UiO : Department of Mathematics
University of Oslo

Portfolio Optimization in a Co-Integrated Asset Market Model

Sejla Ackar

Master's Thesis, Spring 2019

This master's thesis is submitted under the master's programme *Modelling and Data Analysis*, with programme option *Finance, Insurance and Risk*, at the Department of Mathematics, University of Oslo. The scope of the thesis is 60 credits.

The front page depicts a section of the root system of the exceptional Lie group E_8, projected into the plane. Lie groups were invented by the Norwegian mathematician Sophus Lie (1842–1899) to express symmetries in differential equations and today they play a central role in various parts of mathematics.

Abstract

The main objective of this thesis is to investigate the optimal management of portfolios under a co-integrated market model. Given a finite investment horizon, the thesis considers a risk-averse investor who has the limited choice of investing in two risky assets and one risk free, with the aim of maximizing expected utility of wealth. A power utility function represents the investor's appetite for risk. By the dynamic programming approach from stochastic control theory, a semi-explicit solution to the resulting HJB-equation is obtained as the solution to a set of Riccati differential equations. Assuming a solution exist, a verification theorem is presented. By simplifying the market model, a stochastic Feynman-Kac representation is found, and motivated by Merton's constant fraction solution, a naive approach to find constant controls is presented.

Acknowledgements

First and foremost, I would like to thank my supervisor, Professor Fred Espen Benth. Thank you for giving me an interesting project to work on, and for your good guidance. Thank you for your positivity every time my solutions failed, and for encouraging and motivating me in stressful times.

Secondly, I would like to thank my mom and dad for always being my greatest supporters and for dealing with me this past year (plural). I will forever be grateful for your endless support and love. I could not have done this without you.

I would also like to send a special thank you to my dearest grandmother for introducing me to mathematics in my early years and for being such a great inspiration in every possible way.

At last, but certainly not least, I would like to thank my friends and fellow students. A special thanks to Mari Dahl Eggen for an infinitely countable number of coffee breaks, lunches and good and engaging talks, both mathematical and social. Also, thank you to Patrick Henriksen and Vemund S. Schøyen for making me laugh my way through five years at the university with their terrible sense of humor. If «laughter prolongs life», I will live a long time.

Sejla Ackar,
Oslo, May 2019.

Contents

Chapter 1

Introduction

The main objective of portfolio optimization is to find the optimal trading strategy with regards to some prescribed criteria, usually maximization of expected returns or minimization of risk. One way to formulate such a portfolio optimization problem is through the theory of stochastic optimal control. Around the 1970s, Robert C. Merton introduced the concept of dynamic programming to portfolio optimization problems. In [Mer75], Merton considers the optimal investment and consumption of a risk-averse investor who aims at maximizing expected utility of wealth at the end of a finite time horizon. Prior to Merton, portfolio optimization problems were mostly studied in discrete time. In 1952, Markowitz made the first attempt to solve a portfolio optimization problem through the mean-variance approach. In [Mar52], Markowitz tackles the problem of maximizing expected returns, while minimizing the variance of the returns of the resulting portfolio, and the framework for two important aspects of portfolio management is introduced; diversification and risk-reward trade. By spreading the investments of a portfolio in uncorrelated assets, an investor is able to secure oneself against potential risk, while increasing reward. In the long run, one hopes that the downfalls of one investment are neutralized by the success of another. However, one of the clear disadvantages of the Markowitz mean-variance approach is the one-period assumption. The optimal allocation of wealth must be made at the beginning of the investment horizon. Afterwards, the investor becomes a passive agent observing the price fluctuations of the assets without the possibility of reallocating the portfolio. By Merton's problem, through the dynamic programming approach, the investor is allowed to continuously change the allocation of wealth in order to maximize expected utility at some future time point.

Motivated by Merton's portfolio problem, this thesis considers a risk-averse investor in a co-integrated market model. The concept of co-integration was formally introduced in econometrics in 1987 by Granger and Engle. Following [EG87], two stochastic processes are said to be co-integrated if there exists a stationary linear combination of the processes, even if the processes themselves are non-stationary. The asset prices in the co-integrated market in this thesis, are modelled as the exponential of a common non-stationary trend process, a drifted Brownian motion, and two distinct stationary Ornstein-Uhlenbeck processes. On a logarithmic scale, the difference of the asset prices is stationary. As the financial market is divided into several sector, for instance oil, energy, IT or Telecom services, from a financial point of view, the co-integrated market model can be used to represent two companies in the same market segment, where the asset prices follow a common non-stationary trend, whilst still having their idiosyncratic risk. Typically one would spread the investments to diversify risk. Take for instance two offshore drilling companies in the oil sector, where the price of the stocks for the two companies is assumed to simultaneously follow the price of oil, while still being exposed to risk factors distinctive for each of the companies.

The scope of this thesis is to understand co-integration in relation to the optimal management of portfolios. The aim of the investor is to maximize expected utility of wealth at the end of the investment horizon. The optimal portfolio problem is presented as an utility maximization problem from stochastic optimal control theory, where an utility function represents the investors appetite for risk. The problem is addressed by the dynamic programming approach, as presented by Merton, resulting in the non-linear Hamilton-Jacobi-Bellman equation to which a semi-explicit solution is found and verified.

1.1 Outline of the Thesis

The thesis is structured as follows: Chapter 2 gives an overview of the general market model. Descriptions of the stochastic asset price processes are given, with introductions to the stationary CARMA and Ornstein-Uhlenbeck processes, in addition to stating some main properties. The concept of co-integration is explained and a multi-dimensional asset model is presented. In Chapter 3 the main elements of stochastic control theory, relevant to the thesis, are introduced. The basic assumptions and a formulation of a general control problem for Markov processes is stated, the concept of dynamic programming is explained and the Hamilton-Jacobi-Bellman equation is introduced. The chapter is concluded by Merton's well-known standard portfolio optimization problem, which serves as a motivation for Chapter 4. Chapter 4 constitutes the main objective of the thesis. The optimal control problem is presented, along with theorems for existence and verification of optimal control solutions. A semi-explicit solution is derived and a system of equations solving the resulting PDE is given. Two alternative representations are presented, and a verification concludes the findings. The last section serves as an investigator of "what went wrong" and gives a stochastic representation of the solution to a portfolio optimization problem in a simpler market model. Chapter 5 is motivated by Merton's solution to the case of portfolio optimization in a two asset model. An unconventional approach to finding admissible, but not necessarily optimal, controls, is presented. Finally, Chapter 6 gives some concluding remarks and pointers to future work.

The computations of this thesis turned out to be quite tedious, hence notational conventions for constants and functions have been introduced throughout. They are all stated in Appendix A. Appendix B presents preliminaries from stochastic analysis, probability theory and general measure theory. Central theorems, definitions and properties which are stated and used in the thesis, some probably familiar, and some more specific to control theory, are stated. Appendix C presents some of the omitted calculations from Chapters 4 and 5.

Chapter 2

The Market Model

2.1 Asset Market Model

On a complete filtered probability space $(\Omega, \mathcal{F}, \mathcal{P}, \{\mathcal{F}_t\}_{t\geq 0})$ we define three correlated Brownian motions $\widetilde{B}$ and B_i for $i = 1, 2$. We denote by Ω the product space $\Omega_1 \times \Omega_2 \times \Omega_3$, and by $\mathcal{F}$ the σ-algebra $\mathcal{F}_1 \times \mathcal{F}_2 \times \mathcal{F}_3$, where $\mathcal{F}_t$ is the natural filtration generated by $\widetilde{B}(t), B_1(t), B_2(t)$. Assume a complete market where there are no transaction costs when buying or selling assets, and where all investors have perfect and instantaneous information about asset prices. Following [BK15] we let S_1 and S_2 be asset prices, typically stocks, defined by

$$S_1(t) = \exp(c_1X(t) + Y_1(t))S_2(t) \qquad = \exp(c_2X(t) + Y_2(t)) \tag{2.1}$$

where the processes $Y_1(t)$ and $Y_2(t)$ are two stationary CARMA processes (Section 2.2) and $X(t)$ is a non-stationary, drifted Brownian motion with dynamics

$$dX(t) = \mu dt + \sigma d\widetilde{B}(t) \tag{2.2}$$

The parameters μ and $\sigma > 0$ are constants, respectively the drift and diffusion coefficients. c_1 and c_2 can be interpreted as constant conversion factors between two market segments, for instance energy companies and oil producers. The conversion factors are useful if one would like to measure two assets from different segments on the same scale. Letting $c_1 = c_2 = 1$, we assume two segments have the same conversion factor.

The correlation coefficients between the Brownian motions are denoted by $\mathrm{Corr}(\widetilde{B}, B_i) = \rho_i$ for $i = 1, 2$ and $\mathrm{Corr}(B_1, B_2) = \rho$, where the Brownian motions are all assumed to be Gaussian processes with mean zero and variance t. The correlation matrix of the Brownian motions is given by

$$P = \begin{pmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho \\ \rho_2 & \rho & 1 \end{pmatrix} \tag{2.3}$$

where P is symmetric and positive definite. Hence by Cholesky decomposition of P, the correlated Brownian motions can be represented by

$$\begin{bmatrix} d\widetilde{B} \\ dB_1 \\ dB_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ \rho_1 & \sqrt{1-\rho_1^2} & 0 \\ \rho_2 & \dfrac{\rho-\rho_1\rho_2}{\sqrt{1-\rho_1^2}} & \sqrt{1-\rho_2^2-\dfrac{(\rho-\rho_1\rho_2)^2}{1-\rho_1^2}} \end{bmatrix} \begin{bmatrix} dU_1 \\ dU_2 \\ dU_3 \end{bmatrix} \tag{2.4}$$

where $U_1, U_2, U_3 \sim \mathcal{N}(0,1)$ are three independent standard normal Brownian motions. For the matrix in (2.4) to be well defined as the Cholesky decomposition of P, the matrix must have strictly positive diagonal entries, implying

$$\rho^2 + \rho_1^2 + \rho_2^2 - 2\rho\rho_1\rho_2 \leq 1 \tag{2.5}$$

Equation (2.5) implies a condition on the correlation coefficients ρ, ρ_i for the correlated Brownian motions to be well defined, yielding Brownian motions $\widetilde{B}(t)$, $B_i(t)$ with mean value equal to 0 and variance given by t.

2.2 CARMA Processes

Following [BK15], we let $Y_i(t), i = 1,2$ in (2.1) be two continuous autoregressive moving average (CARMA(p_i, q_i), $q_i > p_i$) processes defined by

$$Y_i(t) = \mathbf{b}_i^T \mathbf{Z}_i(t), \quad i = 1,2 \tag{2.6}$$

where $\mathbf{Z}_i(t)$ is a p_i-dimensional Ornstein-Uhlenbeck process and b_i^T is the transpose of the column vector $b_i = [b_{1,i}, b_{2,i}, \ldots, b_{q_i} = 1, 0, \ldots, 0] \in \mathbb{R}^{p_i}$. Equation (2.6) then takes the form

$$Y_i(t) = b_{1,i} Z_i^1(t) + b_{2,i} Z_i^2(t) + \ldots + Z_i^{q_i}(t) \tag{2.7}$$

where the Ornstein-Uhlenbeck processes are defined by

$$d\mathbf{Z}_i(t) = A_i \mathbf{Z}_1(t)dt + \sigma_i \mathbf{e}_{p_i} dB_i(t), \quad i = 1,2 \tag{2.8}$$

for constants $\sigma_i > 0$, where $\mathbf{e}_{p_i}$ the p_ith canonincal coordinatevector and

$$A = \begin{bmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \vdots & & & & \\ -\alpha_{p_i,i} & -\alpha_{p_{i-1},i} & -\alpha_{p_{i-2},i} & \ldots & -\alpha_{1,i} \end{bmatrix} \tag{2.9}$$

a $p_i \times p_i$-matrix. For negative real eigenvalues of A_i, $Y_i(t)$ will be a stationary CARMA process.

Lemma 2.2.1 (2-dimensional OU-process).
Assume $\boldsymbol{Z}(t)$ is a 2-dimensional Ornstein-Uhlenbeck process defined by

$$d\boldsymbol{Z}(t) = A\boldsymbol{Z}(t)dt + \sigma \boldsymbol{e}_2 dB(t)$$

where

$$A = \begin{bmatrix} 0 & 1 \\ -\alpha_1 & -\alpha_2 \end{bmatrix}, \boldsymbol{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Then $\boldsymbol{Z}(t)$ will always be stationary.

Solution to (2.2.1). The characteristic polynomial of A is given by

$$p(\lambda) = \det(A - \lambda I) \tag{2.10}$$

where I is the 2×2 identity matrix. (2.10) equals

$$p(\lambda) = \lambda^2 + \lambda\alpha_2 + \alpha_1 \tag{2.11}$$

for which a solution is given by

$$\lambda = -\frac{1}{2}\alpha_{11} \pm \frac{1}{2}\sqrt{\alpha_{11}^2 - 4\alpha_{21}} \tag{2.12}$$

If $\alpha_{11}^2 < 4\alpha_{21}$, the polynomial has two complex roots, with $\text{Re}(\lambda) = -\frac{1}{2}\alpha_{11} < 0$. If however $\alpha_{11}^2 \geq 4\alpha_{21}$, the polynomial has two real roots where $\alpha_{11}^2 - 4\alpha_{21} < \alpha_{11}^2$, i.e. $-4\alpha_{21} < 0$. In either case, $\lambda(A) < 0$, and hence $\mathbf{Z}(t) \in \mathbb{R}^2$ is always stationary. ■

The CARMA processes are a nice generalization of the Ornstein-Uhlenbeck processes, widely used in for instance weather and energy markets, but in this thesis we consider the case where $p_i = 1, q_i = 0$, for which $Y_i(t) = Z_i(t)$ is simply a 1-dimensional Ornstein-Uhlenbeck process, or a CARMA(1,0), CAR(1) process. Note that for $p_i = 1$, $Z_i(t)$ is a stationary process if $A_i < 0$, for A_i a constant. We will for simplicity throughout denote this *constant* by A_i and assume $A_i < 0$.

2.3 Ornstein-Uhlenbeck Processes

Let the processes $Z_i(t) \in \mathbb{R}, i = 1, 2$ be defined by

$$dZ_i(t) = A_i Z_i(t)dt + \sigma_i dB_i(t) \tag{2.13}$$

i.e. the 1-dimensional case of (2.13).

Proposition 2.3.1 (Ornstein-Uhlenbeck process).
The solution $Z_i(t)$ to (2.13) is given by

$$Z_i(t) = z_i \exp(A_i t) + \int_0^t \sigma_i \exp(A_i(t-s))dB_i(s), \quad i = 1, 2 \tag{2.14}$$

for $Z_i(0) = z_i$, $A_i < 0$ constant, where $Z_i(t)$ is a Ornstein-Uhlenbeck process.

Proof of proposition (2.3.1). Let $f(t, z) = z_i \exp(-A_i t)$, then
$f(t, Z_i(t)) = Z_i(t)\exp(-A_i t)$ is a stochastic process, and by Itô's lemma

$$\begin{aligned} d(Z_i(t)\exp(A_i t)) &= \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial z}dZ_i(t) + \frac{1}{2}\frac{\partial^2 f}{\partial z_i^2}(dZ_i(t))^2 \\ &= -A_i Z_i(t)\exp(-A_i t)dt + \exp(-A_i t)dZ_i(t) \\ &= -A_i Z_i(t)\exp(-A_i t)dt + \exp(-A_i t)\Big(A_i Z_i(t)dt + \sigma_i dB_i(t)\Big) \\ &= \sigma_i \exp(-A_i t)dB_i(t) \end{aligned}$$

Integration on both sides yields

$$Z_i(t)\exp(-A_it) = z_i + \int_0^t \sigma_i \exp(-A_is)dB_i(s) \tag{2.15}$$

Dividing by $\exp(-A_it)$ we obtain (2.14). ■

Proposition 2.3.2 (Statistical properties of the OU-process).
Let $Z_i(t)$ be defined as in (2.14). Then the mean and variance of $Z_i(t)$ is given by

$$E[Z_i(t)] = z_i\exp(A_it) \tag{2.16}$$

$$\mathrm{Var}(Z_i(t)) = \frac{\sigma_i^2}{2A_i}\Big[\exp(2A_it)-1\Big] \tag{2.17}$$

with limiting (stationary) distribution $Z_i(t) \sim \mathcal{N}\Big(0, \frac{\sigma_i^2}{2A_i}\Big)$ for $A_i < 0$. In addition, the covariance between $Z_1(t)$ and $Z_2(t)$ is given by

$$\mathrm{Cov}(Z_1(t), Z_2(t)) = \frac{\rho\sigma_1\sigma_2}{A_1+A_2}\Big[\exp((A_1+A_2)t)-1\Big] \tag{2.18}$$

Proof of proposition (2.3.2). For the expected value, note that

$$\begin{aligned}E[Z_i(t)] =& E\Big[z_i\exp(A_it) + \int_0^t \sigma_i\exp(A_i(t-s))dB_i(s)\Big]\\ =& z_i\exp(A_it)\end{aligned}$$

since $E[\int_0^t g(s)dB_i(s)] = 0$ for all $\mathcal{F}$-measurable functions $g \in \mathcal{V}([0,T])$, see property (iii) of Proposition B.1.4. Following,

$$\begin{aligned}\mathrm{Var}(Z_i(t)) =& \mathrm{Var}\Big(z_i\exp(A_it) + \int_0^t \sigma_i\exp(A_i(t-s))dB_i(s)\Big)\\ =& \mathrm{Var}\Big(\int_0^t \sigma_i\exp(A_i(t-s))dB_i(s)\Big)\\ =& E\Big[\Big(\int_0^t \sigma_i\exp(A_i(t-s))dB_i(s)\Big)^2\Big]\\ =& \int_0^t E\Big[\sigma_i^2\exp(2A_i(t-s))\Big]ds = \frac{\sigma_i^2}{2A_i}[\exp(2A_it)-1]\end{aligned}$$

by Itô isometry (Corollary B.1.3) and the argument used for the expectation above. Letting $t\to\infty$, for $A_i<0$ we see that

$$E[Z_i(t)] = 0, \quad \mathrm{Var}(Z_i(t)) = -\frac{\sigma_i^2}{2A_i} > 0$$

By the fact that $Y = a + bX \sim \mathcal{N}(a, b^2)$ if $X \sim \mathcal{N}(0,1)$. Since the Itô integral in (2.14) is a normal random variable (by being the approximation of the sum of normally distributed random variables), we obtain that the stationary distribution (in the sense that $t\to\infty$) is $\sim \mathcal{N}(0, \frac{\sigma_i^2}{2A_i})$. For the covariance, note that

$$\mathrm{Cov}(Z_1(t), Z_2(t)) = \mathrm{Cov}\Big(\int_0^t \sigma_1\exp(A_1(t-s))dB_1(s),$$

$$\int_0^t \sigma_2 \exp(A_2(t-s))dB_2(s)\Big)$$
$$= E\Big[\rho\sigma_1\sigma_2 \int_0^t \exp((A_1+A_2)(t-s))ds\Big]$$

by the Itô isometry. Then equation (2.18) follows from straight forward integration. ■

A stochastic process $\{Z_i(t)\}_{t\geq 0}$ is said to be stationary if for all n, h the joint probability distribution of $(Z_i(t_1), \ldots, Z_i(t_n))$ and $(Z_i(t_1+h), \ldots, Z_i(t_n+h))$ does not change when shifted in time. In other words, the random vectors $(Z_i(t_1), \ldots, Z_i(t_n))$ and $(Z_i(t_1+h), \ldots, Z_i(t_n+h))$ have the same joint probability distribution [Ros14]. As time goes by, stationary processes tend to drift towards their long-term mean. The drift term of the dynamics of $Z_i(t)$ is dependent on the current state of the process. For a mean-reverting process, the dynamics are expressed by an equilibrium level and some percentage of drift, for instance, for some general OU-process $U(t)$,

$$dU(t) = A_i(\theta - U(t))dt + \sigma dB(t) \tag{2.19}$$

where the sign of A_i determines whether the stationary level attractes or repulses. For $Z_i(t)$, the equilibrium level is at 0, hence $\theta = 0$, and if the current value is less than the long-term mean, the process will be driven upwards towards the equilibrium (note that $A_i < 0$), and vice versa if the current state is less than the long term mean. In the long run, the long-term mean works as an equalizer for the process.

The Ornstein-Uhlenbeck processes are quite special in the sense that they are both stationary, Gaussian and Markov processes. In addition to the stationary property, the random vector of variables of $\{Z_i(t)\}_{t\geq 0}$ has a multivariate normal distribution. By the Markov property, future states of the process only depend on the current value. The current value of $Z_i(t)$ contains all information relevant for the future evolution of the process, see Definition B.2.5.

Remark 2.3.3.
Note that the Ornstein-Uhlenbeck process $Z_i(t)$ is a Markov process, but not a martingale by the fact that

$$\begin{aligned}
E[Z_i(t)|\mathcal{F}_s] =& z_i e^{A_i t} + E\Big[\int_0^t \sigma_i e^{A_i(t-u)} dB_i(u)|\mathcal{F}_s\Big] \\
=& z_i e^{A_i t} + E\Big[\int_0^s \sigma_i e^{A_i(t-u)} dB_i(u) + \int_s^t \sigma_i e^{A_i(t-u)} dB_i(u)|\mathcal{F}_s\Big] \\
=& z_i e^{A_i t} + \int_0^s \sigma_i e^{A_i(t-u)} dB_i(u) + E[\int_s^t \sigma_i e^{A_i(t-u)} dB_i(u)] \\
=& z_i e^{A_i t} + \int_0^s \sigma_i e^{A_i(t-u)} dB_i(u) \\
=& Z_i(s) e^{A_i(t-s)} \neq Z_i(s)
\end{aligned}$$

for $s < t$, where the third equality follows from the fact that $\int_0^s e^{A_i(t-u)} dB_i(u)$ is $\mathcal{F}_s$-measurable and the integral $\int_s^t e^{A_i(t-u)} dB_i(u)$ is independent of $\mathcal{F}_s$, Proposition B.4.2. The fourth equality follows from the fact that $g(u) = \exp A_i(t-u)$ is *in the class of Itô-integrable functions, and by property (iii) of Proposition B.1.4,* $E[\int_s^t g(u)dB_i(u)] = 0$.

2.4 Co-Integration

The concept of co-integration was formally introduced in econometrics by Granger and Engle in 1987 [EG87]. We say that two processes are co-integrated if there exists a linear combination of the processes which is stationary, even if the processes themselves are non-stationary. Formally, in [EG87], co-integration was introduced through the concept of differencing of time series, a commonly used technique for making non-stationary processes stationary by computing the difference of values from one period to the next. For instance, the first difference of a process $\{Y(t)\}_{t\geq 0}$ is denoted by $\{Y(t_i) - Y(t_{i-1})\}_{i\geq 1}$.

Definition 2.4.1 (Integrated process).
A stochastic process $\{X(t)\}_{t\geq 0}$ *is integrated if, after differencing d-times, the process is stationary. Then* $\{X(t)\}_{t\geq 0}$ *is said to be* $I(d)$.

Note that a stationary process, for instance the Ornstein-Uhlenbeck process, is $I(d = 0)$. Formally, we have that co-integration is defined as follows

Definition 2.4.2 (Co-integration).
Two stochastic processes, $\{X(t)\}_{t\geq 0}, \{Y(t)\}_{t\geq 0}$ *are said to be co-integrated if each of the processes is I(d), and there exists a* $I(d-b)$, $b > 0$, *linear combination of* $X(t)$ *and* $Y(t)$.

Note that for $d = b$, the linear combination of the processes is $I(0)$, hence stationary in itself and one do not need to find a stationary difference process of the linear combination.

Remark 2.4.3.
In [EG87], through a series of examples, Granger and Engle find that for instance short and long interest rates are co-integrated due to economic theory imposed by the government. Quite intuitively, they also show that income and consumption, and prices of the same commodity in different markets are co-integrated variables as well.

For the market model introduced in Section 2.1., the asset prices $S_1(t)$ and $S_2(t)$ are co-integrated by the fact that the difference on a logarithmic scale is stationary

$$\log(S_1(t)) - \log(S_2(t)) = Z_1(t) - Z_2(t) \tag{2.20}$$

when $Z_1(t), Z_2(t)$ are assumed to be stationary Ornstein-Uhlenbeck processes as defined in Section 2.2. The combined difference process $\{Z_1(t) - Z_2(t)\}_{t\geq 0}$ will never stray too far from it's long term combined mean, and the deviations from the equilibrium are stationary with finite variance

$$\mathrm{Var}(Z_1(t)) + \mathrm{Var}(Z_2(t)) - 2\mathrm{Cov}(Z_1(t), Z_2(t)) < \infty \tag{2.21}$$

2.5 Multi-Dimensional Asset Model

In real life, a portfolio often consists of more than two investments. The extended multi-dimensional asset price model of (2.1) in n assets is given by

$$S_i(t) = \exp\Big(\sum_{j=1}^{m} a_{ij} X_j(t) + Z_i(t)\Big), \quad i = 1, \ldots, n \tag{2.22}$$

where $m \leq n$, $X_j(t)$ are m non-stationary processes with different drifts and diffusions, and possibly correlated Brownian motions. $Y_i(t)$ are the n stationary CARMA(p_i, q_i) processes driven by correlated p_i-dimensional Brownian motions. We say that $S_1(t), S_2(t), \cdots, S_n(t)$ are co-integrated if there exists a vector $\mathbf{c} \in \mathbb{R}^n$ such that a linear combination of $S_i(t)$ is stationary or at least $I(d-b)$ for some $b \leq d$. On a logarithmic scale, for $\sum_{i=1}^{n} c_i \log(S_i(t))$, $\mathbf{c}$ must satisfy

$$c_1 \sum_{j=1}^{m} a_{1j} X_j(t) + \cdots + c_n \sum_{j=1}^{m} a_{nj} X_j(t) = 0 \tag{2.23}$$

or equivalently

$$X_1(t)\Big(a_{11}c_1 + a_{21}c_2 \cdots + a_{n1}c_n\Big) + \\ \ldots + X_m(t)\Big(a_{1m}c_1 + a_{2m}c_2 \cdots + a_{nm}c_n\Big) = 0 \tag{2.24}$$

Since $X_j(t)$ is non-zero for all j, we get the following system of equations

$$\begin{aligned} a_{11}c_1 + \cdots + a_{n1}c_n &= 0 \\ a_{12}c_1 + \cdots + a_{n2}c_n &= 0 \\ &\vdots \\ a_{1n}c_1 + \cdots + a_{nm}c_n &= 0 \end{aligned}$$

From ([BK15]) we give an example of a 3-dimensional co-integrated asset price model.

Example 2.5.1 (Crush Spreads).
Soybeans can be crushed (processed) into soybean meals and soybean oil. A trading strategy in which the trader places a long position in soybean futures and a short position in soybean meal and soybean oil futures is called a crush spread [Wik19]. The crush spread is the difference between the price of quantity of soybeans, soybean meals and soybean oils. [LL06]. We look at an example with three assets driven by two non-stationary processes, then $n = 3$ *and* $m = 2$ *and* $S_1(t)$, $S_2(t)$, $S_3(t)$ *are co-integrated if the linear combination*

$$c_1 \log(S_1(t)) + c_2 \log(S_2(t)) + c_3 \log(S_3(t)) \tag{2.25}$$

is stationary, i.e. if there exist a non-zero vector of constants, $\mathbf{c} = (c_1, c_2, c_3)$, *such that*

$$\begin{aligned} X_1(t)\Big(a_{11}c_1 + a_{21}c_2 + a_{31}c_3\Big) &= 0 \\ X_2(t)\Big(a_{12}c_1 + a_{22}c_2 + a_{32}c_3\Big) &= 0 \end{aligned}$$

For $n > n_c = 2$, *we might investigate co-integration between two of the assets. For instance we have that* $S_1(t)$ *and* $S_2(t)$ *are co-integrated if for* $\mathbf{c} = (c_1, c_2)$

$$a_{11}c_1 + a_{21}c_2 = 0$$

$$a_{12}c_1 + a_{22}c_2 = 0$$

We see that $\mathbf{c}$ *exists as long as* $a_{11}a_{22} = a_{21}a_{12}$. *In addition, one might work out conditions to ensure additional co-integration between* $S_1(t)$ *and* $S_3(t)$ *or* $S_2(t)$ *and* $S_3(t)$. *This gives us flexibility to modify situations for* $n > 2$ *where some assets are co-integrated and some are not.*

Other possibilities are to include several stationary processes to add additional flexibility in the stationary effects, or to include non-Gaussian features, through for instance Lévy jump-processes or processes with stochastic volatility like the Heston model.

2.6 Properties of the Asset Prices

From now on and throughout the rest of the chapters, we assume we are in the 2-dimensional asset price model where $S_i(t), i = 1, 2$ is modelled by a common non-stationary trend process, $X(t) \in \mathbb{R}$, and the stationary processes are represented by the 1-dimensional Ornstein-Uhlenbeck processes $Z_i(t), i = 1, 2$. This section states some of the properties of the process $S_i(t)$ for $i = 1, 2$.

Proposition 2.6.1 (The dynamics of S_i).
The dynamics of the stock price process $S_i(t)$ *defined in (2.1) are given by*

(i) *For* $\mathrm{Corr}(\widetilde{B}_t, B_i(t)) = 0$, *then*

$$dS_i(t) = S_i(t)\Big(\mu + A_i Z_i(t) + \frac{1}{2}(\sigma^2 + \sigma_i^2)\Big)dt + S_i(t)\Big(\sigma d\tilde{B}_t + \sigma_i dB_i(t)\Big) \tag{2.26}$$

(ii) *For* $\mathrm{Corr}(\widetilde{B}_t, B_i(t)) = \rho_i$, *then*

$$dS_i(t) = S_i(t)\Big(\mu + A_i Z_i(t) + \frac{1}{2}(\sigma^2 + \sigma_i^2 + 2\rho_i\sigma\sigma_i)\Big)dt + S_i(t)\Big(\sigma d\tilde{B}_t + \sigma_i dB_i(t)\Big) \tag{2.27}$$

Proof of proposition (2.6.1). Similarly to the proof of (2.3.1), we let $f(x, z) = \exp(x + z)$, then $S_i(t) = f(X(t), Z_i(t)$ and by Itô's lemma

$$\begin{aligned} dS_i(t) =& \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial x}dX(t) + \frac{\partial f}{\partial z}dZ_i(t) + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(dX(t))^2 + \frac{1}{2}\frac{\partial^2 f}{\partial z^2}(dZ_i(t))^2 \\ &+ \frac{\partial^2 f}{\partial x \partial z}(dX(t)dZ_i(t)) \\ =& S_i(t)\Big(\mu dt + \sigma d\tilde{B}_t\Big) + \frac{1}{2}S_i(t)\sigma^2 dt + S_i(t)\Big(A_i Z_i(t)dt \\ &+ \sigma_i dB_i(t)\Big) + \frac{1}{2}S_i(t)\sigma_i^2 dt + S_i(t)dX(t)dZ_i(t) \\ =& S_i(t)\Big(\mu + A_i Z_i(t) + \frac{1}{2}(\sigma^2 + \sigma_i^2 + 2\rho_i\sigma\sigma_i\Big)dt + S_i(t)\Big(\sigma d\tilde{B}_t \\ &+ \sigma_i dB_i(t)\Big) \end{aligned}$$

where the last equality follows from the fact that $dX_t dZ_i(t) = \rho_i\sigma\sigma_i dt$. ■

In later sections we introduce the notational convention $\Sigma_i = \sigma^2+\sigma_i^2+2\rho_i\sigma\sigma_i$ to shorten the notation for some of the calculations.

Proposition 2.6.2 (Distribution, expected value and variance of $S_i(t)$).
For any t, $S_i(t)$ is a log-normal random variable with mean and variance given by

$$E[S_i(t)] = \exp\Big((\mu+\frac{1}{2}\sigma^2)t+z_ie^{A_it}+\frac{1}{2}\Big(\frac{\sigma_i^2}{2A_i}(e^{2A_it}-1)+\frac{1}{A_i}\rho_i\sigma\sigma_i(e^{A_it}-1)\Big)\Big) \tag{2.28}$$

$$\begin{aligned}\mathrm{Var}(S_i(t)) = &\Big[\exp\Big(\sigma^2t+\frac{\sigma_i^2}{2A_i}(e^{2A_it}-1)+\frac{1}{A_i}\rho_i\sigma\sigma_i(e^{A_it}-1)\Big)-1\Big]\\ &\times\exp\Big((2\mu+\sigma^2)t+2z_ie^{A_it}+\frac{\sigma_i^2}{2A_i}(e^{2A_it}-1)+\frac{1}{A_1}\rho_i\sigma\sigma_i(e^{A_it}-1)\Big)\end{aligned} \tag{2.29}$$

Proof of proposition (2.6.2). Note that for every t, $S_i(t)$ is the exponential of two normally distributed random variables, $X(t) \sim \mathcal{N}(\mu t, \sigma^2 t)$ and $Z_i(t) \sim \mathcal{N}(z_ie^{A_it}, \frac{\sigma_i^2}{2A_i}(e^{2A_it}-1))$, hence $S_i(t)$ is a log-normal random variable with mean and variance given by

$$E[\log(S_i(t))] = E[X(t)] + E[Z_i(t)] = \mu t + z_ie^{A_it}$$

$$\begin{aligned}\mathrm{Var}(\log(S_i(t))) =&\mathrm{Var}(X(t)) + \mathrm{Var}(Z_i(t)) + \mathrm{Cov}(X(t), Z_i(t))\\ =&\sigma^2t+\frac{\sigma_i^2}{2A_i}(e^{2A_it}-1)+\frac{1}{A_i}\rho_i\sigma\sigma_i(e^{A_it}-1)\end{aligned}$$

for

$$\begin{aligned}\mathrm{Cov}(X(t), Z_i(t)) =&\mathrm{Cov}\Big(\mu t+\int_0^t \sigma d\widetilde{B}(s), z_ie^{A_it}+\int_0^t \sigma_ie^{A_i(t-s)}dB_i(s)\Big)\\ =&E\Big[\Big(\int_0^t \sigma d\widetilde{B}(s)\Big)\Big(\int_0^t \sigma_ie^{A_i(t-s)}dB_i(s)\Big)\Big]\\ &-E\Big[\int_0^t \sigma d\widetilde{B}(s)\Big]E\Big[\int_0^t \sigma_ie^{A_i(t-s)}dB_i(s)\Big]\\ =&\int_0^t \rho_i\sigma\sigma_ie^{A_i(t-s)}ds\end{aligned}$$

where the last equation follows from the Itô-isometry and the fact that both integrals are well-defined Itô integrals with zero expectation. By the formulas for the expected value and variance of a log-normal random variable (Definition B.4.1), we obtain (2.28) and (2.29) ■

Note that $S_i(t)$ is non-stationary even for $A_i < 0$. As $t \to \infty$

$$\lim_{t\to\infty} E[S_i(t)] \to \exp\Big((\mu+\frac{1}{2}\sigma^2)t\Big) \to \infty \tag{2.30}$$

$$\lim_{t\to\infty} \mathrm{Var}(S_i(t)) \to \exp\Big((2\mu + \sigma^2)t\Big) \to \infty \tag{2.31}$$

We conclude this chapter by remarking that the non-stationary trend process $X(t)$ dominates the stationary part of $S_i(t), i = 1, 2$ as $t \to \infty$, yielding a non-stationary limiting distribution.

