To Patrick.
Choose well.

Contents

Preface to revised edition

Since this book first appeared in 2000, it has been adopted by a number of universities as a standard text for a graduate course in finance. As a result we have produced this revised edition. The only differences in content between this text and its predecessor are the inclusion of an additional chapter of exercises with solutions, and the corrections of a number of errors.

Many of the exercise are variants of ones given to students of the M.Sc. in Mathematical Finance at the University of Warwick, so they have been tried and tested. Many provide drill in routine calculations in the interest-rate setting.

Since the book first appeared there has been further development in the modelling of interest rate derivatives. The modelling and approximation of market models has progressed further, as has that of Markov-functional models that are now used, in their multi-factor form, in a number of banks. A notable exclusion from this revised edition is any coverage of these advances. Those interested in this area can find some of this in Rebonato (2002) and Bennett and Kennedy (2004).

A few extra acknowledgements are due in this revised edition: to Noel Vaillant and Jørgen Aase Nielsen for pointing out a number of errors, and to Stuart Price for typing the exercise chapter and solutions.

Preface

The growth in the financial derivatives market over the last thirty years has been quite extraordinary. From virtually nothing in 1973, when Black, Merton and Scholes did their seminal work in the area, the total outstanding notional value of derivatives contracts today has grown to several trillion dollars. This phenomenal growth can be attributed to two factors.

The first, and most important, is the natural need that the products fulfil. Any organization or individual with sizeable assets is exposed to moves in the world markets. Manufacturers are susceptible to moves in commodity prices; multinationals are exposed to moves in exchange rates; pension funds are exposed to high inflation rates and low interest rates. Financial derivatives are products which allow all these entities to reduce their exposure to market moves which are beyond their control.

The second factor is the parallel development of the financial mathematics needed for banks to be able to price and hedge the products demanded by their customers. The breakthrough idea of Black, Merton and Scholes, that of pricing by arbitrage and replication arguments, was the start. But it is only because of work in the field of pure probability in the previous twenty years that the theory was able to advance so rapidly. Stochastic calculus and martingale theory were the perfect mathematical tools for the development of financial derivatives, and models based on Brownian motion turned out to be highly tractable and usable in practice.

Where this leaves us today is with a massive industry that is highly dependent on mathematics and mathematicians. These mathematicians need to be familiar with the underlying theory of mathematical finance and they need to know how to apply that theory in practice. The need for a text that addressed both these needs was the original motivation for this book. It is aimed at both the mathematical practitioner and the academic mathematician with an interest in the real-world problems associated with financial derivatives. That is to say, we have written the book that we would both like to have read when we first started to work in the area.

This book is divided into two distinct parts which, apart from the need to cross-reference for notation, can be read independently. Part I is devoted to the theory of mathematical finance. It is not exhaustive, notable omissions being a treatment of asset price processes which can exhibit jumps, equilibrium theory and the theory of optimal control (which underpins the pricing of American options). What we have included is the basic theory for continuous asset price processes with a particular emphasis on the martingale approach to arbitrage pricing.

The primary development of the finance theory is carried out in Chapters 1 and 7. The reader who is not already familiar with the subject but who has a solid grounding in stochastic calculus could learn most of the theory from these two chapters alone. The fundamental ideas are laid out in Chapter 1, in the simple setting of a single-period economy with only finitely many states. The full (and more technical) continuous time theory, is developed in Chapter 7. The treatment here is slightly non-standard in the emphasis it places on numeraires (which have recently become extremely important as a modelling tool) and the (consequential) development of the L^1 theory rather than the more common L^2 version. We also choose to work with the filtration generated by the assets in the economy rather than the more usual Brownian filtration. This approach is certainly more natural but, surprisingly, did not appear in the literature until the work of Babbs and Selby (1998).

An understanding of the continuous theory of Chapter 7 requires a knowledge of stochastic calculus, and we present the necessary background material in Chapters 2–6. We have gathered together in one place the results from this area which are relevant to financial mathematics. We have tried to give a full and yet readable account, and hope that these chapters will stand alone as an accessible introduction to this technical area. Our presentation has been very much influenced by the work of David Williams, who was a driving force in Cambridge when we were students. We also found the books of Chung and Williams (1990), Durrett (1996), Karatzas and Shreve (1991), Protter (1990), Revuz and Yor (1991) and Rogers and Williams (1987) to be very illuminating, and the informed reader will recognize their influence.

The reader who has read and absorbed the first seven chapters of the book will be well placed to read the financial literature. The last part of theory we present, in Chapter 8, is more specialized, to the field of interest rate models. The exposition here is partly novel but borrows heavily from the papers by Baxter (1997) and Jin and Glasserman (1997). We describe several different ways present in the literature for specifying an interest rate model and show how they relate to one another from a mathematical perspective. This includes a discussion of the celebrated work of Heath, Jarrow and Morton (1992) (and the less celebrated independent work of Babbs (1990)) which, for the first time, defined interest rate models directly in terms of forward rates.

Part II is very much about the practical side of building models for pricing derivatives. It, too, is far from exhaustive, covering only topics which directly reflect our experiences through our involvement in product development within the London interest rate derivative market. What we have tried to do in this part of the book, through the particular problems and products that we discuss, is to alert the reader to the issues involved in derivative pricing in practice and to give him a framework within which to make his own judgements and a platform from which to develop further models.

Chapter 9 sets the scene for the remainder of the book by identifying some basic issues a practitioner should be aware of when choosing and applying models. Chapters 10 and 11 then introduce the reader to the basic instruments and terminology and to the pricing of standard vanilla instruments using swaption measure. This pricing approach comes from fairly recent papers in the area which focus on the use of various assets as numeraires when defining models and doing calculations. This is actually an old idea dating back at least to Harrison and Pliska (1981) and which, starting with the work of Geman *et al.* (1995), has come to the fore over the past decade. Chapter 12 is on futures contracts. The treatment here draws largely on the work of Duffie and Stanton (1992), though we have attempted a cleaner presentation of this standard topic than those we have encountered elsewhere.

The remainder of Part II tackles pricing problems of increasing levels of complexity, beginning with single-currency European products and finishing with various Bermudan callable products. Chapters 13–16 are devoted to European derivatives. Chapters 13 and 15 present a new approach to this much-studied problem. These products are theoretically straightforward but the challenge for the practitioner is to ensure the model he employs in practice is well calibrated to market-implied distributions and is easy to implement. Chapter 14 discusses convexity corrections and the pricing of 'convexity-related' products using the ideas of earlier chapters. These are important products which have previously been studied by, amongst others, Coleman (1995) and Doust (1995). We provide explicit formulae for some commonly met products and then, in Chapter 16, generalize these to multi-currency products.

The last three chapters focus on the pricing of path-dependent and American derivatives. Short-rate models have traditionally been those favoured for this task because of their ease of implementation and because they are relatively easy to understand. There is a vast literature on these models, and Chapter 17 provides merely a brief introduction. The Vasicek-Hull-White model is singled out for more detailed discussion and an algorithm for its implementation is given which is based on a semi-analytic approach (taken from Gandhi and Hunt (1997)).

More recently attention has turned to the so-called market models, pioneered by Brace *et al.* (1997) and Miltersen *et al.* (1997), and extended by Jamshidian (1997). These models provided a breakthrough in tackling the issue of model calibration and, in the few years since they first appeared, a vast literature has developed around them. They, or variants of and approximations to them, are now starting to replace previous (short-rate) models within most major investment banks. Chapter 18 provides a basic description of both the LIBOR- and swap-based market models. We have not attempted to survey the literature in this extremely important area and for this we refer the reader instead to the article of Rutkowski (1999).

The book concludes, in Chapter 19, with a description of some of our own recent work (jointly with Antoon Pelsser), work which builds on Hunt and Kennedy (1998). We describe a class of models which can fit the observed prices of liquid instruments in a similar fashion to market models but which have the advantage that they can be efficiently implemented. These models, which we call Markov-functional models, are especially useful for the pricing of products with multi-callable exercise features, such as Bermudan swaptions or Bermudan callable constant maturity swaps. The exposition here is similar to the original working paper which was first circulated in late 1997 and appeared on the Social Sciences Research Network in January 1998. A précis of the main ideas appeared in *RISK Magazine* in March 1998. The final paper, Hunt, Kennedy and Pelsser (2000), is to appear soon. Similar ideas have more recently been presented by Balland and Hughston (2000).

We hope you enjoy reading the finished book. If you learn from this book even one-tenth of what we learnt from writing it, then we will have succeeded in our objectives.

Phil Hunt
Joanne Kennedy
31 December 1999

Acknowledgements

There is a long list of people to whom we are indebted and who, either directly or indirectly, have contributed to making this book what it is. Some have shaped our thoughts and ideas; others have provided support and encouragement; others have commented on drafts of the text. We shall not attempt to produce this list in its entirety but our sincere thanks go to all of you.

Our understanding of financial mathematics and our own ideas in this area owe much to two practitioners, Sunil Gandhi and Antoon Pelsser. PH had the good fortune to work with Sunil when they were both learning the subject at NatWest Markets, and with Antoon at ABN AMRO Bank. Both have had a marked impact on this book.

The area of statistics has much to teach about the art of modelling. Brian Ripley is a master of this art who showed us how to look at the subject through new eyes. Warmest and sincere thanks are also due to another colleague from JK's Oxford days, Peter Clifford, for the role he played when this book was taking shape.

We both learnt our probability while doing PhDs at the University of Cambridge. We were there at a particularly exciting time for probability, both pure and applied, largely due to the combined influences of Frank Kelly and David Williams. Their contributions to the field are already well known. Our thanks for teaching us some of what you know, and for teaching us how to use it.

Finally, our thanks go to our families, who provided us with the support, guidance and encouragement that allowed us to pursue our chosen careers. To Shirley Collins, Harry Collins and Elizabeth Kennedy: thank you for being there when it mattered. But most of all, to our parents: for the opportunities given to us at sacrifice to yourselves.

Part I
Theory

1

Single-Period Option Pricing

1.1 OPTION PRICING IN A NUTSHELL

To introduce the main ideas of option pricing we first develop this theory in the case when asset prices can take on only a finite number of values and when there is only one time-step. The continuous time theory which we introduce in Chapter 7 is little more than a generalization of the basic ideas introduced here.

The two key concepts in option pricing are *replication* and *arbitrage*. Option pricing theory centres around the idea of replication. To price any derivative we must find a portfolio of assets in the economy, or more generally a trading strategy, which is guaranteed to pay out *in all circumstances* an amount identical to the payout of the derivative product. If we can do this we have exactly replicated the derivative.

This idea will be developed in more detail later. It is often obscured in practice when calculations are performed 'to a formula' rather than from first principles. It is important, however, to ensure that the derivative being priced can be reproduced by trading in other assets in the economy.

An arbitrage is a trading strategy which generates profits from nothing with no risk involved. Any economy we postulate must not allow arbitrage opportunities to exist. This seems natural enough but it is essential for our purposes, and Section 1.3.2 is devoted to establishing necessary and sufficient conditions for this to hold.

