Quadratic Criteria for Optimal Martingale Measures in Incomplete Markets

Thomas Andrew McWalter

A dissertation submitted to the Faculty of Science, University of the Witwatersrand, in fulfilment of the requirements for the degree of Master of Science.

April 7, 2006

Programme in Advanced Mathematics of Finance,
University of the Witwatersrand,
Johannesburg.

Declaration

I declare that this dissertation is my own, unaided work. It is being submitted for the Degree of Master of Science in the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination in any other University.

April 7, 2006

Abstract

This dissertation considers the pricing and hedging of contingent claims in a general semimartingale market. Initially the focus is on a complete market, where it is possible to price uniquely and hedge perfectly. In this context the two fundamental theorems of asset pricing are explored. The market is then extended to incorporate risk that cannot be hedged fully, thereby making it incomplete. Using quadratic cost criteria, optimal hedging approaches are investigated, leading to the derivations of the minimal martingale measure and the variance-optimal martingale measure. These quadratic approaches are then applied to the problem of minimizing the basis risk that arises when an option on a non-traded asset is hedged with a correlated asset. Closed-form solutions based on the Black-Scholes equation are derived and numerical results are compared with those resulting from a utility maximization approach, with encouraging results.

For my parents
David and Ingrid.

Acknowledgements

I would like to express my sincere appreciation to my supervisor, Hardy Hulley, for his tireless assistance and encouragement. His ability to unravel my (often convoluted) thinking and help me to present it in the best possible light has been crucial to the outcome.

On the two occasions that I travelled to Australia I was made welcome at the University of Technology, Sydney. I would like to thank Tony Hall and the School of Finance and Economics for providing me with facilities. Thanks also to Eckhard Platen for his interest in my work and for facilitating my participation in the QMF conferences in 2004 and 2005. On my second visit to Australia, Hardy and Louise Hulley were very kind to provide me with accommodation.

Helpful email communications with Norbert Christopeit and Martin Schweizer (who also sent me a copy of his Ph.D. thesis) are gratefully acknowledged, as well as an interesting discussion with Michael Monoyios that motivated my investigation of basis risk. Martin Schweizer also made some valuable comments and pointed out a number of errors.

Finally, I would like to thank everyone at Wits who provided support and encouragement: Grant Lotter (who provided a proof of the "infamous McWalter theorem"), David Taylor, Alex Welte, Graeme West and all my fellow students.

Contents

List of Figures

List of Tables

Index of Notation

Chapter 1

Introduction

In the theory of contingent claim pricing, the primary innovation of Black and Scholes [6] and Merton [64] was to show that, for European calls or puts, a hedging portfolio can be constructed with initial value equal to the fair price of the option and final value equal to its payoff at expiry. Prices determined in this manner are unique and independent of risk preference. This approach makes explicit use of a no-arbitrage assumption, but the link between no-arbitrage and equivalent martingale measures (EMMs) was not yet apparent, nor was it immediately clear that every contingent claim could be priced in this manner (i.e. that the market is complete).

Since the seminal papers of Harrison and Kreps [39] and Harrison and Pliska [40, 41], the modern theory of contingent claim valuation has been developed with the firm mathematical foundation of martingales and stochastic integrals. In these papers the relationship between no-arbitrage and EMMs was made explicit, as was the concept of market completeness. It was shown that a market is complete if and only if its vector price process has a certain martingale representation property. When a market is complete, every contingent claim has a unique preference-independent price, due to the existence of a unique EMM and a self-financing trading strategy which replicates its terminal payoff.

Unfortunately, not many market models are complete; and complete markets, while useful for producing mathematical results, are nonetheless idealizations which model reality imperfectly. In attempting to model the complexities of the real market more closely, it is necessary to consider more complex and necessarily incomplete models. This means that a hedger cannot create a risk-free hedge portfolio for a claim, which in turn translates into a market model where the above-mentioned representation property does not hold and EMMs are not unique. The problem then becomes one of selecting a strategy that minimizes risk, or equivalently of choosing an EMM for pricing that is optimal in some sense.

This dissertation reviews the theory of complete markets and then extends this theory, in a particular direction, to the case where unhedgeable risk is incorporated.

In this introduction we provide a précis of the work, with an indication of the kind of applications that have appeared in the literature.

1.1 Stochastic Integrals and Portfolios

The key concept of stochastic integration is fundamental to the analysis of contingent claims, precisely because of the idea of a replicating portfolio. At inception, an initial endowment, being the amount charged for the claim, is invested in a portfolio with a strategy designed to produce a payoff that best reflects the value of the claim at maturity.

The stochastic integral may be thought of intuitively as the mechanism for calculating the change in value of the portfolio over time, given a strategy and a number of instruments in which to invest. As an example, consider a discrete-time scenario, where an investor is allowed to hold his wealth as an investment in a stock or as cash[1]. If at time t he holds ξ_t shares and an amount η_t in cash, the value of his portfolio will be given by

$$V_t = \eta_t + \xi_t S_t,$$

where S_t is the share price. To simplify the example, it will be assumed that the interest rate is zero, so that the cash portion of the portfolio does not grow in value between time steps[2]. After one time step, the portfolio will have changed in value to

$$V_{t+1} = \eta_t + \xi_t S_{t+1},$$

as a result of the change in value of the share. The investor may then select a new portfolio holding of ξ_{t+1} shares and η_{t+1} in cash. If we assume that he doesn't transfer money to or from his portfolio at any stage and that there are no transaction costs, then we have the self-financing constraint

$$\eta_t + \xi_t S_{t+1} = \eta_{t+1} + \xi_{t+1} S_{t+1}.$$

This condition implies a relationship between the values of the portfolio at successive time steps, namely

$$\begin{aligned} V_{t+1} - V_t &= \eta_{t+1} + \xi_{t+1} S_{t+1} - (\eta_t + \xi_t S_t) \\ &= \xi_t (S_{t+1} - S_t). \end{aligned} \tag{1.1}$$

Given that the portfolio starts with an initial value V_0, (1.1) allows us to express its value at time t as

$$V_t = V_0 + \sum_{i=0}^{t-1} \xi_i (S_{i+1} - S_i). \tag{1.2}$$

[1] Short holdings are allowed.
[2] Alternatively, one may assume the stock price is discounted.

The sum in (1.2), which is called the gain for the strategy ξ, may be interpreted as an integral (in discrete-time) of the portfolio strategy with respect to the stochastic share price. It should be noted that we must insist that the strategy be non-anticipating. In other words, in formulating his strategy at time t, the investor should not take into account any information about the price of the stock at any time later that t. This rules out unfair strategies, such as those which involve insider trading.

We wish to explore a continuous time version of the integral constructed in this example. It is not a simple matter to assign meaning to expressions of the form

$$V_t = V_0 + \int_0^t \xi_u \, dS_u,$$

precisely because we are dealing with stochastic processes which are not of finite variation and, as a result, the Lebesgue-Stieltjes integral is not well-defined. Nevertheless, it turns out that a theory of integration, initiated by Itô [47, 49, 48], is possible for processes that have finite quadratic variation. The idea of insisting on integrands that are non-anticipating (or, more formally, predictable) turns out to be key to the development of the theory.

In this dissertation, risky assets are represented as semimartingales — the most general class of processes for which stochastic integration is defined. This class includes both continuous and discontinuous processes. We keep the exposition manageable by providing results for only a single risky asset and refer the reader to the relevant literature for the full vector formulations of the theory. To simplify the mathematics further, it will be assumed that the price process, denoted by X, is discounted at the riskless rate.

1.2 Complete Markets

The theory of complete markets is encapsulated in two important results known, as the fundamental theorems of asset pricing. The first fundamental theorem links the concept of no-arbitrage to the existence of an equivalent local martingale measure (ELMM) for the price process X (i.e. an equivalent probability measure under which X is a local martingale). If a strategy has no chance of making a loss, but there is a non-zero probability of it making a profit, then it constitutes an arbitrage. Clearly, if any strategy of this form could be found, then it would be exploited in arbitrarily large amounts by market participants wishing to make a riskless profit. Any market allowing this kind of activity cannot be in equilibrium and consequently the market ensures that strategies of this sort are eliminated. It is relatively easy to show that if an ELMM for X exists, then arbitrage of this type cannot exist. However, this simple definition of no-arbitrage is not sufficient to show the reverse implication — a more sophisticated notion is required. The ultimate development of this theory was accomplished by Delbaen and Schachermayer [20, 22], who provided the first fundamental theorem of asset pricing in its most general form. We shall explore these concepts in detail.

The second fundamental theorem of asset pricing links market completeness to the uniqueness of an EMM for X and is intimately connected with the mathematical theory of martingale representation. In summary, a market is complete if and only if every contingent claim, represented by an appropriately chosen random variable H, has a unique representation in terms of X as follows:

$$H = H_0 + \int_0^T \xi_s^H \, dX_s \quad \text{a.s.} \tag{1.3}$$

Here ξ^H is interpreted as the hedging strategy. If such a representation exists for every claim, it can be shown that the EMM for X is unique. The properties of ξ^H ensure that the stochastic integral in (1.3) is also a martingale under the EMM. Thus the unique (discounted) fair price of the claim can be computed by taking expectations on both sides of (1.3):

$$H_0 = \mathbb{E}^{\mathbb{Q}}[H],$$

where $\mathbb{Q}$ is the unique EMM for X.

1.3 Incomplete Markets

Market completeness, though mathematically convenient and extremely useful for producing closed-form pricing formulae for many derivatives, represents an idealization of reality. A market model becomes incomplete when it is expanded to incorporate more realistic assumptions, such as trading restrictions, stochastic volatility and event risk. Contrary to the complete market situation, preference-independent pricing and hedging becomes impossible. This follows directly from the fact that it is impossible to find a self-financing strategy whose terminal wealth exactly replicates the expiry value of the claim.

Another way of stating this is that under any EMM for X, not all martingales can be represented as stochastic integrals with respect to X. In particular, there will always be martingales orthogonal to X. This ensures that there is no unique equivalent martingale measure for X; in fact, there are infinitely many of them.

Now that perfect hedging is impossible, an optimality criterion based on the market participant's attitude to risk is required. Following the approach of Föllmer and Sondermann [33], we shall introduce a cost process C. Due to inexact hedging, the intrinsic value[3] of the contingent claim and the value of the portfolio may deviate from each other over time. The cost process allows the modelling of this deviation and is given by

$$C_t := V_t - \int_0^t \xi_s \, dX_s,$$

where ξ is the trading strategy used and V represents the value of the hedging portfolio. Note that the cost process above should not be confused with the cost of transactions, which we do not consider here. Instead, it is a way of modelling the fact that the cost of hedging, which was constant in the complete market situation, is now a random variable evolving over time. As a result, it will be used as a measure of the effectiveness of hedging strategies.

In the complete market situation there is no deviation between portfolio value and intrinsic value, because hedging is perfect. As a result, the cost process is a constant and equal to H_0. In an incomplete market, equation (1.3) now takes the form

$$H = C_T + \int_0^T \xi_s \, dX_s,$$

where ξ must be chosen to minimize the risk associated with C, based on some criterion. A quadratic criterion is the natural first choice. Two quadratic strategies lead to the minimal martingale measure and the variance-optimal martingale measure, respectively.

The Minimal Martingale Measure

One approach to the problem of hedging under incompleteness is to relax the self-financing constraint and replace it with a new condition that the hedge portfolio remains mean self-financing. This corresponds to the hedger continually adding or removing any deficit or surplus, so that the value of the hedge portfolio is always in agreement with the intrinsic value of the claim. Under this strategy the cost process becomes a martingale.

In the simplest case, where X is a martingale under the real-world measure, this approach corresponds to minimizing the conditional remaining risk, defined by

$$R_t := \mathbb{E}\left[(C_T - C_t)^2 \,\middle|\, \mathscr{F}_t \right],$$

at every time. It was first proposed by Föllmer and Sondermann [33].

When the price process X is not a martingale under the real-world measure, then a mean self-financing strategy can be shown to be consistent with minimizing

[3] We shall give a formal definition of intrinsic value in later chapters, but for now it can be interpreted as the best estimate of the value of the claim given current information.

a quadratic measure of local risk, known as the risk quotient. The risk quotient is a variational concept introduced by Schweizer [83, 84] — its analogue in discrete-time is the local conditional risk, which has the form

$$R_{t_i} := \mathbb{E}\left[(C_{t_{i+1}} - C_{t_i})^2 \,\middle|\, \mathscr{F}_{t_i}\right].$$

A strategy for which the risk quotient is minimized at every time is called locally risk-minimizing. In this situation the claim has a unique representation

$$H = H_0 + \int_0^T \xi_s^H \, dX_s + L_T^H \quad \text{a.s.,} \tag{1.4}$$

where L^H is a martingale strongly orthogonal to X and ξ^H is unique (compare this with (1.3)). This is known as the Föllmer-Schweizer [32] decomposition of the claim. The cost process is then given by $C_t = H_0 + L_t^H$.

In the set of EMMs for X, there exists a particular measure, denoted by $\widehat{\mathbb{P}}$ and called the minimal martingale measure, which is the unique EMM for X with the property that every martingale under the real-world measure, strongly orthogonal to X, is still a martingale under $\widehat{\mathbb{P}}$. In particular, L^H in (1.4) is a martingale under $\widehat{\mathbb{P}}$. Therefore, a price for the claim can be computed by taking expectations under $\widehat{\mathbb{P}}$ on both sides of (1.4), giving

$$H_0 = \mathbb{E}^{\widehat{\mathbb{P}}}[H].$$

This price corresponds with the initial value of the locally risk-minimizing strategy described above.

The Variance-Optimal Martingale Measure

A disadvantage of the previous method is that because the hedging portfolio is not self-financing, it requires constant readjustment. It is possible to insist on the self-financing criterion, resulting in a constant cost process $C_t = c$ over the life of the claim, with a surplus or shortfall occurring at maturity. A strategy that minimizes the variance of this profit and loss is sought.

The self-financing strategy that minimizes the quadratic functional

$$R_0 = \mathbb{E}\left[(C_T - C_0)^2\right],$$

produces the combination of the initial endowment c (known as the approximation price) and strategy ξ for which the variance of the profit and loss is minimized. This can be reformulated as the strategy that minimizes the expression

$$R_0 = \mathbb{E}\left[\left(H - c - \int_0^T \xi_s \, dX_s\right)^2\right],$$

over all real numbers c and all trading strategies ξ. This so-called mean-variance optimal strategy was first proposed in the case where X is a martingale under the real-world measure by Föllmer and Sondermann [33] and Bouleau and Lamberton [7]. Duffie and Richardson [27] subsequently considered this approach in a diffusion framework for hedging futures. Their work was extended to incorporate general claims by Schweizer [87] and Hipp [46]. Finally, mean-variance hedging was investigated in a semimartingale framework by Schweizer [88, 90], Monat and Striker [67] and others [21, 76, 70, 37].

It can be shown that associated with the mean-variance optimal strategy there is a measure, known as the variance-optimal martingale measure $\widetilde{\mathbb{P}}$, under which the approximation price can be computed. Under certain conditions, the variance-optimal measure and the minimal martingale measure coincide.

Other Martingale Measures

Although the dissertation concentrates on quadratic criteria for martingale measures, it should be noted that there are a number of other strategies that have been proposed in the literature to deal with market incompleteness. The minimal entropy martingale measure (see e.g. [34, 38]) and utility-based approaches (see e.g. [56, 16]) offer some alternatives.

1.4 Applications

A number of applications of quadratic criteria for risk minimization to pricing and hedging have been formulated in the literature. In this section we briefly survey some of the sources of market incompleteness and discuss some applications of the theory outlined earlier.

Non-Traded Assets

In certain circumstances a hedger faces a commitment which is contingent on an asset that is not available for trade. If it is possible to trade in another correlated asset, a hedging strategy may be constructed that allows the hedger to minimize risk. This problem was first considered in the context of quadratic criteria by Duffie and Richardson [27], for the problem of hedging futures and later expanded to general claims by Schweizer [87].

A similar problem is faced by the hedger of an index option. Here the problem is the impracticality or impossibility of trading in all the underlying assets. This application is described in the paper of Lamberton and Lapeyre [60].

Transaction Costs

With the addition of frictions, such as transaction costs, a market model becomes incomplete. Although much of the research in this area focuses on the use of utility functions (see e.g. Davis, Panas and Zaraphopolou [15]), some investigations have considered quadratic techniques. The papers by Mercurio and Vorst [63] and Lamberton, Pham and Schweizer [61] are indicative. However, both of these papers feature discrete-time models; the formulation of this type of problem in continuous-time is an open problem at this stage (see [71]).

Stochastic Volatility Models

Stochastic volatility is another well-publicized source of market incompleteness. Market models featuring stochastic volatility are a source of many difficulties. Quadratic techniques have, however, been applied to this problem with promising results. The papers by Heath, Platen and Schweizer [42, 43, 44] compare the two quadratic criteria discussed in this dissertation, when applied to pricing and hedging under stochastic volatility. The paper by Biagini, Guasoni and Pratelli [3] also explores this area.

1.5 The Structure of the Dissertation

In Chapter 2 the mathematical notation and theory required for the rest of the dissertation is introduced. A brief survey of the general theory of stochastic integration is provided, along with further sections dealing with quadratic variation and martingale representation. The emphasis in this chapter is on introducing the concepts as quickly as possible, while pointing the reader to the literature for a full account.

Chapter 3 introduces the basic market model and provides a survey of the theory of complete markets. The concepts of trading strategy, numéraire asset and risk-neutral measure are introduced. Once a formal specification of arbitrage is provided, the first fundamental theorem of asset pricing is stated. Since a complete proof of this theorem would require a body of work at least as large as the one being presented here, we only provide a proof of one of the implications, while referring the reader to the literature for the complete story. A formal definition of market completeness is provided next, along with a statement and proof of the second fundamental theorem of asset pricing. This relates market completeness to martingale representation and the Jacod-Yor [52] theorem. To make this theory useful, we provide a mechanism for constructing EMMs. Finally, we demonstrate the use of the mathematical machinery presented in the chapter, by applying it to the Black-Scholes model. In the process, we make use of Lévy's characterization of Brownian motion and provide a statement of the Feynman-Kač theorem.

In Chapter 4 we introduce the simplest market that incorporates unhedgeable risk. In this market the asset price X is a martingale under the real-world measure. We follow the account of Föllmer and Sondermann [33].

Chapter 5 explores a generalization of the market considered in Chapter 4, to the case where X is no longer a martingale under the real-world measure. The idea of local risk minimization is now introduced. The optimal portfolio choice is shown to possess a martingale representation in terms of the underlying asset and an orthogonal component. Although there is no unique EMM for X in this setup, there is a unique EMM with the property that orthogonal martingales remain martingales under the change of measure. It is known as the minimal martingale measure and is used for pricing in this situation.

In Chapter 6 we explore the idea of minimizing the terminal variance of the portfolio, while insisting on the self-financing condition. A Hilbert projection argument is used to find a mean-variance optimal strategy, which is linked to the previous local risk-minimizing approach. The so called variance-optimal martingale measure is shown to be the measure under which pricing takes place. Under certain conditions it is possible to show that the minimal martingale measure and variance-optimal measure coincide.

Finally, in Chapter 7 we apply the theory developed throughout the dissertation to the problem of hedging basis risk. Here the idea is to hedge an option on an asset in which trading is restricted, using a closely correlated asset. The risk associated with this problem is called basis risk. Both the locally risk-minimizing and mean-variance optimal strategies are applied. Closed-form solutions based on the Black-Scholes formula are derived and compared with the numerical results obtained by Monoyios [68], where a utility indifference approach was employed.

Before continuing, a brief comment on the expository style of this dissertation. One of our aims has been to produce a self-contained document that could be used by relative newcomers as an introduction to the semimartingale theory of financial markets in general and to quadratic approaches to incomplete markets in particular. This has resulted in a style of proof, for example, which a well-informed reader may find somewhat tedious. (For this we apologize; but our approach has been rather to include too much detail than too little.)

Chapter 2

Mathematical Preliminaries

In this chapter the mathematical notation and theory required for the rest of the dissertation is presented. A background in probability theory is assumed, for which a number of good references are available [10, 50, 92, 95]. A small amount of functional analysis is also used, for which the accounts of Kreyzig [58] and Luenberger [62] are useful.

The aim here is to provide a resumé, in the spirit of Jacod and Shiryaev [51], for stochastic integration and the supplementary results that are needed. The emphasis will be on the statement of results and the literature will be cited for proofs.

In constructing a theory of stochastic integration, this chapter follows the classical approach of Jacod and Shiryaev [51], while the section on martingale representation follows the account of Protter [74]. Although these are our primary references, other sources [9, 24, 28, 29, 55, 75, 78, 79] have been consulted and have influenced our definitions and terminology.

The chapter commences with initial sections on stochastic processes, measurability, filtrations, stopping times, martingales and finite variation processes. With the preliminaries under the belt, a summary of the general theory of stochastic integration is presented. Finally, there are also sections on quadratic variation, changes of measure and martingale representation.

2.1 Stochastic Processes

In this section we introduce the basic concepts and definitions of stochastic processes. We start by fixing a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ for the remainder of this chapter. All random variables are assumed to be $\mathscr{F}$-measurable.

Definition 2.1. A *stochastic process*, denoted by $X = (X_t)_{t \in I}$, is a family of real-valued random variables $X_t : \Omega \to \mathbb{R}$, indexed by $t \in I$, where I is some index set. The two cases of most interest are when $I = \mathbb{N}$, in which case X is called a discrete-time process; and when I is a subinterval of $\mathbb{R}_+$, in which case X is called a continuous-time process.

In this dissertation we are only concerned with continuous-time processes. Note that this includes the special case of a finite time-horizon, considered later in the dissertation, where $I = [0, T]$, for some $T \in (0, \infty)$. For the remainder of this chapter, however, we fix $I = \mathbb{R}_+$. From the point of view of mathematical finance, stochastic processes are used to model the time evolution of the prices of financial instruments.

Definition 2.2. For a fixed sample point $\omega \in \Omega$, the map $t \mapsto X_t(\omega)$, for $t \in \mathbb{R}_+$, is called a *sample path* or *trajectory* of a stochastic process X.

Since we consider general processes, the sample paths may include discontinuities, in the form of jumps.

Definition 2.3. A function $x : \mathbb{R}_+ \to \mathbb{R}$ is called *RCLL* or *"right continuous with left limits"*, if the one-sided limits x_{t+} and x_{t-} exist finitely and $x_t = x_{t+}$, for all $t \in \mathbb{R}_+$. Similarly x is called *LCRL* or *"left continuous with right limits"*, if the same one-sided limits exist finitely and $x_t = x_{t-}$, for all $t \in \mathbb{R}_+$. (Following the convention of Jacod and Shiryaev [51, p. 3], we set $x_{0-} := x_0$.) A process is called RCLL (resp. LCRL) if its sample paths are RCLL (resp. LCRL) almost surely.

RCLL (resp. LCRL) processes are also sometimes called R-processes (resp. L-processes) or càdlàg (resp. càglàd), an acronym from the French "continu à droite limites à gauche" (resp. "continu à gauche limites à droite"). Other authors describe them as CORLOL (resp. COLLOR), the acronym for "continuous on the right with limits on the left" (resp. "continuous on the left with limits on the right").

Definition 2.4. An RCLL process X may possess finite jump-discontinuities. The *jump process* of X, denoted by ΔX, is defined by

$$\Delta X_t := X_t - X_{t-},$$

for all $t \in \mathbb{R}_+$.

Since $X_{0-} = X_0$, we have $\Delta X_0 = 0$. It is also clear from this definition that if X has continuous paths, then $\Delta X = 0$.

2.2 Measurability and Filtrations

In the theory of stochastic processes the idea of the flow of information plays a central role. This is formalized through the concept of a filtration.

Definition 2.5. A process X is said to be *measurable* if the map $\Omega \times \mathbb{R}_+ \to \mathbb{R} : (\omega, t) \mapsto X_t(\omega)$ is $\mathscr{F} \otimes \mathscr{B}(\mathbb{R}_+)$-measurable. It is said to be *progressively measurable* if this map is $\mathscr{F}_t \otimes \mathscr{B}([0,t])$-measurable, for each $t \in \mathbb{R}_+$. The σ-algebra on $\Omega \times \mathbb{R}_+$ generated by all progressively measurable processes, namely

$$\mathscr{M} := \sigma \{ X : \Omega \times \mathbb{R}_+ \to \mathbb{R} \,|\, X \text{ is progressively measurable} \},$$

is called the *progressive σ-algebra*.

Definition 2.6. A family of σ-algebras $\mathbf{F} = (\mathscr{F}_t)_{t \in \mathbb{R}_+}$ satisfying $\mathscr{F}_s \subseteq \mathscr{F}_t \subseteq \mathscr{F}$, for all $s \le t \in \mathbb{R}_+$, is called a *filtration*. When we endow the probability space $(\Omega, \mathscr{F}, \mathbb{P})$ with a filtration, we refer to the combined structure $(\Omega, \mathscr{F}, \mathbf{F}, \mathbb{P})$ as a *stochastic basis* or a *filtered probability space*.

Intuitively, a filtration provides a mechanism for accumulating information over time.

Definition 2.7. A stochastic basis $(\Omega, \mathscr{F}, \mathbf{F}, \mathbb{P})$ is called *complete* if the σ-algebra $\mathscr{F}$ is $\mathbb{P}$-complete[1] and $\mathscr{F}_0$ contains all the $\mathbb{P}$-null sets of $\mathscr{F}$. A stochastic basis is said to satisfy the *usual conditions* if it is complete and the filtration is *right continuous*; that is

$$\mathscr{F}_t = \mathscr{F}_{t+} := \bigcap_{t>s} \mathscr{F}_s,$$

for all $t \in \mathbb{R}_+$.

For the remainder of this chapter we fix a filtration $\mathbf{F} = (\mathscr{F}_t)_{t \in \mathbb{R}_+}$ and assume that the stochastic basis $(\Omega, \mathscr{F}, \mathbf{F}, \mathbb{P})$ satisfies the usual conditions.

Definition 2.8. A process X is said to be *adapted* to $\mathbf{F}$ if X_t is $\mathscr{F}_t$-measurable, for every $t \in \mathbb{R}_+$. The space of all RCLL adapted processes is denoted by $\mathcal{R}$, while $\mathcal{L}$ denotes the space of all LCRL adapted processes.

Note that from now on it is implicitly assumed that all processes are adapted, unless specified otherwise. For the next definitions it is useful to think of a stochastic process as a map $\Omega \times \mathbb{R}_+ \to \mathbb{R}$, rather than as an indexed family of random variables.

Definition 2.9. The σ-algebra on $\Omega \times \mathbb{R}_+$ generated by the processes in $\mathcal{L}$, namely

$$\mathscr{P} := \sigma \{ X : \Omega \times \mathbb{R}_+ \to \mathbb{R} \,|\, X \in \mathcal{L} \},$$

is called the *predictable σ-algebra*. Similarly, the *optional σ-algebra* is generated by the processes in $\mathcal{R}$ and is denoted by $\mathscr{O}$.

[1] Recall that $\mathscr{F}$ is $\mathbb{P}$-complete iff $A \subseteq B$, with $B \in \mathscr{F}$ such that $\mathbb{P}(B) = 0$, implies that $A \in \mathscr{F}$.

Processes are called predictable or optional if they are measurable with respect to $\mathscr{P}$ or $\mathscr{O}$, respectively. Stochastic intervals, defined in the next section, may be used to provide alternative formulations of the predictable and optional σ-algebras. It is easily demonstrated that predictable processes are optional; optional processes are progressively measurable; and progressively measurable processes are measurable. In other words, we have the following inclusions of σ-algebras:

$$\mathscr{P} \subseteq \mathscr{O} \subseteq \mathscr{M} \subseteq \mathscr{F} \otimes \mathscr{B}(\mathbb{R}_+)$$

(see Chung and Williams [9, §3.2, p. 57] for details).

2.3 Stopping times

Definition 2.10. A random variable $\tau : \Omega \to \overline{\mathbb{R}}_+$ is called a *stopping time* if $\{\tau \leq t\} \in \mathscr{F}_t$, for all $t \in \mathbb{R}_+$.

If τ and σ are stopping times and $t, c \in \mathbb{R}_+$, with $c \geq 1$, then it can be shown that t, $\sigma \wedge \tau$, $\sigma \vee \tau$, $\tau + \sigma$ and $c\tau$ are all stopping times, while $\tau - \sigma$ is not necessarily a stopping time.

Definition 2.11. Let τ be a stopping time. The *stopping time σ-algebra* $\mathscr{F}_\tau$ is the collection of all events $A \in \mathscr{F}$, such that $A \cap \{\tau \leq t\} \in \mathscr{F}_t$, for all $t \in \mathbb{R}_+$.

Intuitively, a stopping time σ-algebra can be thought of as the collection of events that have occurred (or not occurred) up to the stopping time.

Definition 2.12. Given two stopping times σ and τ, with $\sigma \leq \tau$ a.s., *stochastic intervals* may be defined as follows:

$$\begin{aligned}
[\![\sigma, \tau]\!] &:= \{\, (\omega, t) \in \Omega \times \mathbb{R}_+ \,|\, \sigma(\omega) \leq t \leq \tau(\omega)\}\,; \\
[\![\sigma, \tau[\![&:= \{\, (\omega, t) \in \Omega \times \mathbb{R}_+ \,|\, \sigma(\omega) \leq t < \tau(\omega)\}\,; \\
]\!]\sigma, \tau]\!] &:= \{\, (\omega, t) \in \Omega \times \mathbb{R}_+ \,|\, \sigma(\omega) < t \leq \tau(\omega)\}\,; \\
]\!]\sigma, \tau[\![&:= \{\, (\omega, t) \in \Omega \times \mathbb{R}_+ \,|\, \sigma(\omega) < t < \tau(\omega)\}\,.
\end{aligned}$$

The stochastic interval $[\![\tau, \tau]\!]$ is often written as $[\![\tau]\!]$ and is called the *graph* of the stopping time τ.

For stopping times σ and τ, with $\sigma \leq \tau$ a.s., the stochastic process $\mathbb{1}_{]\!]\sigma,\tau]\!]}$ is clearly a member of $\mathcal{L}$ and hence predictable. In fact, an alternative characterization of $\mathscr{P}$ is as the σ-algebra generated by $\{\, \{0\} \times A \,|\, A \in \mathscr{F}_0\}$ and the stochastic intervals $]\!]0, \tau]\!]$, for all stopping times τ. This follows from the fact that all LCRL processes may be expressed as limits of convergent sequences[2] of elementary processes (see Definition 2.32). These ideas will be formalized later, when we introduce the concept of stochastic integration.

By a similar characterization of RCLL processes, it may also be shown that $\mathscr{O}$ is the σ-algebra generated by the stochastic intervals $[\![0, \tau[\![$, for all stopping times τ. Furthermore, all other stochastic intervals may be constructed from such intervals using only countable operations[3].

Definition 2.13. Given a process X and an a.s. finite stopping time τ, the *process stopped at time* τ, denoted by X^τ, is defined by

$$X_t^\tau := X_{\tau \wedge t} = \mathbb{I}_{\{t<\tau\}} X_t + \mathbb{I}_{\{t \geq \tau\}} X_\tau,$$

for all $t \in \mathbb{R}_+$.

Definition 2.14. Suppose $\mathcal{C}$ is a family of processes. A process X is a member of the corresponding *localized* family of processes, denoted by $\mathcal{C}_{\mathsf{loc}}$, if there exists an increasing sequence of stopping times $(\tau_n)_{n\in\mathbb{N}}$, with $\lim_{n\to\infty} \tau_n = \infty$ a.s., such that $X^{\tau_n} \in \mathcal{C}$, for each $n \in \mathbb{N}$. The sequence of stopping times is called a *reducing* or *localizing sequence*.

2.4 Martingales

The theory of martingales is central to everything we do and consequently we review the basic concepts.

Definition 2.15. Let $p \geq 1$. The family of random variables $X : \Omega \to \overline{\mathbb{R}}$, such that

$$\|X\|_{L^p} := (\mathbb{E}[|X|^p])^{\frac{1}{p}} = \left(\int_\Omega |X|^p \, d\mathbb{P} \right)^{\frac{1}{p}} < \infty,$$

is denoted by $\mathscr{L}^p(\Omega, \mathscr{F}, \mathbb{P})$. A random variable X is called *integrable* (resp. *square integrable*) if $X \in \mathscr{L}^1(\Omega, \mathscr{F}, \mathbb{P})$ (resp. $X \in \mathscr{L}^2(\Omega, \mathscr{F}, \mathbb{P})$). We define an equivalence relation on $\mathscr{L}^p(\Omega, \mathscr{F}, \mathbb{P})$, by setting

$$X \sim Y \text{ iff } X = Y \quad \text{a.s.},$$

for all $X, Y \in \mathscr{L}^p(\Omega, \mathscr{F}, \mathbb{P})$. Then $L^p(\Omega, \mathscr{F}, \mathbb{P})$ is defined as the corresponding family of equivalence classes. For convenience, we abbreviate $L^p(\Omega, \mathscr{F}, \mathbb{P})$ to $L^p(\mathbb{P})$ or L^p if the components of the probability triple are clear from the context.

[2] With a suitable definition of convergence.
[3] For example, $]\!]\sigma, \tau]\!] = (\cap_{n\in\mathbb{N}} [\![0, \tau + 1/n[\![) \cap (\cap_{n\in\mathbb{N}} [\![0, \sigma + 1/n, [\![)^c$.

Note that $\|\cdot\|_{L^p}$ induces a metric on L^p, defined by $d_p(X,Y) := \|X-Y\|_{L^p}$, for each $X, Y \in L^p$, where $p \geq 1$. It can be shown that (L^p, d_p) is a Banach space (see e.g. [50] or [78, p. 100]).

Definition 2.16. If a process X satisfies the condition

$$\lim_{n\to\infty} \sup_{t\in\mathbb{R}_+} \int_{\{|X_t|\geq n\}} |X_t| \, d\mathbb{P} = 0,$$

then it is said to be *uniformly integrable*.

Definition 2.17. A process M is called a *martingale* if

1. M is adapted to $\mathbf{F}$;

2. M_t is integrable, for all $t \in \mathbb{R}_+$; and

3. $\mathbb{E}[M_t \,|\, \mathscr{F}_s] = M_s$ a.s., for all $s \leq t \in \mathbb{R}_+$.

The family of all *uniformly integrable martingales* is denoted by $\mathcal{M}$.

The following useful characterization of martingales is often used in applications to establish the martingale property for a given process.

Lemma 2.18. *Let M be an adapted process such that M_t is integrable, for all $t \in \mathbb{R}_+$. Then M is a martingale iff*

$$\mathbb{E}[\mathbb{1}_A M_t] = \mathbb{E}[\mathbb{1}_A M_s], \tag{2.1}$$

for all $s \leq t \in \mathbb{R}_+$ and all $A \in \mathscr{F}_s$.

Proof. ($\Rightarrow$) Suppose M is a martingale; and choose $s \leq t \in \mathbb{R}_+$ and $A \in \mathscr{F}_s$. Then

$$\mathbb{E}[\mathbb{1}_A M_t] = \mathbb{E}[\mathbb{E}[\mathbb{1}_A M_t \,|\, \mathscr{F}_s]] = \mathbb{E}[\mathbb{1}_A \mathbb{E}[M_t \,|\, \mathscr{F}_s]] = \mathbb{E}[\mathbb{1}_A M_s].$$

($\Leftarrow$) Suppose M satisfies (2.1) and choose $s \leq t \in \mathbb{R}_+$. Then, by the definition of conditional expectations (see Jacod and Protter [50, Def. 23.5, p. 200]), $\mathbb{E}[M_t \,|\, \mathscr{F}_s]$ is the a.s. unique $\mathscr{F}_s$-measurable random variable satisfying

$$\mathbb{E}[\mathbb{1}_A \mathbb{E}[M_t \,|\, \mathscr{F}_s]] = \mathbb{E}[\mathbb{1}_A M_t],$$

for all $A \in \mathscr{F}_s$. But, by assumption, M_s satisfies this condition as well. Hence $\mathbb{E}[M_t \,|\, \mathscr{F}_s] = M_s$. ■

Definition 2.19. Let $M \in \mathcal{M}$. If $\sup_{t\in\mathbb{R}_+} \mathbb{E}\left[M_t^2\right] < \infty$, then M is said to belong to the family of *square integrable martingales*, denoted by $\mathcal{M}^2$.

Using Definition 2.14, the family of local martingales is denoted by $\mathcal{M}_{\mathsf{loc}}$. Furthermore, the family of all uniformly integrable martingales (resp. local martingales) null at time zero is denoted by $\mathcal{M}_0$ (resp. $\mathcal{M}_{0,\mathsf{loc}}$). The same conventions apply to the family of square integrable martingales. It is easily seen that every martingale is a local martingale. The converse is not true, however (see Durrett [28, §2.2, p. 41–42] for counterexamples).

Definition 2.20. A process X is said to admit a *terminal variable*, denoted by X_∞, if $\lim_{t\to\infty} X_t =: X_\infty$ exists a.s. and X_∞ is integrable.

Theorem 2.21. *Let $M \in \mathcal{M}$. Then M admits a terminal variable M_∞ and $M_t \to M_\infty$ in L^1 as well as a.s. In this case, $M_\tau = \mathbb{E}\left[M_\infty \,|\, \mathscr{F}_\tau\right]$, for all stopping times τ. Furthermore, $M \in \mathcal{M}^2$ iff $M_\infty \in L^2$, in which case $M_t \to M_\infty$ in L^2 as well.*

Proof. See Jacod and Shiryaev [51, Prop. 1.42, p. 11]. ■

The last theorem allows us to characterize $\mathcal{M}^2$ as a Hilbert space. Since there is a bijective correspondence between square integrable martingales $M \in \mathcal{M}^2$ and their terminal variables M_∞, we can define an inner product $(\,\cdot\,,\,\cdot\,)_{H^2}$ and a norm $\|\cdot\|_{H^2}$ on $\mathcal{M}^2$, by setting

$$(M,N)_{H^2} := \mathbb{E}\left[M_\infty N_\infty\right]; \quad \text{and}$$
$$\|M\|_{H^2} := \|M_\infty\|_{L^2},$$

for all $M, N \in \mathcal{M}^2$.

Finally we state two fundamental theorems for martingales: the *optional sampling theorem* and *Doob's maximal quadratic inequality*. The latter may be used to show convergence of Cauchy sequences in $\mathcal{M}^2$, using the norm defined above (see Jacod and Shiryaev [51, Lem. 4.7, p. 39] for details).

Theorem 2.22 (Optional Sampling Theorem)**.** *Let $M \in \mathcal{M}$. Then*

$$\mathbb{E}\left[M_\tau \,|\, \mathscr{F}_\sigma\right] = M_\sigma,$$

for all stopping times $\sigma \le \tau$.

Proof. See Dellacherie and Meyer [24, V§2, p. 6–10]. ■

Theorem 2.23 (Doob's Maximal Quadratic Inequality)**.** *Let $M \in \mathcal{M}^2$, with terminal variable M_∞. Then*

$$\mathbb{E}\left[\sup_{t\in\mathbb{R}_+} M_t^2\right] \le 4\mathbb{E}\left[M_\infty^2\right].$$

Proof. See Dellacherie and Meyer [24, V§24, p. 17]. ■

2.5 Finite Variation Processes

The variation (of some order) of a process is defined as the limit of sums of powers of its increments along a sequence of partitions.

Definition 2.24. Let $m \in \mathbb{N}$ and $a < b \in \mathbb{R}_+$. A *partition* of the interval $[a,b]$ is a finite ordered set $\pi_{a,b} := \{t_0, t_1, \ldots, t_m\}$, such that $a = t_0 \leq t_1 \leq \cdots \leq t_m = b$. For a partition[4] π, the quantity

$$\|\pi\| := \sup_{0 \leq i < m} |t_{i+1} - t_i|$$

is called the *mesh* of π.

Definition 2.25. Let $t \in \mathbb{R}_+$. A sequence of partitions $(\pi_t^n)_{n\in\mathbb{N}}$ of the interval $[0,t]$, with $\pi_t^n = \{t_0^n, \ldots, t_{m^n}^n\}$ for each $n \in \mathbb{N}$ and $\lim_{n\to\infty} \|\pi_t^n\| = 0$, is called a *Riemann sequence.*

Definition 2.26. Let $X \in \mathcal{R}$ and for each $t \in \mathbb{R}_+$ choose a Riemann sequence $(\pi_t^n)_{n\in\mathbb{N}}$. For $p > 0$, the p^{th} variation of X, for the partition π_t^n, is defined by

$$S^{(p)}(X, \pi_t^n) := \sum_{i=0}^{m_n - 1} |X_{t_{i+1}^n} - X_{t_i^n}|^p.$$

If, for each $t \in \mathbb{R}_+$, the limit

$$\lim_{n\to\infty} S^{(p)}(X, \pi_t^n) =: V_t^{(p)}(X)$$

exists a.s., then the process $V^{(p)}(X)$ is well-defined and is called the p^{th} *variation* of X.

Definition 2.27. A process $A \in \mathcal{R}$ is called a *finite variation* process if almost all its sample paths are of finite first variation on compacts. In other words, we require

$$\int_0^t |dA|_s := V_t^{(1)}(A) < \infty \quad \text{a.s.},$$

for all $t \in \mathbb{R}_+$. We denote the family of all finite variation processes (resp. finite variation processes null at time zero) by $\mathcal{V}$ (resp. $\mathcal{V}_0$). Processes $A \in \mathcal{V}$ (resp. $A \in \mathcal{V}_0$), such that $\mathbb{E}\left[V_\infty^{(1)}(A)\right] < \infty$, are called processes of *integrable variation* and the family of such processes is denoted by $\mathcal{A}$ (resp. $\mathcal{A}_0$). (Note that for any $X \in \mathcal{R}$, $V^{(1)}(X)$ is an increasing process. Consequently the a.s. limit $V_\infty^{(1)}(X) := \lim_{t\to\infty} V_t^{(1)}(X)$ exists.)

[4] Note that the one or more of the subscripts of π will be omitted when their values are clear from the context.

2.6 Semimartingales and Stochastic Integrals

There are two main approaches to the construction of the stochastic integral. Historically, the first rigorous account of stochastic integration was published in 1944 by Itô [47, 49]. In formulating a theory of integration with respect to Wiener processes, his key insight was to limit the space of integrands to non-anticipating (adapted) processes. This allowed him to establish the famous Itô isometry, which was central to providing a well-defined integral. A little later, in 1951, he published the now-famous Itô formula [49]. In 1953 Doob [26] conjectured that this approach could be extended, by using the full strength of the independent increments property, to define integrals with respect to continuous martingales. This programme was realized by Meyer [65, 66] in 1962–1963, with the proof of the Doob-Meyer decomposition theorem. As a result, a more systematic treatment of integration with respect to square integrable martingales was put forward by Courrège [12] in 1963. He also pioneered the use of the predictable σ-algebra.

