Asymptotic Behaviour of Maximised Expected Utility in a Fast Mean Reverting Environment



Yuze Jin

St John's College

University of Oxford

A thesis submitted for the degree of

MSc by Research

Trinity 2025

Acknowledgements

I would like to thank my supervisors Prof. Samuel Cohen and Prof. Christoph Reisinger for their support and guidance throughout this project. I would also like to thank my parents for all their support, faith, and interest.

Abstract

We consider the stochastic control problem of maximising the expected value of a utility function of wealth in a market with fast mean reverting assets. A step on the way to understanding this problem is to find the asymptotic behaviour of the maximised expected utility as the rate of mean reversion diverges to infinity. To simplify the problem, we mainly focused on the basic setting such that the market contains only a fast mean reverting asset and a constant bank account. In this setting, our problem can be transformed to the problem of finding the long time behaviour of a solution to the relevant Hamilton–Jacobi–Bellman equation. It turns out that asymptotically this solution is separated into a divergent term independent of the initial price of the asset and correction terms, both of which are related to terms of the solution to a related ergodic PDE. Our approach relies on both a probabilistic and a viscosity solution argument. We also discuss how we can extend these results to more general settings including a multi-scale asset setting and a non-linear market setting.

Contents

1 Introduction **6**

1.1 Intuition of the Problem 6

1.2 A Motivating Example 7

1.3 Related Literature 15

1.4 Outline of the Thesis 16

1.5 General Formulation of the Problem 18

2 Maximisation of Expected Power Utility of Wealth Process **22**

3 Maximisation of Expected Logarithmic Utility via Probabilistic Approach **28**

3.1 Setting of the Problem 28

3.2 Basic Tools 30

3.3 Main Result 32

4 Maximisation of Expected Logarithmic Utility via Viscosity Approach **34**

4.1 Brief Idea 34

4.2 Existence and Uniqueness of Viscosity Solution to Cauchy Problem 36

4.3 Existence and Uniqueness of Viscosity Solution to Ergodic Problem 37

4.4 Convergence of Viscosity Solutions to the Cauchy Problem 43

4.5 Probabilistic Representation of Parameters in Theorem 4.4.1 48

5 Extension to Multi-asset Setting **51**

5.1 Extension to Linear Market with Both Fast and Slow Moving Assets 51

5.1.0.1 Extension to More General Case 63

5.2 A Simple Example of Extension to a Non-linear Market 66

5.3 Perpendicular Noise 69

5.4 A More General Extension to a Non-linear Market 72
5.5 The Choice of Trading Strategy 80

6 Conclusion **84**

Bibliography **86**

Chapter 1

Introduction

1.1 Intuition of the Problem

The problem we consider in this paper is motivated by the following phenomenon. Consider a person who obtains a time series of asset prices from a market and wants to deduce a good trading strategy using these data. If he knows or observes that the assets in the market are related, a possible idea is to build a statistical model for the market. However, when building a statistical model, one cares about different aspects of the market from trading in the market, such as ensuring a good fit with observations on average rather than perfect agreement with economic constraints. As a result, a good statistical model is not necessarily a suitable model for deducing a good trading strategy. In particular, people typically find apparent (asymptotic) arbitrage opportunities in such statistical models. In this case, the statistical model suggests that there are trading strategies where the expected optimal wealth explodes, and the optimal trading strategy is unstable or explodes as well. If we use such a trading strategy in the real market, we highly likely cannot gain the wealth we expect. These issues suggest that we cannot use standard statistical techniques to build models, and instead need to use bespoke methods which avoid arbitrage, which are usually more complex. However, the simplicity of such a statistical model is often desirable, so we want to find a new way of working with this problematic model which can avoid these errors and gets a reasonable trading strategy. A first step of this approach is to understand how our wealth process explodes in such a model and recognize the source of divergence so that we can avoid this issue. In this thesis, we will mainly focus on the divergence caused by the fast oscillating behaviour of market factors. We will consider the asymptotic behaviour of the maximised expected utility of wealth process when trading in such a market.

1.2 A Motivating Example

To gain a better understanding of the problems one may encounter when using a statistical model, we consider the following example. Consider a market with a stock whose price S_t is approximately determined by a Heston model, and a European call options written on it with maturity times T and exercise prices K, and a risk-free bank account B_t. For simplicity, we assume that our market model is free of arbitrage and has a zero market price of risk processes.

Now consider a person who observe a time series of the prices of assets in this market but did not know that this market has an underlying Heston model. He can approximate the price of the call option C_t at time t as a function of strike K, stock price S_t and time to maturity $T - t$. This approximation has an error as it does not include the stochastic volatility term, but may be a good approximation if the volatility is mean-reverting. He can also add more factors to this model or consider a market with more assets; this kind of approximation has been well studied in the past decades using different methods from basic linear regression to more advanced methods such as artificial neural networks and is widely used in mathematical finance, for example [7] or [29]. As this model has a dimension lower than the real market and does not fit the data perfectly, there will be an error between the real price data and the approximated price. He can model this using an additional mean-reverting noise factor, to get a better approximation or reduce the arbitrage in the approximated lower-dimensional market.

Consider now that he wants to use this model to deduce a good trading strategy. Intuitively, if he approximates the dynamics of stock in the lower dimensional market model and the noise factors with suitable Itô's processes, e.g. the market factors and noise factors are independent of each other with suitable drift restriction, there is no arbitrage in this market model. Hence, it seems reasonable for him to deduce an optimal trading strategy by maximising the expected value of some utility function of his wealth process. However, this arbitrage-free market model may still cause a problem.

Here we use a concrete example to get some intuitive ideas about the possible statistical behaviour of this market model that leads to the problem.

In this example, we assume our true underlying market model follows a Heston model with dynamics of the stock and volatility:

$$\begin{aligned} dS_t &= \sqrt{V_t} S_t \, dW_t^1, \\ dV_t &= 5.1300(0.0436 - V_t) \, dt + 0.5200 \sqrt{V_t} \, dW_t^2, \end{aligned}$$

where W_t^1 and W_t^2 are Brownian motions with quadratic covariation $[W_t^1, W_t^2] = -0.7540t$ and the risk free interest rate in this market is 0. These parameters were calibrated from the S&P 500 index data from 2nd January 1990 to 30th September 2003 in [8] while time t is counted in years. We also assume that the initial conditions of the stock and volatility are $S_0 = 2.5$ and $V_0 = 0.04$. We start by simulating a time series of the stock price using the Euler-Maruyama method with 1000 steps for $0 < t < 1$. Then we consider a European Call option with strike $K = 3.2$ and maturity time $T = 1.5$ in this market. For simplicity, we assume that the market price of the volatility risk is 0. We calculate the price C_t for this option for each triple (t, S_t, V_t) using the Monte Carlo method with 300 steps for each path and 10000 simulations. This can also be calculated using a Fourier-based method, but this accuracy is not significant here. In Figure 1.1, we give an example of time series of the stock price and the related simulated price of the European call option.

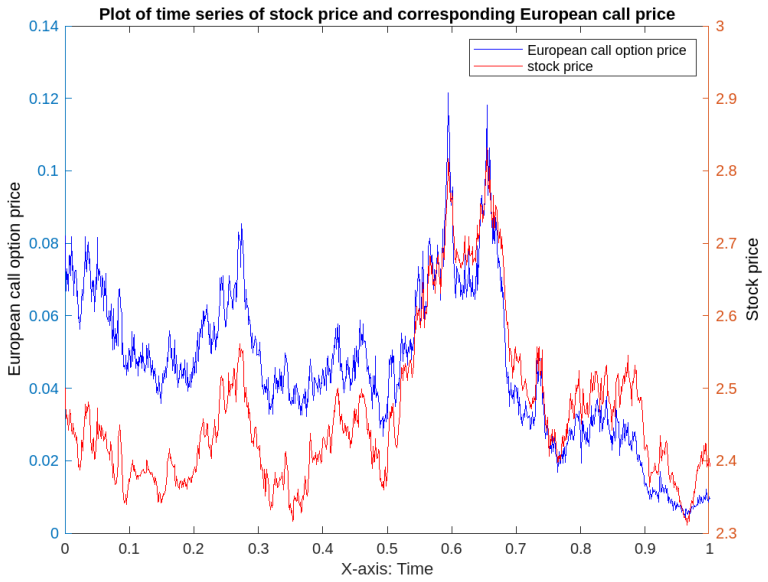


Figure 1.1: The price of the stock and the related European call option

Consider a person who does not know anything about this market, and observes the data of these asset prices. From Figure 1.1 he can observe that the assets in the market are highly related. Hence, he may choose to approximate the real market using a lower-dimensional statistical model with additional noise factor. Here, as an example, we choose a very simple model without additional restrictions. In particular, we assume that the approximation function has the form

$$\hat{C}_t = \beta_1 S_t^2 + \beta_2 S_t + \beta_3 + \beta_4 t S_t^2 + \beta_5 t S_t + \beta_6 t,$$

where $C_t = \hat{C}_t + e_t$ for an error term e_t. These parameters will be obtained by applying the least squares method to our data. The following Figure 1.2 gives an example of such

an approximation. From Figure 1.2 (b) we can see that the the real market price data lie around the plane given by this polynomial.

We next consider the difference between the price of the European call option obtained from the Heston model and the price of the European call option obtained from our approximated models. In Figure 1.3 we give the time series of the error term ϵ_t, that is, the difference between the real European call option price and the approximated European call option price. We can see that the the error term process is small compared with the asset price and has fast oscillating behaviour. In Figure 1.4, we compared the cumulative distribution function of our error data with a suitable normal distribution with almost 0 mean. We see that distributions of errors are close to a suitable normal distribution. Hence, it is superficially reasonable to approximate these errors $e_t := C_t - \hat{C}_t$ using a mean reverting process such as an Ornstein–Uhlenbeck process with zero mean. Here I give a naive approximation such that the parameters of the Ornstein–Uhlenbeck process are calculated using an discrete $AR(1)$ approximation whose parameters are approximated using linear regression.

$$de_t = -561 e_t dt + 0.145 dW_t.$$

Comparing with parameters in the original market, we can see this noise factor has quite fast mean reverting behaviour. In addition, as this noise factor is much smaller than the price of stock and approximated European call option, we can also regard this process as a scaled version of a fast oscillating process. As the scale of the stock is approximately 50 times larger than the error terms, Choosing $\epsilon = 0.0004$ and $\delta = 0.02$ the dynamic of the noise factor can also be written as

$$de_t = -0.2244 \frac{1}{\epsilon} e_t dt + 0.145 \frac{\delta}{\sqrt{\epsilon}} dW_t$$

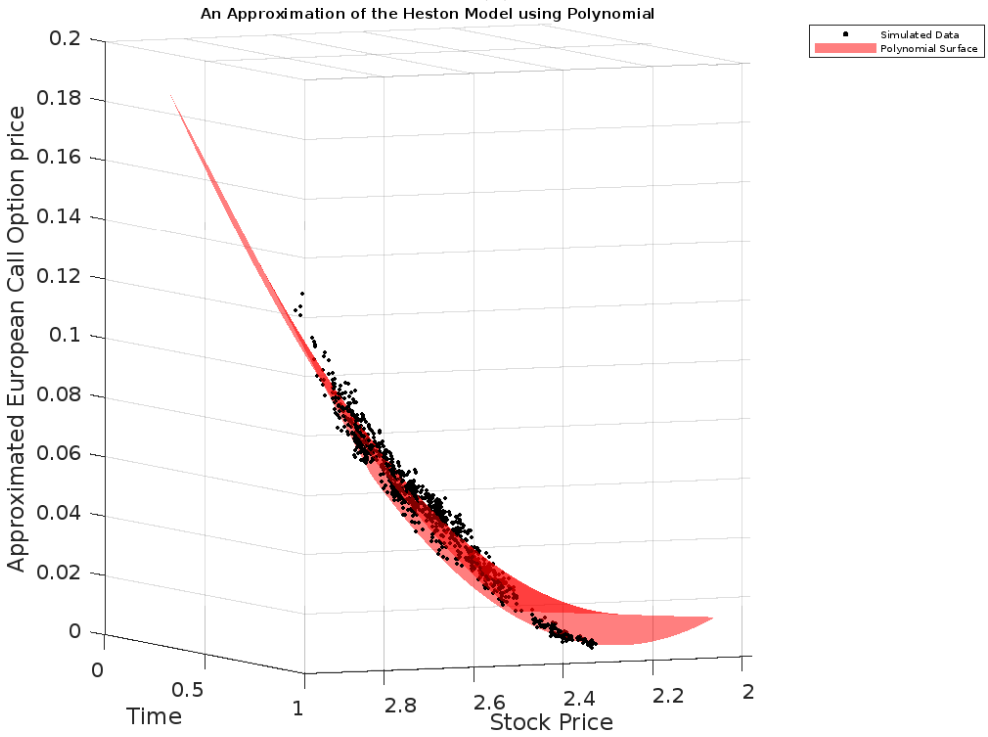
Here $\frac{1}{\epsilon}$ determines the rate of oscillation and δ determines the scale of the noise factor.

From this approximation we can see that, if someone tries to approximate the error terms using an Ornstein–Uhlenbeck process, it is possible for him to get a scaled version of a fast oscillating process.

This example suggests that our real Heston market is visually indistinguishable and may even be statistically indistinguishable from a polynomial model with an additional (scaled) fast oscillating process if we only observe a time series of prices from the market. Although this example is just a quite simple and naive model to approximate the asset price, we expect that a similar phenomenon will occur if we approximate using more advanced



(a)



(b)

Figure 1.2: The approximation of the price of the European option using polynomial $\hat{C}_t = 0.1387S_t^2 - 0.4462S_t + 0.3332 + 0.0540tS_t^2 - 0.4275tS_t + 0.6629t$ where parameters are obtained using least square method. Figure (b) is a rotative view of Figure (a) so that we can see that the data points lie close to the polynomial surface more clearly.

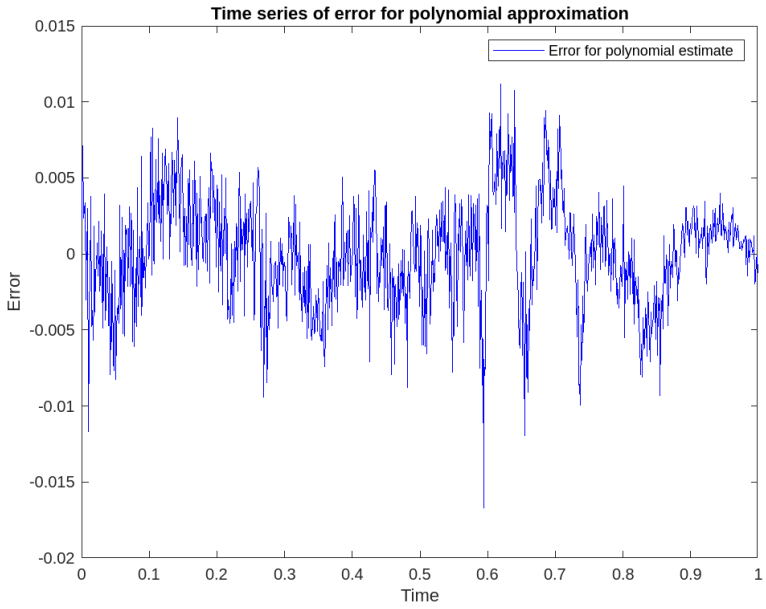


Figure 1.3: A time series for the difference between real European call option prices and approximated European call option price for polynomial approximations.

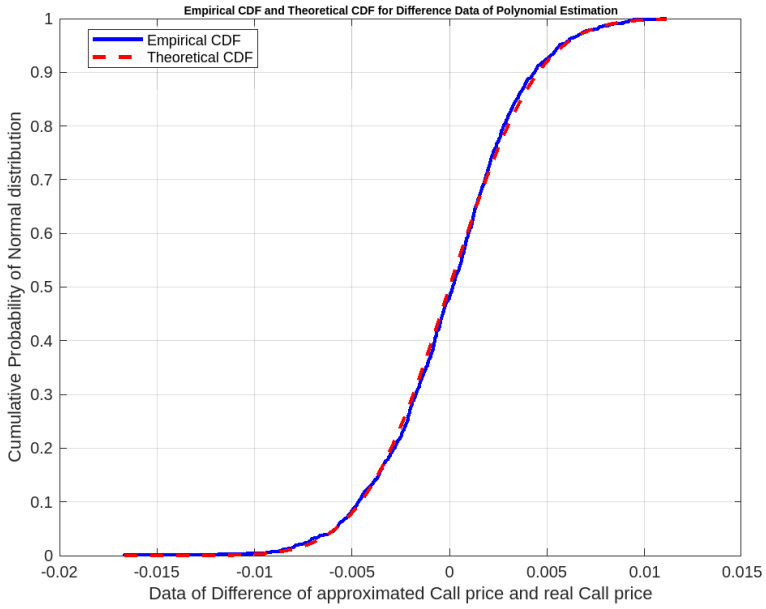


Figure 1.4: A plot of the cumulative distribution function of a time series of model error for a polynomial approximation with a normal distribution with same mean and variance.

methods such as an artificial neural network. This is because neural networks are a class of non-linear regressions.

If we choose to work with this regression-based model, we see that there are two potential sources of arbitrage in our approximated model and that we cannot remove both simultaneously. On the one hand, there could be arbitrage in the lower-dimensional market model if there are no additional noise factors or asymptotic arbitrage as the scale of noise factor converges to 0. For example, if we approximate the price of an asset using a strictly convex and time-independent function of the price of another asset, there is no equivalent measure under which the prices of both assets are martingale due to Jensen's inequality. On the other hand, these fast mean reverting noise factors may induce a statistical arbitrage as the rate of mean reversion diverges to infinity.

These potential arbitrages lead to the problem that people who want to automatically optimize their strategy against this market model may face. If we keep these fast oscillating noise factors, this model suggests that the wealth we can potentially gain in a fixed time interval from trading these noise factors as a linear combination of assets will be much larger than the wealth we can potentially gain from other parts of the market. Hence, if we simply try a large number of trading strategies and choose the optimal one, or if we use reinforcement learning to optimize our strategy, we may focus too much on optimizing the wealth gained from this fast oscillating behaviour and neglect the wealth gained from other parts of the market. What makes this situation worse is that this maximised expected utility and related trading strategy may not be reliable when applied in reality. On the one hand, a small error in these noise factors will make a large difference in the maximised expected utility, and since these noise factors are usually small in magnitude, they are more easily polluted by errors during observation. On the other hand, the optimal trading strategy we obtained in such a market will also be fast oscillating, which is almost impossible to apply in a real market.

There are various ways of addressing this issue. One may neglect the additional noise factors and just consider the (lower-dimensional non-linear) statistical model. However, this approach may also cause problems. As we have mentioned before, this lower-dimensional market model highly likely contains an arbitrage, and hence this model suggests that the maximised expected utility will diverge.

There are a few reasons why we cannot always choose this market model to be arbitrage free.

One reason is that the space of arbitrage free models is much smaller than a general one, so we need to apply sufficient restrictions on the model to grantee an arbitrage free property.

Such restrictions may not be applicable or may require a very high computational cost to solve a constrained optimization problem.

In addition, the statistical accuracy of such an approximation may highly depend on the choice of the lower-dimensional submanifold and is usually lower than that of the arbitrage-containing ones. For example, in our previous example, we can also approximate our market using a Black–Scholes model. In particular, we can approximate our data using an European call option with the same strike and time to maturity in a Black Scholes model with 0 risk free interest rate (here we choose the drift of the stock to be 0 for simplicity). The volatility of the Black–Scholes model is calibrated to be the volatility that minimizes the sum of squared errors between the real European Call option price and approximated European Call option price at each time point. Figure 1.5 gives an example of such an approximation. However, comparing Figure 1.3 and Figure 1.6, we can see that this arbitrage free model gives us a worse approximation of the real market than the simple polynomial model even in the case where the real market follows a Heston model. In addition, in reality, we usually need to find the arbitrage-free price by solving a PDE numerically, and the error of this approximation may lead to an arbitrage in our model. As a result, we cannot always use an arbitrage free model in reality due to computational cost or accuracy issues.



Figure 1.5: The approximation of the price of the European call option using the value with same strike $K = 3.2$ and maturity time $T = 1.5$ in the Black Scholes model with volatility $\sigma = 0.0346$, which is chosen to minimise the sum of square error.

Hence, we can see that regardless of whether we keep the noise factor in our statistical model or not, the model may suggest that there is a trading strategy that leads to exploding

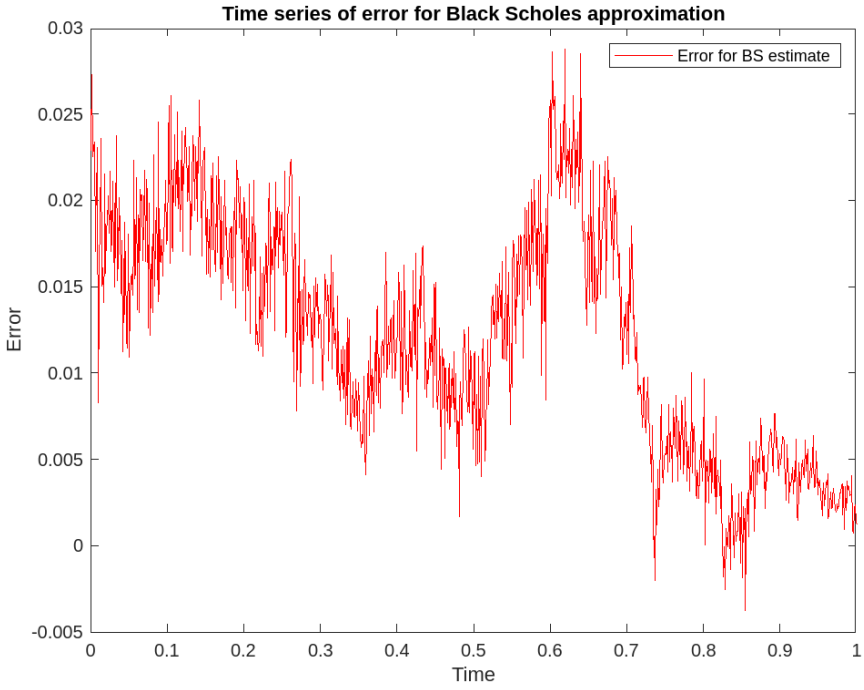


Figure 1.6: A time series of error of approximation of the price of the European call option in Figure 1.5

wealth. However, a person cannot make such amount of wealth by using this trading strategy to trade in the real market. This is because the real market does not contain an arbitrage opportunity and the fast oscillating noise factor. In fact, in our previous Heston model, the fast oscillation behaviour of the noise factor is highly related to the regression-based nature of our approximated model. This implies that an optimal trading strategy obtained from a statistical model through classical methods may not be directly applicable in the real market. However, we still want to use a statistical model to approximate the real market because of its simplicity. Hence, we try to keep the market model simple, but by changing the way we use it, we can deduce a good trading strategy that can be used in the real market. In particular, we want to keep noise factors in the statistical model and find an approach to avoid trying to gain wealth through these unrealistic potential arbitrage opportunities when training the program using such a model. The first step in doing so is to understand the impact of these (asymptotic) arbitrages on the maximised expected utility in such a model. In particular, how will it diverge as the rate of mean reversion of these noise factors diverges to infinity and the magnitude of noise factors converges to 0. This will be the main research purpose of this paper. We will also briefly discuss the possibility of using this problematic market model to obtain a good trading strategy that can be applied in the real market that maximises the expected utility of the wealth process in some sense and is robust to arbitrage-inducing perturbations of a model at the end of this paper.

1.3 Related Literature

As we will see in this paper, our problem is highly related to long-horizon stochastic control, singular perturbation problems, and homogenization problems. We now review the previous literature on these topics.

Consider the simplest version of our problem, where our approximated market model contains a zero-interest bank account, an asset with constant price, and a fast mean reverting noise factor. Using a simple change of variable method, our problem can be transformed into a long time maximised expected utility problem. There are numerous papers related to this problem that can be roughly divided into two directions.

One direction is to seek an asymptotically optimal control process and only focus on the leading order term of the maximised expected utility as time diverges to infinity. Fleming and Sheu ([11] and [10]) dealt with risk-sensitive control, which focuses on maximisation of the expectation of the average per unit time utility of the wealth process for HARA utility functions. Guasoni et al. ([18] and [17]) found the asymptotic optimal control and the leading terms of the related expected utility in a mean-reverting market for the HARA utility function and the CARA utility function, respectively. It is worth noting that, by [17], we expect that the asymptotically optimal expected utility has the leading order e^{T^2} for the CARA utility function instead of e^T for the power utility functions.

Another direction which is more relevant to our research is to seek the optimal control and to give an asymptotic expansion of (the logarithm of) the maximised expected utility as time diverges to infinity. Most of the papers in this direction dealt with problems with periodic assumptions. Namah, Roquejoffre [26] studied the asymptotic behaviour of solutions of parabolic PDEs on the torus. Barles and Souganidis [5] and [6] considered the long time behaviour of space-time periodic solutions of quasi-linear parabolic equations. Namah and Roquejoffre [25] also considered the long time behaviour of solutions to second order PDE with periodic parameters in finite dimension, but they also provided an exponential rate of convergence. Hu, Madec and Richou [20] considered long time behaviour of solutions of the Cauchy problem in an infinite-dimensional real Hilbert space via a probabilistic approach and also got an exponential rate of convergence. Hu and Madec [19] also studied the long time behaviour of viscosity solutions of parabolic equations with Neumann boundary conditions in some bounded domain. Ichihara and Sheu [21] studied the long time behaviour of solutions to the Cauchy problem for semilinear parabolic equations having quadratic nonlinearity in gradients.

In the above papers, the expression of order 1 terms of expansion is usually given with an additional unknown constant, since the solution to the related ergodic PDE is unique up

to the addition of constant. However, this constant is important in our context, especially when the non-fast mean reverting asset we consider does not have constant price. In this case, the additional constant will become a function of the initial price of the non-fast mean reverting asset. Hence, it is crucial for us to find this constant. In order to give an approximation of this constant, we need to use a singular perturbation method. There are also a wide variety of papers about singular perturbation. In papers related to stochastic control theory, singular perturbation problems usually correspond to trading in a market with fast mean reverting stochastic volatility. Fouque and Sircar et al. (see, eg. [14], [13] or [12]) discussed HJB equations with fast mean reverting stochastic volatility and gave the first few terms of an asymptotic expansion for both maximised expected utility and optimal control. Bardi, Cesaroni and Scotti [4] also considered a stochastic volatility model with a fast mean reverting process driven by a pure jump Levy process. However, one of the main differences between our model and a fast mean reverting volatility model is that our noise factor may depend on the asset price. Thus, we need to focus on a singular perturbation problem for fully dependent stochastic systems and may need to invoke homogenization theory. Khasminskii and Yin [23] and [22] gave a full asymptotic expansion of the solution to Cauchy problems under sufficient smoothness assumptions, which is related to maximisation of expected logarithmic utility in a market with fast mean reverting volatility dependent on stock price. In addition, Alvarez and Bardi et al. [2] and [3] showed that the solution to the HJB equation related to a multi-scale stochastic system will converge to the solution to an effective HJB equation that only depends on the slow process.

