About mean-variance hedging with basis risk

Sami Lähdemäki

Master's thesis of mathematics

University of Jyväskylä
Deparment of Mathematics and Statistics
Autumn 2021

Abstract: Sami Lähdemäki, *About mean-variance hedging with basis risk*, Master's Thesis of mathematics, 52p., University of Jyväskylä, Deparment of Mathematics and Statistics, autumn 2021.

In this thesis we introduce a mean-variance hedging problem in an incomplete market. As a main source we follow X. Xue, J. Zhang and C. Weng article *Mean-variance Hedging with Basis Risk*. We assume a time interval $[0, T]$ for some $T > 0$, an arbitrage free financial market, and consider one risk-free asset and $(m + 1)$ risky assets. The dynamics of the assets are given by stochastic differential equations with deterministic and Borel-measurable coefficients. One risky asset is connected to the pay-off function which we want to hedge. We assume that this connected asset can not be used in hedging and this makes the market incomplete. Because of incompleteness perfect hedging is not possible.

We define a profit-and-loss random variable by using the difference between the value of the hedging portfolio and the pay-off function. A mean-variance criterion is used to this random variable and by that the solution is a hedging strategy which maximizes the difference between the expected value and variance of the profit-and-loss random variable.

To find a solution we start by recalling some important results from probability theory and stochastic analysis. We introduce shortly multiple stochastic integrals and properties of them. These integrals are used to define the Malliavin derivative. The mean-variance hedging problem is solved by using Linear-Quadratic theory. We consider an auxiliary problem and show that by solving the auxiliary problem we are able to solve the original problem. The solving method with Linear-Quadratic theory is connected to the backward stochastic differential equations (BSDE) and in the thesis we see also the connection of the BSDEs to the Malliavin derivative. We compute an explicit formula for the Malliavin derivative of a forward contract and an European put and call option.

The pay-off function in this thesis is assumed to be Malliavin differentiable and hence we are able to give an explicit solution for the problem. As a main theorem we formulate an explicit hedging strategy which solves the mean-variance hedging problem in the incomplete market.

Tiivistelmä: Sami Lähdemäki, *About mean-variance hedging with basis risk*, matematiikan pro gradu -tutkielma, 52 s., Jyväskylän yliopisto, Matematiikan ja tilastotieteen laitos, syksy 2021.

Tässä tutkielmassa perehdytään odotusarvo-varianssi -suojausongelmaan (engl. mean-variance hedging problem) epätäydellisillä sijoitusmarkkinoilla. Päälähteenä seuraamme X. Xuen, J. Zhanging ja C. Wengin artikkelia *Mean-variance Hedging with Basis risk*. Oletamme aikavälin $[0, T]$ jollekin $T > 0$, arbitraasivapaan sijoitusmarkkinan, yhden riskittömän sijoituskohteen ja $(m + 1)$ riskillistä sijoituskohdetta. Näiden kohteiden arvon oletetaan noudattavan stokastisia differentiaaliyhtälöitä, joissa kertoimet ovat deterministisiä ja Borel-mitallisia. Yksi näistä riskillisistä sijoituskohteista oletetaan liittyvän vaateeseen, jolle haluamme rakentaa suojaussalkun. Tätä kyseistä sijoituskohdetta ei voida käyttää suojaussalkun rakentamisessa, mikä aiheuttaa sijoitusmarkkinan epätäydellisyyden. Tämän vuoksi myös täydellisen suojaussalkun rakentaminen ei ole mahdollista.

Määrittelemme voittoa/tappiota kuvaavan satunnaismuuttujan käyttämällä suojaussalkun arvon ja vaateen erotusta. Odotusarvo-varianssi -kriteeriä käytetään tähän satunnaismuuttujaan ja tämän johdosta ratkaisu on suojaussalkku, joka maksimoi erotuksen voittoa/tappiota kuvaavan satunnaismuuttujan odotusarvon ja varianssin välillä.

Ratkaisun löytämiseksi aloitamme kertaamalla tärkeitä ja tarpeellisia tuloksia todennäköisyysteoriasta ja stokastisesta analyysistä. Tämän jälkeen esittelemme lyhyesti moninkertaiset stokastiset integraalit ja niiden ominaisuuksia sekä käytämme näitä Malliavin derivaatan määrittelyyn. Odotusarvo-varianssi -ongelman ratkaisun löytämiseksi käytämme "Linear-Quadratic" -teoriaa. Oletamme apuongelman ja osoitamme, että ratkaisemalla apuongelman on mahdollista ratkaista myös alkuperäinen ongelma. Käyttämämme "Linear-Quadratic" -teoria on yhteydessä takaperoisiin stokastisiin differentiaaliyhtälöihin ja tutkielmassa näemme näiden yhteyden Malliavin derivaattaan. Johdamme myös eksplisiittiset ratkaisut suoran sopimuksen ja Euroopalaisen myynti- ja osto-option Malliavin derivaatalle.

Tässä tutkielmassa vaateen oletetaan olevan Malliavin derivoituva ja tämä mahdollistaa eksplisiittisen ratkaisun löytämisen. Pääteoreemana muotoilemme eksplisiittisen suojaussalkun, joka ratkaisee odotusarvo-varianssi -ongelman tilanteessa, jossa sijoitusmarkkina on epätäydellinen.

Contents

Introduction

In this thesis we are interested in optimal hedging strategies for a given pay-off, that means we try to determine a trading strategy which replicates the pay-off. We assume that the financial market is arbitrage free and consists of one risk-free asset earning by constant rate and $(m+1)$ risky assets with dynamics given by stochastic differential equations. In a complete market it is always possible to find a hedging strategy which replicates a given pay-off, and this possibility is actually given as a definition of completeness in [**11**]. However, in the setting of this thesis the asset which is connected to our hedging objective is not allowed to be used in hedging. So we only can use other assets which are stochastically dependent on the one which we can not trade with. This causes the market to be incomplete. So our aim here is to determine trading strategies such that the outcome is close to the given pay-off in some sense.

We use a mean-variance criterion to measure the closeness. The portfolio selection using this criteria has been proposed by Markowitz [**17**], where the variance is assumed to be a measure for risk. We define a profit-and-loss random variable at terminal time $T > 0$ by setting

$$V^\theta(T) = X^\theta(T) - G(S_0; T),$$

where X^θ is the value of the hedging portfolio using the hedging strategy θ, and G is the pay-off function. By the mean-variance criterion the aim is to find a strategy which solves the problem

$$\max_{\theta \in \Theta} \left\{ \mathbb{E}[V^\theta(T)] - \frac{\gamma}{2} \mathrm{Var}[V^\theta(T)] \right\},$$

where $\gamma > 0$ can be interpreted by [**16**] as the weight which the investor puts on the variance.

To solve this problem we use stochastic Linear-Quadratic theory as a tool. This method is used in [**26**] and [**16**] for example. Both papers assume a complete market where all assets are possible to use for hedging. So we can only use the idea. The main source for us which we follow is [**24**].

In Linear-Quadratic theory solving the mean-variance problem is connected to solving two different equations. In our case these are a backward differential equation and a backward stochastic differential equation. It is shown that if these two equations are solvable, then also our problem can be solved.

We will see that solving the backward stochastic differential equation has a connection to Malliavin derivatives so we also shortly introduce Malliavin calculus by using [**19**]. Also the Malliavin derivative of the pay-off function is needed in the optimal hedging strategy so we compute the Malliavin derivatives of a forward contract, a European put and call option and an Asian option.

As a main result we are able to prove an explicit formula for the hedging strategy which solves the mean-variance hedging problem in an incomplete market where the basis risk follows from the setting.

The thesis is organized as follows: In Chapter 1 we recall some important results from Probability theory and Stochastic analysis which are needed later. In Chapter 2 the basics of Malliavin calculus is given and at the end of the chapter the reader gets some useful tools for calculating Malliavin derivatives. Chapter 3 is used to formulate our hedging problem. In Chapter 4 we give the basic idea of stochastic Linear-Quadratic problems and derive important results concerning our problem. Chapter 5 concludes the thesis and contains the main result, the optimal solution to our mean-variance hedging problem. In Appendix A is a collection of results needed in this thesis, and Appendix B contains a list of notations.

CHAPTER 1

Probability theory and stochastic analysis

Here we will give some basic definitions and tools from probability theory and stochastic analysis. In this chapter we will use basically [**7**], [**10**], [**13**] and [**15**].

1.1. Probability space and random variables

DEFINITION 1.1.1. Let Ω be a non-empty set. Then a system $\mathcal{F}$ of subsets $A \subset \Omega$ is called σ-algebra on Ω if the following holds

(1) $\emptyset, \Omega \in \mathcal{F}$
(2) if $A \in \mathcal{F}$ then $A^c \in \mathcal{F}$
(3) if $A_1, A_2, ... \in \mathcal{F}$ then $\cup_{i=1}^{\infty} A_i \in \mathcal{F}$.

In this situation the pair $(\Omega, \mathcal{F})$ is called a measurable space.

For later use we give a definition of the Borel σ-algebra.

DEFINITION 1.1.2 ([**25**]Definition 1.4(ii)). Let $\mathfrak{B}(\mathbb{R}^d)$ be the system of all open subsets in $\mathbb{R}^d$ for all $d \in \mathbb{N}$. Then it is called Borel σ-algebra on $\mathbb{R}^d$.

DEFINITION 1.1.3. Let $(\Omega, \mathcal{F})$ be a measurable space. Then a map $\mu : \mathcal{F} \to [0, \infty]$ is called measure if the following two properties holds

(1) $\mu(\emptyset) = 0$
(2) for all $A_1, A_2, ... \in \mathcal{F}$ with $A_i \cap A_j = \emptyset$ for $i \neq j$ it holds

$$\mu(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i).$$

Then $(\Omega, \mathcal{F}, \mu)$ is called measure space.

DEFINITION 1.1.4. Let $(\Omega, \mathcal{F})$ be a measurable space. If for a measure $\mu : \mathcal{F} \to [0, \infty]$ it holds that $\mu(\Omega) = 1$ then we denote $\mu = \mathbb{P}$ and the measure space $(\Omega, \mathcal{F}, \mathbb{P})$ is called probability space.

A special probability space, called a complete probability space, is used in many cases.

DEFINITION 1.1.5. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Let $A \in \mathcal{F}$ such that $\mathbb{P}(A) = 0$. If $B \subseteq A$ implies that $B \in \mathcal{F}$, then probability space is called complete.

To define random variables we start with simple functions.

DEFINITION 1.1.6. Let $(\Omega, \mathcal{F})$ be a measurable space. A function $f : \Omega \to \mathbb{R}$ is called simple function if there exists $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ and $A_1, \ldots, A_n \in \mathcal{F}$ such that

$$f(\omega) = \sum_{i=1}^{n} \alpha_i \mathbb{1}_{A_i}(\omega),$$

where $\mathbb{1}_{\mathbb{A}}(\omega) = 1$ if $\omega \in A$ and 0 otherwise.

Definition 1.1.7. Let $(\Omega, \mathcal{F})$ be a measurable space and $f : \Omega \to \mathbb{R}$. If there is a sequence of simple functions $(f_n)_{n=1}^{\infty}$ such that

$$f(\omega) = \lim_{n\to\infty} f_n(\omega)$$

for all $\omega \in \Omega$, then function f is called measurable. If the measurable space has a probability measure $\mathbb{P}$, then the measurable function is called random variable.

We do not always have to show that a function is a random variable by using simple functions. For a measurable function there exists an equivalent condition which is often more useful.

Proposition 1.1.8 ([**7**] Proposition 3.1.3). *Let $(\Omega, \mathcal{F})$ be a measurable space and let $f : \Omega \to \mathbb{R}$ be a function. Then f is measurable if and only if*

$$f^{-1}((a,b)) := \{\omega \in \Omega : a < f(\omega) < b \in \mathcal{F}\},$$

for all $-\infty < a < b < \infty$.

An important concept in probability is the independence of random variables. It will be needed later.

Definition 1.1.9. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $f_i : \Omega \to \mathbb{R}$ be random variables for all $i = 1, 2, \ldots, n$. If for all $B_1, \ldots, B_n \in \mathfrak{B}(\mathbb{R})$ one has

$$\mathbb{P}(f_1 \in B_1, \ldots, f_n \in B_n) = \mathbb{P}(f_1 \in B_1) \ldots \mathbb{P}(f_n \in B_n),$$

then the random variables $f_1, \ldots, f_n$ are called independent.

1.2. Lebesgue integral

We assume that the reader is familiar with the Lebesgue integral and recall only some important tools. As a reference we recommend [**7**]. Our first tool is Dominated convergence.

Proposition 1.2.1 (Dominated convergence, ([**7**] Proposition 5.4.5)). *Let $(\Omega, \mathcal{F}, \mu)$ be a measure space and $g, f_1, f_2, \dots : \Omega \to \mathbb{R}$ be measurable functions such that $|f_n| \le g$ for all $n \in \mathbb{N}$. Assume that g is integrable and $\lim_{n\to\infty} f_n = f$. Then f is integrable and*

$$\lim_{n\to\infty} \int_\Omega f_n d\mu = \int_\Omega f d\mu.$$

Another tool that gives a possibility to calculate integrals and especially expectations explicitly is the Change of variable formula which we formulate for a probability space.

Proposition 1.2.2 (Change of variable, ([**7**] Proposition 5.6.1)). *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$ measurable space, $f : \Omega \to \mathbb{R}$ a random variable and $\varphi : \mathbb{R} \to \mathbb{R}$ Borel-measurable function. Assume that $\mathbb{P}_f$ is the distribution of f, meaning*

$$\mathbb{P}_f(B) = \mathbb{P}(\{\omega \in \Omega : f(\omega) \in B\}) = \mathbb{P}(f^{-1}(B)),$$

for all $B \in \mathbb{R}$. Then

$$\int_B \varphi d\mathbb{P}_f = \int_{f^{-1}(B)} \varphi(f) d\mathbb{P},$$

for all $B \in \mathbb{R}$.

1.3. Stochastic processes

Families of random variables play an important role in stochastic analysis and in our case we will give a definition for a continuous time interval $[0, T]$, where $T > 0$.

DEFINITION 1.3.1. Let $T > 0$ and $[0, T]$. Then a family of random variables $X = (X_t)_{t\in[0,T]}$ with $X_t : \Omega \to \mathbb{R}$ is called stochastic process with a continuous interval $[0, T]$.

We can think a σ-algebra as all information what we have. Next we introduce a definition which tells about the information what we have at some time point.

DEFINITION 1.3.2. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Then the family of σ-algebras $(\mathcal{F}_t)_{t\in[0,T]}$ is called filtration if $\mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F}$ for all $0 \le s \le t \le T$ and $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\in[0,T]})$ is called stochastic basis.

Using a filtration one can say something about types of measurability of stochastic processes.

DEFINITION 1.3.3. Let $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\in[0,T]})$ be a stochastic basis and $X = (X_t)_{t\in[0,T]}$, $X_t : \Omega \to \mathbb{R}$ a stochastic process. Then

(1) The process X is called measurable if the function $(\omega, t) \mapsto X_t(\omega)$ seen as a map between $\Omega \times [0, T]$ and $\mathbb{R}$ is measurable with respect to $\mathcal{F} \otimes \mathfrak{B}([0, T])$ and $\mathfrak{B}(\mathbb{R})$.
(2) The process X is called progressively measurable with respect to $(\mathcal{F}_t)_{t\in[0,T]}$ if for all $s \in [0, T]$ the function $(\omega, t) \mapsto X_t(\omega)$ seen as a map between $\Omega \times [0, s]$ and $\mathbb{R}$ is measurable with respect to $\mathcal{F}_s \otimes \mathfrak{B}([0, s])$ and $\mathfrak{B}(\mathbb{R})$.
(3) The process X is called adapted with respect to $(\mathcal{F}_t)_{t\in[0,T]}$ if for all $t \in [0, T]$ the random variable X_t is $\mathcal{F}_t$-measurable.

Between these three different kinds of measurability one has the following connections.

PROPOSITION 1.3.4 ([**10**] Propositions 2.1.10. and 2.1.11.). *The following holds*

(1) A progressively measurable process is measurable and adapted.
(2) An adapted process with left- or right-continuous paths is progressively measurable.

Next we look at one famous stochastic process. It was first observed by Robert Brown when he was looking at pollen in water by a microscope. He realized that particles move incessant and irregular and published papers 1928 and 1929 about this movement. After this observation 1905 Albert Einstein gave a correct explanation for this phenomenon. From the perspective of mathematics in 1900 Louis Bachelier gave a first, but not rigorous, definition for Brownian motion when he studied fluctuation of stock prices. This was without a connection to Brownian motion in physics. The first rigorous mathematical construction was given by Norbert Wiener in 1923.

DEFINITION 1.3.5 (Brownian motion, ([**20**] Definition 1.2.1)). Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space.

(1) A process $W = (W_t)_{t\in[0,T]}$ with $W_0 = 0$ is called standard Brownian motion if
  (a) $(W_t)_{t\in[0,T]}$ is continuous.
  (b) For all $n \in \mathbb{N}$ and $0 \leq t_1 \leq t_2 \leq \cdots \leq t_n = T$ the increments $W_{t_n} - W_{t_{n-1}}, \ldots, W_{t_2} - W_{t_1}$ are independent.
  (c) For all $0 \leq s \leq t \leq T$ it holds $W_{t-s} \sim N(0, t-s)$.
(2) An $\mathbb{R}^d$-valued stochastic process $W = (W_t)_{t\in[0,T]}$, $W_t = (W_t^1, \ldots, W_t^d)$, where $W^1, \ldots, W^d$ are Brownian motions independent from each other, is called a d-dimensional Brownian motion.

Information from Brownian motion can be collected to a special σ-algebra.

Definition 1.3.6. Let $W = (W_t)_{t\in[0,T]}$ be a Brownian motion which generates σ-algebra for all $t \in [0,T]$

$$\mathcal{F}_t^W = \sigma(W_s : 0 \leq s \leq t).$$

If

$$\mathcal{N} = \left\{A \subseteq \Omega : \text{there exist a } B \in \mathcal{F}_T^W \text{ such that } A \subseteq B \text{ and } \mathbb{P}(B) = 0\right\},$$

then $(\mathcal{F}_t)_{t\in[0,T]}$ with $\mathcal{F}_t = \mathcal{F}_t^W \vee \mathcal{N}$ is called augmentation of $(\mathcal{F}_t^W)_{t\in[0,T]}$.

1.4. Conditional expectation and martingales

Assume a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, a random variable $f : \Omega \to \mathbb{R}$ and another σ-algebra $\mathcal{G}$ which is included in $\mathcal{F}$. Then if the random variable f is $\mathcal{F}$-measurable it is not always $\mathcal{G}$-measurable. The following proposition and definition introduce the concept of conditional expectation.

Proposition 1.4.1 ([**7**] Proposition 7.3.1). *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, $\mathcal{G} \subseteq \mathcal{F}$ be a sub-σ-algebra and f a random variable such that $f \in L(\Omega, \mathcal{F}, \mathbb{P})$. Then*
(1) There exists a random variable $g \in L(\Omega, \mathcal{G}, \mathbb{P})$ such that

$$\int_B f d\mathbb{P} = \int_B g d\mathbb{P}, \text{ for all } B \in \mathcal{G}.$$

(2) If there is g and g' such that for both above hold, then

$$\mathbb{P}(g \neq g') = 0.$$

Definition 1.4.2 ([**7**] Definition 7.3.2). The random variable $g \in L(\Omega, \mathcal{G}, \mathbb{P})$ in Proposition 1.4.1 is called conditional expectation of f given $\mathcal{G}$. It is denoted by

$$g = \mathbb{E}[f|\mathcal{G}].$$

It is good to notice that the conditional expectation can be changed on sets with probability zero so it is unique only almost surely. The conditional expectation has many properties which are listed below. For later use especially the property called tower property is useful for us.

