \Delta Z_{i+1}) = \frac{1}{2}(\Delta t)^2. \quad (3.7)$$

If we replace $\Delta Z_{i+1}$ by $\Delta Z_{i+1} = \frac{1}{2}\Delta W_i\Delta t$ as suggested in [27], then

$$\mathbb{E}[\Delta Z_{i+1}] = 0, \quad \text{Var}(\Delta Z_{i+1}) = \frac{1}{4}(\Delta t)^3 \quad \text{and} \quad \text{Cov}(\Delta W_i, \Delta Z_{i+1}) = \frac{1}{2}(\Delta t)^2. \quad (3.8)$$

This replacement has the same moments in first order and simplifies the scheme.

**Remark 3.6.** Similarly as in [29], the Taylor discretization scheme can be written in a general form, given $X_i^\Pi = x$, as follows:

$$\begin{aligned}
X_{i+1}^\Pi &= x + m(t_i, x)\Delta t + s(t_i, x)\Delta W_i + t(t_i, x)(\Delta W_i)^2, \quad \text{for } i = 0, \ldots, M - 1 \\
X_0^\Pi &= x_0.
\end{aligned} \quad (3.9)$$

- For the Euler scheme, we find:
  $$\begin{aligned}
m(t_i, x) &= b(t_i, x), \\
s(t_i, x) &= \sigma(t_i, x), \\
t(t_i, x) &= 0.
\end{aligned}$$

- For the Milstein scheme, we have:
  $$\begin{aligned}
m(t_i, x) &= b(t_i, x) - \frac{1}{2}\sigma(t_i, x)D\sigma(t_i, x), \\
s(t_i, x) &= \sigma(t_i, x), \\
t(t_i, x) &= \frac{1}{2}\sigma(t_i, x)D\sigma(t_i, x).
\end{aligned}$$

- For the 2.0-weak-Taylor scheme, we see that:
  $$\begin{aligned}
m(t_i, x) &= b(t_i, x) - \frac{1}{2}\sigma(t_i, x)D\sigma(t_i, x) + \frac{1}{2}(b(t_i, x)Db(t_i, x) + \frac{1}{2}D^2\sigma(t_i, x)\sigma(t_i, x))\Delta t, \\
s(t_i, x) &= \sigma(t_i, x) + \frac{1}{2}(D\mu(t_i, x)\sigma(t_i, x) + \mu(t_i, x)D\mu(t_i, x)) + \frac{1}{2}D^2\mu(t_i, x)\sigma^2(t_i, x)\Delta t, \\
t(t_i, x) &= \frac{1}{2}\sigma(t_i, x)D\sigma(t_i, x).
\end{aligned}$$

For the discretization schemes above we can determine a characteristic function of $X_{i+1}^\Pi$ (3.9), which is given in Lemma 3.7. For notational simplicity, in Lemma 3.7, we will use that the variables only depend on the space term, i.e. $m(t_i, x) = m(x), s(t_i, x) = s(x), t(t_i, x) = t(x).$
Lemma 3.7. [30] The characteristic function of $X_{i+1}^\Pi$ (3.9), given $X_i^\Pi = x$, is given by

$$
\phi_{X_{i+1}^\Pi}(u) = \mathbb{E}[\exp(iuX_{i+1}^\Pi)|X_i^\Pi = x]
= \exp \left( iux + ium(x)\Delta t - \frac{1}{2}u^2s^2(x)\Delta t \right) \frac{(1 - 2iut(x)\Delta t)^{-\frac{1}{2}}}{1 - 2iut(x)\Delta t}.
$$

Proof. First we show the result when $t(x) = 0$,

$$
\phi_{X_{i+1}^\Pi}(u) = \mathbb{E}\left[ \exp(iuX_{i+1}^\Pi)|X_i^\Pi = x\right]
= \mathbb{E}\left[ \exp(iux + ium(x)\Delta t + ius(x)\Delta W_i)|X_i^\Pi = x\right]
= \exp \left( iux + ium(x)\Delta t - \frac{1}{2}u^2s^2(x)\Delta t \right).
$$

For $t(x) \neq 0$ we find that

$$
\phi_{X_{i+1}^\Pi}(u) = \mathbb{E}\left[ \exp(iuX_{i+1}^\Pi)|X_i^\Pi = x\right]
= \mathbb{E}\left[ \exp(iux + ium(x)\Delta t + ius(x)\Delta W_i + iut(x)(\Delta W_i)^2)\right]
= \mathbb{E}\left[ \exp(iux + ium(x)\Delta t + iut(x)\left(\Delta W_i + \frac{1}{4}t(x)\right)^2 - iu\frac{1}{4}t(x)\right] \mathbb{E}\left[ \exp(t(x)iut(\Delta W_i + \frac{1}{4}t(x))^2)\right]
= \exp \left( iux + ium(x)\Delta t - \frac{1}{4}u^2t(x)\Delta t \right)
$$

where $\Delta W_i + \frac{1}{4}t(x)^2 \sim \mathcal{N}(\frac{1}{4}t(x)^2, \Delta t)$ by (A.11) this is equivalent to

$$
\Delta W_i \sim \mathcal{N}(0, \Delta t),
$$

hence,

$$
\phi_{X_{i+1}^\Pi}(u) = \exp \left( iux + ium(x)\Delta t + iut(\Delta W_i)\right) \phi_{\chi^2_1}(\frac{1}{4}t(x)^2, \Delta t)
= \exp \left( iux + ium(x)\Delta t + iut(\Delta W_i)\right) \frac{1}{\sqrt{2\pi t(x)\Delta t}} \left(1 - 2iut(x)\Delta t\right)^{-\frac{1}{2}}
= \exp \left( iux + ium(x)\Delta t - \frac{1}{2}u^2s^2(x)\Delta t \right) \left(1 - 2iut(x)\Delta t\right)^{-\frac{1}{2}}
$$

and the statement is proved. \qed
3.2. Discretization of the BSDE

In this section we will develop a discretization scheme for the BSDE (3.1). First, the terminal condition $Y_M$ will be approximated by using the Euler scheme (3.6) for $X_t$:

$$Y_M = g(X^M_T).$$

From the Feynman-Kac formula, i.e. Proposition 2.6, we know that the solution of the FBSDE (3.1) $(Y_t, Z_t) = (u(t, X_t), Du(t, X_t))$. At time $T$ we know that $u(t_M, X_M) = g(X_M)$, so we have $Z_M = \sigma(t_M, X_M)Dg(X_M)$. Hence the terminal condition for $Z_t$ is obtained by substituting for $X_T$ its Euler scheme:

$$Z_M = \sigma(t_M, X^\Pi_M)Dg(X^\Pi_M).$$

To develop a numerical scheme for the BSDE, we focus on the discrete version of the BSDE on the interval $[t_i, t_{i+1}]$:

$$Y_i = Y_{i+1} + \int_{t_i}^{t_{i+1}} f(s, X_s, Y_s, Z_s)ds - \int_{t_i}^{t_{i+1}} Z_s dW_s. \quad (3.10)$$