1.4 Outline of the Thesis

In this thesis, we mainly consider the market with one fast mean reverting asset for simplicity. Hence we only need to find long time behaviour of some specific HJB equations. This paper will mainly depend on [15] by Fujita, Ishii and Loreti, which studied long time behaviour of Cauchy problems for semilinear parabolic equations with the Ornstein–Uhlenbeck operator. Assuming the lower-dimensional market model is already arbitrage free, the decision rule we want to choose for the real market should completely neglect the noise factors and do nothing to our wealth. This means that all terms except for $\log(x)$ in the maximised expected utility are the terms we want to remove. Although this problem itself is not very interesting, it provides a basis for extending it to more general cases. The proofs of this problem also give us a general view about how much we expect to gain if we try to trade noise factors and how this will explode. We will also extend these results to the more general cases in Chapter 5.

This paper contributes to this field of research in several ways.

- First, the main technical contribution is the asymptotic behaviour of the maximised expected utility of a market containing both fast mean reverting and slow assets, with a clear separation of wealth from the fast mean reverting asset and the slow asset in the maximised expected utility. One of the major differences between this problem and existing papers is that we consider a fast mean reverting asset instead of fast mean reverting volatility, so we need to find the wealth we gain by trading this asset. Another difference is that our market contains both fast mean reverting and normal assets, so that our problem cannot be easily transformed into a long time horizon problem.
- Second, we also consider the case where the fast mean reverting asset starts from its invariant distribution instead of a deterministic value, by considering the probabilistic representation for the maximised expected utility.
- Third, the effect of the arbitrage due to non-linearity of our approximated model on the maximised expected utility, and how this is addressed by adding additional noise factors, is intuitively discussed.
- Finally, we will discuss a possible trading strategy that optimizes the maximised expected utility only subject to the reliable part of the market.

The remainder of this paper is structured as follows. In Chapter 2 we will briefly consider the problem with only one fast mean reverting asset with a power utility function $u(x) = \frac{x^p}{p}$. We will look for conditions under which the results of [21] can be applied. In Chapter 3, we will focus on the logarithmic utility function $u = \log(x)$ using a probabilistic approach. This will also give us the existence of a classical solution to this problem and its asymptotic limit. In Chapter 4 we will again consider $u = \log(x)$ but consider the viscosity solution to the related PDE. We will show that this solution will coincide with our classical solution and arrive at a more precise result of the asymptotic limit. We will also consider the case where the asset starts from a non-deterministic initial condition. In Chapter 5 we will discuss possible extensions of our current result to multi-scale problems in both linear and non-linear cases. We also give a formal full asymptotic expansion for maximised expected utility in a specific market with a non-fast mean reverting stochastic asset and a fast mean reverting asset. In addition, we will also briefly discuss the choice of a suitable trading strategy.

1.5 General Formulation of the Problem

In this section, we will only give a general formulation of the problem, the choices of parameters, utility function and set of admissible control will be specified in later chapters.

In this paper, we will mainly consider the simplest case of our problem. In particular, we assume that our real market contains a risk free bank account B_t with 0 interest rate and two risky assets: S_t^1 follows a time homogeneous Itô process and $\hat{S}_t^2$ has dynamics $\hat{S}_t^2 = cS_t^1 + S_t^2$ where c is a deterministic constant and S_t^2 denotes the time homogeneous fast oscillating and mean reverting error term whose dynamics will be specified below. To simplify the problem, we assume that our market is complete and frictionless. Then using a linear transformation, trading with this market model is equivalent to trading in a market with risk free bank account B_t with 0 interest rate, stock S_t^1, and fast mean reverting asset S_t^2. For simplicity, we do not assume that the prices of the assets are strictly positive. In particular, we do not expect that it has the dynamics of a geometric Brownian motion. In addition, it is clear that in this simplified case the scaling of the noise factor S_t^2, which is denoted as δ in our previous Heston example, is not important. For different scaling of noise factors, we can always adjust the control so that the optimal wealth process remains the same. Hence we drop the scaling of noise factor in this linear model as well. However, we will see that this scaling is important in a non-linear market in Chapter 5. We will focus on the asymptotic behaviour of maximised expected utility with HARA utility functions, that is, power utility and logarithmic utility, of trading in this market at some deterministic terminal time T.

In Chapters 2, 3, and 4, we consider a further simplified model. We will assume that S_t^1 is a constant for all t. In this case our problem can be simplified to the optimal problem for trading in a market with a constant bank account B_t and an asset S_t with fast mean reverting price process. Then we can easily extend this result to the case that S_t^1 is random but is independent of S_t^2.

In what follows, we consider a reference Itô process with dynamics:

$$dS_t = \mu(S_t)\,dt + \sigma(S_t)\,dW_t. \tag{1.1}$$

We assume that our market contains an asset whose price S_t^ϵ follows the the following sped-up version:

$$dS_t^\epsilon = \frac{1}{\epsilon}\mu(S_t^\epsilon)\,dt + \frac{1}{\sqrt{\epsilon}}\sigma(S_t^\epsilon)\,dW_t$$

and a constant bank account

$$B_t = 1.$$

Here we choose $\mu : \mathbb{R} \to \mathbb{R}$ and $\sigma : \mathbb{R} \to \mathbb{R}$ such that S_t is mean reverting. As both S_t and S_t^ϵ are time-homogeneous, if we assume that $S_0 = S_0^\epsilon$ then $S_{\frac{t}{\epsilon}}$ and S_t^ϵ have the same distribution.

We consider the wealth process X_t^π related to S_t^ϵ. We set π_t as the number of shares of S_t^ϵ that we hold at time t. We assume that our wealth process X_t^π is self financing, with dynamics

$$dX_t^\pi = \pi_t \frac{1}{\epsilon}\mu(S_t^\epsilon)\, dt + \pi_t \frac{1}{\sqrt{\epsilon}}\sigma(S_t^\epsilon)\, dW_t. \tag{1.2}$$

We want to find

$$V^\epsilon(x, s, t) := \sup_{\pi \in \mathcal{A}} \mathbb{E}[u(X_t^\pi) | X_0 = x, S_0^\epsilon = s]. \tag{1.3}$$

such that $\mathcal{A}$ is the space of admissible processes and u is the utility function representing the risk preference of an investor. Note that, since we assumed that μ and σ are time independent, this problem is a time-reversed version of the classical utility maximisation problem. In this paper, we will use a well-established methodology to approach this problem. As we search for an explicit form of the maximised expected utility, we choose to first fix an HARA utility function that describes the agent's preferences. Abusing the notation π, we construct the related Hamilton-Jacobi-Bellman (HJB) equation, which has the form:

$$\begin{aligned} &- \frac{\partial V}{\partial t} + \frac{1}{\epsilon}\mu(s)\frac{\partial V}{\partial s} + \frac{1}{2\epsilon}\sigma(s)^2\frac{\partial^2 V}{\partial s^2} \\ &+ \frac{1}{\epsilon}\sup_{\pi \in \mathbb{R}} \left\{ \mu(s)\pi\frac{\partial V}{\partial x} + \frac{1}{2}\sigma(s)^2\pi^2\frac{\partial^2 V}{\partial x^2} + \sigma(s)^2\pi\frac{\partial^2 V}{\partial x \partial s} \right\} = 0 \end{aligned} \tag{1.4}$$

with initial condition $V(x, s, 0) = u(x)$ where $u(x)$ is the utility function we choose. If we choose suitable parameters μ and σ and admissible set $\mathcal{A}$, we can verify that V^ϵ is a solution to (1.4). Hence, our problem becomes investigating the asymptotic behaviour of this solution to (1.4) as ϵ converges to 0.

Using a simple change of variable of time t our asymptotic problem is equivalent to the

problem of long-time behaviour of the solution to the HJB equation

$$
\begin{aligned}
&-\frac{\partial V}{\partial t}+\mu(s)\frac{\partial V}{\partial s}+\frac{1}{2}\sigma(s)^2\frac{\partial^2 V}{\partial s^2}\\
&+\sup_{\pi\in\mathbb{R}}\left\{\mu(s)\pi\frac{\partial V}{\partial x}+\frac{1}{2}\sigma(s)^2\pi^2\frac{\partial^2 V}{\partial x^2}+\sigma(s)^2\pi\frac{\partial^2 V}{\partial x\partial s}\right\}=0.
\end{aligned}
\tag{1.5}
$$

with initial condition $V(x,s,0)=u(x)$. In particular, if V^ϵ is a solution solution to (1.4), then $V(x,s,t):=V^\epsilon(x,s,\epsilon t)$ is a solution to (1.5).

In addition, (1.5) is the HJB equation corresponding to the stochastic control problem in a market containing assets

$$dS_t=\mu(S_t)\,dt+\sigma(S_t)\,dW_t \tag{1.6}$$

$$B_t=1. \tag{1.7}$$

where the self financing wealth process X_t^π has dynamics

$$dX_t^\pi=\pi_t\mu(S_t)\,dt+\pi_t\sigma(S_t)\,dW_t. \tag{1.8}$$

and

$$V(x,s,t):=\sup_{\pi\in\mathcal{A}}\mathbb{E}[u(X_t^\pi)|X_0=x,S_0=s]. \tag{1.9}$$

If we can verify that (1.9) is a solution for (1.5), then (1.3) is also a solution to (1.4).

If we also assume that V is twice differentiable with respect to x and that $\frac{\partial^2 V}{\partial x^2}$ is always strictly negative, we can use the optimal feedback strategy

$$\pi(x,s,t)=-\frac{\mu(s)\frac{\partial V}{\partial x}+\sigma(s)^2\frac{\partial^2 V}{\partial s\partial x}}{\sigma(s)^2\frac{\partial^2 V}{\partial x^2}}.$$

Substituting it back into the HJB equation (1.5), we obtain

$$-\frac{\partial V}{\partial t}+\mu(s)\frac{\partial V}{\partial s}+\frac{1}{2}\sigma(s)^2\frac{\partial^2 V}{\partial s^2}-\frac{1}{2}\frac{(\mu(s)\frac{\partial V}{\partial x}+\sigma(s)^2\frac{\partial^2 V}{\partial s\partial x})^2}{\sigma(s)^2\frac{\partial^2 V}{\partial x^2}}=0. \tag{1.10}$$

We want to find the asymptotic behaviour of (1.4) with initial condition

$$V(x,s,0)=u(x)$$

as $\epsilon\to 0$ which is equivalent to the limiting behaviour of (1.5) as t diverges to ∞. As we

shall see, this limit is highly related to the solution to a specific kind of PDE called ergodic PDE. Note that in general case if our market contains both a fast mean reverting asset and a non-fast mean reverting asset, a simple changing of variable method no longer works as the HJB equation after transform contains order ϵ terms due to non-fast mean reverting process and more advanced setting and methods will be introduced in Chapter 5.

Chapter 2

Maximisation of Expected Power Utility of Wealth Process

In this chapter, we consider the power utility $u(x) = \frac{x^p}{p}$ with $0 < p < 1$. We will only give a formal argument instead of a detailed proof of the asymptotic behaviour of the maximised expected utility. This is because the method used in this chapter cannot easily be extended to the multidimensional case, due to the nonlinearity of the HJB equation. However, this chapter will give us some intuition about why the maximised expected utility will diverge and what limit we expect the maximised expected utility to diverge to for a general HARA utility function.

Remark that the dynamics of assets in the market is given by

$$dS_t = \mu(S_t)\,dt + \sigma(S_t)\,dW_t \tag{2.1}$$

$$B_t = 1. \tag{2.2}$$

To begin with, we make the following hypothesis on parameters μ and σ

Hypothesis 2.0.1. μ is twice differentiable, globally Lipschitz continuous and asymptotically equal to a decreasing linear function $-as$ where $a > 0$, in particular, $\lim_{|s|\to\infty}(\mu(s) + as) = 0$. In addition, $\lim_{|s|\to\infty} \frac{d\mu(s)}{ds} = -a$, and $\sigma > 0$ is a constant.

In addition, we say a control process $\pi = (\pi_t)_{t>0}$ is admissible if (1.2) has a unique strong solution X_t^π and this solution satisfies

$$\mathbb{E}[\sup_{0<t<T} u(X_t^\pi)] < \infty \quad \forall T.$$

We denote the set of admissible controls π by $\mathcal{A}$. Then (1.5) can be written as

$$-\frac{\partial V}{\partial t}+\mu(s)\frac{\partial V}{\partial s}+\frac{1}{2}\sigma^2\frac{\partial^2 V}{\partial s^2}+\sup_{\pi\in\mathbb{R}}\left\{\mu(s)\pi\frac{\partial V}{\partial x}+\frac{1}{2}\sigma^2\pi^2\frac{\partial^2 V}{\partial x^2}+\sigma^2\pi\frac{\partial^2 V}{\partial x\partial s}\right\}=0. \tag{2.3}$$

We make a further assumption on its solution

Hypothesis 2.0.2. (2.3) has a solution $V\in C^{2,2,1}(\mathbb{R}^+\times\mathbb{R}\times\mathbb{R}^+)$ and the optimal feedback control $\pi_t:=\pi(X_t^\pi,S_t,t)$ where

$$\pi(x,s,t)=-\frac{\mu(s)\frac{\partial V}{\partial x}+\sigma^2\frac{\partial^2 V}{\partial s\partial x}}{\sigma^2\frac{\partial^2 V}{\partial x^2}}. \tag{2.4}$$

is admissible.

Here we will only state the HJB equations, without showing the detailed verification, as similar types of problem have been previously studied in well established literature (see, for example, Chapter 3 of [27]). In particular, we can deduce that this solution V to (2.3) coincides with $\sup_{\pi\in\mathcal{A}}\mathbb{E}[u(X_t^\pi)|X_0=x,S_0=s]$.

The proof in this chapter is intuitively motivated by the following example of the martingale distortion transformation idea, in [13].

Example 2.0.3. *Guess that V has the form $V(x,s,t)=\frac{x^p}{p}(f(s,t))^q$ where q is a constant to be determined. Substituting this and (2.4) into (2.3), we obtain the following PDE for f*

$$\begin{aligned}&-\frac{\partial f}{\partial t}+\mu(s)\frac{\partial f}{\partial s}+\sigma^2\frac{\partial^2 f}{\partial s^2}-\frac{p}{p-1}\gamma(s)\frac{\partial f}{\partial s}-\frac{p}{2q(p-1)}\gamma(s)^2 f\\&+\frac{1}{2}\sigma^2 f^{-1}\left(\frac{\partial f}{\partial s}\right)^2(q-1-\frac{pq}{p-1})=0\end{aligned} \tag{2.5}$$

with initial condition $f(s,0)=1$ where $\gamma(s):=\frac{\mu(s)}{\sigma}$. Choosing $q=1-p$ such that the last term equals 0, (2.5) becomes

$$-\frac{\partial f}{\partial t}+\mu(s)\frac{\partial f}{\partial s}+\sigma^2\frac{\partial^2 f}{\partial s^2}-\frac{p}{p-1}\gamma(s)\frac{\partial f}{\partial s}+\frac{p}{2(p-1)^2}\gamma(s)^2 f=0.$$

This is a linear second-order PDE. Hence, if we assume that $\gamma(s)^2$ is uniformly bounded by $\log(s)$, we apply can apply the Feynman–Kac formula with the Girsanov transformation

to obtain a probabilistic representation of f:

$$f(s,t) = \tilde{\mathbb{E}}\left[e^{\frac{p}{2(p-1)^2}\int_0^t \gamma(S_u)^2 du} | S_0 = s\right].$$

Here $\tilde{\mathbb{E}}$ is the expectation with respect to the probabilistic measure $\tilde{\mathbb{P}}$ under which $\tilde{W}_t = W_t + \int_0^t \frac{p}{p-1}\gamma(S_u)du$ is a Brownian motion.

This result suggests that it is a reasonable assumption that $V(x,s,t) = \frac{x^p}{p}\tilde{f}(s,t)$ for some strictly positive $\tilde{f}$. This result could easily be extended to the case that $\gamma(s)^2 \leqslant \log(s)$ does not necessarily hold but $\gamma(s)$ satisfies some integrability condition using the localization argument. This separation can also be intuitively deduced from the fact that u is homogeneous and that X_t and π_t appear linearly in the dynamics of X_t. Hence, we make the following assumption:

Hypothesis 2.0.4. $\hat{V}(x,s,t) := \log(u^{-1}(V(x,s,t)))$ is well defined and equals $\log(x) + g(s,t)$ for some function $g \in C^{2,1}(\mathbb{R}\times\mathbb{R}^+)$ where u^{-1} denotes the inverse of the function u.

Substituting $\hat{V}$ and the optimal feedback control (2.4) into (2.3) and using chain rule we obtain the following PDE:

$$\begin{aligned} &-\frac{\partial \hat{V}}{\partial t} + \mu(s)\frac{\partial \hat{V}}{\partial s} + \frac{1}{2}\sigma^2\frac{\partial^2 \hat{V}}{\partial s^2} + \frac{p}{2}\sigma^2\left(\frac{\partial \hat{V}}{\partial s}\right)^2 \\ &-\frac{1}{2}\frac{(\mu(s)\frac{\partial \hat{V}}{\partial x} + \sigma^2(\frac{\partial^2 \hat{V}}{\partial s \partial x} + p\frac{\partial \hat{V}}{\partial x}\frac{\partial \hat{V}}{\partial s}))^2}{\sigma^2\left(\frac{\partial^2 \hat{V}}{\partial x^2} + p(\frac{\partial \hat{V}}{\partial x})^2\right)} = 0, \end{aligned} \tag{2.6}$$

with initial condition $\hat{V}(x,s,0) = \log(x)$.

Without loss of generality, we assume that $\sigma = 1$. Then substitute $\hat{V}(x,s,t) = \log(x) + g(s,t)$ into (2.6), we obtain the following PDE for $\hat{g}(s,t) := -g(s,t)$

$$\frac{\partial \hat{g}}{\partial t} - \frac{1}{2}\frac{\partial^2 \hat{g}}{\partial s^2} + H\left(s, \frac{\partial \hat{g}}{\partial s}\right) = 0, \tag{2.7}$$

with $\hat{g}(s,0) = 0$ where

$$H(s,d) = \frac{p}{2(1-p)}d^2 - \left(1 + \frac{p}{1-p}\right)\mu(s)d - \frac{1}{2(1-p)}\mu(s)^2 \tag{2.8}$$

is a quadratic polynomial both in variable s and in gradient $\frac{\partial \hat{g}}{\partial s}$.

We denote by C_p^k the set of functions whose derivatives are polynomial growing up to the

order k derivative. In particular, $f \in C_p^k$ if there exists $C, m > 0$ such that

$$\left| \frac{d^\kappa f}{ds^\kappa} \right| \leqslant C(1 + s^m) \quad \forall 0 \leqslant \kappa \leqslant k$$

Then we have the following Lemma:

Lemma 2.0.5. *Given Hypothesis 2.0.1 and H as in (2.8), the following condition holds for sufficiently small p such that $p > 0$.*

(A1): $H \in C_p^2(\mathbb{R}^2)$ and there exists a function $f \in C_p(\mathbb{R}) := C_p^0(\mathbb{R})$ such that

$$\left| \frac{\partial H(s,d)}{\partial s} \right| \leq f(s)\left(1 + d^2\right), \quad (s,d) \in \mathbb{R}^2,$$

(A2): there exist constants $b_1, b_2 > 0$ such that

$$b_1 \leq \frac{\partial^2 H(s,d)}{\partial d^2} \leq b_2, \quad (s,d) \in \mathbb{R}^2,$$

(A3)*: there exist $C_0 > 0$ and $c_0 \in \mathbb{R}$ such that*

$$|H(s,d)| \leq C_0\left(1 + s^2 + d^2\right), \qquad \limsup_{|s|\to\infty} \frac{H(s, c_0 s)}{1 + s^2} < 0.$$

(A4): the initial condition $\hat{g}(s,0) = 0$ is in the set

$$\Phi_0 := \left\{ v \in C_p(\mathbb{R}) \mid \inf_{\mathbb{R}} \left(v - \frac{c_0}{2} s^2 \right) > -\infty \right\} \tag{2.9}$$

where $C_p(\mathbb{R})$ denotes the set of polynomial growing functions on $\mathbb{R}$.

Proof. It is clear that (A1) holds with any $0 < p < 1$ as both $\mu(s)$ and d have quadratic growth in H and the derivative of μ is continuous and asymptotically equals to $-a$. (A2) is also clearly satisfied. The first condition for (A3)' holds as $x^2 + y^2 \geq 2xy$. Now for the last condition, as $\lim_{|s|\to\infty}(\mu(s) + as) = 0$, we only need that

$$\frac{p}{2(1-p)} c_0^2 + \left(1 + \frac{p}{1-p}\right) a c_0 - \frac{1}{2(1-p)} a^2 = 0$$

has 2 different real solutions and choose c_0 to be any value between these solutions, i.e. the determinant of this quadratic polynomial should be positive. This can be satisfied for sufficiently small p as the determinant equals 1 when $p = 0$ and this function is continuous

with respect to p around 0. Also note that as long as this determinant is non-negative for some $0 < p < 1$ and a, it is non-negative for the same p and any a.

Finally, in order to satisfy (A4), c_0 in (A3) needs to be negative. This can be satisfied by choosing $\mu(s)$ asymptotically equal to $-as$ for some $a > 0$. □

If (A1)-(A4) are satisfied, by direct application of the main result in [21] we obtain the following.

Theorem 2.0.6. *Given Hypothesis 2.0.1 for sufficiently small p depending on a, (2.7) has a unique global solution $\hat{g}(s,t)$ which is differentiable with polynomial growth. In addition, the following 'ergodic' PDE:*

$$-\frac{1}{2}\frac{\partial^2\phi}{\partial s^2} + H\left(s, \frac{\partial\phi}{\partial s}\right) = \lambda \tag{2.10}$$

with H defined in (2.8) also has a unique solution $(\lambda, \phi(s))$, up to addition of a constant in ϕ such that $\phi \in \Phi_0$. Furthermore, there exists $c \in \mathbb{R}$ such that

$$\lim_{t\to\infty} \hat{g}(s,t) - (\phi(s) - \lambda t) \longrightarrow c. \tag{2.11}$$

This convergence holds uniformly on any compact set of $\mathbb{R}$.

Remark 2.0.7. The condition that p is sufficiently small means that the risk aversion of the trader is sufficiently close to the risk aversion of a trader with logarithmic utility. This suggests that we expect a similar convergence result for the logarithmic utility function. As we will show in the following two chapters, the convergence results for the logarithmic utility function will still hold under weaker conditions.

Finally, we can substitute $\hat{g}$ into V and consider the behaviour of $\sup_{\pi\in\mathcal{A}} \mathbb{E}[u(X_t^\pi)|X_0 = x, S_0^\epsilon = s] =: V^\epsilon(x,s,t) = V(x,s,\frac{t}{\epsilon})$ which is a solution to (1.4). We obtain:

$$\lim_{\epsilon\to 0} \log\left(V^\epsilon(x,s,t)) - \lambda p\frac{t}{\epsilon}\right) = p\log(x) - \log(p) - p(c + \phi(s)).$$

This expression suggests that we can separate the logarithm of the maximised expected utility of wealth into the logarithm of the utility of initial wealth $p\log(x) - \log(p)$, long-term asymptotic income $\lambda p\frac{t}{\epsilon}$ independent of the initial price of the asset S_t^ϵ, and income $-p(c + \phi(s))$ which can be regarded as the correction term depending on the initial price of the asset.

As we have claimed before, this result is quite difficult to generalize to the case if we add another slow asset which is correlated to it into our market. In addition, the probabilistic

representations for λ and c are also quite difficult to deduce.

Chapter 3

Maximisation of Expected Logarithmic Utility via Probabilistic Approach

In this and the next chapter we will focus on the case $u(x) = \log(x)$, which can be regarded informally as the special case $p = 0$ of power utilities. That is, we want to find $V(x, s, t) := \sup_\pi \mathbb{E}[\log(X_t^\pi)|X_0 = x, S_0 = s]$.

3.1 Setting of the Problem

We recall that the dynamics of assets in the market is given by

$$dS_t = \mu(S_t)\,dt + \sigma(S_t)\,dW_t \tag{3.1}$$

$$B_t = 1. \tag{3.2}$$

As before, we first state the hypothesis for the parameters and the set of admissible controls.

Hypothesis 3.1.1. $\mu(s)$ is globally Lipschitz continuous and monotonically decreasing, In addition, there exist $a > 0$ such that

$$\lim_{|s|\to\infty} \frac{\mu(s)}{s} = -a$$

and $\sigma > 0$ is constant.

We consider the self financing wealth process defined as

$$dX_t^\pi = \pi_t \, dS_t \tag{3.3}$$

where π_t denotes the number of shares of asset S_t we hold at time t.

We say a control process π is an admissible control if (3.3) has a unique strong solution X_t^π which is strictly positive and

$$\mathbb{E}[\sup_{0<t<T} \log(X_t^\pi)] < \infty \quad \forall T.$$

We will denote the set of all admissible controls by $\mathcal{A}$. Abusing the notation of π, the HJB equation related to our stochastic control problem is

$$-\frac{\partial V}{\partial t} + \mu(s)\frac{\partial V}{\partial s} + \frac{1}{2}\sigma^2\frac{\partial^2 V}{\partial s^2} + \sup_{\pi\in\mathbb{R}}\left\{\mu(s)\pi\frac{\partial V}{\partial x} + \frac{1}{2}\sigma^2\pi^2\frac{\partial^2 V}{\partial x^2} + \sigma^2\pi\frac{\partial^2 V}{\partial x \partial s}\right\} = 0 \tag{3.4}$$

with initial condition $V(x, s, 0) = \log(x)$.