Proposition 1.4.3 ([**7**] Proposition 7.3.3 (1)-(9)). *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $\mathcal{H} \subseteq \mathcal{G} \subseteq \mathcal{F}$ sub-σ-algebras of $\mathcal{F}$. Assume $f, g \in L(\Omega, \mathcal{F}, \mathbb{P})$. Then we have the following properties almost surely:*

(1) *Linearity:* *Let* $\lambda, \mu \in \mathbb{R}$. *Then*
$$\mathbb{E}\left[\lambda f + \mu f | \mathcal{G}\right] = \lambda \mathbb{E}\left[f|\mathcal{G}\right] + \mu \mathbb{E}\left[g|\mathcal{G}\right].$$
(2) *Monotonicity:* *Let* $g \leq f$ *almost surely. Then* $\mathbb{E}\left[g|\mathcal{G}\right] \leq \mathbb{E}\left[f|\mathcal{G}\right]$.
(3) *Positivity:* *Let* $f \geq 0$ *almost surely. Then* $\mathbb{E}\left[f|\mathcal{G}\right] \geq 0$.
(4) *Convexity:* $|\mathbb{E}\left[f|\mathcal{G}\right]| \leq \mathbb{E}\left[|f||\mathcal{G}\right]$.
(5) *Projection property:* *Let* f *be* $\mathcal{G}$*-measurable. Then* $\mathbb{E}\left[f|\mathcal{G}\right] = f$.
(6) *Tower property:* $\mathbb{E}\left[\mathbb{E}\left[f|\mathcal{G}\right]|\mathcal{H}\right] = \mathbb{E}\left[\mathbb{E}\left[f|\mathcal{H}\right]|\mathcal{G}\right] = \mathbb{E}\left[f|\mathcal{H}\right]$.
(7) *Let* $h : \Omega \to \mathbb{R}$ *be* $\mathcal{G}$*-measurable and* $fh \in L(\Omega, \mathcal{F}, \mathbb{P})$. *Then*
$$\mathbb{E}\left[fh|\mathcal{G}\right] = h\mathbb{E}\left[f|\mathcal{G}\right].$$
(8) *Let* $\mathcal{G} = \{\emptyset, \Omega\}$. *Then* $\mathbb{E}\left[f|\mathcal{G}\right] = \mathbb{E}[f]$.
(9) *Let* f *be independent from* $\mathcal{G}$. *Then* $\mathbb{E}\left[f|\mathcal{G}\right] = \mathbb{E}[f]$.

For the conditional expectation we have a similar property as Proposition 1.2.1 states for the Lebesgue integral. It is called Dominated convergence for conditional expectation.

Proposition 1.4.4 ([**1**] Equation (15.14)). *Let* $(\Omega, \mathcal{F}, \mathbb{P})$ *be a probability space,* $\mathcal{G} \subseteq \mathcal{F}$ *be a sub-*σ*-algebra and* $(X_n)_{n\in\mathbb{N}}$ *a sequence of random variables such that* $|X_n| \leq Y$ *for all* $n \in \mathbb{N}$ *and for some integrable random variable* Y*. Assume also that* $X_n \to X$ *almost surely when* $n \to \infty$. *Then*
$$\lim_{n\to\infty} \mathbb{E}\left[X_n|\mathcal{G}\right] = \mathbb{E}\left[X|\mathcal{G}\right].$$

The conditional expectation is used in the definition of martingales, which are important processes in the field of stochastics.

Definition 1.4.5. Let $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\in[0,T]})$ be a stochastic basis. A stochastic process $M = (M_t)_{t\in[0,T]}$ is called martingale with respect to a filtration $(\mathcal{F}_t)_{t\in[0,T]}$ if
(1) M_t is $\mathcal{F}_t$-measurable for all $t \in [0,T]$
(2) $\mathbb{E}|M_t| < \infty$ for all $t \in [0,T]$
(3) $\mathbb{E}[M_t|\mathcal{F}_s] = M_s$ for all $0 \leq s \leq t \leq T$.
Moreover, a martingale $M = (M_t)_{t\in[0,T]}$ belongs to $\mathcal{M}_2^c$ if
(1) $\mathbb{E}|M_t|^2 < \infty$ for all $t \in [0,T]$
(2) the paths $t \to M_t(\omega)$ are continuous for all $\omega \in \Omega$.
The space $\mathcal{M}_2^{c,0}$ consists of all $M \in \mathcal{M}_2^c$ with $M_0 = 0$.

Sometimes the set $\mathcal{M}_2^c$ is not large enough, so we need also a definition for a larger set of processes. For that a map called stopping time is needed.

Definition 1.4.6. Let $(\Omega, \mathcal{F})$ be a measurable space with filtration $(\mathcal{F}_t)_{t\in[0,T]}$. Then the map $\tau : \Omega \to [0,T]$ is called stopping time with respect to the filtration $(\mathcal{F}_t)_{t\in[0,T]}$ if
$$\{\tau \leq t\} \in \mathcal{F}_t, \text{ for all } t \in [0,T].$$

Definition 1.4.7. Let $M = (M_t)_{t\in[0,T]}$ be a continuous and adapted process with $M_0 = 0$. If there exists an increasing sequence $(\tau_n)_{n=0}^{\infty}$ of stopping times with $\lim_{n\to\infty} \tau_n(\omega) = \infty$ for all $\omega \in \Omega$ such that $\mathcal{M}^{\tau_n} = (\mathcal{M}_{t\wedge\tau_n})_{t\in[0,T]}$ is martingale for all $n \in \mathbb{N}$, then the process M is called local martingale. Moreover, the set of local martingales is denoted by $\mathcal{M}_{loc}^{c,0}$.

We have a sufficient condition for a local martingale to be a martingale.

LEMMA 1.4.8. *Let $(M_t)_{t\in[0,T]}$ be a (continuous) local martingale and G such that*

$$\sup_{t\in[0,T]} |M_t| < G \text{ and } \mathbb{E}G < \infty.$$

Then $(M_t)_{t\in[0,T]}$ is martingale.

PROOF. Let $(\tau_N)_{N\geq 1}$ be a localizing sequence. Now we have for $0 \leq s < t \leq T$ that

$$\mathbb{E}\left[M_{t\wedge\tau_N}|\mathcal{F}_s\right] = M_{s\wedge\tau_N}.$$

By dominated convergence for conditional expectation (Proposition 1.4.4) we get

$$\begin{aligned}\lim_{N\to\infty} \mathbb{E}\left[M_{t\wedge\tau_N}|\mathcal{F}_s\right] &= \mathbb{E}\left[\lim_{N\to\infty} M_{t\wedge\tau_N}|\mathcal{F}_s\right] \\ &= \mathbb{E}\left[M_t|\mathcal{F}_s\right].\end{aligned}$$

On the other hand we have

$$\lim_{N\to\infty} M_{s\wedge\tau_N} = M_s.$$

So we conclude

$$\mathbb{E}\left[M_t|\mathcal{F}_s\right] = M_s.$$

□

1.5. Itô integral and Itô's formula

In this section we recall the Itô integral. We assume the usual conditions meaning that we have a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a right continuous filtration $(\mathcal{F}_t)_{t\in[0,T]}$ which means $\cap_{\epsilon>0}\mathcal{F}_{t+\epsilon} = \mathcal{F}_t$ for all $t \in [0,T]$. We also assume that all sets of probability zero are included in $\mathcal{F}_0$ and $W = (W_t)_{t\in[0,T]}$ is a standard Brownian motion. We give only the definitions and main properties of the Itô integral. For more information for example [**10**] is recommend. First we need simple stochastic processes and a stochastic integral for them.

DEFINITION 1.5.1. Let $L = (L_t)_{t\in[0,T]}$ be a stochastic process. It is called simple if there exist

(1) a sequence of time points such that $0 = t_0 < t_1 < \cdots < t_n = T$ and
(2) $\mathcal{F}_{t_i}$-measurable bounded random variables $\upsilon_i : \Omega \to \mathbb{R}$,

such that L has a representation

$$L_t(\omega) = \sum_{i=1}^{n} \upsilon_{i-1}(\omega)\mathbb{1}_{(t_{i-1},t_i]}(t).$$

The set of all simple processes is denoted by $\mathcal{L}_0$.

DEFINITION 1.5.2. Let $L \in \mathcal{L}_0$ and $t \in [0,T]$. Then stochastic integral is defined as

$$I_t(L)(\omega) = \sum_{i=1}^{n} \upsilon_{i-1}(\omega)\left(W_{t_i\wedge t}(\omega) - W_{t_{i-1}\wedge t}(\omega)\right).$$

The stochastic integral of a simple process is a continuous, square integrable martingale as the next proposition states.

Proposition 1.5.3. *Let $L \in \mathcal{L}_0$. Then $I(L) = (I_t(L)_{t\in[0,T]}) \in \mathcal{M}_2^{c,0}$.*

Stochastic integrals of simple processes can be extended to a larger set of integrands.

Definition 1.5.4. The set of all progressively measurable stochastic processes $L = (L_t)_{t\in[0,T]}, L_t : \Omega \to \mathbb{R}$ with

$$\|L\|_{\mathcal{L}_2,t} = \left(\mathbb{E}\int_0^t L_u^2 du\right)^{\frac{1}{2}} < \infty \text{ for all } t \in [0,T],$$

is denoted by $\mathcal{L}_2$.

Proposition 1.5.5 ([**10**]Proposition 3.1.12 (i)-(v)). *The map $I : \mathcal{L}_0 \to \mathcal{M}_2^{c,0}$ can be extended to $J : \mathcal{L}_2 \to \mathcal{M}_2^{c,0}$ with properties:*

(1) Linearity: Let $\alpha, \beta \in \mathbb{R}$ and $L, K \in \mathcal{L}_2$. Then

$$J_t(\alpha L + \beta K) = \alpha J_t(L) + \beta J_t(K) \text{ a.s. for } t \in [0,T].$$

(2) Extension property: Let $L \in \mathcal{L}_0$. Then $I_t(L) = J_t(L)$ a.s. for $t \in [0,T]$.

(3) Itô isometry: Let $L \in \mathcal{L}_2$. Then

$$\left(\mathbb{E}\left[J_t(L)^2\right]\right)^{\frac{1}{2}} = \left(\mathbb{E}\int_0^t L_u^2 du\right)^{\frac{1}{2}} \text{ for } t \in [0,T].$$

(4) Continuity property: Let $(K^{(n)})_{n=1}^\infty$ be a sequence where $K^{(n)} \in \mathcal{L}_2$ for all $n \in \mathbb{N}$. Let $L \in \mathcal{L}_2$. If $d(K^{(n)}, L) = \sum_{m=1}^\infty 2^{-m} \min\left\{1, \|K^{(n)} - L\|_{\mathcal{L}_2,m}\right\} \to 0$ when $n \to \infty$, then

$$\mathbb{E}\left[\sup_{t\in[0,T]} |J_t(L) - J_t(K^{(n)})|^2\right] \to 0$$

as $n \to \infty$.

(5) Uniqueness: If $\hat{J} : \mathcal{L}_2 \to \mathcal{M}_2^{c,0}$ is another mapping for which above properties hold, then

$$\mathbb{P}\left(J_t(L) = \hat{J}_t(L), t \in [0,T]\right) = 1,$$

for all $L \in \mathcal{L}_2$.

For $L \in \mathcal{L}_2$ we have that $J(L)$ is a square integrable martingale. But this property is vanishing in the next extension.

Definition 1.5.6. (1) The set of progressively measurable stochastic process $L = (L_t)_{t\in[0,T]}$ with

$$\mathbb{P}\left(\omega \in \Omega : \int_0^T L_u(\omega)^2 du < \infty\right) = 1$$

is denoted by $\mathcal{L}_2^{loc}$.

(2) Let $(\tau_n)_{n=1}^\infty$ be a sequence of stopping times. It is called localizing for $L = (L_t)_{t\in[0,T]} \in \mathcal{L}_2^{loc}$ if
 (a) $0 \le \tau_0(\omega) \le \tau_1(\omega) \le \cdots \le T$ with $\lim_{n\to\infty} \tau_n = T$ for all $\omega \in \Omega$ and
 (b) $L^{\tau_n} = L_t \mathbb{1}_{t\le\tau_n} \in \mathcal{L}_2$ for all $n = 0, 1, 2, \ldots$

The next lemma states the existence of a stochastic integral for processes in $\mathcal{L}_2^{loc}$.

LEMMA 1.5.7 ([**10**] Lemma 3.1.19). *Let $L \in \mathcal{L}_2^{loc}$. Then there is a unique and adapted process $X = (X_t)_{t\in[0,T]}$ with $X_0 = 0$ such that*

$$\mathbb{P}\left(J_t(L^{\tau_n}) = X_t, t \in [0, \tau_n]\right) = 1,$$

for all localizing sequences $(\tau_n)_{n=0}^{\infty}$ of L and for all $n = 1, 2, \ldots$.

REMARK 1.5.8. The uniquenes in Lemma 1.5.7 means that if there is a $Y = (Y_t)_{t\in[0,T]}$ with the same properties, then $\mathbb{P}\left(X_t = Y_t, t \in [0, T]\right) = 1$.

DEFINITION 1.5.9. Let $L \in \mathcal{L}_2^{loc}$. Then the process X in Lemma 1.5.7 is called Ito integral and it is denoted by

$$X = \left(\int_0^t L_u dW_u\right)_{t\in[0,T]}.$$

For integrands from L_2^{loc} the Itô integral is a local martingale:

PROPOSITION 1.5.10 ([**10**]Proposition 3.1.23 (i)). *Let $L \in \mathcal{L}_2^{loc}$. Then one has that $\left(\int_0^t L_u dW_u\right)_{t\in[0,T]} \in \mathcal{M}_{loc}^{c,0}$.*

In many situations the task is to show that the integrand is in $\mathcal{L}_2$ implying that the Itô integral is a martingale. As a notation we get for the Itô isometry in Proposition 1.5.5

$$\left(\mathbb{E}\left[\left(\int_0^t L_u dW_u\right)^2\right]\right)^{\frac{1}{2}} = \left(\mathbb{E}\int_0^t L_u^2 du\right)^{\frac{1}{2}} \text{ for } t \in [0, T].$$

A very useful tool in stochastic analysis is called Itô's formula which allows us to write stochastic processes in different form. To recall this we need first some definitions.

DEFINITION 1.5.11. Let $A = (A_t)_{t\in[0,T]}$, $A_t : \Omega \to \mathbb{R}$ be a stochastic process such that

$$\sup_{n\in\mathbb{N}} \sup_{0=t_0\leq t_1\leq\ldots\leq t_n=t} \sum_{k=1}^{n} |A_{t_k}(\omega) - A_{t_{k-1}}(\omega)| < \infty \text{ a.s. for all } t \in [0, T].$$

Then the process A is said to be of bounded variation.

DEFINITION 1.5.12. Let $X = (X_t)_{t\in[0,T]}$ be a continuous and adapted stochastic process. If there exist $x_0 \in \mathbb{R}$, $L \in \mathcal{L}_2^{loc}$ and a progressively measurable process $a = (a_t)_{t\in[0,T]}$ with

$$\int_0^t |a_u(\omega)| du < \infty$$

for all $t \in [0, T]$ and $\omega \in \Omega$ such that X can be represented as

$$X_t = x_0 + \left(\int_0^t L_u dW_u\right)(\omega) + \int_0^t a_u du \text{ a.s. for all } t \in [0, T],$$

then X is called Itô process.

Definition 1.5.13. Let $X = (X_t)_{t\in[0,T]}$ be a continuous and adapted process. If there exist $x_0 \in \mathbb{R}$, $M \in \mathcal{M}_2^{loc}$ and a process A of bounded variation with $A_0 = 0$ such that

$$X_t = x_0 + M_t + A_t,$$

then X is called a continuous semi-martingale.

Remark 1.5.14. Especially Itô processes are continuous semi-martingales since

$$\begin{aligned}\sup_{0=t_0\le t_1\le\ldots\le t_n=t}\sum_{k=1}^{n}\left|\int_0^{t_k} a_u du - \int_0^{t_{k-1}} a_u du\right| &= \sup_{0=t_0\le t_1\le\ldots\le t_n=t}\sum_{k=1}^{n}\left|\int_{t_{k-1}}^{t_k} a_u du\right| \\ &\le \sup_{0=t_0\le t_1\le\ldots\le t_n=t}\sum_{k=1}^{n}\int_{t_{k-1}}^{t_k} |a_u| du \\ &= \int_0^t |a_u| du < \infty.\end{aligned}$$

We give a version of Itô's formula which is for continuous semi-martingales.

Proposition 1.5.15 ([**10**]Proposition 3.4.3). *Let $f \in C^2(\mathbb{R}^d)$ and $X_t = (X_t^1, \ldots, X_t^d)$ be a vector of continuous semi-martingales. Then almost surely*

$$f(X_t) = f(X_0) + \sum_{i=1}^{d}\int_0^t \frac{\partial f}{\partial x_i}(X_u)dX_u^i + \frac{1}{2}\sum_{i,j=1}^{d}\int_0^t \frac{\partial^2 f}{\partial x_i \partial x_j}(X_u)d\left\langle M^i, M^j\right\rangle_u,$$

where $dX_u^i = dM_u^i + dA_u^i$ and $\langle M^i, M^j\rangle_u = \frac{1}{4}\left[\langle M^i + M^j\rangle_u - \langle M^i - M^j\rangle_u\right]$ is called cross-variation.

To make it possible to use explicitly above Itô's formula, we need also the following proposition.

Proposition 1.5.16 ([**10**]Proposition 4.4.3.). *Let $L \in \mathcal{L}_2^{loc}$. Then*

$$\left\langle \int_0^{\cdot} L_u dW_u \right\rangle_t = \int_0^t L_u^2 du,$$

for all $t \in [0,T]$ almost surely.

We had before the Lemma 1.4.8 which provided a condition when a local martingale is a martingale. We are able to give a similar condition called Novikov's condition for the exponential martingale.

Proposition 1.5.17 ([**10**]Proposition 4.4.8). *Let $L \in \mathcal{L}_2^{loc}$ and $t \in [0,T]$. Then*

$$e^{\int_0^t L_u dW_u - \frac{1}{2}\int_0^t L_u^2 du}$$

is a martingale if

$$\mathbb{E}\left[e^{\frac{1}{2}\int_0^T L_u^2 du}\right] < \infty.$$

1.6. Stochastic differential equations

In this section we focus on stochastic differential equations, also called SDEs, and state a proposition about existence and uniqueness of solutions. For more information about solving SDEs we recommend for example [**13**]. We assume again that the usual conditions hold as in the previous section but now we assume $(W_t)_{t\in[0,T]}, W_t = (W_t^1, \dots, W_t^d)$ to be a d-dimensional standard Brownian motion adapted to the filtration $(\mathcal{F}_t)_{t\in[0,T]}$.

Let b be a d-dimensional vector of functions and σ be a $d \times k$-dimensional matrix of functions. For b_j and σ_{ij}, $j \in \{1, 2, \dots, d\}$, $i \in \{1, 2, \dots, k\}$ we assume that

(1) $b_j, \sigma_{ij} : [0, T] \times \mathbb{R}^d \times \Omega \to \mathbb{R}$.
(2) b_j and σ_{ij} are $\mathfrak{B}([0, T]) \times \mathfrak{B}(\mathbb{R}^d) \times \mathcal{F}$-measurable.
(3) For all $t \in [0, T]$, $b_j(t, \cdot, \cdot)$ and $\sigma_{ij}(t, \cdot, \cdot)$ are measurable with respect to $\mathfrak{B}(\mathbb{R}^d) \times \mathcal{F}_t$.

Definition 1.6.1. Let $x_0 \in \mathbb{R}^d$. Then

$$\begin{cases} dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t \\ X_0 = x_0 \end{cases} \tag{1.1}$$

is called stochastic differential equation. The d-dimensional stochastic process $X = (X_t)_{t\in[0,T]}$ is called a strong solution if the following holds:

(1) X is $\mathcal{F}_t$ -adapted and has continuous sample paths.
(2) $\int_0^T \left(|b(t, X_t)| + |\sigma(t, X_t)|^2\right) dt < \infty$ a.s. where $|\cdot|$ denotes both the d-dimensional norm and the norm of a matrix.
(3) For all $t \in [0, T]$

$$X_t = x_0 + \int_0^t b(s, X_s)ds + \int_0^t \sigma(s, X_s)dW_s \text{ a.s.}$$

Does an SDE always have a solution? We provide a property when there exist a unique strong solution.

Proposition 1.6.2 ([**13**] Theorem 6.2.1). *Assume SDE* (1.1). *If for b and σ the following holds*

(1) $|b(t, x)|^2 + |\sigma(t, x)|^2 \le K(1 + |x|^2)$ *a.s.*
(2) $|b(t, x) - b(t, y)|^2 + |\sigma(t, x) - \sigma(t, y)|^2 \le K|x - y|^2$ *a.s.*

for all $t \in [0, T]$, $x, y \in \mathbb{R}^d$ and some constant $K > 0$, then the SDE has a unique strong solution $X = (X_t)_{t\in[0,T]}$.

Remark 1.6.3. Uniqueness in the above proposition means that if there exists another strong solution $Y = (Y_t)_{t\in[0,T]}$ then

$$\mathbb{P}(X_t = Y_t, t \in [0, T]) = 1.$$

Remark 1.6.4. From [**13**] (proof of Theorem 6.2.1) we get that

$$\mathbb{E}\left[\sup_{t\in[0,T]} |X_t|^2\right] < \infty.$$

We conclude this section with some useful inequalities called Burkholder-Davis-Gundy inequality and Hölder inequality.

PROPOSITION 1.6.5 (Burkholder-Davis-Gundy, ([**10**]Theorem 4.3.1)). *Let* $L \in \mathcal{L}_2^{loc}$. *Then for any* $0 < p < \infty$ *there exist constants* $\alpha_p, \beta_p > 0$ *such that*

$$\beta_p \left\| \sqrt{\int_0^T L_t^2 dt} \right\|_p \leq \left\| \sup_{t\in[0,T]} \left| \int_0^t L_s dW_s \right| \right\|_p \leq \alpha_p \left\| \sqrt{\int_0^T L_t^2 dt} \right\|_p .$$

Moreover, $\alpha_p \leq c\sqrt{p}$ *for* $2 \leq p < \infty$ *and for some constant* $c > 0$.