Chapter 3

Stochastic Control Problems

3.1 Introduction

Stochastic optimal control is a sub-field of control theory in mathematics, widely used in several areas, including finance, insurance and industry, with applications to portfolio optimization, reinsurance and automatic control, among others. Stochastic control theory deals with dynamic systems exposed to uncertainty, either in the observations of the dynamics or in the underlying processes driving the system. We present a case where the dynamics of the state process are described by a stochastic differential equation (SDE), where the basic source of uncertainty comes from white noise (Brownian motions). In this thesis we consider such a problem where the system is required to be Markovian, allowing us to exploit the dynamic programming approach which deduces the optimal control problem to a non-linear partial differential equation (PDE) with boundary conditions.

This chapter gives a short introduction to the theory of continuous stochastic control in view of the dynamic programming approach. The general set-up of a stochastic control problem will be introduced, and the connection between the dynamic programming principle and the Hamilton-Jacobi-Bellman equation (HJB for short) will be presented. In Section 5, a familiar example from portfolio optimization, often referred to as Merton's problem, will be given. Chapter 4 and the motivation behind it is built on Merton's portfolio optimization problem.

The theory presented in this chapter is mainly based on [YZ99], [Mor10] and [FS06].

3.2 A General Stochastic Control Problem

The purpose of stochastic control problems is to optimize a dynamic system evolving over time, according to certain criteria. Among all feasible controls, the decision maker must choose the optimal one in order to maximize or minimize a pre-given criteria. Denote the controlled state process at time t by $X(t)$, and suppose the dynamics of $X(t)$ are given by

$$dX(t) = b(t, X(t), u(t))dt + \sigma(t, X(t), u(t))dB(t) \tag{3.1}$$

$$X(0) = x_0 \tag{3.2}$$

where $B(t)$ a Brownian motion in $\mathbb{R}^m$, $X(t) \in \mathbb{R}^n$ and $b(t,x,u) : [0,T] \times \mathbb{R}^n \times U \to \mathbb{R}^n$, $\sigma(t,x,u) : [0,T] \times \mathbb{R}^n \times U \to \mathbb{R}^{n \times m}$ are measurable functions for $U \subseteq \mathbb{R}$. The control process $\{u(t)\}_{0 \leq t \leq T}$ represents the decisions of the decision maker at each time instant t. Since the system under study is continuously changing, the decisions controlling it must be adapted to the most recent information given. We assume the controller is familiar with all information obtained up till time t, mathematically speaking, $u(t)$ is $\mathcal{F}_t$-adapted, where $\mathcal{F}_t$ is the natural filtration generated by $B(t)$. Due to the uncertainty of the stochastic state process, the controller is not able to forecast the future of the state process and can therefore only base his decision on information contained in $\mathcal{F}_t$. Problems where the decision maker is familiar with either less information through partially observable systems, or more through insider information, are discussed in for instance [Øks05]. In the case where the information is only partially observable, the controls are adapted w.r.t. a smaller σ-algebra, $\mathcal{G}_t \subseteq \mathcal{F}_t$. Contrary, in the insider information case, the controls $u(t)$ are $\mathcal{H}_t$-adpated, for $\mathcal{F}_t \subseteq \mathcal{H}_t$.

The aim of a controller is to optimize, whether it be a minimization or maximization problem, a certain criteria over a given set of admissible controls. There are two main formulations of a stochastic control problem; the strong and weak formulation. Under strong formulation, one seeks an optimal solution on a a priori fixed probability space on which a filtration is given and a Brownian motion defined. If we allow the probability space to vary, we seek weak solutions $u(t)$ defined as 5-tuples $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathcal{P}}, \widetilde{B}(t), u(t))$ with corresponding filtration $\{\mathcal{F}_t\}_{t \geq 0}$.

The Strong Formulation

On a fixed complete filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathcal{P})$ we define a Brownian motion $B(t)$, and let

$$\mathcal{U}[0,T] = \{[0,T] \times \Omega \to U \subseteq \mathbb{R}\}$$

denote the set of all measurable and $\mathcal{F}_t$-adapted processes $u(t)$, where U is a given complete and separable space. For each $(t,x) \in D$, where D is the domain of the control problem, and $T < \infty$, we define the performance function, or the cost functional, by

$$J(t,x;u) = E^x\Big[\int_t^T f(s,X(s),u(s))ds + g(X(T))\Big] \tag{3.3}$$

where $f : [0,T] \times \mathbb{R}^n \times U \to \mathbb{R}$, $g : [0,T] \times \mathbb{R}^n \to \mathbb{R}$ are two continuous functions, often referred to as the running cost or profit rate function, and the terminal cost or bequest function, depending on the setting of the control problem and whether the aim is to minimize cost or maximize profit. From here on, J is referred to as the performance function, and f and g the profit rate and bequest functions, respectively, in order to coincide with Chapter 4. The performance function in (3.3), where $T < \infty$ is fixed in advance, is defined for a fixed duration or finite horizon problem. In the case that one lets $T \to \infty$, i.e. one looks at an infinite time horizon, see for instance [CCF17], a discounting factor

$$e^{\int_t^T -c(s)ds} \tag{3.4}$$

for some continuous time-dependent function $c : [0, \infty) \to \mathbb{R}$ is included to draw f and g back to 0 as $t \to \infty$. We drop the discounting factor and define the control problem on a finite interval.

The aim of the controller is to optimize the performance function J over a given set of admissible controls, $u(t) \in \mathcal{A}$, where $\mathcal{A}$ is defined as follows

Definition 3.2.1 (Admissible controls).
Given a complete filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathcal{P})$ where $B(t)$ is a given m-dimensional Brownian motion, a control $u(t)$ is said to be admissible if

(i) $u(t) \in \mathcal{U}[0, T]$

(ii) *$u(t)$ is such that a unique solution $X(t)$ to (3.1) exists, i.e. $\{X(t)\}_{t\geq 0}$ is an Itô-diffusion*

(iii) *some prescribed state constraint*

(iv) *$u(t) \in L^2([0, T] \times \mathbb{R}^n)$, i.e.*

$$E[\int_0^T u^2(s)ds] < \infty \tag{3.5}$$

(v) *$u(t)$ is such that $f \in L^1([0, T] \times \mathbb{R})$, i.e.*

$$E\Big[\int_0^T |f(s, X(s), u(s))|ds\Big] < \infty \tag{3.6}$$

and $g \in L^1(\Omega, \mathbb{R})$

$$E\Big[|g(T, X(T), u(T))|\Big] < \infty \tag{3.7}$$

Remark 3.2.2.
Assumption (iii) is some prescribed state constraint depending on the specific control problem. For instance we might require that $u(t) \in [u_0, u_1]$, where u_0 and u_1 are values in $\mathbb{R}$. Assumptions (iv) and (v) ensure that the performance function is well-defined.

Under strong formulation, the stochastic control problem may be formulated as follows

Problem 3.2.3 (Stochastic control problem under strong formulation).
Given any $(t, x) \in D$, we want to find an optimal control $u^*(t) = u^*(t, X(t)) \in \mathcal{A}$ such that

$$\Phi(t, x) = \sup_{u\in\mathcal{A}} J(t, x; u(t)) = J(t, x; u^*(t)) \tag{3.8}$$

where $\Phi(t, x)$ denotes the optimal value function of the stochastic control problem. Equivalently,

$$\Phi(t, x) = \inf_{u\in\mathcal{A}} J(t, x; u(t)) = J(t, x; u^*(t)) \tag{3.9}$$

If there exists a unique and finite solution to problem (3.2.3) under strong formulation, the problem is said to be finite and uniquely solvable. $u^*(t)$ is said to be the optimal control and the corresponding state process $X^*(t) = X(t, u^*(t))$ is said to be optimal.

The Weak Formulation

We focus on the strong formulation, hence only a short description of the weak formulation is given. The weak formulation is mainly a mathematical tool for finding solutions to problems initially formulated under the strong formulation. Under weak formulation we allow $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathcal{P}})$, hence $\{\widetilde{\mathcal{F}}_t\}_{t\geq 0}$ as well, and $\widetilde{B}(t)$ to vary. Given functions $b(\cdot, \cdot, u(\cdot))$ and $\sigma(\cdot, \cdot, u(\cdot))$, and the set $U \subseteq \mathbb{R}$, we consider the filtered probability space and the Brownian motion as parts of the solution. As long as the solution to (3.1) has the same probability distribution for each t under $(\Omega, \mathcal{F}, \mathcal{P})$ and $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{\mathcal{P}})$, the performance function (and expectation) depending on $X(t, u(t))$ w.r.t. respectively $\mathcal{P}$ and $\widetilde{\mathcal{P}}$ will coincide. The weak formulation opens for more flexibility in the solutions to the stochastic control problem.

3.3 The Dynamic Programming Principle

Under strong formulation, given the controlled system (3.1), we make the standing assumptions that

$$|b(t,x,u) - b(t,y,u)| + |\sigma(t,x,u) - \sigma(t,y,u)| \leq C|x-y| \qquad t \in [0,T],\ x,y \in \mathbb{R}^n,\ u \in U \tag{3.10}$$

$$|b(t,x,u)| + |\sigma(t,x,u)| \leq C(1+|x|), \quad t \in [0,T],\ x \in \mathbb{R}^n \tag{3.11}$$

i.e. we assume the functions $b(\cdot, \cdot, u(\cdot))$, $\sigma(\cdot, \cdot, u(\cdot))$ are Lipschitz continuous, and that the controlled process $X(t, u(t))$ does not explode for any $t < \infty$. Conditions (3.10) and (3.11) ensures the existence and uniqueness of a strong solution to $dX(t)$. Furthermore, for any $u(t) \in \mathcal{A}$ the profit rate function f and the bequest function g satisfies assumption (v) in definition (4.2.1). The controlled solution $X(t)$ will be a Markovian Itô diffusion, allowing us to apply the dynamic programming approach to solve the stochastic control problem. Note that since the system itself is Markovian, the control-processes will be Markov decision processes.

In this section we present how the stochastic control problem can be deduced to a non-linear boundary value problem through the value function Φ.

The dynamic programming approach, originally developed by R. Bellman [Bel18] in the 1950s, is a method for solving complex problems through a series of easier sub-problems. In stochastic control, Bellman's principle connects the original problem to a non-linear partial differential equation of the underlying dynamics of the system. We consider the simplest case, where the controlled process $X(t)$ is a Markov process, and where we assume there exists a "smooth" value function Φ solving the optimal control problem. By Bellmans's principle of optimality we obtain the Hamilton-Jacobi-Bellman equation.

Bellman's principle of optimality is defined as follow

$$\Phi(t,x) = \sup_{u \in U} E^x \Big[\int_t^{t+h} f(s, X(s), u(s))ds + \Phi(t+h, X(t+h)) \Big] \tag{3.12}$$

for $t \leq t+h \leq T$. Heuristically, Bellman's principle states that; under the optimal control $u^* \in \mathcal{A}$, the value function for $(t,x) \in D$ must be equal to the profit rate function under optimal u on a smaller time interval $[t, t+h]$ and the value of preceding optimally from the point $t + h$. If an optimal control exists, the principle ensures the property that no matter which initial state the process starts in, if preceding in accordance to the optimal control, optimality will be reached. Hence the problem is deduced to a family of sub-problems with different initial values, and when combining the sub-problems, the more complex and general problem is obtained yielding the same solution. We restate the heuristic approach to show optimality given in [FS06]. A more formal proof for the specific control problem of this thesis is given in Chapter 4. Following Bellman's principle, for $t \leq s \leq t+h$ and a constant control $u(t) = u \in U$, then

$$\Phi(t,x) \leq E^x\Big[\int_t^{t+h} f(s,X(s,u))ds\Big] + E^x[\Phi(t+h,X(t+h))] \tag{3.13}$$

by definition of Φ. We subtract by $\Phi(t,x)$ on both sides

$$0 \leq E^x\Big[\int_t^{t+h} f(s,X(s,u))ds\Big] + E^x[\Phi(t+h,X(t+h))] - \Phi(t,x) \tag{3.14}$$

Dividing by h and letting $h \to 0^+$, we obtain

$$\begin{aligned} 0 &\leq \lim_{h\to 0}\frac{1}{h}\Big(E^x\Big[\int_t^{t+h} f(s,X(s,u))ds\Big] + E^x[\Phi(t+h,X(t+h))] - \Phi(t,x)\Big) \\ 0 &\leq E^x[f(t,x,u)] + (L^u\Phi)(t,x) \end{aligned}$$

where $(L^u\Phi)(t,x) \in \mathcal{D}_A(x)$ is the infinitesimal generator, Definition B.2.2, of Φ for (t,x). We obtain that

$$0 \leq f(t,x,u) + (L^u\Phi)(t,x) \tag{3.15}$$

for any $\Phi \in \mathcal{D}_A(x)$. If however $u(t)$ yields the optimal value u^*, the inequality in (3.13) becomes an equality and

$$0 = f(t,x,u^*) + (L^{u^*}\Phi)(t,x) \tag{3.16}$$

If such an optimal $u^*(t)$ exists, then it must be the control with value u for which (3.16) attains it's maximum, i.e.

$$\sup_{u\in U} f(t,x,u) + (L^u\Phi)(t,x) = 0 \tag{3.17}$$

with boundary value

$$\Phi(T,x) = g(T,X(T)) \tag{3.18}$$

Equation (3.17) is a non-linear second order (in the state variable, first order w.r.t. time) partial differential equation, referred to as the Hamilton-Jacobi-Bellman equation. A solution Φ to the non-linear boundary value problem (3.17)-(3.18) gives a solution to the initial stochastic optimization problem (3.2.3). For this solution to exists, the generator operator must be well-defined, i.e. the controlled process must be a Markov process, specifically an Itô diffusion. Furthermore, we need to impose conditions on the value function for which (3.17)-(3.18) has a solution. We require that the value function Φ is smooth, in the sense that it is continuous and bounded on the domain D, twice continuously differentiable w.r.t. the state variable x and continuously differentiable w.r.t. t.

These are restrictive assumptions and more often than not, the value function is not smooth and there exists no classical solution to the boundary value problem. In such cases, viscosity solutions must be introduced. Such solutions yield a far more complex theory, see for instance [FS06] where Nisio's [Nis15] construction of a non-linear semigroup is used to prove relevance and validation of viscosity solutions.

3.4 A Verification Theorem

If a smooth function $\Phi \in \mathcal{D}_A$ exists, i.e. Φ is a function for which the infinitesimal generator is defined, such that Φ satisfied the HJB-equation (3.17) with corresponding boundary value condition (3.18), then a verification theorem conjoins the initial control problem and the HJB-equation. The verification theorem guarantees that the value function is indeed the supremum of the performance function over the set of admissible controls.

Theorem 3.4.1 (A verification theorem).
Assume the value function $\Phi \in \mathcal{D}$ is a smooth solution to (3.17)-(3.18) for all $(t, x) \in D$. Then

(i) $\Phi(t, x) \geq J(t, x; u)$ *for any* $u(t) \in \mathcal{A}$

(ii) *if there exists a control* $u^*(t) = u^*$, $u \in U$ *such that*

$$f(t, x, u^*) + (L^{u^*}\Phi)(t, x) = 0 \tag{3.19}$$

then $u^*(t) = u^*$ *must be the optimal control and*

$$\Phi(t, x) = J(t, x; u^*) \tag{3.20}$$

A more deductive theorem and the proof of it will be given in Chapter 4.

3.5 Portfolio Optimization

A special application of stochastic optimal control theory is that of portfolio optimization in mathematical finance. It has long been known that a good way to spread risk is through portfolio diversification. By investing in different market segments, investors reduce the exposure to risk of one particular asset. Portfolio optimization problems may be formulated as stochastic optimal control problems where an investor seeks to optimize expected utility of wealth through investing in an optimal portfolio. A utility function represents the investor's aversion towards risk.

Power Utility

In the following section and in Chapters 4 and 5 we consider a risk-averse investor, meaning that the investor is seeking to reduce risk associated to the investments made. The utility function is an expression of the investor's preferences in the financial market, and is defined as follows.

Definition 3.5.1 (Utility function).
An utility function $U(w) : [0, \infty) \to [0, \infty)$ is a continuously differentiable, strictly increasing and strictly concave function, satisfying the following properties

(i)
$$\lim_{w \to 0} U'(w) = \infty \tag{3.21}$$

(ii)
$$\lim_{w \to \infty} U'(w) = 0 \tag{3.22}$$

The function U represents the utility of the investor given the wealth w. All of the assumptions in the definition above have economical interpretations. Strictly increasing implies that additional wealth, in the form of for instance capital, increases utility, and more wealth is preferred to less. Strictly concave means that the marginal utility decreases, i.e. the utility of one additional krone or unit of wealth, decreases the more wealth the investor has. Conditions (i) and (ii) are know as the Inada conditions [Gal], and ensure that the investor does not end up in a situation where the wealth increases infinitely or where wealth is close to zero.

In a portfolio optimization problem the choice of utility function expresses the level of risk-aversion of the investor. One such utility function that satisfies the assumptions in Definition 3.5.1, is a special case of the HARA (hyperbolic absolute risk aversion)-utility functions, namely the power function

$$U(w) = w^{\gamma}, \quad 0 < \gamma < 1, \quad U(0) = 0 \tag{3.23}$$

The absolute risk aversion coefficient of an utility function is defined to be the fraction

$$A(w) = -\frac{U''(w)}{U'(w)}$$

If the wealth of an investor increases, the investor will increase the fraction of wealth invested in the particular asset yielding higher utility if the absolute risk aversion is decreasing.

Merton's Problem

Robert C. Merton was the first to introduce the dynamic programming approach to portfolio optimization problems. In [Mer69], [Mer71], Merton presents a continuous-time consumption-investment problem for a risk-averse investor, where income is generated by returns on assets, driven by Brownian motions. Specifically an explicit solution is found in the two-asset model, where one asset is a risky investment, and the other a risk-free. Following [Øks13], we present an example of the simplest case; we look at an investor aiming at maximizing expected utility of wealth at some future time T, given the two investment opportunities. The investor must choose how to optimally allocate wealth in order to maximize expected utility in the future. We show that the optimal allocation is indeed to invest a constant fraction in the risky asset. The following example serves as a motivation for Chapter 4.

Example 3.5.2 (Merton's portfolio optimization problem).
Given initial wealth w_0 at time t_0, we denote by $W(t)$ the wealth of an investor at each time point t. The investor aims at maximizing expected utility of wealth at some future time T by investing in two assets; one risky and one risk-free asset, denoted as $S(t)$ and $X(t)$ respectively. The risky asset is assumed to satisfy the following stochastic differential equation

$$dS(t) = \mu S(t)dt + \sigma S(t)dB(t) \tag{3.24}$$

where $B(t)$ is a Brownian motion. $X(t)$ is defined as the solution to the ordinary differential equation

$$dX(t) = rX(t)dt \tag{3.25}$$

where r is a constant interest-rate. We let $u(t)$ denote the fraction of wealth invested in the risky asset at each time point t, thereby investing the fraction $1 - u(t)$ in the risk-free asset. Assuming $0 \leq u(t) \leq 1$, and that the wealth is given under a self-financing portfolio, the dynamics of the $W(t)$ are given by

$$dW(t) = W(t)(\mu u(t) + r(1 - u(t)))dt + \sigma u(t)W(t)dB(t) \tag{3.26}$$

Given the power utility function in (3.23), the aim of the investor is to maximize expected utility of wealth at time $T < \infty$. The stochastic control problem is formulated as follows

$$\Phi(t, w) = \sup_{u \in \mathcal{A}} J(t, x; u) = E[U(W(\tau_D))] \tag{3.27}$$

where the supremum is taken over a set of admissible controls, assumed to be measurable and $\mathcal{F}_t$-adapted. τ_D is the first exit time of the domain D, denote by

$$D = \{(s, w) : 0 \leq s \leq T, w > 0\} \tag{3.28}$$

The goal is to find the optimal value $u = u(t)$ such that

$$\sup_{u \in U} \{(L^u \Phi)(t, w)\} = 0 \quad \textit{for } (t, w) \in D \tag{3.29}$$

$$\Phi(t, w) = U(w) \textit{ for } t = T, \quad \Phi(t, 0) = 0 \textit{ for } t < T \tag{3.30}$$

where

$$(L^u \Phi)(t, w) = \frac{\partial \Phi}{\partial t} + w(r + (\mu - r)u)\frac{\partial \Phi}{\partial w} + \frac{1}{2}\sigma^2 u^2 w^2 \frac{\partial^2 \Phi}{\partial w^2} \tag{3.31}$$

Equation (3.29) is the HJB-equation associated to the stochastic control problem. Following the dynamic programming approach presented above, we know that the optimal value u of $u(t)$, if an optimal control exists, ensures that equation (3.29) attains it's maximum for u. We find u by solving the first order condition of a maxima,

$$\frac{\partial (L^u \Phi)(t, w)}{\partial u} = 0, \quad \textit{for} \quad \frac{\partial \Phi}{\partial w} > 0, \quad \frac{\partial^2 \Phi}{\partial w^2} < 0 \tag{3.32}$$

By substituting the solution for u into the HJB-equation, we obtain a non-linear boundary value problem for Φ

$$\Phi_t + rw\Phi_w - \frac{(\mu - r)^2 \Phi_w^2)}{2\sigma^2 \Phi_{ww}} = 0, \quad \textit{for } t < T, w > 0 \tag{3.33}$$

$$\Phi(t,w) = U(w) \quad \textit{for } t = T,\ w = 0 \tag{3.34}$$

The next step is to guess on a solution on the form $\phi = f(t)w^\gamma$. *By substituting the guess for* ϕ *into problem (3.33), the partial differential equation reduces to a ordinary differential equation which can be solved for* $f(t)$. *The solution to Merton's problem is then given by*

$$\Phi(t,w) = \exp(\lambda(T-t_0))w^\gamma, \quad \lambda = \gamma r + \frac{(\mu - r)^2\gamma}{2\sigma^2(1-\gamma)} \tag{3.35}$$

where the optimal control equals

$$u^*(t,w) = \frac{\mu - r}{\sigma^2(1-\gamma)} \tag{3.36}$$

Remark 3.5.3.
Problems where one allows for negative wealth, are for instance studied in [JS19], where one looks at a combined optimal investment and consumption problem, where an investor is allowed to obtain negative wealth by borrowing against future income.

Note that the optimal investment strategy, often referred to as the Merton ratio, of the risk-averse investor is to allocate a constant fraction of wealth in the risky asset. If the drift in the price of the risky asset is higher that the return on the interest rate, the investor will allocate at least a small fraction of wealth in the risky asset (note that $1-\gamma > 0$). The risk-aversion of the investor is clear; if the volatility of the risky asset is large, the fraction of wealth invested in the risky asset decreases, and the investor allocates more wealth to the safe investment.

One of the advantages with the dynamic programming approach to portfolio optimization is the continuity of the stochastic controls; the investor is able to base his decision on current information about the wealth process and underlying drivers, and (in theory) continuously change the fraction of wealth invested in each asset (if disregarding transaction costs). Before Merton's studies, most portfolio optimization problems, for instance the mean-variance approach in Markowitz portfolio theory [Mar52], were studied in discrete time. By relating the control problem to the dynamics of the state process, dynamic programming allows for continuous time controls. The anticlimax of Merton's simple problem is however the fact that the optimal control $u^*(t)$ is a constant fraction, only depending on the volatility and drift of the asset price. The investor chooses the initial allocation of wealth at time t and then becomes a passive agent. The Merton optimal control neither depends on time, nor on the current value of wealth. Chapter 4 is built on Merton's problem and in Chapter 5 we try to attain a constant-fraction investment strategy trough mathematically unconventional, at least, methods, motivated by Merton's solution.

Chapter 4

Optimal Portfolio Selection Problem

4.1 Introduction

In a complete market with co-integrated asset prices, we consider a risk averse investor who has the limited choice of investing in two risky and one risk-free asset. The risky assets are typically stocks, while the safe investment might be a bond or a bank-account. The aim of the investor is to maximize expected utility of wealth at some future time point. At all times the investor can choose how much of the wealth he or she wants to invest in each of the assets with the aim of allocating wealth in an utility maximizing manner, i.e. finding the optimal portfolio investment strategy.

4.2 Formulation of the Control Problem

Let $(\Omega, \mathcal{F}, \mathcal{P})$ be the probability space defined in Section 2.1 on which the three correlated Brownian motions $\widetilde{B}(t)$, $B_1(t)$ and $B_2(t)$ are defined.

We let $W(t)$ denote the wealth of an investor at time t. Recall from Chapter 2 that the risky assets S_1, S_2 are assumed to follow the dynamics

$$dS_i(t) = S_i(t)\Big(\mu + A_i Z_i(t) + \frac{1}{2}(\sigma^2 + \sigma_i^2 + 2\rho_i\sigma\sigma_i)\Big)dt \\ + S_i(t)\Big(\sigma d\widetilde{B}(t) + \sigma_i dB_i(t)\Big) \quad (4.1)$$

where $S_1(t)$ and $S_2(t)$ are co-integrated, as presented in Section 2.4. Further we assume that the price of the risk-free asset satisfies the ordinary differential equation

$$dX_0(t) = rX_0(t)dt \quad (4.2)$$

where r is a continuous compounding rate of return, $r < E[S_i]$ for $i = 1, 2$. From a financial point of view, the latter is a natural assumption, but not necessarily always true in the real market. If the returns on interest rates were higher that the expected value of the price of a risky asset (or the expected value of the returns), it would obviously be wiser to invest all capital in the risk-free asset; yielding higher returns and carrying no risk. However note that, even though the risk-free assets might guarantee a fixed rate of return, there is risk associated with inflation which might decrease the real value of the expected returns by means of for instance purchasing power.

We assume there are no transaction costs in the market and that the investor invests in a self-financing portfolio, meaning that, given initial wealth w_0, no additional wealth is neither added nor withdrawn. All changes in the wealth after time t_0 are due to changes in the price of the underlying assets. As the asset prices fluctuate, the investor has to reallocate the wealth in order to reach the goal of maximizing expected utility of wealth at time T.

We let $u_i(t)$ denote the fraction of wealth invested in risky asset i for $i = 1, 2$ at each time instant t. The remaining fraction of wealth, $1 - u_1(t) - u_2(t)$, is invested in the risk-free asset. Let

$$\mathcal{U}[0,T] = \{u(t) : [0,T] \times \mathbb{R}^3 \to U\} \tag{4.3}$$

denote the set of measurable and $\mathcal{F}_t$-adapted processes representing the decisions of the investor at each time $0 \leq t \leq T$, where $U \in \mathbb{R}$. Then $\mathcal{A} \in \mathcal{U}$ is defined as the set of admissible controls.

Definition 4.2.1 (Admissible controls).
A stochastic control process

$$u(t) = \{u(t) : 0 \leq t \leq T\} : [0,T] \times \mathbb{R}^3 \to U \tag{4.4}$$

is said to be admissible and $u \in \mathcal{A}$ *if*

(i) $u(t)$ *in* $\mathcal{U}[0,T]$

(ii) $u(t)$ *is such that a unique solution* $W(t) \geq 0$ *for all* $t \geq 0$ *to the wealth dynamics* $dW(t)$ *exists*

(iii) $u(t) \in L^2([0,T] \times \mathbb{R}^3)$, *i.e.*

$$E[\int_0^T u^2(s)ds] < \infty \tag{4.5}$$

By definition (4.2.1) we are looking for admissible feedback Markov controls $u_1(t)$ and $u_2(t)$ which are measurable, $\mathcal{F}_t$-adapted and bounded in L^2, such that the controlled process $W(t)$ becomes an Itô-diffusion. We say that $u(t)$ is a Markov feedback control if $u(t) = u_0(t, W(t), Z_1(t), Z_2(t))$ where u_0 is a measurable function from $[0,T] \times \mathbb{R}^3$ to U. The last condition, in connection with the adaptedness of the controls, ensures the Itô-integrability of $W(t)$, i.e. that there exists a solution to the wealth dynamics $dW(t)$, (given in 4.6) under all $u \in \mathcal{A}$. Note that the adaptedness of the control processes ensures the non-anticipativity condition, i.e. at each time t the investor is only familiar with the market information revealed up till time t and must base his decision on this information solely. The controller cannot exercise his decision before t and cannot control $W(s)$ for $s \geq t$. By the fact that the controls $u(t) \in \mathcal{A}$ are not constrained as in Merton's portfolio problem, we allow both borrowing and short selling of the assets.

Remark 4.2.2.
Note that if the set U*, which* $u(t)$ *maps to, is bounded, then condition (iii) in definition (4.2.1) is immediately fulfilled.*

Notation 4.2.3.
From now on, to ease the notation, by $W(t)$ *we mean the controlled process* $W(t, u_1(t), u_2(t))$ *and by* $u_i(t)$ *we refer to the feedback controls* $u_i(t, W(t), Z_1(t), Z_2(t))$ *for* $i = 1, 2$*.*

Under the assumption that the portfolio is self-financing, we obtain the following proposition.

Proposition 4.2.4 (The dynamics of $W(t)$).
Given initial wealth $W(t_0) = w_0$*, the controlled wealth process* $W(t, u_1(t), u_2(t))$ *is the solution to the stochastic differential equation*

$$\begin{aligned} dW(t) = W(t)\Big(&u_1(t)(\mu + A_1 Z_1(t) + \frac{1}{2}\Sigma_1) + u_2(t)(\mu + A_2 Z_2(t) + \frac{1}{2}\Sigma_2) \\ &+ (1 - u_1(t) - u_2(t))r\Big)dt + W(t)\Big(\sigma(u_1(t) + u_2(t))d\widetilde{B} + \sigma_1 u_1(t) dB_1(t) \\ &+ \sigma_2 u_2(t) dB_2(t)\Big) \end{aligned} \tag{4.6}$$

where $Z_1(t), Z_2(t)$ *are Ornstein-Uhlenbeck processes given in (2.14) and* Σ_1, Σ_2 *constants defined in Appendix A.*

Proof of proposition (4.2.4). The total wealth of the investor is given by the position held in the risky assets $i = 1, 2$ and the position held in the risk-free asset, i.e.

$$W(t) = \frac{u_1(t)W(t)}{S_1(t)} S_1(t) + \frac{u_2(t)W(t)}{S_2(t)} S_2(t) + \frac{(1 - u_1(t) - u_2(t))W(t)}{X_0(t)} X_0(t)$$

By the assumption of self-financing portfolio, we have that

$$dW(t) = \frac{u_1(t)W(t)}{S_1(t)} dS_1(t) + \frac{u_2(t)W(t)}{S_2(t)} dS_2(t) + \frac{(1 - u_1(t) - u_2(t))W(t)}{X_0(t)} dX_0(t)$$

for $t_0 \le t \le T$. By substituting (4.1) for $i = 1, 2$ for $dS_1(t), dS_2(t)$ and (4.2) for $dX_0(t)$, we obtain equation (4.6). ■

Clearly, we have a unique and strong solution to the stochastic differential equation (4.6), namely the wealth process $W(t)$. By (4.2), if the controls $u_1(t)$ and $u_2(t)$ are measurable and adapted processes, the controlled wealth process becomes a semimartingale (Definition B.1.6), especially an adapted process itself. All the more, $W(t)$ becomes a Markov process, allowing us to use the theory on controlled Markov diffusion problems given in Chapter 3.

Given initial wealth $W(t_0) = w_0$, the aim of the investor is to maximize expected utility of wealth at some future time $T \ge t_0$ given an utility function $U(W(t))$. We denote by $J(t_0, w_0, z_1, z_2; u_1, u_2)$ the performance function

$$J(t_0, w_0, z_1, z_2; u_1, u_2) = E^{w_0}\Big[U(W(\tau_D))\mathcal{X}_{\{\tau_D < \infty\}}\Big] \tag{4.7}$$

where $U(w) : \mathbb{R} \to \mathbb{R}$ represents our choice of utility function, namely the power utility function given in (3.23), and $E^{w_0}[X]$ denotes the expected value of a random variable X w.r.t. the probability law of $W(t)$ starting at (t_0, w_0, z_1, z_2).

Let τ_D denote the first exit time of $W(t)$ from the fixed domain region $D \in [t_0, T] \times \mathbb{R}^3$

$$\tau_D = \inf\{t > t_0 : W(t) \notin D\} \tag{4.8}$$

for

$$D = \{(t, w, z_1, z_2) : t_0 \leq t < T, w > 0, z_1, z_2\} \tag{4.9}$$

Remark 4.2.5.
Note that, either $\tau_D = T$ or $\tau_D = t$ where t is the time at which $W(t) = 0$, i.e. the investor reaches "bankruptcy" in the sense that the wealth is zero.

Since we do not put any restrictions on the processes $Z_1(t)$ and $Z_2(t)$ in the domain D, we assume that $E^{w_0}[\tau_D] < \infty$, and that the points of the boundary of D are regular, i.e. if $W(t)$ starts at $(t_0, w_0, z_1, z_2) \in \partial D$, the process is immediately stopped and $W(t)$ will never return to D, hence in this case $\tau_D = 0$. We summarize the statements so far in the following problem

Problem 4.2.6 (Optimal control problem).
The optimal stochastic control problem consists of finding the optimal pair of controls $u_1^*(t), u_2^*(t)$ among all admissible controls and the corresponding value function $\Phi(t_0, w_0, z_1, z_2)$ for all $(t_0, w_0, z_1, z_2) \in D$, such that

$$\begin{aligned}\Phi(t_0, w_0, z_1, z_2) :=& \sup_{u_1, u_2 \in \mathcal{A}} J(t_0, w_0, z_1, z_2; u_1, u_2)\\ =& \sup_{u_1, u_2 \in \mathcal{A}} E^{w_0}\Big[U(W(\tau_D))\mathcal{X}_{\{\tau_D < \infty\}}\Big]\\ =& J(t_0, w_0, z_1, z_2; u_1^*, u_2^*)\end{aligned} \tag{4.10}$$

In problem (4.2.6), $\Phi(t_0, w_0, z_1, z_2)$ denotes the optimal value or the optimal performance function introduced in Chapter 3. Recall that if problem (4.10) has a solution, it is said to be solvable, and if the solution is finite, the stochastic control problem is itself said to be finite in the strong sense, and $W(t, u_1^*(t), u_2^*(t)$ is said to be the optimal controlled state process. The next theorem presents the dynamic programming principle in connection to the control problem stated above.