The Itô formula for square integrable martingales was derived by Kunita and Watanabe [59] in 1967. This paper was influential because of its use of the concepts of orthogonality and quadratic variation. Meyer extended their approach, synthesizing a number of other concepts available at the time, in a series of papers published in the Séminaire de Probabilités. This culminated in the classical definition of a semimartingale as the sum of a locally square integrable martingale and a finite variation process. Later the square integrability condition was removed. For an interesting article on the early history of stochastic integration, consult Jarrow and Protter [54].

The second, more modern, approach to stochastic integration, popularized by Protter [72, 74], uses the idea of the stochastic integral defined as a "Riemann-type limit of sums". This analogy is flawed, however, since the classical Riemann-Stieltjes construction only works when either the integrator or the integrand is of finite variation. By considering stochastic integrals of simple predictable processes, semimartingales are defined implicitly as those processes, which when used as integrators, yield well-defined integrals. These simple integrals are then extended to integrals of LCRL processes and then further to integrals of general predictable processes. Finally it is shown, using results from Bichteler [4] and Dellacherie [25], that this definition of semimartingales is equivalent to the classical definition. We provide a summary of the classical approach, as presented in the account of Jacod and Shiryaev [51].

Definition 2.28. A *semimartingale* is a process X of the form

$$X = X_0 + M + A, \tag{2.2}$$

where X_0 is finite and $\mathscr{F}_0$-measurable; $M \in \mathcal{M}_{0,\mathsf{loc}}$ and $A \in \mathcal{V}_0$. The family of all semimartingales is denoted by $\mathcal{S}$.

Definition 2.29. Suppose $X \in \mathcal{S}$ possesses a decomposition (2.2) such that A is predictable. Then X is a member of the family of *special semimartingales*, denoted by $\mathcal{S}_{\mathsf{p}}$. The decomposition with the above property is necessarily unique and is called the *canonical decomposition* of X.

Definition 2.30. Let $X \in \mathcal{S}$, with decomposition (2.2). If $M \in \mathcal{M}_0^2$ and $A \in \mathcal{V}_0$ satisfies

$$\mathbb{E}\left[\left(\int_0^\infty |dA|_s\right)^2\right] < \infty,$$

then X is a member of the family of *square integrable semimartingales*, denoted by $\mathcal{S}^2$.

Semimartingales with bounded jumps are special (see Jacod and Shiryaev [51, Lem. 4.24, p. 44]). Square integrable semimartingales are also special (see Dellacherie and Meyer [24, VII§98, p. 294–297]).

At its heart, the construction of the stochastic integral relies on a convergence argument. The next definition specifies the mode of convergence.

Definition 2.31. A sequence of processes $(\phi_n)_{n\in\mathbb{N}}$ is said to converge *uniformly on compacts in probability* (*UCP*) if

$$\sup_{0\le s\le t} |\phi_s^n - \phi_s| \to 0$$

in probability, for all $t \in \mathbb{R}_+$.

We now define the class of processes used as the building blocks for the construction of the stochastic integral.

Definition 2.32. A stochastic process ϕ is called an *elementary* process if, for some $n \in \mathbb{N}$, there exists a finite sequence of stopping times $0 = \tau_0 \le \tau_1 \le \ldots \le \tau_n < \infty$ and a corresponding sequence of finite-valued $\mathscr{F}_{\tau_i}$-measurable random variables ξ_i, for each $i \in \{0, \ldots, n-1\}$, such that

$$\phi_t(\omega) = \xi_0(\omega)\mathbb{1}_{[\![0]\!]}(\omega,t) + \sum_{i=0}^{n-1} \xi_i(\omega)\mathbb{1}_{]\!]\tau_i,\tau_{i+1}]\!]}(\omega,t), \tag{2.3}$$

for all $(\omega, t) \in \Omega \times \mathbb{R}_+$. The class of elementary processes is denoted by $\mathcal{E}$.

Since elementary processes are LCRL, they are also measurable with respect to $\mathscr{P}$. It is now possible to define the integral of an elementary process with respect to a semimartingale. This is done by direct analogy to the sum (1.2).

Definition 2.33. Let $X \in \mathcal{S}$ and suppose $\phi \in \mathcal{E}$ can be expressed as (2.3). Then

$$I_X(\phi) := \sum_{i=0}^{n-1} \xi_i (X^{\tau_{i+1}} - X^{\tau_i})$$

defines a linear mapping $I_X : \mathcal{E} \to \mathcal{S}$. The process $I_X(\phi)$ is called the *elementary stochastic integral* of ϕ with respect to X.

Theorem 2.34. *Let $X \in \mathcal{S}$. The mapping $I_X : \mathcal{E} \to \mathcal{S}$ of Definition 2.33 has an extension from the space of elementary processes to the space of all locally bounded predictable processes.*

Proof. See Jacod and Shiryaev [51, Thm. 4.31, p. 46-51]. ■

Note that the stochastic integral can be further extend to processes that are not locally bounded, in which case the class of predictable processes integrable with respect to X is denoted by $\mathcal{L}(X)$ (see Dellacherie and Meyer [24, VIII§69–75, p. 377–385] for details).

We now outline the four main steps in the proof of Theorem 2.34. We also employ the suggestive notation $\int_0^{\cdot} \phi_s \, dX_s$ to indicate the semimartingale $I_X(\phi)$. For convenience, we fix a semimartingale X with the decomposition (2.2) and a locally bounded predictable process ϕ.

Step 1: Assume that $X = A$. Then $\int_0^{\cdot} \phi_s \, dA_s$ is just the Lebesgue-Stieltjes integral, defined pathwise.

Step 2: Now assume that $X = M$, where $M \in \mathcal{M}_0^2$. In order to construct a stochastic integral with respect to a square integrable martingale, we require some extra definitions and theorems.

The Doob-Meyer decomposition of a submartingale of class (D) (see [24, VII§8–9, p. 194–195]) establishes the existence of a finite variation process which, when subtracted from the original submartingale, yields a martingale. We now present a special case of the Doob-Meyer decomposition theorem.

Theorem 2.35. *Let $M, N \in \mathcal{M}^2_{\mathsf{loc}}$. There exists a unique non-decreasing predictable process $\langle M \rangle \in \mathcal{V}_0$, such that $M^2 - \langle M \rangle$ is a local martingale. Furthermore, there exists a unique predictable process $\langle M, N \rangle \in \mathcal{V}_0$, defined by the polarization identity*

$$\langle M, N \rangle := \frac{1}{4}\big(\langle M + N \rangle - \langle M - N \rangle\big),$$

such that $MN - \langle M, N \rangle$ is a local martingale. Furthermore, if $M, N \in \mathcal{M}^2$, then $\langle M, N \rangle \in \mathcal{A}_0$ and $MN - \langle M, N \rangle \in \mathcal{M}$.

Proof. See Jacod and Shiryaev [51, Prop. 4.2, p. 38]. ■

Let $M, N \in \mathcal{M}^2_{\mathsf{loc}}$. The process $\langle M \rangle$ is called the *angle brackets process* of M; while $\langle M, N \rangle$ is called the *predictable quadratic covariation* of M and N (for reasons that will become apparent in the next section).

Definition 2.36. Let $M \in \mathcal{M}^2_{\mathsf{loc}}$. The family of all predictable processes ϕ, such that process $\int_0^\cdot \phi_s^2 \, d\langle M \rangle_s$ is integrable (resp. locally integrable), is denoted by $\mathcal{L}^2(M)$ (resp. $\mathcal{L}^2_{\mathsf{loc}}(M)$)

Definition 2.37. Let $M \in \mathcal{M}^2$. The positive finite measure μ_M on $(\Omega \times \mathbb{R}_+, \mathscr{P})$, defined by

$$\mu_M\{B\} := \mathbb{E}\left[\int_0^\infty \mathbb{1}_B(\omega, t) \, d\langle M \rangle_s(\omega)\right],$$

for all $B \in \mathscr{P}$, is called the *Doléans measure*.

The Doléans measure[5] enables a characterization of $\mathcal{L}^2(M)$ as a Hilbert space $L^2(\Omega \times \mathbb{R}_+, \mathscr{P}, \mu_M)$. This in turn lays the foundation for the following fundamental result.

Lemma 2.38 (Itô Isometry)**.** *Let $\phi \in \mathcal{E}$, with representation* (2.3)*. Then*

$$\mathbb{E}\left[(I_X(\phi)_t)^2\right] = \mathbb{E}\left[\left(\int_0^t \phi_s \, dM_s\right)^2\right] = \mathbb{E}\left[\int_0^t \phi_s^2 \, d\langle M \rangle_s\right],$$

for all $t \in \mathbb{R}_+$.

Proof. This follows directly from the fact that

$$\begin{aligned}\left(\int_0^\cdot \phi_s \, dM_s\right)^2 - \int_0^\cdot \phi_s^2 \, d\langle M \rangle_s &= 2\sum_{i=0}^{n-1}\sum_{j=0}^{n-1} \xi_i \xi_j (M^{\tau_{i+1}} - M^{\tau_i})(M^{\tau_{j+1}} - M^{\tau_j}) \\ &\quad + \sum_{i=0}^{n-1} \xi_i^2 \left[(M^{\tau_{i+1}})^2 - \langle M \rangle^{\tau_{i+1}} - (M^{\tau_i})^2 + \langle M \rangle^{\tau_i} \right. \\ &\qquad \left. - 2M^{\tau_i}(M^{\tau_{i+1}} - M^{\tau_i})\right]\end{aligned}$$

is a martingale. ■

Since $\mathcal{E}$ is a dense subspace of $L^2(\Omega \times \mathbb{R}_+, \mathscr{P}, \mu_M)$, the above result allows for a unique continuous extension of I_M to an isometry of $L^2(\Omega \times \mathbb{R}_+, \mathscr{P}, \mu_M)$ and $\mathcal{M}^2$. This extends the definition of I_M to all integrands in $\mathcal{L}^2(M)$.

Step 3: Suppose $X = M$, where $M \in \mathcal{M}^2_{0,\mathsf{loc}}$. The integral $\int_0^\cdot \phi_s \, dM_s$, for $\phi \in \mathcal{L}^2_{\mathsf{loc}}(M)$, can now be constructed from the integral in step 2 using a localization procedure.

[5] See Chung and Williams [9, p. 33] for a justification of the name.

Step 4: Finally, remove all restrictions on X. It can be shown that X possesses at least one decomposition of the form (2.2), with $M \in \mathcal{M}^2_{0,\mathsf{loc}}$ and $A \in \mathcal{V}_0$ (see Jacod and Shiryaev [51, Prop. 4.17, p. 42] for details). Consequently, we can define

$$\int_0^{\cdot} \phi_s \, dX_s := \int_0^{\cdot} \phi_s \, dM_s + \int_0^{\cdot} \phi_s \, dA_s,$$

where the integrals on the right-hand side are defined as in steps 3 and 1, respectively.

This concludes the brief exposé of the proof of Theorem 2.34. Note that having defined the stochastic integral with respect to a locally square integrable martingale as another locally square integrable martingale, it follows from Theorem 2.35 that the stochastic integral possesses an angle brackets process. The following result indicates what this process looks like.

Corollary 2.39. *Let $M, N \in \mathcal{M}^2_{\mathsf{loc}}$ and let $\phi, \varphi \in \mathcal{L}^2(M)$, with representation* (2.3). *Then*

$$\left\langle \int_0^{\cdot} \phi_s \, dM_s \right\rangle = \int_0^{\cdot} \phi_s^2 \, d\langle M \rangle_s.$$

Furthermore,

$$\left\langle \int_0^{\cdot} \phi_s \, dM_s, \int_0^{\cdot} \varphi_s \, dN_s \right\rangle = \int_0^{\cdot} \phi_s \varphi_s \, d\langle M, N \rangle_s.$$

Proof. The first expression follows as a consequence of the uniqueness of the angle brackets process, while the second follows from the polarization identity of Theorem 2.35. ■

The following two theorems provide a number of useful properties of the stochastic integral.

Theorem 2.40. *Let $M \in \mathcal{M}_{\mathsf{loc}}$ and let ϕ be a locally bounded predictable process. Then $\int_0^{\cdot} \phi_s \, dM_s \in \mathcal{M}_{0,\mathsf{loc}}$.*

Proof. See Jacod and Shiryaev [51, 4.34, p. 47]. ■

Theorem 2.41. *Let $X, Y \in \mathcal{S}$, let $\phi, \varphi \in \mathcal{L}(X)$ and fix $\alpha, \beta \in \mathbb{R}$. Then the stochastic integral has the following properties:*

1. $\int_0^0 \phi_s \, dX_s = 0$;
2. $\Delta \left(\int_0^{\cdot} \phi_s \, dX_s \right) = \phi \Delta X$;
3. (*linearity*) $\int_0^{\cdot} (\alpha \phi_s + \beta \varphi_s) \, dX_s = \alpha \int_0^{\cdot} \phi_s \, dX_s + \beta \int_0^{\cdot} \varphi_s \, dX_s$;
4. (*associativity*) $\int_0^{\cdot} \phi_u \, d \left(\int_0^u \varphi_s \, dX_s \right) = \int_0^{\cdot} \phi_s \varphi_s \, dX_s$; *and*
5. *if* $\phi \in \mathcal{L}(Y)$, *then* $\int_0^{\cdot} \phi_s \, d(X+Y)_s = \int_0^{\cdot} \phi_s \, dX_s + \int_0^{\cdot} \phi_s \, dY_s$.

Proof. For the first four items, see Jacod and Shiryaev [51, p. 47–51]. For the last item, see Protter [74, Thm. 17, p. 164]. ∎

Finally, we provide a useful result that allows us to approximate the stochastic integral of an LCRL predictable process with Riemann sums.

Theorem 2.42. *Let $X \in \mathcal{S}$ and fix $\phi \in \mathcal{L}$. For each $t \in \mathbb{R}_+$, let $(\pi_t^n)_{n\in\mathbb{N}}$ be a Riemann sequence, with $\pi_t^n = \{t_0^n, \ldots, t_{m^n}^n\}$, for each $n \in \mathbb{N}$. For each $n \in \mathbb{N}$, define $I_X^n(\phi)$, by setting*

$$I_X^n(\phi)_t := \sum_{i=0}^{m_n-1} \phi_{t_i^n}(X_{t_{i+1}^n} - X_{t_i^n}),$$

for each $t \in \mathbb{R}_+$. Then the Riemann approximations $I_X^n(\phi)$ converge to $\int_0^\cdot \phi_s \, dX_s$ in UCP.

Proof. See Jacod and Shiryaev [51, Prop. 4.44, p. 51]. ∎

2.7 Quadratic Variation

Definition 2.43. Let X be a semimartingale. The *quadratic variation* process of X, denoted by $[X]$, is defined by setting

$$[X]_t := X_t^2 - X_0^2 - 2\int_0^t X_{s-} \, dX_s,$$

for all $t \in \mathbb{R}_+$.

It is clear from the definition that $[X]_0 = 0$. Note that since $[X]$ is non-decreasing (see Theorem 2.44 below), it is a finite variation process (i.e. $[X] \in \mathcal{V}_0$). The next result provides an intuitive understanding of quadratic variation, by linking it to Definition 2.26.

Theorem 2.44. *Let $X \in \mathcal{S}$. Then $V^{(2)}(X) = [X]$.*

Proof. Fix $t \in \mathbb{R}_+$ and let $(\pi_t^n)_{n\in\mathbb{N}}$ be a Riemann sequence, with $\pi_t^n = \{t_0^n, \ldots, t_{m^n}^n\}$, for each $n \in \mathbb{N}$. Then

$$\begin{aligned}
V_t^{(2)}(X) &= \lim_{n\to\infty} S^{(2)}(X, \pi_t^n) \\
&= \lim_{n\to\infty} \sum_{i=0}^{m_n-1} \left(X_{t_{i+1}^n} - X_{t_i^n} \right)^2 \\
&= \lim_{n\to\infty} \sum_{i=0}^{m_n-1} \left(X_{t_{i+1}^n}^2 - 2X_{t_{i+1}^n} X_{t_i^n} + X_{t_i^n}^2 \right) \\
&= \lim_{n\to\infty} \sum_{i=0}^{m_n-1} \left(X_{t_{i+1}^n}^2 - X_{t_i^n}^2 \right) - 2 \lim_{n\to\infty} \sum_{i=0}^{m_n-1} X_{t_i^n} (X_{t_{i+1}^n} - X_{t_i^n}) \qquad (2.4) \\
&= X_t^2 - X_0^2 - 2 \lim_{n\to\infty} I_X^n(X_-)_t \\
&= X_t^2 - X_0^2 - 2 \int_0^t X_{s-}\, dX_s \\
&= [X]_t,
\end{aligned}$$

since the first term in (2.4) is a telescoping sum, together with Theorem 2.42. ■

Definition 2.45 (Stochastic Integration by Parts). Let $X, Y \in \mathcal{S}$. The *covariation* of X and Y is the process $[X,Y]$, defined by

$$[X,Y]_t := X_t Y_t - X_0 Y_0 - \int_0^t X_{s-}\, dY_s - \int_0^t Y_{s-}\, dX_s, \qquad (2.5)$$

for all $t \in \mathbb{R}_+$.

Let $X, Y \in \mathcal{S}$. It is clear from the definition that $[X,Y]_0 = 0$. Furthermore, the *polarization identity*

$$[X,Y] = \frac{1}{2}\left([X+Y] - [X] - [Y] \right),$$

relating covariations and quadratic variations, follows from the linearity of stochastic integrals. Hence $[X,Y]$ is also a finite variation process. A result similar to Theorem 2.44 relates $[X,Y]$ to a constructive notion of covariation. The next two results provide four useful properties of the covariation process.

Proposition 2.46. *Let $X, Y \in \mathcal{S}$ and $A \in \mathcal{V}_0$. Then*

1. $\Delta[X,Y] = \Delta X \Delta Y$;
2. $[X,A]_t = \sum_{s\le t} \Delta X_s \Delta A_s$, *for all* $t \in \mathbb{R}_+$; *and*
3. *if τ is a finite stopping time, then* $[X^\tau, Y] = [X, Y^\tau] = [X^\tau, Y^\tau] = [X,Y]^\tau$.

Proof. For the first two items see Jacod and Shiryaev [51, Thm. 4.47 & Prop. 4.49, p. 52]. For the last item see Protter [74, Thm. 23, p. 68]. ■

Lemma 2.47 (Yoeurp)**.** *Let $M \in \mathcal{M}_{\mathsf{loc}}$ and suppose that $A \in \mathcal{V}$ is predictable. Then $[M, A]$ is a local martingale. (See also item (2) of Proposition 2.46.)*

Proof. See Dellacherie and Meyer [24, VII§36, p. 225]. ■

The next definition and proposition explain the terminology "predictable quadratic covariation" (see the comment after Theorem 2.35).

Definition 2.48. Let $A \in \mathcal{A}_{0,\mathsf{loc}}$. The unique predictable process $\widetilde{A} \in \mathcal{A}_{0,\mathsf{loc}}$, such that $A - \widetilde{A} \in \mathcal{M}_{\mathsf{loc}}$, is called the *compensator* of A (see Jacod and Shiryaev [51, Thm. 3.17, p. 32] for a proof of existence and other properties). It is also known as the *dual predictable projection*.

Proposition 2.49. *Let $M, N \in \mathcal{M}^2_{\mathsf{loc}}$. Then $[M, N] \in \mathcal{A}_{0,\mathsf{loc}}$ and its compensator is $\langle M, N\rangle$. Furthermore, if $M, N \in \mathcal{M}^2$, then $MN - [M, N] \in \mathcal{M}$.*

Proof. See Jacod and Shiryaev [51, Prop. 4.50, p. 53]. ■

Suppose $M, N \in \mathcal{M}^2_{\mathsf{loc}}$. It follows from Definition 2.48 that $[M, N] - \langle M, N\rangle \in \mathcal{M}_{\mathsf{loc}}$. In the case where $M, N \in \mathcal{M}^2$, it follows from Theorem 2.35 and Proposition 2.49 that $[M, N] - \langle M, N\rangle \in \mathcal{M}$. Furthermore, if M and N are continuous, then $[M, N] = \langle M, N\rangle$. In the case where we only have $M, N \in \mathcal{M}_{\mathsf{loc}}$, then $\langle M, N\rangle$ (defined as the compensator of $[M, N]$) only exists if $[M, N] \in \mathcal{A}_{0,\mathsf{loc}}$. A sufficient condition for this is that the jumps of M are bounded, i.e. $|\Delta M| \le c$, for some $c \in \mathbb{R}_+$ (see Jacod and Shiryaev [51, Thm. 3.14, p. 169] for a justification). Note that these considerations provide a way of extending the definition of the angle brackets process to semimartingales, by setting

$$\langle X, Y\rangle := \widetilde{[X, Y]}$$

for $X, Y \in \mathcal{S}$, provided this makes sense.

One needs to be cautious when dealing with angle brackets processes, since they are not invariant under changes of measure (see Protter [74, p. 123]). For this reason we will use the notation $\langle\,\cdot\,,\,\cdot\,\rangle^{\mathbb{Q}}$ to denote angle brackets under any probability measure $\mathbb{Q} \sim \mathbb{P}$.

We now provide a property of quadratic covariation which allows us to prove a useful property that will be used extensively in this dissertation.

Theorem 2.50. *Let $X, Y \in \mathcal{S}$ and let X^c, Y^c denote their respective continuous local martingale parts. Then*

$$\begin{aligned}[X, Y]_t &= [X^c, Y^c]_t + \sum_{0\le s\le t} \Delta X_s \Delta Y_s \\ &= \langle X^c, Y^c\rangle_t + \sum_{0\le s\le t} \Delta X_s \Delta Y_s,\end{aligned}$$

for all $t \in \mathbb{R}_+$

Proof. See Jacod and Shiryaev [51, Thm. 4.52, p. 55]. ■

Theorem 2.51. *Let* $X, Y \in \mathcal{S}$ *and let* $\phi \in \mathcal{L}(X)$*. Then*

$$\left[\int_0^{\cdot} \phi_s \, dX_s, Y\right]_t = \int_0^t \phi_s \, d[X,Y]_s,$$

for all $t \in \mathbb{R}_+$.

Proof. This follows directly from Theorem 2.50, Corollary 2.39 and property 2 of Theorem 2.41. ■

2.8 The Doléans Exponential and the Stochastic Logarithm

We now focus on a very important class of semimartingales which, as we shall see in Section 3.4, play the role of density processes (see Section 2.9) for equivalent martingale measures.

Definition 2.52. Let $Y \in \mathcal{S}$. The *Doléans exponential* or *stochastic exponential* of Y, denoted by $\mathscr{E}(Y)$, is the unique strong solution of the SDE

$$d\mathscr{E}(Y)_t = \mathscr{E}(Y)_{t-} \, dY_t, \tag{2.6}$$

for all $t \in [0,T]$, with $\mathscr{E}(Y)_0 = 1$. It is given explicitly by

$$\mathscr{E}(Y)_t = \exp\left[Y_t - \frac{1}{2}\langle Y^c \rangle_t\right] \prod_{0 \le s \le t} (1 + \Delta Y_s) e^{-\Delta Y_s}, \tag{2.7}$$

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

Two properties of the Doléans exponential are of particular interest to us. Firstly, it is strictly positive if $\Delta Y > -1$ (this follows from the positivity of the exponential and the fact that the product is always positive). The second property is expressed by the following result.

Proposition 2.53. *If* $M \in \mathcal{M}_{\mathsf{loc}}$*, then* $\mathscr{E}(M)$ *is also a local martingale.*

Proof. See Jacod and Shiryaev [51, Thm. 4.61, p. 59]. ■

Although not as widely used as the stochastic exponential, the stochastic logarithm may be defined as the solution of the "inverse" of equation (2.6).

Definition 2.54. Let $X \in \mathcal{S}$ and suppose that X and X_- do not vanish. Then the *stochastic logarithm* of X, denoted by $\mathscr{L}(X)$, is the unique strong solution of the SDE

$$d\mathscr{L}(X)_t = \frac{1}{X_{t-}}\, dX_t,$$

for all $t \in [0, T]$, with $\mathscr{L}(X)_0 = 0$.

Note that the stochastic logarithm also possesses an explicit representation, akin to (2.7); but we will not require it. However, the next two results are sometimes useful.

Theorem 2.55. *Let $X \in \mathcal{S}$ satisfy the conditions of Definition 2.54, so that $\mathscr{L}(X)$ exists. Then $\mathscr{L}(X)$ is the unique semimartingale Y such that $X = X_0 \mathscr{E}(Y)$ and $Y_0 = 0$.*

Proof. See Jacod and Shiryaev [51, Thm. 8.3, p. 134]. ■

Corollary 2.56. *Let $X \in \mathcal{S}$.*

1. *If X satisfies $\Delta X \neq 1$ identically, then $\mathscr{L}(\mathscr{E}(X)) = X - X_0$.*
2. *If X and X_- do not vanish, then $\mathscr{E}(\mathscr{L}(X)) = X/X_0$.*

2.9 Changes of Measure

Changes of measure feature prominently if mathematical finance, with equivalent martingale measures (see Chapter 3) being particularly important.

Definition 2.57. A probability measure $\mathbb{Q}$ on $(\Omega, \mathscr{F})$ is said to be *absolutely continuous* with respect to $\mathbb{P}$, indicated by $\mathbb{Q} \ll \mathbb{P}$, if

$$\mathbb{P}\{A\} = 0 \quad \Rightarrow \quad \mathbb{Q}\{A\} = 0,$$

for all $A \in \mathscr{F}$. If $\mathbb{P} \ll \mathbb{Q}$ and $\mathbb{Q} \ll \mathbb{P}$, then the measures are said to be *equivalent*, indicated by $\mathbb{P} \sim \mathbb{Q}$.

One way of constructing a probability measure on $(\Omega, \mathscr{F})$, absolutely continuous with respect to $\mathbb{P}$, is to obtain a random variable $\zeta \geq 0$, satisfying $\mathbb{E}[\zeta] = 1$. Then $\mathbb{Q} \ll \mathbb{P}$ can be defined as follows:

$$\mathbb{Q}\{A\} := \mathbb{E}\left[\mathbb{1}_A \zeta\right],$$

for all $A \in \mathscr{F}$. If $\zeta > 0$, then $\mathbb{Q} \sim \mathbb{P}$. The following fundamental result asserts that in fact all absolutely continuous probability measures can be obtained in this way[6].

[6] We state the Radon-Nikodým theorem only in the context of probability measures. The result, of course, has a more general measure-theoretic formulation (see e.g. Cohn [11, Thm. 4.2.2, p. 132]).

Theorem 2.58 (Radon-Nikodým)**.** *Let $\mathbb{Q}$ be a probability measure on $(\Omega, \mathscr{F})$. If $\mathbb{Q} \ll \mathbb{P}$, then there exists an a.s. unique random variable $\zeta \geq 0$ satisfying, $\mathbb{E}[\zeta] = 1$, such that*

$$\mathbb{Q}\{A\} = \mathbb{E}[\mathbb{1}_A \zeta],$$

for all $A \in \mathscr{F}$. If $\mathbb{Q} \sim \mathbb{P}$, then $\zeta > 0$.

Proof. See Jacod and Protter [50, Thm. 28.3, p. 246]. ■

Definition 2.59. Let $\mathbb{Q}$ be a probability measure on $(\Omega, \mathscr{F})$, such that $\mathbb{Q} \ll \mathbb{P}$. The random variable ζ in Theorem 2.58 is called the *Radon-Nikodým derivative* of $\mathbb{Q}$ with respect to $\mathbb{P}$ and is often written as $\frac{d\mathbb{Q}}{d\mathbb{P}} := \zeta$.

Suppose that $\mathbb{Q}$ is a probability measure on $(\Omega, \mathscr{F})$ satisfying $\mathbb{Q} \sim \mathbb{P}$. Then $\frac{d\mathbb{Q}}{d\mathbb{P}} > 0$ and so we may perform the following calculation:

$$\begin{aligned}\mathbb{P}\{A\} = \mathbb{E}[\mathbb{1}_A] &= \mathbb{E}\left[\mathbb{1}_A \frac{1}{d\mathbb{Q}/d\mathbb{P}} \frac{d\mathbb{Q}}{d\mathbb{P}}\right] \\ &= \mathbb{E}^{\mathbb{Q}}\left[\mathbb{1}_A \frac{1}{d\mathbb{Q}/d\mathbb{P}}\right],\end{aligned}$$

for all $A \in \mathscr{F}$. By the uniqueness of the Radon-Nikodým derivative, we then have

$$\frac{d\mathbb{P}}{d\mathbb{Q}} = \frac{1}{d\mathbb{Q}/d\mathbb{P}} \quad \text{a.s.}$$

Definition 2.60. Let $\mathbb{Q}$ be a probability measure on $(\Omega, \mathscr{F})$ satisfying $\mathbb{Q} \ll \mathbb{P}$. The *density process* of $\mathbb{Q}$ with respect to $\mathbb{P}$ is the uniformly integrable martingale Z, defined by

$$Z_t := \mathbb{E}\left[\frac{d\mathbb{Q}}{d\mathbb{P}} \,\middle|\, \mathscr{F}_t\right],$$

for all $t \in \mathbb{R}_+$.

Theorem 2.61 (Bayes' Rule)**.** *Let $\mathbb{Q}$ be a probability measure on $(\Omega, \mathscr{F})$ satisfying $\mathbb{Q} \sim \mathbb{P}$. Denote the density process of $\mathbb{Q}$ with respect to $\mathbb{P}$ by Z and suppose that $s \leq t \in \mathbb{R}_+$. Then*

$$\frac{1}{Z_s}\mathbb{E}[Z_t Y \mid \mathscr{F}_s] = \mathbb{E}^{\mathbb{Q}}[Y \mid \mathscr{F}_s],$$

for all $Y \in L^1(\Omega, \mathscr{F}_t, \mathbb{Q})$.

Proof. Let $A \in \mathscr{F}_s$. Then

$$
\begin{aligned}
\mathbb{E}^{\mathbb{Q}}\left[\mathbb{1}_A \frac{1}{Z_s}\mathbb{E}\left[Z_t Y \mid \mathscr{F}_s\right]\right] &= \mathbb{E}\left[\mathbb{1}_A \frac{d\mathbb{Q}}{d\mathbb{P}} \frac{1}{Z_s}\mathbb{E}\left[Z_t Y \mid \mathscr{F}_s\right]\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[\mathbb{1}_A \frac{d\mathbb{Q}}{d\mathbb{P}} \frac{1}{Z_s}\mathbb{E}\left[Z_t Y \mid \mathscr{F}_s\right] \,\middle|\, \mathscr{F}_s\right]\right] \\
&= \mathbb{E}\left[\mathbb{1}_A \mathbb{E}\left[\frac{d\mathbb{Q}}{d\mathbb{P}} \,\middle|\, \mathscr{F}_s\right] \frac{1}{Z_s}\mathbb{E}\left[Z_t Y \mid \mathscr{F}_s\right]\right] \\
&= \mathbb{E}\left[\mathbb{1}_A \mathbb{E}\left[Z_t Y \mid \mathscr{F}_s\right]\right] = \mathbb{E}\left[\mathbb{E}\left[\mathbb{1}_A Z_t Y \mid \mathscr{F}_s\right]\right] \\
&= \mathbb{E}\left[\mathbb{1}_A Z_t Y\right] = \mathbb{E}\left[\mathbb{1}_A \mathbb{E}\left[\frac{d\mathbb{Q}}{d\mathbb{P}} \,\middle|\, \mathscr{F}_t\right] Y\right] \\
&= \mathbb{E}\left[\mathbb{E}\left[\mathbb{1}_A \frac{d\mathbb{Q}}{d\mathbb{P}} Y \,\middle|\, \mathscr{F}_t\right]\right] = \mathbb{E}\left[\mathbb{1}_A \frac{d\mathbb{Q}}{d\mathbb{P}} Y\right] \\
&= \mathbb{E}^{\mathbb{Q}}\left[\mathbb{1}_A Y\right].
\end{aligned}
$$

The result now follows from the $\mathbb{Q}$-a.s. uniqueness of $\mathbb{E}^{\mathbb{Q}}\left[Y \mid \mathscr{F}_s\right]$. ■

We shall most often make use of Theorem 2.61 in the following forms.

Corollary 2.62. *Let $\mathbb{Q}$ be a probability measure on $(\Omega, \mathscr{F})$ satisfying $\mathbb{Q} \sim \mathbb{P}$ and suppose that Z is its density process. Then*

1. *X is a $\mathbb{Q}$-martingale iff XZ is a martingale; and*
2. *X is a local $\mathbb{Q}$-martingale iff XZ is a local martingale.*

Proof. This follows directly from Theorem 2.61. See also Protter [74, Exercise 20, p. 149]. ■

2.10 Martingale Representation and the GKW Decomposition

Recall that it was established in Section 2.4 that $\mathcal{M}^2$ has a Hilbert space structure. The implications of this will play a central role throughout this dissertation. We examine some of them here.

Definition 2.63. Let $M, N \in \mathcal{M}^2$. Then M and N are said to be *weakly orthogonal*, if $\mathbb{E}\left[M_\infty N_\infty\right] = M_0 N_0$. If MN is a uniformly integrable[7] martingale, then M and N are said to be *strongly orthogonal*.

[7] Actually, it is enough if $MN \in \mathcal{M}_{\mathsf{loc}}$. With the aid of Doob's maximal quadratic inequality and the Cauchy–Schwarz inequality, one can then show that $MN \in \mathcal{M}$.

It can be shown that strong orthogonality implies weak orthogonality, but not vice versa (see Protter [74, p. 179–180] for details). Another characterization of strong orthogonality is given by the following theorem.

Theorem 2.64. *Let $M, N \in \mathcal{M}^2$. Then M and N are strongly orthogonal iff $[M,N] \in \mathcal{M}_0$. Equivalently, they are strongly orthogonal iff $\langle M, N\rangle = 0$ a.s.*

Proof. The first statement follows directly from the stochastic integration by parts rule (2.5). For the second statement, see Jacod and Shiryaev [51, Prop. 4.14, p. 41]. ■

We now focus on martingale representation. It is impossible to give a complete exposition of this topic and so we only provide the terminology and results necessary for the rest of the dissertation. The central result in this subject is the celebrated Jacod-Yor theorem [52]. The reader is referred to Protter [74, §IV.3–5, p. 178–205] for the complete story and to an informative summary by Davis [18].

Definition 2.65. A set of probability measures $\mathbf{P}$ is *convex* if $\lambda\mathbb{Q}_1 + (1-\lambda)\mathbb{Q}_2 \in \mathbf{P}$, for all $\mathbb{Q}_1, \mathbb{Q}_2 \in \mathbf{P}$ and all $0 \le \lambda \le 1$.

Definition 2.66. Let $\mathbf{P}$ be a convex set of probability measures and $0 \le \lambda \le 1$. A probability measure $\mathbb{Q} \in \mathbf{P}$ is said to be an *extremal point* of $\mathbf{P}$ if whenever $\mathbb{Q} = \lambda\mathbb{Q}_1 + (1-\lambda)\mathbb{Q}_2$, with $\mathbb{Q}_1, \mathbb{Q}_2 \in \mathbf{P}$ and $\mathbb{Q}_1 \neq \mathbb{Q}_2$, then $\lambda = 0$ or $\lambda = 1$.

Definition 2.67. Let $\mathcal{A} \subseteq \mathcal{M}_0^2$. The set of probability measures $\mathbb{Q}$ on $(\Omega, \mathscr{F})$ satisfying

1. $\mathbb{Q} \ll \mathbb{P}$;
2. $\mathbb{Q} = \mathbb{P}$ on $\mathscr{F}_0$; and
3. $\mathcal{A} \subseteq \mathcal{M}_0^2(\mathbb{Q})$

is denoted by $\mathbf{P}(\mathcal{A})$.

Note that $\mathbf{P}(\mathcal{A})$ is a convex set (see Protter [74, p. 182] for details).

Definition 2.68. A closed[8] subspace $\mathcal{A}$ of $\mathcal{M}_0^2$ is called a *stable subspace* if it is stable under stopping (i.e. if $M \in \mathcal{A}$ and τ is a stopping time, then $M^\tau \in \mathcal{A}$).

Definition 2.69. Let $\mathcal{A} \subseteq \mathcal{M}_0^2$. The stable subspace of $\mathcal{M}_0^2$ generated by $\mathcal{A}$, denoted by $\mathcal{S}(\mathcal{A})$, is the intersection of all closed stable subspaces of $\mathcal{M}_0^2$ containing $\mathcal{A}$.

[8] Closedness here is obviously defined with respect to the metric induced by $\|\cdot\|_{H^2}$.

Theorem 2.70. *Let $n \in \mathbb{N}$ and suppose that $M^1, \ldots, M^n \in \mathcal{M}_0^2$ are mutually strongly orthogonal. Then*

$$\mathcal{S}(M^1, \ldots, M^n) = \left\{ \sum_{i=1}^{n} \int_0^{\cdot} \phi_s^i \, dM_s^i \;\middle|\; \phi^i \in \mathcal{L}^2(M^i), \text{ for each } i = 1, \ldots, n \right\}.$$

Proof. See Protter [74, Thm. 36, p. 180]. ■

Definition 2.71. Let $n \in \mathbb{N}$ and $M^1, \ldots, M^n \in \mathcal{M}_0^2$ and define

$$\mathcal{I}(M^1, \ldots, M^n) := \left\{ \sum_{i=1}^{n} \int_0^{\cdot} \phi_s^i \, dM_s^i \;\middle|\; \phi^i \in \mathcal{L}^2(M^i), \text{ for each } i = 1, \ldots, n \right\}.$$

If $\mathcal{I}(M^1, \ldots, M^n) = \mathcal{M}_0^2$, then $\{M^1, \ldots, M^n\}$ is said to possess the *predictable representation property*.

We now state three crucial theorems which encapsulate the martingale representation theory required for our purposes.

Theorem 2.72. *Let $\mathcal{A} \subseteq \mathcal{M}_0^2$. If $\mathcal{S}(\mathcal{A}) = \mathcal{M}_0^2$, then $\mathbb{P}$ is an extremal point of $\mathbf{P}(\mathcal{A})$.*

Proof. See Protter [74, Thm. 38, p. 183]. ■

Theorem 2.73. *Let $\mathcal{A} \subseteq \mathcal{M}_0^2$. If $\mathbb{P}$ is an extremal point of $\mathbf{P}(\mathcal{A})$, then the only bounded elements of $\mathcal{M}_0^2$ strongly orthogonal to $\mathcal{A}$ are null.*

Proof. See Protter [74, Thm. 39, p. 183]. ■

Theorem 2.74. *Let $n \in \mathbb{N}$ and suppose that $M^1, \ldots, M^n \in \mathcal{M}_0^2$ are continuous and mutually strongly orthogonal. If $\mathbb{P}$ is an extremal point of $\mathbf{P}(M^1, \ldots, M^n)$, then*

1. *every stopping time is accessible;*
2. *every bounded martingale is continuous;*
3. *every uniformly integrable martingale is continuous; and*
4. *$\{M^1, \ldots, M^n\}$ possesses the predictable representation property.*

Proof. See Protter [74, Thm. 40, p. 184]. ■

Finally, we state a result, originally due to Kunita and Watanabe [59] and Galtchouk [35], that will be central to the our treatment of incomplete markets. It shows that an L^2 random variable can be uniquely decomposed into a stochastic integral with respect to a given square integrable martingale and another square integrable martingale orthogonal to the stable subspace generated by the first.

Proposition 2.75 (GKW Decomposition)**.** *Let $M \in \mathcal{M}_0^2$. Every random variable $H \in L^2$ has a unique representation*

$$H = \mathbb{E}\,[H \,|\, \mathscr{F}_0] + \int_0^T \mu_s^H \, dM_s + N^H \quad a.s.,$$

where $\mu^H \in \mathcal{L}^2(M)$ and $N^H \in L^2$ is orthogonal to the space

$$\left\{ \int_0^T \phi_s \, dM_s \,\middle|\, \phi \in \mathcal{L}^2(M) \right\}.$$

Furthermore, $N \in \mathcal{M}_0^2$, defined by $N_t := \mathbb{E}[N^H \,|\mathscr{F}_t]$, for all $t \in \mathbb{R}_+$, is strongly orthogonal to $\mathcal{I}(M)$.

Proof. See Protter [74, Cor. 1, p. 181] (note that this is stated somewhat differently from our statement of the result). ■

2.11 A Result Concerning the Doléans Measure

The following result, though an application of elementary measure theory, is used a number of times in Chapter 4 and Chapter 5. This, together with the fact that it is repeatedly used in the literature without comment, motivates us to present a proof.

Lemma 2.76. *Let $M \in \mathcal{M}^2$ and suppose that $\phi \in \mathcal{L}^2(M)$ satisfies*

$$\int_0^t \phi_s \, d\langle M \rangle_s = 0 \quad a.s., \tag{2.8}$$

for all $t \in [0, T]$. Then $\phi = 0$ μ_M-a.e.

Proof. Firstly, note that

$$\mathbb{E}\left[\int_0^\infty \phi_s^2 \, d\langle M \rangle_s \right] < \infty,$$

by the definition of $\mathcal{L}^2(M)$. Hence

$$\int_0^\infty \phi_s^2 \, d\langle M \rangle < \infty \quad \text{a.s.}$$

In other words,

$$\phi_\cdot(\omega) \in L^2(\mathbb{R}_+, \mathscr{B}(\mathbb{R}_+), \langle M \rangle_\cdot(\omega)) \subseteq L^1(\mathbb{R}_+, \mathscr{B}(\mathbb{R}_+), \langle M \rangle_\cdot(\omega)), \tag{2.9}$$

for a.a. $\omega \in \Omega$. (We let context determine whether $\langle M \rangle_\cdot(\omega)$ refers to a sample path of the angle bracket process of M, or to the Lebesgue-Stieltjes measure on $(\mathbb{R}_+, \mathscr{B}(\mathbb{R}_+))$ induced by that sample path, for a.a. $\omega \in \Omega$.)