By a basic Euler discretization, backward in time, the unknown value $Y_{i+1}$ is required to approximate $Y_i$. However, by taking the adaptability constraints on $Y$ and $Z$ and taking the conditional expectation on both sides with respect to the filtration $\mathcal{F}_{t_i}$ into account, it is possible to get a backward induction scheme. Hence, as $Y_i$ and $Z_i$ are adapted to $\mathcal{F}_{t_i}$, we have

$$E[Z_i|\mathcal{F}_{t_i}] = Z_i,$$
$$E[Y_i|\mathcal{F}_{t_i}] = Y_i.$$

We consider (3.10) with the conditional expectation w.r.t. the filtration $\mathcal{F}_{t_i}$, we get,

$$Y_i = E[Y_{i+1}|\mathcal{F}_{t_i}] + E\left[\int_{t_i}^{t_{i+1}} f(s, X_s, Y_s, Z_s)ds | \mathcal{F}_{t_i}\right] - E\left[\int_{t_i}^{t_{i+1}} Z_s dW_s | \mathcal{F}_{t_i}\right]$$
$$= E[Y_{i+1}|\mathcal{F}_{t_i}] + \int_{t_i}^{t_{i+1}} E[f(s, X_s, Y_s, Z_s)|\mathcal{F}_{t_i}]ds - E\left[\int_{t_i}^{t_{i+1}} Z_s dW_s | \mathcal{F}_{t_i}\right]$$
(by Fubini’s theorem)

Since we assumed that $Z_t \in \mathbb{H}_T^2$, we get that its integral form is a martingale. Hence we get,

$$Y_i = E[Y_{i+1}|\mathcal{F}_{t_i}] + \int_{t_i}^{t_{i+1}} E[f(s, X_s, Y_s, Z_s)|\mathcal{F}_{t_i}]ds. \quad (3.11)$$
Now we will use the theta-time discretization method [13]: a convex combination of an explicit term, i.e. at time \( t_i+1 \) and an implicit term, i.e. at time \( t_i \) to approximate the integral. The equation (3.11) becomes:

\[
Y_i \approx E[Y_{i+1}|\mathcal{F}_{t_i}] + \Delta t \theta_i E[f(t_i, X_i, Y_i, Z_i)|\mathcal{F}_{t_i}] + \Delta t(1 - \theta_1)E[f(t_{i+1}, X_{i+1}Y_{i+1}, Z_{i+1})|\mathcal{F}_{t_i}]
\]

(since \( Y_i, Z_t \) are \( \mathcal{F}_{t_i} \)-measurable)

\[
= E[Y_{i+1}|\mathcal{F}_{t_i}] + \Delta t \theta_i f(t_i, X_i, Y_i, Z_i) + \Delta t(1 - \theta_1)E[f(t_{i+1}, X_{i+1}Y_{i+1}, Z_{i+1})|\mathcal{F}_{t_i}],
\]

with \( \theta_1 \in [0, 1] \).

To get a numerical scheme for the process \( Z_t \), we multiply both sides of equation (3.10) by \( \Delta W_t \) and then take conditional expectations w.r.t the filtration \( \mathcal{F}_{t_i} \). We get the following

\[
0 = E[\Delta W_i Y_{i+1}|\mathcal{F}_{t_i}] + E\left[\int_{t_i}^{t_{i+1}} \Delta W_i f(s, X_s, Y_s, Z_s)ds|\mathcal{F}_{t_i}\right] - E\left[\Delta W_i \int_{t_i}^{t_{i+1}} Z_s dW_s|\mathcal{F}_{t_i}\right]
\]

\[
= E[\Delta W_i Y_{i+1}|\mathcal{F}_{t_i}] + \int_{t_i}^{t_{i+1}} E\left[f(s, X_s, Y_s, Z_s)(W_{i+1} - W_s + W_s - W_i)|\mathcal{F}_{t_i}\right]ds
\]

\[
- E\left[\int_{t_i}^{t_{i+1}} dW_s \int_{t_i}^{t_{i+1}} Z_s dW_s|\mathcal{F}_{t_i}\right]
\]

\[
= E[\Delta W_i Y_{i+1}|\mathcal{F}_{t_i}] + \int_{t_i}^{t_{i+1}} E\left[f(s, X_s, Y_s, Z_s)(W_{i+1} - W_s + W_s - W_i)|\mathcal{F}_{t_i}\right]ds
\]

\[
- \int_{t_i}^{t_{i+1}} E[Z_s|\mathcal{F}_{t_i}]ds
\]

\[
= E[\Delta W_i Y_{i+1}|\mathcal{F}_{t_i}] + \int_{t_i}^{t_{i+1}} E[f(s, X_s, Y_s, Z_s)(W_{i+1} - W_s)|\mathcal{F}_{t_i}]ds
\]

\[
+ \int_{t_i}^{t_{i+1}} E[f(s, X_s, Y_s, Z_s)(W_s - W_i)|\mathcal{F}_{t_i}]ds - \int_{t_i}^{t_{i+1}} E[Z_s|\mathcal{F}_{t_i}]ds
\]

\[
= E[\Delta W_i Y_{i+1}|\mathcal{F}_{t_i}] + \int_{t_i}^{t_{i+1}} E[f(s, X_s, Y_s, Z_s)(W_{i+1} - W_s)|\mathcal{F}_{t_i}]ds
\]

\[
+ \int_{t_i}^{t_{i+1}} E[f(s, X_s, Y_s, Z_s)(W_s - W_i)|\mathcal{F}_{t_i}]ds - \int_{t_i}^{t_{i+1}} E[Z_s|\mathcal{F}_{t_i}]ds
\]

\[
= E[\Delta W_i Y_{i+1}|\mathcal{F}_{t_i}] + \int_{t_i}^{t_{i+1}} E[f(s, X_s, Y_s, Z_s)(W_{i+1} - W_s)|\mathcal{F}_{t_i}]ds
\]

\[
+ \int_{t_i}^{t_{i+1}} E[f(s, X_s, Y_s, Z_s)(W_s - W_i)|\mathcal{F}_{t_i}]ds - \int_{t_i}^{t_{i+1}} E[Z_s|\mathcal{F}_{t_i}]ds,
\]

(3.13)
where we used Fubini’s theorem to switch the order of integration in the first integral. By theta-approximation of both integrals in (3.13), we get:

\[
0 \approx \mathbb{E}[\Delta W_i Y_{i+1} | \mathcal{F}_t] + \Delta t (1 - \theta_2) \mathbb{E}[f(t_{i+1}, X_{i+1}, Y_{i+1}, Z_{i+1}) \Delta W_i | \mathcal{F}_t]
- \Delta t \theta_2 Z_t - \Delta t (1 - \theta_2) \mathbb{E}[Z_{i+1} | \mathcal{F}_t], \quad \theta_2 \in [0, 1],
\]

(3.14)

where we used the Ito-adaptedness property of \( Y_i \) and \( Z_i \) and the fact that \( \Delta W_i | \mathcal{F}_t \sim \mathcal{N}(0, \Delta t) \). The above equations (3.12) and (3.14) lead to the following discrete-time approximation \((Y^\Pi, Z^\Pi)\) for \((Y, Z)\):

\[
\begin{align*}
Y^\Pi_M &= g(X^\Pi_M), \\
Z^\Pi_M &= \sigma Dg(X^\Pi_M), \\
Z^\Pi_i &= -\theta_2^{-1}(1 - \theta_2) \mathbb{E}[Z^\Pi_{i+1} | \mathcal{F}_i] + \frac{1}{\Delta t} \theta_2^{-1} \mathbb{E}[Y^\Pi_{i+1} \Delta W_i | \mathcal{F}_i] \\
& \quad + \Delta t (1 - \theta_2) \mathbb{E}[f(t_{i+1}, X^\Pi_{i+1}, Y^\Pi_{i+1}, Z^\Pi_{i+1}) | \mathcal{F}_i], \\
Y^\Pi_i &= \mathbb{E}[Y^\Pi_{i+1} | \mathcal{F}_i] + \Delta t \theta_1 f(t_i, X^\Pi_i, Y^\Pi_i, Z^\Pi_i) \\
& \quad + \Delta t (1 - \theta_1) \mathbb{E}[f(t_{i+1}, X^\Pi_i, Y^\Pi_i, Z^\Pi_i) | \mathcal{F}_i].
\end{align*}
\]