Theorem 4.2.7 (HJB-equation).
Define the value function Φ by

$$\Phi(t, w, z_1, z_2) = \sup_{u_1, u_2 \in \mathcal{A}} J(t, w, z_1, z_2; u_1, u_2) \quad \forall\ (t, w, z_1, z_2) \in D \tag{4.11}$$

$$J(t, w, z_1, z_2; u_1, u_2) = E^w\Big[U(W(\tau_D))\Big] \tag{4.12}$$

and assume that $\Phi \in C^{1,2,2,2}(D) \cap C(\bar{D})$ where ∂D is regular for $W(t)$, $E^w[\tau_D] < \infty$ and

$$E^w[\Phi(W(\tau_D))] < \infty \tag{4.13}$$

i.e, Φ is bounded on ∂D. Further, assuming that optimal controls $u_1^*, u_2^* \in \mathcal{A}$ exist, the value function Φ satisfies the HJB-equation

$$\sup_{v_1, v_2 \in \mathcal{A}} (L^{v_1, v_2}\Phi)(t, w, z_1, z_2) = 0 \quad \forall\ (t, w, z_1, z_2) \in D \tag{4.14}$$

with boundary values

$$\Phi(t, w, z_1, z_2) = U(W(t)) \quad \forall\ (t, w, z_1, z_1) \in \partial D \tag{4.15}$$

for v_1, v_2 values in U. If $u_1^*(t) = v_1$ and $u_2^*(t) = v_2$, the supremum in equation (4.14) is obtained for u_1^*, u_2^* and

$$(L^{u_1^*,u_2^*}\Phi)(t, w, z_1, z_2) = 0 \quad \forall\ (t, w, z_1, z_2) \in D \tag{4.16}$$

Remark 4.2.8.
The notation $\phi \in C^{1,2,2,2}(D)$ means that the function ϕ is twice continuously differential w.r.t. the state variables w, z_1, z_2 in D and continuously differentiable w.r.t. t. Further, $\phi \in C(\partial D)$ means that the value function is continuous on the boundary of D.

The theorem gives a connection between the stochastic optimal control problem and the non-linear second order partial differential equation referred to as the Hamilton-Jacobi-Bellman equation. We give an adapted version of the proof from [Øks13].

Proof of theorem (4.2.7). We start by proving the boundary values (4.15). Assume $(t, w, z_1, z_2) \in \partial D$. Since the domain D puts no restriction on the state variables z_1 and z_2, by the assumption that ∂D is regular w.r.t. $W(t)$, $\tau_D = 0$ $\mathcal{P}$-a.s. if either $w = 0$ or $t = T$. Assuming controls $u_1^*(t) = v_1$ and $u_2^*(t) = v_2$ exist and are optimal, by definition of Φ

$$\begin{aligned}\Phi(t, w, z_1, z_2) &= E^{(t,w,z_1,z_2)}\Big[U(W(\tau_D))\Big] \\ &= \begin{cases} E^{(t,0,z_1,z_2)}\Big[U(W(0))\Big] = U(0) \text{ for } w = 0 \\ E^{(T,w,z_1,z_2)}\Big[U(W(T))\Big] = U(w) \text{ for } t = T, W(T) = w \end{cases}\end{aligned}$$

Hence $\Phi(t, w, z_1, z_2) = U(w)$ for $(t, w, z_1, z_2) \in \partial D$. Problem (4.15)-(4.16) can be recognized as a Dirichlet problem (Problem B.2.11) to which a solution Φ exists if $\Phi \in C^{1,2,2,2}(D) \cap C(\bar{D})$, Φ bounded on ∂D for ∂D regular and $E^w[\tau_D] < \infty$. By assumption these conditions hold, hence Φ exists and we have proven (4.16) for u_1^*, u_2^* optimal. We have only left to show (4.14), i.e. that the supremum over all values $v_1, v_2 \in U$ of the generator of Φ is zero. For this part we would like to exploit the strong Markov property of $W(t)$, Theorem B.2.8. Recall that

$$E[f(X(\tau + h))|\mathcal{F}_\tau] = E[f(X^y(h))]_{|X(\tau)=y}$$

for some measurable function f, where $\mathcal{F}_\tau$ is the stopping-time σ-algebra and $y = X(\tau)$ is the process evaluated at τ. Let $t \leq \tau_H \leq \tau_D$ define the first exit time from a smaller subset H of the domain, $H \subset D$

$$\tau_H = \inf\{r > t : W(r) \notin H\}$$

where

$$H = \{(s, w, z_1, z_2) : s < T, w > 0, z_1, z_2\}$$

Fix $u_1(t) = v_1$ and $u_2(t) = v_2$ for $u_1, u_2 \in \mathcal{A}$ and a point $(t, w, z_1, z_2) \in D$. Then

$$E^w[J(\tau_H, W(\tau_H), z_1, z_2; u_1, u_2)] = E^w[E^{W(\tau_H)}[U(W(\tau_D)]]$$

$$=E^w[E^w[U(W(\tau_D)|\mathcal{F}_{\tau_H}]]$$
$$=E^w[U(W(\tau_D))]$$
$$=J(t,w,z_1,z_2;v_1,v_2)$$

where the third equality follows from the law of total expectation. By the definition of Φ, this implies that

$$E^w[\Phi(\tau_H)] = J(t,w,z_1,z_2;v_1,v_2) \leq \Phi(t,w,z_1,z_2) \tag{4.17}$$

Now define a new pair of admissible controls

$$u_i(s,w,z_1,z_2) = \begin{cases} (v_1,v_2) & \text{if } (s,w,z_1,z_2) \in H \\ (u_1^*,u_2^*) & \text{if } (s,w,z_1,z_2) \in D \setminus H \end{cases}$$

for arbitrary $v_1, v_2 \in U$ and u_1^*, u_2^* optimal and assumed to exists. By Dynkin's formula, Theorem B.2.4

$$E^w[\Phi(W(\tau_H))] = \Phi(s,w,z_1,z_2) + E^w\Big[\int_t^{\tau_H} (L^{u_1,u_2}\Phi)(r,w(r),z_1(r),z_2(r))dr\Big] \tag{4.18}$$

Inserting (4.18) into (4.17), we obtain

$$E^w\Big[\int_t^{\tau_H} (L^{u_1,u_2}\Phi)(r,w(r),z_1(r),z_2(r))dr\Big] \leq 0$$

Dividing by $E^w[\tau_H]$ (> 0 by the fact that $t \leq \tau_H \leq \tau_D$) and letting $s \to t$, i.e. expanding the subspace H , $\tau_H \to \tau_D$ and we obtain

$$\lim_{\tau_H \to \tau_D} \frac{1}{E^w[\tau_H]} E^w\Big[\int_t^{\tau_H} (L^{u_1,u_2}\Phi)(r,w(r),z_1(r),z_2(r))dr\Big]$$
$$= (L^{v_1,v_2}\Phi)(t,w,z_1,z_2) \leq 0$$

by the fact that the partial derivatives of Φ are continuous for all points in D. If v_1, v_2 are the values to the optimal controls u_1^*, u_2^* respectively, by (4.16) we obtain equality. This concludes the proof. ■

The simplicity of the proof of Theorem 4.2.7 is solely based on the "smooth" assumptions of the value function, i.e. $\Phi \in C^{1,2,2,2}(D) \cap C(\bar{D})$, yielding a classical solution to the non-linear partial differential equation we call the HJB-equation. As mentioned in Chapter 3, more often than not, the value function is not smooth, and viscosity solutions must be introduced. Such solutions yield a far more complex theory and thus fare more complex proofs.

In Theorem 4.2.7, $(L^{v_1,v_2}\Phi)(t,w,z_1,z_2)$ denotes the infinitesimal generator of $W(t)$ for each choice of u_1, u_2 and any $(t,w,z_1,z_2) \in D$. The specific expression for the generator of the wealth process $W(t)$ is given in the following proposition.

Proposition 4.2.9 (The Infinitesimal Generator of $W(t)$).
Assuming $W(t)$ is an Itô diffusion in $\mathbb{R}$, for $(t_0,w_0,z_1,z_2) \in D$, let $\phi(t,w,z_1,z_2) : [0,T] \times \mathbb{R}^3$ be a function such that $\phi \in C^{1,2,2,2}(D) \cap C(\bar{D})$ and

let $u_i(t) = v_i \in U$ for $i = 1, 2$, $u_i \in \mathcal{A}$. The infinitesimal generator of $W(t)$ for each choice of u_1, u_2 is given by

$$\begin{aligned}(L^{v_1,v_2})(t_0, w_0, z_1, z_2) = {} & \frac{\partial\phi}{\partial t} + w\Big(v_1(\mu + A_1 z_1 + \frac{1}{2}\Sigma_1) + v_2(\mu + A_2 z_2 + \frac{1}{2}\Sigma_2) \\ & +(1 - v_1 - v_2)r\Big)\frac{\partial\phi}{\partial w} + A_1 z_1 \frac{\partial\phi}{\partial z_1} + A_2 z_2 \frac{\partial\phi}{\partial z_2} + \frac{1}{2}w^2\Big(\sigma^2(v_1 + v_2)^2 + \sigma_1^2 v_1^2 \\ & +\sigma_2^2 v_2^2 + 2(\sigma(v_1 + v_2)(\rho_1\sigma_1 v_1 + \rho_2\sigma_2 v_2) + \rho\sigma_1\sigma_2)\Big)\frac{\partial^2\phi}{\partial w^2} + \frac{1}{2}\sigma_1^2\frac{\partial^2\phi}{\partial z_1^2} \\ & +\frac{1}{2}\sigma_2^2\frac{\partial^2\phi}{\partial z_2^2} + w(v_1 P_1 + v_2 R_1)\frac{\partial^2\phi}{\partial w\partial z_1} + w(v_1 R_2 + v_2 P_2)\frac{\partial^2\phi}{\partial w\partial z_2} \\ & +\rho\sigma_1\sigma_2\frac{\partial^2\phi}{\partial z_1\partial z_2} \end{aligned} \tag{4.19}$$

where Σ, Σ_i, P_i and R_i for $i = 1, 2$ are constants defined in Appendix A.

Proof of proposition (4.2.9). Let $W(0) = w_0$ and $f(t, w, z_1, z_2) \in C^{1,2,2,2}([0, \infty) \times \mathbb{R}^3)$. Define the Itô process

$$Y(t, w, z_1, z_2) = f(t, W(t), Z_1(t), Z_2(t)), \quad Y(0) = w_0$$

By Itô's lemma

$$\begin{aligned} dY(t) = {} & \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial w}dW_t + \frac{\partial f}{\partial z_1}dZ_1(t) + \frac{\partial f}{\partial z_2}dZ_2(t) + \frac{1}{2}\frac{\partial f}{\partial w}dW_t^2 + \frac{1}{2}\frac{\partial^2 f}{\partial z_1^2}dZ_1(t)^2 \\ & + \frac{1}{2}\frac{\partial^2 f}{\partial z_2^2}dZ_2(t)^2 + \frac{\partial^2 f}{\partial w\partial z_1}dW_t \cdot dZ_1(t) + \frac{\partial^2 f}{\partial w\partial z_2}dW_t \cdot dZ_2(t) \\ & + \frac{\partial^2 f}{\partial z_1\partial z_2}dZ_1(t) \cdot dZ_2(t) \end{aligned}$$

Substituting for $dW(t)$, $dZ_1(t)$, $dZ_2(t)$, then by using lemma (C.1.1) we obtain that

$$\begin{aligned} dY(t) = {} & \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial w}W_t\Big((v_1(t)(\mu + b_1 A_1 Z_1(t) + \frac{1}{2}\Sigma_1) + v_2(t)(\mu + A_2 Z_2(t) \\ & + \frac{1}{2}\Sigma_2) + (1 - v_1(t) - v_2(t))r)dt + \sigma(v_1(t) + v_2(t))d\widetilde{B}(t) + \sigma_1 v_1(t)dB_1(t) \\ & + \sigma_2 v_2(t)dB_2(t)\Big) + \frac{\partial f}{\partial z_1}\Big(A_1 Z_1(t)dt + \sigma_1 dB_1(t)\Big) + \frac{\partial f}{\partial z_2}\Big(A_2 Z_2(t)dt \\ & + \sigma_2 dB_2(t)\Big) + \frac{1}{2}\frac{\partial^2 f}{\partial w^2}\Big(W_t^2(\sigma^2(v_1(t) + v_2(t))^2 + \sigma_1^2 v_1^2(t) + \sigma_2^2 v_2^2(t))dt \\ & + 2W_t^2(\sigma(v_1(t) + v_2(t))(\rho_1\sigma_1 v_1(t) + \rho_2\sigma_2 v_2(t)) + \rho\sigma_1 v_1(t)\sigma_2 v_2(t)\Big)dt \\ & + \frac{1}{2}\frac{\partial^2 f}{\partial z_1^2}\sigma_1^2 dt + \frac{1}{2}\frac{\partial^2 f}{\partial z_2^2}\sigma_2^2 dt + \frac{\partial^2 f}{\partial w\partial z_1}\Big(W_t(v_1(t)P_1 + v_2(t)R_1)\Big)dt \\ & + \frac{\partial^2 f}{\partial w\partial z_2}\Big(W_t(v_1(t)R_2 + v_2(t)P_2)\Big)dt + \frac{\partial^2 f}{\partial z_1\partial z_2}\rho\sigma_1\sigma_2 dt \end{aligned}$$

By integration and by taking expectations on both sides, the expression for $dY(t)$ becomes

$$E^{w_0}[Y(t)] = f(w_0) + E^{w_0}\Big[\int_{t_0}^{t} b(t, w, z_1, z_2)ds\Big] + E^{w_0}\Big[\int_{t_0}^{t} \sigma(s, w, z_1, z_2)d\widetilde{B}(s)$$

$$+ \int_{t_0}^{t} \sigma_1(s, w, z_1, z_2) dB_1(s) + \int_{t_0}^{t} \sigma_2(s, w, z_1, z_2) dB_2(s)\Big] \quad (4.20)$$

where $b(t, w, z_1, z_2)$ is a function of the terms corresponding to the dt-differential, and

$$\sigma = \sigma(v_1(t) + v_2(t))W(t)\frac{\partial f}{\partial w}, \quad \sigma_i = \sigma_i v_i(t)W(t)\frac{\partial f}{\partial w} + \sigma_i \frac{\partial f}{\partial z_i}, \; i = 1, 2 \quad (4.21)$$

generic constants. By assumption (iv) in (4.2.1), $u_1(t)$ and $u_2(t)$ are such that $W(t)$ is the unique solution to (4.6), σ, σ_1 and σ_2 are Itô-integrable functions and the integrals have zero expectation. By Fubini's theorem, Theorem B.3.1, equation (4.20) equals

$$E^{w_0}[Y(t)] - f(w_0) = \int_{t_0}^{t} E^{w_0}[b(t, w, z_1, z_2)] ds$$

Dividing by $t - t_0$ and letting $t \to t_0$, Definition B.2.2 and the continuity of f gives that

$$\lim_{t \to t_0} \frac{E^x[Y(t)] - f(x)}{t - t_0} = \lim_{t \to t_0} \frac{1}{t - t_0} E^{w_0}[b(t, w, z_1, z_2)] = E^{w_0}[b(t_0, w_0, z_1, z_2)]$$
$$= b(t_0, w_0, z_1, z_2)$$

where we recognize that $b(t_0, w_0, z_1, z_2) = (L^{v_1,v_2})(t_0, w_0, z_1, z_2)$ in (4.19). ■

4.3 A Specific Verification Theorem

Theorem 4.2.7 turns the original stochastic control problem into an much simpler maximization problem through the HJB-equation. It states that if optimal controls u_1^*, u_2^* exist, their values v_1, v_2 for any $(t_0, w_0, z_1, z_2) \in D$ must be such that the generator of $W(t)$ attains it maximum at the specific values v_1, v_2 for $(t_0, w_0, z_1, z_2) \in D$. The next theorem ensures that if such maximizing values $v_1 = u_1(t)$ and $v_2 = u_2(t)$ are found for any $(t_0, w_0, z_1, z_2) \in D$, they are indeed values corresponding to the optimal controls. In other words, Theorem 4.3.1 gives a verification of the solution obtained from Theorem 4.2.7.

Theorem 4.3.1 (A verification theorem).
Assume $\phi \in C^{1,2,2,2}(D) \cap C(\bar{D})$ is a solution to the HJB-equation

$$(L^{v_1,v_2}\phi)(t_0, w_0, z_1, z_2) \leq 0 \quad \forall \; (t_0, w_0, z_1, z_2) \in D \quad (4.22)$$

for all $v_1, v_2 \in U$ with boundary values

$$\lim_{t \to \tau_D} \phi(W(t)) = U(W(\tau_D)) \quad (4.23)$$

Assume also that

$$\int_{t_0}^{T} E\Big[\Big(\sigma(s, W(s, v_1, v_2), Z_1(s), Z_2(s))\Big)^2\Big] ds < \infty \quad (4.24)$$

$$\int_{t_0}^{T} E\Big[\Big(\sigma_i(s, W(s, v_1, v_2), Z_1(s), Z_2(s))\Big)^2\Big] ds < \infty \quad (4.25)$$

for σ, σ_i $(i = 1, 2)$ given in (4.21). Then

$$\phi(t_0, w_0, z_1, z_2) \geq J(t_0, w_0, z_1, z_2; u_1, u_2) \tag{4.26}$$

for all Markov controls $u_1(t) = v_1$, $u_2(t) = v_2$ and all $(t_0, w_0, z_1, z_2) \in D$. If in addition, we have that

$$(L^{u_1^*, u_2^*}\phi)(t_0, w_0, z_1, z_2) = 0 \tag{4.27}$$

then the Markov controls $u_1^*(t), u_2^*(t) \in \mathcal{A}$ define the optimal (feedback) controls and $\phi(t_0, w_0, z_1, z_2)$ must be the optimal value function such that

$$\phi(t_0, w_0, z_1, z_2) = J(t_0, w_0, z_1, z_2; u_1^*, u_2^*) = E^{w_0}\Big[U(W(\tau_D)\Big] \tag{4.28}$$

Proof of theorem (4.3.1). Let $(t_0, w_0, z_1, z_2) \in D$ and assume $\phi \in C^{1,2,2,2}(D) \cap C(\bar{D})$. In addition, assume (4.24) and (4.25). We want to show that if, for some values $v_1, v_2 \in U$, condition (4.22) holds, then

$$\phi(t_0, w_0, z_1, z_2) \geq J(t_0, w_0, z_1, z_2; v_1, v_2) \tag{4.29}$$

for all $(t_0, w_0, z_1, z_2) \in D$. By Itô's lemma we have that

$$\begin{aligned} d\phi =& \frac{\partial \phi}{\partial t} + \frac{\partial \phi}{\partial w} dW(t) + \frac{\partial \phi}{\partial z_1} dZ_1(t) + \frac{\partial \phi}{\partial z_2} dZ_2(t) + \frac{\partial^2 \phi}{\partial w^2}(dW(t))^2 \\ &+ \frac{\partial^2 \phi}{\partial z_1^2}(dZ_1(t))^2 + \frac{\partial^2 \phi}{\partial z_2^2}(dZ_2(t))^2 + \frac{\partial^2 \phi}{\partial w \partial z_1} dW(t) dZ_1(t) \\ &+ \frac{\partial^2 \phi}{\partial w \partial z_2} dW(t) dZ_2(t) + \frac{\partial^2 \phi}{\partial z_1 \partial z_2} dZ_1(t) dZ_2(t) \end{aligned} \tag{4.30}$$

Substituting for $dW(t)$, $dZ_1(t)$, $dZ_2(t)$, (expressions for the squared dynamics and the cross-products are given in Appendix B, and collecting all dt-differential terms, $d\widetilde{B}(t)$-differential terms, etc., we obtain the following equation

$$\begin{aligned} \phi(W(t)) =& \phi(t_0, w, z_1, z_2) + \int_{t_0}^{t} \bigg(\frac{\partial \phi}{\partial t} + W(s)\Big(v_1(\mu + A_1 Z_1(s) + \frac{1}{2}\Sigma_1) \\ &+ v_2(\mu + A_2 Z_2(s) + \frac{1}{2}\Sigma_2) + (1 - v_1 - v_2)r\Big)\frac{\partial \phi}{\partial w} + A_1 Z_1(s)\frac{\partial \phi}{\partial z_1} \\ &+ A_2 Z_2(s)\frac{\partial \phi}{\partial z_2} + \frac{1}{2}W^2(s)\Big(\sigma^2(v_1 + v_2)^2 + \sigma_1^2 v_1^2 + \sigma_2^2 v_2^2 + 2\sigma(v_1 + v_2) \\ &\times(\rho_1 \sigma_1 v_1 + \rho_2 \sigma_2 v_2) + 2\rho\sigma_1\sigma_2\Big)\frac{\partial^2 \phi}{\partial w^2} + \frac{1}{2}\Big(\sigma_1^2 \frac{\partial^2 \phi}{\partial z_2} + \sigma_2^2 \frac{\partial^2 \phi}{\partial z_2^2}\Big) + W(s) \\ &\times\Big(P_1 v_1 + R_1 v_2\Big)\frac{\partial^2 \phi}{\partial w \partial z_1} + W(s)\Big(R_2 v_1 + P_2 v_2\Big)\frac{\partial^2 \phi}{\partial w \partial z_2} + \rho\sigma_1\sigma_2 W(s) \\ &\times\Big(P_1 v_1 + R_1 v_2\Big)\frac{\partial^2 \phi}{\partial z_2 \partial z_2}\bigg)ds + \int_{t_0}^{T} W(s)\Big(\sigma(v_1 + v_2)\frac{\partial \phi}{\partial w}\Big)d\widetilde{B}(s) \\ &+ \int_{t_0}^{T} \Big(\sigma_1 v_1 W(s)\frac{\partial \phi}{\partial w} + \sigma_1 \frac{\partial \phi}{\partial z_1}\Big)dB_1(s) + \int_{t_0}^{T} \Big(\sigma_2 v_2 W(s)\frac{\partial \phi}{\partial w} \\ &+ \sigma_2 \frac{\partial \phi}{\partial z_2}\Big)dB_2(s) \end{aligned} \tag{4.31}$$

By assumptions (4.24) and (4.25), we know that the integrals w.r.t. $d\widetilde{B}(s)$, $dB_i(s)$ for $i = 1, 2$ are Itô integrals, hence taking expectations on both sides of (4.31) and recognizing that the Riemann integral is the integral of the generator $(L^{v_1,v_2}\phi)(t_0, w_0, z_1, z_2)$ in (4.31), we obtain

$$E[\phi(W(t))] = \phi(t_0, w_0, z_1, z_2) + E^{w_0}\Big[\int_{t_0}^{t}(L^{v_1,v_2}\phi)(t_0, w_0, z_1, z_2)ds\Big] \\ \leq \phi(t_0, w_0, z_1, z_2) \tag{4.32}$$

by (4.22). Let $t \to \tau_D$, then by (4.32)

$$\lim_{t\to\tau_D} E[\phi(W(t))] = E[U(W(\tau_D))] = J(t_0, w_0, z_1, z_2; v_1, v_2) \\ \leq \phi(t_0, w_0, z_1, z_2) \tag{4.33}$$

which proves (4.26) for $(t_0, w_0, z_1, z_2) \in D$. For the second part of the theorem, since $u_1(t)$ and $u_2(t)$ are assumed to be Markov controls and by the construction of $W(t)$, $W(t, u_1(t), u_2(t))$ is the unique solution to the dynamics $dW(t)$ in (4.6), i.e. $u_1(t)^*$ and $u_2^*(t)$ are admissible. By (4.28) the controls are maximizers to the HJB-equation, and following the same calculations as in the first part of the proof, only with equalities, we obtain

$$\lim_{t\to\tau_D} E[\phi(W(t))] = E[U(W(\tau_D))] = J(t_0, w_0, z_1, z_2; u_1^*.u_2^*) \\ = \phi(t_0, w_0, z_1, z_2) \tag{4.34}$$

for any $(t_0, w_0, z_1, z_2) \in D$. Hence u_1^* and u_2^* must be optimal admissible controls. This completes the prove for values in the domain D. The remaining part is to prove the boundary values, $t = T$ and $w = 0$. By (4.28), for both $t = T$ and $w = 0$, $t = \tau_D$ and we see that

$$\phi(\tau_D, W(\tau_D), Z_1(\tau_D), Z_2(\tau_D)) = E[U(W(\tau_D))] = U(W(\tau_D)) \tag{4.35}$$

This concludes the proof for all values $(t, w, z_1, z_2) \in \bar{D}$. ■

4.4 A Semi-Explicit Solution

In this section we aim at solving the optimal control problem stated in theorem Theorem 4.2.7. That is, our goal is to find the optimal pair of controls $(u_1^*(t), u_2^*(t))$ and the corresponding optimal value function $\Phi(t_0, w_0, z_1, z_2)$ for any (t_0, w_0, z_1, z_2) such that $u_1^*(t)$ and $u_2^*(t)$ maximizes expected utility of wealth at time T. By Theorem 4.2.7, the optimal controls $u_1^*(t)$ and $u_2^*(t)$ are given as the maximizing values to the HJB-equation (4.14). For each $(t_0, w_0, z_1, z_2) \in D$ we try to find the values $v_1 = u_1^*(t)$ and $v_2 = u_2^*(t)$ such that ν attains it maximum

$$\nu(v_1, v_2) = (L^{v_1,v_2})(t_0, w_0, z_1, z_2) = 0 \tag{4.36}$$

By first and second-order conditions for a maximum, the optimum in (4.36) is obtained if $\Phi_x > 0$ and $\Phi_{xx} < 0$. A solution is given by solving the first-order optimality conditions for v_1 and v_2.

Proposition 4.4.1 (Optimal values v_1 and v_2).
The optimal values $v_1 = u_1(t)$ and $v_2 = u_2(t)$ maximizing the HJB-equation (4.14), are given by

$$v_1 = u_1^*(t) = \frac{1}{w\Gamma\Phi_{ww}}\Big(G_1(z_1,z_2)\Phi_w - M_1\Phi_{wz_1} - N_1\Phi_{wz_2}\Big) \tag{4.37}$$

$$v_2 = u_2^*(t) = \frac{1}{w\Gamma\Phi_{ww}}\Big(G_2(z_1,z_2)\Phi_w - M_2\Phi_{wz_1} - N_2\Phi_{wz_2}\Big) \tag{4.38}$$

where $\Gamma, G_i(z_1,z_2), M_i, N_i$ for $i = 1,2$ are as in Appendix A.

Proof of proposition (4.4.1). The values v_1, v_2 are given as the solutions to the first-order optimality equations

$$\frac{\partial \nu}{\partial v_1} = 0 \tag{4.39}$$

$$\frac{\partial \nu}{\partial v_2} = 0 \tag{4.40}$$

The solutions in (4.37) and (4.38) are obtained by straight forward derivation w.r.t. v_1 and v_2 of (4.36), then by substitution of variables to attain explicit expressions for v_1 and v_2. The calculations are comprehensive and require introduction of new variables several times. They are therefore omitted here, but the resulting expressions are presented in Appendix A. ■

Our next mission is to find the optimal performance function $\Phi(t_0, w_0, z_1, z_2)$ for any $(t_0, w_0, z_1, z_2) \in D$, corresponding to the optimal controls u_1^*, u_2^*. The procedure will be as follows: we make a guess on the value function Φ and obtain a new boundary value problem which we try to solve by again making an ansatz on the solution. When or if a solution is obtained, we have to verify that $u_1^*(t)$ and $u_2^*(t)$ are indeed admissible optimal controls under Φ be means of the Theorem 4.3.1.

Notation 4.4.2.
The following notational conventions will be used to denote the partial derivatives of Φ

$$\Phi_t = \frac{\partial\Phi}{\partial t}, \Phi_w = \frac{\partial\Phi}{\partial w}, \Phi_{ww} = \frac{\partial^2\Phi}{\partial w^2}, \Phi_{z_i} = \frac{\partial\Phi}{\partial z_i}, \Phi_{z_iz_j} = \frac{\partial^2\Phi}{\partial z_i^2}, \Phi_{wz_i} = \frac{\partial^2\Phi}{\partial w\partial z_i}$$

for all i,j. In addition, we suppress the dependencies on t and write $W(t) = w$, $Z_1(t) = z_1$ and $Z_2(t) = z_2$, not to be confused with the initial values z_1, z_2.

By substituting the optimal values v_1 and v_2 into the HJB-equation (4.14), we get the following non-linear boundary value problem for Φ

Problem 4.4.3 (Boundary value problem for Φ).

$$\Phi_t + rw\Phi_w + A_1z_1\Phi_{z_1} + A_2z_2\Phi_{z_2} + \frac{1}{2}\sigma_1^2\Phi_{z_1z_1} + \frac{1}{2}\Phi_{z_2z_2} + \rho\sigma_1\sigma_2\Phi_{z_1z_2}$$
$$+ \frac{\Phi_w^2}{\Gamma^2\Phi_{ww}}p(z_1,z_2) + \frac{\Phi_w\Phi_{wz_1}}{\Gamma^2\Phi_{ww}}p_1(z_1,z_2) + \frac{\Phi_w\Phi_{wz_2}}{\Gamma^2\Phi_{ww}}p_2(z_1,z_2)$$

$$+\frac{\Phi_{wz_1}^2}{\Gamma^2\Phi_{ww}}C_1+\frac{\Phi_{wz_2}^2}{\Gamma^2\Phi_{ww}}C_2+\frac{\Phi_{wz_1}\Phi_{wz_2}}{\Gamma^2\Phi_{ww}}C_3=0 \quad \textit{for } t<T, \quad w>0 \tag{4.41}$$

$$\Phi(t,w,z_1,z_2)=U(w) \quad \textit{for } t=T \textit{ or } W(t)=0 \tag{4.42}$$

where $p(z_1,z_2)$, $p_1(z_1,z_2)$ *and* $p_2(z_1,z_2)$ *are polynomials of degree 2, 1 and 1 respectively, and* C_1, C_2 *and* C_3 *are constants.*

The polynomials in problem (4.4.3) are on the form

$$p(z_1,z_2)=\alpha_0+\alpha_1A_1z_1+\alpha_2A_2z_2+\alpha_{11}A_1^2z_1^2+\alpha_{22}A_2^2z_2^2+\alpha_{12}A_1A_2z_1z_2$$

$$p_1(z_1,z_2)=\beta_0+\beta_1A_1z_1+\beta_2A_2z_2$$

$$p_2(z_1,z_2)=\delta_0+\delta_1A_1z_1+\delta_2A_2z_2$$

where the coefficients $\alpha_0,\ldots,\beta_0,\ldots,\delta_0,\ldots$ and the constants C_1, C_2, C_3 are different expressions of the correlations and standard deviations $\rho,\rho_i,\sigma,\sigma_i$ for $i=1,2$. The expressions are cumbersome and are hence omitted here. They are stated in Appendix A.

Given the strictly increasing and strictly concave power utility function presented in (3.23), we try to find a solution to Problem 4.4.3 on the form

$$\phi(t,w,z_1,z_2)=f(t,z_1,z_2)w^\gamma, \quad 0<\gamma<1 \tag{4.43}$$

where w^γ represents our choice of utility function, and $f(t,z_1,z_2)$ is a multi-variable function of t and the space variables z_1 and z_2. We substitute for ϕ in (4.41) and obtain the alternative non-linear boundary value problem for f

Problem 4.4.4 (Boundary value problem for f).

$$f_t+\gamma rf+A_1z_1f_{z_1}+A_2z_2f_{z_2}+\frac{1}{2}\sigma_1^2f_{z_1z_1}+\frac{1}{2}\sigma_2^2f_{z_2z_2}+\rho\sigma_1\sigma_2f_{z_1z_2}$$
$$+\frac{\gamma\cdot p(z_1,z_2)}{\Gamma^2(\gamma-1)}f+\frac{\gamma\cdot p_1(z_1,z_2)}{\Gamma^2(\gamma-1)}f_{z_1}+\frac{\gamma\cdot p_2(z_1,z_2)}{\Gamma^2(\gamma-1)}f_{z_2}+\frac{\gamma\cdot C_1}{\Gamma^2(\gamma-1)}\frac{f_{z_1}^2}{f}$$
$$+\frac{\gamma\cdot C_2}{\Gamma^2(\gamma-1)}\frac{f_{z_2}^2}{f}+\frac{\gamma\cdot C_3}{\Gamma^2(\gamma-1)}\frac{f_{z_1}f_{z_2}}{f}=0, \quad \textit{for } t<T, \quad w>0 \tag{4.44}$$

$$f(t,z_1,z_2)=1 \quad \textit{for } t=T \tag{4.45}$$

where f_t, $f_{z_1}, f_{z_2}, ..$ *are the partial derivatives of* f *w.r.t.* t, z_1 *and* z_2 *respectively.*

Note the boundary value for $f(t,z_1,z_2)$ when $t=T$. For the power utility function, the boundary condition, equation (4.45), implies that

$$\phi(t,w,z_1,z_2)=\begin{cases} f(t,z_1,z_2)\cdot U(0)=0, & \text{if } w=0 \\ f(t,z_1,z_2)\cdot U(w)=U(w), & \text{if } t=T \end{cases}$$

By guessing on a solution for ϕ, we have reduced one of the state variables, w, and are left with a non-linear partial differential equation in time and the variables z_1 and z_2. We try to solve the partial differential equation further by making the ansatz

$$f(t, z_1, z_2) = e^{g(t,z_1,z_2)} \tag{4.46}$$

where $g(t, z_1, z_2)$ is on the form

$$g(t, z_1, z_2) = g_0(t) + g_1(t)z_1 + g_2(t)z_2 + g_{11}(t)z_1^2 + g_{22}(t)z_2^2 + g_{12}(t)z_1z_2 \tag{4.47}$$

i.e., g is a polynomial of degree 2 in two space variables, where the coefficients $g_0(t), \dots, g_{12}(t)$ are functions of t. The boundary value condition for $t = T$, implies that

$$e^{g(T,z_1,z_2)} = 1 \implies g(T, z_1, z_2) = 0 \tag{4.48}$$

By the ansatz on $f(t, z_1, z_2)$, for non-zero state variables, when $t = T$ the coefficients of the polynomial $g(t, z_1, z_2)$ are zero, i.e.

$$g_0(t) = g_1(t) = \dots = g_{12}(t) = 0 \text{ for } t = T \text{ and } z_1, z_2 \neq 0 \tag{4.49}$$

Notation 4.4.5.
We will later denote by $g(t)$ the set of functions $g_0(t)$, $g_1(t)$, $g_2(t)$, $g_{11}(t)$, $g_{22}(t)$ and $g_{12}(t)$, while $g(t, z_1, z_2)$ is the polynomial in (4.47). Furthermore, to spare some notation, we define the constant

$$\zeta = \frac{\gamma}{\Gamma^2(\gamma - 1)}$$

By substituting for $f(t, z_1, z_2)$ and the partial derivatives f_t, f_{z_1}, f_{z_2}, we obtain a new equation for (4.44). Again, suppressing the dependencies on t, z_1, z_2, we obtain

$$\begin{aligned} &e^g g_t + \gamma r e^g + A_1 z_1 e^g g_{z_1} + A_2 z_2 e^g g_{z_2} + \frac{1}{2}\sigma_1^2 e^g g_{z_1z_1} + \frac{1}{2}\sigma_2^2 e^g g_{z_2z_2} + \rho\sigma_1\sigma_2 e^g g_{z_1z_2} \\ &\quad + p(z_1, z_2)e^g\zeta + p_1(z_1, z_2)e^g g_{z_1}\zeta + p_2(z_1, z_2)e^g g_{z_2}\zeta + C_1 e^g g_{z_1}^2\zeta \\ &\quad + C_2 e^g g_{z_2}^2\zeta + C_3 e^g g_{z_1} g_{z_2}\zeta = 0, \quad \text{for } t < T, \quad w > 0 \end{aligned} \tag{4.50}$$

We divide by $e^{g(t,z_1,z_2)}$, substitute for $p(z_1, z_2), p_1(z_1, z_2), p_2(z_1, z_2)$ and obtain the following "messy" equation

$$\begin{aligned} &g_0'(t) + g_1'(t)z_1 + g_2'(t)z_2 + g_{11}'(t)z_1^2 + g_{22}'(t)z_2^2 + g_{12}'(t)z_1z_2 \\ &\quad + A_1z_1\Big(g_1(t) + 2g_{11}(t)z_1 + g_{12}(t)z_2\Big) + A_2z_2\Big(g_2(t) + 2g_{22}(t)z_2 + g_{12}(t)z_1\Big) \\ &\quad + \frac{1}{2}\sigma_1^2\Big(2g_{11}(t) + g_1^2(t) + 4g_{11}^2(t)z_1^2 + g_{12}^2(t)z_2^2 + 4g_1(t)g_{11}(t)z_1 \\ &\quad + 2g_1(t)g_{12}(t)z_2 + 4g_{11}(t)g_{12}(t)z_1z_2\Big) + \frac{1}{2}\sigma_2^2\Big(2g_{22}(t) + g_2^2(t) + g_{12}^2(t)z_1^2 \\ &\quad + 4g_{22}^2(t)z_2^2 + 2g_2(t)g_{12}(t)z_1 + 4g_2(t)g_{22}(t)z_2 + 4g_{12}(t)g_{22}(t)z_1z_2\Big) \\ &\quad + \rho\sigma_1\sigma_2\Big(g_{12}(t) + g_1(t)g_2(t) + z_1(g_1(t)g_{12}(t) + 2g_2(t)g_{11}(t)) + z_2(2g_1(t)g_{22}(t) \\ &\quad + g_2(t)g_{12}(t)) + 2g_{11}(t)g_{12}(t)z_1^2 + 2g_{12}(t)g_{22}(t)z_2^2 + z_1z_2(4g_{11}(t)g_{22}(t) \\ &\quad + g_{12}^2(t))\Big) + \Big(\beta_0 + \beta_1A_1z_1 + \beta_2A_2z_2\Big)\Big(g_1(t) + 2g_{11}(t)z_1 + g_{12}(t)z_2\Big)\zeta \end{aligned}$$

$$
\begin{aligned}
&+ \Big(\delta_0 + \delta_1 A_1 z_1 + \delta_2 A_2 z_2\Big)\Big(g_2(t) + 2g_{22}(t)z_2 + g_{12}(t)z_1\Big)\zeta \\
&+ C_1\Big(g_1(t) + 2g_{11}(t)z_1 + g_{12}(t)z_2\Big)^2\zeta + C_2\Big(g_2(t) + 2g_{22}(t)z_2 + g_{12}(t)z_1\Big)^2 \\
&\times \zeta + C_3\Big(g_1(t) + 2g_{11}(t)z_1 + g_{12}(t)z_2\Big)\Big(g_2(t) + 2g_{22}(t)z_2 + g_{12}(t)z_1\Big) \\
= &-\Big(\alpha_0 + \alpha_1 A_1 z_1 + \alpha_2 A_2 z_2 + \alpha_{11} A_1^2 z_1^2 + \alpha_{22} A_2^2 z_2^2 + \alpha_{12} A_1 A_2 z_1 z_2\Big)\zeta - \gamma r
\end{aligned}
$$

So far our guess on $f(t, z_1, z_2)$ seems reasonable, and we are left with a polynomial of degree 2 where the coefficients are time-dependent functions. However, we need to show that $e^{g(t,z_1,z_2)}$ is indeed a solution to the partial differential equation. Writing out the parentheses and collecting all the terms not including z_1 or z_2 in one equation, all the terms including only z_1 and only z_2 in separate equations, and so on, we obtain a system of six first-order, non-linear (ordinary) differential equations of t.

