TWO STUDIES ON BACKWARD STOCHASTIC DIFFERENTIAL EQUATIONS

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

BY

VİLDAN TUNÇ

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

JULY 2012

Approval of the thesis:

TWO STUDIES ON BACKWARD STOCHASTIC DIFFERENTIAL EQUATIONS

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

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

Assoc. Prof. Dr. Ömür Uğur
Head of Department, **Financial Mathematics** __________

Assoc. Prof. Dr. Ali Devin Sezer
Supervisor, **Department of Financial Mathematics** __________

Examining Committee Members:

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

Assoc. Prof. Dr. Ali Devin Sezer
Institute of Applied Mathematics, METU __________

Assist. Prof. Dr. Özlem İlk
Department of Statistics, METU __________

Date: __________

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

Name, Last Name: VİLDAN TUNÇ

Signature :

ABSTRACT

TWO STUDIES ON BACKWARD STOCHASTIC DIFFERENTIAL EQUATIONS

Tunç, Vildan
M.Sc., Department of Financial Mathematics
Supervisor : Assoc. Prof. Dr. Ali Devin Sezer

July 2012, 48 pages

Backward stochastic differential equations appear in many areas of research including mathematical finance, nonlinear partial differential equations, financial economics and stochastic control. The first existence and uniqueness result for nonlinear backward stochastic differential equations was given by Pardoux and Peng (Adapted solution of a backward stochastic differential equation. System and Control Letters, 1990). They looked for an adapted pair of processes $\{x(t), y(t);\ t \in [0, 1]\}$ with values in $\mathbb{R}^d$ and $\mathbb{R}^{d\times k}$ respectively, which solves an equation of the form: $x(t) + \int_t^1 f(s, x(s), y(s))\ ds + \int_t^1 [g(s, x(s)) + y(s)]\ dW_s = X$. This dissertation studies this paper in detail and provides all the steps of the proofs that appear in this seminal paper. In addition, we review (Cvitanić and Karatzas, Hedging contingent claims with constrained portfolios. The annals of applied probability, 1993). In this paper, Cvitanić and Karatzas studied the following problem: the hedging of contingent claims with portfolios constrained to take values in a given closed, convex set K. Processes intimately linked to BSDEs naturally appear in the formulation of the constrained hedging problem. The analysis of Cvitanić and Karatzas is based on a dual control problem. One of the contributions of this thesis is an algorithm that numerically solves this control problem in the case of constant volatility. The algorithm is based on discretization of time. The convergence proof is also

provided.

Keywords: Backward stochastic differential equations, constrained replication, dual control problem, mathematical finance

ÖZ

GERİYE DOĞRU STOKASTİK DİFERANSİYEL DENKLEMLER ÜZERİNE İKİ ÇALIŞMA

Tunç, Vildan
Yüksek Lisans, Finansal Matematik Bölümü
Tez Yöneticisi : Doç. Dr. Ali Devin Sezer

Temmuz 2012, 48 sayfa

Geriye doğru stokastik diferansiyel denklemler, finansal matematik, doğrusal olmayan kısmi diferansiyel denklemler, finansal ekonomi ve stokastik kontrol alanları dahil olmak üzere birçok uygulama ve teorik çalışmalarda yer almıştır. Doğrusal olmayan geriye doğru diferansiyel denklemlerin çözümü ilk olarak Pardoux ve Peng (Adapted solution of a backward stochastic differential equation. System and Control Letters, 1990) tarafından ortaya koyulmuştur. Pardoux ve Peng, $x(t) + \int_t^1 f(s, x(s), y(s))\ ds + \int_t^1 [g(s, x(s)) + y(s)]\ dW_s = X$ formundaki denklemi çözen ve sırasıyla $\mathbb{R}^d$ ve $\mathbb{R}^{d\times k}$'da değer alan $\{x(t), y(t);\ t \in [0, 1]\}$ sürecinin varlığını ve tekliğini kanıtlamışlardır. Bu tez, bu makalede yer alan ispatların makalede belirtilmeyen tüm adımlarını vermektedir. Bu makaleye ek olarak, Cvitanić ve Karatzas'ın (Hedging contingent claims with constrained portfolios. The annals of applied probability, 1993) makalesi çalışılmıştır. Bu makalede, Cvitanić ve Karatzas, finansal ürünlerin kapalı ve konveks K kümesinde değer alan portföyler kullanılarak replike edilmesi problemini analiz etmişlerdir. Cvitanić ve Karatzas'ın incelemeleri dual kontrol problemine dayanmaktadır. Bu tezin son katkısı, volatilite sabit alındığında dual kontrol sorusunu numerik olarak çözen bir algoritma geliştirmesidir. Bu algoritma, zamanın kesikleştirilmesi (uzunluğu birbirine eş

parçalara bölünmesi) ile elde edilmiştir. Algoritmanın elde ettiği sonucun asıl kontrol sorusunun sonucuna yakınsadığı ispat edilmiştir.

Anahtar Kelimeler: Geriye doğru stokastik diferansiyel denklemler, portföy üzerinde kısıtlar olduğunda replikasyon, dual kontrol sorusu, matematiksel finans

To my family

ACKNOWLEDGMENTS

I would like to express my gratitude to all people who supported and helped me during the work on this thesis. First of all, I would like to acknowledge the efforts of all the members of department who taught and advised me during the M.Sc. courses. I am very grateful to my supervisor Assoc. Dr. Ali Devin Sezer who always supported me in every conceivable way including many mathematical suggestions. I thank him for his invaluable guidance and motivation through this dissertation. I would like to give special thanks to Prof. Dr. Gerhard Wilhelm Weber for his encouragement. I have been supported by the Yurtiçi Yüksek Lisans Bursu (National Masters Degree Scholarship) of TÜBİTAK, for which I am grateful. Last but not least, I would like to thank my family, especially my parents, for their patience and encouragement during my whole studies.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

Stochastic differential equations with equality constraints on the position of the process at terminal time are called backward stochastic differential equations (BSDEs). This thesis consists of a careful study of two important papers [14] (Pardoux and Peng, Adapted Solution of a Backward Stochastic Differential Equation, 1990) and [5] (Cvitanić and Karatzas, Hedging Congtingent Claims with Constrained Portfolios, 1993) in the BSDE literature and a numerical solution of an optimal control problem that arises in the latter.

Let us give a very brief and partial review of the history of BSDEs. To the best of our knowledge linear BSDEs were first observed by Bismut [3] in 1973. BSDEs in this context arise as a stochastic version of the adjoint equation in Pontryagin's maximum principle. In 1990, Pardoux and Peng [14] considered general BSDEs and gave the first existence and uniqueness result for nonlinear BSDEs. They studied a BSDE of the type

$$dY_t = -f(t, Y_t, Z_t)dt + Z_t dW_t; \quad Y_T = \epsilon,$$

where the coefficient $f : [0, T] \times \mathbb{R}^d \times \mathbb{R}^{d\times m} \to \mathbb{R}^d$ (called generator) is uniformly Lipschitz continuous in variable Y , Z and the terminal condition ϵ is square integrable. The unique solution is a pair $(Y, Z) = (Y_t, Z_t)_{0\leq t\leq T}$ of square integrable $\mathcal{F}$-adapted processes where F is the complete σ - algebra generated by m-dimensional Brownian motion W on a probability space $(\Omega, \mathcal{F}, P)$. This paper will be studied in detail in Chapter 2.

Since the paper [14] of Pardoux and Peng, BSDEs have been extensively studied and used in many applied and theoretical areas, particularly in mathematical finance. Based on the theory of BSDEs, Peng [15] derived a general stochastic maximum principle with first and second order adjoint equations. In 1992, Pardoux and Peng [13] gave a generalization of the Feynman Kac formula and showed that a probabilistic solution of a nonlinear partial differential equations (PDEs) corresponds to the solution of the BSDE in Markovian case. Moreover, Peng [16] also indicated the relationship between BSDEs and PDEs. In addition to Pardoux and Peng, Duffie and Epstein [7] studied nonlinear BSDEs to give a stochastic differential formulation of recursive utilities and their properties in the setting of Brownian information.

A new type of BSDEs, namely Forward-Backward Stochastic Differential Equations (FBSDEs in short) was initiated by Antonelli [1]. FBSDEs were also investigated by Ma, Protter and Yong [10] who established the "Four Step Scheme" which is one of the first methods to solve FBSDEs in a Markovian setting. This method is based on Ito's rule to convert the FBSDE system to the corresponding nonlinear PDE and then under some strong regularity and growth conditions on the coefficients solve this PDE [6]. BSDEs with random jumps were studied by Barles, Buckdahn and Pardoux [2] in 1997. They proved that a viscosity solution of a system of parabolic integral-partial differential equations can be provided by the solution of the BSDE. The paper [8] by El Karoui, Peng and Quenez suggests that option pricing problems can be solved by using BSDEs and shows an outline how to apply BSDEs in finance.

In 2000, Rogue and El Karoui [18] used BSDEs to solve the utility maximization problem. Then, weak solutions of BSDEs were introduced and these solutions were used in the field of FBSDEs. In 2008, Ma et al. [11] studied weak solutions of FBSDEs by using martingale problem approach.

The existence and uniqueness of adapted processes that solve a second-order backward stochastic differential equation (2BSDE) is studied by Cheridito, Soner, Touzi and Victoir in 2007 [4]. They suggested in the paper [4] that 2BSDE provides a probabilistic representation for fully nonlinear PDEs and thus opens the door for Monte Carlo methods for high dimensional fully nonlinear PDEs. In 2010, Soner, Touzi and Zhang [19] provided a complete theory of exixtence and uniqueness for 2BSDEs.

BSDEs and backward stochastic differential *inequalities* (BSDI) also arise naturally in hedging problems in mathematical finance. To the best of our knowledge, one of the first papers that observe this is [5] by Cvitanić and Karatzas, which was published in 1993. [5] studies the following problem: in a market driven by Brownian Motion where there are constraints on admissible portfolios, expressed in terms of a convex set K , how to hedge a given financial security that makes a payment at terminal time. The approach of [5] consists of the following steps: 1) formulate the hedging problem as an optimization problem using BSDI; the resulting optimization problem is nonstandard and therefore it is handled by 2) using convex analysis to derive a standard stochastic optimal control problem that is dual to the original hedging problem. The paper then suggests the use of the HJB equation associated with the dual problem to provide a solution to the original hedging problem. These and other ideas and methods that appear in [5] we review in Chapter 4. In the same section, we suggest a different approach for the solution of the dual control problem in the case when volatility and interest rate are constant: namely, discretize time and obtain a discrete time stochastic optimal control problem which can be solved by dynamic programming. We use this method to hedge a standard call and put in the case when borrowing and lending interest rates differ.

The contents of Chapters 2 and 4 have already been outlined above. The stochastic calculus used in [14] of Pardoux and Peng is the one that we are familiar with from books such as Karatzas and Shreeve [9] and Oksendal [12]. However, the notation is different from that of these books. In Chapter 3, we explain this notation in detail and derive formulas for quadratic variations of several processes in terms of this notation. We have not come across these formulas in standard textbooks on the subject and, therefore, found it worth recording them here.

CHAPTER 2

BACKWARD STOCHASTIC DIFFERENTIAL EQUATIONS

In this chapter, we introduce the theory of Backward Stochastic Differential Equations (BSDEs). The content of this chapter is already published in Pardoux and Peng [14]. We supplemented a detailed exposition of the ideas and arguments used in this article.

2.1 Preliminaries

The main result of this article is an existence and uniqueness result for an adapted pair $\{x(t), y(t);\ t \in [0,1]\}$ which solves

$$x(t) + \int_t^1 f(s,\ x(s),\ y(s))\, ds + \int_t^1 [\, g(s, x(s)) + y(s)\,]\, dW(s) = X, \tag{2.1.1}$$

where

- $\{W(t),\ t \in [0,1]\}$ is a standard k-dimensional Wiener process defined on $(\Omega, \mathscr{F}, P)$,
- $\{\mathscr{F}_t,\ t \in [0,1]\}$ is its natural filtration (i.e., $\mathscr{F}_t = \sigma(W_s),\ 0 \le s \le t$),
- X is a given $\mathscr{F}_1$ measurable d-dimensional random vector such that $f : \Omega \times [0,1] \times \mathbb{R}^d \times \mathbb{R}^{d\times k} \to \mathbb{R}^d$,
- f is assumed to be $\mathscr{P} \otimes \mathscr{B}_d \otimes \mathscr{B}_{d\times k} / \mathscr{B}_d$ measurable where $\mathscr{P}$ denotes the σ-algebra of $\mathscr{F}_t$-progressively measurable subsets of $\Omega \times [0,1]$,
- f is uniformly Lipschitz with respect to both x and y.

There are also some other notations that will be used throughout this chapter. They are as follows:

- $M^2(0,1;\mathbb{R}^d)$ denotes the set of $\mathbb{R}^d$-valued processes which are $\mathscr{F}_t$-progressively measurable and are square integrable over $\Omega \times (0,1)$ with respect to $P \times \lambda$, where λ denotes the Lebesgue measure over [0,1],
- $M^2(0,1;\mathbb{R}^{d\times k})$ denotes the set of $\mathbb{R}^{d\times k}$-valued processes which are $\mathscr{F}_t$-progressively measurable and are square integrable over $\Omega \times (0,1)$ with respect to $P \times \lambda$,
- For $x \in \mathbb{R}^d$, $|x|$ denotes its Euclidean norm,
- An element $y \in \mathbb{R}^{d\times k}$ is considered as a $d \times k$ matrix, where its Euclidean norm is given by $|y| = \sqrt{Tr(yy^*)}$ and $(y,z) = Tr(yz^*)$.

2.2 Existence and Uniqueness of Solutions to BSDEs

Before we study equation (2.1.1), we first consider three simplified versions of that equation in this section. There are two simplified ones in the article [14]. Before we study those, we discuss the simplest possible BSDE.

Let X be $\mathscr{F}_T$-measurable random variable in $L^1(\Omega, \mathscr{F}_1, P; \mathbb{R}^d)$ such that $\mathbb{E}\left[|X|\right] < \infty$ and $M_t = \mathbb{E}\left[X|\mathscr{F}_t\right]$. It is obvious that M_t is a martingale.

"Martingale Representation Theorem" [9, page 182, thm 4.15] says that there exists Y such that

$$M_t = \mathbb{E}[X] + \int_0^t Y_s \, dW_s \tag{2.2.1}$$

and

$$M_T = \mathbb{E}[X|\mathscr{F}_T] = X. \tag{2.2.2}$$

On the other hand,

$$M_T = \mathbb{E}[X] + \int_0^T Y_s \, dW_s. \tag{2.2.3}$$

Combining equations (2.2.2) and (2.2.3), we get;

$$X = \mathbb{E}[X] + \int_0^T Y_s \, dW_s. \tag{2.2.4}$$

Substituting (2.2.4) into (2.2.1);

$$M_t = X - \int_0^T Y_s\, dW_s + \int_0^t Y_s\, dW_s. \tag{2.2.5}$$

By changing interval of the second integral, we get a negative one such that

$$M_t = X - \int_0^T Y_s\, dW_s - \int_t^0 Y_s\, dW_s. \tag{2.2.6}$$

Using one of the basic properties of integral, the equation (2.2.6) is of the form

$$M_t = X - \int_t^T Y_s\, dW_s. \tag{2.2.7}$$

The desired Backward Stochastic Differential Equation is

$$X = M_t + \int_t^T Y_s\, dW_s.$$

This result shows us that Backward Stochastic Differential Equations, even the simplest possible one, are generalizations of conditional expectation.

