Simon Pojer, BSc

Portfolio Optimization in Heston's Stochastic Volatility Model

MASTER'S THESIS

to achieve the university degree of

Diplom-Ingenieur

Master's degree programme: Mathematics

submitted to

Graz University of Technology

Supervisor

Assoc.Prof. Dipl.-Ing. Dr.techn. Stefan Thonhauser

Institute of Statistics

Graz, May 2020

Affidavit

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources, and that I have explicitely indicated all material which has been quoted either literally or by content from the sources used. The text document uploaded to TUGRAZonline is identical to the present master‘s thesis.

Date	Signature

Abstract

Extending Merton's portfolio problem to stochastic volatility models, it is not always obvious which methods of stochastic control theory can be applied, and if a closed-form solution can be found. Further, it is still unknown how the inclusion of early announced discrete paid dividends influences optimality in the underlying market model. In this thesis, we show that it is possible to apply the martingale approach to Heston's stochastic volatility model with power utility functions, and that both the optimal portfolio, and the optimal dual process, can be derived explicitly. Incorporating early announced dividend payments, the optimal expected utility from terminal wealth does not change, but the optimal portfolio process does. To achieve optimality, an investor must dedicate a larger fraction of his wealth to risky assets, hence take more risk than in the dividend-free model.

Bei der Verallgemeinerung von Mertons Portfolio-Problem mittels stochastischer Volatilitätsmodelle ist nicht immer ersichtlich, welche Methoden der stochastischen Kontrolltheorie angewandt werden können und ob eine Lösung in geschlossener Form ermittelt werden kann. Weiters ist ungeklärt, wie das Miteinbeziehen von vorzeitig bekanntgegebenen Dividenden die Optimalität im zugrundeliegenden Modell beeinflusst. Wir zeigen, dass die Martingalmethode auf Hestons stochastisches Volatilitätsmodell mit Potenznutzenfunktion angewandt werden kann und sowohl der optimale Portfolioprozess als auch der optimale duale Prozess explizit bestimmt werden können. Die Erweiterung von Hestons Modell mit vorzeitig bekanntgegebenen Dividendenzahlungen bringt zwar keine Veränderung des Erwartungsnutzen durch Endvermögen mit sich, sehr wohl aber eine Veränderung des optimalen Portfolioprozesses. In diesem Fall muss ein Investor einen größeren Teil seines Geldes risikobehaftet veranlagen, um optimal zu agieren.

Contents

1 Introduction

Time-continuous portfolio optimization was led into a new direction as Merton (1969), introduced the problem of optimizing expected utility from terminal wealth for both finite and infinite time horizons. This famous problem was formulated and solved in the framework developed by Black and Scholes (1973), the well-known Black-Scholes model.

Crucial assumptions made by Black and Scholes include the log-normal distribution of the stocks as well as a constant volatility parameter, but empirical data do not support these premises. Those issues can be solved by the usage of stochastic volatility models which use a stochastic process to describe the volatility of the stocks. This approach, however, includes further sources of randomness. Hence, those models are incomplete.

Heston's stochastic volatility model, short Heston's SV model, is one of the best-known stochastic volatility models. Developed by Heston (1993), it consists of a single risky asset and the volatility is modelled by a Cox-Ingersoll-Ross process, a non-negative and mean-reverting stochastic process. Thus, the value of the assets is no longer log-normally distributed nor is the volatility constant. Consequently, two major drawbacks of the Black-Scholes model are removed.

The problem of optimizing utility from terminal wealth in Heston's SV model with power utility has already been solved by Kraft (2003). To obtain a solution, he used stochastic dynamic programming, developed by Bellman (1958), for optimal decision making under uncertainty. The main idea of this approach is, to solve the Hamilton-Jacobi-Bellman equation, a non-linear partial differential equation. This equation is necessarily solved by the optimal value function, but an additional verification process has to be made, to guarantee the optimality of the derived candidate.

The second major path to solve a dynamic portfolio optimization problem is the martingale approach, introduced by Pliska (1986), Karatzas, Lehoczky, and Shreve (1987), and Cox and Huang (1989). This method exploits the martingale representation theorem, hence the optimal portfolio process is only implicitly given and cannot always be derived explicitly. In the case of incomplete markets, optimal processes can be identified, via the solution of a dual optimization problem.

In this thesis, we will discuss the applicability of the martingale approach to Heston's SV model with power utility functions. Furthermore, we will study the solvability of the corresponding dual optimization problem and wether a dual optimum can be derived. To achieve this, we recapitulate some statements from probability theory and stochastic analysis, where the focus lies mainly on affine processes and diffusion market models. Furthermore, the martingale method to time-continuous portfolio optimization is introduced, for both, complete and incomplete market models, including some statements concerning the existence of dual optima. Finally, we apply this approach to Heston's SV model, in its parametrization used by Kraft (2003), with and without early announced dividends.

2 Probabilistic Basics and Affine Processes

In this chapter, we will present some definitions and theorems which are the basis of our further work. Firstly, we will recapture some probabilistic results, like the martingale representation theorem and Girsanov's theorem which are of utmost importance for the martingale approach to portfolio optimization. For this, we will rely on Karatzas and Shreve (1988) since the notation there suits perfectly with the notation of Karatzas and Shreve (2003). Alternatively, one could see most of these well-known findings in Cohen and Elliott (2015). Later, we will introduce the class of affine processes, which are not needed for the theory of portfolio optimization but are helpful in the application onto Heston's market model. Here we will follow Filipovic (2009), who used affine processes in the framework of term structure models.

2.1 Probabilistic Basics

Remark 2.1. For the rest of this thesis $\underline{1}_d$ denotes the d-dimensional vector, consisting only of 1s. Furthermore, for any vector b or matrix A we will write $b^\top$ for the corresponding transposed vector and $A^\top$ for the transposed matrix, respectively. The norm $\|b\|$ denotes the Euclidean norm, if not mentioned differently.

Remark 2.2. For convenience, we will introduce 'Permanent Assumptions' which are assumed to be in force until the end of the section, in which they were stated. 'Assumptions' are only in force if it is explicitly declared they are.

Permanent Assumption 2.3. We assume that there is some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ on which all processes and random variables of this chapter are defined. If not mentioned differently, we further assume that the processes live on a finite time interval $[0, T]$.

Definition 2.4. We say a function is *càdlàg* (*làdcàg*) if it is right-continuous (left-continuous) and the left (right) limit exists in every point.

Remark 2.5. We call a process X càdlàg, làdcàg, continuous, of bounded variation or bounded, if its sample paths satisfy the respective property almost surely.

Definition 2.6. Let $\mathcal{N}$ be the set, consisting of all subsets of zero sets of the measure $\mathbb{P}$, and $\left\{\tilde{\mathcal{F}}_t\right\}$ be a filtration. The filtration defined by

$$\mathcal{F}_t := \sigma\left(\tilde{\mathcal{F}}_t \cup \mathcal{N}\right),$$

is called augmention of the filtration $\{\mathcal{F}_t\}$. Here $\sigma(A)$ denotes the smallest σ-algebra containing A, also known as σ-algebra generated by A.
We call $\left\{\tilde{\mathcal{F}}_t\right\}$ complete, if for all t the equation

$$\tilde{\mathcal{F}}_t = \mathcal{F}_t$$

holds true.

Definition 2.7. We call a filtration right-continuous, if

$$\mathcal{F}_t = \bigcap_{s>t} \mathcal{F}_s,$$

left-continuous if

$$\mathcal{F}_t = \sigma\left(\bigcup_{s<t} \mathcal{F}_s\right)$$

and continuous if it is both left and right continuous.

Definition 2.8. We say a filtration $\{\mathcal{F}_t\}$ satisfies the usual conditions if it is right-continuous and complete.

Lemma 2.9. *The augmented natural filtration of a d-dimensional Brownian motion is continuous and satisfies the usual conditions.*

Definition 2.10. A n-dimensional stochastic process X is called progressively measurable with respect to a filtration $\{\mathcal{F}_t\}$ if the function

$$\iota : \left([0,t] \times \Omega, \mathcal{B}([0,t]) \otimes \mathcal{F}_t\right) \to (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n)),$$

mapping (s,ω) to $X(s,\omega)$, is measurable for all $t \geq 0$.

Proposition 2.11. *Every left-continuous (or right-continuous) process X, adapted to $\{\mathcal{F}_t\}$ is progressively measurable with respect to $\{\mathcal{F}_t\}$.*

Definition 2.12. A one-dimensional, $\{\mathcal{F}_t\}$-adapted process ν on $[0,T]$, is called locally bounded if there exists a monotone sequence of stopping times $(\tau_n)_{n\geq 1}$ and constants $(K_n)_{n\geq 1}$ such that $\tau_n \overset{a.s}{\to} T$ for $n \to \infty$ and for all n, the stopped process ν^{τ_n} defined by $\nu^{\tau_n}(\omega,t) := \nu(\omega, \min(t,\tau_n))$, $\mathbb{P}$-almost surely satisfies

$$|\nu^{\tau_n}(t)| \leq K_n \quad \forall\, t \in [0,T].$$

Lemma 2.13. *Let ν be a one-dimensional, $\{\mathcal{F}_t\}$-adapted, continuous process. Then ν is locally bounded.*

Proof. Since ν is continuous and $\{\mathcal{F}_t\}$-adapted, the random times

$$\tau_n := \inf\left\{T \geq t \geq 0 \,|\, |\nu(t)| \geq n\right\}$$

are $\{\mathcal{F}_t\}$-stopping times. Then $\tau_n \overset{a.s}{\to} T$ and for all τ_n and $t \in [0, T]$ we have

$$|\nu^{\tau_n}(t)| \leq n.$$

□

Definition 2.14. An one-dimensional, $\{\mathcal{F}_t\}$-adapted process X, satisfying

$$\mathbb{E}\left[|X(t)|\right] < \infty$$

for all t, is said to be

- a submartingale, if $\mathbb{E}[X(t)\,|\mathcal{F}_s] \geq X(s)$ holds for every $0 \leq s \leq t$ $\mathbb{P}$-almost surely.
- a supermartingale, if $\mathbb{E}[X(t)\,|\mathcal{F}_s] \leq X(s)$ holds for every $0 \leq s \leq t$ $\mathbb{P}$-almost surely.
- a martingale, if $\mathbb{E}[X(s)|\mathcal{F}_s] = X(s)$ holds for every $0 \leq s \leq t$ $\mathbb{P}$-almost surely.

Definition 2.15. A one-dimensional process X is a local martingale if there exists a monotone increasing sequence of $\{\mathcal{F}_t\}$-stopping times $\{\tau_n\}_{n\in\mathbb{N}}$, with

$$\mathbb{P}[\lim_{n\to\infty} \tau_n = T] = 1,$$

such that for every n the stopped process X^{τ_n} is a martingale.

Lemma 2.16. *Let X be a non-negative local martingale. Then X is a supermartingale.*

Proof. Let X be a local martingale and $\{\tau_n\}_{n\in\mathbb{N}}$ the corresponding sequence of stopping times. For $0 \leq s \leq t$ we get

$$\mathbb{E}\left[X(t) - X(s) \mid \mathcal{F}_s\right] = \mathbb{E}\left[\lim_{n\to\infty} X^{\tau_n}(t) - X^{\tau_n}(s) \mid \mathcal{F}_s\right].$$

Fatou's lemma and the martingale property of X^{τ_n} lead to

$$\mathbb{E}\left[X(t) - X(s) \mid \mathcal{F}_s\right] \leq \lim_{n\to\infty} \mathbb{E}\left[X^{\tau_n}(t) - X^{\tau_n}(s) \mid \mathcal{F}_s\right] = 0.$$

By using that $X(s)$ is $\mathcal{F}_s$ measurable we have

$$\mathbb{E}\left[X(t)\ |\mathcal{F}_s\right] \leq X(s).$$

□

Definition 2.17. A one-dimensional process X is called continuous semimartingale with respect to a filtration $\{\mathcal{F}_t\}$, if there exists a continuous $\{\mathcal{F}_t\}$-local martingale M, with $M(0) = 0$, a continuous $\{\mathcal{F}_t\}$-adapted process of finite variation A, with $A(0) = 0$ and a $\mathcal{F}_0$ measurable random variable X_0 satisfying almost surely

$$X(t) = X_0 + M(t) + A(t), \quad \forall\, t \in [0, T].$$

If A is assumed to be previsible, the decomposition is even unique.

Theorem 2.18. *Let M and N be continuous, local martingales. Then there exists a unique adapted, continuous process $\langle M, N\rangle$ of bounded variation, such that $\langle M, N\rangle(0) = 0$ and the process*

$$MN - \langle M, N\rangle$$

is again a continuous local martingale.

Remark 2.19. In the case of non-continuous local martingales, there exists a similar process which is no longer continuous, and uniqueness is only given if predictability is assumed.

Definition 2.20. We call the process $\langle M, N\rangle$ of Theorem 2.18 cross-variation or quadratic covariation process of M and N. We call $\langle M, M\rangle$ quadratic variation process of M and write for short $\langle M\rangle$.

Theorem 2.21. *(Itô's Lemma)*
Let $X = \left(X^{(1)}, \ldots, X^{(d)}\right)^\top$ be a process, consisting of d continuous semimartingales and $f : [0, T] \times \mathbb{R}^d \to \mathbb{R}$ a function which is continuously differentiable in the first argument and twice continuously differentiable in the second argument. Then

$$\begin{aligned} f(t, X(t)) = f(0, X(0)) &+ \int_0^t \frac{\partial}{\partial t} f(s, X(s))\, \mathrm{d}s + \sum_{i=1}^d \int_0^t \frac{\partial}{\partial x_i} f\left(s, X(s)\right)\, \mathrm{d}X^{(i)}(s) \\ &+ \frac{1}{2} \sum_{i=1}^d \sum_{j=1}^d \int_0^t \frac{\partial^2}{\partial x_i \partial x_j} f\left(s, X(s)\right)\, \mathrm{d}\langle X^{(i)}, X^{(j)}\rangle(s) \end{aligned}$$

holds true almost surely. Here $\langle X^{(i)}, X^{(j)}\rangle$ denotes the quadratic covariation process of the semimartingales $X^{(i)}$ and $X^{(j)}$.

Theorem 2.22. *(Integration by Parts)*
Let X and Y be continuous semimartingales and let $\langle X, Y\rangle$ denote the quadratic covariation process of these processes. Then the equality

$$X(t)Y(t) = X(0)Y(0) + \int_0^t Y(s)\, \mathrm{d}X(s) + \int_0^t X(s)\, \mathrm{d}Y(s) + \langle X, Y\rangle(t)$$

holds true for all t.

Corollary 2.23. *Let $F : [0, T] \to \mathbb{R}$ be a continuously differentiable function, with derivative f and X be a continuous semimartingale. Then for all $t \in [0, T]$, we have the equality*

$$F(t)X(t) = F(0)X(0) + \int_0^t F(s)\,\mathrm{d}X(s) + \int_0^t X(s)f(s)\,\mathrm{d}s.$$

Definition 2.24. We will call a process ψ, satisfying

$$\int_0^T \|\psi(t)\|^2\,\mathrm{d}t < \infty \quad a.s.,$$

square integrable.

Theorem 2.25. *Let M be a one dimensional, càdlàg, $\{\mathcal{F}_t\}$-local martingale, where $\{\mathcal{F}_t\}$ denotes the augmented natural filtration of a Brownian motion B. Then there exists a square integrable, progressively measurable processes ψ, such that*

$$M(t) = M(0) + \int_0^t \psi(s)^\top\,\mathrm{d}B(s).$$

Remark 2.26. This representation theorem is the basis of the martingale approach to portfolio optimization. Even though this representation looks quite nice, we must consider that ψ is generally not known and in many cases, there is no possibility to determine it.

Corollary 2.27. *If $\{\mathcal{F}_t\}$ is the augmented natural filtration of a Brownian motion, then every $\{\mathcal{F}_t\}$-local martingale has a continuous version.*

Definition 2.28. Let B be a standard Brownian motion, $\{\mathcal{F}_t\}$ the augmention of its natural filtration and H a $\{\mathcal{F}_t\}$-progressively measurable, square integrable process. Then we call the process

$$Z(t) := \exp\left(\int_0^t H(s)^\top\,\mathrm{d}B(s) - \frac{1}{2}\int_0^t \|H(s)\|^2\,\mathrm{d}s\right)$$

stochastic exponential process. We sometimes call it stochastic exponential of X, where

$$X(t) := \int_0^t H(s)^\top\,\mathrm{d}B(s).$$

In this case, Z can be rewritten as

$$Z(t) = \exp\left(X(t) - \frac{1}{2}\langle X\rangle(t)\right),$$

where $\langle X\rangle$ denotes the quadratic variation process of X.

Lemma 2.29. *The stochastic exponential process of Definition 2.28 is a positive local martingale, hence a supermartingale. Moreover, it solves the SDE*

$$\mathrm{d}Z(t) = Z(t)H(t)^\top \mathrm{d}B(t), \quad Z(0) = 1,$$

and has consequently the alternative representation

$$Z(t) = 1 + \int_0^t Z(s)H(s)^\top \, \mathrm{d}B(s).$$

Proof. If we define

$$\chi(t) := \int_0^t H(s)^\top \, \mathrm{d}B(s) - \frac{1}{2}\int_0^t \|H(s)\|^2 \, \mathrm{d}s$$

and apply Itô's lemma to the process $Z(t) = \exp\big(\chi(t)\big)$ we get

$$\begin{aligned}
\mathrm{d}Z(t) &= Z(t)\mathrm{d}\chi(t) + \frac{1}{2}Z(t)\mathrm{d}\langle\chi\rangle(t) \\
&= Z(t)H(t)^\top \mathrm{d}B(t) - \frac{1}{2}Z(t)\|H(t)\|^2\mathrm{d}t + \frac{1}{2}Z(t)\|H(t)\|^2\mathrm{d}t \\
&= Z(t)H(t)^\top \mathrm{d}B(t).
\end{aligned}$$

Therefore, Z has a representation as

$$Z(t) = 1 + \int_0^t Z(s)H(s)^\top \, \mathrm{d}B(s).$$

Since all stochastic integrals with respect to the Brownian motions are local martingales, $Z(t)$ is a local martingale and non-negative by definition. By Lemma 2.16, $Z(t)$ is a supermartingale. □

Remark 2.30. The definition of stochastic exponential processes can easily be generalized to processes of the form

$$X(t) = \int_0^t b(s) \, \mathrm{d}s + \int_0^t H(s)^\top \, \mathrm{d}B(s),$$

by the definition

$$Z(t) = \exp\left(X(t) - \frac{1}{2}\langle X\rangle(t)\right).$$

The stochastic exponential of the process X is given by

$$Z(t) = \exp\left(\int_0^t b(s) \, \mathrm{d}s + \int_0^t H(s)^\top \, \mathrm{d}B(s) - \frac{1}{2}\int_0^t \|H(s)\|^2 \, \mathrm{d}s\right).$$

Here b denotes an one-dimensional, integrable $\{\mathcal{F}_t\}$-progressively measurable, process. The stochastic exponential solves the SDE

$$\mathrm{d}Z(t) = Z(t) \, \mathrm{d}X(t).$$

Theorem 2.31. *(Girsanov's Theorem)*
Let B be a standard Brownian motion with respect to the measure $\mathbb{P}$ and $\{\mathcal{F}_t\}$ its augmented natural filtration. Furthermore, let H be a $\{\mathcal{F}_t\}$-progressively measurable, square integrable process and Z be the stochastic exponential process

$$Z(t) = \exp\left(\int_0^t H(s)^\top \, \mathrm{d}B(s) - \frac{1}{2}\int_0^t \|H(s)\|^2 \, \mathrm{d}s\right).$$

Assume Z is a true martingale and define a new measure $\tilde{P}$ by

$$\tilde{P}\,(A) = \mathbb{E}\left[Z(T)\mathbb{1}_A\right],$$

where the expectation is taken under the measure $\mathbb{P}$ and $\mathbb{1}_A$ denotes the indicator function of the event A. Then the process

$$\tilde{B}(t) := B(t) - \int_0^t H(s)\,\mathrm{d}s$$

is a standard Brownian motion with respect to the measure $\tilde{P}$ and the Filtration $\{\mathcal{F}_t\}$.

Remark 2.32. The importance of Girsanov's theorem is due to the usage of equivalent measures for pricing purposes. Furthermore, the martingale method in incomplete markets relies fundamentally on this representation of densities of equivalent measures, since in diffusion market models all equivalent martingale measures are given by Girsanov's theorem. We will discuss this feature in Theorem 4.44.

Theorem 2.33. *Let M be a continuous local martingale with respect to the filtration $\{\mathcal{F}_t\}$ and define*

$$Z(t) := \exp\left(M(t) - \frac{1}{2}\langle M\rangle(t)\right).$$

If

$$\mathbb{E}\left[\exp\left(\frac{1}{2}\langle M\rangle(t)\right)\right] < \infty$$

holds true for all $t \in [0,T]$, then $\mathbb{E}\left[Z(t)\right] = 1$ for all $t \in [0,T]$.

Remark 2.34. The property $\mathbb{E}\left[\exp\left(\frac{1}{2}\langle M\rangle(t)\right)\right] < \infty$ is often called Novikov's condition, and is a nice way to ensure that the local martingale Z is a true martingale. Especially in combination with Girsanov's theorem, this is a powerful tool, to guarantee the existence of exponential measure changes, but it is only sufficient, hence there are exponential martingales which do not satisfy this condition.

Lemma 2.35. [1] *Let H be a left continuous, locally bounded process. For every sequence of partitions of the interval $[0,T]$ satisfying*

$$\lim_{n\to\infty}\sup_{0\le k<n}|\,t^n_{k+1}-t^n_k|=0,$$

the equality

$$\int_0^T H(s)^\top\,\mathrm{d}X(s)=\lim_{n\to\infty}\sum_{0\le k<n}H(t^n_k)^\top(X(t^n_{k+1})-X(t^n_k))$$

holds true for every continuous semimartingale X. Here, the limit denotes convergence in probability.

Remark 2.36. This lemma allows us, to interpret these stochastic integrals as a limit of Riemann sums, which can be used to determine some properties of the integral.

Theorem 2.37. *Let B be a standard Brownian motion with respect to the filtration $\{\mathcal{F}_t\}$ and assume this filtration satisfies the usual conditions. Furthermore, let H be a continuous, square integrable one-dimensional process which is $\{\mathcal{F}_t\}$-adapted, but independent of B. Then*

$$\mathbb{E}\left[\exp\left(x\int_0^T H(s)\,\mathrm{d}B(s)\right)\Bigg|\,\mathcal{G}\right]=\exp\left(\frac{x^2}{2}\int_0^T H^2(s)\,\mathrm{d}s\right),$$

where $\mathcal{G}$ is the σ-algebra generated by $\big\{H(s)\,|\,0\le s\le T\big\}$.

Proof. Due to its continuity, the process H is $\{\mathcal{F}_t\}$-progressively measurable and locally bounded, hence there exists a sequence of partitions of $[0,T]$, such that for $n\to\infty$

$$\sum_{k=0}^{n-1}H(t^n_k)\left(B(t^n_{k+1})-B(t^n_k)\right)\xrightarrow{p}\int_0^T H(s)\,\mathrm{d}B(s),$$

where $\xrightarrow{p}$ denotes convergence in probability. Therefore, we get for the conditional characteristic function

$$\mathbb{E}\left[\exp\left(ix\int_0^T H(s)\,\mathrm{d}B(s)\right)\Bigg|\,\mathcal{G}\right]=\lim_{n\to\infty}\mathbb{E}\left[\exp\left(ix\sum_{k=0}^{n-1}H(t^n_k)\left(B(t^n_{k+1})-B(t^n_k)\right)\right)\Bigg|\,\mathcal{G}\right].$$

If we fix n and take a look at the sum, we see that, conditioned on $\mathcal{G}$, we sum over independent normally distributed random variables with zero mean and variance

[1] Revuz and Yor 1999, p. 142.