Motivated by the result of Chapter 2 we make the following assumption on the solution to (3.4):

Hypothesis 3.1.2. There exists a function $V \in C^{2,2,1}(\mathbb{R}^+ \times \mathbb{R} \times \mathbb{R}^+)$ that satisfies (3.4) with the initial condition $V(x, s, 0) = \log(x)$. Furthermore, $V(x, s, t) = \log(x) + g(s, t)$ for a function $g \in C^{2,1}(\mathbb{R} \times \mathbb{R}^+)$ that satisfies the quadratic growth condition, that is, $\forall t > 0$ there exists a constant C such that

$$|g(s, t)| < C(1 + s^2)$$

We first sketch a proof to verify that the function V in Hypothesis 3.1.2 satisfies

$$V(x, s, t) = \sup_{\pi\in\mathcal{A}} \mathbb{E}[\log(X_t^\pi)|X_0 = x, S_0 = s].$$

Using the assumption $V(x, s, t) = \log(x) + g(s, t)$ and follows a proof similar to Theorem 4.2 in [1], we can show that the corresponding optimal feedback control $\hat{\pi}$ defined as

$$\hat{\pi}(x, s, t) = -\frac{\mu(s)\frac{\partial V}{\partial x} + \sigma^2\frac{\partial^2 V}{\partial x \partial s}}{\sigma^2\frac{\partial^2 V}{\partial x^2}} \tag{3.5}$$

is admissible. Hence, by Hypothesis 3.1.2 we can apply the Itô's formula to V. As for

arbitrary admissible control process π we have

$$-\frac{\partial V}{\partial t}+\mu(S_t)\frac{\partial V}{\partial s}+\frac{1}{2}\sigma^2\frac{\partial^2 V}{\partial s^2}+\mu(S_t)\pi_t\frac{\partial V}{\partial x}+\frac{1}{2}\sigma^2\pi_t^2\frac{\partial^2 V}{\partial x^2}+\sigma^2\pi_t\frac{\partial^2 V}{\partial x\partial s}\leqslant 0, \tag{3.6}$$

substituting (3.6) into the Itô's formula, taking expectation on both sides and using localisation argument we obtain

$$V(x,s,t)\leqslant \sup_{\pi\in\mathcal{A}}\mathbb{E}[\log(X_t^\pi)|X_0=x,S_0=s].$$

As the optimal feedback control is admissible, repeating this proof with the optimal feedback control we can show that

$$V(x,s,t)\geqslant \sup_{\pi\in\mathcal{A}}\mathbb{E}[\log(X_t^\pi)|X_0=x,S_0=s]$$

which finish the verification.

Substituting $V(x,s,t)=\log(x)+g(s,t)$ and the optimal feedback control (3.5) into (3.4) we obtain the following PDE for g:

$$-\frac{\partial g}{\partial t}+\mu(s)\frac{\partial g}{\partial s}+\frac{1}{2}\sigma^2\frac{\partial^2 g}{\partial s^2}+\frac{1}{2}\frac{\mu(s)^2}{\sigma^2}=0 \tag{3.7}$$

with boundary condition $g(s,0)=0$.

Note that compared with (2.10), (3.7) is linear and easier for us to analyse.

From the above argument, to learn the asymptotic behaviour of $V(x,s,t)$, we just need to learn the asymptotic behaviour of $g(s,t)$, the solution to (3.7). By Hypothesis 3.1.1 and 3.1.2, we can apply the Feynman–Kac formula to (3.7) to obtain a probabilistic representation of g:

$$g(s,t)=\mathbb{E}\left[\int_0^t\frac{\mu(S_u)^2}{2\sigma^2}du|S_0=s\right] \tag{3.8}$$

where S_t has the dynamic (3.1). Now we want to show that (3.8) has simple asymptotic behaviour.

3.2 Basic Tools

Before we state the main result of this chapter, we first introduce basic tools involving Lyapunov functions.

A Lyapunov function is a function $L : \mathbb{R} \to [1, \infty)$ such that $\lim_{x\to\infty} L(x) = \infty$ and there exist real numbers $c_1 \in (0, \infty), c_2 \in (0, \infty)$ such that

$$\mu(s)\frac{\partial L}{\partial s} + \frac{1}{2}\sigma^2\frac{\partial^2 L}{\partial s^2} \leqslant -c_1\{L(s)\} + c_2 \tag{3.9}$$

and

$$\frac{\mu(s)^2}{2\sigma^2} \leqslant L(s), \quad \forall s. \tag{3.10}$$

Lemma 3.2.1. *Under Hypothesis 3.1.1 there exists a function L that satisfies (3.9) and (3.10)*

Proof. As μ is asymptotic linear and $\sigma > 0$ is constant, choosing $c_1 \in (0, 2a)$ and c_2 sufficiently large (3.9) is satisfied for any function $L(s) = c_3 s^2 + c_4$ with $c_3 \in (0, \infty)$ and $c_4 \in \mathbb{R}$. Then choosing $c_3 \in (\frac{a^2}{2\sigma^2}, \infty)$ and c_4 sufficiently large (3.10) is also satisfied. Hence, this $L(s) = c_3 s^2 + c_4$ is the Lyapunov function we require. □

In addition to the existence of a Lyapunov function, we also need to show that the process S_t satisfies the following minorization condition:

For any fixed compact set $\mathcal{K} \in \mathbb{R}$ there exist a choice $T \in \mathbb{R}^+$, a constant $c > 0$ and a probability measure $\mathbb{Q}$, with $\mathbb{Q}(\mathcal{K}) = 1$ and $\mathbb{Q}(\mathcal{K}^c) = 0$, such that

$$P_T(s, A) \geqslant c\mathbb{Q}(A) \quad \forall s \in \mathcal{K} \text{ and Borel set } A. \tag{3.11}$$

By Lemma 2.3 of [24] the following lemma implies the minorization condition we require:

Lemma 3.2.2. *The process S_t with transition probability $P_t(s, A) := P(S_t \in A | S_0 = s)$ satisfies for any non-empty fixed compact set $\mathcal{K}$,*

(i) for some $s^* \in \text{int}(\mathcal{K})$ there is, for any $\delta > 0$, a t_1 depending on δ such that

$$P_{t_1}(s, B_\delta(s^*)) > 0, \quad \forall s \in \mathcal{K},$$

(ii) the transition kernel has a density $p_t(s_1, s_2)$, i.e.

$$P_t(s, A) = \int_A p_t(s, s^*)ds^*, \quad \forall s \in \mathcal{K},$$

which is jointly continuous in $(s_1, s_2) \in \mathcal{K} \times \mathcal{K}$.

Sketch of proof of Lemma 3.2.2. Choosing $\mathcal{K}$ to be any compact interval in $\mathbb{R}$, as μ is global Lipschitz, the process S_t is continuous for any initial point $s \in \mathcal{K}$. Then considering the Fokker–Planck equation, (ii) clearly holds as μ is global Lipschitz and $\sigma > 0$. (i) needs further verification. Here I will only put a sketch of the proof of (i). The idea of proving (i) is to consider the associated deterministic control problem:

$$\frac{ds_t}{dt} = \mu(s_t) + \sigma \frac{df}{dt}.$$

Then, for any $t > 0$, $s \in \mathbb{R}$ and $s^* \in R$, we can find a continuous deterministic function $f : [0, t] \times \mathbb{R} \to \mathbb{R}$ such that $s_0 = s$ and $s_t = s^*$. Then, as the event $\sup_{0<s<t} ||W_s - f(s)|| < \epsilon$ occurs with positive probability for all $\epsilon > 0$, where W_t is the Brownian motion in (1.1), and the fact that μ is globally Lipschitz, we can deduce condition (i) holds. □

Remark 3.2.3. This argument actually proves a stronger result than (i). In particular, we show that

$$P_{t_1}(s, B_\delta(s^*)) > 0, \quad \forall s \in \mathcal{K},$$

holds for any $s^* \in \text{int}(\mathcal{K})$ and $t_1 > 0$.

3.3 Main Result

We can now state the main result of this chapter.

Theorem 3.3.1. *Assuming that Hypothesis 3.1.1 holds and S_t has dynamic (3.1), for all s there exists a constant λ such that $\lim_{t\to\infty} \left| \mathbb{E}\left[\int_0^t \frac{\mu(S_u)^2}{2\sigma^2} du | S_0 = s \right] - \lambda t \right| < \infty$. Furthermore, S_t has a unique invariant distribution Π and $\lambda = \int \frac{\mu(s)^2}{2\sigma^2} d\Pi$.*

Proof of Theorem 3.3.1. We follow the idea of [24]. By Lemma 2.3 of [24], Lemma 3.2.2 and Remark 3.2.3 implies that there exist a time $T > 0$, a constant $c > 0$ and a probability measure $\mathbb{Q}$ with $\mathbb{Q}(\mathcal{K}) = 1$ such that

$$P_T(s, A) \geqslant c\mathbb{Q}(A) \quad \forall s \in \mathcal{K} \text{ and Borel set } A. \tag{3.12}$$

Hence, by Theorem 2.5 of [24] if we sample our process S_t using the sampling rate T, that is, if we consider $s_n = S_{nT}$, the process s_n has a unique stationary distribution Π.

Furthermore, for all measurable $f : |f| \leqslant L$, there is $C > 0$ and $k > 0$ such that

$$\left| \mathbb{E}^s f\left(s_n\right) - \int f d\Pi \right| \leqslant C e^{-kn} L(s).$$

By Lemma 3.2.1, our choice of Lyapunov function L satisfies $\frac{\mu(s)^2}{2\sigma^2} \leqslant L$. Hence, this result holds for $f = \frac{\mu(s)^2}{2\sigma^2}$.

Now we can move back to the continuous process and consider an arbitrary $\epsilon \in (0, T]$. By (3.9) we obtain

$$\left| \mathbb{E}^s[f\left(S_{nT+\epsilon}\right)] - \int f d\Pi \right| \leqslant C e^{-kn} \mathbb{E}^s[L(S_\epsilon)] \leqslant C e^{-kn} \left(e^{-c_1 \epsilon} L(s) + \frac{c_2}{c_1} \right) \leqslant \tilde{C} e^{-\tilde{k}(nT+\epsilon)} L(s).$$

Thus for all $t > 0$

$$\left| \mathbb{E}^s[f\left(S_t\right)] - \int f d\Pi \right| \leqslant \tilde{C} e^{-\tilde{k}t} L(s). \tag{3.13}$$

In addition, substituting (3.13) into (3.8) we obtain:

$$\begin{aligned} \left| \mathbb{E}^s \left[\int_0^t \frac{\mu(S_u)^2}{2\sigma^2} - \int \frac{\mu(s)^2}{2\sigma^2} d\Pi du \right] \right| &\leqslant \int_0^t \left| \mathbb{E}^s \frac{\mu(S_u)^2}{2\sigma^2} - \int \frac{\mu(s)^2}{2\sigma^2} d\Pi \right| du \\ &\leqslant \int_0^t \tilde{C} e^{-\tilde{k}t} L(s) du. \end{aligned}$$

In particular, for fixed s, $g(s,t) - t \int \frac{\mu(s)^2}{2\sigma^2} d\Pi$ is uniformly bounded as $t \to \infty$. □

Remark 3.3.2. As the function L is a quadratic polynomial, this proof also shows that g satisfies the quadratic growth condition.

In this chapter, we have discussed the order $\frac{1}{\epsilon}$ term of the maximised expected utility V^ϵ that is a solution to (1.4) with initial condition $V(x, s, 0) = \log(x)$. We have shown that $\lim_{t\to\infty}(V^\epsilon - \log(x) - \frac{t}{\epsilon} \int \frac{\mu(s)^2}{2\sigma^2} d\Pi) < \infty$. One may be curious whether the expression in the brackets converges and, if so, whether we can find an explicit expression of this limit. In the next chapter, we will show that this will converge to a limit function which is highly related to solutions of ergodic PDE.

Chapter 4

Maximisation of Expected Logarithmic Utility via Viscosity Approach

4.1 Brief Idea

In the previous chapter, we showed that $g(s,t) - t\int \frac{\mu(s)^2}{\sigma^2} d\Pi$ is bounded for fixed s as $t \to \infty$. In this chapter, we will show that it is not only bounded, it converges to a function of s independent of t as $t \to \infty$. In this chapter, we will focus on the viscosity solution of (3.7).

We will first outline the main steps of this approach. As we have shown in the last chapter that

$$g(s,t) = \mathbb{E}\left[\int_0^t \frac{\mu(S_u)^2}{2\sigma^2} du | S_0 = s\right]$$

is a classical solution to (3.7), it is also a viscosity solution to (3.7). Thus, we have the existence of the viscosity solution. The first step is to show the comparison result for this Cauchy problem so that we can then prove the uniqueness of the viscosity solutions to (3.7). After that, we will follow the steps in [15]. We will consider the existence of viscosity solution to the related ergodic problem

$$\lambda - \mu(s)\frac{\partial \phi}{\partial s} - \frac{1}{2}\sigma^2 \frac{\partial^2 \phi}{\partial s^2} - \frac{1}{2}\frac{\mu(s)^2}{\sigma^2} = 0 \tag{4.1}$$

as a limit of viscosity solution to the discounted problem

$$\epsilon\phi^\epsilon - \mu(s)\frac{\partial\phi^\epsilon}{\partial s} - \frac{1}{2}\sigma^2\frac{\partial^2\phi^\epsilon}{\partial s^2} - \frac{1}{2}\frac{\mu(s)^2}{\sigma^2} = 0$$

This will be done by the Perron method given in [9]. We will then look for a uniform bound for $\epsilon\phi^\epsilon$ and $\phi^\epsilon(s_1) - \phi^\epsilon(s_2)$ independent of ϵ so there is a convergent subsequence. After that, we will look for a comparison result for (4.1) to show that any viscosity supersolution and subsolution (λ, ϕ) of (3.7) will agree up to the addition of a constant to ϕ. Finally, we will show that $g(s,t)$ converges to the limit $\lambda t + \phi(s) + c$ as t goes to infinity.

Before we start the proof, we will first show intuitively why this convergence holds. Consider $g_\epsilon(s,t) := g(s, \frac{t}{\epsilon})$. By the chain rule, g is a classical solution to (3.7) if and only if g_ϵ is a classical solution to

$$\frac{\partial g_\epsilon}{\partial t} - \frac{1}{\epsilon}\left(\mu(s)\frac{\partial g_\epsilon}{\partial s} + \frac{1}{2}\sigma^2\frac{\partial^2 g_\epsilon}{\partial s^2} + \frac{1}{2}\frac{\mu(s)^2}{\sigma^2}\right) = 0. \tag{4.2}$$

Gaining intuition from the last chapter, we guess g_ϵ has an expansion:

$$g_\epsilon(s,t) = \frac{1}{\epsilon}g^{(-1)}(s,t) + g^{(0)}(s,t) + O(\epsilon). \tag{4.3}$$

Substituting this expression into (3.7) and comparing the orders of ϵ, we obtain:

For order $\frac{1}{\epsilon^2}$

$$\mu(s)\frac{\partial g^{(-1)}}{\partial s} + \frac{1}{2}\sigma^2\frac{\partial^2 g^{(-1)}}{\partial s^2} = 0. \tag{4.4}$$

For order $\frac{1}{\epsilon}$

$$\frac{\partial g^{(-1)}}{\partial t} = \mu(s)\frac{\partial g^{(0)}}{\partial s} + \frac{1}{2}\sigma^2\frac{\partial^2 g^{(0)}}{\partial s^2} + \frac{1}{2}\frac{\mu(s)^2}{\sigma^2}. \tag{4.5}$$

Now we can see that a solution to (4.4) and (4.5) is $g^{(-1)} = \lambda t$ and $g^{(0)} = \phi(s) + f(t)$ where f is a function to be determined and (λ, ϕ) is a solution to the ergodic PDE:

$$\lambda - \mu(s)\frac{\partial\phi}{\partial s} - \frac{1}{2}\sigma^2\frac{\partial^2\phi}{\partial s^2} - \frac{1}{2}\frac{\mu(s)^2}{\sigma^2} = 0. \tag{4.6}$$

In addition, as $g_\epsilon(s,t) := g(s, \frac{t}{\epsilon})$, we need to restrict that as ϵ converges to 0 the limit should be independent of t, which means that $f(t)$ should be a constant c. Hence, we have informally argued that $g_\epsilon(s,t)$ converges to $\lambda\frac{t}{\epsilon} + \phi(s) + c$ as ϵ goes to 0. This expansion will

be extended to apply to the problem of trading in a market with an additional non-fast mean-reverting stock in chapter 5.

4.2 Existence and Uniqueness of Viscosity Solution to Cauchy Problem

First we need the following hypothesis on parameters. Here we make a slightly stronger assumption than Hypothesis 3.1.1.

Hypothesis 4.2.1. $\mu(s)$ is globally Lipschitz continuous and monotonically decreasing, In addition, there exist $a > 0$ such that

$$\lim_{|s|\to\infty} \frac{\mu(s)}{s} = -a,$$
$$|\mu(s_1) - \mu(s_2)| \geqslant a\,|s_1 - s_2|\,,$$

and $\sigma > 0$ is constant.

Without loss of generality, we choose $\sigma = \sqrt{2}$.

Considering the function $\psi_k(x) := e^{kx^2}$ with $0 < k < \frac{a}{2}$, we will focus on the set of functions:

$$\Psi_k(\mathbb{R}) = \left\{ v : \mathbb{R} \to \mathbb{R} \mid \lim_{|x|\to\infty} \frac{v(x)}{\psi_k(x)} = 0 \right\},$$
$$\Psi_k^+(\mathbb{R}) = \left\{ v : \mathbb{R} \to \mathbb{R} \mid \limsup_{|x|\to\infty} \frac{v(x)}{\psi_k(x)} \leqslant 0 \right\},$$
$$\Psi_k^-(\mathbb{R}) = \left\{ v : \mathbb{R} \to \mathbb{R} \mid \liminf_{|x|\to\infty} \frac{v(x)}{\psi_k(x)} \geqslant 0 \right\},$$
$$\Psi_k(\mathbb{R}_T) = \left\{ v : \mathbb{R} \to \mathbb{R} \mid \lim_{|x|\to\infty} \sup_{0\leqslant t\leqslant T} \frac{v(x,t)}{\psi_k(x)} = 0, \quad \lim_{|x|\to\infty} \inf_{0\leqslant t\leqslant T} \frac{v(x,t)}{\psi_k(x)} = 0 \right\},$$
$$\Psi_k^+(\mathbb{R}_T) = \left\{ v : \mathbb{R} \to \mathbb{R} \mid \limsup_{|x|\to\infty} \sup_{0\leqslant t\leqslant T} \frac{v(x,t)}{\psi_k(x)} \leqslant 0 \right\},$$
$$\Psi_k^-(\mathbb{R}_T) = \left\{ v : \mathbb{R} \to \mathbb{R} \mid \liminf_{|x|\to\infty} \inf_{0\leqslant t\leqslant T} \frac{v(x,t)}{\psi_k(x)} \geqslant 0 \right\}.$$

Observing that

$$\frac{d\psi_k(s)}{ds} = 2k\psi_k(s)s, \quad \frac{d^2\psi_k(s)}{ds^2} = 2k\left(2ks^2 + 1\right)\psi_k(s),$$

if we choose $k < \frac{a}{2}$ we then can find a constant $C > 0$ independent of s such that

$$-\mu(s)\frac{\partial \psi_k(s)}{\partial s} - \frac{\partial^2 \psi_k(s)}{\partial s^2} \geqslant \psi_k(s) - C. \tag{4.7}$$

This property is important throughout this chapter. Using this property, we can extend the comparison result of solutions to (3.7) on $[0,T] \times C$ for a fixed compact subset C of $\mathbb{R}$ to the whole space $[0,T] \times \mathbb{R}$ as follows:

Theorem 4.2.2. *For fixed $T > 0$, if the upper semicontinuous $u \in \Psi_k^+(\mathbb{R}_T)$, lower semicontinuous $v \in \Psi_k^-(\mathbb{R}_T)$ are a viscosity subsolution and viscosity supersolution of the Cauchy problem (3.7) respectively, with $u(s,0) \leqslant v(s,0) \quad \forall s$ then $u \leqslant v$ on $\mathbb{R} \times (0,T]$.*

Proof. Consider

$$u_\epsilon := u - \epsilon(\psi_k + Ct),$$
$$v_\epsilon := v + \epsilon(\psi_k + Ct).$$

From (4.7) and the linearity of (3.7) u_ϵ is still a viscosity supersolution and v_ϵ is still a viscosity subsolution. In addition, there exists some $R > 0$ such that $u_\epsilon \leqslant v_\epsilon$ for all s outside $B_R(0)$ as $u \in \Psi_k^+(\mathbb{R}_T), v \in \Psi_k^-(\mathbb{R}_T)$ which implies $u_\epsilon \leqslant v_\epsilon$ for $s = \pm R$ and $t \in (0,T]$. Now we have reduced our problem to proving the comparison result on $[0,T] \times [-R,R]$ and this is standard in viscosity solution theory (see, e.g. Theorem 8.2 of [9]). Hence, we obtain $u_\epsilon \leqslant v_\epsilon$ on $\mathbb{R} \times (0,T]$. Taking ϵ to 0 shows that the comparison result holds for u and v. □

Remark 4.2.3. (3.13) implies that the classical solution g for the Cauchy problem (3.7) with the boundary condition $g(s,0) = 0$ is in $\Psi_k(\mathbb{R}_T)$. By Proposition 7.2 of [9], any classical solution of this Cauchy problem is also a viscosity solution, so we can apply the comparison result to deduce that the unique viscosity solution in $\Psi_k(\mathbb{R}_T)$ is the classical solution whose probabilistic representation is given by (3.8). This is essential to us, as the viscosity solution to (1.5) may not coincide with $\log(x)$ plus the viscosity solution of (3.7).

4.3 Existence and Uniqueness of Viscosity Solution to Ergodic Problem

In this section we will focus on the ergodic problem (4.1). We will first consider the approximation to (4.1):

$$\epsilon\phi^\epsilon - \mu(s)\frac{\partial\phi^\epsilon}{\partial s} - \frac{\partial^2\phi^\epsilon}{\partial s^2} - \frac{\mu(s)^2}{4} = 0 \tag{4.8}$$

with $\epsilon \in (0,1]$. As before, we will first prove the comparison result for solutions to (4.8).

Theorem 4.3.1. *If upper semicontinuous $u \in \Psi_k^+(\mathbb{R})$, lower semicontinuous $v \in \Psi_k^-(\mathbb{R})$ are viscosity subsolution and viscosity supersolution of (4.8) respectively , then $u \leqslant v$ on $\mathbb{R}$*

Proof. The proof is almost identical to the proof of Theorem 4.2.2. Set

$$u_\delta(s) := u(s) - \delta\left(\psi_k(s) + \frac{1}{\epsilon}C\right),$$
$$v_\delta(s) := v(s) + \delta\left(\psi_k(s) + \frac{1}{\epsilon}C\right),$$

where C is defined in (4.7). From (4.7), u_δ and v_δ are still viscosity supersolution and viscosity subsolution for (4.8) and there exists R such that $u \leqslant v$ outside $B_R(0)$. Applying the standard comparison result on $B_R(0)$ for u_δ and v_δ we obtain $u \leqslant v$ on $\mathbb{R}$. Letting δ go to 0 finishes the proof. □

We will now show the existence of a viscosity solution using the Perron method. We follow the steps in [15]. Observe that there exist a constant C and a function $f : (0,1] \to \mathbb{R}^+$ such that, with $\psi_k(s) = e^{ks^2}$ as above and $\delta \in (0,1]$,

$$-\mu(s)\frac{\partial\psi_k(s)}{\partial s} - \frac{\partial^2\psi_k(s)}{\partial s^2} \geqslant \psi_k(s) - C,$$
$$\frac{\mu(s)^2}{4} \leqslant \delta\psi_k(s) + f(\delta).$$

Now we fix ϵ. For $\delta \in (0,1]$, we consider the function

$$h(s,\delta) := \delta\psi_k(s) + \frac{1}{\epsilon}(f(\delta) + \delta C). \tag{4.9}$$

Substituting this into (4.8) we obtain:

$$\epsilon h - \mu(s)\frac{dh}{ds} - \frac{d^2h}{ds^2} - \frac{\mu(s)^2}{4} \geqslant 0.$$

Thus, we see that h is a viscosity supersolution of (4.8). Defining

$$H(s) := \inf_{\delta\in(0,1]} h(s,\delta)$$

gives an upper semicontinuous function. We now consider its upper and lower semicontinuous envelopes H^* and H_* which are defined as

$$H^*(s) = \limsup_{\hat{s}\to s} H(\hat{s})$$
$$H_*(s) = \liminf_{\hat{s}\to s} H(\hat{s})$$

It is clear that the lower semicontinuous envelope H_* satisfies $H_*(s) < \infty \quad \forall s$. Lemma 4.2 of [9] implies that H_* is a viscosity supersolution to (4.8). For the same reason, we have that $(-H)^*$ is a viscosity subsolution to (4.8). So we can apply the Perron method. Defining

$$\Phi(s) := \sup\{v(s) \quad s.t. (-H)^* \leqslant v(s) \leqslant H_*(s) \text{ and } v(s) \text{ is a viscosity subsolution of (4.8)}\},$$

Its upper and lower semicontinuous envelopes Φ^* and Φ_* are a viscosity subsolution and viscosity supersolution to (4.8) respectively.

From our construction we obtain the following uniform bound for Φ:

$$|\Phi(s)| \leqslant \delta\psi_k(s) + \frac{1}{\epsilon}(f(\delta) + \delta C), \quad \forall s \text{ and } \delta.$$

This shows that Φ^* and Φ_* are in $\Psi_k(\mathbb{R})$. Hence, by Theorem 4.3.1 we have $\Phi^* \leqslant \Phi_*$. From the definition of semicontinuous envelopes we have $\Phi^* \geqslant \Phi_*$. Thus, $\phi^\epsilon := \Phi^* = \Phi_*$ and is a continuous viscosity solution to (4.8) on $\mathbb{R}$.

By taking $\delta = 1$ this construction also provides a uniform bound for $|\epsilon\phi^\epsilon(0)|$:

$$|\epsilon\phi^\epsilon(0)| \leqslant 1 + f(1) + C < \infty.$$

To show the convergence of ϕ^ϵ we still need to find a uniform continuity bound on ϕ^ϵ, which is done in the next lemma.