Now, we can turn back to the article [14].

Lemma 2.2.1 *[14, Lemma 2.1, page 56] Given* $X \in L^2(\Omega, \mathscr{F}_1, P; \mathbb{R}^d)$, $f \in M^2(0,1;\, \mathbb{R}^d)$ *and* $g \in M^2(0,1;\, \mathbb{R}^{d\times k})$*, there exists a unique pair* $(x, y) \in M^2(0,1;\, \mathbb{R}^d) \times M^2(0,1;\, \mathbb{R}^{d\times k})$ *such that*

$$x(t) + \int_t^1 f(s)\, ds + \int_t^1 [\, g(s) + y(s)\,]\, dW(s) = X, \quad 0 \le t \le 1. \tag{2.2.8}$$

Proof. Define

$$x(t) = \mathbb{E}\left[X - \int_t^1 f(s)\, ds \,|\mathscr{F}_t\right], \quad 0 \le t \le 1 \tag{2.2.9}$$

and

$$\widetilde{Y}_t = \mathbb{E}\left[X - \int_0^1 f(s)\, ds \,|\mathscr{F}_t\right]. \tag{2.2.10}$$

From "Martingale Representation Theorem" [9, page 182, thm 4.15] there exists $\bar{y} \in M^2(0,1;\, \mathbb{R}^{d\times k})$ such that

$$\widetilde{Y}_t = x(0) + \int_0^t \bar{y}(s)\, dW(s), \tag{2.2.11}$$

where $y(t) = \bar{y}(t) - g(t)$.

By substituting (2.2.9) into the original equation (2.2.8), we get;

$$\mathbb{E}\left[X - \int_t^1 f(s)\,ds\,|\mathscr{F}_t\right] + \int_t^1 f(s)\,ds + \underbrace{\int_t^1 \bar{y}(s)\,dW(s)}_{\tilde{Y}_1 - \tilde{Y}_t}. \tag{2.2.12}$$

To show existence of the unique pair (x, y) that solves (2.2.8), it is enough to show that the equation (2.2.12) is equal to X .

Using the linearity property of expectation and substituting the values of $\widetilde{Y}_1$ and $\widetilde{Y}_t$ into the equation (2.2.12), we get

$$\mathbb{E}\,[X\,|\mathscr{F}_t] - \mathbb{E}\left[\int_t^1 f(s)\,ds\,|\mathscr{F}_t\right] + \int_t^1 f(s)\,ds + X - \int_0^1 f(s)\,ds - \mathbb{E}\left[X - \int_0^1 f(s)\,ds\,|\mathscr{F}_t\right]. \tag{2.2.13}$$

Again using the linearity property of expectation and the basic property of integral, we get;

$$\mathbb{E}\,[X\,|\mathscr{F}_t] - \mathbb{E}\left[\int_t^1 f(s)\,ds\,|\mathscr{F}_t\right] + \int_t^1 f(s)\,ds + X - \int_0^t f(s)\,ds - \int_t^1 f(s)\,ds - \mathbb{E}\,[X\,|\mathscr{F}_t]$$
$$+\,\mathbb{E}\left[\int_0^t f(s)\,ds\,|\mathscr{F}_t\right] + \mathbb{E}\left[\int_t^1 f(s)\,ds\,|\mathscr{F}_t\right].$$

Since the terms, except "X", cancel each other, we 've shown that

$$\mathbb{E}\left[X - \int_t^1 f(s)\,ds\,|\mathscr{F}_t\right] + \int_t^1 f(s)\,ds + \int_t^1 \bar{y}(s)\,dW(s) = X.$$

As a result, there exists the pair (x, y) which solves (2.2.8). ■

We conclude from this lemma that this simplified version is nothing but martingale representation theorem. We see that Backward Stochastic Differential Equations are generalizations of conditional expectation.

Now, we consider the following equation:

$$x(t) + \int_t^1 f(s, y(s))\,ds + \int_t^1 [g(s) + y(s)]\,dW(s) = X. \tag{2.2.14}$$

Proposition 2.2.2 *[14, Proposition 2.2, page 57] Let* $X \in L^2(\Omega, \mathscr{F}_1, P;\, \mathbb{R}^d)$, $g \in M^2(0, 1;\, \mathbb{R}^{d\times k})$ *and* $f : \Omega \times (0, 1) \times \mathbb{R}^{d\times k} \to \mathbb{R}^d$ *be a mapping satisfying:*

- $f : \Omega \times (0, 1) \times \mathbb{R}^{d\times k} \to \mathbb{R}^d$ *is* $\mathscr{P} \otimes \mathscr{B}_{d\times k}/\mathscr{B}_d$ *measurable (*$\mathscr{P}$ *denotes the* σ*-algebra of progressively measurable subsets of* $\Omega \times (0, 1)$*),*

- $f(\cdot\,,0) \in M^2(0,1;\,\mathbb{R}^d)$,
- *There exists* $c > 0$ *such that* $|f(t,y_1) - f(t,y_2)| \leq c\,|y_1 - y_2|$ *for any* $y_1,\ y_2 \in \mathbb{R}^{d\times k}$,
- *Whenever* $y \in M^2(0,1;\,\mathbb{R}^{d\times k}),\ f(\cdot\,,y(\cdot)) \in M^2(0,1;\,\mathbb{R}^d)$.

Then, there exists a unique pair $(x,y) \in M^2(0,1;\,\mathbb{R}^d) \times M^2(0,1;\,\mathbb{R}^{d\times k})$ *which satisfies (2.2. 14).*

Proof. *Uniqueness.* Let (x_1,y_1) and (x_2,y_2) be two solutions such that

$$x_1(t) = X - \int_t^1 f(s,y_1(s))\,ds - \int_t^1 [g(s) + y_1(s)]\,dW_s$$

and

$$x_2(t) = X - \int_t^1 f(s,y_2(s))\,ds - \int_t^1 [g(s) + y_2(s)]\,dW_s.$$

It follows from Ito's formula applied to $|x_1(s) - x_2(s)|^2$ from $s = t$ to $s = 1$. First of all,

$$x_1(t) - x_2(t) = \int_t^1 \left[f(s,y_2(s)) - f(s,y_1(s))\right]\,ds + \int_t^1 \left[y_2(s) - y_1(s)\right]\,dW_s. \tag{2.2.15}$$

Applying Ito formula to $|x_1(s) - x_2(s)|^2$, we get

$$\begin{aligned}(x_1(1) - x_2(1))^2 = (x_1(t) - x_2(t))^2 &+ \int_t^1 2\,(x_1(s) - x_2(s))\;d\,(x_1 - x_2)_s \\ &+ \int_t^1 d\,\langle x_1 - x_2\rangle_s\,.\end{aligned} \tag{2.2.16}$$

Inserting values of $d\,(x_1 - x_2)_s$ and $d\,\langle x_1 - x_2\rangle_s$ and since x_1 and x_2 vanishes at point 1, equation (2.2.16) becomes;

$$\begin{aligned}0 = (x_1(t) - x_2(t))^2 &+ \int_t^1 2\,(x_1(s) - x_2(s))\,[f(s,y_2(s)) - f(s,y_1(s))]\;ds \\ + \int_t^1 2\,(x_1(s) - x_2(s))\,[y_2(s) - y_1(s)]\;dW_s &+ \int_t^1 (y_2(s) - y_1(s))^2\;ds.\end{aligned}$$

Rearranging the terms, the above equation is as follows;

$$\begin{aligned}(x_1(t) - x_2(t))^2 + \int_t^1 (y_1(s) - y_2(s))^2\;ds = &-2\int_t^1 (x_1(s) - x_2(s))\,[f(s,y_2(s)) - f(s,y_1(s))]\,ds \\ &- 2\int_t^1 (x_1(s) - x_2(s))\,[y_2(s) - y_1(s)]\;dW_s.\end{aligned}$$

Since $y_1 - y_2 \in M^2(0,1;\ \mathbb{R}^{d\times k})$, the stochastic integral of the above equation is P - integrable and has zero expectation. Taking expectation of the remaining terms, we get

$$\mathbb{E}\left[(x_1(t) - x_2(t))^2\right] + \int_t^1 \mathbb{E}\left[(y_1(s) - y_2(s))^2\right]\, ds =$$
$$-2\int_t^1 \mathbb{E}\left[(x_1(s) - x_2(s))\left[f(s, y_1(s)) - f(s, y_2(s))\right]\right]\, ds. \tag{2.2.17}$$

Remark 2.2.3 *From general mathematical concepts, we know the following inequalities;*

1. $|\langle X, Y\rangle| \leq |X|\,|Y|$,
2. $\mathbb{E}|XY| \leq \sqrt{\mathbb{E}|X|^2}\ \sqrt{\mathbb{E}|Y|^2}$,
3. $\left|2\int_t^1 f(s)g(s)\,ds\right| \leq \int_t^1 f^2(s)\,ds + \int_t^1 g^2(s)\,ds.$

Before studying equation (2.2.17), we first consider the following inequality:

$$\left|-2\,\mathbb{E}\int_t^1 \left[f(s, y_1(s)) - f(s, y_2(s))\right](x_1(s) - x_2(s))\ ds\right|$$
$$\leq 2\int_t^1 \mathbb{E}\ |f(s, y_1(s)) - f(s, y_2(s))|\ |x_1(s) - x_2(s)|\ ds$$

by Lipschitz continuity of f

$$\leq 2\int_t^1 \mathbb{E}\left[c\,|y_1(s) - y_2(s)|\ |x_1(s) - x_2(s)|\right]\ ds,$$

by Cauchy-Schwarz Inequality

$$\leq 2\int_t^1 \left(\mathbb{E}\left[\frac{1}{2}\,|y_1(s) - y_2(s)|^2\right]\right)^{\frac{1}{2}} \left(\mathbb{E}\left[2c^2\,|x_1(s) - x_2(s)|^2\right]\right)^{\frac{1}{2}}\ ds,$$

and by Remark 2.2.3 (3)

$$\leq \frac{1}{2}\int_t^1 \mathbb{E}\left[|y_1(s) - y_2(s)|^2\right]\ ds + 2c^2\int_t^1 \mathbb{E}\left[|x_1(s) - x_2(s)|^2\right]\ ds.$$

From the above result, equation (2.2.17) becomes,

$$\mathbb{E}\left[|x_1(t) - x_2(t)|^2\right] + \int_t^1 \mathbb{E}\left[|y_1(s) - y_2(s)|^2\right] ds$$
$$\leq \frac{1}{2}\int_t^1 \mathbb{E}\left[|y_1(s) - y_2(s)|^2\right] ds + 2c^2\int_t^1 \mathbb{E}\left[|x_1(s) - x_2(s)|^2\right] ds.$$

Taking the second integral to the left side of the above equation we get,

$$\mathbb{E}\left[|x_1(t) - x_2(t)|^2\right] + \frac{1}{2}\mathbb{E}\int_t^1 |y_1(s) - y_2(s)|^2\, ds \leq 2c^2\,\mathbb{E}\int_t^1 |x_1(s) - x_2(s)|^2\, ds. \quad (2.2.18)$$

Then,

$$\mathbb{E}\left[|x_1(t) - x_2(t)|^2\right] \leq 2c^2\,\mathbb{E}\int_t^1 |x_1(s) - x_2(s)|^2\, ds. \quad (2.2.19)$$

Remark 2.2.4 *One of the results of Gronwall's Inequality says that*

$$f(t) \leq K\int_{t_0}^t f(s)\, ds \Rightarrow f(t) = 0,$$

where f is continuous nonnegative function for $t_0 \leq t$ and K is any nonnegative constant.

It follows from Gronwall's Inequality applied to equation (2.2.19),

$$\mathbb{E}\left[|x_1(t) - x_2(t)|^2\right] = 0.$$

Expectation of a term is zero if the inside of this expectation is zero. Therefore,

$$|x_1(t) - x_2(t)|^2 = 0.$$

From the above equation, it is obvious that

$$x_1(t) = x_2(t).$$

Now, letting $x_1(t) = x_2(t)$ and using equation (2.2.18) we get,

$$\frac{1}{2}\mathbb{E}\int_t^1 |y_1(s) - y_2(s)|^2\, ds \leq 0.$$

The above integral is less than or equal to zero if and only if $y_1(t) = y_2(t)$. As a result,

$$(x_1, y_1) = (x_2, y_2).$$

We've completed the uniqueness part of the proof. Now, we can study the existence part.

Existence. Now, we define an approximating sequence by a kind of Picard iteration. Let $y_0(t) = 0$ and $\{(x_n(t), y_n(t));\ 0 \leq t \leq 1\}_{n\geq 1}$ be a sequence in $M^2(0,1;\ \mathbb{R}^d) \times M^2(0,1;\ \mathbb{R}^{d\times k})$

defined recursively by

$$
\begin{aligned}
&x_1(t) + \int_t^1 f(s, y_0(s))\, ds + \int_t^1 [g(s) + y_1(s)]\, dW_s = X,\\
&x_2(t) + \int_t^1 f(s, y_1(s))\, ds + \int_t^1 [g(s) + y_2(s)]\, dW_s = X,\\
&\vdots\\
&x_n(t) + \int_t^1 f(s, y_{n-1}(s))\, ds + \int_t^1 [g(s) + y_n(s)]\, dW_s = X.
\end{aligned}
$$

Using again Ito formula and the same inequalities as we did in uniqueness part with $K = 2c^2$ we get

$$
\begin{aligned}
&\mathbb{E}\left[|x_{n+1}(t) - x_n(t)|^2\right] + \mathbb{E}\int_t^1 |y_{n+1}(s) - y_n(s)|^2 ds\\
&\le K\,\mathbb{E}\int_t^1 |x_{n+1}(t) - x_n(t)|^2 ds + \frac{1}{2}\mathbb{E}\int_t^1 |y_n(s) - y_{n-1}(s)|^2 ds. \qquad (2.2.20)
\end{aligned}
$$

Now, define

$$
\begin{aligned}
u_n(t) &= \mathbb{E}\int_t^1 |x_n(s) - x_{n-1}(s)|^2\, ds,\\
v_n(t) &= \mathbb{E}\int_t^1 |y_n(s) - y_{n-1}(s)|^2\, ds,
\end{aligned}
$$

for $n \ge 1$ and $[x_0(t) = 0]$.