Given the absence of arbitrage in the economy it follows immediately that the value of a derivative is the value of a portfolio that replicates it. To see this, suppose to the contrary that the derivative costs more than the replicating portfolio (the converse can be treated similarly). Then we can sell the derivative, use the proceeds to buy the replicating portfolio and still be left with some free cash to do with as we wish. Then all we need do is use the income from the portfolio to meet our obligations under the derivative contract. This is an arbitrage opportunity – and we know they do not exist.

All that follows in this book is built on these simple ideas!

1.2 THE SIMPLEST SETTING

Throughout this section we consider the following simple situation. There are two assets, $A^{(1)}$ and $A^{(2)}$, with prices at time zero $A_0^{(1)}$ and $A_0^{(2)}$. At time 1 the economy is in one of two possible states which we denote by ω_1 and ω_2. We denote by $A_1^{(i)}(\omega_j)$ the price of asset i at time 1 if the economy is in state ω_j. Figure 1.1 shows figuratively the possibilities. We shall denote a portfolio of assets by $\phi = (\phi^{(1)}, \phi^{(2)})$ where $\phi^{(i)}$, the holding of asset i, could be negative.

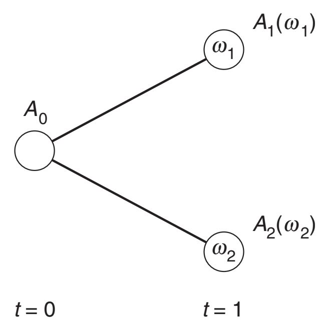

Figure 1.1 Possible states of the economy

We wish to price a derivative (of $A^{(1)}$ and $A^{(2)}$) which pays at time 1 the amount $X_1(\omega_j)$ if the economy is in state ω_j. We will do this by constructing a replicating portfolio. Once again, this is the *only* way that a derivative can be valued, although, as we shall see later, explicit knowledge of the replicating portfolio is not always necessary.

The first step is to check that the economy we have constructed is arbitrage-free, meaning we cannot find a way to generate money from nothing. More precisely, we must show that there is no portfolio $\phi = (\phi^{(1)}, \phi^{(2)})$ with one of the following (equivalent) conditions holding:

(i) $\phi^{(1)}A_0^{(1)} + \phi^{(2)}A_0^{(2)} < 0, \quad \phi^{(1)}A_1^{(1)}(\omega_j) + \phi^{(2)}A_1^{(2)}(\omega_j) \geq 0, \quad j = 1, 2,$

(ii) $\phi^{(1)}A_0^{(1)} + \phi^{(2)}A_0^{(2)} \leq 0, \quad \phi^{(1)}A_1^{(1)}(\omega_j) + \phi^{(2)}A_1^{(2)}(\omega_j) \geq 0, \quad j = 1, 2,$
where the second inequality is strict for some j.

Suppose there exist $\phi^{(1)}$ and $\phi^{(2)}$ such that

$$\phi^{(1)}A_1^{(1)}(\omega_j) + \phi^{(2)}A_1^{(2)}(\omega_j) = X_1(\omega_j), \quad j = 1, 2. \tag{1.1}$$

We say that ϕ is a replicating portfolio for X. If we hold this portfolio at time zero, at time 1 the value of the portfolio will be exactly the value of X, no matter which of the two states the economy is in. It therefore follows that a fair price for the derivative X (of $A^{(1)}$ and $A^{(2)}$) is precisely the cost of this portfolio, namely

$$X_0 = \phi^{(1)} A_0^{(1)} + \phi^{(2)} A_0^{(2)}.$$

Subtleties

The above approach is exactly the one we use in the more general situations later. There are, however, three potential problems:

(i) Equation (1.1) may have no solution.
(ii) Equation (1.1) may have (infinitely) many solutions yielding the same values for X_0.
(iii) Equation (1.1) may have (infinitely) many solutions yielding different values of X_0.

Each of these has its own consequences for the problem of derivative valuation.

(i) If (1.1) has no solution, we say the economy is *incomplete*, meaning that it is possible to introduce further assets into the economy which are not redundant and cannot be replicated. Any such asset is not a derivative of existing assets and cannot be priced by these methods.
(ii) If (1.1) has many solutions all yielding the same value X_0 then there exists a portfolio $\psi \neq 0$ such that

$$\psi \cdot A_0 = 0, \quad \psi \cdot A_1(\omega_j) = 0, \quad j = 1, 2.$$

This is not a problem and means our original assets are not all independent – one of these is a derivative of the others, so any further derivative can be replicated in many ways.
(iii) If (1.1) has many solutions yielding different values for X_0 we have a problem. There exist portfolios ϕ and ψ such that

$$(\phi - \psi) \cdot A_0 < 0$$
$$(\phi - \psi) \cdot A_1(\omega_j) = \big(X_1(\omega_j) - X_1(\omega_j)\big) = 0, \qquad j = 1, 2.$$

This is an arbitrage, a portfolio with strictly negative value at time zero and zero value at time 1.

In this final case our initial economy was poorly defined. Such situations can occur in practice but are not sustainable. From a derivative pricing viewpoint such situations must be excluded from our model. In the presence of arbitrage there is no unique fair price for a derivative, and in the absence of arbitrage the derivative value is given by the initial value of *any* replicating portfolio.

1.3 GENERAL ONE-PERIOD ECONOMY

We now develop in more detail all the concepts and ideas raised in the previous section. We restrict attention once again to a single-period economy for clarity, but introduce many assets and many states so that the essential structure and techniques used in the continuous time setting can emerge and be discussed.

We now consider an economy $\mathcal{E}$ comprising n assets with m possible states at time 1. Let Ω be the set of all possible states. We denote, as before, the individual states by $\omega_j, j = 1, 2, \ldots, m$, and the asset prices by $A_0^{(i)}$ and $A_1^{(i)}(\omega_j)$. We begin with some definitions.

Definition 1.1 *The economy $\mathcal{E}$ admits arbitrage if there exists a portfolio ϕ such that one of the following conditions (which are actually equivalent in this discrete setting) holds:*

(i) $\phi \cdot A_0 < 0$ *and* $\phi \cdot A_1(\omega_j) \geq 0$ for all j,
(ii) $\phi \cdot A_0 \leq 0$ *and* $\phi \cdot A_1(\omega_j) \geq 0$ for all j, *with strict inequality for some* j.

If there is no such ϕ then the economy is said to be arbitrage-free.

Definition 1.2 *A derivative X is said to be attainable if there exists some ϕ such that*

$$X_1(\omega_j) = \phi \cdot A_1(\omega_j) \quad \text{for all } j.$$

We have seen that an arbitrage-free economy is essential for derivative pricing theory. We will later derive conditions to check whether a given economy is indeed arbitrage-free. Here we will quickly move on to show, in the absence of arbitrage, how products can be priced. First we need one further definition.

Definition 1.3 *A pricing kernel Z is any strictly positive vector with the property that*

$$A_0 = \sum_j Z_j A_1(\omega_j). \tag{1.2}$$

The reason for this name is the role Z plays in derivative pricing as summarized by Theorem 1.4. It also plays an important role in determining whether or not an economy admits arbitrage, as described in Theorem 1.7.

1.3.1 Pricing

One way to price a derivative in an arbitrage-free economy is to construct a replicating portfolio and calculate its value. This is summarized in the first part of the following theorem. The second part of the theorem enables us to price a derivative without explicitly constructing a replicating portfolio, as long as we know one exists.

Theorem 1.4 *Suppose that the economy $\mathcal{E}$ is arbitrage-free and let X be an attainable contingent claim, i.e. a derivative which can be replicated with other assets. Then the fair value of X is given by*

$$X_0 = \phi \cdot A_0 \tag{1.3}$$

where ϕ *solves*

$$X_1(\omega_j) = \phi \cdot A_1(\omega_j) \quad \text{for all } j.$$

Furthermore, if Z *is some pricing kernel for the economy then* X_0 *can also be represented as*

$$X_0 = \sum_j Z_j X_1(\omega_j).$$

Proof: The first part of the result follows by the arbitrage arguments used previously. Moving on to the second statement, substituting the defining equation (1.2) for Z into (1.3) yields

$$X_0 = \phi \cdot A_0 = \sum_j Z_j \big(\phi \cdot A_1(\omega_j)\big) = \sum_j Z_j X_1(\omega_j).$$

□

Remark 1.5: Theorem 1.4 shows how the pricing of a derivative can be reduced to calculating a pricing kernel Z. This is of limited value as it stands since we still need to show that the economy is arbitrage-free and that the derivative in question is attainable. This latter step can be done by either explicitly constructing the replicating portfolio or proving beforehand that *all* derivatives (or at least all within some suitable class) are attainable. If all derivatives are attainable then the economy is said to be *complete*. As we shall shortly see, both the problems of establishing no arbitrage and completeness for an economy are in themselves intimately related to the pricing kernel Z.

Remark 1.6: Theorem 1.4 is essentially a geometric result. It merely states that if X_0 and $X_1(\omega_j)$ are the projections of A_0 and $A_1(\omega_j)$ in some direction ϕ, and if A_0 is an affine combination of the vectors $A_1(\omega_j)$, then X_0 is the same affine combination of the vectors $X_1(\omega_j)$. Figure 1.2 illustrates this for the case $n = 2$, a two-asset economy.

1.3.2 Conditions for no arbitrage: existence of Z

The following result gives necessary and sufficient conditions for an economy to be arbitrage-free. The result is essentially a geometric one. However, its real importance as a tool for calculation becomes clear in the continuous time setting of Chapter 7 where we work in the language of the probabilist. We will see the result rephrased probabilistically later in this chapter.

Theorem 1.7 *The economy* $\mathcal{E}$ *is arbitrage-free if and only if there exists a pricing kernel, i.e. a strictly positive* Z *such that*

$$A_0 = \sum_j Z_j A_1(\omega_j).$$

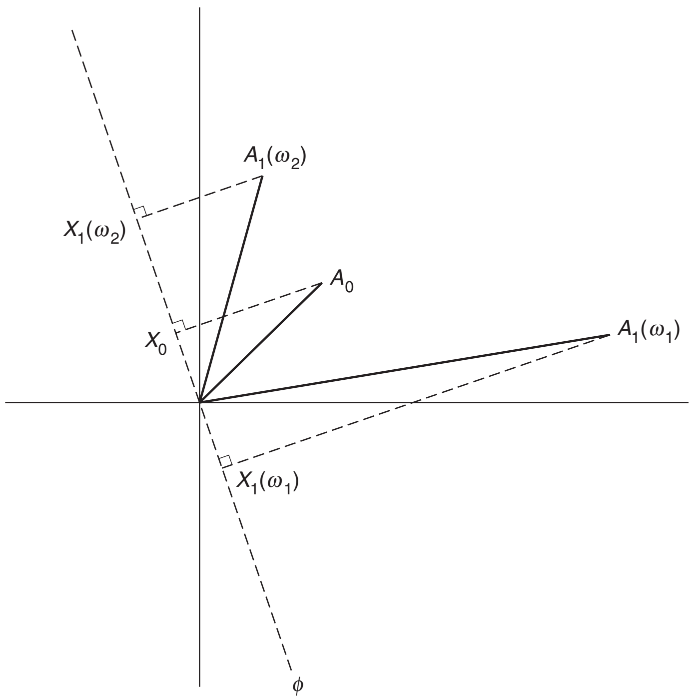

Figure 1.2 Geometry of option pricing

Proof: Suppose such a Z exists. Then for any portfolio ϕ,

$$\phi \cdot A_0 = \phi \cdot \Big(\sum_j Z_j A_1(\omega_j)\Big) = \sum_j Z_j \big(\phi \cdot A_1(\omega_j)\big). \tag{1.4}$$

If ϕ is an arbitrage portfolio then the left-hand side of (1.4) is non-positive, and the right-hand side is non-negative. Hence (1.4) is identically zero. This contradicts ϕ being an arbitrage portfolio and so $\mathcal{E}$ is arbitrage-free.