Next, fix any $\omega \in \Omega$ satisfying (2.8) and define the measures ν_ω^+ and ν_ω^- on $(\mathbb{R}_+, \mathscr{B}(\mathbb{R}_+))$ as follows:

$$\nu_\omega^+(A) := \int_A \phi_s^+(\omega)\, d\langle M\rangle_s(\omega) \quad \text{and} \quad \nu_\omega^-(A) := \int_A \phi_s^-(\omega)\, d\langle M\rangle_s(\omega),$$

for all $A \in \mathscr{B}(\mathbb{R}_+)$. Then

$$\begin{aligned}
\nu_\omega^+(s,t] - \nu_\omega^-(s,t] &= \int_s^t \phi_u^+(\omega)\, d\langle M\rangle_u(\omega) - \int_s^t \phi_u^-(\omega)\, d\langle M\rangle_u(\omega) \\
&= \int_s^t \phi_u(\omega)\, d\langle M\rangle_u(\omega) \\
&= \int_0^t \phi_u(\omega)\, d\langle M\rangle_u(\omega) - \int_0^s \phi_u(\omega)\, d\langle M\rangle_u(\omega) \\
&= 0,
\end{aligned}$$

for all $s \leq t \in \mathbb{R}_+$, according to (2.8). In other words, ν_ω^+ and ν_ω^- agree on the π-system (see Rogers and Williams [78, p. 87])

$$\mathscr{I} := \{\, (s,t] \subseteq \mathbb{R}_+ \mid s \leq t \in \mathbb{R}_+ \,\}.$$

Also, it follows directly from (2.9) that

$$\nu_\omega^+(\mathbb{R}_+) = \int_0^\infty \phi_s^+(\omega)\, d\langle M\rangle_s(\omega) < \infty \quad \text{and} \quad \nu_\omega^-(\mathbb{R}_+) = \int_0^\infty \phi_s^-(\omega)\, d\langle M\rangle_s(\omega) < \infty.$$

The above observations, together with the fact that $\mathscr{B}(\mathbb{R}_+) = \sigma(\mathscr{I})$, allow us to deduce from Rogers and Williams [78, Lem. 4.6, p. 93] that $\nu_\omega^+ = \nu_\omega^-$. In particular, this means that

$$\begin{aligned}
\int_A \phi_s(\omega)\, d\langle M\rangle_s(\omega) &= \int_A \phi_s^+(\omega)\, d\langle M\rangle_s(\omega) - \int_A \phi_s^-(\omega)\, d\langle M\rangle_s(\omega) \\
&= \nu_\omega^+(A) - \nu_\omega^-(A) \\
&= 0,
\end{aligned} \tag{2.10}$$

for all $A \in \mathscr{B}(\mathbb{R}_+)$.

Finally, choose $\varepsilon > 0$. Since $\mathscr{P} \subseteq \mathscr{F} \otimes \mathscr{B}(\mathbb{R}_+)$ and ϕ is predictable, it follows that ϕ is $\mathscr{F} \otimes \mathscr{B}(\mathbb{R}_+)$-measurable, whence its ω-sections $\phi_\cdot(\omega)$ are $\mathscr{B}(\mathbb{R}_+)$-measurable, for a.a. $\omega \in \Omega$. (The latter is a standard result in any presentation of Fubini's theorem — see e.g. Cohn [11, Lem. 5.1.1, p. 155].) Consequently,

$$\{\phi > \varepsilon\} := \{\, (\omega,t) \in \Omega \times \mathbb{R}_+ \mid \phi_t(\omega) > \varepsilon \} \in \mathscr{F} \otimes \mathscr{B}(\mathbb{R}_+)$$

and

$$\{\phi_\cdot(\omega) > \varepsilon\} := \{\, t \in \mathbb{R}_+ \mid \phi_t(\omega) > \varepsilon \} \in \mathscr{B}(\mathbb{R}_+),$$

for a.a. $\omega \in \Omega$. Thus, by Fubini's theorem (we only need the version for σ-finite measure spaces — see e.g. Cohn [11, Prop. 5.2.1, p. 159]),

$$\begin{aligned}
\varepsilon(\mathbb{P} \otimes \langle M \rangle)\{\phi > \varepsilon\} &= \varepsilon \int_{\Omega \times \mathbb{R}_+} \mathbb{1}_{\{\phi > \varepsilon\}}(\omega, t)\, d(\mathbb{P} \otimes \langle M \rangle)(\omega, t) \\
&= \varepsilon \int_\Omega \int_{\mathbb{R}_+} \mathbb{1}_{\{\phi_\cdot(\omega) > \varepsilon\}}(t)\, d\langle M \rangle_t(\omega)\, d\mathbb{P}(\omega) \\
&\leq \int_\Omega \int_{\mathbb{R}_+} \mathbb{1}_{\{\phi_\cdot(\omega) > \varepsilon\}}(t) \phi_t(\omega)\, d\langle M \rangle_t(\omega)\, d\mathbb{P}(\omega) \\
&= 0,
\end{aligned}$$

since (2.10) has already established that the inner integral is a.s. zero. It thus follows that

$$\begin{aligned}
(\mathbb{P} \otimes \langle M \rangle)\{\phi > 0\} &= (\mathbb{P} \otimes \langle M \rangle)\left(\bigcup_{n=1}^{\infty} \left\{ \phi > \frac{1}{n} \right\} \right) \\
&\leq \sum_{n=1}^{\infty} (\mathbb{P} \otimes \langle M \rangle) \left\{ \phi > \frac{1}{n} \right\} \\
&= 0.
\end{aligned}$$

A similar argument works for the set $\{\phi < 0\}$; hence $(\mathbb{P} \otimes \langle M \rangle)\{\phi \neq 0\} = 0$. Since $\mu_M = \mathbb{P} \otimes \langle M \rangle|_{\mathscr{P}}$, the result follows. ■

Chapter 3

Complete Markets

In the introduction we indicated how the concept of stochastic integration is appropriate for modelling the behavior of a portfolio of traded assets. In this chapter these ideas will be developed further, with the presentation of a general semimartingale model for a simple market. We describe a market formulation with two assets for investment — one of these is a traded security, while the other is a bank account. In order to ease the mathematical exposition, the bank account is identified as the numéraire asset and the traded security is expressed as a numéraire-denominated price process.

Investment strategies and portfolio choices are discussed next. For theoretical and practical reasons, we shall limit our attention in this chapter to investment strategies that are self-financing and satisfy an admissibility condition. These concepts are introduced and discussed.

It is probably no exaggeration to say that the most fundamental advances in mathematical finance over the past twenty years have involved the careful analysis of the relationships between the existence of market equilibria, the non-existence of arbitrage opportunities and the existence of equivalent martingale measures. These developments were initiated by Harrison and Kreps [39] and Harrison and Pliska [40, 41] and culminated with the formulation and proof of the fundamental theorems of asset pricing.

The *first fundamental theorem of asset pricing* establishes a relationship between no-arbitrage and the existence of an equivalent measure under which all numéraire-denominated self-financing portfolio processes are local martingales. It has been established in its most general form by Delbaen and Schachermayer [20, 22]. We shall provide a formal and precise statement of this result.

The concept of market completeness expresses the condition that all contingent claims can be replicated with self-financing investment strategies. The *second fundamental theorem of asset pricing* provides a relationship between market completeness and the uniqueness of the equivalent martingale measure for a given numéraire. We shall present a proof of a version this theorem, which is essentially a martingale representation result due to Jacod and Yor [52].

With the fundamental theorems established, we shall concentrate on an approach to constructing equivalent local martingale measures in the general semimartingale framework. This will lead to the formulation of the structure condition, which will be used throughout the rest of the dissertation. The chapter concludes with the familiar example of the standard Black-Scholes model, which serves to illustrate the approach outlined above.

3.1 The Market Model

Throughout this chapter we fix a filtered probability space $(\Omega, \mathscr{F}, \mathbf{F}, \mathbb{P})$, satisfying the usual conditions. The filtration $\mathbf{F} = (\mathscr{F}_t)_{t\in[0,T]}$ is defined over a finite time-horizon $T \in (0, \infty)$ and we make the simplifying assumptions that $\mathscr{F}_0$ is $\mathbb{P}$-trivial (i.e. it only contains $\mathbb{P}$-null sets and their compliments) and $\mathscr{F}_T = \mathscr{F}$. All stochastic processes are defined on $(\Omega, \mathscr{F}, \mathbf{F}, \mathbb{P})$ (in particular, they are defined on the time interval $[0, T]$) and are implicitly understood to be adapted to $\mathbf{F}$.

Two processes specify the assets at an investor's disposal. In order to price instruments in the market, it is necessary to select one of them as a common standard of value; this asset is called the *numéraire*. The natural and usual choice for the numéraire is the *bank account*. In keeping with convention this is an adapted finite variation process B, defined by

$$B_t := \exp\left(\int_0^t r_s\, ds\right),$$

for all $t \in [0, T]$, where r is the non-negative adapted instantaneous *short rate* process. The other process S is called the *stock price* process. As opposed to the case for the bank account, we shall leave the specific dynamics of the stock price unspecified and assume only that it is a positive semimartingale.

It is convenient to work with price processes denominated in units of the numéraire. In the case where the bank account is selected as numéraire, we refer to numéraire-denominated prices as discounted prices.

Definition 3.1. The process X, defined by

$$X_t := \frac{S_t}{B_t},$$

for all $t \in [0, T]$, describes the discounted stock price process. We shall refer to it simply as the *price process*.

Given the assets at our disposal, it is now possible to consider trading strategies and portfolios.

Definition 3.2. A *trading strategy* is a pair of processes (ξ, η), where $\xi \in \mathcal{L}(X)$ and η is adapted. At time $t \in [0, T]$, the component ξ_t represents the *stock-holding* or number of units of stock held in a portfolio, while η_t is the *bank account holding* or number of units of the bank account held in the portfolio.

Definition 3.3. Let (ξ, η) be a trading strategy. The process $V(\xi, \eta)$, defined by

$$V_t(\xi, \eta) := \frac{1}{B_t}(\xi_t S_t + \eta_t B_t) = \xi_t X_t + \eta_t, \tag{3.1}$$

for all $t \in [0, T]$, describes the discounted value of the strategy. We shall refer to it simply as the *value process* (of (ξ, η)).

In the introduction we justified the concept of stochastic integration with a discrete-time trading example. It is now possible to do the same for continuous-time trading. We start by defining the process that describes the gain from a trading strategy.

Definition 3.4. Let (ξ, η) be a trading strategy. The process $G(\xi)$, defined by

$$G_t(\xi) := \int_0^t \xi_s \, dX_s, \tag{3.2}$$

for all $t \in [0, T]$, describes the discounted gain from trade associated with the strategy. We shall refer to it simply as the *gain process* (of (ξ, η)).

Often a value or gain process associated with a trading strategy is the main object of interest, not the strategy itself. In such cases we shall refer to a value process V or a gain process G, without explicitly indicating the underlying strategy.

An important class of portfolio strategies are those for which, after the initial endowment, there is no further investment or withdrawal of funds. The example in the introduction employed a self-financing constraint which required that, at the instant trading took place, the portfolio weights were adjusted to ensure that no net capital gain or loss was incurred. An analogous condition may be specified for the continuous-time case, leading to the definition of a self-financing trading strategy as one for which changes in value are due only to gains from trading in the market and are not the result of any inflow or outflow of funds.

Definition 3.5. A trading strategy (ξ, η) is said to be *self-financing* if the value process $V(\xi, \eta)$ satisfies the condition

$$V_t(\xi, \eta) = V_0(\xi, \eta) + G_t(\xi), \tag{3.3}$$

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

It should be noted that there is a redundancy inherent in the concept of a self-financing strategy. If the initial capital v of such a strategy (ξ, η) is specified, then the amount invested in the riskless asset is determined by

$$\eta_t = v + G_t(\xi) - \xi_t X_t, \tag{3.4}$$

for all $t \in [0, T]$. In other words, the strategy is specified either by the pair (ξ, η), in which case $v = V_0(\xi, \eta)$ can be calculated from (3.1); or by the pair (ξ, v), in which case η is determined by (3.4).

The next result establishes that the self-financing property is preserved irrespective of whether the portfolio is denominated in terms of the numéraire. This explains why our decision to work with discounted asset and portfolio values results in no loss of generality.

Theorem 3.6 (Numéraire Invariance). *The self-financing property is not affected by discounting.*

Proof. Let (ξ, η) be a self-financing trading strategy and V its associated value process. Since B is continuous and of finite variation, the stochastic integration by parts rule gives

$$\begin{aligned} d(BV)_t &= B_t\, dV_t + V_{t-}\, dB_t \\ &= B_t \xi_t\, dX_t + V_{t-}\, dB_t \\ &= \xi_t\, dS_t - \xi_t X_{t-}\, dB_t + V_{t-}\, dB_t \\ &= \xi_t\, dS_t + (V_{t-} - \xi_t X_{t-})\, dB_t \\ &= \xi_t\, dS_t + (V_t - \xi_t X_t)\, dB_t \\ &= \xi_t\, dS_t + \eta_t\, dB_t, \end{aligned}$$

for all $t \in [0, T]$, where the second equality is an expression of the self-financing condition, the third equality follows from

$$dS_t = d(BX)_t = B_t\, dX_t + X_{t-}\, dB_t,$$

and the penultimate equality follows from the fact that $\Delta V_t = \xi_t \Delta X_t$, since (ξ, η) is self-financing. Thus the non-discounted portfolio value may be written as

$$B_t V_t = V_0 + \int_0^t \xi_u\, dS_u + \int_0^t \eta_u\, dB_u, \tag{3.5}$$

for all $t \in [0, T]$. This is precisely the self-financing condition for the non-discounted portfolio[1]. The reverse implication follows by a similar argument. ■

[1] In other words, the change in value of the non-discounted portfolio is due only to the gain from trading in the market assets (bank account and stock) — compare (3.5) with (3.3) and (3.2).

More general versions of the numéraire invariance theorem exist for markets comprising many assets. It is also possible to denominate portfolios in terms of an asset with a stochastic component and obtain a similar result (see e.g. [36, 2]).

For practical and theoretical reasons, we wish to eliminate strategies that lead to certain profit, but possibly incur unbounded interim losses. The classic doubling strategy and the suicide strategy of Harrison and Pliska [40] are examples of strategies we wish to censure. The concept of portfolio admissibility is introduced for this reason.

Definition 3.7. A stock-holding process ξ is said to be *admissible* if there exists some $\alpha \in \mathbb{R}_+$ such that $G_t(\xi) \geq -\alpha$, for all $t \in [0,T]$. A trading strategy that utilizes an admissible stock-holding process will be called admissible.

Admissibility is an economically realistic constraint, since it represents a limited line of credit. Note that, in general, admissibility is not preserved under a change of numéraire — it is a requirement that the numéraire be bounded away from zero for this to be the case. The significance of admissibility is highlighted by the following theorem, due to Ansel and Stricker.

Theorem 3.8. *If $M \in \mathcal{M}_{\mathsf{loc}}$ and ϕ is a predictable process such that the integral $\int_0^{\cdot} \phi_s \, dM_s$ is bounded below, then $\int_0^{\cdot} \phi_s \, dM_s$ is a local martingale.*

Proof. See Ansel and Stricker [1, Cor. 3.5, p. 309]. ■

We now provide a simple lemma that characterizes the integral in the previous theorem as a supermartingale.

Lemma 3.9. *Let $M \in \mathcal{M}_{\mathsf{loc}}$ possess a localizing sequence of stopping times $(\tau_n)_{n\in\mathbb{N}}$ such that*

$$\lim_{n\to\infty} \mathbb{P}\{\tau_n = T\} = 1. \tag{3.6}$$

If M is bounded below, then it is a supermartingale.

Proof. Let $s \leq t \in [0,T]$. Then

$$\begin{aligned}
\mathbb{E}\left[M_t \,|\, \mathscr{F}_s\right] &= \mathbb{E}\left[\lim_{n\to\infty} M_t^{\tau_n} \,\middle|\, \mathscr{F}_s\right] \\
&= \mathbb{E}\left[\liminf_{n\to\infty} M_t^{\tau_n} \,\middle|\, \mathscr{F}_s\right] \\
&\leq \liminf_{n\to\infty} \mathbb{E}\left[M_t^{\tau_n} \,|\, \mathscr{F}_s\right] \\
&= \lim_{n\to\infty} M_s^{\tau_n} = M_s.
\end{aligned}$$

The inequality follows from Fatou's lemma, while the final equality is a consequence of assumption (3.6) on the localizing sequence. ■

The proof of this lemma highlights the fact that when working with a local martingale on a finite time-horizon one must rule out possible end point pathologies. This is the point of (3.6), which can also be found in Harrison and Pliska [40, p. 233]. To see how the result can fail if we only assume that M is a local martingale, consider the following example.

Example 3.10. Define M by setting $M_t := \mathbb{1}_{\{t=T\}}$, for all $t \in [0, T]$. Then M is a (deterministic) local martingale — the sequence of (deterministic) stopping times $(\tau_n)_{n\in\mathbb{N}}$, with $\tau_n := T - 1/n$, is a localizing sequence. However, M is a strict submartingale, since $\mathbb{E}[M_T] = 1 > 0 = \mathbb{E}[M_t]$, for all $t \in [0, T)$.

Note that for any increasing sequence of stopping times $(\tau_n)_{n\in\mathbb{N}}$ taking values in $[0, T]$,

$$\lim_{n\to\infty} \mathbb{P}\{\tau_n = T\} \leq \mathbb{P}\left\{\lim_{n\to\infty} \tau_n = T\right\}. \tag{3.7}$$

This implies that if the sequence satisfies (3.6), then $\tau_n \uparrow T$ a.s. In other words, (3.6) represents a stronger condition on the localizing sequence of stopping times for a local martingale than (the conventional) a.s. convergence to T.

We shall not worry about these technicalities anymore. Henceforth we shall take Lemma 3.9 to say that any bounded below local martingale is a supermartingale (this is the folk theorem as it repeatedly appears in the literature). It is now possible to relate the above results to the market, by providing a corollary to Theorem 3.8.

Corollary 3.11. *If $X \in \mathcal{M}_{\mathsf{loc}}$ and (ξ, η) is an admissible self-financing strategy, then $V(\xi, \eta)$ is a local martingale and consequently also a supermartingale.*

Often the admissibility criterion of Definition 3.7 is difficult to incorporate in a mathematical formalism — this will certainly be the case when we start to consider market incompleteness. An alternative approach to eliminating pathological strategies is provided by the following definition (see e.g. [69]).

Definition 3.12. Let $\mathbb{Q}$ be a probability measure equivalent to $\mathbb{P}$. A process ξ, representing a stock-holding, is said to be $\mathbb{Q}$-*admissible* if $G(\xi)$ is a martingale under $\mathbb{Q}$. A trading strategy that utilizes a $\mathbb{Q}$-admissible stock-holding process will also be called $\mathbb{Q}$-admissible.

This definition does not have the same intuitive economic appeal as the notion of admissibility offered by Definition 3.7. However, it will be seen that portfolios that satisfy Definition 3.12 form a suitable setting for the first fundamental theorem of asset pricing. This is the subject of the next section.

3.2 The First Fundamental Theorem of Asset Pricing

The notion of arbitrage is a key concept and is intrinsically linked to market equilibrium. Following the description of Kreps [57], in a market in equilibrium it should not be possible for an agent to purchase, at zero cost, a portfolio that will strictly increase his utility. Such a portfolio would constitute an arbitrage opportunity, since, in the absence of credit constraints, he would continue to purchase it until either the price of the portfolio increased or it ceased to increase his utility. The absence of arbitrage is thus a necessary condition for economic equilibrium.

We shall not explore the idea of market equilibrium further, but instead turn our attention to another feature of arbitrage. The absence of arbitrage and the existence of an equivalent probability measure under which discounted asset prices are local martingales — a so-called risk-neutral measure — are also linked. These two concepts will now be made more concrete and their relationship explored.

Definition 3.13. Let V be the value process associated with a self-financing trading strategy. If there exist two stopping times σ and τ, with $\sigma < \tau \in [0,T]$ a.s., for which

$$\mathbb{P}\{V_\tau \geq V_\sigma\} = 1 \quad \text{and} \quad \mathbb{P}\{V_\tau > V_\sigma\} > 0,$$

then V is an *arbitrage*.

In other words, an arbitrage is a portfolio that with certainty grows at least as fast as the riskless asset over some time interval $[\sigma, \tau]$ and has a positive probability of delivering a return in excess of the risk-free rate over that interval. Clearly, the existence of an arbitrage allows an investor to make an arbitrarily large profit without assuming any risk. This is achieved by borrowing as much capital as possible, corresponding to a negative bank account holding, and investing it in the arbitrage portfolio over the time interval. At the end of the arbitrage period the portfolio is sold to cover the outstanding bank account liability, returning an excess amount with non-zero probability.

Definition 3.14. A probability measure $\mathbb{Q}$ on $(\Omega, \mathscr{F})$ is called a *risk-neutral measure* if $\mathbb{Q} \sim \mathbb{P}$ and the price process X is a local martingale under $\mathbb{Q}$. We shall also call such a measure an equivalent local martingale measure. If X is in fact a martingale under $\mathbb{Q}$, then we call $\mathbb{Q}$ an equivalent martingale measure.

With the definition of arbitrage and risk-neutral measures in place, it is now possible to present a result which is central to mathematical finance. The first fundamental theorem of asset pricing links the concept of no-arbitrage to the existence of risk-neutral measures. We start by stating an informal version.

Folk Theorem. *A risk-neutral measure exists for the market of Section 3.1 iff no admissible self-financing strategy yields an arbitrage.*

At first the statement of the theorem seems surprising, but it turns out that necessity is quite easy to prove, as is seen below. The reverse implication is however remarkably difficult and requires a slightly modified definition of arbitrage, as well as some sophisticated functional analysis.

Proof of necessity. Assume the existence of a risk-neutral measure $\mathbb{Q}$ and let V be the value process of an admissible self-financing trading strategy. By Corollary 3.11, V is then a supermartingale under $\mathbb{Q}$. Now suppose that V is an arbitrage. By assumption we can find stopping times $\sigma < \tau \in [0, T]$, such that

$$\mathbb{P}\{V_\tau \geq V_\sigma\} = 1 \quad \text{and} \quad \mathbb{P}\{V_\tau > V_\sigma\} > 0.$$

Then $\mathbb{Q} \sim \mathbb{P}$ implies that

$$\mathbb{Q}\{V_\tau \geq V_\sigma\} = 1 \quad \text{and} \quad \mathbb{Q}\{V_\tau > V_\sigma\} > 0.$$

The first condition implies that

$$\mathbb{E}^{\mathbb{Q}}\left[V_\tau\right] = \mathbb{E}^{\mathbb{Q}}\left[\mathbb{1}_{\{V_\tau \geq V_\sigma\}} V_\tau\right] \quad \text{and} \quad \mathbb{E}^{\mathbb{Q}}\left[V_\sigma\right] = \mathbb{E}^{\mathbb{Q}}\left[\mathbb{1}_{\{V_\tau \geq V_\sigma\}} V_\sigma\right],$$

while from the second condition we get

$$\mathbb{E}^{\mathbb{Q}}\left[\mathbb{1}_{\{V_\tau > V_\sigma\}} V_\tau\right] > \mathbb{E}^{\mathbb{Q}}\left[\mathbb{1}_{\{V_\tau > V_\sigma\}} V_\sigma\right].$$

Hence,

$$\begin{aligned}
\mathbb{E}^{\mathbb{Q}}\left[V_\tau\right] &= \mathbb{E}^{\mathbb{Q}}\left[\mathbb{1}_{\{V_\tau \geq V_\sigma\}} V_\tau\right] \\
&= \mathbb{E}^{\mathbb{Q}}\left[\mathbb{1}_{\{V_\tau = V_\sigma\}} V_\tau\right] + \mathbb{E}^{\mathbb{Q}}\left[\mathbb{1}_{\{V_\tau > V_\sigma\}} V_\tau\right] \\
&= \mathbb{E}^{\mathbb{Q}}\left[\mathbb{1}_{\{V_\tau = V_\sigma\}} V_\sigma\right] + \mathbb{E}^{\mathbb{Q}}\left[\mathbb{1}_{\{V_\tau > V_\sigma\}} V_\tau\right] \\
&> \mathbb{E}^{\mathbb{Q}}\left[\mathbb{1}_{\{V_\tau = V_\sigma\}} V_\sigma\right] + \mathbb{E}^{\mathbb{Q}}\left[\mathbb{1}_{\{V_\tau > V_\sigma\}} V_\sigma\right] \\
&= \mathbb{E}^{\mathbb{Q}}\left[\mathbb{1}_{\{V_\tau \geq V_\sigma\}} V_\sigma\right] \\
&= \mathbb{E}^{\mathbb{Q}}\left[V_\sigma\right].
\end{aligned}$$

But this contradicts

$$\mathbb{E}^{\mathbb{Q}}\left[V_\tau\right] = \mathbb{E}^{\mathbb{Q}}\left[\mathbb{E}^{\mathbb{Q}}\left[V_\tau \mid \mathscr{F}_\sigma\right]\right] \leq \mathbb{E}^{\mathbb{Q}}\left[V_\sigma\right],$$

which follows from the supermartingale property of V under $\mathbb{Q}$, together with the optional sampling theorem. ■

Note that the Folk Theorem above — or the first fundamental theorem of asset pricing, in general — cannot be formulated using the notion of $\mathbb{Q}$-admissibility. This is due to the fact that one already needs to specify a candidate EMM in order to define $\mathbb{Q}$-admissibility and, in the event that there are numerous candidates, it is not clear which one should be used. It is, however, possible to show that no $\mathbb{Q}$-admissible strategy is an arbitrage. To see this we can follow the proof of necessity above with the value process of an admissible strategy replaced by the value process of a $\mathbb{Q}$-admissible strategy (ξ, η). Now, $G(\xi)$ is a martingale under $\mathbb{Q}$ and the self-financing property (Definition 3.5) means that $V(\xi, \eta)$ is also a martingale under $\mathbb{Q}$ (and hence also a supermartingale). Following the rest of the proof as before yields the result.

The history of this theorem is rich, with the first versions presented for the discrete-time case. Inspired by the work of Ross [80] and Cox and Ross [13] (in this paper the market was modified so that stocks appreciate at the riskless rate and claims are calculated as expected values), Harrison and Kreps [39] provided the first proof of the theorem for the case where Ω is finite. Later Dalang, Morton and Willinger [14] proved the result for arbitrary Ω. Finally, it was proved in the continuous-time setting in a series of papers by Delbaen and Schachermayer [23, 20, 22].

In order to provide a precise statement of the first fundamental theorem, a more formal definition of arbitrage is now provided. We start by defining the following sets (see Delbaen and Schachermayer [20, p. 473] and Björk [5, p. 139]):

$$
\begin{aligned}
\mathcal{K}_0 &= \{\, G_T(\xi) \,|\, \xi \text{ is an admissible stock-holding strategy}\} \,, \\
\mathcal{K} &= \mathcal{K}_0 \cap L^\infty, \\
L^\infty_+ &= \{ x \in L^\infty \,|\, x \geq 0 \} \,, \\
\mathcal{C} &= \mathcal{K} - L^\infty_+ = \left\{ y - x \,|\, y \in \mathcal{K},\, x \in L^\infty_+ \right\} .
\end{aligned}
$$

The set $\mathcal{K}_0$ contains the terminal values of all admissible self-financing strategies starting with zero initial investment. The set $\mathcal{K}$ is the subset of these terminal portfolio values that are also bounded. The elements of $\mathcal{C}$ are those non-negative bounded terminal portfolio values dominated by elements of $\mathcal{K}$. Alternatively, they may be thought of as the terminal values of all admissible self-financing strategies starting at zero that lead to a bounded outcome, but may lose money through an ineffective strategy (for instance by the addition of a suicide strategy).

It should be clear (by Definition 3.13) that any non-negative element of $\mathcal{C}$, with a non-zero probability of a strictly positive value, constitutes an arbitrage. If arbitrage is not allowed, then the only non-negative elements of $\mathcal{C}$ should be a.s. zero. In other words, the only non-negative terminal portfolio value reachable by an admissible self-financing strategy with zero initial investment should be zero. This leads to the following definition.

Definition 3.15. The market of Section 3.1 satisfies the conditions of

1. *no arbitrage* (*NA*) if $\mathcal{C} \cap L^\infty_+ = \{0\}$; and
2. *no free lunch with vanishing risk* (*NFLVR*) if $\overline{\mathcal{C}} \cap L^\infty_+ = \{0\}$.

Note that $\mathcal{C} \subseteq \overline{\mathcal{C}}$ implies that NFLVR is a stronger condition than NA. Earlier we demonstrated that the existence of a risk-neutral measure implies NA. The reverse implication is, however, false — Schachermayer [81, Prop. 4.5] provides a counterexample. This is why we need the strengthened condition of NFLVR. Under various ancillary assumptions, it is possible to show that NFLVR implies the existence of a measure under which the price process is (in full generality) a sigma-martingale.

Definition 3.16. X is called a *sigma-martingale* if there exists a martingale M and a non negative process $\phi \in \mathcal{L}(M)$, such that

$$X = \int_0^\cdot \phi_s \, dM_s.$$

A probability measure $\mathbb{Q} \sim \mathbb{P}$ under which X is a sigma-martingale is called an *equivalent sigma-martingale measure.*

Note that sigma-martingales are more general objects than local martingales. While all local martingales are sigma-martingales, it is possible to find sigma-martingales that are not local martingales — a famous example is due to Emery [30] (see also [74, p. 176]).

We are now able to summarize the results of Delbaen and Schachermayer [20, 22], which constitute the most general statement of the first fundamental theorem of asset pricing (see also [93, VII§2c, p. 655–659]).

Theorem 3.17 (First Fundamental Theorem of Asset Pricing)**.** *Let EMM (resp. ELMM, EσMM) denote the condition that the price process admits an equivalent martingale measure (resp. equivalent local martingale measure, equivalent sigma-martingale measure).*

1. *If X is a general semimartingale, then*

$$EMM \Rightarrow ELMM \Rightarrow E\sigma MM \Leftrightarrow NFLVR \Rightarrow NA.$$

2. *If X is a locally bounded semimartingale, then*

$$EMM \Rightarrow ELMM \Leftrightarrow E\sigma MM \Leftrightarrow NFLVR \Rightarrow NA.$$

3. *If X is a bounded semimartingale, then*

$$EMM \Leftrightarrow ELMM \Leftrightarrow E\sigma MM \Leftrightarrow NFLVR \Rightarrow NA.$$

3.3 The Second Fundamental Theorem of Asset Pricing

As we have just seen, the first fundamental theorem of asset pricing relates the concept of no-arbitrage to the existence of an equivalent martingale measure for X. The second fundamental theorem of asset pricing relates the concept of market completeness to the uniqueness of such a probability measure. The version of this theorem presented here considers the case of equivalent martingale measures under which X is a square integrable martingale. It is not the most general version, but it is appropriate, since Chapter 4 is devoted to the theory of incomplete markets in the case where X is a square integrable martingale under $\mathbb{P}$.

We start by generalizing some of the terminology and concepts presented in Section 2.10, so that they apply to any measure $\mathbb{Q} \sim \mathbb{P}$, rather than to $\mathbb{P}$ exclusively. We denote by $\mathcal{M}^2(\mathbb{Q})$ (resp. $\mathcal{M}_0^2(\mathbb{Q})$) the class of square integrable $\mathbb{Q}$-martingales (resp. the class of square integrable $\mathbb{Q}$-martingales with initial value zero); $\mathcal{S}(\mathbb{Q};\mathcal{A})$ is the stable subspace of $\mathcal{M}_0^2(\mathbb{Q})$ generated by $\mathcal{A} \subseteq \mathcal{M}_0^2(\mathbb{Q})$; and $\mathbf{P}(\mathbb{Q};\mathcal{A})$ is the family of probability measures $\bar{\mathbb{Q}}$ on $(\Omega, \mathscr{F})$ satisfying properties (1)–(3) of Definition 2.67, but with $\mathbb{P}$ replaced by $\mathbb{Q}$ and $\mathbb{Q}$ replaced by $\bar{\mathbb{Q}}$. We now fix some additional terminology and notation.

Definition 3.18. Set $\mathbf{P}_{\mathsf{e}}(X) := \big\{ \mathbb{Q} \sim \mathbb{P} \,|\, X \in \mathcal{M}^2(\mathbb{Q}) \big\}$.

From the martingale property of X under elements of $\mathbf{P}_{\mathsf{e}}(X)$, it is easily seen that this set of probability measures is convex. Also note that $\mathbf{P}_{\mathsf{e}}(X) \subseteq \mathbf{P}(X)$, by Definition 2.67,

As is standard, a contingent claim is represented by an $\mathscr{F}_T (= \mathscr{F})$-measurable random variable, describing the discounted payoff received by the holder at maturity.

Definition 3.19. Let $\mathbb{Q} \in \mathbf{P}_{\mathsf{e}}(X)$. A contingent claim $H \in L^2(\mathbb{Q})$ is said to be $\mathbb{Q}$-*attainable* if there exists a $\mathbb{Q}$-admissible self-financing strategy (ξ, η), with $\xi \in \mathcal{L}^2(X)$, such that $H = V_T(\xi, \eta)$ $\mathbb{Q}$-a.s.

Definition 3.20. The market is called *complete* if there exists a probability measure $\mathbb{Q} \in \mathbf{P}_{\mathsf{e}}(X)$ such that every claim $H \in L^2(\mathbb{Q})$ is $\mathbb{Q}$-attainable.

We now present a version of the second fundamental theorem of asset pricing, subject to the condition that X is continuous. This constraint is only really required for the implication $(2 \Rightarrow 3)$, where it is necessary in order to apply Theorem 2.74. This version is based on Protter [73, p. 189] (see also Davis [18]).

Theorem 3.21 (Second Fundamental Theorem of Asset Pricing)**.** *Suppose X is continuous. Then the following statements are equivalent:*

1. *the market is complete;*

2. $X - X_0$ *possesses the predictable representation property under some probability measure* $\mathbb{Q} \in \mathbf{P}_{\mathsf{e}}(X)$*; and*

3. $|\mathbf{P}_{\mathsf{e}}(X)| = 1$.

Proof. ($1 \Rightarrow 2$) Let $\mathbb{Q} \in \mathbf{P}_{\mathsf{e}}(X)$ be a probability measure such that every claim in $L^2(\mathbb{Q})$ is $\mathbb{Q}$-attainable and fix $M \in \mathcal{M}_0^2(\mathbb{Q})$. By assumption, there exists a $\mathbb{Q}$-admissible self-financing strategy (ξ, η), with $\xi \in \mathcal{L}^2(X)$, such that $V_T(\xi, \eta) = H := M_T$. $\mathbb{Q}$-admissibility implies that $V(\xi, \eta)$ is a martingale under $\mathbb{Q}$ and so we have

$$M_t = \mathbb{E}^{\mathbb{Q}}\left[M_T \mid \mathscr{F}_t\right] = \mathbb{E}^{\mathbb{Q}}\left[V_T(\xi, \eta) \mid \mathscr{F}_t\right] = V_t(\xi, \eta) = G_t(\xi),$$

for all $t \in [0, T]$, where the last step follows from the self-financing condition and the fact that $V_0 = \mathbb{E}^{\mathbb{Q}}\left[M_T\right] = 0$. So the desired representation for M is given by

$$M_t = \int_0^t \xi_s \, dX_s = \int_0^t \xi_s \, d(X - X_0)_s,$$

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

($2 \Rightarrow 1$) Let $\mathbb{Q} \in \mathbf{P}_{\mathsf{e}}(X)$ be a probability measure under which $X - X_0$ enjoys the predictable representation property and fix an arbitrary claim $H \in L^2(\mathbb{Q})$. Define a martingale $M \in \mathcal{M}_0^2(\mathbb{Q})$, by setting

$$M_t := \mathbb{E}^{\mathbb{Q}}\left[H \mid \mathscr{F}_t\right] - \mathbb{E}^{\mathbb{Q}}\left[H\right],$$

for all $t \in [0, T]$. By assumption, there exists a process $\phi \in \mathcal{L}^2(X)$ satisfying

$$M_t = \int_0^t \phi_s \, d(X - X_0)_s = \int_0^t \phi_s \, dX_s,$$

for all $t \in [0, T]$. Now take $\xi := \phi$ and define the process η, by setting

$$\eta_t := \mathbb{E}^{\mathbb{Q}}\left[H\right] + G(\xi)_t - \xi_t X_t,$$

for all $t \in [0, T]$. This makes (ξ, η) a self-financing $\mathbb{Q}$-admissible strategy, with $V_T(\xi, \eta) = H$ a.s. In other words, H is $\mathbb{Q}$-attainable.

($2 \Rightarrow 3$) Let $\mathbb{Q} \in \mathbf{P}_{\mathsf{e}}(X)$ and suppose that $X - X_0$ possesses the predictable representation property under $\mathbb{Q}$. Then by Definition 2.71 and Theorem 2.70 we have $\mathcal{S}(\mathbb{Q}; X - X_0) = \mathcal{M}_0^2(\mathbb{Q})$. Consequently, by Theorem 2.72, $\mathbb{Q}$ is an extremal point of $\mathbf{P}(\mathbb{Q}; X - X_0)$. Now suppose that $\bar{\mathbb{Q}} \in \mathbf{P}_{\mathsf{e}}(X)$ is another equivalent martingale measure for X. Define the $\mathbb{Q}$-martingale L, by setting

$$L_t := \mathbb{E}^{\mathbb{Q}}\left[\frac{d\bar{\mathbb{Q}}}{d\mathbb{Q}} \,\middle|\, \mathscr{F}_t\right],$$

for all $t \in [0,T]$. Since $\mathbb{Q}$ is extremal in $\mathbf{P}(\mathbb{Q}; X - X_0)$ and L is, by construction, a uniformly integrable $\mathbb{Q}$-martingale, it follows from item (3) of Theorem 2.74 that L is continuous. Now fix $K > 1$ and define the stopping time τ_K as follows:

$$\tau_K := \inf\{t \in [0,T] \,|\, L_t > K\}.$$

Let $s \leq t \in [0,T]$ and choose $A \in \mathscr{F}_s$. Then on the event $\{\tau_K \leq s\}$, we have

$$L_t^{\tau_K} X_t^{\tau_K} = K X_{\tau_K} = L_s^{\tau_K} X_s^{\tau_K},$$

which means that

$$\mathbb{E}^{\mathbb{Q}}\left[\mathbb{1}_{\{\tau_K \leq s\}} \mathbb{1}_A L_t^{\tau_K} X_t^{\tau_K}\right] = \mathbb{E}^{\mathbb{Q}}\left[\mathbb{1}_{\{\tau_K \leq s\}} \mathbb{1}_A L_s^{\tau_K} X_s^{\tau_K}\right]. \tag{3.8}$$

Next, we observe[2] that $A \cap \{\tau_K > s\} \in \mathscr{F}_{\tau_K}$. It then follows that

$$A \cap \{\tau_K > s\} \in \mathscr{F}_s \cap \mathscr{F}_{\tau_K} = \mathscr{F}_{\tau_K \wedge s} \subseteq \mathscr{F}_{\tau_K \wedge t}. \tag{3.9}$$

Consequently,

$$\begin{aligned}
&\mathbb{E}^{\mathbb{Q}}\left[\mathbb{1}_{\{\tau_K > s\}} \mathbb{1}_A L_t^{\tau_K} X_t^{\tau_K}\right] = \mathbb{E}^{\mathbb{Q}}\left[\mathbb{1}_{A \cap \{\tau_K > s\}} \mathbb{E}^{\mathbb{Q}}\left[\frac{d\bar{\mathbb{Q}}}{d\mathbb{Q}} \,\middle|\, \mathscr{F}_{\tau_K \wedge t}\right] X_{\tau_K \wedge t}\right] \\
&= \mathbb{E}^{\mathbb{Q}}\left[\mathbb{E}^{\mathbb{Q}}\left[\mathbb{1}_{A \cap \{\tau_K > s\}} \frac{d\bar{\mathbb{Q}}}{d\mathbb{Q}} X_{\tau_K \wedge t} \,\middle|\, \mathscr{F}_{\tau_K \wedge t}\right]\right] = \mathbb{E}^{\mathbb{Q}}\left[\mathbb{1}_{A \cap \{\tau_K > s\}} \frac{d\bar{\mathbb{Q}}}{d\mathbb{Q}} X_{\tau_K \wedge t}\right] \\
&= \mathbb{E}^{\bar{\mathbb{Q}}}\left[\mathbb{1}_{A \cap \{\tau_K > s\}} X_t^{\tau_K}\right] = \mathbb{E}^{\bar{\mathbb{Q}}}\left[\mathbb{1}_{A \cap \{\tau_K > s\}} X_s^{\tau_K}\right] \\
&= \mathbb{E}^{\mathbb{Q}}\left[\mathbb{1}_{A \cap \{\tau_K > s\}} \frac{d\bar{\mathbb{Q}}}{d\mathbb{Q}} X_{\tau_K \wedge s}\right] = \mathbb{E}^{\mathbb{Q}}\left[\mathbb{E}^{\mathbb{Q}}\left[\mathbb{1}_{A \cap \{\tau_K > s\}} \frac{d\bar{\mathbb{Q}}}{d\mathbb{Q}} X_{\tau_K \wedge s} \,\middle|\, \mathscr{F}_{\tau_K \wedge s}\right]\right] \\
&= \mathbb{E}^{\mathbb{Q}}\left[\mathbb{1}_{A \cap \{\tau_K > s\}} \mathbb{E}^{\mathbb{Q}}\left[\frac{d\bar{\mathbb{Q}}}{d\mathbb{Q}} \,\middle|\, \mathscr{F}_{\tau_K \wedge s}\right] X_{\tau_K \wedge s}\right] = \mathbb{E}^{\mathbb{Q}}\left[\mathbb{1}_{\{\tau_K > s\}} \mathbb{1}_A L_s^{\tau_K} X_s^{\tau_K}\right],
\end{aligned} \tag{3.10}$$

where the second and eighth equalities follow from (3.9); and the fifth equality follows from the fact that X^{τ_K} is a $\bar{\mathbb{Q}}$-martingale. Putting (3.8) and (3.10) together, with an application of Lemma 2.18, establishes that $L^{\tau_K} X^{\tau_K}$ is a $\mathbb{Q}$-martingale.