From the above definition, it is obvious that

$$
u_{n+1}(t) = \mathbb{E}\int_t^1 |x_{n+1}(s) - x_n(s)|^2\, ds. \qquad (2.2.21)
$$

Equation (2.2.21) implies that

$$
\begin{aligned}
\frac{d}{dt}\left(u_{n+1}(t)e^{Kt}\right) &= \frac{d}{dt}\left[\left(\int_t^1 \mathbb{E}|x_{n+1}(s) - x_n(s)|^2\, ds\right)e^{Kt}\right] \qquad (2.2.22)\\
&= \mathbb{E}|x_{n+1}(t) - x_n(t)|^2\, e^{Kt} + K\, e^{Kt}\left(\int_t^1 \mathbb{E}|x_{n+1}(s) - x_n(s)|^2\, ds\right). \qquad (2.2.23)
\end{aligned}
$$

Turning back to equation (2.2.20) and substituting values of $v_n(t)$, $v_{n+1}(t)$ and $u_{n+1}(t)$ we get,

$$\mathbb{E}\left[|x_{n+1}(t)-x_n(t)|^2\right]+\underbrace{\mathbb{E}\int_t^1 |y_{n+1}(s)-y_n(s)|^2 ds}_{v_{n+1}(t)}$$

$$\leq K\underbrace{\mathbb{E}\int_t^1 |x_{n+1}(t)-x_n(t)|^2 ds}_{u_{n+1}(t)}+\frac{1}{2}\underbrace{\mathbb{E}\int_t^1 |y_n(s)-y_{n-1}(s)|^2 ds}_{v_n(t)} \tag{2.2.24}$$

$$\therefore\ \mathbb{E}\left[|x_{n+1}(t)-x_n(t)|^2\right]+v_{n+1}(t)\leq K\,u_{n+1}(t)+\frac{1}{2}v_n(t). \tag{2.2.25}$$

Multiplying both sides by e^{Kt} we get

$$\begin{aligned}
&e^{Kt}\mathbb{E}\left[|x_{n+1}(t)-x_n(t)|^2\right]+e^{Kt}v_{n+1}(t)\leq Ke^{Kt}\,u_{n+1}(t)+\frac{1}{2}e^{Kt}v_n(t)\\
&=-Ke^{Kt}\,u_{n+1}(t)+e^{Kt}\mathbb{E}\left[|x_{n+1}(t)-x_n(t)|^2\right]+e^{Kt}v_{n+1}(t)\leq\frac{1}{2}e^{Kt}v_n(t)\\
&=-\frac{d}{dt}\left(u_{n+1}(t)e^{Kt}\right)+e^{Kt}v_{n+1}(t)\leq\frac{1}{2}e^{Kt}v_n(t).
\end{aligned}$$

Integrating from t to 1 we obtain,

$$\begin{aligned}
&-\int_t^1\left(\frac{d}{dt}\left[u_{n+1}(s)e^{Ks}\right]\right)ds+\int_t^1 e^{Ks}v_{n+1}(s)\,ds\leq\frac{1}{2}\int_t^1 e^{Ks}v_n(s)\,ds\\
&=-\left(u_{n+1}(1)e^{K}-u_{n+1}(t)e^{Kt}\right)+\int_t^1 e^{Ks}v_{n+1}(s)\,ds\leq\frac{1}{2}\int_t^1 e^{Ks}v_n(s)\,ds,
\end{aligned}$$

where $u_{n+1}(1)=0$

$$=u_{n+1}(t)+\int_t^1 e^{K(s-t)}v_{n+1}(s)\,ds\leq\frac{1}{2}\int_t^1 e^{K(s-t)}v_n(s)\,ds.$$

It follows in particular that

$$\int_0^1 e^{Ks}v_{n+1}(s)\,ds\leq\frac{1}{2}\int_0^1 e^{Ks}v_n(s)\,ds,$$

and

$$\begin{aligned}
\left(\frac{1}{2}\right)^n\int_0^1 e^{Ks}v_1(s)\,ds&\leq\left(\frac{1}{2}\right)^n sup_{s\leq 1}|v_1(t)|\int_0^1 e^{Ks}\,ds\\
&\leq\left(\frac{1}{2}\right)^n sup|v_1(t)|\,e^{K},
\end{aligned}$$

where $\quad sup|v_1(t)|=\mathbb{E}\int_0^1 |y_1(t)|^2\,dt.$

Then, $\quad u_{n+1}(0)\leq\left(\frac{1}{2}\right)^n\bar{c}\,e^{K},\quad$ where $\quad\bar{c}=sup|v_1(t)|.$

From equation (2.2.25),

$$
\begin{aligned}
v_{n+1}(t) &\le K\,u_{n+1}(t) + \frac{1}{2}v_n(t),\\
v_{n+1}(0) &\le K\,u_{n+1}(0) + \frac{1}{2}v_n(0),\\
v_{n+1}(0) &\le 2^{-n}K\bar{c}\,e^K\,u_1(0) + \frac{1}{2}v_n(0).
\end{aligned}
$$

It follows immediately,

$$v_{n+1}(0) \le 2^{-n}\left(n\bar{c}Ke^K + v_1(0)\right).$$

These imply

$$u_n = \mathbb{E}\int_0^1 |x_n(s) - x_{n-1}(s)|^2\,ds = |x^n - x^{n-1}|^2 \le 2^{-n}\bar{c}Ke^K < \infty, \tag{2.2.26}$$

$$v_n = \mathbb{E}\int_0^1 |y_n(s) - y_{n-1}(s)|^2\,ds = |y^n - y^{n-1}|^2 \le 2^{-n}\bar{c}Ke^K < \infty. \tag{2.2.27}$$

Remember that $\sum_{n=1}^{\infty}\|x^n - x^{n-1}\|_{L^2} < \infty$ implies that x^n is a Cauchy sequence. This, (2.2.26) and (2.2.27) imply that x_n and y_n are Cauchy sequences in $M^2(0,1;\mathbb{R}^d)$ and $M^2(0,1;\mathbb{R}^{d\times k})$ respectively and hence they must converge. The pair (x,y) defined by $x \doteq \lim_{n\to\infty} x_n$ and $y \doteq \lim_{n\to\infty} y_n$ is the desired solution of the equation (2.2.14). ■

We can now study the equation

$$x(t) + \int_t^1 f(s, x(s), y(s))\,ds + \int_t^1 \left[g(s, x(s)) + y(s)\right]\,dW(s) = X. \tag{2.2.28}$$

Theorem 2.2.5 *[14, Theorem 3.1, page 58] Given* $X \in L^2(\Omega, \mathscr{F}_1, P;\, \mathbb{R}^d)$, $f : \Omega \times (0,1) \times \mathbb{R}^d \times \mathbb{R}^{d\times k} \to \mathbb{R}^d$ *and* $g : \Omega \times (0,1) \times \mathbb{R}^d \to \mathbb{R}^{d\times k}$ *with the properties that*

- $f : \Omega \times (0,1) \times \mathbb{R}^d \times \mathbb{R}^{d\times k} \to \mathbb{R}^d$ *is* $\mathscr{P} \otimes \mathscr{B}_d \otimes \mathscr{B}_{dk}/\mathscr{B}_d$ *measurable and* $g : \Omega \times (0,1) \times \mathbb{R}^d \to \mathbb{R}^{d\times k}$ *is* $\mathscr{P} \otimes \mathscr{B}_d/\mathscr{B}_{dk}$ *measurable,*
- $f(\cdot, 0, 0) \in M^2(0,1;\, \mathbb{R}^d)$, $\quad g(\cdot, 0, 0) \in M^2(0,1;\, \mathbb{R}^{d\times k})$,
- *There exists* $c > 0$ *such that* $|f(t, x_1, y_1) - f(t, x_2, y_2)| \le c\,(|x_1 - x_2| + |y_1 - y_2|)$ $\forall\, x, x_1, x_2 \in \mathbb{R}^d,\ \forall\, y, y_1, y_2 \in \mathbb{R}^{d\times k}$,
- $|g(t, x_1) - g(t, x_2)| \le c\,|x_1 - x_2|$, $\quad \forall\, x, x_1, x_2 \in \mathbb{R}^d,\ \forall\, y, y_1, y_2 \in \mathbb{R}^{d\times k}$,
- *Whenever* $x \in M^2(0,1;\, \mathbb{R}^d)$, $f(\cdot, x(\cdot), y(\cdot)) \in M^2(0,1;\, \mathbb{R}^d)$,

- *Whenever* $y \in M^2(0,1;\, \mathbb{R}^{d\times k})$, $g(\cdot, x(\cdot), y(\cdot)) \in M^2(0,1;\, \mathbb{R}^{d\times k})$.

Then, there exists a unique pair $(x,y) \in M^2(0,1;\, \mathbb{R}^{d}) \times M^2(0,1;\, \mathbb{R}^{d\times k})$ *which solves equation (2.2.28).*

Proof. Again, we have two steps, namely existence and uniqueness, to show existence of unique pair (x,y) which solves equation (2.2.28). We can start with uniqueness part.

Uniqueness. We follow the similar argument as that in Proposition (2.2.2). Let (x_1,y_1) and (x_2,y_2) be two solutions in $M^2(0,1;\, \mathbb{R}^{d}) \times M^2(0,1;\, \mathbb{R}^{d\times k})$ such that

$$x_1(t) = X - \int_t^1 f(s, x_1(s), y_1(s))\, ds - \int_t^1 \left[g(s, x_1(s)) + y_1(s)\right]\, dW_s,$$

$$x_2(t) = X - \int_t^1 f(s, x_2(s), y_2(s))\, ds - \int_t^1 \left[g(s, x_2(s)) + y_2(s)\right]\, dW_s,$$

and

$$\begin{aligned} x_1(t) - x_2(t) = &-\left[\int_t^1 \left(f(s, x_1(s), y_1(s)) - f(s, x_2(s), y_2(s))\right)\, ds\right] \\ &+ \int_t^1 \left(g(s, x_1(s)) - g(s, x_2(s))\right)\, dW_s + \int_t^1 \left(y_1(s) - y_2(s)\right)\, dW_s. \end{aligned}$$

We try to show that $(x_1,y_1) = (x_2,y_2)$. By using Ito formula applied to $|x_1(t) - x_2(t)|^2$ we get,

$$(x_1(1) - x_2(1))^2 = (x_1(t) - x_2(t))^2 + \int_t^1 2\,(x_1(s) - x_2(s))\, d(x_1, x_2)s + \int_t^1 d\,\langle x_1, x_2\rangle_s\,;$$

Hence, $x_1(t)$ and $x_2(t)$ vanish at point 0. Inserting values of $d(x_1,x_2)s$ and $d\,\langle x_1, x_2\rangle_s$ into the above equation we get,

$$\begin{aligned} 0 = &(x_1(t) - x_2(t))^2 + \int_t^1 2\,(x_1(s) - x_2(s))\left[f(s, x_2(s), y_2(s)) - f(s, x_1(s), y_1(s))\right]\, ds \\ &+ \int_t^1 2\,(x_1(s) - x_2(s))\left[g(s, x_2(s)) - g(s, x_1(s))\right] dW_s \\ &+ \int_t^1 2\,(x_1(s) - x_2(s))\left[y_2(s) - y_1(s)\right]\, dW_s \\ &+ \int_t^1 \left[g(s, x_2(s)) - g(s, x_1(s))\right]^2\, ds + \int_t^1 (y_2(s) - y_1(s))^2\, ds. \end{aligned}$$

Rearranging the terms;

$$
\begin{aligned}
&(x_1(t) - x_2(t))^2 + \int_t^1 (y_1(s) - y_2(s))^2\, ds = \\
&-2\int_t^1 (x_1(s) - x_2(s))\left[f(s, x_1(s), y_1(s)) - f(s, x_2(s), y_2(s))\right]\, ds \\
&-2\int_t^1 (x_1(s) - x_2(s))\left[g(s, x_1(s)) - g(s, x_2(s))\right] dW_s \\
&-2\int_t^1 (x_1(s) - x_2(s))\left[y_1(s) - y_2(s)\right]\, dW_s \\
&-\int_t^1 \left[g(s, x_1(s)) - g(s, x_2(s))\right]^2\, ds.
\end{aligned}
$$

Since $y_1 - y_2 \in M^2(0, 1; \mathbb{R}^{d\times k})$, the above stochastic integral is P -integrable and has zero expectation. So, taking expectation of the remaining terms we have,

$$
\mathbb{E}|x_1(t) - x_2(t)|^2 + \mathbb{E}\int_t^1 |y_1(s) - y_2(s)|^2 ds = \tag{2.2.29}
$$

$$
-2\mathbb{E}\int_t^1 \left[f(s, x_1(s), y_1(s)) - f(s, x_2(s), y_2(s))\right](x_1(s) - x_2(s))ds \tag{2.2.30}
$$

$$
-\mathbb{E}\int_t^1 \left[g(s, x_1(s)) - g(s, x_2(s))\right]^2\, ds \tag{2.2.31}
$$

$$
-2\mathbb{E}\int_t^1 (x_1(s) - x_2(s))\left[g(s, x_1(s)) - g(s, x_2(s))\right] ds. \tag{2.2.32}
$$

Now, we consider the right side of the above equation as follows:

$$
\begin{aligned}
&|(2.2.30) + (2.2.31) + (2.2.32)| \\
&\leq\ 2\mathbb{E}\int_t^1 |f(s, x_1(s), y_1(s)) - f(s, x_2(s), y_2(s))|\,|(x_1(s) - x_2(s))|\, ds \\
&\quad + 2\mathbb{E}\int_t^1 |g(s, x_1(s)) - g(s, x_2(s))|\,|(x_1(s) - x_2(s)|\, ds \\
&\quad + \mathbb{E}\int_t^1 |g(s, x_1(s)) - g(s, x_2(s))|^2\, ds \\
&\leq 2\int_t^1 \mathbb{E}\left[c(|x_1(s) - x_2(s)| + |y_1(s) - y_2(s)|)|x_1(s) - x_2(s)|\right]\, ds \\
&\quad + 2\int_t^1 \mathbb{E}\left[c|x_1(s) - x_2(s)|\,|x_1(s) - x_2(s)|\right]\, ds
\end{aligned}
$$

by Lipschitz continuity of f and g

$$
\begin{aligned}
&+\int_t^1 \mathbb{E}\left[c^2|x_1(s)-x_2(s)|^2\right] ds \\
&\leq 2\int_t^1 \mathbb{E}\left[c(|x_1(s)-x_2(s)|^2\right] ds + 2\int_t^1 \mathbb{E}\left[c|x_1(s)-x_2(s)||y_1(s)-y_2(s)|\right] ds \\
&+2\int_t^1 \mathbb{E}\left[c|x_1(s)-x_2(s)|^2\right] ds \\
&+\int_t^1 \mathbb{E}\left[c^2|x_1(s)-x_2(s)|^2\right] ds \\
&\leq \bar{c}\int_t^1 \mathbb{E}|x_1(s)-x_2(s)|^2\, ds + 2\int_t^1 \mathbb{E}\left[c|x_1(s)-x_2(s)||y_1(s)-y_2(s)|\right] ds,
\end{aligned}
$$

where $\bar{c} = 4c + c^2$.