Result 4.4.6 (System of differential equations).
By the procedure described above, we express each of the equations as an expression of the derivatives $g_0'(t), g_1'(t), \ldots$. Then the system of differential equations obtained is given by (1)-(6)

$$
\begin{aligned}
g_0'(t) = &-g_1(t)\beta_0\zeta - g_2(t)\delta_0\zeta + \sigma_1^2 g_{11}(t) - \sigma_2^2 g_{22}(t) - \rho\sigma_1\sigma_2 g_{12}(t) - g_1^2(t)\Big(\frac{1}{2}\sigma_1^2 \\
&-C_1\zeta\Big) - g_2^2(t)\Big(\frac{1}{2}\sigma_2^2 - C_2\zeta\Big) - g_1(t)g_2(t)\Big(\rho\sigma_1\sigma_2 - C_3\zeta\Big) - \gamma r - \alpha_0\zeta \qquad (1)
\end{aligned}
$$

$$
\begin{aligned}
g_1'(t) = &-g_1(t)\Big(A_1 + \beta_1 A_1\zeta\Big) - g_2(t)\delta_1 A_1\zeta - 2g_{11}(t)\beta_0\zeta - g_{12}(t)\delta_0\zeta \\
&-g_1(t)g_{11}(t)\Big(2\sigma_1^2 + 4C_1\zeta\Big) - g_1(t)g_{12}(t)\Big(\rho\sigma_1\sigma_2 + C_3\zeta\Big) - g_2(t)g_{11}(t) \\
&\times\Big(2\rho\sigma_1\sigma_2 + 2C_2\zeta\Big) - g_2(t)g_{12}(t)\Big(\sigma_2^2 + 2C_2\zeta\Big) - \alpha_1 A_1\zeta \qquad (2)
\end{aligned}
$$

$$
\begin{aligned}
g_2'(t) = &-g_1(t)\beta_2 A_2\zeta - g_2(t)\Big(A_2 + \delta_2 A_2\zeta\Big) - g_{12}(t)\beta_0\zeta - g_{22}(t)2\delta_0\zeta \\
&-g_1(t)g_{12}(t)\Big(\sigma_1^2 + 2C_1\zeta\Big) - g_2(t)g_{12}(t)\Big(\rho\sigma_1\sigma_2 - C_3\zeta\Big) - g_1(t)g_{22}(t)\Big(\rho\sigma_1\sigma_2 \\
&+2C_3\zeta\Big) - g_2(t)g_{22}(t)\Big(2\sigma_2^2 + 4C_2\zeta\Big) - \alpha_2 A_2\zeta \qquad (3)
\end{aligned}
$$

$$
\begin{aligned}
g_{11}'(t) = &-g_{11}(t)\Big(2A_1 + 2A_1\beta_1\zeta\Big) + g_{12}(t)\delta_1 A_1\zeta - g_{11}^2(t)\Big(2\sigma_1^2 + 4C_1\zeta\Big) \\
&+g_{12}^2(t)\Big(\frac{1}{2}\sigma_2^2 + C_2\zeta\Big) - g_{11}(t)g_{12}(t)\Big(2\rho\sigma_1\sigma_2 + 2C_3\zeta\Big) - \alpha_{11}A_1^2\zeta \qquad (4)
\end{aligned}
$$

$$
\begin{aligned}
g_{22}'(t) = &-g_{22}(t)\Big(2A_2 + 2\delta_2 A_2\zeta\Big) + g_{12}(t)\beta_2 A_2\zeta - g_{22}^2(t)\Big(2\sigma_2^2 + 3C_2\zeta\Big) \\
&-g_{12}^2(t)\Big(\frac{1}{2}\sigma_1^2 + C_1\zeta\Big) - g_{12}g_{22}(t)\Big(\rho\sigma_1\sigma_2 + 2C_3\zeta\Big) - \alpha_{22}A_2^2\zeta \qquad (5)
\end{aligned}
$$

$$
\begin{aligned}
g_{12}'(t) = &-2g_{11}(t)\beta_2 A_2\zeta - 2g_{22}(t)\delta_1 A_1\zeta - g_{12}(t)\Big(A_1 + A_2 + \beta_1 A_1\zeta + \delta_2 A_2\zeta\Big) \\
&-g_{12}^2(t)\Big(\rho\sigma_1\sigma_2 + C_3\zeta\Big) - g_{11}(t)g_{12}(t)\Big(2\sigma_1^2 + 4C_1\zeta\Big) - g_{11}(t)g_{22}(t)\Big(\rho\sigma_1\sigma_2
\end{aligned}
$$

$$+4C_3\zeta\Big) - g_{12}(t)g_{22}(t)\Big(2\sigma_2^2 + 4C_2\zeta\Big) - \alpha_{12}A_1A_2\zeta \tag{6}$$

By our guess on the value function, we have managed to cancel all state variables w, z_1 and z_2 from the initial HJB-equation, and reduced our control problem to a system of non-linear first order (ordinary) differential equations (1)-(6). In the next section we observe that the system of equations can be argued to be a set of Riccati equations, or equivalently a 6-dimensional vector-Riccati equation.

4.5 Riccati Representation

The non-linear system of ordinary differential equations derived in Result 4.4.6 in the previous sections is quite cumbersome. In this section we recognize (1)-(6) as a system of Riccati-equations, which we represent in two alternative ways.

The Riccati equations are a special type of differential equations, defined in terms of an ordinary differential equation that is quadratic in the unknown function. From [Bar16], we state the following definition

Definition 4.5.1 (Riccati-equation).
The non-linear ordinary differential equation

$$x' = p(t) + q(t)x + r(t)x^2, \quad t \in [0, T] \tag{4.51}$$

where $P(t)$, $Q(t)$ and $R(t) \neq 0$ are continuous functions on an interval $[0, T]$, is called a general Riccati differential equation.

The definition above can be extended to a matrix-valued Riccati equation, often appearing in linear quadratic optimization problems or filtering theory. The algebraic Riccati equations are on the form

$$X' = A^T X + XBR^{-1}B^T X + Q$$

where X is the unknown $n \times n$ symmetric matrix and A,B,R and Q are real-valued $n \times n$- matrices. Each of the equations in (1)-(6) constitute a Riccati equation. In this section we look at two Riccati representations similar to the vector Riccati equation

$$\mathbf{x}^{'} = \mathbf{A}(t) + B(t)\mathbf{x} + (\mathbf{C}(t)\mathbf{x})\mathbf{x}^T + Q$$

where $\mathbf{x}$ is a n-dimensional vector of the unknown function, $\mathbf{A}(t)$ and $\mathbf{C}(t)$ are continuous vector functions, and $B(t)$ is a $n \times n$ continuous matrix function. The vector Riccati equation is for instance presented in [AS10]. The following notation will become convenient.

Notation 4.5.2 (Riccati representation).
Let c_{mn} denote the elements of the following 6×6-matrix.

$$C = \begin{bmatrix} 0 & \beta_0\zeta & \delta_0\zeta & \sigma_1^2 & \sigma_2^2 & \rho\sigma_1\sigma_2 \\ 0 & A_1+\beta_1A_1\zeta & \delta_1A_1\zeta & 2\beta_0\zeta & 0 & \delta_0\zeta \\ 0 & \beta_2A_2\zeta & A_2+\delta_2A_2\zeta & 0 & 2\delta_0\zeta & \beta_0\zeta \\ 0 & 0 & 0 & 2A_1+2A_1\beta_1\zeta & 0 & \delta_1A_1\zeta \\ 0 & 0 & 0 & 0 & 2A_2+2\delta_2A_2\zeta & \beta_2A_2\zeta \\ 0 & 0 & 0 & 2\beta_2A_2\zeta & 2\delta_1A_1\zeta & (A_1+A_2 \\ & & & & & +\beta_1A_1\zeta+\delta_2A_2\zeta) \end{bmatrix}$$

and let G be the 6×6 matrix defined by

$$G = \begin{bmatrix} g_0^2(t) & g_0(t)g_1(t) & \cdots & g_0(t)g_{12}(t) \\ g_1(t)g_0(t) & g_1^2(t) & \cdots & g_1(t)g_{12}(t) \\ \vdots & & \ddots & \\ g_{12}(t)g_0(t) & g_{12}(t)g_1(t) & \cdots & g_{12}^2(t) \end{bmatrix}$$

Furthermore, let $\boldsymbol{N}$ be the column-vector in $\mathbb{R}^6$ with elements N_k, defined by

$$\boldsymbol{N} = \left(\gamma r + \alpha_0\zeta,\ \alpha_1 A_1\zeta,\ \alpha_2 A_2\zeta,\ \alpha_{11}A_1^2\zeta,\ \alpha_{22}A_2^2\zeta,\ \alpha_{12}A_1A_2\zeta\right)^T$$

and denote the constants ϵ_i and ψ by

$$\epsilon_i = \frac{1}{2}\sigma_i^2 + C_i\delta, \quad , \psi = \rho\sigma_1\sigma_2 + C_3\delta,$$

for $i = 1, 2$.

Applying the notation introduced above, the system of differential equations obtained in Result 4.4.6 can be represented as follows

Result 4.5.3 (System (1')-(6')).

$$\begin{aligned} g_0'(t) = &- g_1(t)c_{12} - g_2(t)c_{13} - g_{11}(t)c_{14} - g_{22}(t)c_{15} - g_{12}(t)c_{16} - g_1^2(t)\epsilon_1 \\ &- g_2^2(t)\epsilon_2 - g_1(t)g_2(t)\psi - N_1 \end{aligned} \tag{1'}$$

$$\begin{aligned} g_1'(t) = &- g_1(t)c_{22} - g_2(t)c_{23} - 2g_{11}(t)c_{24} - g_{12}(t)c_{26} - 4g_1(t)g_{11}(t)\epsilon_1 \\ &- g_1(t)g_{12}(t)\psi - 2g_2(t)g_{11}(t)\psi - g_2(t)g_{12}(t)\psi - N_2 \end{aligned} \tag{2'}$$

$$\begin{aligned} g_2'(t) = &- g_1(t)c_{32} - g_2(t)c_{33} - g_{22}(t)c_{35} - g_{12}(t)c_{36} - 2g_1(t)g_{12}(t)\epsilon_1 \\ &- 2g_1(t)g_{22}(t)\psi - 4g_2(t)g_{22}(t)\epsilon_2 - g_2(t)g_{12}(t)\psi - N_3 \end{aligned} \tag{3'}$$

$$\begin{aligned} g_{11}'(t) = &- g_{11}(t)c_{44} - g_{12}(t)c_{46} - 4g_{11}^2(t)\epsilon_1 + g_{12}^2(t)\epsilon_2 - 2g_{11}(t)g_{12}(t)\psi \\ &- N_4 \end{aligned} \tag{4'}$$

$$\begin{aligned} g_{22}'(t) = &- g_{22}(t)c_{55} - g_{12}(t)c_{56} - 4g_{22}^2(t)\epsilon_2 - g_{12}^2(t)\epsilon_1 - 2g_{12}g_{22}(t)\psi \\ &- \alpha_{22}A_2^2\zeta \end{aligned} \tag{5'}$$

$$\begin{aligned} g_{12}'(t) = &- 2g_{11}(t)c_{64} - 2g_{22}(t)c_{65} - g_{12}(t)c_{66} - g_{12}^2(t)\psi \\ &- 4g_{11}(t)g_{22}(t)\psi - 4g_{11}(t)g_{12}(t)\epsilon_1 - 4g_{22}(t)g_{12}(t)\epsilon_2 - N_6 \end{aligned} \tag{6'}$$

Representation Through Linear Operator L

In this subsection we aim at rewriting the shortened system of differential equations (1')-(6') on a form similar to the vector Riccati equation introduced at the beginning of the section. Let

$$\mathbf{g}'(t) = \mathbf{N} + C\mathbf{g}(t) + L(G)$$

where $\mathbf{g}(t)$ is a 6-dimensional column vector and $\mathbf{g}^T(t)$ the transpose. C and $\mathbf{N}$ are the matrix and the column-vector, respectively, defined in Notation 4.5.2. $L(G)$ a linear operator from $\mathbb{R}^{6\times 6} \rightarrow \mathbb{R}^6$ which takes the 6×6-matrix $G = \mathbf{g}(t)\mathbf{g}^T(t)$ to a $6-$dimensional column vector. We argue that (1')-(6'), equivalently (1)-(6), is a system of Riccati-equations, by first noting that each of the equations $(1')$-$(6')$ can be expressed as equations of a constant, a linear term and a quadratic term. Define

$$\mathbf{g}(t)^- = (g_0(t)^-, g_1(t)^-, g_2(t)^-, g_{11}(t)^-, g_{22}(t)^-, g_{12}(t)^-)^T$$

where $\mathbf{g}^-(t) \in \mathbb{R}^6$ denotes the vector containing the quadratic parts of equations $(1')$-$(6')$. Take for instance the expression for $g_0^{'}(t)$, which takes the form

$$g_0^{'}(t) = -N_1 - c_{12}g_1(t) - c_{13}g_2(t) - c_{14}g_{11}(t) - c_{15}g_{22}(t) - c_{16}g_{12}(t) - g_0(t)^-$$

for $g_0^-(t) = g_1^2(t)\epsilon_1 + g_2^2(t)\epsilon_2 + g_1(t)g_2(t)\psi$. We are now ready to state the following proposition.

Proposition 4.5.4 (System of Riccati-equations).
Define $\boldsymbol{g}(t) \in \mathbb{R}^6$ to be the 6-dimensional column-vector

$$\boldsymbol{g}(t) = (g_0(t), g_1(t), g_2(t), g_{11}(t), g_{22}(t), g_{12}(t))^T \tag{4.52}$$

and $\boldsymbol{g}^T(t)$ the transpose. The system of equations $(1')$-$(6')$ can be represented by the vector equation

$$\boldsymbol{g}'(t) = \boldsymbol{N} + C\boldsymbol{g}(t) + L(G) \tag{4.53}$$

where C is the 6×6-matrix of the coefficients of the linear terms and $\boldsymbol{N} \in \mathbb{R}^6$ the column-vector of constants, both defined in Notation 4.5.2. $L(G)$ is a linear operator from $\mathbb{R}^{6\times 6} \rightarrow \mathbb{R}^6$ defined by

$$L(G) = \begin{bmatrix} -e_2^T(G_2\epsilon_1 + G_3\psi) - e_3^T G_3\epsilon_2 \\ -e_2^T(4G_4\epsilon_1 + G_6\psi) - e_3^T(2G_4\psi + 2G_6\epsilon_2) \\ -e_2^T(G_6\psi + 4G_5\epsilon_2) - e_3^T(2G_6\epsilon_1 + G_5\psi) \\ -e_4^T(4G_4\epsilon_1 + 2G_6\psi) - e_6^T G_6\epsilon_2 \\ -e_5^T(4G_5\epsilon_2 + 2G_6\psi) - e_6^T G_6\epsilon_1 \\ -4e_4^T G_5\epsilon_1 - e_6^T(G_3\psi + G_4\psi + 4G_5\epsilon_2) \end{bmatrix} \tag{4.54}$$

where G_k are the columns of G in Notation 4.5.2. Then (4.53) is a system of Riccati-equations, also referred to as a vector-Riccati equation.

Proof of proposition (4.5.4). Define

$$\mathbf{g}'(t) = (g_0'(t), g_1'(t), g_2'(t), g_{11}'(t), g_{22}'(t), g_{12}'(t))^T$$

to be the derivative column-vector of $\mathbf{g}(t) \in \mathbb{R}^6$ and $G = \mathbf{g}(t)\mathbf{g}^T(t)$ to be the 6×6-matrix defined in Notation 4.5.2. By noting the notation in $(1')$-$(6')$, we recognize the coefficients to the linear terms as the elements c_{mn} for $m, n = 1, \ldots, 6$ of the matrix C. Similarly the constants in (1')-(6') denote the elements N_k for $k = 1, \ldots, 6$ of the 6-dimensional vector $\mathbf{N}$. We obtain the following notation for $\mathbf{g}'(t)$

$$\mathbf{g}'(t) = \mathbf{N} + C\mathbf{g}(t) + \mathbf{g}(t)^-$$

where $\mathbf{g}(t)^-$ is the vector of the quadratic terms of $\mathbf{g}^{'}(t)$ defined above. We need to find the operator which maps the matrix G in $\mathbb{R}^{6\times 6}$ to the vector $\mathbf{g}(t)^- \in \mathbb{R}^6$.

Let $e_i = (0,\dots,1,\dots,0)$ denote the unit vector in $\mathbb{R}^6$, i.e. e_i is the $i-th$ vector of the canonical basis of $\mathbb{R}^6$. Then e_i^T is the 6-dimensional row-vector where all elements of e_i^T are 0 except for the i-th element. Further, let G_k denote the k-th column of G. We need to extract exactly those elements from G such that, for instance

$$g_0(t)^- = -\epsilon_1 g_1^2(t) - \epsilon_2 g_2^2(t) - \psi g_1(t)g_2(t)$$

By $e_1^T G_2$ we get $g_1^2(t)$, and by $e_2^T G_3$ and $e_3^T G_3$ we get $g_2^2(t)$ and $g_1(t)g_2(t)$, respectively. Hence

$$g_0(t)^- = -e_2^T(G_2\epsilon_1 + G_3\psi) - e_3^T G_3\epsilon_2$$

By following the same manner for $g_1^{'}(t), g_2^{'}(t), \dots$, we obtain the requested result

$$\mathbf{g}(t)^- = \begin{bmatrix} -e_2^T(G_2\epsilon + G_3\psi) - e_3^T G_3\epsilon_2 \\ -e_2^T(4G_4\epsilon_1 + G_6\psi) - e_3^T(2G_4\psi + 2G_6\epsilon_2) \\ -e_2^T(G_6\psi + 4G_5\epsilon_2) - e_3^T(2G_6\epsilon_1 + G_5\psi) \\ -e_4^T(4G_4\epsilon_1 + 2G_6\psi) - e_6^T G_6\epsilon_2 \\ -e_5^T(4G_5\epsilon_2 + 2G_6\psi) - e_6^T G_6\epsilon_1 \\ -4e_4^T G_5\epsilon_1 - e_6^T(G_3\psi + G_4\psi + 4G_5\epsilon_2) \end{bmatrix}$$

By Definition 4.5.1, it is easily seen that (4.53) is a system of Riccati equations if the functions $g_0(t),\dots,g_{12}(t)$ are continuous (and differentiable) on $[0,T]$. The linearity of $L(G)$ is easily shown by noting that the sum of two $n\times n$ matrices A and B, is a new $n\times n$ matrix C and that the sum of two n-dimensional column-vectors $A_1 + B_1$ is a new column vector C_1. Take for instance the first row in $L(G)$, denoted by $L_1(G)$, and note that

$$\begin{aligned} L_1(A+B) =& -e_2^T((A_2+B_2)\epsilon_1 + (A_3+B_3)\psi) - e_3^T(A_3+B_3)\epsilon_2 \\ =& -e_2^T(A_2\epsilon_1 + B_2\epsilon_1 + A_3\psi + B_3\psi)) - e_3^T(A_3\epsilon_2 + B_3\epsilon_2) \\ =& -e_2^T(A_2\epsilon_1 + A_3\psi) - e_3^T A_3\epsilon_2 - e_2^T(B_2\epsilon_2 + B_3\psi) - e_3^T B_3\epsilon_2 \\ =& L_1(A) + L_1(B) \end{aligned}$$

By same argument for the rest of the rows of G, and by similar argument for $L(cG)$ for any scalar c, we conclude that $L(G)$ is a linear transformation from $\mathbb{R}^{6\times 6} \to \mathbb{R}^6$. ∎

Alternative Representation

An alternative approach would be to represent the system (1′)-(6′) as

$$\mathbf{g}'(t) = \mathbf{N} + C\mathbf{g}(t) + M(\mathbf{g}(t))\cdot\mathbf{g}(t) \tag{4.55}$$

where $\mathbf{g}(t), \mathbf{N} \in \mathbb{R}^6$ and the matrix C are as in Notation 4.5.2. $M(\mathbf{g}(t))$ is a 6×6-matrix where the elements are dependent on the unknown functions $g_0(t),\dots,g_{12}(t)$.

Proposition 4.5.5 (Alternative Riccati-representation).
Define $\boldsymbol{g}(t) \in \mathbb{R}^6$ to be the 6-dimensional column vector (4.52) and $\boldsymbol{g}(t)^T$ the transpose. Then the system of equations $(1')$-$(6')$ can be represented by

$$\boldsymbol{g}'(t) = \boldsymbol{N} + C\boldsymbol{g}(t) + M(\boldsymbol{g}(t)) \cdot \boldsymbol{g}(t) \tag{4.56}$$

where .

$$M(\boldsymbol{g}(t)) = \begin{bmatrix} 0 & -g_1\epsilon_1 & -(g_2\epsilon_2+g_1\psi) & 0 & 0 & 0 \\ 0 & -4g_{11}\epsilon_1 & -2g_{12}(t)\epsilon_2 & -2g_2\psi & 0 & -g_1\psi \\ 0 & -2g_{22}\psi & -g_{12}\psi & 0 & -4g_2\epsilon_2 & -2g_1\epsilon_1 \\ 0 & 0 & 0 & -4g_{11}\epsilon_1 & 0 & -(g_{12}\epsilon_2-2g_{11}\psi) \\ 0 & 0 & 0 & 0 & -4g_{22}\epsilon_2 & -(g_{12}\epsilon_1+2g_{22}\psi) \\ 0 & 0 & 0 & -2g_{12}\epsilon_1 & -(4g_{11}\psi+4g_{12}\epsilon_2) & -g_{12}\psi \end{bmatrix}$$

is a 6×6-matrix where the elements are linear terms of the functions $g_0(t), \ldots, g_{12}(t)$, C is the 6×6-matrix of the coefficients of the linear terms in $(1')$-$(6')$ and $\boldsymbol{N} \in \mathbb{R}^6$ the vector of constants, both given in Proposition 4.5.4. Then (4.56) is a vector-Riccati equation.

Proof of proposition (4.5.5). Define again, as in the preceding subsection, $\mathbf{g}(t)^-$ to be the 6-dimensional vector of the non-linear terms of $\mathbf{g}'(t)$, and let $M(\mathbf{g}(t))$ be defined as in Proposition 4.5.5. Following, by matrix-vector multiplication, we see that $M(\mathbf{g}(t)) \cdot \mathbf{g}^T(t) = \mathbf{g}(t)^-$. If we let $\mathbf{N}$ and C be as in Notation 4.5.2, then

$$\mathbf{g}'(t) = C\mathbf{g}(t) + M(\mathbf{g}(t)) \cdot \mathbf{g}(t) + N$$

By Definition 4.5.1, it is easily seen that each of the equations in $(1')$-$(6')$ is a Riccati equation as long as $g_0(t), \ldots, g_{12}(t)$ are continuous on $[0, T]$. ■

A solution to the two representations of the Riccati equations above, is a set of continuously differentiable functions $g_0(t)$, $g_1(t)$, $g_2(t)$, $g_{11}(t)$, $g_{22}(t)$, $g_{12}(t)$ that satisfy all six equations simultaneously over the interval $[0, T]$. Note that $g(t, z_1, z_2)$ is a polynomial of second degree in z_1, z_2, hence $f \in C^{2,2}(D)$ w.r.t. z_1, z_2. If there exists a continuously differentiable solution $\mathbf{g}(t)$ to (1)-(6), which also satisfies the boundary value condition (4.49), we have successfully found a function $f(t, z_1, z_2)$ for which $\Phi(t, w, z_1, z_2) \in C^{1,2,2,2}(D)$ is a classical solution to the HJB-equation stated in (4.14) and the boundary value problem in Problem 4.4.3.

The initial stochastic optimal control problem has been deduced to a set of Riccati equations, however it is highly questionable whether there exists a set of solutions at all. It might also very well be that the functions explode for some t in $[0, T]$, due to for instance singularities.

4.6 Suggestions to Solutions of the Riccati ODEs

Explicit analytical solutions to general Riccati equations are rare. Finding a set of analytical solutions, if they even exist, to six Riccati equations, is an even more elaborate task.

Numerical methods for Riccati equations and algebraic Riccati equations are studied in numerous articles. See for instance [MH11], for an application of the Legendre wavelet method for solving a single Riccati equation, and for a comparison to other existing methods. The Euler-method or fourth order Runge Kutta method (RK-4), are widely used numerical approximations to solutions of both linear and non-linear differential equations. The RK-4 method is computationally effective and simple to implement, and is proven to give accurate approximations when compared to existing analytical solutions [FA16].

The algebraic Riccati equations often arise in linear quadratic optimal control problems and filtering theory. Newton's method and the Sign Function method, are two among many possible computational solutions. [Bun96] gives a survey of some methods for algebraic Riccati equations explored over the past three decades. However there is, to my knowledge, less work done on numerical solutions for systems of Riccati equations, or the so-called vector Riccati equation. By deriving vector-adjusted algebraic Riccati equations, one could investigate if there is possibility of exploiting the numerical methods for matrix Riccati equations adjusted to vector-form.

However, some computational methods for systems of non-linear differential equations, like the Differential Transform Method (DTM), the He Laplace Method or the Adomian Decomposition Method, have been applied to systems of Riccati equations. [SP17] concludes that among the three mentioned, the DT method, initially introduced by J.K. Zhou in 1986, is the most efficient when solving Riccati equations, both computationally and in terms of "errors " when compared to (a few) analytical solutions. A short suggestion to how the method can be applied in the case of a system of Riccati equations, based on [Mir11], is presented in the following section.

Differential Transform Method

The DTM is a numerical method for solving systems of non-linear differential equations, by transforming the equations into converging series. The method is closely related to the Taylor series expansion, but the derivatives are not found symbolically. Following [Mir11], let $Y(k)$ denote the transformation of the k-th derivative of the unknown function $y(x)$, where

$$Y(k) = \frac{1}{k!}\Big[\frac{d^k(y(x))}{dx^k}\Big]_{x=x_0} \tag{4.57}$$

The inverse differential transform of $Y(k)$ is defined by

$$y(x) = \sum_{k=0}^{\infty} Y(k)x^k \tag{4.58}$$

Combining (4.57) and (4.58), the relation to the Taylor series expansion becomes clear, and the unknown functions are expressed by

$$y(x) = \sum_{k=0}^{\infty} \frac{(x-x_0)^k}{k!}\frac{d^k y(x)}{dx^k}_{|x=x_0} \tag{4.59}$$

Then a set of theorems is required to transform the initial set of differential equations, in our case the system $(1')$-$(6')$, to the form of the converging series. In the case where the non-linear differential equations are of Riccati type, the following 4 theorems are needed, see [Mir11] for complete list of theorems,

Theorem 4.6.1 (Theorem 1).
If $y(x) = y_1(x) \pm y_2(x)$, *then* $Y(k) = Y_1(k) \pm Y_2(k)$.

Theorem 4.6.2 (Theorem 2).
If $y(x) = ay_1(x)$, *then* $Y(k) = aY_1(k)$.

Theorem 4.6.3 (Theorem 3).
If $y(x) = \dfrac{dy_1^n(x)}{dx^n}$, *then* $Y(k) = \dfrac{(k+n)!}{k!}Y_1(k+n)$.

Theorem 4.6.4 (Theorem 4).
If $y(x) = y_1(x)y_2(x)$, *then* $Y(k) = \sum_{k_1=0}^{k} Y_1(k_1)Y_2(k-k_1)$.

[Mir11] refers to [AO05] for proof of Theorems 1 to 4. By applying the DT method to the system of Riccati equations, each derivative of $g(t)$ can be expressed by a sum of functions $Y_i(k)$. For instance, equation (1'), for $g_0'(t)$ can be transformed into

$$\begin{aligned}(k+1)Y_0(k+1) = &-N_1 - c_{12}Y_1(k) - c_{13}Y_2(k) + c_{14}Y_{11}(k) - c_{15}Y_{22}(k)\\ &- c_{16}Y_{12}(k) - \sum_{k_1=0}^{k} Y_1(k_1)Y_1(k-k_1)\epsilon_1 - 2\sum_{k_1=0}^{k} Y_2(k_1)Y_2(k-k_1)\epsilon_2\\ &- \sum_{k_1=0}^{k} Y_1(k_1)Y_2(k-k_1)\psi\end{aligned}$$

By doing so for the remaining 5 equations, in correspondence with an initial value for each function, we obtain the numeric values of the transformed derivatives of the unknown functions. By applying (4.57) and (4.58), the derivatives are expresses by series converging to the true value,

$$y_i(x) = \sum_{k=0}^{\infty} Y_i(k)\frac{(x-x_0)^k}{k!} \tag{4.60}$$

for $i = 1, \dots, 6$. Note that the method requires the initial values of $Y_i(0) = y_i(0)$ to be known. However, for the system $(1') - (6')$ we have deduced the boundary value for $t = T$ (4.49), and by a change of variable for $v(s) = g(T-t)$, we can turn the boundary value problem into an initial value problem.

4.7 Verification

By "educated" guessing on solutions, in Section 5.3 we turned the initial non-linear boundary value problem (4.4.3) into a system of non-linear ordinary differential equations, and in Section 4.5 we argued that $(1')$-$(6')$ constitutes a system of Riccati-equations expressed in two alternative ways. In this section we verify that our solution for the value function Φ is a classical solution to the optimal control problem by means of Theorem 4.3.1, assuming that continuously differentiable solutions to the Riccati-equations exist. We also verify that the controls obtained in Proposition 4.4.1 are indeed optimal admissible controls by Theorem 4.3.1.

Theorem 4.7.1 (Semi-explicit solution).
Assume $g(t)$ *is continuously differentiable, such that* $g_0(t)$, $g_1(t)$, $g_2(t)$, $g_{11}(t)$,

$g_{22}(t)$, $g_{12}(t)$ are solutions to the system of Riccati equation in Result 4.4.6. Furthermore, assume that

$$8(2g_{11}(t) + g_{12}(t)) - \frac{1}{2\widetilde{\sigma}_1^2(t)} < 0 \tag{4.61}$$

and

$$8(2g_{22}(t) - g_{12}(t)) - \frac{1}{2\widetilde{\sigma}_2^2(t)} < 0 \tag{4.62}$$

where $\widetilde{\sigma}_i^2 =$Var(Z_i). Then the value function Φ of the optimal control problem stated in Theorem 4.2.7 is given by

$$\Phi(t, w, z_1, z_1) = f(t, z_1, z_2)w^\gamma, \quad 0 < \gamma < 1 \tag{4.63}$$

for any all $t < T$ and $w > 0$, and

$$\Phi(t, w, z_1, z_2) = w^\gamma, \quad \text{for } t = T, w = 0 \tag{4.64}$$

where $f(t, z_1, z_2)$ is on the form

$$f(t, z_1, z_2) = e^{g(t,z_1,z_2)}$$

for

$$g(t, z_1, z_2) = g_0(t) + g_1(t)z_1 + g_2(t)z_2 + g_{11}(t)z_1^2 + g_{22}(t)z_2^2 + g_{12}(t)z_1z_2$$

The optimal allocation of wealth $u_i^*(t, Z_1(t), Z_2(2))$ for $i = 1, 2$ is given by

$$\begin{aligned} u_1^*(t, z_1, z_2) =& \frac{1}{(1-\gamma)\Gamma}\Big[G_1(z_1, z_2) - M_1\Big(g_1(t) + 2g_{11}(t)z_1 + g_{12}(t)z_2\Big) \\ & - N_1\Big(g_2(t) + 2g_{22}(t)z_2 + g_{12}(t)z_1\Big)\Big] \end{aligned} \tag{4.65}$$

$$\begin{aligned} u_2^*(t, z_1, z_2) =& \frac{1}{(1-\gamma)\Gamma}\Big[G_2(z_1, z_2) - M_2\Big(g_1(t) + 2g_{11}(t)z_1 + g_{12}(t)z_2\Big) \\ & - N_2\Big(g_2(t) + 2g_{22}(t)z_2 + g_{12}(t)z_1\Big)\Big] \end{aligned} \tag{4.66}$$

where $G_i(z_1, z_1)$ is a polynomial in $Z_1(t), Z_2(t)$ of degree 1, and Γ, M_i, N_i, $i = 1, 2$ are constants defined in Appendix A.