Lemma 4.3.2. *Let ϕ^ϵ be the unique solution to (4.8) constructed above. There exists some constant L independent of ϵ such that*

$$|\phi^\epsilon(s_1) - \phi^\epsilon(s_2)| \leqslant L\,|s_1 - s_2|\,(\psi_k(s_1) + \psi_k(s_2)).$$

Proof. Notice that there is a constant C_k such that

$$\left|\mu(s_1)^2 - \mu(s_2)^2\right| \leqslant 4C_k \left|s_1 - s_2\right| \left(\psi_k(s_1) + \psi_k(s_2)\right). \tag{4.10}$$

Consider the function

$$h(s_1, s_2) := C_k(\delta + (|s_1 - s_2| + \delta^2)) \left(\psi_k(s_1) + \psi_k(s_2) + \frac{2C}{a}\right)$$

where C comes from (4.7). For simplicity we use the notation:

$$D_2 h := \frac{\partial^2 h}{\partial s_1^2} + \frac{\partial^2 h}{\partial s_2^2} + 2\frac{\partial^2 h}{\partial s_1 \partial s_2}.$$

If $s_1 \neq s_2$, the following inequality is well defined and holds for $\delta \in (0, \frac{\epsilon}{2C}]$.

$$\begin{aligned}
&\epsilon h - \mu(s_1)\frac{\partial h}{\partial s_1} - \mu(s_2)\frac{\partial h}{\partial s_2} - D_2 h \\
&\geqslant \epsilon h + C_k(|s_1 - s_2| + \delta^2) \left(\frac{d^2\psi_k}{ds_1^2}(s_1) + \mu(s_1)\frac{d\psi_k}{ds_1}(s_1) + \frac{d^2\psi_k}{ds_2^2}(s_2) + \mu(s_2)\frac{d\psi_k}{ds_2}(s_2)\right) \\
&\quad + C_k \left|\mu(s_1) - \mu(s_2)\right| \left(\psi_k(s_1) + \psi_k(s_2) + \frac{2C}{a}\right) \\
&\geqslant \epsilon h + C_k(|s_1 - s_2| + \delta^2)(\psi_k(s_1) + \psi_k(s_2) - 2C) + C_k a \left|s_1 - s_2\right| \left(\psi_k(s_1) + \psi_k(s_2) + \frac{2C}{a}\right) \\
&> C_k \left|s_1 - s_2\right| \left(\psi_k(s_1) + \psi_k(s_2)\right).
\end{aligned}$$

The first inequality holds by the chain rule, and the second inequality holds by applying (4.7) twice. In particular, we have shown that

$$\epsilon h - \mu(s_1)\frac{\partial h}{\partial s_1} - \mu(s_2)\frac{\partial h}{\partial s_2} - D_2 h > C_k \left|s_1 - s_2\right| \left(\psi_k(s_1) + \psi_k(s_2)\right). \tag{4.11}$$

Consider the function $v(s_1, s_2) = \phi^\epsilon(s_1) - \phi^\epsilon(s_2)$. By (4.10) and linearity we have:

$$\epsilon v - \mu(s_1)\frac{\partial v}{\partial s_1} - \mu(s_2)\frac{\partial v}{\partial s_2} - D_2 v \leqslant C_k \left|s_1 - s_2\right| \left(\psi_k(s_1) + \psi_k(s_2)\right).$$

We want to show that $v \leqslant h$ on $\mathbb{R}^2$. We will show this by contradiction. Suppose $\sup(v - h) > 0$. By $\phi^\epsilon \in \Psi_k(\mathbb{R})$ we have $\lim_{|s_1|+|s_2|\to\infty} v(s_1, s_2) - h(s_1, s_2) = -\infty$, there exists an open bounded subset $B \subset \mathbb{R}^2$ outside which $v \leqslant h$. Assuming (s_1^*, s_2^*) is a maximum point of $v - h$ over $\bar{B}$, as $v(s_1^*, s_2^*) - h(s_1^*, s_2^*) > 0$, we know that (s_1^*, s_2^*) must be in the interior

of B. By the necessary conditions for local maximum:

$$\begin{aligned}\frac{\partial(v-h)}{\partial s_1}(s_1^*, s_2^*) &= 0,\\ \frac{\partial(v-h)}{\partial s_2}(s_1^*, s_2^*) &= 0,\\ D_2(v-h)(s_1^*, s_2^*) &\leqslant 0.\end{aligned}$$

Thus we have, at the maximum point (s_1^*, s_2^*),

$$\epsilon h - \mu(s_1)\frac{\partial h}{\partial s_1} - \mu(s_2)\frac{\partial h}{\partial s_2} - D_2 h \leqslant C_k \left|s_1 - s_2\right| (\psi_k(s_1) + \psi_k(s_2)).$$

However, this contradicts (4.11) so we conclude $v \leqslant h$ on $\mathbb{R}^2$.

As δ is arbitrary, and $\psi_k(s_1) + \psi_k(s_2) \geqslant 2$ we obtain

$$\phi^\epsilon(s_1) - \phi^\epsilon(s_2) \leqslant C_k(1 + \frac{2C}{a}) \left|s_1 - s_2\right| (\psi_k(s_1) + \psi_k(s_2))$$

and finish the proof by taking $L = C_k(1 + \frac{2C}{a})$ □

Now we are ready to show the convergence of ϕ^ϵ to the solution to the ergodic PDE (4.6). Define

$$\begin{aligned}\lambda^\epsilon &:= \epsilon\phi^\epsilon,\\ \hat{\phi}^\epsilon(s) &:= \phi^\epsilon(s) - \phi^\epsilon(0).\end{aligned}$$

and let $K = \max\{1 + f(1) + C, L\}$. For all $\epsilon \in (0, 1]$ we have

$$\begin{aligned}\lambda^\epsilon(0) &\leqslant K,\\ \lambda^\epsilon(s) - \lambda^\epsilon(0) &\leqslant K\epsilon \left|s\right| (\psi_k(s) + 1),\\ \phi^\epsilon(s) - \phi^\epsilon(0) &\leqslant K \left|s\right| (\psi_k(s) + 1),\\ \phi^\epsilon(s_1) - \phi^\epsilon(s_2) &\leqslant K \left|s_1 - s_2\right| (\psi_k(s_1) + \psi_k(s_2)).\end{aligned}$$

Hence the family $\phi^\epsilon(s)$ is uniformly bounded and equi-continuous on any compact interval of $\mathbb{R}$. In addition, if ϵ goes to 0, then for all $s \in \mathbb{R}$, λ^ϵ will converge to the same constant. Thus, as $\lambda^\epsilon(0)$ is uniformly bounded, we can find a sequence ϵ_j such that $\epsilon_j \to 0$, $\lambda^{\epsilon_j} \to \lambda$

and $\hat{\phi}^\epsilon \to \phi$ as $j \to \infty$ for some constant λ and function ϕ. Finally, by stability of viscosity solutions, (λ, ϕ) is a solution of (4.6).

Also, note that for any $\frac{a}{2} > \nu > k$, from our construction and Lemma 4.3.2 we have $\phi(s) \in \Psi_\nu(\mathbb{R})$.

Finally, to end this section, we will prove uniqueness of (λ, ϕ) up to addition of a constant in ϕ, as in [15].

Theorem 4.3.3. *Let (λ_1, ϕ_1) and (λ_2, ϕ_2) be two solutions of the ergodic PDE (4.6) with ϕ_1 and $\phi_2 \in \Psi_\nu(\mathbb{R}) \cap \mathbb{C}^2(\mathbb{R})$ for some $\nu < \frac{a}{2}$. Then $\lambda_1 = \lambda_2$ and $\phi_1 - \phi_2$ is constant.*

Proof. Firstly, we will show that $\lambda_1 = \lambda_2$ by contradiction. Assuming without loss of generality that $\lambda_1 > \lambda_2$ and writing $\hat{\phi} := \phi_1 - \phi_2$, we obtain

$$-\mu(s)\frac{\partial \hat{\phi}}{\partial s} - \frac{\partial^2 \hat{\phi}}{\partial s^2} = \lambda_2 - \lambda_1.$$

By adding a constant to ϕ_1 if necessary, we can assume that $\sup_{s\in\mathbb{R}} \hat{\phi} > 0$. Also, let C_ν be the constant in (4.7) with ν instead of k. Considering the function $\hat{\phi}^\delta := \hat{\phi} - \delta\psi_\nu$ where δ is to be determined. By (4.7), we obtain

$$-\mu(s)\frac{\partial \hat{\phi}^\delta}{\partial s} - \frac{\partial^2 \hat{\phi}^\delta}{\partial s^2} \leqslant \lambda_2 - \lambda_1 - \delta(\psi_\nu - C_\nu).$$

Choosing δ small enough such that $\lambda_2 - \lambda_1 + \delta C_\nu < 0$, this becomes

$$-\mu(s)\frac{\partial \hat{\phi}^\delta}{\partial s} - \frac{\partial^2 \hat{\phi}^\delta}{\partial s^2} \leqslant 0.$$

As ϕ_1 and $\phi_2 \in \Psi_\nu(\mathbb{R}) \cap \mathbb{C}^2(\mathbb{R})$, there is an open bounded subset $B \subset \mathbb{R}$ such that $\hat{\phi}^\delta < 0$, outside B. Thus, by the maximum principle applied to $\hat{\phi}^\delta$ on B (see [28]), $\hat{\phi}^\delta$ cannot take its maximum at any point in the interior of B. Thus, we have $\hat{\phi}^\delta \leqslant 0$ on B and hence $\hat{\phi}^\delta \leqslant 0$ on $\mathbb{R}$. However, this contradicts our assumption that $\sup \hat{\phi} > 0$ as δ is arbitrary. Thus, we have shown that $\lambda_1 = \lambda_2$.

Next, we will show that $\phi_1 - \phi_2$ is a constant. By (4.7) there exist a bounded open set $B \subset \mathbb{R}$ such that

$$-\mu(s)\frac{\partial \psi_\nu(s)}{\partial s} - \frac{\partial^2 \psi_\nu(s)}{\partial s^2} > 0$$

outside B. If we consider $\hat{\phi}^\delta := \phi_1 - \phi_2 - \delta\psi_\nu$ again we obtain

$$-\mu(s)\frac{\partial\hat{\phi}^\delta}{\partial s} - \frac{\partial^2\hat{\phi}^\delta}{\partial s^2} < 0$$

outside B. By the maximum principle and ϕ_1 and $\phi_2 \in \Psi_\nu(\mathbb{R}) \cap \mathbb{C}^2(\mathbb{R})$, $\hat{\phi}^\delta$ cannot attain its maximum at any point in the interior of B^c. Also, as B is independent of δ, letting $\delta \to 0$ we deduce $\sup_{s\in\mathbb{R}}(\phi_1 - \phi_2) = \max_{s\in B}(\phi_1 - \phi_2)$.

Writing $\tilde{\phi} := \phi_1 - \phi_2$, we know

$$-\mu(s)\frac{\partial\tilde{\phi}}{\partial s} - \frac{\partial^2\tilde{\phi}}{\partial s^2} = 0. \tag{4.12}$$

We can apply the strong maximum principle as in Section 1 of [28] to $\phi_1 - \phi_2$ on $B_R(0)$, for some sufficiently large R such that $B \subset B_R(0)$, to deduce that $\phi_1 - \phi_2$ is constant on $B_R(0)$. As R is arbitrary, we obtain $\forall s \in \mathbb{R}\ \phi_1(s) - \phi_2(s) = \phi_1(0) - \phi_2(0) = C$ for some constant C. □

Remark 4.3.4. Theorem 4.3.3 not only holds for solutions to (4.6). If (λ, ϕ_1) and (λ, ϕ_2) are a supersolution and a subsolution to (4.6) respectively with $\phi_1 \in \Psi_\nu^+(\mathbb{R})$ and $\phi_2 \in \Psi_\nu^-(\mathbb{R})$, (4.12) becomes

$$-\mu(s)\frac{\partial\tilde{\phi}}{\partial s} - \frac{\partial^2\tilde{\phi}}{\partial s^2} \leqslant 0.$$

We can still apply the strong maximum principle in [28] so that $\phi_1 - \phi_2$ is a constant still holds.

4.4 Convergence of Viscosity Solutions to the Cauchy Problem

We are now ready to find the long time asymptotic behaviour of the unique solution of the Cauchy problem:

$$\frac{\partial g}{\partial t} - \mu(s)\frac{\partial g}{\partial s} - \frac{1}{2}\sigma^2\frac{\partial^2 g}{\partial s^2} - \frac{1}{2}\frac{\mu(s)^2}{\sigma^2} = 0 \tag{4.13}$$

with initial condition $g(s, 0) = 0$. Our main result is the following.

Theorem 4.4.1. *Assuming that Hypothesis 4.2.1 holds, (4.13) has a unique solution g such that for any fixed T, $g \in \Psi_k(\mathbb{R}_T)$ where $0 < k < \frac{a}{2}$. In addition, there is a unique solution (λ, ϕ) to (4.6) up to addition of a constant in ϕ such that $\phi \in \Psi_\nu(\mathbb{R}) \cap C(\mathbb{R})$ for*

any $\frac{a}{2} > \nu > k$ *and there exists a constant* c *such that*

$$\lim_{t\to\infty}(g(s,t)-\lambda t)=\phi(s)+c.$$

In this section, we still assume that $\sigma=\sqrt{2}$. Before we start to prove this theorem we claim the following uniform estimates for g: for any $\nu > k$,

$$|g(s_1,t)-g(s_2,t)| \leqslant L_1\,|s_1-s_2|\,(\psi_\nu(s_1)+\psi_\nu(s_2)), \quad \forall s_1,s_2\in\mathbb{R} \text{ and } t\in[0,T) \tag{4.14}$$

for some constant L_1 independent of s_1 and s_2. Proofs of this estimate can be found in Section 3 of [15] and we will skip these proofs in this paper.

Remark 4.4.2. Given a constant $\hat{\lambda}$, consider the Cauchy problem

$$\frac{\partial g}{\partial t}-\mu(s)\frac{\partial g}{\partial s}-\frac{1}{2}\sigma^2\frac{\partial^2 g}{\partial s^2}-\frac{1}{2}\frac{\mu(s)^2}{\sigma^2}+\hat{\lambda}=0 \tag{4.15}$$

with boundary condition $g(s,0)=0$, and the ergodic problem

$$\lambda-\mu(s)\frac{\partial \phi}{\partial s}-\frac{1}{2}\sigma^2\frac{\partial^2 \phi}{\partial s^2}-\frac{1}{2}\frac{\mu(s)^2}{\sigma^2}+\hat{\lambda}=0. \tag{4.16}$$

If g and (λ,ϕ) are solutions to (4.13) and (4.6) respectively, then $g_{\hat{\lambda}} := g-\hat{\lambda}t$ and $(\lambda-\hat{\lambda},\phi)$ are solutions to (4.15) and (4.16) respectively. Now choosing $\hat{\lambda}=\lambda$, we see $g-\lambda t$ converges to $\phi+C$ if and only if g_λ converges to $\phi+C$. Hence, throughout this section, we can without loss of generality assume that $\lambda=0$.

We shall consider a discounted version of $\psi_\nu - C_\nu$ where C_ν is the constant in the function (4.7) with ψ_ν instead of ψ_k. Consider $h(s,t)=(\psi_\nu-C_\nu)e^{-t}$. By (4.7) we obtain

$$\frac{\partial h}{\partial t}-\mu(s)\frac{\partial h}{\partial s}-\frac{\partial^2 h}{\partial s^2}\geqslant 0. \tag{4.17}$$

We will first prove another boundedness property of g.

Lemma 4.4.3. *For any fixed* $T>0$*, there exist a positive function* M *such that if* g *is the unique solution to (4.13) in Theorem 4.4.1,*

$$|g(s,t)|\leqslant \epsilon\psi_\nu(s)+M(\epsilon) \quad \forall \quad s\in\mathbb{R}, \quad t\in[0,T]. \tag{4.18}$$

Proof. As $g\in\Psi_k(\mathbb{R}_T)$ and $\phi\in\Psi_\nu(\mathbb{R})$, there exists a positive function M_ν, such that

$$|g(s,0)-\phi(s)|\leqslant\frac{\epsilon}{2}\psi_\nu(s)+M_\nu(\epsilon).$$

Thus $g(s,t) - \phi(s) - \frac{\epsilon}{2}h(s,t) - \frac{\epsilon}{2}C_\nu - M_\nu(\epsilon) < 0$ on $B_R(0) \times \{0\}$ for any $R > 0$ and

$$\lim_{|s|\to\infty} \sup_{t\in[0,T]} \left\{ g(s,t) - \phi(s) - \frac{\epsilon}{2}h(s,t) \right\} \to -\infty, \quad \forall T > 0.$$

Hence there exists $R_{T,\epsilon} > 0$ such that $g(s,t) - \phi(s) - \frac{\epsilon}{2}h(s,t) < 0$ on $B^c_{R_{T,\epsilon}}(0) \times [0,T]$, in particular on $\partial B^c_{R_{T,\epsilon}}(0) \times [0,T]$. Combining these results we can now apply the maximum principle in [28] to $g(s,t) - \phi(s) - \frac{\epsilon}{2}h(s,t) - \frac{\epsilon}{2}C_\nu - M_\nu(\epsilon)$. By (4.17) and (4.13), as ϕ is a solution to (4.6) with $\lambda = 0$, it cannot attain a maximum on $B_{R_{T,\epsilon}}(0) \times (0,T]$. We deduce that $g(s,t) - \phi(s) - \frac{\epsilon}{2}h(s,t) - \frac{\epsilon}{2}C_\nu - M_\nu(\epsilon) < 0$ on $B_{R_{T,\epsilon}}(0) \times (0,T]$. As $\phi \in \Psi_\nu(\mathbb{R})$, there exist a function $\hat{M}$ such that

$$\phi(s) \leqslant \frac{\epsilon}{2}\psi_\nu(s) + \hat{M}(\epsilon)$$

Choosing $M(\epsilon) = M_\nu(\epsilon) + \epsilon C_\nu + \hat{M}(\epsilon)$, we obtain

$$g(s,t) \leqslant \epsilon\psi_\nu(s) + M(\epsilon), \quad \forall s \in \mathbb{R}, \quad t \in [0,T].$$

With slight modification we can also prove:

$$g(s,t) \geqslant -\epsilon\psi_\nu(s) - M(\epsilon), \quad \forall s \in \mathbb{R}, \quad t \in [0,T].$$

□

By Theorem 3.2 of [15], Lemma 4.4.3 implies that there exist a non-decreasing, positive upper semicontinuous function L_R, with $L_R(0) = 0$, such that for all $s \in B_R(0)$ and t_1, $t_2 > 0$ we have

$$|g(s,t_1) - g(s,t_2)| \leqslant L_R(|t_1 - t_2|). \tag{4.19}$$

We shall now consider the functions:

$$g^+(s) = \limsup_{t\to\infty} g(s,t),$$
$$g^-(s) = \liminf_{t\to\infty} g(s,t).$$

It is clear that (4.19) and (4.14) still hold for g^+ and g^-. We then have the following lemma.

Lemma 4.4.4. *g^+ and g^- are respectively a viscosity subsolution and a viscosity super-*

solution of the ergodic problem

$$-\mu(s)\frac{\partial \phi}{\partial s} - \frac{\partial^2 \phi}{\partial s^2} - \frac{\mu(s)^2}{4} = 0 \tag{4.20}$$

Proof. For any fixed $T > 0$, $g(s, t + T)$ is a solution to (4.13). Consider $\hat{g}^T(s,t) = \sup_{\tilde{T} \geqslant T}\{g(s, t + \tilde{T})\}$. Then, (4.18) and (4.19) holds for $\hat{g}^T$ for all $T > 0$. Hence, $\hat{g}^T$ are continuous for all $T > 0$ and are viscosity subsolutions to (4.13). Now, as $\hat{g}^T(s,t)$ converges monotonically to g^+, by Dini's lemma, this convergence is uniform on compact subsets of $\mathbb{R} \times [0, T]$. By the stability of viscosity solutions, regarding g^+ as a function of s and t we see that g^+ is a viscosity subsolution of (4.13). As g^+ is independent of t and we have assumed $\sigma^2 = 2$, (4.13) reduces to (4.20) and this finishes the argument. An identical proof shows that g^- is a viscosity supersolution for (4.20) □

We have discussed, in Remark 4.3.4, that if $(0, \phi)$ is a solution to (4.20) and $(0, g^+)$ is a subsolution to (4.20) then $\phi - g^+ = C$ where C is constant. This means that g^+ and g^- are, in fact, viscosity solutions to (4.20) with $g^+ - a = \phi = g^- - b$ for some constant a, b. By (4.14) and (4.18) the family $g(\cdot, t)_{t \geqslant 0}$ is uniformly bounded and equicontinuous on compact subset of $\mathbb{R}$. Thus, by the Ascoli–Arzela theorem, this family is precompact in $C(\mathbb{R})$. If we can show $a = b$, we can claim that $g(\cdot, t)$ converges to $\phi + a$ in $C(\mathbb{R})$ as $t \to \infty$ which finishes the proof of Theorem 4.4.1.

Proof of Theorem 4.4.1. We will first show that there is indeed a sequence T_j such that $g(\cdot, T_j) \to g^-$. From the definition of g^- we have a sequence t_j such that $g(0, t_j) \to g^-(0)$ Notice that the family $u(\cdot, \cdot + t)$ is also precompact in $C(\mathbb{R} \times [0, T])$ so we can find a subsequence of t_j, where we abuse notation and still call it t_j, such that $g(\cdot, \cdot + t_j - 1) \to \hat{g}(\cdot, \cdot)$ for some $\hat{g}$ defined on $\mathbb{R} \times [0, T]$. We see that from the definition of g^-, $g^- \leqslant \hat{g} \quad \forall x \in \mathbb{R}$ and $t \in [0, T]$. Moreover, $g^-(0) = \hat{g}(0, 1)$, so $(0, 1)$ is a maximum of $\hat{g} - g^-$, and by stability of the viscosity solutions, $\hat{g}$ is also a viscosity solution to (4.13). Substituting $\tilde{g} := \hat{g} - g^-$ into (4.13) we obtain

$$\frac{\partial \tilde{g}}{\partial t} - \mu(s)\frac{\partial \tilde{g}}{\partial s} - \frac{\partial^2 \tilde{g}}{\partial s^2} = 0.$$

Applying the strong maximum principle in [28] to $\tilde{g}$ on $\mathbb{R} \times [0, 1)$ we obtain $\tilde{g}(s,t) = 0, \quad \forall (s,t) \in \mathbb{R} \times [0, 1)$. From the definition of $\hat{g}$, there is a sequence t_j such that $g(\cdot, t_j - 1) \to \hat{g}(\cdot, 0) = g^-(\cdot)$; setting $T_j = t_j - 1$ we get the sequence T_j such that $g(\cdot, T_j) \to g^-$.

Now for fixed $T > 0$, consider $g^T(s,t) := g(s, t + T) - \phi(s) - b$. This is a viscosity solution

of

$$\frac{\partial f}{\partial t} - \mu(s)\frac{\partial f}{\partial s} - \frac{\partial^2 f}{\partial s^2} = 0. \tag{4.21}$$

It is clear that, for any fixed $\epsilon > 0$, there is a R_ϵ such that $g^T(s,0) \leqslant \epsilon h(s,0)$ on $B_{R_\epsilon}(0)^c$ for all $T > 0$ where $h(s,t) = (\psi_\nu - C_\nu)e^{-t}$. As there is a sequence T_j such that $g(\cdot, T_j) \to g^-$, we can choose T sufficiently large that $g^T(s,0) \leqslant \epsilon$ for all s in $B_{R_\epsilon}(0)$, to obtain $g^T(s,0) \leqslant \epsilon h(s,0) + \epsilon C_\nu + \epsilon, \quad \forall x \in \mathbb{R}$. By (4.19) and $\phi \in \Psi_\nu(\mathbb{R})$ we deduce that, for any fixed $\tilde{T} > 0$,

$$\lim_{|s|\to\infty} \sup_{t\in(0,\tilde{T})} g^T(s,t) - \epsilon h(s,t) - \epsilon C_\nu - \epsilon \to -\infty.$$

Hence, there exists some $R > 0$ such that $g^T(s,t) \leqslant \epsilon h(s,t) + \epsilon C_\nu + \epsilon$ on B_R^c. As g^T is a viscosity subsolution and by (4.17), $\epsilon h(s,t) + \epsilon C_\nu + \epsilon$ is a viscosity supersolution to (4.21) so we can apply the comparison result to deduce $g^T \leqslant \epsilon h(s,t) + \epsilon C_\nu + \epsilon$ for $(s,t) \in \mathbb{R} \times [0, \tilde{T}]$. As $\tilde{T}$ is arbitrary, this holds for all $t > 0$. Thus for all $t > 0$

$$g(0, t+T) \leqslant \phi(0) + b + \epsilon((1 - C_\nu)e^{-t} + (1 + C_\nu)).$$

Choosing a suitable sequence t_i, the left hand side of the above sequence converges to g^+. Then letting ϵ converge to 0 we obtain $a \leqslant b$. However, the definition of g^+ and g^- implies $a \geqslant b$. Thus, we get $a = b$.

Summarising results above, we obtain

$$\limsup_{t\to\infty} g(s,t) = \liminf_{t\to\infty} g(s,t) = \phi + a$$

for constant a if a solution to (4.6) is $(0, \phi)$. By Remark 4.4.2, this can be generalised to

$$\limsup_{t\to\infty} \{g(s,t) - \lambda t\} = \liminf_{t\to\infty} \{g(s,t) - \lambda t\} = \phi + a$$

for constant a if a solution to (4.6) is (λ, ϕ). Here a depends on ϕ as ϕ is unique up to the addition of a constant. Hence, choosing $c = a$ we can deduce that:

$$\lim_{t\to\infty} (g(s,t) - \lambda t) = \phi(s) + c,$$

where (λ, ϕ) is the solution to (4.6) as above, which finishes the proof. □

Remark 4.4.5. Note that the result in this chapter does not include the result in the last chapter or vice versa. In addition, in this chapter, we did not provide an explicit probability

representation for λ.

4.5 Probabilistic Representation of Parameters in Theorem 4.4.1

From the results in Chapter 3, we have obtained the probabilistic representation of λ, which is $\int \frac{\mu(s)^2}{2\sigma^2} d\Pi$ where Π is the unique invariant distribution of S_t. The only parameter that has not been mentioned is the constant c. Note that ϕ is unique up to the addition of a constant, so our parameter c is dependent on ϕ. In particular, we have the following result:

Theorem 4.5.1. *Assume that ϕ in Theorem 4.4.1 satisfies for all $t > 0$ and all s, $\mathbb{E}[|\phi(S_t)| \, | S_0 = s] < \infty$. Then the constant c has the probabilistic representation $c = \int -\phi(s) d\Pi$.*

Proof. To find a probabilistic representation for c we will use an asymptotic expansion. In particular, we shall consider:

$$g^\epsilon(s,t) := g\left(s, \frac{t}{\epsilon}\right). \tag{4.22}$$

By the chain rule, g^ϵ is the unique solution to

$$\frac{\partial g^\epsilon}{\partial t} - \frac{1}{\epsilon}\left(\mu(s)\frac{\partial g^\epsilon}{\partial s} + \frac{1}{2}\sigma^2 \frac{\partial^2 g^\epsilon}{\partial s^2} + \frac{1}{2}\frac{\mu(s)^2}{\sigma^2}\right) = 0 \tag{4.23}$$

with boundary condition $g^\epsilon(s,0) = 0$.