$H^2(t_k)(t^n_{k+1} - t^n_k)$, hence the sum is again normally distributed conditioned on $\mathcal{G}$ with zero mean and variance $\sum_{k=0}^{n-1} H^2(t^n_k)(t^n_{k+1} - t^n_k)$. Since the characteristic function of a normally distributed random variable is well known, we get

$$\begin{aligned}\mathbb{E}\left[\exp\left(ix\int_0^T H(s)\,\mathrm{d}B(s)\right)\Bigg|\,\mathcal{G}\right] &= \lim_{n\to\infty}\exp\left(-\frac{x^2}{2}\sum_{k=0}^{n-1} H^2(t^n_k)(t^n_{k+1} - t^n_k)\right)\\ &= \exp\left(-\frac{x^2}{2}\int_0^T H^2(s)\,\mathrm{d}s\right),\end{aligned}$$

which gives us that $\int_0^T H(s)\,\mathrm{d}B(s)$ conditioned on $\mathcal{G}$ is normally distributed with zero mean and variance $\int_0^T H^2(s)\,\mathrm{d}s$. Hence, the conditional moment generating function is

$$\mathbb{E}\left[\exp\left(x\int_0^T H(s)\,\mathrm{d}B(s)\right)\Bigg|\,\mathcal{G}\right] = \exp\left(\frac{x^2}{2}\int_0^T H^2(s)\,\mathrm{d}s\right).$$

□

Theorem 2.38. *Let H be a $\{\mathcal{F}_t\}$-adapted, continuous and square integrable process and B be a Brownian motion with respect to the same filtration. Assume that H and B are independent, then the local martingales*

$$\begin{aligned}Z(t) &= \exp\left(\int_0^t H(s)\,\mathrm{d}B(s) - \frac{1}{2}\int_0^t H(s)^2\,\mathrm{d}s\right),\\ \tilde{Z}(t) &= \exp\left(-\int_0^t H(s)\,\mathrm{d}B(s) - \frac{1}{2}\int_0^t H(s)^2\,\mathrm{d}s\right),\end{aligned}$$

are true martingales.

Proof. We will prove this for Z only because the property for $\tilde{Z}$ can be shown analogously. Since Z is a positive local martingale, it is a supermartingale too. Therefore, it is a martingale if and only if $\mathbb{E}\left[Z(T)\right] = 1$. To show this we will use Theorem 2.37 and get

$$\mathbb{E}\left[Z(T)\,|\,\mathcal{G}\right] = \mathbb{E}\left[\exp\left(\int_0^T H(s)\,\mathrm{d}B(s)\right)\Bigg|\,\mathcal{G}\right]\exp\left(-\frac{1}{2}\int_0^T H^2(s)\,\mathrm{d}s\right) = 1,$$

where $\mathcal{G}$ denotes the σ-Algebra generated by $\left\{H(s)\,|\,0 \le s \le T\right\}$. It follows that

$$\mathbb{E}\left[Z(T)\right] = \mathbb{E}\left[\mathbb{E}\left[Z(T)\,|\,\mathcal{G}\right]\right] = 1.$$

□

Corollary 2.39. *Let* $B = \begin{pmatrix} B^1 \\ B^2 \end{pmatrix}$ *be a two-dimensional Brownian motion and* $\{\mathcal{F}_t\}$ *the augmented, natural filtration of* B. *Furthermore, let* H *be a continuous, square integrable,* $\{\mathcal{F}_t\}$*-adapted process which is independent of* B^1 *and* f *a continuous, bounded function. Then for*

$$\nu_f(t) := \begin{pmatrix} H(t) \\ f\left(H(t)\right) \end{pmatrix}$$

the equation

$$\mathbb{E}\left[\exp\left(\int_0^T \nu_f(s)^\top \,\mathrm{d}B(s) - \frac{1}{2}\int_0^T \|\nu_f(s)\|^2 \,\mathrm{d}s\right)\right] =$$
$$\mathbb{E}\left[\exp\left(\int_0^T f(H(s)) \,\mathrm{d}B^2(s) - \frac{1}{2}\int_0^T f(H(s))^2 \,\mathrm{d}s\right)\right],$$

holds true.

Proof. Define

$$\chi(t) := \int_0^t f(H(s)) \,\mathrm{d}B^2(s) - \frac{1}{2}\int_0^t f(H(s))^2 \,\mathrm{d}s,$$

and let $\mathcal{G}$ be the σ-algebra generated by $\left\{H(s) \,|\, 0 \le s \le T\right\}$ and $\mathcal{H}$ the σ-algebra generated by $\mathcal{G}$ and $\chi(T)$. Using Theorem 2.38 and the independence of B^1 and $\chi(T)$ we get

$$\mathbb{E}\left[\exp\left(\int_0^T \nu_f(s)^\top \,\mathrm{d}B(s) - \frac{1}{2}\int_0^T \|\nu_f(s)\|^2 \,\mathrm{d}s\right)\right] =$$
$$\mathbb{E}\left[\exp\left(\int_0^T H(s) \,\mathrm{d}B^1(s) - \frac{1}{2}\int_0^T H(s)^2 \,\mathrm{d}s\right)\exp\left(\chi(T)\right)\right] =$$
$$\mathbb{E}\left[\mathbb{E}\left[\exp\left(\int_0^T H(s) \,\mathrm{d}B^1(s) - \frac{1}{2}\int_0^T H(s)^2 \,\mathrm{d}s\right)\middle|\,\mathcal{H}\right]\exp\left(\chi(T)\right)\right] =$$
$$\mathbb{E}\left[\mathbb{E}\left[\exp\left(\int_0^T H(s) \,\mathrm{d}B^1(s) - \frac{1}{2}\int_0^T H(s)^2 \,\mathrm{d}s\right)\middle|\,\mathcal{G}\right]\exp\left(\chi(T)\right)\right] =$$
$$\mathbb{E}\left[\exp\left(\int_0^T f(H(s)) \,\mathrm{d}B^2(s) - \frac{1}{2}\int_0^T f(H(s))^2 \,\mathrm{d}s\right)\right].$$

□

Remark 2.40. This result is very important, since it allows the identification of martingales, even if Novikov's Condition is not satisfied.

2.2 Affine Processes

Remark 2.41. For the rest of this section, if not mentioned differently, the described processes are defined on a filtered probability space $\left(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P}\right)$, where the filtration satisfies the usual conditions. Furthermore, they are living on a finite time interval $[0, T]$ with some constant time horizon $T \in (0, \infty)$.

Definition 2.42. Let $\mathcal{H} \subset \mathbb{R}^d$ be a closed subspace, $W(t)$ a standard Brownian motion, $b : \mathcal{H} \to \mathbb{R}^d$ a continuous function and $\sigma : \mathcal{H} \to \mathbb{R}^{d\times d}$ a measurable function such that $a(x) = \sigma(x)\sigma(x)^\top$ is continuous. Furthermore, assume that the interior of $\mathcal{H}$ is non-empty and the existence of a unique solution X of the SDE

$$\mathrm{d}X(t) = b(X(t))\mathrm{d}t + \sigma(X(t))\,\mathrm{d}W(t), \quad X(0) = x_0 \in \mathcal{H}.$$

We call $\mathcal{H}$ state space of X.

Definition 2.43. A stochastic process X like in Definition 2.42 is called affine if there exists a complex valued function $\phi(t, u)$ and a function $\psi(t, u)$ which takes values in $\mathbb{C}^d$, such that for all $t \leq T$ and $u \in i\mathbb{R}^d$

$$\mathbb{E}\left[\exp\left(u^\top X(T)\right)\middle|\,\mathcal{F}_t\right] = \exp\left(\phi(T-t, u) + \psi(T-t, u)^\top X(t)\right).$$

Theorem 2.44. *Let X be an affine process. Then the functions $a(x)$ and $b(x)$ of Definition 2.42 are affine linear i.e. there exist matrices $A, A_1, \ldots, A_d$ and vectors $b, b_1, \ldots, b_d$ such that for every* $x = \begin{pmatrix} x_1 \\ \vdots \\ x_d \end{pmatrix}$

$$a(x) = A + \sum_{i=1}^{d} A_i x_i \quad \text{and} \quad b(x) = b + \sum_{i=1}^{d} b_i x_i.$$

Theorem 2.45. *The functions $\psi(t, u) = \left(\psi_1(t, u), \cdots, \psi_d(t, u)\right)^\top$ and $\phi(t, u)$ of Definition 2.43 solve the system of Riccati equations*

$$\frac{\partial}{\partial t}\phi(t, u) = \frac{1}{2}\psi(t, u)^\top A\psi(t, u) + b^\top\psi(t, u), \quad \phi(0, u) = 0,$$
$$\frac{\partial}{\partial t}\psi_i(t, u) = \frac{1}{2}\psi(t, u)^\top A_i\psi(t, u) + b_i^\top\psi(t, u), \quad \psi_i(0, u) = u_i.$$

Theorem 2.46. *Assume there is a process X like in Definition 2.42, whose coefficient functions a and b are affine and the corresponding Riccati equations*

$$\frac{\partial}{\partial t}\phi(t,u) = \frac{1}{2}\psi(t,u)^\top A\psi(t,u) + b^\top\psi(t,u), \quad \phi(0,u) = 0,$$
$$\frac{\partial}{\partial t}\psi_i(t,u) = \frac{1}{2}\psi(t,u)^\top A_i\psi(t,u) + b_i^\top\psi(t,u), \quad \psi_i(0,u) = u_i \quad 1 \le i \le d.$$

have a solution, satisfying $\mathcal{R}\left(\phi(t,u) + \psi(t,u)^\top x\right) \le 0$ for all $t \ge 0$, $u \in i\mathbb{R}^d$ and $x \in \mathcal{H}$. Then X is affine. Here $\mathcal{R}(z)$ denotes the real part of the complex number z.

Definition 2.47. For $m \le d \in \mathbb{N}$, we define the canonical state space by

$$\mathcal{H} := \mathbb{R}^m_{\ge 0} \times \mathbb{R}^{d-m}.$$

Permanent Assumption 2.48. From now on we assume that the state space of our affine processes is the canonical state space.

Definition 2.49. Let $I = \{1,\dots,m\}$ and $J = \{m+1,\dots,d\}$. We say the constants $b, b_1, \dots b_d \in \mathbb{R}^d$ and $A, A_1, \dots, A_d \in \mathbb{R}^{d\times d}$ satisfy the admissibility conditions if

- $A, A_1, \dots, A_d$ are symmetric and positive semi-definite,
- $A(ij) = 0$ for all $i, j \in I$,
- A_i =0 for all $i \in J$,
- $A_i(lk) = 0$ for $k \in I \setminus \{i\}$ and all i, l,
- $b \in \mathbb{R}^m_{\ge 0} \times \mathbb{R}^{d-m}$,
- $b_j(i) = 0$ for all $i \in I$ and $j \in J$,
- $b_i(j) \ge 0$ for all $i, j \in I$ with $j \ne i$.

Here $b_i(j)$ denotes the $j-th$ entry of the vector b_i and $A(ij)$ the entry in the i-th row and j-th column of the matrix A.

Theorem 2.50. *Let $m \le d \in \mathbb{N}$, and X be a process on the canonical state space $\mathcal{H} = \mathbb{R}^m_{\ge 0} \times \mathbb{R}^{d-m}$. Then X is affine if and only if it satisfies the SDE of Remark 2.41 and the corresponding coefficients satisfy the admissibility conditions from Definition 2.49.*

Remark 2.51. These admissibility conditions may seem artificial but are a direct consequence of the fact that the process X must not leave the canonical state space.

Remark 2.52. With Theorem 2.50, we have a powerful tool which we can use to identify an affine process on the canonical state space easily, without calculating its characteristic function.

2.2.1 Moment Generating Function of Affine Processes

Remark 2.53. One property of affine processes is that the corresponding moment generating function can nicely be computed if it exists, since it has the same form as the characteristic function.

Lemma 2.54. *Consider the system of Riccati equations*

$$\frac{\partial}{\partial t}\phi(t,u) = \frac{1}{2}\psi(t,u)^\top A\psi(t,u) + b^\top\psi(t,u), \quad \phi(0,u) = 0,$$
$$\frac{\partial}{\partial t}\psi_i(t,u) = \frac{1}{2}\psi(t,u)^\top A_i\psi(t,u) + b_i^\top\psi(t,u), \quad \psi_i(0,u) = u_i \quad 1 \le i \le d,$$

for some $u \in \mathbb{R}^d$ *Then there exists a lifetime* $T(u) \in (0,\infty]$, *such that there exists a unique solution* $F(\cdot,u) = \begin{pmatrix}\phi(\cdot,u)\\ \psi(\cdot,u)\end{pmatrix} : [0,T(u)) \to \mathbb{R}^{d+1}$. *This lifetime is maximal, i.e. either* $T(u) = \infty$ *or* $\lim_{t\uparrow T(u)} \|\psi(t,u)\| = \infty$. *In the second case we say* $\psi(\cdot,u)$ *explodes.*

Definition 2.55. For $t \ge 0$ we define

$$\mathcal{D}(t) := \left\{ u \in \mathbb{R}^d \,\middle|\, t < T(u) \right\},$$

the set of initial points u, such that there exists a real valued solution to the system of Riccati equations up to time t and

$$\mathcal{M}(t) := \left\{ u \in \mathbb{R}^d \,\middle|\, \mathbb{E}\left[\exp\left(u^\top X(t)\right)\right] < \infty \right\}$$

the set of u, such that the moment generating function of $X(t)$ exists.

Theorem 2.56. *Let* X *be affine with corresponding affine parameter functions* $a(x)$ *and* $b(x)$. *Then* $\mathcal{D}(t) = \mathcal{M}(t)$, *i.e. the moment generating function of* $X(t)$ *exists if and only if the system of Riccati equations has a real valued solution up to time* t.

Corollary 2.57. *Let* $u \in \mathbb{R}^d$, *and* X *be an affine process. If one side of the equation*

$$\mathbb{E}\left[\exp\left(u^\top X(T)\right)\middle|\mathcal{F}_t\right] = \exp\left(\phi(T-t,u) + \psi(T-t,u)^\top X(t)\right)$$

is well defined, the other is too and the equality holds true.

Theorem 2.58. *Consider the one-dimensional Riccati equation*

$$\frac{\partial}{\partial t}\psi(t,z) = A\psi(t,z)^2 + B\psi(t,z) + C, \quad \psi(0,z) = z,$$

with $A \neq 0$, $z \in \mathbb{C}$ and $B^2 - 4AC \neq 0$. Choose Δ to be either $\sqrt{B^2 - 4AC}$ or $-\sqrt{B^2 - 4AC}$ and define $A^{\pm} := \frac{-B \pm \Delta}{2A}$. In the case $z = A^{\pm}$ the solution is given as $\psi(t,z) = z$. For $z \neq A^{\pm}$ the solution is

$$\psi(t,z) = A^{-} - \frac{\Delta g(z) e^{-\Delta t}}{A(1 - g(z)e^{-\Delta t})},$$

with $g(z) = \frac{z - A^{-}}{z - A^{+}}$. This solution is unique and exists until the first time $t > 0$ such that $1 - g(z)e^{-\Delta t} = 0$.

Corollary 2.59. *Assume $A^{\pm} \neq 0$. If $z = 0$, the solution can be rewritten as*

$$\psi(t,0) = -2C \frac{e^{-\Delta t} - 1}{(B + \Delta)e^{-\Delta t} - B + \Delta}.$$

Proof. The constant $g(0) = \frac{A^{-}}{A^{+}}$ can be rewritten as $g(0) = \frac{B+\Delta}{B-\Delta}$. Since $A^{\pm} \neq 0$, this is well defined. If we use this in the representation of $\psi(t) := \psi(t,0)$ we get

$$\begin{aligned}
\psi(t) &= A^{-} - A^{-1} \frac{\Delta g(0) e^{-\Delta t}}{1 - g(0)e^{-\Delta t}} = - \frac{\Delta (B+\Delta) e^{-\Delta t}}{A(B - \Delta - (B+\Delta)e^{-\Delta t})} - \frac{B + \Delta}{2A} \\
&= - \frac{(B+\Delta)(B - \Delta - (B+\Delta)e^{-\Delta t}) + 2\Delta (B+\Delta) e^{-\Delta t}}{2A(B - \Delta - (B+\Delta)e^{-\Delta t})} \\
&= - \frac{B^2 - \Delta^2 - (B+\Delta)^2 e^{-\Delta t} + 2\Delta (B+\Delta) e^{-\Delta t}}{2A(B - \Delta - (B+\Delta)e^{-\Delta t})} \\
&= - \frac{4AC - B^2 e^{-\Delta t} - 2B\Delta e^{-\Delta t} - \Delta^2 e^{-\Delta t} + 2\Delta B e^{-\Delta t} + 2\Delta^2 e^{-\Delta t}}{2A(B - \Delta - (B+\Delta)e^{-\Delta t})} \\
&= -2C \frac{e^{-\Delta t} - 1}{(B+\Delta)e^{-\Delta t} - B + \Delta}.
\end{aligned}$$

□

Corollary 2.60. *Let $u \in \mathbb{R}$. The one-dimensional Riccati equation of Theorem 2.58, with real parameters A, B and C has a real valued solution $\psi(\cdot, u)$ if $B^2 \geq 4AC$. This solution is unique and exists until the first time t such that*

$$1 - g(u)e^{-\Delta t} = 0,$$

where g is defined like in Theorem 2.58.

Example 2.1. A well-known example of an affine process is the Cox-Ingersoll-Ross process, short CIR process, on $[0,T]$, defined by its SDE

$$\mathrm{d}z(t) = \kappa\left(\theta - z(t)\right)\mathrm{d}t + \varsigma\sqrt{z(t)}\,\mathrm{d}W(t), \quad z(0) = z_0 > 0.$$

Here κ, θ and ς are positive constants, and $W(t)$ is a one-dimensional Brownian motion. This process lives on the canonical state space $\mathcal{H} = \mathbb{R}_{\geq 0}$ and, if the Feller conditions $2\kappa\theta > \varsigma^2$ are satisfied, it stays almost surely positive.

The parameters are given by $b = \kappa\theta$, $b_1 = -\kappa$, $A = 0$ and $A_1 = \varsigma^2$ which satisfy the admissibility conditions, hence this process is affine. For $u \in \mathbb{R}$, the corresponding Riccati equations are given by

$$\begin{aligned}
\frac{\partial}{\partial t}\phi(t,u) &= \kappa\theta\psi(t,u), \quad \phi(0,u) = 0,\\
\frac{\partial}{\partial t}\psi(t,u) &= \frac{\varsigma^2}{2}\psi^2(t,u) - \kappa\psi(t,u), \quad \psi(0,u) = u.
\end{aligned}$$

Since the solution

$$\psi(t,u) = \frac{ze^{-\kappa t}}{1 - 2z\tau(t)},$$

with $\tau(t) = \dfrac{\varsigma^2}{4\kappa}\left(1 - e^{-\kappa t}\right)$, does not explode, we can calculate the moment generating function of $z(t)$ for all $t \geq 0$.

Example 2.2. Another example, where the theory of affine processes is useful, is the calculation of the moment generating function of the process

$$\tilde{X}(t) = \int_0^t z(s)\,\mathrm{d}s,$$

where z denotes the CIR process. This process is not affine but the two-dimensional process $X(t) = \begin{pmatrix} z(t) \\ \tilde{X}(t) \end{pmatrix}$ is and the moment generating function of $\tilde{X}(T)$ in the point $\tilde{u}$ is equal to the moment generating function of the affine process X in the point $u = \begin{pmatrix} 0 \\ \tilde{u} \end{pmatrix}$. The affine coefficients of X are given by $b = \begin{pmatrix} \kappa\theta \\ 0 \end{pmatrix}$, $b_1 = \begin{pmatrix} -\kappa \\ 1 \end{pmatrix}$, $b_2 = 0$, $A = 0$, $A_1 = \begin{pmatrix} \varsigma^2 & 0 \\ 0 & 0 \end{pmatrix}$ and $A_2 = 0$, hence the corresponding Riccati equations in the point u are

$$\begin{aligned}
\frac{\partial}{\partial t}\phi(t,u) &= \kappa\theta\psi_1(t,u), \quad \phi(0,u) = 0,\\
\frac{\partial}{\partial t}\psi_1(t,u) &= \frac{\varsigma^2}{2}\psi_1^2(t,u) - \kappa\psi_1(t,u) + \psi_2(t,u), \quad \psi_1(0,u) = 0,\\
\frac{\partial}{\partial t}\psi_2(t,u) &= 0, \quad \psi_2(0,u) = \tilde{u}.
\end{aligned}$$

Since the differential equation of ψ_2 is trivial, this system can be reduced to the one-dimensional case

$$\frac{\partial}{\partial t}\phi(t,0) = \kappa\theta\psi_1(t,u), \quad \phi(0,u) = 0,$$

$$\frac{\partial}{\partial t}\psi_1(t,u) = \frac{\varsigma^2}{2}\psi_1^2(t,u) - \kappa\psi_1(t,u) + \tilde{u}, \quad \psi_1(0,u) = 0,$$

which has for $\tilde{u} < \frac{\kappa^2}{2\varsigma^2}$ a real valued solution. For $\tilde{u} \in (0, \frac{\kappa^2}{2\varsigma^2})$, we have that $\kappa^2 > \kappa^2 - 2\varsigma\tilde{u} > 0$, hence $\kappa > |\Delta|$ and consequently $\frac{-\kappa+\Delta}{\kappa+\Delta} < 0$. This gives us

$$1 - g(0)e^{-\Delta t} = 1 - \frac{-\kappa + \Delta}{\kappa + \Delta}e^{-\Delta t} > 0,$$

hence the real valued solution exists for all $t \geq 0$ and so does the moment generating function, which is of the form

$$\mathbb{E}\left[\exp\left(\tilde{u}\tilde{X}(T)\right)\middle|\mathcal{F}_t\right] = \exp\left(\phi(T-t,u) + \psi_1(T-t,u)z(t) + \tilde{u}\tilde{X}(t)\right).$$

Remark 2.61. We will also call such degenerated high-dimensional Riccati equations one-dimensional since they are solved in the same way.

3 Diffusion Market Models

Since we have now developed the probabilistic theory needed, we will introduce diffusion market models. These well-known market models, often called Brownian market models, are characterized by their special form consisting of two major parameters, the drift term b and the diffusion matrix σ. Following Karatzas and Shreve (2003), we will further specify important features of market models, like arbitrage freeness and completeness, and specify when diffusion market models have these properties.