Conversely, suppose no such Z exists. We will in this case construct an arbitrage portfolio. Let C be the convex cone constructed from $A_1(\cdot)$,

$$C = \big\{a : a = \sum_j Z_j A_1(\omega_j), Z \gg 0\big\}.$$

The set C is a non-empty convex set *not* containing A_0. Here $Z \gg 0$ means that all components of Z are greater than zero. Hence, by the separating hyperplane theorem there exists a hyperplane $H = \{x : \phi \cdot x = \beta\}$ that separates A_0 and C,

$$\phi \cdot A_0 \leq \beta \leq \phi \cdot a \quad \text{for all } a \in C.$$

The vector ϕ represents an arbitrage portfolio, as we now demonstrate.

First, observe that if $a \in \bar{C}$ (the closure of C) then $\mu a \in \bar{C}$, $\mu \geq 0$, and

$$\beta \leq \mu(\phi \cdot a).$$

Taking $\mu = 0$ yields $\beta \leq 0$, $\phi \cdot A_0 \leq 0$. Letting $\mu \uparrow \infty$ shows that $\phi \cdot a \geq 0$ for all $a \in \bar{C}$, in particular $\phi \cdot A_1(\omega_j) \geq 0$ for all j. So we have that

$$\phi \cdot A_0 \leq 0 \leq \phi \cdot A_1(\omega_j) \quad \text{for all } j, \tag{1.5}$$

and it only remains to show that (1.5) is not always identically zero. But in this case $\phi \cdot A_0 = \phi \cdot C = 0$ which violates the separating property for H. $\square$

Remark 1.8: Theorem 1.7 states that A_0 must be in the interior of the convex cone created by the vectors $A_1(\omega_j)$ for there to be no arbitrage. If this is not the case an arbitrage portfolio exists, as in Figure 1.3.

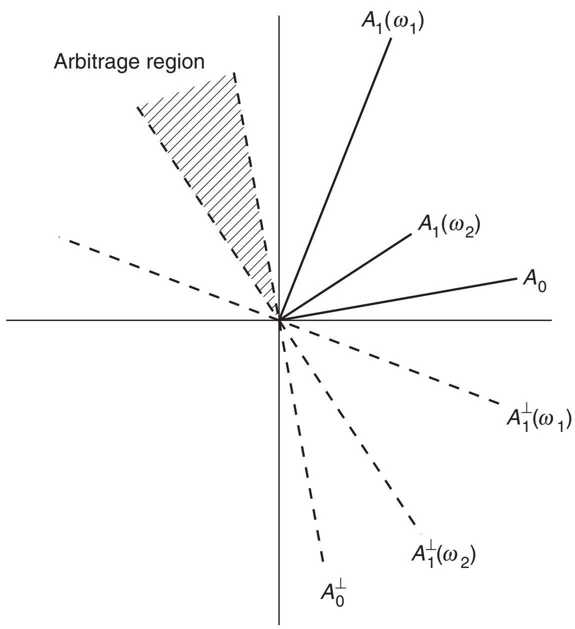

Figure 1.3 An arbitrage portfolio

1.3.3 Completeness: uniqueness of Z

If we are in a complete market all contingent claims can be replicated. As we have seen in Theorem 1.4, this enables us to price derivatives without having to explicitly calculate the replicating portfolio. This can be useful, particularly for the continuous time models introduced later. In our finite economy it is clear what conditions are required for completeness. There must be at least as many assets as states of the economy, i.e. $m \leq n$, and these must span an appropriate space.

Theorem 1.9 *The economy $\mathcal{E}$ is complete if and only if there exists a (generalized) left inverse A^{-1} for the matrix A where*

$$A_{ij} = A_1^{(i)}(\omega_j).$$

Equivalently, $\mathcal{E}$ is complete if and only if there exists no non-zero solution to the equation $Ax = 0$.

Proof: For the economy to be complete, given any contingent claim X, we must be able to solve

$$X_1(\omega_j) = \phi \cdot A_1(\omega_j) \quad \text{for all } j,$$

which can be written as

$$X_1 = A^T \phi. \tag{1.6}$$

The existence of a solution to (1.6) for all X_1 is exactly the statement that A^T has a right inverse (some matrix B such that $A^T B : \mathbb{R}^m \to \mathbb{R}^m$ is the identity matrix) and the replicating portfolio is then given by

$$\phi = (A^T)^{-1} X_1.$$

This is equivalent to the statement that A has full rank m and there being no non-zero solution to $Ax = 0$. □

Remark 1.10: If there are more assets than states of the economy the hedge portfolio will not in general be uniquely specified. In this case one or more of the underlying assets is already completely specified by the others and could be regarded as a derivative of the others. When the number of assets matches the number of states of the economy A^{-1} is the usual matrix inverse and the hedge is unique.

Remark 1.11: Theorem 1.9 demonstrates clearly that completeness is a statement about the rank of the matrix A. Noting this, we see that there exist economies that admit arbitrage yet are complete. Consider A having rank $m < n$. Since $\dim(\mathrm{Im}(A)^\perp) = n - m$ we can choose $A_0 \in \mathrm{Im}(A)^\perp$, in which case there does not exist $Z \in \mathbb{R}^m$ such that $AZ = A_0$. By Theorem 1.7 the economy admits arbitrage, yet having rank m it is complete by Theorem 1.9.

In practice we will always require an economy to be arbitrage-free. Under this assumption we can state the condition for completeness in terms of the pricing kernel Z.

Theorem 1.12 *Suppose the economy $\mathcal{E}$ is arbitrage-free. Then it is complete if and only if there exists a unique pricing kernel Z.*

Proof: By Theorem 1.7, there exists some vector Z satisfying

$$A_0 = \sum_j Z_j A_1(\omega_j). \tag{1.7}$$

By Theorem 1.9, it suffices to show that Z being unique is equivalent to there being no solution to $Ax = 0$. If $\hat{Z}$ also solves (1.7) then $x = \hat{Z} - Z$ solves $Ax = 0$. Conversely, if Z solves (1.7) and $Ax = 0$ then $\hat{Z} = Z + \varepsilon x$ solves (1.7) for all $\varepsilon > 0$. Since $Z_j > 0$ for all j we can choose ε sufficiently small that $\hat{Z}_j > 0$ for all j, yielding a second pricing kernel. □

The combination of Theorems 1.7 and 1.12 gives the well-known result that an economy is complete and arbitrage-free if and only if there exists a unique pricing kernel Z. This strong notion of completeness is not the primary one that we shall consider in the continuous time context of Chapter 7, where we shall say that an economy $\mathcal{E}$ is complete if it is $\mathcal{F}_T^A$-complete.

Definition 1.13 *Let $\mathcal{F}_1^A$ be the smallest σ-algebra with respect to which the map $A_1 : \Omega \to \mathbb{R}^m$ is measurable. We say the economy $\mathcal{E}$ is $\mathcal{F}_1^A$-complete if every $\mathcal{F}_1^A$-measurable contingent claim is attainable.*

The reason why $\mathcal{F}_1^A$ completeness is more natural to consider is as follows. Suppose there are two states in the economy, ω_i and ω_j, for which $A_1(\omega_i) = A_1(\omega_j)$. Then it is impossible, by observing only the prices A, to distinguish which state of the economy we are in, and in practice all we are interested in is derivative payoffs that can be determined by observing the process A. There is an analogue of Theorem 1.12 which covers this case.

Theorem 1.14 *Suppose the economy $\mathcal{E}$ is arbitrage-free. Then it is $\mathcal{F}_1^A$-complete if and only if all pricing kernels agree on $\mathcal{F}_1^A$, i.e. if $Z^{(1)}$ and $Z^{(2)}$ are two pricing kernels, then for every $F \in \mathcal{F}_1^A$,*

$$\mathbb{E}\big[Z^{(1)}\mathbb{1}_F\big] = \mathbb{E}\big[Z^{(2)}\mathbb{1}_F\big].$$

Proof: For a discrete economy, the statement that X_1 is $\mathcal{F}_1^A$-measurable is precisely the statement that $X_1(\omega_i) = X_1(\omega_j)$ whenever $A_1(\omega_i) = A_1(\omega_j)$. If this is the case we can identify any states ω_i and ω_j for which $A_1(\omega_i) = A_1(\omega_j)$. The question of $\mathcal{F}_1^A$-completeness becomes one of proving that this reduced economy is complete in the full sense. It is clearly arbitrage-free, a property it inherits from the original economy.

Suppose Z is a pricing kernel for the original economy. Then, as is easily verified,

$$\hat{Z} := \mathbb{E}[Z|\mathcal{F}_1^A] \tag{1.8}$$

is a pricing kernel for the reduced economy. If the reduced economy is (arbitrage-free and) complete it has a unique pricing kernel $\hat{Z}$ by Theorem 1.12. Thus all pricing kernels for the original economy agree on $\mathcal{F}_1^A$ by (1.8). Conversely, if all pricing kernels agree on $\mathcal{F}_1^A$ then the reduced economy has a unique pricing kernel, by (1.8), thus is complete by Theorem 1.12. □

1.3.4 Probabilistic formulation

We have seen that pricing derivatives is about replication and the results we have met so far are essentially geometric in nature. However, it is standard in the modern finance literature to work in a probabilistic framework. This approach has two main motivations. The first is that probability gives a natural and convenient language for stating the results and ideas required, and it is also the most natural way to formulate models that will be a good reflection of reality. The second is that many of the sophisticated techniques needed to develop the theory of derivative pricing in continuous time are well developed in a probabilistic setting.

With this in mind we now reformulate our earlier results in a probabilistic context. In what follows let $\mathbb{P}$ be a probability measure on Ω such that $\mathbb{P}(\{\omega_j\}) > 0$ for all j. We begin with a preliminary restatement of Theorem 1.7.