We observe that \( Y^\Pi_i \) and \( Z^\Pi_i \) depend on the value \( X^\Pi_i \), so when \( X^\Pi_i = x \), then

\[
\begin{align*}
Y^\Pi_M &= g(X^\Pi_M), \\
Z^\Pi_M &= \sigma Dg(X^\Pi_M), \\
Z^\Pi_i &= -\theta_2^{-1}(1 - \theta_2) \mathbb{E}[Z^\Pi_{i+1} | X^\Pi_i = x] \\
& \quad + \frac{1}{\Delta t} \theta_2^{-1} \mathbb{E}[Y^\Pi_{i+1} (X^\Pi_{i+1}) \Delta W_i | X^\Pi_i = x], \\
Y^\Pi_i &= \mathbb{E}[Y^\Pi_{i+1} (X^\Pi_{i+1}) | X^\Pi_i = x] + \Delta t \theta_1 f(t_i, x, Y^\Pi_i(x), Z^\Pi_i(x)) \\
& \quad + \Delta t (1 - \theta_1) \mathbb{E}[f(t_{i+1}, X^\Pi_i, Y^\Pi_i, Z^\Pi_i) | X^\Pi_i = x],
\end{align*}
\]

for \( i = M - 1, ..., 0 \).
4. The Fourier-Cosine method for Backward Stochastic Differential Equations

In this chapter we present an option pricing method for European options with one underlying asset, based on the Fourier-cosine series. This method developed by Fang and Oosterlee [12] is called the one-dimensional COS method and as it is used for solving BSDEs a different name is BCOS method. The key idea of this method is the relation of the characteristic function with the series coefficients of the Fourier-cosine expansion of the discounted expected payoff.

This method covers a range of application in different underlying dynamics, including Lévy processes and the Heston stochastic volatility model, and various types of option contracts. In this thesis we will only deal with European options in particular. This chapter is based on [26] [27]. To get a better understanding, the theory of Fourier series is introduced in the first section which is based on the book [7].

4.1. Fourier Transform & Fourier Cosine series

Definition 4.1. Let $p(x)$ be a piecewise continuous real function over $\mathbb{R}$ which satisfies the integrability condition

$$
\int_{-\infty}^{+\infty} |p(x)|dx < \infty.
$$

The Fourier transform of $p(x)$ is defined by

$$
\hat{p}(\omega) = \int_{-\infty}^{+\infty} e^{i\omega x} p(x)dx, \quad x \in \mathbb{R}.
$$

(4.1)

The inverse Fourier transform of $\hat{p}(x)$ is given by

$$
p(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-iyx} \hat{p}(y)dy, \quad y \in \mathbb{R}.
$$

(4.2)

Definition 4.2. Let $\hat{R}$ be a random variable having probability density function $f(x)$, such that

$$
\int_{-\infty}^{+\infty} f(x)dx = 1.
$$

42
The Fourier transform of $f(x)$ is the characteristic function, denoted by $\phi$ such that,

$$\phi(u) = \mathbb{E}[e^{iu\hat{R}}] = \int_{-\infty}^{+\infty} e^{iux} f(x) dx, \quad u \in \mathbb{R}. \quad (4.3)$$

The probability density function is the inverse Fourier transform of the characteristic function. The characteristic function and probability density function form a Fourier pair.

A Fourier series is an expansion of a periodic function $f(x)$ in terms of an infinite sum of sines and cosines. A function $f(x)$ supported on the domain $x \in [-\pi, \pi]$ can be written as its Fourier series by

$$f(x) = \frac{1}{2} A_0 + \sum_{n=1}^{\infty} A_n \cos(nx) + \sum_{n=1}^{\infty} B_n \sin(nx),$$

$$A_0 := \frac{1}{\pi} \int_{-\pi}^{\pi} f(y) dy,$$

$$A_n := \frac{1}{\pi} \int_{-\pi}^{\pi} f(y) \cos(ny) dy,$$

$$B_n := \frac{1}{\pi} \int_{-\pi}^{\pi} f(y) \sin(ny) dy,$$

For a symmetric function, i.e. $f(x) = f(-x) \ \forall x$, the coefficients $B_n$ will be zero. The COS method is based on the Fourier cosine series.

A function $f(x)$ supported on the domain $x \in [a, b]$ can be written as its Fourier cosine series by

$$f(x) = \frac{1}{2} A_0 + \sum_{k=1}^{\infty} A_k \cos\left(k\pi \frac{x-a}{b-a}\right)$$

$$= \sum_{k=0}^{\infty} A_k \cos\left(k\pi \frac{x-a}{b-a}\right),$$

where $\sum_{k=0}^{\infty} \'$ indicates that the first term in the summation is weighted by $\frac{1}{2}$ and $A_k$ is defined by

$$A_k := \frac{2}{b-a} \int_{a}^{b} f(y) \cos\left(k\pi \frac{y-a}{b-a}\right) dy. \quad (4.4)$$

For functions supported on any other finite interval, say $[a, b] \in \mathbb{R}$, the Fourier-cosine series expansion can easily be obtained via a change of variables. The coefficients $A_k$ can be approximated using the Fourier transform of $f(y)$ as we explain below. Notice that a cosine-function can be written as follows

$$\cos\left(k\pi \frac{y-a}{b-a}\right) = \mathcal{R}\left\{\exp\left(k\pi \frac{y-a}{b-a}\right)\right\}.$$
where \( R\{\cdot\} \) represents the real part of the argument. Putting the above equality into (4.4) gives

\[
A_k = \frac{2}{b-a} \int_a^b f(y) \mathcal{R} \left\{ \exp \left( \frac{k\pi y - a}{b-a} \right) \right\} dy
\]

\[
A_k = \frac{2}{b-a} \mathcal{R} \left\{ \int_a^b f(y) \exp \left( \frac{k\pi y - a}{b-a} \right) dy \right\}.
\]

Suppose the integral over the whole domain is a good approximation of the integral of the interval. The coefficients \( A_k \) can be written as

\[
A_k = \frac{2}{b-a} \mathcal{R} \left\{ \int_a^b f(y) \exp \left( \frac{i k \pi}{b-a} y - \frac{i k \pi a}{b-a} \right) dy \right\}
\]

\[
\approx \frac{2}{b-a} \mathcal{R} \left\{ \int_a^b f(y) \exp \left( \frac{i k \pi}{b-a} y - \frac{i k \pi a}{b-a} \right) dy \right\}
\]

\[
\approx \frac{2}{b-a} \mathcal{R} \left\{ \exp \left( -i k \pi a \right) \int_a^b f(y) \exp \left( i k \pi y - i \frac{k \pi a}{b-a} \right) dy \right\}.
\]

A method for pricing European options with numerical integration techniques is the risk-neutral valuation formula:

\[
v(t, S_t) = e^{-r(T-t)} \mathbb{E}_Q[v(S_T, T)|S_t] = e^{-r(T-t)} \int_\mathbb{R} v(S_T, T)f(S_T|S_t)dS_T, \quad (4.5)
\]

where \( v \) denotes the option value, \( t \) is the initial time \( T \) is the maturity time, \( \mathbb{E}_Q[\cdot] \) is the expectation under the risk-neutral measure \( Q \), \( S_t \) and \( S_T \) are the price variables of the underlying asset at time \( t \) and \( T \), \( f(S_T|S_t) \) is the density function of \( S_T \) given \( S_t \), \( r \) is the risk-neutral interest rate.