3.1 Market Models

Definition 3.1. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, $T > 0$ be a deterministic time point, called terminal time and $W(t)$ a d-dimensional Brownian motion on this probability space with augmented filtration $\{\mathcal{F}_t\}$. A market model consists of one risk free asset, sometimes called bank account, with price process $S_0(t)$, and $m \leq d$ risky assets, called stocks, with price processes $S_i(t)$. These price processes satisfy the stochastic differential equations

$$\mathrm{d}S_0(t) = r(t)S_0(t)\mathrm{d}t, \quad S_0(0) = 1,$$

$$\mathrm{d}S_i(t) = S_i(t)\left(b_i(t)\mathrm{d}t + \sum_{j=1}^{d} \sigma_{ij}(t)\, \mathrm{d}W^j(t)\right), \quad S_i(0) = S_i > 0 \quad 1 \leq i \leq m.$$

The coefficient processes $r(t)$, $b(t) := \begin{pmatrix} b_1(t) \\ \vdots \\ b_m(t) \end{pmatrix}$ and $\sigma(t) = \begin{pmatrix} \sigma_{11}(t) & \cdots & \sigma_{1d}(t) \\ \vdots & \vdots & \vdots \\ \sigma_{m1}(t) & \cdots & \sigma_{md}(t) \end{pmatrix}$ are assumed to be progressively measurable with respect to the filtration $\{\mathcal{F}_t\}$. Furthermore, they $\mathbb{P}$-almost surely satisfy

$$\int_0^T |r(t)|\, \mathrm{d}t \leq L \in \mathbb{R},$$

$$\int_0^T \|b(t)\|\, \mathrm{d}t < \infty.$$

We sometimes write the combined price process of all m risky assets as $S(t)$. This process satisfies the m-dimensional stochastic differential equation

$$\mathrm{d}S(t) = S(t)\left(b(t)\mathrm{d}t + \sigma(t)\,\mathrm{d}W(t)\right), \quad S(0) = \begin{pmatrix} S_1 \\ \vdots \\ S_m \end{pmatrix}.$$

Remark 3.2. For now, we will not further restrict the volatility process $\sigma(t)$ since these restrictions highly influence the properties of our market.

Remark 3.3. For convenience, we sometimes write almost surely instead $\mathbb{P}$-almost surely. If we want a property to hold almost surely with respect to some different measure, this will be explicitly stated.

Definition 3.4. An absolute portfolio process (φ_0, φ) consists of a $\mathbb{R}$-valued, $\{\mathcal{F}_t\}$-progressively measurable process φ_0 and a $\mathbb{R}^m$-valued, $\{\mathcal{F}_t\}$-progressively measurable process φ satisfying almost surely

$$\int_0^T |\varphi_0(t) + \varphi(t)^\top \underline{1}_m||r(t)|\,\mathrm{d}t < \infty,$$

$$\int_0^T |\varphi(t)^\top (b(t) - r(t)\underline{1}_m)|\,\mathrm{d}t < \infty,$$

$$\int_0^T \|\sigma(t)^\top \varphi(t)\|^2\,\mathrm{d}t < \infty.$$

Here, $\varphi_i(t)$, the i-th entry of $\varphi(t)$, can be interpreted as amount of money invested in asset i at time t and $\varphi_0(t)$ is the money invested in the bank account.

Definition 3.5. For an absolute portfolio process (φ_0, φ) and $t \in [0, T]$ we define the gains process by

$$G(t) := \int_0^t \varphi_0(s)r(s)\,\mathrm{d}s + \int_0^t \varphi(s)^\top b(s)\,\mathrm{d}s + \int_0^t \varphi(s)^\top \sigma(s)\,\mathrm{d}W(s).$$

Definition 3.6. The cumulative income process Γ is a semimartingale. The value $\Gamma(t)$ can be interpreted as the money received from an investor up to time t. The value $\Gamma(0)$ is called initial wealth or starting capital.

Remark 3.7. While an increase of Γ can be seen as further money invested, a decrease can be interpreted as consumption of the original investor.

Definition 3.8. For an absolute portfolio process we define the wealth process by

$$X(t) = G(t) + \Gamma(t).$$

This means the wealth consists of the money we gained by investing, namely $G(t)$, and the money invested, $\Gamma(t)$.

Definition 3.9. An absolute portfolio process (φ_0, φ) is called Γ-financed if the wealth process satisfies

$$X(t) = \varphi_0(t) + \varphi(t)^\top \underline{1}_m.$$

In this case, the dynamics of X are given by

$$\mathrm{d}X(t) = \mathrm{d}\Gamma(t) + X(t)r(t)\mathrm{d}t + \varphi(t)^\top \left(b(t) - r(t)\underline{1}_m\right) \mathrm{d}t + \varphi(t)^\top \sigma(t)\, \mathrm{d}W(t).$$

Definition 3.10. We call an absolute portfolio process self-financed if it is financed by the constant cumulative income process $\Gamma(t) = x$ for some deterministic $x \in \mathbb{R}$. This means we get some initial capital, but after time 0 there is no further investment nor consumption of wealth.

Remark 3.11. Our definition of self-financing slightly differs from the definition of Karatzas and Shreve (2003) since we allow some initial capital $x \neq 0$.

Definition 3.12. An absolute portfolio process is called tame if there is a constant $K \in \mathbb{R}$, such that the discounted gains process satisfies almost surely

$$\frac{G(t)}{S_0(t)} \geq K.$$

Definition 3.13. An arbitrage opportunity is a self-financed, tame absolute portfolio process, such that the corresponding wealth process satisfies

- $X(0) \leq 0$ almost surely,
- $X(T) \geq 0$ almost surely,
- $\mathbb{P}[X(T) > 0] > 0$.

Remark 3.14. We explicitly allow borrowing and short selling.

Remark 3.15. Our market model is frictionless which means there are no taxes or fees. Furthermore, we assume that there are no restrictions on the amount in which stocks can be bought and sold.

Definition 3.16. We call a d-dimensional, $\{\mathcal{F}_t\}$ progressively measurable process θ market price of risk if it satisfies

$$\sigma(t)\theta(t) = b(t) - r(t)\underline{1}_m \quad \textit{almost surely.}$$

Theorem 3.17. *If there is no arbitrage in our market model, then there exists a market price of risk.*

Theorem 3.18. *If there exists a square integrable market price of risk which satisfies*

$$\mathbb{E}\left[\exp\left(-\int_0^T \theta(s)^\top \mathrm{d}W(s) - \frac{1}{2}\int_0^T \|\theta(s)\|^2 \,\mathrm{d}s\right)\right] = 1,$$

then our market model is arbitrage free.

Lemma 3.19. *The matrix-valued process $\sigma(t)$ is invertible for all t almost surely if and only if there exists an almost surely unique market price of risk θ.*

Proof. If $\sigma(t)$ is invertible for all t almost surely then

$$\theta(t) = \sigma(t)^{-1} \left(b(t) - r(t)\underline{1}_m\right)$$

is well defined and almost surely unique, since all processes on the right hand side are. On the other hand if $\sigma(t)$ does not have this property, either because $m < d$ or because there is some positive probability that there exists some t such that $\sigma(t)$ is not invertible, then there exists some process $0 \neq \nu$ in the kernel of σ. The process $\tilde{\theta} = \theta + \nu$ satisfies the equation which defines the market price of risk, hence it is not unique. □

Definition 3.20. Let $\mathbb{Q}$ be a probability measure equivalent to $\mathbb{P}$. We call $\mathbb{Q}$ equivalent martingale measure, or short martingale measure, if for all $1 \leq i \leq m$, the discounted price process $\dfrac{S_i(t)}{S_0(t)}$ is a local martingale with respect to $\mathbb{Q}$.

Theorem 3.21. *The market model is arbitrage free if and only if there exists an equivalent martingale measure $\mathbb{Q}$.*

Remark 3.22. If all assumptions of Theorem 3.18 are fulfilled, we have, by Girsanov's Theorem that there exists an equivalent measure with density

$$Z(t) = \exp\left(-\int_0^t \theta(s)^\top \,\mathrm{d}W(s) - \frac{1}{2}\int_0^t \|\theta(s)\|^2 \,\mathrm{d}s\right)$$

and the discounted price processes are local martingales under this measure.

Definition 3.23. A market model is called complete if for every $\mathcal{F}_T$ measurable random variable B, with $\dfrac{B}{S_0(T)} \geq K$ for some $K \in \mathbb{R}$, there exists a tame, self-financed absolute portfolio process such that the corresponding wealth process satisfies

$$X(T) = B.$$

If this property is not satisfied, we say the market is incomplete.

Theorem 3.24. *A market model is complete if and only if there exists a unique equivalent martingale measure $\mathbb{Q}$.*

Theorem 3.25. *An arbitrage free diffusion market model is complete if and only if $m = d$ and $\sigma(t)$ is invertible for all t almost surely.*

Remark 3.26. This property allows us to identify complete diffusion market models, without calculating an equivalent martingale measure and proving its uniqueness.

3.2 Relative Portfolio Processes

Definition 3.27. We define the relative portfolio process or short portfolio process (π_0, π) of a self-financing absolute portfolio process (φ_0, φ) by

$$\pi_i(t) = \frac{\varphi_i(t)}{X(t)} \quad 0 \le i \le m.$$

The value $\pi_i(t)$ can be interpreted as the fraction of total wealth invested in asset i.

Remark 3.28. We will now switch to relative portfolio processes, because of some nice properties like the representation of the discounted wealth process which is determined in Corollary 3.33.

Lemma 3.29. *A portfolio process satisfies*

$$\pi_0(t) = 1 - \pi(t)^\top \underline{1}_m.$$

In abuse of terminology, we sometimes denote a portfolio process just by π*, which is possible since* π_0 *is uniquely determined by* π*.*

Proof. Let (π_0, π) be a portfolio process, and (φ_0, φ) the corresponding absolute portfolio process which is self-financing, hence the wealth process satisfies

$$X(t) = x + G(t) = \varphi_0(t) + \varphi(t)^\top \underline{1}_m.$$

Therefore, we get

$$\pi_0(t) + \pi(t)^\top \underline{1}_m = \frac{1}{X(t)} \left(\varphi_0(t) + \varphi(t)^\top \underline{1}_m \right) = 1.$$

□

Lemma 3.30. *The dynamics of the wealth process of a portfolio process* π *are given by*

$$\mathrm{d}X(t) = X(t) \left(\pi_0(t) r(t) \mathrm{d}t + \pi(t)^\top \left(b(t) \mathrm{d}t + \sigma(t)\, \mathrm{d}W(t) \right) \right).$$

Proof. Let π be a portfolio process, with corresponding self-financed absolute portfolio process (φ_0, φ). Using the representation of Definition 3.9, and abusing the fact that $\mathrm{d}\Gamma(t) = 0$ we get

$$\begin{aligned} \mathrm{d}X(t) &= X(t) r(t) \mathrm{d}t + \varphi(t)^\top \left(b(t) - r(t) \underline{1}_m \right) \mathrm{d}t + \varphi(t)^\top \sigma(t)\, \mathrm{d}W(t) \\ &= X(t) \left(r(t)(1 - \pi(t)^\top \underline{1}_m) \mathrm{d}t + \pi(t)^\top \left(b(t) \mathrm{d}t + \sigma(t)\, \mathrm{d}W(t) \right) \right) \\ &= X(t) \left(\pi_0(t) r(t) \mathrm{d}t + \pi(t)^\top \left(b(t) \mathrm{d}t + \sigma(t)\, \mathrm{d}W(t) \right) \right). \end{aligned}$$

□

Remark 3.31. To make identification of different wealth process easier, we will from now on denote the wealth process of the portfolio process π by X^π, or sometimes $X^{x,\pi}$ if we want to mention the starting capital x.

Lemma 3.32. *Let π be a portfolio process with initial capital $x > 0$. The corresponding discounted wealth process $\frac{X^{x,\pi}(t)}{S_0(t)}$ satisfies representation*

$$x \exp\left(\int_0^t \pi(s)^\top \sigma(s)\,\mathrm{d}W(s) - \frac{1}{2}\int_0^t \|\sigma(s)^\top \pi(s)\|^2\,\mathrm{d}s + \int_0^t \pi(s)^\top \left(b(s) - r(s)\underline{1}_m\right)\mathrm{d}s\right).$$

In the case of an arbitrage free market this can be rewritten as

$$\frac{X^{x,\pi}(t)}{S_0(t)} = x\exp\left(\int_0^t \pi(s)^\top \sigma(s)\,\mathrm{d}W(s) - \frac{1}{2}\int_0^t \|\sigma(s)^\top \pi(s)\|^2\,\mathrm{d}s + \int_0^t \pi(s)^\top \sigma(s)\theta(s)\,\mathrm{d}s\right),$$

where $\theta(t)$ denotes a market price of risk.

Proof. By Lemma 3.30, we get that the dynamics of $\frac{X^{x,\pi}(t)}{S_0(t)}$ are given by

$$\mathrm{d}\left(X^{x,\pi}(t)S_0(t)^{-1}\right) = X^{x,\pi}(t)S_0(t)^{-1}\left(\pi(t)^\top\left(b(t) - r(t)\underline{1}_m\right)\mathrm{d}t + \pi(t)^\top\sigma(t)\,\mathrm{d}W(t)\right).$$

This is the equation of a stochastic exponential process, hence the solution $\frac{X^{x,\pi}(t)}{S_0(t)}$ is given by

$$x\exp\left(\int_0^t \pi(s)^\top\left(b(s) - r(s)\underline{1}_m\right)\mathrm{d}s + \int_0^t \pi(s)^\top\sigma(s)\,\mathrm{d}W(s) - \frac{1}{2}\int_0^t \|\sigma(s)^\top\pi(s)\|^2\,\mathrm{d}s\right).$$

The result for arbitrage free markets is due to the equation

$$\sigma(t)\theta(t) = b(t) - r(t)\underline{1}_m.$$

□

Corollary 3.33. *The discounted wealth process has the representation*

$$\frac{X^{x,\pi}(t)}{S_0(t)} = x\exp\left(\int_0^t \pi(s)^\top\sigma(s)\,\mathrm{d}\tilde{W}(s) - \frac{1}{2}\int_0^t \|\sigma(s)^\top\pi(s)\|^2\,\mathrm{d}s\right),$$

where $\tilde{W}(t) = W(t) + \int_0^t \theta(t)\,\mathrm{d}t$.

Proof. By Lemma 3.32 we get that

$$\frac{X^{x,\pi}(t)}{S_0(t)} = x\exp\left(\int_0^t \pi(s)^\top\sigma(s)\,\mathrm{d}W(s) - \frac{1}{2}\int_0^t \|\sigma(s)^\top\pi(s)\|^2\,\mathrm{d}s + \int_0^t \pi(s)^\top\sigma(s)\theta(s)\,\mathrm{d}s\right),$$

which can be rewritten as

$$\frac{X^{x,\pi}(t)}{S_0(t)} = x\exp\left(\int_0^t \pi(s)^\top\sigma(s)\,\mathrm{d}\tilde{W}(s) - \frac{1}{2}\int_0^t \|\sigma(s)^\top\pi(s)\|^2\,\mathrm{d}s\right).$$

□

Example 3.1. One of the most basic examples of a diffusion market model is the Black-Scholes market model, described by Black and Scholes (1973). In its easiest form it consists of a bank account, with deterministic interest rate $r \geq 0$ and starting value 1, i.e. its price process is given by

$$S_0(t) = e^{rt}.$$

Furthermore, we have an one-dimensional Brownian motion W and a single risky asset, with price process

$$\mathrm{d}S(t) = S(t)\left(\mu\mathrm{d}t + \sigma\,\mathrm{d}W(t)\right), \quad S(0) = S > 0,$$

where μ is a real constant, σ and S are positive constants. In this case, the process $S(t)$ can be explicitly written as

$$S(t) = S\exp\left(\left(\mu - \frac{\sigma^2}{2}\right)t + \sigma W(t)\right).$$

The unique market price of risk is a deterministic constant and given by

$$\theta = \frac{\mu - r}{\sigma}.$$

Since Novikov's condition is satisfied, we have by Girsanov's theorem, the existence of an equivalent measure $\mathbb{Q}$, with density process

$$Z(t) = \exp\left(-\theta W(t) - \frac{1}{2}\theta^2 t\right)$$

and the dynamics of the discounted price process of the risky asset are

$$\mathrm{d}\left(S(t)S_0(t)^{-1}\right) = S_0^{-1}S(t)\sigma\,\mathrm{d}\tilde{W}(t),$$

where $\tilde{W}$, defined by

$$\tilde{W}(t) = W(t) + \theta t,$$

is a standard Brownian motion under $\mathbb{Q}$. Therefore, the discounted price process is a local martingale under $\mathbb{Q}$. Since we have one Brownian motion, one risky asset and σ is invertible and strictly positive, the market is, by Theorem 3.25, complete.

Remark 3.34. Even though the Black-Scholes model has many advantages, for example a closed formula for pricing European options, it has some major drawbacks, like the constant volatility. This assumption stands in contrast to the volatility smiles observed in real data. One way to fix this issue is, to replace the constant by a stochastic volatility process.

Example 3.2. Another example of a diffusion market model is Heston's stochastic volatility model, or short Heston's model, introduced by Heston (1993). We consider two assets, where the bank account, again has the form $S_0(t) = e^{rt}$ for some constant $r \geq 0$. The risky asset has a price process satisfying the SDE

$$\mathrm{d}S(t) = S(t)\left(\left(\bar{\lambda}z(t) + r\right)\mathrm{d}t + \sqrt{z(t)}\,\mathrm{d}\hat{W}^1(t)\right), \quad S(0) = S > 0.$$

The slope parameter $\bar{\lambda}$ is a real constant and the volatility process z is a so called CIR process following the SDE

$$\mathrm{d}z(t) = \kappa(\theta - z(t))\mathrm{d}t + \varsigma\sqrt{z(t)}\,\mathrm{d}\hat{W}^2(t), \quad z(0) = z_0 > 0,$$

with positive constants κ, θ and ς. Here the two processes $\hat{W}^1$ and $\hat{W}^2$ are dependent Brownian motions with constant correlation $\rho \in (-1, 1)$.
The main idea of this model is, to fix one of the main shortcomings of the Black-Scholes model, namely the constant volatility. It assumes a stochastic, but observable volatility modelled by the CIR process z. Since there is only one risky asset which is influenced by two Brownian motions we can see that this model can never be complete, but it is still arbitrage free.

4 Time Continuous Portfolio Optimization

In this chapter, we will establish the theoretical framework which we will use for the expected utility maximization of terminal wealth in Heston's market model. Our approach will be based on the martingale method mainly introduced by Pliska (1986), Karatzas, Lehoczky, and Shreve (1987), and Cox and Huang (1989).
The chapter roughly separates into two parts. Firstly, we will look at the utility theory and some results for complete markets of diffusion type. Here we refer to Karatzas and Shreve (2003), who in contrast to us do not restrict themselves to the optimization of terminal expected utility, but also consider utility from consumption and expected utility from both terminal wealth and consumption.
Since Heston's model does not fulfil the assumption of completeness, we will further show similar outcomes for incomplete markets in the second part. There we will follow closely to Karatzas, Lehoczky, Shreve, and Xu (1991), who extended the existing theory to incomplete markets of diffusion type. A compact summary of the most important theorems is also given by Pham (2009).

4.1 Utility Theory

Definition 4.1. A utility function is a strictly monotone increasing, strictly concave function $U \in \mathcal{C}^1$ which maps $\mathbb{R}_+$ onto $\mathbb{R}$. Furthermore, U must satisfy the Inada conditions[1]

$$\lim_{x \uparrow \infty} U'(x) = U'(\infty) = 0,$$

$$\lim_{x \downarrow 0} U'(x) = U'(0) = \infty.$$

Remark 4.2. Sometimes we will extend our utility function to a function $U : \mathbb{R} \to \mathbb{R} \cup \{-\infty, \infty\}$ without further mentioning. In this case, we set $U(x) = -\infty$ for all negative x and $U(0) = \lim_{x \downarrow 0} U(x)$. Notice that this function might no longer be continuous, but it is still concave and monotone increasing.

[1]Kramkov and Schachermayer 1999, p. 906.

Lemma 4.3. *Let U be a utility function. The function U' has a monotone decreasing, continuous inverse function $I : (0, \infty) \to (0, \infty)$ satisfying*

$$I(0) := \lim_{x \downarrow 0} I(x) = \infty,$$

$$I(\infty) := \lim_{x \uparrow \infty} I(x) = 0.$$

Proof. The utility function U is strictly concave and continuously differentiable, and consequently U' is strictly monotone decreasing and continuous. Accordingly, there exists a monotone decreasing and continuous inverse function.
Since U is strictly monotone increasing, it follows that $U'(x) \in (0, \infty)$ for all $x > 0$. Hence I maps $(0, \infty)$ to $(0, \infty)$. The limits are an immediate consequence of the corresponding limits of the function U'. □

Definition 4.4. For a convex function $f : \mathbb{R} \to \mathbb{R}$, the conjugate function is given as

$$f^*(y) = \sup_{x \in \mathbb{R}} (xy - f(x)).$$

For a utility function U we define the convex dual as

$$\tilde{U}(y) = (-U)^*(-y).$$

Remark 4.5. Since U is concave, $-U$ is convex. Hence, $\tilde{U}$ is well defined if we allow values in the extended real numbers.
For positive y the convex dual of U has the form

$$\tilde{U}(y) = \sup_{x \in \mathbb{R}} (U(x) - xy),$$

whereas for negative y we have $\tilde{U}(y) = \infty$.

Lemma 4.6. *The convex conjugate of a utility function U is lower semicontinuous and convex.*

Proof. All functions in the set $B = \left\{ f_x(y) := U(x) - xy \,|\, x \geq 0 \right\}$ are affine linear, hence convex, and continuous. Since $U(x) - xy = -\infty$ for all $x < 0$, $\tilde{U}(y)$ is the point-wise supremum of functions in B and consequently convex and lower semicontinuous. □

Lemma 4.7. *For all $y > 0$, the convex conjugate of a utility function has the representation*

$$\tilde{U}(y) = U(I(y)) - yI(y).$$

Proof. By the concavity and differentiability of U on $\mathbb{R}_+$ we have for all $x, y > 0$

$$U(x) \leq U(y) + U'(y)\,[x-y]$$

or equivalently

$$U(x) - U'(y)x \leq U(y) - yU'(y).$$

For $y > 0$ we have that

$$U(x) - xy = U(x) - xU'(I(y)) \leq U(I(y)) - I(y)y,$$

hence the supremum in the definition of the convex dual is attained in $x = I(y)$. Since $U(x) - xy$ is a strictly concave function in x, the maximum is even unique. □

Lemma 4.8. *For every utility function U, the convex conjugate $\tilde{U}$, restricted on $(0, \infty)$, is in $\mathcal{C}^1$ and its derivative is given by*

$$\tilde{U}'(y) = -I(y).$$

Consequently, $\tilde{U}$ is monotone decreasing.

Proof. Let $\partial\tilde{U}(y)$ denote the subdifferential of $\tilde{U}$ in $y \in (0, \infty)$. A convex function is differentiable in a point if and only if the subdifferential in this point contains only a single element. By Theorem 23.5. in Rockafellar (1997), we get that $x \in \partial\tilde{U}(y)$ if and only if $U(z) - zy$ attains its maximum over z in $-x$. By the proof of Lemma 4.7 we have that this function is uniquely maximized at $z = I(y)$. It follows that $\tilde{U}$ is differentiable in every point y and has derivative $-I(y)$. The monotonicity is an immediate consequence of the positivity of I. □

Lemma 4.9. *For all $x \in \mathbb{R}$ we have the relationship*

$$U(x) = \inf_{y\in\mathbb{R}}(\tilde{U}(y) + xy).$$

In the case of $x > 0$ this supremum is attained at $y = U'(x)$ and the equation can be rewritten as

$$U(x) = \tilde{U}(U'(x)) + U'(x)x.$$

Proof. For any convex function f the equality $f^{**} = f$ holds true[2]. It follows that

$$-U(x) = \sup_{y\in\mathbb{R}} xy - (-U)^*(y) = \sup_{y\in\mathbb{R}} -xy - (-U)^*(-y) = -\inf_{y\in\mathbb{R}}\left(\tilde{U}(y) + xy\right).$$

The second equality is due to Lemma 4.7 with $y = U'(x)$. □

[2] Rockafellar 1997, p. 104.