As we have deduced that $g^\epsilon(s,t)$ converges to $\frac{\lambda t}{\epsilon} + \phi(s) + c$ as ϵ converges to 0, if we can apply an expansion to g^ϵ, the expansion must have the form:

$$g^\epsilon(s,t) = \frac{\lambda t}{\epsilon} + \phi(s) + \tilde{g}^{(0)}(s,t,\epsilon), \tag{4.24}$$

where $\tilde{g}^{(0)}$ has order 1 at least pointwisely and converges to constant c as ϵ converges to 0. Substituting this into (4.23) with boundary condition $g^\epsilon(s,0) = 0$, we obtain the following PDE for $\tilde{g}^{(0)}$, with initial condition controlled by ϕ:

$$\frac{\partial \tilde{g}^{(0)}}{\partial t} + \frac{1}{\epsilon}\left(\mu(s)\frac{\partial \tilde{g}^{(0)}}{\partial s} + \frac{1}{2}\sigma^2 \frac{\partial^2 \tilde{g}^{(0)}}{\partial s^2}\right) = 0, \tag{4.25}$$

$$-\tilde{g}^{(0)}(s,0,\epsilon) = \phi(s) \tag{4.26}$$

This is a linear second order PDE. Assuming that $\phi(S_t^\epsilon)$ is integrable for all t, we can apply the Feynman–Kac formula to (4.25) and (4.26) to obtain the probabilistic representation

$$\tilde{g}^{(0)} = \mathbb{E}\left[-\phi(S_t^\epsilon)|S_0^\epsilon = s\right], \tag{4.27}$$

where S_t^ϵ has dynamics

$$dS_t^\epsilon = \frac{1}{\epsilon}\mu(S_t^\epsilon)\,dt + \frac{1}{\sqrt{\epsilon}}\sigma(S_t^\epsilon)\,dW_t.$$

Finally, by Theorem 4.4.1, $\phi \in \Psi_\nu$ for some $\nu < \frac{a}{2}$. Choosing $\hat{\nu}$ such that $\nu < \hat{\nu} < \frac{a}{2}$ and by (4.7), there exists a $\hat{C} > 0$ such that $e^{\hat{\nu}x^2} + \hat{C}$ is a Lyapunov function dominating ϕ. Hence, we can use the same proof as Theorem 3.3.1 to deduce that (4.27) converges to $\int -\phi(s)d\Pi$ as ϵ converges to 0. This gives the probabilistic representation of c in Theorem 4.4.1.

□

Remark 4.5.2. The problem of looking for an asymptotic behaviour of the solution to (4.25) and (4.26) is a singular perturbation problem. What we are actually doing is separating our current problem into a problem related to long time behaviour problem and a singular perturbation. This separation highly relies on the fact that the HJB equation corresponding to log utility can be simplified to a linear parabolic PDE using feedback control using suitable transformation. Thus, it is quite difficult to apply a similar method on a general HARA utility function whose simplified HJB (2.7) is non-linear.

As our fast mean reverting process is used to approximate the behaviour of an iid sequence of noise, it is also a reasonable idea to assume that the initial price of the asset follows some distribution instead of a deterministic value. As in Chapter 3 we have shown that S_t has a unique invariant distribution Π, which is also the unique invariant distribution of S_t^ϵ, we can suppose that S_0^ϵ follows the invariant distribution Π. Then S_t^ϵ will also follow the same distribution, and the Feynman–Kac representation (4.27) leads to the following interesting phenomenon: As $\int \phi(s)d\Pi + \mathbb{E}\left[-\phi(S_t^\epsilon)|S_0^\epsilon \sim \Pi\right] = 0$, all terms in (4.24) related to ϕ vanish and we obtain

$$\int g^\epsilon(s,t)d\Pi = \frac{\lambda t}{\epsilon} \tag{4.28}$$

Hence, we have the following result

Theorem 4.5.3. *Given all conditions in Theorem 4.4.1 and Theorem 4.5.1, the averaged*

maximised expected utility with respect to the unique invariant distribution Π *of* S_t^ϵ *satisfies*

$$\int V^\epsilon(x,s,t)d\Pi := \sup_\pi \mathbb{E}[\log(X_t^\pi)|X_0 = x, S_0^\epsilon \sim \Pi] = \log(x) + \frac{\lambda t}{\epsilon}$$

This means that using the optimal investment strategy, our wealth process has a constant rate of income of wealth asymptotically. In addition, the function $\phi+c$ measures the amount of wealth we gain or lose before the process S_t^ϵ completely mixes up. If we assume that our fast mean-reverting process S_t^ϵ is already mixed up at the beginning of our trading, the asymptotic income is the same as the income due to a fixed interest rate which is proportional to the rate of mean reversion of S_t^ϵ.

Chapter 5

Extension to Multi-asset Setting

In this chapter, we will discuss the possible extension of our result on the asymptotic behaviour of maximised expected utility for a single asset market to a multi-asset market with both fast and slow assets.

5.1 Extension to Linear Market with Both Fast and Slow Moving Assets

The most important extension of the result from this paper is to add another stock that is not fast mean-reverting to the market. The main difference between this setting and the previous one is that we can no longer transform our asymptotic problem to a long time behaviour problem, due to the existence of the slow asset. The general construction of this extension is as follows:

$$dS_t^1 = \mu_1(S_t^1, S_t^2)\,dt + \sigma_1(S_t^1, S_t^2)\,dW_t^1, \tag{5.1}$$

$$dS_t^2 = \frac{1}{\epsilon}\mu_2(S_t^1, S_t^2)\,dt + \frac{1}{\sqrt{\epsilon}}\sigma_2(S_t^1, S_t^2)\,dW_t^2, \tag{5.2}$$

$$B_t = 1 \tag{5.3}$$

where W_t^1 and W_t^2 are Brownian motions. Our self financing wealth process X_t^π has dynamics

$$dX_t^\pi = \pi_t^1\,dS_t^1 + \pi_t^2\,dS_t^2 \tag{5.4}$$

where π_t^1 and π_t^2 are the number of shares of asset S_t^1 and S_t^2 we hold at time t respectively.

We still consider the logarithmic utility function. Our value process for the maximised expected utility is defined by

$$V(x, s_1, s_2, t) := \sup_{\pi \in \mathcal{A}} \mathbb{E}[\log(X_t^\pi) | S_0^1 = s_1, S_0^2 = s_2, X_0^\pi = x], . \tag{5.5}$$

The admissible set $\mathcal{A}$ is still the set of controls $\pi := (\pi_1, \pi_2)$ such that (5.4) has a non-negative strong solution with $\mathbb{E}[\log(X_t^\pi) | S_0^1 = s_1, S_0^2 = s_2, X_0^\pi = x] < \infty$ for all t. Using the following notation for simplicity,

$$\mathcal{L}^1 f := \mu_1(s_1, s_2) \frac{\partial f}{\partial s_1} + \frac{1}{2} \sigma_1(s_1, s_2) \frac{\partial^2 f}{\partial s_1^2},$$
$$\mathcal{L}^2 f := \mu_2(s_1, s_2) \frac{\partial f}{\partial s_2} + \frac{1}{2} \sigma_2(s_1, s_2) \frac{\partial^2 f}{\partial s_2^2},$$

and assuming W_t^1 and W_t^2 are independent, the related HJB equation is

$$\begin{aligned} &- \frac{\partial V}{\partial t} + \mathcal{L}^1 V + \frac{1}{\epsilon} \mathcal{L}^2 V \\ &+ \sup_{\pi_1, \pi_2 \in \mathbb{R}} \left(\pi_1 \mu_1 \frac{\partial V}{\partial x} + (\pi_1)^2 \frac{1}{2} \sigma_1^2 \frac{\partial^2 V}{\partial x^2} + \frac{1}{\epsilon} \pi_2 \mu_2 \frac{\partial V}{\partial x} + (\pi_2)^2 \frac{1}{2\epsilon} \sigma_2^2 \frac{\partial^2 V}{\partial x^2} \right) = 0. \end{aligned} \tag{5.6}$$

Remark 5.1.1. In a simplified version of this multi-scale setting, we assume that μ_i and σ_i depend only on S_t^i. In this case, our utility maximisation problem is equivalent to a combination of a utility maximisation problem in a market with only S_t^1 and B_t and a utility maximisation problem in a market with only S_t^2 and B_t. In particular, we can separate the wealth we obtain from S_t^1 and S_t^2, and the value process V is in the form $V(x, s_1, s_2, t) = \log(x) + f(t, s_1) + g(t, s_2)$ for some functions f and g that can be deduced from the corresponding single asset problems. However, if W_t^1 and W_t^2 are not independent, or μ_i and σ_i not depend only on S_t^i, the stochastic control problem becomes difficult to analyse.

Similarly as before, we make the following assumption on the solution of (5.6).

Hypothesis 5.1.2. There exists a function $V \in C^{2,2,2,1}(\mathbb{R}^+ \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}^+)$ that solves (5.6) with initial condition $V(x, s_1, s_2, 0) = \log(x)$. Furthermore, $V(x, s_1, s_2, t) = \log(x) + g(s_1, s_2, t)$ where g satisfies the quadratic growth condition.

In addition, the optimal feedback control

$$\begin{aligned}\hat{\pi}_t^1 &= X_t^\pi \frac{\mu_1(S_t^1, S_t^2)}{\sigma_1(S_t^1, S_t^2)^2} \\ \hat{\pi}_t^2 &= X_t^\pi \frac{\mu_2(S_t^1, S_t^2)}{\sigma_2(S_t^1, S_t^2)^2}\end{aligned} \tag{5.7}$$

is admissible.

We skip the verification argument, as it is almost identical as before, and claim that this solution V equals to the maximised expected utility $\sup_{\pi\in\mathcal{A}} \mathbb{E}[\log(X_t^\pi)|S_0^1 = s_1, S_0^2 = s_2, X_0^\pi = x]$.

Substituting $V(x, s_1, s_2, t) = \log(x) + g(s_1, s_2, t)$ into (5.6) and applying the optimal feedback control (5.7), we can deduce the following PDE for g:

$$\frac{1}{\epsilon}\left(\mathcal{L}^2 g + \frac{1}{2}\gamma_2^2(s_1, s_2)\right) - \frac{\partial g}{\partial t} + \mathcal{L}^1 g + \frac{1}{2}\gamma_1^2(s_1, s_2) = 0. \tag{5.8}$$

where $\gamma_1(s_1, s_2) := \frac{\mu_1(s_1,s_2)}{\sigma_1(s_1,s_2)}$ and $\gamma_2(s_1, s_2) := \frac{\mu_2(s_1,s_2)}{\sigma_2(s_1,s_2)}$.

As a first attempt to deal with this problem, we assume that μ_2 and σ_2 are functions of s_1 and s_2 while μ_1 and σ_1 are functions of s_1 only. This corresponds to the observation that the dynamic of noise factors depends on the asset but the dynamic of the primary asset is independent of noise factors.

We will first consider in detail a special but useful case. In particular, we assume that the dynamics of the slow moving asset do not depend on the fast mean reverting asset, and the dynamic of fast mean reverting asset is a fast mean reverting process independent of the slow process scaled by the slow process. This case can be understood as where the price of a stock or an option does not depend on the noise factors affecting our observation, but how much noise we observe may depend on the current price of the stock or option.

Formalizing this simplified setting, we assume that the market factors in this market model follow the dynamics

$$\begin{aligned}dS_t^1 &= \mu_1(S_t^1)\, dt + \sigma_1(S_t^1)\, dW_t^1, \\ dS_t^2 &= \frac{1}{\epsilon}\mu_2(S_t^2)\, dt + \frac{1}{\sqrt{\epsilon}}\sigma_2(S_t^2)\, dW_t^2.\end{aligned}$$

The assets we can trade are S_t^1, a bank account B_t and $\tilde{S}_t^2$ with dynamics defined as

$$\begin{aligned} dS_t^1 &= \mu_1(S_t^1)\,dt + \sigma_1(S_t^1)\,dW_t^1, \\ d\tilde{S}_t^2 &= \zeta(S_t^1)\,dS_t^2, \\ B_t &= 1. \end{aligned}$$

Here W_t^1 and W_t^2 are independent Brownian motions, $\tilde{S}_0^2$ is a function of S_0^1 and S_0^2 and ζ is a non-zero deterministic smooth and bounded function. The HJB equation (5.6) becomes

$$\begin{aligned} &-\frac{\partial V}{\partial t} + \mathcal{L}^1 V + \mathcal{L}^2 V \\ &+ \sup_{\pi_1,\pi_2\in\mathbb{R}} \left(\pi_1\mu_1(s_1)\frac{\partial V}{\partial x} + (\pi_1)^2\frac{1}{2}\sigma_1^2(s_1)\frac{\partial^2 V}{\partial x^2} + \pi_2\zeta(s_1)\mu_2(s_2)\frac{\partial V}{\partial x} + (\pi_2)^2\zeta^2(s_1)\frac{1}{2}\sigma_2^2(s_2)\frac{\partial^2 V}{\partial x^2} \right) = 0 \end{aligned}$$

and (5.8) becomes:

$$\frac{1}{\epsilon}\left(\mathcal{L}^2 g + \frac{1}{2}\,\zeta(s_1)\gamma_2^2(s_2)\right) - \frac{\partial g}{\partial t} + \mathcal{L}^1 g + \frac{1}{2}\gamma_1^2(s_1) = 0, \tag{5.9}$$

where

$$\begin{aligned} \mathcal{L}^1 f &:= \mu_1(s_1)\frac{\partial f}{\partial s_1} + \frac{1}{2}\sigma_1(s_1)\frac{\partial^2 f}{\partial s_1^2}, \\ \mathcal{L}^2 f &:= \mu_2(s_2)\frac{\partial f}{\partial s_2} + \frac{1}{2}\sigma_2(s_2)\frac{\partial^2 f}{\partial s_2^2} \\ \gamma_1(s_1) &:= \frac{\mu_1(s_1)}{\sigma_1}, \\ \gamma_2(s_2) &:= \frac{\mu_2(s_2)}{\sigma_2}. \end{aligned}$$

Remark 5.1.3. The maximised expected utility we consider here is

$$V(x, s_1, s_2, t) := \sup_{\pi\in\mathcal{A}} \mathbb{E}[\log(X_t^\pi)|S_0^1 = s_1, S_0^2 = s_2, X_0^\pi = x]$$

instead of

$$V(x, s_1, s_2, t) := \sup_{\pi\in\mathcal{A}} \mathbb{E}[\log(X_t^\pi)|S_0^1 = s_1, \tilde{S}_0^2 = s_2, X_0^\pi = x],$$

In particular, we assume that we can calculate S_0^2 for each pair of fixed S_0^1 and $\tilde{S}_0^2$.

We can then prove the following result.

Theorem 5.1.4. *Assume that $\sigma_1 > 0$ and $\sigma_2 > 0$ are constants and μ_1 is bounded and Lipschitz continuous and μ_2 is a decreasing linear function of s_2. Then $g(s_1, s_2, t)$ in Hypothesis 5.1.2 is the unique quadratic growth function that solves (5.9) with initial condition $g(s_1, s_2, 0) = 0$ and satisfies*

$$\lim_{\epsilon \to 0} \left(g(s_1, s_2, t) - \frac{\lambda}{\epsilon} \mathbb{E} \left[\int_0^t \zeta(S_u^1) du | S_0^1 = s_1 \right] \right) = \zeta(s_1)\phi(s_2) + c(s_1) \tag{5.10}$$

where (λ, ϕ) is a solution to the ergodic PDE

$$\mathcal{L}^2 \phi(s_2) + \frac{1}{2}\gamma_2^2(s_2) = \lambda \tag{5.11}$$

and c is a function only depends on s_1.

Proof. Using the quadratic growth condition of g, we can show that this solution is unique by a comparison result similar to Theorem 4.2.2 with a quadratic test function instead of ψ_k. As $\mathcal{L}^1 g$ has order 1, we expect that this term does not affect the highest order term of the asymptotic behaviour of g. Thus, we expect that as in the one asset case, g has an expansion as before:

$$g(s_1, s_2, t) = \frac{1}{\epsilon} g^{(-1)}(s_1, s_2, t) + \hat{\phi}(s_1, s_2) + g^{(0)}(s_1, s_2, t). \tag{5.12}$$

Substituting this in (5.9) and comparing orders as before we obtain the following set of PDEs: a PDE for $g^{(-1)}$

$$\mathcal{L}^2 g^{(-1)} = 0, \tag{5.13}$$

a PDE for $g^{(-1)}$ and ϕ

$$-\frac{\partial g^{(-1)}}{\partial t} + \mathcal{L}^1 g^{(-1)} = -\mathcal{L}^2 \hat{\phi} - \frac{1}{2}\zeta\gamma_2^2 \tag{5.14}$$

and a PDE for $g^{(0)}$ with parameter depending on ϕ

$$-\frac{\partial g^{(0)}}{\partial t} + \mathcal{L}^1 g^{(0)} + \frac{1}{\epsilon}\mathcal{L}^2 g^{(0)} + \mathcal{L}^1 \hat{\phi} + \frac{1}{2}\gamma_1^2 = 0, \tag{5.15}$$

with initial conditions:

$$g^{(-1)}(s_1, s_2, 0) = 0, \tag{5.16}$$

$$-g^{(0)}(s_1, s_2, 0) = \phi(s_1, s_2). \tag{5.17}$$

First, we consider the PDE (5.13) with initial condition (5.16). A trivial solution to this PDE is $g^{(-1)} = 0$. As this PDE can be regarded as a parametrized parabolic PDE, this solution is unique up to the addition of a function of t and s_1.

With this $g^{(-1)}$, the left-hand side of (5.14) depends only on t and s_1 and its right-hand side only depends on s_1 and s_2, by the separation of variables both sides of (5.14) equal $-\lambda(s_1)$ for some function $\lambda : \mathbb{R} \to \mathbb{R}$.

Considering the RHS first, we have

$$\lambda(s_1) = \mathcal{L}^2\hat{\phi} + \frac{1}{2}\zeta(s_1)\gamma_2^2(s_2). \tag{5.18}$$

As $\mathcal{L}^2$ only involves derivatives with respect to s_2, this equation can be regarded as a parametrized version of an ergodic PDE with respect to s_2. By the result of previous chapters, as μ_2 and σ_2 clearly satisfies Hypothesis 4.2.1, this ergodic PDE has a solution $(\lambda\zeta(s_1), \phi(s_2)\zeta(s_1))$ where (λ, ϕ) is the unique solution up to addition of a constant to ϕ of

$$\lambda = \mathcal{L}^2\phi + \frac{1}{2}\gamma_2^2(s_2) \tag{5.19}$$

such that $\phi \in \Phi_k$. In addition, as we have shown that for each fixed s_1, (5.18) has a unique solution up to the addition of a constant in ϕ, so $(\lambda\zeta(s_1), \phi(s_2)\zeta(s_1))$ is the unique solution up to the addition of a function of s_1 in $\hat{\phi}$ that satisfies the corresponding boundedness property.

Now considering the LHS of (5.14) and (5.16), we have

$$\frac{\partial g^{(-1)}}{\partial t} - \mathcal{L}^1 g^{(-1)} = \lambda\zeta(s_1).$$

with initial condition $g^{(-1)}(s_1, 0) = 0$.

This is a linear second order PDE and it is easy to check that it has a unique solution. Then we can apply the Feynman–Kac formula to obtain a probabilistic representation:

$$g^{(-1)}(s_1, t) = \lambda\mathbb{E}\left[\int_0^t \zeta(S_u^1)du | S_0^1 = s_1\right]. \tag{5.20}$$

The only PDE left is the singular perturbation problem (5.15) with initial condition (5.17). Note that, as we have assumed that ζ is smooth, (5.15) is well defined. If it has a unique classical solution and ϕ is uniformly integrable, we can apply the Feynman–Kac formula

to obtain

$$g^{(0)} = \mathbb{E}\left[\int_0^t \phi(S_u^2)\mathcal{L}^1\zeta(S_u^1) + \frac{1}{2}\gamma_1^2(S_u^1)du - \zeta(S_t^1)\phi(S_t^2)|S_0^1 = s_1, S_0^2 = s_2\right], \tag{5.21}$$

where S_t^1, S_t^2 has dynamics (5.1) and (5.2). Intuitively, this converges to a function that depends only on s_1 using a parametrized version of the ergodic theory. Indeed, as we will see in Lemma 5.1.5 we can not only deduce the asymptotic limit of $g^{(0)}$, we can even expand it to higher order terms and deduce the rate and accuracy of this convergence, which concludes the proof of Theorem 5.1.4.

□

Lemma 5.1.5. *Assume that μ_1 is bounded and Lipschitz continuous and μ_2 is a decreasing linear function and that $\sigma_i > 0$ are constants. Furthermore, assume ζ is smooth and bounded, and the solution ϕ to (5.19) is smooth with $\mathbb{E}[\phi(S_t^2)|S_0^2 = s_2] < \infty \quad \forall\, s_2$ and t. Then $\forall k > 0$ the solution $g^{(0)}$ to (5.15) can be expanded as*

$$\begin{aligned} g^{(0)}(s_1, s_2, t) =& g_0(s_1, t) + h_0\left(s_1, s_2, \frac{t}{\epsilon}\right) + \sum_{i=1}^{k} \epsilon^i g_i(s_1, s_2, t) \\ &+ \sum_{i=1}^{k} \epsilon^i h_i\left(s_1, s_2, \frac{t}{\epsilon}\right) + e_k, \end{aligned} \tag{5.22}$$

where g_i and h_i are the unique solutions to

$$-\frac{\partial g_0}{\partial t} + \mathcal{L}^2 g_1 + \mathcal{L}^1 g_0 + \phi\mathcal{L}^1\zeta + \frac{1}{2}\gamma_1^2 = 0, \tag{5.23}$$

$$-\frac{\partial g_k}{\partial t} + \mathcal{L}^2 g_{k+1} + \mathcal{L}^1 g_k = 0, \tag{5.24}$$

$$-\frac{\partial h_0}{\partial t} + \mathcal{L}^2 h_0 = 0, \tag{5.25}$$

$$-\frac{\partial h_k}{\partial t} + \mathcal{L}^2 h_k + \mathcal{L}^1 h_{k-1} = 0. \tag{5.26}$$

with initial conditions

$$\begin{aligned} g_0(s_1, 0) + h_0(s_1, s_2, 0) &= \zeta(s_1)\phi(s_2) \\ g_k(s_1, s_2, 0) + h_k(s_1, s_2, 0) &= 0 \quad \text{for } k \geqslant 1. \end{aligned}$$

such that $h_i(s_1, s_2, \frac{t}{\epsilon})$ converges to 0 as ϵ converges to 0. In addition, for each fixed triple (s_1, s_2, t), e_k has order $O(\epsilon^{k+1})$.

Remark 5.1.6. Here g_i are the main expansion terms and for each g_i we will choose a specific initial condition, so h_i represents the correction between these chosen initial conditions and the true initial conditions.

Here we choose to deal with this problem using the asymptotic expansion method, as in [22]. We first need the following auxiliary results.

Lemma 5.1.7. *The process S_t^2 has a unique stationary distribution Π.*

Proof. This proof is identical to the proof of Theorem 3.3.1 in Chapter 3. □

With this result we shall use the following notation:

$$\bar{f} = \int f(s) d\Pi(s)$$

We also need the following lemma regarding the related Poisson equations.

Lemma 5.1.8. *For each continuous function ψ such that $\bar{\psi} = 0$, the PDE*

$$\mathcal{L}^2 g = \psi$$

has a solution. Conversely, this PDE has a solution only if $\bar{\psi} = 0$. In addition, the solution is unique up to the addition of a constant in g.

This is a well established result. The proof of a similar result can be found in Lemma 3.1 and Lemma 3.2 of [22].

Now we can start to discuss the existence and uniqueness of solutions to the PDEs for g_i and h_i.

Proof of Lemma 5.1.5. We first consider the leading term (5.23). As g_0 is independent of s_2 we can integrate both sides of (5.23) to obtain

$$-\frac{\partial g_0}{\partial t} + \mathcal{L}^1 g_0 + \left(\int \phi d\Pi \right) \mathcal{L}^1 \zeta + \frac{1}{2}\gamma_1^2 = 0 \tag{5.27}$$

This is because $\int \mathcal{L}^2 g_1 d\Pi = 0$. Now, choosing the initial condition $g_0(s_1, s_2, 0) = \zeta(s_1)\bar{\phi}$, this PDE becomes a parabolic PDE which has a unique solution g_0.

Next, we consider the correction term h_0. Here we have chosen the initial condition $h_0(s_1, s_2, 0) = \zeta(s_1)\phi(s_2) - \zeta(s_1)\bar{\phi}$, so that $g_0(s_1, 0) + h_0(s_1, s_2, 0) = \zeta(s_1)\phi(s_2)$. With

this initial condition the parabolic PDE (5.25) has a unique solution, which has a representation $h_0(s_1, s_2, t) = \hat{h}_0(s_2, t)\zeta(s_1)$ where $\hat{h}_0$ is the unique solution to

$$-\frac{\partial \hat{h}_0}{\partial t} + \mathcal{L}^2 \hat{h}_0 = 0,$$

with initial condition $\hat{h}_0(s_2, 0) = \phi(s_2) - \int \phi(s_2) d\Pi$.

In addition, this choice of initial condition also guarantees that $h_0(s_1, s_2, \frac{t}{\epsilon})$ converges to 0 as ϵ converges to 0. This is because, using the Feynman–Kac theorem, a probabilistic representation of h_0 is given by

$$h_0(s_1, s_2, t) = \zeta(s_1) \left(\mathbb{E}[\phi(S_t^2) | S_0^2 = s_2] - \int \phi d\Pi \right)$$

We can find a Lyapunov function L dominating ϕ as in Section 4.5, and deduce that h_0 converges to 0 as t diverges to infinity.