The Hölder inequality we state for a probability space.

PROPOSITION 1.6.6 (Hölder, ([**7**] Proposition 5.10.5)). *Let* $(\Omega, \mathcal{F}, \mathbb{P})$ *be a probability space and* $X, Y : \Omega \to \mathbb{R}$ *random variables. If* $1 < p, q < \infty$ *and* $\frac{1}{p} + \frac{1}{q} = 1$, *then*

$$\mathbb{E}\left[|XY|\right] \leq \left(\mathbb{E}\left[X^p\right]\right)^{\frac{1}{p}} \left(\mathbb{E}\left[Y^q\right]\right)^{\frac{1}{q}} .$$

CHAPTER 2

Malliavin calculus

2.1. A multiple stochastic integral

Next we will follow Nualart [**18**] and give the basics of Malliavin calculus.

Definition 2.1.1. Let $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\in[0,T]})$ be a filtered probability space and H a real and separable Hilbert space with scalar product $\langle \cdot, \cdot \rangle_H$ and norm $\| \cdot \|_H$. We consider a stochastic process indexed by the elements of H:

$$W = \{W(h); h \in H\}.$$

We say that this process is an isonormal Gaussian process if W is a Gaussian family of random variables such that for all $h, g \in H$

$$\mathbb{E}\left[W(h)W(g)\right] = \langle h, g \rangle_H$$

and

$$\mathbb{E}\left[W(h)\right] = 0.$$

In this thesis we realize such an isonormal Gaussian process as follows. We assume a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ carrying a Brownian motion $W = (W_t)_{t\in[0,T]}$ and a special Hilbert space with

$$L^2\big([0,T], \mathfrak{B}([0,T]), \lambda\big) = \left\{ f : [0,T] \to \mathbb{R}; \left(\int_0^T f(t)^2 d\lambda(t) \right)^{\frac{1}{2}} < \infty \right\}.$$

For all $f, g \in L^2([0,T])$, we let $W(f) = \int_0^T f(s) dW_s$ and $\langle f, g \rangle_{L^2([0,T])} = \mathbb{E}[W(f)W(g)]$. Moreover, for all $A \in \mathfrak{B}([0,T])$, we let $W(A) = W(\mathbb{1}_A)$. In this way, we have an isonormal Gaussian process as we defined above.

Next we define a set of special elementary functions, which will vanish on diagonals.

Definition 2.1.2. Let $A_1, \ldots, A_n \in \mathfrak{B}([0,T])$ such that $A_k \cap A_l = \emptyset$ for all $k \neq l$, where $k, l \in \{1, \ldots, n\}$. Then we define

$$\mathcal{E}_m = \left\{ \begin{array}{l} f : [0,T]^m \to \mathbb{R}; f(t_1, \ldots, t_m) = \sum_{i_1,\ldots,i_m=1}^n a_{i_1\cdots i_m} \mathbb{1}_{A_{i_1}\times\cdots\times A_{i_m}}(t_1, \ldots, t_m) \\ \text{where } a_i \in \mathbb{R} \text{ and } a_{i_1\cdots i_m} = 0, \text{ if } i_k = i_j \text{ for some } k \neq j. \end{array} \right\}$$

We define a multiple stochastic integral first for these elementary functions and after that for all functions in $L^2([0,T]^m)$.

Definition 2.1.3. For $f \in \mathcal{E}_m$ given as above a multiple stochastic integral is defined as

$$I_m(f) = \sum_{i_1,\ldots,i_m=1}^n a_{i_1\cdots i_m} W(A_{i_1}) \cdots W(A_{i_m}).$$

REMARK 2.1.4. One can easily check that this definition is independent from the representation of f. Hence I_m is well-defined.

This defined multiple stochastic integral has three important properties. Before we state the properties, we introduce the symmetrization $\tilde{f}$ of a function f. This is

$$\tilde{f}(t_1,\ldots,t_m)=\frac{1}{m!}\sum_{\sigma\in S_m} f(t_{\sigma(1)},\ldots,t_{\sigma(m)}),$$

where S_m is the set of all permutations of $\{1,\ldots,m\}$.

PROPOSITION 2.1.5. *A multiple stochastic integral with integrands from $\mathcal{E}_m$ has the following properties*

(i) Let $f,g\in\mathcal{E}_m$ and $\alpha,\beta\in\mathbb{R}$. Then

$$I_m(\alpha f+\beta g)=\alpha I_m(f)+\beta I_m(g).$$

(ii) If $\tilde{f}$ is symmetrization of f, then

$$I_m(\tilde{f})=I_m(f).$$

(iii) If $f\in\mathcal{E}_m$ and $g\in\mathcal{E}_n$, then for the product of multiple stochastic integrals it holds

$$\mathbb{E}\left[I_m(f)I_n(g)\right]=\begin{cases}0 & \text{if } m\neq n\\ m!\langle\tilde{f},\tilde{g}\rangle_{L^2_{[0,T]^m}} & \text{if } m=n.\end{cases}$$

PROOF. (i) We can assume that f and g have the same partition $A_1,\ldots,A_n$, because if not, we can make it to be the same by using intersections of sets. Now

$$\begin{aligned}\alpha I_m(f)+\beta I_m(g) &= \alpha\sum_{i_1,\ldots,i_m=1}^{n} a_{i_1\cdots i_m}W(A_{i_1})\ldots W(A_{i_m})\\ &\quad+\beta\sum_{i_1,\ldots,i_m=1}^{n} b_{i_1\cdots i_m}W(A_{i_1})\ldots W(A_{i_m})\\ &= \sum_{i_1,\ldots,i_m=1}^{n}[\alpha a_{i_1\cdots i_m}+\beta b_{i_1\cdots i_m}]W(A_{i_1})\ldots W(A_{i_m})\\ &= I_m(\alpha f+\beta g)\end{aligned}$$

(ii) Because of (i) we can assume a function

$$f(t_1,\ldots,t_m)=\mathbb{1}_{A_{i_1}\times\cdots\times A_{i_m}}(t_1,\ldots,t_m).$$

For this we have by definition

$$I_m(f)=W(A_{i_1})\ldots W(A_{i_m}),$$

and for the symmetrization we have

$$I_m(\tilde{f})=\frac{1}{m!}\sum_{\sigma\in S_m} W(A_{\sigma(1)})\ldots W(A_{\sigma(m)}).$$

Now for all permutations σ we can change order in the product such that we always get

$$W(A_{\sigma(1)})\dots W(A_{\sigma(m)}) = W(A_{i_1})\dots W(A_{i_m}).$$

By that we have

$$I_m(\tilde{f}) = \frac{1}{m!} m! W(A_{i_1})\dots W(A_{i_m}) = I_m(f).$$

(iii) Let $f \in \mathcal{E}_m$ and $g \in \mathcal{E}_k$. Because of (i) and (ii), we can assume that both functions are symmetric and have the same partition $A_1, \dots, A_n$. We have now

$$f(t_1, \dots, t_m) = \sum_{i_1,\dots,i_m=1}^{n} a_{i_1\cdots i_m} \mathbb{1}_{A_{i_1}\times\cdots\times A_{i_m}}(t_1, \dots, t_m)$$

and

$$g(t_1, \dots, t_k) = \sum_{j_1,\dots,j_k=1}^{n} b_{j_1\cdots j_k} \mathbb{1}_{A_{j_1}\times\cdots\times A_{j_k}}(t_1, \dots, t_k).$$

Because the functions are symmetric, we have for all permutations

$$a_{i_1\cdots i_m} = a_{\sigma(i_1)\cdots\sigma(i_m)} \text{ and } b_{j_1\cdots j_k} = b_{\sigma(j_1)\cdots\sigma(j_k)},$$

and because we have $m!$ permutations for the function f and $k!$ for g, we get

$$I_m(f) = m! \sum_{i_1<\cdots<i_m} a_{i_1\cdots i_m} W(A_{i_1})\dots W(A_{i_m})$$

and

$$I_k(g) = k! \sum_{j_1<\cdots<j_k} b_{j_1\cdots j_k} W(A_{j_1})\dots W(A_{j_k}).$$

With these we get

$$\begin{aligned}
\mathbb{E}\left[I_m(f)I_k(g)\right] &= \mathbb{E}\Bigg[m! \sum_{i_1<\cdots<i_m} a_{i_1\cdots i_m} W(A_{i_1})\dots W(A_{i_m}) \\
&\qquad \times k! \sum_{j_1<\cdots<j_k} b_{j_1\cdots j_k} W(A_{j_1})\dots W(A_{j_k})\Bigg] \\
&= m!k! \sum_{i_1<\cdots<i_m} \sum_{j_1<\cdots<j_k} a_{i_1\cdots i_m} b_{j_1\cdots j_k} \\
&\qquad \times \mathbb{E}\left[W(A_{i_1})\dots W(A_{i_m})W(A_{j_1})\dots W(A_{j_k})\right].
\end{aligned}$$

By Definition 2.1.2 we have $A_i \cap A_j = \emptyset$ for all $i \neq j$ so this implies that $W(A_i)$ and $W(A_j)$ are independet for all $i \neq j$. We have now two possibilities:

$$\mathbb{E}\left[W(A_{i_1})...W(A_{i_m})W(A_{j_1})...W(A_{j_k})\right] = \begin{cases} \mathbb{E}[W(A_{i_1})^2]...\mathbb{E}[W(A_{i_m})^2], \text{ if } m = k \\ \qquad \text{and } i_q = j_q \text{ for all } q \in \{1, 2, ..., m\}, \\ 0, \text{ all other cases.} \end{cases}$$

So we conclude if $m \neq k$

$$\mathbb{E}\left[I_m(f)I_k(g)\right] = 0.$$

And for $m = k$ we get by Itô's isometry and the fact that

$$\mathbb{E}\left[W(A_{i_1})\dots W(A_{i_m})W(A_{j_1})\dots W(A_{j_m})\right] = 0$$

unless $i_q = j_q$ for all $q \in \{1, 2, \dots, m\}$ the following:

$$\begin{aligned}
& \mathbb{E}\left[I_m(f)I_m(g)\right] \\
= \ & \mathbb{E}\left[m! \sum_{i_1<\dots<i_m} a_{i_1\cdots i_m} W(A_{i_1})\dots W(A_{i_m}) m! \sum_{j_1<\dots<j_m} b_{j_1\cdots j_m} W(A_{j_1})\dots W(A_{j_m})\right] \\
= \ & \mathbb{E}\left[(m!)^2 \sum_{i_1<\dots<i_m} a_{i_1\cdots i_m} b_{i_1\cdots i_m} W(A_{i_1})^2 \dots W(A_{1_m})^2\right] \\
= \ & (m!)^2 \sum_{i_1<\dots<i_m} a_{i_1\cdots i_m} b_{i_1\cdots i_m} \mathbb{E}\left[W(A_{i_1})^2\right]\dots \mathbb{E}\left[W(A_{i_m})^2\right] \\
= \ & (m!)^2 \sum_{i_1<\dots<i_m} a_{i_1\cdots i_m} b_{i_1\cdots i_m} \lambda(A_{i_1})\dots\lambda(A_{i_m}) \\
= \ & m! \sum_{i_1,\dots,i_m=1}^{n} a_{i_1\cdots i_m} b_{i_1\cdots i_m} \lambda(A_{i_1})\dots\lambda(A_{i_m}) \\
= \ & m! \int_0^T \cdots \int_0^T f(t_1,\dots,t_m) g(t_1,\dots,t_m) d\lambda(t_1)\dots d\lambda(t_m) \\
= \ & m! \langle f, g\rangle_{L^2([0,T]^m)}.
\end{aligned}$$

And because we assumed the functions f and g to be symmetric, we have

$$\mathbb{E}\left[I_m(f)I_m(g)\right] = \mathbb{E}\left[I_m(\tilde{f})I_m(\tilde{g})\right] = m!\langle \tilde{f}, \tilde{g}\rangle_{L^2([0,T]^m)}.$$

□

We also notice that if we have a symmetric function $f \in \mathcal{E}_m$, it holds for all permutations $\sigma \in S_m$

$$\int_{[0,T]^m} |f(t_1,\dots,t_m)|^2 d\lambda^m = \int_{[0,T]^m} |f(t_{\sigma(1)},\dots,t_{\sigma(m)})|^2 d\lambda^m,$$

and thanks to the triangle inequality, this gives us

$$\begin{aligned}
\|\tilde{f}\|_{L^2([0,T]^m)} &= \|\frac{1}{m!}\sum_{\sigma\in S_m} f\|_{L^2([0,T^m])} \\
&\le \frac{1}{m!}\sum_{\sigma\in S_m} \|f\|_{L^2([0,T]^m)} \\
&= \|f\|_{L^2([0,T]^m)}.
\end{aligned}$$

By that we conclude the inequality

$$\|\tilde{f}\|_{L^2([0,T]^m)} \le \|f\|_{L^2([0,T]^m)}. \tag{2.1}$$

Our next step is to extend the multiple stochastic integral to all functions in $L^2([0,T]^m)$.

Proposition 2.1.6. *The set $\mathcal{E}_m$ is dense in $L^2([0,T]^m, \mathfrak{B}([0,T]^m), \lambda^m)$.*

Proof. We do the proof in three steps.

Step 1: Let $\epsilon > 0$. First we show that we can approximate every $\mathbb{1}_A$, where $A = A_1 \times \cdots \times A_m$ and $A_i \in \mathfrak{B}([0,T])$ for all $i \in \{1, \ldots, m\}$, by functions from $\mathcal{E}_m$. Because the Lebesgue measure λ has no atoms, we can find for all $A \in \mathfrak{B}([0,T]^m)$ a measurable set $B \subset A$ such that

$$0 < \lambda(B) < \lambda(A).$$

Now let $\tilde{B} = \{B_1, \ldots, B_n\} \subset \mathfrak{B}([0,T])$, where $B_j \cap B_k = \emptyset$ if $j \neq k$ and $\lambda(B_j) < \epsilon$ for all $i = 1, \ldots, n$. We choose $\tilde{B}$ such that we can express every A_i as union of $B_j \in \tilde{B}$. Since $A \in [0,T]^m$, we let $\lambda_m(A) = \Pi_{i=1}^m \lambda(A_i) = \alpha$ and put $\epsilon_{i_1,\cdots,i_m}$ be 0 or 1. In this case we can write

$$\mathbb{1}_A = \sum_{i_1,\ldots,i_m=1}^{n} \epsilon_{i_1,\cdots,i_m} \mathbb{1}_{B_{i_1}\times\cdots\times B_{i_m}}.$$

If we define a set I which includes all $(i_1, \ldots, i_m)$ where $i_1, \ldots, i_m$ are all different and put $I^c = J$, we get

$$\mathbb{1}_B = \sum_{(i_1,\ldots,i_m)\in I} \epsilon_{i_1,\cdots,i_m} \mathbb{1}_{B_{i_1}\times\cdots\times B_{i_m}}$$

and also $\mathbb{1}_B \in \mathcal{E}_m$. Because in the set J are at least two of the $i_1, \ldots, i_m$ equal, we get

$$\begin{aligned}
\|\mathbb{1}_A - \mathbb{1}_B\|^2_{L^2([0,T]^m)} &= \| \sum_{(i_1,\ldots,i_m)\in J} \epsilon_{i_1,\cdots,i_m} \mathbb{1}_{B_{i_1}\times\cdots\times B_{i_m}} \|^2_{L^2([0,T]^m)} \\
&= \sum_{(i_1,\ldots,i_m)\in J} \epsilon_{i_1,\cdots,i_m} \lambda(B_{i_1})...\lambda(B_{i_m}) \\
&\leq \sum_{(i_1,\ldots,i_m)\in J} \lambda(B_{i_1})...\lambda(B_{i_m}) \\
&= \binom{m}{2} \sum_{j=1}^{n} (\lambda(B_j))^2 \big(\sum_{i=1}^{n} \lambda(B_i)\big)^{m-2} \\
&\leq \binom{m}{2} \epsilon \big(\sum_{i=1}^{n} \lambda(B_i)\big)^{m-1} \\
&\leq \binom{m}{2} \epsilon \alpha^{m-1} \\
&\to 0, \quad \text{if} \quad \epsilon \to 0.
\end{aligned}$$

Step 2: Next we show that every bounded function $f \in L^2([0,T]^m)$ can be approximated by simple functions which are defined by using the sets of the form $A = A_1 \times \cdots \times A_m$. We use the Monotone class theorem for functions (Proposition 1.0.3). First let $\mathcal{H}$ be the set which includes all bounded and measurable functions f such that $\|f - f_n\|_{L^2([0,T]^m)} \to 0$, when $n \to \infty$ and

$$f_n = \sum_{k=1}^{N_n} a_k^n \mathbb{1}_{A_{k_1}^n \times \cdots \times A_{k_m}^n}. \tag{2.2}$$

We define a set $\mathcal{A} = \{A_1 \times \cdots \times A_m, A_i \in \mathfrak{B}([0,T])$ for all $i \in \{1, \dots, m\}\}$. Now we check the properties of the Monotone class theorem for functions.

(i) It is clear that $\mathbb{1}_A \in \mathcal{H}$ for all $A \in \mathcal{A}$ by taking $f_n = \mathbb{1}_A$.
(ii) If we take $f, g \in \mathcal{H}$ and $a, b \in \mathbb{R}$, we find for f and g simple functions f_n and g_n such that $f_n \to f$ and $g_n \to g$ in $L^2([0,T]^m)$. Then $af_n + bg_n$ is a simple function, and $af_n + bg_n \to af + bg$ in $L^2([0,T]^m)$. By this we have $af + bg \in \mathcal{H}$.
(iii) Let $(g_n)_{n=1}^\infty \subseteq \mathcal{H}$ such that $0 \le g_n \uparrow h$, where h is bounded. Because $g_n \in \mathcal{H}$ for all $n \in \mathbb{N}$, there exists a sequence $(f_k^n)_{k=1}^\infty$ of simple functions like (2.2) such that $\|g_n - f_k^n\|_{L^2([0,T]^m)} \le \frac{\epsilon}{2^n}$ for some $k(n) \in \mathbb{N}$. Now we get

$$\begin{aligned} \|h - f_{k(n)}^n\|_{L^2([0,T]^m)} &= \|h - g_n + g_n - f_{k(n)}^n\|_{L^2([0,T]^m)} \\ &\le \|h - g_n\|_{L^2([0,T]^m)} + \|g_n - f_{k(n)}^n\|_{L^2([0,T]^m)} \\ &\le \|h - g_n\|_{L^2([0,T]^m)} + \frac{\epsilon}{2^n} \\ &\to 0, \text{ when } n \to \infty. \end{aligned}$$

This means that $h \in \mathcal{H}$.

With all this we conclude by the Monotone class theorem for functions (Proposition 1.0.3) that $\mathcal{H}$ contains all $\sigma(\mathcal{A}) = \mathfrak{B}([0,T]^m)$-measurable and bounded functions.

Step 3: We show that every function $f \in L^2([0,T]^m)$ can be approximated by bounded functions from $L^2([0,T]^m)$. Now let $N \in \mathbb{N}$. We take a function $f^N = (-N) \vee f \wedge N$ which is bounded for any square integrable f and $\lim_{N\to\infty} f^N = f$. By Dominated convergence (Proposition 1.2.1) we get

$$\begin{aligned} \lim_{N\to\infty} \|f^N - f\|_{L^2([0,T]^m)}^2 &= \lim_{N\to\infty} \int_{[0,T]^m} |f^N - f|^2 d\lambda^m \\ &\le \lim_{N\to\infty} \int_{\{[0,T]^m : |f(t)| > N\}} |f|^2 d\lambda^m \\ &= 0. \end{aligned}$$

□

This above proposition means, if we take $f \in L^2([0,T]^m)$ there exist $f_n \in \mathcal{E}_m$ for all $n = 1, 2, \dots$ such that $f_n \to f$ as $n \to \infty$ in $L^2([0,T]^m)$. By that, Proposition 2.1.5 (iii) and inequality (2.1) we get

$$\begin{aligned} \mathbb{E}\left[I_m(f_n) - I_m(f_k)\right]^2 &= \mathbb{E}\left[I_m(f_n - f_k)\right]^2 \\ &= m!\|\tilde{f}_n - \tilde{f}_k\|_{L^2([0,T]^m)}^2 \\ &\le m!\|f_n - f_k\|_{L^2([0,T]^m)}^2 \\ &\to m!\|f - f\|_{L^2([0,T]^m))}^2 = 0, \end{aligned}$$

when $n, k \to \infty$. This means that the sequence $\big(I_m(f_n)\big)_{n\geq 0}$ is Cauchy sequence in the space $L^2(\Omega, \mathcal{F}, \mathbb{P})$. Because I_m is linear and continuous operator, by Proposition 1.0.1 we can extend the operator I_m uniquely to the space $L^2([0,T]^m)$ with the same properties as given in Proposition 2.1.5. Moreover, if $f \in L^2([0,T]^m, \mathfrak{B}([0,T]^m), \lambda^m)$, then we denote the multiple stochastic integral by

$$\int_0^T \cdots \int_0^T f(t_1, \ldots, t_m) dW_{t_1} \ldots dW_{t_m} = I_m(f).$$

PROPOSITION 2.1.7. *Let $f \in L^2([0,T]^m)$ and $g \in L^2([0,T]^n)$, where $n, m \in \mathbb{N}$. Then the following holds almost surely*

(i) Let $m = n$ and $\alpha, \beta \in \mathbb{R}$. Then

$$I_m(\alpha f + \beta g) = \alpha I_m(f) + \beta I_m(g).$$

(ii) If $\tilde{f}$ is the symmetrization of f, then

$$I_m(\tilde{f}) = I_m(f).$$

(iii) For product of multiple stochastic integrals it holds

$$\mathbb{E}\left[I_m(f) I_n(g)\right] = \begin{cases} 0 & \text{if } m \neq n \\ m! \langle \tilde{f}, \tilde{g} \rangle_{L^2([0,T]^m)} & \text{if } m = n. \end{cases}$$

We can use instead of the multiple stochastic integral the iterated stochastic integral.