We insert in (4.5):

\[
x := S_t, \quad y := S_T, \quad \Delta t := T - t
\]

\[
v(x, t) = e^{-r\Delta t} \mathbb{E}_Q[v(y, T)|x] = e^{-r\Delta t} \int_\mathbb{R} v(y, T)f(y|x)dy. \quad (4.6)
\]

The integral (4.6) will be approximated with the COS method in 5 steps

1. **Truncate the integration part.**
   The density function \( f(y|x) \) in (4.6) decays to zero very fast as \( y \to \pm \infty \). Therefore, \( v(x, t) \) can be approximated by some finite integration range \([a, b] \subset \mathbb{R}\)

\[
v(x, t) \approx v_1(x, t) = e^{-r\Delta t} \int_a^b v(y, T)f(y|x)dy.
\]

2. **Replace the density function by its cosine expansion.**
   The density function is usually not known, so we replace the density function by its cosine expansion in \( y \).

\[
f(y|x) = \sum_{k=0}^{\infty} A_k(x) \cdot \cos \left( k\pi \frac{y-a}{b-a} \right), \quad (4.7)
\]
where the series coefficients $A_k(x)$ are defined as

$$A_k(x) := \frac{2}{b-a} \int_a^b f(y|x) \cos \left( k\pi \frac{y-a}{b-a} \right) \, dy.$$ 

Inserting (4.7) into $v_1(x,t)$, we get

$$v_1(x,t) = e^{-r\Delta t} \int_a^b v(y,T) \sum_{k=0}^{\infty} A_k(x) \cdot \cos \left( k\pi \frac{y-a}{b-a} \right) \, dy. \quad (4.8)$$

3. Interchange summation and integration.

We interchange summation and integration in (4.8)

$$v_1(x,t) = e^{-r\Delta t} \sum_{k=0}^{\infty} \frac{b-a}{2} A_k(x) \frac{2}{b-a} \int_a^b v(y,T) \cdot \cos \left( k\pi \frac{y-a}{b-a} \right) \, dy \quad (4.9)$$

and we define the Fourier-cosine series coefficients $V_k$ of $v(y,T)$ on $[a,b]$ as

$$V_k := \frac{2}{b-a} \int_a^b v(y,T) \cos \left( k\pi \frac{y-a}{b-a} \right) \, dy.$$ 

Insert $V_k$ into (4.9), we get

$$v_1(x,t) = e^{-r\Delta t} \sum_{k=0}^{\infty} \frac{b-a}{2} A_k(x)V_k. \quad (4.10)$$

The integral of the product of $f(y|x)$ and $v(y,T)$ is now written as a sum of the product of their Fourier-cosine coefficients.

4. Truncate the series summation.

As the coefficients $A_k, V_k$ have a fast decay to zero, the summation can be truncated,

$$v_2(x,t) = e^{-r\Delta t} \sum_{k=0}^{N-1} \frac{b-a}{2} A_k(x)V_k. \quad (4.11)$$

5. Insert the characteristic function.

We define the characteristic function of $f(y|x)$ on the interval $[a,b]$ by $\phi_A$. Whether the characteristic function of $f(y|x)$ is known or not known, it will always be defined on the whole domain $\mathbb{R}$. The function $f(y|x)$ decays to zero outside the domain $[a,b]$. Therefore $\phi_A$ does not differ much from $\phi$.

$$A_k(x) = \frac{2}{b-a} \int_a^b f(y|x) \cos \left( k\pi \frac{y-a}{b-a} \right) \, dy \quad (4.12)$$

$$= \frac{2}{b-a} \mathcal{R} \left\{ \phi_A \left( \frac{k\pi}{b-a} \right) \exp \left( -i \frac{ka\pi}{b-a} \right) \right\}$$

$$\approx \frac{2}{b-a} \mathcal{R} \left\{ \phi \left( \frac{k\pi}{b-a} \right) \exp \left( -i \frac{ka\pi}{b-a} \right) \right\}.$$
We introduce $F_k(x)$ defined as
\[
F_k(x) := \mathcal{R} \left\{ \phi \left( \frac{k\pi}{b-a} \right) \exp \left( -i \frac{ka\pi}{b-a} \right) \right\},
\] (4.13)

We insert $F_k(x)$ into (4.11) which gives the COS pricing formula:
\[
v(x, t) \approx v_3(x, t) = e^{-r\Delta t} \sum_{k=0}^{N-1} F_k(x) V_k.
\] (4.14)

4.2. Approximations of the conditional expectations

We wish to approximate the following conditional expectations:
\[
\begin{align*}
\mathbb{E}[Z_{i+1}^\Pi(X_{i+1}^\Pi) | \mathcal{F}_i] \\
\mathbb{E}[Y_{i+1}^\Pi(X_{i+1}^\Pi) \Delta W_i | \mathcal{F}_i] \\
\mathbb{E}[f(t_{i+1}, X_{i+1}^\Pi, Y_{i+1}^\Pi(X_{i+1}^\Pi), Z_{i+1}^\Pi(X_{i+1}^\Pi)) \Delta W_i | \mathcal{F}_i] \\
\mathbb{E}[Y_{i+1}^\Pi(X_{i+1}^\Pi) | \mathcal{F}_i] \\
\mathbb{E}[f(t_{i+1}, X_{i+1}^\Pi, Y_{i+1}^\Pi(X_{i+1}^\Pi), Z_{i+1}^\Pi(X_{i+1}^\Pi)) | \mathcal{F}_i]
\end{align*}
\]

The above conditional expectations can be categorized in two groups, with $v$ a general function:
\[
\begin{align*}
\mathbb{E}[v(t_{i+1}, X_{i+1}^\Pi)|X_i^\Pi = x] & \quad (4.15) \\
\mathbb{E}[v(t_{i+1}, X_{i+1}^\Pi) \Delta W_i|X_i^\Pi = x] & \quad (4.16)
\end{align*}
\]

First we will approximate (4.15). By the Fourier-Cosine method on $v(t_{i+1}, X_{i+1}^\Pi)$, we get
\[
v(t_{i+1}, X_{i+1}^\Pi) = \sum_{k=0}^{\infty} \mathcal{V}_k \cos \left( k\pi \frac{X_{i+1}^\Pi - a}{b-a} \right) \\
\approx \sum_{k=0}^{N-1} \mathcal{V}_k \cos \left( k\pi \frac{X_{i+1}^\Pi - a}{b-a} \right).
\]