Before we proceed with the proof of Theorem 4.7.1, we need the following two lemmas.

Lemma 4.7.2.
Assume

$$8\,(2g_{11}(t) + g_{12}(t)) - \frac{1}{2\widetilde{\sigma}_1^2(t)} < 0 \tag{4.67}$$

and

$$8\,(2g_{22}(t) - g_{12}(t)) - \frac{1}{2\widetilde{\sigma}_2^2(t)} < 0 \tag{4.68}$$

Then

$$\int_{t_0}^{T} E\Big[\Big(\sigma\,(u_1^*(t) + u_2^*(t))\,W(t)\Phi_w\Big)^2\Big]ds < \infty \tag{4.69}$$

Proof of lemma (4.7.2). Recall from Chapter 2.3 that $Z_i(t) \sim \mathcal{N}(\widetilde{\mu}_i(t), \widetilde{\sigma}_i^2(t))$, where $\widetilde{\mu}_i(t) = z_i e^{A_i t}$ and $\widetilde{\sigma}_i^2(t) = \dfrac{\sigma_i^2}{2A_i}(e^{2A_i t} - 1)$. By Hölder's inequality, Proposition B.3.6, we obtain that

$$\begin{aligned}
&\int_{t_0}^{T} E\Big[\Big(\sigma(u_1^*(t) + u_2^*(t))W(t)\Phi_w\Big)^2\Big]ds \\
&\quad = \gamma^2\sigma^2 \int_{t_0}^{T} E\Big[\Big((u_1^*(t) + u_2^*(t))^2 W^{2\gamma}(t) f^2(t, z_1, z_2)\Big)\Big]ds \\
&\quad \le \gamma^2\sigma^2 \int_{t_0}^{T} E\Big[(u_1^*(t) + u_2^*(t))^4 W^{4\gamma}(t)\Big]^{1/2} E\Big[f^4(t, z_1, z_2)\Big]^{1/2} ds \qquad (4.70)
\end{aligned}$$

The $(u_1^*(t) + u_2^*(t))$-term in the expectation to the left imposes no difficulties. Since $u_i^*(t)$ is linear in $Z_i(t)$, where $Z_i(t)$ is the stationary Ornstein-Uhlenbeck process, i.e. a Gaussian random variable for each t. By the properties of a normal random variable, all moments of $Z_i(t)$ are finite, i.e. $E[Z_i^n(t)] < \infty$ for all n. Especially, $E[Z_i(t)] = \widetilde{\mu}$ and $E[Z_i^2(t)] = \widetilde{\sigma}^2 + \widetilde{\mu}^2$. Following, since $Z_i(t)$ is Gaussian, it has existing, i.e. finite, exponential moments as well, and

$$\begin{aligned}
E[e^{Z_i(t)}] &= E\Big[e^{z_i e^{A_i t} + \sigma_i \int_0^t e^{A_i(t-s)} dB_i(s)}\Big] \\
&= e^{z_i e^{A_i t}} E\Big[e^{\sigma_i \int_0^t e^{A_i(t-s)} dB_i(s)}\Big] \\
&= e^{z_i e^{A_i t}} e^{\frac{1}{2}(\frac{\sigma_i^2}{2A_i}(e^{A_i t} - 1))} < \infty
\end{aligned}$$

Since $W^{4\gamma}(t)$ is the exponential of linear terms of $Z_i(t)$, it follows that $E[W^{4\gamma}(t)] < \infty$. Hence we argue that $E[(u_1^*(t) + u_2^*(t))^4 W^{4\gamma}(t)] < \infty$. We have left to show that this also holds for the second expectation in (4.70). Note that

$$\begin{aligned}
f^4(t, z_1, z_2) = \exp\Big(4\Big(&g_0(t) + g_1(t)Z_1(t) + g_2(t)Z_2(t) + g_{11}(t)Z_1^2(t) \\
&+ g_{22}(t)Z_2^2(t) + g_{12}(t)Z_1(t)Z_2(t)\Big)\Big)
\end{aligned}$$

by definition of f. Furthermore, the following inequality

$$0 \le (x - y)^2 \le x^2 - 2xy + y^2 \implies 2xy \le x^2 + y^2 \qquad (4.71)$$

implies that

$$2g_{12}(t)Z_1(t)Z_2(t) \le g_{12}(t)\Big(Z_1^2(t) + Z_2^2(t)\Big)$$

Since e^x is strictly increasing for $x > 0$,

$$\begin{aligned}
f^4(t, z_1, z_2) \le \exp\Big(&4g_0(t) + 4g_1(t)Z_1(t) + 4g_2(t)Z_2(t) + 2(2g_{11}(t) \\
&+ g_{12}(t))Z_1^2(t) + 2(2g_{22}(t) - g_{12}(t)\Big)Z_2^2(t))
\end{aligned}$$

By applying Hölder's inequality again, we obtain that

$$E\Big[\exp\Big(4g_0(t) + 4g_1(t)Z_1(t) + 4g_2(t)Z_2(t) + 2(2g_{11}(t) + g_{12}(t))Z_1^2(t)$$

$$
\begin{aligned}
&+ 2(2g_{22}(t) - g_{12}(t))Z_2^2(t)\Big)\Big]^{1/2} \\
\leq & E\Big[\exp\Big(4\Big(g_0(t) + g_1(t)Z_1(t) + g_2 Z_2(t)\Big)\Big)^2\Big]^{1/4} \\
&\times E\Big[\Big(\exp\Big(2\Big(2g_{11}(t) + g_{12}(t)\Big)Z_1^2(t) + 2\Big(2g_{22}(t) - g_{12}(t)\Big)Z_2^2(t)\Big)\Big)^2\Big]^{1/4} \\
= & E\Big[\exp\Big(8\Big(g_0(t) + g_1(t)Z_1(t) + g_2(t)Z_2(t)\Big)\Big)\Big]^{1/4} \\
&\times E\Big[\exp\Big(4(2g_{11}(t) + g_{12}(t))Z_1^2(t) + 4(2g_{22}(t) - g_{12}(t))Z_2^2(t))\Big)\Big]^{1/4} \qquad (4.72)
\end{aligned}
$$

Note once more that the first expectation in (4.72) is bounded by finite exponential moments of $Z_i(t)$, by the same argument as stated above. Hence we only need to prove that the second expectation is bounded. By applying Hölder's inequality once more, we obtain that

$$
\begin{aligned}
&E\Big[\exp\Big(4(2g_{11}(t) + g_{12}(t))Z_1^2(t)) + 4(2g_{22}(t) - g_{12}(t))Z_2^2(t)\Big)\Big]^{1/2} \\
\leq & E\Big[\exp\Big(8(2g_{11}(t) + g_{12}(t))Z_1^2(t)\Big)\Big]^{1/8} \times E\Big[\exp\Big(8(2g_{22}(t) - g_{12}(t))Z_2^2(t)\Big)\Big]^{1/8}
\end{aligned}
$$

Now let $c(t) = 8(2g_{11}(t) + g_{12}(t))$. Then for each t, $Z = Z_i(t) \sim \mathcal{N}(\widetilde{\mu}, \widetilde{\sigma})$ is a random variable, and the first expectation equals

$$
E\Big[\exp\Big(c(t)Z_1^2(t)\Big)\Big] = \frac{1}{\sqrt{2\pi\widetilde{\sigma}^2}} \int_{\mathbb{R}} \exp\Big(c(t)z^2\Big) \exp\Big(\frac{-(z-\widetilde{\mu})^2}{2\widetilde{\sigma}^2}\Big) dz \qquad (4.73)
$$

Let $x = \dfrac{z - \widetilde{\mu}}{\widetilde{\sigma}} \sim \mathcal{N}(0,1)$, then $z = \widetilde{\sigma}x + \widetilde{\mu}$, and by a change of variable (4.73) becomes

$$
\begin{aligned}
&\frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \exp\Big(c(t)(x\widetilde{\sigma} + \widetilde{\mu})^2\Big) \exp\Big(\frac{-x^2}{2}\Big) dx \\
&\qquad = \frac{1}{\sqrt{2\pi}} \exp\Big(c(t)\widetilde{\mu}^2\Big) \int_{\mathbb{R}} \exp\Big(c(t)(x^2\widetilde{\sigma}^2 + 2x\widetilde{\sigma}\widetilde{\mu} - x^2\Big) dx \qquad (4.74)
\end{aligned}
$$

Completing the square by

$$
c(t)(x^2\widetilde{\sigma}^2 + 2x\widetilde{\sigma}\widetilde{\mu}) - \frac{1}{2}x^2 = \frac{1}{2}\Big((2c(t)\widetilde{\sigma}^2 - 1)\Big(x - \frac{\widetilde{\sigma}\widetilde{\mu}c(t)}{2c(t)\widetilde{\sigma}^2 - 1}\Big)^2 - \frac{\widetilde{\sigma}^2\widetilde{\mu}^2c^2(t)}{2c(t)\widetilde{\sigma}^2 - 1}\Big)
$$

We obtain that (4.74) equals

$$
\begin{aligned}
&\frac{1}{\sqrt{2\pi}} \exp\Big(c(t)\widetilde{\mu}^2 - \frac{\widetilde{\sigma}^2\widetilde{\mu}^2c^2(t)}{2(2c(t)\widetilde{\sigma}^2 - 1)}\Big) \\
&\qquad\qquad \times \int_{\mathbb{R}} \exp\Big(\frac{1}{2}\Big(2c(t)\widetilde{\sigma}^2 - 1\Big)\Big(x - \frac{\widetilde{\sigma}\widetilde{\mu}c(t)}{2c(t)\widetilde{\sigma}^2 - 1}\Big)^2\Big) dx
\end{aligned}
$$

Making another change of variable, for $y = x - \dfrac{\widetilde{\sigma}\widetilde{\mu}c(t)}{2c(t)\widetilde{\sigma}^2 - 1}$, $dy = dx$, we obtain

$$
\frac{1}{\sqrt{2\pi}} \exp\Big(c(t)\widetilde{\mu}^2 - \frac{\widetilde{\sigma}^2\widetilde{\mu}^2c^2(t)}{2(2c(t)\widetilde{\sigma}^2 - 1)}\Big) \int_{\mathbb{R}} \exp\Big(\frac{1}{2}\Big(2c(t)\widetilde{\sigma}^2 - 1\Big)y^2\Big) dy < \infty
$$

for

$$2c(t)\widetilde{\sigma}^2 - 1 < 0 \implies c(t) - \frac{1}{2\widetilde{\sigma}^2} < 0$$

by properties of the Gaussian function. Hence

$$E\Big[\exp\Big(8(2g_{11}(t)+g_{12}(t))Z_1^2(t)\Big)\Big]^{1/8} < \infty$$

as long as $8\,(2g_{11}(t) + g_{12}(t)) - \frac{1}{2\widetilde{\sigma}_1^2} < 0$. By same reasoning, the second expectation is finite as long as

$$8(2g_{22}(t) - g_{12}(t)) - \frac{1}{2\widetilde{\sigma}_2^2(t)} < 0 \tag{4.75}$$

This completes the proof. ■

Lemma 4.7.3.
Assume conditions (4.67) and (4.68) from Lemma 4.7.2, then

$$\int_{t_0}^{T} E\Big[\Big(\sigma_i u_i^*(t)W(t)\Phi_w + \sigma_i \Phi_{z_i}\Big)^2\Big] ds < \infty \tag{4.76}$$

for $i = 1, 2$.

Note the difference between σ_i and $\widetilde{\sigma}_i$, where σ_i is the diffusion coefficient of the Ornstein-Uhlenbeck dynamics, and $\widetilde{\sigma}_i = \mathrm{Var}(Z_i)$.

Proof of lemma (4.7.3). By inequality (4.71)

$$\begin{aligned}
&\int_{t_0}^{T} E\Big[\Big(\sigma_i u_i^*(t)W(t)\Phi_w + \sigma_i \Phi_{z_i}\Big)^2\Big] ds \\
&\quad = \int_{t_0}^{T} E\Big[\Big(\sigma_i u_i^*(t)W^\gamma(t) f(t, z_1, z_2) + \sigma_i W^\gamma(t) f_{z_i}\Big)^2\Big] ds \\
&\quad \leq \sigma_i^2 \int_{t_0}^{T} 2E\Big[\Big(u_i^*(t)W^\gamma(t) f(t, z_1, z_2)\Big)^2\Big] ds + \sigma_i^2 \int_{t_0}^{T} 2E\Big[\Big(W^\gamma(t) f_{z_i}\Big)^2\Big] ds
\end{aligned}$$

From the proof of lemma (4.7.2), we obtain that the first expectation on the right hand side of the equation above is bounded. For the second expectation, note that

$$f_{z_i} = 2g_{ii}(t)Z_i(t) + g_{12}(t)Z_j(t), \quad i \neq j \tag{4.77}$$

i.e., f_{z_i} is linear in Z_i, $i = 1, 2$. Since $Z_i(t)$ has finite moments and finie exponential moments, the same reasoning follows for the second expectation as well, and hence

$$\int_{t_0}^{T} E\Big[\Big(\sigma_i u_i^*(t)W^\gamma(t) f(t, z_1, z_2) + \sigma_i W^\gamma(t) f_{z_i}\Big)^2\Big] ds < \infty \tag{4.78}$$

This completes the proof. ■

We are now ready to state the proof of theorem Theorem 4.7.1.

Proof of theorem (4.7.1). Note that $u_i^*(t)$ only depends on the state variables $Z_i(t)$ for $i = 1, 2$. Since $Z_i(t)$ is defined as a stationary Ornstein-Uhlenbeck process, $u_i^*(t)$ will be measurable and $\mathcal{F}_t$-adapted for all $t \in [0, T]$, and by construction, such that $W(t, u_1(t), u_2(t))$ is the unique solution to (4.2.4). Since $u_i^*(t)$ is linear in $Z_i(t)$, $i = 1, 2$, which is Gaussian with finite first and second moments, $E[u_i^*(t)] < \infty$, $E[(u_i^*)^2(t)] < \infty$, hence $u_i^*(t) \in L^2([0,T] \times \mathbb{R}^3)$. In conclusion, by Definition 4.2.1, $u_i^*(t)$ are admissible controls for $i = 1, 2$.

$\Phi \in C^2(D)$ is obvious w.r.t. w. Assuming a continuously differentiable solution to the system of equations (1)-(6) exists, $f(t, z_1, z_2)$ will be continuously differentiable w.r.t. t. Since $g(t, z_1, z_2)$ is a second degree polynomial in z_1, z_2, g will be twice continuously differentiable w.r.t. z_1, z_2. Hence $\Phi \in C^{1,2,2,2}(D)$. For $t = T$ or $W(t) = 0$, the boundary value (4.41) is satisfied. By $\mathcal{P}$-a.s. continuous sample paths of $\widetilde{B}(t)$ and $B_i(t)$, and since $f(t, z_1, z_2)$ by assumption is continuous for all $t \in [0, T]$,

$$\lim_{t \to \tau_D} \Phi(W(t)) = U(W(\tau_D)) \tag{4.79}$$

hence, $\Phi \in C(\bar{D})$. By lemma (4.7.2) and (4.7.3), for

$$8(2g_{11}(t) + g_{12}(t)) - \frac{1}{2\widetilde{\sigma}_1^2(t)} < 0 \tag{4.80}$$

and

$$8(2g_{22}(t) - g_{12}(t)) - \frac{1}{2\widetilde{\sigma}_2^2(t)} < 0 \tag{4.81}$$

conditions (4.24) and (4.25) in Theorem 4.3.1 are satisfied, i.e.
$\sigma(t, W(t), Z_1(t), Z_2(t))$ and $\sigma_i(t, W(t), Z_1(t), Z_2(t))$ are Itô-integrable functions in $L^2([0,T] \times \mathbb{R}^3)$. Hence, if our assumptions on $g(t, z_1, z_2)$ hold, by Theorem 4.3.1, $\Phi(t, w, z_1, z_2)$ must be a classical solution to the HJB-equation given in Problem 4.4.3 for any $(t, w, z_1, z_2) \in D$. ∎

Note that, in contradiction to Merton's two-asset problem given in Example 3.5.2, where the optimal allocation of wealth is to hold a constant fraction in the risky asset, the optimal controls in this case are stochastic and time-dependent, i.e. the fraction of wealth invested in each of the risky assets, should optimally at each time point t, be updated according to the evolution of the underlying stationary processes $Z_1(t)$ and $Z_2(t)$. However, out in the real market, this would lead to huge transactions costs every time a stock is bought or sold. Hence our solution is probably not optimal in the case where transaction costs are included. Finally, note that the fraction of wealth invested in asset 1, doesn't only depend on the underlying process $Z_1(t)$, it depends on the evolution of the stationary process driving the price of asset 2 as well, and vice versa. Due to the co-integrated market model, the optimal allocations for a particular asset depends on the price of the other asset as well.

4.8 Some Remarks on the Optimal Controls

The optimal controls from Theorem 4.7.1 are stochastic processes, being combinations of $Z_1(t)$ and $Z_2(t)$ with mean-value given by

$$E[u_1^*(t)] = \frac{1}{(1-\gamma)\Sigma}\Big[K_1 - M_1 g_1(t) - N_1 g_2(t) - E[Z_1(t)]\Big(2M_1 g_{11}(t)$$

$$+ N_1 g_2(t) + \Sigma_2 A_1\Big) - E[Z_2(t)]\Big(M_1 g_{12}(t) + 2N_1 g_{22}(t) - \Sigma A_2\Big)\Big]$$

where $K_1 = (r-\mu)(\Sigma_2 - \Sigma) - \frac{1}{2}\Sigma_2(\Sigma_1 - \Sigma))$

$$E[u_2(t)] = \frac{1}{(1-\gamma)\Sigma}\Big[K_2 - M_2 g_1(t) - N_2 g_2(t) - E[Z_1(t)]\Big(2M_2 g_{11}(t) + N_2 g_{12}(t) - \Sigma A_1\Big) - E[Z_2(t)]\Big(M_2 g_{12}(t) + 2N_2 g_{22}(t) + \Sigma_1 A_2\Big)\Big]$$

where $K_2 = (r-\mu)(\Sigma_1 - \Sigma) - \frac{1}{2}\Sigma_1(\Sigma_2 - \Sigma)$, and variance given by

$$\begin{aligned}\mathrm{Var}(u_1^*(t)) = \frac{1}{(1-\gamma)^2\Gamma^2}\Big[&(2M_1 g_{11}(t) + N_1 g_{12}(t) + \Sigma_2 A_1)^2 \mathrm{Var}(Z_1(t)) \\ &+(M_1 g_{12}(t) + 2N_1 g_{22}(t) - \Sigma A_2)^2 \mathrm{Var}(Z_2(t)) - (2M_1 g_{11}(t) + N_1 g_{12}(t) \\ &+\Sigma_2 A_1)(M_1 g_{12}(t) + 2N_1 g_{22}(t) - \Sigma A_2)\mathrm{Cov}(Z_1(t), Z_2(t))\Big]\end{aligned}$$

$$\begin{aligned}\mathrm{Var}(u_2^*(t)) = \frac{1}{(1-\gamma)^2\Gamma^2}\Big[&(2M_2 g_{11}(t) + N_2 g_{12}(t) - \Sigma A_1)^2 \mathrm{Var}(Z_1(t)) \\ &+(M_2 g_{12}(t) + 2N_2 g_{22}(t) + \Sigma_1 A_2)^2 \mathrm{Var}(Z_2(t)) - (2M_2 g_{11}(t) + N_2 g_{12}(t) \\ &-\Sigma A_1)(M_2 g_{12}(t) + 2N_2 g_{22}(t) + \Sigma A_2)\mathrm{Cov}(Z_1(t), Z_2(t))\Big]\end{aligned}$$

Due to the unfamiliar form of $g(t)$, i.e. the set of solutions to the the system of Riccati equations, it is difficult to make any conclusions on the asymptotic form of $u_i(t)$ as the time horizon is expanded. Depending on $g(t)$, $u_i(t)$ could be stationary, being the linear combination of Ornstein-Uhlenbeck processes, if the mean-reverting effect is stronger than the possibly divergence of $g(t)$. Then one could expect fluctuations on both sides of the mean, and simply allocate an amount of wealth in each of the assets equal to the long term mean. From a financial point of view, if the stationary processes have a strong positive correlation or are very positively correlated to the common stationary driver, it seems unwise to invest large amounts of wealth in both risky assets. This would yield high returns in the best case, but even small negative fluctuations would cause a large decrease in wealth. Then perhaps, one could invest a larger fraction in the safe investment (or diversify to other market segments). A strong negative correlation between the stationary drivers could be neutralized by investing equal parts in each asset, assuming the assets yield somewhat equal returns. Nevertheless, by introducing a co-integrated market model with asset prices driven by correlated stationary Ornstein-Uhlenbeck processes and a common non-stationary drifted Brownian motion, the control problem has a much more complex form, yielding an optimal portfolio allocation quite different from Merton's simpler case. In the co-integrated asset case, the optimal allocation is time-dependent and highly dependent on the correlations between the processes and the volatility of each Brownian motion, see the constants in Appendix A.

4.9 A Reduction of Noise

In Section 4.4, the HJB-equation of the stochastic control problem in a co-integrated asset market model, was presented. Notice the difference from the

HJB-equation presented in Section 3.5. For the traditional Merton problem with one risky asset, the HJB-equation is reduced to an ordinary differential equation, to which an explicit solution is known. In contrast, when considering multiple (and correlated) state variables, even after reduction of one of the states, w, the boundary value problem is a non-linear partial differential equation of second order! The complex form of the boundary value problem, and hence the procedure to obtain a solution of the value function, derives from the dynamics of the co-integrated asset prices $S_1(t)$ and $S_2(t)$, through the generator of the wealth process $W(t)$. The control problem presented in the beginning of this chapter must be optimized w.r.t. three sources of uncertainty; $\widetilde{B}(t), B_1(t)$ and $B_2(t)$. Even when removing one "noisy" source, the non-stationary drifted Brownian motion $X(t)$, the reduced boundary value problem is still non-linear. This section works as an example to illustrate some of the (possible) reason to why the stochastic control problem in this chapter resulted in a 6-dimensional system of Riccati equaitons. This section briefly presents a simpler case where the non-stationary "noise" is eliminated, and at last a representation of the solution to the model where $B_1(t)$ and $B_2(t)$ are assumed to be independent, is found. The procedure and technique is the same as in Section 4.4, hence the intermediate steps are omitted and only the resulting equations and expressions are stated.

Let $S_i(t)$ denote the price of risky asset $i = 1, 2$ at time t, given by

$$S_i(t) = \exp(Z_i(t))$$

where $Z_i(t)$ is the Ornstein-Uhlenbeck process from Section 2.3. Then $S_i(t)$ follows the dynamics

$$dS_i(t) = S_i(t)\Big(A_i Z_i(t) + \frac{1}{2}\sigma_i^2\Big)dt + \sigma_i S_i(t) dB_i(t)$$

for $\mathrm{Corr}(B_1(t), B_2(t)) = \rho$. Assuming a self-financing portfolio, as in Proposition 4.2.4, the resulting dynamics of the wealth process are given by

$$\begin{aligned} dW(t) =& W(t)\Big(u_1(t)(A_1 Z_1(t) + \frac{1}{2}\sigma_1^2) + u_2(t)(A_2 Z_2(t) + \frac{1}{2}\sigma_2^2) + (1 - u_1(t) \\ &- u_2(t))r\Big)dt + W(t)\Big(\sigma_1 u_1(t) dB_1(t) + \sigma_2 u_2(t) dB_2(t)\Big) \end{aligned}$$

The infinetisimal generator $(L^{v_1,v_2}\phi)$ of $W(t)$ for each choice of $(t_0, w_0, z_1, z_2) \in D$, is given by

$$\begin{aligned} (L^{v_1,v_2}\phi)(t_0, w_0, z_1, z_2) =& \frac{\partial \phi}{\partial t} + w\Big(v_1(A_1 z_1 + \frac{1}{2}\sigma_1^2) + v_2(A_2 z_2 + \frac{1}{2}\sigma_2^2) \\ &+ (1 - v_1 - v_2)r\Big)\frac{\partial \phi}{\partial w} + A_1 z_1 \frac{\partial \phi}{\partial z_1} + A_2 z_2 \frac{\partial \phi}{\partial z_2} + \frac{1}{2}w^2\Big(\sigma_1^2 v_1^2 + \sigma_2^2 v_2^2 \\ &+ 2\rho\sigma_1\sigma_2\Big)\frac{\partial^2 \phi}{\partial w^2} + \frac{1}{2}\sigma_1^2 \frac{\partial^2 \phi}{\partial z_1^2} + \frac{1}{2}\sigma_2^2 \frac{\partial^2 \phi}{\partial z_2^2} + w\rho\sigma_1\sigma_2\Big(v_2 \frac{\partial^2 \phi}{\partial w \partial z_1} \\ &+ v_1 \frac{\partial^2 \phi}{\partial w \partial z_2}\Big) + \rho\sigma_1\sigma_2 \frac{\partial^2 \phi}{\partial z_1 \partial z_2} \end{aligned}$$

In view of Theorem 4.2.7, if optimal controls $u_1^*(t)$ and $u_2^*(t)$ exist, their values are maximizers of the HJB-equation, i.e. $\max_{v_1,v_2 \in U}(L^{v_1,v_2}\Phi)(t_0, w_0 z_1, z_2) = 0$

for $(t_0, w_0, z_1, z_2) \in D$ and $u_1^* = v_1$, $u_2^* = v_2$. Preceding as in Section 4.4, the optimal values of the reduced noise problem, are given by

$$v_1 = u_1^*(t) = \frac{1}{w\Phi_{ww}\sigma_1^2}\Big((r - A_1 Z_1(t) - \frac{1}{2}\sigma_1^2)\Phi_w - \rho\sigma_1\sigma_2\Phi_{wz_1}\Big)$$

$$v_2 = u_2^*(t) = \frac{1}{w\Phi_{ww}\sigma_2^2}\Big((r - A_2 Z_2(t) - \frac{1}{2}\sigma_2^2)\Phi_w - \rho\sigma_1\sigma_2\Phi_{wz_2}\Big)$$

Note that, at this point, each of the controls u_i is only dependent on the process $Z_i(t)$ driving the asset i and the correlation between $B_1(t)$ and $B_2(t)$, contrary to the controls obtained in Proposition 4.4.1.

The next objective is to find the optimal value function Φ. By substituting the optimal controls $u_1^*(t)$ and $u_2^*(t)$ into the HJB-equation, the following non-linear boundary value problem is obtained

Result 4.9.1 (HJB-eq. Case of reduced noise).

$$\begin{aligned}
&\Phi_t + w\Phi_w + A_1 z_1 \Phi_{z_1} + A_2 z_2 \Phi_{z_2} + w^2\rho\sigma_1\sigma_2\Phi_{ww} + \frac{1}{2}\sigma_1^2\Phi_{z_1z_1} + \frac{1}{2}\sigma_2^2\Phi_{z_2z_2} \\
&+ \rho\sigma_1\sigma_2\Phi_{z_1z_2} + \frac{1}{2}q_1(z_1)\frac{\Phi_w^2}{\Phi_{ww}\sigma_1^2} + \frac{1}{2}q_2(z_2)\frac{\Phi_w^2}{\Phi_{ww}\sigma_2^2} - \frac{1}{2}(\rho\sigma_1\sigma_2)^2\frac{\Phi_{wz_1}^2}{\Phi_{ww}\sigma_2^2} \\
&- \frac{1}{2}(\rho\sigma_1\sigma_2)^2\frac{\Phi_{wz_2}^2}{\Phi_{ww}\sigma_1^2} = 0 \qquad \textit{for } t < T,\ w > 0
\end{aligned}$$

$$\Phi(t, w, z_1, z_2) = U(w) \qquad \textit{for } t = T \textit{ or } w = 0$$

where

$$q_i(z_i) = (\frac{1}{2}\sigma_i^2 + A_i z_i - r)^2$$

Following Section 4.4 by guessing on a solution ϕ on the form

$$\phi(t, w, z_1, z_1) = f(t, z_1, z_2)w^\gamma, \quad 0 < \gamma < 1 \tag{4.82}$$

and substituting ϕ into Result 4.9.1, the following reduced boundary value problem is obtained

Result 4.9.2 (Reduced HJB-eq. Case of reduced noise).

$$\begin{aligned}
&f_t + A_1 z_1 f_{z_1} + A_2 z_2 f_{z_2} + \frac{1}{2}\sigma_1^2 f_{z_1z_1} + \frac{1}{2}\sigma_2^2 f_{z_2z_2} + \rho\sigma_1\sigma_2 f_{z_1z_2} \\
&- (\rho\sigma_1\sigma_2)^2 f_{z_1}^2 \frac{\gamma}{2\sigma_2^2(\gamma-1)}\frac{1}{f} - (\rho\sigma_1\sigma_2)^2 f_{z_2}^2 \frac{\gamma}{2\sigma_1^2(\gamma-1)}\frac{1}{f} - \Big(- q_1(z_1)\frac{1}{2\sigma_1^2(\gamma-1)} \\
&- q_2(z_2)\frac{1}{2\sigma_2^2(\gamma-1)} - \gamma\Big)\gamma f = 0 \qquad \textit{for } t < T
\end{aligned}$$

$$f(t, z_1, z_1) = 1 \qquad \textit{for } t = T$$

Note that Result 4.9.2 is another non-linear boundary value problem, somewhat similar to the one obtain in Section 4.4, requiring numerical computations to obtain a solution. By assuming independent Brownian motions, i.e. cancelling the correlation between $B_1(t)$ and $B_2(t)$, in Result 4.9.2, we obtain the following *linear* boundary value problem

Result 4.9.3 (Reduced HJB-eq. Case of reduced, independent noise).

$$f_t + A_1 z_1 f_{z_1} + A_2 z_2 f_{z_2} + \frac{1}{2}\sigma_1^2 f_{z_1 z_1} + \frac{1}{2}\sigma_2^2 f_{z_2 z_2} - \Big(- \frac{1}{2\sigma_1^2(\gamma-1)} q_1(z_1) - \frac{1}{2\sigma_2^2(\gamma-1)} q_2(z_2) - \gamma \Big)\gamma f = 0 \quad \textit{for } t < T$$

$$f(t, z_1, z_2) = 1 \quad \textit{for } t = T$$

where $q_i(z_i)$ *as in Result 4.9.1.*

The boundary value problem stated above consists of a much simpler linear partial differential equation, for which a stochastic representation of the solution can be obtained by the Feynman-Kac formula. In view of Theorem 5.7.6, page 366, in [KS91], the solution $f(t, z_1, z_2)$ to the boundary value problem has a stochastic representation on the form

$$f(t, z_1, z_2) = E^{t,z_1,z_2}\Big[\exp\Big(- \int_t^T g(s, z_1, z_2) ds\Big)\Big] \tag{4.83}$$

where

$$g(t, Z_1(t), Z_2(t)) = \gamma\Big(- \frac{1}{2\sigma_1^2(\gamma-1)} q_1(Z_1(t)) - \frac{1}{2\sigma_2^2(\gamma-1)} q_2(Z_2(t)) - \gamma\Big) \tag{4.84}$$

if the following assumptions hold:

(i) $f : [0, T] \times \mathbb{R}^2 \to \mathbb{R}$ is continuous and $f \in C^{1,2,2}(D)$

(ii) f satisfies the Cauchy problem

$$-\frac{\partial f}{\partial t} + kf = (Lf)(t, z_1, z_2) \quad \text{for } (t, z_1, z_2) \in D$$

$$f(t, z_1, z_2) = h(z_1, z_2) \quad \text{for } t = T$$

(iii) and f satisfies the polynomial growth condition

$$\max_{0 \le t \le T} |f(t, z_1, z_2)| \le M(1 + \|z_1\|^k + \|z_2\|^k) \tag{4.85}$$

for constants C, k.

Remark 4.9.4.
Actually, the Feynman-Kac result stated in [KS91] is for the case of one state variable, and the conditions stated above are adapted to the multi-variable representation, which by [Che] easily can be adjusted to the multi-state model, through a multi-variable version of Itô's formula.

In [KS91], additional properties to the theorem, are those of the existence and uniqueness of a solution to the stochastic differential equation of the underlying system, in our case $dW(t)$, but we have such a solution, namely the wealth process, hence those assumptions are neglected. Note that (by multiplying the

PDE in Result 4.9.3 by -1 and switching the partial derivatives over to the right hand side of the equality), the differential equation can be recognized as the boundary value problem in condition (ii), for $h(z_1, z_2) = 1$. By assumptions on Φ for $(L^{v_1,v_2}\Phi)$, any solution f is such that (ii) holds. Furthermore, note that $\gamma - 1 < 0$, and since $q_i(z_i) = (\frac{1}{2\sigma_i^2} + A_i z_i - r)^2 \geq 0$, $g(t, z_1, z_2)$ will be non-negative for all t. It follows that $\exp(-\int_{t_0}^{T} g(s, z_1, z_2)ds) \leq 1$ and the polynomial growth condition is satisfied as well. In conclusion, $f(t, z_1, z_2)$ on the form 4.83, must be a valid stochastic representation of the solution to the boundary value problem (4.9.3).

By discarding the non-stationary drifted Brownian motion, $X(t)$, and assuming independent Brownian motions B_1, B_2, this section illustrates a simplification to the control problem under study in the first parts of Chapter 4. However, the market model is changed, and one assumes mean-revering asset prices. From an economic point of view, the mean-reverting effect might be considered as the market impact on price, for instance. If the current market price of an asset is less than the average, an investor will expect the price to rise, hence purchase of this particular asset is viewed as attractive. If many investors start purchasing the particular asset, the price will indeed return to market value (and perhaps above). If the other way around, i.e. an asset is priced above market price, the price is expected to fall, making the asset unattractive.

Chapter 5

"Constant" Control Problem

5.1 Introduction

This chapter presents a naive approach to find a constant fraction investment strategy. The idea is motivated by the optimal solution to Merton's portfolio problem. The controls u_1 and u_2 are initially assumed to be constant, and will be treated as such in the calculations, in order to justify the application of a simple mathematical procedure from linear optimization.