PROPOSITION 2.1.8. *For symmetric $f \in L^2([0,T]^m)$ it holds almost surely*

$$I_m(f) = m! \int_0^T \int_0^{t_m} \cdots \int_0^{t_3} \left(\int_0^{t_2} f(t_1, \ldots, t_m) dW_{t_1} \right) \ldots dW_{t_m}.$$

PROOF. Let S_m be the set of all permutations of $\{1, \ldots, m\}$. Since f is a symmetric function and by Proposition 2.1.6, we can first assume that the function f is zero on diagonals. In this situation we get

$$\begin{aligned}
& \int_0^T \cdots \int_0^T f(t_1, \ldots, t_m) dW_{t_1} \ldots dW_{t_m} \\
= & \int_0^T \cdots \int_0^T \sum_{\pi \in S_m} f(t_1, \ldots, t_m) \mathbb{1}_{\{t_{\pi(1)} \leq \cdots \leq t_{\pi(m)}\}} dW_{t_1} \ldots dW_{t_m} \\
= & \sum_{\pi \in S_m} \int_0^T \cdots \int_0^T f(t_1, \ldots, t_m) \mathbb{1}_{\{t_{\pi(1)} \leq \cdots \leq t_{\pi(m)}\}} dW_{t_{\pi(1)}} \ldots dW_{t_{\pi(m)}} \\
= & \sum_{\pi \in S_m} \int_0^T \cdots \int_0^{t_{\pi(3)}} \int_0^{t_{\pi(2)}} f(t_1, \ldots, t_m) dW_{t_{\pi(1)}} \ldots dW_{t_{\pi(m)}} \\
= & \sum_{\pi \in S_m} \int_0^T \cdots \int_0^{t_{\pi(3)}} \int_0^{t_{\pi(2)}} f(t_{\pi(1)}, \ldots, t_{\pi(m)}) dW_{t_{\pi(1)}} \ldots dW_{t_{\pi(m)}} \\
= & \; m! \int_0^T \cdots \int_0^{t_3} \int_0^{t_2} f(t_1, \ldots, t_m) dW_{t_1} \ldots dW_{t_m}.
\end{aligned}$$

Now we use Itô's isometry to extend the result to all symmetric $f \in L^2([0,T]^m)$.

□

We finish the multiple stochastic integral section with a theorem stating that every square integrable random variable which is $\mathcal{F}^W$-measurable can be expressed by a series of multiple stochastic integrals.

THEOREM 2.1.9 ([**18**] Theorem 1.1.2). *Let $F \in L^2(\Omega, \mathcal{F}^W, \mathbb{P})$, where the σ-algebra $\mathcal{F}^W$ is generated by the Brownian motion $W = (W_t)_{t\in[0,T]}$. Then every F can be uniquely expressed as*

$$F = \sum_{m=0}^{\infty} I_m(f_m),$$

where $f_m \in L^2([0,T]^m)$, and f_m is a symmetric function, for all $m \in \mathbb{N}$.

2.2. The Malliavin derivative

Multiple stochastic integrals can be used to define Malliavin derivatives. First we let

$$C_p^\infty(\mathbb{R}^n) = \left\{ \begin{array}{l} f : \mathbb{R}^n \to \mathbb{R}; \quad \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}} \text{ exists for all } |\alpha| = \alpha_1 + \cdots + \alpha_n, \\ \text{with } \alpha_i \in \mathbb{N}_0, f \text{ and } \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}} \text{ have polynomial growth} \end{array} \right\}.$$

We also need a set of smooth random variables, so we let

$$S = \left\{ \begin{array}{l} F = f(W(h_1), \ldots, W(h_n)); \quad f \in C_p^\infty(\mathbb{R}^n), \\ h_1, \ldots, h_n \in L^2([0,T]^m), n \geq 1 \end{array} \right\}.$$

DEFINITION 2.2.1. For $F \in S$ the Malliavin derivative is a stochastic process $(D_t F)_{t\in[0,T]}$ with

$$D_t F = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(W(h_1), \ldots, W(h_n)) h_i(t).$$

This Malliavin derivative is a linear operator from the space $S \subset L^2(\Omega, \mathcal{F}, \mathbb{P})$ to the space $L^2(\Omega \times [0,T], \mathcal{F} \otimes \mathfrak{B}([0,T]), \mathbb{P} \otimes \lambda)$. Next we introduce for all $p \in \mathbb{N}$ a norm

$$\|F\|_{1,p} = \left(\mathbb{E}(|F|^p) + \mathbb{E}(\|DF\|^p_{L^2([0,T])})\right)^{\frac{1}{p}}.$$

The next proposition states that the Malliavin derivative is a closable operator.

PROPOSITION 2.2.2 ([**19**] Proposition 1.2.1). *The Malliavin derivative $D : S \subset L^p(\Omega, \mathcal{F}, \mathbb{P}) \to L^p(\Omega \times [0,T], \mathcal{F} \otimes \mathfrak{B}([0,T]), \mathbb{P} \otimes \lambda)$ is a closable operator for all $p \geq 1$.*

Now we define $(\bar{D}, \mathbb{D}_{1,p})$ as an extension of (D, S) where

$$\mathbb{D}_{1,p} = \left\{ \begin{array}{l} x \in L^p(\Omega, \mathcal{F}, \mathbb{P}); \text{ there exists a Cauchy sequence} \\ (x_n)_{n=1}^\infty \subseteq S \text{ such that } x_n \to x \text{ in } \| \quad \|_{1,p} \end{array} \right\}$$

and $p \geq 1$. By defining $(\bar{D}, \mathbb{D}_{1,p})$ in this way we have a closed operator since if $(x_n)_{n=1}^\infty$ is a Cauchy sequence in $\| \quad \|_{1,p}$ and $x_n \to x$ in $\| \quad \|_{1,p}$, this means that there exists for a given $\epsilon > 0$ an $n(\epsilon) \in \mathbb{N}$ such that for all $m, n \geq n(\epsilon)$

$$\mathbb{E}|x_n - x_m|^p + \mathbb{E}\left(\int_0^T |D_t x_n - D_t x_m|^2 dt\right)^{\frac{p}{2}} \leq \epsilon.$$

Hence there exist $y \in L^p(\Omega \times [0,T], \mathcal{F} \otimes \mathfrak{B}([0,T]), \mathbb{P} \otimes \lambda)$ such that $Dx_n \to y$ in $L^p(\Omega \times [0,T], \mathcal{F} \otimes \mathfrak{B}([0,T]), \mathbb{P} \otimes \lambda)$. Then $x \in \mathbb{D}_{1,p}$ and $\bar{D}x = y$.

We use again D instead of $\bar{D}$ and get the Malliavin derivative $(D, \mathbb{D}_{1,p})$. Especially we will consider $\mathbb{D}_{1,2}$ i.e. the case $p = 2$, and $\mathbb{D}_{1,\infty} := \cap_{p \geq 1} \mathbb{D}_{1,p}$. Notice that we have

$$\mathbb{D}_{1,\infty} \subset \mathbb{D}_{1,2}.$$

We collect here for later use properties of the Malliavin derivative. The proofs can be found in Nualart [**18**] and [**19**].

Proposition 2.2.3 ([**18**] Proposition 1.2.1). *Let $F \in L^2(\Omega, \mathcal{F}, \mathbb{P})$ have the representation $F = \sum_{m=0}^\infty I_m(f_m)$, where $f_m \in L^2([0,T]^m)$, and f_m are symmetric for all $m \in \mathbb{N}$. Then $F \in \mathbb{D}_{1,2}$ if and only if*

$$\sum_{m=1}^{\infty} m m! \|f_m\|^2_{L^2([0,T]^m)} < \infty.$$

And for these F, we have $DF \in L^2(\Omega \times [0,T], \mathcal{F} \otimes \mathfrak{B}([0,T]), \mathbb{P} \otimes \lambda)$ with

$$D_t F = \sum_{m=1}^{\infty} m I_{m-1}(f_m(\cdot, t)).$$

Proposition 2.2.4 ([**18**]Proposition 1.2.2). *Let the function $f : \mathbb{R} \to \mathbb{R}$ be continuously differentiable such that $|f'| \leq M$ for some $M \in \mathbb{N}$ and assume that $F \in \mathbb{D}_{1,2}$. Then it holds*

(i) $f(F) \in \mathbb{D}_{1,2}$
(ii) $D_t(f(F)) = f'(F) D_t F$.

The above assertion extends to Lipschitz functions.

Proposition 2.2.5 ([**18**]Proposition 1.2.3). *Let $f : \mathbb{R} \to \mathbb{R}$ be a Lipschitz function and $F \in \mathbb{D}_{1,2}$. Then it holds*

(i) $f(F) \in \mathbb{D}_{1,2}$
(ii) There exists a random variable G bounded by the Lipschitz constant $M \in \mathbb{N}$ of f and

$$D_t(f(F)) = G D_t F.$$

We want to apply the Malliavin derivative also to solutions of SDEs. Therefore we assume a stochastic differential equation

$$\begin{cases} dX_t = g(t, X_t)dt + f(t, X_t)dW_t \\ X_0 = x_0 \in \mathbb{R} \end{cases} \tag{2.3}$$

such that for the coefficients the following two properties hold:
(a) $|f(t,x) - f(t,y)| + |g(t,x) - g(t,y)| \leq K|x - y|$ for all $x, y \in \mathbb{R}$ and $t \in [0,T]$
(b) $t \to f(t,0)$ and $t \to g(t,0)$ are bounded for $t \in [0,T]$.
Then the solution is Malliavin differentiable as it is stated in the following proposition.

Proposition 2.2.6 ([**19**]Theorem 2.2.1). *Let* (a) *and* (b) *hold. Then there exists a solution $(X_t)_{t \in [0,T]}$ for (2.3) and $X_t \in \mathbb{D}_{1,\infty}$ for all $t \in [0,T]$. Moreover,*

$$D_r X_t = f(r, X_t) + \int_r^t \bar{f}(s) D_r(X_s) dW_s + \int_r^t \bar{g}(s) D_r(X_s) ds \quad \textit{for } r \leq t \textit{ a.e.}$$

and

$$D_r(X_t) = 0 \qquad \textit{for } r > t \ \textit{a.e.},$$

where $\bar{f}(s)$ *and* $\bar{g}(s)$ *are uniformly bounded and adapted processes. If the coefficents in* (2.3) *are continuously differentiable, then*

$$\bar{f}(s) = \frac{\partial f}{\partial x}(s, X_s)$$

and

$$\bar{g}(s) = \frac{\partial g}{\partial x}(s, X_s).$$

CHAPTER 3

Mean-variance hedging with basis risk: Formulation of the problem

3.1. Continuous financial market

Before formulating our problem we recall some definitions for a continuous financial market on the time interval $[0,T]$, for some $T > 0$, by using [**4**]. Let $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\in[0,T]})$ be a stochastic basis which satisfies the usual conditions i.e. the probability space is complete and the filtration $(\mathcal{F}_t)_{t\in[0,T]}$ is right-continuous with the sets of probability zero included in $\mathcal{F}_0$. Assume that the financial market consists of one riskless asset $\{B(t), t \in [0,T]\}$ which has a constant interest rate $r > 0$ and m risky assets $\{S_i(t), t \in [0,T]\}$, $i = 1, 2, \ldots, m$ which are $\mathcal{F}_t$ -adapted stochastic processes with $S_i(0) = s_i > 0$ for all $i = 1, 2, \ldots, m$. Let $\theta(t) = (\theta_0(t), \theta_1(t), \ldots, \theta_m(t))^T$ be the $\mathcal{F}_t$ -adapted investment strategy at time $t \in [0,T]$. With this we denote the capital of investor which follows from strategy θ by

$$X^\theta(t) = \theta_0(t)B(t) + \theta_1(t)S_1(t) + \cdots + \theta_m(t)S_m(t)$$

at time $t \in [0,T]$. At time $t = 0$ the investor has some initial capital $X^\theta(0) = x_0 \geq 0$.

Definition 3.1.1. Let the capital have a representation

$$X^\theta(t) = X^\theta(0) + \int_0^t \theta_0(u)dB(u) + \int_0^t \theta_1(u)dS_1(u) + \cdots + \int_0^t \theta_m(u)dS_m(u)$$

with almost surely conditions

$$\int_0^t |\theta_0(u)|dB(u) < \infty$$

and

$$\int_0^t [\theta_i(u)S_i(u)]^2\, du < \infty \text{ for i=1,2,...,m.}$$

Then the strategy θ is called self-financing.

Definition 3.1.2. Let θ be a self-financing strategy. If there exists a constant $C > 0$ such that

$$\mathbb{P}\left(X^\theta(t) > -C, t \in [0,T]\right) = 1,$$

then the strategy θ is called admissible.

By using admissible strategies we can define an arbitrage-free market which is needed later.

Definition 3.1.3 ([**4**] Definition 1). Let θ be an admissible strategy such that

(1) $X^\theta(0) \leq 0$
(2) $X^\theta(T) \geq 0$

(3) $\mathbb{P}\left(X^\theta(T) > 0\right) > 0$,

then it is called an arbitrage. If there exists no arbitrage in the market, then the market is called arbitrage-free.

3.2. Formulation of the problem

We will follow [**24**] and start formulating the setting. We assume a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a time interval $[0,T]$ for some $T>0$. For $t \in [0,T]$, we assume $W = \left\{(W^0(t), ..., W^m(t))^T, t \in [0,T]\right\}$ to be an $(m+1)$-dimensional standard Brownian motion. We denote the transpose of a vector or a matrix M by M^T. Let $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ be the augmented filtration generated by the Brownian motion W and $(\mathcal{F}_t^0)_{t\in[0,T]}$ be the augmented filtration generated by W^0.

We assume also an arbitrage free financial market with a risk-free asset earning with a constant rate $r > 0$ and $m+1$ risky assets $\{S_i(t), t \in [0,T]\}$ $i = 1, 2, \dots, m$. We let $\{S_0(t), t \in [0,T]\}$ be the asset connected to the pay-off function with dynamics

$$\begin{cases} dS_0(t) = S_0(t)[\mu_0(t)dt + \sigma_{00}(t)dW^0(t)] \\ S_0(t) = s_0 > 0, \end{cases} \tag{3.1}$$

where $\mu_0(t)$ is the expected return rate and $\sigma_{00}(t)$ is the volatility.

We consider $G(S_0;T)$ be the pay-off function at maturity $T>0$ and assume the following:

Assumptions 3.2.1. $G(S_0;T)$

(1) $G(S_0;T) \in \mathbb{D}_{1,2}$.
(2) The asset S_0 can not be used to hedge the pay-off G.

Remark 3.2.2. As a consequence from the above we have to use the risky assets $\{S_i(t), t \in [0,T]\}$, $i = 1, 2, \dots, m$ and the risk-free asset for hedging.

The price processes of the risky assets we consider are given by the stochastic differential equations

$$\begin{cases} dS_i(t) = S_i(t)[\mu_i(t)dt + \sum_{j=0}^m \sigma_{ij}(t)dW^j(t)] \\ S_i(t) = s_i > 0, \end{cases} \tag{3.2}$$

for $i = 1, 2, \dots, m$. Here $\mu_i(t)$ is the expected return rate and $(\sigma_{i0}(t), \sigma_{i1}(t), \dots, \sigma_{im}(t))$ is the volatility vector of the i-th asset. We will also write

$$\sigma_i(t) = (\sigma_{i0}(t), \dots, \sigma_{im}(t))^T,$$

for all $i = 1, \dots, m$, and use matrix notation for

$$\sigma(t) = \begin{pmatrix} \sigma_1(t)^T \\ \sigma_2(t)^T \\ \vdots \\ \sigma_m(t)^T \end{pmatrix} = (\sigma_{ij}(t))_{m\times(m+1)} \in \mathbb{R}^{m\times(m+1)}.$$

The following is assumed throughout the thesis.

Assumptions 3.2.3. $\mu_i : [0,T] \to \mathbb{R}$ and $\sigma_{ij} : [0,T] \to \mathbb{R}$

(1) The functions μ_i and σ_{ij} are bounded and Borel-measurable for all $i, j = 0, 1, \dots, m$.

(2) There exists a constant ρ such that $\sigma(t)\sigma(t)^T \geq \rho I$ for all $t \in [0,T]$, where I denotes the $(m \times m)$ identity matrix.

The value of our hedging portfolio at time t is denoted by $X^\theta(t)$, where θ denotes the used hedging strategy. This hedging strategy is specified as a vector $\theta(t) = (\theta_1(t), \theta_2(t), \ldots, \theta_m(t))^T$ where every $\theta_i(t)$ is the amount which is invested in the risky asset $S_i(t)$. This holds for all $i = 1, \ldots, m$. The amount which is invested in the risk-free asset, we get by $X^\theta(t) - \sum_{i=1}^m \theta_i(t)$.

For the value process of the hedging portfolio, we have the following SDE

$$\begin{cases} dX^\theta(t) = [rX^\theta(t) + b(t)^T\theta(t)]dt + \theta(t)^T\sigma(t)dW(t) \\ X^\theta(0) = x_0 > 0 \end{cases} \tag{3.3}$$

where $b(t)$ denotes the expected excess return vector in form

$$b(t) = (\mu_1(t) - r, \mu_2(t) - r, \ldots, \mu_m(t) - r)^T,$$

and x_0 is the initial value. We will denote our hedging strategy shortly by $\{X^\theta(t), \theta(t)\}$ at time $t \in [0,T]$. We can measure the hedging error at terminal time T by $G(S_0;T) - X^\theta(T)$. Instead of this, we define

$$V^\theta(T) = X^\theta(T) - G(S_0;T), \tag{3.4}$$

which we call profit-and-loss random variable. This will measure the closeness of our hedging strategy to the pay-off, and negative values indicate a hedging error. We will use a mean-variance criterion whose pioneer was Markowitz [**17**]. By that we try to find an optimal hedging strategy which we denote by θ^*. We assume $\gamma > 0$ be the weight which the investor give for the variance. This strategy will be found by solving

$$\max_{\theta\in\Theta} \left\{ \mathbb{E}[V^\theta(T)] - \frac{\gamma}{2}\mathrm{Var}[V^\theta(T)] \right\}, \tag{3.5}$$

where Θ is the set of all admissible hedging strategies, meaning

$$\Theta = \left\{\theta : \theta \in \mathcal{L}_2\right\},$$

and $\mathcal{L}_2$ was defined in Definition 1.5.4. So we aim to maximize $\mathbb{E}[V^\theta(T)]$ and minimize $\mathrm{Var}[V^\theta(T)]$.

Remark 3.2.4. We assume that the set of admissible hedging strategies consists of progressively measurable and square integrable hedging strategies which differs from [**24**] where they assume square integrable and $\mathcal{F}_t$ -adapted strategies.

CHAPTER 4

The hedging strategy as stochastic linear-quadratic problem

For information about the general stochastic linear-quadratic theory, also called LQ theory, we recommend the reader to [**2**] and especially to Chapter 6 in [**25**]. Our aim is to find a hedging strategy θ^* such that it maximizes the difference, that is

$$\mathbb{E}\left[V^{\theta^*}(T)\right]-\frac{\gamma}{2}\mathrm{Var}\left[V^{\theta^*}(T)\right]=\sup_{\theta\in\Theta}\Big(\mathbb{E}\left[V^{\theta}(T)\right]-\frac{\gamma}{2}\mathrm{Var}\left[V^{\theta}(T)\right]\Big). \tag{4.1}$$

This question was treated in [**24**] where they refer to Section 3 in [**26**] for constructing a problem which can be solved by using the technique from Section 3 in [**16**].