This gives us the following approximation for (4.15)
\[
\mathbb{E}[v(t_{i+1}, X_{i+1}^\Pi)|X_i^\Pi = x] \approx \sum_{k=0}^{N-1} \mathcal{V}_k \mathbb{E} \left[ \cos \left( k\pi \frac{X_{i+1}^\Pi - a}{b-a} \right) | X_i^\Pi = x \right]
\] (4.17)

with $\mathcal{V}_k(t_{i+1}) = \frac{2}{b-a} \int_a^b v(t_{i+1}, x) \cos \left( k\pi \frac{x-a}{b-a} \right) dx$. (4.18)
Now we will rewrite the expectation in (4.17)

\[
\mathbb{E} \left[ \cos \left( \frac{k \pi \text{X}_{i+1}^\Pi - a}{b - a} \right) \mid \text{X}_i^\Pi = x \right]
\]

= \mathcal{R} \left\{ \mathbb{E} \left[ \exp \left( \frac{i k \pi \text{X}_{i+1}^\Pi - a}{b - a} \right) \mid \text{X}_i^\Pi = x \right] \right\}

= \mathcal{R} \left\{ \exp \left( \frac{i k \pi - a}{b - a} \right) \mathbb{E} \left[ \exp \left( \frac{i k \pi \text{X}_{i+1}^\Pi - a}{b - a} \right) \mid \text{X}_i^\Pi = x \right] \right\}

= \mathcal{R} \left\{ \exp \left( \frac{i k \pi - a}{b - a} \right) \phi_{X_{i+1}^\Pi} \left( \frac{k \pi}{b - a} \right) \right\} \cdot \text{(by Lemma 3.7)} \tag{*}

We will insert the equation (*) back to (4.17) and finally for (4.15) we get

\[
\mathbb{E} \left[ v(t_{i+1}, \text{X}_{i+1}^\Pi) \mid \text{X}_i^\Pi = x \right] \approx \sum_{k=0}^{N-1} \mathcal{V}_k(t_{i+1}) \mathcal{R} \left\{ \phi_{X_{i+1}^\Pi} \left( \frac{k \pi}{b - a} \right) \exp \left( \frac{i k \pi - a}{b - a} \right) \right\}
\]

Now we will approximate (4.16) in the same manner as (4.15) but in a slightly more complicated setting. We again use the Fourier-Cosine method on \(v(t_{i+1}, \xi)\) and get

\[
v(t_{i+1}, \text{X}_{i+1}^\Pi) = \sum_{k=0}^{\infty} \mathcal{V}_k \cos \left( \frac{k \pi \text{X}_{i+1}^\Pi - a}{b - a} \right)
\]

\[
\approx \sum_{k=0}^{N-1} \mathcal{V}_k \cos \left( \frac{k \pi \text{X}_{i+1}^\Pi - a}{b - a} \right)
\]

This gives us the following approximation for (4.16),

\[
\mathbb{E} \left[ v(t_{i+1}, \text{X}_{i+1}^\Pi) \Delta W_i \mid \text{X}_i^\Pi = x \right] \approx \sum_{k=0}^{N-1} \mathcal{V}_k \mathbb{E} \left[ \cos \left( \frac{k \pi \text{X}_{i+1}^\Pi - a}{b - a} \right) \Delta W_i \mid \text{X}_i^\Pi = x \right] \tag{4.19}
\]

with \(\mathcal{V}_k(t_{i+1}) = \frac{x}{b - a} \int_a^b v(t_{i+1}, x) \cos \left( \frac{k \pi x - a}{b - a} \right) \, dx\).

From (4.19) we rewrite the expectation,

\[
\mathbb{E} \left[ \cos \left( \frac{k \pi \text{X}_{i+1}^\Pi - a}{b - a} \right) \Delta W_i \mid \text{X}_i^\Pi = x \right]
\]

= \mathcal{R} \left\{ \mathbb{E} \left[ \exp \left( \frac{i k \pi \text{X}_{i+1}^\Pi - a}{b - a} \right) \Delta W_i \mid \text{X}_i^\Pi = x \right] \right\}

= \mathcal{R} \left\{ \mathbb{E} \left[ \exp \left( \frac{i k \pi \text{X}_{i+1}^\Pi - a}{b - a} \right) \Delta W_i \exp \left( \frac{i k \pi - a}{b - a} \right) \mid \text{X}_i^\Pi = x \right] \right\}

= \mathcal{R} \left\{ \exp \left( \frac{i k \pi - a}{b - a} \right) \mathbb{E} \left[ \Delta W_i \exp \left( \frac{i k \pi \text{X}_{i+1}^\Pi - a}{b - a} \right) \mid \text{X}_i^\Pi = x \right] \right\}. \tag{4.20}
Applying integration by parts and substituting $u = \frac{xt}{b-a}$ for the expectation term in equation (4.20), we get
\[
E \left[ \Delta W_t \exp (i u X^\Pi_t) \mid X^\Pi_t = x \right] = E \left[ \Delta W_t \exp \left( i u x + i u b(t_i, x, y(t_i, x), z(t_i, x) \Delta W_t) \Delta t + i u \sigma(t_i, x, y(t_i, x), z(t_i, x)) \right) \mid X^\Pi_t = x \right] (\ast)
\]
by using $b(t_i, x, y(t_i, x), z(t_i, x)) := m(t_i, x)$, $\sigma(t_i, x, y(t_i, x), z(t_i, x)) := s(t_i, x)$ and $\Delta W_t \sim N(0, \Delta t)$
\[
(\ast) = \frac{1}{\sqrt{2\pi \sqrt{\Delta t}}} \int_{-\infty}^{+\infty} \exp (i u x + i u m(t_i, x) \Delta t + i u s(t_i, x) \xi) \exp \left( -\frac{1}{2} \frac{\xi^2}{\Delta t} \right) d\xi
\]
\[
= -\frac{1}{\sqrt{2\pi \sqrt{\Delta t}}} \int_{-\infty}^{+\infty} \exp (i u x + i u m(t_i, x) \Delta t + i u s(t_i, x) \xi) \Delta t \exp \left( -\frac{1}{2} \frac{\xi^2}{\Delta t} \right) \frac{d\xi}{\sqrt{\Delta t}}
\]
\[
+ \frac{1}{\sqrt{2\pi \Delta t}} \int_{-\infty}^{+\infty} i u s(t_i, x) \exp (i u x + i u m(t_i, x) \Delta t + i u s(t_i, x) \xi) \exp \left( -\frac{1}{2} \frac{\xi^2}{\Delta t} \right) d\xi
\]
\[
= 0 + i u s(t_i, x) \Delta t \left[ \exp (i u (x + m(t_i, x) \Delta t + s(t_i, x) \xi)) \exp \left( -\frac{1}{2} \frac{\xi^2}{\Delta t} \right) d\xi \right]
\]
\[
= i u s(t_i, x) \Delta t \mathbb{E} \left[ \exp (i u (x + m(t_i, x) \Delta t + s(t_i, x) \Delta W_t)) \right]
\]
\[
= s(t_i, x) \Delta t \mathbb{E} \left[ i u \exp (i u (X^\Pi_t + m(t_i, x) \Delta t + s(t_i, x) \Delta W_t)) \mid X^\Pi_t = x \right]
\]
\[
= s(t_i, x) \Delta t \mathbb{E} \left[ i u \exp (i u X^\Pi_{t_i+1}) \mid X^\Pi_{t_i} = x \right]
\]
\[
= \sigma(t_i, x, y(t_i, x), z(t_i, x)) \Delta t \mathbb{E} \left[ D \exp (i u X^\Pi_{t_i+1}) \mid X^\Pi_t = x \right]
\]
\[
= \sigma(t_i, x, y(t_i, x), z(t_i, x)) \Delta t \mathbb{E} \left[ i u \phi^x_{X^\Pi_{t_i+1}} (u) \right].
\]
(by Lemma 3.7)