Now, the continuity and non-negativity of L ensure that L^{τ_K} is a bounded $\mathbb{Q}$-martingale, whence $L^{\tau_K} \in \mathcal{M}^2(\mathbb{Q})$. It is also clear that $X^{\tau_K} \in \mathcal{M}^2(\mathbb{Q})$. Consequently, L^{τ_K} and X^{τ_K} are strongly orthogonal. By Theorem 2.64 and Proposition 2.46 we have that $[L^{\tau_K}, X^{\tau_K}] = [L^{\tau_K}, X]$ is a martingale. In other words $L^{\tau_K} - L_0^{\tau_K} \in \mathcal{M}_0^2(\mathbb{Q})$ is strongly orthogonal to $X - X_0$. Since $\mathbb{Q}$ is extremal point in $\mathbf{P}(\mathbb{Q}; X - X_0)$, we conclude by Theorem 2.73 that $L^{\tau_K} - L_0^{\tau_K} = 0$. Finally,

[2] To see this, let $u \in [0,T]$. Then $A \cap \{\tau_K > s\} \cap \{\tau_K \leq u\} = \emptyset \in \mathscr{F}_u$, if $u \leq s$; while, if $u > s$, $A \cap \{\tau_K > s\} \in \mathscr{F}_s \subseteq \mathscr{F}_u$ and $\{\tau_K \leq u\} \in \mathscr{F}_u$ imply that $A \cap \{\tau_K > s\} \cap \{\tau_K \leq u\} \in \mathscr{F}_u$ as well (see Definition 2.11).

the fact that L is a continuous uniformly integrable $\mathbb{Q}$-martingale implies that[3] $\lim_{K\to\infty}\tau_K = T$, $\mathbb{Q}$-a.s. Hence $L - L_0 = \lim_{K\to\infty}(L^{\tau_K} - L_0^{\tau_K}) = 0$, $\mathbb{Q}$-a.s. The assumption that $\mathscr{F}_0$ is $\mathbb{P}$-trivial (and hence also $\mathbb{Q}$-trivial) implies that $L_0 = 1$; and so we have $L_T = \frac{d\bar{\mathbb{Q}}}{d\mathbb{Q}} = 1$, from which it follows that $\bar{\mathbb{Q}} = \mathbb{Q}$.
$(3 \Rightarrow 2)$ Suppose $\mathbf{P}_{\mathsf{e}}(X) = \{\mathbb{Q}\}$. Clearly then $\mathbb{Q} \in \mathbf{P}(\mathbb{Q}; X - X_0)$. Now assume that $\mathbb{Q}$ is not an extremal point of $\mathbf{P}(\mathbb{Q}; X - X_0)$. Then there exist probability measures $\mathbb{Q}_1, \mathbb{Q}_2 \in \mathbf{P}(\mathbb{Q}; X - X_0)$, with $\mathbb{Q}_1 \neq \mathbb{Q}_2$ and some $\lambda \in (0,1)$, such that $\mathbb{Q} = \lambda\mathbb{Q}_1 + (1-\lambda)\mathbb{Q}_2$. Now fix $\gamma \in (0,1)$, with $\gamma \neq \lambda$ and define the probability measure $\bar{\mathbb{Q}} := \gamma\mathbb{Q}_1 + (1-\gamma)\mathbb{Q}_2$. Clearly $\mathbb{Q} \neq \bar{\mathbb{Q}}$. Since $\mathbf{P}(\mathbb{Q}; X - X_0)$ is a convex set, $\bar{\mathbb{Q}} \in \mathbf{P}(\mathbb{Q}; X - X_0)$; so by Definition 2.67 we have $\bar{\mathbb{Q}} \ll \mathbb{Q}$. Also, for any $A \in \mathscr{F}$ with $\mathbb{Q}\{A\} > 0$, the positivity of λ and $1-\lambda$, ensure that either $\mathbb{Q}_1\{A\} > 0$ and/or $\mathbb{Q}_2\{A\} > 0$. This implies that $\bar{\mathbb{Q}}\{A\} > 0$ whence $\mathbb{Q} \ll \bar{\mathbb{Q}}$. Therefore, $\bar{\mathbb{Q}} \sim \mathbb{Q}$, from which it follows that $\bar{\mathbb{Q}} \in \mathbf{P}_{\mathsf{e}}(X)$; a contradiction. Thus $\mathbb{Q}$ is in fact an extremal point of $\mathbf{P}(\mathbb{Q}; X - X_0)$. Consequently, $X - X_0$ has the predictable representation property under $\mathbb{Q}$, according to item (4) of Theorem 2.74. ■

3.4 The Construction of Equivalent Martingale Measures

In this section we derive two results that provide a mechanism for constructing an equivalent local martingale measure for the price process. The first result establishes a relationship between the density process for the local martingale measure and the canonical decomposition of the price process, under the assumption that the price process is a special semimartingale. If we assume that the price process has a certain representation, called the structure condition, then the second result asserts the existence of an equivalent martingale measure for it.

In order to construct an equivalent local martingale measure for the price process, the first step is to identify a candidate density process. Here we follow an approach due to Christopeit and Musiela [8, 69], which relates the density process to the Doléans exponential of another local martingale.

Proposition 3.22. *Suppose that $X \in \mathcal{S}_{\mathsf{p}}$, with canonical decomposition $X = X_0 + M + A$, where X_0 is $\mathscr{F}_0$-measurable, $M \in \mathcal{M}_{0,\mathsf{loc}}$ with $|\Delta M| \leq c$, for $c \in \mathbb{R}_+$, and $A \in \mathcal{V}_0$ is predictable. If $Y \in \mathcal{M}_{0,\mathsf{loc}}$ satisfies $\mathbb{E}\left[\mathscr{E}(Y)_T\right] = 1$, $\Delta Y > -1$ and*

[3] Suppose on the contrary, that $A := \{\lim_{K\to\infty}\tau_K < T\}$ satisfies $\mathbb{Q}\{A\} > 0$. Continuity of L and almost sure convergence to its terminal value mean that $\lim_{t\to T}\mathbb{I}_A L_t = \infty = \mathbb{I}_A L_T$, $\mathbb{Q}$-a.s. But then

$$\mathbb{E}^{\mathbb{Q}}\left[|L_T|\right] = \mathbb{E}^{\mathbb{Q}}\left[L_T\right] > \mathbb{E}^{\mathbb{Q}}\left[\mathbb{I}_A L_T\right] = \infty$$

contradicts the fact that $L \in L^1(\mathbb{Q})$.

$A + \langle M, Y \rangle = 0$, then $\mathscr{E}(Y)$ is the density process of an equivalent local martingale measure for X. Conversely, every equivalent local martingale measure for X can be obtained in this way.

Proof. ($\Leftarrow$) Let $\mathbb{Q}$ be an equivalent local martingale measure for X. Its density process Z is given by

$$Z_t := \mathbb{E}\left[\frac{d\mathbb{Q}}{d\mathbb{P}} \,\middle|\, \mathscr{F}_t \right],$$

for all $t \in [0, T]$. Note that Z is a uniformly integrable martingale, with $Z_0=1$ and $\mathbb{E}[Z_T] = 1$. Furthermore, due to the equivalence of $\mathbb{P}$ and $\mathbb{Q}$, we have

$$\mathbb{P}\left\{ \inf_{t \in [0,T]} Z_t > 0 \right\} = 1, \tag{3.11}$$

according to Jacod and Shiryaev [51, Prop. 3.5, p. 167]. Consequently, Z is strictly positive and the process Z_-^{-1} is well-defined. Now define the sequence of stopping times $(\tau_n)_{n \in \mathbb{N}}$ as follows:

$$\tau_n := \inf \left\{ t \in [0, T] \,\middle|\, Z_t \leq \frac{1}{n} \right\},$$

for all $n \in \mathbb{N}$. It follows from (3.11) that for almost all $\omega \in \Omega$ there exists an $N(\omega) \in \mathbb{N}$ such that $Z_t(\omega) > 1/N(\omega)$, for all $t \in [0, T]$; whence $\tau_{N(\omega)}(\omega) = T$. Thus $\mathbb{1}_{\{\tau_n = T\}} \uparrow 1$ a.s. as $n \to \infty$. A simple application of the monotone convergence theorem then gives

$$\lim_{n \to \infty} \mathbb{P}\{\tau_n = T\} = \lim_{n \to \infty} \mathbb{E}\left[\mathbb{1}_{\{\tau_n = T\}}\right] = \mathbb{E}\left[\lim_{n \to \infty} \mathbb{1}_{\{\tau_n = T\}}\right] = 1.$$

According to (3.7), this implies that $\tau_n \uparrow T$ a.s. (a fact that we could have deduced from (3.11) directly). Finally, note that the left continuity of Z_-^{-1} and the definition of the stopping times ensures that $1/Z_-^{\tau_n} \leq n$, for all $n \in \mathbb{N}$. In other words, Z_-^{-1} is locally bounded. Therefore, by Theorem 2.34, the following stochastic integral is well-defined:

$$Y := \int_0^{\cdot} Z_{s-}^{-1} \, dZ_s. \tag{3.12}$$

It then follows from the associativity of stochastic integrals (see Theorem 2.41 (4)) that

$$Z_t = Z_0 + \int_0^t dZ_s = 1 + \int_0^t Z_{s-} \, dY_s,$$

for all $t \in [0, T]$. Uniqueness of the solution to (2.6) implies that $Z = \mathscr{E}(Y)$. Note that $Y_0 = 0$, $\mathbb{E}[\mathscr{E}(Y)_T] = 1$ and, since Z is strictly positive, $\Delta Y > -1$.

Next, the stochastic integration by parts formula yields

$$X_t Z_t = X_0 Z_0 + \int_0^t X_{s-} \, dZ_s + \int_0^t Z_{s-} \, dX_s + [X, Z]_t,$$

for all $t \in [0, T]$. Expanding the second integral in terms of the canonical decomposition of X gives

$$\begin{aligned} X_t Z_t &= X_0 + \int_0^t X_{s-}\, dZ_s + \int_0^t Z_{s-}\, dM_s + \int_0^t Z_{s-}\, dA_s + [X, Z]_t \\ &= X_0 + N_t + B_t, \end{aligned} \tag{3.13}$$

for all $t \in [0, T]$. Here N, given by

$$N_t := \int_0^t X_{s-}\, dZ_s + \int_0^t Z_{s-}\, dM_s,$$

for all $t \in [0, T]$, is a local martingale by Theorem 2.40; while B is a finite variation process, given by

$$\begin{aligned} B_t &:= \int_0^t Z_{s-}\, dA_s + [X, Z]_t \\ &= \int_0^t Z_{s-}\, dA_s + \left[X, \int_0^{\cdot} Z_{s-} dY_s \right]_t \\ &= \int_0^t Z_{s-}\, dA_s + \int_0^t Z_{s-}\, d[X, Y]_s \\ &= \int_0^t Z_{s-}\, d(A + [X, Y])_s, \end{aligned} \tag{3.14}$$

for all $t \in [0, T]$. Here the third equality is an instance Theorem 2.51, while the fourth follows from item (5) of Theorem 2.41.

Since $X \in \mathcal{M}_{\mathsf{loc}}(\mathbb{Q})$, it follows that $XZ \in \mathcal{M}_{\mathsf{loc}}$, by Corollary 2.62 and so $B \in \mathcal{M}_{0,\mathsf{loc}}$ as well. Rearranging (3.14) (again by associativity of stochastic integrals) yields

$$\int_0^t Z_{s-}^{-1}\, dB_s = A_t + [X, Y]_t,$$

for all $t \in [0, T]$. Since the left hand side is a local martingale, by Theorem 2.40, this means that $A + [X, Y] \in \mathcal{M}_{0,\mathsf{loc}}$. Another application of the semimartingale decomposition of X yields

$$\begin{aligned} A_t + [X, Y]_t &= A_t + [M, Y]_t + [A, Y]_t \\ &= A_t + \langle M, Y \rangle_t + ([M, Y]_t - \langle M, Y \rangle_t) + \sum_{s \le t} \Delta A_s \Delta Y_s, \end{aligned} \tag{3.15}$$

for all $t \in [0, T]$. (Note that $\langle M, Y \rangle$ exists, since the jumps of M are bounded — see the discussion following Theorem 2.49.) Thus the term in brackets is a local martingale by Definition 2.48, while the final term is a local martingale since A is predictable (by Lemma 2.47). Consequently, the predictable finite variation part of (3.15) must be identically zero and we therefore have $A_t + \langle M, Y \rangle_t = 0$, for all

$t \in [0,T]$.
($\Rightarrow$) Let $Y \in \mathcal{M}_{0,\mathsf{loc}}$ with $\mathbb{E}\left[\mathscr{E}(Y)_T\right] = 1$, $\Delta Y > -1$ and $A + \langle M, Y\rangle = 0$ (as mentioned before, our assumptions are sufficient to ensure that $\langle M, Y\rangle$ exists) and set $Z := \mathscr{E}(Y)$. Then Z is strictly positive and Proposition 2.53 informs us that $Z \in \mathcal{M}_{\mathsf{loc}}$. By Lemma 3.9, Z is thus a supermartingale and since $\mathbb{E}\left[Z_0\right] = \mathbb{E}\left[Z_T\right] = 1$, it is in fact a martingale. Moreover, it is uniformly integrable by Jacod and Shiryaev [51, Lem. 1.44, p. 11]. Consequently Z is the density process for a probability measure $\mathbb{Q} \sim \mathbb{P}$. Using the same reasoning as in the first part of the proof (see (3.13)), it can be shown that $XZ = X_0 + N + B$, where N is a local martingale and B is given by (3.14). By the assumptions on Y, we now get

$$\begin{aligned} A_t + [X,Y]_t &= A_t + [M,Y]_t + [A,Y]_t \\ &= A_t + \langle M,Y\rangle_t + ([M,Y]_t - \langle M,Y\rangle_t) + \sum_{s\le t} \Delta A_s \Delta Y_s, \\ &= ([M,Y]_t - \langle M,Y\rangle_t) + \sum_{s\le t} \Delta A_s \Delta Y_s, \end{aligned}$$

for all $t \in [0,T]$. This implies that $A + [X,Y] \in \mathcal{M}_{0,\mathsf{loc}}$, whence $B \in \mathcal{M}_{0,\mathsf{loc}}$. It thus follows that $XZ \in \mathcal{M}_{\mathsf{loc}}$, which establishes that $\mathbb{Q}$ is an equivalent local martingale measure for X. ■

The proof of this theorem could also have been derived directly by an application of Girsanov's theorem. In fact, it is just a special case of Girsanov's theorem. Note that the constraint on the size of the jumps ensures that the Doléans exponential is a strictly positive process; this is necessary if we wish to produce a probability measure. If signed martingale measures are allowed, then this constraint may be relaxed.

We now introduce a representation for the price process known in the literature as the structure condition (see [86, 89]). If a process satisfies this condition, then, under an additional assumption on its jump sizes, it admits an equivalent local martingale measure.

Definition 3.23. Let $X \in \mathcal{S}_{\mathsf{p}}$. We say that X satisfies the *structure condition* if its canonical decomposition takes the form

$$X_t = X_0 + M_t + \int_0^t \alpha_s \, d\langle M\rangle_s, \tag{3.16}$$

for all $t \in [0,T]$, where $M \in \mathcal{M}^2_{0,\mathsf{loc}}$ and $\alpha \in \mathcal{L}^2_{\mathsf{loc}}(M)$.

Theorem 3.24. *Suppose that the canonical decomposition of $X \in \mathcal{S}_{\mathsf{p}}$ satisfies* (3.16) *and that $\alpha \Delta M < 1$. Then X admits an equivalent local martingale measure.*

Proof. Define the process Y, by setting

$$Y_t := -\int_0^t \alpha_s \, dM_s,$$

for all $t \in [0,T]$. Since $\Delta Y = -\alpha \Delta M > -1$, by assumption and

$$\begin{aligned} A_t + \langle M, Y \rangle_t &= \int_0^t \alpha_s \, d\langle M \rangle_s - \left\langle M, \int_0^\cdot \alpha_s \, dM_s \right\rangle_t \\ &= \int_0^t \alpha_s \, d\langle M \rangle_s - \int_0^t \alpha_s \, d\langle M \rangle_s \\ &= 0, \end{aligned}$$

for all $t \in [0,T]$, Y defines an equivalent local martingale measure for X, by Theorem 3.22. ■

We shall see in Chapter 5 that although the structure condition guarantees the existence of an equivalent local martingale measure for X, it does not imply that this measure is unique. Uniqueness is only guaranteed in a complete market.

3.5 A First Application: The Black-Scholes Model

We are now in a position to apply the mathematical machinery developed thus far. We start with a familiar example, by deriving the Black-Scholes partial differential equation (PDE) (see [6, 64]). Suppose the discounted stock price process X has dynamics

$$dX_t = (\mu - r) X_t \, dt + \sigma X_t \, dW_t, \tag{3.17}$$

for all $t \in [0,T]$, where W is a standard Brownian motion and μ, σ and r are constants, with $\mu > r$ and $\sigma > 0$. Note that we have assumed a constant short rate, so the bank account process B is given by

$$B_t = e^{rt},$$

for all $t \in [0,T]$. We do not prove it here, but the market described by these two assets is complete (see e.g. [5, 94]). The canonical decomposition of X is

$$X_t = X_0 + M_t + A_t,$$

where

$$M_t := \int_0^t \sigma X_s \, dW_s \qquad \text{and} \qquad A_t := \int_0^t (\mu - r) X_s \, dt,$$

for all $t \in [0,T]$. Since X is continuous, it satisfies the structure condition (3.16), with the process α featuring there given by

$$\alpha_t := \frac{\mu - r}{\sigma^2 X_t},$$

for all $t \in [0,T]$. As a result of Theorem 3.24, we may easily construct a martingale measure $\mathbb{Q}$ for X, with density process

$$Z_t := \mathbb{E}\left[\frac{d\mathbb{Q}}{d\mathbb{P}}\,\middle|\,\mathscr{F}_t\right] = \mathscr{E}(Y)_t\,,$$

where

$$Y_t := -\int_0^t \alpha_t\,dM_t = -\int_0^t \frac{\mu-r}{\sigma^2 X_t}\,dM_t = -\frac{\mu-r}{\sigma}W_t,$$

for all $t \in [0,T]$. Note that the fraction $\frac{\mu-r}{\sigma}$ in the last expression is the familiar Sharpe ratio, or market price of risk. Now define a new process $\widehat{W}$, by setting

$$\widehat{W}_t := W_t + \frac{\mu-r}{\sigma}t,$$

in which case (3.17) becomes

$$dX_t = \sigma X_t\,d\widehat{W}_t, \tag{3.18}$$

for all $t \in [0,T]$. We shall next verify that the $\widehat{W}$ is a Brownian motion under $\mathbb{Q}$, by appealing to Lévy's characterization of Brownian motion.

Theorem 3.25 (Lévy)**.** *A process M is a Brownian motion iff it is a continuous local martingale with $\langle M\rangle_t = t$, for all $t \in [0,T]$.*

Proof. See Protter [74, Thm. 39, p. 86]. ∎

Clearly $\widehat{W}$ is continuous. To see that it is a local $\mathbb{Q}$-martingale, we note that the stochastic integration by parts rule gives

$$\begin{aligned}
\widehat{W}_t Z_t &= \int_0^t \widehat{W}_s\,dZ_t + \int_0^t Z_s\,d\widehat{W}_t + \big[\widehat{W}, Z\big]_t \\
&= -\int_0^t \widehat{W}_s Z_s\,dY_s + \int_0^t Z_s\left(dW_s + \frac{\mu-r}{\sigma}\,ds\right) \\
&\quad + \left[W - \int_0^\cdot \frac{\mu-r}{\sigma}\,ds\,, -\int_0^\cdot Z_s\,dY_s\right]_t \\
&= -\int_0^t \frac{\mu-r}{\sigma}\widehat{W}_s Z_s\,dW_s + \int_0^t Z_s\,dW_s \\
&\quad + \int_0^t \frac{\mu-r}{\sigma}Z_s\,ds - \int_0^t \frac{\mu-r}{\sigma}Z_s\,d[W]_s \\
&= \int_0^t \left(1 - \frac{\mu-r}{\sigma}\widehat{W}_s\right) Z_s\,dW_s,
\end{aligned}$$

for all $t \in [0,T]$, which is a local martingale. Since $\widehat{W}$ and W are continuous, we see that

$$\big\langle\widehat{W}\big\rangle_t^{\mathbb{Q}} = \big[\widehat{W}\big]_t = \left[W - \int_0^\cdot \frac{\mu-r}{\sigma}\,ds\right]_t = [W]_t = \langle W\rangle_t = t,$$

for all $t \in [0, T]$. By Theorem 3.25, it follows that $\widehat{W}$ is a Brownian motion under $\mathbb{Q}$. It then follows from (3.18) that X is a local $\mathbb{Q}$-martingale. In fact, it is a $\mathbb{Q}$-martingale (see Protter [74, Example, p. 76–77] for justification).

Now consider a European option on X, with maturity T and discounted payoff $h(X_T)$, for some Borel-measurable function $h : \mathbb{R}_+ \to \mathbb{R}_+$ satisfying $h(X_t) \in L^1$, for all $t \in [0, T]$. As mentioned earlier, the market is complete, so by the second fundamental theorem of asset pricing, we know that $\mathbb{Q}$ is unique and that there exists a $\mathbb{Q}$-admissible self-financing strategy (ξ, η), such that

$$V_T(\xi, \eta) = h(X_T).$$

The price of this option is then given by

$$V_0 = \mathbb{E}^{\mathbb{Q}}\left[h(X_T)\right]$$

and is unique and preference-independent.

The next result provides a link between the representation of the price of the claim as an expected value computed under the risk-neutral measure and its representation as the solution of a certain PDE.

Theorem 3.26 (Feynman-Kač). *Consider the stochastic differential equation*

$$dY_t = \mu(t, Y_t)\, dt + \sigma(t, Y_t)\, dW_t,$$

for all $t \in [0, T]$. Let $g : \mathbb{R} \to \mathbb{R}$ be a Borel-measurable function satisfying $g(Y_t) \in L^1$, for all $t \in [0, T]$. Define the function $G : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$, by setting

$$G(t, y) := \mathbb{E}\left[g(Y_T) \,|\, Y_t = y\right],$$

for all $(t, y) \in \mathbb{R}_+ \times \mathbb{R}$. Then G satisfies the following PDE:

$$\frac{\partial G}{\partial t}(t, y) + \mu(t, y)\frac{\partial G}{\partial y}(t, y) + \frac{1}{2}\sigma^2(t, y)\frac{\partial^2 G}{\partial y^2}(t, y) = 0,$$

for all $(t, y) \in \mathbb{R}_+ \times \mathbb{R}$, with the terminal condition $G(T, y) = g(y)$, for all $y \in \mathbb{R}$.

Proof. See Shreve [94, Thm. 6.4.1, p. 268]. ■

The Feynman-Kač theorem facilitates the construction of a PDE representation for the hedging portfolio, in the case where the underlying asset is driven by a Brownian motion. We now define a function $F : [0, T] \times (0, \infty) \to \mathbb{R}_+$, by setting

$$F(t, x) = \mathbb{E}^{\mathbb{Q}}\left[h(X_T) \,|\, X_t = x\right],$$

for all $(t, x) \in [0, T] \times (0, \infty)$. Note that this has the following trivial consequence:

$$F(T, X_T) = h(X_T). \tag{3.19}$$

According to the Feynman-Kač theorem, F satisfies the PDE[4]

$$\frac{\partial F}{\partial t}(t,x) + \frac{1}{2}\sigma^2 x^2 \frac{\partial^2 F}{\partial x^2}(t,x) = 0, \tag{3.20}$$

for all $(t,x) \in [0,T] \times (0,\infty)$, with (3.19) as the appropriate boundary condition. Applying Itô's formula to the process $(F(t,X_t))_{t\in[0,T]}$ yields

$$\begin{aligned} h(X_T) = F(T,X_T) = F(0,X_0) &+ \int_0^T \sigma X_s \frac{\partial F}{\partial x}(s,X_s)\, d\widehat{W}_s \\ &+ \int_0^T \left(\frac{\partial F}{\partial t}(s,X_s) + \frac{1}{2}\sigma^2 X_s^2 \frac{\partial^2 F}{\partial x^2}(s,X_s) \right) ds. \end{aligned}$$

In the light of (3.18) and (3.20), this becomes

$$h(X_T) = F(0,X_0) + \int_0^T \frac{\partial F}{\partial x}(s,X_s)\, dX_s.$$

Since X is a $\mathbb{Q}$-martingale, we obtain $F(0,X_0) = \mathbb{E}^{\mathbb{Q}}[h(X_T)]$ from the above. This corresponds to the unique price of the claim obtained earlier by martingale methods. Consequently, the self-financing strategy (ξ^h, η^h), defined by

$$(\xi_t^h, \eta_t^h) = \left(\frac{\partial F}{\partial x}(t,X_t), \mathbb{E}^{\mathbb{Q}}[h(X_T)] + G(\xi^h)_t - \xi_t^h X_t \right),$$

for all $t \in [0,T]$, hedges the option. The quantity $\frac{\partial F}{\partial x}(t,x)$, for all $(t,x) \in [0,T] \times (0,\infty)$, is known as the option *delta*.

Up to this point, we have only considered the discounted claim and the hedging portfolio with respect to the discounted asset price. In order to express the problem in terms of non-discounted quantities, we transform the PDE (3.20). First we define the function $\bar{F} : [0,T] \times (0,\infty) \to \mathbb{R}_+$, by setting

$$\bar{F}(t,s) := e^{rt} F(t, e^{-rt}s),$$

for all $(t,s) \in [0,T] \times (0,\infty)$. It then follows that

$$\begin{aligned} \frac{\partial \bar{F}}{\partial s}(t,s) &= \frac{\partial F}{\partial x}(t, e^{-rt}s); \\ \frac{\partial^2 \bar{F}}{\partial s^2}(t,s) &= e^{-rt} \frac{\partial^2 F}{\partial x^2}(t, e^{-rt}s); \quad \text{and} \\ \frac{\partial \bar{F}}{\partial t}(t,s) &= r e^{rt} F(t, e^{-rt}s) + e^{rt} \frac{\partial F}{\partial t}(t, e^{-rt}s) - rs \frac{\partial F}{\partial x}(t, e^{-rt}s) \\ &= r\bar{F}(t,s) + e^{rt} \frac{\partial F}{\partial t}(t, e^{-rt}s) - rs \frac{\partial \bar{F}}{\partial s}(t,s), \end{aligned}$$

[4] Note that we are considering the dynamics of X under the measure $\mathbb{Q}$, as determined by (3.18).

for all $(t, s) \in [0, T] \times (0, \infty)$. The PDE (3.20) may now be rewritten as

$$e^{-rt}\frac{\partial \bar{F}}{\partial t}(t,s) - re^{-rt}\bar{F}(t,s) + re^{-rt}s\frac{\partial \bar{F}}{\partial s}(t,s) + \frac{1}{2}\sigma^2(e^{-rt}s)^2 e^{rt}\frac{\partial^2 \bar{F}}{\partial s^2}(t,s) = 0,$$

which may in turn be rearranged to give

$$\frac{\partial \bar{F}}{\partial t}(t,s) + rs\frac{\partial \bar{F}}{\partial s}(t,s) + \frac{1}{2}\sigma^2 s^2 \frac{\partial^2 \bar{F}}{\partial s^2}(t,s) = r\bar{F}(t,s),$$

for all $(t, s) \in [0, T] \times (0, \infty)$. The boundary condition corresponding to (3.19) is now

$$\bar{F}(T,s) = e^{rT}F(T, e^{-rT}s) = e^{rT}h(e^{-rT}s),$$

for all $s \in (0, \infty)$. This is of course the celebrated Black-Scholes PDE. Solving it, subject to the above boundary condition, yields the familiar Black-Scholes option pricing formula initially presented in Black and Scholes [6].

Chapter 4

Incomplete Markets: The Martingale Case

In the previous chapter we considered the pricing and hedging of contingent claims in a complete market setting and it was shown that a self-financing strategy could be constructed to price and hedge claims so that no risk was assumed by the writer. Following the account of Föllmer and Sondermann [33], we shall now incorporate market incompleteness in the form of unhedgeable or intrinsic risk.

With the introduction of intrinsic risk, the problem now becomes one of finding a hedging strategy that minimizes risk in a suitable manner. Since incompleteness imposes a regime where a contingent claim no longer has a unique (risk-free) price, we introduce a cost process that expresses the difference between the value of the hedging portfolio and the gains associated with hedging. This process provides a proxy for the price of the claim and allows us to model the difference between its intrinsic value (which we take to mean the expected value of the claim, conditional on all information available at the present time) and its cost. A risk process is then defined by specifying a quadratic functional of the cost process, thereby providing a measure of the variance of the cost. Strategies that minimize this variance are optimal, in the sense that they ensure the least deviation between the hedging portfolio and the intrinsic value of the claim. In this respect, two approaches to hedging will be considered — the strategy that minimizes the total risk and the strategy that minimizes the conditional remaining risk.

In this chapter we consider the simplest case, where the discounted stock price process is a martingale under the real-world measure. In this setting it turns out that the same stock-holding is used for both strategies outlined above. The simplest total risk-minimizing strategy employs a bank account holding corresponding to a self-financing strategy until maturity, at which time there is a possible shortfall or surplus; while the strategy that minimizes remaining risk is mean self-financing. Subsequent chapters will generalize the situation, leading to the minimal martingale measure and the variance-optimal martingale measure, respectively. We shall then see that the corresponding trading strategies are substantially different, due to a subtlety of the semimartingale representation.

4.1 Market Assumptions

As before, let $T \in (0,\infty)$ and fix a stochastic basis $(\Omega, \mathscr{F}, \mathbf{F}, \mathbb{P})$. We suppose that the filtration $\mathbf{F} = (\mathscr{F}_t)_{t\in[0,T]}$ satisfies the usual conditions, that $\mathscr{F}_T = \mathscr{F}$ and that $\mathscr{F}_0$ contains only the null sets of $\mathscr{F}$ and their complements. All processes are defined on the stochastic basis above (in particular, they are defined over the time interval $[0,T]$) and are implicitly assumed to be adapted to $\mathbf{F}$.

We assume the existence of a discounted stock price process X, a bank account process B and an $\mathscr{F}_T$-measurable claim H (maturing at time T) that we wish to hedge. For now we consider the situation where X is a square integrable martingale[1] and set aside consideration of the more general case where X is only a semimartingale. We assume further that $H \in L^2$ (the reason for this will become clear when we introduce the cost and risk processes).

By the GKW decomposition theorem (Proposition 2.75), H has the following (unique) representation:

$$H = H_0 + \int_0^T \xi_s^H \, dX_s + L_T^H, \tag{4.1}$$

where $H_0 = \mathbb{E}[H]$, $\xi^H \in \mathcal{L}^2(X)$ and $L^H \in \mathcal{M}_0^2$ is strongly orthogonal to X.

The economic interpretation of equation (4.1) is that any claim may be decomposed into a component that can be hedged using a portfolio containing holdings of the stock and the bank account and an unhedgeable component. If $L^H = 0$, then all the risk from H can be hedged and we are back in the complete market of Chapter 3. However, it is a feature of many practical problems — for example, hedging a non-traded asset with a correlated asset, hedging under the assumption of stochastic volatility, hedging an index with a subset of its constituents, etc. — that $L^H \neq 0$, in which case it is impossible to find a self-financing strategy that replicates the claim perfectly.

We now introduce the class of feasible strategies[2]. These are strategies (not

[1] The results in this chapter have been generalized to the situation where X is a locally square integrable martingale [91].

[2] Our account deviates from that of Föllmer and Sondermann [33]. What we call feasible strategies correspond to what they call “admissible” strategies. We do this to avoid a clash of terminology with our definition of admissibility in Chapter 3. Note also that Schweizer [91] calls this class of strategies RM-strategies.

necessarily self-financing) that replicate the claim at maturity and possess certain integrability properties.

Definition 4.1. A strategy (ξ, η) is called *feasible* if $\xi \in \mathcal{L}^2(X)$, $V(\xi, \eta)$ is right continuous with $V_t(\xi, \eta) \in L^2$, for all $t \in [0, T]$ and $V_T(\xi, \eta) = H$ a.s.

Note that since X is a square integrable martingale, the gain process associated with a feasible strategy is also a square integrable martingale and consequently every feasible strategy is $\mathbb{P}$-admissible. Therefore, by the reasoning of the discussion following the Folk Theorem in Chapter 3, we can conclude that no feasible strategy is an arbitrage. This martingale property of the gain process is important for a number of proofs in this chapter and is a consequence of the following lemma.

Lemma 4.2. *If $M \in \mathcal{M}^2$ and $\phi \in \mathcal{L}^2(M)$, then $\int_0^{\cdot} \phi_s \, dM_s \in \mathcal{M}^2$.*

Proof. See Protter [74, p. 171]. ■

Since Definition 4.1 does not insist on the self-financing constraint, it is always possible to construct a feasible strategy that replicates the payoff of the claim. The simplest example is the unhedged strategy, which amounts to doing nothing until maturity, at which point the value of the bank account is set equal to the payoff. In detail, this is the feasible strategy (ξ, η), with $\xi_t = 0$ and $\eta_t = H\mathbb{1}_{\{t=T\}}$, for all $t \in [0, T]$. It is obvious that this strategy incurs significant risk. The task of hedging in the current incomplete market setting involves finding a feasible trading strategy that minimizes risk in an appropriate manner.

4.2 Cost and Risk Processes

Since the writer of a claim is exposed to risk when hedging in an incomplete market, it is appropriate to inquire about his appetite for risk. This is usually expressed mathematically as a utility function. Here we assume that the writer has a quadratic utility of wealth. This leads to measures of risk based on the variance of the total portfolio outcome. Consequently, we define a cost process that models the deviation of the hedge portfolio from the intrinsic value of the claim and a risk process which is a quadratic function of the cost.

There are known objections to the use of quadratic utility, since it penalizes both the losses and the profits associated with bearing risk [87, §1, Remark 3]. However, it does provide a symmetric pricing mechanism which does not depend on whether the claim is held long or short. We now introduce the cost process.

Definition 4.3. The *cost process* $C(\xi, \eta)$ associated with a feasible strategy (ξ, η) is defined by

$$C_t(\xi, \eta) := V_t(\xi, \eta) - G_t(\xi), \tag{4.2}$$

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

It is important to note that the concept of cost in this definition does not refer to transaction costs, which we do not model, but rather relates to the cost of hedging the claim. In the complete market of the previous chapter, the cost of hedging was just the risk-neutral price, which could be calculated at the inception of the claim. With the introduction of a cost process, we are now able to incorporate the shortfall or surplus resulting from imperfect hedging over the life of the claim.

From Definition 4.3 it can be seen that when a feasible strategy (ξ, η) is used, the claim may be represented as

$$H = V_T(\xi, \eta) = C_T(\xi, \eta) + G_T(\xi) \quad \text{a.s.} \tag{4.3}$$

This generalizes the representation of the claim derived in the previous chapter, with the complete market characterized by the trivial case where $C_t(\xi, \eta) = V_0(\xi, \eta)$, for all $t \in [0, T]$. Furthermore, by substituting (4.2) into (3.1) we obtain

$$\eta_t = C_t(\xi, \eta) + G_t(\xi) - \xi_t X_t,$$

for all $t \in [0, T]$. Comparing this with (3.4), we see that (ξ, η) is self-financing only if $C_t(\xi, \eta) = V_0(\xi, \eta)$, for all $t \in [0, T]$.

A number of risk criteria based on the cost process may now be formulated; the most obvious being quadratic criteria. Schäl [82] has identified three quadratic functionals of the cost process that we may choose to minimize in a discrete-time setup. Given the discrete-time price process $\{ X_t \mid t = 0, 1, \ldots, T\}$, these are[3]:

1. the *local conditional risk* $\mathbb{E}\left[(C_{t+1} - C_t)^2 \,\middle|\, \mathscr{F}_t \right]$;
2. the *conditional remaining risk* $\mathbb{E}\left[(C_T - C_t)^2 \,\middle|\, \mathscr{F}_t \right]$; and
3. the *total risk* $\mathbb{E}\left[(C_T - C_0)^2 \right]$.

Analogues of these risk measures exist for continuous-time. In this chapter we explore continuous-time formulations of the last two criteria. For the general case, where X is only a semimartingale, it turns out that no continuous-time analogue for the second criterion can be formulated; however, analogues of the first and third criteria will be investigated in Chapters 5 and 6, respectively. Motivated by these considerations, we now define the risk process.

[3] As in the previous chapter, the arguments of $C(\xi, \eta)$ will be dropped when the strategy is clear from the context or redundant.

Definition 4.4. The *risk process* $R(\xi, \eta)$ associated with a feasible strategy (ξ, η) is defined by

$$R_t(\xi, \eta) := \mathbb{E}\left[(C_T(\xi, \eta) - C_t(\xi, \eta))^2 \,\middle|\, \mathscr{F}_t \right],$$

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

The reason for the square integrability requirements in Definition 4.1 is now apparent; without them the risk process would not be well-defined.

Note that when the market is complete, all the risk measures above are consistent with the fact that perfect hedging is possible and no risk is assumed. As mentioned before, this corresponds to the situation where the cost process is constant through time and equal to the risk-neutral price. In that case[4], $R = 0$ as well.

In the incomplete market case, two hedging approaches seem obvious: we can insist on a self-financing strategy and make up the shortfall at maturity; or we can continually add and remove money from the hedge portfolio to ensure that its value corresponds at all times with the intrinsic value of the claim.[5] The first strategy corresponds with minimizing the total risk

$$R_0 = \mathbb{E}\left[(C_T - C_0)^2 \right].$$

A feasible trading strategy with this property is known as *mean-variance optimal.* In the second case it is natural to insist on minimizing the conditional remaining risk

$$R_t = \mathbb{E}\left[(C_T - C_t)^2 \,\middle|\, \mathscr{F}_t \right],$$

at every time $t \in [0, T]$. A feasible strategy with this property is called a *risk-minimizing* strategy; we will see later that it is also mean self-financing. We now explore these two approaches in detail.

4.3 Minimizing Total risk

The next result links the representation of the claim to the structure of the mean-variance optimal portfolio. In particular, it shows that the optimal hedging strategy is directly related to the stochastic integral in (4.1), while the initial endowment needed to hedge the claim optimally is given by its intrinsic value at inception; that is, $V_0 = H_0 = \mathbb{E}[H]$.

Theorem 4.5. *A feasible strategy (ξ, η) minimizes total risk iff $V_0(\xi, \eta) = H_0$ and $\xi = \xi^H$; in which case*

$$R_0(\xi, \eta) = \mathbb{E}\left[(C_T(\xi, \eta) - H_0)^2 \right] = \mathbb{E}\left[(L_T^H)^2 \right].$$

[4] As before, arguments of $R(\xi, \eta)$ will be dropped when the strategy is clear from the context.

[5] The intrinsic value of the claim, at time $t \in [0, T]$, is given by $\mathbb{E}[H \mid \mathscr{F}_t] = H_0 + \int_0^t \xi_s^H \, dX_s + L_t^H$.

Furthermore, the initial value and stock-holding are unique (in a μ_X-a.e. sense).

Proof. Let (ξ, η) be a feasible trading strategy, with C and R its associated cost and risk processes. By (4.1) and (4.3) we have

$$C_T = H - \int_0^T \xi_s \, dX_s = H_0 + \int_0^T (\xi_s^H - \xi_s) \, dX_s + L_T^H.$$

Since L^H is strongly orthogonal to X, it follows that

$$\begin{aligned} R_0 &= \mathbb{E}\left[\left(H_0 - C_0 + \int_0^T (\xi_s^H - \xi_s) \, dX_s + L_T^H\right)^2\right] \\ &= (H_0 - C_0)^2 + \mathbb{E}\left[\left(\int_0^T (\xi_s^H - \xi_s) \, dX_s\right)^2\right] + \mathbb{E}\left[(L_T^H)^2\right] \\ &\quad + 2(H_0 - C_0)\mathbb{E}\left[\int_0^T (\xi_s^H - \xi_s) \, dX_s\right] + 2(H_0 - C_0)\mathbb{E}\left[L_T^H\right] \\ &\quad + 2\mathbb{E}\left[L_T^H \int_0^T (\xi_s^H - \xi_s) \, dX_s\right] \\ &= (H_0 - V_0)^2 + \mathbb{E}\left[\int_0^T (\xi_s^H - \xi_s)^2 \, d\langle X \rangle_s\right] + \mathbb{E}\left[(L_T^H)^2\right], \end{aligned}$$

where the last step follows as a result of the martingale property of the gain process and L^H. This is minimized iff $V_0 = H_0$ and $\xi = \xi^H$, whence uniqueness of ξ^H follows by Lemma 2.76. ■

Although this theorem specifies the form of the stock-holding ξ, it does not impose conditions on η, other than at the initial and terminal times. These conditions are as follows: $\eta_0 = H_0 - \xi_0 X_0$, in order that $V_0 = H_0$; and $\eta_T = H - \xi_T X_T$, to ensure that the strategy is feasible. It is thus possible to impose the self-financing constraint to determine η during $(0, T)$. The resulting shortfall or excess is added or removed at termination. The resulting strategy $(\xi^\star, \eta^\star)$ is given by

$$(\xi_t^\star, \eta_t^\star) := (\xi_t^H, H_0 + G_t(\xi^H) - \xi_t^H X_t + \mathbb{1}_{\{t=T\}} L_T^H), \tag{4.4}$$

for all $t \in [0, T]$. The consequence of employing a self-financing strategy is that at maturity there is a random shortfall or profit L_T^H.

4.4 Minimizing Conditional Remaining Risk

We now investigate strategies that minimize the conditional remaining risk at each time $t \in [0, T)$. This corresponds to the idea of letting the past be the past and continuously adjusting the portfolio choice so that the remaining risk is always instantaneously minimized.

Definition 4.6. Let (ξ, η) and $\left(\tilde{\xi}, \tilde{\eta}\right)$ be feasible strategies. Then $\left(\tilde{\xi}, \tilde{\eta}\right)$ is called a *feasible continuation* of (ξ, η) at time $t \in [0, T)$, if $\tilde{\xi}_s = \xi_s$, for $s \in [0, t]$ and $\tilde{\eta}_s = \eta_s$, for $s \in [0, t)$.

Note that the time interval for $\tilde{\xi}$ is closed, while for $\tilde{\eta}$ it is half-open. In this respect we follow Schweizer [91] and deviate from the account of Föllmer and Sondermann [33], where half-open intervals are used in the specification of both processes. As Schweizer points out, the formulation of feasible continuations in Definition 4.6 can be applied in more general settings, for instance in discrete-time and in the generalization to local risk minimization (which we explore in Chapter 5). It also ensures that $\tilde{\xi}$ is predictable, while $\tilde{\eta}$ need only be adapted. This has the following important consequence (exploited in the proof of Theorem 4.10):

$$\begin{aligned} V_t(\xi, \eta) - V_t\left(\tilde{\xi}, \tilde{\eta}\right) &= (\xi_t - \tilde{\xi}_t) X_t + (\eta_t - \tilde{\eta}_t) \\ &= \eta_t - \tilde{\eta}_t, \end{aligned} \tag{4.5}$$

at the continuation time $t \in [0, T)$.

It is now possible to define a risk-minimizing strategy by insisting that it minimizes the conditional remaining risk at each time, over all feasible continuations.

Definition 4.7. A feasible strategy (ξ, η) is called *risk-minimizing* if at every time $t \in [0, T)$ we have

$$R_t(\xi, \eta) \leq R_t\left(\tilde{\xi}, \tilde{\eta}\right) \quad \text{a.s.},$$

for any feasible continuation $\left(\tilde{\xi}, \tilde{\eta}\right)$ of (ξ, η) at time t.