Solving (4.1) is the same as if we try to find a hedging strategy θ^* which solves

$$\min_{\theta\in\Theta}\left\{J_1(\theta;\gamma)=\frac{\gamma}{2}\mathrm{Var}\left[V^{\theta}(T)\right]-\mathbb{E}\left[V^{\theta}(T)\right]\right\} \tag{4.2}$$

where $\gamma>0$ is the weight which the investor puts on the variance. By Section 3 in [**26**] this problem can be solved using an auxiliary problem which is defined as

$$\min_{\theta\in\Theta}\left\{J_2(\theta;\gamma,\lambda)=\mathbb{E}\left[\frac{\gamma}{2}V^{\theta}(T)^2-\lambda V^{\theta}(T)\right]\right\} \tag{4.3}$$

with the parameters $\gamma>0$ and $-\infty<\lambda<\infty$. The next lemma states that we can find any optimal solution for (4.2), if it exists, by solving (4.3).

Lemma 4.0.1 ([**24**] Lemma 1). *If $\theta^*(\cdot)$ is the solution to* (4.2)*, then it is the optimal control for the auxiliary problem* (4.3) *when λ is given by*

$$\lambda=1+\gamma\mathbb{E}[V^{\theta^*}].$$

Proof. Let $\theta^*(\cdot)$ be the solution for the problem (4.2). Let us assume that $\theta^*(\cdot)$ is not the optimal control for (4.3). This means that there exists a $\theta(\cdot)$ such that for the corresponding $V^{\theta}(\cdot)$ it holds

$$\begin{aligned}
& J_2(\theta;\gamma,\lambda)-J_2(\theta^*;\gamma,\lambda)\\
= {}& \frac{\gamma}{2}\mathbb{E}\left[V^{\theta}(T)^2\right]-\lambda\mathbb{E}\left[V^{\theta}(T)\right]-\Big(\frac{\gamma}{2}\mathbb{E}\left[V^{\theta^*}(T)^2\right]-\lambda\mathbb{E}\left[V^{\theta^*}(T)\right]\Big)\\
= {}& \frac{\gamma}{2}\left(\mathbb{E}\left[V^{\theta}(T)^2\right]-\mathbb{E}\left[V^{\theta^*}(T)^2\right]\right)-\lambda\left(\mathbb{E}\left[V^{\theta}(T)\right]-\mathbb{E}\left[V^{\theta^*}(T)\right]\right)\\
< {}& 0.
\end{aligned} \tag{4.4}$$

Now we define the function

$$f(x,y)=\frac{\gamma}{2}x-\frac{\gamma}{2}y^2-y \tag{4.5}$$

for $x,y\in\mathbb{R}$. It is easy to check that this function is concave. If we plug $\mathbb{E}\left[V^{\theta}(T)^2\right]$ and $\mathbb{E}\left[V^{\theta}(T)\right]$ into the function f, we get

$$f\left(\mathbb{E}\left[V^{\theta}(T)^2\right],\mathbb{E}\left[V^{\theta}(T)\right]\right) = \frac{\gamma}{2}\mathbb{E}\left[V^{\theta}(T)^2\right]-\frac{\gamma}{2}\left(\mathbb{E}\left[V^{\theta}(T)\right]\right)^2-\mathbb{E}\left[V^{\theta}(T)\right]$$

$$= \frac{\gamma}{2}\mathrm{Var}\left(V^\theta(T)\right) - \mathbb{E}\left[V^\theta(T)\right],$$

which is the same expression as in (4.2). Since $\gamma > 0$ and the only second partial derivative of (4.5) which differs from 0 is $\frac{\partial^2 f(x,y)}{\partial y \partial y} = -\gamma$ we get by Taylor approximation ([**5**] Theorem 2.8.3) that

$$\begin{aligned} f(x,y) &= f(x_0,y_0) + \frac{\gamma}{2}(x-x_0) - (1+\gamma y_0)(y-y_0) - \int_0^1 (1-t)\gamma(y-y_0)^2 dt \\ &\leq f(x_0,y_0) + \frac{\gamma}{2}(x-x_0) - (1+\gamma y_0)(y-y_0). \end{aligned}$$

By that and using (4.4) we have

$$\begin{aligned} & f\left(\mathbb{E}\left[V^\theta(T)^2\right], \mathbb{E}\left[V^\theta(T)\right]\right) \\ \leq\ & f\left(\mathbb{E}\left[V^{\theta^*}(T)^2\right], \mathbb{E}\left[V^{\theta^*}(T)\right]\right) + \frac{\gamma}{2}\left(\mathbb{E}\left[V^\theta(T)^2\right] - \mathbb{E}\left[V^{\theta^*}(T)^2\right]\right) \\ & - \left(1+\gamma\mathbb{E}\left[V^{\theta^*}(T)\right]\right)\left(\mathbb{E}\left[V^\theta(T)\right] - \mathbb{E}\left[V^{\theta^*}(T)\right]\right) \\ <\ & f\left(\mathbb{E}\left[V^{\theta^*}(T)^2\right], \mathbb{E}\left[V^{\theta^*}(T)\right]\right). \end{aligned}$$

This is a contradiction, since $\theta^*(\cdot)$ is the optimal control to (4.2).

□

Now, if we do some manipulations for the expression in (4.3)

$$\begin{aligned} \mathbb{E}\left[\frac{\gamma}{2}V^\theta(T)^2 - \lambda V^\theta(T)\right] &= \mathbb{E}\left[\frac{\gamma}{2}\left(V^\theta(T)^2 - \frac{2\lambda}{\gamma}V^\theta(T)\right)\right] \\ &= \mathbb{E}\left[\frac{\gamma}{2}\left(\left(V^\theta(T) - \frac{\lambda}{\gamma}\right)^2 - \frac{\lambda^2}{\gamma^2}\right)\right] \\ &= \mathbb{E}\left[\frac{\gamma}{2}\left(V^\theta(T) - \frac{\lambda}{\gamma}\right)^2 - \frac{\lambda^2}{2\gamma}\right] \\ &= \mathbb{E}\left[\frac{\gamma}{2}\left(V^\theta(T) - \frac{\lambda}{\gamma}\right)^2\right] - \frac{\lambda^2}{2\gamma} \end{aligned}$$

we can re-formulate our auxiliary problem (4.3) as

$$\min_{\theta\in\Theta}\left\{\mathbb{E}\left[\frac{\gamma}{2}V^\theta(T)^2 - \lambda V^\theta(T)\right]\right\} = \min_{\theta\in\Theta}\left\{\mathbb{E}\left[\frac{\gamma}{2}(V^\theta(T) - \frac{\lambda}{\gamma})^2\right]\right\} - \frac{\lambda^2}{2\gamma}.$$

Using the notation from (3.4) we can write

(4.6)
$$\min_{\theta\in\Theta}\left\{\mathbb{E}\left[\frac{\gamma}{2}\left(V^\theta(T) - \frac{\lambda}{\gamma}\right)^2\right]\right\} - \frac{\lambda^2}{2\gamma} = \min_{\theta\in\Theta}\left\{\mathbb{E}\left[\frac{\gamma}{2}\left(X^\theta(T) - G(S_0;T) - \frac{\lambda}{\gamma}\right)^2\right]\right\} - \frac{\lambda^2}{2\gamma}.$$

By putting

$$\xi = G(S_0;T) + \frac{\lambda}{\gamma}, \tag{4.7}$$

we get a new stochastic LQ problem

$$\min_{\theta\in\Theta}\left\{J_3(\theta;\lambda,\gamma)=\mathbb{E}\left[\left(X^\theta(T)-\xi\right)^2\right]\right\}. \tag{4.8}$$

So after all manipulations and because of Lemma 4.0.1 this means that we can solve the original problem (4.2) by solving (4.8).

We will follow [**24**] to solve (4.8). The method used there originates from Section 3 in [**16**]. We remark here that the variable ξ in problem (4.8) depends from $G(S_0;T)$ and especially on S_0. In Section 3 from [**16**] the problem uses instead of ξ a constant $d\in\mathbb{R}$. This makes the difference between [**16**] and [**24**].

First we introduce some notation. Recall that σ and b were used for the value process of the hedging portfolio (3.3).

Let

$$\Sigma(t)=\sigma(t)\sigma(t)^T \tag{4.9}$$

and

$$\begin{cases}\kappa(t)=b(t)^T\Sigma(t)^{-1}b(t)\\ \zeta(t)=b(t)^T\Sigma(t)^{-1}\sigma(t)e_1,\end{cases} \tag{4.10}$$

where $e_1=(1,0,\dots,0)^T$.

Solving problem (4.8) is connected to the following two equations: the backward differential equation

$$\begin{cases}dP(t)=[\kappa(t)-2r]P(t)dt,\\ P(T)=1\end{cases} \tag{4.11}$$

and the backward stochastic differential equation

$$\begin{cases}d\varphi(t)=\{[\kappa(t)-r]\varphi(t)+\zeta(t)\psi_0(t)\}dt+\psi_0(t)dW^0(t),\\ \varphi(T)=-\xi.\end{cases} \tag{4.12}$$

The next proposition shows that (4.11) and (4.12) have unique solutions in the space where we have an augmented filtration generated by the Brownian motion.

Proposition 4.0.2 ([**10**] Theorem 5.3.3). *Let $(\Omega,\mathcal{F}^W,\mathbb{P},(\mathcal{F}^W_t)_{t\in[0,T]})$ be a filtered probability space with Brownian motion and let f be a function such that $f:[0,T]\times\Omega\times\mathbb{R}\times\mathbb{R}\to\mathbb{R}$. Assume that*

(1) *$\xi\in L^2(\Omega,\mathcal{F},\mathbb{P})$.*
(2) *$f(\cdot,\cdot,y,z)$ is progressively measurable for all $y,z\in\mathbb{R}$*
(3) *There exists an $L_f>0$ such that .*

$$|f(t,\omega,y,z)-f(t,\omega,\hat y,\hat z)|\le L_f\left(|y-\hat y|+|z-\hat z|\right)$$

for all $(t,\omega)\in[0,T]\times\Omega$ and $y,\hat y,z,\hat z\in\mathbb{R}$.

(4) *$\mathbb{E}\int_0^T f(t,0,0)^2dt<\infty$.*

Then

$$Y(t)=\xi+\int_t^T f(s,Y(s),Z(s))ds-\int_t^T Z(s)dW(s)$$

has a unique solution $(Y,Z)\in\mathcal{S}_2\times\mathcal{L}_2$, where $\mathcal{S}_2$ is the set of all adapted and continuous processes $(Y_t)_{t\in[0,T]}$ with $\mathbb{E}\sup_{0\le t\le T}X_t^2<\infty$.

Since (4.11) and (4.12) are linear equations, they can be solved explicitly.

PROPOSITION 4.0.3. *Equations* (4.11) *and* (4.12) *have solutions*

$$P(t) = e^{\int_t^T [2r-\kappa(s)]ds} \tag{4.13}$$

and

$$\varphi(t) = e^{-\int_t^T \left(\kappa(s)-r+\frac{1}{2}[\zeta(s)]^2\right)ds}\mathbb{E}\left[-\xi e^{-\int_t^T \zeta(s)dW^0(s)}\big|\mathcal{F}_t^0\right], \tag{4.14}$$

respectively.

PROOF. Let $t \in [0,T]$. We first solve (4.11). Interpreting (4.11) as a backward stochastic differential equation we may apply Proposition 4.0.2 and conclude that it has a unique solution. Hence it suffices to check that $P(t)$ given in (4.13) solves (4.11):

$$\begin{aligned}
\frac{dP(t)}{dt} &= \frac{de^{\int_t^T [2r-\kappa(s)]ds}}{dt} \\
&= \frac{de^{\int_0^T [\kappa(s)-2r]ds}e^{-\int_0^t [2r-\kappa(s)]ds}}{dt} \\
&= e^{\int_0^T [2r-\kappa(s)]ds}e^{-\int_0^t [2r-\kappa(s)]ds}\left(\kappa(t)-2r\right) \\
&= e^{\int_t^T [2r-\kappa(s)]ds}\left(\kappa(t)-2r\right),
\end{aligned}$$

with

$$P(T) = e^{\int_T^T 2r-\kappa(s)ds} = 1.$$

So $P(t) = e^{\int_t^T 2r-\kappa(s)ds}$ is the unique solution for (4.11).

Now we consider (4.12) with solution (φ, ψ_0). We use solving method from [**6**]. First let

$$\begin{cases} \Psi(t) = -\Psi(t)[\kappa(t)-r]dt - \Psi(t)\zeta(t)dW^0(t) \\ \Psi(0) = 1. \end{cases}$$

This is a linear stochastic differential equation which has a unique solution (see for example [**13**] or [**14**]) given by

$$\Psi(t) = e^{\int_0^t -(\kappa(s)-r)-\frac{1}{2}\zeta(s)^2ds-\int_0^t \zeta(s)dW^0(s)}.$$

By (4.12) we have

$$\varphi(T) = \varphi(t) - \int_t^T [\kappa(s)-r]\varphi(s)+\zeta(s)\psi_0(s)ds - \int_t^T \psi_0(s)dW^0(s),$$

and we can use Itô's formula

$$\begin{aligned}
\Psi(T)\varphi(T) &= \Psi(t)\varphi(t) + \int_t^T \Psi(s)d\varphi(s) + \int_t^T \varphi(s)d\Psi(s) + \int_t^T 1d\left\langle \Psi, \varphi\right\rangle \\
&= \Psi(t)\varphi(t) - \int_t^T \Psi(s)\left(\left[\kappa(s)-r\right]\varphi(s)+\zeta(s)\psi_0(s)\right)ds \\
&\quad - \int_t^T \Psi(s)\psi_0(s)dW^0(s) + \int_t^T \varphi(s)\Psi(s)[\kappa(s)-r]ds \\
&\quad + \int_t^T \varphi(s)\Psi(s)\zeta(s)dW^0(s) - \int_t^T \psi_0(s)\Psi(s)\zeta(s)ds
\end{aligned}$$

$$= \quad \Psi(t)\varphi(t) + \int_t^T \varphi(s)\Psi(s)\zeta(s) - \Psi(s)\psi_0(s)dW^0(s).$$

It can be shown similarly as in the proof of Theorem 4.16, which comes later, by using Lemma 1.4.8, the Burkholder-Davis-Gundy inequality (Proposition 1.6.5), Remark 1.6.4 and Hölder's inequality (Proposition 1.6.6) that the stochastic integral is a martingale. We get by taking the conditional expectation with respect to $\mathcal{F}_t^0$

$$\Psi(t)\varphi(t) = \mathbb{E}\left[\Psi(T)\varphi(T)|\mathcal{F}_t^0\right].$$

Dividing this equation by $\Psi(t)$ gives us

$$\varphi(t) = e^{-\int_t^T \kappa(s)-r+\frac{1}{2}\zeta(s)^2 ds}\mathbb{E}\left[-\xi e^{-\int_t^T \zeta(s)dW^0(s)}|\mathcal{F}_t^0\right].$$

□

If we assume that G is a European option we can use the notation $\xi = G(S_0(T)) + \frac{\lambda}{\gamma}$, and φ can be interpreted as the price. Let

$$d\tilde{\mathbb{P}} = e^{-\int_0^T \zeta(s)dW^0(s)-\frac{1}{2}\int_0^T \zeta(s)^2 ds}d\mathbb{P},$$

where $e^{-\int_0^T \zeta(s)dW^0(s)-\frac{1}{2}\int_0^T \zeta(s)^2 ds}$ is a martingale by Novikov's condition (Proposition 1.5.17) since ζ is bounded. By Girsanov's Theorem [[**10**] Proposition 4.4.6] we have that the process

$$\tilde{W} = \left\{\tilde{W}(t), t \in [0,T]\right\},$$

where $\tilde{W}(t) = W^0(t) + \int_0^t \zeta(s)ds$ for $t \in [0,T]$ is a Brownian motion under the measure $\tilde{\mathbb{P}}$. Using this, we can rewrite the SDE for S_0 as follows

$$\begin{aligned} dS_0(t) &= S_0(t)\left[\mu_0(t)dt + \sigma_{00}(t)dW^0(t)\right] \\ &= S_0(t)\left[\mu_0(t)dt - \sigma_{00}\zeta(t)dt + \sigma_{00}(t)d\tilde{W}(t)\right] \\ &= S_0(t)\left[\tilde{\mu}_0(t)dt + \sigma_{00}(t)d\tilde{W}(t)\right], \end{aligned}$$

where $\tilde{\mu}_0(t) = \mu_0(t) - \sigma_{00}(t)\zeta(t)$ and $S_0(0) = s_0$. For this we have a solution

$$S_0(t) = s_0 e^{\int_0^t \tilde{\mu}_0(s)-\frac{1}{2}\sigma_{00}(s)^2 ds + \int_0^t \sigma_{00}(s)d\tilde{W}(s)}.$$

By the above we get for φ

$$\begin{aligned} \varphi(t) &= -e^{-\int_t^T \kappa(s)-r+\frac{1}{2}\zeta(s)^2 ds}\mathbb{E}\left[\xi e^{-\int_t^T \zeta(s)dW^0(s)}|\mathcal{F}_t^0\right] \\ &= -e^{-\int_t^T \kappa(s)-r ds}\mathbb{E}_{\tilde{\mathbb{P}}}\left[\xi|\mathcal{F}_t^0\right] \\ &= -e^{-\int_t^T \kappa(s)-r ds}\left(\mathbb{E}_{\tilde{\mathbb{P}}}\left[G(S_0(T))|\mathcal{F}_t^0\right] + \frac{\lambda}{\gamma}\right) \\ &= -e^{-\int_t^T \kappa(s)-r ds}\left(\mathbb{E}_{\tilde{\mathbb{P}}}\left[G\left(s_0 e^{\int_0^T \tilde{\mu}_0(s)-\frac{1}{2}\sigma_{00}(s)^2 ds + \int_0^T \sigma_{00}(s)d\tilde{W}(s)}\right)|\mathcal{F}_t^0\right] + \frac{\lambda}{\gamma}\right) \\ &= -e^{-\int_t^T \kappa(s)-r ds}\left(\mathbb{E}_{\tilde{\mathbb{P}}}\left[G\left(S_0(t) e^{\int_t^T \tilde{\mu}_0(s)-\frac{1}{2}\sigma_{00}(s)^2 ds + \int_t^T \sigma_{00}(s)d\tilde{W}(s)}\right)|\mathcal{F}_t^0\right] + \frac{\lambda}{\gamma}\right). \end{aligned}$$

Because σ_{00} and $\tilde{\mu}_0$ are deterministic we have that

$$\int_t^T \tilde{\mu}_0(s) - \frac{1}{2}\sigma_{00}(s)^2 ds + \int_t^T \sigma_{00}(s) d\tilde{W}(s)$$

follows normal distribution with mean $\int_t^T \tilde{\mu}_0(s) - \frac{1}{2}\sigma_{00}(s)^2 ds$ and variance $\int_t^T \sigma_{00}(s)^2 ds$ (see [**10**] Example 3.1.13). Also $S_0(t)$ is $\mathcal{F}_t^0$-measurable and $e^{\int_t^T \tilde{\mu}_0(s) - \frac{1}{2}\sigma_{00}(s)^2 ds + \int_t^T \sigma_{00}(s) d\tilde{W}(s)}$ is independent from $\mathcal{F}_t^0$. This means that it is possible to use the Black-Scholes formula to calculate the price of an European option. For more information about the Black-Scholes formula we recommend for example Section 4 in [**11**].

For ease of the presentation, let

$$\Gamma(t) = b(t)\varphi(t) + \sigma(t)e_1\psi_0(t). \tag{4.15}$$

The following theorem shows the connection between problem (4.8) and equations (4.11) and (4.12).