Finally we get,
\[
E \left[ v(t_{i+1}, X^\Pi_{t_{i+1}}) \Delta W_t | X^\Pi_t = x \right]
\]
\[
\approx \Delta t \sigma(t_i, x, y(t_i, x), z(t_i, x)) \sum_{k=0}^{N-1} \mathcal{V}_k(t_{i+1}) R \left\{ E \left[ D \exp (i u X^\Pi_{t_{i+1}}) \mid X^\Pi_t = x \right] \left( i k \pi - \frac{a}{b-a} \right) \right\}
\]
\[
= \Delta t \sigma(t_i, x, y(t_i, x), z(t_i, x)) \sum_{k=0}^{N-1} \mathcal{V}_k(t_{i+1}) R \left\{ i k \pi \phi^x_{X^\Pi_{t_{i+1}}} \left( \frac{k \pi}{b-a} \right) \exp \left( i k \pi - \frac{a}{b-a} \right) \right\}
\]
with $\mathcal{V}_k(t_{i+1}) = \frac{2}{b-a} \int_a^b v(t_{i+1}, x) \cos \left( \frac{k \pi x - a}{b-a} \right) dx$. 

48
4.2.1. COS approximation of the function \(Z_i^\Pi\)

For the computation of \(Z_i^\Pi\) in (3.17), we need to compute three expectations, 
\[
\mathbb{E}[Z_{i+1}^\Pi(X_{i+1}^\Pi)|F_t] \text{, } \mathbb{E}[Z_{i+1}^\Pi(X_{i+1}^\Pi)\Delta W_t|F_t] \text{ and } \mathbb{E}[f(t_{i+1}, X_{i+1}^\Pi, Y_{i+1}^\Pi(X_{i+1}^\Pi), Z_{i+1}^\Pi(X_{i+1}^\Pi))\Delta W_t|F_t].
\]

With the help of COS formulas we can derive the following approximations for these expectations:

\[
\mathbb{E}[Z_{i+1}^\Pi(X_{i+1}^\Pi)|F_t] \approx \sum_{k=0}^{N-1} Z_k(t_{i+1})\mathcal{R}\left\{\phi_{X_{i+1}^\Pi}^r\left(\frac{k\pi}{b-a}\right) \exp\left(ik\pi\frac{a}{b-a}\right)\right\}.
\]

\[
\mathbb{E}[Y_{i+1}^\Pi(X_{i+1}^\Pi)\Delta W_t|F_t] 
\approx \Delta t\sigma(t_i, x, y(t_i, x), z(t_i, x))\sum_{k=0}^{N-1} Y_k(t_{i+1})\mathcal{R}\left\{i\frac{k\pi}{b-a}\phi_{X_{i+1}^\Pi}^r\left(\frac{k\pi}{b-a}\right) \exp\left(ik\pi\frac{a}{b-a}\right)\right\}.
\]

\[
\mathbb{E}[f(t_{i+1}, X_{i+1}^\Pi, Y_{i+1}^\Pi(X_{i+1}^\Pi), Z_{i+1}^\Pi(X_{i+1}^\Pi))|F_t] 
\approx \Delta t\sigma(t_i, x, y(t_i, x), z(t_i, x))\sum_{k=0}^{N-1} f_k(t_{i+1})\mathcal{R}\left\{i\frac{k\pi}{b-a}\phi_{X_{i+1}^\Pi}^r\left(\frac{k\pi}{b-a}\right) \exp\left(ik\pi\frac{a}{b-a}\right)\right\}.
\]

By using the above equalities we obtain the COS approximation for \(Z_i^\Pi\) in (3.17).

4.2.2. COS approximation of the function \(Y_i^\Pi\)

For the computation of \(Y_i^\Pi\) in (3.18), we need to compute two expectations, 
\[
\mathbb{E}[Y_{i+1}^\Pi(X_{i+1}^\Pi)|F_t] \text{ and } \mathbb{E}[f(t_{i+1}, X_{i+1}^\Pi, Y_{i+1}^\Pi(X_{i+1}^\Pi), Z_{i+1}^\Pi(X_{i+1}^\Pi))|F_t],
\]

that are approximated by the following COS formulas:

\[
\mathbb{E}[Y_{i+1}^\Pi(X_{i+1}^\Pi)|F_t] \approx \sum_{k=0}^{N-1} Y_k(t_{i+1})\mathcal{R}\left\{\phi_{X_{i+1}^\Pi}^r\left(\frac{k\pi}{b-a}\right) \exp\left(ik\pi\frac{a}{b-a}\right)\right\}.
\]

\[
\mathbb{E}[f(t_{i+1}, X_{i+1}^\Pi, Y_{i+1}^\Pi(X_{i+1}^\Pi), Z_{i+1}^\Pi(X_{i+1}^\Pi))|F_t] 
\approx \sum_{k=0}^{N-1} f_k(t_{i+1})\mathcal{R}\left\{\phi_{X_{i+1}^\Pi}^r\left(\frac{k\pi}{b-a}\right) \exp\left(ik\pi\frac{a}{b-a}\right)\right\}.
\]

When \(\theta_1 > 0\) we obtain an implicit part in (3.18), for which we perform \(P\) Picard iterations [26] starting with initial guess \(\mathbb{E}\left[Y_{i+1}^\Pi(X_{i+1}^\Pi)|X_i = x\right]\).

By using the above equalities and the Picard iteration method we obtain the COS approximation for \(Y_i^\Pi\) in (3.18).
4.2.3. Fourier-cosine coefficients

Let $Y_k(t_i)$ be the Fourier-cosine coefficients of $Y_i^\Pi(x)$, i.e.,

$$ Y_k(t_i) = \frac{2}{b-a} \int_a^b Y_i^\Pi(x) \cos \left( k\pi \frac{x-a}{b-a} \right) \, dx, $$

(4.21)

and $Z_k(t_{i+1})$ be the Fourier-cosine coefficients of $Z_i^\Pi(x)$, i.e.,

$$ Z_k(t_i) = \frac{2}{b-a} \int_a^b Y_i^\Pi(x) \cos \left( k\pi \frac{x-a}{b-a} \right) \, dx, $$

(4.22)

and $F_k(t_i)$ the Fourier-cosine coefficients of driver $f(t_i, x, Y_i^\Pi(x), Z_i^\Pi(x))$,

$$ F_k(t_i) = \frac{2}{b-a} \int_a^b f(t_i, x, Y_i^\Pi(x), Z_i^\Pi(x)) \cos \left( k\pi \frac{x-a}{b-a} \right) \, dx. $$

(4.23)

At maturity $T$, we have the following Fourier-cosine coefficients

$$ Y_k(t_M) = \frac{2}{b-a} \int_a^b g(x) \cos \left( k\pi \frac{x-a}{b-a} \right) \, dx, $$

(4.24)

$$ Z_k(t_M) = \frac{2}{b-a} \int_a^b \sigma(x) Dg(x) \cos \left( k\pi \frac{x-a}{b-a} \right) \, dx. $$

(4.25)