By Cauchy-Schwarz Inequality,

$$
\begin{aligned}
|(2.2.30)+(2.2.31)+(2.2.32)| \leq\ & \bar{c}\int_t^1 \mathbb{E}|x_1(s)-x_2(s)|^2 ds \\
&+2\int_t^1 \left(\mathbb{E}\frac{1}{2}|y_1(s)-y_2(s)|^2\right)^2 \left(\mathbb{E}2c^2|x_1(s)-x_2(s)|^2\right)^2 ds \\
\leq\ & \bar{c}\int_t^1 \mathbb{E}|x_1(s)-x_2(s)|^2 ds + \frac{1}{2}\int_t^1 \mathbb{E}|y_1(s)-y_2(s)|^2\, ds
\end{aligned}
$$

by Remark (2.2.3)(3)

$$
\begin{aligned}
&+2c^2\int_t^1 \mathbb{E}|x_1(s)-x_2(s)|^2\, ds \\
&\leq \tilde{c}\,\mathbb{E}\int_t^1 |x_1(s)-x_2(s)|^2\, ds + \frac{1}{2}\int_t^1 \mathbb{E}|y_1(s)-y_2(s)|^2\, ds,
\end{aligned}
$$

where $\tilde{c} = \bar{c} + 2c^2$.

Therefore, we have the following inequality:

$$
\begin{aligned}
&\mathbb{E}|x_1(s)-x_2(s)|^2 + \mathbb{E}\int_t^1 |y_1(s)-y_2(s)|^2\, ds \\
&\leq \tilde{c}\,\mathbb{E}\int_t^1 |x_1(s)-x_2(s)|^2\, ds + \frac{1}{2}\int_t^1 \mathbb{E}|y_1(s)-y_2(s)|^2\, ds.
\end{aligned}
$$

Taking the last term to the left side of the equation we get,

$$
\mathbb{E}|x_1(s)-x_2(s)|^2 + \frac{1}{2}\int_t^1 \mathbb{E}|y_1(s)-y_2(s)|^2\, ds \leq \tilde{c}\,\mathbb{E}\int_t^1 |x_1(s)-x_2(s)|^2\, ds. \tag{2.2.33}
$$

As a result,

$$\mathbb{E}|x_1(s) - x_2(s)|^2 \leq \tilde{c}\,\mathbb{E}\int_t^1 |x_1(s) - x_2(s)|^2\,ds.$$

From Gronwall's Inequality [Remark 2.2.4], $|x_1(t) - x_2(t)| = 0$. Therefore,

$$x_1(t) = x_2(t).$$

Letting $x_1(t) = x_2(t)$ in the equation (2.2.33) we get

$$\frac{1}{2}\int_t^1 \mathbb{E}|y_1(s) - y_2(s)|^2\,ds \leq 0.$$

The above integral is less than or equal to zero if and only if $y_1(t) = y_2(t)$. As a rsult,

$$(x_1, y_1) = (x_2, y_2)$$

We have completed the uniqueness part of the proof. Now, we can study the existence part.

Existence. Now, we define an approximating sequence using a Picard type iteration with the help of Proposition 2.2.2. Let $x_0(t) = 0$ and $\{x_n(t), y_n(t)\};\ 0 \leq t \leq 1\}_{n\geq 1}$ be sequence in $M^2(0,1;\,\mathbb{R}^d) \times M^2(0,1;\,\mathbb{R}^{d\times k})$ defined recursively by

$$X = x_n(t) + \int_t^1 f(s, x_{n-1}(s), y_n(s))\,ds + \int_t^1 \left[g(s, x_{n-1}(s)) + y_n(s)\right]\,dW_s, \quad 0 \leq t \leq 1. \quad (2.2.34)$$

With the same steps that in uniqueness part, namely from equation (2.2.33), we obtain,

$$\begin{aligned} &\mathbb{E}|x_{n+1}(t) - x_n(t)|^2 + \frac{1}{2}\,\mathbb{E}\int_t^1 |y_{n+1}(s) - y_n(s)|^2\,ds \\ &\leq c\left(\mathbb{E}\int_t^1 |x_{n+1}(s) - x_n(s)|^2\,ds + \mathbb{E}\int_t^1 |x_n(s) - x_{n-1}(s)|^2\,ds\right). \end{aligned} \quad (2.2.35)$$

Now, define

$$u_n(t) = \mathbb{E}\int_t^1 |x_n(s) - x_{n-1}(s)|^2\,ds.$$

So,

$$u_{n+1}(t) = \mathbb{E}\int_t^1 |x_{n+1}(s) - x_n(s)|^2\,ds.$$

Taking derivative of $u_{n+1}(t)$ with respect to t we get

$$-\frac{d}{dt}u_{n+1}(t) = \mathbb{E}|x_{n+1}(t) - x_n(t)|^2.$$

From equation (2.2.35), we have the following inequality,

$$\underbrace{\mathbb{E}|x_{n+1}(t)-x_n(t)|^2}_{-\frac{d}{dt}u_{n+1}(t)} \leq c\left(\underbrace{\mathbb{E}\int_t^1 |x_{n+1}(s)-x_n(s)|^2\,ds}_{u_{n+1}(t)} + \underbrace{\mathbb{E}\int_t^1 |x_n(s)-x_{n-1}(s)|^2\,ds}_{u_n(t)}\right). \quad (2.2.36)$$

Therefore,

$$-\frac{d}{dt}u_{n+1}(t) - c\,u_{n+1}(t) \leq c\,u_n(t), \quad \text{where } u_{n+1}(1)=0. \quad (2.2.37)$$

Multiplying both sides of the equation (2.2.37) by e^{Kt} we get,

$$-\frac{d}{dt}u_{n+1}(t)\,e^{Kt} - c\,u_{n+1}(t)\,e^{Kt} \leq c\,u_n(t)\,e^{Kt}.$$

Now, integrating both sides from t to 1 we get,

$$\begin{aligned}
&-\int_t^1\left[\frac{d}{ds}u_{n+1}(s)\,e^{Ks}\right]ds - c\int_t^1 u_{n+1}(s)\,e^{Ks}\,ds \leq c\int_t^1 u_n(s)\,e^{Ks}\,ds \\
&= -\left(u_{n+1}(1)e^{K} - u_{n+1}(t)\,e^{Kt}\right) - c\int_t^1 u_{n+1}(s)\,e^{Ks}\,ds \leq c\int_t^1 u_n(s)\,e^{Ks}\,ds,
\end{aligned}$$

where $u_{n+1}(1)=0$.

Multiplying both sides by e^{-Kt} we get,

$$u_{n+1}(t) - c\int_t^1 u_{n+1}(s)\,e^{K(s-t)}\,ds \leq c\int_t^1 u_n(s)\,e^{K(s-t)}\,ds. \quad (2.2.38)$$

From equation (2.2.38) we obtain,

$$\begin{aligned}
u_{n+1}(t) \leq c\int_t^1 u_n(s)\,e^{K(s-t)}\,ds &\Rightarrow u_{n+1}(t) \leq c\int_t^1 u_n(s)\,e^{Ks}\,ds \\
&\Rightarrow u_{n+1}(0) \leq c^n\,u_1(0)\int_0^1 e^{K}s\,ds \\
&\Rightarrow u_{n+1}(0) \leq \frac{\left(ce^{K}\right)^n}{n!}\,u_1(0),
\end{aligned}$$

where

$$\begin{aligned}
u_n &= \mathbb{E}\int_0^1 |x_n(s)-x_{n-1}(s)|^2\,ds \\
&= |x^n - x^{n-1}|^2 \\
&\leq \frac{\left(ce^{K}\right)^n}{n!}.
\end{aligned}$$

This, together with the equation (2.2.35), implies that $\{x_n\}$ is a Cauchy sequence in $M^2(0,1;\mathbb{R}^d)$ and $\{y_n\}$ is a Cauchy sequence in $M^2(0,1;\mathbb{R}^{d\times k})$.

Then, from equation(2.2.34), $\{x_n\}$ converges also in $L^2(\Omega;\ C(0,1;\ \mathbb{R}^d))$. It then follows from (2.2.34) that

$$x = \lim_{n\to\infty} x_n$$

and

$$y = \lim_{n\to\infty} y_n.$$

As a result, we obtain the pair (x, y) that solves the equation (2.2.28). ∎

CHAPTER 3

VARIANTS OF THE ITO FORMULA

3.1 Introduction

While studying [14] we ended up needing the following derivations and formulas. We have not come across these formulas anywhere in the prior literature and therefore found it worth recording them here.

Let

$$W = \begin{bmatrix} W_1 \\ W_2 \\ \vdots \\ W_d \end{bmatrix}$$

be a d dimensional Brownian motion.

Let

$$X = \int G(s)\, dW \tag{3.1.1}$$

and

$$Y = \int F(s)\, dW, \tag{3.1.2}$$

where

$$G(s) \doteq \begin{bmatrix} g_{1,1} & g_{1,2} & \cdots & g_{1,d} \\ g_{2,1} & g_{2,2} & \cdots & g_{2,d} \end{bmatrix}, \; F(s) \doteq \begin{bmatrix} f_{1,1} & f_{1,2} & \cdots & f_{1,d} \\ f_{2,1} & f_{2,2} & \cdots & f_{2,d} \end{bmatrix}$$

X and Y are two dimensional processes. The first components of X and Y are as follows,

$$\sum_{j=1}^{d} \int_0^t g_{1,j}\, dW_j\,, \quad \sum_{j=1}^{d} \int_0^t f_{1,j}\, dW_j\,.$$

The second components of X and Y can be expressed with similar sums.

Define $Z = X + Y$. Suppose $F : \mathbb{R}^2 \to \mathbb{R}$ is a smooth function. We would like to derive simple expressions for the application of Ito's formula to $F(Z)$.

Let DF denote the gradient of F and HF the Hessian of F. For two processes A and B let $V(A, B)$ denote the matrix of cross variations of A and B. That is:

$$V_{i,j} = \left[A_i, B_j\right]$$

and

$$V = \begin{bmatrix} [A_1, B_1] & [A_1, B_2] \\ [B_2, A_1] & [A_2, B_2] \end{bmatrix}.$$

Here $\left[A_i, B_j\right]$ denotes the cross variation between the one dimensional processes A_i and B_j.

Define

$$V(C) \doteq V(C, C).$$

$V(C)$ can be referred to as the variation matrix of the process C.

Let us now go back to the process $F(Z)$. Ito's formula in terms of V, HF and DF is

$$F(Z_t) = F(Z_0) + \int_0^t \langle DF, dZ\rangle + \frac{1}{2}\int_0^t \langle HF, dV(Z)\rangle . \tag{3.1.3}$$

$\langle a, b\rangle$ denotes the inner product. If the arguments a and b are matrices, they are treated as vectors and their inner product is computed in the usual way.

For example,

$$\left\langle \begin{bmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{bmatrix}, \begin{bmatrix} b_{1,1} & b_{1,2} \\ b_{2,1} & b_{2,2} \end{bmatrix} \right\rangle = \sum_{i,j} a_{i,j} b_{i,j}.$$

Remark 3.1.1 *Note that*

$$\sum_{i,j} a_{i,j} b_{i,j} = \sum_i \sum_j a_{i,j} (b^t)_{j,i} = tr(ab^t).$$

where t and tr denote the transpose and the trace operators. Therefore, another way to write (3.1.3) is

$$F(Z_t) = F(Z_0) + \int_0^t \langle DF, dZ \rangle + \frac{1}{2} \int_0^t tr((HF)^t dV(Z)) \tag{3.1.4}$$

We will not be using this notation in what follows.

Our goal is to write (3.1.3) in terms of X and Y. For this purpose it enough to write $V(Z)$ in terms of $V(X)$, $V(Y)$ and $V(X, Y)$.

Proposition 3.1.2 *For the process X and Y given in (3.1.1) and (3.1.2)*

$$\frac{DV}{ds}(X, Y) = GF^t.$$

Note that this immediately implies:

Proposition 3.1.3

$$V(Z) = V(X) + V(Y) + 2V(X, Y). \tag{3.1.5}$$

Remark 3.1.4 *Note that Proposition 3.1.3 is a generalization of the well known formula for the one dimensional cross variation:*

$$[X + Y, X + Y] = [X, X] + 2 [X, Y] + [Y, Y]. \tag{3.1.6}$$

The key difference between this and (3.1.5) is that (3.1.6) is a scalar equation whereas (3.1.5) is a matrix equation.

(3.1.5) allows us to rewrite (3.1.4) as

$$\begin{aligned} F(Z_t) = F(Z_0) + \int_0^t \langle DF, dZ \rangle + \frac{1}{2} \int_0^t tr((HF)^t dV(X)) \\ + \frac{1}{2} \int_0^t tr((HF)^t dV(Y)) + \int_0^t tr((HF)^t dV(X, Y)). \end{aligned}$$

This fact is used in the proof of Theorem 3.1 in [14].

CHAPTER 4

HEDGING CONTINGENT CLAIMS WITH CONSTRAINED PORTFOLIOS

4.1 Introduction

Backward stochastic differential equations naturally come up in finance. In this section we would like to review an instance of this phenomenon that is reported in [5]. In this introductory section we very briefly introduce the results in this paper. The following sections will reintroduce the notation and review the results derived in the paper. Everything in the following sections except for the last one are from [5].

Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space that supports a d dimensional Brownian motion W. Let $\sigma : \mathbb{R}^d \to \mathbb{R}^{d\times d}$ be a progressively measurable process with respect to $\mathscr{F}_t$, the $\mathbb{P}$ augmentation of the filtration $\mathscr{F}_t^W \doteq \sigma(W(s), s \leq t)$. $\sigma\sigma'$ is assumed to be strictly positive definite, i.e.,

$$\langle \sigma\xi, \sigma\xi \rangle \geq \epsilon \|\xi\|^2 \tag{4.1.1}$$

for some $\epsilon > 0$. Let $r : \mathbb{R} \to \mathbb{R}$, and $b : \mathbb{R} \to \mathbb{R}^d$ to be two progressively measurable processes. b will be the appreciation rates of the stocks and r will be the instantaneous interest rate. σ, r and b are all assumed to be uniformly bounded.

The market model used in [5] is the following standard continuous time model. We focus on the finite time horizon $[0, T]$. There are d risky assets whose prices are modeled with

$$P(t) \doteq P(0) + \int_0^t P(s) \cdot b(s)ds + \int_0^t P(s) \cdot (\sigma(s)dW(s))$$

where $\cdot$ denotes pointwise multiplication. The price process P is d dimensional with components P_i, $i = 1, 2, 3, ..., d$ that satisfy $P_i(0) > 0$. In addition we have the riskless security whose price process is

$$P_0(t) \doteq e^{\int_0^t r(s)ds}$$

The strict positive definiteness (4.1.1) of σ allows one to define

$$\theta(t) \doteq \sigma^{-1}[b(t) - r(t)],$$

where we use the convention that a vector plus a constant means that the constant is added to each of the components of the vector (as in matlab). The discount process is

$$\gamma_0(t) \doteq \frac{1}{P_0(t)}$$

4.1.1 Admissible portfolios

A portfolio invested in this market can be represented by an $\mathbb{R}^d$ valued $\mathscr{F}_t$ progressively measurable process π. [5] also stipulates $\int_0^T \|\pi(t)\|^2 dt < \infty$ almost surely.

A nonnegative, nondecreasing $\mathscr{F}_t$ progressively measurable process c with RCLL paths, $c(0) = 0$ and $c(T) < \infty$ almost surely is called a "consumption process."