5.2 Statement of the Problem

In the solution to Merton's problem, presented in Section 3.5, the optimal allocation in the two-asset model, is to hold a constant fraction of wealth, $u^*(t) = u^*$, in the risky asset. Thereby, the investor also holds a constant fraction $1 - u^*$ in the risk-free assets. Inspired by the constant control solution, this section presents an unconventional approach to find admissible controls u_1, u_2 to the problem studied in the previous chapter. The market model is assumed to be the same, i.e. the asset prices $S_1(t)$ and $S_2(t)$ are modelled by the exponential of a common non-stationary shifted Brownian motion $X(t)$ and by two distinct stationary Ornstein-Uhlenbeck processes $Z_1(t)$ and $Z_2(t)$, generated by correlated Brownian motions $B_1(t)$ and $B_2(t)$, respectively. The aim of the investor is still to maximize expected utility of wealth at the end of the investment horizon, given the power utility function representing the investor's aversion towards risk. We look for admissible, but not necessarily optimal by means of the HJB-equation, controls u_1 and u_2 by assuming the controls are constant fractions of wealth. However, as it turns out, the controls are not nearly constant, they are not even deterministic.

The calculations in this chapter are so-called "quick and dirty," meaning we allow ourselves to use methods and techniques which are not necessarily mathematically correct. The strategy is as follows: Assume there exist constant controls, i.e. constant fractions of wealth, which are solutions to the following unconstrained maximization problem

$$\max_{u_1, u_2} \quad E[U(W(T))|\mathcal{F}(t)] \tag{5.1}$$

We seek the *optimal* constant fractions u_1 and u_2 maximizing the expected value of utility of wealth at the future time point T, given the market information revealed to the investor at the current time t. Note that optimal controls by means of (5.1) are not the same as optimal in the sense of Chapter 4! As before, $\mathcal{F}_t$ denotes the filtration w.r.t. the correlated Brownian motions $\widetilde{B}(t)$, $B_i(t)$ for $i = 1, 2$.

We start by finding the conditional expectation of the wealth process, simultaneously showing that $W(t)$ is Markovian. Under the assumption that u_1, u_2 are constant, we show that the solution to the wealth dynamics $dW(t)$ is a log-normally distributed process, namely the exponential of the correlated Brownian motions $\widetilde{B}(t), B_1(t), B_2(t)$ and the Ornstein-Uhlenbeck processes $Z_1(t)$ and $Z_2(t)$. Thereby applying the formula for the expected value of a log-normal random variable, by standard differentiation, we obtain the *optimal* solutions $u_1^{'}$ and $u_2^{'}$, and check if they are admissible by means of definition (4.2.1).

5.3 The Wealth Process under "Constant" Controls

Assuming u_1 and u_2 are constant, recall the stochastic differential equation for the wealth given in (4.6)

$$\begin{aligned} dW_t = W(t)\Big(&u_1(\mu + A_1 Z_1(t) + \frac{1}{2}\Sigma_1) + u_2(\mu + A_2 Z_2(t) + \frac{1}{2}\Sigma_2) \\ &+ (1 - u_1 - u_2)r\Big)dt + W(t)\Big(\sigma(u_1+u_2)d\widetilde{B}_t + \sigma_1 u_1 dB_1(t) + \sigma_2 u_2 dB_2(t)\Big) \end{aligned} \tag{5.2}$$

The explicit solution $W(t)$ to (5.2) is given by the following proposition.

Proposition 5.3.1 (The Wealth Process).
Assuming the controls u_1 and u_2 are constant and given initial wealth $W(0) = w_0$, the unique solution of the controlled stochastic differential equation (5.2) is given by

$$\begin{aligned} W(t) = w_0 \exp\Bigg\{&\Big(u_1(\mu + \frac{1}{2}\Sigma_1) + u_2(\mu + \frac{1}{2}\Sigma_2) + (1 - u_1 - u_2)r - \frac{1}{2}\Big(\sigma^2(u_1 \\ &+ u_2)^2 + \sigma_1^2 u_1^2 + \sigma_2^2 u_2^2\Big) - \sigma(u_1+u_2)(\rho_1\sigma_1 u_1 + \rho_2\sigma_2 u_2) - \rho\sigma_1\sigma_2 u_1 u_2\Big)t \\ &+ \int_0^t \Big(u_1 A_1 Z_1(s) + u_2 A_2 Z_2(s)\Big)ds + \sigma(u_1+u_2)\widetilde{B}(t) + \sigma_1 u_1 B_1(t) \\ &+ \sigma_2 u_2 B_2(t)\Bigg\} \end{aligned} \tag{5.3}$$

where $Z_1(t), Z_2(t)$ are the Ornstein-Uhlenbeck processes (2.14), and $\widetilde{B}(t)$, $B_1(t)$, $B_2(t)$ are the correlated Brownian motions presented in Chapter 2.1.

Proof of proposition (5.3.1). The idea is to use Itô's formula to rewrite the process (5.3) as a stochastic differential equation. Define a function f by

$$f(t, x, y, z) = w_0 \exp\Big(At + Bx + Cy + Dz\Big)$$

where we assume that

$$W(t) = f(t, \widetilde{B}(t), B_1(t), B_2(t)) = w_0 \exp\Big(At + B\widetilde{B}(t) + CB_1(t) + DB_2(t)\Big)$$

for some functions A, B, C, D. By applying Itô's lemma on f, we obtain that

$$\begin{aligned} df(t,x,y,z) =& f \cdot Adt + f \cdot Bd\widetilde{B}(t) + f \cdot CdB_1(t) + f \cdot DdB_2(t) \\ &+ \frac{1}{2} f \cdot \Big(B^2 + C^2 + D^2\Big)dt + f \cdot \Big(BC\rho_1 + BD\rho_2 + CD\rho\Big)dt \\ =& f \cdot \Big(A + \frac{1}{2}(B^2 + C^2 + D^2) + \rho_1 BC + \rho_2 BD + \rho CD\Big)dt \\ &+ f \cdot \Big(Bd\widetilde{B}(t) + CdB_1(t) + DdB_2(t)\Big) \end{aligned} \tag{5.4}$$

for the correlation coefficients ρ, ρ_i, $i = 1, 2$. Let now $B = \sigma(u_1 + u_2)$, $C = \sigma_1 u_1$, $D = \sigma_2 u_2$ and

$$\begin{aligned} A = & u_1(\mu + A_1 Z_1(t) + \frac{1}{2}\Sigma_1) + u_2(\mu + A_2 Z_2(t) + \frac{1}{2}\Sigma_2) + (1 - u_1 - u_2)r - \frac{1}{2}\Big(\sigma^2 \\ & \times (u_1 + u_2)^2 + \sigma_1^2 u_1^2 + \sigma_2^2 u_2^2\Big) - \sigma(u_1 + u_2)(\rho_1 \sigma_1 u_1 + \rho_2 \sigma_2 u_2) - \rho \sigma_1 \sigma_2 u_1 u_2 \end{aligned}$$

By substituting the expressions into (5.4), we see that

$$df(t, \widetilde{B}(t), B_1(t), B_2(t) = dW(t) \tag{5.5}$$

and hence $W(t) = f(t, \widetilde{B}(t), B_1(t), B_2(t))$ for A, B, C and D defined above. By construction, $W(t)$ must be the unique solution to (5.2). ■

Note that, by very little adjustment, the solution to the dynamics of $W(t)$ with time-dependent controls $u_i(t)$, from the previous chapter, is given by

$$\begin{aligned} W(t) = w_0 \exp\Bigg\{ \Big(\int_0^t & u_1(s)(\mu + \frac{1}{2}\Sigma_1) + u_2(s)(\mu + \frac{1}{2}\Sigma_2) + (1 - u_1(s) - u_2(s))r \\ & - \frac{1}{2}\Big(\sigma^2(u_1(s) + u_2(s))^2 + \sigma_1^2 u_1^2(s) + \sigma_2^2 u_2^2(s)\Big) - \sigma(u_1(s) + u_2(s)) \\ & \times (\rho_1 \sigma_1 u_1(s) + \rho_2 \sigma_2 u_2(s)) - \rho \sigma_1 \sigma_2 u_1(s) u_2(s)\Big)ds + \int_0^t \Big(u_1(s) A_1 Z_1(s) \\ & + u_2(s) A_2 Z_2(s)\Big)ds + \int_0^t \sigma(u_1(s) + u_2(s))\widetilde{B}(t)ds + \int_0^t \sigma_1 u_1(s) B_1(t) ds \\ & + \int_0^t \sigma_2 u_2(s) B_2(t) ds \Bigg\} \end{aligned} \tag{5.6}$$

5.4 The Optimization Problem

Following, the *optimal* controls are assumed to be solutions of a standard unconstrained optimization problem. Under the bold assumptions that the controls are constant fractions, we solve the following optimization problem introduced in the beginning of this chapter

Problem 5.4.1 (Linear optimization problem).

$$\max_{u_1, u_2} \quad E[U(W(T))|\mathcal{F}(t)] \tag{5.7}$$

Note that we for simplicity in this chapter disregard the assumption of non-negative wealth. The following notation will become useful for the calculations of problem (5.4.1).

Notation 5.4.2.
We denote the constants $\nu(u_1,u_2)$ and $\widetilde{\nu}(u_1,u_2)$, depending on the controls u_1 and u_2, by

$$\nu(u_1,u_2) = u_1(\mu + \frac{1}{2}\Sigma_1) + u_2(\mu + \frac{1}{2}\Sigma_2) + (1-u_1-u_2)r - \frac{1}{2}\Big(\sigma^2(u_1+u_2)^2 + \sigma_1^2 u_1^2 + \sigma_2^2 u_2^2\Big) - \sigma(u_1+u_2)(\rho_1\sigma_1 u_1 + \rho_2\sigma_2 u_2) - \rho\sigma_1\sigma_2 u_1 u_2 \quad (5.8)$$

$$\widetilde{\nu}(u_1,u_2) = \gamma\nu(u_1,u_2) + \frac{1}{2}\gamma^2\Big(u_1^2(\sigma^2+2\sigma_1^2) + u_2^2(\sigma^2+2\sigma_2^2) + 2u_1u_2(\rho\sigma_1\sigma_2 + \sigma^2)\Big) \quad (5.9)$$

Furthermore, let $F_i, F_i^2, F_{i,j}$ for $i,j=1,2$ denote the following functions

$$F_i = \frac{1}{A_i}\Big[e^{A_i(T-t)} - 1\Big], \qquad F_i^2 = \frac{1}{2A_i}\Big[e^{2A_i(T-t)} - 1\Big], \quad i=1,2$$

$$F_{i,j} = \frac{1}{A_i + A_j}\Big[e^{(A_i+A_j)(T-t)} - 1\Big], \quad i,j=1,2 \quad i \neq j \quad (5.10)$$

We aim at solving problem (5.4.1) by means of ordinary optimization techniques, i.e. the *optimal* u_1 and u_2 are solutions to the first order maximum equations

$$\frac{\partial E[U(W(T))|\mathcal{F}_t]}{\partial u_i} = 0$$

for $i=1,2$. The conditional expectation of $W(T)$, given the filtration $\mathcal{F}_t$ at time $t \le T$, is required.

Proposition 5.4.3 (Conditional expectation of utility).
Given initial wealth w_0 and assuming constant controls u_1,u_2, the expectation of the controlled process (5.3) conditioned on $\mathcal{F}_t$, is given by

$$E[U(W(T))|\mathcal{F}_t] = W^\gamma(t)\exp\bigg\{\gamma\Big(u_1A_1Z_1(t)\int_t^T e^{A_1 s}ds + u_2A_2Z_2(t)\int_t^T e^{A_2 s}ds\Big) + \tilde{\nu}(u_1,u_2)(T-t) + \frac{1}{2}\gamma^2\bigg(u_1^2\Big(\sigma_1^2F_1^2 + \rho_1\sigma\sigma_1F_1\Big) + u_2^2\Big(\sigma_2^2F_2^2 + \rho_2\sigma\sigma_2F_2\Big) + u_1u_2\Big(2\rho\sigma_1\sigma_2(F_{1,2} - F_1 - F_2) + \rho_1\sigma\sigma_1F_1 + \rho_2\sigma\sigma_2F_{2,1}\Big)\bigg\} \quad (5.11)$$

Before we perceede with the proof, we need the following lemma.

Lemma 5.4.4 (Fubini on $Z_i(t)$).
Let $Z_i(t)$ be the Ornstein-Uhlenbeck process defined in Section 2.3. Then the following equality holds

$$E\Big[\int_0^T Z_i(s)ds\Big] = \int_0^T E[Z_i(s)]ds \quad (5.12)$$

Proof of Lemma 5.4.4. By Jensen's inequality, Theorem B.3.7, for $g(x) = x^2$ note that

$$\Big(E\Big[|Z_i(t)|\Big]\Big)^2 \leq E\Big[|Z_i(t)|^2\Big] \tag{5.13}$$

for any t. Since $g(x)$ is non-negative and measurable for $x = Z_i(t)$, by Tonelli's theorem, Theorem B.3.3, we obtain that

$$\begin{aligned} E[\int_0^T Z_i^2(s)ds] &= \int_\Omega \int_t^T Z_i^2(s)dsdP(\omega) \\ &= \int_t^T \int_\Omega Z_i^2(s)dP(\omega)ds \\ &= \int_t^T E[Z_i^2(s)]ds \end{aligned}$$

Since $Z_i(t) \sim \mathcal{N}(\mu_i, \sigma_i^2)$ for any t, $Z_i(t)$ has finite moments, yielding

$$E[\int_0^T Z_i^2(s)ds] = \int_0^T E[Z_i^2(s)]ds < \infty \tag{5.14}$$

But then it follows that,

$$E[\int_0^T Z_i(s)ds] \leq E[\int_0^T Z_i^2(s)ds] < \infty \tag{5.15}$$

and we may apply Fubini's theorem on $E[\int_0^T Z_i(s)ds]$. ■

Proof of Proposition 5.4.3. Recall Equation (5.3). We aim at finding

$$\begin{aligned} E[W^\gamma(T)|\mathcal{F}_t] = E\Big[w_0^\gamma \exp\Big\{\gamma\Big(\nu(u_1,u_2)T + \int_0^T u_1A_1Z_1(s)ds + \int_0^T u_2A_2Z_2(s)ds \\ + \int_0^T \sigma(u_1+u_2)d\widetilde{B}(s) + \int_0^T \sigma_1u_1dB_1(s) + \int_0^T \sigma_2u_2dB_2(s)\Big)\Big\}|\mathcal{F}_t\Big] \end{aligned} \tag{5.16}$$

We expand the expectation by adding and subtracting the process evaluated at time t

$$\begin{aligned} E\Big[w_0^\gamma \exp\Big\{\gamma\Big(\nu(u_1,u_2)(T-t+t) + u_1A_1\Big(\int_0^T Z_1(s)ds - \int_0^t Z_1(s)ds \\ + \int_0^t Z_1(s)ds\Big) + u_2A_2\Big(\int_0^T Z_2(s)ds - \int_0^t Z_2(s)ds + \int_0^t Z_2(s)\Big)ds \\ + \sigma(u_1+u_2)\Big(\widetilde{B}(T) - \widetilde{B}(t) + \widetilde{B}(t)\Big) + \sigma_1u_1\Big(B_1(T) - B_1(t) + B_1(t)\Big) \\ + \sigma_2u_2\Big(B_2(T) - B_2(t) + B_2(t)\Big)\Big)\Big\}|\mathcal{F}_t\Big] \end{aligned} \tag{5.17}$$

Since $\int_0^t Z_i(s)ds$ and $\widetilde{B}(t), B_i(t)$ are $\mathcal{F}_t$-measurable for $i = 1, 2$, by properties of conditional expectations, Proposition B.4.2, the integral and the Brownian motions can be taken out of the expectation. Note further that

$$\int_t^T Z_i(s)ds = \int_t^T Z_i(t)e^{A_is}ds + \int_t^T \int_t^s \sigma_i e^{A_i(s-u)}dB_i(u)ds$$

$$=Z_i(t)\int_t^T e^{A_i s}ds + \int_t^T\int_t^s \sigma_i e^{A_i(s-u)}dB_i(u)ds$$

by definition of $Z_i(s)$. The process evaluated at time t, $Z_i(t)$, is $\mathcal{F}_t$-measurable. Following the argumentation above, it can be taken out of the expectation. The remaining Itô integrals from t to T are independent of $\mathcal{F}_t$, and hence the conditioning on $\mathcal{F}_t$ is superfluous. For the last integral, since $\sigma_i e^{A_i(s-u)}$ obviously is measurable w.r.t. $\mathcal{B}([0,t])\times\mathcal{F}_t$, by the stochastic Fubini theorem, Theorem B.3.5, we can interchange the order of integration such that

$$\begin{aligned}\int_t^T\int_t^s \sigma_i e^{A_i(s-u)}dB_i(u)ds &= \int_t^T\int_t^T \mathcal{X}_{\{u\leq s\}}\sigma e^{A_i(s-u)}dB_i(u)ds\\ &= \int_t^T\int_u^T \sigma_i e^{A_i(s-u)}dsdB_i(u) \qquad (5.18)\end{aligned}$$

and equation (5.17) then equals

$$\begin{aligned}&W^\gamma(t)\exp\Big\{\gamma\Big(\nu(u_1,u_2)(T-t)+u_1A_1Z_1(t)\int_t^T e^{A_1 s}ds\\ &\qquad + u_2A_2Z_2(t)\int_t^T e^{A_2 s}ds\Big)\Big\}\\ &\times E\Big[\exp\Big\{\gamma\Big(\int_t^T\int_u^T u_1A_1\sigma_1 e^{A_1(s-u)}dsdB_1(u)+\int_t^T\int_u^T u_2A_2\sigma_2\\ &\qquad\times e^{A_2(s-u)}dsdB_2(u)+\int_t^T\sigma(u_1+u_2)d\widetilde{B}(s)+\int_t^T\sigma_1u_1dB_1(s)\\ &\qquad+\int_t^T\sigma_2u_2dB_2(s)\Big)\Big\}\Big] \qquad (5.19)\end{aligned}$$

by recognizing parts of the first exponential as the utility function of the wealth process evaluated at t. Let now $Y(T)$ denote the stochastic process in the exponential inside the expectation in (5.19), i.e.

$$\begin{aligned}&Y(T))=\\ &\quad\gamma\Big(\int_t^T\int_u^T u_1A_1\sigma_1e^{A_1(s-u)}dsdB_1(u)+\int_t^T\int_u^T u_2A_2\sigma_2e^{A_2(s-u)}dsdB_2(u)\\ &\quad+\int_t^T\sigma(u_1+u_2)d\widetilde{B}(s)+\int_t^T\sigma_1u_1dB_1(s)+\int_t^T\sigma_2u_2dB_2(s)\Big)\end{aligned}$$

Following the argument in Section 2.2, by Cholesky decomposition the three correlated Brownian motions $\widetilde{B}(t)$, $B_i(t), i=1,2$ can be expressed as linear combinations of independent Brownian motions $U_i, i=1,2,3$. By that, $Y(T)$ is the sum of independent normally distributed random variables, hence $\exp(Y(t))$ is a log-normal random variable for each t. The problem of finding the explicit expectation in (5.19) falls down to finding the expectation of $Y(T)$ and exploiting the formula for the expected value of a log-normal random variable

$$E[\exp(Y(T))]=\exp(\mu_Y+\frac{1}{2}\sigma_Y^2),\quad Y(T)\sim\mathcal{N}(\mu_Y,\sigma_Y^2) \qquad (5.20)$$

Recall (5.18). By the property of Itô integrals, Proposition B.1.4,

$$E[\int_0^t g(u)dB(u)] = 0 \tag{5.21}$$

if g in $\mathcal{V}$. Since $\int_u^T \exp(A_i(s-u))ds \in \mathcal{V}$,

$$E\Big[\int_t^T \int_u^T \sigma_i e^{A_i(s-u)} ds dB_i(u)\Big] = 0 \tag{5.22}$$

Hence by the previous argument $E[Y(T)] = 0$. We have left to find the variance of $Y(T)$, now obtained by

$$\text{Var}(Y(T)) = E[Y^2(T)] \tag{5.23}$$

Since the calculations of $E[Y^2(T)]$ are cumbersome, they will be omitted omitted here, and we only state the result. The calculations can be found in Appendix C.

$$\begin{aligned}\text{Var}(Y(T)) = \gamma^2\bigg(&\Big(u_1^2(\sigma^2+2\sigma_1^2) + u_2^2(\sigma^2+2\sigma_2^2) + 2u_1u_2(\rho\sigma_1\sigma_2+\sigma^2)\Big)(T-t)\\ &+u_1^2\Big(\sigma_1^2F_1^2 + \rho_1\sigma\sigma_2F_1\Big) + u_2^2\Big(\sigma_2^2F_2^2 + \rho_2\sigma\sigma_2F_2\Big)\\ &+u_1u_2\Big(2\rho\sigma_1\sigma_2(F_{1,2}-F_1-F_2) + \rho_1\sigma\sigma_1F_1 + \rho_2\sigma_2\sigma F_2\Big)\bigg)\end{aligned} \tag{5.24}$$

Combining (5.24) and the fact that $E[Y(T)] = 0$ with equation (5.19), we obtain the conditional expected value of $E[U(W^\gamma(T)|\mathcal{F}_t]$ in (5.11). This completes the proof. ■

Note the Markov property of $W(t)$ in (5.19). $W(t)$ is a Markov process w.r.t. both the filtration generated by $W(t)$ itself, and the filtration generated by the Ornstein-Uhlenbeck processes $Z_i(t)$ for $i = 1, 2$. By definition of $\mathcal{F}$, see for instance Section 2.1, we simply say that $W(t)$ is a Markov process w.r.t. $\mathcal{F}_t$.

5.5 The Controls $u_i^{'}$

In the previous section we derived an explicit expression of the expected value of utility of terminal wealth, conditioned on the market information available up till time t. In this section we derive the optimal, by means of Equation (5.7), controls. We are still under the assumption that the controls are "constant", such that the calculations in the previous section and the following differentiation can be justified.

Proposition 5.5.1 ("Optimal" controls).
The controls $u_1^{'}$ and $u_2^{'}$ maximizing Equation (5.7) *are given by*

$$\begin{aligned}u_1^{'} = &\Big(A_1Z_1(t)\int_t^T e^{A_1 s}ds + \gamma\int_0^t A_1Z_1(s)ds\Big)\vartheta_1 + \Big(A_2Z_2(t)\int_t^T e^{A_2 s}ds\\ &+\gamma\int_0^t A_2Z_2(s)ds\Big)\vartheta_2 + \gamma\sigma\widetilde{B}(t)\vartheta_3 + \gamma\sigma_1B_1(t)\vartheta_4 + \gamma\sigma_2B_2(t)\vartheta_5 + \vartheta_6\end{aligned} \tag{5.25}$$

$$u_2' = \Big(A_1 Z_1(t)\int_t^T e^{A_1 s}ds + \gamma\int_0^t A_1 Z_1(s)ds\Big)\varepsilon_1 + \Big(A_2 Z_2(t)\int_t^T e^{A_2 s}ds$$
$$+\gamma\int_0^t A_2 Z_2(s)ds\Big) + \gamma\sigma\widetilde{B}(t)\varepsilon_2 - \gamma\sigma_1 B_1(t)\varepsilon_3 + \gamma\sigma_2 B_2(t) - \varepsilon_4 \tag{5.26}$$

where $\vartheta_i, k = 1,\dots,6$ and $\varepsilon_l, l = 1,\dots,4$ are constants, defined in Appendix A, depending on the correlations ρ,ρ_1,ρ_2 and the standard deviations σ,σ_1,σ_2.

Proof of Proposition 5.5.1. First note that, by construction of $E[U(W(T))|\mathcal{F}_t]$

$$\frac{\partial W^\gamma(t)}{\partial u_i} + W^\gamma(t)\frac{\partial g(u_i,u_j)}{\partial u_i} = 0 \tag{5.27}$$

where

$$\frac{\partial W^\gamma(t)}{\partial u_i} = W^\gamma(t)\gamma\Big[\Big((\mu\frac{1}{2}\Sigma_i - r) - u_i\Sigma_i - u_j\Sigma\Big)t$$
$$+\int_0^t A_i Z_i(s)ds + \sigma\widetilde{B}(t) + \sigma_i B_i(t)\Big] \tag{5.28}$$

and

$$\frac{\partial g(u_i,u_j)}{\partial u_i} = \gamma\Big(A_i Z_i(t)\int_t^T e^{A_i s}ds\Big) + \gamma\Big(\mu + \frac{1}{2}\Sigma_i - r\Big)(T-t)$$
$$- u_i\gamma\Big[\Big(\Sigma_i - \gamma(\sigma^2 + 2\sigma_i^2)\Big)(T-t) - \gamma\Big(\sigma_i^2 F_i^2 + \rho_i\sigma\sigma_i F_i\Big)\Big]$$
$$- u_j\gamma\Big[\Big(\Sigma - \gamma P_j\Big)(T-t) - \gamma\Big(\rho\sigma_i\sigma_j(F_{i,j} - F_i - F_j) + \frac{1}{2}(\rho_i\sigma\sigma_i F_i$$
$$+ \rho_j\sigma\sigma_j F_j)\Big)\Big] \tag{5.29}$$

Combining (5.28) and (5.29) in (5.27), then solving for u_i, we obtain the following expression

$$u_i\Big[\Sigma_i T - \gamma\Big((\sigma^2 + 2\sigma_i^2)(T-t) + \sigma_i^2 F_i^2 + \rho_i\sigma\sigma_i F_i\Big)\Big]$$
$$= A_i Z_i(t)\int_t^T e^{A_i s}ds + (\mu + \frac{1}{2}\Sigma_i - r)T + \gamma\int_0^t A_i Z_i(s)ds + \gamma\sigma\widetilde{B}(t)$$
$$+ \gamma\sigma_i B_i(t) - u_j\Big[\Sigma T - \gamma\Big(+ P_j(T-t)$$
$$\rho\sigma_i\sigma_j(F_{i,j} - F_i - F_j) + \frac{1}{2}(\rho_i\sigma\sigma_i F_i + \rho_j\sigma\sigma_j F_j)\Big)\Big] \tag{5.30}$$

By substitution of the expression of u_1 into u_2, we obtain the explicit expression (5.25) and (5.26) for $\vartheta_k, k = 1,\dots,6$ and $\varepsilon_l, l = 1,\dots,4$ in the proposition above. Some of the intermediate calculations are given in Appendix C. ■

Note that, by (5.5.1), we have reached a contradiction. Under the initial assumption the controls were constant, and we were allowed to optimize by means of ordinary maximization. However, as it turns out, the optimal controls are not constant, neither are they deterministic. Both u_1 and u_2 are time-dependent and stochastic processes of the stationary Ornstein-Uhlenbeck processes and the correlated Brownian motions. Our initial attempt to find constant fractions of wealth as in Merton's original case, failed, however, as we will see by the following theorem, we have found a set of admissible controls. In conclusion, $u_1^{'}$ and $u_2^{'}$ are non-stationary stochastic control processes. We summarize in the following theorem.

Theorem 5.5.2 (Admissible controls).
Initially assuming constant controls u_1 and u_2, the solution to the optimization problem (5.7) w.r.t. u_1 and u_2, given conditional expectation (5.11), are two admissible stochastic controls

$$
\begin{aligned}
u_1^{'} = &\Big(A_1 Z_1(t)\int_t^T e^{A_1 s}ds + \gamma\int_0^t A_1 Z_1(s)ds\Big)\vartheta_1 + \Big(A_2 Z_2(t)\int_t^T e^{A_2 s}ds \\
&+\gamma\int_0^t A_2 Z_2(s)ds\Big)\vartheta_2 + \gamma\sigma\widetilde{B}(t)\vartheta_3 + \gamma\sigma_1 B_1(t)\vartheta_4 + \gamma\sigma_2 B_2(t)\vartheta_5 + \vartheta_6
\end{aligned}
$$

$$
\begin{aligned}
u_2^{'} = &\Big(A_1 Z_1(t)\int_t^T e^{A_1 s}ds + \gamma\int_0^t A_1 Z_1(s)ds\Big)\varepsilon_1 + \Big(A_2 Z_2(t)\int_t^T e^{A_2 s}ds \\
&+\gamma\int_0^t A_2 Z_2(s)ds\Big) + \gamma\sigma\widetilde{B}(t)\varepsilon_2 - \gamma\sigma_1 B_1(t)\varepsilon_3 + \gamma\sigma_2 B_2(t) - \varepsilon_4
\end{aligned}
$$

Proof of Theorem 5.5.2. By construction, $u_1^{'}$ and $u_2^{'}$ are such that $W(t, u_1^{'}, u_2^{'})$ is the unique solution to the wealth dynamics given in (5.2). Under questionable mathematical methods, we have shown that $u_1^{'}$ and $u_2^{'}$ are solutions to the optimization problem in (5.7). We have left to show that $u_1^{'}(t)$ and $u_2^{'}(t)$ are indeed admissible controls. Note that,

$$
A_i Z_i(t)\int_t^T e^{A_i s}ds = Z_i(t)[e^{A_i T} - e^{A_i t}], \quad i = 1, 2
$$

which is obviously measurable and $\mathcal{F}_t$-adapted by $Z_i(t)$. The same follows for the Brownian motions $\widetilde{B}(t), B_i(t)$ for $i = 1, 2$. Furthermore,

$$
\begin{aligned}
\gamma\int_0^t A_i Z_i(s)ds =&\gamma\Big(\int_0^t z_i e^{A_i s}ds + \int_0^t\int_0^s \sigma_i e^{A_i(s-u)}dB_i(u)ds\Big) \\
=&\gamma\Big(z_i[e^{A_i t} - 1] + \sigma_i A_i\int_0^t\int_u^t e^{A_i(s-u)}dsdB_i(u)\Big)
\end{aligned}
$$

by the stochastic Fubini theorem, Theorem B.3.5. Following, we see that

$$
\begin{aligned}
&\gamma\Big(z_i[e^{A_i t} - 1] + \sigma_i\int_0^t\int_u^t e^{A_i(s-u)}dsdB_i(u)\Big) \\
&\qquad = \gamma\Big(z_i[e^{A_i t} - 1] + \int_0^t e^{A_i(t-u)}dB_i(u) - B_i(t)\Big)
\end{aligned}
$$

Hence $\int_0^t A_i Z_i(t)ds$ is $\mathcal{F}_t$-measurable and adapted as well, and the controls inherit both measurability and adaptedness from $Z_i(t)$ and the Brownian motions. The Itô integrals are by construction normally distributed. They may be approximated by the sum of normally distributed random variables, and since the limit of a convergent sequence of Gaussian random variables still is a Gaussian random variable, the integral itself is normally distributed, and

$$\int_0^t e^{A_i(t-u)} dB_i(u) \sim \mathcal{N}(0, \frac{1}{2A_i^2}[e^{2A_i t} - 1])$$

Since the controls are linear in $Z_i(t), \widetilde{B}(t), B_i(t)$, similar to the proof of Theorem 4.7.1, $u_i^{'}(t)$ will be a normally distributed random variable for any t, with existing and finite first and second moments, hence $E[\int_0^T (u_i^{'}(t))^2 ds] < \infty$, and condition (iii) of Definition 4.2.1 is fulfilled. By definition (4.2.1), $u_i^{'}$ are admissible on the interval $[0, T]$. This concludes the proof. ■

Note that, similarly to the control problem solved in the previous chapter, the controls obtained as the optimal solution to Equation (5.7) depend on the Ornstein-Uhlenbeck processes $Z_i(t)$, but in addition, through this naive approach, we now have the non-stationary dependency from the correlated $\widetilde{B}(t)$ and $B_i(t)$. The expected value of $u_1^{'}$ and $u_2^{'}$ is given by

$$\begin{aligned} E[u_1^{'}(t)] =& \Big(z_1 e^{A_1 t}(e^{A_1 T} - e^{A_1 t}) + \gamma z_1 (e^{A_1 t} - 1)\Big)\vartheta_1 + \Big(z_2 e^{A_2 t}(e^{A_2 T} - e^{A_2 t}) \\ &+ \gamma z_2 (e^{A_2 t} - 1)\Big)\vartheta_2 + \vartheta_6 \end{aligned}$$

by noting that

$$E[A_i Z_i(t) \int_t^T e^{A_i s} ds] = z_i \int_0^T e^{A_i s} ds \tag{5.31}$$

Also, for $u_2^{'}(t)$

$$\begin{aligned} E[u_2^{'}(t)] =& \Big(z_1 e^{A_1 t}(e^{A_1 T} - e^{A_1 t}) + \gamma z_1 (e^{A_1 t} - 1)\Big)\varepsilon_1 + \Big(z_2 e^{A_2 t}(e^{A_2 T} - e^{A_2 t}) \\ &+ \gamma z_2 (e^{A_2 t} - 1)\Big)\varepsilon_2 - \varepsilon_4 \end{aligned}$$

Recall the notation of $\vartheta_i, \varepsilon_j$, and note their dependency on T and t. The controls obtained in this chapter are directly dependent on time, both of the current time t and the future time T at the end of the investment horizon. In addition, the controls are depending on the current value of the non-stationary Brownian motions, which is completely different to case of Merton. Note however that the controls are not optimal by means of the dynamic programming approach or the HJB-equation, they are simply optimal in the sense that their values are maximizers to the optimization problem stated at the beginning of this chapter.