Theorem 1.15 *The economy $\mathcal{E}$ is arbitrage-free if and only if there exists a strictly positive random variable Z such that*

$$A_0 = \mathbb{E}[ZA_1]. \tag{1.9}$$

Extending Definition 1.3, we call Z a pricing kernel *for the economy $\mathcal{E}$.*

Suppose, further, that $\mathbb{P}(A_1^{(i)} \gg 0) = 1$, $A_0^{(i)} > 0$ for some i. Then the economy $\mathcal{E}$ is arbitrage-free if and only if there exists a strictly positive random variable κ with $\mathbb{E}[\kappa] = 1$ such that

$$\mathbb{E}\left[\kappa \frac{A_1}{A_1^{(i)}}\right] = \frac{A_0}{A_0^{(i)}}\,. \tag{1.10}$$

Proof: The first result follows immediately from Theorem 1.7 by setting $Z(\omega_j) = Z_j/\mathbb{P}(\{\omega_j\})$. To prove the second part of the theorem we show that (1.9) and (1.10) are equivalent. This follows since, given either of Z or κ, we can define the other via

$$Z(\omega_j) = \kappa(\omega_j)\frac{A_0^{(i)}}{A_1^{(i)}(\omega_j)}\,.$$

□

Remark 1.16: Note that the random variable Z is simply a weighted version of the pricing kernel in Theorem 1.7, the weights being given by the probability measure $\mathbb{P}$. Although the measure $\mathbb{P}$ assigns probabilities to 'outcomes' ω_j, these probabilities are arbitrary and the role played by $\mathbb{P}$ here is to summarize which states may occur through the assignment of positive mass.

Remark 1.17: The random variable κ is also a reweighted version of the pricing kernel of Theorem 1.7 (and indeed of the one here). In addition to being positive, κ has expectation one, and so $\kappa_j := \kappa(\omega_j)/\mathbb{P}(\{\omega_j\})$ defines a probability measure. We see the importance of this shortly.

Remark 1.18: Equation (1.10) can be interpreted geometrically, as shown in Figure 1.4 for the two-asset case. No arbitrage is equivalent to $A_0 \in C$ where

$$C = \big\{a : a = \sum_j Z_j A_1(\omega_j), Z \gg 0\big\}.$$

Rescaling A_0 and each $A_1(\omega_j)$ does not change the convex cone C and whether or not A_0 is in C, it only changes the weights required to generate A_0 from the $A_1(\omega_j)$. Rescaling so that A_0 and the $A_1(\omega_j)$ all have the same (unit) component in the direction i ensures that the weights κ_j satisfy $\sum_j \kappa_j = 1$.

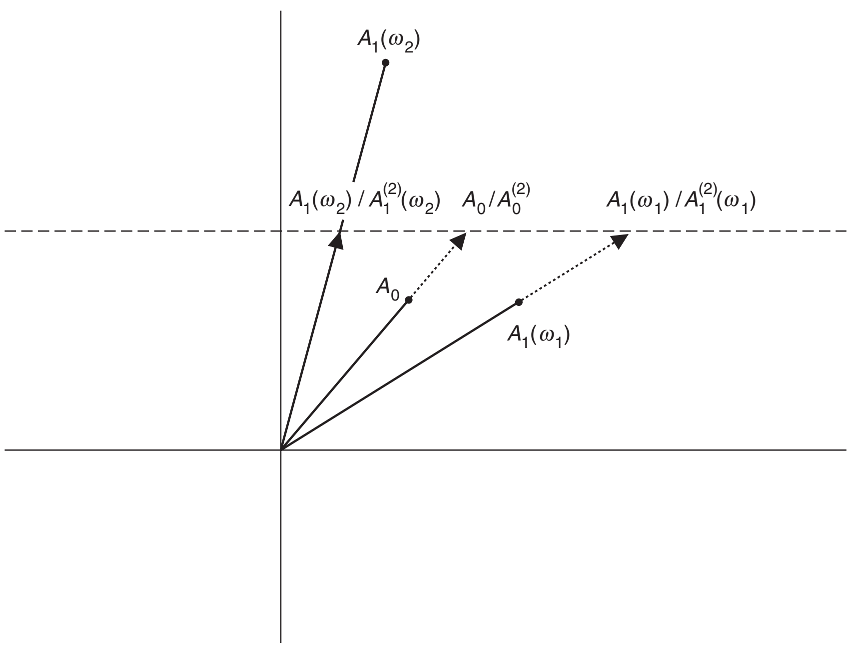

Figure 1.4 Rescaling asset prices

We have one final step to take to cast the results in the standard probabilistic format. This format replaces the problem of finding Z or κ by one of finding a probability measure with respect to which the *process* A, suitably rebased, is a *martingale*. Theorem 1.20 below is the precise statement of what we mean by this (equation (1.11)). We meet martingales again in Chapter 3.

Definition 1.19 *Two probability measures $\mathbb{P}$ and $\mathbb{Q}$ on the finite sample space Ω are said to be equivalent, written $\mathbb{P} \sim \mathbb{Q}$, if*

$$\mathbb{P}(F) = 0 \;\Leftrightarrow\; \mathbb{Q}(F) = 0,$$

for all $F \subseteq \Omega$. If $\mathbb{P} \sim \mathbb{Q}$ we can define the Radon–Nikodým derivative of $\mathbb{P}$ with respect to $\mathbb{Q}$, $\frac{d\mathbb{P}}{d\mathbb{Q}}$, by

$$\frac{d\mathbb{P}}{d\mathbb{Q}}(S) = \frac{\mathbb{P}(S)}{\mathbb{Q}(S)}\,.$$

Theorem 1.20 *Suppose $\mathbb{P}(A_1^{(i)} \gg 0) = 1$, $A_0^{(i)} > 0$ for some i. Then the economy $\mathcal{E}$ is arbitrage-free if and only if there exists a probability measure $\mathbb{Q}_i$ equivalent to $\mathbb{P}$ such that*

$$\mathbb{E}_{\mathbb{Q}_i}[A_1/A_1^{(i)}] = A_0/A_0^{(i)}. \tag{1.11}$$

The measure $\mathbb{Q}_i$ is said to be an equivalent martingale measure for $A^{(i)}$.

Proof: By Theorem 1.15 we must show that (1.11) is equivalent to the existence of a strictly positive random variable κ with $\mathbb{E}[\kappa] = 1$ such that

$$\mathbb{E}\left[\kappa \frac{A_1}{A_1^{(i)}}\right] = \frac{A_0}{A_0^{(i)}}\,.$$

Suppose $\mathcal{E}$ is arbitrage-free and such a κ exists. Define $\mathbb{Q}_i \sim \mathbb{P}$ by $\mathbb{Q}_i(\{\omega_j\}) = \kappa(\omega_j)\mathbb{P}(\{\omega_j\})$. Then

$$\begin{aligned}\mathbb{E}_{\mathbb{Q}_i}[A_1/A_1^{(i)}] &= \mathbb{E}[\kappa A_1/A_1^{(i)}]\\ &= A_0/A_0^{(i)}\end{aligned}$$

as required.

Conversely, if (1.11) holds define $\kappa(\omega_j) = \mathbb{Q}_i(\{\omega_j\})/\mathbb{P}(\{\omega_j\})$. It follows easily that κ has unit expectation. Furthermore,

$$\begin{aligned}\mathbb{E}\left[\kappa \frac{A_1}{A_1^{(i)}}\right] &= \mathbb{E}\left[\frac{d\mathbb{Q}_i}{d\mathbb{P}} \frac{A_1}{A_1^{(i)}}\right]\\ &= \mathbb{E}_{\mathbb{Q}_i}[A_1/A_1^{(i)}]\\ &= A_0/A_0^{(i)},\end{aligned}$$

and there is no arbitrage by Theorem 1.15. □

We are now able to restate our other results on completeness and pricing in this same probabilistic framework. In our new framework the condition for completeness can be stated as follows.

Theorem 1.21 *Suppose that no arbitrage exists and that* $\mathbb{P}(A_1^{(i)} \gg 0) = 1$, $A_0^{(i)} > 0$ *for some* i. *Then the economy* $\mathcal{E}$ *is complete if and only if there exists a unique equivalent martingale measure for the 'unit'* $A^{(i)}$.

Proof: Observe that in the proof of Theorem 1.20 we established a one-to-one correspondence between pricing kernels and equivalent martingale measures. The result now follows from Theorem 1.12 which establishes the equivalence of completeness to the existence of a unique pricing kernel. □

Our final result, concerning the pricing of a derivative, is left as an exercise for the reader.

Theorem 1.22 *Suppose* $\mathcal{E}$ *is arbitrage-free,* $\mathbb{P}(A_1^{(i)} \gg 0) = 1$, $A_0^{(i)} > 0$ *for some* i, *and that* X *is an attainable contingent claim. Then the fair value of* X *is given by*

$$X_0 = A_0^{(i)} \mathbb{E}_{\mathbb{Q}_i}[X_1 / A_1^{(i)}],$$

where $\mathbb{Q}_i$ *is an equivalent martingale measure for the unit* $A^{(i)}$.

1.3.5 Units and numeraires

Throughout Section 1.3.4 we assumed that the price of one of the assets, asset i, was positive with probability one. This allowed us to use this asset price as a *unit*; the operation of dividing all other asset prices by the price of asset i can be viewed as merely recasting the economy in terms of this new unit.

There is no reason why, throughout Section 1.3.4, we need to restrict the unit to be one of the assets. All the results hold if the unit $A^{(i)}$ is replaced by some other unit U which is strictly positive with probability one and for which

$$U_0 = \mathbb{E}_{\mathbb{P}}(ZU_1) \tag{1.12}$$

where Z is some pricing kernel for the economy. Note, in particular, that we can always take $U_0 = 1$, $\mathbb{Q}(\{\omega_j\}) = 1/n$ and $U_1(\omega_j) = 1/(nZ_j\mathbb{P}(\{\omega_j\}))$.

Observe that (1.12) automatically holds (assuming a pricing kernel exists) when U is a derivative and thus is of the form $U = \phi \cdot A$. In this case we say that U is a *numeraire* and then we usually denote it by the symbol N in preference to U. In general there are more units than numeraires.

The ideas of numeraires, martingales and change of measure are central to the further development of derivative pricing theory.

1.4 A TWO-PERIOD EXAMPLE

We now briefly consider an example of a two-period economy. Inclusion of the extra time-step allows us to develop new ideas whilst still in a relatively simple framework. In particular, to price a derivative in this richer setting we shall need to define what is meant by a *trading strategy* which is *self-financing*.

For our two-period example we build on the simple set-up of Section 1.2. Suppose that at the new time, time 2, the economy can be in one of four states which we denote by ω'_j, $j = 1, \ldots, 4$, with the restriction that states ω'_1 and ω'_2 can be reached only from ω_1 and states ω'_3 and ω'_4 can be reached only from ω_2. Figure 1.5 summarizes the possibilities.