The wealth of an investor who uses the (π, c) pair for her investment and consumption will have the following dynamics:

$$\begin{aligned} X(t) = X(0) &+ \int_0^t X(s) \langle \pi(s), (b(s)ds + \sigma(s)dW(s)) \rangle \\ &+ \int_0^t X(s)(1 - \sum_{i=1}^d \pi_i(s))ds - c(t) \end{aligned} \tag{4.1.2}$$

Define

$$W_0(t) \doteq W(t) + \int_0^t \theta(s)ds$$

(4.1.2) in terms of W_0 is

$$X(t) = X(0) + \int_0^t X(s) \left(r(s)ds + \langle \pi(s), \sigma(s)dW_0(s) \rangle\right) - c(t).$$

(π, c) is called admissible if $X^{x,\pi,c}(t) \geq 0$ for all $t \in [0, T]$. Admissibility of (π, c) will be denoted with $(\pi, c) \in \mathscr{A}_0(x)$.

Define

$$Z_0(t) \doteq \exp\left(-\int_0^t \langle \theta(s), dW(s) \rangle - \frac{1}{2}\int_0^t \|\theta(s)\|^2 ds\right)$$

and define the measure

$$P^0(A) \doteq E[Z_0(T)1_A].$$

W_0 is a standard Brownian motion under P^0.

A contingent claim in this market is an $\mathscr{F}_T$ measurable random variable B that satisfies

$$0 < \mathbb{E}^0[\gamma_0(T)B] < \infty.$$

Now suppose that we would like to replicate this contingent claim by investing in the market. [5] assumes that there is the following restriction on this problem: the portfolio at any time has to take values in the convex set K, i.e., we require that $\pi(s,\omega) \in K$ $l \otimes P$ almost surely. If an admissible portfolio π satisfies this constraint, we will write $\pi \in \mathscr{A}_0'(x)$. We can represent this problem as the following constrained control problem:

$$h(0) \doteq \inf\{x \in (0,\infty) : \exists(\pi, c) \in \mathscr{A}'(x) \text{ such that } X^{x,\pi,c}(T) \geq B\} \tag{4.1.3}$$

Note that h can be thought of as the value function of this control problem. A natural question is: what is the state of this control problem? Note that the given "initial condition" is actually the contingent claim B, h is a function of this "final state." The cost to be optimized is the initial state x of the controlled process. Therefore, the state of this control problem evolves backward in time. The chief accomplishment of [5] is to find an optimal control problem that is dual to (4.1.3) that is in a completely standard form. By "a standard form" we mean in particular the following: 1) dynamics evolving forward in time 2) the cost to be optimized is in the form of an accumulated running cost based on the chosen control and and a final cost. A review of this development is in the following subsection.

An important remaining problem is the solution of the dual problem. Section 4.7 proposes a simple algorithm that only discretizes time in the case of constant volatility and interest rate.

4.2 Preliminaries

Let $\mathscr{M}$ be a financial market that consists of one bond and (d) stocks. The prices of which are as follows:

$$dP_0(t) = P_0(t)\, r(t)\, dt, \quad P_0(0) = 1 \tag{4.2.1}$$

and

$$dP_i(t) = P_i(t)\left[b_i(t)\,dt + \sum_{j=1}^{d}\sigma_{ij}(t)\,dW^{(j)}(t)\right], \quad P_i(0) = p_i \in (0,\infty), \;\; i = 1,\ldots,d. \quad (4.2.2)$$

Here;

- $W = (W^{(1)},\ldots,W^{(d)})^*$ is a standard Brownian Motion in $\mathscr{R}^d$, defined on a complete probability space $(\Omega, \mathscr{F}, \mathbb{P})$,
- The $\mathbb{P}$-augmentation of the filtration $\mathscr{F}_t^W = \sigma(W(s);\; 0 \le s \le t)$ generated by W shall be denoted by $\mathscr{F}_t$,
- The process $r(t)$-scalar interest rate, $b(t) = (b_1(t),\ldots,b_d(t))^*$-vector of appreciation rates and $\sigma(t) = \{\sigma_{ij}(t)\}_{1\le i,\, j\le d}$ -volatility matrix are progressively measurable with respect to $\{\mathscr{F}_t\}$ and bounded uniformly in $(t,w) \in [0,T]\times\Omega$.

We introduce also some more processes;

1. $\theta(t)$: market price of risk ≡ relative risk or sharpe ratio:

$$\theta\,(t) \doteq \sigma^{-1}(t)[b(t) - r(t)].$$

2. The exponential martingale:

$$Z_0(t) \doteq \exp\left[-\int_0^t \theta^*(s)\,dW(s) \;-\frac{1}{2}\int_0^t \|\theta(s)\|^2\,ds\right].$$

3. The discount process:

$$\gamma_0(t) \doteq \exp\left[-\int_0^t r(s)\,ds\right].$$

4.3 Portfolio, Consumption and Wealth Processes

In this section, we introduce some basic concepts and terms that are used throughout the paper. We know that there are (d) stocks and one bond in the market $\mathscr{M}$. Now, it is important to decide for an investor at any time $t \in [0,T]$ that what proportion $\pi_i(t)$ of his wealth $X(t)$ to invest in the ith stock $(1 \le i \le d)$ and what amount of money $c(t+h) - c(t) \ge 0$ to withdraw for consumption during $(t, t+h], h > 0$. The amount $X(t)[1-\Sigma_{i=1}^{d}\pi_i(t)]$ is invested in the bond. Therefore, the wealth process $X(t)$ is as follows:

$$\begin{aligned} dX(t) &= \sum_{i=1}^{d} \pi_i(t)\, X(t) \left[b_i(t)\, dt + \sum_{j=1}^{d} \sigma_{ij}(t)\, dW^j(t) \right] \\ &\quad + \left[1 - \sum_{i=1}^{d} \pi_i(t) \right] X(t)\, r(t)\, dt - dc(t) \\ &= r(t)\, X(t)\, dt - dc(t) + X(t)\, \pi^*(t)\, \sigma(t)\, dW_0(t), \quad X(0) = x > 0, \end{aligned}$$

where

$$W_0(t) \doteq W(t) + \int_0^t \theta(s)\, ds, \quad 0 \le t \le T.$$

4.3.1 Definitions

1. *Portfolio Process* : An $\mathscr{R}^d$-valued, $\{\mathscr{F}_t\}$-progressively measurable process $\pi = \{\pi(t),\ 0 \le t \le T\}$ with $\int_0^T \|\pi(t)\|^2\, dt < \infty$ a.s..

2. *Consumption Process* : A nonnegative, nondecreasing, $\mathscr{F}_t$-progressively measurable process $c = \{c(t),\ 0 \le t \le T\}$ with RCLL paths, $c(0) = 0$ and $c(T) < \infty$ a.s..

3. *Wealth Process* : The solution $X \equiv X^{x,\pi,c}$ corresponding to the portfolio-consumption pair (π, c) and initial capital $x \in (0, \infty)$.

4. A portfolio-consumption process pair (π, c) is called *admissible* for the initial capital $x \in (0, \infty)$, if the corresponding wealth process is always nonnegative. The set of admissible pairs (π, c) is denoted by $\mathscr{A}_0(x)$.

5. A *contingent claim* is a nonnegative, $\mathscr{F}_T$-measurable random variable B satisfying

$$0 < \mathbb{E}^0[\gamma_0(T)\, B] < \infty.$$

6. The *hedging price* of the corresponding contingent claim is defined by

$$u_0 \doteq inf\{x > 0;\ \exists\, (\pi, c) \in \mathscr{A}_0(x)\ s.t.\ X^{x,\pi,c}(T) \ge B\ a.s.\}.$$

4.4 Convex Sets and Constrained Portfolios

In this section, we fix a nonempty, closed, convex set K in $\mathscr{R}^d$. For this convex set K,

$$\delta(x) \;\equiv\; \delta(x|K) \;\doteq\; sup_{\pi\in K}\,\langle -\pi, x\rangle : \mathscr{R}^d \to \mathscr{R}\cup\{+\infty\}. \tag{4.4.1}$$

Here, δ is the "support function" of the convex set $-K$. Note that if $-K$ is interpreted to be the set of subgradients of a function g at a point, δ is the best convex approximation of g at that point. Furthermore,

$$\begin{aligned}\tilde{K} &\doteq \{x\in\mathscr{R}^d;\ \delta(x|K)<\infty\}\\ &= \{x\in\mathscr{R}^d;\ \exists\,\beta\in\mathscr{R}\ s.t.\ -\pi^* x\le\beta,\ \forall\,\pi\in K\}.\end{aligned} \tag{4.4.2}$$

$\tilde{K}$ is a convex cone (the barrier cone of $-K$) such that $\delta\,(\cdot\,|K)$ is continuous on $\tilde{K}$ and bounded from below on $\mathscr{R}^d$:

$$\delta(x|K)\ge\delta_0,\quad \forall\ x\in\mathscr{R}^d\ \text{ for some }\delta_0\ \in\mathscr{R}.$$

We denote the set of admissible pairs (π, c) as $\mathscr{A}_0$. From now on, we replace the set of admissible policies $\mathscr{A}_0(x)$ with

$$\mathscr{A}'(x)\doteq\{(\pi,c)\in\mathscr{A}_0(x)\,;\pi(t,w)\in K for\ \ell\otimes\mathbb{P}-a.e.\ (t,w)\}.$$

In other words, $\mathscr{A}'(x)$ is the set of admissible portfolios that also stay in the set K at all times. Now, the class $\mathscr{H}$ is $\tilde{K}$ valued $\mathscr{F}_t$ progressively measurable processes $\nu:[0,T]\to\tilde{K}$, that satisfy

$$\mathbb{E}\left[\int_0^T \|\nu(t)\|^2\,dt+\int_0^T\delta(\nu(t))\,dt\right]<\infty.$$

For each $\nu\in\mathscr{H}$, we define a market as follows:

1. ν modifies the returns additively:

$$b_\nu = b+\nu+\delta(\nu).$$

   This implies, in particular, $\quad\theta_\nu(t)\doteq\theta(t)+\sigma^{-1}(t)\,\nu(t).$

2. $\delta(\nu)$ modifies the interest rate additively:

$$\gamma_\nu\doteq\exp\left[-\int_0^t[r(s)+\delta(\nu(s))]\,ds\right],$$

3. The Radon Nikodyn derivative:

$$Z_\nu \doteq \exp\left[-\int_0^t \langle\theta_\nu(s),\ dW_s\rangle - \frac{1}{2}\|\theta_\nu(s)\|^2\ ds\right],$$

4. A new Brownian Motion:

$$W_\nu \doteq W(t) + \int_0^t \theta_\nu(s)\ ds.$$

5. The new measure:

$$dP_\nu = Z_\nu dP.$$

6. $\mathscr{D} \subset \mathscr{H}$: the subset such that the exponential local martingale $Z_\nu(\cdot)$ is a martingale for $\nu \in \mathscr{D}$.

7. $V(0)$:

$$V(0) \doteq sup_{\nu\in\mathscr{D}}\ \mathbb{E}^\nu\ [\gamma_\nu(T)B] < \infty.$$

A contingent claim B is $K-hedgeable$ if it satisfies the above equation. Note that the optimization is over $\mathscr{D}$, not over $\mathscr{H}$.

8. The dynamics of the price processes under P_ν in terms of W_ν:

$$dP = \langle P, [r(t) + \delta(\nu(t))dt + \sigma dW_\nu(t)]\rangle.$$

It is easy to show how this equation is driven. First of all, we know that

$$W_\nu(t) = W(t) + \int_0^t \theta_\nu(s)\ ds.$$

Therefore,it is obvious that

$$dW(t) = dW_\nu(t) - \theta_\nu(t)\ dt.$$

The price of the instruments in the new market is

$$dP(t) = P(t)\left[\{b(t) + \nu(t) + \delta(\nu(t))\}\ dt + \sigma dW_t\right].$$

Inserting $dW(t)$ into the above equation, we get

$$dP(t) = P(t)\left[\{b(t) + \nu(t) + \delta(\nu(t))\}\ dt + \sigma(dW_\nu(t) - \theta_\nu(t)\ dt)\right],$$

where

$$\theta_\nu(t) = \sigma^{-1}[b(t) + \nu(t) - r(t)].$$

After rearranging the equation by inserting the equivalence of $\theta_\nu(t)$ and canceling the terms with opposite signs we get our result:

$$dP(t) = P(t)[\{r(t) + \delta(\nu(t))\}dt + \sigma dW_\nu(t)].$$

4.5 Hedging With Constrained Portfolios

In this section, we introduce the *hedging price* of a contingent claim B. The constraint in portfolios is to take values in the set K. We see that the hedging price coincides with $V(0) = sup_{\nu\in\mathscr{D}}\ \mathbb{E}^{\nu}\ [\gamma_{\nu}(T)B]$.

The *hedging price with K-constraint portfolios* of a contingent claim B is defined by

$$h(0) \doteq inf\{x \in (0,\infty);\ \exists\ (\pi, c) \in \mathscr{A}'(x),\ s.t.\ X^{x,\pi,c}(T) \geq B\ a.s.\}. \tag{4.5.1}$$

Let $\mathscr{I}$ denote the set of all $\{\mathscr{F}_t\}$-stopping times τ with values on $[0,T]$ and by $\mathscr{I}_{\rho,\sigma}$ the subset of $\mathscr{I}$ consisting of stopping times τ s.t. $\rho(w) \leq \tau(w) \leq \sigma(w),\ \forall\ w \in \Omega$, for any two $\rho \in \mathscr{I}$, $\sigma \in \mathscr{I}$ such that $\rho \leq \sigma$ a.s. Now, let us consider that

$$V(\tau) \doteq \operatorname{ess\,sup}_{\nu\in\mathscr{D}} \mathbb{E}^{\nu}\left[B\gamma_0(T)\exp\left\{-\int_{\tau}^{T}\delta(\nu(s))\,ds\right\}|\mathscr{F}_{\tau}\right]. \tag{4.5.2}$$

Definition 4.5.1 *[17, Definition 2.1, page 3] Let $f : D \to \mathbb{R}$ be a measurable function defined on a measurable set D of $\mathbb{R}^N$ with respect to the Lebesgue measure μ, where $0 < \mu(D) < +\infty$. B is said to be an essential upper bound for f iff $f(x) \leq B$ for almost all $x \in D$, i.e.,*

$$\mu\{x \in D : f(x) > B\} = 0.$$

The least essential upper bound is called essential supremum of f.

$$\operatorname{ess\,sup} f \doteq \inf\{B : B \text{ is an essential upper bound for } f\}.$$

■

Proposition 4.5.2 *[5, Proposition 6.2, page 662] For any contingent claim that satisfies (7), the family of (4.5.2) of random variables $\{V(\tau)\}_{\tau\in\mathscr{I}}$ satisfies the equation of dynamic programming:*

$$V(\tau) = \operatorname{ess\,sup}_{\nu\in\mathscr{D}_{\tau,\theta}} \mathbb{E}^{\nu}\left[V(\theta)\exp\left\{-\int_{\tau}^{\theta}\delta(\nu(u))\,du\right\}|\,\mathscr{F}_{\tau}\right],\quad \forall\ \theta \in \mathscr{I}_{\tau,T}. \tag{4.5.3}$$

The definiton of $V(\tau)$ on given in (4.5.2) involves an unusual discounting:

$$V(\tau) \doteq \operatorname{ess\,sup}_{\nu\in\mathscr{D}} \mathbb{E}^{\nu}\left[B\gamma_0(T)\exp\left\{-\int_{\tau}^{T}\delta(\nu(s))\,ds\right\}|\mathscr{F}_{\tau}\right].$$

This is the same as

$$V(\tau) = \operatorname{ess\,sup}_{\nu\in\mathscr{D}} \mathbb{E}^{\nu}\left[\exp\left(-\int_0^T r(s)\,ds\right) B\,\frac{\gamma_\nu(T)}{\gamma_\nu(\tau)}\middle|\,\mathscr{F}_\tau\right]$$
$$= \exp\left(-\int_0^T r(s)\,ds\right) esssup_{\nu\in\mathscr{D}}\,\mathbb{E}^{\nu}\left[B\,\frac{\gamma_\nu(T)}{\gamma_\nu(\tau)}\middle|\,\mathscr{F}_\tau\right].$$

Thus, the discounting after time τ is made using the discount factor under ν measure, but the discounting between 0 and τ is done using the original discount factor. The main reason for this choice seems to be the simple dynamic programming equation that it leads to. In a DPE we would like to relate future values to present values. Note that in this formulation of the control problem at each time we are choosing a new market. To relate these markets to each other we need a common currency; the currency is the value of money at the original market at time 0.