THEOREM 4.0.4 ([**24**] Theorem 1). *Let $P \in C([0,T])$ and $(\varphi, \psi_0) \in \mathcal{S}_2 \times \mathcal{L}_2$ be the solutions of (4.11) and (4.12). Then the problem (4.8) is solvable and the unique feedback control is given by*

$$\theta_\lambda(t) = -\Sigma(t)^{-1}\left[b(t)X^{\theta_\lambda}(t) + \frac{1}{P(t)}\Gamma(t)\right]. \tag{4.16}$$

The associated optimal value is

$$J_3(\theta_\lambda;\lambda) = P(0)x_0^2 + 2\varphi(0)x_0 + 2\mathbb{E}[\xi^2] - \mathbb{E}\left[\int_0^T \frac{1}{P(t)}\Gamma(t)^T\Sigma(t)^{-1}\Gamma(t)dt\right]. \tag{4.17}$$

PROOF. Let $\theta \in \Theta$ and X^θ the solution to (3.3) corresponding to θ. We can write the function of problem (4.8) as

$$J_3(\theta;\lambda) = \mathbb{E}\left[\left(X^\theta(T) - \xi\right)^2\right] = \mathbb{E}\left[\left(X^\theta(T)\right)^2\right] - 2\mathbb{E}\left[\xi X^\theta(T)\right] + \mathbb{E}\left[\xi^2\right]. \tag{4.18}$$

Now we get by Itô's formula

$$\begin{aligned}(X^\theta(t))^2 &= x_0^2 + \int_0^t 2X^\theta(s)\theta(s)^T\sigma(s)dW(s)\\ &\quad + \int_0^t 2X^\theta(s)\left(rX^\theta(s) + b(s)^T\theta(s)\right) + \theta(s)^T\sigma(s)\sigma(s)^T\theta(s)ds.\end{aligned}$$

Recall that we used notations

$$\Sigma(t) = \sigma(t)\sigma(t)^T$$

and

$$\begin{cases}\kappa(t) = b(t)^T\Sigma(t)^{-1}b(t)\\ \zeta(t) = b(t)^T\Sigma(t)^{-1}\sigma(t)e_1.\end{cases}$$

By using (4.11) and Itô's formula again, we get

$$\begin{aligned}P(t)\left(X^\theta(t)\right)^2 &= P(0)x_0^2 + \int_0^t P(s)d\left(X^\theta(s)\right)^2 + \int_0^t \left(X^\theta(s)\right)^2 dP(s)\\ &= P(0)x_0^2 + \int_0^t P(s)2X^\theta(s)\theta(s)^T\sigma(s)dW(s)\end{aligned}$$

$$
\begin{aligned}
& + \int_0^t P(s)\left[2X^\theta(s)\left(rX^\theta(s)+b(s)^T\theta(s)\right)\right]+\theta(s)^T\Sigma(s)\theta(s)ds \\
& + \int_0^t \left(X^\theta(s)\right)^2(\kappa(s)-2r)\,P(s)ds \\
= \;& P(0)x_0^2+\int_0^t 2P(s)X^\theta(s)\theta(s)^T\sigma(s)dW(s) \qquad (4.19)\\
& + \int_0^t P(s)\left[2X^\theta(s)b(s)^T\theta(s)+\theta(s)^T\Sigma(s)\theta(s)+\left(X^\theta(s)\right)^2\kappa(s)P(s)\right]ds.
\end{aligned}
$$

We want to show that

$$
\left\{\int_0^t 2P(s)X^\theta(s)\theta(s)^T\sigma(s)dW(s); \quad t\in[0,T]\right\}
$$

is martingale. We use Lemma 1.4.8. Let $G=\sup_{0\le t\le T}|\int_0^t 2P(s)X^\theta(s)\theta(s)^T\sigma(s)dW(s)|$. Now it is clear that G is a bound for the martingale and by the Burkholder-Davis-Gundy inequality (Proposition 1.6.5) and Hölder inequality (Proposition 1.6.6) we get

$$
\begin{aligned}
\mathbb{E}|G| &\le \alpha\mathbb{E}\left[\left(\int_0^T (2P(s)X^\theta(s)\theta(s)^T\sigma(s))^2ds\right)^{\frac{1}{2}}\right]\\
&\le 2\alpha\mathbb{E}\left[\sup_{0\le t\le T}|P(t)X^\theta(t)\sigma(t)|\left(\int_0^T(\theta(s)^T)^2ds\right)^{\frac{1}{2}}\right]\\
&= 2\alpha\sup_{0\le t\le T}|P(t)\sigma(t)|\mathbb{E}\left[\sup_{0\le t\le T}|X^\theta(t)|\left(\int_0^T(\theta(s)^T)^2ds\right)^{\frac{1}{2}}\right]\\
&\le 2\alpha\sup_{0\le t\le T}|P(t)\sigma(t)|\left(\mathbb{E}\sup_{0\le t\le T}|X^\theta(t)|^2\right)^{\frac{1}{2}}\left(\mathbb{E}\int_0^T(\theta(s)^T)^2ds\right)^{\frac{1}{2}}\\
&< \infty,
\end{aligned}
$$

where the last inequality follows by Remark 1.6.4 and assumptions of P,σ and θ. By Lemma 1.4.8 the Itô-integral $\left\{\int_0^t 2P(s)X^\theta(s)\theta(s)^T\sigma(s)dW(s); t\in[0,T]\right\}$ is a martingale and

$$
\mathbb{E}\left[\int_0^t 2P(s)X^\theta(s)\theta(s)^T\sigma(s)dW(s)\right]=0.
$$

So by setting $t=T$ and taking expectation of (4.19), we get

$$
\begin{aligned}
\mathbb{E}\left[P(T)\left(X^\theta(T)\right)^2\right] &= \mathbb{E}\left[\left(X^\theta(T)\right)^2\right]\\
&= P(0)x_0^2+\mathbb{E}\left\{\int_0^T\left[P(t)\left(2X\theta(t)b(t)^T\theta(t)+\theta(t)^T\Sigma(t)\theta(t)\right)\right.\right. \qquad (4.20)\\
&\qquad \left.\left.+\left(X^\theta(t)\right)^2\kappa(t)P(t)\right]dt\right\}.
\end{aligned}
$$

Similarly, we use Itô's formula to $\varphi(t)X^\theta(t)$ and get

$$
\begin{aligned}
\varphi(t)X^\theta(t) &= \varphi(0)x_0 + \int_0^t X^\theta(s)d\varphi(s) + \int_0^t \varphi(s)dX^\theta(s) \\
&\quad + \int_0^t 1d\left\langle \int_0^\cdot \psi_0(z)dW^0(z), \int_0^\cdot \theta(z)^T\sigma(z)dW(z)\right\rangle_s .
\end{aligned} \tag{4.21}
$$

For the cross-variation we have

$$
\begin{aligned}
&\left\langle \int_0^\cdot \psi_0(z)dW^0(z), \int_0^\cdot \theta(z)^T\sigma(z)dW(z)\right\rangle_s \\
= &\left\langle \int_0^\cdot \psi_0(z)dW^0(z), \sum_{l=0}^m \sum_{k=1}^m \int_0^\cdot \theta_k(z)\sigma_{kl}(z)dW^l(z)\right\rangle_s \\
= &\sum_{l=0}^m \left\langle \int_0^\cdot \psi_0(z)dW^0(z), \sum_{k=1}^m \int_0^\cdot \theta_k(z)\sigma_{kl}(z)dW^l(z)\right\rangle_s .
\end{aligned}
$$

Because of independence of the Brownian motions $W^0, W^1, \dots, W^m$ we get

$$
\left\langle \int_0^\cdot \psi_0(z)dW^0(z), \sum_{k=1}^m \int_0^\cdot \theta_k(z)\sigma_{kl}(z)dW^l(z)\right\rangle_s = 0, \qquad \text{if } l \neq 0.
$$

By the above computations

$$
\begin{aligned}
&\left\langle \int_0^\cdot \psi_0(z)dW^0(z), \int_0^\cdot \theta(z)^T\sigma(z)dW(z)\right\rangle_s \\
= &\left\langle \int_0^\cdot \psi_0(z)dW^0(z), \sum_{k=1}^m \int_0^\cdot \theta_k(z)\sigma_{k0}(z)dW^0(z)\right\rangle_s \\
= &\int_0^s \psi_0(z)\theta(z)^T\sigma(z)e_1dz.
\end{aligned}
$$

Substituting this into (4.21) and using (4.12) for φ and (3.3) for X^θ yields

$$
\begin{aligned}
\varphi(t)X^\theta(t) &= \varphi(0)x_0 + \int_0^t X^\theta(s)d\varphi(s) + \int_0^t \varphi(s)dX^\theta(s) \\
&\quad + \int_0^t 1d\left\langle \int_0^\cdot \psi_0(z)dW^0(z), \int_0^\cdot \theta(z)^T\sigma(z)dW(z)\right\rangle_s \\
&= \varphi(0)x_0 + \int_0^t X^\theta(s)d\varphi(s) + \int_0^t \varphi(s)dX^\theta(s) \\
&\quad + \int_0^t \psi_0(s)\theta(s)^T\sigma(s)e_1ds \\
&= \varphi(0)x_0 + \int_0^t X^\theta(s)\left[(\kappa(s)-r)\varphi(s) + \zeta(s)\psi_0(s)\right]ds \\
&\quad + \int_0^t X^\theta(s)\psi_0(s)dW^0(s) + \int_0^t \varphi(s)\left[rX^\theta(s) + b(s)^T\theta(s)\right]ds
\end{aligned}
$$

$$
\begin{aligned}
& + \int_0^t \varphi(s)\theta(s)^T\sigma(s)dW(s) + \int_0^t \psi_0(s)\theta(s)^T\sigma(s)e_1 ds \\
= \; & \varphi(0)x_0 + \int_0^t X^\theta(s)\psi_0(s) + \varphi(s)\theta(s)^T\sigma(s)dW^0(s) \\
& + \int_0^t X^\theta(s)\kappa(s)\varphi(s) + X^\theta(s)\zeta(s)\psi_0(s) + \varphi(s)b(s)^T\theta(s) \\
& + \psi_0(s)\theta(s)^T\sigma(s)e_1 ds \\
= \; & \varphi(0)x_0 + \int_0^t X^\theta(s)\psi_0(s) + \varphi(s)\theta(s)^T\sigma(s)dW^0(s) \\
& + \int_0^t X^\theta(s)b(s)^T\Sigma(s)^{-1}b(s)\varphi(s) + X^\theta(s)b(s)^T\Sigma(s)^{-1}\sigma(s)e_1\psi_0(s) \\
& + \varphi(s)b(s)^T\theta(s) + \psi_0(s)\theta(s)^T\sigma(s)e_1 ds \\
= \; & \varphi(0)x_0 + \int_0^t X^\theta(s)\psi_0(s) + \varphi(s)\theta(s)^T\sigma(s)dW^0(s) \\
& + \int_0^t X^\theta(s)b(s)^T\Sigma(s)^{-1}\left[b(s)\varphi(s) + \sigma(s)e_1\psi_0(s)\right] \\
& + \theta(s)^T\left[b(s)\varphi(s) + \sigma(s)e_1\psi_0(s)\right] ds \\
= \; & \varphi(0)x_0 + \int_0^t X^\theta(s)\psi_0(s) + \varphi(s)\theta(s)^T\sigma(s)dW^0(s) \\
& + \int_0^t X^\theta(s)b(s)^T\Sigma(s)^{-1}\Gamma(s) + \theta(s)^T\Gamma(s)ds,
\end{aligned}
\tag{4.22}
$$

where (4.15) was used for the last equality. It can be shown similarly as before that

$$
\left\{\int_0^t \left[X^\theta(s)\psi_0(s) + \varphi(s)\theta(s)^T\sigma(s)\right] dW^0(s); t \in [0,T]\right\}
$$

is a martingale and hence

$$
\mathbb{E}\left[\int_0^t \left[X^\theta(s)\psi_0(s) + \varphi(s)\theta(s)^T\sigma(s)\right] dW^0(s)\right] = 0.
$$

By choosing $t = T$ and taking expectation, we get

$$
\begin{aligned}
\mathbb{E}\left[\varphi(T)X^\theta(T)\right] &= -\mathbb{E}\left[\xi X^\theta(T)\right] \\
&= \varphi(0)x_0 + \mathbb{E}\left\{\int_0^T \left[b(s)^T\Sigma(s)^{-1}\Gamma(s)X^\theta(s) + \theta(s)^T\Gamma(s)\right] ds\right\}.
\end{aligned}
\tag{4.23}
$$

We write (4.18) by using (4.20) and (4.23) for some $\theta \in \Theta$. We separate the terms which depend on θ and from those which do not. Then we notice that if we put (4.16) instead of θ we get the inequality

$$
\begin{aligned}
J_3(\theta;\lambda) &= \mathbb{E}\left[\left(X^\theta(T)\right)^2\right] - 2\mathbb{E}\left[\xi X^\theta(T)\right] + \mathbb{E}\left[\xi^2\right] \\
&= P(0)x_0^2 + 2\varphi(0)x_0 + \mathbb{E}[\xi^2] \\
&\quad + \mathbb{E}\left[\int_0^T P(t)\left[2X^\theta(t)b(t)\theta(t) + \theta(t)^T\Sigma(t)\theta(t) + \left(X^\theta(t)\right)^2\kappa(t)\right] dt\right]
\end{aligned}
$$

$$
\begin{aligned}
& +2\mathbb{E}\left[\int_0^T \left[b(t)^T\Sigma(t)^{-1}\Gamma(t)X^\theta(t) + \theta(t)^T\Gamma(t)\right] dt\right] \\
= \;& P(0)x_0^2 + 2\varphi(0)x_0 + \mathbb{E}\left[\xi^2\right] - \mathbb{E}\left[\int_0^T \frac{1}{P(t)}\Gamma(t)^T\Sigma(t)^{-1}\Gamma(t)dt\right] \\
& +\mathbb{E}\int_0^T \left(X^\theta(t)b(t) + \Sigma(t)\theta(t) + \frac{1}{P(t)}\Gamma(t)\right)^T \\
& \times\Sigma(t)^{-1}\left(X^\theta(t)b(t) + \Sigma(t)\theta(t) + \frac{1}{P(t)}\Gamma(t)\right) P(t)dt \\
\geq \;& P(0)x_0^2 + 2\varphi(0)x_0 + \mathbb{E}\left[\xi^2\right] - \mathbb{E}\left[\int_0^T \frac{1}{P(t)}\Gamma(t)^T\Sigma(t)^{-1}\Gamma(t)dt\right] \\
= \;& J_3(\theta_\lambda;\lambda),
\end{aligned}
$$

where one can see the third equality by multiplying the terms. This calculation shows that (4.16) is the optimal control for (4.8) and the associated optimal value is given by $J_3(\theta_\lambda;\lambda)$. Our inequality can be an equality only when $\theta = \theta_\lambda$ so (4.16) is also unique.

□

We have the explicit formula for φ but as we see above, the optimal θ_λ depends also on Γ and in that way on ψ_0. Assume a backward stochastic differential equation

$$
Y(t) = \xi + \int_t^T f(s, Y(s), Z(s))ds - \int_t^T Z(s)dW(s), \tag{4.24}
$$

such that it has a solution. The next proposition shows the connection between the Malliavin derivative and the solution (Y, Z).

Proposition 4.0.5 ([**6**] Proposition 5.3. and [**9**] Theorem 3.12). *Let the BSDE (4.24) have a solution (Y, Z). Assume the following*

(1) $\xi \in \mathbb{D}_{1,2}$, $\int_0^T \mathbb{E}[|D_u\xi|^2]du < \infty$, $\mathbb{E}[\xi^4] < \infty$
(2) $f : \Omega \times [0, T] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ *is continuously differentiable in* (y, z) *with uniformly bounded and continuous derivatives*
(3) for each $(y, z) \in \mathbb{R} \times \mathbb{R}$
  (a) $f(\cdot, y, z)$ *is progressively measurable*
  (b) $f(\cdot, y, z) \in \mathbb{D}_{1,2}$
  (c) $Df(\cdot, y, z) \in L^2([0, T])$
  (d) $\mathbb{E}\left[\int_0^T f(t, y, z)^2dt + \int_0^T\int_0^T D_uf(t, y, z)^2dudt\right] < \infty$
(4) $\mathbb{E}\int_0^T f(t, 0, 0)^4dt < \infty$
(5) $\int_0^T \mathbb{E}\left[(D_uf(t, Y(t), Z(t)))^2\right] du < \infty$
(6) for all $t \in [0, T]$ *and* $y, \hat{y}, z, \hat{z} \in \mathbb{R}$

$$
|D_uf(t, \omega, y, z) - D_uf(t, \omega, \hat{y}, \hat{z})| \leq K_u(t, \omega)\left(|y - \hat{y}| + |z - \hat{z}|\right),
$$

where $K_u : [0, T] \times \Omega \to [0, \infty)$ *is adapted process with* $\int_0^T \mathbb{E}[|K_u|^p]du < \infty$ *for some* $p > 0$.

Then $(Y, Z) \in \mathbb{D}_{1,2} \times \mathbb{D}_{1,2}$ *and*

(i) $f(t, Y(t), Z(t)) \in \mathbb{D}_{1,2}$
(ii) $D_u Y(t) = D_u Z(t) = 0$ *for* $0 \leq t < u \leq T$
(iii)

$$\begin{aligned} D_u Y(t) &= D_u \xi + \int_t^T \left[\frac{df}{dy}(s, Y(s), Z(s)) D_u Y(s) + \frac{df}{dz}(s, Y(s), Z(s)) D_u Z(s) \right. \\ &\quad \left. + D_u f(s, Y(s), Z(s)) \right] ds - \int_t^T D_u Z(s) dW(s), \end{aligned}$$

for $u \leq t \leq T$*. Moreover,* $Z(t) = D_t Y(t)$*, for* $0 \leq t \leq T$*.*

Remark 4.0.6. The assumption $\int_0^T \mathbb{E}[K_u^4] du < \infty$ from [**6**] Proposition 5.3 can be relaxed to $\int_0^T \mathbb{E}[|K_u|^p] du < \infty$ for some $p > 1$. This is shown in [**9**]. Also a proof of (i) one can find in [**9**].

In our case we get a simplified version as a corollary:

Corollary 4.0.7 ([**24**] Proposition 2.). *Let* (4.12) *have the solution* (φ, ψ_0) *and assume that* $\xi \in \mathbb{D}_{1,2}$ *with* $\mathbb{E}[\xi^4] < \infty$*. Then*
(i) $D_u \varphi(t) = D_u \psi_0(t) = 0$*, for* $0 \leq t < u \leq T$*,*
(ii) $D_u \varphi(t) = -D_u \xi - \int_t^T [\kappa(s) - r] D_u \varphi(s) + \zeta(s) D_u \psi_0(s) ds - \int_t^T D_u \psi_0(s) dW^0(s)$*,*
for $u \leq t \leq T$*. Moreover,* $\psi_0(t) = D_t \varphi(t)$*, for* $0 \leq t \leq T$*.*

By the above corollary, we are able to give an explicit formula for ψ_0. For later use, we give also explicit formulas for the expectations of φ and ψ_0.

Proposition 4.0.8 ([**24**] Proposition 3.). *Let* (φ, ψ_0) *be the solution to the BSDE* (4.12)*. Then*

$$\psi_0(t) = D_t \varphi(t) = e^{-\int_t^T \left(\kappa(s) - r + \frac{1}{2}[\zeta(s)]^2\right) ds} \mathbb{E}\left[-e^{-\int_t^T \zeta(s) dW^0(s)} D_t \xi | \mathcal{F}_t^0 \right], \tag{4.25}$$

$$\mathbb{E}[\varphi(t)] = e^{-\int_t^T \left(\kappa(s) - r + \frac{1}{2}[\zeta(s)]^2\right) ds} \mathbb{E}\left[-\xi e^{-\int_t^T \zeta(s) dW^0(s)} \right], \tag{4.26}$$

$$\mathbb{E}[\psi_0(t)] = \mathbb{E}[D_t \varphi(t)] = e^{-\int_t^T \left(\kappa(s) - r + \frac{1}{2}[\zeta(s)]^2\right) ds} \mathbb{E}\left[-e^{-\int_t^T \zeta(s) dW^0(s)} D_t \xi \right]. \tag{4.27}$$

Proof. To show (4.25) we put $Y(t) = D_u \varphi(t)$ and $Z(t) = D_u \psi_0(s)$ in

$$D_u \varphi(t) = -D_u \xi - \int_t^T [\kappa(t) - r] D_u \varphi(s) + \zeta(s) D_u \psi_0(s) ds - \int_t^T D_u \psi_0(s) dW^0(s).$$

Then we have

$$Y(t) = -D_u \xi - \int_t^T [\kappa(t) - r] Y(s) + \zeta(s) Z(s) ds - \int_t^T Z(s) dW^0(s) \tag{4.28}$$

which is a backward stochastic differential equation. Defining a linear stochastic differential equation

$$\begin{cases} d\Psi(t) = -\Psi(t)[\kappa(t) - r] dt - \Psi(t) dW^0(t) \\ \Psi(T) = 1, \end{cases}$$

we can use the same method as in the proof of Proposition 4.0.3 to get the solution.

To show (4.26) we take the expectation of the formula from Proposition 4.0.3 and use tower property for the conditional expectation. It can be done similarly also for (4.27).

□

We introduced the notation $\xi = G(S_0; T) + \frac{\lambda}{\gamma}$ in (4.7) to get formula (4.8). By the above proposition we see that ψ_0 depends on $D_t\xi$ so we have to find a formula also for that. By Assumptions 3.2.1 we have $G(S_0; T) \in \mathbb{D}_{1,2}$ and we can use Proposition 2.2.4 to get $D_t\xi = D_tG(S_0; T)$. We are able to find this Malliavin derivative explicitly for some common financial instruments.