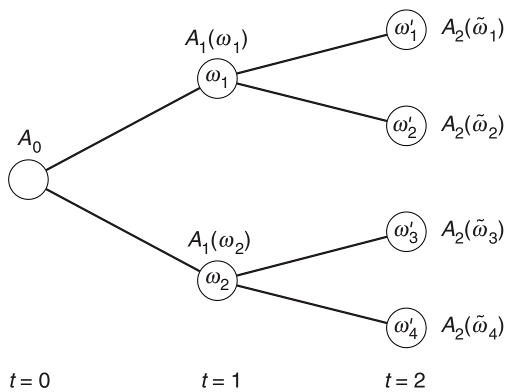

Figure 1.5 Possible states of a two-period economy

Let Ω now denote the set of all possible *paths*. That is, $\Omega = \{\tilde{\omega}_k, k = 1, \ldots, 4\}$ where $\tilde{\omega}_k = (\omega_1, \omega'_k)$ for $k = 1, 2$ and $\tilde{\omega}_k = (\omega_2, \omega'_k)$ for $k = 3, 4$. The asset prices follow one of four paths with $A_t^{(i)}(\tilde{\omega}_k)$ denoting the price of asset i at time $t = 1, 2$.

Consider the problem of pricing a derivative X which at time 2 pays an amount $X_2(\tilde{\omega}_k)$. In order to price the derivative X we must be able to replicate over all paths. We cannot do this by holding a static portfolio. Instead, the portfolio we hold at time zero will in general need to be changed at time 1 according to which state ω_j the economy is in at this time. Thus, replicating X in a two-period economy amounts to specifying a *process* ϕ which is non-anticipative, i.e. a process which does not depend on knowledge of a future state at any time. Such a process is referred to as a *trading strategy*. To find the fair value of X we must be able to find a trading strategy which is *self-financing*; that is, a strategy for which, apart from the initial capital at time zero, no additional influx of funds is required in order to replicate X. Such a trading strategy will be called *admissible*.

We calculate a suitable ϕ in two stages, working backwards in time. Suppose we know the economy is in state ω_1 at time 1. Then we know from the one-period example of Section 1.2 that we should hold a portfolio $\phi_1(\omega_1)$ of assets satisfying

$$\phi_1(\omega_1) \cdot A_2(\omega'_j) = X_2(\tilde{\omega}_j), \quad j = 1, 2.$$

Similarly, if the economy is in state ω_2 at time 1 we should hold a portfolio $\phi_1(\omega_2)$ satisfying

$$\phi_1(\omega_2) \cdot A_2(\omega'_j) = X_2(\tilde{\omega}_j), \quad j = 3, 4.$$

Conditional on knowing the economy is in state ω_j at time 1, the fair value of the derivative X at time 1 is then $X_1(\omega_j) = \phi_1(\omega_j) \cdot A_1(\omega_j)$, the value of the replicating portfolio.

Once we have calculated ϕ_1 the problem of finding the fair value of X at time zero has been reduced to the one-period case, i.e. that of finding the fair value of an option paying $X_1(\omega_j)$ at time 1. If we can find ϕ_0 such that

$$\phi_0 \cdot A_1(\omega_j) = X_1(\omega_j), \quad j = 1, 2,$$

then the fair price of X at time zero is

$$X_0 = \phi_0 \cdot A_0.$$

Note that for the ϕ satisfying the above we have

$$\phi_0 \cdot A_1(\omega_j) = \phi_1(\omega_j) \cdot A_1(\omega_j), \quad j = 1, 2,$$

as must be the case for the strategy to be self-financing; the portfolio is merely rebalanced at time 1.

Though of mathematical interest, the multi-period case is not important in practice and when we again take up the story of derivative pricing in Chapter 7 we will work entirely in the continuous time setting. For a full treatment of the multi-period problem the reader is referred to Duffie (1996).

2

Brownian Motion

2.1 INTRODUCTION

Our objective in this book is to develop a theory of derivative pricing in continuous time, and before we can do this we must first have developed models for the underlying assets in the economy. In reality, asset prices are piecewise constant and undergo discrete jumps but it is convenient and a reasonable approximation to assume that asset prices follow a continuous process. This approach is often adopted in mathematical modelling and justified, at least in part, by results such as those in Section 2.3.1. Having made the decision to use continuous processes for asset prices, we must now provide a way to generate such processes. This we will do in Chapter 6 when we study stochastic differential equations. In this chapter we study the most fundamental and most important of all continuous stochastic processes, the process from which we can build all the other continuous time processes that we will consider, Brownian motion.

The physical process of Brownian motion was first observed in 1828 by the botanist Robert Brown for pollen particles suspended in a liquid. It was in 1900 that the mathematical process of Brownian motion was introduced by Bachelier as a model for the movement of stock prices, but this was not really taken up. It was only in 1905, when Einstein produced his work in the area, that the study of Brownian motion started in earnest, and not until 1923 that the existence of Brownian motion was actually established by Wiener. It was with the work of Samuelson in 1969 that Brownian motion reappeared and became firmly established as a modelling tool for finance.

In this chapter we introduce Brownian motion and derive a few of its more immediate and important properties. In so doing we hope to give the reader some insight into and intuition for how it behaves.

2.2 DEFINITION AND EXISTENCE

We begin with the definition.

Definition 2.1 *Brownian motion is a real-valued stochastic process with the following properties:*

(BM.i) *Given any* $t_0 < t_1 < \ldots < t_n$ *the random variables* $\{(W_{t_i} - W_{t_{i-1}}),$ $i = 1, 2, \ldots, n\}$ *are independent.*
(BM.ii) *For any* $0 \le s \le t, W_t - W_s \sim N(0, t-s)$.
(BM.iii) W_t *is continuous in* t *almost surely (a.s.).*
(BM.iv) $W_0 = 0$ *a.s.*

Property (BM.iv) is sometimes omitted to allow arbitrary initial distributions, although it is usually included. We include it for definiteness and note that extending to the more general case is a trivial matter.

Implicit in the above definition is the fact that a process W exists with the stated properties and that it is unique. Uniqueness is a straightforward matter – any two processes satisfying (BM.i)–(BM.iv) have the same finite-dimensional distributions, and the finite-dimensional distributions determine the law of a process. In fact, Brownian motion can be characterized by a slightly weaker condition than (BM.ii):

(BM.ii$'$) For any $t \ge 0, h \ge 0$, the distribution of $W_{t+h} - W_t$ is independent of $t, \mathbb{E}[W_t] = 0$ and $\text{var}[W_t] = t$.

It should not be too surprising that (BM.ii$'$) in place of (BM.ii) gives an equivalent definition. Given any $0 \le s \le t$, the interval $[s, t]$ can be subdivided into n equal intervals of length $(t-s)/n$ and

$$W_t - W_s = \sum_{i=1}^{n} W_{s+(t-s)i/n} - W_{s+(t-s)(i-1)/n},$$

i.e. as a sum of n independent, identically distributed random variables. A little care needs to be exercised, but it effectively follows from the central limit theorem and the continuity of W that $W_t - W_s$ must be Gaussian. Breiman (1992) provides all the details. The moment conditions in (BM.ii$'$) now force the process to be a standard Brownian motion (without them the conditions define the more general process $\mu t + \sigma W_t$).

The existence of Brownian motion is more difficult to prove. There are many excellent references which deal with this question, including those by Breiman (1992) and Durrett (1984). We will merely state this result. To do so, we must explicitly define a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a process W on this space which is Brownian motion. Let $\Omega = \mathbb{C} \equiv \mathbb{C}(\mathbb{R}^+, \mathbb{R})$ be the set of continuous functions from $[0, \infty)$ to $\mathbb{R}$. Endow $\mathbb{C}$ with the metric of uniform convergence, i.e. for $x, y \in \mathbb{C}$

$$\rho(x,y) = \sum_{n=1}^{\infty} \left(\frac{1}{2}\right)^n \frac{\rho_n(x,y)}{1+\rho_n(x,y)}$$

where

$$\rho_n(x,y) = \sup_{0\le t\le n} |x(t) - y(t)|.$$

Now define $\mathcal{F} = \mathcal{C} \equiv \mathcal{B}(\mathbb{C})$, the Borel σ-algebra on $\mathbb{C}$ induced by the metric ρ.

Theorem 2.2 *For each $\omega \in \mathbb{C}, t \ge 0$, define $W_t(\omega) = \omega_t$. There exists a unique probability measure $\mathbb{W}$ (Wiener measure) on $(\mathbb{C}, \mathcal{C})$ such that the stochastic process $\{W_t, t \ge 0\}$ is Brownian motion (i.e. satisfies conditions (BM.i)–(BM.iv)).*

Remark 2.3: The set-up described above, in which the sample space Ω is the set of sample paths, is the *canonical set-up*. There are others, but the canonical set-up is the most direct and intuitive to consider. Note, in particular, that for this set-up *every* sample path is continuous.

In the above we have defined Brownian motion without reference to a filtration. Adding a filtration is a straightforward matter: it will often be taken to be $\{\mathcal{F}_t^W\}^\circ := \sigma(W_s, s \le t)$, but it could also be more general. Brownian motion relative to a filtered probability space is defined as follows.

Definition 2.4 *The process W is Brownian motion with respect to the filtration $\{\mathcal{F}_t\}$ if:*
(i) it is adapted to $\{\mathcal{F}_t\}$;
(ii) for all $0 \le s \le t, W_t - W_s$ is independent of $\mathcal{F}_s$;
(iii) it is a Brownian motion as defined in Definition 2.1.

2.3 BASIC PROPERTIES OF BROWNIAN MOTION

2.3.1 Limit of a random walk

If you have not encountered Brownian motion before it is important to develop an intuition for its behaviour. A good place to start is to compare it with a simple symmetric random walk on the integers, S_n. The following result roughly states that if we speed up a random walk and look at it from a distance (see Figure 2.1) it appears very much like Brownian motion.