Next, consider (5.24) and (5.26). We have chosen the initial condition such that $g_k(s_1, s_2, 0) + h_k(s_1, s_2, 0) = 0$. We will consider case $k = 1$ and case $k > 1$ separately. Subtracting (5.27) from (5.23) we get

$$\mathcal{L}^2 g_1 = \left(\int \phi d\Pi \right) \mathcal{L}^1 \zeta - \phi \mathcal{L}^1 \zeta \tag{5.28}$$

Using the notation $\Phi := (\int \phi d\Pi)\mathcal{L}^1\zeta - \phi\mathcal{L}^1\zeta$, as $\bar{\Phi} = 0$, by Lemma 5.1.8 this PDE has a unique solution g_1 up to addition of a function which only depends on t and s_1. Thus, we write this function $g_1 = g_{1,1}(s_1, t) + g_{1,2}(s_1, s_2, t)$ where $g_{1,2}$ is a particular solution of (5.28) satisfying $\bar{g}_{1,2} = 0$. We still need to find the corresponding $g_{1,1}$. Substituting $g_1 = g_{1,1}(s_1, t) + g_{1,2}(s_1, s_2, t)$ into (5.24) with $k = 1$ and integrating both sides with respect to Π we obtain

$$-\frac{\partial g_{1,1}}{\partial t} + \mathcal{L}^1 g_{1,1} + \tilde{\Phi} = 0$$

where $\tilde{\Phi}$ is a function that can be calculated using $g_{1,2}$. To find $g_{1,1}$ we still need to find the initial condition. We integrate (5.26) with $k = 1$ first with respect to Π then integrate from $t = 0$ to $t = \infty$. As we restrict that $h_1(s_1, s_2, t)$ converges to 0 as t diverges to infinity and $\int \mathcal{L}^2 h_1 d\Pi(s_2) = 0$ we obtain

$$\int h_1(s_1, s_2, 0) d\Pi(s_2) = \int_0^\infty \int \mathcal{L}^1 h_0(s_1, s_2, t) d\Pi(s_2)\, dt \tag{5.29}$$

Then using $g_1(s_1, s_2, 0) + h_1(s_1, s_2, 0) = 0$ and $\bar{g}_{1,2} = 0$ we get

$$g_{1,1}(s_1, 0) = \int_0^\infty \int \mathcal{L}^1 h_0(s_1, s_2, t) d\Pi(s_2)\, dt = \mathcal{L}^1 \zeta(s_1) \int_0^\infty \int \hat{h}_0(s_2, t) d\Pi(s_2)\, dt \tag{5.30}$$

Then we can find a unique solution $g_{1,1}$ of (5.28) with initial condition (5.30).

Substituting h_0 into (5.26) and using the initial condition $g_1(s_1, s_2, 0) + h_1(s_1, s_2, 0) = 0$ we obtain a parabolic PDE for h_1 that has a unique solution.

The final step for deducing the asymptotic expansion is the rest of the terms g_k and h_k for $k > 1$. This is almost identical to the case $k = 1$. To find g_k and h_k, we assume that g_{k-1} and h_{k-1} are known. By (5.24) we have

$$-\frac{\partial g_{k-1}}{\partial t} + \mathcal{L}^2 g_k + \mathcal{L}^1 g_{k-1} = 0. \tag{5.31}$$

Integrating (5.31) with respect to Π and subtracting the result from (5.31) we obtain

$$\mathcal{L}^2 g_k = \Phi_k. \tag{5.32}$$

Here Φ_k is a function that can be calculated from μ_1, σ_1, Π and g_{k-1} with $\bar{\Phi}_k = 0$. By Lemma 5.1.8 we deduce that this PDE has a unique solution g_k with $g_k = g_{k,1}(s_1, t) + g_{k,2}(s_1, s_2, t)$ where $\bar{g}_{k,2} = 0$. Substituting this expression into (5.24) and integrate both sides with respect to Π we obtain

$$-\frac{\partial g_{k,1}}{\partial t} + \mathcal{L}^1 g_{k,1} = \hat{\Phi}_k \tag{5.33}$$

where $\hat{\Phi}_k$ can be calculated from μ_1, σ_1, Π and $g_{k,2}$. Finally similarly to before, we integrate (5.26) first with respect to Π then from 0 to infinity and using $\lim_{t\to\infty} h_k(s_1, s_2, t) = 0$, $\int \mathcal{L}^2 h_k d\Pi = 0$ and initial condition $g_k(s_1, s_2, 0) + h_k(s_1, s_2, 0) = 0$ we get the following initial condition for $g_{k,1}$

$$g_{k,1}(s_1, 0) = \int_0^\infty \int \mathcal{L}^1 h_{k-1}(s_1, s_2, t) d\Pi(s_2)\, dt. \tag{5.34}$$

Combining (5.33) and (5.34) we can find a unique solution $g_{k,1}$. Thus, we have the initial condition for h_k, and using (5.26) we can solve the unique solution h_k.

Now the final step is to show that the asymptotic expansion indeed converges to $g^{(0)}$. In

particular, we define the error term e_k as

$$e_k := g^{(0)}(s_1, s_2, t) - \left(g_0(s_1, t) + h_0\left(s_1, s_2, \frac{t}{\epsilon}\right) + \sum_{i=1}^{k} \epsilon^i g_i(s_1, s_2, t) + \sum_{i=1}^{k} \epsilon^i h_i\left(s_1, s_2, \frac{t}{\epsilon}\right)\right).$$

We will show that for fixed (s_1, s_2, t), e_k has order $O(\epsilon^{k+1})$.

We first show that

$$-\frac{\partial e_k}{\partial t} + \mathcal{L}^1 e_k + \frac{1}{\epsilon}\mathcal{L}^2 e_k \tag{5.35}$$

has order $O(\epsilon^k)$ pointwise.

If $k = 0$, we have

$$-\frac{\partial e_0}{\partial t} + \mathcal{L}^1 e_0 + \frac{1}{\epsilon}\mathcal{L}^2 e_0 = -\mathcal{L}^1\hat{\phi} - \frac{1}{2}\gamma_1^2 - \frac{\partial g_0}{\partial t} - \mathcal{L}^1 g_0 - \mathcal{L}^1 h_0.$$

To show that this has order $O(1)$ pointwise, the only problematic term is $\mathcal{L}^1 h_0$. By the representation $h_0(s_1, s_2, \frac{t}{\epsilon}) = \hat{h}_0(s_2, \frac{t}{\epsilon})\zeta(s_1)$, this term has the same order as $\hat{h}_0(s_2, \frac{t}{\epsilon})$ which is $O(1)$ due to the restriction $h_0(s_1, s_2, \frac{t}{\epsilon})$ converges to 0 as ϵ converges to 0.

If $k \geqslant 1$, we have

$$-\frac{\partial e_k}{\partial t} + \mathcal{L}^1 e_k + \frac{1}{\epsilon}\mathcal{L}^2 e_k = \epsilon^k\left(\frac{\partial g_k}{\partial t} - \mathcal{L}^1 g_k - \mathcal{L}^1 h_k\right). \tag{5.36}$$

The only problematic term is still $\mathcal{L}^1 h_k$. We will show this inductively. Assume that h_{k-1} has a representation $h_{k-1}(s_1, s_2, \frac{t}{\epsilon}) = H_{k-1}(s_1)\hat{h}(s_2, \frac{t}{\epsilon})$ for some smooth function H, $\hat{h}(s_2, \frac{t}{\epsilon})$ has order $O(1)$. Then by (5.28) and (5.32) we can choose $g_{k,2}(s_1, s_2, t) = \mathcal{L}^1 H_{k-1}(s_1)\hat{g}_{k,2}(s_2, t)$ for some function $\hat{g}_{k,2}$. Combining this with the initial condition (5.30) and (5.34) and the PDE (5.26) we can deduce that h_k has a representation $h_k(s_1, s_2, \frac{t}{\epsilon}) = \mathcal{L}^1 H_{k-1}(s_1)\hat{h}(s_2, \frac{t}{\epsilon})$. Hence, $\mathcal{L}^1 h_k$ is well defined and has order $O(1)$ pointwise.

Then using the probabilistic representation of e_k given by the Feynman–Kac formula applied to (5.36) with initial condition $e_k(s_1, s_2, 0) = 0$ we can deduce that $e_k(s_1, s_2, t)$ has order $O(\epsilon^k)$.

Finally from the definition we have

$$e_k = e_{k+1} + \epsilon^{k+1} g_{k+1} + \epsilon^{k+1} h_{k+1}.$$

As we have shown that all three terms on the right-hand side have order $O(\epsilon^{k+1})$ pointwise, we have shown that $e_k(s_1, s_2, t)$ has order $O(\epsilon^{k+1})$. □

Remark 5.1.9. For simplicity, we assume that μ_2 is a decreasing linear function throughout this chapter. The results in this chapter can be extended to the case where μ_2 satisfies Hypothesis 4.2.1 with a slight modification of the proofs.

In addition, we can also consider the case where $\tilde{S}_t^2$ starts from a distribution as before and obtain the special case of Theorem 5.1.15. In this case, if $S_t^1 \equiv s_1$, given that S_t^2 has a unique invariant distribution Π, it is clear that $\tilde{S}_t^2$ has a unique invariant distribution Π_{s_1} which is a scaled version of Π.

Theorem 5.1.10. *If we further assume that $S_0^1 = s_1$ and $\tilde{S}_0^2 \sim \Pi_{s_1}$, the maximised expected utility $\bar{V}(x, s_1, t) := \int V(x, s_1, s_2, t) d\Pi_{s_1} = \sup_\pi \mathbb{E}[\log(X_t^\pi) | S_0^1 = s_1, S_0^2 \sim \Pi]$ satisfies*

$$\lim_{\epsilon \to 0} \left(\bar{V}(x, s_1, t) - \frac{1}{\epsilon} \mathbb{E}\left[\int_0^t \lambda(S_u^1) du | S_0^1 = s_1 \right] \right) = \log(x) + E\left[\int_0^t \gamma_1(S_u^1) du | S_0^1 = s_1 \right]. \tag{5.37}$$

The proof of this theorem will be separated into several lemmas.

Lemma 5.1.11. *S_t^2 converges to its unique invariant distribution Π in distribution $\forall t > 0$ as ϵ converges to 0.*

Proof. This is identical to the one-dimensional problem as the dynamics of S_t^2 is independent of S_t^1. □

Using the notation $\bar{\phi}(S_t^1) := \int \phi(S_t^1, s_2) d\Pi(s_2)$, we have the following result:

Lemma 5.1.12. *$\mathbb{E}[\hat{\phi}(S_t^1, S_t^2) | S_0^1 = s_1, \hat{S}_0^2 \sim \Pi_{s_1}]$ converges to $\mathbb{E}[\bar{\phi}(S_t^1) | S_0^1 = s_1]$ and $\mathbb{E}[\int_0^t \mathcal{L}^1 \hat{\phi}(S_u^1, S_u^2) du | S_0^1 = s_1, \hat{S}_0^2 \sim \Pi_{s_1}]$ converges to $\mathbb{E}[\int_0^t \mathcal{L}^1 \bar{\phi}(S_u^1) du | S_0^1 = s_1]$ as ϵ converges to 0.*

Proof. We have shown that $\hat{\phi}(s_1, s_2) = \zeta(s_1)\phi(s_2)$. Also, as S_t^1 and S_t^2 are independent, we can write

$$\mathbb{E}[\hat{\phi}(S_t^1, S_t^2) | S_0^1 = s_1, \hat{S}_0^2 \sim \Pi_{s_1}] = \mathbb{E}[\zeta(S_t^1) \mathbb{E}[\phi(S_t^2) | S_0^2 \sim \Pi] | S_0^1 = s_1] \tag{5.38}$$

and

$$\mathbb{E}\left[\int_0^t \mathcal{L}^1\hat{\phi}(S_u^1, S_u^2)du\Big|S_0^1 = s_1, \hat{S}_0^2 \sim \Pi_{s_1}\right] = \mathbb{E}\left[\int_0^t \mathcal{L}^1\zeta(S_u^1)\mathbb{E}[\phi(S_u^2)\Big|S_0^2 \sim \Pi]du|S_0^1 = s_1\right] \tag{5.39}$$

Hence, this lemma follows directly from the result of the one-dimensional case. □

Lemma 5.1.13. $\bar{\phi}(S_t^1) := \int \phi(S_t^1, s_2)d\Pi(s_2)$ *satisfies the following equality:*

$$\mathbb{E}\left[\int_0^t \mathcal{L}^1\bar{\phi}(S_u^1)du + \bar{\phi}(S_0^1) - \bar{\phi}(S_t^1)|S_0^1 = s_1\right] = 0 \tag{5.40}$$

Proof. This clearly holds as $\bar{\phi}$ equals a constant times the smooth function ζ, and $\mathcal{L}^1$ is the generator of the process S_t^1. □

Combining these lemmas, we have proved Theorem 5.1.10.

5.1.0.1 Extension to More General Case

We will next extend this result to the general case where $\mu_2(s_1, s_2)$ no longer equals $\zeta(s_1)\mu_2(s_2)$. Instead, we assume that for each fixed s_1, $\mu_2(s_1, s_2)$ is a linear decreasing function of s_2. In this case, we only sketch the proof of extended version of Theorem 5.1.4 and Theorem 5.1.10 and point out the parts of the proof that are still open problems. In particular, we expect the following results:

Hypothesis 5.1.14. Assume that $\sigma_1 > 0$ and $\sigma_2 > 0$ are constants and μ_1 is smooth and bounded. Furthermore, assume that for each fixed s_1, μ_2 is a decreasing linear function of s_2. Then there is a unique solution $g(s_1, s_2, t)$ which solves (5.8) and satisfies

$$\lim_{\epsilon\to 0}\left(g(s_1, s_2, t) - \frac{1}{\epsilon}\mathbb{E}\left[\int_0^t \lambda(S_u^1)du|S_0^1 = s_1\right]\right) = \phi(s_1, s_2) + c(s_1). \tag{5.41}$$

where (λ, ϕ) is a solution to parametrized ergodic PDE

$$\mathcal{L}^2\phi(s_1, s_2) + \frac{1}{2}\gamma_2^2(s_1, s_2) = \lambda(s_1). \tag{5.42}$$

Hypothesis 5.1.15. Assume that, for each fixed $S_t^1 \equiv s_1$, S_t^2 has a unique invariant distribution Π_{s_1}. Define

$$\bar{V}(x, s_1, t) := \int V(x, s_1, s_2, t)d\Pi_{s_1} = \sup_\pi \mathbb{E}[\log(X_t^\pi)|S_0^1 = s_1, S_0^2 \sim \Pi_{s_1}]$$

where X_t^π, S_t^1 and S_t^2 have dynamics (5.4), (5.1) and (5.2).

Then

$$\begin{aligned}&\lim_{\epsilon\to 0}\left(\bar{V}(x,s_1,t)-\frac{1}{\epsilon}\mathbb{E}\left[\int_0^t \lambda(S_u^1)du|S_0^1=s_1\right]\right)\\ =&\log(x)+E\left[\int_0^t \gamma_1(S_u^1)du|S_0^1=s_1\right].\end{aligned}\tag{5.43}$$

Sketch of proof of Hypothesis 5.1.14. The proof of this theorem is almost identical to the proof of Theorem 5.1.4. We shall still assume that the function g contains an asymptotic term with order $\frac{1}{\epsilon}$, an ergodic term and a correction term independent of s_2. In particular, we expect that g has the expansion:

$$g(s_1,s_2,t)=\frac{1}{\epsilon}g^{(-1)}(s_1,s_2,t)+\hat{\phi}(s_1,s_2)+g^{(0)}(s_1,s_2,t)$$

Substituting this into (5.8) we still get (5.13)-(5.17).

The main difficulty in this case is (5.18) becomes

$$\lambda(s_1)=\mathcal{L}^2\hat{\phi}+\frac{1}{2}\gamma_2^2(s_1,s_2)\tag{5.44}$$

In this case we can no longer separate a function of s_1 from the solution $(\lambda,\hat{\phi})$. As a result, we need to find a condition such that this solution behaves well with respect to the parameter s_1. Otherwise, (5.15) is not well defined if ϕ is not twice differentiable with respect to s_1. Assuming this can be overcome, we apply the Feynman–Kac theorem to (5.15) with initial condition (5.17) to obtain

$$g^{(0)}=\mathbb{E}\left[\int_0^t \mathcal{L}^1\hat{\phi}(S_u^1,S_u^2)+\frac{1}{2}\gamma_1^2(S_u^1)du-\hat{\phi}(S_t^2,S_t^2)|S_0^1=s_1,S_0^2=s_2\right],\tag{5.45}$$

and apply parametrized ergodic theory to deduce this converges to $c(s_1)$ if $\hat{\phi}$ satisfies suitable integrability and boundedness conditions. The suitable choice of such parameters is still an open problem. □

Sketch of proof of Hypothesis 5.1.15. This proof is much more difficult than the proof of Theorem 5.1.10. Write $\bar{V}(x,s_1,t)=\log(x)+\bar{g}(s_1,t)$ where $\bar{g}(s_1,t):=\mathbb{E}[g(s_1,s_2,t)|s_2\sim$

$\Pi_{s_1}]$. By (5.41) and (5.45) we have

$$\bar{g}(s_1,t) = \frac{1}{\epsilon}\mathbb{E}\left[\int_0^t \lambda(S_u^1)du|S_0^1 = s_1\right] + E\left[\int_0^t \frac{1}{2}\gamma_1^2(S_u^1)du|S_0^1 = s_1\right] + \mathbb{E}\left[\int_0^t \mathcal{L}^1\hat{\phi}(S_u^1,S_u^2)du + \hat{\phi}(S_0^1,S_0^2) - \hat{\phi}(S_t^1,S_t^2)|S_0^1 = s_1, S_0^2 \sim \Pi_{s_1}\right].$$

The first step is to show that the distribution of S_t^2 converges to $\Pi_{S_t^1}$ uniformly in time t as ϵ converges to 0.

We choose to show the convergence of the distribution of S_t^2 using discretization. Consider a discontinuous discretized version ζ_t^2 of S_t^2 with dynamics

$$\begin{aligned}\zeta_t^2 &= S_t^2 \text{ for } t = n\delta, n \in \mathbb{N},\\ d\zeta_t^2 &= \frac{1}{\sqrt{\epsilon}}\sigma_2(S_{n\delta}^1,\zeta_t^2)\,dW_t^2 + \frac{1}{\epsilon}\mu_2(S_{n\delta}^1,\zeta_t^2)\,dt \text{ for } t \in [n\delta,(n+1)\delta).\end{aligned}$$

If the probability density of the distribution Π_{s_1} is uniformly Lipschitz continuous in s_1 and S_t^1 had uniformly Lipschitz continuous paths, we expect that the distribution of ζ_t^2 should converge to $\Pi_{S_t^1}$ as both ϵ and δ converge to 0 with $\delta \gg \epsilon$. This is because, for sufficiently small δ, S_t^1 will not change much and, for even smaller ϵ, ζ_u^2 will be close to $\Pi_{S_{n\delta}^1}$ on every interval $[n\delta,(n+1)\delta)$.

Remark 5.1.16. A detailed proof of a similar result can be found in [16] and a sufficient condition for convergence in distribution of S_t^2 typically includes the following conditions: μ_1 and σ_1 are Lipschitz continuous, σ_2 is bounded, μ_2 and σ_2 are smooth and have bounded derivatives of all orders, there exist positive constants c_1, c_2 such that $s_2\mu_2(s_1,s_2) \le c_1(s_2)^2$ for all s_1 (i.e. a recurrence condition for S_t^2) and $\mu_2(s_1,s_2)^2 > c_2 > 0$. These guarantee the non-degeneracy of the process S_t^2. Note that these conditions also imply the existence of a unique invariant probability measure $\Pi_{S_t^1}$.

Then we can show that $\hat{\phi}(S_t^1,S_t^2)$ converges to $\bar{\phi}(S_t^1) := \int \hat{\phi}(S_t^1,S_t^2)d\Pi_{S_t^1}$ using ergodic theory if we can show that $\hat{\phi}$ is sufficiently smooth and bounded. Hence, we can rewrite the last term of the probabilistic representation of $\bar{g}$ as

$$\mathbb{E}\left[\int_0^t \mathcal{L}^1\bar{\phi}(S_u^1)du + \bar{\phi}(S_0^1) - \bar{\phi}(S_t^1)|S_0^1 = s_1\right]. \tag{5.46}$$

If we can further show that this limit function $\bar{\phi}$ is also smooth and bounded in s_1, then,

as $\mathcal{L}^1$ is the generator of S_t^1, (5.46) equals 0 from the definition of the generator, and we obtain

$$\lim_{\epsilon \to 0} \left(\bar{g}(t, s^1) - \frac{1}{\epsilon} \mathbb{E} \left[\int_0^t \lambda(S_u^1) du | S_0^1 = s_1 \right] \right) = E \left[\int_0^t \frac{1}{2} \gamma_1^2(S_u^1) du | S_0^1 = s_1 \right].$$

which finishes the proof. □

Note that if S_t^1 is a constant, this result coincides with (4.28) in the one-dimensional case.

Remark 5.1.17. Consider that we are trading in a market with only an asset S_t^1 with dynamics (5.1) and a bank account $B_t = 1$. Our controlled wealth process X_t^π is self-financing and has dynamics

$$dX_t^\pi = \pi_t \, dS_t^1.$$

Then for a natural choice of admissible set $\mathcal{A}$, i.e. such that our wealth process X_t^π is non negative, the maximised expected utility

$$\hat{V}(x, s_1, t) := \sup_{\pi \in \mathcal{A}} \mathbb{E}[\log(X_t^\pi) | S_0^1 = s_1],$$

equals the RHS of (5.43). This means that, asymptotically, we get a clear separation of the wealth from trading the asset S_t^1 and the wealth we asymptotically gain due to the effect of noise, if we suitably average the maximised expected utility.

5.2 A Simple Example of Extension to a Non-linear Market

In this and the next two sections, we briefly discuss the most interesting case in which the assets in our approximated market are not linearly dependent and hence our lower-dimensional market without noise possibly contains arbitrage. In particular, in this section, we assume that the dynamics of stocks are defined by

$$S_t^3 = f_1(S_t^1) + \zeta_1(S_t^1) S_t^2,$$
$$S_t^4 = f_2(S_t^1) + \zeta_2(S_t^1) S_t^2,$$

where the underlying process S_t^1, S_t^2 have dynamics:

$$dS_t^1 = \sigma_1(S_t^1)\, dW_t^1 + \mu_1(S_t^1)\, dt, \tag{5.47}$$

$$d\hat{S}_t^2 = \frac{1}{\sqrt{\epsilon}}\sigma_2(\hat{S}_t^2)\, dW_t^2 + \frac{1}{\epsilon}\mu_2(\hat{S}_t^2)\, dt, \tag{5.48}$$

$$S_t^2 = \delta \hat{S}_t^2, \tag{5.49}$$

$$B_t = 1. \tag{5.50}$$

Here, the vector $(f_1(S_t^1), f_2(S_t^1))'$ is the approximated asset price, moving along a lower dimensional submanifold, and $(\zeta_1(S_t^1)S_t^2, \zeta_2(S_t^1)S_t^2)'$ is the additional noise factor. The constant ϵ controls the rate of mean reversion and δ controls the magnitude of the noise factor. This kind of market model can be applied when approximating market containing two European call options written on the same stock. We will see, in this section, that the utility maximisation problem in such a market is related to the utility maximisation problem in a linear market whose assets have the same dynamics as S_t^1 and S_t^2, plus an additional drift independent of the rate of mean reversion ϵ.

Before we start to consider the general case, we first consider the special case where $f_1(x) = x$, to get intuitive understanding. Let

$$\begin{aligned} S_t^3 &= S_t^1 + \zeta_1(S_t^1)S_t^2, \\ S_t^4 &= f(S_t^1) + \zeta_2(S_t^1)S_t^2. \end{aligned}$$

Assume that f_i and ζ_i are twice continuously differentiable. By Itô's formula, we obtain

$$\begin{aligned} dS_t^3 &= dS_t^1 + \zeta_1\, dS_t^2 + \zeta_1' S_t^2\, dS_t^1 + \frac{1}{2}\zeta_1'' S_t^2\, d[S_t^1], \\ dS_t^4 &= f'\, dS_t^1 + \frac{1}{2}f''\, d[S_t^1] + \zeta_2\, dS_t^2 + \zeta_2' S_t^2\, dS_t^1 + \frac{1}{2}\zeta_2'' S_t^2\, d[S_t^1]. \end{aligned}$$

Our wealth process $X_t^{\hat{\pi}}$ is a self-financing process with dynamics

$$dX_t^{\hat{\pi}} = \hat{\pi}_t^1\, dS_t^3 + \hat{\pi}_t^2\, dS_t^4 \tag{5.51}$$

where $\hat{\pi}_t^1$ and $\hat{\pi}_t^2$ are the number of shares of asset S_t^1 and S_t^2 we hold at time t. Our goal is still to maximise the expected log utility, that is, finding

$$V(s_1, \hat{s}_2, t) = \sup_{\hat{\pi}\in\mathcal{A}} \mathbb{E}[\log(X_t^{\hat{\pi}}) | S_0^1 = s_1, \hat{S}_0^2 = \hat{s}_2],$$

where $\mathcal{A}$ is the set of admissible controls π, such that (5.51) has a unique strong solution $X_t^{\hat{\pi}}$ and $\mathbb{E}[\sup_{0<t<T} \log(X_t^\pi)|S_0^1 = s_1, \hat{S}_0^2 = \hat{s}_2] < \infty \quad \forall T$.

For simplicity, we use the notation

$$\begin{aligned} h_{11}(s_1, \hat{s}_2) &= 1 + \delta \hat{s}_2 \zeta_1'(s_1), \\ h_{12}(s_1) &= \zeta_1(s_1), \\ h_{21}(s_1, \hat{s}_2) &= f'(s_1) + \delta \hat{s}_2 \zeta_2'(s_1), \\ h_{22}(s_1) &= \zeta_2(s_1). \end{aligned}$$

Hence, if we define

$$\begin{aligned} \pi_t^1 &= \hat{\pi}_t^1 h_{11}(S_t^1, \hat{S}_t^2) + \hat{\pi}_t^2 h_{21}(S_t^1, \hat{S}_t^2) \\ \pi_t^2 &= \hat{\pi}_t^1 h_{12}(S_t^1) + \hat{\pi}_t^2 h_{22}(S_t^1) \end{aligned}$$

we can write

$$\begin{aligned} dX_t^{\hat{\pi}} =& \pi_t^1 \sigma_1 \, dW_t^1 + \frac{\delta}{\sqrt{\epsilon}} \pi_t^2 \sigma_2 \, dW_t^2 + \pi_t^1 \mu_1 \, dt + \frac{\delta}{\epsilon} \pi_t^2 \mu_2 \, dt \\ &+ \frac{1}{2} (\hat{\pi}_t^2 f'' \sigma_1^2 + \hat{\pi}_t^1 \delta \hat{S}_t^2 \zeta_1'' \sigma_1^2 + \hat{\pi}_t^2 \delta \hat{S}_t^2 \zeta_2'' \sigma_1^2) \, dt \end{aligned} \tag{5.52}$$

For notation simplicity we drop the argument of h_{ij} in the rest of this chapter.