Lemma 4.0.9. *Assume a forward contract with a payoff $G(S_0; T) = S_0(T)$, where S_0 has the dynamics from* (3.1)*. Then*

$$D_tG(S_0; T) = D_tS_0(T) = S_0(T)\sigma_{00}(t).$$

Proof. By Proposition 2.2.6 we have that

$$\begin{aligned} D_tS_0(T) &= S_0(t)\sigma_{00}(t) + \int_t^T D_t(S_0(s))\sigma_{00}(s)dW^0(s) + \int_t^T D_t(\sigma_{00}(s))S_0(s)dW^0(s) \\ &\quad + \int_t^T D_t(S_0(s))\mu_0(s) + D_t(\mu_0(s))S_0(s)ds \\ &= S_0(t)\sigma_{00}(t) + \int_t^T D_t(S_0(s))\sigma_{00}(s)dW^0(s) + \int_t^T D_t(S_0(s))\mu_0(s)ds. \end{aligned}$$

Since σ_{00} and μ_0 are deterministic, hence their Malliavin derivative is zero. By putting $Y(T) = D_tS_0(T)$ we get a linear stochastic differential equation

$$Y(T) = S_0(t)\sigma_{00}(t) + \int_t^T Y(s)\sigma_{00}(s)dW^0(s) + \int_t^T Y(s)\mu_0(s)ds, \tag{4.29}$$

which we are able to solve. The solving method can be found for example in [**13**]. We get as a solution

$$Y(T) = S_0(t)\sigma_{00}(t)e^{\int_t^T \sigma_{00}(s)dW^0(s) + \int_t^T \mu_0(s) - \frac{1}{2}\sigma_{00}^2(s)ds}. \tag{4.30}$$

Similarly we can solve (3.1) and this gives us

$$S_0(t) = s_0 e^{\int_0^t \sigma_{00}(s)dW^0(s) + \int_0^t \mu_0(s) - \frac{1}{2}\sigma_{00}^2(s)ds}. \tag{4.31}$$

By substituting (4.31) to (4.30) we get

$$\begin{aligned} Y(T) = D_tS_0(t) &= s_0 e^{\int_0^t \sigma_{00}(s)dW^0(s) + \int_0^t \mu_0(s) - \frac{1}{2}\sigma_{00}^2(s)ds} \\ &\quad \times \sigma_{00}(t) e^{\int_t^T \sigma_{00}(s)dW^0(s) + \int_t^T \mu_0(s) - \frac{1}{2}\sigma_{00}^2(s)ds} \\ &= \sigma_{00}(t) s_0 e^{\int_0^T \sigma_{00}(s)dW^0(s) + \int_0^T \mu_0(s) - \frac{1}{2}\sigma_{00}^2(s)ds} \\ &= \sigma_{00}(t)S_0(T). \end{aligned}$$

□

Lemma 4.0.10. *Assume European put and call options with payoffs $G(S_0; T) = (K - S_0(T))_+$ and $G(S_0; T) = (S_0(T) - K)_+$ respectively. Then*

(1) $D_t(K - S_0(T))_+ = -\mathbb{1}_{\{S_0(T) \le K\}}S_0(T)\sigma_{00}(t)$

(2) $D_t(S_0(T) - K)_+ = \mathbb{1}_{\{S_0(T) \ge K\}}S_0(T)\sigma_{00}(t)$.

PROOF. (1) Let $K \in [0,\infty)$ and $F : [0,\infty) \to \mathbb{R}$ with $F(x) = (K-x)_+$. To use Proposition 2.2.4, we need to have a continuously differentiable function. Let $n \in \mathbb{N}$ and

$$f_n(x) = \begin{cases} -1, \text{ if } x \leq K - \frac{1}{n} \\ \frac{x-(K-\frac{1}{n})}{\frac{2}{n}} - 1, \text{ if } K - \frac{1}{n} \leq x \leq K + \frac{1}{n} \\ 0, \text{ if } x \geq K + \frac{1}{n}. \end{cases} \tag{4.32}$$

With f_n we get a continuously differentiable and bounded function

$$F_n(x) = K + \int_0^x f_n(y)dy.$$

By Lemma 4.0.9 and Proposition 2.2.4

$$F_n(S_0(T)) \in \mathbb{D}_{1,2}$$

and

$$\begin{aligned} D_t F_n(S_0(T)) &= f_n(S_0(T)) D_t S_0(T) \\ &= f_n(S_0(T)) S_0(T) \sigma_{00}(t). \end{aligned}$$

Now we want to show that

$$F_n(S_0(T)) \to F(S_0(T) \text{ in } \mathbb{D}_{1,2}.$$

Let us recall that for all $G \in \mathbb{D}_{1,2}$ we have

$$\|G\|_{1,2} = \left(\mathbb{E}[G^2] + \mathbb{E}\|DG\|^2_{L_2([0,T])}\right)^{\frac{1}{2}} < \infty.$$

We will show the convergence in two parts:

(i) Since F and F_n are bounded, and for all $x \in \mathbb{R}$ it holds $F_n(x) \to F(x)$, when $n \to \infty$, we get by the Dominated convergence (Proposition 1.2.1)

$$\lim_{n\to\infty} \mathbb{E}|F_n(S_0(T)) - F(S_0(T))|^2 = 0.$$

(ii) Recall that D is a closed operator. Now

$$\begin{aligned} & \mathbb{E}\|D_{.}F_n(S_0(T)) + \mathbb{1}_{\{S_0(T)\leq K\}} S_0(T)\sigma_{00}\|^2_{L_2([0,T])} \\ = \ & \mathbb{E}\int_0^T |f_n(S_0(T))S_0(T)\sigma_{00}(t) + \mathbb{1}_{\{S_0(T)\leq K\}} S_0(T)\sigma_{00}(t)|^2 dt \\ = \ & \mathbb{E}\int_0^T |(f_n(S_0(T)) + \mathbb{1}_{\{S_0(T)\leq K\}}) S_0(T)\sigma_{00}(t)|^2 dt \\ = \ & \mathbb{E}\left[|f_n(S_0(T) + \mathbb{1}_{\{S_0(T)\leq K\}}|^2 S_0(T)^2\right] \int_0^T \sigma_{00}(t)^2 dt. \end{aligned}$$

We use (4.32) for f_n and estimate by using the inequality $(a+b)^2 \leq 2a^2 + 2b^2$ and $|f_n(x)| \leq 1$ for all $x \in \mathbb{R}$

$$\begin{aligned} & \mathbb{E}\left[|f_n(S_0(T)) + \mathbb{1}_{\{S_0(T)\leq K\}}|^2 S_0(T)^2\right] \\ = \ & \mathbb{E}\Bigg[\Bigg| -\mathbb{1}_{\{S_0(T)\leq K-\frac{1}{n}\}} + \left(\frac{S_0(T)-(K-\frac{1}{n})}{\frac{2}{n}} - 1\right)\mathbb{1}_{\{K-\frac{1}{n}\leq S_0(T)\leq K+\frac{1}{n}\}} + \mathbb{1}_{\{S_0(T)\leq K\}}\Bigg|^2 \end{aligned}$$

$$\times S_0(T)^2\Bigg]$$

$$\begin{aligned}
&= \mathbb{E}\Bigg[\Bigg|\Bigg(\frac{S_0(T)-(K-\frac{1}{n})}{\frac{2}{n}}-1\Bigg)\mathbb{1}_{\{K-\frac{1}{n}\le S_0(T)\le K+\frac{1}{n}\}}+\mathbb{1}_{\{K-\frac{1}{n}<S_0(T)\le K\}}\Bigg|^2 S_0(T)^2\Bigg]\\
&\le 2\mathbb{E}\left[\mathbb{1}_{\{K-\frac{1}{n}<S_0(T)\le K\}}S_0(T)^4\right]+2\mathbb{E}\left[\mathbb{1}_{\{K-\frac{1}{n}\le S_0(T)\le K+\frac{1}{n}\}}S_0(T)^4\right].
\end{aligned}$$

By Change of variable (Proposition 1.2.2) we have

$$\begin{aligned}
&2\mathbb{E}\left[\mathbb{1}_{\{K-\frac{1}{n}<S_0(T)\le K\}}S_0(T)^4\right]+2\mathbb{E}\left[\mathbb{1}_{\{K-\frac{1}{n}\}\le S_0(T)\le K+\frac{1}{n}}S_0(T)^4\right]\\
&= 2\int_{\{K-\frac{1}{n}<x\le K\}}x^4 d\mathbb{P}_{S_0(T)}(x)+2\int_{\{K-\frac{1}{n}\le x\le K+\frac{1}{n}\}}x^4 d\mathbb{P}_{S_0(T)}(x).
\end{aligned}$$

If we take intersections of sets we get

$$\cap_{n=1}^{\infty}\{K-\frac{1}{n}<x\le K\}=\emptyset$$

and

$$\cap_{n=1}^{\infty}\{K-\frac{1}{n}\le x\le K+\frac{1}{n}\}=\{K\}.$$

Now because S_0 is a geometric Brownian motion (exponential Brownian motion), S_0 has a continuous distribution which has all moments and

$$2\int_{\emptyset}x^4 d\mathbb{P}_{S_0(T)}(x)+2\int_{\{K\}}x^4 d\mathbb{P}_{S_0(T)}(x)=0.$$

By (i) and (ii) $F_n(S_0(T))\to F(S_0(T))$ in $\mathbb{D}_{1,2}$ and since the operator D is closed we conclude that

$$D_tF(S_0(T))=-\mathbb{1}_{\{S_0(T)\le K\}}S_0(T)\sigma_{00}(t).$$

(2) We get from the put-call parity by writing $(x-K)_+=(x-K)\mathbb{1}_{\{x\ge K\}}$ and $(K-x)_+=(K-x)\mathbb{1}_{\{x\le K\}}$

$$\begin{aligned}
(x-K)_+-(K-x)_+ &= (x-K)\mathbb{1}_{\{x\ge K\}}-(K-x)\mathbb{1}_{\{x\le K\}}\\
&= x\mathbb{1}_{\{x\ge K\}}-K\mathbb{1}_{\{x\ge K\}}-K\mathbb{1}_{\{x\le K\}}+x\mathbb{1}_{\{x\le K\}}\\
&= x-K.
\end{aligned}$$

By putting $x=S_0(T)$ we get

$$(S_0(T)-K)_+-(K-S_0(T))_+=S_0(T)-K$$

which can be written as

$$(S_0(T)-K)_+=S_0(T)-K+(K-S_0(T))_+.$$

Since $(S_0(T)-K)_+=(K-S_0(T))_+=0$, when $S_0(T)=K$, we get

$$\begin{aligned}
D_t(S_0(T)-K)_+ &= D_tS_0(T)-D_tK+D_t(K-S_0(T))_+\\
&= S_0(T)\sigma_{00}(t)-\mathbb{1}_{\{S_0(T)\le K\}}S_0(T)\sigma_{00}(t)\\
&= \mathbb{1}_{\{S_0(T)\ge K\}}S_0(T)\sigma_{00}(t).
\end{aligned}$$

□

LEMMA 4.0.11. *Assume an Asian option with payoff* $G(S_0;T) = \left(\frac{1}{T}\int_0^T S_0(r)dr - K\right)_+$. *Then*

$$D_t\left(\frac{1}{T}\int_0^T S_0(r)dr - K\right)_+ = \mathbb{1}_{\{\frac{1}{T}\int_0^T S_0(r)dr \geq K\}}\frac{\sigma_{00}(t)}{T}\int_t^T S_0(r)dr.$$

PROOF. This can be done similarly as the proof of Lemma 4.0.10. The only thing we need is the Malliavin derivative

$$D_t\left(\frac{1}{T}\int_0^T S_0(r)dr\right)$$

which we get by Proposition 2.2.6

$$D_t(\frac{1}{T}\int_0^T S_0(r)dr) = \frac{\sigma_{00}(t)}{T}\int_t^T S_0(r)dr.$$

$\square$

CHAPTER 5

The optimal hedging strategy

The aim of this chapter is to find the optimal hedging strategy which solves the mean-variance hedging problem. For this we need an explicit formula for λ.

PROPOSITION 5.0.1. *Assume the optimal hedging strategy θ_λ as given in Theorem 4.0.4. Then*

$$\begin{aligned}\lambda &= \frac{1}{N}\left\{1-\gamma\mathbb{E}\left[G(S_0;T)\right]+\gamma e^{\int_0^T r-\kappa(t)dt}x_0\right\}\\ &\quad +\frac{\gamma}{N}\int_0^T \zeta(t)\mathbb{E}[\psi_0(t)]e^{-r(T-t)}dt\\ &\quad +\frac{\gamma}{N}\int_0^T \kappa(t)\Sigma(t)^{-1}e^{-\int_t^T \kappa(s)+\frac{1}{2}\zeta(s)^2ds}\mathbb{E}\left[G(S_0;T)e^{-\int_t^T\zeta(s)dW^0(s)}\right]dt,\end{aligned}$$

if $N = 1-\int_0^T \kappa(t)\Sigma(t)^{-1}e^{\int_t^T -\kappa(s)ds}dt > 0$.

PROOF. Let θ_λ be a hedging strategy given in (4.16). We plug this into (3.3) and get

$$X^{\theta_\lambda}(t) = x_0 + \int_0^t rX^{\theta_\lambda}(s)+b(s)^T\theta_\lambda(s)ds + \int_0^t \theta_\lambda(s)^T\sigma(s)dW(s).$$

Since $\theta_\lambda \in \Theta$ and σ is bounded,

$$\left\{\int_0^t \theta_\lambda(s)\sigma(s)dW(s); t\in[0,T]\right\}$$

is a martingale and

$$\mathbb{E}\left[\int_0^t \theta_\lambda(s)\sigma(s)dW(s)\right] = 0,$$

for all $t\in[0,T]$. Let $t=T$. Taking the expectation and using (4.16) and (4.10) leads to

$$\begin{aligned}\mathbb{E}\left[X^{\theta_\lambda}\right] &= x_0+\mathbb{E}\int_0^T rX^{\theta_\lambda}(s)+b(s)^T\theta_\lambda(s)ds\\ &= x_0+\mathbb{E}\int_0^T rX^{\theta_\lambda}(s)-b(s)^T\Sigma(s)^{-1}\left[b(s)X^{\theta_\lambda}(s)+\frac{1}{P(s)}\Gamma(s)\right]ds\\ &= x_0+\int_0^T (r-\kappa(s))\,\mathbb{E}\left[X^{\theta_\lambda}(s)\right]-b(s)^T\Sigma(s)^{-1}\frac{1}{P(s)}\mathbb{E}\left[\Gamma(s)\right]ds.\end{aligned}$$

This can be seen as a differential equation by putting $y(T)=\mathbb{E}[X^{\theta_\lambda}(T)]$. For this we get the solution

$$\mathbb{E}\left[X^{\theta_\lambda}(T)\right] = x_0e^{\int_0^T r-\kappa(s)ds} - \int_0^T b(s)\Sigma(s)^{-1}\frac{1}{P(s)}\mathbb{E}\left[\Gamma(s)\right]e^{\int_s^T r-\kappa(u)du}ds. \tag{5.1}$$

We use the explicit formula for P which was given in Proposition 4.0.3 and get

$$\begin{aligned}\mathbb{E}\left[X^{\theta_\lambda}(T)\right] &= x_0 e^{\int_0^T r-\kappa(s)ds} - \int_0^T b(s)\Sigma(s)^{-1}\mathbb{E}\left[\Gamma(s)\right] e^{\int_s^T(-r)du}ds \\ &= x_0 e^{\int_0^T r-\kappa(s)ds} - \int_0^T b(s)\Sigma(s)^{-1}\mathbb{E}\left[\Gamma(s)\right] e^{-r(T-s)}ds.\end{aligned}$$

Using first the notation (4.15) and then (4.10) leads to

$$\begin{aligned}\mathbb{E}\left[X^{\theta_\lambda}(T)\right] &= x_0 e^{\int_0^T r-\kappa(s)ds} - \int_0^T b(s)\Sigma(s)^{-1}\mathbb{E}\left[\Gamma(s)\right] e^{-r(T-s)}ds \\ &= x_0 e^{\int_0^T r-\kappa(s)ds} - \int_0^T b(s)\Sigma(s)^{-1}\mathbb{E}\left[b(s)\varphi(s)+\sigma(s)e_1\psi_0(s)\right] e^{-r(T-s)}ds \\ &= x_0 e^{\int_0^T r-\kappa(s)ds} - \int_0^T \kappa(s)\mathbb{E}\left[\varphi(s)\right] e^{-r(T-s)} + \zeta(s)\mathbb{E}\left[\psi_0(s)\right] e^{-r(T-s)}ds.\end{aligned}$$

By Lemma 4.0.1 we know that the solution for (4.2) can be found by using the solution of (4.3) with

$$\lambda = 1+\gamma\mathbb{E}\left[V^{\theta_\lambda}(T)\right]. \tag{5.2}$$

So by (3.4), (4.7) and the above calculations for $\mathbb{E}\left[X^{\theta_\lambda}(T)\right]$ we get

$$\begin{aligned}\lambda &= 1+\gamma\mathbb{E}\left[X^{\theta_\lambda}(T)-G(S_0;T)\right] \\ &= 1+\gamma x_0 e^{\int_0^T r-\kappa(s)ds} - \gamma\int_0^T \kappa(s)\Sigma(s)^{-1}\mathbb{E}\left[\varphi(s)\right] e^{-r(T-s)}ds \\ &\quad -\gamma\int_0^T \zeta(s)\mathbb{E}\left[\psi_0(s)\right] e^{-r(T-s)}ds - \gamma\mathbb{E}\left[G(S_0;T)\right].\end{aligned} \tag{5.3}$$

In order to find an expression for λ we need to make sure which terms on the right-hand side of (5.3) depend on λ. Since $D_s\xi = D_s(G(S_0;T)+\frac{\lambda}{\gamma}) = D_s(G(S_0;T))$ in the formula for $\mathbb{E}[\psi_0]$ given in Proposition 4.0.8, we can see that the only part depending on λ is $\mathbb{E}[\varphi(s)]$ which is given in Proposition 4.0.8. By using (4.7) we get

$$\begin{aligned}\mathbb{E}[\varphi(s)] &= e^{-\int_s^T \kappa(u)-r+\frac{1}{2}\zeta(u)^2du}\mathbb{E}\left[-\xi e^{-\int_s^T\zeta(u)dW^0(u)}\right] \\ &= -e^{-\int_s^T \kappa(u)-r+\frac{1}{2}\zeta(u)^2du}\mathbb{E}\left[G(S_0;T)e^{-\int_s^T\zeta(u)dW^0(u)}\right] \\ &\quad -\frac{\lambda}{\gamma}e^{-\int_s^T \kappa(u)-r+\frac{1}{2}\zeta(u)^2du}\mathbb{E}\left[e^{-\int_s^T\zeta(u)dW^0(u)}\right].\end{aligned}$$

Now we notice that by Novikov's condition (Proposition 1.5.17) we have an exponential martingale with

$$\mathbb{E}\left[e^{-\int_s^T\zeta(u)dW^0(u)-\frac{1}{2}\zeta(u)^2du}\right] = 1. \tag{5.4}$$

With this we get $\mathbb{E}[\varphi(s)]$ simplified to

$$\mathbb{E}[\varphi(s)] = -e^{-\int_s^T \kappa(u)-r+\frac{1}{2}\zeta(u)^2du}\mathbb{E}\left[G(S_0;T)e^{-\int_s^T\zeta(u)dW^0(u)}\right] - \frac{\lambda}{\gamma}e^{-\int_s^T\kappa(u)-rdu}.$$

Substituting this into (5.3) leads to

$$\begin{aligned}
\lambda &= 1-\gamma\mathbb{E}\left[G(S_0;T)\right]+\gamma x_0 e^{\int_0^T r-\kappa(s)ds}-\gamma\int_0^T \zeta(s)\mathbb{E}\left[\psi_0(s)\right]e^{-r(T-s)}ds \\
&\quad -\gamma\int_0^T \kappa(s)\Sigma(s)^{-1}e^{-r(T-s)}\Big\{-e^{-\int_s^T \kappa(u)-r+\frac{1}{2}\zeta(u)^2du}\mathbb{E}\left[G(S_0;T)e^{-\int_s^T \zeta(u)dW^0(u)}\right] \\
&\quad -\frac{\lambda}{\gamma}e^{-\int_s^T \kappa(u)-rdu}\Big\}ds \\
&= 1-\gamma\mathbb{E}\left[G(S_0;T)\right]+\gamma x_0 e^{\int_0^T r-\kappa(s)ds}-\gamma\int_0^T \zeta(s)\mathbb{E}\left[\psi_0(s)\right]e^{-r(T-s)}ds \\
&\quad +\gamma\int_0^T \kappa(s)\Sigma(s)^{-1}e^{-\int_s^T \kappa(u)+\frac{1}{2}\zeta(u)^2du}\mathbb{E}\left[G(S_0;T)e^{-\int_s^T \zeta(u)dW^0(u)}\right]ds \\
&\quad +\lambda\int_0^T \kappa(s)\Sigma(s)^{-1}e^{-\int_s^T \kappa(u)du}ds.
\end{aligned}$$

By moving all terms depending on λ to the left-hand side we get

$$\begin{aligned}
&\lambda-\lambda\int_0^T \kappa(s)\Sigma(s)^{-1}e^{-\int_s^T \kappa(u)du}ds \\
&= 1-\gamma\mathbb{E}\left[G(S_0;T)\right]+\gamma x_0 e^{\int_0^T r-\kappa(s)ds}-\gamma\int_0^T \zeta(s)\mathbb{E}\left[\psi_0(s)\right]e^{-r(T-s)}ds \\
&\quad +\gamma\int_0^T \kappa(s)\Sigma(s)^{-1}e^{-\int_s^T \kappa(u)+\frac{1}{2}\zeta(u)^2du}\mathbb{E}\left[G(S_0;T)e^{-\int_s^T \zeta(u)dW^0(u)}\right]ds.
\end{aligned}$$

Now taking λ as common divisor and dividing the equation with

$$N := 1-\int_0^T \kappa(s)\Sigma(s)^{-1}e^{-\int_s^T \kappa(u)du}ds$$

we get provided that $N>0$

$$\begin{aligned}
\lambda &= \frac{1}{N}\Big\{1-\gamma\mathbb{E}\left[G(S_0;T)\right]+\gamma e^{\int_0^T r-\kappa(s)ds}x_0\Big\} \\
&\quad +\frac{\gamma}{N}\int_0^T \zeta(s)\mathbb{E}[\psi_0(s)]e^{-r(T-s)}ds \\
&\quad +\frac{\gamma}{N}\int_0^T \kappa(s)\Sigma(s)^{-1}e^{-\int_s^T \kappa(u)+\frac{1}{2}\zeta(u)^2du}\mathbb{E}\left[G(S_0;T)e^{-\int_s^T \zeta(u)dW^0(u)}\right]ds.
\end{aligned}$$

□

Before we formulate our main result let us recall the setting and summarize the outcome we have so far. We assumed a complete probability space $(\Omega,\mathcal{F},\mathbb{P})$, a time interval $[0,T]$ for some $T>0$ and an arbitrage free financial market with one risk-free asset earning with constant rate $r>0$, $(m+1)$ risky assets $\{S_i(t), t\in[0,T]\}$ for $i=1,2,\dots,m$ with dynamics given in (3.2) and an asset $\{S_0(t), t\in[0,T]\}$ with dynamics (3.1). The asset S_0 was assumed to be not allowed to use in hedging. The expected return rates and volatilities were assumed to be deterministic and Borel-measurable.