Theorem 2.5 *Let $\{X_i, i \ge 1\}$ be independent and identically distributed random variables with $\mathbb{P}(X_i = 1) = \mathbb{P}(X_i = -1) = \frac{1}{2}$. Define the simple*

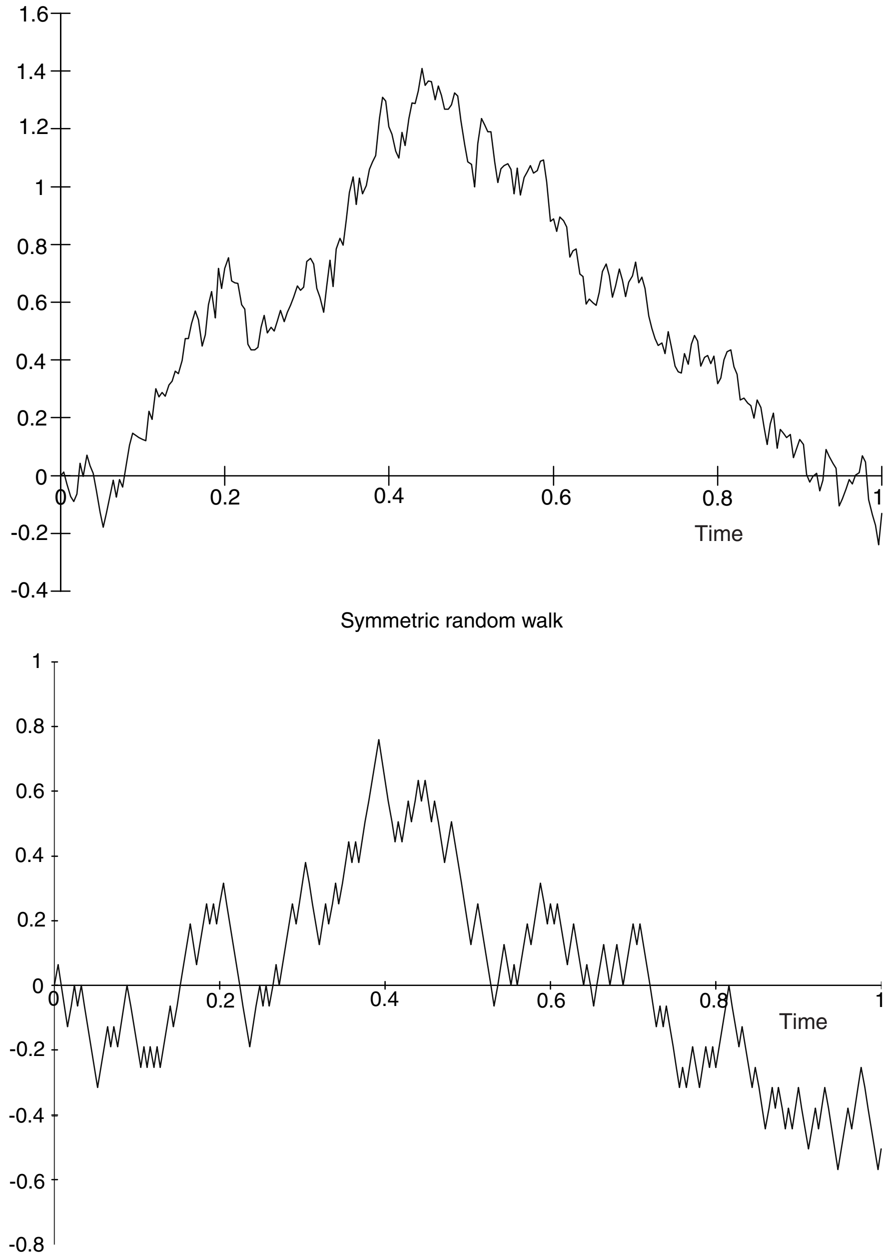

Figure 2.1. Brownian motion and a random walk

symmetric random walk $\{S_n, n \geq 1\}$ *as* $S_n = \sum_{i=1}^n X_i$, *and the rescaled random walk* $Z_n(t) := \frac{1}{\sqrt{n}} S_{\lfloor nt \rfloor}$ *where* $\lfloor x \rfloor$ *is the integer part of* x. *As* $n \to \infty$,

$$Z_n \Rightarrow W.$$

In the above, W *is a Brownian motion and the convergence* $\Rightarrow$ *denotes weak convergence. In this setting this is equivalent to convergence of finite-dimensional distributions, meaning, here, that for any* $t_1, t_2, \ldots, t_k$ *and any* $z \in \mathbb{R}^k$,

$$\mathbb{P}(Z_n(t_i) \leq z_i,\ i = 1, \ldots, k) \to \mathbb{P}(W_{t_i} \leq z_i,\ i = 1, \ldots, k)$$

as $n \to \infty$.

Proof: This follows immediately from the (multivariate) central limit theorem for the simple symmetric random walk. □

2.3.2 Deterministic transformations of Brownian motion

There is an extensive theory which studies what happens to a Brownian motion under various (random) transformations. One question people ask is what is the law of the process X defined by

$$X_t = W_{\tau(t)},$$

where the random clock τ is specified by

$$\tau(t) = \inf\{s \geq 0 : Y_s > t\},$$

for some process Y. In the case when Y, and consequently τ, is non-random the study is straightforward because the Gaussian structure of Brownian motion is retained. In particular, there are a well-known set of transformations of Brownian motion which produce another Brownian motion, and these results prove especially useful when studying properties of the Brownian sample path.

Theorem 2.6 *Suppose* W *is Brownian motion. Then the following transformations also produce Brownian motions:*

(i) $\widetilde{W}_t := cW_{t/c^2}$ *for any* $c \in \mathbb{R}\backslash\{0\}$; *(scaling)*
(ii) $\widetilde{W}_t := tW_{1/t}$ *for* $t > 0, \widetilde{W}_0 := 0$; *(time inversion)*
(iii) $\widetilde{W}_t := \{W_{t+s} - W_s : t \geq 0\}$ *for any* $s \in \mathbb{R}^+$. *(time homogeneity)*

Proof: We must prove that (BM.i)–(BM.iv) hold. The proof in each case follows similar lines so we will only establish the time inversion result which is slightly more involved than the others.

Given any fixed $t_0 < t_1 < \dots < t_n$ it follows from the Gaussian structure of W that $(\widetilde{W}_{t_0}, \dots, \widetilde{W}_{t_n})$ also has a Gaussian distribution. A Gaussian distribution is completely determined by its mean and covariance and thus (BM.i) and (BM.ii) now follow if we can show that $\mathbb{E}[\widetilde{W}_t] = 0$ and $\mathbb{E}[\widetilde{W}_s \widetilde{W}_t] = s \wedge t$ for all s, t. But this follows easily:

$$\mathbb{E}[\widetilde{W}_t] = \mathbb{E}[tW_{1/t}] = t\mathbb{E}[W_{1/t}] = 0,$$
$$\mathbb{E}[\widetilde{W}_s \widetilde{W}_t] = \mathbb{E}[stW_{1/s}W_{1/t}] = st\big((1/t) \wedge (1/s)\big) = s \wedge t\,.$$

Condition (BM.iv), that $\widetilde{W}_0 = 0$, is part of the definition and, for any $t > 0$, the continuity of $\widetilde{W}$ at t follows from the continuity of W. It therefore only remains to prove that $\widetilde{W}$ is continuous at zero or, equivalently, that for all $n > 0$ there exists some $m > 0$ such that $\sup_{0<t\leq 1/m} |\widetilde{W}_t| < 1/n$. The continuity of $\widetilde{W}$ for $t > 0$ implies that

$$\sup_{0<t\leq 1/m} |\widetilde{W}_t| = \sup_{q\in\mathbb{Q}\cap(0,1/m]} |\widetilde{W}_q|\,,$$

and so

$$\begin{aligned}\mathbb{P}\Big(\lim_{t\to 0} \widetilde{W}_t = 0\Big) &= \mathbb{P}\Big(\bigcap_n \bigcup_m \Big\{ \sup_{q\in\mathbb{Q}\cap(0,1/m]} |\widetilde{W}_q| < 1/n \Big\}\Big) \\ &= \mathbb{P}\Big(\bigcap_n \bigcup_m \bigcap_{q\in\mathbb{Q}\cap(0,1/m]} \big\{ |\widetilde{W}_q| < 1/n \big\}\Big)\,. \qquad (2.1)\end{aligned}$$

There are a countable number of events in the last expression of equation (2.1) and each of these involves the process $\widetilde{W}$ at a single strictly positive time. The distributions of W and $\widetilde{W}$ agree for $t > 0$ and therefore this last probability is unaltered if the process $\widetilde{W}$ is replaced by the original Brownian motion W. But W is a.s. continuous at zero and so the probability in (2.1) is one and $\widetilde{W}$ is also a.s. continuous at zero. □

2.3.3 Some basic sample path properties

The transformations just introduced have many useful applications. Two of them are given below.

Theorem 2.7 *If W is a Brownian motion then, as $t \to \infty$,*

$$\frac{W_t}{t} \to 0 \qquad a.s.$$

Proof: Writing $\widetilde{W}_t$ for the time inversion of W as defined in Theorem 2.6, $\mathbb{P}(\lim_{t\to\infty} \frac{W_t}{t} = 0) = \mathbb{P}(\lim_{s\to 0} \widetilde{W}_s = 0)$ and this last probability is one since $\widetilde{W}$ is a Brownian motion, continuous at zero. □

Theorem 2.8 *Given a Brownian motion W,*

$$\mathbb{P}\left(\sup_{t\geq 0} W_t = +\infty, \inf_{t\geq 0} W_t = -\infty\right) = 1.$$

Proof: Consider first $\sup_{t\geq 0} W_t$. For any $a > 0$, the scaling property implies that

$$\mathbb{P}\left(\sup_{t\geq 0} W_t > a\right) = \mathbb{P}\left(\sup_{t\geq 0} cW_{t/c^2} > a\right)$$
$$= \mathbb{P}\left(\sup_{s\geq 0} W_s > a/c\right).$$

Hence the probability is independent of a and thus almost every sample path has a supremum of either 0 or ∞. In particular, for almost every sample path, $\sup_{t\geq 0} W_t = 0$ if and only if $W_1 \leq 0$ and $\sup_{t\geq 1}(W_t - W_1) = 0$. But then, defining $\widetilde{W}_t = W_{1+t} - W_1$,

$$\begin{aligned} p &:= \mathbb{P}\left(\sup_{t\geq 0} W_t = 0\right) \\ &= \mathbb{P}\left(W_1 \leq 0, \sup_{t\geq 1}(W_t - W_1) = 0\right) \\ &= \mathbb{P}(W_1 \leq 0)\mathbb{P}\left(\sup_{t\geq 0} \widetilde{W}_t = 0\right) \\ &= \frac{1}{2}p. \end{aligned}$$

We conclude that $p = 0$. A symmetric argument shows that $\mathbb{P}(\inf_{t\geq 0} W_t = -\infty) = 1$ and this completes the proof. □

This result shows that the Brownian path will keep oscillating between positive and negative values over extended time intervals. The path also oscillates wildly over small time intervals. In particular, it is nowhere differentiable.

Theorem 2.9 *Brownian motion is nowhere differentiable with probability one.*

Proof: It suffices to prove the result on [0, 1]. Suppose $W_\cdot(\omega)$ is differentiable at some $t \in [0, 1]$ with derivative bounded in absolute value by some constant $N/2$. Then there exists some $n > 0$ such that

$$h \leq 4/n \quad \Rightarrow \quad |W_{t+h}(\omega) - W_t(\omega)| \leq Nh. \tag{2.2}$$

Denote the event that (2.2) holds for some $t \in [0, 1]$ by $A_{n,N}$. The event that Brownian motion is differentiable for some $t \in [0, 1]$ is a subset of the event $\bigcup_N \bigcup_n A_{n,N}$ and the result is proven if we can show that $\mathbb{P}(A_{n,N}) = 0$ for all n, N.