4.2 Portfolio Optimization in Complete Markets

Permanent Assumption 4.10. For this section, we assume that the market is complete, and the market price of risk almost surely satisfies

$$\int_0^T \|\theta(t)\|^2 \, \mathrm{d}t < \infty.$$

Remark 4.11. The completeness of the market ensures the invertibility of our volatility coefficient matrix $\sigma(t)$ and the representation of the market price of risk as

$$\theta(t) = \sigma(t)^{-1}(b(t) - r(t)\underline{1}_d).$$

We further define the positive local martingale

$$Z(t) = \exp\left(-\int_0^t \theta(s)^\top \, \mathrm{d}W(s) - \frac{1}{2}\int_0^t \|\theta(s)\|^2 \, \mathrm{d}s \right).$$

For convenience we will from now on denote the discounting factor as $D(t) := \dfrac{1}{S_0(t)}$.

Problem 4.12. Let U be a utility function and let $\mathcal{A}(x)$ denote the set of portfolio processes π with starting capital $x > 0$ which satisfy $\mathbb{E}\left[U(X^\pi(T))^-\right] < \infty$, where $x^- := \max(0, -x)$. Then the value function of the portfolio optimization problem is given as

$$V(x) = \sup_{\pi \in \mathcal{A}(x)} \mathbb{E}\left[U(X^\pi(T))\right].$$

A portfolio process $\hat{\pi} \in \mathcal{A}(x)$ is called optimal or solution to our problem if it satisfies

$$V(x) = \mathbb{E}\left[U(X^{\hat{\pi}}(T))\right].$$

Remark 4.13. Our problem is similar to the optimization problems in Karatzas and Shreve (2003), who optimize over all admissible pairs consisting of a portfolio process π and a consumption process c. This discrepancy is because we only consider utility from terminal wealth. Since consumption reduces our wealth, a pair with $c \neq 0$ can never be optimal to our problem which is why we can simply ignore consumption.

Definition 4.14. For a utility function U we define the function $\mathcal{X}$ which maps $(0, \infty) \to (0, \infty]$ by

$$\mathcal{X}(y) := \mathbb{E}\left[D(T)Z(T)I\left(yD(T)Z(T)\right)\right].$$

Lemma 4.15. *Assume there exist some $\hat{\alpha} \in (0,1)$ and $\hat{\beta} \in (1,\infty)$ such that*

$$U'(\hat{\beta}x) \leq \hat{\alpha}U'(x) \quad \forall x > 0.$$

Then $I(\hat{\alpha}y) \leq \hat{\beta}I(y)$ for all $y > 0$. Additionally it holds that for all $\alpha \in (0,1)$ there exists some $\beta \in (1,\infty)$ such that

$$U'(\beta x) \leq \alpha U'(x) \quad \forall x > 0.$$

Proof. Substituting $U'(x) = y$ and applying I leads to the inequality $I(\hat{\alpha}y) \leq \hat{\beta}I(y)$ for all $y > 0$, since I is monotone decreasing.
For $\alpha > \hat{\alpha}$ we have

$$\alpha U'(x) \geq \hat{\alpha}U'(x) \geq U'(\hat{\beta}x).$$

Let $\alpha < \hat{\alpha}$ then there exists a $k \in \mathbb{N}$ such that $\alpha > \hat{\alpha}^k$. Using this leads to

$$\alpha U'(x) \geq \hat{\alpha}^k U'(x) \geq U'(\hat{\beta}^k x).$$

□

Remark 4.16. This property is satisfied by the power utility functions $U(x) = \frac{1}{\gamma}x^\gamma$, with parameter $\gamma \in (0,1)$ and the logarithmic utility function given by $U(x) = \log(x)$, for $x > 0$.

Lemma 4.17. *If there exist some constants $\hat{\alpha} \in (0,1)$ and $\hat{\beta} \in (1,\infty)$ such that $\hat{\alpha}U'(x) \geq U'(\hat{\beta}x)$ for all $x > 0$ and there is some $\hat{y} > 0$ such that $\mathcal{X}(\hat{y}) < \infty$, then $\mathcal{X}(y) < \infty$ holds true for all $y > 0$.*

Proof. By Lemma 4.15, we get that for all $\alpha \in (0,1)$ there exist a $\beta \in (1,\infty)$ such that $I(\alpha y) \leq \beta I(y)$ for all $y > 0$. In the case $y > \hat{y}$, we get by monotonicity of I and the expected value that

$$\mathcal{X}(y) \leq \mathcal{X}(\hat{y}) < \infty.$$

For $y < \hat{y}$ it holds that $\frac{y}{\hat{y}} \in (0,1)$, and consequently there exists a $\tilde{\beta} \in (0,1)$ satisfying $\frac{y}{\hat{y}}U'(x) > U'(\tilde{\beta}x)$. Consequently,

$$\mathbb{E}\Big[D(T)Z(T)I\left(yD(T)Z(T)\right)\Big] = \mathbb{E}\left[D(T)Z(T)I\left(\frac{y}{\hat{y}}\hat{y}D(T)Z(T)\right)\right] \leq \tilde{\beta}\mathcal{X}(\hat{y}) < \infty.$$

□

Permanent Assumption 4.18. From now on we will assume that $\mathcal{X}(y) < \infty$ holds for all $y > 0$.

Proposition 4.19. *The function $\mathcal{X}$ is monotone decreasing and continuous on $(0,\infty)$. Moreover, there exists an $a \in (0,\infty]$ such that $\mathcal{X}$ is strictly monotone on $(0,a)$.*

Remark 4.20. The strict monotonicity implies the existence of an inverse function $\mathcal{Y}$ of $\mathcal{X}$ which maps the interval $(\mathcal{X}(\infty), \infty)$ to $(0, a)$.

Proposition 4.21. *For every $\pi \in \mathcal{A}(x)$ the process $(D(t)Z(t)X^\pi(t))_{t\in[0,T]}$ is a $\mathbb{P}$-supermartingale satisfying*

$$\mathbb{E}\left[D(t)Z(t)X^\pi(t)\right] \le x.$$

Proof. Using the representation of Lemma 3.32 and defining the process μ by

$$\begin{aligned}\mu(t) = &\int_0^t \pi(s)^\top \sigma(s)\,\mathrm{d}W(s) - \frac{1}{2}\int_0^t \|\pi(s)^\top\sigma(s)\|^2\,\mathrm{d}s - \int_0^t \theta(s)^\top\,\mathrm{d}W(s)\\ &- \frac{1}{2}\int_0^t \|\theta(s)\|^2\,\mathrm{d}s + \int_0^t \pi(t)^\top\sigma(t)\theta(t)\,\mathrm{d}t\end{aligned}$$

we get

$$D(t)X^\pi(t)Z(t) = xe^{\mu(t)}.$$

Since we can rewrite μ as

$$\mu(t) = \int_0^t \left(\pi(s)^\top\sigma(s) - \theta(s)^\top\right)\,\mathrm{d}W(s) - \frac{1}{2}\int_0^t \|\sigma(s)^\top\pi(s) - \theta(s)\|^2\,\mathrm{d}s,$$

this is a positive local martingale, hence a supermartingale with

$$\mathbb{E}\left[D(t)X^\pi(t)Z(t)\right] \le x.$$

□

Lemma 4.22. *Let U be a utility function, I the inverse of its derivative, and $\mathcal{X}$ and $\mathcal{Y}$ defined as before. Define*

$$\xi(x) := I\left(\mathcal{Y}(x)D(T)Z(T)\right),$$

then $\xi(x)$ satisfies the following properties

1. $\mathbb{E}\left[D(T)Z(T)\xi(x)\right] = x$,
2. $\mathbb{E}\left[U(\xi(x))^-\right] < \infty$,
3. $\mathbb{E}\left[U(X^\pi(T))\right] \le \mathbb{E}\left[U(\xi(x))\right]$ *for all* $\pi \in \mathcal{A}(x)$.

Proof. The first equality is an immediate consequence of the definition of $\mathcal{Y}$, since we have

$$\mathbb{E}\left[D(T)Z(T)\xi(x)\right] = \mathbb{E}\left[D(T)Z(T)I\left(\mathcal{Y}(x)D(T)Z(T)\right)\right] = \mathcal{X}(\mathcal{Y}(x)) = x.$$

By concavity of U we get that

$$U(z) \geq U(y) + U'(z)\,[z-y]\,.$$

Using this property with $z = \xi(x)$ and $y = 1$ leads to

$$U(\xi(x)) \geq U(1) + \mathcal{Y}(x)D(T)Z(T)\left[\xi(x) - 1\right] \geq -|U(1)| - \mathcal{Y}(x)D(T)Z(T).$$

For $z \leq y$, we have $-z \geq -y$ and $z^- \geq y^-$. Consequently, we get

$$U(\xi(x))^- \leq |U(1)| + \mathcal{Y}(x)D(T)Z(T) \leq \mathcal{Y}(x)LZ(T).$$

Taking expectations and exploiting the supermartingale property of $Z(T)$, leads to

$$\mathbb{E}\left[U(\xi(x))^-\right] < \infty.$$

The last inequality is again a consequence of the property of concave functions we used before. Let now $\pi \in \mathcal{A}(x)$ be arbitrary. Then we have

$$U(\xi(x)) \geq U(X^\pi(T)) + \left[\mathcal{Y}(x)D(T)Z(T)\xi(x) - \mathcal{Y}(x)D(T)Z(T)X^\pi(T)\right].$$

Again we simply take expectations and use a result of Proposition 4.21, namely that $\mathbb{E}\left[D(t)Z(t)X^\pi(t)\right] \leq x$ holds for every $t \in [0,T]$. Combined with the first equality shown in this Lemma, we get

$$\mathbb{E}\left[U(\xi(x))\right] \geq \mathbb{E}\left[U(X^\pi(T))\right].$$

□

Remark 4.23. A direct consequence of the proof of the previous Lemma is that every $\hat{\pi} \in \mathcal{A}(x)$ satisfying $\mathbb{E}\left[D(T)Z(T)X^{\hat{\pi}}(T)\right] = x$ is optimal for our problem. This property ensures that if $\hat{\pi}$ is optimal, $D(T)Z(T)X^{\hat{\pi}}(T)$ is even a true martingale.

Lemma 4.24. *There exists a portfolio process $\hat{\pi} \in \mathcal{A}(x)$ such that*

$$X^{\hat{\pi}}(T) = \xi(x)$$

holds almost surely.

Proof. By the martingale representation theorem we have that there exists a $\{\mathcal{F}_t\}$-adapted, square integrable process $\varphi(t)$ satisfying

$$M(t) := \mathbb{E}\left[D(T)Z(T)\xi(x) \mid \mathcal{F}_t\right] = x + \int_0^t \varphi(s)^\top \,\mathrm{d}W(s).$$

Define now the processes $\hat{X}(t) := \dfrac{M(t)}{D(t)Z(t)}$ and $\mathrm{d}\tilde{W}(t) := \mathrm{d}W(t) + \theta(s)\,\mathrm{d}s$. By Girsanov's theorem we know that $\tilde{W}$ is a standard Brownian Motion with respect to our measure $\mathbb{Q}$. Due to Lemma 4.22 we have that $\hat{X}(0) = x$ and $\hat{X}(T) = \xi(x)$. Furthermore, we get through Itô's Lemma that

$$\hat{X}(t) = \frac{1}{D(t)} \left(x + \int_0^t \frac{1}{Z(s)} (\varphi(s) + M(s)\theta(s))^\top \, \mathrm{d}\tilde{W}(s) \right).$$

Therefore, we have that $\hat{X}$ is the wealth process of the portfolio process

$$\hat{\pi}(t) := \frac{1}{M(t)} \left(\sigma(t)^\top \right)^{-1} (\varphi(t) + M(t)\theta(t)).$$

We still must show that $\hat{\pi} \in \mathcal{A}(x)$. To get this we have to show that

$$\mathbb{E}\left[U(X^{\hat{\pi}}(T))^- \right] < \infty,$$

which is holds true by Lemma 4.22. □

Theorem 4.25. *The portfolio process $\hat{\pi}$ in Lemma 4.24 is a solution to our utility maximization problem and*

$$V(x) = \mathbb{E}\left[U(\xi(x)) \right].$$

Proof. By Lemma 4.24 we have that there exists an admissible portfolio process $\hat{\pi}$ with $X^{\hat{\pi}}(T) = \xi(x)$. Lemma 4.22 shows us that for every $\pi \in \mathcal{A}(x)$ the inequality

$$\mathbb{E}\left[U(X^{\pi}(T)) \right] \leq \mathbb{E}\left[U(\xi(x)) \right]$$

holds. Accordingly we have

$$\mathbb{E}\left[U(\xi(x)) \right] = \mathbb{E}\left[U(X^{\hat{\pi}}(T)) \right] \leq \sup_{\pi \in \mathcal{A}(x)} \mathbb{E}\left[U(X^{\pi}(T)) \right] \leq \mathbb{E}\left[U(\xi(x)) \right].$$

□

Example 4.1. Consider a one-dimensional Black-Scholes model, consisting of the bank account $S_0(t) = e^{rt}$, with $r \geq 0$ and a risky asset, whose price process is of the form

$$S(t) = S(0) \exp\left(\left(\mu - \frac{\sigma^2}{2} \right) t + \sigma W(t) \right),$$

with positive constants $S(0)$ and σ, and a real valued constant μ. The unique risk neutral measure is given by its density process

$$Z(t) = \exp\left(-\frac{\mu - r}{\sigma} W(t) - \frac{1}{2} \frac{(\mu - r)^2}{\sigma^2} t \right).$$

Furthermore, assume that the utility function is of the form $U(x) = \log(x)$, hence $U'(x) = \frac{1}{x}$ and $I(y) = \frac{1}{y}$.
The function $\mathcal{X}$ is given by

$$\mathcal{X}(y) = \mathbb{E}\left[D(T)Z(T)I\left(yD(T)Z(T)\right)\right] = \frac{1}{y},$$

and its inverse is $\mathcal{Y}(x) = \frac{1}{x}$. The optimizer $\xi(x)$ satisfies

$$\xi(x) = x\frac{1}{D(T)Z(T)} = x\exp\left(\frac{\mu - r}{\sigma}W(T) + \left(\frac{1}{2}\frac{(\mu - r)^2}{\sigma^2} + r\right)T\right).$$

Using the representation of Corollary 3.33, we easily see that the optimal portfolio process is

$$\hat{\pi}(t) = \frac{\mu - r}{\sigma^2}$$

and the optimal expected utility from terminal wealth is

$$V(x) = \log(x) + \left(\frac{1}{2}\frac{(\mu - r)^2}{\sigma^2} + r\right)T.$$

Remark 4.26. We see that in some markets, we get a closed form for the optimal portfolio process. In the Black-Scholes market model, with logarithmic utility function, it is even constant. Since π is a relative portfolio process, this means that the fraction of total wealth invested into the stock is constant, but the absolute value changes, as the total wealth does. This is also the case if we use power utility, defined by $U(x) = \frac{1}{\gamma}x^\gamma$, with a parameter $\gamma \in (0, 1)$, which has already been shown by Merton (1969).

4.3 Portfolio Optimization in Incomplete Markets

Remark 4.27. If we want to use the martingale method in incomplete markets, we have the problem that the risk-neutral measure is no longer unique, but there is even a convex set of them. Roughly speaking, the main idea of this generalized approach is, to formulate an optimality criterion for the change of measure and apply the results of the previous chapter. Since it could be that no equivalent martingale measure satisfies this optimality condition, we must extend the set over which we optimize slightly, to guarantee the existence of a solution.

Permanent Assumption 4.28. Additionally, to the assumptions made in Chapter 3, we assume that

1. the market is incomplete i.e. $m < d$.

2. the matrix $\sigma(t)$ has full rank for every t almost surely, i.e. $rank(\sigma(t)) = m$ for all t almost surely.

Problem 4.29. Analogously to Problem 4.12 we consider a utility function U, the set $\mathcal{A}(x)$ and define the value function of the portfolio optimization problem

$$V(x) = \sup_{\pi \in \mathcal{A}(x)} \mathbb{E}\left[U(X^{\pi}(T))\right].$$

Again we will call any process $\hat{\pi} \in \mathcal{A}(x)$ optimal or solution if it satisfies

$$V(x) = \mathbb{E}\left[U(X^{\hat{\pi}}(T))\right].$$

Lemma 4.30. *The process*

$$\theta(t) = \sigma(t)^{\top}(\sigma(t)\sigma(t)^{\top})^{-1}(b(t) - r(t)\underline{1}_m),$$

is a market price of risk.

Proof. To prove the statement, we simply have to check that this process satisfies

$$\sigma(t)\theta(t) = b(t) - r(t)\underline{1}_m$$

almost surely. Plugging in, leads to

$$\sigma(t)\theta(t) = \sigma(t)\sigma(t)^{\top}(\sigma(t)\sigma(t)^{\top})^{-1}(b(t) - r(t)\underline{1}_m) = b(t) - r(t)\underline{1}_m.$$

□

Remark 4.31. Since $\sigma(t)$ is assumed to have full rank for all t a.s., it follows that the inverse of $\sigma(t)\sigma(t)^{\top}$ is well defined for all t almost surely.

Remark 4.32. Even though we know that in the incomplete case the market price of risk is not unique, we will from now on denote the process $\theta(t)$ defined in Lemma 4.30 as **the** market price of risk.

Permanent Assumption 4.33. From now on we assume that $\int_0^T \|\theta(t)\|^2 \,\mathrm{d}t < \infty$ holds almost surely.

Definition 4.34. For $i \in \{m+1, \dots, d\}$ we define the fictitious stocks as

$$\mathrm{d}S_i(t) = S_i(t)\left(\alpha_i(t)\,\mathrm{d}t + \sum_{j=1}^{d} \rho_{ij}(t)\,\mathrm{d}W^{(j)}(t)\right), \quad S_i(0) = S_i > 0.$$

Here the volatility coefficient process $\rho(t) \in \mathbb{R}^{d-m \times d}$ is uniformly bounded, has orthonormal rows, full rank, is $\{\mathcal{F}_t\}$-progressively measurable and satisfies $\sigma(t)\rho(t)^{\top} = 0$ almost surely.
The drift term $\alpha(t) \in \mathbb{R}^{d-m}$ is $\{\mathcal{F}_t\}$-progressively measurable too and $\int_0^T \|\alpha(t)\| \,\mathrm{d}t < \infty$ holds almost surely.

Remark 4.35. Even though the volatility coefficient process $\rho(t)$ doesn't have to be unique we choose one process satisfying our needed properties and keep it fixed, whereas we use the drift as a parameter which we use to change the assets we add to our market model.

Lemma 4.36. *The market price of risk of our completed market model is of the form*

$$\theta_{com}(t) = \theta(t) + \nu(t).$$

Proof. Let $\sigma_{com}(t)$ and $b_{com}(t)$ be the coefficients of the completed market. Using the same steps as in the proof of Lemma 4.30 we can easily show that

$$\theta_{com} = \sigma_{com}(t)^\top (\sigma_{com}(t)\sigma_{com}(t)^\top)^{-1}(b_{com}(t) - r(t)\underline{1}_d)$$

holds almost surely. Defining $\nu(t) = \rho(t)^\top(\alpha(t) - r(t)\underline{1}_{d-m})$ and exploiting that $\sigma(t)\rho(t)^\top = 0$ and $\rho(t)\rho(t)^\top = \underline{1}_{d-m\times d-m}$ leads directly to the desired decomposition of $\theta_{com}(t)$. □

Remark 4.37. Every completed market model is, by its construction, complete, hence the market price of risk is unique for every completion, but since ν is orthogonal on σ, we can see that any market price of risk of a completed model, is a market price of risk for the incomplete model too.

Lemma 4.38. *The process ν is almost surely orthogonal on the market price of risk $\theta(t)$ of the incomplete market. Consequently,*

$$\|\theta_{com}(t)\|^2 = \|\theta(t) + \nu(t)\|^2 = \|\theta(t)\|^2 + \|\nu(t)\|^2$$

holds almost surely.

Proof. Using the definition of θ in Lemma 4.30 and the equality $\sigma(t)\rho(t)^\top = 0$, we have

$$\theta(t)^\top \nu(t) = (b(t) - r(t)\underline{1}_m)^\top (\sigma(t)\sigma(t)^\top)^{-1}\sigma(t)\rho(t)^\top(\alpha(t) - r(t)\underline{1}_{d-m}) = 0.$$

The equality of the norms is a direct consequence of the orthogonality. □

Permanent Assumption 4.39. We assume that $\int_0^T \|\nu(t)\|^2 \,\mathrm{d}t < \infty$ holds almost surely.

Definition 4.40. For every completion of our market we can define the positive local martingale

$$Z_\nu(t) = \exp\left(-\int_0^t (\theta(s)^\top + \nu(s)^\top)\,\mathrm{d}W(s) - \frac{1}{2}\int_0^t (\|\theta(s)\|^2 + \|\nu(s)\|^2)\,\mathrm{d}s\right).$$

Remark 4.41. Because we fixed ρ, the market completions can be parametrized by square-integrable processes ν, which are connected to the risk-neutral measures of our incomplete market. We will sometimes call a market to be ν-completed if it is completed by assets with drift α_ν.

Definition 4.42. Like in Section 4.2 we define for every completion the function

$$\mathcal{X}_\nu(y) := \mathbb{E}\left[D(T)Z_\nu(T)I\left(yD(T)Z_\nu(T)\right)\right].$$

Under the assumption of finiteness of $\mathcal{X}_\nu(y)$ we define the inverse function $\mathcal{Y}_\nu(x)$. Moreover, we set

$$\xi_\nu(x) := I\left(\mathcal{Y}_\nu(x)D(T)Z_\nu(T)\right).$$

Definition 4.43. For convenience we will denote the set of $\{\mathcal{F}_t\}$-progressively measurable, square integrable processes ν, taking values in $\mathbb{R}^d$ and satisfying $\sigma(t)\nu(t) = 0$ for all t almost surely by $\mathcal{K}(\sigma)$.
Additionally we define the set $\mathcal{K}^1(\sigma) := \left\{\nu \in \mathcal{K}(\sigma) \,|\, \mathcal{X}_\nu(y) < \infty \text{ for all } y > 0\right\}$.

Theorem 4.44. [3] *Consider the set $\mathcal{K}_m(\sigma) \subset \mathcal{K}(\sigma)$ given by*

$$\nu \in \mathcal{K}_m(\sigma) \;\Leftrightarrow\; \mathbb{E}\left[Z_\nu(T)\right] = 1.$$

A measure $\mathbb{Q}$ is an equivalent martingale measure if and only of its density is given by Z_ν for some $\nu \in \mathcal{K}_m(\sigma)$.

Remark 4.45. The set of equivalent martingale measures, represented by their density processes is a subset of the set of our dual variables $\mathcal{K}(\sigma)$.

Definition 4.46. For $\nu \in \mathcal{K}(\sigma)$ we define

$$V_\nu(x) = \sup_{\pi \in \mathcal{A}_\nu(x)} \mathbb{E}\left[U(X^\pi(T))\right],$$

where $\mathcal{A}_\nu(x)$ is defined analogously to $\mathcal{A}(x)$, but for the ν-completed market.

Remark 4.47. A portfolio process $\pi \in \mathcal{A}_\nu(x)$, is allowed to invest money into the stocks of our incomplete market, as well into the fictitious stocks of the ν-completed market. Hence, every $\pi \in \mathcal{A}(x)$ is in $\mathcal{A}_\nu(x)$ for all completions ν. A consequence of this is Corollary 4.48.

Corollary 4.48. *For every $\nu \in \mathcal{K}^1(\sigma)$ we have*

$$V(x) \le V_\nu(x) = \mathbb{E}\left[U(\xi_\nu(x))\right].$$

[3] Pham 2009, pp. 178f.