As mentioned previously, we are no longer constrained to use self-financing strategies. However, it is useful to define a new class of strategies, called mean self-financing.

Definition 4.8. A strategy is called *mean self-financing* if its associated cost process is a martingale.

The optimal risk-minimizing portfolio will be expressed in terms of a process that represents the intrinsic risk of the claim.

Definition 4.9. The *intrinsic risk process* R^H associated with the contingent claim H, with representation (4.1), is defined by

$$R_t^H := \mathbb{E}\left[(L_T^H - L_t^H)^2 \,\middle|\, \mathscr{F}_t \right],$$

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

Theorem 4.10. *The strategy* (ξ^*, η^*)*, defined by*

$$(\xi_t^*, \eta_t^*) := (\xi_t^H, H_0 + G_t(\xi^H) - \xi_t^H X_t + L_t^H), \tag{4.6}$$

for all $t \in [0, T]$*, is feasible and risk-minimizing. Its associated risk process* R^* *is given by*

$$R_t^* = R_t^H \quad a.s.,$$

for all $t \in [0, T]$*. Furthermore, this strategy is unique (in a* μ_X*-a.s. sense, for the stock-holding).*

Proof. First, we note that

$$V_t(\xi^*, \eta^*) = \xi_t^* X_t + \eta_t^* = H_0 + G_t(\xi^H) + L_t^H, \tag{4.7}$$

for all $t \in [0, T]$. It follows from (4.1) that $V_T(\xi^*, \eta^*) = H$. Furthermore, since $G(\xi^H), L^H \in \mathcal{M}^2$, it follows that $V(\xi^*, \eta^*) \in \mathcal{M}^2$ as well. In particular, we have established that (ξ^*, η^*) is feasible.

It follows immediately from (4.7) that the cost process C^* associated with (ξ^*, η^*) is determined by

$$C_t^* = H_0 + L_t^H, \tag{4.8}$$

for all $t \in [0, T]$. Its risk process R^* is thus given by

$$R_t^* = \mathbb{E}\left[(C_T^* - C_t^*)^2 \,\middle|\, \mathscr{F}_t\right] = \mathbb{E}\left[(L_T^H - L_t^H)^2 \,\middle|\, \mathscr{F}_t\right] = R_t^H,$$

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

Now let $(\tilde{\xi}, \tilde{\eta})$ be a feasible continuation of (ξ^*, η^*) at some time $t \in [0, T)$ and denote its associated cost and risk processes by $\tilde{C}$ and $\tilde{R}$, respectively. Then, using (4.1), we obtain

$$\begin{aligned}
\tilde{C}_T - \tilde{C}_t &= \tilde{V}_T(\tilde{\xi}, \tilde{\eta}) - \int_0^T \tilde{\xi}_s \, dX_s - V_t(\tilde{\xi}, \tilde{\eta}) + \int_0^t \tilde{\xi}_s \, dX_s \\
&= H - \int_t^T \tilde{\xi}_s \, dX_s - V_t(\tilde{\xi}, \tilde{\eta}) \\
&= H_0 + \int_0^T \xi_s^H \, dX_s + L_T^H - \int_t^T \tilde{\xi}_s \, dX_s - V_t(\tilde{\xi}, \tilde{\eta}) \\
&= H_0 + \int_0^t \xi_s^H \, dX_s + L_t^H + \int_t^T \xi_s^H \, dX_s - \int_t^T \tilde{\xi}_s \, dX_s \\
&\quad - V_t(\tilde{\xi}, \tilde{\eta}) + L_T^H - L_t^H \\
&= \int_t^T (\xi_s^H - \tilde{\xi}_s) \, dX_s + \Big(V_t(\xi^*, \eta^*) - V_t(\tilde{\xi}, \tilde{\eta})\Big) + (L_T^H - L_t^H) \\
&= \int_t^T (\xi_s^* - \tilde{\xi}_s) \, dX_s + (\eta_t^* - \tilde{\eta}_t) + (L_T^H - L_t^H),
\end{aligned}$$

where the last equality follows from (4.5) and the discussion after Definition 4.6. Finally,

$$
\begin{aligned}
\tilde{R}_t &= \mathbb{E}\left[(\tilde{C}_T - \tilde{C}_t)^2 \,\middle|\, \mathscr{F}_t \right] \\
&= \mathbb{E}\left[\left(\int_t^T (\xi_s^* - \tilde{\xi}_s)\, dX_s \right)^2 \,\middle|\, \mathscr{F}_t \right] + (\eta_t^* - \tilde{\eta}_t)^2 + \mathbb{E}\left[(L_T^H - L_t^H)^2 \,\middle|\, \mathscr{F}_t \right] \\
&\quad + 2(\eta_t^* - \tilde{\eta}_t)\mathbb{E}\left[\int_t^T (\xi_s^* - \tilde{\xi}_s)\, dX_s \,\middle|\, \mathscr{F}_t \right] + 2(\eta_t^* - \tilde{\eta}_t)\mathbb{E}\left[(L_T^H - L_t^H) \,\middle|\, \mathscr{F}_t \right] \\
&\quad + 2\mathbb{E}\left[(L_T^H - L_t^H) \int_t^T (\xi_s^* - \tilde{\xi}_s)\, dX_s \,\middle|\, \mathscr{F}_t \right] \\
&= \mathbb{E}\left[\int_t^T (\xi_s^* - \tilde{\xi}_s)^2 \, d\langle X \rangle_s \,\middle|\, \mathscr{F}_t \right] + (\eta_t^* - \tilde{\eta}_t)^2 + R_t^H \\
&\geq R_t^H.
\end{aligned}
$$

The last equality above follows from the martingale property of $\int_0^{\cdot} (\xi_s^* - \tilde{\xi}_s)\, dX_s$ and L^H and the fact that these two martingales are strongly orthogonal. We can now immediately read off from the above that (ξ^*, η^*) is both risk-minimizing and unique (by Lemma 2.76), as advertised. ■

Corollary 4.11. *The risk-minimizing strategy is mean self-financing.*

Proof. Since $L^H \in \mathcal{M}^2$, it is clear from (4.8) that C^* is a (square integrable) martingale, and consequently mean self-financing. ■

Note that in the original account of Föllmer and Sondermann [33], Theorem 4.10 was proved in two steps: firstly it was shown from first principles (i.e. without explicitly constructing the strategy) that a risk-minimizing strategy is mean self-financing; then the portfolio in equation (4.6) was shown to be risk-minimizing by an argument similar to that of Theorem 4.10. The extra difficulty encountered results from the way continuations were defined there (see the discussion following Definition 4.6).

By comparing (4.4) and (4.6), we see that the mean-variance optimal and risk-minimizing strategies differ only in the bank account holding. Since the risk-minimizing strategy is mean self-financing, any surplus between the hedging portfolio and the intrinsic value of the claim is instantaneously withdrawn and any shortfall is immediately funded.

4.5 The Hedge Ratio

We may now provide a more explicit characterization of the stock-holding component of the hedging strategies described above. This is done in the risk-minimizing case by computing the covariation with respect to the price process on both sides of (4.7). In the light of the strong orthogonality of L^H and X, this gives

$$\langle V^*, X\rangle_t = \left\langle \int_0^t \xi_s^H \, dX_s, X \right\rangle_t + \left\langle L^H, X \right\rangle_t = \int_0^t \xi_s^H \, d\langle X\rangle_s,$$

for all $t \in [0, T]$. Consequently, we may express ξ^H as the pathwise Radon-Nikodým derivative of the Lebesgue-Stieltjes measure induced by $\langle V^*, X\rangle$ with respect to the Lebesgue-Stieltjes measure induced by $\langle X\rangle$; i.e.

$$\xi^H = \frac{d\langle V^*, X\rangle}{d\langle X\rangle}.$$

As mentioned previously, ξ^H is the stock-holding for both the mean-variance and risk-minimizing strategies. This is easily seen by comparing (4.4) and (4.6). As in the previous chapter, a PDE representation of the value of the claim is possible if the price process is driven by a Brownian motion. We do not pursue this here, however. Once we have explored the more general situation where X is a semimartingale, we shall revisit the issue of PDE representations.

Chapter 5

Incomplete Markets: Local Risk Minimization

In this chapter we generalize the idea of a risk-minimizing strategy to the situation where the price process X is a square integrable semimartingale with a canonical decomposition. In this more general setting an unfortunate complication arises — Schweizer [83, 91] showed through counter-examples that a compatibility problem exists when attempting risk minimization in a semimartingale market; in general it is not possible to find a risk-minimizing strategy. This problem stems from the fact that if at each time one computes the optimal strategy over the remaining time interval, then for $s < t \in [0, T]$, it is possible to find examples where the risk-minimizing strategy over the interval $(s, T]$ is inconsistent with the risk-minimizing strategy over the interval $(t, T]$. It is therefore necessary to use another measure of risk.

Based on the idea of sequential hedging in a discrete-time framework developed in Föllmer and Schweizer [31], Schweizer [83, 84, 85] introduced the idea of local risk minimization; thereby providing a new measure of risk, suitable for optimization. This approach generalizes the second quadratic measure of risk in Schäl's classification (see Section 4.2) to continuous time and corresponds to minimizing the local conditional risk over an infinitesimal time increment, at each time.

In presenting these ideas, the first task is to provide a precise definition of local risk minimization. This is a variational concept involving the definition of a risk quotient. In an analogous manner to the risk-minimizing strategy of Chapter 4, where the remaining risk increases over every continuation, the local risk-minimizing strategy is characterized by a risk quotient that increases over all instantaneous deviations (or perturbations) of the strategy. Associated with these definitions we state a result linking locally risk-minimizing strategies to mean self-financing strategies.

Based on the above result, an alternative characterization of local risk minimization is derived, illustrating the added complication introduced by the finite variation component of the price process. It also provides insight into the measure under which pricing should occur. This is the so-called minimal martingale measure, which we construct explicitly. Finally, it is shown that pricing under the minimal martingale measure satisfies the alternative characterization of local risk minimization mentioned above.

Although we do not provide exhaustive references, it should be noted that our account borrows heavily from the original papers of Schweizer [83, 84, 85], in terms of concepts and notation.

5.1 Market Assumptions

As in previous chapters, let $T \in (0, \infty)$ and fix a stochastic basis $(\Omega, \mathscr{F}, \mathbf{F}, \mathbb{P})$. We suppose further that the filtration $\mathbf{F} = (\mathscr{F}_t)_{t\in[0,T]}$ satisfies the usual conditions, that $\mathscr{F}_T = \mathscr{F}$ and that $\mathscr{F}_0$ contains only the null sets of $\mathscr{F}$ and their complements. All processes are assumed to be defined on the above stochastic basis (in particular, they are defined over the finite time interval [0,T]) and are implicitly understood to be adapted to $\mathbf{F}$.

We assume the existence of a discounted stock price process X, a bank account process B and an $\mathscr{F}_T$-measurable claim $H \in L^2$ (maturing at time T) that we wish to hedge. However, we now only require that $X \in \mathcal{S}^2$, with canonical decomposition

$$X_t = X_0 + M_t + A_t, \tag{5.1}$$

for all $t \in [0, T]$, where $X_0 > 0$ is $\mathscr{F}_0$-measurable, $M \in \mathcal{M}_0^2$ and $A \in \mathcal{V}_0$. Furthermore, we assume that X satisfies the structure condition (Definition 3.23), so that

$$A_t = \int_0^t \alpha_s \, d\langle M\rangle_s, \tag{5.2}$$

for all $t \in [0, T]$, for some predictable process α. In other words, $A \ll \langle M\rangle$, from which it follows that α can be written as the pathwise Radon-Nikodým derivative

$$\alpha = \frac{dA}{d\langle M\rangle}.$$

With an extra condition on $\alpha\Delta M$, (5.2) also ensures that an equivalent martingale measure exists for X, by Theorem 3.24.

We now define a process that is related to the structure condition and is used in a number of theorems.

Definition 5.1. The *mean-variance tradeoff (MVT) process*, denoted by $\widehat{K}$, is defined by

$$\widehat{K}_t := \int_0^t \alpha_s^2 \, d\langle M\rangle_s,$$

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

Since $\mathbb{P}$ is no longer a martingale measure for X, we must revisit the concept of a feasible strategy, to ensure integrability with respect to the semimartingale X.

Definition 5.2. Let Θ denote the family of predictable processes ϕ, satisfying $G(\phi) \in \mathcal{S}^2$.

Definition 5.3. A strategy (ξ, η) is called *feasible* if $\xi \in \Theta$; $V(\xi, \eta)$ is right continuous with $V_t(\xi, \eta) \in L^2$, for all $t \in [0, T]$; and $V_T(\xi, \eta) = H$ a.s.

Definition 5.3 is a straightforward refinement of Definition 4.1. Note that, for a feasible strategy (ξ, η), this definition in combination with (4.2) implies that $C_t(\xi, \eta) \in L^2$, for all $t \in [0, T]$. Under an additional assumption that $C(\xi, \eta)$ is a martingale, the second part of Theorem 2.21 ensures that $C(\xi, \eta) \in \mathcal{M}^2$.

5.2 Local Risk Minimization

In the previous chapter we considered ways of hedging an L^2 claim in a market incorporating unhedgeable risk, when the price process X was a martingale. This relied on the GKW decomposition of the claim. This decomposition was shown to be related to the risk-minimizing strategy. Since X is now only a semimartingale, it is impossible to apply the GKW decomposition directly. Furthermore, the notion of risk minimization is also incompatible with the current setup. As a result, we introduce a new local risk function, called a risk quotient, based on a variational approach.

Given a feasible strategy, we may perturb it and inquire whether this increases or decreases the risk quotient. A locally risk-minimizing strategy is one for which the risk quotient increases or stays the same under all perturbations. A theorem relates such an optimal strategy to an orthogonality property of its associated cost process. This in turn yields a certain decomposition of the claim itself. We start by defining the small perturbations of a feasible strategy that preserve its feasibility.

Definition 5.4. A strategy (δ, ε) is called a *small perturbation* if $\delta \in \Theta$; $V(\delta, \varepsilon)$ is right continuous with $V_t(\delta, \varepsilon) \in L^2$, for all $t \in [0, T]$; δ and $\int_0^T |\delta_s||dA|_s$ are bounded; and $\delta_T = \varepsilon_T = 0$.

Note that a small perturbation (δ, ε) is not feasible, since $V_T(\delta, \varepsilon) = 0$. However, for any feasible strategy (ξ, η), it follows that $(\xi + \delta, \eta + \varepsilon)$ is again a feasible strategy. Since $\int_0^T \delta_s \, dA_s$ represents the systematic part of trading gains from (δ, ε), the boundedness condition in Definition 5.4 ensures that these gains are limited. The condition $\delta_T = \varepsilon_T = 0$ also ensures that any restriction of (δ, ε) to a subinterval of $[0, T]$ is again a small perturbation.

Definition 5.5. The *R-quotient* of a feasible strategy (ξ, η) and a small perturbation (δ, ε) with respect to a partition $\pi = \{0 = t_0, t_1, \ldots, t_m, t_{m+1} = T\}$ of $[0, T]$, is the following process

$$r^{\pi}[\xi, \eta; \delta, \varepsilon] := \sum_{i=0}^{m} \frac{R_{t_i}(\xi + \delta \mathbb{1}_{]\!]t_i, t_{i+1}]\!]}, \eta + \varepsilon \mathbb{1}_{[\![t_i, t_{i+1}[\![}) - R_{t_i}(\xi, \eta)}{\mathbb{E}\left[\langle M \rangle_{t_{i+1}} - \langle M \rangle_{t_i} \,\middle|\, \mathscr{F}_{t_i}\right]} \mathbb{1}_{]\!]t_i, t_{i+1}]\!]}.$$

Note that the asymmetry of the time intervals in the restrictions of δ and ε reflects the fact that $\xi + \delta$ should be predictable, while $\eta + \varepsilon$ need only be adapted.

Definition 5.6. A feasible strategy (ξ, η) is called *locally risk-minimizing* if

$$\liminf_{n \to \infty} r^{\pi^n}[\xi, \eta; \delta, \varepsilon] \geq 0 \quad \mu_M\text{-a.e.},$$

for every small perturbation (δ, ε) and every sequence of partitions $(\pi^n)_{n \in \mathbb{N}}$ with the property $\lim_{n \to \infty} \|\pi^n\| = 0$.

We shall now state a result of Schweizer (see [85, Lem. 2.2, p. 351] and [91, Thm. 3.3, p. 14]), linking local risk minimization to the mean self-financing condition. We prepare the way for it with the following technical assumptions.

Assumption 5.7.

1. $\langle M \rangle$ is a.s. strictly increasing;

2. A is a.s. continuous; and

3. $\mathbb{E}\big[\widehat{K}_T\big] < \infty$.

Note that Assumption 5.7 and the assumption that X satisfies the structure condition are requirements for the result we now state.

Proposition 5.8. *Let (ξ, η) be a feasible trading strategy with associated cost process C. The following statements are equivalent:*

1. *(ξ, η) is locally risk-minimizing;*

2. *(ξ, η) is mean self-financing and the (square integrable[1]) martingale C is strongly orthogonal to M.*

Using the above result, it is now possible to show that the existence of a locally risk-minimizing strategy implies a certain representation property for the claim, called the Föllmer-Schweizer decomposition.

[1] See the discussion following Definition 5.3.

Definition 5.9. The claim H is said to admit a *Föllmer-Schweizer* (FS) *decomposition* if it can be expressed as

$$H = H_0 + \int_0^T \xi_s^H \, dX_s + L_T^H \quad \text{a.s.}, \tag{5.3}$$

where $H_0 \in \mathbb{R}$, $\xi^H \in \Theta$ and $L^H \in \mathcal{M}_0^2$ is strongly orthogonal to M.

The following result establishes the connection between the FS decomposition and the existence of locally risk-minimizing strategies (see Schweizer [91, Prop. 3.4, p. 15]).

Proposition 5.10. *There exists a locally risk-minimizing strategy iff H admits an FS decomposition.*

Proof. ($\Rightarrow$) Let (ξ, η) be a locally risk-minimizing strategy with associated cost process C and value process V. According to (4.2), we have

$$H = V_T = C_0 + \int_0^T \xi_s \, dX_s + (C_T - C_0).$$

By Proposition 5.8, $C - C_0 \in \mathcal{M}_0^2$ is strongly orthogonal to M.
($\Leftarrow$) Let H be represented by (5.3) and define the feasible strategy $(\widehat{\xi}, \widehat{\eta})$, by setting

$$(\widehat{\xi}_t, \widehat{\eta}_t) := (\xi_t^H, H_0 + G_t(\xi^H) - \xi_t^H X_t + L_t^H), \tag{5.4}$$

for all $t \in [0, T]$. Then by (3.1) and (4.2), the cost process $\widehat{C}$ of this strategy is given by

$$\widehat{C}_t = H_0 + L_t^H,$$

for all $t \in [0, T]$. This is a square integrable martingale strongly orthogonal to M. Consequently, by Proposition 5.8, $(\widehat{\xi}, \widehat{\eta})$ is locally risk-minimizing. ■

The importance of the representation (5.3) cannot be overstated. As we shall see in later chapters, in any application the first task is to provide a characterization of the claim in terms of it. Thereafter, the price of the claim and the optimal hedging strategy follow as a consequence of the results presented below.

5.3 An Optimality Condition

We now provide an alternative characterization of a locally risk-minimizing strategy, in the form of an optimality condition. To start with, let

$$H = \mathbb{E}\,[H] + \int_0^T \mu_s^H \, dM_s + N_T^H, \tag{5.5}$$

where $\mu^H \in \mathcal{L}^2(M)$ and $N^H \in \mathcal{M}_0^2$ is strongly orthogonal to M, be the GKW decomposition of H with respect to M. Next, for any feasible strategy (ξ, η), let

$$\int_0^T \xi_s \, dA_s = \mathbb{E}\left[\int_0^T \xi_s \, dA_s\right] + \int_0^T \mu_s^{\xi \cdot A} \, dM_s + N_T^{\xi \cdot A}, \tag{5.6}$$

where $\mu^{\xi \cdot A} \in \mathcal{L}^2(M)$ and $N^{\xi \cdot A} \in \mathcal{M}_0^2$ is strongly orthogonal to M, be the GKW decomposition of $\int_0^T \xi_s \, dA_s$ with respect to M. Consequently, by (4.2) and (5.1), the maturity value of the cost process C associated with (ξ, η) is

$$C_T = H - \int_0^T \xi_s \, dM_s - \int_0^T \xi_s \, dA_s. \tag{5.7}$$

Combining (5.5), (5.6) and (5.7) yields

$$C_T = C_0 + \int_0^T \left(\mu_s^H - \xi_s - \mu_s^{\xi \cdot A}\right) dM_s + L_T^H, \tag{5.8}$$

where

$$C_0 := \mathbb{E}\left[H\right] - \mathbb{E}\left[\int_0^T \xi_s \, dA_s\right] \quad \text{and} \quad L^H := N^H - N^{\xi \cdot A}.$$

A locally risk-minimizing strategy (ξ, η) can now be characterized in terms of an optimality condition involving the processes μ^H and $\mu^{\xi \cdot A}$, defined above (see Schweizer [85, Thm. 2.4, p. 353]).

Theorem 5.11. *A feasible strategy (ξ, η) is locally risk-minimizing iff it is mean self-financing and the optimality equation*

$$\mu^H - \xi - \mu^{\xi \cdot A} = 0 \quad \mu_M\text{-a.e.} \tag{5.9}$$

is satisfied.

Proof. $(\Rightarrow)$ Suppose (ξ, η) is locally risk-minimizing. Then by Theorem 5.8, C is a square integrable martingale strongly orthogonal to M. Applying the martingale property of C to (5.8) gives

$$C_t = \mathbb{E}\left[C_T \,|\, \mathscr{F}_t\right] = C_0 + \int_0^t \left(\mu_s^H - \xi_s - \mu_s^{\xi \cdot A}\right) dM_s + L_t^H, \tag{5.10}$$

for all $t \in [0, T]$. From this expression and the strong orthogonality of L^H and M, it follows that

$$\langle C, M \rangle_t = \int_0^t \left(\mu_s^H - \xi_s - \mu_s^{\xi \cdot A}\right) d\langle M \rangle_s, \tag{5.11}$$

for all $t \in [0, T]$. Since C and M are strongly orthogonal, however, we must have $\langle C, M \rangle = 0$. Applying Lemma 2.76 to (5.11) then yields (5.9).
$(\Leftarrow)$ Suppose (ξ, η) is mean self-financing and (5.9) holds. By assumption, the cost process C associated with (ξ, η) is a (square integrable) martingale; combining this fact with (5.8) yields (5.10) again. Substituting (5.9) into (5.10) gives $C = C_0 + L^H$, which is strongly orthogonal to M. So by Theorem 5.8, (ξ, η) is locally risk-minimizing. ■

5.4 The Minimal Martingale Measure

Since we do not have a complete market, the set of equivalent martingale measures for X has infinitely many elements. In this section we characterize a certain subset of this set of measures. Following Föllmer and Schweizer [32, 85], we also identify one of the measures in this subset with the additional feature that it preserves the martingale property of all martingales orthogonal to M. This is the so-called minimal martingale measure. We start with a formal definition of the set of equivalent martingale measures for X that will be considered here.

Definition 5.12. Let

$$\mathbf{P}^2_{\mathsf{e}}(X) := \left\{ \mathbb{Q} \in \mathbf{P}_{\mathsf{e}}(X) \,\middle|\, \frac{d\mathbb{Q}}{d\mathbb{P}} \in L^2 \right\}$$

denote the set of martingale measures for X equivalent to $\mathbb{P}$ and with an L^2-density.

We next describe what it means for an element of $\mathbf{P}^2_{\mathsf{e}}(X)$ to be a minimal martingale measure (see Föllmer and Schweizer [32]). Thereafter we make two structural assumptions that enable us not only to infer the existence and uniqueness of such a measure, but also to obtain a concrete description of its density process.

Definition 5.13. A probability measure $\mathbb{Q} \in \mathbf{P}^2_{\mathsf{e}}(X)$ is called a *minimal martingale measure* (*for* X) if every square integrable martingale strongly orthogonal to M is also a $\mathbb{Q}$-martingale; i.e.

$$L \in \mathcal{M}^2 \text{ and } \langle L, M \rangle = 0 \quad \Rightarrow \quad L \text{ is a } \mathbb{Q}\text{-martingale.} \tag{5.12}$$

Assumption 5.14. There exists a martingale $N \in \mathcal{M}^2_0$, strongly orthogonal to M in (5.1), so that

$$L^2 = \left\{ c + \int_0^T \mu_s \, dM_s + \int_0^T \nu_s \, dN_s \,\middle|\, c \in \mathbb{R},\, \mu \in \mathcal{L}^2(M),\, \nu \in \mathcal{L}^2(N) \right\}. \tag{5.13}$$

Note that this assumption ensures that $\{M, N\}$ possesses the predictable representation property for $\mathcal{M}^2_0$ and forms what Schweizer [85] calls a $\mathbb{P}$-basis for L^2. Furthermore, the strong orthogonality of M and N ensures that the representations of elements in L^2, given by (5.13), are unique. Assumption 5.14 also means any martingale $L \in \mathcal{M}^2$, strongly orthogonal to M, may be expressed as follows:

$$L_t = L_0 + \int_0^t \nu_s \, dN_s, \tag{5.14}$$

for all $t \in [0, T]$, where $\nu \in \mathcal{L}^2(N)$. To see this, use (5.13) to write

$$L_T = L_0 + \int_0^T \mu_s \, dM_s + \int_0^T \nu_s \, dN_s,$$

for some $\mu \in \mathcal{L}^2(M)$ and $\nu \in \mathcal{L}^2(N)$. The martingale property of L then gives

$$\begin{aligned} L_t &= \mathbb{E}\left[L_0 + \int_0^T \mu_s \, dM_s + \int_0^T \nu_s \, dN_s \,\middle|\, \mathscr{F}_t \right] \\ &= L_0 + \int_0^t \mu_s \, dM_s + \int_0^t \nu_s \, dN_s, \end{aligned}$$

for all $t \in [0,T]$ Computing the angle bracket covariation of both sides of this equation with respect to M, yields

$$\begin{aligned} \langle L, M \rangle_t &= \left\langle \int_0^\cdot \mu_s \, dM_s, M \right\rangle_t + \left\langle \int_0^\cdot \nu_s \, dN_s, M \right\rangle_t \\ &= \int_0^t \mu_s \, d\langle M \rangle_s + \int_0^t \nu_s \, d\langle M, N \rangle_s, \end{aligned}$$

for all $t \in [0,T]$. Since M and N are strongly orthogonal, the second integral above vanishes. According to Lemma 2.76, this implies that $\mu = 0$ μ_M-a.e., which in turn gives (5.14).

Assumption 5.15. There exists a probability measure $\widehat{\mathbb{P}} \in \mathbf{P}^2_{\mathsf{e}}(X)$ such that $X, N \in \mathcal{M}^2(\widehat{\mathbb{P}})$ are strongly orthogonal under $\widehat{\mathbb{P}}$ (where N is the same process in Assumption 5.14) and

$$L^2(\widehat{\mathbb{P}}) = \left\{ c + \int_0^T \widehat{\mu}_s \, dX_s + \int_0^T \widehat{\nu}_s \, dN_s \,\middle|\, c \in \mathbb{R}, \, \widehat{\mu} \in \mathcal{L}^2(X), \, \widehat{\nu} \in \mathcal{L}^2(N) \right\}.$$

Again, this assumption ensures that $\{X - X_0, N\}$ possesses the predictable representation property for $\mathcal{M}^2_0(\widehat{\mathbb{P}})$. It also implies that $\mathbf{P}^2_{\mathsf{e}}(X) \neq \emptyset$, thereby implicitly ensuring that the market is arbitrage-free.

The next theorem (which is similar to Schweizer [89, Thm. 1, p. 576]) characterizes the probability measures in $\mathbf{P}^2_{\mathsf{e}}(X)$. Before proving it, we take a moment formally to defuse the structure condition assumption (5.2). As we see in the statement of Theorem 5.16, the fact that X satisfies the structure condition is now a consequence of the two assumptions introduced in this section.

Theorem 5.16. *Let $\mathbb{Q} \in \mathbf{P}^2_{\mathsf{e}}(X)$. Then X satisfies the structure condition and the density process Z for $\mathbb{Q}$ can be expressed as*

$$Z_t := \mathbb{E}\left[\frac{d\mathbb{Q}}{d\mathbb{P}} \,\middle|\, \mathscr{F}_t \right] = \mathscr{E}\left(-\int_0^\cdot \alpha_s \, dM_s + \int_0^\cdot \nu_s \, dN_s \right)_t, \tag{5.15}$$

for all $t \in [0,T]$, where α and ν are predictable processes. Furthermore, α is unique μ_M-a.e.

Proof. Firstly, by Assumption 5.14, we know that

$$Z_T = 1 + \int_0^T \mu_s^Z \, dM_s + \int_0^T \nu_s^Z \, dN_s,$$

for some $\mu^Z \in \mathcal{L}^2(M)$ and $\nu^Z \in \mathcal{L}^2(N)$. By the definition of a density process,

$$Z_t = \mathbb{E}\left[Z_T \,|\, \mathscr{F}_t \right] = 1 + \int_0^t \mu_s^Z \, dM_s + \int_0^t \nu_s^Z \, dN_s, \tag{5.16}$$

for all $t \in [0, T]$. Secondly, according to Theorem 3.22 there exists a local martingale $Y \in \mathcal{M}_{0,\mathsf{loc}}$ satisfying the conditions $\mathbb{E}\left[\mathscr{E}(Y)_T\right] = 1$, $\Delta Y > -1$ and $A + \langle M, Y \rangle = 0$, such that $Z = \mathscr{E}(Y)$. It then follows from Corollary 2.56 that

$$Y = Y - Y_0 = \mathscr{L}(\mathscr{E}(Y)) = \mathscr{L}(Z)\,.$$

In other words, Y satisfies the SDE

$$Y_t = \mathscr{L}(Z)_t = \int_0^t Z_{s-}^{-1} \, dZ_s, \tag{5.17}$$

for all $t \in [0, T]$ (see Definition 2.54). Combining (5.16) and (5.17), using the associativity property of stochastic integrals (Theorem 2.41 (4)), gives

$$Y_t = \int_0^t Z_{s-}^{-1} \mu_s^Z \, dM_s + \int_0^t Z_{s-}^{-1} \nu_s^Z \, dN_s, \tag{5.18}$$

for all $t \in [0, T]$. Taking $\alpha := Z_-^{-1} \mu^Z$ and $\nu := Z_-^{-1} \nu^Z$, establishes the representation (5.15) for $Z = \mathscr{E}(Y)$.

Next, we use the relation $A + \langle M, Y \rangle = 0$, together with (5.18) and the strong orthogonality of M and N, to verify the structure condition, as follows:

$$\begin{aligned} A_t = -\langle M, Y \rangle_t &= -\left\langle M, \int_0^\cdot Z_{s-}^{-1} \mu_s^Z \, dM_s + \int_0^\cdot Z_{s-}^{-1} \nu_s^Z \, dN_s \right\rangle_t \\ &= -\int_0^t Z_{s-}^{-1} \mu_s^Z \, d\langle M \rangle_s - \int_0^t Z_{s-}^{-1} \nu_s^Z \, d\langle M, N \rangle_s \\ &= -\int_0^t Z_{s-}^{-1} \mu_s^Z \, d\langle M \rangle_s = \int_0^t \alpha_s \, d\langle M \rangle_s, \end{aligned}$$

for all $t \in [0, T]$. Finally, to see that α is unique, suppose there exists a predictable process $\bar{\alpha}$ satisfying

$$A_t = \int_0^t \bar{\alpha}_s \, d\langle M \rangle_s,$$

for all $t \in [0, T]$, as well. Then

$$\int_0^t (\alpha_s - \bar{\alpha}_s) \, d\langle M \rangle_s = 0,$$

for all $t \in [0, T]$, from which it follows that $\alpha = \bar{\alpha}$ μ_M-a.e., by Lemma 2.76. ∎

This theorem illustrates why there is no unique equivalent martingale measure for X in an incomplete market — the process ν in (5.15) is essentially a free parameter. We now use Assumption 5.15 and the representation (5.15) to construct the minimal martingale measure for X.

Theorem 5.17. *The density process $Z^{\widehat{\mathbb{P}}}$ of the probability measure $\widehat{\mathbb{P}}$ in Assumption 5.15 is given by*

$$Z_t^{\widehat{\mathbb{P}}} := \mathbb{E}\left[\frac{d\widehat{\mathbb{P}}}{d\mathbb{P}}\,\middle|\,\mathscr{F}_t\right] = \mathscr{E}\left(-\int_0^{\cdot} \alpha_s\, dM_s\right)_t, \tag{5.19}$$

for all $t \in [0,T]$. In particular $\widehat{\mathbb{P}}$ is unique. Furthermore, it is the minimal martingale measure.

Proof. By Theorem 5.16, $Z^{\widehat{\mathbb{P}}}$ has the representation (5.15) for some predictable process ν (recall that α in (5.15) is unique). According to Definition 2.52, $Z^{\widehat{\mathbb{P}}}$ is then the unique strong solution of the following SDE:

$$\begin{aligned} Z_t^{\widehat{\mathbb{P}}} &= 1 + \int_0^t Z_{s-}^{\widehat{\mathbb{P}}}\, d\left(-\int_0^{\cdot} \alpha_u\, dM_u + \int_0^{\cdot} \nu_u\, dN_u\right)_s \\ &= 1 - \int_0^t Z_{s-}^{\widehat{\mathbb{P}}} \alpha_s\, dM_s + \int_0^t Z_{s-}^{\widehat{\mathbb{P}}} \nu_s\, dN_s, \end{aligned} \tag{5.20}$$

for all $t \in [0,T]$. (The second equality above follows from the associativity property of stochastic integrals.) The strong orthogonality of M and N then gives

$$\begin{aligned} \left\langle Z^{\widehat{\mathbb{P}}}, N\right\rangle_t &= \left\langle -\int_0^{\cdot} Z_{s-}^{\widehat{\mathbb{P}}}\alpha_s\, dM_s, N\right\rangle_t + \left\langle \int_0^{\cdot} Z_{s-}^{\widehat{\mathbb{P}}}\nu_s\, dN_s, N\right\rangle_t \\ &= -\int_0^t Z_{s-}^{\widehat{\mathbb{P}}}\alpha_s\, d\langle M, N\rangle_s + \int_0^t Z_{s-}^{\widehat{\mathbb{P}}}\nu_s\, d\langle N\rangle_s \\ &= \int_0^t Z_{s-}^{\widehat{\mathbb{P}}}\nu_s\, d\langle N\rangle_s, \end{aligned} \tag{5.21}$$

for all $t \in [0,T]$. Now, by Assumption 5.15, N is a square integrable martingale under $\widehat{\mathbb{P}}$, which in turn implies that $Z^{\widehat{\mathbb{P}}}N$ is a martingale (see Corollary 2.62). In other words, $\mathbb{Z}^{\widehat{\mathbb{P}}}$ and N are strongly orthogonal, from which we conclude that (5.21) must be zero. By Lemma 2.76 and the fact that $Z^{\widehat{\mathbb{P}}} > 0$, this implies that $\nu = 0$ μ_N-a.e.; and so (5.19) follows.

To verify that $\widehat{\mathbb{P}}$ is in fact the minimal martingale measure, let $L \in \mathcal{M}^2$ satisfy $\langle L, M\rangle = 0$. Then there exists a process $\kappa \in \mathcal{L}^2(N)$ such that[2]

$$L_t = L_0 + \int_0^t \kappa_s\, dN_s,$$

[2] See the discussion after Assumption 5.14.

for all $t \in [0,T]$. Then, since $Z^{\widehat{\mathbb{P}}}$ satisfies (5.20) with $\nu = 0$, we have

$$\begin{aligned}
\left\langle L, Z^{\widehat{\mathbb{P}}} \right\rangle_t &= \left\langle \int_0^{\cdot} \kappa_s \, dN_s, -\int_0^{\cdot} Z_{s-}^{\widehat{\mathbb{P}}} \alpha_s \, dM_s \right\rangle_t \\
&= -\int_0^t Z_{s-}^{\widehat{\mathbb{P}}} \alpha_s \kappa_s \, d\langle M, N \rangle_s \\
&= 0,
\end{aligned}$$

for all $t \in [0,T]$, by the strong orthogonality of M and N. Thus L and $Z^{\widehat{\mathbb{P}}}$ are also strongly orthogonal; whence $LZ^{\widehat{\mathbb{P}}}$ is a martingale. By Corollary 2.62, L is thus a $\widehat{\mathbb{P}}$-martingale; which verifies that $\widehat{\mathbb{P}}$ is a minimal martingale measure, as described in Definition 5.13. ∎

5.5 The Optimality of Pricing under the Minimal Martingale Measure

Now that we have an optimality condition characterizing the locally risk-minimizing strategy and have constructed the minimal martingale measure, it is possible to express the locally risk-minimizing strategy explicitly. In this section we show how it can be obtained from the minimal martingale measure.

Since X is a $\widehat{\mathbb{P}}$-martingale, we may consider the GKW decomposition of the claim H in terms of X under $\widehat{\mathbb{P}}$. With some manipulation, a comparison with the GKW decomposition of H in terms of M under $\mathbb{P}$ verifies the optimality condition. Finally, this leads to the crucial observation that the FS decomposition of H under $\mathbb{P}$ corresponds with the GKW decomposition of H under $\widehat{\mathbb{P}}$.

Assumption 5.15 allows us to express the GKW decomposition of H with respect to $X - X_0$ under $\widehat{\mathbb{P}}$ as

$$\begin{aligned}
H &= \mathbb{E}^{\widehat{\mathbb{P}}}[H] + \int_0^T \widehat{\mu}_s^H \, d(X_s - X_0) + \int_0^T \widehat{\nu}_s \, dN_s \\
&= \mathbb{E}^{\widehat{\mathbb{P}}}[H] + \int_0^T \widehat{\mu}_s^H \, dX_s + \int_0^T \widehat{\nu}_s \, dN_s,
\end{aligned} \tag{5.22}$$

for some $\widehat{\mu}^H \in \mathcal{L}^2(X)$ and $\widehat{\nu} \in \mathcal{L}^2(N)$. We now show that this decomposition is related to the local risk-minimizing portfolio, by verifying the optimality condition (see Schweizer [85, Thm. 3.2, p. 357]).

Theorem 5.18. *The mean self-financing feasible strategy (ξ, η), with $\xi := \widehat{\mu}^H$, satisfies the optimality condition of Theorem 5.11.*

Proof. Using the canonical decomposition of X, we may write (5.22) as

$$H = \mathbb{E}^{\widehat{\mathbb{P}}}[H] + \int_0^T \xi_s \, dM_s + \int_0^T \xi_s \, dA_s + \int_0^T \widehat{\nu}_s \, dN_s. \tag{5.23}$$

Taking an expectation on both sides yields

$$\mathbb{E}[H] = \mathbb{E}^{\widehat{\mathbb{P}}}[H] + \mathbb{E}\left[\int_0^T \xi_s \, dA_s\right], \tag{5.24}$$

since M and N are martingales. Next, we expand the second integral in (5.23) in terms of its GKW decomposition with respect to M, to get

$$\int_0^T \xi_s \, dA_s = \mathbb{E}\left[\int_0^T \xi_s \, dA_s\right] + \int_0^T \mu_s^{\xi \cdot A} \, dM_s + L_T,$$

where $\mu^{\xi \cdot A} \in \mathcal{L}^2(M)$ and $L \in \mathcal{M}_0^2$ is strongly orthogonal to M. According to the discussion following Assumption 5.14, there exists a process $\nu \in \mathcal{L}^2(N)$ satisfying (5.14). Thus

$$\int_0^T \xi_s \, dA_s = \mathbb{E}\left[\int_0^T \xi_s \, dA_s\right] + \int_0^T \mu_s^{\xi \cdot A} \, dM_s + \int_0^T \nu_s \, dN_s. \tag{5.25}$$

Substituting (5.24) and (5.25) into (5.23) yields

$$H = \mathbb{E}[H] + \int_0^T \left(\xi_s + \mu_s^{\xi \cdot A}\right) dM_s + \int_0^T (\widehat{\nu}_s + \nu_s) \, dN_s.$$

Comparing this expression with the GKW decomposition of H in (5.5) allows us conclude that

$$\mu^H = \xi + \mu^{\xi^H \cdot A} \quad \mu_M\text{-a.e.},$$

by the uniqueness of the GKW decomposition. This verifies the optimality condition (5.9). ■

We shall now demonstrate that the FS decomposition (5.3) corresponds with the representation (5.22). Taking expectations under $\widehat{\mathbb{P}}$ on both sides of (5.3) gives

$$\mathbb{E}^{\widehat{\mathbb{P}}}[H] = \mathbb{E}^{\widehat{\mathbb{P}}}\left[H_0 + \int_0^T \xi_s^H \, dX_s + L_T^H\right] = H_0,$$

since X is by definition a $\widehat{\mathbb{P}}$-martingale and the fact that $L^H \in \mathcal{M}_0^2$ is strongly orthogonal to M implies that L^H is a $\widehat{\mathbb{P}}$-martingale, by (5.12). Also, according to the discussion following Assumption 5.14, L^H may be expressed as (5.14), for some process $\nu \in \mathcal{L}^2(N)$. Since X and N are strongly orthogonal under $\widehat{\mathbb{P}}$, by Assumption 5.15, we then obtain

$$\begin{aligned}
\langle L^H, X\rangle_t^{\widehat{\mathbb{P}}} &= \left\langle \int_0^{\cdot} \nu_s \, dN_s, X \right\rangle_t^{\widehat{\mathbb{P}}} \\
&= \int_0^t \nu_s \, \langle N, X\rangle_s^{\widehat{\mathbb{P}}} \\
&= 0,
\end{aligned}$$

for all $t \in [0,T]$. In other words, X and L^H are strongly orthogonal under $\widehat{\mathbb{P}}$. This means that the FS decomposition (5.3), which can be rewritten as

$$H = \mathbb{E}^{\widehat{\mathbb{P}}}[H] + \int_0^T \xi_s^H \, d(X_s - X_0) + \int_0^T \nu_s \, dN_s,$$

is nothing other then the GKW decomposition (5.22) under $\widehat{\mathbb{P}}$ (by the uniqueness of the GKW decomposition). The uniqueness of the GKW decomposition also allows us to conclude that the FS decomposition is unique and that $\xi^H = \widehat{\mu}^H$.

5.6 The Hedge Ratio

In a manner analogous to the previous chapter, we now define the hedge ratio for the locally risk-minimizing portfolio in terms of the intrinsic value of the claim.