Fix n and N. Let $\Delta_n(k) = W_{(k+1)/n} - W_{k/n}$ and define $k_t^n = \inf\{k : k/n \geq t\}$. If (2.2) holds at t then it follows from the triangle inequality that $|\Delta_n(k_t^n + j)| \leq 7N/n$, for $j = 0, 1, 2$. Therefore

$$A_{n,N} \subseteq \bigcup_{k=0}^{n} \bigcap_{j=0}^{2} \{|\Delta_n(k+j)| \leq 7N/n\},$$

$$\begin{aligned}
\mathbb{P}(A_{n,N}) &\leq \mathbb{P}\Big(\bigcup_{k=0}^{n} \bigcap_{j=0}^{2} \{|\Delta_n(k+j)| \leq 7N/n\}\Big) \\
&\leq (n+1)\mathbb{P}\Big(\bigcap_{j=0}^{2} \{|\Delta_n(j)| \leq 7N/n\}\Big) \\
&= (n+1)\mathbb{P}\Big(\{|\Delta_n(0)| \leq 7N/n\}\Big)^3 \\
&\leq (n+1)\left(\frac{14N}{\sqrt{2\pi n}}\right)^3 \\
&\to 0 \qquad \text{as } n \to \infty,
\end{aligned}$$

the last equality following from the independence of the Brownian increments. If we now note that $A_{n,N}$ is increasing in n we can conclude that $\mathbb{P}(A_{n,N}) = 0$ for all N, n and we are done. □

The lack of differentiability of Brownian motion illustrates how irregular the Brownian path can be and implies immediately that the Brownian path is not of finite variation. This means it is not possible to use the Brownian path as an integrator in the classical sense. The following result is central to stochastic integration. A proof is provided in Chapter 3, Corollary 3.81.

Theorem 2.10 *For a Brownian motion W, define the doubly infinite sequence of (stopping) times T_k^n via*

$$T_0^n \equiv 0,\ T_{k+1}^n = \inf\{t > T_k^n : |W_t - W_{T_k^n}| > 2^{-n}\},$$

and let

$$[W]_t(\omega) := \lim_{n\to\infty} \sum_{k\geq 1} [W_{t\wedge T_k^n}(\omega) - W_{t\wedge T_{k-1}^n}(\omega)]^2.$$

The process $[W]$ is called the quadratic variation of W, and $[W]_t = t$ a.s.

2.4 STRONG MARKOV PROPERTY

Markov processes and strong Markov processes are of fundamental importance to mathematical modelling. Roughly speaking, a Markov process is one for which 'the future is independent of the past, given the present'; for any fixed $t \geq 0$, X is Markovian if the law of $\{X_s - X_t : s \geq t\}$ given X_t is independent of the law of $\{X_s : s \leq t\}$. This property obviously simplifies the study of Markov processes and often allows problems to be solved (perhaps numerically) which are not tractable for more general processes. We shall, in Chapter 6, give a definition of a (strong) Markov process which makes the above more precise. That Brownian motion is Markovian follows from its definition.

Theorem 2.11 *Given any $t \geq 0$, $\{W_s - W_t : s \geq t\}$ is independent of $\{W_s : s \leq t\}$ (and indeed $\mathcal{F}_t$ if the Brownian motion is defined on a filtered probability space).*

Proof: This is immediate from (BM.i). □

The strong Markov property is a more stringent requirement of a stochastic process and is correspondingly more powerful and useful. To understand this idea we must here introduce the concept of a stopping time. We will reintroduce this in Chapter 3 when we discuss stopping times in more detail.

Definition 2.12 *The random variable T, taking values in $[0, \infty]$, is an $\{\mathcal{F}_t\}$ stopping time if*

$$\{T \leq t\} = \{\omega : T(\omega) \leq t\} \in \mathcal{F}_t\,,$$

for all $t \leq \infty$.

Definition 2.13 *For any $\{\mathcal{F}_t\}$ stopping time T, the pre-T σ-algebra $\mathcal{F}_T$ is defined via*

$$\mathcal{F}_T = \{F : \text{ for every } t \leq \infty, F \cap \{T \leq t\} \in \mathcal{F}_t\}\,.$$

Definition 2.12 makes precise the intuitive notion that a stopping time is one for which we know that it has occurred at the moment it occurs. Definition 2.13 formalizes the idea that the σ-algebra $\mathcal{F}_T$ contains all the information that is available up to and including the time T.

A strong Markov process, defined precisely in Chapter 6, is a process which possesses the Markov property, as described at the beginning of this section, but with the constant time t generalized to be an arbitrary stopping time T. Theorem 2.15 below shows that Brownian motion is also a strong Markov process. First we establish a weaker preliminary result.

Proposition 2.14 *Let W be a Brownian motion on the filtered probability space $(\Omega, \{\mathcal{F}_t\}, \mathcal{F}, \mathbb{P})$ and suppose that T is an a.s. finite stopping time taking on one of a countable number of possible values. Then*

$$\widetilde{W}_t := W_{t+T} - W_T$$

is a Brownian motion and $\{\widetilde{W}_t : t \geq 0\}$ is independent of $\mathcal{F}_T$ (in particular, it is independent of $\{W_s : s \leq T\}$).

Proof: Fix $t_1, \ldots, t_n \geq 0, F_1, \ldots, F_n \in \mathcal{B}(\mathbb{R}), F \in \mathcal{F}_T$ and let $\{\tau_j, j \geq 1\}$ denote the values that T can take. We have that

$$\begin{aligned}\mathbb{P}(\widetilde{W}_{t_i} \in F_i, i = 1, \ldots, n; F) &= \sum_{j=1}^{\infty} \mathbb{P}(\widetilde{W}_{t_i} \in F_i, i = 1, \ldots, n; F; T = \tau_j) \\ &= \sum_{j=1}^{\infty} \mathbb{P}(W_{t_i+\tau_j} - W_{\tau_j} \in F_i, i=1, \ldots, n)\mathbb{P}(F; T = \tau_j) \\ &= \mathbb{P}(W_{t_i} \in F_i, i = 1, \ldots, n)\mathbb{P}(F),\end{aligned}$$

the last equality following from the Brownian shifting property and by performing the summation. This proves the result. □

Theorem 2.15 *Let W be a Brownian motion on the filtered probability space $(\Omega, \{\mathcal{F}_t\}, \mathcal{F}, \mathbb{P})$ and suppose that T is an a.s. finite stopping time. Then*

$$\widetilde{W}_t := W_{t+T} - W_T$$

is a Brownian motion and $\{\widetilde{W}_t : t \geq 0\}$ is independent of $\mathcal{F}_T$ (in particular, it is independent of $\{W_s : s \leq T\}$).

Proof: As in the proof of Proposition 2.14, and adopting the notation introduced there, it suffices to prove that

$$\mathbb{P}(\widetilde{W}_{t_i} \in F_i, i = 1, \ldots, n; F) = \mathbb{P}(W_{t_i} \in F_i, i = 1, \ldots, n)\mathbb{P}(F).$$

Define the stopping times T_k via

$$T_k = \frac{q}{2^k} \quad \text{if} \quad \frac{q-1}{2^k} \leq T < \frac{q}{2^k}.$$

Each T_k can only take countably many values so, noting that $T_k \geq T$ and thus $F \in \mathcal{F}_{T_k}$, Proposition 2.14 applies to give

$$\mathbb{P}(\widetilde{W}^k_{t_i} \in F_i, i = 1, \ldots, n; F) = \mathbb{P}(W^k_{t_i} \in F_i, i = 1, \ldots, n)\mathbb{P}(F), \tag{2.3}$$

where $\widetilde{W}^k_t = W_{t+T_k} - W_{T_k}$. As $k \to \infty, T_k(\omega) \to T(\omega)$ and thus, by the almost sure continuity of Brownian motion, $\widetilde{W}^k_t \to W_t$ for all t, a.s. The result now follows from (2.3) by dominated convergence. □

2.4.1 Reflection principle

An immediate and powerful consequence of Theorem 2.15 is as follows.

Theorem 2.16 *If W is a Brownian motion on the space $(\Omega, \{\mathcal{F}_t\}, \mathcal{F}, \mathbb{P})$ and $T = \inf\{t > 0 : W_t \geq a\}$, for some a, then the process*

$$\widetilde{W}_t := \begin{cases} W_t, & t < T \\ 2a - W_t, & t \geq T \end{cases}$$

is also Brownian motion.

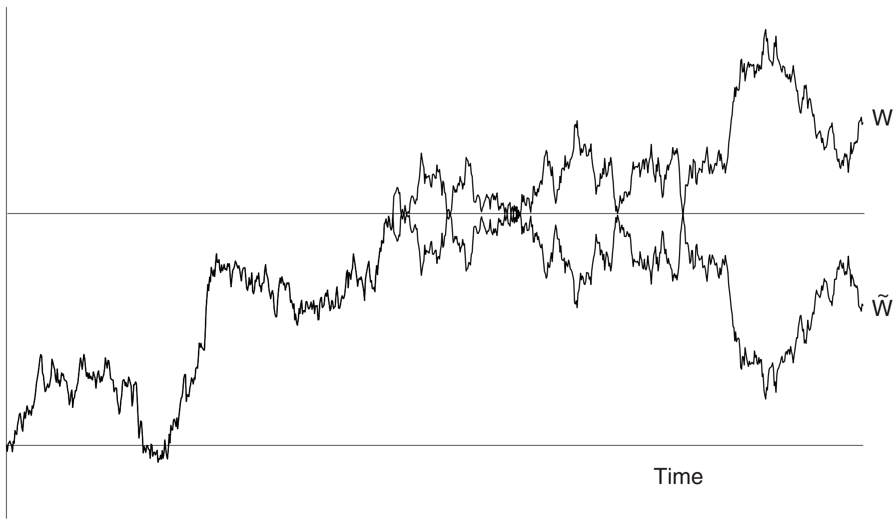

Figure 2.2 Reflected Brownian path

Proof: It is easy to see that T is a stopping time (look ahead to Theorem 3.40 if you want more details here), thus $W_t^* := W_{t+T} - W_T$ is, by the strong Markov property, a Brownian motion independent of $\mathcal{F}_T$, as is $-W_t^*$ (by the scaling property). Hence the following have the same distribution:

(i) $W_t \mathbb{1}_{\{t<T\}} + \big(a + W_{t-T}^*\big)\mathbb{1}_{\{t \geq T\}}$
(ii) $W_t \mathbb{1}_{\{t<T\}} + \big(a - W_{t-T}^*\big)\mathbb{1}_{\{t \geq T\}}$.

The first of these is just the original Brownian motion, the second is $\widetilde{W}$, so the result follows. □

A 'typical' sample path for Brownian motion W and the reflected Brownian motion $\widetilde{W}$ is shown in Figure 2.2.