4.2.4. Discretization of the Fourier-cosine coefficient

In this subsection we will refer to the general Fourier-cosine coefficient $V_k(t_{i+1})$ in (4.18), but the same holds for $F_k(t_{i+1}), Z_k(t_{i+1}), Y_k(t_{i+1})$. The Fourier-cosine coefficient $V_k(t_{i+1})$ of $v$ at time points $t_i$ where $i = 0, ..., M$ and $j = 0, ..., N-1$ is not always available to us. So the integral needs to be approximated with a discrete Fourier-cosine transform. To obtain such an approximation, we take $N$ equidistant grid-points $x_n$ on the grid $[a, b]$ and compute the value of $v(t_{i+1}, x)$ on these grid-points.

$$ V_k(t_{i+1}) = \frac{2}{b-a} \int_a^b v(t_{i+1}, x) \cos \left( k\pi \frac{x-a}{b-a} \right) \, dx. $$

$$ \approx \frac{2}{b-a} \sum_{n=0}^{N-1} v(t_{i+1}, x_n) \cos \left( k\pi \frac{x_n-a}{b-a} \right) \delta x $$

$$ = \frac{2}{N} \sum_{n=0}^{N-1} v(t_{i+1}, x_n) \cos \left( k\pi \frac{2n+1}{2N} \right). $$

(4.26)

The approximation (4.26) is known as the discrete Fourier-cosine transform.
4.3. BCOS summarized

<table>
<thead>
<tr>
<th>BCOS method</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Initial step:</strong> Compute, or approximate, the terminal coefficients $\mathcal{Y}_k(t_M)$, $\mathcal{Z}_k(t_M)$ with formulas (4.24) and (4.25).</td>
</tr>
<tr>
<td><strong>Loop step:</strong> For $i = M - 1$ to $i = 1$ approximate the necessary conditional expectations with Section 4.3 and compute the functions $Y_i^\Pi(x)$, $Z_i^\Pi(x)$ and $f(t_i, x, Y_i^\Pi(x), Z_i^\Pi(x))$ for $x \in [a, b]$ with formulas (3.17) and (3.18). Determine the corresponding Fourier-cosine coefficients $Z_k(t_i)$, $\mathcal{F}_k(t_i)$, $\mathcal{Y}_k(t_i)$ with formulas (4.21)-(4.23). Those integrals can be approximated by the discrete Fourier-cosine transform, as in section 4.3.4.</td>
</tr>
<tr>
<td><strong>Terminal step:</strong> Compute $Y_0^\Pi(X_0)$ and $Z_0^\Pi(X_0)$.</td>
</tr>
</tbody>
</table>

51
5. Numerical experiment: Hedging in the Black-Scholes model

We use the COS method for performing a numerical example taken from [26]. The calculations are performed in MATLAB.

We compute the price \( v(t, S_t) \) of a call option by a BSDE where the underlying asset follows a geometric Brownian motion,

\[
\mathrm{d}S_t = \mu S_t \, \mathrm{d}t + \sigma S_t \, \mathrm{d}W_t.
\]

The exact solution can be computed analytically by solving the Black-Scholes equation with the help of the Feynman-Kac theorem for linear PDEs.

For the derivation of the Black-Scholes PDE we set up a self-financing portfolio \( Y_t \) with \( a_t \) assets and \( Y_t - a_t S_t \) bonds as well as \( r \) the risk-free return.

In this model we assume that the market is complete, there are no trading restrictions and the option can be exactly replicated by the hedging portfolio \( Y_T = \max(S_T - K, 0) \).

Then, the option value at initial time should be equal to the initial value of the portfolio. The portfolio follows the following SDE

\[
\mathrm{d}Y_t = r(Y_t - a_t S_t) \, \mathrm{d}t + a_t S_t \, \mathrm{d}W_t,
\]

If we set \( Z_t = \sigma a_t S_t \), then \((Y, Z)\) solves the BSDE

\[
\begin{align*}
\mathrm{d}Y_t &= f(t, Y_t, Z_t) \, \mathrm{d}t + Z_t \, \mathrm{d}W_t \\
\mathrm{d}Z_t &= -f(t, Y_t, Z_t) \, \mathrm{d}t + \frac{\mu - r}{\sigma} \, \mathrm{d}W_t,
\end{align*}
\]

\(Y_T = \max(S_T - K, 0)\).

\(Y_t\) corresponds to the value of the portfolio and \(Z_t\) is related to the hedging strategy.

In this case, the option value is given by \( v(t, S_t) = Y_t \) and \( \sigma S_t \, \mathrm{d}v(t, S_t) = Z_t \).

For the numerical approximation, we need to switch to log asset domain \( X_t = \log S_t \), with

\[
\mathrm{d}X_t = (\mu - \frac{1}{2} \sigma^2) \, \mathrm{d}t + \sigma \, \mathrm{d}W_t.
\]

We set \( \bar{\mu} = \mu - \frac{1}{2} \sigma^2 \), thus

\[
\mathrm{d}X_t = \bar{\mu} \, \mathrm{d}t + \sigma \, \mathrm{d}W_t.
\]
and $Y_t = v(t, e^{X_t})$, $Z_t = \sigma e^{X_t} v_s(t, e^{X_t})$.

We take the following parameter values

$$X_0 = \log(100), \ K = 100, \ r = 0.1, \ \mu = 0.2, \ \sigma = 0.25, \ T = 0.1,$$

with the exact solutions $Y_0 = v(t_0, S_0) = 3.65997$ and $Z_0 = \sigma S_0 v_s(t_0, S_0) = 14.14823$.

For the results of the BCOS method, we refer to the Figure 8.2 for the approximated value

![Graph](image)

**Figure 5.1.** In this example, we take $N = 2^9$, timesteps $M = 10$. On the left graph, we see the error of $\hat{Y}(t_0, x_0)$ and on the right, the error of $\hat{Z}(t_0, x_0)$. 

53
A. Appendix

Theorem A.1. [14] Burkholder-Davis-Gundy inequalities
This gives bounds for the maximum of a martingale in terms of the quadratic variation. For a local continuous martingale $M_t = \int_0^t \phi_s dW_s$ starting at zero, with maximum denoted by $M^*_t = \sup_{s \leq t} |M_s|$, and any real number $p \geq 1$, the inequality is

$$c_p \mathbb{E}\left( \int_0^t \phi^2_s ds \right)^{\frac{p}{2}} \leq \mathbb{E}(M^*_t)^p \leq C_p \mathbb{E}\left( \int_0^t \phi^2_s ds \right)^{\frac{p}{2}}.$$ 

Here, $c_p < C_p$ are constants depending on the choice of $p$, but not depending on the martingale $M$ or time $t$ used and $p > 0$.

Theorem A.2. [10] Martingale’s Representation Theorem
For $0 \leq t \leq T$, let $W_t$ be a Brownian motion on $(\Omega, \mathcal{F}, \mathbb{P})$ and $\mathcal{F}_t$ be the filtration generated by this Brownian motion. Let $M_t$ be a square-integrable martingale (under $\mathbb{P}$) relative to this filtration. Then there exists a unique adapted process $Z_t$ such that

$$M_t = M_0 + \int_0^t Z_s dW_s, \quad 0 \leq t \leq T.$$ 

In particular, the paths of $M$ are continuous.