The object for which we wanted to find the hedging strategy was the pay-off function $G(S_0;T)$ with maturity $T > 0$. The assumptions for G are given in Assumption 3.2.1. The value process of the hedging portfolio was denoted by X^θ, and the dynamics for this was given by the SDE (3.3) where θ was our hedging strategy.

We consider a profit-and-loss random variable

$$V^\theta(T) = X^\theta(T) - G(S_0;T) \tag{5.5}$$

to describe our hedging error at terminal time T. The mean-variance criterion was used to determine an optimal hedging strategy, meaning that we tried to find a strategy which solves the problem

$$\max_{\theta\in\Theta}\left\{\mathbb{E}\left[V^\theta(T)\right] - \frac{\gamma}{2}\mathrm{Var}\left[V^\theta(T)\right]\right\}, \tag{5.6}$$

where $\gamma > 0$ is assumed to be the weight which the investor give for the variance.

This is the same as solving

$$\min_{\theta\in\Theta}\left\{J_1(\theta;\gamma) = \frac{\gamma}{2}\mathrm{Var}\left[V^\theta(T)\right] - \mathbb{E}\left[V^\theta(T)\right]\right\} \tag{5.7}$$

and finding a solution for this was connected to the auxiliary problem

$$\min_{\theta\in\Theta}\left\{J_2(\theta;\gamma,\lambda) = \mathbb{E}\left[\frac{\gamma}{2}V^\theta(T)^2 - \lambda V^\theta(T)\right]\right\} \tag{5.8}$$

where $\gamma > 0$ and $-\infty < \lambda < \infty$. The connection between the solutions of these two problems was shown in Lemma 4.0.1.

We re-formulated the auxiliary problem to

$$\min_{\theta\in\Theta}\left\{J_3(\theta;\lambda,\gamma) = \mathbb{E}\left[(X^\theta(T) - \xi)^2\right]\right\}, \tag{5.9}$$

where $\xi = G(S_0;T) + \frac{\lambda}{\gamma}$. Solving (5.9) was connected to equations for P and φ given in (4.11) and (4.12). These had explicit solutions

$$P(t) = e^{\int_t^T 2r-\kappa(s)ds} \tag{5.10}$$

$$\varphi(t) = e^{-\int_t^T\left(\kappa(s)-r+\frac{1}{2}\zeta(s)^2\right)ds}\mathbb{E}\left[-\xi e^{\int_t^T\zeta(s)dW^0(s)}|\mathcal{F}_t^0\right] \tag{5.11}$$

$$\psi_0(t) = e^{-\int_t^T\left(\kappa(s)-r+\frac{1}{2}[\zeta(s)]^2\right)ds}\mathbb{E}\left[-e^{-\int_t^T\zeta(s)dW^0(s)}D_t\xi|\mathcal{F}_t^0\right]. \tag{5.12}$$

Especially we got for the expectation of $\psi_0(t)$

$$\mathbb{E}[\psi_0(t)] = e^{-\int_t^T\left(\kappa(s)-r+\frac{1}{2}[\zeta(s)]^2\right)ds}\mathbb{E}\left[-e^{-\int_t^T\zeta(s)dW^0(s)}D_t\xi\right]. \tag{5.13}$$

With the above solutions, Theorem 4.0.4 showed that (5.9) is solvable with the solution

$$\theta_\lambda(t) = -\Sigma(t)^{-1}\left[b(t)X^{\theta_\lambda}(t) + P(t)\Gamma(t)\right] \tag{5.14}$$

and in Proposition 5.0.1 we derived the explicit formula for the λ connected to the optimal hedging strategy.

Collecting all above together we are able to formulate the optimal hedging strategy which solves the mean-variance hedging problem:

THEOREM 5.0.2. *Assume the above setting and consider the mean-variance hedging problem*

$$\max_{\theta\in\Theta}\left\{\mathbb{E}\left[V^{\theta}(T)\right]-\frac{\gamma}{2}Var\left[V^{\theta}(T)\right]\right\},$$

where $\gamma>0$, $V^{\theta}(T)=X^{\theta}(T)-G(S_0;T)$ *is the profit-and-loss random variable at terminal time* $T>0$ *with the hedging portfolio* X^{θ} *and a Malliavin differentiable pay-off function* $G(S_0;T)\in\mathbb{D}_{1,2}$. *Then the optimal hedging strategy is*

$$\theta^*(\cdot)=\theta_\lambda(\cdot)=-\Sigma(\cdot)^{-1}\left[b(\cdot)X^{\theta_\lambda}(\cdot)+P(\cdot)\gamma(\cdot)\right],$$

where P *is given in* (5.10), φ *in* (5.11), $\mathbb{E}[\psi_0]$ *in* (5.13) *and*

$$\begin{aligned}\lambda &= \frac{1}{N}\left\{1-\gamma\mathbb{E}\left[G(S_0;T)\right]+\gamma e^{\int_0^T r-\kappa(u)du}x_0\right\}\\ &+\frac{\gamma}{N}\int_0^T\zeta(t)\mathbb{E}[\psi_0(t)]e^{-r(T-t)}dt\\ &+\frac{\gamma}{N}\int_0^T\kappa(t)\Sigma(t)^{-1}e^{-\int_t^T\kappa(u)+\frac{1}{2}\zeta(u)^2du}\mathbb{E}\left[G(S_0;T)e^{-\int_t^T\zeta(u)dW^0(u)}\right]dt\end{aligned}$$

with $N=1-\int_0^T\kappa(t)\Sigma(t)^{-1}e^{\int_t^T-\kappa(s)ds}dt>0$.

PROOF. The result will follow by Theorem 4.0.4, Lemma 4.0.1 and Proposition 5.0.1.

□

REMARK 5.0.3. By varying the weight γ we get a family of solutions.

APPENDIX A

Here we collect some results which are needed in the calculations of this thesis.

Proposition 1.0.1 ([**22**]Theorem 4.2.14). *Let Z be a normed linear space and Y a complete normed linear space. Assume a linear set X such that $X \subseteq Z$ and let $A : X \to Y$ be a bounded linear operator. If X is dense in Z, then A can be extended to space Z with preserving the norm.*

Definition 1.0.2. Let $\Omega \neq \emptyset$. Then non-empty system $\mathcal{P}$ of subsets of Ω is called π-system if

$$A \cap B \in \mathcal{P} \text{ for all } A, B \in \mathcal{F}.$$

The following proposition states a powerful tool called Monotone class theorem for functions.

Proposition 1.0.3 (Monotone class for functions, ([**7**] Proposition 9.3.13)). *Let $\mathcal{A} \subseteq 2^{\Omega}$ be a π-system which contains Ω and assume $\mathcal{H} \subseteq \{f; f : \Omega \to \mathbb{R}\}$ such that*

(1) $\mathbb{1}_A \in \mathcal{H}$ for all $A \in \mathcal{A}$,
(2) linear combinations of elements of $\mathcal{H}$ are also in $\mathcal{H}$,
(3) if $(f_n)_{n=1}^{\infty} \subseteq \mathcal{H}$ such that $f_n \uparrow f$ and f is bounded, this implies $f \in \mathcal{H}$.

Then $\mathcal{H}$ contains all bounded functions f, which are $\sigma(\mathcal{A})$-measurable.

Next we assume spaces X, Y to be Banach spaces and an operator A to be linear. We denote the domain of A by $D(A) \subseteq X$ and the image set by $R(A) \subseteq Y$. Then we have next definition adapted from [**23**] (where only the case $X = Y$ was considered):

Definition 1.0.4. A linear operator A with $D(A) \subseteq X$ and $R(A) \subseteq Y$ is called a closed operator if for every Cauchy sequence $(x_n)_{n=1}^{\infty} \subseteq D(A)$ such that $x_n \to x$ and $Ax_n \to y$ it follows that $x \in D(A)$ and $Ax = y$.

Definition 1.0.5 ([**8**], Definition 1.3.5). Let X and Y be Banach spaces and let $S \subseteq X$ be a linear subspace. A linear operator $A : S \to Y$ is called closable if for any $(x_n)_{n=1}^{\infty} \subseteq S$ such that $x_n \to 0$ and $Ax_n \to y$ it follows that $y = 0$.

A closable linear operator can be extended in the following way:

Proposition 1.0.6. *If a linear operator $A : D(A) \to Y$ is closable, then it does have a closed extension $\bar{A} : D(\bar{A}) \to Y$ such that $A = \bar{A}$ on $D(A)$ and $D(A) \subseteq D(\bar{A})$.*

Proof. We define an extension

$$D = \{x \in X : \text{ there exists } (x_n)_{n=1}^{\infty} \subseteq D(A) \text{ with } x_n \to x \text{ and } Ax_n \to y\}$$

and $\bar{A} : D \to Y$ by $\bar{A}x \to y$. We show that $(\bar{A}, D)$ is well-defined, linear and closed.

(1) $\bar{A}$ is well-defined: Let us assume that $x_n \to x$ and $Ax_n \to y$ and another sequence $(z_n)_{n=1}^{\infty} \subseteq D(A)$ which also converges to x, but $Az_n \to w$. Then $x_n - z_n \to 0$ and $A(z_n - x_n) \to y - w$. Since A is closable we have that $y - w = 0$ which is equal to $y = w$.

(2) $\bar{A}$ is linear: Let $\alpha, \beta \in \mathbb{R}$ and x_1, x_2. By definition of D there exists sequences $(x_1^{(n)})_{n=1}^{\infty}$ and $(x_2^{(n)})_{n=1}^{\infty}$ such that $x_1^{(n)} \to x_1$ and $x_2^{(n)} \to x_2$ with $Ax_1^{(n)} \to \bar{A}x_1$ and $Ax_2^{(n)} \to \bar{A}x_2$, when $n \to \infty$. Now

$$\begin{aligned} \alpha\bar{A}x_1 + \beta\bar{A}x_2 &= \alpha \lim_{n\to\infty} Ax_1^{(n)} + \beta \lim_{n\to\infty} Ax_2^{(n)} \\ &= \lim_{n\to\infty} A(\alpha x_1^{(n)} + \beta x_2^{(n)}) \\ &= \bar{A}(\alpha x_1 + \beta x_2), \end{aligned}$$

since A is a linear operator.

(3) $\bar{A}$ is closed: Assume that $(x_n)_{n=1}^{\infty} \subseteq D$ such that $x_n \to x$ and $\bar{A}x_n \to y$. We show that then $x \in D$ and $\bar{A}x = y$. If $(x_n)_{n=1}^{\infty} \subseteq D$ then there exists $(z_n)_{n=1}^{\infty} \in D(A)$ with

$$\|x_n - z_n\|_X \to 0$$

and

$$\|\bar{A}x_n - \bar{A}z_n\|_Y \to 0$$

when $n \to \infty$. Now since

$$\begin{aligned} \|x - z_n\|_X &\leq \|x - x_n\|_X + \|x_n - z_n\|_X \\ &\to 0, \text{ as } n \to \infty \end{aligned}$$

and

$$\begin{aligned} \|y - Az_n\|_Y &\leq \|y - \bar{A}x_n\|_Y + \|\bar{A}x_n - Az_n\|_Y \\ &\to 0 \text{ as } n \to \infty \end{aligned}$$

it follows that $x \in D$ and $\bar{A}x = y$.

□

APPENDIX B

Notations

$L^2([0,T]^m)$ The set of Borel functions $f:[0,T]^m \to \mathbb{R}$ with $\|f\|^2_{L^2([0,T]^m)} = \int_{[0,T]^m} f^2 d\lambda^m < \infty$, also denoted by $L^2([0,T]^m, \mathfrak{B}([0,T]^m), \lambda^m)$.

$\mathcal{L}_0$ The set of simple stochastic processes given in Definition 1.5.1, $L = (L_t)_{t\in[0,T]}$, $L_t : \Omega \to \mathbb{R}^m$.

$\mathcal{L}_2$ The set of progressively measurable stochastic processes $L = (L_t)_{t\in[0,T]}$, $L_t : \Omega \to \mathbb{R}^m$ with $\left(\mathbb{E}\int_0^T L_t^2 dt\right)^{\frac{1}{2}} < \infty$.

$\mathcal{L}_2^{loc}$ The set of progressively measurable stochastic processes $L = (L_t)_{t\in[0,T]}$, $L_t : \Omega \to \mathbb{R}^m$ with $\mathbb{P}\left(\omega \in \Omega : \int_0^T L_u(\omega)^2 du < \infty\right) = 1$.

$L^p(\Omega, \mathcal{F}, \mathbb{P})$ The space of random variables $f : \Omega \to \mathbb{R}$ with $(\mathbb{E}[f^p])^{\frac{1}{p}} < \infty$ for $p \geq 1$.

$\mathcal{M}_2^c$ The set of continuous martingales $M = (M_t)_{t\in[0,T]}$, $M_t : \Omega \to \mathbb{R}$ with $\mathbb{E}\left[M_t^2\right] < \infty$.

$\mathcal{M}_2^{c,0}$ The set of continuous martingales $M = (M_t)_{t\in[0,T]}$, $M_t : \Omega \to \mathbb{R}$ with $M_0 = 0$ and $\mathbb{E}\left[M_t^2\right] < \infty$.

$\mathcal{M}_{loc}^{c,0}$ The set of continuous local martingales given in Definition 1.4.7.

$\mathcal{E}_m$ The set of elementary functions $f : [0,T]^m \to \mathbb{R}$ of the form $f(t_1, ..., t_m) = \sum_{i_1,\ldots,i_m=1}^n a_{i_1\cdots i_m} \mathbb{1}_{A_{i_1}\times\cdots\times A_{i_m}}(t_1, ..., t_m)$ where $a_i \in \mathbb{R}$ and $a_{i_1\cdots i_m} = 0$, if $i_k = i_j$ for some $k \neq j$ and $A_k \cap A_l = \emptyset$ for all $k \neq l$ with $k, l \in \{1, 2, ..., n\}$.

$\mathcal{S}_2$ The set of all adapted and continuous stochastic processes $L = (L_t)_{t\in[0,T]}$, $L_t : \Omega \to \mathbb{R}$, with $\mathbb{E}\left[\sup_{0\leq t\leq T} L_t^2\right] < \infty$.

$C([0,T])$ The set of continuous functions $f : [0,T] \to \mathbb{R}$.

$C_p^\infty(\mathbb{R}^n)$ The set of functions $f : \mathbb{R}^n \to \mathbb{R}$ such that $\frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1}\cdots\partial x_n^{\alpha_n}}$ exists for all $|\alpha| = \alpha_1 + ... + \alpha_n$, where $\alpha_i \in \mathbb{N}_0$, f and $\frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1}\cdots\partial x_n^{\alpha_n}}$ have polynomial growth.

S	The set of smooth random variables of the form $F = f(W(h_1), \ldots, W(h_n))$ where $f \in C_p^\infty(\mathbb{R}^n)$ and $h_1, ..., h_n \in L^2([0,T]^m)$.
$\mathbb{D}_{1,p}$	The domain of the Malliavin derivative operator in space $L^p(\Omega, \mathcal{F}, \mathbb{P})$.
Θ	The set of hedging strategies $\theta : [0,T] \to \mathbb{R}^m, \theta(t) = (\theta_1(t), ..., \theta_m(t))^T$, with $\theta \in \mathcal{L}_2$.

Bibliography

[1] Bauer, H. (1995) *Probability Theory.* De Gruyter, Inc. Proquest Ebook https://ebookcentral.proquest.com/lib/jyvaskyla-ebooks/detail.action?docID=928950.

[2] Chen, S., Li, X., Zhou, X. Y. (1998). *Stochastic linear quadratic regulators with indefinite control weight costs.* SIAM J. Control Optim, 36(5), 1685-1702.

[3] Curtain, R. F. and Pritchard, A. J. (1977) *Functional Analysis in Modern Applied Mathematics.* Academic Press Inc., London.

[4] Dhaene, J., Kukush, A. and Linders, D. (2012). *The Multivariate Black & Scholes Market: Conditions for Completeness and No-Arbitrage.* Theory of Probability and Mathematical Statistics, 88, 1-14, 2013 Available at SSRN: https://ssrn.com/abstract=2186830 or http://dx.doi.org/10.2139/ssrn.2186830

[5] Duistermaat, J.J. and Kolk, J.A.C. (2004) *Multidimensional Real Analysis 1: Differentiation.* Cambridge University Press.

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

[7] Geiss, C. and Geiss, S. (28.10.2020). *An Introduction to Probability Theory.* University of Jyväskylä, Department of Mathematics and Statistics.

[8] Geiss, S. (5.8.2021). *Gaussian Hilbert spaces.* University of Innsbruck, Department of Mathematics.

[9] Geiss, C. and Steinicke, A. (2016). *Malliavin derivative of random functions and applications to Lévy driven BSDEs.* Electron. J. Probab., 21(10), 1-28.

[10] Geiss, S. (23.2.2021). *Stochastic Differential Equations.* University of Jyväskylä, Department of Mathematics and Statistics.

[11] Geiss, S. (24.7.2021). *Stochastic Finance in discrete and continuous time.* University of Jyväskylä, Department of Mathematics and Statistics.

[12] Itô, K. (1951). *Multiple Wiener Integral.* Journal of the Mathematical Society of Japan, 3(1), 157-169.

[13] Kallianpur, G. and Sundar, P. (2014). *Stochastic Analysis and Diffusion Processes.* Oxford University Press.

[14] Karatzas, I. and Shreve, S. E. (1997). *Brownian motion and stochastic calculus.* Springer, New York.

[15] Lamberton, D. and Lapeyre, B. (1996) *Stochastic Calculus Applied to Finance.* Chapman and Hall.

[16] Lim, A. E., Shou, X. Y. (2002). *Mean-variance portfolio selection with random parameters in a complete market.* Mathematics of Operations Research, 27, 101-120.

[17] Markowitz, H. (1952). *Portfolio Selection.* The Journal of Finance, 7(1), 77-91.

[18] Nualart, D. (1995). *The Malliavin Calculus and Related Topics.* Springer Science+Business Media.

[19] Nualart, D. (2006). *The Malliavin Calculus and Related Topics.* Springer, Berlin Heidelberg.

[20] Nualart, D. and Nualart, E. (2018). *Introduction to Malliavin Calculus.* Cambridge: Cambridge University Press.

[21] Schilling, R. L. and Partzsch, L. (2012). *Brownian Motion: An Introduction to Stochastic Processes.* De Gruyter, Berlin.

[22] Sen, R. (2012) *A First Course in Functional Analysis: Theory and Applications.* Anthem Press, ProQuest Ebook https://ebookcentral.proquest.com/lib/jyvaskyla-ebooks/detail.action?docID=1130116.

[23] Triebel, H. (1992). *Higher Analysis*. Johan Ambrosius Barth Verlag.

[24] Xue, X., Zhang, J., Weng, C. (2019). *Mean-variance Hedging with Basis Risk*. Applied Stochastic Models in Business and Industry, Forthcoming, Available at SSRN: https://ssrn.com/abstract=3384769.

[25] Yong, J.M., Zhou, X. Y. (1999). *Stochastic controls: Hamiltonian systems and HJB equations*. Springer-Verlag, New York.

[26] Zhou, X. Y., Li, D. (2000). *Continuous-Time Mean-Variance Portfolio Selection: A Stochastic LQ Framework*. Applied Mathematics & Optimization, 42, 19-33.