Example 2.17 The reflection principle can be used to derive the joint distribution of Brownian motion and its supremum, a result which is used for the analytic valuation of barrier options. Let W be a Brownian motion and $\widetilde{W}$ be the Brownian motion reflected about some $a > 0$. Defining $M_t = \sup_{s \leq t} W_s$

(and $\widetilde{M}$ similarly), it is clear that, for $x \le a$,

$$\{\omega : W_t(\omega) \le x, M_t(\omega) \ge a\} = \{\omega : \widetilde{W}_t(\omega) \ge 2a - x, \widetilde{M}_t(\omega) \ge a\},$$

and thus, denoting by Φ the normal distribution function,

$$\begin{aligned}\mathbb{P}(W_t \le x, M_t \ge a) &= \mathbb{P}(\widetilde{W}_t \ge 2a - x, \widetilde{M}_t \ge a)\\ &= \mathbb{P}(\widetilde{W}_t \ge 2a - x)\\ &= 1 - \Phi\left(\frac{2a - x}{\sqrt{t}}\right),\end{aligned}$$

the second equality holding since $2a - x \ge a$. For $x > a$,

$$\begin{aligned}\mathbb{P}(W_t \ge x, M_t \ge a) &= \mathbb{P}(W_t \ge x)\\ &= 1 - \Phi\left(\frac{x}{\sqrt{t}}\right).\end{aligned}$$

From these we can find the density with respect to x:

$$\mathbb{P}(W_t \in dx, M_t \ge a) = \begin{cases} \dfrac{\exp(-(2a - x)^2/2t)}{\sqrt{2\pi t}}, & x \le a\\ \dfrac{\exp(-x^2/2t)}{\sqrt{2\pi t}}, & x > a. \end{cases}$$

The density with respect to a follows similarly.

3

Martingales

Martingales are amongst the most important tools in modern probability. They are also central to modern finance theory. The simplicity of the martingale definition, a process which has mean value at any future time, conditional on the present, equal to its present value, belies the range and power of the results which can be established. In this chapter we will introduce continuous time martingales and develop several important results. The value of some of these will be self-evident. Others may at first seem somewhat esoteric and removed from immediate application. Be assured, however, that we have included nothing which does not, in our view, aid understanding or which is not directly relevant to the development of the stochastic integral in Chapter 4, or the theory of continuous time finance as developed in Chapter 7.

A brief overview of this chapter and the relevance of each set of results is as follows. Section 3.1 below contains the definition, examples and basic properties of martingales. Throughout we will consider general continuous time martingales, although for the financial applications in this book we only need martingales that have continuous paths. In Section 3.2 we introduce and discuss three increasingly restrictive classes of martingales. As we impose more restrictions, so stronger results can be proved. The most important class, which is of particular relevance to finance, is the class of uniformly integrable martingales. Roughly speaking, any uniformly integrable martingale M, a stochastic process, can be summarized by M_∞, a random variable: given M we know M_∞, given M_∞ we can recover M. This reduces the study of these martingales to the study of random variables. Furthermore, this class is precisely the one for which the powerful and extremely important optional sampling theorem applies, as we show in Section 3.3. A second important space of martingales is also discussed in Section 3.2, square-integrable martingales. It is not obvious in this chapter why these are important and the results may seem a little abstract. If you want motivation, glance ahead to Chapter 4, otherwise take our word for it that they are important.

Section 3.3 contains a discussion of stopping times and, as mentioned above, the optional sampling theorem ('if a martingale is uniformly integrable then the mean value of the martingale at any stopping time, conditional on the present, is its present value'). The quadratic variation for continuous martingales is introduced and studied in Section 3.4. The results which we develop are interesting in their own right, but the motivation for the discussion of quadratic variation is its vital role in the development of the stochastic integral. The chapter concludes with two sections on processes more general than martingales. Section 3.5 introduces the idea of localization and defines *local* martingales and *semimartingales*. Then, in Section 3.6, *supermartingales* are considered and the important Doob–Meyer decomposition theorem is presented. We will call on the latter result in Chapter 8 when we study term structure models.

3.1 DEFINITION AND BASIC PROPERTIES

Definition 3.1 *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability triple and $\{\mathcal{F}_t\}$ be a filtration on $\mathcal{F}$. A stochastic process M is an $\{\mathcal{F}_t\}$ martingale (or just martingale when the filtration is clear) if:*

(M.i) *M is adapted to $\{\mathcal{F}_t\}$;*
(M.ii) *$\mathbb{E}\big[|M_t|\big] < \infty$ for all $t \geq 0$;*
(M.iii) *$\mathbb{E}[M_t|\mathcal{F}_s] = M_s$ a.s., for all $0 \leq s \leq t$.*

Remark 3.2: Property (M.iii) can also be written as $\mathbb{E}[(M_t - M_s)\mathbb{1}_F] = 0$ for all $s \leq t$, for all $F \in \mathcal{F}_s$. We shall often use this representation in proofs.

Example 3.3 Brownian motion is a rich source of example martingales. Let W be a Brownian motion and $\{\mathcal{F}_t\}$ be the filtration generated by W. Then it is easy to verify directly that each of the following is a martingale:

(i) $\{W_t, t \geq 0\}$;
(ii) $\{W_t^2 - t, t \geq 0\}$;
(iii) $\{\exp(\lambda W_t - \frac{1}{2}\lambda^2 t), t \geq 0\}$ for any $\lambda \in \mathbb{R}$ (Wald's martingale).

Taking Wald's martingale for illustration, property (M.i) follows from the definition of $\{\mathcal{F}_t\}$, and property (M.ii) follows from property (M.iii) by setting $s = 0$. To prove property (M.iii), note that for $t \geq 0$,

$$\begin{aligned}\mathbb{E}[M_t] &= \int_{-\infty}^{\infty} \exp(\lambda u - \tfrac{1}{2}\lambda^2 t)\frac{\exp(-u^2/2t)}{\sqrt{2\pi t}}\, du \\ &= \int_{-\infty}^{\infty} \frac{\exp(-(u-\lambda t)^2/2t)}{\sqrt{2\pi t}}\, du \\ &= 1\,.\end{aligned}$$

Thus, appealing also to the independence of Brownian increments, $\mathbb{E}[M_t - M_s|\mathcal{F}_s] = \mathbb{E}[M_t - M_s] = 0$, which establishes property (M.iii).

We will meet these martingales again later.

Example 3.4 Most martingales explicitly encountered in practice are Markov processes, but they need not be, as this example demonstrates. First define a *discrete time martingale* M via

$$M_0 = 0$$
$$M_n = M_{n-1} + \alpha_n, \qquad n \geq 1,$$

where the α_n are independent, identically distributed random variables taking the values ± 1 with probability 1/2. Viewed as a discrete time process, M is Markovian. Now define $M^c_t = M_{\lfloor t \rfloor}$ where $\lfloor t \rfloor$ is the integer part of t. This is a continuous time martingale which is not Markovian. The process (M_t, t) is Markovian but it is not a martingale.

Example 3.5 Not all martingales behave as one might expect. Consider the following process M. Let T be a random exponentially distributed time, $\mathbb{P}(T > t) = \exp(-t)$, and define M via

$$M_t = \begin{cases} 1 & \text{if } t - T \in \mathbb{Q}^+, \\ 0 & \text{otherwise,} \end{cases}$$

$\mathbb{Q}^+$ being the positive rationals. Conditions (M.i) and (M.ii) of Definition 3.1 are clearly satisfied for the filtration $\{\mathcal{F}_t\}$ generated by the process M. Further, for any $t \geq s$ and any $F \in \mathcal{F}_s$,

$$\mathbb{E}[M_t \mathbb{1}_F] \leq \mathbb{E}[\mathbb{1}_{\{t-T \in \mathbb{Q}^+\}}] = 0 = \mathbb{E}[M_s \mathbb{1}_F].$$

Thus $\mathbb{E}[M_t|\mathcal{F}_s] = M_s$ a.s., which is condition (M.iii), and M is a martingale.

Example 3.5 seems counter-intuitive and we would like to eliminate from consideration martingales with behaviour such as this. We will do this by imposing a (right-) continuity constraint on the paths of martingales that we will consider henceforth. We shall see, in Theorem 3.8, that this restriction is not unnecessarily restrictive.

Definition 3.6 *A function x is said to be càdlàg (continu à droite, limites à gauche) if it is right-continuous with left limits,*

$$\lim_{h \downarrow 0} x_{t+h} = x_t$$
$$\lim_{h \downarrow 0} x_{t-h} \text{ exists.}$$

In this case we define $x_{t-} := \lim_{h \downarrow 0} x_{t-h}$ (which is left-continuous with right limits).

We say that a stochastic process X is càdlàg if, for almost every ω, $X_\cdot(\omega)$ is a càdlàg function. If $X_\cdot(\omega)$ is càdlàg for every ω we say that X is entirely càdlàg.

The martingale in Example 3.5 is clearly not càdlàg, but it has a càdlàg *modification* (the process $M \equiv 0$).

Definition 3.7 *Two stochastic processes X and Y are modifications (of each other) if, for all t,*

$$\mathbb{P}(X_t = Y_t) = 1 .$$

We say X and Y are indistinguishable if

$$\mathbb{P}(X_t = Y_t, \text{ for all } t) = 1 .$$

Theorem 3.8 *Let M be a martingale with respect to the right-continuous and complete filtration $\{\mathcal{F}_t\}$. Then there exists a unique (up to indistinguishability) modification M^* of M which is càdlàg and adapted to $\{\mathcal{F}_t\}$ (hence is an $\{\mathcal{F}_t\}$ martingale).*

Remark 3.9: Note that it is not necessarily the case that every sample path is càdlàg, but there will be a set (having probability one) of sample paths, in which set every sample path will be càdlàg. This is important to bear in mind when, for example, proving several results about stopping times which are results about filtrations and not about probabilities (see Theorems 3.37 and 3.40).

Remark 3.10: The restriction that $\{\mathcal{F}_t\}$ be right-continuous and complete is required to ensure that the modification is adapted to $\{\mathcal{F}_t\}$. It is for this reason that the concept of completeness of a filtration is required in the study of stochastic processes.

A proof of Theorem 3.8 can be found, for example, in Karatzas and Shreve (1991). Càdlàg processes (or at least piecewise continuous ones) are natural ones to consider in practice and Theorem 3.8 shows that this restriction is exactly what it says and no more – the càdlàg restriction does not in any way limit the finite-dimensional distributions that a martingale can exhibit. The 'up to indistinguishability' qualifier is, of course, necessary since we can always modify any process on a null set. Henceforth we will restrict attention to càdlàg processes. The following result will often prove useful.

Theorem 3.11 *Let X and Y be two càdlàg stochastic processes such that, for all t,*

$$\mathbb{P}(X_t = Y_t) = 1 ,$$

i.e. they are modifications. Then X and Y are indistinguishable.

Proof: By right-continuity,

$$\{X_t \neq Y_t, \text{ some } t\} = \bigcup_{q \in \mathbb{Q}} \{X_q \neq Y_q\} ,$$