Definition 5.19. The *intrinsic value* of H is the process $\widehat{V}$, defined by

$$\widehat{V}_t := \mathbb{E}^{\widehat{\mathbb{P}}}[H \mid \mathscr{F}_t] = H_0 + G_t(\xi^H) + L_t^H, \tag{5.26}$$

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

Using the intrinsic value process and (5.4), it is now possible to express the locally risk-minimizing portfolio $(\widehat{\xi}, \widehat{\eta})$ as

$$(\widehat{\xi}_t, \widehat{\eta}_t) := (\xi_t^H, \widehat{V}_t - \xi_t^H X_t),$$

for all $t \in [0,T]$. Taking the angle brackets covariation under $\widehat{\mathbb{P}}$ with respect to X on both sides of (5.26) gives

$$\langle \widehat{V}, X \rangle_t^{\widehat{\mathbb{P}}} = \left\langle \int_0^{\cdot} \widehat{\xi}_s \, dX_s, X \right\rangle_t^{\widehat{\mathbb{P}}} + \langle L^H, X \rangle_t^{\widehat{\mathbb{P}}} = \int_0^t \widehat{\xi}_s \, d\langle X \rangle_s^{\widehat{\mathbb{P}}},$$

for all $t \in [0,T]$. Consequently, we may express ξ^H as the pathwise Radon-Nikodým derivative of the Lebesgue-Stieltjes measure induced by $\langle \widehat{V}, X \rangle^{\widehat{\mathbb{P}}}$ with respect to the Lebesgue-Stieltjes measure induced by $\langle X \rangle^{\widehat{\mathbb{P}}}$; i.e.

$$\widehat{\xi} = \frac{d\langle \widehat{V}, X \rangle^{\widehat{\mathbb{P}}}}{d\langle X \rangle^{\widehat{\mathbb{P}}}}.$$

This expression looks similar to the hedge ratio derived at the end of Chapter 4. Its form lends itself to PDE representations, which we explore in Chapter 7.

Chapter 6

Incomplete Markets: Mean-Variance Optimization

The quadratic approach in the previous chapter employed a measure of local risk. Essentially this entailed the correction of hedge errors at each instant by minimizing the conditional variance of the cost process and relaxing the self-financing condition. In contrast, we now focus on quadratic measures of the global risk. While insisting on the self-financing condition, the goal is to find strategies that minimize the variance of the difference between the final value of the hedge portfolio and the terminal value of the claim.

Given an arbitrary fixed initial amount with which to hedge the claim, the first theorem provides a recursive characterization of the self-financing strategy that minimizes the variance of the terminal hedging error. Its proof relies on a certain condition being imposed on the market. In particular, a process, called the extended mean-variance tradeoff process, is defined. The optimal strategy is determined under the rather restrictive condition that this process is deterministic.

A natural extension of the above problem is to find the combination of initial endowment and strategy that minimize the variance of the hedging error. The optimal initial endowment — which is no longer pre-determined, but rather part of the solution to the problem — is called the approximation price of the claim. We establish the link between this problem and the so-called variance-optimal martingale measure, for which a definition is provided. In particular, the approximation price is determined by computing the expected value of the claim under the variance-optimal martingale measure. Furthermore, subject to the assumption of a deterministic mean-variance tradeoff process, the variance-optimal martingale measure and the (possibly signed) minimal martingale measure coincide. When this is the case, the optimal strategy is the recursively generated strategy of the previous result.

Finally, we cite some results showing that the condition of a deterministic mean-variance tradeoff process can be relaxed, if we assume that the price process is continuous.

6.1 Market Assumptions

In this section we specify an incomplete market very similar to the market prescribed in the previous chapter (but employing slightly different technical assumptions). In order to make the current chapter self-contained, we risk being repetitious, by recalling the setup in Section 5.1.

As before, we assume a fixed finite time-horizon $T \in (0, \infty)$ and a stochastic basis $(\Omega, \mathscr{F}, \mathbf{F}, \mathbb{P})$. The filtration $\mathbf{F} = (\mathscr{F}_t)_{t\in[0,T]}$ satisfies the usual conditions, with $\mathscr{F}_T = \mathscr{F}$ and $\mathscr{F}_0$ trivial. Let X be a special semimartingale with canonical decomposition

$$X_t = X_0 + M_t + A_t, \tag{6.1}$$

for all $t \in [0, T]$, where $X_0 > 0$ is $\mathscr{F}_0$-measurable, $M \in \mathcal{M}_0^2$ and $A \in \mathcal{V}_0$. Again, we regard X as describing the discounted value of a risky asset. Finally, let $H \in L^2$ be the $\mathscr{F}_T$-measurable claim that we wish to hedge.

In the previous two chapters we employed feasible strategies. Here we revert to self-financing strategies and accept the potential shortfall or excess in the value of the hedge portfolio at maturity. Since a feasible strategy can be constructed from a self-financing strategy by making up the hedging error at maturity, we do not consider feasibility any further. However, we shall require that $\xi \in \Theta$ (see Definition 5.2), for any self-financing strategy under consideration. In particular, this means that

$$\mathbb{E}\left[\int_0^T \xi_s^2 \, d\langle M\rangle_s + \left(\int_0^T |\xi_s| \, |dA|_s\right)^2\right] < \infty; \tag{6.2}$$

a condition that will be verified for the optimal strategy.

As in the previous chapter, we impose the structure condition on X. This means that

$$A_t = \int_0^t \alpha_s \, d\langle M\rangle_s, \tag{6.3}$$

for all $t \in [0, T]$, where α is some predictable process. Furthermore, we assume that the claim admits an FS decomposition. Formally, this is expressed as follows:

Assumption 6.1. The claim H admits an FS decomposition, so that

$$H = H_0 + \int_0^T \xi_s^H \, dX_s + L_T,$$

where $H_0 \in \mathbb{R}$, $\xi^H \in \Theta$ and $L \in \mathcal{M}_0^2$ is strongly orthogonal to M.

Note that, in contrast to the previous chapter, we no longer impose continuity on A. As a result we extend the idea of a MVT process (see Definition 5.1) as follows:

Definition 6.2. The *extended mean-variance tradeoff* ($EMVT$) *process*, denoted by $\widetilde{K}$, is defined by

$$\widetilde{K}_t := \int_0^t \frac{\alpha_s^2}{1+\alpha_s^2 \Delta\langle M\rangle_s}\, d\langle M\rangle_s,$$

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

It should be noted that when A (or X) has continuous paths, then by (6.3) we have $\alpha\Delta\langle M\rangle = 0$. In this case the EMVT process coincides with the MVT process of Definition 5.1.

6.2 Mean-Variance Hedging

Given an initial endowment with which to hedge the claim, it is reasonable to enquire as to which self-financing strategy offers the best hedging performance, in a quadratic sense. This question may be posed formally as follows:

$$\text{Given } c \in \mathbb{R}, \text{ minimize } \mathbb{E}\left[(H - c - G_T(\xi))^2\right], \text{ over all } \xi \in \Theta. \tag{6.4}$$

This problem was originally considered by Duffie and Richardson [27] and Schweizer [87], in the case where X is a geometric Brownian motion. Later Schweizer [88] (whose account we follow closely) and Monat and Stricker [67] extended the analysis to a general semimartingale framework.

According to the next result, problem (6.4) can be solved if we impose a rather strong condition on the EMVT process. The resulting optimal stock-holding strategy is specified by a recursive relation (see Schweizer [88, Thm. 3, p. 1543]).

Theorem 6.3. *Suppose $\widetilde{K}$ is deterministic. Then problem* (6.4) *has a solution $\xi^{(c)} \in \Theta$, determined in feedback form by*

$$\xi_t^{(c)} = \xi_t^H + \widetilde{\alpha}_t\Big(\widehat{V}_{t-} - c - G_{t-}\left(\xi^{(c)}\right)\Big), \tag{6.5}$$

for all $t \in [0,T]$. Here $\widetilde{\alpha}$ is defined by

$$\widetilde{\alpha}_t := \frac{\alpha_t}{1+\alpha_t^2 \Delta\langle M\rangle_t}, \tag{6.6}$$

for all $t \in [0,T]$ and $\widehat{V}$ is the intrinsic value of the claim, given by

$$\widehat{V}_t := H_0 + \int_0^t \xi_s^H\, dX_s + L_t^H,$$

for all $t \in [0,T]$ (see also Definition 5.19).

Before proving this theorem, we should check that (6.5) determines a well-defined strategy, in the sense that $\xi^{(c)} \in \Theta$. This is the subject of the next lemma; the proof of which relies on the repeated use of four inequalities which we now provide.

Note that since $\alpha^2 \Delta\langle M\rangle \geq 0$, it follows from (6.6) that $|\widetilde{\alpha}| \leq |\alpha|$. Consequently, for any predictable process ϕ, we have

$$\begin{aligned}
\int_0^t \widetilde{\alpha}_s^2 \phi_s^2 \, d\langle M\rangle_s &\leq \int_0^t |\widetilde{\alpha}_s||\alpha_s|\phi_s^2 \, d\langle M\rangle_s \\
&= \int_0^t \widetilde{\alpha}_s \alpha_s \phi_s^2 \, d\langle M\rangle_s = \int_0^t \phi_s^2 \, d\widetilde{K}_s \qquad (6.7)\\
&\leq \Big(\sup_{s\in[0,T]} \phi_s^2\Big) \int_0^T d\widetilde{K}_s = \Big(\sup_{s\in[0,T]} \phi_s^2\Big) \widetilde{K}_T, \qquad (6.8)
\end{aligned}$$

for $t \in [0,T]$, where the last inequality follows since $\widetilde{K}$ is increasing. We also have

$$\begin{aligned}
\Big(\int_0^t \widetilde{\alpha}_s \phi_s \, dA_s\Big)^2 &\leq \Big(\int_0^t |\widetilde{\alpha}_s \phi_s| \, |dA|_s\Big)^2 \\
&= \Big(\int_0^t \widetilde{\alpha}_s \alpha_s |\phi_s| \, d\langle M\rangle_s\Big)^2 = \Big(\int_0^t |\phi_s| \, d\widetilde{K}_s\Big)^2 \\
&\leq \Big(\int_0^t |\phi_s|^2 \, d\widetilde{K}_s\Big)\Big(\int_0^t d\widetilde{K}_s\Big) \leq \widetilde{K}_T \int_0^t \phi_s^2 \, d\widetilde{K}_s \qquad (6.9)\\
&\leq \widetilde{K}_T \Big(\sup_{s\in[0,T]} \phi_s^2\Big) \int_0^T d\widetilde{K}_s = \Big(\sup_{s\in[0,T]} \phi_s^2\Big) \widetilde{K}_T^2 \qquad (6.10)
\end{aligned}$$

for all $t \in [0,T]$, where the second inequality is an instance of the Cauchy-Schwartz inequality. Note that the right-hand sides of the inequalities are not necessarily finite — however, we can ensure that they are finite by choosing ϕ appropriately.

Lemma 6.4. *If $\widetilde{K}$ is deterministic, then there exists a $\xi^{(c)} \in \Theta$ such that*

$$\xi_t^{(c)} = \xi_t^H + \widetilde{\alpha}_t\Big(\widehat{V}_{t-} - c - G_{t-}\big(\xi^{(c)}\big)\Big),$$

for all $t \in [0,T]$ (with equality μ_M-a.e.).

Proof. Since $\widetilde{K}_T < \infty$, inequalities (6.8) and (6.10), with $\phi := 1$, imply that

$$\mathbb{E}\left[\int_0^T \widetilde{\alpha}_s^2 \, d\langle M\rangle_s + \Big(\int_0^T |\widetilde{\alpha}_s||dA|_s\Big)^2\right] < \infty.$$

Thus $\widetilde{\alpha} \in \Theta$. Now define the processes Z and Y, by setting

$$Z_t := -\int_0^t \widetilde{\alpha}_s \, dX_s; \quad \text{and} \qquad (6.11)$$

$$Y_t := \int_0^t \Big(\xi_s^H + \widetilde{\alpha}_s\big(\widehat{V}_{s-} - c\big)\Big) \, dX_s, \qquad (6.12)$$

for all $t \in [0,T]$. Since $\widetilde{\alpha} \in \Theta$, we have $\sup_{t\in[0,T]} \mathbb{E}\left[Z_t^2\right] < \infty$. We now verify that $\sup_{t\in[0,T]} \mathbb{E}\left[Y_t^2\right] < \infty$. Iterated applications of the inequality

$$(a+b)^2 \le (a+b)^2 + (a-b)^2 = 2a^2 + 2b^2 \tag{6.13}$$

to (6.12) give

$$Y_t^2 \le 2\left(\int_0^t \xi_s^H \, dX_s\right)^2 + 4\left(\int_0^t \widetilde{\alpha}_s\big(\widehat{V}_{s-} - c\big)\, dM_s\right)^2 + 4\left(\int_0^t \widetilde{\alpha}_s\big(\widehat{V}_{s-} - c\big)\, dA_s\right)^2,$$

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

$$\begin{aligned}
\sup_{t\in[0,T]} \mathbb{E}\left[Y_t^2\right] &\le 2 \sup_{t\in[0,T]} \mathbb{E}\left[\left(\int_0^t \xi_s^H \, dX_s\right)^2\right] + 4 \sup_{t\in[0,T]} \mathbb{E}\left[\int_0^t \widetilde{\alpha}_s^2\big(\widehat{V}_{s-} - c\big)^2 \, d\langle M\rangle_s\right] \\
&\quad + 4 \sup_{t\in[0,T]} \mathbb{E}\left[\left(\int_0^t \widetilde{\alpha}_s\big(\widehat{V}_{s-} - c\big)\, dA_s\right)^2\right] \\
&\le 2\mathbb{E}\left[\sup_{t\in[0,T]} \left(\int_0^t \xi_s^H \, dX_s\right)^2\right] + 4\big(\widetilde{K}_T + \widetilde{K}_T^2\big)\mathbb{E}\left[\sup_{t\in[0,T]} \big(\widehat{V}_t - c\big)^2\right],
\end{aligned}$$

where the second inequality follows from (6.8) and (6.10). Since $\xi^H \in \Theta$, $\widetilde{K}_T < \infty$ and $\sup_{t\in[0,T]} \big|\widehat{V}_t - c\big| \in L^2$, we conclude that

$$\sup_{t\in[0,T]} \mathbb{E}\left[Y_t^2\right] < \infty. \tag{6.14}$$

Now, the process U, given by

$$U_t := Y_t + \int_0^t U_{s-}\, dZ_s, \tag{6.15}$$

for all $t \in [0,T]$, has a unique strong solution which is a semimartingale, according to Protter [74, Thm. V.7, p. 253]. Once again, by iterated applications of (6.13), we obtain

$$\begin{aligned}
\mathbb{E}\left[U_{t-}^2\right] &\le 2\mathbb{E}\left[Y_{t-}^2\right] + 4\mathbb{E}\left[\int_0^{t-} \widetilde{\alpha}_s^2 U_{s-}^2 \, d\langle M\rangle_s\right] + 4\mathbb{E}\left[\left(\int_0^{t-} \widetilde{\alpha}_s U_{s-}\, dA_s\right)^2\right] \\
&\le 2\mathbb{E}\left[Y_{t-}^2\right] + 4\mathbb{E}\left[\int_0^{t-} U_{s-}^2 d\widetilde{K}_s\right] + 4\widetilde{K}_T\mathbb{E}\left[\int_0^{t-} U_{s-}^2 d\widetilde{K}_s\right] \\
&\le 2\mathbb{E}\left[Y_{t-}^2\right] + 4\big(1 + \widetilde{K}_T\big)\int_0^t \mathbb{E}\left[U_{s-}^2\right] d\widetilde{K}_s \\
&\le 2\exp\Big(4\big(1 + \widetilde{K}_T\big)\widetilde{K}_T\Big) \sup_{s\in[0,t]} \mathbb{E}\left[Y_{s-}^2\right] < \infty,
\end{aligned} \tag{6.16}$$

for all $t \in [0,T]$. The second inequality above follows from (6.9); the third inequality follows by an application of Fubini's theorem and the fact that $\widetilde{K}$ is deterministic;

the penultimate inequality is an application of Gronwall's inequality; and finally the last inequality follows by (6.14).

Next, define the predictable process ϑ by setting

$$\vartheta_t := \widetilde{\alpha}_t\big(\widehat{V}_{t-} - c - U_{t-}\big),$$

for all $t \in [0,T]$. Then, another application of (6.13) yields

$$\begin{aligned}
\mathbb{E}\left[\int_0^T \vartheta_s^2\, d\langle M\rangle_s\right] &\leq 2\mathbb{E}\left[\int_0^T \widetilde{\alpha}_s^2\big(\widehat{V}_{s-} - c\big)^2\, d\langle M\rangle_s + \int_0^T \widetilde{\alpha}_s^2 U_{s-}^2\, d\langle M\rangle_s\right] \\
&\leq 2\widetilde{K}_T\mathbb{E}\left[\sup_{t\in[0,T]} \big(\widehat{V}_t - c\big)^2\right] + 2\int_0^T \mathbb{E}\left[U_{s-}^2\right]\, d\widetilde{K}_s \\
&\leq 2\widetilde{K}_T\mathbb{E}\left[\sup_{t\in[0,T]} \big(\widehat{V}_t - c\big)^2\right] + 2\widetilde{K}_T \sup_{t\in[0,T]} \mathbb{E}\left[U_{t-}^2\right] < \infty.
\end{aligned}$$

Here the second inequality follows from (6.7), (6.8) and Fubini's theorem; while the final inequality follows from (6.16) and the fact that $\sup_{t\in[0,T]}\big|\widehat{V}_t - c\big| \in L^2$. Similarly, by the triangle inequality and (6.13), it follows that

$$\begin{aligned}
\mathbb{E}\left[\left(\int_0^T |\vartheta_s|\,|dA|_s\right)^2\right] &\leq 2\mathbb{E}\left[\left(\int_0^T \big|\widetilde{\alpha}_s\big(\widehat{V}_{s-} - c\big)\big|\,|dA|_s\right)^2 + \left(\int_0^T |\widetilde{\alpha}_s U_{s-}|\,|dA|_s\right)^2\right] \\
&\leq 2\widetilde{K}_T^2\mathbb{E}\left[\sup_{t\in[0,T]} \big(\widehat{V}_t - c\big)^2\right] + 2\widetilde{K}_T \int_0^T \mathbb{E}\left[U_{s-}^2\right]\, d\widetilde{K}_s \\
&\leq 2\widetilde{K}_T^2\mathbb{E}\left[\sup_{t\in[0,T]} \big(\widehat{V}_t - c\big)^2\right] + 2\widetilde{K}_T^2 \sup_{t\in[0,T]} \mathbb{E}\left[U_{t-}^2\right] < \infty.
\end{aligned}$$

Again, the second inequality follows from (6.9), (6.10) and Fubini's theorem; while the final inequality follows from (6.16) and the fact that $\sup_{t\in[0,T]}\big|\widehat{V}_t - c\big| \in L^2$. Consequently, $\vartheta \in \Theta$, by (6.2).

Now define the strategy $\xi^{(c)} \in \Theta$, by setting

$$\xi_t^{(c)} := \xi_t^H + \vartheta_t = \xi_t^H + \widetilde{\alpha}_t\big(\widehat{V}_{t-} - c - U_{t-}\big),$$

for all $t \in [0,T]$. Then by (6.11), (6.12) and (6.15) we have

$$G_t\big(\xi^{(c)}\big) = Y_t + \int_0^t U_{s-}\, dZ_s = U_t \quad \text{a.s.},$$

for all $t \in [0,T]$. Therefore, $G\big(\xi^{(c)}\big)$ satisfies

$$\begin{aligned}
G_t\big(\xi^{(c)}\big) &= Y_t + \int_0^t G_{s-}\big(\xi^{(c)}\big)\, dZ_s \\
&= G_t\big(\xi^H\big) + \int_0^t \widetilde{\alpha}_s\Big(\widehat{V}_{s-} - c - G_{s-}\big(\xi^{(c)}\big)\Big)dX_s,
\end{aligned}$$

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

$$\int_0^t \left(\xi_s^{(c)} - \xi_s^H - \widetilde{\alpha}_s\left(\widehat{V}_{s-} - c - G_{s-}(\xi^{(c)})\right)\right) dX_s = 0$$

for all $t \in [0,T]$. The result now follows from the uniqueness of the stochastic integral. ■

In order to simplify the proof Theorem 6.3, we provide the following lemma.

Lemma 6.5. *Let $\varphi, \vartheta \in \Theta$. Then*

$$\mathbb{E}\big[[G(\varphi), G(\vartheta)]_t\big] = \mathbb{E}\left[\int_0^t \varphi_s\vartheta_s(1+\alpha_s^2\Delta\langle M\rangle_s)\, d\langle M\rangle_s\right],$$

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

Proof. Using the bilinearity of the covariation process and the decomposition of X in (6.1), we have

$$\begin{aligned}[G(\varphi), G(\vartheta)]_t = &\left[\int_0^\cdot \varphi_s\, dM_s, \int_0^\cdot \vartheta_s\, dM_s\right]_t + \left[\int_0^\cdot \varphi_s\, dA_s, \int_0^\cdot \vartheta_s\, dA_s\right]_t \\ &+ \left[\int_0^\cdot \varphi_s\, dM_s, \int_0^\cdot \vartheta_s\, dA_s\right]_t + \left[\int_0^\cdot \varphi_s\, dA_s, \int_0^\cdot \vartheta_s\, dM_s\right]_t\end{aligned} \tag{6.17}$$

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

$$\begin{aligned}\left[\int_0^\cdot \varphi_s\, dM_s, \int_0^\cdot \vartheta_s\, dM_s\right]_t = &\left(\left[\int_0^\cdot \varphi_s\, dM_s, \int_0^\cdot \vartheta_s\, dM_s\right]_t\right. \\ &\left. - \left\langle\int_0^\cdot \varphi_s\, dM_s, \int_0^\cdot \vartheta_s\, dM_s\right\rangle_t\right) + \int_0^t \varphi_s\vartheta_s\, d\langle M\rangle_s,\end{aligned} \tag{6.18}$$

for all $t \in [0,T]$. Next we note that

$$\begin{aligned}\left[\int_0^\cdot \varphi_s\, dA_s, \int_0^\cdot \vartheta_s\, dA_s\right]_t &= \sum_{0<s\le t} \Delta\left(\int_0^\cdot \varphi_s\, dA_s\right)_t \Delta\left(\int_0^\cdot \vartheta_s\, dA\right)_t \\ &= \sum_{0<s\le t} \varphi_s\vartheta_s(\Delta A_s)^2 \\ &= \sum_{0<s\le t} \varphi_s\vartheta_s\alpha_s^2(\Delta\langle M\rangle_s)^2 \\ &= \int_0^t \varphi_s\vartheta_s\alpha_s^2\Delta\langle M\rangle_s\, d\langle M\rangle_s,\end{aligned} \tag{6.19}$$

for all $t \in [0,T]$. Note that the term in brackets in (6.18) is a martingale by Proposition 2.49 (see also the discussion thereafter) while the last two terms in

(6.17) are martingales by Lemma 2.47. Finally, substituting (6.18) and (6.19) into (6.17) we get

$$\mathbb{E}\big[[G(\varphi), G(\vartheta)]_t\big] = \mathbb{E}\left[\int_0^t \varphi_s \vartheta_s (1 + \alpha_s^2 \Delta\langle M\rangle_s)\, d\langle M\rangle_s\right],$$

for all $t \in [0, T]$. ■

Proof of Theorem 6.3. In order to prove optimality, we show that

$$\mathbb{E}\left[\Big(H - c - G_T\big(\xi^{(c)}\big)\Big)\, G_T(\vartheta)\right] = 0, \tag{6.20}$$

for all $\vartheta \in \Theta$ and then appeal to the Hilbert projection theorem (see Luenberger [62, Thms. 1&2, p. 50]). We start by defining the function $f : [0, T] \to \mathbb{R}$, by setting

$$f(t) := \mathbb{E}\left[\Big(\widehat{V}_t - c - G_t\big(\xi^{(c)}\big)\Big)\, G_t(\vartheta)\right],$$

for all $t \in [0, T]$ (note that $f(T)$ is equal to the left-hand side of (6.20)). Using the stochastic integration by parts rule we obtain

$$\begin{aligned}
d\Big\{\Big(\widehat{V}_t - c - G_t\big(\xi^{(c)}\big)\Big) G_t(\vartheta)\Big\}&\\
&= \Big(\widehat{V}_{t-} - c - G_{t-}\big(\xi^{(c)}\big)\Big)\, dG_t(\vartheta) + G_{t-}(\vartheta)\, d\widehat{V}_t - G_{t-}(\vartheta)\, dG_t\big(\xi^{(c)}\big)\\
&\quad + d\Big[H_0 + G\big(\xi^H\big) + L - c - G\big(\xi^{(c)}\big), G(\vartheta)\Big]_t\\
&= \Big(\widehat{V}_{t-} - c - G_{t-}\big(\xi^{(c)}\big)\Big)\vartheta_t\, dX_t + \big(\xi_t^H - \xi_t^{(c)}\big) G_{t-}(\vartheta)\, dX_t\\
&\quad + G_{t-}(\vartheta)\, dL_t + d\Big[G\big(\xi^H - \xi^{(c)}\big), G(\vartheta)\Big]_t,
\end{aligned}$$

for all $t \in [0, T]$. Integrating, using the decomposition of X in (6.1), taking expectations and using Lemma 6.5, yields

$$\begin{aligned}
f(t) = \mathbb{E}\bigg[&\int_0^t \Big(\widehat{V}_{s-} - c - G_{s-}\big(\xi^{(c)}\big)\Big)\vartheta_s\, dA_s + \int_0^t \big(\xi_s^H - \xi_s^{(c)}\big) G_{s-}(\vartheta)\, dA_s\\
&+ \int_0^t \big(\xi_s^H - \xi_s^{(c)}\big)\vartheta_s (1 + \alpha_s^2 \Delta\langle M\rangle_s)\, d\langle M\rangle_s\bigg],
\end{aligned}$$

for all $t \in [0, T]$. Substituting the expression for $\xi^{(c)}$ in (6.5) into the last two terms of the above equation gives

$$\begin{aligned}
f(t) = \mathbb{E}\bigg[&\int_0^t \Big(\widehat{V}_{s-} - c - G_{s-}\big(\xi^{(c)}\big)\Big)\vartheta_s \alpha_s\, d\langle M\rangle_s\\
&- \int_0^t \Big(\widehat{V}_{s-} - c - G_{s-}\big(\xi^{(c)}\big)\Big) G_{s-}(\vartheta)\widetilde{\alpha}_s \alpha_s\, d\langle M\rangle_s\\
&- \int_0^t \Big(\widehat{V}_{s-} - c - G_{s-}\big(\xi^{(c)}\big)\Big)\vartheta_s \alpha_s\, d\langle M\rangle_s\bigg],
\end{aligned}$$

for all $t \in [0,T]$. The first and third terms cancel and the definition of $\widetilde{K}$ yields

$$\begin{aligned} f(t) &= \mathbb{E}\left[\int_0^t \left(\widehat{V}_{s-} - c - G_{s-}\big(\xi^{(c)}\big)\right) G_{s-}(\vartheta)\, d\widetilde{K}_s\right] \\ &= \int_0^t \mathbb{E}\left[\left(\widehat{V}_{s-} - c - G_{s-}\big(\xi^{(c)}\big)\right) G_{s-}(\vartheta)\right] d\widetilde{K}_s \\ &= \int_0^t f(s-)\, d\widetilde{K}_s, \end{aligned}$$

for all $t \in [0,T]$. Here the second equality follows from Fubini's theorem and the fact that $\widetilde{K}$ is deterministic. Consequently, since $f(0) = 0$, Gronwall's lemma [75, p. 543] indicates that $f(t) = 0$ for all $t \in [0,T]$. This verifies (6.20). ∎

6.3 The Variance-Optimal Martingale Measure

In the previous section we provided the strategy $\xi^{(c)}$ associated with a given initial endowment c. We now pose a related optimization problem: what combination of c and $\xi \in \Theta$ minimizes the variance of the hedging error at maturity? This section follows the structure of Schweizer [90] closely. Mathematically, the problem may be formulated as:

$$\text{Minimize } \mathbb{E}\left[(H - c - G_T(\xi))^2\right], \text{ over all } (c,\xi) \in \mathbb{R} \times \Theta. \tag{6.21}$$

Definition 6.6. If a solution $(v, \widetilde{\xi})$ to problem (6.21) exists, with $v \in \mathbb{R}$ and $\widetilde{\xi} \in \Theta$, then v is called the *approximation price* for H.

Note that if the EMVT process of X is deterministic and the approximation price of the claim is known, then $\widetilde{\xi} = \xi^{(v)}$ by Theorem 6.3.

As in Chapter 5, where we showed that local risk minimization corresponded to pricing the claim under the minimal martingale measure for X, we now show that a solution to the optimization problem (6.21) corresponds to pricing under the so-called variance-optimal martingale measure for X. In contrast to the previous chapter, we now consider signed martingale measures for X, for which we provide the following definition[1]:

Definition 6.7. A *signed L^2-martingale measure* for X is a signed measure on $(\Omega, \mathscr{F})$ satisfying $\mathbb{Q}\{\Omega\} = 1$, $\mathbb{Q} \ll \mathbb{P}$ with $\frac{d\mathbb{Q}}{d\mathbb{P}} \in L^2$ and

$$\mathbb{E}\left[\frac{d\mathbb{Q}}{d\mathbb{P}}(X_t - X_s) \,\middle|\, \mathscr{F}_s\right] = 0 \quad \text{a.s.},$$

[1] We have deviated slightly from the account of Schweizer in our definition of signed martingale measures. There the set of measures $\mathbf{P}_{\mathsf{s}}(\Theta)$ is used; but later it is shown that $\mathbf{P}_{\mathsf{s}}(\Theta) = \mathbf{P}^2_{\mathsf{s}}(X)$, if the structure condition holds and $X \in \mathcal{S}^2(\mathbb{P})$ (see [90, Lem. 12, p. 222]).

for all $s \le t \in [0,T]$. The convex set of all signed L^2-martingale measures for X is denoted $\mathbf{P}^2_{\mathsf{s}}(X)$.

Assumption 6.8. The set $\mathbf{P}^2_{\mathsf{s}}(X)$ contains at least one element.

This assumption ensures that no feasible strategies are arbitrages, in the light of the discussion following the Folk Theorem in Chapter 3.

Definition 6.9. A signed measure in $\mathbf{P}^2_{\mathsf{s}}(X)$ is called *variance-optimal* if it minimizes

$$\operatorname{Var}\left[\frac{d\mathbb{Q}}{d\mathbb{P}}\right] = \mathbb{E}\left[\left(\frac{d\mathbb{Q}}{d\mathbb{P}} - 1\right)^2\right] = \mathbb{E}\left[\left(\frac{d\mathbb{Q}}{d\mathbb{P}}\right)^2\right] - 1,$$

over all $\mathbb{Q} \in \mathbf{P}^2_{\mathsf{s}}(X)$.

Questions of the existence of the variance-optimal martingale measure are deferred to the next section, where we shall construct it explicitly. For the remainder of this section we prove a couple of results, conditional upon its existence. We start by deriving an alternative characterization of this measure, used in the subsequent theorem which shows how the approximation price of the claim can be computed (see Schweizer [90, Lem. 1, p. 210]).

Lemma 6.10. *Let $\widetilde{\mathbb{P}} \in \mathbf{P}^2_{\mathsf{s}}(X)$ be variance-optimal. Then*

$$\mathbb{E}\left[\frac{d\widetilde{\mathbb{P}}}{d\mathbb{P}}\left(\frac{d\widetilde{\mathbb{P}}}{d\mathbb{P}} - \frac{d\mathbb{Q}}{d\mathbb{P}}\right)\right] = 0,$$

for all $\mathbb{Q} \in \mathbf{P}^2_{\mathsf{s}}(X)$.

Proof. Let

$$\mathscr{D} := \left\{ \frac{d\mathbb{Q}}{d\mathbb{P}} \,\middle|\, \mathbb{Q} \in \mathbf{P}^2_{\mathsf{s}}(X) \right\}$$

be the family of densities of the measures in $\mathbf{P}^2_{\mathsf{s}}(X)$. To start with, let $x \in \mathbb{R}$ and $D_1, D_2 \in \mathscr{D}$. Then

$$xD_1 + (1-x)D_2 =: \bar{D} \in \mathscr{D}. \tag{6.22}$$

This follows by verifying the four properties of Definition 6.7. To do this, let $\mathbb{Q}_1, \mathbb{Q}_2 \in \mathbf{P}^2_{\mathsf{s}}(X)$ be the measures associated with D_1 and D_2 respectively and define the signed measure $\bar{\mathbb{Q}}$ on $(\Omega, \mathscr{F})$, by setting $\bar{\mathbb{Q}}\{A\} := \mathbb{E}\left[\bar{D}A\right]$, for all $A \in \mathscr{F}$. Then

1. $\bar{\mathbb{Q}}\{\Omega\} = x\mathbb{Q}_1\{\Omega\} + (1-x)\mathbb{Q}_2\{\Omega\} = x + (1-x) = 1$;
2. Let $A \in \mathscr{F}$ with $\mathbb{P}\{A\} = 0$. Then $\bar{\mathbb{Q}}\{A\} = x\mathbb{Q}_1\{A\} + (1-x)\mathbb{Q}_2\{A\} = 0$. Consequently $\bar{\mathbb{Q}} \ll \mathbb{P}$;
3. Since $D_1, D_2 \in L^2$, we have $\bar{D} \in L^2$; and

4. Let $s \le t \in [0,T]$. Then

$$\begin{aligned}\mathbb{E}\left[\bar{D}(X_t - X_s) \,\middle|\, \mathscr{F}_s\right] &= x\mathbb{E}\left[D_1(X_t - X_s) \,|\, \mathscr{F}_s\right] \\ &\quad + (1-x)\mathbb{E}\left[D_2(X_t - X_s) \,|\, \mathscr{F}_s\right] = 0.\end{aligned}$$

Now, for any $x \in \mathbb{R} \setminus \{0\}$, define the map $\varphi_x : \mathscr{D} \setminus \{\widetilde{D}\} \to \mathscr{D} \setminus \{\widetilde{D}\}$, where $\widetilde{D} := \frac{d\widetilde{\mathbb{P}}}{d\mathbb{P}}$, by setting

$$\varphi_x(D) := xD + (1-x)\widetilde{D}.$$

By (6.22), φ_x is well-defined. Next, define the map $\phi_x : \mathscr{D} \setminus \{\widetilde{D}\} \to \mathscr{D} \setminus \{\widetilde{D}\}$, by setting

$$\phi_x(D) := \frac{1}{x}D + \left(1 - \frac{1}{x}\right)\widetilde{D},$$

for all $D \in \mathscr{D} \setminus \{\widetilde{D}\}$. Again, it follows from (6.22) that ϕ_x is well defined and the following calculations establish the surjectivity of φ_x:

$$\begin{aligned}\varphi_x(\phi_x(D)) &= x\left(\frac{1}{x}D - \frac{1-x}{x}\widetilde{D}\right) + (1-x)\widetilde{D} \\ &= D - (1-x)\widetilde{D} + (1-x)\widetilde{D} \\ &= D.\end{aligned}$$

Note that φ_x is also injective, since for $D_1, D_2 \in \mathscr{D} \setminus \{\widetilde{D}\}$, we have

$$\begin{aligned} & \varphi_x(D_1) = \varphi_x(D_2) \\ \Rightarrow \quad & xD_1 + (1-x)\widetilde{D} = xD_2 + (1-x)\widetilde{D} \\ \Rightarrow \quad & D_1 = D_2.\end{aligned}$$

Thus φ_x is a bijection. Finally, for any $x \in \mathbb{R} \setminus \{0\}$ and any $D \in \mathscr{D} \setminus \{\widetilde{D}\}$, we have

$$\begin{aligned}\mathbb{E}\left[\varphi_x(D)^2\right] &= \mathbb{E}\big[(xD + (1-x)\widetilde{D})^2\big] = \mathbb{E}\big[(\widetilde{D} + x(D - \widetilde{D}))^2\big] \\ &= \mathbb{E}\big[\widetilde{D}^2\big] + 2x\mathbb{E}\big[\widetilde{D}(D - \widetilde{D})\big] + x^2\mathbb{E}\big[(D - \widetilde{D})^2\big].\end{aligned}$$

Therefore,

$$x^2\mathbb{E}\big[(D - \widetilde{D})^2\big] + 2x\mathbb{E}\big[\widetilde{D}(D - \widetilde{D})\big] = \mathbb{E}\left[\varphi_x(D)^2\right] - \mathbb{E}\big[\widetilde{D}^2\big] \ge 0, \tag{6.23}$$

since $\widetilde{\mathbb{P}}$ is variance-optimal. It follows that the coefficient of x in the left-hand side of (6.23) must be zero; otherwise a prudent choice of x would contradict the inequality (6.23)[2]. Thus $\mathbb{E}\big[\widetilde{D}(D - \widetilde{D})\big] = 0$, for all $D \in \mathscr{D}$, as required. ■

---

[2] To see this, note that the left-hand side of (6.23) is a parabola of the form $y(x) = ax^2 + bx$, with $a \ge 0$. To ensure that $y(x) \ge 0$, for all $x \in \mathbb{R} \setminus \{0\}$, we require that the discriminant satisfies $b^2 \le 0$ which in turn implies that $b = 0$.

We are now ready to establish the relationship between the variance-optimal martingale measure for X and the approximation price for H (see Schweizer [90, Prop. 2, p. 211]).

Theorem 6.11. *Suppose that $G_T(\Theta) \subseteq L^2$ is a linear space, that the solution $(v, \widetilde{\xi})$ to problem (6.21) exists and that $\widetilde{\mathbb{P}} \in \mathbf{P}^2_{\mathsf{s}}(X)$ is variance-optimal. Then the approximation price for H is $v = \mathbb{E}^{\widetilde{\mathbb{P}}}[H]$.*

Proof. Since $(v, \widetilde{\xi})$ is the solution to problem (6.21), by the projection theorem we have that

$$\mathbb{E}\left[\left(H - c - G_T\big(\widetilde{\xi}\big)\right) G_T(\vartheta)\right] = 0, \tag{6.24}$$

for every $\vartheta \in \Theta$ and

$$\mathbb{E}\left[H - v - G_T\big(\widetilde{\xi}\big)\right] = 0. \tag{6.25}$$

The last equation follows since $\mathbb{R} \times G_T(\Theta)$ is a linear space and therefore we can find $\psi \in \Theta$ with $G_T(\psi) = H - c - G_T\big(\widetilde{\xi}\big)$. Consequently, $\mathbb{E}\left[G_T(\psi)^2\right] = 0$ and by the properties of the L^2-norm, this implies that $\mathbb{E}\left[G_T(\psi)\right] = 0$ (see Luenberger [62, Lem. 2, p. 48]).

We now construct a new measure $\mathbb{Q}$ on $(\Omega, \mathscr{F})$, by setting

$$\frac{d\mathbb{Q}}{d\mathbb{P}} := \frac{d\widetilde{\mathbb{P}}}{d\mathbb{P}} + H - v - G_T\big(\widetilde{\xi}\big). \tag{6.26}$$

To verify that $\mathbb{Q} \in \mathbf{P}^2_{\mathsf{s}}(X)$, we show that it obeys the three properties of Definition 6.7. Note that

$$\mathbb{Q}\{A\} = \mathbb{E}^{\mathbb{Q}}\left[\mathbb{1}_A\right] = \mathbb{E}\left[\frac{d\mathbb{Q}}{d\mathbb{P}}\mathbb{1}_A\right] = \mathbb{E}\left[\frac{d\widetilde{\mathbb{P}}}{d\mathbb{P}}\mathbb{1}_A\right] + \mathbb{E}\left[\left(H - v - G_T\big(\widetilde{\xi}\big)\right)\mathbb{1}_A\right],$$

for all $A \in \mathscr{F}$. Consequently, by (6.25) we have $\mathbb{Q}\{\Omega\} = 1$. We also have $\mathbb{Q}\{A\} = 0$ for all $A \in \mathscr{F}$ such that $\mathbb{P}\{A\} = 0 = \widetilde{\mathbb{P}}\{A\}$, which implies that $\mathbb{Q} \ll \mathbb{P}$. Now let $s \leq t \in [0, T]$ and define $\vartheta \in \Theta$, by setting $\vartheta := \mathbb{1}_{[s,t]}$; then $G_T(\vartheta) = X_t - X_s$. Consequently, by (6.26) we have

$$\begin{aligned}
&\mathbb{E}\left[\frac{d\mathbb{Q}}{d\mathbb{P}}(X_t - X_s)\,\middle|\,\mathscr{F}_s\right] \\
&\qquad = \mathbb{E}\left[\frac{d\widetilde{\mathbb{P}}}{d\mathbb{P}}(X_t - X_s)\,\middle|\,\mathscr{F}_s\right] + \mathbb{E}\left[\left(H - v - G_T\big(\widetilde{\xi}\big)\right) G_T(\vartheta)\,\middle|\,\mathscr{F}_s\right] \\
&\qquad = \mathbb{E}\left[\frac{d\widetilde{\mathbb{P}}}{d\mathbb{P}}(X_t - X_s)\,\middle|\,\mathscr{F}_s\right] + \mathbb{E}\left[\left(H - v - G_T\big(\widetilde{\xi}\big)\right) G_T(\vartheta)\right] = 0,
\end{aligned}$$

where the second equality follows by the fact that $G_u(\vartheta) = 0$, for all $u \in [0, s]$ (which means the right hand term is independant of $\mathscr{F}_s$); and the final equality follows by virtue of $\widetilde{\mathbb{P}} \in \mathbf{P}^2_{\mathsf{s}}(X)$ and (6.24). Therefore $\mathbb{Q} \in \mathbf{P}^2_{\mathsf{s}}(X)$.

Multiplying (6.26) by $\frac{d\widetilde{\mathbb{P}}}{d\mathbb{P}}$ and taking expectations, yields

$$\mathbb{E}\left[\frac{d\mathbb{Q}}{d\mathbb{P}}\frac{d\widetilde{\mathbb{P}}}{d\mathbb{P}}\right] = \mathbb{E}\left[\left(\frac{d\widetilde{\mathbb{P}}}{d\mathbb{P}}\right)^2\right] + \mathbb{E}\left[\frac{d\widetilde{\mathbb{P}}}{d\mathbb{P}}\left(H - v - G_T(\widetilde{\xi})\right)\right],$$

which, according to Lemma 6.10, implies that

$$0 = \mathbb{E}\left[\frac{d\mathbb{Q}}{d\mathbb{P}}\frac{d\widetilde{\mathbb{P}}}{d\mathbb{P}}\right] - \mathbb{E}\left[\left(\frac{d\widetilde{\mathbb{P}}}{d\mathbb{P}}\right)^2\right] = \mathbb{E}\left[\frac{d\widetilde{\mathbb{P}}}{d\mathbb{P}}\left(H - v - G_T(\widetilde{\xi})\right)\right] = \mathbb{E}^{\widetilde{\mathbb{P}}}[H] - v,$$

yielding the result. ∎

6.4 The Variance-Optimal Measure for a Deterministic MVT Process

In Section 6.2 we assumed that the EMVT process was deterministic in order to find the optimal portfolio. With the assumption of a deterministic MVT process we now show that it is possible to characterize the variance-optimal measure as being the same as the minimal martingale measure.