Proposition 4.5.3 *[5, Proposition 6.3, page 662] The process $V = \{V(t),\ \mathscr{F}_t;\ 0 \le t \le T\}$ of the above proposition can be considered in its RCLL modification and for every $\nu \in \mathscr{D}$,*

$$Q_\nu(t) \doteq V(t)\exp\left(-\int_0^t \delta(\nu(u))\,du\right),\ \mathscr{F}_t;\ 0 \le t \le T. \tag{4.5.4}$$

Here, Q_ν can be thought of as the value of V at time 0 in the market corresponding to ν (the discount with respect to r is already included in the definition of V); so we only need to discount with respect to the residual discount rate of this market.

Q_ν is a $\mathbb{P}_\nu$ supermartingale; let ψ_ν^* be the process that is integrated with respect to W_ν in Q_ν's stochastic integral representation. The optimal portfolio is defined in terms of ψ_ν^* as follows:

$$\hat{\pi}(t) = \frac{\exp\left(\int_0^t \delta(\nu(s))\,ds\right)}{V(t)}\psi_\nu\sigma^{-1}(t).$$

Theorem 4.5.4 *[5, Theorem 6.4, page 662] For an arbitrary contingent claim B*

$$h(0) = V(0).$$

If $V(0) < \infty$, there exists a pair $(\hat{\pi}, \hat{c}) \in \mathscr{A}'(V(0))$ such that

$$X^{V(0),\hat{\pi},\hat{c}}(T) = B,\ a.s..$$

Here, V is the value function of the dual maximization problem. h is the direct definition (the value function of the primal problem.) Note that this theorem says that there is an optimal solution to the primal problem.

Proof. First, we show that $h(0) \leq V(0)$. We may assume that $V(0) < \infty$. From (4.5.4), the martingale representation theorem and the Doob-Meyer decomposition theorem, we have for every $\nu \in \mathscr{D}$:

$$Q_\nu(t) = V(0) + \int_0^t \psi_\nu^*(s)\, dW_\nu(s) - A_\nu(t), \quad 0 \leq t \leq T, \tag{4.5.5}$$

where $\psi_\nu^*(\cdot)$ is an $\mathscr{R}^d$-valued $\{\mathscr{F}_t\}$-progressively measurable and square-integrable process i.e. $\int_{-\infty}^{\infty} |\psi_\nu(x)|^2\, dx < \infty$ and $A_\nu(\cdot)$ is adapted with increasing RCLL paths and $A_\nu(0) = A_\nu(T) < \infty$. Now, we consider the positive, adapted RCLL process

$$\hat{X}(t) \doteq \frac{V(t)}{\gamma_0(t)} = \frac{Q_\nu(t)}{\gamma_\nu(t)}, \quad 0 \leq t \leq T, \ \ \forall\, \nu \in \mathscr{D}, \tag{4.5.6}$$

with $\hat{X}(0) = V(0)$, $\hat{X}(T) = B$ a.s. To show that $h(0) \leq V(0)$, we have to find a pair $(\hat{\pi}, \hat{c}) \in \mathscr{A}'(V(0))$ such that $\hat{X}(\cdot) = X^{V(0),\hat{\pi},\hat{c}}(\cdot)$. We have from (4.5.4) that

$$Q_\nu(t) = V(t) \exp\left(-\int_0^t \delta(\nu(s))\, ds\right). \tag{4.5.7}$$

From this equation, for any $\mu \in \mathscr{D}$ and $\nu \in \mathscr{D}$ we have,

$$Q_\mu(t) = V(t) \exp\left(-\int_0^t \delta(\mu(s))\, ds\right). \tag{4.5.8}$$

From (4.5.7), it is obvious to see that

$$V(t) = Q_\nu(t) \exp\left(\int_0^t \delta(\nu(s))\, ds\right). \tag{4.5.9}$$

Writing $V(t)$ into the equation (4.5.8), we get

$$Q_\mu(t) = Q_\nu(t) \exp\left(\int_0^t \{\delta(\nu(s)) - \delta(\mu(s))\}\ ds\right). \tag{4.5.10}$$

Taking derivative of both sides, we have

$$\begin{aligned}
dQ_\mu(t) &= dQ_\nu(t) \exp\left(\int_0^t \{\delta(\nu(s)) - \delta(\mu(s))\}\ ds\right) \\
&\qquad + Q_\nu(t)\{\delta(\nu(t)) - \delta(\mu(t))\}\, dt \exp\left(\int_0^t \{\delta(\nu(s)) - \delta(\mu(s))\}\ ds\right) && (4.5.11) \\
&= \exp\left(\int_0^t \{\delta(\nu(s)) - \delta(\mu(s))\}\ ds\right) \\
&\qquad \times [Q_\nu(t)\ \{\delta(\nu(t)) - \delta(\mu(t))\}\, dt + \psi_\nu^*(t) dW_\nu(t) - dA_\nu(t)] && (4.5.12) \\
&= \exp\left(\int_0^t \{\delta(\nu(s)) - \delta(\mu(s))\}\ ds\right) \\
&\qquad \times \Big[\hat{X}(t)\gamma_\nu(t)\{\delta(\nu(t)) - \delta(\mu(t))\}\ dt - dA_\nu(t) \\
&\qquad\quad + \psi_\nu^*(t)\left(dW_\mu(t) + \sigma^{-1}(t)(\nu(t) - \mu(t))\, dt\right)\Big]. && (4.5.13)
\end{aligned}$$

We find dW_ν in terms of dW_μ by using the following method;

$$\begin{aligned} W_\nu(t) = W(t) + \int_0^t \theta_\nu(s)\,ds \Rightarrow W_\mu(t) &= W(t) + \int_0^t \theta_\mu(s)\,ds \\ &\Rightarrow W_\nu(t) = W_\mu(t) - \int_0^t \theta_\mu(s)\,ds + \int_0^t \theta_\nu(s)\,ds \\ &\Rightarrow dW_\nu(t) = dW_\mu(t) - \underbrace{\theta_\mu(t)\,dt}_{\theta(t)dt+\sigma^{-1}(t)\mu(t)dt} + \underbrace{\theta_\nu(t)\,dt}_{\theta(t)dt+\sigma^{-1}(t)\nu(t)dt} \end{aligned}$$

$$\therefore\ dW_\nu(t) = dW_\mu(t) + \sigma^{-1}(t)[\nu(t) - \mu(t)]\,dt.$$

Comparing (4.5.13) with

$$dQ_\mu(t) = \psi_\mu^*(t)dW_\mu(t) - dA_\mu(t),$$

we conclude that

$$\psi_\nu^*(t)\exp\left(\int_0^t \delta(\nu(s))\,ds\right) = \psi_\mu^*(t)\exp\left(\int_0^t \delta(\mu(s))\,ds\right),$$

and, hence, that this expression is independent of $\nu \in \mathscr{D}$:

$$\psi_\nu^*(t)\exp\left(\int_0^t \delta(\nu(s))\,ds\right) = \hat{X}(t)\,\gamma_0(t)\,\hat{\pi}^*(t)\,\sigma(t), \quad \forall\ 0 \le t \le T,\ \nu \in \mathscr{D}, \tag{4.5.14}$$

for some adapted, $\mathscr{R}^d$-valued, a.s. square-integrable process $\hat{\pi}$.

Similarly, we conclude from (4.5.13) and (4.5.14) that

$$\begin{aligned} &\exp\left(\int_0^t \delta(\nu(s))\,ds\right)dA_\nu(t) - \gamma_0(t)\hat{X}(t)[\delta(\nu(t)) + \hat{\pi}^*(t)\nu(t)]\,dt \\ &= \exp\left(\int_0^t \delta(\mu(s))\,ds\right)dA_\mu(t) - \gamma_0(t)\hat{X}(t)[\delta(\mu(t)) + \hat{\pi}^*(t)\mu(t)]\,dt, \end{aligned}$$

this expression is also independent of $\nu \in \mathscr{D}$:

$$\hat{c}(t) \doteq \int_0^t \gamma_\nu^{-1}(s)dA_\nu(s) - \int_0^t \hat{X}(s)[\delta(\nu(s)) + \nu^*(s)\hat{\pi}(s)]\,ds \tag{4.5.15}$$

for every $0 \le t \le T,\ \nu \in \mathscr{D}$. From (4.5.15) with $\nu \equiv 0$ we obtain

$$\hat{c}(t) = \int_0^t \gamma_0^{-1}(s)\,dA_0(s), \quad 0 \le t \le T,$$

hence, $\hat{c}(\cdot)$ is an increasing, adapted, RCLL process with $\hat{c}(0) = 0$ and $\hat{c}(T) < \infty$ a.s. Next, we claim that

$$\delta(\nu(t,w)) + \nu^*(t,w)\,\hat{\pi}(t,w) \ge 0, \quad \ell \otimes \mathbb{P} - \text{a.e.}, \tag{4.5.16}$$

holds for every $\nu \in \mathscr{D}$. To verify (4.5.16), notice that from (4.5.15) we obtain

$$
\begin{aligned}
A_\nu(t) &= \int_0^t \gamma_\nu(s)\{d\hat{c}(s) + \hat{X}(s)\{\delta(\nu(s)) + \nu^*(s)\hat{\pi}(s)\}\, ds\} \\
&\le k\left[\hat{c}(t) + \int_0^t \{\delta(\nu(s)) + \nu^*(s)\hat{\pi}(s)\}\hat{X}(s)\, ds\right], \;\; 0 \le t \le T, \;\; \nu \in \mathscr{D},
\end{aligned}
$$

for some $k > 0$. Fix $\nu \in \mathscr{D}$ and define the set $F_t \doteq \{w \in \Omega;\; \delta(\nu(t,w)) + \nu^*(t,w)\hat{\pi}(t,w) < 0\}$ for every $t \in [0,T]$. Let $\mu(t) \doteq [\nu(t)1_{F_t^c} + n\nu(t)1_{F_t}](1 + \|\nu(t)\|)^{-1}$, $n \in \mathbb{N}$. Then, $\mu \in \mathscr{D}$ and, assuming that (4.5.16) does not hold, we get for n large enough,

$$
\begin{aligned}
\mathbb{E}\left[A_\mu(T)\right] &\le \mathbb{E}\left[k\hat{c}(T) + k\int_0^T (1 + \|\nu(t)\|)^{-1}\hat{X}(t)1_{F_t^c} \times \{\delta(\nu(t)) + \nu^*(t)\hat{\pi}(t)\}\, dt\right] \\
&\quad + n\mathbb{E}\left[k\int_0^T (1 + \|\nu(t)\|)^{-1}\hat{X}(t)1_{F_t}\{\delta(\nu(t)) + \nu^*(t)\hat{pi}(t)\}\, dt\right] < 0,
\end{aligned}
$$

a contradiction. Now we can put together (4.5.5)-(4.5.15) to deduce

$$
\begin{aligned}
d(\gamma_\nu(t)\hat{X}(t)) &= dQ_\nu(t) = \psi_\nu^*(t)dW_\nu(t) - dA_\nu(t) \\
&= \gamma_\nu(t)[-d\hat{c}(t) - \hat{X}(t)\{\delta(\nu(t)) + \nu^*(t)\hat{\pi}(t)\}\, dt \\
&\qquad + \hat{X}(t)\hat{\pi}^*(t)\sigma(t)dW_\nu(t)]
\end{aligned}
\tag{4.5.17}
$$

for any given $\nu \in \mathscr{D}$. As a consequence, the process

$$
\begin{aligned}
\hat{M}_\nu(t) &\doteq \gamma_\nu(t)\hat{X}(t) + \int_0^t \gamma_\nu(s)\, d\hat{c}(s) \\
&\quad + \int_0^t \gamma_\nu(s)\hat{X}(s)[\delta(\nu(s)) + \nu^*(s)\hat{\pi}(s)]\, ds \\
&= V(0) + \int_0^t \gamma_\nu(s)\hat{X}(s)\hat{\pi}^*(s)\sigma(s)dW_\nu(s), \quad 0 \le t \le T,
\end{aligned}
\tag{4.5.18}
$$

is a nonnegative, $\mathbb{P}^\nu$-local martingale. In particular, for $\nu \equiv 0$, (4.5.17) gives

$$
d(\gamma_0(t)\hat{X}(t)) = -\gamma_0(t)d\hat{c}(t) + \gamma_0(t)\hat{X}(t)\hat{\pi}^*(t)\sigma(t)dW_0(t),
$$

$$
\hat{X}(0) = V(0), \qquad \hat{X}(T) = B.
$$

This shows $\hat{X}(\cdot) \equiv X^{V(0),\hat{\pi},\hat{c}}(\cdot)$ and $h(0) \le V(0) < \infty$. It remains to show that $h(0) \ge V(0)$ to complete the proof. We may assume $h(0) < \infty$, then there exists a number $x \in (0,\infty)$ such that $X^{x,\pi,c}(T) \ge B$ a.s. for some $(\pi, c) \in \mathscr{A}'(x)$. Then, (4.5.17) holds and it follows from the supermartingale property that

$$x \geq \mathbb{E}^{\nu}\Bigg[\gamma_\nu(T)X^{x,\pi,c}(T) + \int_0^T \gamma_\nu(t)\, dc(t) + \int_0^T \gamma_\nu(t)X^{x,\pi,c}(t)\{\delta(\nu(t)) + \nu^*(t)\pi(t)\}\, dt\Bigg] \geq \mathbb{E}^{\nu}[B\gamma_\nu(T)], \tag{4.5.19}$$

$\forall\ \nu \in \mathscr{D}$. Therefore, $x \geq V(0)$ and thus $h(0) \geq V(0)$. As a result, we've shown that $h(0) = V(0)$. ■

Definition 4.5.5 *[5, Definition 6.5, page 665] K-hedgeable contingent claim B is K−attainable if there exists a portfolio process π with values in K such that $(\pi, 0) \in \mathscr{A}'(V(0))$ and $X^{V(0),\pi,0}(T) = B$ a.s.* ■