Proof. Every admissible strategy of our original market is also admissible in the completed market. Therefore, the supremum over all admissible strategies of the incomplete case is less or equal the supremum in any completion, which is why the inequality hold. The equality $V_\nu(x) = \mathbb{E}\left[U(\xi_\nu(x))\right]$ is a direct consequence of Theorem 4.25. □

Remark 4.49. The condition $\mathcal{X}_\nu(y) < \infty$ for all $y > 0$ is not necessary for the inequality to hold but is only needed to ensure finiteness of $V_\nu(x)$ and the representation according to Theorem 4.25.

Lemma 4.50. *For every $\nu \in \mathcal{K}(\sigma)$ and $\pi \in \mathcal{A}(x)$ we have that $D(t)X^\pi(t)$ is a positive local martingale with respect to the measure $\mathbb{P}_\nu$, where $\mathbb{P}_\nu$ denotes the unique equivalent martingale measure of the ν-completed market.*

Proof. The proof is exactly like the proof of Proposition 4.21. □

Permanent Assumption 4.51. From now on we will assume that the utility function U satisfies $U(0) > -\infty$.

Remark 4.52. This assumption is needed to guarantee the equivalence of the primal and the dual problem, but it excludes some well-known examples of utility functions like the logarithmic utility function, we have already seen in Example 4.1.

Theorem 4.53. [4] *For $\hat{\nu} \in \mathcal{K}^1(\sigma)$ the following statements are equivalent:*

1. *There exists a portfolio process $\hat{\pi} \in \mathcal{A}(x)$ with $X^{\hat{\pi}}(T) = \xi_{\hat{\nu}}(x)$ almost surely.*
2. $\mathbb{E}\left[U(\xi_{\hat{\nu}}(x))\right] \leq \mathbb{E}\left[U(\xi_\nu(x))\right]$ *for all* $\nu \in \mathcal{K}^1(\sigma)$.
3. $\mathbb{E}\left[\tilde{U}(\mathcal{Y}_{\hat{\nu}}(x)D(T)Z_{\hat{\nu}}(T))\right] \leq \mathbb{E}\left[\tilde{U}(\mathcal{Y}_{\hat{\nu}}(x)D(T)Z_\nu(T))\right]$ *for all* $\nu \in \mathcal{K}^1(\sigma)$.
4. $\mathbb{E}\left[D(T)Z_\nu(T)\xi_{\hat{\nu}}(x)\right] \leq x$ *for all* $\nu \in \mathcal{K}^1(\sigma)$.

Theorem 4.54. [5] *Let $\hat{\pi} \in \mathcal{A}(x)$ be a solution to our optimization problem and assume there are some constants $\alpha \in (0,1)$ and $\beta \in (1,\infty)$ such that $U'(\beta x) \leq \alpha U'(x)$ holds true for all $x > 0$. Then there exists a $\hat{\nu} \in \mathcal{K}^1(\sigma)$ satisfying $X^{\hat{\pi}}(T) = \xi_{\hat{\nu}}(x)$ almost surely.*

Remark 4.55. Theorem 4.54 shows us, some similarities between the complete and the incomplete case. By Theorem 4.25 we have $X^{\hat{\pi}}(T) = \xi(x)$ for the complete case which can be seen as minimum over all $\xi_\nu(x)$.

Theorem 4.56. *If there exist some $\hat{\nu} \in \mathcal{K}^1(\sigma)$ and $\hat{\pi} \in \mathcal{A}(x)$ which satisfy $X^{\hat{\pi}}(T) = \xi_{\hat{\nu}}(x)$ almost surely, then $\hat{\pi}$ is a solution to our optimization problem.*

[4] Karatzas, Lehoczky, Shreve, and Xu 1991, pp. 716f.
[5] Karatzas, Lehoczky, Shreve, and Xu 1991, pp. 716f.

Proof. By Corollary 4.48 we have that

$$\mathbb{E}\left[U(X^{\hat{\pi}}(T))\right] \leq V(x) \leq \mathbb{E}\left[U(\xi_{\hat{\nu}}(x))\right] = \mathbb{E}\left[U(X^{\hat{\pi}}(T))\right].$$

Consequently, $\hat{\pi}$ is optimal. □

Remark 4.57. The restriction, to $\mathcal{K}^1(\sigma)$ is necessary, to ensure we can use the results of the previous chapter, since we assumed the additional property $\mathcal{X}_\nu(y) < \infty$ in Permanent Assumption 4.18.

Remark 4.58. Theorem 4.54 and Theorem 4.56 allow us to avoid a direct optimization. Instead, we can use a minimizing problem, since we know that under the right circumstances there is a relation between the two optimal values. The benefit is that we can minimize over a set of nicely parameterized processes.

4.4 Dual Problem

Definition 4.59. For our primal problem

$$V(x) = \sup_{\pi \in \mathcal{A}(x)} \mathbb{E}\left[U(X^\pi(T))\right] \quad \forall x > 0,$$

we define the dual problem as

$$\tilde{V}(y) = \inf_{\nu \in \mathcal{K}(\sigma)} \mathbb{E}\left[\tilde{U}(yD(T)Z_\nu(T))\right] \quad \forall y > 0.$$

Assumption 4.60. Assume that $\tilde{V}(y) < \infty$ for all $y > 0$.

Remark 4.61. Since the dual problem is a minimizing problem, it would be equivalent to demand that for every $y > 0$ there exists a process $\nu \in \mathcal{K}(\sigma)$ such that

$$\mathbb{E}\left[\tilde{U}(yD(T)Z_\nu(T))\right] < \infty.$$

Lemma 4.62. *For all $x, y > 0$, $\pi \in \mathcal{A}(x)$ and $\nu \in \mathcal{K}(\sigma)$ we have*

$$U(X^\pi(T)) \leq \tilde{U}\left(yD(T)Z_\nu(T)\right) + yD(T)Z_\nu(T)X^\pi(T),$$

and consequently also the inequality

$$\mathbb{E}\left[U(X^\pi(T))\right] \leq \mathbb{E}\left[\tilde{U}(yD(T)Z_\nu(T))\right] + xy$$

holds.

Proof. The convex dual for $z > 0$ is defined as $\tilde{U}(z) = \sup_{\zeta\in\mathbb{R}} \big(U(\zeta) - z\zeta\big)$, hence for all $\zeta \geq 0$ we have the inequality

$$\tilde{U}(z) + z\zeta \geq U(\zeta).$$

Using this with $\zeta = X^\pi(T)$ and $z = yD(T)Z_\nu(T)$ leads to

$$U(X^\pi(T)) \leq \tilde{U}(yD(T)Z_\nu(T)) + yD(T)Z_\nu(T)X^\pi(T).$$

Taking expectation and exploiting the supermartingale property of $D(T)Z_\nu(T)X^\pi(T)$ gives us

$$\mathbb{E}\left[\tilde{U}(yD(T)Z_\nu(T))\right] + xy \geq \mathbb{E}\left[\tilde{U}(yD(T)Z_\nu(T))\right] + y\mathbb{E}\left[D(T)Z_\nu(T)X^\pi(T)\right]$$

and consequently

$$\mathbb{E}\left[\tilde{U}(yD(T)Z_\nu(T))\right] + xy \geq \mathbb{E}\left[U(X^\pi(T))\right].$$

□

Lemma 4.63. *Assume there exist some $x, y > 0$, $\hat{\pi} \in \mathcal{A}(x)$ and $\hat{\nu} \in \mathcal{K}(\sigma)$ which satisfy*

$$\mathbb{E}\left[U(X^{\hat{\pi}}(T))\right] = \mathbb{E}\left[\tilde{U}(yD(T)Z_{\hat{\nu}}(T))\right] + xy.$$

Then $\hat{\pi}$ is a solution to the primal problem $V(x)$ and $\hat{\nu}$ is a solution to the dual problem $\tilde{V}(y)$.

Proof. By Lemma 4.62 we know that for all $\pi \in \mathcal{A}(x)$

$$\mathbb{E}\left[U(X^{\hat{\pi}}(T))\right] = \mathbb{E}\left[\tilde{U}(yD(T)Z_{\hat{\nu}}(T))\right] + xy \geq \mathbb{E}\left[U(X^\pi(T))\right]$$

holds. Consequently, $\hat{\pi}$ must be optimal. Further we have for all $\nu \in \mathcal{K}(\sigma)$

$$\mathbb{E}\left[\tilde{U}(yD(T)Z_{\hat{\nu}}(T))\right] + xy = \mathbb{E}\left[U(X^{\hat{\pi}}(T))\right] \leq \mathbb{E}\left[\tilde{U}(yD(T)Z_\nu(T))\right] + xy.$$

This proves the optimality of $\hat{\nu}$. □

Lemma 4.64. *Let $y > 0$ and assume there exists a solution $\nu_y \in \mathcal{K}^1(\sigma)$ to the dual problem $\tilde{V}(y)$. Additionally, let Assumptions 4.60 and 4.51 hold. Then there exists a solution $\hat{\pi}$ to the primal problem $V(\mathcal{X}_{\nu_y}(y))$. Furthermore, we have the equality*

$$\tilde{V}(y) = \sup_{\eta>0}(V(\eta) - y\eta).$$

Proof. By Theorem 4.53 we know there exists a portfolio process $\hat{\pi} \in \mathcal{A}(\mathcal{X}_{\nu_y}(y))$ such that $X^{\hat{\pi}}(T) = \xi_{\nu_y}\left(\mathcal{X}_{\nu_y}(y)\right)$ almost surely. Due to Theorem 4.56, we have that $\hat{\pi}$ is a solution to the primal problem $V(\mathcal{X}_{\nu_y}(y))$. Lemma 4.7 gives us

$$\mathbb{E}\left[\tilde{U}(yD(T)Z_{\nu_y}(T))\right] = \mathbb{E}\left[U(X^{\hat{\pi}}(T))\right] - y\mathcal{X}_{\nu_y}(y),$$

hence by Lemma 4.62, we get

$$\tilde{V}(y) = \mathbb{E}\left[\tilde{U}(yD(T)Z_{\nu_y}(T))\right] = \mathbb{E}\left[U(X^{\hat{\pi}}(T))\right] - \mathcal{X}_{\nu_y}(y)y = \sup_{\eta>0}(V(\eta) - y\eta).$$

□

Remark 4.65. The function $V(x)$ is only defined for positive x. If we extend it in a similar way as we did for our utility functions i.e. $V(x) = -\infty$ for $x < 0$ and $V(0) = \lim_{x\downarrow 0} V(x)$, we can rewrite the last equation of the previous lemma as

$$\tilde{V}(y) = \sup_{\eta\in\mathbb{R}}(V(\eta) - y\eta).$$

As we can see, if all requirements of the lemma are fulfilled, we can represent $\tilde{V}$ as convex dual of V for this particular $y > 0$, hence if this holds true for all y, then $\tilde{V}$ is convex.

Assumption 4.66. Assume for every $y > 0$ there exists an process $\nu_y \in \mathcal{K}^1(\sigma)$ satisfying

$$\tilde{V}(y) = \mathbb{E}\left[\tilde{U}(yD(T)Z_{\nu_y}(T))\right].$$

Proposition 4.67. *Let U be a utility function satisfying $U(\infty) = \infty$ and assume there exist some constants $\alpha \in (0,1)$ and $\beta \in (1,\infty)$ such that*

$$U'(\beta x) \leq \alpha U'(x) \quad \forall x > 0.$$

Additionally let all conditions of Lemma 4.64 and Assumption 4.66 hold. Then for every $x > 0$ there exists a $y > 0$ such that

$$\inf_{\eta}(\tilde{V}(\eta) + x\eta) = \tilde{V}(y) + xy$$

This y satisfies the equality

$$x = \mathcal{X}_{\nu_y}(y) = \mathbb{E}\left[D(T)Z_{\nu_y}(T)I\left(yD(T)Z_{\nu_y}(T)\right)\right],$$

where ν_y denotes the optimal process of Assumption 4.66.

Proof. By Jensen's Inequality and the monotonicity of $\tilde{U}$ we get that

$$\mathbb{E}\left[\tilde{U}(yD(T)Z_\nu(T))\right] \geq \tilde{U}\left(ye^L\mathbb{E}\left[Z_{\nu_y}(T)\right]\right).$$

Using the fact that $Z_\nu(t)$ is a supermartingale leads to

$$\tilde{V}(y) \geq \tilde{U}(ye^L),$$

hence $\tilde{V}(0) = \infty$. For $x > 0$ define the convex function $g_x(y) = \tilde{V}(y) + xy$ which satisfies $g_x(0) = g_x(\infty) = \infty$. Therefore, g_x attains a minimum at some $y(x) \in (0, \infty)$. Defining $G_y(\eta) := \mathbb{E}\left[\tilde{U}(\eta y D(T) Z_{\nu_y}(T))\right]$ and $D_x(\eta) := \eta x y(x) + G_{y(x)}(\eta)$ for $\eta > 0$, one can derive that

$$D'_x(\eta) = xy(x) - y(x)\mathcal{X}_{\nu_{y(x)}}\left(y(x)\right) = 0.$$

Thus the equality $\mathcal{X}_{\nu_{y(x)}}\left(y(x)\right) = x$ holds true[6]. □

Theorem 4.68. *Suppose that all conditions of Proposition 4.67 hold. Then for every $x > 0$, there exists a solution $\hat{\pi} \in \mathcal{A}(x)$ to our primal optimization problem.*

Proof. Proposition 4.67 gives us that we can apply Lemma 4.64 to every $x > 0$. Therefore, we get an optimizer for our primal optimization problem $V(x)$. □

Remark 4.69. Theorem 4.68 is the main result of this subsection but needs Assumption 4.66 to hold. Now we will discuss some circumstances under which this is always fulfilled.

4.5 Existence of a Dual Solution

Permanent Assumption 4.70. Assume that the market price of risk satisfies

$$\int_0^T \|\theta(t)\|^2 \, \mathrm{d}t < \infty$$

almost surely.

Permanent Assumption 4.71. The mapping $x \to xU'(x)$ is monotone increasing.

Remark 4.72. We have already used Permanent Assumption 4.70 in previous chapters, so it is no further restriction, but Permanent Assumption 4.71 considerably constrains the choice of our utility functions.

[6]Karatzas, Lehoczky, Shreve, and Xu 1991, pp. 725f.

Remark 4.73. If U is twice continuously differentiable, Assumption 4.71 is equivalent to

$$-\frac{xU''(x)}{U'(x)} \leq 1.$$

This can be derived by differentiating the function $xU'(x)$ and using that its derivative has to be non-negative. The term $-\frac{xU''(x)}{U'(x)}$ is called coefficient of relative risk aversion.

Lemma 4.74. *It holds that* $U(\infty) = \infty$*.*

Proof. The function $xU'(x)$ is monotone increasing so we have for $x \geq 1$ that $U'(x) \geq \frac{U'(1)}{x}$. Integrating both sides from 1 to $y > 1$ leads us to

$$U(y) - U(1) \geq U'(1) \log y.$$

The limit $y \to \infty$ gives us $U(\infty) = \infty$. □

Lemma 4.75. *If Assumption 4.51 holds, the mapping* $x \to \tilde{U}(e^x)$ *is convex,* $\tilde{U}(0) = \infty$ *and* $\tilde{U}(\infty) > -\infty$*.*

Proof. To show the equalities we use the fact that

$$\tilde{U}(y) = \sup_{\eta > 0}(U(\eta) - \eta y) \geq U(x) - xy \quad \forall x > 0.$$

If we let y tend to 0, we get $\tilde{U}(0) \geq U(x)$ for all $x > 0$. Now, for $x \to \infty$ we see

$$\infty = U(\infty) \leq \tilde{U}(0).$$

If we let x go to 0, we get

$$\tilde{U}(y) \geq U(0) > -\infty,$$

hence by continuity

$$\tilde{U}(\infty) \geq U(0) > -\infty.$$

The function is continuously differentiable, since $\tilde{U}(y)$ and e^x are, and has derivative

$$\frac{\partial}{\partial x}\tilde{U}(e^x) = -I(e^x)e^x.$$

The function I is monotone decreasing, $-I(e^x)e^x$ is increasing, hence $\tilde{U}(e^x)$ convex. □

Lemma 4.76. *The set* $\mathcal{K}(\sigma)$ *is a closed, convex subset of the Hilbert space*

$$\mathcal{H} := \left\{ \nu : \Omega \times [0,T] \to \mathbb{R}^d \;\middle|\; \int_0^T \|\nu(t)\|^2 \,\mathrm{d}t < \infty \ a.s. \right\},$$

with inner product $\langle \nu, \mu \rangle := \mathbb{E}\left[\int_0^T \nu(t)^\top \mu(t) \,\mathrm{d}t\right]$.

Proof. Let $\nu, \mu \in \mathcal{K}(\sigma)$ and $\lambda \in (0,1)$ be arbitrary but fixed. The set

$$\big\{\omega \in \Omega \,|\, \exists t \in [0,T] \; \textit{such that}\; \sigma(\omega,t)\nu(\omega,t) \neq 0 \; \textit{or}\; \sigma(\omega,t)\mu(\omega,t) \neq 0\big\}$$

is a union of two zero sets and consequently again a zero set. For all ω which are not in this set, we have for all $t \in [0,T]$

$$\sigma(\omega,t)\left(\lambda\nu(\omega,t) + (1-\lambda)\mu(\omega,t)\right) = \lambda\sigma(\omega,t)\nu(\omega,t) + (1-\lambda)\sigma(\omega,t)\mu(\omega,t) = 0,$$

hence $\mathcal{K}(\sigma)$ is convex.
Let now $(\nu_n)_{n\geq 0}$ be a convergent sequence with $\nu_n \in \mathcal{K}(\sigma)$ for all n and $\lim_{n\to\infty} \nu_n = \nu \in \mathcal{H}$. Define $\mathcal{N} := \big\{\omega \in \Omega \,|\, \sigma(\omega,t)\nu_n(\omega,t) \neq 0 \; \textit{for some}\; n \geq 0 \; \textit{and}\; t \in [0,T]\big\}$. This set is a countable union of zero sets and therefore, again a zero set. For all $\omega \notin \mathcal{N}$ we have

$$\sigma(\omega,t)\nu(\omega,t) = \sigma(\omega,t) \lim_{n\to\infty} \nu_n(\omega,t) = \lim_{n\to\infty} \sigma(\omega,t)\nu_n(\omega,t) = 0.$$

We can interchange the limit and $\sigma(\omega,t)$, due to the continuity of matrix multiplication. □

Remark 4.77. $\mathcal{K}(\sigma)$ is even a closed subspace and consequently a Hilbert space itself, but for the existence of a minimum, it is sufficient to show it is a closed and convex subset.

Definition 4.78. For $y > 0$ we define a functional from $\mathcal{H}$ to $\mathbb{R} \cup \{\infty\}$ by

$$\tilde{J}_y(\nu) := \mathbb{E}\left[\tilde{U}\left(yD(T)e^{-\chi_\nu(T)}\right)\right] \quad \forall \nu \in \mathcal{K}(\sigma),$$

where

$$\chi_\nu(s) = \int_0^s (\theta(u) + \nu(u))^\top \,\mathrm{d}W(u) + \frac{1}{2}\int_0^s \left(\|\theta(u)\|^2 + \|\nu(u)\|^2\right) \mathrm{d}u,$$

and

$$\tilde{J}_y(\nu) = \infty \quad \forall \nu \notin \mathcal{K}(\sigma).$$

Lemma 4.79. *Let Assumption 4.60 hold. The functional $\tilde{J}_y$ is convex, lower semicontinuous and proper. Moreover we have that*

$$\lim_{\|\nu\|\to\infty} \tilde{J}_y(\nu) = \infty.$$

Here $\|\nu\|$ denotes the induced norm of the Hilbert space $\mathcal{H}$.

Proof. The lower semicontinuity is a direct consequence of Fatou's lemma in combination of the continuity of $\tilde{U}$ and the exponential function. Assumption 4.60 gives us the existence of a ν such that $\tilde{J}_y(\nu) < \infty$. The function $\tilde{U}$ is monotone decreasing and by Lemma 4.75 $\tilde{U}(y) > -\infty$ for all $y \geq 0$, hence the function $\tilde{J}_y$ is proper. The convexity is also due to Lemma 4.75.
By Jensen's Inequality we get

$$\tilde{J}_y(\nu) \geq \tilde{U}\left(\exp\left(\mathbb{E}\left[\log(y) + L - \chi_\nu(T)\right]\right)\right).$$

We know that $\int_0^t (\theta(u) + \nu(u))^\top \,\mathrm{d}W(u)$ is a martingale with expectation 0. This leads to

$$\tilde{J}_y(\nu) \geq \tilde{U}\left(\exp\left(\log(y) + L - \frac{1}{2}(\|\theta\|^2 + \|\nu\|^2)\right)\right),$$

which tends to infinity for $\|\nu\| \to \infty$ by Lemma 4.75. □

Theorem 4.80. *Let Assumptions 4.51 and 4.60 hold. Additionally assume there are some constants $\alpha \in (0,1)$ and $\beta \in (1,\infty)$, such that $\alpha U'(x) \geq U'(\beta x)$ holds for every $x > 0$. Then for every $y > 0$ there exists a solution $\hat{\nu}_y \in \mathcal{K}^1(\sigma)$ to our dual problem.*

Proof. The functional $\tilde{J}_y(\nu)$ is proper, convex, lower semicontinuous and $\mathcal{K}(\sigma)$ is a convex and closed subset of a Hilbertspace and therefore, of a reflexive Banach space. Moreover, the functional tends to infinity as $\|\nu\|$ does. Therefore, it attains at least one solution $\hat{\nu} \in \mathcal{K}(\sigma)$[7]. By monotonicity of $\tilde{U}$ we get

$$\tilde{U}(z) - \tilde{U}(\infty) \geq \tilde{U}(z) - \tilde{U}\left(\frac{z}{\alpha}\right).$$

Due to its convexity it follows that

$$\tilde{U}(z) - \tilde{U}\left(\frac{z}{\alpha}\right) \geq I\left(\frac{z}{\alpha}\right)\left(\frac{z}{\alpha} - z\right) = I\left(\frac{z}{\alpha}\right)\left(\frac{z(1-\alpha)\beta}{\alpha\beta}\right) \geq I(z)\frac{z(1-\alpha)}{\alpha\beta}.$$

Combining these inequalities, choosing $z = yD(T)Z_{\hat{\nu}}(T)$ and taking expectations leads to

$$\mathcal{X}_{\hat{\nu}}(y) \leq \frac{\alpha\beta}{y(1-\alpha)}\mathbb{E}\left[\tilde{U}(yD(T)Z_{\hat{\nu}}(T)) - \tilde{U}(\infty)\right] < \infty,$$

and therefore, $\hat{\nu} \in \mathcal{K}^1(\sigma)$. □

Remark 4.81. We have now developed the theoretical framework, we need to apply the martingale method in incomplete markets, but we still must check if Heston's SV market model satisfies all assumptions needed. Since this technique is highly restrictive, it is also important to choose the right utility function.

[7] Ekeland and Temam 1999, p. 35.

5 Portfolio Optimization in Hestons Market Model

Now we will apply the martingale method to Heston's market model with power utility functions. The portfolio optimization problem itself has already been solved through different approaches, for example by Kallsen and Muhle-Karbe (2010), who used semimartingale characteristics to obtain a solution to the optimization problem. Our notation of the optimization problem, and of some variables, is based on Kraft (2003), who enhanced the findings of Zariphopoulou (2001) to achieve a representation result, which allows identification of an optimal portfolio process in Heston's market model, derived by stochastic dynamic programming.