5.6 T-dependent Controls at $t = 0$

For this section, we assume the initial time point is $t = 0$, and take another look at the result obtained in the previous section. Still under the initial assumption that u_1 and u_2 are constant, for $t = 0$, the conditional expected value turns out to be

$$E[U(W(T))|\mathcal{F}_0] = w_0^{\gamma} \exp\Big\{\gamma\Big(u_1 A_1 z_1 \int_0^T e^{A_1 s} ds + u_2 A_2 z_2 \int_0^T e^{A_2 s} ds\Big)$$

$$+ \tilde{\nu}(u_1, u_2)T + \frac{1}{2}\gamma^2\Big(u_1^2(\sigma_1^2\bar{F}_1^2 + \rho_1\sigma\sigma_1\bar{F}_1) + u_2^2(\sigma_2^2\bar{F}_2^2 + \rho_2\sigma\sigma_2\bar{F}_2)$$
$$+ u_1u_2(2\rho\sigma_1\sigma_2(\bar{F}_{1,2} - \bar{F}_1 - \bar{F}_2) + \rho_1\sigma\sigma_1\bar{F}_1 + \rho_2\sigma\sigma_2\bar{F}_{2,1})\Big)\Big\} \quad (5.32)$$

for

$$\bar{F}_i = \frac{1}{A_i}\Big[e^{A_iT} - 1\Big], \quad \bar{F}_i^2 = \frac{1}{2A_i}\Big[e^{2A_i(T-t)} - 1\Big]$$
$$\bar{F}_{i,j} = \frac{1}{A_i + A_j}\Big[e^{(A_i+A_j)T} - 1\Big]$$

and $\nu(u_1, u_2), \widetilde{\nu}(u_1, u_2)$ as in Notation 5.4.2. The optimal allocation of wealth, by means of Equation (5.7) is given by

$$\bar{u}_1 = \Big(A_1z_1\int_0^T e^{A_1s}ds\Big)\vartheta_1 + \Big(A_2z_2\int_0^T e^{A_2s}ds\Big)\vartheta_2 + \vartheta_6 \quad (5.33)$$

$$\bar{u}_2 = \Big(A_1z_1\int_0^T e^{A_1s}ds\Big)\varepsilon_1 + A_2z_2\int_0^T e^{A_2s}ds - \varepsilon_4 \quad (5.34)$$

The controls are simply obtained by setting $t = 0$ in Proposition 5.5.1. Note that in the case where the initial starting point is 0, $\bar{u}_1$ and $\bar{u}_2$ only depend on the future time T. Following the same argumentation as in Theorem 5.5.2, we see that $\bar{u}_1$ and $\bar{u}_2$ are indeed admissible. Since each of the components of ϑ_1 and ϑ_2 are linear in T (see Appendix A), as the investment horizon is expanded, i.e. $T \to \infty$, the exponential terms converge to 1 faster than T goes to infinity (recall that $A_i < 0$). Note that

$$\vartheta_6 = \frac{\varrho_1T^3 - T^2\varrho_2}{T^3(\varrho_3 + \varrho_4)} = \frac{\varrho_1 + \varrho_2/T}{\varrho_3 + \varrho_4} \quad (5.35)$$

for some generic constants $\varrho_i, i = 1, \ldots, 4$. (These are extensive constants depending on the volatility of the Brownian motions and the correlations among them). Asymptotically, the "optimal" control $\bar{u}_1$ seems to tend to a finite value as $T \to \infty$

$$\lim_{T\to\infty}\vartheta_6 = \frac{\varrho_1}{\varrho_3 + \varrho_4} \quad (5.36)$$

For instance, in an independent market model, where there are no correlations, note that

$$\lim_{T\to\infty}\vartheta_6 = \frac{\sigma^4(1-\gamma)^2}{\sigma^4(1-\gamma)^2(\sigma^2(1-\gamma) + \sigma_1^2(1-2\gamma)) + \sigma^2(1-\gamma) + \sigma_2^2(1-2\gamma) + \sigma_1^2(1-2\gamma)}$$

As in Merton's case, higher volatilities, i.e. a higher risk, causes the risk averse investor to allocate smaller "optimal" fractions of wealth in risky asset 1. (Note that $1 - \gamma > 0$). For $\bar{u}_2$ however, the case is completely different since

$$\varepsilon_4 = \frac{\varrho_1}{\varrho_2T} + \varrho_3T$$

for $\varrho_i, i = 1, 2, 3$ some other constants. Asymptotically, according to the trading strategy obtained, one would invest infinitely large fractions of wealth in the second asset, while holding a constant fraction of wealth in asset 1. Again, noting that these controls are not optimal by means of Chapter 4, they are only feasible controls, and due to the unconventional, at least, approach of finding them, the solutions are not necessarily financially or mathematically reasonable. In conclusion, we have been able to find a set of controls, not constant as hoped for, but admissible nevertheless, which conceivably imply that, as the time horizon broadens, one should invest a constant fraction of wealth in risky asset 1 and a time-dependent fraction in risky asset 2.

Chapter 6

Concluding Remarks

This thesis has considered the optimal management of portfolios under a co-integrated market model. The investor is assumed to be risk-averse and has the limited choice of investing in two risky assets and one risk free, with the aim of maximizing expected utility of wealth at the end of a finite investment horizon. The market models consists of two risky assets, modelled by a common non-stationary trend process, and to correlated stationary Ornstein-Uhlenbeck processes. The portfolio management problem is presented as an optimization problem from stochastic control theory. By the dynamic programming approach, a semi-explicit solution to the resulting HJB-equation is obtained and presented as the solution to a set of Riccati differential equations. After simplifications of the model, a stochastic Feynman-Kac representation is found. Motivated by Merton's constant fraction solution, a naive approach to find constant controls is presented (with failure, resulting in stochastic controls).

Because the solution of the Riccati system is unknown (might even by non-existing) and because the expressions obtained for the optimal controls in this thesis are so tedious, it is difficult to draw exact conclusions like in the case of Merton's portfolio problem, where the effect of risk is clear from the form of the constant optimal allocation. However, there is clearly a dependency on the correlations and the volatilities from all three sources of uncertainty, and the resulting optimal controls are directly dependent on the current value of the stationary Ornstein-Uhlenbeck processes from the co-integrated market model. The optimal allocation seems to be one which should be updated as the time passes, corresponding to the information revealed by the distinct stationary processes.

A natural continuation of the problem addressed in this thesis, is to explore whether there exists a set of solutions to the Riccati equations, and if so, to try finding them numerically, for instance by the DT method presented in Section 4.6. Analytical solutions seem unlikely, but the DT method is efficient both computationally and in terms of convergence toward the true values of the differential equations. If a set of solutions exists, an interesting extension would be to simulate the optimal portfolio, for instance by fitting to actual data and comparing against existing trading strategies or the Markowitz mean-variance approach.

With regards to the attempt at finding constant controls, one could try to find the closets constant allocation by means of minimizing expected squared error, and then by averaging over the investment horizon to obtain time-independent controls.

Further extensions could be to include stochastic volatilities or several stationary distinctive processes, allowing for more characteristic risk to each company within a market sector. Or on the complete contrary, add multiple common risky sources. However, due to extensive expressions and calculations, this could preferably be done numerically such that clear conclusions on the optimal allocations can be drawn.

Appendices

Appendix A

List of Notational Conventions

A.1 List of Constants and Functions

Stochastic Control Problem

General notation
For $i = 1, 2$

$$\Sigma_i = \sigma^2 + \sigma_i^2 + 2\rho_i\sigma\sigma_i \qquad P_i = \rho_i\sigma\sigma_i + \sigma^2$$
$$\Sigma = \sigma^2 + \rho_1\sigma\sigma_1 + \rho_2\sigma\sigma_2 + \rho\sigma_1\sigma_2 \qquad R_i = \rho_i\sigma\sigma_i + \rho\sigma_1\sigma_2$$
$$\Gamma = \Sigma_1\Sigma_2 - \Sigma^2$$

The optimal controls

$$M_1 = \Sigma_2 P_1 - \Sigma R_1 \qquad M_2 = \Sigma_1 R_1 - \Sigma P_1$$
$$N_1 = \Sigma_2 R_2 - \Sigma P_2 \qquad N_2 = \Sigma_1 P_2 - \Sigma R_2$$

$$G_1(z_1, z_2) = (r - \mu)(\Sigma_2 - \Sigma) - \frac{1}{2}\Sigma_2(\Sigma_1 - \Sigma) - \Sigma_2 A_1 Z_1(t) + \Sigma A_2 Z_2(t)$$
$$G_2(z_1, z_2) = (r - \mu)(\Sigma_1 - \Sigma) - \frac{1}{2}\Sigma_1(\Sigma_2 - \Sigma) + \Sigma A_1 Z_1(t) - \Sigma_1 A_2 Z_2(t)$$

In connection to Problem 4.4.3
For $p(z_1, z_2)$:

$$p(z_1, z_2) = \alpha_0 + \alpha_1 A_1 z_1 + \alpha_2 A_2 z_2 + \alpha_{11} A_1^2 z_1^2 + \alpha_{22} A_2^2 z_2^2 + \alpha_{12} A_1 z_1 A_2 z_2$$

where

$$\begin{aligned}\alpha_0 =& \Gamma\Big((\mu + \frac{1}{2}\Sigma_1 - r)(\Sigma_2\Pi_1 - \Sigma\Pi_2) + (\mu + \frac{1}{2}\Sigma_2 - r)(\Sigma_1\Pi_2 - \Sigma\Pi_1)\Big)\\ &+ K_1(\Sigma_2\Pi_1 - \Sigma\Pi_2)^2 + K_2(\Sigma_1\Pi_2 - \Sigma\Pi_1)^2 + \Sigma(\Sigma_2\Pi_1 - \Sigma\Pi_2)\\ &\times (\Sigma_1\Pi_2 - \Sigma\Pi_1)\end{aligned}$$

$$\begin{aligned}\alpha_1 =& \Gamma\Big(- \Sigma_2(\mu + \frac{1}{2}\Sigma_1 - r) + \Sigma_2\Pi_1 - \Sigma\Pi_2 + \Sigma(\mu + \frac{1}{2}\Sigma_2 - r)\Big)\\ &- 2K_1\Sigma_2(\Sigma_2\Pi_1 - \Sigma\Pi_2) + 2K_2\Sigma(\Sigma_1\Pi_2 - \Sigma\Pi_1) + \Sigma(\Sigma(\Sigma_2\Pi_1\end{aligned}$$

$$- \Sigma\Pi_2) - \Sigma_2(\Sigma_1\Pi_2 - \Sigma\Pi_1))$$

$$\begin{aligned}
\alpha_2 =& \Gamma\Big(\Sigma(\mu + \frac{1}{2}\Sigma_1 - r) - \Sigma_1(\mu + \frac{1}{2}\Sigma_2 - r) + \Sigma\Pi_2 - \Sigma\Pi_1\Big) + 2K_1\Sigma(\Sigma_2\Pi_1 \\
& - \Sigma\Pi_2) - 2K_2\Sigma_1(\Sigma_1\Pi_2 - \Sigma\Pi_1) + \Sigma\Big(\Sigma(\Sigma_1\Pi_2 - \Sigma\Pi_1) \\
& - \Sigma_1(\Sigma_2\Pi_1 - \Sigma\Pi_2)\Big) \\
\alpha_{11} =& -\Gamma\Sigma_2 + K_1\Sigma_2^2 + K_2\Sigma^2 - \Sigma^2\Sigma_2 \\
\alpha_{22} =& -\Gamma\Sigma_1 + K_1\Sigma^2 + K_2\Sigma_1^2 - \Sigma^2\Sigma_1 \\
\alpha_{12} =& \Gamma\Sigma^2 - 2K_1\Sigma_2\Sigma - 2K_2\Sigma_1\Sigma + \Sigma(\Sigma^2 - \Sigma_1\Sigma_2)
\end{aligned}$$

For $p_1(z_1, z_2)$

$$p_1(z_1, z_2) = \beta_0 + \beta_1 A_1 z_1 + \beta_2 A_2 z_2$$

where

$$\begin{aligned}
\beta_0 =& \Gamma\Big(P_1(\Sigma_2\Pi_1 - \Sigma\Pi_2) + R_1(\Sigma_1\Pi_2 - \Sigma\Pi_1) - (\Sigma_2 P_1 - \Sigma R_1)(\mu + \frac{1}{2}\Sigma_1 - r) \\
& - (\Sigma_1 R_1 - \Sigma P_1)(\mu + \frac{1}{2}\Sigma_2 - r)\Big) - 2K_1(\Sigma_2 P_1 - \Sigma R_1)(\Sigma_2\Pi_1 - \Sigma\Pi_2) \\
& - 2K_2(\Sigma_1 R_1 - \Sigma P_1)(\Sigma_1 P_2 - \Sigma\Pi_1) - \Sigma\Big((\Sigma_1 R_1 - \Sigma P_1)(\Sigma_2\Pi_1 - \Sigma\Pi_2) \\
& + (\Sigma_2 P_1 - \Sigma R_1)(\Sigma_1\Pi_2 - \Sigma\Pi_1)\Big) \\
\beta_1 =& -\Sigma_2\Big(\Gamma P_1 - 2K_1(\Sigma_2 P_1 - \Sigma R_1) - \Sigma(\Sigma_1 R_1 - \Sigma P_1)\Big) \\
& + \Sigma\Big(\Gamma R_1 - 2K_2(\Sigma_1 R_1 - \Sigma P_1) - \Sigma(\Sigma_2 P_1 - \Sigma R_1)\Big) - \Gamma(\Sigma_2 P_1 - \Sigma R_1) \\
\beta_2 =& \Sigma\Big(\Gamma P_1 - 2K_1(\Sigma_2 P_1 - \Sigma R_1) - \Sigma(\Sigma_1 R_1 - \Sigma P_1)\Big) \\
& - \Sigma_1(\Gamma R_1 - 2K_2(\Sigma_1 R_1 - \Sigma P_1) - \Sigma(\Sigma_2 P_1 - \Sigma R_1)) - \Gamma(\Sigma_1 R_1 - \Sigma P_1)
\end{aligned}$$

For $p_2(z_1, z_2)$

$$p_2(z_1, z_2) = \delta_0 + \delta_1 A_1 z_1 + \delta_2 A_2 z_2$$

where

$$\begin{aligned}
\delta_0 =& \Gamma\Big(R_2(\Sigma_2 Pi_1 - \Sigma\Pi_2) + P_2(\Sigma_1\Pi_2 - \Sigma\Pi_1) - (\Sigma_2 R_2 - \Sigma P_2)(\mu + \frac{1}{2}\Sigma_1 - r) \\
& - (\Sigma_1 P_2 - \Sigma R_2)(\mu + \frac{1}{2}\Sigma_2 - r)\Big) - \Sigma\Big((\Sigma_1 P_2 - \Sigma R_2)(\Sigma_2\Pi_1 - \Sigma\Pi_2) \\
& + (\Sigma_2 R_2 - \Sigma P_2)(\Sigma_1\Pi_2 - \Sigma\Pi_1)\Big) - 2K_1(\Sigma_2\mathbb{R}_2 - \Sigma P_2)(\Sigma_2\Pi_1 - \Sigma\Pi_2) \\
& - 2K_2(\Sigma_1 P_2 - \Sigma R_2)(\Sigma_1\Pi_2 - \Sigma\Pi_1) \\
\delta_1 =& -\Sigma_2\Big(\gamma R_2 - 2K_1(\Sigma_2 P_2 - \Sigma P_2) - \Sigma(\Sigma_1 P_2 - \Sigma R_2)\Big)
\end{aligned}$$

$$+\Sigma\Big(\Gamma P_2 - 2K_2(\Sigma_1 P_2 - \Sigma R_2) - \Sigma(\Sigma_2 R_2 - \Sigma P_2)\Big) - \Gamma(\Sigma_2 R_2 - \Sigma P_2)$$

$$\begin{aligned}\delta_2 =& \Sigma\Big(\Gamma R_2 - 2K_1(\Sigma_2 P_2 - \Sigma P_2) - \Sigma(\Sigma_1 P_2 - \Sigma R_2)\Big) \\ &- \Sigma_1(\Gamma P_2 - 2K_2(\Sigma_1 P_2 - \Sigma R_2) - \Sigma(\Sigma_2 R_2 - \Sigma P_2)\Big) - \Gamma(\Sigma_1 P_2 - \Sigma R_2)\end{aligned}$$

We also have the following constants

$$\begin{aligned}C_1 =& P_1^2\Big(K_1\Sigma_2^2 - \Gamma\Sigma_2 + K_2\Sigma^2 - \Sigma^2\Sigma_2\Big) + R_1^2\Big(K_1\Sigma^2 - \Gamma R_1^2 + K_2\Sigma^2 - \Sigma^2\Sigma_1\Big) \\ &+ P_1 R_1 \Sigma\Big(\Gamma\Sigma - 2K_1\Sigma_2 - 2K_2\Sigma_1 + \Sigma_1\Sigma_2 - \Sigma^2\Big) \\ C_2 =& P_2^2\Big(K_1\Sigma^2 + K_2\Sigma_1^2 - \Gamma\Sigma_1 - \Sigma^2\Sigma_1\Big) + R_2^2\Big(K_1\Sigma_2^2 - \Gamma\Sigma_2 - K_2\Sigma_2^2 - \Sigma^2\Sigma_2\Big) \\ &+ P_2 R_2 \Sigma\Big(\Gamma\Sigma - 2K_1\Sigma_2 - 2K_2\Sigma_1 + \Sigma_1\Sigma_2 - \Sigma^2\Big) \\ C_3 =& P_1 P_2 \Sigma\Big(\Gamma\Sigma - 2K_1\Sigma_2 - 2K_2\Sigma_1 + \Sigma_1\Sigma_2 + \Sigma^2\Big) + 2P_1 R_2\Big(-\Gamma\Sigma_2 + K_1\Sigma_2^2 \\ &+ K_2\Sigma^2 - \Sigma^2\Sigma_2\Big) + 2R_1 P_2\Big(-\Gamma\Sigma_1 + K_1\Sigma^2 + K_2\Sigma_1^2 - \Sigma^2\Sigma_1\Big)\end{aligned}$$

Riccati representation

$$\zeta = \frac{\gamma}{\Gamma^2(\gamma - 1)}$$

.

$$C = \begin{bmatrix} 0 & \beta_0\zeta & \delta_0\zeta & \sigma_1^2 & \sigma_2^2 & \rho\sigma_1\sigma_2 \\ 0 & A_1+\beta_1 A_1\zeta & \delta_1 A_1\zeta & 2\beta_0\zeta & 0 & \delta_0\zeta \\ 0 & \beta_2 A_2\zeta & A_2+\delta_2 A_2\zeta & 0 & 2\delta_0\zeta & \beta_0\zeta \\ 0 & 0 & 0 & 2A_1+2A_1\beta_1\zeta & 0 & \delta_1 A_1\zeta \\ 0 & 0 & 0 & 0 & 2A_2+2\delta_2 A_2\zeta & \beta_2 A_2\zeta \\ 0 & 0 & 0 & 2\beta_2 A_2\zeta & 2\delta_1 A_1\zeta & (A_1+A_2 \\ & & & & & +\beta_1 A_1\zeta+\delta_2 A_2\zeta) \end{bmatrix}$$

$$G = \begin{bmatrix} g_0^2(t) & g_0(t)g_1(t) & \cdots & g_0(t)g_{12}(t) \\ g_1(t)g_0(t) & g_1^2(t) & \cdots & g_1(t)g_{12}(t) \\ \vdots & & \ddots & \\ g_{12}(t)g_0(t) & g_{12}(t)g_1(t) & \cdots & g_{12}^2(t) \end{bmatrix}$$

$$\mathbf{N} = \left(\gamma r + \alpha_0\zeta,\ \alpha_1 A_1\zeta,\ \alpha_2 A_2\zeta,\ \alpha_{11} A_1^2\zeta,\ \alpha_{22} A_2^2\zeta,\ \alpha_{12} A_1 A_2\zeta\right)^T$$

$$\epsilon_i = \frac{1}{2}\sigma_i^2 + C_i\delta, \quad , \psi = \rho\sigma_1\sigma_2 + C_3\delta,$$

for $i = 1, 2$.

$$M(\mathbf{g}(t)) = \begin{bmatrix} 0 & -g_1\epsilon_1 & -(g_2\epsilon_2+g_1\psi) & 0 & 0 & 0 \\ 0 & -4g_{11}\epsilon_1 & -2g_{12}(t)\epsilon_2 & -2g_2\psi & 0 & -g_1\psi \\ 0 & -2g_{22}\psi & -g_{12}\psi & 0 & -4g_2\epsilon_2 & -2g_1\epsilon_1 \\ 0 & 0 & 0 & -4g_{11}\epsilon_1 & 0 & -(g_{12}\epsilon_2-2g_{11}\psi) \\ 0 & 0 & 0 & 0 & -4g_{22}\epsilon_2 & -(g_{12}\epsilon_1+2g_{22}\psi) \\ 0 & 0 & 0 & -2g_{12}\epsilon_1 & -(4g_{11}\psi+4g_{12}\epsilon_2) & -g_{12}\psi \end{bmatrix}$$

"Constant" Controls

For the constants $\nu(u_1, u_2)$ and $\widetilde{\nu}(u_1, u_2)$

$$\begin{aligned} \nu(u_1, u_2) =& u_1(\mu + \frac{1}{2}\Sigma_1) + u_2(\mu + \frac{1}{2}\Sigma_2) + (1 - u_1 - u_2)r - \frac{1}{2}\Big(\sigma^2(u_1 + u_2)^2 \\ &+ \sigma_1^2 u_1^2 + \sigma_2^2 u_2^2\Big) - \sigma(u_1 + u_2)(\rho_1\sigma_1 u_1 + \rho_2\sigma_2 u_2) - \rho\sigma_1\sigma_2 u_1 u_2 \\ \widetilde{\nu}(u_1, u_2) =& \gamma\nu(u_1, u_2) + \frac{1}{2}\gamma^2\Big(u_1^2(\sigma^2 + 2\sigma_1^2) + u_2^2(\sigma^2 + 2\sigma_2^2) + 2u_1u_2(\rho\sigma_1\sigma_2 \\ &+ \sigma^2)\Big) \end{aligned}$$

For F_i, F_i^2 and $F_{i,j}$ for $i = 1, 2$, $i \neq j$

$$\begin{aligned} F_i =& \frac{1}{A_i}\Big[e^{A_i(T-t)} - 1\Big] \\ F_i^2 =& \frac{1}{2A_i}\Big[e^{2A_i(T-t)} - 1\Big] \\ F_{i,j} =& \frac{1}{A_i + A_j}\Big[e^{(A_i+A_j)(T-t)} - 1\Big] \end{aligned}$$

Section 5.4

For the constants, U_i, L_i, K_i, $i = 1, 2$, $i \neq j$

$$\begin{aligned} U_i =& \Sigma_i T - \gamma\Big((\sigma^2 + 2\sigma_i^2)(T - t) + \sigma_i^2 F_i^2 + \rho_i\sigma\sigma_i F_i\Big) \\ L_i =& \Big(\Sigma T - \gamma\Big(P_j(T - t) + \rho\sigma_i\sigma_j(F_{i,j} - F_i - F_j) + \frac{1}{2}(\rho_i\sigma\sigma_i F_i + \rho_j\sigma\sigma_j F_j)\Big)\Big) \\ K_i =& (\mu + \frac{1}{2}\Sigma_i - r)T \end{aligned}$$

For ϑ_i, $i = 1, \ldots, 6$

$$\vartheta_1 = \frac{2}{U_1} \qquad \vartheta_2 = -\frac{L_2}{(L_1L_2 - U_1U_2)}$$

$$\vartheta_3 = \frac{1}{U_1} - \frac{L_2^2 U_1 - L_2}{U_1(L_1L_2 - U_1U_2)} \qquad \vartheta_4 = \frac{1}{U_1} + \frac{L_2^2}{U_1(L_1L_2 - U_1U_2)}$$

$$\vartheta_5 = -\frac{L_2}{L_1L_2 - U_1U_2} \qquad \vartheta_6 = \frac{L_2(L_2K_1 - K_2)}{U_1(L_1L_2 - U_1U_2)}$$

For ε_i, $i = 1, \ldots, 4$

$$\varepsilon_1 = -\frac{L_2}{L_1L_2 - U_1U_2} \qquad \varepsilon_2 = \frac{U_1(1 - L_2)}{L_1L_2 - U_1U_2}$$

$$\varepsilon_3 = -\frac{L_2}{L_1L_2 - U_1U_2} \qquad \varepsilon_4 = -\frac{L_2}{L_1L_2 - U_1U_2} + K_2$$

Appendix B

Preliminaries

This appendix provides some stochastic preliminaries covered in the thesis. A selection of definitions, theorems and properties are presented. The main references for this chapter are [Øks13], [KS91] and [YZ99]. All statements are made in the 1-dimensional case, but can easily by expanded to include processes in $\mathbb{R}^n$.

B.1 General Preliminaries

Let $(\Omega, \mathcal{F}, \mathcal{P})$ denote a probability space for which a filtration $\{\mathcal{F}\}_{t\geq 0}$ is given. The following are standing assumption

(i) $(\Omega, \mathcal{F}, \mathcal{P})$ is complete, meaning that for any $\mathcal{P}$-null set A, i.e. $\mathcal{P}(A) = 0$, in $\mathcal{F}$, $B \subset \mathcal{F}$ whenever $B \subset A$

(ii) $\mathcal{F}_0$ contains all null sets of $\mathcal{F}$

(iii) the filtration $\{\mathcal{F}_t\}_{t\geq 0}$ is right-continuous, i.e. $\mathcal{F}_t = \bigcap_{s>t} \mathcal{F}_s$ for all $t \in [0, T)$.

We denote by $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathcal{P})$ a filtered probability space. To "set the scene", we recall the definition of a Brownian motion, even though the reader is assumed to be familiar with it already.

Definition B.1.1.
A stochastic process $\{B(t)\}_{t\geq 0}$ is said to be a standard Brownian motion, if for $0 \leq s < t < \infty$ it satisfies the following properties

(i) $B(0) = 0$ *$\mathcal{P}$-a.s.*

(ii) $\{B(t)\}_{t\geq 0}$ *has independent increments, i.e. $B(t) - B(s)$ is independent of $\mathcal{F}_s$*

(iii) $\{B(t)\}_{t\geq 0}$ *has normally distributed increments with mean 0 and variance given by (t-s), i.e. the random variable $B(t) - B(s) \sim \mathcal{N}(0, t-s) \forall t > s$*

From now on we denote by $B(t)$ both the process and the random variable. Since the Itô calculus is the cornerstone of stochastic analysis, we recall some properties of the Itô integrals.

Definition B.1.2 (Itô integrals).
The Itô integral

$$\int_0^T f(s,\omega)dB(s,\omega) \tag{B.1}$$

is well-defined for functions $f : [0,T] \times \Omega \to \mathbb{R}$ *satisfying the following assumptions*

(i) $(t,\omega) \to f(t,\omega)$ *is* $\mathcal{B}([0,T]) \times \mathcal{F}$*-measurable, where* $\mathcal{B}([0,T])$ *denotes the Borel* σ*-algebra on* $[0,T]$

(ii) $f(t,\omega)$ *is adapted w.r.t. the filtration* $\mathcal{F}_t$

(iii) $f \in L^2((\Omega,\mathcal{F},\mathcal{P}))$*, i.e.*

$$\int_0^T E[f^2(s,\omega)]ds < \infty$$

We denote by $\mathcal{V}([0,T])$ *the class of functions* f *for which the Itô integral is well defined.*

Corollary B.1.3 (Itô isometry).
Let $f \in \mathcal{V}([0,T])$*, then*

$$E\Big[\Big(\int_0^T f(s,\omega)dB(s,\omega)\Big)^2\Big] = \int_0^T E\Big[f^2(s,\omega)ds\Big]ds \tag{B.2}$$

We refer to [Øks13] for proof of Corollary B.1.3. Following are some properties that hold for all Itô integrable functions $f, g \in \mathcal{V}([0,T])$.

Proposition B.1.4 ().

(i) *the Itô integrals satisfy the linearity property, i.e.*

$$\int_0^T \alpha(f(s,\omega) + g(s,\omega))dB(s) = \alpha \int_0^T f(s,\omega)dB(s) + \int_0^T g(s,\omega)dB(s)$$

(ii)

$$\int_0^T f(s,\omega)dB(s) = \int_0^t f(s,\omega)dB(s) + \int_t^T f(s,\omega)dB(s)$$

for $0 < t < T$

(iii)

$$E\Big[\int_0^T f(s,\omega)dB(s)\Big] = 0$$

(iv) *the Itô integral*

$$\int_t^T f(s,\omega)dB(s), \quad \textit{for } 0 < t < T \tag{B.3}$$

if $\mathcal{F}_T$*-measurable and independent of* $\mathcal{F}_t$

The properties above follow from the fact that they all hold for elementary functions. By taking limits in $L^2([0,T]\times\Omega)$, they must hold for all $f \in \mathcal{V}([0,T])$.

Definition B.1.5 (Martingale).
A $\mathcal{F}_t$-adapted stochastic process $\{M(t)\}_{t\geq 0}$ is called a martingale if

(i) $E[|M(t)|] < \infty$ *for all* t

(ii) $E[M(t)|\mathcal{F}_s] = M(s)$ *for* $s < t$

Note that the Itô integral, $\int_0^t f(s,\omega)dB(s)$ for $f \in \mathcal{V}([0,T])$ is a martingale.

Definition B.1.6 (Semimartingale).
A stochastic process $\{X(t)\}_{t\geq 0}$ is called a continuous semimartingale w.r.t. the filtration $\{\mathcal{F}\}_{t\geq 0}$ if $X(t)$ can be decomposed as

$$X(t) = x + \int_0^t Z(s)ds + \int_0^t Y(s)dB(s) \tag{B.4}$$

for $Z(t)$ and $Y(t)$ $\mathcal{F}_t$-adapted processes.

Note that, for $Z(t) = 0$, the process $X(t)$ is a martingale. The following lemma is used repeatedly throughout the thesis.

Lemma B.1.7 (Itô's lemma).
Assume $X(t)$ is a semimartingale and let $f(t,x) \in C^{1,2}$, i.e. f is twice continuously differentiable w.r.t. x and continuously differentiable w.r.t. t. Then

$$df(X(t)) = \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial x}dX(t) + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(dX(t))^2 \tag{B.5}$$

$$= \frac{\partial f}{\partial t}dt + Z(s)\frac{\partial f}{\partial x}ds + \frac{1}{2}Y^2(t)\frac{\partial^2 f}{\partial x^2}dt + Y(t)\frac{\partial f}{\partial x}dB(t) \tag{B.6}$$

A proof of Itô's lemma can be found in [KS91].

B.2 Preliminaries Relevant to Stochastic Control Theory

The following are some preliminaries relevant to the theory of stochastic control applied in the thesis. We begin with the notion of an Itô diffusion. Roughly speaking, an Itô diffusion is a solution to a stochastic differential equation (SDE).

Definition B.2.1 (Itô diffusions).
Given a stochastic differential equation on the form

$$dX(t) = b(t, X(t))dt + \sigma(t, X(t))dB(t), \quad t \geq s, \quad X(s) = x \tag{B.7}$$

where $b : [0,T] \to \mathbb{R}$, $\sigma : [0,T] \to \mathbb{R}$ satisfying the Lipschitz condition

$$|b(x) - b(y)| + |\sigma(x) - \sigma(y)| \leq K|x - y| \quad \text{for } x, y \in \mathbb{R}$$

then the stochastic process $\{X(t)\}_{t\geq 0}, X(t,\omega) : [0,T] \times \Omega \to \mathbb{R}$ satisfying (B.7) is defined as an Itô diffusion.

Assuming a solution exists, the Lipschitz condition ensures taht the Itô diffusion is the unique solution. To each Itô diffusion there can be associated a second order partial differential generator A.

Definition B.2.2 (Infinitesimal Generator).
Let $X(t)$ be the Itô diffusion satisfying the stochastic differential eqution (B.7). The infinitesimal generator A of $X(t)$ is defined by

$$(Af)(x) = \lim_{t\to 0} \frac{E^x[f(X(t))] - f(x)}{t} \tag{B.8}$$

for $X(0) = x$ and where E^x denotes the expectation w.r.t. the probability measure $\mathcal{P}$ for the Itô diffusion starting at $x \in \mathbb{R}$.

Definition B.2.3 ($\mathcal{D}_A(x)$).
The set of functions f for which the generator of $X(t)$ is well-defined for all $x \in \mathbb{R}$, is denoted $\mathcal{D}_A$.

It can be shown that any function $f \in C^{1,2}(\mathbb{R})$ with compact support, meaning $f(y) = 0$ for some y outside a compact set, is in $\mathcal{D}_A$. See for instance [Øks13]. Then the generator $(Af)(x)$ can be represented on a partial differential form

$$(Af)(x) = b(x)\frac{\partial f}{\partial x} + \frac{1}{2}\sigma^2\frac{\partial^2 f}{\partial x^2} \tag{B.9}$$

The following formula is relevant for the verification proof of the HJB-equation.

Theorem B.2.4 (Dynkin's formula).
Let $f \in C_0^2(\mathbb{R})$, i.e. f is twice continuously in $\mathbb{R}$ and has compact support, and assume τ is a stopping time, $E[\tau] < \infty$. Then the following holds

$$E^x[f(X(\tau))] = f(x) + E^x\Big[\int_0^\tau (Af)(X(s))ds\Big] \tag{B.10}$$

The Itô diffusions satisfy two important properties; the Markov property and the strong Markov property. A stochastic process satisfying the Markov property "looses it's memory," meaning the process only "remembers" the current state. Stated in another way, the current value $X(t)$ of a Markov process contains all information need for the future evolution of the process. The future state of $X(s)$, for $s > t$, only depends on the current state $X(t)$. Formally, the Markov process is defined as follows

Definition B.2.5 (The Markov property).
A stochastic process $X(t,\omega) : [0,T] \times \Omega \to \mathbb{R}$ is called a Markov process if

$$\mathcal{P}(X(t) \in O | X(t_1), \ldots, X(t_n)) = \mathcal{P}(X(t) \in O | X(t_n)), \quad \mathcal{P} - a.s. \tag{B.11}$$

for some open set $O \subset \mathcal{B}(\mathbb{R})$, and times $0 \leq t_1 < t_2 < \ldots < t_n < t \leq T$.

If the stochastic process $X(t)$ is a $\mathcal{F}_t$ Markov process, then the following is another characterization of the Markov property

$$E^x[f(X(t))|\mathcal{F}_s] = E^x[f(X(t))|X(s)] \tag{B.12}$$

for $t > s$ and f bounded and measurable. Itô diffusions are some times called time-homogeneous strong Markov processes. Since the Itô diffusions are important Markov processes, we give the following Markov property special for Itô diffusion

Theorem B.2.6.
Let $X(t)$ be an Itô diffusion and $\mathcal{F}_t$ the filtration generated by the Brownian motion $B(t)$. Then the following holds for $t, h \geq 0$

$$E^x[f(X(t))|\mathcal{F}_s] = E^{x_s}[f(X(t))]_{|X(s)=x_s} \tag{B.13}$$

where $f : \mathbb{R} \to \mathbb{R}$ bounded and measurable.

A proof of this and the following theorem can be found in [Øks13]. As the future evolution of $X(t)$ only depends on the current value of the process, the future behaviour is independent of the initial starting point of the process. There is a similar property for stopping times, called the strong Markov property, but first we need the definition of a stopping time.

Definition B.2.7 (Stopping time).
A random variable $\tau : \Omega \to [0, T]$ is called a stopping time w.r.t. the filtration $\{\mathcal{F}_t\}_{t\geq 0}$ if

$$\{\omega : \tau(\omega) \leq t\} \in \mathcal{F}_t \tag{B.14}$$

Theorem B.2.8 (The strong Markov property).
Let $X(t)$ be an Itô diffusion and $\mathcal{F}_t$ the filtration generated by $B(t)$. Assume $\tau < \infty$ is a stopping time w.r.t. $\mathcal{F}_t$. Then the following holds for $t, h \geq 0$

$$E^x[f(X(\tau + h))|\mathcal{F}_\tau] = E^{x_\tau}[f(X(h))]_{|X(\tau)=x_\tau}] \tag{B.15}$$

From the definition of a stopping time, we obtain the following notion of the first exit time of a domain D

Definition B.2.9 (First exit time).
Let D denote the domain of some process $X(t)$. The first exit time τ_D of D for$X(t)$ is defined by

$$\tau_D = \inf\{t > 0 : X(t) \neq D\} \tag{B.16}$$

In connection to the domain D, we have the following definition of regular points.

Definition B.2.10 (Regular points).
Let D be a open connected set. Then a point $x \in \partial D$, i.e. a point in the boundary of D, is called regular w.r.t. the Itô diffusion $X(t)$, if $\mathcal{P}^{\S}(\tau_D = 0) = 1$, where $\mathcal{P}$ is the probability law of $X(t)$ starting at x.