Remark 5.2.1. The first four terms correspond to the wealth process of X_t^π when $dX_t^\pi = \pi_t^1 \, dS_t^1 + \pi_t^2 \, dS_t^2$, which is the wealth process if we trade in a linear market with assets S_t^1, S_t^2 and bank account B_t. Each of the three remaining terms adds an additional drift to our wealth process. The fifth term $\frac{1}{2} \hat{\pi}_t^2 f'' \sigma_1^2 \, dt$ corresponds to the wealth we obtain from non-linearity of our approximated market. In addition, this is also the term that makes our wealth diverge if there is no noise. The last two terms come from the wealth we gain due to changes of direction of noise. These two terms will vanish if we assume that the direction of the noise is independent of the factor S_t^1.

If there is no additional noise, that is, $\delta = 0$, the second, fourth and last two terms of (5.52) vanish. Using the relationship $\hat{\pi}_t^2 = \frac{h_{11}\pi_t^2 - h_{12}\pi_t^1}{h_{11}h_{22} - h_{12}h_{21}}$ with $\delta = 0$, (5.52) becomes

$$dX_t^{\hat{\pi}} = \pi_t^1 \sigma_1 \, dW_t^1 + \pi_t^1 \mu_1 \, dt - \frac{1}{2} \frac{h_{12}\pi_t^1}{h_{11}h_{22} - h_{12}h_{21}} f'' \sigma_1^2 \, dt + \frac{1}{2} \frac{h_{11}\pi_t^2}{h_{11}h_{22} - h_{12}h_{21}} f'' \sigma_1^2 \, dt. \tag{5.53}$$

From this we see that there is no randomness controlled by π_t^2, so a stochastic control π_t^2 such that $\frac{h_{11}\pi_t^2}{h_{11}h_{22}-h_{12}h_{21}}f''\sigma_1^2 > 0$ can be made as large as possible, which will cause our wealth process to explode. Therefore, the corresponding maximised expected utility will also diverge.

Now we start to analyse the HJB equation related to this problem. Assume that we can still write $V(s_1, \hat{s}_2, x, t) = \log(x) + g(s_1, s_2, t)$. Using the notation

$$\mathcal{L}^1 f := \mu_1 \frac{\partial f}{\partial s_1} + \frac{1}{2}\sigma_1 \frac{\partial^2 f}{\partial s_1^2},$$
$$\mathcal{L}^2 f := \frac{1}{\epsilon}\mu_2 \frac{\partial f}{\partial \hat{s}_2} + \frac{1}{2\epsilon}\sigma_2 \frac{\partial^2 f}{\partial \hat{s}_2^2},$$

and assuming the optimal feedback control exist and is admissible, we can simplify the HJB equation related to this problem to

$$-\frac{\partial g}{\partial t} + \mathcal{L}^1 g + \mathcal{L}^2 g + H = 0, \tag{5.54}$$

for some Hamiltonian function H which is a deterministic function of s_1, s_2 and t.

5.3 Perpendicular Noise

To further simplify the problem, we can choose the noise factor to be perpendicular to the $\hat{S}_t^1$ axis, that is $\zeta_1 = 0$ and $\zeta_2 = 1$. This assumption is reasonable if we want to consider an approximated market which has some fast oscillation behaviour. In this case, there may not be a bijection between (S_t^1, S_t^2) and (S_t^3, S_t^4) if the noise factor depends on S_t^1 (i.e. ζ_1 and ζ_2 are not constants).

Another advantage of this choice of noise factor is that from (5.52) we can see that the nonlinearity of the model induces a drift perpendicular to the $\hat{S}_t^1$ axis. Choosing noise in the same direction, we can balance all of this drift so that there is no remaining drift projected in the direction of S_t^1.

Using

$$\hat{\pi}_t^1 = \pi_t^1 - f'\pi_t^2,$$
$$\hat{\pi}_t^2 = \pi_t^2,$$

the function H in (5.54) has the representation

$$H = \frac{\mu_1^2}{2\sigma_1^2} + \frac{(\frac{\delta}{\epsilon}\mu_2 + f''\sigma_1^2)^2}{2\frac{\delta^2}{\epsilon}\sigma_2^2}.$$

We can expand H to different orders of δ and ϵ. First, (5.54) can be simplified to

$$-\frac{\partial g}{\partial t} + \mathcal{L}^1 g + \mathcal{L}^2 g + \frac{1}{2\epsilon}\frac{\mu_2^2}{\sigma_2^2} + \frac{1}{\delta}\frac{\mu_2}{\sigma_2^2} f''\sigma_1^2 + \frac{1}{2}\frac{\mu_1^2}{\sigma_1^2} + \frac{\epsilon(f'')^2\sigma_1^2}{2\delta^2} = 0. \tag{5.55}$$

If we further assume that ϵ is much smaller than δ^2, which means that the rate of mean reversion of the noise is much larger than the inverse of the magnitude of the noise, the last term converges to 0 as ϵ and δ converges to 0.

Remark 5.3.1. If we only consider the PDE without higher order term $\frac{\epsilon(f'')^2\sigma_1^2}{2\delta^2}$, we obtain the HJB equation corresponding to the following stochastic control problem:

Maximise expected log utility problem in a linear market with assets $\tilde{S}_t^1$ and $\tilde{S}_t^2$ with dynamic

$$\begin{aligned}
d\tilde{S}_t^1 &= \sigma_1(\tilde{S}_t^1)\, dW_t^1 + \mu_1(\tilde{S}_t^1)\, dt, \\
d\hat{\tilde{S}}_t^2 &= \frac{1}{\sqrt{\epsilon}}\sigma_2(\hat{\tilde{S}}_t^2)\, dW_t^2 + \frac{1}{\epsilon}\mu_2(\hat{\tilde{S}}_t^2)\, dt + \frac{1}{\delta} r\, dt, \\
\tilde{S}_t^2 &= \delta\hat{\tilde{S}}_t^2,
\end{aligned}$$

where $r := f''\sigma_1^2$ is a function only dependent of $\tilde{S}_t^1$. This suggests that the arbitrage due to the nonlinearity of lower dimensional market has the same effect as adding an additional drift in our market, at least in a limiting sense.

We expect that we can find the asymptotic behaviour of solution to (5.55) using an expansion method similar to before. In particular, we expect that the function g has an expansion

$$g = \hat{g} + \frac{1}{\delta} g_\delta = \frac{1}{\epsilon} g_\epsilon + \phi + g_0 + \frac{1}{\delta} g_\delta + (\text{higher order terms})$$

As $\mathcal{L}^1$ and $\mathcal{L}^2$ are independent of δ, comparing order of δ we obtain

$$-\frac{\partial g_\delta}{\partial t} + \mathcal{L}^1 g_\delta + \mathcal{L}^2 g_\delta + \frac{\mu_2}{\sigma_2^2} f''\sigma_1^2 = 0 \tag{5.56}$$

and

$$-\frac{\partial \hat{g}}{\partial t} + \mathcal{L}^1 \hat{g} + \mathcal{L}^2 \hat{g} + \frac{1}{2\epsilon}\frac{\mu_2^2}{\sigma_2^2} + \frac{1}{2}\frac{\mu_1^2}{\sigma_1^2} + \frac{\epsilon (f'')^2 \sigma_1^2}{2\delta^2} = 0. \tag{5.57}$$

(5.56) is a classical singular perturbation problem. If we assume that we can apply the Feynman–Kac representation to this formula with initial condition $g_\delta(s_1, \hat{s}_2, 0) = 0$, we can see that a probabilistic representation of the solution is

$$g_\delta = \mathbb{E}\left[\int_0^t \frac{\mu_2}{\sigma_2^2} f'' \sigma_1^2 du | S_0^1 = s_1, \hat{S}_0^2 = \hat{s}_2\right]. \tag{5.58}$$

By the ergodic theory, we expect that the distribution of S_t^2 converges to its invariant distribution in a short time; hence (5.58) converges to $c_1 E[\int_0^t f'' \sigma_1^2 du | S_0^1 = s_1]$ as ϵ converges to 0 for some constant c_1.

(5.57) is almost identical to the HJB equation (5.9) we have discussed before, except for an additional term which will vanish as ϵ and δ converge to 0 with $\epsilon \ll \delta^2$. Due to the linearity of (5.57) by applying the Feynman–Kac formula we can at least formally deduce that $\hat{g}$ converges to $\frac{\lambda t}{\epsilon} + \phi + c_2 + \mathbb{E}[\frac{1}{2}\frac{\mu_1^2}{\sigma_1^2} | S_0^1 = s_1]$ where (λ, ϕ) is a solution to the ergodic PDE

$$\lambda + \mathcal{L}^2 \phi + \frac{1}{2}\frac{\mu_2^2}{\sigma_2^2} = 0 \tag{5.59}$$

and c_2 is a constant dependent of ϕ.

Combining these results, we expect that our maximised expected value of log utility satisfies

$$\begin{aligned} &\lim_{\epsilon \to 0, \delta \to 0, \epsilon \ll \delta^2} \left(V(x, s_1, \hat{s}_2, t) - \frac{\lambda t}{\epsilon} - \frac{1}{\delta} c_1 E\left[\int_0^t f'' \sigma_1^2 du \middle| S_0^1 = s_1\right]\right) \\ &= \log(x) + \phi + c_2 + \mathbb{E}\left[\frac{1}{2}\frac{\mu_1^2}{\sigma_1^2}\middle| S_0^1 = s_1\right]. \end{aligned} \tag{5.60}$$

In addition, defining $\bar{V}(x, s_1, t) := \int V(x, s_1, \hat{s}_2, t) d\Pi = \sup_\pi \mathbb{E}[\log(X_t^\pi) | S_0^1 = s_1, S_0^2 \sim \Pi]$

where Π is the unique invariant distribution of $\hat{S}_t^2$, we expect that

$$\lim_{\epsilon\to 0,\delta\to 0,\epsilon\ll\delta^2}\left(\bar{V}(x,s_1,\hat{s}_2,t)-\frac{\lambda t}{\epsilon}-\frac{1}{\delta}c_1 E\left[\int_0^t f''\sigma_1^2 du\middle| S_0^1=s_1\right]\right) \tag{5.61}$$

$$=\log(x)+\mathbb{E}\left[\frac{1}{2}\frac{\mu_1^2}{\sigma_1^2}\middle| S_0^1=s_1\right]. \tag{5.62}$$

This means that if we make a suitable choice of our noise factor and remove the exploding terms from the maximised expected utility of this model, we asymptotically gain the maximised expected utility of trading only S_t^1, which corresponds to the only arbitrage-free market in the time-independent market where the assets S_t^1 and S_t^2 linearly depend on each other.

5.4 A More General Extension to a Non-linear Market

In this section we move to the general case where both assets depend on the market factor S_t^1 non-linearly. Suppose

$$S_t^3 = f_1(S_t^1)+\zeta_1(S_t^1)S_t^2,$$
$$S_t^4 = f_2(S_t^1)+\zeta_2(S_t^1)S_t^2.$$

The main aim of this section is to find a suitable way to model the noise in our statistical model, i.e. the choice of functions ζ_1 and ζ_2. We will see that the asymptotic behaviour of maximised expected utility highly depends on our choice of these functions. As a result, we cannot generally obtain a similar result as before, where removing the exploding terms and averaging the remaining terms with the invariant distribution of S_t^2, we obtained the maximised expected utility from trading in a linear market. However, we can see that if we choose to model the noise in a certain way, this result still holds with an error of order $O(1)$, and this error vanishes for some specific choices of f_1 and f_2. In particular, such model of noise is convenient for us to remove the effect of arbitrage in our statistical model.

Again using notation

$$\begin{aligned} h_{11}(s_1, \hat{s}_2) &= f_1'(s_1) + \delta \hat{s}_2 \zeta_1'(s_1), \\ h_{12}(s_1) &= \zeta_1(s_1), \\ h_{21}(s_1, \hat{s}_2) &= f_2'(s_1) + \delta \hat{s}_2 \zeta_2'(s_1), \\ h_{22}(s_1) &= \zeta_2(s_1), \end{aligned}$$

and

$$\begin{aligned} \pi_t^1 &= \hat{\pi}_t^1 h_{11}(S_t^1, \hat{S}_t^2) + \hat{\pi}_t^2 h_{21}(S_t^1, \hat{S}_t^2) \\ \pi_t^2 &= \hat{\pi}_t^1 h_{12}(S_t^1) + \hat{\pi}_t^2 h_{22}(S_t^1) \end{aligned}$$

(5.52) becomes

$$\begin{aligned} dX_t^{\hat{\pi}} =& \pi_t^1 \sigma_1 \, dW_t^1 + \frac{\delta}{\sqrt{\epsilon}} \pi_t^2 \sigma_2 \, dW_t^2 + \pi_t^1 \mu_1 \, dt + \frac{\delta}{\epsilon} \pi_t^2 \mu_2 \, dt \\ &+ \frac{1}{2} (\hat{\pi}_t^1 f_1'' \sigma_1^2 + \hat{\pi}_t^2 f'' \sigma_1^2 + \hat{\pi}_t^1 S_t^2 \zeta_1'' \sigma_1^2 + \hat{\pi}_t^2 S_t^2 \zeta_2'' \sigma_1^2) \, dt. \end{aligned} \tag{5.63}$$

Using this expression and Figure 5.1 we can see how our wealth process evolves as the control and noise factor changes.

If we set the noise factor $(\zeta_1(S_t^1)S_t^2, \zeta_2(S_t^1)S_t^2)$ to 0, changing S_t^1 will make our asset prices move along (f_1', f_2'), the tangent direction of the lower-dimensional submanifold. We also need to consider the dependence of S_t^1 of the noise factor. For fixed S_t^2, changing S_t^1 will make the noise factor $(\zeta_1(S_t^1)S_t^2, \zeta_2(S_t^1)S_t^2)$ move along (ζ_1', ζ_2') and the magnitude of this move is scaled by $\delta \hat{S_t^2}$. Hence, if only S_t^1 changes while S_t^2 is constant, the prices in our market will move along the blue vector V_3 which is exactly the vector (h_{11}, h_{22}).

Then we consider the additional drift terms.

If there is no additional noise factor $(\zeta_1(S_t^1)S_t^2, \zeta_2(S_t^1)S_t^2)$, the non-linear dependence of S_t^3 and S_t^4 on S_t^1 induces additional drift with magnitude $(f_1'' \sigma_1^2, f_2'' \sigma_1^2)$ given by the quadratic variation term from Itô's formula. Finally, the non-linear dependence of the noise factor on S_t^1 also induces an additional small variation in the direction of (ζ_1'', ζ_2'') and scaled by $\delta \hat{S}_t^2 \sigma_1^2$. Hence, compared to the linear case, the market model contains an additional drift

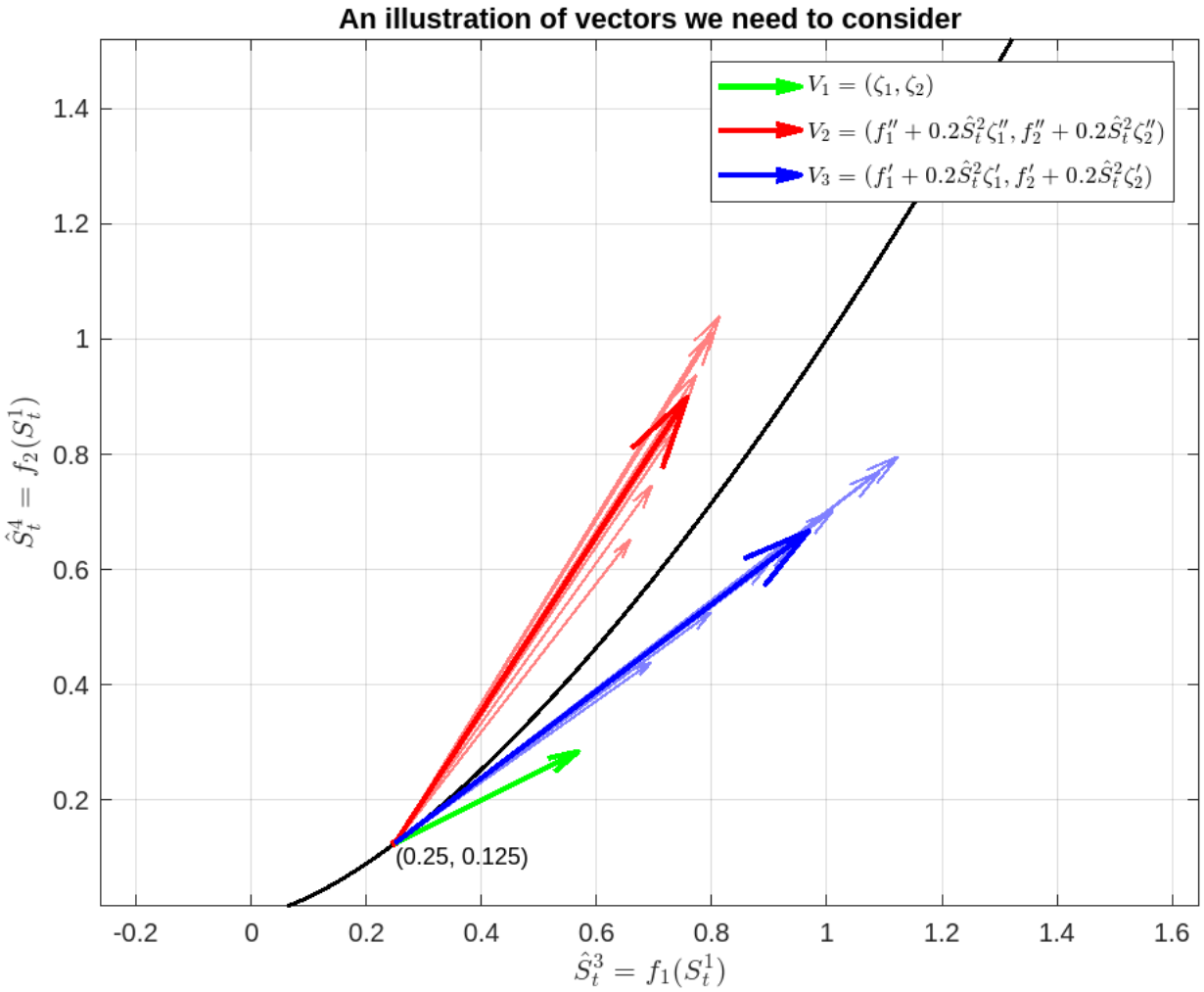


Figure 5.1: An example of the directions vectors we need to consider in a market with $f_1(s) = s^2$, $f_2(s) = s^3$, $\zeta_1(s) = s^{1.5}$ and $\zeta_2(s) = s^{2.5}$. $S_t^1 = 0.5$, $\delta = 0.15$ and I chose 10 realizations for $\hat{S}_t^2 \sim \mathcal{N}(0,1)$ and highlighted the realization $\hat{S}_t^2 = 1.0289$. The black curve is the underlying submanifold. The green arrow V_1 is the noise. The red arrow V_2 is the drift caused by non-linearity of the market. The blue arrow V_3 is the the direction our prices changes if we change S_t^1. Both V_2 and V_3 are perturbed by the non-zero noise factor. Scales of vectors have been adjusted so that they are visually comparable on the same plot.

in the direction of the red vector V_2 and this drift due to nonlinearity will result in an arbitrage if not treated properly.

We hope that in a similar way to the previous case we can choose to model the noise in the same direction as this drift $\frac{1}{2}(\hat{\pi}_t^1 f_1'' \sigma_1^2 + \hat{\pi}_t^2 f'' \sigma_1^2 + \hat{\pi}_t^1 \delta \hat{S}_t^2 \zeta_1'' \sigma_1^2 + \hat{\pi}_t^2 \delta \hat{S}_t^2 \zeta_2'' \sigma_1^2)\, dt$ so that this drift does not affect how we trade with respect to S_t^1. In particular, we want to choose (ζ_1, ζ_2) so that this drift term is independent of π_t^1, which means that this drift does not affect how we choose π_t^1. Using

$$\begin{aligned}\hat{\pi}_t^1 &= \frac{h_{22}\pi_t^1 - h_{21}\pi_t^2}{h_{11}h_{22} - h_{12}h_{21}},\\ \hat{\pi}_t^2 &= \frac{h_{11}\pi_t^2 - h_{12}\pi_t^1}{h_{11}h_{22} - h_{12}h_{21}}.\end{aligned}$$

Substituting this into (5.63), the drift term becomes

$$\begin{aligned}&\frac{\pi_t^1((f_1''\sigma_1^2 + \delta\hat{S}_t^2\zeta_1''\sigma_1^2)\zeta_2 - (f_2''\sigma_1^2 + \delta\hat{S}_t^2\zeta_2''\sigma_1^2)\zeta_1)}{(f_1' + \delta\hat{S}_t^2\zeta_1')\zeta_2 - (f_2' + \delta\hat{S}_t^2\zeta_2')\zeta_1}\\ &+\frac{\pi_t^2((f_2''\sigma_1^2 + \delta\hat{S}_t^2\zeta_2''\sigma_1^2)(f_1' + \delta\hat{S}_t^2\zeta_1') - (f_1''\sigma_1^2 + \delta\hat{S}_t^2\zeta_1''\sigma_1^2)(f_2' + \delta\hat{S}_t^2\zeta_2'))}{(f_1' + \delta\hat{S}_t^2\zeta_1')\zeta_2 - (f_2' + \delta\hat{S}_t^2\zeta_2')\zeta_1}\end{aligned} \tag{5.64}$$

As π_t^1 represents the trading of a linear combination of assets in the direction V_3, and π_t^2 represents trading a linear combination of assets in the direction V_1, we decompose the additional drift into two components.

One component has magnitude

$$\frac{(f_1''\sigma_1^2 + \delta\hat{S}_t^2\zeta_1''\sigma_1^2)\zeta_2 - (f_2''\sigma_1^2 + \delta\hat{S}_t^2\zeta_2''\sigma_1^2)\zeta_1}{(f_1' + \delta\hat{S}_t^2\zeta_1')\zeta_2 - (f_2' + \delta\hat{S}_t^2\zeta_2')\zeta_1}$$

and is in direction V_3. Another component has magnitude

$$\frac{(f_2''\sigma_1^2 + \delta\hat{S}_t^2\zeta_2''\sigma_1^2)(f_1' + \delta\hat{S}_t^2\zeta_1') - (f_1''\sigma_1^2 + \delta\hat{S}_t^2\zeta_1''\sigma_1^2)(f_2' + \delta\hat{S}_t^2\zeta_2')}{(f_1' + \delta\hat{S}_t^2\zeta_1')\zeta_2 - (f_2' + \delta\hat{S}_t^2\zeta_2')\zeta_1}.$$

and is in the direction V_1, which is the direction of the additional noise.

Our aim is to choose ζ_1 and ζ_2 such that the first component equals 0. In particular,

$$(f_1''\sigma_1^2 + \delta\hat{S}_t^2\zeta_1''\sigma_1^2)\zeta_2 - (f_2''\sigma_1^2 + \delta\hat{S}_t^2\zeta_2''\sigma_1^2)\zeta_1 \tag{5.65}$$

equals 0.

If we can find ζ_1 and ζ_2 such that this holds, for example

$$\zeta_1 = (f_1''\sigma_1^2 + \delta\hat{S}_t^2\zeta_1''\sigma_1^2), \tag{5.66}$$

$$\zeta_2 = (f_2''\sigma_1^2 + \delta\hat{S}_t^2\zeta_2''\sigma_1^2), \tag{5.67}$$

(5.64) becomes π_t^2 and (5.63) becomes

$$dX_t^{\hat{\pi}} = \pi_t^1\sigma_1\,dW_t^1 + \frac{\delta}{\sqrt{\epsilon}}\pi_t^2\sigma_2\,dW_t^2 + \pi_t^1\mu_1\,dt + \frac{\delta}{\epsilon}\pi_t^2\mu_2\,dt + \pi_t^2\,dt. \tag{5.68}$$

We see that in this case our wealth process is independent of f_1 and f_2 and the wealth process of trading in this market equals that from trading in a market with assets $\tilde{S}_t^1$ and $\tilde{S}_t^2$ with dynamic

$$\begin{aligned} d\tilde{S}_t^1 &= \sigma_1(\tilde{S}_t^1)\,dW_t^1 + \mu_1(\tilde{S}_t^1)\,dt, \\ d\tilde{S}_t^2 &= \frac{1}{\sqrt{\epsilon}}\sigma_2(\tilde{S}_t^2)\,dW_t^2 + \frac{1}{\epsilon}\mu_2(\tilde{S}_t^2)\,dt + \frac{1}{\delta}\,dt. \end{aligned}$$

Using our previous results in linear market it is easy to show that as ϵ and δ converge to 0 with $\epsilon \ll \delta^2$, our maximised expected log utility $V(x, s_1, \hat{s}_2, t) := \mathbb{E}[\log(X_t^\pi)|S_t^1 = s_1, \hat{S}_t^2 = \hat{s}_2]$ satisfies

$$\lim_{\epsilon\to 0,\delta\to 0,\epsilon\ll\delta^2}\left(V(x, s_1, \hat{s}_2, t) - \frac{\lambda t}{\epsilon} - \frac{1}{\delta}c_1\right) = \log(x) + \phi + c_2 + \mathbb{E}\left[\frac{1}{2}\frac{\mu_1^2}{\sigma_1^2}\middle| S_0^1 = s_1\right]. \tag{5.69}$$

for suitable constants λ, c_1, c_2 and function ϕ depending on s_1 and $\hat{s}_2$. This result perfectly coincides with the result from the previous section.

However, since the RHS of (5.66) and (5.67) are $\hat{S}_t^2$ dependent, while the LHS of (5.66) and (5.67) are independent of S_t^2, these equalities cannot hold unless S_t^2 is constant. Hence, such a choice of noise factor is not possible. We can also see that for any choice of ζ_1 and ζ_2, (5.65) has order at least $O(\delta)$ locally due to dependence on $\hat{S}_t^2$.

Here we choose an approximation such that (5.65) has order $O(\delta)$ locally, in particular we choose

$$\begin{aligned} \zeta_1 &= f_1'' \\ \zeta_2 &= f_2'' \end{aligned} \tag{5.70}$$

We will see that with such choice although (5.65) converges to 0 as δ converges to 0, it makes an order $O(1)$ difference on the maximised expected utility of our wealth process

X_t^π when comparing with (5.69).