5.1 Power Utility

Definition 5.1. For $\gamma \in (0,1)$ we define the power utility function by

$$U(x) = \frac{1}{\gamma} x^\gamma \quad \forall\, x > 0.$$

Lemma 5.2. *For every $\gamma \in (0,1)$ we have that the corresponding power utility function U is a utility function.*

Proof. The function U is continuously differentiable, and even in C^∞. The first two derivatives are given by

$$U'(x) = x^{\gamma-1},$$
$$U''(x) = (\gamma - 1)x^{\gamma-2}.$$

Since U' is strictly positive, U is strictly monotone increasing and the negativity of U'' shows that U is strictly concave. The Inada conditions are fulfilled, because $\gamma - 1 < 0$. □

Lemma 5.3. *Let U be the power utility function with parameter $\gamma \in (0,1)$. Then, the functions U', I and $\tilde{U}$ are given by*

$$U'(x) = x^{\gamma-1},$$
$$I(y) = y^{\frac{1}{\gamma-1}},$$
$$\tilde{U}(y) = \frac{1-\gamma}{\gamma} y^{\frac{\gamma}{\gamma-1}}.$$

Proof. The functions U' and I are derived by differentiating U and inverting respectively. To determine $\tilde{U}$ we use Lemma 4.7 and conclude

$$\tilde{U}(y) = U(I(y)) - yI(y) = \frac{1}{\gamma} y^{\frac{\gamma}{\gamma-1}} - y y^{\frac{1}{\gamma-1}} = \frac{1-\gamma}{\gamma} y^{\frac{\gamma}{\gamma-1}}.$$

□

Lemma 5.4. *The coefficient of relative risk aversion is constant* $1 - \gamma < 1$

Proof. The second derivative is given by $U''(x) = (\gamma - 1)x^{\gamma-2}$, hence

$$-\frac{xU''(x)}{U'(x)} = -\frac{(\gamma-1)x^{\gamma-1}}{x^{\gamma-1}} = 1 - \gamma.$$

□

Lemma 5.5. *Let* $\gamma \in (0,1)$ *and* U *be the power utility function with parameter* γ. *For all* $\alpha \in (0,1)$ *there exists a* $\beta \in (1,\infty)$ *such that*

$$\alpha U'(x) \geq U'(\beta x).$$

Proof. Let $\alpha \in (0,1)$ be arbitrary and set $\beta = \alpha^{\frac{1}{\gamma-1}}$. The constant β is in $(1,\infty)$, because $0 < \gamma < 1$. With this choice we get

$$U'(\beta x) = \alpha x^{\gamma-1} = \alpha U'(x).$$

□

Lemma 5.6. *Let* U *be a power utility function with parameter* $\gamma \in (0,1)$. *Then the mapping* $x \to xU'(x)$ *is monotone increasing.*

Proof. The function U is twice continuously differentiable hence $xU'(x)$ is continuously differentiable with derivative

$$U'(x) + xU''(x) = x^{\gamma-1} + (\gamma-1)x^{\gamma-1} = \gamma.$$

Since $\gamma \in (0,1)$ it is positive hence $xU'(x)$ is even strictly monotone increasing. □

Remark 5.7. An alternative way to prove this property, is to use Lemma 5.4 and Remark 4.73 which give us the same result.

Remark 5.8. Lemma 5.5 and Lemma 5.6 are important, because they show that power utility satisfies all assumptions we have made for a utility function, to guarantee the existence of a unique solution in our space $\mathcal{K}^1(\sigma)$.

Permanent Assumption 5.9. For the rest of this chapter, we will assume our utility function to be of power utility type.

5.2 Heston's Stochastic Volatility Model

Definition 5.10. We consider Heston's Market model. It consists of a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, one risk free asset $S_0(t)$ given by

$$\mathrm{d}S_0(t) = rS_0(t)\mathrm{d}t, \quad S_0(0) = 1,$$

for some constant $r \geq 0$ and a risky asset is given by the following SDE

$$\mathrm{d}S(t) = S(t)\left((\bar{\lambda}z(t) + r)\mathrm{d}t + \sqrt{z(t)}\,\mathrm{d}\hat{W}^1(t)\right), \quad S(0) = S > 0,$$

where $\bar{\lambda} \in \mathbb{R}$. The process $z(t)$ is a Cox-Ingersoll-Ross process

$$\mathrm{d}z(t) = \kappa(\theta - z(t))\mathrm{d}t + \varsigma\sqrt{z(t)}\,\mathrm{d}\hat{W}^2(t), \quad z(0) = z_0 > 0,$$

where the constants κ, θ and ς are assumed to be positive. The processes $\hat{W}^1(t)$ and $\hat{W}^2(t)$ are Brownian motions under the measure $\mathbb{P}$ with constant correlation $\rho \in (-1, 1)$.

Permanent Assumption 5.11. For the rest of the chapter we will assume the Feller conditions

$$2\kappa\theta > \varsigma^2$$

to hold. Additionally we assume

$$\frac{\kappa^2}{2\varsigma^2} > \frac{\gamma}{1-\gamma}\bar{\lambda}\left(\frac{\kappa\rho}{\varsigma} + \frac{\bar{\lambda}}{2}\right).$$

Remark 5.12. The Feller conditions ensure the existence of a strictly positive, unique solution to the *SDE* of the CIR process, whereas the second inequality ensures the existence of the moment generating functions of some affine processes, which will appear later in this chapter.

Proposition 5.13. *The CIR process is affine on the canonical state space* $\mathbb{R}_{\geq 0}$ *and, defining* $\tau(t) := \frac{\varsigma^2}{4\kappa\theta}(1 - e^{-t\kappa})$*, the solution of the corresponding Riccati equations is given by*

$$\psi(t, z) = \frac{ze^{-\kappa t}}{1 - 2z\tau(t)},$$

$$\phi(t, z) = -\frac{2\kappa\theta}{\varsigma^2}\log(1 - 2z\tau(t)),$$

where log *denotes the main branch of the logarithm.*

Proof. If we check the corresponding functions $b(z)$ and $A(z)$ are affine linear which is the case since

$$b(z) = \kappa\theta - \kappa z,$$
$$A(z) = \varsigma^2 z.$$

we get $b = \kappa\theta$, $b_1 = -\kappa$ and $A_1 = \varsigma^2$. These coefficients satisfy the admissibility conditions, and consequently the characteristic function is given by the solution of the one-dimensional Riccati equations

$$\frac{\partial}{\partial t}\psi(t,z) = \frac{\varsigma^2}{2}\psi(t,z)^2 - \kappa\psi(t,z), \quad \psi(0,z) = z$$
$$\frac{\partial}{\partial t}\phi(t,z) = \kappa\theta\psi(t,z), \quad \phi(0,z) = 0.$$

Theorem 2.58 gives us

$$\psi(t,z) = \frac{ze^{-\kappa t}}{1 - 2z\tau(t)},$$

and by $\gamma(t) = 1 - 2z\tau(t)$ we get

$$\phi(t,z) = \kappa\theta \int_0^t \frac{\gamma'(u)}{\gamma(u)}\,\mathrm{d}u = -\frac{2\kappa\theta}{\varsigma^2}\log(1 - 2z\tau(t)).$$

The last equation holds only if $\gamma(t)$ does not leave $\mathbb{C} \setminus \mathbb{R}_-$, which is true for all $z \in i\mathbb{R}$. □

Corollary 5.14. *The random variable*

$$Z = \frac{z(t)}{\tau(t)}$$

is noncentral Chi-squared distributed with $\frac{4\kappa\theta}{\varsigma^2}$ *degrees of freedom and non-centrality parameter* $\frac{z_0 e^{-\kappa t}}{\tau(t)}$*, where* $\tau(t)$ *is given as in Proposition 5.13.*

Proof. Let $\varphi_Y(x)$ denote the characteristic function of a random variable Y in the point x. Since the process $z(t)$ is affine we can calculate its characteristic function through the Riccati equations like we did in Proposition 5.13. The characteristic function of $Z = \frac{z(t)}{\tau(t)}$ is given by

$$\varphi_Z(x) = \mathbb{E}\left[\exp\left(ixZ\right)\right] = \mathbb{E}\left[\exp\left(i\frac{x}{\tau(t)}z(t)\right)\right] = \varphi_{z(t)}\left(\frac{x}{\tau(t)}\right)$$
$$= \exp\left(\phi(t, x\tau(t)^{-1}) + \psi(t, x\tau(t)^{-1})z_0\right).$$

If we use our functions ϕ and ψ we calculated in Proposition 5.13 we get

$$\varphi_Z(x) = \exp\left(\frac{ixe^{-\kappa t}}{\tau(t)(1-2ix)} z_0\right)(1-2ix)^{-\frac{2\kappa\theta}{\varsigma^2}}.$$

This is the characteristic function of a noncentral Chi-squared distributed random variable with $\frac{4\kappa\theta}{\varsigma^2}$ degrees of freedom and non-centrality parameter $\frac{z_0 e^{-\kappa t}}{\tau(t)}$.

□

Remark 5.15. Most of the time we will consider the same model, but with respect to a two-dimensional standard Brownian motion $W = \begin{pmatrix} W^1 \\ W^2 \end{pmatrix}$ instead of the correlated ones. If we do this, we must change the SDEs which describe our processes into

$$\mathrm{d}S(t) = S(t)\left(\left(\bar{\lambda} z(t) + r\right)\mathrm{d}t + \sqrt{z(t)}\left(\sqrt{1-\rho^2}\,\mathrm{d}W^1(t) + \rho\,\mathrm{d}W^2(t)\right)\right), \quad S(0) = S > 0,$$

$$\mathrm{d}z(t) = \kappa(\theta - z(t))\mathrm{d}t + \varsigma\sqrt{z(t)}\,\mathrm{d}W^2(t), \quad z(0) = z_0 > 0.$$

Lemma 5.16. *The process* $\bar{\lambda} z(t) + r$ *satisfies*

$$\int_0^T |\bar{\lambda} z(t) + r|\,\mathrm{d}t < \infty \ a.s.$$

Proof. We will not directly prove this property but show that

$$\mathbb{E}\left[\int_0^T |\bar{\lambda} z(t) + r|\,\mathrm{d}t\right] < \infty.$$

This implies almost sure finiteness by Markov's inequality. To get this result we use Tonelli's theorem which tells us that, if we integrate a positive function, we may interchange two integrals, hence here we may interchange expectation and integration. Applying this leads to

$$\mathbb{E}\left[\int_0^T |\bar{\lambda} z(t) + r|\,\mathrm{d}t\right] \leq rT + |\bar{\lambda}| \int_0^T \mathbb{E}[z(t)]\,\mathrm{d}t.$$

Due to Corollary 5.14 we know that $\mathbb{E}[z(t)] = \frac{4\kappa\theta}{\varsigma^2}\tau(t) + z_0 e^{-\kappa t}$. Therefore,

$$\int_0^T \mathbb{E}[z(t)]\,\mathrm{d}t = \theta T + (z_0 - \theta)\int_0^T e^{-\kappa t}\,\mathrm{d}t = \theta T + (z_0 - \theta)\frac{1 - e^{-\kappa T}}{\kappa} < \infty$$

holds true. □

Lemma 5.17. *The diffusion matrix* $\sigma(t) = \sqrt{z(t)}\begin{pmatrix}\sqrt{1-\rho^2} & \rho\end{pmatrix}$ *has full rank for all* t *almost surely.*

Proof. Since the Feller Conditions are assumed to hold, the CIR process z stays positive almost surely. Therefore, also $\sqrt{z(t)} > 0$ for all t almost surely. Hence $\sigma(t)$ has rank 1 for all t almost surely. □

Proposition 5.18. *The market price of risk is of the form*

$$\theta(t) = \bar{\lambda}\sqrt{z(t)}\begin{pmatrix}\sqrt{1-\rho^2} \\ \rho\end{pmatrix}$$

and satisfies

$$\int_0^T \|\theta(t)\|^2 \, \mathrm{d}t < \infty \text{ a.s.}$$

Proof. To get the representation we use Lemma 4.30. Since $\sigma(t)\sigma(t)^\top = z(t)$ we have

$$\theta(t) = \sigma(t)^\top(\sigma(t)\sigma(t)^\top)^{-1}[b(t) - r(t))] = \sigma(t)^\top\bar{\lambda}.$$

For the integral we get

$$\int_0^T \|\theta(t)\|^2 \, \mathrm{d}t = \bar{\lambda}^2 \int_0^T z(t) \, \mathrm{d}t,$$

which is almost surely finite by the proof of Lemma 5.16. □

Proposition 5.19. [1] *Let* $\lambda \in \left[-\frac{\kappa}{\varsigma}, \infty\right)$, *then*

$$\mathbb{E}\left[\exp\left(-\lambda\int_0^T \sqrt{z(t)}\,\mathrm{d}W^2(t) - \frac{\lambda^2}{2}\int_0^T z(t)\,\mathrm{d}t\right)\right] = 1.$$

Remark 5.20. This property is not strong enough to ensure that for all parametrizations of the model the process

$$Z(t) = \exp\left(-\int_0^t \theta(s)^\top \, \mathrm{d}W(s) - \frac{1}{2}\int_0^t \|\theta(s)\|^2 \, \mathrm{d}s\right)$$

is necessarily a martingale, for this specific choice of the market price of risk $\theta(t)$. Nevertheless, the model is always arbitrage free and incomplete.

Lemma 5.21. *Heston's market model is arbitrage free.*

[1] Wong and Heyde 2006, p. 8.

Proof. To see this, we define the equivalent measure $\tilde{\mathbb{P}}$ by its density

$$\tilde{Z}(t) = \exp\left(- \int_0^t \frac{\bar{\lambda}}{\sqrt{1-\rho^2}} \sqrt{z(s)}\, \mathrm{d}W^1(s) - \frac{1}{2} \int_0^t \frac{\bar{\lambda}^2}{1-\rho^2} z(s)\, \mathrm{d}s \right).$$

This process is a true martingale by Theorem 2.38 which implies that the equivalent measure exists. Hence, the process

$$B(t) = W(t) + \int_0^t \frac{\bar{\lambda}}{\sqrt{1-\rho^2}} \sqrt{z(t)} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \mathrm{d}s$$

is a standard Brownian motion under $\tilde{\mathbb{P}}$.

$$D(t)S(t) \left(\bar{\lambda} z(t) \mathrm{d}t + \sqrt{z(t)} \begin{pmatrix} \sqrt{1-\rho^2} & \rho \end{pmatrix} \mathrm{d}W(t) \right) =$$

$$D(t)S(t)\sqrt{z(t)} \begin{pmatrix} \sqrt{1-\rho^2} & \rho \end{pmatrix} \left(\mathrm{d}W(t) + \frac{\bar{\lambda}}{\sqrt{1-\rho^2}} \sqrt{z(t)} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \mathrm{d}t \right) =$$

$$D(t)S(t)\sqrt{z(t)} \begin{pmatrix} \sqrt{1-\rho^2} & \rho \end{pmatrix} \mathrm{d}B(t).$$

Therefore, we have that $D(t)S(t)$ is a local martingale with respect to the equivalent measure. Hence, the model is arbitrage free. □

Lemma 5.22. *Heston's model is incomplete.*

Proof. To prove this, we just must find a second martingale measure $\mathbb{P}_2$. Let ε be such that $\bar{\lambda}\varepsilon \in \left[\frac{\kappa}{2\sigma^2}, \infty\right)$ and define

$$\tilde{\nu}(t) := \frac{\varepsilon - \rho}{\sqrt{1-\rho^2}} \bar{\lambda} \sqrt{z(t)} \begin{pmatrix} -\rho \\ \sqrt{1-\rho^2} \end{pmatrix}.$$

Then

$$Z_2(t) = \exp\left(\int_0^t (\theta(s) + \tilde{\nu}(s))^\top \mathrm{d}W(s) - \frac{1}{2} \int_0^t \|\theta(s) + \tilde{\nu}(s)\|^2 \mathrm{d}s \right)$$

is a true martingale, due to Corollary 2.39 and Proposition 5.19, hence there exists an equivalent measure $\mathbb{P}_2$ such that

$$\tilde{W}(t) = W(t) + \bar{\lambda} \int_0^t \sqrt{z(s)} \begin{pmatrix} \frac{1-\rho\varepsilon}{\sqrt{1-\rho^2}} \\ \varepsilon \end{pmatrix} \mathrm{d}s$$

is a standard Brownian motion. By easy computations we get that the discounted stock price satisfies

$$\mathrm{d}\left(D(t)S(t)\right) = D(t)S(t)\sigma(t)\, \mathrm{d}\tilde{W}(t)$$

which implies that it is a local martingale under this measure. Therefore, the martingale measure is not unique, and the market cannot be complete. □

Remark 5.23. To prove the incompleteness of this market model, we can also use the fact that the only risky asset depends on a two-dimensional Brownian motion. By Theorem 3.25, Heston's SV model is incomplete.

5.3 Portfolio Optimization in Heston's SV Model

Remark 5.24. In the previous section, we have introduced Heston's stochastic volatility model and proved some of its main properties. Since it is our aim, to calculate the optimal expected utility from terminal wealth in this model, using the main results of Chapter 4, we must check if we can apply them.

Proposition 5.25. *The set $\mathcal{K}(\sigma)$ is of the form*

$$\mathcal{K}(\sigma) = \left\{ \nu(t) = \tilde{\nu}(t) \begin{pmatrix} -\rho \\ \sqrt{1-\rho^2} \end{pmatrix} \middle| \, \tilde{\nu} : \Omega \times [0,T] \to \mathbb{R}, \int_0^T |\tilde{\nu}(t)|^2 \, \mathrm{d}t < \infty \; a.s. \right\}.$$

Proof. Let $\nu \in \mathcal{K}(\sigma)$ and $\nu(t) = \begin{pmatrix} \nu_1(t) \\ \nu_2(t) \end{pmatrix}$. The process must satisfy $\sigma(t)\nu(t) = 0$ for all t almost surely. If we look at the product we get

$$\sigma(t)\nu(t) = \sqrt{z(t)} \left(\sqrt{1-\rho^2}\nu_1(t) + \rho\nu_2(t) \right) \stackrel{!}{=} 0.$$

Since $z(t) > 0$ almost surely we can divide by $\sqrt{z(t)}$ and get $\rho\nu_2(t) = -\sqrt{1-\rho^2}\nu_1(t)$ for all t almost surely. By defining $\tilde{\nu}(t) := \nu_2(t)$ we get the proposed form. The square integrability follows from the definition of $\mathcal{K}(\sigma)$. □

Lemma 5.26. *The functions $\mathcal{X}_\nu$ and $\mathcal{Y}_\nu$ of Definition 4.42 have closed forms. They are given by*

$$\mathcal{X}_\nu(y) = y^{\frac{1}{\gamma-1}} D(T)^{\frac{\gamma}{\gamma-1}} \mathbb{E}\left[Z_\nu(T)^{\frac{\gamma}{\gamma-1}} \right],$$

$$\mathcal{Y}_\nu(x) = x^{\gamma-1} D(T)^{-\gamma} \mathbb{E}\left[Z_\nu(T)^{\frac{\gamma}{\gamma-1}} \right]^{1-\gamma}.$$

Proof. To prove this, we will use the fact that in the case of power utility we have that $I(xy) = I(x)I(y)$. Plugging in leads to

$$\mathcal{X}_\nu(y) = \mathbb{E}\left[D(T)Z_\nu(T) \left(yD(T)Z_\nu(T) \right)^{\frac{1}{\gamma-1}} \right] = y^{\frac{1}{\gamma-1}} D(T)^{\frac{\gamma}{\gamma-1}} \mathbb{E}\left[Z_\nu(T)^{\frac{\gamma}{\gamma-1}} \right].$$

Since this is a simple function in y it is easy to calculate its inverse function

$$\mathcal{Y}_\nu(x) = x^{\gamma-1} D(T)^{-\gamma} \mathbb{E}\left[Z_\nu(T)^{\frac{\gamma}{\gamma-1}} \right]^{1-\gamma}.$$

□

Remark 5.27. This special form is a direct consequence of the choice of our utility function U and the deterministic bank account. A stochastic interest rate would give us

$$\mathcal{X}_\nu(y) = y^{\frac{1}{\gamma-1}}\mathbb{E}\left[D(T)^{\frac{\gamma}{\gamma-1}} Z_\nu(T)^{\frac{\gamma}{\gamma-1}}\right].$$

Even in this case, $\mathcal{Y}_\nu$ has a closed form.

Theorem 5.28. *If there exists some optimal process $\hat{\nu} \in \mathcal{K}^1(\sigma)$ to the dual problem $\tilde{V}(y)$, then it is independent of y. Moreover*

$$\tilde{V}(y) = y^{\frac{\gamma}{\gamma-1}}\tilde{V}(1).$$

Proof. Let $y > 0$ be arbitrary but fixed and assume there exists a $\hat{\nu}$ such that

$$\mathbb{E}\left[\tilde{U}(yD(T)Z_{\hat{\nu}}(T))\right] \le \mathbb{E}\left[\tilde{U}(yD(T)Z_\nu(T))\right] \quad \forall \nu \in \mathcal{K}^1(\sigma).$$

By Lemma 5.3 we know that $\tilde{U}(y) = \frac{1-\gamma}{\gamma} y^{\frac{\gamma}{\gamma-1}}$. Therefore,

$$\mathbb{E}\left[\tilde{U}(yD(T)Z_\nu(T))\right] = y^{\frac{\gamma}{\gamma-1}}\mathbb{E}\left[\tilde{U}(D(T)Z_\nu(T))\right].$$

Combined with the optimality of $\hat{\nu}$ we get

$$\mathbb{E}\left[\tilde{U}(D(T)Z_{\hat{\nu}}(T))\right] \le \mathbb{E}\left[\tilde{U}(D(T)Z_{\hat{\nu}}(T))\right].$$

This means that $\hat{\nu}$ is also optimal for $\tilde{V}(1)$. On the other hand, is every solution to $\tilde{V}(1)$ a solution of $\tilde{V}(y)$. This can easily be verified through multiplication with the positive factor $y^{\frac{\gamma}{\gamma-1}}$. □

Remark 5.29. This feature is a direct consequence of the choice of U too, and does not depend on the underlying market model.

Lemma 5.30. *The process*

$$Z_0(t) := \exp\left(-\int_0^t \theta(s)^\top \,\mathrm{d}W(s) - \frac{1}{2}\int_0^t \|\theta(s)\|^2 \,\mathrm{d}s\right),$$

where $\theta(t) := \bar{\lambda}\sqrt{z(t)}\begin{pmatrix}\sqrt{1-\rho^2} \\ \rho\end{pmatrix}$ *denotes the market price of risk, satisfies the equation*

$$\mathbb{E}\left[\tilde{U}\left(Z_0(T)\right)\right] = \mathbb{E}\left[Z_0(T)^{\frac{\gamma}{\gamma-1}}\right] = e^{\phi(T)+\psi(T)z_0} < \infty.$$

Here the functions ψ and ϕ are given by

$$\psi(t) = -2c\frac{1-e^{-\Delta t}}{(b-\Delta)-(b+\Delta)e^{-\Delta t}},$$

$$\phi(t) = \kappa\theta \int_0^t \psi(s)\,\mathrm{d}s$$

with constants $a = \frac{\varsigma^2}{2}$, $b = -\kappa + \frac{\gamma}{1-\gamma}\bar{\lambda}\rho\varsigma$, $c = \frac{\gamma\bar{\lambda}^2}{2(1-\gamma)^2}$ *and* $\Delta = \sqrt{b^2 - 4ac}$.