Let $T > 0$, $c \geq 0$, $u(\cdot)$ a measurable non negative function in $[0, T]$, and $v(\cdot)$ an integrable non negative function in $[0, T]$. If

$$u(t) \leq c + \int_0^t v(s)u(s)ds, \quad \forall t \in [0, T],$$

then

$$u(t) \leq c \exp\left( \int_0^t v(s)ds \right).$$

Theorem A.4. Young’s inequality
Let $a, b$ be non negative real numbers and $p, q \in (0, \infty]$ such that $\frac{1}{p} + \frac{1}{q} = 1$. Then

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}.$$ 

For $p = q = 2$, we have

$$ab \leq \frac{a^2}{2} + \frac{b^2}{2}$$

and for an $\epsilon > 0$ we have the so-called Peter–Paul inequality.

$$ab \leq \frac{a^2}{2\epsilon} + \frac{\epsilon b^2}{2}.$$ 

54
Lemma A.5. [3] For $a, b \in \mathbb{R}$ it holds

$$|a + b|^2 \leq 2|a|^2 + 2|b|^2$$

Proof. This result is a direct conclusion from Young’s inequality (A.4):

$$|a + b|^2 \leq |a|^2 + 2|a||b| + |b|^2 \leq 2|a|^2 + 2|b|^2$$

Lemma A.6. [3] For positive, real numbers $\lambda, \mu > 0$ and $C, y, z, t \geq 0$, the inequality

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

is valid.

Proof. By Young’s inequality (A.4), it follows that

$$2yt \leq \frac{t^2}{\mu^2} + y^2\mu^2$$

$$2yz \leq \frac{Cz^2}{\lambda^2} + Cy^2\lambda^2,$$

thus

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

Theorem A.7. [5] Itô isometry

For a local continuous martingale $M_t = \int_0^t \phi_s dW_s$ starting at zero. The stochastic integral satisfies the Itô isometry

$$\mathbb{E} \left( \int_0^t \phi_s dW_s \right)^2 = \mathbb{E} \int_0^t \phi_s^2 ds,$$

which holds when $H$ is bounded or, more generally, when the integral on the right hand side is finite.


Let $S_t = \sup_{0 \leq s \leq t} X_s$ and for $p \geq 1$ and

$$\|X_t\|_p = (\mathbb{E} [|X_t|^p])^{\frac{1}{p}}.$$

where $X$ is a martingale, for $p > 1$,

$$\|X_T\|_p \leq \|S_T\|_p \leq \frac{p}{p-1} \|X_T\|_p.$$
Theorem A.9. [16] Banach Fixed Point Theorem
Let $\mathcal{H}$ be a non-empty Banach space and $f : \mathcal{H} \to \mathcal{H}$ be a contraction on $\mathcal{H}$ (i.e. there is a real number $c$ such that $c \in (0, 1)$ and for any $x, y \in \mathcal{H}$: $||f(x) - f(y)|| \leq c||x - y||$). Then $f$ has only one fixed point.

Theorem A.10. [17] Let $Z_1, Z_2, ..., Z_n$ be independent random variables with $Z_i \sim \mathcal{N}(0, 1)$. If $Y = \sum_{i=1}^{n} z_i^2$ then $Y$ follows the chi-square distribution with $n$ degrees of freedom. We write

$$Y \sim \chi^2_n$$

Theorem A.11. [17] Let $X_1, X_2, ..., X_n$ be independent random variables with $X_i \sim \mathcal{N}(\mu, \sigma)$. If $Y = \sum_{i=1}^{n} \left( \frac{X_i - \mu}{\sigma} \right)^2$ then $Y$ follows the chi-square distribution with $n$ degrees of freedom. We write

$$Y \sim \chi^2_n$$

Theorem A.12. [14] Suppose that the coefficients $b(t, x), \sigma(t, x)$ of the SDE $dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t$ are locally Lipschitz-continuous in the space variable; i.e., there exists a constant $C > 0$ such that for every $t \geq 0$:

$$|b(t, x) - b(t, x')| + |\sigma(t, x) - \sigma(t, x')| \leq C|x - x'|$$

$$|\sigma(t, x)| + |b(t, x)| \leq C(1 + |x|)$$

Then the SDE has indeed a strong solution.
Conclusion

In this thesis we considered a Backward SDEs (BSDEs), where the generator satisfies the Lipschitz condition. In order to find a $\mathcal{F}_t$-adapted solution to the BSDE an additional $\mathcal{F}_t$-adapted process $Z$ has to be introduced, which results in the fact that the solution of a BSDE consists of a pair of processes $(Y, Z)$. By the existence and uniqueness theorem we proved that this pair of processes indeed exists and is unique. Solving BSDEs are used in stochastic optimal control problems and hedging problems in finance where $Y$ represents the dynamics of the value of the replicating portfolio and $Z$ corresponds to the hedging strategy. When the solution of a BSDE is associated with a state process satisfying a forward SDE, we showed that this Forward-Backward SDE satisfies the Markovian property. By the famous Feynman-Kac formula Markovian BSDEs yield a representation formula for semi-linear parabolic PDEs. Conversely, under smoothness assumptions (i.e. viscosity solution), the solution of the BSDE corresponds to the solution of a semi-linear parabolic PDEs.

A drawback of BSDEs is that they can rarely be solved explicitly, so simulation of BSDEs is of big importance. We introduce a probabilistic numerical method for solving backward stochastic differential equation, whereas the design of numerical schemes for BSDEs is an important task. Since BSDEs are terminal value problems for stochastic differential equations, a natural time discretization for BSDEs works backwards in time. However, the solution must be adapted to the information, which makes the construction of numerical solutions to BSDEs a more challenging problem compared to initial value problems for stochastic differential equations. In order to keep the time discretized solution adapted to the filtration, the time discretization scheme requires nestings of the conditional expectation. In the next step, an estimator for the conditional expectation had to be applied. We studied the method, called BCOS method, developed in [26]. This method is based on Fourier-cosine series expansions and relies on the characteristic function of the transitional probability density function. This approach allowed us to approximate the conditional expectations in an efficient way. When the characteristic function of the underlying process cannot be derived easily, we can use the characteristic function of a discrete forward process to approximate the solution, where we approximated the underlying forward SDE by the Euler scheme. The applicability of the resulting method is therefore quite general. A numerical test demonstrates the applicability of the BCOS method for BSDEs in economic and financial problems.
Popular summary

Whereas the theory and applications of classical forward stochastic differential equations (FSDE), with a known initial value, is traditional, in this thesis we studied backward stochastic differential equations (BSDEs). A BSDE is a stochastic differential equation for which a terminal condition, instead of an initial condition, is known beforehand.

In recent years, BSDEs have received more attention in mathematical finance and economics, for example, the Black-Scholes formula for pricing options can be represented by a system of (decoupled) Forward-Backward SDE. The solution of a BSDE consists of a pair of processes. However as in most of the cases it is not possible to analytically solve BSDEs, numerical methods are in great demand. These so-called probabilistic numerical methods for solving BSDEs can be divided into two steps: a time discretization of the stochastic process and approximations for the appearing conditional expectations. In this thesis we employ a general theta-method for the time-integration and propose an approach based on the Fourier series, more precise on Fourier-Cosine series, to approximate the solution backwards in time. This method links the computation of the expectations to the characteristic function of the transitional probability density function, which is either given or easy to approximate. We call the method the BCOS method, short for BSDE-COS method.
Bibliography