Remark 5.4.1. Note that in our market model, we need to restrict that

$$(f_1' + \delta \hat{S}_t^2 \zeta_1')\zeta_2 - (f_2' + \delta \hat{S}_t^2 \zeta_2')\zeta_1 \neq 0$$

for all possible S_t^1 and S_t^2. If at some point $(f_1' + \delta \hat{S}_t^2 \zeta_1')\zeta_2 = (f_2' + \delta \hat{S}_t^2 \zeta_2')\zeta_1$ holds, the direction of noise is in V_3 which is the direction that our asset prices change with respect to S_t^1. This means no matter how we choose $\hat{\pi}_t^1$ and $\hat{\pi}_t^2$, there always exists a deterministic function $K : \mathbb{R}^2 \to \mathbb{R}$ such that $\pi_t^2 = K(S_t^1, S_t^2)\pi_t^1$ and (5.63) becomes

$$\begin{aligned} dX_t^\pi = \pi_t^1 \left(\sigma_1 \, dW_t^1 + K\frac{\delta}{\sqrt{\epsilon}}\sigma_2 \, dW_t^2 + \mu_1 \, dt + K\frac{\delta}{\epsilon}\mu_2 \, dt \right) \\ + \frac{1}{2}(\hat{\pi}_t^1 f_1'' \sigma_1^2 + \hat{\pi}_t^2 f_2'' \sigma_1^2 + \hat{\pi}_t^1 \delta \hat{S}_t^2 \zeta_1'' \sigma_1^2 + \hat{\pi}_t^2 \delta \hat{S}_t^2 \zeta_2'' \sigma_1^2) \, dt \end{aligned}$$

Unless the additional drift $\hat{\pi}_t^1 f_1'' \sigma_1^2 + \hat{\pi}_t^2 f'' \sigma_1^2 + \hat{\pi}_t^1 \delta \hat{S}_t^2 \zeta_1'' \sigma_1^2 + \hat{\pi}_t^2 \delta \hat{S}_t^2 \zeta_2'' \sigma_1^2$ is also a multiple of π_t^1, we can always choose suitable $\hat{\pi}_t^1$ and $\hat{\pi}_t^2$ so that this wealth process explodes, and hence, the maximised expected utility will also explode.

For our choice of modelling the noise factors (5.70), this means we need to additionally assume that $f_1' f_2'' \neq f_2 f_1''$. In particular, such choice of noise factor is not applicable if our approximated market is linear. In fact, in the linear market our lower-dimensional market is already arbitrage free, and hence we cannot identify a direction and magnitude of additional noise using this idea.

Using this choice to model the noise, consider the maximised expected value of log utility. As before, we can write the value process $V(s_1, \hat{s}_2, x, t) = \log(x) + g(s_1, \hat{s}_2, t)$ and deduce a PDE for g:

$$-\frac{\partial g}{\partial t} + \mathcal{L}^1 g + \frac{1}{\epsilon}\mathcal{L}^2 g + H = 0, \tag{5.71}$$

where

$$\begin{aligned} \mathcal{L}^1 f &:= \mu_1 \frac{\partial f}{\partial s_1} + \frac{1}{2}\sigma_1 \frac{\partial^2 f}{\partial s_1^2}, \\ \mathcal{L}^2 f &:= \mu_2 \frac{\partial f}{\partial s_2} + \frac{1}{2}\sigma_2 \frac{\partial^2 f}{\partial s_2^2}, \end{aligned}$$

and

$$H = \frac{\left(\mu_1 + \frac{(f_1''\sigma_1^2+\delta\hat{S}_t^2\zeta_1''\sigma_1^2)\zeta_2+(f_2''\sigma_1^2+\delta\hat{S}_t^2\zeta_2''\sigma_1^2)\zeta_1}{(f_1'+\delta\hat{S}_t^2\zeta_1')\zeta_2-(f_2'+\delta\hat{S}_t^2\zeta_2')\zeta_1}\right)^2}{2\sigma_1^2} + \frac{\left(\frac{\delta}{\epsilon}\mu_2 + \frac{(f_2''\sigma_1^2+\delta\hat{S}_t^2\zeta_2''\sigma_1^2)(f_1'+\delta\hat{S}_t^2\zeta_1')-(f_1''\sigma_1^2+\delta\hat{S}_t^2\zeta_1''\sigma_1^2)(f_2'+\delta\hat{S}_t^2\zeta_2')}{(f_1'+\delta\hat{S}_t^2\zeta_1')\zeta_2-(f_2'+\delta\hat{S}_t^2\zeta_2')\zeta_1}\right)^2}{2\frac{\delta^2}{\epsilon}\sigma_2^2}.$$

Substituting ζ_1 and ζ_2 into this expression, assuming ϵ converges to 0 faster than δ^2 and applying a Taylor expansion up to order 1 terms we obtain

$$H = \frac{1}{\epsilon}\frac{\mu_2^2}{2\sigma_2^2} + \frac{1}{\delta}\frac{\mu_2\sigma_1^2}{2\sigma_2^2} + \frac{\mu_1^2}{2\sigma_1^2} + \frac{\mu_2\hat{S}_t^2\sigma_1^2(f_2''''f_1' + f_2''f_1''' - f_1''''f_2' - f_1''f_2''')}{2\sigma_2^2(f_1'f_2'' - f_2'f_1'')} + \frac{\mu_2\hat{S}_t^2\sigma_1^2(f_1'''f_2'' - f_2'''f_1'')}{2\sigma_2^2(f_1'f_2'' - f_2'f_1'')} + \text{higher order terms}$$

From this expression we can see that our choice of noise not only removes the effect of drift due to nonlinearity, we even have asymptotically removed the effect of the drift and the effect of changing direction of noise from all exploding terms of H.

As before, we consider the expansion of g with respect to ϵ and δ, that is

$$g = \hat{g} + \frac{1}{\delta}g_\delta = \frac{1}{\epsilon}g_\epsilon + \phi + g_0 + \frac{1}{\delta}g_\delta \tag{5.72}$$

Substituting (5.72) into (5.71) and considering the orders of ϵ and δ we obtain the following set of PDEs:

$$-\frac{\partial g_\delta}{\partial t} + \mathcal{L}^1 g_\delta + \frac{1}{\epsilon}\mathcal{L}^2 g_\delta + \frac{\mu_2\sigma_1^2}{\sigma_2^2} = 0, \tag{5.73}$$

$$\mathcal{L}^2 g_\epsilon = 0, \tag{5.74}$$

$$-\frac{\partial g_\epsilon}{\partial t} + \mathcal{L}^1 g_\epsilon + \mathcal{L}^2\phi + \frac{\mu_2^2}{2\sigma_2^2} = 0, \tag{5.75}$$

$$-\frac{\partial g_0}{\partial t} + \mathcal{L}^1 g_0 + \frac{1}{\epsilon}\mathcal{L}^2 g_0 + \mathcal{L}^1\phi + \frac{\mu_1^2}{2\sigma_1^2} + F + \text{higher order terms} = 0, \tag{5.76}$$

where

$$F = \frac{\mu_2 \hat{S}_t^2 \sigma_1^2 (f_2'''' f_1' + f_2'' f_1''' - f_1'''' f_2' - f_1'' f_2''')}{2\sigma_2^2 (f_1' f_2'' - f_2' f_1'')} + \frac{\mu_2 \hat{S}_t^2 \sigma_1^2 (f_1''' f_2'' - f_2''' f_1'')}{2\sigma_2^2 (f_1' f_2'' - f_2' f_1'')}.$$

Note that, in general, this function is not zero. However, in some special and useful cases, such as both f_1 and f_2 being arbitrary quadratic polynomials with $f_1' f_2'' - f_2' f_1'' \neq 0$, this additional function F vanishes.

Observe that (5.73) is a parabolic PDE for g_δ and (5.74), (5.75) and (5.76) are almost identical to the PDEs (5.13), (5.14) and (5.15) which we have dealt with before. Hence, if we choose $g_\delta(s_1, \hat{s}_2, 0) = g_\epsilon(s_1, \hat{s}_2, 0) = 0$, we expect that as ϵ and δ converge to 0, g_δ converges to $\mathbb{E}[\sigma_1^2 | S_0^1 = s_1] \int \frac{\mu_2}{\sigma_2^2} d\Pi$, g_ϵ converges to λt where (λ, ϕ) is the unique solution to the ergodic PDE

$$\lambda + \mathcal{L}^2 \phi + \frac{\mu_2^2}{2\sigma_2^2} = 0, \tag{5.77}$$

and λ has a probabilistic representation $\lambda = \int \frac{\mu_2^2}{2\sigma_2^2} d\Pi$, where Π is the unique invariant distribution for the process $\hat{S}_t^2$ with dynamics

$$d\hat{S}_t^2 = \frac{1}{\sqrt{\epsilon}} \sigma_2(\hat{S}_t^2)\, dW_t^2 + \frac{1}{\epsilon} \mu_2(\hat{S}_t^2)\, dt.$$

The only PDE that remains to be solved is (5.76), with the initial condition $g_0(s_1, \hat{s}_2, 0) = -\phi(\hat{s}_2, 0)$.

If we can apply the Feynman–Kac formula to this PDE, we can get the following probabilistic representation for g_0:

$$\begin{aligned} g_0 =& \mathbb{E}\left[\int_0^t \frac{\mu_1^2}{2\sigma_1^2} du \middle| S_0^1 = s_1 \right] - \mathbb{E}[\phi(\hat{S}_t^2) | \hat{S}_0^2 = \hat{s}_2] \\ &+ \mathbb{E}\left[\int_0^t F du \middle| S_t^1 = s_1, \hat{S}_t^2 = \hat{s}_2 \right] + \text{higher order terms} \end{aligned} \tag{5.78}$$

Hence, we expect that the value process V for maximised expected log utility satisfies

$$\lim_{\epsilon\to 0,\delta\to 0,\epsilon\ll\delta^2}\left(V(x,s_1,\hat{s}_2,t)-\frac{t\int\frac{\mu_2^2}{2\sigma_2^2}d\Pi}{\epsilon}-\frac{1}{\delta}\mathbb{E}\left[\int_0^t\sigma_1^2du|S_0^1=s_1\right]\int\frac{\mu_2}{\sigma_2^2}d\Pi\right)$$
$$=\log(x)+\mathbb{E}\left[\int_0^t\frac{\mu_1^2}{2\sigma_1^2}du\middle|S_0^1=s_1\right]+\phi-\int\phi d\Pi+\mathbb{E}\left[\int_0^t\int Fd\Pi du\middle|S_t^1=s_1\right]$$

Note that, as before, the first two terms on the right hand side are exactly the maximised expected utility of trading in a market with only S_t^1 and bank account. In addition, if we again consider $S_0^2\sim\Pi$, the averaged maximised expected utility $\bar{V}(x,s_1,t):=\int V(x,s_1,s_2,t)d\Pi=\sup_\pi\mathbb{E}[\log(X_t^\pi)|S_0^1=s_1,S_0^2\sim\Pi]$ satisfies:

$$\lim_{\epsilon\to 0,\delta\to 0,\epsilon\ll\delta^2}\left(\bar{V}(x,s_1,s_2,t)-\frac{t\int\frac{\mu_2^2}{2\sigma_2^2}d\Pi}{\epsilon}-\frac{1}{\delta}\mathbb{E}\left[\int_0^t\sigma_1^2du\middle|S_0^1=s_1\right]\int\frac{\mu_2}{\sigma_2^2}d\Pi\right)$$
$$=\log(x)+\mathbb{E}\left[\int_0^t\frac{\mu_1^2}{2\sigma_1^2}du\middle|S_0^1=s_1\right]+\mathbb{E}\left[\int_0^t\int Fd\Pi du\middle|S_t^1=s_1\right]$$

Compared with the results in the previous section, we see that there is an additional $\mathbb{E}\left[\int_0^t\int Fd\Pi du|S_t^1=s_1\right]$ term that we cannot remove. This term can be regarded as the cost of simplification such that the direction of the additional noise depends only on the slow market factor S_t^1.

This result also gives us an idea about how we choose to model the noise factors. It suggests that if we project the data points along the direction (f_1'',f_2'') instead of perpendicular to the submanifold which is in the direction (f_1',f_2') or perpendicular to the x-axis which is in the direction of $(0,1)$, we can more easily recognize and deal with the exploding terms due to the fast oscillating noise factors.

5.5 The Choice of Trading Strategy

In the previous sections, we explored the asymptotic behaviour of our maximised expected log utility. In this section, building on the intuition from this behaviour, we briefly discuss what trading strategy can be obtained when working in an error-prone and arbitrage containing market. In particular, we want to choose a trading strategy that can avoid exploding terms in the asymptotic maximised expected utility, while still approximately maximising the wealth from other parts of the market. In this case, we do not try to gain wealth from oscillation behaviour of the noise factors or the curvature of our market model.

The intuition for such a choice comes from the expression (5.63). We have argued in the previous section that if there exists (ζ_1, ζ_2) such that (5.65) equals 0, the sum of the last four additional drift terms becomes π_t^2. In this case if we choose π_t^2 to be 0, the wealth process $X_t^{\hat{\pi}}$ becomes

$$dX_t^{\hat{\pi}} = \pi_t^1 \sigma_1 \, dW_t^1 + \pi_t^1 \mu_1 \, dt. \tag{5.79}$$

This is the wealth process in a market with only S_t^1 and bank account and controlled by trading process π_t^1. Hence choosing π_t^1 to be the process maximising the expected log utility of (5.79), and using the relationship

$$\hat{\pi}_t^1 = \frac{h_{22}\pi_t^1}{h_{11}h_{22} - h_{12}h_{21}}$$
$$\hat{\pi}_t^2 = -\frac{h_{12}\pi_t^1}{h_{11}h_{22} - h_{12}h_{21}}$$

where h_{ij} are functions of S_t^1 and S_t^2 defined as

$$h_{22}(s_1) = \zeta_2(s_1)$$
$$h_{12}(s_1) = \zeta_1(s_1)$$
$$h_{11}(s_1, \hat{s}_2)h_{22}(s_1) - h_{12}(s_1)h_{21}(s_1, \hat{s}_2) = (f_1'(s_1) + \delta \hat{s}_2 \zeta_1'(s_1))\zeta_2(s_1) - (f_2'(s_1) + \delta \hat{s}_2 \zeta_2'(s_1))\zeta_1(s_1)$$

we can construct a trading strategy that neglects exploding terms and maximises the expected utility with respect to the market factor S_t^1, which we can regarded as the only reliable part of the market.

However, as we have argued that such a choice of (ζ_1, ζ_2) is not possible unless S_t^2 is constant, we choose an approximation as before. Using the noise

$$\zeta_1 = f_1'',$$
$$\zeta_2 = f_2'',$$

our choice of the trading strategy is

$$\hat{\pi}_t^1 = \frac{f_2''\pi_t^1}{(f_1' + \delta \hat{S}_t^2 f_1''')f_2'' - (f_2' + \delta \hat{S}_t^2 f_2''')f_1''}, \tag{5.80}$$
$$\hat{\pi}_t^2 = -\frac{f_1''\pi_t^1}{(f_1' + \delta \hat{S}_t^2 f_1''')f_2'' - (f_2' + \delta \hat{S}_t^2 f_2''')f_1''}. \tag{5.81}$$

where π_t^1 is the optimal trading strategy of trading in a market with only S_t^1 and B_t.

Remark 5.5.1. It is quite important to note that this trading strategy is not simply removing the fast oscillation terms from the optimal trading strategy corresponding to the stochastic control problem in the previous sections. For logarithmic utility, the optimal trading process is to put a suitable ratio of current wealth in each available asset. Hence, since the wealth processes corresponding to these two trading strategies are quite different, these trading strategies will be quite different. However, in our new trading strategy the ratio of wealth we used to buy assets almost equals the ratio of wealth we used to buy assets if we use the optimal trading strategy in the full market but remove the fasting oscillating terms.

Remark 5.5.2. π_t^1, f_1 and f_2 only depend on S_t^1, and we can also write $\hat{\pi}_t^1$ and $\hat{\pi}_t^2$ as functions of S_3 and S_4 if f_1 and f_2 are bijective functions.

Finally, we will briefly argue why this trading strategy is a reasonable choice. Substituting this into (5.63) we obtain

$$dX_t^{\hat{\pi}} = \pi_t^1\sigma_1\,dW_t^1 + \pi_t^1\mu_1\,dt + \delta\pi_t^1\hat{S}_t^2\frac{f_2''f_1''''\sigma_1^2 + f_1''f_2''''\sigma_1^2}{2((f_1' + \delta\hat{S}_t^2f_1''')f_2'' - (f_2' + \delta\hat{S}_t^2f_2''')f_1'')}\,dt. \tag{5.82}$$

The last term has order δ, which means that the difference between our wealth process and the wealth process of trading in a market with only asset S_t^1 and bank account B_t has order δ. Hence if we choose π_t^1 to be the optimal trading strategy in this one asset market, by Taylor expansion of log function we can deduce that we achieve a good approximation of the maximised expected utility we want to achieve by using this trading strategy. This also suggests that, by using this trading strategy, the wealth process we obtain in markets in different submanifolds close to each other will be similar, as long as these submanifolds do not oscillate too much. This somehow solves the problem mentioned in Chapter 1 that the maximised expected utility for trading in two arbitrage-free markets support on submanifolds close to each other may be quite different.

The other reason for this choice of trading strategy is that this choice is consistent with more basic cases:

Consider the market where one asset depends on S_t^1 linearly, and the dynamic of assets is given by

$$\begin{aligned} S_t^3 &= S_t^1, \\ S_t^4 &= f(S_t^1) + \zeta_2(S_t^1)S_t^2. \end{aligned}$$

Our choice of trading strategy is reduced to

$$\hat{\pi}_t^1 = \pi_t^1,$$
$$\hat{\pi}_t^2 = 0,$$

which means we only trade asset S_t^3 using the trading strategy π_t^1. Clearly, in this case our wealth process coincides with the wealth process of trading in the market with single asset S_t^1.

Another case is when we assume our market model does not contain noise, that is $\delta = 0$, and the dynamics of assets follows

$$S_t^3 = f_1(S_t^1),$$
$$S_t^4 = f_2(S_t^1).$$

In this case our trading strategy reduces to

$$\hat{\pi}_t^1 = \frac{f_2'' \pi_t^1}{f_1' f_2'' - f_2' f_1''},$$
$$\hat{\pi}_t^2 = -\frac{f_1'' \pi_t^1}{f_1' f_2'' - f_2' f_1''}.$$

Using this trading strategy, our wealth process is

$$dX_t^{\hat{\pi}} = \pi_t^1 \sigma_1 \, dW_t^1 + \pi_t^1 \mu_1 \, dt,$$

which is the wealth process we expect.

From these arguments, we can see that (5.80) and (5.81) are a reasonable trading strategy.

Chapter 6

Conclusion

In this paper, we found the asymptotic behaviour of the maximised expected value of the logarithmic utility of wealth process in a market approximated by an underlying lower-dimensional market plus additional fast mean-reverting factor. These behaviours are different for the linear and nonlinear underlying markets.

If the underlying lower-dimensional submanifold is linear, and we choose the noise to be normal to the submanifold, the maximised expected utility would asymptotically be equal to a divergent term proportional to the rate of mean reversion of noise plus some finite terms. If we further integrate these finite terms with respect to the invariant distribution of the fast mean reverting noise, we obtain the maximised expected log utility of trading in the underlying lower-dimensional market.

If the underlying lower-dimensional submanifold is nonlinear, the situation will be much more complicated. In this case, the maximised expected utility depends on our choice of noise factor. We first need to choose a noise factor with a suitable scale and direction related to the curvature of the submanifold. We showed that, in general, for any choice of noise factors we can make, the maximised expected utility will asymptotically be equal to the sum of a divergent term proportional to the rate of mean reversion of noise, a divergent term proportional to the inverse of magnitude of noise, some finite terms that can be integrated to the maximised expected log utility of trading in the market with only noise factor and the bank account, and an error term with at $O(1)$. This suggests that the impact of curvature of the underlying market goes beyond exploding terms, no matter how we choose the direction and magnitude of the additional noise factor. In addition to the exploding terms, lower order terms depend on the curvature as well. Hence, when we try to approximate the real market and use the approximated market to approximate the maximised expected utility of trading in the real market, simply minimizing the squared

distance of error of the approximation is not sufficient to get a good approximation, even if we can remove exploding terms from maximised expected utility.

In addition, we obtained a trading strategy that only optimizes wealth with respect to the reliable part of the market. In particular, using this strategy means that we believe that the only thing worth trading in this market is the underlying slow market factor. All other components of the market, such as noise factors and curvature of the lower-dimensional submanifold, are not reliable. This trading strategy suggests a reasonable way of projecting data in the real market to the lower-dimensional submanifold we approximate as well.

Bibliography

[1] Sahar Albosaily and Serguei Pergamenshchikov. Optimal investment and consumption for Ornstein–Uhlenbeck spread financial markets with logarithmic utility. *arXiv preprint arXiv:1809.08139*, 2018.

[2] Olivier Alvarez and Martino Bardi. Singular perturbations of nonlinear degenerate parabolic PDEs: a general convergence result. *Archive for rational mechanics and analysis*, 170(1):17–61, 2003.

[3] Olivier Alvarez, Martino Bardi, and Claudio Marchi. Multiscale problems and homogenization for second-order Hamilton–Jacobi equations. *Journal of Differential Equations*, 243(2):349–387, 2007.

[4] Martino Bardi, Annalisa Cesaroni, and Andrea Scotti. Convergence in multiscale financial models with non-Gaussian stochastic volatility. *ESAIM. Control, optimisation and calculus of variations*, 22(2):500–518, 2016.

[5] G. Barles and P.E. Souganidis. On the large time behavior of solutions of Hamilton–Jacobi equations. *SIAM journal on mathematical analysis*, 31(4):925–939, 2000.

[6] G. Barles and P.E. Souganidis. Space-time periodic solutions and long-time behavior of solutions to quasi-linear parabolic equations. *SIAM journal on mathematical analysis*, 32(6):1311–1323, 2001.

[7] Samuel N. Cohen, Christoph Reisinger, and Sheng Wang. Arbitrage-free neural-SDE market models. *Applied mathematical finance.*, 30(1):1–46, 2023.

[8] Andrei Cozma and Christoph Reisinger. A mixed Monte Carlo and PDE variance reduction method for foreign exchange options under the Heston–CIR model. *arXiv preprint arXiv:1509.01479*, 2016.

[9] Michael G. Crandall, Hitoshi Ishii, and Pierre-Louis Lions. User's guide to viscosity solutions of second order partial differential equations. *Bulletin (new series) of the American Mathematical Society*, 27(1):1–67, 1992.

[10] Wendell H. Fleming and Shuenn-Jyi Sheu. Optimal long term growth rate of expected utility of wealth. *The Annals of applied probability*, 9(3):871–903, 1999.

[11] Wendell H. Fleming and Shuenn-Jyi Sheu. Risk-sensitive control and an optimal investment model. *Mathematical finance*, 10(2):197–213, 2000.

[12] Jean-Pierre Fouque. *Multiscale stochastic volatility for equity, interest rate, and credit derivatives* . Cambridge University Press, Cambridge, 2011.

[13] Jean-Pierre Fouque and Ruimeng Hu. Portfolio optimization under fast mean-reverting and rough fractional stochastic environment. *Applied mathematical finance.*, 25(4):361–388, 2018.

[14] Jean-Pierre Fouque, Ronnie Sircar, and Thaleia Zariphopoulou. Portfolio optimization and stochastic volatility asymptotics. *Mathematical finance*, 27(3):704–745, 2017.

[15] Yasuhiro Fujita, Hitoshi Ishii, and Paola Loreti. Asymptotic solutions of viscous Hamilton–Jacobi equations with Ornstein–Uhlenbeck operator. *Communications in partial differential equations*, 31(6):827–848, 2006.

[16] Dror Givon. Strong convergence rate for two-time-scale jump-diffusion stochastic differential systems. *Multiscale modeling & simulation*, 6(2):577–594, 2007.

[17] Paolo Guasoni, Lóránt Nagy, and Miklós Rásonyi. Young, timid, and risk takers. *Mathematical finance*, 31(4):1332–1356, 2021.

[18] Paolo Guasoni, Antonella Tolomeo, and Gu Wang. Should commodity investors follow commodities' prices? *SIAM journal on financial mathematics*, 10(2):466–490, 2019.

[19] Ying Hu and Pierre-Yves Madec. A probabilistic approach to large time behaviour of viscosity solutions of parabolic equations with Neumann boundary conditions. *Applied mathematics & optimization*, 74(2):345–374, 2016.

[20] Ying Hu, Pierre-Yves Madec, and Adrien Richou. A probabilistic approach to large time behavior of mild solutions of Hamilton–Jacobi–Bellman equations in infinite dimension. *SIAM journal on control and optimization*, 53(1):378–398, 2015.

[21] Naoyuki Ichihara and Shuenn-Jyi Sheu. Large time behavior of solutions of Hamilton–Jacobi–Bellman equations with quadratic nonlinearity in gradients. *SIAM Journal on Mathematical Analysis*, 45(1):279–306, 2013.

[22] R.Z. Khasminskii and G. Yin. On averaging principles: An asymptotic expansion approach. *SIAM journal on mathematical analysis*, 35(6):1534–1560, 2004.

[23] R.Z. Khasminskii and G. Yin. Limit behavior of two-time-scale diffusions revisited. *Journal of Differential Equations*, 212(1):85–113, 2005.

[24] J.C. Mattingly, A.M. Stuart, and D.J. Higham. Ergodicity for SDEs and approximations: locally Lipschitz vector fields and degenerate noise. *Stochastic processes and their applications*, 101(2):185–232, 2002.

[25] Gawtum Namah and Jean-Michel Roquejoffre. Convergence to periodic fronts in a class of semilinear parabolic equations. *Nonlinear differential equations and applications*, 4(4):521–536, 1997.

[26] Gawtum Namah and Jean-Michel Roquejoffre. Remarks on the long time behaviour of the solutions of Hamilton–Jacobi equations. *Communications in partial differential equations*, 24(5-6):883–893, 1999.

[27] Huyen. Pham. *Continuous-time stochastic control and optimization with financial applications*. Stochastic modelling and applied probability, 61. Springer, Berlin, 1st ed. 2009. edition, 2009.

[28] Murray H. Protter and Hans F. Weinberger. *Maximum principles in differential equations*. Prentice-Hall partial differential equations series. Prentice-Hall, Englewood Cliffs, 1967.

[29] Johannes Ruf and Weiguan Wang. Neural networks for option pricing and hedging: a literature review. *arXiv preprint arXiv:1911.05620*, 2020.