Theorem 4.5.6 *[5, Theorem 6.6, page 666] For a given K-hedgeable contingent claim B and any given $\lambda \in \mathscr{D}$, the conditions*

$$\left\{Q_\lambda(t) = V(t)\exp\left(-\int_0^t \delta(\lambda(u))\, du\right),\ \mathscr{F}_t;\ \ 0 \leq t \leq T\right\} is\ a\ \mathbb{P}^\lambda - martingale, \tag{4.5.20}$$

$$\lambda\ achieves\ the\ supremum\ in\ V(0) = sup_{\nu\in\mathscr{D}}\mathbb{E}^{\nu}\ [B\gamma_\nu(T)], \tag{4.5.21}$$

$$\left\{B\ is\ K-attainable\ (by\ a\ portfolio\ \pi)\ and\ the\ corresponding\ \gamma_\lambda(\cdot)X^{V(0),\pi,0}(\cdot)\ is\ a\ \mathbb{P}^\lambda - martingale\right\} \tag{4.5.22}$$

are equivalent and imply

$$\hat{c}(t,w) = 0,\ \ \delta(\lambda(t,w)) + \lambda^*(t,w)\hat{\pi}(t,w) = 0,\ \ \ell \otimes P - a.e. \tag{4.5.23}$$

for the pair $(\hat{\pi}, \hat{c}) \in \mathscr{A}'(V(0))$ of Theorem 4.5.4

Theorem 4.5.7 *[5, Theorem 6.7, page 666] Let B be a K-hedgeable contingent claim. Suppose that for any $\nu \in \mathscr{D}$ with $\delta(\nu) + \nu * \hat{\pi} \equiv 0$,*

$$Q_\nu(\cdot)\ in\ (4.5.4)\ is\ of\ class\ D[0,T],\ under\ \mathbb{P}^\nu. \tag{4.5.24}$$

Then, for any given $\lambda \in \mathscr{D}$, the conditions (4.5.20), (4.5.21) and (4.5.23) are equivalent, and imply

$$\left\{B\ is\ K-attainable\ (by\ a\ portfolio\ \pi)\ and\ the\ corresponding\ \gamma_0(\cdot)X^{V(0),\pi,0}(\cdot)\ is\ a\ \mathbb{P}^0 - martingale\right\} \tag{4.5.25}$$

4.5.1 Examples [5, Examples, page 669]

1. (No short selling) $K = [0, \infty)$ with r, $\sigma \equiv \sigma_{11}$ positive constants. Then $\tilde{K} = K$, $\delta(x) = 0$ for $x \geq 0$, $\delta(x) = \infty$ for $x < 0$. So, $x + \delta(x) = x \geq 0$ on $\tilde{K}$. Take $B = \varphi(P_1(T))$, where $\varphi : \mathscr{R}^+ \to [0, \infty)$ is continuous, increasing, piecewise continuously differentiable and satisfies $\varphi(p) \leq \alpha p$ for some real $\alpha > 0$. Then, we have $V(0) < \infty$. We know that

$$X(t) = e^{rt}V(t) = e^{-r(T-t)}U(T-t, P_1(t)), \tag{4.5.26}$$

$$\hat{\pi}(t) = \frac{Q(T-t, P_1(t))}{U(T-t, P_1(t))} = P_1(t)\frac{(\partial/\partial p)U(T-t, P_1(t))}{U(T-t, P_1(t))} \geq 0, \tag{4.5.27}$$

where

$$U(t, p) \doteq \int_{-\infty}^{\infty} \varphi\left(pe^{\sigma(\xi+\delta t)}\right)\frac{e^{-\xi^2/2t}}{\sqrt{2\pi t}}\, d\xi, \tag{4.5.28}$$

$$Q(t, p) \doteq \int_{-\infty}^{\infty} \psi(\left(pe^{\sigma(\xi+\delta t)}\right)\frac{e^{-\xi^2/2t}}{\sqrt{2\pi t}}\, d\xi = p\frac{\partial}{\partial p}U(t, p) \tag{4.5.29}$$

with $\delta = (r/\sigma) - (\sigma/2)$ and $\psi(p) \doteq p\varphi'(p) \geq 0$.

As a result, this example is the case of the European call option $\varphi(p) = (p - q)^+$ with exercise price $q > 0$. Thus, the unconstrained hedging portfolio does not require short-selling and the constraint $K = [0, \infty)$ makes no difference.

2. (No borrowing) Let $K = (-\infty, 1]$, then $\tilde{K} = (-\infty, 0]$, $\delta(\nu) = \nu$ and consider the contingent claim $B = (P_1(T) - q)^+$. The process $\exp(\int_0^t \nu(s)\, ds)\gamma_0(t)P_1(t)$ is a $\mathbb{P}^\nu$-martingale, for every $\nu \in \mathscr{D}$. Consequently,

$$\begin{aligned} V(t) &\leq \text{ess sup}_{\nu\in\mathscr{D}}\exp\left(-\int_0^t \nu(s)\, ds\right)\mathbb{E}^\nu\left[\exp\left(\int_0^T \nu(s)\, ds\right)\gamma_0(T)P_1(T)|\mathscr{F}_t\right] \\ &= \gamma_0(t)P_1(t), \quad 0 \leq t \leq T. \end{aligned} \tag{4.5.30}$$

On the other hand, by Jensen's inequality,

$$\begin{aligned} V(t) &\geq \text{ess sup}_{\nu\in\mathscr{D}}\left\{\exp\left(-\int_0^t \nu(s)\, ds\right)\mathbb{E}^\nu\left[\exp\left(\int_0^T \nu(s)\, ds\right)\gamma_0(T)P_1(T)|\mathscr{F}_t\right]\right. \\ &\quad \left. - \mathbb{E}^\nu\left[\exp\left(\int_t^T \nu(s)\, ds\right)\gamma_0(T)q|\mathscr{F}_t\right]\right\}^+ \\ &\geq \text{ess sup}_{\nu\in\mathscr{D}_d}\left\{\gamma_0(t)P_1(t) - \exp\left(\int_t^T \nu(s)\, ds\right)q\mathbb{E}^\nu[\gamma_0(T)|\mathscr{F}_t]\right\}^+ \\ &= \gamma_0(t)P_1(t) \end{aligned} \tag{4.5.31}$$

for $0 \le t \le T$. The inequalities (4.5.30) and (4.5.31) imply

$$V(t) = \begin{cases} \gamma_0(t)P_1(t), & 0 \le t < T, \\ \gamma_0(T)(P_1(T) - q)^+, & t = T, \end{cases} \tag{4.5.32}$$

or equivalently

$$dV(t) = \gamma_0(t)P_1(t)\sigma(t)dW_0(t) - dA_0(t), \tag{4.5.33}$$

where

$$A_0(t) = \begin{cases} 0, & 0 \le t < T, \\ \gamma_0(T)[P_1(T) - (P_1(T) - q)^+], & t = T. \end{cases} \tag{4.5.34}$$

In particular, (4.5.33) implies $X \doteq V/\gamma_0 \equiv X^{V(0),\hat{\pi},\hat{c}}$ with

$$\hat{\pi} \equiv 1, \quad \hat{c}(t) = \int_0^t \gamma_0^{-1}(s)\, dA_0(s). \tag{4.5.35}$$

In other words, in order to replicate $B = (P_1(T) - q)^+$ without borrowing, one has to invest all the wealth in the stock, not consume before the expiration date T, and consume at time $t = T$ the amount

$$\hat{c}(T) = P_1(T) - (P_1(T) - q)^+ = \min(P_1(T), q). \tag{4.5.36}$$

This example resolves two questions that can be raised in the context of Theorem 4.5.4. First, it shows the process $V(\cdot)$ is not, in general, a regular $\mathbb{P}^0$-supermartingale, for if it were, $A_0(\cdot)$ would be continuous. Second, it shows that, in general, the supremum of $\gamma_\nu = \exp\left[-\int_0^t [r(s) + \delta(\nu(s))]\, ds\right]$ is not attained. Indeed, one has to let $\nu \equiv -\infty$ in order to achieve equality in (4.5.31).

3. (Option with a ceiling on a stock that cannot be traded.) Let $K = \{x \in \mathscr{R}^d;\ x_1 = 0\}$, $B = (P_1(T) - q)^+ \wedge L$ for some real $q > 0,\ L > 0$. Then, $\tilde{K} = \{x \in \mathscr{R}^d;\ x_2 = x_3 = \cdots = x_d = 0\}$ and $\delta \equiv 0$ on $\tilde{K}$. Assume deterministic market coefficients. We want to verify

$$V(0) = \gamma_0(T)L \tag{4.5.37}$$

by first showing $V(0) \ge \gamma_0(T)L$ and then providing the opposite inequality by constructing a consumption process c such that the wealth process corresponding to the triple $(\gamma_0(T)L, 0, c)$ satisfies $X(T) = B$ a.e.. We have

$$V(0) \ge \gamma_0(T)L \text{ ess sup}_{\nu \in \mathscr{D}_d} \mathbb{E}^\nu 1_{\{P_1(T) - q > L\}}. \tag{4.5.38}$$

Define an $\mathscr{R}^d_+$-valued process $\tilde{P}^{(\nu)}(\cdot) = \{\tilde{P}^{\nu}_i(\cdot)\}^d_{i=1}$ by

$$d\tilde{P}^{(\nu)}_i(t) = \tilde{P}^{(\nu)}_i(t)[r(t) - \nu_i(t)]\, dt + \tilde{P}^{(\nu)}_i(t)\, \Sigma^d_{j=1}\sigma_{ij}(t)\, dW^{(j)}_0(t), \quad \tilde{P}^{(\nu)}_i(0) = P_i(0), \tag{4.5.39}$$

for $i = 1, \ldots, d$ and $\nu \in \mathscr{D}_d$. We have

$$\mathbb{E}^{\nu} 1_{\{P_1(T)-q>L\}} = \mathbb{E}^0 1_{\{\tilde{P}^{\nu}_1(T)-q>L\}}. \tag{4.5.40}$$

Letting $\nu \to -\infty$, (4.5.38)-(4.5.40) imply

$$V(0) \geq \gamma_0(T)L. \tag{4.5.41}$$

Next, we define a consumption process c by

$$c(t) = \begin{cases} 0, & t < T \text{ or } t = T,\ P_1(T) - q > L, \\ L - (P_1(T) - q)^+, & t = T,\ P_1(T) - q \leq L. \end{cases} \tag{4.5.42}$$

Then the wealth process $X(\cdot)$ associated with the policy $(\gamma_0(T)L, 0, c)$ is given by $X(t) = (\gamma_0(T)/\gamma_0(t))L$ for $t < T$ and by

$$X(T) = L - c(T) = B \tag{4.5.43}$$

for $t = T$. This implies $V(0) \geq \gamma_0(T)L$ by Theorem 4.5.4. Consequently, the way to hedge a bounded option on a stock that is not available for investment is to replicate the upper bound of the option by investing in the bond only, and then to consume the difference at the expiration date.

4.6 Numerical Solution of the Dual Problem

For the purposes of this section we will concentrate on the problem of replication when the interest rate for borrowing and lending is different.

We make the following assumptions. There is only one risky security with constant volatility σ. The interest rate for borrowing and lending are R and r respectively and both are assumed to be constant. The security to be hedged is assumed to be of the form

$$B = \phi(P(T))$$

for some positive measurable and finite function ϕ. Then as developed in [5, Example 9.5] the hedging problem of (4.1.3) is equivalent to the following stochastic optimal control problem:

$$V(t,p) \doteq \sup_{\nu\in\mathscr{D}} \mathbb{E}\left[\phi(P(T))e^{-\int_t^T \nu(s)ds}|P(t)=p\right], \tag{4.6.1}$$

where P satisfies

$$P(t) = P(0) + \int_0^t P(s)\nu(s)ds + \int_0^t \sigma P(s)dW(s) \tag{4.6.2}$$

and $\mathscr{D}$ is the set of progressively measurable processes taking values in $[r,R]$.

Remark 4.6.1 It is interesting to note that the state process of the dual problem is the price process (which can be d dimensional) whereas the state process of the primal problem is intimately related to the one dimensional wealth process. Therefore, in general, by going to the dual problem we increase the dimension of the problem.

We begin by noting that Ito's formula implies that the explicit solution of (4.6.2) is

$$P(t) = e^{Y(t)}$$

with

$$Y(t) = y + \int_0^t \nu(s)ds - \frac{1}{2}\sigma^2 t + \sigma W(t). \tag{4.6.3}$$

Then, instead of P we can use Y as the state process and rewrite (4.6.1) as

$$V(t,y) \doteq \sup_{\nu\in\mathscr{D}} \mathbb{E}\left[\phi(\exp(Y(T))e^{-\int_t^T \nu(s)ds}|Y(t)=y\right], \tag{4.6.4}$$

with Y given in (4.6.3).

Our goal is to solve (4.6.4) by discretizing time and thereby obtaining a discrete time optimal control problem.

For ease of notation let $t = 0$; (4.6.3) says that $Y(0) = y$ a constant. Let n be the level of discretization and let $\Delta_n = T/2^n$ be the size of the discrete time step at level n. Let $\mathscr{D}_n$ denote the class of simple processes taking values in $[r,R]$ that are piecewise constant on the intervals $[k,k+1)\Delta_n$ and define

$$V_n(0,y) \doteq \sup_{\nu\in\mathscr{D}_n} \mathbb{E}_y\left[\phi(\exp(Y(T))e^{-\int_t^T \nu(s)ds}\right], \tag{4.6.5}$$

where the subscript y denotes $Y(0) = y$. We would like to show

1. $V_n(0,y)$ can be computed recursively using dynamic programming.

2. $V_n(0,y) \to V(0,y)$.