We start by providing a result that shows the relationship between a (possibly signed) martingale density for X and a certain stochastic differential equation. This is essentially a generalization of Theorem 5.16. Then we state and prove the main result, which is essentially Schweizer [89, Thm. 8, p. 589].

Proposition 6.12. *Let $Z \in \mathcal{M}^2$. Then Z is a density function for $\mathbb{Q} \in \mathbf{P}^2_{\mathsf{s}}(X)$ iff it satisfies the SDE*

$$Z_t = 1 - \int_0^t \alpha_s Z_{s-}\, dM_s + L_t, \tag{6.27}$$

for all $t \in [0, T]$ and for some $L \in \mathcal{M}^2_0$, strongly orthogonal to M.

Proof. ($\Rightarrow$) Let $\mathbb{Q} \in \mathbf{P}^2_{\mathsf{s}}(X)$. Suppose $Z \in \mathcal{M}^2$ is its associated density function and define $L \in \mathcal{M}^2_0$ by

$$L_t := Z_t - 1 + \int_0^t \alpha_s Z_{s-}\, dM_s,$$

for all $t \in [0, T]$. Then, using the stochastic integration by parts rule, we have

$$\begin{aligned} d(XZ)_t &= X_{t-}\, dZ_t + Z_{t-}\, dM_t + Z_{t-}\, dA_t + d[M, Z]_t + d[A, Z]_t \\ &= X_{t-}\, dZ_t + Z_{t-}\, dM_t + \alpha_t Z_{t-}\, d\langle M\rangle_t + d[M, Z]_t + \Delta A_t \Delta Z_t \\ &= X_{t-}\, dZ_t + Z_{t-}\, dM_t + \alpha_t Z_{t-}(d\langle M\rangle_t - d[M]_t) + \Delta A_t \Delta Z_t \\ &\quad + \alpha_t Z_{t-}\, d[M]_t + d[M, Z]_t \\ &= X_{t-}\, dZ_t + Z_{t-}\, dM_t + \alpha_t Z_{t-}(d\langle M\rangle_t - d[M]_t) + \Delta A_t \Delta Z_t + d[L, M]_t, \end{aligned}$$

for all $t \in [0,T]$. Note that all the terms in this expression are martingales except the last (by Theorem 4.2, Proposition 2.49 and Lemma 2.47). Since XZ is a martingale by assumption, we require the last term above to be a martingale, which implies that L is strongly orthogonal to M, yielding the representation (6.27).
($\Leftarrow$) Suppose that Z satisfies (6.27). Then using the stochastic integration by parts rule we have

$$\begin{aligned} d(XZ)_t &= X_{t-}\,dZ_t + Z_{t-}\,dM_t + Z_{t-}\,dA_t + d[M,Z]_t + d[A,Z]_t \\ &= X_{t-}\,dZ_t + Z_{t-}\,dM_t + \alpha_t Z_{t-}(d\langle M\rangle_t - d[M]_t) + \Delta A_t \Delta Z_t, \end{aligned}$$

for all $t \in [0,T]$, which shows that XZ is a martingale in which case Z is a density function associated with some $\mathbb{Q} \in \mathbf{P}^2_{\mathsf{s}}(X)$. ■

Theorem 6.13. *Suppose $\widehat{K}$ is deterministic. Then the minimal martingale measure and the variance-optimal martingale measure coincide.*

In order to prove this result, we need a result pertaining to the Doléans exponential, which we provide as a lemma.

Lemma 6.14. *Suppose $\widehat{K}$ is deterministic and let $V \in \mathcal{S}$. Then the solution of the SDE*

$$U_t = V_t + \int_0^t U_{s-}\,d\widehat{K}_s, \tag{6.28}$$

is given by

$$U_t = \mathscr{E}\big(\widehat{K}\big)_t \left(V_0 + \int_0^t \frac{1}{\mathscr{E}\big(\widehat{K}\big)_s}\,dV_s \right)$$

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

Proof. Following Protter [74, Thm. V.52, p. 322] and Jacod [53, Thm. 6.8][3], we assume that the solution to this equation is of the form

$$U = \mathscr{E}\big(\widehat{K}\big) C, \tag{6.29}$$

for some process C. Then the stochastic integration by parts rule gives

$$\begin{aligned} dU_t &= C_{t-}\,d\mathscr{E}\big(\widehat{K}\big)_t + \mathscr{E}\big(\widehat{K}\big)_{t-}\,dC_t + d\big[C, \mathscr{E}\big(\widehat{K}\big)\big]_t \\ &= C_{t-}\mathscr{E}\big(\widehat{K}\big)_{t-}\,d\widehat{K}_t + \mathscr{E}\big(\widehat{K}\big)_{t-}\,dC_t + \mathscr{E}\big(\widehat{K}\big)_{t-}\,d\big[C, \widehat{K}\big]_t \\ &= U_{t-}\,d\widehat{K}_t + \mathscr{E}\big(\widehat{K}\big)_{t-}(1 + \Delta\widehat{K}_t)\,dC_t \\ &= U_{t-}\,d\widehat{K}_t + \mathscr{E}\big(\widehat{K}\big)_t\,dC_t, \end{aligned} \tag{6.30}$$

[3] Protter treats a less general case than is necessary for this application, while Jacod considers a more general case.

for all $t \in [0,T]$. (The penultimate line follows from the fact that $\widehat{K}$ is deterministic and hence of finite variation.) Comparing the differential form of (6.28) with (6.30) yields

$$\mathscr{E}\big(\widehat{K}\big)_t \, dC_t = dV_t,$$

for all $t \in [0,T]$. Since $\widehat{K}$ is increasing, we have $\Delta\widehat{K}_t \geq 0$ and therefore $\mathscr{E}\big(\widehat{K}\big)_t > 0$, for all $t \in [0,T]$. Dividing through by $\mathscr{E}\big(\widehat{K}\big)$ and using (6.29) gives

$$d\left(\frac{U_t}{\mathscr{E}\big(\widehat{K}\big)_t}\right) = \frac{1}{\mathscr{E}\big(\widehat{K}\big)_t}\, dV_t.$$

The result follows by expressing the above in integral form. ■

Proof of Theorem 6.13. Let $\mathbb{Q} \in \mathbf{P}^2_{\mathsf{s}}(X)$, with density process $Z \in \mathcal{M}^2$, so that

$$Z_t := \mathbb{E}\left[\frac{d\mathbb{Q}}{d\mathbb{P}} \,\middle|\, \mathscr{F}_t\right],$$

for all $t \in [0,T]$. By Proposition 6.12, Z solves the SDE:

$$Z_t = 1 - \int_0^t Z_{s-}\alpha_s \, dM_s + L_t,$$

for all $t \in [0,T]$ and some $L \in \mathcal{M}^2_0$, strongly orthogonal to M. As a consequence of this strong orthogonality we obtain

$$\begin{aligned}
\langle Z\rangle_t &= \left\langle -\int_0^{\cdot} Z_{s-}\alpha_s \, dM_s \right\rangle_t + \langle L\rangle_t \\
&= \int_0^t Z_{s-}^2 \alpha_s^2 \, d\langle M_s\rangle_s + \langle L\rangle_t \\
&= \int_0^t Z_{s-}^2 \, d\widehat{K}_s + \langle L\rangle_t,
\end{aligned}$$

for all $t \in [0,T]$. Since $Z^2 - \langle Z\rangle$ is a martingale and $Z_0 = 1$, we get

$$\begin{aligned}
\mathbb{E}\left[Z_t^2\right] &= 1 + \mathbb{E}\left[\langle Z\rangle_t\right] \\
&= 1 + \mathbb{E}\left[\langle L\rangle_t\right] + \int_0^t \mathbb{E}\left[Z_{s-}^2\right] \, d\widehat{K}_s,
\end{aligned} \tag{6.31}$$

for all $t \in [0,T]$, where the last line follows by an application of Fubini's theorem, since $\widehat{K}$ is deterministic. Now define two new processes U and V, by setting

$$U_t := \mathbb{E}\left[Z_t^2\right] \qquad \text{and} \qquad V_t := 1 + \mathbb{E}\left[\langle L\rangle_t\right],$$

for all $t \in [0,T]$. This allows (6.31) to be rewritten as

$$U_t = V_t + \int_0^t U_{s-} \, d\widehat{K}_s,$$

for all $t \in [0, T]$. Then by Lemma 6.14,

$$\mathbb{E}\left[Z_t^2\right] = U_t = \mathscr{E}\big(\widehat{K}\big)_t \left(1 + \int_0^t \frac{1}{\mathscr{E}\big(\widehat{K}\big)_s} \, dV_s\right), \tag{6.32}$$

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

Now, in the case of the minimal martingale measure $\widehat{\mathbb{P}}$, $L = 0$, by Theorem 5.17 and therefore $V = 1$. Consequently,

$$\mathbb{E}\left[\left(Z_t^{\widehat{\mathbb{P}}}\right)^2\right] = \mathscr{E}\big(\widehat{K}\big)_t,$$

for all $t \in [0, T]$, according to (6.32). Finally, since $\widehat{K}$ and V are increasing and non-negative, we have

$$\mathbb{E}\left[\left(\frac{d\widehat{\mathbb{P}}}{d\mathbb{P}} - 1\right)^2\right] = \mathscr{E}\big(\widehat{K}\big)_T - 1 \leq \mathscr{E}\big(\widehat{K}\big)_T + \int_0^T \frac{\mathscr{E}\big(\widehat{K}\big)_T}{\mathscr{E}\big(\widehat{K}\big)_s} \, dV_s - 1 = \mathbb{E}\left[\left(\frac{d\mathbb{Q}}{d\mathbb{P}} - 1\right)^2\right].$$

Since $\mathbb{Q}$ was chosen arbitrarily, this establishes that $\widetilde{\mathbb{P}}$ is variance-optimal, by Definition 6.9. ■

6.5 Mean-Variance Hedging for Continuous Processes

For completeness, we summarize a few results that generalize Theorem 6.3, without providing any detailed proofs. For a full account, consult the survey of Schweizer [91] and the original references [76, 37, 70, 77].

Firstly, under the assumption of a continuous price process X (note that processes orthogonal to X need not be continuous), the mean-variance optimal strategy can be determined without the requirement of a deterministic EMVT process. This can be derived using the so-called weighted norm inequalities (see [19]).

The next two results provide representation properties for the variance-optimal measure. Note that all results are stated under the assumptions that X is continuous and that $\mathbf{P}_{\mathsf{e}}^2(X) \neq \emptyset$.

Theorem 6.15. $\widetilde{\mathbb{P}} \in \mathbf{P}_{\mathsf{e}}^2(X)$.

Proof. See Delbaen and Schachermayer [21, Thm. 1.3]. ■

In the light of Theorem 6.15, the density process $Z^{\widetilde{\mathbb{P}}}$ for the variance-optimal martingale measure, defined by

$$Z_t^{\widetilde{\mathbb{P}}} := \mathbb{E}\left[\frac{d\widetilde{\mathbb{P}}}{d\mathbb{P}} \,\middle|\, \mathscr{F}_t\right],$$

for all $t \in [0,T]$, is strictly positive. Now define a new process $\widetilde{Z}$, by setting

$$\widetilde{Z}_t := \mathbb{E}^{\widetilde{\mathbb{P}}}\left[\frac{d\widetilde{\mathbb{P}}}{d\mathbb{P}}\,\middle|\,\mathscr{F}_t\right] = \frac{1}{Z_t^{\widetilde{\mathbb{P}}}}\mathbb{E}\left[\left(Z_T^{\widetilde{\mathbb{P}}}\right)^2\,\middle|\,\mathscr{F}_t\right],$$

for all $t \in [0,T]$ (the equality follows from Theorem 2.61). The following lemma provides a representation for $\widetilde{Z}$.

Lemma 6.16. *There exists a process $\widetilde{\zeta} \in \Theta$ such that*

$$\widetilde{Z}_t = \widetilde{Z}_0 + \int_0^t \widetilde{\zeta}_s \, dX_s,$$

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

Proof. See Delbaen and Schachermayer [21, Lem. 2.2]. ■

Under the standing assumptions of this section, the next result provides the recipe for mean-variance optimal hedging.

Theorem 6.17. *Write the* GKW *decomposition of H with respect to X under $\widetilde{\mathbb{P}}$ as*

$$H = \mathbb{E}^{\widetilde{\mathbb{P}}}[H] + \int_0^T \widetilde{\xi}_s^H \, dX_s + \widetilde{L}_T^H,$$

where $\widetilde{\xi}^H \in \mathcal{L}^2(X)$ and $L^H \in \mathcal{M}_0^2(\widetilde{\mathbb{P}})$ is strongly orthogonal to X under $\widetilde{\mathbb{P}}$. Now define the process $\widetilde{V}$, by setting

$$\widetilde{V}_t^H := \mathbb{E}^{\widetilde{\mathbb{P}}}[H \mid \mathscr{F}_t] = \mathbb{E}^{\widetilde{\mathbb{P}}}[H] + \int_0^t \widetilde{\xi}_s^H \, dX_s + \widetilde{L}_t^H,$$

for all $t \in [0,T]$. Then the mean-variance optimal strategy for H is given in feedback form by

$$\widetilde{\xi}_t = \widetilde{\xi}_t^H - \frac{\widetilde{\zeta}_t}{\widetilde{Z}_t}\left(\widetilde{V}_{t-}^H - \mathbb{E}^{\widetilde{\mathbb{P}}}[H] - G_t(\widetilde{\xi})\right),$$

for all $t \in [0,T]$. Furthermore, the approximation price of the claim is given by $v = \mathbb{E}^{\widetilde{\mathbb{P}}}[H]$.

Proof. See Schweizer [91, Thm. 4.6, p. 567]. ■

Finally, if we again assume that the MVT process is deterministic then of course Theorem 6.13 ensures that the minimal martingale measure and the variance-optimal measure coincide. The following theorem summarizes the results relating to continuous processes.

Theorem 6.18. *If $\widehat{K}_T$ is deterministic, then $\widetilde{\mathbb{P}} = \widehat{\mathbb{P}}$ and*

$$\begin{aligned}
Z_t^{\widetilde{\mathbb{P}}} &= Z_t^{\widehat{\mathbb{P}}} = \mathscr{E}\left(-\int_0^{\cdot} \alpha\, dM\right)_t; \\
\widetilde{Z}_t &= e^{\widehat{K}_T} Z_t^{\widetilde{\mathbb{P}}}; \\
\widetilde{\zeta}_t &= -e^{\widehat{K}_T} Z_t^{\widetilde{\mathbb{P}}} \alpha_t = -\widetilde{Z}_t \alpha_t; \quad \text{and} \\
\frac{Z_t^{\widetilde{\mathbb{P}}}}{\widetilde{Z}_t} &= e^{-(\widehat{K}_T - \widehat{K}_t)},
\end{aligned}$$

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

Proof. See Schweizer [91, Lem. 4.7, p. 569]. ■

The condition of a deterministic MVT process in Theorem 6.18 ensures that the mean-variance optimal strategy can be easily determined once the minimal martingale measure is constructed. If this condition is met (as is the case for our basis risk application in Chapter 7), then the strategy may be derived merely by inspection. If not, however, then any application of the mean-variance optimal strategy faces significant hurdles (see e.g. Heath, Platen and Schweizer [42, 43, 44], where quadratic approaches are applied to a stochastic volatility model).

Chapter 7

An Application to Basis Risk

In previous chapters the theory of quadratic hedging was developed, with local risk minimization and mean-variance hedging being the two approaches explored. We now apply this theory to the practical problem of hedging basis risk.

When a contingent claim is written on an underlying asset in which trading is not possible, it is natural to enquire about the effectiveness of hedging with a correlated asset. In this situation the market is incomplete and the risk that arises as a result of imperfect hedging is known as basis risk. Examples include weather derivatives, real options, options on illiquid stocks and options on very large baskets of stocks.

A simple basis risk model comprising two correlated assets is specified in the first section. We assume that it is not possible to trade in the asset on which the option is written; we do, however, require that the price of this non-traded asset is observable. The second asset is available for trade and will be used as a proxy to hedge the option. Since the theory in the previous chapters was developed in terms of discounted assets, we specify the discounted dynamics of these assets as geometric Brownian motions.

The FS decomposition of a claim is derived in the second section. This is achieved by expressing the non-traded asset in terms of the traded asset and an orthogonal process. By using a drift-adjusted representation of the non-traded asset, it is possible to construct the minimal martingale measure and employ the Feynman-Kač theorem to express the discounted claim price as the solution of a PDE boundary-value problem.

In Section 3 we present the hedging strategies for the two quadratic approaches. The FS decomposition makes it easy to specify the locally risk-minimizing strategy, with prices determined by taking expectations under the minimal martingale measure. Furthermore, since the mean-variance tradeoff process is deterministic under the chosen model assumptions, the minimal martingale measure and the variance-optimal martingale measure coincide. The mean-variance optimal self-financing strategy is thus easily constructed.

Having obtained a PDE representation of the price of the claim and its hedge parameters in discounted terms in Section 2, the fourth section does the same in non-discounted terms, by employing a simple transformation of variables. Remarkably, the PDE that emerges, for both local risk minimization and mean-variance optimization, is the familiar Black-Scholes equation. Consequently, both approaches yield classical closed-form derivative pricing formulas (which is a boon, from a computational point of view). What do change, however, are the hedge ratios — reflecting the different attitudes to hedging risk implied by the two quadratic criteria.

Finally, in Section 5 we conduct some numerical experiments to demonstrate the efficacy of the quadratic hedging approaches and compare these results with those obtained using a utility indifference approach. Utility indifference pricing and hedging has been explored by a number of authors (see e.g. [17, 96, 45, 68]). In particular, Monoyios [68] has developed a hedging algorithm based on perturbation expansions. A series of Monte Carlo experiments compare the numerical efficiency of this approach to that of the quadratic approaches.

7.1 Market Assumptions

We fix a finite time-horizon $T \in (0, \infty)$ and a stochastic basis $(\Omega, \mathscr{F}, \mathbf{F}, \mathbb{P})$, which supports two orthogonal Brownian motions W^1 and W^2. All processes are defined on the above stochastic basis (in particular, they exist over the time interval $[0, T]$) and are adapted to the filtration $\mathbf{F} = (\mathscr{F}_t)_{t \in [0,T]}$, which we take to be the augmentation of the filtration generated by W^1 and W^2 and consequently satisfies the usual conditions.

We specify a bank account process B, as follows:

$$B_t = e^{rt},$$

for all $t \in [0, T]$, where $r > 0$ is a constant short rate. Two processes U and S represent the risky assets at our disposal; U is not traded, while (the correlated asset) S is available for trade. Now consider a European option on U, with maturity T and payoff $h(U_T)$, for some Borel-measurable function $h : \mathbb{R}_+ \to \mathbb{R}_+$ satisfying $h(U_t) \in L^1$, for all $t \in [0, T]$. The objective is to hedge this instrument using the traded asset S, in such a way that the basis risk[1] is minimized.

Since the analysis in the previous chapters was carried out using the discounted assets, we introduce two discounted assets — $\overline{U}$ representing the discounted non-

[1] The risk that offsetting investments in a hedging strategy will not experience price changes in entirely opposite directions from each other. This imperfect correlation between the two investments creates the potential for excess gains or losses in a hedging strategy, thus adding risk to the position. (Source: www.investopedia.com.)

traded asset and X representing the discounted traded asset — by setting

$$\overline{U}_t := \frac{U_t}{B_t} \quad \text{and} \quad X_t := \frac{S_t}{B_t},$$

for all $t \in [0,T]$. Furthermore, we assume that the discounted assets are driven by the Brownian motions W^1 and W^2, as follows:

$$d\overline{U}_t = (\mu_{\mathsf{u}} - r)\overline{U}_t\,dt + \sigma_{\mathsf{u}}\overline{U}_t(\rho\,dW_t^1 + \sqrt{1-\rho^2}\,dW_t^2), \tag{7.1}$$

$$dX_t = (\mu - r)X_t\,dt + \sigma X_t\,dW_t^1 \tag{7.2}$$

for all $t \in [0,T]$, where $\sigma_{\mathsf{u}}, \sigma > 0$, $\mu_{\mathsf{u}}, \mu > r$ and $-1 \leq \rho \leq 1$ are constants. We now wish to hedge the discounted European claim $\bar{h}(\overline{U}_T)$, where $\bar{h} : \mathbb{R}_+ \to \mathbb{R}_+$ is defined by

$$\bar{h}(x) = e^{-rT}h(e^{rT}x),$$

for all $x \in \mathbb{R}_+$, using the correlated asset X.

For convenience, we define the market prices of risk for the non-traded and traded assets by

$$\lambda_{\mathsf{u}} := \frac{\mu_{\mathsf{u}} - r}{\sigma_{\mathsf{u}}} \quad \text{and} \quad \lambda := \frac{\mu - r}{\sigma},$$

respectively.

In the case where the assets are perfectly correlated (i.e. $\rho = 1$), it is well known (see e.g. [17]) that the absence of arbitrage implies that their market prices of risk should be equal (i.e. $\lambda_{\mathsf{u}} = \lambda$). Under this condition, the (non-discounted) price of a European call or put on the non-traded asset, given that its (non-discounted) value is s at time $t \in [0,T]$, is determined by the standard Black-Scholes formula

$$\mathsf{BS}(t,s,q,\sigma) := \delta\big(se^{-q(T-t)}N(\delta d_1) - Ke^{-r(T-t)}N(\delta d_2)\big),$$

with

$$d_1 := \frac{\ln(s/K) + (r - q + \sigma^2/2)(T-t)}{\sigma\sqrt{T-t}} \quad \text{and} \quad d_2 := d_1 - \sqrt{T-t}.$$

Here $\delta = 1$ for a call and $\delta = -1$ for a put, while $K > 0$ is the strike price and $q > 0$ is the dividend yield of the non-traded asset[2]. Then,

$$\Delta_{\mathsf{BS}}(t,s,q,\sigma) := \delta e^{-q(t-T)}N(\delta d_1)$$

is the usual Black-Scholes delta and perfect hedging is achieved by holding

$$\frac{\sigma_{\mathsf{u}}U_t}{\sigma S_t}\Delta_{\mathsf{BS}}(t,U_t,0,\sigma_{\mathsf{u}}) \tag{7.3}$$

units of the traded asset S. It will be shown that the quadratic hedging approaches are consistent with this limiting regime.

[2] Note that we have not modelled dividend yield in our basis risk model. We will however require this general Black-Scholes formula later.

7.2 The Föllmer-Schweizer Decomposition

To start with, note that X satisfies the structure condition, since its canonical decomposition takes the form

$$X_t = X_0 + M_t + \int_0^t \alpha_s \, d\langle M \rangle_s, \tag{7.4}$$

with

$$M_t := \int_0^t \sigma X_s \, dW_s^1 \quad \text{and} \quad \alpha_t := \frac{\mu - r}{\sigma^2 X_t}, \tag{7.5}$$

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

Our task in this section is to find the FS decomposition of the discounted claim, which has the form

$$\bar{h}(\overline{U}_T) =: H = H_0 + \int_0^T \xi_s^H \, dX_s + L_T^H \tag{7.6}$$

for some $\xi^H \in \mathcal{L}(\overline{U})$ and $L^H \in \mathcal{M}_0^2$, strongly orthogonal to M. Having done so, it is a simple matter to identify the price and the hedging parameters for both the local risk-minimizing and the mean-variance optimal strategies, by inspection. (Note that the approach taken is similar to that of Schweizer [87].)

By rearranging (7.2), we get

$$dW_t^1 = \frac{dX_t}{\sigma X_t} - \lambda \, dt,$$

for all $t \in [0, T]$. Substituting this into (7.1) yields

$$\frac{d\overline{U}_t}{\overline{U}_t} = (\mu_{\mathsf{u}} - r - \rho \sigma_{\mathsf{u}} \lambda) \, dt + \sigma_{\mathsf{u}} \left(\frac{\rho}{\sigma X_t} \, dX_t + \sqrt{1 - \rho^2} \, dW_t^2 \right), \tag{7.7}$$

for all $t \in [0, T]$. We now specify the drift-adjusted process $\widetilde{U}$ as the unique strong solution of the SDE

$$d\widetilde{U}_t = \widetilde{U}_t \left(\frac{d\overline{U}_t}{\overline{U}_t} + \gamma \, dt \right),$$

for all $t \in [0, T]$, with

$$\gamma := \sigma_{\mathsf{u}} \left(\rho \lambda - \lambda_{\mathsf{u}} \right) \tag{7.8}$$

and subject to the terminal condition $\widetilde{U}_T = \overline{U}_T$. A simple calculation shows that

$$\widetilde{U}_t = e^{-\gamma (T - t)} \overline{U}_t,$$

for all $t \in [0, T]$; which, when substituted into (7.7), yields

$$d\widetilde{U}_t = \sigma_{\mathsf{u}} \widetilde{U}_t \left(\frac{\rho}{\sigma X_t} \, dX_t + \sqrt{1 - \rho^2} \, dW_t^2 \right), \tag{7.9}$$

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

We now construct the minimal martingale measure for X. By Theorem 5.17 and the canonical decomposition (7.4), the density process for the minimal martingale measure $\widehat{\mathbb{P}}$ is given by

$$\widehat{Z}_t := \mathbb{E}\left[\frac{d\widehat{\mathbb{P}}}{d\mathbb{P}}\,\middle|\,\mathscr{F}_t\right] = \mathscr{E}\Big(-\int_0^{\cdot} \alpha_s\, dM_s\Big)_t = \mathscr{E}\big(-\lambda W^1\big)_t,$$

for all $t \in [0,T]$. Since X is a martingale under $\widehat{\mathbb{P}}$, we can define a new process $\widehat{W}$ as follows:

$$d\widehat{W}_t = dW_t^1 + \lambda\, dt,$$

for all $t \in [0,T]$. Since $\widehat{W}\widehat{Z}$ is a martingale and $\langle \widehat{W}\rangle_t^{\widehat{\mathbb{P}}} = \big[\widehat{W}\big]_t = t$, for all $t \in [0,T]$, Lévy's characterization of Brownian motion (Theorem 3.25) informs us that $\widehat{W}$ is a Brownian motion under $\widehat{\mathbb{P}}$. Rewriting (7.2) and (7.9) in terms of $\widehat{W}$ gives

$$dX_t = \sigma X_t\, d\widehat{W}_t$$

and

$$d\widetilde{U}_t = \sigma_{\mathsf{u}}\widetilde{U}_t\big(\rho\, d\widehat{W}_t + \sqrt{1-\rho^2}\, dW_t^2\big), \tag{7.10}$$

for all $t \in [0,T]$. Note that W^2 is strongly orthogonal to M which means that its martingale property is preserved under the minimal martingale measure (see Definition 5.13); and $\langle U\rangle_t^{\widehat{\mathbb{P}}} = [U]_t = t$. Then U is a Brownian motion under $\widehat{\mathbb{P}}$, again by Lévy's characterization of Brownian motion, which in turn means that the expression in brackets in (7.10) is a Brownian motion under $\widehat{\mathbb{P}}$.

We now use the Feynman-Kač theorem (Theorem 3.26) to infer a PDE representation for the claim. Define $F : [0,T] \times (0,\infty) \to \mathbb{R}_+$, by setting

$$F(t,x) := \mathbb{E}^{\widehat{\mathbb{P}}}\left[\bar{h}(\widetilde{U}_T)\,\middle|\,\widetilde{U}_t = x\right],$$

for all $(t,x) \in [0,T] \times (0,\infty)$. Obviously we then have

$$F(T,x) = \bar{h}(x), \tag{7.11}$$

for all $x \in (0,\infty)$. According to the Feynman-Kač theorem, F satisfies the following PDE:

$$\frac{\partial F}{\partial t}(t,x) + \frac{1}{2}\sigma_{\mathsf{u}}^2 x^2 \frac{\partial^2 F}{\partial x^2}(t,x) = 0, \tag{7.12}$$

for all $(t,x) \in [0,T] \times (0,\infty)$, with terminal condition (7.11). Applying Itô's formula to the process $(F(t,\widetilde{U}_t))_{t\in[0,T]}$ yields

$$\begin{aligned}\bar{h}(\overline{U}_T) = \bar{h}(\widetilde{U}_T) = F(T,\widetilde{U}_T) = F(0,\widetilde{U}_0) &+ \int_0^T \frac{\partial F}{\partial x}(t,\widetilde{U}_s)\, d\widetilde{U}_s \\ &+ \int_0^T \left(\frac{\partial F}{\partial t}(s,\widetilde{U}_s) + \frac{1}{2}\sigma_{\mathsf{u}}^2 \widetilde{U}_s^2 \frac{\partial^2 F}{\partial x^2}(s,\widetilde{U}_t)\right) ds.\end{aligned}$$

Substituting (7.9) and (7.12) into this expression gives

$$\bar{h}(\overline{U}_T) = F(0, \widetilde{U}_0) + \int_0^T \frac{\rho\sigma_{\mathrm{u}}\widetilde{U}_s}{\sigma X_s}\frac{\partial F}{\partial x}(s, \widetilde{U}_s)\, dX_s + \int_0^T \sigma_{\mathrm{u}}\widetilde{U}_s\sqrt{1-\rho^2}\frac{\partial F}{\partial x}(s, \widetilde{U}_s)\, dW_s^2.$$

This is the FS decomposition we have been looking for. Comparing terms with (7.6), we obtain

$$\begin{aligned} H_0 &= F(0, \widetilde{U}_0) = \mathbb{E}^{\widehat{\mathbb{P}}}\left[\bar{h}(\widetilde{U}_T)\right]; \\ \xi_t^H &= \frac{\rho\sigma_{\mathrm{u}}\widetilde{U}_t}{\sigma X_t}\frac{\partial F}{\partial x}(t, \widetilde{U}_t); \quad \text{and} \\ L_t^H &= \int_0^t \sigma_{\mathrm{u}}\widetilde{U}_s\sqrt{1-\rho^2}\frac{\partial F}{\partial x}(s, \widetilde{U}_s)\, dW_s^2, \end{aligned} \tag{7.13}$$

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

7.3 Hedging Strategies

Now that we have the FS decomposition, it is an easy matter to determine the locally risk-minimizing strategy. By Proposition 5.10 it is the mean self-financing strategy $(\widehat{\xi}, \widehat{\eta})$ determined by

$$(\widehat{\xi}_t, \widehat{\eta}_t) := \left(\xi_t^H, \widehat{V}_t - \xi_t^H X_t\right),$$

where ξ^H is given by (7.13) and

$$\widehat{V}_t := \mathbb{E}^{\widehat{\mathbb{P}}}\left[\bar{h}(\widetilde{U}_T)\,\middle|\,\mathscr{F}_t\right] = F(t, \widetilde{U}_t),$$

for $t \in [0, T]$. (Here $\widehat{\xi}$ specifies the holding in the discounted traded asset X and $\widehat{\eta}$ is the bank account holding.)

It is also an easy matter to find the mean-variance optimal strategy. Since X satisfies the structure condition, we can use (7.5) to obtain the mean-variance trade-off process $\widehat{K}$ as follows:

$$\widehat{K}_t = \int_0^t \alpha_s^2 d\langle M\rangle_s = \int_0^t \left(\frac{\mu - r}{\sigma}\right)^2 ds = \lambda^2 t,$$

for all $t \in [0, T]$. This is a deterministic quantity and thus by Theorem 6.3, Theorem 6.11 and Theorem 6.13, we can express the self-financing mean-variance optimal strategy $(\widetilde{\xi}, \widetilde{\eta})$ as follows:

$$(\widetilde{\xi}_t, \widetilde{\eta}_t) := \left(\xi_t^{(v)}, v + G_t\big(\xi^{(v)}\big) - \xi_t^{(v)} X_t\right),$$

where

$$v := \mathbb{E}^{\widehat{\mathbb{P}}}\left[\bar{h}(\widetilde{U}_T)\right] = F(0, \widetilde{U}_0) \quad \text{and} \quad \xi_t^{(v)} := \xi_t^H + \alpha_t\left(\widehat{V}_t - v - G_t\big(\xi^{(v)}\big)\right),$$

for all $t \in [0, T]$. (Here $G\big(\xi^{(v)}\big)$ is the gain from trading the discounted asset X, using $\xi^{(v)}$.)

7.4 Closed-Form Expressions for Pricing and Hedging

In the previous two sections we manipulated the discounted assets to obtain the FS decomposition of the discounted claim and the hedge portfolios (in terms of the discounted assets) for the two quadratic approaches. Now, by transforming variables, we consider the situation without discounting. It is interesting to note that (7.12) looks similar to the discounted Black-Scholes PDE of Section 3.5. By performing a similar transformation of variables as was utilized there, an expression for the non-discounted intrinsic value of the claim can be derived.

Define the function $V : [0,T] \times (0,\infty) \to \mathbb{R}_+$, by setting

$$V(t,s) := e^{rt} F(t, e^{-rt} s e^{-\gamma(T-t)}),$$

for all $(t,s) \in [0,T] \times (0,\infty)$. Then

$$\begin{aligned}
\frac{\partial V}{\partial s}(t,s) &= e^{-\gamma(T-t)} \frac{\partial F}{\partial x}\big(t, e^{-rt} s e^{-\gamma(T-t)}\big) \\
\frac{\partial^2 V}{\partial s^2}(t,s) &= e^{-2\gamma(T-t)} e^{-rt} \frac{\partial^2 F}{\partial x^2}\big(t, e^{-rt} s e^{-\gamma(T-t)}\big) \\
\frac{\partial V}{\partial t}(t,s) &= r e^{rt} F\big(t, e^{-rt} s e^{-\gamma(T-t)}\big) + e^{rt} \frac{\partial F}{\partial t}\big(t, e^{-rt} s e^{-\gamma(T-t)}\big) \\
&\quad + (\gamma - r) s e^{-\gamma(T-t)} \frac{\partial F}{\partial x}\big(t, e^{-rt} s e^{-\gamma(T-t)}\big) \\
&= rV(t,s) + e^{rt} \frac{\partial F}{\partial t}\big(t, e^{-rt} s e^{-\gamma(T-t)}\big) + (\gamma - r) s \frac{\partial V}{\partial s}(t,s),
\end{aligned}$$

for all $(t,x) \in [0,T] \times (0,\infty)$. The PDE (7.12) may now be rewritten as

$$\begin{aligned}
&e^{-rt} \frac{\partial V}{\partial t}(t,s) - r e^{-rt} V(t,s) + (r-\gamma) e^{-rt} s \frac{\partial V}{\partial s}(t,s) \\
&\qquad + \frac{1}{2} \sigma_{\mathsf{u}}^2 (e^{-rt} s e^{-\gamma(T-t)})^2 e^{2\gamma(T-t)} e^{rt} \frac{\partial^2 V}{\partial s^2}(t,s) = 0,
\end{aligned}$$

which may in turn be rearranged to give

$$rV(t,s) = \frac{\partial V}{\partial t}(t,s) + (r-\gamma) s \frac{\partial V}{\partial s}(t,s) + \frac{1}{2} \sigma_{\mathsf{u}}^2 s^2 \frac{\partial^2 V}{\partial s^2}(t,s),$$

for all $(t,x) \in [0,T] \times (0,\infty)$. The boundary condition corresponding to (7.11) is

$$V(T,s) = h(s)$$

for all $s \in (0,\infty)$.

When $h(s)$ is the payoff of a put or a call, then the solution of this PDE is given by the Black-Scholes option pricing formula for a stock with a continuous dividend yield γ; i.e. $V(t,s) = \mathsf{BS}(t,s,\gamma,\sigma_{\mathsf{u}})$. Then, with the relevant substitutions, we are able to calculate the hedge parameters for the optimal strategies. For the local risk-minimizing strategy we have

$$\widehat{\xi}_t = \frac{\rho\sigma_{\mathsf{u}} U_t}{\sigma S_t}\frac{\partial V}{\partial s}(t, U_t) = \frac{\rho\sigma_{\mathsf{u}} U_t}{\sigma S_t}\Delta_{\mathsf{BS}}(t, U_t, \gamma, \sigma_{\mathsf{u}}), \tag{7.14}$$

for all $t \in [0, T]$; and for the mean-variance optimal strategy we get

$$\begin{aligned}\widetilde{\xi}_t &= \widehat{\xi}_t + \frac{\mu - r}{\sigma^2 e^{-rt} S_t}\left(\widehat{V}_t - v - G_t\big(\widetilde{\xi}\big)\right) \\ &= \widehat{\xi}_t + \frac{\mu - r}{\sigma^2 e^{-rt} S_t}\left(\widehat{V}_t - v - \int_0^t \widetilde{\xi}_u \, d\big(e^{-ru} S_u\big)\right),\end{aligned} \tag{7.15}$$

with

$$\begin{aligned}\widehat{V}_t &= e^{-rt} V(t, U_t) & \text{and} \qquad v &= V(0, U_0) \\ &= e^{-rt}\mathsf{BS}(t, U_t, \gamma, \sigma_{\mathsf{u}}) & &= \mathsf{BS}(0, U_0, \gamma, \sigma_{\mathsf{u}}).\end{aligned}$$

for all $t \in [0, T]$. (Recall that v in this context is interpreted as the approximation price of the claim.)

When U and S are perfectly correlated (i.e. when $\rho = 1$), arbitrage considerations ensure that their respective market prices of risk are equal. Consequently, by (7.8) we have $\gamma = 0$, which implies that (7.14) is the same as (7.3); thereby demonstrating that the local risk minimization approach is consistent with the standard Black-Scholes hedge for a complete market.

7.5 Hedge Simulation Results

To evaluate the effectiveness of hedging using the quadratic techniques, we now analyze the results of some hedge simulations. Initially, a comparison of the quadratic techniques with the numerical results obtained by Monoyios [68][3] was undertaken for a European put option. The put was written on the non-traded asset U and the risk was hedged by trading in S. Table 7.1 lists the model parameters that were chosen.

A Monte Carlo experiment was undertaken to test hedging performance. Ten-thousand paths for U and S were generated and rebalancing was allowed to take place 200 times, at equal intervals, over the life of the option. At the end of the period a profit or loss was recorded as the difference between the accumulated gain from hedging and the expiry value of the option. The approximation price of the option was used as the initial endowment.

[3] Monoyios applies a utility maximization approach to the problem of hedging. We do not present his approach here, but refer the reader to [68] for full details and the hedging algorithm.

U_0	S_0	K	r	μ_u	σ_u	μ	σ	T
100	100	100	5%	0.12	0.30	0.10	0.25	1 year

Tab. 7.1: Model Parameters employed by the hedge simulations (for purposes of comparison, these are the same parameter values as in Monoyios [68]).

Four strategies were used to compute the hedge parameters. These were a naive strategy (as suggested by Monoyios [68]), which simply employs the standard Black-Scholes hedge ratio (7.3); the local risk-minimizing hedge ratio (7.14); the mean-variance optimal hedge ratio (7.15); and the hedge ratio proposed by Monoyios [68, §4.1.1, p. 250]. The algorithm of Monoyios employs a risk-aversion constant[4]. For the purposes of our simulations, we set the value of this parameter to $\mathsf{rac} = 0.001$.

Two values for ρ were used, namely 0.65 and 0.85. Histograms of the resulting simulated hedge errors are given in Figure 7.1, with Tables 7.2 and 7.3 providing summary statistics.

Strategy	Max	Min	Mean	SD	Median
Naive	38.55	-53.53	-0.8492	10.7644	0.4715
Local risk	27.12	-45.11	-0.0149	9.4936	2.6151
Mean variance	30.62	-52.48	0.0014	9.3690	2.4631
Monoyios	27.15	-45.11	-0.0184	9.4933	2.6004

Tab. 7.2: Summary statistics for $\rho = 0.65$

Strategy	Max	Min	Mean	SD	Median
Naive	30.69	-39.99	-0.4004	7.0660	0.1125
Local risk	26.00	-33.37	-0.0200	6.6839	1.1064
Mean variance	28.74	-36.91	-0.0092	6.5943	1.0149
Monoyios	26.03	-33.36	-0.0222	6.6837	1.0958

Tab. 7.3: Summary statistics for $\rho = 0.85$

The results are encouraging, with the local risk-minimizing strategy performing almost as well as Monoyios' algorithm. This is not surprising, since his algorithm is based on a utility maximization paradigm for which expectations are taken under the minimal martingale measure. The mean-variance optimal strategy performed

[4] Note that the risk aversion constant in Monoyios [68] is represented by γ — this should not be confused with our use of the constant γ.