Essentially, a regular point x on the boundary of some domain D, is a point for which the process $X(t)$ leaves D immediately if starting at x. A point for which $\mathcal{P}(\tau_D = 0) < 1$, is called irregular.

The Dirichlet Problem

The following problem is denoted as the Dirichlet problem and becomes useful in the proof of Theorem 4.2.7

Problem B.2.11 (Dirichlet Problem).
Given a bounded domain $D \in R$, the operator generator $L = A$ and a continuous function defined on ∂D, i.e. $\phi \in C(\partial D)$, we wish to find a continuous solution $u \in C^2(D)$ satisfying

(i) $(Lu)(x) = 0$ *for all* $x \in D$

(ii) $u = \phi$ *on* ∂D

It can be shown, see for instance [Øks13] or [KS91] that the function

$$u(x) = E^x[\phi(X(\tau_D))] \tag{B.17}$$

is a solution to the Dirichlet problem if u is continuous, $u \in C^2(D)$ and $E^x[\tau_D] < \infty$. For the stochastic representation of the function f in Section 5.9, the following important theorem is crucial.

Theorem B.2.12 (Feynman-Kac representation).
Suppose $f(t,x) : [0,T] \times \mathbb{R} \to \mathbb{R}$ is continuous, $f \in C^{1,2}([0,T] \times \mathbb{R}$ and satisfies the following Cauchy problem

$$-\frac{\partial v}{\partial t} + kv = \mathcal{A}_t v + g, \qquad \text{in } [0,T) \times \mathbb{R}^d \tag{B.18}$$

$$v(T,x) = f(x), \qquad \text{for } x \in \mathbb{R}^d \tag{B.19}$$

as well as the polynomial growth condition

$$\max_{0 \le t \le T} |v(t,x)| \le M(1 + \|x\|^{2\mu}), \qquad \text{for } x \in \mathbb{R}^d \tag{B.20}$$

for some $M > 0, \mu \ge 1$. If so, then the solution $v(t,x)$ must be representable by the following stochastic representation

$$v(t,x) = E^{t,x}\Big[f(X(T)) \exp\Big(-\int_t^T k(\theta, X(\theta)) d\theta\Big) \tag{B.21}$$

$$+ \int_t^T g(s, X(s)) \exp\Big(-\int_t^s k(\theta, X(\theta))\Big) ds\Big] \tag{B.22}$$

on $[0,T] \times \mathbb{R}^d$.

B.3 Some Preliminaries from Measure Theory

The following theorem is taken from [McD13].

Theorem B.3.1 (Fubini's theorem).
Let μ_1, μ_2 be probability measures on $(\Omega_1, \mathcal{F}_1)$ and $(\Omega_2, \mathcal{F}_2)$, respectively. Furthermore, let $f : \Omega_1 \times \Omega_2 \to \mathbb{R}$ be a $\mathcal{F}_1 \times \mathcal{F}_2$-measurable function and assume that at least one of the following integrals is finite

(i)
$$\int_{\Omega_1 \times \Omega_2} |f(x,y)| d(\mu_1 \times \mu_2)(x,y)$$

(ii)

$$\int_{\Omega_1}\Big(\int_{\Omega_2}|f(x,y)|d\mu_2(y)\Big)d\mu_1(x)$$

(iii)

$$\int_{\Omega_2}\Big(\int_{\Omega_1}|f(x,y)|d\mu_1(x)\Big)d\mu_2(y)$$

then the following equality holds

$$\begin{aligned}\int_{\Omega_1\times\Omega_2} f(x,y)d(\mu_1\times\mu_2)(x,y) &= \int_{\Omega_2}\Big(\int_{\Omega_1} f(x,y)d\mu_1(x)\Big)d\mu_2(y)\\ &= \int_{\Omega_1}\Big(\int_{\Omega_2} f(x,y)d\mu_2(y)\Big)d\mu_1(x)\end{aligned}$$

Remark B.3.2 (Fubini on expectations).
Note that the following is a direct consequence of Fubini's theorem for measurable functions f, if either $E[\int_0^T f(s,\omega)ds]$ or $\int_0^T E[f(s,\omega)]ds < \infty$

$$E\Big[\int_0^T f(s,\omega)ds\Big] = \int_\Omega\int_0^T f(s,\omega)dsdP(\omega) \tag{B.23}$$

$$= \int_0^T\int_\Omega f(s,\omega)dP(\omega)ds = \int_0^T E\Big[f(s,\omega)\Big]ds \tag{B.24}$$

Another version of Fubini's theorem, with the same conclusion, but with the assumption that f is a non-negative measurable functions, instead of the assumption that $|f|$ is integrable, is the following Tonelli theorem

Theorem B.3.3 (Tonelli's theorem).
Let $(\Omega_1, \mathcal{F}_1, \mu_1)$ and $(\Omega_2, \mathcal{F}_2, \mu_2)$ be the two probability spaces from Theorem B.3.1, and let $f : \Omega_1 \times \Omega_2 \to \mathbb{R}$ be a non-negative measurable function. Then the following equality holds

$$\begin{aligned}\int_{\Omega_1\times\Omega_2} f(x,y)d(\mu_1\times\mu_2)(x,y) &= \int_{\Omega_2}\Big(\int_{\Omega_1} f(x,y)d\mu_1(x)\Big)d\mu_2(y)\\ &= \int_{\Omega_1}\Big(\int_{\Omega_2} f(x,y)d\mu_2(y)\Big)d\mu_1(x)\end{aligned}$$

The stochastic Fubini theorem, states that, under suitable assumptions, we are allowed to interchange the order of a Lebesgue integral and an integral w.r.t. a semimartingale, but first we need the notion of a progressively measurable process.

Definition B.3.4 (Progressively measurable process).
A stochastic process $X(t,\omega)$ in $\mathbb{R}$ is said to be progressively *measurable w.r.t. the filtration $\mathcal{F}_t$ if for each fixed $t \geq 0$ the mapping $([0,t]\times\Omega) \to \mathcal{B}([0,t]) \times \mathcal{F}_\sqcup$ defined by $(s,\omega) \to X(s,\omega)$ for $s \leq t$ is $\mathcal{B}([0,t]) \times \mathcal{F}_t$-measurable. We say that X is* Prog_T*-measurable, where*

$$\text{Prog}_T = \{A \in \mathcal{B}([0,T]) \times \mathcal{F}_T\} \tag{B.25}$$

The following theorem can be found in [Fil09].

Theorem B.3.5 (Stochastic Fubini).
Let $\phi(\omega, t, s)$ *be a stochastic process for* $0 \leq t,\ s \leq T$*, satisfying the following properties*

(i) ϕ *is* $\mathrm{Prog}_T \times \mathcal{B}([0,T])$*-measurable*

(ii) $\sup_{t,s}\|\phi(t,s)\| < \infty$

then

$$\int_0^T \Big(\int_0^T \phi(t,s)dB(t)\Big)ds = \int_0^T \Big(\int_0^T \phi(t,s)ds\Big)dB(t) \tag{B.26}$$

for $B(t)$ *a Brownian motion.*

Proposition B.3.6 (Hölder's inequality).
Let f *and* g *be two measurable functions. Then*

$$\int_\Omega |f(\omega)g(\omega)|d\mathcal{P}(\omega) \leq \Big(\int_\Omega |f(\omega)|^p d\mathcal{P}(\omega)\Big)^{1/p}\Big(\int_\Omega |g(\omega)|^q d\mathcal{P}(\omega)\Big)^{1/q} \tag{B.27}$$

for $p, q > 0$ *if* $\frac{1}{p} + \frac{1}{q} = 1$.

Note that for $p = q = 2$, we have the Cauchy-Schwarz inequality. [Øks13] refers to [Chu74] for the following result.

Theorem B.3.7 (Jensen's inequality).
Let X *be a random variable, and suppose* $g(X) : \mathbb{R} \to \mathbb{R}$ *is a convex and* $E[g(X)] < \infty$*, then the following inequality holds*

$$g(E[X|\mathcal{H}]) \leq E[g(X)|\mathcal{H}] \tag{B.28}$$

for any σ*-algebra* $\mathcal{H}$*.*

B.4 Some Statistical Properties

Definition B.4.1 (A log-normal random variable).
A random variable $X = \exp(Y)$ *is said to be a log-normal random variable if* Y *is normal with mean* μ *and variance* σ^2*. The expectation and variance of a log-normal random variable are given by*

$$E[X] = \exp(\mu + \frac{1}{2}\sigma^2) \tag{B.29}$$

$$Var(X) = [\exp(\sigma^2) - 1]\exp(2\mu + \sigma^2) \tag{B.30}$$

We state some properties of the expected value of a random variable from [Øks13], which become handy in Chapter 6.

Proposition B.4.2 (Properties of the conditional expectation).
Let $X(t,\omega) : [0,T] \times \Omega \to \mathbb{R}$ *be a random variable with* $E[|X|] < \infty$*. The conditional expectation* X *given* $\mathcal{G} \subset \mathcal{F}$ *is given by the new random variable* $Y = E[X|\mathcal{G}]$*. The following properties hold for the conditional expectations*

(i) $E[X|\mathcal{G}]$ *is* $\mathcal{G}$*-measurable*

(ii) $E[aX + bZ|\mathcal{G}] = E[aX|\mathcal{G}] + E[bZ|\mathcal{G}]$, *for* $a, b \in \mathbb{R}$ *and* $Z(t, \omega) : [0, T] \times \Omega \to \mathbb{R}$ *a random variable*

(iii) $E[E[X|\mathcal{G}]] = E[X]$

(iv) $E[X|\mathcal{G}] = X$ *if* X *is* $\mathcal{G}$*-measurable*

(v) $E[X|\mathcal{G}] = E[X(t)]$ *if* $X(t)$ *is independent of*$\mathcal{G}$

(vi) $E[E[X|\mathcal{H}]|\mathcal{G}]$ *for some* $\mathcal{H} \supseteq \mathcal{G}$.

Appendix C

Calculations and Proofs

C.1 The Differentials

Lemma C.1.1 (The differentials).

$$
\begin{aligned}
dW_t \cdot dW_t = {} & W_t^2\big(\sigma^2(u_1(t)+u_2(t))^2 + \sigma_1^2 u_1(t)^2 + \sigma_2^2 u_2(t)^2\Big)dt \\
& + 2W_t^2\Big(\sigma(u_1+u_2)(\rho_1\sigma_1 u_1(t) + \rho_2\sigma_2 u_2(t)) \\
& + \rho\sigma_1 u_1(t)\sigma_2 u_2(t)\Big)dt
\end{aligned} \tag{C.1}
$$

$$
\begin{aligned}
dW_t \cdot dZ_1(t) = {} & W_t\Big(\rho_1\sigma(u_1(t)+u_2(t))\sigma_1 + \sigma_1^2 u_1(t) \\
& + \rho\sigma_1\sigma_2 u_2(t)\Big)dt
\end{aligned} \tag{C.2}
$$

$$
\begin{aligned}
dW_t \cdot dZ_2(t) = {} & W_t\Big(\rho_2\sigma(u_1(t)+u_2(t))\sigma_2 + \sigma_2^2 u_2(t) \\
& + \rho\sigma_1\sigma_2 u_1(t)\Big)dt
\end{aligned} \tag{C.3}
$$

$$
dZ_1(t) \cdot dZ_2(t) = \sigma_1\sigma_2\rho dt \tag{C.4}
$$

C.2 Short Sketch of the Computations for the Optimal Controls

By solving the conditions of a optimum for the HJB-equation w.r.t. both v_1 and v_2, the following expressions are obtained

$$
v_1^* = \frac{\Phi_w}{w\Phi_{ww}\Sigma_1}\Big(r - \mu - A_1 z_1 - \frac{1}{2}\Sigma_1\Big) - \frac{\Phi_{wz_1}}{w\Phi_{ww}\Sigma_1}P_1 - \frac{\Phi_{wz_2}}{w\Phi_{ww}\Sigma_1}R_2 - v_2\Sigma
$$

$$
v_2^* = \frac{\Phi_w}{w\Phi_{ww}\Sigma_2}\Big(r - \mu - A_2 z_2 - \frac{1}{2}\Sigma_2\Big) - \frac{\Phi_{wz_1}}{w\Phi_{ww}\Sigma_2}R_2 - \frac{\Phi_{wz_2}}{w\Phi_{ww}\Sigma_2}P_2 - v_1\Sigma
$$

When substituting for v_2 into v_1, the following expression is obtained

$$
v_1^* = \frac{1}{\Sigma_1\Sigma_2(1-\Sigma^2)}\Bigg(\frac{\Phi_w}{w\Phi_{ww}}\Big(\Sigma_2(r - \mu - \frac{1}{2}\Sigma_1 - A_1 z_1) - \Sigma\Sigma_1(r - \mu - \frac{1}{2}\Sigma_2
$$

$$- A_2 z_2)\Big) - \frac{\Phi_{wz_1}}{w\Phi_{ww}}\Big(\Sigma_2 P_1 - \Sigma_1 \Sigma R_1\Big) - \frac{\Phi_{wz_2}}{w\Phi_{ww}}\Big(\Sigma_2 R_2 - \Sigma\Sigma_1 P_2\Big)\Big)$$

Similarly, for v_1^* in v_2, we get

$$v_2^* = \frac{1}{\Sigma_1\Sigma_2(1-\Sigma^2)}\Big(\frac{\Phi_w}{w\Phi_{ww}}\Big(\Sigma_1(r-\mu-\frac{1}{2}\Sigma_2 - A_2 z_2) - \Sigma\Sigma_2(r-\mu-\frac{1}{2}\Sigma_1$$
$$- A_1 z_1)\Big) - \frac{\Phi_{wz_1}}{w\Phi_{ww}}\Big(R_1\Sigma_1 - \Sigma\Sigma_2 P_1\Big) - \frac{\Phi_{wz_2}}{w\Phi_{ww}}\Big(P_2\Sigma_1 - \Sigma\Sigma_2 R_2\Big)\Big)$$

By introducing the abbreviating constants M_i, N_i and the functions G_i for $i = 1, 2$, in Appendix A, we obtain the optimal controls in Proposition 4.4.1.

The calculations done for the substitution of the optimal controls into the HJB-equation would require pages with basic calculations, hence they are omitted. But the resulting expressions were achieved by first rewriting the HJB-equation by collecting all linear and quadratic terms of v_1 and v_2, then by computing the squared expressions, thereby substituting all the expression into the HJB-equation and collecting all the terms corresponding to equal partial derivatives in the functions obtained in Problem 4.4.3 and Appendix A.

C.3 Calculations of Var($Y(T)$) for Proposition 5.4.3

In this section we derive the expression for the variance of $Y(T)$ in Chapter 5. By previous calculations, we have found that $E[Y(T)] = 0$, hence

$$\text{Var}(Y(T)) = E\Big([Y(T)]\Big)^2$$

By Theorem B.3.5, Theorem B.3.1 and Corollary B.1.3, it follows that, for $i, j = 1, 2$

Calculation nr.1

$$E\Big[\Big(\int_t^T\int_u^T u_i A_i \sigma_i e^{A_i(s-u)} ds dB_i(u)\Big)\Big(\int_t^T\int_u^T u_j A_j \sigma_j e^{A_j(s-u)} ds dB_j(u)\Big)\Big]$$
$$= \begin{cases} u_i^2 A_i^2 \sigma_i^2 \int_t^T \Big(\int_u^T e^{A_i(s-u)} ds\Big)^2 du = u_i^2\sigma_i^2\Big(F_i^2 - F_i + T - t\Big) & i = j \\ \rho\sigma_i\sigma_j u_i u_j A_i A_j \int_t^T \Big(\int_u^T e^{A_i(s-u)} ds\Big)\Big(\int_u^T e^{A_j(s-u)} ds\Big) du & \\ \qquad = \rho\sigma_i\sigma_j u_i u_j\Big(F_{i,j} - F_i - F_j + T - t\Big) & i \neq j \end{cases}$$

Calculation nr.2

$$E\Big[\Big(\int_t^T\int_u^T u_i A_i \sigma_i e^{A_i(s-u)} ds dB_i(u)\Big)\Big(\int_t^T \sigma(u_1+u_2) d\widetilde{B}(u)\Big)\Big]$$
$$= \rho_i u_i A_i \sigma_i \sigma(u_i + u_j)\int_t^T(\int_u^T e^{A_i(s-u)} ds) du$$
$$= \rho_i u_i \sigma_i \sigma(u_1+u_2)\Big(F_i - (T-t)\Big) \quad i = 1, 2$$

Calculation nr.3

$$E\Big[\Big(\int_t^T\int_u^T u_i A_i \sigma_i e^{A_i(s-u)} ds dB_i(u)\Big)\Big(\int_t^T \sigma_j u_j dB_j(u)\Big)\Big]$$

$$= \begin{cases} u_i^2 A_i \sigma_i^2 \int_t^T \int_u^T e^{A_i(s-u)} ds du = u_i^2 \sigma_i^2 \Big(F_i - (T-t)\Big) & i = j \\ \rho\sigma_i\sigma_j u_i u_j A_i \int_t^T \int_u^T e^{A_i(s-u)} ds du = \rho\sigma_i\sigma_j u_i u_j \Big(F_i - (T-t)\Big) & i \neq j \end{cases}$$

Calculation nr. 4

$$E\Big[\Big(\int_t^T \sigma(u_i + u_j) d\widetilde{B}(s)\Big)\Big(\int_t^T \sigma_i u_i dB_i(u)\Big)\Big] = \int_t^T \rho_i \sigma \sigma_i u_i (u_1 + u_2) du \qquad \text{for } i = 1, 2$$

Calculation nr. 5

$$E\Big[\Big(\int_t^T \sigma_i u_i dB_i(u)\Big)\Big(\int_t^T \sigma_j u_j dB_j(u)\Big)\Big] = \begin{cases} \int_t^T u_i^2 \sigma_i^2 du & \text{if } i = j \\ \int_t^T \rho\sigma_i\sigma_j u_i u_j du & \text{if } i \neq j \end{cases}$$

Calculation nr. 6

$$E\Big[\Big(\int_t^T \sigma(u_1 + u_2) d\widetilde{B}(u)\Big)^2\Big] = \int_t^T \sigma^2 (u_1 + u_2)^2 du \tag{C.5}$$

Combining calculations (1) to (6), we obtain that the variance of $Y(T)$ is given by

$$\begin{aligned} \text{Var}(Y(T)) = {} & \gamma^2 \Big(u_1^2(\sigma^2 + 2\sigma_1^2) + u_2^2(\sigma^2 + 2\sigma_2^2) + 2u_1u_2(\rho\sigma_1\sigma_2 + \sigma^2)\Big)(T-t) \\ & + \gamma^2 \Big(u_1^2 \Big(\sigma_1^2 F_1^2 + \rho_1 \sigma \sigma_2 F_1\Big) + u_2^2 \Big(\sigma_2^2 F_2^2 + \rho_2 \sigma_2 \sigma F_2\Big) \\ & + u_1 u_2 \Big(2\rho\sigma_1\sigma_2(F_{1,2} - F_1 - F_2) + \rho_1 \sigma \sigma_1 F_1 + \rho_2 \sigma_2 \sigma F_2\Big)\Big) \end{aligned} \tag{C.6}$$

We are finally able to conclude

$$\begin{aligned} E[U(W(T))|\mathcal{F}_t] = {} & W^\gamma(t) \exp\Bigg[\gamma\Big(\nu(u_1, u_2)(T-t) + u_1 A_1 Z_1(t) \int_t^T e^{A_1 s} ds \\ & + u_2 A_2 Z_2(t) \int_t^T e^{A_2 s} ds\Big) + \frac{1}{2}\gamma^2 \Big(u_1^2(\sigma^2 + 2\sigma_1^2) \\ & + u_2^2(\sigma^2 + 2\sigma_2^2) + 2u_1u_2(\rho\sigma_1\sigma_2 + \sigma^2)\Big)(T-t) \\ & + \gamma^2 \Big(u_1^2 \Big(\sigma_1^2 F_1^2 + \rho_1 \sigma \sigma_2 F_1\Big) + u_2^2 \Big(\sigma_2^2 F_2^2 + \rho_2 \sigma_2 \sigma F_2\Big) \\ & + u_1 u_2 \Big(2\rho\sigma_1\sigma_2(F_{1,2} - F_1 - F_2) + \rho_1 \sigma \sigma_1 F_1 + \rho_2 \sigma_2 \sigma F_2\Big)\Bigg]\Big) \end{aligned} \tag{C.7}$$

by combining (5.19) and the formula of the expected value of a log-normal random variable.

C.4 Calculations of $u_i^{'}$ in Proposition 5.5.1

In this section we give some of the calculation for the partial derivatives of $E[U(W(T))|\mathcal{F}_t]$. Recall equation

$$\frac{\partial W^\gamma(t)}{\partial u_i} + W^\gamma(t) \frac{\partial g(u_i, u_j)}{\partial u_i} = 0 \tag{C.8}$$

where

$$\frac{\partial W^\gamma(t)}{\partial u_i} = W^\gamma(t)\gamma\Big[\Big((\mu\frac{1}{2}\Sigma_i - r) - u_i\Sigma_i - u_j\Sigma\Big)t + \int_0^t A_i Z_i(s)ds + \sigma\widetilde{B}(t) + \gamma\sigma_i B_i(t)\Big] \quad \text{(C.9)}$$

and

$$\begin{aligned}\frac{\partial g(u_i,u_j)}{\partial u_i} =& \gamma\Big(A_i Z_i(t)\int_t^T e^{A_i s}ds\Big) + \gamma\Big(\mu + \frac{1}{2}\Sigma_i - r\Big)(T-t) \\ &- u_i\gamma\Big[\Big(\Sigma_i - \gamma(\sigma^2 + 2\sigma_i^2)\Big)(T-t) - \gamma\Big(\sigma_i^2 F_i^2 + \rho_i\sigma\sigma_i F_i\Big)\Big] \\ &- u_j\gamma\Big[\Big(\Sigma - \gamma P_j\Big)(T-t) - \gamma\Big(\rho\sigma_i\sigma_j(F_{i,j} - F_i - F_j) + \frac{1}{2}(\rho_i\sigma\sigma_i F_i \\ &+ \rho_j\sigma\sigma_j F_j)\Big)\Big] \qquad \text{(C.10)}\end{aligned}$$

Combining (C.9) and (C.10), we obtain

$$\begin{aligned}&\gamma\Big[\Big((\mu + \frac{1}{2}\Sigma_i - r) - u_i\Sigma_i - u_j\Sigma\Big)t + \gamma\int_0^t A_i Z_i(s)ds + \gamma\sigma\widetilde{B}(t) + \gamma\sigma_i B_i(t)\Big] \\ &+ \gamma A_i Z_i(t)\int_t^T e^{A_i s}ds + \gamma\Big(\mu + \frac{1}{2}\Sigma_i - r\Big)(T-t) \\ &- u_i\gamma\Big[\Big(\Sigma_i - \gamma(\sigma^2 + 2\sigma_i^2)\Big)(T-t) - \gamma\Big(\sigma_i^2 F_i^2 + \rho_i\sigma\sigma_i F_i\Big)\Big] \\ &- u_j\gamma\Big[\Big(\Sigma - \gamma P_j\Big)(T-t) - \gamma\Big(\rho\sigma_i\sigma_j(F_{i,j} - F_i - F_j) + \frac{1}{2}(\rho_i\sigma\sigma_i F_i \\ &+ \rho_j\sigma\sigma_j F_j)\Big)\Big] = 0\end{aligned}$$

Solving for u_i we obtain the following equation

$$\begin{aligned}&u_i\Big[\Sigma_i T - \gamma\Big((\sigma^2 + 2\sigma_i^2)(T-t) + \sigma_i^2 F_i^2 + \rho_i\sigma\sigma_i F_i\Big)\Big] \\ &= A_i Z_i(t)\int_t^T e^{A_i s}ds + (\mu + \frac{1}{2}\Sigma_i - r)T + \gamma\int_0^t A_i Z_i(s)ds + \gamma\sigma\widetilde{B}(t) \\ &+ \gamma\sigma_i B_i(t) - u_j\Big[\Sigma T - \gamma\Big(+ P_j(T-t) \\ &\rho\sigma_i\sigma_j(F_{i,j} - F_i - F_j) + \frac{1}{2}(\rho_i\sigma\sigma_i F_i + \rho_j\sigma\sigma_j F_j)\Big)\Big]\end{aligned}$$

Notation C.4.1.
Denote

$$U_i = \Sigma_i T - \gamma\Big((\sigma^2 + 2\sigma_i^2)(T-t) + \sigma_i^2 F_i^2 + \rho_i\sigma\sigma_i F_i\Big)$$

Then

$$\begin{aligned}u_1 = \frac{1}{U_1}\Big[&A_1 Z_1(t)\int_t^T e^{A_1 s}ds + (\mu + \frac{1}{2}\Sigma_1 - r)T + \gamma\int_0^t A_1 Z_1(s)ds + \gamma\sigma\widetilde{B}(t) \\ &+ \gamma\sigma_1 B_1(t) - u_2\Big[\Sigma T - \gamma\Big(P_2(T-t) + \rho\sigma_1\sigma_2(F_{1,2} - F_1 - F_2)\end{aligned}$$

$$+\frac{1}{2}(\rho_1\sigma\sigma_1 F_1+\rho_j\sigma\sigma_2 F_2)\Big)\Big]\Big]$$

$$u_2=\frac{1}{U_2}\Big[A_2Z_2(t)\int_t^T e^{A_2 s}ds+(\mu+\frac{1}{2}\Sigma_2-r)T+\gamma\int_0^t A_2Z_2(s)ds+\gamma\sigma\widetilde{B}(t)$$
$$+\gamma\sigma_2B_2(t)-u_1\Big[\Sigma T-\gamma\Big(P_1(T-t)$$
$$\rho\sigma_1\sigma_2(F_{1,2}-F_1-F_2)+\frac{1}{2}(\rho_1\sigma\sigma_1F_1+\rho_2\sigma\sigma_2F_2)\Big)\Big]\Big]$$

Introduce the shorter version

$$u_1=\frac{1}{U_1}\Big[A_1Z_1(t)\int_t^T e^{A_1 s}ds+K_1+\gamma\int_0^t A_1Z_1(s)ds+\gamma\sigma\widetilde{B}(t)$$
$$+\gamma\sigma_1B_1(t)-u_2L_1\Big]$$

$$u_2=\frac{1}{U_2}\Big[A_2Z_2(t)\int_t^T e^{A_2 s}ds+K_2+\gamma\int_0^t A_2Z_2(s)ds+\gamma\sigma\widetilde{B}(t)$$
$$+\gamma\sigma_2B_2(t)-u_1L_2\Big]$$

Substitute u_1 into u_2 and we obtain that the optimal u_2 is

$$u_2^{'}=\frac{U_1}{L_1L_2-U_1U_2}\Big[-\frac{L_2}{U_1}\Big(A_1Z_1(t)\int_t^T e^{A_1 s}ds+\gamma\int_0^t A_1Z_1(s)ds\Big)$$
$$+\Big(A_2Z_2(t)\int_t^T e^{A_2 s}ds+\gamma\int_0^t A_2Z_2(s)ds\Big)+\gamma\sigma\widetilde{B}(t)\Big(1-\frac{L_2}{U_1}\Big)$$
$$-\gamma\sigma_1\frac{L_2}{U_1}B_1(t)+\gamma\sigma_2B_2(t)-\frac{L_2}{U_1}K_1+K_2\Big]$$

$$u_1^{'}=\frac{1}{U_1(L_1L_2-U_1U_2)}\Big[(2L_1L_2-U_1U_2)\Big(A_1Z_1(t)\int_t^T e^{A_1 s}ds+\gamma\int_0^t A_1Z_1(s)ds\Big)$$
$$-L_2U_1\Big(A_2Z_2(t)\int_t^T e^{A_2 s}ds+\gamma\int_0^t A_2Z_2(s)ds\Big)+\gamma\sigma\widetilde{B}(t)\Big(L_1L_2-U_1U_2$$
$$-L_2U_1-L_2^2\Big)+\gamma\sigma_1B_1(t)\Big(L_1L_2-U_1U_2+L_2^2)-U_1L_2\gamma\sigma_2B_2(t)+L_2^2K_1$$
$$-L_2K_2\Big)$$

The optimal controls

$$u_1^{'}=\Big(A_1Z_1(t)\int_t^T e^{A_1 s}ds+\gamma\int_0^t A_1Z_1(s)ds\Big)\vartheta_1+\Big(A_2Z_2(t)\int_t^T e^{A_2 s}ds$$
$$+\gamma\int_0^t A_2Z_2(s)ds\Big)\vartheta_2+\gamma\sigma\widetilde{B}(t)\vartheta_3+\gamma\sigma_1B_1(t)\vartheta_4+\gamma\sigma_2B_2(t)\vartheta_5+\vartheta_6$$

$$u_2' = \Big(A_1 Z_1(t) \int_t^T e^{A_1 s} ds + \gamma \int_0^t A_1 Z_1(s) ds\Big)\varepsilon_1 + \Big(A_2 Z_2(t) \int_t^T e^{A_2 s} ds + \gamma \int_0^t A_2 Z_2(s) ds\Big) + \gamma\sigma \widetilde{B}(t)\varepsilon_2 - \gamma\sigma_1 B_1(t)\varepsilon_3 + \gamma\sigma_2 B_2(t) - \varepsilon_4\Big]$$

for ϑ_i, ε_j, $i = 1, .., 6$, $j = 1, ..4$ in Appendix A.

Bibliography

[AO05] Arikoglu, A. and Ozkol, I. "Solution of boundary value problems for integro-differential equations by using differential transform method". In: *Applied Mathematics and Computation* 168.2 (2005), pp. 1145–1158.

[AS10] Andersen, K. M. and Sandqvist, A. "A generalized Allwright Formula and the Vector Riccati Equation". In: *arXiv preprint arXiv:1010.5702* (2010).

[Bar16] Barbu, V. *Differential Equations.* eng. Springer Undergraduate Mathematics Series. Cham, 2016.

[Bel18] Bellman, R. "An Introduction to the Theory of Dynamic Programming". In: Rand report R-245 (2018).

[BK15] Benth, F. E. and Koekebakker, S. "Pricing of Forwards and Other Derivatives in Cointegrated Commodity Markets". In: *Energy Economics* 52 (2015), pp. 104–117.

[Bun96] Bunse-Gerstner, A. "Computational Solution of the Algebraic Riccati Equation." eng. In: *Journal of The Society of Instrument and Control Engineers* 35.8 (1996), pp. 632–639.

[CCF17] Confortola, F., Cosso, A., and Fuhrman, M. "Backward SDEs and Infinite Horizon Stochastic Optimal Control". In: (2017).

[Che] Chen, N. *Lecture 5: Connections with Partial Differential Equations and The Feynman-Kac Theorem.* URL: http://www1.se.cuhk.edu.hk/~seem5670/Lecture/Lecture5.pdf. (accessed: 29.04.2019).

[Chu74] Chung, K. L. *A Course in Probability Theory.* eng. 2nd ed. Vol. 21. Probability and Mathematical Statistics. New York: Academic Press, 1974.

[EG87] Engle, R. F. and Granger, C. W. J. "Co-Integration and error correction: representation, sstimation, and testing". English. In: *Econometrica (1986-1998)* 55.2 (1987), pp. 251–276.

[FA16] File, G. and Aga, T. "Numerical solution of quadratic Riccati differential equations". eng. In: *Egyptian Journal of Basic and Applied Sciences* 3.4 (2016), pp. 392–397.

[Fil09] Filipovivic, D. *Term-Structure Models : A Graduate Course.* eng. Springer finance. Berlin: Springer, 2009.

[FS06] Fleming, W. H. and Soner, H. *Controlled Markov Processes and Viscosity Solutions.* eng. Second Edition. Vol. 25. Stochastic Modelling and Applied Probability. New York, NY: Springer New York, 2006.

[Gal] Gale, T. *Inada Conditions, International Encyclopedia of the Social Sciences*. URL: https://www.encyclopedia.com/social-sciences/applied-and-social-sciences-magazines/inada-conditions.

[JS19] Jeon, J. and Shin, Y. H. "Finite Horizon Portfolio Selection With a Negative Wealth Constraint". In: *Journal of Computational and Applied Mathematics* 356 (2019), pp. 329–338.

[KS91] Karatzas, I. and Shreve, S. E. *Brownian Motion and Stochastic Calculus*. eng. 2nd ed. Vol. 113. Springer study edition. New York: Springer, 1991.

[LL06] Lee, C. F. and Lee, A. C. *Encyclopedia of finance*. Springer, 2006, pp. 78–78.

[Mar52] Markowitz, H. "Portfolio Selection". In: *Journal of Finance* 7.1 (1952), pp. 77–91.

[McD13] McDonald, J. N. *A Course in Real Analysis*. eng. Amsterdam: Academic Press, 2013.

[Mer69] Merton, R. C. "Lifetime Portfolio Selection under Uncertainty: The Continuous-Time Case". eng. In: *The Review of Economics and Statistics* 51.3 (1969), pp. 247–257.

[Mer71] Merton, R. C. "Optimum consumption and portfolio rules in a continuous-time model". eng. In: *Journal of Economic Theory* 3.4 (1971), pp. 373–413.

[Mer75] Merton, R. C. "Optimum consumption and portfolio rules in a continuous-time model". In: *Stochastic Optimization Models in Finance*. Elsevier, 1975, pp. 621–661.

[MH11] Mohammadi, F. and Hosseini, M. "A comparative study of numerical methods for solving quadratic Riccati differential equations". eng. In: *Journal of the Franklin Institute* 348.2 (2011), pp. 156–164.

[Mir11] Mirzaee, F. "Differential transform method for solving linear and nonlinear systems of ordinary differential equations". In: *Applied Mathematical Sciences* 5.70 (2011), pp. 3465–3472.

[Mor10] Morimoto, H. *Stochastic control and Mathematical Modeling : Applications in Economics*. eng. Vol. 131. Encyclopedia of mathematics and its applications. Cambridge: Cambridge University Press, 2010.

[Nis15] Nisio, M. *Stochastic Control Theory : Dynamic Programming Principle*. eng. 2nd ed. Vol. 72. Probability Theory and Stochastic Modelling. Tokyo, 2015.

[Øks05] Øksendal, B. "The Value of Information in Stochastic Control and Finance". In: *Australian Economic Papers* 44.4 (2005), pp. 352–364.

[Øks13] Øksendal, B. "Stochastic differential equations: An introduction with applications". In: *Stochastic Differential Equations*. Springer Science & Business Media, 2013.

[Ros14] Ross, S. M. *Introduction to Probability Models*. Academic press, 2014.

[SP17] Suresh, P. and Piriadarshani, D. "Numerical Analysis of Riccati Equations using Differential Transform Method, He Laplace Method and Adomian Decomposition Method". In: *Global Journal of Mathematical Sciences: Theory and Practical* 9.1 (2017), pp. 31–50.

[Wik19] Wikipedia, c. *Crush Spread*. Mar. 10, 2019. URL: https://en.wikipedia.org/w/index.php?title=Crush_spread&oldid=816281939.

[YZ99] Yong, J. and Zhou, X. Y. *Stochastic Controls: Hamiltonian Systems and HJB Equations*. Vol. 43. New York: Springer Science & Business Media, 1999.