Proof. The random variable $Z_0(T)^{\frac{\gamma}{\gamma-1}}$ can be rewritten as

$$Z_0(T)^{\frac{\gamma}{\gamma-1}} = \exp\big(\chi(T)\big)\,,$$

where

$$\chi(t) := \frac{\gamma}{1-\gamma}\left(\int_0^t \bar{\lambda}\sqrt{z(s)}\left(\sqrt{1-\rho^2} \quad \rho\right)\mathrm{d}W(s) + \frac{\bar{\lambda}^2}{2}\int_0^t z(s)\,\mathrm{d}s\right).$$

The process $X(t) = \begin{pmatrix} z(t) \\ \chi(t) \end{pmatrix}$ is affine with constants $b = \begin{pmatrix} \kappa\theta \\ 0 \end{pmatrix}$, $A = 0$, $b_2 = 0$, $A_2 = 0$ and

$$b_1 = \begin{pmatrix} -\kappa \\ \frac{\bar{\lambda}^2\gamma}{2(1-\gamma)} \end{pmatrix},$$

$$A_1 = \begin{pmatrix} \varsigma^2 & \frac{\gamma}{1-\gamma}\bar{\lambda}\rho\varsigma \\ \frac{\gamma}{1-\gamma}\bar{\lambda}\rho\varsigma & \frac{\gamma^2}{(1-\gamma)^2}\bar{\lambda}^2 \end{pmatrix}.$$

So $\mathbb{E}\left[Z_0(T)^{\frac{\gamma}{\gamma-1}}\right]$ can be rewritten as moment generating function of our affine process in the point $x = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Additionally we have that the corresponding Riccati equations simplify to the one-dimensional case, since $b_2 = 0$ and $A_2 = 0$, hence we have the ODEs

$$\frac{\partial}{\partial t}\psi_2(t,z) = 0, \quad \psi_2(0,z) = z_2,$$

$$\frac{\partial}{\partial t}\phi(t,z) = \kappa\theta\psi_1(t,z), \quad \phi(0,z) = 0$$

with trivial solutions $\psi_2(t,z) = z_2$ and $\phi(t,z) = \int_0^t \psi_1(s,z)\,\mathrm{d}s$, hence the differential equation of ψ_1 simplifies to

$$\frac{\partial}{\partial t}\psi_1(t,z) = \frac{\varsigma^2}{2}\psi_1(t,z)^2 + \left(\frac{\gamma}{1-\gamma}\bar{\lambda}\rho\varsigma z_2 - \kappa\right)\psi_1(t,z) + \left(\frac{\bar{\lambda}^2\gamma}{2(1-\gamma)} + \frac{\bar{\lambda}^2\gamma^2}{(1-\gamma)^2}z_2\right)z_2, \quad \psi_1(0,z) = z_1.$$

Our Permanent Assumptions ensure $\Delta \in \mathbb{R}$, hence by Theorem 2.58, we have a real valued solution. Since $|\Delta| < -b$ we have that $\frac{b+\Delta}{-b+\Delta} < 0$ and

$$1 - g(0)e^{-\Delta t} = 1 - \frac{b+\Delta}{-b+\Delta} e^{-\Delta t} > 0.$$

Therefore, the solution does not explode and by Corollary 2.60, the expectation of $Z_0(T)^{\frac{\gamma}{\gamma-1}}$ is well defined and of the desired form.

□

Corollary 5.31. *For every $y > 0$ there exists a process ν such that*

$$\mathbb{E}\left[\tilde{U}(yD(T)Z_\nu(T))\right] < \infty.$$

Proof. By Lemma 5.30 we have finiteness of $\mathbb{E}\left[\tilde{U}(Z_0(T))\right]$. Using Theorem 5.28, we only have to show that $\mathbb{E}\left[\tilde{U}(yD(T)Z_0(T))\right] < \infty$. Since y and $D(T)$ are deterministic we get

$$\mathbb{E}\left[\tilde{U}(yD(T)Z_0(T))\right] = y^{\frac{\gamma}{\gamma-1}} D(T)^{\frac{\gamma}{\gamma-1}} \mathbb{E}\left[\tilde{U}(Z_0(T))\right] < \infty.$$

□

Theorem 5.32. *For all $y > 0$, there exists a solution $\hat{\nu} \in \mathcal{K}^1(\sigma)$ to the dual problem $\tilde{V}(y)$. This solution does not depend on y.*

Proof. Since we consider power utility $U(0) = 0 > -\infty$ is satisfied, and by Corollary 5.31 we have that $\tilde{V}(y) < \infty$ for all $y > 0$. The function $x \to xU'(x)$ is monotone increasing and the market price of risk is square integrable by Proposition 5.18. Furthermore, we have by Lemma 5.5 the existence of constants $\alpha \in (0,1)$, $\beta \in (1,\infty)$ such that $\alpha U'(x) \geq U'(\beta x)$ for all $x > 0$, so all conditions of Theorem 4.80 are satisfied. Hence, for all $y > 0$ there exists a solution $\hat{\nu} \in \mathcal{K}^1(\sigma)$ to the dual problem $\tilde{V}(y)$. The independence of y is a consequence of Theorem 5.28. □

Remark 5.33. Since the optimal process for the dual problem $\tilde{V}(y)$ does not depend on y, we will call it from now on the optimal dual process.

Proposition 5.34. *Let $x > 0$ be arbitrary. There exists a solution $\hat{\pi} \in \mathcal{A}(x)$ to our primal problem $V(x)$.*

Proof. By Theorem 5.32 we know a dual solution exists for every $y > 0$. Using Theorem 4.53 and Theorem 4.56 we get the existence of a solution to our primal problem. □

Remark 5.35. Now we have transferred all results from Chapter 4 to Heston's SV model. The remaining question is if we can explicitly calculate the optimal dual and primal process of our market model. Since the martingale method does not provide any mechanics which allow us to find optimal processes, but only gives us a way to verify the optimality of a dual and a primal process, we have to guess the correct processes. Since the discounted wealth process of the optimal portfolio must be a martingale, this guess is not completely arbitrary.

Definition 5.36. For the constants $c = \dfrac{1-\gamma}{1-\gamma+\gamma\rho^2}$, $\beta := -\dfrac{\bar{\lambda}^2\gamma}{2c(1-\gamma)}$, $a := \sqrt{\tilde{\kappa}^2 + 2\beta\varsigma^2}$ and $\tilde{\kappa} := \kappa - \dfrac{\gamma}{1-\gamma}\rho\bar{\lambda}\varsigma$ we define the functions

$$\hat{\psi}(t) := 2\beta \frac{e^{at}-1}{e^{at}(\tilde{\kappa}+a) - \tilde{\kappa} + a},$$

$$\hat{\phi}(t) := \kappa\theta \int_0^t \hat{\psi}(t)\, \mathrm{d}t.$$

Remark 5.37. The functions $\hat{\psi}$ is a solution to the one-dimensional Riccati equation i.e.

$$\frac{\partial}{\partial t}\hat{\psi}(t) = -\frac{\varsigma^2}{2}\hat{\psi}(t)^2 - \tilde{\kappa}\hat{\psi}(t) + \beta.$$

Additionally $\hat{\phi}(t)$ solves

$$\frac{\partial}{\partial t}\hat{\phi}(t) = \kappa\theta\hat{\psi}(t).$$

Together they solve a special kind of Riccati equations we already met in Lemma 5.30.

Definition 5.38. We define

$$\hat{\nu}(t) := c\sqrt{1-\rho^2}\varsigma\hat{\psi}(T-t)\sqrt{z(t)} \begin{pmatrix} -\rho \\ \sqrt{1-\rho^2} \end{pmatrix} \in \mathcal{K}^1(\sigma),$$

where c is as in Definition 5.36.

Lemma 5.39. *The function $\hat{\psi}(T-t)$ satisfies*

$$\frac{\partial}{\partial t}\hat{\psi}(T-t) = \frac{\varsigma^2}{2}\hat{\psi}(T-t)^2 + \tilde{\kappa}\hat{\psi}(T-t) - \beta.$$

Proof. We get this result, simply by differentiating and using the fact that $\hat{\psi}$ solves the Riccati equation with constants $-\frac{\varsigma^2}{2}$, $-\tilde{\kappa}$ and β. □

Lemma 5.40. *The equality*

$$\hat{\psi}(T)z_0 + \hat{\phi}(T) + \frac{\varsigma^2}{2}\int_0^T \hat{\psi}(T-t)^2 z(t)\,\mathrm{d}t + (\tilde{\kappa}-\kappa)\int_0^T \hat{\psi}(T-t)z(t)\,\mathrm{d}t - \beta\int_0^T z(t)\,\mathrm{d}t$$
$$+\varsigma\int_0^T \hat{\psi}(T-t)\sqrt{z(t)}\,\mathrm{d}W^2(t) = 0$$

holds true.

Proof. Integration by parts of the process $\hat{\psi}(T-t)z(t)$ and the fact that $\hat{\psi}(0)=0$ leads to

$$0 = \hat{\psi}(T)z_0 + \int_0^T \frac{\partial}{\partial t}\hat{\psi}(T-t)z(t)\,\mathrm{d}t + \int_0^T \hat{\psi}(T-t)\,\mathrm{d}z(t).$$

Using $\mathrm{d}z(t) = (\kappa\theta - \kappa z(t))\,\mathrm{d}t + \varsigma\sqrt{z(t)}\,\mathrm{d}W^2(t)$ and the definition of $\hat{\phi}(t)$, gives us the equation. □

Remark 5.41. Even though this equality holds we have to be careful if we use it, while we are calculating something under a different measure as $\mathbb{P}$ since $W^2(t)$ is not necessarily a Brownian Motion under this new measure.

Theorem 5.42. *Let $\hat{\nu}$ be the process from Definition 5.38. Then*

$$\mathbb{E}\left[Z_{\hat{\nu}}(T)^{\frac{\gamma}{\gamma-1}}\right] = \exp\left(-\frac{c}{1-\gamma}(\hat{\phi}(T) + \hat{\psi}(T)z_0)\right),$$

where the random variable $Z_{\hat{\nu}}(T)$ is given by

$$Z_{\hat{\nu}}(T) = \exp\left(-\int_0^T \big(\theta(t)+\hat{\nu}(t)\big)^\top\,\mathrm{d}W(t) - \frac{1}{2}\int_0^T \left(\|\theta(t)\|^2 + \|\hat{\nu}(t)\|^2\right)\,\mathrm{d}t\right)$$

with

$$\theta(t) + \hat{\nu}(t) = \sqrt{z(t)}\begin{pmatrix}\bar{\lambda}\sqrt{1-\rho^2} - c\sqrt{1-\rho^2}\rho\varsigma\hat{\psi}(T-t)\\ \bar{\lambda}\rho + (1-\rho^2)c\varsigma\hat{\psi}(T-t)\end{pmatrix}$$

and

$$\|\theta(t)\|^2 + \|\hat{\nu}(t)\|^2 = z(t)\left(\bar{\lambda}^2 + (1-\rho^2)c^2\varsigma^2\hat{\psi}(T-t)^2\right).$$

Proof. To show this we want to apply our knowledge about affine processes and its moment generating functions. The process $Z_{\hat{\nu}}(T)^{\frac{\gamma}{\gamma-1}}$ does not fulfil our definition of such an affine process, since $\hat{\nu}(t)$ depends on $\hat{\psi}(T-t)$, so we want to change the measure in the right way, to obtain a suitable process.
Firstly we define

$$\tilde{Z}(t) := \exp\left(-\frac{\gamma}{1-\gamma}\int_0^t \hat{\nu}_1(t)\,\mathrm{d}W^1(s) - \frac{\gamma^2}{2(1-\gamma)^2}\int_0^t \hat{\nu}_1(t)^2\,\mathrm{d}t\right),$$

where $\hat{\nu}_1(t)$ denotes the first component of the two-dimensional process $\hat{\nu}(t)$. This process is a martingale by Theorem 2.38, hence it is a density process of an equivalent measure $\tilde{P}$ under which the process

$$\tilde{W}(t) = \begin{pmatrix} W^1(t) + \frac{\gamma}{1-\gamma}\int_0^t \hat{\nu}_1(t)\,\mathrm{d}t \\ W^2(t) \end{pmatrix},$$

is a standard Brownian motion. So, we get

$$\mathbb{E}\left[Z_{\hat{\nu}}(T)^{\frac{\gamma}{\gamma-1}}\right] = \mathbb{E}^{\tilde{\mathbb{P}}}\left[\exp\left(\mu(T)\right)\right].$$

The process $\mu(t)$ has the form

$$\begin{aligned}\mu(t) = {} & \frac{\gamma\bar{\lambda}}{1-\gamma}\int_0^t \sqrt{z(s)}\left(\sqrt{1-\rho^2} \quad \rho\right)\mathrm{d}\tilde{W}(s) + \frac{\gamma}{2(1-\gamma)}\bar{\lambda}^2\int_0^t z(s)\,\mathrm{d}s \\ & + \frac{1}{2}\frac{\gamma}{1-\gamma}c(1-\rho^2)\varsigma^2\int_0^t \hat{\psi}(T-s)^2 z(t)\,\mathrm{d}s \\ & + \frac{\gamma}{1-\gamma}c(1-\rho^2)\varsigma\int_0^t \hat{\psi}(T-s)\sqrt{z(s)}\,\mathrm{d}W^2(s) \\ & - \frac{\gamma^2}{(1-\gamma)^2}\bar{\lambda}\rho c(1-\rho^2)\varsigma\int_0^t \hat{\psi}(T-s)z(s)\,\mathrm{d}s.\end{aligned}$$

The first two integrals are nice since they have the right form to ensure that $\mu(t)$ and $z(t)$ are affine. The other three are of different form, but pulling out the factor $\frac{\gamma}{1-\gamma}c(1-\rho^2)$ and using Lemma 5.40 leads us to

$$\begin{aligned}\mu(T) = {} & \frac{\gamma\bar{\lambda}}{1-\gamma}\int_0^T \sqrt{z(s)}\left(\sqrt{1-\rho^2} \quad \rho\right)\mathrm{d}\tilde{W}(s) + \frac{\gamma}{2(1-\gamma)}\bar{\lambda}^2\int_0^T z(s)\,\mathrm{d}s \\ & + \frac{\gamma}{1-\gamma}c(1-\rho^2)\left(-\hat{\phi}(T) - \hat{\psi}(T)z_0 + \beta\int_0^T z(s)\,\mathrm{d}s\right).\end{aligned}$$

Since $\hat{\psi}(T)$, $\hat{\phi}(T)$ and z_0 are constants we can simply pull them outside the expected value. Rearranging the integrals, the definition of c and $\tilde{\kappa}$, and the observation that

$$\begin{aligned}\frac{\gamma}{1-\gamma}c(1-\rho^2)\beta + \frac{\gamma}{2(1-\gamma)}\bar{\lambda}^2 &= \frac{\gamma}{2(1-\gamma)}\left(-(1-\rho^2)\frac{\gamma}{1-\gamma}\bar{\lambda}^2 + \bar{\lambda}^2\right) \\ &= \frac{\gamma}{2(1-\gamma)}\left(\frac{1-\gamma+\gamma\rho^2}{1-\gamma}\bar{\lambda}^2 - \frac{\bar{\lambda}}{1-\gamma} + \frac{(1-\gamma)\bar{\lambda}^2}{1-\gamma}\right) \\ &= -\beta - \frac{\gamma^2}{2(1-\gamma)^2}\bar{\lambda}^2\end{aligned}$$

leads to the following equation

$$\begin{aligned}\mu(T) &= -\frac{\gamma}{1-\gamma}c(1-\rho^2)\left(\hat{\phi}(T)+\hat{\psi}(T)z_0\right) - \beta\int_0^T z(s)\,\mathrm{d}s \\ &+ \frac{\gamma\bar{\lambda}}{1-\gamma}\int_0^T \sqrt{z(s)}\left(\sqrt{1-\rho^2}\quad \rho\right)\mathrm{d}\tilde{W}(s) - \frac{\gamma^2}{2(1-\gamma)^2}\bar{\lambda}^2\int_0^T z(s)\,\mathrm{d}s.\end{aligned}$$

The process

$$\bar{Z}(T) = \exp\left(\frac{\gamma\bar{\lambda}}{1-\gamma}\int_0^T \sqrt{z(s)}\left(\sqrt{1-\rho^2}\quad \rho\right)\mathrm{d}\tilde{W}(s) - \frac{\gamma^2}{2(1-\gamma)^2}\bar{\lambda}^2\int_0^T z(s)\,\mathrm{d}s\right),$$

is again a martingale and leads to a third equivalent measure $\mathbb{Q}$. The martingale property is due to Proposition 5.19 and the fact that $\tilde{\kappa} = \kappa - \frac{\gamma}{1-\gamma}\rho\varsigma\bar{\lambda} > 0$, hence $-\frac{\gamma\bar{\lambda}\rho}{1-\gamma} > -\frac{\kappa}{\varsigma}$. The process

$$\bar{W}(t) = \begin{pmatrix}\bar{W}^1(t)\\ \bar{W}^2(t)\end{pmatrix},$$

with components

$$\bar{W}^1(t) := W^1(t) + \frac{\gamma}{1-\gamma}\int_0^t\left(\hat{\nu}_1(s) - \sqrt{1-\rho^2}\,\bar{\lambda}z(s)\right)\mathrm{d}s,$$

$$\bar{W}^2(t) = W^2(t) - \frac{\gamma}{1-\gamma}\rho\bar{\lambda}\int_0^t z(s)\,\mathrm{d}s,$$

is now a Brownian motion under the new measure. By this, we have the equation

$$\mathbb{E}\left[Z_{\hat{\nu}}(T)^{\frac{\gamma}{\gamma-1}}\right] = \exp\left(-\frac{\gamma}{1-\gamma}c(1-\rho^2)\left(\hat{\phi}(T)+\hat{\psi}(T)z_0\right)\right)\mathbb{E}^{\mathbb{Q}}\left[e^{-\beta\int_0^T z(t)\,\mathrm{d}t}\right].$$

Since we have met this form of process in Example 2.2, we now know how to apply the theory about affine processes, to calculate this expectation, but have to be careful, since the dynamics of $z(t)$ are different under our new measure $\mathbb{Q}$. They are given by

$$\mathrm{d}z(t) = \left(\kappa\theta - \tilde{\kappa}z(t)\right)\mathrm{d}t + \varsigma\sqrt{z(t)}\,\mathrm{d}\bar{W}^2(t).$$

The process

$$X = \begin{pmatrix} z(t) \\ -\beta\int_0^t z(s)\,\mathrm{d}s\end{pmatrix}$$

is affine with coefficients $b = \begin{pmatrix} \kappa\theta \\ 0 \end{pmatrix}$, $b_1 = \begin{pmatrix} -\tilde{\kappa} \\ -\beta \end{pmatrix}$ and $A_1 = \begin{pmatrix} 0 & \varsigma^2 \\ 0 & 0 \end{pmatrix}$. All other coefficients are 0. Since we are only interested in the moment generating function of this process in $x = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, we get as solution of our Riccati equations

$$\begin{aligned} \psi_2(t) &= 1, \\ \psi_1(t) &= -\hat{\psi}(t), \\ \phi(t) &= -\hat{\phi}(t). \end{aligned}$$

This leads us to

$$\begin{aligned} \mathbb{E}\left[Z_{\hat{\nu}}(T)^{\frac{\gamma}{\gamma-1}}\right] &= \exp\left(-\left(1+\frac{\gamma}{1-\gamma}c\right)(1-\rho^2)\left(\hat{\phi}(T)+\hat{\psi}(T)z_0\right)\right) \\ &= \exp\left(-\frac{c}{1-\gamma}\left(\hat{\phi}(T)+\hat{\psi}(T)z_0\right)\right). \end{aligned}$$

□

Remark 5.43. Instead of changing a second time we also could apply the theory of affine processes to the expectation under the measure $\tilde{\mathbb{P}}$. Here $z(t)$ would still have its original dynamics, since $W^2(t)$ is a Brownian motion under $\tilde{\mathbb{P}}$.

Theorem 5.44. *The optimal expected utility from terminal wealth with initial capital $x > 0$ has the form*

$$V(x) = x^\gamma V(1) = \frac{1}{\gamma} x^\gamma \mathbb{E}\left[Z_{\hat{\nu}}(T)^{\frac{\gamma}{\gamma-1}}\right]^{1-\gamma}.$$

Proof. Let $x > 0$ be arbitrary but fixed. By Proposition 5.34 and Proposition 4.67, we know there exists a solution to our primal optimization problem

$$V(x) := \mathbb{E}[U(\xi_{\hat{\nu}}(x))],$$

where $\hat{\nu}$ is the optimal dual process for some $y(x)$. Lemma 5.26 and the definition of $\xi_{\hat{\nu}}(x)$ leads to

$$\begin{aligned} V(x) &= \mathbb{E}\left[U\left(I\left(x^{\gamma-1}D(T)^{1-\gamma}\mathbb{E}\left[Z_{\hat{\nu}}(T)^{\frac{\gamma}{\gamma-1}}\right]^{1-\gamma} Z_{\hat{\nu}}(T)\right)\right)\right] \\ &= x^\gamma \mathbb{E}\left[U\left(\xi_{\hat{\nu}}(1)\right)\right] = x^\gamma V(1). \end{aligned}$$

The last equality holds since the optimal dual process is independent of y. □

Remark 5.45. As we can see, it is enough to calculate the optimal expected utility from terminal wealth with initial capital 1 to determine $V(x)$ for all $x > 0$.