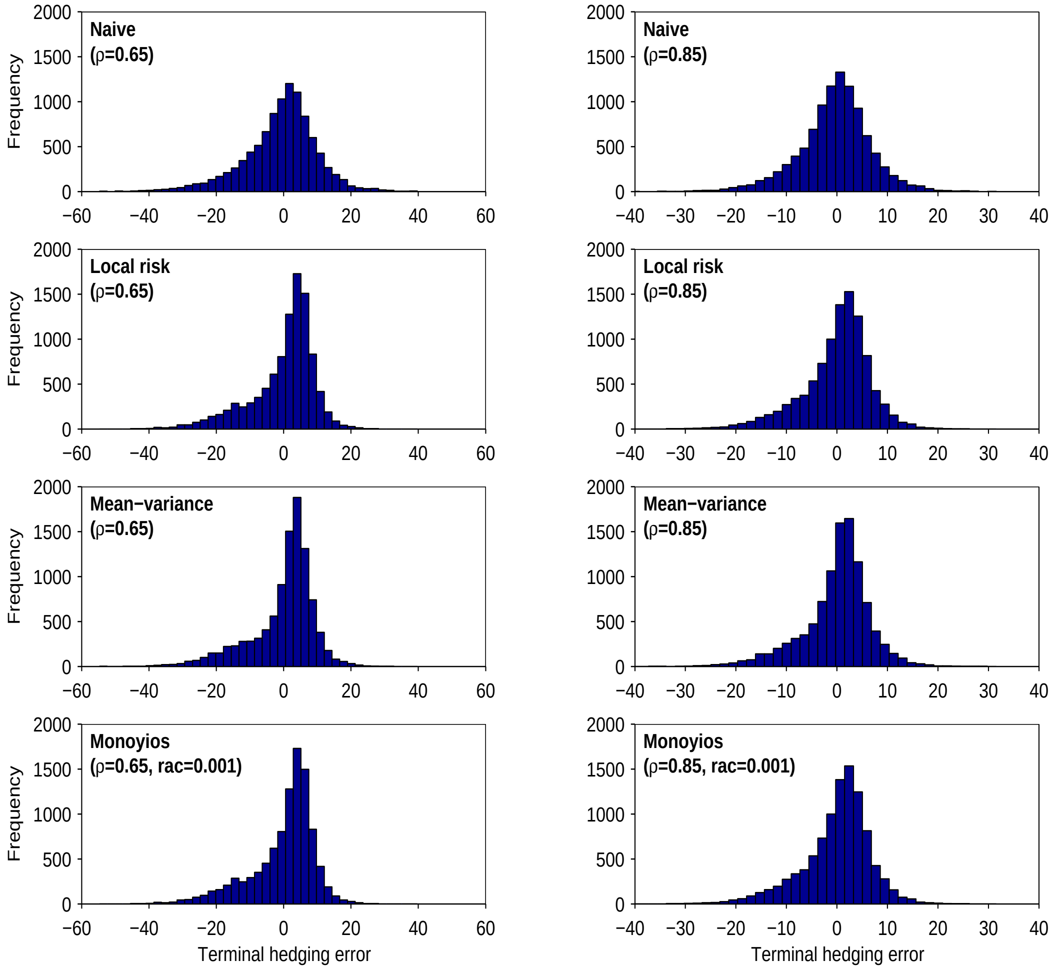

Fig. 7.1: Histograms of the hedging errors for the put option, based on 10000 sample paths. The approximation prices were 8.2324 and 8.6564, corresponding to correlation coefficients of 0.65 and 0.85, respectively.

slightly better than the other two, with a standard deviation that was 1%–2% lower. (This can be seen as an enhanced peak around the mean in the relevant histograms.) One slight drawback of this method is that its largest losses exceeded the largest losses of the other methods; but in general, it did perform better.

It should be noted that due to a constraint in his utility maximization formulation, the algorithm proposed by Monoyios [68] cannot be used directly for pricing and hedging a call option. To overcome this shortcoming, Henderson and Hobson [45, p. 74] suggest modelling the call using a static hedge consisting of put options. In contrast, our approach is applicable for both puts and calls, without modification.

Table 7.4 shows the approximation prices for various values of the correlation

ρ	Put	Call
-0.95	5.3127	23.7315
-0.75	5.6321	22.6965
-0.50	6.0493	21.4435
-0.25	6.4870	20.2358
0	6.9451	19.0730
0.25	7.4238	17.9549
0.50	7.9231	16.8812
0.75	8.4428	15.8514
0.95	8.8733	15.0588
1	9.3542	14.2312

Tab. 7.4: Put and call option approximation prices for various values of ρ. (When $\rho = 1$, they are the standard Black-Scholes prices.)

coefficient, for both put and call options, based on the parameters presented in Table 7.1. It is interesting to note that the approximation prices for the put are lower than the Black-Scholes prices, while the converse is true for the call. One should not interpret these prices as being the premiums charged for the options, since not all risk is hedged due, to incompleteness. It is therefore necessary to estimate the standard deviation (SD) of the hedging error, so that the option writer can charge an appropriate risk premium. Figure 7.2 shows the approximation prices and the standard deviations of the hedging errors for the put, using both the local risk-minimizing and mean-variance optimal hedging strategies. Figure 7.3 shows the same results for a call. The standard deviations were estimated based on Monte Carlo samples of 10000 paths.

7.6 Conclusions

Although the performance of the quadratic techniques is at least as good as that achieved with the utility approach, it should be noted that the hedging algorithms are considerably simpler. This is a pleasant consequence of having access to closed-form formulas. In contrast, the utility indifference approach requires perturbation expansions to solve the relevant PDEs which are originally derived using a “distortion” technique.

There is scope for further research. It would be nice to obtain an estimate of the variance of the hedging error, expressed in terms of the parameters of the model. This would allow one to estimate the risk premium that should be charged (over and above the approximation price) and may perhaps lead to a quantile-based formulation of the claim price.

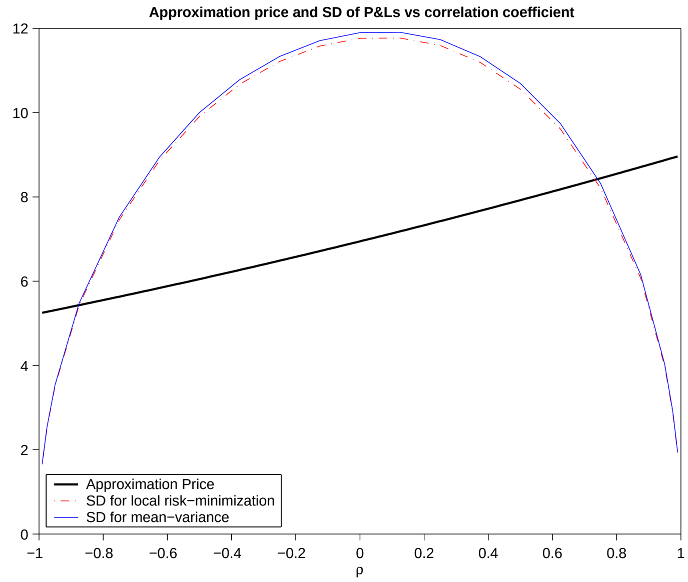

Fig. 7.2: Approximation price and standard deviation of hedging error vs correlation, for the put option with parameters given by Table 7.1.

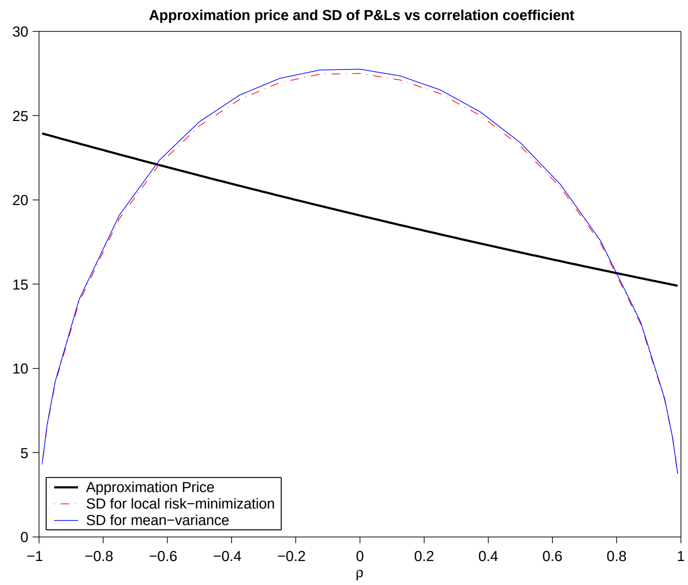

Fig. 7.3: Approximation price and standard deviation of hedging error vs correlation, for the call option with parameters given by Table 7.1.

Appendix A

Matlab Code for Chapter 7

A.1 BasisHist.m

```
%
% Program to produce Figure 7.1 and generate data for Table 7.2 and Table 7.3.
% This program calls BasisRisk.m to generate the histogram data
%

rho=0.85
BasisRisk;
rmPnLs85=rmPnLs;
mvPnLs85=mvPnLs;
nvPnLs85=nvPnLs;
monPnLs85=monPnLs;

rho=0.65
BasisRisk;
rmPnLs65=rmPnLs;
mvPnLs65=mvPnLs;
nvPnLs65=nvPnLs;
monPnLs65=monPnLs;

maxPnL85=max([rmPnLs85 mvPnLs85 nvPnLs85 monPnLs85]);
minPnL85=min([rmPnLs85 mvPnLs85 nvPnLs85 monPnLs85]);
maxPnL65=max([rmPnLs65 mvPnLs65 nvPnLs65 monPnLs65]);
minPnL65=min([rmPnLs65 mvPnLs65 nvPnLs65 monPnLs65]);
deltaPnL85=(maxPnL85-minPnL85)/nbins;
deltaPnL65=(maxPnL65-minPnL65)/nbins;

subplot(4,2,1)
hist(nvPnLs65,[minPnL65:deltaPnL65:maxPnL65]);
set(gca,'TickDir','out','FontSize',8,'Box','off')
title 'Naive (\rho=0.65)';
axis([-60 60 0 2000]);
ylabel('Frequency','FontSize',8);
set(get(gca,'Title'),'FontWeight','bold')
line([-60,60],[2000,2000],'Color','k');
line([60,60],[0,2000],'Color','k');

subplot(4,2,2)
hist(nvPnLs85,[minPnL85:deltaPnL85:maxPnL85]);
set(gca,'TickDir','out','FontSize',8,'Box','off');
set(gca,'xTick',[-40:10:40]);
title 'Naive (\rho=0.85)';
axis([-40 40 0 2000]);
set(get(gca,'Title'),'FontWeight','bold')
line([-40,40],[2000,2000],'Color','k');
line([40,40],[0,2000],'Color','k');

subplot(4,2,3)
hist(rmPnLs65,[minPnL65:deltaPnL65:maxPnL65]);
set(gca,'TickDir','out','FontSize',8,'Box','off')
title 'Local risk (\rho=0.65)';
axis([-60 60 0 2000]);
ylabel('Frequency','FontSize',8);
set(get(gca,'Title'),'FontWeight','bold')
line([-60,60],[2000,2000],'Color','k');
line([60,60],[0,2000],'Color','k');

subplot(4,2,4)
hist(rmPnLs85,[minPnL85:deltaPnL85:maxPnL85]);
set(gca,'TickDir','out','FontSize',8,'Box','off')
set(gca,'xTick',[-40:10:40]);
title 'Local risk (\rho=0.85)';
axis([-40 40 0 2000]);
set(get(gca,'Title'),'FontWeight','bold')
```

```
line([-40,40],[2000,2000],'Color','k');
line([40,40],[0,2000],'Color','k');

subplot(4,2,5)
hist(mvPnLs65,[minPnL65:deltaPnL65:maxPnL65]);
set(gca,'TickDir','out','FontSize',8,'Box','off')
title 'Mean-variance (\rho=0.65)';
axis([-60 60 0 2000]);
ylabel('Frequency','FontSize',8);
set(get(gca,'Title'),'FontWeight','bold')
line([-60,60],[2000,2000],'Color','k');
line([60,60],[0,2000],'Color','k');

subplot(4,2,6)
hist(mvPnLs85,[minPnL85:deltaPnL85:maxPnL85]);
set(gca,'TickDir','out','FontSize',8,'Box','off')
set(gca,'xTick',[-40:10:40]);
title 'Mean-variance (\rho=0.85)';
axis([-40 40 0 2000]);
set(get(gca,'Title'),'FontWeight','bold')
line([-40,40],[2000,2000],'Color','k');
line([40,40],[0,2000],'Color','k');

subplot(4,2,7)
hist(monPnLs65,[minPnL65:deltaPnL65:maxPnL65]);
set(gca,'TickDir','out','FontSize',8,'Box','off')
title 'Monoyios (\rho=0.65, \gamma=0.001)';
axis([-60 60 0 2000]);
xlabel('Terminal hedging error','FontSize',8);
ylabel('Frequency','FontSize',8);
set(get(gca,'Title'),'FontWeight','bold')
line([-60,60],[2000,2000],'Color','k');
line([60,60],[0,2000],'Color','k');

subplot(4,2,8)
hist(monPnLs85,[minPnL85:deltaPnL85:maxPnL85]);
set(gca,'TickDir','out','FontSize',8,'Box','off')
set(gca,'xTick',[-40:10:40]);
title 'Monoyios (\rho=0.85, \gamma=0.001)';
axis([-40 40 0 2000]);
xlabel('Terminal hedging error','FontSize',8);
set(get(gca,'Title'),'FontWeight','bold');
line([-40,40],[2000,2000],'Color','k');
line([40,40],[0,2000],'Color','k');
```

A.2 BasisRisk.m

```
%
% Program to produce histogram data
%

tic;

% Initialize constants

Y0=100;                 % Initial stock price of untraded stock
X0=100;                 % Initial stock price of traded stock
mux=0.10;               % drift rate of untraded stock
muy=0.12;               % drift rate of traded stock
sigmax=0.25;            % volatility of untraded stock
sigmay=0.30;            % volatility of traded stock
r=0.05;                 % NACC rate
%rho=?;                 % correlation constant set by BasisHist.m
eps=sqrt(1-rho^2);
eta=-1;                 % 1 for calls, -1 for puts
T=1;                    % maturity of the call in years
K=100;                  % Strike price
Kprime=K*exp(-r*T);     % Discounted strike price
n=200;                  % number of rebalance points
ntrials=10000;          % number of trials
nbins=40;               % number of bins for the histogram
rac=0.001;              % Risk aversion constant for Monoyios algorithm
randn('state',10);      % Initialize the random seed (for deterministic results)


%
% Calculate some reusable constants and arrays
%

deltaT=T/n;
tau=[T:-deltaT:deltaT];
cumt=[0:deltaT:T];

driftx=(mux-r-0.5*sigmax^2)*cumt;                       % Discounted drift per unit time
drifty=(muy-r-0.5*sigmay^2)*cumt;

varx=sigmax*sqrt(deltaT);                               % Variance per unit time
```

```
vary=sigmay*sqrt(deltaT);

gamma=rho*sigmay*(mux-r)/sigmax-(muy-r);                 % Calculate BS delta multiplier for RM strategy
%deltamult=exp(tau.*-gamma).*rho.*sigmay./sigmax;
deltamult=rho.*sigmay./sigmax;
gammafact=exp(tau.*-gamma);

d1bs=0.5*sigmay^2.*tau;                                   % d1 constants
d1lrm=(0.5*sigmay^2-gamma).*tau;
s=sigmay.*sqrt(tau);

lambda=(mux-r)/sigmax;                                    % Calculate constants and arrays for Monoyios algorithm
q=r-(muy-sigmay*rho*lambda);

d1mon=(r-q+0.5*sigmay^2).*tau;                            % discounted d1 constant

My2term=exp((2*(r-q)+sigmay^2).*tau);                     % Reusable arrays
My3term=exp(3*(r-q+sigmay^2).*tau);
My4term=exp(2*(2*(r-q)+3*sigmay^2).*tau);
dMy1term=exp((r-q+sigmay^2).*tau);
dMy2term=exp((2*(r-q)+3*sigmay^2).*tau);
dMy3term=exp(3*(r-q+2*sigmay^2).*tau);

erq=exp((r-q).*tau);
beta=exp(r.*tau);

%
% Calculate Black-Scholes price
%

d1=(log(Y0/K)+(r+0.5*sigmay^2)*T)/(sigmay*sqrt(T));
d2=d1-sigmay*sqrt(T);
BSprice=eta*(Y0*cnd(eta*d1)-K*exp(-r*T)*cnd(eta*d2))

%
% Calculate LRM price
%

d1=(log(Y0/K)+(r-gamma+0.5*sigmay^2)*T)/(sigmay*sqrt(T));
d2=d1-sigmay*sqrt(T);
H0=eta*(Y0*exp(-gamma*T)*cnd(eta*d1)-K*exp(-r*T)*cnd(eta*d2))

endow=H0;                                                 % Initial endowment

%
% Perform numerical experiment
%

nvPnLs=[];
rmPnLs=[];
mvPnLs=[];
monPnLs=[];

for loop=1:ntrials

   zx=[0 randn(n,1).'];                                   % Calculate discounted stock price paths
   zy=rho*zx+eps.*[0 randn(n,1).'];
   x=X0*exp(driftx+varx*cumsum(zx));
   y=Y0*exp(drifty+vary*cumsum(zy));

   d1=(log(y(1:n)./Kprime)+d1bs)./s;                      % Calculate BS deltas
   delta=eta*cnd(eta*d1);

   nvstrat=delta.*sigmay.*y(1:n)./(sigmax.*x(1:n));          % Calculate naive hedge ratios
   banknv=endow+sum(([0 nvstrat(1:(n-1))]-nvstrat).*x(1:n));  % Calculate portfolio gain for naive strategy

   d1=(log(y(1:n)./Kprime)+d1lrm)./s;                     % Calculate Local Risk Minimization deltas
   d2=d1-s;
   delta=eta*gammafact.*cnd(eta*d1);
   intrinsic=y(1:n).*delta-eta*Kprime.*cnd(eta*d2);          % Calculate discounted Intrinsic values

   rmstrat=delta.*y(1:n).*deltamult./x(1:n);                 % Calculate local risk minimizing hedge ratios
   bankrm=endow+sum(([0 rmstrat(1:(n-1))]-rmstrat).*x(1:n));  % Calculate portfolio gain for rm strategy

   bankmv=endow;                                          % Calculate portfolio gain for mv strategy
   prevdeltamv=0;
   for t=1:n
      newdeltamv=rmstrat(t)+(intrinsic(t)-H0-(bankmv+prevdeltamv*x(t)-endow))*(mux-r)/(sigmax^2*x(t));
      bankmv=bankmv+(prevdeltamv-newdeltamv)*x(t);
      prevdeltamv=newdeltamv;
   end

%
% Now compute hedging strategy for Monoyios algorithm
%
   y1=exp(r*cumt(1:n)).*y(1:n);                           % Calculate components of Monoyios expansion
   y2=y1.*y1;
   y3=y2.*y1;
```

```
    y4=y3.*y1;

    d1=(log(y1./K)+d1mon)./s;

    Nd1=cnd(-d1);                                              % Calculate cumulative normal functions
    Nd1ps=cnd(-d1+s);
    Nd1ms=cnd(-d1-s);
    Nd1m2s=cnd(-d1-2*s);
    Nd1m3s=cnd(-d1-3*s);

    M1=K*Nd1ps-y1.*erq.*Nd1;
    M2=K^2.*Nd1ps-2*K*y1.*erq.*Nd1+y2.*My2term.*Nd1ms;
    M3=K^3*Nd1ps-3*K^2*y1.*erq.*Nd1+3*K*y2.*My2term.*Nd1ms-y3.*My3term.*Nd1m2s;
    M4=K^4*Nd1ps-4*K^3*y1.*erq.*Nd1+6*K^2*y2.*My2term.*Nd1ms-4*K*y3.*My3term.*Nd1m2s+y4.*My4term.*Nd1m3s;

    dM1=-erq.*Nd1;
    dM2=-2*erq.*(K*Nd1-y1.*dMy1term.*Nd1ms);
    dM3=-3*erq.*(K^2*Nd1-2*K*y1.*dMy1term.*Nd1ms+y2.*dMy2term.*Nd1m2s);
    dM4=-4*erq.*(K^3*Nd1-3*K^2*y1.*dMy1term.*Nd1ms+3*K*y2.*dMy2term.*Nd1m2s-y3.*dMy3term.*Nd1m3s);

    monstrat=dM1+rac*eps^2/2*(dM2-2*M1.*dM1)+rac^2*eps^4/6*(dM3-3*M2.*dM1-3*M1.*dM2+6*M1.^2.*dM1);
    monstrat=monstrat+rac^3*eps^6/24*(dM4-6*M2.*dM2+12*M1.^2.*dM2+24*M1.*M2.*dM1-4*M1.*dM3-4*M3.*dM1-24*M1.^3.*dM1);
    monstrat=(rho*sigmay/sigmax.*y(1:n)./x(1:n)).*(monstrat./beta);

    bankmon=endow+sum(([0 monstrat(1:(n-1))]-monstrat).*x(1:n));  % Calculate portfolio gain for Monoyios strategy

    %
    % Calculate profit and Loss for strategies
    %

    nvPnL=exp(r*T)*(banknv+nvstrat(n)*x(n+1)-max(eta*(y(n+1)-Kprime),0));
    nvPnLs=[nvPnLs nvPnL];
    rmPnL=exp(r*T)*(bankrm+rmstrat(n)*x(n+1)-max(eta*(y(n+1)-Kprime),0));
    rmPnLs=[rmPnLs rmPnL];
    mvPnL=exp(r*T)*(bankmv+prevdeltamv*x(n+1)-max(eta*(y(n+1)-Kprime),0));
    mvPnLs=[mvPnLs mvPnL];
    monPnL=exp(r*T)*(bankmon+monstrat(n)*x(n+1)-max(eta*(y(n+1)-Kprime),0));
    monPnLs=[monPnLs monPnL];

end
toc;

%
% Print out summary stats
%

disp(sprintf('Strategy      Max    Min     Mean     SD     Median'))
disp(sprintf('Naive         %4.2f  %4.2f  %2.4f  %2.4f %2.4f',max(nvPnLs),min(nvPnLs),mean(nvPnLs),std(nvPnLs,1),
      median(nvPnLs)))
disp(sprintf('Local Risk    %4.2f  %4.2f  %2.4f  %2.4f %2.4f',max(rmPnLs),min(rmPnLs),mean(rmPnLs),std(rmPnLs,1),
      median(rmPnLs)))
disp(sprintf('Mean variance %4.2f  %4.2f  %2.4f  %2.4f %2.4f',max(mvPnLs),min(mvPnLs),mean(mvPnLs),std(mvPnLs,1),
      median(mvPnLs)))
disp(sprintf('Monoyios      %4.2f  %4.2f  %2.4f  %2.4f %2.4f',max(monPnLs),min(monPnLs),mean(monPnLs),std(monPnLs,1),
      median(monPnLs)))

return
```

A.3 ApproximationPrice.m

```
%
% Program to generate data for Table 7.4 and produce Figure 7.2 and Figure 7.3.
% This program calls QuadraticMethods.m to generate the data for the graphs.
%

tic;

% Initialize constants

Y0=100;                 % Initial stock price of untraded stock
X0=100;                 % Initial stock price of traded stock
mux=0.10;               % drift rate of untraded stock
muy=0.12;               % drift rate of traded stock
sigmax=0.25;            % volatility of untraded stock
sigmay=0.30;            % volatility of traded stock
r=0.05;                 % NACC rate
eta=-1;                 % 1 for calls, -1 for puts
T=1;                    % maturity of the call in years
K=100;                  % Strike price

%
% Calculate approximation price vs correlation constant
%

rhoarray=[-0.95 -0.75:0.25:0.75 0.95]
ap=[];
```

```
for loop=1:length(rhoarray)
gamma=rhoarray(loop)*sigmay*(mux-r)/sigmax-(muy-r);
d1=(log(Y0/K)+(r-gamma+0.5*sigmay^2)*T)/(sigmay*sqrt(T));
d2=d1-sigmay*sqrt(T);
ap=[ap eta*(Y0*exp(-gamma*T)*cnd(eta*d1)-K*exp(-r*T)*cnd(eta*d2))];
end

ap

%
% Calculate approximation price and SDs for quadratic methods vs rho
%

cc=[];
ap=[];

for rho=-0.99:0.01:0.99
cc=[cc rho];
gamma=rho*sigmay*(mux-r)/sigmax-(muy-r);
d1=(log(Y0/K)+(r-gamma+0.5*sigmay^2)*T)/(sigmay*sqrt(T));
d2=d1-sigmay*sqrt(T);
ap=[ap eta*(Y0*exp(-gamma*T)*cnd(eta*d1)-K*exp(-r*T)*cnd(eta*d2))];
end

rhoarray=[-0.99 -0.975 -0.95 -0.875:0.125:0.875 0.95 0.975 0.99];
stdrm=[];
stdmv=[];

for loop=1:length(rhoarray)
rho=rhoarray(loop)
QuadraticMethods;
stdrm=[stdrm std(rmPnLs)];
stdmv=[stdmv std(mvPnLs)];
end

%
% Produce graph
%

plot(cc,ap,'k','Linewidth',1.5);
hold
plot(rhoarray,stdmv,'r','LineStyle','-.');
plot(rhoarray,stdrm,'b');
set(gca,'TickDir','out','Box','off')
title('Approximation price and SD of P&Ls vs correlation coefficient');
set(get(gca,'Title'),'FontWeight','bold')
xlabel('\rho');
legend('Approximation Price','SD for local risk-minimization','SD for mean-variance',3);
```

A.4 QuadraticMethods.m

```
%
% Program to compare the local risk minimizing stratey with the mean-variance optimal strategy
% and produce a histogram showing the distribution of profits and losses incured.
%

tic;

% Initialize constants

eps=sqrt(1-rho^2);
Kprime=K*exp(-r*T);    % Discounted strike price
n=200;                 % number of rebalance points
ntrials=10000;         % number of trials
randn('state',10);     % Initialize the random seed (for deterministic results)

%
% Calculate some reusable constants and arrays
%

deltaT=T/n;
tau=[T:-deltaT:deltaT];
cumt=[0:deltaT:T];

driftx=(mux-r-0.5*sigmax^2)*cumt;                       % Discounted drift per unit time
drifty=(muy-r-0.5*sigmay^2)*cumt;

varx=sigmax*sqrt(deltaT);                               % Variance per unit time
vary=sigmay*sqrt(deltaT);

gamma=rho*sigmay*(mux-r)/sigmax-(muy-r);                % Calculate BS delta multiplier for RM strategy
deltamult=rho.*sigmay./sigmax;
gammafact=exp(tau.*-gamma);

d1lrm=(0.5*sigmay^2-gamma).*tau;                        % d1 constants
s=sigmay.*sqrt(tau);
```

```
%
% Calculate Black-Scholes price
%

d1=(log(Y0/K)+(r+0.5*sigmay^2)*T)/(sigmay*sqrt(T));
d2=d1-sigmay*sqrt(T);
BSprice=eta*(Y0*cnd(eta*d1)-K*exp(-r*T)*cnd(eta*d2))

%
% Calculate LRM price
%

d1=(log(Y0/K)+(r-gamma+0.5*sigmay^2)*T)/(sigmay*sqrt(T));
d2=d1-sigmay*sqrt(T);
H0=eta*(Y0*exp(-gamma*T)*cnd(eta*d1)-K*exp(-r*T)*cnd(eta*d2))
endow=BSprice;                                                % Initial endowment

%
% Perform numerical experiment
%

rmPnLs=[];
mvPnLs=[];
monPnLs=[];

for loop=1:ntrials

   zx=[0 randn(n,1).'];                                       % Calculate discounted stock price paths
   zy=rho*zx+eps.*[0 randn(n,1).'];
   x=X0*exp(driftx+varx*cumsum(zx));
   y=Y0*exp(drifty+vary*cumsum(zy));

   d1=(log(y(1:n)./Kprime)+d1lrm)./s;                         % Calculate Local Risk Minimization deltas
   d2=d1-s;
   delta=eta*gammafact.*cnd(eta*d1);
   intrinsic=y(1:n).*delta-eta*Kprime.*cnd(eta*d2);          % Calculate discounted Intrinsic values

   rmstrat=delta.*y(1:n).*deltamult./x(1:n);                  % Calculate local risk minimizing hedge ratios
   bankrm=endow+sum(([0 rmstrat(1:(n-1))]-rmstrat).*x(1:n));  % Calculate portfolio gain for rm strategy

   bankmv=endow;                                              % Calculate portfolio gain for mv strategy
   prevdeltamv=0;
   for t=1:n
      newdeltamv=rmstrat(t)+(intrinsic(t)-H0-(bankmv+prevdeltamv*x(t)-endow))*(mux-r)/(sigmax^2*x(t));
      bankmv=bankmv+(prevdeltamv-newdeltamv)*x(t);
      prevdeltamv=newdeltamv;
   end

   %
   % Calculate profit and Loss for strategies
   %

   rmPnL=exp(r*T)*(bankrm+rmstrat(n)*x(n+1)-max(eta*(y(n+1)-Kprime),0));
   rmPnLs=[rmPnLs rmPnL];
   mvPnL=exp(r*T)*(bankmv+prevdeltamv*x(n+1)-max(eta*(y(n+1)-Kprime),0));
   mvPnLs=[mvPnLs mvPnL];

end
toc;

return
```

Bibliography

[1] J.-P. Ansel and C. Stricker, *Couverture des actifs contingents et prix maximum*, Annales de l'Institut Henri Poincaré **30** (1994), no. 2, 303–315.

[2] S. Benninga, T. Björk, and Z. Wiener, *On the use of numeraires in option pricing*, Journal of Derivatives **10** (2002), no. 2, 43–58.

[3] F. Biagini, P. Guasoni, and M. Pratelli, *Mean-variance hedging for stochastic volatility models*, Mathematical Finance **10** (2000), no. 2, 109–123.

[4] K. Bichteler, *Stochastic integration and L^p-theory of semimartingales*, Annals of Probability **9** (1981), no. 1, 49–89.

[5] T. Björk, *Arbitrage Theory in Continuous Time*, 2nd ed., Oxford University Press, 2004.

[6] F. Black and M. Scholes, *The pricing of options and corporate liabilities*, Journal of Political Economy **81** (1973), no. 3, 637–659.

[7] N. Bouleau and D. Lamberton, *Residual risks and hedging strategies in markovian markets*, Stochastic Processes and Their Applications **33** (1989), no. 1, 131–150.

[8] N. Christopeit and M. Musiela, *On the existence and characterization of arbitrage-free measures in contingent claim valuation*, Stochastic Analysis and Applications **12** (1994), no. 1, 41–63.

[9] K. L. Chung and R. J. Williams, *Introduction to Stochastic Integration*, 2nd ed., Birkhäuser, 1990.

[10] K. L. Chung, *A Course in Probability Theory*, 3rd ed., Academic Press, 2001.

[11] D. L. Cohn, *Measure Theory*, Birkhäuser, 1980.

[12] P. Courrège, *Intégrales stochastiques et martingales de carré intégrable*, Séminaire Brélot-Choquet-Dény (Théorie du Potentiel) **7e année 1962/63** (1963), 7–01–7–20.

[13] J. C. Cox and S. A. Ross, *The valuation of options for alternative stochastic processes*, Journal of Financial Economics **3** (1976), no. 1, 145–166.

[14] C. Dalang, A. Morton, and W. Willinger, *Equivalent martingale measures and no arbitrage in stochastic securities market models*, Stochastics and Stochastics Reports **29** (1990), no. 2, 185–201.

[15] M. H. A. Davis, V. G. Panas, and T. Zariphopoulou, *European option pricing with transaction costs*, SIAM Journal on Control and Optimization **31** (1993), no. 2, 470–493.

[16] M. H. A. Davis, *Option pricing in incomplete markets*, Mathematics of Derivative Securities (M. A. H. Dempster and S. R. Pliska, eds.), Cambridge University Press, 1997, pp. 227–254.

[17] ———, *Option valuation and hedging with basis risk*, System Theory: Modeling, Analysis and Control (T. E. Djaferis and I. C. Schuck, eds.), Kluwer, 1999, pp. 245–254.

[18] ———, *Martingale representation and all that*, Festschrift for Professor Pravin Varaiya (E. H. Abed, ed.), Birkhauser, to appear.

[19] F. Delbaen, P. Monat, W. Schachermayer, M. Schweizer, and C. Stricker, *Weighted norm inequalities and hedging in incomplete markets*, Finance and Stochastics **1** (1997), no. 3, 181–227.

[20] F. Delbaen and W. Schachermayer, *A general version of the fundamental theorem of asset pricing*, Mathematische Annalen **300** (1994), no. 1, 463–520.

[21] ———, *The variance-optimal martingale measure for continuous processes*, Bernoulli **2** (1996), no. 1, 81–105.

[22] ———, *The fundamental theorem of asset pricing for unbounded stochastic processes*, Mathematische Annalen **312** (1998), no. 2, 215–250.

[23] F. Delbaen, *Representing martingale measures when asset prices are continuous and bounded*, Mathematical Finance **2** (1992), no. 2, 107–130.

[24] C. Dellacherie and P.-A. Meyer, *Probabilities and Potential B*, North Holland, 1980.

[25] C. Dellacherie, *Un survol de la theorie de l'intégrale stochastique*, Stochastic Processes and Their Applications **10** (1980), no. 2, 115–144.

[26] J. L. Doob, *Stochastic Processes*, John Wiley and Sons, New York, 1953.

[27] D. Duffie and H. R. Richardson, *Mean-variance hedging in continuous time*, Annals of Applied Probability **1** (1991), no. 1, 1–15.

[28] R. Durrett, *Stochastic Calculus: A Practical Introduction*, CRC Press, 1996.

[29] R. J. Elliott, *Stochastic Calculus and Applications*, Springer-Verlag, 1982.

[30] M. Emery, *Compensation de processus à variation finie non localement intégrables*, Séminaire de Probabilités XIV, Lecture Notes in Mathematics, vol. 784, Springer-Verlag, 1980, pp. 152–160.

[31] H. Föllmer and M. Schweizer, *Hedging by sequential regression: An introduction to the mathematics of option trading*, ASTIN Bulletin **19** (1989), no. S, 147–160.

[32] ______, *Hedging of contingent claims under incomplete information*, Applied Stochastic Analysis (M. H. A. Davis and R. J. Elliott, eds.), Gorden and Breach Science Publishers, 1991, pp. 389–414.

[33] H. Föllmer and D. Sondermann, *Hedging of non-redundant contingent claims*, Contributions to Mathematical Economics (W. Hildenbrand and A. Mas-Colell, eds.), North-Holland, 1986, pp. 205–223.

[34] M. Frittelli, *The minimal entropy martingale measure and the valuation problem in incomplete markets*, Mathematical Finance **10** (2000), no. 1, 39–52.

[35] L. I. Galtchouk, *Représentation des martingales engendrées par un processus à accroissements indépendants (cas des martingales de carré intégrable)*, Annales de l'Institut Henri Poincaré (B) **12** (1976), no. 3, 199–211.

[36] H. Geman, N. El Karoui, and J.-C. Rochet, *Changes of numéraire, changes of probability measure and option pricing*, Journal of Applied Probability **32** (1995), no. 2, 443–458.

[37] C. Gouriéroux, J. P. Laurent, and H. Pham, *Mean-variance hedging and numéraire*, Mathematical Finance **8** (1998), no. 3, 179–200.

[38] P. Grandits and T. Rheinländer, *On the minimal entropy martingale measure*, Annals of Probability **30** (2002), no. 3, 1003–1038.

[39] J. M. Harrison and D. M. Kreps, *Martingales and arbitrage in multiperiod securities markets*, Journal of Economic Theory **20** (1979), no. 3, 381–408.

[40] J. M. Harrison and S. R. Pliska, *Martingales and stochastic integrals in the theory of continuous trading*, Stochastic Processes and Their Applications **11** (1981), no. 3, 215–260.

[41] ______, *A stochastic calculus model of continuous trading: Complete markets*, Stochastic Processes and Their Applications **15** (1983), no. 3, 313–316.

[42] D. Heath, E. Platen, and M Schweizer, *Comparison of some key approaches to hedging in incomplete markets*, Proceedings: Risk Conference on Computational and Quantitative Finance, New York, 10–11 September 1998.

[43] D. Heath, E. Platen, and M. Schweizer, *A comparison of two quadratic approaches to hedging in incomplete markets*, Mathematical Finance **11** (2001), no. 4, 385–413.

[44] ______, *Numerical comparison of local risk-minimisation and mean-variance hedging*, Option Pricing, Interest Rates and Risk Management (E. Jouini, J. Cvitanić, and M. Musiela, eds.), Cambridge University Press, 2001, pp. 509–537.

[45] V. Henderson and D. G. Hobson, *Substitute hedging*, Risk **15** (2002), 71–75, (May edition).

[46] C. Hipp, *Hedging general claims*, Proceedings of the 3rd AFIR Colloquium, Rome, vol. 2, 1993, pp. 603–613.

[47] K. Itô, *Stochastic integral*, Proceedings of the Imperial Academy of Tokyo **20** (1944), 519–524.

[48] ______, *Multiple Wiener integral*, Journal of the Mathematical Society of Japan **3** (1951), no. 1, 157–169.

[49] ______, *On a formula concerning stochastic differentials*, Nagoya Mathematical Journal **3** (1951), 55–65.

[50] J. Jacod and P. Protter, *Probability Essentials*, 2nd ed., Springer-Verlag, 2003.

[51] J. Jacod and A. N. Shiryaev, *Limit Theorems for Stochastic Processes*, 2nd ed., Springer-Verlag, 2003.

[52] J. Jacod and M. Yor, *Étude de solutions extrémales et représentation intégrale des solutions pour certains problèmes de martingales*, Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete **38** (1977), no. 2, 83–125.

[53] J. Jacod, *Calcul stochastique et problèmes de martingales*, Lecture Notes in Mathematics, vol. 714, Springer-Verlag, 1979.

[54] R. Jarrow and P. Protter, *A short history of stochastic integration and mathematical finance: The early years, 1880–1970*, The Herman Rubin Festschrift, IMS Lecture Notes, vol. 45, 2004, pp. 75–91.

[55] I. Karatzas and S. E. Shreve, *Brownian Motion and Stochastic Calculus*, 2nd ed., Springer-Verlag, 1991.

[56] I. J. Karazas, J. P. Lehoczky, S. E. Shreve, and G.-L. Xu, *Martingale and duality methods for utility maximization in an incomplete market*, SIAM Journal on Control and Optimization **29** (1991), no. 3, 702–730.

[57] D. M. Kreps, *Arbitrage and equilibrium in economics with infinitely many commodities*, Journal of Mathematical Economics **8** (1981), no. 1, 15–35.

[58] E. Kreyszig, *Introductory Functional Analysis with Applications*, John Wiley & Sons, 1978.

[59] H. Kunita and S. Watanabe, *On square integrable martingales*, Nagoya Mathematical Journal **30** (1967), 209–245.

[60] D. Lamberton and B. Lapeyre, *Hedging index options with few assets*, Mathematical Finance **3** (1993), no. 1, 25–41.

[61] D. Lamberton, H. Pham, and M. Schweizer, *Local risk-minimization under transaction costs*, Mathematics of Operations Research **23** (1998), no. 3, 585–612.

[62] D. G. Luenberger, *Optimization by Vector Space Methods*, John Wiley & Sons, 1969.

[63] F. Mercurio and T. Vorst, *Option pricing and hedging in discrete time with transaction costs and incomplete markets*, Mathematics of Derivative Securities (M. A. H. Dempster and S. R. Pliska, eds.), Cambridge University Press, 1997, pp. 227–254.

[64] R. C. Merton, *Theory of rational option pricing*, Bell Journal of Economics and Managment Science **4** (1973), no. 1, 141–183.

[65] P.-A. Meyer, *A decomposition theorem for supermartingales*, Illinois Journal of Mathematics **6** (1962), no. 2, 193–205.

[66] ———, *Decomposition of supermartingales: The uniqueness theorem*, Illinois Journal of Mathematics **7** (1963), no. 1, 1–17.

[67] P. Monat and C. Stricker, *Föllmer-Schweizer decomposition and mean-variance hedging for general claims*, Annals of Probability **23** (1995), no. 2, 605–628.

[68] M. Monoyios, *Performance of utility-based strategies for hedging basis risk*, Quantitative Finance **4** (2004), no. 3, 245–255.

[69] M. Musiela and M. Rutkowski, *Martingale Methods in Financial Modelling*, Springer-Verlag, 1997.

[70] H. Pham, T. Rheinländer, and M. Schweizer, *Mean-variance hedging for continuous processes: New proofs and examples*, Finance and Stochastics **2** (1998), no. 2, 173–198.

[71] H Pham, *On quadratic hedging in continuous time*, Mathematical Methods of Operations Research **51** (2000), no. 2, 315–339.

[72] P. Protter, *Stochastic integration without tears (with apology to P. A. Meyer)*, Stochastics **16** (1986), no. 3-4, 295–325.

[73] ———, *A partial introduction to financial asset pricing theory*, Stochastic Processes and Their Applications **91** (2001), no. 2, 169–203.

[74] ———, *Stochastic Integration and Differential Equations*, 2nd ed., Springer-Verlag, 2004.

[75] D. Revuz and M. Yor, *Continuous Martingales and Brownian Motion*, 3rd ed., Springer-Verlag, 1999.

[76] T. Rheinländer and M. Schweizer, *On $\mathcal{L}^2$-projections on a space of stochastic integrals*, Annals of Probability **25** (1997), no. 4, 1810–1831.

[77] T. Rheinländer, *Optimal Martingale Measures and Their Applications in Mathematical Finance*, Ph.D. thesis, Technische Universität Berlin, 1999.

[78] L. C. G. Rogers and D. Williams, *Diffusions, Markov Processes and Martingales*, 2nd ed., vol. 1, Cambridge University Press, 1994.

[79] ———, *Diffusions, Markov Processes and Martingales*, 2nd ed., vol. 2, Cambridge University Press, 1994.

[80] S. A. Ross, *The arbitrage theory of capital asset pricing*, Journal of Economic Theory **13** (1976), no. 3, 341–360.

[81] W. Schachermayer, *Introduction to the mathematics of financial markets*, Lectures on Probability Theory and Statistics, Saint-Flour Summer School 2000 (P. Bernard, ed.), Lecture Notes in Mathematics, vol. 1816, Springer-Verlag, 2003, pp. 111–177.

[82] M. Schäl, *On quadratic cost criteria for option hedging*, Mathematics of Operations Research **19** (1994), no. 1, 121–131.

[83] M. Schweizer, *Hedging of Options in a General Semimartingale Model*, Ph.D. thesis (8615), ETH Zürich, 1988.

[84] ———, *Risk-minimality and orthogonality of martingales*, Stochastics and Stochastics Reports **30** (1990), no. 2, 123–131.

[85] ———, *Option hedging for semimartingales*, Stochastic Processes and Their Applications **37** (1991), no. 2, 339–363.

[86] ———, *Martingale densities for general asset prices*, Journal of Mathematical Economics **21** (1992), no. 4, 363–378.

[87] ———, *Mean-variance hedging for general claims*, Annals of Applied Probability **2** (1992), no. 1, 171–179.

[88] ———, *Approximating random variables by stochastic integrals*, Annals of Applied Probability **22** (1994), no. 3, 1536–1575.

[89] ———, *On the minimal martingale measure and the Föllmer-Schweizer decomposition*, Stochastic Analysis and Applications **13** (1995), no. 5, 573–599.

[90] ———, *Approximation pricing and the variance-optimal martingale measure*, Annals of Probability **24** (1996), no. 1, 206–236.

[91] ———, *A guided tour through quadratic hedging approaches*, Option Pricing, Interest Rates and Risk Management (E. Jouini, J. Cvitanić, and M. Musiela, eds.), Cambridge University Press, 2001, pp. 538–574.

[92] A. N. Shiryaev, *Probability*, 2nd ed., Springer-Verlag, 1995.

[93] ______, *Essentials of Stochastic Finance: Facts, Models, Theory*, Advanced Series on Statistical Science & Applied Probability, vol. 3, World Scientific, 1999.

[94] S. E. Shreve, *Stochastic Calculus for Finance II: Continuous-Time Models*, Springer-Verlag, 2004.

[95] D. Williams, *Probability with Martingales*, Cambridge University Press, 1991.

[96] T. Zariphopoulou, *A solution approach to valuation with unhedgeable risks*, Finance and Stochastics **5** (2001), no. 1, 61–82.