4.7 Solution of the Discretized Problem

Let us begin with the first. For $\nu \in \mathscr{D}_n$ Y can be written as

$$Y(k\Delta_n) = y + \sum_{j=0}^{k} \nu^n(j)\Delta_n - \frac{1}{2}\sigma^2 k\Delta_n + \sigma \sum_{j=0}^{k} W(k\Delta_n) - W((k-1)\Delta_n),$$

where $\nu^n(j)$ is an $\mathscr{F}_{j\Delta n}$ measurable random variable taking values in $[r, R]$. Now we argue that it is sufficient to consider $\nu^n(j)$ that depend only on $\{W(i\Delta_n), i \leq j\}$. Let $\tilde{V}_n(0, y)$ denote the sup in (4.6.4) over such ν. Because we are taking a sup over a smaller class we have

$$\tilde{V}_n(0, y) \leq V_n(0, y).$$

To see that the opposite inequality, imagine that W is sampled as follows: first the $W(i\Delta_n)$ are sampled and then the Brownian bridges that connect these values are sampled. Apriori, the controls $\nu^n(j)$ that define the ν in (4.6.4) are a function of both the values $W(i\Delta n)$ and the Brownian bridges that connect them. Now, condition the expectation in (4.6.5) on the Brownian bridges. The resulting conditional expectation (because the values of the Brownian bridges are now fixed) is of the same form that occurs in the definition of $\tilde{V}_n$. Then the conditioned expectation is less then $\tilde{V}_n(0, y)$. This constant term that doesn't depend on the bridges comes out of the expectation and gives

$$V_n(0, y) \leq \tilde{V}_n(0, y).$$

Combining the last inequalities we get

$$V_n(0, y) = \tilde{V}(n, y).$$

In this way we see that (4.6.5) is equivalent to the following finite step problem. Let $\{\epsilon^n(k)\}$ is an iid sequence with common distribution $N(0, 1)$. Let

$$\mathscr{F}_k^n \doteq \sigma(\epsilon^n(i), i \leq k).$$

$\nu^n(k) \in \mathscr{F}_k^n$.

$$Y^n(k) = Y^n(k-1) + \nu^n(k)\Delta_n + \sigma\sqrt{\Delta_n}\epsilon^n(k), \quad Y^n(0) = y.$$

Note that this problem is defined in terms of a finite iid sequence of random variables rather than a Brownian motion. It is a standard discrete time finite optimal stopping problem. It is well known for these types of problems that the value function satisfies a dynamic programming equation (DPE). In the present case this equation turns out to be

$$V_n(y,(k-1)\Delta_n) \tag{4.7.1}$$
$$= \sup_{\nu\in[r,R]} e^{-\nu\Delta_n}\frac{1}{\sqrt{2\pi}}\int V_n\left(y+\left(\nu-\frac{1}{2}\sigma^2\right)\Delta_n+\sigma\sqrt{\Delta_n}x, k\Delta_n\right)e^{-x^2/2}dx$$

along with the base case

$$V_n(y,T)=\phi(e^y).$$

From the last DPE several facts are immediate. If $\phi(e^y)$ is convex in y then so are all $V_n(\cdot,\Delta_n k)$, $k=0,1,2,3,...,2^n$ and in particular V_n are all differentiable, i.e., $\frac{dV_n}{dy}$ is well defined. Let us assume that $\phi(e^y)$ is convex. Then we can simply differentiate the function of ν on the right side of (4.7.1) to compute the optimal ν. This derivative is

$$e^{-\nu\Delta_n}\Delta_n\left(\mathbb{E}\left[V\left(y+\left(\nu-\frac{1}{2}\sigma^2\right)\Delta+\epsilon,k\Delta_n\right)\right]\right.$$
$$\left.-\mathbb{E}\left[\frac{dV}{dy}\left(y+\left(\nu-\frac{1}{2}\sigma^2\right)\Delta+\epsilon,k\Delta_n\right)\right]\right), \tag{4.7.2}$$

where ϵ is $N(0,\sigma^2\Delta_n)$. The optimal ν can be obtained by solving this equation. (4.7.2) also suggests the following result.

Proposition 4.7.1 *Define $B(y)\doteq\phi(e^y)$. If*

$$B(y)\geq\frac{dB}{dy} \tag{4.7.3}$$

for almost all y then the optimizer ν in (4.7.1) *is always $\nu^*=R$.*

Proof. We will prove by induction that

$$V(y,k\Delta_n)\geq\frac{dV}{dy}(y,k\Delta_n) \tag{4.7.4}$$

for all k; as shown below, this will in particular establish the result of the proposition. For $k=2^n$ (4.7.4) is exactly the assumption (4.7.3). Let us now assume that (4.7.4) is true for k. Then (4.7.2) implies that the right side of (4.7.1) is increasing in ν and thus it is maximized for

$$\nu^*=R. \tag{4.7.5}$$

Substituting this value in (4.7.1) gives

$$V(y,(k-1)\Delta_n)=e^{-R\Delta_n}\mathbb{E}\left[V\left(y+\left(R-\frac{1}{2}\sigma^2\right)\Delta_n+\epsilon,k\Delta_n\right)\right]$$

where ϵ is $N(0, \sigma^2\Delta_n)$. Now differentiating the right side of this display with respect to y gives

$$
\begin{aligned}
&V(y, (k-1)\Delta_n) - \frac{dV}{dy}(y, (k-1)\Delta_n) \\
&= e^{-R\Delta_n}\Delta_n\left(\mathbb{E}\left[V\left(y + \left(R - \frac{1}{2}\sigma^2\right)\Delta + \epsilon, k\Delta_n\right)\right]\right. \\
&\quad \left. - \mathbb{E}\left[\frac{dV}{dy}\left(y + \left(R - \frac{1}{2}\sigma^2\right)\Delta + \epsilon, k\Delta_n\right)\right]\right).
\end{aligned}
$$

(4.7.4) again implies that this expression is greater than zero, which proves (4.7.4) for $k-1$. This completes the induction. (4.7.5) implies the statement of the theorem. ■

Reversing the argument in the previous proposition gives

Proposition 4.7.2 *Define* $B(y) \doteq \phi(e^y)$. *If*

$$
B(y) \leq \frac{dB}{dy} \tag{4.7.6}
$$

for almost all y *then the optimizer* v *in* (4.7.1) *is always* $v^* = r$.

Example 4.7.3 Take $\phi(e^y) = (e^y - K)^+$. This choice corresponds to a call option with strike K. This is clearly a convex function in y and it satisfies (4.7.6). Proposition 4.7.2 implies that for call options the optimal v always turns out to be r in the dual problem.

Remark 4.7.4 Our propositions 4.7.1 and 4.7.2 are similar to the argument on page 676 of [5], which is based on the assumption

$$
\langle p, D\phi \rangle \geq \phi(p).
$$

The key difference is that ϕ is assumed to be a C_1 function in [5] whereas our results need differentiability only almost everywhere. For example, to the best of our understanding the argument of [5, page 676] is not directly applicable to the call option treated in the previous example.

4.7.1 Convergence of V_n to V

Now let us go back to (4.6.5) and (4.6.4) and prove

Proposition 4.7.5 *If*

$$\mathbb{E}\left[\phi\left(e^{y+(R-\frac{1}{2}\sigma^2)t+|W_T|}\right)\right] < \infty. \tag{4.7.7}$$

then

$$\lim_{n\to} V_n(t,y) = V(t,y). \tag{4.7.8}$$

Remark 4.7.6 Note that (4.7.7) is satisfied by any function which has polynomial growth and in particular by Lipschitz continuous functions.

Proof. We begin by noting that $\mathscr{D}_n \subset \mathscr{D}_{n+1}$. Therefore, the sup in the definition of $V_{n+1}(t,y)$ is over a larger set than the sup in the definition of $V_n(t,y)$. This implies that $V_n(t,y)$ is increasing in n and therefore has to converge to a limit $\bar{V}(t,y)$. Similarly, $\mathscr{D} \supset \mathscr{D}_n$ implies $V(t,y) \geq V_n(t,y)$. As a result of these one gets

$$\lim_{n\to\infty} V_n(t,y) = \bar{V}(t,y) \leq V(t,y). \tag{4.7.9}$$

It remains to prove the reverse inequality. For this purpose fix an arbitrary control $\nu \in \mathscr{D}$. By assumption ν is progressively measurable and bounded. [9, Lemma 2.4, page 132] says that there is a sequence of piecewise constant controls ν^n such that $\nu^n \in \mathscr{D}_n$ and

$$\lim_{n\to\infty} \mathbb{E}\left[\int_0^T |\nu^n(t) - \nu(t)|^2 dt\right] = 0. \tag{4.7.10}$$

The expected difference in using ν^n instead of ν in (4.6.4) is

$$D_n \doteq \left|\mathbb{E}_y\left[\phi(\exp(Y^n(T))e^{-\int_t^T \nu^n(s)ds}\right] - \mathbb{E}_y\left[\phi(\exp(Y(T))e^{-\int_t^T \nu(s)ds}\right]\right|, \tag{4.7.11}$$

where

$$Y^n(t) = y + \int_0^t \nu^n(s)ds - \frac{1}{2}\sigma^2 t + \sigma W(t).$$

We would like to show that $\lim D_n = 0$. To do so, let us take any subsequence D_{n_k} of D_n. (4.7.10) implies that n_k has a further subindex $\{n_{k_j}\}$ for which

$$\nu^{n_{k_j}} \to \nu \text{ almost surely.} \tag{4.7.12}$$

This, the continuity of the exponential function and the boundedness of ν and ν^n imply that

$$e^{-\int_t^T \nu^n(s)ds} \text{ converges to } e^{-\int_t^T \nu(s)ds} \text{ almost surely.} \tag{4.7.13}$$

Furthermore note that

$$\max(|Y^n(t)|, |Y(t)|) \leq y + (R - \frac{1}{2}\sigma^2)T + \sigma W_T.$$

We will use (4.7.13) and (4.7.7) in the invocation of the dominated convergence theorem below.

To continue our analysis, add and subtract

$$\mathbb{E}_y\left[\phi(\exp(Y^n(T))e^{-\int_t^T \nu(s)ds}\right]$$

to (4.7.11) and get

$$\begin{aligned} D_{n_{k_j}} \leq \mathbb{E}_y & \left[\left|\phi(\exp(Y^{n_{k_j}}(T))\left(e^{-\int_t^T \nu^{n_{k_j}}(s)ds} - e^{-\int_t^T \nu(s)ds}\right)\right|\right] \\ & + \mathbb{E}_y\left[|\phi(\exp(Y^{n_{k_j}}(T)) - \phi(\exp(Y(t))|e^{-\int_t^T \nu(s)ds}\right] \end{aligned}$$

(4.7.12), (4.7.13), (4.7.7) and the dominated convergence theorem imply that $D_{n_{j_k}}$ converges to 0. Because this process does not depend on the particular subsequence, we have that

$$\lim_n D_n = 0. \tag{4.7.14}$$

(4.7.14) implies that for any $\epsilon > 0$ we can find $\nu^n \in \mathscr{D}_n$ such that

$$\mathbb{E}_y\left[\phi(\exp(Y(T))e^{-\int_t^T \nu(s)ds}\right] \leq \mathbb{E}_y\left[\phi(\exp(Y^n(T))e^{-\int_t^T \nu^n(s)ds}\right] + \epsilon.$$

Note that the last expression is less than $\bar{V}(y,t) + \epsilon$. Taking the sup of the left side of the last inequality over all $\nu \in \mathscr{D}$ gives

$$V(y,t) \leq \bar{V}(y,t) + \epsilon.$$

Because ϵ in the last display can be chosen arbitrarily small we have

$$V(y,t) \leq \bar{V}(y,t).$$

The last display and (4.7.9) imply (4.7.8).

■

CHAPTER 5

CONCLUSION

This thesis consists of a careful study of two important papers in backward stochastic differential equations literature: [14, Pardoux and Peng, 1990] and [5, Cvitanić and Karatzas, 1993]. The main result of [14] is the existence and uniqueness for an adapted pair $\{x(t), y(t);\ t \in [0,1]\}$ which solves the equation:

$$x(t) + \int_t^1 f(s, x(s), y(s))ds + \int_t^1 [g(s, x(s)) + y(s)]dW_s = X,$$

where

$$f : \Omega \times [0,1] \times \mathbb{R}^d \times \mathbb{R}^{d\times k} \to \mathbb{R}^d,$$

$$g : \Omega \times [0,1] \times \mathbb{R}^d \times \mathbb{R}^{d\times k} \to \mathbb{R}^{d\times k}.$$

The details of all the steps of the proofs in the paper by Pardoux and Peng are provided in our thesis.

In [5] Cvitanić and Karatzas studied the following problem: the hedging of contingent claims with portfolios constrained to take values in a given closed, convex set. The analysis of this paper is based on a dual control problem. We provide a numerical solution of the dual problem when the volatility term is assumed to be constant. Let us now briefly compare the approach taken in this thesis to the solution of the dual problem to the approach suggested in [5, Example 9.5]. In the mentioned reference the authors suggest the following solution method: solve the HJB equation associated with the stochastic optimal control problem. To find the solution they invoke an existence and uniqueness result from the classical PDE theory. Of course, to get a more concrete solution one can solve the PDE numerically and for this there are many algorithms. In our approach, we directly discretize time and obtain a discrete time finite step stochastic optimal control problem whose dynamic programming equation is an integral equation rather than a differential equation. Hence, its solution is much simpler (in fact, it doesn't need a solution, because it defines the solution recursively.) The drawback of the argument presented above is the constancy assumptions on r and σ. Future work may try to extend our approach to cases where these parameters are not necessarily constant.

REFERENCES

[1] F. Antonelli. Backward-forward stochastic differential equations. *The Annals of Applied Probability*, pages 777–793, 1993.

[2] G. Barles, R. Buckdahn, and E. Pardoux. Backward stochastic differential equations and integral-partial differential equations. *Stochastics and Stochastics Reports*, 60(1):57–84, 1997.

[3] J.M. Bismut. Intégrales convexes et probabilités. *Journal of Mathematical Analysis and Applications*, 42(3):639–673, 1973.

[4] P. Cheridito, H.M. Soner, N. Touzi, and N. Victoir. Second-order backward stochastic differential equations and fully nonlinear parabolic pdes. *Communications on Pure and Applied Mathematics*, 60(7):1081–1110, 2007.

[5] J. Cvitanić and I. Karatzas. Hedging contingent claims with constrained portfolios. *The Annals of Applied Probability*, pages 652–681, 1993.

[6] J. Cvitanić and J. Ma. Hedging options for a large investor and forward-backward sde's. *The annals of applied probability*, 6(2):370–398, 1996.

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

[8] N. El Karoui, S. Peng, and M.C. Quenez. Backward stochastic differential equations in finance. *Mathematical finance*, 7(1):1–71, 1997.

[9] I. Karatzas and S.E. Shreve. *Brownian motion and stochastic calculus*. Springer, 1991.

[10] J. Ma, P. Protter, and J. Yong. Solving forward-backward stochastic differential equations explicitlya four step scheme. *Probability Theory and Related Fields*, 98(3):339–359, 1994.

[11] J. Ma, J. Zhang, and Z. Zheng. Weak solutions for forward–backward sdesa martingale problem approach. *The Annals of Probability*, 36(6):2092–2125, 2008.

[12] B.K. Øksendal. *Stochastic differential equations: an introduction with applications*. Springer Verlag, 2003.

[13] E. Pardoux and S. Peng. Backward stochastic differential equations and quasilinear parabolic partial differential equations. *Stochastic partial differential equations and their applications*, pages 200–217, 1992.

[14] E. Pardoux and SG Peng. Adapted solution of a backward stochastic differential equation. *Systems & Control Letters*, 14(1):55–61, 1990.

[15] S. Peng. A general stochastic maximum principle for optimal control problems. *SIAM Journal on control and optimization*, 28:966, 1990.

[16] S. Peng. Probabilistic interpretation for systems of quasilinear parabolic partial differential equations. *Stochastics and stochastics reports*, 37(1-2):61–74, 1991.

[17] HX Phu and A. Hoffmann. Essential supremum and supremum of summable functions. *Numerical functional analysis and optimization*, 17(1-2):161–180, 1996.

[18] R. Rouge and N. El Karoui. Pricing via utility maximization and entropy. *Mathematical Finance*, 10(2):259–276, 2000.

[19] H.M. Soner, N. Touzi, and J. Zhang. Wellposedness of second order backward sdes. *Probability Theory and Related Fields*, pages 1–42, 2011.