Theorem 5.46. *The optimal dual process is given by $\hat{\nu}$ of Definition 5.38, the optimal portfolio process is given by*

$$\hat{\pi}(t) := \frac{1}{1-\gamma}\bar{\lambda} - \frac{1}{1-\gamma}c\rho\varsigma\hat{\psi}(T-t).$$

Proof. To prove the optimality of $\hat{\nu}$ and $\hat{\pi}$ we have to show that

$$D(T)\xi_{\hat{\nu}}(1) = D(T)X^{\hat{\pi}}(T)$$

holds true, where

$$\xi_{\hat{\nu}}(1) = I\left(\mathcal{Y}_{\hat{\nu}}(1)D(T)Z_{\hat{\nu}}(T)\right) = D(T)^{-1}Z_{\hat{\nu}}(T)^{\frac{1}{\gamma-1}}\mathbb{E}\left[Z_{\hat{\nu}}(T)^{\frac{\gamma}{\gamma-1}}\right]^{-1}$$

$$= D(T)^{-1}Z_{\hat{\nu}}(T)^{\frac{1}{\gamma-1}}\exp\left(\frac{c}{1-\gamma}\left(\hat{\phi}(T) + \hat{\psi}(T)z_0\right)\right),$$

and $X^{\hat{\pi}}(T)$ denotes the wealth process with initial capital 1. The discounted terminal wealth $D(T)X^{\hat{\pi}}$ has, by Corollary 3.33, the representation

$$D(T)X^{\hat{\pi}}(T) = \exp\left(\int_0^T \hat{\pi}(t)\bar{\lambda}\sqrt{z(t)}\left(\sqrt{1-\rho^2} \quad \rho\right)\mathrm{d}B(t) - \frac{1}{2}\int_0^T \hat{\pi}(t)^2\bar{\lambda}^2 z(t)\,\mathrm{d}t\right),$$

where B denotes the process

$$\mathrm{d}B(t) = \mathrm{d}W(t) + \bar{\lambda}\sqrt{z(t)}\begin{pmatrix}\sqrt{1-\rho^2}\\ \rho\end{pmatrix}\mathrm{d}t.$$

We must be careful, since

$$Z_0(t) = \exp\left(-\int_0^s \theta(s)^\top\,\mathrm{d}W(s) - \frac{1}{2}\int_0^t \|\theta(s)\|^2\,\mathrm{d}s\right)$$

is not necessarily a martingale, hence it could be that there is no equivalent measure such that B is a standard Brownian motion, but there is no need of this property in our calculations.
We start with $D(T)\xi_{\hat{\nu}}(1)$ and use Lemma 5.40 to get

$$\begin{aligned}\hat{\phi}(T) + \hat{\psi}(T)z_0 = &-\frac{\varsigma^2}{2}\int_0^T \hat{\psi}(T-t)^2 z(t)\,\mathrm{d}t - (\tilde{\kappa}-\kappa)\int_0^T \hat{\psi}(T-t)z(t)\,\mathrm{d}t\\ &+\beta\int_0^T z(t)\,\mathrm{d}t - \varsigma\int_0^T \hat{\psi}(T-t)\sqrt{z(t)}\,\mathrm{d}W^2(t).\end{aligned}$$

Moreover, we have

$$\varsigma \int_0^T \hat{\psi}(T-t)\sqrt{z(t)}\,\mathrm{d}W^2(t) = \varsigma \int_0^T \hat{\psi}(T-t)\sqrt{z(t)}\,\mathrm{d}B^2(t) - \rho\varsigma\bar{\lambda}\int_0^T \hat{\psi}(T-t)z(t)\,\mathrm{d}t$$

and

$$-\int_0^T \hat{\nu}(t)^\top\,\mathrm{d}B(t) + \varsigma c\int_0^T \hat{\psi}(T-t)\sqrt{z(t)}\,\mathrm{d}B^2(t) = \rho\varsigma c\int_0^T \hat{\psi}(T-t)\sigma(t)\,\mathrm{d}B(t).$$

As we can see, the integrand of the above integral in no longer orthogonal on $\sigma(t)$, but instead in the range. Summing up the integrals leads finally to

$$\begin{aligned} D(T)\xi_{\hat{\nu}}(1) = &\exp\left(\int_0^T \frac{1}{1-\gamma}\left(\hat{\lambda} - \rho\varsigma c\hat{\psi}(T-t)\right)\sigma(t)\,\mathrm{d}B(t)\right) \\ &\exp\left(-\frac{1}{2(1-\gamma)^2}\int_0^T \left(\hat{\lambda} - \rho\varsigma c\hat{\psi}(T-t)\right)^2\mathrm{d}t\right), \end{aligned}$$

which is exactly the discounted wealth process of $\hat{\pi}$ with initial capital 1. Since we now have a dual process $\hat{\nu}$ and portfolio process $\hat{\pi}$ with $\xi_{\hat{\nu}}(1) = X^{\hat{\pi}}(T)$, we know that both are optimal. □

Remark 5.47. If we consider Heston's SV model, with stock price

$$\mathrm{d}S(t) = S(t)\left(\left(\bar{\lambda}z(t) + r\right)\mathrm{d}t + \sqrt{z(t)}\,\mathrm{d}\hat{W}^1(t)\right)$$

and treat $z(t)$ as a constant, we end up in the Black-Scholes setting with $\mu = \bar{\lambda}z(t) + r$ and $\sigma = \sqrt{z(t)}$. Using the solution of Merton's portfolio problem, we see that the optimal portfolio process, in this model is given by

$$\pi^*(t) = \frac{\mu - r}{(1-\gamma)\sigma^2} = \frac{\bar{\lambda}}{1-\gamma},$$

which is exactly the first term of the process

$$\hat{\pi}(t) = \frac{\bar{\lambda}}{1-\gamma} - \frac{c\rho\varsigma}{1-\gamma}\hat{\psi}(T-t).$$

Hence, the optimal portfolio process in Heston's SV model, separates into two parts. The first part, which is similar to the solution in the Black-Scholes model, and a second term, compensating the additional uncertainty, caused by the correlation between the Brownian motions $\hat{W}^1$ and $\hat{W}^2$.

Theorem 5.48. *The optimal expected utility from terminal wealth with initial capital x is given by*

$$V(x) = \frac{1}{\gamma}x^\gamma \exp\left(-c\left(\hat{\phi}(T) + \hat{\psi}(T)z_0\right)\right).$$

Proof. We know that

$$V(x) = \mathbb{E}[U(\xi_{\hat{\nu}}(x)] = x^{\gamma}\mathbb{E}[U(\xi_{\hat{\nu}}(1)].$$

If we plug $\xi_{\hat{\nu}}(1)$ and use Theorem 5.42 in we get

$$V(x) = \frac{1}{\gamma}x^{\gamma}\mathbb{E}\left[Z_{\hat{\nu}}(T)^{\frac{\gamma}{\gamma-1}}\right]^{1-\gamma} = \frac{1}{\gamma}x^{\gamma}\exp\left(-c(\hat{\phi}(T) + \hat{\psi}(T)z_0)\right).$$

□

5.4 Including Discrete Dividends

The diffusion markets we introduced in the previous chapters, do not include any dividend payments. Profit from stocks is only gained via the rise of the price itself, but in reality most companies pay dividends to their shareholders. These payments are naturally discrete, and the payment times are generally known in advance. To model this, we will include the theory of early announced dividends, described by Grün (2017), to Heston's SV model and study how this extension influences the results we already derived.

5.4.1 Basic Notation

Definition 5.49. For every dividend payment of a stock, there are four important dates. The dividend is paid on the *payment date*. On the *announcement date*, the exact worth of the dividend will be announced. This is naturally some days before the payment date. The *cum-dividend date* or short *cum-date* is the last date on which a buyer of a unit of the stock still receives the dividend. The *ex-dividend date* or *ex-date* is the date after the cum-date, so everyone who buys a stock on the ex-date will not receive the corresponding dividend.

Remark 5.50. Since everyone who buys the stock on cum-date gets the next dividend, but the buyer on ex-date does not, there is usually a drop in the stock price on ex-date.

Permanent Assumption 5.51. The ex-date coincides with the payment date.

Permanent Assumption 5.52. There are deterministic time points $0 < T_1 < \ldots < T_n \leq T$ on which the dividends are paid. The amount of the dividend paid at T_i is given by D_i.

Permanent Assumption 5.53. The drop of the stock price at the ex-dividend date for dividend i is exactly the amount of the dividend D_i.

5.4.2 Existing Results

Remark 5.54. Here, we will recapitulate some results regarding portfolio optimization with early announced dividends in the Black-Scholes market, derived by Grün (2017). Alternatively, a compact summary can also be found in Desmettre et al. (2018).

Definition 5.55. The market model consists of a risk-free asset B and a dividend paying stock S. The corresponding price processes are given by

$$\mathrm{d}B_t = rB(t)\mathrm{d}t, \quad B(0) = B_0 > 0$$

$$S(t) = S_D(t) + S_E(t), \quad S(0) = S_0 > 0,$$

where $S_D(t)$ denotes the price process of the dividend part of the stock and $S_E(t)$ is the ex-dividend price process which is related to the a fictitious dividend free stock, modelled by a geometric Brownian motion

$$\mathrm{d}\tilde{S}(t) = \tilde{S}(t)\left(\mu\mathrm{d}t + \sigma\,\mathrm{d}W(t)\right), \quad \tilde{S}(0) = \tilde{S}(0) > 0.$$

Here the constants satisfy $r \geq 0$, $\mu \in \mathbb{R}$ and $\sigma > 0$.

Problem 5.56. Again we want to solve the portfolio optimization problem

$$V(x) = \sup_{\pi \in \mathcal{A}(x)} \mathbb{E}\left[U(X^\pi(T))\right],$$

where $\mathcal{A}(x)$ is defined as in Problem 4.12.

Model 1. *We assume that all future dividends up to time T are already known at time $t = 0$. The price process $S_D(t)$ is given as*

$$S_D(t) = \sum_{i;t<T_i} e^{-r(T_i - t)} D_i.$$

Furthermore, the ex-dividend price is given by

$$\mathrm{d}S_E(t) = \mathrm{d}\tilde{S}(t), \quad S_E(0) = S(0) - S_D(0) > 0.$$

Theorem 5.57. *Between two payment dates, the wealth process X^π of a portfolio process π in Model 1 can be rewritten as wealth process $X^{\tilde{\pi}}$ of a different price process $\tilde{\pi}$ in the standard Black-Scholes model, where*

$$\tilde{\pi}(t) := \pi(t)\frac{S_E(t)}{S(t)}.$$

Proof. The main idea of the proof is that, between two payment dates, the stock price with dividends can be replicated by $\tilde{S}$ and the bank account $B(t)$, hence every portfolio process consisting of S and B can be reproduced by a portfolio process which invests only in $\tilde{S}$ and B. For further details, we refer to Grün (2017).

□

Theorem 5.58. *The optimization problem with known dividends is equivalent to the portfolio optimization problem without dividends in the sense that the optimal expected utilities from terminal wealth coincide. The optimal portfolio processes are related by*

$$\hat{\pi}_{DIV}(t) = \hat{\pi}_{BS}(t)\frac{S(t)}{S_E(t)},$$

where $\hat{\pi}_{BS}(t)$ denotes the optimal portfolio process in the Black-Scholes market model without dividends.

Proof. Because of Theorem 5.57 we know that between two payment times, the markets evolve in the same way. Since the dividend payment at T_i does not change the wealth, which is due to Permanent Assumption 5.53, we simply have to reinvest the paid dividends like Korn and Rogers (2005) proposed. The relation between the optimal portfolios is a result of the way S is replicated in the original Black-Scholes Market. Again, we refer to Grün (2017) for more details. □

Remark 5.59. Since naturally not all dividend payments are known at time $t = 0$, we now replace this model for discrete dividends by a model which includes early announcement. This means that the dividend payment D_i is known at some time $\tilde{T}_i < T_i$ but not necessarily at time $t = 0$.

Model 2. *We assume the dividends are announced at some date $\tilde{T}_i < T_i$ and are proportional to the stock price before the announcement, i.e. $e^{-r(T_i - \tilde{T}_i)} D_i = \alpha_i S(\tilde{T}_i -)$. Let $\mathcal{N}(t) := \left\{ 1 \le i \le n \,\middle|\, \tilde{T}_i \le t < T_i \right\}$, then the price processes are given by*

$$S_D(t) = \sum_{i \in \mathcal{N}(t)} \left(\prod_{j > i; j \in \mathcal{N}(t)} (1 - \alpha_j) \right) D_i e^{-r(T_i - t)},$$

$$S_E(t) = \prod_{i; \tilde{T}_i \le t} (1 - \alpha_i) \tilde{S}(t).$$

Lemma 5.60. *Model 2 fulfils our permanent assumption only if every announcement date is after all payment dates of previously announced dividends.*

Remark 5.61. Even though in reality it is often the case that payments are paid before new ones are announced, this model is not completely satisfactory, since it excludes the possibility to announce multiple dividends at the same time.

Model 3. *This model is like Model 2, but now we assume that the dividends are proportional to the ex-dividend price process $e^{-r(T_i - \tilde{T}_i)} D_i = \alpha_i S_E(\tilde{T}_i -)$. The price process can then again be separated into S_D and S_E but now*

$$S_D(t) = \sum_{i \in \mathcal{N}(t)} D_i e^{-r(T_i - \tilde{T}_i)}$$

$$S_E(t) = \prod_{i;\tilde{T}_i \le t} (1-\alpha_i)\tilde{S}(t).$$

Lemma 5.62. *Model 3 always satisfies Permanent assumption 5.53.*

Theorem 5.63. *The optimal utility from terminal wealth of Model 2 and Model 3 coincide and are equal to the optimal utility from terminal wealth in the standard Black-Scholes market model.*

Proof. In Model 3 we can express the dividend free stock $\tilde{S}$ by S and B via

$$\tilde{S}(t) = \left(\prod_{i;\tilde{T}_i \le t} (1-\alpha_i)^{-1}\right) S(t) - \left(\prod_{i;\tilde{T}_i \le t} (1-\alpha_i)^{-1} \sum_{i\in\mathcal{N}(t)} \frac{D_i}{B(T_i)}\right) B(t).$$

So again we can express any portfolio process in this model by a portfolio process in the dividend free model and vice versa. Hence, the portfolio optimization problems are equivalent. The property for Model 2 can be proven analogously, but here we have

$$\tilde{S}(t) = \left(\prod_{i;\tilde{T}_i \le t} (1-\alpha_i)^{-1}\right) S(t) - \left(\prod_{i;\tilde{T}_i \le t} (1-\alpha_i)^{-1} \sum_{i\in\mathcal{N}(t)} \prod_{j>i;j\in\mathcal{N}(t)} (1-\alpha_j)\frac{D_i}{B(T_i)}\right) B(t).$$

□

Theorem 5.64. *The optimal portfolio process in Model 2 and Model 3 is given by*

$$\hat{\pi}(t) = \hat{\pi}_{BS}(t)\frac{S(t)}{S_E(t)},$$

where $\hat{\pi}_{BS}(t)$ denotes the optimal portfolio process in the dividend free model.

Proof. Let $\hat{\pi}_{BS}(t)$ denote the optimal portfolio process in the dividend free Black-Scholes model. We know that we must invest this part of the money into $\tilde{S}$ to get our optimal expected utility at T. If we rewrite $\tilde{S}$ by S and B, we get that the wealth processes of $\hat{\pi}_{BS}(t)$ in the dividend free market and $\hat{\pi}(t) = \hat{\pi}_{BS}(t)\frac{S(t)}{S_E(t)}$ in the market model with dividends coincide. Since the optimal expected utilities from terminal wealth are the same, we get that $\hat{\pi}(t)$ is optimal. □

5.4.3 Application to Heston's Stochastic Volatility model

Remark 5.65. We now want to include discrete dividends into the portfolio optimization problem in Heston's model. For this, we will work with the dividend models introduced in the previous section and just change the dynamics of the fictitious dividend free stock $\tilde{S}$.

Definition 5.66. The market model consists again of a risk-free asset B and a dividend paying stock S. The corresponding price processes are given by

$$\mathrm{d}B_t = rB(t)\mathrm{d}t, \quad B(0) = B_0 > 0,$$

$$S(t) = S_D(t) + S_E(t), \quad S(0) = S_0 > 0,$$

where $S_D(t)$ like in Definition 5.55, but now the fictitious dividend free stock follows the dynamics

$$\mathrm{d}\tilde{S}(t) = \tilde{S}(t)\left((\bar{\lambda}z(t) + r)\mathrm{d}t + \sqrt{z(t)}\left(\sqrt{1-\rho^2}\,\mathrm{d}W^1(t) + \rho\,\mathrm{d}W^2(t)\right)\right),$$

where the volatility process $z(t)$ satisfies

$$\mathrm{d}z(t) = \kappa(\theta - z(t))\mathrm{d}t + \varsigma\sqrt{z(t)}\,\mathrm{d}W^2(t).$$

The constants are assumed to satisfy all assumptions of Section 5.2.

Theorem 5.67. *The optimal expected utility from terminal wealth of Model 1, Model 2 and Model 3 does not differ from the optimal value of the same optimization problem without dividends.*

Proof. The proof is exactly like the corresponding proofs in the Black Scholes setting. Again, we can express the dividend free stock by S and B and the dividend-paying stock between the payment dates can be rewritten in terms of $\tilde{S}$ and B. Since the wealth process, X does not jump at payment dates T_i, we can immediately reinvest the paid dividend and continue our strategy. Hence, the portfolio optimization problems are equivalent. □

Theorem 5.68. *The optimal portfolio process in the market model with dividends is related, to the dividend free optimal portfolio by*

$$\hat{\pi}(t) = \hat{\pi}_{HS}(t)\frac{S(t)}{S_E(t)},$$

where $\hat{\pi}_{HS}$ denotes the optimal portfolio process in Heston's stochastic volatility model without dividends.

Proof. Similar to the proofs before we take the optimal portfolio of the dividend free model and rewrite $\tilde{S}$ in terms of S and B to get a portfolio process in the market with discrete dividends which has almost surely the same terminal wealth. Since the optimal values coincide, this portfolio process must be optimal. □

Remark 5.69. Surprisingly, the inclusion of early announced dividends does not change the optimal expected utility from terminal wealth, although the stock price becomes discontinuous. This is mainly due to the preserved continuity of the wealth process, which allows replication through portfolio processes in the dividend free market model. The only remarkable difference of the models with discrete dividends and the one without is that we must invest a larger proportion of our wealth into the stock, hence take more risk, to achieve optimality.

6 Conclusion

In contrast to Merton's problem, it is not always obvious how the expected utility from terminal wealth in stochastic volatility models can be maximized. To apply the martingale method of portfolio optimization to Heston's stochastic volatility model, we have to consider two major restrictions. Firstly, the coefficients of the squared volatility process have to satisfy the Feller condition. This guarantees the positivity of the Cox-Ingersoll-Ross process over the whole time period and subsequently the full rank of the volatility process.

The second crucial assumption is

$$\frac{\kappa^2}{2\varsigma^2} > \frac{\gamma}{1-\gamma}\bar{\lambda}\left(\frac{\kappa\rho}{\varsigma} + \frac{\bar{\lambda}}{2}\right),$$

because it ensures, that the primal optimization problem is well defined for all finite time horizons. If the inequality is satisfied, the optimal expected utility from terminal wealth can be calculated, using the structure of the moment generating functions of affine processes.

We demonstrate that under these conditions, for every positive y, the dual optimization problem has a minimal value in the form of

$$\tilde{V}(y) = \frac{1-\gamma}{\gamma} y^{\frac{\gamma}{\gamma-1}} \mathbb{E}\left[\left(D(T)Z_{\hat{\nu}}(T)\right)^{\frac{\gamma}{\gamma-1}}\right],$$

where $D(T)$ denotes the discounting factor and $\hat{\nu}$ a dual process, which is independent of y. The structure is a direct consequence of the usage of power utility functions. These functions also ensure the existence and uniqueness of the optimal dual process

$$\hat{\nu}(t) := c\sqrt{1-\rho^2}\varsigma\hat{\psi}(T-t)\sqrt{z(t)}\begin{pmatrix} -\rho \\ \sqrt{1-\rho^2} \end{pmatrix}.$$

Using the relation between dual and primal optimization problem, the optimal expected utility from terminal wealth with positive initial capital x can be determined as

$$V(x) = x^{\gamma}V(1) = \frac{1}{\gamma}x^{\gamma}\exp\left(-c\left(\hat{\phi}(T) + \hat{\psi}(T)z_0\right)\right).$$

In contrast to Merton's portfolio problem, the optimal portfolio process

$$\hat{\pi}(t) = \frac{\bar{\lambda}}{1-\gamma} - \frac{c\rho\varsigma}{1-\gamma}\hat{\psi}(T-t)$$

is time dependent, but still non-random.

The theory of early announced dividends can be directly included into Heston's SV model. Like in the Black-Scholes setting, the optimal expected utility does not change in any of the presented models, mainly because of the continuity of the corresponding wealth processes. Similarly to the results derived by Grün 2017, the optimal portfolio processes change and to achieve optimality, an investor must invest a greater proportion of his wealth into the risky asset.

Bibliography

Bellman, Richard (1958). “Dynamic programming and stochastic control processes”. In: *Information and Control* 1.3, pp. 228–239. DOI: 10.1016/S0019-9958(58)80003-0.

Black, Fischer and Myron Scholes (1973). “The pricing of options and corporate liabilities”. In: *Journal of political economy* 81.3, pp. 637–654.

Cohen, Samuel N. and Robert J. Elliott (2015). *Stochastic Calculus and Applications*. Second edition. Probability and Its Applications. New York et al.: Springer. DOI: 10.1007/978-1-4939-2867-5.

Cox, John C. and Chi-fu Huang (1989). “Optimal consumption and portfolio policies when asset prices follow a diffusion process”. In: *Journal of Economic Theory* 49.1, pp. 33–83. DOI: 10.1016/0022-0531(89)90067-7.

Desmettre, Sascha, Sarah Grün, and Ralf Korn (2018). “Portfolio optimization with early announced discrete dividends”. In: *Operations Research Letters* 46.5, pp. 548–552. DOI: 10.1016/j.orl.2018.09.001.

Ekeland, Ivar and Roger Temam (1999). *Convex analysis and variational problems*. Vol. 28. Classics in applied mathematics. Philadelphia: Society for Industrial and Applied Mathematics.

Filipovic, Damir (2009). *Term-Structure Models*. Berlin, Heidelberg: Springer. DOI: 10.1007/978-3-540-68015-4.

Grün, Sarah (2017). “Discrete Dividends: Modeling, Estimation and Portfolio Optimization”. Dissertation. Kaiserslautern: Technische Universität Kaiserslautern. URL: https://kluedo.ub.uni-kl.de/frontdoor/deliver/index/docId/5213/file/Discrete_Dividends_Sarah_Gruen.pdf.

Heston, Steven L. (1993). “A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options”. In: *Review of Financial Studies* 6.2, pp. 327–343. DOI: 10.1093/rfs/6.2.327.

Kallsen, Jan and Johannes Muhle-Karbe (2010). “Utility Maximization in Affine Stochastic Volatility Models”. In: *International Journal of Theoretical and Applied Finance* 13.03, pp. 459–477. DOI: 10.1142/S0219024910005851.

Karatzas, Ioannis, John P. Lehoczky, and Steven E. Shreve (1987). “Optimal Portfolio and Consumption Decisions for a “Small Investor” on a Finite Horizon”. In: *SIAM Journal on Control and Optimization* 25.6, pp. 1557–1586. DOI: 10.1137/0325086.

Karatzas, Ioannis, John P. Lehoczky, Steven E. Shreve, and Gan-Lin Xu (1991). “Martingale and Duality Methods for Utility Maximization in an Incomplete Market”. In: *SIAM Journal on Control and Optimization* 29.3, pp. 702–730. DOI: 10.1137/0329039.

Karatzas, Ioannis and Steven E. Shreve (1988). *Brownian motion and stochastic calculus.* Vol. 113. Graduate texts in mathematics. New York: Springer.

– (2003). *Methods of mathematical finance.* Vol. 39. Applications of mathematics. New York: Springer.

Korn, Ralf and L. C. G. Rogers (2005). "Stocks Paying Discrete Dividends". In: *The Journal of Derivatives* 13.2, pp. 44–48. DOI: 10.3905/jod.2005.605354.

Kraft, Holger (2003). *Optimal Portfolios and Heston's Stochastic Volatility Model.* URL: https://www.cofar.uni-mainz.de/dgf2003/paper/paper127.pdf.

Kramkov, D. and W. Schachermayer (1999). "The asymptotic elasticity of utility functions and optimal investment in incomplete markets". In: *The Annals of Applied Probability* 9.3, pp. 904–950. DOI: 10.1214/aoap/1029962818.

Merton, Robert C. (1969). "Lifetime portfolio selection under uncertainty: The continuous-time case". In: *The review of economics and statistics* 51.3, pp. 247–257.

Pham, Huyên (2009). *Continuous-time Stochastic Control and Optimization with Financial Applications.* Vol. 61. Stochastic Modelling and Applied Probability. Berlin, Heidelberg: Springer. DOI: 10.1007/978-3-540-89500-8.

Pliska, Stanley R. (1986). "A Stochastic Calculus Model of Continuous Trading: Optimal Portfolios". In: *Mathematics of Operations Research* 11.2, pp. 371–382. DOI: 10.1287/moor.11.2.371.

Revuz, Daniel and Marc Yor (1999). *Continuous Martingales and Brownian Motion.* Corrected Third Printing of the Third Edition. Vol. 293. Grundlehren der mathematischen Wissenschaften, A Series of Comprehensive Studies in Mathematics. Berlin and Heidelberg: Springer. DOI: 10.1007/978-3-662-06400-9.

Rockafellar, R. Tyrrell (1997). *Convex analysis.* Princeton paperbacks. Princeton, NJ: Princeton University Press. DOI: 10.2307/j.ctt14bs1ff.

Wong, Bernard and C. C. Heyde (2006). "On changes of measure in stochastic volatility models". In: *Journal of Applied Mathematics and Stochastic Analysis* 2006, pp. 1–13. DOI: 10.1155/JAMSA/2006/18130.

Zariphopoulou, Thaleia (2001). "A solution approach to valuation with unhedgeable risks". In: *Finance and Stochastics* 5.1, pp. 61–82. DOI: 10.1007/PL00000040.